"Tuning and porting activities on Low-power multicore platforms"

Istituto Nazionale
di Fisica Nucleare

| TTLab http://ttlab.infn.it

Andrea Ferraro
Daniele Cesini
T3LAB (BOLOGNA)
18/10/2017

Istituto Nazionale di Fisica Nucleare

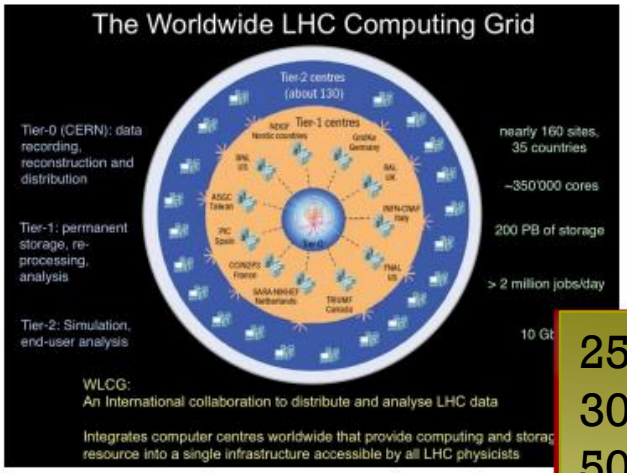
INFN

TTLab

- **5000 ricercatori** (2000 propri e 3000 universitari associati)
- **1300 giovani** tra laureandi, borsisti e dottorandi
- **270 milioni** di euro (circa) il bilancio annuale
- **2500 in un anno** le pubblicazioni scientifiche dei ricercatori INFN
- **Presenza qualificante nelle principali imprese scientifiche mondiali:** spesso gli scienziati dell'INFN dirigono o co-dirigono gli esperimenti e i programmi internazionali più importanti

L'Istituto Nazionale di Fisica Nucleare (INFN) è un Ente pubblico di ricerca, con sedi e laboratori in tutta Italia. Svolge attività in vari campi della fisica fondamentale: fisica delle particelle agli acceleratori (come il nuovo LHC al CERN di Ginevra) e nello spazio; onde gravitazionali; fisica dei nuclei; fisica teorica. La ricerca di base è affiancata da attività tecnologiche ed applicative in vari settori.

INFN
A network of data centers for BigData



25,000 cores
30 PB HDD
50 PB TAPE

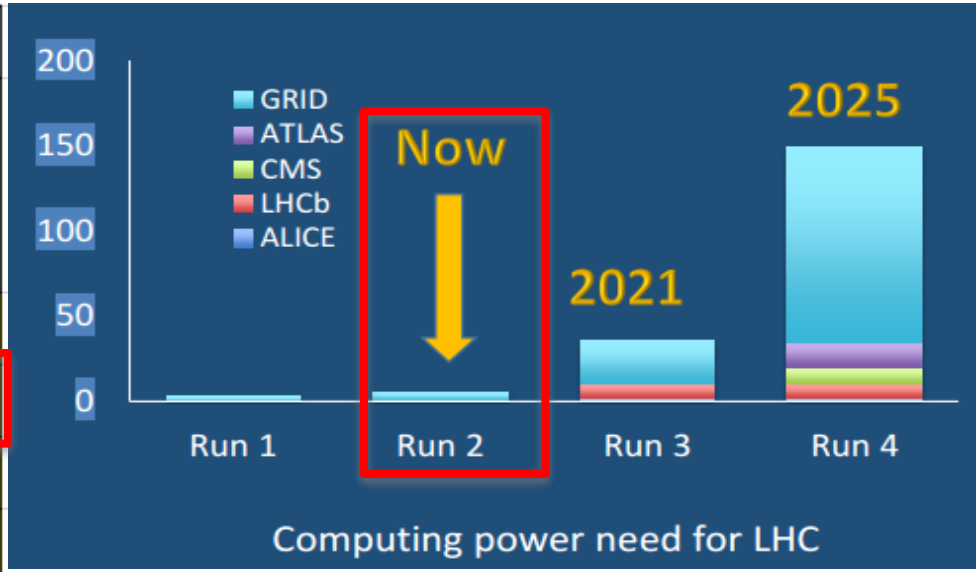
60 Gbit/s link to Geant

INFN TTLab

- INFN TTLab è un laboratorio di ricerca industriale che si prefigge l'obiettivo di tradurre i risultati di ricerca ed il know-how dell'INFN in applicazioni di possibile interesse per l'innovazione del tessuto industriale regionale,
- Il laboratorio TTLab si caratterizza sulle seguenti linee di ricerca:
  - ICT
  - Meccatronica ed Elettronica
  - Sistemi, Dispositivi e Nanotecnologie

The real challenge for every datacenter: the electrical bill !!!

INFN DATACENTER (Bologna)			
YEAR	Power (MW)	Electrical costs (€/kWh)	Electrical bill (€)
2010	993	0,14	1,2 M
2015	1050	0,20	1,8 M
2020	?	?	?
2025	???	???	???

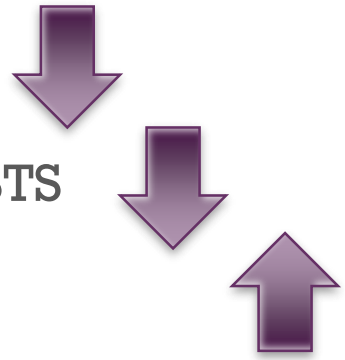


INFN is investigating on low-power multicore solutions

- Involved projects:
  - INFN COSA project (www.cosa-project.it)
  - OPEN-NEXT project (www.crit-research.it/it/projects/open-next)
- Acquiring know-how
  - Technology tracking on SoC (System on Chip)
  - Software porting and benchmarking on SoC
  - Operations of real Linux system on SoCs
  - Benchmarking hybrid architectures (CPU/GPU/DSP/etc,)
- Technology Transfer
  - Collaboration with companies and suppliers

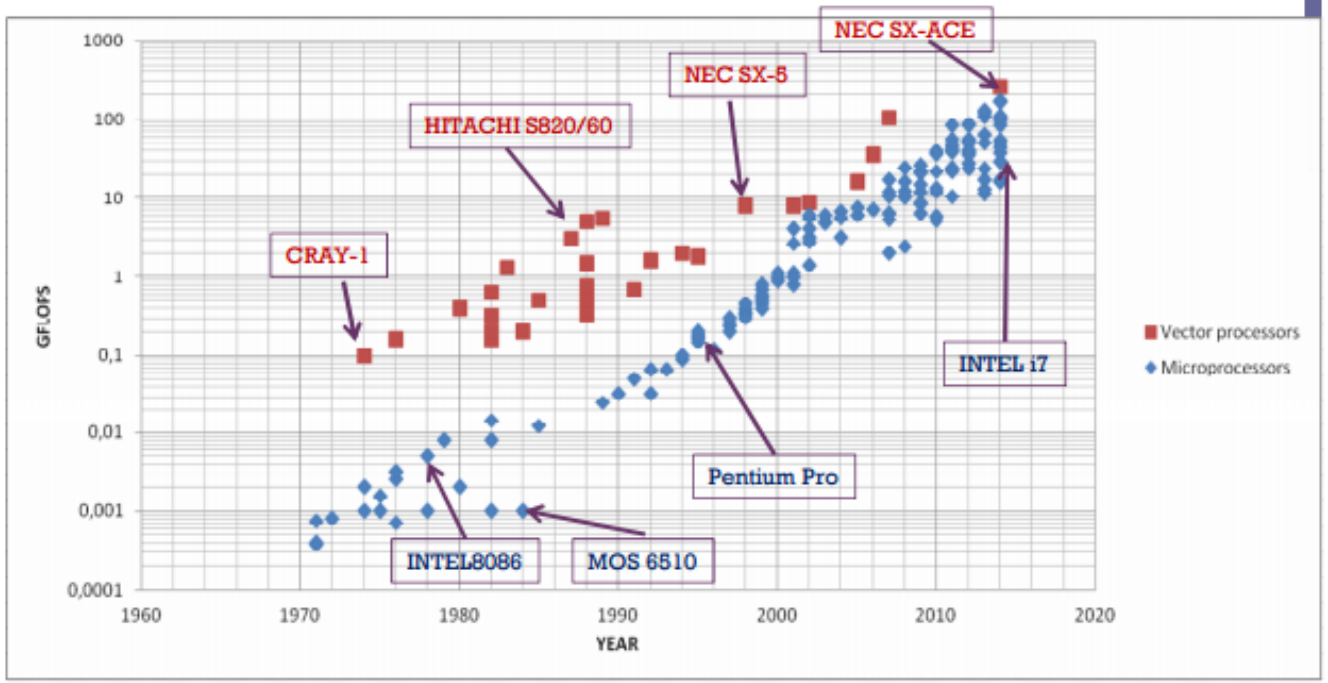
3 GOALS: OPTIMIZATION, OPTIMIZATION, OPTIMIZATION

- BOM COSTS
- ELECTRICAL COSTS
- PERFORMANCE



- Analyze the source algorithm
- Choose the right HW
- Choose the right SW program model
- Benchmark

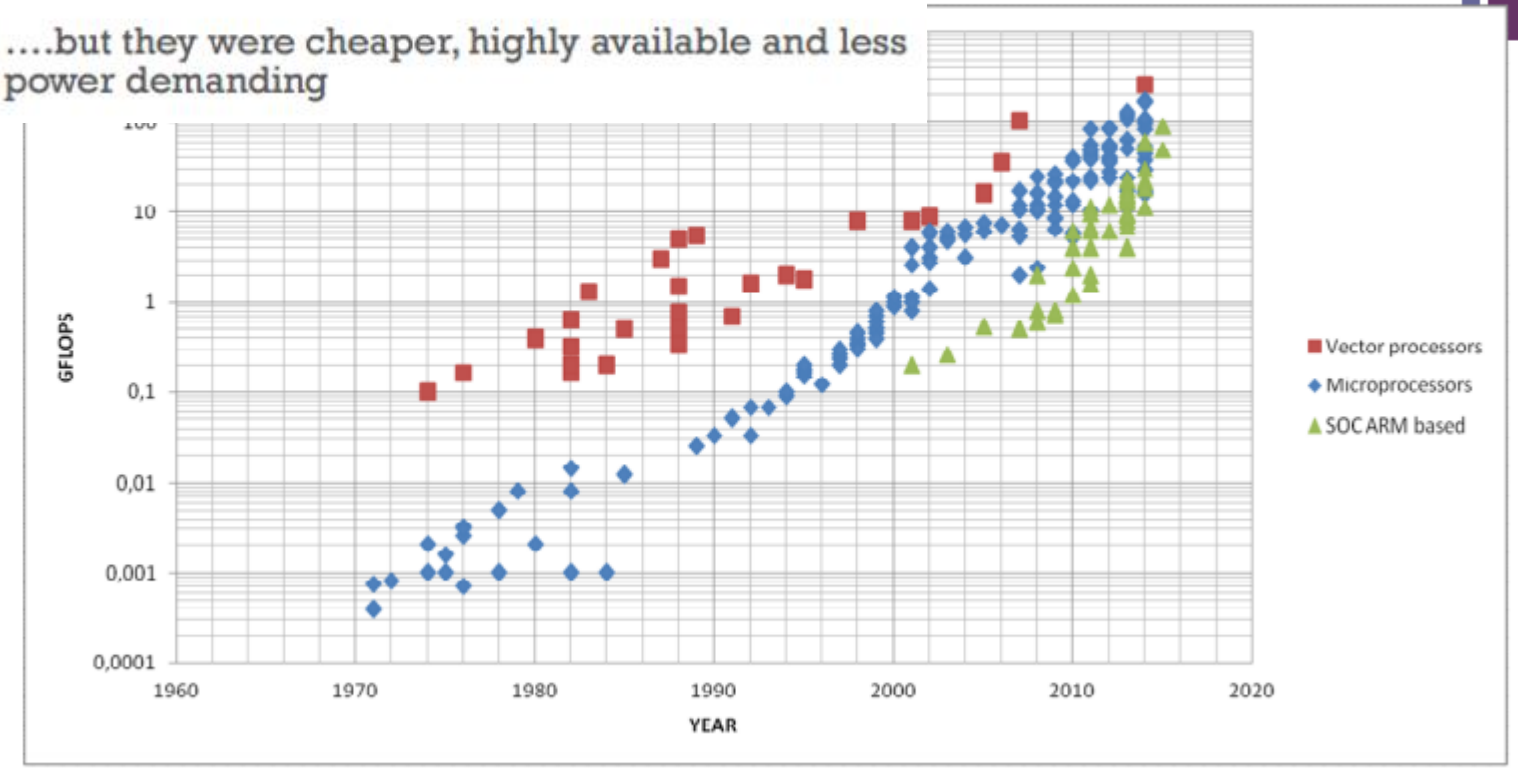
Vector vs Micro computing power



- Why did microprocessors take over?
  - They have never been more powerful…
  - ….but they were cheaper, highly available and less power demanding

Vector vs Micro vs ARM based

- Why did microprocessors take over?
  - They have never been more powerful…
  - ….but they were cheaper, highly available and less power demanding



Low-power SoC….

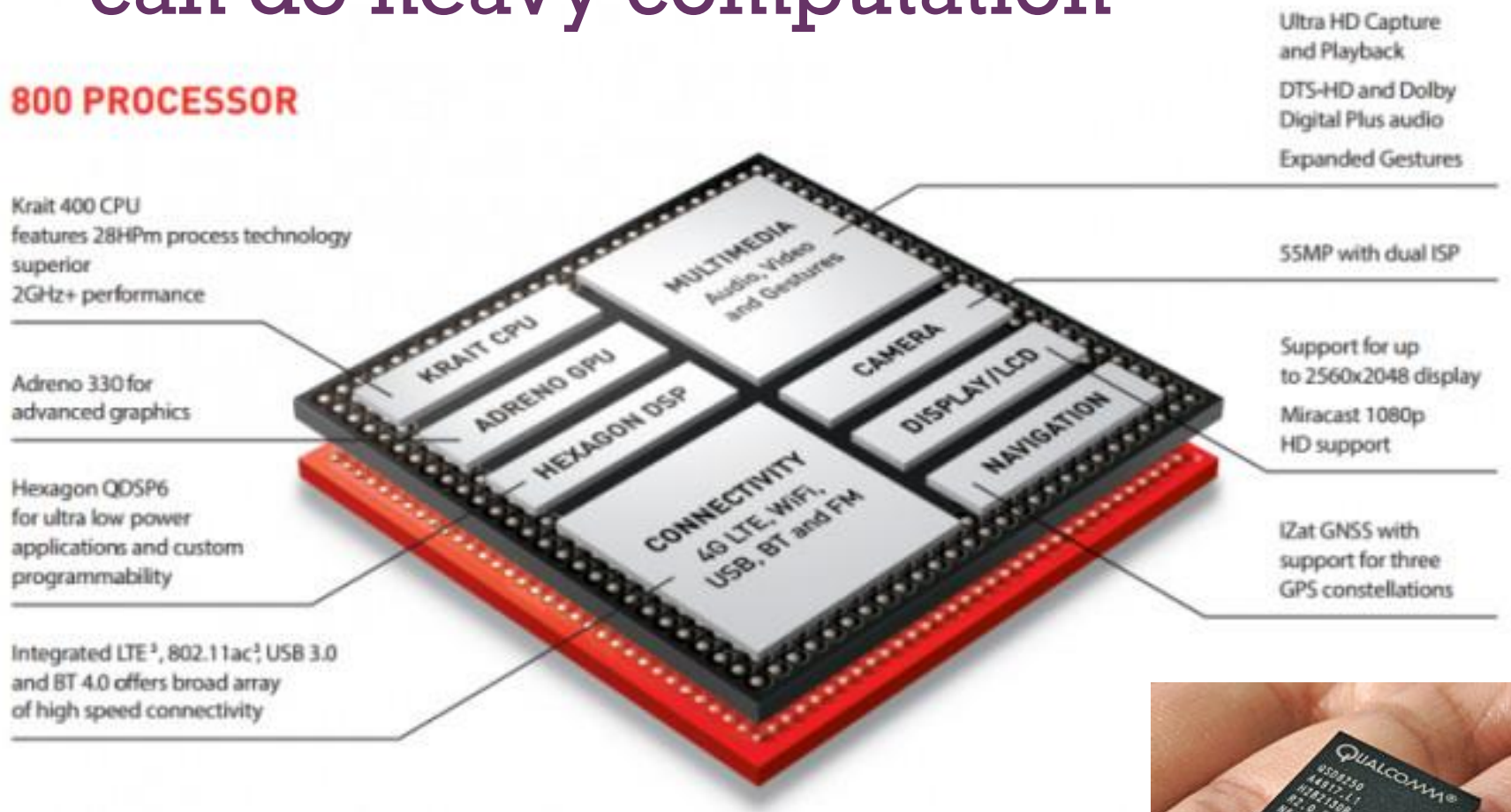
- Is history repeating?

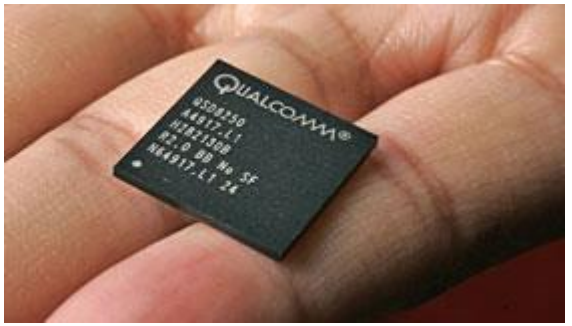
Ok, but then….an iPhone cluster?

- NO, we are not thinking to build an iPhone cluster
- We want to use SoC processors in a standard computing center configuration
  - Rack mounted
  - Linux powered
  - Running scientific application mostly in a batch environment
- ….. Use development board…

INFN Istituto Nazionale di Fisica Nucleare

Low-Power System on Chip (SoCs) can do heavy computation





INFN Istituto Nazionale di Fisica Nucleare

Before 2016: only 32bit ARM boards..

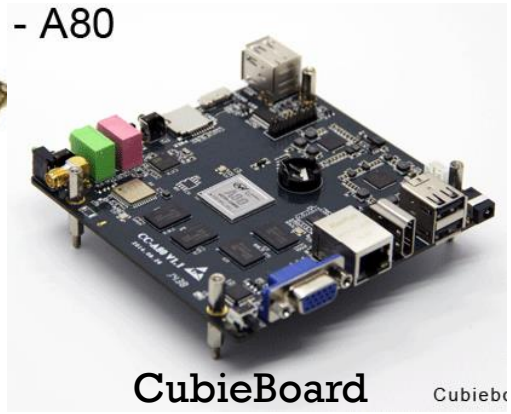
WandBoard

Rock2Board

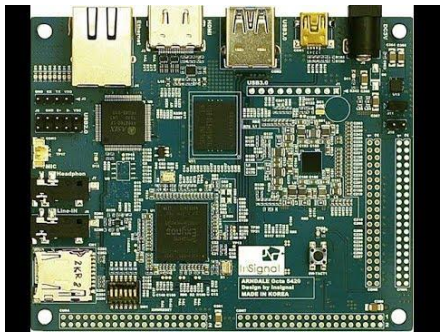
PandaBoard

DragonBoard

SabreBoard



CubieBoard



Arndale OCTA Board

Texas Instruments EVMK2H

http://elinux.org/Development_Platforms

- ...and counting...

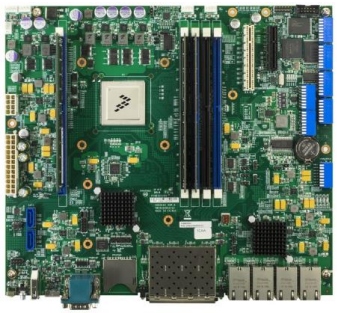
INFN Istituto Nazionale di Fisica Nucleare

Since 2016: nice 64bit ARM boards…

Server grade

ARM Juno Board
r1: 2xA57 + 4xA53
r2: 2xA72 + 4xA53
DRAM: 8 Gbytes
4 PCI-E (Gen.2, 4x)
r1: 5000$
r2:7000$

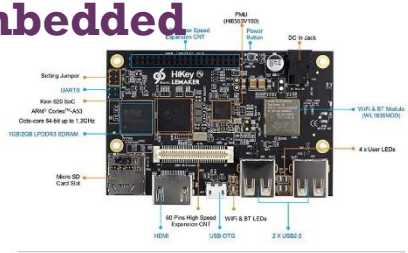
Gigabyte MP30-AR0
AppliedMicro X-Gene1 8core
DRAM:max128GB
2 x 10GbE SFP+
2 x 1GbE LAN ports
2 x PCI-Express slots (Gen.3, 8x)
700eu



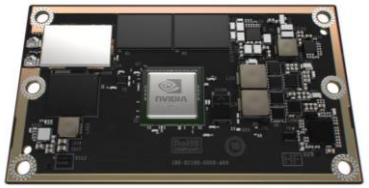
FreescaleQorIQ LS2085A
8 x Cortex-A57 cores
DRAM:max 16GB
PCI Gen3 (x8)
4 x 10 GbE SFP
4 x 10 GbE RJ45
About 3000$

AMD Opteron A1100
16GB RAM
2x10Gbs
Cost 2000$

Embedded



HiKey 96boards
1/2GB LPDDR3 SDRAM
8 x Cortex-A53 cores
Cost: $100 (2GB)

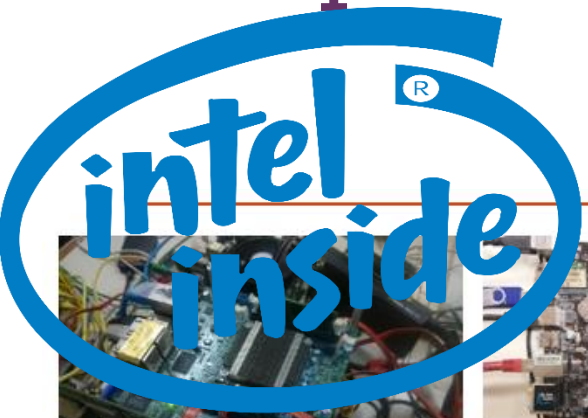


NVIDIA Jetson TX1
4x A57 2 MB di L2; 4x A53 512 KB di L2
256 core di GPU NVIDIA Maxwell
600$

ODROID-C2 64-Bit ARM
4xA53@2GHz
Mali™-450 GPU
2GB RAM
1Gbs ETH

INFN Istituto Nazionale di Fisica Nucleare

ARM is not the only player in the low-power multicore industry



C2750 ATOM

- 8 cores
- **TDP: 20 W**
- Silvermont
- 64 bit
- SATA ports
- PCIe 2.0
- Fallese
- 100 euro

N3700/N3710

- 4 cores, Intel HD Graphics
- **TDP: 6 W**
- Airmont
- 64 bit
- SATA ports
- PCIe 2.0
- Fanless
- 100 euro

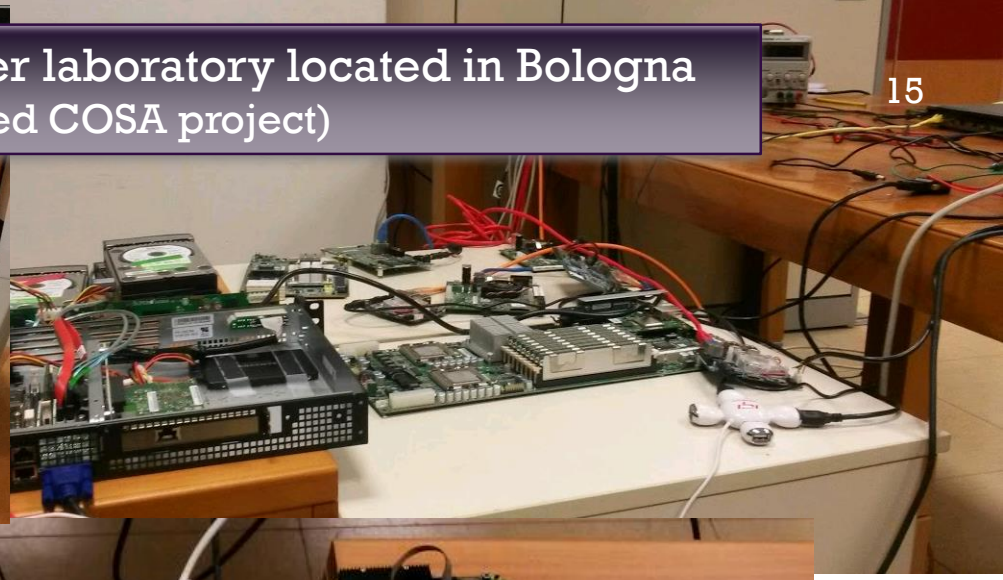
J4205

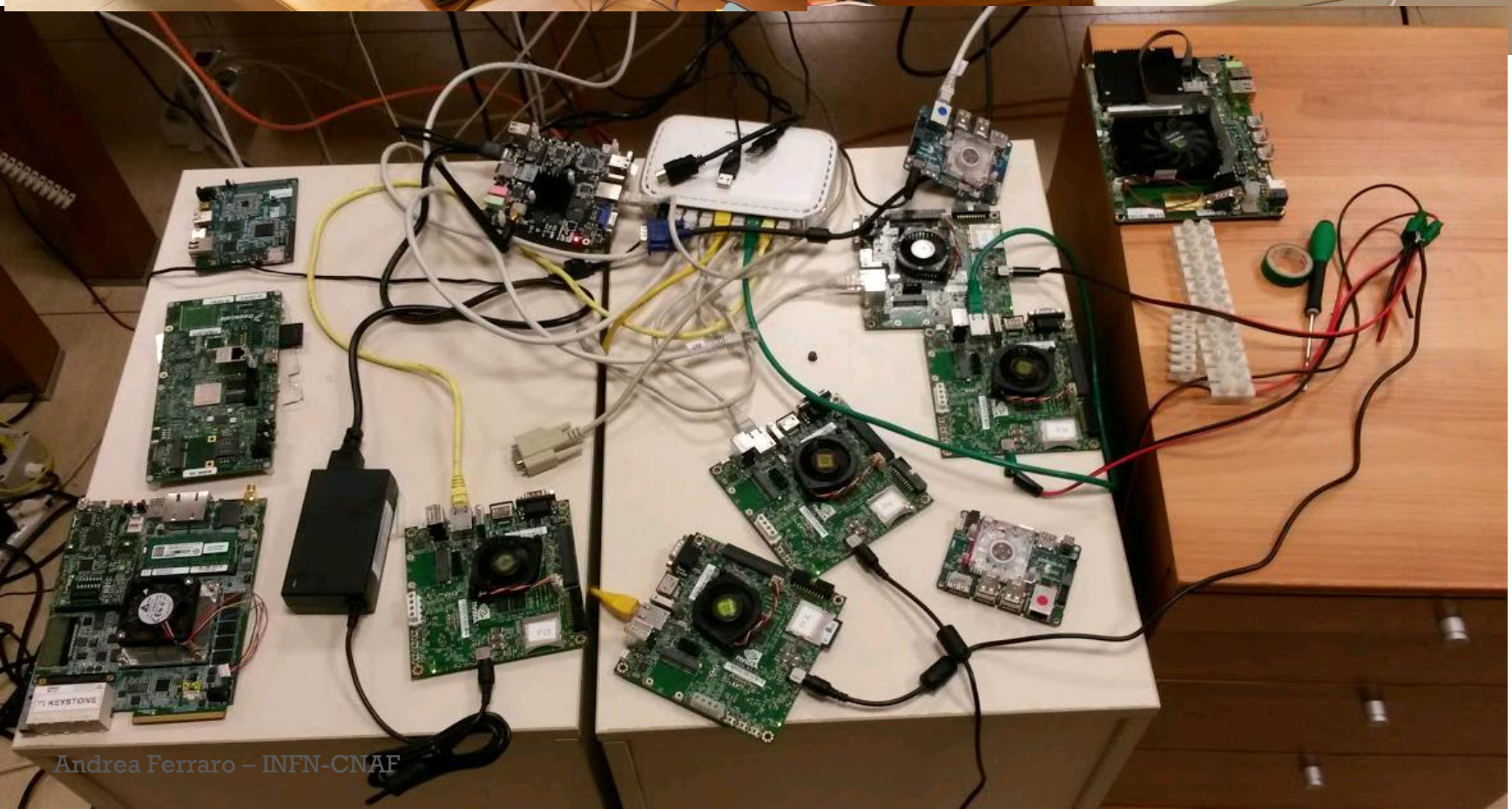
- 4 cores, Intel HD Graphics
- **TDP: 10 W**
- Pentium Apollo Lake
- 64 bit
- SATA ports
- PCIe 2.0
- Fanless
- 150 euro

XEON D

- 16 cores (HT)
- **TDP: 45 W**
- Broadwell
- 64 bit, AVX
- 6 SATA ports
- PCIe 2.0
- 900 euro

The INFN low-power laboratory located in Bologna
(assets by INFN-funded COSA project)





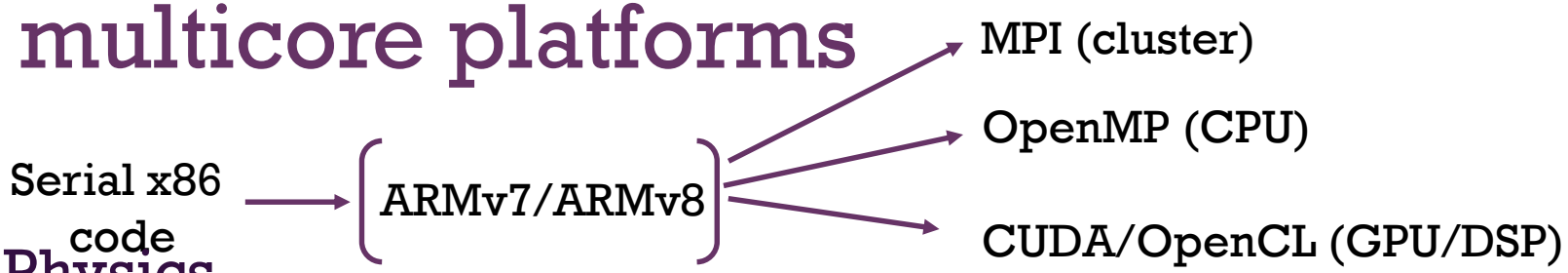
Clusters (assets by INFN-founded COSA project)

16xARMv7
2xARMv8

4xINTEL AVOTON C-2750
4xINTEL XEOND-1540

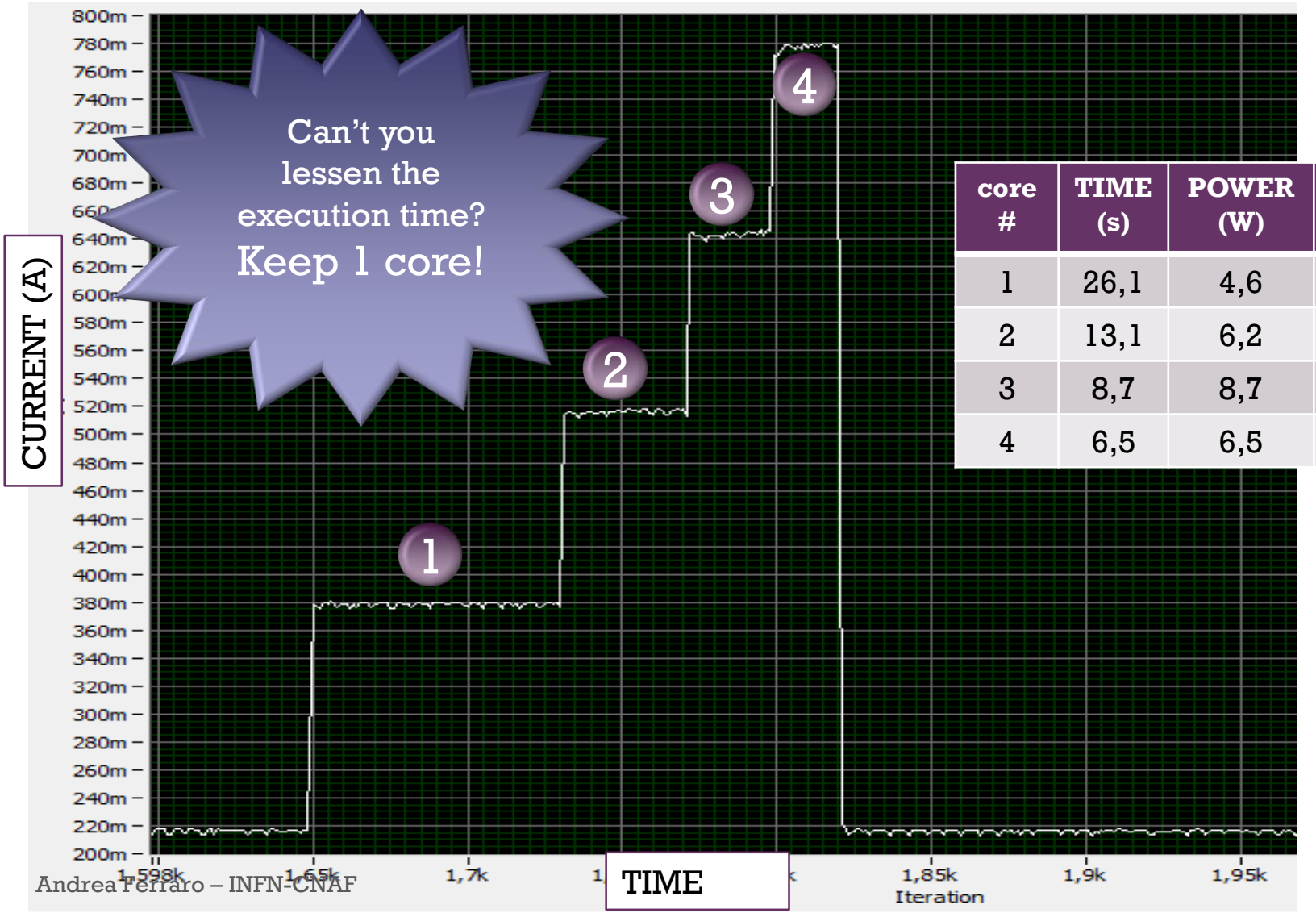
2xINTEL N3700
4xINTEL N3710
2XINTEL J4205

Applications ported to low-power multicore platforms



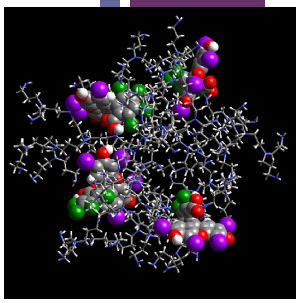
- Physics
  - Montecarlo and analysis of LHC experiments
  - HEP experiments High Level Trigger and Data Acquisition applications
  - Parallel applications usually run in HPC environments (Lattice Quantum ChromoDynamics simulations)
- Biomedical applications
  - Computer tomography
  - Bioinformatic pipelines
  - Space-aware stochastic simulator
- Deep learning and neural networks
  - Image classification and segmentation

Multicore means a lot of energy… Goal: lessen the execution time!!!

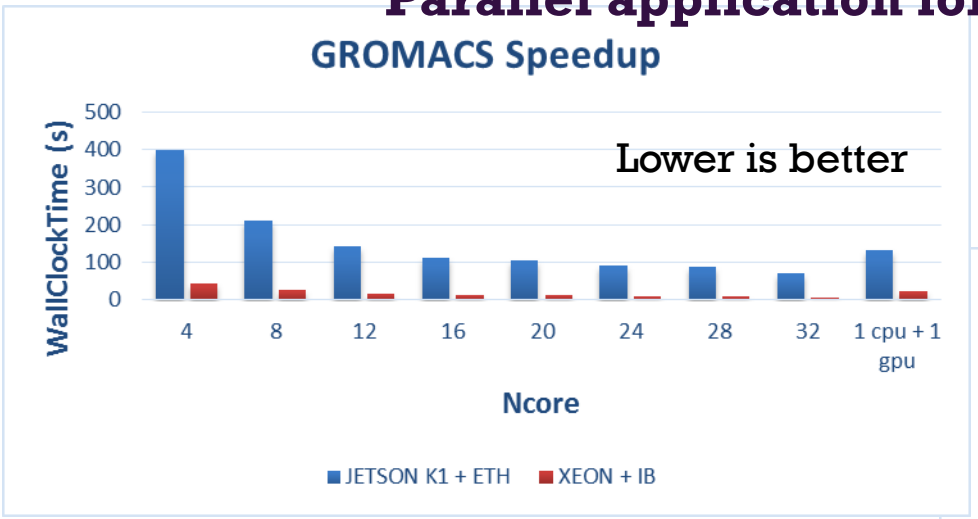


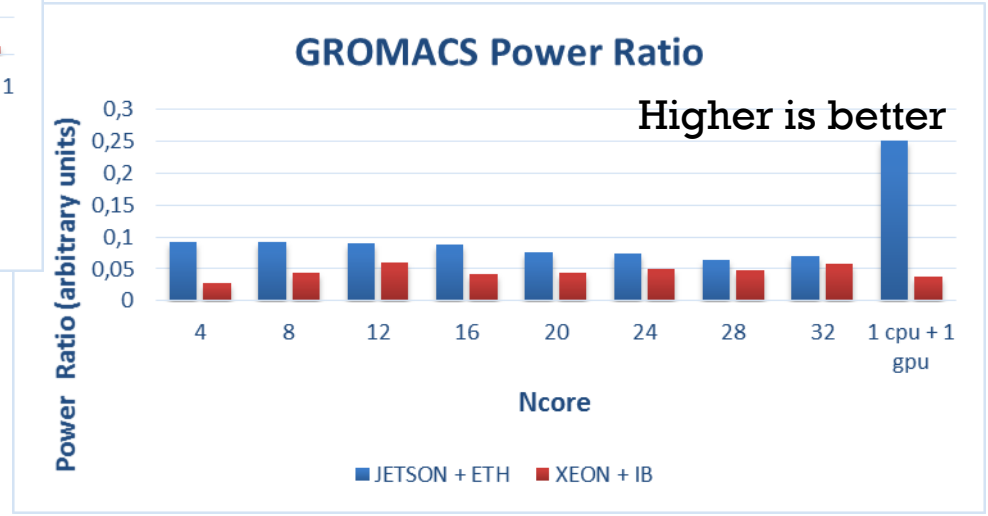
core #	TIME (s)	POWER (W)	ENERGY (J)
1	26,1	4,6	120,1
2	13,1	6,2	81,2
3	8,7	8,7	75,7
4	6,5	6,5	42,2

Molecular Dynamics on ARM Nvidia Jetson-TK1



Parallel application for CPU and GPU



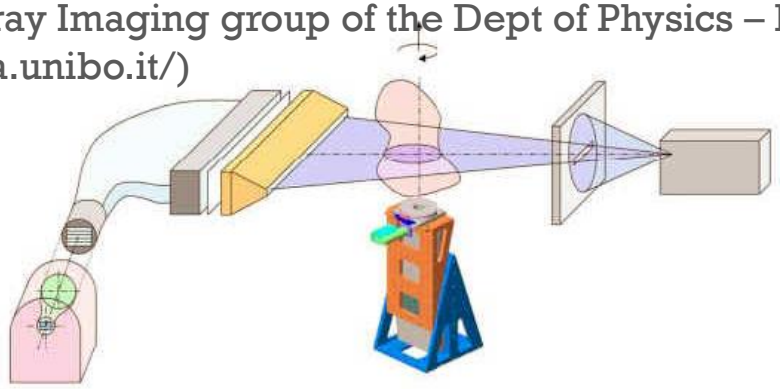


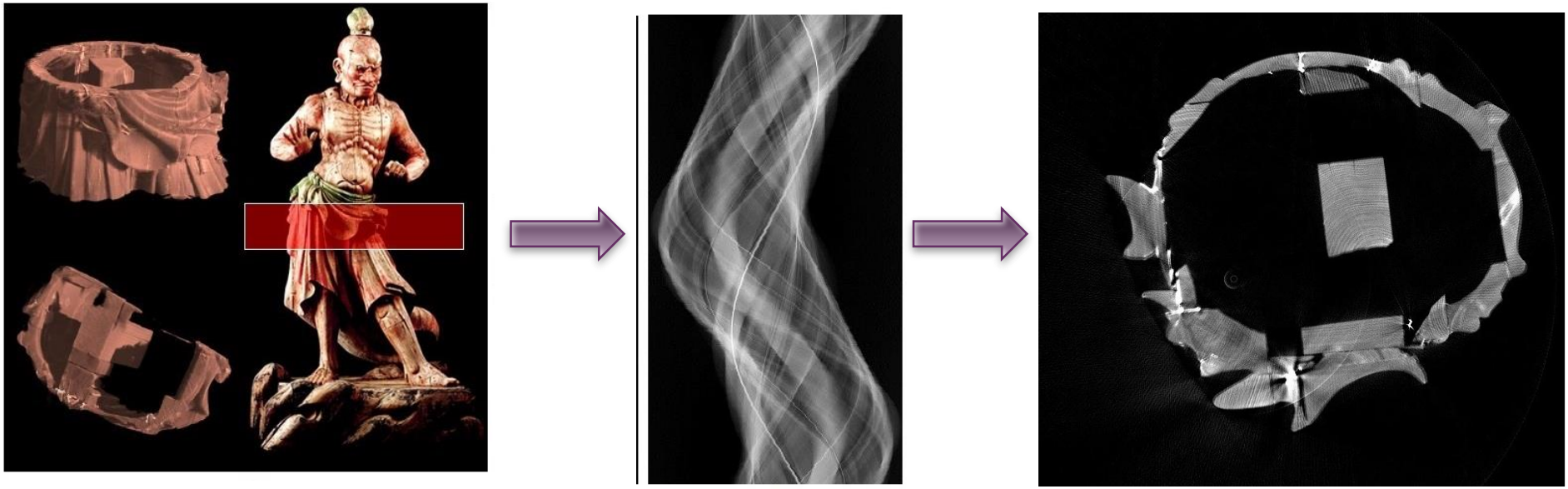
- Jetson-TK1 about 10X slower using the same number of cores
- Jetson-TK1 about 10X slower using the GPU (vs. an NVIDIA Tesla K20)
  - Jetson-TK1 13.5Watt
  - Xeon+K20 ~320Watt

Computer tomography

Filtered Backprojection Algorithm

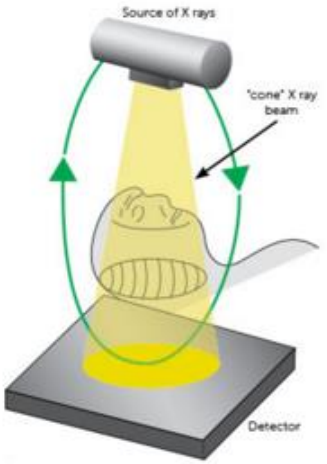
In collaboration with the X-ray Imaging group of the Dept of Physics – Bologna University (http://xraytomography.difa.unibo.it/)



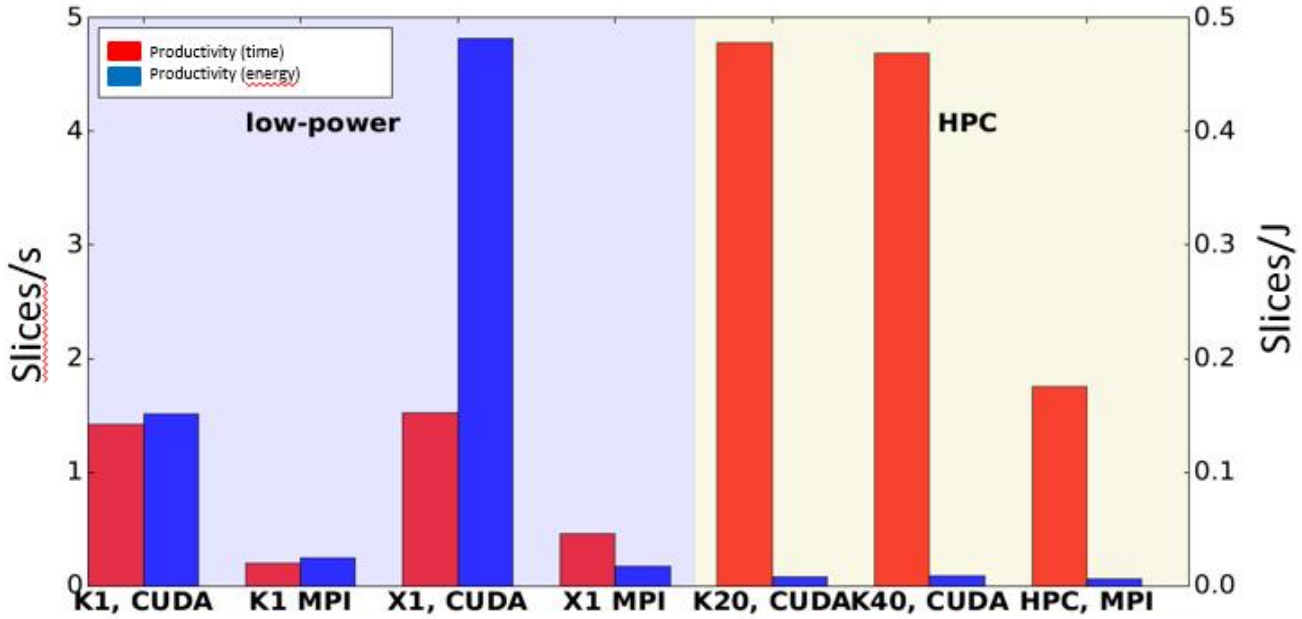


Real-Time Reconstruction for 3-D CT Applied to Large Objects of Cultural Heritage, R. Brancaccio, M. Bettuzzi, F. Casali, M. P. Morigi, G. Levi, A. Gallo, G. Marchetti, and D. Schneberk, IEEE TRANSACTIONS ON NUCLEAR SCIENCE, VOL. 58, NO. 4, AUGUST 2011

Computed Tomography



Feldkamp *cone beam* reconstruction

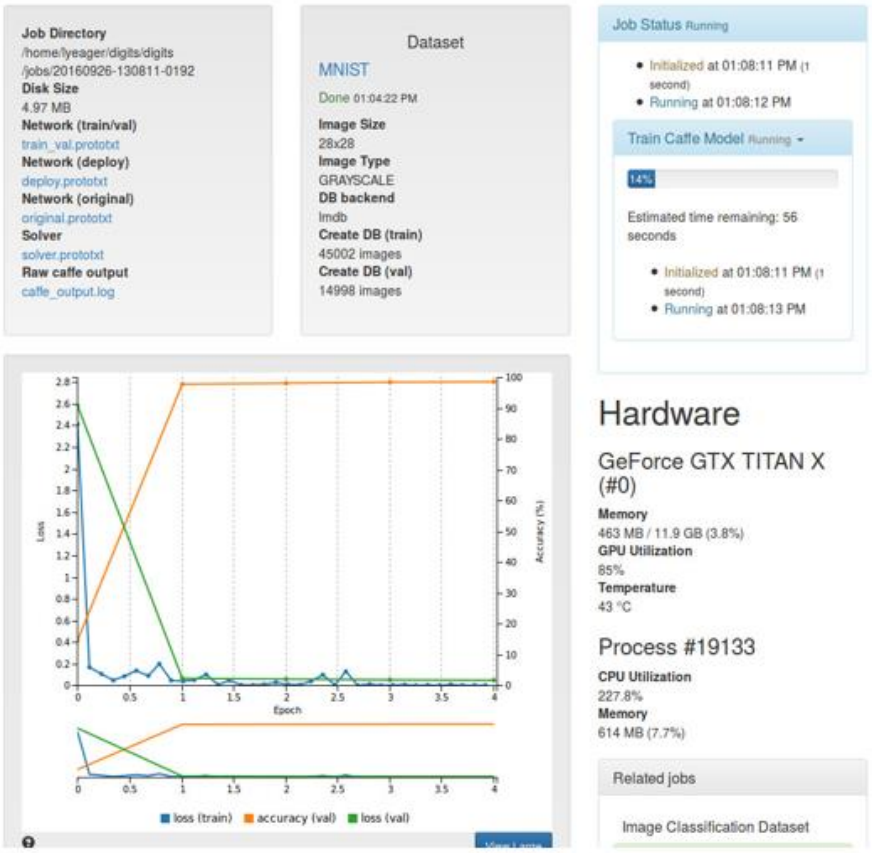


4 TX1:
21600 slices in 1 hour
60W

2xE5-2620+1K20:
21000 slices in 1 hour
350W

INFN Istituto Nazionale di Fisica Nucleare

Deep learning and neural networks: image classification



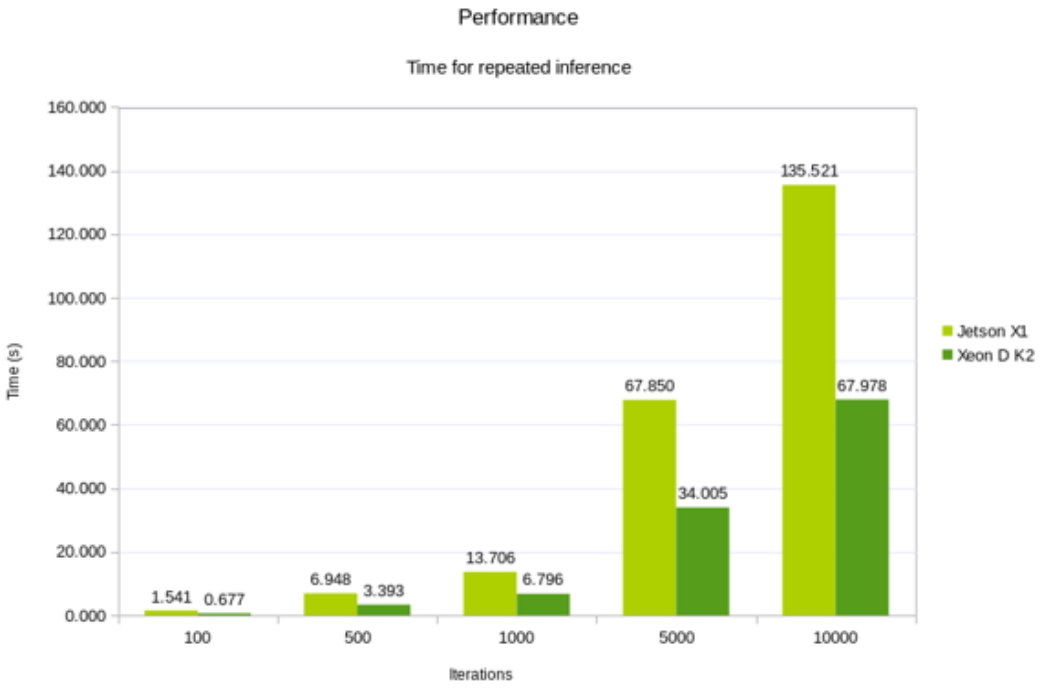
- **Dataset creation using ImageNet**
- **Training using DIGITS + Caffe on 2 K40 or XeonD+K20**

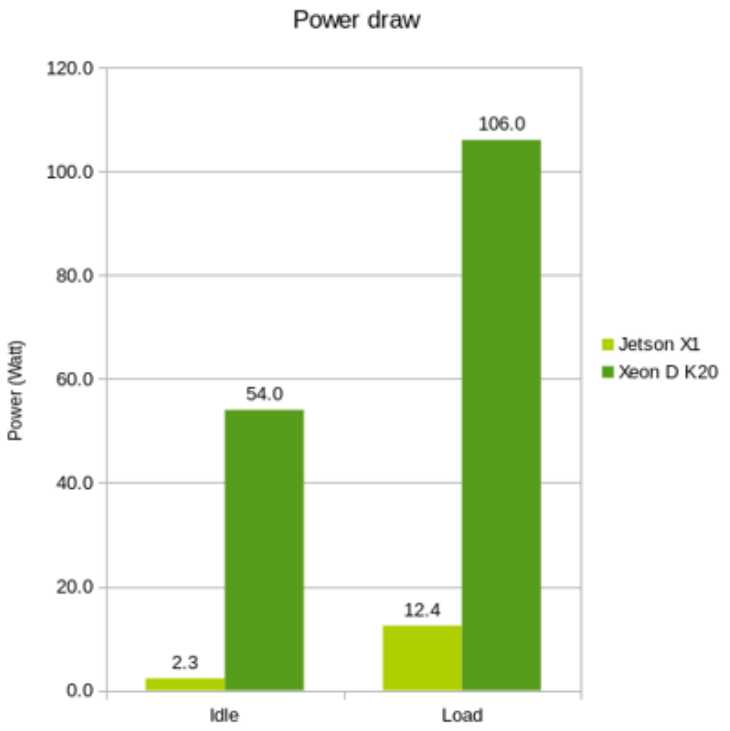
For comparison:

Training on CPU (Xeon E5-2640 v2, 16HT) -
> 4 cycles in 1 day, then stopped

Training on GPU for the same model
-> 90 cycles in 3h27m

Deep learning and neural networks: image classification

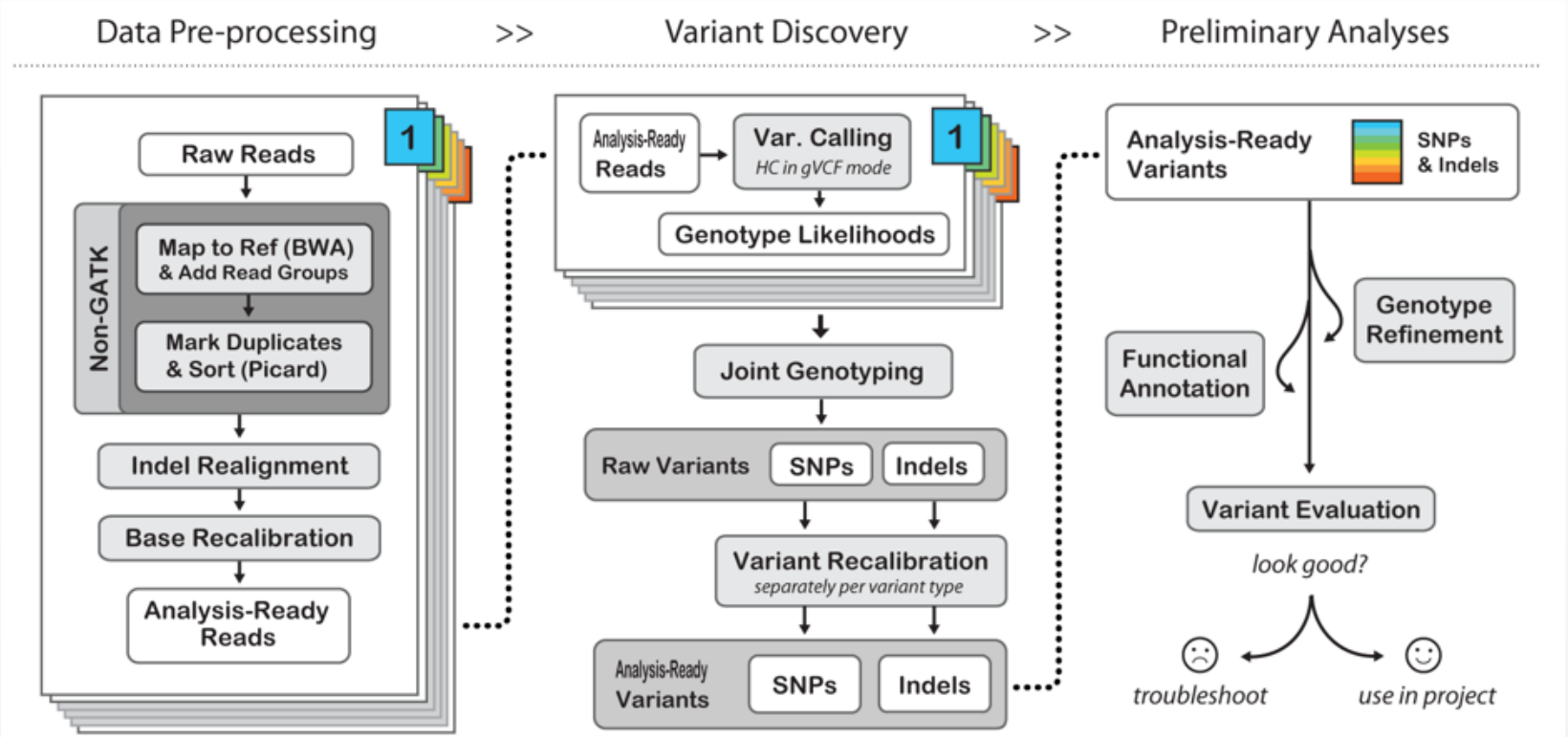




- XEOND + K20 is 2x faster, but…

… Tegra X1 is 10 times less power hungry

Analysis workflow



gatk

BROAD INSTITUTE

© Broad Institute 2012

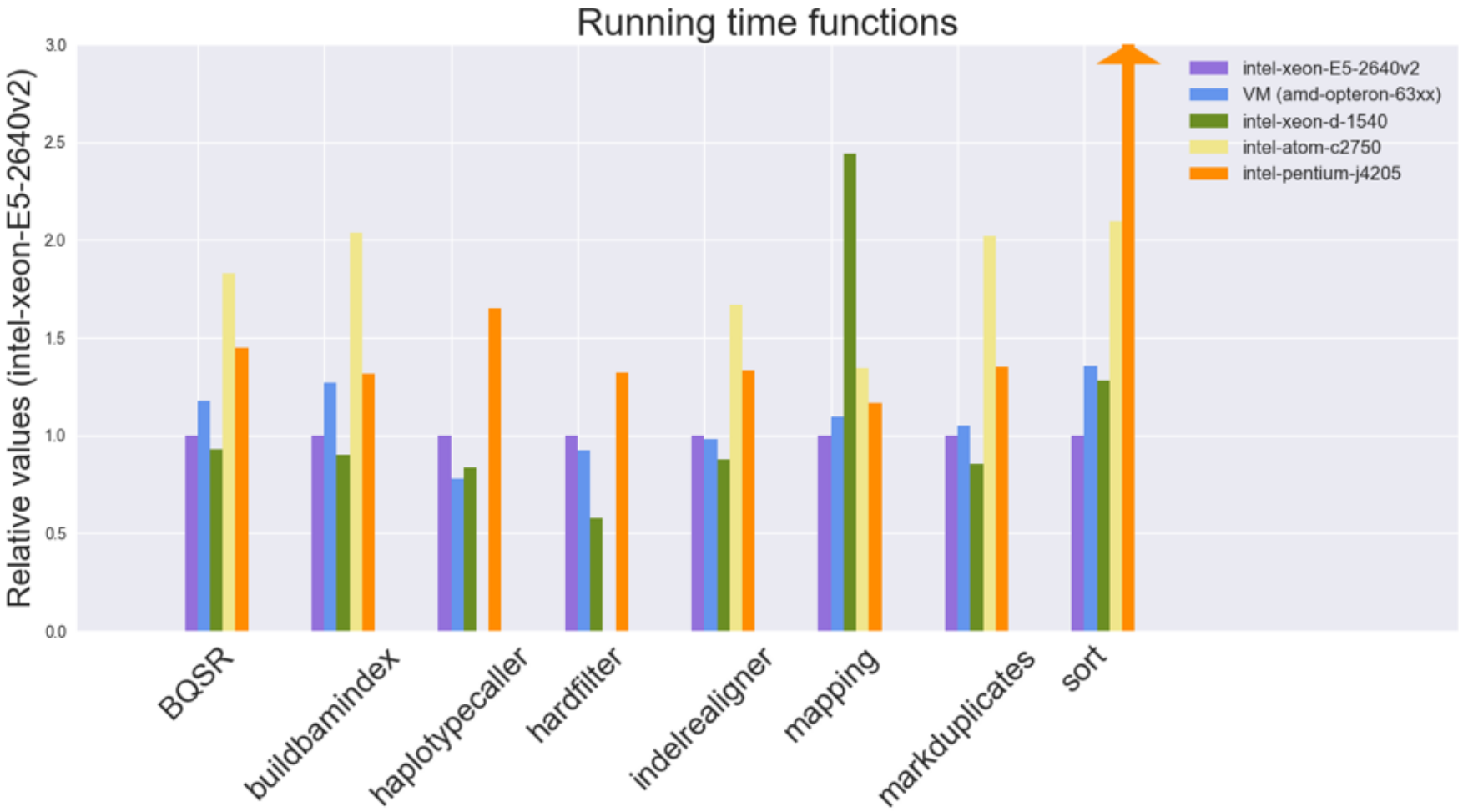
INFN Istituto Nazionale di Fisica Nucleare

Low-power multicore for bioinformatics pipelines

	Server-grade nodes		Low-power multicore nodes			Virtual machines
CPU	**Intel Xeon E5-2683v3**	**Intel Xeon E5-2640v2**	**Intel Pentium J4205**	**Intel Xeon D-1540**	**Intel Atom C2750**	**AMD Opteron 6386 SE**
Microarchitecture	Haswell	Ivy Bridge EP	Apollo Lake	Broadwell	Avoton	Piledriver
Launch Date	Q3'14	Q3'13	Q4'16	Q1'15	Q3'13	Q3'12
Lithography	22 nm	22 nm	14 nm	14 nm	22 nm	32 nm
Cores/threads	14/28	8/16	4/4	8/16	8/8	16
Base/Max Freq (GHz)	2.00/3.00	2.00/2.50	1.50/2.60	2.00/2.60	2.40/2.60	2.80/3.50
L2 Cache	35 MB	20 MB	2 MB	12 MB	4 MB	16 MB
TDP	120 W	95 W	10 W	45 W	20 W	115 W
Total CPUs	2	2	1	1	1	1
Total cores/threads	28/56	16/32	4/4	8/16	8/8	16
Total Memory	256 GB	128 GB	8 GB	32 GB	16 GB	63 GB
System power	240 W + 60 W	190 W + 60 W	10 W + 2 W	45 W + 10 W	20 W +10 W	115 W + 10 W
Electrical costs (0,25 €/kWh)	650 €/year	550 €/year	26 €/year	120 €/year	65 €/year	273€ /year
System price	4000-6000 €	3000-5000 €	100-130 €	900-1200 €	500-700 €	2000-3000€

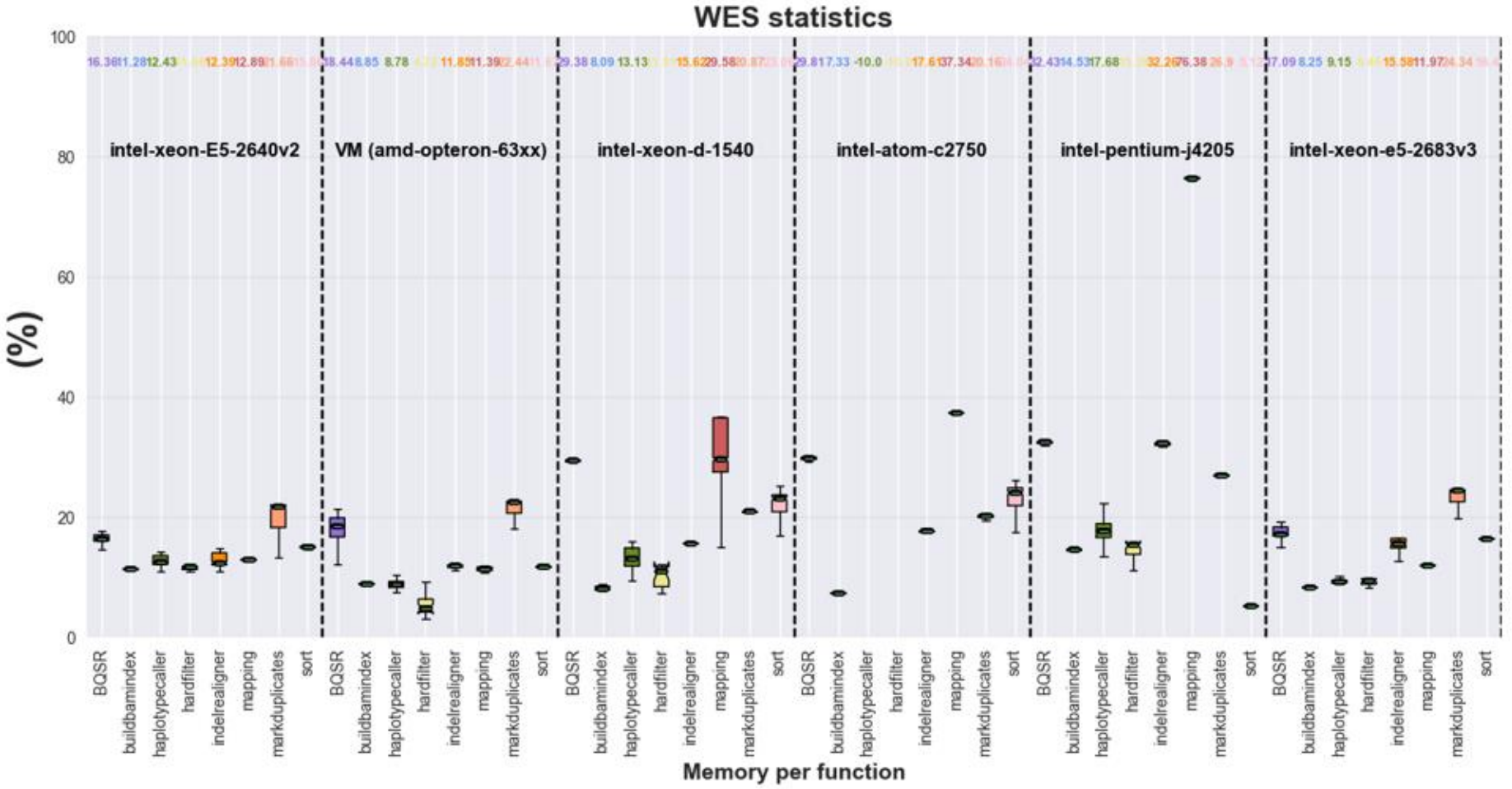
Low-power multicore for bioinformatics pipelines

Performance tests



Low-power multicore for bioinformatics pipelines

Memory tests

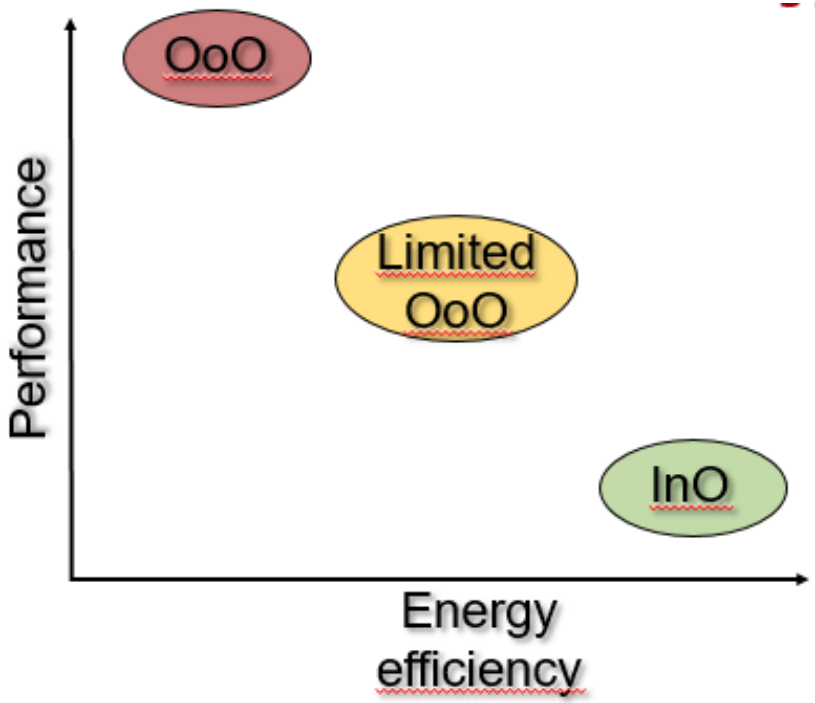


Conclusions

- Now bioinformatics scientists buy big servers (128GB/256GB)
- 95% of tasks require less than 32GB
- Optimize software pipes of genomics data for low-power multicore nodes is the right approach
  - E.g. BWA can use a cluster of low-power multicore nodes with less than 8GB

Collaboration with Montblanc Project and Department of Information Technology, Uppsala University

Compiler techniques to deliver high performance at low energy costs!



- Software Decoupled Access Execute (DAE) for energy efficiency
  - Goal: reduce energy
  - Large power-hungry OoO cores
- Clairvoyance
  - Goal: increase performance, memory- & instruction-level-parallelism
  - Limited OoO cores
- SWOOP
  - Goal: Make IO cores run like OoO cores
  - In-Order (IO) cores

SCADA and low-power IT

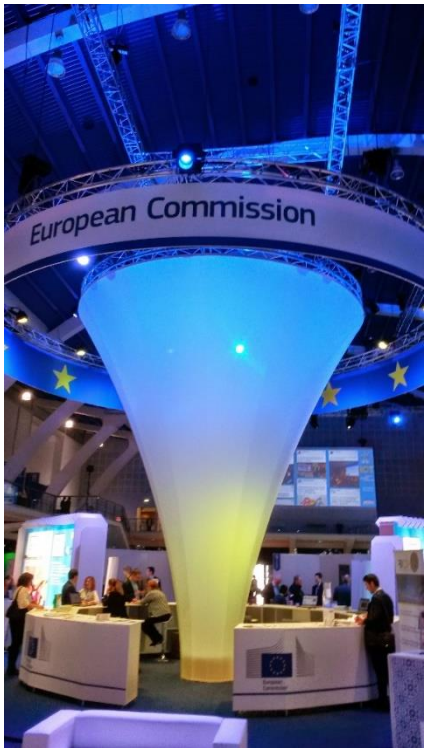
- Porting OPC-UA stacks to a BigData low-power cluster
  - OPC-UA messages fired by PLCs
  - OPC-UA server in a low-power server (Intel n3700)
  - Collecting, gathering, data analytics frameworks (Hadoop/Spark/InfluxDB) in a low-power cluster
- BENEFITS
  - Joining IT BigData experience + SCADA industrial experience
  - A low-cost BigData cluster (up to 10 Hadoop/Spark nodes cluster 40cores/160TB) for SCADA tests

Conclusion

- Embedded multicore SoCs are becoming attractive for real life scientific and industrial applications
  - Easy to program if developers use the appropriate programming paradigms (OpenMP, OpenACC, OpenCL, CUDA, etc.)
  - Great results if you manage to extract power from the integrated GPU
  - ARM dominated until last year, now INTEL is becoming competitive in this segment
- INFN has a proven competence in optimization of hybrid low-power embedded architectures and experience porting applications to multi-core/hybrid platforms

Horizon2020: We participated (not funded) to a consortium for Low Power and Customized Computing

  - HW+SW software prototype



INFN Istituto Nazionale di Fisica Nucleare