

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

CISCO SYSTEMS, INC.,
Petitioner

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

**PETITION FOR *INTER PARTES* REVIEW
UNDER 35 U.S.C. §312 AND 37 C.F.R. §42.104**

(Claims 3-6 and 10)

TABLE OF CONTENTS

Petitioner’s Exhibit List	5
I. Introduction.....	7
II. Grounds for standing	7
III. Note.....	7
IV. Background.....	8
V. Summary of the ’099 patent	8
VI. Effective priority date of the ’099 patent	9
VII. Level of ordinary skill in the art	9
VIII. Claim construction.....	9
IX. Relief requested and reasons for the relief	10
X. Identification of how the claims are unpatentable.....	10
A. Challenged claims	10
B. Statutory grounds for challenges.....	11
C. Ground 1	12
1. Summary of Setton.....	12
2. Summary of Baker	14
3. Reasons to Combine Setton and Baker.....	16
4. Claim 3	19
5. Claim 10.....	37
D. Ground 2	40

1.	Summary of Wong.....	40
2.	Reasons to Combine Setton, Baker, and Wong.....	41
3.	Claim 4.....	44
4.	Claim 5.....	46
5.	Claim 6.....	47
E.	Ground 3.....	49
1.	Summary of Apostolopoulos	49
2.	Reasons to Combine Apostolopoulos and Baker.....	51
3.	Claim 3.....	53
4.	Claim 10.....	69
F.	Ground 4.....	71
1.	Reasons to Combine Apostolopoulos, Baker, and Wong.....	71
2.	Claim 4.....	73
3.	Claim 5.....	75
4.	Claim 6.....	75
XI.	Discretionary denial is not appropriate.....	77
A.	No basis for 35 U.S.C. §325(d) denial.....	77
B.	No basis for 35 U.S.C. §314/ <i>Fintiv</i> denial.....	77
1.	Whether evidence exists that a stay may be granted if a proceeding is instituted	78

2.	Proximity of the court’s trial date to the board’s projected statutory deadline for a final written decision	78
3.	Investment in the parallel proceeding	79
4.	Overlap between issues raised in the petition and in the parallel proceeding	79
5.	Whether the petitioner is a defendant	80
6.	Other circumstances that impact the Board’s exercise of discretion, including the merits	80
C.	Discretionary denial under <i>General Plastic</i> is not appropriate	81
XII.	Conclusion	81
XIII.	Mandatory notices	82
	Real party-in-interest	82
	Related matters	82
	Lead and back-up counsel and service information	83
	Claims Appendix	84
	Certificate of Word Count	86
	Certificate of Service	87

PETITIONER’S EXHIBIT LIST

Ex.1001	U.S. 9,204,099
Ex.1002	Prosecution History of U.S. 9,204,099
Ex.1003	Declaration of Narasimha Reddy, Ph.D. under 37 C.F.R. §1.68
Ex.1004	<i>Curriculum Vitae</i> of Narasimha Reddy, Ph.D.
Ex.1005	U.S. Pub. 2010/0287238
Ex.1006	U.S. 8,581,958
Ex.1007	U.S. Pub. 2005/0080894
Ex.1008	U.S. 7,692,683
Ex.1009	U.S. Pub. 2009/0232129
Ex.1010	U.S. 8,601,097
Ex.1011	U.S. 9,124,757
Ex.1012	Andrew S. Tanenbaum, “Computer Networks,” 4th ed. (2002)
Ex.1013	U.S. 9,009,225
Ex.1014	U.S. 7,877,489
Ex.1015	Michael Arrington, “Facebook Will Launch In-Browser Video Chat Next Week In Partnership with Skype,” TechCrunch (Jul. 1, 2011).
Ex.1016	Ian Paul, “Facebook Video Calling vs. Google+ Hangouts,” PCWorld (July 7, 2011).
Ex.1017	U.S. Pub. 2006/0215765
Ex.1018	U.S. Pub. 2009/0015661
Ex.1019	U.S. Pub. 2010/0157017
Ex.1020	U.S. Pub. 2010/0333004

Ex.1021	Complaint, Video Solutions Pte. Ltd. V. Cisco Systems, Inc., 2-23-cv-00222 (E.D. Tex.)
Ex.1022	Docket Control Order, Video Solutions Pte. Ltd. V. Cisco Systems, Inc., 2-23-cv-00222 (E.D. Tex.)
Ex.1023	Calendar Events Set for 2/10/2025 - 2/11/2025, Judge Rodney Gilstrap
Ex.1024	Interim Procedure for Discretionary Denials in AIA Post-Grant Proceedings with Parallel District Court Litigation (June 21, 2022)
Ex.1025	United States Copyright Office Catalog Record for Andrew S. Tanenbaum, Computer Networks, 4th ed. (Ex.1012)

I. INTRODUCTION

Pursuant to 35 U.S.C. §§311, 314(a), and 37 C.F.R. §42.100, Cisco Systems, Inc. (“Petitioner”) respectfully requests that the Board review and cancel as unpatentable under (pre-AIA) 35 U.S.C. §103(a) claims 3-6 and 10 (the “Challenged Claims”) of U.S. 9,204,099 (the “’099 patent,” Ex.1001).

The ’099 patent describes videoconferencing in mixed topology networks where participants may be at different locations. Ex.1001, Abstract. The ’099 patent proposes a user interface that includes a “filmstrip consist[ing] of a frame strip containing reduced size video feeds from the various participants.” Ex.1001, 4:63-65. As shown below and confirmed in the Declaration of Narasimha Reddy, Ph.D. (Ex.1003) these concepts were well known. Therefore, the references presented in this Petition teach and render obvious the Challenged Claims.

II. GROUNDS FOR STANDING

Petitioner certifies that the ’099 patent is eligible for IPR and that Petitioner is not barred or estopped from requesting IPR challenging the patent claims. 37 C.F.R. §42.104(a).

III. NOTE

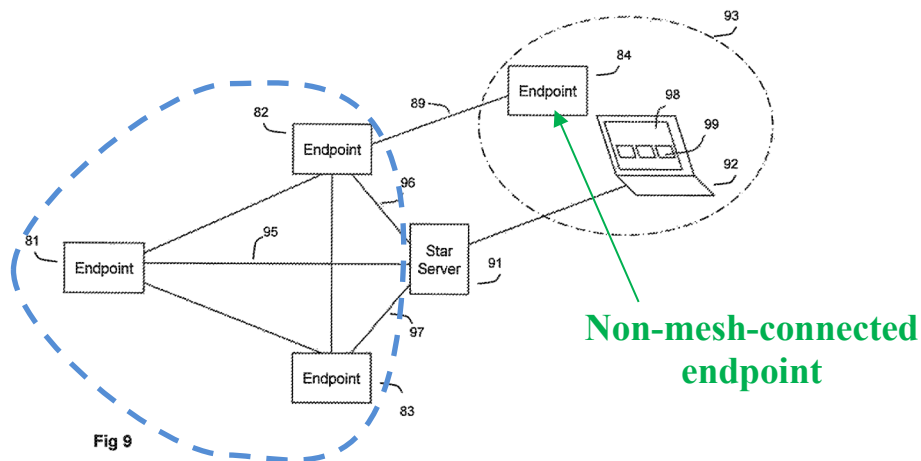
Petitioner cites to exhibits’ original page numbers. **Emphasis** in quoted material has been added. Claim terms are presented in *italics*.

IV. BACKGROUND

The '099 patent relates to videoconferencing techniques that were well-known as of the earliest claimed priority date. Petitioner's expert Dr. Reddy describes relevant background predating the '099 patent. *See* Ex.1003, ¶¶25-27.

V. SUMMARY OF THE '099 PATENT

The '099 patent proposes a video teleconferencing system in which “multiple parties participate in a teleconference from different locations.” Ex.1001, Abstract. In mixed topology networks, the '099 patent describes a mechanism where “non-mesh endpoint(s) user(s) can use [a] filmstrip user interface...to exercise the same control over [the] video presentation as would a mesh connected user.” Ex.1001, 15:46-49. Fig. 9 shows an example of such a mixed topology network.



Mesh connected endpoints

Ex.1001, Fig. 9 (annotated)

As explained below, these concepts were well known as of the priority date of the '099 patent. Ex.1003, ¶¶28-34.

VI. EFFECTIVE PRIORITY DATE OF THE '099 PATENT

The '099 patent claims priority to a U.S. provisional patent application filed on February 1, 2012. All of the cited and relied-upon prior art predates February 2012. Accordingly, Petitioner has not undertaken a priority date analysis. Petitioner does not waive any right or opportunity it may have to dispute the priority date of the '099 patent in this or another forum where the issue is relevant.

VII. LEVEL OF ORDINARY SKILL IN THE ART

A Person of Ordinary Skill in The Art (“POSITA”) in February 2012 would have had a working knowledge of video conference techniques, mixed topology networks that are pertinent to the '099 patent, including mesh and non-mesh (e.g., star, tree-based, and hybrid) networks, internet-based software development, video and audio compression, and user interface design. A POSITA would have had a bachelor’s degree in computer science, computer engineering, electrical engineering, or equivalent, and approximately three years of work experience relating to video teleconferencing systems. Lack of professional experience can be remedied by additional education, and vice versa. Ex.1003, ¶¶18-20.

VIII. CLAIM CONSTRUCTION

Claim terms in IPR are construed according to their “ordinary and customary meaning” to those of skill in the art. 37 C.F.R. §42.100(b). Petitioner submits that, for the purposes of this proceeding and the grounds presented herein, no claim

term requires express construction. *Nidec Motor Corp. v. Zhongshan Broad Ocean Motor Co.*, 868 F.3d 1013, 1017 (Fed. Cir. 2017).

IX. RELIEF REQUESTED AND REASONS FOR THE RELIEF

Petitioner asks that the Board institute a trial for *inter partes* review and cancel the Challenged Claims in view of the analysis below.

X. IDENTIFICATION OF HOW THE CLAIMS ARE UNPATENTABLE

A. Challenged claims

Petitioner challenges claims 3-6 and 10. Claims 3-6 are asserted against Petitioner in the co-pending litigation. A finding that the Challenged Claims are unpatentable in this proceeding will resolve the parties' dispute and obviate the need for a trial regarding the '099 patent in the co-pending litigation, substantially reducing the time and expense of that litigation for all parties.

B. Statutory grounds for challenges¹

Grounds	Claims	Basis
#1	3 and 10	35 U.S.C. §103 (Pre-AIA) over Setton and Baker
#2	4-6	35 U.S.C. §103 (Pre-AIA) over Setton, Baker, and Wong
#3	3 and 10	35 U.S.C. §103 (Pre-AIA) over Apostolopoulos and Baker
#4	4-6	35 U.S.C. §103 (Pre-AIA) over Apostolopoulos, Baker, and Wong

U.S. Pub. 2010/0287238 (Ex.1005, “**Setton**”); published on Nov. 11, 2010.

U.S. 8,581,958 (Ex.1006, “**Baker**”); filed on April 18, 2011.

U.S. Pub. 2005/0080894 (Ex.1007, “**Apostolopoulos**”); published on April 14, 2005.

U.S. Pub. 2009/0232129 (Ex.1009, “**Wong**”); published on Sept. 17, 2009.

Setton, Apostolopoulos, and Wong are prior art under 35 U.S.C. §102(b).²

Baker is prior art under 35 U.S.C. §102(e).

¹ Petitioner relies on the combined teachings of the references and not on a physical incorporation of elements. *See In re Mouttet*, 686 F.3d 1322, 1332 (Fed. Cir. 2012); Ex.1003, ¶¶37-39.

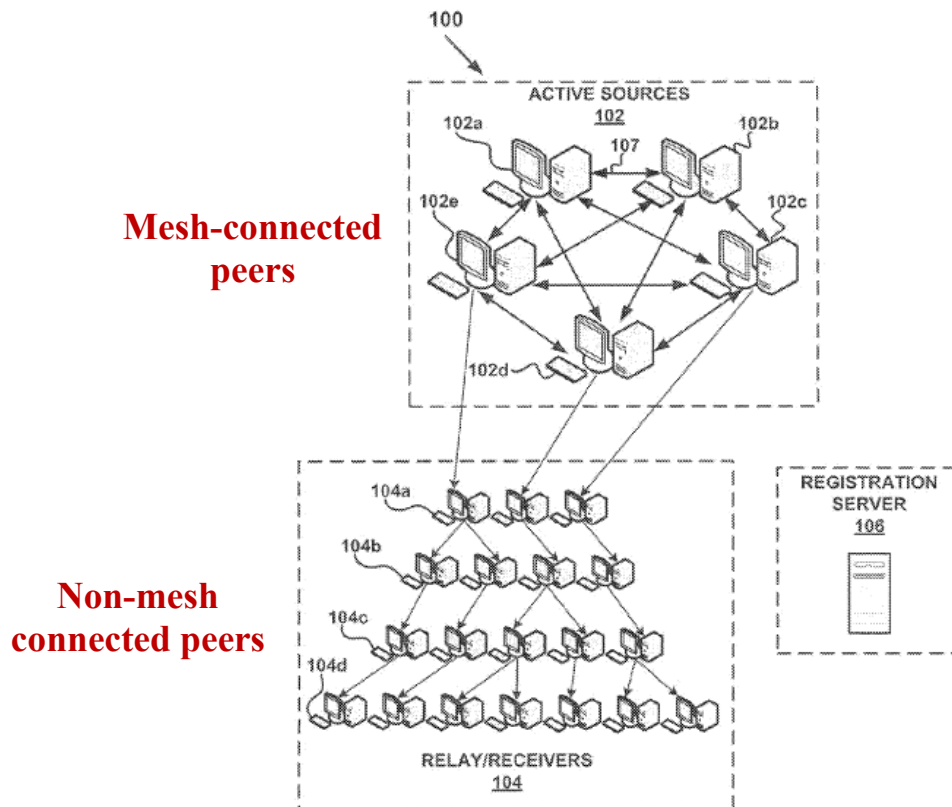
² If Patent Owner argues that the ’099 patent is an AIA patent, Setton, Baker, Apostolopoulos, and Wong would still qualify as prior art under post-AIA 35 U.S.C. §102(a).

Petitioner also cites to additional prior art (Ex.1008, 1010-1020) as evidence of the background knowledge of a POSITA and to provide contemporaneous context to support Petitioner’s assertions regarding what a POSITA would have understood from the prior art in the grounds. *Yeda Research v. Mylan Pharm. Inc.*, 906 F.3d 1031, 1041-1042 (Fed. Cir. 2018); 37 C.F.R. §42.104(b).

C. Ground 1

1. Summary of Setton

Setton relates to “peer-to-peer media streaming,” such as “streaming of video data” for a “video teleconference.” Ex.1005, [0014], [0019]. Fig. 1 shows a peer-to-peer network that includes “a first portion comprised of active sources 102 and a second portion comprised of relay/receivers 104.” Ex.1005, [0016]. Setton’s active sources 102 and relay/receivers 104 are called “peers.” Ex.1003, ¶¶40-46.



Ex.1005, Fig. 1 (annotated)

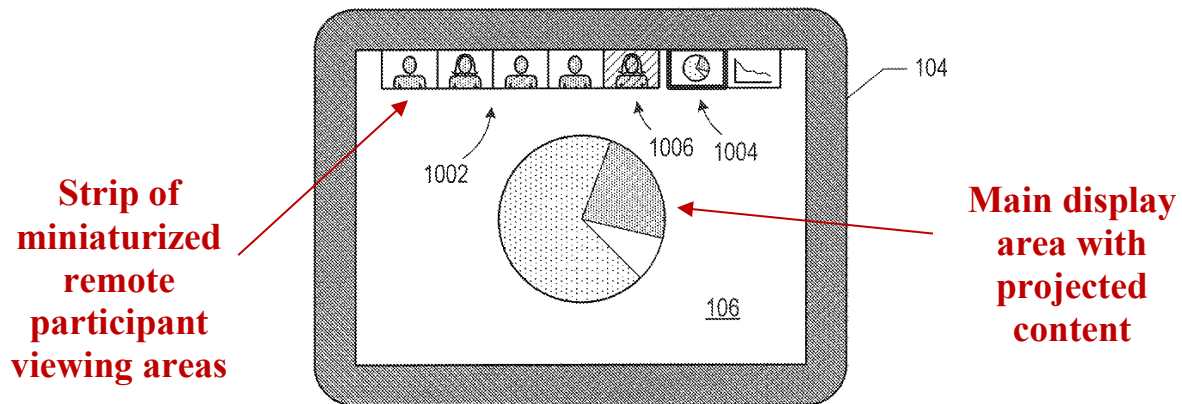
The active sources 102 are connected in a “**mesh-based**” configuration and used by teleconference participants to “generate content (e.g., a video stream, an audio stream, etc.).” Ex.1005, [0018], [0016]. The “relay/receivers 104 use a **tree-based** video multicast protocol.” Ex.1005, [0018], [0026]; Ex.1003, ¶¶47-48. In Fig. 1, relay/receivers 104 are connected to mesh-connected active sources 102 by a tree-type connection through active sources 102. Active sources “aggregate the media streams using a peer-to-peer data-driven streaming protocol.” Ex.1005, [0014]. Then, the “aggregated stream is transmitted **to all the relay/receiver peers** via application layer multicast tree(s).” Ex.1005, [0014]; Ex.1003, ¶¶49-50.

Setton is analogous art to the '099 patent and is from the same field of endeavor. The '099 patent is directed to a video teleconferencing system for use by endpoints connected in a “mesh” and “non-mesh” configuration. Ex.1001, Abstract, 15:42-45. Setton similarly describes a “media streaming” system for “streaming of video data” for a “video teleconference” among peers. Ex.1005, [0014], [0019]. Setton’s peers are connected in “mesh-based” and “tree-based” (non-mesh) configurations. Ex.1005, [0018], [0016], [0026]. Setton is reasonably pertinent to a particular problem of the '099 patent, which is video teleconferencing in mixed topology networks (e.g., Ex.1001, 15:46-49; Ex.1005, [0014], [0019], [0018], [0016], [0026]). Ex.1003, ¶40.

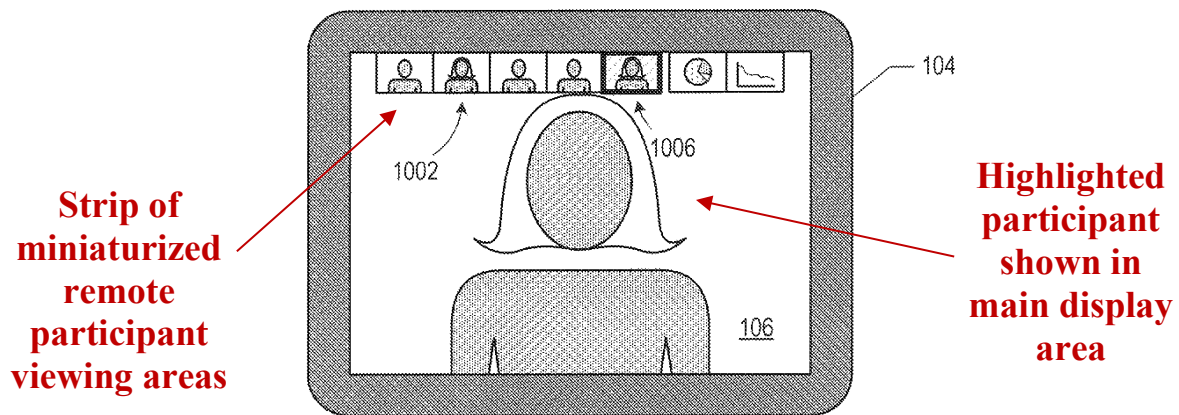
2. Summary of Baker

Baker “us[es] portable electronic devices (‘PEDs’) in video conferences.” Ex.1006, 2:31-32. “[V]ideo and audio streams generated by each of the remote participants” are “**presented as separate viewing areas on a touchscreen/display 106** of the PED 104.” Ex.1006, 2:43-48, Fig. 1; Ex.1003, ¶¶51-52.

Baker’s Figs. 10A and 10B illustrate a main display area (e.g., video of person speaking) and a “strip of miniaturized remote participant viewing areas” along an edge of the display. Ex.1006, 8:40-41, 55-57. The miniaturized video streams may show “non-speaking participants” and have “a lower bit-rate/quality” than the main display area. Ex.1006, 9:11-14, 19-26; Ex.1003, ¶¶53-54.



Ex.1006, Fig. 10A (annotated)



Ex.1006, Fig. 10B (annotated)

A local participant viewing the video conference via the display “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.” Ex.1006, 8:57-62; Ex.1003, ¶55.

Baker is analogous art to the ’099 patent and is from the same field of endeavor. The ’099 patent is directed to a video teleconferencing system with a

“filmstrip user interface” that allows a user to control what is displayed on their screen. Ex.1001, 4:5-11, 14:46-49. Baker similarly describes video teleconferencing system that uses a filmstrip to organize the user’s display area. Ex.1006, 8:27-32, 40-44. The main display area shows the “video of [the] person speaking” and a strip at the edge of the display shows a “strip of miniaturized remote participant viewing areas.” Ex.1006, 8:40-41, 55-57. Baker is also reasonably pertinent to a particular problem of the ’099 patent, which is a “video navigation system” for controlling a videoconference display. *See* Ex.1001, 3:27-52, 4:5-11, 14:46-49; Ex.1005, 8:40-62; Ex.1003, ¶51.

3. Reasons to Combine Setton and Baker

Incorporating Baker’s features into Setton’s peer-to-peer media streaming system would be combining prior art elements according to known methods to yield predictable results. Ex.1003, ¶56. The incorporated features would be Baker’s display, including the “strip of miniaturized remote participant viewing areas,” and the scroll, touch, and tap UI functionalities that allow the local participant to promote a miniaturized video stream to a full-sized video stream, and vice versa. Ex.1003, ¶56.

Combining Setton and Baker would be achieved through known methods. Ex.1003, ¶57. These known methods include routine software application programming to specify how Setton’s peers display speaking participants and non-

speaking participants, and to provide UI objects (e.g., buttons, clickable areas) that would modify the presentation of the video streams to enlarge one or more video streams. Ex.1003, ¶56. The results of the combination of Setton and Baker would have been predictable, because adding Baker's display and display control features simply specify the relative sizes of different video streams on the display, and that the relative sizes of the video streams can be updated by common UI inputs, such as scrolling, touch, or tap. Ex.1003, ¶57.

Further, combining the features Baker with Setton's peer-to-peer system would be applying a known technique to a known device ready for improvement to yield predictable results. Ex.1003, ¶58. Setton discloses active sources generating and aggregating video streams to transmit to passive relay/receivers in a video conference, but Setton provides limited details about how different video streams are displayed or controlled. Ex.1005, [0019], [0034]; Ex.1003, ¶58. Baker also discloses that multiple videoconferencing streams are combined, and provides additional implementation details for displaying multiple video streams on the same display, with some of the video streams having a reduced size. Ex.1006, Figs. 10A, 10B, 8:40-62; Ex.1003, ¶59. In this way, Baker's features would further define Setton's video stream display and interfaces while also accommodating network constraints of different participants. Using Baker's miniaturized strip of participants may reduce the size of the data traffic that would have otherwise been

required from full-size video feeds. The lower data stream requirements would result in reduced delay for the receiving peers. Ex.1003, ¶59. These benefits provided by Baker are consistent with the purposes of Setton's peer-to-peer network system to accommodate different quality of service requirements (e.g., communication stability) and constraints (e.g., network load, delay). Ex.1005, [0014]; Ex.1003, ¶59. The results of the combination would have been predictable for the same reasons explained above (e.g., achieved by routine software programming). Ex.1003, ¶59.

A POSITA would have been motivated to incorporate Baker's film strip of reduced-size video feeds in Setton's peer-to-peer network. Ex.1003, ¶60. The incorporation would provide flexibility for users (e.g., non-speakers) with network constraints while prioritizing network resources for higher-priority conference participants (e.g., the active speaker). Ex.1005, [0017], [0026]; Ex.1003, ¶60. Baker's film strip features would also enhance Setton's video conference experience. Ex.1003, ¶60. Higher-priority participants can be allotted more network resources (e.g., bandwidth) for high-quality video streams, but lower-priority participants (e.g., audience members, non-speaking panel members) can still be monitored. Ex.1005, [0017]; Ex.1003, ¶60.

The proposed combination would have been predictable and a POSITA would have had a reasonable expectation of success, because the combination of

Setton's peer-to-peer network system with Baker's film strip would have been well-within the skill level of a POSITA. Ex.1003, ¶61. Setton's peers already stream "video data" for a "video teleconference," but Setton provides limited detail on how those streams are displayed, indicating that a POSITA would have had the requisite background knowledge. Ex.1005, [0019], [0034]; Ex.1003, ¶61.

Modifying Setton's display of video teleconferencing data into a film strip format, per Baker, would have been a straight-forward exercise of programming per Baker's well-known techniques. Ex.1006, Figs. 10A, 10B, 8:40-62; Ex.1003, ¶61.

4. Claim 3

[3.0] A teleconferencing system comprising:

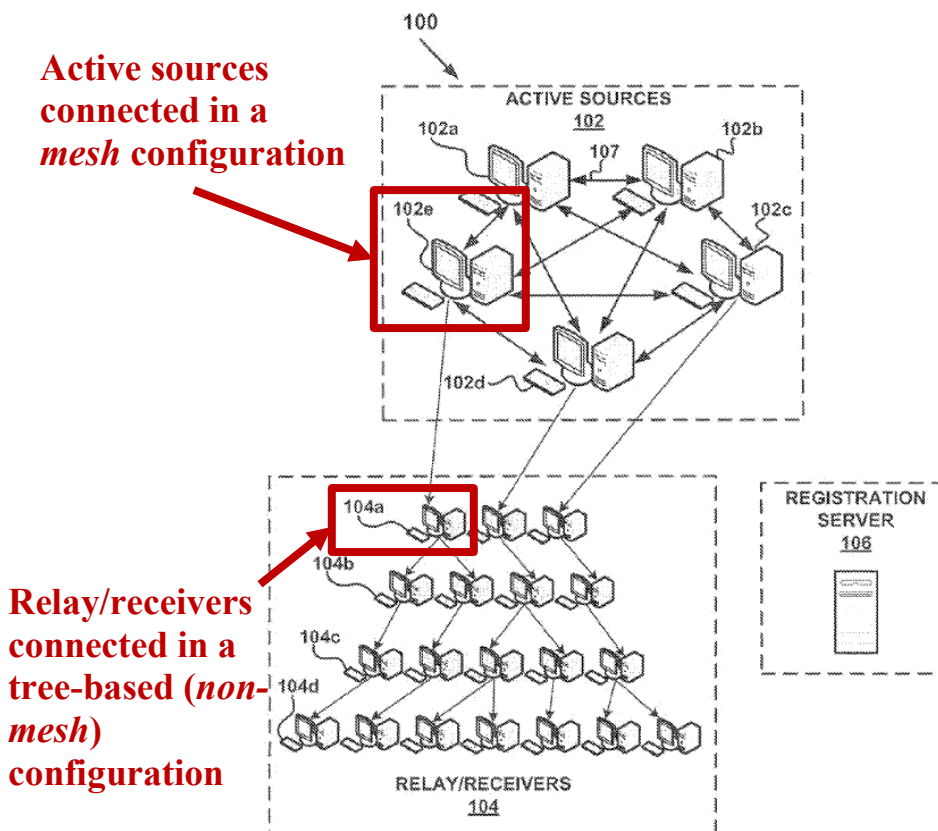
Setton discloses a "peer-to-peer media streaming" network that is used for "video teleconferenc[ing]." Ex.1005, Abstract, [0014], [0019]. The peers of the network include "active sources" and "relay/receivers." Ex.1005, [0014]. "[T]he active sources 102 use a mesh-based data aggregation peer-to-peer protocol, while the relay/receivers 104 use a tree-based video multicast protocol." Ex.1005, [0018]. Setton's active sources generate an aggregated media stream which is "transmitted to all the relay/receiver peers via application layer multicast tree(s)." Ex.1005, [0014]; Ex.1003, ¶63.

Setton's "multiway peer-to-peer media streaming" network, including the "active sources" and "relay/receivers," discloses the claimed "*teleconferencing*

system.” Ex.1003, ¶64. Thus, to the extent the preamble is limiting, Setton discloses this limitation. Ex.1003, ¶62.

[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

Setton’s active sources 102 are *endpoints...connected in a mesh configuration*. The relay/receivers 104 are *endpoints...connected in a non-mesh configuration*. Ex.1003, ¶¶65-66. Setton’s Fig. 1, below, depicts multiple active sources 102 and relay/receivers 104. Ex.1003, ¶67.



Ex.1005, Fig. 1 (annotated)

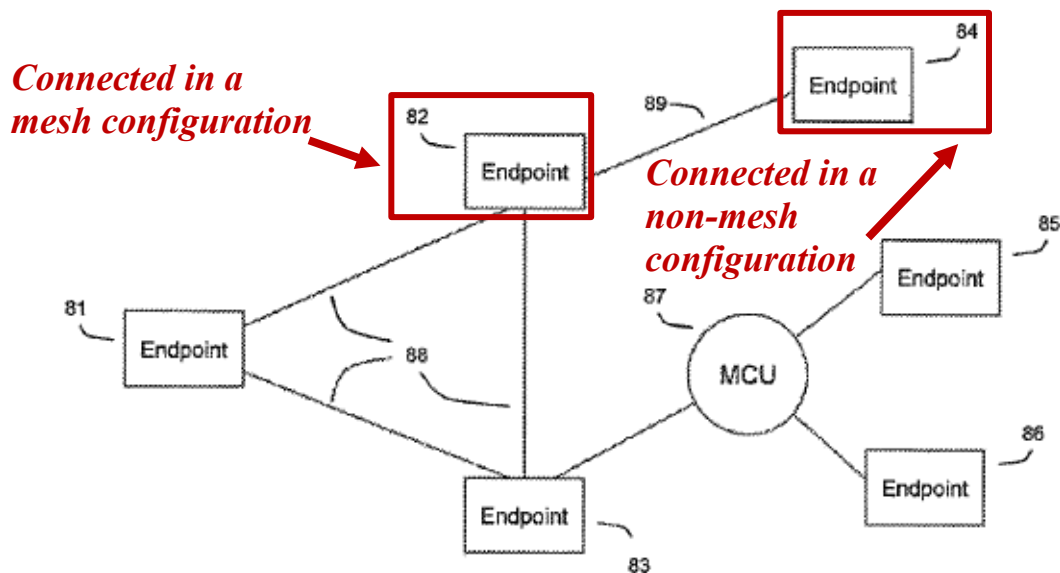
The active sources 102 and relay/receivers 104 may include computers. Ex.1005, Figs. 1 and 4, [0016], [0014], [0032]. The active sources 102 may be used by participants in a panel discussion, and the relay/receivers may be used by “member[s] of the audience.” Ex.1005, [0026]. The active sources 102 and relay/receivers 104 are like the ’099 patent’s *endpoints*. For instance, the ’099 patent explains that endpoints include “video cameras,” “monitors,” and/or “display controller[s]” that have a “network address (e.g., IP address or URL) used for call set-up.” Ex.1001, 5:4-7, 5:13-17; Ex.1003, ¶¶68-69. A POSITA would have understood that Setton’s peers are *endpoints*. Ex.1005, [0032]; Ex.1003, ¶69.

As shown in Fig. 1 above, the active sources 102 (*endpoints*) communicate directly with one another using “a **mesh-based** data aggregation peer-to-peer protocol” to aggregate video streams in a teleconference to transmit to the relay/receivers 104 (*endpoints*). Ex.1005, [0018]; Ex.1003, ¶¶70-71. The aggregated streams are transmitted through the relay/receivers using “a **tree-based** peer-to-peer video multicast protocol.” Ex.1005, [0023]; Ex.1003, ¶74.

Setton’s mesh-based active sources 102 communicate with each other directly, with each having a direct connection to each other active source. Ex.1005, [0023]. This is consistent with the ’099 patent’s description of the “mesh configuration” as one where “each endpoint is connected directly to each other endpoint in the conference.” Ex.1001, 2:34-37; Ex.1003, ¶¶72-73. Thus, the active

sources 102 are *endpoints...connected in a mesh configuration*.

The relay/receivers 104 communicate with each other using a tree-based protocol; they are *connected in a non-mesh configuration*. See Ex.1005, [0014], [0022], [0023]; Ex.1003, ¶¶74-75. Setton's tree-based connections are consistent with the '099 patent's description of an endpoint connected in a non-mesh configuration, which are also not directly connected to each endpoint of a mesh configuration. Ex.1001, 15:46-16:35; Ex.1003, ¶76. As shown in Fig. 8, non-mesh endpoint 84 (*non-mesh connected*) is connected only to the endpoint 82:



Ex.1001, Fig. 8 (annotated)

In Setton's Fig. 1, above, relay/receiver 104a (*endpoint...connected in a non-mesh configuration*) is connected to the mesh-networked active sources 102 (*endpoints connected in a mesh configuration*) via active source 102e. Ex.1003, ¶77. An active source 102 is "*selected*" as a "*root[]*" of the application layer

multicast tree” to deliver the aggregated media stream to the relay/receivers 104:

[T]he aggregated media stream is transmitted to the relay/receivers 104 (e.g., 104 a, 104 b, 104 c, and 104 d) via a tree-based peer-to-peer video multicast protocol. Such a tree is depicted in FIG. 1. In this embodiment, **the roots of the application layer multicast tree are selected** from the active sources (102 c, 102 d, and **102 e**) which aggregate the media stream as described above.

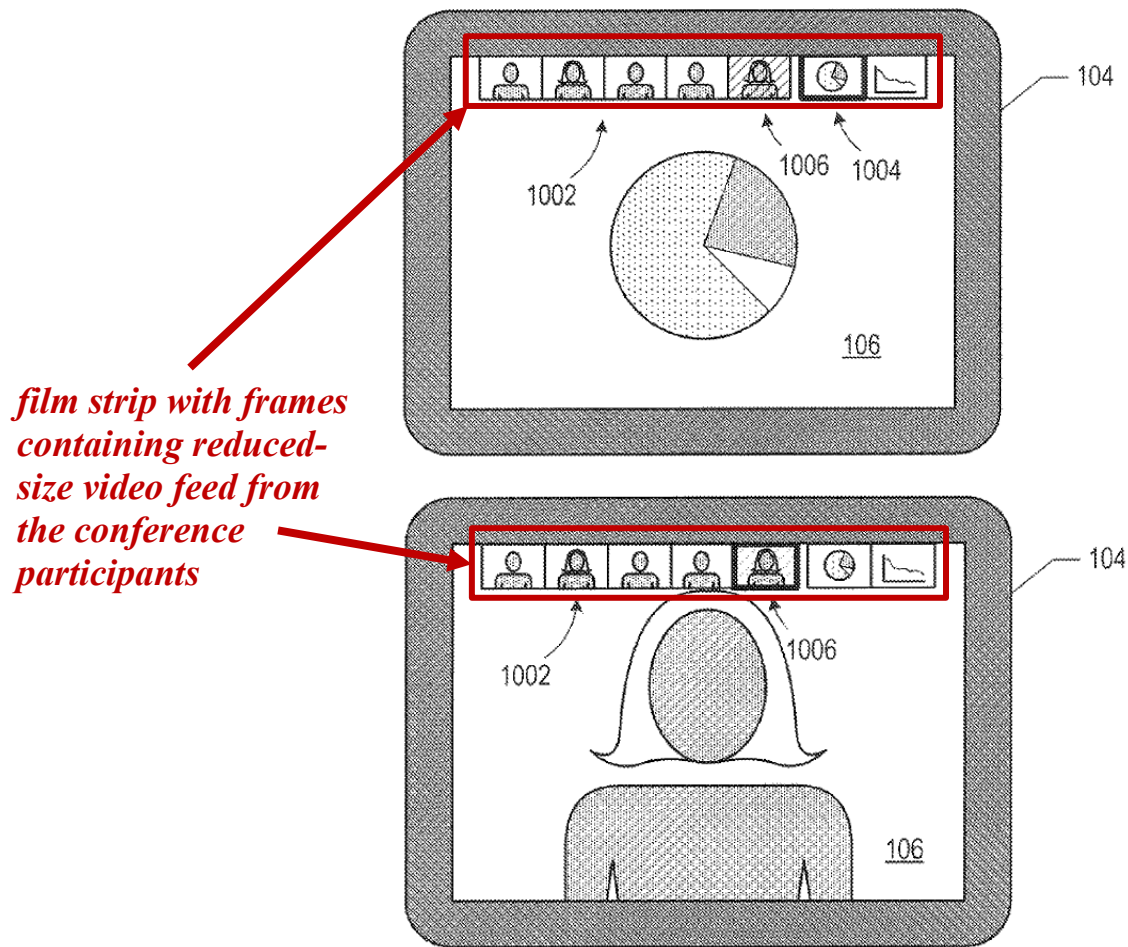
Ex.1005, [0023]; Ex.1003, ¶78.

Active source 102e was therefore *selected* to provide the connection point in the multicast tree for relay/receiver 104a to the receiver the aggregated media stream from the active sources 102. Thus, Setton discloses that the relay/receiver 104a is connected in a non-mesh configuration *to a selected one of the mesh-connected endpoints*. Ex.1003, ¶¶79-80.

[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;

Setton’s peer-to-peer network “provide[s] a real-time multi-way video communication protocol based on peer-to-peer technologies.” Ex.1005, [0014]. Setton describes aggregating and transmitting “video stream[s]...generated by [attendees] of a video conference for attendees in a video teleconference” (*video feed from the conference participants*). Ex.1005, [0019]; Ex.1003, ¶82.

Like Setton, Baker also describes displaying “video and audio streams” from teleconference participants. Ex.1006, 2:40-43; Ex.1003, ¶83. Baker “permit[s] the local participant 102 to control the manner in which video data streams are to be **received** and displayed” to “**reduce the size of the data streaming**...from the network 100.” Ex.1006, 8:63-66. In Baker, “video streams” of participants in a teleconference are displayed as “a strip of miniaturized images,” which Baker also refers to as “**reduced size**.” Ex.1006, 9:23-25. “[A]n example strip 1002” is shown in Figs. 10A and 10B, below. Ex.1006, 8:45-47; Ex.1003, ¶¶83-84.



Ex.1006, Figs. 10A and 10B (annotated)

As shown in Figs. 10A and 10B, strip 1002 includes the “**reduced size images of the live remote participant viewing areas** and two content viewing areas located along the top of the display.” Ex.1006, 8:40-47; Ex.1003, ¶84.

Baker explains that the video streams of non-speaker participants “can be sent with a lower bit-rate/quality” to save on processing power and may be “decoded at a slower frame [rate].” Ex.1006, 9:11-14, 19-26. Baker’s “strip of miniaturized remote participant viewing areas” on the display is the claimed *film*

strip with frames containing reduced-size video feeds from the conference participants. Ex.1006, 8:40-44; Ex.1003, ¶¶85-86.

A POSITA would have found it obvious to display, on the screens of the active sources 102 and relay/receiver 104a (*mesh-connected endpoints* and *non-mesh connected endpoint*), Setton’s aggregated video streams of the teleconference participants (*video feed from the conference participants*) as a strip of miniaturized images (*film strip with frames containing reduced-size video feeds from the conference participants*), per Baker. Ex.1003, ¶87. The ’099 patent acknowledges that presenting smaller video feeds from remote endpoints in a “*film strip*” was known in the art. *See* Ex.1001, Figs. 3b-3c, 4:42 (“typical prior art screen layouts”); Ex.1003, ¶89. Other prior art patent applications confirm that use of a “*film strip*”-style user interface was common. *See, e.g.*, Ex.1019, [0047], Figs. 4 & 9; Ex.1018, Fig. 10; Ex.1017, [0060], Figs. 9, 17, 18; Ex.1003, ¶88. Thus, a POSITA would have found it obvious to combine Baker’s miniaturized images with Setton’s videoconferencing system because showing miniaturized images of videoconferencing participants was a common UI technique. Ex.1003, ¶89.

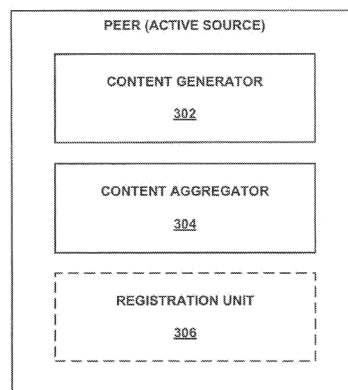
The combination supports Setton’s objective to “accommodate the peers’ different quality of service or delay constraints.” Ex.1005, [0014], [0028]; Ex.1003, ¶87. A POSITA would also have recognized the benefits of using Baker’s strip of miniaturized images to display non-speaking participants or audience

members in order to “maintain very low latency” and allow speaking participants to “converse naturally.” Ex.1005, [0004]; Ex.1003, ¶90; Section X.C.3.

Thus, the combination of Setton and Baker renders this limitation obvious.
Ex.1003, ¶¶81, 91.

[3.2] a proxy at the selected mesh-connected endpoint for the non-mesh connected endpoint and connected to each of the mesh-connected endpoints to receive video feeds from the mesh-connected endpoints for the non-mesh connected endpoint;

Setton describes “a content aggregator 304 for aggregating the content generated by the peer with content generated by other peers” (*a proxy*). Ex.1005, [0021]. Setton explains that content aggregator 304, and content distribution functionalities, may be incorporated into an active source 102 (*mesh-connected endpoint*) to distribute the aggregated content to the other peers, including the relay/receivers 104 (*for the non-mesh connected endpoint*). Ex.1005, [0021], Fig. 3; Ex.1003, ¶¶93-95.

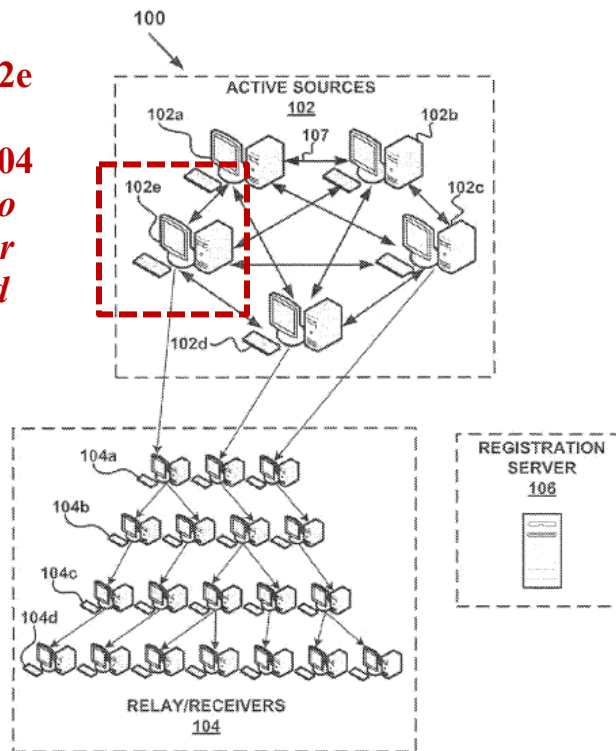


Ex.1005, Fig. 3

Setton's Fig.1 shows active source 102e connected to the relay/receiver 104a. The relay/receivers 104 receive aggregated streams from the active sources 102. Ex.1005, [0023], [0019]; Ex.1003, ¶95. In this architecture, a POSITA would have recognized that implementing the content aggregator 304 on the active source 102e would provide an efficient communication pathway to distribute the aggregated streams to the relay/receivers 104. Ex.1003, ¶98. Thus, to the extent Setton does not specifically disclose that the active source 102e (*selected mesh-connected endpoint*) includes the content aggregator 304 (*proxy*), this would have been obvious to a POSITA. Ex.1003, ¶¶97-98, 103.

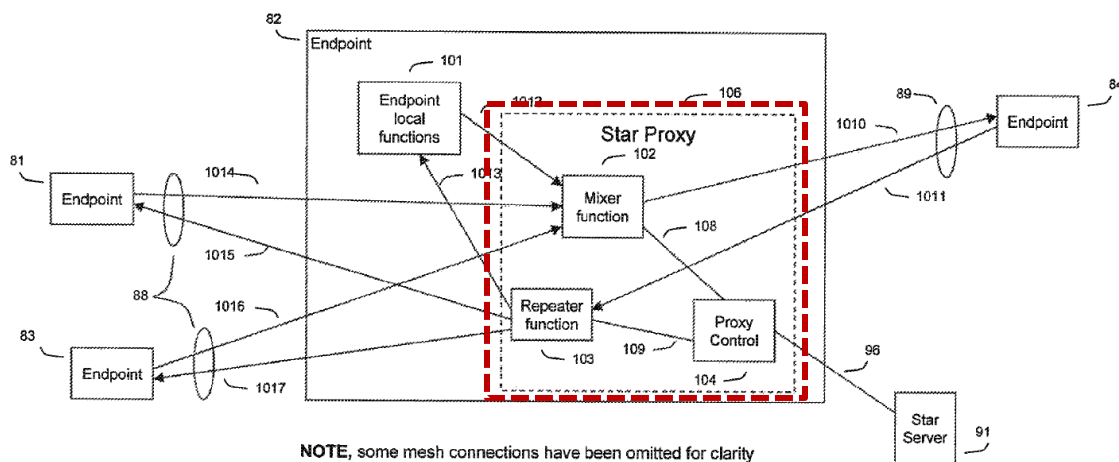
Setton's peers may include, in addition to the content aggregator 304, a content distribution feature "which establishes a bridge between active sources 102 and relay/receivers 104." Ex.1005, [0021]; Ex.1003, ¶¶99, 102. Fig. 1 shows that active source 102e's content aggregator 304 (*proxy*) is connected to each of the *mesh-connected* endpoints. Ex.1003, ¶99. Fig. 1 also shows active source 102e providing the connection point for the relay receivers 104 to the rest of the active sources 102. It would have been obvious for the active source 102e's content aggregator 304 to aggregate content for distribution to the relay/receiver 104a (*proxy...for the non-mesh connected endpoint*). Ex.1003, ¶¶98-102.

**Active source 102e
supporting
relay/receivers 104
and connected to
each of the other
mesh-connected
endpoints**



Ex.1005, Fig. 1 (annotated)

This is consistent with the '099 patent's description of the proxy as including software that performs "Mixer and Repeater" functionality. Ex.1001, 16:40-46, Fig. 10; Ex.1003, ¶96.



'099 patent, Fig. 10 (annotated)

Thus, Setton renders obvious the content aggregator 304 running on the active source 102e to aggregate and distribute content to the relay/receivers 104, such as 104a (*proxy...for the non-mesh connected endpoint*). Ex.1003, ¶103.

The active sources 102, executing content aggregators 304, “generate an aggregated media stream” from “a plurality of media streams,” including those generated the other active sources 102 (*proxy...receive[s] video feeds from the mesh-connected endpoints*). Ex.1005, [0019]; Ex.1003, ¶104. The aggregated media stream is “transmitted to all the relay receiver peers via application layer multicast trees.” (*...for the non-mesh connected endpoint*). Ex.1005, [0014]; Ex.1003, ¶¶100, 104.

Thus, Setton renders obvious that the content aggregator 304 at the active source 102e (*proxy*) *receive[s] video feeds from* the other active sources 102 (*the mesh-connected endpoints*) to be distributed to the relay/receiver 104a (*for the non-mesh connected endpoint*). Ex.1003, ¶¶92, 105-106.

[3.3] a server associated with the proxy for serving the film strip with frames containing reduced-size video feeds from the mesh-connected conference participants to the non-mesh connected endpoint; and

As discussed in [3.1.2], Baker’s “strip of miniaturized remote participant viewing areas” discloses the *film strip with frames containing reduced-size video feeds from conference participants*. As discussed in [3.2], Setton’s content

aggregator 304 running on active source 102e discloses or renders obvious the *proxy at the selected mesh-connected endpoint*. Ex.1003, ¶¶108-109.

The active sources 102 transmit aggregated streams “to all relay/receiving peers.” Ex.1005, [0014]. A peer (e.g., active source 102e) may include a feature “for providing content distribution and which establishes a bridge between active sources 102 and relay/receivers 104.” Ex.1005, [0021]. A POSITA would have understood that the content distribution feature, which transmits the aggregated streams, is a *server*. Ex.1003, ¶¶109-110.

Additionally, or alternatively, it would have been obvious to a POSITA for active source 102e to run server software (*server*) to accomplish the content distribution and transmitting function. Ex.1003, ¶111. Setton contemplates incorporating additional server functionality into an active source. For instance, Setton contemplates “incorporat[ing] the functionality of [registration] server 106” into an active source peer. Ex.1005, [0027]; *see also* [0021]; Ex.1003, ¶¶111-112.

Further, Setton’s peer-to-peer media streaming “can operate on or within a number of different computer systems including general purpose networked computer systems, [and] server devices.” Ex.1005, [0032]; Ex.1003, ¶113. This is consistent with the Star Server 91 of the ’099 patent that “could be a program running on a computer within one of the mesh endpoints e.g. 11 in Fig. 1.”

Ex.1001, 17:3-5. A POSITA would have found it obvious for the active source

102e to perform the role of a *server* via server software. Ex.1003, ¶113.

It was well-known to a POSITA to run multiple software servers and functions on a single device. Ex.1003, ¶¶113-114. A POSITA would have recognized the benefits (e.g., flexible network configurations) of incorporating server software and functionality (*server*) at the active source 102e for transmitting the aggregated video media stream from the active sources to the relay/receiver peers (*for serving the film strip with frames containing reduced-size video feeds from the mesh-connected conference participants to the non-mesh connected endpoint*). Ex.1003, ¶114. It would have been obvious to a POSITA for the active source 102e to run the server software (e.g., for content distribution) (*server*) in addition to the content aggregator 304 (*associated with the proxy*). Ex.1003, ¶114.

Further, it would have been obvious for the content distribution feature (*server*) to *serve* Baker's strip of miniaturized images as "reduced size" images, in addition to the images being displayed as miniaturized images. Ex.1003, ¶¶115-118. Baker's methods "permit[] the local participant 102 to control the manner in which video data streams are to be **received** and displayed" in order to "**reduce the size of the data streaming**...from the network 100." Ex.1006, 8:63-67. It would have been obvious to a POSITA for the content distribution to serve Baker's miniaturized images in a reduced data size, as this would "reduce the size of data streaming" and allow video streams to "be sent with a lower bit-rate/quality" to

save on processing power, and may be “decoded at a slower frame [*sic*].” Ex.1006, 9:11-14, 19-26; Ex.1003, ¶¶85, 116-117. It was generally known in the art to reduce the size of images placed into a “*film strip*”-style display of remote conference participants. *See, e.g.*, Ex.1018, [0049], [0051] (describing “scalers” that combine “plurality of pixels in a video image 555 for each pixel scaled”); Ex.1017, [0061] (describing “down-sampling” or “cutting/cropping out some video pixels” to produce a “reduced-resolution video”); Ex.1003, ¶¶115-116.

Thus, the combination of Setton and Baker renders this limitation obvious. Ex.1003, ¶¶107, 117-118.

[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.

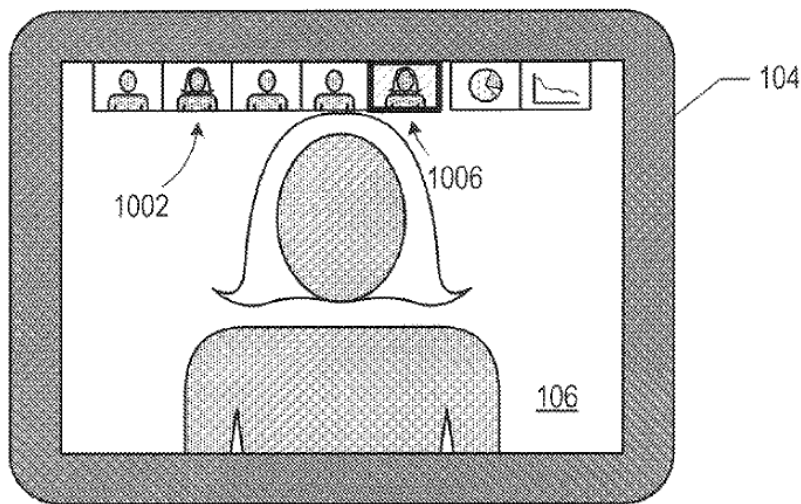
As discussed in [3.3], it would have been obvious for Setton’s active source 102e to include or implement a content aggregator and distributor to distribute the aggregated video streams to the relay/receivers 104. Ex.1003, ¶¶119-121.

Further, Setton’s registration server, which may be incorporated into an active source peer (Ex.1005, [0027]), assists peers in joining a teleconferencing session and “transition[ing] between being a member of active sources 102 and relay/receivers 104,” such as when a member of the audience joins the panel discussion between the active peers. Ex.1005, [0026]. Setton’s registration server

functionality at active source 102 allows users at a relay/receiver peer (*user located at the non-mesh connected endpoint*), through a connection to the active source 102, to change their membership to an active source. Ex.1003, ¶¶122-123.

Referring to Fig. 10B, Baker discloses that a local participant can “touch or tap on [a] **highlighted image 1006 of participant**” in a reduced-sized video feed on the strip (*select a video feed on the film strip*) to bring up that conference participants viewing area on the display (*for promotion to full size*).

Ex.1006, 8:40-62; Ex.1003, ¶124.



Ex.1006, Fig. 10B

It would have been obvious to a POSITA to make this touch-to-enlarge functionality available for all conference attendees, as it would predictably improve the attendee’s experience by permitting the attendee to control which available video stream is the focus of the attendee’s display. Ex.1003, ¶125. As discussed in [3.3], it would have been obvious for Setton’s active source to provide

“*reduced-size video feeds*” for the videos in the “*film strip*” (e.g., item 1002 in Baker’s Fig.10B, above). To facilitate displaying a selected video feed (e.g., 1006 in Baker’s Fig. 10B, above) at full size in the main part of the display (e.g., 106 in Baker’s Fig. 10B, above), it would have been obvious to a POSITA that the endpoint should receive the selected video stream as a *full-size* video feed. Ex.1003, ¶¶125-127. Hence, it would have been obvious for the active source 102e to further include a *control connection* to request transmission of the selected video stream as a full-size video feed. Ex.1003, ¶123. Sending a command from user’s touchscreen to control the size and delivery of video streams in a videoconference was known in the art. *See, e.g.*, Ex.1020, [0037], [0039]; Ex.1003, ¶125. Thus, a participant at the relay/receiver 104a would be able to touch or tap on images in the miniaturized strip and cause an enlarged version of the selected video feed to show up on the participant’s large main display area (*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*). Ex.1003, ¶¶126-128.

Setton explains that the active source peers have more “stringent” network constraints (e.g., delay requirements), compared to the relay/receiver peers “who do not suffer from stringent constraints.” Ex.1005, [0014]. Baker explains that the local participant viewing the conference can “move the display 106 to view

participant 5, as shown in FIG. 10.” Ex.1006, 8:52-56. Baker also contemplates transmitting video frames of speaking participants with a higher bit-rate/quality than the non-speaking participants. Ex.1006, 9:9-26, 45-49. A POSITA would have appreciated that Baker’s reduced-quality miniaturized strip of non-speaking participants would benefit Setton’s peer-to-peer network because of the different network constraints and requirements imposed on different peers. *See, e.g.*, Ex.1005, [0014]; Ex.1003, ¶¶128-131.

In addition, a POSITA would have been motivated to allow a non-mesh participant to select which video feed to be promoted to full-size, per Baker. This would provide participants the flexibility to focus on a speaker or shared content of interest to the participant. Ex.1003, ¶132. Further, Baker’s display interface of the full-size video with miniaturized film strips reduces visual complexity for larger meetings with many participants. Ex.1003, ¶132. Allowing participants to select a video for full-size streaming increases accessibility of the video conference platform by allowing participants to modify the display to meet their needs. Ex.1003, ¶132. For example, participants may view particular feeds in full-size, such as the video feed of an interpreter. Ex.1003, ¶132; Section X.C.3.

A POSITA would have understood that incorporating Baker’s touch or tap control functionality into Setton’s peers would involve software modifications to provide the UI objects and corresponding actions (e.g., promote to main display

area). Implementing the *control connection* by software is consistent with the '099 patent's description of "control connections" as including "software 'handles' [to] control the connection and operation of the Mixer and Repeater respectively using well-known MCU technology." Ex.1001, 16:43-46; Ex.1003, ¶¶125-126.

Computer programming, including web application programming, would have been a core part of a POSITA's technical experience and expertise. Ex.1003, ¶61.

In particular, a POSITA would have had ample experience developing software and/or web pages that include UI objects that define how video streams are displayed, and how the display may have been controlled. Ex.1003, ¶61.

Thus, the combination of Setton and Baker render this limitation obvious, ¶133.

5. Claim 10

[10.0], [10.1], [10.2]

Limitation [10.0] is substantively identical to limitations [3.0], [3.1.1], and [3.1.2], where the claimed *endpoint for use as a mesh-connected endpoint in a teleconferencing system* is Setton's active source 102e. This limitation is rendered obvious for the same reasons explained in [3.0], [3.1.1], and [3.1.2]; Ex.1003, ¶134.

Limitation [10.1] is substantively identical to limitation [3.2] and is rendered obvious for the same reasons. Ex.1003, ¶135.

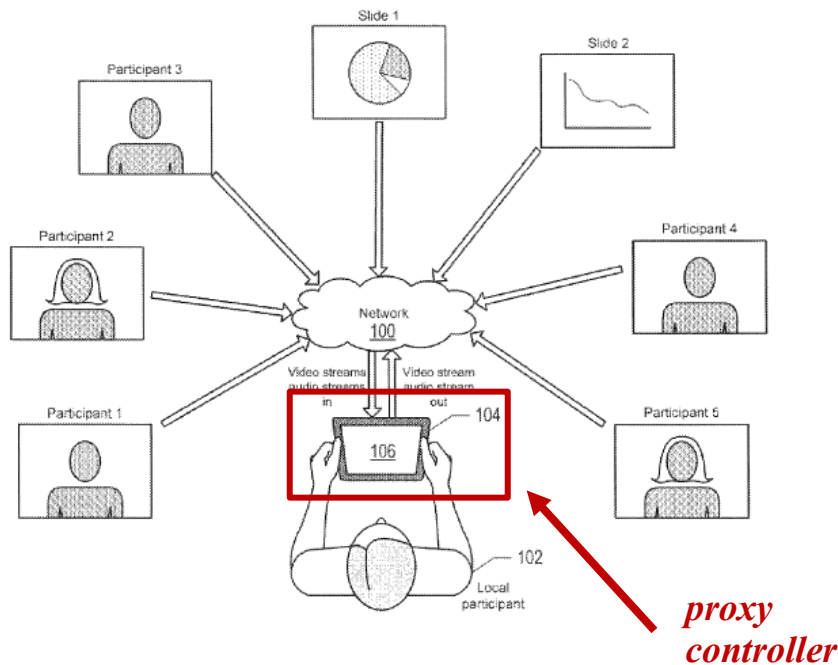
Limitation [10.2] is substantively identical to limitation [3.3] and is rendered obvious for the same reasons. Ex.1003, ¶136.

[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.

As explained in [3.2], Setton’s active sources 102 (*mesh-connected endpoint*) include “a content aggregator 304 for aggregating the content generated by the peer with content generated by other peers” (*proxy*). Ex.1005, [0021]; Ex.1003, ¶138.

As explained in [3.4], it would have been obvious to a POSITA that Baker’s “touch or tap” functionality would be facilitated by a *control connection* because the act of touching or taping controls what content is displayed on the local participant’s display and, as a result, triggers the full size video feed to be transmitted to and displayed on the local participant’s device. Ex.1003, ¶¶137, 139. Thus, to the extent the *proxy controller* is interpreted to be the same as the *control connection*, the combination of Setton and Baker renders this limitation obvious for the same reasons explained in [3.4]. Ex.1003, ¶139.

If the *proxy controller* is interpreted as a feature distinct from the *control connection*, Setton and Baker also render this limitation obvious. Ex.1003, ¶140. Baker discloses a “PED [portable electronic device] 104”: a *proxy controller*.



Ex.1006, Fig. 1

The user of PED 104 selects which video stream to be displayed in full size on the display 106 using touch commands. Ex.1006, 8:52-62; Ex.1003, ¶141. A POSITA would have understood that these touch commands cause the full-size video stream to be transmitted from active source 102e (*mesh connected endpoint*) to the relay/receiver 104 (*non-mesh connected endpoint*) by active source 102e's "content aggregator 304 for aggregating the content generated by the peer with content generated by other peers" (*proxy*). Ex.1005, [0021]; Ex.1003, ¶141.

Thus, the combination of Setton and Baker render this limitation obvious. Ex.1003, ¶¶137,142.

D. Ground 2

1. Summary of Wong

Wong describes a video share casting service that includes “streaming video casting from a first group to a first video portal,...and streaming the first video portal to a web-browser through a flash proxy component.” Ex.1009, [0010].

Wong further describes “streaming [] one or more video portals **to a web-browser** accessing the web server via a proxy of **web-browser** plug-in media.” Ex.1009, [0013]; Ex.1003, ¶¶144-145.

For live video blogging, Wong describes a web browser automatically downloading plug-ins to provide a user interface for viewing video casts. Wong, [0041]. Wong’s video share casting service “can be further extended to a multi-casting service from single casting service in **a conferencing call or chat**,” where “multiple users can broadcast their video, and other users can select the cast to view.” Wong, [0100]; Ex.1003, ¶¶146-148.

Wong is analogous art to the ’099 patent and is from the same field of endeavor. The ’099 patent is directed to a video teleconferencing system that utilizes a web browser to display the user-side video teleconferencing interface. Ex.1001, 15:50-53. Wong similarly describes video casting, which includes video conferencing calls, “**to a web-browser** accessing the web server via a proxy of **web-browser** plug-in media.” Ex.1009, [0013], [0100]; *see also* Ex.1009, [0041]

(web browser “automatically download[s] a plug-in that can implement the multimedia communication to show the user the blog or the video cast” where the plug-in is software that “execute[s] within the browser[.]”). Wong is reasonably pertinent to a particular problem of the ’099 patent, which is using a web browser to display a videoconferencing interface. Ex.1001, 15:50-53; Ex.1009, [0013], [0041]; Ex.1003, ¶143.

2. Reasons to Combine Setton, Baker, and Wong

A POSITA would have been motivated to combine Setton and Baker for the reasons explained above. Section X.C.3. Further, a POSITA would have been motivated to combine Wong with Setton and Baker, and doing so would be combining prior art elements according to known methods to yield predictable results. Ex.1003, ¶¶149-152.

The incorporated feature would be Wong’s web browser for accessing and controlling video share casting services. Ex.1009, [0013], [0041], [0106], [0107]; Ex.1003, ¶¶149-150, 154. Combining Wong with Setton and Baker would be achieved through well-known and routine computer programming techniques. Ex.1003, ¶151. The results of combining Wong with Setton and Baker would have been predictable, because web development and web-based programming was commonly used for communications applications, and a web browser would facilitate the functionalities of Setton and Baker described above. Ex.1003, ¶152.

Combining Wong's web browser conference access technique with Setton and Baker would be applying a known technique to a known device ready for improvement to yield predictable results. Ex.1003, ¶¶152, 153, 155. Setton discloses techniques for aggregating video streams, and Baker describes how the aggregated video streams are displayed on the same display (e.g., reduced-size film strip). Ex.1006, Figs. 10A, 10B, 8:40-62; Ex.1003, ¶153. Setton and Baker provide limited details on the type of software or application platform used to power the display features, which is disclosed in Wong.

Wong describes using a web browser to participate in and control a video teleconference. Ex.1009, [0013], [0106], [0107]. Wong even describes using a browser for similar display techniques as Baker, such as displaying "thumbnails of live video casts", and that a "user can click on a thumbnail to view that particular blog or cast." Ex.1009, [0041]. Web browsers like Wong's provided a versatile, well-known platform for internet applications and services. Ex.1003, ¶156. Internet browsers would have advantageously promoted cross-platform compatibility, supporting multiple operating systems, and would be used by a variety of video teleconferencing service providers. Browsers were common and familiar to a wide range of users, and operable on a wide range of computers. Ex.1015; Ex.1016. Teleconferencing providers used standardized web programming languages to provide the display and user interface that is used by participants. Ex.1003, ¶156.

The proposed combination would have been predictable and a POSITA would have had a reasonable expectation of success, because the combination of Wong's web browser-based video share casting technique with Setton's peer-to-peer network and Baker's film strip would have been well-within the skill level of a POSITA. Ex.1003, ¶¶152, 155, 157. Setton's peers already stream "video data" for a "video teleconference," but Setton and Baker provide limited detail on the types of application platforms that could be used to view and control the conference, indicating that a POSITA would have had the requisite background knowledge. Ex.1005, [0019], [0034]; Ex.1003, ¶157. Baker describes storing "software applications including...browsers." Ex.1006, 3:29-34. Modifying Setton's distribution of video streams, and Baker's film strip, to be served to a web browser, would have been a straight-forward exercise of web application development per Wong's well-known techniques. Ex.1009, [0013], [0041], [0100]; Ex.1003, ¶¶156-157. This is another reason a POSITA would have been motivated to incorporate Wong's web browser features in the system resulting from the combination of Setton and Baker. Ex.1003, ¶157. The combination would have resulted in a greater number of people able to use the teleconferencing service, since web browsers were particularly common. Ex.1003, ¶¶156-157.

3. Claim 4

[4.0] A teleconferencing system as claimed in claim 3, wherein the server is configured to serve the filmstrip to a browser co-located with the non-mesh connected endpoint.

As discussed in [3.3], it would have been obvious to a POSITA that active source 102e would include and run a content distribution feature (*server*) to accomplish Setton's transmitting function. Ex.1003, ¶159. It would also have been obvious, as discussed in [3.3], that Setton's *server* is used to *serve the filmstrip to...the non-mesh connected endpoint*. Ex.1003, ¶160. As explained below, it would have been obvious for the content distribution or server software (*server*) on Setton's active source 102e to *serve* Baker's strip of video media to a *browser* on Setton's relay/receiver 104 (*non-mesh connected endpoint*). Ex.1003, ¶160.

Setton discloses that its system may be used to stream both "video data" and "web page-based data." Ex.1005, [0014]; Ex.1003, ¶160. Baker discloses a "video conference application" for transmitting video and viewing a "virtual meeting space." Ex.1006, 3:50-54. Ex.1003, ¶161. It was common to access and control video conferences via internet browsers. Ex.1013, 5:3-17; Ex.1014, 2:21-29. Video call services such as Facebook's Messenger Skype service and Google Hangouts were accessed through web browsers. Ex.1015; Ex.1016; Ex.1003, ¶¶162-163.

Wong describes a video streaming service that includes "streaming the one or more video portals to a web-browser accessing the web server via a proxy of

web-browser plug-in media.” Ex.1009, [0013]. For instance, “[t]he web browser can automatically download a plug-in that can implement the multimedia communication to **show the user the blog or the video cast.**” Ex.1009, [0041]. Wong’s browser-based video streaming technology supports “**a conferencing call.**” Ex.1019, [0100]; Ex.1003, ¶164.

It would have been obvious to a POSITA for the relay/receiver peers to use Wong’s web browser to participate in the video conferencing system taught by the combination of Setton and Baker, because it would eliminate the need for participants to download and install specialized programs at each endpoint. Ex.1009, [0103]. Internet browsers advantageously promote cross-platform compatibility, support multiple operating systems, and can be used by a variety of video teleconferencing service providers. Ex.1003, ¶¶165, 168; Section X.D.2.

A POSITA would have found it obvious for the *browser* to be installed and running on the relay/receiver 104 (*co-located with the non-mesh connected endpoint*). Ex.1003, ¶¶166, 169. Setton’s relay/receivers 104 may be audience members and more numerous than the active speakers connected in the mesh-based portion of the network. Ex.1005, [0014]. Baker also describes storing “software applications including...browsers.” Ex.1006, 3:29-34. It was well-known that internet browsers installed at the PC (*co-located*) were one of the principal mechanisms for users to communicate over the internet. Ex.1003, ¶¶166-167, 170.

Thus, the combination of Setton, Baker, and Wong render this claim obvious. Ex.1003, ¶¶158, 171.

4. Claim 5

[5.0] A teleconferencing system as claimed in claim 4, wherein the browser is displayed on a computer co-located with the non-mesh connected endpoint.

The '099 patent explains that an endpoint can be both “co-located” with a computer and “run on that same computer.” Ex.1001, 16:9-21, Fig. 9; Ex.1003, ¶174. As explained in [4.0], the combination of Setton, Baker, and Wong renders obvious Setton’s relay/receiver 104 (*non-mesh connected endpoint*) executing Wong’s browser to display Baker’s strip of video media (*co-located with the non-mesh connected endpoint*). Setton explains that the peers include computer systems. Ex.1005, [0032], Fig. 4. It was well-known that browsers provided a graphical user interface (*browser is displayed...*) for a user to view and interact with the content, including teleconferencing communications. A browser running on a relay/receiver peer computer system would be *displayed on a computer co-located with the relay/receiver 104*. Ex.1003, ¶¶173, 175-177.

Thus, the combination of Setton, Baker, and Wong renders this limitation obvious. Ex.1003, ¶¶172, 178.

5. Claim 6

[6.0] A teleconferencing system as claimed in claim 5, wherein the proxy and server are merged into a virtualized environment.

A POSITA would have found it obvious for Setton's content aggregator (*proxy*) and a content distribution feature (*server*) running on Setton's active source 102e to operate in conjunction with, and be supported by, a *virtualized environment*.

For instance, it would have been obvious for the proxy and server to be instantiated in a virtual server (*merged into a virtualized environment*) at the mesh-connected node implementing the central server 365. In other words, it would have been obvious to implement the central server 365's *proxy* and *server* functions into the mesh-connected node as a virtual server. Virtual servers and virtual machines were commonly used in internet communications, and a POSITA would have recognized the benefits of doing so. Ex.1003, ¶180. Virtual servers allow for more scalability and flexibility than dedicated, standalone servers. Ex.1003, ¶181. Virtual servers can also be modified or configured dynamically based on demand. Ex.1003, ¶181. This would help to achieve Setton's goal of "reduc[ing] the infrastructure" required for media streaming. Ex.1005, [0027].

In another example, it would have been obvious for at least some of the *proxy* and *server* functions to be performed or supported by the cloud (*merged into*

a virtualized environment). Ex.1003, ¶¶180-183. For instance, the *proxy* and *server* may distribute aggregated streams based on teleconference participant information and permissions managed by a cloud server in communication with the peers. It was well known to administer software and computing services virtually through a cloud. *See* Ex.1010, 5:50-51 (“cloud server system 220 runs virtualization software to provide a virtual machine”); Ex.1011, 4:40-46 (“a computing device can be...a server machine that is a physical or **virtual server** and hosted in an Internet public or private data center by a service provider or a third party for the service provider”). Ex.1003, ¶¶180, 185.

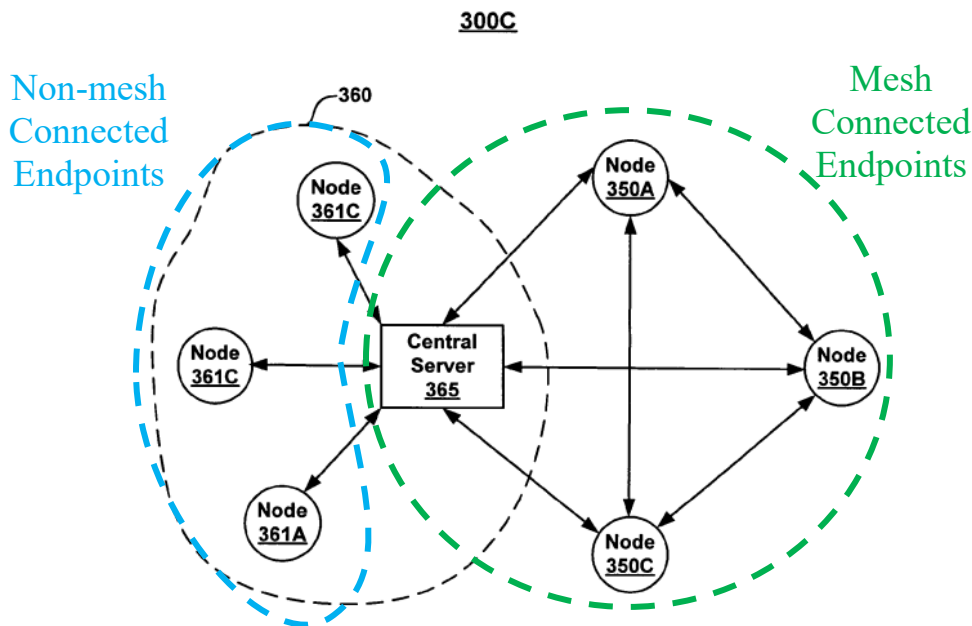
As explained above, it would have been obvious to implement the *proxy* and *server* as software. Ex.1003, ¶184. A POSITA would have understood that software could be at least partially administered or supported by a cloud (*merged into a virtualized environment*). Ex.1003, ¶185. This would help to achieve Setton’s goal of “reduc[ing] the infrastructure” required for media streaming (Ex.1005, [0027]), and free up computational resources to “accommodate the different quality of service requirements of” the active sources 102. Ex.1005, [0028], [0014]; Ex.1003, ¶¶186-187.

Thus, the combination of Setton, Baker, and Wong render this claim obvious. Ex.1003, ¶¶179, 188.

E. Ground 3

1. Summary of Apostolopoulos

Apostolopoulos relates to a “communication network” for an “N-way collaborative video conference.” Ex.1007, [0021], [0032]. Each participant can “transmit data streams (e.g., **audio and video data streams**) to some or all of the other participants.” Ex.1007, [0033], [0025]; Ex.1003, ¶¶189, 191. Apostolopoulos describes a “hybrid topology,” shown in Fig. 3C, that combines full-mesh and star topologies. Ex.1007, [0052]; Ex.1003, ¶192.



Ex.1007, Fig. 3C (annotated)

Apostolopoulos explains that “the full-mesh topology comprises node 350A, node 350B, node 350C, as well as central server 365 of the star topology 360.” Ex.1007, [0069]. In a full-mesh topology, “**a plurality of communication paths**

couple each of the plurality of collaborators together for transmitting and receiving [video streams]. Ex.1007, [0054], Fig. 3A; Ex.1003, ¶193. In a star topology, each node is connected to a “central server” (also called a central node) that routes data streams to other nodes. Ex.1007, [0056]; Ex.1003, ¶194.

In the hybrid topology, the “[star] nodes 361A-C are communicatively coupled to the central node 365,” which “communicate[s] in a full-mesh topology” with the full-mesh nodes 350A-C. Ex.1007, [0070]. The “central server/node 365 provides the processing power for generating communication data streams for the collaborators at nodes 361A-C to communicate amongst themselves, as well as with the collaborators at nodes 350A-C.” Ex.1007, [0070]; Ex.1003, ¶¶195-196.

Apostolopoulos is analogous art to the '099 patent and is “from the same field of endeavor.” *In re Bigio*, 381 F.3d at 1325. The '099 patent is directed to a video teleconferencing system for use by endpoints connected in a “mesh” and “non-mesh” configuration. Ex.1001, Abstract, 15:42-45. Apostolopoulos describes a “communication network” used to transmit audio and video data streams to conference participants. Ex.1007, [0021], [0033], [0025]. Apostolopoulos' conference participants are connected in a “full-mesh” and “star” (non-mesh) hybrid topology. Ex.1007, [0052], [0069], Fig. 3C. Apostolopoulos is also reasonably pertinent to the particular problem of the '099 patent, which is videoconferencing in mixed topology networks (e.g., Ex.1001, 15:46-49; Ex.1007,

[0021], [0033], [0025], [0052], [0069], [0070], Fig. 3C); Ex.1003, ¶190.

2. Reasons to Combine Apostolopoulos and Baker

Incorporating Baker's features into Apostolopoulos' network would have been simply combining prior art elements according to known methods to yield predictable results. The incorporated features of Baker would include the "strip of miniaturized remote participant viewing areas" as well as the scroll, touch, and tap control functionalities to control how the participant viewing areas appear on the screen. Ex.1006, 8:40-41, 57-62; Ex.1003, ¶197. The combination would involve implementing these features from Baker into Apostolopoulos' central node 365. Ex.1003, ¶197. The incorporation of Baker's functionalities in the central node 365 would be achieved through well-known, routine software application programming methods. Ex.1003, ¶197. A POSITA would have been capable of constructing application user interfaces including display areas and control objects using computer programming. Ex.1003, ¶197. The results would have been predictable, since the additional software features simply modify how the participant viewing areas are displayed and controlled, and would not affect the architecture of Apostolopoulos' peer-to-peer system. Ex.1003, ¶197.

The combination described above also amounts to applying a known technique to a known device ready for improvement to yield predictable results. Ex.1003, ¶198. Apostolopoulos describes computer networks for video

conferencing but provides fewer details on what the UI looks like, or how to control the display at a given device. Ex.1003, ¶198. Baker's disclosure provides additional disclosure related to how participants in a video teleconference are displayed, and how participants can control or modify that display (e.g., a single strip of multiple reduced-sized video streams of non-speaking participants). Ex.1003, ¶¶198-199. Baker's features would improve Apostolopoulos' video conferencing technology by providing a user-friendly display and accompanying UI control mechanisms. Ex.1003, ¶199. Further, Baker's miniaturized strip would advantageously accommodate network and power constraints of different participants, and reduce delay, which are consistent with Apostolopoulos' stated goals. Ex.1007, [0039], [0040], [0048]; Ex.1003, ¶¶199-200. The combination would involve routine software programming that would have been within the skill set of a POSITA. Thus, the results would have been predictable. Ex.1003, ¶199.

A POSITA would be motivated to provide the film strip of reduced-size video feeds to allow for some video feeds deemed more "important" (e.g., active speaker) to be prioritized while operating within network and power use limitations. *See, e.g.*, Ex.1007, [0082], [0083]; Ex.1003, ¶200. A POSITA would also have been motivated to provide the film strip of reduced-size video feeds because the film-strip advantageously allows participants to simultaneously focus on the speaking participant, while also monitoring all participants at once for

various purposes. Ex.1003, ¶200.

The proposed combination would have been predictable and a POSITA would have had a reasonable expectation of success, because the combination of Apostolopoulos' videoconferencing nodes with Baker's film strip would have been well-within the skill level of a POSITA. Ex.1003, ¶201. Apostolopoulos' nodes already transmit video data streams but Apostolopoulos provides limited detail on how those streams are displayed, indicating that a POSITA would have had the requisite background knowledge. Ex.1007, [0033], [0025]; Ex.1003, ¶201. Modifying Apostolopoulos' display of video streams into a film strip format, per Baker, would have been a straight-forward exercise of programming per Baker's well-known techniques. Ex.1006, Figs. 10A, 10B, 8:40-62; Ex.1003, ¶201.

3. Claim 3

[3.0] A teleconferencing system comprising:

Apostolopoulos discloses a “communication network” that is used for “N-way collaborative video conferenc[ing].” Ex.1007, [0021], [0032]. In the communication network, “each participant can possibly transmit data streams (e.g., audio and video data streams) to some or all of the other participants.” Ex.1007, [0033]; Ex.1003, ¶¶203-205. Apostolopoulos' “communication network,” including “nodes,” teaches the claimed “*teleconferencing system*.” Ex.1003, ¶206.

Thus, to the extent the preamble is limiting, Apostolopoulos discloses this

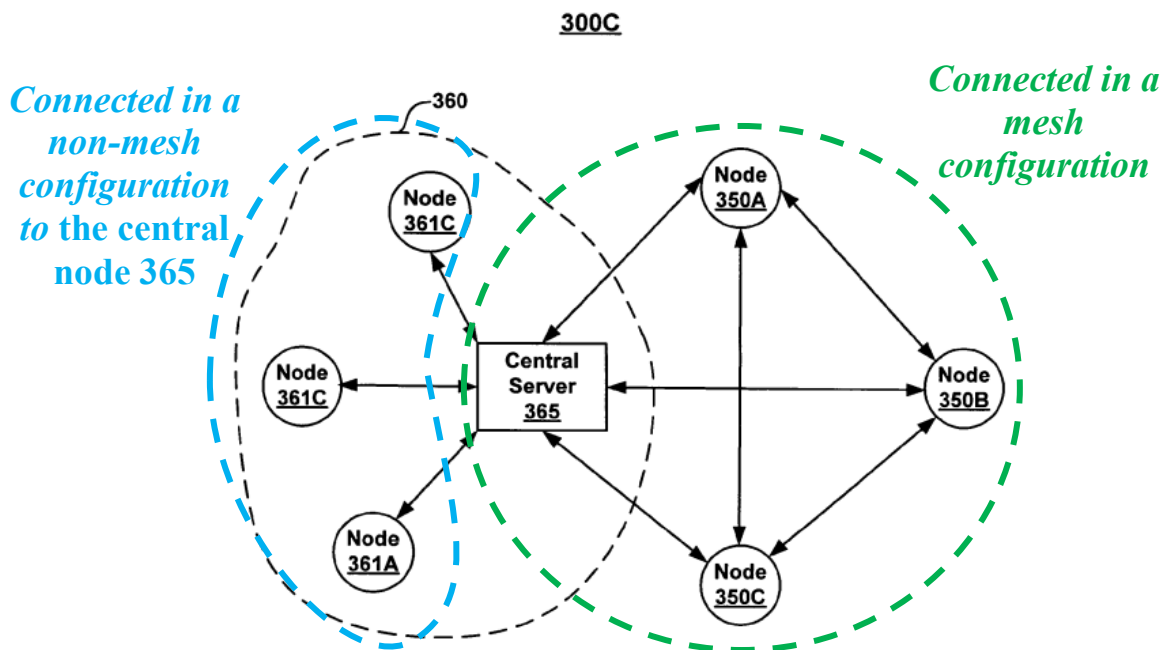
limitation. Ex.1003, ¶¶202, 207.

[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

Apostolopoulos' network comprising multiple "nodes" (*endpoints*) that produce "audio and video data streams." Ex.1007, [0025], [0033]. These nodes can be used by participants in a "communicative environment." Ex.1007, [0033].

Apostolopoulos' nodes are the claimed *plurality of endpoints*. Ex.1003, ¶¶208-209.

Hybrid topology 300C "combines features of the full-mesh topology of Fig. 3A and the star topology of Fig. 3B." Ex.1007, [0069]; Ex.1003, ¶¶210-211.



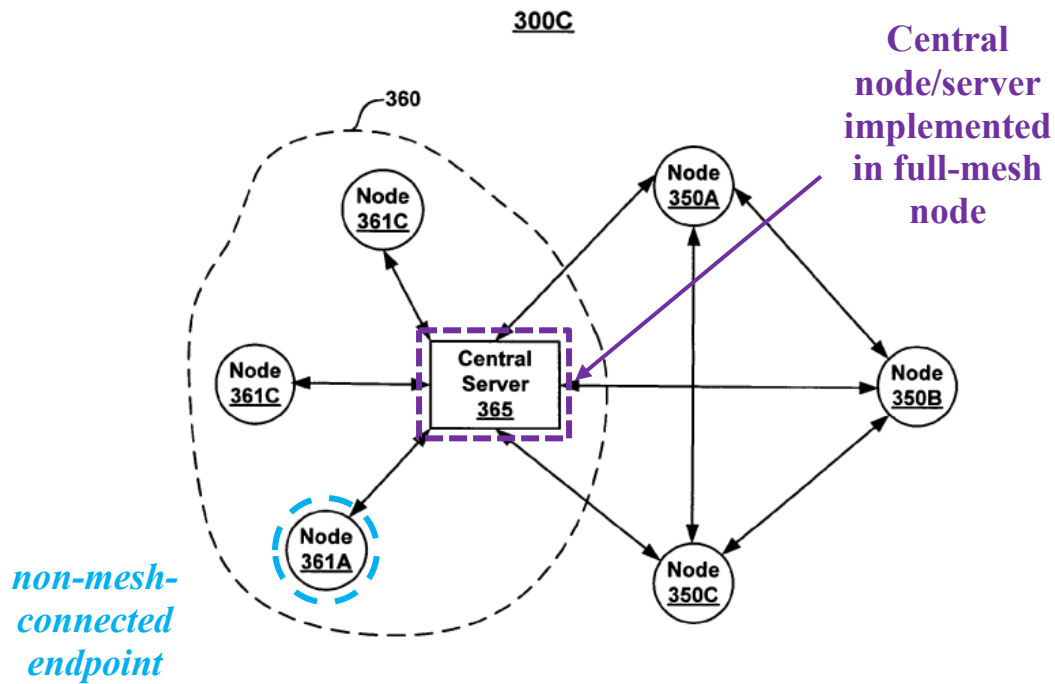
Ex.1007, Fig. 3C (annotated)

Apostolopoulos' full-mesh nodes are consistent with the '099 patent's

description of a “mesh configuration” where “each endpoint is connected directly to each other endpoint in the conference.” Ex.1001, 2:34-37; Ex.1003, ¶¶212-213. Apostolopoulos’ star topology nodes are consistent with the ’099 patent’s description of the non-mesh configuration where each node is connected to a server that routes data streams to each node, rather than being directly connected to each of the other nodes. ’099 patent, Fig. 9, 16:6-18; Ex.1003, ¶¶214-215.

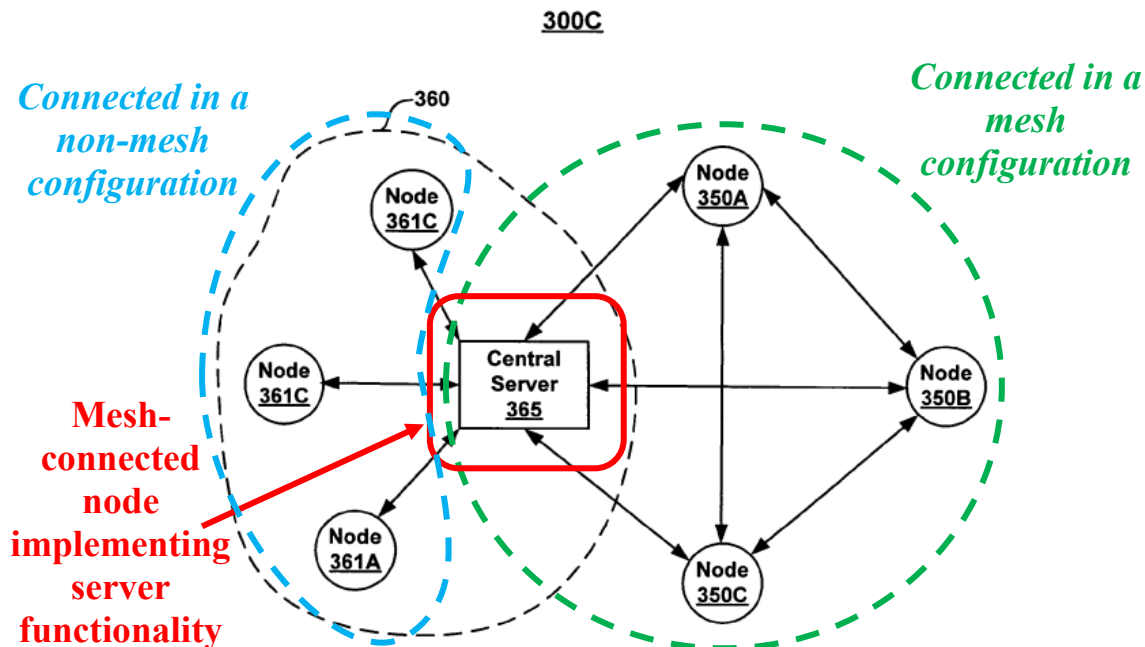
Thus, the full-mesh topology nodes 350A, 350B, 350C (*endpoints*) are *connected in a mesh configuration* and, the star topology nodes 361C and 361A (*endpoints*) are *connected in a non-mesh configuration*. Ex.1003, ¶216.

It would have been obvious to a POSITA for a star node 361, such as 361A, to be *connected in a non-mesh configuration to a selected one of the mesh-connected endpoints*. The node 361A is connected to central node/server 365, which is *mesh-connected* to the rest of the full-mesh nodes 350. Ex.1003, ¶217. Apostolopoulos refers to the central server 365 as a “central **node** 365,” and states that the central node/server is part of the “full-mesh topology.” Ex.1007, [0069], [0070]; Ex.1003, ¶218. This configuration is depicted in Apostolopoulos’ Fig. 3C:



Ex.1007, Fig. 3C (annotated)

A POSITA would have found it obvious for Apostolopoulos' central node/server 365 to be implemented as a full-mesh node similar to the other full-mesh nodes 350A-350C. Ex.1003, ¶219. Consistent with the other nodes (e.g., full-mesh nodes 350), Apostolopoulos explains that “central server 365” “generat[es] **communication data streams** for the collaborators at nodes 361A-C to communicate amongst themselves, as well as with the collaborators at nodes 350A-C.” Ex.1007, [0070]; Ex.1003, ¶220. Apostolopoulos' network provides for dynamic topology changes such that “computation is moved dynamically, as necessary, to meet various constraints, such as, computation, bandwidth, or low-power constraints.” Ex.1007, [0048]; Ex.1003, ¶¶220-221.



Ex.1007, Fig. 3C (annotated)

Implementing server functionality in an endpoint or end node was also well-known in the art. Ex.1003, ¶222. Baker discloses a “centralized computing device 302” that “can be **operated by one of the remote participants**” and “receives the video and audio streams from each of the participants and includes a server-side video conferencing application that processes each of the video and audio streams supplied by the participants.” Ex.1006, 3:58-4:4; Ex.1003, ¶222. And Kenoyer describes a “Multipoint Control Unit (MCU),” which “may be embedded in one of the video conferencing endpoints” for transmitting and receiving data streams from “video conferencing endpoints.” Ex.1008, 3:55-65; Ex.1003, ¶223.

In view of this disclosure, it would have been obvious to a POSITA to implement Apostolopoulos’ central server 365 as a full-mesh-connected node,

similar to nodes 350A-350C, with the central node 365 performing the role of a server. Ex.1003, ¶221. This would provide a more flexible network configuration that could be updated or reconfigured dynamically based on network constraints and the requirements of the conference participants. Ex.1003, ¶222.

As shown in Fig. 3C, the central node 365 is positioned in the network as a connection point between the star nodes 361 and the mesh nodes 350. The central node 365 was, at some point, *selected* as the node connecting the star nodes (*at least one of which is connected in a non-mesh configuration to a selected one of the mesh-connected endpoints*). A POSITA would also have found it obvious for the mesh node implementing the central server 365 to be selected by a network administrator or administration system, as Apostolopoulos itself explains that the network topology is selected to provide efficient communications. Ex.1007, [0052]; Ex.1003, ¶224. A POSITA would have found it obvious to select the full-mesh node implementing the central server to be part of the dynamic network topology selection process. Ex.1003, ¶225.

Thus, Apostolopoulos renders obvious *at least one* star node 361 *connected in a non-mesh configuration to a selected one of the mesh-connected endpoints*. Ex.1003, ¶226.

[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;

As explained in Ground 1, [3.1.2], Baker’s “strip of miniaturized remote participant viewing areas” is the claimed *film strip with frames containing reduced-size video feeds from the conference participants*. Ex.1003, ¶229.

It would have been obvious to a POSITA for Apostolopoulos’ nodes 350 (including mesh-connected node implementing the central server 365) and 361 to display Baker’s film strip of miniaturized participant viewing areas (*a film strip with frames containing reduced-size video feed from the conference participants is displayed*). Ex.1003, ¶¶228, 230. A POSITA would have been motivated to make the combination because, for example, the use of a “*film strip*”-style user interface was common in videoconferencing systems. *See, e.g.*, Ex.1019, [0047], Figs. 4 & 9; Ex.1018, Fig. 10; Ex.1017, [0060], Figs. 9, 17, 18; Ex.1003, ¶230.

Apostolopoulos’ technologies offer a “low complexity” video collaboration accommodating bandwidth constraints, and a collaborative visual environment. Ex.1007, [0062], [0037], [0052]; Ex.1003, ¶231. Apostolopoulos notes that “the person who is speaking may be identified as having higher importance than someone in the back row who is simply listening.” Ex.1007, [0082]. Therefore, Apostolopoulos gives data traffic priority to objects deemed “important” so that their images have “higher resolution than data traffic for objects with lower

importance.” Ex.1007, [0082], [0083]; Ex.1003, ¶232.

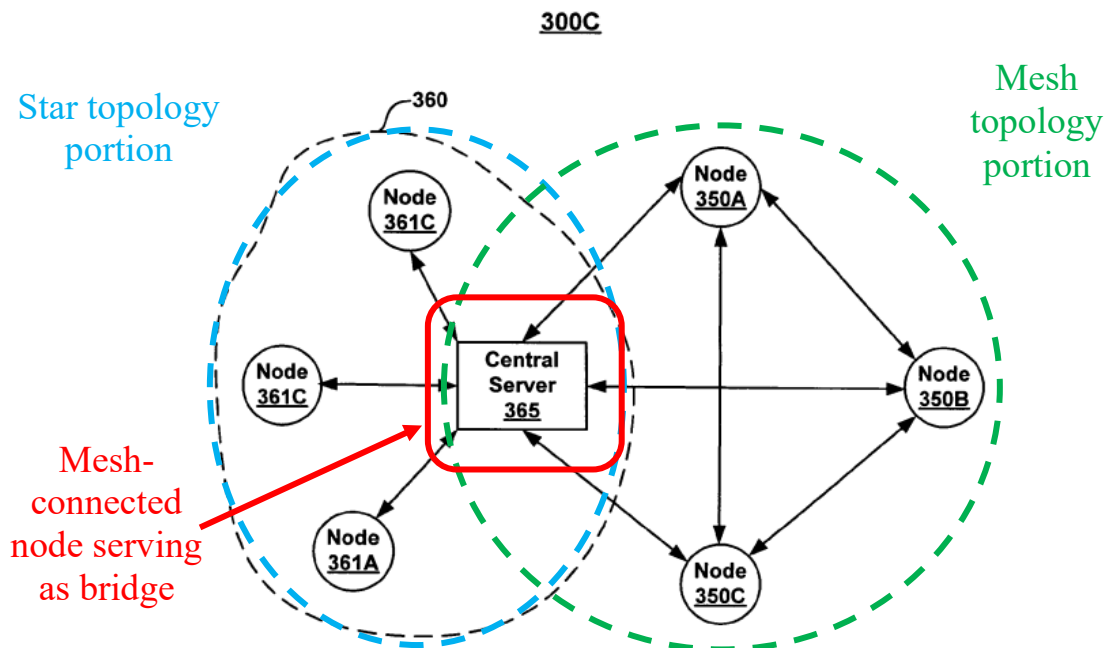
It would have been obvious to a POSITA to present Apostolopoulos’ compressed video stream in the form of Baker’s strip of miniaturized remote participant viewing areas at the receiving nodes to address bandwidth and processing constraints while prioritizing important objects/speakers in the conference. Ex.1003, ¶233; Section X.D.2.

Thus, the combination of Apostolopoulos and Baker renders this limitation obvious. Ex.1003, ¶¶227, 234.

[3.2] a proxy at the selected mesh-connected endpoint for the non-mesh connected endpoint and connected to each of the mesh-connected endpoints to receive video feeds from the mesh-connected endpoints for the non-mesh connected endpoint;

It would have been obvious to a POSITA to incorporate a *proxy* at the mesh-connected node implementing the central server/node 365 (*at the selected mesh-connected endpoint; see supra* [3.1.1]) to perform server functionality in the hybrid topology 300C. Ex.1003, ¶¶236, 242. The ’099 patent describes a proxy as performing “local Mixer and Repeater functions” and being “similar to an embedded MCU.” Ex.1001, 16:36-46. An MCU “**consolidates** all the video feeds from the end points and presents appropriate video images to each of the end points of all the video camera and collaboration material video feeds in the conference.” Ex.1001, 1:34-39; Ex.1003, ¶¶240-241.

Apostolopoulos' server-related functionalities include “receiv[ing]...communication traffic between the plurality of nodes associated with the plurality of collaborators that are participating in the N-way collaborative environment.” Ex.1007, [0057]. The central node 365 bridges the star topology portion of the hybrid network to the full-mesh topology portion and, as shown in Fig. 3C, is *connected to each of the mesh-connected endpoints*. Ex.1003, ¶237.



Ex.1007, Fig. 3C (annotated)

Further, the central node/server 365 has additional functionality compared to the other nodes 350. For instance, the central node “provides the processing power for generating communication data streams for the collaborators at nodes 361A-C to communicate amongst themselves, as well as with the collaborators at nodes 350A-C.” Ex.1007, [0070]; Ex.1003, ¶237. Similarly, Apostolopoulos explains that

a central node/server in a star communication topology “receiv[es] and send[s] communication traffic between the plurality of nodes, and N connections coupling the plurality of nodes to the central server 330.” Ex.1007, [0056]; Ex.1003, ¶214.

Given the central node 365’s location in the network and additional functionalities in the full-mesh portion of the network, a POSITA would have found it obvious to install and operate, at the mesh-connected node implementing the central node/server 365, software that incorporates a *proxy* acting as an intermediary to receive and send data traffic between the full-mesh nodes 350 and the star node 361A (*...for the non-mesh connected endpoint*). Ex.1003, ¶237.

Further, Apostolopoulos explains that “within a star topology, **some computation may be changed from an end-node to a central node**...without changing the star topology[.]” Ex.1007, [0051]. Ex.1003, ¶¶238-239. Thus, Apostolopoulos contemplates the central node 365 performing computational tasks on behalf of other nodes in the network, and acting as an intermediary between the full-mesh and star portions of the hybrid network. Ex.1003, ¶239.

To the extent Apostolopoulos does not disclose it explicitly or implicitly, it would have at least been obvious to a POSITA for the functionalities of the mesh-connected node implementing the central node/server 365 to include receiving video streams from the full-mesh-connected collaborators (*proxy...receive[s] video feeds from the mesh-connected endpoints*). Ex.1003, ¶243.

Finally, it would have been obvious for *proxy* at the mesh-connected node to receive, mix, and repeat the video streams from the full mesh nodes for the star nodes 361 (*proxy...receive[s] video feeds from the mesh-connected endpoints for the non-mesh connected endpoints*). Ex.1003, ¶243. Apostolopoulos' central node/server 365 streams video and audio data traffic between nodes 350A-C (*mesh-connected endpoints*) and nodes 361A-C (*non-mesh connected endpoints*):

In this case, the central server/node 365 provides the processing power for generating communication data streams for the collaborators at nodes 361A-C to communicate amongst themselves, as well as with the collaborators at nodes 350A-C.

Ex.1007, [0070]; Ex.1003, ¶243.

It would have been obvious to a POSITA for the *proxy* software of the mesh-connected node implementing the central node/server 365 to receive and mix the video streams from the full-mesh nodes 350 for transmission (e.g., repeating) to the star nodes (*receive[s] video feeds from the mesh-connected endpoints for the non-mesh connected endpoints*). Ex.1003, ¶¶243-244.

Thus, Apostolopoulos renders obvious this limitation. Ex.1003, ¶¶235, 244.

[3.3] a server associated with the proxy for serving the film strip with frames containing reduced-size video feeds from the mesh-connected conference participants to the non-mesh connected endpoint; and

As explained above in [3.2], the *proxy* implements functionalities at the central node for receiving and “[m]ixing” video data traffic for transmission to the other nodes. Further, as explained in Ground 1, [3.1.2], Baker’s “strip of miniaturized remote participant viewing areas” is the claimed *film strip with frames containing reduced-size video feeds from the...conference participants*. Ex.1003, ¶¶246-247. It would have been obvious for Apostolopoulos’ “mixing” to include reducing the size of videos destined for the *film strip*, since it was generally known in the art to reduce the size of images placed into a “*film strip*”-style display of remote conference participants. See, e.g., Ex.1018, Fig. 10; Ex.1017, [0060], Figs. 9, 17, 18; Ex.1003, ¶¶248-249.

The central server/node 365 “provides the processing power for generating communication data streams for the collaborators at nodes 361A-C to communicate...with the collaborators at nodes 350A-C.” Ex.1007, [0070]. Further, “[t]he central node is capable of **streaming** communication traffic to a group of **other collaborators** that communicate amongst themselves.” Ex.1007, [0070]. Ex.1007, [0070]; Ex.1003, ¶¶250, 252. A POSITA would have understood that the portion of the central server/node 365 (e.g., software module, communication port)

that performs the streaming or transmitting of the video streams to the other node includes a *server*. Apostolopoulos refers to the central node 365 as a “central server,” indicating that the central node/server 365 has server capabilities. Thus, the central node/server 365 includes a *server*. Ex.1003, ¶¶251-253.

Additionally, or alternatively, Apostolopoulos renders obvious that the central node 365, implemented in a full mesh-connected node, includes a *server*. Along with the intermediary proxy functions (receiving and mixing) discussed above at [3.2], it would have been obvious for the central node 365 to also incorporate server software (*server*) to **transmit** the resulting video stream(s) from nodes 350 to nodes 361A-C (*serv[e]...video feeds from the mesh-connected conference participants to the non-mesh connected endpoint*). It was common and well-understood to facilitate server functionalities by server software installed on a computer (*endpoint*). Further, implementing the *server* as software on the central node is consistent with the '099 patent's description of a *server*, which explains that the Star Server 91 “could be a program running on **a computer within one of the mesh endpoints** e.g. 11 in Fig. 1.” Ex.1001, 17:3-5; Ex.1003, ¶254.

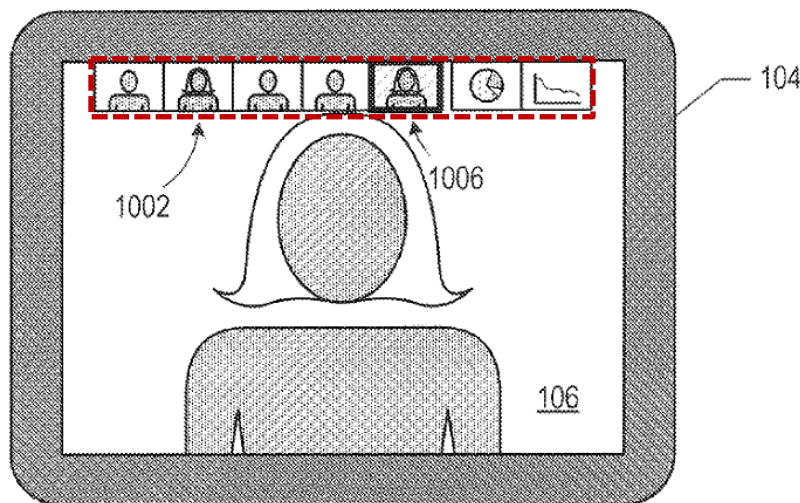
Thus, a POSITA would have found it obvious to run a software server (*server*) on the mesh-connected node implementing the central node/server 365, along with the proxy (*associated with the proxy*) to transmit streams to the nodes 361A-C (*for serving the film strip with frames containing reduced-size video feeds*

from the mesh-connected conference participants to the non-mesh connected endpoint). Ex.1003, ¶255.

Thus, the combination of Apostolopoulos and Baker render this limitation obvious. Ex.1003, ¶¶245, 256.

[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.

As discussed in Ground 1, [3.4], Baker discloses that a local participant can “touch or tap” a conference participant’s reduced-sized video feed on the strip (*select a video feed on the film strip*) to bring up that conference participant’s viewing area on the display (*for promotion to full size*). In Baker’s Fig. 10B, the strip miniaturization of participant 5 remains in the strip even when participant 5 is selected and moved to the center of display 106. Ex.1003, ¶¶259, 261.



Ex.1006, Fig. 10B

Further, Baker explains that the local participant viewing the conference can “move the display 106 to view participant 5, as shown in FIG. 10.” Ex.1006, 8:52-56. Baker contemplates transmitting video frames of speaking participants with a higher bit-rate/quality than the non-speaking participants. Ex.1006, 9:9-26, 45-49. Ex.1003, ¶¶259, 262. Baker describes users making touch- or tap-based commands via the PED to control how videos are received and displayed. Ex.1006, 8:32-67.

It would have been obvious for the full-mesh node implementing the central server 365 to include a *control connection* that permits participants at the star node 361A (*at the non-mesh connected endpoint*) to touch or tap on images in the miniaturized strip to bring them to the main display (*...to select a video feed on the film strip for promotion to full size*). Ex.1003, ¶260. Sending a command from user’s touchscreen to control the size and delivery of video streams in a videoconference was generally known in the art. Ex.1020, [0037], [0039]; Ex.1003, ¶¶259-260. As explained in Ground I, [3.4], it would have been obvious to incorporate software objects (*control connection*) at the central node 365 and the star node 361A to allow for the exchange of command messages to control the display based on Baker’s touch and tap interface controls. Ex.1001, 16:43-46; Ex.1003, ¶260. Tapping or touching an image in the miniaturized film strip would cause the central node 365, which implements the *proxy* and *server* functionalities, to bring up the full size video feed for the selected participant on the main display

(*initiate transfer of the full size video feed*). Ex.1003, ¶¶262, 265.

Baker's reduced-quality miniaturized strip of non-speaking participants in Apostolopoulos' hybrid network would have been useful to a POSITA because of the different network constraints and requirements imposed on the different kinds of nodes. Ex.1007, [0048]; Ex.1003, ¶263. It would have been obvious to provide Baker's strip display and selection UI to Apostolopoulos' nodes 361A-C (*non-mesh connected endpoints*), while keeping the video processing and serving functionalities at the central node 365. Ex.1003, ¶¶258, 263. Apostolopoulos explains that star nodes 361A-C "may comprise a group of low-power cellphone clients" that lack the processing power of nodes 350A-C. Ex.1007, [0070]. Further, Apostolopoulos contemplates providing higher resolution images for objects or people with higher importance. Ex.1007, [0082], [0083]; Ex.1003, ¶¶263-265.

Implementing a *control connection* between star node 361A and the central node using Baker's strip UI to *select a video feed on the film strip for promotion to full size and initiate transfer of the full size video feed* would allow a participant with lower computational power or network resources to select a higher priority video feed to promote to full-size (e.g., active speaker) without unnecessary visual complexity at the display. Ex.1007, [0070], [0082]; Ex.1003, ¶¶263, 264.

Thus, the combination of Apostolopoulos and Baker render this limitation obvious. Ex.1003, ¶¶257, 266.

4. Claim 10

[10.0], [10.1], [10.2]

Limitation [10.0] is substantively identical to limitations [3.0], [3.1.1], and [3.1.2], where the claimed *endpoint for use as a mesh-connected endpoint in a teleconferencing system* is Apostolopoulos' central node/server 365. This limitation is rendered obvious for the same reasons explained above in [3.0], [3.1.1], and [3.1.2]; Ex.1003, ¶267.

Limitation [10.1] is substantively identical to limitation [3.2] and is rendered obvious for the same reasons. Ex.1003, ¶268.

Limitation [10.2] is substantively identical to limitation [3.3] and is rendered obvious for the same reasons. Ex.1003, ¶269.

[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.

As explained in [3.4], it would have been obvious to a POSITA for the central node 365 to include Baker's film strip display and "touch or tap" functionality to allow participants at star nodes 361 to control, through a *control connection*, how video streams are displayed, including triggering a full size video feed to be displayed on the local participant's device. Ex.1003, ¶¶271-272. Thus, if the claimed *proxy controller* is interpreted to include the *control connection* of

[3.4] (e.g., software handles), this limitation is obvious for the same reasons explained above with respect to the *control connection*. Ex.1003, ¶272.

Alternatively, if the *proxy controller* is interpreted as distinct from the *control connection*, Apostolopoulos and Baker render this limitation obvious. Ex.1003, ¶272. For instance, it would have been obvious to a POSITA for the central node 365's *proxy* (e.g., communication traffic “**receiv**[ing] and send[ing]” functionality) to include a software submodule (*proxy controller*) to receive and respond to the touch or tap commands from the other connected nodes (e.g., star node 361A). Ex.1007, [0057]. Ex.1003, ¶273. With the central node 365's *proxy* already handling the video stream receiving functions, it would have been obvious for the *proxy* to also assume the responsibility for handling the touch or tap commands there at the central node 365, which could respond to such commands by triggering changes to the film strip to be transmitted. Ex.1003, ¶273.

When the conference participant at star node 361A selects a viewing area to show up on the main display area (*select[s] a video feed on the film strip for promotion to full size*), the central node 365's proxy controller would respond by causing the full size video feed for the selected feed to be displayed at the star node 361A (*initiate transfer of the full size video feed to the non-mesh connected endpoint*). Ex.1003, ¶274. As explained in [3.4], a POSITA would have appreciated the advantages of Baker's reduced quality miniaturization of non-

speaking participants in Apostolopoulos' hybrid topology because Apostolopoulos' nodes 361A-C "may comprise a group of low-power cellphone clients" that lack the processing power of nodes 350A-C. Ex.1007, [0070]; Ex.1003, ¶274. Further, incorporating Baker's display control features into Apostolopoulos' nodes would prioritize important objects/speakers in the conference and allow participants to modify the display to meet their needs. Ex.1003, ¶¶275-276.

Thus, the combination of Apostolopoulos and Baker render this limitation obvious. Ex.1003, ¶¶270, 277.

F. Ground 4

1. Reasons to Combine Apostolopoulos, Baker, and Wong.

A POSITA would have been motivated to combine Apostolopoulos and Baker for the reasons explained above. Section X.E.2. Further, a POSITA would have been motivated to combine Wong with Apostolopoulos and Baker, and doing so would be combining prior art elements according to known methods to yield predictable results. Ex.1003, ¶¶278-281.

The incorporated feature would be Wong's web browser for accessing and controlling video share casting services. Ex.1009, [0013], [0041], [0100]; Ex.1003, ¶278. As explained in Ground 2, combining Wong with Apostolopoulos and Baker would be achieved through well-known and routine computer programming techniques, and the results of the combination would have been predictable.

Ex.1003, ¶¶278, 281, 150; Section X.D.2.

Combining Wong’s web browser-based video share casting techniques with Apostolopoulos and Baker would be applying a known technique to a known device ready for improvement to yield predictable results. Ex.1003, ¶282.

Apostolopoulos discloses techniques for mesh and star-based networks for video conferencing and streaming video to conference participants, and Baker further describes how the aggregated video streams can be displayed on the same display (e.g., reduced-size film strip). Ex.1006, Figs. 10A, 10B, 8:40-62; Ex.1003, ¶282. Apostolopoulos and Baker provide limited details on the type of software or application platform used to power the display features.

Wong describes using a web browser to participate in video teleconference. Ex.1009, [0013], [0106], [0107]; Ex.1003, ¶282. Wong even describes using a browser for similar display techniques as Baker, such as displaying “thumbnails of live video casts”, and that a “user can click on a thumbnail to view that particular blog or cast.” Ex.1009, [0041]; Ex.1003, ¶282; Section X.D.2.

The result of the proposed combination would have been predictable and a POSITA would have had a reasonable expectation of success, because the combination of Wong’s web browser-based video share casting technique with Apostolopoulos’ communication network and Baker’s film strip would have been well-within the skill level of a POSITA. Ex.1003, ¶¶238-284. Apostolopoulos’

nodes already transmit video data streams, but Apostolopoulos and Baker provide limited detail on the types of application platforms that could be used to view and control the conference, indicating that a POSITA would have had the requisite background knowledge. Ex.1007, [0033], [0025]; Ex.1003, ¶284. Baker describes storing “software applications including...browsers.” Ex.1006, 3:29-34. Modifying Apostolopoulos’ video data stream transmission, and Baker’s film strip, to be served to a web browser, therefore, would have been a straight-forward exercise of web application development per Wong’s well-known techniques. Ex.1009, [0013], [0041], [0100]; Ex.1003, ¶284. This is another reason a POSITA would have been motivated to incorporate Wong’s web browser features in the system resulting from the combination of Apostolopoulos and Baker. Ex.1003, ¶284. The combination would have resulted in a greater number of people able to use the teleconferencing service, since web browsers were particularly common. *See, e.g.*, Ex.1015, Ex.1016 (describing Google Hangouts’ and Facebook Skype’s browser-based video call services); Ex.1003, ¶284.

2. Claim 4

***[4.0] A teleconferencing system as claimed in claim 3,
wherein the server is configured to serve the filmstrip to a
browser co-located with the non-mesh connected endpoint.***

It would have been obvious to a POSITA for the server software (*server*) at Apostolopoulos’ central node 365 to *serve* Baker’s filmstrip to a *browser* on

Apostolopoulos' nodes 361A-C (*non-mesh connected endpoint*). Ex.1003, ¶¶286-287. Apostolopoulos' nodes may be lower-power devices such as cell-phones, laptops, or PDAs. Ex.1007, [0063], [0070]. Such, low-power, low-complexity devices “only require[] **a simple video decoder** to view the synthetic N-way collaborative environment.” Ex.1007, [0063]; Ex.1003, ¶¶288-289.

As explained in Ground 2, [4.0], browsers, such as the web browser disclosed by Wong, were one of the principal mechanisms for users to access the internet and internet-based applications. Web browsers were versatile software platforms used to facilitate a host of internet-based applications, including real time communications. Ex.1013, 1:47-60; Ex.1014, 2:21-29; Ex.1003, ¶290. Video call services such as Facebook's Messenger Skype service and Google Hangouts were accessed through web browsers. *See, e.g.*, Ex.1015; Ex.1016; Ex.1003, ¶290.

It would have been obvious for each of the nodes 361A-C to have installed Wong's web browser (Ex.1009, [0013], [0041], [0106]) on local memory (*co-located with the non-mesh connected endpoint*), and for the central node's server software to serve the filmstrip to the star node 361A's browser. Ex.1003, ¶291; Section X.F.1.

Thus, the combination of Apostolopoulos, Baker, and Wong renders this limitation obvious. Ex.1003, ¶¶285, 292.

3. Claim 5

[5.0] A teleconferencing system as claimed in claim 4, wherein the browser is displayed on a computer co-located with the non-mesh connected endpoint.

As explained in [3.1.1], Apostolopoulos' nodes include laptops, PDAs, or cellphones. Ex.1007, [0060], [0063], [0070] (cellphones). Ex.1003, ¶295. The '099 patent explains that an endpoint 84 can be both "co-located" with a computer and "run on that same computer." Ex.1001, 16:10-12, 20-21, Fig. 9; Ex.1003, ¶294.

As explained in [4.0], it would have been obvious for the star node 361A (*non-mesh connected endpoint*) to use Wong's *browser* to display Baker's strip of miniaturized images. It was well-known that browsers provided a graphical user interface to view and interact with the content. Thus, a POSITA would have understood that one of Apostolopoulos' star-connected nodes 361 running the *browser* would include the browser being *displayed on a computer co-located with* star node 361A. Ex.1003, ¶296.

Thus, the combination of Apostolopoulos, Baker, and Wong renders this limitation obvious. Ex.1003, ¶¶293, 297.

4. Claim 6

[6.0] A teleconferencing system as claimed in claim 5, wherein the proxy and server are merged into a virtualized environment.

As explained in [3.2] and [3.3], it would have been obvious to a POSITA for

Apostolopoulos' central node 365 to implement the functionalities of a *proxy* and a *server* using software. Ex.1003, ¶299. Further, it was well-known to use virtualized environments such as clouds to provide computational assistance and other administrative tasks for programs accessed through a computer. Ex.1010, 5:50-52; Ex.1011, 4:40-46; Ex.1003, ¶¶300-301. A POSITA would have found it obvious for Apostolopoulos' *proxy* software and server software, both running on the central node 365 (*server*), to be supported by, or communicate with, a cloud (*merged into a virtualized environment*). Ex.1003, ¶¶300, 302.

For instance, it would have been obvious for the *proxy* and *server* to perform receive, aggregate, and distribute aggregated streams based on teleconference participant information managed by a cloud server in communication with the peers. A POSITA would have appreciated that merging the *proxy* and *server* into a virtualized environment in this way would reduce processing demands at the nodes and free up computational resources of the central node 365 to meet video streaming quality demands. Ex.1003, ¶302.

Thus, the combination of Apostolopoulos and Baker render this limitation obvious. Ex.1003, ¶¶298, 303.

XI. DISCRETIONARY DENIAL IS NOT APPROPRIATE

A. No basis for 35 U.S.C. §325(d) denial

Setton, Baker, Apostolopoulos, and Wong were not cited or considered by the Examiner during prosecution of the '099 patent. Discretionary denial under 35 U.S.C. §325(d) is inappropriate at least because the first part of the *Advanced Bionics* framework is not met. *Advanced Bionics, LLC v. MED-EL Elektromedizinische Geräte GmbH*, IPR2019-01469, Paper 6 at 8 (PTAB Feb. 13, 2020) (precedential).

B. No basis for 35 U.S.C. §314/*Fintiv* denial

The six factors for considering denial under §314 strongly favor institution. *Apple Inc. v. Fintiv, Inc.*, IPR2020-00019, Paper 11 (Mar. 20, 2020) (precedential). The district court case is in its early stages. Patent Owner filed its complaint on May 19, 2023 (Ex.1021) and served it on May 22, 2023. Other than a pre-answer partial motion to dismiss, little has happened with regard to invalidity. And Petitioner has worked diligently since service of the complaint to prepare this Petition, which is being filed well within the one-year timeframe allowed by Congress.

1. Whether evidence exists that a stay may be granted if a proceeding is instituted

A motion to stay has not yet been filed and the Board should not infer or attempt to predict the outcome if such a motion is later filed. *Sand Revolution II LLC v. Continental Intermodal Group – Trucking LLC*, IPR2019-01393, Paper 24 at 7 (June 16, 2020) (informative); *see also Dish Network L.L.C. v. Broadband iTV, Inc.*, IPR2020-01359, Paper 15 (Feb. 12, 2021) (“It would be improper to speculate...what the Texas court might do regarding a motion to stay...”). Thus, this factor is neutral.

2. Proximity of the court’s trial date to the board’s projected statutory deadline for a final written decision

The operative Docket Control Order sets August 13, 2024, for the *Markman* hearing and February 10, 2025, for Jury Selection. Ex.1022, 1, 3. The Final Written Decision is expected in this case in May 2025. This factor nevertheless is neutral, because the Petitioner has worked expeditiously to prepare and file this petition within just six months of being served with Patent Owner’s complaint on May 22, 2023. *Fintiv*, Paper 11 at 11-12.

Furthermore, while trial is currently scheduled to begin in February 2025, “scheduled trial dates are often subject to change.” *In re Apple Inc.*, 979 F.3d 1332, 1344, n.5 (Fed. Cir. 2020). Petitioner notes that sixteen other cases are scheduled for Jury Selection on February 10, 2025—the same Jury Selection date as the co-

pending litigation—and have not settled. Ex.1023 (listing all cases having Jury Selection scheduled for February 10, 2025).

As such, this factor is neutral. *See Sand Revolution II, LLC v. Continental Intermodal Grp.—Trucking LLC*, IPR2020-01393, Paper 24 (June 16, 2020) (panel granting institution on rehearing based on finding that uncertainty over district court’s trial date weighed against discretionary denial).

3. Investment in the parallel proceeding

The co-pending litigation is in its early stages, and the investment has been minimal. Fact discovery is not set to close until September 23, 2024, and expert discovery is not set to close until October 28, 2024. Ex.1022, 3. The Docket Control Order sets August 13, 2024, as the date for the *Markman* Hearing. Ex.1022, 3. Thus, this factor weights against discretionary denial. *See PEAG LLC v. Varta Microbattery GMBH*, IPR2020-01214, Paper 8 at 17 (Jan. 6, 2021) (little investment in the parallel proceeding weighed against discretionary denial).

4. Overlap between issues raised in the petition and in the parallel proceeding

There will be no overlap of prior art issues. Petitioner stipulates that, if the Board institutes trial, Petitioner will cease asserting in the co-pending litigation the combination of references on which trial is instituted for the claims, to the extent Petitioner even asserts the same combination in the district court. This factor thus

weighs against discretionary denial. *See Verizon v. Huawei*, IPR2020-01079, Paper 10 at 38 (Jan. 14, 2021) (finding a similar stipulation weighs “against discretionary denial of the Petition.”).

5. Whether the petitioner is a defendant

Petitioner is a defendant in the litigation. Ex.1021. This is true of most Petitioners in IPR proceedings. While the *Fintiv* case indicates that a difference between the district court defendant and the petitioner may weigh against discretionary denial, nothing within the *Fintiv* case suggests that the same party between the proceedings weighs in favor of discretionary denial. *See HP Inc. v. Slingshot Printing LLC*, IPR2020-01084, Paper 13 at 9. Accordingly, this factor is neutral and should not be a basis for denying institution.

6. Other circumstances that impact the Board’s exercise of discretion, including the merits

As discussed above, the prior art presented in this Petition renders the Challenged Claims unpatentable as obvious. In view of the striking similarities between the ’099 patent claims and Setton, Baker, Apostolopoulos, and Wong, this petition “presents compelling evidence of unpatentability.” Ex.1024, 2. This factor weighs against discretionary denial.

As such, the *Fintiv* factors are either neutral or weigh against discretionary denial. Because this Petition was filed more than six months before the statutory bar date, institution should not be denied on discretionary factors.

C. Discretionary denial under *General Plastic* is not appropriate

The '099 patent has not been challenged in any prior IPR petition, so no *General Plastic* discretionary institution factors apply here. *See General Plastic Indus. Co., Ltd. v. Canon Kabushiki Kaisha*, IPR2016-01357, Paper 19 at 16 (Sept. 6, 2016) (Section II.B.4.i. precedential).

XII. CONCLUSION

Petitioner has established a reasonable likelihood that the Challenged Claims are unpatentable.

Respectfully submitted,

Dated: November 21, 2023
HAYNES AND BOONE, LLP
2323 Victory Avenue, Suite 700
Dallas, Texas 75201
Customer No. 27683

/Theodore M. Foster/
Theodore M. Foster
Lead Counsel for Petitioner
Registration No. 57,456

XIII. MANDATORY NOTICES

Real party-in-interest

Pursuant to 37 C.F.R. §42.8(b)(1), Petitioner certifies that the real party-in-interest is Cisco Systems, Inc.

Related matters

Pursuant to 37 C.F.R. §42.8(b)(2), to the best knowledge of the Petitioner, the '099 patent is or was involved in the following cases:

Case Heading	Number	Court	Filed
<i>Video Solutions Pte. Ltd. V. Cisco Systems, Inc.</i>	2-23-cv-00222	E.D. Tex.	May 19, 2023

Lead and back-up counsel and service information

Lead Counsel

Theodore M. Foster
HAYNES AND BOONE, LLP
2801 N. Harwood Street, Suite 2300
Dallas, TX 75201

Phone: (303) 382-6205
Fax: (214) 200-0853
ipr.theo.foster@haynesboone.com
USPTO Reg. No. 57,456

Back-up Counsel

David L. McCombs
HAYNES AND BOONE, LLP
2801 N. Harwood Street, Suite 2300
Dallas, TX 75201

Phone: (972) 739-8636
Fax: (214) 200-0853
david.mccombs.ipr@haynesboone.com
USPTO Reg. No. 32,271

Eugene Goryunov
HAYNES AND BOONE, LLP
2801 N. Harwood Street, Suite 2300
Dallas, TX 75201

Phone: (312) 216-1630
Fax: (214) 200-0853
eugene.goryunov.ipr@haynesboone.com
USPTO Reg. No. 61,579

Brett Bostrom
HAYNES AND BOONE, LLP
2801 N. Harwood Street, Suite 2300
Dallas, TX 75201

Phone: (972) 739-6973
Fax: (214) 200-0853
brett.bostrom.ipr@haynesboone.com
USPTO Reg. No. 73,009

Please address all correspondence in this proceeding to lead and back-up counsel. Petitioner consents to service in this proceeding by email at the addresses above.

CLAIMS APPENDIX

- [3.0] A teleconferencing system comprising:
- [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;
- [3.2] a proxy at the selected mesh-connected endpoint for the non-mesh connected endpoint and connected to each of the mesh-connected endpoints to receive video feeds from the mesh-connected endpoints for the non-mesh connected endpoint;
- [3.3] a server associated with the proxy for serving the film strip with frames containing reduced-size video feeds from the mesh-connected conference participants to the non-mesh connected endpoint; and
- [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.
- [4.0] A teleconferencing system as claimed in claim 3, wherein the server is configured to serve the filmstrip to a browser co-located with the non-mesh connected endpoint.
- [5.0] A teleconferencing system as claimed in claim 4, wherein the browser is displayed on a computer co-located with the non-mesh connected endpoint.
- [6.0] A teleconferencing system as claimed in claim 5, wherein the proxy and server are merged into a virtualized environment.
- [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:

[10.1] a proxy for connection to each of the mesh-connected endpoints to receive video feeds from the mesh-connected endpoints for the non-mesh connected endpoint;

[10.2] a server associated with the proxy for serving the film strip with frames containing reduced-size video feeds from the mesh-connected conference participants to the non-mesh connected endpoint; and

[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.

CERTIFICATE OF WORD COUNT

Pursuant to 37 C.F.R. §42.24(d), Petitioner hereby certifies, in accordance with and reliance on the word count provided by the word-processing system used to prepare this Petition, that the number of words in this paper is 13,907. Pursuant to 37 C.F.R. §42.24(d), this word count excludes the table of contents, table of authorities, mandatory notices under §42.8, certificate of service, certificate of word count, appendix of exhibits, and any claim listing.

Dated: November 21, 2023

/Theodore M. Foster/

Theodore M. Foster
Lead Counsel for Petitioner
Registration No. 57,456

CERTIFICATE OF SERVICE

The undersigned certifies that, in accordance with 37 C.F.R. §42.6(e) and 37 C.F.R. §42.105, service was made on Patent Owner as detailed below.

Date of service November 21, 2023

Manner of service FEDERAL EXPRESS

Documents served Petition for *Inter Partes* Review Under 35 U.S.C. §312 and 37 C.F.R. §42.104 of U.S. 9,204,099;
Petitioner's Power of Attorney; Petitioner's Exhibit List;
Exhibits Ex.1001-Ex.1025

Persons served Marks & Clerk
180 Kent Street, Suite 1900
Ottawa, ON K1P 0B6
Canada

/Theodore M. Foster/
Theodore M. Foster
Lead Counsel for Petitioner
Registration No. 57,456