Errata to
ENERGETIC: Final Report

Authors: Michael K. Bane, Andrew Kibats. Manchester Metropolitan University.
Date: 21 December 2023
Data and scripts: https://e-space.mmu.ac.uk/633613/

1. Introduction

During further work on the EXCALiBUR H&ES FPGA Testbed looking at FPGA performance it has come to light that the approach for monitoring the power of an Xilinx U280 FPGA card only looked at the PCI-express power rail and, regrettably, overlooked the AUX power rail. The only data where this is an issue in the original Final Report is regarding SGEMM energy consumption i.e. Figure 6 and its discussion. The correct methodology was applied to the U50 FPGA card (hosted at Newcastle) in all cases.

We have developed and applied an alternative approach, which we have confirmed with AMD/Xilinx as to its correctness. Whilst applying this new method in order to correct FPGA energy consumption, we also examined the host CPU energy consumption. We made use of the amd_energy kernel module [1] installed so could read counters directly for FPGA energy consumption and discovered significant difference to the previously published data.

Section 2 describes our approaches to measuring energy consumption of FPGA and host CPU during the execution of the SGEMM bitstream. Section 3 gives our updated results (see https://e-space.mmu.ac.uk/633613/), with Figure 2 illustrating the observed differences. Section 4 discusses how this new data amends some findings with Sections 5 and 6 giving our conclusions and plans for further work.

2. Measuring energy correctly for Xilinx Alveo FPGA cards

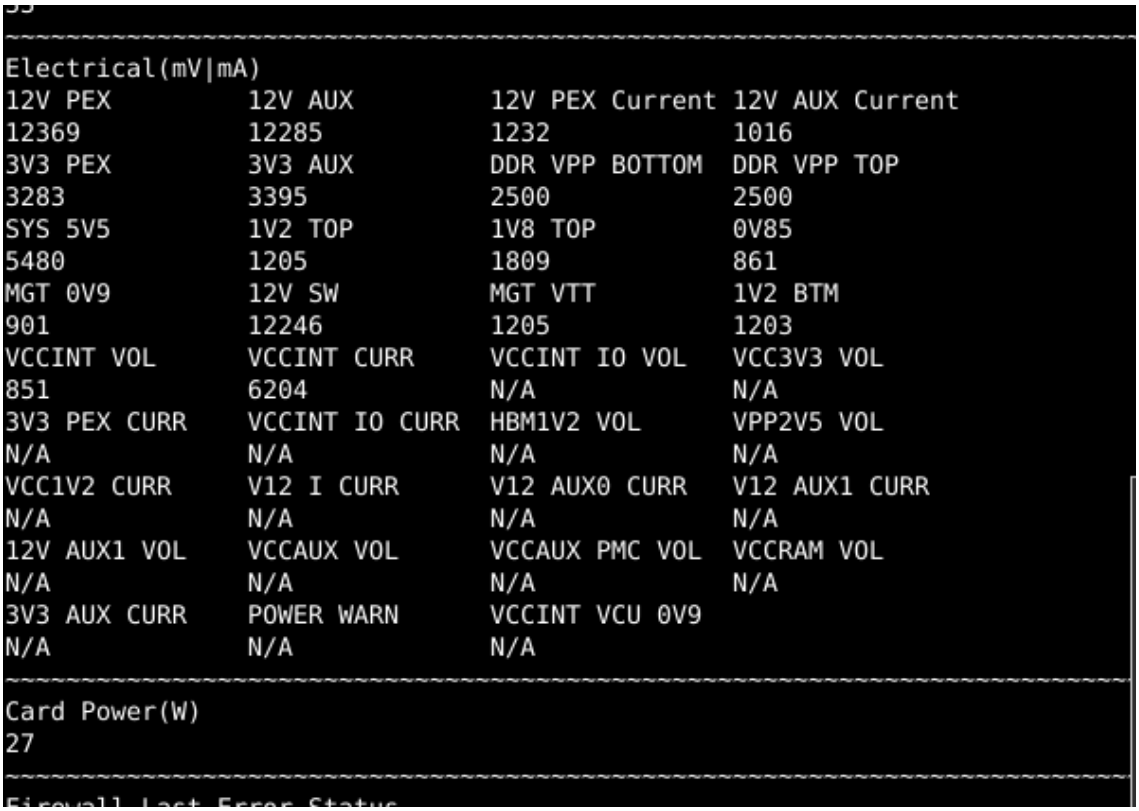
The AMD/Xilinx Alveo series of PCI-express mounted FPGA cards each have two power supplies, one directly via the PCI-e socket and another via an “auxillary” power lead. The Xilinx utility “xbutil” can be used to examine various electrical properties of an Alveo cards. This includes the voltage and current for the PCI-express and the voltage and current for the auxillary power supply. These are the PEX and AUX fields of the xbutil output, and output as milliVolts and milliAmps. xbutil also provides a power field, given as an integer value in Watts. These values are also reflected via counters within the relevant /sys/devices directory of the host. For example, on the nextgenio-amd01 node used for ENERGETIC experiments, the relevant directory is /sys/devices/pci0000:c0/0000:c0:01.1/0000:c3:00.1/xmc.u.19922944

The system used in this Errata is that used in the original Final Report, namely node “nextgenio-amd01” of the EXCALiBUR H&ES FPGA Testbed. This comprises a node of two AMD EPYC 7502 32-core processors hosting a Xilinx U280 card over PCI-express. We use version 2020.2 of the Xilinx Vitis Development Environment and version 2.11.634 (git branch 2021.1) of the Xilinx Run Time (XRT).

Figure 1 is example output from “xbutil” where we see the PCI-express has 12.369 Volts and draws 1.232 Amps, whereas the auxillary power supply is at 12.285 Volts and drawing 1.016 Amps. In the original Energetic Final Report [2], the overall power supply was computed by reading only the PCI-express values from the counters and multiplying to give the power. In the example of Figure 1, this would give 12.369*1.232=15.24 Watts which is clearly lower than the 27 Watts reported at the bottom of Figure 1. Adding the power draw of the auxillary supply (12.238*1.016=12.43 Watts) brings the total FPGA power draw to 27.67 Watts, which when truncated gives the 27 Watts reported by “xbutil”. It is thus clear that the

auxillary supply needs to be taken into account and that the total power draw of the FPGA card is significantly higher than reported in [2].

We further note that the counter
/sys/devices/pci0000:c0/0000:c0:01.1/0000:c3:00.1/xmc.u.19922944/xmc_power
gives the total power draw in microWatts. For our revised data, we use a script that records this counter at a given interval, and the time (in nanoseconds) this was read. This series of instantaneous power readings is then integrated (using trapezoidal rule) over the time measurements to give the energy consumption. Data is given to 3 decimal places.



```
Electrical(mV|mA)
12V PEX          12V AUX          12V PEX Current 12V AUX Current
12369            12285            1232            1016
3V3 PEX          3V3 AUX          DDR VPP BOTTOM  DDR VPP TOP
3283             3395             2500            2500
SYS 5V5          1V2 TOP          1V8 TOP         0V85
5480             1205             1809            861
MGT 0V9          12V SW           MGT VTT         1V2 BTM
901              12246            1205            1203
VCCINT VOL       VCCINT CURR      VCCINT IO VOL   VCC3V3 VOL
851              6204             N/A             N/A
3V3 PEX CURR     VCCINT IO CURR   HBM1V2 VOL      VPP2V5 VOL
N/A              N/A              N/A             N/A
VCC1V2 CURR      V12 I CURR       V12 AUX0 CURR   V12 AUX1 CURR
N/A              N/A              N/A             N/A
12V AUX1 VOL     VCCAUX VOL       VCCAUX PMC VOL  VCCRAM VOL
N/A              N/A              N/A             N/A
3V3 AUX CURR     POWER WARN       VCCINT VCU 0V9
N/A              N/A              N/A
Card Power(W)
27
```

Figure 1: Example output from from "xbutil" command

We also took this opportunity to review, for SGEMM only, the methodology of measuring the energy of the node hosting the FPGA card. We present energy efficiency of the various hardware implementations of SGEMM used within the original ENERGETIC Final Report in Figure 2. This is a composite diagram with Figure 2a being the original histogram from the report (i.e. Figure 6 of [2]), Figure 2b being a corrected plot of the original data from the project, and Figure 2c being a plot using corrected data.

The raw data used to construct the original histogram (Figure 2a) is available from the dataset [3] and had been entered into an intermediary Excel spreadsheet (not published) for which the relevant rows are given in Figure 3. Data from this spreadsheet was then put into CSV files for each benchmark and a Python script used to create the histograms given in [2]. Examination of the data in Figure 3 with respect to Figure 2 shows also that the CPU and accelerator energy consumption data has been transposed. We provide a corrected histogram (before we correctly measure the U280 energy) as Figure 4.

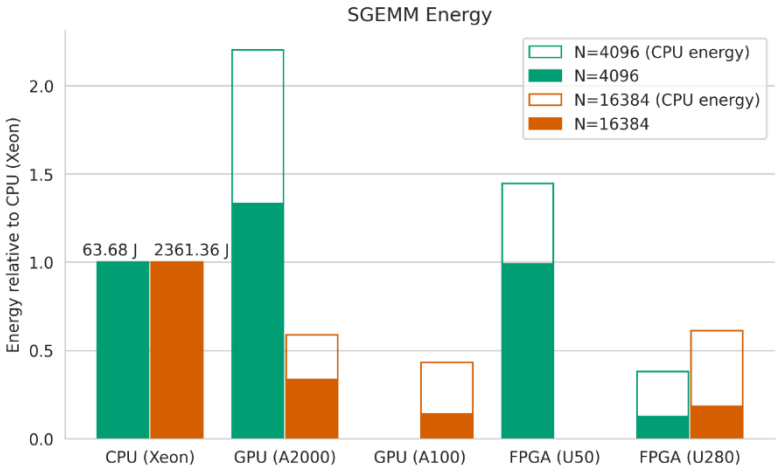


Figure 2a: Figure 6 of ENERGETIC Final Report showing relative energy consumption for the SGEMM benchmark using original energy measurement methodology

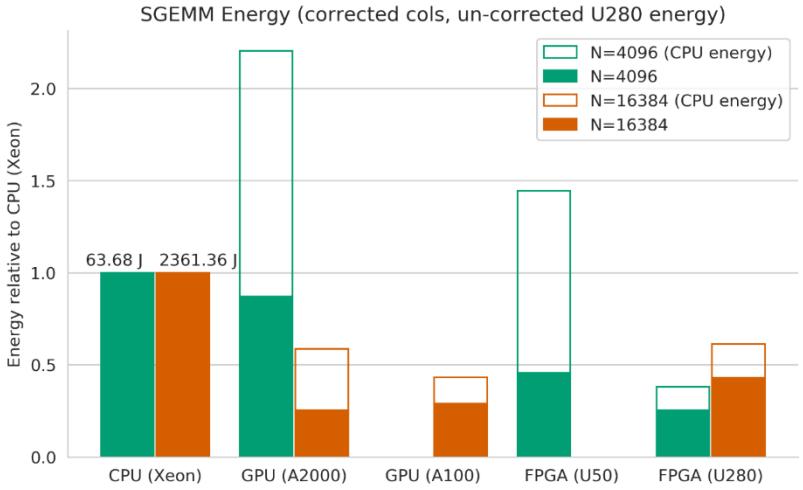


Figure 2b: Corrected histogram for SGEMM energy consumed as recorded by ENERGETIC project (prior to correcting U280 energy consumption)

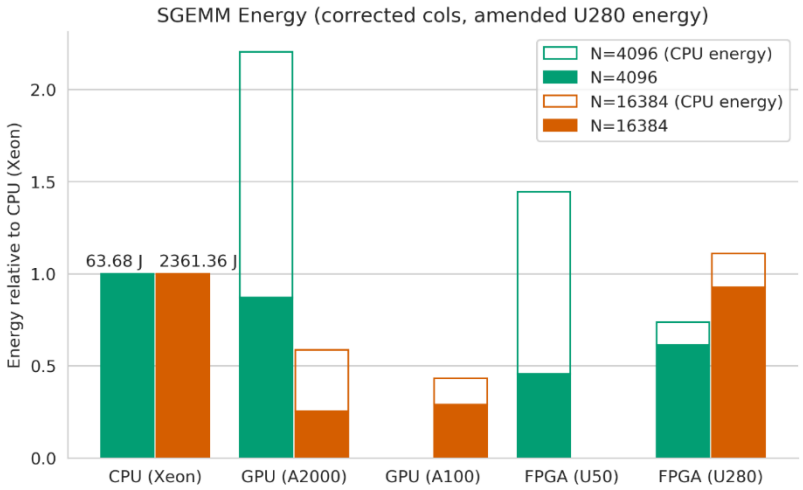


Figure 2c: Histogram of relative energy consumption of SGEMM with correct U280 energy measurements.

2				Total consumption data				
3	Benchmark	Accelerator	Other info	Total wallclock time / s	Measured at wall total energy / J	Sum of CPU & accelerator / J	Total CPU energy / J	Total accelerator energy / J
4	SGEMM (4096)	i9-11900KF	Matrix Size (m,n,k) = 4096 (RAPL)	0.93	130.79	80.89	80.89	0.00
5	SGEMM (4096)	A2000	Matrix Size (m,n,k) = 4096 (NVML)	1.24	139.74	140.27	84.82	55.45
6	SGEMM (4096)	Alveo U50	Matrix Size (m,n,k) = 4096 (xbutil)	1.71	245.06	92.02	62.97	29.05
7	SGEMM (4096)	(Xeon)	Matrix size (m,n,k) = 4096	0.78		63.68	63.68	0.00
8	**SGEMM**	**Alveo U28**	**Matrix size (m,n,k) = 4096**	**0.97**		**24.20**	**8.00**	**16.20**
9	SGEMM (16384)	i9-11900KF	Matrix Size (m,n,k) = 16384 (RAPL)	28.70	6312.90	3353.53	3353.53	0.00
10	SGEMM (16384)	(Xeon)	Matrix size (m,n,k) = 16384	19.56		2361.36	2361.36	0.00
11	SGEMM (16384)	A2000	Matrix Size (m,n,k) = 16384 (NVML)	12.27	1705.25	1387.66	790.72	596.94
12	SGEMM (16384)	NVIDIA A100	Matrix size (m,n,k) = 16384	0.50		1018.31	332.64	685.67
13	**SGEMM**	**Alveo U28**	**Matrix size (m,n,k) = 16384**	**58.15**		**1446.32**	**431.47**	**1014.85**

Figure 3: Spreadsheet data of CPU and accelerator energy consumption for SGEMM for each of 4K and 16K test cases.

Further, as we revised our methodology to measure the U280 energy consumption, we identified the energy consumption reported for the U280 host CPU to be significantly lower than our contemporary measurements for the total CPU host. Our contemporary measurements make use of amd_energy to obtain the per-socket energy consumption during execution. Typically for the 4K data set, we observer 120-130 Joules, compared to 8 Joules reported in Figure 3 which was obtained by use of LIKWID[4]. Discussion with AMD clarifies that the socket data provided via amd_energy approach is the sum of the energy used by "cores, I/O and memory control" [5]. This figure represents the total energy consumed by the host node. We also examined per-core energy consumption via amd_energy approach, and for 3 runs for the 4K data set (with 1 millisecond polling) we observed per-core energies ranging from 0 to 1224 milliJoules, with sum of the core energies being 6425, 6858 and 6674 milliJoules. The previous approach using LIKWID was a polling method with some overhead but (presumably) only measured the energy of cores and not the sockets.

Data available at https://e-space.mmu.ac.uk/633613/

3. Results using corrected energy measurements of FPGA

Using our revised methodology to calculate the power of the U280 FPGA card, we measured the energy consumption of SGEMM for each of the 4K and 16K input data set test cases of the original report, using a 1 millisecond polling, as per original report. For each measurement, we reset the FPGA card and ran SGEMM. This initial run includes ~10 seconds to load the bitstream so we ignore this data point. We then ran the given test 3 times and recorded the energy consumption of the CPU sockets and of the U280 FPGA card, as well as the overall time and kernel times taken. Table 1 shows our revised results (to be compared to Figure 3). If we presume fluctuations in CPU energy are due to (e.g.) OS operations, we take the minimum of these. Rounding to the nearest Joule, we have host CPU energy of 135 and 6999 Joules, FPGA energy of 39 and 2188 Joules with total energy of 173 and 97187 Joules for the 4K and 16K cases respectively.

Test ID	Matrix Size (N)	Run Time (S)	Kernel Time (S)	Total CPU Energy (J)	FPGA Energy (J)	Total Energy (J)
4k_1_l_1	4096	1.114	0.908	134.101	38.756	172.857
4k_1_l_2	4096	1.114	0.908	135.120	38.831	173.951
4k_1_l_3	4096	1.115	0.908	134.969	38.339	173.308
16k_1_l_1	16384	58.284	57.556	7053.616	2184.440	9238.056
16k_1_l_2	16384	58.286	57.556	7020.006	2185.639	9205.645
16k_1_l_3	16384	58.294	57.556	6998.936	2188.228	9187.164

Table 1: Revised energy consumption figures for SGEMM on U280

The FPGA energy consumption is roughly double that erroneously reported in the ENERGETIC final report, whereas the CPU energy consumption given there appears correct for the method employed in the original

project. We therefore, in this Errata can correct Figure 2b by keeping the original CPU energy data but using the corrected U280 energy data, and this is given in Figure 2c (using FPGA energy of 39 and 2188 Joules but all other data points remaining as per original report).

4. Discussion of Corrected Results

Firstly, we consider Figure 2b and Figure 2c. For these, the only change is for the FPGA energy consumption (all other data points remain unchanged) and we observe the FPGA energy consumption is noticeably higher, which is as expected since we were previously omitting to measure ~50% of the power supply to the card.

In terms of the meaning of these Figures, the original report commented that "when using the U280 FPGA that the CPU uses nearly three times as much energy as the FPGA itself, despite not performing any of the computation. Despite this, and despite the relatively long runtime of the U280 benchmarks, it still manages to use far less energy than the CPU-only benchmark." However, with revised data it is **not** true that for the U280 FPGA run that the CPU (as measured i.e. core-based) uses more than the FPGA. However, it is still true that for N=4K the CPU+FPGA uses less energy that the CPU Xeon, although this is no longer true for the N=16K case. Therefore, we now conclude that this FPGA implementation is not always more energy efficient than the baseline CPU (Xeon) implementation.

The U280 data for SGEMM was not quoted elsewhere in the original report. Figure 5 has the GPU A100 as the most energy efficient and thus the comparative figure (Figure 10) remains the same in the original report, as do the general conclusions.

5. Conclusions

We have shown a corrected methodology for measuring energy consumption of AMD/Xilinx Alveo FPGA cards. We have applied this corrected methodology to the sole benchmark ("SGEMM") run on the U280 during the ENERGETIC project and, as expected, observe a significantly higher energy consumption for the FPGA bitstream. This does not affect the overall conclusions, since U50 FPGA measurements had been correctly recorded for all benchmarks and the GPU A100 had been noted in the original report as the most energy efficient implementation for SGEMM at N=16K. The authors would emphasize the original report findings that the energy efficiency of codes depends partly upon the quality of the port, and this is most noticeably so for FPGA than for GPU than for CPU.

6. Further Work

During the revision of the data for this Errata a comprehensive study has been undertaken regarding measurement of energy consumption of AMD/Xilinx Alveo cards. As outlined above, we poll the xmc_power counter and integrate to find the energy consumed. We have examined the effect of varying the poll interval. Further, we have measured the host CPU energy consumption at both the per-core and per-socket level and found that per-socket is significantly greater than the sum of the per-core energy consumption. Clearly there is a philosophical discussion to be had regarding which is more appropriate for any given comparison. We give comprehensive data, for SGEMM and further benchmarks on both U280 and 250, and explore these points in-depth, in our future paper (in prep) to be made available from https://helward.mmu.ac.uk/STAFF/M.Bane/index.html.

7. References

[1] AMD. (2023) ‘amd_energy’, https://github.com/amd/amd_energy, accessed 28/11/2023.
[2] Bane, M.K. *et al*. (2023) ‘ENERGETIC: Final Report’, version 1.1, DOI:10.5281/zenodo.6787467, accessed 28/11/2023.
[3] Bane, M.K. *et al*. (2023) ‘ENERGETIC project’, DOI: https://doi.org/10.23634/MMU.00631226, accessed 28/11/2023.
[4] NHR, Likwid Performance Tools, https://hpc.fau.de/research/tools/likwid/, (2022), accessed 03/01/2023.
[5] AMD. (2023). Private correspondence. 17/11/2023.