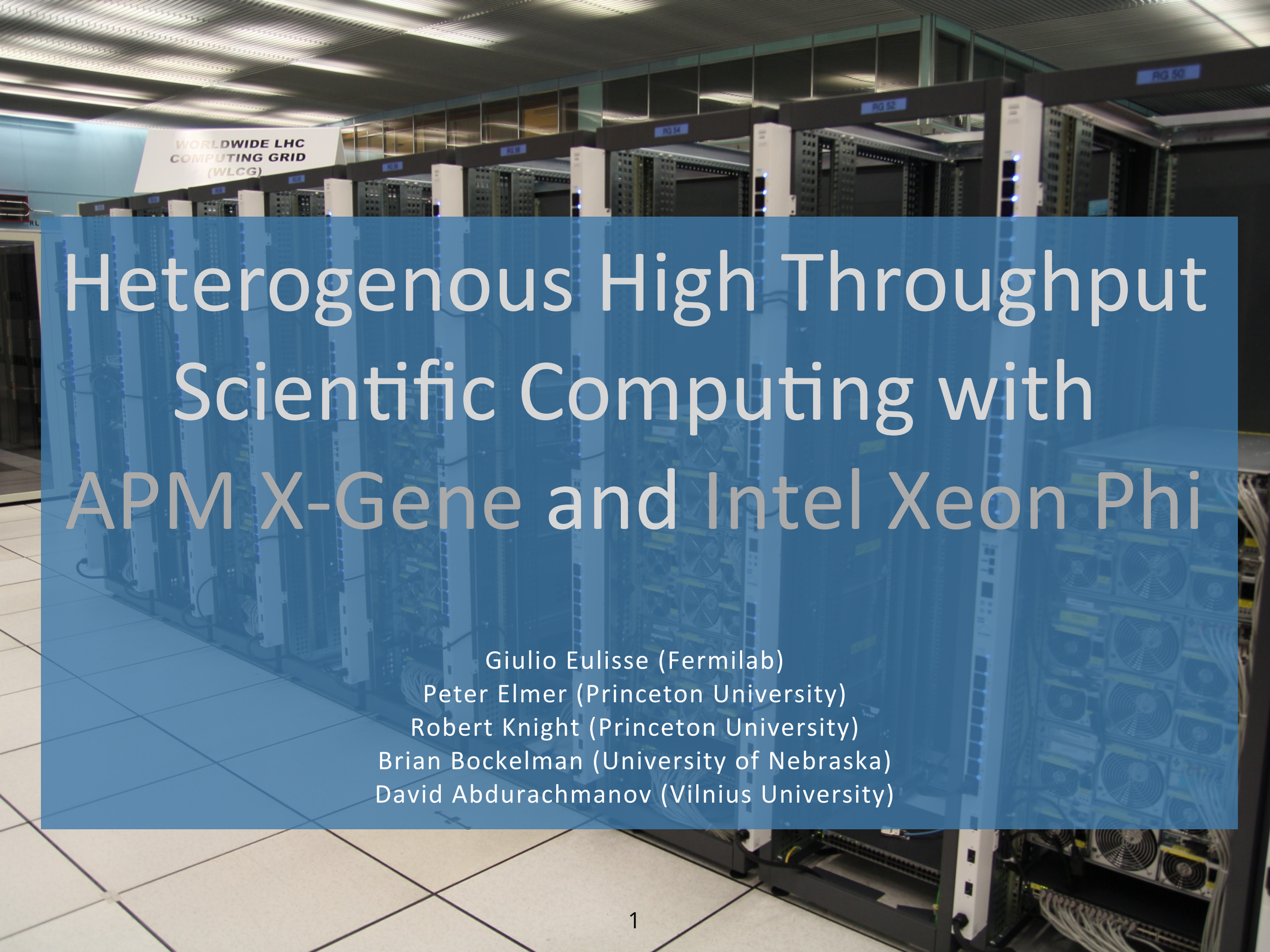


Heterogenous High Throughput Scientific Computing with APM X-Gene and Intel Xeon Phi

Giulio Eulisse (Fermilab)
Peter Elmer (Princeton University)
Robert Knight (Princeton University)
Brian Bockelman (University of Nebraska)
David Abdurachmanov (Vilnius University)

Power in Data Centers

Why is the power important?

> Energy-related costs account for approximately 12 percent of overall data center expenditure and are the **fastest-rising cost in the data center**, according to Gartner, Inc. (September 29, 2010)

CMS for 2012 data used **80 000** - **100 000** x86_64 cores from **350 000** cores at Worldwide LHC Computing Grid (WLCG)

Scaling up from the mix of machines at FNAL we estimate WLCG aggregate power consumption for machines at **10MW**

CMS expects 2 to 3 orders of magnitude increase in data produced in 15 years

Two directions became common on the market

> **Local green or/and cheaper power source**, e.g., Princeton University
>
> Princeton energy plant (15MW) combines electricity, heat and cooling. When electricity cost increased gas, diesel or/and bio-diesel fuel is used to power local generators. Hot water and steam is provided from waste energy.
>
> **Low-power and highly efficient hardware**, e.g., Xeon Phi and X-Gene (ARMv8 64-bit)

A single 7U system consumes
full rack power (3.5kW)

Investigation Goals

Provide CMS Software (CMSSW) for **non-x86_64** platforms

> CMSSW is software distribution with CMS code and other projects incl. ROOT, GEANT4, FastJet, Python, Boost and similar

Understand how our software scales on highly parallel machines as Intel Xeon Phi (formerly Knights Corner), and how it behaves in low-power high-computing-density **general purpose** systems as based on APM X-Gene 1 ARMv8 64-bit silicon

Understand not only the obtainable performance, but also the power utilization

Showcasing that alternative **general purpose** silicon (64-bit) could be a drop-in replacement for x86_64

Each iteration, i.e., porting to a new architecture or/and Linux distribution, allows us to locate issues within our build tools and code base

Xeon Phi

Intel Xeon Phi (aka MIC) is a highly parallel machine with 512-bit long vectors. Twice of what is provided in current gen Intel Xeon silicon

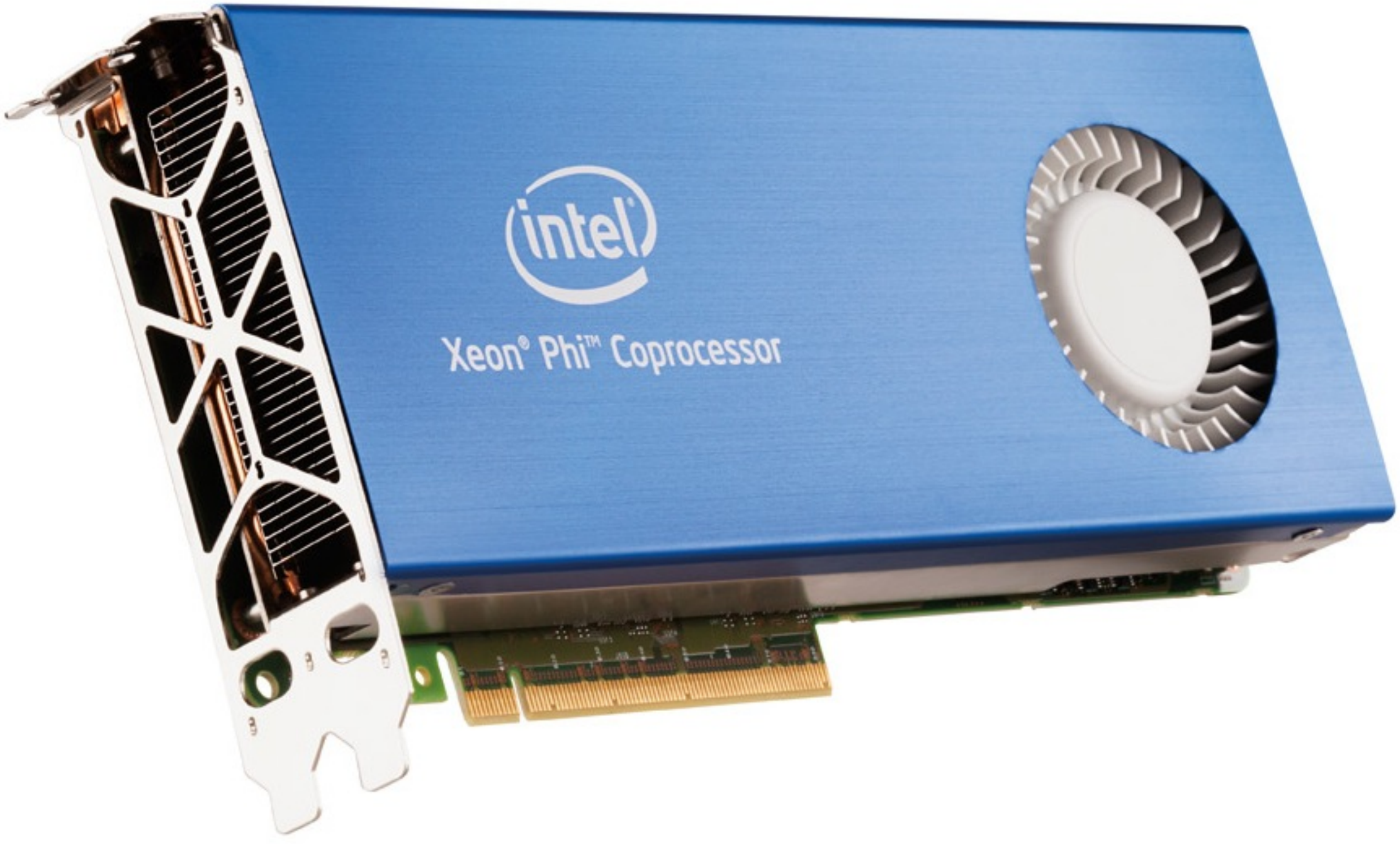
MIC is capable of 4-way multi-threading with at least 2+ threads required to fully utilize the hardware

No official support in upstream GCC and **Intel C++ Compiler (ICC)** is required for auto-vectorization

MIC is available as PCI-E add-on card and host Intel Xeon silicon is required

Two modes supported

- **Native** execution using it as a standard Linux machine
- **Offloading** similar to GPU via multiple software solutions


intel
Xeon® Phi™ Coprocessor

CMSSW for Xeon Phi

We have partial and working port of CMSSW for Xeon Phi

- CMS builds tools were updated to support building Xeon Phi native software
- CMSSW development environment is provided for Xeon Phi
- Intel C++ Compiler (ICC) does not support all C++11 features needed for CMSSW
- Xeon Phi is based on experimental GCC 4.7 and 4.8 is required for full C++11
- ICC is an optimized compiler for Intel architecture

Agreement was reached with Intel to have CMSSW tested with upcoming ICC releases

- Intel was provided with latest CMSSW code base and they are working on solving issues in their compiler
- A major update comes with ICC 15.0 released on 2014 August 15 (few days ago)

We have managed to compile and successfully execute a few years old CMSSW version with ICC for x86_64

- This CMSSW does not use C++11
- Bugs located in ICC and workaround were required

ARMv8 64-bit

ARMv8 introduces 64-bit ISA and is capable executing older 32-bit code

ARM Ltd. is a fabless* semiconductor company with licensing their intellectual property (IP) to partners

- This is a major difference comparing to Intel, a semiconductor foundry
- ARM provides ISA licenses for partners to create a custom silicon solutions for wide market applications (IoT, smartphones, servers and others)
- Additionally ARM delivers already designed and verified building blocks, e.g., CPU or GPU for partners

The industry is interested into a pure 64-bit only Linux distribution

It will be as a standard PC

- You will be able to buy hardware and then load your desired Linux distribution
- Multiple standards were established by ARM Ltd. and partners to make experience consistent between vendors

Linaro, a non-profit organization, is responsible for porting and optimizing free and open-source software for ARMv7 and ARMv8

***** Does not have their own semiconductor fabrication plant

X-Gene 1 ARMv8 64-bit Server-on-Chip

The first **server-grade ARMv8 64-bit** silicon on the market, **X-Gene 1** (40 nm), released by **AppliedMicro (APM)**

Upstream Linux support available since 1H 2013 and first patchset adding X-Gene 1 support to GCC was published late last year

APM is a group member of Linaro

We have been working with APM for more than a year

APM has provided Princeton and CERN an XC-1 ("Mustang") development platform for porting efforts on the X-Gene 1 processor.

APM Mustang is a development board intended primarily for speeding up development efforts


applied micro
MAC:00:01:73:01

CMSSW for ARMv8 64-bit

The initial port to ARMv8 64-bit was done on emulation software followed by a real hardware solution from APM

CMSSW for ARMv8 64-bit packages are available in the official CMS repository and standard installation instructions apply

It is available for GRID sites on CernVM File System (CVMFS) standard location (**/cvmfs/cms.cern.ch**)

Working is underway on a demo cluster located at Princeton

We have successfully used CVMFS and GRID-required software on APM Mustang

- CMSSW jobs are able to fetch input data from remote location via xrootd
- We can schedule a job via x86_64 master to ARMv8 64-bit working node with HTCondor software

Ported **IgProf** (memory and performance profiler): **igprof.org**

- Details on IgProf port for ARMv8 64-bit and energy profiling is provided in the next talk, **"Techniques and Tools for Measuring Energy Efficiency of Scientific Software Applications"**

Ongoing ARMv8 64-bit Ecosystem Development

No **proprietary** Oracle Instant Client is available for ARMv8 64-bit

- Oracle Instant Client is only provided as binary blobs, no source code is available
- None of standard CMS workflows depends on Oracle

We added ARMv8 64-bit support to ROOT5

- Only feature not supported is ROOT Reflex dictionary generation
- Pre-generated dictionaries from x86_64 can be successfully compiled and executed on ARMv8 64-bit

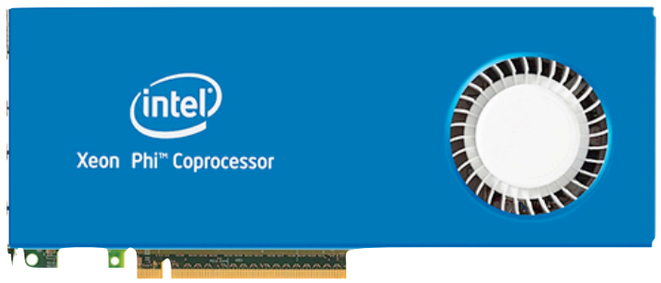
ROOT6 does not provide ARMv8 64-bit support yet, but work is undergoing to move to a new LLVM/Clang

Minor external packages updates were required

- Outdated configuration script did not know of ARMv8 64-bit
- Some packages were updated to a newer versions

Majority of issues were resolved by our efforts porting to ARMv7 32-bit and Fedora Linux distribution

Silicon Specifications

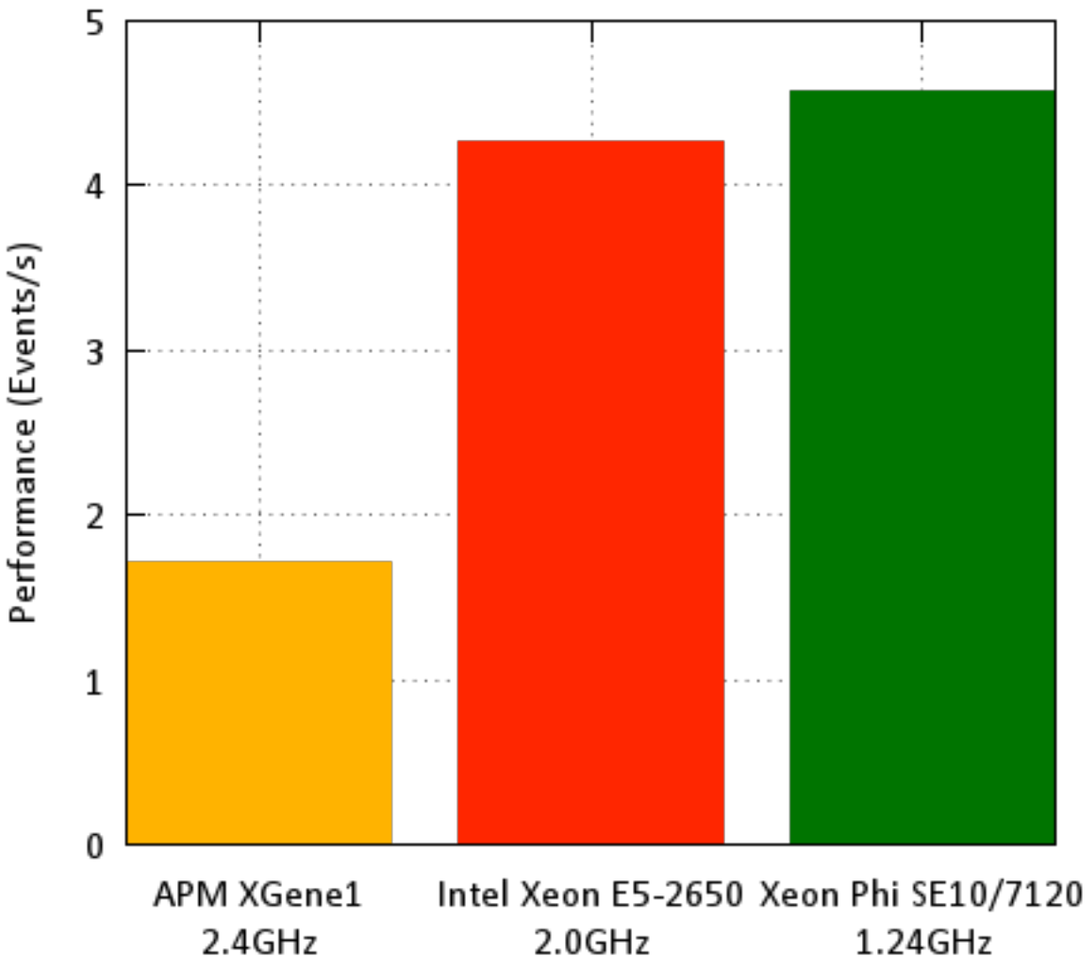
			
	X-Gene 1*	**Xeon E5-2650****	**Xeon Phi SE10/7120**
Physical cores	8	8	61
Threads per core	1	2	4
Total threads	8	16	244
L1d/L1i	32K/32K	32K/32K	32K/32K
L2	256K	256K	512K
L3	8192K	20480K	Don't have
Frequency	2.4GHz	2.0GHz	1.24GHz
Memory	16GB (DDR3)	256GB (DDR3)	16GB (GDDR5)

We are measuring power levels at silicon level via **on-board sensors**, e.g., **RAPL**

***** Measured on the XC-1 platform, which provides 2 Memory Channels (X-Gene 1 supports a total of 4 Memory Channels and higher memory capacities)

****** Machine is dual-socket system, but only a single socket was measured for a more direct comparison

Performance per Architecture



We used multi-threaded version of the standalone **GEANT4** benchmark application, **FullCMS**

FullCMS uses **complex CMS detector geometry**

No 3-way comparison is possible using CMS Software (CMSSW)

- Xeon Phi does not have a full port of CMSSW

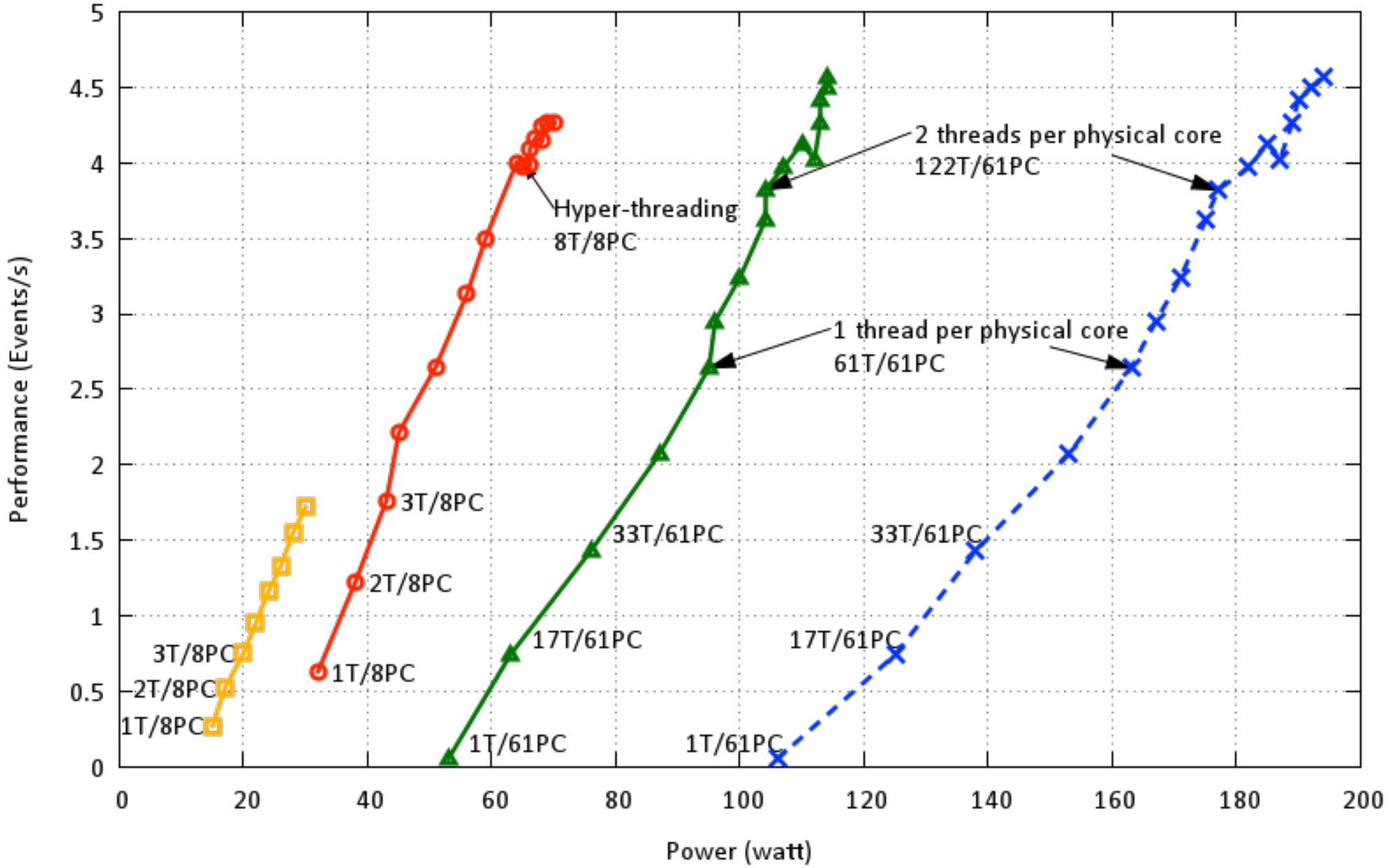
Default configuration options were set

X-Gene 1/Xeon uses **GCC** and Xeon Phi uses **ICC**

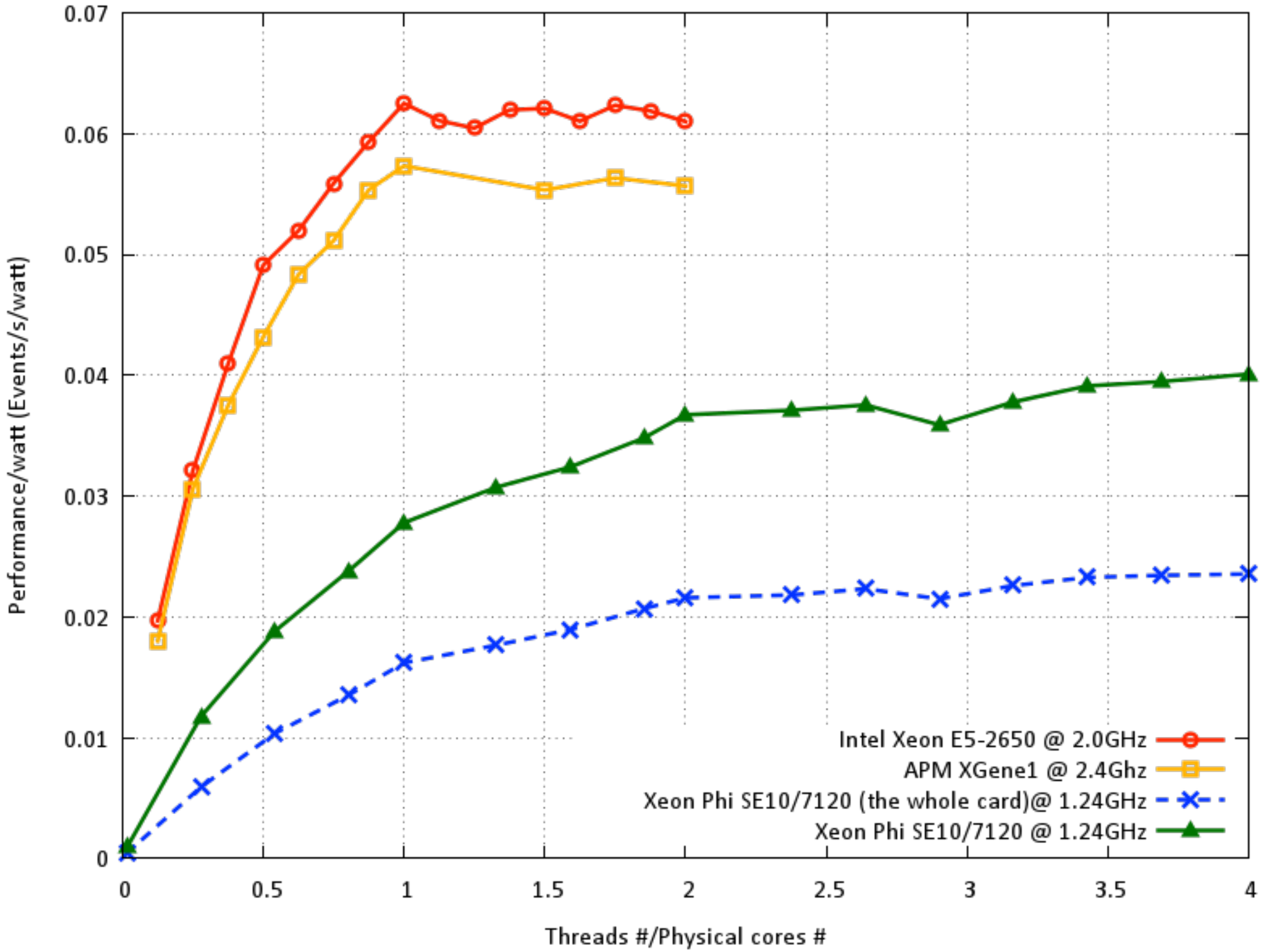
The benchmark does not fully utilize Xeon Phi hardware long vectors

- See poster **"Traditional Tracking with Kalman Filter on Parallel Architectures"** in poster session (starts tomorrow) for Xeon Phi optimized software benchmarks

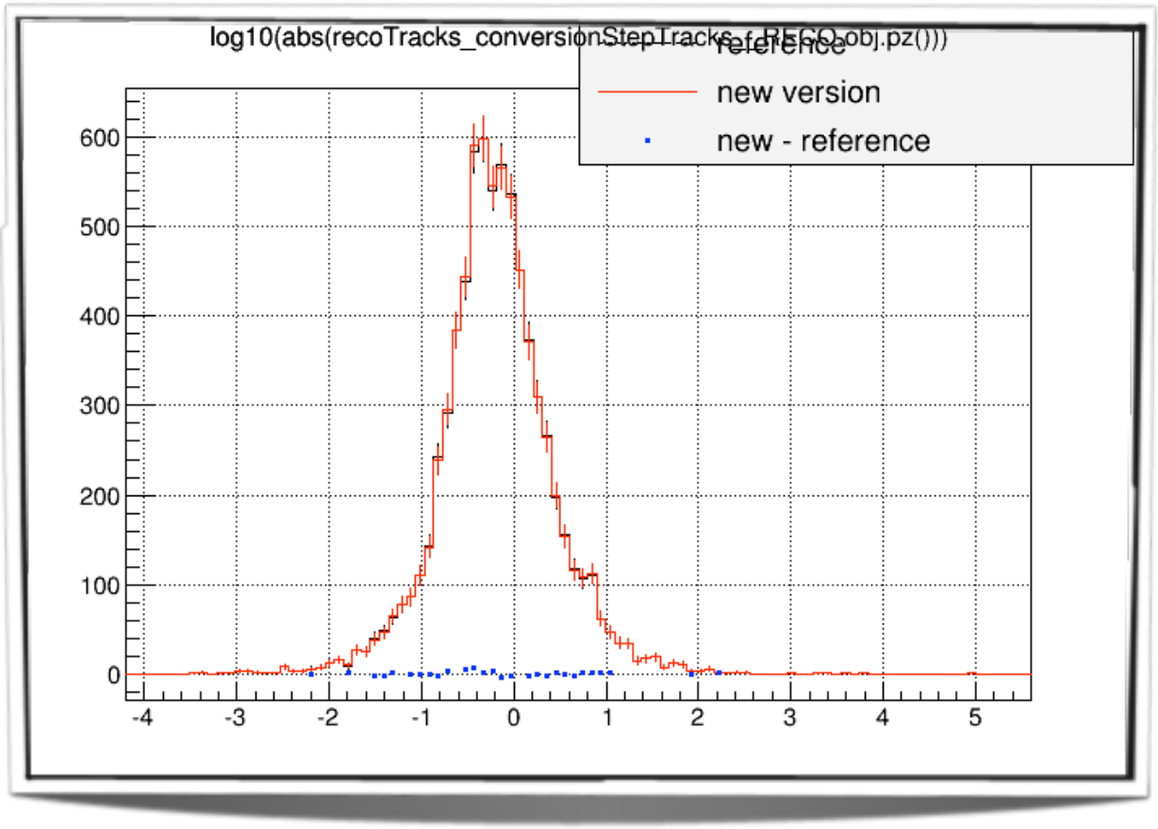
Perf. Scalability / Power

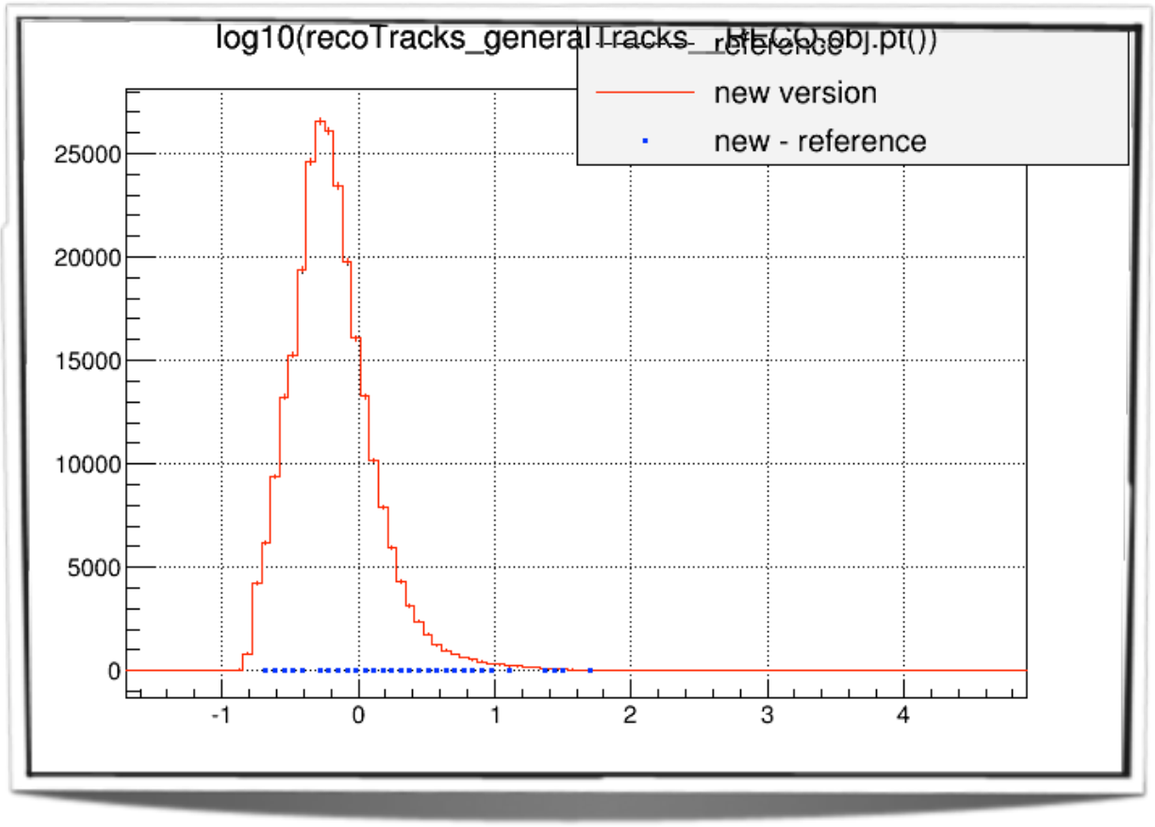


Perf./watt / Threads



ARMv8 64-bit Validation





Initial validation of CMSSW for ARMv8 64-bit was done using **reconstruction (RECO)** software workflow

We have **observed small differences**, which are being investigated

What's next?

Provide regular CMSSW for ARMv8 64-bit nightly integration builds

Expand ARMv8 64-bit physics content validation by using higher statistics and larger set of standard samples

Repeat power consumption measurements for the X-Gene 1 with an updated firmware

- Support for some ACPI processor states was not available in time for this study
- Performance/watt results are anticipated to improve with the use of ACPI

Repeat performance measurements using optimized compiler for X-Gene 1 ARMv8 64-bit

Repeat performance and power consumption measurements with **X-Gene 2***

Provide a full port of OSG software stack

Submit first CMSSW job to a ARMv8 64-bit worker node

Profile performance differences between x86_64 and ARMv8 64-bit using updated **IgProf (igprof.org)**

***** X-Gene 2 (28nm) silicon is currently sampling

Summary

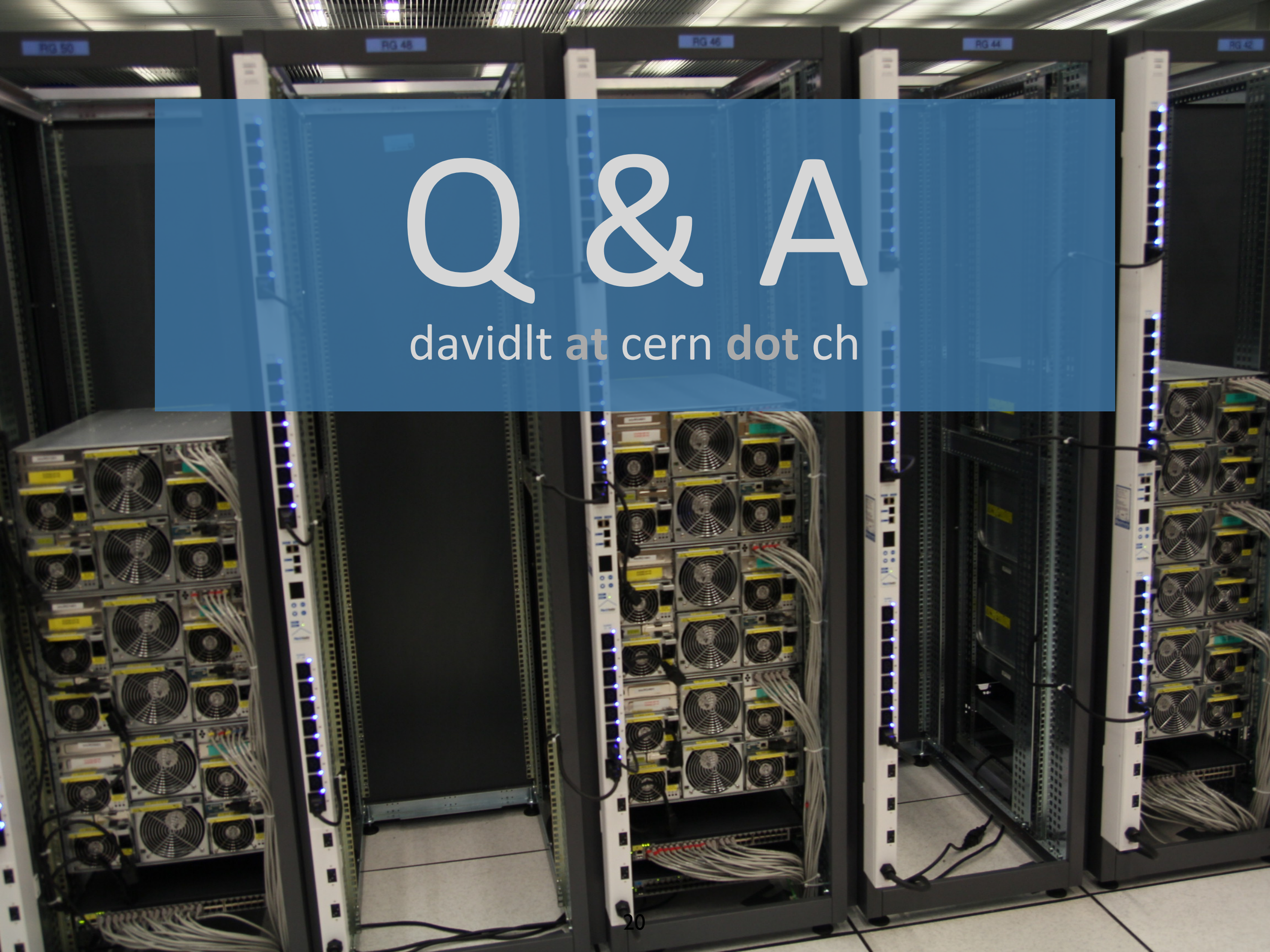
We have built a functional CMSSW multi-core aware framework for Intel Xeon Phi

We have CMSSW for ARMv8 64-bit available in the official CMS packages repository and CVMFS distributed file system used by GRID sites

The initial validation proves APM X-Gene 1 ARMv8 64-bit is a relevant platform for Heterogeneous High-density-computing

We have ported part of OSG software stack required for a worker node to ARMv8 64-bit

Using the current software Intel Xeon is the most energy efficient hardware and Intel Xeon Phi significantly behind



Q & A

davidlt at cern dot ch