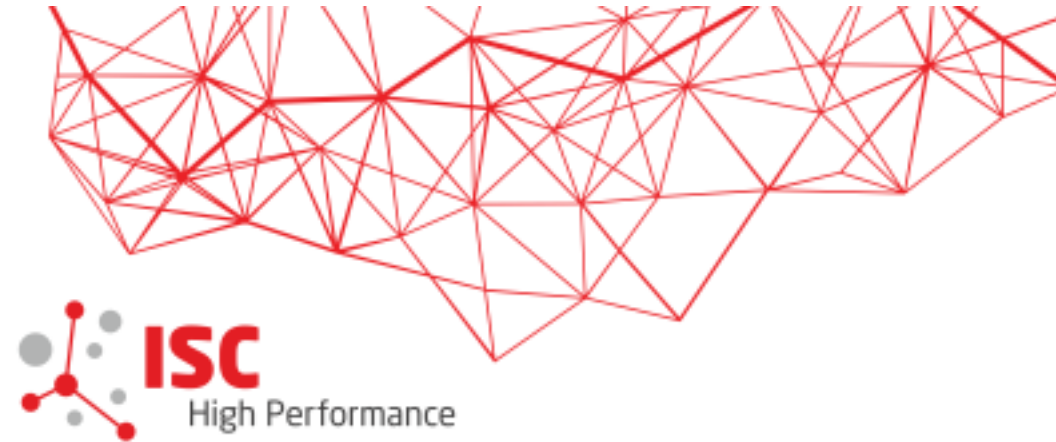


ISC HIGH
PERFORMANCE

JUNE 18 - 22, 2017
FRANKFURT,
GERMANY



Trip Report

Sofia Vallecorsa

SFT group meeting - July 2017

2017 Focus Topics



“ISC High Performance focuses on HPC technological development and its application in scientific fields, as well as its adoption in commercial environments”

HPC Systems

- Exascale System Developments (Architecture & Concepts)
- Processor Technologies for HPC and AI
- Memory Outlook
- Interconnects for HPC Systems
- Programming Models (Concepts for Exascale)

HPC Applications & Algorithms

- Life Sciences
- Energy Exploration
- Turbulences & Combustion
- Advanced Material Science
- Algorithms for Extreme Scale in Practice
- Large Scale Engineering & Cloud Computing

HPC Trends & Challenges

- High Performance Visualization
- Big Data Experiments & Big Data Analysis
- Quantum Annealing for Combinatorial Optimization Problems

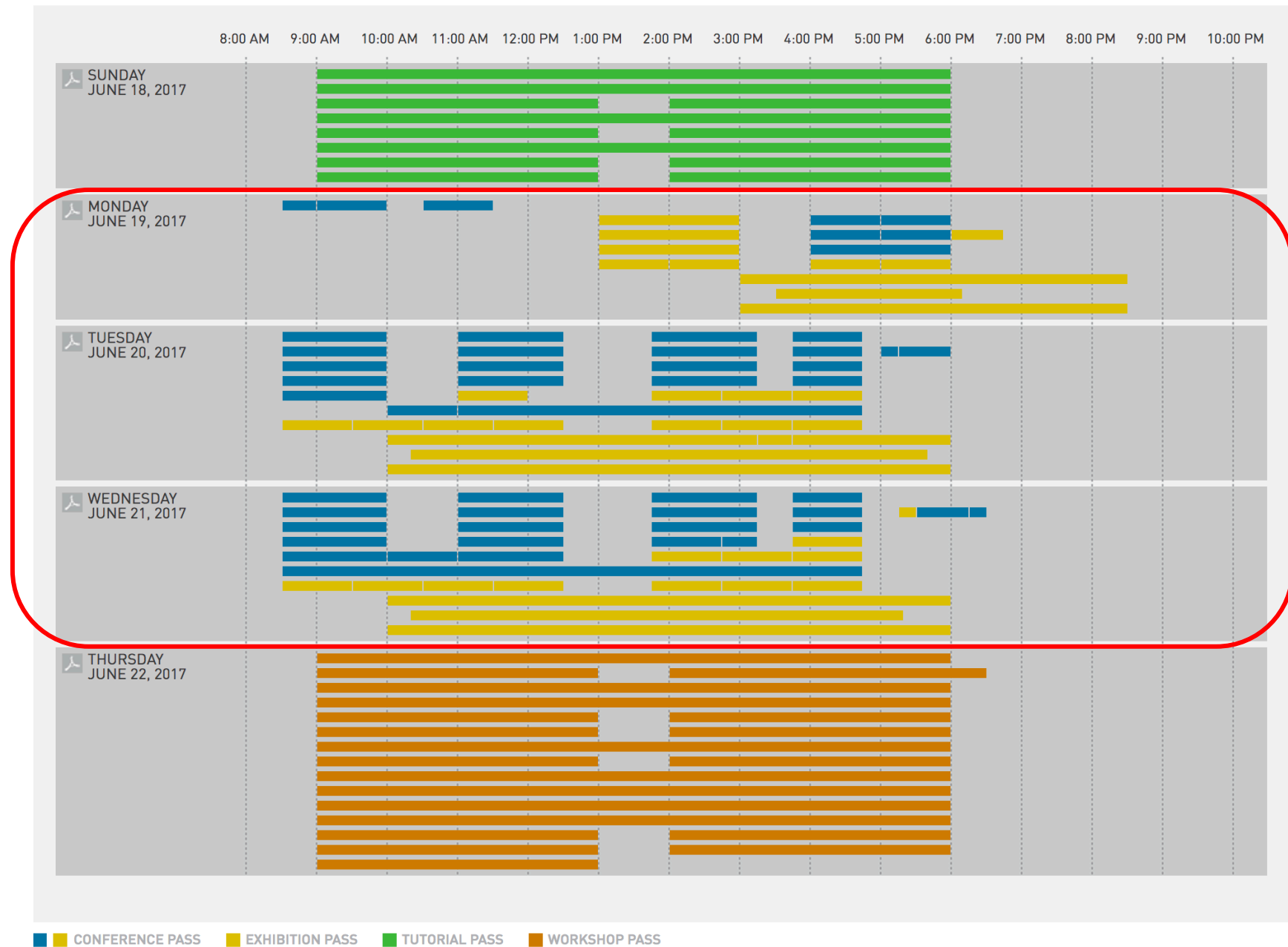
Deep Learning Day (Changing the HPC Landscape)

Industrial Day (HPC for Industry)

Overview

- 3 days conference and exhibition
- Tutorials
- Workshops

- 450 speakers
- 150 exhibitors
- 3200 attendees



41.58%

Academia

43.91%

Industry

10.86%

Students

3.64%

Media

Technical program

ISC 2017 TECH PROGRAM

- ISC 2017 had 275 submissions in total:
 - 65 research papers; 22 accepted.
 - 37 research poster submissions; 21 accepted.
 - 35 project poster submissions; 25 accepted.
 - 23 PhD forum submissions; 13 accepted.
 - 47 BOF submissions; 26 accepted.
 - 42 Tutorial submissions; 13 accepted.
 - 26 workshop submissions; 21 accepted.
- Of the 275 submissions 141 accepted.

GeantV - Next generation simulation prototype

Sofia Vallecorsa¹ for the GeantV Project CERN, Sofia.Vallecorsa@cern.ch

The project

Introduction and Motivation

Simulation of particle transport through matter is fundamental for understanding the physics of High Energy Physics (HEP) experiments, as the ones at the Large Hadron Collider (LHC) at CERN. Such experiments have dedicated so far most of their worldwide distributed CPU budget – in the range of half a million CPU-years equivalent – to simulation. In particular, the most computing-intensive components are geometry modelling, navigation through millions of objects and physics models.

Parallel particle transport

The scope of the project is the development of a community supported, open-source, next generation particle transport code for HEP integrating both detailed and fast simulation algorithms, optimized for the emerging parallel and vector architectures. Beta release by end of 2018.

- Group particles by locality into vectors (baskets)
- Invoke geometry to determine particle position
- Invoke physics to predict stochastically a process location
- Validate proposed physics step against geometry
- Propagate and regroup baskets

Geometry redesign for vectorization

VecGeom is a geometry modeler evolved from legacy geometry libraries (Genie4, Isolids, ROOT). It introduces a many-particle API besides the standard single one, and relies on templated backend abstraction to enable both platform/architecture specific optimizations and vector/scalar API polymorphism.

VecGeom = Evolved Isolids + many-particle API + geometry model/navigation

GeantV is a collaboration among several research institutes. It is also partially funded by Intel Parallel Computing Center program (geant.cern.ch)

Achievements

Geometry Performance

VecGeom can run chains of algorithms in parallel on multiple processors.

Scalar navigation performance in full CMS (Compact Muon Solenoid, one of the major LHC experiments) geometry. Left, real time for the simulation of 10 pp events at 7TeV using VecGeom instead of legacy existing geometry. Right, the resident memory of the full application. (Platform: Intel Haswell)

Platform	AVX2	AVX512	AVX-512
Intel Ivy Bridge (AVX2)	-2.9x	-4x	4x
Intel Haswell (AVX2)	-1x	-5x	4x
Intel Xeon Phi (AVX-512)	-4.1	-4.8	8x

Intel Xeon Phi for detector simulations

The X-Ray benchmark tests geometry navigation in a real detector geometry: a detector volume is scanned with virtual rays along a given direction. Each ray is propagated from boundary to boundary and the number of crossings is counted.

Scalar case: Simple loop over pixels, generating a ray ideal vectorization case: Fill vectors with N times the same X-Ray, using this as reference for the maximum achievable vectorization

Realistic (basket) case: Fill baskets per geometry volume as particles are entering (as in GeantV)

X2 gain from vectorization when filling all vector pipelines for AVX512 wrt AVX2

Real time VecGeom versus ROOT geometry

Resident memory VecGeom versus ROOT

Intel Xeon Phi for detector simulations (continued)

Simplified geometry emulating a detector tracking system.

A pixel is produced for each ray having a grey value proportional to the number of crossings.

Dispatching one full scan (image) per task

The full prototype

- Exercise at the scale of LHC experiments
- Full geometry + uniform magnetic field
- Tabulated physics, 1MeV energy threshold
- Full track transport and basketization procedure
- Compare scalability to classical approach single-thread

ETL throughput (events/second) vs. threads

Development

Sub-node clustering

- Known scalability issues due to re-basketizing synchronization
- Deploy several propagators to cluster resources at sub-node level
- Improve scalability
- Address many-node and multi-socket (HPC) modes + non-homogenous resources

First tests on KNL

GeantV for HPC environment

Standard mode (1 process per node)

- Need work balancing
- Check output granularity [merging may be required]

Multi-tier mode (event servers)

- Gets events from file, handles merging and workload balancing
- Communication with event servers via MPI

Machine Learning for fast simulation

A faster approach is to treat traditional simulation as a black-box and replace it by a deep learning algorithm trained on different particle quantities. We are testing several techniques such as generative adversarial networks (GANs) to replace the Monte Carlo approach. We expect to achieve a significant speedup (x25) with respect to GeantV full simulation approach. Development of such tool can further benefit other fields, such as radioactivity protection, environmental modeling and medicine.

Single particle energy deposits in the Linear Collider Detector calorimeters generated by a GAN

Networks

- Conference keynote by Jennifer Tour Chayes, Microsoft Research
- Dedicated Deep Learning day:
 - How Deep Learning is Changing the HPC Landscape
 - Autonomous driving an Connected Vehicles



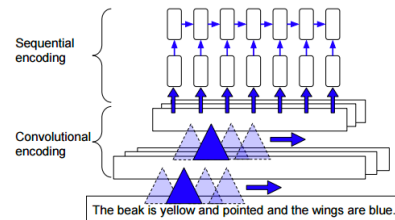
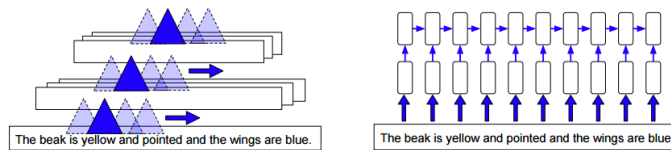
Generative Adversarial Networks Architecture for Image Synthesis from Text

Zeynep Akata

Max-Planck Institute for Informatics

June 21, 2017

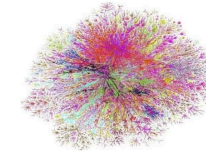
Deep Representations of Visual Descriptions



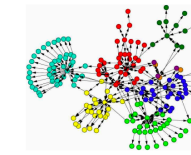
[Reed and Akata et.al. CVPR'16]



Motivation: The Age of Networks



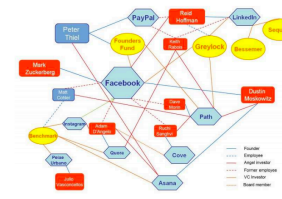
Internet



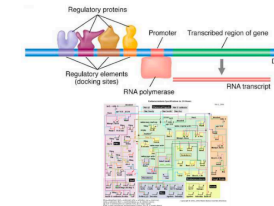
WWW



Social network



Economic network



Gene regulatory network



Neural network



Mathematical and Algorithmic Problems on Networks

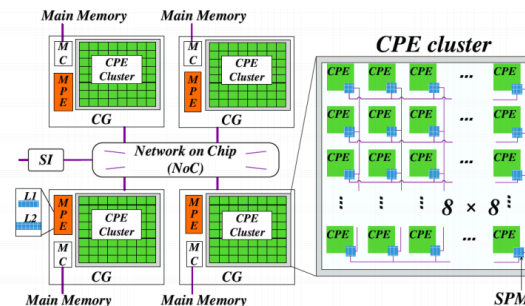
- Modeling networks
- Sampling from and machine learning of networks
- Processes on networks
- Algorithms on networks
- Network reconstruction algorithms

HPC

- The HPC event in Europe
- Showcase of most powerful systems

SUNWAY TAIHULIGHT

- SW26010 processor (Chinese design, ISA, & fab)
- 1.45 GHz
- Node = 260 Cores (1 socket)
 - 4 – core groups
 - 32 GB memory
- 40,960 nodes in the system
- 10,649,600 cores total
- 1.31 PB of primary memory (DDR3).
- 125.4 Pflop/s theoretical peak
- 93 Pflop/s HPL, 74% peak
- 15.3 Mwatts water cooled
- 3 of the 6 finalists for Gordon Bell Award@SC16



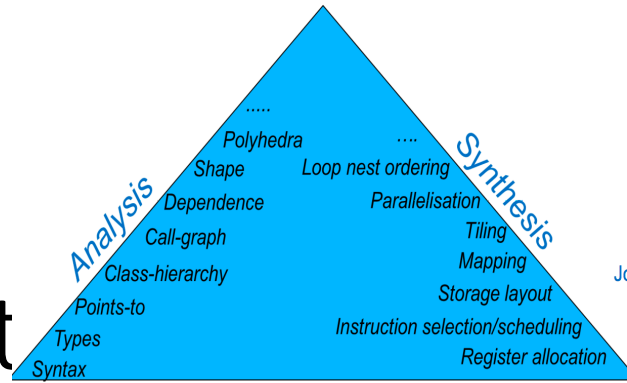
#	Site	Manufacturer	Computer	Country	Cores	Rmax [Pflops]	Power [MW]
1	National Supercomputing Center in Wuxi	NRCPC	Sunway TaihuLight NRCPC Sunway SW26010, 260C 1.45GHz	China	10,649,600	93.0	15.4
2	National University of Defense Technology	NUDT	Tianhe-2 NUDT TH-IVB-FEP, Xeon 12C 2.2GHz, IntelXeon Phi	China	3,120,000	33.9	17.8
3	Swiss National Supercomputing Centre (CSCS)	Cray	Piz Daint Cray XC50, Xeon E5 12C 2.6GHz, Aries, NVIDIA Tesla P100	Switzerland	361,760	19.6	2.27
4	Oak Ridge National Laboratory	Cray	Titan Cray XK7, Opteron 16C 2.2GHz, Gemini, NVIDIA K20x	USA	560,640	17.6	8.21
5	Lawrence Livermore National Laboratory	IBM	Sequoia BlueGene/Q, Power BQC 16C 1.6GHz, Custom	USA	1,572,864	17.2	7.89
6	Lawrence Berkeley National Laboratory	Cray	Cori Cray XC40, Intel Xeons Phi 7250 68C 1.4 GHz, Aries	USA	622,336	14.0	3.94
7	JCAHPC Joint Center for Advanced HPC	Fujitsu	Oakforest-PACS PRIMERGY CX1640 M1, Intel Xeons Phi 7250 68C 1.4 GHz, OmniPath	Japan	556,104	13.6	2.72
8	RIKEN Advanced Institute for Computational Science	Fujitsu	K Computer SPARC64 VIIIfx 2.0GHz, Tofu Interconnect	Japan	795,024	10.5	12.7
9	Argonne National Laboratory	IBM	Mira BlueGene/Q, Power BQC 16C 1.6GHz, Custom	USA	786,432	8.59	3.95
10	Los Alamos NL / Sandia NL	Cray	Trinity Cray XC40, Xeon E5 16C 2.3GHz, Aries	USA	301,0564	8.10	4.23



Software: 65 years of compiler development

Compiler construction as tax evasion

Two topics: *taxation and representation*



Paul Kelly
Group Leader, Software
Performance Optimisation
Department of Computing
Imperial College London
Joint work with David Ham (Imperial Maths),
Lawrence Mitchell (Imperial Computing)
And many others...

- Engaging with applications to exploit domain-specific optimisations can be incredibly fruitful
 - Compiling general purpose languages is worthy but usually incremental
- Compiler architecture is all about designing intermediate representations – that make hard things look easy
 - Tools to deliver domain-specific optimisations often have domain-specific representations
 - Premature lowering is the constant enemy (appropriate lowering is great)
- Along the way, we learn something about building better general-purpose compilers and programming abstractions
 - Drill vertically, expand horizontally

Turing tax evasion

Turing tax

- Alan Turing (1912-1954) realised we could use digital technology to implement any computable function
- Proposed “universal” computing device –
 - a single device which, with the right program, can implement any computable function without further configuration

	Perf. (fps)	Area (mm ²)	Enrgy/frame (mJ)
Intel (720x480 SD)	30	122	742
Intel (1280x720 HD)	11	122	2023
ASIC	30	8	4

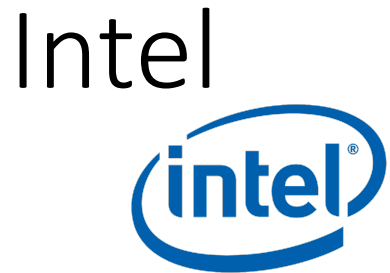
Intel's highly optimized, 2.8GHz Pentium 4 implementation of a 480p H.264 encoder versus a 720p HD ASIC. The second row presents Intel's SD data scaled to HD H.264. ASIC numbers have been scaled from 180nm to 90nm. (Hameed et al, ISCA 2010)

- The “**Turing Tax**” is a term for the overhead (performance, cost, or energy) of universality in this sense
- That is, the performance difference between a special-purpose device and a general-purpose one
- One of the fundamental questions of computer architecture is to how to reduce the Turing Tax

- FPGAs
 - Fetch-execute
 - But since it doesn't involve communication, it's not the important thing
- Registers
 - Because if we know the program's dataflow, we can use wires and latches to pass data from functional unit to functional unit
- Cache
 - If we know exactly when the reuse will occur, we can program movement to and from local fast memory explicitly
- Vectorisation
- Message aggregation
- Inspector-executor
- Dependence, points-to, parallelization

Exhibition





A long fruitful relation with Intel

- Intel has just funded our second IPCC on a ML based tool for Fast Simulation in GeantV
- We presented our first results on GAN application to calorimeter simulation
- Meeting with several ML specialists
 - Study scaling on KNL cluster (MARCONI system @CINECA)
 - Port our code to NEON framework

“New IPCC at CERN Pushes Code Modernization”

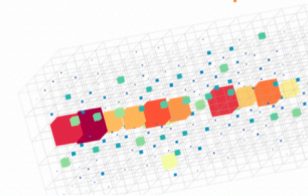
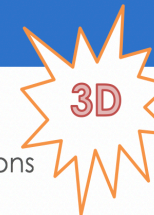
“CERN Modernizes Code with IPCC”

InsideHPC
GOParallel



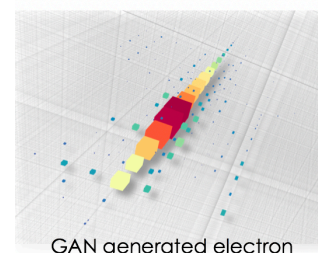
3D GAN for particle detectors

- Treat energy deposits in cells as 3D image
 - Generator and Discriminator based on 3D convolutions
- Explored several “tips&tricks”
 - No batch normalisation in the last step, LeakyRelu, no hidden dense layers 😊, Adam optimiser 😊
 - Batch training

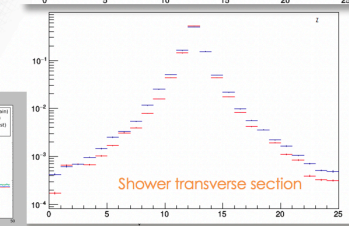
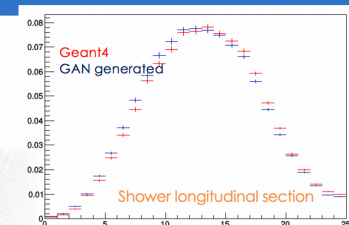
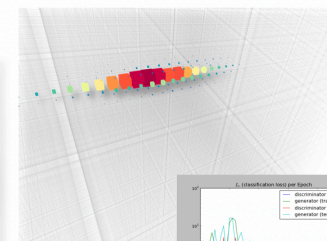


Some generated images

- First results look very promising!
- Qualitative results show no collapse problem



GAN generated electron





- Leader in energy efficient processors, extending to the HPC environment
 - Fujitsu Post-K to replace (#5 top500) computer-K at RIKEN supercomputer center in 2020
 - Mont-Blanc phase3 Exascale project uses Cavium's ThunderX2
- Longer vector extensions to effectively provide massive computational throughput for HPC
- Scalable Vector Extension (SVE: 128 to 2,048 bit) with vector-length-agnostic programming model
 - Orthogonal to existing 128-bit Neon SIMD
- New Cortex-A systems targeted to ML workloads

www.montblanc-project.eu

MONT-BLANC

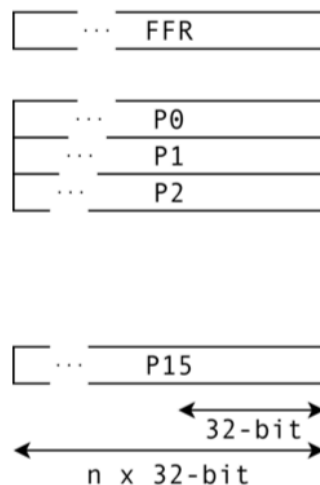
EUROPEAN APPROACH TOWARDS ENERGY EFFICIENT HIGH PERFORMANCE

ARM SVE

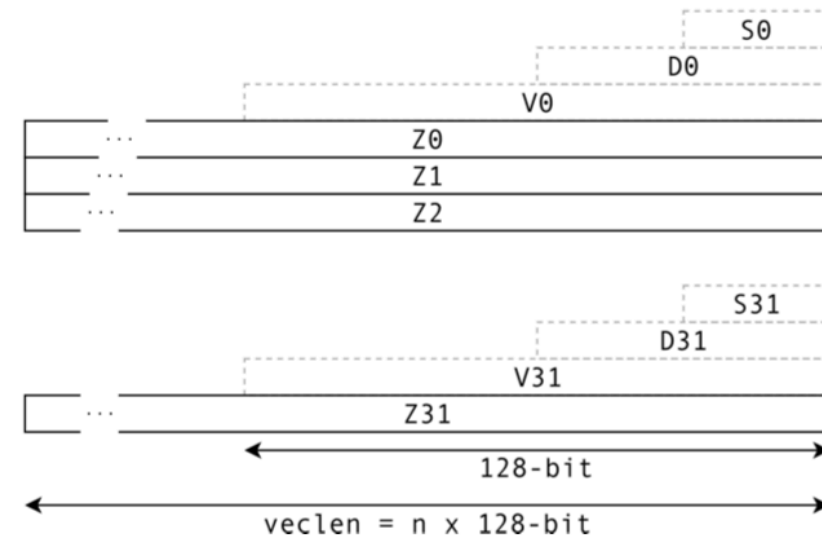


A “universal” vector extension

Predicate registers



Data Registers (SVE/NEON/VFP)

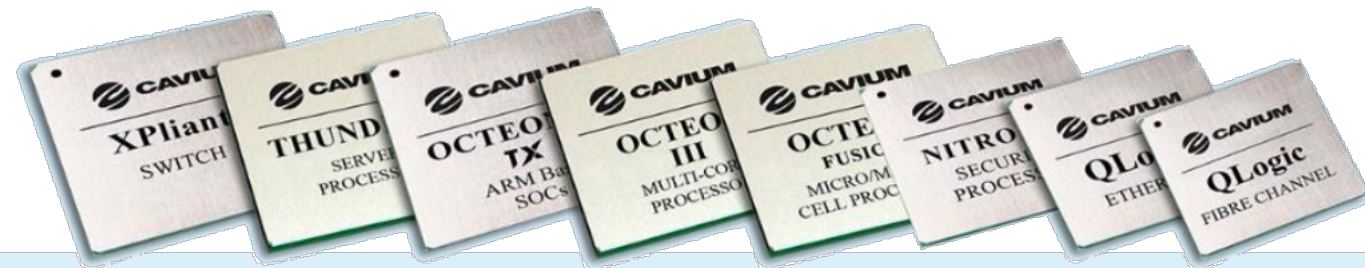


- Supports vector-length agnostic (VLA) programming that adapts to available vector length
- Introduces novel features to improve compiler vectorization (per-lane predication, gather-load/scatter-load,..)
- A unique challenge for software!
- Very interesting meeting with ARM senior engineers
 - Benchmark compiler vectorisation on GeantV code
 - Test our 3dGAN network training on the new architectures

Cavium



- A well know processors manufacturer
- We met Avinash Sodani, a former senior Intel engineer, head of the KNL design team, now at Cavium
- Co-development of a custom accelerator targeted to GeantV simulation workload
 - Hardware acceleration for Machine Learning
 - Further meetings to discuss GeantV code design (data layout and processing flow) and the machine learning approach



HPC Redefined – Join the AI Revolution



- IBM to deliver new hardware to CERN (openlab)
- GeantV is part of the physics WG of the OpenPOWER foundation (resp. Marilena)
- Proposal to fund a doctoral student to work on GeantV optimisation for Power architectures



IBM PowerAI
Simplified Development and Deployment
Get to work faster, better for HPC or AI workloads

- Fast, Performant AI deployment: single command install
- Complete developer workflow tools for DL or HPC
- Rapid cluster deployment for fast time to value

Fast, Performant AI deployment

Deploy open source deep learning frameworks rapidly & experiment with optimized frameworks and tools

- Deploy Caffe, Torch, Theano

Complete developer workflow tools for DL or HPC

Leverage a complete set of AI development tools that addresses every step of your AI pipeline

- Simplified AI development with

Rapid cluster deployment for fast time to value

Deliver rapid deployment for fast time to value with a comprehensive cluster stack

- Rapidly deploy a cluster in minutes with Spectrum Cluster

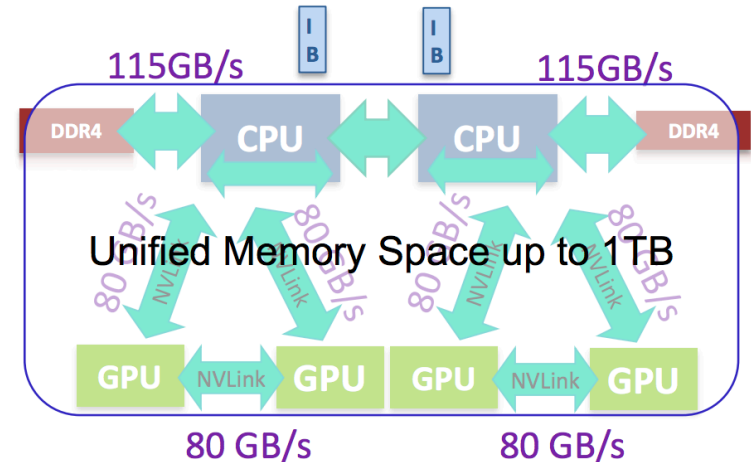
Better System Design for AI: Flat and Fat

S822LC for HPC is engineered both **flat** and **fat**

- Data flows freely across system
- Nearly as broad from CPU: GPU as System Memory: CPU
- Big pipes between GPUs on the same socket

Designed for AI, Deep Learning

- GPU:GPU communication
- Unified memory + CPU:GPU NVLink carry data large deep learning models
- Superior Tensorflow, Caffe performance






THE CONVERGENCE OF HPC & DEEP LEARNING

TESLA P100

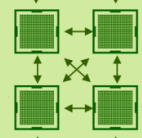
- NVIDIA is still leading the way on DL hw
- Demand for GPU @CERN is high!

Pascal Architecture



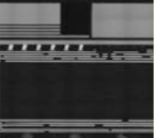
Highest Compute Performance

NVLink



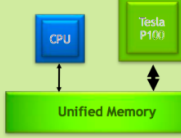
GPU Interconnect for Maximum Scalability

HBM2 Stacked Memory



Unifying Compute & Memory in Single Package

Page Migration Engine



Simple Parallel Programming with 512 TB of Virtual Memory

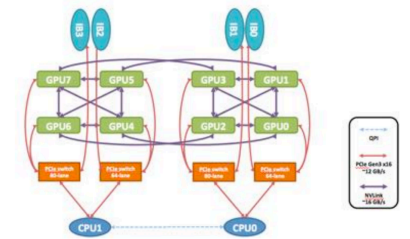


NVIDIA DGX-1
AI supercomputer-in-a-box

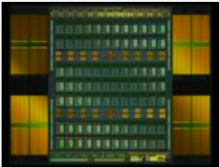


Meeting to discuss collaboration on our new ML project and GeantV deployment on GPU

42.5 TF FP64, 85 TF FP32, 170 TF FP16 perf
 8x Tesla P100 16GB
 NVLink Hybrid Cube Mesh
 Optimized Deep Learning Software Stack
 Dual Xeon
 7 TB SSD Deep Learning Cache
 Dual 10GbE, Quad IB 100Gb
 3RU - 3200W

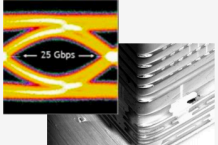


Volta Architecture



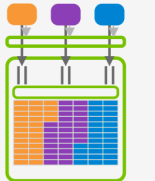
Most Productive GPU

Improved NVLink & HBM2



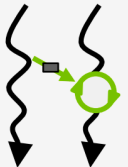
Efficient Bandwidth

Volta MPS



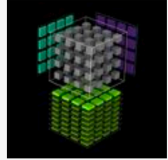
Inference Utilization

Improved SIMT Model



New Algorithms

Tensor Core

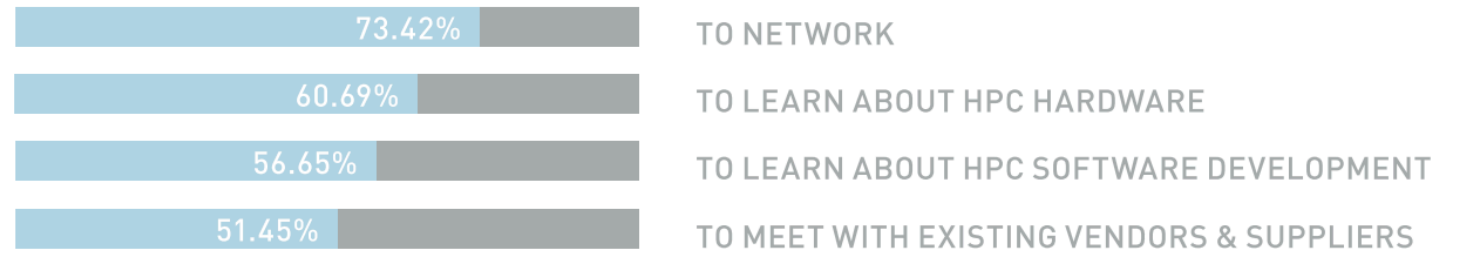


120 Programmable TFLOPS Deep Learning

TESLA V100

Conclusions

TOP FOUR REASONS FOR ATTENDING



- ISC is a huge event bringing academia and industry together
 - High level technical program including innovative applications in many different fields
- Networking is one of the main reasons for attending ISC
- This edition was particularly useful for us
 - Recognition of the work done
 - Collaboration with experts
 - Building collaboration

Thanks!