

Observations from the Run II Task Force

Rick Snider

Fermilab Computing Division

Outline

- Run II Task Force
- Computing analysis
- Observations on risks
- Strategies to control risks
- Summary and status

Run II Computing Review

Sept. 14, 2005

Origin of the Task Force

- HEPAP manpower survey
 - Observed a possibly significant gap between required and available manpower to Run II experiments after 2006
- Director charged a task force
 - Understand potential staffing problems
 - Develop remedies to identified shortfalls
 - Submit interim report before P5 meeting

Full charge, summaries of current work on P5 meeting homepage
www.fnal.gov/directorate/program_planning/P5/P5_Sept2005/HEPAP_Subpanel_P5_Meeting.html

See talks by Montgomery, Roser, Wyatt

Collider Experiment Task Force

- Members from Fermilab, CDF/DØ, HEPAP study co-chair
- Formed four sub-groups:
 - Collaboration resource availability
 - Detector operations resource requirements
 - Offline computing resource requirements
 - Physics and algorithm resource requirements

Collider Experiment Task Force

- Offline computing sub-group
 - Members
 - CD division management
 - Run II Dept. management
 - Experiment spokes
 - Experiment offline leaders
 - Goals of computing analysis
 - Evaluate required computing resources through end of Run II
 - Focus on data processing
 - Exclude reconstruction algorithms
 - Identify risks and vulnerabilities in estimated requirements
 - Suggestions to mitigate risks, reduce long-term resource needs

Context of computing analysis

- Computing at the experiments
 - Require increased computing capacity as datasets grow
 - Increasing demand for off-site computing through end of run
 - Scalability, operational workload are central issues
 - Expect declining dedicated resources
 - Dedicated computer pools off-site not likely to grow indefinitely
 - Collaborators move to other projects
- Long-term strategy of the experiments
 - Adapt to OSG/LCG grid infrastructure
 - Maintains current capacity, provides growth potential
 - **Very different approaches at CDF and DØ**

Context of computing analysis

- Long-term strategy of the CD for Run II experiments
 - Adopt common solutions for common problems
 - Eliminate redundant efforts
 - Move key infrastructure into centrally supported services
 - Makes best use of Division resources
 - Insulates important systems from contraction of collaboration
 - Leverage LHC manpower devoted to grid technologies
 - Minimize net effort required to support expansion of remote computing

Context of computing analysis

- Goals of the Task Force
 - Take a broad view of existing computing program
 - Find issues that systematically introduce risk over the long term
 - Suggest actions or strategies that will position the experiments, the CD to mitigate or control that risk
 - Not charged with solving today's problems

Computing analysis

- Activities examined
 - Operations
 - Central farm, MC production, re-processing, data handling
 - System administration
 - Databases
 - Code support and builds
 - Data handling development
 - Distributed computing
 - Job control/tools
 - CD centrally supported services available to both experiments

Observations: areas of risk

- Identified areas of risk in estimates of future needs
 - Risk in implementing grid-based computing strategy
 - Deployment of SAMgrid and other grid technologies
 - SAMgrid interoperability
 - Development of grid-based analysis models
 - Vulnerabilities in staffing
 - Staffing sub-project leadership
 - Desktop, trigger farm and database support
 - DØ online systems / CDF and DØ offline code management
 - Scientific positions within the CD

Risks in grid strategy

- Deployment of SAMgrid and other grid technologies
 - Current products are labor intensive, require experts
 - Efficient installs become more important as grid needs grow
- SAMgrid interoperability
 - The cornerstone of grid technologies
 - Required for success of DØ computing
- Development of viable grid-based analysis models
 - Considerable use of remote resources for analysis already
 - No integrated grid-based solution
 - Continued expansion will require better technology

Risks in staffing

- Core computing system support
 - Critical systems supported by collaboration effort
 - Desktop system administration (DØ only)
 - One of primary analysis platforms at DØ
 - Trigger farm system administration
 - Database application support
 - Maintaining collaboration support may become difficult
- Staffing sub-project leadership
 - Need qualified experts to lead sub-projects
 - Limited, declining number of people in the collaborations

Risks in staffing

- DØ online systems / CDF and DØ code management
 - Critical tasks supported by single high-skilled individuals
- Scientific positions within the CD
 - Important resource for development sub-projects
 - Has been difficult to attract high caliber physicists to open RA positions
 - Feedback indicates that the CD has a perception problem
 - The view is that CD jobs are not good for RA's interested in later faculty appointments

Strategies to control risks

- CD strategy aimed at increasing efficiency of Run II support workforce, or reducing risk to key services
 - Focus on common solutions, consolidation of support, leveraging grid resources is sensible
 - Achieving goals may require short-term influx in manpower
- Specific actions to reduce long-term needs
 - Grid deployment team
 - Interoperability
 - Consolidation of key central services
 - CD guest scientists
 - Research Associate and Associate Scientist positions

Remedies

- Grid deployment team

- Streamline installation procedures
- Assist sites with grid-based technology deployments

Free developers to work on things such as grid-based analysis model

- Interoperability

- Aggressively pursue implementing interoperable computing solutions for distributed computing by the Run II experiments

- First make SAMgrid interoperable on OSG / LCG resources

Maximizes leveraging potential

Remedies

- Interoperability (cont'd)
 - Note some recent successes
 - SAMgrid co-exists with OSG on common head node
 - Share underlying VTD components
 - Direction for further interoperability with OSG well defined
 - Progress on making SAMgrid interoperable with LCG
 - Goal to be in production by start of calendar 2006
 - Fermigrid
 - $D\bar{0}$ re-processing on gen. purpose farm, CMS resources via SAMgrid
 - CDF glide-CAF utilizing general-purpose farm, CMS resources
 - Efforts outside CD
 - Working to make CAF functionality interoperable on LCG (CDF/INFN)
 - Working to deploy CDF glide-CAF at U. Wisconsin Grid Lab (CDF)

Remedies

- Consolidation of key central services
 - Many collaboration-supported services are within scope of work performed by computing professionals
 - Desktop and Level-3 trigger farm system administration
 - Good match to existing administration responsibilities held by CD
 - Will stabilize a fraction of the operational load for experiments
 - Database application support, etc.
 - Reduce risk of losing effort for these tasks
 - Cross-train teams to support similar services on both experiments

Improve efficiency, reduce risk by expanding central-services model

This is one of the primary goals of the Run II department

Remedies

- CD Guest Scientists

- Fill critical managerial, technical roles within experiments
- GS appointment is a tool to attract high-caliber physicists
 - Broadens pool of candidates for important jobs

Term positions to meet short-term needs

- Research Associate and Associate Scientist position

- Fill critical roles within experiments
- Strengthen scientific program of lab, provide mentors and role-models
 - Aimed in part to combat perception that computing specialization is contrary to strong physics orientation

Summary and status

- There are risks in resource availability through 2009
 - Run II Task Force has identified some particular areas
- Suggested actions to mitigate risks and improve workforce efficiency
 - Consistent with overall CD strategy
 - Considerable success in some areas already
 - Short-term implications of suggested remedies
 - Addition of a few computing professionals and scientists
 - Long-term implications
 - Efficiency gains result in modest reduction in net effort
 - Unanticipated problems could again require short-term increase

Summary and status

- Run II Task Force will continue working, per charge
 - Perform more detailed analysis of resource balance
 - Refine risk assessments, suggested remedies
 - Submit final report to Director