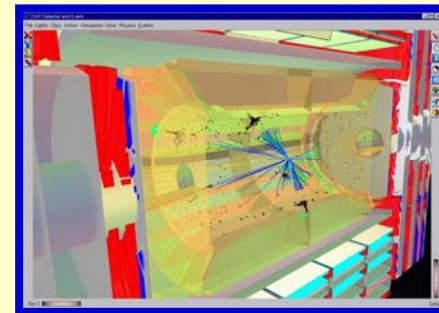
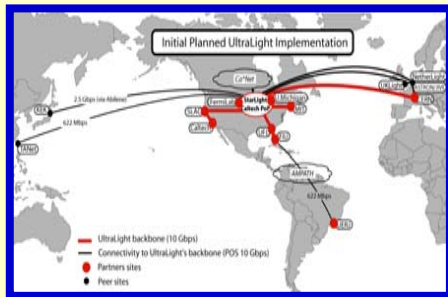


CMS Data Analysis using Grid and UltraLight Services

Richard Cavanaugh
University of Florida



- **CMS Physics Requirements**
- **CMS Application Software Stack**
- **UltraLight and Grid Services**
- **CMS Integrated System**

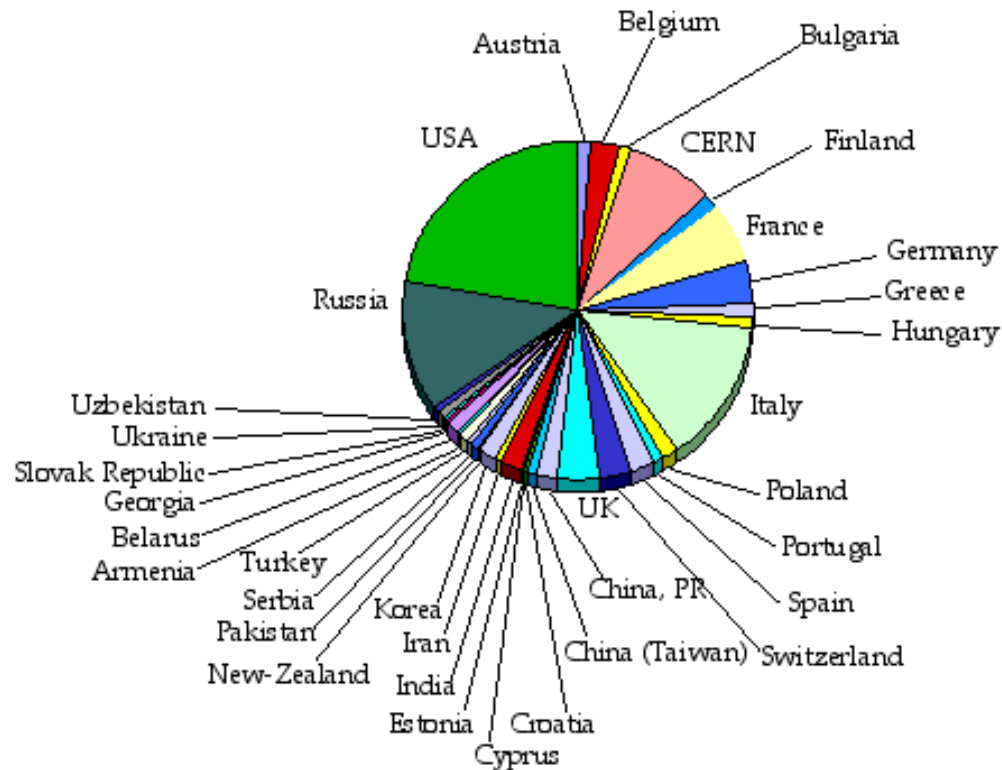
Physics Requirements : CMS Collaboration size & distribution

	Number of Laboratories
Member States	58
Non-Member States	53
USA	36
Total	147

	Number of Scientists
Member States	969
Non-Member States	467
USA	413
Total	1849

Associated Institutes	
Number of Scientists	64
Number of Laboratories	9

Nov. 18, 2002/gm
<http://cms.cern.ch/pictures/energy/overview.html>



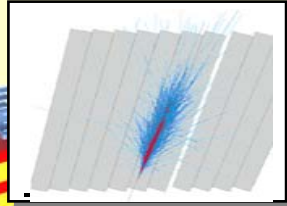
1849 Physicists and Engineers
34 Countries
147 Institutions

Physics Requirements : CMS Detector

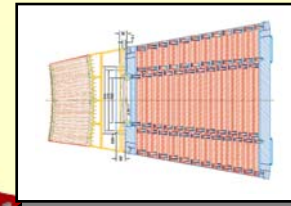
SUPERCONDUCTING COIL

Total weight : 12,500 t
 Overall diameter : 15 m
 Overall length : 21.6 m
 Magnetic field : 4 Tesla

CALORIMETERS
ECAL Scintillating $PbWO_4$ Crystals

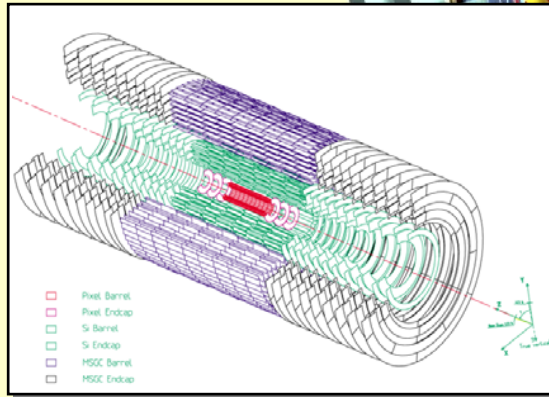


HCAL Plastic scintillator copper sandwich



IRON YOKE

TRACKERS

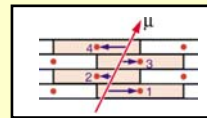


Silicon Microstrips
 Pixels

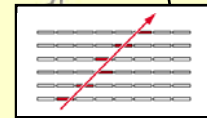
- Pixel Barrel
- Pixel Endcap
- Si Barrel
- Si Endcap
- MSQC Barrel
- MSQC Endcap

MUON ENDCAPS

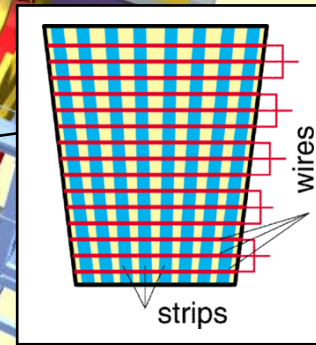
MUON BARREL



Drift Tube Chambers

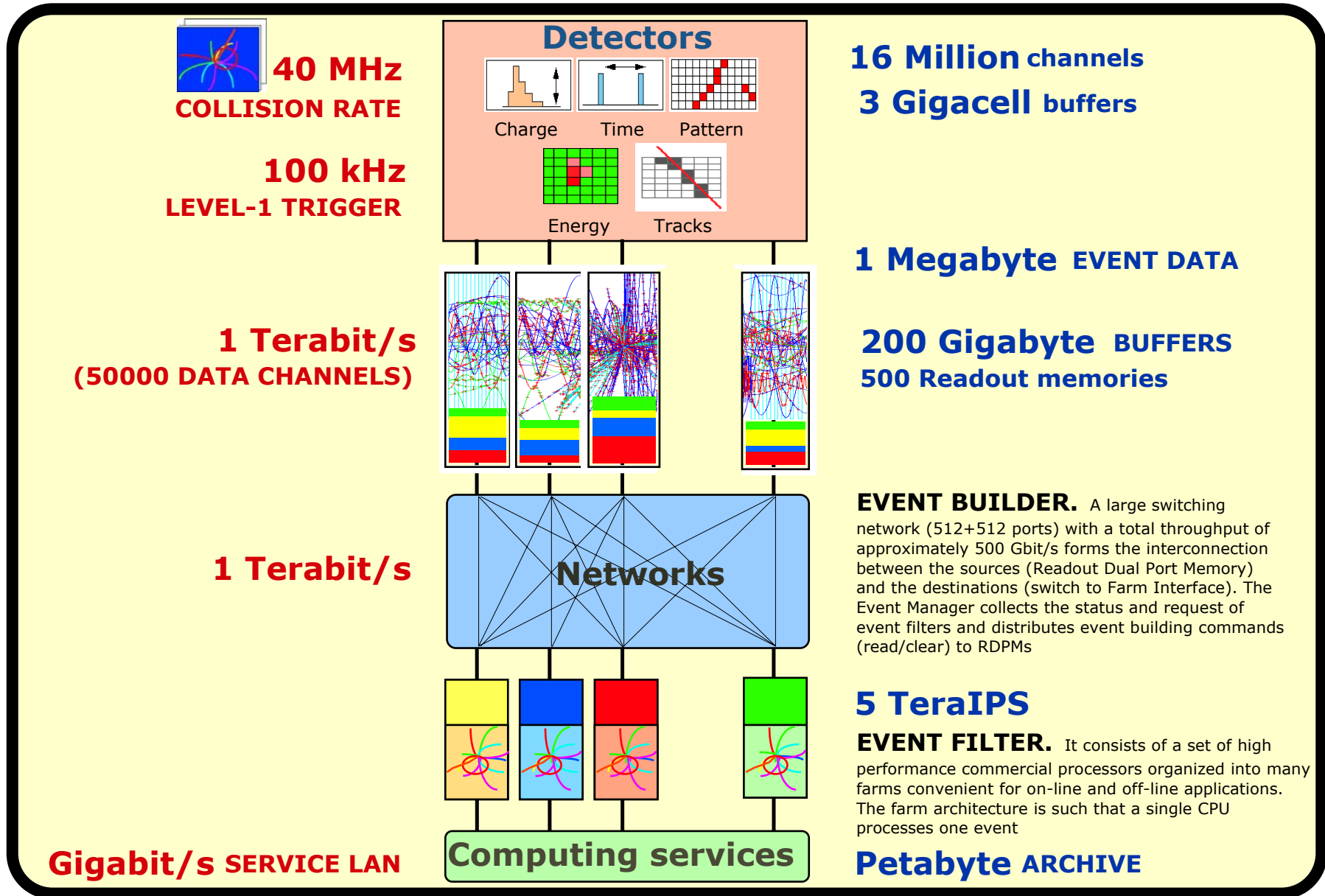


Resistive Plate Chambers



Cathode Strip Chambers
 Resistive Plate Chambers

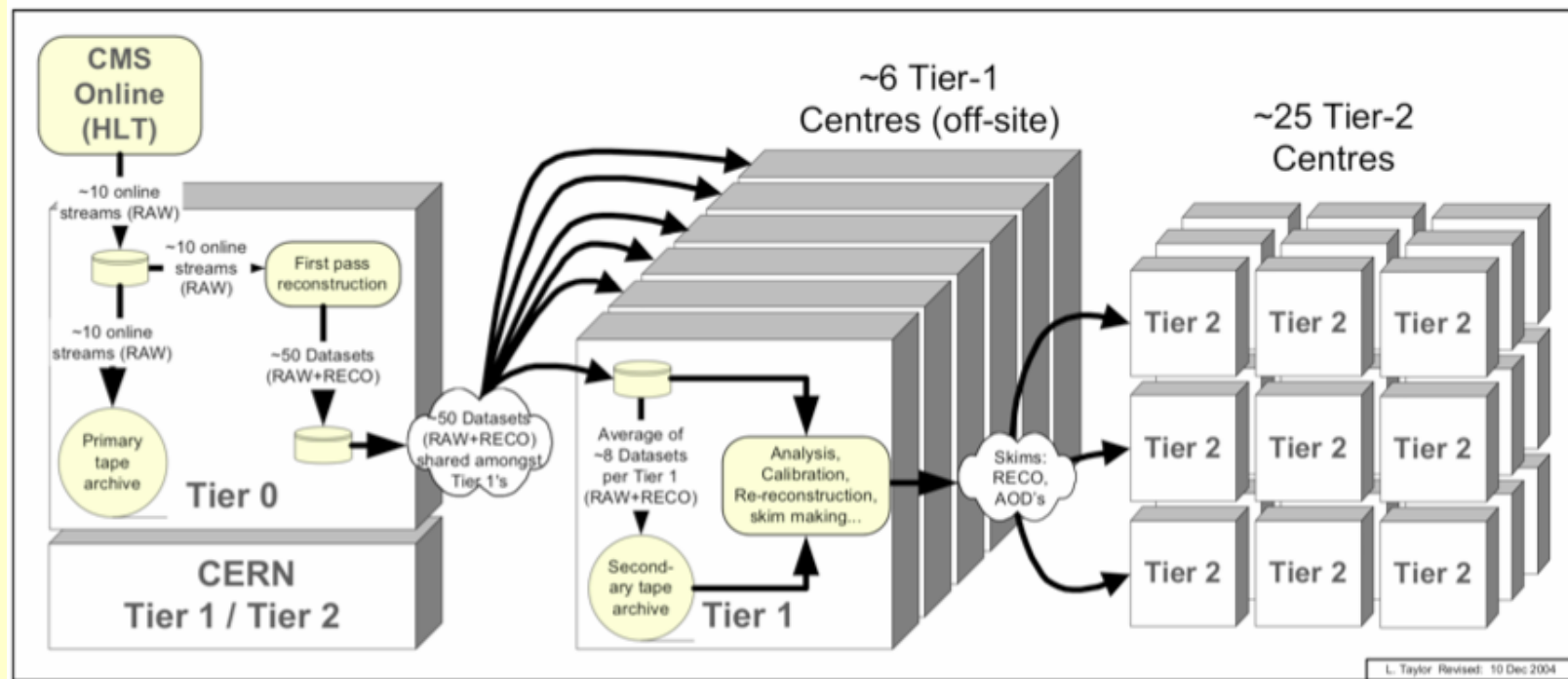
Physics Requirements : Designed Data Rates



Physics Requirements : Expected Scheduled Processing, Flows

- **Online processing**
- **Offline Re-processing**
- **Monte Carlo Simulations**

See talk by L.A.T.Bauerdick



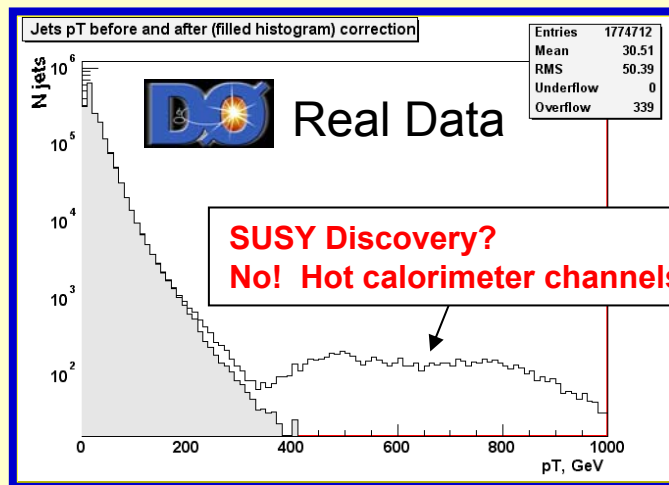
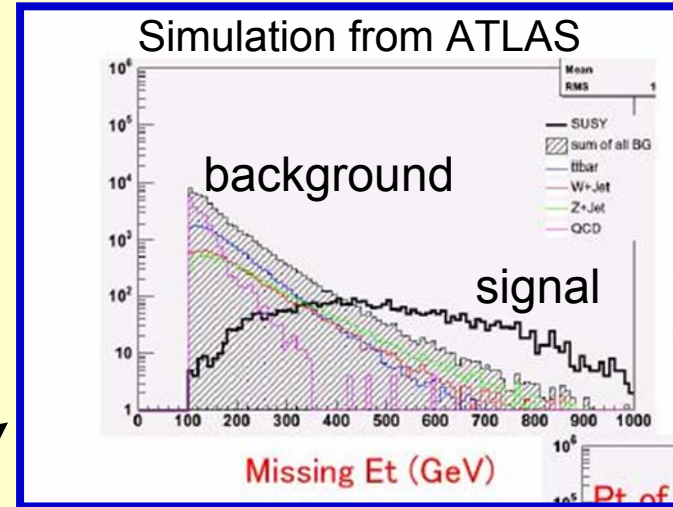
- **Hierarchical Data Flow (Grid) Strategy**

Physics Requirements : Expected Data Analysis Behaviour

- **Coherent “Peak” Patterns → Resonant Behaviour**
 - Driven by conference deadlines, rumours, etc
- **Data Placement & Replication**
 - Strategic decisions by the Collaboration
 - **Petabyte transfers**
 - Tactical decisions by users
 - **Terabyte transfers**
- **Job Placement & Parallelisation**
 - Collaboration-level scheduled/sustained workflows
 - User-level ad-hoc/bursty workflows
- **Data Access (based on trigger streams)**
 - RAW (1.5MB) → RECO (0.25 MB) → AOD (0.05 MB) → TAG (0.01 MB)
 - Planned: RAW (only scheduled), RECO (some scheduled, some ad-hoc), AOD (primarily ad-hoc), TAG (only ad-hoc)
 - Early Reality: RAW (significant access), RECO (significant access)
- **Collaboration controlled prioritisation required very early on!**

Physics Requirements : Discovery Mode Early On...

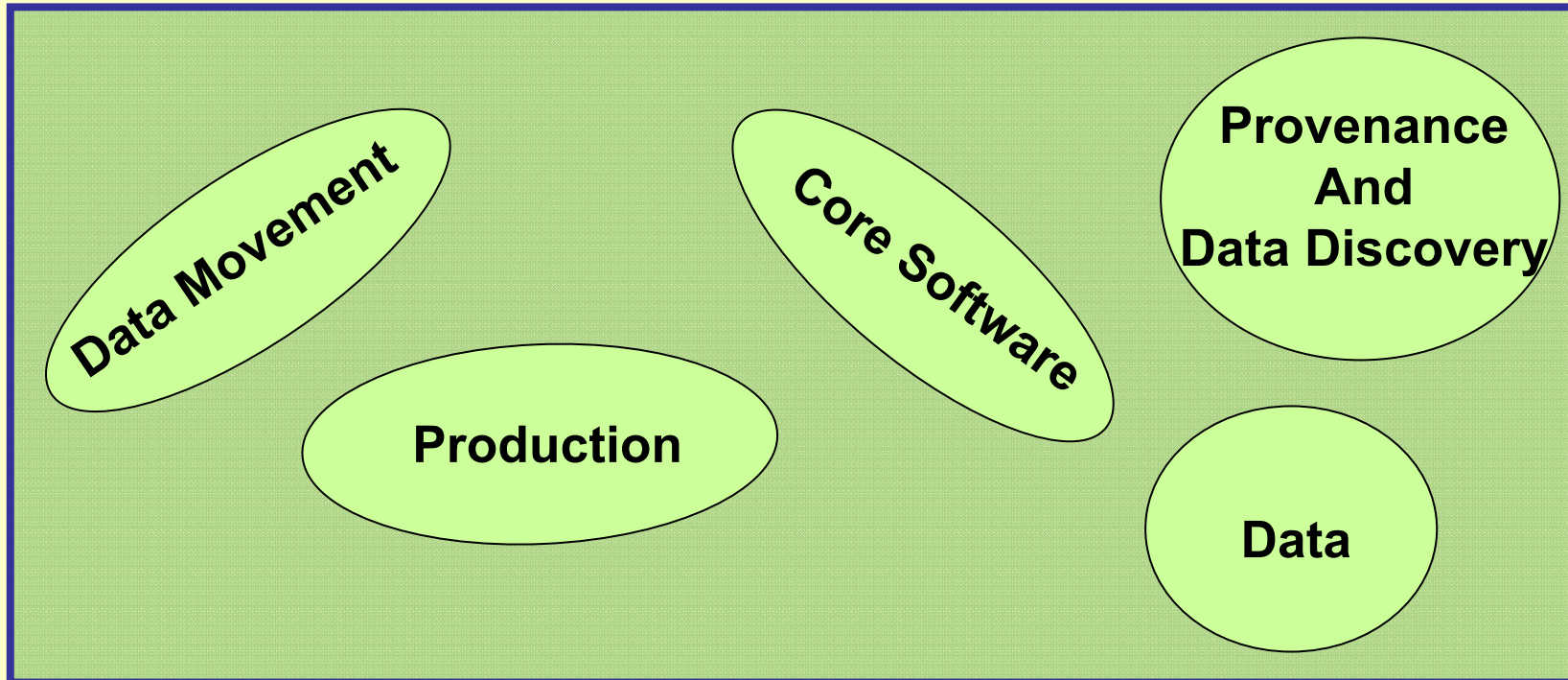
- **Example: Inclusive SUSY**
 - **Signal :** $O(\text{pb} - \text{fb})$
 - **Depends on mass scale**
 - **Background :** $O(\text{mb})$
 - **Signal/Background**
 - $\approx 10^{-9} - 10^{-12}$ (No selection)
 - $\approx 10^{-3} - 10^1$ (Use Trigger & Missing ET)



- **Data analysis requires a good understanding of the CMS Detector**
 - The early "hunt" is not so much searching for the "needle in a haystack"
 - It will be to "hunt" down detector problems!

CMS Software Stack

- **CMS has a rich and sophisticated set of software tools**

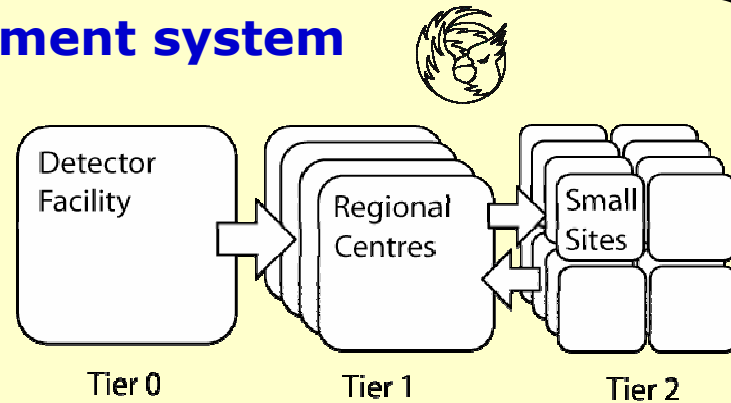


- **Main point for the slides following...**
 - **Data Analysis involves working a complex software environment!**

CMS Software Stack : Data Flow Framework

- **PhEDEx is a data replica management system**

- Multiple sources to Multiple sinks
- Monitors cost, latency, and rate
 - Enable future scheduling
 - Prioritize transfers

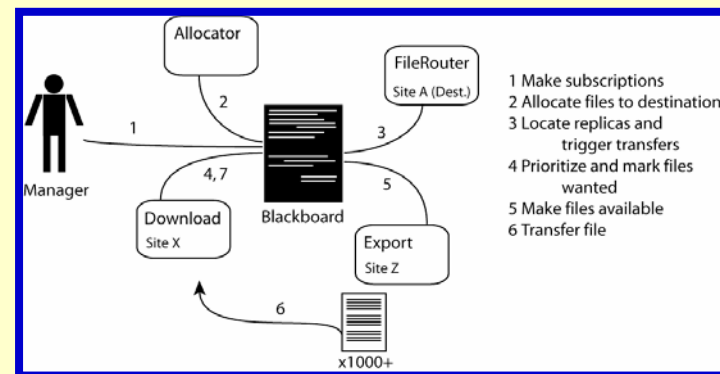


- **Manages distribution at "dataset" level**

- Where dataset is a set of files

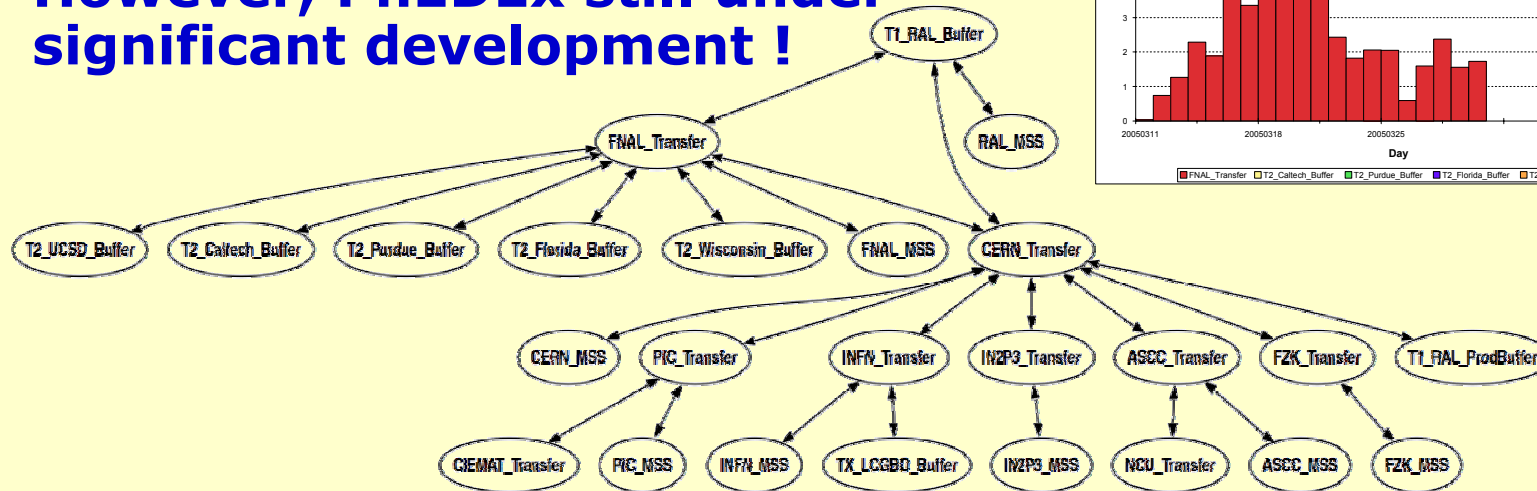
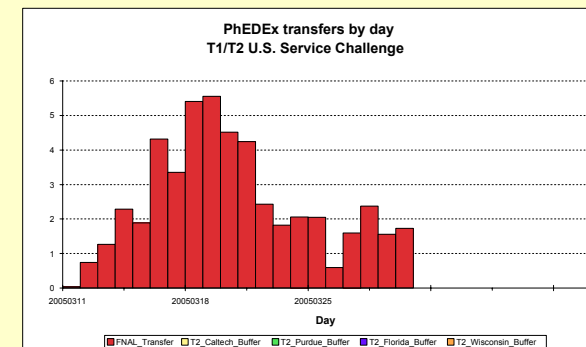
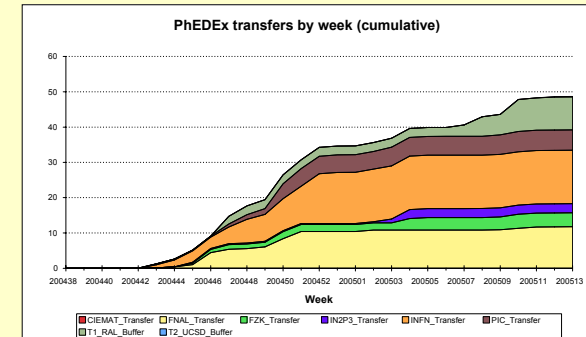
- **Attempts to bridge "traditional" and "Grid" data distribution models**

- traditional \Rightarrow large transfers to large sites, managed by hand
- grid \Rightarrow replication of data in response to user demand



CMS Software Stack : Data Flow Framework

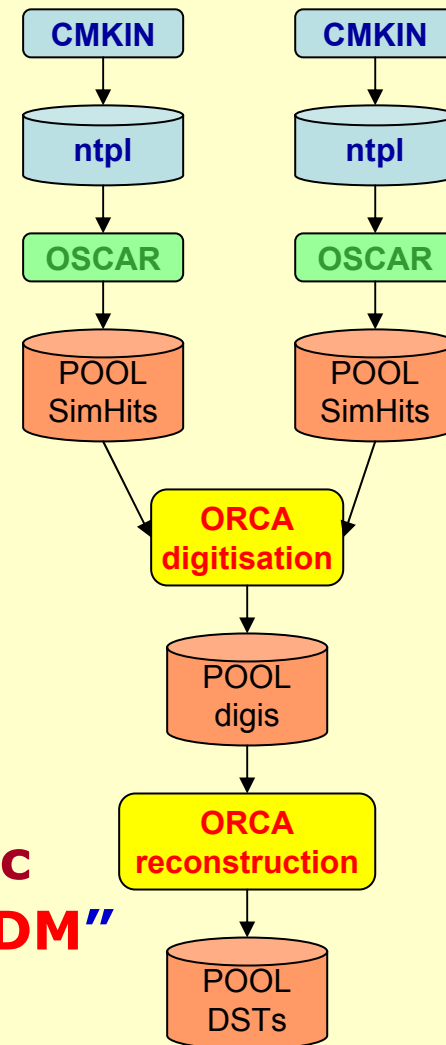
- **PhEDEx currently manages large-scale transfers for CMS**
 - Large-scale \Rightarrow $O(1000+)$ files per dataset
- **In production operation now**
 - ~ 70 TB known to PhEDEx, ~ 150 TB total replicated
 - Reaching nearly 20TB a month
 - Reaching up to 5TB a day
- **However, PhEDEx still under significant development !**



PhEDEx Routing Table

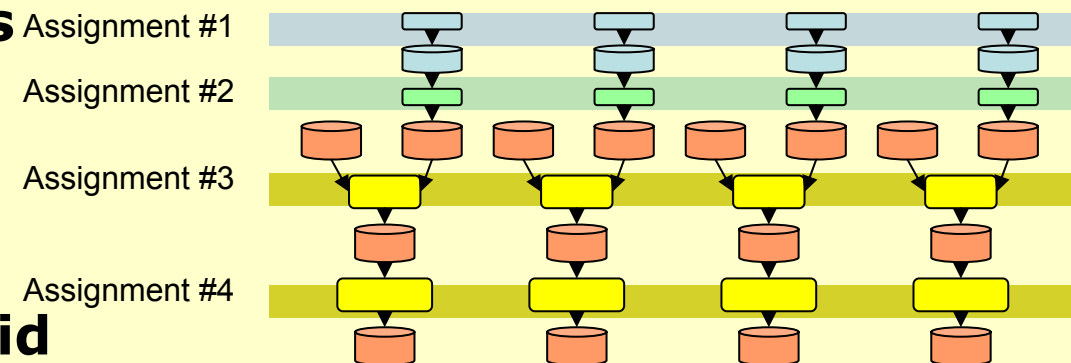
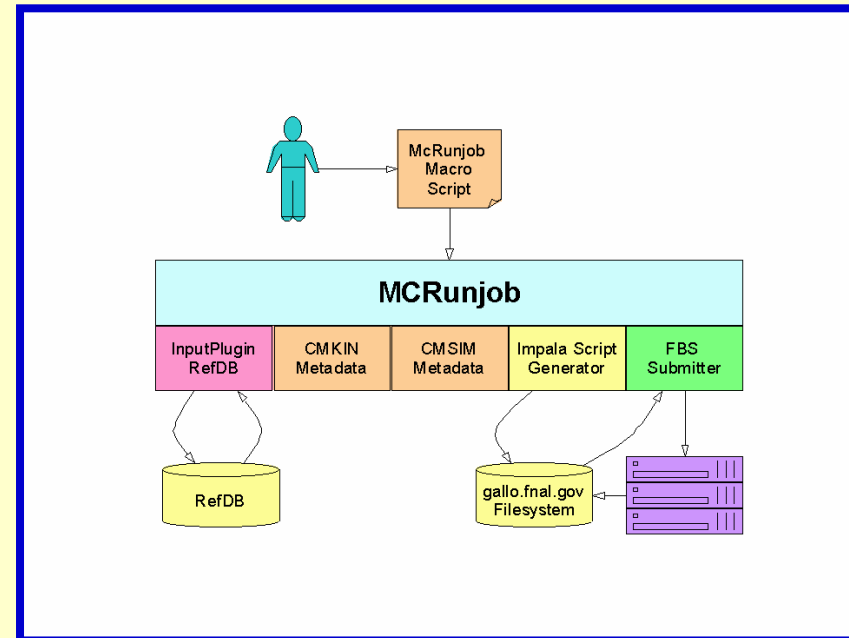
CMS Software Stack : Core Software

- **Environment Management : SCRAM**
- **Generation : CMKIN**
 - Interface to Monte Carlo p-p Generators
 - CPU bound (depends on physics process), small output size
- **Simulation : OSCAR**
 - GEANT 4 based full simulation of CMS Detector
 - CPU intensive, large output size
- **Reconstruction : ORCA**
 - Sophisticated SDK
 - Many applications:
 - **Digitisation, Reconstruction, etc**
- **Analysis Framework : ORCA, "New EDM"**
- **Visualisation: IGUANA**



CMS Software Stack : Distributed Processing Environment

- **RefDB**
 - Records production requests
 - Describes production assignments
- **MCRunJob**
 - Queries RefDB for assignment
 - Parallelizes production workflow
- **MOP**
 - Maps & submits workflow to grid
- **BOSS**
 - Monitors workflow on grid



CMS Software Stack : Analysis Development Environment

- **ORCA**

- **Software Development Kit**

- **Not a single application**
- **Collection of shared object libraries**

- **Provides sole access to Event Data**

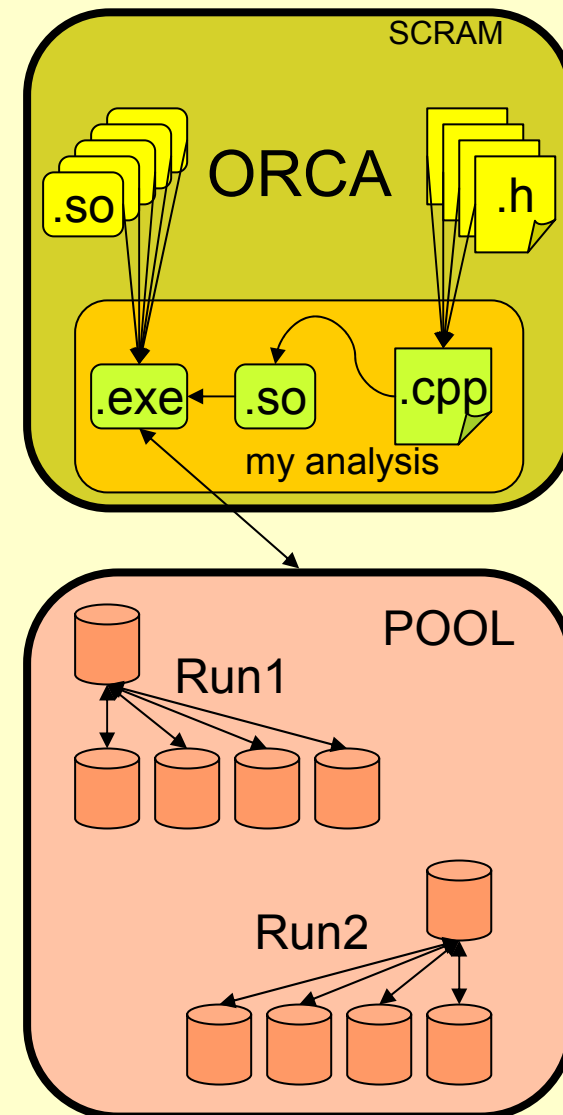
- **Event Data Model**

- **Separate file containers for**

- **Metadata**
- **Event data (object collections)**
- **Each type of event data**
- **Many small files: Write once, read many times**

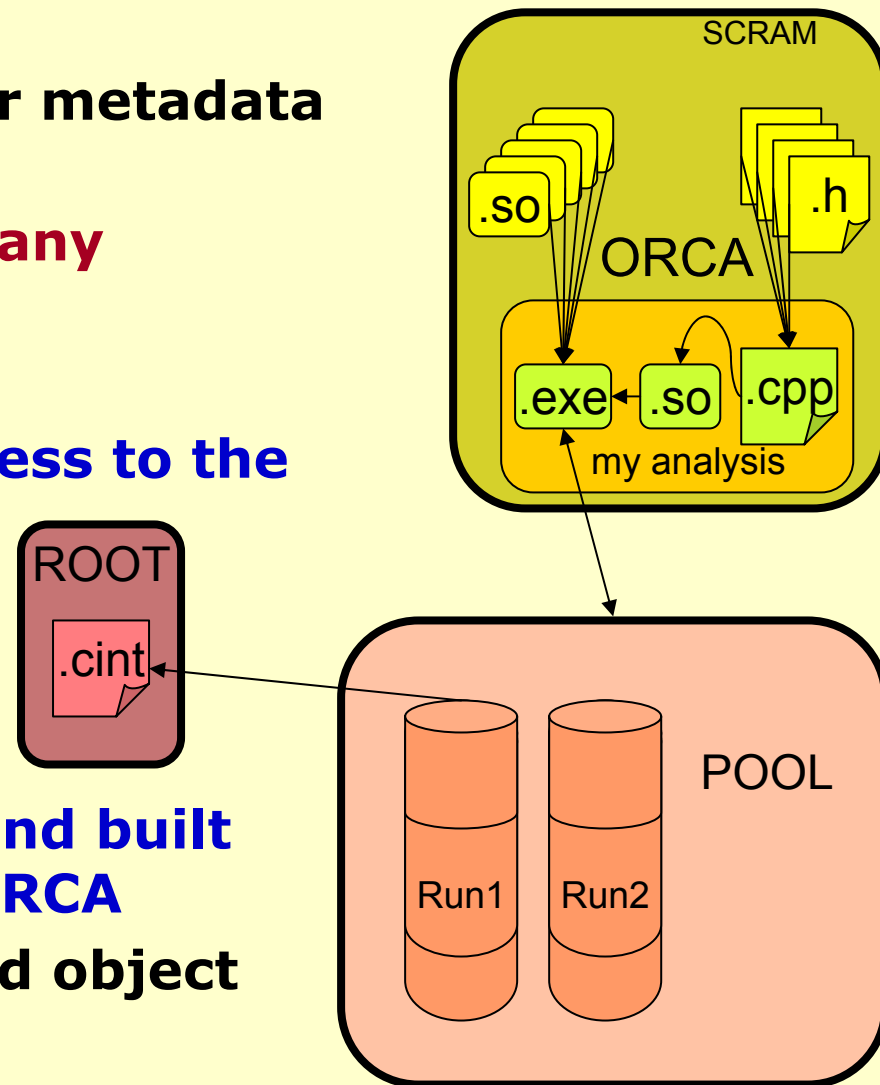
- **All analysis code is written and built leveraging ORCA**

- **Analysis specific shared object library & executable**



CMS Software Stack : Analysis Development Environment

- **New Event Data Model**
 - **Single file container for metadata and data**
 - **Write many, read many times**
 - **Far fewer files**
- **Enables more flexible access to the event data**
 - **Bare Root level**
 - **"ORCA" Lite level**
 - **Full "ORCA" level**
- **Analysis code is written and built leveraging Root and/or ORCA**
 - **Analysis specific shared object library & executable**



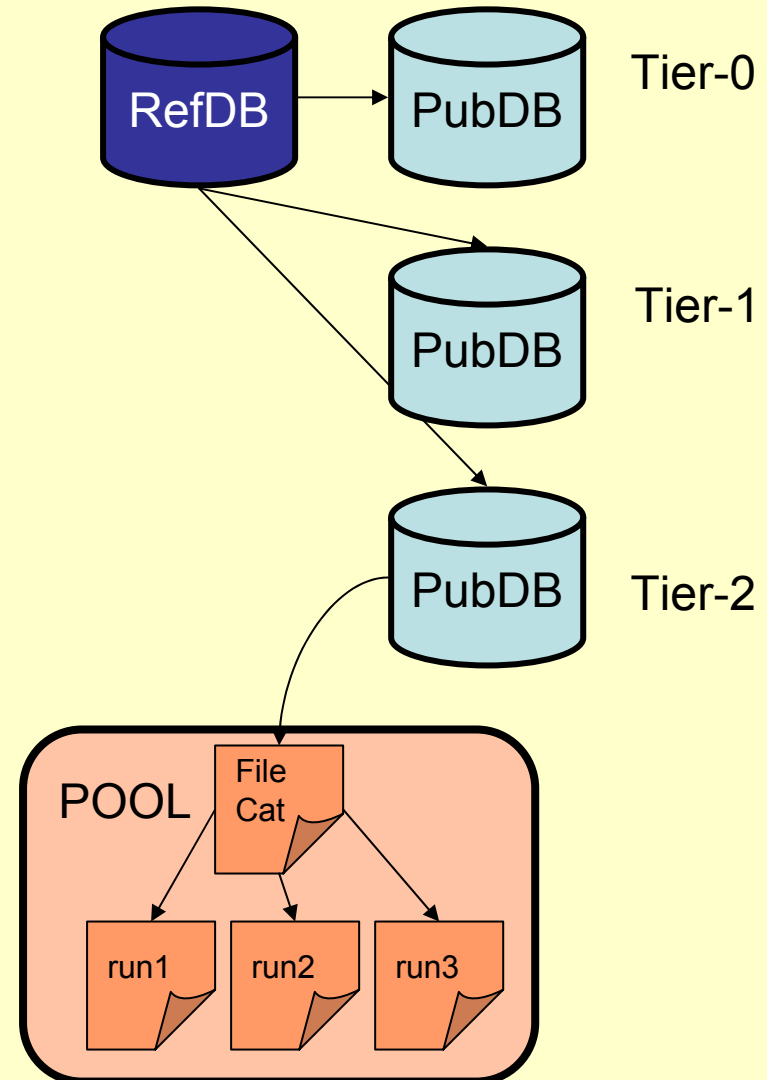
CMS Software Stack : Metadata and Bookkeeping

- **RefDB**

- Contains dataset provenance information
- Contains PubDB locations for all datasets
- Currently single instance located at the T0

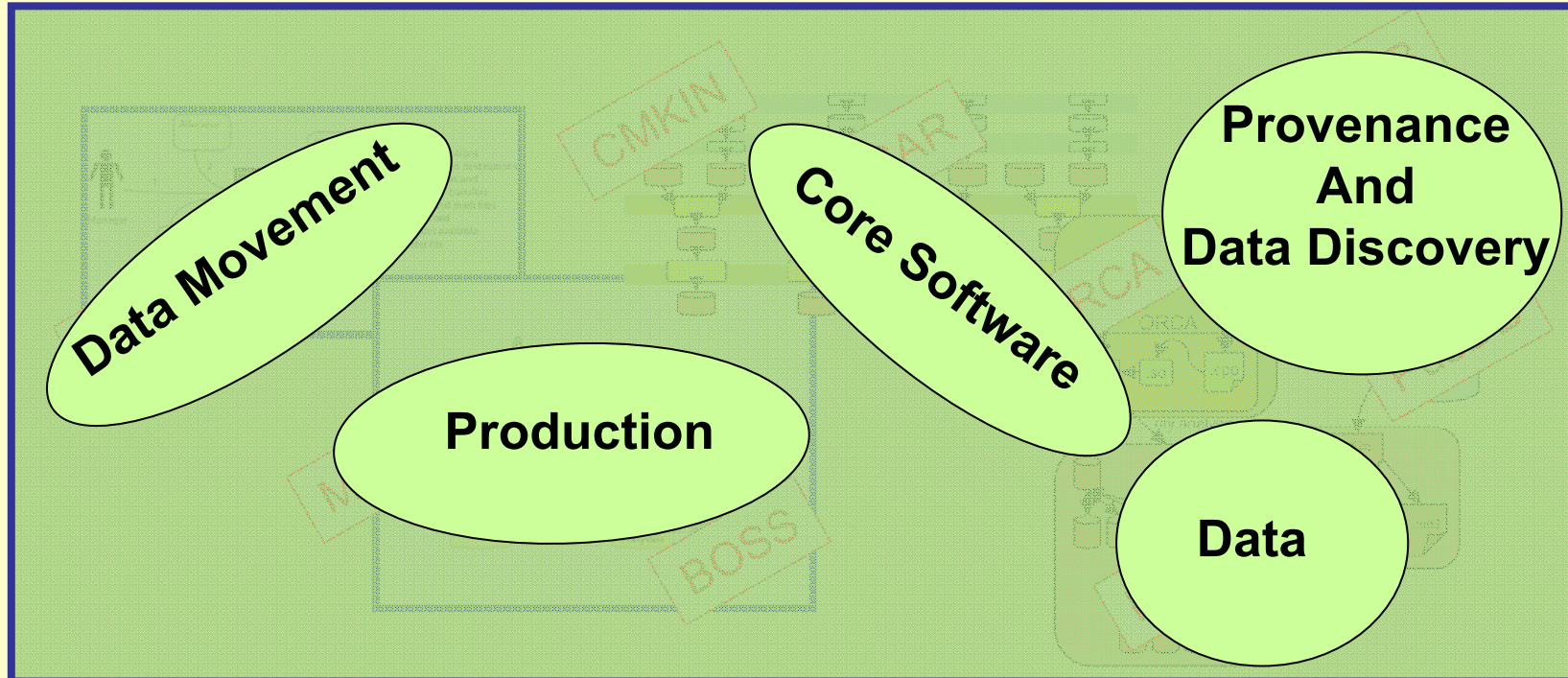
- **PubDB**

- Contains the path (contact string) to a local replica catalogue
 - **PoolFileCatalogue**
- Located at each site which publishes "official" CMS datasets



Interlude : Statement of the Problem

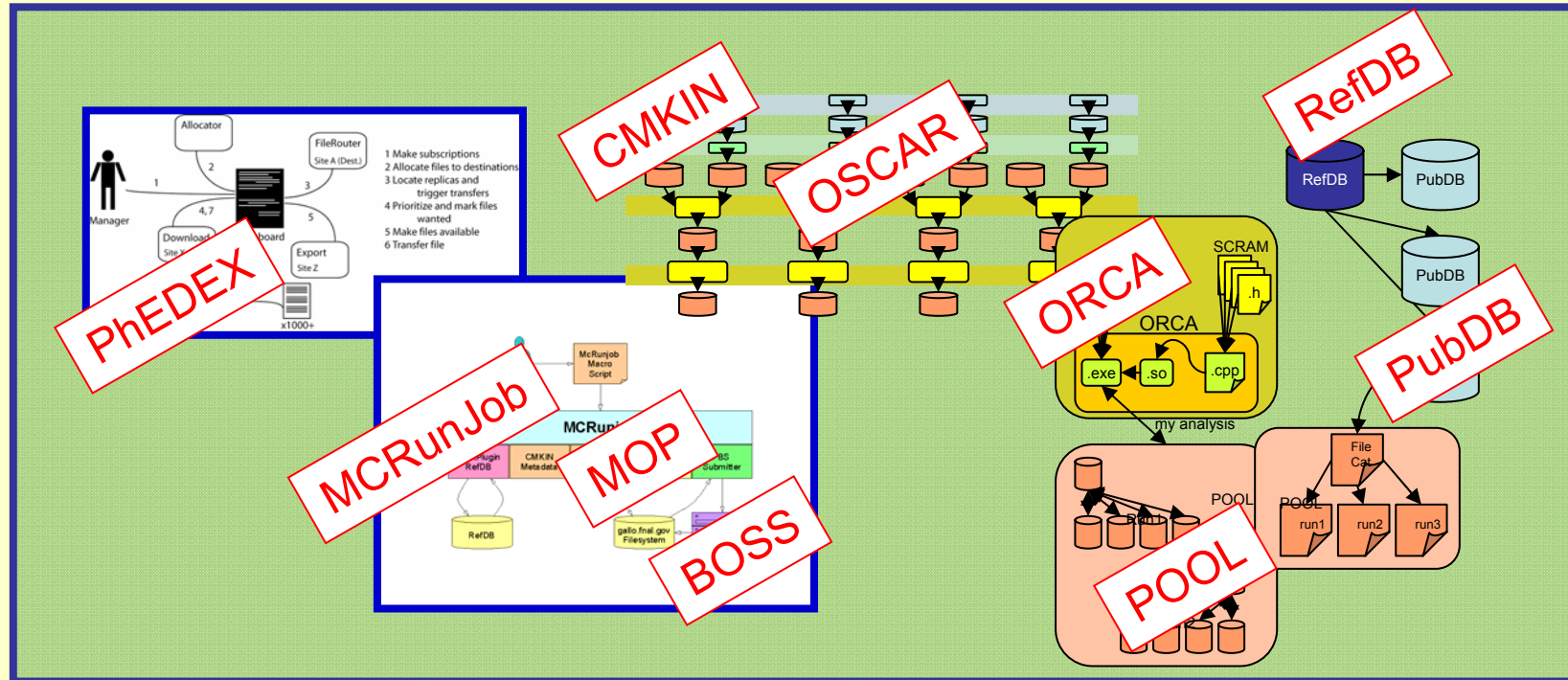
- **CMS has a rich and sophisticated set of software tools**



- **Many/Most essential components exist for Local Data Analysis**

Interlude : Statement of the Problem

- CMS has a rich and sophisticated set of software tools



- Many/Most essential components exist for Local Data Analysis
- But the tools address many individual requirements, not the system as a whole
- **E2E Distributed Data Analysis**
 - Requires service "glue" for creating a global, coherent system

UltraLight Services : Clarens

- **Backbone**

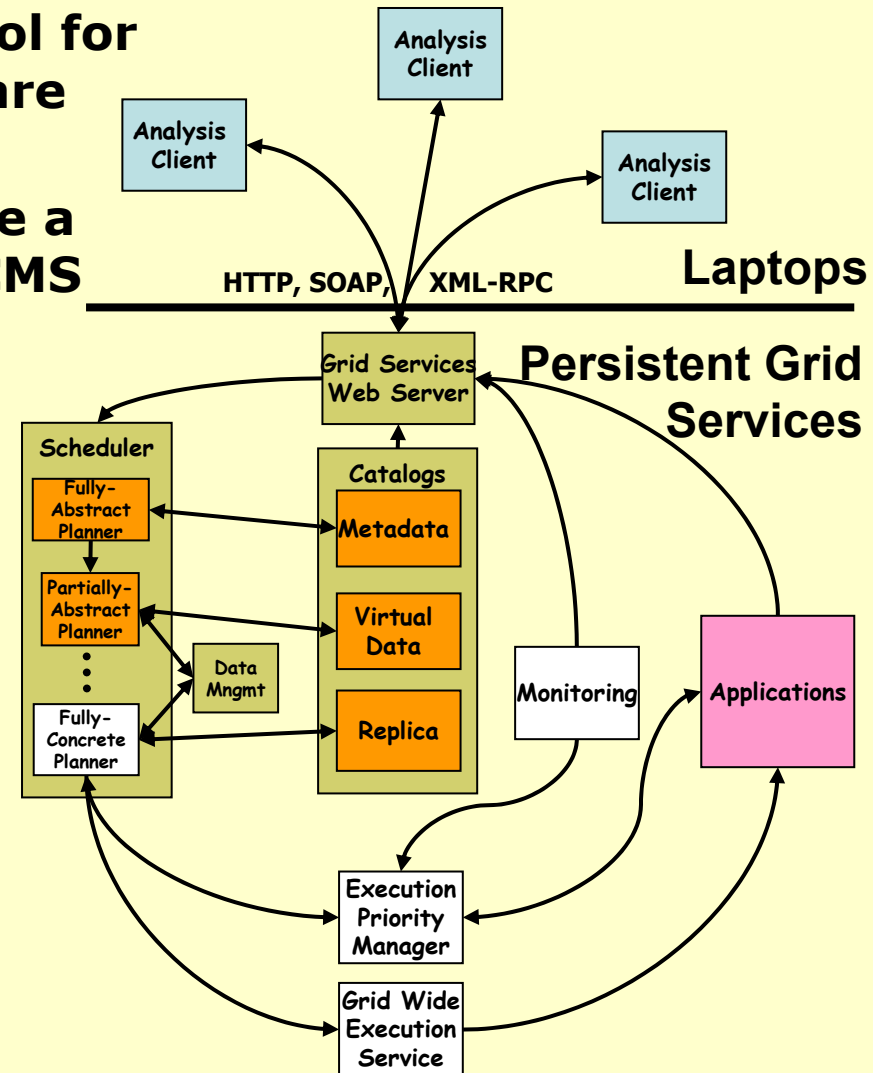
- Provides a systematic tool for exposing the CMS software stack as "services"
- Allows one to orchestrate a system from individual CMS components

- **Generic Service Oriented Architecture**

- Service discovery
- Software discovery
- File transfer
- VO-management
- Scheduling

- **Specific Clients**

- Root
- CODESH



Integrated System : Conceptual Phases See talk from F.Wuerthwein

Laptop

Me

Persistent
Services @ T2

My friends

Global
Computing
Grids

*Anonymous
world*

- **Physicists will rely on services at T2 they work with no matter where they are working from.**
- **All the “heavy lifting” is done by my friends.**

Integrated System : Workload Management : MCPS

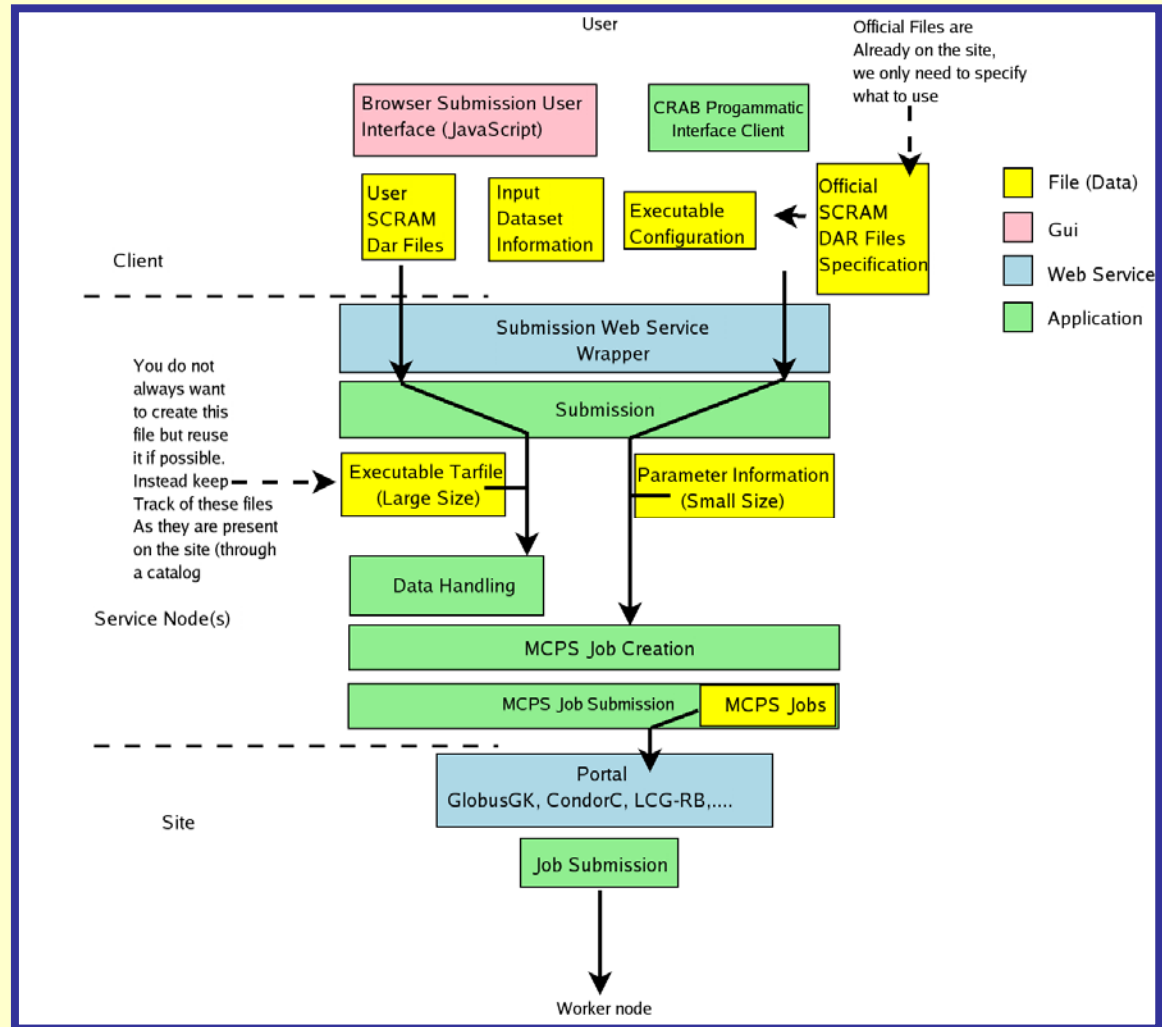
- Monte Carlo Production Service

See talk from G.Graham

Laptop

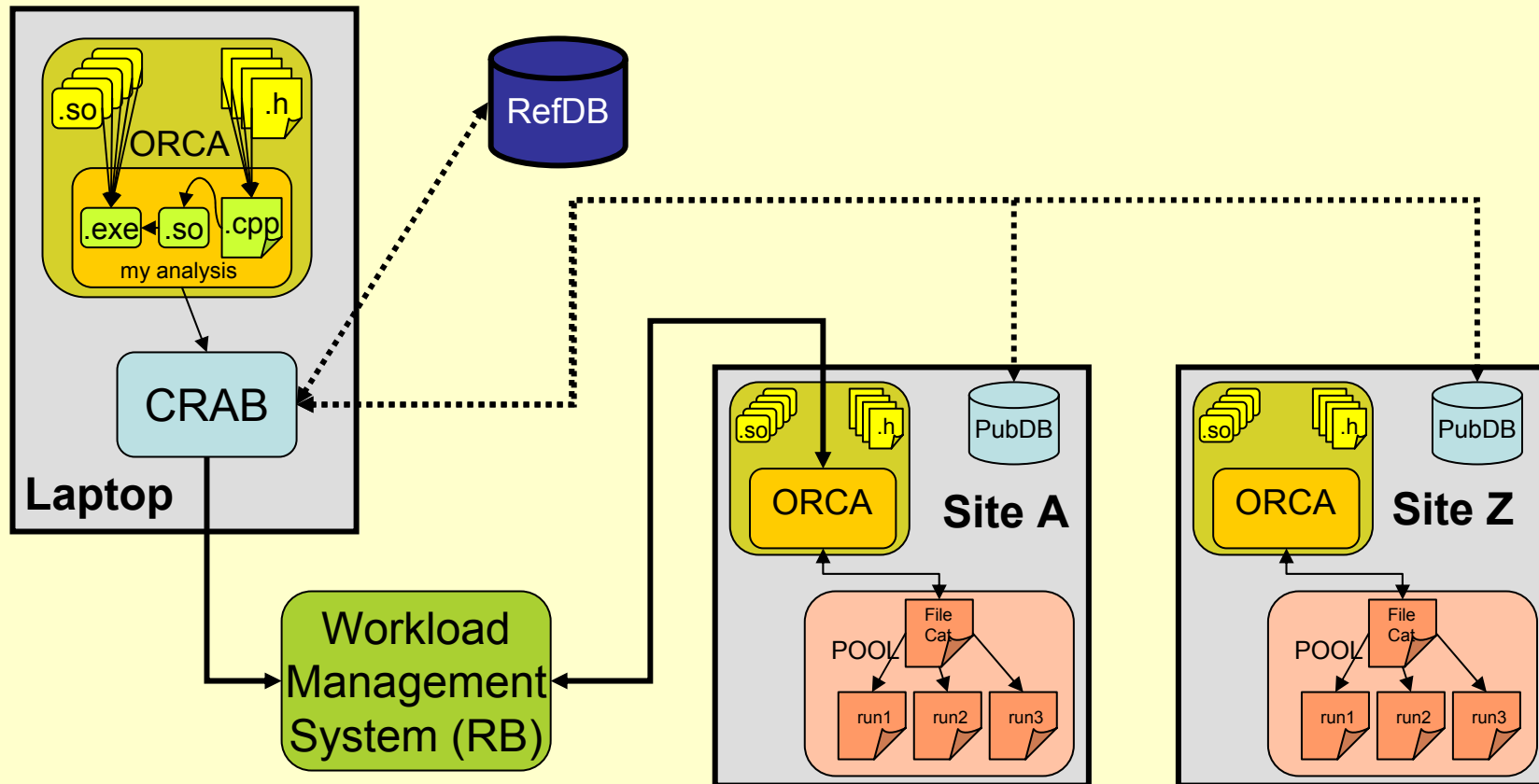
Tier-2

Anonymous
Grid site
Anywhere.



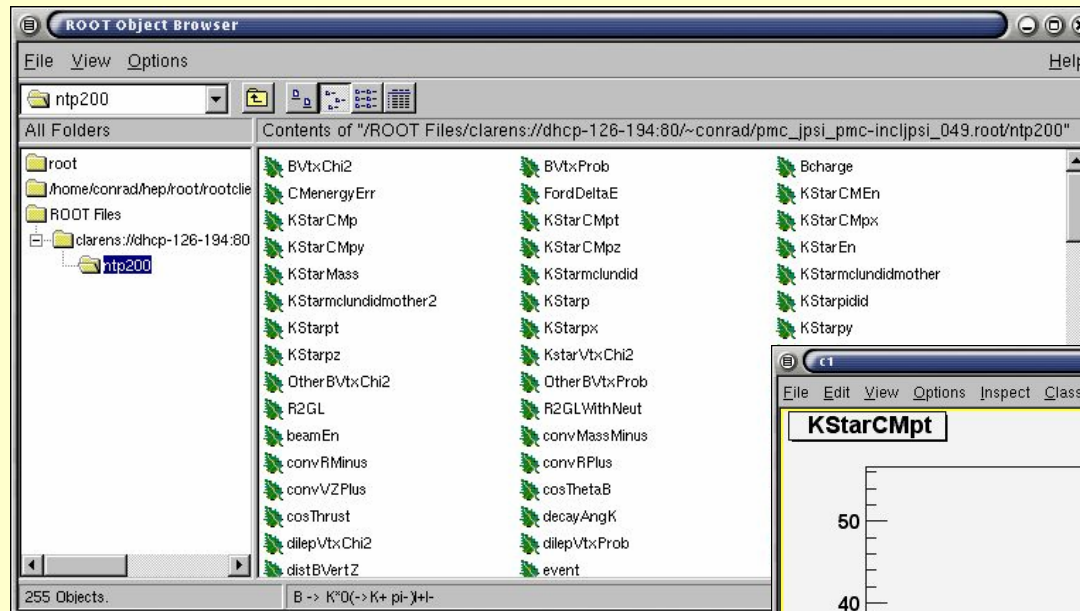
Integrated System : Workload Management : CRAB

- **CMS Remote Analysis Builder (CRAB)**
 - **Allows FULL use of CMS analysis software**
- **Requires that ORCA be remotely installed on each CE**



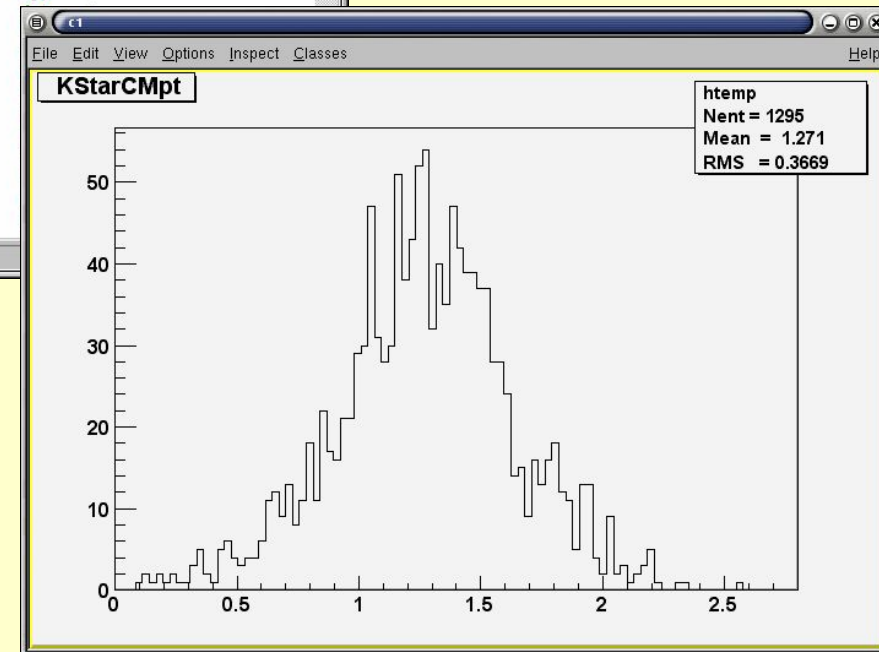
Integrated System : Root Clarens-Client

- Provides interactive remote access to Root files via Clarens File Service



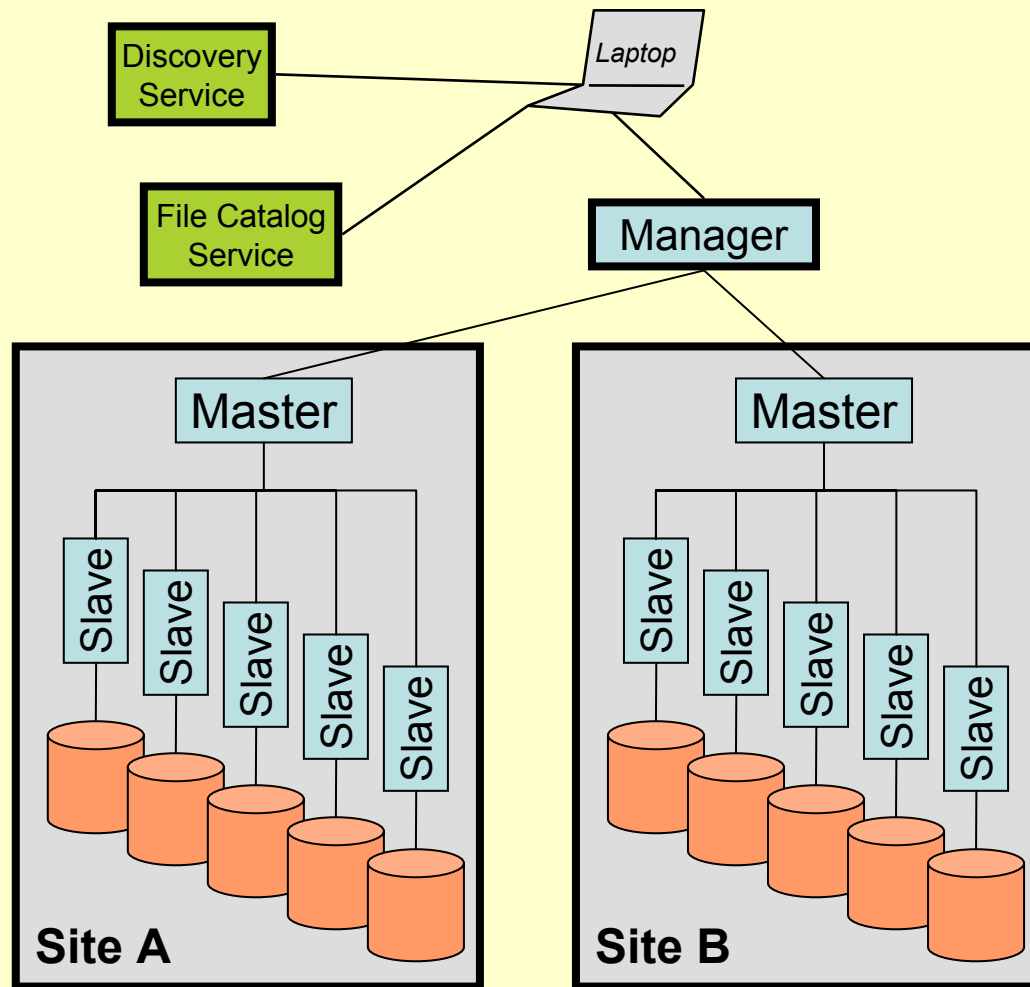
Now possible in CMS
via "New" EDM !

- No need for any specific CMS Software on CE !
- Provides standard Grid Security Infrastructure Authentication



Integrated System : Proof Enabled Analysis Cluster - PEAC

- **Interactive Physics Analysis of Large Datasets Using Parallel Processing and I/O**



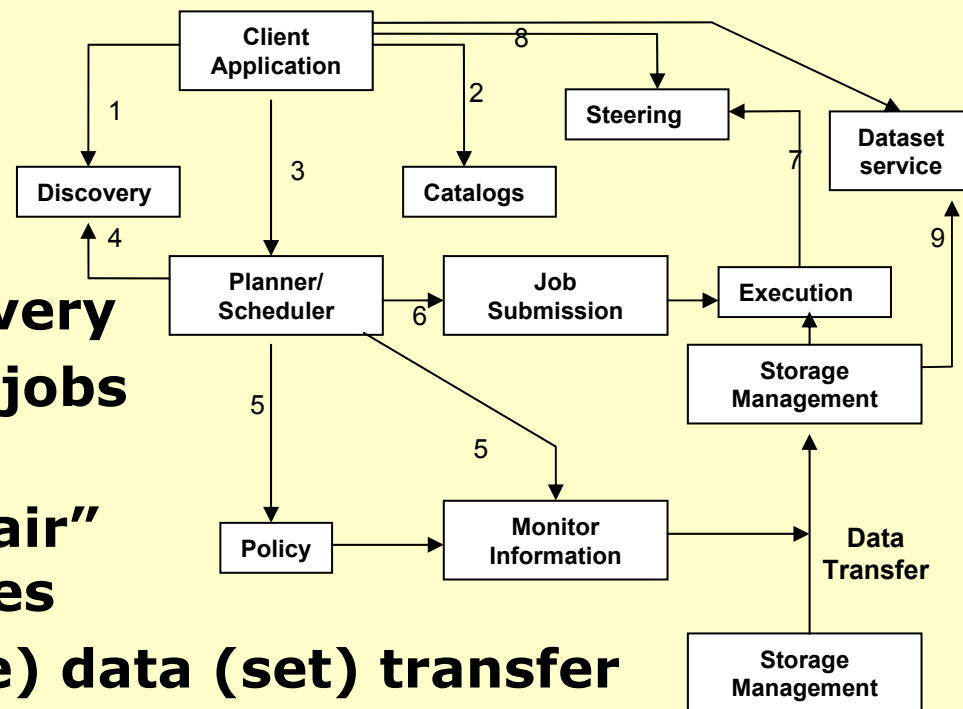
- **Transparency, Adaptability, Scalability**
 - **Adapt to Network and System Performance**
 - **Support Heterogeneous Systems**
 - **Efficient and Reliable**
- **Multi Tier architecture, Data Locality**
- **“New” EDM enables PEAC for CMS!**

Integrated System : End-to-End

See talk from F.v.Lingen

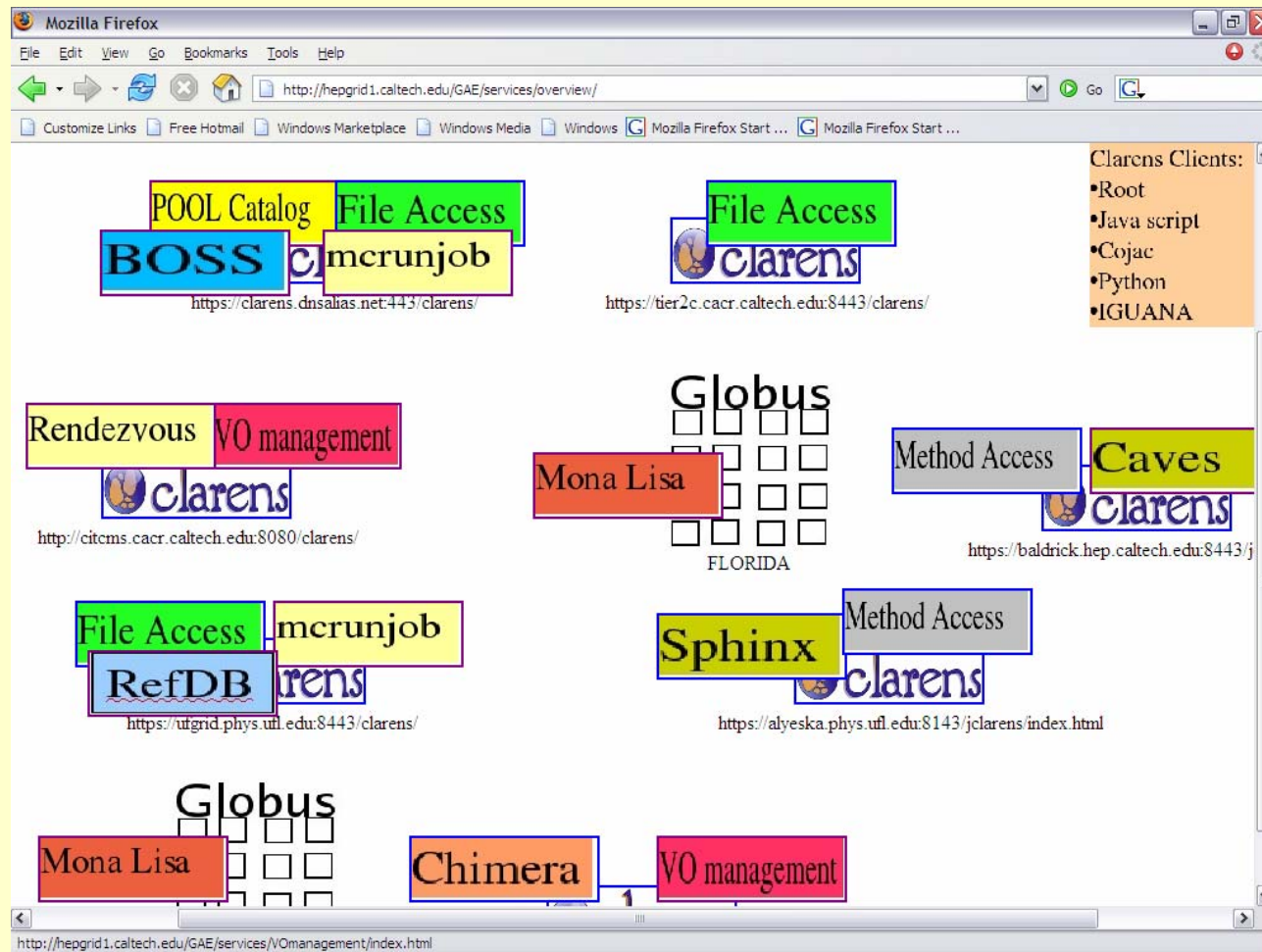
- **CMS Software stack is complex and still developing**
 - **Integration work is challenging**
- **Generic Service Oriented Architecture crucial for integration**

- **Catalogs to select datasets,**
- **Resource & Application Discovery**
- **Schedulers guide jobs to resources**
- **Policies enable "fair" access to resources**
- **Robust (large size) data (set) transfer**



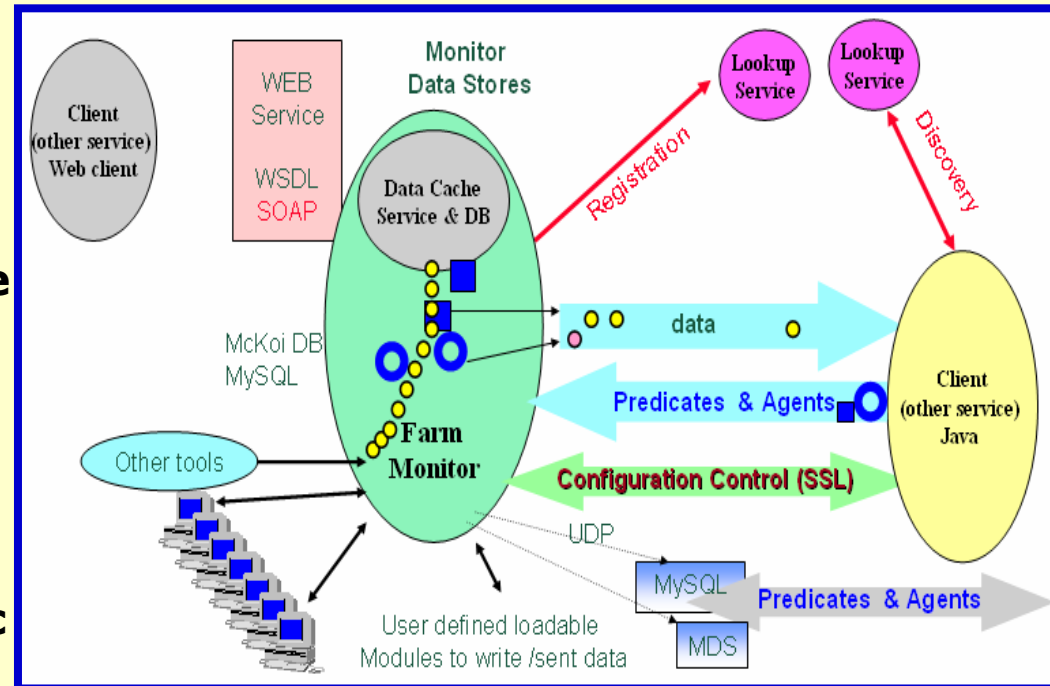
Integrated System : Grid-enabled Analysis Testbed

- **Distributed CMS services using the Clarens Web-service Framework**



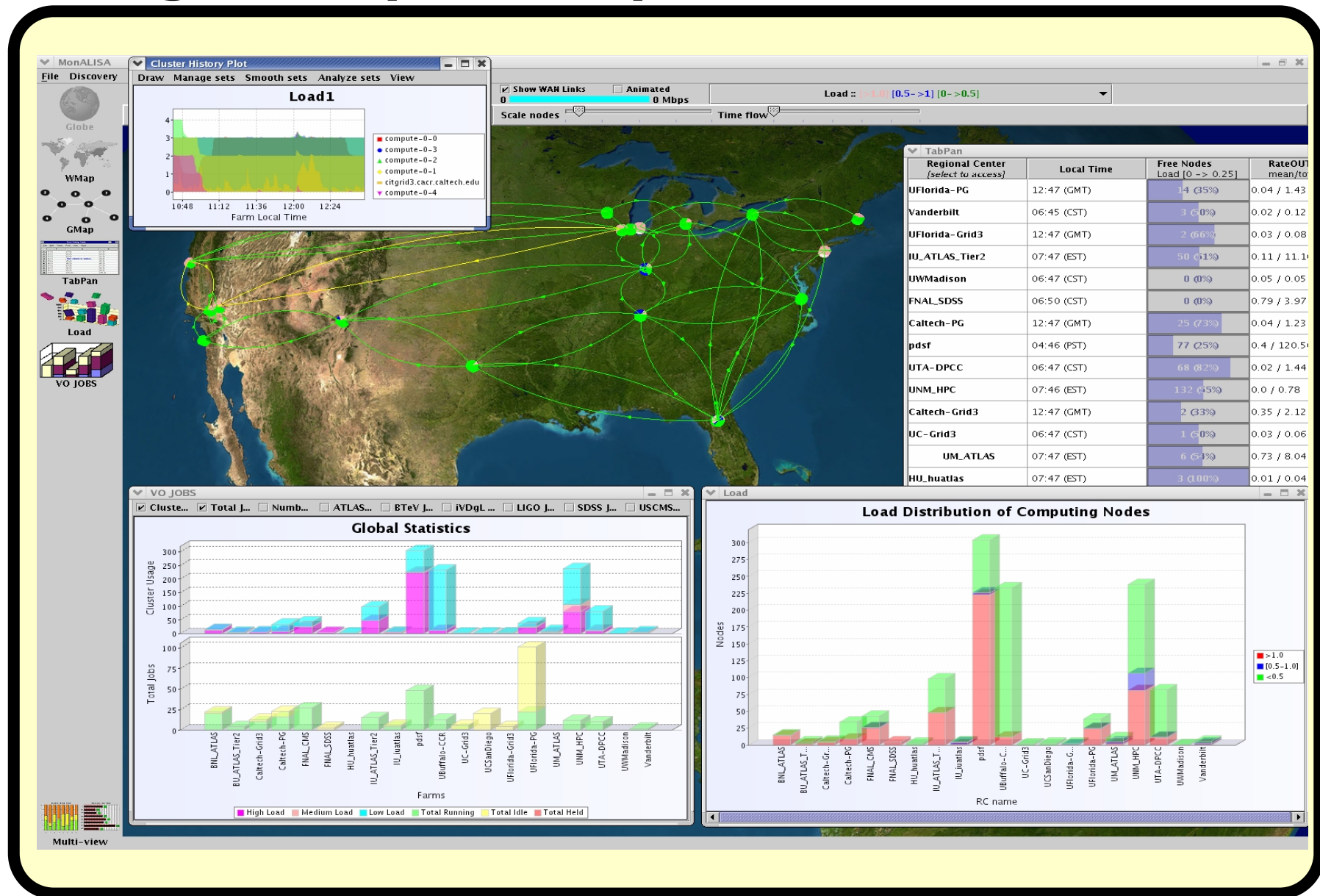
UltraLight Services : MonALISA

- **MonALISA able to dynamically**
 - **register & discover**
- **Services are self describing**
- **Code updates**
 - **Automatic & secure**
- **Dynamic config for services**
 - **Secure Admin Interface**
- **Active filter agents**
 - **Process data**
 - **Application specific monitoring**
- **Mobile agents**
 - **decision support**
 - **global optimisations**
- **Personal agent: LISA**
 - **Decision support**
 - **Automatic optimisation**



- **Integrated with Clarens WS Backbone**
- **Provides a "Dashboard" for User Activity**
 - **Resource status**
 - **Network status**
 - **Application status**

Integrated System : System "Dashboard"



Integrated System : Application Monitoring : ApMON



Summary

- **Data Analysis is a “search and discovery” process**
 - Entails the effective use of a complex set of components
 - CMS applications are rich in functionality and diversity
 - **Still rapidly evolving !**
- **Grid-enabled CMS Analysis requires an E2E Integrated Approach**
 - **Generic System Design**
 - **Un-affected by (inevitable) evolution of applications**
 - **Service Oriented Architecture exists !**
 - **Robust System Implementation**
 - **Ability to set priorities**
 - **Command and control features**
 - **Majority of effort is now in Integration**
 - **Expose CMS Software as high-level System Services**