

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

CEREBRAS SYSTEMS INC.,
Petitioner,

v.

REX COMPUTING, INC,
Patent Owner.

IPR2022-00741
Patent 10,355,975 B2

Before MICHAEL R. ZECHER, MICHAEL T. CYGAN, and
JULIET MITCHELL DIRBA, *Administrative Patent Judges*.

CYGAN, *Administrative Patent Judge*.

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

I. BACKGROUND

A. Background and Summary

Cerebras Systems Inc. (“Petitioner”) filed a Petition requesting *inter partes* review (“IPR”) of claims 1–19 of U.S. Patent No. 10,355,975 B2 (Ex. 1001, “the ’975 patent”). Paper 1 (“Pet.”). Rex Computing, Inc. (“Patent Owner”) filed a Preliminary Response. Paper 7 (“Prelim. Resp.”). Institution of an *inter partes* review is authorized by statute when “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). An *inter partes* review may not be instituted on fewer than all claims challenged in the Petition. 37 C.F.R. § 42.108(a).

Having considered the Petition and the Preliminary Response, we conclude that the information presented in the Petition establishes that there is a reasonable likelihood that Petitioner would prevail in challenging at least one of claims 1–19 of the ’975 patent as unpatentable under 35 U.S.C. § 103(a). Pursuant to § 314, we institute an *inter partes* review as to claims 1–19 of the ’975 patent on all grounds raised in the Petition.

B. Real Party in Interest

The parties identify themselves as the real parties in interest. Pet. 1; Paper 6, 1.

C. Related Matters

The parties indicate that the ’975 patent is involved in *Rex Computing, Inc. v. Cerebras Systems Inc.*, No. 1:21-cv-00525-MN (D. Del.), filed April 13, 2021. Pet. 2; Paper 6, 2.

D. The '975 Patent

The '975 patent (Ex. 1001) generally relates to architectures for multi-core computer processors. Ex. 1001, 1:19–37. The '975 patent identifies a need for architectures that are “faster, smaller, less power consuming, more reliable, and easier to use.” *Id.* at 1:37–39. The '975 patent describes a system in which each processing core corresponds to a different router of a set of routers, and through adjacent routers of the set, communicates to other processing cores. *Id.* at 1:43–53. The system may be arranged as a set of tiles in a grid configuration, with each tile including a processing core and its corresponding router. *Id.* at 1:63–67.

Each router receives data packets from the adjacent routers based on a physical destination address of the data packet. *Id.* at 1:53–57. Each router sends data packets to adjacent routers, or to the processing cores connected thereto, wherein each router is operable to retain a data packet in the event of a traffic condition. *Id.* at 1:57–61. Each router implements a deterministic, i.e., static, routing policy. *Id.* at 1:61–62; Prelim. Resp. 2. Execution of an application may be optimized by an optimization module, which assigns functions of the application to tiles. *Id.* at 15:51–61.

E. Challenged Claims

Petitioner challenges claims 1–19 of the '975 patent. Claims 1 and 13 are independent. Claim 1, reproduced below (with bracketed identifying labels added), is an independent claim:

1. [1pre] A system comprising:
[1a] a set of processor cores;
[1b] a set of routers each including a set of input ports and a set of output ports, wherein:

[1c] each processor core of the set of processor cores corresponds to a different router of the set of routers, wherein each processor core and corresponding router form a tile,

[1d] each processor core is communicatively coupled with a corresponding router via the router's set of input ports and set of output ports,

[1e] each router is communicatively coupled with one or more adjacent routers via the set of input ports and the set of output ports, and wherein each processor core is communicatively coupled with the other processor cores via the set of routers,

[1f] each router is operable to receive one or more data packets from the one or more adjacent routers or the processor core corresponding to the router,

[1g] based on a physical destination address of a data packet, each router is operable to send one or more data packets to the one or more adjacent routers or the processor core corresponding to the router, wherein each router is operable to retain a data packet in the event of a traffic condition, and

[1h] each router implements a static priority routing policy;
and

[1i] an optimization module configured to:

[1j] determine optimal function assignment configurations for groups of tiles, and

[1k] assign two or more functions, which communicate at least unilaterally more frequently with one another than with other functions, to groups of adjacent tiles based on an optimal function assignment configuration determination, wherein the two or more functions are assigned to groups of tiles communicatively coupled in square configurations when the function executes optimally when executed by the groups of tiles communicatively coupled in the square configurations and are assigned to groups of tiles communicatively coupled in linear configurations when the function executes optimally when executed by the groups of tiles communicatively coupled in the linear configurations.

Ex. 1001, 28:24–67.

F. Prior Art Relied Upon

Petitioner relies upon the prior art references listed below. Pet. 4.

U.S. Patent No. 8,531,943 B2 to Olofsson, filed October 29, 2009, and issued on September 10, 2013 (Ex. 1004);

U.S. Patent Application Publication No. 2016/0179728 A1 to Kaul *et al.* (“Kaul”), filed December 17, 2014, and published on June 23, 2016 (Ex. 1005);

“The Raw Microprocessor: A Computational Fabric for Software Circuits and General-Purpose Programs,” Taylor et al., IEEE Micro 25, March-April 2002 (“MIT2002”; Ex. 1006);

“Scalar Operand Networks: On-Chip Interconnect for ILP in Partitioned Architectures,” Taylor et al., (“MIT2003”; Ex. 1007);

“Evaluation of the Raw Microprocessor: An Exposed-Wire-Delay Architecture for ILP and Streams,” Taylor et al., (“MIT2004”; Ex. 1008);

“A Modified NoC Router Architecture with Fixed Priority Arbiter,” Ansari et al. (“Ansari”), International Journal of Science and Research, Vol. 4, n.10, 923, October 2015 (Ex. 1009).

Based on the current record, we determine Petitioner has made a threshold showing that these references are prior art patents or printed publications under 35 U.S.C. § 102. *See* Pet. 3–4; Ex. 1005, codes (22), (43); Ex. 1006, codes (22), (45); Ex. 1007, codes (22), (63). With respect to MIT2002, MIT2003, and MIT2004, Petitioner relies upon a declaration of Dr. Sylvia Hall-Ellis to establish the provenance of these references. Pet. 10. Dr. Hall-Ellis states that MIT2002 was “available on March 1, 2002, in the *ACM Digital Library* and on August 7, 2002, in the *IEEE Xplore*

database.” Ex. 1016 ¶ 51. Dr. Hall-Ellis further states that the IEEE Micro article publication became publicly available on, or shortly after, April 17, 2002. *Id.* ¶ 53.

With respect to MIT2003, Petitioner asserts that it is referenced in MIT2004; we understand the corresponding citation to be the same-titled paper in “2003 HPCA” at pages 341–353. Pet. 10; Ex. 1008, 12. Dr. Hall-Ellis states that MIT2003 is “a copy of a conference paper published in the volume titled *HPCA-9 2003: the Ninth International Symposium on High-Performance Computer Architecture: Proceedings* published by the IEEE Computer Society with a 2003 copyright date.” Ex. 1016 ¶ 58. Dr. Hall-Ellis states that this conference took place on February 8–12, 2003. *Id.* Dr. Hall-Ellis further states that MIT2003 “was added to the *IEEE Xplore* database on February 28, 2003, and to the *ACM Digital Library* on February 8, 2003.” *Id.* ¶ 59. Dr. Hall-Ellis further states “that the book titled *HPCA-9 2003: the Ninth International Symposium on High-Performance Computer Architecture: Proceedings*, including the MIT2003 reference, was publicly available in the Linda Hall Library as of March 5, 2003.” *Id.* ¶ 61.

With respect to MIT2004, Dr. Hall-Ellis states that it “is a copy of a conference paper published in the volume titled *Proceedings of the 31st Annual International Symposium on Computer Architecture* published by the IEEE Computer Society with a 2004 copyright date.” *Id.* ¶ 65. Dr. Hall-Ellis states that this conference took place on June 19–23, 2004. *Id.* Dr. Hall-Ellis further states that MIT2004 was added to the IEEE Xplore database on July 12, 2004. *Id.* ¶ 66. Dr. Hall-Ellis states “that the book titled *Proceedings of the 31st Annual International Symposium on Computer*

Architecture was publicly available in the Linda Hall Library as of July 16, 2004.” *Id.* ¶ 68.

For the above described reasons, we determine that Petitioner has met its burden to show, by a reasonable likelihood, that the cited references are available as prior art to the ’975 patent.

G. Asserted Grounds of Unpatentability

Petitioner challenges the patentability of claims 1–19 of the ’975 patent on the following grounds (Pet. 4):

Claim(s) Challenged	35 U.S.C. § ¹	Reference(s)
1, 3, 4, 7–9, 11–13, 15–19	103	MIT2002, MIT2003, MIT2004
1–4, 7–19	103	MIT2002, MIT2003, MIT2004, Ansari
5, 6	103	MIT2002, MIT2003, MIT2004, Kaul
5, 6	103	MIT2002, MIT2003, MIT2004, Ansari, Kaul
1, 3–9, 11–13, 15–19	103	Olofsson, Kaul
1–19	103	Olofsson, Kaul, Ansari

Petitioner also relies on the Declaration of Dr. Thomas Conte to support its positions. Ex. 1002.

¹ The Leahy-Smith America Invents Act (“AIA”), Pub. L. No. 112-29, 125 Stat. 284, 287–88 (2011), revised 35 U.S.C. §§ 102 and 103, effective March 16, 2013. Because the application from which the ’975 patent issued was not filed before this date, and does not claim the benefit of an application that was filed before this date, the post-AIA versions of §§ 102 and 103 apply.

II. DISCUSSION

A. Legal Standards

1. § 103

A claim is unpatentable under § 103 if the differences between the claimed subject matter and the prior art are such that the subject matter, as a whole, would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 406 (2007). The question of obviousness is resolved on the basis of 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 skill in the art and (4) when in evidence, objective indicia of non-obviousness (i.e., secondary considerations).² *Graham v. John Deere Co.*, 383 U.S. 1, 17–18 (1966). We analyze the asserted grounds based on obviousness with these principles in mind.

B. Level of Skill in the Art

Petitioner offers an assessment as to the level of skill in the art as of the time the ’975 patent was filed. Pet. 6. Relying upon the Conte Declaration, Petitioner asserts a person of ordinary skill in the art would have had “a Bachelor of Science degree in electrical engineering, computer engineering, or a related subject matter, plus at least two years of professional experience in the field of microprocessor design or a similar field. . . . and more experience would compensate for less formal education,

² Patent Owner does not identify any secondary considerations to be considered. Accordingly, we do not consider this aspect of the obviousness analysis at this preliminary stage of the proceeding.

and vice versa.” *Id.* (citing Ex. 1002 ¶¶ 27–30). At this juncture, Patent Owner has not provided an assessment of the level of ordinary skill in the art. *See* Prelim. Resp. To the extent necessary, and for purposes of this Decision, we accept the assessment offered by Petitioner as it is supported by the Conte Declaration and is consistent with the ’975 patent and the asserted prior art.

C. Claim Construction

In this proceeding, we apply the same standard that is used to construe a patent claim in a civil action under 35 U.S.C. § 282(b), and accordingly, we construe each claim in accordance with the ordinary and customary meaning of such claim, as understood by one of ordinary skill in the art in the context of the patent, and the prosecution history pertaining to the patent, as articulated in *Phillips v. AWH Corp.*, 415 F.3d 1303 (Fed. Cir. 2005) (en banc). 37 C.F.R. § 42.100(b) (2021).

Petitioner submits that no constructions are necessary because the prior art renders the claims obvious under either the Petitioner’s or the Patent Owner’s proposed litigation constructions. Pet. 6. In particular, Petitioner points to Patent Owner’s proposed litigation construction of “static priority routing policy,” and “destination address.” *Id.* at 6–7. With respect to “optimization module,” Petitioner asserts that it is a means-plus-function limitation, but that Patent Owner characterizes it as not a means-plus function limitation because the “‘linear’ and ‘square’ configurations of tiles [are] the definite, structural way to perform the optimization,” and that Patent Owner points to corresponding structure in the form of “a multi-core microprocessor having linear and square configurations of tiles.” *Id.* at 7

(citing Ex. 1018, 8, 10). Petitioner states that it adopts Patent Owner’s construction of this term. *Id.*

Petitioner further adopts Patent Owner’s construction in the litigation of “optimal,” and “optimally” as the comparison and binary choice between tiles in a square or linear configuration. *Id.* at 8. Petitioner asserts that “data packet” should be given its plain and ordinary meaning of “a unit of data sent over a network.” *Id.* (citing Ex. 1002 ¶¶ 65, 157–158).

In its Preliminary Response, Patent Owner asserts that no construction of any claim term is necessary at the institution stage. Prelim. Resp. 7–8. Patent Owner asserts that the different litigation constructions need not be resolved because institution should be denied for other reasons, relying on the plain language of the claims. However, because Patent Owner addresses the interpretation of “data packet” in contesting Petitioner’s application of Olofsson, further briefing on the construction of that term is recommended. *Id.* at 8.

“[W]e need only construe terms ‘that are in controversy, and only to the extent necessary to resolve the controversy.’” *Nidec Motor Corp. v. Zhongshan Broad Ocean Motor Co. Matal*, 868 F.3d 1013, 1017 (Fed. Cir. 2017) (quoting *Vivid Techs., Inc. v. Am. Sci. & Eng’g, Inc.*, 200 F.3d 795, 803 (Fed. Cir. 1999)). We agree that no constructions are required at this preliminary stage for the above-disputed terms; therefore, we do not directly address the parties’ proposed constructions, except as discussed, *infra*, with respect to the asserted contentions of obviousness.

D. Asserted Obviousness of Claims over MIT2002, MIT2003, and MIT2004

Petitioner contends that claims 1, 3, 4, 7–9, 11–13, and 15–19 of the

'975 patent are unpatentable as obvious under 35 U.S.C. § 103, over the combined teachings of MIT2002, MIT2003, and MIT2004. Pet. 11.

Petitioner also relies upon the Conte Declaration to support its positions. Ex. 1002. At this preliminary stage of the proceeding, we are persuaded by Petitioner's explanations and supporting evidence. We begin our analysis with a brief overview of MIT2002, MIT2003, and MIT2004.

1. *Overview of MIT2002*

MIT2002 relates to a “scalable instruction set architecture [] to attack the emerging wire-delay problem by providing a parallel, software interface to the gate, wire, and pin resources of the chip.” Ex. 1006, 1.³ The “Raw” processor exhibiting such architecture “divides the usable silicon area [of a chip] into 16 identical, programmable tiles.” *Id.* at 2. Each tile contains a processor, static router, two dynamic routers, floating point unit, data cache, and software-managed instruction cache. *Id.* at 2–3. Each tile connects only to its four neighbors, thereby reducing the length of the wires to nearest-neighbor distances, with consequent reduction to wire-delay. *Id.* at 3. The wiring resources of the processor are directly programmed by the software. *Id.* The “Raw” operating system “allows both space and time multiplexing of processes,” and can run multiple independent processes simultaneously and “switch them in and out as on a conventional processor.” *Id.* at 5. The operating system “allocates a rectangular-shaped number of tiles (corresponding to physical threads that can themselves be virtualized)

³ We cite to the numbering of the Exhibit provided by the Petitioner in the bottom right-hand corner of each page; any use of the numbering of the journal pages by either party has been converted to the corresponding page number of the Exhibit.

proportional to the amount of computation that is required by that process.”
Id. In switching, the operating system “finds a contiguous region of tiles that corresponds to the dimension of the process, and resumes the execution of the physical threads.” *Id.*

2. Overview of MIT2003

MIT2003 relates to scalar operand networks, defined as “interconnects optimized for scalar data transport, whether centralized or distributed,” including the algorithms that manage them. Ex. 1007, 1. The described scalar operand networks may be implemented in the “Raw” microprocessor. *Id.* at 2. Scalar operand networks, which can be designed to have short wire lengths, can scale with increasing transistor counts and provide transport for multiple forms of data. *Id.* at 1. Implemented in the “Raw” processor, the scalar operand network “addresses the *bandwidth scalability* challenge by replacing buses with a point-to-point mesh interconnect” that “is programmed to route operands only to those tiles that need them,” through the combination of a “static network, an intelligent compiler, and bypass path integrated network interfaces.” *Id.* at 8. Further, “[f]or each word sent between tiles on the static network, there is a corresponding instruction in the instruction memory of each router that the word will travel through.” *Id.* The static routers “collectively reconfigure the entire communication pattern of the network on a cycle-by-cycle basis.” *Id.*

3. Overview of MIT2004

MIT2004 is an evaluation of the “Raw” microprocessor. Ex. 1008, 1. MIT2004 provides an overview of its architecture, its implementation, results of testing, and analysis of the results. *Id.* at 1–11.

4. *Analysis*

Upon review of Petitioner’s explanations and supporting evidence, and the arguments presented in Patent Owner’s Preliminary Response, we are persuaded that Petitioner has demonstrated sufficiently, for purposes of institution, that the challenged claims would have been obvious over the combined teachings of MIT2002, MIT2003, and MIT2004. Petitioner provides claim mappings of each individual limitation of the challenged claims of the ground to the combined teachings. Pet. 11–34.

a) *Claim 1*

(1) *[1pre]*

The preamble of claim [1pre] recites, “[a] system comprising.” Petitioner contends this claim language, if limiting, is met by teachings of the combination. Petitioner points to MIT2002’s description of a “Raw” microprocessor. Pet. 12 (citing Ex. 1006, 1–2; Ex. 1002 ¶ 49). We agree with Petitioner that the preamble, to the extent limiting, has been shown to be taught by MIT2002.

(2) *[1a]*

Limitation [1a] recites, “a set of processor cores.” Petitioner points to the “Compute resources” of the “Raw” microprocessor described in MIT2002. Pet. 12–13 (citing Ex. 1006, Fig. 1). Petitioner also points to the computer processor in each tile of the processor described in MIT2003. *Id.* at 13 (citing Ex. 1007 § 3; Ex. 1002 ¶ 51). Consequently, we determine that Petitioner has identified teachings in the asserted combination of references for each element of limitation [1a].

(3) [1b], [1c], [1d], [1e]

Limitation [1b] recites, “a set of routers each including a set of input ports and a set of output ports, wherein.” Limitation [1c] recites, “each processor core of the set of processor cores corresponds to a different router of the set of routers, wherein each processor core and corresponding router form a tile.” Limitation [1d] recites, “each processor core is communicatively coupled with a corresponding router via the router’s set of input ports and set of output ports.” Limitation [1e] recites, “each router is communicatively coupled with one or more adjacent routers via the set of input ports and the set of output ports, and wherein each processor core is communicatively coupled with the other processor cores via the set of routers.”

Petitioner asserts that the “Raw” system describes each tile in the set of 16 tiles having a dynamic router, with each router having input and output ports. Pet. 14 (citing Ex. 1006, Fig. 1; Ex. 1008 § 6). Petitioner provides an annotated version of Figure 1 illustrating this assertion, reproduced below:

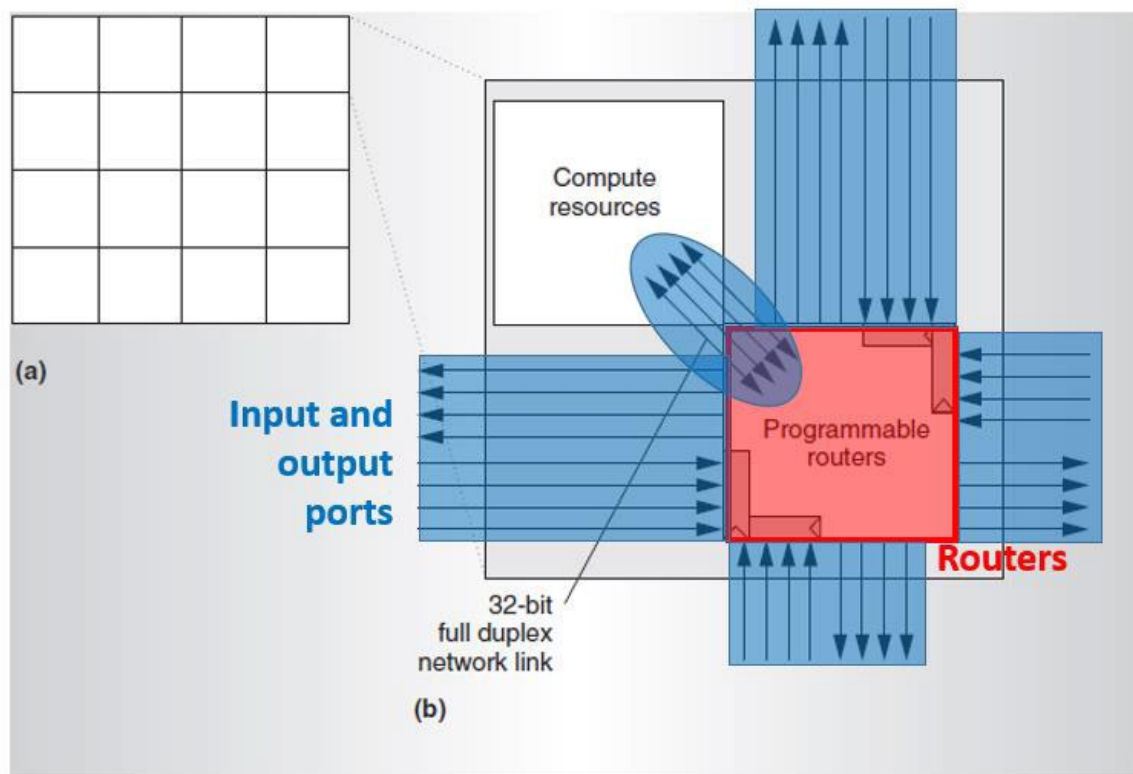


Figure 1. On-chip interconnects in Raw. The Raw microprocessor comprises 16 tiles (a). Each tile (b) has computational resources and four networks, each with eight point-to-point 32-bit buses to neighbor tiles.

Id. Annotated Figure 1, reproduced above, depicts a set of 16 tiles (1a), wherein each tile contains “Compute resources” connected, by arrows described as a “32-bit full duplex network link,” to a square box labeled “Programmable routers,” wherein the routers having multiple double-headed arrows extending from each side of the square, which Petitioner annotates as “Input and output ports.” *Id.* Consequently, we agree with Petitioner that the limitations of [1b]–[1e] are taught by the applied combination of references.

(4) [1f]

Limitation [1f] recites, “each router is operable to receive one or more data packets from the one or more adjacent routers or the processor core

corresponding to the router.” Petitioner points to MIT2002’s description of dynamic routers using data packets, including a destination address and data payload. Pet. 18–19 (citing Ex. 1006, 8 (“To send a message on one of these networks, the user injects a single header word that specifies the destination tile (or I/O port), a user field, and the length of the message. The user then sends up to 31 data words.”))).

With respect to “data packets,” Petitioner asserts that the “Raw” system describes a data packet in the form of an “operand message,” which consists of “the message header with the addressing information, an operand tag, and the operand itself.” *Id.* at 19 (citing Ex. 1007 § 6). Petitioner also points to the “Raw” system’s “low-latency scalar operand network.” *Id.* (citing Ex. 1008 § 7). Relying on the declaration testimony of Dr. Conte, Petitioner asserts that the description of the “Raw” system in the three applied references renders obvious the “data packet” limitation. *Id.* (citing Ex. 1002 ¶ 64). Petitioner asserts that the “Raw” system’s operand messages are data packets “because they are ‘a unit of data sent over a network,’ and they also contain fields including a destination address and other data fields,” which is “the precise way the ’975 patent specification discusses the format of the claimed data packets.” *Id.* at 20 (citing Ex. 1002 ¶ 65; Ex. 1001, 4:50–54; 11:20–22; 14:50–52). Consequently, we determine that Petitioner has identified teachings in the asserted combination of references for each element of limitation [1f].

(5) [1g]

Limitation [1g] recites, “based on a physical destination address of a data packet, each router is operable to send one or more data packets to the one or more adjacent routers or the processor core corresponding to the

router, wherein each router is operable to retain a data packet in the event of a traffic condition.” Petitioner asserts that the “Raw” system uses a single header word to specify a destination tile; i.e., a particular physical tile. *Id.* at 20 (citing Ex. 1002 ¶ 67). Petitioner asserts that the claimed retention of a data packet in the event of a traffic condition is taught by the “Raw” system’s retention of data in the event of “deadlock,” or network retention. *Id.* at 20–21. (citing Ex. 1006, 8–9). Petitioner further points to retention of data by the “Raw” system’s buffering system, and how the system detects “if words have waited too long on the input.” *Id.* at 21 (citing Ex. 1002 ¶ 68).

Patent Owner argues that these descriptions of the “Raw” system do not teach limitation [1g]. Patent Owner argues that the “Raw” system addresses deadlock by draining the network (data) into DRAM (Dynamic Random Access Memory). Prelim. Resp. 12–13 (citing Ex. 1006, 9). Patent Owner argues that Petitioner does not explain how such a draining operation teaches a router retaining a data packet. *Id.* at 13. Patent Owner contrasts this operation with the system of the ’975 patent, in which the packet is retained by buffers, included in each processing core, “corresponding to the router.” *Id.* at 13–14 (citing Ex. 1001, 13:32–35 (“[I]f a router indicates that a particular input port is stalled, the neighboring router will retain a packet that is to be transmitted to that input port.”)).

For purposes of institution, Petitioner has presented sufficient evidence that would support a finding that the router of the “Raw” system is operable to retain a data packet in the event of a traffic condition. Patent Owner clarifies that the claim requires only that a packet be retained by buffers that correspond to the router. Prelim. Resp. 13–14. The ’975 patent supports this reading, describing each router has having corresponding

buffers on its input and output ports, as shown in Figure 4, reproduced below.

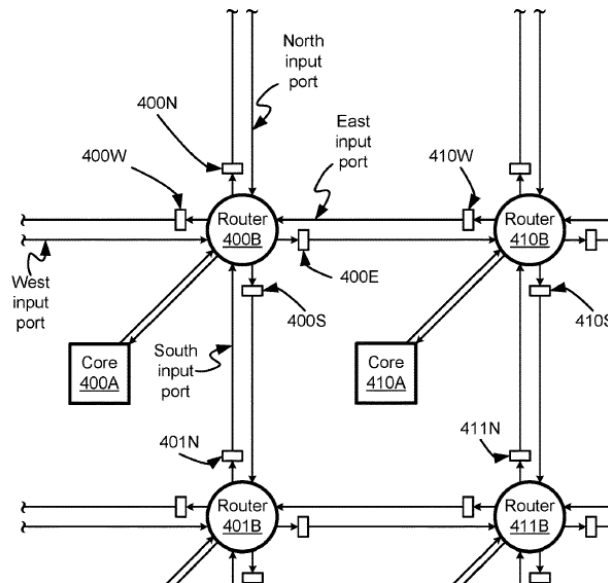


Figure 4, reproduced above, shows router 400B connected through input and output ports to neighboring routers, and to core 400A. Each connection to a neighboring router includes a buffer, such as buffer 400E that sends a packet from router 400B to router 410B. Ex. 1001, 14:25–26. Stalled packets “remain stalled at the buffer.” *Id.* at 12:43–67. Thus, the router need not itself provide storage, but need only have a corresponding buffer (or similar feature) that retains the packet.

Although Petitioner does not point to a specific feature that retains a data packet, Petitioner does point to its declarant’s statement that MIT2002’s buffering system “retains a data packet in the event of a traffic condition.” Pet. 21 (citing Ex. 1002 ¶ 68). The record shows that MIT2002 describes the buffering system of its dynamic networks as having “buffering resources on the path to its destination.” Ex. 1006, 8; *see also* Ex. 1002 ¶ 70. MIT2002 also states that those buffering resources may be “over

commit[ted],” and it expressly assumes “an uncongested network” when providing a calculation for the number of cycles required to reach a destination tile on the dynamic networks. Ex. 1006, 8. MIT2002 further describes that each tile has input and output buffers connected to registers that send data onto the on-chip physical networks, and that each tile has “the option of keeping local copies of transmitted values.” *Id.* at 6–7. In this context, we view MIT2002’s statement that the processor detects “if words have waited too long on the input” as teaching that data packets are retained on the input of the processor. Ex. 1006, 9;⁴ Pet. 21. Because the processor corresponds to a router, the input of the processor also corresponds to the router. Ex. 1006, Fig. 1. Accordingly, the current record appears to support Petitioner’s assertion that data packets are retained by a buffer corresponding to the router; i.e., the router associated with the input of that particular processor.

We emphasize that the current record supports Petitioner’s assertions under the reasonable likelihood standard applicable for institution. The parties will have further opportunity to supplement the record during trial, and the final determination will be made at the conclusion of the trial as to whether Petitioner has proven its proposition of unpatentability by a preponderance of the evidence on the final record.

(6) [1h]

Limitation [1h] recites, “each router implements a static priority routing policy.” Petitioner asserts that MIT2002 describes a static

⁴ We note that Exhibit 1006 footnotes this quotation with a reference not currently of record. Ex. 1006 n.6.

(unchanging) priority routing policy in at least two ways: “1) by pre-allocating buffers that allow particular prioritized threads to communicate even if there is a potential communications deadlock; and 2) by prioritizing ‘trusted clients’ and allowing them to use a faster ‘memory network’ while untrusted clients must use a ‘general network’ that may deadlock.” Pet. 21 (citing Ex. 1006, 8–9).

Patent Owner asserts that a routing policy, as used in the ’975 patent, relates to the forwarding or routing of data packets, not buffer allocation. Prelim. Resp. 17. Patent Owner argues that MIT2002 contains a static router, using pre-determined routes, and two dynamic routers, which “wait for a header word to arrive before routing the data.” *Id.* (citing Ex. 1006, 8). Patent Owner argues that, although the dynamic routers pre-allocate buffers, this does not demonstrate a static priority routing policy based on these buffers. *Id.* at 19. Patent Owner also argues that prioritizing trusted clients on one network over untrusted clients using another network does not relate to a routing policy for routing particular data packets. *Id.* Patent Owner characterizes Petitioner’s expert’s contrary opinion as merely conclusory, and unsupported by factual corroboration. *Id.* at 20.

We find Petitioner’s assertions sufficiently supported by the record. Dr. Conte characterizes the allocation of buffers “on the path to its destination” as pre-allocating buffers to allow particular prioritized threads to communicate. Ex. 1002 ¶ 70. Patent Owner argues that allocation of buffers or prioritization of clients does not relate to a routing policy for particular data packets. However, at this preliminary stage Patent Owner does not focus on the asserted prioritization of threads, which seems (unlike buffers or clients) to relate to a routing policy. Based on the relied-upon

description in MIT2002, and the statements of Dr. Conte, the current record supports Petitioner’s assertion that static routing policies are employed, including by the dynamic routers, to prioritize particular threads and their associated data packets.

Petitioner provides an alternate contention, pointing to the routing policy for static routers in MIT2002 that “knows what route will be performed long before the word arrives.” Pet. 21 (citing Ex. 1006, 7). Petitioner asserts that such policies could also have been used to improve its dynamic routers. *Id.* at 22 (citing Ex. 1002 ¶ 70).

Patent Owner argues that Petitioner has not provided sufficient rationale to use the static routing policy of MIT2002’s static router in MIT2002’s dynamic routers. Prelim. Resp. 20–21. Patent Owner points to differences in the described purposes of MIT2002’s static (“exploiting instruction-level parallelism in scalar codes”) and dynamic (“better suited for long data streams and not scalar transport”) routers. *Id.* at 21–22 (citing Ex. 1006, 7–8).

Patent Owner’s arguments regarding this alternative contention appear to have merit. Petitioner has not persuasively shown why a person of ordinary skill in the art would modify MIT2002’s dynamic routers to have features already present in MIT2002’s static routers. *See Polaris Indus., Inc. v. Arctic Cat, Inc.*, 882 F.3d 1056, 1068 (Fed. Cir. 2018) (faulting the Board for “focus[ing] on what a skilled artisan would have been *able* to do, rather than what a skilled artisan would have been *motivated* to do at the time of the invention”); *Intelligent Bio-Sys., Inc. v. Illumina Cambridge Ltd.*, 821 F.3d 1359, 1367–68 (Fed. Cir. 2016) (“It was [Petitioner’s] burden to demonstrate both that a skilled artisan would have been motivated to

combine the teachings of the prior art references to achieve the claimed invention, and that the skilled artisan would have had a reasonable expectation of success in doing so.” (quotations omitted)). Moreover, Petitioner has not explained how the dynamic routers would be thereby improved, and whether such alteration would retain the advantages of the dynamic routers, which handle “unpredictable” operations. Ex. 1006, 6.

(7) [1i], [1j]

Limitation [1i] recites, “an optimization module configured to.” Limitation [1j] recites, “determine optimal function assignment configurations for groups of tiles.” Petitioner asserts that these limitations are taught by the “Raw” microprocessor, which selects between linear and square configurations of programmable tiles, and assigns operations “in a manner that minimizes congestion and then configure[s] the network routes between these operations.” Pet. 22–23 (citing Ex. 1006, 2, 5, 6, 11; Ex. 1002 ¶ 72). Consequently, we determine that Petitioner has identified teachings in the asserted combination of references for each element of limitations [1i] and [1j].

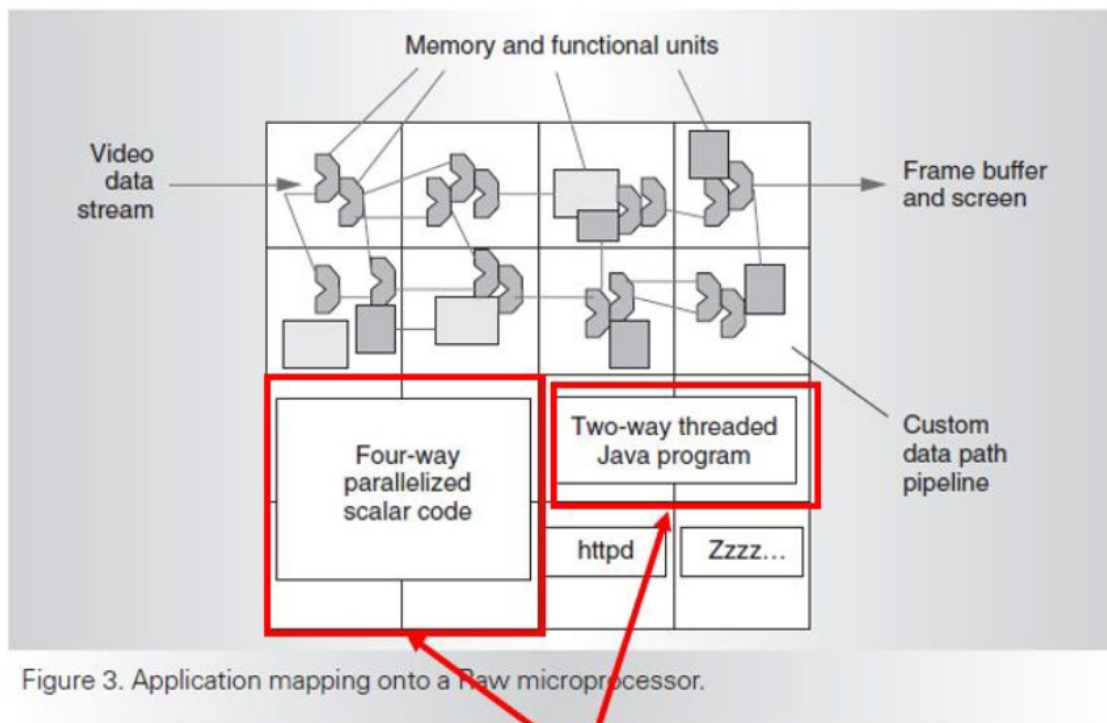
(8) [1k]

Limitation [1k] further recites that the optimization module is configured to

assign two or more functions, which communicate at least unilaterally more frequently with one another than with other functions, to groups of adjacent tiles based on an optimal function assignment configuration determination, wherein the two or more functions are assigned to groups of tiles communicatively coupled in square configurations when the function executes optimally when executed by the groups of tiles communicatively coupled in the square configurations and are assigned to groups of tiles communicatively coupled in linear

configurations when the function executes optimally when executed by the groups of tiles communicatively coupled in the linear configurations

Petitioner asserts that the “Raw” system assigns multiple functions to adjacent tiles, including linear adjacent tiles and square adjacent tiles. Pet. 23–24 (citing Ex. 1006, Fig. 3; Ex. 1002 ¶ 76). Petitioner provides an annotated version of Figure 3, reproduced below:



Two different functions assigned to two different groups of adjacent tiles

Id. at 24. As shown above, MIT2002’s annotated Figure 3 illustrates application mapping onto a “Raw” microprocessor consisting of 16 tiles in a 4-by-4 square array, with some individual tiles including memory and functional units, a two-block linear array containing a “Two-way threaded Java program,” and a four-block square containing a “Four-way parallelized scalar code.” The top eight tiles “work together, parallelizing the

computation” of individual functions, including video input, filtering, buffering, and streaming output. Ex. 1006, 6. Output from one tile can be routed to the input of a neighboring tile “in three cycles,” or order of magnitude less than “state-of-the-art networks hav[ing] latencies of 30 cycles.” *Id.* Petitioner highlights the two-block and four-block arrays, annotating them as “[t]wo different functions assigned to two different groups of adjacent tiles.” Pet. 24.

Petitioner asserts that functions, such as the two-block and four-block functions, are assigned to tiles based on optimization in light of whether “physical threads are likely to communicate with each other” and in order to “minimize[] congestion.” *Id.* at 25 (quoting Ex. 1006, 5–6; Ex. 1002 ¶ 76). Petitioner further asserts that tile configurations are performed by “[t]he operating system allocat[ing] *a rectangular-shaped number of tiles* (corresponding to physical threads . . .) proportional to the amount of computation that is required by that process.” *Id.* (quoting Ex. 1006, 5). Petitioner further points to description that “[w]e assign operations to tiles in a manner that minimizes congestion and then configure the network routes between those operations.” *Id.* at 26 (quoting Ex. 1006, 6; citing Ex. 1002 ¶ 77).

Patent Owner first argues that Petitioner has failed to consider the claim as a whole. Prelim. Resp. 24. Patent Owner argues that Petitioner has neither alleged nor shown that functions assigned to adjacent tiles in Figure 3 of MIT2002 are functions that communicate with each other more frequently than with other functions. *Id.* at 25–26.

We determine that Petitioner has shown, and that the record supports, that MIT2002’s assignment of functions involves determining whether the

functions are likely to communicate with each other, and that such assignment considers which functions are to be placed adjacent to one another. Ex. 1006, 5 (“allocates a rectangular-shaped number of tiles[,] corresponding to physical threads” because “the physical threads of the process are likely to communicate with each other”), 5–6 (the “top eight tiles in Figure 3 illustrate . . . [v]ideo data is streamed in over the pins[and] transformed by a filtering operation” in which “the tiles work together, parallelizing the computation,” and operations are assigned to tiles “in a manner that minimizes congestion and then configure[s] the network routes between these operations”).

Patent Owner next argues that Petitioner does not show that functions are assigned to square or linear adjacent tile arrangements based on optimization. Prelim. Resp. 27–31. Patent Owner describes the relevant support in the ’975 patent:

A function may be assigned to a group of tiles with a particular configuration or arrangement so as to optimize execution of the application and/or the function in accordance with the static priority routing policy. For example, a function may execute optimally when executed by a group of tiles in a square configuration (e.g., like the group of tiles 650). In another example, a function may execute optimally when executed by a group of tiles in a linear configuration (e.g., like the group of tiles 660 or 670).

Id. at 30 (citing Ex. 1001, 17:47–56). Patent Owner argues that MIT2002 merely describes assigning a function to a rectangular-shaped number of tiles, not between a square or linear tile configuration based on optimal execution of the function. *Id.*

We find Petitioner’s assertion that MIT2002 teaches such assignment supported by the record. MIT2002 discloses assigning functions to tiles so

that “tiles work together, parallelizing the computation.” Ex. 1006, 6. Figure 3 shows examples of square (an automatically parallelized scalar code) and linear parallelization (a two-way threaded Java application). The Raw processor “exploit[s] parallelism by automatically partitioning the program graph, placing the operations, and programming the routes for the static router.” *Id.* at 10. Based on those descriptions, MIT2002 supports connecting tiles to reduce latency and to best parallelize the computations, including in the square and linear configurations shown in Figure 3. Because tile placements are taught to advantageously be optimized, we are persuaded, on this record, by Petitioner’s assertion that Figure 3’s square and linear configurations are the result of such optimization.

(9) *Reasons to Combine*

Petitioner argues that it would have been obvious to combine the teachings of MIT2002, MIT2003, and MIT2004. Pet. 11–12. Petitioner asserts that a person having ordinary skill in the art would have naturally been motivated to combine references that describe the same system. *Id.* at 12 (citing *BMW of North Am., LLC v. Stragent, LLC*, IPR2017-01520, Paper 32 at 33 (PTAB Dec. 6, 2018) (Final Written Decision)). Petitioner argues that MIT2003 and MIT2004 further describe the system of MIT2002, rather than a modification of MIT2002, and a person having ordinary skill in the art would seek to apply their teachings together “to obtain a more fulsome view of the technical capabilities of that system.” *Id.* (citing Ex. 1002 ¶ 46).

Patent Owner argues that Petitioner has not provided any reason to combine the “technical details of the Raw microprocessor architecture” with either the “compar[ison of] the Raw microprocessor with other scalar operand networks” of MIT2003 or the “evaluat[ion of] the Raw

microprocessor's specialization, exploitation of parallel resources, management of wires and wire delay, and management of pins" of MIT2004. Prelim. Resp. 32. Patent Owner argues that *BMW* involved documents that were part of a single standard, that would be "naturally consider[ed] . . . together in order to gain a complete understanding of the standard and the capabilities." *Id.* at 33–34.

We agree with Petitioner that one having ordinary skill in the art would have been motivated to look to the interrelated teachings of MIT2002, MIT2003, and MIT2004 because they describe the same system. Patent Owner characterizes MIT2003 and MIT2004 as comparing and evaluating the Raw microprocessor of MIT2002. Prelim. Resp. 32. Consequently, the teachings of MIT2003 and MIT2004 are understood only with reference to the "Raw" microprocessor of MIT2002, and we agree with Petitioner that one having ordinary skill in the art would consider the teachings of all three in combination.

b) *Claim 13*

Petitioner relies upon the same teachings as for claim 1. Petitioner additionally cites to a processor coupled to non-transitory computer readable media, and to reception of the first RAN profile indexing message. *See, e.g.*, Pet. 31–33.

c) *Dependent claims*

With respect to claims 3, 4, 7–9, 11, 12, and 15–19, Petitioner points to specific teachings from MIT2002, MIT2003, and MIT2004 as meeting the additional limitations set forth in those claims. Pet. 26–31, 34.

5. *Conclusion*

Accordingly, based on the current record and for purposes of

institution, we are persuaded that Petitioner has shown a reasonable likelihood that it would prevail on its assertions that the subject matter of claims 1, 3, 4, 7–9, 11–13, and 15–19 would have been obvious over the combination of MIT2002, MIT2003, and MIT2004.

E. Asserted Obviousness of Claims over MIT2002, MIT2003, MIT2004, and Ansari, over MIT2002, MIT2003, MIT2004, and Kaul, or over MIT2002, MIT2003, MIT2004, Ansari, and Kaul

Although we have determined that Petitioner has shown a reasonable likelihood that at least one claim is unpatentable under the asserted obviousness ground over MIT2002, MIT2003, and MIT2004, we address Petitioner’s assertions for the remaining obviousness grounds, based in part, over these three references in furtherance of clarifying the issues for trial.

Petitioner asserts that claims 2, 7, 8, 17, 22, and 23 would have been obvious in view of MIT2002, MIT2003, and MIT2004, in view of the teachings of Ansari for limitation [1h]. Pet. 35–36. Petitioner asserts that it would have been obvious for a person having ordinary skill in the art to use Ansari’s “fixed priority arbiter” to provide a static priority routing policy. Pet. 36. Petitioner characterizes Ansari’s arbiter as choosing between requesters, for example, “a maintenance port may always be lower priority than the functional port.” *Id.* (citing Ex. 1009, 925). Dr. Conte states that Ansari’s arbiter therefore “assigns unchanging relative priorities to input ports of a router.” Ex. 1003 ¶ 117. Petitioner asserts that a person having ordinary skill in the art would have been motivated to implement such an arbiter in the system of MIT2002 because it is “a common arbitration scheme,” and Dr. Conte states that the arbiter “would be able to efficiently manage traffic” of certain types of workflows. Pet. 36; Ex. 1002 ¶ 118.

Petitioner further notes that such a modification would have been “simple to implement,” particularly in MIT2002, which “share[s] a common network-on-chip architecture” with Ansari and “no hardware or minimal hardware modification of the Raw system would be required.” Pet. 36 (citing Ex. 1002 ¶ 118).

Patent Owner argues that Petitioner’s asserted motivation is no more than an assertion that the reference are directed to the same workflows, and this is insufficient. Prelim. Resp. 36–37. Patent Owner further argues that making such modifications would “increase the size of tiles and degrades th[e] high efficiency” of MIT2002’s system “by causing signals to travel a longer distance, which undercut[s] Petitioner’s unsupported contentions on the compatibility of the two systems.” *Id.* at 37.

On this record, we do not agree with Patent Owner’s argument, because there is sufficient evidence in the record that would support that Petitioner’s proposed modification of MIT2002 with Ansari’s arbiter “would be able to efficiently manage traffic.” Pet. 36; Ex. 1002 ¶ 118. Ansari provides an example of maintenance and functional ports that has no direct analogue in MIT2002 (on the current record), and therefore appears to add a desired efficient traffic management functionality to the system described in MIT2002. Consequently, we determine that Petitioner has met its burden in showing a sufficient reason to combine Ansari with the system taught by MIT2002, MIT2003, and MIT2004.

With respect to the remaining obviousness grounds based, in part, on MIT2002, MIT2003, and MIT2004, Petitioner asserts that Kaul provides teachings of the limitations of claims 5 and 6 not otherwise found in the combination of MIT2002, MIT2003, and MIT2004, or the combination of

MIT2002, MIT2003, MIT2004, and Ansari. Pet. 38–42. Patent Owner contests these grounds for the same reasons already discussed above. Prelim. Resp. 37–38. We determine that Petitioner’s assertions are supported by the record, as set forth in our discussion above, and in Petitioner’s discussion of Kaul. Consequently, we determine that Petitioner has met its burden with respect to the asserted obviousness grounds over MIT2002, MIT2003, MIT2004, and Kaul, and over MIT2002, MIT2003, MIT2004, Ansari, and Kaul.

F. Asserted Obviousness of Claims over Olofsson and Kaul, or over Olofsson, Kaul, and Ansari

Although we have determined that Petitioner has shown a reasonable likelihood that at least one claim is unpatentable under the asserted obviousness ground over MIT2002, MIT2003, and MIT2004, we address Petitioner’s assertions based, in part, upon Olofsson in furtherance of clarifying the issues for trial.

1. Overview of Olofsson

Olofsson is directed to a mesh network in the form of a network on a chip having a three-dimensional grid of elements addressable by dedicated x, y, and z coordinates that reflects physical positions of processors in that grid. Ex. 1004, 2:16–22. Each transaction between processors consists of an address and data. *Id.* at 2:23–24. Olofsson’s network employs a selectable routing scheme that can be random, fixed or adaptive. *Id.* at 2:33–34. The routing algorithm determines how data is routed from source to destination, attempting to minimize traffic without adding significant complexity. *Id.* at 1:55–61. Flow control is applied through buffers at each routing node. *Id.* at 1:62–64.

2. *Overview of Kaul*

Kaul is directed to a high bandwidth core to network-on-chip interface. Ex. 1005 ¶ 1. Each physical processor includes two cores, and each core include two hardware threads, so as to permit each processor to execute four software threads concurrently. *Id.* at ¶¶ 17–18. An allocator and renamer block provides for reserving resources, such as reorder buffers, to support out-of-order execution. *Id.* at ¶ 21.

3. *Overview of Ansari*

Ansari is directed to a network on a chip router architecture having an arbiter that detects buffer status of input ports and changes the priorities of the input ports dynamically to enhance router performance. Ex. 1009, 1. An arbiter is used when a number of inputs are requested for the same output port, trapping the source and destination address from the buffer output and generating a control signal so that input data is sent to the output port. *Id.* at 3. Ansari describes a fixed priority arbiter, in which grant is given to the active requester with the highest priority. *Id.* Ansari describes the fixed priority arbiter as being very common when choosing between just a few requesters. *Id.*

4. *Analysis of Claim 1*

a) *[1pre]*

The preamble of claim [1pre] recites, “[a] system comprising.” Petitioner contends this claim language, if limiting, is met by Olofsson’s teachings of a mesh network with a system for routing messages between processors. Pet. 42 (citing Ex. 1004, 2:12–15; Ex. 1002 ¶ 137). We agree with Petitioner that the preamble, to the extent limiting, has been shown to be taught by Olofsson.

b) [1a]–[1e]

Limitation [1a] recites, “a set of processor cores.” Petitioner points to Olofsson’s processing nodes 130 having processing element 140. Pet. 42–43 (citing Ex. 1004, Fig. 1A, 3:16–18, 3:30–33; Ex. 1002 ¶ 139).

Limitation [1b] recites, “a set of routers each including a set of input ports and a set of output ports, wherein.” Limitation [1d] recites, “each processor core is communicatively coupled with a corresponding router via the router's set of input ports and set of output ports.” Limitation [1e] recites, “each router is communicatively coupled with one or more adjacent routers via the set of input ports and the set of output ports, and wherein each processor core is communicatively coupled with the other processor cores via the set of routers.” Petitioner points to Olofsson’s crossbars 110 connected via input and output ports to associated processing node and adjacent crossbars. *Id.* at 43–45, 49–53 (citing Ex. 1004, Figs. 1A, 3, 3:18–29, 4:32–43, 4:50–54, 5:62–67; Ex. 1002 ¶¶ 140–143). Petitioner further points to ports, such as port 220, for each router. *Id.* at 49–50 (citing Ex. 1004, 4:35–37, Fig. 7; Ex. 1002 ¶149).

Limitation [1c] recites, “each processor core of the set of processor cores corresponds to a different router of the set of routers, wherein each processor core and corresponding router form a tile.” Petitioner points to Olofsson’s individual nodes P0–P15 in its mesh grid, each containing processing node and crossbar. *Id.* at 45–46 (citing Ex. 1004, Figs. 1A, 12, 2:61–3:21; 3:30–33, 7:62–8:7; Ex. 1002 ¶¶ 145–146). Petitioner asserts that, because each node contains processor and router, each node forms a “tile.” *Id.* at 48 (citing Ex. 1001, 1:65–67, 4:55–57, 5:48–50; Ex. 1002 ¶ 147).

Consequently, we agree with Petitioner that the limitations of [1a]–[1e] are taught by the applied combination of references.

c) [1f], [1g]

Limitation [1f] recites, “each router is operable to receive one or more data packets from the one or more adjacent routers or the processor core corresponding to the router.” Petitioner points to Olofsson’s “mesh transactions” containing data fields and a destination address, in which the destination address contains coordinates used to route transactions in the mesh network between routers and processors. *Id.* at 53 (citing Ex. 1004, 3:38–42, 3:53–60, 5:41–6:9).

Limitation [1g] recites, “based on a physical destination address of a data packet, each router is operable to send one or more data packets to the one or more adjacent routers or the processor core corresponding to the router, wherein each router is operable to retain a data packet in the event of a traffic condition.” Petitioner asserts that, in the related litigation, Patent Owner construed “destination address” as “an identifier used to determine a unique location within the system.” Pet. 57 (citing Ex. 1014). Petitioner asserts that Olofsson teaches this because “[e]ach processor is addressable through a unique x, y, and z coordinate that reflects the processor’s physical position.” *Id.* (citing Ex. 1004, 2:20–22). Petitioner further asserts that when congestion exists, the router retains data packets in a buffer attached to the router. *Id.* at 57–58 (citing Ex. 1004, 4:43–46, 5:65–67, 6:63–7:1).

Patent Owner argues that Petitioner has not shown Olofsson teaches the recited “data packets” because Olofsson “explicitly teaches that it *does not* use data packets.” Prelim. Resp. 38.

At this preliminary stage, we question whether Olofsson teaches the claimed “router” that, in limitations [1f] and [1g], is operable to receive, send, and retain one or more data packets. Petitioner points to the underlying nature of the data transmitted in Olofsson, asserting that Olofsson’s flat packet structure is a data packet because it is a unit of data sent over a network that contains data and a destination address. Pet. 53–54; Ex. 1004, 2:16–25. Petitioner’s position is supported by its declarant, and by a definition of “packet” as “a block of data transmitted between one location and another within a communication network.” Ex. 1002 ¶¶ 154–156; Ex. 1020.

Patent Owner instead points to Olofsson’s statement that it does not use “packets.” Prelim. Resp. 38 (citing Ex. 1004, 2:16–18 (“mesh routing scheme that *does not use packets.*”)). Patent Owner’s position appears to be further supported by alternate definitions of “packet,” including “information divided into small segments.” Ex. 1020.

Olofsson’s statement that it “does not use packets” is relevant to our inquiry, but is not dispositive. Specifically, in order to determine whether Olofsson teaches a “data packet” as required by claim 1 of the ’975 patent, we must first determine the proper claim construction of this term. Petitioner contends that a “data packet” is “a unit of data sent over a network” (Pet. 8), and at this preliminary stage, Patent Owner does not propose a construction for this claim term (*see* PO Resp. 7–8). During trial, the parties should specifically address the proper construction of the term “data packet” to a person of ordinary skill in the art in view of the intrinsic record of the ’975 patent. As noted above, the parties have identified extrinsic evidence regarding the meaning of this term, namely dictionary

definitions (*see* Exs. 1019, 1020) and Olofsson’s disclosure.⁵ “However, while extrinsic evidence can shed useful light on the relevant art, [the Federal Circuit has] explained that it is less significant than the intrinsic record in determining the legally operative meaning of claim language.” *Phillips*, 415 F.3d at 1317 (internal quotation marks omitted).

Consequently, the issue of whether the claimed “data packet” is taught by Olofsson requires further development during the trial stage.

Petitioner also argues, without supporting citations, that Olofsson uses data packets because it “describes a packet-switched system, not a circuit-switched one.” Pet. 54. Patent Owner argues that this argument is unsupported by evidence, would require packetizing data, and that Olofsson does not necessarily describe a packet-switched system. Prelim. Resp. 39–40. We agree with Patent Owner’s concerns, and do not find Petitioner to have met its burden on this argument.

In the alternative, Petitioner asserts that a person having ordinary skill in the art would have been motivated to use data packets as taught by Kaul in the system of Olofsson. Pet. 54–55. Petitioner asserts that “Kaul describes a ‘source-synchronous hybrid packet-circuit-switched NoC.’” *Id.* at 55 (citing Ex. 1005 ¶ 5). Petitioner further asserts that “Kaul describes how a packet-switched network may be used for communication in a grid-based system where processors are associated with routers, and where the packet data fields may be the precise types of fields that both Olofsson and the ’975 patent contemplate using.” *Id.* (citing Ex. 1005 ¶ 33). Petitioner

⁵ In particular, Olofsson’s use of the term “packet” provides evidence of how this term would have been understood by an ordinarily skilled artisan.

asserts that using “Kaul’s data packets in Olofsson would have reduced implementation cost and allowed re-use of standard networking circuitry and algorithms.” *Id.*

Petitioner acknowledges that Olofsson states that its system “could be problematic with a packet based network-on-a-chip because of ordering.” *Id.* at 56. However, Petitioner asserts that Kaul teaches that “non-packet systems may be easily adapted to packet systems,” and that any needed modification could be performed with off-the-shelf software. *Id.*

Patent Owner argues that the aforementioned statements of Olofsson teach away from use of packets. Prelim. Resp. 41–42. Patent Owner further argues that Petitioner’s motivation of reduced cost and re-use are generic, and insufficient to overcome the express statements in Olofsson that packetized data is “not use[d]” and “could be problematic . . . because of ordering.” *Id.* Patent Owner asserts that Olofsson’s full data messages consisting of full address and data is noted to permit software programs to “directly access other cores by simply addressing that part of the address space.” *Id.* at 41 (citing Ex. 1004, 2:22–25).

We agree with Patent Owner that Olofsson states that it does not use packets, including specific reasons why packets would be detrimental. *See* Ex. 1004, 3:67–4:5 (“The flat packet structure removes the need for packing and unpacking logic at the network interfaces. For nearest neighbor communication, the power and/or latency associated with packing and unpacking data can be significant when compared to the power of the links.”). Olofsson’s disparagement of data packet use would appear to have discouraged one having ordinary skill in the art from using a packet-based system such as Kaul. *See In re Gurley*, 27 F.3d 551, 553 (Fed. Cir. 1994);

Depuy Spine, Inc. v. Medtronic Sofamor Danek, 567 F.3d 1314, 1327 (Fed. Cir. 2009). As for motivation to use Kaul’s teachings, Petitioner has asserted reduced implementation cost and ease of modifying Olofsson with additional packet structure. Pet. 55–56. On balance, we do not find Petitioner to have asserted a sufficient motivation for the proposed combination.

d) [1h]

Limitation [1h] recites, “each router implements a static priority routing policy.” Petitioner asserts that this is taught by Olofsson’s “simple fixed priority arbiter.” *Id.* at 58 (citing Ex. 1004, 4:46–48). We agree with Petitioner that limitation [1h] has been shown to be taught by Olofsson.

e) [1i], [1j]

Limitation [1i] recites, “an optimization module configured to.” Limitation [1j] recites, “determine optimal function assignment configurations for groups of tiles.” Petitioner asserts that Olofsson describes assigning particular functions to particular nodes, i.e., tiles. Pet. 59. (citing Ex. 1004, 7:24–27, 7:62–66). Petitioner asserts that, by comparing and choosing between square and linear configurations for the function assignment, Olofsson’s system maximizes computational or network efficiency. *Id.* (citing Ex. 1004, 7:24–27, 7:30–33, 8:4–7, 8:13–16; Ex. 1002 ¶ 171). We agree with Petitioner that limitations [1i] and [1j] have been shown to be taught by Olofsson.

f) [1k]

Limitation [1k] further recites that the optimization module is configured to

assign two or more functions, which communicate at least unilaterally more frequently with one another than with other functions, to groups of adjacent tiles based on an optimal function assignment configuration determination, wherein the two or more functions are assigned to groups of tiles communicatively coupled in square configurations when the function executes optimally when executed by the groups of tiles communicatively coupled in the square configurations and are assigned to groups of tiles communicatively coupled in linear configurations when the function executes optimally when executed by the groups of tiles communicatively coupled in the linear configurations.

With respect to assigning frequently communicating functions to adjacent tiles, Petitioner asserts that Olofsson teaches this because,

Olofsson's system assigns two or more processing functions to tiles based on the frequency of communication (including unilaterally) between adjacent tiles. Olofsson at 7:24-27 ("reduce communication bottlenecks"); 7:42-45 (functions located based on "nearest neighbor communication"); 7:66-8:2 (functional assignment made such that only adjacent tiles communicate); 8:4-7. Olofsson's sequencer thus co-locates functions "which communicate at least unilaterally more frequently with one another than with other functions."

Pet. 59–60. Petitioner further cites to Figures 10–13, and column 7, line 34–column 8, line 18, as exemplifying such assignment. *Id.* at 60.

With respect to assigning two or more functions to groups of adjacent tiles, Petitioner asserts that "no node needs to directly communicat[e] with any node that is not adjacent, which is a functional assignment to groups of adjacent tiles." *Id.* (citing Ex. 1004, 7:43–45, 7:66–8:2, 8:4–7).

With respect to assigning two or more functions based on an optimal function assignment configuration determination, Petitioner points to Olofsson's goals of "high computational efficiency," "flexibility and ease of

implementation,” and to “optimize utilization of network links.” *Id.* at 61 (citing Ex. 1004, 7:24–33, 8:4–7, 8:13–16).

Patent Owner argues that none of the cited sections, or goals therein, related to assigning functions based on an optimal assignment configuration determination. Prelim. Resp. 44–45. We agree with Patent Owner that the cited sections do not explicitly discuss any configuration determination. On this record, Petitioner’s citation to general goals of efficiency and optimization of network links does not appear to show that functions are assigned to tiles in an optimized manner as required by the claim.

With respect to assigning in either square or linear configurations of tiles based on optimal execution of the function, Petitioner points to Olofsson’s use of linear tile blocks for a Fast Fourier Transform function (Figure 10) and for matrix multiplication (Figure 11). Pet. 61–63. Petitioner also points to Figure 13, which is described by its declarant as a “group[] of tiles communicatively coupled in square configurations when the function executes optimally when executed by the groups of tiles communicatively coupled in the square configurations.” *Id.* at 64 (citing Ex. 1002 ¶ 179).

Patent Owner argues that Petitioner has simply drawn boxes without any support in Olofsson for doing so. Prelim. Resp. 46–47. Patent Owner argues that Olofsson is not shown to select either a linear or square configuration based on an optimal configuration. *Id.*

We agree with Patent Owner that Petitioner has insufficiently explained how its asserted configurations, shown in its annotations of Figures 10 and 11, are reflected in the description of Olofsson as the result of an optimal function assignment. For example, Petitioner asserts that Figure 11 “demonstrates both linear and square assignment of functions”

because “the overall arrangement of rows and columns is square” while exhibiting “a linear workflow proceeding from left to right and from bottom to top.” Pet. 63 (citing Ex. 1002 ¶ 177). Although we credit Dr. Conte’s characterization of process flows in these figures, we are unable to discern whether Petitioner sufficiently shows that Olofsson teaches that these processes are assigned to square configurations when optimally executed in that configuration, and are assigned to linear configurations when optimally executed in that configuration.

5. *Summary*

As detailed above, certain issues relating to the proper construction of claim 1 and to the teachings of Olofsson remain. Because we have found that Petitioner has established a reasonable likelihood that at least one claim is unpatentable under the asserted obviousness ground over MIT2002, MIT2003, and MIT2004, we need not determine whether Petitioner has shown any claim to be unpatentable over the remaining obviousness grounds based, in part, on Olofsson. However, the parties should address these grounds at trial.

III. CONCLUSION

Taking into account the Preliminary Response, we conclude that the information presented in the Petition demonstrates that there is a reasonable likelihood that Petitioner would prevail in challenging at least one of claims 1–19 of the ’975 patent as unpatentable. Given this determination, we institute trial on all challenged claims and all grounds raised in the Petition. 37 C.F.R. § 42.108(a). At this stage of the proceeding, we have not made a final determination with respect to the patentability of any challenged claim or the construction of any claim term.

IV. ORDER

Accordingly, it is

ORDERED that, pursuant to 35 U.S.C. § 314(a) and 37 C.F.R. § 42.4, an *inter partes* review is hereby instituted based on all claims and all grounds raised in the Petition, which follow:

Claims 1, 3, 4, 7–9, 11–13, and 15–19 under 35 U.S.C. § 103 over MIT2002, MIT2003, and MIT2004;

Claims 1–4 and 7–19 under 35 U.S.C. § 103 over MIT2002, MIT2003, MIT2004, and Ansari;

Claims 5–6 under 35 U.S.C. § 103 over MIT2002, MIT2003, MIT2004, and Kaul;

Claims 5–6 under 35 U.S.C. § 103 over MIT2002, MIT2003, MIT2004, Kaul, and Ansari;

Claims 1, 3–9, 11–13, and 15–19 under 35 U.S.C. § 103 over Olofsson and Kaul.

Claims 1–19 under 35 U.S.C. § 103 over Olofsson, Kaul, and Ansari.

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.

IPR2022-00741
Patent 10,355,975 B2

PETITIONER:

James Glass
Nima Hefazi
QUINN EMANUEL URQUHART & SULLIVAN, LLP
jimglass@quinnemanuel.com
nimahefazi@quinnemanuel.com

PATENT OWNER:

David A. Caine
Christopher T. Marando
ARNOLD & PORTER KAYE SCHOLER LLP
david.caine@arnoldporter.com
chris.marando@arnoldporter.com