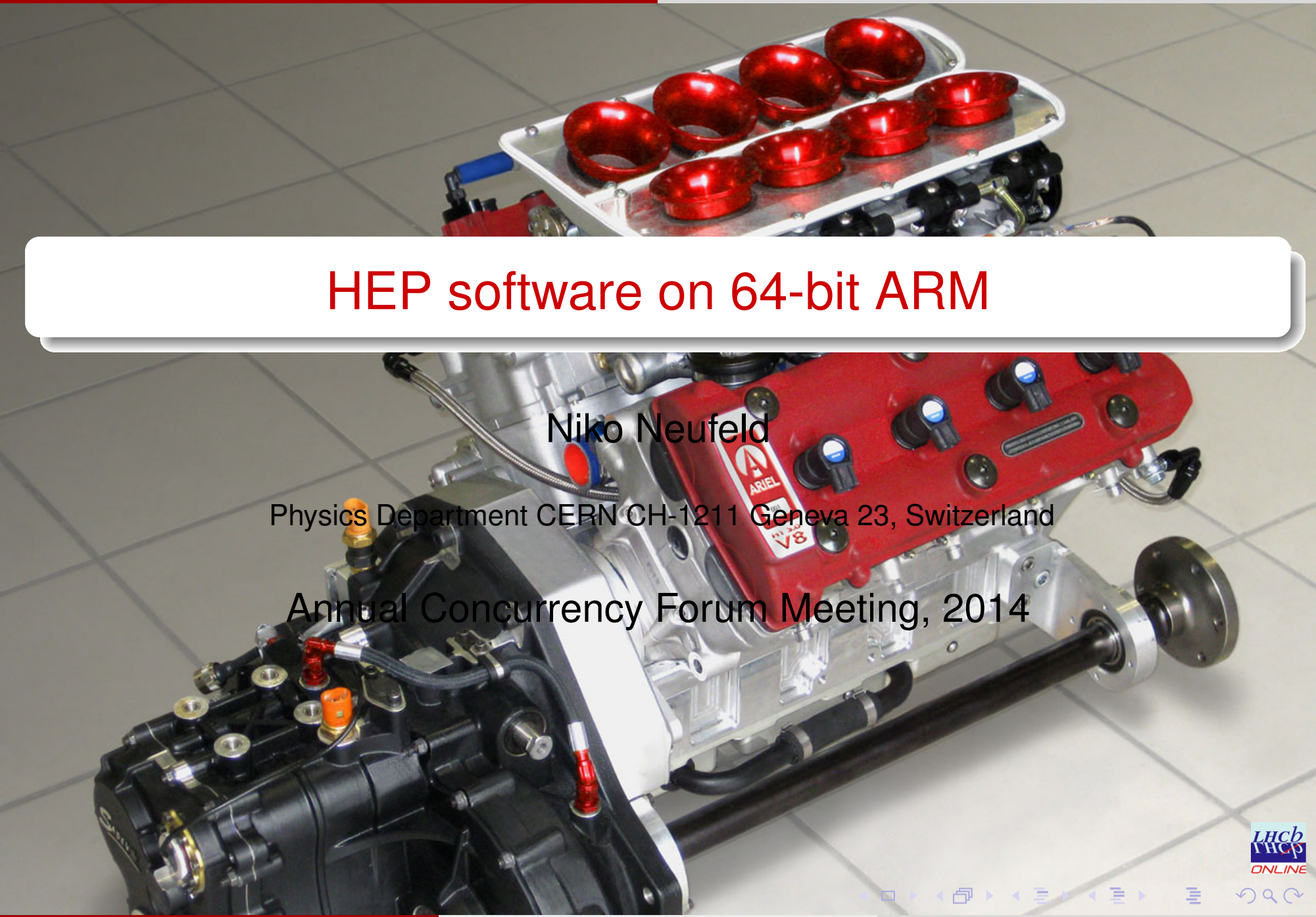


HEP software on 64-bit ARM

Niko Neufeld

Physics Department CERN CH-1211 Geneva 23, Switzerland

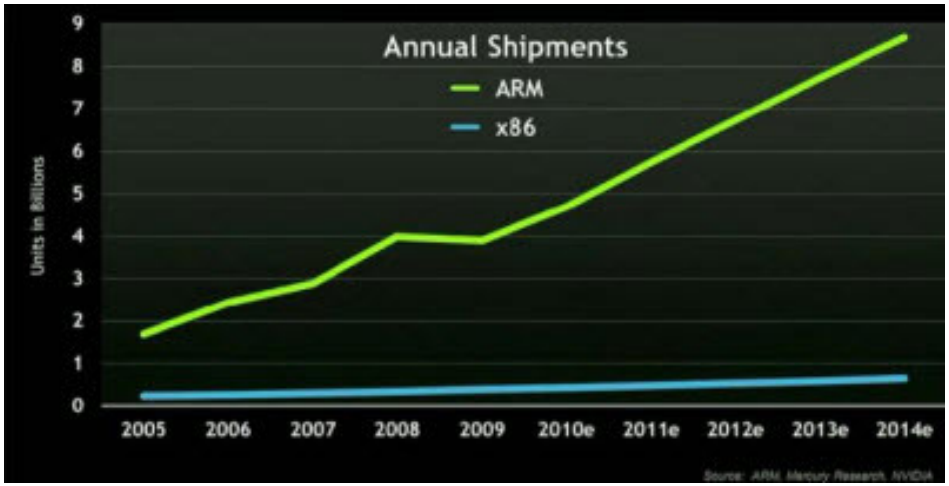
Annual Concurrency Forum Meeting, 2014

Overview

1. Why ARM?
2. Measurements on ARM v7
3. V8

The Market

- Most tablets and smart-phones powered by ARM processors
- 2013 will see more shipments of tablets than laptops
- Server market has virtually no growth, desktops declining



ARM interest at CHEP 2013

- *Measurements of the LHCb software stack on the ARM architecture*
- *HS06 benchmark values for an ARM based server*
- *Explorations of the viability of ARM and Intel Xeon Phi for Physics Processing*
- *Optimization of Italian CMS Computing Centers via MIUR funded Research Projects*
- . . . + a couple of honorable mentionings :-)

All measurements where done on ARMv7 (32-bit architecture)

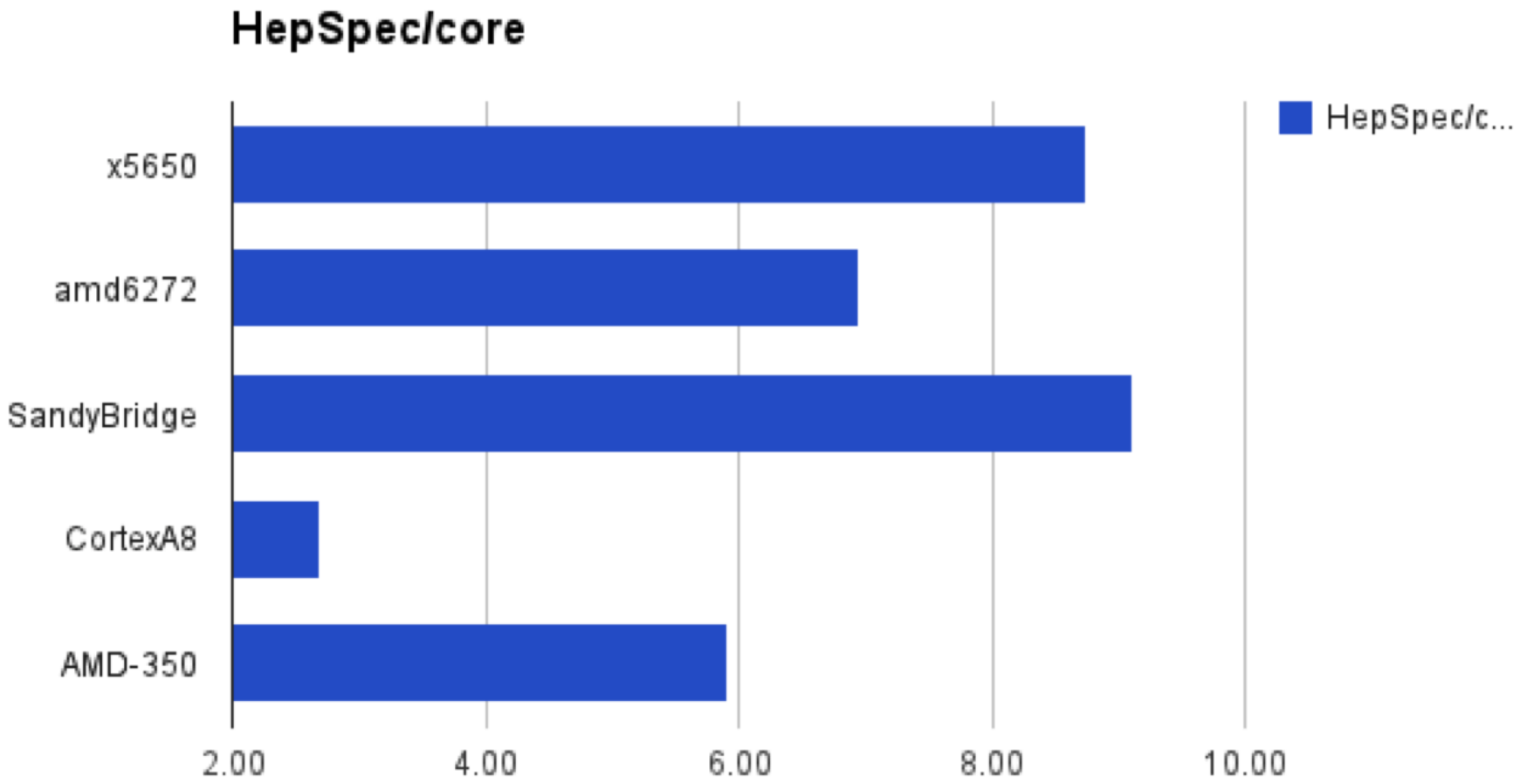
ARM v7

- ARM-based SoCs are widely available now
- Lots of interest in micro-servers (cloud, web-shops)
- Still glue-logic is needed (no fast PCIe/SATA on current SoCs)
- Dense packaging of multiple SoCs
- Example (and used in some of these tests) Boston Viridis (based on Calxeda)

- 48 SoC, 4 cores 4 GB RAM
- ARM A9 Cortex 1.4 GHz (v7 architecture 32-bit)
- 80 Gb Ethernet switch (10 GigE external)
- Total 192 cores / 192 GB RAM /300 Watt
- redundant power, etc. . .

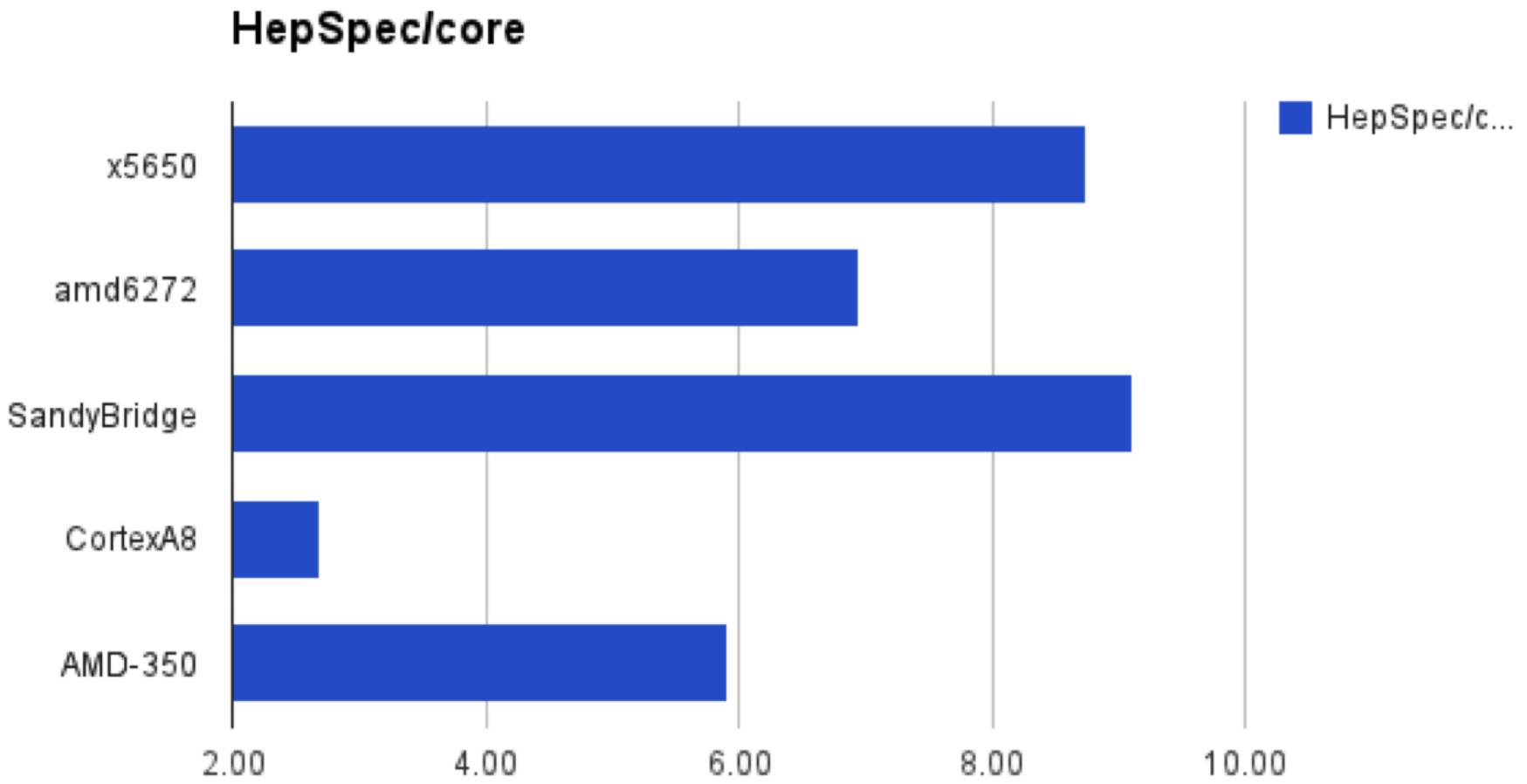
But. . .

First look: HEPSpec



So we need many

First look: HEPSpec



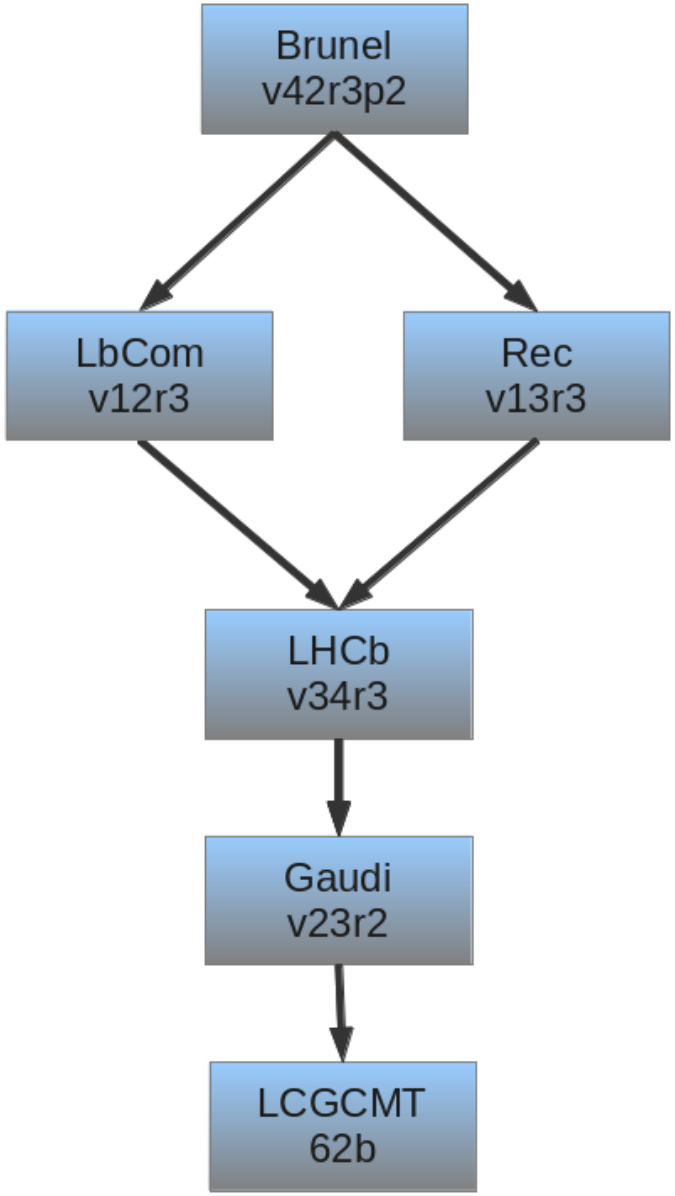
So we need many

HEP spec is not enough - do the real test

- HepSPEC is not necessarily a good test for Online usage
  - Online we (currently) run n instances of the same application in parallel, where n is $\geq$ number of cores/hyperthreads
  - In such a scenario hyperthreading typically adds overproportionally (up to 40% of total machine performance) compared to mixed work-loads
- Need to benchmark a real LHCb work-load: Brunel (the reconstruction program)

The task

- 3.6 MLOC of code (excluding LCG_CMT)
- ROOT v5.34.05
- gcc 4.7.2, Boost 1.51



The platform

- Remote test
  - CPU = Calxeda EnergyCore (SoC), ARM Cortex-A9 CPU, 4 cores, $\approx$ 1.1 GHz, 4GB RAM
  - `Linux cloud12 3.6.10-8.fc18.armv7hl.highbank 1 SMP Tue Jan 29 14:01:38 EST 2013 armv7l armv7l armv7l GNU/Linux`
- Local development, CARMA
  - CPU = NVIDIA Tegra 3, a Quad-core ARM Cortex-A9 CPU $\approx$ 1.3 GHz
  - Ubuntu 11.04
  - `Linux carma-devkit 3.1.10-carma 2 SMP PREEMPT Fri Aug 31 15:28:42 PDT 2012 armv7l armv7l armv7l GNU/Linux`
  - Not a hard-float kernel - (not good)

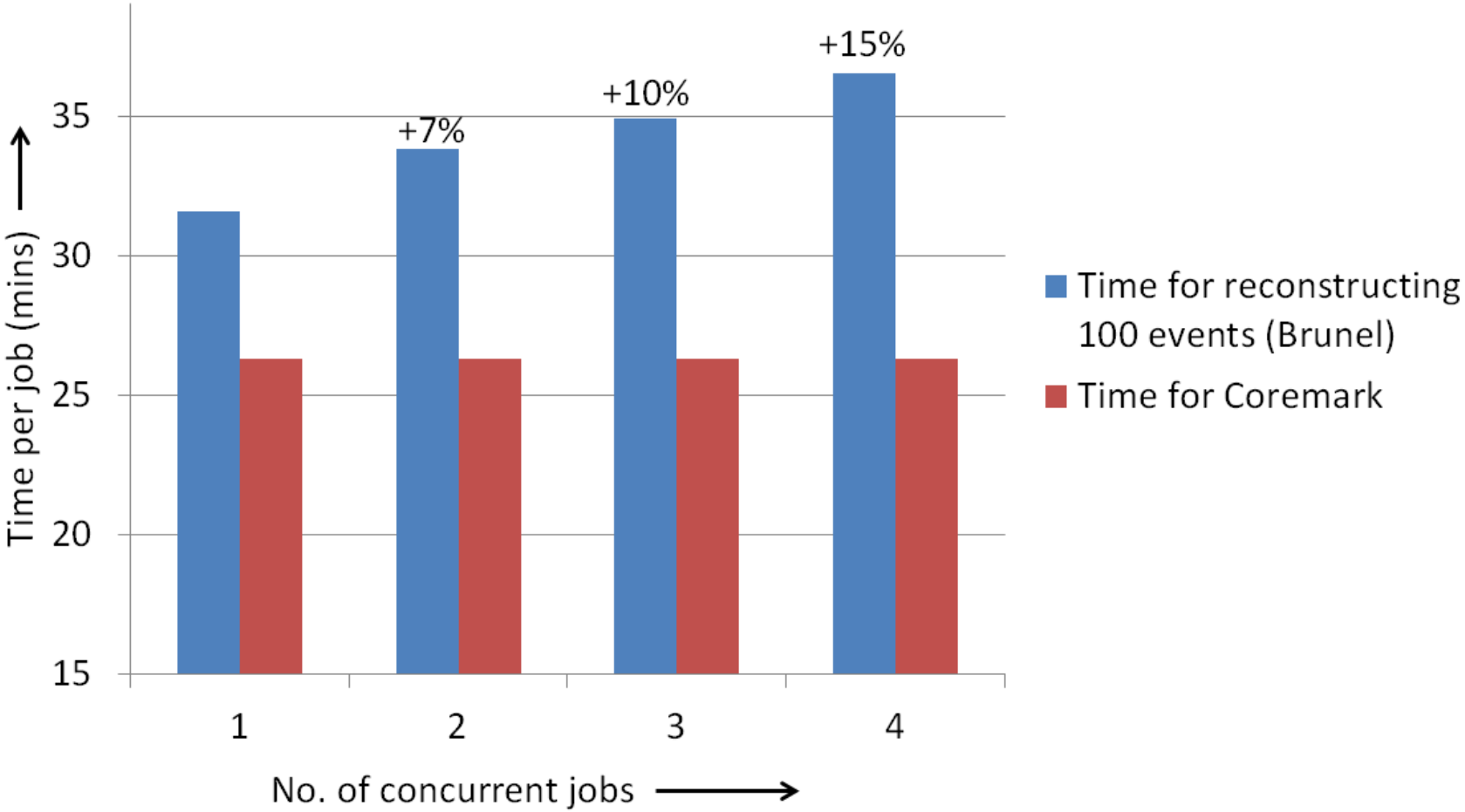
Single core performance

	CARMA	Viridis	x86 L5520
freq. (GHz)	1.3	1.1	2.27
# cores	4	4	8 (with HT)
total RAM (GB)	4	4	48
time (h)	~ 10	~ 2.5	~ 0.5

- Time for reconstructing the same 1000 events in Brunel
- Events read from network share (NFS/AFS)
- Test results stable

Scaling

- Loading the entire machine with identical jobs
- x86 scales **linearly** for real cores, HT add about 40%
- ARM memory bandwidth does not scale



Correctness

"Brem Match"	sum		mean/eff^*		rms/err^*	
	ARMv7	x86_64	ARMv7	x86_64	ARMv7	x86_64
#calos	50085	50085	60.489	60.489	30.140	30.140
#chi2	2.737109e+09	2.737105e+09	5009.1	5009.1	2866.4	2866.4
#links	546430	546415	659.94	659.92	611.38	611.33
#overflow	403843**4**	403843**0**	4877.**3**	4877.**2**	5074.**2**	5074.**0**
#tracks	5861**0**	5860**9**	70.78**5**	70.78**4**	48.51**3**	48.51**1**

On the way

- Many small problems with very recent OS releases (FC18, Ubuntu 12) - however ARM is a primary architecture for Fedora as of FC20
- Kernel updates a bit more complicated, every platform needs specific patches, because there is no "ARM-PC"
- Delicious architecture specific problems for the conaisseur
  - x86-icisms (e.g. sizeof empty struct)
  - Bad instructions issued by compiler (refused by assembler), toolchain problem?
- ROOT (v5.34.05) cintex not working completely (test fails), but subset required by LHCb seems to work. CMS still has problems with root persistance in ROOT5

Now

- Micro-servers based on v8 are coming
- Lessons learned
- Seem to have shed the tablet heritage

*

Goals for ARM v8 server SoCs

- Single core performance comparable to IvyBridge at comparable clock-speed
- Power consumption significantly lower (SoC power budget 20 - 40 W)
- SoC cost significantly lower than Intel
- no turbo, no HT
- virtualization support

V8 players

- X-Gene by Applied Micro
  - 10/40 G Ethernet
  - ECC memory (DDR3)
  - BMC on SoC
  - Shipping: now
- AMD Opteron A “Seattle”
  - 4MB L2 cache 8MB L3 cache (shared)
  - up to 128GB DDR3 or DDR4
  - 2 x 10G Ethernet
  - PCIe Gen3 x8
  - 8x Sata 3
  - Shipping: Q3 2014
- **Rumours** abound: Nvidia (project “Denver”), Samsung (with Google?), your favourite conspiracy here...

In lieu of conclusions: What's next?

- Let's port (again) - redo CMS and LHCb exercises from v7 on v8
- Fix remaining issues
- Share (and enjoy) results
- Try to leverage support from vendors - **this is the moment for early show-case projects!**

Thanks

- to my co-authors in the original LHCb study: V. Kartik, B. Couturier, M. Clemencic
- to Pete Elmer for interesting input about the CMS work
- to all members of the SFT group who supported us
- and to Boston HPC, UK, who kindly provided remote access to their Viridis platform