

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

CISCO SYSTEMS, INC.,

Petitioner,

v.

VIDEO SOLUTIONS PTE. LTD.,

Patent Owner.

IPR2024-00194
U.S. Patent No. 9,204,099

DECLARATION OF DR. JACOB SHARONY

TABLE OF CONTENTS

I. INTRODUCTION	1
II. QUALIFICATIONS	1
III. BASIS OF MY OPINIONS AND MATERIALS CONSIDERED	5
IV. PATENT PRINCIPLES	6
V. SUMMARY OF CONCLUSIONS FORMED	10
VI. LEVEL OF ORDINARY SKILL IN THE ART	12
VII. SUMMARY OF THE '099 PATENT	12
VIII. SUMMARY OF THE ASSERTED PRIOR ART	20
a. Setton (EX1005)	20
b. Baker (EX1006)	30
c. Apostolopoulos (EX1007)	39
IX. CLAIM CONSTRUCTION	44
X. ANALYSIS OF PETITION GROUNDS	45
a. The Motivations to combine Setton and Baker are inherently deficient	47
b. The Motivations to Combine Apostolopoulos and Baker are Deficient	52
c. Elements [3.1.1], [3.1.2], [10.0]	55
d. Elements [3.4], [10.3]	63
XI. CONCLUSION	68

I, Dr. Jacob Sharony, hereby declare as follows:

I. INTRODUCTION

1. I have been retained by counsel for Video Solutions Pte. Ltd., the Patent Owner, as an independent technical expert in the above-identified Inter Partes Review (“IPR”) brought by Cisco Systems, Inc. (the “Petitioner”) challenging the validity of U.S. Patent No. 9,204,099 (“the ’099 Patent”).

2. I have been asked to provide opinions regarding the ’099 Patent and claims 3-6 and 10 (the “Challenged Claims”) that are challenged in the Petition for this IPR. This Declaration provides the opinions that I have reached to date regarding these matters.

3. I expect to testify regarding the matters set forth in this Declaration if asked to do so.

4. I am being compensated at an hourly rate of \$600 with reimbursement for actual expenses for my work related to this matter. I have received no additional compensation for my work in this case.

5. My compensation is not dependent on, and in no way affects, the substance of my statements in this Declaration, any testimony I may provide, or the ultimate outcome of the case.

II. QUALIFICATIONS

6. I am a founder and principal consultant at Mobius Consulting in Dix Hills, New York. In this role, I provide professional consulting services in mobile wireless strategy, technologies, systems, and applications. I am also an Adjunct Professor of Electrical Engineering at Columbia University in New York. In addition, I have provided technical consulting services on intellectual property matters, during which I have written expert reports and provided depositions and trial testimonies.

7. My academic credentials include a Bachelor of Science Degree in Electrical Engineering, magna cum laude, from Tel Aviv University, Israel, in 1979 and a Master of Science Degree in Electrical Engineering from Tel Aviv University, Israel, in 1984. In 1989, I received a Master of Business Administration (MBA) from Recanati Business School, Tel Aviv University, Israel. I went on to receive a Master of Philosophy in Electrical Engineering from Columbia University, New York, in 1991, and a PhD in Electrical Engineering from Columbia University, New York, in 1993.

8. I have worked or consulted for over 35 years in the fields of information communication technologies (ICT), wireless networks, mobile devices/applications, and data capture and sensing. I have authored and co-authored numerous technical papers. I hold over fifty patents in various fields of communications and networks including fiberoptics and mobile and wireless technologies. My employment history

following my graduation from Columbia University began at BAE systems in 1994, developing tactical mesh-based wireless network systems for the U.S. Department of Defense. During my time at BAE, I was an inventor on several U.S. patents.

9. From 1997 to 2005, I held positions at Symbol Technologies (acquired by Motorola Solutions). As Senior Director of Research and Development from 1997 to 2003, I initiated and led several research and development programs in wireless local area network (LAN) technologies, including mobile device management and security. As Senior Director of Technology Strategy and Development from 2003 to 2005, I was responsible for research and development of new mobile applications for delivering multimedia-rich content (video, audio, data) to mobile devices connected over heterogeneous networks. My work at Motorola/Symbol resulted in several U.S. patents.

10. In 2004, I founded Mobius Consulting, a consulting firm providing professional services in mobile wireless strategy, technologies, systems, and applications, including enterprise mobility, wireless communication networks, mobile embedded devices, device management, and mobile applications and services. In this capacity, I have worked with many companies in the mobile and wireless ecosystem, including service providers and operators, equipment vendors, and semiconductor companies. I am an inventor on numerous patents stemming from the work I have done at Mobius Consulting.

11. Between 2006 and 2016, I have held numerous roles as consultant, advisor, and director in organizations including Mobile-Access (acquired by Corning), Time Warner Cable, Center of Excellence in Wireless & IT, Powerwave Technologies, Sibeam, and Lattice Semiconductor.

12. My research centers around several fields, including wireless and mobile systems, ICT, and communication networks. Some of the representative publications and patents from such fields are listed in the following (a full list of my publications and patents is included in my *Curriculum Vitae (CV)*, attached):

- a. US 5,652,751 - Architecture for mobile radio networks with dynamically changing topology using virtual subnets;
- b. US 5,742,593 - On-line distributed TDMA/FDMA/CDMA link assignment in mobile radio networks with flexible directivity;
- c. Y. Li, T. Wang, J. Sharony, "Free-Space Optical Interconnects Using Connectivity-Enhanced Mesh-Based Networks," *Optical Engineering*, vol. 33, no. 5, 1994 pp. 1532-1542;
- d. J. Sharony, "A Multi-Star Optical Backbone Network Suitable for Distributed Microcellular Wireless Systems," *Proc. of IEEE GLOBECOM'96*, London, UK, 1996;
- e. US6,577,609 - Local addressing of mobile units in a WLAN with multicast packet addressing;
- f. US 7,668,201 - Bandwidth management in wireless networks;
- g. US 9,474,500 - Method and system for transfer of cardiac medical image data files; and
- h. US 9,807,657 - Small-cell gateway configured for multiple air interfaces.

13. I am a senior member of the Institute of Electrical and Electronic Engineers (IEEE). I have served on several government agency expert panels, including the National Science Foundation (NSF) and the National Institutes of Health (NIH).

14. I have delivered numerous presentations on industry-sponsored webinars and conferences.

15. Additional information regarding my education, technical experience and publications, including a list of the U.S. and international patents of which I am an inventor/coinventor, is included in my *Curriculum Vitae*.

III. BASIS OF MY OPINIONS AND MATERIALS CONSIDERED

16. In forming my opinion, I have relied upon my education, knowledge and experience in the information and communications field, and in particular my experience in research of multimedia communication and data network systems, including mobile and wireless systems and multimedia applications.

17. For my analysis, I reviewed and considered the following materials:

- The above-identified IPR Petition, along with all the ancillary documents and exhibits;
- The '099 Patent (EX1001);
- Prosecution history of the '099 Patent (EX1002);
- Declaration of A.L. Narasimha Reddy (EX1003);
- The Setton reference (US Pub. 2010/00287238) (EX1005);
- The Baker reference (US Pat. 8,581,958) (EX1006);
- The Apostolopoulos reference (US Pub. 2005/0080894) (EX1007); and

- The Wong reference (US Pub. 2009/0232129) (EX1009).

IV. PATENT PRINCIPLES

18. The opinions I express in this Declaration involve the application of my engineering and technical knowledge, and experience to the evaluation of certain prior art with respect to the '099 Patent.

19. I am not a lawyer, I have not been trained in the law of patents, and I do not opine on matters of law.

20. For the purposes of my analysis, I have applied the below understanding of patent law and current legal principles related to patent validity to my analysis, as have been provided to me by counsel.

21. I have been informed by counsel that in determining whether a patent claim under post-grant review before the United States Patent Office (PTO) is anticipated or obvious in view of the prior art, the PTO must construe the claim by giving the claim its broadest reasonable interpretation consistent with the specification as the claim terms and specification would be have been understood by a person of ordinary skill in the art ("POSITA").

22. I have been informed by counsel that the broadest reasonable interpretation is the plain meaning, i.e., the ordinary and customary meaning, given to the term by a POSITA at the time of the invention, taking into account whatever guidance, such as through definitions, may be provided by the written description

in the patent, without importing limitations from the specification. For the purposes of this review, unless noted otherwise, I have construed each claim term in accordance with its plain meaning, i.e., its ordinary and customary meaning under the required broadest reasonable interpretation.

23. I have been informed by counsel that a claim is anticipated under 35 U.S.C. § 102 if each and every limitation of the claim is disclosed in a single prior art reference, either expressly or inherently. I have been informed by counsel that inherent disclosure means that the claim feature necessarily flows from the disclosure of the prior art reference. I have been informed by counsel that a claim is unpatentable under 35 U.S.C. § 103 if the claimed subject matter as a whole would have been obvious to a POSITA at the time of the invention, which I have been instructed to treat at present as the priority date of the '099 Patent, which is February 1, 2012.

24. I have been informed by counsel that an obviousness analysis takes into account the scope and content of the prior art, the differences between the claimed subject matter and the prior art, and the level of ordinary skill in the art at the time of the invention.

25. Finally, I understand that I must consider any known objective evidence that might show nonobviousness of the application (sometimes referred to as secondary considerations of nonobviousness), such as long felt but unfulfilled

need for the claimed invention, failure by others to come up with the claimed invention, commercial success of the claimed invention, praise of the invention by others in the field, unexpected results achieved by the invention, the taking of licenses under the patent by others, expressions of surprise by experts and those skilled in the art at the making of the invention, and the patentee proceeded contrary to the conventional wisdom of the prior art. This secondary evidence must be tied specifically to claim features that are argued to be patentable, and not those already in the public domain. I have been informed by counsel that secondary considerations must be assessed as part of the overall obviousness analysis.

26. In determining the scope and content of the prior art, I have been informed by counsel that a reference is considered appropriate prior art if it falls within the field of the inventor's endeavor. In addition, a reference is prior art if it is reasonably pertinent to the particular problem with which the inventor was involved. A reference is reasonably pertinent if it logically would have commended itself to an inventor's attention in considering his problem. The fact that a reference relates to the same problem as the claimed invention supports use of the reference as prior art in an obviousness analysis.

27. To assess the differences between prior art and the claimed subject matter, I have been informed by counsel that 35 U.S.C. § 103 requires the claimed invention to be considered as a whole. This "as a whole" assessment requires

considering whether one of ordinary skill in the art at the time of invention, confronted by the same problems as the inventor and with no knowledge of the claimed invention, would have selected the elements from the prior art and combined them in the claimed manner.

28. In determining whether the subject matter as a whole would have been considered obvious at the time that the patent application was filed, by a POSITA, I have been informed by counsel of several principles regarding the combination of elements of the prior art. First, a combination of familiar elements according to known methods is likely to be obvious when it yields predictable results. Likewise, combinations involving simple substitution of one known element for another to obtain predictable results, a predictable use of prior art elements according to their established functions, applying a known technique to a known device (method or product) ready for improvement to yield predictable results, and choosing from a finite number of identified, predictable solutions to solve a problem are likely to be obvious. Thus, if a POSITA can implement a “predictable variation” in a prior art device, and would see the benefit from doing so, such a variation would be obvious. Also, when there is pressure to solve a problem and there are a finite number of identifiable, predictable solutions, it would be reasonable for a POSITA to pursue those options that fall within his or her technical grasp. If such a process

leads to the claimed invention, then the latter is not an innovation, but more the result of ordinary skill and common sense.

29. I have been informed by counsel that the “teaching, suggestion, or motivation” test is a useful guide in establishing a rationale for combining elements of the prior art. This test poses the question as to whether there is an explicit teaching, suggestion, or motivation in the prior art to combine prior art elements in a way that realizes the claimed invention. Though useful to the obviousness inquiry, I understand that this test should not be treated as a rigid rule. It is not necessary to seek out precise teachings; it is permissible to consider the inferences and creative steps that a POSITA (who is considered to have an ordinary level of creativity and is not an “automaton”) would employ.

30. I have been informed by counsel that when interpreting the claims of the '099 Patent, I must do so based on the perspective of one of ordinary skill in the art at the relevant priority date. I have been requested by counsel to assume that the earliest claimed priority date of the '099 Patent is February 1, 2012.

V. SUMMARY OF CONCLUSIONS FORMED

31. As part of my analysis of the documents of this case, I have arrived at certain conclusions, some of which I list below. The list is not exclusive; for example, further conclusions are provided throughout this Declaration.

32. The '099 Patent discloses a network topology that allows non-mesh connected endpoints and mesh-connected endpoints to experience the same functionality during a teleconference, which is not taught or suggested by the references, whether taken alone or in combination. Setton describes a network in which endpoints are either active participants or receivers (i.e., observers). Since the receivers are not active, they do not have control functionalities disclosed by the '099 Patent, nor do they have bidirectional communication with the active participants to select a full size video feed (rather, they receive aggregation of video feeds from the active participants rather than a selected feed), *unless* they request, from a registration server, to be promoted to the mesh network of active participants (and thus are no longer considered receivers or non-mesh connected). And while the Petition attempts to rely on a combination of Baker and Wong to supply missing elements, both Baker and Wong fail to supply such teaching. Indeed, Baker is silent about a control connection that communicates with an endpoint or proxy server to initiate transfer of the full size video feed to the non-mesh connected endpoint. In Baker, all endpoint participants send their full size videos to a network, and a participant endpoint can request, from the network, which of those full size videos to receive. The participant endpoint can then *locally* process and promote one of the received full size videos in their display. Moreover, Apostolopoulos provides a system for generating avatars and facilitating communication between endpoints in

a three-dimensional virtual environment. Apostolopoulos describes different network topologies, such as a full-mesh topology and a star topology. In the full-mesh topology, each collaborator (an endpoint) can generate its avatar. In the star topology, a central server is connected to all collaborators (endpoints), receives data from the collaborators, and *generates or renders an avatar* of a collaborator (one of the endpoints) using vantage points/perspectives of the data from each collaborator, then distributes the avatars to the collaborators. Although, Apostolopoulos suggests dynamically moving between the mesh and star topologies when a number of collaborators changes by using a hybrid topology.

VI. LEVEL OF ORDINARY SKILL IN THE ART

33. Consistent with the '099 Patent and the Challenged Claims, a person of ordinary skill in the art ("POSITA") at the time of the claimed invention would have had a degree in computer science, electrical engineering, or a similar field with at least 2-3 years of experience with data networks and video streaming systems or other comparable professional experience. Additional education could substitute for professional experience, or vice versa.

VII. SUMMARY OF THE '099 PATENT

34. The technology covered by the '099 Patent pertains to a teleconferencing system that allows multiple parties to participate in a teleconference regardless of their location. EX1001, Abstract. More specifically the

'099 patent discloses a system in which non-mesh connected endpoints and mesh-connected endpoints are treated equally and can participate in the teleconference as if they are all part of a mesh network. In other words, "non-mesh endpoint(s) user(s) can use [a] film strip user interface...to exercise the same control over...their video presentation as would a mesh connected user." *Id.*, 15:46-49.

35. As further recited in the '099 patent,

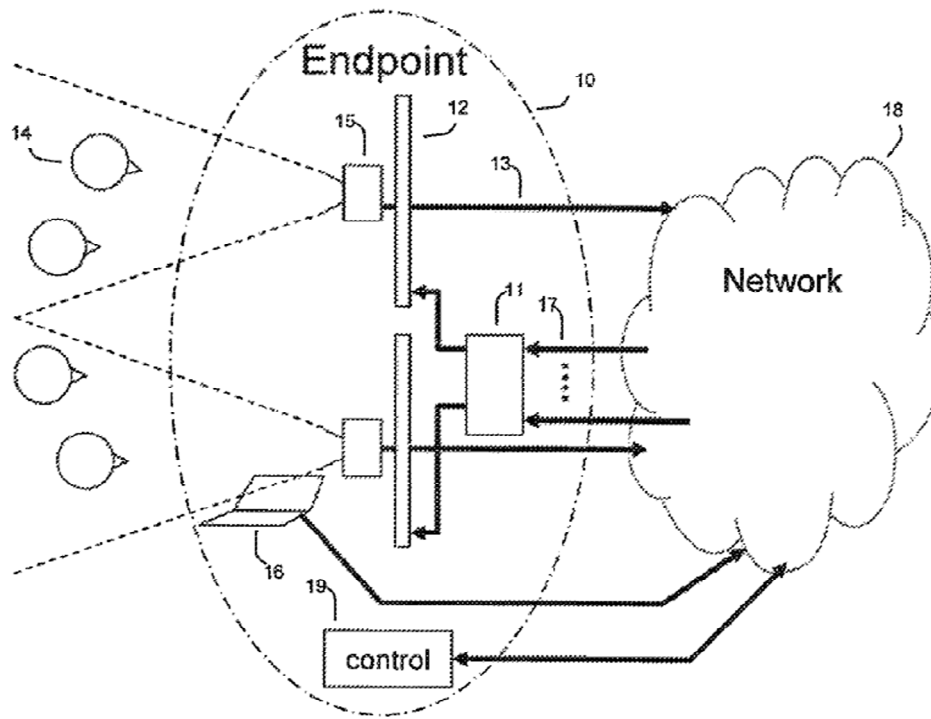
The proxy allows the endpoint 84 to participate in the conference in much the same way as if it were part of the mesh topology. For bandwidth considerations, it is undesirable to stream high definition video to the endpoint 84 from each mesh-connected endpoint. In accordance with embodiments of the invention it is only necessary to stream the full size feed from the endpoint currently promoted to the main window. Only reduced size feeds need [be] streamed to create the film strip, either to the endpoint itself or to a browser running on a computer co-located at the endpoint. The user can then select a frame on the film strip for promotion, and the promoted feed is then substituted for the current full sized frame feed.

Id., 17:47-58.

36. In response the user selection of a frame on the film strip for promotion, "a control connection" can "initiate transfer of the full size video feed to the non-mesh connected endpoint." *Id.*, 18:44-46. The non-mesh connected endpoint can therefore communicate with a mesh-connected endpoint to receive full size videos only when the non-mesh connected endpoint receives user input indicating a desire to promote one of the many video feeds in the film strip. Otherwise, the non-mesh connected endpoint may receive and display reduced-size video feeds until the user

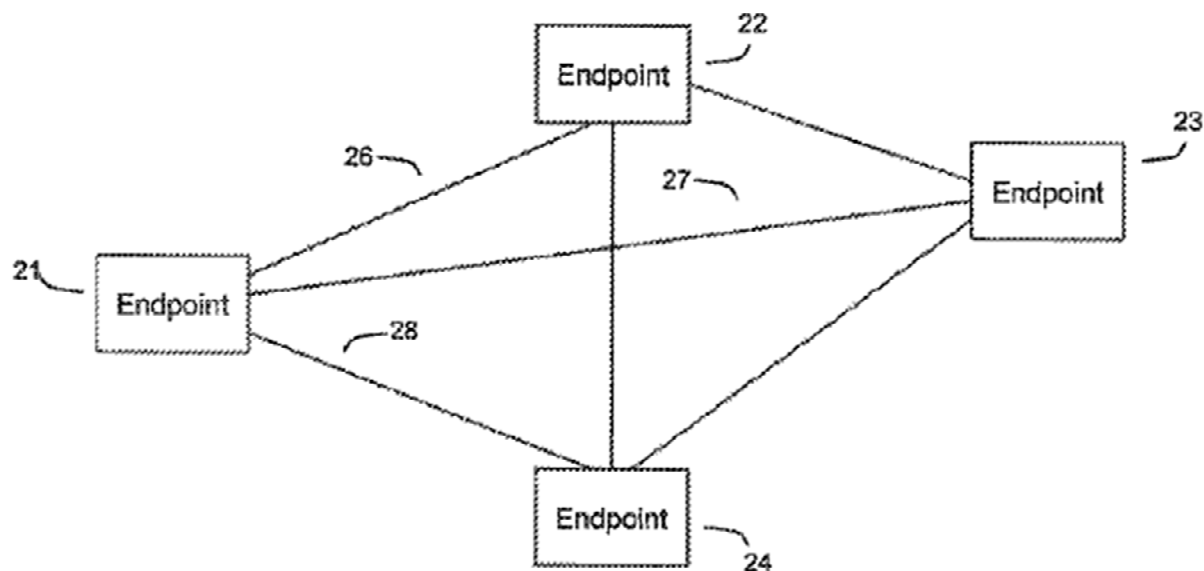
input is received. Such a configuration advantageously can reduce strain on network bandwidth as well as communication lines between any of the endpoints.

37. An example endpoint is illustrated in FIG. 1 (*see* endpoint 10), reproduced below. The endpoint can include a “collection of devices managed by a common Call Control software 19 having a single network address (e.g. IP address or URL) used for call set-up.” *Id.*, 5:4-8. In a teleconference, “Conference participants 14 sit before video cameras 15 and in front of monitors 12 [at the respective endpoint 10]. The monitors are controlled by display controller 11 and receive multiple video streams and collaboration data 17 from all other endpoints participating in a given conference.” *Id.*, 5:14-20. Each endpoint can have “a single display screen intended to show video or collaboration material” from all the conference participants, regardless of whether those participants are connected in a mesh or non-mesh network. *Id.*, 5:36-38.



Id., Fig. 1

38. One or more of the “endpoints in a multipoint call may be interconnected in a mesh configuration,” as shown in FIG. 2, reproduced below. *Id.*, 5:21-22. In a mesh configuration, “each endpoint is connected directly to each other endpoint in the conference. In such an arrangement each other endpoint has complete freedom to present video and other collaboration material uniquely and independently of the way it is presented at other endpoints,” as recited in the background section of the ’099 Patent. *Id.*, 2:36-40.

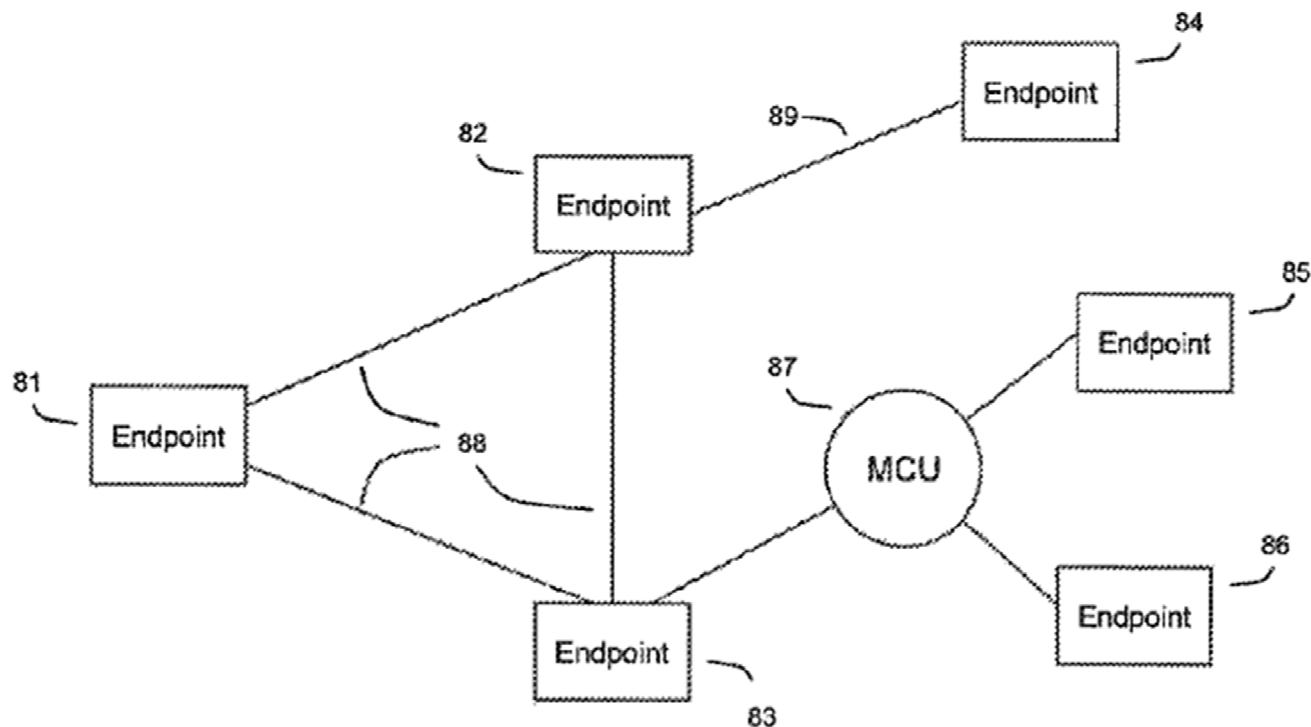


Id., Fig. 2.

39. Typical Multipoint Conference Units, or MCUs, would “consolidate...the video feeds from the end points and present...appropriate video images to each of the end points of all the video camera and collaboration material video feeds in the conference,” resulting in “limited independent control of what images are promoted to full-scale size” at each of the endpoints. *Id.*, 1:34-41. In contrast to these typical MCUs, the ’099 patent “facilitates endpoints interconnected in a mesh configuration, eliminating the capital and management cost of an MCU, although does not preclude star configuration.” *Id.*, 4:1-10. The ’099 patent therefore provides a solution for mixed topology networks in which “a number of endpoints are connected in a mesh connection, and at least one endpoint is connected to a mesh-connected endpoint with a non-mesh connection” so that “the non-mesh connected endpoint [can] enjoy the same features as the mesh-connected endpoints.” *Id.* Prior

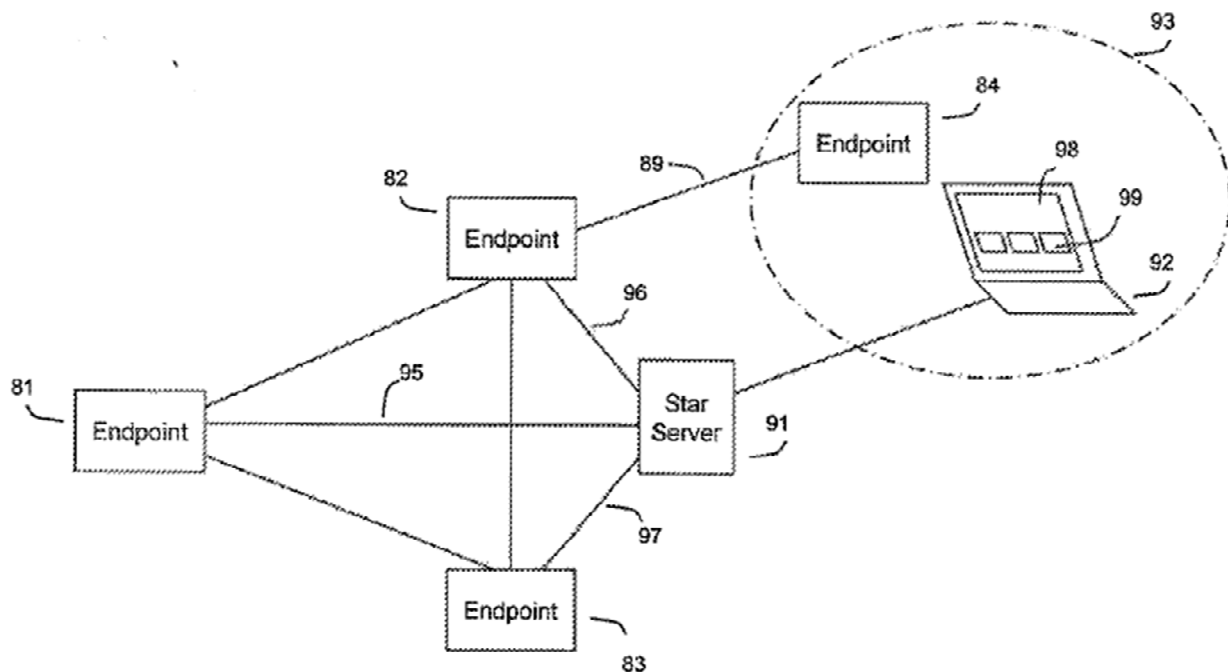
to the '099 patent, “non-mesh parties could participate in a mesh conference but without the personal control over video content and other benefits offered by the film strip [provided herein] in a mesh conference.” *Id.*, 15:38-41, 46-49.

40. FIG. 8 (reproduced below) illustrates a typical mixed topology conference call in which “endpoints 81, 82, and 83 are mesh connected” such that “each endpoint transmits all its media (video, audio and data) to all other mesh endpoints via multiple connections 88.” *Id.*, 15:58-61. A “star endpoint 84 connects to any one of the mesh endpoints” “or multiple star end points, e.g. 85 and 86, may connect to an MCU (multipoint control unit) which has one leg connected to one mesh endpoint.” *Id.*, 15:63-67, 16:1-2.



Id., Fig. 8.

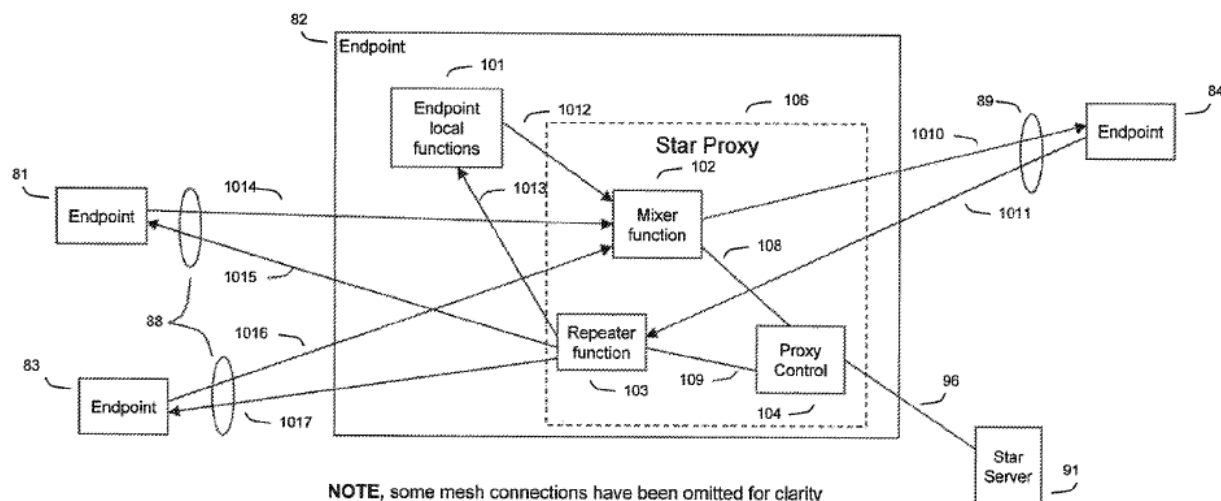
41. Similarly, FIG. 9 (reproduced below) illustrates “a multiparty conference” comprising “mesh-connected endpoints 81, 82 and 83 and a star endpoint 84 [(e.g., a non-mesh endpoint)] connected to mesh endpoint 82.” *Id.*, 16:6-8.



Id., Fig. 9.

42. FIG. 10, reproduced below, embodies the technology disclosed by the '099 patent: a mixed, hybrid topology allowing the same controls and functionality to both mesh-connected and non-mesh connected endpoints. Here, “a mesh endpoint 82 is shown” and “adapted to include a Star Proxy function...Once an endpoint, in this case endpoint 82, has been designated as a proxy for the star endpoint 84 it will instantiate local Mixer and Repeated functions which will act as a mesh proxy for the star endpoint. Control connections 108 and 109, for example software ‘handles’,

control the connection and operation of the Mixer and Repeater respectively.” *Id.*, 16:36-45.



Id., Fig. 10.

43. In operation, “[t]he Proxy Control function 104 will communicate with other Mesh Endpoints in the conference,” the “Mixer Function 102 will be connected to the transmit media connections (video, audio and data) of endpoints 81, 83 and local endpoint 101 via connections 1014, 1016 and 1012 respectively,” the “Repeater function 103 will be connected to the receive media connections (video, audio and data) of endpoints 81, 83 and local endpoint 101 via connections 1015, 1017 and 1013 respectively,” “the output of Mixer functions 102 is connected to star endpoint 84 via connections 1010,” and “the input of Repeater function 103 is connected to star endpoint 84 via connections 1011.” *Id.*, 16:47-67. Although FIG. 10 illustrates “connection of a single star endpoint to a mesh conference,” “any

number of similar Star Proxies 106...could be integrated into a single mesh endpoint or multiple mesh endpoints allowing any number of star endpoints to participate in a conference that has a mesh topology at the core.” *Id.*, 17:17-22.

44. The configuration and integration of mesh and non-mesh topology, as shown by FIG. 10 and described throughout the '099 patent, allow for *all endpoints*, both mesh-connected and non-mesh connected, “to participate in the conference in much the same way as if it were part of the mesh topology” such that the “non-mesh connected endpoint is able to benefit from the advantage of being able to view all the video feeds at reduced size in the film strip and select any one of the feeds for promotion without consuming the bandwidth on the link 89 that would be required to carry all the high definition feeds to the endpoint 84 simultaneously.” *Id.*, 17:47-66.

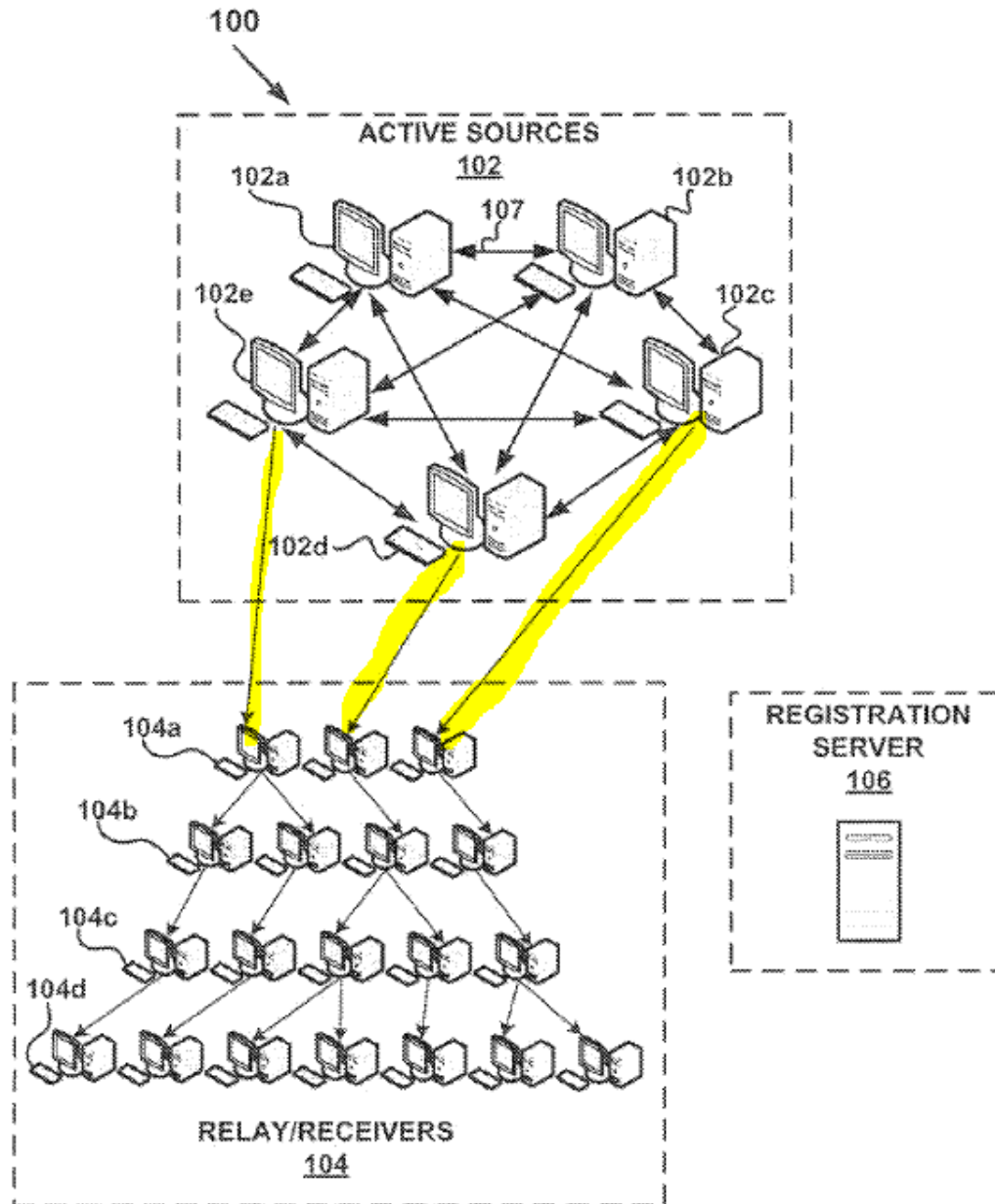
VIII. SUMMARY OF THE ASSERTED PRIOR ART

a. Setton (EX1005)

45. In contrast to the '099 Patent (*supra*, Sec. VII), the aim of the Setton reference is to permit for video feeds from mesh-connected endpoints to be rebroadcasted to non-mesh connected endpoints. Setton demonstrates that a subset of all users in a conference may generate content that is then distributed to another subset of all the users, where the other subset of users is merely observing the conference. In other words, the observers are not speaking or sharing content in the

conference. The subset of users generating content is therefore separate from the subset of users receiving and observing the content during the conference. Setton designates these content creators as active sources 102 and the mere observers or recipients of the content as relay/receivers 104, as illustrated in FIG. 1, reproduced below. *See* EX1005, [0014], [0016], [0018].

46. The ability of the relay/receivers 104 to merely receive content from the active sources 102 is represented in Setton's FIG. 1, by the highlighted unidirectional arrows. In my opinion, a POSITA would understand that because Setton's FIG. 1 clearly shows bidirectional arrows between and amongst the active sources 102 but not with any of the relay/receivers 104, Setton intended for communication between the active sources 102 and the relay/receivers 104 to be only unidirectional.



Id., Fig. 1 (Annotated).

47. As Setton describes, “the bi-directional arrows, typically shown as 107, denote packet exchanges between the active sources [102]” and “the aggregated media stream [from the active sources 102] is transmitted to the relay/receivers 104.” *Id.*, [0023]. The active sources 102 communicate back and

forth with each other, but the relay/receives 104 lack the capability to have open, bidirectional communication with any of the active sources 102, which teaches away from a network topology envisioned by the '099 patent in which all endpoints, whether mesh-connected (Setton's active sources 102) or non-mesh connected (Setton's relay/receivers 104), communicate with each other and control their respective video displays during the conference as if they are all part of a mesh-connected network topology.

48. Moreover, although Setton describes an aggregated stream being transmitted to the relay/receivers 104, once the aggregated stream is transmitted, the relay/receiver 104 cannot make any additional requests for conference data to any of the active sources 102 (as demarcated by the unidirectional arrows between the active sources 102 and the relay/receivers 104). Setton is also silent as to *how* the aggregated stream is displayed at each of the relay/receivers 104. Because the relay/receivers 104 receive the aggregated stream only once (and cannot request the conference data from any of the active sources 102 during the conference), Setton suggests, absent explicit description, that the aggregated stream is in full, high resolution. Even if a user at the relay/receiver 104 were able to select a video for promotion at their endpoint (relay/receiver 104) using the Setton network topology, the user's endpoint would have no communicative ability to request a full size version of the video from any of the active sources 102. The user's

endpoint would have to rely on the aggregated stream initially sent by an active source to carry out the user's intent of promoting the selected video at the endpoint's display. Thus, Setton's arrangement would consume large bandwidth (and therefore be wasteful) as the unidirectional links between the active sources 102 and the relay/receivers 104 are required to carry all the high-definition feeds to the relay/receivers 104 simultaneously. This is in stark contrast to the '099 Patent, where only the promoted feed is sent at high definition. *See* EX1001, 17:59-64.

49. Setton's network topology is framed in a way that ensures minimal to no network burdens on only those active sources 102 (e.g., mesh-connected endpoints) and thus leaves the relay/receivers 104 (e.g., non-mesh connected endpoints) at the mercy of potential lag or poor network communication and connection quality. Setton explains that "sharing in real time of K distributed media streams to a large set of users, using only one control plane" is "more efficient than using K independent multicast sessions, and such an approach avoids competition among the sources for the peer-to-peer network resources." EX1005, [0030]. Furthermore, according to Setton,

By having members of active sources 102 aggregate the media stream, embodiments in accordance with the present invention create several versions of a "super-stream", which can be transmitted as a unique stream to the rest of the active sources. Additionally, by limiting the active sources to only those peers who generate content, are sufficiently constraint burdened, or who opt to be included as an active source, embodiments in accordance with the present invention

drastically limit the number of hops a packet needs to go through before reaching any of the peers. As a result, embodiments in accordance with present invention limit the delay required transmit content through the peer-to-peer network (e.g., the first portion of the peer-to-peer network [which includes the active sources 102]).

Id., [0019].

50. Because Setton's non-mesh connected endpoints, the relay/receivers 104, are treated differently than mesh-connected endpoints, the active sources 102, Setton's framework does not permit for all conference participants to be treated equally, as contemplated by the '099 Patent. The '099 Patent, on the other hand, repeatedly describes all participants in the conference as being displayed on a film strip such that any participant, whether they are mesh-connected or non-mesh connected, can control how and when a video feed from the film strip is promoted and rendered in full size/quality/resolution at the respective endpoint. EX1001, Abstract, 3:11-22.

51. Setton has the opposite goal for non-mesh connected endpoints, in that it describes participants at those endpoints as only an "audience" that *observes* the "active" or "speaker" participants at the mesh-connected endpoints. EX1005, [0014]. Setton treats non-mesh connected endpoints differently from mesh-connected endpoints such that the non-mesh connected endpoints *do not have the same controls* as the mesh-connected endpoints, which counters at least one of the goals of the '099 Patent. As a result of treating the non-mesh connected endpoints

differently from the mesh-connected endpoints, “content generating or constraint burdened peers are defined to active sources and, as a result, do not suffer from unacceptable delay issues” whereas “listeners of the teleconference (relay/receivers) can tolerate a higher delay, on the order of a few seconds, as long as the media streams originating from the different speakers remain synchronized.” *Id.*, [0028]. Setton teaches different quality of service requirements for active, mesh-connected endpoints versus relay/receivers, non-mesh connected endpoints, which is not contemplated by the '099 Patent.

52. Instead of treating all users as equal participants from the onset of the conference, Setton allows for a relay/receiver to be promoted to an active source if that user desires to become a participant. If the user desires to stop participating and to merely observe the conference, they can also request to be demoted from active source (e.g., mesh-connected) to relay/receiver (e.g., non-mesh connected) and thus experience lesser quality and limited or no controls during the conference. Setton makes clear that a relay/receive *may never* generate content unless it requests to change its designated role to become an active source. *See Id.*, [0014], [0026]. In Setton’s system, “members of the peer-to-peer network [can] change their membership between being a member of said first portion of said peer-to-peer network (e.g., active sources 102) and being a member of the second portion of said peer-to-peer network (e.g., relay/receivers 104).” *Id.*, [0020]. Setton’s system

includes “a registration server 106...for maintaining a list of active sources 102” and “an approximate list of receiver/relay peers 104.” *Id.*, [0026]. The registration server 106 is akin to a gatekeeper for the active sources, and thus governs who, when, and how a relay/receiver may transition into an active source.

53. For example,

When a new peer sends a request to join the session, the new peer sends a request to registration server 106 indicating whether or not it is an active source. If the new peer is an active source, it will receive, as a response, information allowing the new peer to connect to the other active sources 102 as a member of the active sources 102. If the new peer is not an active source, it will receive, as a response, information allowing it to connect to relay/receivers 104. Registration server 106 also assists peers, in a similar fashion, when they transition between being a member of active sources 102 and relay/receivers 104.

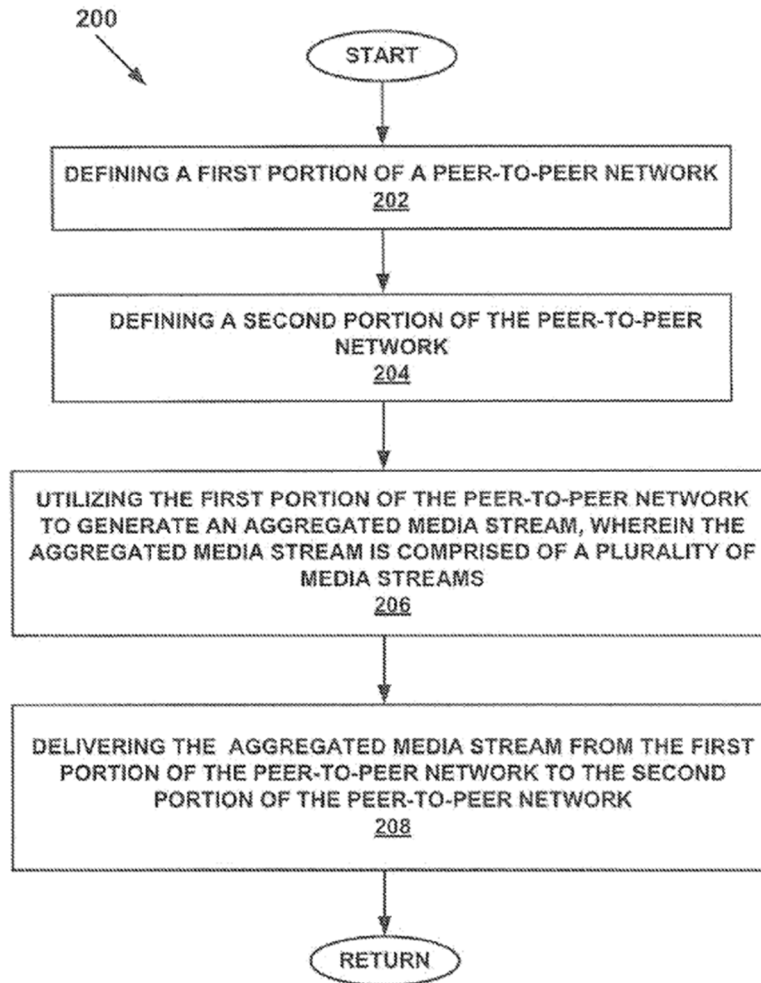
Id.

54. The techniques and framework of Setton’s network topology are intended to “accommodate the different quality of service requirements of the peers” to ensure that *only the active sources avoid unintentional delay issues or other burdens on their participation in the conference*, while the mere observers to the conference (e.g., the non-mesh connected endpoints) are at the mercy of delays and bandwidth constraints throughout the conference. *Id.*, [0028]. Such shifting of burdens from one group of users to another was not contemplated by the ’099 Patent. In fact, the ’099 Patent serves to solve a different problem in which *all users*, regardless of whether they are actively participating (e.g., mesh-connected)

or merely observing (e.g., non-mesh connected) in the conference, are treated equally in a mixed network topology. Moreover, the '099 Patent's topology allows for bidirectional communication between the non-mesh connected endpoints and the mesh-connected endpoints to alleviate any burden that may arise from delays, limited bandwidth, or other network constraints on any of the conference participants, which is seemingly not addressed by Setton.

55. Moreover, Setton describes “the aggregated media stream” as a single stream that is generated and transmitted by Setton's active sources to the relay/receivers. *Id.*, [0014], [0022]-[0023]. Setton is silent about whether any of the active sources receive requests for, generate, and/or transmit individualized aggregated media streams to any of the relay/receivers. *Id.*, [0018]-[0019], [0022]-[0023]. For example, FIG. 2 (reproduced below) provides an example of Setton's “process [200] for peer-to-peer media streaming,” thereby identifying the active sources (“first portion”) and the relay/receivers (“second portion”) (202-204). In this process, the active sources generate an aggregated media stream (206) and deliver such stream to the relay/receivers (208). *Id.*, [0015]-[0016], [0022]-[0023], FIG. 2. The process 200 is silent about whether any input, controls, or requests are generated by the relay/receivers (“second portion”) related to the content transmitted in the aggregated media stream. The process 200 is also silent about

whether any of the relay/receivers transmit media streams to any of the active sources.



Id., FIG. 2.

56. Although Setton describes using different protocols to generate and transmit varied versions of the aggregated stream, this discussion highlights differences between prioritized media streams transmitted among the active sources and a lower priority media stream broadcast to the relay/receivers. *Id.*, [0014], [0019], [0028]-[0030]. For example, Setton describes that versions of an

aggregated “super-stream” are generated by “members of active sources 102” and “transmitted as a unique stream to the rest of the active sources.” *Id.*, [0019].

57. Setton also does not describe the actual content of the aggregated media stream being transmitted to the relay/receivers, just that it “is comprised of a plurality of media streams.” *Id.*, Abstract, [0019]. As indicated above, Setton also does not describe how the aggregated media stream is presented at the relay/receivers.

58. As mentioned, Setton requires a relay/receiver to be promoted to an active source if that relay/receiver desires to participate in the conference. *Id.*, [0026]. To achieve this, Setton discloses a registration server 106 that manages the network configuration 100 and designation of endpoints as the active sources 102 and the relay/receivers 104. *Id.*, [0014], [0026], FIG. 1. A “speaker” who stops participating in the conference can transition from being an active source to a relay/receiver via the registration server 106. *Id.*, [0026]. Likewise, “a member of the audience” seeking to “join[] the discussion” would need to transition from being a relay/receiver to being an active source via the registration server 106. *Id.*

b. Baker (EX1006)

59. Although the '099 patent is generally directed to a technical solution for videoconferencing in mixed network topologies with mesh and non-mesh endpoints, the Baker reference is focused on presenting and navigating between

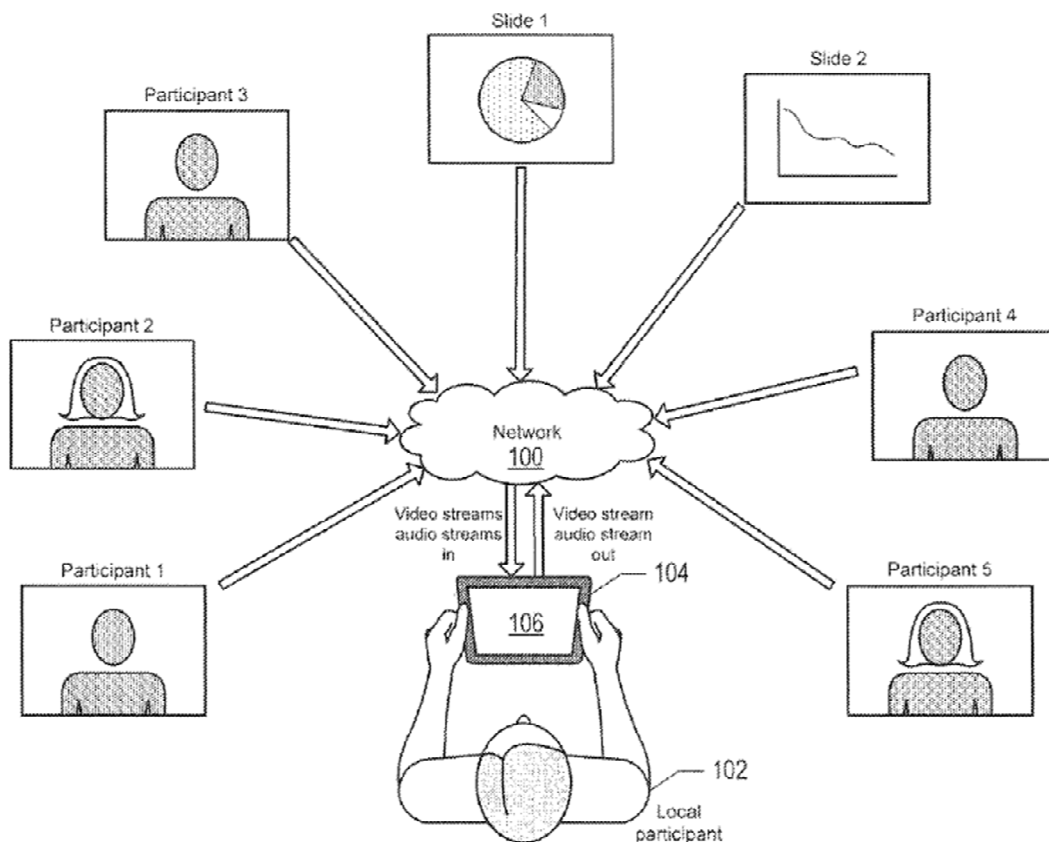
streams from multiple videoconference participants on a portable electronic device (“PED”). *See* EX1006, Abstract.

60. The Baker reference discloses a system in which each participant endpoint in a conference sends full rate/resolution video streams to a network and then a participant endpoint (e.g., the PED) requests only specific videos from the network to locally process and display. The data is received at full rate/resolution and resides at the participant endpoint so that the endpoint can locally select and promote a video for display *without* requesting a full rate/resolution version of the video from another endpoint. The ’099 Patent, on the other hand, sends low rate video and audio data to all of the participant endpoints and then permits for any of the endpoints to select one of the low rate videos to cause initiation of a request to a mesh-connected endpoint to transmit back a full rate/high resolution version of the selected video.

61. Contrary to the ’099 Patent, Baker’s participant endpoints do not have the ability to request high-resolution versions of selected video feeds from another endpoint, primarily because each endpoint receives full versions of the video feeds (from Network 100, *see id.*, FIG. 1) and locally manipulates the full versions of the video feeds for presentation at the endpoint. Therefore, Baker teaches changing a display at the endpoint using the data that is readily and locally accessible at the endpoint. Baker further does not shed light on *how* it locally changes the display at

an endpoint, which further does not support an assumption that Baker can provide functionality for the endpoint to *request* a full version of a video feed for promotion at the endpoint's display.

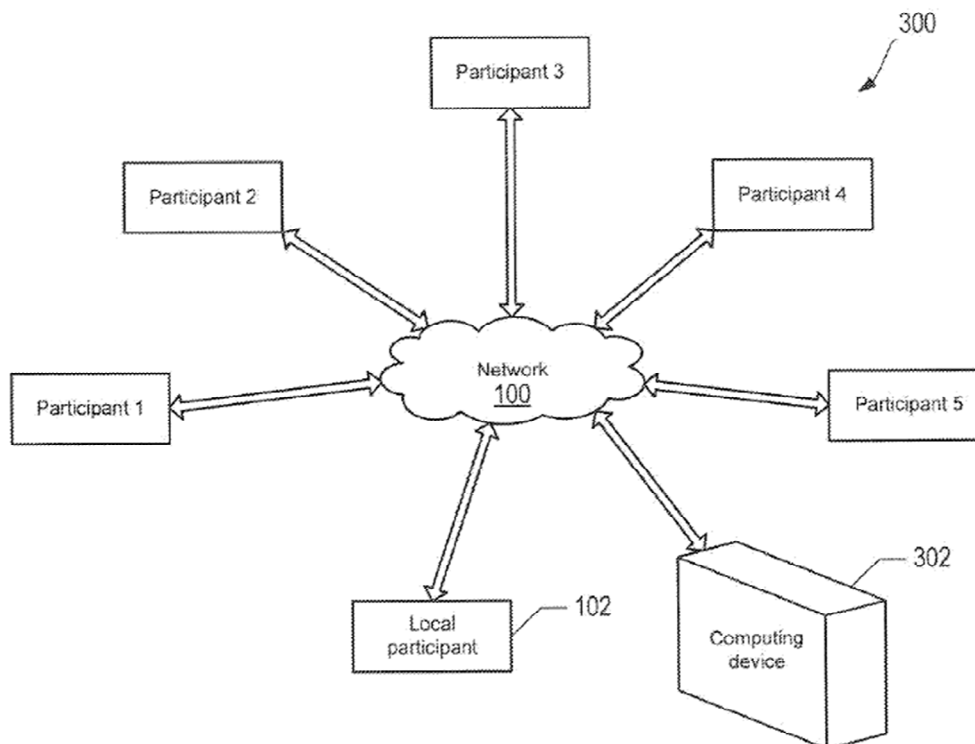
62. Baker's FIG. 1, reproduced below, illustrates how "video and audio streams generated by each of the remote participants are sent to the network 100 and received by the PED 104" of an example participant 102. EX1006, 2:43-46. Baker's primary embodiment appears in FIGs. 1-10, showing a PED 104 of a participant 102 that receives full resolution streams (e.g., video and audio streams) for each other conference participant over network 100 and then locally determines how to present the streams on the display 106. *Id.*, 2:31-49, 3:6-8, 4:2-4, 4:25-30, 4:48-57, FIGS. 1-4.



Id., Fig. 1.

63. Although Baker is ambiguous as to whether a server, another participant endpoint, and/or any other computer system is part of the network 100 to facilitate transfer of the video and audio streams amongst all the conference participants, Baker's system can include, in some embodiments, a centralized computing device to "receive...the video and audio streams from each of the participants" then "sends to each participant the video and audio streams of the other participants." *Id.*, 3:62-67, 4:1-4. This configuration offered by Baker suggests a client-server network, not a hybrid network topology as contemplated by the '099 Patent.

64. FIG. 3 (reproduced below) illustrates another embodiment of the Baker technology: a client-server network 300. In this embodiment of Baker, “the server-side application sends the video and audio streams generated by the remote participants 1-5 to the local participant 102 and sends the video and audio streams generated by the local participant 102 to the participants 1-5. The client-side application run by each remote participant displays the video streams, or a subset of the video streams, and plays the combined audio of the other participants and may allow the participant to arrange the viewing areas of the other participants as desired.” *Id.*, 4:5-13. In this client-server network topology a participant endpoint requests the data streams from the a central server (e.g., computing device 302) then locally changes *what and how* the received data streams are displayed at the endpoint.



Id., Fig. 3.

65. Baker further suggests a “video conference...conducted using a peer-to-peer network,” in which the computing device 302 shown above in FIG. 3 is not part of the system. *Id.*, 4:19-20. In the peer-to-peer topology, all participants exchange data with each other at the onset of joining the conference, such that the conference data is locally maintained (at full rate) and available for promotion at each of the participant endpoints.

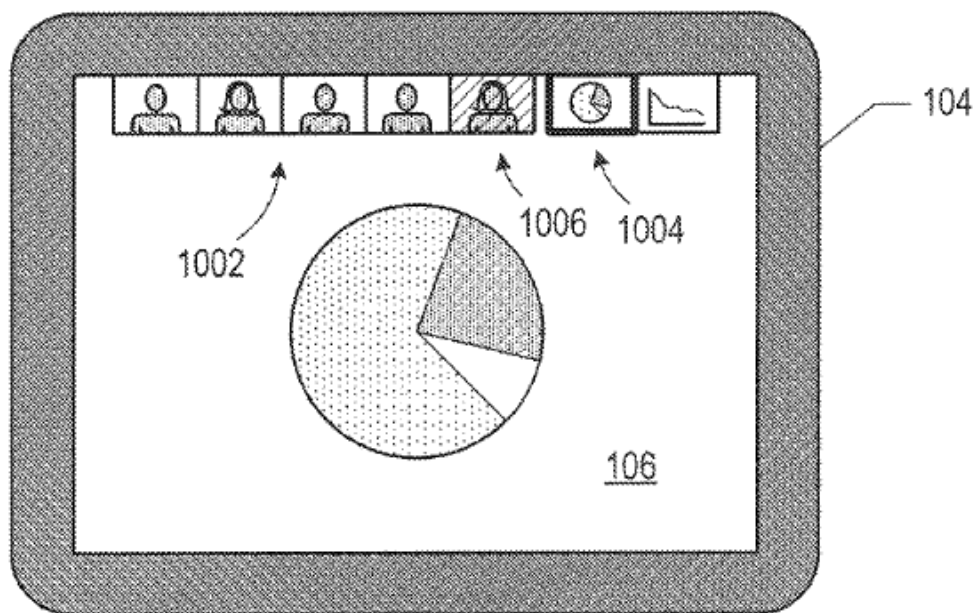
66. Baker describes:

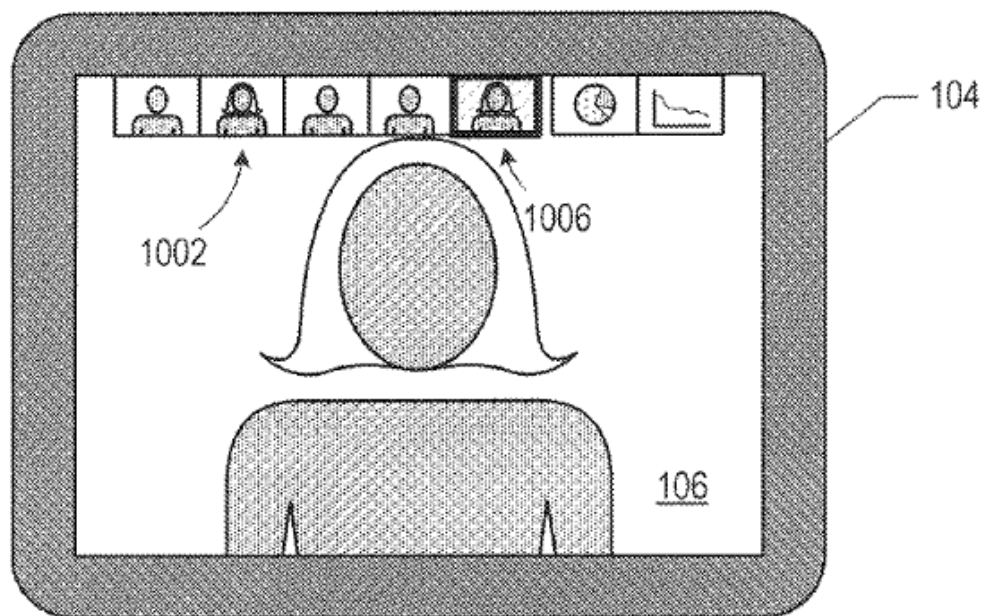
when a video conference is established...local participant 102 has access to each of the remote participants computing devices to download the video and audio streams generated by the remote participants....When the second participant accepts the conference with the first participant, the video conferencing applications exchange

video and audio streams. A third participant can join the video conference where each participant receives video and audio streams from the other two participants and sends locally generated video and audio streams to the other two participants.

Id., 4:30-47.

67. FIGs. 10A and 10B (reproduced below) of Baker further illustrate a display of the data from each of the participants from the perspective of the local participant's device 104 (e.g., PED). As shown, the participant locally modifies the display of participants at their device 104: "[t]he local participant can...move the display 106 to view participant 5....Alternatively, the local participant can scroll with his finger along the strip 1002 to participant 5's viewing area to move participant 5's viewing area into the display 106, or the local participant can touch or tap on the highlighted image 1006 of participant 5 to bring up participant 5's viewing area in the display 106." *Id.*, 8:55-62.

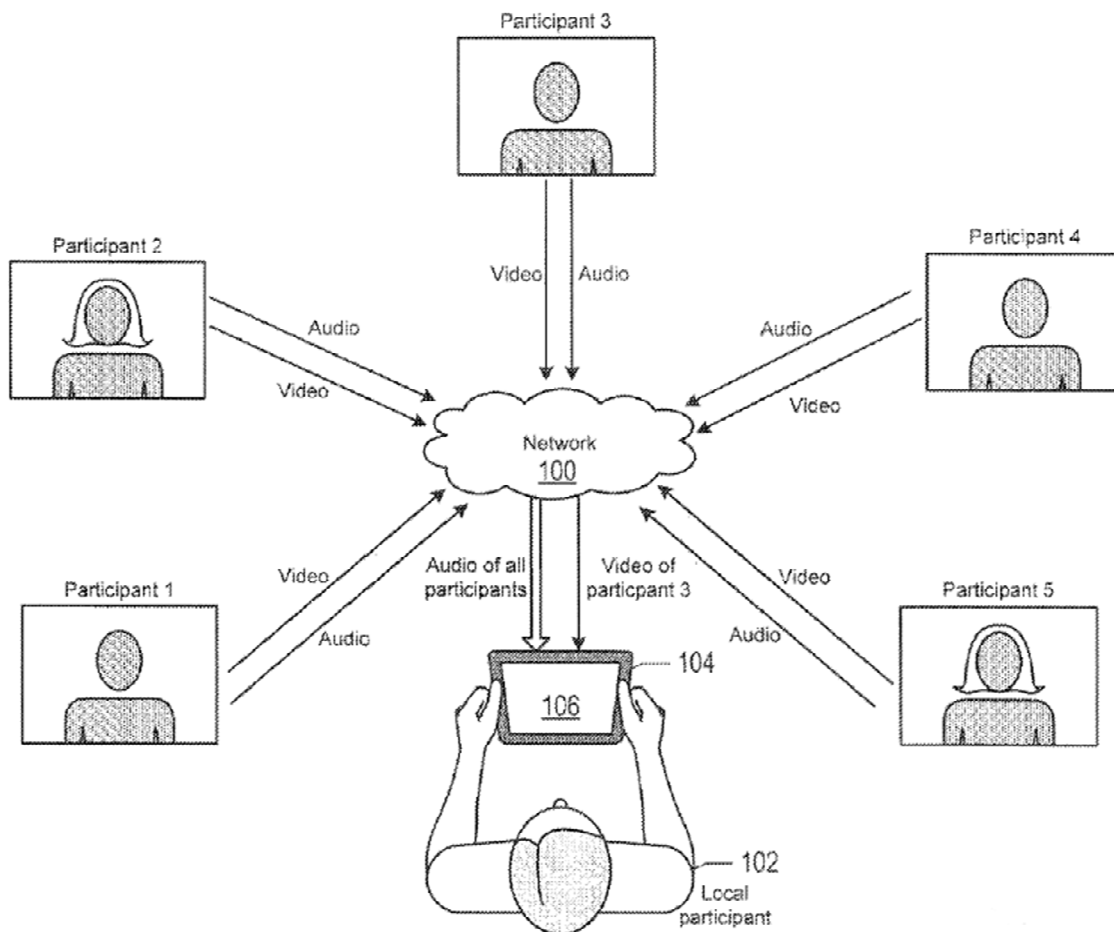


Id., Fig. 10A.*Id.*, Fig. 10B.

68. As described above, the local participant has all the data readily available at their device 104 (locally stored) so that the participant can toggle between viewing the different video feeds during the conference. Baker's description is limited in that it merely describes that the local participant can select an icon in the strip 1002 to promote that selected icon for display. Baker's description of FIGs. 10A and 10B is ambiguous as to what operations Baker's system performs to promote the selected icon. Granted Baker's description in FIGs. 10A and 10B is limited to user interface features on the display 106, it is in my opinion that a POSITA would have understood that Baker's description of "bringing up" another participant's video stream at the display 106 involved the

device 104 locally toggling between which streams from the other participants are to be presented in the viewing area on the display 106, and which of those streams are to be selected, and subsequently highlighted, in the strip 1002. Such local toggling is made possible because all of the participant streams have already been received and made locally available at the device 104.

69. FIG. 11 of Baker (reproduced below) does not cure such deficiencies, and in fact lends more to the ambiguities present throughout Baker. FIG. 11 illustrates another embodiment in which less than a complete set of video streams from other participants are sent to, received by, processed, and/or displayed by the device 104. *Id.*, 8:63-9:26, FIG. 11. Given that, specific mechanisms and network topology used to provide features described in Baker's FIG. 11 embodiment are unclear. For example, Baker describes that the local participant 102 can choose "to have only the video streams of visible remote participants to be sent to the PED 104." *Id.* Baker does not describe how such a high-level election by the participant 102 results in "only the video stream of the participant (i.e. participant 3) [being] received, processed, and presented on the display 106." *Id.* Baker also does not shed light on the network components and control processing needed for the FIG. 11 embodiment. Rather, Baker provides description of high-level prospective features without detailing underlying components to achieve such features.

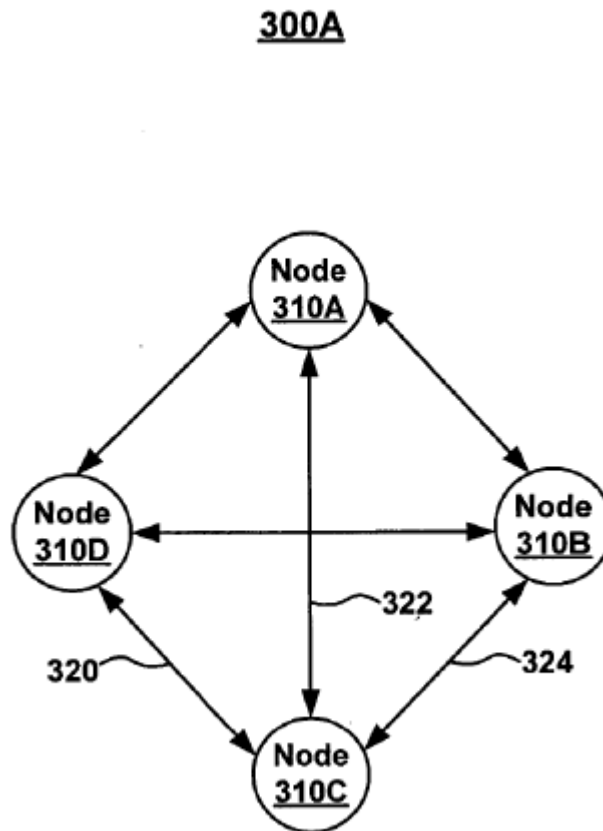


Id., Fig. 11.

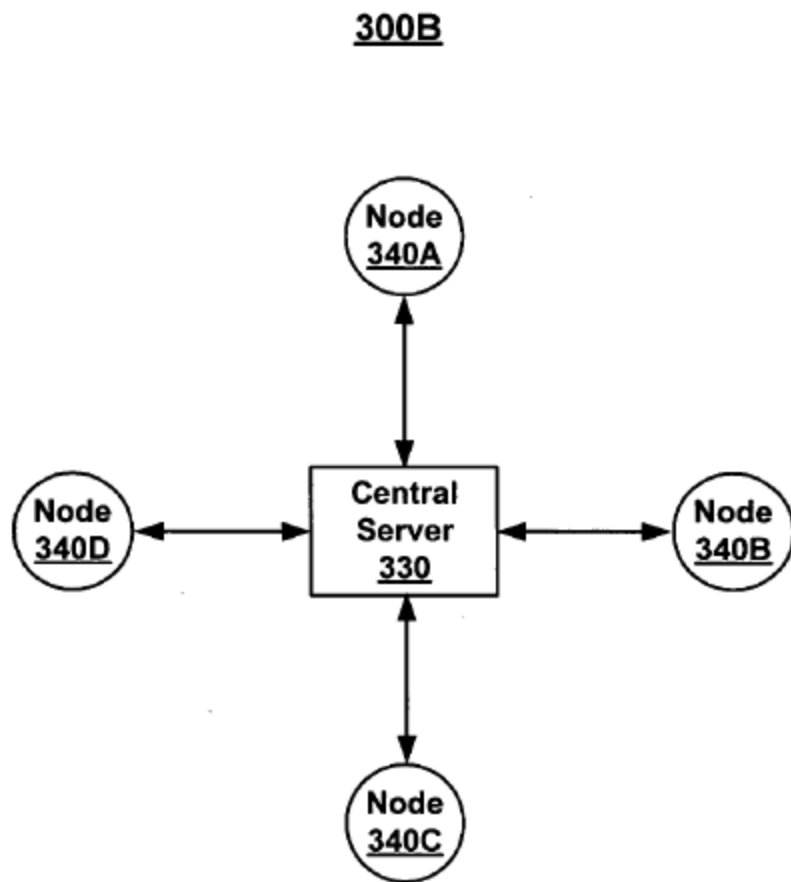
c. Apostolopoulos (EX1007)

70. The Apostolopoulos reference is directed to providing communication amongst users in a virtual environment and dynamically adapting a network topology to adapt to a number of users participating in the virtual environment. EX1007, [0002], [0018], [0022]. This virtual environment can be provided in mesh topologies or star topologies, depending on the total number of participating users. For smaller numbers of participating users, the full-mesh topology can be employed, whereas for larger numbers of users, the star topology is employed.

FIG. 3A, reproduced below, illustrates the virtual environment in a full-mesh topology. *See Id.*, [0053]-[0054]. The virtual environment is illustrated in an alternative embodiment of the star topology, as shown in FIG. 3B below. *See Id.*, [0057]-[0058].



Id., Fig. 3A.

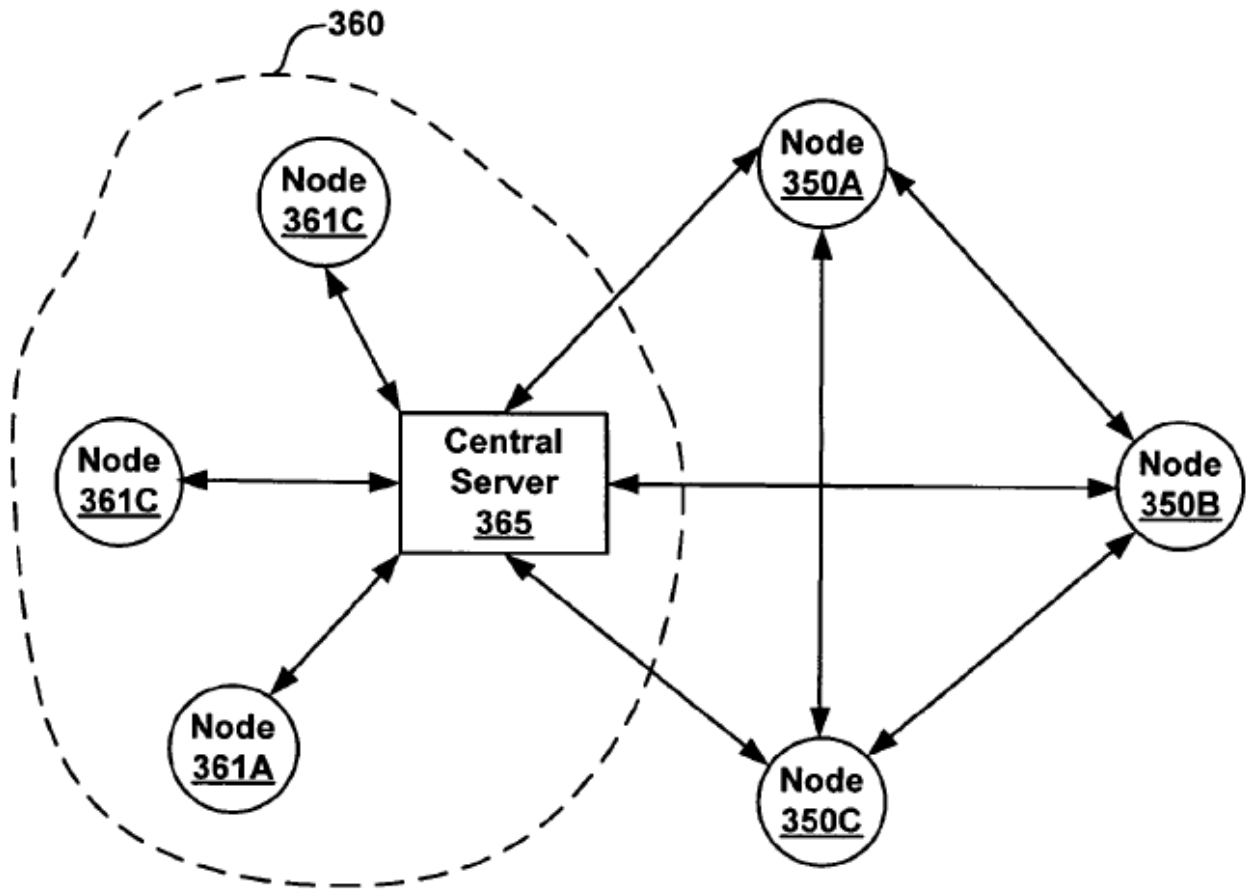


Id., Fig. 3B.

71. Apostolopoulos describes a “view dependent three dimensional virtual environment,” where instead of displaying video feeds that are part of a video conference, as disclosed by the ’099 Patent, “images in real-time of a sending collaborator...are generated from the viewpoints of receiving collaborator,” including “rendering a three dimensional model of the sending collaborator.” *Id.*, [0026]-[0028], [0055]. Apostolopoulos refers to a *three dimensional virtual environment* in which collaborators engage with each other, rather than a conference where 2D video feeds are arranged in a 2D display layout, as

contemplated by the '099 Patent. *Id.*, [0031]-[0032]. In other words, Apostolopoulos generates avatars of collaborators using data of the collaborators from various viewpoints and provides those avatars to a receiving collaborator in a three dimensional virtual environment.

72. Apostolopoulos further leverages a “central server” to provide processing capabilities for collaborator nodes to generate such three dimensional renderings in the virtual environment and deploy those renderings across the collaborator nodes. *Id.*, [0023]. The central server does “generate view dependent video streams” for the virtual environment based on a viewpoint of each receiving node in that virtual environment. *Id.*, [0026]-[0028], [0031], [0059]-[0064], [0067]. The central server can offload some processing to the collaborator nodes, such as viewpoint-independent processing, described further below, to reduce computational complexity and/or usage at the central server. The Apostolopoulos configuration of the central server and nodes is illustrated in FIG. 3C, reproduced below.

300C

Id., Fig. 3C.

73. As an example of the processing, the central server can receive a video stream from one of the nodes, process that video stream, and then pass the processed stream to other nodes in the environment. *Id.*, [0056]. The central server “performs most or all of the processing on the captured real-time videos of the local collaborator.... As such, the central node...would perform the new view synthesis techniques to render output video image streams from the perspectives of

collaborators that are viewing the local collaborator...in the N-way collaborative environment.” *Id.*, [0061]. The new view synthesis techniques can “construct, from the various real-time video streams of the sending collaborator taken from the multiple sample perspectives, a new view taken from a new and arbitrary perspective, such as, the perspective of a receiving collaborator in the virtual environment.” *Id.*, [0027]. The central server can further “render...a three dimensional model of the sending collaborator, from which the new view of the sending collaborator is generated.” *Id.*, [0028].

74. In the Apostolopoulos three dimensional virtual environment, some of the collaborator nodes can perform processing tasks before sending information to the central server to reduce complexity and processing requirements at the central server. *Id.*, [0065]. For example, “viewpoint-independent” collaborator nodes can perform “viewpoint-independent processing...before sending the processed information to the central [server]” such that “the central [server only] performs the viewpoint-dependent processing for each collaborator [node].” *Id.* The viewpoint-independent processing techniques can include “image analysis, background segmentation,” and other pre-processing techniques, which can then be used by the central server “for generating the view dependent video streams.” *Id.*, [0066], [0067].

IX. CLAIM CONSTRUCTION

75. For the purposes of my analysis, I note that constructions have not been proposed by either the Petitioner or Patent Owner.

X. ANALYSIS OF PETITION GROUNDS

76. In my opinion, a POSITA would have understood and appreciated that Setton, Baker, and Apostolopoulos, whether taken alone or in combination, do not disclose or suggest all of the elements of the challenged claims of the '099 Patent. Additionally, a POSITA would not have been motivated to combine Setton with Baker or Apostolopoulos with Baker, as suggested by the Petition and the Reddy Declaration.

77. The claims of the '099 Patent are directed to a network for video- or tele- conferencing that allows any user to participate in the conference, regardless of whether they are at a mesh-connected or non-mesh connected endpoint in a network topology, in a same or similar manner. As such, the challenged claims of the '099 Patent provide a framework for permitting non-mesh connected users to participate in the conference in very much same or similar ways as mesh-connected users participate. Each of the asserted references, whether alone or in combination, are deficient to teach how the mesh-connected and non-mesh connected users communicate in the conference as if they are all mesh-connected.

78. Setton describes a network topology in which a conference feed from a group of active sources or participants is broadcast to a group of inactive users (e.g.,

observers, relay/receivers). Only the active participants are mesh-connected with each other. Communication between these groups is unidirectional, meaning that the group of inactive users may not experience the same functionality or control during the conference as the active participants. Nor can the group of inactive users request full size data from any of the active participants, unless an inactive user first requests to be promoted to an active participant. Setton's network topology is therefore deficient in suggesting the challenged claims of the '099 Patent.

79. Baker provides description about user interfaces and systems to present conferences at portable electronic devices (PEDs). Baker provides minimal discussion of network topologies. In Baker's system, conference participants receive all data (e.g., full size/resolution video and audio streams) associated with the conference. Although the Petition relies on Baker for its user interface features, Baker lacks discussion of the network topology that is necessary for implementing user controls that are similar to the challenged claim limitations.

80. Apostolopoulos provides a network for establishing a three dimensional virtual environment where avatars of endpoint collaborators are generated and transmitted to endpoints. In this network, a central server can perform some computationally-low processes on data streams from endpoints to generate the avatars before dispersing the avatars to the endpoints. One or more of the endpoints may also shoulder some of the burden of pre-processing the data streams before the

central server generates the data streams that are transmitted to all of the endpoints. For low-capability endpoints (e.g., low-capability star nodes), the central server may still perform a majority of the processing (e.g., generating and distributing the avatars). Apostolopoulos does not describe film strips, full size video feeds, or reduced-size video feeds, let alone subject matter of the challenged claims of the '099 Patent that allow for non-mesh endpoints to participate in a conference in a same or similar manner as mesh-connected endpoints. In my opinion, Baker is incompatible with Apostolopoulos and also does not cure the deficiencies of Apostolopoulos to teach the claimed limitations of the '099 Patent, especially since Apostolopoulos relies upon some central server (for a star node participant) to process a virtual meeting of avatars while Baker relies upon a local participant endpoint for processing a virtual meeting space.

a. The Motivations to combine Setton and Baker are inherently deficient

81. In Ground 1 of the Petition, the Petition incorrectly alleges that Setton and Baker teach the challenged claims of the '099 Patent. Tacking Baker onto Setton does not cure the deficiencies in Setton's network topology. Setton distinguishes active participants (which the Petition correlates to mesh-connected endpoints) from mere observers (which the Petition correlates to non-mesh connected endpoints) in a conference. As a result, Setton requires only unidirectional flow of data *from* the

active participants *to* the observers endpoints. Should the observers endpoints want to *request* data, they first must request to be moved up into the mesh-connected network and thus promoted to an active participant endpoint. It does not make sense to then layer Baker onto the Setton framework, because Baker requires a flow of the data *between participants*, where each participant receives the full data (size and resolution) to then parse through and locally present.

82. The Reddy Declaration posits that a POSITA would have a reasonable expectation of success in combining Setton and Baker. EX1003, ¶61. However, both the Reddy Declaration and the Petition are deficient in explaining *why* there would be success in combining the two references, which champion different network topologies, different designations of types of participants in a conference, and different techniques for presenting and promoting videos in the conference for each type of participant. The Petition alleges that it would be obvious to make touch-to-enlarge functionality of videos in Baker available for all conference participants as it would improve the participants' experiences by permitting each participant to control which video is the focus of their display. *See* Pet., pp.33-34. In my opinion, a POSITA would not find it obvious to provide Baker's functionality to all conference participants in Setton's framework, however, because Setton only provides such interactive functionality to *active* participants, and those participants are part of the mesh-connected network, not the non-mesh connected network. If

combined, Baker's functionality would only be provided to the active participants that are mesh-connected and thus would not achieve at least one of the goals of the '099 Patent. Moreover, although participants in Baker and receivers in Setton both receive full size of all video streams, the deficiency of *requesting* a full size video by Setton's receivers is not cured by Baker's teachings.

83. The Petition further alleges that the POSITA is motivated to combine Setton and Baker because the combination would allow non-speaking users (e.g., the non-mesh connected endpoints) with network constraints to still be part of the conference *and also monitored* while allowing the active speakers (e.g., the mesh-connected endpoints) to have prioritized network access. *Id.*, p.18. Prioritized network access can ensure that the active speakers do not experience lag or reduction in quality of communications and streaming during the conference. The Petition is deficient in that it lacks reasoning behind this alleged motivation to combine. In my opinion, the Petition is deficient in providing this motivation to combine because it does not take into consideration the facts that (i) Baker provides *all data to all users* so that each user, regardless of whether they are active or not, can decide how to *locally* display or promote the data and (ii) Setton only permits control and bidirectional communication amongst the active users.

84. The Reddy Declaration incorrectly asserts that Baker's features further Setton's purposes of accommodating "network constraints" of different

user participants. “Network constraints” here refer to Setton accommodating active users who need delays of less than a few hundreds of milliseconds along with observers, who can accommodate delays of a few seconds. EX1003, ¶¶59-60, 87. Since Baker provides functionality to *all* users, it would not further Setton’s goal to *only provide enhanced functionality to active users*. The Reddy Declaration as well as the Petition allege that it would be obvious to combine the two references to achieve Setton’s purpose of reducing network load, however the declaration is deficient because (i) it does not describe *why* it would be obvious to combine (other than it being done with well-known software programming), nor does it (ii) describe *how* the combination would disclose the purpose of ’099 Patent. *Id.*, ¶59.

85. Combining Setton and Baker would force Baker to act in a certain way that frustrates the intent of the ’099 Patent. Under this combination, a user at a non-mesh connected endpoint cannot simply select a video in a film strip to then cause the non-mesh connected endpoint to transmit a request to a mesh-connected endpoint for a full version of the selected video. This is because Setton’s network topology clearly prevents bidirectional transfer of information between mesh-connected (active participants) and non-mesh connected (observers) endpoints in a conference system. Even if Baker could be combined with Setton, Baker would not cure the inability of Setton’s non-mesh connected endpoints to request or select

specific videos to receive in full size. Based on reading Setton as a whole, if one does not want to burden the network to allow for selection of one of the videos, then the references cannot be combined because there would be no capability in the combination to allow for transmitting a request for the full version of the selected video from the non-mesh connected endpoint to the mesh-connected endpoint. Setton's unidirectional broadcast or multicast simply does not allow for feedback and thus does not permit for any motivation or obviousness to combine with the Baker reference.

86. The Reddy Declaration is further incorrect to assert that the combination of Setton and Baker renders the claimed film strip obvious, which is discussed further below. *Id.*, ¶¶82, 86, 90, 108. First, Setton does not display video feeds "from the conference participants," as claimed in the '099 Patent. Rather, the Setton topology, as illustrated in Setton's FIG. 1 and discussed above, shows that only the peer-to-peer active endpoints (e.g., mesh-connected) display video feeds from other peer-to-peer active endpoints. *Supra*, Sec. VIII.a. No video feeds are displayed from the participants at the non-mesh connected endpoints in Setton's topology because the non-mesh connected endpoints do not have bidirectional communication with the mesh-connected endpoints. *Id.* second, Baker only discloses displaying videos from client-server nodes *or* peer-to-peer nodes, *not both*. *Supra*, Sec. VIII.b. Even if the references can be combined, Setton and Baker

fail to disclose the claimed film strip, and both the Petition and Reddy Declaration lack sufficient reasoning as to how and why, if combined, the Setton and Baker references clearly disclosed the claimed film strip for *both mesh-connected and non-mesh connected endpoints*.

87. Accordingly, it is my opinion that the Petition and Reddy Declaration's motivations to combine are not rooted in the teachings and intent of the references.

b. The Motivations to Combine Apostolopoulos and Baker are Deficient

88. Since the Baker reference does not address challenges associated with a hybrid network topology, it would not be obvious to combine with Apostolopoulos, nor, in my opinion, would a POSITA be motivated to consider such a combination. Apostolopoulos's topology adaptation method and system, has a goal of generating avatars and providing them in a *three dimensional virtual environment* for users, rather than simply providing multiple 2D video feeds on a type of 2D display taught by Baker. *See* EX1003, ¶233.

89. In my opinion, the Reddy Declaration incorrectly ascertains that the combination of Apostolopoulos and Baker would yield predictable results. *See Id.*, ¶¶198-201. To the contrary. Apostolopoulos is aimed at three dimensional virtual rendered environments while Baker is aimed at a 2D display of video feeds. Baker's features would not further define Apostolopoulos's main purpose,

which is to support three dimensional virtual environment communication (not video- and tele- conferencing techniques as discussed herein). Moreover, in my opinion, a POSITA would not have been motivated to inject Baker's film strip into Apostolopoulos's virtual environment to allow users to promote video feeds of interest. Apostolopoulos, as asserted in the Reddy Declaration, describes increasing resolution of an object of interest in the virtual environment, but does not teach or suggest adjusting presentation of a *2D video feed* because Apostolopoulos never contemplated implementation in a 2D conference environment. *Id.* Accordingly, in my opinion, a POSITA also would not have had a reasonable expectation of success in adding Baker's film strip to the three dimensional virtual environment of Apostolopoulos because the POSITA would not have considered Baker a candidate for improving the three dimensional virtual environment of Apostolopoulos. In my opinion, neither the Petition nor the Reddy Declaration provide guidance as to *how and why* (i) the POSITA would be motivated to combine these references, (ii) the combination would yield predictable results, and (iii) the POSITA would have had a reasonable expectation of success in combining the references. In addition, the processing means in the two references are incompatible because Apostolopoulos relies on a central server to process the virtual meeting of avatars, whereas Baker relies on the local participant's device for processing the virtual meeting space. For at least these

reasons, it is my opinion that a POSITA would not have been prompted to modify Apostolopoulos in view of Baker.

90. In my opinion, the Petition and the Reddy Declaration ignore the fundamental teachings of Apostolopoulos. Apostolopoulos teaches a different network topology than the '099 Patent in which all star endpoints must go through a central server as a main contact point to any mesh endpoints. *See, supra* Sec. VIII.c. Baker does not remedy the network topology deficiencies of Apostolopoulos. Even if the Apostolopoulos central server becomes an endpoint, Baker would not disclose means for receiving reduced-size video feeds at an endpoint and, based on selection, obtaining transmission of a full size video feed. Furthermore, Apostolopoulos does not provide means for receiving the reduced-size video feeds at the endpoint and, based on selection, obtaining transmission of the full size video feed. *See* EX1007, [0055]. Apostolopoulos rather generates three dimensional models and avatars then distributes those to endpoints. For example, each collaborator endpoint can generate a three dimensional avatar and sends it to receiving collaborator endpoints. *Id.*, [0028]. For a star connected endpoint, a central server/endpoint distributes the three dimensional avatars (from the collaborator endpoints) to the star endpoint. That is, the “central node performs the view-dependent processing for a local collaborator.” *Id.*, [0088]. In Baker, the endpoint itself (i.e., local participant) receives an audio/video stream from each

remote participant (via network 100), arranges the streams in a data structure describing the locations of the viewing areas within a virtual meeting place, and presents on the display a view of a portion of the virtual meeting space. *See supra*, Sec. VIII.b. In sum, Apostolopoulos relies on a central server/node (for a star endpoint participant) to process the virtual meeting of avatars, whereas Baker relies on the local participant for processing the virtual meeting place (irrespective of the topology of network 100, client-server or peer-to-peer). Therefore, it is my opinion that a POSITA would not be motivated to combine Apostolopoulos and Baker.

c. Elements [3.1.1], [3.1.2], [10.0]

91. In this section, I address Elements [3.1.1], [3.1.2], and [10.0] as a group. Elements [3.1.1] and [3.1.2] collectively correspond to or recite similar limitations as Element [10.0], which are detailed in the following table:

[3.1.1] a plurality of endpoints, at least some of which are connected in a mesh configuration and at least one of which is connected in a non-mesh configuration to a selected one of the mesh-connected endpoints, and [3.1.2] which are configured such that a film strip with frames containing reduced-size video feeds from the conference participants is displayed at the mesh endpoints;	[10.0] An endpoint for use as a mesh-connected endpoint in a teleconferencing system comprising a plurality of endpoints, at least some of which are connected in a mesh configuration and at least one of which is connected in a non-mesh configuration to a selected one of the mesh-connected endpoints, wherein a film strip with frames containing reduced-size video feeds from the conference participants is
--	--

	displayed at the mesh endpoints, the mesh-connected endpoint comprising:
--	--

92. Contrary to the Petition’s allegations (Pet., 20-27), it is my opinion that Setton in combination with Baker do not disclose or suggest Elements [3.1.1] and [3.1.2], in combination, or [10.0]. For example, Setton does not disclose “a plurality of endpoints, at least some of which are connected in a mesh configuration and at least one of which is connected in a non-mesh configuration to a selected one of the mesh-connected endpoints, and which are configured such that a film strip with frames containing reduced-size video feeds from the conference participants is displayed at the mesh endpoints” of Elements [3.1.1] and [3.1.2], at least because Setton does not teach the film strip, and even when combined with Baker, Setton’s active participants (e.g., mesh endpoints) would not have relay/receivers (e.g., non-mesh endpoints) included in the film strip due to Setton’s unidirectional communication between active participants and relay/receivers.

93. In my opinion, the Petition’s Ground 1 does not consider Elements [3.1.1] and [3.1.2] as a whole, which read together, indicate that all of the endpoints, whether they are mesh-connected or non-mesh connected, receive the film strip with reduced-size video feeds from all the conference participants. The

conference participants refer to all of the endpoints in the conference (i.e., mesh or non-mesh endpoints), which is a construction of the claimed limitations that the Petition has omitted from its allegations.

94. In my opinion, when read in the context of the entire claim 3, “the conference participants” in Element [3.1.2] has a broad reading of *all* conference participants, whether or not those conference participants are mesh-connected and non-mesh connected. This interpretation is further contrasted to “the mesh-connected conference participants” in Element [3.3], which defines *a group of the conference participants* as being mesh-connected. Claim 3 intentionally calls out “the conference participants” and “the mesh-connected conference participants,” thereby indicating that the conference participants include all participants, from both the mesh-connected and the non-mesh connected endpoints. The specification further describes, in the context of FIG. 1 that illustrates an endpoint: “conference participants 14 sit before video cameras...and in front of monitors....The monitors are controlled by display controller...and receive multiple video streams and collaboration data...from all other endpoints participating in a given conference.” EX1001, 5:13-17. The endpoint, as described throughout the specification, is either mesh-connected or non-mesh connected. The specification sheds light into the notion that “the conference participants” include all participants from both mesh-connected and non-mesh connected endpoints. Therefore, in my opinion, “the

conference participants,” read in the context of the claim 3 and the specification, apprises a POSITA of the scope of the claim limitations. The combination of Elements [3.1.1] and [3.1.2] is also consistent with the ’099 Patent’s stated objective to allow non-mesh connected endpoints to participate in conferences in same or similar ways as mesh-connected endpoints in the conference. In summary, the combination of Elements [3.1.1] and [3.1.2] ensures that the non-mesh endpoints are included in the film strip displayed on the mesh endpoints, while Element [3.3] ensures that the mesh endpoints are included in the film strip displayed on the non-mesh endpoint(s). As a result, all conference participants can see each other.

95. As indicated above, Setton also does not teach or suggest the claimed framework because Setton’s topology lacks bidirectional communication between mesh-connected and non-mesh connected endpoints. Setton requires all observers in a conference (which the Petition correlates to non-mesh connected endpoints) to request to be promoted to active participants (which the Petition correlates to mesh-connected endpoints) if the film strip is to include all participants in the conference so that everyone can see each other. Setton therefore inherently does not allow for mesh-connected *and* non-mesh connected endpoints to share data and view the data for *all of the conference participants* during the conference.

96. As indicated above, the Petition turns to Baker to remedy deficiencies in Setton, specifically Baker's FIGs. 10A and 10B, to teach the film strip having the reduced size video feeds for all conference participants. *See* Pet., p.25. The Petition further asserts portions of Baker that simply are deficient in rendering it obvious to a POSITA to combine Baker with Setton to reach the claimed limitations of Elements [3.1.1] and [3.1.2]. *Id.*; *see also* EX1006, 9:11-14, 9:19-26, 8:40-44; *supra* Sec. III.B. As described herein, the Petition misses the mark – Baker's framework is not compatible with Setton because Setton does not allow any type of bidirectional communication between mesh-connected and non-mesh connected endpoints.

97. Additionally, the Petition does not identify or allege that the combination of Setton and Baker would include non-mesh connected endpoints transmitting "video feeds," let alone that these non-mesh endpoint "video feeds" would be displayed at "mesh endpoints" as part of "a film strip with frames containing reduced-size video feeds from the conference participants." Elements [3.1.1]-[3.1.2]. At best, the Petition alleges with regard to Element [3.4] that Setton's relay/receivers could be modified to transmit requests for full size video feeds, but does not allege that Setton's relay/receivers would have been modified to transmit "video feeds" for display in "a film strip" on Setton's active sources.

98. Baker's description alludes to remote participants in the conference sending their full-rate video streams to a network, and based on a request/control from a local participant (e.g., a non-mesh connected endpoint) the network sends only specific videos and/or different qualities of the videos. The local participant can then *locally* control what is sent to it from network 100, as well as what to display and how to display. EX1006 8:40-67, 9:1-26. Both Baker and Setton stay away from describing a topology that allows every mesh-connected and non-mesh connected node to interact directly with each other. Baker does not cure Setton's topology deficiencies, especially since Baker sends all conference data to each endpoint, through the network, and allows the endpoint to locally control display and size/quality/resolution of what is being displayed. Based on the innate and inherently different topologies and teachings of Setton and Baker, it would not have been obvious to combine these references to reach the Elements [3.1.1] and [3.1.2], or [10.0]. For at least the above mentioned reasons, it is my opinion that the Petition's Ground 1 is insufficient. Moreover, it is my opinion that Ground 2 of the Petition is also insufficient for the same reasons mentioned herein regarding the Petition's Ground 1.

99. Moreover, Apostolopoulos discloses a different network topology that is more narrowing than that contemplated by the Elements [3.1.1] and [3.1.2], or [10.0]. *See supra*, Sec. VIII.c. First, Apostolopoulos requires all star (i.e., non-

mesh) endpoints in a network to communicate through a central server to mesh endpoints. *Id.* It is my opinion, a POSITA would understand that the central server is not an endpoint, and therefore, a non-mesh endpoint is not connected to a mesh endpoint as required by Element [3.1.1].. Second, Apostolopoulos does not consider a hybrid network topology *unless* an amount of endpoints changes. In other words, “the preferred topology varies as a function of the number of nodes. For example if there are only 2-3 collaborators then a full-mesh may be appropriate, while if there are 10-15 collaborators, then a star topology may be appropriate. In another embodiment, it is valuable to move from a full-mesh to a star topology through hybrid topologies as the number of collaborators increases, as well as the reverse (if beneficial) if the number of collaborators decreases on an ongoing basis.” EX1007, [0076]. Apostolopoulos only considers a hybrid topology *during an interim when shifting between a mesh topology* (see FIG. 3A in Sec. VIII.c) *and a star topology* (see FIG. 3B in Sec. VIII.c).

100. In my opinion, a POSITA would not find it obvious to combine Apostolopoulos with Baker for the same reasons mentioned above that it would not be obvious to combine Setton with Baker. Baker simply does not cure the network topology deficiencies of Apostolopoulos in order to teach or suggest the Elements [3.1.1] and [3.1.2], or [10.0]. Furthermore, Baker and Apostolopoulos are not compatible as discussed earlier – Apostolopoulos relies upon a central server to

process a virtual meeting of avatars (in a star topology, for example), whereas Baker relies upon a local participant to process a virtual meeting space. In my opinion, the Petition incorrectly alleges that a POSITA would have been prompted to modify the Apostolopoulos central server in the hybrid topology to be a full mesh node without adequately considering the teachings of Apostolopoulos as a whole or how such modifications would impact the Apostolopoulos system. Moreover, the Petition does not consider whether a full-mesh node participating in Apostolopoulos's three dimensional virtual environment would be able to also serve the same function of Apostolopoulos's central server to offload three dimensional model generation from other nodes as processing complexity increases (as the number of participating nodes increases).

101. Moreover, in my opinion, a POSITA would understand and appreciate that Apostolopoulos's environment has different architecture and network features than the '099 Patent as well as the both the Setton and Baker references. Therefore, it would not necessarily be well-suited to combine Baker with Apostolopoulos's virtual environment, and may instead degrade performance of Apostolopoulos's virtual environment. In my opinion, a POSITA considering both Apostolopoulos and Baker in their entirety would have found them to be non-analogous and would not have been motivated to apply features from Baker to Apostolopoulos's virtual environment.

102. In my opinion, a POSITA would not have been motivated to combine Baker with Apostolopoulos to incorporate Baker's film strip into Apostolopoulos's virtual environment. When reading Apostolopoulos and Baker in their entirety, it is my opinion that the POSITA would not have incorporated the film strip into the three dimensional virtual environment because the film strip would detract from and be inconsistent with Apostolopoulos's goal to permit users to interact with each other in a *virtual environment* through *dependent rendered avatars*. For similar reasons as the Petition's Ground 1 regarding the combination of Setton and Baker, it is my opinion that the Petition's Ground 3 is insufficient. It is also my opinion that the Petition's Ground 4 is insufficient for the same reasons mentioned above regarding the Petition's Ground 3.

d. Elements [3.4], [10.3]

103. In this section, I address Elements [3.4] and [10.3] as a group. Elements [3.4] and [10.3] correspond to each other and recite similar limitations, which are detailed in the following table:

[3.4] a control connection permitting a user located at the non-mesh connected endpoint to select a video feed on the film strip for promotion to full size and to initiate transfer of the full size video feed to the non-mesh connected endpoint.	[10.3] a proxy controller permitting a user located at the non-mesh connected endpoint to select a video feed on the film strip for promotion to full size and to initiate transfer of the full size video feed to the non-mesh connected endpoint.
--	---

104. Contrary to the Petition's allegations (Pet., 33-37), Setton combined with Baker do not disclose or suggest Elements [3.4] or [10.3]. The references, whether alone or in combination, do not allude to "a control connection permitting a user located at the non-mesh connected endpoint to select a video feed on the film strip for promotion to full size and to initiate transfer of the full size video feed to the non-mesh connected endpoint," as recited in Element [3.4]. In other words, the claim limitations provide for the non-mesh connected endpoint to *request a full size video feed* in response to selecting a video feed on the film strip. *See supra*, Sec. VII. As described above, Setton does not permit for this bidirectional flow of information between mesh-connected and non-mesh connected endpoints. *See supra*, Sec. VIII.a. Baker does not cure the deficiencies of Setton because Baker requires receiving all video feeds from a network at the non-mesh connected endpoint to allow the non-mesh connected endpoint to *locally* create icons of the film strip. *See supra*, Sec. VIII.b. In contrast, Elements [3.4] and [10.3] require that the film strip be generated by a server, which is outside of the non-mesh connected endpoint. Furthermore, in my opinion, a POSITA would not have been motivated to combine the references because Baker simply does not address the pitfalls of Setton's network topology.

105. The Petition merely states that it would be obvious to combine the references to improve all conference attendees experience by allowing them to control what videos are the focus of their respective displays. *See Pet.*, p.34. The

Petition neglects to explain *why* it would be obvious to combine references that simply may not work together, and thus, the Petition does not provide any reasoning or explanation behind a motivation to combine such references. In my opinion, the Petition appears to combine features from Setton and Baker at will without providing reasons or motivations to combine the references. Such combinations appear to disrupt the core operation of Setton's system, which the Petition does not appreciate or address. For example, combining the references as the Petition suggests can increase the processing and serving burden for Setton's active sources, which would have to assemble and serve multiple different video streams to individual relay/receivers instead of assembling a single video stream of the conference that is broadcast to all relay/receivers. Consequently, such increased processing and serving burden would have competed with and added delays to the active source's transmission and receipt of video feeds with the other active sources.

106. In addition to the reasons mentioned above (*supra* Sec. X.c), in my opinion, a POSITA would not be motivated to combine Setton and Baker because the combination would cause costly burdens on Setton's network. Setton champions its network of segregated active participants and observers on its ability to reduce network load and ensure that active participants, not the observers, experience minimal delays or other quality issues during the conference. *See supra*, Sec. VIII.a; *see also Pet.*, pp.35-36. Setton will aggregate and send all the video feeds to the

active participants (the mesh-connected nodes) but not the observers, in order to achieve this objective. Adding in Baker's transmission of all the video feeds to all the participants for local processing and promotion would frustrate the objective of Setton (assuming Setton can even be modified in the first place to provide bidirectional communication between the active participants and the observers) and thus cause network burdens and increased costs on communication amongst the participants.

107. Under a mixed framework of Setton and Baker, it would undeniably be costly to send a request back from a non-mesh connected endpoint to a mesh-connected endpoint for transmission of a full video feed to the requesting non-mesh connected endpoint. Moreover, if Setton does not want to burden the network to allow selection of one of the video feeds and transmission of the full selected video feed at a non-mesh connected endpoint, then Setton cannot be combined with Baker. There is no capability in this combination to transmit requests back and forth between the mesh-connected and non-mesh connected endpoints.

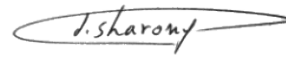
108. Moreover, in my opinion, a POSITA would not find it obvious to combine Apostolopoulos with Baker to teach or suggest the Elements [3.4] or [10.3]. In Apostolopoulos, each collaborator endpoint generates a three dimensional avatar and sends it to receiving endpoints. *See* EX1007, [0028]. Further, for a star connected endpoint in Apostolopoulos's network topology, a central server/node

distributes the three dimensional avatars (from the collaborator endpoints) to the star endpoint. That is, the “central node performs the view-dependent processing for a local collaborator.” *Id.*, [0088]. In Baker, on the other hand, an endpoint itself (i.e., local participant) receives an audio/video stream from each remote participant (via network 100), arranges the streams in a data structure describing locations of viewing areas within a virtual meeting space, and presents on a display a view of a portion of the virtual meeting space. *See* EX1006, claim 11. In my opinion, it would not be obvious to a POSITA to consider combining these references because Apostolopoulos relies on a central server/node (for a star endpoint) to process a virtual meeting of avatars, whereas Baker relies on a local participant endpoint for processing a virtual meeting space, irrespective of the topology of network 100. Apostolopoulos and Baker, whether alone or in combination, do not suggest or teach the Elements [3.4] or [10.3], as indicated by the discussion above.

XI. CONCLUSION

109. I declare that all statements made in this Declaration are true, and all statements made on information and belief are believed to be true, and that the statements were made with the knowledge that willful false statements and the like are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Dated: February 29, 2024

A handwritten signature in cursive script, appearing to read "J. Sharony", is written above a horizontal line.

Dr. Jacob Sharony