PTO/SB/08b (07-09)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 1 of 1

Complete if Known	
Application Number	12/947,393-Conf. #1801
Filing Date	November 16, 2010
First Named Inventor	Philip J. ABERCROMBIE
Art Unit	2189
Examiner Name	R. G. Bragdon
Attorney Docket Number	2203828.00130US1

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
	AA*	US-7,065,619	06-20-2006	Zhu et al.	

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document Country Code[3] -Number[4]-Kind Code[5] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages Or Relevant Figures Appear	T[6]

NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
	CA	"EMC Data Domain Replicator," White Paper, EMC Corporation 2011 (25 pages)	
	CB	Guo, F. and Efstathopoulos, P., "Building a High-performance Deduplication System," Symantec Research Labs, Symantec Corporation, Culver City, CA, USA (14 pages) (2011)	
	CC	Silverberg, S. "SDFS Overview," April 2010 (17 pages)	
	CD	Tridgell, A., "Efficient Algorithms for Sorting and Synchronization," Thesis, The Australian National University (Feb. 1999) (115 pages)	
	CE	What is Deduplication and Why Does It Matter?, dated May 7, 2010 (2 pages)	
	CF	Zhu, B. et al., "Avoiding the Disk Bottleneck in the Data Domain Deduplication File System," Fast '08: 6th USENIX Conference on File and Storage Technologies, USENIX Association, pp: 269-282 (2008)	

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. * CITE NO.: Those application(s) which are marked with an single asterisk (*) next to the Cite No. are not supplied (under 37 CFR 1.98(a)(2)(iii)) because that application was filed after June 30, 2003 or is available in the IFW. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

Building a High-performance Deduplication System

Fanglu Guo *Petros Efstathopoulos*
Symantec Research Labs
Symantec Corporation, Culver City, CA, USA

Abstract

Modern deduplication has become quite effective at eliminating duplicates in data, thus multiplying the effective capacity of disk-based backup systems, and enabling them as realistic tape replacements. Despite these improvements, single-node raw capacity is still mostly limited to tens or a few hundreds of terabytes, forcing users to resort to complex and costly multi-node systems, which usually only allow them to scale to single-digit petabytes. As the opportunities for deduplication efficiency optimizations become scarce, we are challenged with the task of designing deduplication systems that will effectively address the capacity, throughput, management and energy requirements of the petascale age.

In this paper we present our high-performance deduplication prototype, designed from the ground up to optimize overall single-node performance, by making the best possible use of a node's resources, and achieve three important goals: *scale* to large capacity, provide good *deduplication efficiency*, and near-raw-disk *throughput*. Instead of trying to improve duplicate detection algorithms, we focus on system design aspects and introduce novel mechanisms—that we combine with careful implementations of known system engineering techniques. In particular, we improve single-node scalability by introducing *progressive sampled indexing* and *grouped mark-and-sweep*, and also optimize throughput by utilizing an event-driven, multi-threaded client-server interaction model. Our prototype implementation is able to scale to billions of stored objects, with high throughput, and very little or no degradation of deduplication efficiency.

1 Introduction

For many years, tape-based backup solutions have dominated the backup landscape. Most of their users have been eager to replace them with disk-based solutions that are faster, easier to use (search, restore, etc.) and less fragile. In the past few years, disk-based backup systems have gained significant momentum, and today most enterprises are rapidly adopting such solutions, especially when the data volume is moderate.

One of the most important factors enabling the recent success of disk-based backup is data *deduplication* ("dedupe")—a form of compression that detects and eliminates duplicates in data, therefore storing only a single copy of each data unit. By using dedupe in a disk-based backup system one can multiply the effective capacity by 10-50 times, rendering the system a realistic tape replacement, whose cost is on par with tape-based systems, while also 1) making backup data always available online (for indexing, data mining, etc.), 2) enabling effective remote backups by minimizing network traffic, and 3) reducing client side I/O overhead by eliminating the need to read unchanged, previously backed-up files.

The explosive increase in the amount of data corporations are required to store, however, puts great pressure on the storage and backup systems, creating immediate demand for new ways to address the capacity, performance and cost challenges, and generally increase their overall effectiveness.

The effectiveness of a deduplication system is determined by the extent to which it can achieve three mutually competing goals: *deduplication efficiency*, *scalability*, and *throughput*. Deduplication efficiency refers to how well the system can detect and share duplicate data units—which is its primary compression goal. Scalability refers to the ability to support large amounts of raw storage with consistent performance. Throughput refers to the rate at which data can be transferred in and out of the system, and constitutes the main performance metric.

All three metrics are important. Good dedupe efficiency reduces the storage cost. Good scalability reduces the overall cost by reducing the total number of nodes since each node can handle more data. High throughput is particularly important because it can enable fast backups, minimizing the length of a backup window. Among

the three goals, it is easy to optimize any two of them, but not all. To get good deduplication efficiency, it is necessary to perform data indexing for duplicate detection. The indexing metadata size grows linearly with the capacity of the system. Keeping this metadata in memory, would yield good throughput. But the amount of available RAM would set a hard limit to the scalability of the system. Moving indexing metadata to disk would remove the scalability limit, but significantly hurt performance. Finally, we can optimize for both throughput and scalability, as in regular file servers, but then we lose deduplication. Achieving all three goals is a non-trivial task.

Another less obvious but equally important problem is duplicate reference management: duplicate data sharing introduces the need to determine who is using a particular data unit, and when it can be reclaimed. The computational and space complexity of these reference management mechanisms grows with the amount of supported capacity. Our field experience, from a large number of deduplication product deployments, has shown that the cost of reference management (upon addition and deletion of data) has become one of the biggest real-world bottlenecks, involving operations that take many hours per day, and force a hard limit to scalability.

A lot of the research in the area has focused on optimizing deduplication efficiency and index management, without being able to sufficiently boost single-node capacity: with the current state-of-the-art a single node is limited to a few tens, or hundreds, of terabytes—which is far from sufficient for the petascale. Consequently, scalability has been addressed mostly through the deployment of complex, multi-node systems, that aggregate the limited capacity of each node in order to provide a few petabytes of storage at very high (acquisition, management, energy, etc.) cost. Surprisingly, the problem of reference management performance is largely ignored.

As the rate at which data are generated is rapidly increasing, the pressure for high-performance, scalable and cost-effective deduplication systems becomes more evident. We advocate that single-node performance is of key importance to next-generation deduplication systems: by making the most of a single node's resources, it is possible to build a high-performance deduplication system that will be able to scale to billions of objects. Based on our field experience, we know that such a system would be valuable to a very large number of users (e.g., small/medium businesses) where simplicity is also a top priority. Additionally, we believe that improving single-node performance is essential for multi-node systems as well, since a lot of our techniques can be used to provide more efficient building blocks for these systems, or even collapse them into a single node.

This paper presents a *complete*, single-node deduplication system that covers indexing, reference management, and end-to-end throughput optimization. We contribute new mechanisms to address dedupe challenges and combine them with well-known engineering techniques in order to design and evaluate the system considering all three dedupe goals. *Progressive sampled indexing* removes scalability limitations imposed by indexing, while serving most lookup requests in $O(1)$ time complexity from memory. Our index uses sampling to perform fine-grained indexing, and greatly improves scalability by requiring significantly less memory resources. We address the problem of reference management by introducing *grouped mark-and-sweep*, a mechanism that minimizes disk accesses and achieves near-optimal scalability. Finally, we present a modular, event-driven, client pipeline design that allows the client to make the most of its resources and process backup data at a rate that can fully utilize the dedupe server. As a result, our prototype can achieve high backup (1 GB/sec for unique data and 6 GB/sec for duplicate data) and restore throughput (1 GB/sec for single stream and 430 MB/sec for multiple streams) and good deduplication efficiency (97%), at high capacities (123 billion objects, 500 TB of data per 25 GB of system memory).

The rest of the paper is organized as follows: Section 2 gives a detailed description of the major challenges we had to address. In Section 3 we describe how we address them through our prototype's novel mechanisms, and in Section 4 we present our evaluation results.

2 Challenges

2.1 Indexing

Most deduplication systems operate at the sub-file level: a file or a data stream is divided into a sequence of fixed or variable sized *segments*. For each segment, a cryptographic hash (MD5, SHA-1/2, etc.) is calculated as its *fingerprint* (*FP*), and it is used to uniquely identify that particular segment. A *fingerprint index* is used as a catalog of FPs stored in the system, allowing the detection of duplicates: during backup, if a tuple of the form $<$ FP, location_on_disk $>$ exists in the index for a particular FP, then a reference to the existing copy of the segment is created. Otherwise, the segment is considered new, a copy is stored on the server and the index is updated accordingly. In many systems, the FP index is also crucial for the restore process, as index entries are used to locate the exact storage location of the segments the backup consists of.

The index needs to have three important properties: 1) scale to high capacities, 2) achieve good indexing throughput, and 3) provide high duplicate detection rate—i.e., high deduplication efficiency. Table 1 demonstrates how these goals become very challenging for a

Item	Scale	Remarks
Physical capacity C	C = 1,000 TB	
Segment size S	S = 4 KB	
Number of segments N	$N = 250*10^9$ segs	$N = C/S$
Segment FP size E	E = 22 B	
Segment index size I	I = 5,500 GB	$I = N*E$
Disk speed Z	400 MB/sec	
Block lookup speed goal	100 Kops/sec	Z/S

Table 1: An example system configuration, illustrating some of the challenges involved.

Petascale system. If the system capacity is 1 PB, and the segment size is 4 KB (for fine-granularity duplicate detection), indexing capacity will need to be at least 5,500 GB to support all 250 billion objects in the system. Such an index is impossible to maintain in memory Storing it on disk, however, would greatly reduce query throughput. To achieve a rate of 400 MB/sec, would require the index—and the whole dedupe system for that matter—to provide a query service throughput of at least 100 Kops/sec. Trying to scale to 1 PB by storing the index on disk would make it impossible to achieve this level of performance[1]. Making the segment size larger (e.g., 128 KB) would make deduplication far more coarse and severely reduce its efficiency, while still requiring no less than 172 GB of RAM for indexing.

It becomes obvious that efficient, scalable indexing is a hard problem. On top of all other indexing challenges, one must point out that segment FPs are cryptographic hashes, randomly distributed in the index. Adjacent index entries share no locality and any kind of simple read-ahead scheme could not amortize the cost of storing index entries on disk.

2.2 Reference Management

Contrary to a traditional backup system, a dedupe system shares data among files by default. Reference management is necessary to keep track of segment usage and reclaim freed space. In addition to scalability and speed, reliability is another challenge for reference management. If a segment gets freed while it is still referenced by files, data loss occurs and files cannot be restored. On the other hand, if a segment is referenced when it is actually no longer in use, it causes storage leakage.

Previous work [12, 19] mainly focused on indexing and largely ignored reference management. Some recent work [4, 18] started to acknowledge the difficulty of the problem. But, for simplicity, only simple reference counting was investigated without considering reliability and recoverability. Reference counting, however, suffers from low reliability, since it is vulnerable to lost or repeated updates: when errors occur some segments may be updated and some may not. Complicated transaction rollback logic is required to make reference counts consistent. Moreover, if a segment becomes corrupted, it is important to know which files are using it so as to recover the lost segment by backing up the file again. Unfortunately, reference counting cannot provide such information. Finally, there is almost no way to verify if the reference count is correct or not in a large dynamic system. Our field feedback indicates that power outages and data corruption are really not that rare. In real deployments, where data integrity and recoverability directly affect product reputation, simple reference counting is unsatisfactory.

Maintaining a reference list is a better solution: it is immune to repeated updates and it can identify the files that use a particular segment. However, some kind of logging is still necessary to ensure correctness in the case of lost operations. More importantly, variable length reference lists need to be stored on disk for each segment. Every time a reference list is updated, the whole list (and possibly its adjacent reference lists—due to the lists' variable length) must be rewritten. This greatly hurts the speed of reference management.

Another potential solution is mark-and-sweep. During the mark phase, all files are traversed so as to mark the used segments. In the sweep phase all segments are swept and unmarked segments are reclaimed. This approach is very resilient to errors: at any time the process can simply be restarted with no negative side effects. Scalability, however, is an issue. Going back to the example of Table 1, we would need to deal with $N = 250$ billion segments. If a segment FP is E = 22 bytes, that would be $I = N * E$ = 5,500 GB of data. If we account for an average deduplication factor of 10 (i.e., each segment is referenced by 10 different files), the total size of files that need to be read during the mark phase will be 55,000 GB. This alone will take almost 4 hours on a 400 MB/sec disk array. Furthermore, marking the in-use bits for 250 billion entries is no easy task. There is no way to put the bit map in memory. Once on disk, the bit map needs to be accessed randomly multiple times. This also takes significant amount of time. One might want to mitigate the poor performance of mark-and-sweep by doing it less frequently. But in practice this is not a viable option: customers always want to keep the utilization of the system close to its capacity so that a longer history can be stored. With daily backups taking place, systems rarely have the luxury to postpone deletion operations for a long time. In our field deployment, deletion is done twice a day. More than 4 hours in each run is too much. In a large production-oriented dedupe system reference management needs to be very reliable and have good recoverability. It should tolerate errors and always ensure correctness. Although mark-and-sweep provides

[1] Our measurements show that even high-end SSDs cannot achieve more than 60 Kops/sec

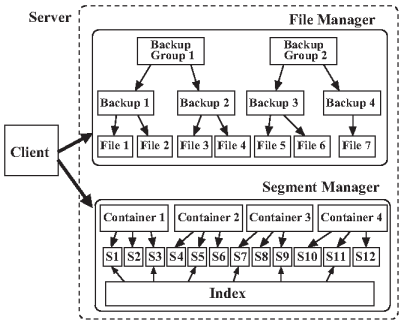

Figure 1: Client and deduplication server components. The server components may be hosted on the same or different nodes.

these properties, its performance is proportional to the capacity of the system, thus limiting its scalability.

2.3 Client-Server Interaction

Even if we solve the indexing and reference management problems, high end-to-end throughput is not guaranteed. An optimized client-server interface is necessary to reap the benefits of deduplication. The typical dedupe client performs the following steps during backup: 1) read data from files, 2) form segments and calculate FPs, 3) send FPs to the server and wait for index lookup results, and 4) for each index miss, transmit the relevant data to the server—otherwise create references to the existing segments. This process may suffer from three different types of bottlenecks. First, reading files from disk is an I/O-bound operation. Second, calculating cryptographic hashes is a very CPU-intensive task, and the client may not be able to compute FPs at the necessary rate. Finally, high latency and low communication throughput may become the main bottleneck for overall performance.

3 Prototype Design

3.1 Goals and System Architecture

We set our performance goals as follows:

- **Scalability**: store and index hundreds of billions of segments.
- **Deduplication efficiency**: best-effort deduplication: if resources are scarce, sacrifice some deduplication for speed and scale.
- **Throughput**: near-raw-disk throughput for data backup, restore, and delete.

To that end, we have implemented a prototype of our scalable duplication system aiming to validate the effectiveness of the proposed mechanisms. Our implementation uses C++ and pthreads, on 64-bit Linux, and it is based on the architecture shown in Figure 1.

The server side component consists of two main modules—the *File Manager* and the *Segment Manager*—that implement all the deduplication and backup management logic.

The File Manager (*FM*) is responsible for keeping track of files stored on the deduplication server. The FM manages file information using a three level hierarchy, visible in Figure 1. The bottom level consists of *files*, each represented by a set of metadata and identified by a file FP, calculated over all segment FPs that the file consists of. The middle level consists of *backups*, that group files belonging to the same backup session. At the top level, multiple backups are aggregated to a *backup group*, allowing the FM to perform coarse-granularity tracking of file/backup changes in the system, so as to assist our reference management mechanism.

The Segment Manager (*SM*) is responsible for the indexing and storage of raw data segments, and may run on the same or a different server than the FM. Segments are stored on disk in large (e.g., 16 MB) storage units, called *containers*. Containers consist of raw data and a catalog which lists all FPs stored in the container. All disk accesses are performed in the granularity of containers. Storing adjacent segments in the same container greatly improves dedupe performance, by reducing container I/O and by improving indexing efficiency (as discussed in Section 3.2.1). The SM also incorporates the dedupe index, and updates it when segments are added/removed.

The client component reads file contents or receives data streams (e.g., data from *tar*), performs segmentation, and calculates segment FPs. After querying the SM index, the client creates references to the existing copies of FPs located in the SM, and initiates data transfers for new FPs. Once a file has been fully processed, the File Manager is updated with file metadata.

Without loss of generality, we use fix-sized, 4 KB segments, for fine-granularity dedupe—although none of the mechanisms relies on this assumption.

3.2 Progressive Sampled Indexing

Most dedupe systems, when performing backup restore, rely on the index—or a similar catalog-like structure—in order to determine the disk location of each segment. This forces the strict requirement for at least one *complete index* containing location information for all FPs, that the system will have to maintain and protect against crashes, corruption etc., because errors cannot be tolerated. If a segment's disk location cannot be determined due to index failure, the whole file or backup gets corrupted. Maintaining such a data structure is a difficult and resource consuming task, that almost certainly impacts system scalability and performance, since the index typically needs to be stored both in memory, for performance, and on disk, for durability.

In order to address the indexing challenges and scale to billions of objects with high performance we had to remove this restriction by introducing *directly locatable objects*: when a file is stored in the system, file segment location information is stored with the file metadata, therefore removing the need to consult the index for the exact location of file segments. For example, if file F consists of segments with FPs A, B and C, stored at disk locations 1, 2 and 3 respectively, F would be represented by the list "$A,1,B,2,C,3$"—instead of just "A,B,C". The increased file metadata size is not a problem, since metadata are stored on disk, while the indexing freedom we get in exchange is extremely valuable.

By decoupling indexing and restore we no longer need to maintain a full index. Instead, we introduce *sampled indexing*, that is based on the observation that given certain amounts of memory and raw capacity, we can calculate the index size, and determine the number of entries that need to be dropped. In particular, if M is the amount of memory available for indexing (in GB), S is the dedupe segment size (in KB), E is the memory entry size (in bytes), and C is the total supported storage (in TB), then we can support M/E billion entries, while the system consists of a total of C/S billion segments. Therefore, if we assume a sampling period T, signifying that we maintain "1 out T" fingerprints in memory, we can define a sampling rate R as follows:

$$R = 1/T = (M/E)/(C/S) = (M*S)/(E*C) \qquad (1)$$

In the example of Table 1, using 22 bytes per index entry, with 4 KB segments and 64 GB of memory for indexing, we can support 11.6 TB of data with a sampling rate of 1 (i.e., a full index). Scaling to 1,000 TB, would require a sampling rate of 0.0116—i.e., insert in the index one out of 86 FPs. Using an 8 KB segment, we could double the raw capacity, or double the rate to $1/43$, sacrificing some dedupe accuracy for higher index density. Increasing the indexing capacity of the system by adding more RAM is rewarded with higher sampling rates (i.e., better dedupe efficiency), while increasing only the storage capacity results in a lower sampling rate, but this is often acceptable, in return for "infinite" system scalability.

3.2.1 *Dedupe efficiency: pre-fetching and caching.*

Since "1 out of T" FPs is inserted in the index, index hits—and, consequently, dedupe efficiency—would be reduced by a factor of T. However, when a lookup operation hits on a sampled FP (also referred to as a "hook"), we locate the container it belongs to and pre-fetch all FPs from that container's catalog into a memory cache. It has been shown [19] that the likelihood of subsequent lookups hitting on the FP cache is high, due to spacial locality: if hook FP A was followed by dropped FP B, then it is very likely that A and B will reappear in order in the future, in which case A will have seeded pre-fetching of its container catalog, resulting in a cache hit for B.

Container catalog pre-fetching can be extremely effective in improving the deduplication efficiency of a sampled index. However, pre-fetching introduces a minimum sampling rate: at least one FP per container (e.g., the first FP stored in the container) must be in the memory index as a hook, in order to seed pre-fetching. Because of this, if container size is K MB, then $R \geq R_{min} = S/(K*2^{10})$ and, subsequently, scalability is no longer "unlimited": the maximum supported capacity is now $C \leq (M*K*2^{10})/E$. For 4 KB segments and 16 MB containers, at least 1 out 4096 FPs needs to be sampled, and with 64 GB of RAM, as in the example of Table 1, $C \leq 47,662$ TB—which is still very high.

Deduplication efficiency. Although the combination of sampling and FP pre-fetching can often yield up to 100% duplicate detection, random eviction of cache entries may reduce deduplication. Using a simplified model we can estimate the dedupe efficiency of the system. Each container catalog contains at most $(K*2^{10})/S = 1/R_{min} = T_{min}$ entries. If we want to achieve deduplication efficiency f%, and we suffer x misses from one container, then:

$$f/100 = 1-(x/T_{min}) \Rightarrow x = T_{min}*(1-(f/100)).$$

If a particular container suffers one eviction during a large time frame (most likely scenario, especially when LRU is used), then all x misses will fall between two consecutive hooks hitting on the index, and therefore:

$$T = 1/R = x+1 \Rightarrow T = T_{min}*(1-(f/100))+1 \Rightarrow$$

$$\Rightarrow (E*C)/(S*M) = T_{min}*(1-(f/100))+1 \qquad (2)$$

Using Equation 2 we can calculate that in the example of Table 1, with 64 GB of memory, the deduplication efficiency will be $f = 97.9\%$. Alternatively, for a given target dedupe efficiency, we can calculate the necessary values to achieve it: for example, if we want $f \geq 95\%$, and given E, C and S, the amount of memory required is $M \geq 26.7$ GB.

3.2.2 *Progressive Sampling.*

A simple, yet important, optimization to sampled indexing is based on the observation that Equation 1 is using the total storage capacity of the system, and, therefore, calculates the value of R_{tot}, required to support all C/S billions of objects. However, at any given time, only the amount of data that are actually stored in the system need to be indexed, which allows us to utilize a *progressive*

sampling rate that calculates R using the amount of storage *used*, as opposed to the maximum raw storage. Initially we set $R = 1$, and gradually decrease it as more storage gets used. In our working example, with 64 GB of RAM, $R = 1$ can index 11 TB of storage. As we approach the 11 TB limit, we can set $R = 0.5$ and down-sample the index (e.g., drop index FPs with $FP \bmod 2 \neq 0$), thus doubling the indexing capacity. Eventually, as usage approaches 1,000 TB, R will converge to $R_{tot} = 0.0116$.

3.2.3 Implementation

The index and cache have been implemented in C++ using a highly parametrizable hash table design, which we call *dhash*, optimized for high performance and efficient memory usage. The M GB of memory available for indexing are divided to fixed size buckets (1 KB by default), allowing us to have a maximum of $Y = M/bucket\_size\_in\_KB$ millions of buckets. No pointers are used in a dhash structure, and all operations use offsets, allowing us to 1) perform custom memory management (bucket slab allocator), 2) get memory savings by replacing each 8-byte pointer with 6 bytes of offset data, and 3) make the dhash easily serializable (e.g., when checkpointing to disk at system shutdown).

If a dhash is used at the role of the index, we aim to accommodate as many sampled FP entries as possible. We utilize 2^b buckets for the hash table, where $b = log_2(Y * 2^{20}) - k$. The system parameter k determines the number of buckets reserved for collision handling. Each index entry contains a partial FP (since the b least significant bits of the FP are encoded in the hash table position), and the container number the FP belongs to. For simplicity we use 128-bit MD5 (which is not strong enough for production, but adequate for our testing purposes), leading to a typical entry size of 18 bytes[2]. Each index dhash also utilizes a Bloom filter, to avoid unnecessary lookup operations, which greatly improves performance.

A cache dhash is optimized mainly for performance: it will use all buckets for the hash table, and handle collisions by running a cache eviction algorithm. A cache dhash can employ one of three eviction policies when collisions for a particular bucket Q occur: *Immediate eviction* will empty Q, and consider all the containers of Q's previous entries as evicted from the cache. This policy is very fast since it performs lazy eviction of FPs, allowing for subsequent lookups to hit on those entries. On the downside, this policy penalizes multiple containers at once. *Eviction by threshold* is similar to immediate eviction, but the containers whose entries are being removed from Q will not be considered as evicted until a certain percentage of their total entries has been removed from all cache buckets. This imposes less of a penalty to containers with entries in Q, but may lead to poor deduplication if the threshold is high, since a particular container may not be pre-fetched even though many of its entries have been evicted. *Container LRU* will evict the entries of the least recently pre-fetched container. If that does not free up space in Q, the process is repeated. Although this is the policy that yields maximum dedupe efficiency, it is also the one with the most overhead. Our default policy is immediate eviction, which provides good deduplication efficiency, and performance only slightly lower than eviction by threshold.

In order to provide high dedupe efficiency after system reboots or crashes, we must ensure that a relatively recent index checkpoint is stored persistently[3]. Bucket change-tracking combined with our pointer-free implementation make checkpointing efficient (only a few seconds per checkpoint). Our current policy creates checkpoints every few minutes, and on system shutdown.

SSD indexing. Although sampling provides an efficient way around scalability restrictions imposed by memory limitations, we wanted to also provide a way to improve scalability even with modest amounts of memory, and without having to resort to very low sampling rates. To that end we have also implemented a (persistent) SSD-based version of our sampled index. Sampled fingerprints are stored on sorted SSD blocks and all available memory is used for three performance optimizations: 1) create an SSD summary data structure *SSD_sum*, 2) maintain a Bloom Filter for the SSD index, and 3) maintain an FP cache of pre-fetched containers—similar to that used for the memory index. The *SSD_sum* data structure keeps track of the first FP in each of the SSD's (sorted) blocks, thus allowing us to perform any lookup with at most one SSD block read: when a *lookup(X)* operation is performed, X may be found in the cache, or it may be found by reading the SSD block i, where $SSD\_sum(i) \leq X < SSD\_sum(i+1)$. The SSD index is read-only, eliminating the need for shared locking during accesses. All SSD index updates are cached and logged. Eventually, index updates are performed in batches (and with the SSD exclusively locked): for our 128 GB SSD a full update takes less than 9 minutes, and we can afford to update the SSD many times per day.

3.3 Grouped Mark-and-Sweep

The challenge in reference management, as discussed in Section 2.2, is to ensure reliability while ensuring that the reference management mechanism is also both scalable and fast enough to keep up with the backup speed. A mark-and-sweep approach is very reliable, but offers

[2] With a stronger 160-bit hash, the entry size becomes 22 bytes.

[3] Notice that even if we lose all index index entries, correctness is preserved.

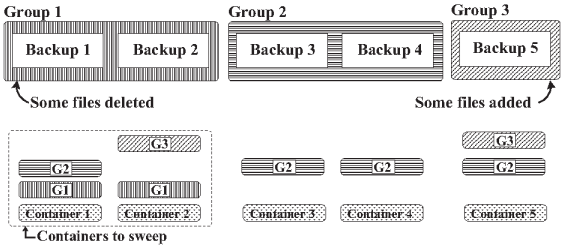

Figure 2: Example illustrating the scalability of grouped mark-and-sweep.

poor scalability because it needs to touch every file in the system. To address this challenge we propose the *grouped mark-and-sweep* (GMS) mechanism, which is reliable, scalable, and fast. The key idea is to avoid touching every file in the mark phase and every container in the sweep phase. GMS achieves scalability because its workload becomes proportional to the changes—instead of the capacity of the system.

The operation of GMS is based on change-tracking within the File Manager. As presented in Figure 1, the File Manager keeps track of files, backups, and backup groups. A file can be a regular file, a database backup stream, an email, etc. A backup is a set of files, e.g., all files under a set of directories. The creation and contents of backups are in the control of the user.

Backup groups aim to control the number of entries that GMS needs to manage, and are created and managed by the File Manager. When backups are small, we aggregate multiple small backups to one bigger backup group. The File Manager tracks changes to each backup group, and for each changed backup group, it further tracks whether files have been added to or deleted from it. During a GMS run, the following steps take place:

1. **Mark changed groups.** Only mark the changed backup groups and do nothing for unchanged backup groups. As an example in Figure 2, assume that File Manager's change tracking shows that, since the last GMS cycle, we deleted some files from group Group1, added some files to group Group3, and made no modifications to Group2. In this case we only need to touch files in backup groups Group1 and Group3. Usually, most backup groups (e.g., Group2) are not changed and files in those groups don't need to be marked. The mark results of G1 and G3 are recalculated by traversing all files in Group1 and Group3 and recalculating G1 and G3 for all containers that have segments used by those files. A group's mark results, say G1, is a bitmap implemented as a file for each container.
2. **Add affected containers to the sweep list.** Only containers used by groups that have deleted files need to be swept because only those containers may have segments freed. In the example of Figure 2, Group1 has files deleted and it has used containers 1 and 2. So we put these two containers in the sweep list. The segments in other containers are either still referenced by files in the unchanged groups (say Group2), or referenced by new files in new groups (say Group3).
3. **Merge, sweep, and reclaim freed space.** For each container in the sweep list, we merge the mark results of all groups using that container. If a segment is not used, it can be reclaimed. In the example of Figure 2, for Container 1, we merge (the old) G2 and (the new) G1, to determine potentially unused segments. Similarly, we merge (the new) G3 and (the new) G1, to determine potentially unused segments in Container 2.

As it becomes clear from the example of Figure 2, GMS provides two important scalability benefits. First, old mark results (e.g., G2) can be reused, without having to re-generate them in every mark-and-sweep cycle. Each set of mark results is stored and reused in the future, making the mark phase scalable by avoiding to touch the majority of the unchanged backup groups. Secondly, unlike conventional mark-and-sweep where all the entries are swept to determine the unused entries, in GMS we know which containers have reference removal operations, and the system only needs to sweep that subset of containers. Therefore the majority of containers in the system are usually not touched in the sweep step.

One drawback of GMS is that a group needs to be remarked even if just one file has been deleted from it. Fortunately the overhead is surprisingly small: segments can be marked at a rate of 26 GB/sec. Since most bitmaps are not changed, there are little work in the sweep phase.

Overall, GMS makes mark-and-sweep scalable by only touching the changed objects, while maintaining the reliability of mark-and-sweep. If errors occur, the whole process can start over and all operations are idempotent. Finally, the mark results (e.g., G1 and G2 for Container 1) serve as a coarse reference list for segments in the containers. When data corruption occurs in a container, the mark results can give us a complete list of backup groups that use that particular container. This limits the set of affected files significantly, and greatly enhances recoverability. Otherwise, we would need to go through all files in the system to determine which files are using that container.

Discussion. An interesting issue related to reference management is concurrent reference updates (data deletion) and data backup. In the example of Figure 2, Backup 5 may still be active when it gets marked, and after all changed backup groups are marked, GMS determines that segment x can be deleted. If Backup 5 uses

x between the time Backup 5 was marked and the time that GMS deleted segment x, data loss will occur as a backup uses deleted/non-existent segments. HYDRAstor [4] uses a read-only phase to freeze the system while updating segment reference counts. In practice, the viability is dubious. On a busy system, there are always some active backups. It is very unlikely to find a time window when the system can be frozen.

Our system uses an in-memory protection map to address this problem: after GMS begins, all segments used by current active backups are protected by storing their segment fingerprints in a protection map in memory. GMS only deletes segments whose fingerprint is not in the protection map. This way GMS can be certain that segments in use will never get deleted. The protection map grows while GMS is running and gets deleted once GMS completes. This is another reason why GMS needs to be fast enough to prevent the protection map from using too much memory. To mitigate the time spent in GMS, and limit the growth of the protection map, GMS can be done more frequently.

3.4 Client-Server Interaction

Even with high-performance server components, it is impossible to achieve high throughput, unless the client is able to push data to the server at a high-enough rate. To that end, our client component is based on an event-driven, pipelined design, that utilizes a simple, fully asynchronous RPC implementation.

Our RPC protocol is implemented via message passing over TCP streams or system IPC mechanisms (e.g., named pipes), depending on whether communication is remote or local. The TCP implementation utilizes multiple TCP connections to keep up with the throughput requirements. All RPC requests are asynchronous and batched in order to minimize the round-trip overheads and improve throughput. A client can register different callback functions for each type of RPC. The callback functions are used to deliver the RPC results to the caller as they become available.

Based on our asynchronous RPC protocol, we have implemented an event-driven client pipeline, presented in Figure 3, where each backup step is implemented as a separate pipeline stage.

First, the reader thread R receives the backup schedule, reads large chunks of data (e.g., 256 segments), and enqueues requests to the hash queue HQ. The hashing thread H dequeues requests from HQ, performs segmentation for each data chunk, and calculates FPs. Calculating cryptographic hashes is a computationally expensive operation, and, in order to fully utilize multiple CPU cores, H employs n MD5 worker threads ($H_1, H_2, \ldots, H_n$) that calculate FPs asynchronously. Once a chunk's segment FPs have been calculated, callback function $CB1$ enqueues the updated request to the lookup queue LQ.

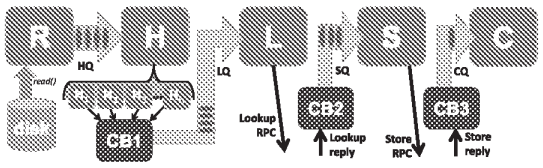

Figure 3: Client pipeline, consisting of five main event-handling threads connected using queues.

The lookup thread L receives requests from LQ and issues one single, batched, asynchronous lookup RPC to the server, incurring a single RPC round-trip for all 256 FPs. Callback function $CB2$ delivers the RPC reply and creates references to the containers of the FPs that were found on the server. If one or more FPs were not found, CB2 enqueues the updated request in the store queue SQ.

The store thread S receives requests from SQ, and sends raw data blocks to the back-end through one single, batched, asynchronous RPC. Callback function $CB3$ ensures that the write operation was successful, and forwards the last request for each file to the close queue CQ.

Finally, close thread C, receives the final request from CQ, performs cleanup, calculates file metadata, and updates the File Manager.

Client queues allow us to better understand system behavior. For instance, on a client with low hash calculation throughput, we can observe HQ to be full most of the time, while low network performance will lead to LQ and SQ being mostly full. In such cases, more than one threads can be used for each pipeline stage. By using two store threads, for example, we can consume requests from SQ at a higher rate.

4 Evaluation

Our main test-bed is an 8-core Xeon E5450 at 3 GHz with 32 GB RAM, running Linux. Our 24 TB disk array consists of 12 disks, 2 TB each, and uses RAID 0[4] to stripe all physical disks to a single logical volume.

We used two main data sets for testing. Our synthetic data set consists of multiple 3 GB files, each with globally unique data segments. Our second data set consists of virtual machine images, which are a very common real-world enterprise use-case, that takes advantage of deduplication. We use a VMware "gold" disk image ($VM0$), hosting a Microsoft Windows XP installation, and created three additional versions of it ($VM1$, $VM2$, and $VM3$), each with incremental changes: VM1 is VM0 with all Microsoft updates and service packs, VM2 is

[4]RAID 0 is not recommended for a high-availability system, but we used it to achieve maximum performance and mitigate the disk bottleneck—thus emulate a high-end array.

VM1 with a large anti-virus suite installed, and VM3 is VM2 after the installation of various utilities (document readers, compression tools, etc.). This data set aims to measure the "real-world" dedupe performance of our system, using a file type of great importance for the enterprise.

For both data sets we configured the system to use a sampling rate of $R = 1/101$, which is low enough to stress the system. For the synthetic tests performed on our current test-bed, the index uses 25 GB memory to hold 1.23 billion FPs. With a sampling rate of 1/101, this is equivalent to a full index of 124 billion FPs, or 500 TB of raw storage—given that our segment size is 4 KB[5].

4.1 Throughput

4.1.1 Backup Throughput

Index throughput. Before performing any macro-benchmarks, we used micro-benchmarks to ensure that the index can support our goals—e.g., in the example of Table 1, at least 400 MB/sec. In all the micro-benchmarks the index could easily handle the desired rates: insert/lookup/remove cost does not exceed 7,619/12,020/16,836 cycles, respectively, even when index occupancy is more than 97%. For instance, on a 3 GHz CPU, and in the worst-case scenario where all incoming FPs exist in the system (and the Bloom filter is of no help), the index can sustain a backup rate of around $975 * T$ MB/sec, where T is the sampling period. For our test configuration, $T = 101$, and the index can sustain a rate of about 98.5 GB/sec.

Unique data: baseline vs. prototype. Figure 4 shows the backup throughput using the synthetic data set. We vary the number of concurrent backups, in steps of 1, 4, 16, 32, 64, 128 and 256, in order to evaluate the system's capability for concurrency. For consistency, all backups consist of multiple 3 GB files that add up to 768 GB.

The unique data throughput test aims to measure the prototype's behavior in the absence of duplicates. Unique data can be significant when a client performs the initial backup or a lot of changes have been made. This test stresses the disk and the network systems as large amounts of data need to be transferred.

To get a sense of the performance of raw hardware, we first measured a baseline throughput. The baseline throughput of the disk array ("Baseline" in Figure 4), is measured by writing the same synthetic workload to the file system. For a single backup, the baseline throughput is around 1 GB/sec. This is the maximum throughput of the storage system. The baseline throughput quickly drops to around 300 MB/sec for storing multiple backups concurrently because disk contention increases with the number of concurrent backups.

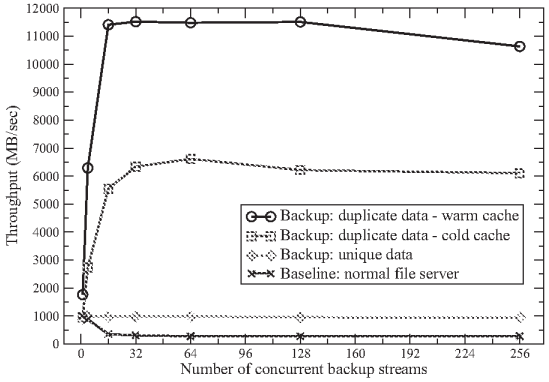

Figure 4: Aggregate throughput for our synthetic data set, with varying number of concurrent backups. Our system is capable of 6 GB/sec for duplicate data backup, and close to 1 GB/sec for concurrent backups of unique data. Dedupe efficiency is 97%, and we support 200 TB storage for every 10 GB of system memory (500 TB for 25 GB in this test).

Backing up the same data set ("Unique data" in Figure 4) using our prototype achieves a steady throughput of about 950 MB/sec as we scale to multiple concurrent backups, which is significantly better than the regular file server. This is mainly because our prototype performs segmentation on all incoming data, and manages the serialization of containers to disk (regardless of content source), therefore decreasing concurrent disk accesses.

Duplicate data backup throughput. After backing up the unique data workload using our prototype, we backup the same files again ("Duplicate data - cold cache" line in Figure 4). This time, all segments are duplicates, and we aim to observe how our prototype performs when it only needs to reference existing data, instead of physically storing new data. This test mainly stresses the index lookup and disk pre-fetching operations.

Initially, for low levels of concurrency, the penalty for small random disk reads, for container FP catalog pre-fetching, dominates performance. Throughput improves steadily as we increase the level of concurrency and duplicate elimination pays off, with aggregate disk throughput reaching over 6.6 GB/sec for 64 concurrent backup streams. When disk accesses are already random, concurrent access doesn't introduce more randomness. On the other hand, concurrent accesses can fully utilize every disks in the disk array. Thus the aggregate throughput increases. After 64 concurrent streams, the disk ar-

[5]Testing our system with a configuration that supports a raw capacity of 500 TB per node may seem inadequate at first. One should keep in mind, however, that 1) We are stressing the system by using 4 KB segments. Most systems use significantly larger segments, leading to higher raw capacities. 2) This is *single-node* capacity with only 25 GB memory for indexing. As such, it is higher than that of most systems we know of (as presented in Section 4.4). Unfortunately we don't have access to servers with more memory or larger disk arrays so as to test higher capacities.

Backup streams	Unique data	Duplicates (cold cache)	Duplicates (warm cache)
1	840 (-4.9%)	699 (-26.4%)	1,989 (12.9%)
4	992 (-0.5%)	2,556 (-6.3%)	6,326 (0.6%)
16	999 (1.9%)	4,802 (-0.2%)	11,992 (5.1%)
32	985 (0.3%)	6,420 (1.3%)	12,134 (5.3%)
64	984 (-0.2%)	6,621 (0.1%)	11,865 (3.3%)
128	988 (3.2%)	6,315 (1.6%)	11,755 (2.1%)
256	955 (1.9%)	6,041 (-1.1%)	11,946 (12.3%)

Table 2: We repeated the experiments of Figure 4 using the SSD index. Results are in MB/sec. The percentages in parentheses show how much faster/slower the SSD index is from the memory index.

ray's capacity for pre-fetching is saturated and mild effects from concurrency overhead (index/cache locking, disk accesses etc.) are becoming obvious: duplicate data backup throughput falls to 6 GB/sec and remains mostly constant.

To verify our conjecture that duplicate data backup throughput limitations are mainly due to disk bottleneck (container FP catalog pre-fetching) instead of CPU, we backup the same files a third time immediately after the second backup. In this case, many FPs are already in the cache and fewer disk pre-fetches will be necessary. The throughput is shown as "Duplicate data - warm cache" in Figure 4. First we observe that overall throughput is much higher, reaching 11.5 GB/sec at around 16 streams, confirming that the bottleneck in our previous tests was in the disk random access performance, which determines the duplicate backup throughput. Additionally, we observe that the effects of concurrency are barely visible: aggregate throughput is stable up to 128 concurrent backups, but at 256 concurrent streams the overhead of pthread shared locks used for protected accesses to the FP cache buckets, as well as a few cache evictions that render the cache less "warm", take their toll—slightly lowering the aggregate throughput (10.6 GB/sec).

SSD indexing throughput. Using SSD index implementation on an 128 GB SSD drive, we repeated the throughput experiments of Figure 4 in order to 1) test the efficiency of our SSD indexing design, and 2) verify the effects of shared locking to duplicate data backups—since the SSD index is read-only and uses no shared locks. For our tests, we maintained the same sampling rate ($R = 1/101$) and used the same amount of memory for caching as before (2 GB)—so as to make a fair comparison. Notice that with this setup we are now using a total of only 10 GB and the amount of raw storage the system can support rose from 500 to 1,600 TB. Due to our efficient SSD index design and the lack of shared locking, most throughput results were similar or superior to those of the memory index. Table 2 summarizes the results and difference between the SSD index and memory index throughput. Notice, however, that these results include the cost of updating the SSD every time 65,536 new sampled entries have accumulated. A less (more) frequent SSD update policy would yield faster (slower) throughput results.

CPU cores	Unique data	100% Duplicates (cold cache)	100% Duplicates (warm cache)
1	347	354	356
2	599	612	612
4	900	1,167	1,172
8	907	1,983	2,004
14	925	2,373	2,485

Table 3: End-to-end backup throughput using a varying number of CPU cores. All numbers in MB/sec.

End-to-end throughput. Our next test attempted to include client performance in our evaluation, in an end-to-end system test, using a single 25 GB backup stream of unique segments. As presented in Table 3, we varied the number of CPU cores dedicated to MD5 calculation, and performed three tests for each configuration: an initial backup, a second backup of the same data with cold caches, and a third run with warm caches. All backups were performed using a 16-core Intel Xeon E5520 "client", with 32 GB of RAM, running RedHat Enterprise Linux 5. The results of Table 3 show that backing up unique data does not get much faster with more than 4 cores. Careful observation revealed two reasons for this behavior. First, even when using the Linux loopback interface, we could not get throughput higher than 10 Gbps, on that particular host. Notice that when bulk data transfers become unnecessary, the performance reaches 2.49 GB/sec. Second, we realized that careful optimization of our simple RPC mechanism might be able to yield better performance. However, optimizing network behavior and the RPC implementation is beyond the scope of this study. In order to evaluate the real throughput of our client design we made the assumption of an infinitely fast network/RPC infrastructure, and, temporarily, eliminated the network performance bottleneck. This revealed the client's full potential: running on our (slower) main Intel Xeon 5450 server, the client was able to push 360/697/1,023/1,319 MB/sec of unique data, with 1/2/3/4 cores dedicated to MD5, respectively.

Backup throughput conclusions. In summary, our backup throughput experiments show that, when backing up unique data, our system is nearly as efficient as a normal file server for single stream backup (no penalty for deduplication) and several times faster for multi-stream backups. This shows that our system can better organize the data on disks to achieve high throughput even with concurrent backups. When data are mostly duplicates, we can achieve 950 MB/sec for single stream backup and 6 GB/sec for multi-stream backups. Multiple

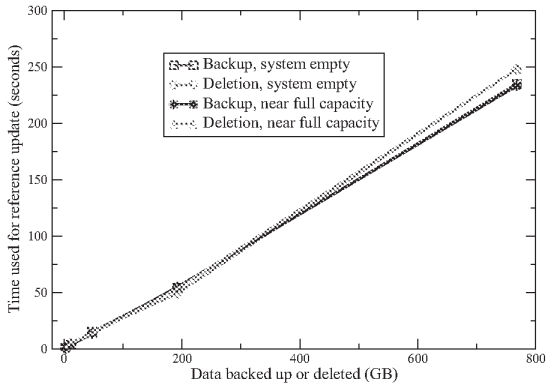

Figure 5: The reference update time for a given amount of data backed up or deleted when the system is empty and nearly full. The time is proportional to the data changed, and the slope shows the update throughput (3 GB/sec). Notice that the throughput is stable regardless of the capacity of the system or the amount of changed data.

streams help improve the aggregate throughput because they maximize the throughput of container FP catalog pre-fetching.

The major limitations that we observed are due to hardware restrictions: limited container pre-fetching throughput and CPU/networking bottlenecks in our end-to-end performance tests. On a production system equipped with hundreds of fast-seeking physical disks, and utilizing faster network connectivity, we expect to see much higher throughput. The only software limitation we observed was due to pthread locks, and is considered of secondary importance since it only impacts throughput minimally for more than 128 concurrent backup streams.

4.1.2 Reference Update Throughput

A critical property that is not often tested in deduplication systems, is the performance of reference updates, especially when we need to delete data—an operation that happens almost daily. Figure 5 shows reference update times measured when the synthetic data set gets backed up or deleted, both when the system is empty and near full capacity. The time is linear with the size of data backed up or deleted, because we need to update the reference of each segment that gets used.

The slope of the line corresponds to the throughput of the reference update, which is 3.2 GB/sec for data addition, and 3.1 GB/sec for data deletion. Deletion is slightly slower because when segments get deleted, they also need to be removed from the index. Contrary to a regular file-system, the deletion throughput of the deduplication system is slow because we pay the price of data sharing. However, it is still faster than the backup throughput of new data, which prevents the backup process from having to stall and wait for the deletion mechanism to free up space.

	Unique segs	Total unique	Ideal MBs	Real MBs	De-dupe
VM0	518,326	518,326	2,123	2,211	96%
VM1	733,267	921,522	3,775	3,938	96%
VM2	904,579	1,189,230	4,871	5,085	96%
VM3	1,145,029	1,616,585	6,621	6,860	97%

Table 4: Deduplication efficiency results for subsequent backups of four different versions of a Windows XP VMware image file.

4.1.3 Restore Throughput

Deduplication system benchmarks are dominated by backup testing and testing of restore is mostly ignored—probably because the restore process is usually slow, and correctness is the main concern. However, restore is an important operation and we wanted to ensure that our prototype provides sufficient performance. During our tests all data were restored correctly. Our single stream restore throughput was measured around 1 GB/sec, and 430 MB/sec for two or more concurrent restore streams. Single stream restore is fast because most accesses are sequential, while multiple concurrent restore streams introduce disk seeking. The use of directly locatable objects allows us to perform restore without using the index, making the whole process very scalable.

4.2 Deduplication Efficiency

Although we are willing to sacrifice some dedupe accuracy for high scalability, we still want to make sure the system provides adequate duplicate detection. In particular, since sampling provides the desired scalability, dedupe efficiency will be mostly determined by the effectiveness of pre-fetching.

In our synthetic data set, the true ("ground truth") duplication is 100%. Our prototype consistently eliminates no less than 97% of duplicates. This is consistent with the theoretical expectation, based on Equation 2: when we pre-fetch FPs from the container catalog, and because the sampling rate is 1 out of 101, the first 100 FPs may not be found. After the first hit, (101st FP in the worst case), we pre-fetch all FPs in that container. So theoretically we may fail to detect 100 over 4096 FPs, i.e., 2.4%.

For our VMWare data set we used our test sampling rate of 1/101, and a small FP cache (256 MB) in order to ensure that the cache cannot hold the whole working set. We performed multiple backups of each VM image, observing 100% dedupe efficiency for each run, with very high throughput (2.4 GB/sec). A more interesting experiment, however, presented in Table 4, is the dedupe efficiency achieved when backing up VM0, VM1, VM2, and VM3 back-to-back. Image VM0 has 518,326 4 KB segments, taking up 2,211 MB of disk space, instead

of 2,123 MB, giving us 96% of the ideal dedupe efficiency. Backing up VM1 introduced 403,196 new segments (330,071 of VM1's segments were also in VM0), taking up 3,938 MB, for a steady dedupe efficiency of 96%. Similarly, VM2 and VM3 were deduplicated at 96% and 97% of the optimal dedupe rate, which is a very satisfying result for a cache of only 256 MB. These results are particularly encouraging, since field experience has demonstrated that VM image backups are one of the most common and effective uses of dedupe.

4.3 Scalability

In order to test the scalability of the system we first populated it to near-full capacity (480 out of 500 TB i.e., 95.5%) with unique data. Because our disk array is only 24 TB, we stored everything except the actual segment data. As the code mainly operates on the metadata, discarding segment data has no impact on the correctness of the test. After the system was populated we repeated the same throughput tests, during which everything was stored on disk (including segment data).

Figure 6 presents a throughput comparison between an empty and near-full system. For multi-stream throughput, the system occupancy has negligible performance impact because for both unique and duplicate data the throughput is, once again, bounded by the disk's sequential write and random read performance, respectively. When the system is near full capacity, the index lookup and update time increase slightly. But the main bottleneck is still disk I/O—overshadowing the effects of CPU overhead. This means that the throughput of the system will scale well in terms of system capacity while disk I/O is the main bottleneck—which is probably going to be true in the foreseeable future.

The index overhead does show up for single stream throughput. The throughput of single stream backup near full capacity is slower than that of the empty system because single stream throughput is CPU bound and accessing a "fuller" index takes a little bit more CPU time.

Figure 5 also compares reference update performance when the system is empty and near-full. As expected, the time for reference update is almost the same, since the grouped mark-and-sweep algorithm only touches the changed backup groups. The majority of the references, regardless of how many they are when the system is near full capacity, are not touched by the grouped mark-and-sweep. Finally, we also checked the deduplication efficiency for both the synthetic and real data sets and observed no degradation in a near-full system.

Our results demonstrate that all parts of our prototype are able to scale to high capacity, with almost no performance decrease. We are confident that our system would scale to higher capacities, given more resources. Moreover, the raw capacity supported by our system (200 TB for every 10 GB of memory) is higher than the capacity of any other single-node system presented in Section 4.4.

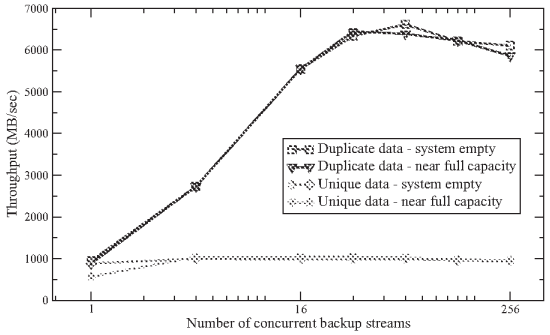

Figure 6: Throughput scalability tests show that there is no significant throughput drop when we get close to full capacity. We incur O(1) cost for most index operations, and throughput is disk-bound for both unique and duplicate data backups.

4.4 Comparison to State-of-the-art

When evaluating dedupe systems it is often the case that custom methods and private workloads are used to quantify the effectiveness of the proposed mechanisms (e.g., [19] and [12]). Comparisons to other systems are usually difficult, and limited to references to results reported by vendors, mostly because there is no agreed deduplication benchmark that would make benchmarking and comparisons fair and meaningful. Furthermore, when aiming to top the performance of state-of-the art systems, it is almost impossible to justify the cost and effort of obtaining, deploying and benchmarking even a single one of them. In our evaluation we tried to use data sets that will exercise the system in interesting ways, and that are relatively easy to be recreated and tested by other systems.

Table 4.4 presents some of the most popular high-performance deduplication solutions available as of April 2011. Assuming that all systems provide adequate deduplication efficiency (specifications do not provide precise numbers), we can see that our prototype's peak performance is similar to or better than that of all systems, with the exception of NEC's HydraStor. However, notice that HydraStor utilizes a large distributed system (55 "accelerator" and 110 "storage" nodes) in order to achieve its maximum throughput, and yet its raw capacity is limited to only 1.3 PB. Our prototype's single-node scalability competes with that of all systems and surpasses most of them, especially considering the limited amount of resources we have used (e.g., only 25 GB of RAM per 500 TB for $R = 1/101$, on an older 8-core server). Notice, however, that our goal to increase single-node scalability is not meant to replace all multi-node systems, but to potentially improve them by enabling each node to make better use of its resources and increase

Product	Backup (MB/sec)	Capacity (TB)	Nodes
DataDomain DD890 [3]	4,083	384	1
HP D2D4324 [7]	1,110	18	1
IBM ProtecTier [8]	1,000	1,000	2
Greenbytes GB4000 [6]	950	216	1
NEC HydraStor HS8-3110R [14]	41,250	1,300	55 + 110
Our prototype	6,000	500	1

Table 5: Summary: state-of-the art dedupe products as of April 2011.

data density per node. By doing so we could decrease the number of nodes necessary for a particular deployment, thus significantly decreasing the overall (acquisition, management, energy, etc) cost.

5 Related Work

Since the days of early deduplication systems, that performed mostly file-level or naive block-level deduplication [1, 11, 16], a lot of effort has been put into optimizing duplicate detection. In particular, many systems have investigated methods to perform content-aware segment boundary calculation, aiming to improve better duplicate coverage. Any degradation in dedupe efficiency was considered unacceptable. Such variable-size segmentation algorithms, utilize different variations of byte-level approaches, such as sliding window approaches (e.g., [5]), rolling hashes (e.g., [15]), Rabin fingerprints [2], and bimodal chunking [10]. For instance, systems like MAD2 [18], HYDRAstor [4, 17], as well as deduplication solutions by DataDomain [19] and Hewlett-Packard [12], utilize variable-size segments, in an attempt to achieve maximum compression. However, even if these algorithms make the best of raw storage (which is not always the case, as observed by [9]), single-node capacity is limited. Our work takes a different approach: we are willing to sacrifice some deduplication efficiency in order to achieve higher single-node scalability.

A sampling method is used in [12] to address indexing scalability restrictions. However, that approach is significantly different from ours, since it uses sampling to probabilistically identify "super-segments" that are used to perform coarse-granularity deduplication. Our segmentation algorithm operates at fine granularity at all times, and sampling is not used for pattern-matching, but for indexing actual file segments. Additionally, our approach is significantly more scalable, and can operate under heavy memory constraints, with good sampling rates: in a setting similar to the experiments presented in [12], our sampled index would require about 74% less memory (4.4 GB instead of 17 GB, with $R = 1/101$).

A lot of systems have used spacial locality to perform some type of caching (e.g., [18, 19]), but, to our knowledge, it has not been used before in combination with an aggressive sampling approach, such as the one we are proposing.

Our key assumption difference from previous efforts is that we are willing to relax our duplicate detection efficiency requirements, in order to address *all three* major challenges of single-node deduplication. Most other systems have provided good solutions for a subset of problems, usually excluding single-node scalability and reference management. For instance, DataDomain [19] addressed the disk bottleneck, by introducing a series of optimizations, including a Bloom filter, and spacial locality. However, their system can support a limited amount of raw storage, and is limited by network performance, since duplicate detection is performed only at the server. Additionally, it is not clear whether DataDomain's system can perform truly scalable resource reclamation.

HYDRAstor [4] on the other hand, achieves good scalability, but it does so by using a highly distributed, hierarchical model, with each node holding only a few tens of TB of storage. This design yields a high backup throughput, but at the cost of a highly distributed, costly system. Deletions in HYDRAstor, are implemented with a distributed reference counting method, which is difficult to maintain correctly, and scale without a large performance hit.

MAD2 [18] also uses a distributed storage system to provide scalability, as well as a number of optimizations that include spacial locality caching, and Bloom filters. Deletions are a very challenging operation in this system as well: they are performed only at the file level, and they are also handled by a variant of reference counting, with all the scalability and correctness problems discussed in Section 2.2. To our knowledge, our grouped mark-and-sweep approach is the only truly scalable, documented reference management implementation, that is also very resilient to errors.

Many scalable systems have adopted the event-driven design, however it is interesting that the nature of our application requires that we utilize it for the *client*, rather than the server. A pipelined client design was also proposed by [13], but it is significantly different from our design: it assumes pipeline stages whose operation requires a fixed amount of time, making it unrealistic for network operation. It also uses disk-based, client-side indexing, it implements a lot of functionality in the kernel, and it achieves scalability and throughput that is orders of magnitude lower than those of our client design.

6 Conclusion

Important engineering challenges need to be addressed in order to achieve the scalability, throughput and dedu-

plication efficiency necessary to provide next-generation deduplication support. We have presented a clean-slate design that aims to maximize overall single-node effectiveness, and introduces new mechanisms that address the most pressing of these challenges. Our directly locatable objects enable the use of progressive sampled indexing—in memory or on SSD—which provides superior single-node scalability and memory usage efficiency—unlike any other system we know of. Our grouped mark-and-sweep mechanism attacks the difficult, and often neglected, resource management and reclamation problem, in a truly scalable and efficient manner. Additionally, we have proposed an asynchronous interface to the server back-end, capable of pushing data to the server at a high-enough rate.

The performance of our prototype validates the effectiveness of our design. Progressive sampled indexing achieves very good deduplication efficiency, while using only 10 GB of memory per 200 TB of raw storage (25 GB for 500 TB in our tests). Additionally, we were able to achieve backup throughput ranging from 950 (all unique data) to 6,000 MB/sec (all duplicate data), with deduplication efficiency no less than 97%, while our grouped mark-and-sweep approach can process data with speeds higher than 3.1 GB/sec, demonstrating that single-node dedupe effectiveness can be greatly improved by making good use of available resources.

Acknowledgments

The authors would like to thank Weibao Wu, Joe Pasqua, Sanjay Sawhney, Kent Griffin, Tzi-cker Chiueh, Vish Janakiraman, Vic Pantaleon, and the Symantec PureDisk team for their input and their help on evaluating the system. We would also like to thank the anonymous reviewers for their valuable comments.

References

[1] ADYA, A., BOLOSKY, W. J., CASTRO, M., CERMAK, G., CHAIKEN, R., DOUCEUR, J. R., HOWELL, J., LORCH, J. R., THEIMER, M., AND WATTENHOFER, R. P. Farsite: federated, available, and reliable storage for an incompletely trusted environment. *SIGOPS Oper. Syst. Rev. 36*, SI (2002), 1–14.

[2] BRODER, A. Z. Some applications of rabin's fingerprinting method. In *Sequences II: Methods in Communications, Security, and Computer Science* (1993), Springer-Verlag, pp. 143–152.

[3] DATADOMAIN. EMC data domain dd890. `http://bit.ly/eXL6Uc`.

[4] DUBNICKI, C., GRYZ, L., HELDT, L., KACZMARCZYK, M., KILIAN, W., STRZELCZAK, P., SZCZEPKOWSKI, J., UNGUREANU, C., AND WELNICKI, M. Hydrastor: a scalable secondary storage. In *FAST '09: Proccedings of the 7th conference on File and storage technologies* (Berkeley, CA, USA, 2009), USENIX Association, pp. 197–210.

[5] FORMAN, G., ESHGHI, K., AND CHIOCCHETTI, S. Finding similar files in large document repositories. In *KDD '05: Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining* (New York, NY, USA, 2005), ACM, pp. 394–400.

[6] GREENBYTES. GB-X Series. `http://bit.ly/fTOSXd`.

[7] HP. StorageWorks D2D Backup Systems. `http://bit.ly/bfhyiU`.

[8] IBM. ProtecTIER Deduplication Solutions. `http://bit.ly/dVg3Z5`.

[9] JIN, K., AND MILLER, E. L. The effectiveness of deduplication on virtual machine disk images. In *SYSTOR '09: Proceedings of SYSTOR 2009: The Israeli Experimental Systems Conference* (New York, NY, USA, 2009), ACM, pp. 1–12.

[10] KRUUS, E., UNGUREANU, C., AND DUBNICKI, C. Bimodal content defined chunking for backup streams. In *FAST'10: Proceedings of the 8th USENIX conference on File and storage technologies* (Berkeley, CA, USA, 2010), USENIX Association, pp. 18–18.

[11] KULKARNI, P., DOUGLIS, F., LAVOIE, J., AND TRACEY, J. M. Redundancy elimination within large collections of files. In *ATEC '04: Proceedings of the annual conference on USENIX Annual Technical Conference* (Berkeley, CA, USA, 2004), USENIX Association, pp. 5–5.

[12] LILLIBRIDGE, M., ESHGHI, K., BHAGWAT, D., DEOLALIKAR, V., TREZISE, G., AND CAMBLE, P. Sparse indexing: large scale, inline deduplication using sampling and locality. In *FAST '09: Proccedings of the 7th conference on File and storage technologies* (Berkeley, CA, USA, 2009), USENIX Association, pp. 111–123.

[13] LIU, C., XUE, Y., JU, D., AND WANG, D. A novel optimization method to improve de-duplication storage system performance. In *ICPADS '09: Proceedings of the 2009 15th International Conference on Parallel and Distributed Systems* (Washington, DC, USA, 2009), IEEE Computer Society, pp. 228–235.

[14] NEC. HYDRAstor HS8-3000 Series. `http://bit.ly/hCHxhn`.

[15] NETAPP. Deduplicating Backup Data Streams with the NetApp VTL, 2009. `http://bit.ly/buXcaV`.

[16] QUINLAN, S., AND DORWARD, S. Venti: A new approach to archival storage. In *FAST '02: Proceedings of the Conference on File and Storage Technologies* (Berkeley, CA, USA, 2002), USENIX Association, pp. 89–101.

[17] UNGUREANU, C., ATKIN, B., ARANYA, A., GOKHALE, S., RAGO, S., CALKOWSKI, G., DUBNICKI, C., AND BOHRA, A. Hydrafs: a high-throughput file system for the hydrastor content-addressable storage system. In *FAST'10: Proceedings of the 8th USENIX conference on File and storage technologies* (Berkeley, CA, USA, 2010), USENIX Association, pp. 17–17.

[18] WEI, J., JIANG, H., ZHOU, K., AND FENG, D. Mad2: A scalable high-throughput exact deduplication approach for network backup services. In *MSST'10: Proceedings of the IEEE 26th Symposium on Mass Storage Systems and Technologies* (2010), pp. 1–14.

[19] ZHU, B., LI, K., AND PATTERSON, R. H. Avoiding the disk bottleneck in the datadomain deduplication file system. In *FAST '08: Proceedings of the Conference on File and Storage Technologies* (2008), pp. 269–282.

Electronic Acknowledgement Receipt

EFS ID:	12611925
Application Number:	12947393
International Application Number:	
Confirmation Number:	1801
Title of Invention:	SYSTEM AND METHOD FOR PERFORMING BACKUP OR RESTORE OPERATIONS UTILIZING DIFFERENCE INFORMATION AND TIMELINE STATE INFORMATION
First Named Inventor/Applicant Name:	Philip J. ABERCROMBIE
Customer Number:	23483
Filer:	Michael Yasuhiro Saji/Christina Schroeter
Filer Authorized By:	Michael Yasuhiro Saji
Attorney Docket Number:	2203828.00130US1
Receipt Date:	26-APR-2012
Filing Date:	16-NOV-2010
Time Stamp:	15:22:21
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		130US1_IDS_and_SB08_24April2012.pdf	161534 3848870ecc5db95bf9f909ca56a0185fe75cba2d	yes	2

	Multipart Description/PDF files in .zip description		
	Document Description	**Start**	**End**
	Transmittal Letter	1	1
	Information Disclosure Statement (IDS) Form (SB08)	2	2

Warnings:					
Information:					
2	Non Patent Literature	ISR_PCTUS2011060417_03162012.pdf	178301 6d0028f4531fce9d17efb274aec5f16cc87af9c0	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	339835		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference 2203828.0013WO1	**FOR FURTHER ACTION**	see Form PCT/ISA/220 as well as, where applicable, item 5 below.
International application No. PCT/US2011/060417	International filing date *(day/month/year)* 11 November 2011	(Earliest) Priority Date *(day/month/year)* 16 November 2010
Applicant ACRIFIO, INC.		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 2 sheets.

☐ It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**

   a. With regard to the **language**, the international search was carried out on the basis of:

   ☒ the international application in the language in which it was filed.

   ☐ a translation of the international application into _____ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

   b. ☐ This international search report has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43.6*bis*(a)).

   c. ☐ With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, see Box No. I.

2. ☐ **Certain claims were found unsearchable** (see Box No. II).

3. ☐ **Unity of invention is lacking** (see Box No. III).

4. With regard to the **title**,

   ☒ the text is approved as submitted by the applicant.

   ☐ the text has been established by this Authority to read as follows:

5. With regard to the **abstract**,

   ☒ the text is approved as submitted by the applicant.

   ☐ the text has been established, according to Rule 38.2, by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. With regard to the **drawings**,

   a. the figure of the **drawings** to be published with the abstract is Figure No. 4

   ☒ as suggested by the applicant.

   ☐ as selected by this Authority, because the applicant failed to suggest a figure.

   ☐ as selected by this Authority, because this figure better characterizes the invention.

   b. ☐ none of the figures is to be published with the abstract.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2011/060417

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - G06F 12/00 (2012.01)
USPC - 707/649
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8) - G06F 7/00; 12/00; 12/02; 12/08 (2012.01)
USPC - 707/640, 644-45, 648-54, 674-79, 790-92; 711/100, 161-62, E12.103

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
MicroPatent, Google Patents, ProQuest, Google Scholar

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2002/0049778 A1 (BELL et al) 25 April 2002 (25.04.2002) entire document	1-49
A	US 2008/0034016 A1 (CISLER et al) 07 February 2008 (07.02.2008) entire document	1-49
A	US 6,883,073 B2 (ARAKAWA et al) 19 April 2005 (19.04.2005) entire document	1-49
A	US 2010/0077013 A1 (CLEMENTS et al) 25 March 2010 (25.03.2010) entire document	1-49
A	US 2010/0138827 A1 (FRANK et al) 03 June 2010 (03.06.2010) entire document	1-49

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
09 March 2012	16 MAR 2012
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

PTO/SB/08b (07-09)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Substitute for form 1449/PTO

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 1 of 1

Complete if Known	
Application Number	12/947,393-Conf. #1801
Filing Date	November 16, 2010
First Named Inventor	Philip J. ABERCROMBIE
Art Unit	2189
Examiner Name	R. G. Bragdon
Attorney Docket Number	2203828.00130US1

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
	AA*	US-20020049778	04-25-2002	Bell et al.	
	AB*	US-20080034016	02-07-2008	Cisler et al.	
	AC*	US-20090307251	12-10-2009	Heller et al.	
	AD*	US-20100077013	03-25-2010	CLEMENTS et al.	
	AE*	US-20100088277	04-08-2010	Rao et al.	
	AF*	US-20100138827	06-03-2010	Frank et al.	
	AG*	US-20100276744	11-04-2010	Lee	
	AH*	US-20110179341	07-21-2011	Falls et al.	
	AI*	US-20110307447	12-15-2011	Sabaa et al.	
	AJ*	US-20110307683	12-15-2011	SPACKMAN	
	AK*	US-20120017060	01-19-2012	Kapanipathi et al.	
	AL*	US-6,883,073	04-19-2005	Arakawa et al.	
	AM*	US-7,814,149	10-12-2010	Stringham	

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document Country Code[3]-Number[4]-Kind Code[5] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages Or Relevant Figures Appear	T[6]

NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
	CA	International Search Report issued for PCT/US2011/060417, dated March 16, 2012 (2 pages)	

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. * CITE NO.: Those application(s) which are marked with an single asterisk (*) next to the Cite No. are not supplied (under 37 CFR 1.98(a)(2)(iii)) because that application was filed after June 30, 2003 or is available in the IFW. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

Docket No.: 2203828.00130US1
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	Philip J. Abercrombie et al.	Confirmation No.:	1801
Application No.:	12/947,393	Art Unit:	2189
Filed:	November 16, 2010	Examiner:	R. G. Bragdon
Title:	SYSTEM AND METHOD FOR PERFORMING BACKUP OR RESTORE OPERATIONS UTILIZING DIFFERENCE INFORMATION AND TIMELINE STATE INFORMATION		

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT (SIDS)

Dear Sir:

This Supplemental Information Disclosure Statement is being filed prior to the mailing date of a first Office Action on the merits. No fee is required. Applicants request that the Examiner initial and return a copy of the enclosed Form PTO SB-08 with the next communication.

Applicant believes no fee is due with this response. However, if a fee is due, please charge our Deposit Account No. 08-0219, under Order No. 2203828.00130US1 from which the undersigned is authorized to draw.

Respectfully submitted,

Dated: April 24, 2012

/Michael Saji/
Michael Y. Saji
Registration No.: 66,291
Attorney for Applicant(s)

Wilmer Cutler Pickering Hale and Dorr LLP
60 State Street
Boston, Massachusetts 02109
(617) 526-6000 (telephone)
(617) 526-5000 (facsimile)

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
12/947,393	11/16/2010	Philip J. ABERCROMBIE	2203828.00130US1

CONFIRMATION NO. 1801

23483
WILMERHALE/BOSTON
60 STATE STREET
BOSTON, MA 02109

PUBLICATION NOTICE

OC000000054334234

Title:SYSTEM AND METHOD FOR PERFORMING BACKUP OR RESTORE OPERATIONS UTILIZING DIFFERENCE INFORMATION AND TIMELINE STATE INFORMATION

Publication No.US-2012-0124306-A1
Publication Date:05/17/2012

NOTICE OF PUBLICATION OF APPLICATION

The above-identified application will be electronically published as a patent application publication pursuant to 37 CFR 1.211, et seq. The patent application publication number and publication date are set forth above.

The publication may be accessed through the USPTO's publically available Searchable Databases via the Internet at www.uspto.gov. The direct link to access the publication is currently http://www.uspto.gov/patft/.

The publication process established by the Office does not provide for mailing a copy of the publication to applicant. A copy of the publication may be obtained from the Office upon payment of the appropriate fee set forth in 37 CFR 1.19(a)(1). Orders for copies of patent application publications are handled by the USPTO's Office of Public Records. The Office of Public Records can be reached by telephone at (703) 308-9726 or (800) 972-6382, by facsimile at (703) 305-8759, by mail addressed to the United States Patent and Trademark Office, Office of Public Records, Alexandria, VA 22313-1450 or via the Internet.

In addition, information on the status of the application, including the mailing date of Office actions and the dates of receipt of correspondence filed in the Office, may also be accessed via the Internet through the Patent Electronic Business Center at www.uspto.gov using the public side of the Patent Application Information and Retrieval (PAIR) system. The direct link to access this status information is currently http://pair.uspto.gov/. Prior to publication, such status information is confidential and may only be obtained by applicant using the private side of PAIR.

Further assistance in electronically accessing the publication, or about PAIR, is available by calling the Patent Electronic Business Center at 1-866-217-9197.

Office of Data Managment, Application Assistance Unit (571) 272-4000, or (571) 272-4200, or 1-888-786-0101

Electronic Acknowledgement Receipt

EFS ID:	13213283
Application Number:	12947393
International Application Number:	
Confirmation Number:	1801
Title of Invention:	SYSTEM AND METHOD FOR PERFORMING BACKUP OR RESTORE OPERATIONS UTILIZING DIFFERENCE INFORMATION AND TIMELINE STATE INFORMATION
First Named Inventor/Applicant Name:	Philip J. ABERCROMBIE
Customer Number:	23483
Filer:	Michael Yasuhiro Saji/Kathleen Bastarache
Filer Authorized By:	Michael Yasuhiro Saji
Attorney Docket Number:	2203828.00130US1
Receipt Date:	10-JUL-2012
Filing Date:	16-NOV-2010
Time Stamp:	16:02:41
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Transmittal Letter	2203828_130US1_IDS_10July2012.pdf	62959 6ff2bec3c1496d5af027bfb5293fe13b39f47ae3	no	1

Warnings:

Information:

ACT6002305

2	Information Disclosure Statement (IDS) Form (SB08)	2203828_130US1_SB08_10July 2012.pdf	96454 e13d61fb49595fb0c7d27aaa6746110ed914c75e	no	1

Warnings:

Information:

This is not an USPTO supplied IDS fillable form

Total Files Size (in bytes):	159413

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/08b (07-09)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Substitute for form 1449/PTO

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 1 of 1

Complete if Known	
Application Number	12/947,393-Conf. #1801
Filing Date	November 16, 2010
First Named Inventor	Philip J. ABERCROMBIE
Art Unit	2189
Examiner Name	G. Bansal
Attorney Docket Number	2203828.00130US1

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
	AA*	US-20110252198	10-13-2011	Ogasawara et al.	

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document Country Code[3] -Number[4]-Kind Code[5] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages Or Relevant Figures Appear	T[6]

NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. * CITE NO.: Those application(s) which are marked with an single asterisk (*) next to the Cite No. are not supplied (under 37 CFR 1.98(a)(2)(iii)) because that application was filed after June 30, 2003 or is available in the IFW. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

Docket No.: 2203828.00130US1
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	Philip J. Abercrombie et al.	Confirmation No.:	1801
Application No.:	12/947,393	Art Unit:	2189
Filed:	November 16, 2010	Examiner:	G. Bansal
Title:	SYSTEM AND METHOD FOR PERFORMING BACKUP OR RESTORE OPERATIONS UTILIZING DIFFERENCE INFORMATION AND TIMELINE STATE INFORMATION		

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT (IDS)

Dear Sir:

This Supplemental Information Disclosure Statement is being filed prior to the mailing date of a first Office Action on the merits. No fee is required.

Applicants request that the Examiner initial and return a copy of the enclosed Form PTO SB-08 with the next communication.

Applicant believes no fee is due with this response. However, if a fee is due, please charge our Deposit Account No. 08-0219, under Order No. 2203828.00130US1 from which the undersigned is authorized to draw.

Respectfully submitted,

Dated: July 10, 2012

/Michael Saji/
Michael Y. Saji
Registration No.: 66,291
Attorney for Applicant(s)

Wilmer Cutler Pickering Hale and Dorr LLP
60 State Street
Boston, Massachusetts 02109
(617) 526-6000 (telephone)
(617) 526-5000 (facsimile)

PTO/SB/20PCT-US (09-10)
Approved for use through 01/31/2012. OMB 0651-0058
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

REQUEST FOR PARTICIPATION IN THE PATENT COOPERATION TREATY – PATENT PROSECUTION HIGHWAY (PCT-PPH) PILOT PROGRAM IN A U.S. APPLICATION WHERE THE USPTO WAS THE ISA OR IPEA

Application No.:	12/947,393-Conf. #1801	Filing Date:	November 16, 2010
First Named Inventor:	Philip J. ABERCROMBIE		
Title of the Invention:	SYSTEM AND METHOD FOR PERFORMING BACKUP OR RESTORE OPERATIONS UTILIZING DIFFERENCE INFORMATION AND TIMELINE STATE INFORMATION		

THIS REQUEST FOR PARTICIPATION IN THE PCT-PPH PILOT PROGRAM ALONG WITH THE REQUIRED DOCUMENTS MUST BE SUBMITTED VIA EFS-WEB. INFORMATION REGARDING EFS-WEB IS AVAILABLE AT HTTP://WWW.USPTO.GOV/EBC/EFS_HELP.HTML.

APPLICANT HEREBY REQUESTS PARTICIPATION IN THE PATENT PCT-PPH PROGRAM AND PETITIONS TO MAKE THE ABOVE-IDENTIFIED APPLICATION SPECIAL UNDER THE PCT-PPH PROGRAM.

The above-identified application is (1) a national stage entry of the corresponding PCT application, or (2) a national stage entry of another PCT application which claims priority to the corresponding PCT application, or (3) a national application that claims domestic/ foreign priority to the corresponding PCT application, or (4) a national application which forms the basis for the priority claim in the corresponding PCT application, or (5) a continuing application of a U.S. application that satisfies one of (1) to (4) above, or (6) a U.S. application that claims domestic benefit to a U.S. provisional application which forms the basis for the priority claim in the corresponding PCT application.

The corresponding PCT application number(s) is/are: PCT/US11/60417

The international date of the corresponding PCT application(s) is/are: November 11, 2011

I. List of Required Documents:

a. **A copy of the latest international work product (WO/ISA, WO/IPEA, or IPER) in the above-identified corresponding PCT application(s)**

☐ Is attached.

☒ Is not attached because the document is already in the U.S. application.

b. **A copy of all claims which were indicated as having novelty, inventive step and industrial applicability in the above-identified corresponding PCT application(s).**

☐ Is attached.

☒ Is not attached because the document is already in the U.S. application.

c. **English translations of the documents in a. and b. are attached (if the documents are not in the English language). A statement that the English translation is accurate is attached for the document in b. above.**

d. **(1) An information disclosure statement listing the documents cited in the international work products (ISR, WO/ISA, WO/IPEA, IPER) of the corresponding PCT application.**

☐ Is attached.

☒ Has already been filed in the above-identified U.S. application on April 26, 2012

(2) Copies of all documents (except) for U.S. patents or U.S. patent application publications)

☐ Are attached.

☒ Have already been filed in the above-identified U.S. application on April 26, 2012

100398398

PTO/SB/20PCT-US (09-10)
Approved for use through 01/31/2012. OMB 0651-0058
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

REQUEST FOR PARTICIPATION IN THE PCT-PPH PILOT PROGRAM BETWEEN THE EPO AND THE USPTO
(continued)

Application No.:	12/947,393-Conf. #1801
First Named Inventor:	Philip J. ABERCROMBIE

II. Claims Correspondence Table:

Claims in US Application	Patentable Claims in the corresponding PCT Application	Explanation regarding the correspondence
1-8	15-22	Claims are identical

III. All the claims in the US application sufficiently correspond to the patentable claims in the corresponding PCT application.

Signature	/Michael Saji/	Date	August 24, 2012
Name (Print/Typed)	Michael Y. Saji	Registration Number	66,291

Electronic Acknowledgement Receipt

EFS ID:	13583396
Application Number:	12947393
International Application Number:	
Confirmation Number:	1801
Title of Invention:	SYSTEM AND METHOD FOR PERFORMING BACKUP OR RESTORE OPERATIONS UTILIZING DIFFERENCE INFORMATION AND TIMELINE STATE INFORMATION
First Named Inventor/Applicant Name:	Philip J. ABERCROMBIE
Customer Number:	23483
Filer:	Michael Yasuhiro Saji/Kim LaRocca
Filer Authorized By:	Michael Yasuhiro Saji
Attorney Docket Number:	2203828.00130US1
Receipt Date:	24-AUG-2012
Filing Date:	16-NOV-2010
Time Stamp:	16:09:01
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Petition to make special under Patent Prosecution Hwy	2203828_00130US1_PPH_Request.PDF	73623 0fdbc66f916954653295e244c2adee3ecd7d224f	no	2

Warnings:	
Information:	

ACT6002311

Total Files Size (in bytes):	73623

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

MAILED

SEP 21 2012

PCT LEGAL ADMINISTRATION

UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

WILMERHALE/BOSTON
60 STATE STREET
BOSTON MA 02109

In re Application of:	:	
ABERCROMBIE, Philip, J., et al.	:	**DECISION ON REQUEST TO**
Application No.: 12/947,393	:	**PARTICIPATE IN THE PATENT**
Filing Date: 16 November 2010	:	**PROSECUTION HIGHWAY**
Attorney's Docket No.: 2203828.00130US1	:	**AND PETITION TO MAKE**
For: SYSTEM AND METHOD FOR	:	**SPECIAL UNDER 37 CFR**
PERFORMING BACKUP OR	:	**1.102(a)**
RESTORE OPERATIONS ...	:	

This is a decision on the request to participate in the PCT Patent Prosecution Highway (PCT-PPH) pilot program and the petition under 37 CFR 1.102(a), filed on 24 August 2012, to make the above-identified application special.

The request and petition are **DISMISSED**.

DISCUSSION

A grantable request to participate in the PCT-PPH pilot program and petition to make special require:

(1) The U.S. application must have an eligible relationship to one or more PCT applications where the ISA or IPEA are the JPO, EPO, KIPO, IPAU, Russia, Spain, Finland, Austria, or USPTO;

(2) At least one claim in the PCT application has novelty, inventive step, and industrial applicability and must be free of any observations in Box VIII in the latest work product in the international stage or applicant must identify and explain why the claim(s) is/are not subject to the observation in Box VIII;

(3) Applicant must submit a copy of the claim(s) from the PCT application(s) that have novelty, inventive step, and industrial applicability along with an English translation thereof and a statement that the English translation is accurate, if the claims are not in the English language;

(4) All the claims in the U.S. application must sufficiently correspond or be amended to sufficiently correspond to the claim(s) that have novelty, inventive step, and industrial applicability in the PCT application(s);

(5) Examination of the U.S. application has not begun;

(6) Applicant must submit a copy of the latest international work product from the PCT application indicating that the claim(s) have novelty, inventive step, and industrial applicability along with an English translation thereof.

(7) Applicant must submit an IDS listing the documents cited by the PCT examiner in the international work product along with copies of documents except U.S. patents or U.S. patent application publications.

(8) Applicant is required to submit a claims correspondence table in English which indicates how all the claims in the U.S. application correspond to the claims indicated as having novelty, inventive step and industrial applicability in the latest international work product.

Requirements (1), (2), (5), and (7-8) above are considered to have been met. However, the request to participate in the PPH pilot program and petition fails to meet requirements (3), (4), and (6).

Regarding the requirement of condition (3), applicant has failed to submit a copy of the claims from the PCT (the petition asserts that the claims are already in the present U.S. application; however, there is no document in the present application file identified as the claims from PCT/US2011/060417).

Regarding the requirement of condition (4), it cannot be determined if requirement (4) has been met since requirement (3) has not been met. Applicant is responsible for ensuring that the claims in the U.S. application sufficiently correspond to the claims that have novelty, inventive step, and industrial applicability in the PCT application.

Regarding the requirement of condition (6), applicant has failed to submit a copy of the latest international work product indicating that claims in the PCT have novelty, inventive step, and industrial applicability (the petition asserts that the latest international work product is already in the U.S. application; however, the present application file contains only the International Search Report from the international application; a copy of the Written Opinion of the International Searching Authority has not been submitted).

Applicant is given **ONE** opportunity within a time period of **ONE MONTH or THIRTY DAYS**, whichever is longer, from the mailing date of this decision to correct the deficiencies. **NO EXTENSION OF TIME UNDER 37 CFR 1.136 IS PERMITTED.** If the deficiencies are not corrected with the time period given, the application will await action in its regular turn.

Response must be filed via the Electronic Filing System (EFS) using the document description: Petition to make special under PCT – Patent Pros Hwy. Any preliminary amendments and IDS submitted with the PPH documents must be separately indexed as a preliminary amendment and IDS, respectively.

Telephone inquiries concerning this decision should be directed to the undersigned.

All other inquiries concerning the examination or status of the application is accessible in the PAIR system at http://www.uspto.gov/ebc/index.html.

/RichardMRoss/

Richard M. Ross
Attorney Advisor
Office of PCT Legal Administration
Tel.: (571) 272-3296

Docket No.: 2203828.00130US1
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	Philip J. Abercrombie et al.	Confirmation No.:	1801
Application No.:	12/947,393	Art Unit:	2189
Filed:	November 16, 2010	Examiner:	G. Bansal
Title:	SYSTEM AND METHOD FOR MANAGING DATA WITH SERVICE LEVEL AGREEMENTS THAT MAY SPECIFY NON-UNIFORM COPYING OF DATA		

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

PETITION TO MAKE SPECIAL UNDER 37 C.F.R. § 1.102(a)

Dear Sir or Madam:

INTRODUCTORY COMMENTS

In response to the Decision on Request to Participate in the Patent Prosecution Highway and Petition to Make Special Under 37 C.F.R. § 1.102(a), dated September 21, 2012, applicants request reconsideration of the Petition in light of the Remarks and supporting documents found herein.

Remarks begin on page 2 of this paper.

A **copy of the claims from the PCT application** is submitted herewith.

A **copy of the latest international work product from the PCT application** is also submitted herewith.

REMARKS

In response to the Decision on Request to Participate in the Patent Prosecution Highway and Petition to Make Special Under 37 C.F.R. § 1.102(a), dated September 21, 2012, Applicants thank the Office for review of applicants' Petition, filed August 24, 2012. The Office has indicated that the application failed to meet the requirements numbered (3), (4), and (6) in the Decision.

With respect to requirement (3), applicants submit herewith a copy of the claims from the PCT application.

With respect to requirement (4), applicants refer the Office to the correspondence chart filed with the initial Petition of August 24, 2012, for evidence that all of the U.S. claims correspond to the PCT claims.

With respect to requirement (6), applicants submit herewith a copy of the latest international work product from the PCT application. Applicants note that the included document includes a Written Opinion of the International Searching Authority dated March 16, 2012.

In view of the documents submitted herewith, applicants request reconsideration of the Petition. Applicant believes no fees are due with this response. However, if a fee is due, please charge our Deposit Account No. 08-0219, under Order No. 2203828.00131US1 from which the undersigned is authorized to draw.

Respectfully submitted,

Dated: October 10, 2012

/Michael Saji/
Michael Y. Saji
Registration No.: 66,291
Attorney for Applicant(s)

Wilmer Cutler Pickering Hale and Dorr LLP
60 State Street
Boston, Massachusetts 02109
(617) 526-6000 (telephone)
(617) 526-5000 (facsimile)

1. A system for performing a plurality of prescribed data management functions in a manner that reduces redundant access operations to primary storage, said system comprising:

a data management engine for performing data management functions, including at least a snapshot function operable to create a point-in-time image of primary storage data to secondary storage, and at least one back-up function operable to create at least one back-up copy of data, said data management engine being responsive to an electronic service level agreement (SLA) that specifies a schedule for performing data management functions,

wherein point-in-time images of data include a reference to a complete baseline image of data at a specific point in time and difference data indicating changes to the data at a later, specific point in time, and

wherein, in response to the schedule requiring at least some data management functions to be performed concurrently, the data management engine creates a point-in-time image of the primary storage data and communicates the difference information of that point-in-time image to the secondary storage to update the at least one of a back-up copy of the primary data, such that the primary storage is accessed only once for all corresponding sets of updates to the secondary storage.

2. The system of claim 1, wherein the point-in-time image of primary storage data at secondary storage is stored on performance optimized secondary storage.

3. The system of claim 1, wherein said back-up copy of the point-in-time image of primary storage data is stored on remote storage.

4. The system of claim 1, wherein said back-up copy of the point-in-time image of primary storage data is stored at capacity optimized storage.

5. The system of claim 4, wherein said back-up copy of the point-in-time image of primary storage data is stored as a deduplicated image on capacity optimized storage.

6. The system of claim 1, wherein difference data includes bitmap information with each bit of the bitmap corresponding to a portion of primary storage data, and including new data for those portions of the bitmap which are set to indicate that data has changed.

7. The system of claim 1, wherein difference data includes extent information.

8. The system of claim 1, wherein the data management engine includes logic to invoke the primary storage to provide a point-in-time image of data and includes logic to retrieve the point-in-time image from the primary storage.

9. A system for managing data in accordance with service level agreements (SLAs) that specify schedules on a calendar basis for performing prescribed data management functions and for reducing inter-function redundancy, said system comprising:

a data management engine for performing data management functions, including at least a snapshot function, and at least one back-up function, said data management engine including a service level policy engine that receives SLAs in electronic form and which controls the scheduling of the data management functions in accordance therewith,

wherein each electronic SLA is associated with a corresponding application that uses data, and wherein each SLA specifies at least one service level policy, each policy specifying a source pool for data, a destination pool where a copy should be made of the source pool data, copy frequency indicating the frequency of operation for that policy, retention period indicating how long a given copy should be retained before being allowed to expire, and schedule information indicating hours and days of operation when the policy is in place, such that a collection of policies within a SLA is capable of expressing a non-uniform schedule for when a given function should be performed and is capable of expressing multiple data management functions that should be performed on a given source of data, and

wherein said data management engine is operable to perform preparatory operations with the application and with the source pool so that the source pool of data has a coherent image of data to be copied and wherein said preparatory operations are

performed once even if the SLA specifies multiple data management functions to be performed on that source pool at the current time.

10. The system of claim 9, wherein if two or more copy operations are scheduled to occur at the same instant between the same source pool and destination pool, only one of the two or more copy operations is performed by the data management engine and that copy is associated with the longest retention time corresponding to the two or more scheduled copy operations.

11. The system of claim 9, wherein preparatory operations include the data management engine collecting metadata about the application to store in conjunction with application data.

12. The system of claim 9, wherein preparatory operations include application quiescing operations.

13. The system of claim 12, wherein the application quiescing operations include freezing the application from further updating application data.

14. The system of claim 12, wherein the application quiescing operations include flushing the I/O cache of the application server of application data.

15. A system for backing-up data from a first storage pool to a second storage pool using difference information between time states, said system comprising:

a data management engine for performing data management functions, including at least a back-up function to create a back-up copy of data,

said data management engine operable to execute a sequence of snapshot operations to create point-in-time images of application data on a first storage pool, each successive point-in-time image corresponding to a specific, successive time-state of the application data, and each snapshot operation creating difference information indicating which application data has changed and the content of the changed application data for the corresponding time state;

said data management engine operable to execute at least one back-up function for the application data wherein the backup operation is scheduled for execution at non-consecutive time-states

wherein said data management engine is operable to maintain history information having time-state information indicating the time-state of the last back-up function performed on the application data for a corresponding back-up copy of data; and

wherein the data management engine is operable to create composite difference information from the difference information for each time-state between the time-state of the last back-up function performed on the application data and the time-state of the currently-scheduled back-up function to be performed on the application data, and wherein the data management engine is operable to send the composite difference information to a second storage pool to be compiled with the back-up copy of data at the last time-state to create a back-up copy of data for the current time-state.

16. The system of claim 15, wherein difference information includes bitmap information with each bit of the bitmap corresponding to a portion of primary storage data, and including new data for those portions of the bitmap whisch are set to indicate that data has changed.

17. The system of claim 15, wherein difference information includes extent information.

18. The system of claim 15, wherein multiple back-up functions are scheduled to occur simultaneously, each with different gaps of non-consecutive time-states, and each with different composite difference information generated corresponding to the different gaps

19. A system for restoring data in a storage pool from a back-up copy of the data using difference information between time states, said system comprising:

a data management engine wherein said data management engine is operable to maintain history information indicating the time-states for which storage pools have point-in-time images of application data; and

wherein said data management engine includes logic for restoring application data in a storage pool to a point-in-time image of the data for a specified time-state;

said data management engine operable to identify the existence of a point-in-time image of the data at the storage pool for a time-state prior to the specified time-state and sending difference information from the back-up copy of data to the storage pool, said difference information indicating which application data has changed and the content of the changed application data for the time between the specified time state and the time state prior to the specified time-state.

20. The system of claim 19, wherein difference information includes bitmap information with each bit of the bitmap corresponding to a portion of primary storage data, and includes back-up data for those portions of the bitmap which are set to indicate that data has changed.

21. The system of claim 19, wherein difference information includes extent information.

22. The system of claim 19, wherein the prior time-state and the specified time-state are non-consecutive time-states.

23. A method of forming deduplicated images of a data object that changes over time using difference information between temporal states of the data object, said method comprising:

organizing the content of the data object for a first temporal state as a plurality of content segments and storing the content segments in a data store;

creating an organized arrangement of hash structures to represent the data object in its first temporal state wherein for a subset of the hash structures, each structure includes a hash signature for a corresponding content segment and is associated with a reference to the corresponding content segment, and wherein the logical organization of the arrangement represents the organization of the content segments as they are represented within the data object;

receiving difference information for the data object, said difference information indicating the changed content for the data object for a second temporal state relative to

the first temporal state, and said difference information indicating the location of the changed content within the data object;

forming at least one hash signature for the changed content;

storing the changed content that is unique in the data store as content segments

modifying the organized arrangement of hash structures to incorporate new structures for the at least one hash signature for the changed content, incorporating the new structures in the organized arrangement of structures at a position corresponding to the location of the changed content within the data object as indicated within said difference information, and associating the hash signatures for the new structures with references to the corresponding content segments for the changed content; and

associating the new structures with the second temporal state, whereby a deduplicated image of the data object for a second temporal state is stored without requiring reception of a complete image of the data object for the second temporal state.

24. The method of claim 23, wherein after forming the at least one hash signature for changed content, the formed signature is compared to at least one hash signature in the organized arrangement of hash structures to determine if the formed structure already exists in the organized arrangement.

25. The method of claim 23, wherein the comparison first occurs with a hash structure in the organized arrangement at a position corresponding to the location of the changed content as indicated in the difference information.

26. The method of claim 23, wherein the organized arrangement of hash structures is an organized tree structure.

27. The method of claim 23, wherein an organized arrangement of temporal structures are maintained, each temporal structure associated with a time state and each including information indicative of the hash structures corresponding to the associated time state.

28. A method of managing deduplicated images of data objects that change over time, said method comprising:

organizing unique content of each data object as a plurality of content segments and storing the content segments in a data store;

for each data object, creating an organized arrangement of hash structures, wherein each structure, for a subset of the hash structures, includes a hash signature for a corresponding content segment and is associated with a reference to the corresponding content segment, wherein the logical organization of the arrangement represents the logical organization of the content segments as they are represented within the data object, and wherein another subset of the hash structures includes a hierarchy of hash signatures of said hash signatures for corresponding content segments so that the organized arrangement may be traversed to determine if content is represented by said organized arrangement of hash structures; and

for each data object, maintaining an organized arrangement of temporal structures to represent a corresponding data object over time, wherein each structure is associated with a temporal state of the data object and wherein the logical arrangement of structures is indicative of the changing temporal states of the data object, and wherein each temporal state is associated with the hash structures representing the content of the data object during that temporal state.

29. The method of claim 28, wherein the temporal structures for a data object at a given temporal state is associated with hash structures for data object content that has changed relative to a prior temporal state of the data object.

30. The method of claim 28, wherein the hash structures for the changed data content is organized as a graph separate from the organized arrangement of hash structures for a prior temporal state.

31. The method of claim 28, wherein differences in content of a data object from one temporal state to another is determined by referencing the organized arrangement of temporal structures for the other temporal state and for all other temporal states intervening between the other temporal state and the one state so that differences may be determined over multiple temporal states.

32. A method of storing deduplicated images in which a portion of the image is stored in encoded form directly in a hash table, the method comprising:

organizing unique content of each data object as a plurality of content segments and storing the content segments in a data store;

for each data object, creating an organized arrangement of hash structures, wherein each structure, for a subset of the hash structures, includes a field to contain a hash signature for a corresponding content segment and is associated with a reference to the corresponding content segment, wherein the logical organization of the arrangement represents the logical organization of the content segments as they are represented within the data object;

receiving content to be included in the deduplicated image of the data object;

determining if the received content may be encoded using a predefined non-lossy encoding technique and in which the encoded value would fit within the field for containing a hash signature;

if so, placing the encoding in the field and marking the hash structure to indicate that the field contains encoded content for the deduplicated image;

if not, generating a hash signature for the received content and placing the hash signature in the field and placing the received content in a corresponding content segment in said data store if it is unique.

33. The method of claim 32, wherein said non-lossy encoding is a run-length encoding.

34. The method of claim 32, wherein hash signatures for each data object are created using the SHA-1 cryptographic hash function.

35. The method of claim 32, further comprising subsequently reconstituting the content from the encoded content.

36. A method for using a first deduplicating store to update a second deduplicating store with information representing how data objects change over time, said method comprising:

at a first deduplicating store, organizing unique content of each data object as a plurality of content segments and storing the content segments in a data store;

at a first deduplicating store, for each data object, creating an organized arrangement of hash structures, wherein each structure, for a subset of the hash structures, includes a hash signature for a corresponding content segment and is associated with a reference to the corresponding content segment, wherein the logical organization of the arrangement represents the logical organization of the content segments as they are represented within the data object;

at a first deduplicating store, for each data object, maintaining an organized arrangement of temporal structures to represent a corresponding data object over time, wherein each structure is associated with a temporal state of the data object and wherein the logical arrangement of structures is indicative of the changing temporal states of the data object, and wherein each temporal state is associated with the hash structures representing the content of the data object that has changed relative to a prior temporal state;

at a second dedeuplicating store, organizing unique content of each data object as a plurality of content segments and storing the content segments in a data store;

at a second deduplicating store, for each data object, maintaining an organized arrangement of hash structures, that is at least a subset of the hash structures at said first deduplicating store;

at a second deduplicating store, for each data object, maintaining an organized arrangement of temporal structures to represent a corresponding data object over time, wherein said organized arrangement of temporal structures is at least a subset of the temporal structures at said first deduplicating store, representing a subset of the temporal states;

in response to a request to update the second deduplicating store with information from said first deduplicating store, finding a temporal state that is common to the first and second deduplicating stores and that is in temporal proximity to the current state of the second deduplicating store; and

compiling a set of hash signatures for the content that has changed from the common state to the current temporal state of the first deduplicating store and sending that set of hash signatures to the second deduplicating store so it can update its organized arrangement of hash structures to represent the content of the data object up to the current temporal state of the first deduplicating store.

37. The method of claim 36, further including maintaining a history of the hash signatures that each deduplicating store contains, and for hash signatures in the set of hash signatures that are new to the second deduplicating store, sending the corresponding content segments from the first deduplicating store so that the second deduplicating store may update its data store with the new content.

38. The method of claim 36, wherein the temporal state that is in proximity to the current state of the second deduplicating store is a nearest neighbor state of the current state.

39. The method of claim 36, wherein the temporal state that is in proximity to the current state of the second deduplicating store is an ancestor state of the current state.

40. The method of claim 36, wherein the temporal state that is in proximity to the current state of the second deduplicating store is a child state of the current state.

41. The method of claim 38, wherein the nearest neighbor state is the state which is connected by a set of edges whose sum is lower than the sum of any other set of edges.

42. The method of claim 36, wherein the logical arrangement of structures contains branching.

43. The method of claim 37, further comprising recording, at the current state, to what states the content segments corresponding to the current temporal state has been sent.

44. A method of performing garbage collection to identify content segments no longer referenced in a deduplicating storage system in which redundant mark operations in a mark-and-sweep technique are avoided, the method comprising:

organizing unique content of each data object as a plurality of content segments in the deduplicating storage system;

for each data object, creating an organized arrangement of hash structures, wherein each structure, for a subset of the hash structures, includes a hash signature for a corresponding content segment and is associated with a reference to the corresponding content segment, wherein the logical organization of the arrangement represents the logical organization of the content segments as they are represented within the data object, and wherein another subset of the hash structures includes a hierarchy of hash signatures of said hash signatures for corresponding content segments so that the organized arrangement may be traversed to determine if content is represented by said organized arrangement of hash structures;

for each data object, maintaining an organized arrangement of temporal structures to represent a corresponding data object over time, wherein each structure is associated with a temporal state of the data object and wherein the logical arrangement of structures is indicative of the changing temporal states of the data object, and wherein each temporal state is associated with the hash structures representing the content of the data object that has changed relative to an immediately prior temporal state of the data object;

for each content segment in the deduplicating storage system, clearing its corresponding garbage collection state;

iterating over the temporal structures and, for each temporal structure, marking the garbage collection state for the associated content segments for only the content segments that have changed relative to an immediately prior temporal state of the data object; and

returning any content segments to a free pool of storage that has a cleared garbage collection state after the iteration step.

45. The method of claim 44, further comprising iterating over the temporal structures using a depth first search.

46. The method of claim 44, further comprising repeating the method at periodic intervals.

47. The method of claim 44, further comprising performing the method subsequent to the addition of a new temporal state of a data object.

48. The method of claim 44, further comprising performing the method subsequent to the removal of a temporal state of a data object.

49. The method of claim 44, further comprising maintaining a global reference list of all content segments that have been allocated in the deduplicated storage system.

PATENT COOPERATION TREATY

From the INTERNATIONAL SEARCHING AUTHORITY

To: PETER DICHIARA
WILMER CUTLER PICKERING HALE AND DORR LLP
60 STATE STREET
BOSTON, MA 02109

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL SEARCH REPORT AND THE WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY, OR THE DECLARATION

(PCT Rule 44.1)

Applicant's or agent's file reference: 2203828.0013WO1	Date of mailing *(day/month/year)*
	FOR FURTHER ACTION See paragraphs 1 and 4 below
International application No. PCT/US2011/060417	International filing date *(day/month/year)* 11 November 2011
Applicant ACRIFIO, INC.	

1. ☒ The applicant is hereby notified that the international search report and the written opinion of the International Searching Authority have been established and are transmitted herewith.

   Filing of amendments and statement under Article 19:
   The applicant is entitled, if he so wishes, to amend the claims of the international application (see Rule 46):

   When? The time limit for filing such amendments is normally two months from the date of transmittal of the international search report.

   Where? Directly to the International Bureau of WIPO, 34 chemin des Colombettes 1211 Geneva 20, Switzerland, Facsimile No.: +41 22 338 82 70

   For more detailed instructions, see *PCT Applicant's Guide*, International Phase, paragraphs 9.004 – 9.011.

2. ☐ The applicant is hereby notified that no international search report will be established and that the declaration under Article 17(2)(a) to that effect and the written opinion of the International Searching Authority are transmitted herewith.

3. ☐ **With regard to any protest** against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

   ☐ the protest together with the decision thereon has been transmitted to the International Bureau together with any request to forward the texts of both the protest and the decision thereon to the designated Offices.

   ☐ no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

4. **Reminders**

   The applicant may submit comments on an informal basis on the written opinion of the International Searching Authority to the International Bureau. The International Bureau will send a copy of such comments to all designated Offices unless an international preliminary examination report has been or is to be established. Following the expiration of 30 months from the priority date, these comments will also be made available to the public.

   Shortly after the expiration of **18 months** from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau before the completion of the technical preparations for international publication (Rules 90*bis*.1 and 90*bis*.3).

   Within **19 months** from the priority date, but only in respect of some designated Offices, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase **until 30 months** from the priority date (in some Offices even later); otherwise, the applicant must, **within 20 months** from the priority date, perform the prescribed acts for entry into the national phase before those designated Offices.

   In respect of other designated Offices, the time limit of **30 months** (or later) will apply even if no demand is filed within 19 months.

   For details about the applicable time limits, Office by Office, see www.wipo.int/pct/en/texts/time_limits.html and the *PCT Applicant's Guide*, National Chapters.

Name and mailing address of the ISA/ Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450	Authorized officer Blaine R. Copenheaver
Facsimile No. 571-273-3201	Telephone No. PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/220 (July 2010)

PATENT COOPERATION TREATY

From the INTERNATIONAL SEARCHING AUTHORITY

To: PETER DICHIARA
WILMER CUTLER PICKERING HALE AND DORR LLP
60 STATE STREET
BOSTON, MA 02109

PCT

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL SEARCH REPORT AND THE WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY, OR THE DECLARATION

(PCT Rule 44.1)

Date of mailing *(day/month/year)*	16 MAR 2012
Applicant's or agent's file reference: 2203828.0013WO1	**FOR FURTHER ACTION** See paragraphs 1 and 4 below
International application No.: PCT/US2011/060417	International filing date *(day/month/year)*: 11 November 2011
Applicant: ACRIFIO, INC.	

1. ☒ The applicant is hereby notified that the international search report and the written opinion of the International Searching Authority have been established and are transmitted herewith.

   Filing of amendments and statement under Article 19:
   The applicant is entitled, if he so wishes, to amend the claims of the international application (see Rule 46):

   When? The time limit for filing such amendments is normally two months from the date of transmittal of the international search report.

   Where? Directly to the International Bureau of WIPO, 34 chemin des Colombettes 1211 Geneva 20, Switzerland, Facsimile No.: +41 22 338 82 70

   For more detailed instructions, see *PCT Applicant's Guide*, International Phase, paragraphs 9.004 – 9.011.

2. ☐ The applicant is hereby notified that no international search report will be established and that the declaration under Article 17(2)(a) to that effect and the written opinion of the International Searching Authority are transmitted herewith

3. ☐ **With regard to any protest** against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

   ☐ the protest together with the decision thereon has been transmitted to the International Bureau together with any request to forward the texts of both the protest and the decision thereon to the designated Offices.

   ☐ no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

4. **Reminders**

   The applicant may submit comments on an informal basis on the written opinion of the International Searching Authority to the International Bureau. The International Bureau will send a copy of such comments to all designated Offices unless an international preliminary examination report has been or is to be established. Following the expiration of 30 months from the priority date, these comments will also be made available to the public.

   Shortly after the expiration of **18 months** from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau before the completion of the technical preparations for international publication (Rules 90*bis*.1 and 90*bis*.3).

   Within **19 months** from the priority date, but only in respect of some designated Offices, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase **until 30 months** from the priority date (in some Offices even later); otherwise, the applicant must, **within 20 months** from the priority date, perform the prescribed acts for entry into the national phase before those designated Offices.

   In respect of other designated Offices, the time limit of **30 months** (or later) will apply even if no demand is filed within 19 months.

   For details about the applicable time limits, Office by Office, see www.wipo.int/pct/en/texts/time_limits.html and the *PCT Applicant's Guide*, National Chapters.

Name and mailing address of the ISA/ Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 Telephone No. PCT OSP: 571-272-7774

Form PCT/ISA/220 (July 2010)

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference 2203828.0013WO1	FOR FURTHER ACTION	see Form PCT/ISA/220 as well as, where applicable, item 5 below.
International application No. PCT/US2011/060417	International filing date *(day/month/year)* 11 November 2011	(Earliest) Priority Date *(day/month/year)* 16 November 2010
Applicant ACRIFIO, INC.		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 2 sheets.

☐ It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**

a. With regard to the **language**, the international search was carried out on the basis of:

☒ the international application in the language in which it was filed.

☐ a translation of the international application into _____ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

b. ☐ This international search report has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43.6*bis*(a)).

c. ☐ With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, see Box No. I.

2. ☐ **Certain claims were found unsearchable** (see Box No. II).

3. ☐ **Unity of invention is lacking** (see Box No. III).

4. With regard to the **title**,

☒ the text is approved as submitted by the applicant.

☐ the text has been established by this Authority to read as follows:

5. With regard to the **abstract**,

☒ the text is approved as submitted by the applicant.

☐ the text has been established, according to Rule 38.2, by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. With regard to the **drawings**,

a. the figure of the **drawings** to be published with the abstract is Figure No. 4

☒ as suggested by the applicant.

☐ as selected by this Authority, because the applicant failed to suggest a figure.

☐ as selected by this Authority, because this figure better characterizes the invention.

b. ☐ none of the figures is to be published with the abstract.

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - G06F 12/00 (2012.01)
USPC - 707/649
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8) - G06F 7/00; 12/00; 12/02; 12/08 (2012.01)
USPC - 707/640, 644-45, 648-54, 674-79, 790-92; 711/100, 161-62, E12.103

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
MicroPatent, Google Patents, ProQuest, Google Scholar

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2002/0049778 A1 (BELL et al) 25 April 2002 (25.04.2002) entire document	1-49
A	US 2008/0034016 A1 (CISLER et al) 07 February 2008 (07.02.2008) entire document	1-49
A	US 6,883,073 B2 (ARAKAWA et al) 19 April 2005 (19.04.2005) entire document	1-49
A	US 2010/0077013 A1 (CLEMENTS et al) 25 March 2010 (25.03.2010) entire document	1-49
A	US 2010/0138827 A1 (FRANK et al) 03 June 2010 (03.06.2010) entire document	1-49

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
09 March 2012	16 MAR 2012
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To: PETER DICHIARA
WILMER CUTLER PICKERING HALE AND
DORR LLP
60 STATE STREET
BOSTON, MA 02109

PCT

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43*bis*.1)

	Date of mailing *(day/month/year)* 16 MAR 2012
Applicant's or agent's file reference 2203828.0013WO1	FOR FURTHER ACTION See paragraph 2 below

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US2011/060417	11 November 2011	16 November 2010

International Patent Classification (IPC) or both national classification and IPC
IPC(8) - G06F 12/00 (2012.01)
USPC - 707/649

Applicant ACRIFIO, INC.

1. This opinion contains indications relating to the following items:

- ☒ Box No. I Basis of the opinion
- ☐ Box No. II Priority
- ☐ Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- ☐ Box No. IV Lack of unity of invention
- ☒ Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- ☐ Box No. VI Certain documents cited
- ☐ Box No. VII Certain defects in the international application
- ☐ Box No. VIII Certain observations on the international application

2. FURTHER ACTION

If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1*bis*(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220

Name and mailing address of the ISA/US	Date of completion of this opinion	Authorized officer:
Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	09 March 2012	Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/237 (cover sheet) (July 2011)

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US2011/060417

Box No. I Basis of this opinion

1. With regard to the **language**, this opinion has been established on the basis of:

☒ the international application in the language in which it was filed.

☐ a translation of the international application into _____ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

2. ☐ This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43*bis*.1(a))

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of a sequence listing filed or furnished:

a. (means)

☐ on paper

☐ in electronic form

b. (time)

☐ in the international application as filed

☐ together with the international application in electronic form

☐ subsequently to this Authority for the purposes of search

4. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

5. Additional comments:

Box No. V **Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

1. Statement

Novelty (N)	Claims	1-49	YES
	Claims	None	NO
Inventive step (IS)	Claims	1-49	YES
	Claims	None	NO
Industrial applicability (IA)	Claims	1-49	YES
	Claims	None	NO

2. Citations and explanations:

Claims 1, 9, 15, 23, 28, 32, 36 and 44 meet the criteria set out in PCT Article 33(2)-(3), because the prior art does not teach or fairly suggest:

In regards to claim 1, a system for performing a plurality of prescribed data management functions in a manner that reduces redundant access operations to primary storage, the system comprising: at least a snapshot function operable to create a point-in-time image of primary storage data to secondary storage, and at least one back-up function operable to create at least one back-up copy of data, the data management engine being responsive to an electronic service level agreement [SLA] that specifies a schedule for performing data management functions, wherein point-in-time images of data include a reference to a complete baseline image of data at a specific point in time and difference data indicating changes to the data at a later, specific point in time, and wherein, in response to the schedule requiring at least some data management functions to be performed concurrently, the data management engine creates a point-in-time image of the primary storage data and communicates the difference information of that point-in-time image to the secondary storage to update the at least one of a back-up copy of the primary data, such that the primary storage is accessed only once for all corresponding sets of updates to the secondary storage.
The prior art (Bell et al., Cisler et al., Arakawa et al., Clements et al., and Frank et al.), while teaching some features and aspects of the claim limitations such as database management systems with database management engines in association data transfer/back-up systems and using snapshot copying of portiotions of the primary date depending on the applicable service level aggreement at scheduled backup processes (see figs. 1 and 11 as well as pgs. 1-2, para. 0010-0020, and pgs 4, 6 and 7, para. 0044, 65-68 of Bell et al.; figs. 1-4 as well as pgs. 1-2, para. 0018-0023 of Cisler et al.; fig. 1 as well as col. 2, line 66-col. 3, line 52 of Arakawa et al.; figs. 1-4 as well as pg. 1, para. 0003-0005 of Clements et al.; and figs. 1-3 as well as pgs. 2-4, para. 0018-0032 of Frank et al., respectively), yet, fails to teach the claim limitations in their entirety and as specifically recited in the claim.

In regards to claim 9, a system for managing data in accordance with service level agreements [SLAs] that specify schedules on a calendar basis for performing prescribed data management functions and for reducing inter-function redundancy, the system comprising: each policy specifying a source pool for data, a destination pool where a copy should be made of the source pool data, copy frequency indicating the frequency of operation for that policy, retention period indicating how long a given copy should be retained before being allowed to expire, and schedule information indicating hours and days of operation when the policy is in place, such that a collection of policies within a SLA is capable of expressing a non-uniform schedule for when a given function should be performed and is capable of expressing multiple data management functions that should be performed on a given source of data, and wherein the data management engine is operable to perform preparatory operations with the application and with the source pool so that the source pool of data has a coherent image of data to be copied and wherein the preparatory operations are performed once even if the SLA specifies multiple data management functions to be performed on that source pool at the current time.
The prior art (Bell et al., Cisler et al., Arakawa et al., Clements et al., and Frank et al.), while teaching some features and aspects of the claim limitations such as database management systems with database management engines in association data transfer/back-up systems with database management engines in association data transfer/back-up systems and using snapshot copying of portiotions of the primary date depending on the applicable service level aggreement at scheduled backup processes (see figs. 1 and 11 as well as pgs. 1-2, para. 0010-0020, and pgs 4, 6 and 7, para. 0044, 65-68 of Bell et al.; figs. 1-4 as well as pgs. 1-2, para. 0018-0023 of Cisler et al.; fig. 1 as well as col. 2, line 66-col. 3, line 52 of Arakawa et al.; figs. 1-4 as well as pg. 1, para. 0003-0005 of Clements et al.; and figs. 1-3 as well as pgs. 2-4, para. 0018-0032 of Frank et al., respectively), yet, fails to teach the claim limitations in their entirety and as specifically recited in the claim.

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.
Continuation of:

In regards to claim 15, a system for backing-up data from a first storage pool to a second storage pool using difference information between time states, the system comprising: a data management engine operable to execute a sequence of snapshot operations to create point-in-time images of application data on a first storage pool, each successive point-in-time image corresponding to a specific, successive time-state of the application data, and each snapshot operation creating difference information indicating which application data has changed and the content of the changed application data for the corresponding time state; the data management engine operable to execute at least one back-up function for the application data wherein the backup operation is scheduled for execution at nonconsecutive time-states wherein the data management engine is operable to maintain history information having time-state information indicating the time-state of the last back-up function performed on the application data for a corresponding back-up copy of data; and wherein the data management engine is operable to create composite difference information from the difference information for each time-state between the time-state of the last back-up function performed on the application data and the time-state of the currently-scheduled back-up function to be performed on the application data, and wherein the data management engine is operable to send the composite difference information to a second storage pool to be compiled with the back-up copy of data at the last time-state to create a back-up copy of data for the current time-state.
The prior art (Bell et al., Cisler et al., Arakawa et al., Clements et al., and Frank et al.), while teaching some features and aspects of the claim limitations such as database management systems with database management engines in association data transfer/back-up systems and using snapshot copying of portiotions of the primary date depending on the applicable service level aggreement at scheduled backup processes (see figs. 1 and 11 as well as pgs. 1-2, para. 0010-0020, and pgs 4, 6 and 7, para. 0044, 65-68 of Bell et al.; figs. 1-4 as well as pgs. 1-2, para. 0018-0023 of Cisler et al.; fig. 1 as well as col. 2, line 66-col. 3, line 52 of Arakawa et al.; figs. 1-4 as well as pg. 1, para. 0003-0005 of Clements et al.; and figs. 1-3 as well as pgs. 2-4, para. 0018-0032 of Frank et al., respectively), yet, fails to teach the claim limitations in their entirety and as specifically recited in the claim.

In regards to claim 19, a system for restoring data in a storage pool from a back-up copy of the data using difference information between time states, the system comprising: a data management engine operable to maintain history information indicating the time-states for which storage pools have point-in-time images of application data; and wherein the data management engine includes logic for restoring application data in a storage pool to a point-in-time image of the data for a specified time-state; the data management engine operable to identify the existence of a point-in-time image of the data at the storage pool for a time-state prior to the specified time-state and sending difference information from the back-up copy of data to the storage pool, the difference information indicating which application data has changed and the content of the changed application data for the time between the specified time state and the time state prior to the specified time-state.
The prior art (Bell et al., Cisler et al., Arakawa et al., Clements et al., and Frank et al.), while teaching some features and aspects of the claim limitations such as database management systems with database management engines in association data transfer/back-up systems and using snapshot copying of portiotions of the primary date depending on the applicable service level aggreement at scheduled backup processes (see figs. 1 and 11 as well as pgs. 1-2, para. 0010-0020, and pgs 4, 6 and 7, para. 0044, 65-68 of Bell et al.; figs. 1-4 as well as pgs. 1-2, para. 0018-0023 of Cisler et al.; fig. 1 as well as col. 2, line 66-col. 3, line 52 of Arakawa et al.; figs. 1-4 as well as pg. 1, para. 0003-0005 of Clements et al.; and figs. 1-3 as well as pgs. 2-4, para. 0018-0032 of Frank et al., respectively), yet, fails to teach the claim limitations in their entirety and as specifically recited in the claim.

In regards to claim 23, a method of forming deduplicated images of a data object that changes over time using difference information between temporal states of the data object, the method comprising: organizing the content of the data object for a first temporal state as a plurality of content segments and storing the content segments in a data store; creating an organized arrangement of hash structures to represent the data object in its first temporal state wherein for a subset of the hash structures, each structure includes a hash signature for a corresponding content segment and is associated with a reference to the corresponding content segment, and wherein the logical organization of the arrangement represents the organization of the content segments as they are represented within the data object; receiving difference information for the data object, the difference information indicating the changed content for the data object for a second temporal state relative to the first temporal state, and the difference information indicating the location of the changed content within the data object; forming at least one hash signature for the changed content; storing the changed content that is unique in the data store as content segments modifying the organized arrangement of hash structures to incorporate new structures for the at least one hash signature for the changed content, incorporating the new structures in the organized arrangement of structures at a position corresponding to the location of the changed content within the data object as indicated within the difference information, and associating the hash signatures for the new structures with references to the corresponding content segments for the changed content; and associating the new structures with the second temporal state, whereby a deduplicated image of the data object for a second temporal state is stored without requiring reception of a complete image of the data object for the second temporal state.
The prior art (Bell et al., Cisler et al., Arakawa et al., Clements et al., and Frank et al.), while teaching some features and aspects of the claim limitations such as database management systems in association with data management engines and corresponding data management method that involve the ability to access multiple data storage locations in order to transfer data from one data storage location to another location for storage and back-up purposes (see figs. 1 and 11 as well as pgs. 1-2, para. 0010-0020 of Bell et al.; fig. 5 as well as pgs. 1-2, para. 0018-0023 of Cisler; figs. 2-3 as well as col. 2, line 66-col. 3, line 52 of Arakawa et al.; figs. 5-9 as well as pg. 1, para. 0003-0005 of Clements et al.; and fig. 4 as well as pgs. 4-5, para. 0035-0039 of Frank et al., respectively), yet, fails to teach the claim limitations in their entirety and as specifically recited in the claim.

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.
Continuation of:

In regards to claim 28, a method of managing deduplicated images of data objects that change over time, the method comprising: organizing unique content of each data object as a plurality of content segments and storing the content segments in a data store; for each data object, creating an organized arrangement of hash structures, wherein each structure, for a subset of the hash structures, includes a hash signature for a corresponding content segment and is associated with a reference to the corresponding content segment, wherein the logical organization of the arrangement represents the logical organization of the content segments as they are represented within the data object, and wherein another subset of the hash structures includes a hierarchy of hash signatures of the hash signatures for corresponding content segments so that the organized arrangement may be traversed to determine if content is represented by the organized arrangement of hash structures; and for each data object, maintaining an organized arrangement of temporal structures to represent a corresponding data object over time, wherein each structure is associated with a temporal state of the data object and wherein the logical arrangement of structures is indicative of the changing temporal states of the data object, and wherein each temporal state is associated with the hash structures representing the content of the data object during that temporal state.
The prior art (Bell et al., Cisler et al., Arakawa et al., Clements et al., and Frank et al.), while teaching some features and aspects of the claim limitations such as database management systems in association with data management engines and corresponding data management method that involve the ability to access multiple data storage locations in order to transfer data from one data storage location to another location for storage and back-up purposes (see figs. 1 and 11 as well as pgs. 1-2, para. 0010-0020 of Bell et al.; fig. 5 as well as pgs. 1-2, para. 0018-0023 of Cisler; figs. 2-3 as well as col. 2, line 66-col. 3, line 52 of Arakawa et al.; figs. 5-9 as well as pg. 1, para. 0003-0005 of Clements et al.; and fig. 4 as well as pgs. 4-5, para. 0035-0039 of Frank et al., respectively); fails to teach the claim limitations in their entirety and as specifically recited in the claim.

In regards to claim 32, a method of storing deduplicated images in which a portion of the image is stored in encoded form directly in a hash table, the method comprising: organizing unique content of each data object as a plurality of content segments and storing the content segments in a data store; for each data object, creating an organized arrangement of hash structures, wherein each structure, for a subset of the hash structures, includes a field to contain a hash signature for a corresponding content segment and is associated with a reference to the corresponding content segment, wherein the logical organization of the arrangement represents the logical organization of the content segments as they are represented within the data object; receiving content to be included in the deduplicated image of the data object; determining if the received content may be encoded using a predefined non-lossy encoding technique and in which the encoded value would fit within the field for containing a hash signature; if so, placing the encoding in the field and marking the hash structure to indicate that the field contains encoded content for the deduplicated image; if not, generating a hash signature for the received content and placing the hash signature in the field and placing the received content in a corresponding content segment in the data store if it is unique.
The prior art (Bell et al., Cisler et al., Arakawa et al., Clements et al., and Frank et al.), while teaching some features and aspects of the claim limitations such as database management systems in association with data management engines and corresponding data management method that involve the ability to access multiple data storage locations in order to transfer data from one data storage location to another location for storage and back-up purposes (see figs. 1 and 11 as well as pgs. 1-2, para. 0010-0020 of Bell et al.; fig. 5 as well as pgs. 1-2, para. 0018-0023 of Cisler; figs. 2-3 as well as col. 2, line 66-col. 3, line 52 of Arakawa et al.; figs. 5-9 as well as pg. 1, para. 0003-0005 of Clements et al.; and fig. 4 as well as pgs. 4-5, para. 0035-0039 of Frank et al., respectively); yet, fails to teach the claim limitations in their entirety and as specifically recited in the claim.

In regards to claim 36, a method for using a first deduplicating store to update a second deduplicating store with information representing how data objects change over time, the method comprising:
at a first deduplicating store, organizing unique content of each data object as a plurality of content segments and storing the content segments in a data store; at a first deduplicating store, for each data object, creating an organized arrangement of hash structures, wherein each structure, for a subset of the hash structures, includes a hash signature for a corresponding content segment and is associated with a reference to the corresponding content segment, wherein the logical organization of the arrangement represents the logical organization of the content segments as they are represented within the data object; at a first deduplicating store, for each data object, maintaining an organized arrangement of temporal structures to represent a corresponding data object over time, wherein each structure is associated with a temporal state of the data object and wherein the logical arrangement of structures is indicative of the changing temporal states of the data object, and wherein each temporal state is associated with the hash structures representing the content of the data object that has changed relative to a prior temporal state; at a second dedeuplicating store, organizing unique content of each data object as a plurality of content segments and storing the content segments in a data store; at a second deduplicating store, for each data object, maintaining an organized arrangement of hash structures, that is at least a subset of the hash structures at the first deduplicating store; at a second deduplicating store, for each data object, maintaining an organized arrangement of temporal structures to represent a corresponding data object over time, wherein the organized arrangement of temporal structures is at least a subset of the temporal structures at the first deduplicating store, representing a subset of the temporal states; in response to a request to update the second deduplicating store with information from the first deduplicating store, finding a temporal state that is common to the first and second deduplicating stores and that is in temporal proximity to the current state of the second deduplicating store; and compiling a set of hash signatures for the content that has changed from the common state to the current temporal state of the first deduplicating store and sending that set of hash signatures to the second deduplicating store so it can update its organized arrangement of hash structures to represent the content of the data object up to the current temporal state of the first deduplicating store.
The prior art (Bell et al., Cisler et al., Arakawa et al., Clements et al., and Frank et al.), while teaching some features and aspects of the claim limitations such as database management systems in association with data management engines and corresponding data management method that involve the ability to access multiple data storage locations in order to transfer data from one data storage location to another location for storage and back-up purposes (see figs. 1 and 11 as well as pgs. 1-2, para. 0010-0020 of Bell et al.; fig. 5 as well as pgs. 1-2, para. 0018-0023 of Cisler; figs. 2-3 as well as col. 2, line 66-col. 3, line 52 of Arakawa et al.; figs. 5-9 as well as pg. 1, para. 0003-0005 of Clements et al.; and fig. 4 as well as pgs. 4-5, para. 0035-0039 of Frank et al., respectively); yet, fails to teach the claim limitations in their entirety and as specifically recited in the claim.

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.

Continuation of:

With regards to claim 44, a method of performing garbage collection to identify content segments no longer referenced in a deduplicating storage system in which redundant mark operations in a mark-and-sweep technique are avoided, the method comprising: organizing unique content of each data object as a plurality of content segments in the deduplicating storage system; for each data object, creating an organized arrangement of hash structures, wherein each structure, for a subset of the hash structures, includes a hash signature for a corresponding content segment and is associated with a reference to the corresponding content segment, wherein the logical organization of the arrangement represents the logical organization of the content segments as they are represented within the data object, and wherein another subset of the hash structures includes a hierarchy of hash signatures of the hash signatures for corresponding content segments so that the organized arrangement may be traversed to determine if content is represented by the organized arrangement of hash structures; for each data object, maintaining an organized arrangement of temporal structures to represent a corresponding data object over time, wherein each structure is associated with a temporal state of the data object and wherein the logical arrangement of structures is indicative of the changing temporal states of the data object, and wherein each temporal state is associated with the hash structures representing the content of the data object that has changed relative to an immediately prior temporal state of the data object; for each content segment in the deduplicating storage system, clearing its corresponding garbage collection state; iterating over the temporal structures and, for each temporal structure, marking the garbage collection state for the associated content segments for only the content segments that have changed relative to an immediately prior temporal state of the data object; and returning any content segments to a free pool of storage that has a cleared garbage collection state after the iteration step.
The prior art (Bell et al., Cisler et al., Arakawa et al., Clements et al., and Frank et al.), while teaching some features and aspects of the claim limitations such as database management systems in association with data management engines and corresponding data management method that involve the ability to access multiple data storage locations in order to transfer data from one data storage location to another location for storage and back-up purposes (see figs. 1 and 11 as well as pgs. 1-2, para. 0010-0020 of Bell et al.; fig. 5 as well as pgs. 1-2, para. 0016-0023 of Cisler; figs. 2-3 as well as col. 2, line 66-col. 3, line 52 of Arakawa et al.; figs. 5-9 as well as pg. 1, para. 0003-0005 of Clements et al.; and fig. 4 as well as pgs. 4-5, para. 0035-0039 of Frank et al., respectively); yet, fails to teach the claim limitations in their entirety and as specifically recited in the claim.

Claims 2-8, 10-14, 17-22, 24-27, 29-31, 33-35, 37-43 and 45-49 meet the criteria set out in PCT Article 33(2)-(3), because each claim depends either directly or indirectly from the novel base claim 1, 9, 15, 22, 28, 32, 36 and 44, respectively.

Claims 1-49 meet the criteria set out in PCT Article 33(4), and thus have industrial applicability because the subject matter claimed can be made or used in industry.

Search History:

Complete Classification Search

The Patent Analyst performed a complete classification search within the following US, IPC, ECLA, or F-Term classification areas:

U.S. Class/Subclasses:
707/640, 644-45, 648-54, 674-79, 790-92

IPC (8) Class/Subclasses: (2012.01)

See global search results.

Limited Classification Search

The Patent Analyst performed a limited classification search within the following US, IPC, ECLA, or F-Term classification areas:

US Classes and Subclasses:
370; 707; 710; 711; 711/100, 161-62, E12.103

IPC(8) Classes and Subclasses:
G06F; G06F 7/00, 12/00, 12/02, 12/08 (2012.01)

See global search results.

Global Patent Literature Text Search

Google Patent http://www.google.com/advanced_patent_search

No. of Hits	Document Part: Text String
375	(data OR information) management virtualization
0	(data OR information) management virtualization ((data-engine) OR (management-engine)) (snapshot OR snapshots OR clone OR copy OR clones OR copies OR "point-in-time images") (difference OR changes) ((time-state) OR (time-states) OR (time-event) OR (time-events) OR (timeline-state) OR (timeline-states))
0	(data OR information) management virtualization ((data-engine) OR (management-engine)) (snapshot OR snapshots OR clone OR copy OR clones OR copies OR "point-in-time images") (difference OR changes) (timeline OR time) (state OR states OR event OR events)
0	(data OR information) management virtualization ((data-engine) OR (management-engine)) (snapshot OR snapshots OR clone OR copy OR clones OR copies OR "point-in-time images") (difference OR changes) (state OR states OR event OR events)
0	(data OR information) management virtualization ((data-engine) OR (management-engine)) (snapshot OR snapshots OR clone OR copy OR clones OR copies OR "point-in-time images") (difference OR changes OR differences) (time OR timeline)
0	(data OR information) management virtualization ((data-engine) OR (management-engine)) (snapshot OR snapshots OR clone OR copy OR clones OR copies) (transformation OR transformations) (time OR timeline) (event OR events OR state OR states)

0	(data OR information) management virtualization ((data-engine) OR (management-engine)) (snapshot OR snapshots OR clone OR copy OR clones OR copies) (transformation OR transformations) (time OR timeline)
2	(data OR information) management virtualization ((data-engine) OR (management-engine)) (snapshot OR snapshots OR clone OR copy OR clones OR copies) (difference OR differences OR changes OR transformation OR transformations)
1	(data OR information) management ((data-engine) OR (management-engine)) (snapshot OR snapshots OR clone OR copy OR clones OR copies) (difference OR differences OR changes OR transformation OR transformations) ((time-state) OR (time-states) OR (time-event) OR (time-events) OR (timeline-state) OR (timeline-states))
56	(data OR information) management ((data-engine) OR (management-engine)) (snapshot OR snapshots OR clone OR copy OR clones OR copies) (difference OR differences OR changes OR transformation OR transformations) (time OR times) (event OR events OR state OR states)
1	(data OR information) management ((data-engine) OR (management-engine)) (snapshot OR snapshots OR clone OR copy OR clones OR copies) (difference OR differences OR changes OR transformation OR transformations) (time OR times) (event OR events OR state OR states) (service=agreement)
1	time OR times) (event OR events OR state OR states) (service=agreements)
3	(data OR information) management ((data-engine) OR (management-engine)) (snapshot OR snapshots OR clone OR copy OR clones OR copies) (difference OR differences OR changes OR transformation OR transformations) (time OR times) (event OR events OR state OR states) (business-requirement)
5	(data OR information) management virtualization ((data-engine) OR (management-engine)) ((back-up) OR backup)
430	(data OR information) management virtualization (snapshot OR snapshots OR clone OR copy OR clones OR copies) (difference OR differences OR changes OR transformation OR transformations)
160	(data OR information) management virtualization (snapshot OR snapshots OR clone OR copy OR clones OR copies) (difference OR differences OR changes OR transformation OR transformations) (backup OR (back-up)) (first OR second OR primary OR secondary) (storage OR memory OR disk OR disks OR tape OR (optical-media) OR (optical-medium) OR clouds)
8	(data OR information) management virtualization (snapshot OR snapshots OR clone OR copy OR clones OR copies) (difference OR differences OR changes OR transformation OR transformations) (backup OR (back-up)) (first OR second OR primary OR secondary) cloud
371	(data OR information) management (snapshot OR snapshots OR clone OR copy OR clones OR copies) (difference OR differences OR changes OR transformation OR transformations) (backup OR (back-up)) (first OR second OR primary OR secondary) (storage OR memory OR disk OR disks OR tape OR (optical-media) OR (optical-medium) OR clouds)
0	(data OR information) management engine ((service-agreement) OR (business-requirement)) (baseline OR reference OR initial OR source) (snapshot OR snapshots OR clone OR copy OR clones OR copies OR image OR images) (difference OR differences OR changes) (back-up) (storage OR memory OR pool OR pools) (update OR updates) schedule
0	(data OR information) management engine ((service-agreement) OR (business-requirement)) (baseline OR reference OR initial OR source) ((point-time) OR realtime OR (real-time)) (image OR images) (difference OR differences OR changes) (back-up) (storage OR memory OR pool OR pools) (update OR updates) schedule
14	(data OR information) management engine (sla OR slas OR (service-level) OR (business-requirement)) (snapshot OR snapshots OR clone OR clones OR copy OR copies OR image OR images) (difference OR differences OR changes) (redundancy

	OR (back-up)) (storage OR memory OR pool OR pools)
1	(data OR information) management engine (sla OR slas OR (service-level) OR (business-requirement)) (snapshot OR snapshots OR clone OR clones OR copy OR copies OR image OR images) (difference OR differences OR changes) (redundancy OR (back-up)) (storage OR memory OR pool OR pools) frequency lifecycle
0	(data OR information) management engine (sla OR slas OR (service-level) OR (business-requirement)) (snapshot OR snapshots OR clone OR clones OR copy OR copies OR image OR images) (difference OR differences OR changes) (redundancy OR (back-up)) (storage OR memory OR pool OR pools) frequency (schedule OR schedules)
0	(data OR information) management engine (sla OR slas OR (service-level) OR (business-requirement)) (snapshot OR snapshots OR clone OR clones OR copy OR copies OR image OR images) (difference OR differences OR changes) (redundancy OR (back-up)) (storage OR memory OR pool OR pools) frequency hours days
7	(data OR information) management engine (sla OR slas OR (service-level) OR (business-requirement)) (snapshot OR snapshots OR clone OR clones OR copy OR copies OR image OR images) (difference OR differences OR changes) (redundancy OR (back-up)) (storage OR memory OR pool OR pools) (time OR timing) (state OR states)
7	(data OR information) management engine (sla OR slas OR (service-level) OR (business-requirement)) (snapshot OR snapshots OR clone OR clones OR copy OR copies OR image OR images) (difference OR differences OR changes) (redundancy OR (back-up)) (storage OR memory OR pool OR pools) (time OR timing) (event OR events)
2	(data OR information) management engine (sla OR slas OR (service-agreement) OR (business-requirement)) (snapshot OR snapshots OR clone OR clones OR copy OR copies OR image OR images) (difference OR differences OR changes) (redundancy OR (back-up)) (storage OR memory OR pool OR pools) (time OR timing) (event OR events)
1	(data OR information) management engine (sla OR slas OR (service-agreement) OR (business-requirement)) (snapshot OR snapshots OR clone OR clones OR copy OR copies OR image OR images) (difference OR differences OR changes) (redundancy OR (back-up)) (storage OR memory OR pool OR pools) (time OR timing) (state OR states)
56	(data OR information OR content) (management OR organization OR organized) (snapshot OR snapshots OR clone OR copy OR clones OR copies OR "point-in-time images") (difference OR changes) (time OR times OR temporal) (state OR states OR event OR events) hash (marks OR structures)
5	(data OR information OR content) (management OR organization OR organized) (snapshot OR snapshots OR clone OR copy OR clones OR copies OR "point-in-time images") (difference OR changes) (time OR times OR temporal) (state OR states OR event OR events) hash deduplicated
12	(data OR information OR content) (management OR organization OR organized) (snapshot OR snapshots OR clone OR copy OR clones OR copies OR "point-in-time images") (difference OR changes) (time OR times OR temporal) (state OR states OR event OR events) (deduplicated OR (eliminate-duplicate))
120	(data OR information OR content) (management OR organization OR organized) (snapshot OR snapshots OR clone OR copy OR clones OR copies OR image OR images) (difference OR changes) (time OR times OR temporal) (state OR states OR event OR events) hash-(signature OR signatures)
20	(data OR information OR content) (management OR organization OR organized) (snapshot OR snapshots OR clone OR copy OR clones OR copies OR image OR images) (difference OR changes OR transformation OR transformations) (time OR times OR temporal) hash-deduplicated (first OR primary OR second OR secondary) (store OR storage)

20	(data OR information OR content) (management OR organization OR organized) (snapshot OR snapshots OR clone OR copy OR clones OR copies OR image OR images) (difference OR changes OR transformation OR transformations) (time OR times OR temporal) hash-deduplicated (first OR primary OR second OR secondary) (store OR storage)
0	(garbage OR expired OR trash)collection (data OR information OR content) (management OR organization OR organized) (snapshot OR snapshots OR clone OR copy OR clones OR copies OR image OR images) (difference OR changes OR transformation OR transformations) (time OR times OR temporal) hash (structure OR structures) (dedupliated OR dedupliating) (store OR storage)
0	(garbage OR expired OR trash)collection (data OR information OR content) (management OR organization OR organized) (snapshot OR snapshots OR clone OR copy OR clones OR copies OR image OR images) (difference OR changes OR transformation OR transformations) (time OR times OR temporal) hash (structure OR structures) (dedupliated OR dedupliating) (store OR storage)
0	(garbage OR expired OR trash)collection (data OR information OR content) (management OR organization OR organized) (snapshot OR snapshots OR clone OR copy OR clones OR copies OR image OR images) (difference OR changes OR transformation OR transformations) (time OR times OR temporal) hash (marking OR markings) (dedupliated OR dedupliating) (store OR storage)
0	(garbage OR expired OR trash)collection (data OR information OR content) (management OR organization OR organized) (snapshot OR snapshots OR clone OR copy OR clones OR copies OR image OR images) (difference OR changes OR transformation OR transformations) (time OR times OR temporal) hash (signature OR signatures) (dedupliated OR dedupliating) (store OR storage)

MicroPat

101	48 hits	Full patent spec.	(data or information or content or collection*) and (data same management engine) and (management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation* or modification* or modified or modify or modifies) and (timeline or time or hours or days) and (backup or back-up or (back adj5 up) or redundan*) and (storage or store or pool) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update* or updating) and (baseline or base or reference or initial) and calendar and frequency
		Current IPC-R	G06F
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
100	14 hits	Full patent spec.	(data or information or content or collection*) and (garbage or trash or expire*) and (management or manage or manages or managing or organiz*) and (time* or temporal) and (state* or event*) and (store* or storage or pool*) and deduplicat* and (hash same (signature* or structure* or mark*))
		Current IPC-R	G06F
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
99	59 hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or

			clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation* or modification* or modified or modify or modifies) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store or pool or memory or tape or (optical adj media) or (optical adj medium) or disk or disks) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update or updates or updating)
		Current IPC-R	G06F
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
98	no hits	Full patent spec.	(data or information or content or collection*) and (data same management engine) and (management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation* or modification* or modified or modify or modifies) and (timeline or time or hours or days) and (backup or back-up or (back adj5 up) or redundan*) and (storage or store or pool) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update* or updating) and (baseline or base or reference or initial) and calendar and frequency
		Current IPC-R	G06F001208
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
97	no hits	Full patent spec.	(data or information or content or collection*) and (garbage or trash or expire*) and (management or manage or manages or managing or organiz*) and (time* or temporal) and (state* or event*) and (store* or storage or pool*) and deduplicat* and (hash same (signature* or structure* or mark*))
		Current IPC-R	G06F001208
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
96	no hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation* or modification* or modified or modify or modifies) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store or pool or memory or tape or (optical adj media) or (optical adj medium) or disk or disks) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update or updates or updating)
		Current IPC-R	G06F001208
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
95	21683 hits	Current IPC-R	G06F001208
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012

94	no hits	Full patent spec.	(data or information or content or collection*) and (data same management engine) and (management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation* or modification* or modified or modify or modifies) and (timeline or time or hours or days) and (backup or back-up or (back adj5 up) or redundan*) and (storage or store or pool) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update* or updating) and (baseline or base or reference or initial) and calendar and frequency
		Current IPC-R	G06F001202
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
93	1 hit	Full patent spec.	(data or information or content or collection*) and (garbage or trash or expire*) and (management or manage or manages or managing or organiz*) and (time* or temporal) and (state* or event*) and (store* or storage or pool*) and deduplicat* and (hash same (signature* or structure* or mark*))
		Current IPC-R	G06F001202
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
92	no hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation* or modification* or modified or modify or modifies) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store or pool or memory or tape or (optical adj media) or (optical adj medium) or disk or disks) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update or updates or updating)
		Current IPC-R	G06F001202
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
91	14706 hits	Current IPC-R	G06F001202
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
90	no hits	Full patent spec.	(data or information or content or collection*) and (data same management engine) and (management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation* or modification* or modified or modify or modifies) and (timeline or time or hours or days) and (backup or back-up or (back adj5 up) or redundan*) and (storage or store or pool) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update* or updating) and (baseline or base or reference or initial) and calendar and frequency
		Current IPC-R	G06F001200
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA

		Years	1836-2012
89	2 hits	Full patent spec.	(data or information or content or collection*) and (garbage or trash or expire*) and (management or manage or manages or managing or organiz*) and (time* or temporal) and (state* or event*) and (store* or storage or pool*) and deduplicat* and (hash same (signature* or structure* or mark*))
		Current IPC-R	G06F001200
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
88	no hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation* or modification* or modified or modify or modifies) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage^ or store or pool or memory or tape or (optical adj media) or (optical adj medium) or disk or disks) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update or updates or updating)
		Current IPC-R	G06F001200
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
87	69624 hits	Current IPC-R	G06F001200
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
86	no hits	Full patent spec.	(data or information or content or collection*) and (garbage or trash or expire*) and (management or manage or manages or managing or organiz*) and (time* or temporal) and (state* or event*) and (store* or storage or pool*) and deduplicat* and (hash same (signature* or structure* or mark*))
		Current IPC-R	G06F000700
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
85	1 hit	Full patent spec.	(data or information or content or collection*) and (data same management engine) and (management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation* or modification^ or modified or modify or modifies) and (timeline or time or hours or days) and (backup or back-up or (back adj5 up) or redundan*) and (storage or store or pool) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update* or updating) and (baseline or base or reference or initial) and calendar and frequency
		Current IPC-R	G06F000700
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
84	3 hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation* or

			modification* or modified or modify or modifies) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store or pool or memory or tape or (optical adj media) or (optical adj medium) or disk or disks) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update or updates or updating)
		Current IPC-R	G06F000700
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
83	15145 hits	Current IPC-R	G06F000700
		Databases	EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
82	2 hits	Full patent spec.	(data or information or content or collection*) and (garbage or trash or expire*) and (management or manage or manages or managing or organiz*) and (time* or temporal) and (state* or event*) and (store* or storage or pool*) and deduplicat* and (hash same (signature* or structure* or mark*))
		Current US Class	711E12103
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
81	no hits	Full patent spec.	(data or information or content or collection*) and (data same management engine) and (management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation* or modification* or modified or modify or modifies) and (timeline or time or hours or days) and (backup or back-up or (back adj5 up) or redundan*) and (storage or store or pool) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update* or updating) and (baseline or base or reference or initial) and calendar and frequency
		Current US Class	711E12103
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
80	1 hit	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation* or modification* or modified or modify or modifies) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store or pool or memory or tape or (optical adj media) or (optical adj medium) or disk or disks) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update or updates or updating)
		Current US Class	711E12103
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
79	4 hits	Full patent spec.	(data or information or content or collection*) and (garbage or

			trash or expire*) and (management or manage or manages or managing or organiz*) and (time* or temporal) and (state* or event*) and (store* or storage or pool*) and deduplicat* and (hash same (signature* or structure* or mark*))
		Current US Class	711162
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
78	no hits	Full patent spec.	(data or information or content or collection*) and (data same management engine) and (management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation* or modification* or modified or modify or modifies) and (timeline or time or hours or days) and (backup or back-up or (back adj5 up) or redundan*) and (storage or store or pool) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update* or updating) and (baseline or base or reference or initial) and calendar and frequency
		Current US Class	711162
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
77	1 hit	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation* or modification* or modified or modify or modifies) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store or pool or memory or tape or (optical adj media) or (optical adj medium) or disk or disks) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update or updates or updating)
		Current US Class	711162
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
76	5094 hits	Current US Class	711162
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
75	1 hit	Full patent spec.	(data or information or content or collection*) and (garbage or trash or expire*) and (management or manage or manages or managing or organiz*) and (time* or temporal) and (state* or event*) and (store* or storage or pool*) and deduplicat* and (hash same (signature* or structure* or mark*))
		Current US Class	711161
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
74	no hits	Full patent spec.	(data or information or content or collection*) and (data same management engine) and (management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation* or modification* or modified or modify or modifies) and (timeline or time or hours or days) and (backup or

			back-up or (back adj5 up) or redundan*) and (storage or store or pool) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update* or updating) and (baseline or base or reference or initial) and calendar and frequency
		Current US Class	711161
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
73	no hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation* or modification* or modified or modify or modifies) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store or pool or memory or tape or (optical adj media) or (optical adj medium) or disk or disks) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update or updates or updating)
		Current US Class	711161
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
72	1681 hits	Current US Class	711E12103
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
71	1 hit	Full patent spec.	(data or information or content or collection*) and (garbage or trash or expire*) and (management or manage or manages or managing or organiz*) and (time* or temporal) and (state* or event*) and (store* or storage or pool*) and deduplicat* and (hash same (signature* or structure* or mark*))
		Current US Class	711100
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
70	no hits	Full patent spec.	(data or information or content or collection*) and (data same management engine) and (management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation* or modification* or modified or modify or modifies) and (timeline or time or hours or days) and (backup or back-up or (back adj5 up) or redundan*) and (storage or store or pool) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update* or updating) and (baseline or base or reference or initial) and calendar and frequency
		Current US Class	711100
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
69	no hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation* or

		modification* or modified or modify or modifies) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store or pool or memory or tape or (optical adj media) or (optical adj medium) or disk or disks) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduing or update or updates or updating) Current US Class 711100 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
68	2883 hits	Current US Class 711100 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
67	372 hits	Current US Class 707792 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
66	450 hits	Current US Class 707791 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
65	133 hits	Current US Class 707790 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
64	98 hits	Current US Class 707679 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
63	26 hits	Current US Class 707678 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
62	15 hits	Current US Class 707677 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
61	18 hits	Current US Class 707676 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
60	36 hits	Current US Class 707675 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
59	242 hits	Current US Class 707674 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
58	110 hits	Current US Class 707654 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
57	31 hits	Current US Class 707653 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012

56	77 hits	Current US Class 707652 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
55	30 hits	Current US Class 707651 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
54	37 hits	Current US Class 707650 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
53	181 hits	Current US Class 707649 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
52	162 hits	Current US Class 707648 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
51	42 hits	Current US Class 707645 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
50	35 hits	Current US Class 707644 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
49	443 hits	Current US Class 707640 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
48	4 hits	Full patent spec. (data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (modification* or modified or modify or modifies) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and ((primary or first) same (store or storage)) and ((second or secondary) same (storage* or store)) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update* or updating) and (baseline or base or reference or initial) and virtualization and bitmap* Current US Class 370 or 710 or 711 or 707 Databases USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA Years 1836-2012
47	11 hits	Full patent spec. (data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (modification* or modified or modify or modifies) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and ((primary or first) same (store or storage)) and ((second or secondary) same (storage* or store)) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update* or updating) and (baseline or base or reference or initial) and virtualization Current US Class 370 or 710 or 711 or 707

		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
46	16 hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (modification* or modified or modify or modifies) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduing or update* or updating) and (baseline or base or reference or initial) and virtualization
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
45	165 hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (modification* or modified or modify or modifies) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduing or update* or updating)
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
44	550 hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (modification* or modified or modify or modifies) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduing or update* or updating)
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
43	6 hits	Patent/Publication No.	US20080034016 or US6883073 or US7657582 or US20090019535 or US20100077013 or US20100138827
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
42	3 hits	Patent/Publication No.	US20060235715 or US20020049778 or US20050267878
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
41	78 hits	Full patent spec.	(data or information or content or collection*) and (garbage or trash or expire*) and (management or manage or manages or managing or organiz*) and (time* or temporal) and (state* or event*) and (store* or storage or pool*) and deduplicat* and

			(hash same (signature* or structure* or mark*))
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
40	59 hits	Full patent spec.	(data or information or content or collection*) and (management or manage or manages or managing or organiz*) and (time* or temporal) and (state* or event*) and ((first or primary) same deduplicat* same (store* or storage or pool*)) and ((second or secondary) same deduplicat* same (store* or storage or pool*)) and (hash same (signature* or structure* or mark*)) and (update* or updating)
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
39	67 hits	Full patent spec.	(data or information or content or collection*) and (management or manage or manages or managing or organiz*) and (time* or temporal) and ((first or primary) same deduplicat* same (store* or storage or pool*)) and ((second or secondary) same deduplicat* same (store* or storage or pool*)) and (hash same (signature* or structure* or mark*)) and (update* or updating)
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
38	84 hits	Full patent spec.	(data or information or content or collection*) and (management or manage or manages or managing or organiz*) and (time* or temporal) and ((first or primary) same deduplicat* same (store* or storage or pool*)) and ((second or secondary) same deduplicat* same (store* or storage or pool*)) and (hash same (signature* or structure* or mark*))
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
37	62 hits	Full patent spec.	(data or information or content or collection*) and (management or manage or manages or managing or organiz*) and image* and (time* or temporal) and deduplicat* and (store* or storage or pool*) and (hash same (signature* or structure* or mark*)) and (encode* or encoding)
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
36	82 hits	Full patent spec.	(data or information or content or collection*) and (management or manage or manages or managing or organiz*) and image* and (difference* or changes or change or transformation*) and (time* or temporal) and (state or states or event or events) and deduplicat* and (store* or storage or pool*) and (hash same (signature* or structure* or mark*))
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
35	36 hits	Full patent spec.	(data or information or content or collection*) and (management or manage or manages or managing or organiz*) and image* and (difference* or changes or change or transformation*) and (time* or temporal) and (state or states or event or events) and deduplicat* and (store* or storage or pool*) and (hash same (signature* or structure* or mark*)) and hierarchy
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
34	no hits	Full patent spec.	(data or information or content or collection*) and (management or manage or manages or managing or organiz*) and image* and (difference* or changes or change or transformation*) and

			(time* or temporal) and (state or states or event or events) and deduplicat* and (store* or storage or pool*) and (hash same (signature* or structure* or mark*)) and (hash same hierarchy)
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
33	183 hits	Full patent spec.	(data or information or content or collection*) and (management or manage or manages or managing or organiz*) and (difference* or changes or change or transformation*) and (timeline or time or temporal) and (state or states or event or events) and deduplicat* and (store* or storage or pool*) and (hash same (signature* or structure* or mark*))
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
32	663 hits	Full patent spec.	(data or information or content or collection*) and (data same management engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and (timeline or time) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduing) and history and ((first or primary) same (pool or store or storage)) and ((second or secondary) and (pool or store or storage))
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
31	130 hits	Full patent spec.	(data or information or content or collection*) and (data same management engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and (timeline or time) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduing) and history and ((first or primary) same (pool or store or storage)) and ((second or secondary) and (pool or store or storage))
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
30	47 hits	Full patent spec.	(data or information or content or collection*) and (data same management engine) and (management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and (timeline or time or hours or days) and (backup or back-up or (back adj5 up) or redundan*) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduing or update* or updating) and ((baseline or base or reference or initial) same (pool or store or storage)) and (pool or store or storage) and calendar and frequency
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
29	62	Full patent spec.	(data or information or content or collection*) and (data same

	hits		management engine) and (management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and (timeline or time or hours or days) and (backup or back-up or (back adj5 up) or redundan*) and (storage or store or pool) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update* or updating) and (baseline or base or reference or initial) and calendar and frequency
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
28	75 hits	Full patent spec.	(data or information or content or collection*) and (data same management engine) and (management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and (timeline or time or hours or days) and (backup or back-up or (back adj5 up) or redundan*) and (storage or store or pool) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update* or updating) and (baseline or base or reference or initial) and calendar
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
27	139 hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and ((snapshot* or clone* or copy or copies or (point near5 time near5 image*)) same (data or information or content)) and ((difference* or changes or change or transformation*) same (data or collection or information)) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and ((primary or first) same (storage or store)) and ((second or secondary) same (storage* or store)) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling) or (update* or updating) and (baseline or base or reference or initial) and virtualization and deduplicat*
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
26	644 hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and ((snapshot* or clone* or copy or copies or (point near5 time near5 image*)) same (data or information or content)) and ((difference* or changes or change or transformation*) same (data or collection or information)) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and ((primary or first) same (storage or store)) and ((second or secondary) same (storage* or store)) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling) or (update* or updating) and (baseline or base or reference or initial) and

			virtualization and bitmap*
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
25	4324 hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and ((snapshot* or clone* or copy or copies or (point near5 time near5 image*)) same (data or information or content)) and ((difference* or changes or change or transformation*) same (data or collection or information)) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and ((primary or first) same (storage or store)) and ((second or secondary) same (storage* or store)) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduing) or (update* or updating) and (baseline or base or reference or initial) and virtualization
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
24	12 hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and ((snapshot* or clone* or copy or copies or (point near5 time near5 image*)) same (data or information or content)) and ((difference* or changes or change or transformation*) same (data or collection or information)) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduing or update* or updating) and (baseline or base or reference or initial) and virtualization
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
23	16 hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduing or update* or updating) and (baseline or base or reference or initial) and virtualization
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
22	47 hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and

			(timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update* or updating) and (baseline or base or reference or initial) and virtualization
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
21	558 hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) and (schedule or schedules or scheduling or update* or updating)
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
20	175 hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas)
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
19	570 hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas)
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
18	155 hits	Full patent spec.	(data or information or content or collection*) and management and virtualization and (backup or back-up or (back adj5 up) or redundan*) and (storage* or store or stores) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and ((business near5 requirement*) or (service near2 level near2 agreement*) or (service near5 agreement*) or sla or slas) ...
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA

		Years	1836-2012
17	2227 hits	Full patent spec.	(data or information or content or collection*) and management and virtualization and (backup or back-up or (back adj5 up) or redundan*) and (storage* or store or stores) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*)
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
16	3383 hits	Full patent spec.	(data or information or content or collection*) and management and virtualization and (backup or back-up or (back adj5 up) or redundan*) and (storage* or store or stores)
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
15	270 hits	Full patent spec.	(data or information or content or collection*) and management and virtualization and (data same management same engine) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and (storage* or store or stores)
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
14	622 hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and ((primary or first) same (store or storage)) and ((second or secondary) same (storage* or store))
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
13	1680 hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and ((primary or first) same (store or storage)) and ((second or secondary) same (storage* or store))
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
12	947 hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store)
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA

		Years	1836-2012
11	3279 hits	Full patent spec.	(data or information or content or collection*) and management and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*) and image* and (primary or first or second or secondary) and (storage* or store)
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
10	221 hits	Full patent spec.	(data or information or content or collection*) and management and virtualization and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*)
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
9	582 hits	Full patent spec.	(data or information or content or collection*) and management and virtualization and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and (timeline or time or timing) and (state or states or event or events) and (backup or back-up or (back adj5 up) or redundan*)
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
8	251 hits	Full patent spec.	(data or information or content) and management and virtualization and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and (timeline or time or timing) and (state or states or event or events)
		Current US Class	370 or 710 or 711 or 707
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
7	701 hits	Full patent spec.	(data or information or content) and management and virtualization and (data same management same engine) and (snapshot* or clone* or copy or copies or (point near5 time near5 image*)) and (difference* or changes or change or transformation*) and (timeline or time or timing) and (state or states or event or events)
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
6	no hits	Inventor(s)	actifio
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
5	no hits	Assignee/Applicant	actifio
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012

4	417 hits	Assignee/Applicant	ashutosh or (provenzano near5 christopher) or (chang near5 david) or (abercrombie near5 philip) or (mutalik near5 madhav) or (roman near5 mark)
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
3	130 hits	Full patent spec.	(data or information or content) and (management or organiz*) and (backup or back-up or (back adj3 up) or redundan*)
		Inventor(s)	ashutosh or (provenzano near5 christopher) or (chang near5 david) or (abercrombie near5 philip) or (mutalik near5 madhav) or (roman near5 mark)
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
2	583 hits	Full patent spec.	(data or information or content) and (management or organiz*)
		Inventor(s)	ashutosh or (provenzano near5 christopher) or (chang near5 david) or (abercrombie near5 philip) or (mutalik near5 madhav) or (roman near5 mark)
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012
1	2306 hits	Inventor(s)	ashutosh or (provenzano near5 christopher) or (chang near5 david) or (abercrombie near5 philip) or (mutalik near5 madhav) or (roman near5 mark)
		Databases	USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA
		Years	1836-2012

Computer Accessed Text Databases Searched

The Patent Analyst searched the following computer accessed text databases:

ProQuest www.proquest.com

No. of Hits	Document Part: Text String
1	(data OR information OR collection OR content) AND management AND engine AND (snapshot OR snapshots) AND (difference OR differences OR changes OR modification) AND update AND schedule AND deduplicated
4	(data OR information OR collection OR content) AND management AND engine AND (snapshot OR snapshots) AND (difference OR differences OR changes OR modification) AND deduplicated
145	(data OR information OR collection OR content) AND management AND engine AND (snapshot OR snapshots) AND (difference OR differences OR changes OR modification) AND update AND schedule AND virtualization
290	(data OR information OR collection OR content) AND management AND engine AND (snapshot OR snapshots) AND (difference OR differences OR changes OR modification) AND update AND schedule AND hash
690	(data OR information OR collection OR content) AND management AND engine AND (snapshot OR snapshots) AND (difference OR differences OR changes OR modification) AND (event OR state OR events OR states) AND (time OR times) AND hash

252	(garbage OR trash OR expired) AND (data OR information OR collection OR content) AND management AND engine AND (snapshot OR snapshots) AND (difference OR differences OR changes OR modification) AND (event OR state OR events OR states) AND (time OR times) AND hash

GOOGLE SCHOLAR www.googlescholar.com

No. of Hits	Document Part: Text String
1580	(data OR information) management engine (snapshot OR snapshots OR clone OR copy OR clones OR copies) (difference OR differences OR changes OR modification) (event OR state) (sla OR (service-agreement)) update schedule -pat
357	(data OR information) management engine (snapshot OR snapshots OR clone OR copy OR clones OR copies) (difference OR differences OR changes OR modification) (event OR state) (sla OR (service-agreement)) update schedule deduplicated -pat
440	(data OR information) management engine (snapshot OR snapshots OR (point-image)) (difference OR differences OR changes OR modification) (event OR state) (sla OR (service-agreement)) update schedule (store OR storage OR pool) -pat
233	(data OR information) (management OR organization) (snapshot OR snapshots OR image OR images) (difference OR changes) (time OR temporal) (state OR event) ((hash-structure) OR (hash-structures)) -pat
194	(collection OR content) (management OR organization) (snapshot OR snapshots OR image OR images) (difference OR changes) (time OR temporal) (state OR event) ((hash-structure) OR (hash-structures)) -pat
49	(garbage OR trash OR expired) (management OR organization) (snapshot OR snapshots OR image OR images) (difference OR changes) (time OR temporal) (state OR event) ((hash-structure) OR (hash-structures)) -pat

Date search was completed: 09 March 2012

M.N.H./JG

Electronic Acknowledgement Receipt

EFS ID:	13947711
Application Number:	12947393
International Application Number:	
Confirmation Number:	1801
Title of Invention:	SYSTEM AND METHOD FOR PERFORMING BACKUP OR RESTORE OPERATIONS UTILIZING DIFFERENCE INFORMATION AND TIMELINE STATE INFORMATION
First Named Inventor/Applicant Name:	Philip J. ABERCROMBIE
Customer Number:	23483
Filer:	Michael Yasuhiro Saji/Kim LaRocca
Filer Authorized By:	Michael Yasuhiro Saji
Attorney Docket Number:	2203828.00130US1
Receipt Date:	10-OCT-2012
Filing Date:	16-NOV-2010
Time Stamp:	15:51:36
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Petition to make Special under PCT-Patent Pros Hwy	2203828_00130US1_Petition_to_Make_Special.PDF	6247641 aa9971845cb846d34a470e4be2de9571badc6114	no	46

Warnings:	
Information:	

ACT6002362

Total Files Size (in bytes):	6247641

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/947,393	11/16/2010	Philip J. ABERCROMBIE	2203828.00130US1	1801

23483 7590 10/12/2012

WILMERHALE/BOSTON
60 STATE STREET
BOSTON, MA 02109

EXAMINER
BANSAL, GURTEJ

ART UNIT	PAPER NUMBER
2189	

NOTIFICATION DATE	DELIVERY MODE
10/12/2012	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

teresa.carvalho@wilmerhale.com
whipusptopairs@wilmerhale.com

Office Action Summary	Application No.	Applicant(s)	
	12/947,393	ABERCROMBIE ET AL.	
	Examiner	Art Unit	
	GURTEJ BANSAL	2189	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *16 November 2010*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

4)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

5)☒ Claim(s) *1-8* is/are pending in the application.

5a) Of the above claim(s) _____ is/are withdrawn from consideration.

6)☐ Claim(s) _____ is/are allowed.

7)☒ Claim(s) *1-8* is/are rejected.

8)☐ Claim(s) _____ is/are objected to.

9)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

10)☐ The specification is objected to by the Examiner.

11)☒ The drawing(s) filed on *16 November 2010* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

12)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

13)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date *02/08/2012; 04/26/2012; 07/10/2012*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

1. Claims 1, 4, 5 and 8 are rejected under 35 U.S.C. 102(e) as being anticipated by Botes (US 7,640,454).

As per claim 1, Botes teaches a system for backing-up data from a first storage pool to a second storage pool using difference information between time states, said system comprising:

a data management engine for performing data management functions, including at least a back-up function to create a back-up copy of data (col. 6, lines 25-34 describes backing up application data and also see fig. 4a illustrating how data is backed up),

said data management engine operable to execute a sequence of snapshot operations to create point-in-time images of application data on a first storage pool (As illustrated in fig. 4a-4c and described in col. 7, line 33- col. 8, line 19 wherein changes to the application data state is backed up to persistent memory; col. 6, describes the image/data structure being stored as the configuration stored as a data structure

loadable from persistent memory, see IEEE Dictionary for definition of image), each successive point-in-time image corresponding to a specific, successive time-state of the application data (As illustrated in figs. 4a-4d, describes backing up at specific time states), and each snapshot operation creating difference information indicating which application data has changed and the content of the changed application data for the corresponding time state (col. 6, lines 26-30 describing how the configuration information backed up each time provides difference information on how application data has changed over time; also see col. 7, lines 64-67 describing how configuration information is stored at each state change);

said data management engine operable to execute at least one back-up function for the application data wherein the backup operation is scheduled for execution at non-consecutive time-states (col. 6, lines 24-26 describes wherein the backup is done in response to changes in states and not in response to recurring or consecutive time states);

wherein said data management engine is operable to maintain history information having time-state information indicating the time-state of the last back-up function performed on the application data for a corresponding back-up copy of data (col. 6, lines 39-45 and also see col. 6, line 33 describing wherein the backup occurs every hour and wherein the configuration information backed up occurred at T0, T1, T2,…Tn); and wherein the data management engine is operable to create composite difference information from the difference information for each time-state between the time-state of the last back-up function performed on the application data and the time-

state of the currently-scheduled back-up function to be performed on the application data (As described in col. 6, lines 1-15 describes wherein the configuration information stores the configuration of the application data at a specific point in time and the volatile memory stores the current configuration. The data stored in the volatile memory is composite difference information because it stores the changes which occurred between successive backups), and wherein the data management engine is operable to send the composite difference information to a second storage pool to be compiled with the back-up copy of data at the last time-state to create a back-up copy of data for the current time-state (col. 6, lines 7-10 describing wherein the composite information is sent to a different location in persistent storage).

As per claim 4, Botes teaches wherein multiple back-up functions are scheduled to occur simultaneously (col. 5, line 66-col. 6, line 10 describes how periodic backups occur and in between the time the periodic backups occur, backups with respect to state changes also occur thereby indicated that multiple backups functions are occurring simultaneously, i.e. periodic and state change), each with different gaps of non-consecutive time-states (As illustrated in col. 5, line 66- col. 6, line 10 and col. 6, lines 25-35, periodic backups occur hourly and state changes occur when a state is going to be changed), and each with different composite difference information generated corresponding to the different gaps (As described in col. 6, lines 5-10, the difference information being stored in volatile memory is only the configuration changes which occurred since the last backup and periodic backup only stores the configuration

information as it stands when the backup occurs and not the intermediate state changes).

As per claim 5, Botes teaches a system for restoring data in a storage pool from a back-up copy of the data using difference information between time states, said system comprising:

a data management engine wherein said data management engine is operable to maintain history information indicating the time-states for which storage pools have point-in-time images of application data (As illustrated in fig. 4a-4c and described in col. 7, line 33- col. 8, line 19 wherein changes to the application data state is backed up to persistent memory; col. 6, describes the image/data structure being stored as the configuration stored as a data structure loadable from persistent memory, see IEEE Dictionary for definition of image); and

wherein said data management engine includes logic for restoring application data in a storage pool to a point-in-time image of the data for a specified time-state (As illustrated in fig. 5);

said data management engine operable to identify the existence of a point-in-time image of the data at the storage pool for a time-state prior to the specified time-state and sending difference information from the back-up copy of data to the storage pool (As illustrated in figs. 4a-4d, describes backing up at specific time states), said difference information indicating which application data has changed and the content of the changed application data for the time between the specified time state and the time state prior to the specified time-state (col. 6, lines 26-30 describing how the

configuration information backed up each time provides difference information on how application data has changed over time; also see col. 7, lines 64-67 describing how configuration information is stored at each state change).

As per claim 8, Botes teaches wherein the prior time-state and the specified time-state are non-consecutive time-states (col. 6, lines 24-26 describes wherein the backup is done in response to changes in states and not in response to recurring or consecutive time states).

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

2. Claims 2 and 6 are rejected under 35 U.S.C. 103(a) as being unpatentable over Botes (US 7,640,454) as applied to claims 1 and 5, in view of Yagisawa et al. (US 2007/0162716) (hereafter Yagisawa).

As per claim 2, Botes teaches all the limitations of claim 1. Botes does not explicitly teach wherein difference information includes bitmap information with each bit of the bitmap corresponding to a portion of primary storage data, and including new data for those portions of the bitmap which are set to indicate that data has changed.

However, Yagisawa teaches wherein difference information includes bitmap information with each bit of the bitmap corresponding to a portion of primary storage

data, and including new data for those portions of the bitmap which are set to indicate that data has changed ([0081], bitmap table).

It would have been obvious to a person of ordinary skill in the art at the time of the invention to have combined the bitmap of Yagisawa with the system of Botes because it provides an indication of the position of data which has been updated and needs to be backed up ([0081]).

As per claim 6, Botes teaches all the limitations of claim 5. Botes does not explicitly teach wherein difference information includes bitmap information with each bit of the bitmap corresponding to a portion of primary storage data, and includes back-up data for those portions of the bitmap which are set to indicate that data has changed.

However, Yagisawa teaches wherein difference information includes bitmap information with each bit of the bitmap corresponding to a portion of primary storage data, and including new data for those portions of the bitmap which are set to indicate that data has changed ([0081], bitmap table).

It would have been obvious to a person of ordinary skill in the art at the time of the invention to have combined the bitmap of Yagisawa with the system of Botes because it provides an indication of the position of data which has been updated and needs to be backed up ([0081]).

3. Claims 3 and 7 are rejected under 35 U.S.C. 103(a) as being unpatentable over Botes as applied to claims 1 and 5 above, and further in view of Beatty et al. (US 2012/0078855) (hereafter Beatty).

As per claim 3, Botes teaches all the limitations of claim 1. Botes does not explicitly teach wherein difference information includes extent information.

However, Beatty teaches wherein difference information includes extent information ([0030]).

It would have been obvious to a person of ordinary skill in the art at the time of the invention to have combined the extent information of Beatty with the system of Botes because it enables the system to know which discrete blocks of storage within a database have been updated ([0030]).

As per claim 7, Botes teaches all the limitations of claim 5. Botes does not explicitly teach wherein difference information includes extent information.

However, Beatty teaches wherein difference information includes extent information ([0030]).

It would have been obvious to a person of ordinary skill in the art at the time of the invention to have combined the extent information of Beatty with the system of Botes because it enables the system to know which discrete blocks of storage within a database have been updated ([0030]).

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to GURTEJ BANSAL whose telephone number is (571)270-5588. The examiner can normally be reached on Monday - Friday, 7:30 a.m. - 5:00 p.m., EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Reginald Bragdon can be reached on (571)272-4204. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/GB/

/Jared I Rutz/
Primary Examiner, Art Unit 2187

EAST Search History

EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	1	"20120078855".pn.	US-PGPUB; USPAT	OR	OFF	2012/10/04 17:03
S1	6250	711/161.ccls. 711/162.ccls.	US-PGPUB; USPAT	OR	OFF	2012/09/24 16:21
S2	745	point$1in$1time with snapshot and back$up	US-PGPUB; USPAT	OR	OFF	2012/09/24 16:23
S3	28	point$1in$1time with snapshot same difference and back$up	US-PGPUB; USPAT	OR	OFF	2012/09/24 16:23
S4	30	point$1in$1time same snapshot same difference and back$up	US-PGPUB; USPAT	OR	OFF	2012/09/24 16:31
S5	88	point$1in$1time same snapshot and difference with information and back$up	US-PGPUB; USPAT	OR	OFF	2012/09/24 16:35
S6	14	"7065619".pn. "20020049778".pn. "20080034016".pn. "20090307251".pn. "20100077013".pn. "20100088277".pn. "20100138827".pn. "20100276744".pn. "20110179341".pn. "20110307447".pn. "20110307683".pn. "20120017060".pn. "6883073".pn. "7814149".pn.	US-PGPUB; USPAT	OR	OFF	2012/09/24 16:54
S7	15	"7065619".pn. "20020049778".pn. "20080034016".pn. "20090307251".pn. "20100077013".pn. "20100088277".pn. "20100138827".pn. "20100276744".pn. "20110179341".pn. "20110307447".pn. "20110307683".pn. "20120017060".pn. "6883073".pn. "7814149".pn. "20110252189".pn.	US-PGPUB; USPAT	OR	OFF	2012/09/24 16:54
S8	157	snapshot and difference with bitmap	US-PGPUB; USPAT	OR	OFF	2012/09/27 00:18
S9	17	snapshot and difference with bitmap and deduplication	US-PGPUB; USPAT	OR	OFF	2012/09/27 00:20
S10	11	snapshot and update with bitmap and deduplication	US-PGPUB; USPAT	OR	OFF	2012/09/27 00:25
S11	11	snapshot with bitmap and deduplication	US-PGPUB; USPAT	OR	OFF	2012/09/27 00:26
S12	1	"20120124307".pn.	US-PGPUB; USPAT	OR	OFF	2012/09/27 12:04

S13	1	"20120124307".pn. and extent	US-PGPUB; USPAT	OR	OFF	2012/09/27 12:05
S14	10	deduplication with capacity with performance	US-PGPUB; USPAT	OR	OFF	2012/09/27 12:28
S15	53	deduplication same capacity same performance	US-PGPUB; USPAT	OR	OFF	2012/09/27 12:29
S16	130	deduplication with capacity	US-PGPUB; USPAT	OR	OFF	2012/09/27 12:30
S17	8	deduplication with capacity same (efficiency speed)	US-PGPUB; USPAT	OR	OFF	2012/09/27 12:31
S18	25	difference with extent same (backup back$1up back adj up)	US-PGPUB; USPAT	OR	OFF	2012/09/27 13:02
S19	6	bitmap with extent same (backup back$1up back adj up)	US-PGPUB; USPAT	OR	OFF	2012/09/27 13:02
S20	1	"7640454".pn.	US-PGPUB; USPAT	OR	OFF	2012/10/04 13:01

EAST Search History (Interference)

< This search history is empty>

10/4/2012 5:06:38 PM
C:\Users\gbansal\Documents\EAST\Workspaces\12947393.wsp

Atmospheric Correction at AERONET Locations: A New Science and Validation Data Set

Yujie Wang; Lyapustin, A.I.; Privette, J.L.; Morisette, J.T.; Holben, B.

Geoscience and Remote Sensing, IEEE Transactions on

Volume: 47 , Issue: 8 , Part: 1

Digital Object Identifier: 10.1109/TGRS.2009.2016334

Publication Year: 2009 , Page(s): 2450 - 2466

Cited by 7

IEEE JOURNALS & MAGAZINES

| Quick Abstract | PDF (1624 KB)

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

BIB DATA SHEET

CONFIRMATION NO. 1801

SERIAL NUMBER	FILING or 371(c) DATE	CLASS	GROUP ART UNIT	ATTORNEY DOCKET NO.
12/947,393	11/16/2010 RULE	711	2189	2203828.00130US1

APPLICANTS

Philip J. ABERCROMBIE, Belmont, MA;
Madhav MUTALIK, Southborough, MA;
Christopher A. PROVENZANO, Somerville, MA;
Mark A. ROMAN, Arlington, MA;

** CONTINUING DATA **********

** FOREIGN APPLICATIONS **********

** IF REQUIRED, FOREIGN FILING LICENSE GRANTED ** ** SMALL ENTITY **
11/30/2010

Foreign Priority claimed	35 USC 119(a-d) conditions met	Verified and Acknowledged	Met after Allowance	STATE OR COUNTRY	SHEETS DRAWINGS	TOTAL CLAIMS	INDEPENDENT CLAIMS
☐ Yes ☑ No	☐ Yes ☑ No	/GURTEJ BANSAL/ Examiner's Signature	☐ Initials	MA	17	8	2

ADDRESS

WILMERHALE/BOSTON
60 STATE STREET
BOSTON, MA 02109
UNITED STATES

TITLE

SYSTEM AND METHOD FOR PERFORMING BACKUP OR RESTORE OPERATIONS UTILIZING DIFFERENCE INFORMATION AND TIMELINE STATE INFORMATION

FILING FEE RECEIVED		
462	FEES: Authority has been given in Paper No._____ to charge/credit DEPOSIT ACCOUNT No._____ for following:	☐ All Fees ☐ 1.16 Fees (Filing) ☐ 1.17 Fees (Processing Ext. of time) ☐ 1.18 Fees (Issue) ☐ Other _____ ☐ Credit

BIB (Rev. 05/07).

ACT6002377

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	12947393	ABERCROMBIE ET AL.
	Examiner	Art Unit
	GURTEJ BANSAL	2189

SEARCHED			
Class	Subclass	Date	Examiner
711	161; 162	10/04/2012	GB

SEARCH NOTES		
Search Notes	Date	Examiner
Text Search in EAST (See Attached)	10/04/2012	GB
NPL Search in IEEE (See Attached)	10/04/2012	GB
Inventor Search in eDAN	10/04/2012	GB

INTERFERENCE SEARCH			
Class	Subclass	Date	Examiner

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	12947393	ABERCROMBIE ET AL.
	Examiner	**Art Unit**
	GURTEJ BANSAL	2189

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	10/04/2012								
	1	✓								
	2	✓								
	3	✓								
	4	✓								
	5	✓								
	6	✓								
	7	✓								
	8	✓								

PTO/SB/08b (07-09)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449/PTO

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 1 of 1

Complete if Known	
Application Number	12/947,393-Conf. #1801
Filing Date	November 16, 2010
First Named Inventor	Philip J. ABERCROMBIE
Art Unit	2189
Examiner Name	G. Bansal
Attorney Docket Number	2203828.00130US1

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
	AA*	US-20110252198	10-13-2011	Ogasawara et al.	

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document Country Code[3]-Number[4]-Kind Code[5] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages Or Relevant Figures Appear	T[6]

NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]

Examiner Signature	/Gurtej Bansal/ (10/04/2012)	Date Considered	10/04/2012

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. * CITE NO.: Those application(s) which are marked with an single asterisk (*) next to the Cite No. are not supplied (under 37 CFR 1.98(a)(2)(iii)) because that application was filed after June 30, 2003 or is available in the IFW. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

ACTIVEUS 98331009v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /G.B./

ACT6002380

PTO/SB/08b (07-09)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 1 of 1

Complete if Known	
Application Number	12/947,393-Conf. #1801
Filing Date	November 16, 2010
First Named Inventor	Philip J. ABERCROMBIE
Art Unit	2189
Examiner Name	R. G. Bragdon
Attorney Docket Number	2203828.00130US1

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
	AA*	US-7,065,619	06-20-2006	Zhu et al.	

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document Country Code[3]-Number[4]-Kind Code[5] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages Or Relevant Figures Appear	T[6]

NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
	CA	"EMC Data Domain Replicator," White Paper, EMC Corporation 2011 (25 pages)	
	CB	Guo, F. and Efstathopoulos, P., "Building a High-performance Deduplication System," Symantec Research Labs, Symantec Corporation, Culver City, CA, USA (14 pages) (2011)	
	CC	Silverberg, S. "SDFS Overview," April 2010 (17 pages)	
	CD	Tridgell, A., "Efficient Algorithms for Sorting and Synchronization," Thesis, The Australian National University (Feb. 1999) (115 pages)	
	CE	What is Deduplication and Why Does It Matter?, dated May 7, 2010 (2 pages)	
	CF	Zhu, B. et al., "Avoiding the Disk Bottleneck in the Data Domain Deduplication File System," Fast '08: 6th USENIX Conference on File and Storage Technologies, USENIX Association, pp: 269-282 (2008)	

Examiner Signature	/Gurtej Bansal/ (10/04/2012)	Date Considered	10/04/2012

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. * CITE NO.: Those application(s) which are marked with an single asterisk (*) next to the Cite No. are not supplied (under 37 CFR 1.98(a)(2)(iii)) because that application was filed after June 30, 2003 or is available in the IFW. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

ACTIVEUS 92474953v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /G.B./

ACT6002381

PTO/SB/08b (07-09)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Substitute for form 1449/PTO

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 1 of 1

Complete if Known	
Application Number	12/947,393-Conf. #1801
Filing Date	November 16, 2010
First Named Inventor	Philip J. ABERCROMBIE
Art Unit	2189
Examiner Name	R. G. Bragdon
Attorney Docket Number	2203828.00130US1

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
	AA*	US-20020049778	04-25-2002	Bell et al.	
	AB*	US-20080034016	02-07-2008	Cisler et al.	
	AC*	US-20090307251	12-10-2009	Heller et al.	
	AD*	US-20100077013	03-25-2010	CLEMENTS et al.	
	AE*	US-20100088277	04-08-2010	Rao et al.	
	AF*	US-20100138827	06-03-2010	Frank et al.	
	AG*	US-20100276744	11-04-2010	Lee	
	AH*	US-20110179341	07-21-2011	Falls et al.	
	AI*	US-20110307447	12-15-2011	Sabaa et al.	
	AJ*	US-20110307683	12-15-2011	SPACKMAN	
	AK*	US-20120017060	01-19-2012	Kapanipathi et al.	
	AL*	US-6,883,073	04-19-2005	Arakawa et al.	
	AM*	US-7,814,149	10-12-2010	Stringham	

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document Country Code[3]-Number[4]-Kind Code[5] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages Or Relevant Figures Appear	T[6]

NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
	CA	International Search Report issued for PCT/US2011/060417, dated March 16, 2012 (2 pages)	

Examiner Signature	/Gurtej Bansal/ (10/04/2012)	Date Considered	10/04/2012

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. * CITE NO.: Those application(s) which are marked with an single asterisk (*) next to the Cite No. are not supplied (under 37 CFR 1.98(a)(2)(iii)) because that application was filed after June 30, 2003 or is available in the IFW. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

ACTIVEUS 95222698v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /G.B./

ACT6002382

Notice of References Cited	Application/Control No. 12/947,393	Applicant(s)/Patent Under Reexamination ABERCROMBIE ET AL.	
	Examiner GURTEJ BANSAL	Art Unit 2189	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
*	A	US-7,640,454	12-2009	Botes, Par	714/19
*	B	US-2007/0162716	07-2007	Yagisawa et al.	711/162
*	C	US-2012/0078855	03-2012	Beatty et al.	707/676
	D	US-			
	E	US-			
	F	US-			
	G	US-			
	H	US-			
	I	US-			
	J	US-			
	K	US-			
	L	US-			
	M	US-			

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	IEEE 100, The Authoritative Dictionary of IEEE Standards Terms, Seventh Edition, 2000, the Institute of Electrical and Electronics Engineering, Inc., page 532
	V	
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

IEEE 100
The Authoritative Dictionary of IEEE Standards Terms

Seventh Edition

Published by
Standards Information Network
IEEE Press

Trademarks and disclaimers

IEEE believes the information in this publication is accurate as of its publication date; such information is subject to change without notice. IEEE is not responsible for any inadvertent errors.

Other tradenames and trademarks in this document are those of their respective owners.

The Institute of Electrical and Electronics Engineering, Inc.
3 Park Avenue, New York, NY, 10016-5997, USA

To order IEEE Press publications, call 1-800-678-IEEE.

Print: ISBN 0-7381-2601-2 *SP1122*

See other standards and standards-related product listings at: http://standards.ieee.org/

The publisher believes that the information and guidance given in this work serve as an enhancement to users, all parties must rely upon their own skill and judgement when making use of it. The publisher does not assume any liability to anyone for any loss or damage caused by any error or omission in the work, whether such error or omission is the result of negligence or any other cause. Any and all such liability is disclaimed.

This work is published with the understanding that the IEEE is supplying information through this publication, not attempting to render engineering or other professional services. If such services are required, the assistance of an appropriate professional should be sought. The IEEE is not responsible for the statements and opinions advanced in this publication.

ignitron A single-anode pool tube in which an ignitor is employed to initiate the cathode spot before each conducting period. *See also:* electronic controller. (ED) 161-1971w

ignored Used to describe an instruction field, the contents of which are arbitrary and have no effect on the execution of the instruction. The contents of an ignored field will continue to be ignored in future versions of the architecture. *See also:* unused; reserved. (C/MM) 1754-1994

ignored conductor *See:* isolated conductor.

IH *See:* intermediate hub.

ihandle A cell-sized datum identifying a particular package instance. (C/BA) 1275-1994

IIL *See:* integrated injection logic.

IITRAN A programming language similar to PL/1; designed for use as an educational tool. (C) 610.13-1993w

ILD *See:* injection laser diode.

illegal character A character or combination of bits that is not valid according to some criteria; for example, a character that is not a member of some specified alphabet. *Synonyms:* forbidden character; improper character. *Contrast:* forbidden combination. (C) 610.5-1990w

Illegal Command (ILC) bit A bit in the Bus Error register of all S-modules. An S-module sets this bit to indicate that the module has received an illegal command. (TT/C) 1149.5-1995

Illegal Port Selected (IPS) bit A bit in the Bus Error register of all S-modules. An S-module sets this bit to indicate that the module has received a command addressed to an unsupported port. (TT/C) 1149.5-1995

illuminance The unit density of light flux (lm/unit area) that is incident on a surface. (IA/PSE) 241-1990r

illuminance, $E = d\Phi/dA$ (illuminating engineering) The density of the luminous flux incident at a point on a surface. Average illuminance is the quotient of the luminous flux incident on a surface by the area of the surface. (EEC/IE) [126]

illuminance (footcandle or lux) meter (1) (illuminating engineering) An instrument for measuring illuminance on a plane. Instruments that accurately respond to more than one spectral distribution are color corrected; that is, the spectral response is balanced to $V(\lambda)$ or $V'(\lambda)$. Instruments that accurately respond to more than one spatial distribution of incident flux are cosine corrected; that is, the response to a source of unit luminous intensity, illuminating the detector from a fixed distance and from different directions decreases as the cosine of the angle between the incident direction and the normal to the detector surface. The instrument is comprised of some form of photodetector with or without a filter driving a digital or analog readout through appropriate circuitry. (EEC/IE) [126]

(2) (television) *See also:* illumination. (BT/AV) 201-1979w

illumination* (1) (illuminating engineering) An alternate, but deprecated, term for illuminance. It is frequently used since illuminance is subject to confusion with luminance and illuminants, especially when not clearly pronounced. *Note:* The term illumination also is commonly used in a qualitative or general sense to designate the act of illuminating or the state of being illuminated. Usually the context will indicate which meaning is intended, but occasionally it is desirable to use the expression level of illumination to indicate that the quantitative meaning is intended. (EEC/IE) [126]

(2) (A) (television) (general) The density of the luminous flux incident on a surface; it is the quotient of the luminous flux by the area of the surface when the latter is uniformly illuminated. **(B) (television)** (at a point of a surface) The quotient of the luminous flux incident on an infinitesimal element of surface containing the point under consideration by the area of that element. *Notes:* 1. The term illumination also is commonly used in a qualitative or general sense to designate the act of illuminating or the state of being illuminated. Usually the context will indicate which meaning is intended, but occasionally it is desirable to use the expression level of illumination to indicate that the quantitative meaning is intended. The term illuminance, which sometimes is used in place of illumination, is subject to confusion with luminance and illuminates, especially when not clearly pronounced. 2. The units of measurements are: footcandle (lumen per square foot, lm/ft^2 lux (lumen per square meter, lx or lm/m^2). This unit of illumination is recommended by the IEC phot (lumen per square centimeter, lm/cm^2). (BT/ED/AV) 201-1979, [127]

(3) *See also:* aperture illumination. (ANT)

* Deprecated.

illumination (footcandle) meter An instrument for measuring the illumination on a surface. *Note:* Most such instruments consist of barrier-layer cells connected to a meter calibrated in footcandles. *See also:* photometry. (EEC/IE) [126]

illumination sensitivity (camera tubes or phototubes) The quotient of signal output current by the incident illumination, under specified conditions of illumination. *Notes:* 1. Since illumination sensitivity is not an absolute characteristic but depends on the spectral distribution of the incident flux, the term is commonly used to designate the sensitivity to radiation from a tungsten-filament lamp operating at a color temperature of 2870 K. 2. Illumination sensitivity is usually measured with a collimated beam at normal incidence. *See also:* transfer characteristic. (ED) 161-1971w

illuminator That part of a semiactive guidance missile weapon system that radiates electromagnetic waves in the direction of a designated target so that echo signals reflected from the illuminated target can be used by another sensor (the missile seeker) for purposes of homing. (AES) 686-1997

illustration Material that is labeled, numbered, set apart from the main body of text, and, normally, cited within the main text. (C/SE) 1063-1987r

illustrative diagram A diagram whose principal purpose is to show the operating principle of a device or group of devices without necessarily showing actual connections or circuits. Illustrative diagrams may use pictures or symbols to illustrate or represent devices or their elements. Illustrative diagrams may be made of electric, hydraulic, pneumatic, and combination systems. They are applicable chiefly to instruction books, descriptive folders, or other media whose purpose is to explain or instruct. *See also:* control. (IA/ICTL/IAC) 270-1966w, [60]

ILS *See:* instrument landing system.

ILS reference point A point on the centerline of the ILS runway designated as the optimum point of contact for landing: in International Civil Aviation Organization standards this point is from 150 to 300 meters (500 to 1000 feet) from the approach end of the runway. (AES/RS) 686-1982s

image (1) (optoelectronic device) A spatial distribution of a physical property, such as radiation, electric charge, conductivity, or reflectivity, mapped from another distribution of either the same or another physical property. *Note:* The mapping process may be carried out by a flux of photons, electric charges, or other means. *See also:* optoelectronic device. (ED) [46]

(2) (computer graphics) A displayed or drawn representation. (C) 610.6-1991w

(3) (image processing and pattern recognition) A two-dimensional representation of a scene. *Synonym:* picture. *See also:* digital image. (C) 610.4-1990w

(4) (A) In image processing, a two-dimensional representation of a scene. **(B)** In graphics, a displayed or drawn representation. (C) 610.10-1994

(5) *See also:* card image.

Image The data structure contained in the Load Server that the Loadable Device wishes to load. (C) 15802-4-1994

image analysis The process of describing or evaluating an image in terms of its parts, properties, and relationships. (C) 610.4-1990w

UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

MAILED

NOV 20 2012

PCT LEGAL ADMINISTRATION

WILMERHALE/BOSTON
60 STATE STREET
BOSTON MA 02109

In re Application of:	:	
ABERCROMBIE, Philip, J., et al.	:	**DECISION ON RENEWED**
Application No.: 12/947,393	:	**REQUEST TO PARTICIPATE IN**
Filing Date: 16 November 2010	:	**THE PATENT PROSECUTION**
Attorney's Docket No.: 2203828.00130US1	:	**HIGHWAY AND PETITION TO**
For: SYSTEM AND METHOD FOR	:	**MAKE SPECIAL UNDER 37**
PERFORMING BACKUP OR	:	**CFR 1.102(a)**
RESTORE OPERATIONS ...	:	

This is a decision on the renewed request to participate in the PCT Patent Prosecution Highway (PCT-PPH) pilot program and the petition under 37 CFR 1.102(a), filed on 10 October 2012, to make the above-identified application special.

The renewed request and petition are **DISMISSED**.

Examination of the above-identified U.S. application has begun and therefore the renewed request from applicant to participate in the PPH program is appropriately dismissed. A non-final Office action was mailed 12 October 2012 in the above-identified application.

Telephone inquiries concerning this decision should be directed to the undersigned.

All other inquiries concerning the examination or status of the application is accessible in the PAIR system at http://www.uspto.gov/ebc/index.html.

/RichardMRoss/

Richard M. Ross
Attorney Advisor
Office of PCT Legal Administration
Tel.: (571) 272-3296

Docket No.: 2203828.00130US1
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	Philip J. Abercrombie et al.	Confirmation No.:	1801
Application No.:	12/947,393	Art Unit:	2189
Filed:	November 16, 2010	Examiner:	G. Bansal
Title:	SYSTEM AND METHOD FOR PERFORMING BACKUP OR RESTORE OPERATIONS UTILIZING DIFFERENCE INFORMATION AND TIMELINE STATE INFORMATION		

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT (IDS)

Dear Sir:

This Supplemental Information Disclosure Statement is being filed after the mailing date of the first Office Action on the merits and before the mailing date of a final Office Action or a Notice of Allowance. The $180.00 fee is included herewith. The Information Disclosure Statement is not a response to the outstanding Office Action, a proper response will be filed shortly.

Applicants request that the Examiner initial and return a copy of the enclosed Form PTO SB-08 with the next communication. Please charge the required fee to our Deposit Account No. 08-0219, under Order No. 2203828.00130US1 from which the undersigned is authorized to draw.

Respectfully submitted,

Dated: February 14, 2013

/Michael Saji/
Michael Y. Saji
Registration No.: 66,291
Attorney for Applicant(s)

Wilmer Cutler Pickering Hale and Dorr LLP
60 State Street
Boston, Massachusetts 02109
(617) 526-6000 (telephone)
(617) 526-5000 (facsimile)

Electronic Acknowledgement Receipt

EFS ID:	14958746
Application Number:	12947393
International Application Number:	
Confirmation Number:	1801
Title of Invention:	SYSTEM AND METHOD FOR PERFORMING BACKUP OR RESTORE OPERATIONS UTILIZING DIFFERENCE INFORMATION AND TIMELINE STATE INFORMATION
First Named Inventor/Applicant Name:	Philip J. ABERCROMBIE
Customer Number:	23483
Filer:	Michael Yasuhiro Saji/Kathleen Bastarache
Filer Authorized By:	Michael Yasuhiro Saji
Attorney Docket Number:	2203828.00130US1
Receipt Date:	14-FEB-2013
Filing Date:	16-NOV-2010
Time Stamp:	16:45:51
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Credit Card
Payment was successfully received in RAM	$180
RAM confirmation Number	4044
Deposit Account	080219
Authorized User	SAJI, MICHAEL Y.

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

Charge any Additional Fees required under 37 C.F.R. Section 1.16 (National application filing, search, and examination fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.19 (Document supply fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.20 (Post Issuance fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.21 (Miscellaneous fees and charges)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Non Patent Literature	IEEE100_TheAuthoritativeDictionary_2000.PDF	368324 f1ab77d213facf518c8ba02b02bd40cf5ec59bbd	no	3
Warnings:					
Information:					
2	Transmittal Letter	130US1_IDS_14Feb2013.PDF	85683 2bf07886ecbf8c107ab0de28a5cceafb72d94115	no	1
Warnings:					
Information:					
3	Information Disclosure Statement (IDS) Form (SB08)	130US1_SB08_14Feb2013.PDF	110098 51ce372a727155050c4ad2b239b004fcc49ace3e	no	1
Warnings:					
Information:					
This is not an USPTO supplied IDS fillable form					
4	Fee Worksheet (SB06)	fee-info.pdf	30659 000dbaab40d2ffde26ca1758fd8696355b2f36b2	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	594764		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

Electronic Patent Application Fee Transmittal

Application Number:	12947393
Filing Date:	16-Nov-2010
Title of Invention:	SYSTEM AND METHOD FOR PERFORMING BACKUP OR RESTORE OPERATIONS UTILIZING DIFFERENCE INFORMATION AND TIMELINE STATE INFORMATION
First Named Inventor/Applicant Name:	Philip J. ABERCROMBIE
Filer:	Michael Yasuhiro Saji/C. Schroeter
Attorney Docket Number:	2203828.00130US1

Filed as Small Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
Submission- Information Disclosure Stmt	1806	1	180	180
		Total in USD ($)		**180**

PTO/SB/08b (07-09)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 1 of 1

Complete if Known	
Application Number	12/947,393-Conf. #1801
Filing Date	November 16, 2010
First Named Inventor	Philip J. ABERCROMBIE
Art Unit	2189
Examiner Name	G. Bansal
Attorney Docket Number	2203828.00130US1

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
	AA*	US-20060080367	04-13-2006	Pudipeddi	
	AB*	US-20070162716	07-12-2007	Yagisawa et al.	
	AC*	US-20080243769	10-02-2008	Arbour et al.	
	AD*	US-20120078855	03-29-2012	Beatty et al.	
	AE*	US-7,640,454	12-29-2012	Botes	

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document Country Code[3]-Number[4]-Kind Code[5] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages Or Relevant Figures Appear	T[6]

NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
	CA	IEEE 100, The Authoritative Dictionary of IEEE Standards Terms, Seventh Edition, 2000, the Institute of Electrical and Electronics Engineering, Inc. page 532	

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. * CITE NO.: Those application(s) which are marked with an single asterisk (*) next to the Cite No. are not supplied (under 37 CFR 1.98(a)(2)(iii)) because that application was filed after June 30, 2003 or is available in the IFW. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

IEEE 100
The Authoritative Dictionary of IEEE Standards Terms

Seventh Edition

Published by
Standards Information Network
IEEE Press

The Institute of Electrical and Electronics Engineering, Inc.
3 Park Avenue, New York, NY, 10016-5997, USA

To order IEEE Press publications, call 1-800-678-IEEE.

Print: ISBN 0-7381-2601-2 *SP1122*

See other standards and standards-related product listings at: http://standards.ieee.org/

ignitron A single-anode pool tube in which an ignitor is employed to initiate the cathode spot before each conducting period. *See also:* electronic controller. (ED) 161-1971w

ignored Used to describe an instruction field, the contents of which are arbitrary and have no effect on the execution of the instruction. The contents of an ignored field will continue to be ignored in future versions of the architecture. *See also:* unused; reserved. (C/MM) 1754-1994

ignored conductor *See:* isolated conductor.

IH *See:* intermediate hub.

ihandle A cell-sized datum identifying a particular package instance. (C/BA) 1275-1994

IIL *See:* integrated injection logic.

IITRAN A programming language similar to PL/1; designed for use as an educational tool. (C) 610.13-1993w

ILD *See:* injection laser diode.

illegal character A character or combination of bits that is not valid according to some criteria; for example, a character that is not a member of some specified alphabet. *Synonyms:* forbidden character; improper character. *Contrast:* forbidden combination. (C) 610.5-1990w

Illegal Command (ILC) bit A bit in the Bus Error register of all S-modules. An S-module sets this bit to indicate that the module has received an illegal command. (TT/C) 1149.5-1995

Illegal Port Selected (IPS) bit A bit in the Bus Error register of all S-modules. An S-module sets this bit to indicate that the module has received a command addressed to an unsupported port. (TT/C) 1149.5-1995

illuminance The unit density of light flux (lm/unit area) that is incident on a surface. (IA/PSE) 241-1990r

illuminance, $E = d\Phi/dA$ (illuminating engineering) The density of the luminous flux incident at a point on a surface. Average illuminance is the quotient of the luminous flux incident on a surface by the area of the surface. (EEC/IE) [126]

illuminance (footcandle or lux) meter (1) (illuminating engineering) An instrument for measuring illuminance on a plane. Instruments that accurately respond to more than one spectral distribution are color corrected; that is, the spectral response is balanced to $V(\lambda)$ or $V'(\lambda)$. Instruments that accurately respond to more than one spatial distribution of incident flux are cosine corrected; that is, the response to a source of unit luminous intensity, illuminating the detector from a fixed distance and from different directions decreases as the cosine of the angle between the incident direction and the normal to the detector surface. The instrument is comprised of some form of photodetector with or without a filter driving a digital or analog readout through appropriate circuitry. (EEC/IE) [126]

(2) (television) *See also:* illumination. (BT/AV) 201-1979w

illumination* (1) (illuminating engineering) An alternate, but deprecated, term for illuminance. It is frequently used since illuminance is subject to confusion with luminance and illuminants, especially when not clearly pronounced. *Note:* The term illumination also is commonly used in a qualitative or general sense to designate the act of illuminating or the state of being illuminated. Usually the context will indicate which meaning is intended, but occasionally it is desirable to use the expression level of illumination to indicate that the quantitative meaning is intended. (EEC/IE) [126]

(2) (A) (television) (general) The density of the luminous flux incident on a surface; it is the quotient of the luminous flux by the area of the surface when the latter is uniformly illuminated. **(B) (television)** (at a point of a surface) The quotient of the luminous flux incident on an infinitesimal element of surface containing the point under consideration by the area of that element. *Notes:* 1. The term illumination also is commonly used in a qualitative or general sense to designate the act of illuminating or the state of being illuminated. Usually the context will indicate which meaning is intended, but occasionally it is desirable to use the expression level of illumination to indicate that the quantitative meaning is intended. The term illuminance, which sometimes is used in place of illumination, is subject to confusion with luminance and illuminates, especially when not clearly pronounced. 2. The units of measurements are: footcandle (lumen per square foot, lm/ft^2 lux (lumen per square meter, lx or lm/m^2). This unit of illumination is recommended by the IEC phot (lumen per square centimeter, lm/cm^2). (BT/ED/AV) 201-1979, [127]

(3) *See also:* aperture illumination. (ANT)

* Deprecated.

illumination (footcandle) meter An instrument for measuring the illumination on a surface. *Note:* Most such instruments consist of barrier-layer cells connected to a meter calibrated in footcandles. *See also:* photometry. (EEC/IE) [126]

illumination sensitivity (camera tubes or phototubes) The quotient of signal output current by the incident illumination, under specified conditions of illumination. *Notes:* 1. Since illumination sensitivity is not an absolute characteristic but depends on the spectral distribution of the incident flux, the term is commonly used to designate the sensitivity to radiation from a tungsten-filament lamp operating at a color temperature of 2870 K. 2. Illumination sensitivity is usually measured with a collimated beam at normal incidence. *See also:* transfer characteristic. (ED) 161-1971w

illuminator That part of a semiactive guidance missile weapon system that radiates electromagnetic waves in the direction of a designated target so that echo signals reflected from the illuminated target can be used by another sensor (the missile seeker) for purposes of homing. (AES) 686-1997

illustration Material that is labeled, numbered, set apart from the main body of text, and, normally, cited within the main text. (C/SE) 1063-1987r

illustrative diagram A diagram whose principal purpose is to show the operating principle of a device or group of devices without necessarily showing actual connections or circuits. Illustrative diagrams may use pictures or symbols to illustrate or represent devices or their elements. Illustrative diagrams may be made of electric, hydraulic, pneumatic, and combination systems. They are applicable chiefly to instruction books, descriptive folders, or other media whose purpose is to explain or instruct. *See also:* control. (IA/ICTL/IAC) 270-1966w, [60]

ILS *See:* instrument landing system.

ILS reference point A point on the centerline of the ILS runway designated as the optimum point of contact for landing; in International Civil Aviation Organization standards this point is from 150 to 300 meters (500 to 1000 feet) from the approach end of the runway. (AES/RS) 686-1982s

image (1) (optoelectronic device) A spatial distribution of a physical property, such as radiation, electric charge, conductivity, or reflectivity, mapped from another distribution of either the same or another physical property. *Note:* The mapping process may be carried out by a flux of photons, electric charges, or other means. *See also:* optoelectronic device. (ED) [46]

(2) (computer graphics) A displayed or drawn representation. (C) 610.6-1991w

(3) (image processing and pattern recognition) A two-dimensional representation of a scene. *Synonym:* picture. *See also:* digital image. (C) 610.4-1990w

(4) (A) In image processing, a two-dimensional representation of a scene. **(B)** In graphics, a displayed or drawn representation. (C) 610.10-1994

(5) *See also:* card image.

Image The data structure contained in the Load Server that the Loadable Device wishes to load. (C) 15802-4-1994

image analysis The process of describing or evaluating an image in terms of its parts, properties, and relationships. (C) 610.4-1990w

Docket No.: 2203828.00130US1
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	Philip J. Abercrombie et al.	Confirmation No.:	1801
Application No.:	12/947,393	Art Unit:	2189
Filed:	November 16, 2010	Examiner:	G. Bansal
Title:	SYSTEM AND METHOD FOR PERFORMING BACKUP OR RESTORE OPERATIONS UTILIZING DIFFERENCE INFORMATION AND TIMELINE STATE INFORMATION		

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

AMENDMENT IN RESPONSE TO NON-FINAL OFFICE ACTION UNDER 37 C.F.R. 1.111

Dear Madam:

INTRODUCTORY COMMENTS

In response to the Office Action dated October 12, 2012, please amend the above-identified U.S. patent application as follows:

Amendments to the Specification begin on page 2 of this paper.

Amendments to the Claims are reflected in the listing of claims which begins on page 3 of this paper.

Remarks/Arguments begin on page 6 of this paper.

AMENDMENTS TO THE SPECIFICATION

Please amend Paragraph [0001] as follows:

[0001] This application is related to the following applications, filed herewith and hereby incorporated by reference:

- "System and Method for Managing Data with Service Level Agreements That May Specify Non-Uniform Copying of Data" (U.S. Application No. ~~TBD~~ 12/947,385);

- "System and Method for Performing a Plurality of Prescribed Data Management Functions in a Manner That Reduces Redundant Access Operations to Primary Storage" (U.S. Application No. ~~TBD~~ 12/947,436);

- "System and Method for Creating Deduplicated Copies of Data by Tracking Temporal Relationships Among Copies and by Ingesting Difference Data " (U.S. Application No. ~~TBD~~ 12/947,418);

- "System and Method for Managing Deduplicated Copies of Data Using Temporal Relationships Among Copies " (U.S. Application No. ~~TBD~~ 12/947,375);

- "System and Method for Creating Deduplicated Copies of Data by Sending Difference Data Between Two Near-Neighbor Temporal States" (U.S. Application No. ~~TBD~~ 12/947,513);

- "System and Method for Creating Deduplicated Copies of Data Storing Non-Lossy Encodings of Data Directly in a Content Addressable Store" (U.S. Application No. ~~TBD~~ 12/947,438); and

- "System and Method for Improved Garbage Collection Operations in a Deduplicated Store by Tracking Temporal Relationships Among Copies" (U.S. Application No. ~~TBD~~ 12/947,383).

AMENDMENTS TO THE CLAIMS

1. (Cancelled)

2. (Currently Amended) The system of claim 9[[1]], wherein difference information includes bitmap information with each bit of the bitmap corresponding to a portion of primary storage data, and including new data for those portions of the bitmap which are set to indicate that data has changed.

3. (Currently Amended) The system of claim 9[[1]], wherein difference information includes extent information.

4. (Currently Amended) The system of claim 9[[1]], wherein multiple back-up functions are scheduled to occur simultaneously, each with different gaps of non-consecutive time-states, and each with different composite difference information generated corresponding to the different gaps.

5. (Currently Amended) A system for restoring data in a storage pool from a back-up copy of the data using difference information between time states, said system comprising:

a data management engine wherein said data management engine is operable to

maintain history information indicating ~~the~~ time-states for which storage pools have point-in-time images of application data; ~~and~~

~~wherein said data management engine includes logic for restoring~~

restore application data in a storage pool to a point-in-time image of the data for a specified time-state;

~~said data management engine operable to~~ identify the existence of a prior point-in-time image of the data at the storage pool for a time-state prior to the specified time-state,

create composite difference information from difference information for each point-in-time image of the data with a time-state between the specified time-state and the prior time-state, and ~~sending~~

send the composite difference information from the back-up copy of data to the storage pool, said difference information indicating which application data has changed and the content of the changed application data for the time between the specified time state and the time state prior to the specified time-state.

6. (Original) The system of claim 5, wherein difference information includes bitmap information with each bit of the bitmap corresponding to a portion of primary storage data, and includes back-up data for those portions of the bitmap which are set to indicate that data has changed.

7. (Original) The system of claim 5, wherein difference information includes extent information.

8. (Original) The system of claim 5, wherein the prior time-state and the specified time-state are non-consecutive time-states.

9. (New) A system for backing-up data from a first storage pool to a second storage pool using difference information between time states, said system comprising a data management engine operable to:

identify a first set of point-in-time images on a first storage pool and a second set of point-in-time images on a second storage pool, each point-in-time image:

corresponding to a time-state of when the point-in-time image was made for application data; and

comprising difference information indicating a portion of changed application data from a previous point-in-time image in the set of point-in-time images, and content of the portion of changed application data for the time-state;

identify a common point-in-time image between the first set of point-in-time images and the second set of point-in-time images, the common point-in-time image comprising a most recent common time-state between the first set of point-in-time images and the second set of point-in-time images;

identify a third set of point-in-time images on the first storage pool comprising time-states between:

the time-state of the common point-in-time image; and

a time-state of a back-up function scheduled to back up the application data to the second storage pool;

generate composite difference information based on difference information for each point-in-time image in the third set of point-in-time images; and

transmit the composite difference information to the second storage pool to be compiled with the second point-in-time image to create a back-up copy of the application data for the time-state of the scheduled back-up function.

REMARKS

Claims 1-8 are pending. Claim 1 is cancelled, claim 9 is newly added, and claims 2-5 are amended. Support for the amendments can be found in the Present Application, as filed, at least at the original claims and ¶ [00123]. Applicants respectfully submit that no new matter is added.

Rejections under 35 U.S.C. 102 and 103

Claims 1, 4, 5 and 8 are rejected under 35 U.S.C. 102(e) as being anticipated by Botes (US 7,640,454). Claims 2 and 6 are rejected under 35 U.S.C. 103(a) as being unpatentable over Botes (US 7,640,454) as applied to claims 1 and 5, in view of Yagisawa et al. (US 2007/0162716) (hereinafter Yagisawa). Claims 3 and 7 are rejected under 35 U.S.C. 103(a) as being unpatentable over Botes as applied to claims 1 and 5 above, and further in view of Beatty et al. (US 2012/0078855) (hereinafter Beatty).

Botes discloses recovery software for restoring configurations of, e.g., databases, file systems, logical volumes and physical disks. In Botes, recovery software is used to "combine configuration information in volatile storage with configuration information from various locations in persistent storage," and to "store a previous state of the combined configuration information." Botes at Col. 6, lines 5-10. Botes is not directed to storing snapshots of application data.

Yagisawa discloses a storage controller in a storage system in which snapshots are made of two logical volumes. Each pair of logical volumes is assigned a pool volume, to which the appropriate snapshots are stored. The snapshots include difference data, and the difference data of each pair of logical volumes is stored in the associated pool volume.

Beatty discloses a method of performing a granular restore of a database from a differential backup using a bitmask. A differential backup file and bitmask is stored for each point in time at which a differential backup is performed. The bitmask is used to determine whether to read from a differential backup file or from a full backup file.

However, none of the cited prior art discloses "transmit[ting] the composite difference information to the second storage pool to be compiled with the second point-in-time image to create a back-up copy of the application data for the time-state of the scheduled back-up function," as is claimed by Claim 9. As well, neither does the cited prior art disclose "identify[ing] a third set of point-in-time images on the first storage pool comprising time-states between: the time-state of the common point-in-time image; and a time-state of a back-up function scheduled to back up the application data to the second storage pool; [and] generat[ing] composite difference information based on difference information for each point-in-time image in the third set of point-in-time images," as is claimed by Claim 9.

The present application teaches the use of difference information from multiple time-states to create a full back-up copy of data. By compositing the differences from multiple time-states, and by using the differences to create another full back-up copy, the present application teaches a system whereby the need for creating full back-up copies by directly copying application data is greatly reduced.

While Botes combines configuration information from various locations, Botes fails to teach or suggest "identify[ing] a third set of point-in-time images on the first storage pool comprising time-states between: the time-state of the common point-in-time image; and a time-state of a back-up function scheduled to back up the application data to the second storage pool," as is recited in claim 9. Without configuration information from multiple time-states, Botes therefore cannot transmit the composite difference information to a second storage pool to be compiled with a point-in-time image to create a back-up copy of the application data for a time-state of a scheduled back-up function.

Yagisawa and Beatty also do not disclose the claimed limitations. Although Yagisawa discloses snapshots with difference data, and although Beatty discloses differential backups using bitmasks, neither discloses compositing difference information from multiple point-in-time images, or creating full back-ups from composited difference information. Therefore none of Botes,

Yagisawa and Beatty, either alone or in combination, teach, suggest, or otherwise render obvious all the elements of claim 9.

Claim 5 recites that the data management engine is operable to "identify the existence of a prior point-in-time image of the data at the storage pool for a time-state prior to the specified time-state, create composite difference information from difference information for each point-in-time image of the data with a time-state between the specified time-state and the prior time-state, and send the composite difference information from the back-up copy of data to the storage pool." Therefore Botes, Yagisawa and Beatty, either alone or in combination, fail to teach, suggest, or otherwise render obvious all the elements of claim 5 for at least the reasons described above for claim 9.

In view of the above amendment, applicant believes the pending application is in condition for allowance.

Applicant respectfully requests Examiner to return the signed Form SB-08 submitted on February 14, 2013.

A three month extension of time, and appropriate fee, is filed herewith.

Applicant believes no additional fee is due with this response. However, if a fee is due, please charge our Deposit Account No. 08-0219, under Order No. 2203828.00130US1 from which the undersigned is authorized to draw.

Respectfully submitted,

Dated: April 12, 2013

/Zachary P. Piccolomini/
Zachary P. Piccolomini
Registration No.: 63,390
Attorney for Applicant(s)

Wilmer Cutler Pickering Hale and Dorr LLP
60 State Street
Boston, Massachusetts 02109
(617) 526-6000 (telephone)
(617) 526-5000 (facsimile)

PTO/SB/22 (10-12)
Approved for use through 01/31/2013. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

PETITION FOR EXTENSION OF TIME UNDER 37 CFR 1.136(a)	Docket Number (Optional) 2203828.00130US1
Application Number 12/947,393-Conf. #1801	Filed November 16, 2010
For SYSTEM AND METHOD FOR PERFORMING BACKUP OR RESTORE OPERATIONS UTILIZING DIFFERENCE INFORMATION AND TIMELINE STATE INFORMATION	
Art Unit 2189	Examiner G. Bansal

This is a request under the provisions of 37 CFR 1.136(a) to extend the period for filing a reply in the above identified application.

The requested extension and fee are as follows (check time period desired and enter the appropriate fee below):

		Fee	Small Entity Fee	
☐	One month (37 CFR 1.17(a)(1))	$150	$75	$
☐	Two months (37 CFR 1.17(a)(2))	$570	$285	$
☒	Three months (37 CFR 1.17(a)(3))	$1,290	$645	$ 700.00
☐	Four months (37 CFR 1.17(a)(4))	$2,010	$1,005	$
☐	Five months (37 CFR 1.17(a)(5))	$2,730	$1,365	$

☒ Applicant claims small entity status. See 37 CFR 1.27.

☐ A check in the amount of the fee is enclosed.

☒ Payment by credit card. ~~Form PTO-2038 is attac~~hed.

☐ The Director has already been authorized to charge fees in this application to a Deposit Account.

☒ The Director is hereby authorized to charge any fees which may be required, or credit any overpayment, to Deposit Account Number 08-0219 .

☒ Payment made via EFS-Web.

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

I am the

☐ applicant/inventor.

☐ assignee of record of the entire interest. See 37 CFR 3.71. 37 CFR 3.73(b) statement is enclosed (Form PTO/SB/96).

☒ attorney or agent of record. Registration Number 63,390 .

☐ attorney or agent under 37 CFR 1.34. Registration number _____ .

/Zachary P. Piccolomini	April 12, 2013
Signature	Date
Zachary P. Piccolomini	(617) 526-6000
Typed or printed name	Telephone Number

NOTE: This form must be signed in accordance with 37 CFR 1.33. See 37 CFR 1.4 for signature requirements and certifications. Submit multiple forms if more than one signature is required, see below*.

☐ *Total of 1 forms are submitted.

107740631

ACT6002406

Electronic Patent Application Fee Transmittal

Application Number:	12947393
Filing Date:	16-Nov-2010
Title of Invention:	SYSTEM AND METHOD FOR PERFORMING BACKUP OR RESTORE OPERATIONS UTILIZING DIFFERENCE INFORMATION AND TIMELINE STATE INFORMATION
First Named Inventor/Applicant Name:	Philip J. ABERCROMBIE
Filer:	Zachary Paul Piccolomini/Miriam Brooks
Attorney Docket Number:	2203828.00130US1

Filed as Small Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				
Extension - 3 months with $0 paid	2253	1	700	700

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
			Total in USD ($)	**700**

Electronic Acknowledgement Receipt

EFS ID:	15500625
Application Number:	12947393
International Application Number:	
Confirmation Number:	1801
Title of Invention:	SYSTEM AND METHOD FOR PERFORMING BACKUP OR RESTORE OPERATIONS UTILIZING DIFFERENCE INFORMATION AND TIMELINE STATE INFORMATION
First Named Inventor/Applicant Name:	Philip J. ABERCROMBIE
Customer Number:	23483
Filer:	Zachary Paul Piccolomini/Miriam Brooks
Filer Authorized By:	Zachary Paul Piccolomini
Attorney Docket Number:	2203828.00130US1
Receipt Date:	12-APR-2013
Filing Date:	16-NOV-2010
Time Stamp:	13:18:18
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Credit Card
Payment was successfully received in RAM	$700
RAM confirmation Number	10281
Deposit Account	
Authorized User	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)

1		Response_to_nonfinal.PDF	105867 e67409a5a5ad817365668ac4296e1012fdd39e0c	yes	9

Multipart Description/PDF files in .zip description		
Document Description	**Start**	**End**
Amendment/Req. Reconsideration-After Non-Final Reject	1	1
Specification	2	2
Claims	3	5
Applicant Arguments/Remarks Made in an Amendment	6	9

Warnings:

Information:

2	Extension of Time	Extension_of_time.PDF	78107 e94a80a13147a77cd91f954319aa95a8dbfb336e	no	1

Warnings:

Information:

3	Fee Worksheet (SB06)	fee-info.pdf	30768 9c9501418b0fe4ae4b34642250eaccecb9a562ba	no	2

Warnings:

Information:

Total Files Size (in bytes):	214742

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/06 (07-06)
Approved for use through 1/31/2007. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

PATENT APPLICATION FEE DETERMINATION RECORD

Substitute for Form PTO-875

Application or Docket Number	Filing Date	☐ To be Mailed
12/947,393	11/16/2010	

APPLICATION AS FILED – PART I

	(Column 1)	(Column 2)	SMALL ENTITY ☒		OR	OTHER THAN SMALL ENTITY	
FOR	NUMBER FILED	NUMBER EXTRA	RATE ($)	FEE ($)		RATE ($)	FEE ($)
☐ BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A			N/A	
☐ SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A			N/A	
☐ EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A			N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	minus 20 =	*	X $ =		OR	X $ =	
INDEPENDENT CLAIMS (37 CFR 1.16(h))	minus 3 =	*	X $ =			X $ =	
☐ APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $250 ($125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).						
☐ MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))							
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL			TOTAL	

APPLICATION AS AMENDED – PART II

AMENDMENT 04/12/2013

	(Column 1)		(Column 2)	(Column 3)	SMALL ENTITY		OR	OTHER THAN SMALL ENTITY	
	CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)		RATE ($)	ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	* 8	Minus	** 20	= 0	X $40 =	0	OR	X $ =	
Independent (37 CFR 1.16(h))	* 2	Minus	***3	= 0	X $210 =	0	OR	X $ =	
☐ Application Size Fee (37 CFR 1.16(s))									
☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
					TOTAL ADD'L FEE	0	OR	TOTAL ADD'L FEE	

AMENDMENT

	(Column 1)		(Column 2)	(Column 3)					
	CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)		RATE ($)	ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	*	Minus	**	=	X $ =		OR	X $ =	
Independent (37 CFR 1.16(h))	*	Minus	***	=	X $ =		OR	X $ =	
☐ Application Size Fee (37 CFR 1.16(s))									
☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
					TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.

** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".

*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".

The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

Legal Instrument Examiner:
/VANESSA BARBER/

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/947,393	11/16/2010	Philip J. ABERCROMBIE	2203828.00130US1	1801

23483 7590 05/06/2013

WILMERHALE/BOSTON
60 STATE STREET
BOSTON, MA 02109

EXAMINER	
BANSAL, GURTEJ	
ART UNIT	**PAPER NUMBER**
2189	

NOTIFICATION DATE	DELIVERY MODE
05/06/2013	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

teresa.carvalho@wilmerhale.com
whipusptopairs@wilmerhale.com

Office Action Summary	Application No. 12/947,393	Applicant(s) ABERCROMBIE ET AL.	
	Examiner GURTEJ BANSAL	Art Unit 2189	AIA (First Inventor to File) Status No

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *12 April 2013*.

☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

4)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

5)☒ Claim(s) *2-9* is/are pending in the application.

5a) Of the above claim(s) _____ is/are withdrawn from consideration.

6)☐ Claim(s) _____ is/are allowed.

7)☒ Claim(s) *2-9* is/are rejected.

8)☐ Claim(s) _____ is/are objected to.

9)☐ Claim(s) _____ are subject to restriction and/or election requirement.

* If any claims have been determined allowable, you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see http://www.uspto.gov/patents/init_events/pph/index.jsp or send an inquiry to PPHfeedback@uspto.gov.

Application Papers

10)☐ The specification is objected to by the Examiner.

11)☒ The drawing(s) filed on *16 November 2010* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

Certified copies:

a)☐ All b)☐ Some * c)☐ None of the:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Interim copies:

a)☐ All b)☐ Some c)☐ None of the: Interim copies of the priority documents have been received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date *02/14/2013*.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

4) ☐ Other: _____.

DETAILED ACTION

Remarks

In response to Non-Final Office Action dated October 12, 2012, Applicant presents claims 2-9 for examination, whereby claim 5 has been amended, claim 1 has been cancelled and claim 9 has been newly added.

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

1. Claims 2-3 and 6-9 are rejected under 35 U.S.C. 102(e) as being anticipated by Monckton (US 8,099,391).

As per claims 5 and 9, Monckton teaches a system for backing-up data from a first storage pool (virtual disk file 5 is construed to be a first storage pool) to a second storage pool (the storage region where backup data is held is construed to be a second storage pool) using difference information between time states, said system comprising a data management engine operable to:

identify a first set of point-in-time images on a first storage pool and a second set of point- in-time images on a second storage pool (fig. 2 which illustrates a backup of a image file on 5 and also 8), each point-in-time image:

corresponding to a time-state of when the point-in-time image was made for application data (As illustrated in fig. 2, the backup is at T1, fig. 3 is at T2); and

comprising difference information indicating a portion of changed application data from a previous point-in-time image in the set of point-in-time images, and content of the portion of changed application data for the time-state (col. 8, lines 5-40 describing how a bitmap is used to describe the sectors which have changed and also see col. 9, lines 1-15 describing fingerprint information);

identify a common point-in-time image between the first set of point-in-time images and the second set of point-in-time images (fig, 3, incremental backup), the common point-in-time image comprising a most recent common time-state between the first set of point-in-time images and the second set of point-in-time images (As illustrated in fig. 3 and described in col. 7, lines 17-35, this incremental backup is the most recent data as of time T2);

identify a third set of point-in-time images on the first storage pool (col. 8, lines5-35 describes how a bitmap is used to identify a third set between backups which have been updated and by comparing fingerprints as described in fig. 7; see also col. 9, lines 30-40) comprising time-states between:

the time-state of the common point-in-time image ; and

a time-state of a back-up function scheduled to back up the application data to the second storage pool (this information is applicable between every backup including the common point in time image and next back up; see fig. 7 which describes how the bitmap is used);

generate composite difference information based on difference information for each point- in-time image in the third set of point-in-time images (as illustrated in fig. 3-6 and described in col. 7, lines 20-35 the sectors which have been updated are found); and

transmit the composite difference information to the second storage pool to be compiled with the second point-in-time image to create a back-up copy of the application data for the time- state of the scheduled back-up function (As illustrated in the figures, this incremental backup with updated sectors is then stored in the backup area).

As per claims 2 and 6, Monckton teaches wherein difference information includes bitmap information with each bit of the bitmap corresponding to a portion of primary storage data, and including new data for those portions of the bitmap which are set to indicate that data has changed (col. 8, lines 20-25).

As per claims 3 and 7, Monckton teaches wherein difference information includes extent information (col. 8, lines 20-25 wherein the sectors are construed to be extents of the file which have been updated).

As per claim 8, Monckton teaches wherein the prior time-state and the specified time- state are non-consecutive time-states (As illustrated in fig. 5 with the differential backup not being consecutive to the full backup).

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

2. Claim 4 is rejected under 35 U.S.C. 103(a) as being unpatentable over Monckton as applied to claim 9 above, in view of Scheid (US 2006/0271622).

As per claim 4, Monckton teaches all the limitations of claim 9 and further teaches wherein multiple back-up functions are to occur each with different gaps of non-consecutive time-states (As described in col. 4, lines 25-50 with examples being given showing that each full backup is not consecutive since there multiple backups there between), and each with different composite difference information generated corresponding to the different gaps (As illustrated in figs 2-7, the amount backed up changes).

Monckton does not explicitly teach wherein multiple back-up functions are scheduled to occur simultaneously.

However, Scheid teaches wherein multiple back-up functions are scheduled to occur simultaneously ([0010]).

t would have been obvious to a person of ordinary skill in the art at the time of the invention to have combined the simultaneous backup of Scheid with the system of Monckton because it is known in the art that performing multiple tasks in parallel is faster than performing them in parallel and also it allows for multiple copies of the data to be stored in different places to increase security/protection.

Response to Arguments

Applicant's arguments with respect to claims 2-9 have been considered but are moot because the arguments do not apply to any of the references being used in the current rejection.

Conclusion

Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to GURTEJ BANSAL whose telephone number is (571)270-5588. The examiner can normally be reached on Monday - Friday, 7:30 a.m. - 5:00 p.m., EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Reginald Bragdon can be reached on (571)272-4204. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/G. B./
Examiner, Art Unit 2189

/Reginald G. Bragdon/
Supervisory Patent Examiner, Art Unit 2189

PTO/SB/08b (07-09)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 1 of 1

Complete if Known	
Application Number	12/947,393-Conf. #1801
Filing Date	November 16, 2010
First Named Inventor	Philip J. ABERCROMBIE
Art Unit	2189
Examiner Name	G. Bansal
Attorney Docket Number	2203828.00130US1

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
	AA*	US-20060080367	04-13-2006	Pudipeddi	
	AB*	US-20070162716	07-12-2007	Yagisawa et al.	
	AC*	US-20080243769	10-02-2008	Arbour et al.	
	AD*	US-20120078855	03-29-2012	Beatty et al.	
	AE*	US-7,640,454	12-29-2012	Botes	

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document Country Code[3]-Number[4]-Kind Code[5] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages Or Relevant Figures Appear	T[6]

NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
	CA	IEEE 100, The Authoritative Dictionary of IEEE Standards Terms, Seventh Edition, 2000, the Institute of Electrical and Electronics Engineering, Inc. page 532	

Examiner Signature	/Gurtej Bansal/ (04/30/2013)	Date Considered	04/30/2013

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. * CITE NO.: Those application(s) which are marked with an single asterisk (*) next to the Cite No. are not supplied (under 37 CFR 1.98(a)(2)(iii)) because that application was filed after June 30, 2003 or is available in the IFW. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /G.B./

ACT6002420

Notice of References Cited	Application/Control No. 12/947,393	Applicant(s)/Patent Under Reexamination ABERCROMBIE ET AL.	
	Examiner GURTEJ BANSAL	Art Unit 2189	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
*	A	US-8,099,391	01-2012	Monckton, Russell David	707/647
*	B	US-2006/0271622	11-2006	Scheid, William BJ	709/203
	C	US-			
	D	US-			
	E	US-			
	F	US-			
	G	US-			
	H	US-			
	I	US-			
	J	US-			
	K	US-			
	L	US-			
	M	US-			

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	
	V	
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	12947393	ABERCROMBIE ET AL.
	Examiner	**Art Unit**
	GURTEJ BANSAL	2189

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	10/04/2012	04/30/2013							
	1	✓	-							
	2	✓	✓							
	3	✓	✓							
	4	✓	✓							
	5	✓	✓							
	6	✓	✓							
	7	✓	✓							
	8	✓	✓							
	9		✓							

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	12947393	ABERCROMBIE ET AL.
	Examiner	Art Unit
	GURTEJ BANSAL	2189

CPC- SEARCHED		
Symbol	Date	Examiner

CPC COMBINATION SETS - SEARCHED		
Symbol	Date	Examiner

US CLASSIFICATION SEARCHED			
Class	Subclass	Date	Examiner
711	161; 162	10/04/2012	GB

SEARCH NOTES		
Search Notes	Date	Examiner
Text Search in EAST (See Attached)	10/04/2012	GB
NPL Search in IEEE (See Attached)	10/04/2012	GB
Inventor Search in eDAN	10/04/2012	GB
Updated Text Search in EAST (See Attached)	04/30/2013	GB

INTERFERENCE SEARCH			
US Class/ CPC Symbol	US Subclass / CPC Group	Date	Examiner

/GURTEJ BANSAL/ Examiner.Art Unit 2189	

EAST Search History

EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L5	1718	incremental near2 back$1up	US-PGPUB; USPAT	OR	OFF	2013/04/29 21:43
L6	15	incremental near2 back$1up with time with difference	US-PGPUB; USPAT	OR	OFF	2013/04/29 21:43
L7	2	incremental near2 back$1up with time with difference and bitmap	US-PGPUB; USPAT	OR	OFF	2013/04/29 21:43
L8	16	simultaneous$2 with incremental with back$1up	US-PGPUB; USPAT	OR	OFF	2013/04/29 23:29
L9	2	concurrent with incremental with back$1up	US-PGPUB; USPAT	OR	OFF	2013/04/29 23:40
L10	3	simulataneous$2 with back$1up	US-PGPUB; USPAT	OR	OFF	2013/04/29 23:41
L11	3671	simultaneous$2 with back$1up	US-PGPUB; USPAT	OR	OFF	2013/04/29 23:41
L12	784	simultaneous$2 near2 back$1up	US-PGPUB; USPAT	OR	OFF	2013/04/29 23:41
L13	57	simultaneous$2 near2 back$1up with time	US-PGPUB; USPAT	OR	OFF	2013/04/29 23:41
L14	67	parallel near2 back$1up with time	US-PGPUB; USPAT	OR	OFF	2013/04/29 23:43
L15	67	parallel near2 back$3up with time	US-PGPUB; USPAT	OR	OFF	2013/04/29 23:45
L16	156	parallel near2 back$3up same time	US-PGPUB; USPAT	OR	OFF	2013/04/29 23:46
L17	24	parallel with (partial incremental) with back$3up	US-PGPUB; USPAT	OR	OFF	2013/04/29 23:47
L18	8	concurrent$2 with (partial incremental) with back$3up	US-PGPUB; USPAT	OR	OFF	2013/04/29 23:49
L19	0	paralle with back$3up with array	US-PGPUB; USPAT	OR	OFF	2013/04/30 00:05
L20	1	paralle with back$3up	US-PGPUB;	OR	OFF	2013/04/30 00:05

			USPAT			
L21	4636	parallel with back$3up	US-PGPUB; USPAT	OR	OFF	2013/04/30 00:05
L22	77	parallel with back$3up with remote	US-PGPUB; USPAT	OR	OFF	2013/04/30 00:05
L23	3238	simultaneously with back$1up	US-PGPUB; USPAT	OR	OFF	2013/04/30 00:10
L24	64	simultaneously with back$1up with copies	US-PGPUB; USPAT	OR	OFF	2013/04/30 00:11
S1	6250	711/161.ccls. 711/162.ccls.	US-PGPUB; USPAT	OR	OFF	2012/09/24 16:21
S2	745	point$1in$1time with snapshot and back$up	US-PGPUB; USPAT	OR	OFF	2012/09/24 16:23
S3	28	point$1in$1time with snapshot same difference and back$up	US-PGPUB; USPAT	OR	OFF	2012/09/24 16:23
S4	30	point$1in$1time same snapshot same difference and back$up	US-PGPUB; USPAT	OR	OFF	2012/09/24 16:31
S5	88	point$1in$1time same snapshot and difference with information and back$up	US-PGPUB; USPAT	OR	OFF	2012/09/24 16:35
S6	14	"7065619".pn. "20020049778".pn. "20080034016".pn. "20090307251".pn. "20100077013".pn. "20100088277".pn. "20100138827".pn. "20100276744".pn. "20110179341".pn. "20110307447".pn. "20110307683".pn. "20120017060".pn. "6883073".pn. "7814149".pn.	US-PGPUB; USPAT	OR	OFF	2012/09/24 16:54
S7	15	"7065619".pn. "20020049778".pn. "20080034016".pn. "20090307251".pn. "20100077013".pn. "20100088277".pn. "20100138827".pn. "20100276744".pn. "20110179341".pn. "20110307447".pn. "20110307683".pn. "20120017060".pn. "6883073".pn. "7814149".pn. "20110252189".pn.	US-PGPUB; USPAT	OR	OFF	2012/09/24 16:54
S8	157	snapshot and difference with bitmap	US-PGPUB; USPAT	OR	OFF	2012/09/27 00:18
S9	17	snapshot and difference with bitmap and deduplication	US-PGPUB; USPAT	OR	OFF	2012/09/27 00:20
S10	11	snapshot and update with bitmap and deduplication	US-PGPUB; USPAT	OR	OFF	2012/09/27 00:25
S11	11	snapshot with bitmap and deduplication	US-PGPUB; USPAT	OR	OFF	2012/09/27 00:26
S12	1	"20120124307".pn.	US-PGPUB; USPAT	OR	OFF	2012/09/27 12:04

S13	1	"20120124307".pn. and extent	US-PGPUB; USPAT	OR	OFF	2012/09/27 12:05
S14	10	deduplication with capacity with performance	US-PGPUB; USPAT	OR	OFF	2012/09/27 12:28
S15	53	deduplication same capacity same performance	US-PGPUB; USPAT	OR	OFF	2012/09/27 12:29
S16	130	deduplication with capacity	US-PGPUB; USPAT	OR	OFF	2012/09/27 12:30
S17	8	deduplication with capacity same (efficiency speed)	US-PGPUB; USPAT	OR	OFF	2012/09/27 12:31
S18	25	difference with extent same (backup back$1up back adj up)	US-PGPUB; USPAT	OR	OFF	2012/09/27 13:02
S19	6	bitmap with extent same (backup back$1up back adj up)	US-PGPUB; USPAT	OR	OFF	2012/09/27 13:02
S20	1	"7640454".pn.	US-PGPUB; USPAT	OR	OFF	2012/10/04 13:01
S21	1	"20120078855".pn.	US-PGPUB; USPAT	OR	OFF	2012/10/04 17:03

EAST Search History (Interference)

< This search history is empty>

4/ 30/ 2013 12:43:20 AM
C:\ Users\ gbansal\ Documents\ EAST\ Workspaces\ 12947393.wsp

PTO/SB/06 (09-11)
Approved for use through 1/31/2014. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PATENT APPLICATION FEE DETERMINATION RECORD Substitute for Form PTO-875	Application or Docket Number 12/947,393	Filing Date 11/16/2010	☐ To be Mailed

ENTITY: ☐ LARGE ☒ SMALL ☐ MICRO

APPLICATION AS FILED – PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	RATE ($)	FEE ($)
☐ BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A	
☐ SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A	
☐ EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	minus 20 =	*	X $ =	
INDEPENDENT CLAIMS (37 CFR 1.16(h))	minus 3 =	*	X $ =	
☐ APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $310 ($155 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).			
☐ MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))				
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL	

APPLICATION AS AMENDED – PART II

AMENDMENT	08/16/2013	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)
	Total (37 CFR 1.16(i))	* 19	Minus	** 21	= 0	X $40 =	0
	Independent (37 CFR 1.16(h))	* 2	Minus	***5	= 0	X $210 =	0
	☐ Application Size Fee (37 CFR 1.16(s))						
	☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))						
						TOTAL ADD'L FEE	0

AMENDMENT		(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)
	Total (37 CFR 1.16(i))	*	Minus	**	=	X $ =	
	Independent (37 CFR 1.16(h))	*	Minus	***	=	X $ =	
	☐ Application Size Fee (37 CFR 1.16(s))						
	☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))						
						TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".
The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

LIE
/KIM DOWNING/

Electronic Acknowledgement Receipt

EFS ID:	17317672
Application Number:	12947393
International Application Number:	
Confirmation Number:	1801
Title of Invention:	SYSTEM AND METHOD FOR PERFORMING BACKUP OR RESTORE OPERATIONS UTILIZING DIFFERENCE INFORMATION AND TIMELINE STATE INFORMATION
First Named Inventor/Applicant Name:	Philip J. ABERCROMBIE
Customer Number:	23483
Filer:	Zachary Paul Piccolomini/Miranda Mitchell
Filer Authorized By:	Zachary Paul Piccolomini
Attorney Docket Number:	2203828.00130US1
Receipt Date:	05-NOV-2013
Filing Date:	16-NOV-2010
Time Stamp:	15:32:20
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Credit Card
Payment was successfully received in RAM	$1300
RAM confirmation Number	2237
Deposit Account	080219
Authorized User	PICCOLOMINI, ZACHARY

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

Charge any Additional Fees required under 37 C.F.R. Section 1.16 (National application filing, search, and examination fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees)

ACT6002428

Charge any Additional Fees required under 37 C.F.R. Section 1.19 (Document supply fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.20 (Post Issuance fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.21 (Miscellaneous fees and charges)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		130US1_ResponseFinal.PDF	88264 9742768a3b3c8fde70937a384630cb09053fecd4	yes	7
	Multipart Description/PDF files in .zip description				
	Document Description		**Start**	**End**	
	Amendment Submitted/Entered with Filing of CPA/RCE		1	1	
	Claims		2	4	
	Applicant Arguments/Remarks Made in an Amendment		5	7	
Warnings:					
Information:					
2	Extension of Time	130US1_EOT.PDF	78169 b00bef0757077f841ed3b23c35b6e34b085556bf	no	1
Warnings:					
Information:					
3	Request for Continued Examination (RCE)	130US1_RCE.PDF	697759 324e093dc75cb8f6918061e0828842bc2381a042	no	3
Warnings:					
Information:					
4	Fee Worksheet (SB06)	fee-info.pdf	32445 9ae33e96523de70d4fd061e9bbdbbe5b307d1293	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	896637		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

Electronic Patent Application Fee Transmittal

Application Number:	12947393
Filing Date:	16-Nov-2010
Title of Invention:	SYSTEM AND METHOD FOR PERFORMING BACKUP OR RESTORE OPERATIONS UTILIZING DIFFERENCE INFORMATION AND TIMELINE STATE INFORMATION
First Named Inventor/Applicant Name:	Philip J. ABERCROMBIE
Filer:	Zachary Paul Piccolomini/Miranda Mitchell
Attorney Docket Number:	2203828.00130US1

Filed as Small Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				
Extension - 3 months with $0 paid	2253	1	700	700

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
Request for Continued Examination	2801	1	600	600
		Total in USD ($)		**1300**