

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

REALTEK SEMICONDUCTOR CORP.,
Petitioner,

v.

ATI TECHNOLOGIES ULC
Patent Owner.

Case No. IPR2023-00564
U.S. Patent No. 7,742,053

**DECLARATION OF JOHN HART REGARDING
U.S. PATENT NO. 7,742,053**

TABLE OF CONTENTS

	<u>Page</u>
I. INTRODUCTION	1
II. Ground 1 – Stuttard	1
III. Ground 3 – Williams and Whittaker	4
IV. CONCLUSION.....	6

I, John Hart, being of legal age, hereby declare, affirm, and state the following:

I. INTRODUCTION

1. This declaration is made in support of Patent Owner's Sur-Reply. I respond to the technical issues raised in Petitioner's Reply and Dr. Pedram's Reply Declaration, Ex. 1015. Nothing in Petitioner's Reply or Dr. Pedram's Reply Declaration has changed any of my opinions that the challenged claims of the '053 patent are valid and non-obvious.

II. Ground 1 – Stuttard

2. Dr. Pedram asserts that Stuttard's phased operation is contradictory to Stuttard's disclosure that "at any particular moment in time each thread may be assigned to one of any number of tasks." Ex. 1015, ¶10 (citing Ex. 1005, 11:19-25). What Dr. Pedram fails to recognize is that this disclosure relates only to Stuttard's storage of the threads in the thread manager, and not their execution by the array of processing elements. Ex. 1005, 11:16-25. Further, this disclosure regards the cache unit of the thread manager, which is a separate component from the processing elements and processing blocks.

3. There are multiple reasons why Stuttard would store instructions from different threads despite only executing them one at a time, such as preventing downtime. As Stuttard describes, the system will want to switch from one thread to another thread if the current active thread stalls. Ex. 1006, 23:35-25:31. This switch

prevents downtime. As does storing multiple threads at the same time, and processing them via the thread processors, ensures the next thread is ready to start executing when the active thread stalls. Further, Stuttard time-interleaves instructions from different threads for execution. A POSITA would have understood the benefits of doing so. Having more thread processors enables the interleaving of instructions for multiple threads, which also helps avoid stalls.

4. Dr. Pedram opines that each of Stuttard's processing blocks can execute distinct threads independently. Ex. 1015, ¶14. If he is correct (and I don't agree that he is) that each processing block and its binning unit operated independently, each processing block would need to process geometry separately, negating the benefits of parallelism. This is because all of the primitives need to be processed, and those results binned into region tiles, before rasterization can be performed on any tile. Ex. 1005, 41:12-24, 21:15-22:10. Dr. Pedram seems to be under the mistaken understanding that the input primitives can somehow be pre-sorted so that each processing block could magically receive only the input primitives that (after vertex processing) ended up overlapping that processing block's bin. A POSITA would have understood at the time that such an approach was neither disclosed by Stuttard nor would be feasible in the approach taught by Stuttard. Even if geometry processing was somehow divided across different independent blocks, then a triangle in one block may overlap a portion of the screen processed by another

processing block, which had already made it to the pixel processing phase. This would hinder the process of determining which triangles show up on the screen (done via binning across all the processing blocks) and may result in unnecessary processing of triangles which do not show up, a waste of resources.

5. Dr. Pedram provides some examples claiming that pixel processing being more intensive ignores obvious and common real-world considerations. Ex. 1015, ¶11. But the provided examples do not contradict Stuttard's phased operation. First, Dr. Pedram hypothesizes some processing blocks will be idle due to the lack of any pixel command threads. Stuttard provides it resolves this specific issue. First, Stuttard notes that a SIMD system may be inefficient when processing "a large number of relatively small data item groups." Ex. 1005, 2:1-10. Stuttard solves this by dividing the processing elements array into processing blocks, "the processing blocks being operative to process respective groups of data items." Ex. 1005, 1:24-2:23. Since the processing is spread evenly throughout the entire system and Stuttard's processing blocks operate on different groups of data items, the blocks will not be idle. Second, Dr. Pedram provides an example of a video frame with updates in two different regions, one with more intensive vertex processing and one with less intensive vertex processing, claiming this also will render some processing blocks idle. But this is not so, because Stuttard does not process the primitives by region (since the region a primitive belongs in is not determined until *after* vertex

processing). A POSITA would have understood that the entirety of Stuttard's system would process finish vertex processing including any primitives overlapping either area before switching to pixel processing. Thus both geometry phase and rasterization phases are block parallel, and Stuttard's phased operation prevents a load imbalance in one from affecting the performance of the other.

6. Dr. Pedram opines that I am wrong that Stuttard's multiple processing blocks are just for redundancy. Ex. 1015, ¶15. First, that was not my opinion. My opinion is that one of the benefits of having processing blocks is redundancy, I did not say that was the only reason. Ex. 2009, ¶¶79-81. Further, even if Stuttard already discloses additional processing elements (PEs) for redundancy, that does not undermine additional redundancy provided by the division of the array of PEs into processing blocks. Stuttard provides for each panel of 32 PEs, one additional PE is provided. Ex. 1005, 9:30-10:3. This would not resolve a much larger failure where an entire processing block fails. Thus, grouping blocks of PEs into processing blocks also provides additional redundancy in case the failure is greater than a few PEs. Having this multiple redundancy is beneficial to Stuttard's system.

III. Ground 3 – Williams and Whittaker

7. Dr. Pedram opines that by "sorting" and "assigning," Williams' GBS discloses "operable to select a command thread from either of the plurality of pixel command threads and the plurality of vertex commands" required by all claims. Ex.

1015, ¶38. I disagree. Williams' GBS does not disclose the claimed arbiter because it does not choose one or more items from a group of items, which is the plain and ordinary meaning of "select." The "sorting" and "assigning" identified by Dr. Pedram is actually the GBS receiving the next singular graphics command which exits the input buffer and putting it on one of the bus lines. Ex. 1006, 18:26-28. This is not selection.

8. Petitioner argues that the arbiter is not required to "select from" the memory device. Pet., 21. To the contrary, a POSITA would understand the language of the claim requires the arbiter to "select from" the claimed memory device. Specifically, the claims state the arbiter must "select a command thread *from* either of *the* plurality of pixel command threads and *the* plurality of vertex command threads." Ex. 1001, Claims 1, 5. The reference of "the" refers back to the command threads stored by the memory device. Meanwhile, Williams' different assignment schemes, *e.g.*, assigning "according to a type of command" (*e.g.*, geometry, texture) or assigning "based on the availability of a bus, to balance bus load, to even distribute commands, or to unevenly distribute traffic over the busses 116A-116N" do not disclose "select[ing]." Ex. 1006, 12:3-14. Rather, the assignment schemes provide what the GBS does after it receives the graphics command (*i.e.*, if there was selection, what would happen after the selection).

IV. CONCLUSION

9. In signing this declaration, I recognize that it will be filed as evidence in a contested case before the Patent Trial and Appeal Board of the United States Patent and Trademark Office. I also recognize that I may be subject to cross-examination in the case and that cross-examination will take place in the United States. If cross-examination is required of me, I will appear for cross-examination within the United States during the time allotted for cross-examination.

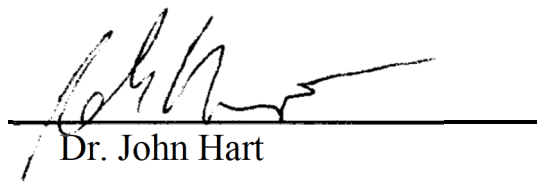
10. The statements and opinions in this declaration comprise and are based upon my review, analysis and understanding of the source materials that I have reviewed as discussed herein, based on the disclosure and statements made within such source materials.

11. I reserve the right to supplement my opinions in the future to respond to any arguments that the Petitioner raises and to take into account new information as it becomes available to me.

12. I declare under penalty of perjury under the laws of the United States of America that the foregoing is true and correct.

May 22, 2024

Respectfully Submitted,



Dr. John Hart