# HSF WLCG Workshop Survey Feedback and Outcomes

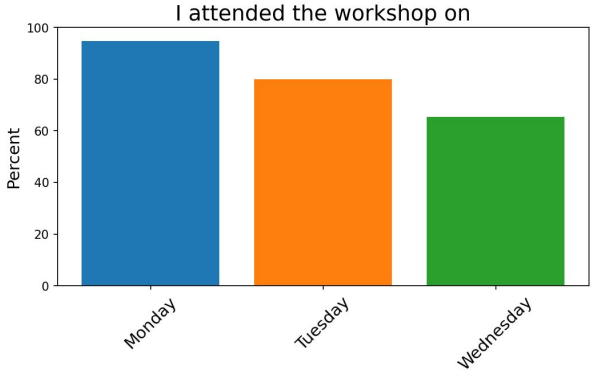
Graeme Stewart, for the workshop organisers

2020-06-17

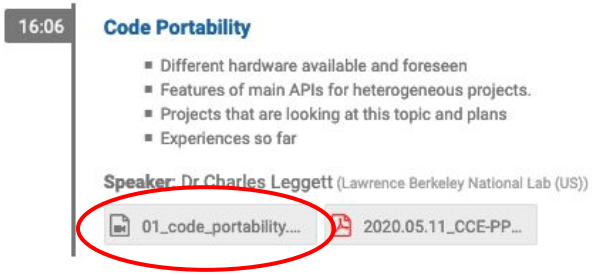
# WLCG-HSF Workshop

- 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
- New Architectures, Portability, and Sustainability theme
  - Monday: Application Software
  - Tuesday: Processing Frameworks
  - Wednesday: Validation and Accounting
  - See, e.g., GDB presentation for summary slides
- 221 people registered
- Attendance peaked at 175 Monday, 150 Tuesday, 110 Wednesday
  - We had a clash with LHCOPN/LHCONE meeting on Wednesday :-(

From the survey...



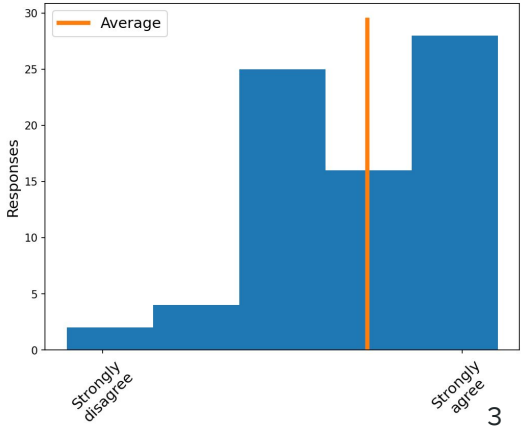
# Post-Workshop



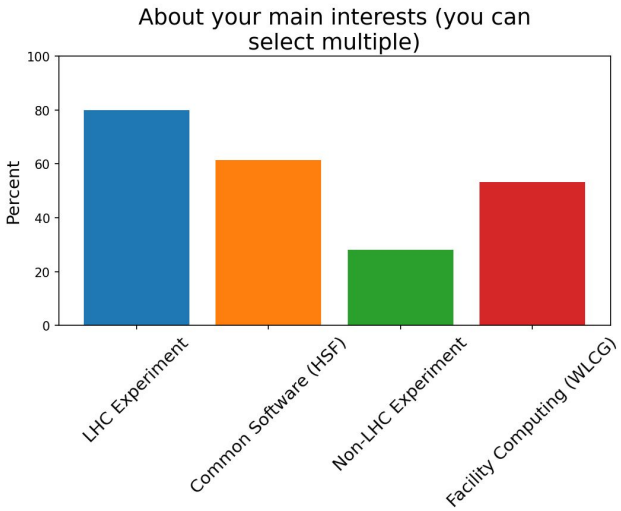

- Video
  - The recorded sessions were posted to Indico
  - Experience with recording via Zoom is quite positive
    - Quality is decent and easy to edit the video into separate files per talk
  - Bug in Indico statistics gathering make popularity hard to assess
    - Could also upload to YouTube (better platform, but ads…)
- 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)
    - Thanks to those who responded
- Outcomes and Summary
  - See document attached, in addition to these slides

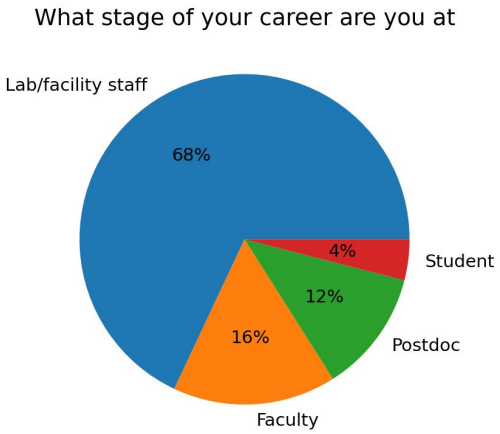
From the survey…

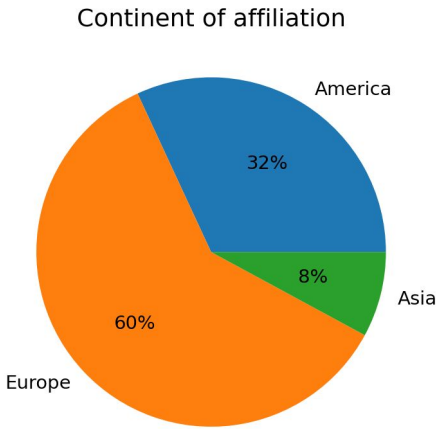
It is a good idea to record and post video of each of the sessions



# Who came?



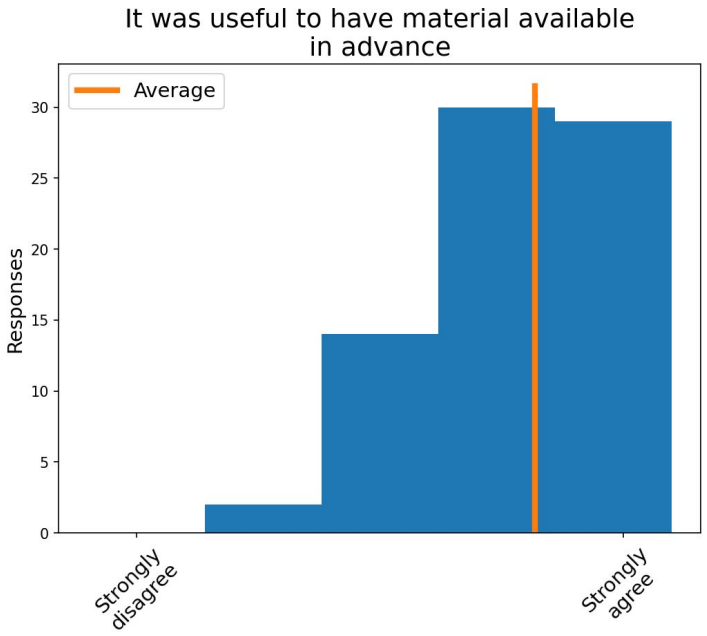



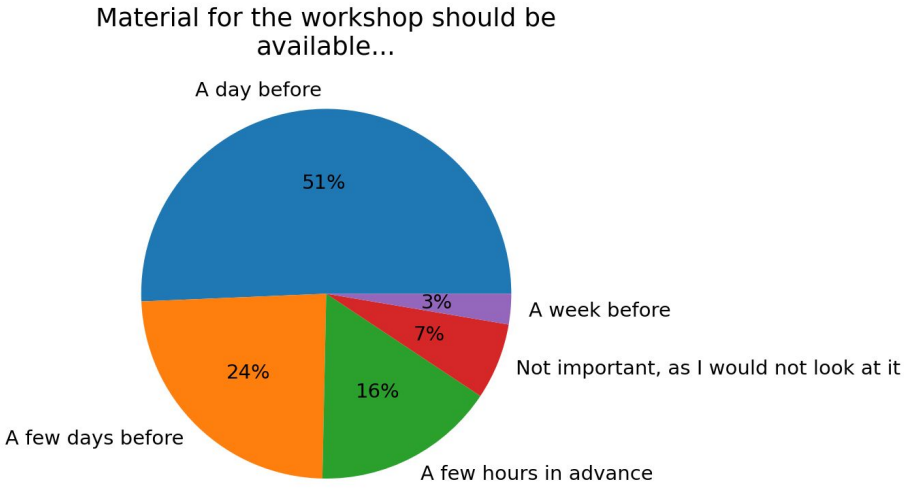




- Strong turnout from LHC experiments, as expected
  - Balanced across software and facility interests (non-exclusive!)
- Lab and facility staff made up ⅔ of respondents
  - *Would like to encourage more students and postdocs*
- Mostly European institutes (18% CERN)

N.B. Doesn't necessarily reflect the *timezone* where people joined from

# Pre-Workshop Matters I

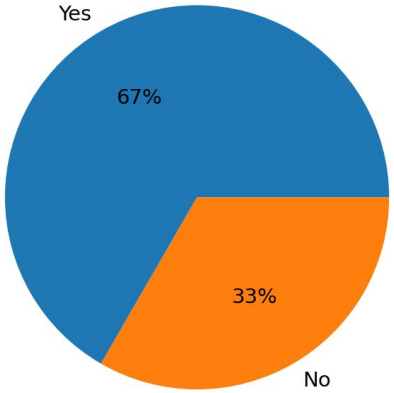


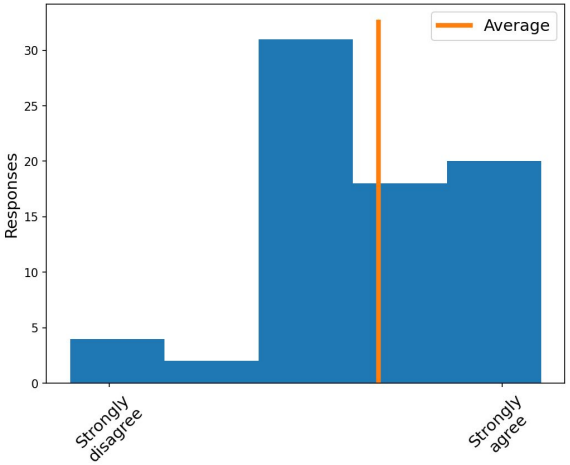

- People like material available in advance
- A day before is enough, a few days would be better

# Pre-Workshop Matters II

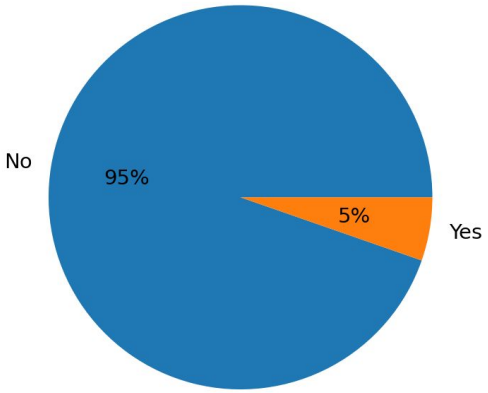
I reviewed the pre-workshop material before the conference




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

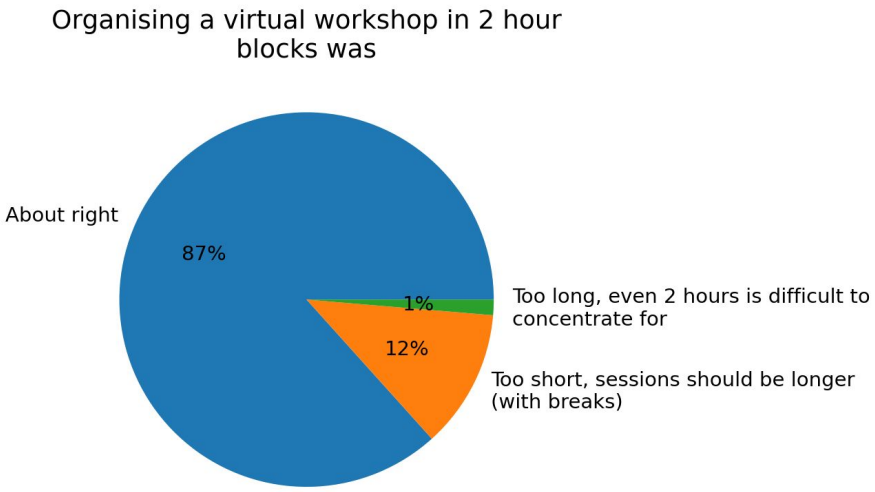


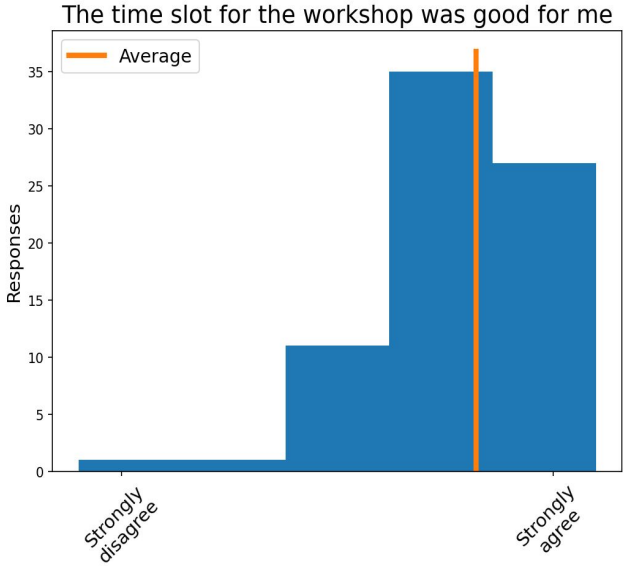

I posted questions in the notebook before the workshop




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

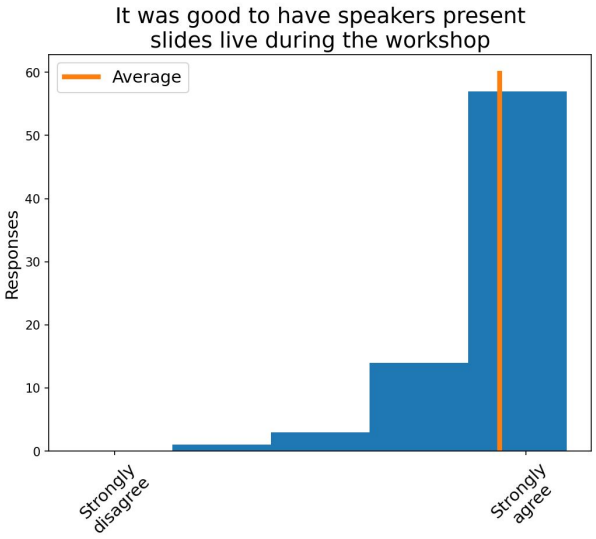
# Workshop Schedule

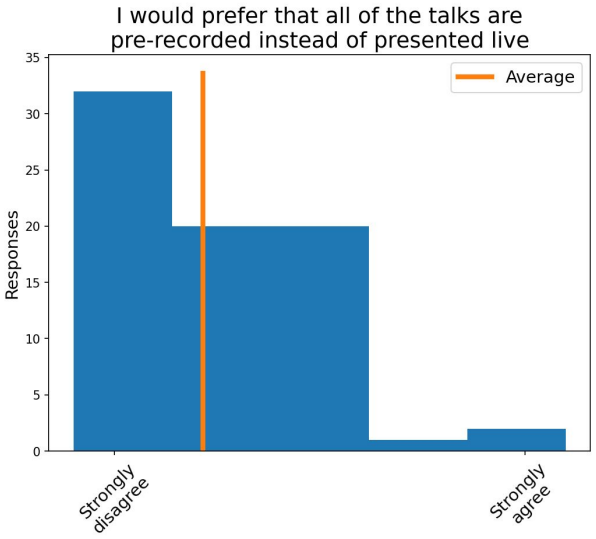



- 2 hour session blocks are a good length
- Timeslot worked for the people who attended... but selection bias!

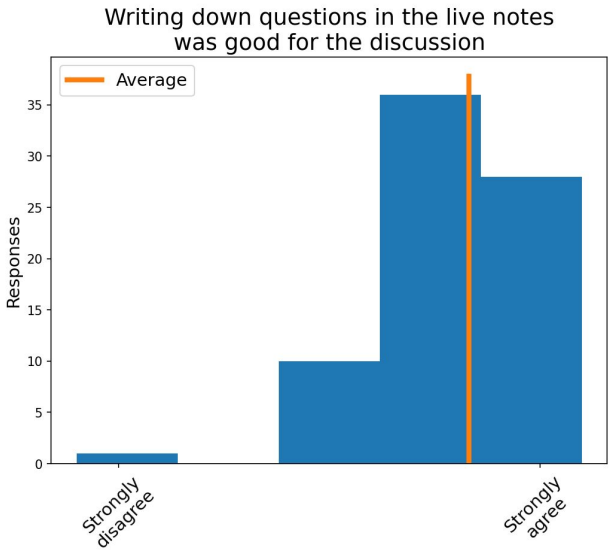
# Running the Sessions I

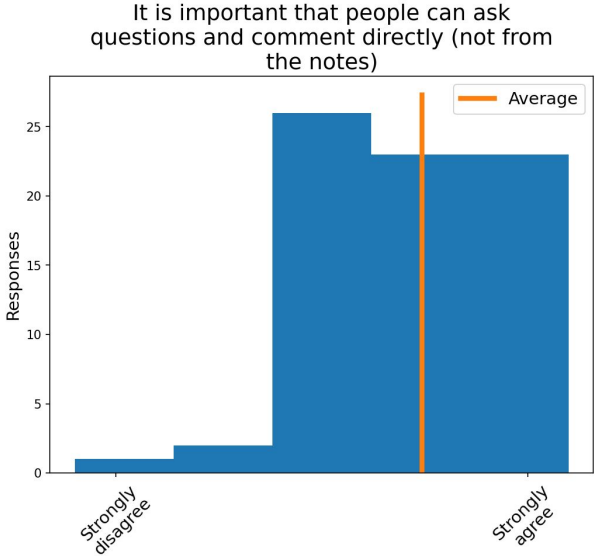


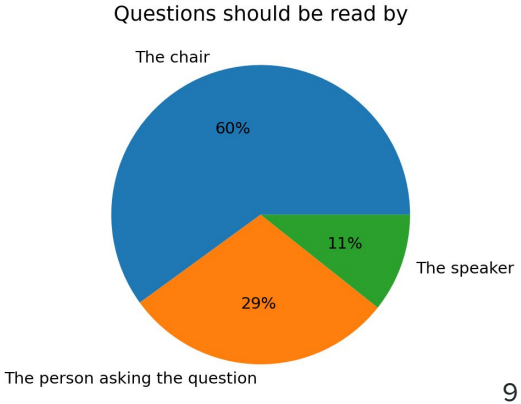




- Strong support for live presentations over pre-recordings
  - Possible reasons: People like the discipline of a set time? The interactivity of a live talk?

# Running the Sessions II



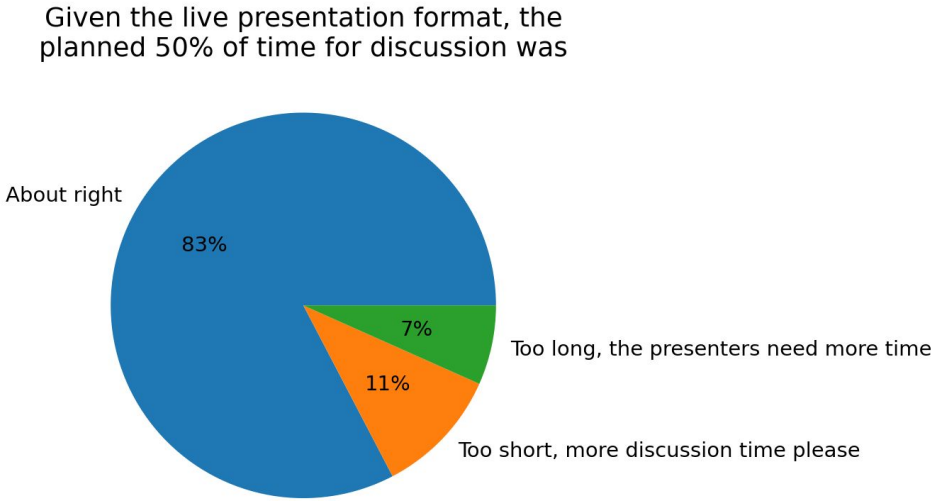



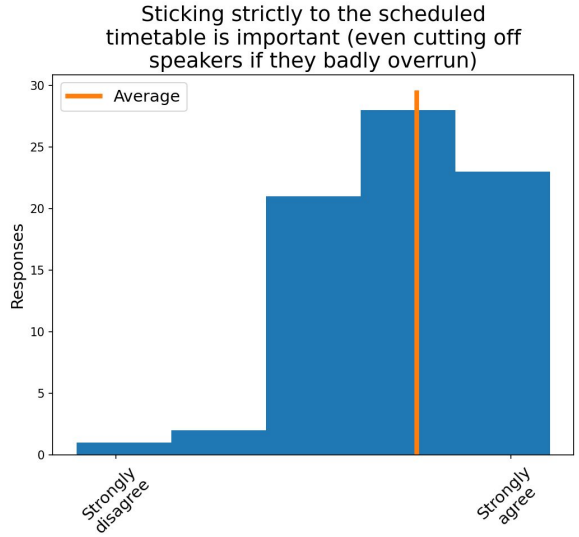




- Notebook is good for questions and discussion
  - It was challenging for the chair to keep track of the notebook during the sessions, but having a notetaker helped a lot
  - But the dynamism of direct questions is supported

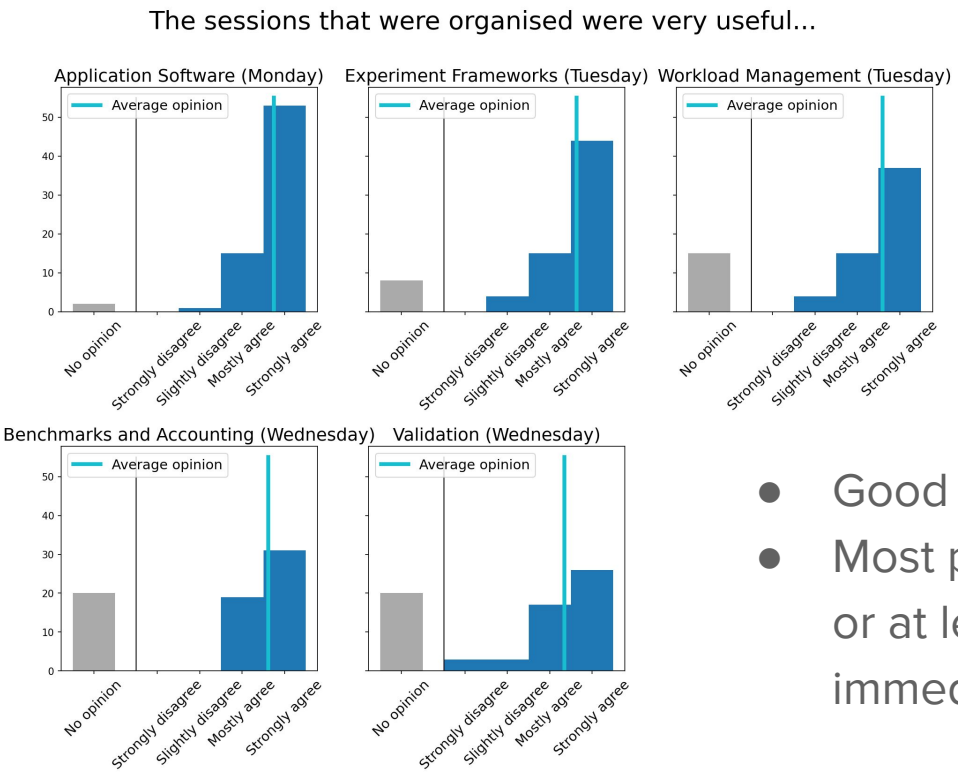
# Running the Sessions III

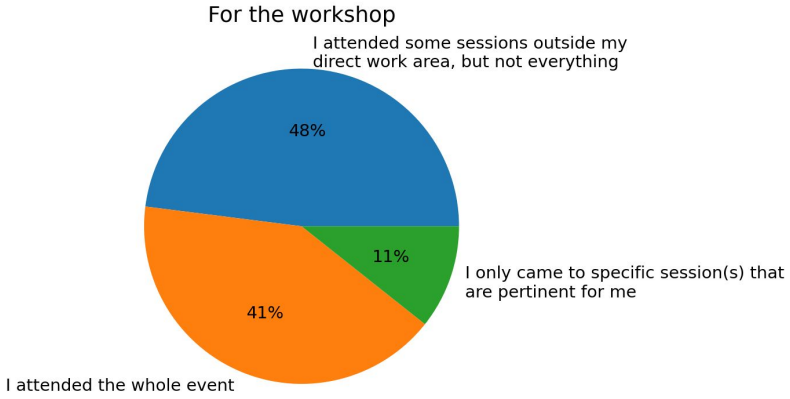
- 50/50 talk/discussion time planning was good
- People want better timekeeping!
  - We need good channels of communication between the chair and the speaker
  - We did get better at this on Tuesday and Wednesday
    - Dry-run and debugging 30 minutes before the sessions started helps a lot



# Did people like the topics?

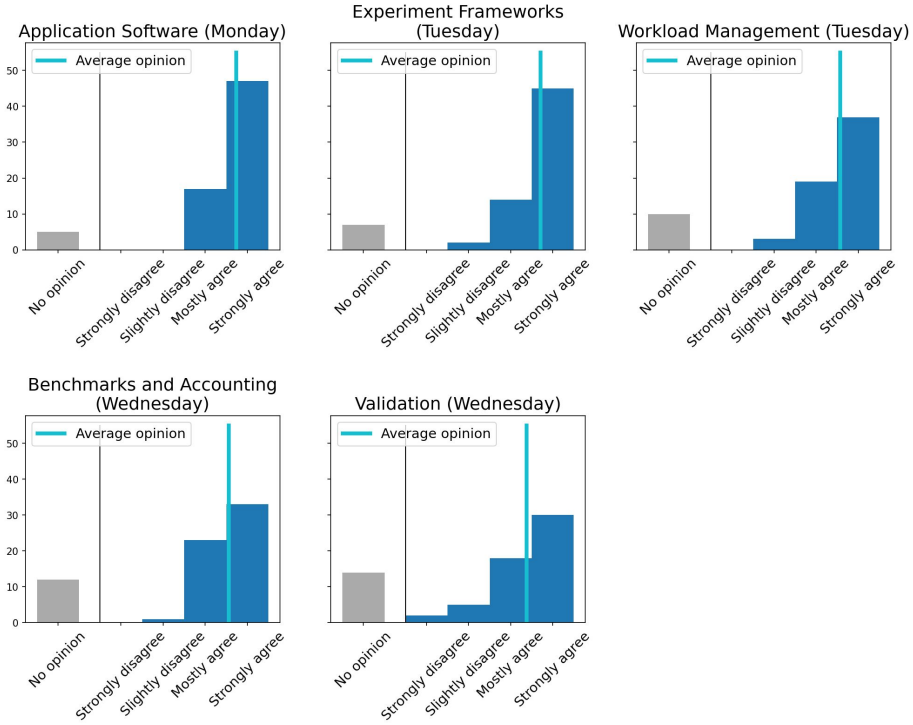



- Good support for the topics covered
- Most people did attend the whole workshop or at least some sessions beyond their immediate work areas

# Follow-ups and Other Topics

The sessions that were organised in this workshop should also be addressed again...

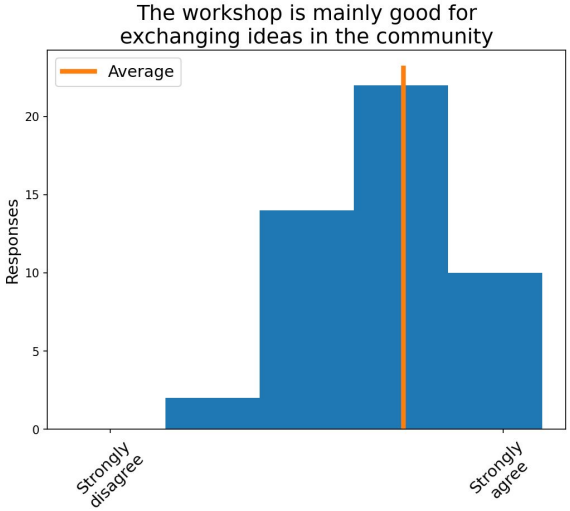


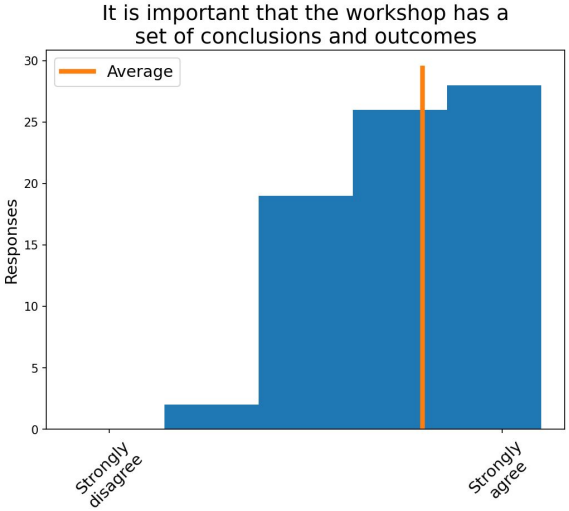

Other topic suggestions

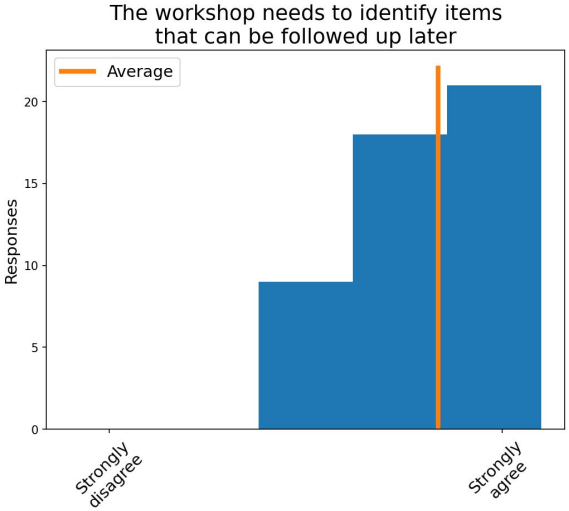
- Storage, data formats and data access (DOMA and Data Lakes)
- Analysis software and languages
- Deep learning
- Simulation
- Generators
- Authentication
- Virtual conferences

Plus *failures as well as successes* (avoid reinventing bad wheels)

# Outcomes and Follow-up

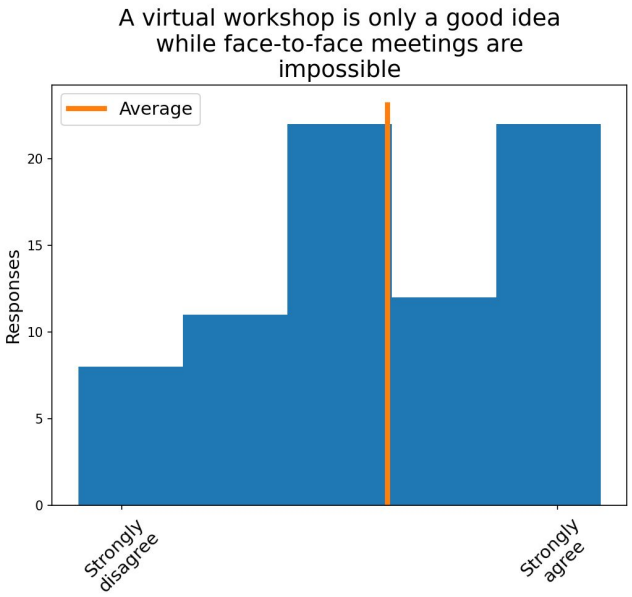


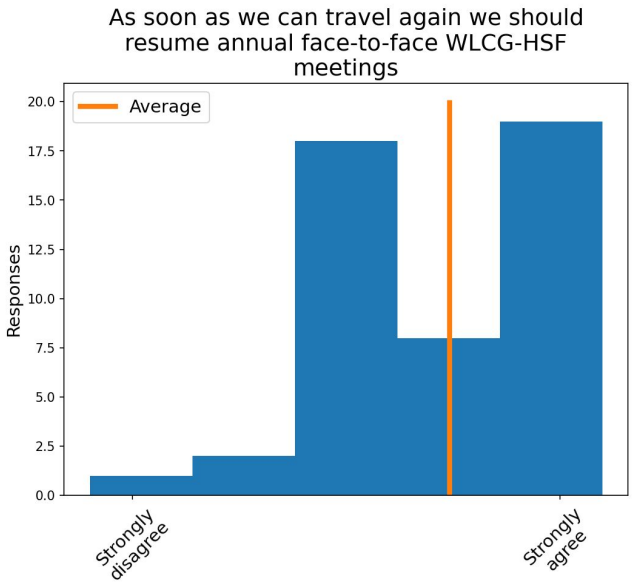




- Workshops are not an end in themselves!
  - Strongest support for follow-ups...

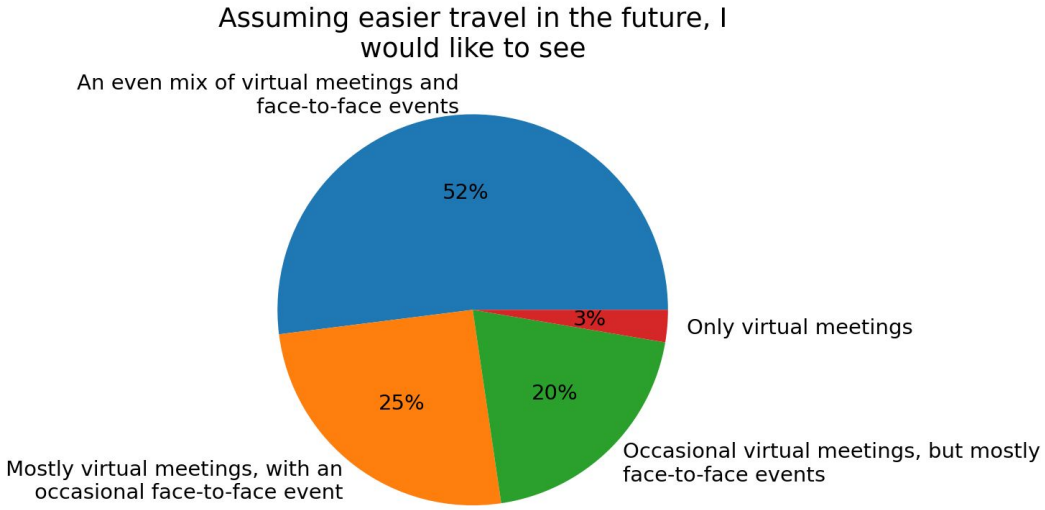
# Virtual or Face-to-Face Events?




- Support for resumption of face-to-face meetings when we can, but people see a place for virtual events

# Future Events




- High level of support for a mix of virtual events with face-to-face workshops

# Specific Feedback Points

A mixture of items we received directly or from the survey (this usually samples what people loved or hated the most)

- Virtual workshops can be shorter and more frequent
  - Lower barrier to entry allows people who would not have f2f meeting funding
  - Lower environmental costs
- Virtual workshops shouldn't replace f2f meetings
  - Some people didn't like virtual meetings at all
- Time keeping is **very important**
- Notebook is helpful to allow more people to contribute

Overall, general thanks for having organised a successful workshop

# Follow-up and Outcomes

- We have tried to summarise what we felt were the most important follow-on items from each of the sessions
  - See attached document
  - Of course this is our view and we'd be very happy to hear yours

# Application Software I

- HSF is a forum for Portable Parallelisation results and discussions
  - Many groups looking at this, HEP-CCE has well focused work package
  - As many examples as possible based on real HEP codes
- There are advanced training opportunities here
  - There are quite a lot of events going on (openlab, SIDIS, CASUS)
  - Online Alpaka Tutorial: https://indico.cern.ch/category/11733/
  - HSF Training WG is working on ways to make training sustainable
- HSF Reconstruction WG covers many code examples and projects right now and has been active in soliciting talks and encouraging discussion

# Application Software II

- HSF Simulation WG have started a series of R&D lightning talks with talks from simulation accelerator projects
  - Ray tracing R&D
  - Managing particle populations on device (generation, propagation and extinction)
  - Porting modern physics models as GPU kernels
- Management of simulation workflow between device and CPU will need development
  - Link to Frameworks WG
- TensorFlow - is an ideal subject for PyHEP (and analysis WG)
  - Many different packages: TFA/AmpliTF, zfit, ComPWA/tensorwaves, VegasFlow for Monte-Carlo techniques, internal BESIII/LHCb library by UCAS Beijing
  - Move to TensorFlow2 brings up common problems
  - Differentiable Computing HSF activity is also proposed and matches this kind of library.

# Experiment Frameworks

- Frameworks area has a lot of R&D going on:
  - async offloading
  - event batching
  - interface exposed to developers
  - backend switching (CPU, GPU flavours)
- These are all topics for the Frameworks WG
- Allied issues, particularly for the Reconstruction WG (algorithms)
  - There will be implications for generators and simulation as R&D develops in this area
  - Training continues to be an issue, with progress here (e.g. upcoming Alpaka workshop)
- Links to Portable Parallelisation as well

# Experiment Frameworks and Workload Management

- Interactions between WMS and Framework, is a cross-cutting topic with more complex information needing to be transmitted to help broker jobs
  - Further discussions needed, but not clear in which forum? Each experiment has different combinations of software. Interaction with resource providers to help standardise resource descriptions (GDB) and methods of probing.
- Code bases being flexible to different devices will help a lot
  - Back to HSF WGs and generic coding approaches - know what targets you can use
- Flexibility: we love containers, continue to promote evolution towards them (GDB) and share site experiences in managing them through Kubernetes
  - This is also relevant for managing GPUs at sites

# Validation

- Assess technical needs within experiments (results could be discussed in an HSF meeting). HSF Forum meeting.
  - No common projects foreseen yet
  - Would need to be some up-front interest/commitment from experiments and projects
- Assessing numerical stability is a topic for both algorithmic design and technical improvements
  - Some work on this with the verrou tool, HSF Tools group
  - Patatrack validation was an excellent example of best practice for a specific component

# Benchmarking and Accounting

- Benchmarking - impressive results from HEPiX WG
  - Wider questions to be followed up in GDB (technical), and WLCG MB
- Different approaches between benchmarking for procurements and pledging based on *usable capacity*
  - Test usable capacity on guinea-pig HPC site
  - Anticipate changes based on development of experiment software and different versions of drivers and API layers
- Can there be one number (that rules them all) for resource capacity?
  - Different metrics per experiment are a problem for sites supporting multiple experiments
- What about power efficiency (climate impact) as well as computing power?
  - This is not an ignorable responsibility for HEP

# 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)
- Many follow-ups identified from the subjects and many other subjects could be addressed

Organisers feeling is that an appropriate time for a next workshop would be November: tentative dates 19-20 + 23-25 November