

CD-Doc 2597

Run II Resource Project

Rob Kutschke, CD/IDS
Feb 7, 2008

Agenda

10 min	Rob K.	Charge
10 min	Rob K.	Status and next steps
5 min	G.P. Yeh	Status
5 min		Next meeting

Charge

Develop and document a computing model for Run II that can be used to inform the division about computing requirements for Run II, both hardware and personnel, starting in FY08 and continuing until the end of the main analysis period, presumably the end of FY 2012. This includes but is not limited to:

1. Learn and document the existing CDF and D0 Run II computing models.
2. Work with the CD Running Experiments Department and with the collaborations to understand how the experiments are actually using existing resources in their different operating modes, reconstruction, skimming, MC, analysis, etc
 - Some of the information is already available and just needs to be collected.
 - Identify what metrics are missing and work with REX to collect them.
 - Identify any resources that are being used inefficiently.
3. Develop a single tool to express the computing models of both experiments. Vicky wants it in Excel.

Charge ...

4. Pay particular attention to several aspects that, in the past, received only cursory treatment:
 - Are the disk caches used in an optimal way?
 - What are the true drivers for doing physics analysis/
5. Ideally this model will be predictive, not just descriptive. On the time scale foreseen, however, this may be limited by the available metrics.

Deliverables

1. The combined tool for Vicky.
 - Prototype version due ASAP.
2. A short weekly report and a biweekly presentation.
 - Received by REX and offline management of CDF and D0.
3. A draft final report due by March 21, 2008(?)
 - Received by REX and offline mgt of CDF and D0.
4. A final report due ???
 1. Received by CD, CDF and D0

After the draft final report we can discuss if further work on improved metrics and a more sophisticated model merits additional effort.

The charge does NOT say that I should make any recommendations. I should provide data that allows others to make decisions.

Status

- Started Jan 21.
 - Jan 28 ... Feb 1: SiD group meeting and P5.
- Acquiring the background information, jargon etc
 - Attending the weekly SAM and CDF-Offline ops meetings
 - [Is there a D0 equivalent that I should attend?](#)
 - Goal to learn background information
 - Met with Adam Lyon, Stephan Lammel, Steve Timm to discuss what their groups do.
 - With Margaret Votava and Jerry Guglielmo.
 - I have the big picture now and am consolidating details.
- CDF and D0 models in hand and I am documenting them.

Next Steps

- Goal for Monday Feb 11:
 - Written document describing the two models.
- A few days later:
 - Prototype of spreadsheet tool to do both models.
 - Discuss with Vicky exactly what she wants.
- Work with G.P. and others to dig out the **current values** of the numbers that enter into the models.
- Work with REX and CDF/D0 offline to figure out how to improve accuracy of numbers, improve metrics etc.
 - Data mining the existing logs to find the biggest users that are under our radar and learn what they are doing.
 - Propose new metrics (I am aware that the group is overloaded ...)
 - How do physics analyses really use resources?
 - How are caches used? Are they over provisioned?
 - Amber observed of 800E9 events delivered by SAM to D0.
 - Reconcile this with a few E9 events to tape and known MC rates.

Information

- Weekly status reports: CD-Doc-2596
- Biweekly presentations: CD-Doc-2597
- All documents via keyword: RunIIResourceProject