

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

**AMAZON.COM, INC.,
AMAZON.COM SERVICES LLC,
AMAZON WEB SERVICES, INC., and
AUDIBLE, INC.,**
Petitioners,

v.

AUDIO POD IP, LLC,
Patent Owner.

Case No. IPR2025-00774
U.S. Patent No. 8,738,740

DECLARATION OF PROFESSOR KETAN MAYER-PATEL, Ph.D.

Amazon v. Audio Pod
US Patent 8,738,740
Amazon EX-1002

TABLE OF CONTENTS

I.	BACKGROUND -----	1
A.	Experience and Qualifications-----	1
B.	Materials Considered -----	3
II.	APPLICABLE LEGAL STANDARDS -----	6
A.	Claim Construction -----	6
B.	Obviousness -----	8
III.	PERSON OF ORDINARY SKILL IN THE ART -----	12
IV.	TECHNOLOGY BACKGROUND -----	13
A.	Selecting Servers Based on Performance Statistics Was Known -----	13
B.	Storing Media Content in Segments Was Known -----	14
V.	THE '740 PATENT -----	15
A.	Overview -----	15
B.	Claims -----	16
VI.	CLAIMS 1-6, 10-11, AND 18 OF THE '740 PATENT WOULD HAVE BEEN OBVIOUS-----	17
A.	Claims 1 and 6 Would Have Been Obvious in View of Young-----	17
1.	Claim 1 -----	20
a.	Preamble -----	20
b.	Element 1[a]: Sending a Request -----	20
c.	Element 1[b]: Loading a List of Servers-----	22
d.	Element 1[c]: Maintaining Server Statistics-----	23

e.	Element 1[d][i]: Selecting a First Server -----	25
i.	First Stage Server Selection -----	25
ii.	Second Stage Server Selection -----	26
f.	Element 1[d][ii]: First Server Having a Plurality of Files -----	27
g.	Element 1[e]: Downloading a First Audio File -----	28
i.	First Stage Server -----	29
ii.	Second Stage Server -----	29
h.	Element 1[f][i]: Selecting a Second Server -----	30
i.	First Stage Server Selection -----	31
ii.	Second Stage Server Selection -----	31
i.	Element 1[f][ii]: Second Server Having a Copy of the Files -----	32
j.	Element 1[g]: Downloading a Second Audio File from the Second Server -----	33
i.	First Stage Server -----	33
ii.	Second Stage Server -----	34
2.	Claim 6 -----	35
B.	Claims 1, 6, and 11 Would Have Been Obvious in View of Young, Yoshimura, and Copley -----	36
1.	Claim 1 -----	37
a.	Elements 1[c], 1[d][i], and 1[e] -----	37
b.	Elements 1[d][ii]-1[g] -----	39
2.	Claim 6 -----	42

3.	Claim 11 -----	43
	a. Element 11[a] -----	43
	b. Element 11[b] -----	46
C.	Claim 2 Would Have Been Obvious in View of Young and Kate, or in View of Young, Yoshimura, Copley, and Kate-----	49
D.	Claim 5 Would Have Been Obvious in View of Young and Sull, or in View of Young, Yoshimura, Copley, and Sull -----	51
E.	Claim 18 Would Have Been Obvious in View of Young and Ganesan, or in View of Young, Yoshimura, Copley, and Ganesan -----	54
F.	Claim 3 Would Have Been Obvious in View of Young, Kate, and Ganesan, or in View of Young, Yoshimura, Copley, Kate, and Ganesan -----	57
G.	Claim 4 Would Have Been Obvious in View of Young, Kate, Ganesan, and Sull or in View of Young, Yoshimura, Copley, Kate, Ganesan, and Sull-----	58
H.	Claim 10 Would Have Been Obvious in View of Young and Shapiro, or in View of Young, Yoshimura, Copley, and Shapiro -----	58
I.	Claim 11 Would Have Been Obvious in View of Young, Yoshimura, Copley, and Lindahl or in View of Young, Yoshimura, Copley, and Sull -----	61
	1. Lindahl-----	61
	2. Sull -----	65
J.	Claims 1 and 6 Would Have Been Obvious in View of Leighton and Seed -----	65
	1. Claim 1 -----	66
	a. Preamble -----	66
	b. Element 1[a]: Sending a Request-----	66

c.	Element 1[b]: Loading a List of Servers-----	67
d.	Element 1[c]: Maintaining Server Statistics-----	68
e.	Element 1[d][i]: Selecting a First Server -----	70
f.	Element 1[d][ii]: First Server Having a Plurality of Files -----	70
g.	Element 1[e]: Downloading a First Audio File -----	74
h.	Element 1[f][i]: Selecting a Second Server -----	75
i.	Element 1[f][ii]: Second Server Having a Copy of the Files -----	75
j.	Element 1[g]: Downloading a Second Audio File from the Second Server-----	76
2.	Claim 6 -----	77
K.	Claims 1, 2, 6, and 10 Would Have Been Obvious in View of Leighton, Seed, and Lindahl-----	78
1.	Claim 1 -----	79
2.	Claim 2 -----	82
3.	Claim 6 -----	84
4.	Claim 10-----	84

L.	Claim 2 Would Have Been Obvious in View of Leighton, Seed, and Kate-----	85
M.	Claim 5 Would Have Been Obvious in View of Leighton, Seed, and Sull or in View of Leighton, Seed, Lindahl, and Sull-----	86
N.	Claim 18 Would Have Been Obvious in View of Leighton, Seed, and Ganesan or in View of Leighton, Seed, Lindahl, and Ganesan -----	88
O.	Claim 3 Would Have Been Obvious in View of Leighton, Seed, Lindahl, and Ganesan, or in View of Leighton, Seed, Kate, and Ganesan -----	89
P.	Claim 4 Would Have Been Obvious in View of Leighton, Seed, Lindahl, Ganesan, and Sull or in View of Leighton, Seed, Kate, Ganesan, and Sull-----	89
Q.	Claim 10 Would Have Been Obvious in View of Leighton, Seed, and Shapiro or in View of Leighton, Seed, Lindahl, and Shapiro -----	90
R.	Claim 11 Would Have Been Obvious in View of Leighton, Seed, Lindahl, and Yoshimura -----	91
	1. Motivation to Use Yoshimura’s SMIL File-----	92
	2. Motivation to Use Lindahl’s Bookmark-----	94
VII.	SECONDARY CONSIDERATIONS OF NONOBVIOUSNESS -----	95
VIII.	CONCLUSION -----	96

I, Ketan Mayer-Patel, do hereby declare:

1. I am making this declaration at the request of Petitioners Amazon.com, Inc., Amazon.com Services LLC, Amazon Web Services, Inc., and Audible, Inc. (“Petitioners”). I have been retained by Petitioners as a technical expert in this matter.

2. I am being compensated for my work on this case. My compensation does not depend on the content of this Declaration or the outcome of these proceedings.

I. BACKGROUND

A. Experience and Qualifications

3. I received Bachelor of Arts degrees in Computer Science and Economics in 1992, a Master of Science in 1997 from the Department of Electrical Engineering and Computer Science, and a Ph.D. in 1999 from the Department of Electrical Engineering and Computer Science, all from the University of California, Berkeley.

4. I have been involved in the research and development of multimedia computing systems for nearly 30 years. I have been a faculty member at the University of North Carolina since January 2000, where I perform research and teach in the areas of networking, web programming, and multimedia computing. I also have

expertise other areas, including distributed systems, networking devices, and the general operation of computer systems.

5. I am a member of the Association for Computing Machinery (ACM) and the Institute of Electrical and Electronics Engineers (IEEE). These are the two leading professional societies for both academic and practicing computer scientists.

6. I have authored or co-authored over 30 papers in peer-reviewed journals and conference proceedings. I have served as an Associate Editor for both IEEE Transactions on Multimedia and ACM Transactions on Multimedia Computing, Communications, and Applications, which are the two leading journals in the field. I regularly serve as a member of the technical program committee for a number of different conferences and workshops including ACM Multimedia, The International Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV), IFIP Networking, ACM Multimedia Systems (MMSys), MMEDIA, and SIGMAP. I am also currently chair of the standing executive committee for both NOSSDAV and MMSys. A complete listing of all my publications can be found in my CV, which I understand is being submitted as Exhibit 1096. I am also a named inventor or co-inventor on multiple issued patents, which are also listed in my CV.

7. My research has been supported by both government agencies as well as private industry. I received the National Science Foundation (NSF) CAREER Award in 2003 while an Assistant Professor. I have been a principal investigator for

grants awarded by the NSF, the Office of Naval Research, and the Laboratory of Analytic Sciences. I have also served on several NSF reviewing panels for funding recommendations.

8. In my research and teaching I have considered problems of video streaming, dynamic adaptation and transcoding of media, adaptive streaming transport protocols, telepresence, and scalable display architectures, among others.

9. In the classroom, I have regularly taught classes on Data Structures, Foundations of Programming, Modern Web Programming, Files and Databases, and Multimedia Computing and Networking. I also serve as the Director of Undergraduate Studies for the Department of Computer Science.

B. Materials Considered

10. In preparing this Declaration, I have considered the following materials:

Exhibit No.	Description
1001	U.S. Patent No. 8,738,740 (“the ’740 patent”)
1003	NATIONAL INFORMATION STANDARDS ORGANIZATION, SPECIFICATIONS FOR THE DIGITAL TALKING BOOK (ANSI/NISO Z39.86-2002) (2002) (“DTB”)
1004	European Patent Publication No. EP 1463258 A1 (“Lindahl”)
1005	U.S. Patent Publication No. 2002/0069218 (“Sull”)

Amazon.com, Inc. v. Audio Pod IP, LLC
 Declaration of Ketan Mayer-Patel – U.S. Pat. No. 8,738,740

Exhibit No.	Description
1006	Yoshimura et al., <i>Content Delivery Network Architecture for Mobile Streaming Service Enabled by SMIL Modification</i> , 86 IEICE TRANSACTIONS ON COMMUNICATIONS 1778 (2003) (“Yoshimura”)
1007	Excerpts from DICK C.A. BULTERMAN & LLOYD RUTLEDGE, SMIL 2.0, INTERACTIVE MULTIMEDIA FOR WEB AND MOBILE DEVICES (2004) (“Bulterman”)
1008	U.S. Patent No. 6,477,522 (“Young”)
1009	U.S. Patent Publication No. 2003/0061305 (“Copley”)
1010	U.S. Patent No. 7,191,215 (“Ganesan”)
1011	PCT Patent Publication No. WO2003/069437 (“Seed”)
1028	U.S. Patent Publication No. 2002/0184189 (“Hay”)
1030	U.S. Patent No. 6,260,011 (“Heckerman”)
1032	U.S. Patent No. 6,108,703 (“Leighton”)
1033	U.S. Patent No. 5,922,045 (“Hanson”)
1035	U.S. Patent Publication No. 2004/0148638 (“Weisman”)
1041	PCT Patent Publication No. WO2005/010776 (“Kate”)
1042	U.S. Patent No. 5,991,810 (“Shapiro”)
1052	U.S. Patent Publication No. 2006/0236219 (“Grigorovitch”)
1080	U.S. Patent Publication No. 2003/0091338 (“Snow”)
1081	John Dilley et al., <i>Globally Distributed Content Delivery</i> , in 6 IEEE INTERNET COMPUTING 50 (2002) (“Dilley”)
1082	Crovella et al., <i>Dynamic Server Selection in the Internet</i> , PROC. 3 RD WORKSHOP ON HIGH PERFORMANCE SUBSYSTEMS (HPCS '95) (1995) (“Crovella”)

Amazon.com, Inc. v. Audio Pod IP, LLC
 Declaration of Ketan Mayer-Patel – U.S. Pat. No. 8,738,740

Exhibit No.	Description
1083	U.S. Patent Publication No. 2001/0041062 (“Ottesen”)
1084	U.S. Patent Publication No. 2004/0052371 (“Watanabe”)
1085	U.S. Patent Publication No. 2002/0147979 (“Corson”)
1086	PCT Patent Publication No. WO2001/24474 (“Shteyn”)
1087	U.S. Patent No. 6,175,869 (“Ahuja”)
1089	Fei et al., <i>A Novel Server Selection Technique for Improving the Response Time of a Replicated Service</i> , PROC. IEEE CONF. ON COMPUTER COMM’NS (INFOCOM ’98) (“Fei”)
1090	Excerpts from T. KENNEDY & M. SLOWINSKI, SMIL: ADDING MULTIMEDIA TO THE WEB (2001) (“Kennedy”)
1092	A. Maino, <i>Providing X.509-Based User Access Control to Web Servers</i> , in 14 TH INT’L INFO. SECURITY CONFERENCE (IFIP/SEC ’98) (1998) (“Maino”)
1093	U.S. Patent No. 6,226,752 (“Gupta”)
1094	U.S. Patent No. 6,505,238 (“Tran”)
1095	File History for U.S. Patent No. 8,738,740

11. In addition, I have reviewed the full file history of the ’740 patent. I have also relied on my education, training, and experience, and my knowledge of pertinent literature in the field of the ’740 patent.

II. APPLICABLE LEGAL STANDARDS

12. I have been asked to provide my opinion as to whether the claims of the '740 patent would have been obvious to a person of ordinary skill in the art at the time of the alleged invention, in view of the prior art.

13. I am a computer scientist by training and profession. The opinions I am expressing in this report involve the application of my training and technical knowledge and experience to the evaluation of certain prior art with respect to the '740 patent.

14. Although I have been involved as a technical expert in patent matters before, I am not an expert in patent law. Therefore, the attorneys from Knobbe, Martens, Olson & Bear, LLP have provided me with guidance as to the applicable patent law in this matter. The paragraphs below express my understanding of how I must apply current principles related to patent validity to my analysis.

A. Claim Construction

15. It is my understanding that in determining whether a patent claim is obvious in view of the prior art, the Patent Office construes the claim by giving the claim terms their plain and ordinary meaning, as they would have been understood by a person of ordinary skill in the art at the time of the invention in view of the intrinsic record (patent specification and file history). For the purposes of this review, and to the extent necessary, I have interpreted each claim term in accordance

with its plain and ordinary meaning as it would have been understood by a person of ordinary skill in the art at the time of the invention, in view of the intrinsic record. I understand that the time of the invention is December 13, 2005.

16. I understand that a patent and its prosecution history are considered “intrinsic evidence” and are the most important sources for interpreting claim language in a patent. I also understand that in reading the claim, I must not import limitations from the specification into the claim terms; in other words, I must not narrow the scope of the claim terms by implicitly adding disclosed limitations that have no express basis in the claims. The prosecution history of related patents and applications can also be relevant.

17. I understand that sources extrinsic to a patent and its prosecution history (such as dictionary definitions and technical publications) may also be used to help interpret the claim language, but that such extrinsic sources cannot be used to contradict the unambiguous meaning of the claim language that is evident from the intrinsic evidence.

18. Unless expressly stated herein, I have applied the plain and ordinary meaning of the claim terms, which I understand is the meaning that a person of ordinary skill in the art would have given to terms in December 2005 based on a review of the intrinsic evidence.

B. Obviousness

19. It is my understanding that a claim is “obvious” if the claimed subject matter as a whole would have been obvious to a person of ordinary skill in the art at the time of the alleged invention. I understand that an obviousness analysis involves a number of considerations. I understand that the following factors must be evaluated to determine whether a claim would have been obvious: (i) the scope and content of the prior art; (ii) the differences, if any, between each claim of the ’740 patent and the prior art; (iii) the level of ordinary skill in the art in December 2005; and (iv) additional considerations, if any, that indicate that the invention was obvious or not obvious. I understand that these “additional considerations” are often referred to as “secondary considerations” of non-obviousness or obviousness.

20. I also understand that the frame of reference when evaluating obviousness is what a hypothetical person of ordinary skill in the pertinent art would have known in December 2005. I understand that the hypothetical person of ordinary skill is presumed to have knowledge of all pertinent prior art references.

21. It is my understanding that something is “inherent in,” and therefore taught by, the prior art, if it necessarily flows from the explicit disclosure of the prior art. I understand that the fact that a certain result or characteristic *may be* present in the prior art is not sufficient to establish inherency. However, if the result or

characteristic is *necessarily* present based upon the explicit disclosure in the prior art, it is inherent in the prior art and is therefore disclosed.

22. I understand that a prior art reference may be a pertinent prior art reference (or “analogous art”) if it is in the same field of endeavor as the patent or if it is pertinent to the problem that the inventors were trying to solve. A reference is reasonably pertinent if it logically would have commended itself to an inventor’s attention in considering the problem at hand. If a reference relates to the same problem as the claimed invention, that supports use of the reference as prior art in an obviousness analysis. Here, all of the references relied on in my obviousness analysis below are from the same field of endeavor as the ’740 patent, e.g., content distribution and/or rendering. The references are also pertinent to a particular problem the inventor was focused on, e.g., efficient and effective distribution and/or rendering of content.

23. It is my understanding that the law recognizes several rationales for combining references or modifying a reference to show obviousness of claimed subject matter. Some of these rationales include:

- combining prior art elements according to known methods to yield predictable results;
- simple substitution of one known element for another to obtain predictable results;

- a predictable use of prior art elements according to their established functions;
- using known techniques to improve similar devices (methods, or products) in the same way;
- applying a known technique to a known device (method, or product) ready for improvement to yield predictable results;
- choosing from a finite number of identified, predictable solutions, with a reasonable expectation of success (in which case a claim would have been obvious to try);
- known work in one field of endeavor may prompt variations of it for use in either the same field or a different one based on design incentives or other market forces if the variations would have been predictable to one of ordinary skill in the art; and
- some teaching, suggestion, or motivation in the prior art that would have led one of ordinary skill to modify the prior art reference or to combine prior art reference teachings to arrive at the claimed invention.

24. I understand that “secondary considerations” must be considered as part of the obviousness analysis when present. I further understand that the secondary considerations may include: (1) a long-felt but unmet need in the prior art that was satisfied by the claimed invention; (2) the failure of others; (3) skepticism by experts;

(4) commercial success of a product covered by the patent; (5) unexpected results achieved by the claimed invention; (6) industry praise of the claimed invention; (7) deliberate copying of the invention; and (8) teaching away by others. I also understand that evidence of the independent and nearly simultaneous “invention” of the claimed subject matter by others is a secondary consideration supporting an obviousness determination and may support a conclusion that a claimed invention was within the knowledge of a person of ordinary skill as of December 13, 2005. I am not aware of any evidence of secondary considerations that would suggest that the claims of the ’740 patent would have been nonobvious in December 2005.

25. I understand that when assessing obviousness, using hindsight is impermissible; that is, what is known today or what was learned from the teachings of the patent should not be considered. The patent should not be used as a road map for selecting and combining items of prior art. Rather, obviousness must be considered from the perspective of a person of ordinary skill at the time the alleged invention was made – December 2005 in this case.

26. I also understand that an obviousness analysis must consider the invention as a whole, as opposed to just a part or element of the invention. I understand this “as a whole” assessment to require showing that one of ordinary skill in the art at the time of invention, confronted by the same problems as the inventor and with

no knowledge of the claimed invention, would have selected the elements from the prior art and combined them in the claimed manner.

III. PERSON OF ORDINARY SKILL IN THE ART

27. It is my understanding that when interpreting the claims of the '740 patent and evaluating whether a claim would have been obvious, I must do so based on the perspective of a person of ordinary skill in the art at the relevant priority date. I understand that the relevant priority date of the '740 patent is December 13, 2005.

28. I understand that in determining the level of ordinary skill in the art, several factors are considered. Those factors may include: (i) the type of problems encountered in the art; (ii) prior art solutions to those problems; (iii) the rapidity with which innovations are made; (iv) the sophistication of the technology; and (v) the educational level of active workers in the field. A person of ordinary skill in the art must have the capability of understanding the scientific and engineering principles applicable to the pertinent art.

29. The '740 patent describes the use of well-known technologies for the rendering and/or distribution of digital content. Based on my review of the specification and claims of the '740 patent, it is my opinion that a person of ordinary skill in the art would have had a minimum of a bachelor's degree in electrical engineering, computer engineering, or computer science, and at least three years of industry or academic experience in the design, development, and/or implementation of content

rendering and/or distribution systems. Work experience could substitute for formal education and additional formal education could substitute for work experience.

30. My conclusions below that the claims of the '740 patent would have been obvious would remain the same even if the priority date, field of endeavor, or level of ordinary skill were slightly different.

31. I meet the above definition of a person of ordinary skill in the art, and did so as of December 13, 2005. Also, I have worked with persons of ordinary skill in the art through my professional and academic experiences, and I have an understanding of their skill level around December 2005.

IV. TECHNOLOGY BACKGROUND

A. Selecting Servers Based on Performance Statistics Was Known.

32. By the 1990s, content distribution networks (“CDNs”) enabled the distribution of content from multiple servers, which improved performance and reliability. A CDN routes a client request to the optimal server using server performance statistics such as the time for information to travel between a server and client, packet loss, bandwidth, or server load. (EX-1081 (Dilley), 51; EX-1082 (Crovella), 1.) This process was often handled by domain name system (“DNS”) servers that map an http request to the identified server or servers. (EX-1081 (Dilley), 52.)

33. In 2002, Young disclosed a CDN that tracks server performance and routes media content requests to an optimal server based on performance statistics.

In response to a client’s request for content, the client obtains a list of servers having the requested content and ranks those servers based on performance statistics. (EX-1008 (Young), Abstract.) The client downloads a portion of the requested content from each server in the list, re-ranks the servers based on performance during these downloads, and then chooses the optimal server for downloading the remainder of the content. (*Id.*) If the chosen server’s performance drops below a set threshold during the download, the next best server is selected. (*Id.*)

34. Similar server-selection processes were described in Leighton (EX-1032), Copley (EX-1009, Abstract), and Seed (EX-1011 ¶[0024]), among many other references. The ’740 patent acknowledges that automatically selecting a second server was known. (EX-1001 (’740 patent), 10:58-67.)

B. Storing Media Content in Segments Was Known.

35. Segmenting a media file into smaller files was well known by 2005, as the ’740 patent admits. (EX-1001 (’740 patent), 2:16-18 (“The Godfather” audiobook segmented into 24 MP3 files).) Such segmentation was described in many references. (*E.g.*, EX-1080 (Snow) ¶¶[0027] (audio file divided into two smaller files), [0029] (file is “sliced into smaller playback files”); EX-1030 (Heckerman), 13:14-27 (audio file divided into “an audio ... file for each sentence or paragraph”), claim 22 (“segmenting the audio and text data into multiple audio and data files”); EX-1083 (Ottesen) ¶¶[0043] (audio presentation may be segmented), [0056]-[0057]

(content, including audio, may be “segmented or divided into a plurality of discrete ... program segments”); EX-1084 (Watanabe) ¶[0002]; EX-1085 (Corson), claim 1; EX-1086 (Shteyn), Abstract.)

36. As one example, Yoshimura disclosed a CDN in which original media content was divided into segments that were stored as separate files on content servers:

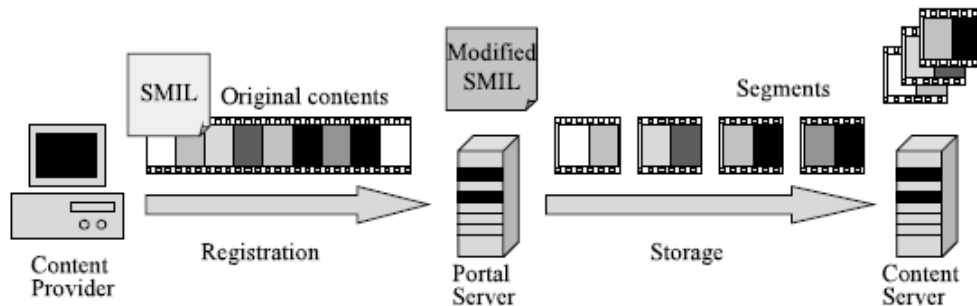


Fig. 3 Content segmentation and SMIL modification at portal server.

(EX-1006 (Yoshimura), Fig. 3; *id.*, 1781 (audio file “content-A.mp4” divided into three separate files, “content-A- $\{1,2,3\}$.mp4”).)

V. THE '740 PATENT

A. Overview

37. The '740 patent discloses a system for delivering digital audio data. (EX-1001 ('740 patent), 1:18-19.) Its purported advance is “segmenting an audio stream into a plurality of small digital audio files” that can be “transmitted, loaded, and played, in a specific order[.]” (*Id.*, 2:33-41.) The files are stored on multiple servers. (*Id.*, 16:52-62.) A client can download a server list and select a “primary

server” (e.g., fastest available) from which to download files. (*Id.*, 10:35-36.) If the primary server encounters performance issues or fails, the client selects another server. (*Id.*, 10:4-39, 10:58-67.) The patent admits that (i) segmenting audio content into smaller files was known (*id.*, 2:16-19), and (ii) “server replacement” was known (*id.*, 10:60-61).

B. Claims

38. I have been asked to consider claims 1-6, 10-11, and 18 in this Declaration. Of those, claim 1 is independent and recites a “non-transitory computer readable storage medium including computer readable code, which when executed by a computer, causes said computer to” perform several steps:

- [a] send a request to a network-based server, the request including a unique identifier for identifying an audio stream;
- [b] load a list of library servers received from the network-based server, the list of library servers determined in dependence upon the unique identifier;
- [c] maintain service level statistics for each library server in the list of library servers;
- [d] select a first library server from the list of library servers in dependence upon the service level statistics, the first library server having a plurality

of digital audio files, each digital audio file in the plurality of digital audio files including a different segment of the audio stream;

[e] download a first digital audio file from the plurality of digital audio files for playback with a media player,

[f] select a second library server from the list of library servers in dependence upon the service level statistics, the second library server having a copy of the plurality of digital audio files; and

[g] download a second other digital audio file from the second library server for playback with the media player.

VI. CLAIMS 1-6, 10-11, AND 18 OF THE '740 PATENT WOULD HAVE BEEN OBVIOUS.

A. Claims 1 and 6 Would Have Been Obvious in View of Young.

39. For at least the reasons I discuss below, Young renders claims 1 and 6 obvious.

40. Young discloses a CDN that tracks server performance and routes media content requests to an optimal server. In response to a client request for a media file, the client obtains a list of servers that host the requested content. (EX-1008 (Young), Abstract.) The client then performs a series of steps relating to server selection and downloads. First, the client pings each of the servers on the list and prioritizes them based on performance statistics (e.g., latency). (*Id.*) The client then downloads a first portion of the requested file from the highest priority server, a

second portion of the file from the second highest priority server, and so on until it has received a portion of the file from every server in the list. Those steps alone, which I refer to as the “First Stage” for ease of reference, disclose or render obvious claims 1 and 6.

41. The client, having monitored the throughput of each server during the downloading of the portions during the First Stage, determines the best server to complete the download based on the servers’ performance statistics. (EX-1008 (Young), Abstract.) But if the chosen server’s performance drops below a set threshold during the download, the client selects the next best server based on the previously obtained statistics. (*Id.*) This “Second Stage” also discloses or renders obvious claims 1 and 6.

42. An overview of Young’s process is shown in Figure 2:

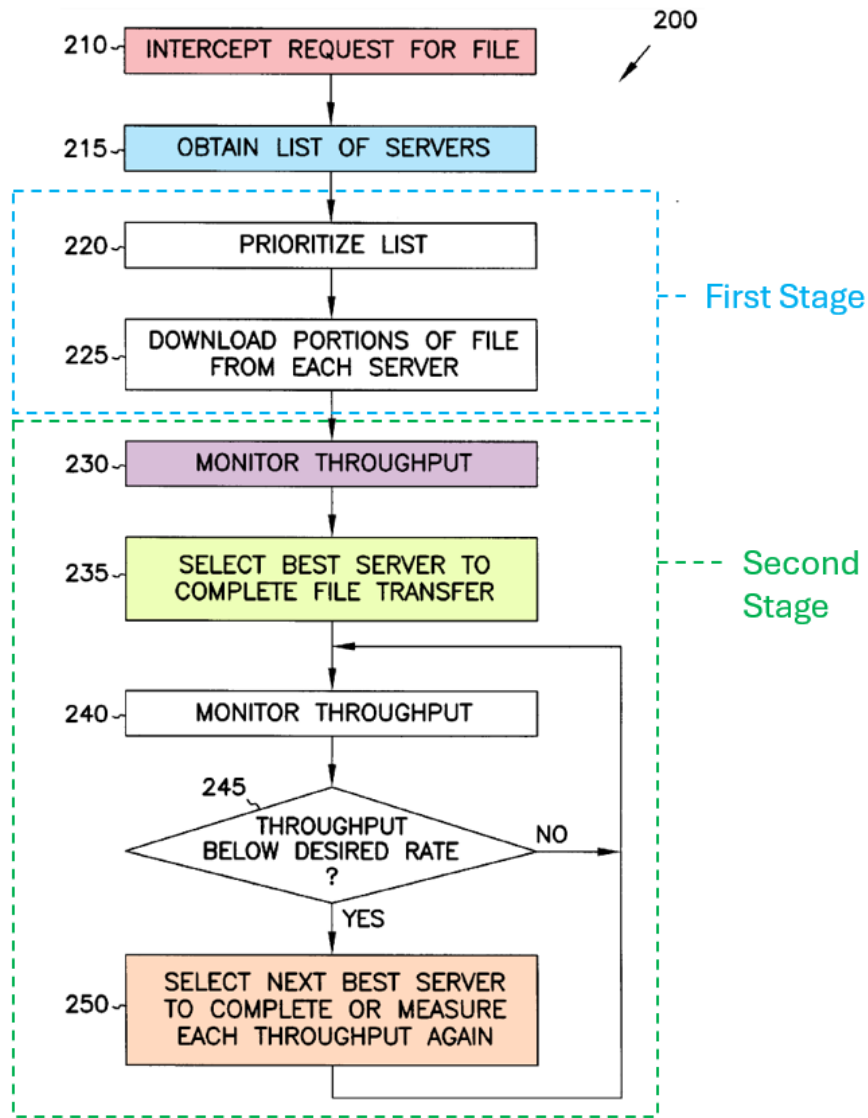


FIG. 2

(EX-1008 (Young), Fig. 2¹.)

¹ Figures in this Declaration may be colored and/or annotated for clarity.

1. Claim 1

a. Preamble

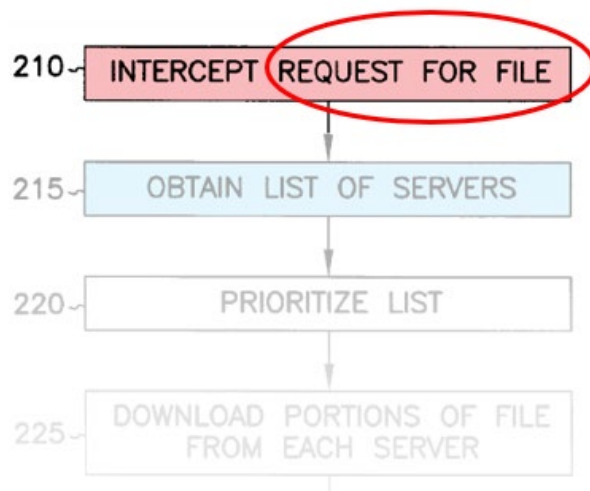
43. The preamble recites “[a] non-transitory computer readable storage medium including computer readable code” that can be executed by a computer. Young discloses this.

44. Young discloses a conventional computer (e.g., a PC) that contains storage and executes code (an “applet”). (EX-1008 (Young), 2:54-55 (PC), 5:29-36 (“software stored on computer readable media” as part of “the browser itself or as a separate application, or even as part of the operating system.”), 2:54-3:17, 3:66-4:15, Fig. 1.) Accordingly, Young discloses the preamble.

b. Element 1[a]: Sending a Request

45. Element 1[a] recites “send[ing] a request to a network-based server, the request including a unique identifier for identifying an audio stream.” Young discloses or renders obvious this claim element.

46. Young discloses that a user can request a file stored on an internet-accessible server. (EX-1008 (Young), 1:64-66 (“request”), 4:9, Abstract.) Young’s Figure 2 shows the user’s “request,” which the applet intercepts at step 210:



(*Id.*, Fig. 2 (excerpt).)

47. The request includes a uniform resource locator (URL) for the file, and each file has a different URL. (EX-1008 (Young), 1:12-19, 3:66-4:7.) Thus, the request includes a unique identifier for the requested file.

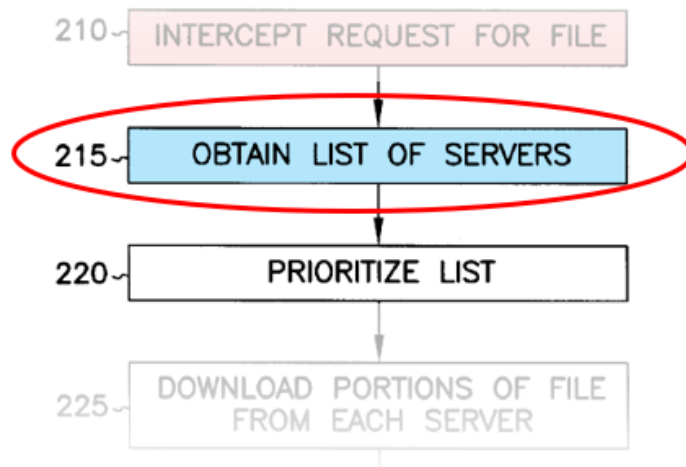
48. Young discloses that the requested file (information) can be “multimedia data.” (EX-1008 (Young), 3:66-4:5.) A person of ordinary skill in the art would have understood that an audio stream is one type of “multimedia data.”

49. Young therefore discloses or renders obvious sending a request to a network-based server (e.g., internet servers), the request including a unique identifier (e.g., URL) for identifying an audio stream (e.g., multimedia audio file). Thus, Young discloses or renders obvious claim element 1[a].

c. Element 1[b]: Loading a List of Servers

50. Element 1[b] recites “load[ing] a list of library servers received from the network-based server, the list of library servers determined in dependence upon the unique identifier.” Young discloses or renders obvious this claim element.

51. Young’s applet obtains a list of library servers based on the user’s request, as shown at step 215:



(EX-1008, Fig. 2 (“obtain list of servers”); *id.*, 4:28-30 (applet 200 “obtains the list of possible servers), 1:66-2:1, claim 7 (“obtains a list of servers having a copy of [the] desired file”).) The list is received from the network-based server. (*Id.*, 4:28-30 (applet obtains list “from the server identified by the initial link”).)

52. Young further discloses that the list is determined in dependence upon the unique identifier (URL) because the list includes only library servers that contain the requested file identified by the URL. (EX-1008 (Young), 4:16-21 (list identifies servers “which contain the desired file”), 1:66-2:1 (list contains “locations from

which to download the file”), claim 7 (obtain list of servers “having a copy of a desired file”).)

53. Accordingly, Young discloses or renders obvious loading a list of library servers (e.g., those having a copy of the requested file) received from the network-based server (e.g., server identified by initial link), the list of library servers determined in dependence upon the unique identifier (URL). Thus, Young discloses or renders obvious claim element 1[b].

d. Element 1[c]: Maintaining Server Statistics

54. Element 1[c] recites “maintain[ing] service level statistics for each library server in the list of library servers.” Young discloses or renders obvious this claim element.

55. Young discloses maintaining service level server statistics. Young prioritizes the list of servers (step 220) by measuring the latency (response time) of each server. (EX-1008 (Young), 4:31-38, Fig. 2.) Latency is measured by sending a “ping” to each server on the list and “keeping track of the amount of time for each server to respond.” (*Id.*) Young then downloads a portion of the requested file from each server in the list (step 225) and measures the throughput of each server during the download (step 230). (*Id.*, 4:39-57, 2:5-13, Fig. 2.) Each server’s ping is used to improve the measured throughput’s accuracy and “obtain a better indication of

true bit rate.” (*Id.*, 4:41-45.) Young then selects the best server to download the rest of the requested file (step 235). (*Id.*, 4:54-57, 2:5-13, Fig. 2.)

56. Young continues to monitor the selected server’s throughput (step 240) and determines a “desired [throughput] rate” for it. (EX-1008 (Young), 4:58-5:3.) The desired rate is a minimum threshold for server performance and is a percentage of the server’s throughput (determined at step 230). (*Id.*, 5:1-6.) If the throughput of the selected server falls below the desired rate (step 245), a second server (as determined from steps 225-235) is selected (step 250). (*Id.*, 4:60-62, 2:16-22.) The download continues using the second server as its throughput is monitored. (*Id.*, 4:60-65.) The desired rate for the second server, based on the throughput measured in step 230, is established. (*Id.*, 5:3-6.)

57. The service level statistics of each server (e.g., ping, throughput, and bit rate) are maintained by the applet. Ping is maintained to provide the preliminary ranking and further to obtain a “better indication of true bit rate.” (EX-1008 (Young), 4:41-45.) Throughput and bit rate are maintained to rank servers, calculate the “desired rate” for selected servers, and select the next best server if necessary. (*Id.*, 4:40-57, 5:1-9.) Thus, the service level statistics for each server in the list are maintained.

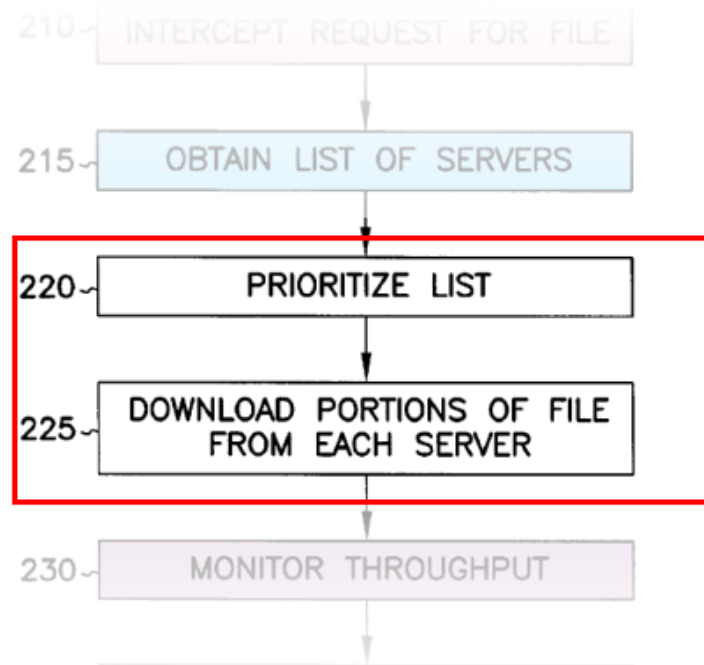
58. Young therefore discloses or renders obvious maintaining service level statistics (e.g., ping, throughput, and/or bit rate) for each library server in the list of library servers. Thus, Young discloses or renders obvious claim element 1[c].

e. Element 1[d][i]: Selecting a First Server

59. Element 1[d][i] recites “select[ing] a first library server from the list of library servers in dependence upon the service level statistics.” Young discloses or renders obvious this claim element in two ways.

i. First Stage Server Selection

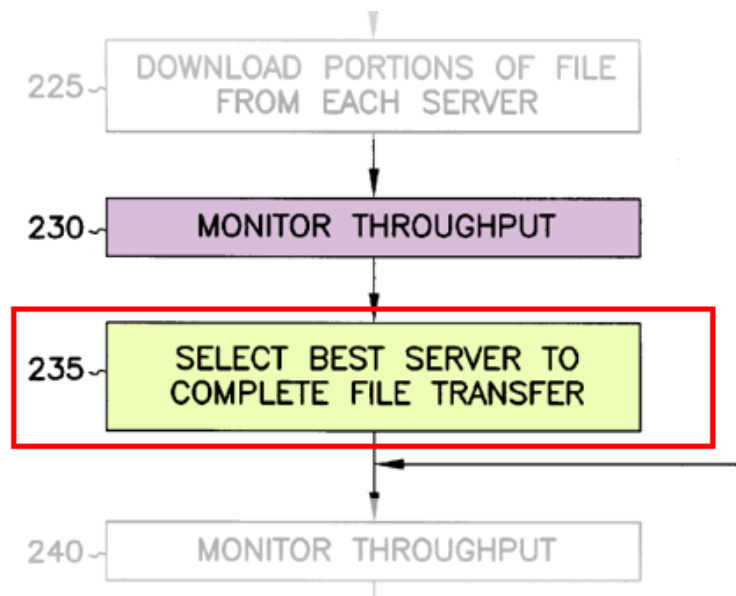
60. First, Young discloses selecting a first library server when it selects the “highest priority server” and downloads a portion of the requested file from that server at step 225:



(EX-1008 (Young), Fig. 2 (excerpt), 4:39-40, 2:10-13, 5:10-19, claim 11.) That server is selected based on the service level statistics, e.g., ping. (*Id.*, 4:31-38.)

ii. Second Stage Server Selection

61. Second, Young discloses selecting a first library server when it selects the “best server” for downloading the rest of the requested file at step 235:



(EX-1008 (Young), Fig. 2 (excerpt), 4:54-57 (the “highest throughput server which is selected as the optimal server”), 5:10-19, claim 11 (“selecting an optimal server”).) This server is selected based on the service level statistics, e.g., throughput, ping, and bit rate. (*See* ¶¶54-58, above.)

62. Young therefore discloses or renders obvious selecting a first library server (first server at step 225 or “best server” at step 235) from the list of library servers in dependence upon the service level statistics (e.g., ping, throughput, bit rate). Thus, Young discloses or renders obvious claim element 1[d][i].

f. Element 1[d][ii]: First Server Having a Plurality of Files

63. Element 1[d][ii] recites “the first library server having a plurality of digital audio files, each digital audio file in the plurality of digital audio files including a different segment of the audio stream.” Young discloses or renders obvious this claim element.

64. Young discloses downloading “information” via a network (e.g., the Internet). (EX-1008 (Young), 3:66-4:5.) Young discloses the information “may be referred to as a file,” and the file may contain “multimedia data.” (*Id.*) Young further discloses downloading different segments of the file/information from different servers. (*Id.*, Abstract, 1:59-63 (content stored on and “downloaded from different servers”), 4:58-5:9 (downloading portions from multiple servers), 5:19-21, Fig. 2 (step 225 (“download portions of file from each server”))).

65. As I discuss above, a person of ordinary skill in the art would have understood Young’s “multimedia data” to include an audio stream. (*See* ¶¶45-49, above.) A person of ordinary skill also would have understood that the different segments or portions of the audio stream would be stored as individual files. A person of ordinary skill would have had so understood because storing multimedia (e.g., audio) file segments as individual media files was conventional, as the ’740 patent admits and many references demonstrate. (EX-1001 (’740 patent), 2:16-18 (audiobook segmented into 24 audio files); *see* ¶¶35-36, above.)

66. It also would have been obvious to a person of ordinary skill that the separately downloadable segments of the multimedia data (audio) would be stored as individual files. A person of ordinary skill would have been motivated to store the individually downloadable portions of the multimedia files as smaller files because such segmentation was conventional and a person of ordinary skill understood the benefits of smaller files (e.g., faster download speeds and allowing content to be provided to devices with less storage). (EX-1004 (Lindahl) ¶¶[0006]-[0007] (smaller files provide faster download speeds and allow content to be downloaded to devices with small storage space); EX-1011 (Seed) ¶[0057] (segmentation enhances network performance because many users do not need the complete large audio file); EX-1006 (Yoshimura), 1780 (segmentation of larger files into smaller segments “leads to efficient cache memory and network resource utilization”), 1778-81; *see* ¶¶35-36, above.) And, because this was well known, a person of ordinary skill would have reasonably expected success in providing a multi-file audio stream using Young’s system.

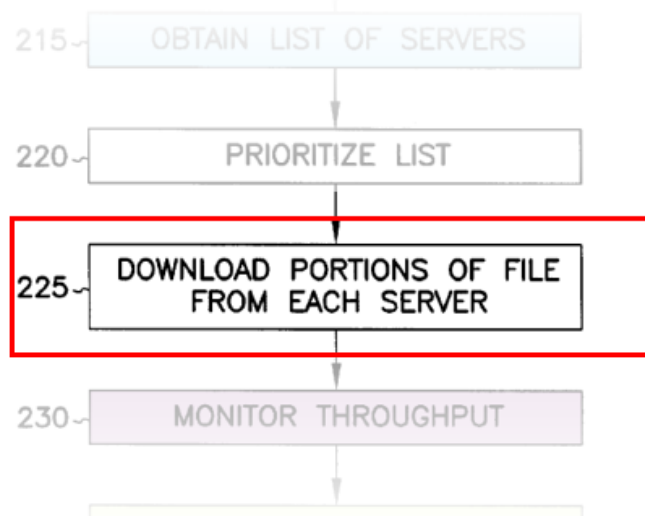
67. Thus, Young discloses or renders obvious claim element 1[d][ii].

g. Element 1[e]: Downloading a First Audio File

68. Element 1[e] recites “download[ing] a first digital audio file from the plurality of digital audio files for playback with a media player.” Young discloses or renders obvious this claim element in two ways.

i. First Stage Server

69. First, Young discloses downloading a first portion of the file from the highest priority server at step 225:

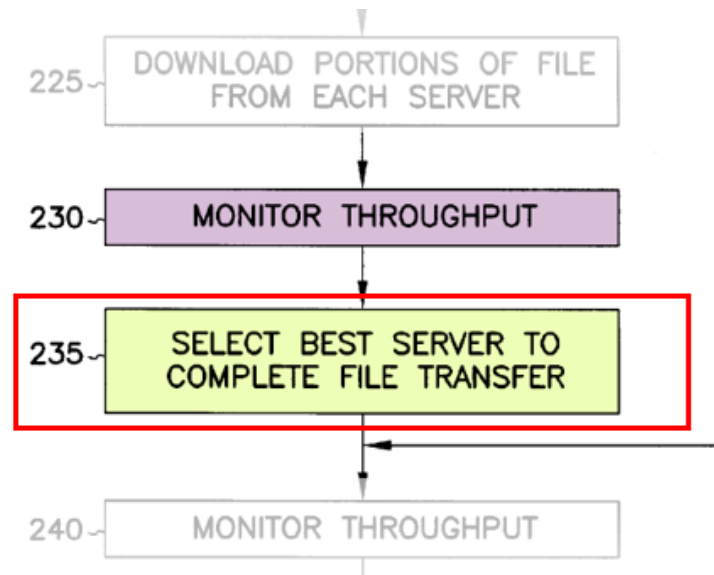


(EX-1008 (Young), Fig. 2 (excerpt), 4:39-46.)

ii. Second Stage Server

70. Second, Young discloses downloading a portion of the multimedia file from the first server, i.e., the “server which is selected as the optimal server.”

(EX-1008 (Young), 4:54-57, 5:10-28, claim 11 (download desired file).)



(*Id.*, Fig. 2 (excerpt).)

71. As I discuss above, a person of ordinary skill in the art would have understood each segment to be stored as a digital multimedia (e.g., audio) file, or this would have been obvious. (See ¶¶63-67, above.) Accordingly, Young discloses or renders obvious downloading a first digital audio file (“first portion” at step 225 or rest of the file at step 235) from the plurality of digital audio files (all files in audio stream) for playback with a media player (on Young’s PC). (*Id.*)

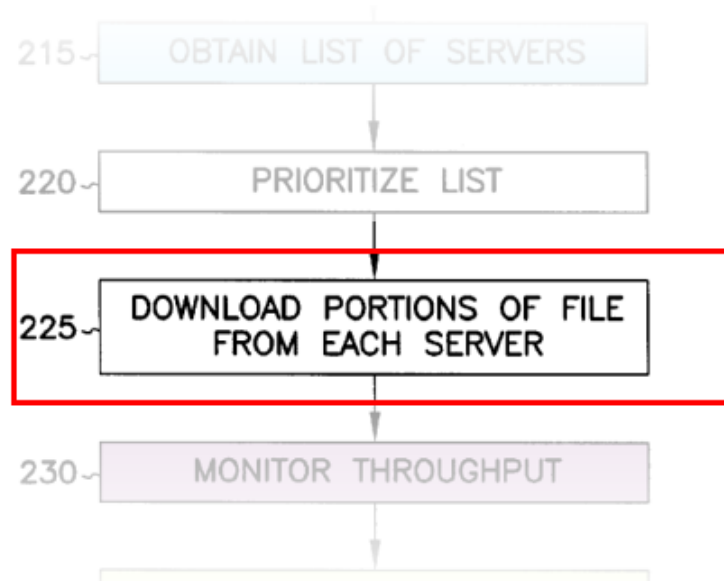
72. Thus, Young discloses or renders obvious claim element 1[e].

h. Element 1[f][i]: Selecting a Second Server

73. Element 1[f][i] recites “select[ing] a second library server from the list of library servers in dependence upon the service level statistics.” Young discloses or renders obvious this claim element in two ways.

i. First Stage Server Selection

74. First, Young discloses selecting a “second server” from the list. (EX-1008 (Young), 4:48-54.) That server is selected second based on the service level statistics, e.g., it had the second-fastest ping. (*Id.*, 4:31-38, 2:7-10.) This is shown in Step 225 (download a portion from “each server,” which would include a second server):

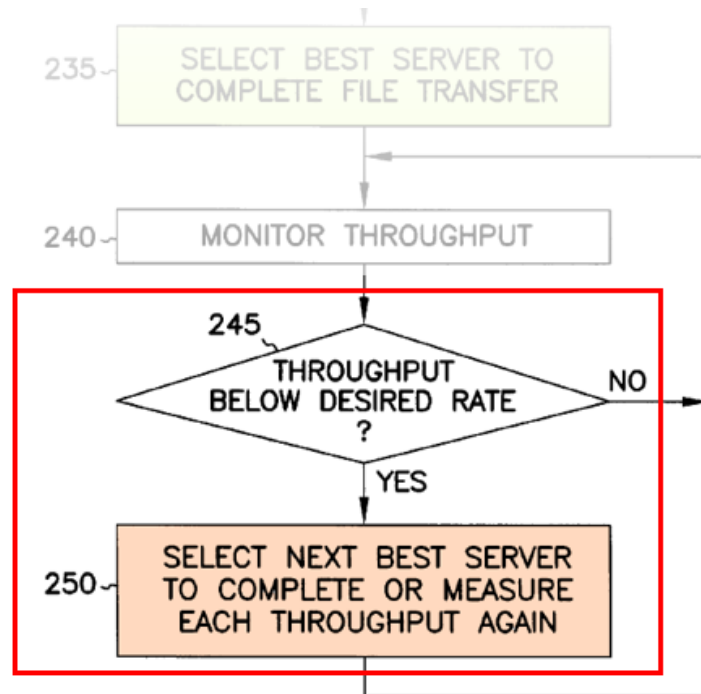


(*Id.*, Fig. 2 (excerpt), 4:45-51.)

ii. Second Stage Server Selection

75. Second, Young discloses that, “[i]f the performance [of the optimal server] falls below a desired rate ..., a different server may be selected.” (EX-1008 (Young), 4:60-65.) The second server may be the “next server on the previously generated performance ranked list[.]” (*Id.*) Thus, this server is selected based on

the maintained service level statistics (e.g., ping, throughput, bit rate, and/or a percentage of the throughput). (*Id.*, 4:58-62, 5:1-6.)



(*Id.*, Fig. 2 (excerpt).)

76. Thus, Young discloses or renders obvious claim element 1[f][i].

i. Element 1[f][ii]: Second Server Having a Copy of the Files

77. Element 1[f][ii] recites “the second library server having a copy of the plurality of digital audio files.” Young discloses or renders obvious this claim element.

78. Young discloses that each library server in the list, and therefore either second library server I discuss above, contains copies of the requested multimedia content. (See ¶¶63-67, above; EX-1008 (Young), 4:58-5:9, claim 7.)

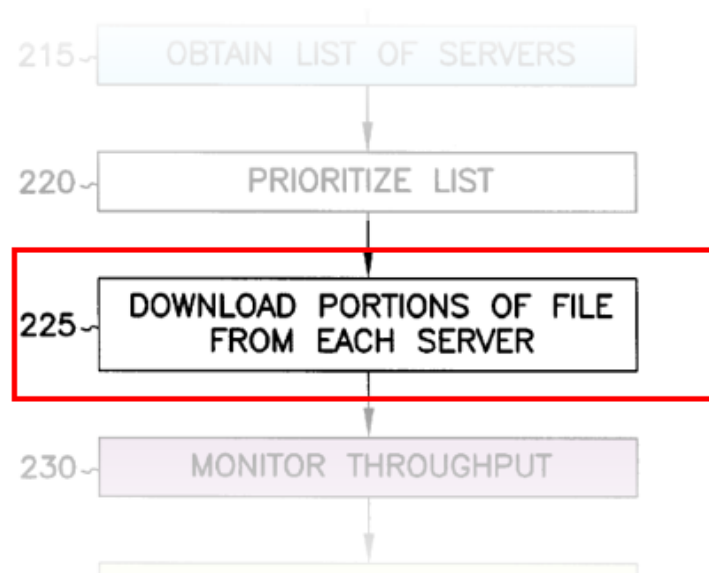
79. Thus, Young discloses or renders obvious claim element 1[f][ii].

j. Element 1[g]: Downloading a Second Audio File from the Second Server

80. Element 1[g] recites “download[ing] a second other digital audio file from the second library server for playback with the media player.” Young discloses or renders obvious this claim element in two ways.

i. First Stage Server

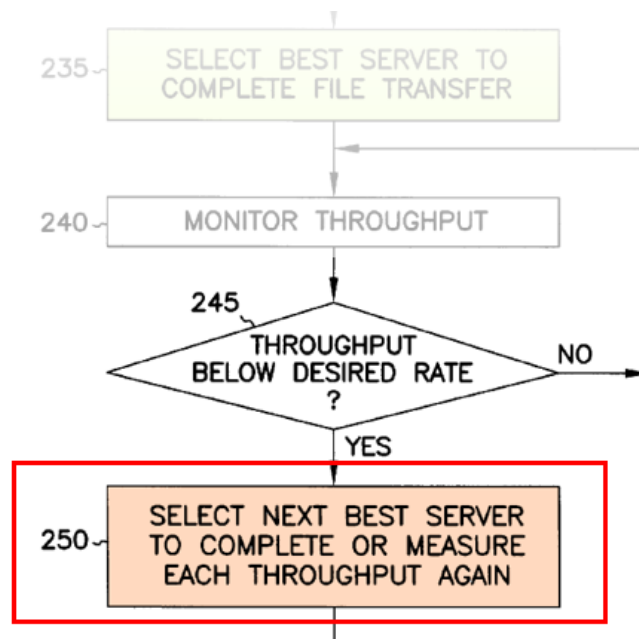
81. First, Young discloses downloading a second portion (e.g., file) from the second-highest priority server. (EX-1008 (Young), 4:48-54, 2:7-10, Fig. 2 (step 225 (“download portions of file from each server”).)



(*Id.*, Fig. 2 (excerpt).)

ii. Second Stage Server

82. Young discloses that, after selecting a second server (“next best server” in Figure 2), the client downloads subsequent portions of the multimedia content from that server. (EX-1008 (Young), 4:60-5:6 (new server is “selected to complete the transmission”), 5:10-28, Fig. 2 (step 250).)



(*Id.*, Fig. 2 (excerpt).)

83. As I discuss above, a person of ordinary skill in the art would have understood the portions to be audio files (*see* ¶¶63-67, above) for playback with the media player (e.g., on Young’s client device) (*see* ¶¶68-72, above).

84. Thus, Young discloses or renders obvious claim element 1[g].

85. Accordingly, for at least the reasons I discuss above, Young renders claim 1 obvious as a whole.

2. Claim 6

86. Claim 6 depends from claim 1 and further recites “wherein the service level statistics include historical transfer rates for each library server in the list of library servers.” Young discloses or renders obvious the additional claim element of claim 6.

87. Young’s service level statistics include “transfer rates” for each server in the list because Young’s statistics include “throughput” measured in “bits per second.” (EX-1008 (Young), 4:40-41, 4:54-57.) Young’s service level statistics are historical because they are obtained, stored, and later used to rank and select servers, and to determine performance thresholds. (See ¶¶54-62, 80-85, above.)

88. This limitation also would have been obvious because Young teaches using server statistics to select a server, and using historical transfer rates to do so was widely known in the art by 2005. Moreover, a person of ordinary skill would have been motivated to use historical transfer rates (in addition to or instead of the performance statistics Young discloses) because doing so would help improve the performance of the system. Further, because using historical transfer rates was well known, using such statistics to select the server in Young represents nothing more than a simple substitution of one known element (historical transfer rates) for another (Young’s statistics, e.g., ping) to obtain predictable results (selecting the historically fastest server). It also represents nothing more than using a known technique (server

selection based on historical transfer rates) to improve a similar device and method (Young's CDN) in the same way. It also applies a known technique to a known device and method that is ready for improvements and yields predictable results (server selected at least in part based on historical transfer rates).

89. Thus, Young discloses or renders obvious the additional claim element of claim 6.

90. Accordingly, for at least the reasons I discuss above, Young renders claim 6 obvious as a whole.

B. Claims 1, 6, and 11 Would Have Been Obvious in View of Young, Yoshimura, and Copley.

91. As I discuss above, Young discloses or renders obvious each limitation of claim 1. (See ¶¶43-85, above.) For at least those reasons and the reasons I discuss below, claims 1, 6, and 11 would have been obvious in further view of Yoshimura and Copley. As I discuss below, Yoshimura discloses segmenting audio files into a plurality of smaller files. Copley discloses maintaining server performance statistics and selecting a content server based thereon. I incorporate my above discussion of Young here, and only discuss claim elements for which Yoshimura and/or Copley are relevant in combination with Young.

1. Claim 1

a. Elements 1[c], 1[d][i], and 1[e]

92. These claim elements recite maintaining service level statistics for each server and selecting the first and second servers in dependence upon those statistics. Element 1[c] recites “maintain[ing] service level statistics for each library server in the list of library servers.” Element 1[d][i] recites “select[ing] a first library server from the list of library servers in dependence upon the service level statistics.” Element 1[e] recites “download[ing] a first digital audio file from the plurality of digital audio files for playback with a media player.” Copley renders obvious these claim elements.

93. Copley discloses a CDN with a client-side computer that accesses media content on “provider” servers. (EX-1009 (Copley), Abstract, ¶¶[0006], [0014].) The client includes “a [media] player for playing content to a user.” (*Id.* ¶[0006].) The client “collect[s] usage and performance data,” and “information about failures, latency, and other information” for each server. (*Id.* ¶¶[0015], [0074]; *id.* ¶[0016].) Copley stores these statistics to rank and reprioritize servers. (*Id.* ¶¶[0015]-[0016], [0051], [0073]-[0074].) Accordingly, Copley discloses maintaining service level statistics (e.g., usage and performance data) for each library server in the list of library servers and selecting a server based on those statistics.

94. A person of ordinary skill in the art would have been motivated to maintain service level statistics, and select the appropriate server based on those statistics, as disclosed in Copley, in Young's system.

95. First, Young suggests the combination because it seeks to optimize the download of media content by selecting an optimal server. (EX-1008 (Young), Abstract, 1:58-2:25, 2:48-53, 5:11-15.) Copley discloses maintaining performance data to determine an optimal server while achieving improved performance and reduced costs. (EX-1009 (Copley) ¶¶[0013]-[0019], [0051], [0073]-[0074].)

96. Second, because the use of service level statistics in CDNs was well known (*see* ¶¶32-34, above), maintaining service level statistics (as taught by Copley) represents nothing more than a simple substitution of one known element (Copley's performance data) for another (Young's server statistics) to obtain predictable results (maintaining server performance data used to select servers). It also represents nothing more than using a known technique (maintaining performance data, as in Copley) to improve a similar device and method (Young's CDN and server selection) in the same way. The combination further applies a known technique (Copley's maintenance of performance data) to a known device and method (Young's CDN and server selection) that is ready for improvement and yields predictable results.

97. A person of ordinary skill would have had a reasonable expectation of success in maintaining the service level statistics as in Copley because it was widely

known and routine to do so. (EX-1009 (Copley) ¶¶[0015]-[0016], [0051], [0073]-[0074]; EX-1087 (Ahuja), Abstract; EX-1089 (Fei), 1, 3-5.) Further, Young discloses using server performance information to rank servers and determine an optimal server, which would have given a person of ordinary skill a reasonable expectation of success in doing so using the conventional performance data disclosed in Copley. (EX-1008 (Young), 4:31-5:9.)

98. Thus, Copley renders obvious claim elements 1[c], 1[d][i], and 1[e].

b. Elements 1[d][ii]-1[g]

99. Element 1[d][ii] recites that the first server has “a plurality of digital audio files, each digital audio file in the plurality of digital audio files including a different segment of the audio stream.” Element 1[e] recites downloading a first “digital audio file from the plurality of audio files for playback with a media player.” Element 1[f][ii] recites that the second server has “a copy of the plurality of digital audio files.” Element 1[g] recites downloading a second “digital audio file from the second library server for playback with the media player.” Yoshimura discloses or renders obvious these claim elements.

100. Yoshimura discloses or renders obvious that the requested multimedia content is an audio stream stored on the servers as a plurality of smaller digital audio files.

101. Yoshimura’s CDN divides original media content including audio streams into multiple digital files, each file containing a segment of the content:

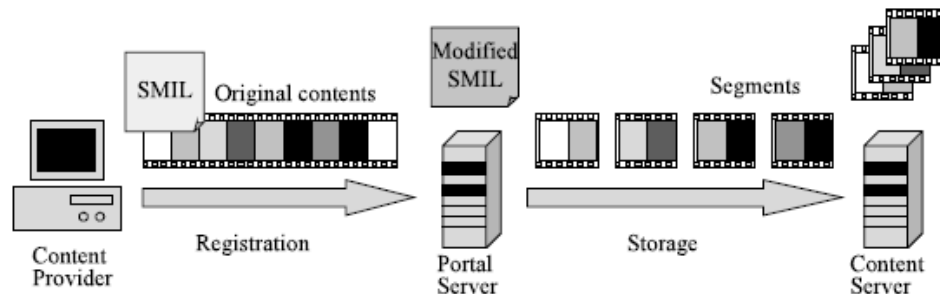


Fig. 3 Content segmentation and SMIL modification at portal server.

(EX-1006 (Yoshimura), Fig. 3; *id.*, 1780 (segmentation “increases the number of [] files”).) Yoshimura provides an example in which an original audio file “content-A.mp4” is segmented and stored in three audio files (“content-A- $\{1,2,3\}$.mp4”). (*Id.*, 1780-81.) The segmented audio files are provided to the content servers. (*Id.*, 1780 (system “sends the segments to a content server and requires it to store them”), Fig. 3.)

102. It would have been obvious to a person of ordinary skill in the art, in view of Yoshimura, that the requested multimedia “information” described in Young would be an audio stream and that the “file” described in Young would be a segmented audio stream stored as a plurality of audio files. Such a modification of Young would render obvious the limitations in claim 1 that recite an “audio stream” and “a plurality of digital audio files.”

103. A person of ordinary skill in the art would have been motivated to segment the media content in Young’s system, and therefore store on the servers a plurality of digital audio files as taught by Yoshimura, for many reasons.

104. First, Yoshimura teaches that by “dividing content into several segments, surrogates [servers] can store and remove the content at the granularity of the segments, which leads to efficient cache memory and network resource utilization.” (EX-1006 (Yoshimura), 1780.) A person of ordinary skill would have been motivated to segment the multimedia content described in Young into smaller files to achieve these benefits.

105. Second, Young recognizes that throughput is monitored “because performance may change over time for very large files.” (EX-1008 (Young), 4:59-60.) Yoshimura explains that content segmentation is “effective especially for large [files].” (EX-1006 (Yoshimura), 1780.) Thus, a person of ordinary skill would have been motivated to segment Young’s media files, especially larger audio files, to reduce download times and the likelihood that the selected server’s performance decreases during such downloads.

106. Third, segmenting media content into smaller files was widely known and supported, as I discuss above and as the ’740 patent admits. (See ¶¶35-36, above.) Accordingly, modifying Young’s system to store an audio stream as multiple digital audio files represents nothing more than a simple substitution of one

known element (audio content stored in multiple smaller files) for another (one larger file) to obtain predictable results (more efficient storage and faster content delivery). This combination represents nothing more than using a known technique (storing audio as multiple files) to improve a similar device and method (Young’s CDN and storage of multimedia files) in the same way. Similarly, this combination applies a known technique (storing segmented media content in multiple files) to a known device and method (Young’s CDN) that is ready for improvements and yields predictable results.

107. Because content segmentation was well known and widely used, a person of ordinary skill would have had a reasonable expectation of success in segmenting an audio file into a plurality of audio files (as taught by Yoshimura) and providing such audio via Young’s system.

108. Thus, Yoshimura discloses or renders obvious claim elements 1[d][ii]-1[g].

109. Accordingly, for at least the reasons I discuss above, Young, Yoshimura, and Copley render claim 1 obvious as a whole.

2. Claim 6

110. Claim 6 depends from claim 1 and further recites “wherein the service level statistics include historical transfer rates for each library server in the list of library servers.” Copley discloses the additional claim element of claim 6 because

Copley discloses maintaining “historical” server statistics. (EX-1009 (Copley) ¶¶[0013]-[0017], [0047]-[0050].)

111. Thus, for at least the reasons I discuss above, Young, Yoshimura, and Copley render claim 6 obvious as a whole.

3. Claim 11

a. Element 11[a]

112. Claim 11 depends from claim 1. Element 11[a] recites “download[ing] a descriptor file from the first library server, the descriptor file for ordering the plurality of digital audio files, the descriptor file including at least one of a start time, an end time, and a play time of each digital audio file in the plurality of digital audio files within the audio stream.” A descriptor file contains information about media content, such as relative positioning and timing of content segments. (EX-1001 (’740 patent), 2:48-57.) Young and Yoshimura together render obvious claim element 11[a].

113. The use of such descriptor files was well known by 2005. One well-known descriptor file was called a “Synchronized Multimedia Integration Language” (“SMIL”) file. (EX-1006 (Yoshimura), 1778-79; EX-1003 (DTB), 6-7; EX-1052 (Grigorovitch) ¶[0041].) SMIL is “a standard for definition and playback of multimedia presentations over the Internet” that “defines the sequence of playback” of media. (EX-1003 (DTB), 21; EX-1090 (Kennedy), 85-114; EX-1007

(Bulterman), 125-55.) A SMIL file represents media content as a sequence of segments that play out sequentially and allows each segment to include different types of media (e.g., video, audio, and text) rendered simultaneously. (EX-1003 (DTB), 21, 26; EX-1090 (Kennedy), 21, 40, 85-114, 115-31; EX-1007 (Bulterman), 88, 92, 125-55.)

114. Yoshimura discloses downloading a SMIL file from the content server. (EX-1006 (Yoshimura), 1779, Fig. 1.) Yoshimura’s SMIL file orders a plurality of digital files (“content-A-1.mp4,” “content-A-2.mp4,” and “content-A-3.mp4”) by placing them in a defined sequence, as shown in Figure 5:

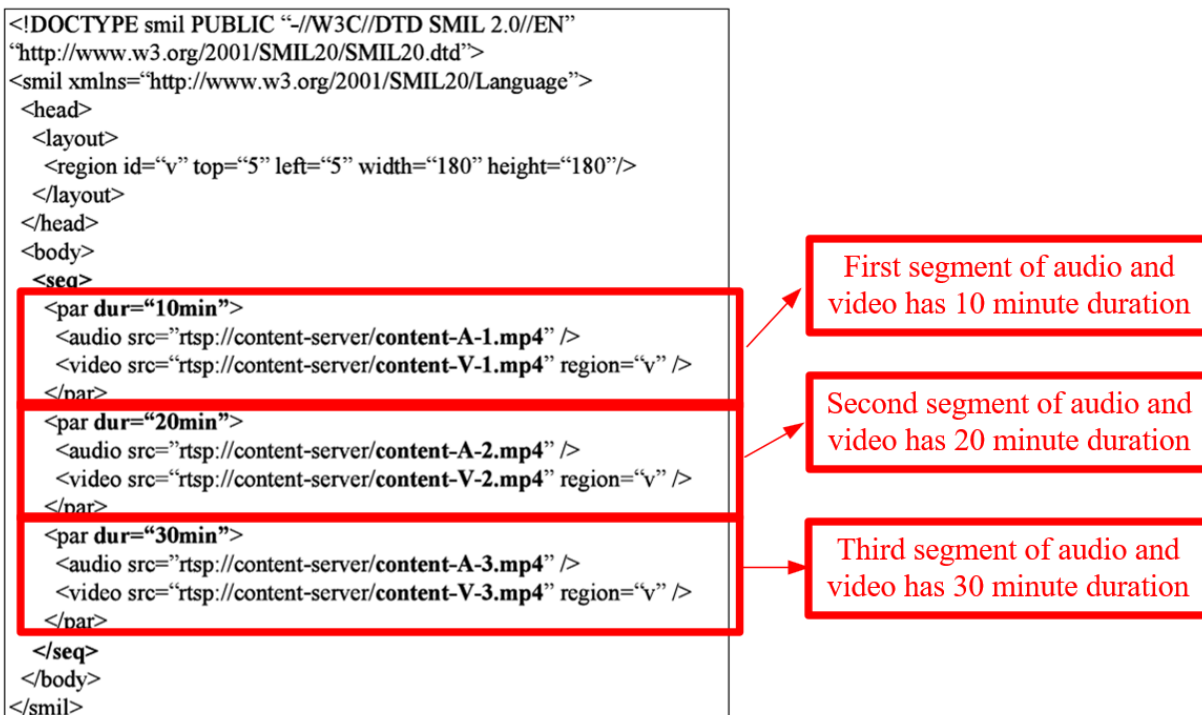


Fig. 5 Modified SMIL file after content segmentation.

(*Id.*, Fig. 5, 1780-81.) In this example, the audio files are “sequentially played.” (*Id.*, 1781; *id.* (“the part located between <seq> and </seq> indicates that these segments are sequentially played”).)

115. As shown in Figure 5 above, the SMIL file provides a duration (“dur” attribute) of each media content segment (<par> element). (EX-1006 (Yoshimura), Fig. 5, 1781 (“the durations of the segments are ten, twenty, and thirty minutes long”).) Because the duration of each segment is provided and the segments are played sequentially, “the timing relation among them is evident.” (*Id.*; EX-1090 (Kennedy), 88-90, 98-99; EX-1007 (Bulterman), 125, 127, 141 (“In general, begin+dur=end”), 142-47.) Thus, the SMIL structure and duration of each segment also provides a start and end time for each segment. (*Id.*) For example, the first segment with a start time of 0 minutes has a duration of 10 minutes and therefore an end time of 10 minutes. The second segment’s start time is therefore 10 minutes and, because it has a duration of 20 minutes, its end time is 30 minutes. Thus, the start and end time of each segment is either expressly disclosed by, or least inherent in, Yoshimura’s SMIL file. Alternatively, it would have been obvious to a person of ordinary skill to add the start and end times to Yoshimura’s SMIL file as was widely known. (EX-1007, 141-42 (<par> elements can have begin/end attributes).)

116. Together, Yoshimura and Young render obvious downloading a descriptor (SMIL) file from the first library server (e.g., Young’s optimal server) for

ordering the plurality of digital audio files (e.g., audio segments) that includes at least one of a start time, an end time, and a play time (e.g., duration) of each digital audio file in the plurality of digital audio files within the audio stream. Thus, Yoshimura and Young render obvious claim element 11[a].

b. Element 11[b]

117. Element 11[b] recites “wherein the computer determines the first digital audio file for playback using a time offset external to the descriptor file and the at least one of the start time, end time, and play time of each digital audio file in the plurality of digital audio files.” Yoshimura discloses or renders obvious claim element 11[b].

118. Yoshimura discloses that the user may invoke conventional media controls—pause, resume, skip forward, and skip backwards. (EX-1006 (Yoshimura), 1783 (client may pause, skip, or “go[] back”).) A person of ordinary skill in the art would have understood that these controls provide a time offset to the client device, as the ’740 patent confirms. (EX-1001 (’740 patent), 2:55-57 (explaining that “time offsets” are “typically provided via ... rewind/fast-forward functions”).)

119. Additionally, because “SMIL provides the timing information for playing or displaying, it can easily manage the timing relations among the segments,” and a client can, “[b]y parsing the SMIL file, ... determine which segments are the next to the segments currently served and how long the segments will last.”

(EX-1006 (Yoshimura), 1780.) Accordingly, a person of ordinary skill would have understood Yoshimura to disclose that a client determines a first segment (the next or previous segment) using a time offset provided by a user (e.g., skipping forward/backward) and timing information in a SMIL file (e.g., segment start time, end time, and duration).

120. Yoshimura therefore discloses or renders obvious determining the first digital audio file for playback using a time offset external to the descriptor file (e.g., time offset from user input) and the at least one of the start time, end time, and play time of each digital audio file. Thus, Yoshimura discloses or renders obvious claim element 11[b].

121. A person of ordinary skill would have been motivated to implement Yoshimura's SMIL-file-based controls in Young's system for several reasons.

122. First, a person of ordinary skill would have been motivated to store on servers, and download to clients, segmented audio files stored as a plurality of smaller files, as Yoshimura teaches. (See ¶¶99-109, above.) Given that motivation, a person of ordinary skill also would have been motivated to use Yoshimura's descriptor file to organize and synchronize the files, as well as provide a means for navigating the files. Thus, Yoshimura's SMIL files would provide content access, synchronization, and navigation controls. (EX-1006 (Yoshimura), 1778, 1783.)

123. Second, Young sought to optimize the download of content by selecting the most efficient server. (EX-1008 (Young), 1:58-59, 5:10-15.) This would have motivated a person of ordinary skill to look to Yoshimura, which uses SMIL files to efficiently provide segmented media content that a user can navigate through while also “enhanc[ing] streaming media quality” and improving network efficiency. (EX-1006 (Yoshimura), 1778, 1783.)

124. Third, as I discuss above, the use of CDNs and SMIL files was widely known, as were conventional navigation controls. (See ¶¶32-34, 112-116, above; EX-1006 (Yoshimura), 1779; EX-1007 (Bulterman); EX-1005 (Sull) ¶[0205], Fig. 9; EX-1035 (Weisman) ¶[0125], Fig. 3; EX-1085 (Corson) ¶[0014]; EX-1083 (Ottesen) ¶[0044]-[0045], [0048].) Thus, using a SMIL file to access and navigate segments of media content represents the simple addition of one known element (SMIL file) to another (segmented media content) to obtain predictable results (using a SMIL file to access and navigate media content segments). This combination also reflects the use of a known technique (using SMIL file to access and play media content segments) to improve a similar device and method (Young’s CDN) in the same way. Moreover, this combination applies a known technique (using SMIL file to access and play media content segments) to a known device and method (Young’s) that is ready for improvements and yields predictable results.

125. A person of ordinary skill would have had a reasonable expectation of success using a SMIL file and conventional navigation controls as taught by Yoshimura in Young’s system because using descriptor files and navigating media content were well known. (See ¶¶112-116; EX-1083 (Ottesen) ¶¶[0044]-[0045], [0048].) Moreover, Young teaches storing multimedia content on servers and downloading them for playback, and Yoshimura teaches segmenting the content and using SMIL files to access and navigate the segmented content on servers in a system similar to Young’s. (EX-1008 (Young), Abstract, 1:59-63, 4:39-5:9, 5:19-21; EX-1006 (Yoshimura), 1778-80.)

126. Accordingly, for at least the reasons I discuss above, Young, Yoshimura, and Copley render claim 11 obvious as a whole.

C. Claim 2 Would Have Been Obvious in View of Young and Kate, or in View of Young, Yoshimura, Copley, and Kate.

127. Claim 2 depends from claim 1 and further recites “wherein the unique identifier is an ISBN number.” As I discuss above, Young, alone or in combination with Yoshimura and Copley, discloses or renders obvious each limitation of claim 1. (See ¶¶43-85, 92-109, above.) Kate discloses the additional claim element of claim 2.

128. As I discuss above, Young discloses using a URL as a unique identifier of media content. (See ¶¶45-49, above.) Kate discloses accessing media content on a server using “server content locators, such as a URL.” (EX-1041 (Kate), Abstract,

2:21-3:16.) Kate further discloses that the server content locator can be a URN with an ISBN (e.g. “urn:isbn:088663475046569”). (*Id.*, 12:30-13:2.) Kate therefore discloses requesting media content using a unique identifier that is an ISBN.

129. Thus, Kate discloses the additional claim element of claim 2.

130. A person of ordinary skill in the art would have been motivated to combine Young and Kate, and to use an ISBN as a unique identifier in Young’s system.

131. First, Young suggests doing so because it discloses requesting content with a unique identifier. (*See* ¶¶45-49, above.) Kate discloses using an ISBN as a unique identifier for the same purpose, and a person of ordinary skill would have recognized the use of an ISBN number—an established standard—would be an efficient way to accurately and consistently identify the requested content (e.g., although content’s URL may change, an ISBN would still uniquely identify it). (EX-1041 (Kate), 12:30-13:2.)

132. Second, an ISBN was a well-established standard for identifying content and was therefore one of a finite number of options for the unique identifier in Young. (EX-1041 (Kate), 12:30-13:2 (use of URL, URN including an ISBN, or URI).)

133. Third, because ISBNs were a standard for identifying published media content such as audiobooks, the combination represents a simple substitution of one known element (URN with ISBN) for another (URL without an ISBN) to obtain

predictable results (identifying content). Fourth, the combination represents using a known technique (identifying content via ISBN) to improve a similar device and method (Young’s content retrieval) in the same way. The combination also applies a known technique (identifying content via ISBN) to a known device and method (Young’s) that is ready for improvement and yields predictable results.

134. A person of ordinary skill would have had a reasonable expectation of success because ISBNs were well known as a standard for identifying content. A person of ordinary skill would have also had reasonable expectation of success because the alteration would be trivial (replacing a URL with a URN comprising an ISBN). (EX-1041 (Kate), 12:30-13:2.)

135. Accordingly, for at least the reasons I discuss above, (i) Young and Kate and/or (ii) Young, Yoshimura, Copley, and Kate render claim 2 obvious as a whole.

D. Claim 5 Would Have Been Obvious in View of Young and Sull, or in View of Young, Yoshimura, Copley, and Sull.

136. As I discuss above, Young, alone or in combination with Yoshimura and Copley, discloses or renders obvious each limitation of claim 1. (See ¶¶43-85, 92-109, above.) Sull discloses the additional claim element of claim 5.

137. Claim 5 depends from claim 1 and further recites “wherein the unique identifier is obtained from one of a bookmark structure, a card catalog structure, and an advertising structure.”

138. Bookmarks were well known and widely used in media content. (*See* EX-1041 (Kate), 11:15-17:7; EX-1005 (Sull), Abstract, ¶¶[0005]-[0007], [0172]-[0175], [0188]-[0193]; EX-1004 (Lindahl) ¶¶[0015], [0035]; EX-1035 (Weisman), Abstract, ¶¶[0049]-[0054], [0060], [0135]-[0138]; EX-1033 (Hanson), 1:6-11, 2:17-30; EX-1028 (Hay) ¶¶[0027], [0040]-[0048].) Bookmarks were commonly used to save a user’s location in media content, often as a time offset from the beginning of the content, so that users could resume content at their last position. (*Id.*) Bookmarks typically included a unique identifier that identified the content to which it applied.

139. Sull discloses a “multimedia bookmark” that indexes a position in media content and uniquely identifies the content. (EX-1005 (Sull), Abstract, ¶¶[0005]-[0007] (conventional bookmarks include identifiers such as a file name, URL, or URI), [0172]-[0175] (“The positional information 212 may be composed of a URI, a URL, or the like[.]”), [0188]-[0193] (disclosing a “metadata identification (ID)” identifier), [0252]-[0261], Fig. 2.) When a bookmark is used, Sull’s system obtains a unique identifier from the bookmark, sends it to the server, and then renders content from a bookmarked position. (*Id.* ¶¶[0211] (“When the play-bookmark control is selected ... the URI or the like, bookmarked position, and metadata ID for the multimedia content to be played back are read from persistent storage.”), [0230]-[0231] (position information sent to server), Fig. 12.)

140. Accordingly, Sull discloses obtaining the requested content's unique identifier from a bookmark structure. Thus, Sull discloses the additional claim element of claim 5.

141. A person of ordinary skill in the art would have been motivated to use Sull's bookmarks in Young's system for several reasons.

142. First, doing so would have been desirable to provide the benefits of a bookmark in Young's system rather than playing content from the beginning when accessing content. (EX-1005 (Sull) ¶[0172].)

143. Second, Young discloses requesting content using an identifier such as a URL and obtaining media content from a position at which the media content was interrupted. (EX-1008 (Young), 1:12-42, 4:60-65.) Sull discloses a bookmark that can provide a unique identifier and a position in media content to efficiently identify and access media content at the appropriate position.

144. Third, because bookmarks were well known, the use of Sull's bookmark in Young's system represents nothing more than using a known technique (obtaining a unique identifier from a bookmark) to improve a similar device and method (Young's system and method of identifying content) to obtain predictable results (obtaining a unique identifier from a bookmark in Young's system). Similarly, this combination merely applies a known technique (obtaining a unique identifier from a bookmark) to a known device and method (obtaining a unique identifier from a

bookmark in Young’s system) that is ready for improvement and yields predictable results.

145. A person of ordinary skill would have had a reasonable expectation of success because bookmarks were well understood, and further because both Young and Sull use a URL or URI as a unique identifier of media content. (See ¶¶45-49, above.)

146. Accordingly, for at least the reasons I discuss above, (i) Young and Sull and/or (ii) Young, Yoshimura, Copley, and Sull render claim 5 obvious as a whole.

E. Claim 18 Would Have Been Obvious in View of Young and Ganesan, or in View of Young, Yoshimura, Copley, and Ganesan.

147. Claim 18 depends from claim 1 and further recites “wherein a size of each digital audio file in the plurality of digital audio files is selected in dependence upon network throughput rates.” As I discuss above, Young, alone or in combination with Yoshimura and Copley, discloses or renders obvious each limitation of claim 1. (See ¶¶43-85, 92-109, above.) Ganesan discloses the additional claim element of claim 18.

148. Ganesan discloses a CDN in which a “large file is fragmented intelligently” and the “number of segments are computed or determined periodically in accordance with the required transmission rate of the title, the minimum available network speeds, etc.” (EX-1010 (Ganesan), 4:35-41; *id.*, 12:31-40, 13:47-62, Fig. 2E.) This is shown in Figure 2E:

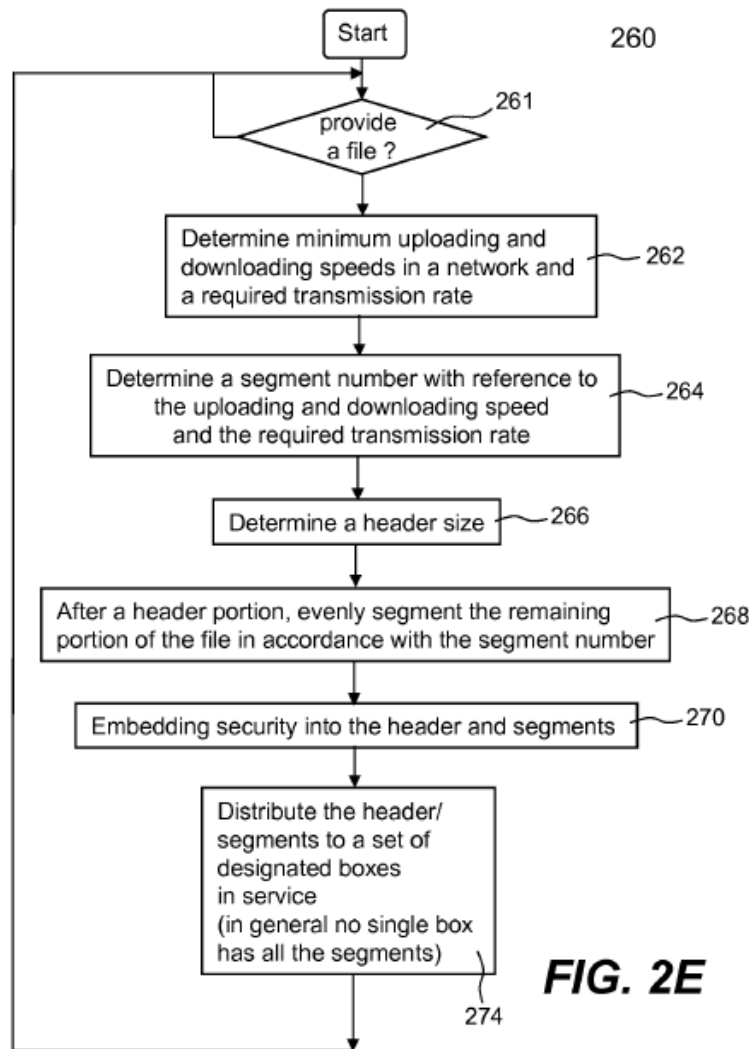


FIG. 2E

(*Id.*, Fig. 2E.)

149. Thus, Ganesan discloses the additional claim element of claim 18.

150. A person of ordinary skill in the art would have been motivated to select the size of audio segments in Young, and/or the size of the segmented audio files in the Young-Yoshimura combination, based on network throughput rates as taught by Ganesan.

151. First, Young and Yoshimura disclose storing and downloading media content in segments, but do not disclose how each segment's size is determined. A person of ordinary skill would have been motivated to look to other references, such as Ganesan, to understand how to determine the size, and Ganesan teaches doing so based on throughput rates to improve performance. (EX-1010 (Ganesan), 4:35-41 (such segmentation “best utilize[s] the network bandwidth and maximize[s] quality of service”), 13:56-61 (enables “efficient use of network speeds”).)

152. Second, Young sought to optimize the download of data such as media content by tracking throughput. (EX-1008 (Young), Abstract, 1:58-59.) This disclosure would have motivated a person of ordinary skill to look to Ganesan, which discloses improved efficiency transmitting content over a network by determining segment size based on the same measurement—throughput. (EX-1010 (Ganesan), 4:35-41.)

153. Third, because segmenting media content and measuring conventional server statistics such as throughput were well known, determining a segment size based on throughput represents nothing more than a simple substitution of one known element (determining segment size based on throughput) for another (arbitrarily determining segment size) to obtain predictable results (segments that can be efficiently transported over the network). This combination also represents nothing more than using a known technique (determining segment size based on throughput)

to improve a similar device and method (storing media segments in Young’s CDN) in the same way. Indeed, other references disclosed determining the size of a segment based on communications bandwidth. (EX-1086 (Shteyn), 2:11-16.)

154. A person of ordinary skill would have had a reasonable expectation of success in segmenting the audio files I discuss above based on throughput (as taught by Ganesan) because such segmentation was widely known, as was choosing the size of those files based on network performance.

155. Accordingly, for at least the reasons I discuss above, (i) Young and Ganesan and/or (ii) Young, Yoshimura, Copley, and Ganesan render claim 18 obvious as a whole.

F. Claim 3 Would Have Been Obvious in View of Young, Kate, and Ganesan, or in View of Young, Yoshimura, Copley, Kate, and Ganesan.

156. Claim 3 depends from claim 2 and further recites “wherein a size of each digital audio file in the plurality of digital audio files is selected in dependence upon network throughput rates.” I note that this is the same additional limitation as recited in claim 18. (See ¶147, above.) As I discuss above, (i) Young and Kate and/or (ii) Young, Yoshimura, Copley, and Kate disclose or render obvious each limitation of claim 2. (See ¶¶127-135, above.) Ganesan discloses the additional claim element of claim 3 for the same reasons I discuss above regarding claim 18. (See ¶¶147-155, above.) Accordingly, for at least the reasons I discuss above, (i)

Young, Kate, and Ganesan and/or (ii) Young, Yoshimura, Copley, Kate, and Ganesan render claim 3 obvious as a whole.

G. Claim 4 Would Have Been Obvious in View of Young, Kate, Ganesan, and Sull or in View of Young, Yoshimura, Copley, Kate, Ganesan, and Sull.

157. Claim 4 depends from claim 3 and further recites “wherein the unique identifier is obtained from one of a bookmark structure, a card catalog structure, and an advertising structure.” I note that this is the same additional limitation as recited in claim 5. (See ¶137, above.) As I discuss above, (i) Young, Kate, and Ganesan and/or (ii) Young, Yoshimura, Copley, Kate, and Ganesan disclose or render obvious each limitation of claim 3. (See ¶156, above.) Sull discloses the additional claim element of claim 4 for the same reasons I discuss above regarding claim 5. (See ¶¶136-146, above.) Accordingly, for at least the reasons I discuss above, (i) Young, Kate, Ganesan, and Sull and/or (ii) Young, Yoshimura, Copley, Kate, Ganesan, and Sull render claim 4 obvious as a whole.

H. Claim 10 Would Have Been Obvious in View of Young and Shapiro, or in View of Young, Yoshimura, Copley, and Shapiro.

158. Claim 10 depends from claim 1 and further recites that “the request includes login information.” As I discuss above, Young, alone or in combination with Yoshimura and Copley, discloses or renders obvious each limitation of claim 1. (See ¶¶43-85, 92-109, above.) Shapiro discloses the additional claim element of claim 10.

159. It was well known to include login information such as a username and password in network requests before the '740 patent's earliest priority date. (EX-1092 (Maino), 3 (request including "username and [] password"); EX-1093 (Gupta), 5:34-6:51 (username and password in "cookies" attached to requests); EX-1094 (Tran), 9:15-20 (automatic login using "userid and password" in a request); EX-1042 (Shapiro), 2:1-33 ("a user name for the particular client is appended to the request's message header").) Accordingly, it would have been obvious to a person of ordinary skill in the art, in view of Young, Yoshimura, Copley, and/or their own knowledge, to include login information in the request.

160. It also would have been obvious to include login information in the request in view of Shapiro. Shapiro discloses including login information such as a username in a request. (EX-1042 (Shapiro), 2:11-13 ("a user name for the particular client is appended to the request's message header" to determine access), 4:38-56, 4:64-5:6 (request is denied if username is absent).)

161. Thus, Shapiro discloses the additional claim element of claim 10.

162. It would have been obvious to combine the teachings of Young and Shapiro, and to include login information such as a username in Young's request. A person of ordinary skill in the art would have been motivated to do so for several reasons.

163. First, the combination would allow Young's system to control access to the requested content.

164. Second, using login information when accessing files or information including media content stored on servers was well known. (EX-1092 (Maino), 3; EX-1093 (Gupta), 5:34-6:51; EX-1094 (Tran), 9:15-20; EX-1042 (Shapiro), 2:1-33.) Accordingly, this combination represents nothing more than using a known technique (login information) to improve a similar device and method (content delivery) in the same way (controlling access). Moreover, this combination applies a known technique (including login information in requests) to a known device and method (Young's system) that is ready for improvement and yields predictable results (the ability to control access).

165. A person of ordinary skill would have had a reasonable expectation of success in making the combination because including login information in a request was trivial, well-known, and explained by Shapiro.

166. Accordingly, for at least the reasons I discuss above, (i) Young and Shapiro and/or (ii) Young, Yoshimura, Copley, and Shapiro render claim 10 obvious as a whole.

I. Claim 11 Would Have Been Obvious in View of Young, Yoshimura, Copley, and Lindahl or in View of Young, Yoshimura, Copley, and Sull.

167. Claim 11 depends from claim 1. As I discuss above, Young, Yoshimura, and Copley disclose or render obvious each limitation of claim 1. (See ¶¶92-109, above.)

168. Element 11[a] recites “download[ing] a descriptor file from the first library server, the descriptor file for ordering the plurality of digital audio files, the descriptor file including at least one of a start time, an end time, and a play time of each digital audio file in the plurality of digital audio files within the audio stream.” As I discuss above, Yoshimura’s SMIL file renders this claim element obvious. (See ¶¶112-116, above.)

169. Element 11[b] recites “wherein the computer determines the first digital audio file for playback using a time offset external to the descriptor file and the at least one of the start time, end time, and play time of each digital audio file in the plurality of digital audio files.” Lindahl and Sull each render obvious claim element 11[b].

1. Lindahl

170. Lindahl discloses using a bookmark with an “offset relative to the beginning” of media content to “resume playback starting from [a] stopped point as desired.” (EX-1004 (Lindahl) ¶¶[0015], [0035], [0001].) This allows the “client

[to] initiate[] play ... from the point as determined by the stored bookmark[.]” (*Id.* ¶¶[0035]; *id.* ¶¶[0040].) When combined with Yoshimura’s multi-file system, Lindahl discloses determining the first digital audio file for playback (e.g., the file containing the bookmarked position) using a time offset (in the bookmark) external to the descriptor file (SMIL file) and the start/end/play times of the audio files (in the SMIL file).

171. It would have been obvious to a person of ordinary skill in the art to use Yoshimura’s time information as intended—to determine an audio file for download—by comparing the time offset of Lindahl’s bookmark to Yoshimura’s time information. For example, it would have been obvious to use a time offset in Lindahl’s bookmark (e.g., 15 minutes into the media stream), and then use Yoshimura’s SMIL file, as shown in Figure 5, to determine the audio file to play (e.g., content-A-2.mp4):

```
<!DOCTYPE smil PUBLIC "-//W3C//DTD SMIL 2.0//EN"  
"http://www.w3.org/2001/SMIL20/SMIL20.dtd">  
<smil xmlns="http://www.w3.org/2001/SMIL20/Language">  
  <head>  
    <layout>  
      <region id="v" top="5" left="5" width="180" height="180"/>  
    </layout>  
  </head>  
  <body>  
    <seq>  
      <par dur="10min">  
        <audio src="rtsp://content-server/content-A-1.mp4" />  
        <video src="rtsp://content-server/content-V-1.mp4" region="v" />  
      </par>  
      <par dur="20min">  
        <audio src="rtsp://content-server/content-A-2.mp4" />  
        <video src="rtsp://content-server/content-V-2.mp4" region="v" />  
      </par>  
      <par dur="30min">  
        <audio src="rtsp://content-server/content-A-3.mp4" />  
        <video src="rtsp://content-server/content-V-3.mp4" region="v" />  
      </par>  
    </seq>  
  </body>  
</smil>
```

Fig. 5 Modified SMIL file after content segmentation.

(EX-1006 (Yoshimura), Fig. 5.) Indeed, this is precisely how SMIL-file-based players worked.

172. A person of ordinary skill in the art would have been motivated to combine Yoshimura and Lindahl’s teachings, and to use a time offset in a bookmark (as in Lindahl) and the time information in the SMIL file to determine the audio file for playback, for several reasons.

173. First, that is how bookmarks were used with SMIL-file-based content.

174. Second, Young’s system provides the user the ability to download media content stored on servers, and Yoshimura discloses doing so via smaller files.

(See ¶¶63-72, above.) A person of ordinary skill would have been motivated to add Lindahl's bookmarking functionality to enable users to resume playback starting from a bookmarked point, which would require downloading the appropriate file.

175. Second, because bookmarks were well understood and described by Lindahl, the addition of bookmark functionality to the Young-Yoshimura system represents a simple substitution of one known element (an offset obtained from a bookmark) for another (a time offset obtained from a user or elsewhere) to obtain predictable results (identifying the media file corresponding to the bookmarked position). (EX-1004 (Lindahl) ¶¶[0015], [0035], [0040].) Similarly, this combination represents nothing more than using a known technique (obtaining time offset from bookmark) to improve a similar device and method (Young's system) in the same way. And, this combination represents nothing more than applying a known technique (obtaining time offset from bookmark) to a known device and method (Young's downloading media segments from content delivery system) that is ready for improvements and yields predictable results.

176. A person of ordinary skill would have had a reasonable expectation of success in using a time offset (as taught by Lindahl) in the Young-Yoshimura system at least because the use of such bookmarks was routine, trivial to implement, and widely used.

2. Sull

177. Sull’s audio bookmark provides “positional information” in the form of “elapsed time” (a time offset). (EX-1005 (Sull) ¶[0172].) The bookmark is external to the descriptor file (SMIL file) and enables determination of the relevant file so the content can be resumed at the bookmarked position. Thus, Sull discloses claim element 11[b].

178. A person of ordinary skill would have been motivated to implement Sull’s bookmark to enable the user to start playback at a particular position and would have reasonably expected success for the same reasons I discuss above for Lindahl. (See ¶¶172-176, above.)

179. Accordingly, for at least the reasons I discuss above, (i) Young, Yoshimura, Copley, and Lindahl and/or (ii) Young, Yoshimura, Copley, and Sull render claim 11 obvious as a whole.

J. Claims 1 and 6 Would Have Been Obvious in View of Leighton and Seed.

180. For at least the reasons I discuss below, Leighton and Seed render claims 1 and 6 obvious.

1. Claim 1

a. Preamble

181. The preamble recites “[a] non-transitory computer readable storage medium including computer readable code” that can be executed by a computer. Leighton discloses this.

182. Leighton’s content distribution system includes servers and a client machine. (EX-1032 (Leighton), 3:4-16, 5:14-22.) A person of ordinary skill would have understood that Leighton’s client machine has a computer readable storage medium including computer readable code (e.g., operating system and application code) executed by the client computer. Accordingly, Leighton discloses the preamble.

b. Element 1[a]: Sending a Request

183. Element 1[a] recites “send[ing] a request to a network-based server, the request including a unique identifier for identifying an audio stream.” Leighton discloses or renders obvious this claim element.

184. Leighton’s client machine “issues an HTTP request” to access a webpage that includes “embedded objects (e.g., images, audio, video, or the like).” (EX-1032 (Leighton), 3:24-28, 5:23-26; 12:53-58 (downloading audio and video content), claim 1.) The client fetches the webpage (an HTML document) and then “immediately fetch[es] the embedded objects” from servers. (*Id.*, 5:29-41; *see also*

id., 7:49-57 (HTML “page with embedded object URLs” provided in response to a request).)

185. The client fetches embedded objects by sending a request based on the object’s URL. (*See* EX-1032 (Leighton), 6:35-36 (“each embedded object that may be served in a page has its own URL”), 8:4-7, 9:20-30 (“After receiving the initial [HTML] page ... the browser needs to load the embedded URLs[.]”).)

186. Each request is sent to a DNS server that provides a list of provider (ghost) servers based on a URL, as explained below. (EX-1032 (Leighton), 9:20-23; *see also id.*, 9:48-10:41, 10:54-12:25.)

187. Leighton therefore discloses or renders obvious sending a request (e.g., for an embedded audio object) to a network-based server (e.g., DNS server), the request including a unique identifier (e.g., URL) for identifying an audio stream. Thus, Leighton discloses or renders obvious claim element 1[a].

c. Element 1[b]: Loading a List of Servers

188. Element 1[b] recites “load[ing] a list of library servers received from the network-based server, the list of library servers determined in dependence upon the unique identifier.” Leighton discloses or renders obvious this claim element.

189. When a client device sends a request for an embedded object to a DNS server, the object’s URL is interpreted using a “standard feature” of DNS servers whereby a DNS server provides to the client a list of ghost servers hosting the

content. (EX-1032 (Leighton), 3:17-37, 3:57-65 (“the top level DNS server returns a list of low-level DNS servers”), 12:13-16, 12:18-25 (“The low level DNS servers ... [return] a list of [server] names”), 5:54-57 (functionality of top-level and low-level DNS servers can be implemented on a single server), 10:36-41 (DNS server provides a list of servers).)

190. Leighton’s ghost servers are library servers because they store media content that clients request. (EX-1032 (Leighton), 3:5-7 (“The actual content to be served is preferably supported on a set of hosting servers (sometimes referred to as ghost servers).”), 2:59-62, 12:37-39 (“Once a copy has been retrieved it is returned to the user, and preferably it is also stored on the ghost for answering future requests.”), 6:26-30 (objects “served from the hosting servers”), 5:42-6:3 (describing servers), claim 1 (content servers).)

191. Leighton therefore discloses or renders obvious loading a list of library servers (e.g., list of ghost servers) from the network-based server (e.g., DNS server), the list of library servers determined in dependence upon the unique identifier (e.g., URL). Thus, Leighton discloses or renders obvious claim element 1[b].

d. Element 1[c]: Maintaining Server Statistics

192. Element 1[c] recites “maintain[ing] service level statistics for each library server in the list of library servers.” Leighton discloses or renders obvious this claim element.

193. Leighton discloses tracking service level statistics of ghost servers to update a “network map” that top-level DNS servers use to resolve content requests. (EX-1032 (Leighton), 10:1-11 (network map “determines where to direct the request” and “is updated continually based on network conditions and traffic”).) Leighton further discloses that low-level DNS servers also use a table to resolve requests, which include “virtual ghost names,” by directing them to “real ghost addresses using a table lookup, where the table is continually updated based on network conditions and traffic in such a way as to [e]nsure load balancing and fault tolerance.” (*Id.*, 11:7-11, 11:35-37 (“low-level DNS servers monitor the various ghost servers to take into account their loads”), 10:54-11:6, 11:23-26.) This functionality can be implemented on the client. (EX-1032 (Leighton), 5:54-57 (DNS functionality may be separated into a “hierarchy” and implemented on remote servers or the client device), 9:57-59, 12:53-67 (client software monitors the status of the network).)

194. Leighton therefore discloses or renders obvious maintaining service level statistics (e.g., network conditions, network traffic, and ghost server loads) for each library server (e.g., ghost server) in the list of library servers (e.g., list of ghost servers). Thus, Leighton discloses or renders obvious claim element 1[c].

e. Element 1[d][i]: Selecting a First Server

195. Element 1[d][i] recites “select[ing] a first library server from the list of library servers in dependence upon the service level statistics.” Leighton discloses or renders obvious this claim element.

196. Leighton discloses selecting a ghost server “close to the client machine, that is not overloaded, and that is most likely to already have ... the required file.” (EX-1032 (Leighton), 5:37-41.) Leighton selects a server from the list of ghost servers using maps and lookup tables that are based on service level statistics. (*See* ¶¶192-194, above; EX-1032 (Leighton), 5:54-57 (DNS functionality may be separated and implemented on server or client), 10:54-61, 11:35-37, 12:8-13.)

197. Leighton therefore discloses or renders obvious selecting a first library server (e.g., ghost server) from the list of library servers (e.g., list of ghost servers) in dependence upon the service level statistics (e.g., using mappings based on network conditions, traffic, and ghost server loads). Thus, Leighton discloses or renders obvious claim element 1[d][i].

f. Element 1[d][ii]: First Server Having a Plurality of Files

198. Element 1[d][ii] recites “the first library server having a plurality of digital audio files, each digital audio file in the plurality of digital audio files including a different segment of the audio stream.” Leighton and/or Seed disclose or render obvious this claim element.

199. Leighton discloses that the content servers have audio files. (EX-1032 (Leighton), 5:23-26, 12:55-58.) I note that Leighton does not expressly disclose that each audio file includes a different segment of an audio stream. However, dividing audio streams into multiple files containing segments of the stream was a conventional approach to storing audio on servers. (EX-1001 ('740 patent), 2:16-18; *see* ¶¶35-36, above.) Accordingly, it would have been obvious, and a routine design choice, to store Leighton's audio files as smaller files representing segments of an audio stream.

200. This limitation also would have been obvious in view of Seed. Seed discloses a CDN that delivers objects, including "an audio file," from servers to end-users (clients). (EX-1011 (Seed) ¶[0010].) Seed discloses that "objects can be segmented into chunks" to "enhance the performance of the network." (*Id.* ¶[0057].) "By partitioning streams in this manner, a first part of an object can be served from edge servers quickly." (*Id.*) The remaining chunks can also be stored on edge servers. (*Id.* ¶¶[0058] ("some or all of the objects may be permanently retained in edge server storage"), [0032].) Because Seed "partition[s] streams" into "chunks" that can be stored and downloaded individually, it would have been obvious to a person of ordinary skill in the art that the chunks could be stored as conventional files containing a segment of an audio stream. (*Id.* ¶¶[0057], [0010], [0054]; EX-1030 (Heckerman), 13:14-34.)

201. Thus, Leighton and Seed render obvious a first library server having a plurality of digital audio files (e.g., Seed’s chunks stored in a conventional file format), each digital audio file in the plurality of digital audio files including a different segment of the audio stream.

202. A person of ordinary skill in the art would have been motivated to combine the teachings of Leighton and Seed, and to store Leighton’s audio objects as a plurality of smaller files, for several reasons.

203. First, Leighton’s goal is to provide media content to clients “efficiently, effectively, and reliably,” and to “speed-up the delivery” of content. (EX-1032 (Leighton), 2:26-31; 2:43-45.) Seed teaches that segmentation “enhance[s] the performance of the network” and avoids providing entire objects unnecessarily (e.g., when users exit before an object is completely served). (EX-1011 (Seed) ¶[0057].) A person of ordinary skill therefore would have been motivated to segment the audio files served in Leighton to obtain the benefits Seed describes.

204. Second, because separating media content into smaller files (e.g., chunks or segments) was widely known, and further given the similarities between Leighton’s and Seed’s CDN systems, storing a media stream in Leighton as several files would represent a simple substitution of one known element (audio stream stored in multiple smaller files) for another (audio stream stored as a single large file) to obtain predictable results (improved performance and ability to serve smaller

files). (EX-1032 (Leighton), 2:26-67; EX-1011 (Seed) ¶¶[0057]; *see* ¶¶[35-36, above.]) This combination also represents nothing more than using a known technique (Seed’s chunking) to improve a similar device and method (Leightons CDN and storing content on ghost servers) in the same way. Moreover, this combination applies a known technique (file division) to a known device and method (Leightons CDN and audio content) that is ready for improvement and yields predictable results (enhancing network performance by serving smaller individual audio files).

205. Third, it would have been obvious to try storing Leighton’s audio files as smaller files (as taught by Seed) because it was one of a finite number of identified, predictable solutions for storing content on servers. Only two options existed for storing Leighton’s audio objects: (1) store it as a single larger file; or (2) store it as a plurality of smaller files. Both would have been obvious.

206. A person of ordinary skill would have reasonably expected success in storing Leighton’s audio content in smaller files (as in Seed) because the systems and their purposes are similar, storing content as multiple smaller files was known, and the modification would have been trivial. (*See* ¶¶[35-36, above.])

207. Thus, Leighton and/or Seed disclose or render obvious claim element 1[d][ii].

g. Element 1[e]: Downloading a First Audio File

208. Element 1[e] recites “download[ing] a first digital audio file from the plurality of digital audio files for playback with a media player.” Leighton and/or Seed render obvious this claim element.

209. Leighton discloses downloading content, including audio files, to be rendered on a client device. (EX-1032 (Leighton), 2:64-67 (“content is automatically sent to the location where it is requested”), 4:17-22, 5:23-32 (objects include audio), 12:26-56, 12:59-62 (“software [] is placed in the client’s browser or media player”).) Because it would have been obvious to store Leighton’s audio objects as a plurality of files, this limitation would have been obvious over Leighton.

210. This limitation also would have been obvious in view of Seed. Seed discloses downloading digital audio files (“chunks”) that are segments of an audio stream. (See ¶¶198-207, above.) These files are downloaded for playback by the client, including through a media player. (EX-1011 (Seed) ¶[0051] (software for object delivery includes software to “process[] requests from a Windows Media Player ... for Windows Media objects”).)

211. Thus, Leighton and/or Seed render obvious claim element 1[e].

h. Element 1[f][i]: Selecting a Second Server

212. Element 1[f][i] recites “select[ing] a second library server from the list of library servers in dependence upon the service level statistics.” Leighton discloses or renders obvious this claim element.

213. Leighton discloses selecting a second library (ghost) server from the list when “it is deemed that the client’s connection can be improved by changing the server.” (EX-1032 (Leighton), 12:59-67; *id.* (“user can be directed to an alternate server in mid-stream”); *id.*, 12:53-55 (performance “improved by dynamically changing the server to which a client is connected based on changing network conditions”).) Leighton also discloses selecting a second server (e.g., “buddy” server) “if the ghost goes down.” (*Id.*, 13:1-4, 12:18-25 (if buddy system fails, client contacts “one of the other ghosts on the list”).)

214. Leighton therefore discloses and renders obvious selecting a second library server from the list of servers in dependence upon the service level statistics. Thus, Leighton discloses or renders obvious claim element 1[f][i].

i. Element 1[f][ii]: Second Server Having a Copy of the Files

215. Element 1[f][ii] recites “the second library server having a copy of the plurality of digital audio files.” Leighton discloses or renders obvious this claim element.

216. Leighton discloses replicating content across multiple ghost servers. (EX-1032 (Leighton), 3:42-49 (“content can be distributed and replicated through a collection of servers”), 12:37-39.) Moreover, the second library server would have a copy of the media content stored on the first library server because it serves as a backup to provide the same content.

217. Accordingly, Leighton discloses or renders obvious that the selected second server has a copy of the plurality of digital audio files. Thus, Leighton discloses or renders obvious claim element 1[f][ii].

j. Element 1[g]: Downloading a Second Audio File from the Second Server

218. Element 1[g] recites “download[ing] a second other digital audio file from the second library server for playback with the media player.” Leighton and Seed render obvious this claim element.

219. Leighton discloses that the user can be “directed to an alternate server in mid-stream.” (EX-1032 (Leighton), 12:53-67; *id.*, 12:53-55 (change server), 13:1-4 (second “buddy” server).) As I discuss above, it would be obvious for the stream to comprise multiple audio files. (See ¶¶192-194, 215-217, above.) Thus, downloading the remainder of the stream from a second (“alternate”) server would include downloading the second digital audio file as recited in element 1[g].

220. Accordingly, Leighton and Seed render obvious downloading a second other digital audio file (in the second portion of the media stream) from the second

library server (alternate server) for playback with the media player. Thus, Leighton and Seed render obvious claim element 1[g].

221. Accordingly, for at least the reasons I discuss above, Leighton and Seed render claim 1 obvious as a whole.

2. Claim 6

222. Claim 6 depends from claim 1 and further recites “wherein the service level statistics include historical transfer rates for each library server in the list of library servers.” Leighton discloses or renders obvious the additional claim element of claim 6.

223. A person of ordinary skill in the art would have understood that Leighton’s service level statistics include historical transfer rates because they provide “network conditions and traffic” for each server (*see* ¶¶192-194, above), and transfer rates (e.g., speed in bits per second) were a conventional measurement of network and server conditions. (*See* ¶¶86-90, above.)

224. Alternatively, it would have been obvious to include historical transfer rates in Leighton’s service level statistics for the reasons I discuss above. (*See* ¶¶86-90, above.) A person of ordinary skill would have been motivated to include server historical transfer rates in Leighton’s statistics because Leighton sought “fast and reliable [] access” to content and servers’ historical transfer rates would indicate their speed and reliability. (EX-1032 (Leighton), 13:55-58.) And, because measuring

transfer rates was well known, this would represent nothing more than the substitution of one known element (historical transfer rates) for another (Leighton's statistics) to achieve predictable results (measuring server speed). (*See* ¶¶86-90, above.) A person of ordinary skill would have had a reasonable expectation of success in maintaining historical transfer rates and using them in Leighton's system because maintaining and using such statistics for this purpose was well known. (*Id.*)

225. Thus, Leighton discloses or renders obvious the additional claim element of claim 6.

226. Accordingly, for at least the reasons I discuss above, Leighton and Seed render claim 6 obvious as a whole.

K. Claims 1, 2, 6, and 10 Would Have Been Obvious in View of Leighton, Seed, and Lindahl.

227. As I discuss above, Leighton and Seed disclose or render obvious each limitation of claim 1. (*See* ¶¶181-221, above.) For at least those reasons and the reasons I discuss below, claims 1, 2, 6, and 10 would have been obvious in further view of Lindahl. In particular, Lindahl discloses or renders obvious claim elements that require a plurality of digital audio files representing segments of the audio stream (e.g., elements 1[d][ii] and 1[f][ii]). I incorporate my above discussion of Leighton and Seed here, and only discuss claim elements for which Lindahl is relevant in combination with Lindahl and Seed.

1. Claim 1

228. Lindahl discloses a client-server system for delivering audiobooks. (EX-1004 (Lindahl) ¶[0001], Fig. 1.) Lindahl explains that audiobooks “typically contain a large amount of data -for example it would not be unusual for an audio book to include data for nine hours of listening.” (*Id.* ¶[0007].) This makes it “difficult to download it to small devices with a small storage space” and the “download time is too long to provide a practical implementation for audio books, especially over wireless narrowband links.” (*Id.*)

229. To overcome these problems, Lindahl discloses that each audiobook datafile is “sub-dividable,” which enables the “download of smaller portions of the larger datafile.” (EX-1004 (Lindahl) ¶[0010]; *id.* ¶[0018] (individual data files segmented into multiple “sub-portions”), claims 1 (original file “may be subdivided into two or more transmission portions, each portion being independently transmittable to the client device”), 3 (audio file).) Lindahl expressly discloses that the smaller portions may be individual files. (*Id.* ¶¶[0022] (each chapter may be its own audio file), [0018] (segmented sub-portions are “individually searchable and retrievable”).) This permits the downloading of audio books “on a chapter per chapter basis ... so that the system may be used in conjunction with small client devices.” (*Id.* ¶¶[0023], [0033] (download on chapter-by-chapter basis), [0045] (system enables retrieval of “individual sub-portions”).)

230. Thus, Lindahl discloses or renders obvious a server having a plurality of digital audio files (e.g., all chapters for an audiobook), each digital audio file (e.g., chapter file) in the plurality of digital audio files including a different segment (e.g., chapter) of the audio stream.

231. A person of ordinary skill in the art would have been motivated to combine the teachings of Leighton, Seed, and Lindahl, and to store Leighton’s audio streams (and Seed’s “chunks”) as a plurality of smaller files as taught by Lindahl, for several reasons.

232. First, Lindahl suggests doing so because Lindahl teaches that storing audio books on servers in a plurality of smaller files: (a) makes it easier to download large audio files (e.g., audiobooks) to small devices with small storage space (EX-1004 (Lindahl) ¶¶[0022]); (b) reduces download times, especially over wireless narrowband links (*id.*); and (c) makes smaller portions of the audio file searchable and downloadable (*id.* ¶¶[0018]). A person of ordinary skill would have been motivated to obtain these benefits in Leighton’s system.

233. Second, Leighton sought to “speed-up the delivery” of content and to provide content “reliably and economically.” (EX-1032 (Leighton), 2:43-48, 2:26-31.) A person of ordinary skill would have understood that segmenting large audio files into smaller files for storage and transmission, as taught by Lindahl, would further these goals.

234. Third, because segmenting media content into smaller files was widely known and expressly disclosed by Lindahl, storing Leighton's audio content as several smaller files represents nothing more than a simple substitution of one known element (audio stream stored as multiple smaller files) for another (audio stream stored as one large file) to obtain predictable results (efficient storage and transmission of content across various networks and to clients with less storage). (See ¶¶35-36 (widely known to segment files); EX-1001 ('740 patent), 2:16-19.) This combination also represents using a known technique (storing and sending content as smaller files) to improve a similar device and method (Leighton's content delivery) in the same way (e.g., faster downloads, access by smaller devices, and individually searchable and retrievable portions of the content). Moreover, this combination applies a known technique (storing and delivering content as smaller files) to a known device and method (Leighton's content delivery) that is ready for improvement and yields predictable results (increased performance and flexibility, especially in low bandwidth transmission).

235. Fourth, storing and delivering Leighton's audio files as several smaller files would have been obvious to try because it is one of a finite number of identified, predictable solutions (storing the content as one large file or several smaller files) for storing content on servers.

236. For these same reasons, a person of ordinary skill would have reasonably expected success in implementing this modification to Leighton’s system.

237. Accordingly, for at least the reasons I discuss above, Leighton, Seed, and Lindahl render claim 1 obvious as a whole.

2. Claim 2

238. Claim 2 depends from claim 1 and further recites “wherein the unique identifier is an ISBN number.” Lindahl discloses the additional claim element of claim 2.

239. Lindahl discloses that a request for media content from a server can use an ISBN as a unique identifier of requested content. (EX-1004 (Lindahl) ¶¶[0039] (user may “request an audio book by providing the server with a suitable unique identifier associated with the chosen book”), [0030] (ISBN is a “unique identifier associated with a stored audio book”), [0017] (media content stored with ISBN), [0035] (“part or the entire book is transmitted to the client device”).)

240. Accordingly, Lindahl discloses sending a request to a server where the request includes an ISBN. Thus, Lindahl discloses the additional claim element of claim 2.

241. A person of ordinary skill in the art would have been motivated to combine the teachings of Leighton and Lindahl, and to use an ISBN as a unique identifier in the client request, for several reasons.

242. First, Leighton suggests the combination because it discloses that the audio stream's unique identifier (e.g., modified URL) may contain a "serial number." (EX-1032 (Leighton), 8:2-15.) Lindahl's ISBN is a serial number for an audiobook, and Lindahl teaches that the ISBN can be used to retrieve the content. (EX-1004 (Lindahl) ¶¶[0039], [0030], [0017].)

243. Second, Leighton discloses requesting content using an HTTP request and Lindahl discloses that such a request can include an ISBN to identify the requested content. (EX-1032 (Leighton), 1:15-34 (well known to use HTTP requests to request content for a web page), 8:3-12 (URLs identifying content using HTTP); EX-1004 (Lindahl) ¶¶[0024] (selected download is accomplished using the HTTP protocol), [0030] (request can identify content by ISBN).)

244. Third, because ISBNs were disclosed by Lindahl and well-understood, the combination of Leighton and Lindahl represents a simple substitution of one known element (a request for content that includes an ISBN) for another (a request for content as disclosed in Leighton) to obtain predictable results (accessing content identified by ISBN). (EX-1004 (Lindahl) ¶[0030].) Similarly, the combination of Leighton and Lindahl represents the same simple substitution of a known element (standardized identifier) for another (Leighton's identifier).

245. Fourth, the combination applies a known technique (identifying content using an ISBN) to a known device and method (Leighton’s content distribution system) that is ready for improvement and yields predictable results.

246. For these same reasons, a person of ordinary skill would have reasonably expected success in incorporating an ISBN (as taught by Lindahl) into Leighton’s modified URL.

247. Accordingly, for at least the reasons I discuss above, Leighton, Seed, and Lindahl render claim 2 obvious as a whole.

3. Claim 6

248. Claim 6 depends from claim 1 and further recites “wherein the service level statistics include historical transfer rates for each library server in the list of library servers.” Because Leighton, Seed, and Lindahl disclose or render obvious each element of claim 1 (*see* ¶¶181-221, 228-237, above) and because Leighton discloses the additional claim element of claim 6 (*see* ¶¶222-226, above), Leighton, Seed, and Lindahl render claim 6 obvious as a whole.

4. Claim 10

249. Claim 10 depends from claim 1 and further recites that “the request includes login information.” Lindahl discloses the additional claim element of claim 10.

250. Lindahl discloses sending “authenticators” such as the device IMEI number or the SIM card MSISDN to the server so that only registered users can access the audiobook content. (EX-1004 (Lindahl) ¶¶[0020], [0028]-[0029], [0039] (server receives authentication request).) A person of ordinary skill in the art would have been motivated to implement this feature in Leighton’s CDN to provide access control, and would have reasonably expected success in doing so given the similarity of the systems and that such access control was widely known.

251. Accordingly, for at least the reasons I discuss above, Leighton, Seed, and Lindahl render claim 10 obvious as a whole.

L. Claim 2 Would Have Been Obvious in View of Leighton, Seed, and Kate.

252. Claim 2 depends from claim 1 and further recites “wherein the unique identifier is an ISBN number.” As I discuss above, Leighton and Seed disclose or render obvious each limitation of claim 1. (See ¶¶181-221, above.) As I also discuss above, Kate discloses the additional claim element of claim 2. (See ¶¶127-135, above.)

253. A person of ordinary skill in the art would have been motivated to combine the teachings of Leighton and Kate to use an ISBN as a unique identifier (as taught by Kate) for the same reasons that a person of ordinary skill would have been motivated to modify Young to use an ISBN, as I discuss above. (See ¶¶127-135, above.) For example, Leighton discloses requesting content using a modified URL

and Kate discloses using a unique identifier with an ISBN for the same purpose, the combination therefore represents a simple substitution of one known element (identifying content using an ISBN) for another (modified URL) to obtain predictable results (accessing content identified by the identifier that includes the ISBN), it represents the same simple substitution of a known element for another to yield predictable results, and it represents using a known technique (identifying content via ISBN) to improve a similar device and method (identifying content via Leighton's modified URL) in the same way.

254. A person of ordinary skill would have reasonably expected success because ISBNs were well understood as identifiers of content and because Kate describes using the ISBN to identify content in a system similar to Leighton's.

255. Accordingly, for at least the reasons I discuss above, Leighton, Seed, and Kate render claim 2 obvious as a whole.

M. Claim 5 Would Have Been Obvious in View of Leighton, Seed, and Sull or in View of Leighton, Seed, Lindahl, and Sull.

256. Claim 5 depends from claim 1 and further recites "wherein the unique identifier is obtained from one of a bookmark structure, a card catalog structure, and an advertising structure." As I discuss above, Leighton and Seed disclose or render obvious each limitation of claim 1. (See ¶¶181-221, above.) As I also discuss above, Leighton, Seed, and Lindahl disclose or render obvious each limitation of claim 1.

(See ¶¶228-237, above.) As I further discuss above, Sull discloses the additional claim element of claim 5. (See ¶¶136-146, above.)

257. A person of ordinary skill in the art would have been motivated to combine the teachings of Leighton and Sull to obtain Leighton's unique identifier (e.g., modified URL) from a bookmark for several reasons.

258. First, Leighton discloses accessing and rendering media content (e.g., audio files). As I discuss above, Sull discloses using bookmarks to identify media content and immediately access it at a desired position rather than restarting at the beginning, thereby providing an improved user experience. (See ¶¶136-146, above.)

259. Second, because bookmarks were well known and taught by Sull, the combination represents nothing more than using a known technique (accessing media content using a bookmark) to improve a similar device and method (accessing media content as in Leighton) to obtain predictable results (accessing a specific position within the requested media content using a bookmark). Similarly, this combination simply applies a known technique (bookmarking) to a known device and method (obtaining media content in Leighton's system) that is ready for improvement and yields predictable results (playing requested content from a specific position).

260. A person of ordinary skill would have had a reasonable expectation of success because bookmarks were well understood, and because both Leighton and

Sull disclose using similar unique identifiers (URIs and URLs) to identify media content. (*See* ¶¶183-187, 136-146, above.)

261. Accordingly, for at least the reasons I discuss above, (i) Leighton, Seed, and Sull and/or (ii) Leighton, Seed, Lindahl, and Sull render claim 5 obvious as a whole.

N. Claim 18 Would Have Been Obvious in View of Leighton, Seed, and Ganesan or in View of Leighton, Seed, Lindahl, and Ganesan.

262. Claim 18 depends from claim 1 and further recites “wherein a size of each digital audio file in the plurality of digital audio files is selected in dependence upon network throughput rates.” As I discuss above, Leighton and Seed disclose or render obvious each limitation of claim 1. (*See* ¶¶181-221, above.) As I also discuss above, Leighton, Seed, and Lindahl disclose or render obvious each limitation of claim 1. (*See* ¶¶228-237, above.) As I further discuss above, Ganesan discloses the additional claim element of claim 18. (*See* ¶¶147-155, above.)

263. A person of ordinary skill in the art would have been motivated to combine the teachings of Leighton and Ganesan to determine the size of media content segments based on throughput in Leighton’s system, and would have had a reasonable expectation of success in doing so, for the same reasons that such a person would have been motivated to do so for Young, as I discuss above. (*See* ¶¶147-155, above.)

264. Accordingly, for at least the reasons I discuss above, (i) Leighton, Seed, and Ganesan and/or (ii) Leighton, Seed, Lindahl, and Ganesan render claim 18 obvious as a whole.

O. Claim 3 Would Have Been Obvious in View of Leighton, Seed, Lindahl, and Ganesan, or in View of Leighton, Seed, Kate, and Ganesan.

265. Claim 3 depends from claim 2 and further recites “wherein a size of each digital audio file in the plurality of digital audio files is selected in dependence upon network throughput rates.” I note that this is the same additional limitation as recited in claim 18. (*See* ¶¶262, above.) As I discuss above, (i) Leighton, Seed, and Lindahl and/or (ii) Leighton, Seed, and Kate disclose or render obvious each limitation of claim 2. (*See* ¶¶238-247, 252-255, above.) Ganesan discloses the additional claim element of claim 3 for the same reasons I discuss above regarding claim 18. (*See* ¶¶147-155, above.) Accordingly, for at least the reasons I discuss above, (i) Leighton, Seed, Lindahl, and Ganesan and/or (ii) Leighton, Seed, Kate, and Ganesan render claim 3 obvious as a whole.

P. Claim 4 Would Have Been Obvious in View of Leighton, Seed, Lindahl, Ganesan, and Sull or in View of Leighton, Seed, Kate, Ganesan, and Sull.

266. Claim 4 depends from claim 3 and further recites “wherein the unique identifier is obtained from one of a bookmark structure, a card catalog structure, and an advertising structure.” I note that this is the same additional limitation as recited

in claim 5. (See ¶256, above.) As I discuss above, (i) Leighton, Seed, Lindahl, and Ganesan and/or (ii) Leighton, Seed, Kate, and Ganesan disclose or render obvious each limitation of claim 3. (See ¶265, above.) Sull discloses the additional claim element of claim 4 for the same reasons I discuss above regarding claim 5. (See ¶¶256-261, above.) Accordingly, for at least the reasons I discuss above, (i) Leighton, Seed, Lindahl, Ganesan, and Sull and/or (ii) Leighton, Seed, Kate, Ganesan, and Sull render claim 4 obvious as a whole.

Q. Claim 10 Would Have Been Obvious in View of Leighton, Seed, and Shapiro or in View of Leighton, Seed, Lindahl, and Shapiro.

267. Claim 10 depends from claim 1 and further recites that “the request includes login information.” As I discuss above, Leighton and Seed disclose or render obvious each limitation of claim 1. (See ¶¶181-221, above.) As I also discuss above, Leighton, Seed, and Lindahl disclose or render obvious each limitation of claim 1. (See ¶¶228-237, above.) As I further discuss above, Shapiro discloses the additional claim element of claim 10. (See ¶¶158-166, above.)

268. A person of ordinary skill would have been motivated to include login information such as a username in the user’s request in Leighton’s system for several reasons.

269. First, Leighton provides media content. Including login information in the request, as taught by Shapiro, would allow Leighton’s system to control access to the requested content.

270. Second, using login information when accessing files or information including media content stored on servers was well known, as I discuss above. (*See* ¶¶158-166, above.) Accordingly, the combination represents a simple substitution of one known element (request including login information) for another (network request without login information) to obtain predictable results (controlling content access based on login information). Similarly, this combination represents using a known technique (login information) to improve a similar device and method (content delivery) in the same way. It also applies a known technique (including login information in requests) to a known device and method (Leighton’s system) that is ready for improvement and yields predictable results (the ability to control access).

271. A person of ordinary skill would have had a reasonable expectation of success in making the combination because including login information in a request was trivial, well-known, and explained by Shapiro and Lindahl.

272. Accordingly, for at least the reasons I discuss above, (i) Leighton, Seed, and Shapiro and/or (ii) Leighton, Seed, Lindahl, and Shapiro render claim 10 obvious as a whole.

R. Claim 11 Would Have Been Obvious in View of Leighton, Seed, Lindahl, and Yoshimura.

273. Claim 11 depends from claim 1. Element 11[a] recites “download[ing] a descriptor file from the first library server, the descriptor file for ordering the

plurality of digital audio files, the descriptor file including at least one of a start time, an end time, and a play time of each digital audio file in the plurality of digital audio files within the audio stream.” Element 11[b] recites “wherein the computer determines the first digital audio file for playback using a time offset external to the descriptor file and the at least one of the start time, end time, and play time of each digital audio file in the plurality of digital audio files.”

274. As I discuss above, Leighton, Seed, and Lindahl disclose or render obvious each limitation of claim 1. (*See* ¶¶228-238, above.) As I also discuss above, Yoshimura and Lindahl together render obvious the additional claim elements of claim 11. (*See* ¶¶112-126, above.)

1. Motivation to Use Yoshimura’s SMIL File

275. A person of ordinary skill in the art would have been motivated to combine the teachings of Yoshimura with Leighton, Seed, and Lindahl for several reasons.

276. First, because Leighton (alone or with Seed and Lindahl) discloses that the audio content can be provided as multiple files, a person of ordinary skill in the art would have been motivated to look to methods for accessing, navigating, and rendering such multi-file streams. Yoshimura discloses using a SMIL file for this purpose, including providing conventional controls like skipping forwards and backwards. (*See* ¶¶117-126, above.)

277. Second, the use of SMIL files to access media content on the internet was widely known and supported. (EX-1006 (Yoshimura), 1779; EX-1007 (Bulterman), 5 (showing SMIL media content accessed in web browsers), 18 (“SMIL 2.0’s release makes full-fledged multimedia accessible ... to the infrastructure of the Web itself” and “SMIL 2.0 is widely supported and distributed”), 63; EX-1090 (Kennedy), 1-2, 8-16.) Downloading and using a SMIL file as taught in Yoshimura would have therefore represented nothing more than a simple substitution of one known element (SMIL file identifying media content) for another (HTML file identifying media content). This combination represents the use of a known technique (accessing and synchronizing media content using a SMIL file) to improve a similar device and method (accessing content using an HTML file) in the same way. Moreover, this combination applies a known technique (using a SMIL file to access and synchronize media) to a known device and method (accessing content as in Leighton) that is ready for improvements and yields predictable results.

278. A person of ordinary skill would have reasonably expected success in using a SMIL file in Leighton’s system because SMIL files were a conventional file type for accessing media content via the internet. (EX-1006 (Yoshimura), 1779; EX-1007 (Bulterman), 5, 63; EX-1090 (Kennedy), 1-2, 8-16.) Moreover, because HTML and SMIL are similar markup languages that identify media content on the internet, a person of ordinary skill would have reasonably expected success in using

a SMIL file to access media content in addition to or in place of the HTML file of Leighton. (EX-1090 (Kennedy), 1 (“SMIL versions 1.0 and 2.0 allow Web authors to weave together multimedia by writing XML-based markup similar to HTML”), 7-8, 19, 50-51, 54.) Indeed, the widely available Internet Explorer web browser allowed for the use of HTML and SMIL together. (*E.g.*, EX-1007 (Bulterman), 18 (“XHTML+SMIL is supported in Internet Explorer, the world’s most widely distributed network browser.”).)

2. Motivation to Use Lindahl’s Bookmark

279. A person of ordinary skill in the art also would have been motivated to determine the audio file for playback using a time offset from Lindahl’s bookmark for several reasons.

280. First, Leighton discloses or renders obvious downloading segments of media content. (EX-1008 (Leighton), 3:66-4:4, 1:59-63, 4:39-57, 4:58-5:9, 5:19-21.) Lindahl discloses using a time offset from a bookmark to identify the segment a client should download and the position at which it should begin playback. (EX-1004 (Lindahl) ¶¶[0015], [0035], [0040].) A person of ordinary skill would have recognized that Lindahl’s bookmark provides desirable functionality—quickly accessing a desired position in media content such as audiobooks. (*Id.* ¶[0035]; EX-1005 (Sull) ¶[0005] (“Most users would want to restart accessing the content from the point where they had left off.”).)

281. Second, because bookmarks were well understood, the combination of Leighton and Lindahl's bookmarks represents nothing more than the addition of one known element (a bookmark containing a time offset) to another (Leighton's system) to obtain predictable results (playback from a marked position). Similarly, this combination represents using a known technique (obtaining a time offset from a bookmark as in Lindahl) to improve a similar device and method (Leighton's system and determining a file for download). And, this combination represents nothing more than applying a known technique (obtaining a time offset from a bookmark as in Lindahl) to a known device and method (Leighton's content delivery system) that is ready for improvements and yields predictable results (the ability to identify the segment and position for resuming playback).

282. A person of ordinary skill would have had a reasonable expectation of success when combining Leighton and Lindahl because the use of multimedia bookmarks was well known.

283. Accordingly, for at least the reasons I discuss above, Leighton, Seed, Lindahl, and Yoshimura render claim 11 obvious as a whole.

VII. SECONDARY CONSIDERATIONS OF NONOBVIOUSNESS

284. I am not aware of any secondary considerations of nonobviousness. If Patent Owner identifies any alleged evidence of secondary considerations in the future, I reserve the right to respond to that information.

VIII. CONCLUSION

285. For the foregoing reasons, it is my opinion that claims 1-6, 10-11, and 18 of the '740 patent would have been obvious to a person of ordinary skill in the art at the time of the alleged invention in view of the prior art discussed above.

286. I reserve the right to supplement my opinions in the future to address or respond to any issues that the Patent Owner may raise, as well as new information including, but not limited to, any claim constructions advanced by the Patent Owner or adopted by the Board in the Institution Decision, and respond to any alleged secondary considerations as they become available to me.

Amazon.com, Inc. v. Audio Pod IP, LLC
Declaration of Ketan Mayer-Patel – U.S. Pat. No. 8,738,740

I declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Executed on March 24th, 2025

at Chapel Hill, NC.

A handwritten signature in black ink, appearing to read 'KMP', written over a horizontal line.

Professor Ketan Mayer-Patel, Ph.D.