

DOE/NSF of the
Software and Computing Efforts
of the U.S. LHC Research Program

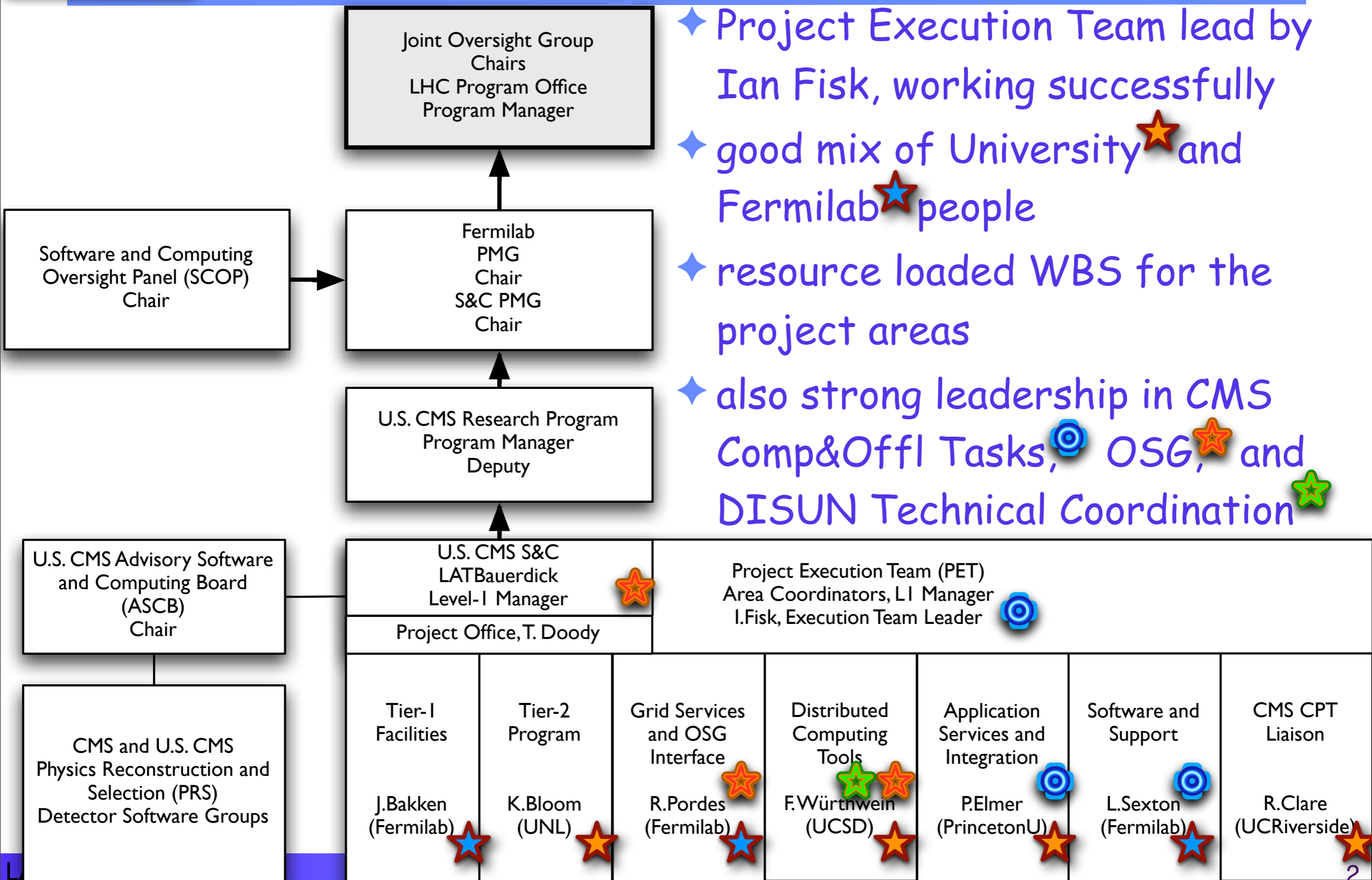
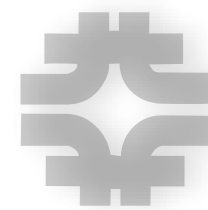
Introduction
U.S. CMS Software and Computing

LATBauerdick/Fermilab

University of Texas at Arlington, January 18, 2007



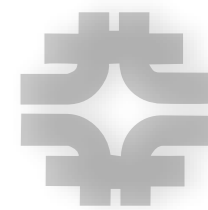
U.S. Organization and Management Team



- ◆ Project Execution Team lead by Ian Fisk, working successfully
- ◆ good mix of University★ and Fermilab★ people
- ◆ resource loaded WBS for the project areas
- ◆ also strong leadership in CMS Comp&Offl Tasks, OSG, and DISUN Technical Coordination



A quick review of 2006...



◆ Previous Review Feb-2006:

- ◆ prototypes of initial computing components being released, initial release of Framework/EDM, CMS computing integration program established as major driver, preparing for SC4 after SC3 done with problems

◆ This Review Jan-2007:

- ◆ P-TDR Vol1 and Vol2 published, with some 200M events produced and being analyzed for 85 notes, 75 of which were included in PTDR Vol.II
- ◆ new software CMSSW, Data Management, MC Systems operational
 - ◆ very successfully tested in CSA06 data challenge -- all metric met at $>\sim 25\%$ of 2008
 - ◆ 60M events produced, 200M events processed T0-T1-T2, Grid based analysis
 - ◆ 12th major CMSSW software release supporting the whole analysis chain, incl. calibration/alignment, HLT, pile-up simulation, ... being verified for physics
- ◆ Fermilab Tier-1 center plus all seven Tier-2/Tier-2C sites in the U.S.
 - ◆ commissioned at $\sim 30-50\%$, strong partner with ~ 25 validated worldwide CMS sites
 - ◆ OSG is operational, interoperability between OSG and EGEE is a reality,
- ◆ CMS transitioning to “computing, offline and physics” organization



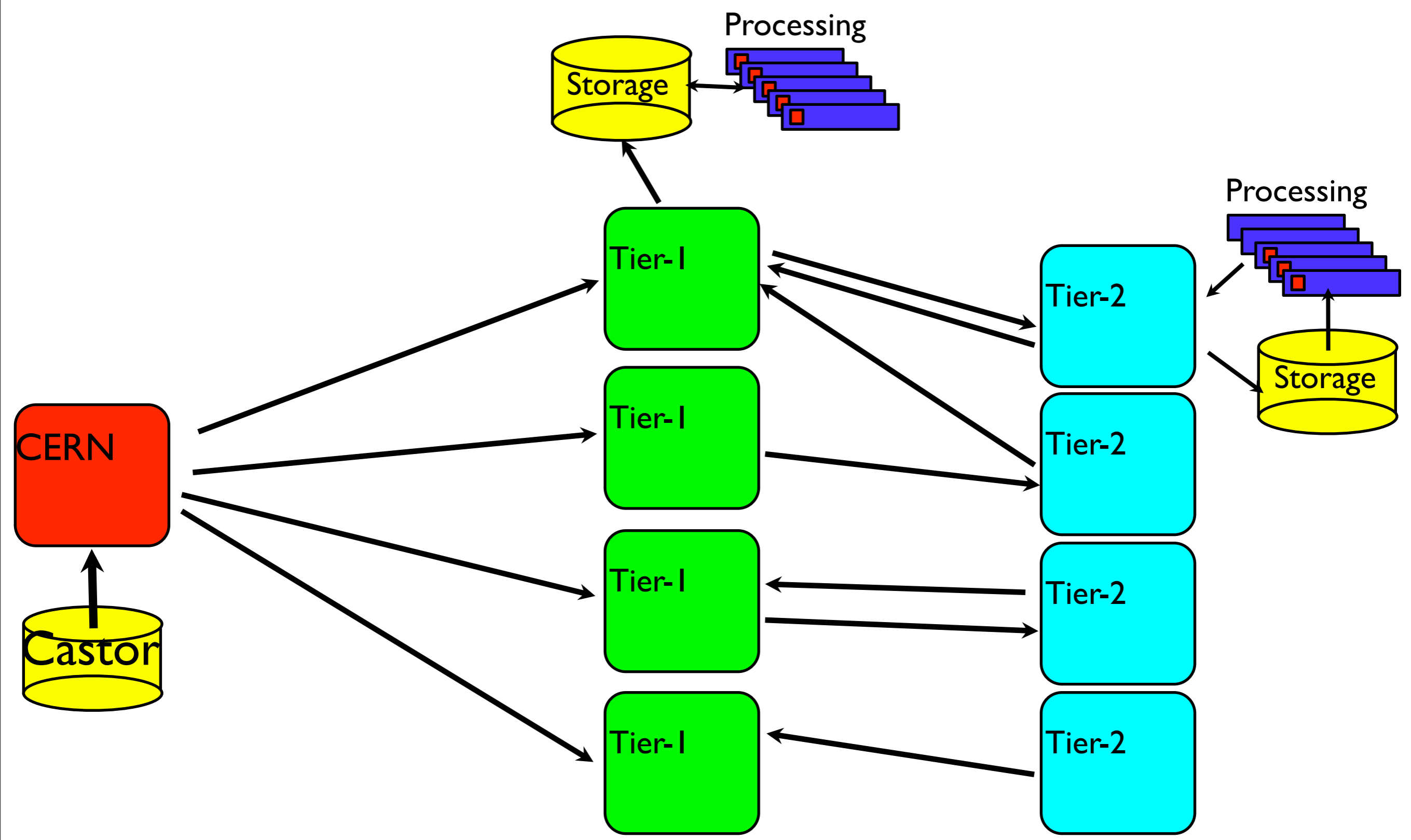
CMS Offline/Computing Successful



- ◆ Software project successfully evolved into offline organization
 - ◆ Defined procedures for release integration, building & distribution.
 - ◆ release tools and work flows, validation, schedule, structure of code repository
 - ◆ developed tools set: HyperNews, Nightly Builds, Tag collector, WorkBook, ...
 - ◆ Control centrally all interfaces and data formats
 - ◆ Uniform touch and feel, consistent definition of data tiers, etc.
 - ◆ Guide specific algorithmic implementations working with detector components
 - ◆ factorization of reconstruction algorithms, coherence L1/HLT/offline, etc
- ◆ Focussed offline and computing on CMS integration&commissioning
 - ◆ Computing Integration driving computing - focus on useable scalable system
 - ◆ Validation software packages and validation procedures - crucial for start-up prep
 - ◆ Strong connections with various detector groups - key for commissioning.
- ◆ Achieved all milestones → success proven in CSA06 test
- ◆ CMS offline and computing well positioned for efficient startup toward Physics!

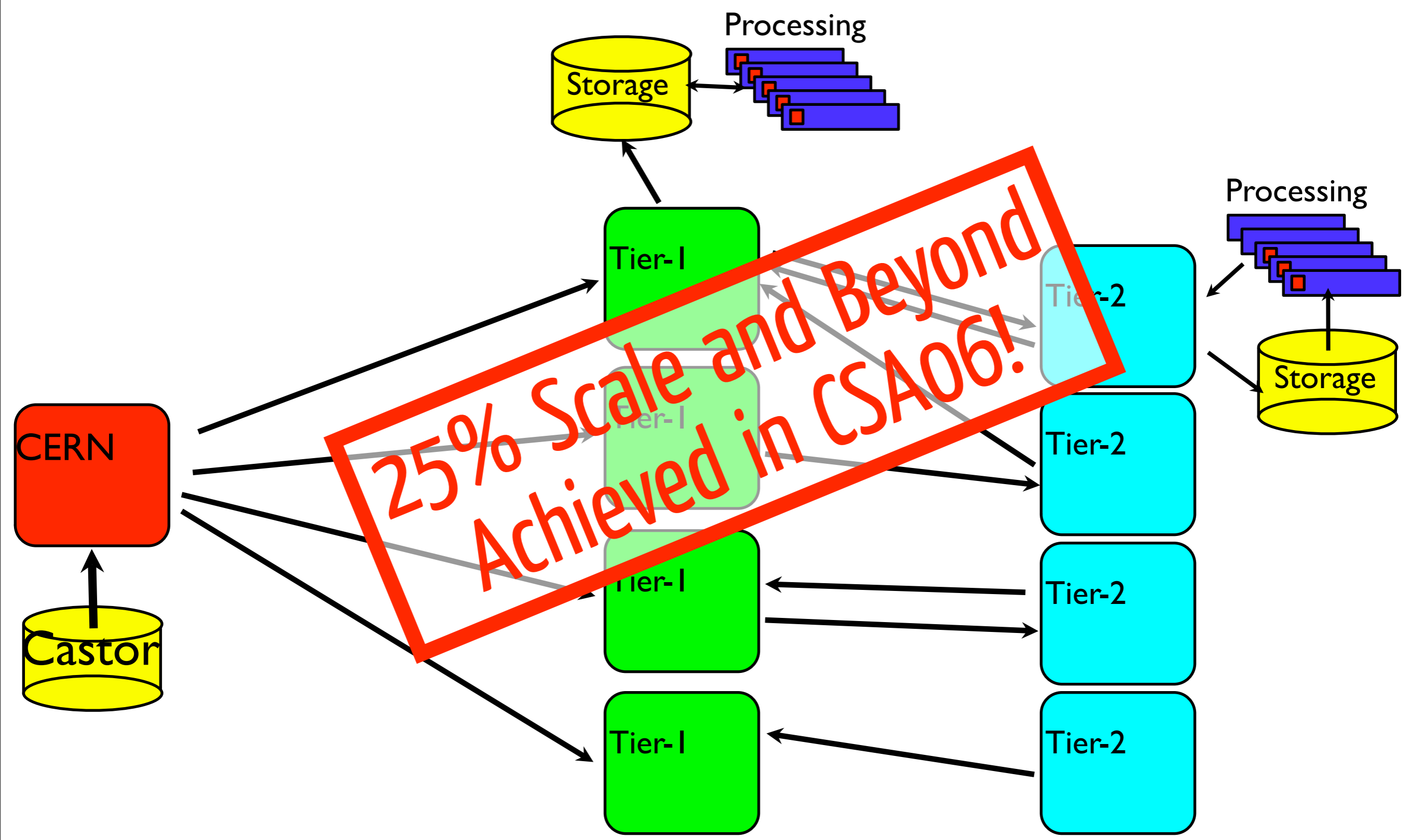


Computing Model Tested "end" to "end" during SC4 and CSA06





Computing Model Tested "end" to "end" during SC4 and CSA06





Large Aggregate Computing Resources Required—and Pledged!



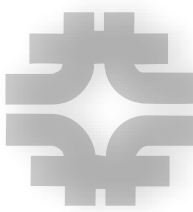
◆ for 2008 CMS requests ~ 35 MSI2k CPU, 12 PB disk, 18 PB tape

	2008	request	pledged	U.S.
T0+CAF	CPU [MSI2k]	7.7	7.4	
	Disk [PB]	1.6	1.6	
	Tape [PB]	5.1	3.8	
T1s	CPU [MSI2k]	12.4	11.6	4.3
	Disk [PB]	5.6	5.5	2.0
	Tape [PB]	13.1	9.6	4.7
T2s	CPU [MSI2k]	15.2	18.3	7.0
	Disk [PB]	4.2	4.8	1.4

- ◆ Tier-0, Tier-1, Tier-2, CAFs, Tier-3s all are essential for success
- ◆ CERN plans to build Tier-0 and CAF at the required level



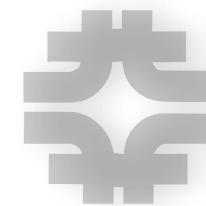
Tier-2/DISUN, OSG Successful



- ◆ Tier-2 and DISUN Tier-2C setup to play their part in CMS model
 - ★ Seven Tier-2s ramping to ~50% of 2008 capacity
 - ★ data gets transferred to sites, physicists can find data, run jobs
- ◆ OSG has strong focus on preparing Grid for LHC startup
 - ★ OSG facility active and supportive, solving real-world problems
- ◆ Strategy for including Univ. computing facility relies on OSG
 - Tier-3 computing has started, OSG and DISUN are enabler
 - ★ several sites already active, profiting on physics side (e.g. testbeam etc)
 - ★ OSG is providing what sites need to succeed on the Grid
 - ★ setting up OSG enables sites to get access to data and resources of Tier-2/C and Tier-1
 - ★ CMS/DISUN provides help with software installs, CMS specifics
 - ★ first U.S. CMS Tier-3 workshop at OSG meeting March
 - ★ teamwork between Tier-1, Tier-2s, Universities → DISUN took an important role for enabling Universities to run their science on the Grid



Cost and Funding

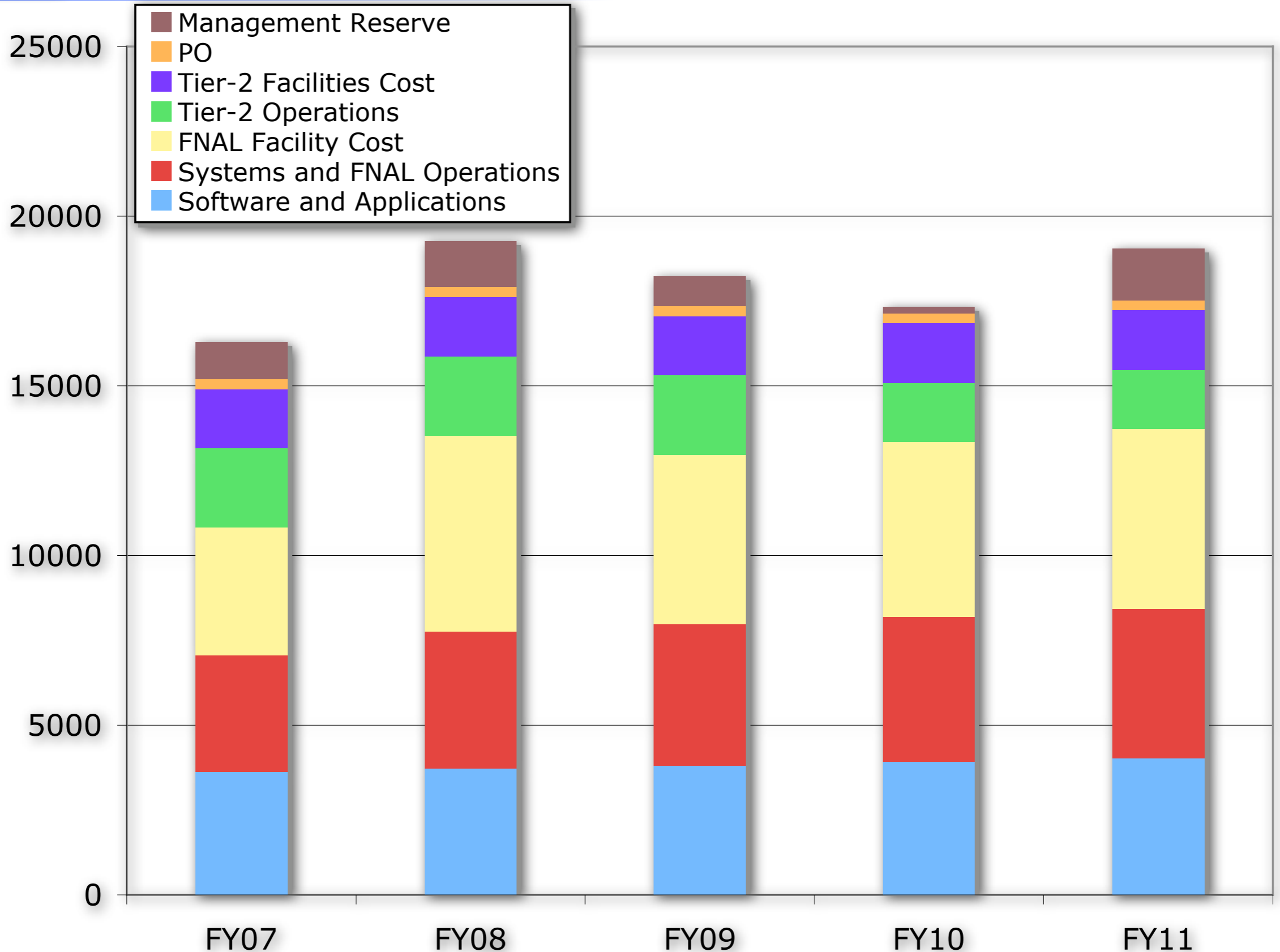
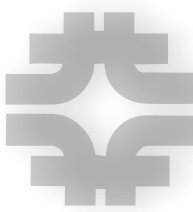


◆ Change in Tier-1/LPC-CAF cost developments, push part of 2008 facility into FY08

	FY07	FY08	FY09	FY10	FY11
S&C budgeted total	\$16,288k	\$19,255k	\$18,232k	\$17,318k	\$19,032k
DOE allocated	\$10,900k	\$12,000k	\$12,000k	\$12,000k	\$12,000k
NSF allocated (incl. DISUN)	\$5,230k	\$5,230k	\$5,230k	\$4,830k	\$4,830k
NSF Research Program	\$3,230k	\$3,230k	\$3,230k	\$4,830k	\$4,830k
NSF DISUN	\$2,000k	\$2,000k	\$2,000k		
carryover from previous FY	\$3,673k	\$3,515k	\$1,490k	\$488k	\$0k
		Carry FY06 \$\$ into FY08			
Total Labor Costs	\$9,385k	\$10,087k	\$10,313k	\$9,944k	\$10,176k
Software and Applications	\$3,625k	\$3,724k	\$3,821k	\$3,921k	\$4,023k
Systems and FNAL Operations	\$3,432k	\$4,029k	\$4,149k	\$4,274k	\$4,402k
Tier-2 Operations	\$2,328k	\$2,335k	\$2,343k	\$1,750k	\$1,750k
Total Facility Costs	\$5,518k	\$7,529k	\$6,740k	\$6,900k	\$7,055k
FNAL Facility Cost	\$3,778k	\$5,789k	\$5,000k	\$5,150k	\$5,305k
Tier-2 Facilities Cost	\$1,740k	\$1,740k	\$1,740k	\$1,750k	\$1,750k
Management and Reserve	\$1,385k	\$1,639k	\$1,179k	\$474k	\$1,802k
PO	\$288k	\$295k	\$302k	\$279k	\$286k
Reserve	\$1,097k	\$1,344k	\$877k	\$195k	\$1,516k
			M.R. below 10%		

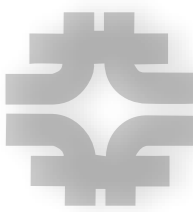


Cost Breakdown Plot





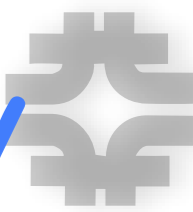
Tier-1 Costs and Deployment Plan



- ◆ "2008 facility": Tier-1 + LPC-CAF planned to finish with FY07
 - ★ at 7.3 MSI2k, 2.5PB disk, 4.7PB tape (no change)
- ◆ with slow LHC startup, push last phase procurements to FY08
 - ★ facility still ready for 2007 running and start of physics data spring 2008
- ◆ delays of procurement help with slow technology developments
 - ★ costs don't drop as fast, but delay compensates total costs to ~\$200k
 - ◆ reasons are thresholds of 1TB disk, 4-core CPU later than anticipated
 - ★ Crossing of FY boundary helps
- ◆ in new plan, facility budget is balanced, 10% M.R. re-instated
- ◆ Beyond 2008 we go into the "replacement mode"
 - ★ studied cost and capacity gains by replacing outdated equipment
 - ★ this allows Tier-1/LPC-CAF to follow expected resource needs of CMS
 - ★ budgeted flat profile at \$5M FY09 dollars (incl. infrastructure/media)
 - ◆ details of estimates and capacities see facility session



Focus on Usability and Getting Ready



- ◆ Agenda oriented toward demonstrating how we get ready for operation and physicists using S&C for their physics analysis
- ◆ 10:45 -- 12:15 CMS Plenary Session I (1.5h)
 - ◆ Introduction -- LATBauerdick (15+5)
 - ◆ CMS Offline and Computing Status and Plans -- Patty McBride (15+5)
 - ◆ US CMS S&C Project Status and Plans -- Ian Fisk (40+5)
- ◆ 13:15 -- 14:15 CMS Plenary II (1h)
 - ◆ CSA06 -- using the (U.S.) CMS Computing and Software Systems -- Darin Acosta
 - ◆ Readiness and Operation of Grid Infrastructure for US CMS -- Ruth Pordes
 - ◆ Physics Readiness -- Sarah Eno
- ◆ Parallel Sessions (2h)
 - ◆ Facilities and Grids (includes T2s, DISUN, WAN) -- Jon Bakken et al
 - ◆ Software and Support -- Liz Sexton-Kennedy et al
 - ◆ Management -- LATBauerdick et al
- ◆ Usability Demo
 - ◆ during first 30mins of parallel sessions -- Oliver Gutsche