UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

DISNEY ENTERTAINMENT & SPORTS LLC, Petitioner

v.

ADEIA GUIDES INC.,
Patent Owner.

Case No. IPR2026-00052
U.S. Patent No. 8,542,705

DECLARATION OF DR. BRIAN W. ANTHONY IN SUPPORT OF PETITION FOR *INTER PARTES* REVIEW OF U.S. PATENT NO. 8,542,705

TABLE OF CONTENTS

Page

TABLE OF CONTENTS
(continued)

Page

TABLE OF CONTENTS
(continued)

Page

TABLE OF CONTENTS
(continued)

Page

TABLE OF CONTENTS
(continued)

Page

TABLE OF CONTENTS
(continued)

Page

TABLE OF CONTENTS
(continued)

Page

TABLE OF CONTENTS
(continued)

Page

TABLE OF CONTENTS
(continued)

Page

TABLE OF CONTENTS
(continued)

Page

I. INTRODUCTION

1. I have been retained by Disney Entertainment & Sports LLC ("Petitioner"), as an independent expert in this proceeding before the Patent Trial and Appeal Board ("PTAB" or "Board"). I understand that Petitioner is requesting that the Board institute an inter partes review ("IPR") proceeding of U.S. Patent No. 8,542,705 ("the '705 Patent") (Ex 1001), currently assigned to Adeia Guides Inc. ("Patent Owner" or "PO").

2. I am not and have never been an employee of Petitioner. I am being compensated at my usual and customary rate of $500 per hour. No part of my compensation depends on the outcome of this proceeding, and I have no other interest in this proceeding.

3. I have been asked to provide my independent analysis of the '705 Patent in light of the prior art publications cited below. I have also been asked to consider the state of the art and prior art available as of January 23, 2007. Based on the prior art discussed in this declaration, it is my opinion that Claims 1, 2, 4-9, 11-16, and 18-20 of the '705 Patent are unpatentable for the reasons provided below.

A. Qualifications

4. As indicated in my *curriculum vitae* ("CV"), included as Ex-1003, I am currently a Principal Research Scientist at the Massachusetts Institute of Technology ("MIT"). My CV includes additional information about my professional history and

contains further details on my experience, publications, patents, and other qualifications to render an expert opinion. Herein, I highlight experiences relevant to the technology of the patent at issue.

5. I earned a Bachelor of Science in Engineering from Carnegie Mellon University in 1994 and a Master's degree in Engineering from MIT in 1998. My thesis topic related to anisotropic wave guides, image processing, timeseries analysis, applied to non-destructive testing. In 2006, I earned my Ph.D. in Engineering from MIT researching in the AI Lab. My research focused on high-performance computation, signal processing, and video-instrumentation system design, video search, human-interfaces for algorithm interaction, and video compression.

6. In 1997, I co-founded Xcitex Inc., a company that specialized in video-acquisition and motion-analysis software. As CTO and first engineer at Xcitex, I designed and built video analytics and streaming products for scientific markets, including GUIs, compression, and key frame technologies and led team to do the same. I served as the Chief Technology Officer and directed and managed product development until 2006. I served as the technical and scientific advisor to eXstream Solutions, an early video streaming company. My work there focused on digital media application development, multimedia content and rights management, hardware and network architecture, and digital design and consulting.

7. I joined MIT in 2006 and have variously served as the Director of the Master of Engineering in Advance Manufacturing and Design Program, Co-Director of the Medical Electronics Device Realization Center. Many of the courses, projects, and papers my students undertake involve technologies relevant to the '705 Patent, for example video and image processing applied to manufacturing system and medical imaging systems.

8. Furthermore, my research focuses on imaging, video processing, human interface systems. I teach or have taught courses in Electrical Engineering, Controls, Optics, Communications and Network Protocols, Signal Processing, Video Processing, Imaging, Smart Systems Design, all pertinent subject matter to the '705 Patent.

9. For decades I have applied principles of video analytics, system design, and compression to domains such as manufacturing imaging, motion capture, and medical imaging. I have extensive experience in video processing, video analytics, GUIs, compression, and key frame detection, and have published widely on video processing and transmission across multiple fields. I was awarded an Emmy Award from the National Academy of Television Arts & Sciences for technical innovation in Broadcast television related to video processing and analysis. I have published approximately 100 papers and have received a number of best paper and

distinguished paper awards. I am a co-author of a number of papers that relate to the technology in the ’705 Patent, such as:

- Anthony, Brian W., and Fitriani Chua. “Computationally efficient optimal video comparison for machine monitoring and process control.” Journal of Manufacturing Science and Engineering 139, no. 10 (2017): 101007.
- Sun, S.Y., Gilbertson, M. and Anthony, B.W., 2014, September. Probe localization for freehand 3D ultrasound by tracking skin features. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 365-372). Cham: Springer International Publishing.
- Magana-Salgado, Uriel, Praneeth Namburi, Micha Feigin-Almon, Roger Pallares-Lopez, and Brian Anthony. “A comparison of point-tracking algorithms in ultrasound videos from the upper limb.” BioMedical Engineering 22, no. 1 (2023): 52.
- Ljubicic, Dean M., and Brian W. Anthony. “3D high-speed profilometer for inspection of micro-manufactured transparent parts” In Optical Measurement Systems for Industrial Inspection VII, vol. 8082, pp. 320-331. SPIE, 2011

B. Materials Considered

10. In forming my opinions, I reviewed the following documents, in addition to others I have cited in my declaration. I also relied on my own knowledge of and experience in the fields video processing and instrumentation, system and user interface design, human machine interfaces, web and communication technology, interactive streaming technologies, user interfaces, software engineering and algorithm design.

Ex. No.	Description
Ex-1001	U.S. Patent No. 8,542,705 (“the ’705 Patent)
Ex-1002	Declaration of Dr. Brian Anthony
Ex-1003	Curriculum Vitae of Dr. Brian Anthony
Ex-1004	Prosecution History of U.S. Patent No. 8,542,705
Ex-1005	U.S. Patent No. 6,728,965 to Mao (“Mao”)
Ex-1006	U.S. Patent No. 8,630,306 to Liu (“Liu”)
Ex-1007	U.S. Patent No. 7,904,581 to Sherer et al. (“Sherer”)
Ex-1008	U.S. Patent No. 7,965,771 to Wu (“Wu-771”)
Ex-1009	U.S. Patent Publication No. 20020133247A1 to Smith et. al. (“Smith”)
Ex-1010	U.S. Patent No. 8,571,111 to Levy (“Levy”)
Ex-1011	U.S. Patent No. 8,396,113 to Choi (“Choi”)
Ex-1012	U.S. Patent Publication No. 2002/0034374 to Barton (“Barton”)
Ex-1013	U.S. Patent Publication No. 2003/0021346 to Bixby (“Bixby”)
Ex-1014	U.S. Patent No. 9,712,581 to Tinsman (“Tinsman”)
Ex-1015	INTENTIONALLY LEFT BLANK
Ex-1016	INTENTIONALLY LEFT BLANK
Ex-1017	INTENTIONALLY LEFT BLANK
Ex-1018	INTENTIONALLY LEFT BLANK
Ex-1019	INTENTIONALLY LEFT BLANK
Ex-1020	Claim Mapping Table

Ex. No.	Description
Ex-1021	*Adeia Techs., Inc. v. The Walt Disney Co.*, No. 1:24-cv-01231, Defendants' Supplemental Disclosure of Proposed Claim Constructions (D. Del. Sept. 16, 2025)
Ex-1022	*Adeia Techs., Inc. v. The Walt Disney Co.*, No. 1:24-cv-01231, Plaintiffs' Initial Proposed Claim Constructions (D. Del. Sept. 5, 2025)

II. LEGAL STANDARDS

11. In forming my opinions and considering the subject matter of the '705 Patent and its claims in light of the prior art, I am relying on certain legal principles that counsel in this case explained to me. My understanding of these concepts is summarized below.

12. I understand that earlier publications and patents may act to render a patent unpatentable for one of two reasons: (1) anticipation, and (2) obviousness.

A. Anticipation

13. As explained to me by counsel, the claims of a patent are anticipated by a prior art reference if each and every element of the claim is found either explicitly or inherently in the reference. I understand that inherency requires a showing that the missing descriptive matter in the claim is necessarily present in the allegedly anticipating reference, and that it would have been so recognized by a person of ordinary skill in the art ("POSITA").

14. I understand that when a challenged claim covers several structures, either generically or as alternatives, the claim is deemed anticipated if any of the structures within the scope of the claim is found in the prior art reference.

15. Although anticipation typically involves the analysis of a single prior art reference, I understand that additional references may be used to show that the prior art reference has enabling disclosure (i.e., allows a POSITA to make the invention without undue experimentation), to explain the meaning of a term used in the prior art reference, and/or to show that a characteristic is inherent in the prior art reference.

B. Obviousness

16. As explained to me by counsel, a claim is invalid as obvious if it would have been obvious to a person of ordinary skill in the art at the time the alleged invention was made. This means that even if all of the elements of the claim cannot be found in a single prior art reference that would anticipate the claim, a person of ordinary skill in the art who was aware of the prior art would have been able to come up with the claimed invention. This may be the case, for example, where the missing element represents only an insubstantial different over the prior art or a reconfiguration of a known system. I understand that in an obviousness determination, the person of ordinary skill in the art is presumed to have knowledge of all material prior art.

17. I understand that an obviousness analysis requires an understanding of the scope and content of the prior art, any differences between the alleged invention and the prior art, and the level of ordinary skill in evaluating the pertinent art.

18. I understand that when a product is available, design incentives and other market forces can prompt variations of it, either in the same field or a different one. If a person of ordinary skill in the art can implement a predictable variation, obviousness likely bars its patentability. For the same reason, if a technique has been used to improve one device and a person of ordinary skill in the art would recognize that it would improve similar devices in the same way, using the technique would have been obvious.

19. I understand that whether a prior art reference renders a patent claim unpatentable as obvious is determine from the perspective of a person of ordinary skill in the art at the time of the alleged invention. I have been told that there is not requirement that the prior art contain an express suggestion to combine known elements to achieve the claimed invention, but a suggestion to combine known elements to achieve the claimed invention may come from the prior art, as filtered through the knowledge of one skilled in the art. In addition, I have been told that the inferences and creative steps a person of ordinary skill in the art would employ are relevant to the determination of obviousness.

20. I understand that one may consider, e.g., whether (1) the change was merely the predictable result of using prior art elements according to their known functions, or whether it was the result of true inventiveness; (2) there is some teaching or suggestion in the prior art to make the modification or combination of elements claimed in the patent; (3) the claimed innovation applies a known technique that had been used to improve a similar device or method in a similar way; (4) the claimed invention would have been obvious to try, meaning that the claimed innovation was one of a relatively small number of possible approaches to the problem with a reasonable expectation of success by those skilled in the art; (5) the invention merely substituted one known element for another known element in order to obtain predictable results; (6) the invention merely applies a known technique to a known device, method, or product to yield predictable results; or (7) known work in one field of endeavor may have prompted variations of it for use in either the same field or a different one based on design incentives or other market forces that would have been predictable to a person of ordinary skill in the art.

21. I further understand that certain factors may support or rebut the obviousness of a claim. I understand that such secondary considerations include, among other things, commercial success of the patented invention, skepticism of those having ordinary skill in the art at the time of the invention, unexpected results of the invention, any long-felt but unsolved need in the art that was satisfied by the

alleged invention, the failure of others to make the alleged invention, praise of the alleged invention by those having ordinary skill in the art, and copying of the alleged invention by others in the field. I understand that there must be a nexus—that is, a connection—between any such secondary considerations and the alleged invention. I also understand that contemporaneous and independent invention by others is a secondary consideration tending to show obviousness.

22. I am not aware of any allegations by the named inventor of the '705 Patent or any assignee of the '705 Patent that any secondary considerations tend to rebut the obviousness of the '705 Patent.

23. Additionally, I understand that in considering obviousness, it is important not to use the benefit of hindsight derived from the patent under consideration.

C. Level of Ordinary Skill in the Art

24. When interpreting a patent, I understand that it is important to identify the relevant art pertaining to that patent, as well as the level of ordinary skill in that art at the time of the claimed invention. The "art" is the field of technology to which the patent is related.

25. I am informed and understand that the POSITA is a hypothetical person who is presumed to know the relevant prior art. I understand that the actual inventor's skill is not determinative of the level of ordinary skill. I further

understand that factors that may be considered in determining level of skill include: (i) the types 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. I understand that not all such factors may be present in every case, and one or more of them may predominate.

26. I understand that a POSITA is one who is presumed to be aware of all pertinent art, thinks along conventional wisdom in the art, and is a person of ordinary creativity.

27. A person of ordinary skill in the art at the relevant time, which I have been told to consider as January 23, 2007 ("POSITA") would have had at least a four-year degree in electrical engineering, computer science, computer engineering, or a related field, and two-to-three years of experience in the research, design, development, or testing of interactive streaming technologies, computing and software systems user interfaces, and platforms, with additional education substituting for experience and vice versa.

28. Based on my knowledge, skill, and experience, I have an understanding of the capabilities of a person of ordinary skill in the relevant art. For example, from my industry experience, designing and developing dynamic and reconfigurable user interfaces for commercial software applications used in the control and configuration

of electronics, cameras, and data acquisition and distribution system and in designing and developing dynamic and reconfigurable user interfaces for the control and configuration of data and video processing and information extraction algorithms I am familiar with what an engineer designing video-acquisition and distribution and video-analysis software would have known and found predictable in the art. From teaching and supervising my post-graduate students, I also have an understanding of the knowledge that a person with this academic experience possesses. Furthermore, I possessed the knowledge of a POSITA myself at least as of 1997.

D. Claim Construction

29. I understand that the United States Patent and Trademark Office interprets claim terms in an *inter partes* review proceeding under the same claim construction standard that is used in a United States federal court. I understand that under this standard, the meaning of claim terms is considered from the viewpoint of a POSITA at the time of the alleged invention.

30. I have been informed that claim terms are generally given their ordinary and customary meaning as understood by one of ordinary skill in the art in light of the specification and the prosecution history pertaining to the patent. I understand, however, that claims terms are generally not limited by the embodiments described in the specification.

31. I understand that in addition to the claims, specification, and prosecution history, other evidence may be considered to ascertain the meaning of claim terms, including textbooks, encyclopedias, articles, and dictionaries. I have been informed that this other evidence is often less significant and less reliable than the claims, specification, and prosecution history.

III. SUMMARY OF GROUNDS

32. I understand that Petitioner challenges the validity of Claims 1, 2, 4-9, 11-16, and 18-20 of the '705 Patent. It is my opinion that the following grounds demonstrate that the claims of the '705 Patent are invalid.

Ground	Summary
1	Claims 1, 2, 4-9, 11-16, and 18-20 are obvious over Mao (Ex-1005)
2	Claims 1, 2, 4-9, 11-16, and 18-20 are obvious over Mao in view of Liu (Ex-1006)
3	Claims 1, 2, 4-6, 8, 9, 11-13, 15, and 18-20 are obvious over Sherer (Ex-1007) in view of Wu-771 (Ex-1008)

IV. THE '705 PATENT

A. Overview of the '705 Patent

33. The '705 Patent relates to a method, system, and non-transitory computer readable medium that allegedly "improv[es] the ability of a streaming

server to efficiently provide media streams to client devices." Ex-1001, 1:36-39. As shown in Figure 5, a streaming server receives multiple media streams (501) and identifies key frames in the media streams (505) so that when a "connection request is received from a client device, a key frame is selected and provided to the client device." *Id.*, 1:45-48.

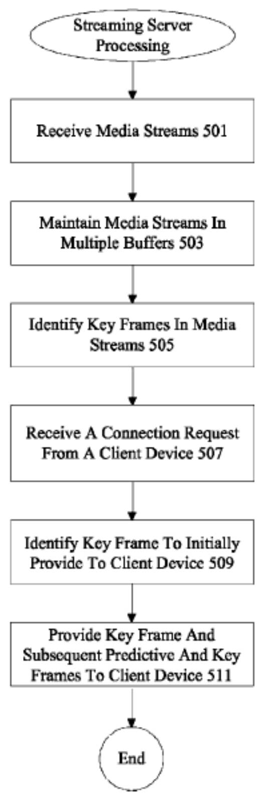


Figure 5

34. The streaming server receives media streams from a media aggregation server, which is "[a]ny mechanism for obtaining media content from a large number of sources in order to provide the media content to mobile devices in live broadcast streams." Ex-1001, 6:54-58. As shown in Figure 4 below, the streaming server

(purple[1]) maintains the media streams in multiple buffers (blue) in step 503. The "buffer to hold media streams," may be a first in first out (FIFO) buffer per channel of video received. *Id.*, 3:8-11. Buffers, such as a FIFO buffer, maintains the sequence of frames for the media streams. *Id.*, 4:48-50; *see also id.*, 8:16-24 ("may be individual physical buffers, portions of buffers, or combinations of multiple physical buffers"). The '705 patent explains that the sequence of video frames associated with a particular channel may be maintained in a buffer assigned to each channel, such that buffers or portions of buffers are maintained for separate video streams and separate channels. *Id.*, 4:13-18.

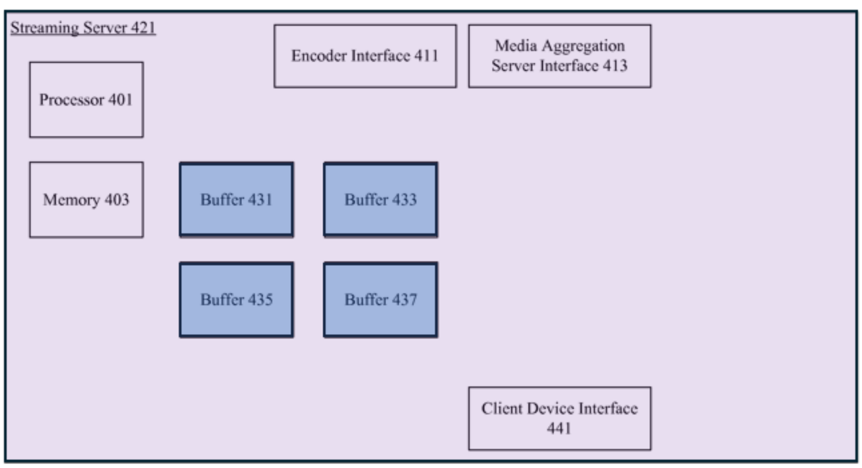


Figure 4

35. Figure 1 (below) shows a sequence of frames that are received at a streaming server, which include "key frame 101" (red) and predictive frames 103,

[1] Throughout, emphasis and annotations are added unless otherwise specified.

105, 107, 109, 11, 113, 115, and 117 (purple). *Id.*, 4:9-12, 4:20-23. "[K]ey frame 101 is an I frame," which is a frame "that includes substantially all of the data needed for a client device to display a frame of video." *Id.*, 4:20-23. Predictive frames are B-frames and P-frames (purple), which are frames that rely on information from a reference key frame. Ex-1001, 2:67-3:7. The sequence of frames, beginning with a key frame and ending with a subsequent key frame is a Group of Pictures (GOP). *Id.*, 4:24-27. Once a media stream is received, the streaming server "identifies key frames in media streams maintained in one or more buffers." *Id.*, 3:61-63.

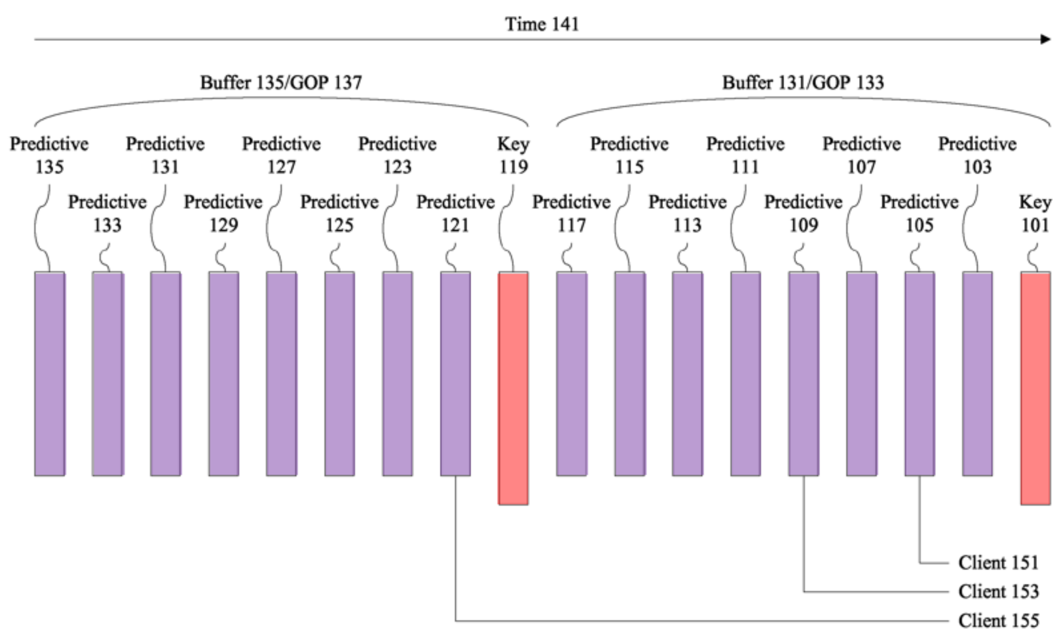


Figure 1

36. When a connection request such as a channel change, an application launch, or a program purchase (Ex-1001, 8:5-7) is received, the streaming server provides the key frame to be displayed without distortion on the client device. Ex-

1001, 3:61-4:4. The ’705 patent’s technique allegedly provides an “intelligent streaming server that identifies key frames in video streams and provides a key frame initially to a client device.” Ex-1001, 8:10-15.

37. As explained below, all subject matter of the challenged claims is described in the prior art and the ’705 claims are obvious in view of the prior art.

B. Prosecution History of the ’705 Patent

38. The ’705 Patent was filed May 26, 2006, as a continuation of International Application PCT/SE04/001540 filed on Oct. 25, 2004. Ex-1001. For purposes of my analysis, I applied the earliest alleged priority date of the ’705 Patent: January 23, 2007.

39. The ’705 Patent was allowed after 6 office actions, including rejections under 35 U.S.C. §§ 101, 102, 103 and an interview with the examiner. The applicant continued to argue that the asserted prior art reference, U.S. Patent Publication No. 2007/0044128 A1 to Witt “converted” the frame that was received, while the claims require the “same key frame received at the streaming server is the key frame provided to the client device.” Ex-1004, 107; *see also id.*, 147, 154. With its request for continued examination, the applicant argued that none of the references cited taught or suggested “disregarding the plurality of predictive frames preceding the key frame maintained in the channel buffer,” “decoding a frame to identify a key frame,” or “transmitting the key frame to the client device in response to the

connection request corresponding to the channel change." *Id*., 167-168; *see also id*., 210-211.

40. On July 29, 2011, the Examiner rejected the amended claims using a combination of Witt, U.S. Patent Publication No. 20110131622 to Wu et al. ("Wu"), and U.S. Patent Publication No. 20070107026 to Sherer et al. ("Sherer"). Ex-1004, 272. The applicant argued that Wu was not applicable prior art because of its filing date (*id*., 287), further amended the claims (*see, e.g., id*., 346, 399) for the independent claims to recite "a plurality of buffers… allocated on a per group of pictures (GOP) basis," and argued that allocating buffers on a GOP basis "allows efficient access of key frames at the beginning of each GOP." *Id*., 464-467. On May 22, 2013, the claims were allowed after an Examiner's amendment reciting that the plurality of buffers are at the streaming server and "determin[e] positions of the plurality of key frames includes determining at the streaming server a group of pictures (GOP) size N and identifying Nth frames following the first key frame as subsequent key frames." *Id*., 480. The '705 Patent issued on September 24, 2013. *Id*., 500.

41. The Examiner materially erred by not considering the Sherer and earlier Wu reference (Wu-771) during prosecution because, as demonstrated in Ground 3 below, Sherer in view of Wu-771 renders obvious each limitation of Claims 1, 2, 4-6, 8, 9, 11-13, 15, and 18-20. It was a material error to not apply Sherer and Wu-

771's teachings to the amended claims during prosecution, especially in view of the Applicant's silence as to Sherer's teachings and the applicant's erroneous arguments regarding Wu-771's applicability as prior art.

42. In addition to the Sherer-Wu-771 grounds, the prior art I apply to the claims in my Declaration includes new arguments and prior art not before the USPTO during prosecution of the '705 Patent. And as I discuss below, this art discloses the '705 Patent's Claims 1, 2, 4-9, 11-16, and 18-20, including the features the applicant contended were new during prosecution.

V. TECHNOLOGY BACKGROUND

A. Media streaming problems and solutions were known prior to 2007

43. Media streaming, along with the various design challenges and the solutions proposed to address these challenges, were well known to a POSITA by 2007. For example, it was known by 2007 that the Internet has caused a significant shift in traditional broadcasting of media content. Ex-1009 ¶5. Specifically, "Internet content providers offer users access to multimedia content online, as an alternative to traditional media such as television, radio, prerecorded video, and audio." *Id*. "Traditionally, users played multimedia content by accessing entire multimedia content files from removable media or downloading them via a network. Downloading and storing entire multimedia content files, however, consume valuable network and storage resources." *Id*. ¶6. Media streaming was developed to

"address these problems" by providing a "technique for transferring multimedia content data from a server to a client device and processing it at the client device as a steady and continuous media stream." *Id*. ¶6. Media streaming allows "[l]ive events, such as concerts and sports events may be streamed to client devices as they occur, rather than creating a monolithic multimedia content file once the event has ended." *Id*. ¶7. For example, client devices play streamed multimedia content as it received from the server, rather than waiting to download an entire content file.

44. One problem that POSITAs realized is the "unacceptably long delay and its accompanying discontinuity in switching between different media streams, often leaving the user to stare at a blank or frozen media player for several seconds" due to the "time it takes for a media player to stop playing a first media stream, request a second media stream from the server, wait to receive the second media stream, and begin playback of the second media stream." Ex-1009 ¶9. POSITAs thus recognized the importance of providing a system, method, and computer-readable medium for switching between multiple media streams that provide the user with continuous playback of multimedia content." *Id*. ¶10.

45. U.S. Patent Publication No. 20020133247A1 to Smith et. al. ("Smith," Ex-1009) was filed November 13, 2001 and published September 19, 2002. Smith provides a method and system for "seamlessly switching between media streams during playback, comprising: receiving a first media stream; playing said first media

stream; receiving a second media stream; playing said second media stream; and ending play of said first media stream after said playing said second media stream." Ex-1009, Cl. 1. Smith's Figure 5 (reproduced below) shows a process for "seamless switching" of media streams. Beginning in step 502, Smith's media server begins providing the first media stream to a client. *Id*. ¶60. At step 504, Smith's delivery layer 104 receives a request for a second media stream from the client and in step 506, the media server streams the second media stream to the client. *Id*. ¶¶61-62. Smith explicitly teaches that the media server 102 stops sending the media stream in step 510. *Id*. ¶62.

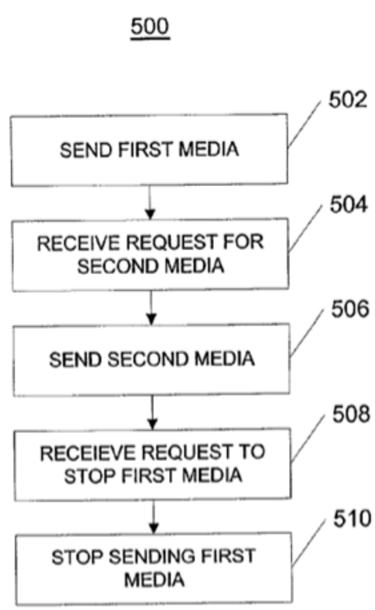


Fig. 5

46. Smith specifically explains that after receiving the request for the second media stream, the media server streams the second media stream to the client.

For example, Smith's background media player 302 "buffer[s] the second media stream before playing it." Ex-1009 ¶62. "Given present bandwidth constraints on network 108, in order to enable media stream switching, the first media stream, the second media stream, or both are buffered at media client 110. Buffering a media stream allows media player 204 to play the buffered media stream even in the absence of the media stream itself." *Id*. ¶34.

47. Similarly, U.S. Patent No. 8,571,111 to Levy ("Levy," Ex-1010) also discusses how channel switch time causes significant delay. Levy explains that "[c]hannel switch time delay may be caused by several factors such as radio frequency (RF) tuner frequency change, quadrature amplitude modulation (QAM) lock, symbol lock, packet identification (PID) decoding and program acquisition, intra-coding frame (I-frame) delay, video frame buffer build, and program selection lag." Ex-1010, 1:18-24. Levy specifically states that "there typically may be as many as 15 frames between I-frames in a GOP. Consequently, when decoding MPEG-2 video streams, there may be uncertainty as to the starting point for the decoder when frames are picked off the video stream after the channel is changed and demodulation begins. In some applications, ***there may be as much as a 500 ms delay to get an I-frame and begin proper decoding after a channel switch occurs***. For example, refresh periods of a half-second are common in a DTV environment, and longer refresh periods may be experienced in some situations. ***This causes serious lag, and***

a blank or frozen screen is presented until an I-frame is detected." Ex-1010, 7:52-63. Therefore, Levy also described the well-known problem of the delay in providing the requested media. Levy provides a remedy by teaching a system "for switching programs streams using a fixed speed program bugger coupled to a decoder" in order to receive "multiple program streams, such as video program streams, and to buffer historical frames comprising at least one anchor frame." *Id*., 1:43-49.

48. U.S. Patent No. 8,396,113 to Choi ("Choi," Ex-1011) is another reference that describes the well-known problems of time delay when providing the requested media. Choi teaches that "since traffic information encoded by an encoder is interleaved and transmitted on the very large unit, a receiving device cannot perform decoding until it receives all the interleaved traffic information. Further the receiver cannot perform the decoding until it completes the deinterleaving. Accordingly, in the satellite DMB service, the receiving device cannot decode broadcast service traffic until it receives data of a known amount interleaved before decoding the received traffic. Therefore, a significant delay is caused in providing the service." Ex-1011, 2:14-24. In particular, Choi discusses how a "video extraction key frame delay 103 refers tot eh time for waiting until a key frame, for example, an instantaneous decoder refresh (IDR) frame…, is reached to decode a compressed image in a normal fashion. *Id*., 2:32-35. Choi thus provides "a data receiving device and method for shortening a channel switching time of a digital multimedia

broadcasting receiving device by minimizing a video key frame extraction delay of a codex delay during a delate time generated at the time of channel switching of a digital multimedia broadcast service." *Id*., 3:33-38.

49. Indeed, the time delay was also noted in U.S. Patent No. 6,728,965 to Mao ("Mao," Ex-1005). Mao explains that emerging technologies have made digital delivery systems economically feasible. Mao is directed "to an apparatus for the distribution and delivery of video in a digital system and, in particular, to a method and apparatus that rapidly changes the channel in such a system." Ex-1005, 1:7-10. Mao teaches that in typical digital video systems, "all of the video services offered by the video service company are again delivered to the [user's] set-top unit," such that the set-top unit "performs the actual switching and also descrambles the digital signals." *Id*., 1:52-58. Mao provides that channel switching is provided "upstream," but that "a drawback of this system is that the subscriber experiences a time delay between the period of time it takes for the subscriber to select a channel and for the newly selected channel to be viewed on the television." *Id*., 2:4-7. Mao teaches that "[a] significant portion of the delay is caused by the time it takes for the video signal to synchronize. This portion accounts for about half of the overall delay (i.e., about a half second)." *Id*., 2:15-18. Mao notes that users find the delay "annoying" and "a one second delay is unacceptable, which would make such a system competitively unattractive." *Id*., 2:15-23.

50. Mao discloses a "method and apparatus [for] rapidly changing the channel in a digital video delivery system." Ex-1005, 2:26-28; *see also id.*, 1:7-10. Mao teaches a "rapid channel changer" that receives a "subscriber's channel change request" and assess and directs the corresponding video signal to the subscriber. Ex-1005, Abstract. For example, when a subscriber makes a request for a change of channel, for example to channel X, using its set-top unit 19, the request for the channel change is sent to broadband digital terminal 12 (BDT) over the broadband access system. Ex-1005, 5:51-57.

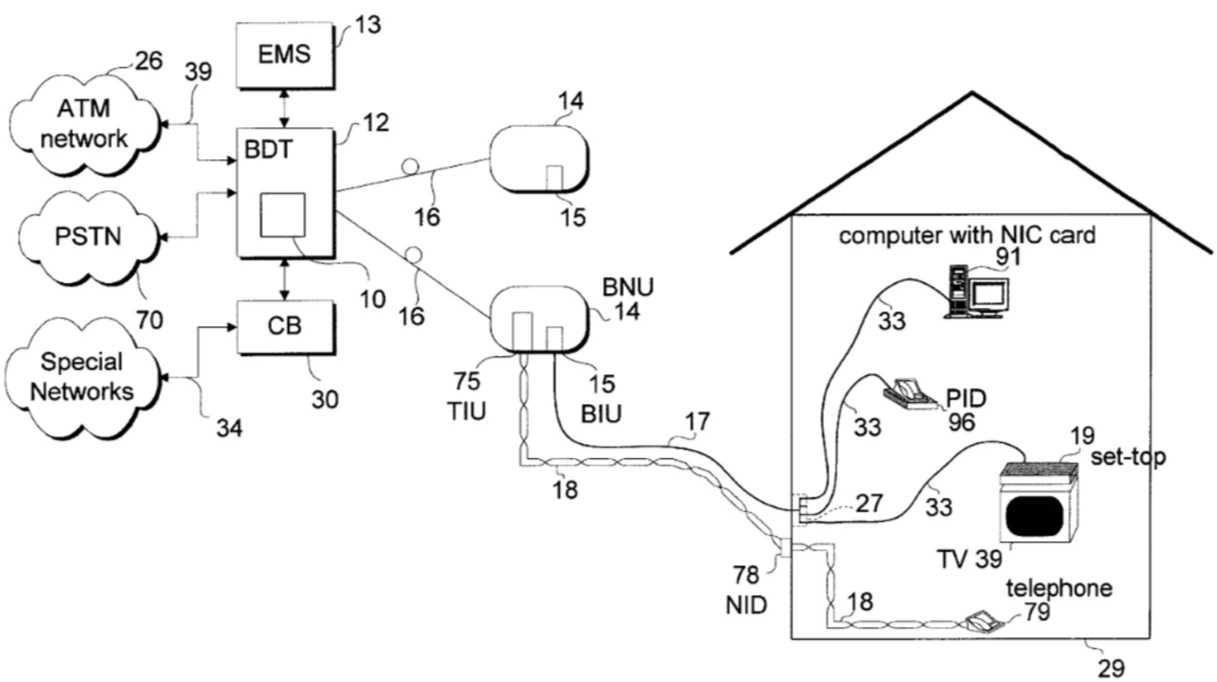


FIG. 1

51. Mao's digital video system includes a broadband digital terminal 12 (BDT) connected to a broadband network unit 14 (BNU), which "bundles all of the

received control signals from the… set-top units it serves and transmits them to the BDT." BDT 12 is "connected to a video source (ATM network 26 or special network 32) that provide a plurality of compressed video signals." Ex-1005, 7:62-64. The BDT receives video signals comprising "a continuous stream of a predetermined number of consecutive data frames." *Id.*, 10:12-13 (Cl. 5). Mao explains that the Group of Pictures (GOP) starting point is coded by an "I" or "intracoded frame[]" and followed by "predicted frames and bidirectional frames (P and B frames, respectively)." *Id.*, 7:29-31. As shown in Figure 7 below, Mao's BDT 12 includes a "first-in-first-out (FIFO) buffer 50 for storing the compressed video data of each video channel." Ex-1005, 8:14-16. Mao provides an example that "forty FIFO buffers 50 will be used" when a subscriber has "forty channels from which to choose." *Id.*, 8:59-62; *see also id.*, 8:17-21 ("BDT 12 "utilizes buffers to store the video data for each channel it is supposed to receive."), 8:39-31 ("multiple buffer units may be used, e.g., one for each video channel").

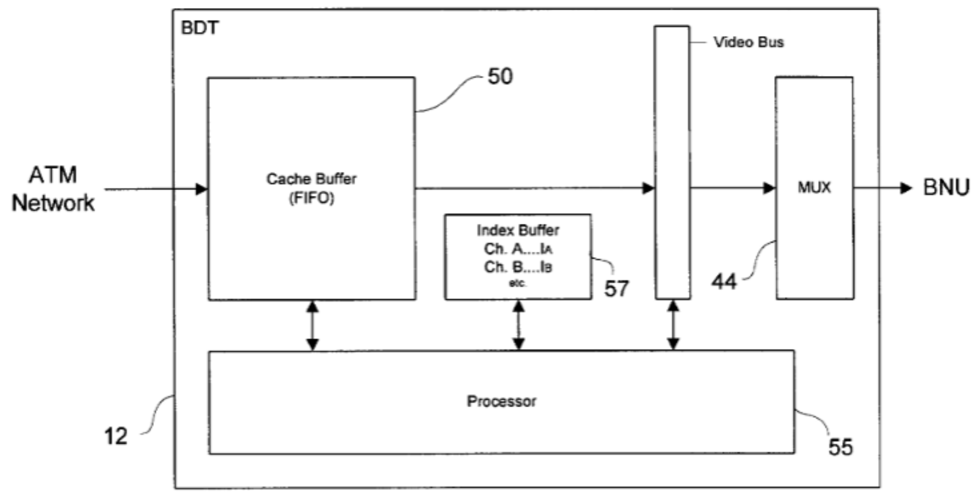


FIG. 7

52. Therefore, once the set-top unit's request to change the channel and the BDT determines that the request is proper, the BDT then "direct[s] the corresponding video data downstream to the appropriate set top unit 19." Ex-1005, 5:55-65. BDT synchronizes with "channel X as it is received from the ATM network" by decompressing and multiplexing the channel. *Id.*, 6:20-26. The BDT "must then wait for the Group of Picture (GOP) start point of the requested channel." Ex-1005, 7:8-12. The GOP start point is coded in the "I" or intracoded frames. *Id.*, 7:29-31. Thus, when "a change of channel request reaches BDT 12, the processor can immediately look up the location of the I frame from the pointer location buffer 57 and immediately access the appropriate video data." Ex-1005, 8:52-55.

53. Mao's "subject rapid channel changer 10 is designed to minimize any delay between a subscriber's request to change a television channel and the actual

delivery of the signal corresponding to the desired channel." Ex-1005, 8:10-14. Specifically, Mao's processor that is "always 'pointing'" to an I frame for each respective channel, eliminates the "the largest portion of the channel changing delay—up to a ½ second that the processor had to wait for the next synchronization frame." *Id*., 8:44-51.

54. Similarly, U.S. Patent No. 7,965,771 to Wu et al. ("Wu-771," Ex-1008) discusses the time delay known to POSITAs. Wu-771 was filed February 27, 2006 and issued June 21, 2011. Wu-771 explains that a drawback to video compression schemes is when a "subscriber switches channels, there can be a significant delay incurred in acquiring and displaying the selected new channel on the television screen." Ex-1008, 2:9-13. Specifically:

> If the subscriber happens to switch channels right after an I frame was just transmitted for the new channel, the TV cannot display any video until the next I frame is received. Although a series of P and B frames are received, they are useless because a corresponding previous I frame is needed as a reference. Essentially, the viewer can occasionally experience a momentary blank or black screen right after switching channels. This can be quite disconcerting to viewers who are accustomed to traditional TV whereby switching channels instantaneously brings up the new channel for viewing.

Ex-1008, 2:13-22.

55. In summary, POSITAs were well aware of the time delays in switching channels and aimed to provide solutions by buffering portions of the video stream to avoid unnecessary delays.

B. Video compression schemes use Group of Pictures as part of the video coding.

56. Mao (discussed above) teaches that a "compression scheme is preferably utilized to pack more video data into each information field. Virtually all of the compression schemes require a starting or synchronization frame usually referred to as the Group of Picture (GOP) start point. In most compression schemes the GOP start point is found proximate the beginning of the information field." Ex-1005, 7:1-6. Mao discusses that its broadband digital terminal 12 receives a subscriber's channel change request and "must then wait for the Group of Picture (GOP) start point of the requested channel. Even if the ATM system is dedicated strictly to video, the time between GOP start points is significant. As illustrated in FIG. 4, the time delay between GOP_{B1} and GOP_{B2} may be ½ second or more depending on the compression scheme and other factors." *Id.*, 7:7-15.

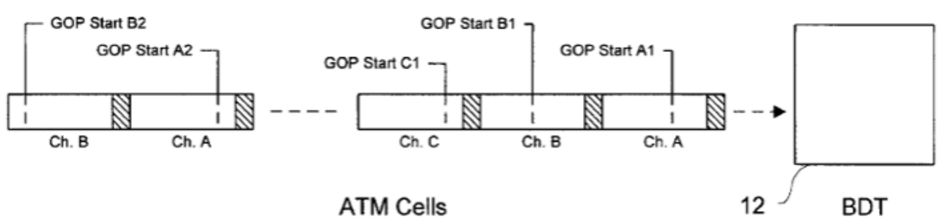


FIG. 4

57. Mao teaches a rapid channel changer, which used "Moving Pictures Experts Group 2 format, known as MPEG-2" to compress and encode the video data received. Ex-1005, 7:26-28. As shown in Figure 5, the data is transmitted as "one of three basic frames":

> The GOP start point is coded in the "I" or intracoded frames. In addition to they frames, there are predicted frames and bidirectional frames (P and B frames, respectively). The P and B frames normally contain the video and audio content. Each synchronization frame is separated from the next synchronization frame by a pre-determined number of other frames. This predetermined number can be set to accommodate a specific requirement but in one preferred embodiment is fifteen frames.

Ex-1005, 7:28-38.

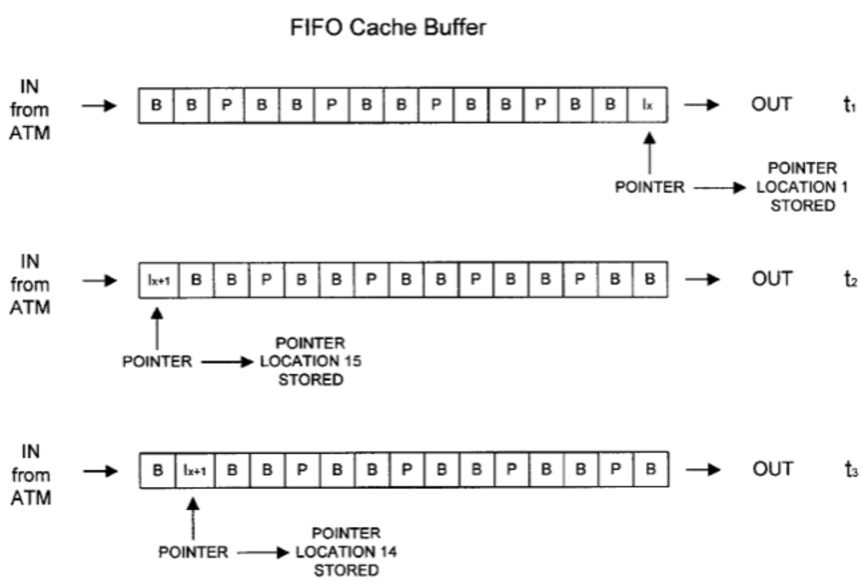


FIG. 5

58. Indeed, the MPEG-2 standard is a widely adopted compression standard. U.S. Patent No. 7,965,771 to Wu et al. ("Wu-771," Ex-1008) was filed February 27, 2006 and issued June 21, 2011. Wu-771 emphasized that video compression allows for "minimizing the bandwidth requirements associated with transmitting video entails compressing the video …, transmitting the smaller compressed video data, and then decompressing the video data before being displayed." Ex-1008, 1:54-60. Wu-771 explains the MPEG-2 standard as follows:

> For instance, the MPEG-2 standard converts a video stream into I, P and B frames. The I frame is an intra-frame which contains all the data required to display the frame. An I frame is followed by a series of P and B frames. The P frame is a predictive frame which contains only the data that has changed from the preceding I frame. P frames rely on I frames to fill in most of its data. The B frame is a bidirectional frame which contains data that have changed from the preceding frame or are different from the data in the very next frame. By converting the frames of the original video stream into much smaller I, P, and B frames, the video stream comprised of I, P and B frames can be transmitted at a fraction of the bandwidth compared to non-compressed video streams. Once received, the I, P, and B frames are then used to reconstruct the original video stream for playback.

Ex-1008, 1:58-2:9; *see also id*., 3:7-17.

59. Wu-771 explains that most carriers deliver video to subscribers using MPEG2. Ex-1008, 3:7-8, 1:58-60.

> More specifically, MPEG2 encodes video into I, P and B frames or elementary streams. I frames are intra-coded only. An I frame serves as a reference frame for future predictions. Moderate compression (on order of 10:1), limits the propagation of transmission of errors, supports random access and fast forward/fast reverse. P frames are forward prediction from either previous I frames or previous P frames. P frames serve as reference for future P or B frames. P frames give good compression savings (20:1). B Frames are bi-directional interpolated prediction from two sources. B frames serve as previous reference I or P frames (forward prediction) or as future reference I or P frames (backwards prediction). B frames confer the highest compression (50:1).
>
> I Frames=Intra-coded only—reference frame for future predictions.
>
> P Frames=Forward prediction from either previous I frames or previous P frames. Reference for future P or B frames.
>
> B Frames=Bi-directional interpolated prediction from two sources. Previous reference I or P frames (forward prediction). Future reference I or P frames (backwards prediction).

Ex-1008, 3:18-37.

60. Similarly, Levy teaches that "program streams received by the fixed speed program stream buffer module 110 and/or the buffers 112-1-N may comprise video streams, such as MPEG-2 encoded video streams. In such implementations, each of the program streams may comprise as a sequence of pictures, such as a group of pictures (GOP) or long GOP (LGOP). When encoded using MPEG-2, for

example, each picture may be segmented into macroblocks, and the sequence of pictures may be implemented by three types of compressed frames: intra-coded frames (I-frames), predictive-coded frames (P-frames), and bi-directionally predictive-coded frames (B-frames)." Ex-1010, 7:5-16. For example, in Levy's Figures 2a and 2b, Levy illustrates a program stream switching apparatus that "comprise MPEG-2 video program streams, including a sequence of pictures including I-frames, P-frames, and B-frames." *Id.*, 9:16-23. In the switching apparatus 200, two "fixed speed program stream buffers" are used to receive the video programs streams from a transport de-multiplexer and coupled to a MPEG-2 decoder. *Id.*, 9:24-33. Each of the buffers comprise "at least sixteen frames including at least one I-frame at all times. It can be appreciated that the I-frames are not synchronized. At a time immediately preceding a channel switch point, a B-frame is sent from the first buffer 202 (e.g., active channel program stream buffer) to the MPEG-2 decoder 206." *Id.*, 9:33-39; *see also id.*, 9:40-52.

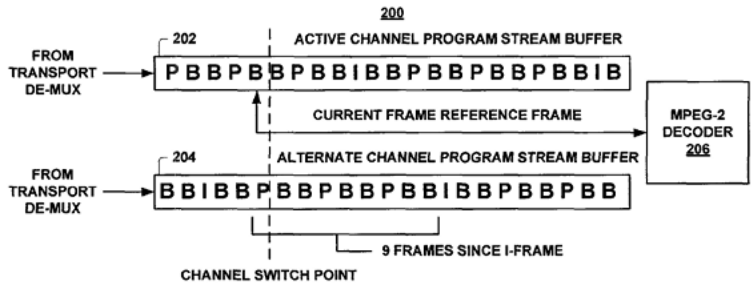


FIG. 2A

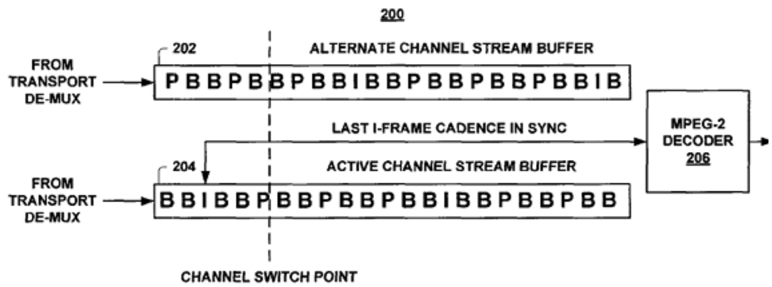


FIG. 2B

61. U.S. Patent Publication No. 2002/0034374 ("Barton," Ex-1012) also teaches that "[l]imited storage capacity is dealt with by compressing the video stream using an algorithm, typically one of the MPEG (Moving Pictures Experts Group) standard algorithms, which can achieve a useful compression of 100:1 in many instances. MPEG video is represented as a sequence of Groups Of Pictures (GOPS), in which each GOP begins with an index frame, called the I-frame." Ex-1012 ¶12. Barton further explains that the GOP represents "up to 15 additional frames by providing a much smaller block of digital data that indicates how small portions of the I-frame, referred to as macroblocks, move over time. Thus, MPEG achieves its compression by assuming that only small portions of an image change over time, making the representation of these additional frames extremely compact." *Id.* ¶13.

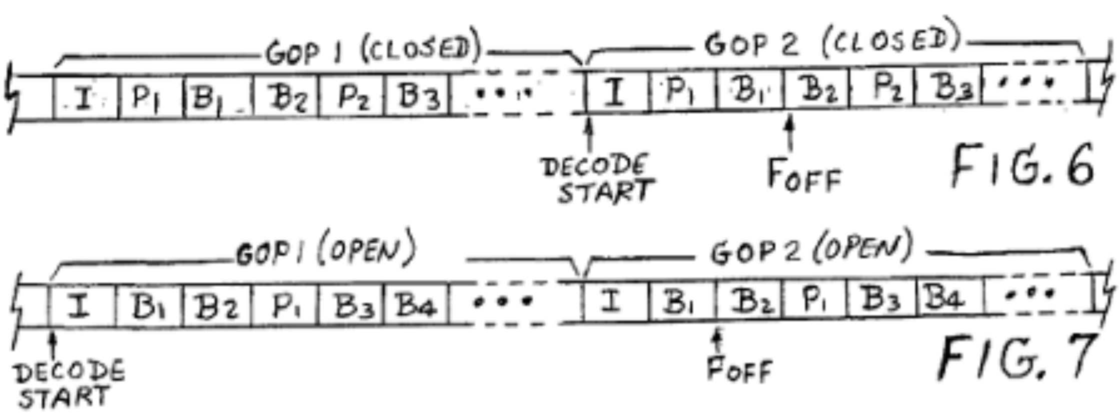
62. Choi (discussed above), further discuses how "in a compression method like H.264, there are three frame types;[] intra-coded (I), predictive coded (P), and bi-directionally predicted coded (B) frames depending on a method for deciding a reference frame." Ex-1011, 2:35-39. Choi further explains that:

> The I frame refers to a frame for simply spatially converting only the frame without a motion compensation. The P frame refers to a frame for performing the motion compensation in a forward direction or in a reverse direction with reference to the I frame or the P frame and then, spatially converting a residual. The B frame refers to a frame for performing the motion compensation like the P frame but performing the motion compensation by the I and P frames, that is, in both directions on a time axis.

Id., 2:39-47. Choi discusses that the "key frame is a complete picture used for video compression. Referring to video group of pictures (GOP) structure, a frame is selected one by one at a constant interval from a temporal video flow, and is designated as the key frame. The key frame is an independently restorable video, and enables an arbitrary video access." *Id*., 2:55-60. In particular, "MPEG, H.26a, and H.264, the key frame is inserted at a constant interval into a GOP, and designates the I fame enabling independent video production." *Id*., 61-63.

63. U.S. Patent Publication No. 2003/0021346 ("Bixby," Ex-1013) similarly teaches that "[f]or the video data, MPEG provides a high degree of compression (up to 200:1) by encoding 8×8 blocks of pixels into a set of discrete

cosine transform (DCT) coefficients, quantizing and encoding the coefficients, and using motion compensation techniques to encode most video frames as predictions from or between other frames. In particular, the encoded MPEG video stream is comprised of a series of groups of pictures (GOPs), and each GOP begins with an independently encoded (intra) I frame and may include one or more following P-frames and B-frames." Ex-1013 ¶6. For example, Figures 6 and 7 (reproduced below) show an MPEG transport Stream including a series of GOPs and showing a decode start point for a particular frame offset of a frame to be decoded. *Id.* ¶¶27-28.



C. Multiple streaming buffers were commonly used in video streaming systems and methods.

64. Smith (discussed above) also teaches a media streaming method and apparatus, which includes multiple streaming buffers. For example, Smith discloses

buffering a "first media stream in a first media stream buffer" and buffering a "second media stream in a second media stream buffer." Ex-1009, Cls. 14, 17.

65. Mao also teaches that its broadband digital terminal 12 includes a "first-in-first-out (FIFO) buffer 50 for storing the compressed video data of each video channel." Ex-1005, 8:14-16. For example, "a subscriber may have forty channels from which to choose; therefore, forty channels may be sent from the ATM network 26 to the BDT 12" and "forty FIFO buffers 50 will be used for this example of the rapid channel changer 10." Ex-1005, 8:59-62. Indeed, multiple buffer units may be used, e.g., one for each video channel." *Id.*, 8:39-31. Mao also teaches that if the BDT 12 "already utilizes buffers to store the video data for each channel it is supposed to receive, and a processor, the subject invention can be implemented in existing hardware, which minimizes additional costs." *Id.*, 8:17-21.

66. Wu-771 also teaches a "video acceleration module VAM" which buffers the frames. *Id.*, 4:66-67. Wu-771's VAM has "separate outgoing client buffers for each client port (PBUF) 401" and "***fixed buffers for each multicast stream*** it is sourcing to clients, referred to herein as Streaming Buffers (StrmBuf)." Ex-1008, 4:67-5:3. Wu-771's StrmBuf 402 buffers "fixed number of GOP's aligning the start of the buffer to the start of a GOP." *Id.*, 5:3-6. Wu-771 teaches that "the higher number of I frames sent, the faster the decoder can start displaying the picture

(i.e., faster channel change time form a user perspective), since the decoder needs an I frame to initially sync to; again." Ex-1008, 3:59-61.

67. Levy discloses multiple buffers as well: "The media system 100 may comprise a fixed speed program stream buffer module 110. The fixed speed program stream buffer module 110 may be coupled to the program stream select module 108 and may be implemented by a plurality of buffers 112-1-N, where N may represent any positive integer value consistent with the described embodiments. In various implementations, ***the number of buffers 112-1*** employed may be designed to accommodate frequent switching from buffer to buffer that may occur when rapidly changing channels." Ex-1010 (Levy) 6:46-55. As shown in Figure 1 below, there are distinct and separate channel stream buffers.

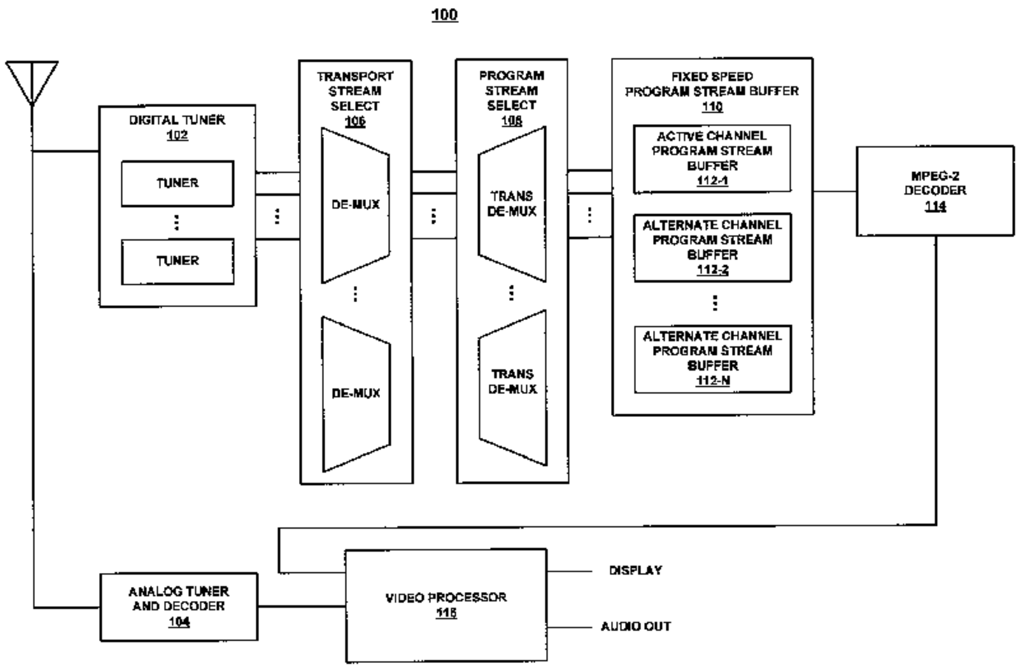


FIG. 1

68. Bixby is another example that discusses having multiple buffers. For example, as show in Figure 2 (below).

> A first clip of MPEG-2 encoded video data 61 is contained in the cached disk storage 26. The data mover 28 fetches blocks of data from the first clip 61 and stores them in a random access memory buffer 62 in the data mover 28. The decoder 32 includes a decoder data buffer 63 for storing MPEG-2 encoded video data, MPEG-2 decoder logic 64 for decoding the MPEG-2 encoded video data, and a video frames buffer 65 for storing frames of video data. The decoder 32 fetches the blocks of data from the buffer 62 in the data mover 28 and stores them in the decoder data buffer 63. The MPEG-2 decoder logic 64 decodes the blocks of data in the decoder data buffer 63 to produce frames of video

> data and stores the frames of video data in the video frames buffer 65. Data of a video frame having a current time code is presented in real-time to the A/B switch 38, and when the A/B switch is in the "A" channel selected state, the A/B switch outputs the real-time video stream from the decoder 32. In a similar fashion, FIG. 2 shows a "B" channel data path including a second clip 66 in cached disk storage 26, a buffer 67 in the data mover 28, a decoder data buffer 68 in the decoder 33, MEPG-2 decoder logic 69 in the decoder 33, and a video frames buffer 70. The buffer 62 and the buffer 67 share random access memory 71 in the data mover 28, and generally speaking, when the buffer 62 is substantially full, the buffer 67 is substantially empty."

Ex-1013 ¶60. Therefore, frames of video stream A are stored in one buffer and frames if video stream B are stored in another buffer.

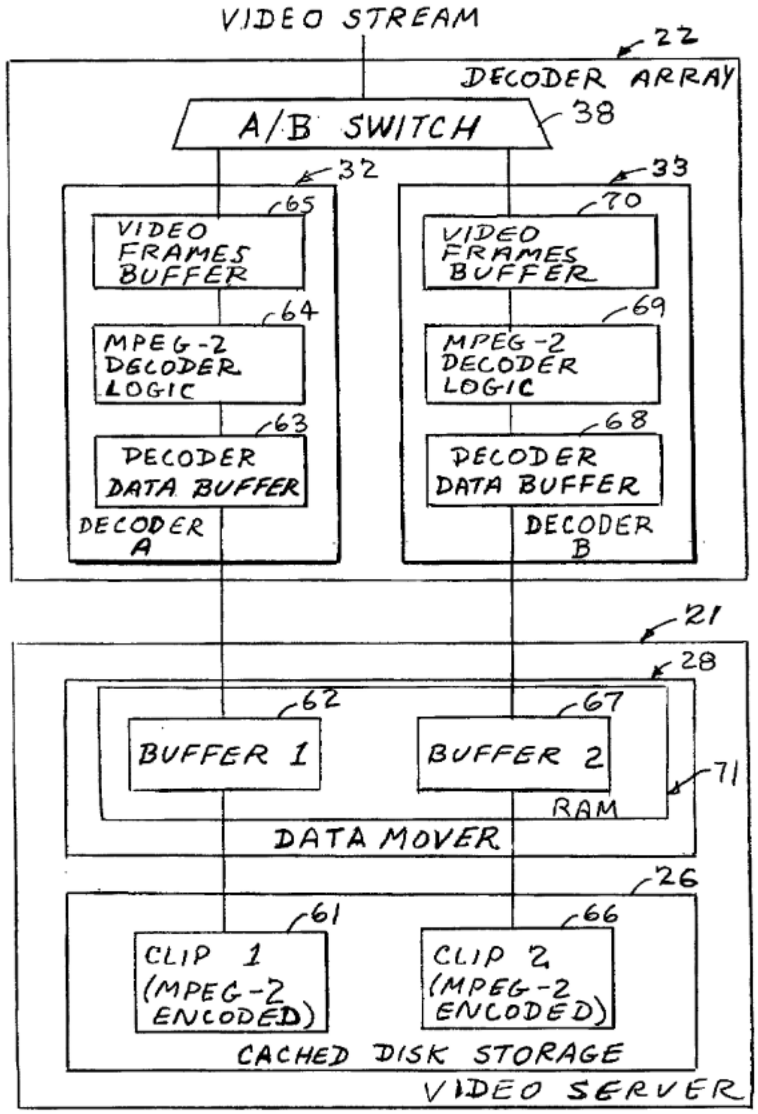


FIG. 2

69. Similarly, U.S. Patent No. 9,712,581 ("Tinsman," Ex-1014) also discloses a switch/router being built with multiple video stream buffers: 2406.1, 2406.2 …, 2406.N: "FIG. 24 is a block diagram of an example switch/router 2400, which may serve as the switch/router 2104 of FIG. 21, the switch/router 104 of FIG. 1, and/or the switch/router 229 of FIG. 2. As shown, the switch/router 2400 may

include a multicast interface 2402, a point-to-point interface 2404, at ***least one video stream buffer 2406.1, 2406.2, . . . 2406.N***, and control logic 2408. The switch/router 2104 may be embodied in hardware, software, firmware, or some combination thereof, as one or more machines, such as the computer system 3400 of FIG. 34 described below." Ex-1014, 21:41–50. Tinsman further claims "receiving a plurality of data streams originating from one or more media content sources into one or more buffers of a router, each of the plurality of data streams comprising a plurality of contiguous, non-overlapping data sets of a first media content, and each of the plurality of data streams having a unique data rate relative to one or more other data streams of the plurality of data streams." *Id*., Cl. 1.

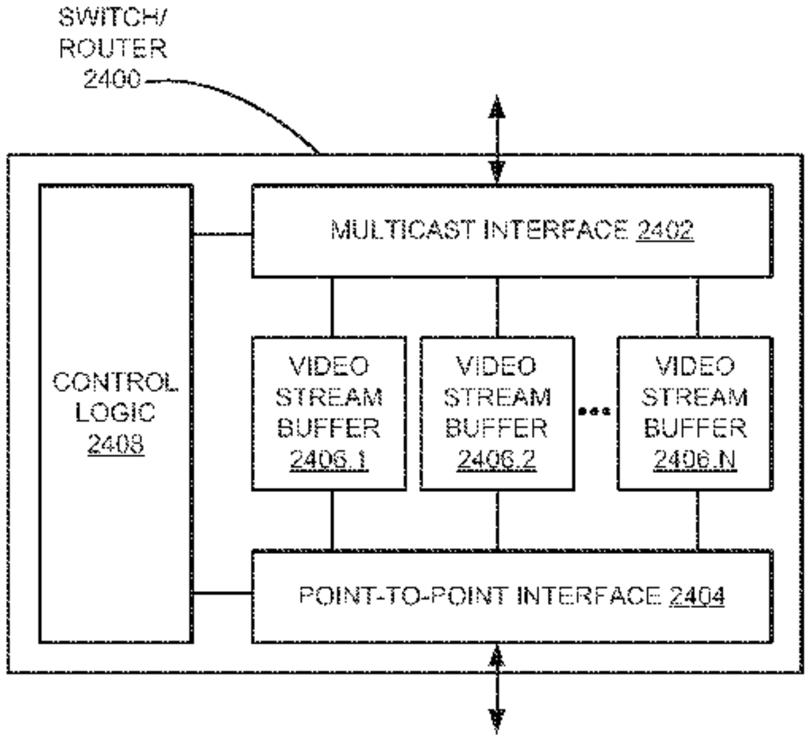


FIGURE 24

70. Similarly, Choi teaches allocating buffers per GOP basis. Specifically, Choi discloses " a GOP buffer for receiving the MPEG2-TS packet" (Ex-1011, Abstract) and that "[t]he GOP buffer 302-1 ... 302-n always stores the MPEG2 TS packet corresponding to the video key frame at its start position, and sequentially stores MPEG2-TS packets not corresponding to a subsequently received key frame, which are video P-frames, or MPEG2-TS packets corresponding to audio data." Ex-1011, 6:32-37.

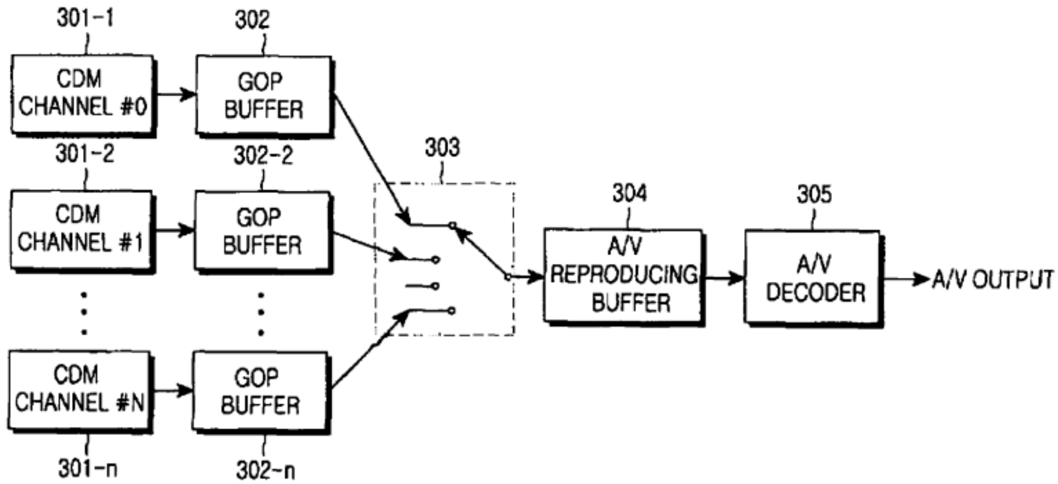


FIG 3

VI. BRIEF DESCRIPTION OF THE APPLIED PRIOR ART

A. Mao (Ex-1005)

71. Mao (U.S. Patent No. 6,728,965) was filed on August 20, 1997, issued on April 27, 2004. As explained to me by counsel, I understand that Mao qualifies prior art to the '705 Patent under 35 U.S.C. §102(a), (b), (e).

72. Mao discloses a "method and apparatus [for] rapidly changing the channel in a digital video delivery system." Ex-1005, 2:26-28; *see also id.*, 1:7-10. Mao teaches a "rapid channel changer" that receives a "subscriber's channel change request" and assess and directs the corresponding video signal to the subscriber. Ex-1005, Abstract. For example, when a subscriber makes a request for a change of channel, for example to channel X, using its set-top unit 19, the request for the

channel change is sent to broadband digital terminal 12 (BDT) over the broadband access system. Ex-1005, 5:51-57. Mao's digital video system includes a broadband digital terminal 12 (BDT) connected to a broadband network unit 14 (BNU), which "bundles all of the received control signals from the… set-top units it serves and transmits them to the BDT." Ex-1005, 6:10-12. BDT 12 is "connected to a video source (ATM network 26 or special network 32) that provide a plurality of compressed video signals." Ex-1005, 7:62-64.

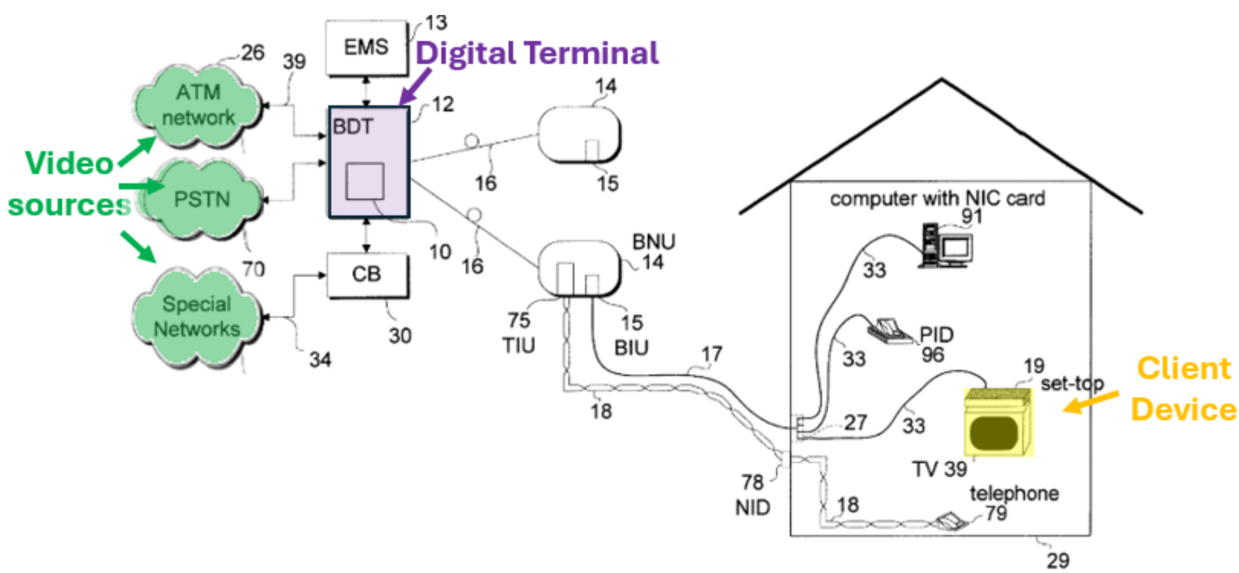


FIG. 1

73. The BDT receives video signals comprising "a continuous stream of a predetermined number of consecutive data frames." *Id.*, 10:12-13 (Cl. 5). Mao explains that the Group of Pictures (GOP) starting point is coded by an "I" or "intracoded frame[]" and followed by "predicted frames and bidirectional frames (P and B frames, respectively)." *Id.*, 7:29-31. As shown in Figure 7 below, Mao's BDT

12 includes a "first-in-first-out (FIFO) buffer 50 for storing the compressed video data of each video channel." Ex-1005, 8:14-16. Mao provides an example that "forty FIFO buffers 50 will be used" when a subscriber has "forty channels from which to choose." *Id.*, 8:59-62; *see also id.*, 8:17-21 ("BDT 12 utilizes buffers to store the video data for each channel it is supposed to receive."), 8:39-41 ("multiple buffer memory units may be used, e.g., one for each video channel").

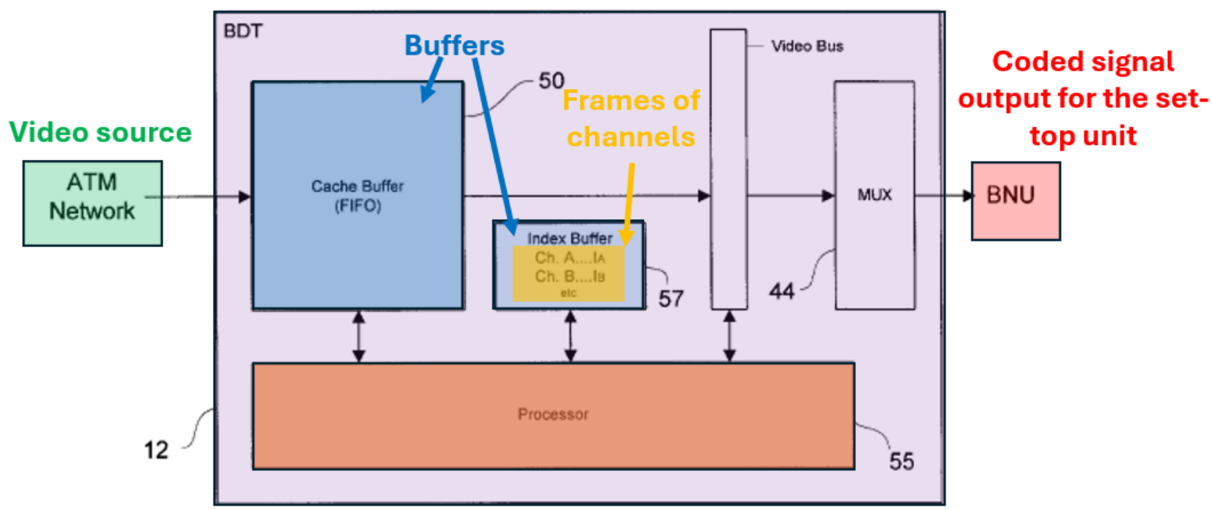


FIG. 7

74. BDT 12's processor 55 receives the frames for each channel and detects, establishes, and records the GOP start frame for each channel. *Id.*, 8:15-26. Each I frame (start frame) is separated from the next I frame "by a pre-determined number of other frames." Ex-1005, 7:34-36. Mao further teaches that two or more synchronization frames per video signal can be determined and maintained. Ex-

1005, 9:2-6. Mao specifically provides that in a MPEG-2 scheme with a GOP size of fifteen, up to thirty frames can be stored. *Id*.

75. Therefore, once the set-top unit's request to change the channel and the BDT determines that the request is proper, the BDT then "direct[s] the corresponding video data downstream to the appropriate set top unit 19." Ex-1005, 5:55-65. BDT synchronizes with "channel X as it is received from the ATM network" by decompressing and multiplexing the channel. *Id*., 6:20-26. The BDT "must then wait for the Group of Picture (GOP) start point of the requested channel." Ex-1005, 7:8-12. The GOP start point is coded in the "I" or intracoded frames. *Id*., 7:29-31. Thus, when "a change of channel request reaches BDT 12, the processor can immediately look up the location of the I frame from the pointer location buffer 57 and immediately access the appropriate video data." Ex-1005, 8:52-55.

B. Liu (Ex-1006)

76. Liu (U.S. Patent No. 8,630,306) was filed on January 9, 2006 and issued January 14, 2014. As explained to me by counsel, I understand that Liu qualifies prior art to the '705 Patent under 35 U.S.C. §102(e).

77. Liu teaches "a computer program embedded in a computer readable medium that contains instructions operative to be executed by a processor," "a method for providing video content," and "a network element that is operative to provide content to a user in a network." Ex-1006, 2:25-27, 2:50-52, 2:65-66. As

shown in Figure 1 below, Liu teaches an Internet Protocol Television (IPTV) Service Architecture 100 with a Digital Subscriber Line Access Multiplexer (DSLAM) 110 (purple), which "perform[s] traffic aggregation, grooming, policing, access multiplexing and other tasks," such as receiving media streams from the video hub office. *Id.*, 3:46-49. Liu's "video hub office" (green) includes server 105, router 107, and transport links 109 that sends media content through the network to the channel buffer device 116 before being distributed to the user via the DSLAM 108. Ex-1006, 3:62-65, 4:57-66, 5:19-27. DLSAM 108 provides the I-frames, B-frames, and/or P-frames to each user. *Id.*, 5:21-22.

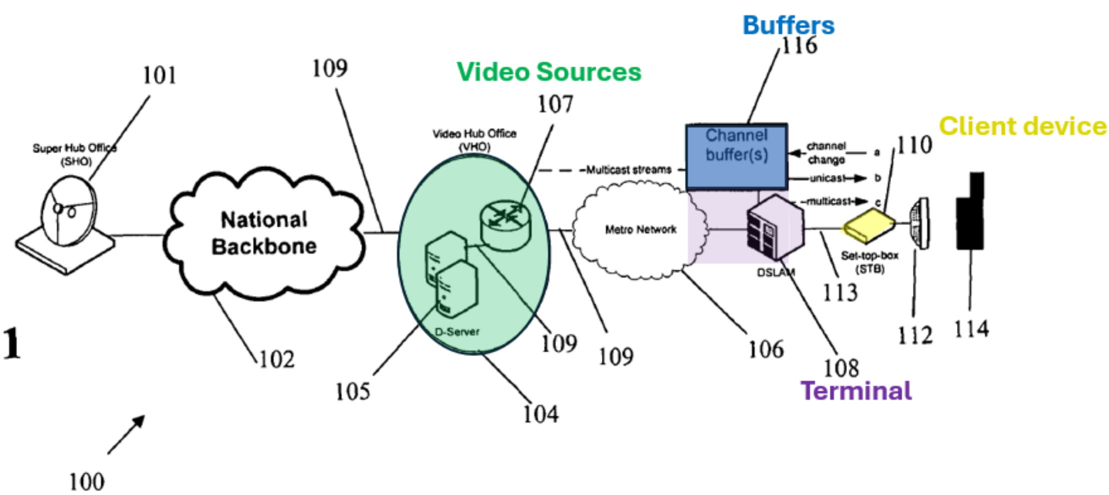


FIG. 1

78. Liu's IPTV Service Architecture 100 also comprises "channel buffer device 116" (blue), which includes "channel buffer(s)." Ex-1006, 4:1-3, Figure 1. Liu's channel buffer device 116 "include[s] memory for buffering data streams including I-frames" (*id.*) and may "utilize any suitable storing scheme." *Id.*, 4:66-5:1. The plurality of buffers "may receive and buffer the latest frames for a channel

up to the most recent 'N' I-frames, where N may be as small as 1." *Id*., 4:46-48. The channel buffer device 116 "may further include … programs or instructions that detect channel changes received via a client premise equipment, such as STB 110, a processor and computer programs or instructions that provide the stored I-frame to the DSLAM in response to a customer event, such as a change of a channel." *Id*., 4:1-7. The channel buffer device's functions may be integrated into the DSLAM. *Id*., 4:7-10.

79. Liu also teaches detecting a user's selection of a particular channel and then determine whether a stored I-frame corresponding to the channel resides within the memory. Ex-1006, 9:12-23. If an I-frame for the requested channel has been stored, Liu provides a signal based upon the stored I-frame to the user. *Id*.

C. Sherer (Ex-1007)

80. Sherer (U.S. Patent No. 7,904,581) was filed on February 23, 2006 and issued March 8, 2011. As explained to me by counsel, I understand that Sherer qualifies prior art to the '705 Patent under 35 U.S.C. §102(e).

81. Sherer teaches a "method for servicing channel change requests in a digital video network that is capable of distributing digital video content to clients via multicasting and unicasting." Ex-1007, 6:35-38, 14:62-65 (Cl. 1). As shown below in Figure 1A and Figure 3, Sherer's server system 104 (purple) includes a stream server 106 and channel change server 108 (orange), which includes a

processor 120, memory 143, stream buffer 134 (blue), and buffer controller 136. Ex-1007, 7:6-15. Sherer's server system 104 ingests the digital video content from the video source (green) and controls the distribution of the digital video content into the distribution network. Ex-1007, 4:29-31. Stream buffer 134 "buffers up to 1 group of pictures (GOP) worth of each ingested stream, where a GOP is defined as a series of multiple consecutive frames beginning with an I-frame and continuing until the last frame before the next I-frame." *Id.*, 8:60-64.

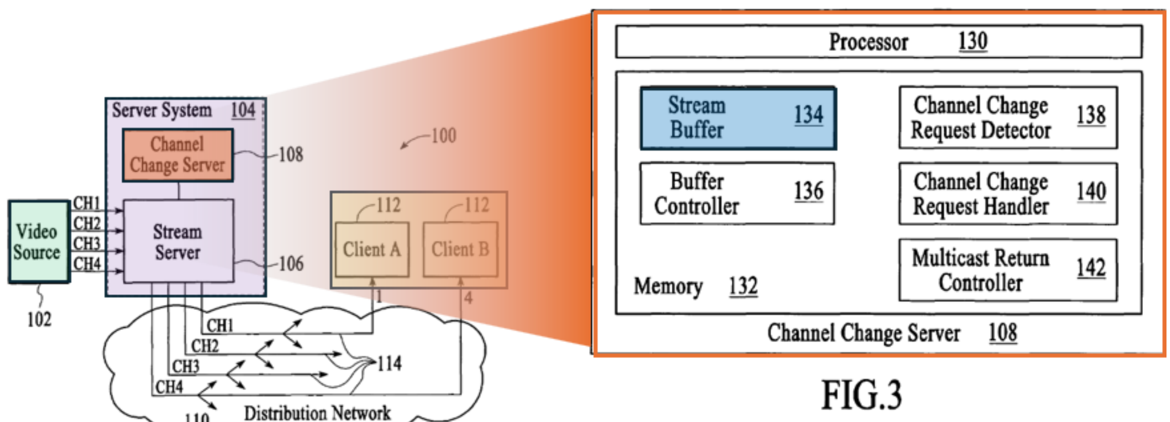


FIG.1A

FIG.3

82. As shown below in Figure 5, Sherer's streaming server "ingests" the digital video content, block 601 (green), and "portions of each channel are buffered by the channel change server for use in servicing channel change requests." *Id.*, 7:56-61. Specifically, at block 602, digital video content is provided to the client via multicasting and, at block 604 (blue), "multiple consecutive frames of digital video content, including at least one I-frame are buffered." *Id.*, 7:59-63. Once a channel request is initiated by the client device, block 608, and received by the channel

change server in the stream server, block 610 (red), "buffered frames for the requested channel are provided to the client via unicasting," block 612 (yellow). Ex-1007, 7:66-8:2. Sherer teaches that "the buffered frames are provided starting at the first buffered I-frame and the digital video content continues to be provided to the client via unicasting until a pre-established condition is met." *Id.*, 8:2-5.

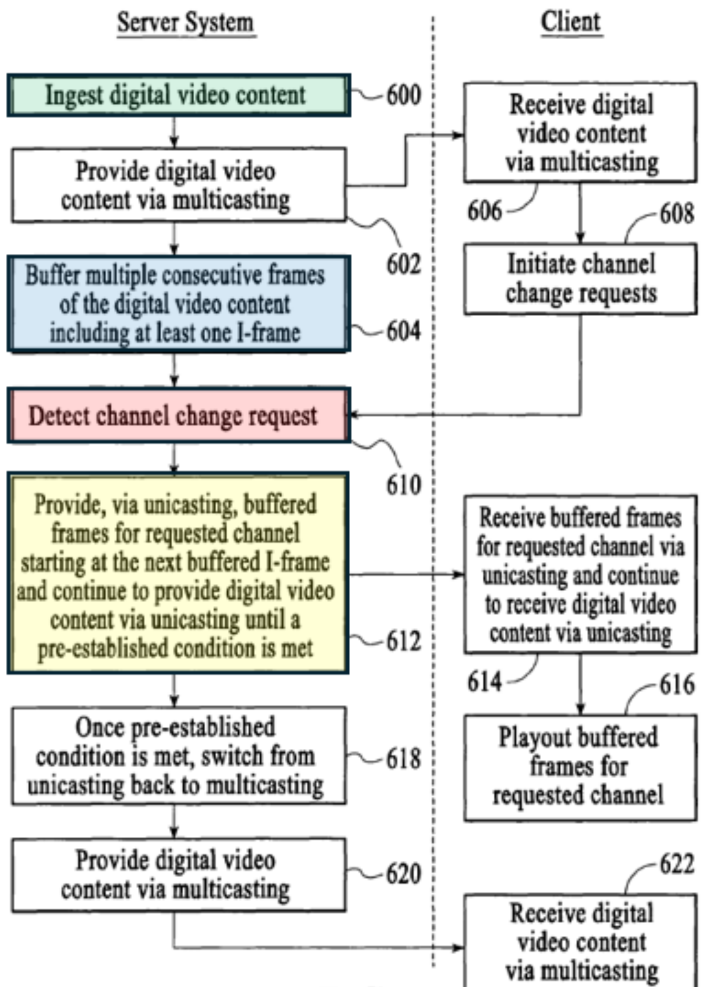


FIG.5

D. Wu-771 (Ex-1008)

83. Wu-771 (U.S. Patent No. 7,965,771) was filed February 27, 2006 and issued June 21, 2011. As explained to me by counsel, I understand that Wu-771 qualifies prior art to the '705 Patent under 35 U.S.C. §102(e).

84. Wu-771 teaches a method and apparatus for minimizing delays when displaying video for a channel change. Ex-1008, Cls. 1, 11. Wu-771 teaches that a commonly used approach for minimizing the bandwidth requirements associated with transmitting video entails compressing the video through a video compression scheme, transmitting the smaller compressed video data, and then decompressing the video data before being displayed. Ex-1008, 1:54-58. Wu-771 teaches that the Moving Pictures Expert Group (MPEG) video compression standard converts a video stream into I, P and B frames. *Id*., 1:58-62. Wu-771 further teaches that "Group of pictures is defined as the number of I, P, B frames that are sent prior to the next I frame." Ex-1008, 3:63-64. For example, a "typical Group of Pictures (GOP) that contains I, P and B frames sent" such that an I frame is followed by fourteen B and P frames. *Id*., 3:65-4:2.

85. Wu-771includes a "video acceleration module VAM" which buffers the frames. *Id*., 4:66-67. As shown in Figure 3 below, Wu-771's VAM has "separate outgoing client buffers for each client port (PBUF) 401" and "fixed buffers for each multicast stream it is sourcing to clients, referred to herein as Streaming Buffers

(StrmBuf).” Ex-1008, 4:67-5:3. Wu-771’s StrmBuf 402 buffers “fixed number of GOP’s aligning the start of the buffer to the start of a GOP.” *Id.*, 5:3-6. Wu-771 teaches that “the higher number of I frames sent, the faster the decoder can start displaying the picture (i.e., faster channel change time form a user perspective), since the decoder needs an I frame to initially sync to; again.” Ex-1008, 3:59-61.

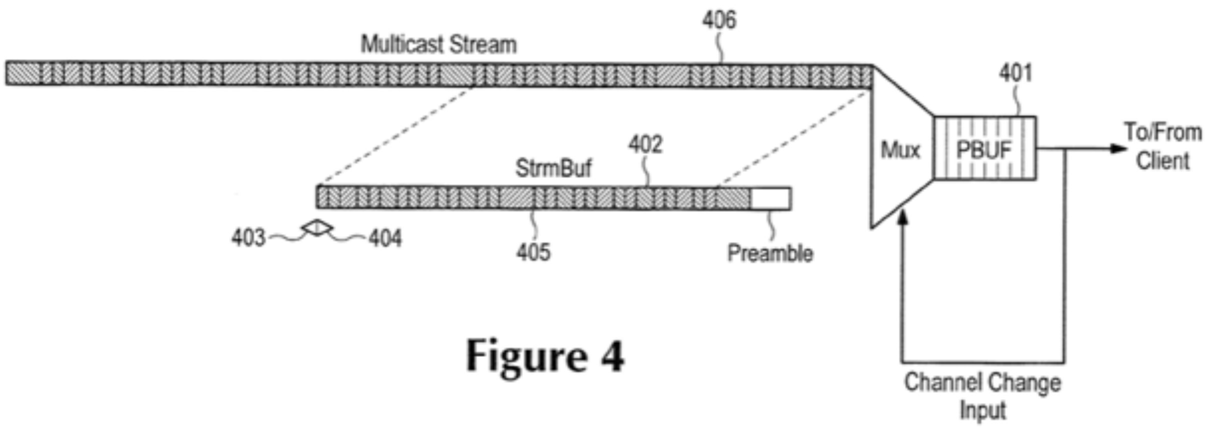


Figure 4

VII. CLAIM CONSTRUCTION

86. For the purposes of my Declaration, I do not believe that any term requires explicit construction to resolve the issues I present in my Declaration because the prior art discloses or renders obvious the claim limitations under any potential construction of the claims.

87. I understand that the Patent Owner and Petitioner in the related litigation have argued the following the claim constructions:

A. "media aggregation server" (Claims 1, 2, 5-7)

88. I understand from counsel that the parties have agreed that the "media aggregation server" recited in claims 1, 2, and 5-7 is "any mechanism for obtaining media content from a large number of sources in order to provide the media content to mobile devices in live broadcast streams." Ex-1021, at 2; Ex-1022; Ex-1001, 6:54-58.

B. "key frame" (Claims 1, 2, 5-7)

89. I understand from counsel that Petitioner has proposed that the recited "key frame" in claims 1, 2, 5-7 is "a frame that can be displayed substantially on its own." Ex-1021, 2; *see also* Ex-1001, 2:65-67. PO has proposed that the "key frame" is a "frame that can be displayed substantially on its own (i.e., without reference to other frames)." Ex-1022, 2. It is my opinion that there is no meaningful difference between these constructions as applied to the prior art references at issue in my declaration.

C. "group of pictures (GOP)" (Claims 1, 5)

90. I understand from counsel that the parties have agreed that "GOP" is a "sequence of different frame types, beginning with a key frame and ending just before a subsequent key frame." Ex-1022, 2.

D. "predictive frame" (Claim 1)

91. I understand from counsel that the parties in the related litigation have agreed that "predictive frame" is "a frame that relies on information from a reference key frame." 1021, 2; Ex-1022; Ex-1001, 2:67-3:2.

E. "wherein a plurality of buffers at the streaming server are allocated on a per group of pictures (GOP) basis" (Claim 1)

92. I understand from counsel, Petitioner has proposed, in the related litigation, that "a plurality of buffers at the streaming server are allocated on a per group of pictures (GOP) basis" means "the streaming server allocates one or more different buffers for each group of pictures (GOP)." Ex-1021, 2. PO proposed that a "plain and ordinary meaning" should apply. Ex-1022. It is my opinion that there is no meaningful difference between Petitioner's constructions is consistent with the "plain and ordinary meaning" of the terms and as applied to the prior art references at issue in my declaration.

F. "wherein the next available key frame is provided as the second key frame in response to the channel change request, wherein no predictive frames are transmitted at the beginning of the channel change request" (Claim 1)

93. I understand from counsel, Petitioner has proposed, in the related litigation, that these limitations are indefinite. *See* Ex-1021. PO has proposed that the limitations are not indefinite. *See* Ex-1022. It is my opinion that there is no meaningful difference between these constructions as applied to the prior art

references at issue in my declaration. I have applied these limitations to include, at least, that in response to a change request, the next available key frame is provided and there are not any predictive frames transmitted at the beginning of the channel change request.

G. "identifying a next available key frame associated with the second media stream" (Claim 1)

94. I understand from counsel, Petitioner has proposed, in the related litigation, that this limitation means "identifying, by the streaming server, by decoding or partially decoding the frame, a next available key frame associated with the second media stream." Ex-1021, 3. PO argues that this limitation should be given its plain and ordinary meaning. Ex-1022. It is my opinion that there is no meaningful difference between these constructions as applied to the prior art references at issue in my declaration because Petitioner's proposed construction is consistent with the "plain and ordinary meaning" of the term.

H. Order of steps (Claim 1)

95. I understand from counsel, Petitioner has proposed, in the related litigation, that elements 1[a] and 1[b] must precede elements 1[c], and 1[c], 1[d], and element 1[e] must be performed in order and PO has proposed that Claim 1's steps do not need to be performed in sequential order. *See* Exs-1021, 1022. It is my opinion that there is no meaningful difference between these constructions as applied to the

prior art references at issue in my declaration because the prior art teaches performing the steps in the order recited by the claims.

I. "wherein the first media stream is a first video stream and the second media stream is a second video stream" (Claim 2)

96. I understand from counsel, Petitioner has proposed, in the related litigation, that claim 2 is invalid under 35 U.S.C. § 112, ¶4. *See* Ex-1021. PO has proposed that the plain and ordinary meaning should apply. *See* Ex-1022. I have applied this claim to mean that the first media stream is a first video stream and the second media stream is a second video stream because the media stream, as recited in claim 1, includes a plurality of frames and therefore is a video stream.

97. I reserve the right to offer opinions on any claim constructions proposed in this proceeding

VIII. DETAILED EXPLANATION OF THE UNPATENTABILITY GROUNDS

A. Grounds 1/2: Claims 1, 2, 4-9, 11-16, and 18-20 are obvious over Mao or Mao further in view of Liu.

1. A POSITA would have been motivated to combine Mao's teachings with Liu's teachings and would have had a reasonable expectation of success.

98. A POSITA would have been motivated to combine the teachings of Mao and Liu and had a reasonable expectation of success in doing so because each

relates to providing a video stream to a client device in response to a channel change request.[2] Ex-1005, Title, Abstract; Ex-1006, Title, Abstract.

99. Mao teaches a method and apparatus that rapidly changes the channel in a digital video delivery system. Ex-1005, 2:26-28; *see also id.*, 1:7-10. Mao's digital video system includes a broadband digital terminal 12 (BDT, purple) that is "connected to a video source (ATM network 26 or special network 32) that provide a plurality of compressed video signals." Ex-1005, 7:62-64. As shown in Figure 7 below, BDT 12's processor 55 (orange) receives the frames for each channel and detects, establishes, and records the GOP start frame (I frames, yellow) for each channel. *Id.*, 8:15-26. Compressed video data of each video channel is stored in a buffer on BDT 12 (blue), where a buffer unit is used for each video channel. Ex-1005, 8:14-26, 8:37-42. BDT 12 includes a "first-in-first-out (FIFO) buffer 50 for storing the compressed video data of each video channel." Ex-1005, 8:14-16. Mao provides an example in which "forty FIFO buffers 50 will be used" when a subscriber has "forty channels from which to choose." *Id.*, 8:59-62*; see also id.*, 8:17-21 ("BDT 12 "utilizes buffers to store the video data for each channel it is

---

[2] I also note that Liu's front page lists Mao (U.S. Patent No. 6,728,965) as prior art, signaling to the POSITA that Mao addresses the same field and problem, and that its mechanisms are considered pertinent and compatible. *See* Ex-1006.

supposed to receive."), 8:39-41 ("multiple buffer memory units may be used, e.g., one for each video channel").

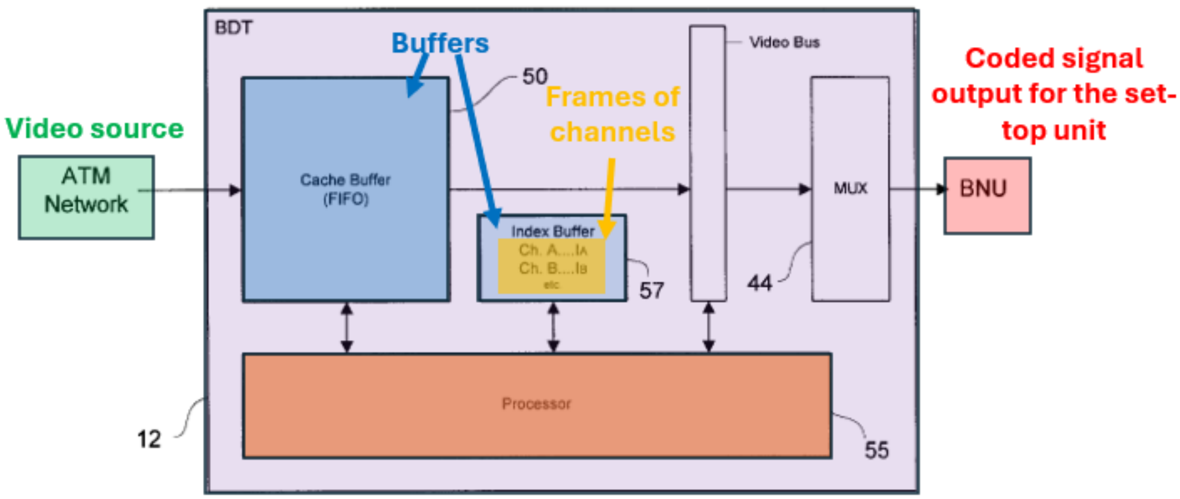


FIG. 7

100. Additionally, Liu teaches an Internet Protocol Television (IPTV) Service Architecture 100, which includes a "video hub office" (green) that includes multiple servers 105, router 107, and transport links 109 to send media content through the network to the channel buffer device 116 (blue) before being distributed to the user's set top box 110 (yellow) via the Digital Subscriber Line Access Multiplexer (DSLAM) 108 (purple). Ex-1006, 3:62-65, 4:57-66, 5:19-27.

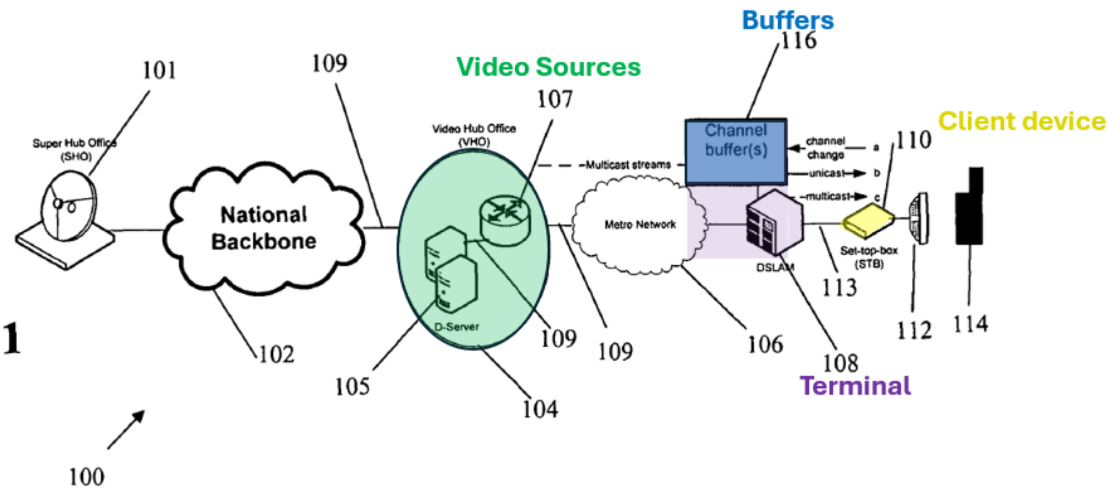


FIG. 1

101. Liu's "channel buffer device 116" (blue) includes "channel buffer(s)" (Ex-1006, 4:1-3, Figure 1) and may be included in the DSLAM. *Id.*, 4:7-10. The channel buffer device 116 "includes memory for buffering data streams including I-frames" (*id.*) and may "utilize any suitable storing scheme." *Id.*, 4:66-5:1. The plurality of buffers "***may receive and buffer the latest frames for a channel up to the most recent 'N' I-frames***, where N may be as small as 1." *Id.*, 4:46-48. The channel buffer device 116 "may further include … programs or instructions that detect channel changes received via a client premise equipment, such as STB 110, a processor and computer programs or instructions that provide the stored I-frame to the DSLAM in response to a customer event, such as a change of a channel." *Id.*, 4:1-7.

102. First, a POSITA would have been motivated to receive video streams from one or more servers, such as those in Liu's video hub office 104, in order to have the video streams from the national backbone and super hub office stored closer

to the terminal. This would, for example, minimize transmission delays in providing the requested channel to the user. While a POSITA would have understood Mao's ATM network 26 or special network 32 video sources (Ex-1005, 7:62-64) include a server, Liu provides a specific example of a well-known teaching of one or more video servers providing the video stream.

103. Second, a POSITA would have been motivated to maintain Mao's video frames for each video stream in separate buffers at its BDT on a per GOP basis, as taught by Liu. A POSITA would have understood that Mao's BDT 12 ("streaming server") allocates one or more different buffers for each group of pictures (GOP) when maintaining the starting frame (I frame) for each channel in a buffer. A POSITA would have further understood that Liu provides an example of this well-known approach. Liu teaches that each channel has its own buffer to maintain its frames according to the encoding scheme to ensure that an I frame is maintained with its proceeding predictive frames. Ex-1006, 4:35-40. For example, Liu teaches that its channel buffers store "the latest frames for a channel up to the most recent 'N' I-frames, where N may be as small as 1." *Id.*, 4:46-48.

104. A POSITA would have been motivated to store each channel as a sequence of different frame types in an allocated buffer, beginning with a key frame and ending just before a subsequent key frame, as taught by Liu. This would ensure that each channel's most recent I frames are stored and can be provided to the user

upon request, thus eliminating delays that would otherwise occur to fetch, identify, and further process the video stream to be provided to the user's device and prevent distorted videos from being presented to the user. Ex-1006, 1:42-52. Liu teaches deploying "channel buffers" at the VHO/DSLAM, implemented as circular FIFO memories that retain the most-recent I-frame per channel, detecting a user channel-change event, and unicast that stored I-frame to the user before handing off to the live multicast stream. Liu adds scalable refinements (e.g., hashing/predicted-channel selection, cluster memory shared across access elements). Those teachings map directly onto Mao's cache-and-pointer model: Liu's user-event detector supplies the trigger Mao already uses; Liu's circular FIFO/cluster memory supplies the I-frame (and subsequent frames) that Mao's architecture is poised to transmit; and both systems operate on the same I/P/B-frame and GOP constructs to solve the same latency problem. Given that Mao contemplates placing the buffers/processor at the access node (analogous to Liu's DSLAM), integrating Liu's VHO/DSLAM channel buffers into Mao's upstream switching architecture would be a straightforward substitution yielding the same channel-change function with Liu's additional benefits (edge scaling, predicted-channel readiness, and clean unicast-to-multicast handoff). A POSITA would thus expect the combination to work for its intended purpose without undue experimentation.

105. A POSITA would have had a reasonable expectation of success in combining the teachings of Mao and Liu. For example, a POSITA would have recognized that both, Mao and Liu, operate on MPEG GOPs and I frame detection and buffering, which are simple tasks in digital video systems. The combination has a reasonable expectation of success because the references discuss how the placement of its components are flexible. For example, Liu states the channel buffer functions may be integrated into the DSLAM or a similar customer facing node, which would readily host Mao's indexing logic.

106. A POSITA would have understood that the combination of Mao-Liu would yield predictable results because Mao's ATM/FTTC context doesn't limit its I frame indexing/caching to ATM because the mechanism is CODEC level (frame indexing), not transport specific. Liu's IPTV applies the same indexing. A POSITA would therefore expect the combination to work without undue experimentation.

107. Additionally, a POSITA would have had a clear, reasonable expectation of success in combining Mao's rapid-channel-change architecture with Liu's fast-channel-change mechanisms because Mao already teaches the core engine: per-channel FIFO caching at the headend/access node and continual indexing of the MPEG-2 synchronization (I-) frame so that, upon a channel-change request, the system immediately starts delivery at the indexed I-frame, thereby eliminating the wait for the next GOP start and cutting the dominant portion of

channel-change latency. Mao also expressly allows the buffer/processor to reside downstream (e.g., in the access node), underscoring that the location is an implementation choice rather than a constraint. As I modify below in Mao's Figure 7 ("Modified Figure 7'"), Liu's VHO and channel buffers would have been implemented in conjunction with Mao's architecture to carry out the same functions that Mao is already performing, while achieving the additional benefits taught by Liu, which was a well-documented advantage known by POSITAs (as I described in Sections V.A and V.C). *See, e.g.*, Ex-1006, 1:51-53 ("provide the video content faster to the customers").

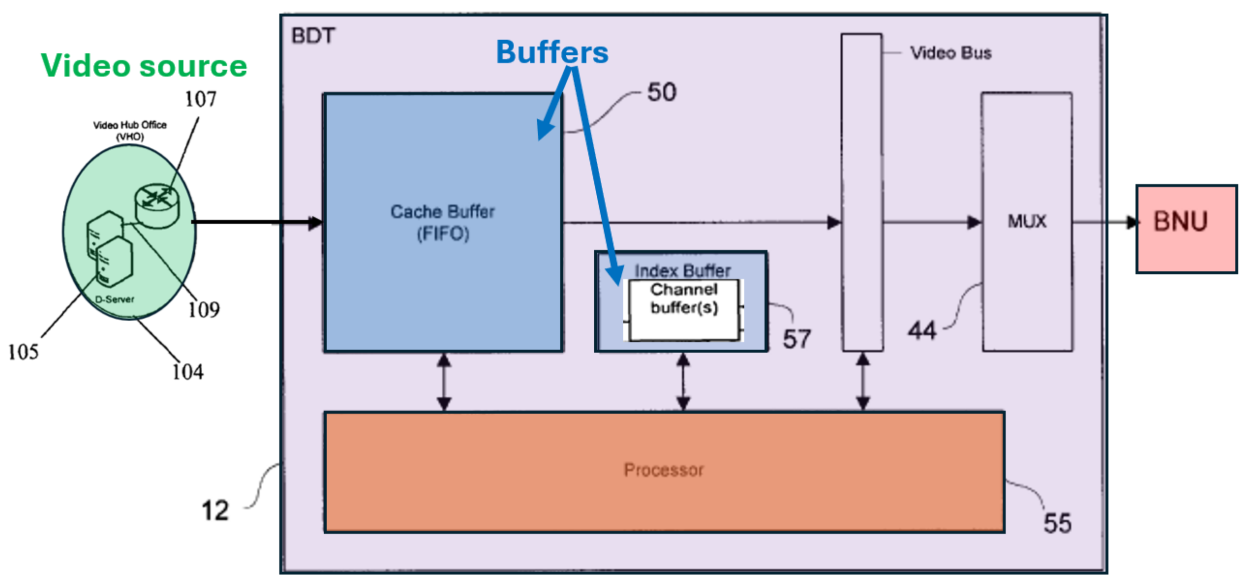


FIG. 7'

108. Additionally, both references describe methods for receiving video data streams, identifying and storing the I-frames and predictive frames, and providing

the channel requested by the user using similar hardware and software components. Ex-1005, Title, Abstract; Ex-1006, Title, Abstract. The combination is simple. Mao teaches how to find the right I frame fast, and Liu teaches where to store the I frame. Mao contributes a concrete indexing/caching mechanism: a FIFO cache buffer plus an index buffer that tracks I frame locations across channels so the system can "immediately synchronize" for the requested channel. That is directly reusable at any network node handling channel change requests. Liu contributes the system placement and switchover logic: maintain per channel I frame buffers (optionally the last N I frames) at a DSLAM/channel buffer device, send the cached I frame by unicast on a request, and switch to multicast when unicast and multicast are aligned. Liu also notes these buffer functions may be integrated into the DSLAM or another interface device, which is exactly where Mao's I frame indexing can occur.

109. A POSITA would thus have understood that these references disclose interrelated teachings based on well-understood technologies that would have been amenable to various well-understood and predictable combinations, which combinations achieve the benefits taught by each reference. For example, both references solve the same bottleneck with the same mechanism (supply an I frame from a buffer), a POSITA would have been naturally motivated to combine them. Mao explains that much of the channel change delay in switched digital video comes from waiting to synchronize to the next I frame; it proposes caching video and

"detecting and pointing to the synchronization [I] frames" so a change request can be served immediately without waiting for the next GOP boundary. Liu describes the identical problem in IPTV: even if a stream is delivered immediately, a viewer "may still have to wait … up to 5 seconds until the next I frame," while typical TV expectations are 0.5–2 seconds. Liu's solution is to buffer recent I frames per channel at the edge (e.g., the DSLAM) and send one via unicast on a change, then switch back to multicast once streams are in sync. Both references emphasize the need to decrease time delays, such as less than 2 seconds. *See, e.g.*, Ex-1005, 7:13-15, 7:44-51, 8:4-13, 8;48-51; Ex-1006, 1:45-50, 4:40-45. A POSITA would have been motivated to combine Mao's and Liu's teachings with a reasonable expectation of success because combining them yields exactly the predictable improvement a POSITA would expect-faster first picture with resource savvy return to multicast-without inventing new codecs or protocols.

2. Independent Claims 1 and 15

a. Element 1[pre]: A method:

Element 15[pre]: A non-transitory computer readable medium, comprising:

110. **<u>Ground 1:</u>** Mao teaches the preambles of claims 1 and 15 to the extent that they are limiting. Mao teaches a "method and apparatus that rapidly changing the channel in a digital video delivery system." Ex-1005, 2:26-28; *see also id.*, 1:7-10. Mao's method, which is executed by a computer processor, would have been

implemented using non-transitory computer readable medium that provide the instructions executed by the processor. For example, Mao teaches that its broadband access system is "for delivery of video, data (for a computer network interface) and telephony services." *Id*., 3:18-21. Mao also explains that its "BDT 12 can be equipped with an element management system (EMS) 13 that provisions services and equipment, and for handling certain video signal controls on the digital broadband access system. The EMS 13 is usually software based and can be implemented on either a personal computer or a workstation. A personal computer based EMS can support one BDT 12 and its associated broad access network equipment. A workstation can support multiple BDT's 12 and their respective access network equipment." *Id*., 3:33-42. Mao thus provides examples of its method being run on a computer, which would be implemented using computer code via a non-transitory computer readable medium. Accordingly, Mao's system includes a computer-readable media (i.e., computer code instructions) that instructs the terminal's processor to perform the recited method.

111. **Ground 2:** Mao, in view of Liu ("Mao-Liu"), teaches the preamble of claim 15, to the extent it is limiting. Liu expressly discloses an example of such a well-known teaching by explaining that "a computer readable medium containing instructions operative to be executed in a computer containing a processor." Ex-1006, 7:28-31; *see also id*., 10:60-64 (claiming a "non-transitory computer readable

medium including processor executable program instructions that, when executed by the processor, cause the processor to perform operations”). For example, the instructions may include, “at least one instruction 310 that can cause the processor to buffer an I-frame in a network from a server to a user, possibly including at least one instruction 320 that can cause the processor to circularly FIFO buffer a most recent I-frame for each channel of the group of channels. The instructions may also include at least one instruction that can cause the processor to buffer an I-frame for each channel in a group of channels.” *Id.*, 7:31-39. A POSITA thus would have implemented Mao’s processor to execute appropriate computer readable instructions, as taught by Liu.

b. Elements 1[A]/15[A]: [(i)] [receiving]/[computer code for receiving] a first media stream including a first plurality of frames, [(ii)] the first media stream received at a streaming server and maintained in a first buffer, the streaming server configured to receive the first media stream from a media aggregation server, [(iii)] wherein a plurality of buffers at the streaming server are allocated on a per group of pictures (GOP) basis;

112. **<u>Ground 1:</u>** Mao teaches elements 1[A] and 15[A].

113. Mao teaches “receiving a first media stream including a first plurality of frames” as claimed. For example, Mao teaches a “digital video system receiving video data corresponding to a plurality of video channels.” Ex-1005, 10:8-9 (Cl. 5). Mao’s broadband digital terminal (BDT 12) receives “a plurality of compressed

video signals" from a video service provider. Ex-1005, 7:62-64, 3:38-42. BDT 12 receives this media stream that is comprised of "continuous stream of a predetermined number of consecutive data frames." *Id*., 10:12-13 (Cl. 5). Mao explains that the received media stream includes a plurality of frames by being grouped into Groups of Picture (GOPs), starting with an "I" or "intracoded frames" followed by "predicted frames and bidirectional frames (P and B frames, respectively)." *Id*., 7:29-31. Mao teaches that "[v]irtually all of the compression schemes require a starting or synchronization frame usually referred to as the Group of Picture (GOP) start point. In most compression schemes the GOP start point is found proximate the beginning of the information field." *Id*., 7:2-7.

114. Mao's I frame are key frames, because, as I explained above in Section V.B, I frames were well known to be "key frames." For example, Wu-771 explains how an "I frame is an intra-frame which contains all of the data required to display the frame," which would have been understood as a key frame. Ex-1008, 1:58-2:9. Barton also described how its I-frame is an "index" frame which starts the GOP. Ex-1012 ¶12. Liu also explicitly describes how "I frames or 'infraframes,'" are "also known as '***key frames***' or 'index frames'" and "are independently decodable, and store all the data required to display a frame." Ex-1006, 4:14-18. Choi also discusses that the "key frame is a complete picture used for video compression. Referring to video group of pictures (GOP) structure, a frame is selected one by one at a constant

interval from a temporal video flow, and is designated as the key frame. The key frame is an independently restorable video, and enables an arbitrary video access." Ex-1011, 2:55-60. In particular, "MPEG, H.26a, and H.264, the key frame is inserted at a constant interval into a GOP, and designates the I fame enabling independent video production." *Id*., 2:61-63. Indeed, the '705 patent itself explains that ("MPEG and its various incarnations use I-frames as key frames" (Ex-1001, 3:3-7) and specifically claims "wherein the key frame is an I-frame." *Id*., 9:58-59.

115. In Mao, "the first media stream received at a streaming server [is] maintained in a first buffer." Mao's BDT is a streaming server with a plurality of buffers. As shown in Figure 7 below, Mao's BDT 12 includes a "first-in-first-out (FIFO) buffer 50 for storing the compressed video data of each video channel." Ex-1005, 8:14-16. Mao provides an example that "forty FIFO buffers 50 will be used" when a subscriber has "forty channels from which to choose." *Id*., 8:59-62.

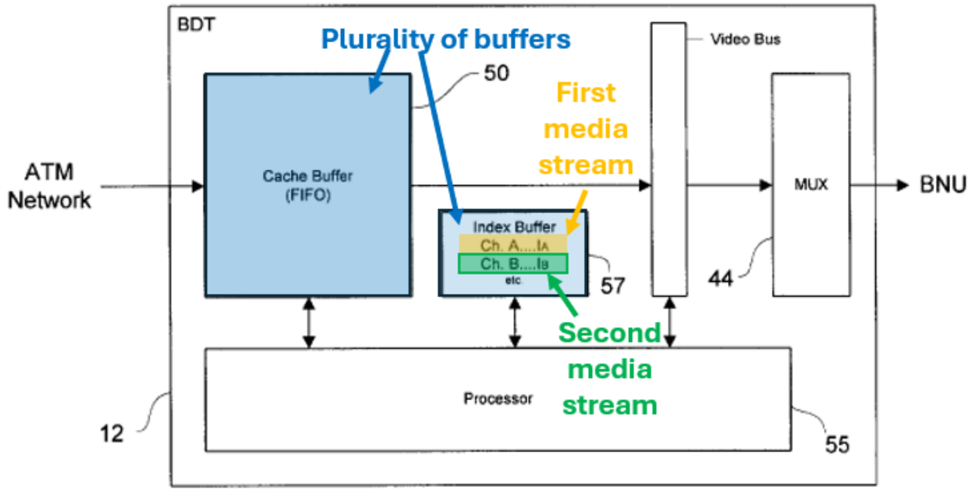


FIG. 7

116. Mao's BDT is streaming server because it includes a processor 55, memory, buffers, and interface operable to receive media streams, such as video streams, and a client device interface that provides the channels to the client device. Ex-1005, 3:34-42, 8:14-27. Indeed, this is just like the '705 patent, which explains that its streaming server includes:

> a processor 401, memory 403, buffers 431, 433, 435, and 437, and a number of interfaces. In some examples, the interfaces include an encoder interface 411, a media aggregation server interface 413, and a client device interface 441. The encoder interface 411 and the media aggregation server interface 413 are operable to receive media streams such as video streams. In some examples, hundreds of video streams associated with hundreds of channels are continuously being received and maintained in buffers 431, 433, 435, and 437 before being provided to client devices through client device interface 441.

Ex-1001, 7:58-8:2.

117. Mao's streaming server is "configured to receive the first media stream from a media aggregation server." As discussed in Section VII, above, a "media aggregation server" is defined in the specification as "any mechanism for obtaining media content from a large number of sources in order to provide the media content to mobile devices in live broadcast streams is referred to herein as a media content aggregation server." Ex-1001, 6:54-58. Mao's BDT 12 (the streaming server) is "connected to a video source ([Asynchronous Transfer Mode (ATM)] network 26 or special network 32) that ***provide a plurality of compressed video signals***." Ex-1005, 7:62-64. Mao's ATM or special network is a "media aggregation server" because it is the "mechanism for obtaining media content from a large number of sources in order to provide the media content … in live broadcast streams" as defined in the specification. Ex-1001, 6:54-58.

118. Additionally, Mao's "plurality of buffers at the streaming server are allocated on a per group of pictures (GOP) basis." As discussed in Section VII, above, Petitioner's proposed construction is "wherein the streaming server allocates one or more different buffers for each group of pictures (GOP)." Ex-1021, 2. "GOP" is a "sequence of different frame types, beginning with a key frame and ending just before a subsequent key frame."

119. Mao's BDT 12 stores the pointer for each GOP start frame (I frame) for each channel "in a second buffer 57." Ex-1005, 8:22-26. The GOP starting point is

coded as "I" or "intracoded frames," which are then followed by "predicted frames and bidirectional frames (P and B frames, respectively)." *Id*., 7:29-31. While Mao's Figure 1 shows a single FIFO and second buffer 57, Mao teaches that BDT 12 "***utilizes buffers*** to store the video data ***for each channel*** it is supposed to receive." *Id*., 8:17-21. Indeed, "***multiple buffer memory units may be used***, e.g., one for each video channel." *Id*., 8:39-31. Mao thus teaches one or more different buffers for the multiple channels received on a GOP basis because Mao teaches that, based on the minimum number of frames in the compression scheme, "***at least one GOP frame for each video channel*** is stored at all times in the buffer." *Id*., 8:29-33. For example, Mao teaches that its buffer 50 "stores, at a minimum, fifteen frames (i.e., one I frame and fourteen P or B frames) for each channel that is received by the BDT." *Id*., 8:37-42. Moreover, as shown in Figure 5, as new video information is received, previous frames are counted out so that "at least one GOP frame for each video channel is stored." Ex-1005, 8:28-33. Specifically, as one I frame leaves the buffer, a new I frame must enter the buffer. *Id*., 7:29-31, 8:33-36.

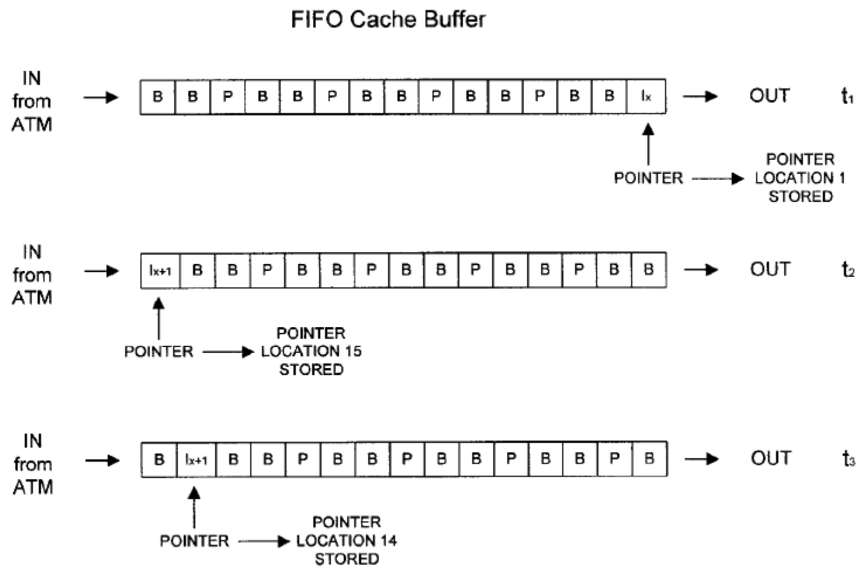


FIG. 5

120. This is similar to how the ’705 patent explains its streaming server’s buffer is allocated on a GOP basis. For example, the ’705 patent explains that “a streaming server includes a first in first out (FIFO) buffer per channel of video received.” Ex-1001, 3:9-11. Therefore, Mao’s BDT 12 (“streaming server”) allocates one or more different buffers for each group of pictures (GOP) by maintaining a sequence of a frames starting with the pointer starting frame (I frame) followed by other frames (B and P frames) for each channel in a buffer, even under PO’s proposed construction.[3]

121. **Ground 2:** Mao-Liu teaches elements 1[A] and 15[A].

---

[3] As discussed in Section VII, I understand that PO proposed the plain and ordinary meaning for the streaming server’s plurality of buffers and that the GOP is a “sequence of different frame types, beginning with a key frame and ending just before a subsequent key frame.” Ex-1022.

122. While Mao teaches implementing buffers on a per GOP basis as described in the preceding section, Liu teaches further implementation details relating to allocating buffers on a per GOP basis. A POSITA would have been motivated to implement Mao's BDT's multiple buffers to be allocated on a per group of pictures (GOP) basis as taught by Liu. As shown in Figure 1 below, Liu teaches "channel buffer device 116," which includes "channel buffer(s)." Ex-1006, 4:1-3, Figure 1. Liu's channel buffer device 116 "includes memory for buffering data streams including I-frames" (*id.*) and may "utilize any suitable storing scheme." *Id.*, 4:66-5:1. The plurality of buffers "may receive and buffer the latest frames for a channel up to the most recent 'N' I-frames, where N may be as small as 1." *Id.*, 4:46-48. In the example where N=1, Liu's buffer would thus include a single GOP (a single I-frame and associated P and B frames).

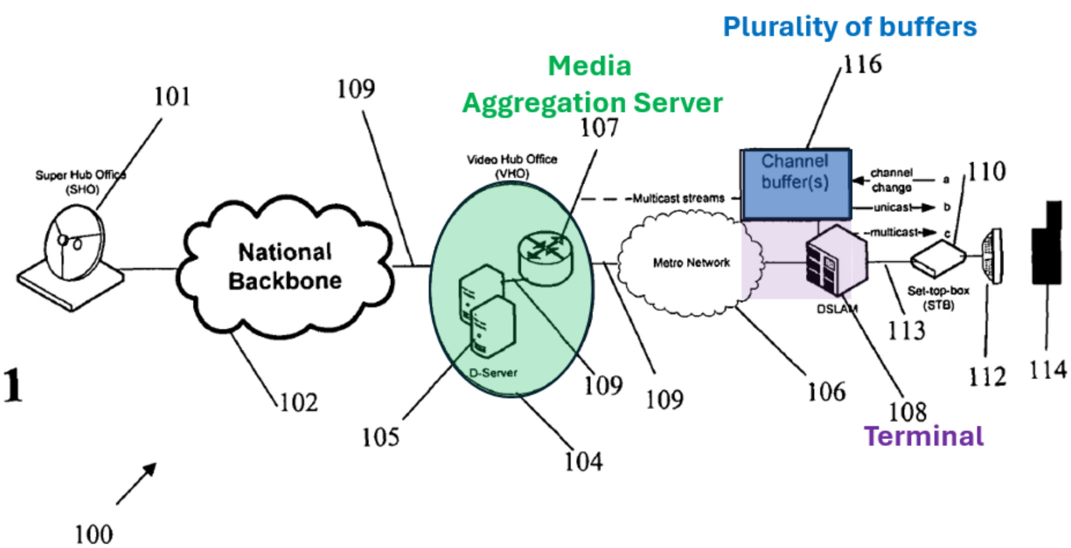


FIG. 1

123. While Mao's streaming server receives video signals from a media aggregation network, which would have been understood as a "media aggregation server" as defined in the specification as explained in the prior section, Liu explicitly teaches that the video streams are provided by one or more "servers." *See, e.g.*, Ex-1001, 6:54-58. Liu teaches a "video hub office" that includes servers 105, router 107, and transport links 109 to send media content through the network to channel buffer device 116 before being distributed to the user via the DSLAM 108. Ex-1006, 3:62-65, 4:57-66, 5:19-27.

124. The Mao-Liu combined system would thus have been implemented as shown in Figure 7' below. Liu's VHO's servers provide the video streams to Mao's BDT just like Mao's video sources, Asynchronous Transfer Mode network 26 or special network 32, provide a plurality of compressed video signals. Ex-1005, 7:62-

64.

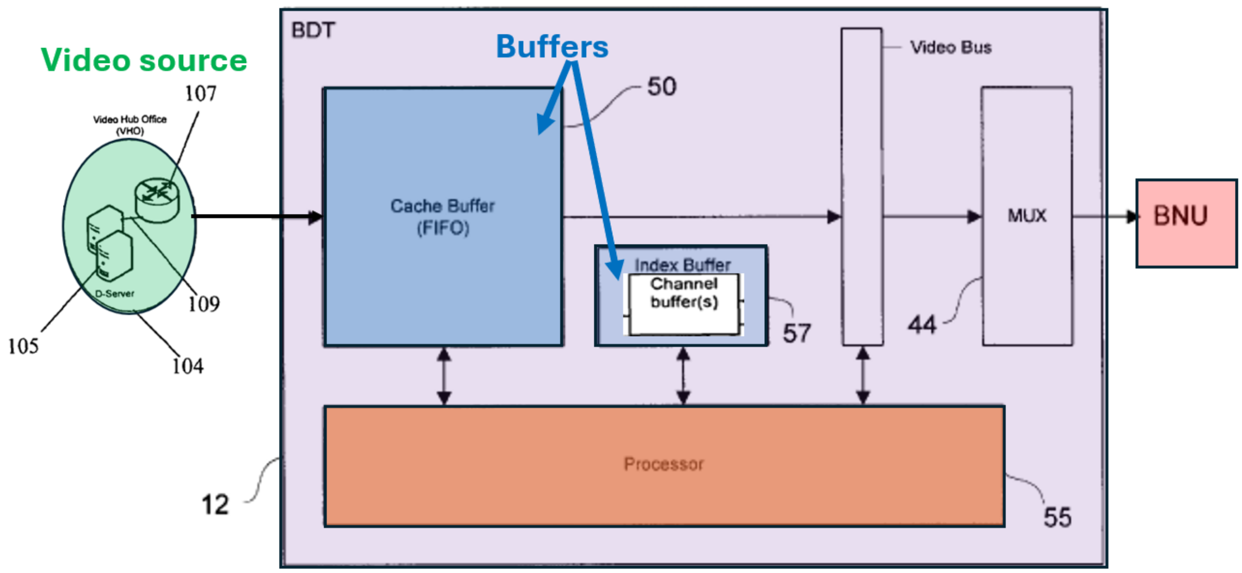


FIG. 7'

125. Additionally, Liu's channel buffers would have been implemented on Mao's index buffer 57 as to maintain Mao's I frame for each channel, including a single GOP in some examples (e.g., when N=1) as taught by Liu. *Id*., 8:17-21, 8:39-31.

126. Therefore, Liu's VHO and channel buffers would have been implemented in conjunction with Mao's architecture to carry out the same functions that Mao is already performing, while achieving the additional benefits taught by Liu. Specifically, in practice, that server side, key frame aware buffer lets the system begin every video stream with a ready I frame, eliminating the need for clients to drop unusable P/B frames and cutting the black screen interval that typically occurs

when joins happen mid-GOP, thereby reducing channel change latency and improving perceived quality. This buffer also bounds startup to (at most) one GOP while allowing immediate picture display followed, where needed on constrained last mile links, by a short, server managed catch up that can safely drop or pad non key frames to re synchronize without artifacts. Because the buffered, per channel feeds sit behind Mao's BDT, the system can swap among sources (e.g., alternate camera angles or redundant encodes) seamlessly by continuing to play from one buffer while the next is acquired and primed, yielding stream switches that are uninterrupted to the viewer. And at a network scale, Liu's VHO servers enable a fast join/return to multicast pattern, briefly unicasting the buffered key frame sequence to get the viewer on screen quickly, then opportunistically switching back to multicast once a simple condition is met. This reduces IGMP churn and uses bandwidth more efficiently without changing Mao's downstream behavior. Indeed, this was well known as I described above in Sections V.A and V.C. Therefore, Liu's VHO and channel buffers could have been implemented in conjunction with Mao's architecture to carry out the same functions that Mao is already performing, while achieving these additional benefits.

127. A POSITA would have been motivated, and have had a reasonable expectation of success, to implement Mao-Liu this way as explained in Section VIII.A.1 (motivation).

c. Elements 1[B]/15[B]: [receiving]/[computer code for receiving] a second media stream including a second plurality of frames, the second media stream received at the streaming server and maintained in a second buffer;

128. **Ground 1**: Mao teaches elements 1[B] and 15[B].

129. Mao teaches "receiving a second media stream including a second plurality of frames," which is "maintained in a second buffer." As explained in the previous section, Mao's BDT 112 (streaming server) receives video signals for multiple video channels (e.g., Ch. A, Ch. B., etc.). Ex-1005, 7:62-64, 3:38-42, 10:8-9 (Cl. 5). BDT 12's processor 55 receives the frames for each channel and detects, establishes, and records the GOP start frame (I frame) for each channel. Specifically, "processor 55 detects the GOP start frame (or I frame in the preferred embodiment) for each channel received and establishes a pointer for each GOP start frame. The processor 55 then records the I frame pointer location for each channel in a second buffer 57." *Id.*, 8:15-26. Compressed video data of each video channel is stored in a buffer on BDT 12, where each buffer unit is used for each video channel. Ex-1005, 8:14-26, 8:37-42. The second video stream made of a plurality of frames (green) is therefore maintained in a second buffer specific for that video channel, because as I

explained in the previous section, Mao's BDT "utilizes ***buffers*** to store the video data for each channel." *Id.*, 8:14-26. As I described above in Section V.C. multiple streaming buffers were commonly used in video streaming systems and methods.

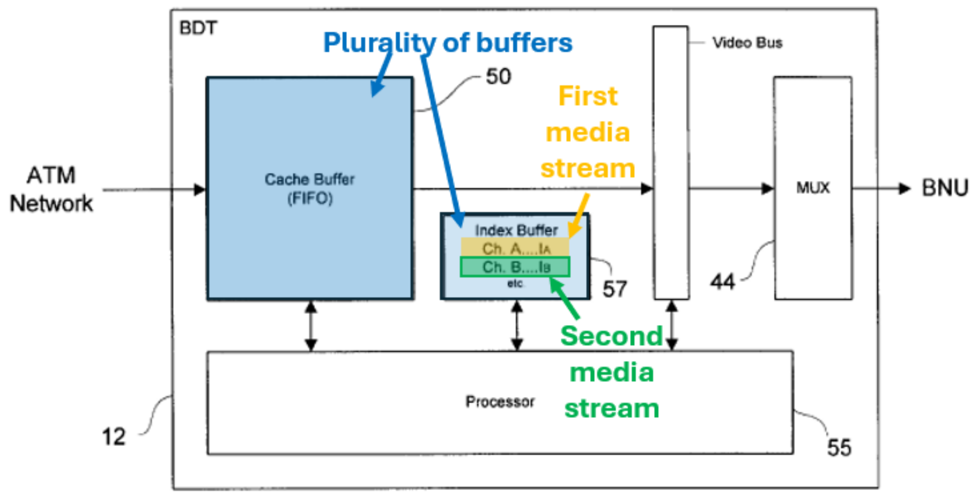


FIG. 7

130. **Ground 2:** Mao-Liu teaches elements 1[B] and 15[B].

131. While Mao teaches implementing buffers for each video stream (Ex-1005, 8:14-26), Liu teaches further implementation details relating to storing a second stream in a second set of buffers. Liu teaches to use "channel buffers" to store video streams for different channels. Ex-1006, 4:1-3, Figure 1. A POSITA would have been motivated to maintain Mao's second video stream, including its plurality of frames, in a separate channel buffer, as taught by Liu, as explained in Section VIII.A.1.

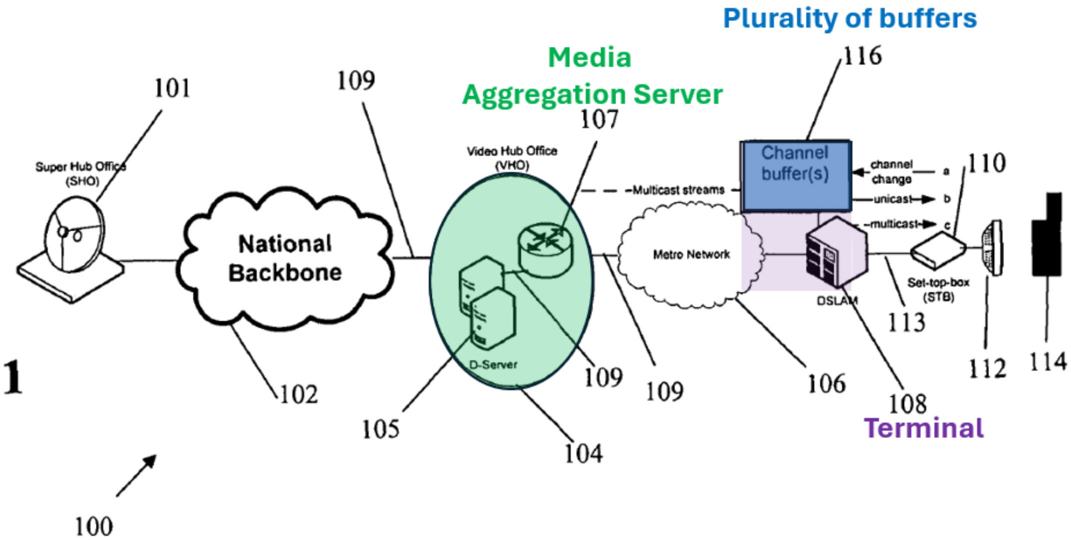


FIG. 1

d. **Elements 1[C]/15[C]: [(i)] [determining]/[computer code for determining] positions of a plurality of key frames in the first buffer and the second buffer, the plurality of key frames including a first key frame in the first buffer and a second key frame in the second buffer, [(ii)] wherein determining positions of the plurality of key frames includes determining at the streaming server a group of pictures (GOP) size N and identifying Nth frames following the first key frame as subsequent key frames;**

132. **Ground 1:** Mao teaches Elements 1[C] and 15[C].

133. Mao determines "positions of a plurality of key frames in the first and second buffer." As discussed in Section VIII, the "key frame" is "a frame that can be displayed substantially on its own." Ex-1001, 2:65-67. Mao's BDT 12 processor 55 receives the frames for each channel and detects, establishes, and records the GOP start frame (I frame) for each channel. Ex-1005, 8:15-26. The GOP start frame

(I frame) corresponds to the claimed "key frame" because it is the starting pointer of each video, "I frame" is a key frame because I frames were well known key frames. For example, Liu, as I discussed above, explicitly states: "I frames or 'infraframes,' (also known as '***key frames***' or 'index frames.') are independently decodable, and store all the data required to display a frame." Ex-1006, 4:14-18. For example, Wu-771 explains how an "I frame is an intra-frame which contains all of the data required to display the frame," which would have been understood as a key frame. Ex-1008, 1:58-2:9. Barton also described how its I-frame is an "index" frame which starts the GOP. Ex-1012 ¶12. As explained above, Liu explicitly describes how "I frames or 'intraframes,'" are "also known as '***key frames***' or 'index frames'" and "are independently decodable, and store all the data required to display a frame." Ex-1006, 4:14-18. Choi also discusses that the "key frame is a complete picture used for video compression. Referring to video group of pictures (GOP) structure, a frame is selected one by one at a constant interval from a temporal video flow, and is designated as the key frame. The key frame is an independently restorable video, and enables an arbitrary video access." Ex-1011, 2:55-60. In particular, "MPEG, H.26a, and H.264, the key frame is inserted at a constant interval into a GOP, and designates the I fame enabling independent video production." *Id.*, 2:61-63; *see also* Ex-1001, 3:3-7 ("MPEG and its various incarnations use I-frames as key frames"), 9:58-59 (Cl. 6) ("wherein the key frame is an I-frame.").

134. Mao stores a first key frame in the first buffer and a second key frame in the second buffer. For example, Mao teaches that "the minimum number of frames in the compression scheme" is held "so that at least one GOP frame for each video channel is stored at all times in the buffer." Ex-1005, 8:29-33. In a MPEG-2 scheme, fifteen frames are maintained in the buffer: one synchronization frame (I frame) and fourteen predictive or bidirectional frames (P or B) for each channel are received. *Id.*, 8:37-39. Since each of the stored GOPs have a key frame, and each GOP is stored in its own buffer, then each buffer will similarly have a first and second key frame, associated with the first and second GOPs stored in the first and second buffers, respectively.

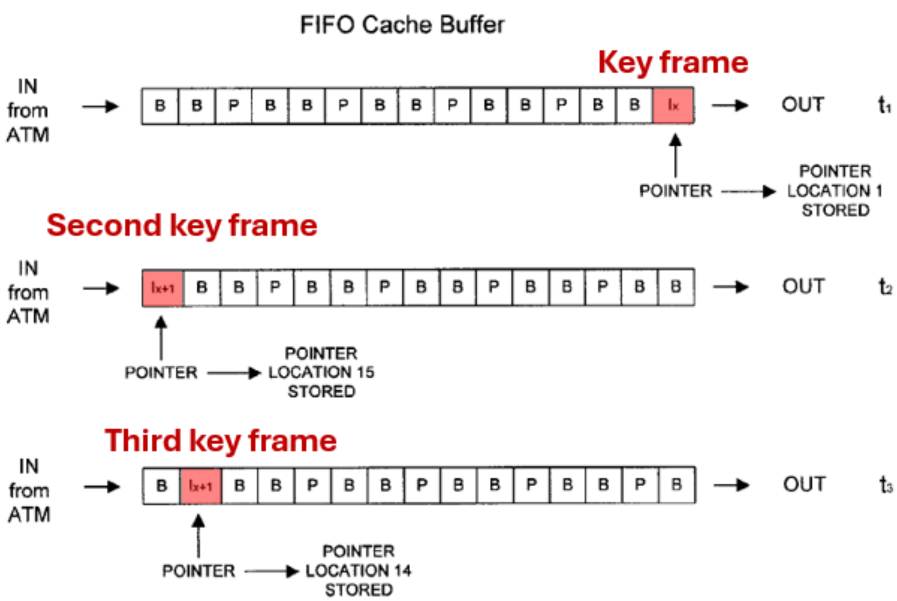


FIG. 5

135. Mao further teaches determining "a group of pictures (GOP) size N and identifying Nth frames following the first key frame as subsequent key frames." Mao teaches that two or more synchronization frames per video signal can be determined and maintained. Ex-1005, 9:2-6. Mao explains that in a MPEG-2 scheme with a GOP size of fifteen, up to thirty frames can be stored. *Id*. Each I frame is separated from the next I frame "by a pre-determined number of other frames." *Id*., 7:33-38. For example, as shown in Figure 5, the position of the first key frame pointer (I frames) and the subsequent key frames are determined by determining that the next pointer is fourteen frames, or ½ second in a MPEG-2 scheme, from the previous key frame pointer (I frame). Ex-1005, 8:37-39, 7:44-45. As I described above in Section V.B, video compression schemes, such as MPEG, used a GOP size of 15.

136. **Ground 2:** Mao-Liu teaches Elements 1[C] and 15[C].[4]

137. While Mao teaches that its processor determines the positions of a plurality of key frames, determines a GOP size N, and identifies Nth frames following the first key frame, as described in the preceding section, Liu teaches further implementation details relating to identifying key frames as a frame that can be displayed substantially on its own and determining a GOP size N and Nth frames.

[4] As explained above, Liu teaches "computer code" because Liu teaches "a computer readable medium containing instructions operative to be executed in a computer containing a processor." Ex-1006, 7:28-31.

For example, Liu teaches that its channel buffers 116 "may receive and buffer the latest frames for a channel up to the most recent 'N' I-frames." Ex-1006, 4:46-48. Liu describes that "I frames or 'infraframes,' (also known as 'key frames' or 'index frames') are independently decodable, and store all the data required to display a frame." Ex-1006, 4:14-18. Liu further teaches that N may be as small as one, such that a single buffer stores a single GOP. Ex-1007, 8:17-21, 8:39-31. A POSITA would implement Mao-Liu for example by (i) using Mao's processor indexed I frame pointers to determine the positions of key frames for each channel and to recognize the GOP interval (e.g., the N frame separation between I frames described by Mao), and (ii) adopting Liu's explicit key frame definition and N I frame buffering so that the channel buffer stores at least the most recent GOP boundary (N=1 I frame) and, where desirable, multiple GOPs (N>1), ensuring that the system can identify successive key frames at the N frame cadence and start playout from a self contained frame on every change.

138. A POSITA would have been motivated, and have had a reasonable expectation of success, to implement Mao-Liu this way as explained in Section VIII.A.1 (motivation).

e. Elements 1[D]/15[D]: [receiving]/[computer code for receiving] a channel change request from a client device receiving the first media stream, the channel change request corresponding to a switch from the first media stream to the second media stream;

139. **Ground 1:** Mao teaches Elements 1[D] and 15[D].

140. Mao teaches "receiving a channel change request from a client device receiving the first media stream… to [] switch from the first media stream to the second media stream." Mao's "rapid channel changer" receives a "subscriber's channel change request" and assesses and directs the corresponding video signal to the subscriber. Ex-1005, Abstract. When a subscriber requests a channel change, for example to channel X, using its set-top unit 19 (the client device), the channel change request is sent to the BDT over the broadband access system as shown in Figure 2A. Ex-1005, 5:51-57. For example, a subscriber's channel change request to channel X is a channel change request corresponding to a switch from the first media steam (a channel) to the second media stream (channel X).

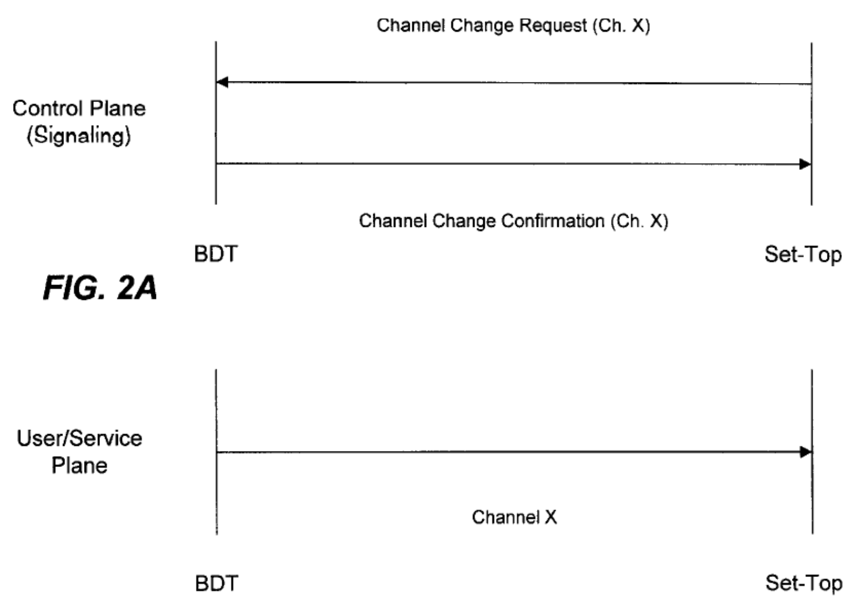


FIG. 2A

FIG. 2B

141. Mao further teaches the recited order of steps. As discussed in Section VIII, Petitioner in the related litigation proposed that elements 1[A] and 1[B] must precede 1[C], and 1[C], 1[D], and 1[E] must be performed in order. Ex-1021. Mao receives media streams (elements 1[A]-1[B]/15[A]-15[B]) and then determines the positions of a plurality of key frames in the first buffer and the second buffer (elements 1[C]/15[C]) prior to receiving a channel change request (elements 1[D]/15[D]). Mao teaches receiving video streams (Ex-1005, 7:62-64, 3:38-42), and as new video information is received, previous frames are counted out so that "at least one GOP frame for each video channel is stored." *Id*., 8:28-33. Specifically, as one I frame leaves the buffer, a new I frame must enter the buffer. *Id*., 7:29-31, 8:33-36. Thus, when a channel change is requested, "the requested signal can instantly be transmitted downstream to the subscriber since the processor 55 is always 'pointing' to an I frame for each respective channel." *Id*., 8:43-47. Therefore, Mao receives the video stream, identifies the I frames, and maintains the I frames in a buffer prior to receiving a channel change request.

142. **Ground 2:** Mao-Liu teaches Elements 1[D] and 15[D].[5]

[5] As explained above, Liu teaches "computer code" because Liu teaches "a computer readable medium containing instructions operative to be executed in a computer containing a processor." Ex-1006, 7:28-31.

143. While Mao teaches receiving a channel change request from a client device to switch from the first media stream to the second media stream, Liu teaches further implementation details relating to channel changes from one channel to another. A POSITA would have been motivated to implement Mao's channel change request as taught by Liu. Liu's channel buffer device 116 "include[s] programs or instructions that detect channel changes received via a client premise equipment, such as STB 110, a processor and computer programs or instructions that provide the stored I-frame to the DSLAM in response to a customer event, such as a change of a channel." Ex-1006, 4:1-7. As shown in Figure 1, Liu explicitly teaches that the video streams are received, the I frames for each channel are identified and determined, and maintained in a buffer, before receiving a channel request.

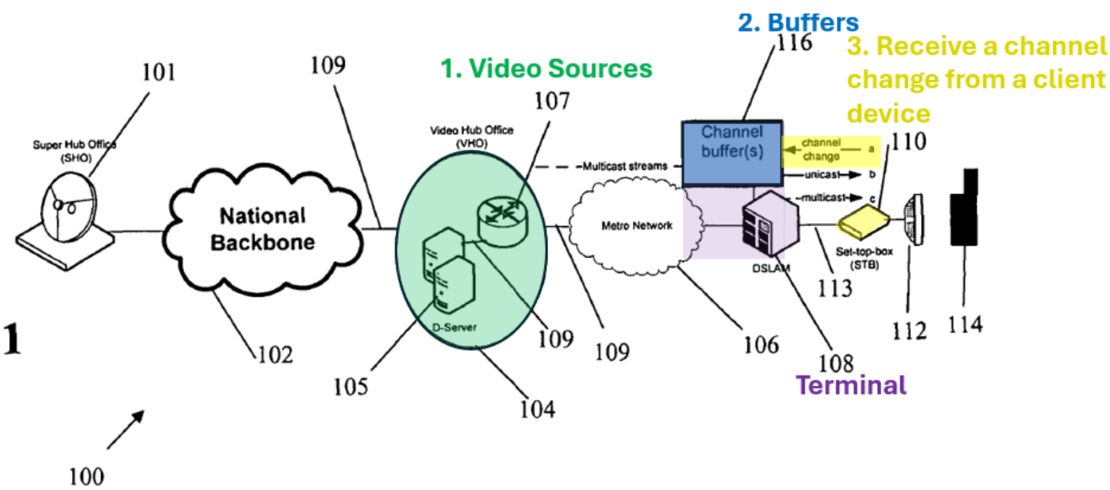


FIG. 1

144. A POSITA would have been motivated and have had a reasonable expectation of success to implement Mao-Liu this way as explained in Section VIII.A.1 (motivation).

f. Elements 1[E]/15[E]: [identifying]/[computer code for identifying] a next available key frame associated with the second media stream, wherein the next available key frame is provided as the second key frame in response to the channel change request, wherein no predictive frames are transmitted at the beginning of the channel.

145. **Ground 1:** Mao teaches Elements 1[E] and 15[E].

146. Mao's processor "identif[ies] a next available key frame associated with the second media stream." As I discussed in Section VIII, Petitioner in the related litigation proposed that "identifying a next available key frame associated with a second media stream" means "identifying, by the streaming server, by decoding or partially decoding the frame, a next available key frame associated with the second media stream." Ex-1021. Once Mao's set-top unit requests to change the channel from the first channel (video stream) to another channel (video stream) and the BDT determines that the request is proper, the BDT then "direct[s] the corresponding video data downstream to the appropriate set top unit 19." Ex-1005, 5:55-65. BDT (the streaming server) synchronizes with "channel X as it is received from the ATM network" by decompressing and multiplexing the channel. *Id.*, 6:20-26. Mao's BDT's decompressing and multiplexing is "decoding or partially decoding the frame" because BDT 12 (the streaming server) receives the media stream and converts the compressed video file into a format that can be displayed on various devices. Specifically, a POSITA would have understood that (i) detecting

and indexing I frames at ingress involves parsing MPEG headers to determine frame type (a form of partial decoding), and (ii) the server's explicit decompression step is decoding outright, both squarely within "decoding or partially decoding the frame" by the streaming server.

147. Mao identifies the next available key frame after receiving a channel change. Mao's BDT receives the channel change and "***must then wait*** for the Group of Picture (GOP) start point of the requested channel." Ex-1005, 7:8-12. The GOP start point is coded in the "I" or intracoded frames. *Id.*, 7:29-31. Thus, when "a change of channel request reaches BDT 12, the processor can immediately look up the location of the I frame from the pointer location buffer 57 and immediately access the appropriate video data." Ex-1005, 8:52-55. The key frame associated with the second media stream is identified as the next available key frame in response to the channel change request because it is the I frame determined as the GOP starting point.

148. Mao provides the "next available key frame… as the second key frame in response to the channel change request, wherein no predictive frames are transmitted at the beginning of the channel."[6] As discussed in Section VIII, a

[6] As discussed in Section VIII, Petitioner in the related litigation has proposed that this limitation is indefinite. PO has proposed that the limitations are not indefinite. *See* Ex-1022. For the purposes of this proceeding, I have applied these limitations

"predictive frame" is "a frame that relies on information from a reference key frame." Ex-1001, 2:67-3:2. While Mao teaches that predicted frames and bidirectional frames (P and B frames) are between "I" frames, in response to the channel change request, the BDT ***waits*** for the starting point (i.e., the next I-frame) when transmitting new channel information. Ex-1005, 7:8-12. After a channel change is requested, "the requested signal can instantly be transmitted downstream to the subscriber since the processor 55 is always 'pointing' to an I frame for each respective channel." *Id.*, 8:43-47. Therefore, transmission of the requested channel begins at the start of a GOP frame, (i.e., the I frame), meaning that the transmission does not include any B or P frames before the starting I-frame is transmitted of ***either*** the new channel or original channel.

149. Mao further teaches the sequential order of elements recited in elements 1[D]/15[D] and 1[E]/15[E] because BDT 12 identifies the next available key frame after receiving the channel change request. Specifically***, after a channel change is requested***, "the requested signal can instantly be transmitted downstream to the subscriber since the processor 55 is always 'pointing' to an I frame for each respective channel." Ex-1005, 8:43-47.

---

to mean that in response to a change request, the next available key frame is provided and there are not any predictive frames transmitted at the beginning of the channel.

150. **<u>Ground 2:</u>** Mao-Liu teaches Elements 1[E] and 15[E].[7]

151. While Mao teaches identifying a next available key frame in response to receiving a channel change request as described in the preceding section, Liu teaches further implementation details relating to identifying the key frame after receiving a channel change. As I described in Section VIII.A.2.e (elements 1[D]/15[D]), Liu's channel buffer device 116 (blue) "may also include programs or instructions that detect channel changes received via a client premise equipment, such as STB 110, a processor and computer programs or instructions that provide the stored I-frame to the DSLAM in response to a customer event, such as a change of a channel." Ex-1006, 4:1-7.

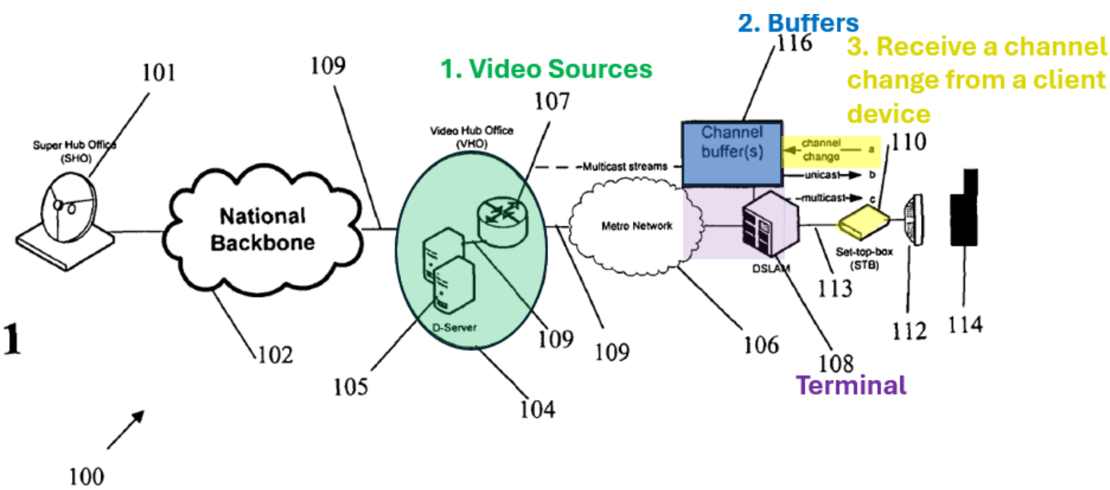
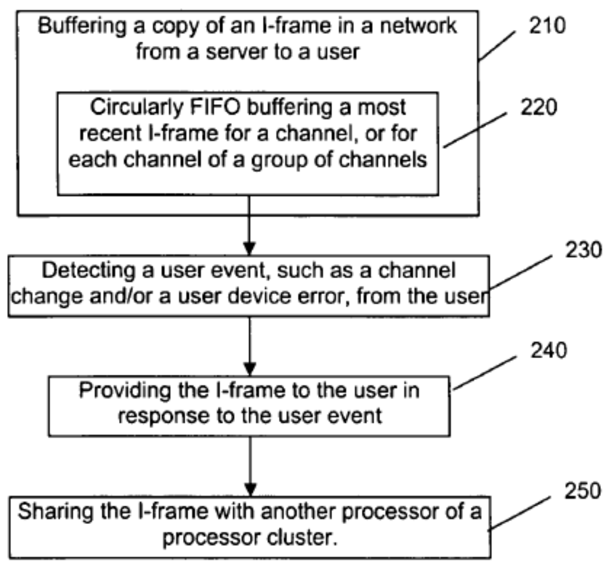


FIG. 1

[7] As explained above, Liu teaches "computer code" because Liu teaches "a computer readable medium containing instructions operative to be executed in a computer containing a processor." Ex-1006, 7:28-31.

152. As shown below in Figure 2, Liu explicitly teaches that in response to a channel change request, the next available key frame is provided to the client device, before any associated P or B frames. Ex-1006, 7:24-25 ("providing 240 the I-frame to the user in response to the user event"), 9:12-23 ("The I-frame is used as a reference to predict and encode B-frames or P-frames that follow the I-frame.").

FIG. 2



153. A POSITA would have been motivated, and have had a reasonable expectation of success, to implement Mao-Liu this way as explained in Section VIII.A.1 (motivation).

3. Dependent Claims 2, 4-7, 16, and 18-20

a. Claims 2/16: [The method of claim 1]/[The computer readable medium of claim 15], wherein the first media stream is a first video stream and the second media stream is a second video stream.

154. **<u>Grounds 1 and 2:</u>** Mao/Mao-Liu teaches Claims 2 and 16.

155. Mao teaches that the first and second media streams are video streams.[8] Mao teaches a "digital video system receiving video data corresponding to a plurality of video channels." Ex-1005, 10:8-9 (Cl. 5). Mao's broadband digital terminal (BDT 12) is provided "a plurality of compressed ***video signals***" from a video service provider for multiple channels. Ex-1005, 7:62-64, 3:38-42, 10:8-9 (Cl. 5). Each video channel is a video stream.

b. Claims 4/18: [The method of claim 1]/[The computer readable medium of claim 15], wherein the key frames are identified by detecting the start of a GOP.

156. **<u>Grounds 1 and 2:</u>** Mao/Mao-Liu teaches Claims 4 and 18.

157. Mao teaches that the key frames are identified by detecting the start of a GOP. Mao teaches that "compression schemes ***require a starting or synchronization frame*** usually referred to as the ***Group of Picture (GOP) start***

[8] As discussed in Section VII, Petitioner, in the related litigation, proposed that claim 2 is invalid under 35 U.S.C. § 112, ¶4. PO has proposed that the plain and ordinary meaning should apply. For the purposes of this proceeding, I have applied this claim to mean that the first media stream is a first video stream and the second media stream is a second video stream because the media stream, as recited in claim 1, includes a plurality of frames.

point." Ex-1005, 7:1-5. Mao "detects the ***GOP start frame*** (or I frame in the preferred embodiment) for each channel received and establishes a pointer for each GOP start frame." Ex-1005, 8:22-24.

c. Claims 5/19: [The method of claim 1]/[The computer readable medium of claim 15], wherein the key frames are identified by flagging each start of a GOP as a key frame.

158. **Grounds 1 and 2:** Mao/Mao-Liu teaches Claims 5 and 19.

159. Mao teaches that the key frames are identified by flagging each start of a GOP as a key frame. Mao "detects the GOP start frame (or I frame in the preferred embodiment) for each channel received and establishes a pointer for each GOP start frame." Ex-1005, 8:22-24. As I described above in Section V.B, GOP begin with an I frame, which would have been understood as a "key frame." *See* Ex-1008, 1:58-2:9; *see also id.*, 3:7-17; Ex-1010, 7:5-16, 9:33-52; Ex-1012 ¶¶12, 13. For example, Choi discusses that the "key frame is a complete picture used for video compression. Referring to video group of pictures (GOP) structure, a frame is selected one by one at a constant interval from a temporal video flow, and is designated as the key frame. The key frame is an independently restorable video, and enables an arbitrary video access." Ex-1011, 2:55-60. In particular, "MPEG, H.26a, and H.264, the key frame is inserted at a constant interval into a GOP, and designates the I fame enabling independent video production." *Id.*, 2:61-63. Bixby also discusses how "the encoded

MPEG video stream is comprised of a series of groups of pictures (GOPs), and each GOP begins with an independently encoded (intra) I frame and may include one or more following P-frames and B-frames." Ex-1013 ¶6.

d. Claims 6/20: [The method of claim 1]/[The computer readable medium of claim 15], wherein the key frame is an I-frame.

160. __**Grounds 1 and 2:**__ Mao/Mao-Liu teaches Claims 6 and 20.

161. Mao teaches that the key frame is an I-frame. Mao "detects the ***GOP start frame*** (***or I frame*** in the preferred embodiment) for each channel received and establishes a pointer for each GOP start frame." Ex-1005, 8:22-24. As I described above in Section V.B, GOP begin with an I frame, which would have been understood as a "key frame." See Ex-1008, 1:58-2:9; see also id., 3:7-17; Ex-1010, 7:5-16, 9:33-52; Ex-1012 ¶¶12, 13. For example, Choi discusses that the "key frame is a complete picture used for video compression. Referring to video group of pictures (GOP) structure, a frame is selected one by one at a constant interval from a temporal video flow, and is designated as the key frame. The key frame is an independently restorable video, and enables an arbitrary video access." Ex-1011, 2:55-60. In particular, "MPEG, H.26a, and H.264, the key frame is inserted at a constant interval into a GOP, and designates the I fame enabling independent video production." *Id.*, 2:61-63. Bixby also discusses how "the encoded MPEG video stream is comprised of a series of groups of pictures (GOPs), and each GOP begins

with an independently encoded (intra) I frame and may include one or more following P-frames and B-frames." Ex-1013 ¶6.

e. Claim 7: The method of claim 1, wherein channel change does not occur until the second key frame can be provided to the client device.

162. **Grounds 1 and 2:** Mao/Mao-Liu teaches Claim 7.

163. Mao-Liu teaches that the channel change does not occur until the second key frame can be provided to the client device. Mao teaches that it is "desirable for the FIFO buffer to ***store two or more synchronization frames per video signal*** (i.e., each buffer stores at least thirty frames in the … MPEG-2 embodiment)." Ex-1005, 9:2-7. Thus, when a subscriber's channel change request is received, the request signal can instantly be transmitted downstream to the subscriber since the processor 55 is always "pointing" to an I frame for each respective channel.

164. A POSITA would have understood that Mao's storage of two synchronization frames is storing the first and second key frames before allowing the channel change to be provided to the client device because doing so was well known as taught in Liu. For example, Liu teaches that the channel buffer device 116 buffers "***the latest frames for a channel up to the most recent 'N' I-frames***, where N may be as small as 1." Ex-1006, 4:46-48. As described above, N may be 1 but it can also store two I-frames ("key frames"). If a user changes a channel, the channel's

I-frame is provided and the "DSLAM 108 retrieves an I-frame from the memory of the channel change buffer and provides that I-frame to the STB 110." Ex-1006, 5:35-44. A POSITA would have been motivated to retrieve two I-frames before providing the channel change to Mao's subscriber because this allows the set-top box to reconstruct a full image from the subsequently provided frames.

4. Independent Claim 8

a. Element 8[pre]: A system:

165. **Grounds 1/2:** Mao/Mao-Liu teaches the preamble of Claim 8, to the extent it is limiting, for the same reasons explained in Section VIII.A.2.a (1[pre]/15[pre]). Mao teaches a "method and ***apparatus*** for rapidly changing the channel in a digital video delivery system." Ex-1005, 2:26-28; *see also id.*, 1:7-10.

b. Element 8[A]: [(i)] an interface configured to receive a first media stream including a first plurality of frames, [(ii)] a second media stream including a second plurality of frames, [(iii)] the first media stream received at a streaming server and maintained in a first buffer and [(iv)] and the second media stream received at the streaming server and maintained in a second buffer, [(v)] the streaming server configured to receive the first media stream from a media aggregation server, wherein a plurality of buffers at the streaming server are allocated on a per group of pictures (GOP) basis;

166. **Grounds 1/2:** Mao/Mao-Liu teaches elements 8[A(i), A(iii), A(v)] for the same reasons explained in Section VIII.A.2.b (1[A]/15[A]) and elements 8[A(ii), A(iv)] for the same reasons explained in Section VIII.A.2.c (1[B]/15[B]). While

element 8[A] recites the same limitations in a slightly different order than elements 1[A]-1[B]/15[A]-15[B] as shown in the table below, the scope and meaning of these elements are not substantively different. For clarity, I have highlighted the differences between the claims in the table below.

<table>
<tr><th>Claim 1</th><th>Claim 15</th><th>Element 8[A]</th></tr>
<tr><td>1[A] [(i)] receiving a first media stream including a first plurality of frames, [(ii)] the first media stream received at a streaming server and maintained in a first buffer, [(iii)] the streaming server configured to receive the first media stream from a media aggregation server, [(iv)] wherein a plurality of buffers at the streaming server are allocated on a per group of pictures (GOP) basis;</td><td>15[A] [(i)] computer code for receiving a first media stream including a first plurality of frames, [(ii)] the first media stream received at a streaming server and maintained in a first buffer, [(iii)] the streaming server configured to receive the first media stream from a media aggregation server, [(iv)] wherein a plurality of buffers at the streaming server are allocated on a per group of pictures (GOP) basis;</td><td rowspan="2">8[A] [(i)] an interface configured to receive a first media stream including a first plurality of frames and [(ii)] a second media stream including a second plurality of frames, [(iii)] the first media stream received at a streaming server and maintained in a first buffer and [(iv)] the second media stream received at the streaming server and maintained in a second buffer, [(v)] the streaming server configured to receive the first media stream from a media aggregation server, wherein a plurality of buffers at the streaming server are allocated on a per group of pictures (GOP) basis;</td></tr>
<tr><td>1[B] receiving a second media stream including a second plurality of frames, the second media stream received at the streaming server and maintained in a second buffer;</td><td>15[B] computer code for receiving a second media stream including a second plurality of frames, the second media stream received at the streaming server and maintained in a second buffer;</td></tr>
</table>

167. Additionally, Mao teaches "an interface configured to receive" the streams because, as shown in Figure 1, Mao's broadband digital terminal (BDT 12) receives "a plurality of compressed video signals" from a video service provider. Ex-1005, 7:62-64, 3:38-42. As shown in Mao teaches that "BDT 12 has two OC-12c broadcast ports, which can only receive signals carrying ATM cells, and one OC-12c interactive port which can receive and transmit signals." Ex-1005, 3:56-59. The "***interface to the ATM network-BDT interface 39*** can be realized using an OC-3 or OC-12 optical interface[] carrying ATM cells." *Id.*, 3:53-56.

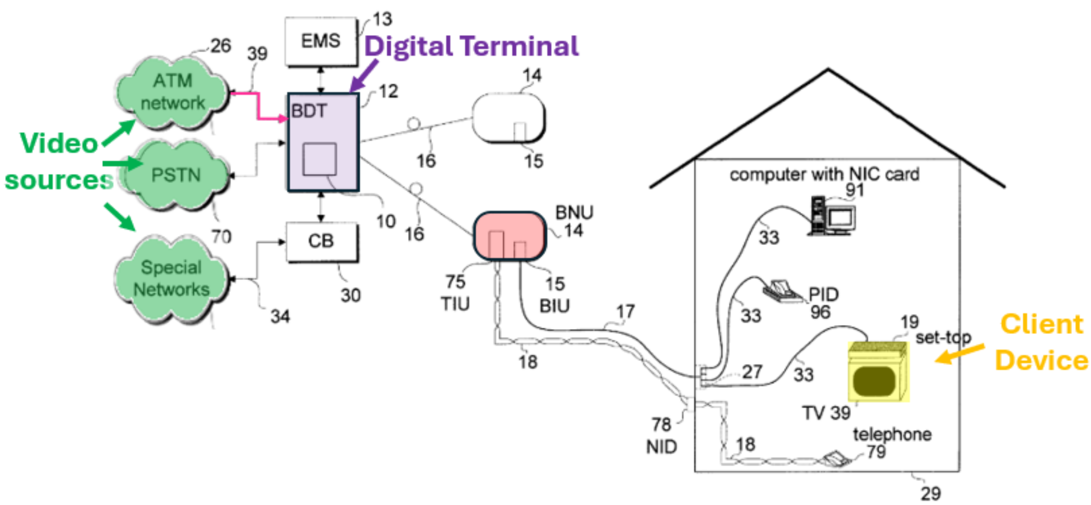


FIG. 1

c. **Element 8[B]: a processor configured to determine positions of a plurality of key frames in the first buffer and the second buffer, the plurality of key frames including a first key frame in the first buffer and a second key frame in the second buffer, wherein determining positions of the plurality of key frames includes determining at the streaming server a group of pictures (GOP) size N and identifying Nth frames**

following the first key frame as subsequent key frames;

168. **<u>Grounds 1/2:</u>** Mao/Mao-Liu teaches element 8[B] for the same reasons explained in Section VIII.A.2.d (1[C]/15[C]). While element 8[B] use slightly different language as shown in the table below, the scope and meaning of these elements are not substantively different. For clarity, I have highlighted the differences between the claims in the table below.

Element 1[C]	***Element 15[C]***	***Element 8[B]***
[(i)] determining positions of a plurality of key frames in the first buffer and the second buffer, the plurality of key frames including a first key frame in the first buffer and a second key frame in the second buffer, [(ii)] wherein determining positions of the plurality of key frames includes determining at the streaming server a group of pictures (GOP) size N and identifying Nth frames following the first key frame as subsequent key frames;	[(i)] ***computer code for*** determining positions of a plurality of key frames in the first buffer and the second buffer, the plurality of key frames including a first key frame in the first buffer and a second key frame in the second buffer, [(ii)] wherein determining positions of the plurality of key frames includes determining at the streaming server a group of pictures (GOP) size N and identifying Nth frames following the first key frame as subsequent key frames;	***a processor configured to*** determine positions of a plurality of key frames in the first buffer and the second buffer, the plurality of key frames including a first key frame in the first buffer and a second key frame in the second buffer, wherein determining positions of the plurality of key frames includes determining at the streaming server a group of pictures (GOP) size N and identifying Nth frames following the first key frame as subsequent key frames;

169. Additionally, Mao teaches a "processor configured to determine," because Mao's BDT 12 processor 55 receives the frames for each channel and

detects, establishes, and records the GOP start frame (I frame) for each channel. Ex-1005, 8:15-26.

d. Element 8[C]: wherein the processor is further configured to identify a next available key frame associated with the second media stream, wherein the next available key frame is provided as the second key frame in response to the channel change request, wherein no predictive frames are transmitted at the beginning of the channel change request.

170. **Grounds 1/2:** Mao/Mao-Liu teaches element 8[C] for the same reasons explained in Section VIII.A.2.f (1[E]/15[E]). While element 8[C] use slightly different language as shown in the table below, the scope and meaning of these elements are not substantively different. For clarity, I have highlighted the differences between the claims in the table below.

Element 1[E]	***Element 15[E]***	***Element 8[C]***
identifying a next available key frame associated with the second media stream, wherein the next available key frame is provided as the second key frame in response to the channel change request, wherein no predictive frames are transmitted at the beginning of the channel change request.	***computer code for*** identifying a next available key frame associated with the second media stream, wherein the next available key frame is provided as the second key frame in response to the channel change request, wherein no predictive frames are transmitted at the beginning of the channel change request.	***wherein the processor is further configured to*** identify a next available key frame associated with the second media stream, wherein the next available key frame is provided as the second key frame in response to the channel change request, wherein no predictive frames are transmitted at the beginning of the channel change request.

5. Dependent Claims 9 and 11-14[9]

a. Claim 9: The method of claim 8, wherein the first media stream is a first video stream and the second media stream is a second video stream.

171. **Grounds 1 and 2:** Mao/Mao-Liu teaches Claim 9 for the same reasons explained in Section VIII.A.3.a (claims 2 and 16).

b. Claim 11: The method of claim 8, wherein the key frames are identified by detecting the start of a GOP.

172. **Grounds 1 and 2:** Mao/Mao-Liu teaches Claim 11 for the same reasons explained in Section VIII.A.3.b (claims 4 and 18).

c. Claim 12: The method of claim 8, wherein the key frames are identified by flagging each start of a GOP as a key frame.

173. **Grounds 1 and 2:** Mao/Mao-Liu teaches Claim 12 for the same reasons explained in Section VIII.A.3.c (claims 5 and 19).

d. Claim 13: The method of claim 9, wherein the key frame is an I-frame.

174. **Grounds 1 and 2:** Mao/Mao-Liu teaches Claim 13 for the same reasons explained in Section VIII.A.3.d (claims 6 and 20).

---

[9] I understand that claims 9 and 11-14 recite the "method of claim 8," but claim 8 recites a system. I have treated claims 9 and 11-14 as reciting the system of claim 8.

e. Claim 14: The method of claim 8, wherein channel change does not occur until the second key frame can be provided to the client device.

175. **Grounds 1 and 2:** Mao/Mao-Liu teaches Claim 14 for the same reasons explained in Section VIII.A.3.e (claim 7).

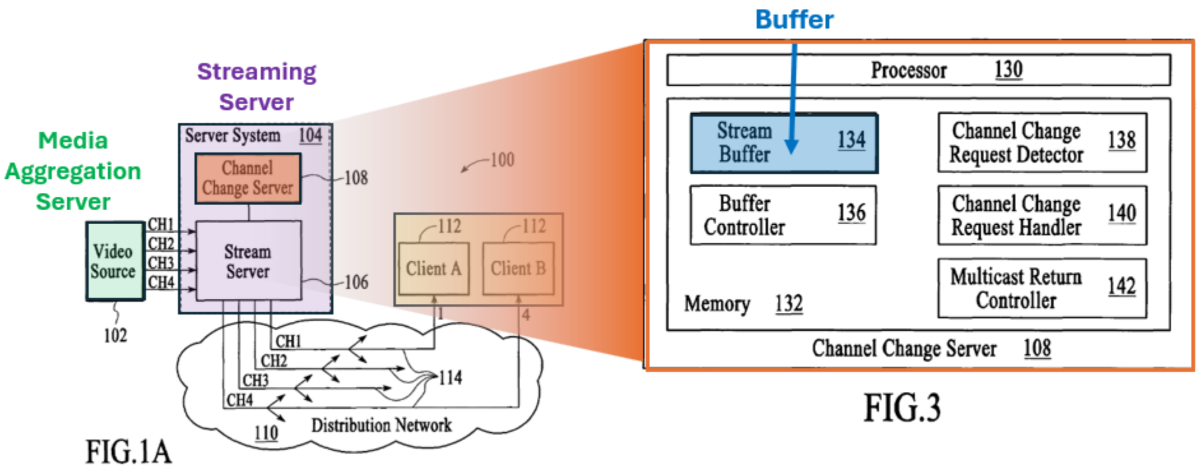
B. Ground 3: Claims 1, 2, 4-6, 8, 9, 11-13, 15, and 18-20 are obvious in view of Sherer in view of Wu-771.

1. A POSITA would have been motivated to combine Sherer's teachings and Wu-771's teachings and would have had a reasonable expectation of success.

176. A POSITA would have been motivated to combine the teachings of Sherer and Wu-771 and had a reasonable expectation of success in doing so because each relates to providing a video stream to a client device in response to a channel change request. Ex-1007, Title, Abstract; Ex-1008, Abstract.

177. Sherer teaches a "method for servicing channel change requests in a digital video network that is capable of distributing digital video content to clients via multicasting and unicasting." Ex-1007, 6:35-38, 14:62-65 (cl. 1). As shown in Figure 1A and Figure 3, Sherer's server system 104 (purple) includes a stream server 106 and channel change server 108 (orange), which includes a processor 120, memory 143, stream buffer 134 (blue), buffer controller 136. Ex-1007, 7:6-15. Sherer's server system 104 ingests the digital video content from the video source (green) and controls the distribution of the digital video content into the distribution network. Ex-1007, 4:29-31. Stream buffer 134 "buffers up to 1 group of pictures

(GOP) worth of each ingested stream, where a GOP is defined as a series of multiple consecutive frames beginning with an I-frame and continuing until the last frame before the next I-frame." *Id.*, 8:60-64.



178. Additionally, Wu-771 teaches to encode a video stream and provide encoded frames of video streams. Wu-771 teaches that "Group of pictures is defined as the number of I, P, B frames that are sent prior to the next I frame." Ex-1008, 3:63-64. For example, a "typical Group of Pictures (GOP) that contains I, P and B frames [are] sent" such that an I frame is followed by fourteen B and P frames. *Id.*, 3:65-4:2. "This is commonly referred to as a broadcast channel with a GOP size of 15 (i.e., 15 frames in the GOP)." *Id.*, 4:3-4. Wu-771 includes a "video acceleration module VAM" which buffers the frames. *Id.*, 4:66-67. As shown in Figure 3 below, Wu-771's VAM has "separate outgoing client buffers for each client port (PBUF) 401" and "fixed buffers for each multicast stream it is sourcing to clients, referred

to herein as Streaming Buffers (StrmBuf).” Ex-1008, 4:67-5:3. Wu-771’s StrmBuf 402 buffers “fixed number of GOP’s aligning the start of the buffer to the start of a GOP.” *Id*., 5:3-6.

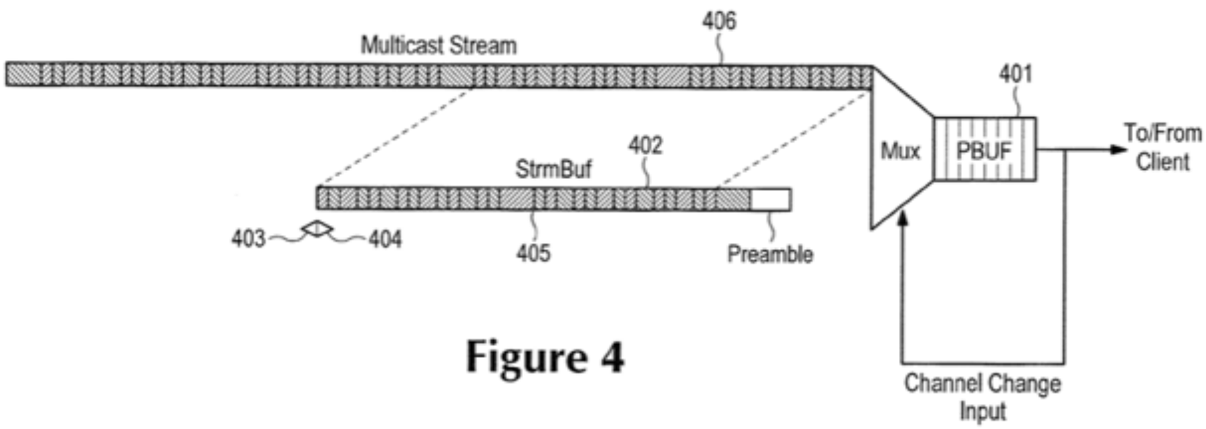


Figure 4

179. A POSITA would have been motivated to maintain Sherer’s “group of pictures (GOP) … of ***each ingested stream***,” in separate buffers at its streaming server, as taught by Wu-771. Wu-771 teaches that its separate outgoing client buffers allow for the video to be sourced to clients faster. Wu-771 teaches that “the higher number of I frames sent, the faster the decoder can start displaying the picture (i.e., faster channel change time from a user perspective), since the decoder needs an I frame to initially sync to; again.” Ex-1008, 3:59-61. This would ensure that each channel’s most recent I frames are stored and can be provided to the user upon request, thus eliminating delays that would otherwise occur to fetch, identify, and further process the video stream to be provided to the user’s device and prevent distorted videos from being presented to the user. While Sherer does not detail how

to guarantee an immediate I frame (it assumes buffered content exists and focuses on when to switch and how to return), Wu-771 delivers the immediate I frame and deterministic catch up under bandwidth limits. Both references target IPTV channel change delay in multicast environments and operate on the same building blocks: MPEG I/P/B frames and GOPs; server/edge buffers; unicast and multicast delivery; and standard transport signaling (PAT/PMT/PCR, IGMP join). There is no functional conflict. Wu's per channel, buffered I, P, and B frames of the user's first channel, while Sherer's logic optimizes the time and conditions to hand back to multicast and scales the system under surf behavior. Using Wu 771's burst as Sherer's unicast content is a straight substitution of known, compatible means to achieve a better fast start. The combination would thus yield a predictable result of faster initial video and controlled, scalable return.

180. A POSITA would have had a reasonable expectation of success in combining the teachings of Sherer and Wu-771. As I have shown below in Modified Figure 3' below, Wu-771's separate client buffers (blue) buffer Sherer's GOPs for each channel. Therefore, Wu-771's channel buffers would have been implemented in conjunction with Sherer's architecture to carry out the same functions that Sherer is already performing, while achieving the benefits taught by Wu-771. For example, both uses CODECs and MPEG compressing schemes. Specifically, both references operate on MPEG style GOPs and name the same control data (PAT/PMT/PCR) and

client behaviors (decoder needs an I frame; IGMP join). No unknown algorithms are required: Wu's frame drop/blank insert math is explicit; Sherer's return mechanisms (rate increase or duplicate frame detection) are spelled out with buffer diagrams. Both also use edge/server buffers, unicast sender, multicast group join, timestamp adjustments, and switch criteria. A POSITA implementing Wu's VAM as the source of Sherer's unicast burst, then using Sherer's (a) rate boost or (b) dual uni to multi cutover, would reasonably expect a working system.

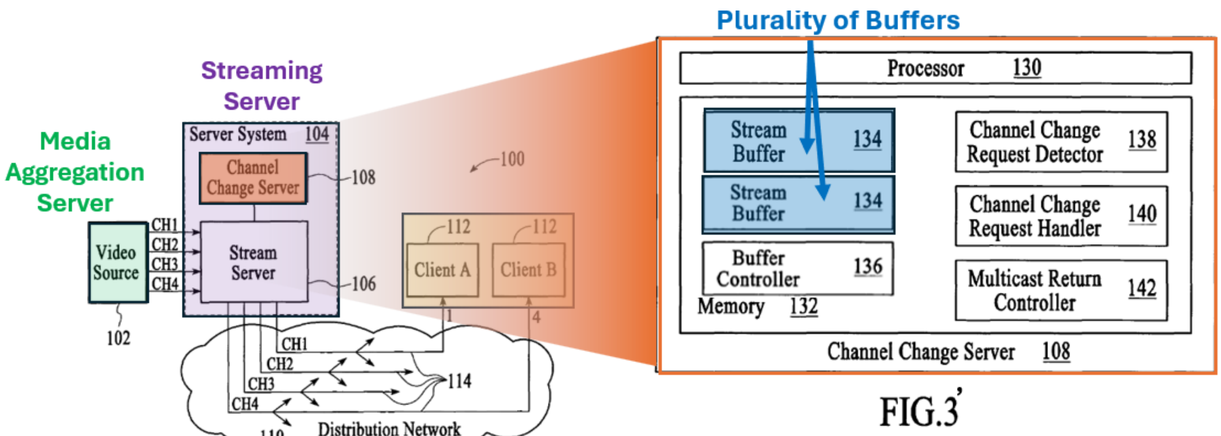


FIG.1A FIG.3

181. Moreover, both references describe methods for receiving video data streams, identifying and storing the I-frames and predictive frames, and providing the channel requested by the user using similar hardware and software components. Ex-1007, Abstract; Ex-1008, Abstract. A POSITA would thus have understood that these references disclose interrelated teachings based on well-understood technologies that would have been amenable to various well-understood and predictable combinations. Taken together, these teachings create a clear design

incentive: (1) get an I frame on screen immediately; (2) manage last mile bandwidth on catch up; (3) avoid IGMP churn and scale by returning to multicast under sensible conditions; these are the split responsibilities of Wu-771 and Sherer. A POSITA would have every reason to integrate them to meet product requirements for both zapping speed and network scalability.

2. Independent Claims 1 and 15

a. Element 1[pre]: A method:

Element 15[pre]: A non-transitory computer readable medium, comprising:

182. Sherer-Wu-771 teaches the preambles of claims 1 and 15 to the extent that they are limiting.

183. Sherer teaches a "method for servicing channel change requests in a digital video network that is capable of distributing digital video content to clients via multicasting and unicasting." Ex-1007, 6:35-38, 14:62-65 (cl. 1). Sherer also teaches "[a] processor accessible non-transitory storage media comprising processor-executable instructions." *Id.*, 15:15-20 (cl. 18); *see also id.*, 12:50-61. Sherer also teaches that the functions of its components "may be embodied as processor-executable instructions that are stored on processor-accessible memory media." *Id.*, 12:62-13:25.

b. Elements 1[A]/15[A]: [(i)] [receiving]/[computer code for receiving] a first media stream including a first plurality of frames, [(ii)] the first media stream

received at a streaming server and maintained in a first buffer, the streaming server configured to receive the first media stream from a media aggregation server, [(iii)] wherein a plurality of buffers at the streaming server are allocated on a per group of pictures (GOP) basis;

184. Sherer-Wu-771 teaches elements 1[A] and 15[A].

185. Sherer teaches "receiving a first media stream including a first plurality of frames" as claimed. For example, as shown in Figure 1A, "multiple channels of digital video content" are provided to the stream server 106 (purple). Ex-1007, 4:56-57. Each channel is made of a plurality of frames because Sherer's "stream buffer buffers up to 1 group of pictures (GOP) worth of each ingested stream, where a GOP is defined as a series of multiple consecutive frames beginning with an I-frame and continuing until the last frame before the next I-frame." *Id.*, 8:60-64. Therefore, Sherer's server system 104 receives a first media stream comprising a plurality of frames because stream server 104 buffers up to 1 GOP worth for each channel.

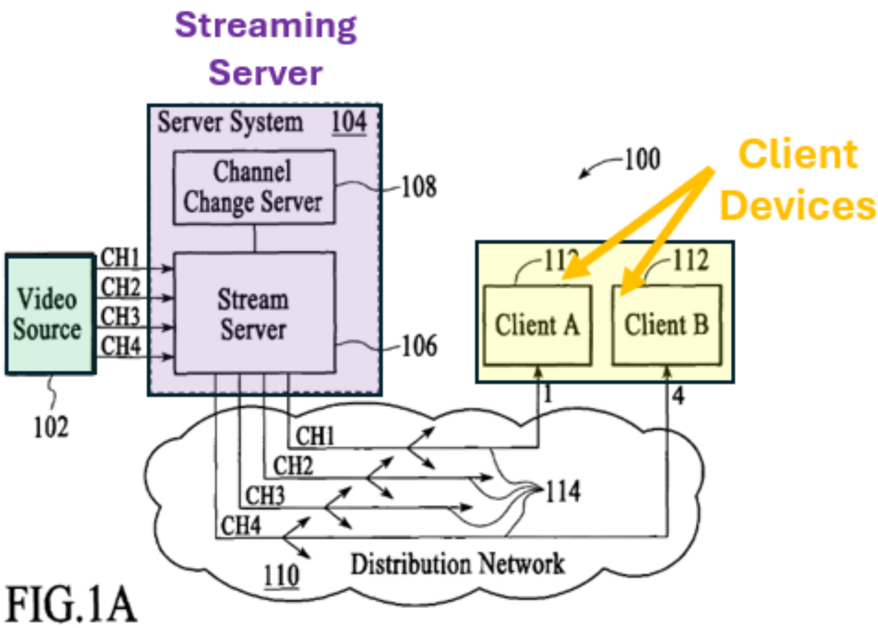


FIG.1A

186. Sherer teaches that "the first media stream received at a streaming server [is] maintained in a first buffer." Sherer's server system 104 ingests the digital video content for the video source and controls the distribution of the digital video content into the distribution network. Ex-1007, 4:29-31. As shown below in Figure 1A and Figure 3, Sherer's server system 104 ("streaming server," purple) includes a stream server 106 a channel change server 108 (orange), which includes a processor 120, memory 143, stream buffer 134 (blue), buffer controller 136. *Id.*, 7:6-15. The buffer controller manages the buffering of digital video content that is streamed into the stream buffer. *Id.*, 7:37-38. The stream buffer "buffers up to 1 group of pictures (GOP) worth of each ingested stream, where a GOP is defined as a series of multiple consecutive frames beginning with an I-frame and continuing until the last frame

before the next I-frame." *Id.*, 8:60-64. Therefore, each channel's GOP is received by the stream server and maintained in a buffer, stream buffer 134 (blue).

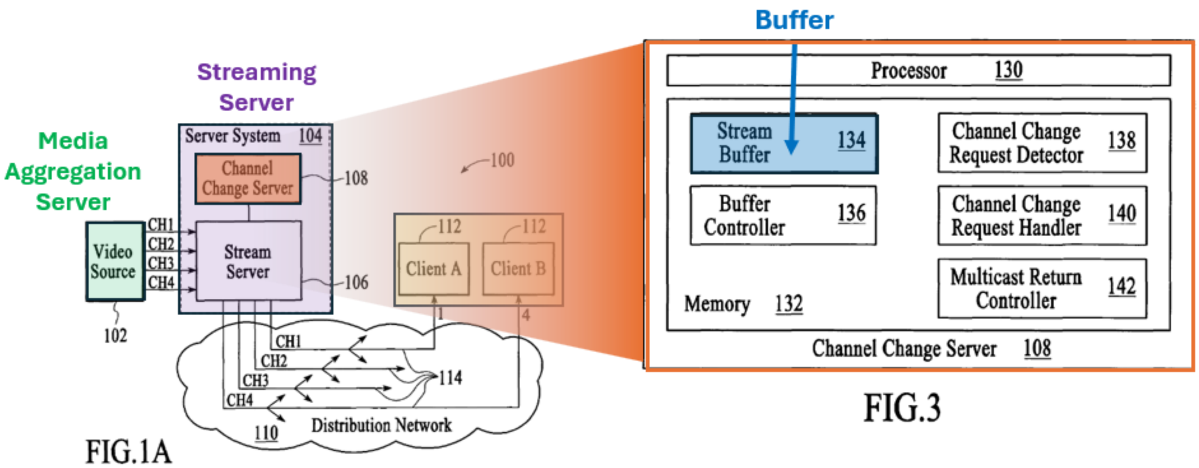


FIG.1A
FIG.3

187. In Sherer, "the streaming server [is] configured to receive the first media stream from a media aggregation server." As discussed in Section VII, above, a "media aggregation server" is defined in the specification as "any mechanism for obtaining media content from a large number of sources in order to provide the media content to mobile devices in live broadcast streams." Ex-1001, 6:54-58. Sherer's streaming server 104, receives "multiple channels of digital video content" from video source 102 (green). Ex-1007, 4:56-57. Sherer teaches that "the digital video content includes real-time content from the content producers and/or distributors" either live or pre-recorded. *Id.*, 4:56-5:1. As shown above in Figure 1A, Sherer's video source 102 (green) is a "media aggregation server" because it is the "mechanism for obtaining media content from a large number of sources in order to

provide the media content … in live broadcast streams" as defined in the specification. Ex-1001, 6:54-58.

188. Additionally, Sherer's "plurality of buffers at the streaming server are allocated on a per group of pictures (GOP) basis." As discussed in Section VII, above, Petitioner's proposed construction is "wherein the streaming server allocates one or more different buffers for each group of pictures (GOP)." Ex-1021, 2. "GOP" is a "sequence of different frame types, beginning with a key frame and ending just before a subsequent key frame." Ex-1022, 2. Sherer's stream buffer 134 "buffers up to 1 group of pictures (GOP) worth of each ingested stream, where a GOP is defined as a series of multiple consecutive frames beginning with an I-frame and continuing until the last frame before the next I-frame." Ex-1007, 8:60-64. For example, as shown below in Figure 7A, stream buffer 134 is allocated on a per GOP basis by allocating one buffer for each GOP received, consisting of a sequence of different frames, beginning with an I-frame and ending just before a subsequent I-frame.

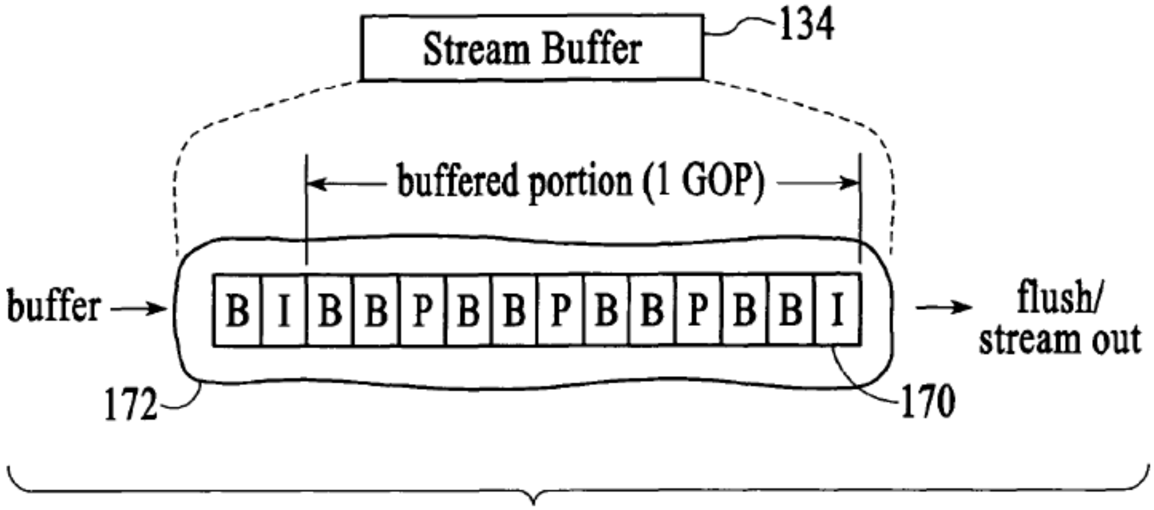
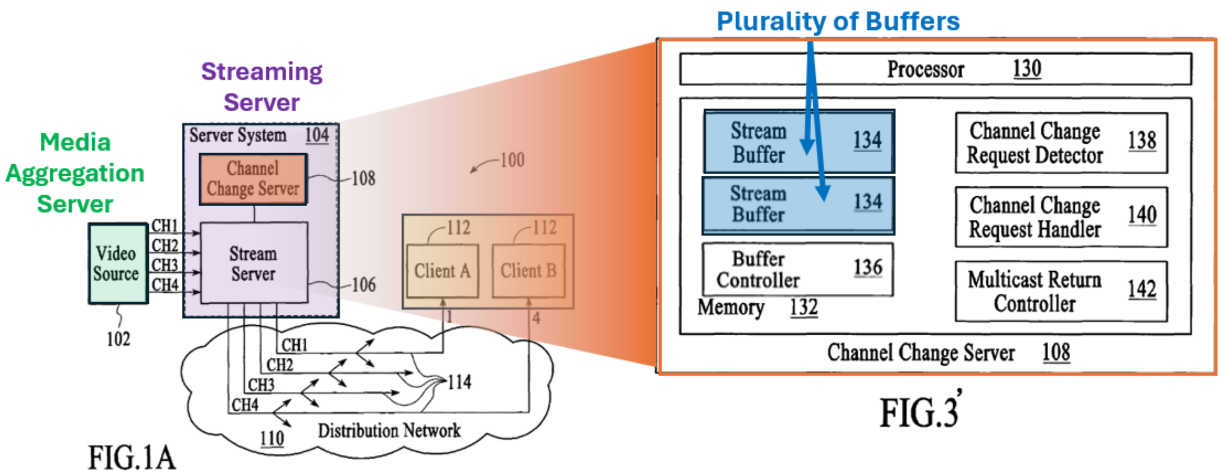


FIG.7A

189. While Sherer teaches implementing buffers on a per GOP basis as described in the preceding section, Wu-771 teaches further implementation details relating to allocating multiple buffers on a per GOP basis. A POSITA would have been motivated to implement Sherer's stream buffer to include multiple buffers, as taught by Wu-771. Wu-771 teaches that its Video/Audio Acceleration Mechanism (VAM) has "as ***separate outgoing client buffers*** for each client port (PBUF) 401" and "***fixed buffers for each multicast stream*** it is sourcing to clients, referred to herein as Streaming Buffers (StrmBuf)." Ex-1008, 4:67-5:3. Wu-771's StrmBuf 402 buffers "fixed number of GOP's aligning the start of the buffer to the start of a GOP." *Id*., 5:3-6.

190. The combined Sherer-Wu-771 system would thus have been implemented as shown in Figure 3' below. Wu-771's separate client buffers buffer

Sherer's GOPs for each channel. Therefore, Wu-771's channel buffers would have been implemented in conjunction with Sherer's architecture to carry out the same functions that Sherer is already performing of buffering the multiple media streams. As I explained above in Section V.C, it was well known in the art to implement one or more streaming buffers.



191. A POSITA would have been motivated, and have had a reasonable expectation of success, to implement Sherer-Wu-771 this way as explained in Section VIII.B.1 (motivation).

c. Elements 1[B]/15[B]: [receiving]/[computer code for receiving] a second media stream including a second plurality of frames, the second media stream received at the streaming server and maintained in a second buffer;

192. Sherer-Wu-771 teaches elements 1[B] and 15[B].

193. Sherer teaches "receiving a second media stream including a second plurality of frames." As explained in the previous section, Sherer's stream server

106 (purple) receives "multiple channels of digital video content." Ex-1007, 4:56-57. Each channel is made of a plurality of frames because Sherer teaches that its "stream buffer buffers up to 1 group of pictures (GOP) worth of each ingested stream, where a GOP is defined as a series of multiple consecutive frames beginning with an I-frame and continuing until the last frame before the next I-frame." *Id.*, 8:60-64. Therefore, Sherer's server system (streaming server) receives a second video stream of a plurality of frames because it receives multiple channels and buffers up to 1 GOP worth.

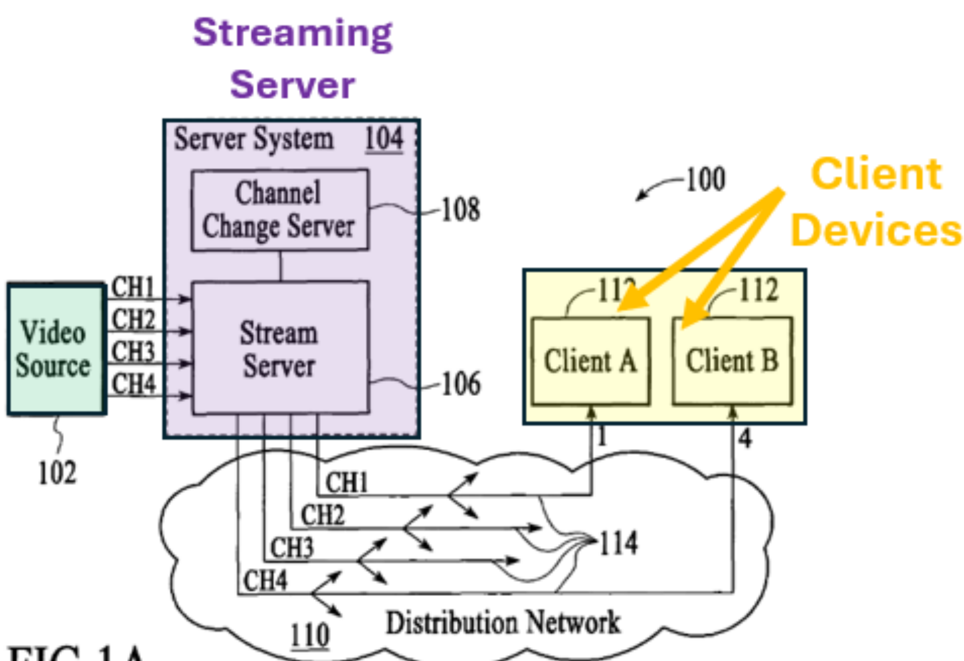


FIG.1A

194. Sherer's second media stream is maintained in a second buffer. Sherer teaches that "stream buffer buffers up to 1 group of pictures (GOP) worth of each ingested stream, where a GOP is defined as a series of multiple consecutive frames

beginning with an I-frame and continuing until the last frame before the next I-frame." Ex-1007, 8:60-64. The second video stream made of a plurality of frames is therefore maintained in a second buffer specific for that video channel, as was a well known technique, as I described above in Section V.C.

195. While Sherer teaches implementing buffers for each video stream, Wu-771 teaches further implementation details relating to storing a second stream in a second set of buffers. Wu-771 teaches that its Video/Audio Acceleration Mechanism (VAM) has "as ***separate outgoing client buffers*** for each client port (PBUF) 401" and "***fixed buffers for each multicast stream*** it is sourcing to clients, referred to herein as Streaming Buffers (StrmBuf)." Ex-1008, 4:67-5:3. Wu-771's StrmBuf 402 buffers "fixed number of GOP's aligning the start of the buffer to the start of a GOP." *Id*., 5:3-6. A POSITA would have been motivated to maintain Sherer's second video stream, including its plurality of frames, in a separate buffer, as taught by Wu-771, as explained in Section VIII.B.1.

d. Elements 1[C]/15[C]: [(i)] [determining]/[computer code for determining] positions of a plurality of key frames in the first buffer and the second buffer, the plurality of key frames including a first key frame in the first buffer and a second key frame in the second buffer, [(ii)] wherein determining positions of the plurality of key frames includes determining at the streaming server a group of pictures (GOP) size N and identifying Nth frames following the first key frame as subsequent key frames;

196. Sherer-Wu-771 teaches Elements 1[C] and 15[C].

197. Sherer determines "positions of a plurality of key frames in the first and second buffer." As discussed in Section VII, the "key frame" is "a frame that can be displayed substantially on its own." Ex-1001, 2:65-67. Sherer teaches that "[a]n I-frame (MPEG2) on the other hand is a frame that presents a complete new image" and thus can be displayed substantially on its own. Ex-1007, 1:46-48. As I explained above in Section V.B, an I frame is a key frame. *See* Ex-1006, 4:14-18; Ex-1001, 3:3-7, 9:58-59; *see also* Section VIII.A.2.d.

198. As shown below in Figure 5, Sherer's streaming server "ingests" the digital video content, block 601 (green), and "***portions of each channel*** are buffered by the channel change server for use in servicing channel change requests." Ex-1007, 7:56-61. Specifically, at block 602, digital video content is provided to the client via multicasting and, at block 604 (blue), "***multiple consecutive frames of digital video content, including at least one I-frame are buffered***." *Id.*, 7:59-63. Sherer's streaming server determines the plurality of I-frames for the first and second media streams because Sherer receives multiple channels and buffers "portions of each channel."

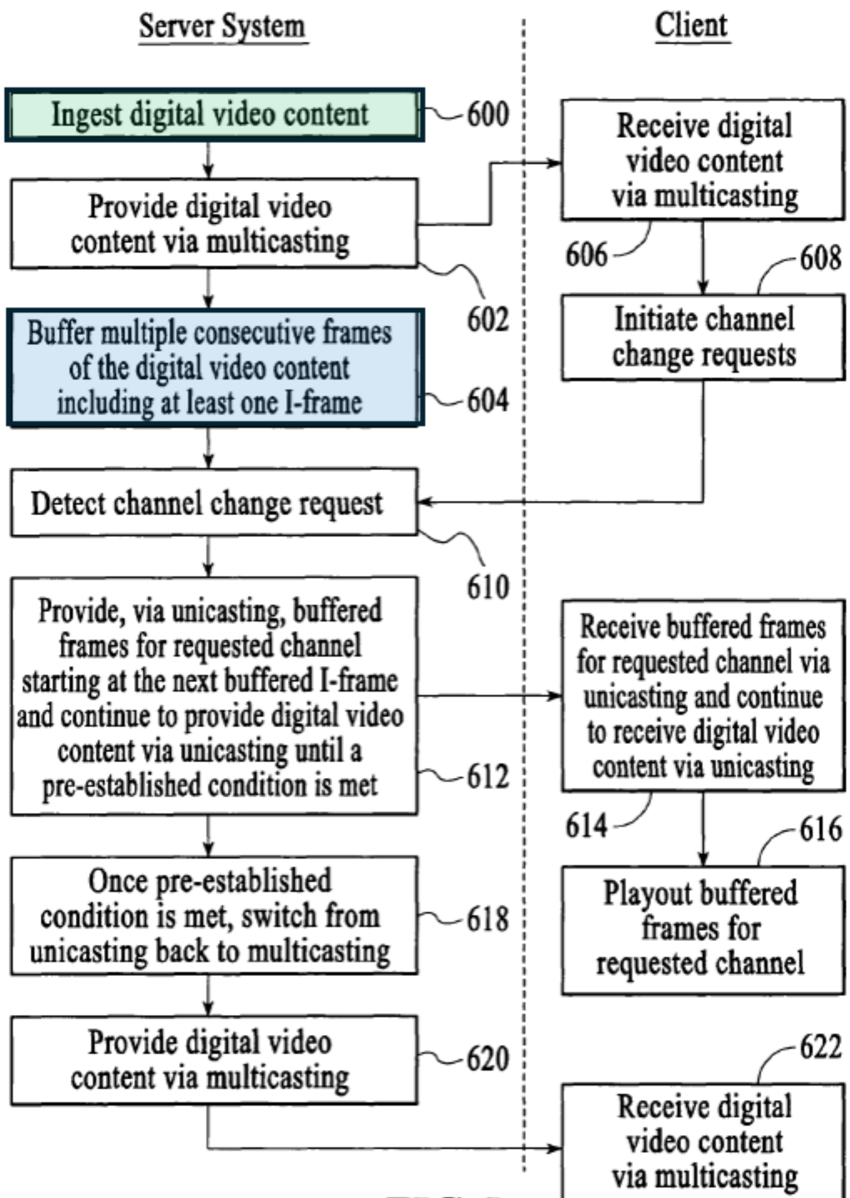


FIG.5

199. Sherer further teaches determining "a group of pictures (GOP) size N and identifying Nth frames following the first key frame as subsequent key frames." Sherer's "stream buffer buffers up to 1 group of pictures (GOP) worth of each ingested stream, where a GOP is defined as a series of multiple consecutive frames

beginning with an I-frame and continuing until the last frame before the next I-frame." Ex-1007, 8:60-64. For example, as shown below in Figure 7A, when determining the plurality of key frames (I-frames, red), Sherer determines that the GOP size is 12 (yellow) and buffers "12 frames 170 consisting of one I-frame [(red)], three P-frames, and eight B frames [(purple)]." *Id.*, 8:64-67. Sherer thus determines the GOP size (e.g., 12) and identifies the Nth frames following the first key frame (e.g., 11 frames following the first I frame), as subsequent frames before identifying the next key frame, as the next I-frame.

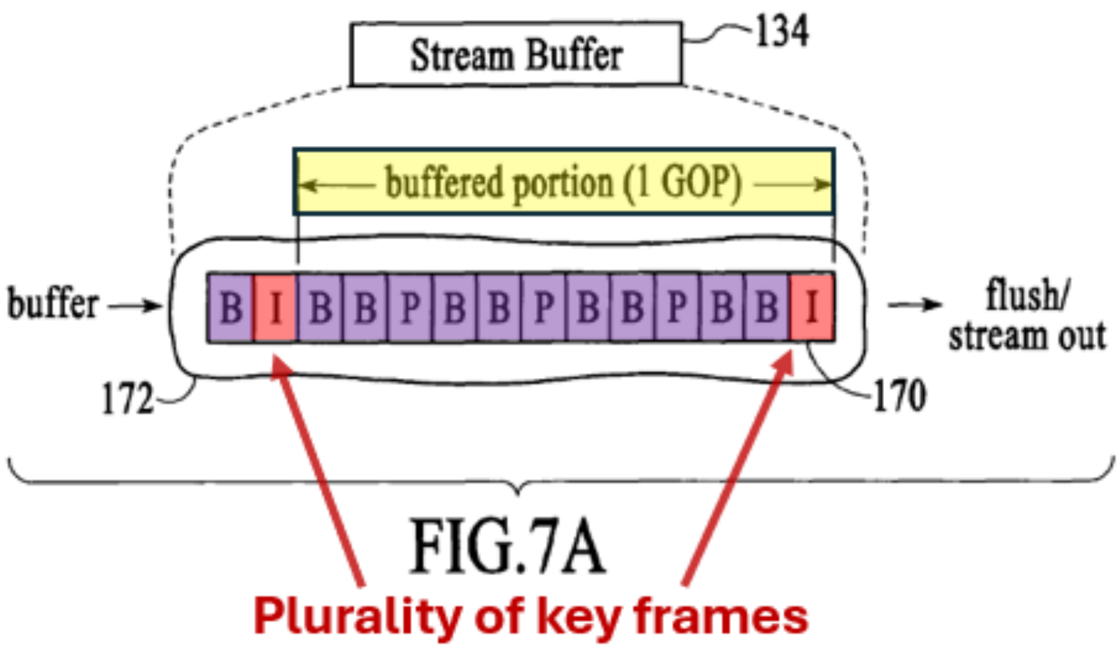


FIG.7A

e. Elements 1[D]/15[D]: [receiving]/[computer code for receiving] a channel change request from a client device receiving the first media stream, the channel change request corresponding to a switch from the first media stream to the second media stream;

200. Sherer-Wu-771 teaches Elements 1[D] and 15[D].

201. Sherer teaches "receiving a channel change request from a client device receiving the first media stream… to [] switch from the first media stream to the second media stream." Sherer teaches "servicing a channel change request from a client involves switching from providing the digital video content to the client via multicasting to providing the digital video content to the client via unicasting and continuing to provide digital video content to the client via unicasting until a pre-established condition is met." Ex-1007, Abstract. As shown in Figure 1B, a "channel change request (CCR) 116 [(red)] is made by client A [(yellow)] to change from channel 1 to channel 2." Ex-1007, 5:22-26.

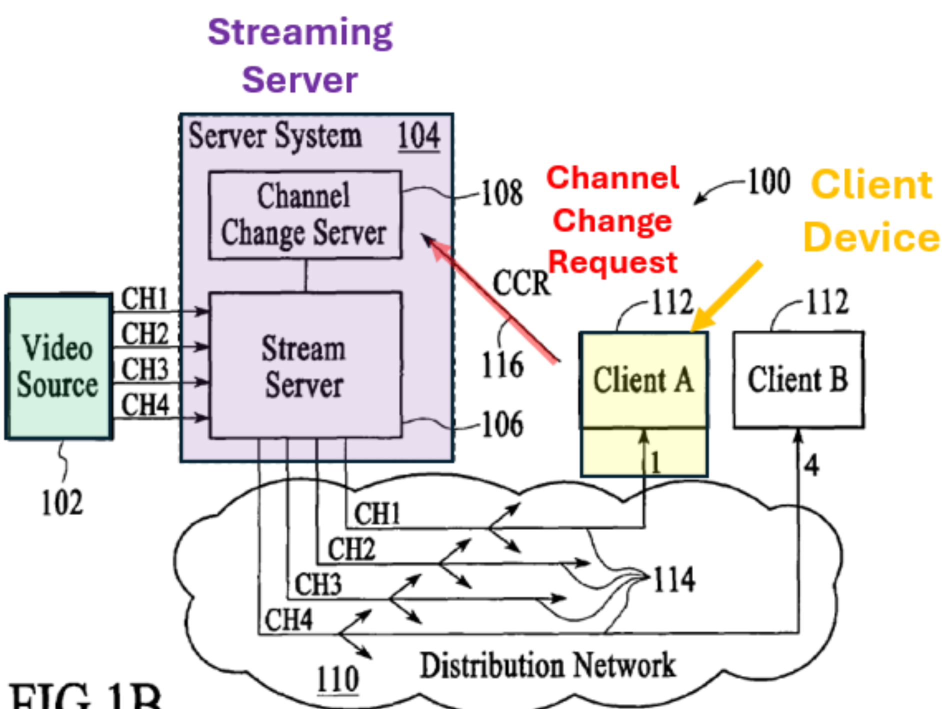


FIG.1B

202. As shown in Figure 5, the streaming server "detects" a channel change request initiated from the client device (red, step 610). Thus, Sherer's server system 104 receives a channel change request from a client device receiving the first media stream of channel 1, to switch to the second media stream of channel 2.

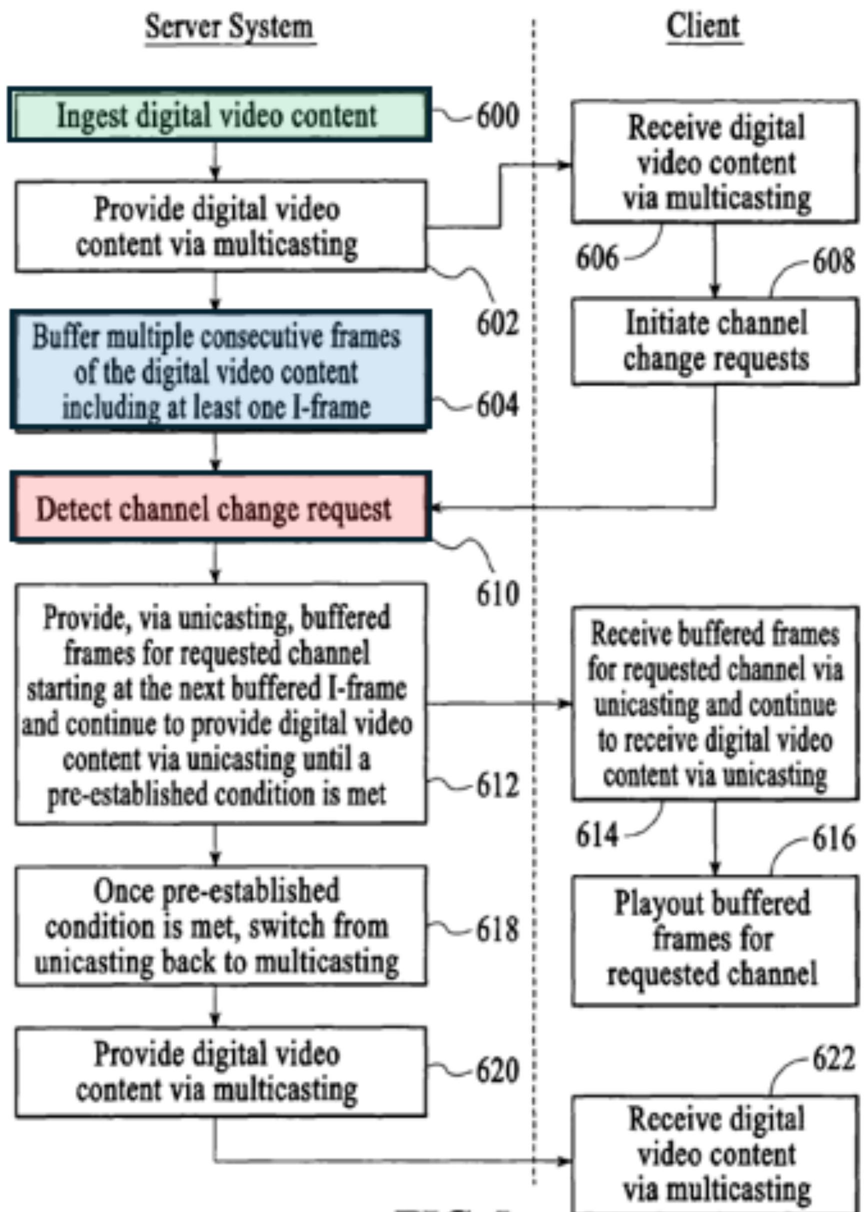


FIG.5

f. Elements 1[E]/15[E]: [identifying]/[computer code for identifying] a next available key frame associated with the second media stream, wherein the next available key frame is provided as the second key frame in response to the channel change request, wherein no predictive frames are transmitted at the beginning of the channel.

203. Sherer-Wu-771 teaches Elements 1[E] and 15[E].

204. Sherer "identif[ies] a next available key frame associated with the second media stream." As discussed in Section VII, Petitioner in the related litigation has proposed that "identifying a next available key frame associated with a second media stream" means "identifying, by the streaming server, by decoding or partially decoding the frame, a next available key frame associated with the second media stream." As shown in Figure 5, once a channel request is initiated by the client device, block 608, and received by the channel change server in the stream server, block 610 (red), "buffered frames for the requested channel are provided to the client via unicasting," block 612 (yellow). Ex-1007, 7:66-8:2.

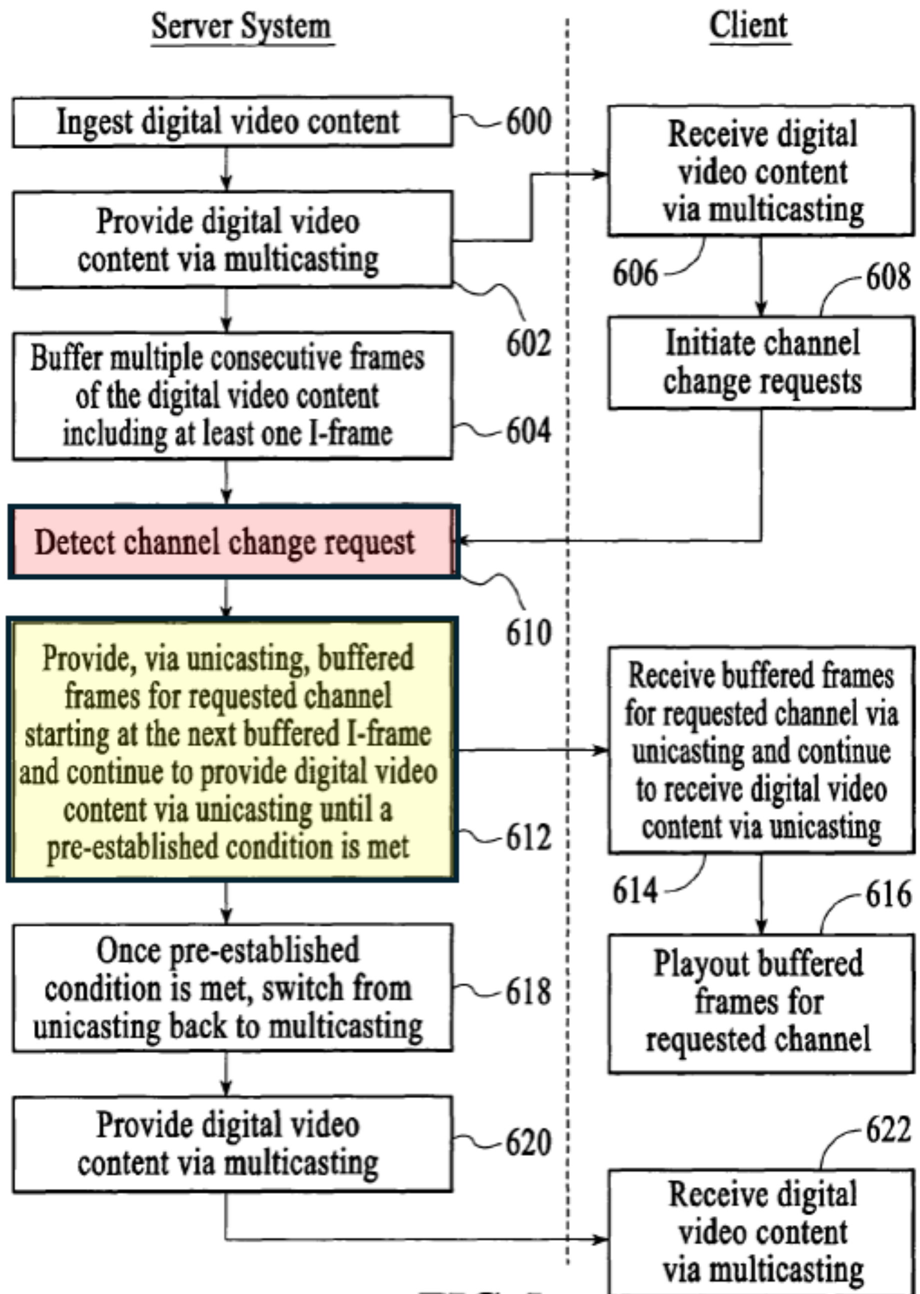


FIG.5

205. Sherer teaches that "the buffered frames are provided starting at the first buffered I-frame and the digital video content continues to be provided to the client via unicasting until a pre-established condition is met." Ex-1007, 8:2-5.

Sherer's streaming server identifies a next available key frame associated with a second media stream of the requested channel by buffering the first I-frame of the requested channel. A POSITA would have understood that Sherer's streaming server "decod[es] or partially decod[es] the frame" because the channel change server receives the media stream and converts the compressed video file into a format that can be displayed on various devices.

206. Sherer provides the "next available key frame… as the second key frame in response to the channel change request, wherein no predictive frames are transmitted at the beginning of the channel."[10] As discussed in Section VII, a "predictive frame" is "a frame that relies on information from a reference key frame." Ex-1001, 2:67-3:2. In response to the channel change request, Sherer's streaming server provides the second stream "***starting at the first buffered I-frame*** and the digital video content continues to be provided to the client." Ex-1007, 8:2-5. Sherer, as shown below in Figure 7B, the channel change request "begins with the first buffered I-frame and continues with the frames 170 (including B, P, and I frames) that consecutively follow the first I-frame." *Id.*, 9:23-26. For example,

[10] As discussed in Section X, Petitioner, in the related litigation, proposed that this limitation is indefinite. For the purposes of this proceeding, I have applied these limitations to mean that in response to a change request, the next available key frame is provided and there are not any predictive frames transmitted at the beginning of the channel.

Sherer's streaming server indicates the channel change boundary by "using a marker bit or an explicit signal ***at the first I-frame of a new unicast stream***" such that "once the client detects the start of the new unicast stream, the client ***resets its decoder and flushes from its buffer undecoded frames related to the previous unicast stream***." *Id.*, 6:18-24. Therefore, at the beginning of the second media stream channel the starting GOP frame (I frame) is the only frame transmitted in response to the channel change because the frames related to the previous channel are "flushed" from the buffer and not transmitted. Therefore, upon a channel-change request, the system immediately starts delivery at the indexed I-frame of the new channel requested.

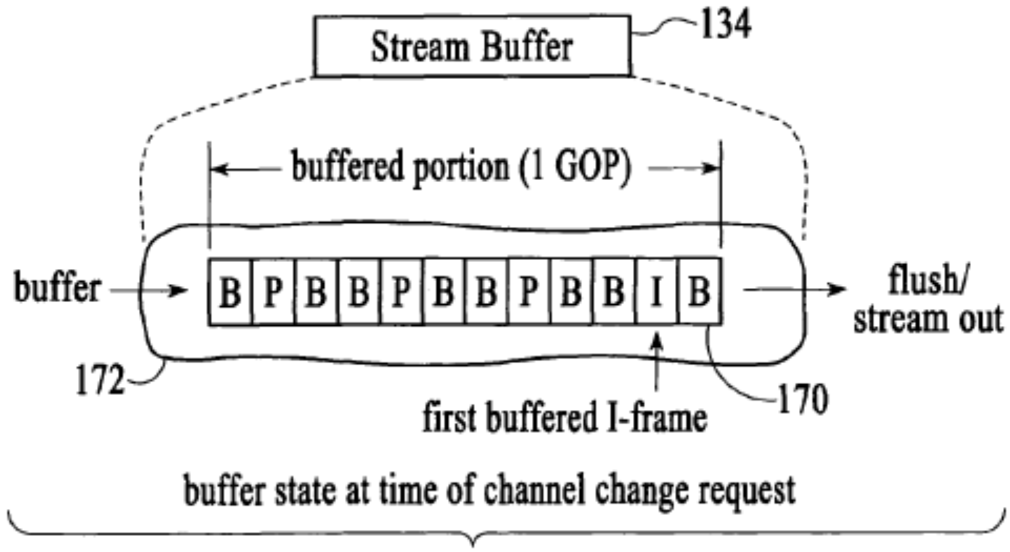


FIG.7B

207. Sherer teaches the recited order. As discussed in Section VII, Petitioner in the related litigation has proposed that elements 1[A] and 1[B] must precede 1[C], and 1[C], 1[D], and 1[E] must be performed in order. Ex-1021. As shown below in

Figure 5, Sherer receives media streams (elements 1[A]-1[B]/15[A]-15[B], green) and then determines the positions of a plurality of key frames in the first buffer and the second buffer (elements 1[C]/15[C], blue) prior to receiving a channel change request (elements 1[D]/15[D], red). Upon receiving the channel change request, Sherer's streaming server identifies and buffers the next available key frame for the requested second media stream (elements 1[E]/15[E], yellow).

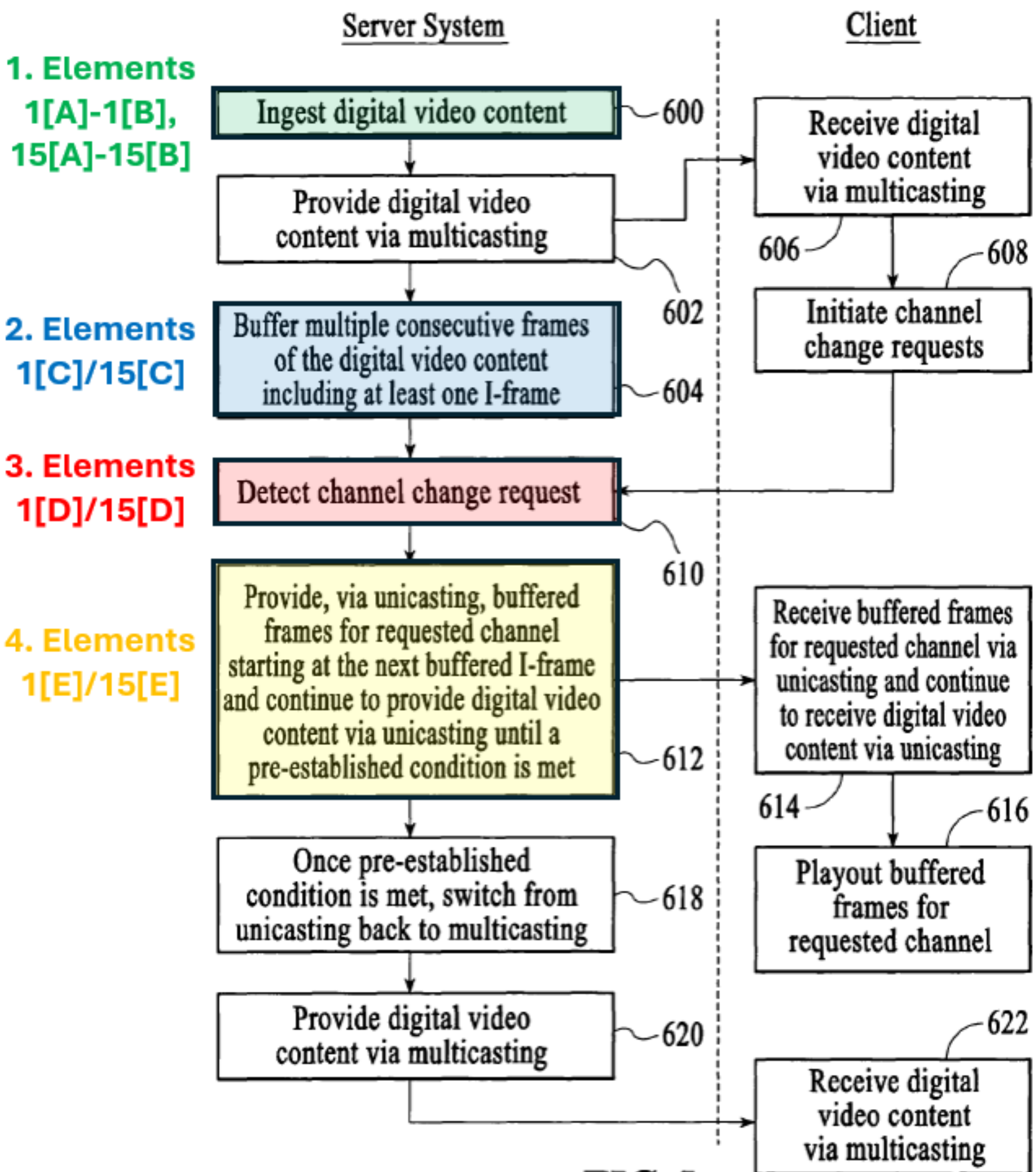


FIG.5

3. Dependent Claims 2, 4-6, 16, and 18-20

a. Claims 2/16: [The method of claim 1]/[The computer readable medium of claim 15], wherein the first media stream is a first video stream and the second media stream is a second video stream.

208. Sherer-Wu-771 teaches Claims 2 and 16.

209. Sherer teaches that the first and second media streams are video streams.[11] Sherer's streaming server 104, receives "multiple channels of ***digital video content***" from video source 102. Ex-1007, 4:56-57. Sherer teaches that "the digital video content includes real-time content from the content producers and/or distributors" either live or pre-recorded. *Id.*, 4:56-5:1. Sherer's digital video channels are first and second video streams.

b. Claims 4/18: [The method of claim 1]/[The computer readable medium of claim 15], wherein the key frames are identified by detecting the start of a GOP.

210. Sherer-Wu-771 teaches Claims 4 and 18.

211. Sherer teaches that the key frames are identified by detecting the start of a GOP. Sherer's "stream buffer buffers up to 1 group of pictures (GOP) worth of each ingested stream, where a GOP is defined as a series of multiple consecutive frames ***beginning with an I-frame*** and continuing until the last frame before the next I-frame." Ex-1007, 8:60-64. Sherer detects the key frames as the start of the GOP because Sherer's streaming server identifies and buffers a GOP "beginning with an I-frame."

[11] As discussed in Section VII, Petitioner proposed that claim 2 is invalid under 35 U.S.C. § 112, ¶4. PO has proposed that the plain and ordinary meaning should apply. For the purposes of this proceeding, I have applied this claim to mean that the first media stream is a first video stream and the second media stream is a second video stream because the media stream, as recited in claim 1, includes a plurality of frames.

c. Claims 5/19: [The method of claim 1]/[The computer readable medium of claim 15], wherein the key frames are identified by flagging each start of a GOP as a key frame.

212. Sherer-Wu-771 teaches Claims 5 and 19.

213. Sherer teaches that the key frames are identified by flagging each start of a GOP as a key frame. Sherer's "stream buffer buffers up to 1 group of pictures (GOP) worth of each ingested stream, where a GOP is defined as a series of multiple consecutive frames ***beginning with an I-frame*** and continuing until the last frame before the next I-frame." Ex-1007, 8:60-64. For example, Sherer's streaming server "us[es] a marker bit or an explicit ***signal at the first I-frame*** of a new unicast stream." *Id*., 6:18-24.

d. Claims 6/20: [The method of claim 1]/[The computer readable medium of claim 15], wherein the key frame is an I-frame.

214. Sherer-Wu-771 teaches Claims 6 and 20.

215. Sherer teaches that the key frame is an I-frame. Sherer's "stream buffer buffers up to 1 group of pictures (GOP) worth of each ingested stream, where a GOP is defined as a series of multiple consecutive frames ***beginning with an I-frame*** and continuing until the last frame before the next I-frame." Ex-1007, 8:60-64. Sherer teaches that "[a]n I-frame (MPEG2) … is a frame that presents a complete new image." Ex-1007, 1:46-48. As I described above in Section V.B, GOP begin with an I frame, which would have been understood as a "key frame."

4. Independent Claim 8

a. Element 8[pre]: A system:

216. Sherer-Wu-771 teaches the preamble of claim 8, to the extent it is limiting, for the same reasons explained in Section VIII.A.2.a (1[pre]/15[pre]). Sherer teaches a "***system*** for servicing channel change requests in a digital video network that is capable of distributing digital video content to clients via multicasting and unicasting." Ex-1007, 16:12-15 (Cl. 20).

b. Element 8[A]: [(i)] an interface configured to receive a first media stream including a first plurality of frames, [(ii)] a second media stream including a second plurality of frames, [(iii)] the first media stream received at a streaming server and maintained in a first buffer and [(iv)] and the second media stream received at the streaming server and maintained in a second buffer, [(v)] the streaming server configured to receive the first media stream from a media aggregation server, wherein a plurality of buffers at the streaming server are allocated on a per group of pictures (GOP) basis;

217. Sherer-Wu-771 teaches elements 8[A(i), A(iii), A(v)] for the same reasons explained in Section VIII.B.2.b (1[A]/15[A]) and elements 8[A(ii), A(iv)] for the same reasons explained in Section VIII.B.2.c (1[B]/15[B]). As I explained above, while element 8[A] recites the same limitations in a slightly different order than elements 1[A]-1[B]/15[A]-15[B] as shown in the table below, the scope and meaning of these elements are not substantively different. For clarity, I have highlighted the differences between the claims in the table below.

Claim 1	***Claim 15***	***Element 8[A]***
1[A] [(i)] receiving a first media stream including a first plurality of frames, [(ii)] the first media stream received at a streaming server and maintained in a first buffer, [(iii)] the streaming server configured to receive the first media stream from a media aggregation server, [(iv)] wherein a plurality of buffers at the streaming server are allocated on a per group of pictures (GOP) basis;	15[A] [(i)] computer code for receiving a first media stream including a first plurality of frames, [(ii)] the first media stream received at a streaming server and maintained in a first buffer, [(iii)] the streaming server configured to receive the first media stream from a media aggregation server, [(iv)] wherein a plurality of buffers at the streaming server are allocated on a per group of pictures (GOP) basis;	8[A] [(i)] an interface configured to receive a first media stream including a first plurality of frames and [(ii)] a second media stream including a second plurality of frames, [(iii)] the first media stream received at a streaming server and maintained in a first buffer and [(iv)] the second media stream received at the streaming server and maintained in a second buffer, [(v)] the streaming server configured to receive the first media stream from a media aggregation server, wherein a plurality of buffers at the streaming server are allocated on a per group of pictures (GOP) basis;
1[B] receiving a second media stream including a second plurality of frames, the second media stream received at the streaming server and maintained in a second buffer;	15[B] computer code for receiving a second media stream including a second plurality of frames, the second media stream received at the streaming server and maintained in a second buffer;	

218. Additionally, Sherer teaches an "interface to receive." As shown below in Figure 1A and Figure 3, Sherer's server system 104 ("streaming server," purple) includes a stream server 106 (brown), a channel change server 108 (orange), which

includes a processor 120 (pink), memory 143, stream buffer 134 (blue), buffer controller 136. Ex-1007, 7:6-15. Stream server 106 (brown) is the interface configured to receive the media streams because it "ingests the digital video content from the video source and controls the distribution of the digital video content into the distribution network." *Id.*, 4:29-31.

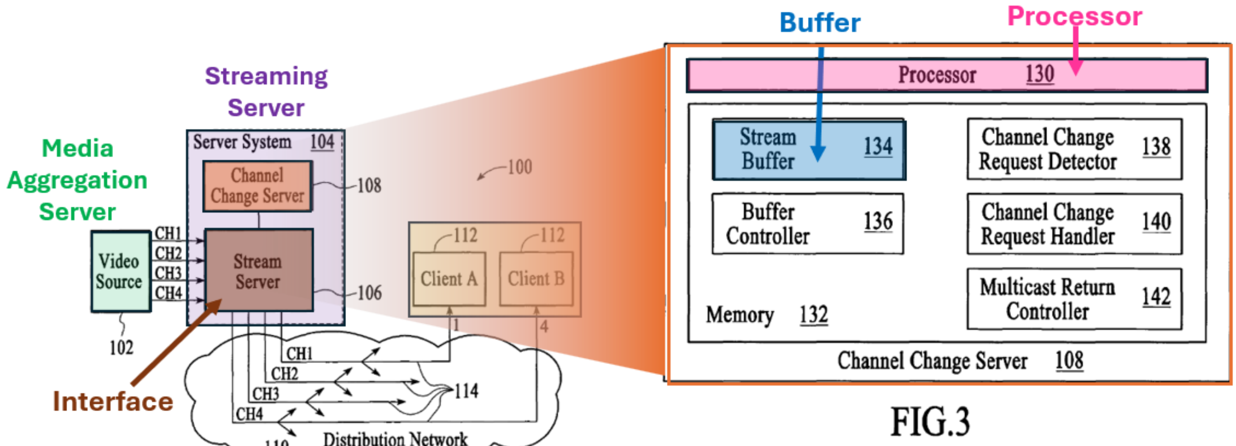


FIG.1A FIG.3

c. Element 8[B]: a processor configured to determine positions of a plurality of key frames in the first buffer and the second buffer, the plurality of key frames including a first key frame in the first buffer and a second key frame in the second buffer, wherein determining positions of the plurality of key frames includes determining at the streaming server a group of pictures (GOP) size N and identifying Nth frames following the first key frame as subsequent key frames;

219. Sherer-Wu-771 teaches element 8[B] for the same reasons explained in Section VIII.B.2.d (1[C]/15[C]). As I explained above, while element 8[B] use slightly different language as shown in the table below, the scope and meaning of

these elements are not substantively different. For clarity, I have highlighted the differences between the claims in the table below. Additionally, Mao teaches a "processor configured to determine," because Mao's BDT 12 processor 55 receives the frames for each channel and ***detects, establishes, and records*** the GOP start frame (I frame) for each channel. Ex-1005, 8:15-26.

Element 1[C]	***Element 15[C]***	***Element 8[B]***
[(i)] determining positions of a plurality of key frames in the first buffer and the second buffer, the plurality of key frames including a first key frame in the first buffer and a second key frame in the second buffer, [(ii)] wherein determining positions of the plurality of key frames includes determining at the streaming server a group of pictures (GOP) size N and identifying Nth frames following the first key frame as subsequent key frames;	[(i)] ***computer code for*** determining positions of a plurality of key frames in the first buffer and the second buffer, the plurality of key frames including a first key frame in the first buffer and a second key frame in the second buffer, [(ii)] wherein determining positions of the plurality of key frames includes determining at the streaming server a group of pictures (GOP) size N and identifying Nth frames following the first key frame as subsequent key frames;	***a processor configured to*** determine positions of a plurality of key frames in the first buffer and the second buffer, the plurality of key frames including a first key frame in the first buffer and a second key frame in the second buffer, wherein determining positions of the plurality of key frames includes determining at the streaming server a group of pictures (GOP) size N and identifying Nth frames following the first key frame as subsequent key frames;

220. Additionally, Sherer teaches a "processor configured to determine." As shown below in Figure 1A and Figure 3, Sherer's server system 104 ("streaming server," purple) includes a stream server 106 (brown), a channel change server 108

(orange), which includes a processor 120 (pink), memory 143, stream buffer 134 (blue), buffer controller 136. Ex-1007, 7:6-15. The buffer controller manages the buffering of digital video content that is streamed into the stream buffer. *Id.*, 7:37-38. Sherer teaches that "buffer controller 136, channel change request detector 138, channel change request handler 140, and multicast return controller 142 may be embodied as processor-executable instructions that are stored on processor-accessible memory media." *Id.*, 12:62-65. Therefore, the processor determines positions of a plurality of key frames in the first buffer and the second buffer.

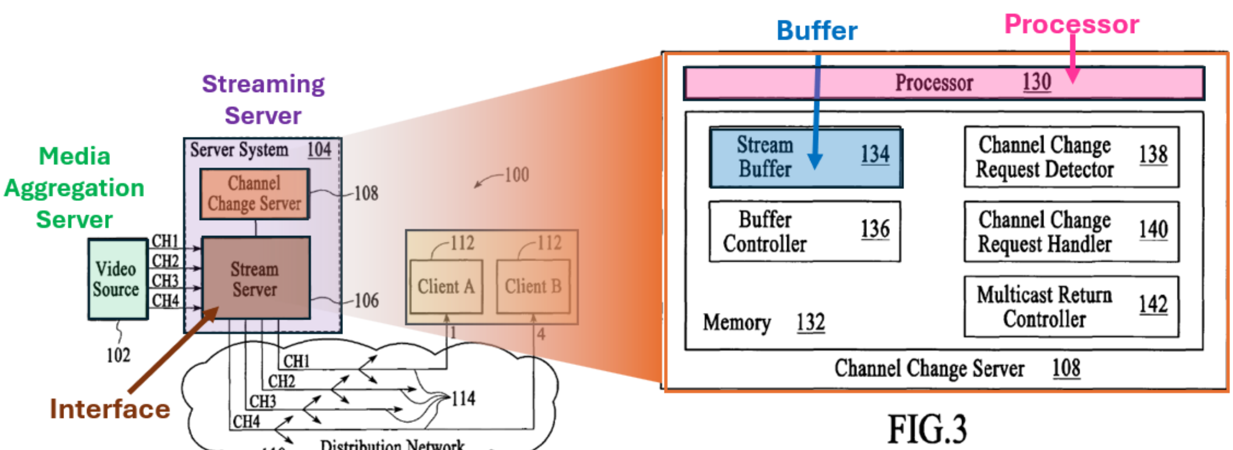


FIG.1A

FIG.3

d. Element 8[C]: wherein the processor is further configured to identify a next available key frame associated with the second media stream, wherein the next available key frame is provided as the second key frame in response to the channel change request, wherein no predictive frames are transmitted at the beginning of the channel change request.

221. Sherer-Wu-771 teaches element 8[C] for the same reasons explained in Section VIII.B.2.f (1[E]/15[E]). As I explained above, while element 8[C] use

slightly different language as shown in the table below, the scope and meaning of these elements are not substantively different. For clarity, I have highlighted the differences between the claims in the table below.

Element 1[E]	***Element 15[E]***	***Element 8[C]***
identifying a next available key frame associated with the second media stream, wherein the next available key frame is provided as the second key frame in response to the channel change request, wherein no predictive frames are transmitted at the beginning of the channel change request.	***computer code for*** identifying a next available key frame associated with the second media stream, wherein the next available key frame is provided as the second key frame in response to the channel change request, wherein no predictive frames are transmitted at the beginning of the channel change request.	***wherein the processor is further configured to*** identify a next available key frame associated with the second media stream, wherein the next available key frame is provided as the second key frame in response to the channel change request, wherein no predictive frames are transmitted at the beginning of the channel change request.

5. Dependent Claims 9 and 11-13

a. Claim 9: The method of claim 8, wherein the first media stream is a first video stream and the second media stream is a second video stream.

222. Sherer-Wu-771 teaches Claim 9 for the same reasons I explained in Section VIII.B.3.a (claims 2 and 16).

b. Claim 11: The method of claim 8, wherein the key frames are identified by detecting the start of a GOP.

223. Sherer-Wu-771 teaches Claim 11 for the same reasons I explained in Section VIII.B.3.b (claims 4 and 18) .

c. Claim 12: The method of claim 8, wherein the key frames are identified by flagging each start of a GOP as a key frame.

224. Sherer-Wu-771 teaches Claim 12 for the same reasons I explained in Section VIII.B.3.c (claims 5 and 19).

d. Claim 13: The method of claim 9, wherein the key frame is an I-frame.

225. Sherer-Wu-771 teaches Claim 13 for the same reasons I explained in Section VIII.B.3.d (claims 6 and 20).

IX. CONCLUSION

226. For the reasons set forth above, I believe that Claims 1, 2, 4-9, 11-16, and 18-20 of the '705 Patent are unpatentable in view of the prior art.

227. In signing this declaration, I recognize that it will be filed as evidence in a contested case before the Patent Trial and Appeal Board of the United States Patent and Trademark Office. I also recognize that I may be subject to cross-examination in the case and that cross-examination will take place in the United States. If cross-examination is required of me, I will appear for cross-examination within the United States during the time allotted for cross-examination.

228. I reserve the right to supplement my opinions in the future to respond to any arguments that the Patent Owner raises and to take into account new information as it becomes available to me.

229. I declare under penalty of perjury under the laws of the United States of America that the foregoing is true and correct.

Dated: November 5 2025

Respectfully submitted,

Brian W. Anthony, Ph.D.