Mu2e-doc-5262-v6

Fermilab

Managed by Fermi Research Alliance, LLC for the U.S. Department of Energy Office of Science

Answers to Questions from the Committee

Rob Kutschke

Mu2e Computing Review

March 6, 2015

Question 1

- Mu2e requested event display development support from SCD. Please provide more details and concrete justification.

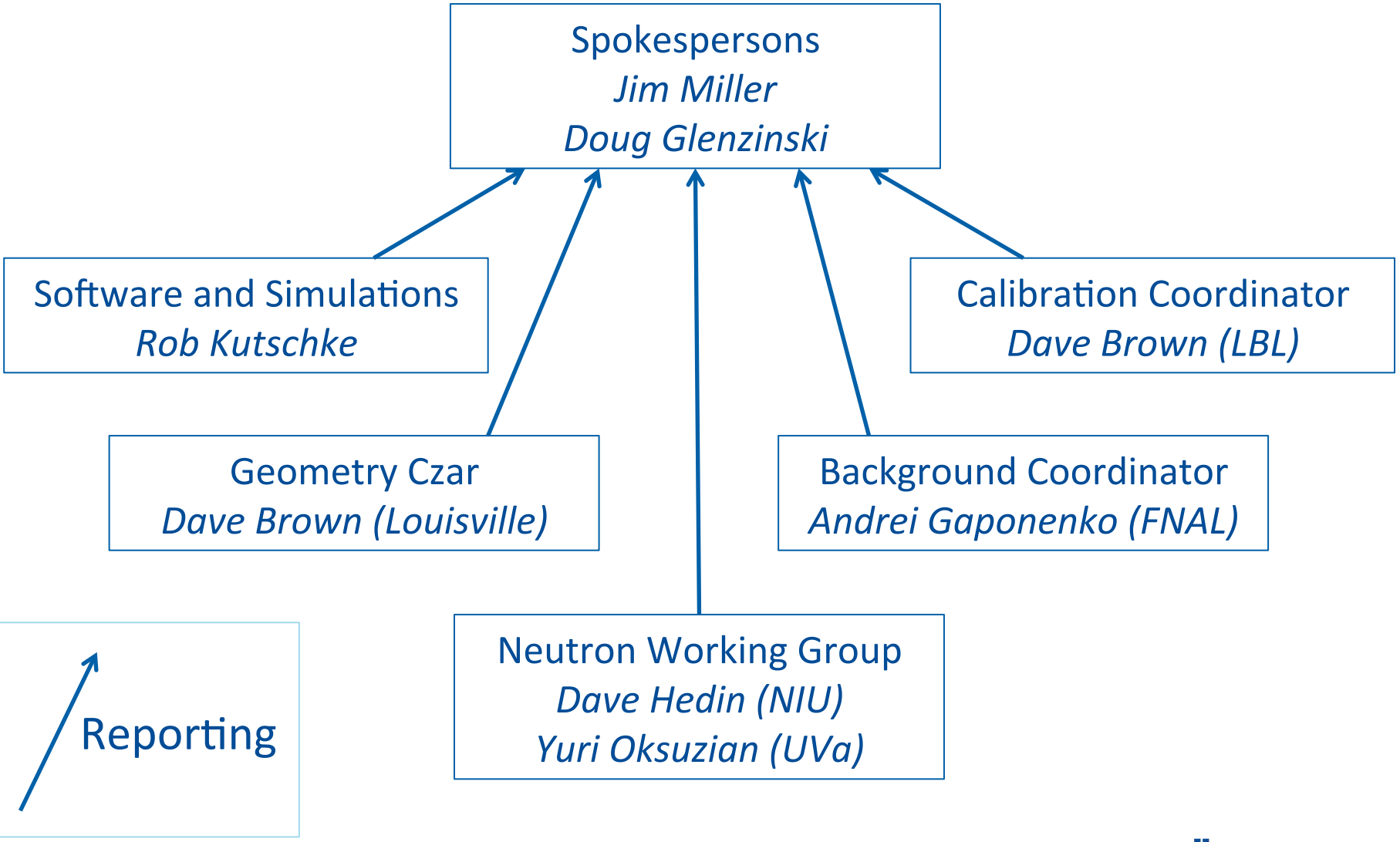
- Several collaborations report that a professionally developed event display improved the overall performance of their experiment.
- Our tracking code does not find all of the tracks that we believe it should. The right event display will be a powerful tool to investigate.
- Similarly for understanding reconstruction in the Calorimeter and CRV.
- A graphics expert will understand how to develop a tool that best exploits the features of the underlying toolkit such as how to design for good scaling behaviour and how to maximize OS and hardware independence.
- Our current event display maintainer has responsibilities as a physicist which limit his time to work on it.
- If time permits we will demo some of our existing packages at the end of this session.

Question 2

- Give an indication of how many people and how many FTE are behind the management boxes. Where are SCD people involved?

- Software and Simulations: 18 people, 7 FTE
  - Includes developers and users
  - Includes Krzysztof Genser (SCD) ~0.5 FTE on Geant4
- Background: 7 people, 2.6 FTE
- Calibration: 5 people, 1.8 FTE
- Geometry: 4 people 0.6 FTE
  - Includes Krzysztof Genser (SCD) ~0.25 FTE
- Neutron WG: 8 people, 2.5 FTE

Management Structure (1)



Question 3

- CD3 MC production is planned to start on April 1st; many of the ingredients are new compared to the former TDR production. Please present:
  a) timeline and scope for development tasks that still have to be done
  b) timeline for integration tests and scale tests that still have to be performed
  c) timeline and plan for operating the infrastructure during the 5 months of the campaign
- These should include manpower estimates from mu2e and SCD for the items listed above.

Question 3a)

- Mu2e Code tag for Apr 1 processing; development and testing.
  - 3 FTE*3weeks (Rob,Andrei,Yuri,Ralf, Dave (Louisville),Bertrand,Zhengyun...) SCD=0
- CVMFS Apr 1 configuration
  - 0.3*1 week (Ray) SCD=0
- CVMFS load tests
  - 0.1*3 weeks (Ray) SCD=0
- Modify mu2egrid workflow for SAM and dCache
  - 1FTE*3 days(Andrei), 0.2 FTE * 3 days (SCD)
- Testing new grid workflow
  - 0.2 FTE*2 weeks (Ray, Andrei), SCD=0
- Data flow optimization
  - 0.05*3 weeks (Ray) 0.05*3 weeks (SCD)

Totals after answer 4)

Question 3b)

- timeline for integration tests and scale tests that still have to be performed

- Both will be underway next week.
  - It's hard to predict what problems will be found
  - If it goes reasonably well, should wrap up in week 2.
- Numbers included in 3a)

Totals after answer 4)

Question 3c)

- Timeline and plan for operating the infrastructure during the 5 months of the campaign

- 0.2 FTE total from Andrei, Yuri, Ralf and Ray
- 0.05 FTE SCD
- 0.5 FTE operators

Totals after answer 4)

Question 4

- Please include manpower estimates from mu2e and SCD for the commissioning and operation of production on the Open Science Grid.

- OSG commissioning
  - This has big unknowns, and is a rough projection based on the previous weeks experience
  - 0.2*3 weeks (Ray) 0.2*3 weeks (SCD)
- OSG load testing
  - 0.1*3 weeks (Ray)

Totals next page

Effort totals for questions 3 and 4 for the April 1 target:

- Ray: 1.85 FTE weeks
- Andrei: 0.8 FTE weeks
- SCD: 1.35 FTE weeks
- The ensemble:
  - Rob, Andrei, Yuri, Ralf, Dave (Louisville), Bertrand, Zhengyun, and others
  - 9 FTE weeks
    - This is high estimate
    - There is scope contingency in what is ready for stage 1 vs stage 2 vs stage 3 vs analysis

Question 5

- Explain in detail the expectations and timeline for analyses based on CD3 MC production, including workflows and number of users.

- About 15 topics will be investigated by about 15 Mu2e collaborators.
- Most work done between Sept 1 and Dec 1
  - Some development work done will start earlier
- Use "mu2egrid" to run on Fermigrid (no offsite running).
- Many jobs will read pre-mixed background files.
- A 10-40 TB pre-mixed dataset will be read about 200 times
- Most jobs will be single stage, producing small size outputs.

Question 6

- What is the impact on the CD3 deliverables if mu2e only achieves half the CPU hours that are required? What is the contingency plan in case of a deficit?

- We believe that this would delay the CD3c Review.
  - The emergency plan is to explore for additional resources that are available at a cost of no more than our own time; failing that we need to accept the delay.
    - Roser is asking for cycles on ASCR HPC
  - If we were told that we had about ~70% of the available cycles we have a plan to triage.
    - Reduce targeted cosmic study to 50%
    - Reduce all others to 80%

Question 7

- Where are we meeting for beer?

- Users Center 5:00 PM today.