# UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

QUALCOMM INCORPORATED, Petitioner,

v.

DAEDALUS PRIME LLC, Patent Owner.

Case No. IPR2023-00550 U.S. Patent No. 8,775,833

PATENT OWNER'S PRELIMINARY RESPONSE PURSUANT TO 37 C.F.R. § 42.107(A)

# TABLE OF CONTENTS

I.	Intro	Introduction		
II.	Technology Background			
	A.	Prior Art Failures	7	
III.	Over	view of the '833 Patent	9	
IV.	The Petition		11	
	A.	Improper Reliance on Mudge Declaration	11	
	B.	Overview of Finkelstein		
	C.	Overview of Conroy		
	D.	Overview of Nussbaum		
	E.	Overview of Bose	20	
V.	Ground 1 Fails			
	A.	Finkelstein does not render obvious Claim 1(A)	21	
	B.	Finkelstein does not render obvious Claim 1(D.i)		
	C.	Finkelstein does not render obvious Claim 1(D.ii)		
	D.	Finkelstein does not render obvious Claim 2		
	E.	Finkelstein does not render obvious Claim 3	27	
	F.	Finkelstein does not render obvious Claim 4	29	
	G.	Finkelstein does not render obvious Claim 7	29	
VI.	Ground 2 Fails		30	
	A.	Finkelstein would not have been combined with Conroy	30	
	B.	Finkelstein in view of Conroy does not render obvious Claim 7		
VII.	Ground 3 Fails			
	A.	Nussbaum does not render obvious Claim 13(A.i)	34	
	B.	Nussbaum does not render obvious Claim 13(A.ii(4))		
	C.	Nussbaum does not render obvious Claim 15		
	D.	Nussbaum does not render obvious Claim 17		
	E.	Nussbaum does not render obvious Claim 18		
VIII.	Ground 4 Fails			
	A.	Nussbaum would not have been combined with Bose	40	
	B.	Nussbaum in view of Bose does not render obvious Claim 16		
IX	Grou	nd 5 Fails	43	

	A.	Conroy does not render obvious Claim 13(A.ii.4)	43
	B.	Conroy does not render obvious Claim 17	45
	C.	Conroy does not render obvious Claim 18	45
X.	Grou	nd 6 Fails	46
	A.	Conroy would not have been combined with Bose	46
	B.	Conroy in view of Bose does not render obvious claim 16	
XI.	Grou	nd 7 Fails	48
	A.	Conroy would not have been combined with Finkelstein	48
	B.	Conroy in view of Finkelstein does not render obvious Claim 1(a)	50
	C.	Conroy in view of Finkelstein does not render obvious Claim 1(D.i)	
	D.	Conroy in view of Finkelstein does not render obvious Claim 1(D.ii)	
	E.	Conroy in view of Finkelstein does not render obvious Claim 2	
	F.	Conroy in view of Finkelstein does not render obvious Claim 3	55
	G.	Conroy in view of Finkelstein does not render obvious Claim 4	56
	H.	Conroy in view of Finkelstein does not render obvious Claim 7	56
	I.	Conroy in view of Finkelstein does not render obvious Claims	
		13-15 and 17-18	56
XII.	Conc	clusion	57

# TABLE OF AUTHORITIES

	Page(s)
Cases	
KSR Int'l Co. v. Teleflex Inc., 550 U.S. 398 (2007)	29, 37, 43
TQ Delta, LLC v. Cisco Sys., 942 F.3d 1352 (Fed. Cir. 2019)	8
Xerox Corp. v. Bytemark, Inc., IPR2022-00624, Paper No. 12 (Feb. 10, 2023)	8, 9

# TABLE OF EXHIBITS

Exhibit	Description
2001	U.S. Patent Publication No. 2010/0162006
2002	Comparison of Petition to Mudge Declaration
2003	Side-by-Side comparison of Petition to Mudge Declaration
2004	U.S. Patent No. 8,924,758
2005	U.S. Patent No. 6,711,691

### I. Introduction

The '833 Patent revolutionized how power is managed in computing systems by assigning processing components to separate power domains so that a power budget can be dynamically allocated between them. Whereas previously processors either received a uniform power supply across components or a set power supply for each component, the inventors of the '833 Patent devised a system in which a centralized power budget is allocated between separate power domains according to sharing policy values for each of the domains. This allows the power budget to be intelligently allocated to components that need processing power, and away from components that do not.

The prior art disclosed by Qualcomm Incorporated ("Petitioner") fails to disclose or render obvious the claims of the '833 Patent. Thus, the Petition should be denied in its entirety.

# II. Technology Background

Modern computing devices make use of integrated processors and systemson-a-chip ("SoCs") that combine previously-independent circuits within a single

<sup>&</sup>lt;sup>1</sup> As filed, the Petition identified Samsung Electronics Co., Ltd. and Samsung Electronics America, Inc. (collectively "Samsung") along with Qualcomm Incorporated as Petitioners. However, on June 2, 2023, Patent Owner and Petitioner Samsung jointly moved to terminate as to Samsung. Paper No. 10. On June 27, 2023, The Board granted the motion and terminated the proceedings as to Samsung, and noted that Qualcomm Incorporated would remain as Petitioner. Paper No. 11

package. Ex. 1001, 1:11-16. Components of processors and SoCs include central processing units ("CPUs"), graphics processing unit ("GPUs"), cache memories, system memory controller(s), and interconnects. Ex. 1001, 1:60-62.

Processors and SoCs can have multiple CPU "cores." Ex. 1001, 1:13-24. Multiple CPU cores can be used, for example, to provide separate functional units, which can increase the amount of computations that can be performed at one time. *Id.* Processors and SoCs can also have other circuity, such as circuitry specialized for completing graphics tasks, such as a GPU. *See* Ex. 1001, 4:25-34. However, this can result in issues, such as different circuitry consuming differing amounts of power based on the different workloads. Ex. 1001, 1:22-24.

The amount of power consumed by a CPU core or other component of a processor or SoC is impacted by, *inter alia*, the voltage levels supplied to the component, and the frequency at which the component operates. Ex. 1002, ¶39. In general, higher voltages and operating frequencies correspond with higher power consumption and processing capabilities. Ex. 1002, ¶¶39-40. In general, increased power consumption also generally corresponds to increased waste heat. *See* Ex. 1005, [0002]. SoCs using significant power are susceptible to overheating, which can damage them. *See* Ex. 1005, [0002]. Because heat tolerance is sometimes the deciding factor in maximum power consumption (e.g., rather than the power supply), power limits are sometimes called "thermal" limits.

Power management within SoCs is important because processing circuitry is often the largest power consumer in computing systems. In general, power considerations can be a major limitation of the performance, power, form factor, and mobility of a device. For example, a device's power supply may be directly limited by battery capabilities, or indirectly limited by the device's thermal capabilities. *See* Ex. 1005, [0002]. For these reasons, how power is provided to and shared between components of an SoC is of significant importance.

#### A. Prior Art Failures

As set forth in the '833 Patent, prior art approaches to power management did not suitably budget power across multi-domain processors while balancing performance. Ex. 1001, 1:11-26 ("One issue that arises is that the different circuitry can consume differing amounts of power based on their workloads. However, suitable mechanisms to ensure that these different units have sufficient power do not presently exist").

At the time of the '833 Patent's invention, power was managed locally, preventing efficient reallocation of unused power. For example, the prior art taught that system-wide power management was inefficient, counseling against central dynamic power allocation like that later claimed by the '833 Patent:

[A] graphics subsystem might at certain times be unable to utilize additional power resources ... Similarly, the graphics subsystem might use the additional power but not provide a result that is valuable. In either case, it may be impractical to have a system-wide power management device determine the most efficient way to allocate an overall thermal budget.

Ex. 2001 (U.S. Pat. Pub. No. 2010/0162006) at [0002]. Because it was considered "impractical to have [] system-wide power management," prior art approaches were limited to localized power management. For example, prior art systems delivered set amounts of power to SoC components, and power management was limited to considering how much of that power to use. In such systems, localized power management can reduce power used by a component, but cannot intelligently reallocate that unused power.

The prior art also identified difficulties associated with dynamic power allocation. For example, dividing the SoC into separate power domains introduces interactions between components and delays associated with intelligent management:

The on-die interaction of different processors and computational blocks within an integrated circuit such as an SOC may be complex and lead to inefficient communication and resource management. Power

management and performance may each degrade due to these inefficiencies. In some cases, software control may be used for management of on-die resources. .... However, software includes an inherent delay, such as tens of milliseconds, when adjusting operating modes and states of on-die resources.

Ex. 2004 (U.S. Patent No. 8,924,758) at 1:44–53. In short, the prior art had failed to address the power management problem:

[C]onventional power management for multi-processor computer systems has not been efficient. Hence, large amounts of power are consumed by these multiprocessor computer systems even when there is no activity.

Ex. 2005 (U.S. Patent No. 6,711,691) at 1:46–49. As explained in the '833 Patent, "conventional power management control of processing engines, [] simply acts to select one or more engines to be placed into a low power state, but does not provide for the dynamic power sharing of a power budget between domains...." Ex.1001, 10:8-13.

#### III. Overview of the '833 Patent

The '833 Patent claims techniques for managing power consumption within integrated circuits such as those used for processors and SoCs. The '833 Patent

describes a system in which different components, such as CPU cores and a GPU, can be divided into power "domains," each having a dynamically allocated portion of a power budget. Ex. 1001, 1:55-2:31. The '833 Patent's solution improves power management techniques and power efficiency by centrally and dynamically allocating a power budget to processor components. "Embodiments thus may dynamically redistribute power between core domain and graphics domain, enabling flexibility to handle various different workload requirements." Ex. 1001, 10:25-28.

For example, Figure 7 illustrates an embodiment in which a multi-domain processor 400 may include a first domain that includes processor cores 410. Ex. 1001, 9:12-13. Processor 400 may also include a domain that includes a graphics engine 420. Ex. 1001, 9:13-15. In addition, processor 400 may also include a third domain that includes system agent circuitry 450. Ex. 1001, 9:15-25.

The '833 Patent describes that a processor can include logic to dynamically allocate a power budget between the various domains at run time. Ex. 1001, 5:65-6:60. In particular, power management logic can allocate power according to sharing policy values that identify preferences for power allocation. Ex. 1001, 9:48-54. For example, one policy sharing value can indicate an amount of a power budget to supply to the CPU core domain, while a second sharing policy value can indicate an amount for the GPU domain. Ex. 1001, 4:21-56, 5:8-15, 6:31-39. The use of a central power management logic and policy sharing values allows the claimed processors

to shift power budget allocations between domains efficiently. For example, rather than components having set power allocations that are either used or unused, the '833 Patent's power management logic can reallocate unused or inefficiently allocated power (e.g., in one domain) to other components (e.g., in a different domain). Therefore, "embodiments provide a mechanism to dynamically share the power budget between different compute components in the same die. As a result, a power budget or power headroom can be reallocated between cores and graphics engine when they are both integrated on the same die." Ex. 1001, 10:16-18.

The '833 Patent's power allocation can account for not only the policy sharing values, but also the domains' workload. Ex. 1001, 2:11-14; 5:2-12. Additionally, the policy sharing values can be controllable by user-level software "to enable the software to determine how package power budget is shared between different domains." Ex. 1001, 4:1-7.

### **IV.** The Petition

# A. Improper Reliance on Mudge Declaration

The Petition relies extensively on the mere *ipse dixit* of the Mudge Declaration, and in doing so fails as a matter of law. For example, the Petition does not assert that any claim is anticipated, and instead relies on the Mudge Declaration to support otherwise largely unsubstantiated claims regarding what a "POSITA would have understood" and alleged motivations to combine references. *See, e.g.*,

Pet., p. 17 (asserting what a "POSITA would have understood" by citation only to the Mudge Decl.), 32 (same), 34 (same), 36 (same), 49-50 (same), 51 (same), 54 (same), 66 (same), 67 (same), 70 (same), 78 (same); see also id. at 23 ("a POSITA would have recognized"), 39 (same); id. at 28 ("a POSITA seeking to implement ... would have found it obvious"), 32 (same), 36 (same), 40-41 (same), 44 (same), 61 (same), 73 (same), 78 (same). In fact, the Mudge Declaration often provides no substantiation for the alleged understandings of a POSITA. For example, on page 17, the Petition asserts that "[a] POSITA would have understood that each 'power plane' 135 corresponds to a 'domain,'" and cites to the Mudge Declaration "Ex-1002 ¶68," which merely makes the same unsubstantiated claim. See Ex. 1002, ¶68. The same is true for the Petition's assertions on page 32 (citing to Ex. 1002 at ¶104), 36 (citing to Ex. 1002 at ¶118), 49-50 (citing to Ex. 1002 at ¶153) for example.

Reliance on unsubstantiated expert declarations is insufficient to support a finding of obviousness as a matter of law, and has been routinely rejected by the Board. 37 C.F.R. § 42.65(a) ("Expert testimony that does not disclose the underlying facts or data on which the opinion is based is entitled to little or no weight"); *Xerox Corp. v. Bytemark, Inc.*, IPR2022-00624, Paper No. 12 (Feb. 10, 2023) (Director Review affirming "little weight to Petitioner's expert because the expert declaration merely offered conclusory assertions without underlying factual support and repeated, *verbatim*, Petitioner's conclusory arguments."); *see also TQ Delta, LLC v.* 

*Cisco Sys.*, 942 F.3d 1352, 1359 (Fed. Cir. 2019)("This court's opinions have repeatedly recognized that conclusory expert testimony is inadequate to support an obviousness determination on substantial evidence review").

Moreover, the Mudge Declaration fails to provide any probative value to the analysis, as it "does not provide any technical detail, explanation, or statements supporting why the expert determines that the feature in question was required or would have been obvious based on the prior art disclosure," and merely "sets forth Petitioner's conclusory assertions as though they are facts, rather than setting forth facts and evidence in support of Petitioner's assertions." IPR2022-00624, Paper No. 12. For example, Ex. 2002 is a comparison of the Mudge Declaration to the Petition, and shows the insignificant additions (in underlining) made between them.<sup>2</sup> The only substantive contribution made by the Mudge Declaration is the addition of the "Technology Background" section. See Ex. 2002 (Mudge Comparison), pp.25-28; see also Ex. 1002, pp.0025-0028. Where an expert's declaration is merely used as a mouthpiece for a Petitioner's attorney argument, it is entitled to no weight and cannot support a finding of obviousness.

<sup>&</sup>lt;sup>2</sup> Ex. 2002 shows the differences between the two, formatted as additions made to the Mudge Declaration. Ex. 2003 shows a full, side-by-side comparison of the Petition to the Mudge Declaration.

### **B.** Overview of Finkelstein

U.S. Patent Application Publication No. 2010/0115304 to Finkelstein et al. ("Finkelstein") is titled "Power Management For Multiple Processor Cores," and is directed to local power management techniques, for example, wherein "one or more throttling techniques may be utilized locally (e.g., on a per core basis) for one or more processor cores (e.g., in a multiple core processor), for example, that share a single power plane, in response to detection of a thermal event (e.g., detection of excessive temperature at one or more of the cores)." Ex. 1005, [0009].

Finkelstein's focus is in addressing "thermal events," "thermal issues," and "thermal emergencies" by throttling performance. Ex. 1005, [0009] (discussing cores "causing a thermal emergency," and explaining that "throttling techniques may be utilized ... in response to detection of a thermal event"); [0023] ("at an operation 302, it may be determined which (and/or how many) core(s) in a multicore processor are causing thermal issues"); [0019-0023] (discussing throttling as the solution to thermal events).

Finkelstein's approach includes reducing power only for cores that need to be reduced, rather than reducing the power of cores uniformly within a domain. Finkelstein touts that this allows the total system slowdown to be less than it otherwise would be: "[w]e measure the total slowdown of the system as a weighted sum of slowdowns of each one of the processors. For example, if one core out of 4

is slowed down by 20%, the effective frequency factor may be determined to be (3+0.8)/4 or 95%." Ex. 1005, [0017]. Thus, Finkelstein encourages the use of single-core power management (and therefore slowdowns), rather than centralized, domain-level power management.

# **C.** Overview of Conroy

U.S. Patent Application Publication No. 2007/0049133 to Conroy et al. ("Conroy") is titled "Methods and Apparatuses for Dynamic Power Control," and is directed to data processing systems in which "power is redistributed between the sub-systems while tracking [a] load profile." Ex. 1006, [0006].

Conroy describes example load profiles in which the power utilization ratio is used to determine the load profile:

the load profile is identified based on a utilization factor for each of the subsystems. In one embodiment, to determine the utilization factor, the actual power used by each of the subsystems is measured. ... the utilization factor is calculated as a ratio of the actual power used by the subsystem to the power allocated to the subsystem.

Ex. 1006, [0006].

Conroy coordinates the power throttling of certain components, and is based on the history of power usage. For example, Conroy explains a situation in which a history of power usage affects future power usage:

After time instance (911), a high demand task is received. The system becomes busy in processing (905) the high demand task. Since the system had low power consumption before the reception of the high demand task, the power management allows the system to remain in the top performance setting for a short period of time. Thus, the actual power usage (917) increases to above the limit for average power usage. However, in average, the actual power usage is still below the limit.

Ex. 1006, [0135]. Conroy explains: "the subsystems may operate at bursts of substantially high power if there is a considerable low power operation, e.g., a [sic] idle time, during the averaging period." Ex. 1006, [0195]. This is based on the estimation that prior low power consumption means the temperature is low, so temporary substantial power use will not overheat the component. Overall, therefore, Conroy is directed to short term over-performance based on power usage history during "the averaging period." This means Conroy is only suitable for short term changes in power allocation. For example, Conroy states "under the control of the

dynamic throttling system, the system processes the high demand task as fast as possible within the limit of power constraint." Ex. 1006, [0144].

#### D. Overview of Nussbaum

U.S. Patent Application Publication No. 2011/0022356 to Nussbaum et al. ("Nussbaum") is titled "Determining Performance Sensitivities of Computational Units," and is directed to "analysis of workload executed on computational units to help identify those computational units that are more performance sensitive to a change in performance capability caused by, e.g., a frequency change." Ex. 1007, [0005].

Nussbaum explains that "[s]electively distributing power among the active cores or other computational units based on frequency sensitivity allows for greater overall system throughput on heterogeneous workloads or multi-threaded workloads with heterogeneous threads." Ex. 1007, [0019]. Nussbaum is therefore directed to determining the frequency sensitivity of computational units to make use of sensitivity-based power allocation. As Nussbaum states, "[t]he rationale for that is to not waste power by boosting cores whose performance will not be increased." Ex. 1007, [0040].

Nussbaum describes its approach to calculating the performance sensitivity:

the computational units are operated at a first performance
level, and respective first utilization metrics are

determined. The computational units are then operated at a second performance level and respective second utilization metrics are determined. The sensitivity to performance capability change, e.g., a frequency change, is determined.

Ex. 1007, Abstract; [0006].

Nussbaum distinguishes its approach from changing power consumption of CPU cores collectively. For example, Nussbaum identifies approaches that "opportunistically raise the performance level (e.g., raise the frequency) of CPU cores on multi-core processors, when the processor integrated circuit is running below its thermal design point (TDP)":

Such approaches allow the operating frequency of the CPU cores to be raised together when there is estimated power, current or thermal headroom in order to improve performance under a given TDP, and decreased when the operation is exceeding those limits. Such approaches have assumed all active CPU cores operate in their maximum performance state when their frequency is raised in a coordinated fashion.

Ex. 1007, [0017]. Nussbaum then distinguishes those approaches from its own, explaining that its approach not only differs, but "allows for greater overall system throughput":

The above approaches for homogenously increasing power to all cores or to one or more cores based on performance states of the cores, allow for power to be reallocated from idle computational units, such as the CPU or graphical processing unit (GPU), but treat all active units homogeneously when dithering frequency or boosting steady state frequency. However, some active cores or other computational units may be gaining little or no performance increase from a higher core frequency, while other cores or computational units may be running workloads with a higher sensitivity to an increase in core frequency. Selectively distributing power among the active cores or other computational units based on frequency sensitivity allows for greater overall system throughput on heterogeneous workloads or multi-threaded workloads with heterogeneous threads. That requires an

effective approach to identify workload sensitivity to changes in core frequency.

Ex. 1007, [0019]. Nussbaum is therefore directed to a different approach than those that allocate power collectively for example.

#### E. Overview of Bose

U.S. Patent Application Publication No. 2009/0089602 to Bose et al. ("Bose") is titled "Method And System Of Peak Power Enforcement Via Autonomous Token-Based Control And Management," and is directed to software-controlled power management using "tokens" of power allocation.

Bose describes using a token allocation map across connected components, wherein each component is assigned a power budget as determined by a number of allocated tokens. Ex. 1008, [0013]. Bose also includes localized token-based power allocation:

control at a component level to adjust a token count of a component includes monitoring an idle counter, which is preset with a fixed value P, associated with the component, while it is decremented on each cycle of successive inactivity of the component, resetting the component's token counter to zero, donating the unused tokens to

central token storage and transitioning the component to a low power state

Ex. 1008, [0014]. Thus, in Bose, following lower power usage by a component, it is allocated fewer tokens, and therefore less power.

#### V. Ground 1 Fails

The Petition asserts in Ground 1 that Claims 1-5 and 7 are rendered obvious by Finkelstein.

# A. Finkelstein does not render obvious Claim 1(A)

Finkelstein does not render obvious at least "each of the first and second domains to operate at an independent voltage and frequency" as recited in claim 1. For example, Petitioner maps Finkelstein's "power planes 135" with the claimed first and second domains, but fails to demonstrate that the "power planes 135" operate at an independent voltage and frequency. Pet., p.17.

For example, Petitioner asserts that because Finkelstein states that its Logic 140 can "throttle or modify an operational characteristic of one or more of the cores 106," that "[a] POSITA implementing Finkelstein would thus have been motivated to operate Core 1 (in a first domain) and Core 2 (in a second domain) at independent voltages and frequencies." Pet., p.19. However, merely operating at *different* voltages or frequencies does not demonstrate that voltages or frequencies operate *independently* as claimed. Two variables can be intertwined (i.e., not independent)

yet have different values, or vice versa (independent with the same value). Neither the Petition nor the Mudge Declaration provides any support for rewriting the claim language from "operate at an *independent* voltage and frequency" to "operate at a *different* voltage and frequency," as Petitioner's assertions would require. Rather, Petitioner engages in impermissible hindsight, using the '833 Patent's disclosures as justification. For example, Petitioner asserts that "independent operation would have allowed for a greater degree of control over one core, while still allowing for increased performance on the other," just as the '833 Patent explains that "Embodiments thus may dynamically redistribute power ... enabling flexibility to handle various different workload requirements." *See* Pet., p.19; Ex. 1001, 10:25-27.

# B. Finkelstein does not render obvious Claim 1(D.i)

Finkelstein does not render obvious at least a "first logic to dynamically allocate a power budget for the processor between the first and second domains" as recited in claim 1.

First, the Petition does not identify what it maps with the claimed "first logic." Petitioner asserts "a first set of processor cores, e.g., Core 1 (red) in a first power plane, is a 'first domain' and another set of cores, e.g., Core 2 (blue) in a second power plane, is a 'second domain'," but fails to demonstrate that they share a power budget. Pet., p. 18. The mere existence of multiple cores does not demonstrate that

they share a power budget, let alone that there exists a "first logic to dynamically allocate a power budget for the processor between" them as claimed.

Second, Petitioner merely uses impermissible hindsight to assert that the alleged first and second domains share a power budget. Petitioner maps the claimed "power budget" with "a variable  $(W_{ki}E^k - E^i_{high})$ ," but Finkelstein does not disclose or suggest that variable being "dynamically allocate[d] ... between the first and second domains" as required. Pet., p. 23. For example, Petitioner relies on the following theory regarding Finkelstein:

[power] planes may 'share a common package power/energy budget' so that, for example, 'unused processor core power' of a core in the first domain (e.g., Core 1) can be used for 'more Graphics Effect(s) (GFX) performance in a GFX intensive workload' being performed by a core in the second domain (e.g., Core 2)."

Pet., pp.22-23. However, Petitioner provides no support for the assumption that "Core 2," (allegedly the "second domain"), performs the graphics processing described in Finkelstein, or that it can be the destination for any "unused processor core power" of Core 1 (allegedly the "first domain"). Instead, Finkelstein differentiates between processor cores and graphics circuits. Ex. 1005, [0050] (referring to processors 702 and 704 in contrast to "high performance graphics

circuit 734"), FIG. 7 (illustrating them separately, divided by chipset 720). Finkelstein's description at most relates to the graphics circuits, not Core 2.

Moreover, the variable that Petitioner maps to the claimed "power budget" is merely an expression that demonstrates that when the power allocated to plane i  $(W_{ki}E^k)$  is more than the maximum turbo energy that can be used by that plane  $(E^i_{high})$ , the plane does not use all the power allocated to it.

## C. Finkelstein does not render obvious Claim 1(D.ii)

Finkelstein does not render obvious at least the use of "a first sharing policy value for the first domain and a second sharing policy value for the second domain" as recited in claim 1. Petitioner identifies "two complimentary approaches" for the claimed first and second sharing policy values; both fail.

Petitioner's "First Approach" alleges that "vectors  $W_{ki}$  ... [are] used to determine  $E_{ki}$ , which represents the portion of the power budget plane 'i' may obtain." Pet., p. 24. Petitioner expressly states "the power budget itself is a variable  $(W_{ki}E^k - E^i_{high})$ ." Pet., p.23. However, that variable is not a budget that gets allocated according to the alleged sharing policy values (allegedly "vectors  $W_{ki}$ "), but rather is a function of those vectors. For example, Finkelstein explains that "we may have that for a power plane i, its portion  $W_{ki}E^k$  is high ... the 'unused' budget for this power plane [is]  $W_{ki}E^k - E^i_{high}$ ." Ex. 1005, [0035]. The

variable  $(W_{ki}E^k - E^i_{high})$  merely demonstrates that when the power allocated  $(W_{ki}E^k)$  is more than the maximum power that can be used  $(E^i_{high})$ , the plane does not use all the power allocated to it. Thus, Petitioner's "First Approach" does not identify a "power budget" allocated according to the first and second sharing policy values as claimed.

Petitioner's "Second Approach" alleges that "TDP constraints on a first power plane (the power plane that supplies power to the first domain) represents a first sharing policy value and TDP constraints on a second power plane (the power plane that supplies power to the second domain) represents a second sharing policy value." Pet., p.26. However, the equation identified,  $E_{n+1} = \alpha E_n + (TDP_n - P_n)\Delta t_n$  (Finkelstein's equation 1), merely expresses an iterative energy budget calculation. Ex. 1005 (Finkelstein) at [0029]. The equation calculates the energy budget ( $E_{n+1}$ ) for a next step (n+1) based on values associated with the current step (n) and changes thereto: the energy budget ( $E_n$ ), the decay component ( $\alpha$ ), and any change from the course of the step (( $TDP_n - P_n$ ) $\Delta t_n$ ).

The "TDP constraints" that Petitioner maps to the claimed sharing policy values are not used to allocate a power budget, but rather to create one. For example, the "TDP constraints" that Petitioner relies on are the "the Thermal Design Power (TDP) power limit on Step n" in Finkelstein ( $TDP_n$ ). Ex. 1005, [0029]. The TDP

constraints are the power limits, that when not used, create an energy surplus. This is distinct from the claimed first and second sharing policy values, which are used to allocate a power budget for the processor between the first and second domains.

#### D. Finkelstein does not render obvious Claim 2

Finkelstein does not render obvious "the first logic is to determine a portion of the power budget to allocate to the first domain based on the first sharing policy value" as recited in claim 2. Petitioner asserts that " $W_{k1}$ , corresponds to [the first plane's] portion of the power determined by logic 140." Pet., p.26. However, the alleged "power budget," variable  $(W_{ki}E^k - E^i_{high})$ , does not have a portion allocated "according to vector  $W_{ki}$ ," but rather is dependent on the value of  $W_{ki}$ . Nor is any alleged allocation disclosed as being performed by the first logic.

Additionally, Petitioner is internally inconsistent. For claim 1, Petitioner alleges that " $W_{k1}$  represents a first sharing policy value for the first domain," see Pet., p.24, but for claim 2, Petitioner alleges that " $W_{k1}$ , corresponds to [the first plane's] portion of the power," see Pet., p.27. The same variable  $W_{k1}$  cannot be both the claimed "first sharing policy value" and the claimed "portion of the power budget to allocate to the first domain based on the first sharing policy value."

Petitioner does not allege that this claim is rendered obvious under the "second approach" taken for claim 1.

#### E. Finkelstein does not render obvious Claim 3

Petitioner misinterprets the claimed "minimum reservation values." The '833 Patent's "minimum reservation values" "indicate a floor level corresponding to a minimum amount of power budget to be allocated to the given domain." Ex. 1001, 6:7-11, 5:33-38, 5:49-64. The minimum reservation values ensure the domains are allocated sufficient power to reach a capability level, before additional power budget is allocated. For example, if the minimum reservation value for the second domain cannot be met, all power should be allocated to the second domain (e.g., to get as close as possible to the desired capability level). *See* Ex. 1001, 6:14-21. If the minimum reservation value can be met, power supply above that value can be allocated, e.g., to the first domain. *See* Ex. 1001, 6:40-49.

Petitioner's apparent interpretation is that the claimed "minimum reservation values" can be satisfied by any value. For example, Petitioner asserts that merely because Finkelstein discloses operation at different power levels, the lowest level necessarily satisfies the reservation values. For example, Petitioner states that " $P_n^i$  represents the minimum reservation value for power plane i" because " $P_n^i$  is the lowest power state." Pet., p.28. However,  $P_n^i$  is not used by the claimed first logic (which petitioner aligns with Finkelstein's "logic 140," see Pet., p. 27) to allocate the claimed power budget (which Petitioner aligns with Finkelstein's  $W_{ki}E^k$  —

 $E_{high}^{i}$ ). In fact,  $P_{n}^{i}$  is a power consumption operational state, not a reservation value for power allocation.

Finkelstein also does not disclose at least that the first and second minimum reservation values are controllable by user-level software. Petitioner acknowledges this shortcoming, alleging merely that a "POSITA seeking to implement the power management logic of Finkelstein would have found it obvious that 'a software application or a user' may control/provide the value of  $P_n^i$ ." Pet., p.28. However, nothing in Finkelstein (or any other cited art) suggests that P-states or their ranges can be changed by user level software. Being operational states, P-states such as Finkelstein's  $P_n^i$  are foundational to the power management of a processor and affect how and when processors cores enter and exit various power states. Merely because Finkelstein says that "TDP values may be provided by a software application or user" does not mean that the *P-states* could be. TDP values relate to the thermal power consumption limits, which can be affected by e.g., cooling,<sup>3</sup> so rationally should be changeable a user. P-states, by contrast, relate to operational states of cores and their capabilities, and are dependent on a core's internal architecture. P-states are not externally affected, and thus do not need to be user changeable.

<sup>&</sup>lt;sup>3</sup> TDP values are the "Thermal Design Power" limits, which are set because, "heat may damage an IC chip." Ex. 1005, ¶ 2. Adding external cooling mechanisms dissipates heat and increases the power level at which cores can operate without overheating. *See generally* Ex. 1006, [0203] (referring to cooling mechanism capacity).

## F. Finkelstein does not render obvious Claim 4

Additionally, Petitioner alleges that "Finkelstein discloses providing a first minimum reservation value,  $P_n^1$ , to the first power plane," but  $P_n^1$  cannot be both the claimed "minimum reservation value" and the power supplied corresponding to the minimum reservation value. *See* Pet., p.29. Therefore, Petitioner has not identified or explained how Finkelstein allegedly discloses or suggests "provid[ing] at least a first portion of the power budget to the first domain," as required.

### G. Finkelstein does not render obvious Claim 7

Petitioner alleges that Finkelstein discloses that there are situations where "a single power plane may obtain the entire budget." Pet., p.32. However, Finkelstein does not render obvious allocating the claimed "power budget," allegedly the variable  $(W_{ki}E^k - E^i_{high})$ , as claimed (e.g., "allocate substantially all of the power budget to the first domain for a first workload…"). *See* Pet., p.23.

Petitioner asserts that "a POSITA seeking to implement the teachings of Finkelstein would have found it obvious that Finkelstein's power management logic 140 would dynamically allocate substantially all of the power budget to the second domain for a second workload executed after the first workload." Pet., p.32. However, Petitioner does not demonstrate that merely executing a "graphics intensive workload" for example, would result in allocating "substantially all of the power budget," allegedly the variable  $(W_{ki}E^k - E^i_{high})$ , to one domain, as claimed,

let alone that thereafter there would be a switch to allocating "substantially all of the power budget" to another domain, as claimed.

#### VI. Ground 2 Fails

The Petition asserts in Ground 2 that Claim 7 is rendered obvious by Finkelstein in view of Conroy.

# A. Finkelstein would not have been combined with Conroy

Petitioner has not demonstrated that a person of ordinary skill in the art would have been motivated to combine Finkelstein with Conroy. *See* Pet., pp.33-34. For example, Petitioner merely baldly states there would have been a motivation because processors can execute applications:

A POSITA would have understood that Finkelstein's processors may also execute applications that 'present an alternating high workload to the CPU and to the GPU.'

Thus, a POSITA would have been motivated to combine the teachings of Conroy with the teachings of Finkelstein...

Pet., p.34 (citations omitted). However, the mere fact that a processor can execute applications does not provide a motivation to combine references allegedly directed to managing power consumption. For example, neither Finkelstein nor Conroy is related to specific processing of applications such that this would provide a

motivation to combine them. Petitioner does not allege that Conroy provides a solution to a problem identified in Finkelstein. Nor does the proposed combination represent a substitution of any known elements from Finkelstein for an element in Conroy. Nor is the proposed combination the use of a technique disclosed in Conroy to improve Finkelstein. Nor does Petitioner identify any market or other factors creating incentive for a variation of Finkelstein.

In fact, Finkelstein and Conroy would not have been combined at least because they operate at different levels within a system: Finkelstein at the processor, single chip, level; and Conroy at the system level made up of distinct subsystems or components, e.g. discrete integrated circuits ("ICs"). Modifying power management techniques based on system level algorithms for use within a single device (and vice versa) would require a POSITA to undertake significant additional research and development and would lead to unpredictable results. For example, whether the sensors provide the needed information, accurate information, and whether that information is processed correctly, accurately, and effectively under the time constrains required for effective power management would not be straightforwardly known.

Moreover, Finkelstein and Conroy relate to different and generally incompatible approaches to power control. Finkelstein's approach is focused on the use of throttling techniques locally (e.g., on a per core basis) using specific equations

and methods, whereas Conroy's approach is focused on using a load profile in a system comprised of discrete components, e.g., ICs.

Petitioner's allegation that "a POSITA would have been motivated to ... use Finkelstein's processors to execute a 'CPU-heavy load profile' after execution of a 'GPU-heavy load profile,' and vice versa" is founded in hindsight, and assumes the combination. *See* Pet., p.34. However, there was no reason to allow for such backand-forth prioritization; only the '833 Patent claims give justification, and Petitioner does not identify any reasons or explanations otherwise.

Petitioner's allegation that provided a motivation to combine Finkelstein with Conroy squarely conflicts with the assertion that "Finkelstein discloses that the first logic is to dynamically allocate substantially all of the power budget to the first domain for a first workload." Pet., p.31. The same disclosure in Finkelstein cannot both disclose something and provide a motivation for a person of ordinary skill in the art to first seek out an additional reference, then find Conroy, and then combine them to allegedly disclose the very same thing.

Petitioner alleges that there would have been a reasonable expectation of success

because Finkelstein explicitly discloses that its power planes may 'share a common package power/energy budget' so that, for example, 'unused processor core

power' of a first domain can be used for a second domain that may be performing 'intensive workload' in 'Graphics Effect(s) (GFX).'

Pet., p.34. However, Finkelstein's statement relates to power sharing of "unused processor core power," not to the removal of power supply from the processor cores for use in performing graphics operations. A POSITA would not have had a reasonable expectation of success in removing power from the processor (to perform graphics workload) and then having the processing capability to retake that power. For example, as Finkelstein discloses, the power management logic is in the processor: "the processor 102 may further include a power management logic 140." Ex. 1005, [0016]. There cannot a reasonable expectation of success without analysis of how that would have been performed.

Petitioner has also failed to demonstrate that, even if these two references were combined, they would have been combined in the specific way so as to arrive at the claimed invention. *See KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 418 (2007) (requiring "a reason that would have prompted a [POSITA] to combine the elements *in the way* the claimed new invention does"). Petitioner does not explain how or why, in combining Finkelstein with Conroy, the other teachings of Finkelstein would have been unchanged so as to not affect the features allegedly rendering obvious claim 1.

## B. Finkelstein in view of Conroy does not render obvious Claim 7

Petitioner alleges that Finkelstein already discloses the additional limitations of claim 7 itself, which goes against any alleged motivation to combine it with Conroy. Pet., p.31-32. Additionally, Petitioner does not explain the teachings of Finkelstein in view of Conroy in combination, nor does it explain how together they render obvious claim 7.

#### VII. Ground 3 Fails

The Petition asserts in Ground 3 that Claims 13-15 and 17-18 are rendered obvious by Nussbaum.

# A. Nussbaum does not render obvious Claim 13(A.i)

Nussbaum does not render obvious at least "a multicore processor having a first domain including a plurality of cores" as recited in claim 13. Petitioner asserts that "Nussbaum's SOC 100 includes multiple CPU processing cores," and "a plurality of cores ... are equally boosted to a higher F, V." Pet., pp.35-36. However, merely having multiple cores does not render them a "first domain." Nussbaum's central theme causes CPU cores to operate at different voltage and frequency points: its approach is based on "[s]electively distributing power among the active cores or other computational units based on frequency sensitivity." Ex. 1007, [0019]; FIG. 6 (allocating power based on sensitivity information).

Petitioner annotates Tables 1 and 2, but Nussbaum explains that "Table 1 itemizes the power budget allocated" not the operating V/F point. Ex. 1007, [0021]; Table 1, 3 ("Allocated Power"). Some cores will not use allocated power because "client applications mostly utilize 1-2 processor cores ... and only few of them have sufficient parallelism for utilizing all 4 cores." Ex. 1007, [0022]. Nussbaum's power allocations limit the operating V/F points, but do not set them. Nussbaum distinguishes between "Allocated Power" and "operational point (F,V)," and adjusts operational points within allocated power constraints. Ex.1007, [0024]. In Nussbaum, the "power budget [in Table 1 for example] is a limit that defines the highest nominal operational point (F,V) of the core... That allocation, however, is conservative and only a nominal maximum since it assumes simultaneous utilization of all on-die components." Ex. 1007, [0022]. Even operating at the same power level (Wattage) does not require the same V/F point: "the power of a particular computational unit (voltage x current) is based on the frequency of the clock signal, the supply voltage, and the amount of activity." Ex. 1007, [0037].

Petitioner implicitly recognizes Nussbaum's failure to disclose a "first domain including a plurality of cores," asserting that "a POSITA seeking to implement the teachings of Nussbaum would also have found it obvious that multiple cores may operate in the same power domain." Pet., p.36 (citing Ex. 1002 – Finkelstein).

However, that assertion is unsupported and therefore rooted in impermissible hindsight.

## B. Nussbaum does not render obvious Claim 13(A.ii(4))

Nussbaum does not render obvious at least "a power sharing logic to dynamically allocate a variable power budget ... based at least in part on a first power sharing value ...," as recited in claim 13.

Petitioner aligns Nussbaum's "boost sensitivity" with the claimed first and second policy values. Pet., p.43. However, Nussbaum's cores do not share a boost sensitivity, so there is not "a first power sharing value for the first domain" as claimed (e.g., because "first domain includ[es] a plurality of cores"). In Nussbaum, each individual core has its own boost sensitivity: "performance sensitivity of each computational unit to frequency change, and/or other change in performance capability, also referred to herein as boost sensitivity, is determined and stored on a computational unit basis." Ex. 1007, [0025]. Nussbaum criticizes allocating power uniformly across cores: "treat[ing] all active units homogeneously when dithering frequency or boosting steady state frequency" is undesirable because "some active cores or other computational units may be gaining little or no performance increase." Ex. 1007, [0019]. Nussbaum discourages using a collective boost sensitivity for multiple cores, and instead details how to calculate per-core boost sensitivities. Ex. 1007, [0025]-[0031].

Additionally, whereas Petitioner maps Nussbaum's "Thermal Design Point (TDP) power headroom" to the claimed "power budget," *see* Pet. at p.41, the Mudge Declaration concedes that "Nussbaum does not subdivide this 'power headroom.'" Ex. 1002, ¶146. Therefore, the claimed power budget is not allocated between domains at all, let alone according to first and second sharing policy values as claimed.

#### C. Nussbaum does not render obvious Claim 15

Nussbaum does not render obvious at least "the power sharing logic is to dynamically allocate substantially all of the variable power budget to the first domain for a first workload," as recited in claim 15. Nussbaum does not allocate TDP headroom (which Petitioner aligns with the claimed "power budget," *see* Pet. at p.41) on a per-domain basis, but instead does so on a per-computational unit basis (i.e., per-core for the CPU). For example, even the portion of Nussbaum that Petitioner cites details "performance throttling based on boost sensitivity information," (which is itself per-core) such that "a subset of computational units, e.g., processing cores, are identified to throttle" and "the power headroom...is provided in 609 to the computational unit(s) executing" the application. Ex. 1007, [0047].

## D. Nussbaum does not render obvious Claim 17

Nussbaum does not render obvious at least the claimed "minimum reservation" value[s]" recited in Claim 17. Petitioner's attempted mapping of both the Deep Cstate and the P3 state to the "first minimum reservation value" (for the CPU) fails. For example, Petitioner does not explain how Core\_DeepCState\_Pwr (or anything else) serves as the claimed "first minimum reservation value for the first domain." Pet., p.48. In fact, rather than serving as a target threshold power allocation (as in discussion claim 3), the '833 Patent, supra, of Nussbaum's see Core\_DeepCState\_Pwr serves as a power-off value. See Ex. 1007, FIG. 3 ("Operational Point (F,V)" for "Deep Cstate" has "Clocks Off/Power Off," i.e., no frequency and no voltage).

		TABLE 3	
Core Performance States	Operational point (F, V)	Power (dynamic and static) consumed in this point	Remarks
P-boost	F-boost/V- boost	CoreBoostPwr	Boost point. Power budget of the Core has been exceeded
PO	F0/V0	Core_Pwr0	Core Power
P1	F1/V1	Core_Pwr1	Budget
P2	F2/V2	Core_Pwr2	
P3	F3/V3	Core_Pwr3	
Idle	Clocks	Core_Idle_Pwr	
	Off/Low		
	voltage		
Deep Cstate	Clocks	Core_DeepCstate_Pwr	Core is either
	Off/Power Off		power gated or
			deep voltage is applied

Nussbaum explains that the Deep Cstate is not just an "idle" state, it is an even deeper shutdown state. Ex. 1007, [0033]. Petitioner also does not explain how

Core\_DeepCState\_Pwr satisfies "the power sharing logic is to further allocate the first portion of the variable power budget according to a first minimum reservation value for the first domain."

Petitioner next alleges that "a POSITA would have understood that the operational point corresponding to the lowest performance state 'P3' represents a minimum reservation value for the first domain when the cores are utilized to perform their operations." Pet., p.48. However, Petitioner does not explain how any value associated with Nussbaum's P3 state satisfies the requirements that "the power sharing logic is to further allocate the first portion of the variable power budget according to a first minimum reservation value," or show any use of such a value for a power budget allocation. Petitioner's allegations with respect to the "second minimum reservation value" and "GPU\_P3" similarly fail.

#### E. Nussbaum does not render obvious Claim 18

Nussbaum does not render obvious at least the claimed "power sharing logic is to further allocate the variable power budget according to a preference value" recited in Claim 18. For example, Petitioner asserts that "if the output value of step 601 identifies an application as GPU-bounded, then CPU cores may be throttled." Pet., p.49. However, the output value of step 601 is not used to "allocate the variable power budget," but rather to adjust the amount of power budget that is available. For example, as Petitioner asserts, the output value is used to throttle CPU cores and

create more power budget, but Nussbaum's budget does not allocate the power using the output value.

#### VIII. Ground 4 Fails

The Petition asserts in Ground 4 that Claim 16 is rendered obvious by Nussbaum in view of Bose.

#### A. Nussbaum would not have been combined with Bose

Petitioner has not demonstrated that a person of ordinary skill in the art would have been motivated to combine Nussbaum with Bose. *See* Pet., pp. 50-53. For example, Petitioner merely baldly states there would have been a motivation:

A POSITA would have been motivated to implement Nussbaum's power allocation operations with Bose's teaching to increment a boost sensitivity of a particular domain ... when a request for a higher frequency for that particular domain is not granted, especially if that particular domain is operating below an expected performance level, so as to avoid continued denial of that particular domain's request for a higher frequency and to avoid starvation of that domain.

Pet., p.53. However, Petitioner does not identify any teaching or suggestion that would have provided that alleged motivation in the art, and merely relies on

hindsight. Petitioner does not allege that Bose provides a solution to a problem identified in Nussbaum. Nor does the proposed combination represent a substitution of any known elements from Nussbaum for an element in Bose. Nor is the proposed combination an example of use of a technique disclosed in Bose to improve Nussbaum. Nor does Petitioner identify any market or other factors creating incentive for a variation of Nussbaum.

Nussbaum and Bose would not have been combined at least because they relate to different and generally incompatible approaches to power control. For example, Nussbaum is rooted in physical solutions and uses "boost sensitivities" that depend on physical characteristics of computational units, whereas Bose is rooted in software solutions and uses arbitrary "tokens" of power allocation. Ex. 1007, [0029]-[0030]; Ex. 1008, [0044]-[0046]. Nussbaum's "boost sensitivities" are incompatible with Bose's tokenizing power allocation because the sensitivities are calculated (not expected to be whole numbers), and therefore not capable of being assigned a whole number of tokens. *See* Ex., 1008, [0068] ("Each token connotes a pre-specified quantum of power").

Petitioner has also failed to demonstrate that, even if these two references were combined, they would have been combined in the specific way so as to arrive at the claimed invention. *See KSR* (requiring motivation to "combine the elements *in the way* the claimed new invention does"). Petitioner does not explain how or

why, in combining Nussbaum with Bose, the other teachings of Nussbaum would have been unchanged so as to not affect the features allegedly rendering obvious claim 14.

#### B. Nussbaum in view of Bose does not render obvious Claim 16

Nussbaum in view of Bose does not disclose or suggest "increment[ing] the first power sharing value" ever, let alone "when a request for a higher frequency for the first domain is not granted." For example, Petitioner maps Nussbaum's "boost values" to the claimed first and second power sharing values, but as Nussbaum explains, the "boost values" are physical characteristics that are calculated based on the performance level and utilization metrics. *See* Ex. 1007, [0006]. Although Nussbaum describes recalculating the boost sensitives, it does not describe intentionally "increment[ing]" them (i.e., raising the value). *See* Ex. 1001, 4:67-5:1 ("as a given entity's requests for a higher frequency are not granted, the policy values can be raised").

Bose also does not disclose or suggest "increment[ing] the first power sharing value when a request for a higher frequency for the first domain is not granted." For example, as Petitioner acknowledges, Bose explicitly describes the opposite—"Bose explains that if no additional tokens can be allocated to the particular resource at the moment (i.e., a request for a higher frequency is not granted), that particular resource can wait for its token count to be increased in the next iteration of the process." Pet.,

pp.52-53 (citing Bose, annotating Bose FIG. 13). Thus, Bose teaches when a request for a higher token count is not granted, the component must "wait," and cannot increase its token count. Moreover, Petitioner has not explained how, if at all, the token count relates to the claimed "request for a higher frequency" or the claimed "first power sharing value."

#### IX. Ground 5 Fails

The Petition asserts in Ground 5 that Claims 13-15 and 17-18 are rendered obvious by Conroy.

## A. Conroy does not render obvious Claim 13(A.ii.4)

Conroy does not render obvious dynamically allocating a variable power budget "based at least in part on a first power sharing value for the first domain stored in a first storage and a second power sharing value for the second domain stored in a second storage" as claimed. Petitioner does not explain what it alleges serve as the claimed first and second sharing values, instead referring generally to Conroy's load profile tables and the "system power settings" values within them. Pet., pp.59-60.

To the extent Petitioner points to the load profile tables, Petitioner concedes that in Conroy, power is distributed by "tracking the workloads of each of the subsystems," not by sharing values for the domains as claimed. Pet., p.59. In Conroy, the workload determines which load profile table is applied, but the same load profile

table is used for the CPU and GPU (allegedly the first and second domains), and therefore does not serve as both the first and second sharing value. Petitioner quotes that the load profile table may be selected from a plurality, but ignores that once one table is selected, it is used for both the CPU and the GPU.

To the extent Petitioner points to the "system power settings" values within the profile tables, Conroy does not render obvious storing the first and second sharing values in "a first storage" and "a second storage" respectively as claimed. For example, Conroy describes storing the load values within a load profile table. Petitioner provides no explanation for why the load values within a single load profile table would be stored in separate "first" and "second" storages. Instead, Petitioner baldly asserts that Conroy's figures "depict power distribution tables that store power settings for various system components. Thus, a POSITA seeking to implement the teachings of Conroy would have found it obvious to use registers." Pet., p.61. Petitioner and the Mudge Declaration provide no justification or explanation for this assertion, nor how it relates to storing the values within the tables in separate registers. In fact, Conroy states that the use of its power distribution tables is beneficial because it is simplistic, therefore counseling away from introducing complexity like separate storage of values. Ex. 1006, [0203] (referring to management "without using complicated software").

## B. Conroy does not render obvious Claim 17

Conroy does not render obvious "allocate the first portion of the variable power budget according to a first minimum reservation value for the first domain and allocate the second portion of the variable power budget according to a second minimum reservation value for the second domain" as claimed.

As explained above, the '833 Patent's minimum reservation values ensure the domains are allocated sufficient power to reach a capability level, before additional power budget is allocated. Petitioner alleges that "even when a subsystem has the lowest power allocation utilization of 0, Conroy's power manager still allocates power to that subsystem," but does not identify an alleged minimum reservation value. In fact, as there is no power allocated by the first logic, the power allocated is, as Petitioner notes, "0."

Petitioner also refers to "the 'minimum possible power consumption'." Pet., p.65. However, evaluating subsystems' power consumption does not mean that every subsystem has non-zero power consumption, let alone a dedicated minimum reservation value.

# C. Conroy does not render obvious Claim 18

Petitioner asserts that "a "GPU-heavy" load profile, such as K3, represents a preference value to favor the second domain," but does not explain how the load

profile can serve as the claimed "preference value" and the claimed first and second sharing policy values. *See* Pet., p.66.

#### X. Ground 6 Fails

The Petition asserts in Ground 6 that Claim 16 is rendered obvious by Conroy in view of Bose.

## A. Conroy would not have been combined with Bose

Petitioner fails to demonstrate that a POSITA would have been motivated to combine Conroy with Bose. The entire alleged motivation is a single sentence asserting that "A POSITA would have been motivated to implement Conroy's power manager with Bose's teaching to not grant a subsystem's request to increase the system's setting immediately and, instead, wait till the next iteration to do so in order to keep the power supplied to the various subsystems stable within each integration." Pet., pp.67-68. Neither the Petition nor the Mudge Declaration provide any explanation or justification for this assertion. Nor does Petitioner explain how any of the other teachings of the references affect the analysis. For example, Petitioner does not explain why a POSITA would have combined the power allocation tables of Conroy with the token-based allocation of Bose. Petitioner does not allege that Bose provides a solution to a problem identified in Conroy. Nor does the proposed combination represent a substitution of any known elements from Conroy for an element in Bose. Nor is the proposed combination an example of use of a technique disclosed in Bose to improve Conroy. Nor does Petitioner identify any market or other factors creating incentive for a variation of Conroy.

Moreover, Conroy conflicts with Bose. For example, whereas Conroy describes a system enacting a burst of high power following a period of low power, Bose describes the opposite—following low power usage, token count is decremented. *Compare* Ex. 1006,[0135] *with* Ex. 1008, [0014].

Petitioner has also failed to demonstrate that, even if these two references were combined, they would have been combined in the specific way so as to arrive at the claimed invention. *See KSR* (requiring motivation to "combine the elements *in the way* the claimed new invention does"). Petitioner does not explain how or why, in combining Conroy with Bose, the other teachings of Conroy would have been unchanged so as to not affect the features allegedly rendering obvious claim 13.

# B. Conroy in view of Bose does not render obvious claim 16

Conroy in view of Bose does not render obvious "increment the first power sharing value when a request for a higher frequency for the first domain is not granted" as claimed. For example, Petitioner asserts that Conroy discloses moving within a load profile table if the "performance of the system requires an increase in *total power*." Pet., p.66. In addition to failing to articulate whether the load tables or the power setting values within them serve as the claimed first and second sharing

values (as discussed above with respect to Conroy in Ground 5), Petitioner fails to explain how moving within the table satisfies the claim requirements. For example, Petitioner does not explain how increasing the "total power" of the system relates to "increment[ing] the first power sharing value."

Nor does Petitioner explain how Conroy allegedly teaches incrementing the first sharing value "when a request for a higher frequency for the first domain is not granted." For example, Petitioner asserts that the selected throttle setting is used until the next iteration of throttling, but does not dispute that the next iteration grants the request.

The same is true for Petitioner's arguments regarding Bose – Petitioner does not dispute that Bose grants the request on the next iteration of the allocation process. Nor does Petitioner explain how Bose allegedly renders obvious incrementing the token count to the first domain following a denied request for higher frequency, let alone how this relates to the claimed "first sharing value."

#### XI. Ground 7 Fails

The Petition asserts in Ground 7 that Claims 1-5, 7, 13-15, and 17-18 are rendered obvious by Conroy in view of Finkelstein.

# A. Conroy would not have been combined with Finkelstein

Petitioner fails to demonstrate that a POSITA would have been motivated to combine Finkelstein with Conroy at least because they operate at different levels

within a system as explained above in Ground 2 (Finkelstein is at the processor, single chip, level; Conroy is at the system level).

Petitioner alleges that a POSITA would have been motivated to combine Finkelstein with Conroy because "a POSITA would have been motivated to implement Conroy's system on a chip, as in Finkelstein." Pet., p.69. However, Petitioner provides no justification for this assertion beyond the Mudge Declaration's *ipse dixit*, particularly given Conroy's statement that it "is not intended to represent any particular architecture or manner of interconnecting the components as such details are not germane to the present invention." Ex. 1006, [0186].

Even a motivation "to implement Conroy's system on a chip" would not be not specific to Finkelstein. There is nothing in Finkelstein that makes it possible to implement processing systems on a single chip, as its disclosure is not directed to ways to do so. Instead, Finkelstein assumes that "processor cores 106 may be implemented on a single integrated circuit (IC) chip," and its disclosure does not contribute to how to implement that. Ex. 1005, [0011].

Even modifying Conroy in view of Finkelstein would not lead a POSITA to the claimed invention. For example, modifying Conroy with Finkelstein would merely add another (Finkelstein-like) component within Conroy's system, but would not change the overall power management characteristics of Conroy. Because Conroy is agnostic as to how each component/subsystem is internally power managed, there would be no need to alter it to add a Finkelstein-like subsystem.

Petitioner also does not explain why there would have been a reasonable expectation of success in combining Conroy with Finkelstein. For example, Petitioner (and the Mudge Declaration) merely assert without explanation that "each of these solutions involves routine hardware and software functionalities," but does not explain how the specific hardware and software designs described in the references would interact. For example, Petitioner does not explain how a POSITA would have combined the features of Conroy with Finkelstein, or how such a combination would change the other disclosures of the references.

## B. Conroy in view of Finkelstein does not render obvious Claim 1(a)

Conroy in view of Finkelstein does not render obvious at least "each of the first and second domains to operate at an independent voltage and frequency" as claimed. For example, Petitioner alleges that Conroy "subsystem A" and "subsystem B" serve as the claimed first and second domains, but Conroy does not disclose or suggest that they operate at independent V/f points. Pet., p.71.

Despite Conroy's statement that "components have independent throttle settings," Conroy's throttle settings are merely variable, not separately controlled. In fact, Conroy discloses throttle settings as intertwined (not independent):

As shown in FIGS. 21A-21C, tables 2100, 2110, and 2120 are built such a way that if a system moves from one system setting, e.g. from setting 1, to another system setting, e.g., setting 2, the power needed to operate each of the subsystems increase or decrease at the same time.

Ex. 1006, [0203]. Additionally, Petitioner's reliance on Conroy's tables in FIG. 21 is misplaced because they illustrate "Allocated Power," not "operat[ing]" power as claimed. In fact, Conroy distinguishes between power "allocation" and power "consumption" (*i.e.*, operating power). For example, Conroy describes a "utilization factor" that compares the power allocated to the power consumed. Ex. 1006, [0199].

Petitioner implicitly recognizes this shortcoming, stating that "a POSITA seeking to implement the teachings of Conroy would have found it obvious to operate the CPU (in a first domain) and the GPU (in a second domain) at independent voltages and frequencies." Pet., p.73. However, such an unsupported assertion (other than the Mudge Declaration *ipse dixit*) is entitled to "little or no weight." *See* 37 C.F.R. § 42.65(a).

Petitioner does not allege any disclosures of Finkelstein would be used in a combination with Conroy to render obvious this limitation in Ground 7.

# C. Conroy in view of Finkelstein does not render obvious Claim 1(D.i)

Conroy does not render obvious "dynamically allocate[ing] a power budget ... at run time" as claimed. For example, in Conroy, the power distribution table is selected, and then the power setting values within the distribution table are used. Conroy states that "when performance of the system transitions between system power settings 1-3 within the same power distribution table, the load profile of the system does not change." Ex. 1006, [0203]. Thus, Petitioner's assertion that Conroy discloses "the power may be 'redistributed in an asymmetric fashion, tracking the workloads of each of the subsystems" does not explain how the allocation is done at "run time."

Petitioner also does not identify any alleged "power budget," or "first logic" as claimed.

Petitioner does not allege any disclosures of Finkelstein would be used in a combination with Conroy to render obvious this limitation in Ground 7.

# D. Conroy in view of Finkelstein does not render obvious Claim 1(D.ii)

Conroy in view of Finkelstein does not render obvious at least "the first and second sharing policy values controllable by user-level software" as claimed. Petitioner alleges that Finkelstein discloses that the claimed first and second sharing

policy values would be controllable by user-level software, but does not evaluate or explain how the two references would be combined.

For example, Conroy states that the use of its power distribution tables is beneficial because it is simplistic, and thus counsels against introducing complexity like user-level software control:

By just following a power usage pattern, the system's character can be dynamically changed with a substantially high accuracy, so as to be a balanced system, a CPU-heavy machine, a GPU-heavy system, or any other subsystem-heavy system, without using complicated software. As a result, efficient, dynamic power management for the system is provided, where portions of the hardware (processors, buses, memories, and other subsystems) may have their performance, and, as a side effect, their power, adjusted fairly transparently to the software.

Ex. 1006, [0203]. In fact, Conroy warns that when the power control is executed on the processor (as making power distribution tables controllable by user level software would require), complications arise:

[when] execut[ing] the dynamic budgeting in the main processor... it is difficult to make an arrangement such

that dynamic power management operates in all situations, including when the software in the main processor fails or is replaced with some other software which has no knowledge of the power management algorithm. Further, when the computer system is in an otherwise idling state, the periodic power management task may prevent the system from entering a low power state, or may periodically wake the system from the low power state. When the ability to load throttle settings is reserved to the main processor of the computer system and the dynamic power throttle determination is not performed in the main processor, making the computed throttle setting the current throttle setting may become complicated. The throttle settings need to be communicated to the main processor, and in some situations, it may be necessary to implement fail-safe mechanisms to deal with the (unlikely) case that the Software running in the main processor ignores the command to load the throttles.

Ex. 1006, [0148]-[0149].

Conroy's statement that "the settings can be arranged 'manually' by 'the designer of the system" has no bearing on the reasonable expectation of success. *See* Pet., p.79. The "designer of the system" refers to entities like manufacturers that make processing systems who control firmware used, not users.

Petitioner's assertion that "a POSITA would have had a reasonable expectation of success because Finkelstein already teaches such implementations with user-level control of power settings" is circular and does not demonstrate a reasonable expectation of success in combination with Conroy. Pet., p.79. Saying that a person of ordinary skill would have had a reasonable expectation of success in combining two disclosures because one of them *is a disclosure* ignores the specifics of how they would have to be combined.

# E. Conroy in view of Finkelstein does not render obvious Claim 2

Petitioner does not allege that this claim is rendered obvious for any specific reasons. In fact, Petitioner does not explain in any way how the additional requirements of claim 2 are allegedly rendered obvious.

# F. Conroy in view of Finkelstein does not render obvious Claim 3

As discussed above with respect to Ground 3, in claim 17, Conroy does not render obvious allocating a power budget according to first and second minimum reservation values. Additionally, because claim 17 does not recite that the minimum reservation values are controllable by user-level software, Petitioner nowhere asserts

that Conroy renders obvious that "the first and second minimum reservation values controllable by the user-level software" as recited in claim 3. And as discussed above with respect to Ground 1, in claim 3, Finkelstein does not render obvious allocating a power budget according to first and second minimum reservation values.

Nor does Petitioner assert that the combination of Conroy with Finkelstein renders obvious more than their individual disclosures.

## G. Conroy in view of Finkelstein does not render obvious Claim 4

Petitioner does not allege that this claim is rendered obvious for any specific reasons. Instead, Petitioner merely refers back to discussion of other claims in Ground 7, which have different claim language and therefore different claim scope.

# H. Conroy in view of Finkelstein does not render obvious Claim 7

Petitioner does not allege that this claim is rendered obvious for any specific reasons. Instead, Petitioner merely refers back to discussion of claim 7 in Ground 5, which asserts Conroy not in combination with Finkelstein.

## I. Conroy in view of Finkelstein does not render obvious Claims 13-15 and 17-18

Petitioner does not allege that any of these claims is rendered obvious for any specific reasons. Instead, Petitioner merely refers back to discussion of these claims in Ground 5, which asserts Conroy not in combination with Finkelstein.

## XII. Conclusion

For at least the foregoing reasons, the Petition does not meet its burden of demonstrating a reasonable likelihood that Petitioner would prevail with respect to any of the challenged claims. Thus, Patent Owner respectfully requests that the Board deny institution.

Dated: July 17, 2023 Respectfully submitted,

## /Peter F. Snell/

Peter F. Snell (Reg. No. 52,235) MINTZ, LEVIN, COHN, FERRIS, GLOVSKY AND POPEO, P.C. 919 Third Avenue New York, New York 10022

Telephone: 212-935-3000 Facsimile: 212-983-3115

E-mails:

PFSnell@mintz.com

# **CERTIFICATE OF WORD COUNT**

The undersigned certifies pursuant to 37 C.F.R. § 42.24(d) that the foregoing Patent Owner Preliminary Response, excluding any table of contents, table of authorities, certificates of service or word count, or appendix of exhibits or claim listing, contains 11,881 words according to the word-processing program used to prepare this paper (Microsoft Word). Patent Owner certifies that this Patent Owner Preliminary Response does not exceed the applicable type-volume limit of 37 C.F.R. § 42.24(a).

Dated: July 17, 2023 /Peter F. Snell/

Peter F. Snell (Reg. No. 52,235)

## **CERTIFICATE OF SERVICE**

I certify that a copy of Patent Owner's Preliminary Response is being served

by electronic mail on the following counsel of record:

#### **Lead Counsel**

William M. Fink, (Reg. No. 72,332) O'MELVENY & MYERS LLP. 1625 Eye Street, NW Washington, DC 20006 Telephone: (202) 383-5300

Facsimile: (202) 383-5414 Email: tfink@omm.com

## **Backup Counsel**

Daniel Leventhal (Reg. No. 59,576) Richard Zembek (Reg. No. 43,306) Darren Smith (Reg No. 64,261) NORTON ROSE FULBRIGHT

NORTON ROSE FULBRIGHT 1301 McKinney, Suite 5100

Houston, TX 77010

Telephone: (713) 651-5151 Facsimile: (713) 651-5246

Emails:

daniel.leventhal@nortonrosefulbright.com richard.zembek@nortonrosefulbright.com Darren.smith@nortonrosefulbright.com

Eagle H. Robinson (Reg No. 61,361)

NORTON ROSE FULBRIGHT

98 San Jacinto Blvd., Suite 1100

Austin, TX 78701

Telephone: (512) 474-5201 Facsimile: (512) 536-4598

Email:

eagle.robinson@nortonrosefulbright.com

Counsel for Petitioner Qualcomm Incorporated

Dated: July 17, 2023 /Peter F. Snell/

Peter F. Snell (Reg. No. 52,235)