

COMPUTE & ACCELERATOR FORUM

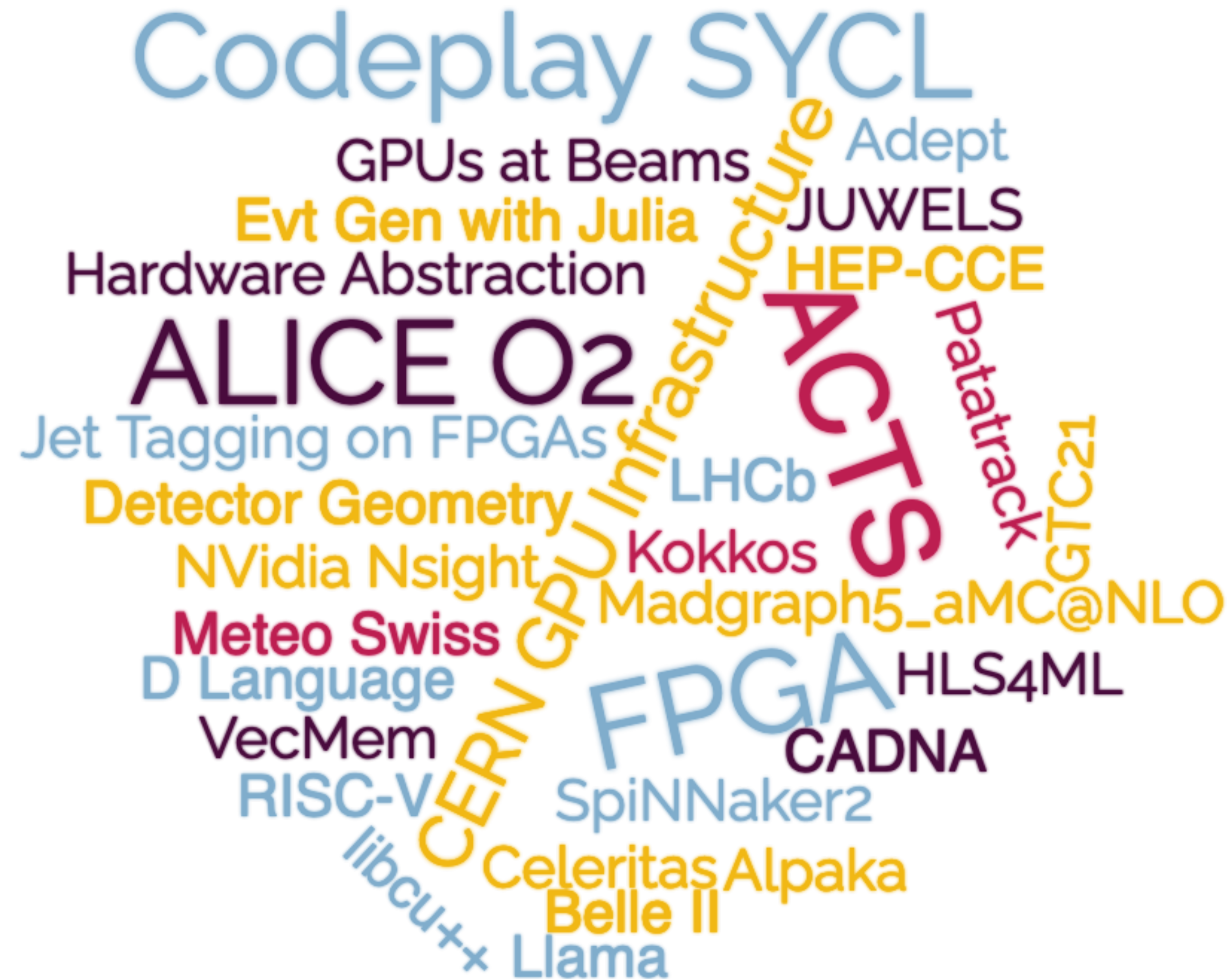
STEFAN ROISER, CERN

GDB, 10 JAN 2024

WHY, WHAT, WHEN, WHO

- ▶ Initially devised as a series of seminar style presentations for in-depth presentations and discussions on latest advances in compute acceleration and heterogeneous computing
- ▶ Second Wednesday of the month, 16:30, max 1 ½ hours
 - ▶ Usually right after GDB in the same room (but different zoom)
 - ▶ Presentations also recorded and linked at the event pages
- ▶ Agenda category @ <https://indico.cern.ch/category/12741/>
- ▶ Co-organised by Maria Girone, Graeme Stewart, SR (all CERN), Ben Morgan (Univ Warwick), Michael Bussmann (Helmholtz)
- ▶ Contact us at compute-accelerator-forum-organizers@cern.ch

TOPICS SO FAR



In order to reflect on the very diverse set of topics being touched on the series title changes as of this year to “Compute & Accelerator Forum”

FLASHING A FEW PRESENTATIONS WITH POSSIBLE CONNECTIONS / OVERLAP TO THE GDB

Please see the indico links for details and video recording

DATA PROCESSING INFRASTRUCTURES

JUWELS Booster Overview

Node Configuration

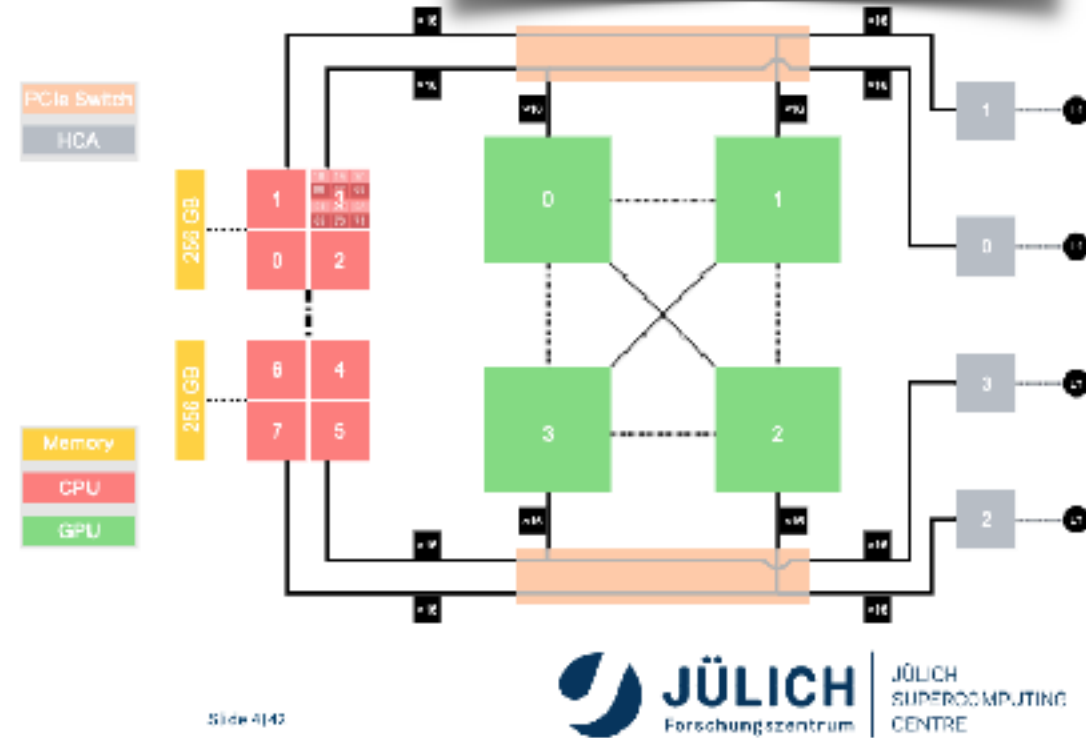
Arch Atos Bull Sequana XH2000

CPU 2 × AMD EPYC 7402:
2_{Socket} × 24_{Core} × 2_{SMT},
2 × 256 GB DDR4-3200 RAM;
NPS-4

GPU 4 × NVIDIA A100 40 GB, NVLink3
73 PFLOP/s, 1.16 EFLOP/s_{FP16TC}
18.7 EOP/s_{BitTC}

HCA 4 × Mellanox HDR200 (200 Gbit/s)
InfiniBand ConnectX 6

etc 2 × PCIe Gen 4 switch



Top500 Nov-2020:

- #1 Europe
- #7 World
- #1* Green500

Member of the Leibniz Association

9 June 2021

Slide 4/42

JÜLICH
Forschungszentrum
JÜLICH SUPERCOMPUTING
CENTRE

A Herten (Jülich), 9 Jun 2021, <https://indico.cern.ch/event/975011/>

latest updates on of the CERN GPU infrastructure

What's New

Action items from the last update in this forum (Oct 2021)

<https://indico.cern.ch/event/975015/>

Upgrade Nvidia drivers (470.82.x) for CUDA 11.4 (done)

Support for GPU profiling in vGPU nodes (done)

Reminder: vGPU is not physical partitioning but time sharing up to 4x

Profiling possible with vGPUs - with new drivers (done)

General availability of vGPU setup (instabilities detected)

R Rocha (CERN), CERN GPU Infrastructure Updates
8 June 2022, <https://indico.cern.ch/event/1073643/>

13 Oct 2021, <https://indico.cern.ch/event/975015/>

20 Oct 2020, <https://indico.cern.ch/event/950196/>

Recent development and status
of beam physics codes for
heterogenous platform

Code structure and technology choices

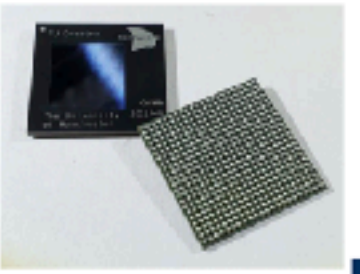
- Code organized in Python packages.
- User interactions via Python scripts.
- Data structures described, allocated and fully exposed (r/w) in Python including GPU (avoid unnecessary copy).
- Performance critical code written in C with using automatically generated C-API from Python.
- Code compiled at run time and on-demand:
 - Allows problem specific optimizations essential on GPU
 - Reduce writing, testing, executing cycle
- Dependencies:
 - numpy: allocate and exchange memory
 - cffi (and a C compiler): generates binary Python modules that can be imported at run time and prepare arguments
 - cupy (and a cuda driver): implements rich numpy-like array on device and compiles cuda kernels
 - pyopencl (and OpenCL drivers): wraps OpenCL API and implement a basic numpy-like array

R De Maria (CERN),
8 Dec 2021, <https://indico.cern.ch/event/975017/>

HARDWARE PLATFORMS

Bio-Inspired Processing from Edge to Cloud: SpiNNaker 2 and Beyond

SpiNNaker2



Hybrid design for deep neural networks, spiking neural networks and symbolic AI

$$T_{ij} = a_i \left(\sum_k T_{jk} - T_{ij} \right) - \theta_j$$

Outperforming Intel, Nvidia, Google on real time AI

Brain-inspired **sparsity**, i.e. highly-parallel operations algorithms on streaming data

Largest real-time AI platform worldwide, **10¹⁴ parameters**
3 PFLOPS CPU
0.4 ExaOPS in AI accelerator

Physical: 10⁷ processors, 70.000 chips, 14 racks, 100.000.000.000.000 transistors

SpiNNaker2 Chip:

- 153 ARM cores
- >100 person design team
- 22FDX Global Foundries
- Developed in EU flagship Human Brain Project
- Development cost: >20Mio
- Deployment cost: >10Mio

Brain-like capabilities for autonomous systems

Drug Screening, fast medical data analysis

TECHNISCHE UNIVERSITÄT DRESDEN

C Mayr (TU Dresden), 14 Sep 2022, <https://indico.cern.ch/event/1073646/>

What is RISC-V? (on one slide)

- ▶ Open Standard Instruction Set Architecture (ISA)
 - ▶ Specifications are open source, no royalty fees
 - ▶ RISC-V cores can be open or proprietary
- ▶ Started at the University of California, Berkley, in 2010
- ▶ Since 2020 published by RISC-V International located in Switzerland
- ▶ Modular design: base ISA with very few (integer) instructions
 - ▶ Many standard extensions and possibility for custom instructions



Code using SSE instructions (omitting setup code)

```

vectorized:
    movups  xmm2, xmmword ptr [rsi + r8]    # load x
    mulps   xmm2, xmm1                    # multiply with a (in xmm1)
    movups  xmm3, xmmword ptr [rdx + r8]    # load y
    addps   xmm3, xmm2                    # ax + y
    movups  xmmword ptr [rdx + r8], xmm3    # store y
    add     r8, 16                         # compute next offset (+ 4 elements)
    cmp     rdi, r8                        # compare to final offset to process
    jne     vectorized                    # with vectorized loop (pre-computed)
    cmp     rcx, rax                        # check if elements remain
    je      done                           # otherwise done

scalar:
    movss   xmm1, dword ptr [rsi + 4*rcx]  # load x
    mulss   xmm1, xmm0                    # multiply with a
    addss   xmm1, dword ptr [rdx + 4*rcx]  # load and add y
    movss   dword ptr [rdx + 4*rcx], xmm1  # store y
    inc     rcx
    cmp     rax, rcx
    jne     scalar
done:     ret
    
```


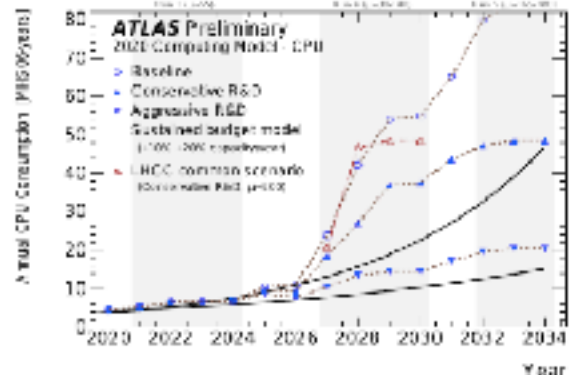
J Hahnfeld (CERN), 13 Sep 2023, <https://indico.cern.ch/event/1264300/>

GPU DEVELOPMENTS IN THE CONTEXT OF LHC EXPERIMENTS

Overview

ACTS - A Common Tracking Software

- Experiment independent tracking toolkit, written in modern C++
- Efficient implementation of common tasks (e.g seed finding, track fitting ...)
- Thread-safety for further parallelization
- Uses a *tracking geometry* with a simplified material description to save compute resources.


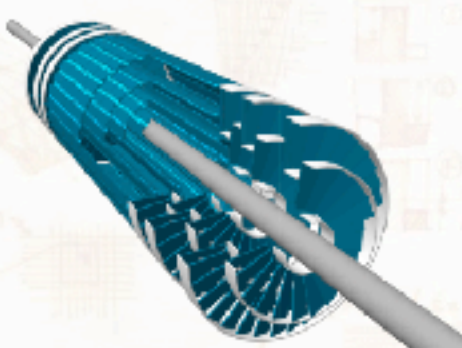
Tracking on GPU

- To tackle the combinatorics in a high luminosity environment, investigate tracking on GPU.
- For this to succeed, define a suitable EDM.
- develop a toolchain that supports e.g. CUDA kernels
- and provide GPU friendly implementations of the geometry and magnetic field.

Image: Annual CPU Consumption [6]
The ACTS Project
Compute Accelerator Forum 1 / 17

what is... Patatrack?

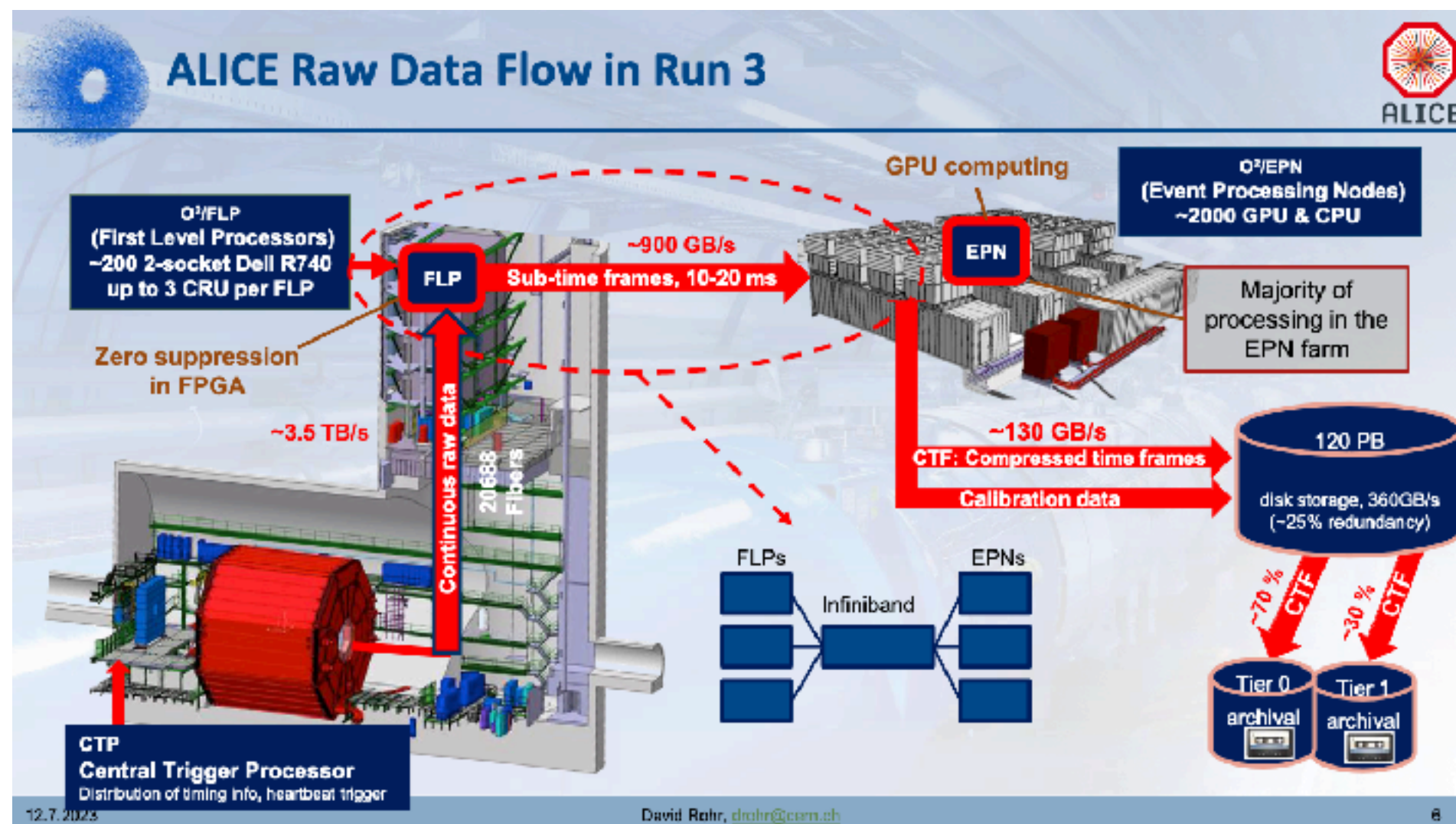
- loosely defined group of people working on R&D and the use of accelerators in CMS
 - based at CERN, Fermilab, CASUS, Aachen, ...
 - overlap with HEP-CCE
- incubator of ideas and R&D efforts
 - algorithms and implementations
 - performance portability
 - development of standalone demonstrators
 - solutions and adoption in CMS and CMSSW
- promote the use of GPUs and accelerators in CMS
 - Patatrack Hackathons, documentation and knowledge transfer
 - collaboration with CASUS
- CMS pixel-only track and vertex reconstruction
 - developed for use at HLT
 - targetting both Phase-1 (ready) and Phase-2 (under development)

March 9th, 2022 Patatrack! - Compute Accelerator Forum 3 / 20

J Niermann (Univ Goettingen), 14 Dec 2022, <https://indico.cern.ch/event/1160623/>

A Bocci (CERN), 9 Mar 2022, <https://indico.cern.ch/event/1073640/>



D Rohr (CERN), 12 July 2023, <https://indico.cern.ch/event/1264298/>

MadEvent with vectorized C++ for $gg \rightarrow t\bar{t}gg$ (on a single CPU core)

$gg \rightarrow t\bar{t}gg$	MEs precision	ACAT2022		madevent		standalone
		$t_{TOT} = t_{Mad} + t_{MEs}$ [sec]	N_{events}/t_{TOT} [events/sec]	N_{events}/t_{MEs} [MEs/sec]		
Fortran (scalar)	double	37.3 = 1.7 + 35.6	2.20E3 (=1.0)	2.30E3 (=1.0)	—	
C++/none (scalar)	double	37.8 = 1.7 + 36.0	2.17E3 (x1.0)	2.28E3 (x1.0)	2.37E3	
C++/sse4 (128-bit)	double	19.4 = 1.7 + 17.8	4.22E3 (x1.9)	4.62E3 (x2.0)	4.75E3	
C++/avx2 (256-bit)	double	9.5 = 1.7 + 7.8	8.63E3 (x3.9)	1.05E4 (x4.6)	1.09E4	
C++/512y (256-bit)	double	8.9 = 1.8 + 7.1	9.29E3 (x4.2)	1.16E4 (x5.0)	1.20E4	
C++/512z (512-bit)	double	6.1 = 1.8 + 4.3	1.35E4 (x6.1)	1.91E4 (x8.3)	2.06E4	
C++/none (scalar)	float	36.6 = 1.8 + 34.9	2.24E3 (x1.0)	2.35E3 (x1.0)	2.45E3	
C++/sse4 (128-bit)	float	10.6 = 1.7 + 8.9	7.76E3 (x3.6)	9.28E3 (x4.1)	9.21E3	
C++/avx2 (256-bit)	float	5.7 = 1.8 + 3.9	1.44E4 (x6.6)	2.09E4 (x9.1)	2.13E4	
C++/512y (256-bit)	float	5.3 = 1.8 + 3.6	1.54E4 (x7.0)	2.30E4 (x10.0)	2.43E4	
C++/512z (512-bit)	float	3.9 = 1.8 + 2.1	2.10E4 (x9.6)	3.92E4 (x17.1)	3.77E4	

512y = AVX512, ymm registers
512z = AVX512, zmm registers
The latter is only better on nodes with 2 FMA units (here an Intel Gold 6148)

Scalar: FLOAT DOUBLE
SSE4: FLOAT FLOAT FLOAT DOUBLE DOUBLE
AVX2: FLOAT FLOAT FLOAT DOUBLE DOUBLE DOUBLE
AVX512: FLOAT FLOAT FLOAT DOUBLE DOUBLE DOUBLE DOUBLE DOUBLE DOUBLE DOUBLE DOUBLE DOUBLE DOUBLE DOUBLE

ME speedup ~ x8 (double) and x16 (float) over scalar Fortran
Our ME engine reaches the maximum theoretical SIMD speedup!
Overall speedup so far~ x6 (double) and x10 (float) over scalar Fortran (Amdahl's law)

A. Valassi - Data parallelism in Madgraph5_aMC@NLO: vectorization and GPUs
Compute Accelerator Forum - CERN, 8 February 2023 31

A Valassi (CERN), 8 Feb 2023, <https://indico.cern.ch/event/1207838/>

MORE HEP PROJECTS

HEP-CCE

HEP-CCE

Portable Parallelization Strategies (PPS)

	CUDA	Kokkos	SYCL	HIP	OpenMP	alpaka	std::par
NVIDIA GPU	Green	Green	Green	hipcc	nvcc LLVM, Dey GCC, XL	Green	nvcc++
AMD GPU	Red	Green	openSYCL intelSYCL	hipcc	nvcc LLVM Gcc	Green	Red
Intel GPU	Red	Green	oneAPI intelSYCL	oneAPI-SPV early prototype	Intel OneAPI oneapi	prototype	oneapi::dpl
x86 CPU	Red	Green	oneAPI intelSYCL openSYCL	via HIP-CPU Runtime	nvcc LLVM, GCC, GCC, XL	Green	Green
FPGA	Red	Green	Green	via XRT Runtime	prototype oneAPI (OpenACC, Intel, etc.)	prototype via SYCL	Red

circa today



Office of
Science



C Legget (LBNL), 14 June 2023, <https://indico.cern.ch/event/1264297/>

HEPIX Benchmarking

SCHEDULED 14 FEB 2024

D Giordano (CERN), 14 Feb 2024, <https://indico.cern.ch/event/1329686/>

C&A FORUM AND GDB?

- ▶ Last year organised a combined event "[CA Forum & HSF Reco WG](#)" → also w/ GDB?
- ▶ Would you be interested in more topics with a GDB connection?
 - ▶ Individual sites adoption of GPUs, FPGAs
 - ▶ WLCG strategy towards hardware acceleration
 - ▶ Scheduling of hardware accelerated applications
 - ▶ ...
- ▶ Please contact any of us (*) or at compute-accelerator-forum-organizers@cern.ch

(*) Stefan Roiser, Graeme Stewart, Maria Girone, Ben Morgan, Michael Bussmann

WHAT DID WE LEARN IN THE PAST 3 YEARS

- ▶ Many in-depth presentations on interesting tools, languages, etc. GPU abstraction layers, CADNA,
- ▶ The computing world is getting ever more diverse with more architectures and chip types coming up
- ▶ No clear winner to heard the cats: oneAPI/SYCL, Alpaka/Llama, Kokkos, pragmas, ...
 - ▶ Changing the code conceptually is the major effort, how to abstract across platforms is less of an issue
 - ▶ C++ standard is also / will provide ways to cope
- ▶ New languages to cope with the new hardware landscape: D, Julia, ...

THANK YOU!!

Info on upcoming meetings: compute-accelerator-forum-announce@cern.ch

10 Jan 2024, J Pivarski (Princeton), "Garbage Collectors: Java, Python, Julia",
<https://indico.cern.ch/event/1329685/>

14 Feb 2024, D. Giordano, "HEPIX Benchmarking",
<https://indico.cern.ch/event/1329686/>

BACKUP

TOPICS POSSIBLY OVERLAPPING / INTERESTING FOR THE GDB

- ▶ Data processing infrastructures, e.g. JUWELS, GPUs @ CERN, GPUs at CERN Beams
- ▶ Hardware platforms: SpiNNaker2, RISC-V,
- ▶ Status of GPU developments at the LHC: ACTS, Patatrack, Madgraph, Adept, Celeritas, ALICE O2, LHCb Allen, Belle II
- ▶ More HEP projects: HEP-CCE, HEPIX Benchmarking,