

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

CISCO SYSTEMS, INC.,
Petitioner

v.

VIDEO SOLUTIONS, PTE. LTD.,
Patent Owner.

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

**SUPPLEMENTAL DECLARATION OF DR. NARASIMHA REDDY
IN SUPPORT OF PETITIONER'S REPLY**

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	SUPPLEMENTAL OPINIONS	4
A.	Claim Construction.....	4
1.	“conference participants” ([3.1.2] and [10.0])	4
B.	Ground I – Setton and Baker.....	6
1.	Limitations [3.4] and [10.3].	6
C.	Ground III – Apostolopoulos and Baker	9
1.	Limitations [3.1.1] and [10.0]	9
III.	CONCLUSION.....	11

I, Narasimha Reddy, Ph.D., do hereby declare as follows:

I. INTRODUCTION

1. I am making this Supplemental Declaration at the request of Cisco Systems, Inc. (“Petitioner”) in support of Petitioner’s Reply in the matter of the *Inter Partes* Review of U.S. Patent No. 9,204,099 (“the ’099 patent”).

2. I am being compensated for my work in this matter at my standard hourly rate. I am also being reimbursed for reasonable and customary expenses associated with my work and testimony in this proceeding. My compensation is not contingent on the outcome of this matter or the specifics of my testimony.

3. I have been asked to provide my supplemental opinions responding to certain arguments Patent Owner and its expert have made regarding the obviousness of claims 3-6 and 10 (“the Challenged Claims”) of the ’099 patent.

4. In the preparation of this declaration, I have studied:

- a. the ’099 patent, Ex.1001;
- b. the prosecution history of the ’099 patent (“’099 File History”), Ex.1002;
- c. U.S. Patent Publication No. 2010/0287238 to Setton, et al. (“Setton”), Setton;
- d. U.S. Patent No. 8,581,958 to Baker, et al. (“Baker”), Ex.1006;

e. U.S. Patent Publication No. 2005/0080894 to Apostolopoulos, et al. (“Apostolopoulos”), Ex.1007; and

f. U.S. Patent Publication No. 2009/0232129 to Wong, et al. (“Wong”), Ex.1009.

5. In forming the opinions expressed below, I have considered: the documents listed above; the relevant legal standards, including the standard for obviousness; and my own knowledge and experience based upon my work in the field of network communications and security as described below, as well as portions of the following additional materials:

a. U.S. Patent No. 7,692,683 to Kenoyer et al. (“Kenoyer”), Ex.1008;

b. U.S. Patent No. 8,601,097 to Sartori (“Sartori”), Ex.1010;

c. U.S. Patent No. 9,124,757 to Weber (“Weber”), Ex.1011;

d. Andrew S. Tanenbaum, “Computer Networks,” 4th ed. (2002) (“Tanenbaum”), Ex.1012;

e. U.S. 9,009,225 to Kjesbu et al. (“Kjesbu”), Ex.1013;

f. U.S. 7,877,489 to Salesky et al., (“Salesky”), Ex.1014;

g. Michael Arrington, “Facebook Will Launch In-Browser Video Chat Next Week In Partnership with Skype,” TechCrunch, Jul. 1, 2011 (archived July 4, 2011) (“TechCrunch”), Ex.1015;

- h.** Ian Paul, “Facebook Video Calling vs. Google+ Hangouts,” PCWorld.com, July 7, 2011 (archived August 7, 2011) (PCWorld), Ex.1016;
- i.** U.S. Pub. No. 2006/0215765 to Hwang et al. (“Hwang”), Ex.1017;
- j.** U.S. Pub. No. 2009/0015661 to King et al. (“King”), Ex.1018;
- k.** U.S. Pub. No. 2010/0157017 to Pepperell (“Pepperell”), Ex.1019;
- l.** U.S. Pub. No. 2010/0333004 to Kristiansen et al. (“Kristiansen”), Ex.1020;
- m.** Andrew S. Tanenbaum, “Computer Networks,” 4th ed. (2002), pp. 529-532, 683-692 (“Tanenbaum II”), Ex.1033;
- n.** Request for Comments (RFC) 3550 – RTP: A Transport Protocol for Real-Time Applications, July 2003, Ex.1034;
- o.** RFC 4123 – Session Initiation Protocol (SIP) – H.323 Interworking Requirements, July 2005, Ex.1035;
- p.** RFC 3261 – SIP: Session Initiation Protocol, June 2002, Ex.1036;
and
- q.** My previously provided declaration in this case (“Reddy Dec.”), Ex.1003.

6. Unless otherwise noted, all **emphasis** in any quoted material has been added. Claim terms are *italicized*.

II. SUPPLEMENTAL OPINIONS

7. I have reviewed the Patent Owner Response filed in this case (Paper No. 16) and it is my opinion that Patent Owner's claim constructions are incorrect because they attempt to add limitations from the specification and are contrary to the plain and ordinary meaning of the terms being construed. It is my opinion that these additional limitations are not supported by the claims or the specification of the '099 patent.

A. Claim Construction

1. “*conference participants*” ([3.1.2] and [10.0])

8. Claim 3 of the '099 patent recites “*a plurality of endpoints...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.*” From my review of the claim language and the specification of the '099 patent, nothing supports Patent Owner's proposed construction that “*the conference participants*” must include at least one conference participant at the “*non-mesh connected endpoint.*”

9. The claim language itself disagrees with Patent Owner's construction for at least two reasons. First, the claim language only recites that the “*mesh-connected endpoints*” provide video feeds. For instance, the “*proxy at the selected mesh-connected endpoint*” receives “*video feeds **from the mesh-connected endpoints.***” Further, the “film strip” contains “*reduced size video feeds from **the***

mesh-connected conference participants.” By contrast, the claim language suggests that the “*non-mesh connected endpoint*” only receives the video feeds, and does not provide them. Thus, the claim feature including “*reduced size video feeds from the conference participants*” is consistent with the scenario in which there are only “conference participants” at the mesh connected endpoints providing video feeds.

10. Second, claim 3 recites a “*a user*”—not a conference participant—“*located at the non-mesh connected endpoint.*” The claim recites that the “*user*” at the non-mesh connected endpoint selects a video feed to be transferred to the non-mesh connected endpoint. The claim language therefore describes the individuals at the mesh-connected endpoints providing video feeds as “*conference participants,*” and to the individual at the non-mesh connected endpoint receiving video feeds as a “*user.*”

11. A review of the specification confirms that the claimed “*user*” does not need to be a “conference participant.” For instance, the ’099 patent specification explains:

The use of a filmstrip is a feature of the user interface of a video conferencing endpoint in accordance with embodiments of the invention. The filmstrip consists of a frame strip containing reduced size video feeds from **the various participants**. Its purpose is to allow **users** to see

video and collaboration material **from all conferenced parties** and to determine which particular video streams or collaboration materials will be “promoted”, that is to say displayed at a larger scale, and how promotions will be initially placed on the screen(s) of the endpoint.

’099 patent, 4:61-5:3; (emphasis added); POR, 35, 36.

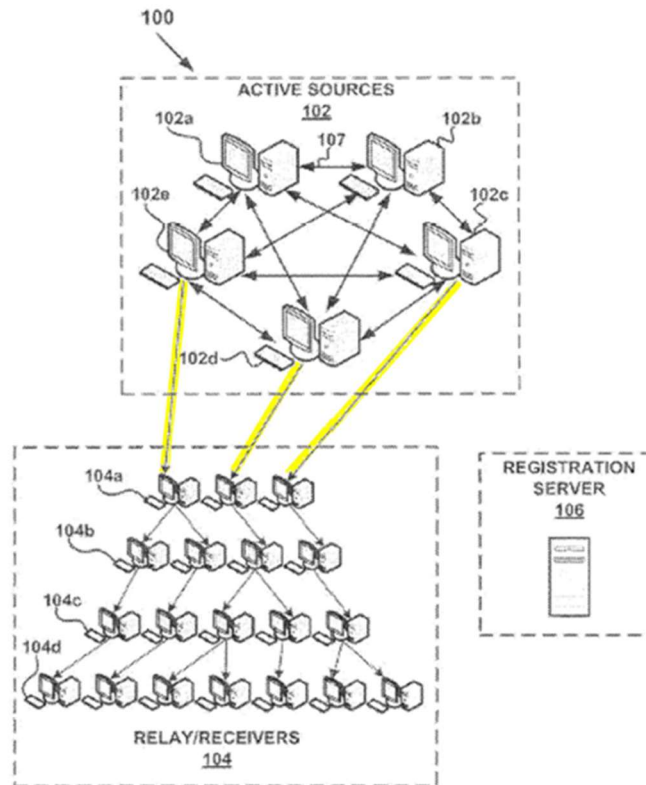
12. Thus, like the claim language, the ’099 specification describes the individuals providing video feeds as “participants,” and the individuals receiving or viewing video feeds as “users.” ’099 patent, 4:61-5:3. A POSITA would have understood this distinction between “*conference participants*” and “*users*” to be intentional. Patent Owner’s proposed construction would confuse the distinctions drawn by the claims and specification between a “*conference participant*” and a “*user*.” Accordingly, I do not believe that the claim should be construed to require the user at the non-mesh endpoint to be a “*conference participant*” or that there be a “*conference participant*” providing a video feed from a non-mesh endpoint.

B. Ground I – Setton and Baker

1. Limitations [3.4] and [10.3].

13. Patent Owner mischaracterizes Setton’s network as being wholly “unidirectional,” such that communications can only flow from the active sources to the relay/receivers. POR, 56, 12, 13, 47, 48. Fig. 1 of Setton, for instance, shows a multicast of an aggregated media stream in a tree-based network of

relay/receivers. *See* Setton, [0023], [0025]. Nothing in Setton suggests that the direction of the arrows represents an overarching limitation on the direction of all communications.



Setton, Fig. 1 (annotated); POR, 13.

14. Setton does not limit its network to any specific type of communication protocol, but it discusses a “tree-based video multicast protocol” as an example. Setton, [0018], [0023]. And Setton does not say that the protocol being cannot allow “upstream” communications from the relay/receivers to the active sources.

15. A POSITA would have understood that the vast majority of multicast communication protocols and/or media streaming protocols are not unidirectional in the way Patent Owner suggests. At the time Setton was filed, for instance, common multicasting and media streaming communication protocols included Real-Time Transport Protocol (RTP), H.323, and Session Initiation Protocol (SIP). *See, e.g.*, Tanenbaum II, 529-532, 683-692; Wong, [0040], [0129]; King, [0033]; RFC 3550 (RTP, RTCP); RFC 4123 (H.323); RFC 3261 (SIP). Each of these protocols involves bi-directional control signaling, and allows for a variety of content types to be transmitted in either direction.

16. For instance, RTP uses a Real-Time Control Protocol (RTCP) for providing feedback on the quality of the transmission. Tanenbaum II, 529-532; RFC 3550, 17-18. A device receiving a media stream provides the feedback as control signaling to the device providing the media stream. RFC 3550, 29. SIP is the signaling protocol that carries Session Description Protocol (SDP) messages, including invite and update requests from devices receiving a media stream. Wong, [0129]; Tannenbaum II, 690, 691, Fig. 7-68; RFC 3261, 9-12, 78-81. These requests are used to renegotiate media stream parameters if a receiver's conditions change. RFC 3261, 9-12. The invite and update requests also include control signaling transmitted from the devices receiving (or requesting to receive) a media

stream. *See also*, RFC 4123, 9, 10 (describing state messages for H.323, including SIP INVITE).

17. It would be highly unusual and impractical for a video conference system like Setton's to use a communication protocol that does not even support bidirectional control signaling, for example. If Setton's network did not support bidirectional control signaling, there would be no suitable mechanism for error and quality reporting, flow control, or dynamic group management. Without some explicit explanation in Setton for using a truly unidirectional communication protocol, a POSITA would not have understood Setton's network to be so constrained.

C. Ground III – Apostolopoulos and Baker

1. Limitations [3.1.1] and [10.0]

18. I disagree with Patent Owner's conclusion that combining central server capabilities with Apostolopoulos's collaborator nodes would cut against Apostolopoulos's goal of having flexible, dynamic networks. Combining central server functionality with one or more collaborator nodes would only enhance the overall flexibility and resilience of the network. Redundancy and fallbacks are common and beneficial features of dynamic networks. A POSITA would have appreciated that having multiple options of nodes capable of performing the role of

central server would allow conference sessions to more quickly recover and continue in the event of server or network failures.

19. Further, combining central server and collaborator node functionalities would provide additional options for collaboration and network configuration. Pet., 57, 58. One available approach would be to install an MCU, like the one described in Kenoyer, at one or more of Apostolopoulos's nodes. Kenoyer, 3:63-65. These nodes would therefore allow for a user to access a video conference, while also being able to take on the role of the server (e.g., encoding, generating composite video data, distributing) based on network conditions and the parameters of the video conference. *See* Kenoyer, 5:4-39 (MCU performing video encoding/decoding, composite video generation, scaling, etc.).

20. Finally, I also disagree with Patent Owner's argument that a POSITA would not have used Baker's graphical interface features with Apostolopoulos's network. POR, 73. I see nothing in Apostolopoulos that would limit its network architecture to virtual 3D environments. Apostolopoulos explains that its networks and methods can "reduc[e] both computational complexity and communication bandwidth traffic." Apostolopoulos, [0019]. Some of Apostolopoulos's embodiments include a central server with reduced "complexity requirements." Apostolopoulos, [0065]. Thus, a POSITA would not have understood Apostolopoulos to be limited to the high-complexity use cases of 3D virtual

environments. Rather, a POSITA would have understood that Apostolopoulos's networks would have been suitable for 2D video conferencing and collaboration environments, including an environment that uses Baker's film strip interface.

III. CONCLUSION

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

A. C. Narasimha Reddy

Dated: October 10, 2024

Narasimha Reddy, Ph.D.