CMS Software and Computing

M. Stavrianakou
On behalf of the CMS collaboration
IEEE/NSS 2005 - LHC Software: Crunch Time
26.10.2005

Outline

- **External (Common LHC, HEP) Software**
- **CMS Software**
  - **Organization, Framework and Software Domains**
  - **Reconstruction and Detector Performance**
  - **Trigger**
  - **Fast Simulation**
- **CMS Computing**
  - **CMS Computing TDR**
  - **LCG Service Challenge 3**
  - **Software/Computing Integration**
- **Major Milestones in CMS Software and Computing**

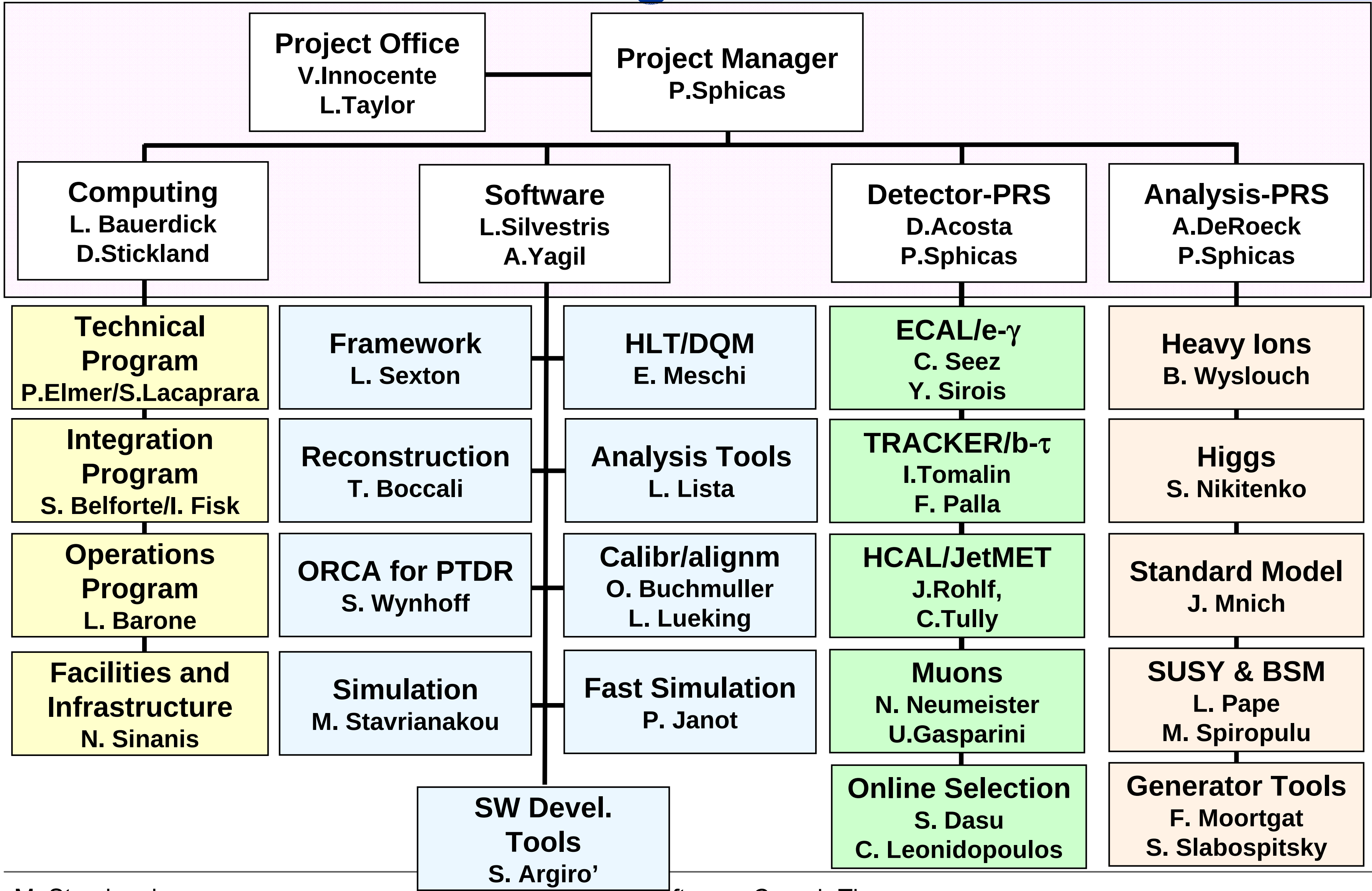
External (common LHC/ HEP) software

External software currently used in CMS

- **Persistency framework: POOL**
- **Core Libraries and Services: SEAL**
- **Data analysis: ROOT**
- **Generators: GENSER**
- **Simulation: Geant4**
- **HEP class libraries: CLHEP, CLHEP/HepMC**
- **Misc. libraries: BOOST**
- **Parsers: Xerces**
- **Visualization: Qt, Coin3D,…**
- **Scripting: Python, Perl**
- **Miscellaneous:**
  - **CppUnit, RuleChecker…**
  - **Oracle, MySQL…**

CMS Software

Computing, Physics, Trigger (CPT) organization



Software Project: two major lines

- **Physics TDR – Current thread**
  - Goal: Maintain and support all current SW projects, in coordination with Physics Reconstruction and Selection (PRS) groups, until April/May 2006
- **Migration thread**
  - Goal: Software Components for the Cosmic Challenge
    - Framework (FW)
    - Event Filter and integration with DAQ and FW
    - Geometry DB/ Calibration and Alignment infrastructure (online/HLT/Offline)
    - Basic Reconstruction (Local)
    - Additional reconstruction needed for Muon Chambers
    - Detector Monitoring and Data Quality Monitoring
    - Basic Visualization
- **Cosmic Challenge will use new EDM and Framework**
  - Essential test toward the integration of the full experiment

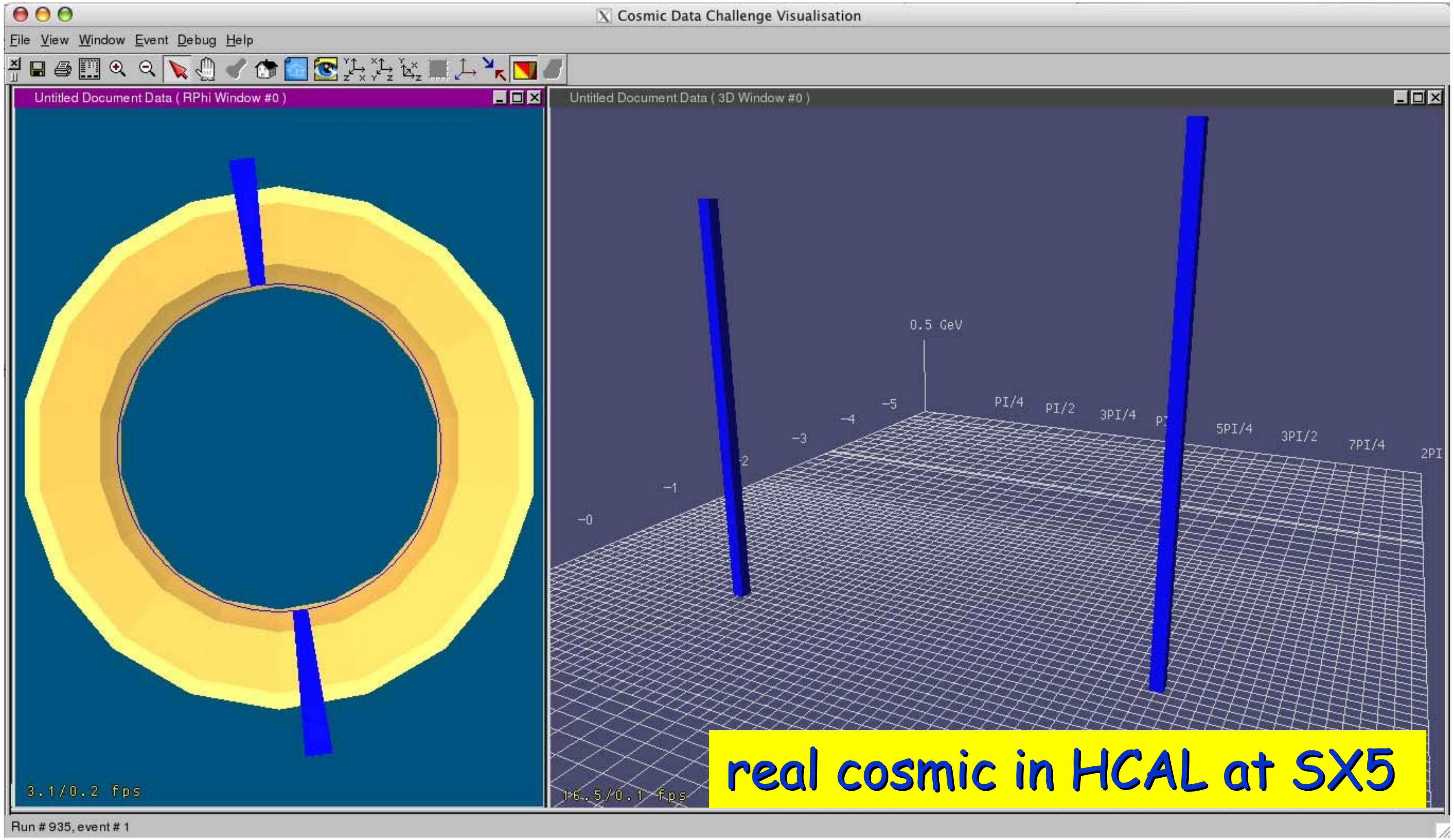
New Software: executive summary (I)

- **Framework in place**
- **Event Filter (EvF)**
  - **Well tested in old FW; migration to new FW in progress**
- **Calibration and Alignment**
  - **Calibration objects defined in collaboration with detector/reconstruction groups**
  - **First-round implementation of conditions service complete**
  - **Interval Of Validity (IOV) object prototyped**
  - **End-to-end (E2E) test well advanced**

New Software: executive summary (II)

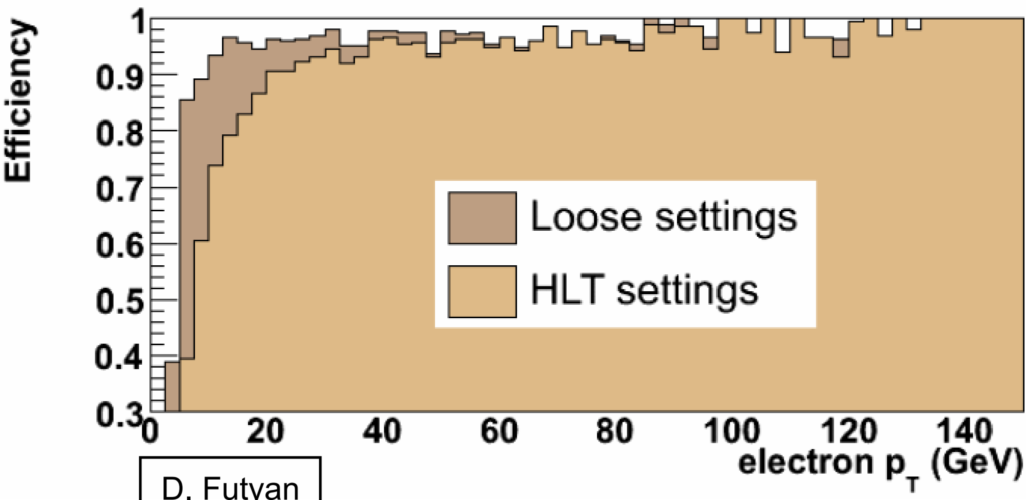
- **Simulation**
  - **Most components in place; integration for production testing scheduled mid-November 2005**
- **Geometry service in new FW**
  - **Done, being tested**
- **Local Reconstruction**
  - **Well underway**
- **Data Quality Monitoring (DQM)**
  - **In excellent shape; porting to new FW in progress**
- **Visualization**
  - **Basic visualization for cosmic data in place**

CMS data


Cosmic Data Challenge Visualisation
File View Window Event Debug Help
Untitled Document Data (RPhi Window #0)
Untitled Document Data (3D Window #0)
0.5 GeV
PI/4
PI/2
3PI/4
5PI/4
3PI/2
7PI/4
2PI
3.1/0.2 fps
Run # 935, event # 1
real cosmic in HCAL at SX5

Reconstruction and detector performance

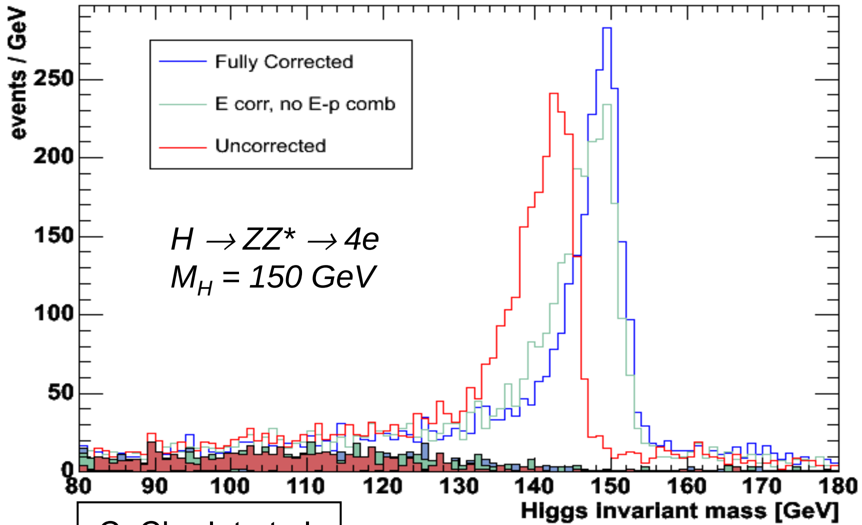
Electron reconstruction



D. Futyan

Effect of loosening seed cluster E_T threshold and pixel matching on electrons in $H \rightarrow ZZ^* \rightarrow 2e2\mu$

Combining Ecal E and track P measurements for best 4-vector estimates
(E scale corrections for Ecal, depending on e classification, full errors in weighted means)

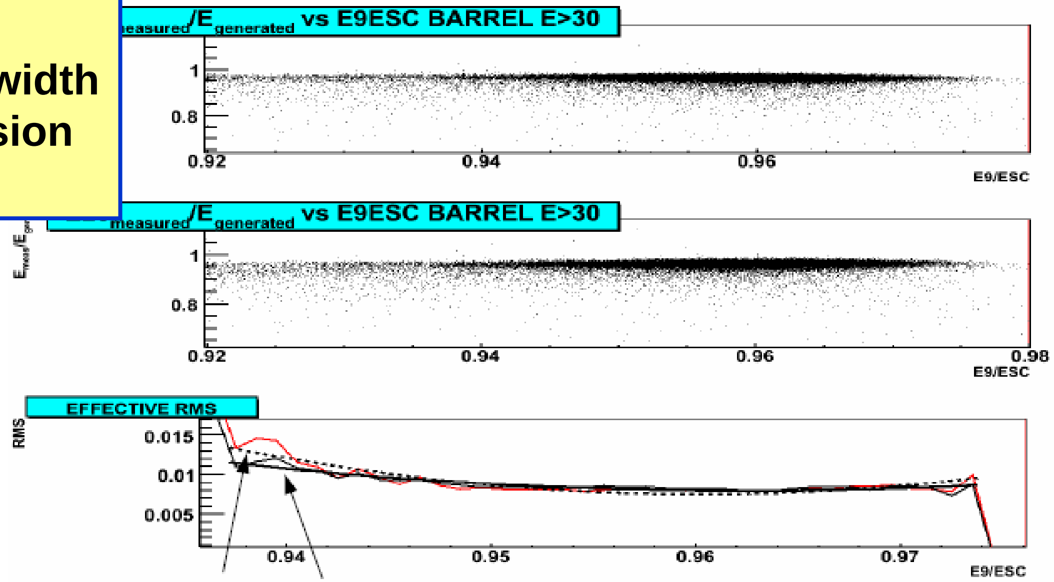


C. Charlot et al.

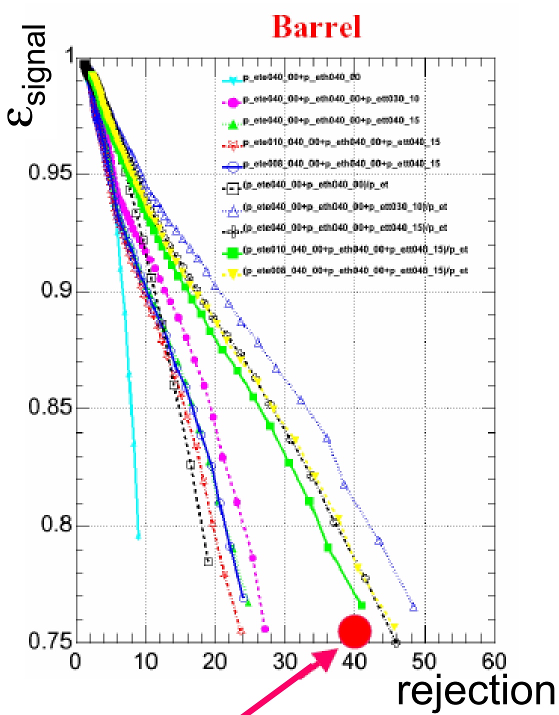
Photon reconstruction

Photon error estimation

Here shown as a function of shower width (sensitive to conversion, and conversion distance from ECAL)



T. Aaltonen, K. Lassila



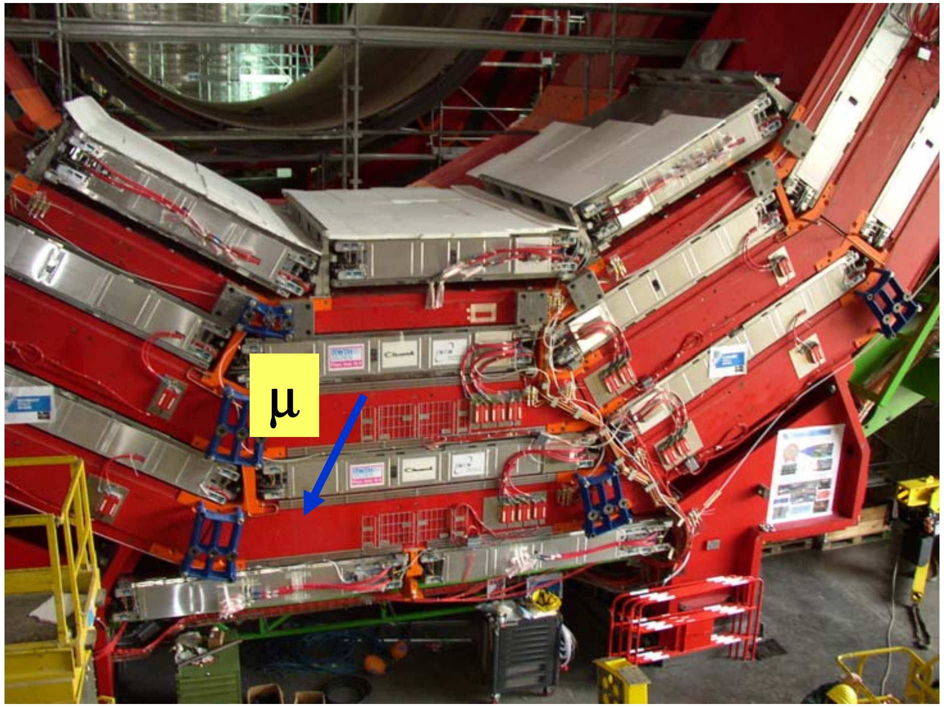
Previous result

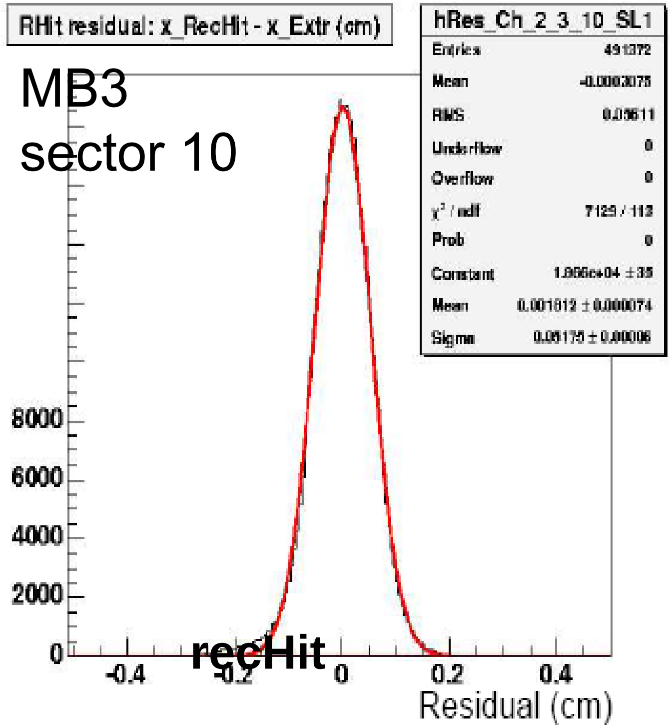
M. Pieri

Photon isolation studies

Here looking for very large rejection factors needed for H→γγ channel, and using combination of ECAL, HCAL and tracker

Muons: reconstruction on cosmics data

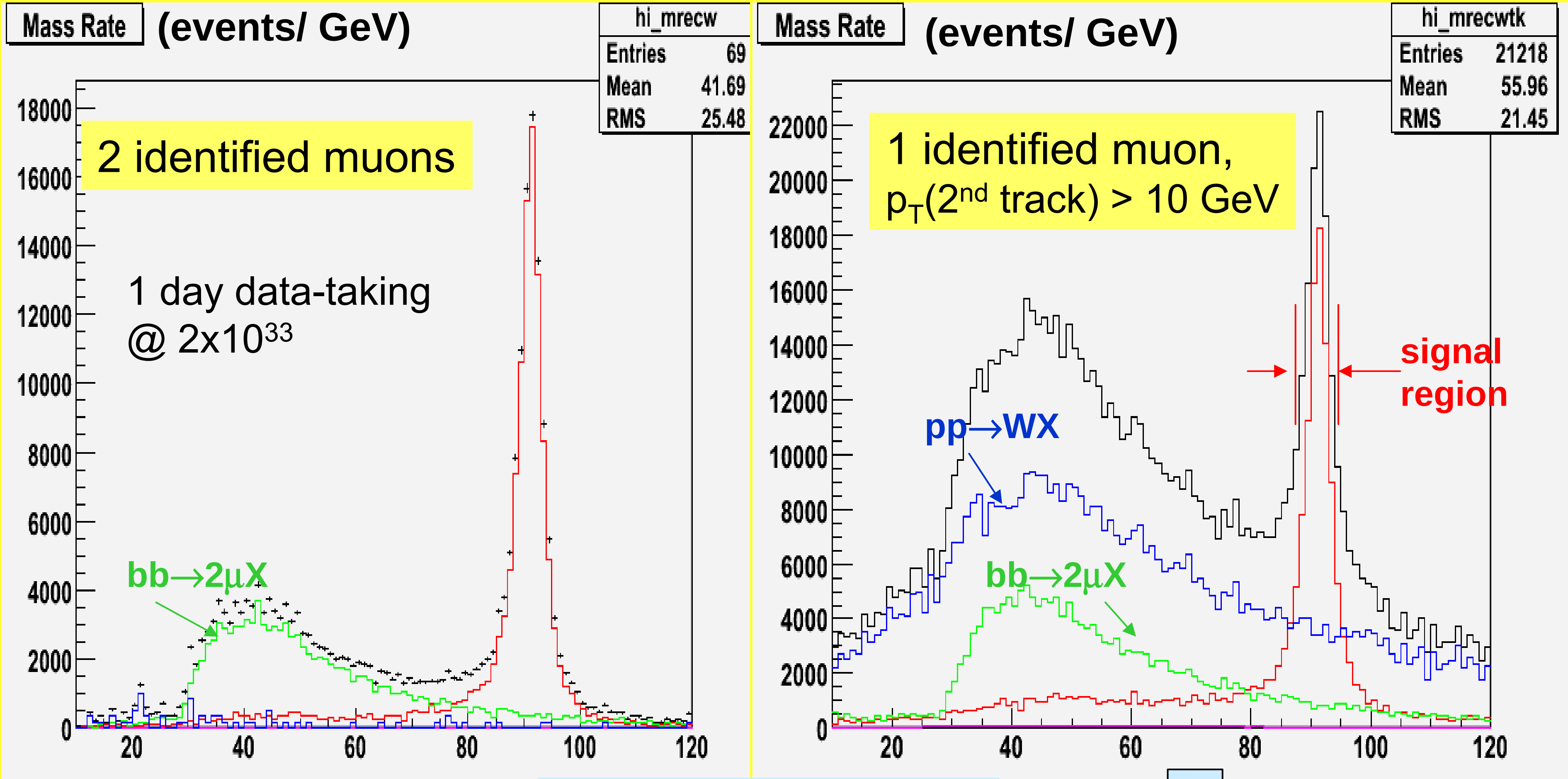




σ=510 μm
(expected from cosmics
(not bunched!)

12 points track segments in one chamber

Z reconstruction



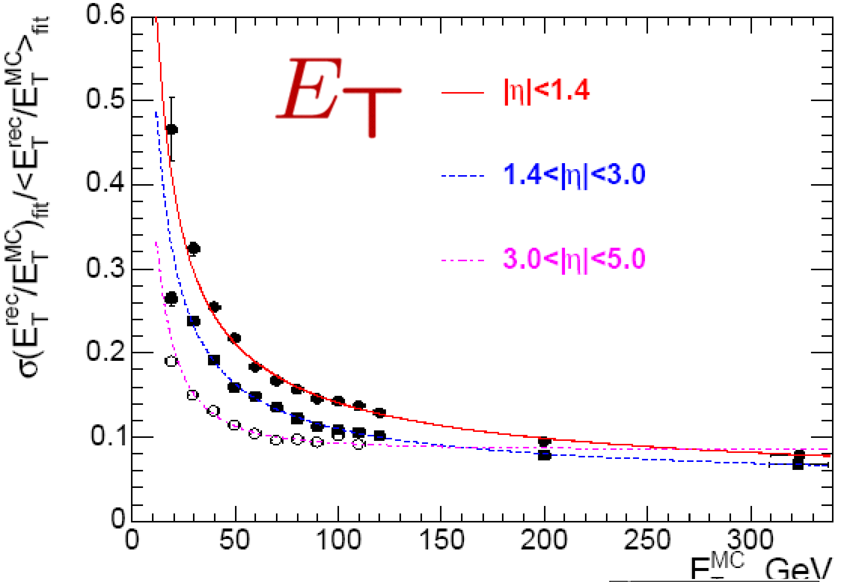
reconstructed $M(\mu^+\mu^-)$

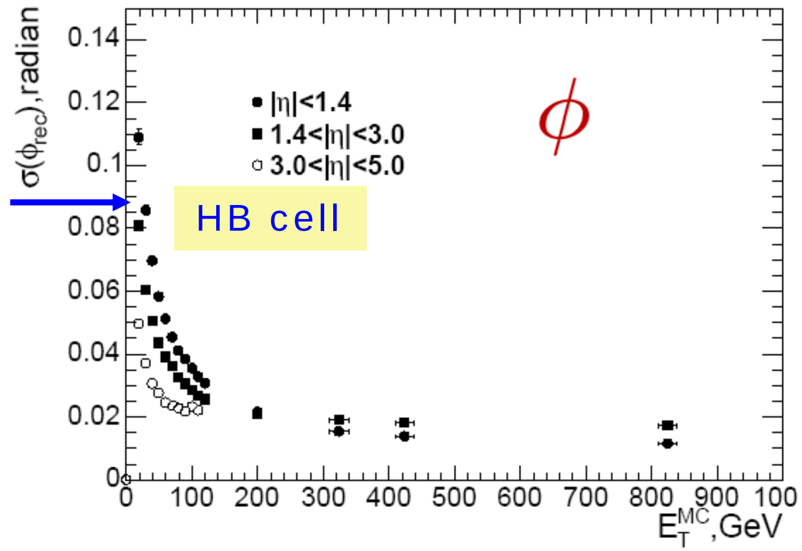
HLT+offline cuts:
$P_T^{1,2}$>20,10 GeV/c

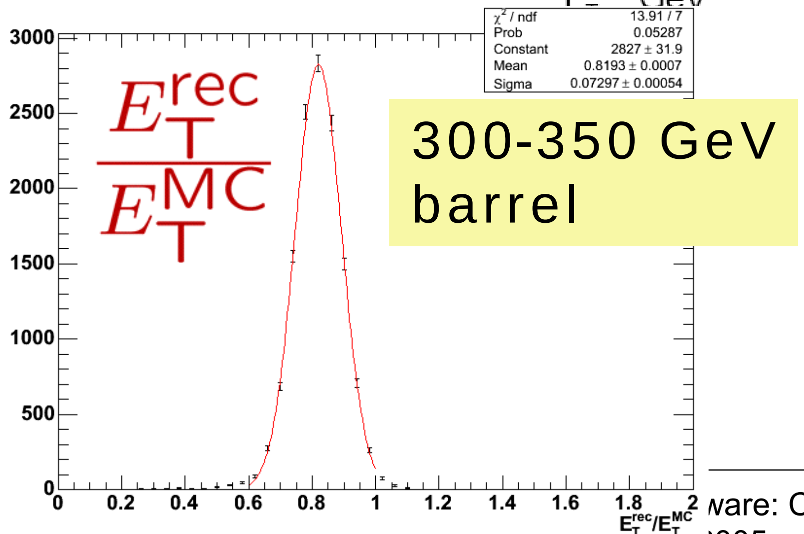
suitable for efficiency study from real data

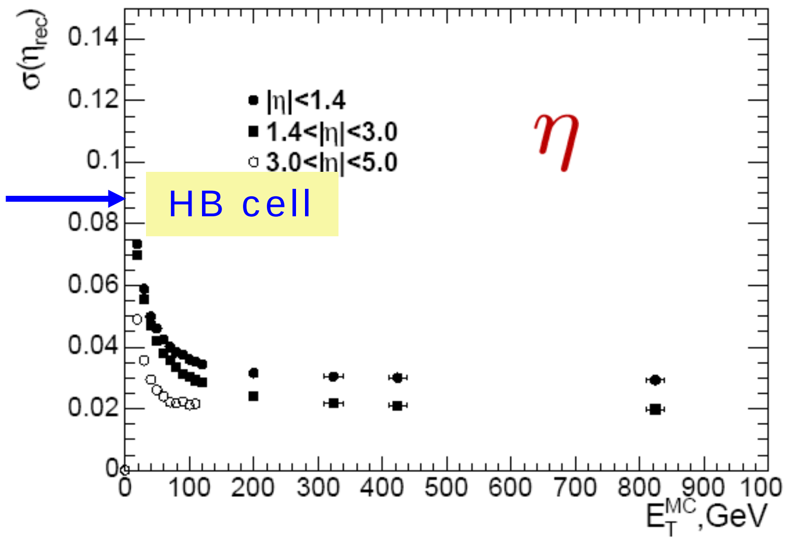
Jet Measurement

- **A. Heister et al. AN 2005/05**



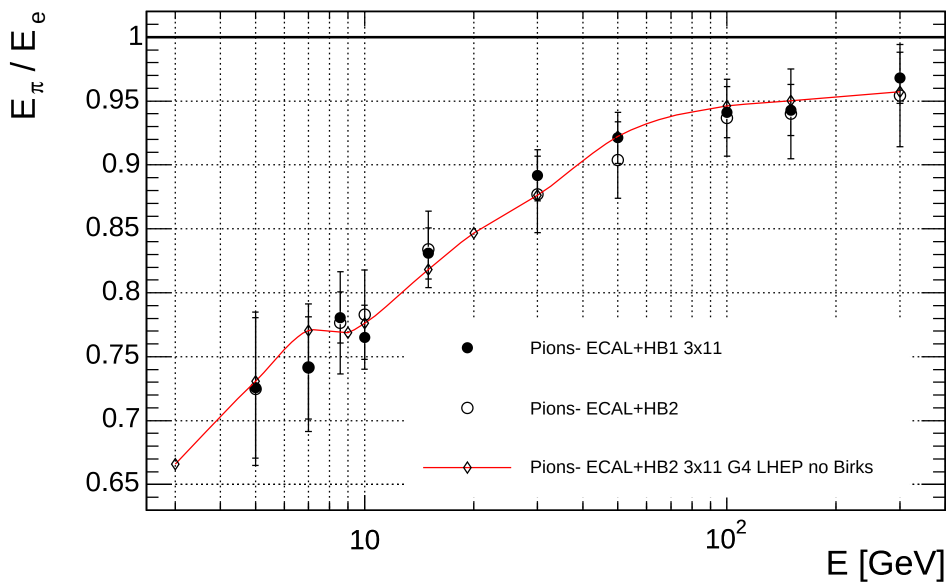


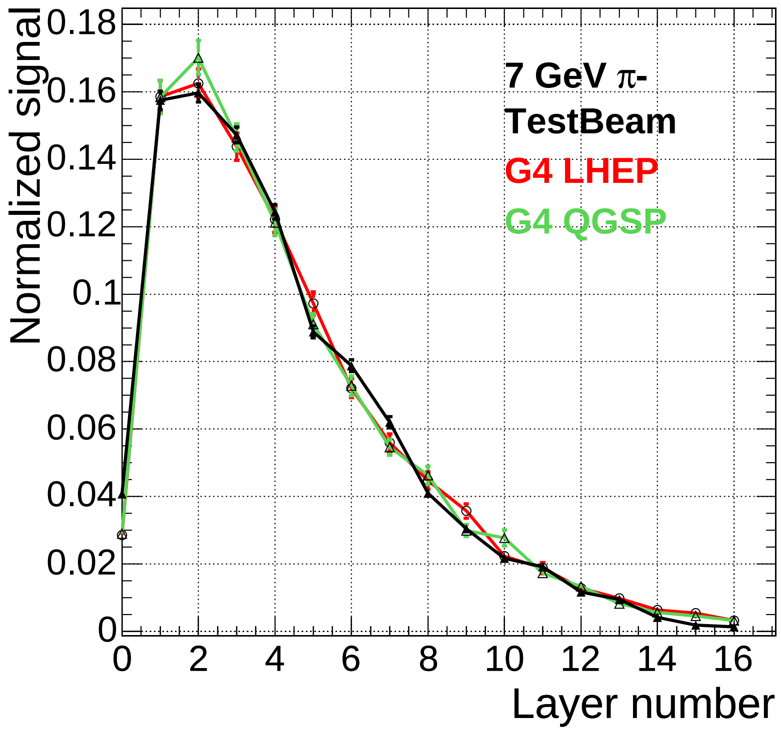




HCAL response at low energy

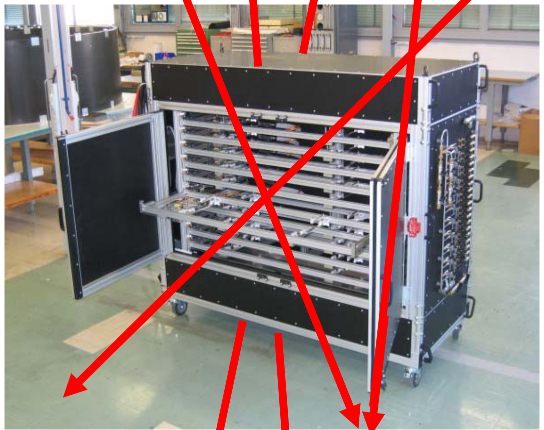
- **test beam 2004 results**

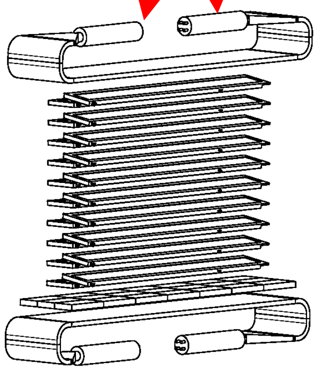




J. Damgov *et al.*

Cosmic data of June 05 used in track-based alignment

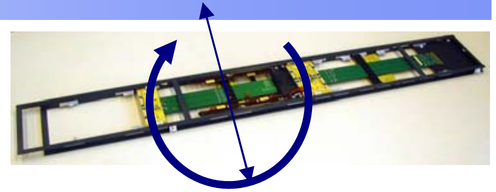




Cosmic Rack / FinnCRack
corresponds to a slice of TOB,
4 fully equipped rods were used

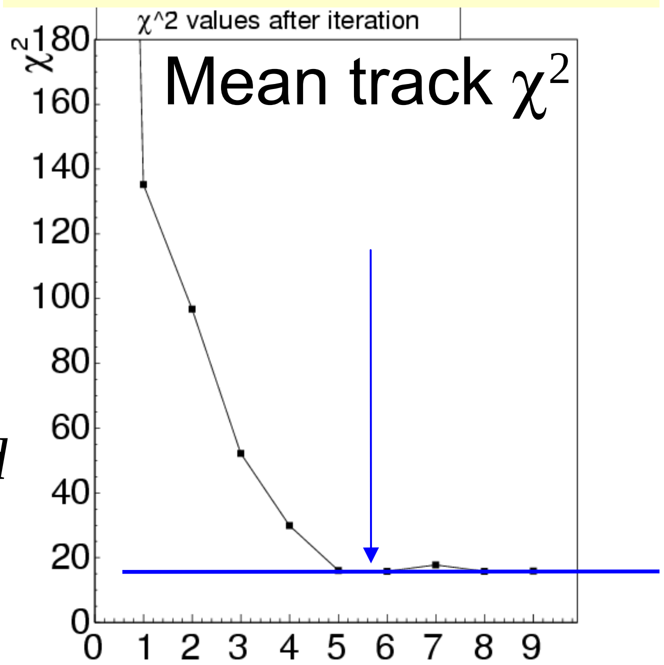
- **Proof of principle: track-based alignment with CMS software tools works**
- **Serves development platform for alignment algorithms**
- **Future: large cosmic runs to validate TOB rod survey measurements**
- **$\rightarrow$ Defines pre-data taking geometry!**

First application of track based alignment using real CMS data



Individual modules aligned

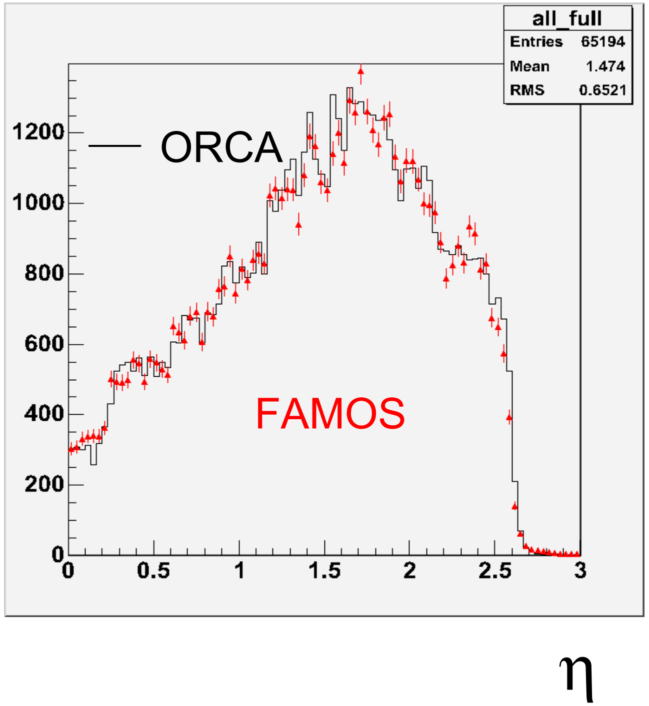
convergence & good alignment in 5 iterations



Fast Simulation: FAMOS

FAMOS highlights

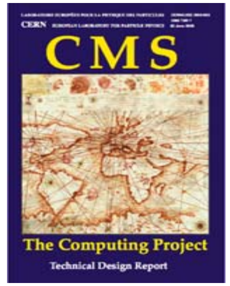
- Tracker material tuned layer by layer, as a function of η, up to 1% accuracy wrt full geometry
- Pileup simulation
- Decays of long-lived particles (K0s, charged pions, …)
- Muons: efficiency/resolution tables, muon isolation, global muons reconstruction
- Electrons/Photons
- HLT: JetMET, e/γ and muons
- Jets: shower parametrization, calibration, reconstruction…
- Root-based analysis



Number of Brem photons (>500MeV) for electrons with p_T = 35 GeV/c

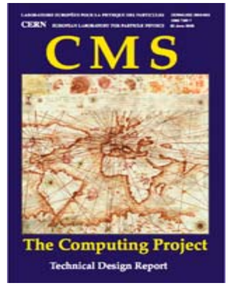
CMS Computing

Computing TDR (C-TDR)



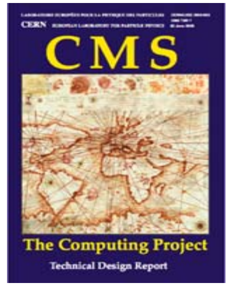
- **C-TDR goals**
  - Summarize, refine the Computing Model (CM), expand on analysis and role of Tier-n centres
  - Explain the architecture of the CMS computing system
  - Detail the project organization and technical planning
  - Form the basis for further internal discussion
- **C-TDR - baseline of CMS computing services and workflows**
  - Specify “baseline” targets and development strategy
    - Alongside LCG TDR, describing facilities and Grid baseline services
  - Not an engineering “blueprint” for the computing system
- **LHCC review of C-TDR (07.10.2005) successfully concluded**

Computing Model Principles



- **Structured Data + Structured Data Grid**
  - **Primary Datasets, Data Tiers — optimizes sequential data access**
  - **Tiered Grid — facilitates well-defined roles/responsibilities for centers**
- **Data Granularity — Primary Datasets**
  - **Data always needs to be considered in its trigger context -> trigger paths**
  - **O(2PB)/yr raw data split into O(50) (40TB) trigger-determined datasets**
- **Data Tiers — RAW, RECO, AOD**
  - **Keep RAW and RECO close together (initially at least) —> FEVT**
  - **AOD, full copy at each Tier-1, partial (even full) copies at many Tier-2**
- **Computing Tiers**
  - **CMS-Tier0: prompt reconstruction, close connection to online**
  - **CMS-CAF: high-priority, short-latency data processing, analysis support**
  - **Tier-1s: FEVT data custody, selection, distribution, re-reco**
  - **Tier-2s: MC production, data analysis under physicist “control”**

Technical Baseline Principles



- **Optimize for the common case:**
  - **Optimize for read access**
    - **Most data is write-once, read-many**
  - **Optimize for bulk processing, but without limiting single user**
- **Decouple parts of the system:**
  - **Minimize job dependencies**
    - **Allow parts of the system to change while jobs are running**
  - **Site-local information stays site-local**
- **Know what you did to get here:**
  - **‘Provenance tracking’ is required to understand data origin**
- **Keep it simple!**
- **e.g.: initially use explicit data placement**
  - **Data does not move around in response to job submission**
  - **All data is placed at a site through explicit CMS policy**

Computing Organization Principles

- **new computing organization based on projects and programs**
  - **realize the need for (sub-)projects that need to be well-scoped**
  - **require institutional responsibility to provide effort to projects**
- **close coordination across management team**
  - **incl. weekly meetings w/ Grids: Integration Task Force, LCG PEB/MB**
- **up-front: address the issue of "getting computing ready"**
  - **have computing model and technical baseline in hand**
  - **use integration milestones as Computing task drivers**
  - **system continually 'up', DC/SC, incrementally improving capabilities**
- **establishing integration, (analysis support) operations, facilities, programs**
  - **a paradigm shift towards focusing on making things work end-to-end**
  - **together with WLCG, EGEE/LCG, OSG, CERN-IT, T1s, T2s et al.**

LCG Service Challenge 3 (SC3)

- **Goal: demonstrate basic functionality expected at each computing tier simultaneously**
  - Data can be transferred from CERN to Tier-1 sites
  - Data can be validated and existence of data can be published
  - Remote applications can be executed against local data to simulate skimming and selection jobs
  - Data can be transferred to Tier-2 centers, while validation and publication proceeds and remote analysis applications execute against data
  - At the same time, simulation jobs are executed at Tier-2 centers and results transferred to Tier-1 centers for archiving
- **Sites have validated CMS components for transfer during the throughput phase, publication tools are tested at most sites, grid interfaces are validated at all sites**

Integration program

- **To ensure Computing delivers functional computing systems**
  - **preparing for the major service and data challenges**
  - **integrating CMS data management and workload management systems with EGEE and OSG middleware**
- **Covers**
  - **Production-Related Activities**
  - **Analysis-Related Activities**
  - **Database Deployment Activities**
  - **Participation in Service Challenges**

CMS-LCG Integration Task Force

- Goal: Satisfy CMS needs making EGEE middleware work for CMS
- "Make it work": not stop at problems, fix them working with EGEE/LCG
- TF is a CMS project, lead by CMS, part of Integration Program
- TF is build around specific items that CMS needs
- TF will not address all items in CMS agenda, just those that can be dealt with using EGEE middleware
- TF will steer EGEE/LCG development and deployment to have them take over as much as possible of CMS needs

CPT: major milestones in 2005/06

- **Computing TDR – done**
- **Magnet test (cosmic challenge)**
  - **ready by Dec 2005; "slice" tests in progress**
  - **start taking data in January, for ~5 months**
- **Physics TDR – Volume I**
  - **Draft 2 ready**
- **Physics TDR – Volume II**
  - **Detailed outline ready**
- **Data/computing/analysis challenge in 2006**
  - **Currently the most important goal for 2006**