

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

GOOGLE LLC,
Petitioner,

v.

VALTRUS INNOVATIONS LIMITED,
Patent Owner.

IPR2022-01418
Patent 7,523,454 B2

Before NORMAN H. BEAMER, SHARON FENICK, and
PAUL J. KORNICZKY, *Administrative Patent Judges*.

KORNICZKY, *Administrative Patent Judge*.

DECISION
Granting Institution of *Inter Partes* Review
35 U.S.C. § 314

I. INTRODUCTION

Google LLC (“Petitioner”) filed a Petition requesting an *inter partes* review of claims 1–21 of U.S. Patent No. 7,523,454 B2 (Ex. 1001, “the ’454 patent”). Paper 2 (“Pet.”). Valtrus Innovations Limited (“Patent Owner”) filed a Preliminary Response opposing institution. Paper 6 (“Prelim. Resp.”). With the Board’s permission, Petitioner filed a Reply (Paper 7, “Reply”) and Patent Owner filed a Sur-Reply (Paper 8, “Sur-Reply”).

We have authority to institute an *inter partes* review if “the information presented in the petition . . . and any response . . . shows that there is a reasonable likelihood that the petitioner would prevail with respect to at least 1 of the claims challenged in the petition.” 35 U.S.C. § 314(a) (2018).

After considering the Petition, Preliminary Response, and other evidence of record, we determine that Petitioner has demonstrated a reasonable likelihood of showing the unpatentability of at least one of the challenged claims. Thus, we institute an *inter partes* review on all challenged claims on all asserted grounds. 37 C.F.R. § 42.108(a). Specifically, we grant Petitioner’s request to institute an *inter partes* review of claims 1–21 of the ’454 patent.

II. BACKGROUND

A. *Real Parties-in-Interest*

As required by 37 C.F.R. § 42.8(b)(1), each party identifies the real party-in-interest. Petitioner identifies itself as a real party-in-interest. Pet. x; *see id.* at n.1 (noting that Petitioner is a subsidiary of XXVI Holdings Inc., which is itself a subsidiary of Alphabet Inc., but asserting that neither of

these are real parties-in-interest). Patent Owner states that itself, Key Patent Innovations Limited, and Valinvest LLC “are financially interested in the outcome of the litigation between Valtrus and Google.” Paper 4, 1.

B. Related Proceedings

As required by 37 C.F.R. § 42.8(b)(2), Petitioner and Patent Owner identify the judicial or administrative matters that would affect or be affected by a decision in this proceeding. Petitioner and Patent Owner state the ’454 patent is the subject of (1) *Valtrus Innovations, Ltd. v. Google LLC*, No. 3:22-cv-00066-N (N.D. Tex.) (the “Texas district court litigation”) and (2) *Valtrus Innovations Ltd. v. Google LLC*, No. 4:22-cv-00020 (N.D. Tex.) (transferred to the Dallas Division of N.D. Texas and captioned 3:22-cv-00066-N). Pet. x; Paper 4, 1.

C. Overview of the ’454 Patent (Ex. 1001)

The ’454 patent is titled “Apparatus and Method for Routing a Transaction to a Partitioned Server.” Ex. 1001, code (54). The ’454 patent describes “an apparatus and method for routing a transaction to a partition on a server based at least in part on the configuration thereof.” *Id.* at 1:7–9.

According to the ’454 patent, individual servers may be used to process transactions or “multiple servers may be pooled, and the incoming transactions directed to an available server in the server pool using a load balancer. As such, the traffic is balanced across all of the servers in the pool, and the transactions are processed more quickly.” Ex. 1001, 1:22–26.

The ’454 patent recognizes that using multiple servers may be inefficient and one solution to this problem “is a partitioned server.” Ex. 1001, 1:27–34. The ’454 patent states that a “partitioned server is a

single server or aggregation of server resources subdivided to perform as multiple servers.” *Id.* at 1:34–36. “[A] partitioned server continues to serve multiple locations, multiple departments, and/or multiple transactions.” *Id.* at 1:36–38. With partitioned servers, “fewer individual servers are needed,” “operations are centralized,” multiple individual server characteristics are maintained, and “utilization of server resources is enhanced.” *Id.* at 1:38–42.

The ’454 patent states that because “[t]he partitions of partitioned servers . . . may not be equally weighted,” some partitions may be better able to process transactions than other partitions, based on the configuration of the partition. Ex. 1001, 1:43–44. Some partitions, for example, may be allocated more server resources (such as CPUs or number of processors) and the allocation may be dynamic. *Id.* at 1:44–56.

To address these issues, the ’454 patent describes a load balancer which recognizes each partition of a partitioned server and its associated configuration, and routes transactions to partitions based at least in part on those configurations of the partitions. Ex. 1001, Abstract. “The configuration may be determined in response to an event, on a continuous basis, etc. In addition, the load balancer may also use other factors, in combination with the configuration, to route the transaction to the partition, such as a characteristic of the transaction.” *Id.* The patent also states configuration comprises the configuration parameters of the respective partitions, including, for example, the resource allocation, the resource utilization, the average response time, or any other suitable information regarding the partitions. *Id.* at 7:34–39.

The partitions and the associated configuration may be stored in a configuration profile provided to the load balancer. Ex. 1001, Fig. 3.

Figure 3 is reproduced below:

FIG. 3

PARTITION ID	NETWORK ADDRESS	CORRESPONDING CONFIGURATION	ASSIGNED RANK
A	1.1.1.1	RESOURCE ALLOCATION = 20% RESOURCE UTILIZATION = 50% AVG RESPONSE TIME = 0.5 ms	3
B	1.1.1.2	RESOURCE ALLOCATION = 50% RESOURCE UTILIZATION = 35% AVG RESPONSE TIME = 0.2 ms	1
C	1.1.1.3	RESOURCE ALLOCATION = 30% RESOURCE UTILIZATION = 40% AVG RESPONSE TIME = 0.3 ms	2
.	.	.	.
.	.	.	.
PARTITION n	1.1.1.n	CONFIGURATION n	RANK n

Figure 3 illustrates a partition profile for maintaining the configuration of the partitions. *Id.* at 3:13–14. The partitions can be identified by “a name, serial number, etc.” *Id.* at 7:46–48. Based at least in part on the associated configurations, the partitions may be assigned a rank. *Id.* at 7:53–56. The ranks can be indicated by “a numerical assignment (e.g., ‘1’, ‘2’, ‘3’), an empirical assignment (e.g., ‘Fast’, ‘Slow’), a specific parameter (e.g., ‘0.5 ms’), etc.” *Id.*

D. Illustrative Claim

As mentioned above, Petitioner challenges claims 1–21 of the ’454 patent. Independent claim 1 is reproduced below.¹

1. [1PRE] A method for routing a transaction to a partitioned server, comprising:

¹ For ease of reference, we use Petitioner’s claim recitation numbering scheme as indicated by the bracketed numbers.

[1A] identifying a plurality of partitions on said partitioned server;

[1B] determining a configuration of each of said plurality of partitions;

[1C](1) providing said configurations to a load balancer,

[1C](2) wherein said transaction is routed by said load balancer to one of said plurality of partitions based at least in part on said configurations;

[1D] assigning a rank to each of said plurality of partitions based at least in part on said configurations, wherein said transaction is routed based on said ranking; and

[1E] storing said configurations in a partition profile, wherein said configurations are provided to said load balancer via said partition profile.

Ex. 1001, 8:62–9:10.

E. Evidence and Asserted Grounds

Petitioner relies upon the following evidence:

- (1) US 6,006,264, issued December 21, 1999 (“Colby” Ex. 1005);
- (2) US 6,587,938 B1, issued July 1, 2003 (“Eilert” Ex. 1006); and
- (3) US 6,092,178, issued July 18, 2000 (“Jindal” Ex. 1007).

Petitioner submits a declaration from Vijay K. Madiseti, Ph.D. (Ex. 1003).

Petitioner challenges the patentability of the '454 patent claims based on the following grounds (Pet. 2):

Ground	Claim(s) Challenged	35 U.S.C. § ²	Reference(s)/Basis
1	1–6, 8, 10,13,16-21	103	Colby
2	1-6, 8-10, 13-21	103	Colby, Eilert
3	1-21	103	Colby, Eilert, Jindal

III. ANALYSIS

A. Legal Standards

Petitioner bears the burden of persuasion to prove unpatentability, by a preponderance of the evidence, of the claims challenged in the Petition. 35 U.S.C. § 316(e). Except in limited circumstances not present here, this burden of persuasion does not shift to Patent Owner. *Dynamic Drinkware, LLC v. Nat’l Graphics, Inc.*, 800 F.3d 1375, 1378 (Fed. Cir. 2015). The Board may authorize an *inter partes* review if we determine that the information presented in the Petition and Patent Owner’s Preliminary Response shows that there is a reasonable likelihood that Petitioner will prevail with respect to at least one of the claims challenged in the Petition. 35 U.S.C. § 314(a).

As mentioned above, Petitioner’s challenges are based on obviousness. Pet. 2. A claim is unpatentable under 35 U.S.C. § 103 if the differences between the claimed invention and the prior art are such that the

² The relevant sections of the Leahy-Smith America Invents Act (“AIA”), Pub. L. No. 112–29, 125 Stat. 284 (Sept. 16, 2011), took effect on March 16, 2013. Because the '454 patent claims priority to an application filed before this date, our citations to 35 U.S.C. § 103 in this Decision are to the pre-AIA version. Our decision is not impacted, however, by which version of the statute applies.

claimed invention as a whole would have been obvious before the effective filing date of the claimed invention to a person having ordinary skill in the art to which the claimed invention pertains. *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 406 (2007). The question of obviousness is resolved based on underlying factual determinations including: (1) the scope and content of the prior art; (2) any differences between the claimed subject matter and the prior art; (3) the level of ordinary skill in the art; and (4) when in the record, objective evidence of nonobviousness.³ *Graham v. John Deere Co.*, 383 U.S. 1, 17–18 (1966).

B. Level of Ordinary Skill in the Art

The level of ordinary skill in the art is “a prism or lens” through which we view the prior art and the claimed invention. *Okajima v. Bourdeau*, 261 F.3d 1350, 1355 (Fed. Cir. 2001). The person of ordinary skill in the art is a hypothetical person presumed to have known the relevant art at the time of the invention. *In re GPAC Inc.*, 57 F.3d 1573, 1579 (Fed. Cir. 1995). In determining the level of ordinary skill in the art, we may consider certain factors, including the “type of problems encountered in the art; prior art solutions to those problems; rapidity with which innovations are made; sophistication of the technology; and educational level of active workers in the field.” *Id.*

Petitioner states a person of ordinary skill in the art would have had at least a bachelor’s degree in computer science, computer engineering, or a related field, with 3 years of experience in the area of distributed computer systems, including performance

³ At this stage of the proceeding, Patent Owner has not directed us to any objective evidence of non-obviousness. *See* Prelim. Resp.

and/or resource optimizations. A higher level of education may substitute for less experience.

Pet. 6 (citing Ex. 1003 ¶¶ 38–42). Patent Owner applies the skill level proposed by Petitioner. Prelim. Resp. 16.

We adopt Petitioner’s definition of the level of ordinary skill for the purposes of institution.

C. *Claim Construction*

In an *inter partes* review, the claims are construed using the same claim construction standard that would be used to construe the claim in a civil action under 35 U.S.C. § 282(b). *See* 37 C.F.R. § 42.100(b) (2021). This claim construction standard includes construing the claim in accordance with the ordinary and customary meaning of such claims as would have been understood by one of ordinary skill in the art. *Id.*; *see Phillips v. AWH Corp.*, 415 F.3d 1303, 1312–13 (Fed. Cir. 2005). In construing claims in accordance with their ordinary and customary meaning, we consider intrinsic evidence such as the specification and the prosecution history of the patent. *Phillips*, 415 F.3d at 1315–17. Extrinsic evidence, including expert and inventor testimony, dictionaries, and treatises, may also be used but is less significant than the intrinsic record. *Id.* at 1315. Usually, the specification is dispositive, and it is the single best guide to the meaning of a disputed term. *Id.* Any special definitions for claim terms must be set forth in the specification with reasonable clarity, deliberateness, and precision. *See In re Paulsen*, 30 F.3d 1475, 1480 (Fed. Cir. 1994).

In its Petition, Petitioner states that it construes the claim terms “in accordance with the ordinary and accustomed meaning as understood by a [person of ordinary skill in the art] and the patent’s prosecution history,”

except for the means plus function terms in claim 16 which it construes under 35 U.S.C. § 112 ¶ 6. Pet. 6–7.

1. Whether the preamble is limiting

The preamble of independent claims 1, 6, 16, and 17 all recite “routing a transaction to a partitioned server.” Ex. 1001, 8:62 (claim 1), 9:20–21 (claim 6), 10:15–16 (claim 16), 10:28–29 (claim 17).

Petitioner, in this proceeding, does not state whether the preamble is limiting but contends that Colby discloses the preamble. Pet. 16.

Patent Owner states that the preamble is limiting and that Petitioner agreed that it is limiting in the Texas district court litigation. Prelim. Resp. 17 (citing Ex. 2002, 2). At this stage, Patent Owner does not contest that Colby discloses the preamble. See Prelim. Resp.

Based on the present record, we make no determination at this stage of the proceeding whether the preamble of claims 1, 6, 16, and 17 is limiting. We suggest that Petitioner clearly indicate whether the preamble is limiting.

2. Construction of “configuration”

Independent claims 1, 6, 16, and 17 all recite “determining a configuration of each of said plurality of partitions.” Ex. 1001, 8:66–67 (claim 1), 9:27–28 (claim 6), 10:19–20 (claim 16), 10:32–33 (claim 17). The parties dispute the construction of a partition’s “configuration.”

Petitioner states that, according to the ’454 patent, a partition’s “configuration” includes “‘any suitable information’ about the partition, including ‘resource allocation,’ ‘resource utilization,’ and ‘average response

time” and “information regarding a server’s load, health, and content.” Pet. 19 (citing Ex. 1001, 2:22, 7:34–39, 7:50–53; Ex. 1015 ¶ 133 (complaint asserting that the “health” of a partition is a configuration of that partition); Madisetti Decl. ¶¶ 147, 151).

Patent Owner’s position about the meaning of “configuration” is unclear. Patent Owner does not state what “configuration” means, and only asserts what “configuration” does not mean. Prelim. Resp. 17–21. As we understand Patent Owner’s position, Patent Owner contends that the ’454 patent differentiates between a “configuration” and “other factors” used to route the transaction to a partition (*id.* at 17) and that “configuration” should not be construed so broadly to include parameters that the ’454 patent expressly describes as such “other factors” (*id.* at 21). Patent Owner, however, does not explain how one would differentiate between the claimed “configuration” and “other factors.”

First, we begin our analysis by examining the portions of the ’454 patent cited by Petitioner to support its proposed definition of “configuration.” Petitioner relies on the following portions of the ’454 patent.

The corresponding configuration 330 comprises the configuration parameters of the respective partition 160–162. For example, the corresponding configuration 330 may include the resource allocation . . . , the resource utilization . . . , the average response time . . . , or any other suitable information regarding the partitions 160–162.

Ex. 1001, 7:34–39.

[A]ny suitable information may be included as part of the corresponding configuration 300, including specifics of the various resources and the corresponding allocation thereof.

Id. at 7:50–53.

[T]he partition manager 150 may reside on a single partition 160, or have modules on each partition 160 for monitoring *the configuration thereof (e.g., the resource utilization, **load**, etc.)*.

Id. at 6:4–7 (emphasis added).

From these portions of the '454 patent, we understand that the '454 patent uses the term “configuration” and “suitable information” very broadly. These portions of the '454 patent also support explicitly the first portion of Petitioner’s definition that

a partition’s “configuration” includes “any suitable information” about the partition, including “resource allocation,” “resource utilization,” and “average response time”

and

information regarding a server’s load.

Pet. 19.

As to the remaining portion of Petitioner’s definition (i.e., information regarding a server’s “health and content”), Petitioner does not identify any portion of the '454 patent which addresses explicitly a server’s “health and content.” However, Petitioner, relying on Dr. Madisetti’s testimony, argues that a server’s “health and content” is “suitable information” about a partition’s configuration. Pet. 19 (citing Madisetti ¶ 151, Ex. 1001, 7:34–39, 7:50–53). Dr. Madisetti states that a person of ordinary skill in the art would have understood that “the '454 patent refers to a “configuration” of a partition as including “any suitable information” about the partition, including “resource allocation,” “resource utilization,” and “average response time.” Madisetti ¶ 147 (citing Ex. 1001, 2:22, 7:34–39, 7:50–53). He also states that his opinion is confirmed by Patent Owner’s complaint in the Texas district court litigation which states, “A partition’s ‘health’ is a configuration of the partition.” *Id.* (citing Ex. 1015 ¶ 133). He further states

that “information regarding a server’s load, health, and content, is independently a configuration, as that term is used in the ’454 patent.” *Id.* at ¶ 151 (citing Ex. 1001, 7:34–39, 7:50–53).

For purposes of institution, Petitioner supports its construction that configuration means information relating to a server’s health and content, but we encourage the parties to address this issue at trial. And, even if Petitioner’s use of “health” and “content” is not adequately supported, Petitioner adequately shows that information relating to load is supported by the ’454 patent and disclosed in Colby for purposes of institution.

Second, we examine Patent Owner’s support for its contention that “configuration” should not be construed so broadly to include parameters that the ’454 patent expressly describes as “other factors.” Prelim. Resp. 21; Sur-Reply 2–3. Patent Owner identifies the following portions of the ’454 patent which purported state that its load balancer may use “other factors”:

[T]he load balancer may also use ***other factors***, in combination with the configuration, to route the transaction to the partition, such as a characteristic of the transaction.

Ex. 1001, Abstract (emphasis added).

[T]he load balancer routes the transaction to one of partitions based at least in part on the configuration thereof. Optionally, one or more characteristics of the transaction may also be determined and used for routing the transaction to the appropriate portion.

Id. at 1:66–2:4; *see id.* at 5:26–30 (“For example, the characteristic(s) may include a type of transaction (e.g., video streams, backup requests, etc.), a requested level of service (e.g., high priority, low priority), a combination thereof, etc.).

The apparatus . . . comprises program code for determining a configuration (e.g., the allocation of resources) for each

partition. . . . The load balancer routes the transaction to one of the partitions of the server based at least in part on the configuration thereof. The load balancer may also use *other factors* (e.g., response time of each partition, network proximity, etc.) in combination with the configuration to route the transaction to the partition.

Id. at 2:21–33 (summary of the inventions, emphasis added).

[T]he load balancer 130 may route 180, 181, 182 the transaction 110 to one of the partitions 160-162 based [on] any number of *factors*, such as the configuration (e.g., resource allocation) of the partition 160 and/or other system metrics (e.g., response time, load).

Id. at 3:39–43 (emphasis added). From these portions of the ’454 patent cited by Patent Owner, we understand that the load balancer uses *other factors*, in combination with the configuration, to route the transaction to the partition. *See, e.g., id.* at Abstract. These factors include, for example, response time of each partition, network proximity, and system metrics such as response time and load. *See, e.g., id.* at 2:21–33, 3:39–43.

As we understand Patent Owner’s contention, because the ’454 patent states that the load balancer uses these “other factors,” the partition configuration cannot also use these “other factors” and should not be construed so broadly to use them. Prelim. Resp. 21.

Based on the record at this stage of the proceeding, we do not agree with Patent Owner’s arguments for two reasons. First, Patent Owner’s citation to the Summary of the Invention leaves out a critical portion:

The load balancer routes the transaction to one of the partitions of the server based at least in part on the configuration thereof. The load balancer may also use other factors (e.g., response time of each partition, network proximity, etc.) in combination with the configuration to route the transaction to the partition. **These factors may be included with the configuration,**

separately provided, and/or determined by the load balancer.

Ex. 1001, 2:27–35 (emphasis added to highlight the sentence missing from Patent Owner’s quotation). Contrary to Patent Owner’s argument that the “factors” used by the load balancer cannot also be used by the partition’s configuration, the sentence missing from Patent Owner’s quotation states explicitly the “factors may be included with the configuration” and undercuts Patent Owner’s arguments. *Id.*

Second, as explained above, Petitioner’s proposed definition is supported explicitly by the ’454 patent except for information regarding a server’s “health and content.” Even assuming that Patent Owner’s position is correct, none of its citations to the ’454 patent indicate that the “other factors” include information relating to the server’s health and content. As noted above, we encourage the parties to address at trial whether configuration may be construed to include information relating to a server’s health and content. We also encourage the parties to whether information about the server’s health and content is suitable information about the partitions. *See, e.g.*, Ex. 1001, 7:34–39 (“the corresponding configuration 330 may include . . . any other suitable information regarding the partitions 160–162”).

For purposes of institution, we determine that the only claim term that needs to be construed is the partition’s “configuration” and, for the reasons above, we adopt Petitioner’s definition:

a partition’s “configuration” includes “any suitable information” about the partition, including “resource allocation,” “resource utilization,” and “average response time” and information regarding a server’s load, health, and content.

D. Ground 1: Asserted Obviousness of Claims 1–6, 8, 10, 13, and 16–21 Over Colby

Petitioner asserts that claims 1–6, 8, 10, 13, and 16–21 are unpatentable under 35 U.S.C. § 103 as being obvious over Colby (Ex. 1005). Pet. 7–50. Patent Owner filed a Preliminary Response opposing Petitioner’s challenge. Prelim. Resp. 21–38. Based on the present record, and for the reasons identified below, we determine Petitioner has shown a reasonable likelihood of prevailing on its obviousness challenge at least as to claim 1.

Below, we present a brief overview of Colby and then we address Petitioner’s and Patent Owner’s contentions.

1. Overview of Colby (Ex. 1005)

Colby is a U.S. patent titled “Method and system for Directing a Flow between a Client and a Server.” Ex. 1005, code (54). Colby describes a content-aware flow switch (sometimes referred to as CAFS) for directing content requests to a best-fit server among a selection of servers. *Id.* at code (57). The best-fit server is chosen based on the type of content requested, the quality of service requirements implied by the content request, the degree of load on available servers, network congestion information, and the proximity of the client to available servers. *Id.* The flow switch also provides the functionality of multiple physical web servers on a single web server in a way that is transparent to the client, through the use of virtual web hosts (VWH) and flow pipes. *Id.* Each virtual web host provide the functionality of a single server, and many virtual web hosts may reside on a single physical server. *Id.* at 2:27–33, 16:20–24.

2. *Claim 1*

Petitioner, relying on Dr. Madiseti's testimony, provides a limitation-by-limitation comparison of Colby to independent claim 1. Pet. 15–29.

a) *[IPRE] A method for routing a transaction to a partitioned server,*

Petitioner contends that, to the extent the preamble is limiting, Colby teaches the preamble. Pet. 15–17. According to Petitioner, Colby's "method for servicing a content request using a [content aware flow switch]" is the recited "method for routing a transaction." *Id.* at 16 (citing Ex. 1005, 4:20-22, Fig. 3; Madiseti Decl. ¶ 132). Petitioner explains that Colby's "content" is "loosely defined as any information that a client application is interested in receiving," including "an HTML page." *Id.* at 16 (citing Ex. 1005, 1:59–65). Petition further explains that a person of ordinary skill in the art would have understood that the "content request" referenced in Colby's Figure 3 and related discussion is an example of the recited "transaction." *Id.* (citing Madiseti Decl. ¶¶ 133–136; Ex. 1005, 16:5–9; Ex. 1001, 3:61–4:4 ("transaction" may be "a uniform resource locator (URL) requesting a hypertext markup language (HTML) page")).

Petitioner contends that Colby routes transactions the "partitioned server," as claimed. Pet. 17 (citing Madiseti Decl. ¶¶ 138–141). Petitioner explains that the '454 patent says, "a partitioned server is a single server or aggregation of server resources subdivided to perform as multiple servers," and that this is "the same as Colby's physical web server with multiple VWHs," which "are implemented on and 'share a single physical web server,' with each VWH 'provid[ing] the functionality of a single physical web server.'" *Id.* (citing Ex. 1001, 1:34–37; Madiseti ¶¶ 139; Ex. 1005, 2:29–44).

b) [1A] identifying a plurality of partitions on said partitioned server

Petitioner contends that Colby teaches this limitation

[b]ecause each of Colby's VWHs "provides the functionality of a single physical web server" and "has one or more public virtual IP address[es] that clients use to access content on the [VWH]," and because multiple VWHs may "share a single physical web server" . . . , a [person of ordinary skill in the art] would have understood each VWH is an example of what the '454 patent calls a *partition* on a *partitioned server*, as claimed.

Pet. 17–18 (citing Ex. 1005, 2:23–44; Madisetti Decl. ¶¶ 142–143).

According to Petitioner, Colby's virtual web host (VWH) implementation is a physical web server with partitions for two reasons. First, Petitioner contends that it is expressly taught by Colby. Pet. 12. Colby states that a "virtual web host" ("VWH")—referred to as a "logical entity"—may "provide[] the functionality of a single physical web server," and that "*several* virtual web hosts . . . may share a single physical web server." *Id.* (citing Ex. 1005, 1:8, 2:22–36; *see also id.*, 16:20–24 ("Each of the physical web servers 100a–c may embody one or more [VWHs]."); Madisetti Decl. ¶ 119). Each "virtual web host has one or more public virtual IP address that clients use to access content," and is "uniquely identified by its public IP address." *Id.* (citing Ex. 1005, 2:36–39; Madisetti Decl. ¶ 120). Petitioner states that, given these disclosures, a person of ordinary skill in the art would have understood that (1) "Colby teaches a switch that is *not* restricted to operating solely with physical servers, and discloses a CAFS used with 'web servers' (e.g., in Fig. 1b) implemented as multiple VWHs on a 'single physical web server'" and (2) "the 'single physical web server' acts as a 'partitioned server' for the VWHs." *Id.* (citing Ex. 1001, 1:34–36 (describing "partitioned server" as "a single server or

aggregation of server resources subdivided to perform as multiple servers”); Ex. 1005, 2:29–44; Madisetti Decl. ¶¶ 121–122).

Second, Petitioner argues that, alternatively, a person of ordinary skill in the art would have been motivated to implement Colby’s web servers as multiple VWHs on a single physical web server. Pet. 13 (citing Madisetti Decl. ¶ 123). Petitioner presents the following reasons:

First, Colby expressly teaches using its CAFS with physical web servers that “embody one or more virtual web hosts” and discloses that VWHs may each “provide[] the functionality of a single physical web server.” Colby, 16:21–24, 2:27–33; Madisetti, ¶ 124.

Second, such an implementation would have conferred the many well-known benefits (acknowledged in the ’454 patent)—*e.g.*, flexibility, efficiency, ease of management—of having multiple servers implemented on a single physical server. Madisetti, ¶ 125; Kauffman (Ex. 1010), 2:7–18; Shahabuddin (Ex. 1011), 2:31–40; ’454 patent, 1:34–42.

Third, such an implementation would have been nothing more than a simple substitution of one known element (multiple physical web servers) for another (multiple VWHs on a single physical web server each operating as a web server), yielding the predictable result of providing the functionality of multiple web servers in a content delivery network environment like the one described in Colby. Colby, 2:26–37; Madisetti, ¶ 126. *See KSR Int’l v. Teleflex*, 550 U.S. 398, 416–17 (2007).

Pet. 13.

c) *[1B] determining a configuration of each of said plurality of partitions*

Petitioner contends that Colby teaches this limitation by disclosing that its “ICP [Intelligent Content Probe] ‘probes the servers’ to determine ‘information regarding server status and content’ for each server” and then

“stores this information — ‘status’ data like ‘load metrics’ and server ‘health,’ as well as ‘content’ data — about each server.” Pet. 19 (citing Ex. 1005, 6:58–63, 7:4–15, 8:19–25; Madisetti Decl. ¶¶ 148–149). Relying on the testimony of Dr. Madisetti, Petitioner asserts that “[e]ach piece of information regarding a server’s load, health, and content is independently *a configuration* as claimed in the ’454 patent.” *Id.* (citing Madisetti Decl. ¶ 151; Ex. 1001, 7:34–39, 7:50–53). According to Petitioner, a person of ordinary skill in the art would have understood “each piece of information (load, health, content) is *determined* by Colby’s [content-aware flow switch] when the ICP ‘probes’ the servers and ‘updates the CSD’ with information for use in Colby’s best-fit server selection process.” *Id.* (citing Ex. 1005, 6:58–63, 7:4–15, 8:19–25; Madisetti Decl. ¶ 151).

Finally, Petitioner contends that because servers 100a-100c are the “plurality of partitions” in Colby’s VWH Implementation, the ICP’s collection of server “status and content” information (citing Ex. 1005, 8:19–21) used during Colby’s method “for servicing a content request using a [content-aware flow switch]” (citing *id.* at 4:20–22), constitutes “determining a configuration of each of said plurality of partitions” as recited in limitation 1B. Pet. 20 (citing Madisetti Decl. ¶ 152).

d) [1C](1) providing said configurations to a load balancer

Petitioner contends that Colby teaches this limitation. Pet. 20–21. According to Petitioner, Colby’s “content-aware flow *switch* [CAFS]” includes a Web Flow Redirector (WFR), Content Server Database (CSD), and Flow Admission Control (FAC), “which together ‘are responsible for selecting a server to service a content request based on a variety of criteria,’

including ‘load’ on the servers.” *Id.* at 20 (citing Ex. 1005, 6:24–28, 6:35–44, 6:65–67, 8:33–54; 2:48–58, claims 1–2; Madisetti Decl. ¶¶ 155–157). Petitioner explains that process uses “load metrics indicating [server] load,” and “load balance algorithm[s] to use when accessing [particular] content.” *Id.* (citing Ex. 1005, 6:42–63; Madisetti Decl. ¶ 158). Thus, Petitioner contends that Colby’s WFR, CSD, and FAC collectively comprise the recited “load balancer.” *Id.* at 20–21 (citing Madisetti Decl. ¶ 159).

Petitioner further explains that “Colby’s ICP collects information comprising *said configurations*—*i.e., a configuration of each of said plurality of partitions* recited in [1B], such as load, health, or content” and the “‘ICP’s job is to ‘update[] the CSD’ with these configurations — that is, to ‘populate the CSD with server and content information,’ *e.g.*, server ‘status’ information like ‘load’ or ‘health,’ and ‘content’ information.” Pet. 21 (citing Ex. 1005, 6:58–63, 7:4–12, 8:19–26; Madisetti Decl. ¶¶ 160–161). Because the CSD is part of the load balancer, Petitioner contends that this process constitutes providing the configurations to a load balancer. *Id.* (citing Madisetti Decl. ¶ 161).

e) *[1C](2) wherein said transaction is routed by said load balancer to one of said plurality of partitions based at least in part on said configurations*

Petitioner contends that Colby teaches this limitation. Pet. 21–24. According to Petitioner, Colby’s content-aware flow switch “services” content requests received from clients and “forwards” them to a best-fit server (VWH) for each request and each “content request” is the recited “transaction.” *Id.* at 21 (citing Ex. 1005, Abstract, 2:48–53, 4:21–22, Fig. 3; Madisetti Decl. ¶ 162). Petitioner explains that forwarding the transaction

constitutes “routing” the transaction, as the term “routing” is used in the ’454 patent, and that Colby’s transactions (i.e., “content requests”) are routed by Colby’s load balancer (i.e., Colby’s WFR, CSD, and FAC). *Id.* (citing Madisetti Decl. ¶ 163). Thus, Petitioner contends that “said transaction is routed by said load balancer” to Colby’s VWH (i.e., a claimed partition) on Colby’s multiple-VWH web server (i.e., a claimed partitioned server). *Id.* (citing Madisetti Decl. ¶ 169).

According to Petitioner, a person of ordinary skill in the art would have understood that each of Colby’s “content requests” (i.e., transactions) is routed based at least in part on the configurations, as claimed. Pet. 22–23 (citing Madisetti Decl. ¶ 170). Petitioner explains that (1) the “content” information, “health” information, and “load metrics” collected for each partition are each configurations (i.e., a configuration of each of said plurality of partitions recited in limitation 1B) and (2) each content, health, load from each partition is used in various steps of Colby’s method “for servicing a content request using a [CAFS].” *Id.* (citing Ex. 1005, 4:20–22).

f) [1D] assigning a rank to each of said plurality of partitions based at least in part on said configurations, wherein said transaction is routed based on said ranking

Petitioner contends that a person of ordinary skill in the art would have understood Colby’s method for servicing a content request involves “assigning a rank to each of said plurality of partitions based at least in part on said configurations” for at least two independent reasons. Pet. 24 (citing Madisetti Decl. ¶ 178).

First, Petitioner contends that Colby’s method includes selecting a “best-fit server” based at least in part on the “server load metrics” of each of

the “candidate servers” and the skilled artisan would have understood that “selecting a ‘best’ server from among multiple candidates based on load metrics is an ‘empirical assignment’ (Ex. 1001, 7:52-56), *i.e.*, **rank**, based on evaluating such metrics from all candidates (to determine the ‘best’).” Pet. 23–24 (citing Colby, Abstract, claims 1–2; Madisetti Decl. ¶ 179; Ex. 1019, 507). Thus, Petitioner asserts that because Colby’s various “load metrics” are each a “configuration” of a partition, Colby’s rankings are “based at least in part on said configurations,” as claimed. *Id.* at 24 (citing Madisetti Decl. ¶ 179).

g) *[1E] storing said configurations in a partition profile, wherein said configurations are provided to said load balancer via said partition profile*

Petitioner contends that Colby teaches this limitation. Pet. 28–29. Petitioner explains that the recited “configurations” are the server “load,” “health,” and “content” information collected by the ICP, and that Colby expressly states “load” information for each server is stored in the CSDs “server record” for each server. *Id.* at 28 (citing Ex. 1005, 6:58–63). Petitioner also explains that a person of ordinary skill in the art would have understood that “health” information is also stored in “server records,” or would have found it obvious to implement Colby this way, “because: (1) Colby’s ‘server record’ stores information about individual servers; and (2) Colby says the ICP ‘probes’ servers to ‘monitor’ their ‘health,’ and then ‘updates the CSD [Content Server Database]’—which stores the ‘server records’ — with the data it collects.” *Id.* (citing Ex. 1005, 6:56–58, 7:7–12, 8:19–25; Madisetti Decl. ¶¶ 192–194).

Finally, Petitioner explains that Colby's Content Server Database (CSD), which contains server and content records with information regarding specific servers (i.e., VWHs), is the recited "partition profile," and Colby's request servicing process includes storing the configurations in that profile. Pet. 29 (citing Ex. 1005, 6:38; Madisetti Decl. ¶¶ 195–196). Petitioner contends that, "[l]ike the 'partition profile' in the '454 patent, which is 'a computer readable file, such as a database' 'for maintaining the configuration of the partitions'" (Ex. 1001, 7:16–20), Colby's CSD "maintains several databases" (Colby, 6:41–42) that contain configuration information. Thus, because Colby's CSD is part of the "load balancer" in limitation [1C](1), Petitioner states that the "configurations are provided to said load balancer via said partition profile," as claimed. *Id.* (citing Madisetti Decl. ¶¶ 197–199; Ex. 1022, 13:15 ("via" means "by way of" or "by means of").

h) Patent Owner's Arguments

Patent Owner contends that Petitioner's contentions are erroneous for several reasons. First, Patent Owner contends that Colby, alone or in combination, does not disclose "determining a configuration of each of said plurality of partitions" as recited in all independent claims because Colby teaches a configuration associated with a webserver, not a configuration for a partition on a partitioned server. Prelim. Resp. 21–26. According to Patent Owner, Petitioner incorrectly asserts that each of Colby's web servers 100a-c is a virtual web host that "may share a single physical web server." *Id.* at 23 (citing Ex. 1005, 2:29–30 (cited on Pet. 17)). Patent Owner explains that Figure 21a and Colby's discussion at column 16, lines 30–35

disclose that “VWHs and web servers are distinct and not co-extensive.” *Id.* at 23. Figure 21a, as annotated by Patent Owner, is reproduced below.

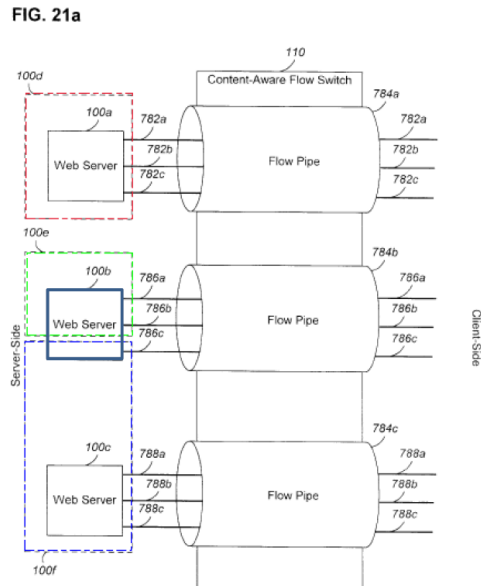


Figure 21a illustrates a block diagram of flow pipe traffic through a content-aware flow switch. Ex. 1005, 4:50–51. Patent Owner annotated Figure 21 with colored boxes to illustrate VWHs 100d, 100e, 100f.

The '454 patent states:

Referring to FIG. 21a, web servers 100a-c provide service to VWHs 100d-f as follows. Web server 100a provides all services to VWH 100d. Web server 100b provides service to VWH 100e and a portion of the services to VWH 100f. Web server 100c provides service to the remainder of VWH 100f.

Ex. 1005, 16:30–34. In light of this discussion and Figure 21a, Patent Owner argues that, contrary to Petitioner’s position that Colby’s web servers 100a-c is a virtual web host that may share a single physical web server (Pet. 17), Colby’s “the web servers 100a-c are not the same as VWHs 100c-d.” Prelim. Resp. 25.

In a related argument, Patent Owner also argues that Colby, alone or in combination, does not disclose additional elements reciting the claimed

“configuration.” Prelim. Resp. 29–34. These additional elements require, among others, “providing said configuration to a load balancer,” “wherein said transaction is routed ... to one of said plurality of partitions based at least in part on said configuration(s),” and “assigning a rank to each of said plurality of partitions based at least in part on said configurations.”

Ex. 1001, 9:1–7, 9:30–38, 10:21–27, 10:37–40. Patent Owner contends that, because the configurations that Petitioner points to are associated with a server and not those for partitions on a partitioned server, Colby does not disclose the additional limitations that recite the “configuration” element. *Id.* at 29.

In another related argument, Patent Owner argues that Colby, alone or in combination, does not disclose “identifying a plurality of partitions on said partitioned server” as recited in the claims. Prelim. Resp. 34–38. Patent Owner contends that Petitioner erroneously “equates identifying a subset of VWHs via their public IP addresses with identifying of a plurality of partitions on a partitioned server” (citing Pet. 18) because Colby’s “server records store the IP addresses of the servers 100a-c, not the IP address of VWH 100d-f.” *Id.* at 34–35.

Based on the present record of this proceeding, we find Petitioner has sufficiently shown a likelihood of success with respect to these issues. First, Petitioner adequately explains why Colby’s virtual web host (VWH) implementation is a physical web server with partitions because (1) it is expressly taught by Colby (see, e.g., Pet. 12) and (2) a person of ordinary skill in the art would have been motivated to implement Colby’s web servers as multiple VWHs on a single physical web server (*id.* at 13–15 (citing Madisetti Decl. ¶¶ 119–129)). Second, we understand that Figure 21a’s

embodiment is only one way of implementing multiple VWHs in a server environment. Reply, 2–3.

Second, Patent Owner argues that Petitioner has not offered competent evidence to show why each of “load, health, content” is a claimed configuration. Prelim. Resp. 27–29. According to Patent Owner, Petitioner erroneously argues that “a server’s load, health, and content is independently a configuration as claimed in the ’454 patent” because, as Patent Owner’s claim construction argues, the ’454 patent distinguishes between “configuration” and “other system metrics (e.g., response time, load).” *Id.* at 27. We do not agree with Patent Owner’s claim construction arguments for the reasons discussed in Section III.C.2 — determining that a partition’s “configuration” includes “any suitable information” about the partition, including “resource allocation,” “resource utilization,” and “average response time” and information regarding a server’s load, health, and content.

Third, Patent Owner argues that Petitioner fails to comply with 37 C.F.R. § 42.104(b)(3)–(4) because “Petitioner fails to inform the Board of its proposed constructions for claim terms in the district court or how the references disclose the limitation under its own definition.” Prelim. Resp. 38; *see id.* at 38–42. As discussed above, Petitioner supports its present claim constructions and obviousness arguments for purposes of institution but the parties may address at trial contradictory statements between this proceeding and the district court.

i) *Conclusion*

Based on the record before us, we determine Petitioner demonstrates a reasonable likelihood of prevailing in showing that the subject matter of independent claim 1 would have been obvious over Colby.

3. *Independent Claims 6, 16, and 17*

Independent claims 6, 16, and 17 recite limitations similar to the limitations recited in independent claim 1 and Petitioner relies on its analysis addressing independent claim 1 to address independent claim 13. *See* Ex. 1001, 9:30–38 (claim 6), 10:15–40 (claims 16, 17); Pet. 37–38 (comparing limitations in claims 1 and 6), 41–42 (comparing limitations in claims 6 and 16), 42–44 (comparing limitations in claims 1 and 17). Patent Owner’s arguments addressing independent claims 6, 16,⁴ and 17 rely on the same arguments addressed above in the context of independent claim 1. *See* Prelim. Resp. 21–38. Accordingly, for the same reasons as those explained above addressing independent claim 1, for purposes of institution and based on the record before us, Petitioner establishes sufficiently that the subject

⁴ Patent Owner argues that Petitioner fails to identify the specific portions of the Specification that describe the structure, material or acts corresponding to each claimed function of claim 16. Prelim. Resp. 38 (citing 37 C.F.R. § 42.104(b)(3)). Petitioner states, in part, that “[l]imitations [16A]–[16D] recite ‘means for’ performing the same functions performed by ‘program code’ recited in limitations [6C]–[6F], respectively, and ‘the load balancer that routes transactions as recited in [6E] performs the required function of **routing said transaction** in [16C].” Pet. 41 (citing Madisetti Decl. ¶¶ 232–233). We understand that Petitioner contends that the structure, material, and acts in limitations 6C–6F and limitations 6E correspond to the functions in limitations 16A–16D and limitation 16C, respectively. In light of Petitioner’s detailed discussion of the claim 6 limitations, this is sufficient for purposes of institution but we encourage the parties to develop this issue at trial.

matter of independent claims 6, 16, and 17 would have been obvious over Colby.

4. *Dependent Claims 2–5, 8, 10, 13, and 18–21*

Claims 2–5, 8, 10, 13, and 18–21 depend directly from independent claims 1, 6 and 17. *See* Ex. 1001, 9:11–19, 9:39–10:14, 10:41–51. We have reviewed Petitioner’s contentions and cited supporting evidence addressing how Colby teaches or suggests each of the additional limitations recited in these dependent claims. *See* Pet. 29–36, 39–40, 45–50. As we preliminarily conclude that Petitioner demonstrates a reasonable likelihood of prevailing with respect to its obviousness challenge to independent claims 1, 6, and 17, we institute review on all challenged claims on all grounds set forth in the Petition, including Petitioner’s challenge to dependent claims 2–5, 8, 10, 13, and 18–21 as unpatentable over Colby. 37 C.F.R. § 42.108(a).

E. *Grounds 2–3: Asserted Obviousness of Claims 1–21*

Petitioner asserts that (1) claims 1–6, 8–10, 13–21 are unpatentable as obvious over Colby and Eilert (Pet. 50–58); and (2) claims 1–21 are unpatentable as obvious over Colby, Eilert, and Jindal (*id.* at 58–70). As we preliminarily conclude that Petitioner demonstrates a reasonable likelihood of prevailing with respect to its patentability challenge to independent claims 1, 6, and 16, and 17, we institute review on all challenged claims on all grounds set forth in the Petition, including Petitioner’s challenge to claims 1–21. 37 C.F.R. § 42.108(a).

IV. DISCRETIONARY DENIAL UNDER 35 U.S.C. § 325(d)

Patent Owner argues that we should exercise discretion to deny institution under 35 U.S.C. § 325(d) because “the examiner has already considered whether the claims are allowable over prior art — such as that described in the background of the ‘454 patent — in combination with references such as Kubala (which has substantial disclosure as Eilert).”

Prelim. Resp. 42. Patent Owner contends that (1) “Colby is no different than the partitioned servers described in the background where the load balancing is done blind to ‘the partitions of a server and the various configurations thereof’” (*id.* at 42–43 (citing Ex. 1001, 1:56–58); (2) “Eilert has essentially the same disclosures as the Kubala and Schmidt that the examiners cited in same the office action where he allowed the claims” (*id.* at 43 (citing Ex. 1002, 299); and (3) “Jindal is merely cited for additional information to be included for ranking (Pet. 63–64) and still relies on combining general load balancing and Eilert in the first instance” (*id.* at 43–44).

We decline to exercise our discretion to deny the Petition because the Examiner did not consider Colby, singly or in combination with Eilert and Jindal. For the reasons discussed above in Section III.D.2.b-c, Petitioner adequately shows that Colby’s load balancing is not done blind to the partitions of a server and the various configurations thereof, as argued by Patent Owner. We agree with Petitioner that “Patent Owner does not meaningfully explain any overlap between arguments made during examination and the manner in which Petitioner relies on the prior art.” Reply 5 (quoting *Snap v. SRK Tech.*, IPR2020-00820, Paper 15, at 28 (Oct. 21, 2020)).

V. CONCLUSION

After considering the evidence and arguments presented in the current record, we determine that Petitioner demonstrates a reasonable likelihood of prevailing in establishing that at least one of the challenged claims of the '454 patent is unpatentable. We, therefore, institute trial on all challenged claims and all grounds raised in the Petition.

At this stage of the proceeding, we have not made a final determination as to the patentability of any challenged claim or as to the construction of any claim term. Any final determination will be based on the record developed during trial. We place Patent Owner on express notice that any argument not asserted in a timely-filed Response to the Petition, or in another manner permitted during trial, shall be deemed waived, even if that argument was presented in the Preliminary Response. Similarly, if the parties do not present their proposed claim constructions in their briefs, or in another manner permitted during trial, they will be precluded from presenting such arguments and their arguments shall be deemed waived.

VI. ORDER

In consideration of the foregoing, it is hereby:

ORDERED that, *inter partes* review of claims 1–21 of U.S. Patent No. 7,523,454 B2 is instituted on all grounds raised in the Petition; and

FURTHER ORDERED that pursuant to 35 U.S.C. § 314(c) and 37 C.F.R. § 42.4, notice is hereby given of the institution of a trial; the trial will commence on the entry date of this decision.

IPR2022-01418
Patent 7,523,454 B2

For PETITIONER:

Anant Saraswat
Adam Wichman
WOLF, GREENFIELD, & SACKS, P.C.
Asaraswat-ptab@wolfgreenfield.com
Awichman-ptab@wolfgreenfield.com

For PATENT OWNER:

Hong Zhon
Amy Proctor
Blair Silver
IRELL & MANELLA LLP
hzhong@irell.com
aproctor@irell.com
bsilver@irell.com