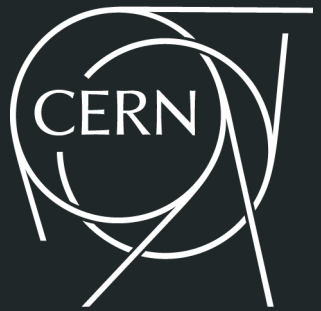
HSF WLCG Workshop Report

Graeme Stewart, for the workshop organisers



2020-06-10

WLCG-HSF Workshop

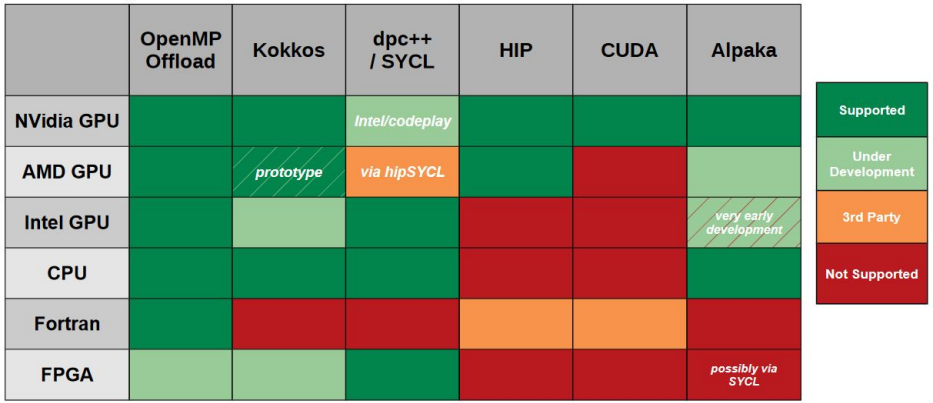
- Planned face-to-face workshop in Lund from 11-15 May was unable to take place
- Decided not to lose the slot, but organise a virtual workshop, taking place over 3 days, 2 hours a day
  - 16-18h CERN time - fairly good for Europe and US, but impossible for Asia and Australia
  - Strong feeling that virtual meetings demand higher concentration and cannot run for as long as normal workshops
- New Architectures, Portability, and Sustainability theme
  - Would have been one of the plenary days in Lund
    - Monday: Application Software
    - Tuesday: Processing Frameworks
    - Wednesday: Validation and Accounting

A Success!

- 221 people registered
- Slides were posted in advance for review
  - We had aimed for a week, but in practice it was more like a day for most talks
- Workshop notebook was available in advance
  - A sort of Live Notes++
- Attendance peaked at 175 Monday, 150 Tuesday, 110 Wednesday
  - We had a clash with LHCOPN/LHCONE meeting on Wednesday :-(
- As this was the first event of this type we hosted we put effort into
  - Post-workshop survey
  - Identifying outcomes and follow-ups
  - Learning how to run these kind of events most effectively - do more virtual workshops have a role to play in the future?

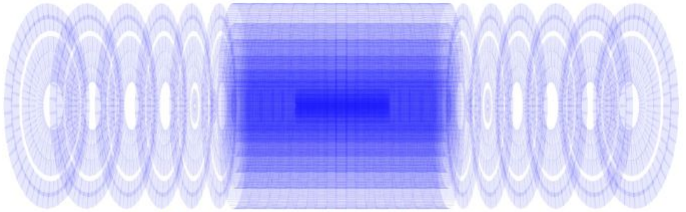
Monday - Application Software I

- Code Portability
  - Increasingly large number of possible non-CPU devices available
  - Clear that the community cannot support N codes for N platforms
  - Industry knows this too, hence proliferation of toolkits and projects
- How to assess the best?
  - This is an orthogonal question to redesigning code for at least one parallel architecture
- DOE HEP-CCE Project
  - *Portable Parallelization Strategies*
  - Assess metrics for toolkits on real HEP examples:
    - Patatrack (CMS),
    - FastCaloSim (ATLAS)
    - WireCell (Neutrino)
  - Will produce recommendations taking into account the nature of HEP workflows



All this changes rapidly

Monday - Application Software II



GPU ray tracing of Track ML geometry using VecGeom

- Heterogeneous Architectures and Detector Simulation
  - Simulation a very significant part of HEP computing; common engine in Geant4
    - Increased luminosity and trigger rates only increase the pressure
  - GeantV (arXiv:2005.00949) taught us valuable lessons about how to optimise
    - Optimisation from SIMD far less than hoped for - data preparation costs to use vector registers is high (see Andrei's talk, https://indico.cern.ch/event/818702/)
    - Modernising and reducing code size can bring up to x2
      - We think mainly from more optimal use of data and instruction caches
    - Dedicated libraries to do pieces of HEP specific code can be reused (VecGeom)
      - These need to develop and adapt to these new architectural challenges
    - Ideas for the future: ray tracing on GPUs?
      - The HSF Detector Simulation group will have a set of lightning talks in the coming weeks (first one this Wednesday)

Monday - Application Software II

Amplitude analyses

- Large amounts of data
- Complex models
- … which depend on optimisable parameters
- Optimise by minimising neg. log. likelihood (NLL)
- Need tools which allow
  - Convenient description of models
  - Efficient computations

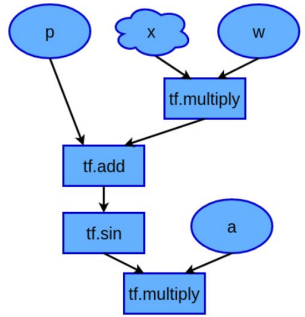
and don't require deep low-level hardware knowledge.

Machine learning

- Large amounts of data
- Complex models
- … which depend on optimisable parameters
- Optimise by minimising cost function
- Need tools which allow
  - Convenient description of models
  - Efficient computations

and don't require deep low-level hardware knowledge.

- TensorFlow as a Compute Engine
  - Using highly optimised libraries with built in GPU support
    - Many of these are developed to support machine learning, but are a good fit for some problems in HEP, e.g., amplitude analyses
  - TensorFlow is a declarative programming environment - describe *what* you want to do, not *how* to do it
    - Graph with nodes as operations, edges as data flow
  - Need to layer HEP concepts on top: Dalitz plots, four-body phase space, etc.
  - Quite a few HEP projects: TensorFlowAnalysis, zfit, pyhf, VegasFlow, PDFFlow
  - Engine is designed for different purpose to ours, impedance matching can be awkward
    - Then major library upgraded can be a real perturbation… TensorFlow 1 to 2 is a significant change




```
f = a*tf.sin(w*x + p)
```

Tuesday - Application Frameworks

- Heterogeneous Experimental Frameworks
  - Goal is *optimal use* of heterogeneous resources
    - Easier on owned resources (HLT) as opposed to HPCs or other sites
  - Separate process spaces (ALICE-FAIR approach in O2 - message passing)
    - Great code separation, dynamically balance CPU resource use at process level
  - Accelerator only approach
    - Ideal for R&D projects; optimal performance for the target devices
    - Can leave CPUs idle when they could do useful work (may be balanced by other tasks)
  - Hybrid approach
    - Asynchronous execution, so most complex for framework
      - But maybe the biggest prize
    - Smart underlying schedulers (TBB, HPX) help maximise CPU usage
    - CMSSW is a good example of implementing this
      - Can switch between CPU/GPU version according to resource availablity
  - In all cases there are hurdles for the experiment developers and some steep learning curves (cf. portable parallelisation strategies talk)

Tuesday - Workload Management

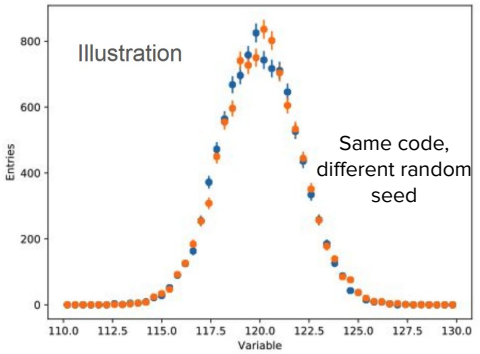
Testing Queues

Queue	Resource type	CE/BS	N GPUs	GPU Models	GPU memory	Host Memory	Outbound connectivity
ANALY_BNL_GPU_ARC	HPC/grid	ARC-CE/slurm	12 (guaranteed for ATLAS), 200 available	P100	16GB	256GB	No
ANALY_OU_OSCER_GPU_TEST	HPC/grid	HTCondor-CE/slurm	80	K40	12GB	12GB	yes
ANALY_QMUL_GPU_TEST	grid	CREAM-CE/slurm	2,2*2	K40,K80	12GB	12GB	yes
ANALY_MANC_GPU_TEST	grid	ARC-CE/HTCondor	6, 4	VT100, K40	12GB	12 GB	yes
ANALY_MWT2_GPU	grid	HTCondor-CE/HTCondor	8	1080Ti	11GB	24 GB	yes
ANALY_INFN-T1_GPU	grid	HTCondor-CE/HTCondor	2	K40	12GB	12GB	yes
ANALY_SLAC_GPU	grid	ARC-CE & Slurm	326 (opportunistic)	1080Ti, 2080Ti & v100	11 & 32GB	192GB	yes

Current ‘generic’ grid resources are tiny overall, but growing

- Challenge here is to seamlessly incorporate heterogeneous resources
  - And outside of HLT farms this can be *really heterogeneous* - many different GPU and CPU combinations
- Workloads
  - Traditional CPU jobs (for now the vast majority)
  - Jobs that *require* a GPU
  - Jobs that could take advantage of a GPU, but can run CPU only
    - HEP might require a lot of these jobs for high efficiency use of our global resource pool
- Issues
  - Don’t auto-discover resources, negotiate with the LRMS
  - Tag resources properly for matchmaking (more complex than CPU cases)
- ATLAS and CMS users can submit payloads that require GPUs via PanDA and CMS Connect
  - Use of containers for the software stack is ubiquitous (sites also using k8s to manage their resources)
- At the moment no large scale production workflows using GPUs
  - Open question as to how much machine learning HEP might use

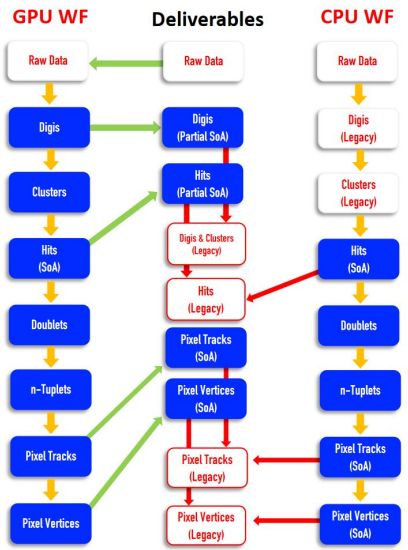
Wednesday - Validation

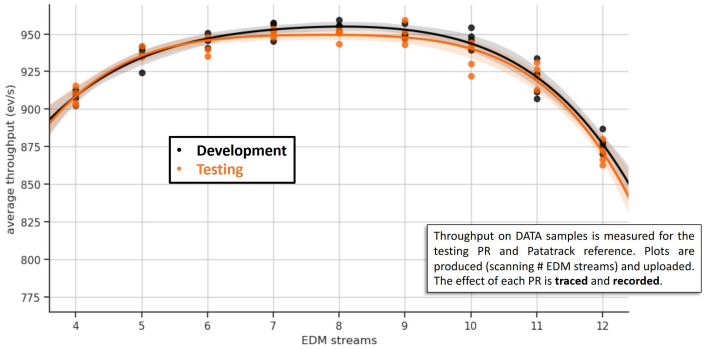


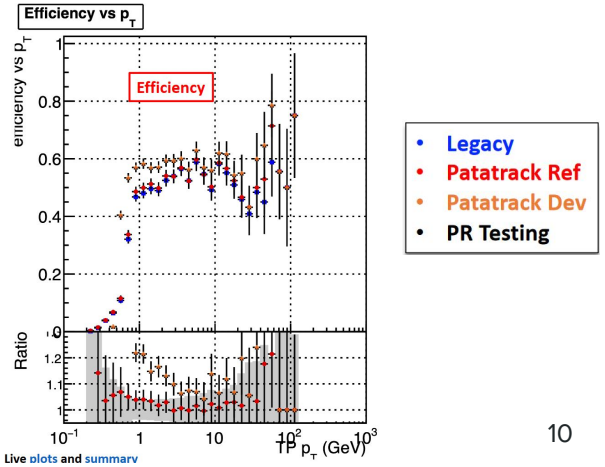
- Process of Physics Validation
  - Running a HEP workflow and checking results against a known output (usually a suite of histograms for jets, electrons, photons, muons, etc.)
    - Using new version of software, using different (new) hardware
  - Technical validations should not change the output *significantly*
    - What that means is often *not as clear as one might imagine* - ulp difference can cause a cut to pass or fail, with significant knock on effects; change in software may perturb PRNG
      - Community has become (too) used to homogeneous x86_64 environment
    - Technical improvements could be foreseen, but each experiment has its own machinery
  - Physics changes always require expert input
    - Irreducible part of the problem - experts usually in short supply

Wednesday - Validation

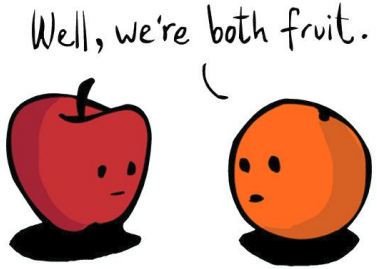
- Patatrack Validation Process
  - CMS project running pixel tracking and calorimeter reconstruction on GPUs
  - Completely new code base on new hardware, so certainly no expectation of the same results
  - Complex workflow - can validate at many intermediate steps
- Trigger validation based on triggers from GitHub PRs
  - Aim for as much automation as possible
  - Measure physics performance and computational performance





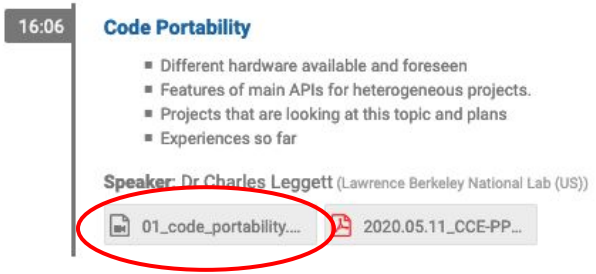


Wednesday - Accounting and Benchmarking



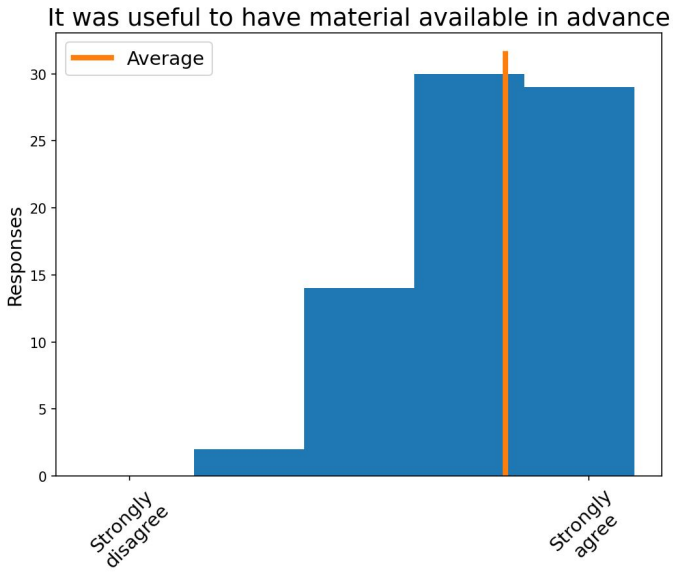
- HEP has valued CPUs using the HEP-SPEC06 benchmark for many years
  - Subset of SPEC® CPU2006, matched HEP application performance well
- Increasing problems with this benchmark
  - Divergence between values and performance of HEP applications
  - No way to take into account heterogeneous resources (GPU v. CPU)
- Now much easier to run HEP workloads encapsulated - *containers!*
  - By definition gives a ‘score’ correlated with real HEP throughput in events/sec.
  - After running HEP-benchmarks, generate a score by geometrically weighting different workloads
- Can develop additional workloads for GPUs (Patatrack, SixTrack)
- Then possible to account for the value of a resource (*usable capacity*) by its total throughput in events/s
  - Unused hardware elements count for nothing
  - Can also discuss *realised potential*, as a way to measure how efficiently we can use a platform

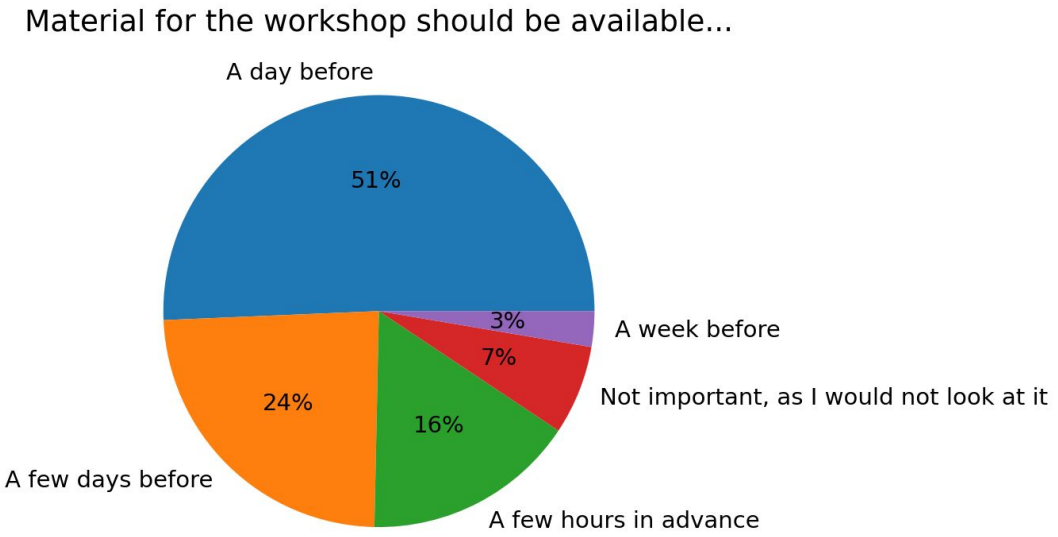
Post-Workshop



- Video
  - The recorded sessions were posted to Indico
  - Experience with recording via Zoom is quite positive
    - Quality is decent and easy to edit the video into separate files per talk
  - Bug in Indico statistics gathering make popularity hard to assess
    - Could also upload to YouTube
- Survey
  - Survey posted in the same week as the workshop
    - Try to get responses while the workshop was fresh in people's minds
  - 75 responses were received (so about ⅓ of registrants)
  - Still (fully) digesting the results, but here are some highlights…

Pre-Workshop Matters I

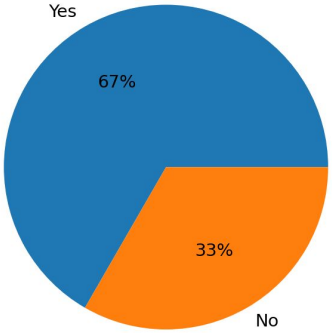




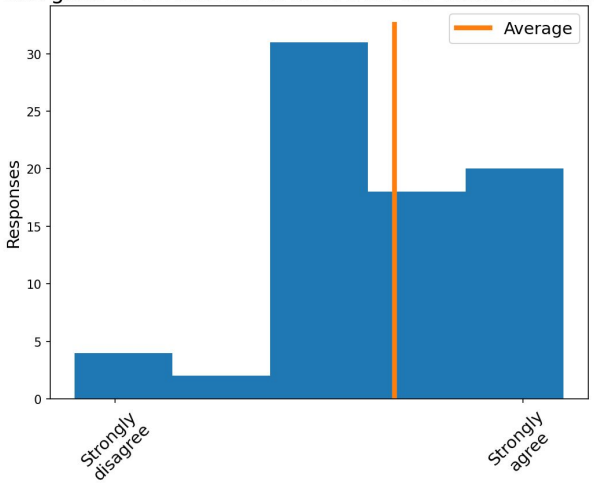
- People like material available in advance
- A day before is enough, a few days would be better

Pre-Workshop Matters II

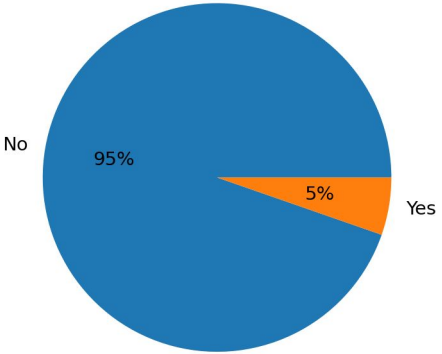
I reviewed the pre-workshop material before the conference



It was good to be able to comment on the material in advance

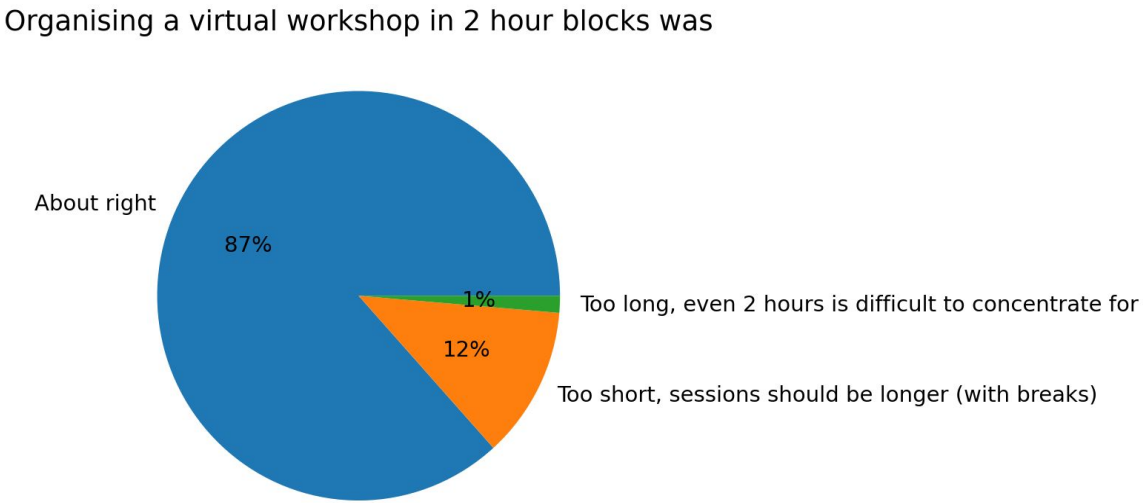


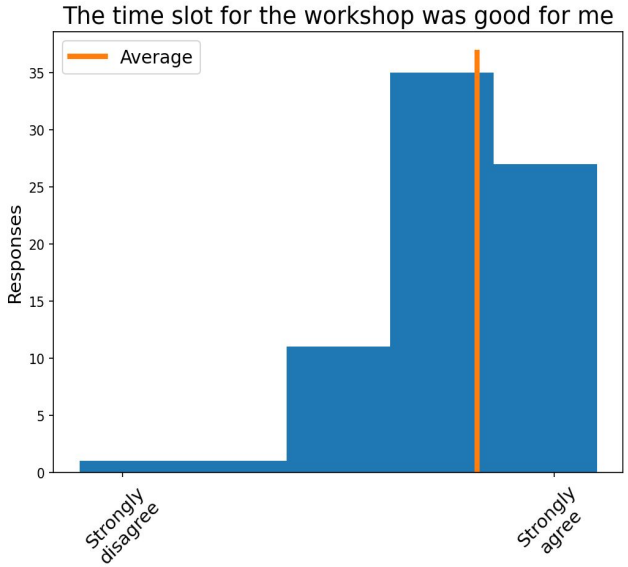
I posted questions in the notebook before the workshop



- Material was looked at a lot
- Few people posted comments, but the ability to was supported

Workshop Schedule

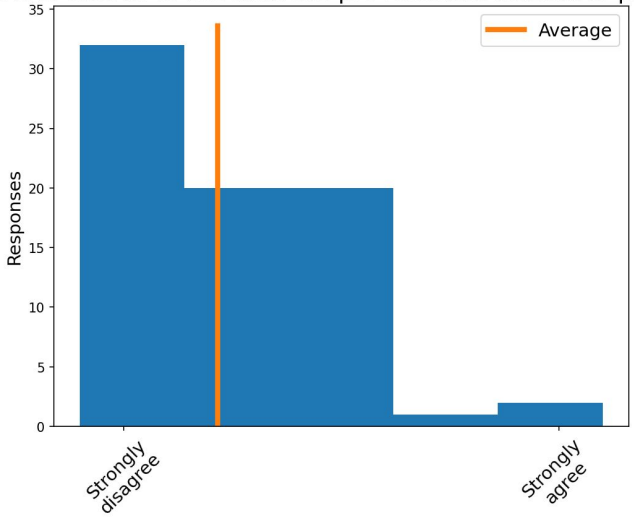




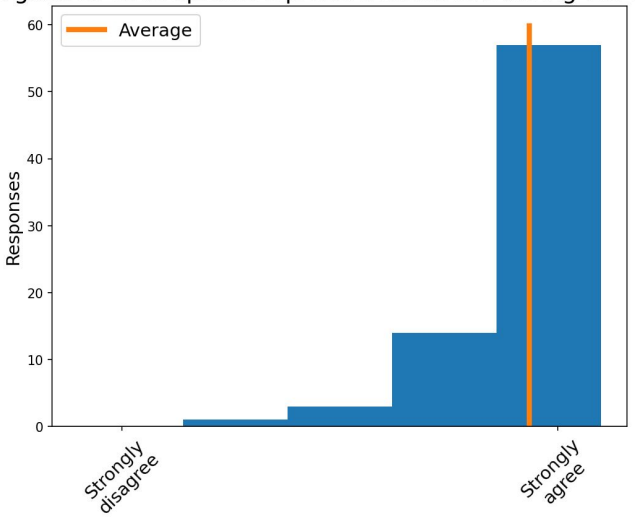
- 2 hour session blocks are a good length
- Timeslot worked for the people who attended... but selection bias!

Running the Sessions I

I would prefer that all of the talks are pre-recorded instead of presented live



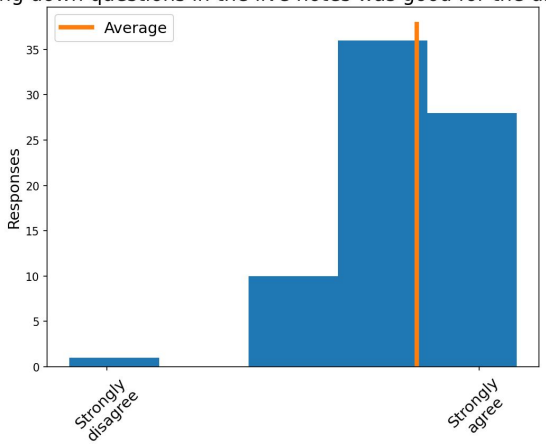
It was good to have speakers present slides live during the workshop



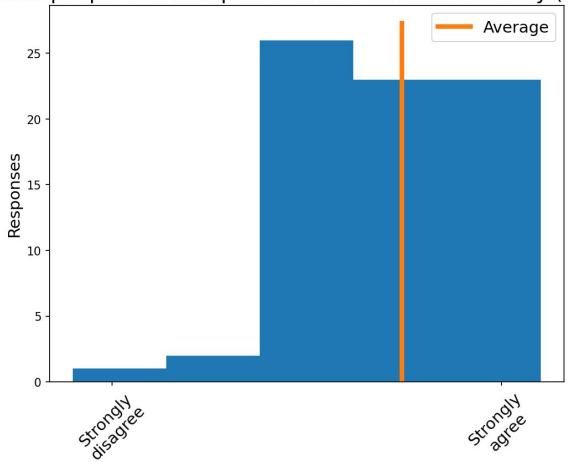
- Strong support for live presentations over pre-recordings

Running the Sessions II

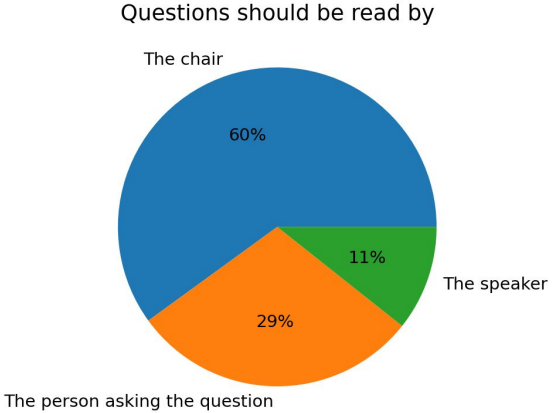
Writing down questions in the live notes was good for the discussion



It is important that people can ask questions and comment directly (not from the notes)



Questions should be read by

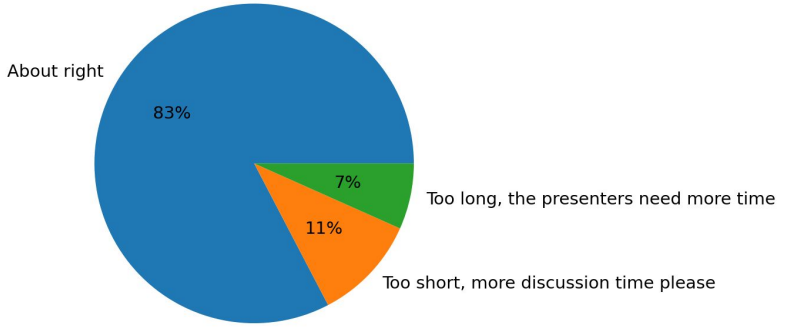


- Notebook is good for questions and discussion
  - It was challenging for the chair to keep track of the notebook during the sessions, but having a notetaker helped a lot

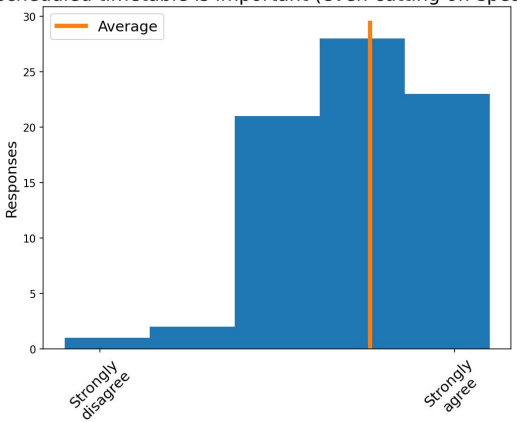
Running the Sessions III

- 50/50 talk/discussion time was good
- People want better timekeeping!
  - We need good channels of communication between the chair and the speaker
  - We did get better at this on Tuesday and Wednesday
    - Dry-run and debugging 30 minutes before the sessions started

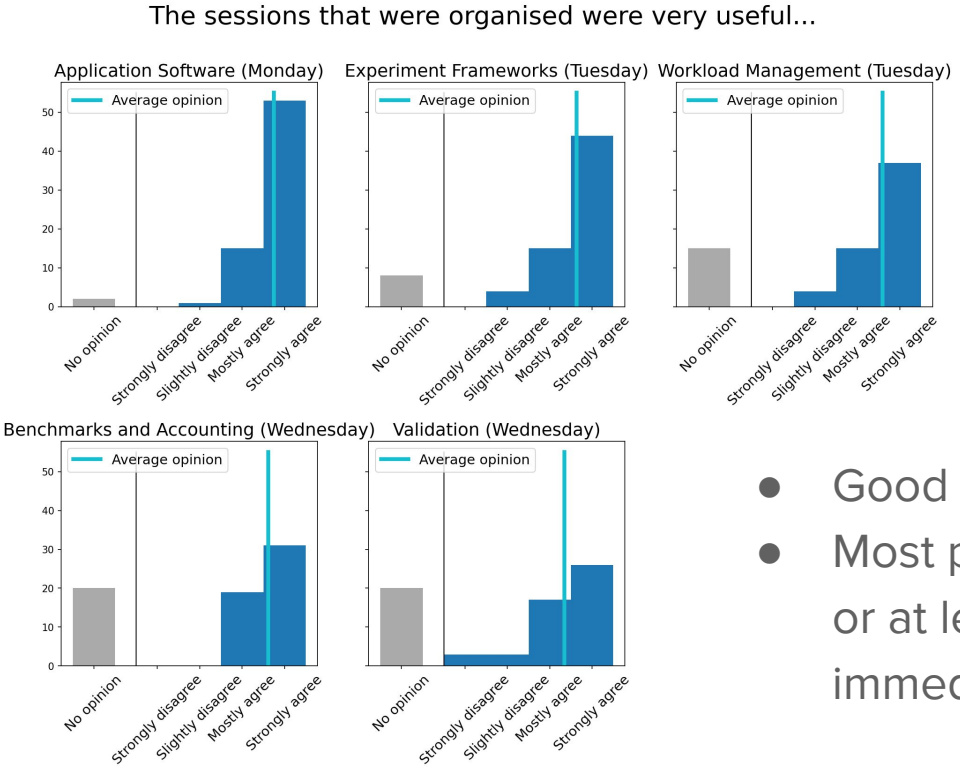
Given the live presentation format, the planned 50% of time for discussion was

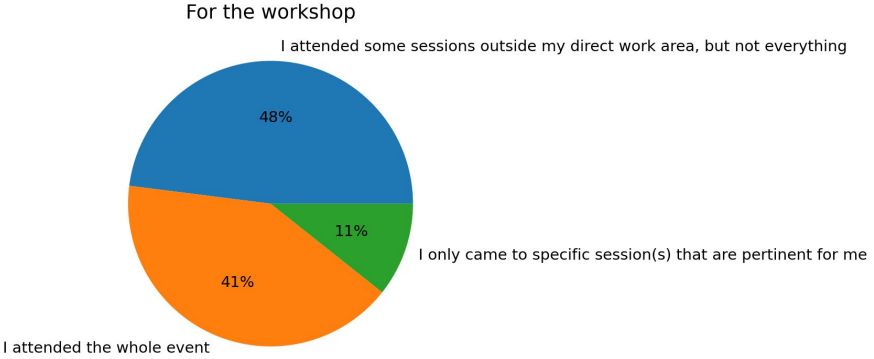


Sticking strictly to the scheduled timetable is important (even cutting off speakers if they badly overrun)



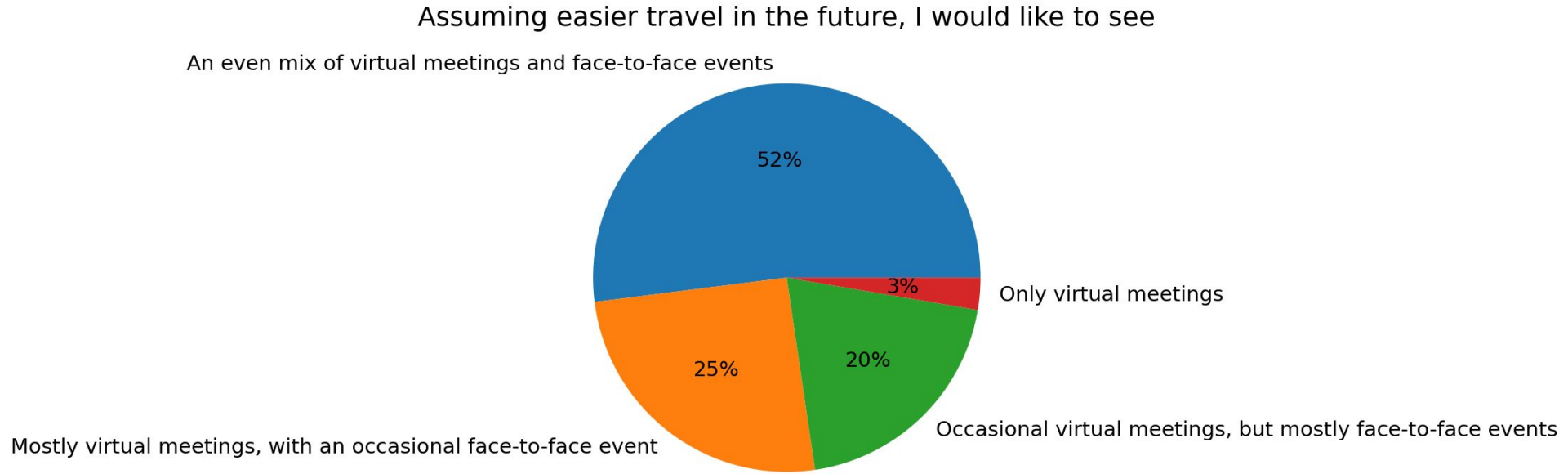
Did people like the topics?





- Good support for the topics covered
- Most people did attend the whole workshop or at least some sessions beyond their immediate work areas

Future Events



- High level of support for virtual events like this
  - But people also want to have a mixture with face-to-face workshops as well

Conclusions

- HSF-WLCG *Virtual* Workshop was a success
  - People attended in significant numbers for all of the sessions
  - Even outside their immediate work areas
    - Thus vindicating this as a workshop, instead of a series of topical meetings
- Organising a virtual event is helped by...
  - Material and notebook available in advance, but also during the sessions
  - Having restricted timeslots to help with focus and attendance
  - Generous time for discussion, with live focused introductions that keep to time
  - *Virtual workshops can be part of our suite of collaboration tools in the future, even when more normal travel can restart*
    - A few people noted the lower barrier to entry for virtual meetings (travel time and money)
- Topics of accelerators continues to be a key area of R&D for HEP
  - In addition to many other issues we did not manage to cover
  - *We hope to organise a further virtual meeting later this year*

More feedback, planning and discussion HSF-WLCG meeting 17 June