The impact of Moore's Law and loss of Dennard scaling: Are DSP SoCs an energy efficient alternative to x86 SoCs?

The impact of Moore's Law and loss of Dennard scaling: Are DSP SoCs an energy efficient alternative to x86 SoCs?

L Johnsson[1], **G Netzer**[2]

[1] Department of Computer Science, University of Houston, Houston, TX 77204, USA
[2] PDC, Royal Institute of Technology, SE 100 40 Stockholm, Sweden

E-Mail: johnsson@cs.uh.edu

Abstract Moore's law, the doubling of transistors per unit area for each CMOS technology generation, is expected to continue throughout the decade, while Dennard voltage scaling resulting in constant power per unit area stopped about a decade ago. The semiconductor industry's response to the loss of Dennard scaling and the consequent challenges in managing power distribution and dissipation has been leveled off clock rates, a die performance gain reduced from about a factor of 2.8 to 1.4 per technology generation, and multi-core processor dies with increased cache sizes. Increased caches sizes offers performance benefits for many applications as well as energy savings. Accessing data in cache is considerably more energy efficient than main memory accesses. Further, caches consume less power than a corresponding amount of functional logic. As feature sizes continue to be scaled down an increasing fraction of the die must be "underutilized" or "dark" due to power constraints. With power being a prime design constraint there is a concerted effort to find significantly more energy efficient chip architectures than dominant in servers today, with chips potentially incorporating several types of cores to cover a range of applications, or different functions in an application, as is already common for the mobile processor market. Digital Signal Processors (DSPs), largely targeting the embedded and mobile processor markets, typically have been designed for a power consumption of 10% or less of a typical x86 CPU, yet with much more than 10% of the floating-point capability of the same technology generation x86 CPUs. Thus, DSPs could potentially offer an energy efficient alternative to x86 CPUs. Here we report an assessment of the Texas Instruments TMS320C6678 DSP in regards to its energy efficiency for two common HPC benchmarks: STREAM (memory system benchmark) and HPL (CPU benchmark)

1. Introduction.

Energy efficiency has become a primary concern in all stages of computer systems design and operation, from solid-state device design, see, e.g. [1,2,3,4] to data-center operations, see e.g. [5,6,7,8,9,10,11,12], in regards to cost, environmental impact, and sustainability of energy supply. With an annual growth rate estimated at 6.6% for ICT (Information and Computing Technologies) [13] and an estimated 4.7% share of the total electric energy consumption in 2012, by 2040 ICT would consume 30% of todays' total energy consumption. Even assuming a continued growth rate of total electric energy consumption at the current level, about 3%/year [14], a substantial portion of the total electric energy consumption would be consumed by ICT. These concerns have engaged the semiconductor industry, the platform industry as well as large consumers of computing, such as the

internet companies, in finding more energy efficient solutions to computing. Facebook [15,16,17], Google [18], and Microsoft [9,10,11,19] as well as others have opted to design their own platforms, data centers [6,7,20,21,22], and software systems [23,24]. The CMOS (Complementary Metal-Oxide Semiconductor) processor design challenges in light of continued Moore's law scaling [25] without Dennard scaling [26] are illustrated well in [27] that also describes how the exponential growth in transistors per die have been used historically and the resulting decrease in energy efficiency. The impact of the energy challenges on computer architecture and System-on-Chip (SoC) design has also been highlighted in, e.g. [28,29,30], whereas [31] highlights the impact of the energy efficiency challenges on software.

Dennard scaling [26] through which the electric field of a transistor remains constant while reducing the physical dimension, allowed for a doubling of the number of transistors per unit area for each CMOS technology generation while keeping the power dissipation per unit area constant. Further, the combination of the scaling of physical dimensions and voltage the switching speed was also reduced so the performance per unit area increased by about a factor of about 2.8 for each technology generation. However, by about 2004 the feature sizes for CMOS had scaled down to a size where leakage power reached a level where further reduction in voltage was no longer feasible, Figures 1 and 2, and heat density reached unsustainable levels, Figure 3. The industry's response was multi-core processors, with the IBM Power4 in 2001 [34] being the first followed by dual-core x86 based CPUs by AMD in 2004 [35] and Intel in 2005 [36]. The change in CMOS processor characteristics is well captured in Figure 4.

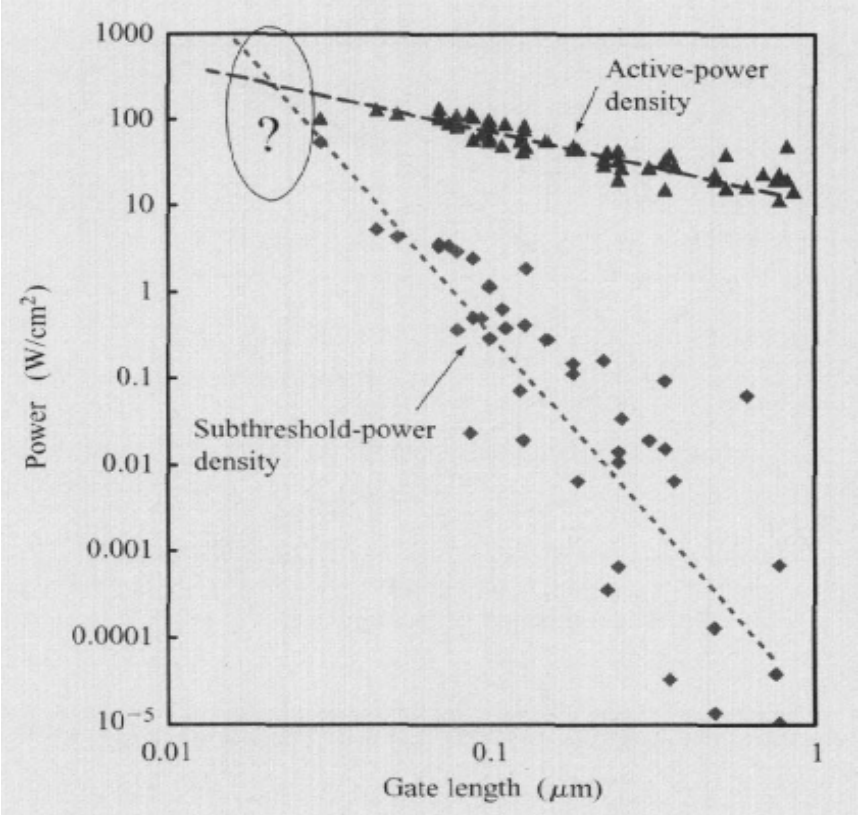

Figure 1. Dennard scaling of active power and leakage [32].

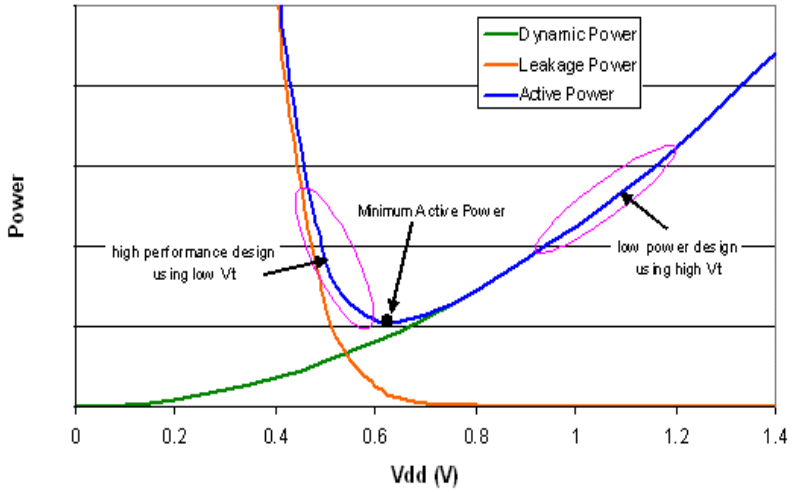

Figure 2. Power consumption as a function of threshold and supply voltage [33].

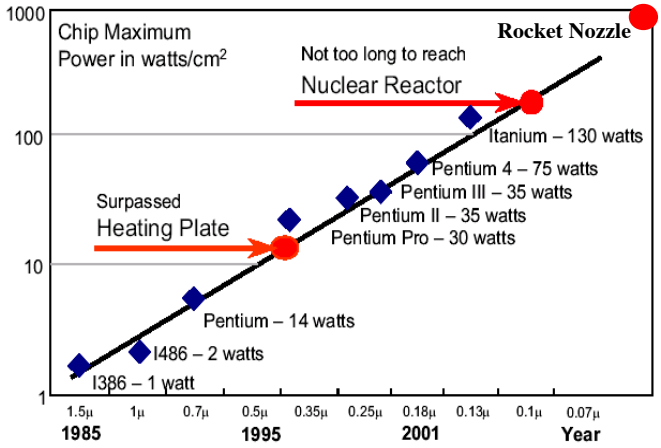

Figure 3. Power densities for Intel CPUs. Source: Shekhar Borkar, Intel.

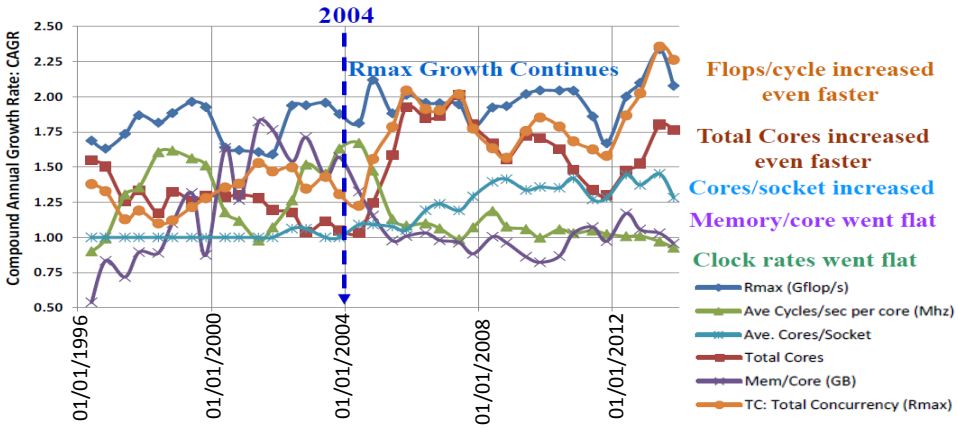

Figure 4. The evolution of CPU characteristics 1996 – 2012, P. Kogge [37].

The power constraints led to the concept of "dark silicon" [38], i.e., silicon that is underutilized. In todays' CPUs (22 nm technology) about 20% is dark [39], a percentage that is rapidly increasing as feature sizes keep being scaled down [40] such that by the end of the decade predictions range from 50% [39,41] to 90+% dark [38,40] silicon for processor dies. This fact poses interesting challenges and opportunities in processor design. The energy requirements for computations carried out on "big", "small" and specialized cores can be quite different. Further, on die memory (cache) tend to consume less energy than corresponding amount of functional logic since only small fractions of memory are actively used at any given time. Caches can not only lead to increased performance but also reduced need for very energy consuming main memory accesses. Though L1 and L2 cache sizes per core have remained relatively unchanged over time, L3 caches have increased significantly with some current processor dies having over 100MB of cache [42]. The difference in average power density of functional logic and caches is illustrated in Figure 5.

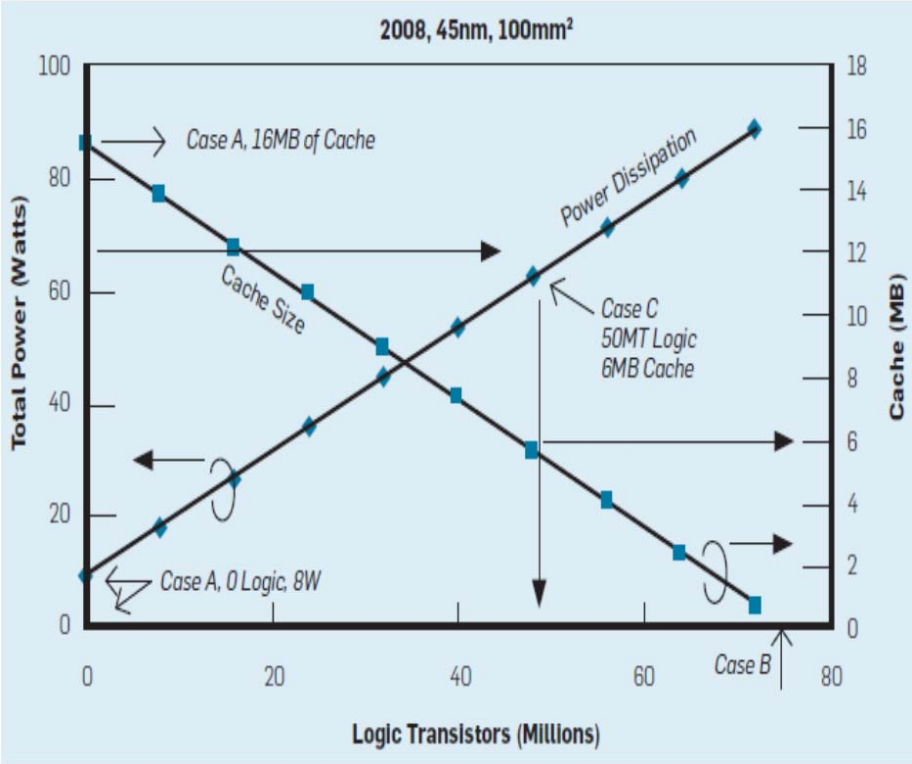

Figure 5. Functional logic and cache trade-offs for constant power dissipation [27].

The energy efficiency benefits of specialization can yield several orders of magnitude improvement in performance, power consumption, or both as seen in Figure 6. For the embedded and mobile markets the use of specialization to achieve high energy efficiency is common as exemplified by the schematic drawing of a Texas Instruments (TI) System-on-Chip (SoC), Figure 7, and the Qualcomm Snapdragon 820 SoC, Figure 8. A comparison of the nominal 64-bit floating-point energy efficiency and computational density for processors considered for evaluation in the PRACE project is shown in Table 1.

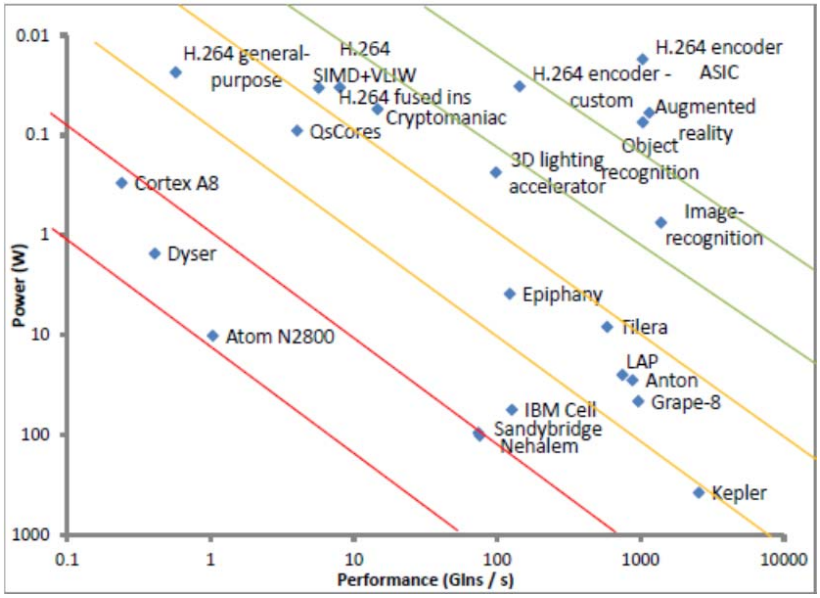

Figure 6. Power and performance of a wide range of accelerators (scaled to 45 nm) [43].

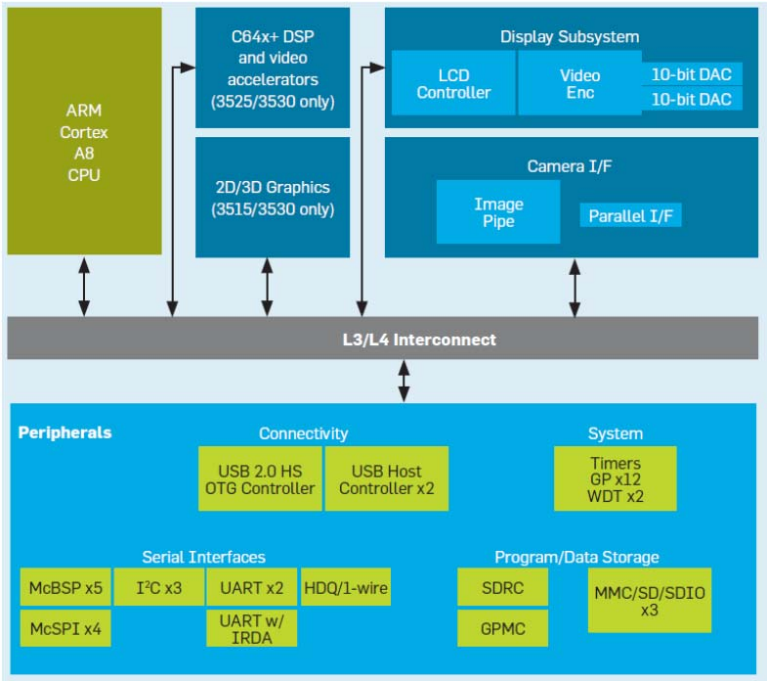

Figure 7. Schematic of a Texas Instruments System-on-Chip [27].

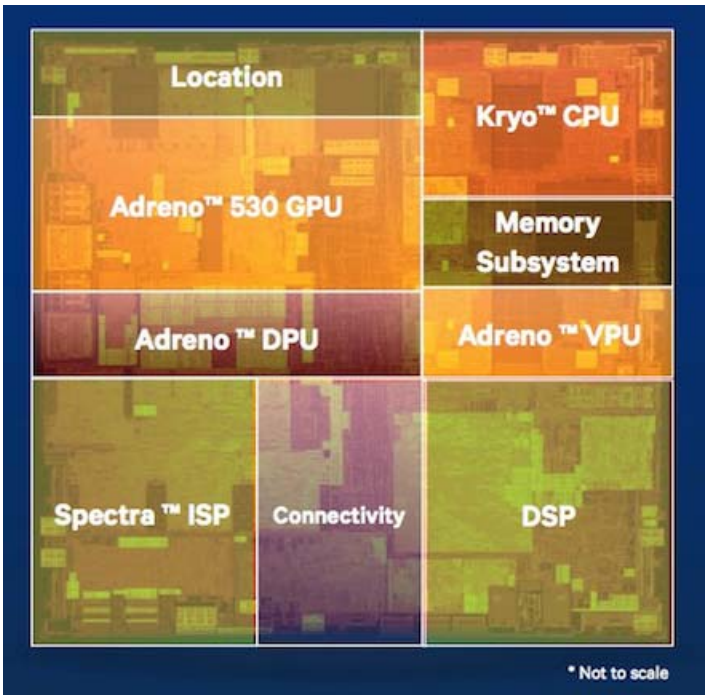

Figure 8. Qualcomm Snapdragon 820 mobile power and thermal optimized SoC [44].

Table 1. Nominal double precision floating-point performance of processors evaluated or considered for evaluation in the PRACE 1st Implementation Phase project [45].

CPU	Current [2012]			Next Generation [2012]		
	Lith. [nm]	E. Eff. [GF/J]	Comp. Dens. [GF/mm^2]	Lith. [nm]	E. Eff. [GF/J]	Comp. Dens. [GF/mm^2]
AMD Interlagos (16C, 2.7 GHz)	32	1.3	0.55			
AMD FirePro 7900	40	3.0	1.19			
AMD S9000				28	3.6	2.21
AMD Trinity	32	1.0	0.40			
IBM Blue Gene/Q	45	3.7	0.57			
IBM Power7	45	1.4	0.48			
Intel Sandy Bridge (8C, 3.1 GHz)	32	1.3	0.46			
Intel Ivy Bridge (4C, 3.5 GHz)				22	1.45	0.70
NVIDIA Fermi	40	2.66	1.26			
NVIDIA Kepler				28	5.2	2.36
NVIDIA Tegra 2	40	≈1	0.04			
TI TMS320C6678	40	6	≈3			
Xilinx Vertex-6	40	5-10				
Xilinx Vertex-7				28	≈13	0.236

Below, we report briefly on an assessment of the energy efficiency of a DSP based SoC in the context of High-Performance Computing (HPC) applications. The assessment used the TI TMS320C6678 DSP, a 10W SoC manufactured in 40 nm CMOS technology with a potential peak 64-bit floating-point performance of 60 GF/s and hence a nominal peak energy efficiency of 6GF/J. As a comparison the contemporary x86 CPU by Intel, the 32nm Sandy Bridge, has a nominal peak energy efficiency of 1.3 GF/J, and the 45nm Blue Gene/Q a nominal energy efficiency of 3.7 GF/J. In section 2 we describe the features of the TI SoC essential for our benchmark implementations, and in Section 3 the instrumentation used to measure power consumption. Section 4 describes the outcome and techniques used to achieve high efficiency and energy efficiency for the STREAM memory system benchmark [46] and Section 5 the result of and techniques used in the High-Performance Linpack (HPL) compute intensive benchmark [47]. Section 6 gives the conclusions.

2. The Texas Instruments DSP TMS320C6678

The TMS320C6678 multi-core SoC, referred to as the 6678 below, contains eight TMS320C66x DSP cores. These 8-wide Very Large Instruction Word (VLIW) processors can execute four IEEE 754 64-bit floating-point additions and two multiplications, together with two 64-bit data accesses per clock cycle in a single hardware thread [48]. The 16-stage instruction pipeline is almost free of interlocks and makes extensive use of *delay slots* to hide instruction latencies. At a maximum common core clock frequency of 1.25 GHz, the SoC delivers up to 60 GF/s, double precision, within the 10W TDP (Thermal Design Power) envelope, and a nominal energy efficiency of 6 GF/J. Figure 9 shows a functional block diagram of the 6678.

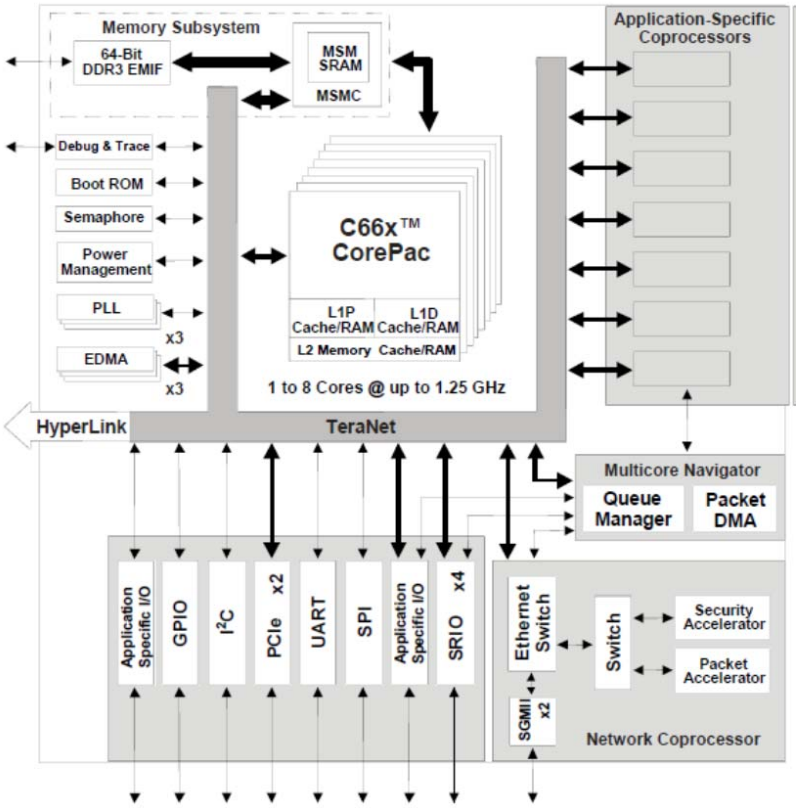

Figure 9. A functional block diagram of the TI TMS320C6678 SoC [49].

A single 64-bit DDR3 interface supports up to 8GiB at up to 1600 MT/s. Of the SoC communication features the proprietary 50 Gb/s Hyperlink [50] is the most interesting for clustering of DSPs for HPC and offers a 0.1 B/Flop off-chip communication to double-precision floating-point capability. This communication-to-compute ratio is almost identical to that of the Blue Gene/Q [51]. The 6678 core private L1 and L2 memories, Table 2, can be configured as cache or fast scratch-pad memory, referred to as SRAM below, or a combination thereof [52]. The memory latencies in Table 2 are from load to use, with L1 latencies fully overlapped by five delay slots for branches and four for loads. Stores do not require a delay slot.

Table 2. TI 6678 core private memory characteristics [53].

	Unit	L1I	L1D	L2
Total Size	KiB	32	32	512
Cache Block Size	B	32	64	128
Associativity		Direct	2-way	4-way
Replacement Policy			LRU	LRU
Allocation Policy		Read	Read	Read/Write
Clock		f_{ck}	f_{ck}	$f_{ck}/2$
Latency SRAM	Cycles	6	5	8-15.5
Latency Cache	Cycles	6	5	12-17.5
Bandwidth	B/Cycle		16	16

Core-initiated memory operations pass via its eXtended Memory Controller (XMC), which also contains a simple hardware prefetcher, to the 6678 Multi-core Shared Memory Controller (MSMC) that manages accesses to the 4MiB on-chip Multi-core Shared Memory (MSM), as well as to the external DDR3 memory. Figure 10 illustrates the data paths of a core (one shown) as well as the 6678 interconnect fabric, the 2 Tbps non-blocking TeraNet (which in fact consists of three networks). The digit after the name in Figure 10 denotes the fraction of the core clock frequency, f_{CK}, of a particular TeraNet (TeraNet 2 runs at $f_{CK}/2$, TeraNet 3 at $f_{CK}/3$). Communication between a core, other cores and on-chip peripheral units use the TeraNets via a core's External Memory Controller (EMC). The TeraNet also connects to the MSM and external memory and hence provides an alternate path to the one via the XMC to MSM and external memory. However, while the XMC port provides a bandwidth of $32*f_{CK}/2$ B/s, the EMC port is connected to the TeraNet 3 and hence is limited to a bandwidth of $16*f_{CK}/3$ B/s. Despite the limitation of the path via the EMC it can offer higher realized bandwidth than the XMC path as will be shown in sections four and five. Using the EMC path a single core cannot fully utilize the DDR3-1333 memory bandwidth of 10.67 GB/s, but two cores could potentially do so.

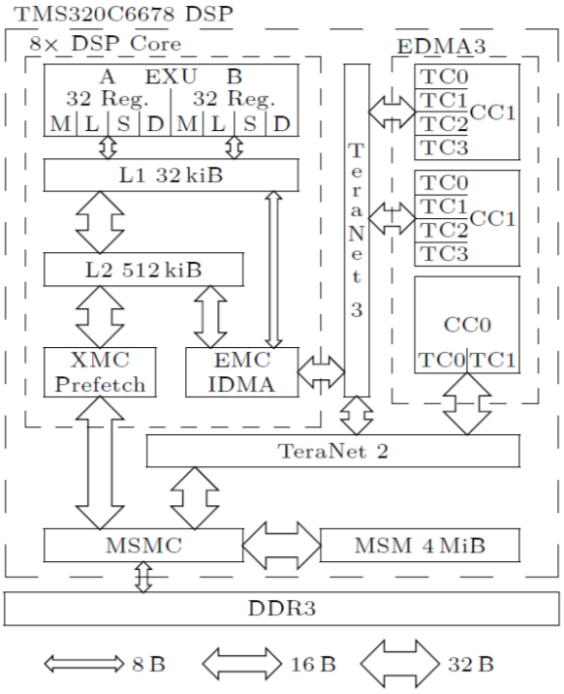

Figure 10. A schematic of the data paths of the 6678 SoC [54].

The 6678 has two DMA engines: 1) an Internal DMA (IDMA) engine [52] that can transfer data between the core private L1 and L2 memories, and between those and the EMC and 2) an Enhanced DMA engine, EDMA3 [55]. The EDMA3 controller can move data between all memories including the core private memories. The EDMA3 engine is very versatile and supports strided block transfers as shown in Figure 11. It has three independent Channel Controllers (CC) that each can serve multiple Transfer Controllers (TC) that execute the data transfers. The CCs support *chaining* by which the completion of a transfer can trigger another transfer, possibly on another channel. *Linking* enables the automatic reload of new transfer specifications from the controller's Parameter RAM (PaRAM) making implementations of linked lists of transfers possible. Strided block transfers, chaining and linking are used to achieve high efficiency and competitive energy efficiency in the benchmarks reported below.

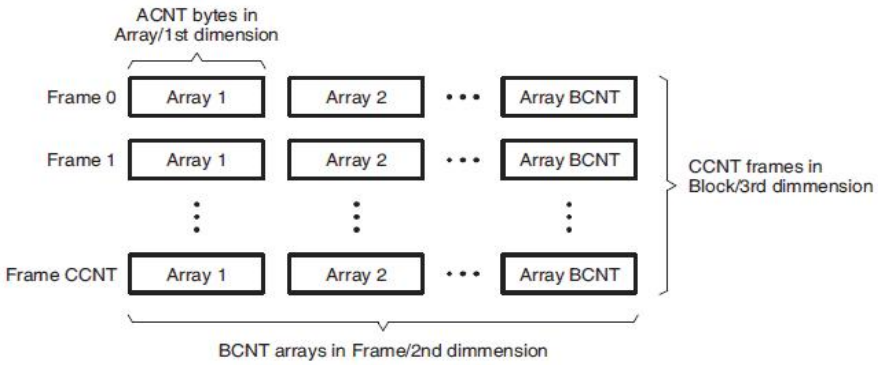

Figure 11. The EDMA3 strided block transfer capabilities [55].

Our experimental setup consists of an Advantech Evaluation Module (EVM) containing the 6678 and 1GiB DDR3 memory [56]. The four 2Gb Samsung K4B2G1646C-HCH9 DDR3 memories run at 666MHz (1333MT/s) giving a peak bandwidth of 10.67 GB/s. The eight bank DDR3 memory has a page size of 8KiB. Knowledge of the page size is essential to understand some aspects of the performance behavior of the benchmarks.

3. The Power Measurement setup

The 6678 has 16 power domains: one for each of the cores, one for the MSM, one for the Hyperlink, one for the SRIO (Serial Rapid I/O), one for the PCIe, one for Peripheral Logic, one for the Packet Engine, and one for the Trace Logic and one for remaining logic. There is no voltage control. Each core can be powered off or on independently, a feature we used in the STREAM benchmark in assessing the energy efficiency of the 6678. We also powered off logic not used in the benchmarks, such as the SRIO, Peripheral Logic, and the Packet Engine.

The different power domains are supplied by three power rails that were instrumented as shown in Figure 12. The logic associated with each of the three power rails is shown in Table 3 and depicted in Figure 13. With our instrumentation we estimate the error to less than 1% for the measured parts. For details about the instrumentation and the error estimation see [57].

Table 3. Logic associated with the three power rails supplying the SoC (Variable, Fixed and Memory)

Rail	Name	Voltage [V]	Usage
Variable	CVDD	0.9 – 1.1	All SoC logic
Fixed	CVDD1	1.0	On-chip memories
	VDDT1		HyperLink termination
	VDDT2		Other termination
Memory	DVDD15	1.5	DDR3 Memory and I/O
	VDDR1		HyperLink SerDes
	VDDR2		PCIe SerDes
	VDDR3		SGMII SerDes
	VDDR4		SRIO SerDes
Input		12	DC input to EVM
Other			Uninstrumented parts
			Calculated from above
			DC-DC converter losses

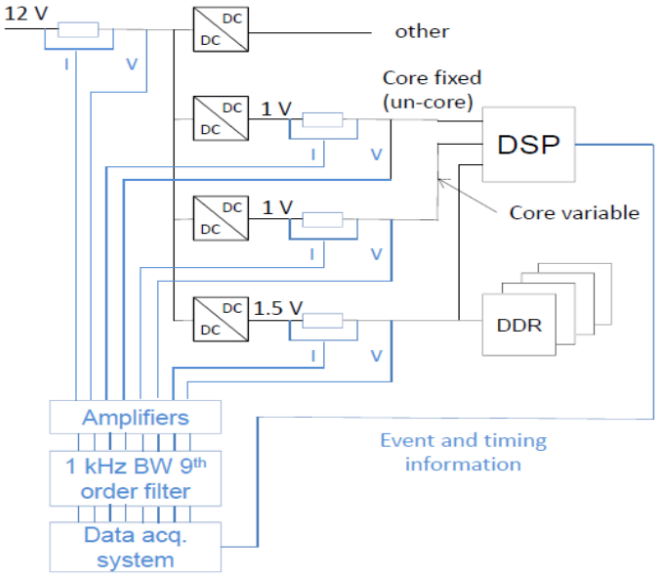

Figure 12. Instrumentation of the EVM for power measurements.

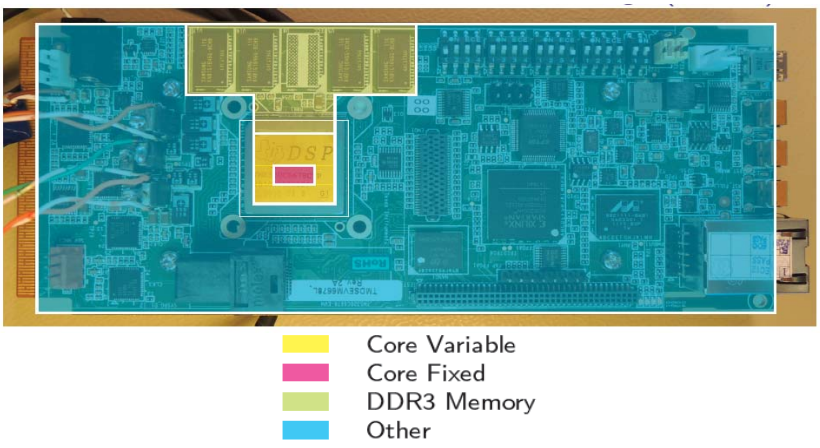

Figure 13. The EVM parts associated with the three power rails.

4. The STREAM Benchmark

The Stream memory system benchmark [46] commonly used to assess memory system performance for stride one access with an emphasis on main memory access consists of the operations shown in Table 4. Despite their low stride that is ideal for the way data is assigned to memory, their regularity, and simplicity, the operations present serious challenges for modern cache based computer architectures. Since there is no data reuse the main objective of caches is not met. Further, caches typically use a read-modify-write policy in writing to main memory causing an extra load not accounted for in the STREAM benchmark hence limiting the peak achievable bandwidth for COPY and SCALE to 2/3rd of the peak main memory bandwidth and to 3/4th of the peak for SUM and TRIAD. The lack of data reuse makes STREAM operations very latency sensitive. To address the memory performance challenges for applications with limited or no data reuse most modern processors employ prefetching, some very sophisticated prefetching techniques that tracks memory

access patterns and adapt the prefetching accordingly to avoid stalls caused by waiting for main memory loads. Some also employ cache bypass mechanism for store operations to avoid the cache problems on writes to main memory. Despite these mechanisms, high STREAM performance is difficult to achieve as is evident from Table 5 based on reported performance data and estimates of energy efficiencies when we did not find any reported power consumption data. We estimated the power consumption for STREAM to 50% of the processor TDP. STREAM is not CPU intensive.

Table 4. The STREAM Benchmark operations.

Name	Operation	Data Size [B]	Flops
COPY	$a_i = b_i$	16	0
SCALE	$a_i = qb_i$	16	1
SUM	$a_i = b_i+c_i$	24	1
TRIAD	$a_i = b_i+qc_i$	24	2

Table 5. STREAM data from literature with estimates of energy consumption.

	Bandw. (BW) [GB/s]		Eff. [%]	E. Eff. [GB/J]	Lith. [nm]		Ref.
	Peak	Meas.					
IBM Power 7	409.6	122.8	30	≈0.10	45	Power est.	[58][59],[60]
Intel Xeon Phi	352.0	174.8	50	≈1.20	22	at 50% of	[58],[61]
Intel E5-2697v2	119.4	101.5	85	≈0.30	22	TDP	[62]
NVIDIA Tegra 3	6.0	1.6	27	0.05	40	Power and BW meas.	[63]
TI TMS320C6678 Cache	10.7	3.0	28	0.40	40	Power and	[45]
TI TMS320C6678 DMA	10.7	10.2	96	1.26	40	BW meas. at 1 GHz	[53]

As seen from Table 5, the efficiency of the "6678 cache" STREAM implementation yields an efficiency at the low end of the range among processors in the table, though the energy efficiency of the 40 nm 6678 is about the same as the 22 nm Intel E5-2697v2 (Ivy Bridge). Thus, for the 6678 to represent an interesting alternative from an energy efficiency point of view the memory bandwidth must be utilized close to 100% without a comparable increase in power consumption. If that could be achieved, then the 6678 would compete with the 22 nm Intel Xeon Phi in regards to energy efficiency. As seen in Table 5 the 6678 DMA alternative indeed realizes this potential.

As shown in Table 6 much of the poor performance of the "6678 cache" can be attributed to latency [53]. This fact also implies that the XMC prefetching is not very effective. The only poor performance prediction from measured latencies is that for the DDR3 bandwidth. The reason for the poor prediction is that the DDR3 measured stall time did not incur a significant number of DDR3 bank conflicts. Table 7 shows the impact on DDR3 memory latency as a function of memory stride and resulting bank conflicts.

Table 6. Stall times: estimated and measured. Bandwidth: per core bandwidth estimated based on stall times, bandwidth measured for 8 cores and average measured bandwidth/core [53].

	Stall Time [ns]		Bandwidth [GB/s]		
	Spec. [64]	Meas.	Est./core	Meas./8	Meas.
L1D	0	0	16.0	15.5	124.0
L2 SRAM	3-10.5	7	5.8	6.0	47.6
L2 Cache	7-12.5	9	4.9	5.8	46.5
MSM	-	10	4.6	4.8	38.7
DDR3	-	23	1.7	0.38	3.0

Table 7. Measured latencies as a function of stride for the EVM.

Memory	Stride in Bytes								
	64	128	256	512	1ki	2ki	4ki	8ki	16ki
L1D SRAM	5	5	5	5	5	5	5	5	5
L2 SRAM	12	12	12	12	12	12	12	12	12
MSM	15	28	28	28	28	28	28	28	28
DDR3-1333	28	49	92	93	95	99	106	121	122

By using the EDMA3 and configuring the on-die L1 and L2 memories mostly as scratchpad memory, not only can prefetching be managed such that stalls are avoided, but the DDR access by the eight cores managed such that bank conflicts are minimized and not affecting the performance. In our EDMA3 implementation it is in fact the master of the benchmark execution that signals the cores when data is available in their caches. Conversely, they signal the EDMA3 when data is ready to be copied back to main memory. No communication between cores is necessary. Communication between a core and the EDMA3 is made via flags local to the core. To further decouple core and EDMA3 operations multi-buffering is used so that for each core the core and the EDMA3 use different buffers at any given time. To overlap overhead and latencies in setting up and executing data transfers two transfer controllers are used. The chaining feature of EDMA3 is used to reduce latencies by having the completion of one transfer event directly start another. The linking feature is used to avoid waiting time for loading transfer descriptors. Details of the implementation can be found in [53].

Despite the fact that a single core bandwidth to DDR3 memory is limited to $16*f_{CK}/3$ B/s due to the TeraNet 3, the EDMA3 bandwidth for a single core in our implementation is still higher than what was achieved using the XMC $32*f_{CK}/2$ B/s path. The limited success of the XMC pre-fetcher, the read-modify-write cache policy, and the DDR3 bank conflicts caused by the accesses from the eight cores all contributes to this less than stellar performance (though the efficiency is comparable to some very sophisticated designs). Using the EDMA3 theoretically two cores operating at 1 GHz suffices to saturate the DDR3-1333 bandwidth. Table 8 and Figure 14 give some of the results of our EDMA3 implementation. Our implementation achieves an efficiency of 96.8% for a single core for which the bandwidth is limited by the TeraNet 3. For two cores for which the TeraNet 3 and DDR3 bandwidths are the same the implementation achieves an efficiency of 83.0%. For four active cores the efficiency is 89.0%, for six active cores 92.1% and for eight active cores 95.6%. The "dips" in performance for an odd number of cores is due to load imbalance between the two transfer-controllers used in our EDMA3 STREAM implementation. Increasing the core clock frequency improves the performance somewhat, largely due to the fact that the TeraNet bandwidth scales with the core clock frequency.

Table 8. Performance and energy efficiency for STREAM on 128 MiB vectors and different active core counts with non-active cores powered off for f_{CK} = 1 GHz [53]. DDR3 bandwidth 10.67 GB/s.

Cores	Measured BW [GB/s]				Energy Efficiency [GB/J]			
	COPY	SCALE	SUM	TRIAD	COPY	SCALE	SUM	TRIAD
1	5.16	5.16	5.12	5.12	1.01	1.01	1.01	1.00
2	8.86	8.86	8.41	8.41	1.47	1.47	1.43	1.42
3	6.99	6.99	6.89	6.89	1.15	1.15	1.15	1.14
4	9.50	9.50	9.11	9.11	1.41	1.41	1.38	1.37
5	7.99	7.99	7.85	7.85	1.17	1.17	1.16	1.16
6	9.83	9.83	9.58	9.58	1.33	1.33	1.31	1.30
7	8.67	8.67	8.66	8.66	1.15	1.15	1.15	1.14
8	10.20	10.20	10.04	10.04	1.26	1.26	1.25	1.24

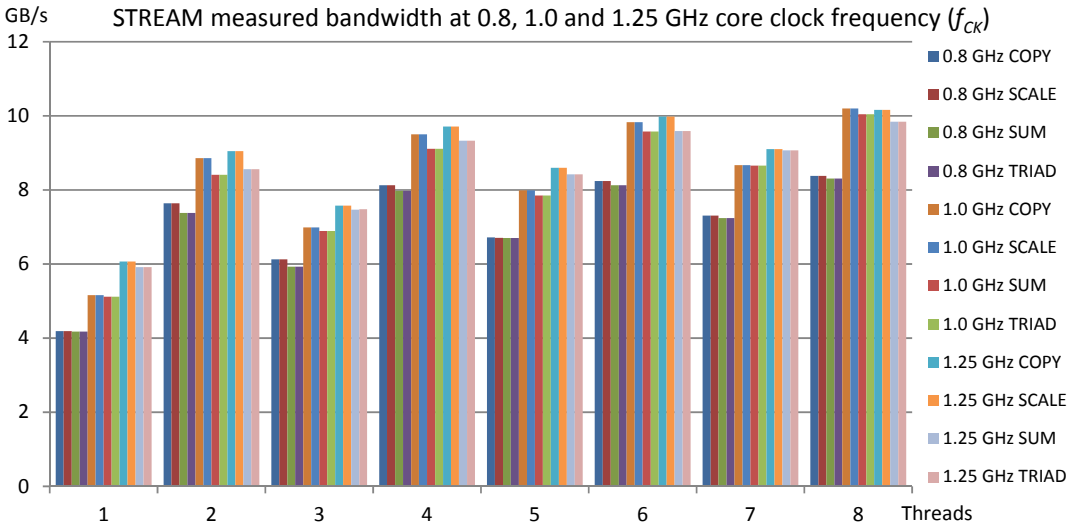

Figure 14. EDMA3 STREAM performance data as a function of core clock frequency and active cores for DDR3-1333 (10.67 GB/s).

Table 8 and Figure 15 shows that the energy efficiency peaks at 1.47 GB/J for COPY and SCALE, 1.43 GB/J for SUM and 1.42 GB/J for TRIAD with two active cores. As additional cores are activated and used the energy consumption increases more than what is gained in DDR3 memory bandwidth utilization. It is also clear that 1 GHz core clock frequency for two or more active cores is optimum. For a single core the increased bandwidth offered by the increased TeraNet bandwidth at the higher clock frequency yields a higher energy efficiency, i.e., the realized bandwidth increases more than the power consumption.

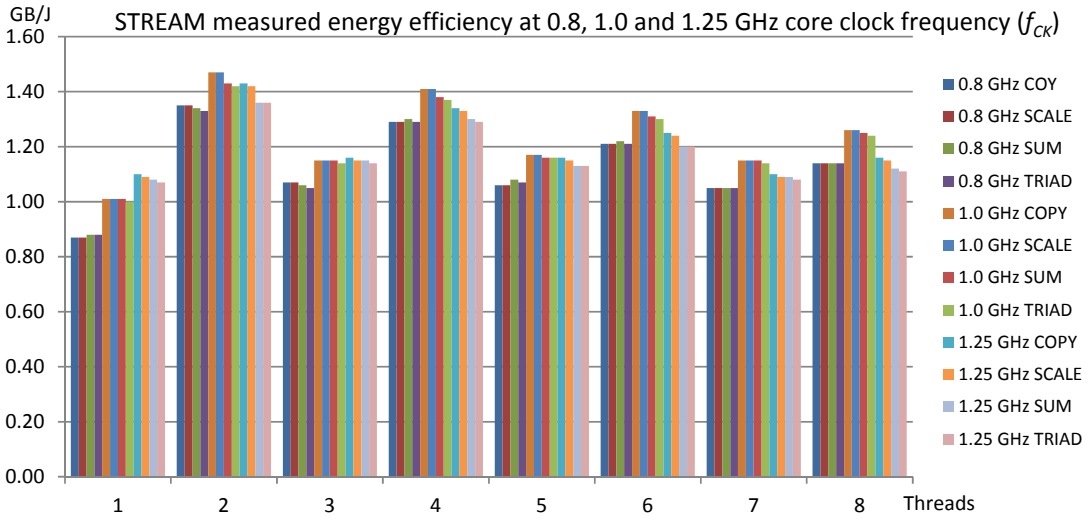

Figure 15. EDMA3 STREAM energy efficiency as a function of core clock frequency and active cores.

Our measurements show that an idle 6678 core consumes about 110 mW. For COPY the power consumption going from idle to fully active increases by about 200 mW at f_{CK} = 1 GHz.

Thus, for STREAM using the EDMA3 feature the 40nm 6678 is competitive with the 22 nm Intel Xeon Phi in regards to energy efficiency, and considerably more efficient in utilizing the main memory bandwidth (about 95% of peak compared to about 50% of peak).

5. The HPL Benchmark

The HPL benchmark (used for the Top500 computer system ranking [65]) is a compute intensive benchmark that on x86 architectures typically achieves in excess of 90% efficiency [66]. For most current architectures high efficiency is required for high energy efficiency (as shown above for STREAM). Like for STREAM a straightforward compilation of the HPL benchmark for the 6678 resulted in poor efficiency, less than 25% at f_{CK} = 1 GHz. To get an efficiency that could make the 6678 a viable competitor in regards to energy efficiency, an efficiency comparable to that of x86 platforms is required. To achieve this goal efficient use of the EDMA3 engine, as for STREAM, was necessary, but not only for DDR3 memory access, but also for overlapping memory accesses with computation. In addition, EDMA3 was used to carry out data transposition in moving data from main memory to L2 SRAM using the MSM as a staging memory and the strided block transfer capability of the EDMA3 [54]. Unlike STREAM, HPL has the potential for significant data reuse since it requires about $\approx 2/3N^3$ operations for N^2 data. For our HPL implementation the 6678 L1 and L2 core private memories are largely configured as SRAM for storing matrix blocks that are reused. The MSM also serves a critical role, mostly as a staging memory in allocating matrix blocks retrieved from main memory to cores in a cyclic fashion, and when needed used for rearranging data from block row-major to block column major ordering. Out HPL implementation also make use of the IDMA engine. It is used to pre-fetch matrix blocks from L2 SRAM to L1 SRAM such that all data loads are from L1 SRAM without stalls. Stores are directed directly to L2 SRAM via a write-buffer bypassing L1 SRAM.

As for most efficient HPL implementation ours also depend on an efficient matrix-multiplication (DGEMM [67]) implementation. Our matrix-multiplication routine achieves an efficiency of about

95% at f_{CK} =1 GHz and hence is comparable to that observed for Intel CPUs and higher than the efficiency on some other x86 CPUs as well as other architectures [66]. With un-optimized forward and backsolve our HPL implementation achieves a peak efficiency of 77%, and a peak energy efficiency of 2.92 GF/J [54] which compares favorably with Intel Sandy Bridge systems, 0.97 GF/J [69] and Blue Gene/Q systems 2.10 GF/J [68].

The 6678 computational rate as a function of matrix size for three core clock rates are shown in Figure 16, the corresponding efficiencies in Figure 17 and energy efficiencies in Figure 18

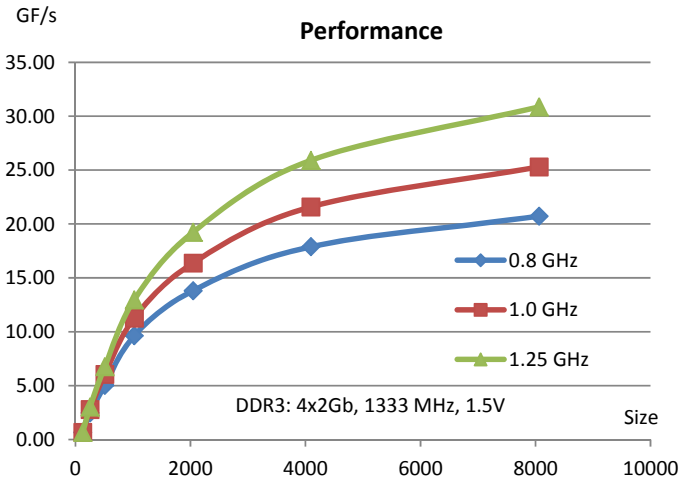

Figure 16. The 6678 HPL performance as a function of matrix size and core clock rate.

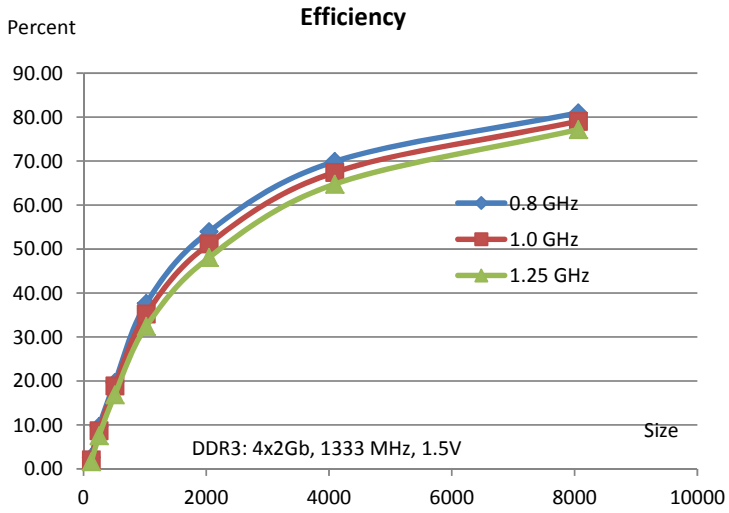

Figure 17. The 6678 HPL efficiency as a function of matrix size and core clock rate.

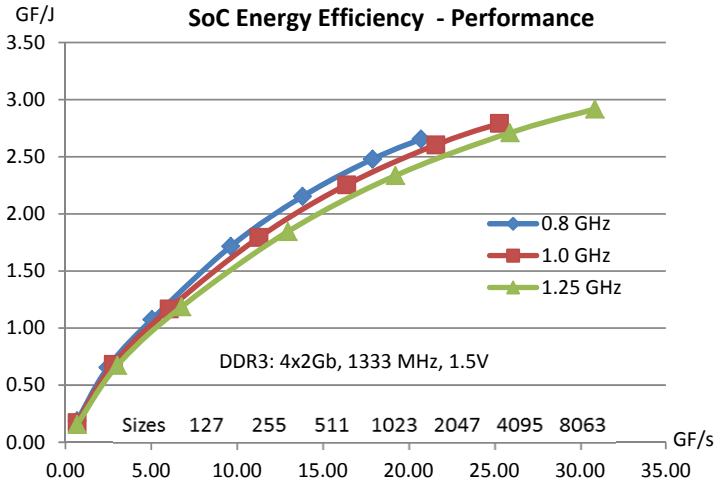

Figure 18. The 6678 HPL energy efficiencies as a function of matrix size and core clock rate.

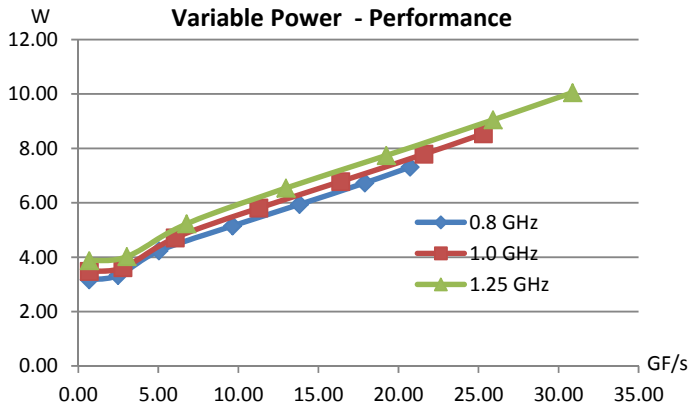

Figure 19. The 6678 HPL Variable (core) power consumption as a function of computational rate and clock rate.

Figure 19 shows, as expected, that for HPL the power consumption, except at low rates and small matrices increases linearly with the computational rate. Figure 20 supports the claim that HPL is limited by the computational rate since as the matrix size increases the DDR3 power dissipation in fact decreases once the matrix and computational demands increases beyond a certain size.

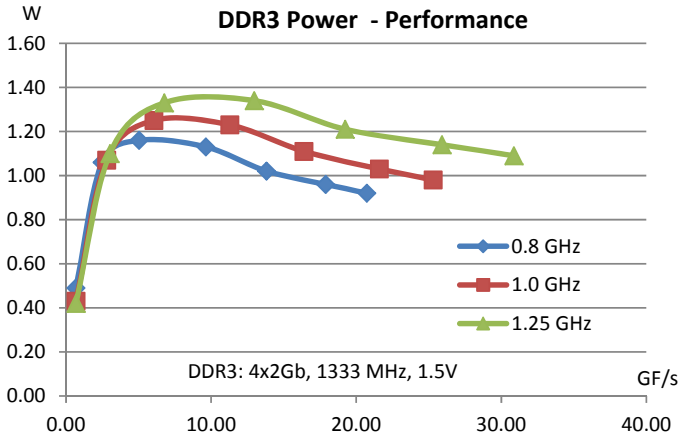

Figure 20. The EVM DDR3 memory power consumption as a function of computational and clock rates.

6. Conclusion

Compared to the highly successful and feature rich x86 architecture with a TDP of up to about 150W the 10W TMC320C6678 DSP from Texas Instruments does offer competitive energy efficiencies for the two common HPC benchmarks reported here; STREAM and HPL. In fact, the 40 nm 6678 SoC offered comparable or better energy efficiency than 22 nm x86 based designs and even better energy efficiency than the proprietary Blue Gene/Q. Compared to the latter the 6678 is a high volume product with a cost even significantly below high end x86 CPUs from Intel. But, in regards to HPC applications the software eco-system is quite immature compared to that for x86 architecture based systems. Further, some features in the x86 architecture that simplifies achieving high efficiency are not present in the 6678 SoC requiring a greater effort in achieving comparable efficiencies, and a more detailed understanding of the architecture and how to use it effectively. However, TI provided tools for the 6678 are sufficient to achieve efficiencies comparable to that on x86 architectures for STREAM and HPL and we conjecture also for many HPC applications.

References

[1] W.M. Holt. Moore's Law: A path going forward. *2016 IEEE International Solid-State Circuits Conference (ISSCC)*, pp 8-13 January 2016. http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=7417888&url=http%3A%2F%2Fieeexplore.ieee.org%2Fstamp%2Fstamp.jsp%3Ftp%3D%26arnumber%3D7417888

[2] K. Bourzac. Intel: Chips will have to sacrifice speed gains for energy savings. *MIT Review*, February 5 2016. https://www.technologyreview.com/s/600716/intel-chips-will-have-to-sacrifice-speed-gains-for-energy-savings

[3] R. Merritt. Moore's Law goes Post-CMOS. *EE-Times*, February 1 2016. http://www.eetimes.com/document.asp?doc_id=1328835

[4] F. von Trapp. From the Ashes of More's Law: More than Moore has Arrived. 3DInCites, February 12 2016. http://www.3dincites.com/2016/02/from-the-ashes-of-moores-law-more-than-more-has-arrived/

[5] J. Gao. Machine Learning Applications for Data Center Optimization. May 28 2014. http://static.googleusercontent.com/media/www.google.com/en//about/datacenters/efficiency/internal/assets/machine-learning-applicationsfor-datacenter-optimization-finalv2.pdf
[6] S. Levy. Google Throws Open Doors to Its Top-Secret Data Center. *Wired,* October 17 2012. http://www.wired.com/2012/10/ff-inside-google-data-center/
[7] L.A. Barroso and U. Hölzle. The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines, *In Synthesis Lectures on Computer Architecture* ed Mark D Hill (Morgan & Claypool Publishers) 2009.
[8] L.A. Barroso and U. Hölzle. The case for energy proportional computing. *IEEE Computer* **vol 40** No. 12 pp 33 – 37 2007. http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=4404806&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D4404806
[9] A.M. Putnam, A.M. Caulfield, E.S. Chung, D. Chiou, K. Constantinides, J. Demme, H. Esmaeilzadeh, J. Fowers, G.P. Gopal, J. Gray, M. Haselman, S. Hauck, S. Heil, A. Hormati, J.-Y. Kim, S. Lanka, J. Larus, E. Peterson, S. Pope, A. Smith, J. Thong, P.Y. Xiao, and D. Burger. A Reconfigurable Fabric for Accelerating Large-Scale Datacenter Services. *41st Annual International Symposium on Computer Architecture (ISCA)*, pp. 13-24 June 2014. http://research.microsoft.com/apps/pubs/default.aspx?id=212001
[10] K. Ovtcharov, R. Olatunji, J.-Y. Kim, J. Fowers, K. Strauss, and E.S. Chung. Accelerating Deep Convolutional Neural Networks Using Specialized Hardware. February 23, 2015. http://research.microsoft.com/apps/pubs/default.aspx?id=240715
[11] A. Putnam. Large-Scale Reconfigurable Computing in a Microsoft Datacenter. *HotChips* August 2014. http://www.hotchips.org/wp-content/uploads/hc_archives/hc26/HC26-12-day2-epub/HC26.12-5-FPGAs-epub/HC26.12.520-Recon-Fabric-Pulnam-Microsoft-Catapult.pdf
[12] The Open Compute Project. http://www.opencompute.org/
[13] B. Lannoo, S. Lambert, W. Van Heddeghem, M. Pickavet, F. Kuipers, G. Koutitas, H. Niavis, A. Satsiou, M.T. Beck, A. Fischer, H. de Meer, P. Alcock, T. Papaioannou, N.H. Viet, T. Plagemann, and J. Aracil. Information and Communication Technologies, Network of Excellence in Internet Science, D8.1. Overview of ICT energy consumption. February 5 2013. http://www.internet-science.eu/sites/eins/files/biblio/EINS_D8%201_final.pdf
[14] B. Lannoo. Energy Consumption of ICT Networks. TREND Final Workshop, Brussels, October 24 2013.
[15] H. Li. Server Board Design. *HotChips,* August 2011. http://www.hotchips.org/wp-content/uploads/hc_archives/hc23/HC23.17.2-tutorial2/HC23.17.230.Open%20Compute%20Motherboard.pdf
[16] E. Tal. Facebook Storage Tiers. *HotChips*, August 2011. http://www.hotchips.org/wp-content/uploads/hc_archives/hc23/HC23.17.2-tutorial2/HC23.17.220.Open%20Compute%20Storage.pdf
[17] J. Brodkin. Who needs HP and Dell? Facebook now designs all its own servers. *Arstechnica*, February 14 2013. http://arstechnica.com/information-technology/2013/02/who-needs-hp-and-dell-facebook-now-designs-all-its-own-servers/
[18] S. Shankland. Google uncloaks once-secret server. *CNET*, December 11 2009. http://www.cnet.com/news/google-uncloaks-once-secret-server-10209580/
[19] T.P. Morgan. How Microsoft Is Using FPGAs To Speed Up Bing Search. *EnterpriseTech,* September 3 2014. http://www.enterprisetech.com/2014/09/03/microsoft-using-fpgas-speed-bing-search/

[20] A. Michael. The Open Computer Project: Project Motivation and high level system design *HotChips*, August 2011. http://www.hotchips.org/wp-content/uploads/hc_archives/hc23/HC23.17.2-tutorial2/HC23.17.210.Open%20Compute%20Overview.pdf
[21] A. Singh, J. Ong, A. Agarwal, G. Anderson, A. Armistead, R. Bannon, S. Boving, G. Desai, B. Felderman, P. Germano, A. Kanagala, J. Provost, J. Simmons, E. Tanda, J. Wanderer, U. Hölzle, S. Stuart, and A. Vahdat. Jupiter Rising: A Decade of Clos Topologies and Centralized Control in Google's Datacenter Network, *SIGCOMM*, August 2015. http://conferences.sigcomm.org/sigcomm/2015/pdf/papers/p183.pdf
[22] R. Miller. Inside Microsoft's Chicago Data Center. *Data Center Knowledge*, 2009. http://www.datacenterknowledge.com/inside-microsofts-chicago-data-center
[23] J. Dean and S. Ghemawat. MapReduce: Simplified Data Processing on Large Clusters. *CACM*, **vol 51** No.1 pp 107-113 2008. http://dl.acm.org/citation.cfm?id=1327492
[24] F. Chang, J. Dean, S. Ghemawat, W.C. Hsieh, D.A. Wallach, M. Burrows, T. Chandra, A. Fikes, and R.E. Gruber. Bigtable: A Distributed Storage System for Structured Data. ACM Transactions on Computer Systems, **vol 26** No.2 article No. 4 June 2008. http://dl.acm.org/citation.cfm?id=1365816
[25] G.E. Moore. Cramming more components onto integrated circuits. *Electronics*, **vol 38** No 8 April 1965.
[26] R.H. Dennard, F.H. Gaensslen, H.-N. Yu, V.L. Rideout, E. Bassous, and A.R. Leblanc. Design of ion-implanted MOSFETs with very small physical dimensions *IEEE J. Solid- State Circ* **vol 9** p.256 1974. http://ieeexplore.ieee.org/xpl/abstractAuthors.jsp?tp=&arnumber=752522&url=http%3A%2F%2Fieeexplore.ieee.org%2Fiel4%2F5%2F16251%2F00752522.pdf%3Farnumber%3D752522
[27] S. Borkar and A. Chien. The Future of Microprocessors *CACM* **vol 54** No 5 pp 67-77 May 2011.
[28] R Hameed, W. Qadeer, M. Wachs, O. Azizi, A. Solomatnikov, B.C. Lee, S. Richardson, C. Kozyrakis, and M. Horowitz. Understanding Sources of Inefficiency in General-Purpose Chips. *ISCA '10 Proceedings of the 37th annual international symposium on Computer architecture* pp 37-47 2010. http://dl.acm.org/citation.cfm?id=1815968
[29] W.J. Dally, J. Balfour, D. Black-Shaffer, J. Chen, R.C. Harting, V. Parikh, J. Park, and D. Sheffield. Efficient Embedded Computing. *Computer* **vol 41** no 7 pp 27-32 July 2008.
[30] W.J. Dally. Challenges for Future Computing Systems. *HiPEAC* 2015 Keynote. https://www.hipeac.net/2015/amsterdam/keynotes/
[31] J. Larus. It's the End of the World as we know it. *HiPEAC* 2015 Keynote. https://www.hipeac.net/2015/amsterdam/keynotes/
[32] E.J. Nowak. Maintaining the benefits of CMOS scaling when scaling bogs down *IBM Journal of Research and Development* **vol 46** no 2/3 pp 169-180 2002.
[33] D. Hillman. Integrated Power Management, Leakage Control and Process Compensation Technology for Advanced Processes. *Design & Reuse*, March 16 2009. http://www.design-reuse.com/articles/20296/power-management-leakage-control- process-compensation.html
[34] Power4, The first Multi-core, 1 GHz processor http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/power4/
[35] J. Kovacs. AMD Demonstrates World's First x86 Dual-Core Processor. August 2004. http://ir.amd.com/phoenix.zhtml?c=74093&p=irol-newsArticle&ID=608198
[36] Dual Core Era Begins, PC Makers Start Selling Intel-Based PCs April, 2005. http://www.intel.com/pressroom/archive/releases/2005/20050418comp.htm

[37] P. Kogge. Reading the Entrails: How Architecture Has Evolved at the High End IPDPS Keynote 2014.
http://www.ipdps.org/ipdps2014/IPDPS2014keynote-Kogge.pdf
[38] R. Merritt. ARM CTO: power surge could create 'dark silicon'. *EE-Times* October 22 2009.
http://www.eetimes.com/document.asp?doc_id=1172049
[39] H. Esmaeilzadehy, E. Blemz, R. St. Amantx, K. Sankaralingamz, and D. Burger. Dark Silicon and the End of Multicore Scaling *ISCA '11 Proceedings of the 38th annual International Symposium on Computer Architecture* pp 365-376 2011 also
IEEE Micro **vol 32** no 3 pp 122-134 2012.
http://dl.acm.org/citation.cfm?id=2000108
[40] M.B. Taylor. A Landscape of the New Dark Silicon Design Regime. *IEEE Micro* **vol 33** no 5 pp 8-19 2013.
http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6583151
[41] International Technology Roadmap for Semiconductors
www.itrs2.net
[42] J. Stuecheli. Power8. *HotChips*, August 2013.
http://www.hotchips.org/wp-content/uploads/hc_archives/hc25/HC25.20-Processors1-epub/HC25.26.210-POWER-Studecheli-IBM.pdf
[43] A. Guha, Y. Zhang,cR.Yao ur Rasool, and A. Chien. Systematic Evaluation of Workload Clustering for Extremely Energy-Efficient Architectures. *ACM SIGARCH Computer Architecture News*, **vol 41** Issue 2 pp 22-29 May 2013.
[44] **A.** Cunningham. Snapdragon 820 is official: A look at its GPU (and how much the chip matters). *Arstechnica* August 12 2015
http://arstechnica.com/gadgets/2015/08/snapdragon-820-is-official-a-look-at-its-gpu-and-how-much-the-chip-matters
[45] L. Johnsson and G. Netzer. Report on Prototypes Evaluation D9.3.3 PRACE April 2013.
http://www.prace-ri.eu/prace-1ip-deliverables/
[46] J. McCalpin. Memory Bandwidth and Machine Balance in Current High Performance Computers. *IEEE Computer Society Technical Committee on Computer Architecture (TCCA) Newsletter* pp 19-25 December 1995.
[47] A. Petitet, R.C. Whaley, J. Dongarra, and A.Cleary. HPL - A Portable Implementation of the High-Performance Linpack Benchmark for Distributed Memory Computers version 2.2, February 2016.
http://www.netlib.org/benchmark/hpl/
[48] Texas Instruments. TMS320C66x DSP CPU and Instruction Set Reference Guide Literature Number: SPRUGH7. November 2010.
[49] Texas Instruments. Multicore Fixed and Floating-Point Digital Signal Processor. March 2014.
http://www.ti.com.cn/cn/lit/ds/symlink/tms320c6678.pdf
[50] Texas Instruments. KeyStone Architecture, Hyperlink *User Guide* Literature Number: SPRUGW8C. June 2013.
http://www.ti.com/lit/ug/sprugw8c/sprugw8c.pdf
[51] R. Haring. The Blue Gene/Q Compute Chip, *HotChips*, August 2011.
http://www.hotchips.org/wp-content/uploads/hc_archives/hc23/HC23.18.1-manycore/HC23.18.121.BlueGene-IBM_BQC_HC23_20110818.pdf
[52] Texas Instruments. TMS320C66x DSP CorePac *User Guide*, Literature Number: SPRUGW0C. July 2013.
http://www.ti.com/lit/ug/sprugw0c/sprugw0c.pdf
[53] G. Netzer, L. Johnsson, D. Ahlin, E. Stotzer, P. Varis, and E. Laure. Exploiting DMA for Performance and Energy Optimized STREAM on a DSPIEEE *2014 International Parallel Distributed Processing Symposium Workshop* pp 805-814 May 2014.

http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=6969464&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D6969464

[54] G. Netzer. Efficient LU-Factorization for Texas Instruments Keystone Architecture Digital Signal Processors, M.Sc. Thesis, KTH June 2015. http://www.divaportal.org/smash/get/diva2:837145/FULLTEXT01.pdf

[55] Texas Instruments. KeyStone Architecture Enhanced Direct Memory Access (EDMA3) Controller *User Guide* Literature Number:SPRUGS5B. May 2015. http://www.ti.com.cn/cn/lit/ug/sprugs5b/sprugs5b.pdf

[56] TMDXEVM6678LE. TMX320C6678 Evaluation Module [Online] Available: http://www2.advantech.com/Support/TI-EVM/6678le_of.aspx

[57] G. Netzer, L. Johnsson, D. Ahlin, and E. Laure. Instrumentation for accurate energy-to-solution measurements of a Texas Instruments TMS320C6678 digital signal processor and its DDR3 memory. *E2SC '14 Proceedings of the 2nd International Workshop on Energy Efficient Supercomputing* (IEEE Press) pp 89-98 November 2014. http://dl.acm.org/citation.cfm?id=2689720

[58] J.D. McCalpin. STREAM: Sustainable memory bandwidth in high performance computers. https://www.cs.virginia.edu/stream/

[59] IBM. IBM Power7 data from IBM Power 750 and 760 Technical Overview and Introduction Literature Number 4985. May 2013.

[60] W.J. Starke. Power7: IBMs next generation, balanced POWER server chip. *HotChips*, August 2009. http://www.hotchips.org/wp-content/uploads/hc_archives/hc21/3_tues/HC21.25.800.ServerSystemsII-Epub/HC21.25.835.Starke-IBM-POWER7SystemBalancev13_display.pdf

[61] R. Krishnaiyer, E. Kultursay, P. Chawla, S. Preis, A. Zvezdin, and H. Saito. Compiler-based Data Prefetching and Streaming Non-temporal Store Generation for the Intel R Xeon Phi Coprocessor. *IEEE 27th International Parallel and Distributed Processing Symposium Workshops & PhD Forum* (NewYork, NY) IEEE pp 1575–1586 May 2013. http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=6651054&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D6651054

[62] Dell Performance and Power Efficiency of Dell Power Edge servers with E5-2600v2. October 2013.

[63] PRACE Final Report on Prototype Evaluation D9.3.4. December 2013.

[64] Texas Instruments. TMS320C66x DSP Cache *User Guide* Literature Number: SPRUGY8. November 2010. http://www.ti.com/lit/ug/sprugy8/sprugy8.pdf

[65] Top500 www.top500.org

[66] The HPC Challenge Benchmark http://icl.cs.utk.edu/hpcc/

[67] DGEMM http://www.netlib.org/lapack/explore-html/d7/d2b/dgemm_8f_source.html

[68] The Green 500 www.green500.org