CMS Software and Computing Status

DOE/NSF Baseline Review of
US CMS Software and Computing
BNL November 14, 2000

David Stickland, Princeton University
(Deputy PM, CMS Core SW&C)

Status Overview

- Functional Prototype stage being completed now
- We are <u>using</u> these SW prototypes for simulation, reconstruction, visualization...
- Two major "productions" have been completed, simulating and reconstructing ~2 million events at 10^{34} LHC luminosity
- A third worldwide production is currently underway
  - ~ 5 million events
  - Simulation and Reconstruction with Objectivity DB population happening worldwide with GLOBUS based replication between DB producing sites
- Tier-1 and Tier-2 regional centers are being defined and prototyped

Deliverables for each of the Software Functional Prototypes

- Documented set of use-cases / scenarios / requirements
  - These are not (yet) very formal
- Suite of software packages / prototypes
  - Public releases of associated software
  - Documentation of components (user-level and reference)
- Software infrastructure
  - Public software repository
  - Build / release / distribution / documentation systems
  - Multi-platform support both centrally and in remote institutes, …etc...
- Proposal for a baseline
  - recommendations on main design and technology choices
- Proposed project plan for next phase

Simulation

- Current simulation by our GEANT3 product: CMSIM
- Migration underway to GEANT4
  - We require GEANT4 for its Physics processes and OO implementation
- FAMOS project underway for Fast Simulation
  - Will be vital to complete many topics in the Physics TDR
  - The FAMOS project is preparing for its first release this autumn

OSCAR (CMS G4 Simulation)

- OSCAR project for CMS implementation of GEANT4
  - Proof of concept reached in July 1998
    - Built barrel geometry
  - Functional Prototype: originally due Dec 1999, slipped by 6 months
    - Delayed due to manpower problems
      - Persistent hits not yet implemented
      - Geometries for most sub-detectors ready
      - Many performance tune-ups completed
    - June 2000, milestone passed but with severely limited functionality

CMS recognizes that this project will become a critical path item and is addressing these issues now

ORCA

Has been major focus of SW activity for last two years

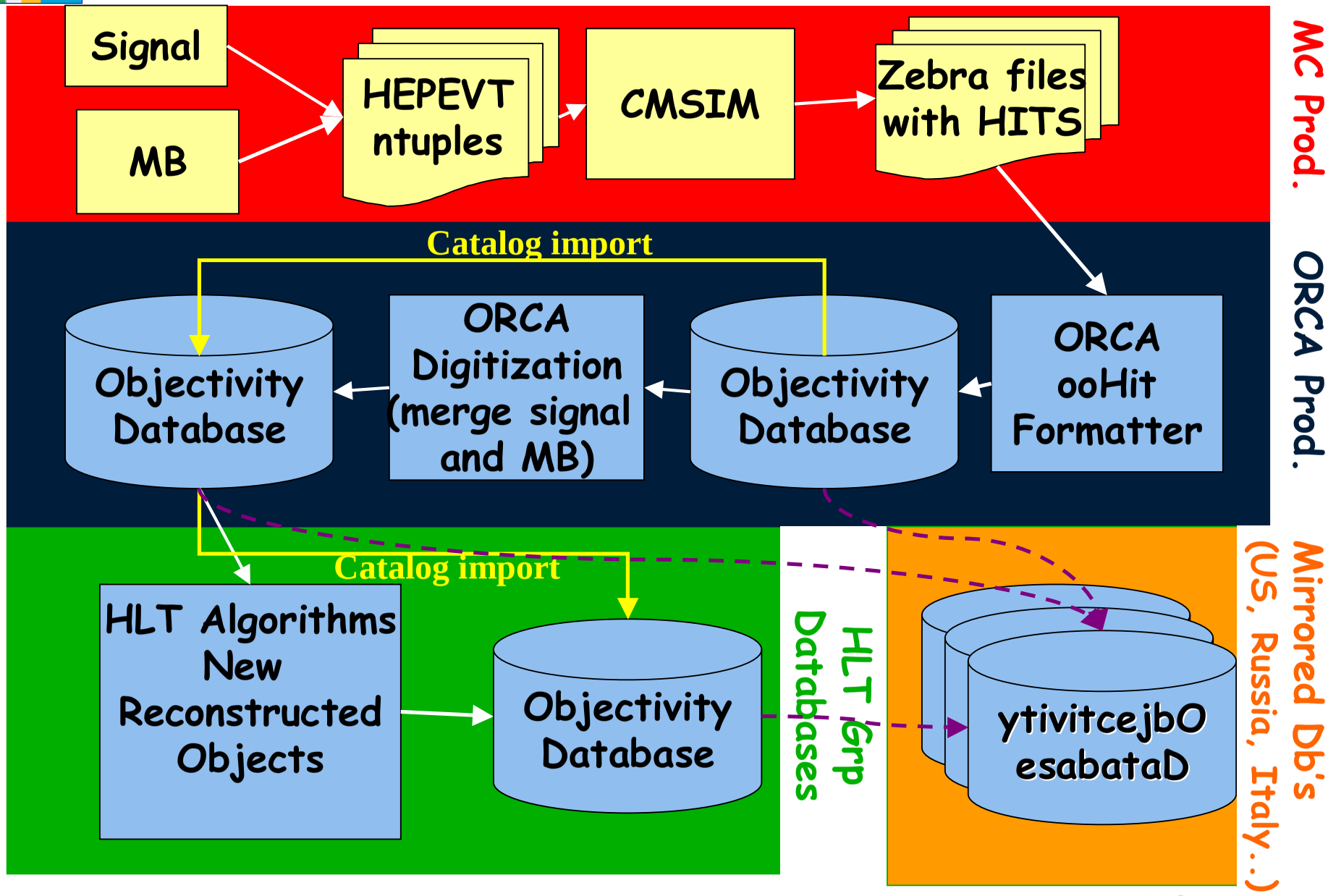
- Object Reconstruction for CMS Analysis
  - Project started in Sept. '98
  - Currently in fourth major release (4_3_0)
- Based on CARF framework
- Large base of detector code
  - Detailed digitization with Pileup. Persistent storage
  - L1 trigger simulation
  - Clusters, Tracks, Muons, Jets, Vertices
    - Being used by physicists in their studies
    - Not yet persistent
    - Much tuning remains to be done, but the groundwork is in place
  - Very little Fortran code remaining (mostly associated with the GEANT3/CMSIM interface)

Use of ORCA

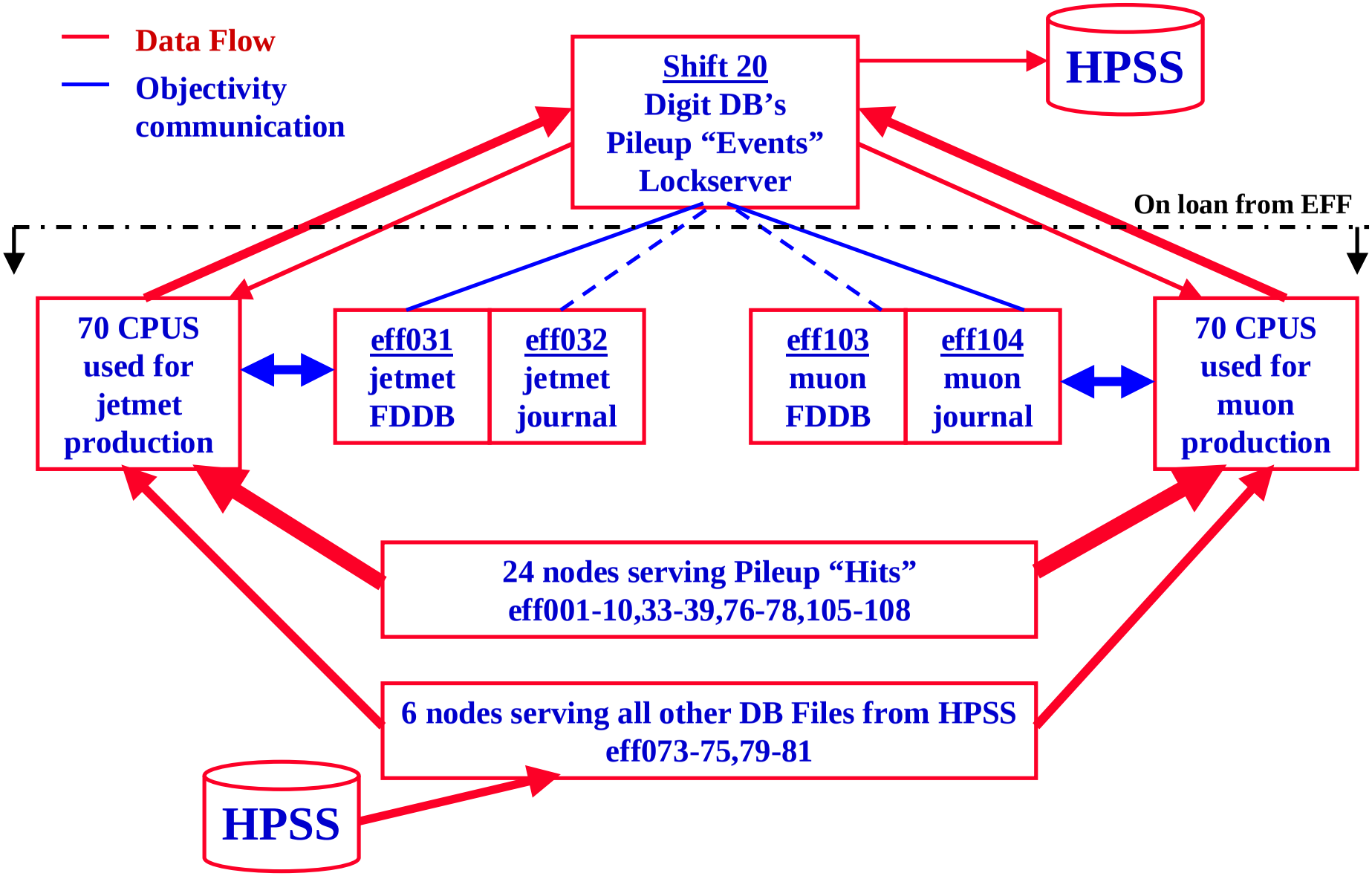
- Physics/Trigger results related to HLT.
  - Digitization is resource intensive stage
    - 200 events pileup for each signal event
  - Trial production in Oct 99
  - Major Production in Spring 2000
    - Simulation Worldwide
    - Digitization and reconstruction at CERN
      - 200 CPUs, 2 weeks production, 70TB data transferred
    - 2 million events with full pileup at 10^{34} luminosity
    - ~4TB of data stored and analyzed
  - Third production now underway
    - 5 million events (ie. Pileup: 1 Billion events!)
    - Simulation and Reconstruction Worldwide
    - GLOBUS/GRID based tools for Database import/export

Digitization of one CMS event at full luminosity requires (from first principles) about 400 times more computing than one Run II event

ORCA Production 2000



Production Farm (Spring 2000)



Performance

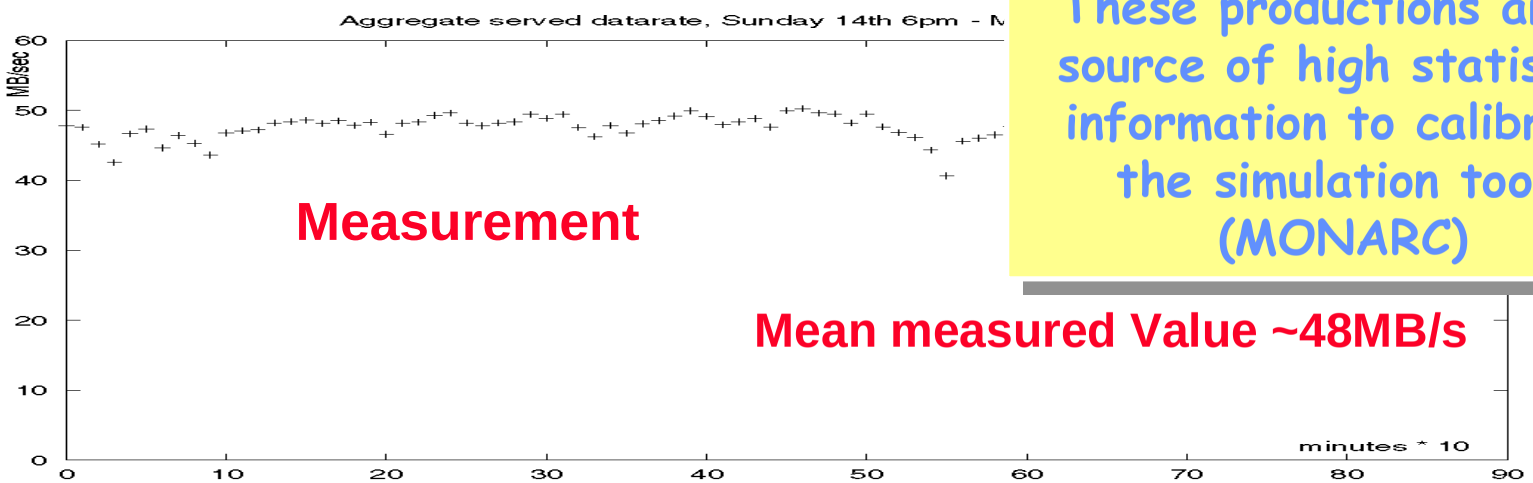
- More like an Analysis facility than a DAQ facility
  - 140 jobs reading asynchronously and chaotically from 30 AMS server's, writing to a high speed SUN server
  - Non disk-resident data being staged from tape
  - 70 jetmet jobs at ~60 seconds/event and 35MB IO/event
  - 70 muon jobs at ~90 seconds/event and 20MB IO/event
  - Best Reading rate out of Objectivity ~ 70MB/sec
  - Continuous 50MB/sec reading rate
  - 1million jetmet events, ~ 10 days
  - 1million muon events, ~15 days

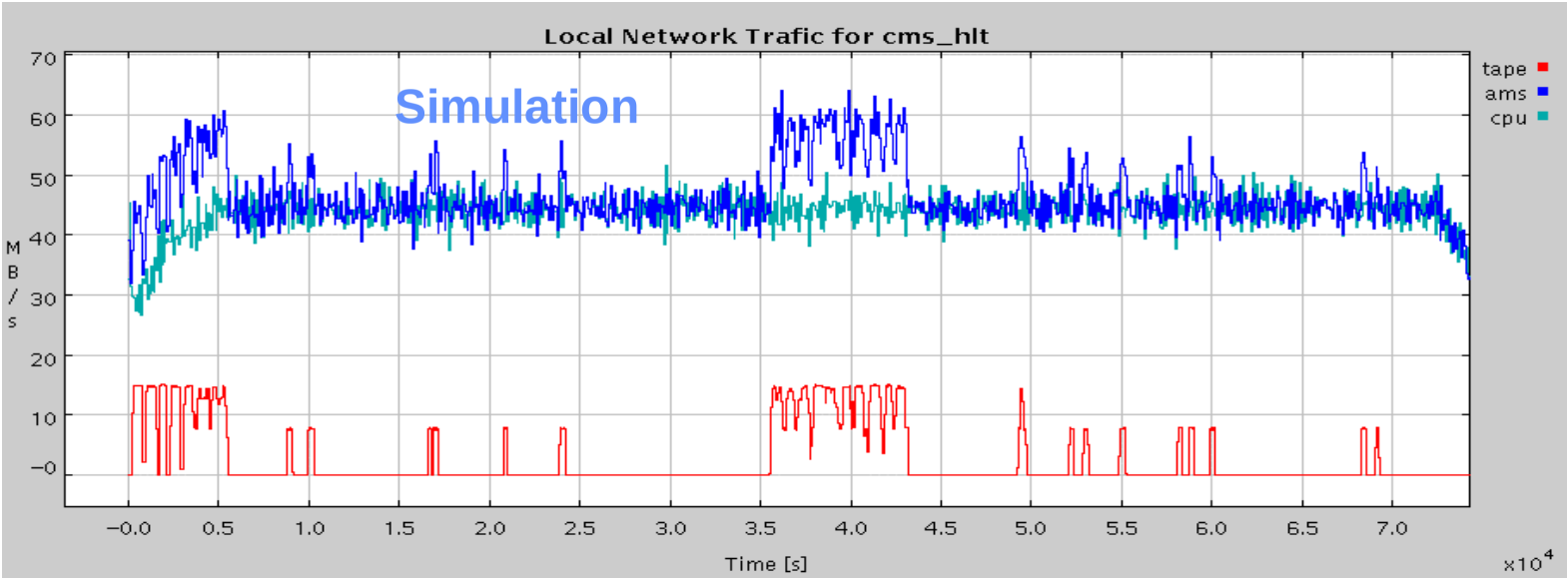
} in parallel

- Extensive monitoring was deployed, significant statistics were accumulated and analyzed

Feedback to MONARC Simulations

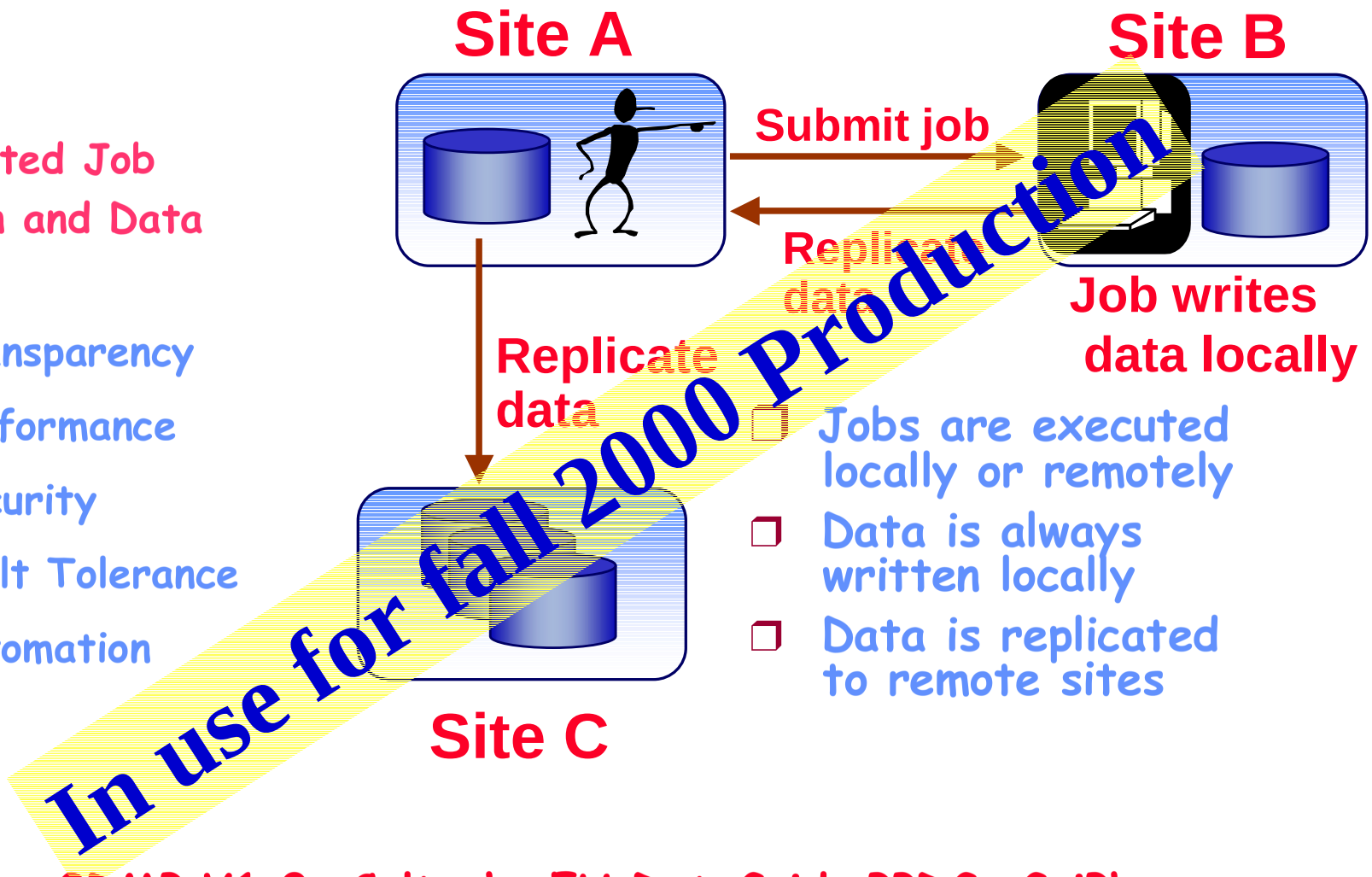
These productions are a source of high statistics information to calibrate the simulation tools (MONARC)





Grid Data Management Prototype

- Distributed Job Execution and Data Handling:
  - Transparency
  - Performance
  - Security
  - Fault Tolerance
  - Automation



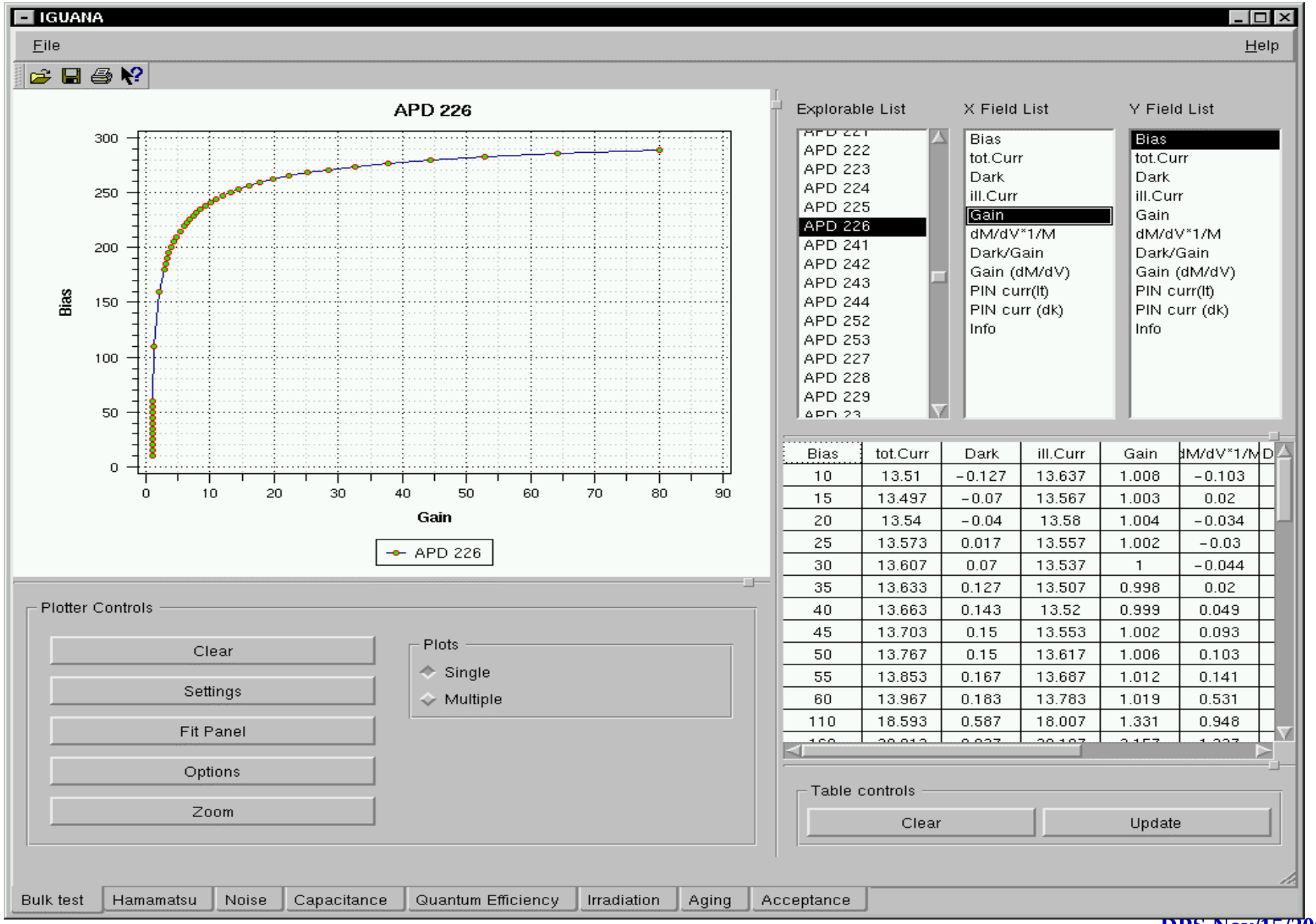
- Jobs are executed locally or remotely
- Data is always written locally
- Data is replicated to remote sites

GDMP V1.0: Caltech, EU DataGrid, PPDG, GriPhyn

User Analysis Software Strategy

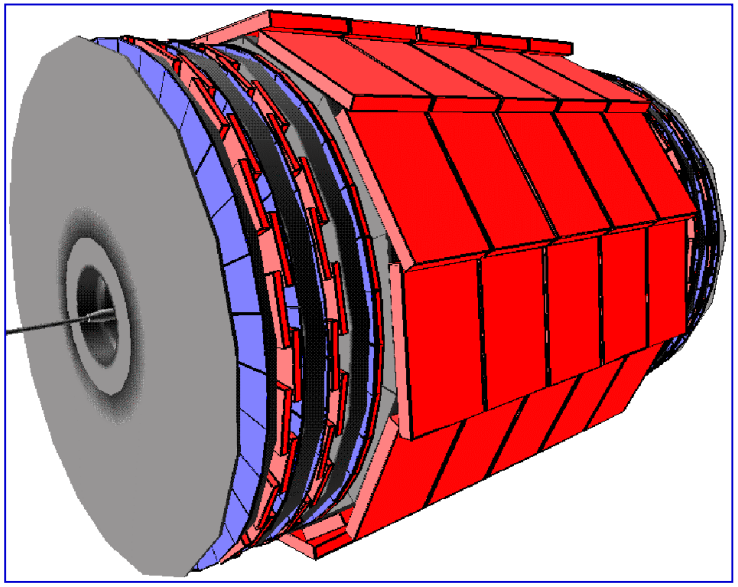
- Core software "User Analysis Environment" should:
  - Ensure coherence with framework and data handling
    - Coherent across all CMS software as much as possible
    - Filtering of data sets, user collections, user tags, etc.
  - Emphasize generic toolkits (not monolithic applications)
    - Histogramming, fitting, tag analysis,….
    - Graphical user interfaces, plotting, etc.
    - Interactive detector and event visualization
  - Include supporting infrastructure
    - Support for remote deployment and installation
    - Building against CMS, HEP, and non-HEP software
    - Documentation,...

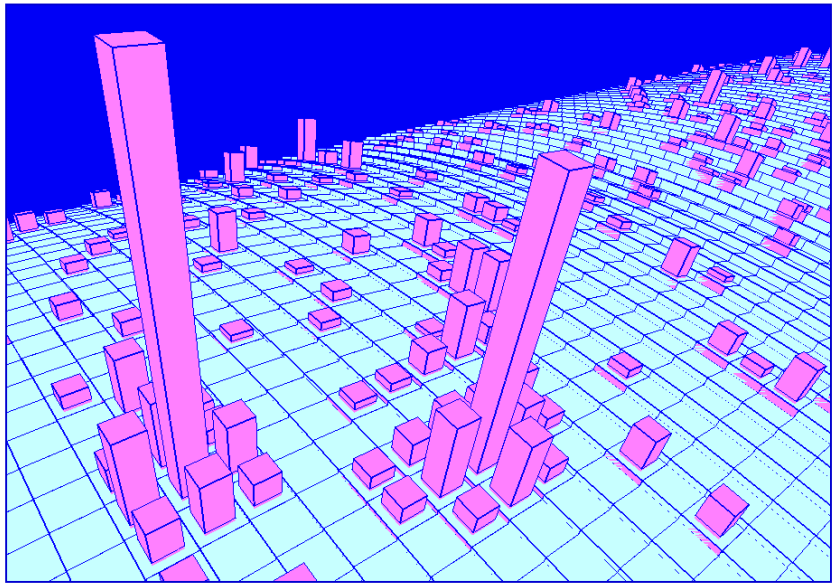
IGUANA: GUI / Analysis Components

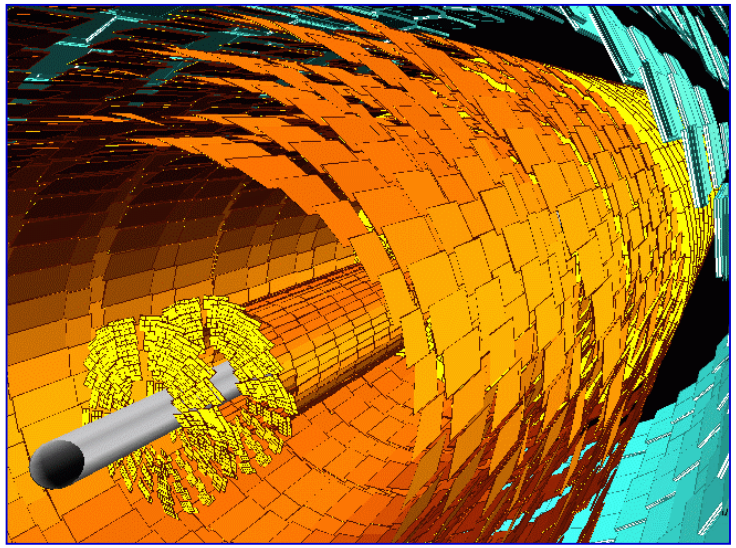


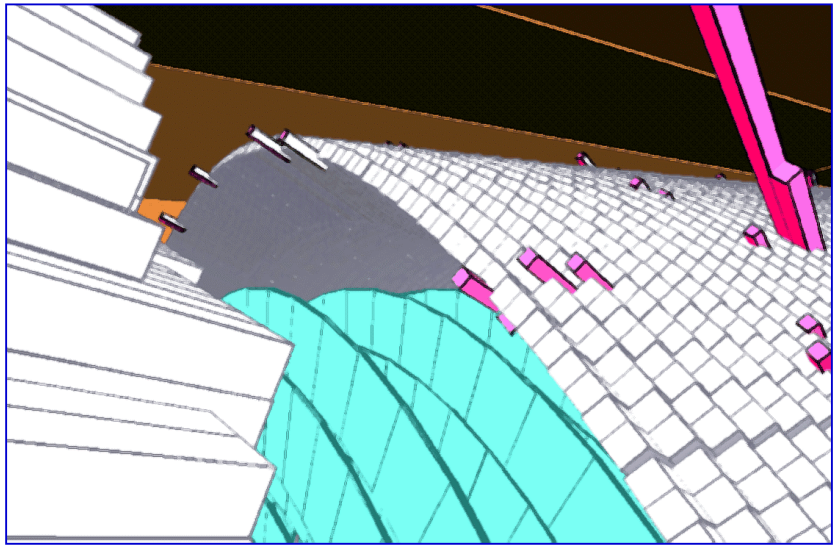
Bias	tot.Curr	Dark	ill.Curr	Gain	dM/dV*1/M
10	13.51	-0.127	13.637	1.008	-0.103
15	13.497	-0.07	13.567	1.003	0.02
20	13.54	-0.04	13.58	1.004	-0.034
25	13.573	0.017	13.557	1.002	-0.03
30	13.607	0.07	13.537	1	-0.044
35	13.633	0.127	13.507	0.998	0.02
40	13.663	0.143	13.52	0.999	0.049
45	13.703	0.15	13.553	1.002	0.093
50	13.767	0.15	13.617	1.006	0.103
55	13.853	0.167	13.687	1.012	0.141
60	13.967	0.183	13.783	1.019	0.531
110	18.593	0.587	18.007	1.331	0.948

IGUANA Event Display with ORCA









Current Analysis Chain

Prod.

- CMSIM (Geant3) Simulation
- ORCA Hit formatting into ODBMS
- ORCA Pileup and Digitization of some subset of detectors

Users

- Selection of events for further processing
- Create new collection
  - shallow (links), or deep (data) to existing objects
  - add new objects (ie. Tk Digits)
  - replace existing objects if required
- All combinations of
  - Transient - (Persistent)
  - Persistent - Transient
  - Persistent - Transient - Persistent

User collections of events of interest

- Collections can span datasets, navigation back from event data to run data can be used to get correct setups

Full ODBMS functionality being used

Organization of the Meta-data and production issues, are typically much more technically complex than the *simple* issues of persistent objects!

Data Handling

- We already have more data (10TB) than disk space
  - use MSS systems (HPSS at CERN, ENSTORE at FNAL)
  - automatic staging by hooks in Objectivity code
- Tools are in place to replicate federations
  - Shallow (catalog and schema files only)
  - Meta-deep (plus meta-data)
  - Deep (plus Database files)
  - LAN and WAN

Leverage products like Objectivity to reduce the amount of SW we have to write

- "Request Redirection Protocols" being prototyped and field tested
  - allows to "hide" actual data location from users
    - disk goes down, single change on a central server redirects user requests to a new server
    - A powerful place to hook into Objectivity to give us the required control

Product- and Project- ize

Moving from system-scale to enterprise-scale.

- Productize
  - Make Software products that:
    - are capable of running at all regional centers,
    - give consistent, repeatable and verifiable results
    - have well understood quality
    - have been tested rigorously
    - have controlled change...
  - Rate of increase of new features will reduce as more work goes onto quality

- Projectize
  - General milestones were satisfactory for the prototype stage
  - Now require a more detailed planning,
  - Attachment of milestones to critical (in-project and off-project) items
  - Filling in of short-term milestones
  - More detailed manpower planning

5% and 20% Data Challenges

- In 2002/2004 we have 5% and 20% data challenges specified
  - In fact these are just part of a steady process
- The % we refer to is the % of complexity, rather than the genuine 5% of the data, or 5% of the CPU…
- Use Farms sized (in boxes) at 5/20%
  - processing speed whatever it is
  - run for a period of a month or so to see the whole system in operation:
    - first pass reconstruction
    - roll data out to users continuously (no scheduled downtimes)
    - selected "streams" (collections) in operation
    - user offline selection -> user collections
    - replication between sites
    - timely "results"

Summary

We are on course

Many aspects of the computing model are already being tested in detail.

User acceptance is significant and growing.

The collaboration is fully engaged in this process.