Fermilab

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

Summary

Ray Culbertson

Mu2e Computing Review

6 Mar 2015

Overview

- Summary response to the charge
  - Part 1 – Current Computing Infrastructure
  - Part 2 – Simulation Campaign for CD3
  - Part 3 – Developing the Offline Infrastructure

- Summary of Notable Issues

Part 1 – Current Computing Infrastructure

The current offline computing infrastructure and tools, including build and release tools, simulation tools, framework, database, workflow, workflow management, data management, and operations. Are the tools, infrastructure, and established processes appropriate for this stage in the experiment? Are the tools, infrastructure, and established processes sufficient to engage non-expert resources from the collaboration? Are best practices being employed as appropriate? Is the experiment appropriately leveraging tools and expertise provided by SCD?

Part 1 – Current Computing Infrastructure – details (1)

Are the tools, infrastructure, and established processes appropriate for this stage in the experiment?

- A proven baseline ability – successful TDR, CD2
- All needed functionality is there
  - All backgrounds have been simulated
  - All detector design questions can be investigated
- Decisions can be based on best metrics - final sensitivity
- Next steps in development are desirable and soon underway
  - Code and build cleanups
  - Geometry cleanup
  - Validation inserted several places
  - Improvements in specific object persistency
  - Data off bluearc and in dCache
  - Databases will be coordinated with SCD, probably in 2017

Part 1 – Current Computing Infrastructure – details (2)

Are the tools, infrastructure, and established processes sufficient to engage non-expert resources from the collaboration?

- Art manual is a comprehensive introduction that will bring a user into the main framework, but takes a while
- Considering a very fast start process with a simple ntuple
- Would like to coordinate this with an analysis ntuple plan
- That ntuple plan will evolve so stage it
  - Prototype solutions are available
  - Many demands of efficiency, maintenance, personal styles
  - First stage, consulting with the collaboration, within six months

Part 1 – Current Computing Infrastructure – details (3)

Are best practices being employed as appropriate? Is the experiment appropriately leveraging tools and expertise provided by SCD?

- Well-aligned with SCD practices, no notable conflicts
- UPS, git, gcc, command line tools – latest, standard
- G4, Root, CLHEP, boost, etc are latest, standards
- Jobsub_client, dCache are latest, standard usage
- Full buy-in for art+fhicl
  - persistency adequate, but needs updates
- Will coordinate with SCD on database access in 2017

Part 2 – Simulation Campaign for CD3

The requirements for the simulation campaign leading to the late 2015 CD review and the plan for meeting these requirements. Is the plan reasonable and achievable? Are the current infrastructure and tools capable of meeting these requirements in a timely manner? If not, what additional infrastructure and tools are required? Have adequate personnel resources, both from the experiment and SCD, been identified? If not, where are the personnel shortfalls?

Part 2 – Simulation Campaign for CD3 – details (1)

Is the plan reasonable and achievable? Are the current infrastructure and tools capable of meeting these requirements in a timely manner? If not, what additional infrastructure and tools are required?

- In general, offline code is ready
- Some specific simulation code and simulation plan details are still being developed, three FTE over the next three weeks
- Operations infrastructure needs approximately one FTE, over the next three weeks, of collaboration development
  - Grid workflows to and from dCache
  - OSG operations

Part 2 – Simulation Campaign for CD3 – details (2)

Have adequate personnel resources, both from the experiment and SCD, been identified? If not, where are the personnel shortfalls?

- Physicists are stretched, but we expect to keep the schedule
  - Preparation of jobs (many collaborators)
  - Operations (Rob, RLC, Andrei, Yuri, Ralf)
  - Projects are being prioritized, and not all need to start Apr 1
- May require 0.5 FTE of SCD operators
- May require OSG issues to be treated as high priority

Part 3 – Developing the Offline Infrastructure

The plan for developing the offline computing infrastructure, tools, and processes (as defined in (1) above), leading to Mu2e data taking and beyond. Is the plan reasonable and achievable? Does the plan appropriately leverage tools and services provided by SCD? Are best practices employed? Do the required personnel resources, both from the experiment and SCD, seem reasonable? Will the plan result in tools, infrastructure, and processes that are capable of producing "analysis ready" data in a timely manner and allow for significant engagement of non-expert resources from the collaboration?

Part 3 – Developing the Offline Infrastructure – details (1)

Is the plan reasonable and achievable? Does the plan appropriately leverage tools and services provided by SCD? Are best practices employed?

- We are staying aligned with SCD tools and see no notable future conflict (art, G4, jobsub_client, ifdh …)
- We are employing best practices wherever possible, no notable conflicts (dCache, FTS, CVMFS…)

Do the required personnel resources, both from the experiment and SCD, seem reasonable?

- Physicists are stretched, but we expect to keep the schedule
- May require 0.5 FTE of SCD operators

Part 3 – Developing the Offline Infrastructure – details (2)

Will the plan result in tools, infrastructure, and processes that are capable of producing "analysis ready" data in a timely manner and allow for significant engagement of non-expert resources from the collaboration?

- Systems are well-developed and have proven success
- No major issues identified
- Will implement internal review and planning procedures
  - Better code review
  - Analysis workflow planning – datasets, loads, features, working style – will be discussed with the collaboration
- Non-expert engagement has been brought to the foreground
  - Investigations starting now
- Full timeline and details in Rob's roadmap

Summary of Notable Issues

Mu2e expects to meet the CD3 challenge!

Projects for the collaboration:

- Several areas marked for improvement
- Will perform analysis workflow review, including newcomers

Requests to SCD:

- CPU for CD3
  - Support for 2000 Fermigrid, 2000 OSG slots for CD3 challenge
  - High priority for jobsub and OSG issues
- 0.5 FTE from operators, when needed
- Graphics expert help for development of displays
- Maintain G4Beamline package, if orphaned