Behind the scenes of Big Science

Amber Boehnlein
Department of Energy
And
Fermi National Accelerator Laboratory

What makes Big Science Big?

- The scientific questions being asked and answered
- The complexity of the instruments and instrumentation
- The amount of data generated and accessed
- The number of people involved in the research: the collection of the data, the analysis and interpretation of the data
  - In the case of some disciplines such as astronomy, this can include the general public
- The timescales
- The cost to the taxpayer and the amount of oversight required.

High Energy Physics experiments can serve as examples to hopefully illustrate general principles of Big Science Projects(BSP)

The Large Hadron Collider (LHC)

- It is located in Switzerland at the European Organization For Nuclear Research (CERN)

- Planning began in the late '80s
  - Concurrent with the Superconducting Supercollider
- LHC collides two beams of protons at four interactions regions.
  - Steady operations began in earlier this year (2010)
  - The machine will require repairs in 2012 in order to reach design energy.
- There is a particle detector at each interaction region
  - Each particle detector is designed, constructed and operated by an international collaboration of O(1000) scientists
  - The detectors are scientific facilities capable of tackling a broad range of physics questions.
- The previous generation of experiments are called CDF & DO (Fermilab) and BaBar (SLAC)

CMS Phase 1 Upgrade Technical Proposal

- Preliminary draft given to LHCC for initial comments
  - Only about half the chapters were declared ready for even a preliminary reading
    - But they are the main projects and main resource drivers
      - Muons
      - HCAL
      - Pixels
    - Main chapter that was incomplete is Trigger, which will be done in a few days
    - chapter on Beam Monitoring Instrumentation (not US and not a big cost driver)
    - There will also be an appendix on Phase 2 R&D
- Very similar to the document DOE/NSF requested in November of 2009 but for all CMS (not just US)
  - Also includes all improvements and repairs foreseen for the detector, not just luminosity-related upgrades
- Expect to finish document in early November after October Upgrade Workshop
- Daniela and I would like to brief DOE/NSF on this

CMS Phase 1 Upgrade Technical Proposal

- Preliminary draft given to LHCC for initial comments
  - Only about half the chapters were declared ready for even a preliminary reading
    - But they are the main projects and main resource drivers
      - Muons
      - HCAL
      - Pixels
    - Main chapter that was incomplete is Trigger, which will be done in a few days
    - chapter on Beam Monitoring Instrumentation (not US and not a big cost driver)
    - There will also be an appendix on Phase 2 R&D
- Very similar to the document DOE/NSF requested in November of 2009 but for all CMS (not just US)
  - Also includes all improvements and repairs foreseen for the detector, not just luminosity-related upgrades
- Expect to finish document in early November after October Upgrade Workshop
- Daniela and I would like to brief DOE/NSF on this

Big Science Projects are typically global endeavors, funded by multiple agencies

Big Science Projects have an identity, an organization, a personality and are made up of individuals

Big Science Projects have phases and long timelines.

- **Scientific and Technical R&D**
- **Design**
- **Construction**
- **Operations**
- **Scientific Research**

These phases co-exist in time

The projects have history and constraints are imposed by that history

Big Science Projects are process-intensive. They are well documented, required constant communication and subject to intense scrutiny in peer and financial reviews.

CMS Phase 1 Upgrade Technical Proposal

- Preliminary draft given to LHCC for initial comments
  - Only about half the chapters were declared ready for even a preliminary reading
    - But they are the main projects and main resource drivers
      - Muons
      - HCAL
      - Pixels
    - Main chapter that was incomplete is Trigger, which will be done in a few days
    - chapter on Beam Monitoring Instrumentation (not US and not a big cost driver)
    - There will also be an appendix on Phase 2 R&D
- Very similar to the document DOE/NSF requested in November of 2009 but for all CMS (not just US)
  - Also includes all improvements and repairs foreseen for the detector, not just luminosity-related upgrades
- Expect to finish document in early November after October Upgrade Workshop
- Daniela and I would like to brief DOE/NSF on this

Big Science Projects live within constrained funding

Big Science Projects have many components and are driven by internal and external requirements and factors

Big Science Projects are always a work in progress

Big Science Collaborations tend towards bottom up

We love color for emphasis! We also have a common vocabulary and context

So where's the Science?

- Prioritization of the science comes before the detailed design, construction and operation of any research facility or instrument
- Decadal Survey is a recent example

Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond

Authors:
Committee on Earth Science and Applications from Space: A Community Assessment and Strategy for the Future, National Research Council
Authoring Organizations
Description:

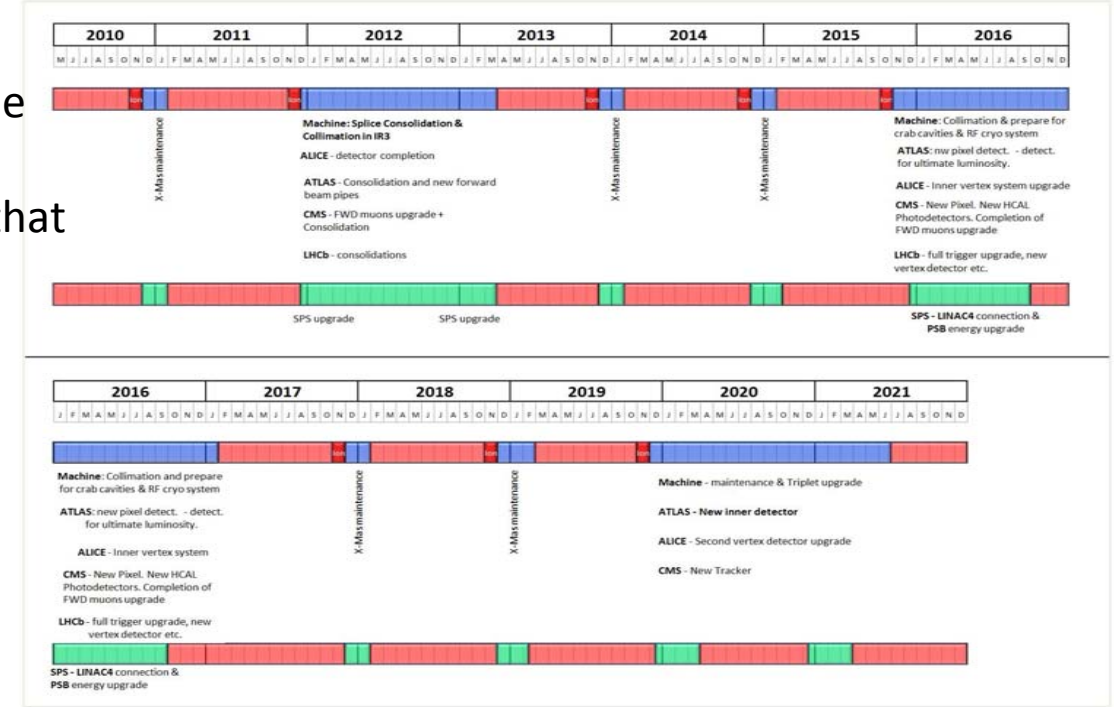
Natural and human-induced changes in Earth's interior, land surface, biosphere, atmosphere, and oceans affect all aspects of life. Understanding these changes requires a range of observations acquired from land-, sea-, air-, and space-based platforms. To assist NASA, NOAA, USGS….

Phases and long timelines.

- The LHC at CERN is twenty years into forty year program of work

This is the accelerator schedule that is driving the detector upgrade proposals-a process that began before the LHC had circulating beam

The 10 year technical Plan



Implications

- The cutting edge nature of the research insures innovation and pushing technology
- The peer-review and funding processes insure inherent conservatism (including fiscal conservatism)
- Endeavors built around the long haul imply that dealing with obsolescence (or dealing with the consequences of not dealing with it) are a major fact of life
- Component obsolescence implies component wise replacements
  - The time lag between design and operation can lead to obsolescence of electronics before an instrument takes data
  - A significant amount of an operations budget is used to buy or build spares, or to eventually to replace components.
  - Computing and software not exempt.
- Decoupling of construction funding and operations funding can have interesting consequences
- Long term collaborative/competitive environment rely on trust relationships
- Timescales are long; Memories are longer
  - A cat that walks on a hot stove will never walk on another hot stove; he will also never walk on a cold one.

The Killer Requirement

- A challenge for BSP is that they typically have anomalous requirements-it's the nature of the business
  - The instruments often operate in extreme conditions-space, the ocean, Antarctic ice, particle accelerators...
- The Killer requirements can be relatively mundane
  - Facebook would make a splendid online logbook for HEP experiments except that most experiments are required to be able to take data without network connectivity.
- There are many potential killer requirements for using a commercial products
  - Scalability
  - Longevity
  - Recent example from CERN attempt to replace a custom application with a relatively cheap commercial application: Where's the linux client?

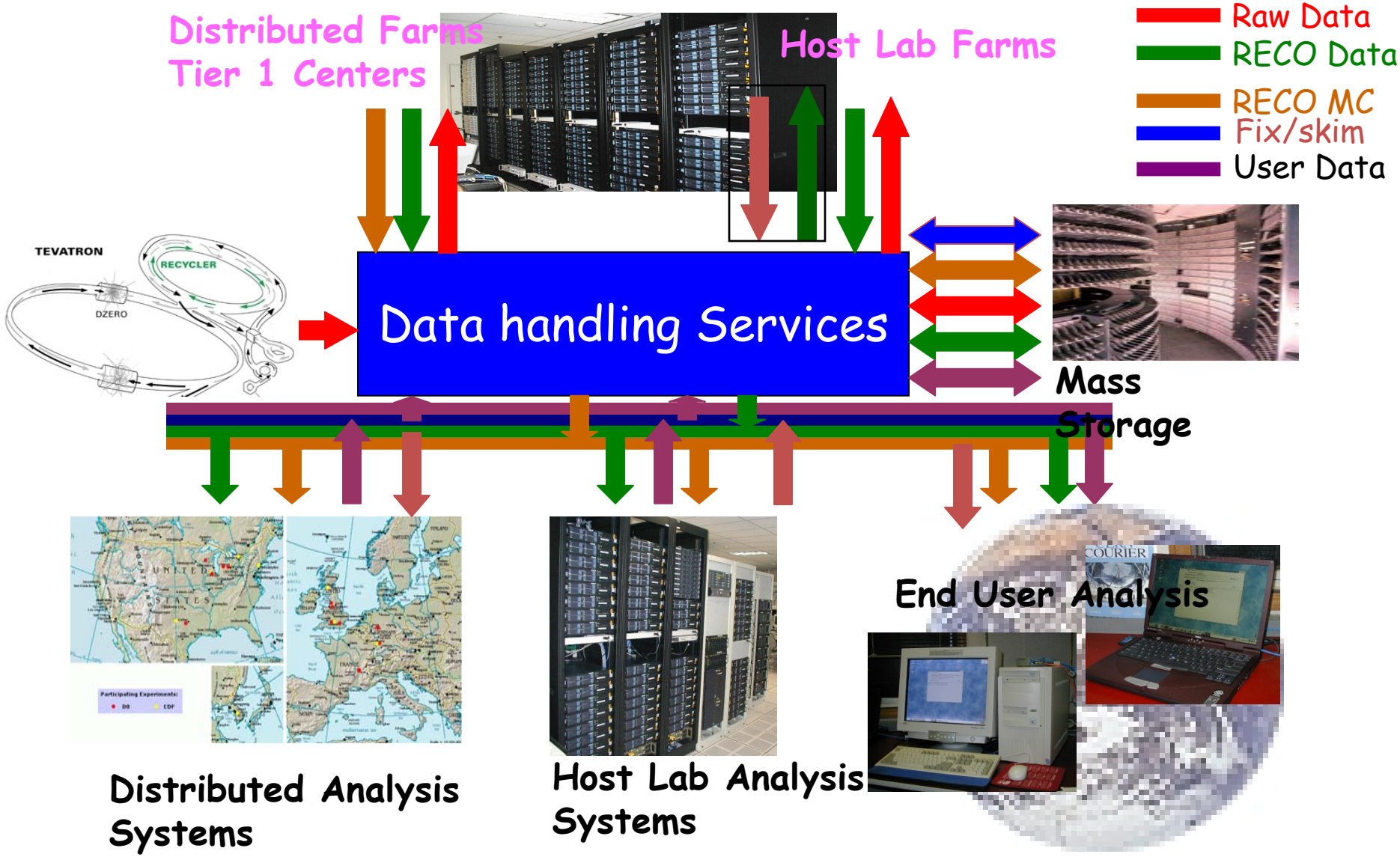
Planning for Computing in HEP

- The technical design phase for an instrument usually includes the planning for software and computing planning.
  - Data centers are scoped as part of the program of work.
- The planning starts with vision of what a "modern" computing and analysis system should do and how users should interact with the data.
  - The point of view of the typical computing model is data-centric
    - How big is the data, what operations will be performed on it, where will it reside, who can access it.
    - Data coming from an HEP experiment is self-integrated—all event data is read out of the detector at the same time and time stamped.
    - Conditions data which is outside of the event stream are logged in databases, times tamped and associated with event data
  - New techniques are usually required meet the anticipated needs within in anticipated budgets
  - Have to account for a wide dynamic range of abilities and knowledge in the user community

Workflow Terminology

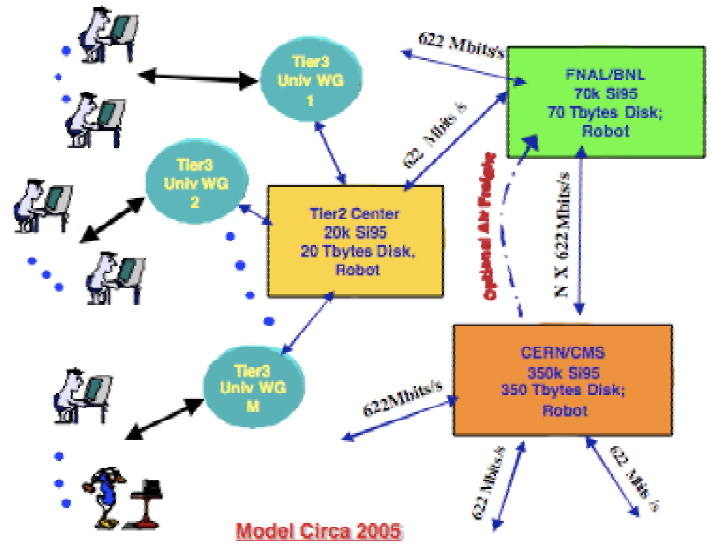
- Reconstruction is the process of applying calibrations and algorithms to digitized detector (raw) data to make physics objects
  - Complicated—data intensive; accessing calibration
  - Re-reconstruction is doing it again because of significant improvements in alignment, calibration and algorithms
  - "Fixing" is a limited re-reconstruction
  - Data "tiers" refers to the output contents/format " "raw", "dst", "tmb"
- Monte Carlo Data (MC) involves simulated physics processes processed through a detector representation and then reconstructed.
  - Relatively straightforward although time intensive
- Activities are performed in common for the collaboration, are centrally managed and are called "production" activities
- Analysis activities are performed by individuals for themselves or small groups of people. Analysis activities are not managed
- A set of parameters determine the computing resources needed
  - the size of the data tiers
  - the amount of live time for data collection
  - Reconstruction time

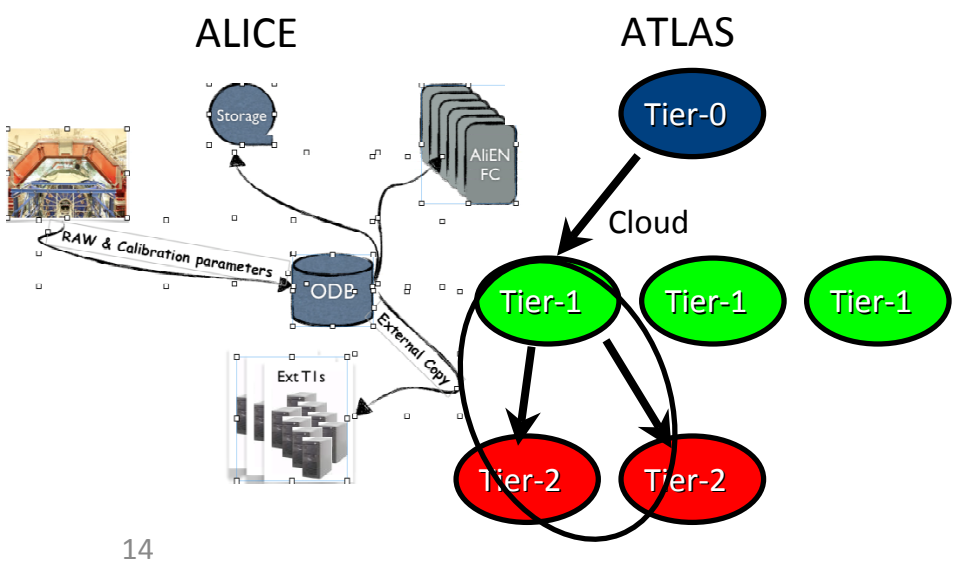
Computing Model

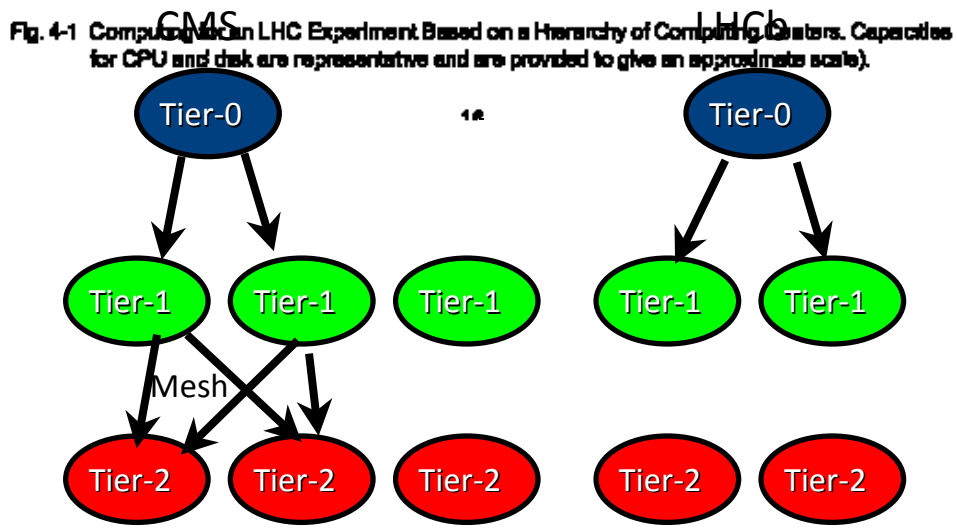


LHC Architectures

- To greater and lesser extents LHC Computing model are based on the MONARC model (MOdels of Networked Analysis at Regional Centers)
  - Developed more than a decade ago
  - Foresaw Tiered Computing Facilities to meet the needs of the LHC Experiments







Planning example: DO Experiment 1997

D0 Vital Statistics	
	1997(projections)
Peak (Average) Data Rate(Hz)	50(20)
Events Collected	600M/year
Raw Data Size (kbytes/event)	250
Reconstructed Data Size (kbytes/event)	100 (5)
User format (kbytes/event)	1
Tape storage	280 TB/year
Tape Reads/writes (weekly)	
Analysis/cache disk	7TB/year
Reconstruction Time (Ghz-sec/event)	2.00
Monte Carlo Chain (GHz-sec/event)	150
user analysis times (Ghz-sec/event)	?
user analysis weekly reads	?
Primary Reconstruction farm size (THz)	0.6
Central Analysis farm size (GHz)	0.6
Remote resources(GHz)	?

In "then year" costs, much computing was a formidable challenge!

Prior to commodity systems not in general use.

Had to generate MC data
Using distributed Resources

Anticipating LHC-Grid/distributed Computing

DO as a function of time

D0 Vital Statistics			
	1997(projections)	2006	2010
Peak (Average) Data Rate(Hz)	50(20)	100(35)	120(80)
Events Collected	600M/year	2 B	2 B
Raw Data Size (kbytes/event)	250	250	270
Reconstructed Data Size (kbytes/event)	100 (5)	80	100
User format (kbytes/event)	1	80	40
Tape storage	280 TB/year	1.6 pb on tape	5.8 pb on tape
Tape Reads/writes (weekly)		30TB/7TB	110TB/40TB
Analysis/cache disk	7TB/year	220 TB	750 TB
Reconstruction Time (Ghz-sec/event)	2.00	50 (120)	120
Monte Carlo Chain (GHz-sec/event)	150	240	240
user analysis times (Ghz-sec/event)	?	1	wide variation
user analysis weekly reads	?	8B events	18B events
Primary Reconstruction farm size (THz)	0.6	2.4 THz	7.5 THz
Central Analysis farm size (GHz)	0.6	2.2 THz	13.3 THz
Remote resources(GHz)	?	~ 2.5 THz(grid)	?

'97 planning dramatically underestimated the needs in many key areas.
Moore's law and a delayed start of the Tevatron run == Good fortune.
All systems scaled up and have proven operationally robust.

CMS Input Parameters

- Input Parameters are similar to previous years.
  - Some items evolve with experience and were pushed from 2010 to 2011
    - Primary Dataset Overlap
  - Some items like RAW event size were allowed to change as expected in 2011
- AOD has increased based on what's in the AOD currently
  - Unlikely to reduce before the transition from RECO to AOD

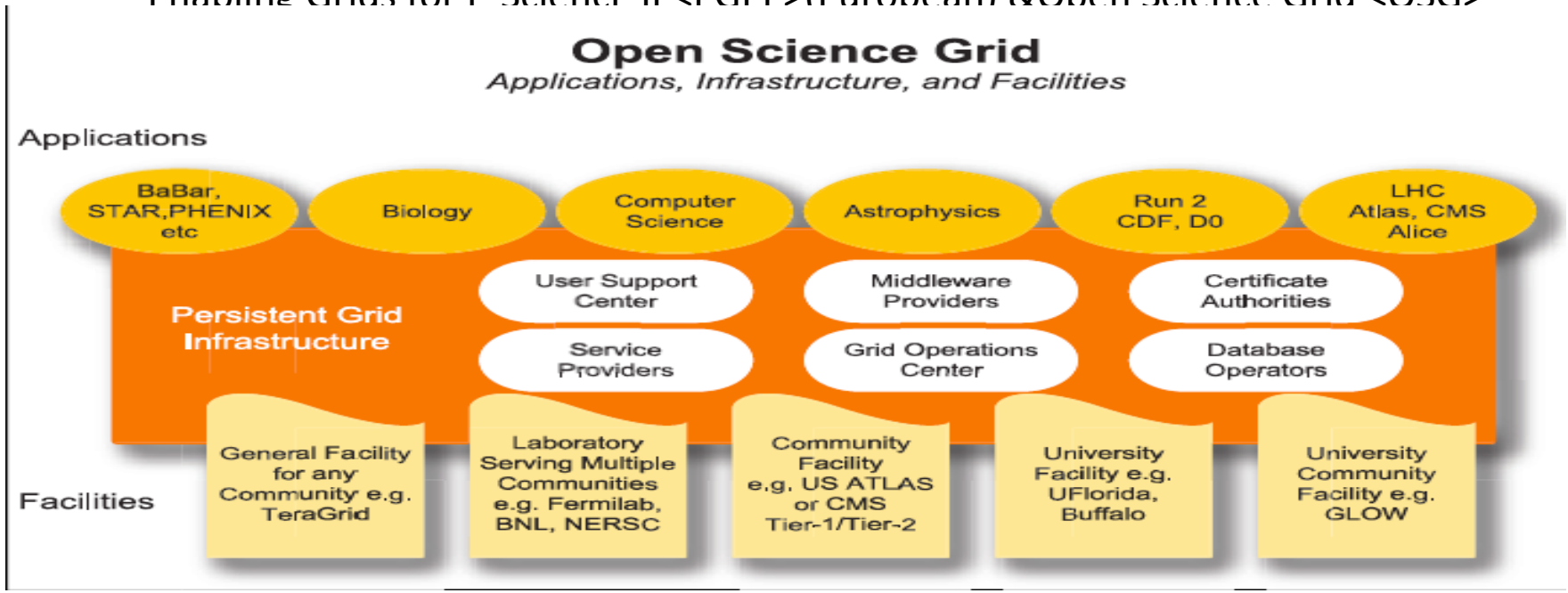
Parameter	Year		
	2009	2010	2011
Trigger(Hz)	300Hz	300Hz	300Hz
PDS Overlap Factor	1.4	1.4	1.25
Tier-0 Recovery	0.5	0.75	0.75
RAW Data (MB)	1.5	1.5	1.0
RAW MC (MB)	2.0	1.5	1.5
RECO Data (MB)	0.5	0.5	0.5
RECO MC (MB)	0.6	0.6	0.6
AOD Data(MB)	0.2	0.2	0.2
AOD MC (MB)	0.25	0.25	0.25
RECO(HS06)	100	100	125
SIM(HS06)	700	700	900
Days of Re-RECO	30	60	90

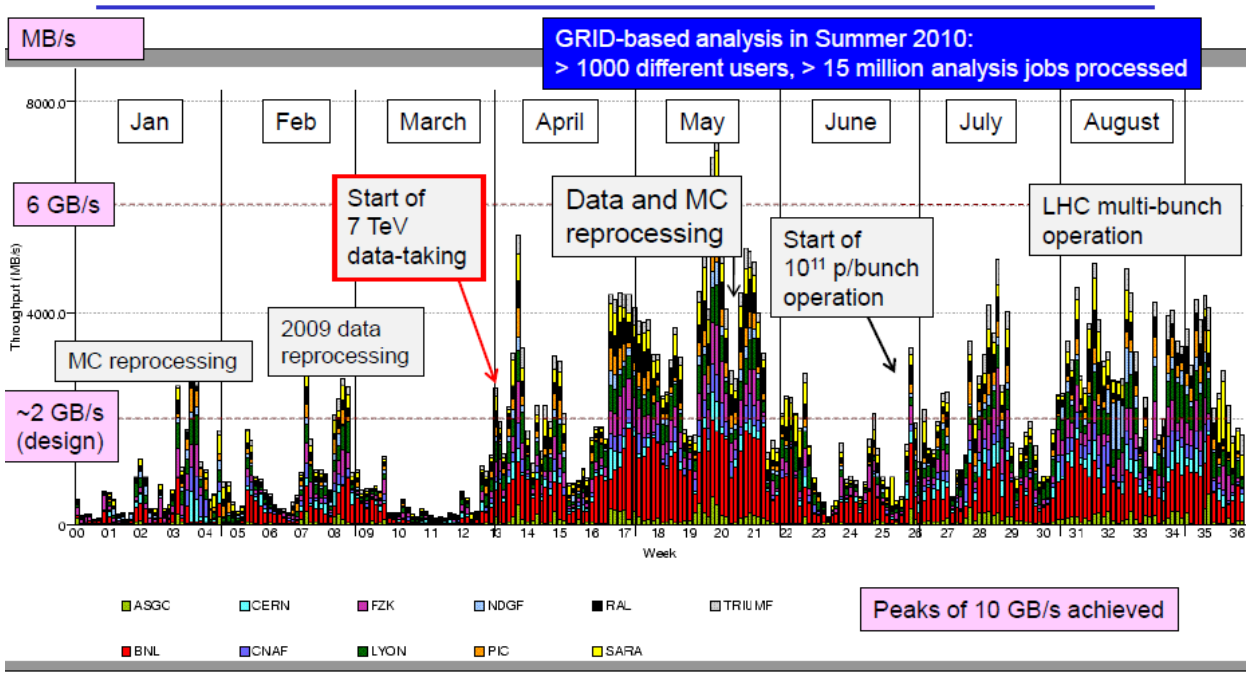
How do technical decisions get made?

- Part of the planning process is looking at the technical requirements necessary to meet the data volume needs and how well those solutions are meeting the needs.
  - Hardware roadmaps
  - Data management and access technologies, process management, security.
- The planning and evaluation is on-going process over the life of an experiment
- Survey of available commercial and labware solutions
  - R&D, prototyping
  - Internal collaboration decision processes and external reviews
  - Generally detailed criteria on all technical aspects including scalability and long term maintenance costs
  - Evaluation of the costs of change and the work required
    - Migrating small amounts of complicated data might take more effort than large amounts of straightforward data.
- Common alternative path-the voting of feet
  - Bottoms up alternatives to the officially sanctioned solutions pop up all time
  - Generally homegrown with all the issues that can imply
- As the experiments become more expensive, there has been a push toward joint projects in areas of obvious commonality

Grid Computing for LHC

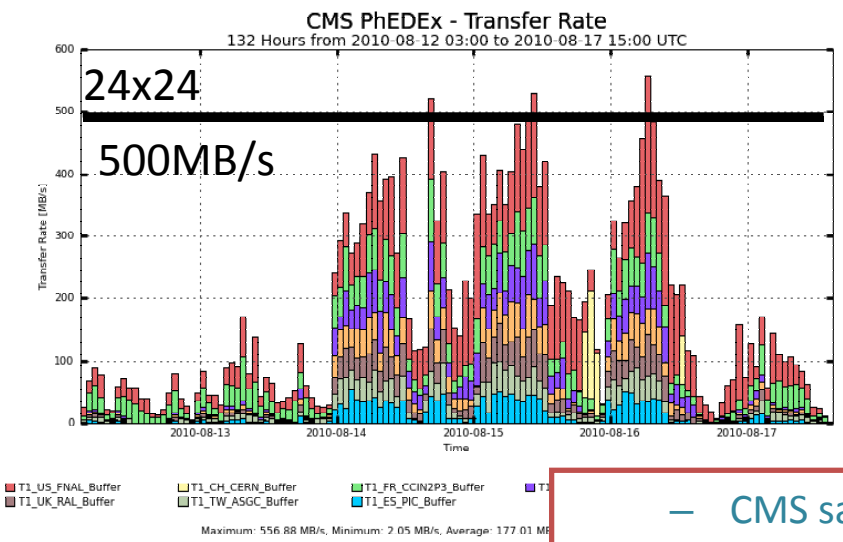
- Based on the long range planning, large hardware installations are required for LHC computing
- For political, practical and financial reasons, the hardware needed to be distributed
- Physical resources located at Cern, at 11 Laboratory computing centers around the world and at 0(150) collaborating universities
- Ten years of Grid projects have linked the physical resources in a robust and operationally sustainable way
- Enabling Grids for E-SciencE II <EGEE>(European) &Open Science Grid <OSG>

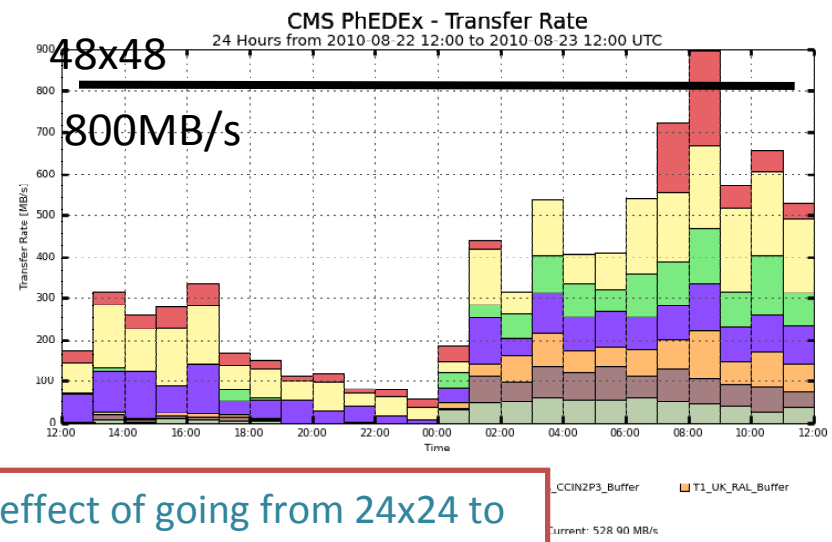




ATLAS Throughput

For all experiments: early data has been available for analysis within hours of data taking





– CMS saw effect of going from 24x24 to 48x48 bunches: same trigger

- Operational problems at sites requiring an incident report (i.e. That have a user-visible service degradation), remain largest stability issue.
  - ~5 -6 per month; some few hours, some more extended
  - Databases, power/cooling, hardware failures, etc.
  - Not grid middleware
- Last 2 weeks have had a security problem – not grid specific - requiring urgent reactions from sites to protect themselves
  - Net result was significant reduction in overall capacity
  - This is unavoidable at some level
- Have to become more resilient to these;
  - evolution of computing models (cannot assume all sites will always be available)
  - Ensure lessons are learned by other sites

Data Curation in HEP

- Growing interest in long term curation of HEP data.
  - A number of experiments have ended data collection recently.
- The data collected by an HEP experiment is expensive to produce, can be impossible to reproduce, the data already resides at national laboratories. Curation is a minor incremental cost.
  - So why don't we do it?
- HEP collaborations tend to be very proprietary about "their" data-even after they think it is no use to anyone else.
- Technical issues
  - The human expertise required to make sense of the data
  - Lack of "management" of analysis
  - A large software suite remains necessary (MC)

Opportunities for Collaboration in Big Science Computing

- There is considerable intellectual similarity between the problems faced in different domains— within a domain, across the scientific domains and with industry.
  - Success of the Grid computing projects has increased confidence that HEP experiments can collaborate and yet maintain autonomy.
- There is frequent and fruitful collaboration between vendors and those working in HEP and there is definitely opportunities for more.
  - Venues such as this provide cross-pollination that can lead to collaboration
  - Understanding needs and culture of the community

Summary

- Big Science projects have long timescales, involve long term relationships and have funding constraints
- They are in the public eye and under tremendous pressure to succeed.
  - Scientific success and a public perception of success are not always the same.
- There are many areas of commonality in large scale computing projects across the spectrum of public and private projects.
- With the LHC entering operations AND planning for upgrades, there are opportunities for collaboration