

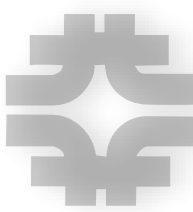
# Establishing Cost and Benefits of U.S. CMS Tier-2s

LATBauerdick/Fermilab  
U.S. CMS Software and Computing Manager

U.S. CMS Tier-2 Meeting at the  
OSG All-Hands Meeting,  
U. Nebraska Lincoln, March 19, 2012



# U.S. CMS needs to perform a Cost/Benefit Analysis of T2s and T1



- ◆ This could and probably should have happened at the time of the new Collaborative Agreement with the NSF
  - ★ however, it did not fit into the time scale of writing the CA proposal
    - ◆ U.S. Atlas did such an analysis and “confirmed” the cost effectiveness of their Tier-2 program w/r to Tier-1 cost, and also added additional sites to their 5 existing Tier-2 centers, for a total of 10 universities
- ◆ the 2011 DOE/NSF review recommended U.S.CMS to perform
  - ★ “U.S. CMS should perform a cost benefit analysis of each Tier-2 site to determine how reductions would proceed if required”
  - ★ “U.S. CMS should perform a cost/benefit analysis of the Tier-1 and Tier-2 activities”
- ◆ both in 2011 and 2012 DOE/NSF reviews we got probing questions
  - ★ about the relative performance and relative cost/benefit of T2 sites
  - ★ relative cost of computing equipment to be operated at T1 and T2 sites
- ◆ we promised to perform the cost/benefit analysis in particular to decide on eventual supplemental funds



# How things can go awry...

- ◆ in parallel session we told a pretty complete story highlighting the benefits of individual sites, and got very positive feedback from reviewers
- ◆ however, at the Q&A session, in plenary, we were asked a very pointed question:
  - ★ “what is the cost to U.S. CMS to operate a CPU at each Tier2 and Tier1”
  - ★ ... and in the end I had to provide just a table, without being able to give the context, and it wasn't pretty:  
the #CPU cores/site vary from 1662 to 5200, and the <cost>/FTE from \$124k/yr to \$184k/yr (including overheads etc)
- ◆ Let's try to control our story and how we tell it and that means to be fully in the picture!



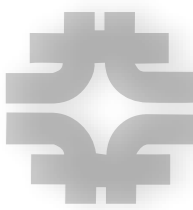
# Goals for the Cost/Benefit Analysis



- ◆ stated U.S. CMS goal: have all sites perform well and look good!
- ◆ assessing metric will enable us to initiate corrective measures in case of deviations
  - ★ will make an initial cost/benefit assessment of costs (see later)
  - ★ review other criteria and document them
  - ★ the iterate and come to an assessment
- ◆ We want the process to be fair and open
  - ★ expect that most of the input comes from the site itself
  - ★ make information available cross sites
- ◆ We also want to allow a “sportive” element
  - ★ sites should be able to compare themselves, learn from each other etc
- ◆ At the end we expect that all sites will be judged as well performing
  - ★ goal is to have all sites “above threshold”, if needed with iteration
  - ★ eventual ranking for \*specific\* criteria
    - ◆ example: where would be place a CPU and/or storage upgrade above baseline?



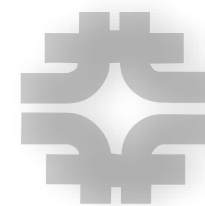
# Information Requested from Sites



- ◆ A self-assessment from each sites
  - ★ a narrative (3 pages maximum), see next slide
  - ★ a “facilities plan” addressing a deployment of T2 resources with a “fixed” yearly increase as a baseline, and outlining the opportunities for additions using “build to cost” at \$250k/yr
  - ★ a detailed budget breakdown for 3 years
- ◆ In order to have a common basis for projecting the cost of equipment at Tier 2 centers, we will provide a spreadsheet with typical cost estimates per HS06 and TB.
  - ★ These estimates include unit cost for “standard” servers - they do not include ancillary expenses like rack, power supplies, networking, and cooling.
- ◆ Use these numbers for your projected budget, if possible
  - ★ If different equipment will be used, please provide justification and basis.
  - ★ Use Moore's law doubling every 3 yrs for cost projection in future yrs



# “Standardized” cost estimate

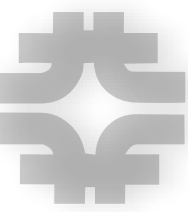


- ◆ example spreadsheet for cost of standardized components used by US Atlas in Nov 2010

Dell Pricing Examples									
NodeType	Cost	ProcType	JobSlots	TB (Useable;R6)	HS06Total	\$/HS06	Power(W)	Watts/HS06	\$/TB
R410	\$3,817	X5670	24		219	\$17.41	301	1.37	
R410	\$3,632	X5660	24		209	\$17.34	301	1.44	
R410	\$3,382	X5650	24		199	\$16.99	301	1.51	
R410	\$2,239	E5620	16		120	\$18.71	268	2.24	
R610	\$3,945	X5670	24		219	\$18.00	314	1.43	
R610	\$3,706	X5660	24		209	\$17.69	314	1.50	
R610	\$3,456	X5650	24		199	\$17.37	314	1.58	
R610	\$2,347	E5620	16		120	\$19.61	302	2.52	
M610	\$4,039	X5670	24		219	\$18.43	264	1.20	
C6100	\$15,202	X5670	96		877	\$17.34	1037	1.18	
C6100	\$11,391	E5620	64		479	\$23.79	857	1.79	
MD1000, SATA, 5.4K 2TB, no ctrl	\$3,743			26			462		
MD1000, SATA, 7.2K 2TB, no ctrl	\$4,532			26			462		
MD1200 with 12x2TB NL SAS 7.2K rpm + H800 3YWar	\$4,811			20			287		
MD1200 with 12x2TB NL SAS 7.2K rpm 3YWar	\$4,635			20			287		
R710 Storage, 48GB, E5620, 2x146 15K, red p/s, 5YrWar	\$3,282	E5620			120		319		
R710 Storage, 64GB, X5650, 2x146 15K, red p/s, 5YrWar	\$4,564	X5650			199		374		
R710 Storage, 64GB, X5670, 2x73 15K, red p/s, 5YrWar	\$5,039	X5670			219		374		
Storage Node Examples	Cost	Memory(GB)	Network	TB (Useable;R6)	HS06Total	\$/HS06	Power(W)	Watts/HS06	\$/TB
R710(Fast)+6*MD1200(7.2K NLSAS 2TB disks)	\$32,725.52	64	2x10GE(SFP+)	120			2096		\$273
R710(Default)+6*MD1200(7.2K NLSAS 2TB disks)	\$31,443.68	48	2x10GE(SFP+)	120			2041		\$262
R710(Default)+8*MD1200(7.2K NLSAS 2TB disks)	\$40,713.22	48	2x10GE(SFP+)	160			2615		\$254
R710(Default)+4*MD1000(5.4K SATA 2TB disks)	\$18,252.34	48	2x10GE(SFP+)	104			2167		\$176
R710(Default)+4*MD1000(7.2K SATA 2TB disks)	\$21,410.14	48	2x10GE(SFP+)	104			2167		\$206
R710(Default)+6*MD1000(5.4K SATA 2TB disks)	\$25,737.38	48	2x10GE(SFP+)	156			3091		\$165
R710(Default)+6*MD1000(7.2K SATA 2TB disks)	\$30,474.08	48	2x10GE(SFP+)	156			3091		\$195
R710(Fast)+6*MD1000(7.2K SATA 2TB disks)	\$31,755.92	64	2x10GE(SFP+)	156			3146		\$204



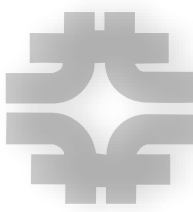
# Site “Self-Assessment”



- ◆ A max 3-pages narrative should include the following information:
  - ★ short history and current status of the Tier 2 center
  - ★ description of infrastructure to house Tier 2, including space, power, cooling
  - ★ management and staffing plans, including local experts not included in budget
  - ★ table showing leveraged local resources
  - ★ networking infrastructure for Tier 2 center
  - ★ projection of delivered resources (operational and usable by CMS) on April 1st of every year of the funding cycle
  - ★ provide any other details that may be helpful in evaluating your site
- ◆ Budget breakdown for coming 4 years
  - ★ Salaries and Fringe
  - ★ Infrastructure costs
  - ★ Cost of CPU
  - ★ Cost of Storage
  - ★ Overhead charged
  - ★ Anything else (networking, racks etc)



# Each Site to fill a Spreadsheet like



## ◆ Income

- ★ Income from ops pgm (yearly)
- ★ Contributed equipment funds
- ★ Funds provided by university
- ★ Funds provided for networking
- ★ Cost for Contributed Personnel
- ★ Cost for contrib space, power, cool
- ★ Contributed compute cycles
- ★ Income per site

## ◆ Cost

- ★ Personnel FTE (total)
- ★ Personnel FTE (funded by ops pgm)
- ★ Personnel Cost (total)
- ★ Personnel Cost (funded by ops pgm)
- ★ Computer Room, space, power, cool.
- ★ Space, power, cool. (funded by ops)
- ★ Indirects
- ★ Networking & Communications
- ★ Networking (funded by ops pgm)
- ★ WAN Network Bandwidth (10/2010)
- ★ Rack space in Data Center
- ★ Total Equipment Funds





# Beyond resources, there are criteria for “quality” of service



- ◆ most of these are being assessed by operations teams, such as
  - ★ CPU efficiency for MC jobs
  - ★ analysis throughput statistics
    - ◆ #events processed, #files accessed, #analysis jobs “terminated”
  - ★ relative resource allocation MC vs Analysis
  - ★ #users supported, /store/user accounting
  - ★ #tickets from CSPs, Sites, users, and response time
  - ★ which datasets used, pro-activeness of data managers, AAA overflow
  - ★ reliability of storage elements, how often do we lose files
  - ★ I/O performance metric
    - ◆ gives indication of performance of site setup
- ◆ these criteria need to be assessed very carefully
  - ★ we do not want to “judge” sites on criteria that are out of their control
  - ★ a lot of these metric depend on WMA finally interfacing to Dashboard
- ◆ We will probably assess these, review them and iterate with sites



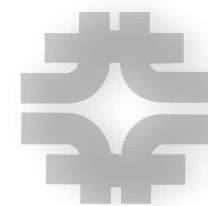
# This should start within a few weeks



- ◆ Will send a formal letter with instructions
  - ★ collect the information within a couple of weeks
- ◆ Will set up a small review committee
  - ★ reporting to the S&C manager
- ◆ In case of “outliers” will initiate “corrective actions” working with sites
  - ★ e.g. sites can plan for ramping up CPU in case they fell behind, etc
  - ★ these may require supplemental funds, if there is a case for optimizing the output of each and every site!
- ◆ Will run specific audits for “site performance”, led by operations group
  - ★ putting together the information for perusal
- ◆ probably will foresee an “appeal process”
  - ★ to make sure sites feel they’re fully engaged in the process



# We Request Your Help With This!



- ◆ We want that U.S. CMS and each site come out of this process even stronger, and that we can tell the success story of CMS Tier-2s better!
- ◆ We count on your constructive and pro-active participation
- ◆ comments, critiques, requests and any kind of input now or in the next days are very welcome!