

Homework from GDM April 4th

a) Line management to find out what actions and projects people say are needed to reduce the current and/or projected operational load in the GDM area by a factor of ~2. Plan to report these to GDMs as the information is available.

Can we make a schedule for this?

05/30/06

06/13/06

07/11/06

07/25/06

b) Make a table of stakeholder requirements metrics for Job Submission and Brokering.

		Minos	D0	CMS	CDF	LQCD	Other
Number of Simultaneous (CDF, D0) Jobs per day(CMS, prod jobs run >1 day)							
2006	Total	500	5,000	12,500	5,000		
	Offsite /Grid Fraction		50%	100%	50%		
	Offsite Applications:		MC, Prod	MC, Prod, Analysis	MC Analysis [1]		
	Number of jobs from single submit site						
2007	Total		6500	25,000	6500		
	Offsite/Grid Fraction		50%	100%	50%		
	Offsite Applications:		MC, Prod	MC, Prod, Analysis	MC, Analysis		
	Number of jobs from single submit site						
2008			8450	50,000	8450		
	Offsite Fraction		50%	100%	50%		
	Offsite Applications:		MC, Prod	MC, Prod, Analysis	MC, Analysis		
	Number of jobs from single submit site						
2009			10985	?	10985		
	Offsite Fraction		50%		50%		
	Offsite Applications:		MC, Analysis	MC, Prod, Analysis	MC, Analysis		
	Number of jobs from single submit site						

More detail on CDF:

these are current, measured, CDF numbers:

	AVG	MAX
Running sections:	4500	6000
Started/hour:	650	1200
Started/day:	16k	24k

For FNAL alone:

	AVG	MAX
Running sections:	3200	3800
Started/hour:	500	1200
Started/day:	12k	20k

For offsite:

	AVG	MAX
Running sections:	1500	3000
Started/hour:	150	500
Started/day:	4k	7k

I know the max numbers don't sum up, but this is fine, since not all the peaks happen at the same time.

The averages are close enough, though.

There are a few more comments on the table you sent us:

1) For us offsite != grid != non FNAL money

Our current situation is

50% non-FNAL-bought CPU

~35% offsite CPU

~35% grid resources

long term:

50% non-FNAL-bought CPU

~100% grid resources

2)

We are already doing analysis offsite, at least at one site.

And that site is a Grid site, exploited via a GlideCAF.

Missing Capabilities:

Meeting the Job metrics

Support for offsite analysis

User Priorities

Jobs/Submission Point

Range of Job CPUs

Resource Selection

Monitoring

Common user interface for onsite/offsite

User level accounting?

<.25 FTE support for software and operations?

c) Make a table of stakeholder data caching current and future performance metrics.

Talked to CDF and contacted CMS, SAM, D0. Not much response to date.

Need information about current usage for background.

Then for each of 2007, 2008, 2009 need:

Size of disk cache on site (off-site?)

Cache refresh rate expected

Write access rates

Read access rates and patterns