Energy-efficient data centers with Millicomputing opportunity and challenges

Pascal Bouvry, Frédéric Pinel

- Context
- Energy optimisation for data center
- Introduction to millicomputing
- Benchmarking millicomputing for data centers
- Application choice
- Conclusion

ICT in Luxembourg:

- An European hub (Dark fiber)
- Large set of data centers (5 Tier 4)
- Specific Laws and Tax policies
- Fund management/Banking Center
- EU HQ of Amazon, iTunes, Paypal, Ebay, AoL, Netflix, etc.
- Other: Skype, Goodyear, gaming, etc.

ICT – Security, Reliability, Trust is part of the country strategic priorities at national and UL levels.
Systems Biomedicine is also part of the strategic priorities.

FNR funded projects in the related to GreenIT:

- GreenIT (2009-2012)
- Green@cloud (2012-2015), joint funding with CNRS (Polytech Lille)
- EcoCloud (2013-2015)

CAPEX:	OPEX:
Building	Personnel
Racks	Electricity
Machines	
Software	

Green is more than energy/cost savings,
but cost saving is the easiest way to buy corporates in.

1. Optimising the PUE (Power Usage Efficiency)
   - Free cooling/heat reuse
   - Fluid dynamics analysis
2. Optimising the dynamic allocation
   - Virtualisation/Server consolidation
   - Smart schedulers using
     . On-Off
     . DVFS (Dynamic Voltage and Frequency Scaling)
3. Modeling, simulation, testbed
   . Holistic approach (computing, storage, communications)
   . Greencloud
   . Grid5000

Re-use low-power solutions from another industry

Smartphones

(http://blogs.norman.com/)

A questionnable approach

- “It’s a toy, not enterprise-ready”
- “It can’t do big I/O”, “It doesn’t have big memory”
- “Its more efficient to manage fewer bigger machines”

Miniaturization is on its way:

- Mainframes replaced by Minicomputers
- Minicomputers replaced by RISC servers
- RISC servers replaced by PC servers

acockcroft@netflix.com, HPTS 2007

- 1 billion computers* shipped since 1975...

→ Over 2 billion ARM processors shipped last quarter !!!

- Large market volumes trigger innovation
- Large market volumes lower processor costs
- Price/performance ratio favors *clusters of commodity computers*

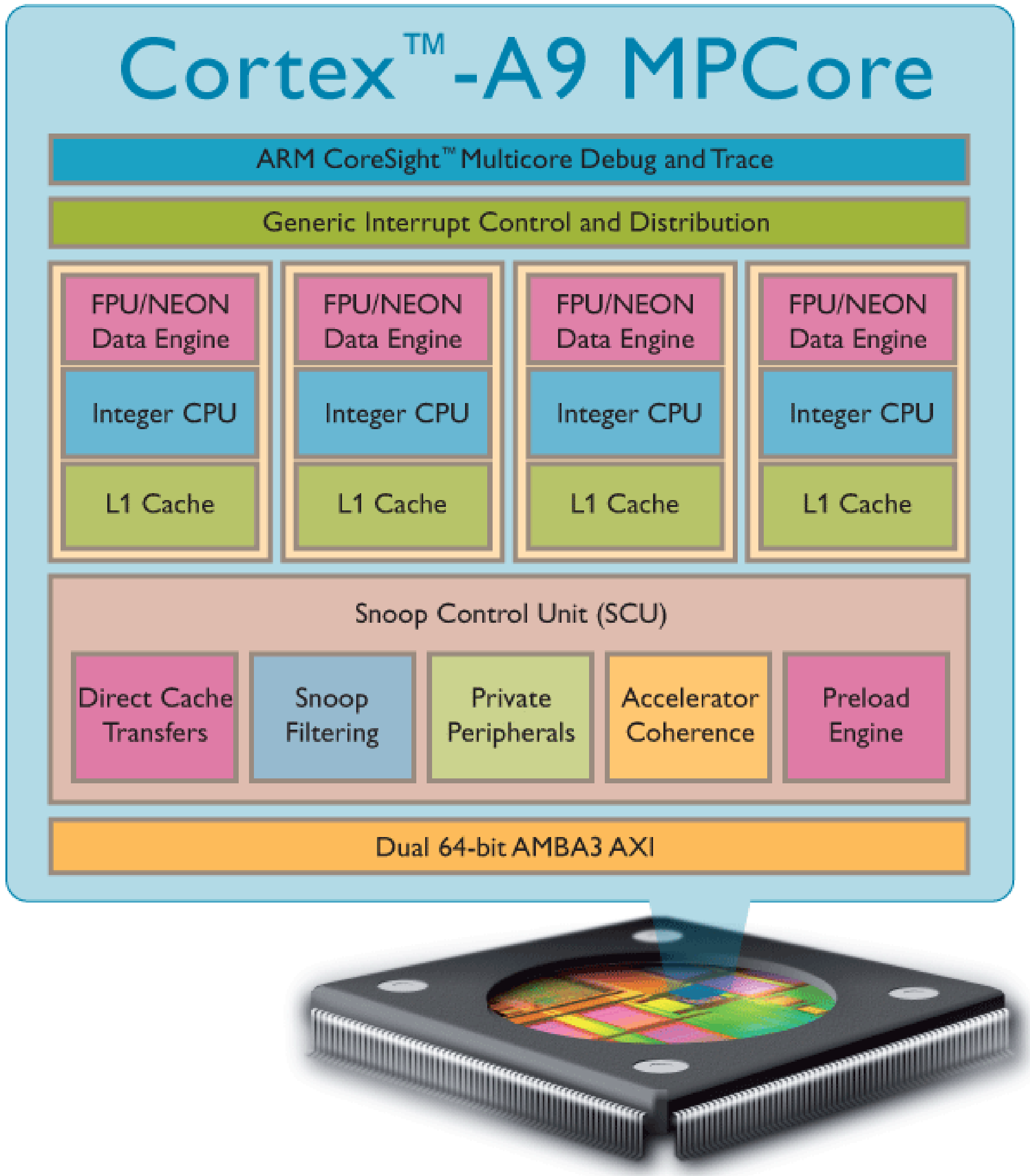
- ARM processors can be found in products ranging from the LG Viewty, Nokia N95 and Sony Ericsson P1i smart phones; the iPhone and the iPod; Garmin, Navman and Tom Tom portable navigation systems; Kodak still cameras; Sony video cameras; and the Nintendo DS handheld gaming device, up to Toshiba HD digital televisions; hard disk drives from Samsung and Seagate; automotive braking systems from Bosch; HP printers and wireless routers from Linksys and Netgear.

Millicomputing aims at

- Diminishing the CAPEX, cutting down machine expenditures
- Diminishing the OPEX, cutting down the electrical bill

- ARM supplies majority of 32-bit processors
- 32-bit RISC
- Low power components
- ARM Cortex A9
  0.4 – 1.9 W (2 GHz)

- Common software stack,
  Gnu/Linux environment



(http://www.usporedi.hr)

Raspberry Pi (40 Euros)
ARM 11 (76JZFS) 700 MHz RAM 512 MB

Switched On Tech Design
(www.sotechdesign.com.au)

ARM computer

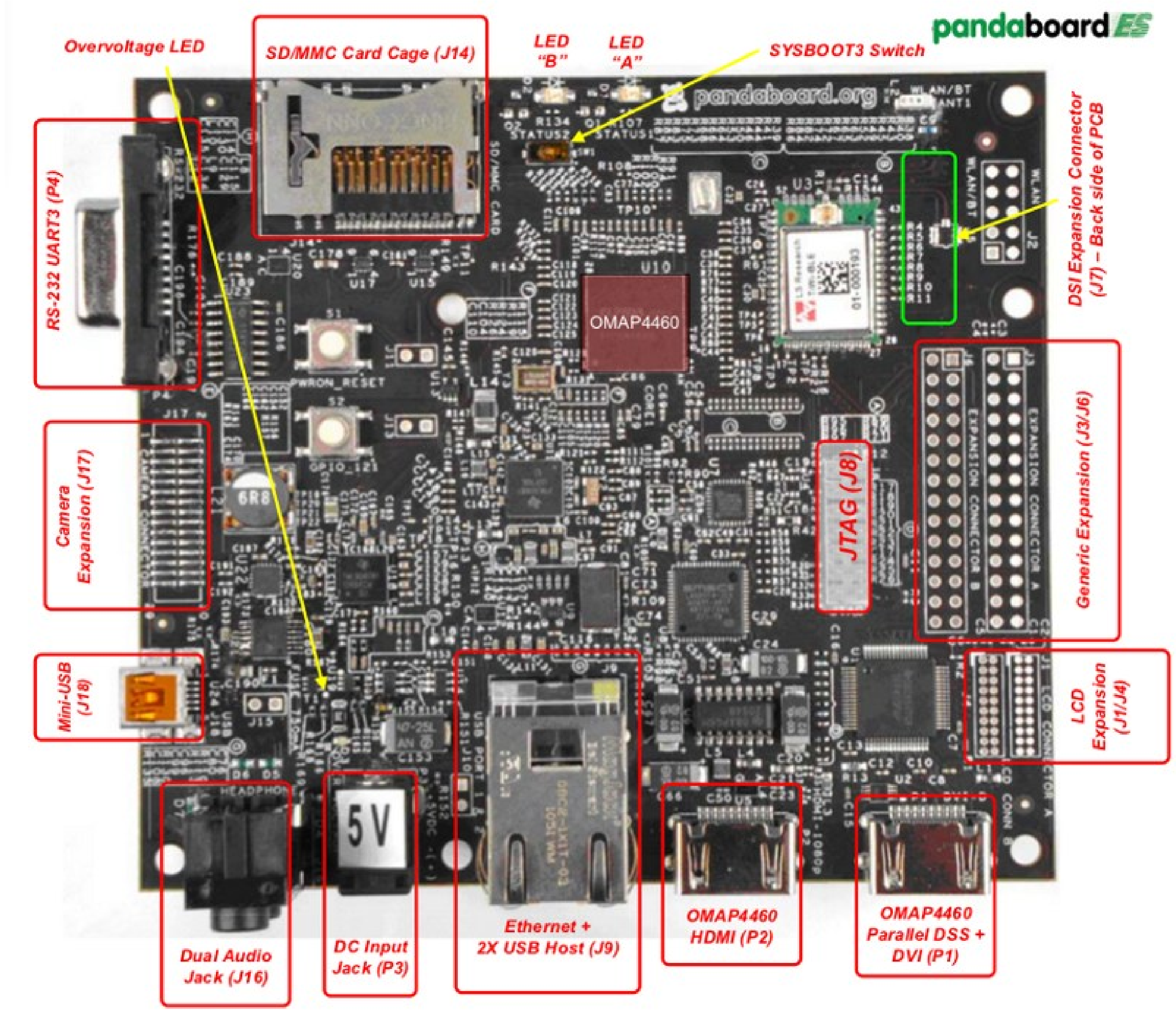
- Pandaboard (ES)
- Open mobile software development platform
- TI OMAP 4460
  Dual core A9

1.2 GHZ

RAM 1 GB DDR2
(low power)

10/100 Mb ethernet

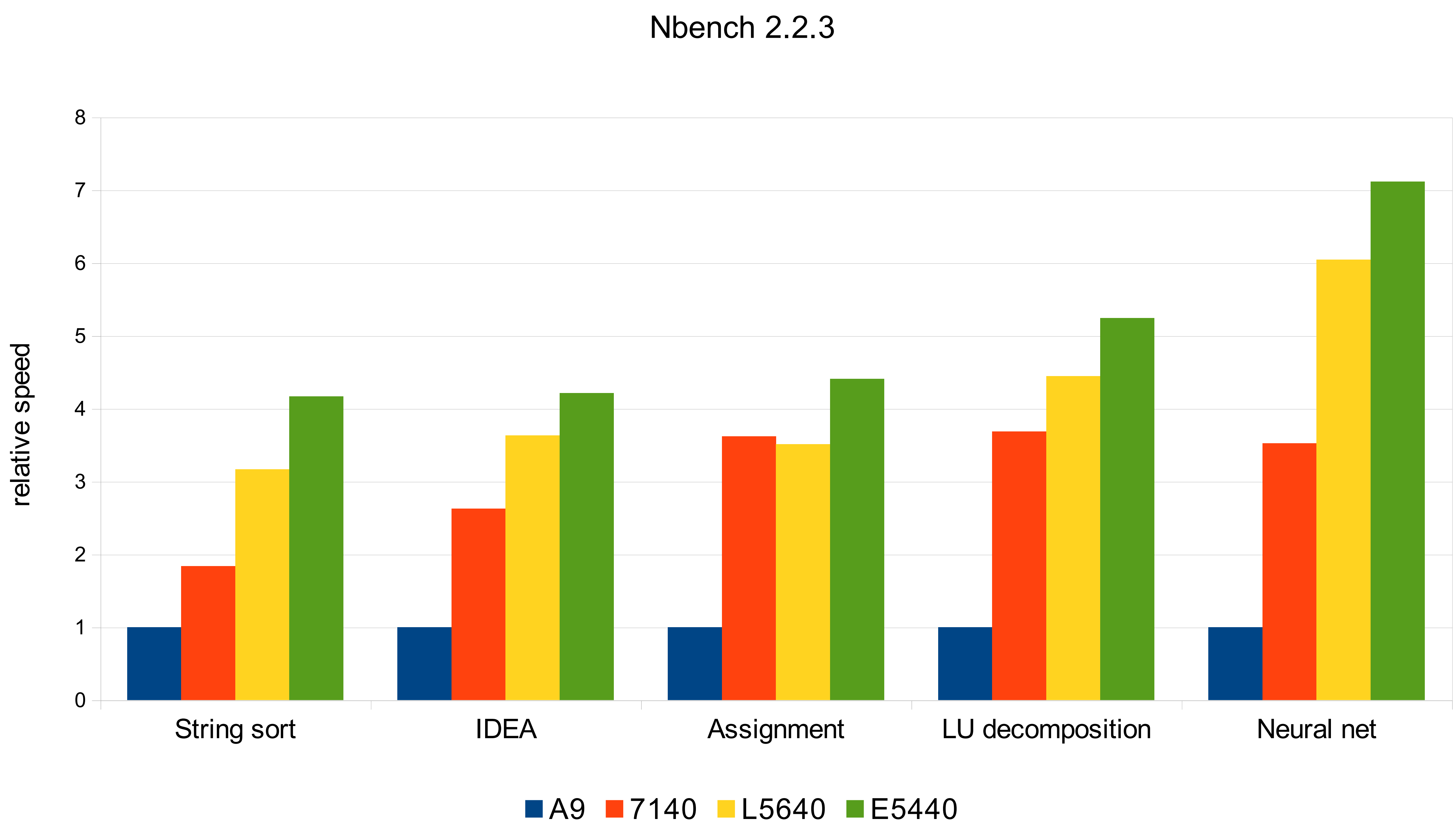
- 5 W power supply
- No cooling



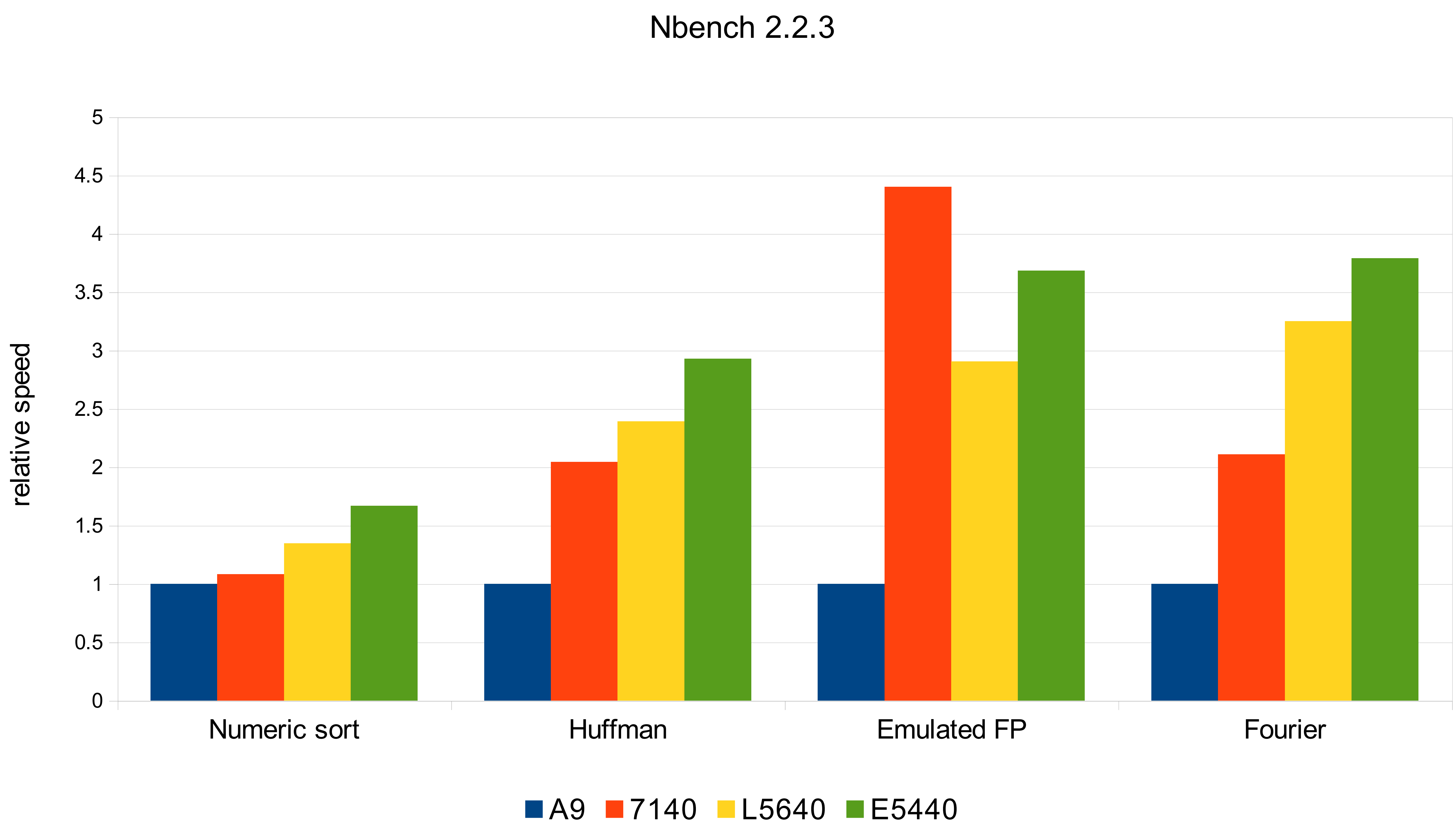
Pandaboard.org

Processor	CPU freq (GHz)	Cores	Processor power (W)	Machine power (W)
A9 (pandaboard)	1.2	2	~ 1	< 5 (1 CPU)
Xeon 7140 (2006)	3.4	2	< 212	< 1470 (4 CPU)
Xeon E5440 (2007)	2.83	4	< 133	146 - 220 (2 CPU)
Xeon L5640 (2011)	2.26 - 2.8	6	< 90	- (2 CPU)

Sequential performance 1/2



Sequential performance 2/2

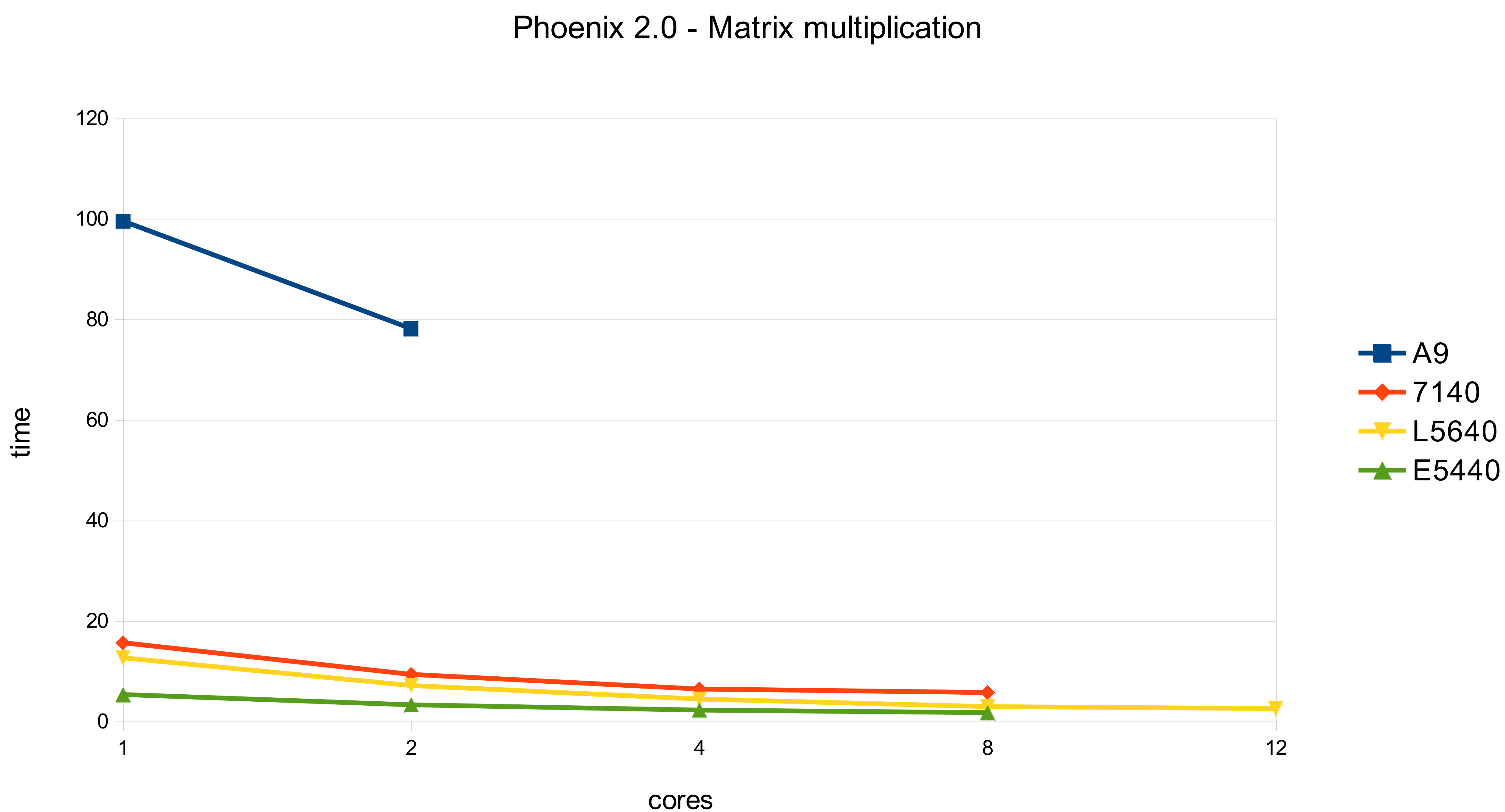


- Objective
  - test multicore performance &
  - inter-thread communication

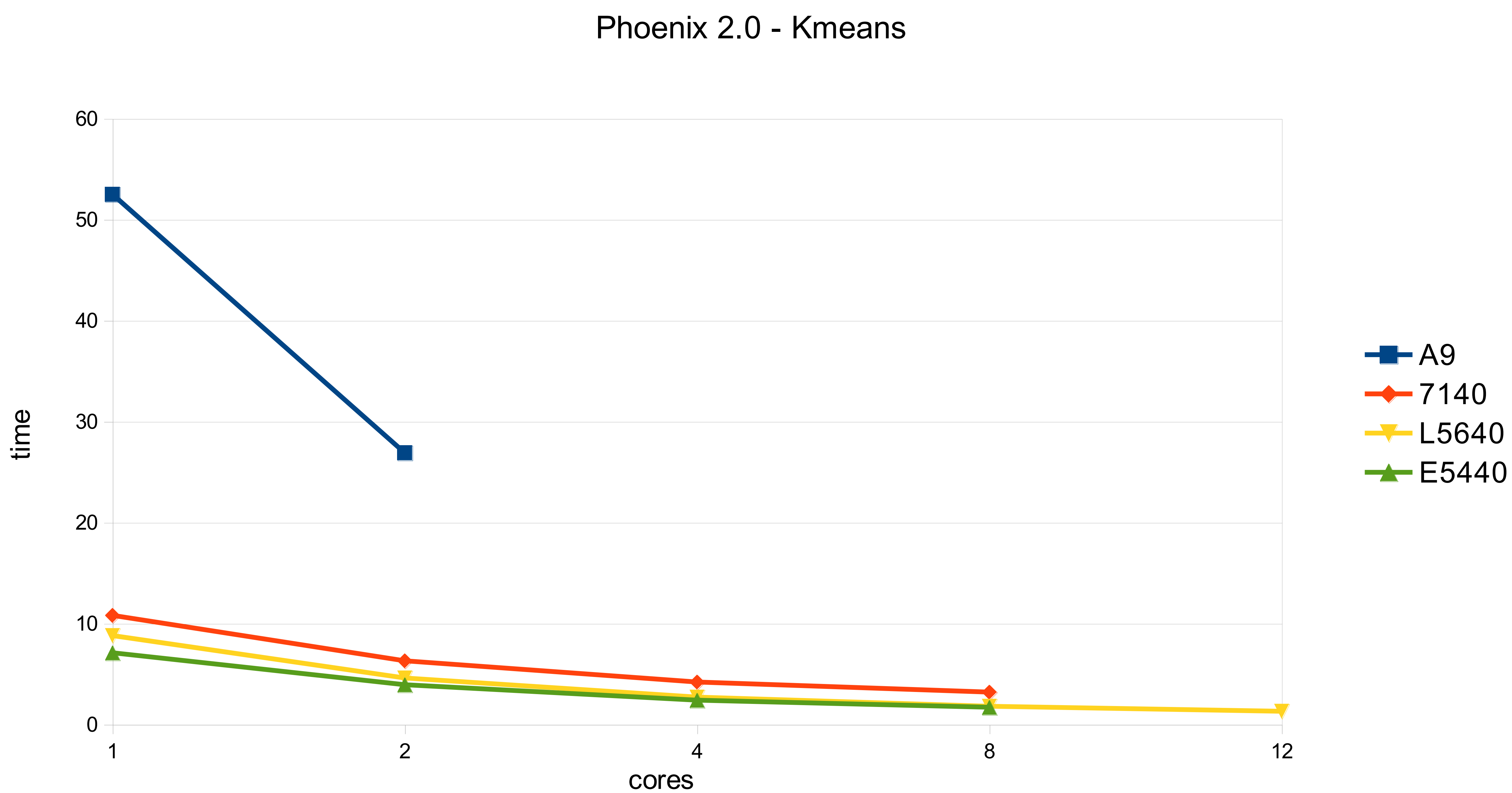
- Phoenix Benchmark : **http://mapreduce.stanford.edu/**

- Enterprise computing (Word Count, Reverse Index, String Match),
- Scientific computing (Matrix Multiply),
- Artificial intelligence (Kmeans, PCA, Linear Regression), and
- Image processing (Histogram).

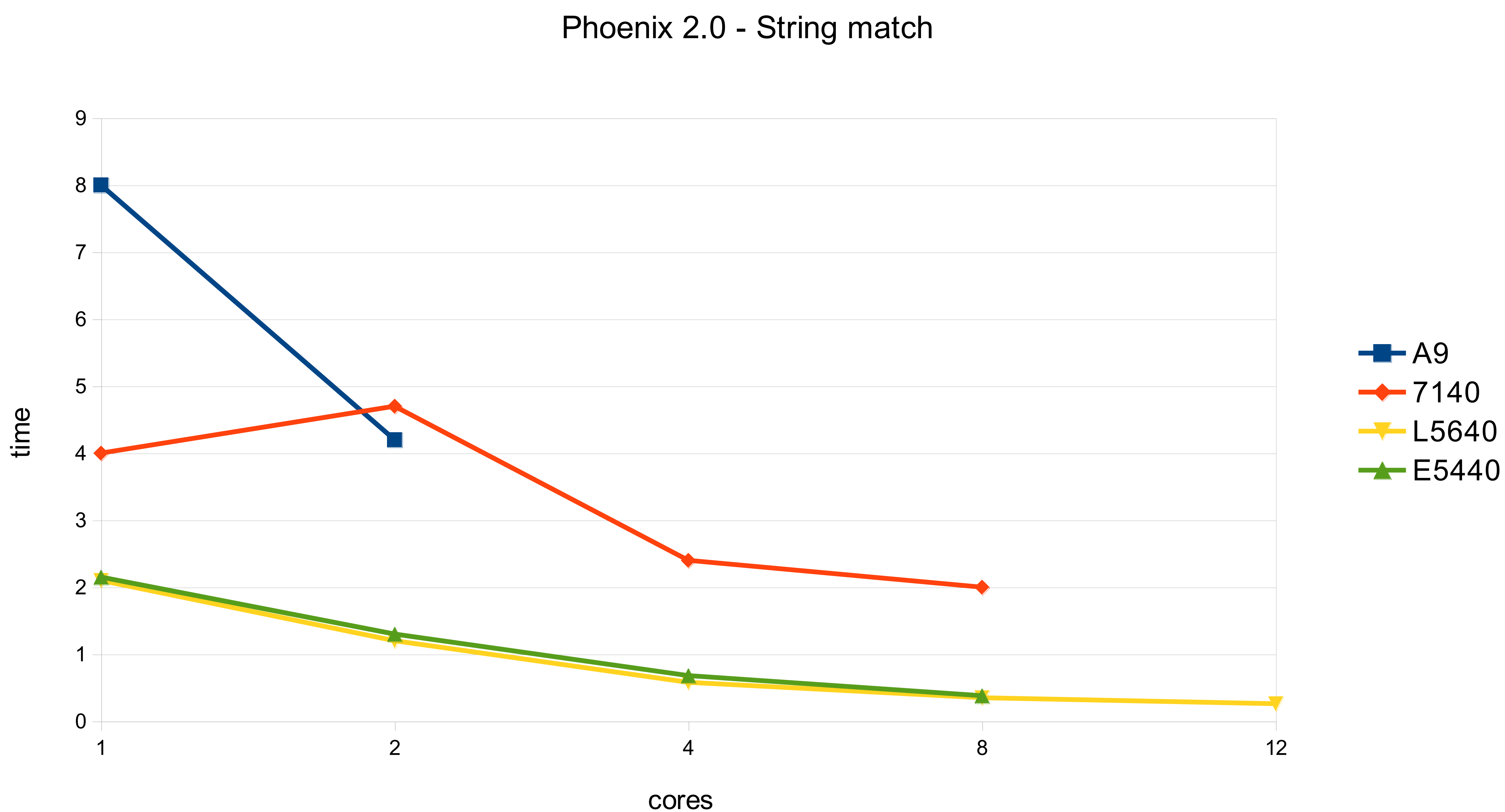
Multicore performance 1/4



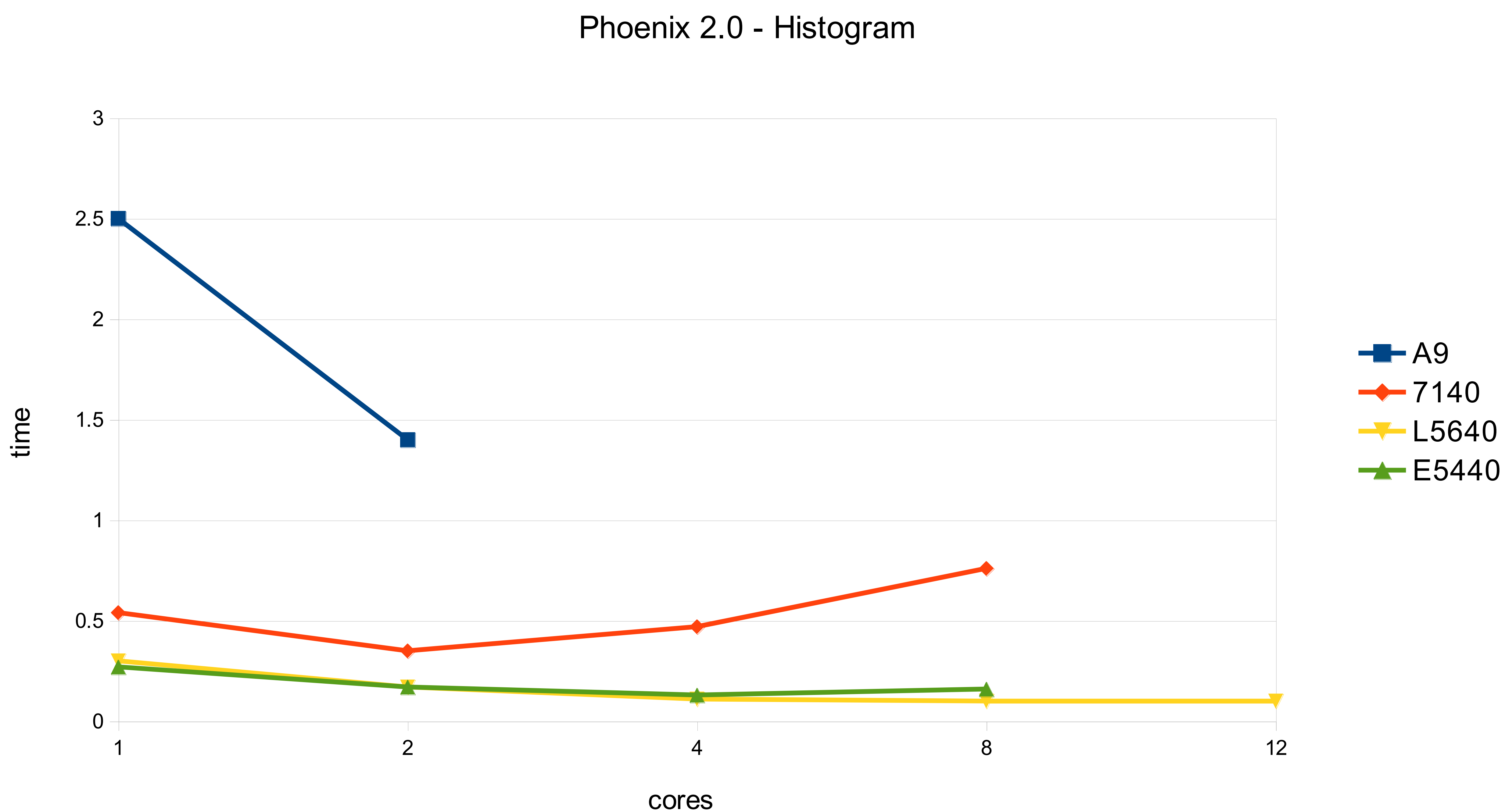
Multicore performance 2/4



Multicore performance 3/4



Multicore performance 4/4



Let’s cluster ARMs

Ex: 96 cores @ 1.2 GHz, 200 W Pandaboard-based at MITp owered by solar panel.

http://www.phoronix.com

Millicluster based on Calxeda

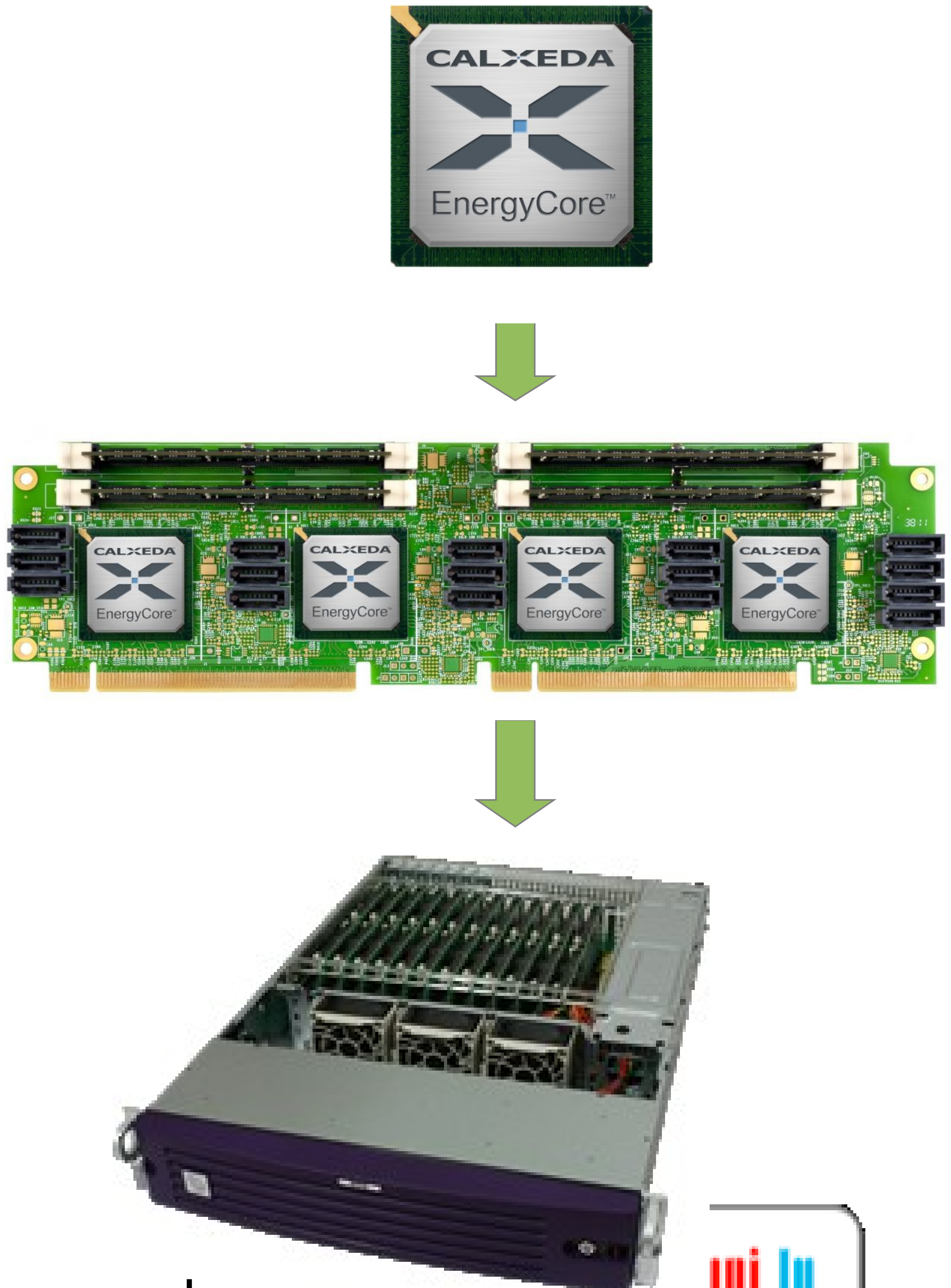
SoC:

- Quadcore A9 @ 1.4 GHz
  SIMD, FPU, 4MB L2
- 80 Gb/s crossbar switch
- Latency 200 ns

Card: 4 SoC

Rack:

- 12 cards → 192 cores
- 2U rack
- Max 300 W



http://boston.co.uk

Boston Viridis – Power "at the Wall"

BOSTON
Servers · Storage · Solutions

Workload (on 24 Nodes & SSDs)	Total System* Power	~Power per node
Linux at rest	130w	5.5w
phpbench	155w	6.5w
Coremark (4 threads/node)	169w	7.0w
Website @ 70% utilization	172w	7.2w
Linpack	191w	7.9w
STREAM	205w	8.5w

*All measurements performed on a 24 node system @ 1.1Ghz with 24 SSDs and 96GB DRAM

What are millicomputing clusters good for ?

"FAWN: A fast array of wimpy nodes", Andersen *et al.* 2008

Distributed key-value storage service (with more reads)

Tiny nodes:

- 800Mhz XScale CPU Marvell PXA 320
- Flash storage

Heterogeneous and balanced architecture
Load balancing of independent requests
→ Web applications, DNS service

Potential Application : DBMS

“Wimpy Node Clusters: What about non-wimpy workloads?” Lang *et al*. 2010

Parallel DBMS (full fledge database including ACID properties, not just key-value)

Atom vs Xeon, SSD vs HDD, single vs cluster

Raised question: scale-out improves performance?

→ Performance, price, energy are *worse*

Cause: parallel DBMS (startup, interference, skew)

Warning: simulation study only, Atom and not ARM, application specific

- Independent parallel requests:

  → multiple cores / multiple CPUs / multiple hosts
- Independent processing across dimensions:
  - Storage:

    Geographic distribution, caching (CDN)

    Independent data (Facebook)
  - Networking:

    balanced system design (match network performance)
  - Computing:

    customers are independent (stateless components)

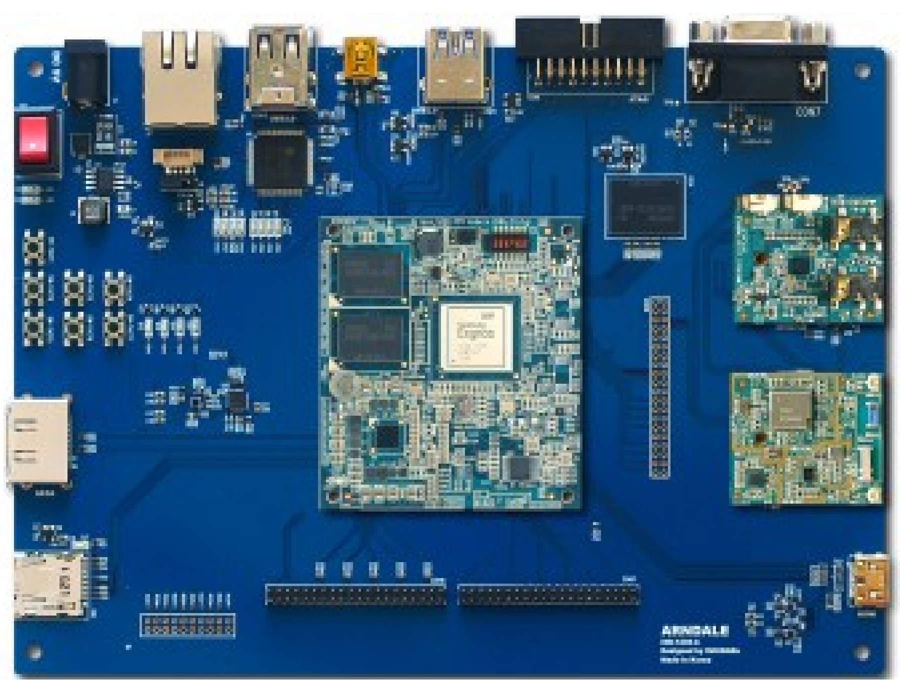
Potential Application: Cloud Computing

- Web considerations apply: independent tasks
- Specific cloud services:
  - Virtualization ?
  - Storage (key/value), archives
  - CDN (latency based routing)
  - Database service (RDBMS ? NoSQL)
  - Messaging API
  - Map-reduce, e.g. companies like Heroku, Iron.io

Potential application HPC: http://www.montblanc-project.eu/

Millicomputing for HPC Exascale, Montblanc project

Samsung Exynos 5 Dual is built on 32nm low-power HKMG (High-K Metal Gate), and features a dual-core 1.7GHz mobile CPU built on ARM® Cortex™-A15 architecture plus an integrated ARM Mali™-T604 GPU for increased performance density and energy efficiency.



Short/Medium terms:
- Millicomputing is a reality:
  main manufacturers and some key customers are in

Long term/perspectives
- Hybrid architectures
- Convergence
- Software engineering needs to adapt

GreenIT:

Dynamic multi-agent energy efficient load balancing

Green@cloud:

Cloud brokers

Millicomputing and VMs

Evocloud:

Energy efficient communications

To come:

Hybrid architectures and opportunistic computing

Advert: greencloud.gforge.uni.lu ☺

Thank you

(Questions)

Fonds National de la Recherche Luxembourg

TriICT
BEYOND ADVISORY

cnrs
dépasser les frontières

uni.lu
UNIVERSITÉ DU LUXEMBOURG