Ongoing research towards Energy Efficient Performant Computing

Michael Bane
(December 2021)

Energy Efficient Computing Research at STFC Hartree Centre

Dr. Michael K. Bane, STFC Hartree WA4 4AD, UK

Bootstrapping EEC Research

ICT has a significant impact on the environment

- Operational emissions: Carbon released due to running data centres, PCs & embedded devices. Estimated that US-based cloud data centres to consume 73 *billion* kW-hrs by 2020 [1]
- Embodied emissions: Carbon released during manufacture and disposal
- Other environmental footprint aspects such as high use of water (washing PCBs) and use of exotic materials

Hartree's EEC Research

	Funder & Time	Project Aims	Hartree EEC Aims
COMPAT [2]	H2020 2016-18	energy efficient placement of components of multi-scale codes	Requires ability to MEASURE & PREDICT
TSERO [3]	Innovate UK 2014-17	use of Machine Learning to determine compiler flags leading to lower energy-to-solution; instrumentation of data centre	Requires ability to MEASURE; aims to REDUCE
VINEYARD [4]	H2020 2016-19	quantification of emerging tech alternatives to CPU for lowering energy-to-solution	Requires ability to MEASURE & MONITOR
Energy Efficient HPC Working Group [5] & ETP4HPC [6]	-	working with leading HPC providers to understand & tackle challenges of energy efficient supercomputers & data centres	Requires ability to MONITOR & PREDICT
EuroEXA [7]	H2020 2017-20	aiming to build an exascale prototype	Requires ability to PREDICT & REDUCE

The EEC research group aims to work with manufacturers and industry to provide solutions that enable

1. Processor/chipset to run any given code with lower amounts of energy
2. Any given code to run with least amount of energy on any given platform
3. Every data centre to be more efficient in the running of user codes

In order to tackle these, the EEC group is exploring each level of its mantra

measure – monitor – predict – reduce

as applied to energy consumed.

Expected energy savings: due to the wide nature of codes (& their current energy optimisation) etc it is hard to quantify predicted savings but we look to save 20%.

References

[1] Muhammad Zakarya, Lee Gillam, "Energy efficient computing, clusters, grids and clouds: A taxonomy and survey", Sustainable Computing: Informatics and Systems, 14, 2017

[2] http://www.compat-project.eu/

[3] http://tsero.org/

[4] http://vineyard-h2020.eu/en/

[5] https://eehpcwg.llnl.gov/

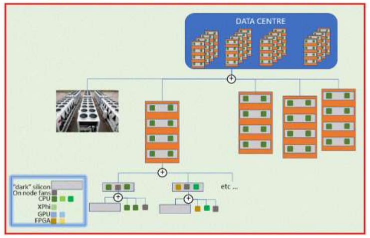
[6] https://euroexa.eu/

[7] http://www.etp4hpc.eu/

Hartree Centre
Science & Technology Facilities Council
Dr. Michael K. Bane
michael.bane@stfc.ac.uk

http://community.hartree.stfc.ac.uk/portal/site/eecrp

Provision of *HIERARCHICAL* Measurements...

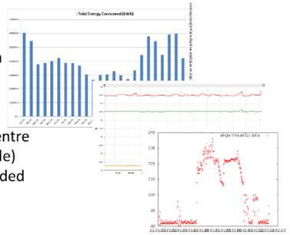


..for compute, network and storage, supplemented by an array of temperature and humidity sensors and data from schedulers.

The hierarchy of measurements needs to consider two axes:

- **Functional Resolution** - where energy is being consumed; represented as a tree with the 'root' measurement being the total energy consumed (by the data centre)
- **Accuracy & Temporal Resolution** - the accuracy of a given "leaf" measurement and its temporal resolution; summing over leaves gives error at any required level

We will empower researchers and data centre managers to *"drill down"* to the most appropriate level eg from data centre (blue), to rack (middle) and to a single threaded application (red).



A trusted measurement system

- ✓ to ensure monitoring is reliable for its chosen purpose
- ✓ to enhance models to accurately predict energy consumed under various scenarios
- ✓ to provide faith that *in silico* explorations of energy savings within a chip, a rack or a data centre
- ✓ to empower policy makers (energy caps & charging models)

A hierarchy of measurements but also of energy savings

Current research ideas to reduce energy consumed at each level of the hierarchy include

- modelling of data transfer and cache coherence protocols: to reduce energy consumed at the chip-set level
- extending TSERO by inclusion of expert knowledge to reduce ML requirements (number of input data points), increase the average energy saved: to reduce energy for any given code
- extension of batch schedulers to include DVFS, automatic power off of "unwanted" nodes, and migration to more energy efficient platforms.
- lessons from social science to ensure acceptance of energy caps

- STFC Hartree Centre (2016 – 2018) Energy Efficient Computing (EEC) Research
  - "TSERO"
  - "Vineyard"
- University of Liverpool (2018-2021)
  - teaching focussed
  - research
    - energy measure of local HPC (Barkla, 160 nodes)
    - Ryan L, ML to predict energy measurements
    - use of FPGA in EEC
    - potential role of quantum computing

https://helward.mmu.ac.uk/STAFF/M.Bane/MSc/

Energy Efficient Performant Computing

- Performant computing
  - Doing simulations & analyses
    {faster, higher resolution, larger domains, …}

- Energy Efficient Performant Computing
  - Undertaking performant computing whilst
    reducing energy consumption
    (without un-acceptable adverse implications on
    e.g. execution times)

- Involves
  - HPC, energy measure/predict/reduce,
  - social science

ukri.org/news/ukri-targets-net-zero-computing/

EEC FPGA grants EMiT HEC MKB mail Azure AI DL ML languages Ipool pol

UKRI targets net zero computing

ipcc.ch

MMU EEC FPGA grants EMiT HEC MKB mail Azure AI DL ML languages Ipool pollensa2021

MENU ABOUT DATA DOCUMENTATION FOCAL POINTS PORTAL LINKS HELP BUREAU PORTAL LIBRARY LANGUAGES SEARCH

ipcc REPORTS SYNTHESIS REPORT WORKING GROUPS ACTIVITIES NEWS CALENDAR FOLLOW SHARE

The Intergovernmental Panel on Climate Change

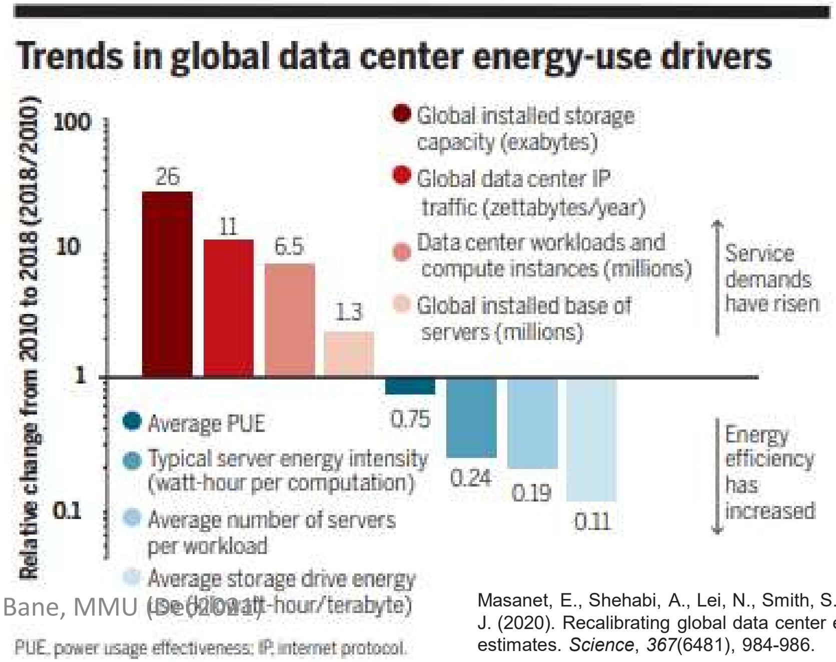
The Intergovernmental Panel on Climate Change (IPCC) is the United Nations body for assessing the science related to climate change.

mmu.ac.uk/sustainability/strategy

MMU EEC FPGA grants EMiT HEC MKB mail Azure AI DL ML languages

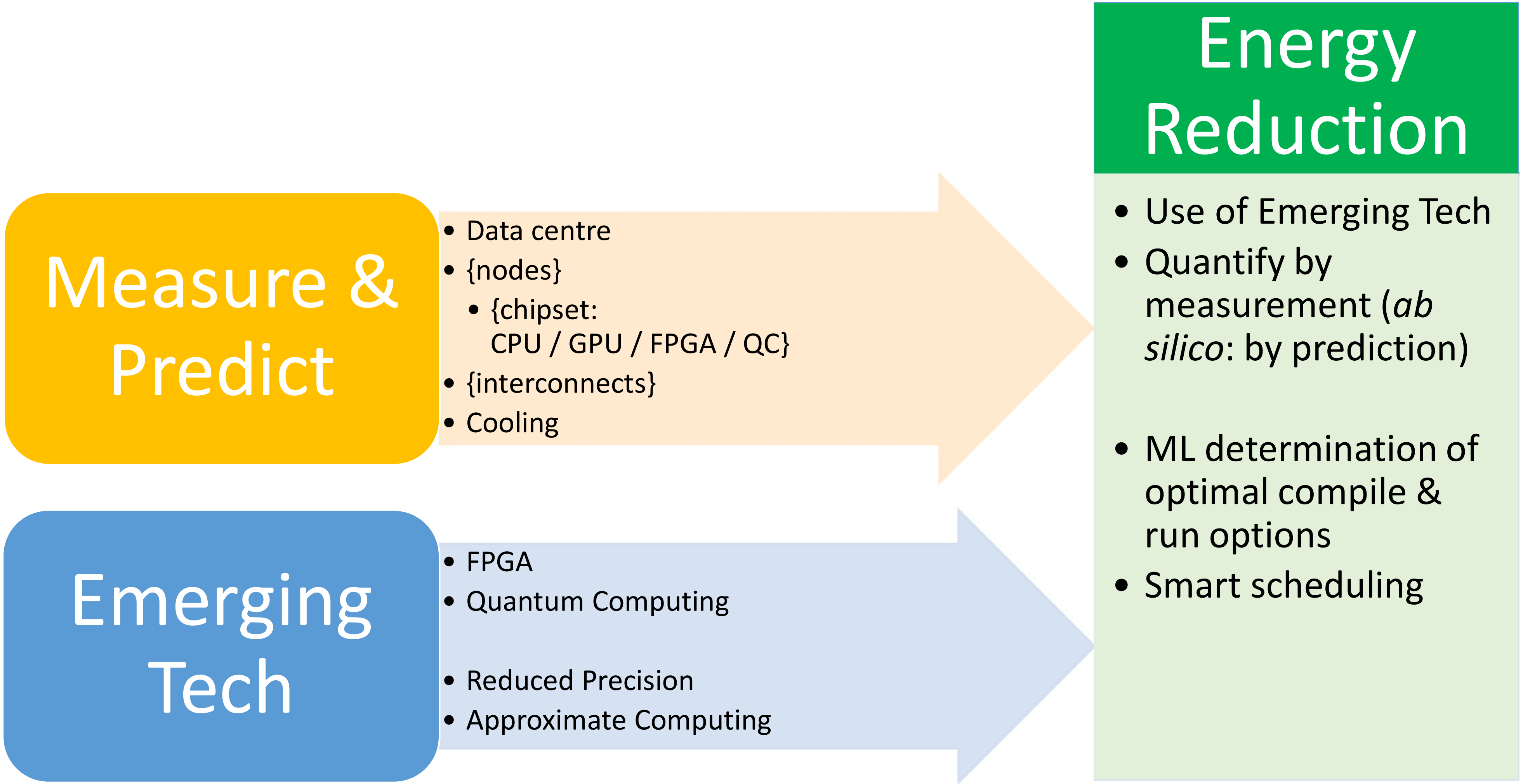
We're committed to becoming a **zero carbon University by 2038** for our scope 1 and 2 carbon emissions.

This not only mirrors Manchester City Council's own zero carbon pledge, but also supports the wider aim to make Manchester one of the world's leading cities for responding to the climate emergency. Because we believe that it's when we work together – every organisation and every individual – that we can make the biggest impact. As such, we have developed the first of three plans - our **Carbon Management Plan 2020-2026**, which sets out the steps we're taking towards being a zero carbon University by 2038.



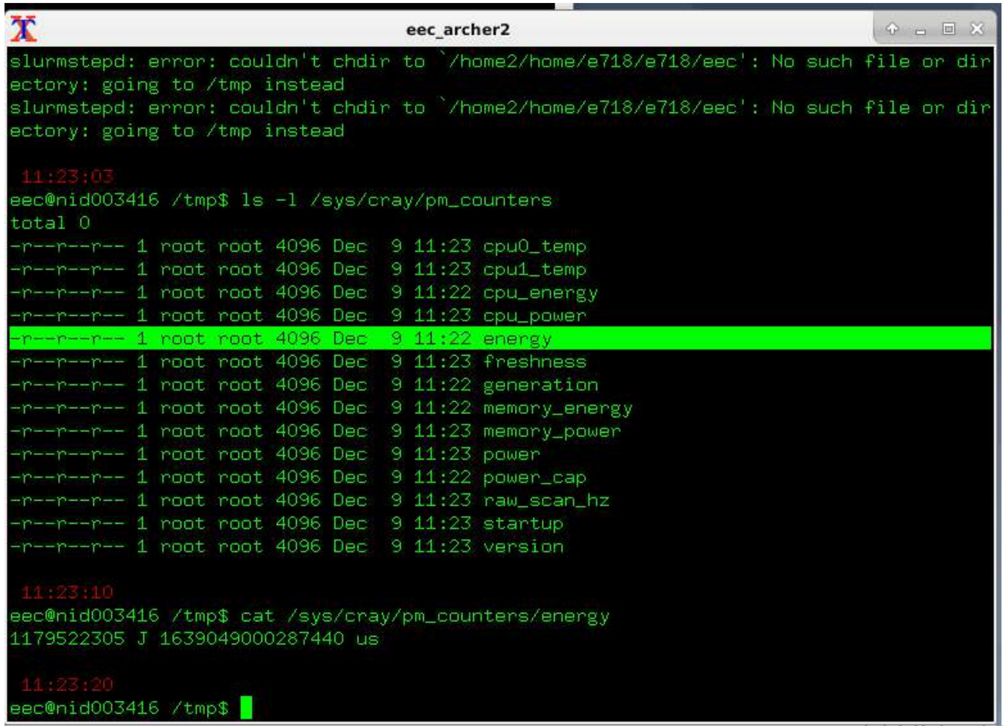
2010	194 TWh
2018	205 TWh (1% global electricity use)

Masanet, E., Shehabi, A., Lei, N., Smith, S., & Koomey, J. (2020). Recalibrating global data center energy-use estimates. *Science*, *367*(6481), 984-986.


Measure & Predict
Data centre
{nodes}
{chipset: CPU / GPU / FPGA / QC}
{interconnects}
Cooling
Emerging Tech
FPGA
Quantum Computing
Reduced Precision
Approximate Computing
Energy Reduction
Use of Emerging Tech
Quantify by measurement (*ab silico*: by prediction)
ML determination of optimal compile & run options
Smart scheduling

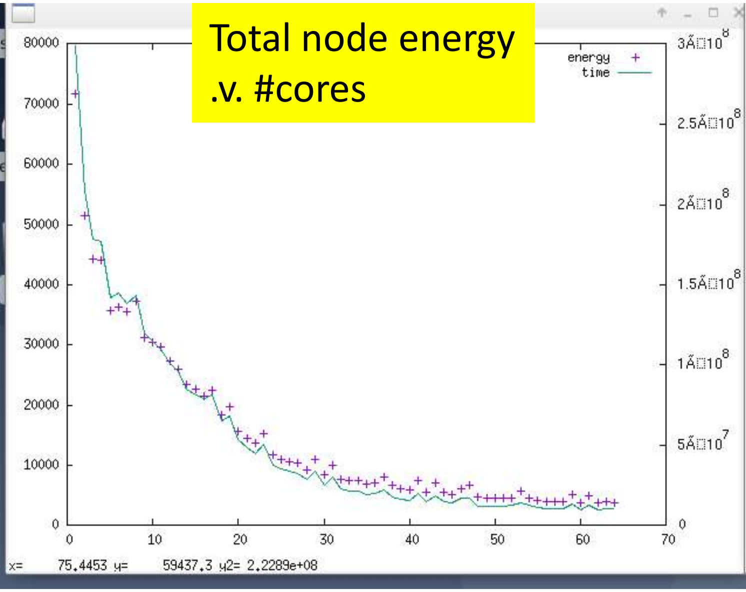
Measure & Predict

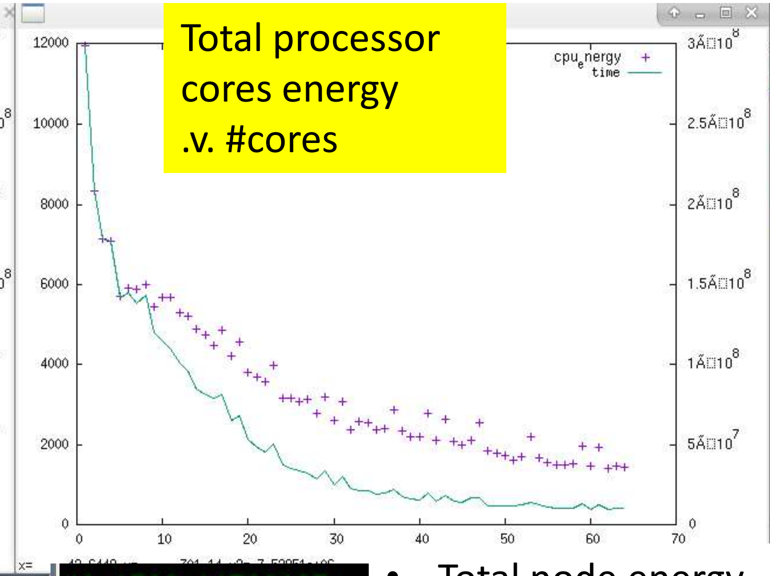
- Data centre
- {nodes}
  - {chipset:
    CPU / GPU / FPGA / QC}
- {interconnects}

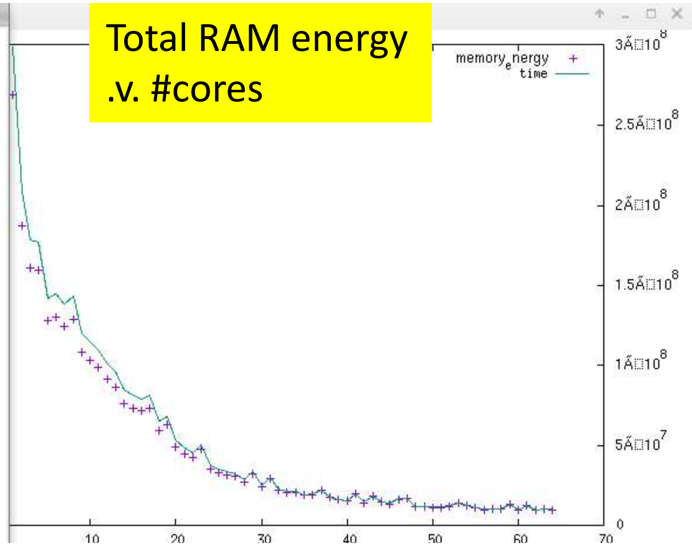


- Example:
  - Archer2 compute node
    2* AMD Rome EPYC chips (each of 64c)
  - Access to
    Performance Monitor Counters (PMCs)
  - Many tools to profile *time*
  - No tools to profile *energy*
  - Energy consumed, $E = \int P(t)dt$

“cloverleaf” Energy consumption on Archer2







For given optimisation, this code

- generally goes faster with more cores
- generally uses less energy with more cores

BUT

- least energy is not the fastest

49	1788	11790887
50	1738	11705346
51	1630	11887457
52	1691	12504948
53	2198	14185991
54	1672	13004578
55	1557	11439211
56	1506	10398060
57	1512	10503680
58	1539	10730837
59	1970	13083094
60	1464	9793602
61	1928	12605467
62	1420	10075778
63	1465	10793133
64	1448	10406323

(END)

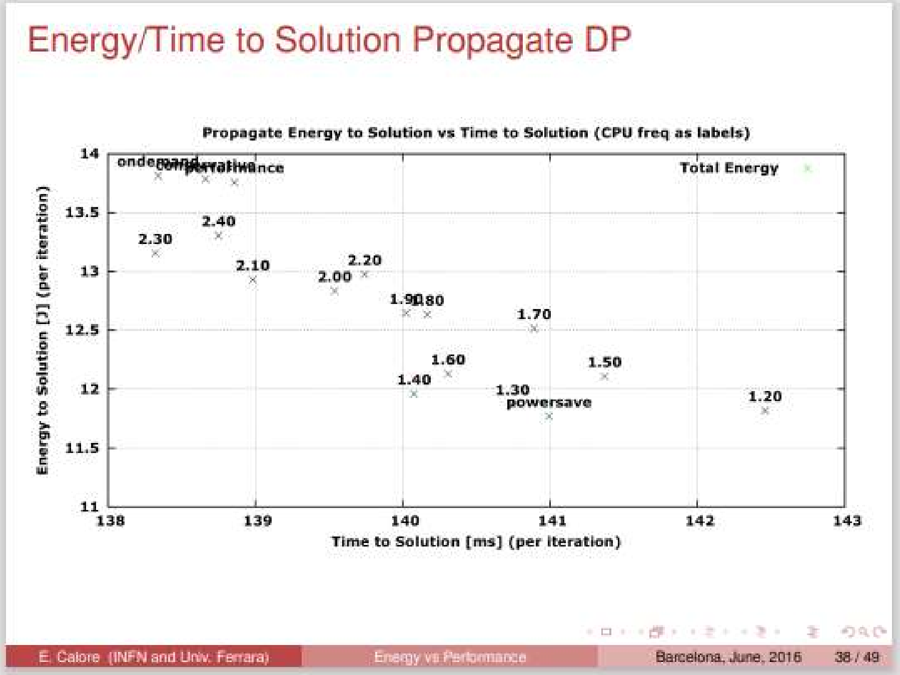
- Total node energy
  = Processor cores + RAM + “dark silicon”

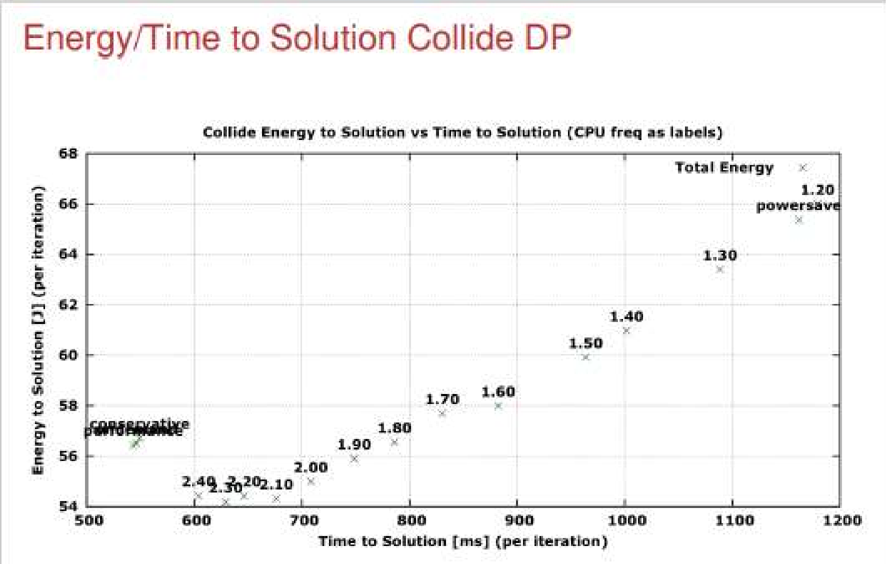
For 62c

- Node energy 3.83 kJ
- Processor cores 1.42 kJ, RAM energy 1.36 kJ

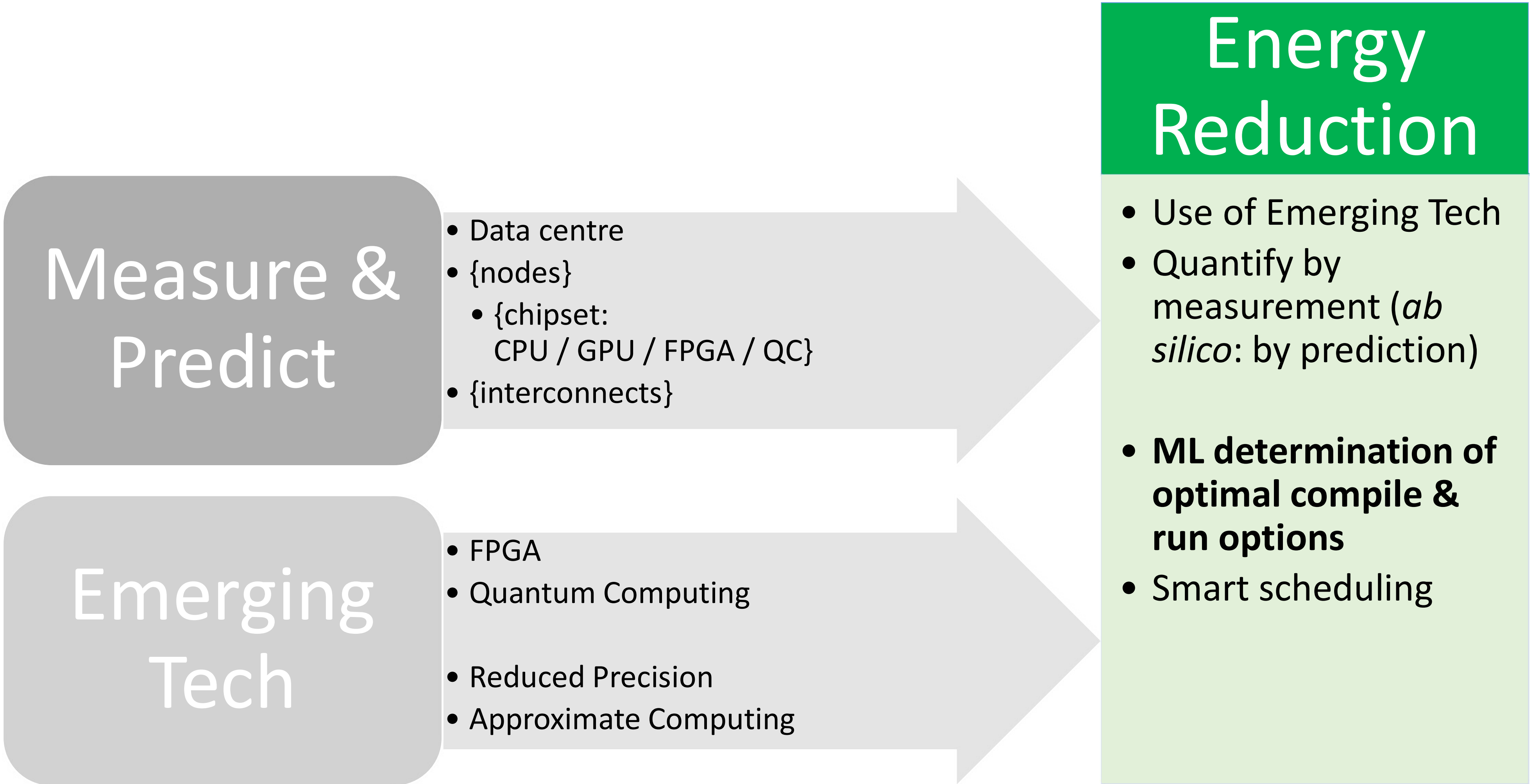
==> “dark silicon” 1.05kJ

Average power (for node) = 3,830 / 10.08 = 380 Watts
Archer2 has 5,860 nodes ==> 2.2 MW
At 10 pence per kW-hr ==> compute electricity costs c. £2M/year (+ cooling at ~10-20%). *All savings help!*





- Calore et al (2016)
  - Energy-performance trade-offs for HPC applications on high-end and low power systems, EMiT2016
  - Haswell chip, LBM
  - Freq giving least energy varies by problem type
    - Not the fastest in either case


Measure & Predict
Data centre
{nodes}
{chipset: CPU / GPU / FPGA / QC}
{interconnects}
Emerging Tech
FPGA
Quantum Computing
Reduced Precision
Approximate Computing
Energy Reduction
Use of Emerging Tech
Quantify by measurement (ab silico: by prediction)
ML determination of optimal compile & run options
Smart scheduling

ML Tool to Improve EEC

- For given code, coarse controls: compiler options, run time options
  - GCC .v. Intel
  - Level of optimisations (-O0, -O1, ...)
  - #cores, OpenMP .v. MPI implementations

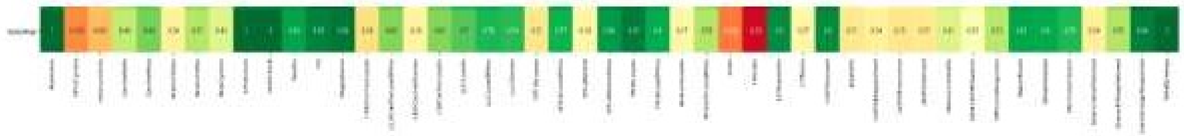
- AIM: for given input code, determine set of compiler & run time options (for given ISA) that gives lowest energy-to-solution

ML Tool to Improve EEC

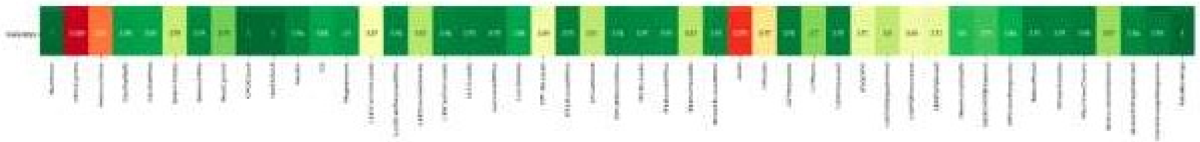
- Need: energy-to-solution
  - Currently, require ability to measure
  - Future: develop accurate predictor

- Training
  - Run set of benchmarks for various compiler & run time options, recording energy to solution; 13 benchmarks
  - 50 features of code (via ‘perf’)
  - Use PCA to select 20 most relevant ({Pearson, Spearman, Kendall} rank coeffs)
    - Feature selection → reduce overfitting, reduce time to train
  - ML options: linear regression, random forest regression
    - Python, sklearn

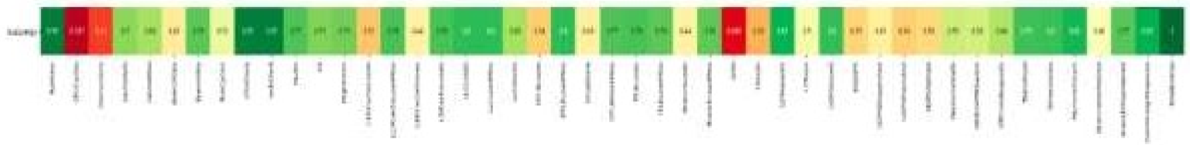
Pearson's Correlation Coefficient heatmap:



Spearman's Rank Coefficient heatmap:



Kendall Rank Coefficient heatmap:



- Red (low correlation) to green (+correlation)

- Considered perf features:
  - Run time
  - CPU clock [walltime]
  - Task clock
  - CPU (task) Migrations
  - I TLB loads
  - Context Switches
  - Micro Ops Issues
  - L2 requests
  - D TLB store miss
  - Faults
  - CPU cycles
  - Arith
  - Instructions
  - L1 D Cache Loads

ML Tool to Improve EEC

- Initial testing
  - 20% of initial dataset
  - Average diff of energy consumed (predicted .v. actual)
    Linear regression = 10.0%
    Random Forest reg = 1.8%

- Implementation
  - Comparison of best {compiler options, run time options} to baseline of {{GCC, “-O0”}, 40 OpenMP threads}

N.B. previous work, on 2*20c Intel Skylake processors per node of U/Liverpool “Barkla” (160 nodes + 20 GPUs)

ID	Benchmark	Total energy consumption		Suggested optimal settings	Total energy consumption	Energy efficiency gain/loss
39	Mantevo Cloverleaf	27.15 Joules		GCC, O0, 4, OpenMP	11.43 Joules	57.90%
359	HPC Challenge Linpack	497.20 Joules		ICC, O2, 40, OpenMP	497.20 Joules	0.0%
399	Mantevo MiniAero	3099.58 Joules		ICC, O3, 27, MPI	3086.47 Joules	0.42%
439	Mantevo MiniMD	276.08 Joules		GCC, O2, 39, OpenMP	198.78 Joules	28.00%
599	NASAPB DC.A	23337.55 Joules		ICC, O2, 21, OpenMP	20550.26 Joules	11.94%
2013	NASAPB IS.C	438.50 Joules		GCC, O3, 39, OpenMP	276.07 Joules	37.04%

Not full node

Not Intel compiler

Not full optimisation

Average: 22.5% energy saving

ML Tool to Improve EEC

- Next steps
  - Improvements to code base [CfACS seed funding]
    - Modularise (csv input);
    - Investigate/implement *static* code analysis for given code
    - Automate prediction of settings that give least energy to solution
  - Current results from Intel Skylake platform
    - Training on more platforms
    - Test/implement per-platform
    - Test/implement x-platform *including* *GPU & FPGA options*

Recent bid (with Glasgow) to UKRI Excalibur "h/w & enabling s/w"

ML Tool to Improve EEC

- Next steps
  - Current results from Intel Skylake platform
    - Training on more platforms
    - Test/implement per-platform
    - Test/implement x-platform *including GPU & FPGA options*
  - More fine grained compiler options

Measure & Predict

- Data centre
- {nodes}
  - {chipset: CPU / GPU / FPGA / QC}
- {interconnects}

- **Predictors**
  - Current predictors focus on time
  - Work with collabs (Glasgow) to predict energy consumed

- Incorporate within training of ML tool

➔Ability to predict (for given code) what would be best arch and compiler options for least energy (==> smart x-platform scheduler) *without having to expend compute energy in doing so*

- super optimisation for energy reduction
  - exhaustive search of all possible ISA instructions (for given basic block of code), using predictor to evaluate energy cost of each option ==> global minimum of energy-to-solution
- *selected* super optimisation for energy reduction
  - S.O. for E.R *and* make use of ML to sensibly prune search tree

Measure & Predict

- Data centre
- {nodes}
  - {chipset: CPU / GPU / FPGA / QC}
- {interconnects}

Emerging Tech

- FPGA
- Quantum Computing

- Reduced Precision
- Approximate Computing

Energy Reduction

- Use of Emerging Tech
- Quantify by measurement (*ab silico*: by prediction)

- ML determination of optimal compile & run options
- **Smart scheduling**

Data Centres

- How reduce carbon footprint?
  - Use of renewables & making use of waste heat
    - Location location location
  - LUMI

  - Smart scheduling
    - Don’t run what don’t need to run
      (re-use data, reproducibility / repro repositories,
      AI checking on job)
    - Only run vital jobs during ‘peak power’ times (e.g. standard jobs run when ambient temp drops so less cooling required)
    - *Social science element*
    - How integrate “between” data centres?

  - How green is the cloud…?

Emerging Tech

- **FPGA**
- Quantum Computing
- Reduced Precision
- Approximate Computing

- FPGA
  - Low power
  - Not easy to program
    - C/C++ with pragmas, Verilog, VHDL
  - Previous research
    - Porting linear algebra & fintech to FPGA
    - [energy results?]
  - Next steps
    - Reduced / variable / mixed precision (w. Manchester, Sorbonne)

  - Xilinx University Programme: card, training, workshop

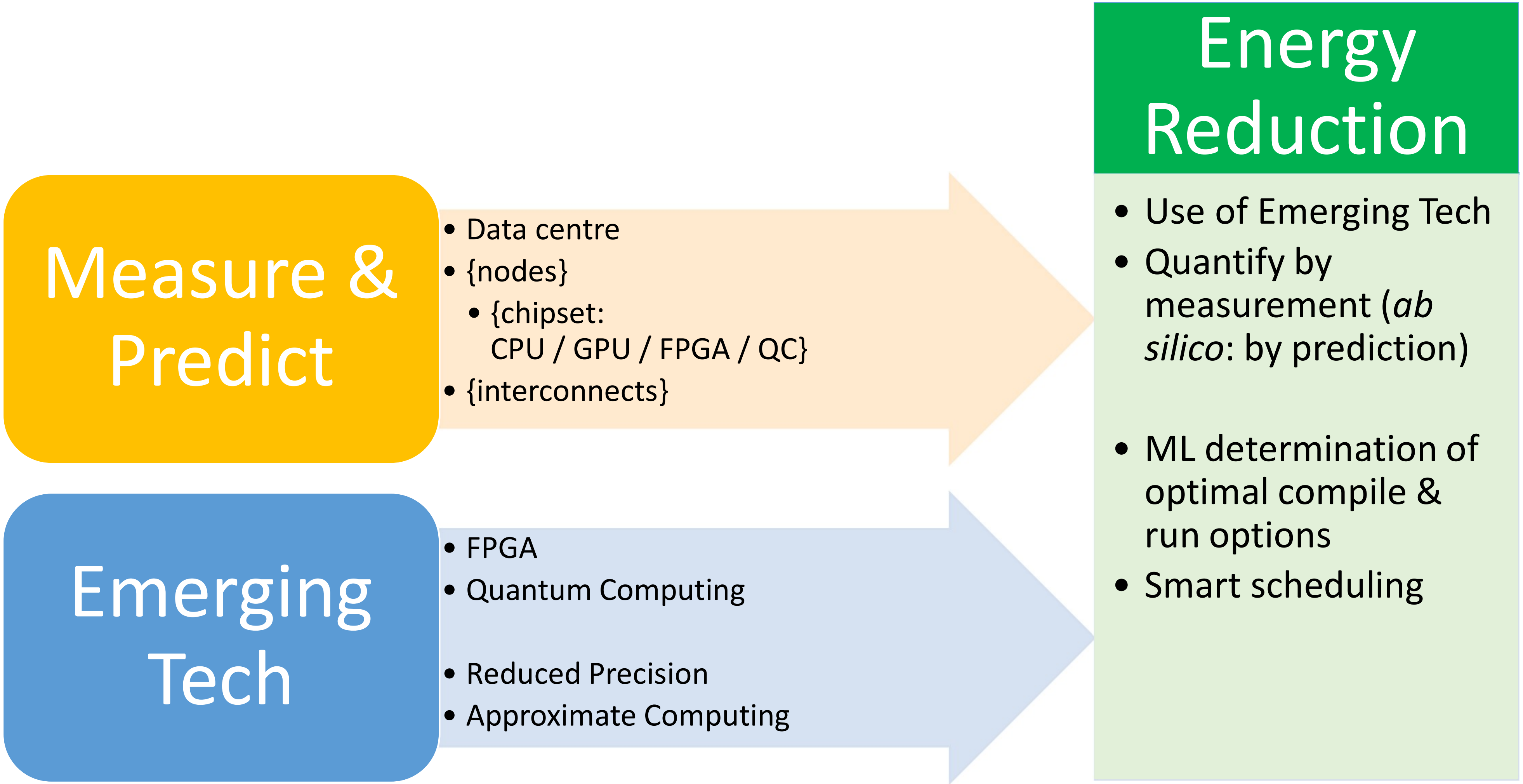
Emerging Tech

- FPGA
- **Quantum Computing**
- Reduced Precision
- Approximate Computing

- Quantum Computing (QC)
  - Hype or reality?
  - Energy efficient or vastly inefficient [explain]

- Recent bid
  - QCS to evaluate use of QC to simulate gas/liquid phase changes on atmospheric aerosol
  - Potential partnership: Zapata Computing
  - If want access, contact me

One of a few ongoing collaborations with U/Manchester (other e.g. use of Big Data to analyse aerosol from coughs)


Measure & Predict
Data centre
{nodes}
{chipset:
CPU / GPU / FPGA / QC}
{interconnects}
Emerging Tech
FPGA
Quantum Computing
Reduced Precision
Approximate Computing
Energy Reduction
Use of Emerging Tech
Quantify by measurement (*ab silico*: by prediction)
ML determination of optimal compile & run options
Smart scheduling

Michael Bane
m.bane@mmu.ac.uk
E140, John Dalton East