CMS at the LHC
&
Data Grid for Data Analysis

Shuichi Kunori

U. of Maryland / CMS

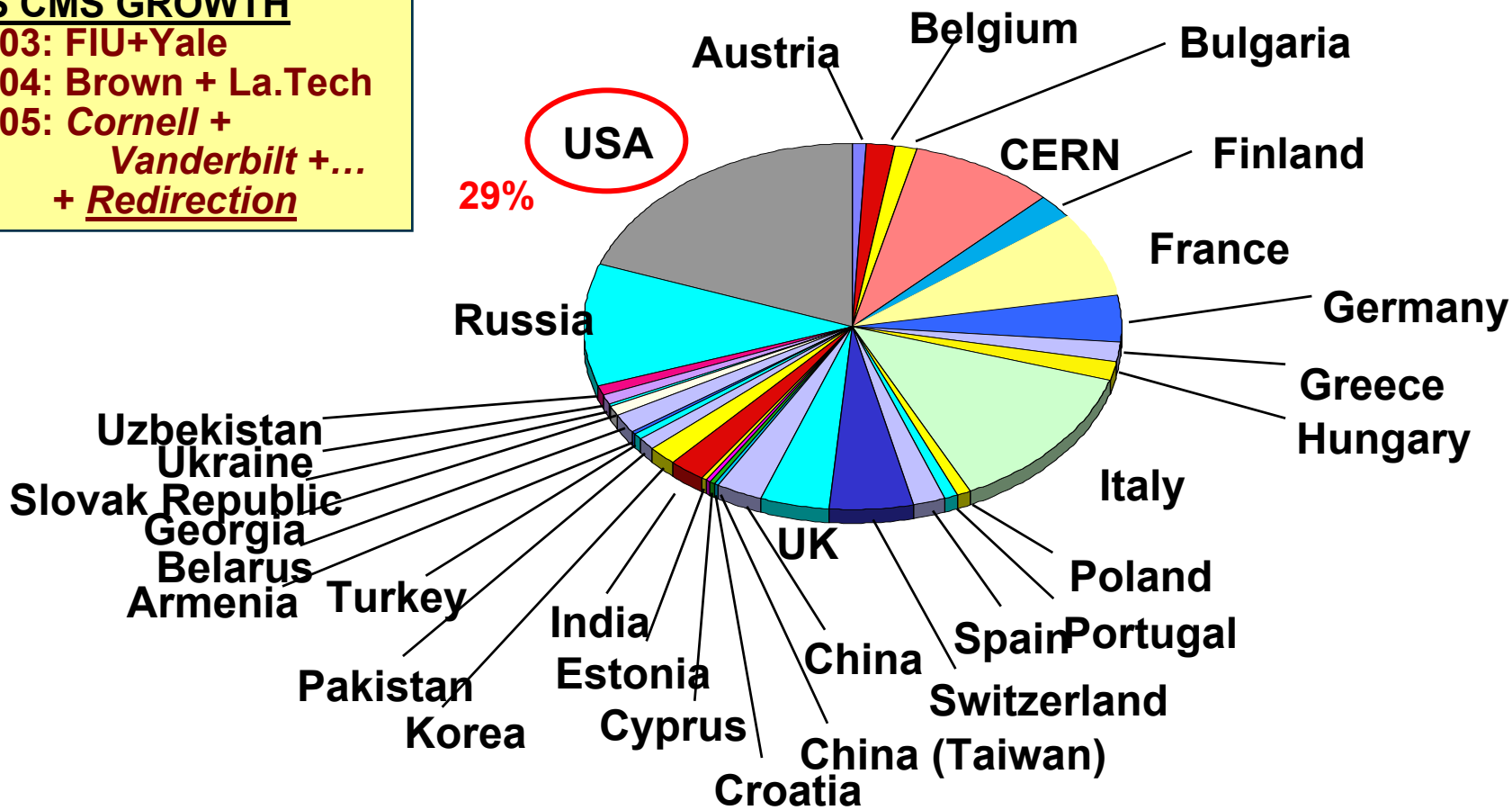
28-April-2005

Outline

CMS Collaboration
LHC
CMS Detector
Physics at LHC
Grid for CMS
Conclusion

The CMS Collaboration

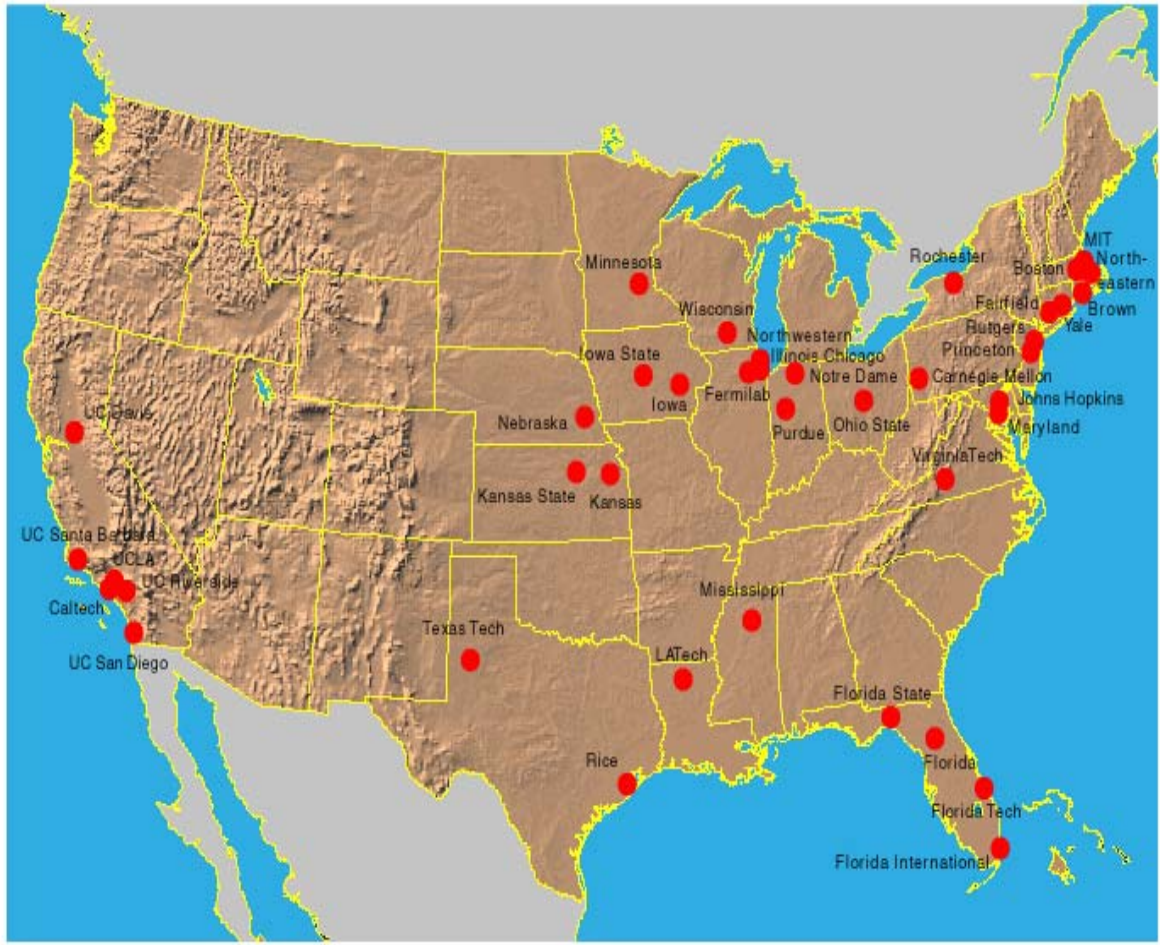
US CMS GROWTH
2003: FIU+Yale
2004: Brown + La.Tech
2005: *Cornell* +
Vanderbilt +…
+ *Redirection*



CERN + 31 Countries, ~2000 Collaborators

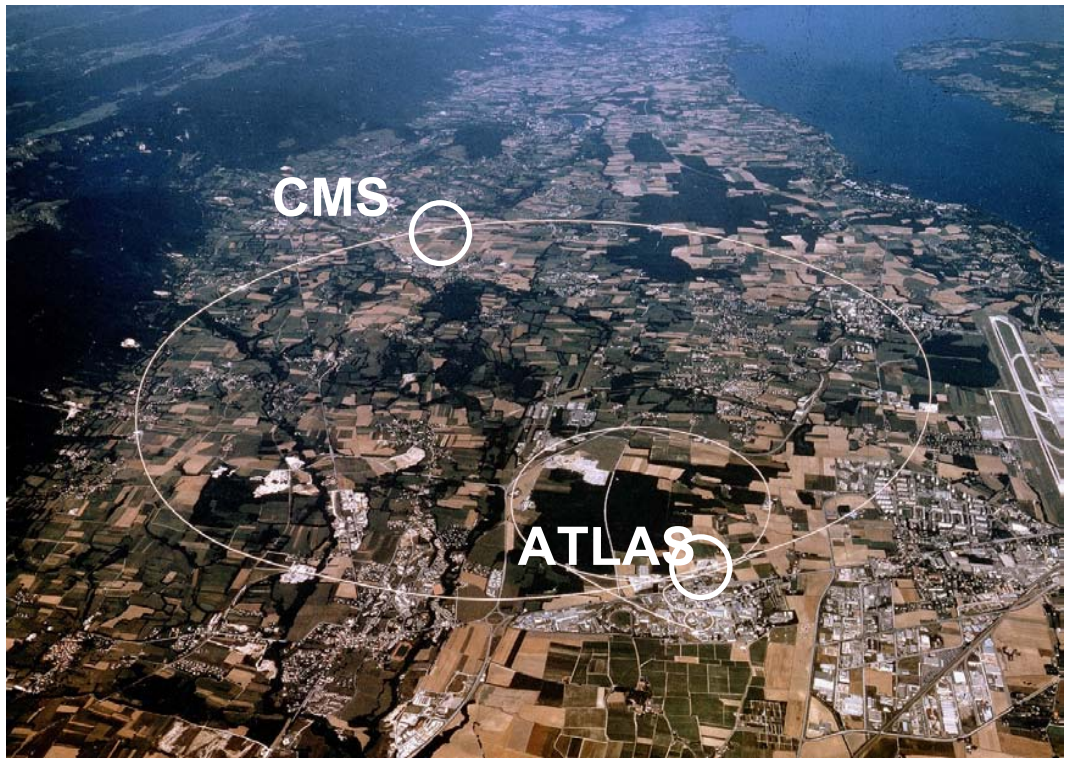
US CMS Collaboration

Subsystem	*Institutions*
Endcap Muon	UC Davis, UCLA, UC Riverside, Carnegie Mellon, Fermilab, Florida, Northeastern, Ohio State. Purdue, Rice, Wisconsin
HCAL	Boston, Fairfield, Fermilab, Florida Int'l, Florida State, Florida Inst. of Tech, Illinois-Chicago, Iowa, Iowa State, Maryland, Minnesota, Mississippi, Nebraska, Notre Dame, Princeton, Purdue, Rochester, Texas Tech
Trigger / DAQ	Boston, UC Davis, UCLA, UC San Diego, Caltech, Fermilab, Florida, Iowa, Maryland, MIT, Ohio State, Princeton, Rice Wisconsin
ECAL	Caltech, Fermilab, Minnesota, Northeastern, Princeton, Yale
Tracking	Brown, UC Davis, UC Riverside, UC Santa Barbara, Fermilab, Illinois-Chicago, Johns Hopkins, Kansas, Kansas State, Louisiana Tech, Mississippi, Nebraska, Northwestern, Rochester, Purdue, Rutgers, Texas Tech



40+ institutions, 500+ members

The LHC



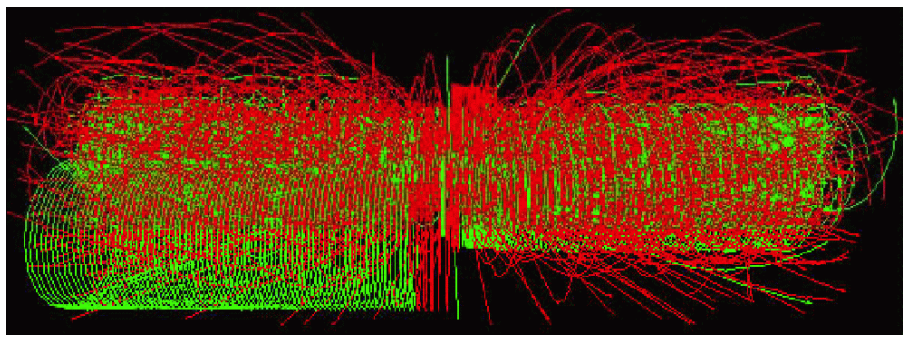
R = 4.5 Km
E = 7+7 TeV (pp)

crossing rate
=40MHz
(25nsec)

design luminosity
= $10^{34} cm^{-2} s^{-1}$

~20 pp interactions
per crossing
at design luminosity

First beams summer in 2007

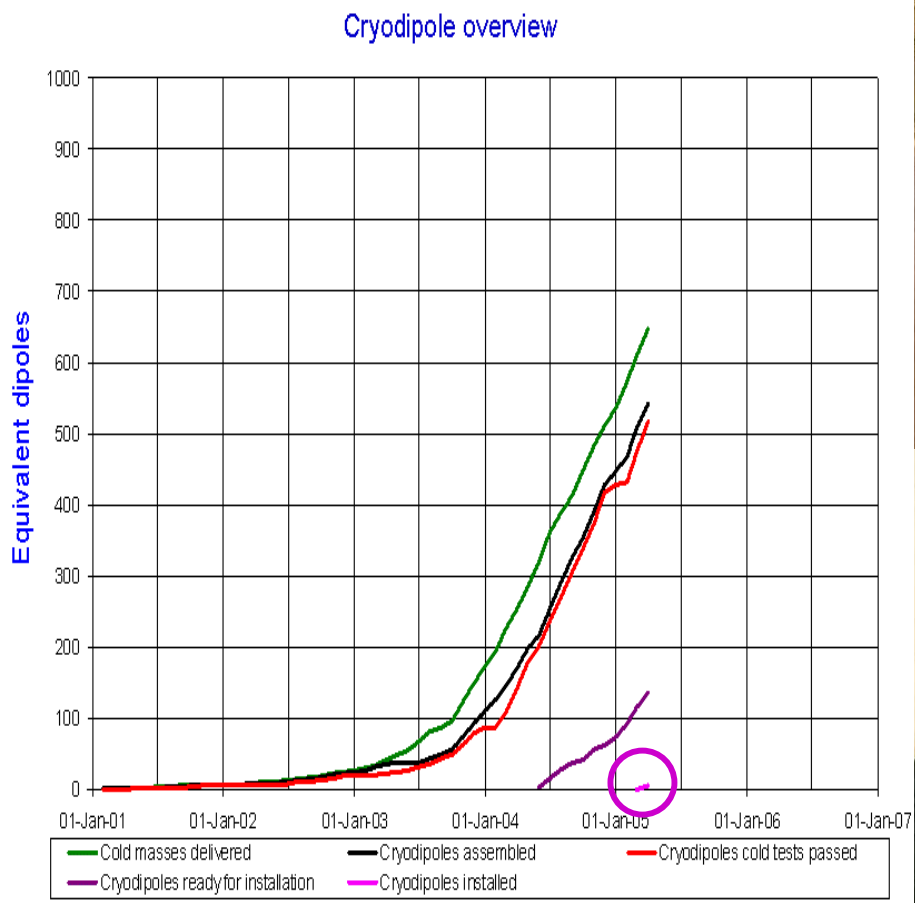


h → 4 μ with 20 min. bias evt.

LHC Dipole Magnet

LHC Progress Dashboard

Accelerator Technology Department

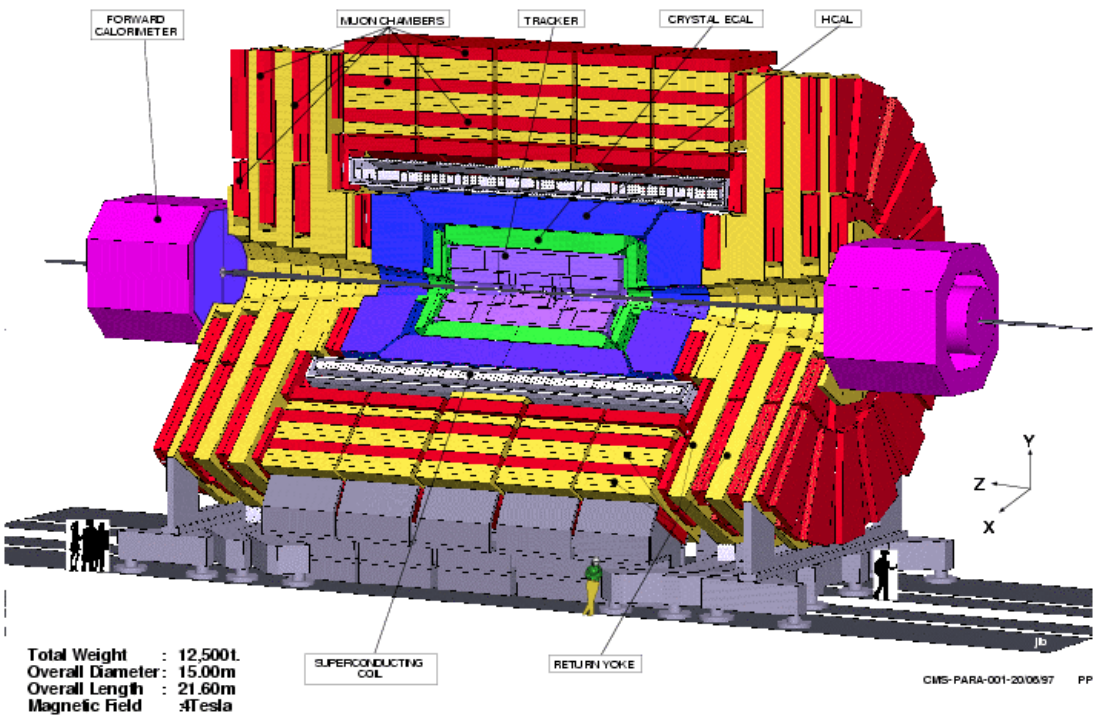


Updated 31 Mar 2005

Data provided by D. Tommasini AT-MAS

March 7, 2005

The CMS detector



Total weight	12500 t
Overall diameter	15 m
Overall length	21.6 m

All silicon tracker
- micro strips (10M ch)
- pixel (60M ch)
- (5.4m long, 2.4m Φ: $|\eta| < 2.4$)

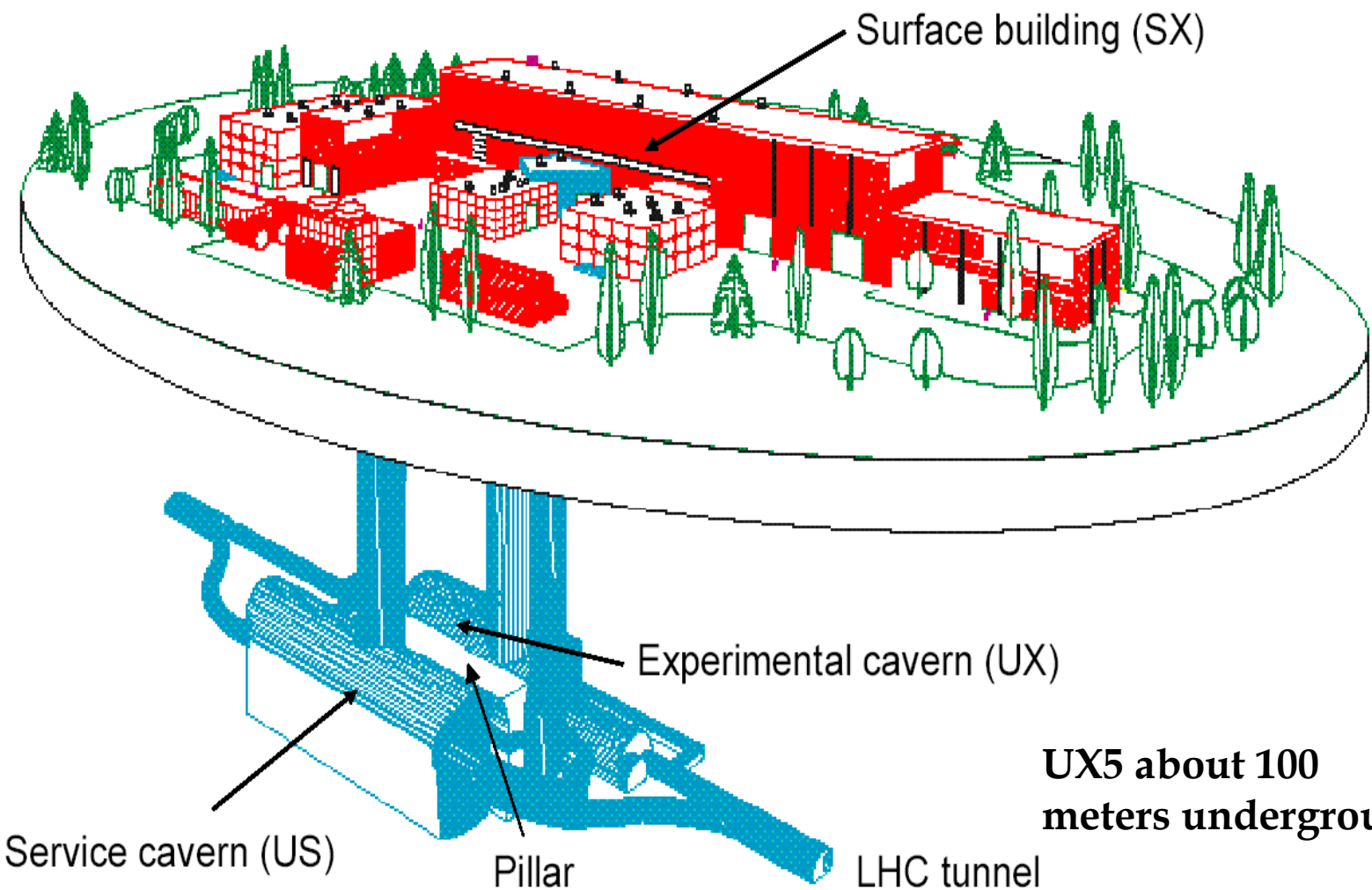
Hermetic calorimeter
- ECAL: PbWO4 crystal
- HCAL: brass+scinti.
- ($|\eta| < 3.0$)

in 4 Tesla solenoid
- (12.5m long, 6m Φ in)

Robust muon system
- DT+RPC (barrel)
- CSC+RPC (endcap)
- (in iron yoke: $|\eta| < 2.4$)

Fast cerenkov calor.
- quartz fibber
- ($3 < |\eta| < 5$)

CMS at LHC Point 5



UX5 about 100 meters underground

Surface buildings and main shaft

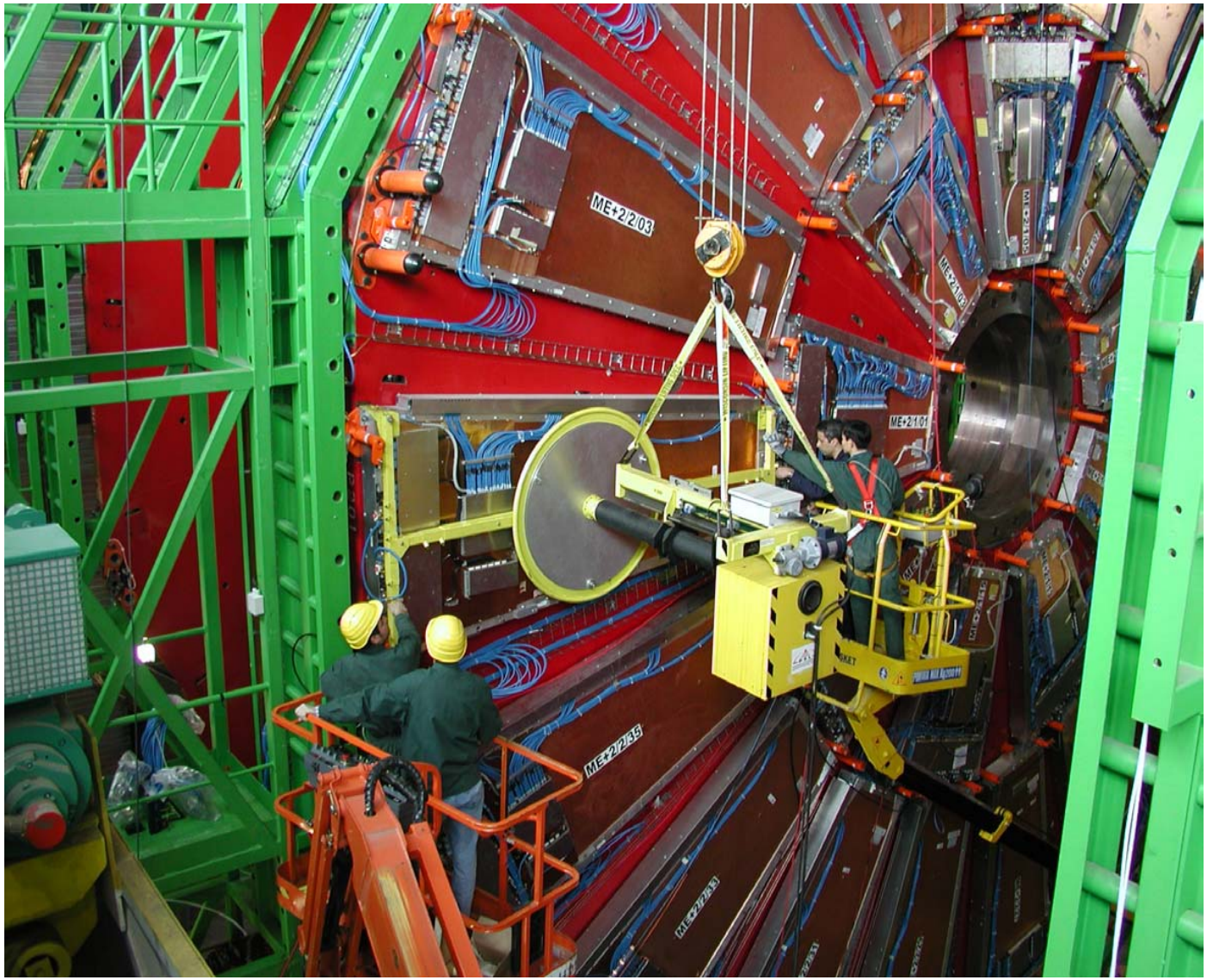
My slide from 2003

Installation of the first muon chamber

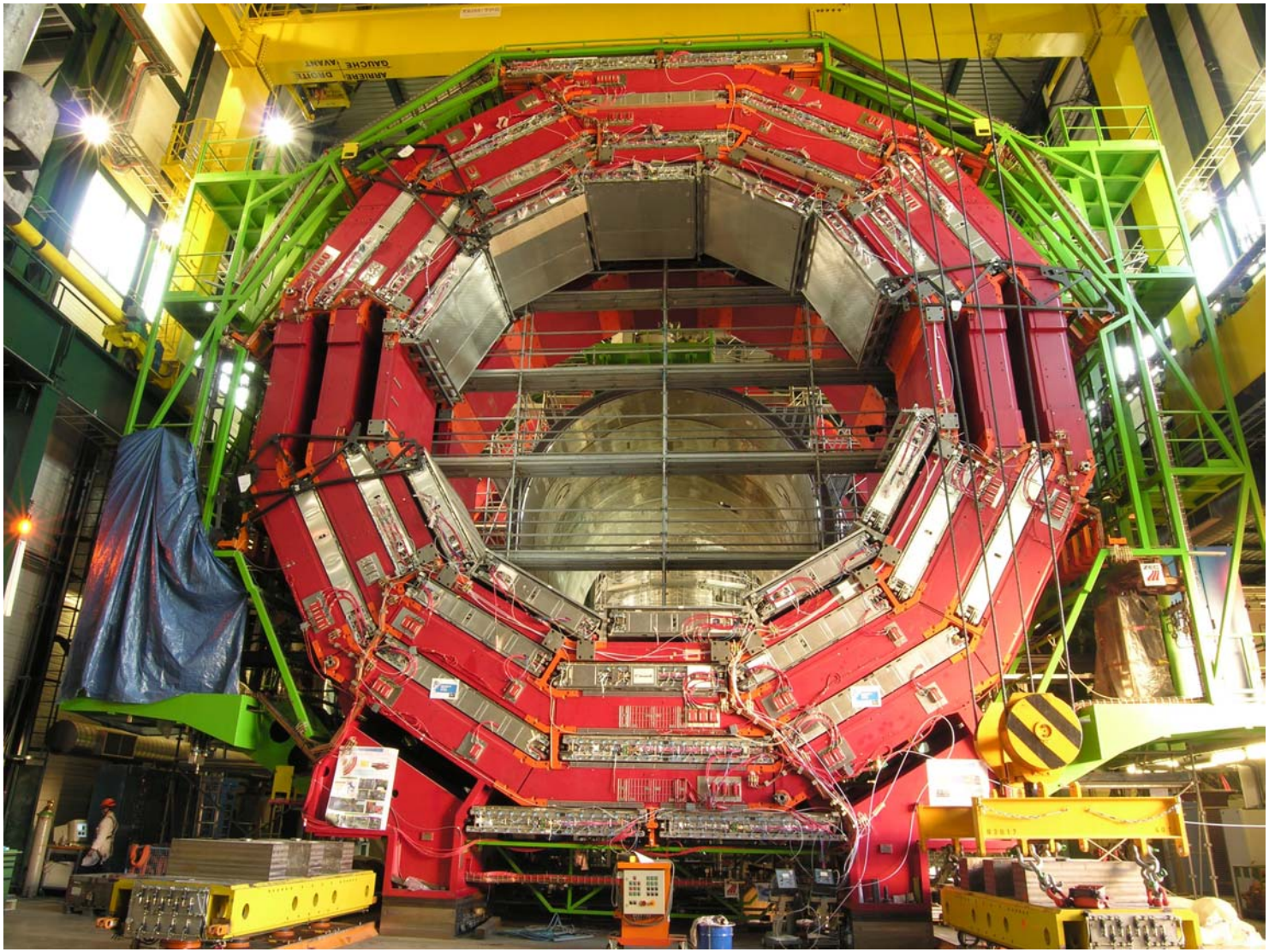
My slide from 2003

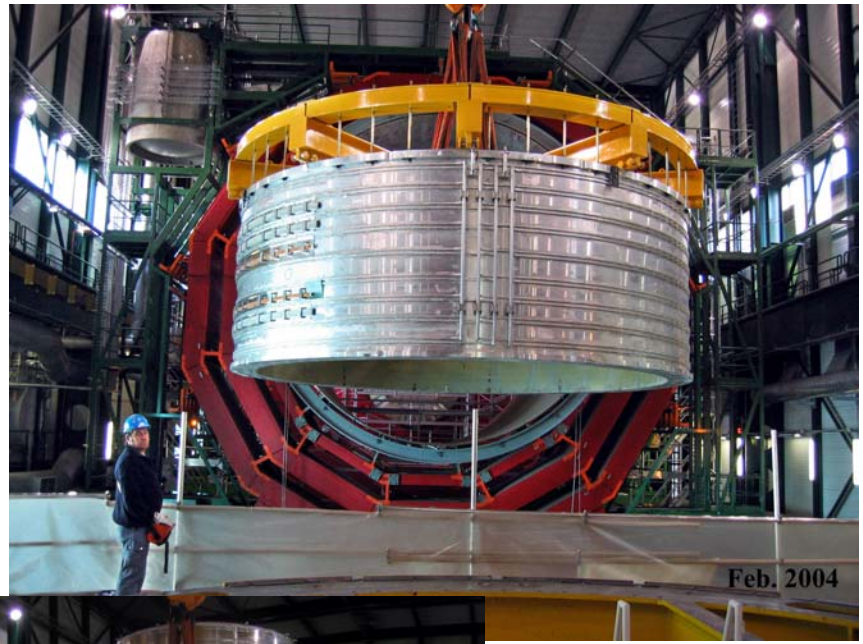
February 1st, 2005

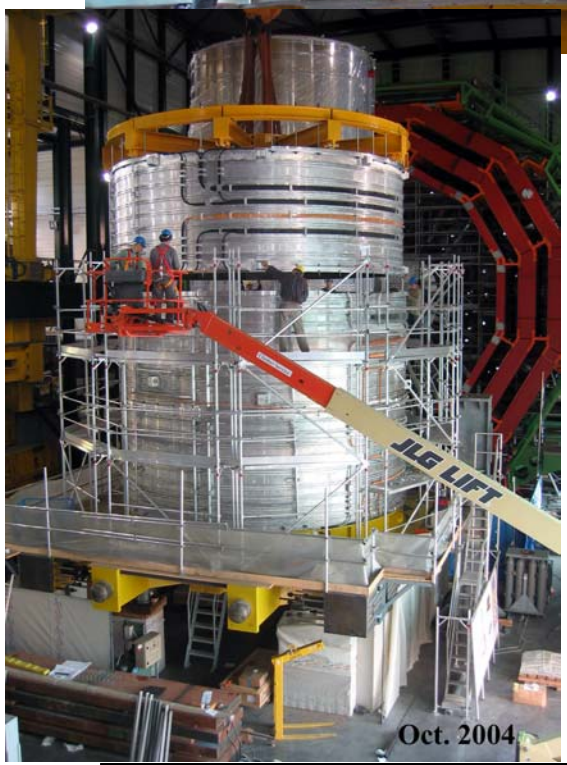
Endcap Muon: CSC Installation



Barrel Muon

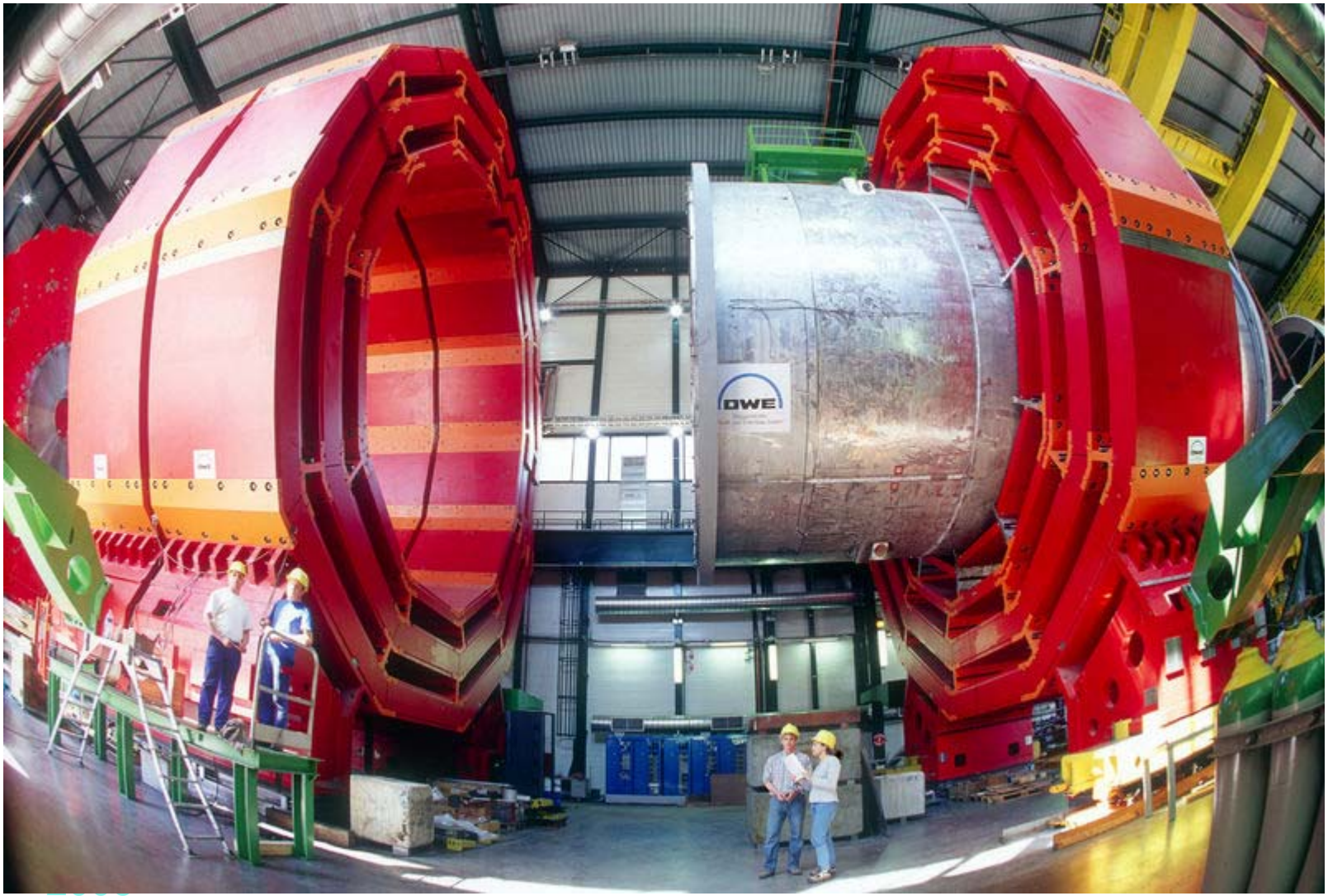






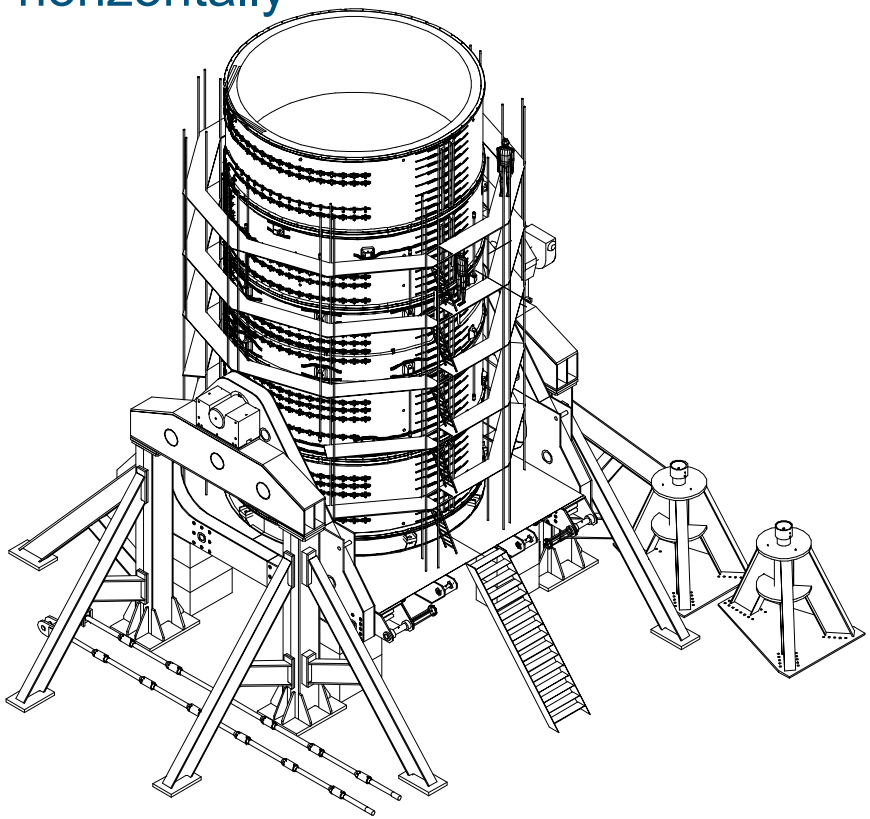
Superconducting Solenoid

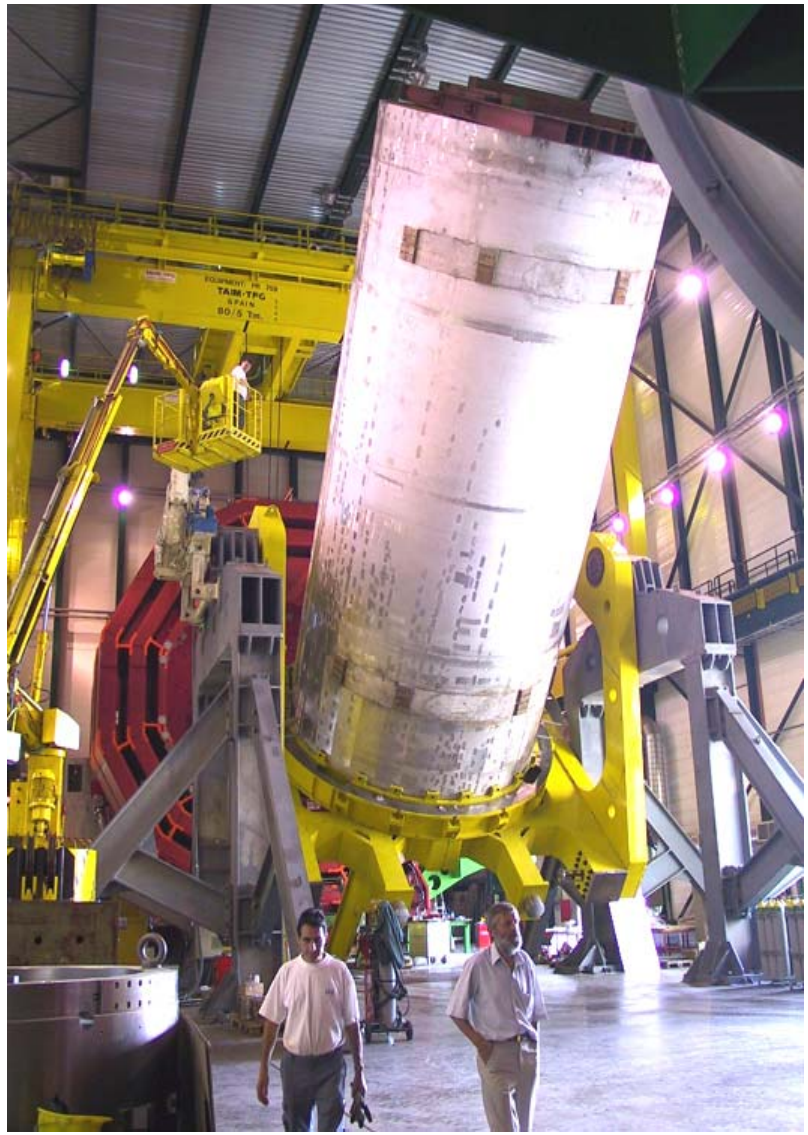
Muon Ring-0 and Vacuum Tank



Test Coil Swiveling

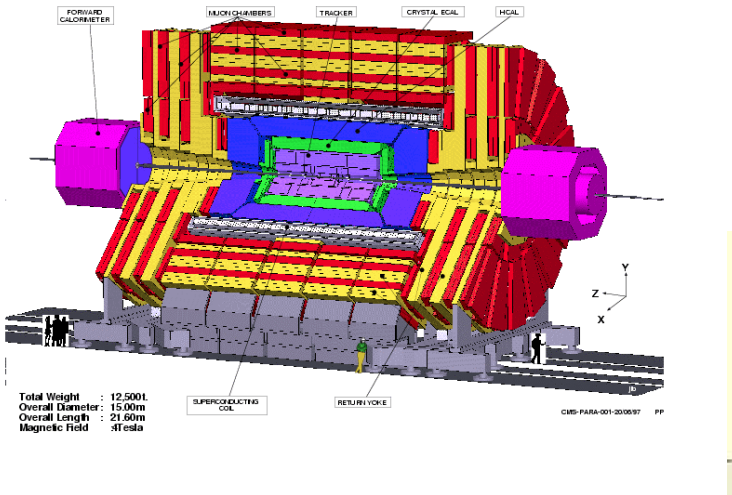
The 5 coil modules will be stacked vertically on a Swiveling Platform. The coil will then be inserted horizontally

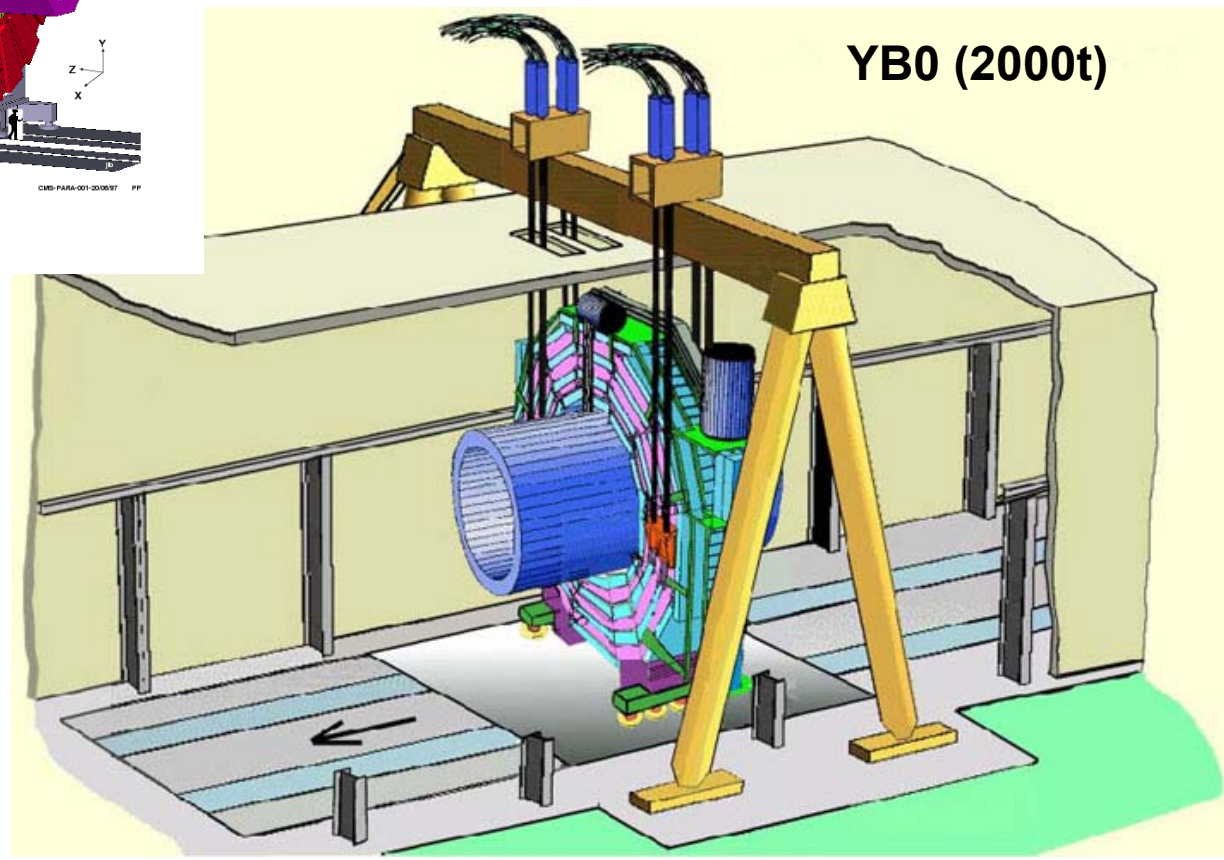




2003 Test

Heavy Lowering

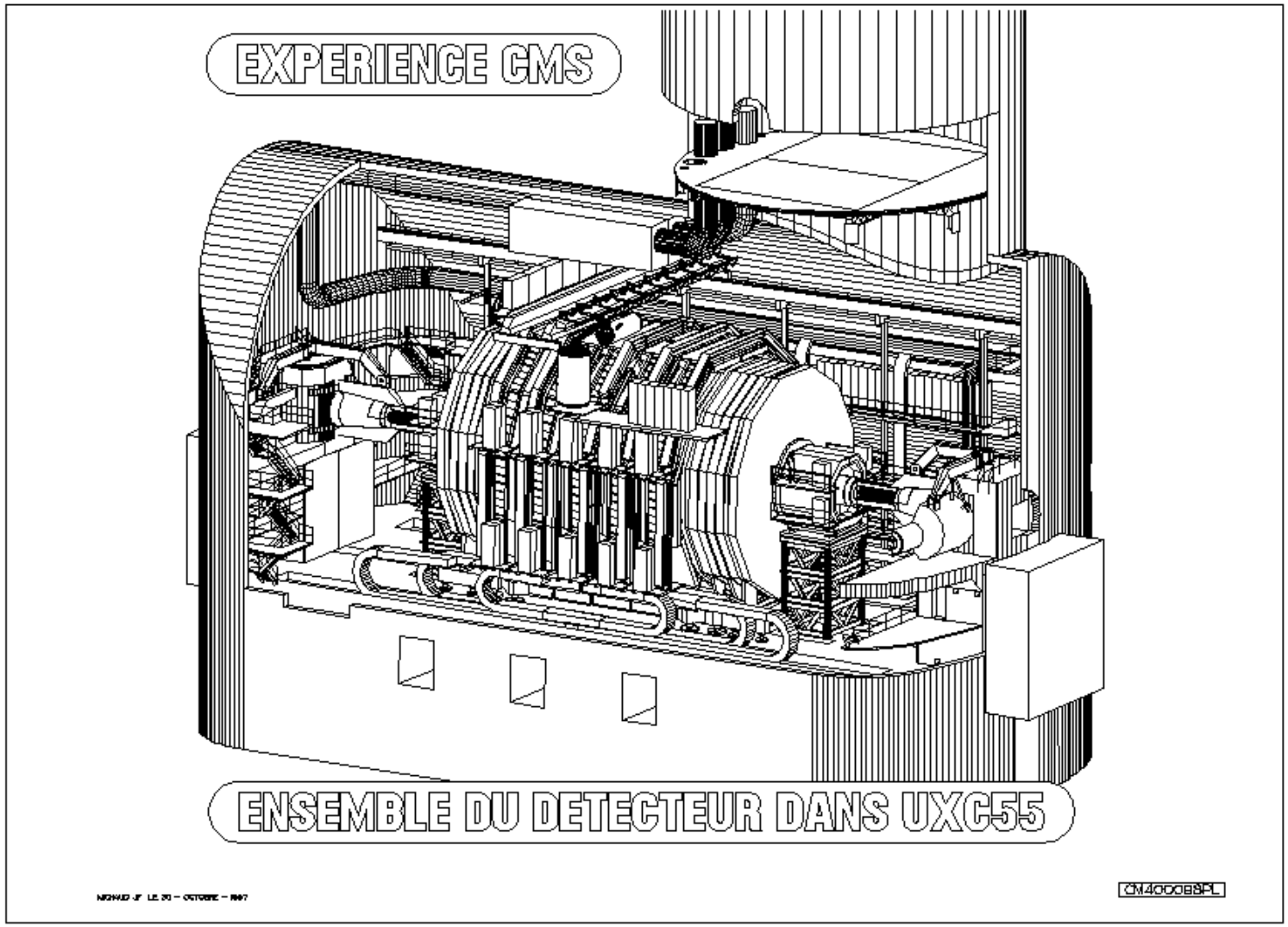




Heavy Lowering starts Feb.206.

15 Major lifts.

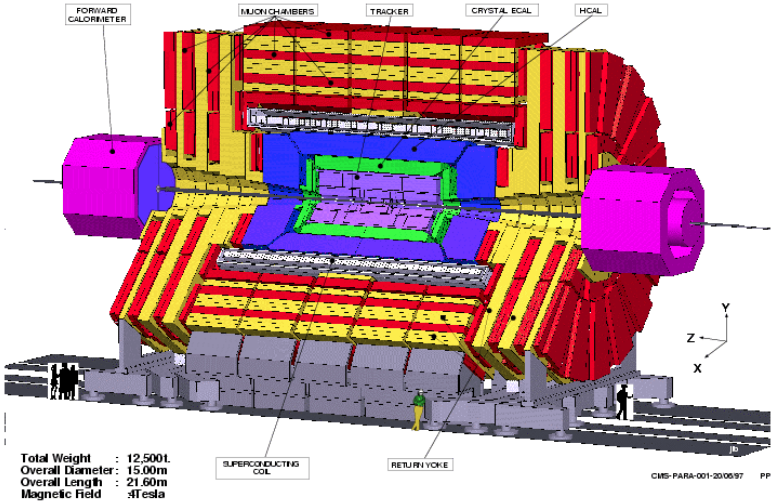
CMS Collision Hall (UX5)



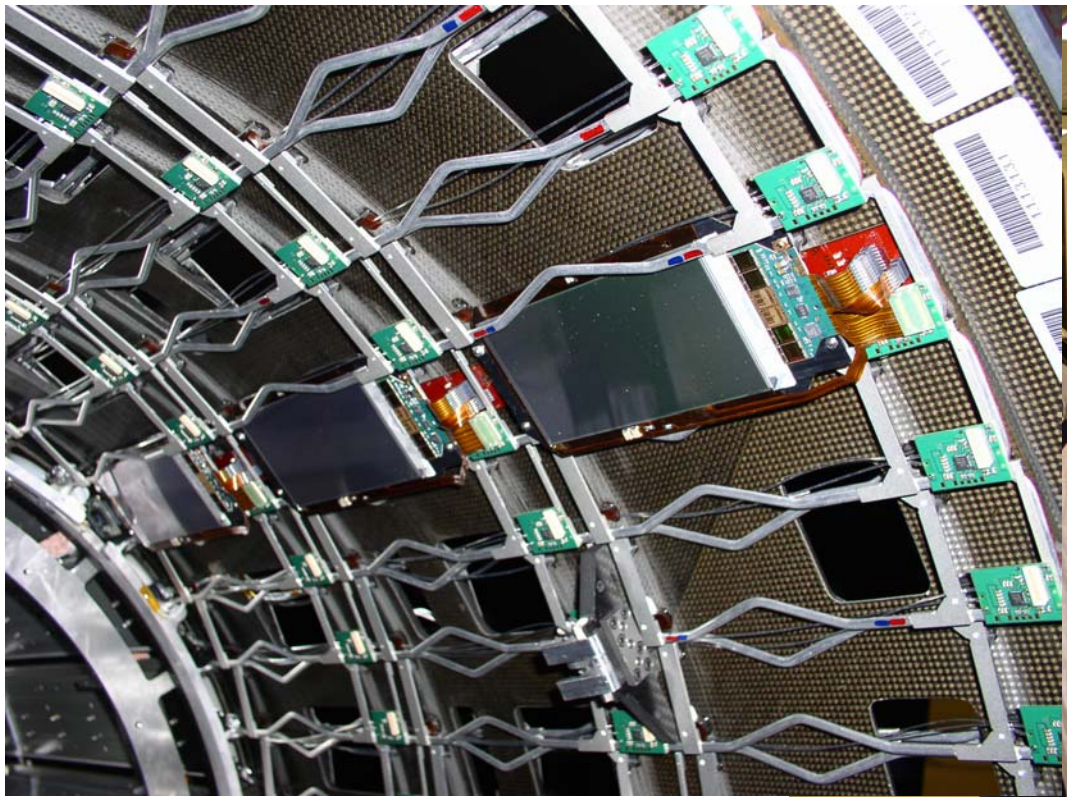
HF (quartz fiber calorimeter)

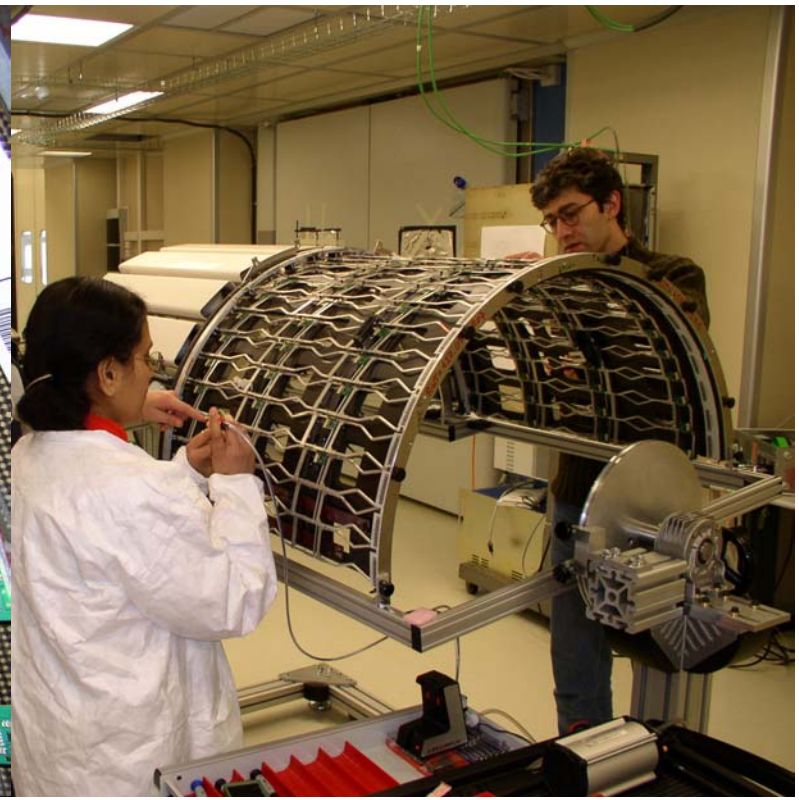
Texas Tech

HF are first items to be lowered in Jan 2006.



Silicon Tracker (Inner Barrel)





Standard Model

Force

Electromagnetic

Photon

Atoms
Light
Chemistry
Electronics

Weak

Bosons (W,Z)

Neutron decay
Beta radioactivity
Neutrino interactions
Burning of the sun

Strong

Gluons (8)

Quarks

Mesons
Baryons

Nuclei

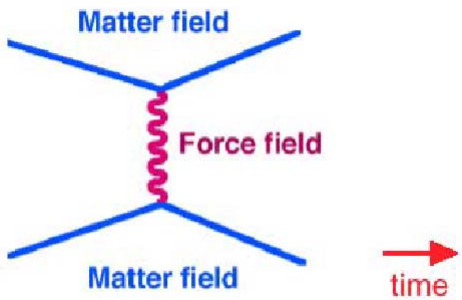
Gravitational

Graviton ?

Solar system
Galaxies
Black holes

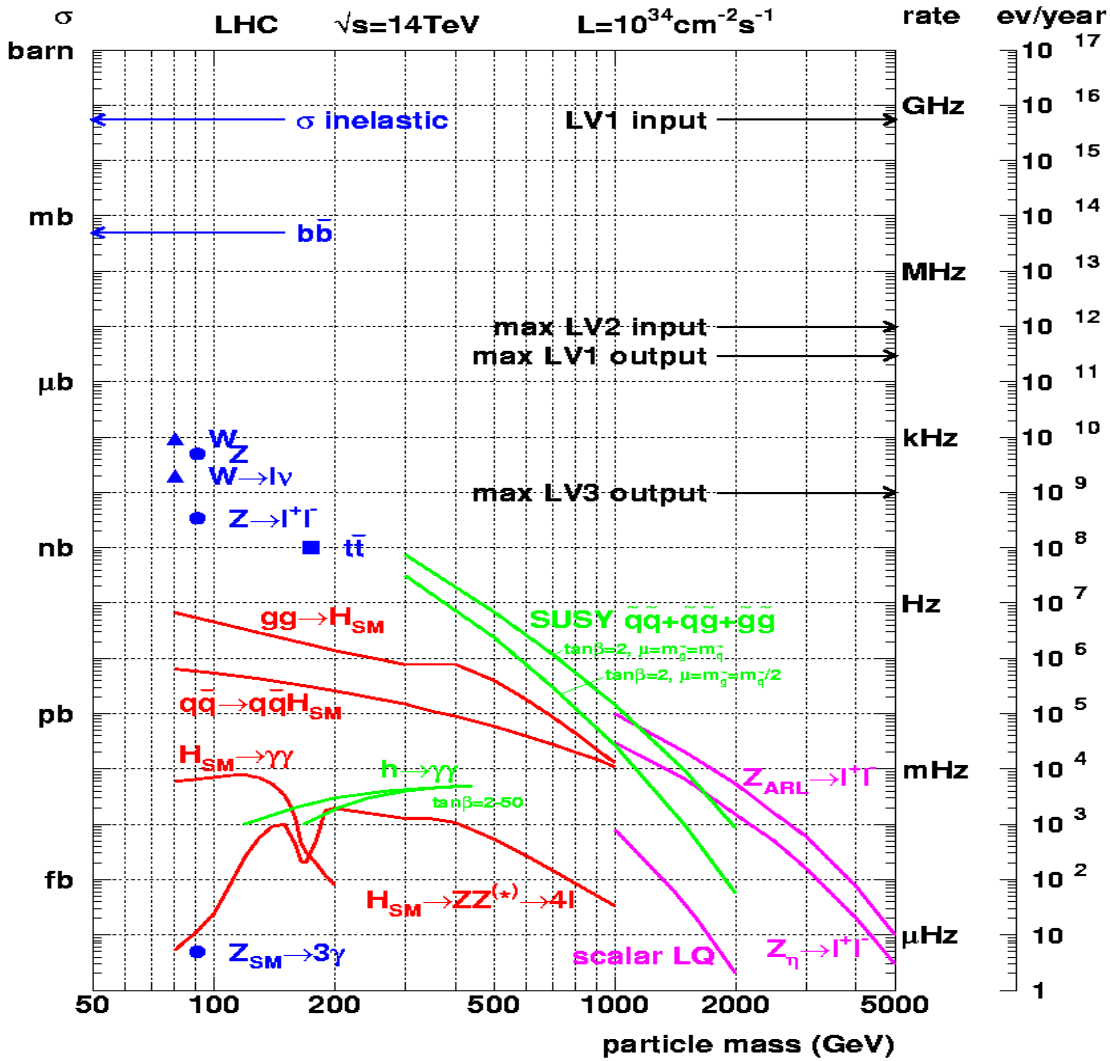
Matter

	Quarks		Leptons	
Generation 3	t Top	b Bottom	τ Tau	ν_τ Tau-neutrino
Generation 2	c Charm	s Strange	μ Muon	ν_μ Muon-neutrino
Generation 1	u Up	d Down	e Electron	ν_e Electron-neutrino



One missing piece: Higgs – Origin of electroweak symmetry breaking?

Physics at LHC



Standard Model

Higgs
- Origon of EW symmetry breaking? Mass?

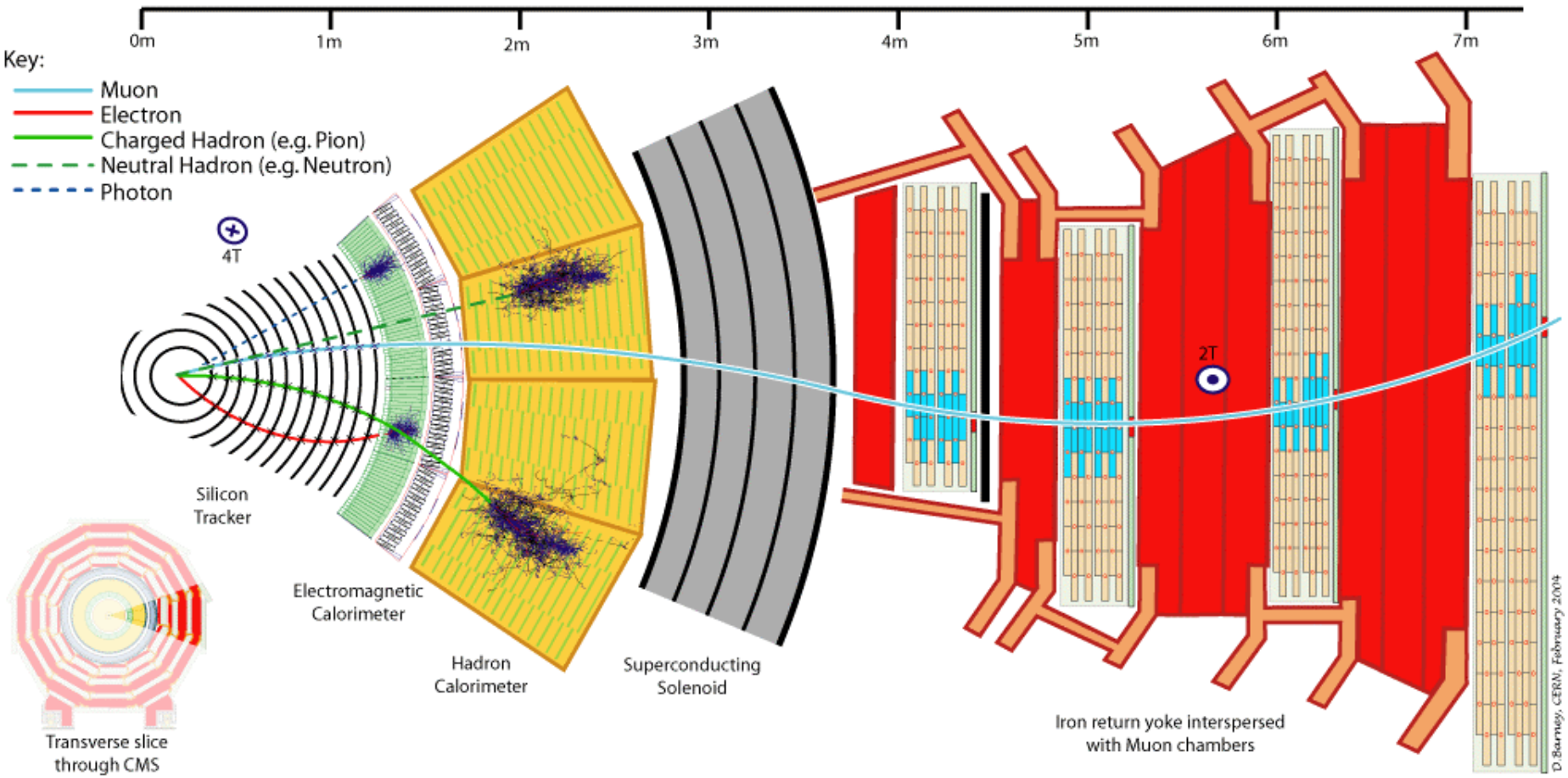
Top, b-quark, …

Beyond Standard Model

Supersymmetry (SUSY)
- Dark matter
- Unify gauge couplings

Extra dimension
Little Higgs
Technicolor
Heavy Z', W'
Black Hole
Quark substructure
:

The CMS detector

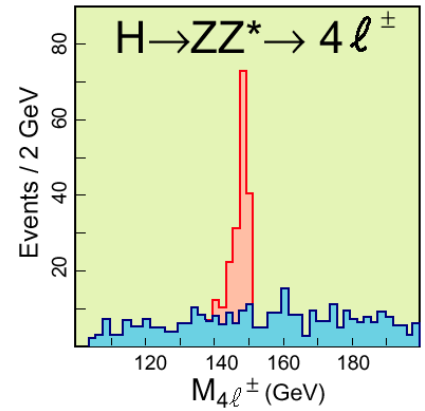


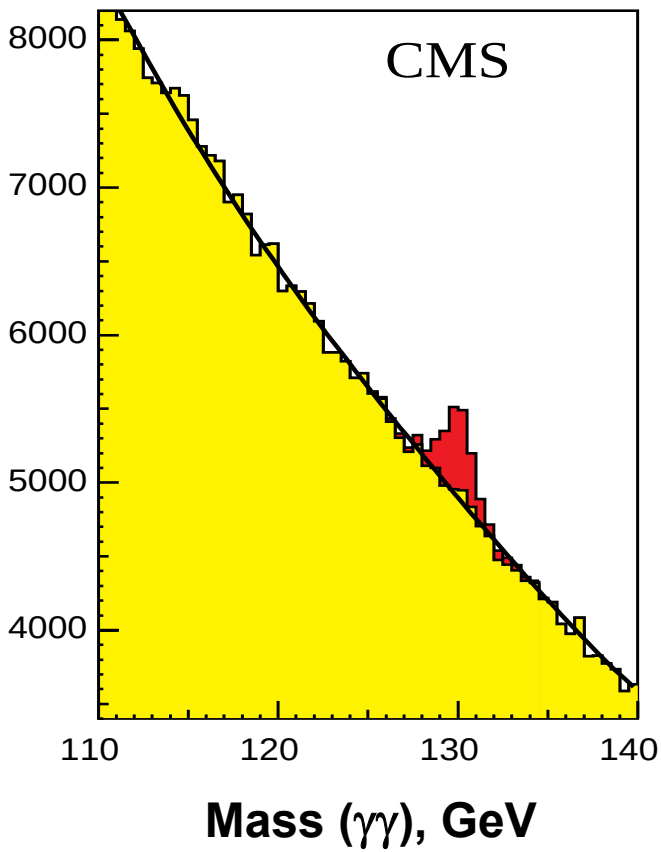
CMS_Slice.mov

Higgs

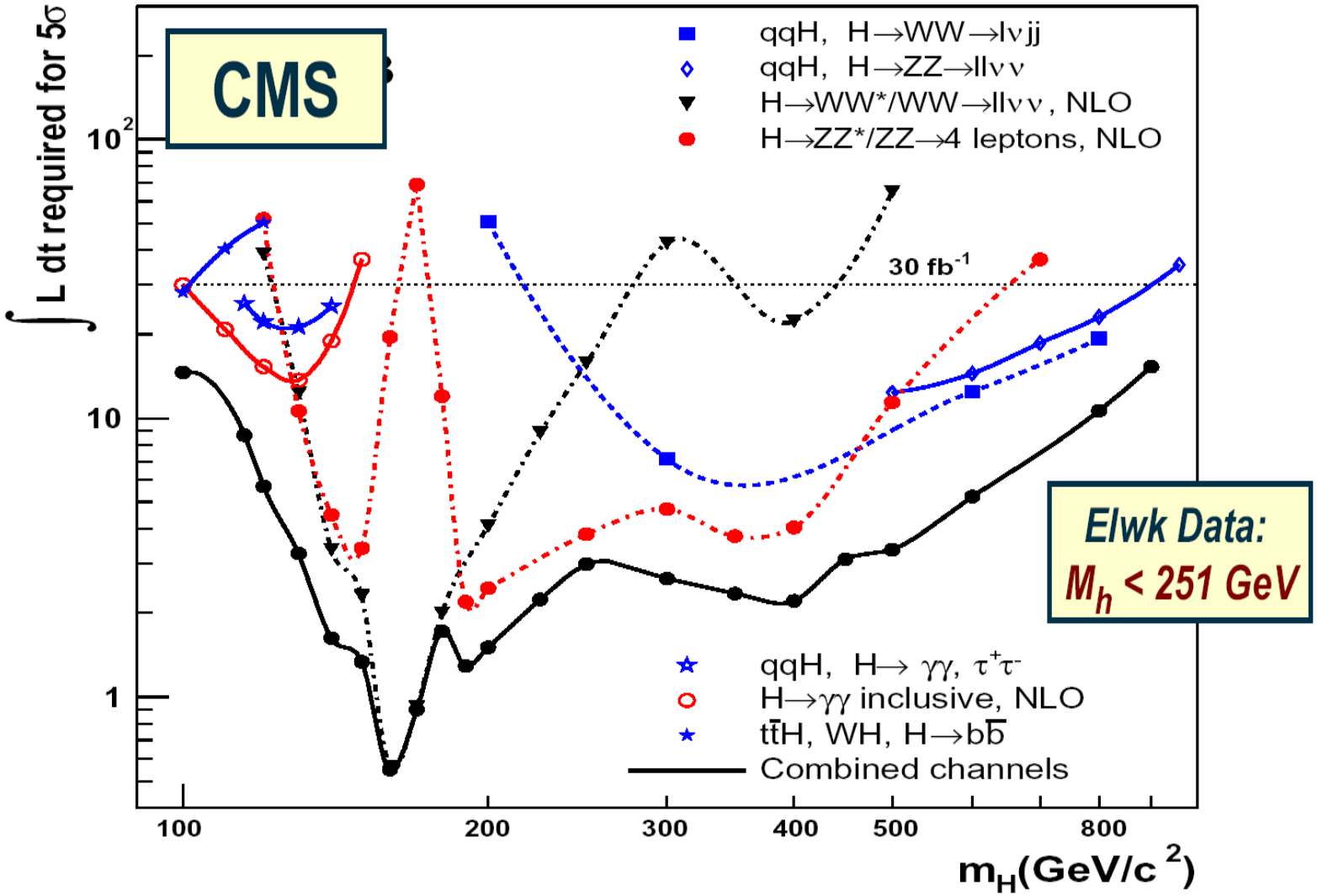
$H \rightarrow \gamma\gamma$

$H \rightarrow ZZ \rightarrow \mu\mu\mu\mu$





Higgs

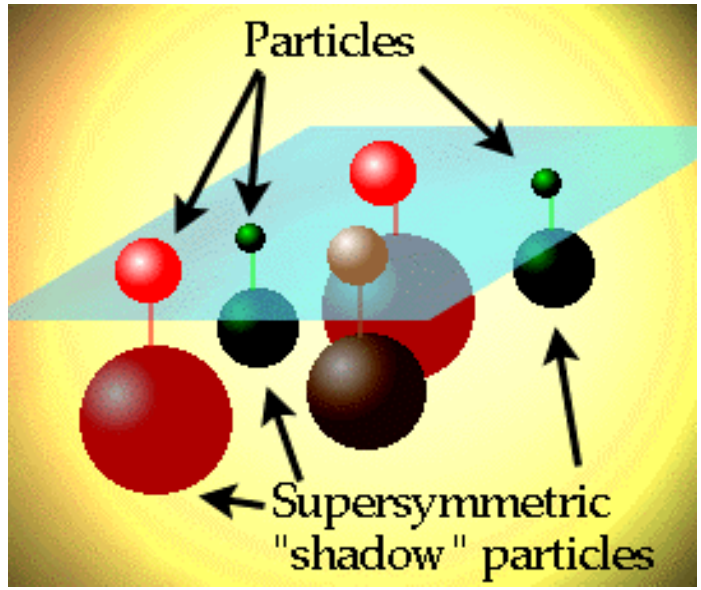

CMS
qqH, H→WW→lνjj
qqH, H→ZZ→llνν
H→WW*/WW→llνν, NLO
H→ZZ*/ZZ→4 leptons, NLO
30 fb-1
qqH, H→ γγ, τ+τ-
H→γγ inclusive, NLO
tt̄H, WH, H→bb̄
Combined channels
∫L dt required for 5σ
10^2
10
1
100
200
300
400
500
800
mH(GeV/c 2)
Elwk Data:
Mh < 251 GeV

SUSY

An extension of Standard Model:

- Symmetry between bosons and fermions
- Many new super particles.

Heavy s-particles →

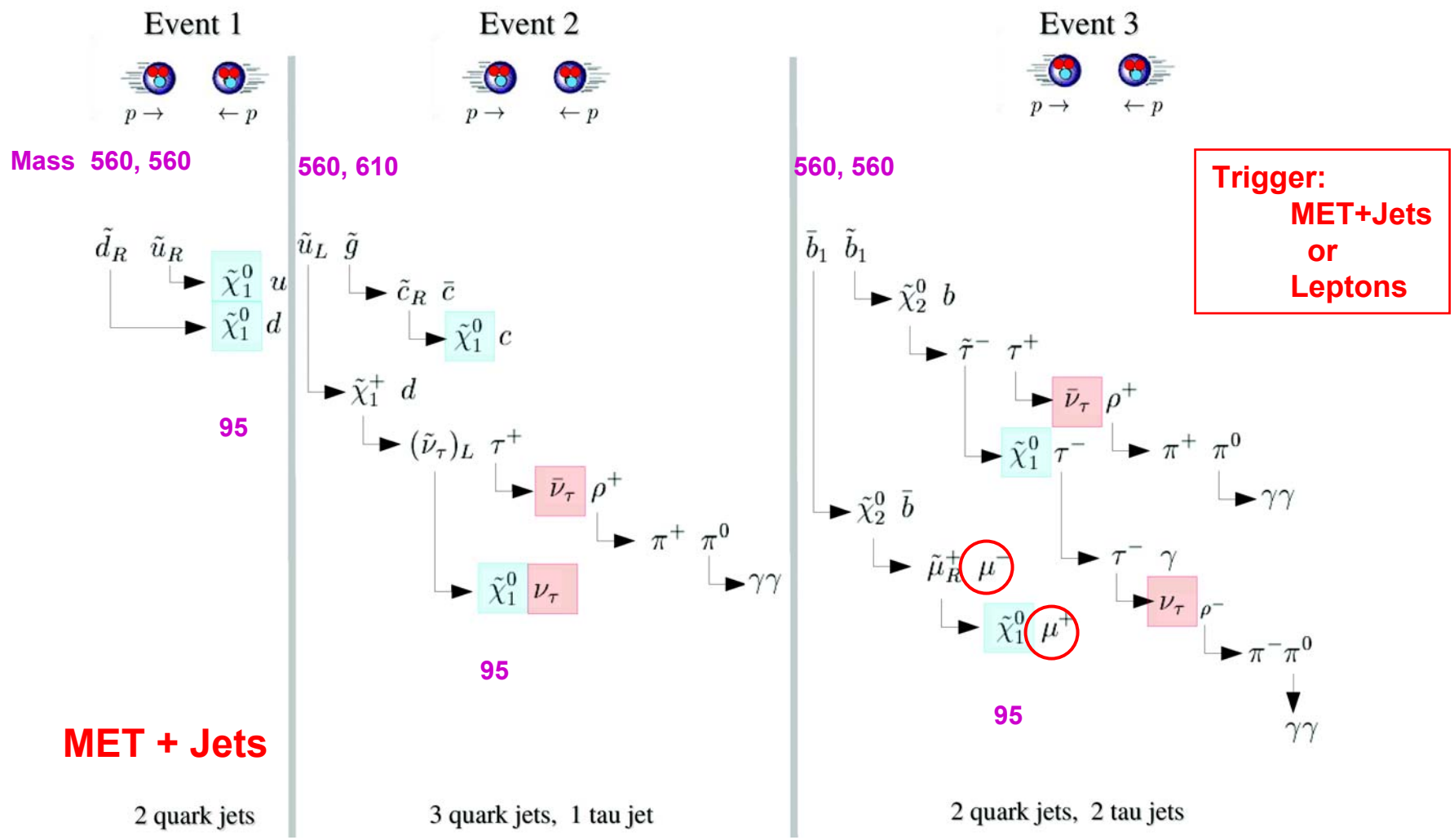


Solve several problem at once:

- Dark matter candidate (=lightest supersymmetric particle)
- Opening towards a theory of gravity
- Unification of gauge couplings
- Hierarchy problem
- Allows to explain why the Higgs mechanism works.

SUSY Particle Decay Chain (example)

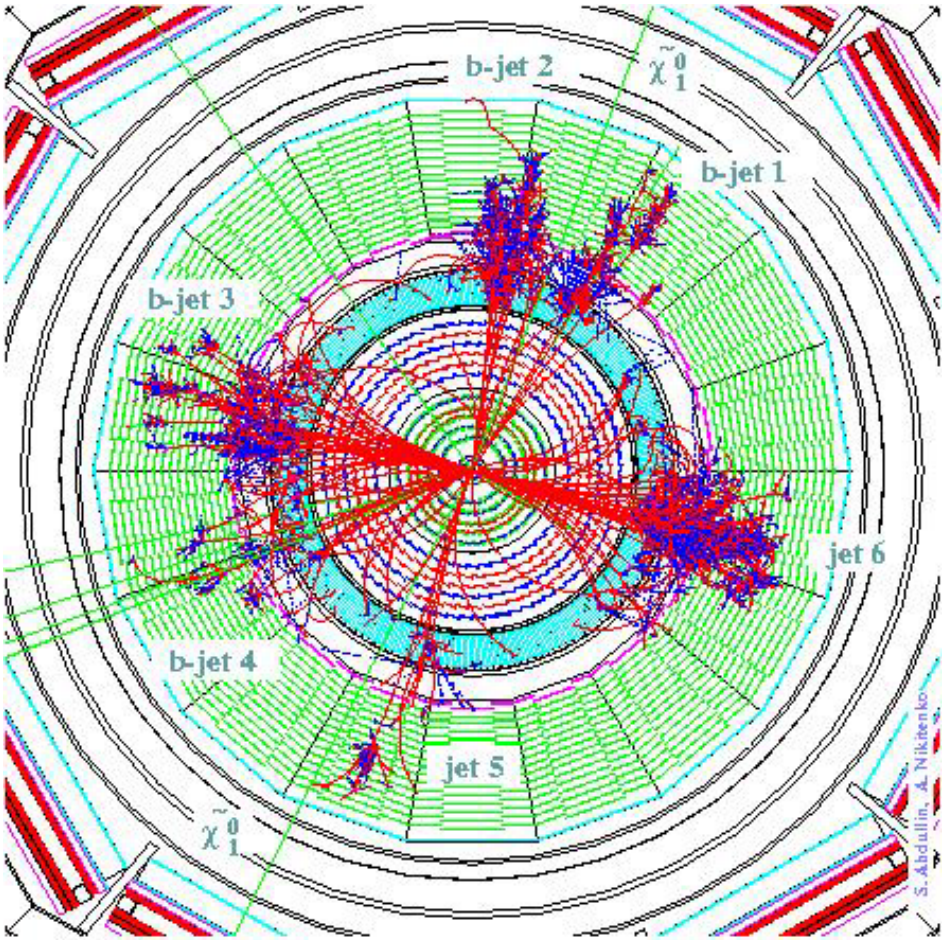
M_0 = 60, $M_{1/2}$=250, A_0=0, μ=+, tanβ=10



MET = Missing Transverse Energy

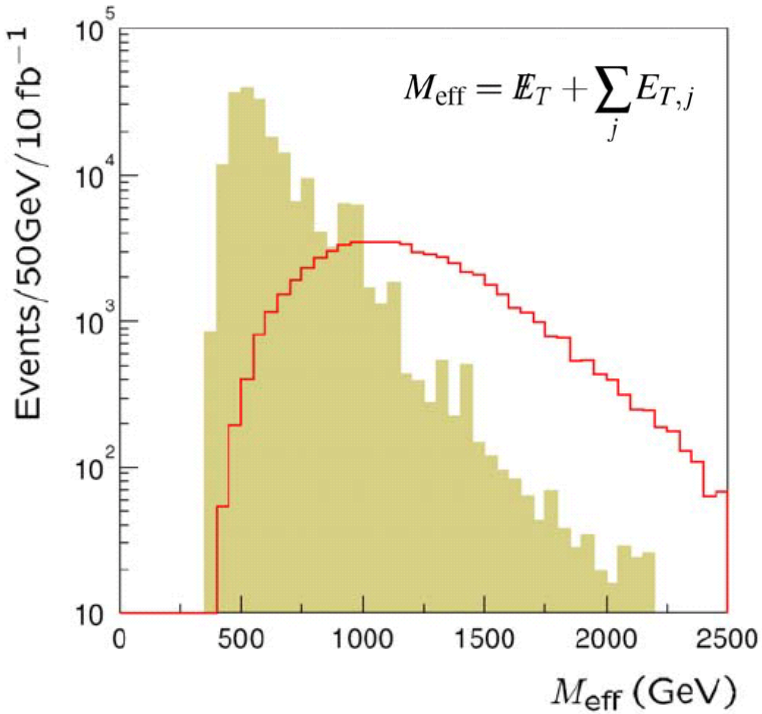
SUSY Signature

Complex Topology



First Event Selection

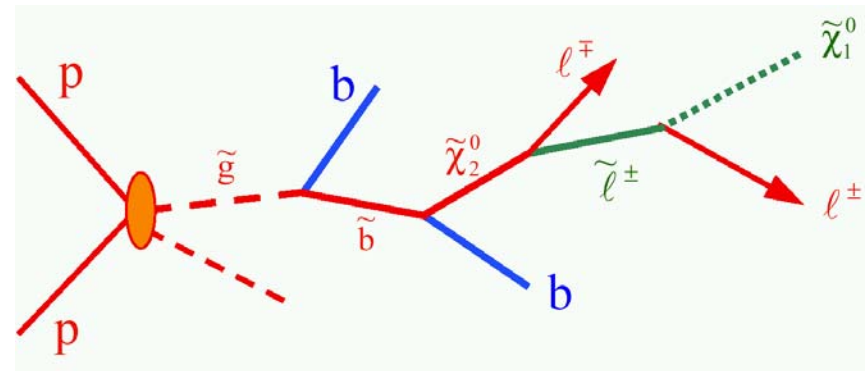
Multiple jets
+
MET (Missing)



Reconstruction of Mass

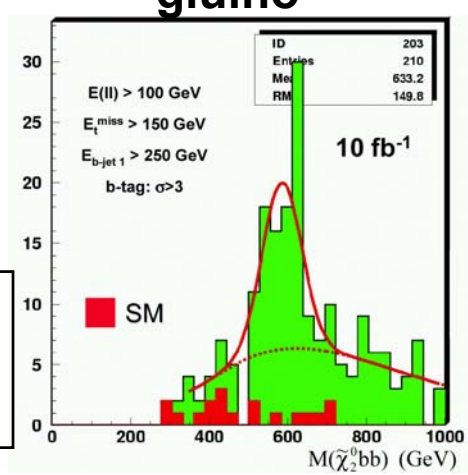
g	595.1	t_L	392.9
b_L	496.0	t_R	575.9
b_R	524.0	χ_4^0	361.1
q_L	540	χ_3^0	339.9
q_R	520	χ_2^0	174.4
l_L	196.5	$\chi_2^\pm$	361.6
l_R	136.2	$\chi_1^\pm$	173.8
		χ_1^0 = LSP	95.6

$$\sigma_{SUSY}^{TOT} = 57.77 \text{ pb}$$



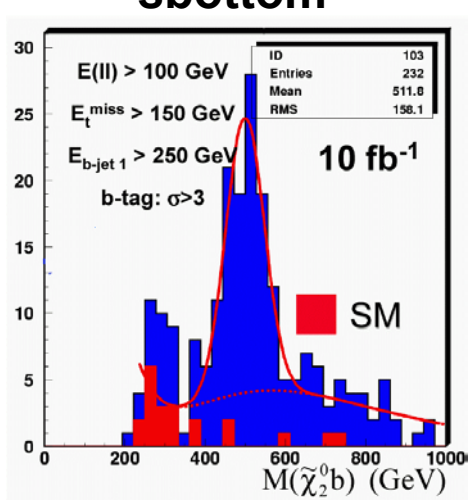
$$M_{\ell^+\ell^-}^{max} = \frac{\sqrt{(M_{\tilde{\chi}_2^0}^2 - M_{\tilde{\ell}}^2)(M_{\tilde{\ell}}^2 - M_{\tilde{\chi}_1^0}^2)}}{M_{\tilde{\ell}}}$$

gluino

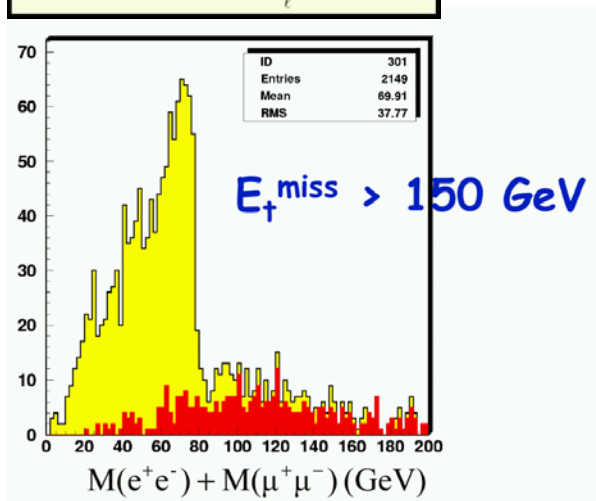


585.1 +/- 11.1 GeV

sbottom



499.4 +/- 6.6 GeV

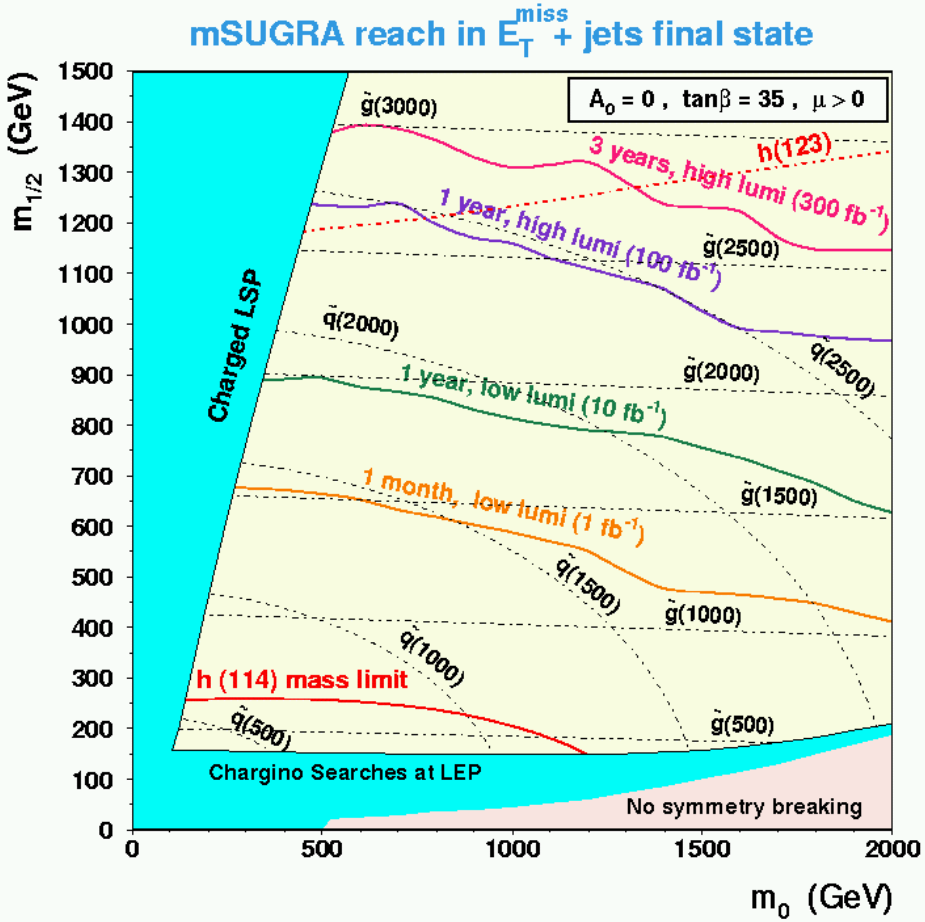


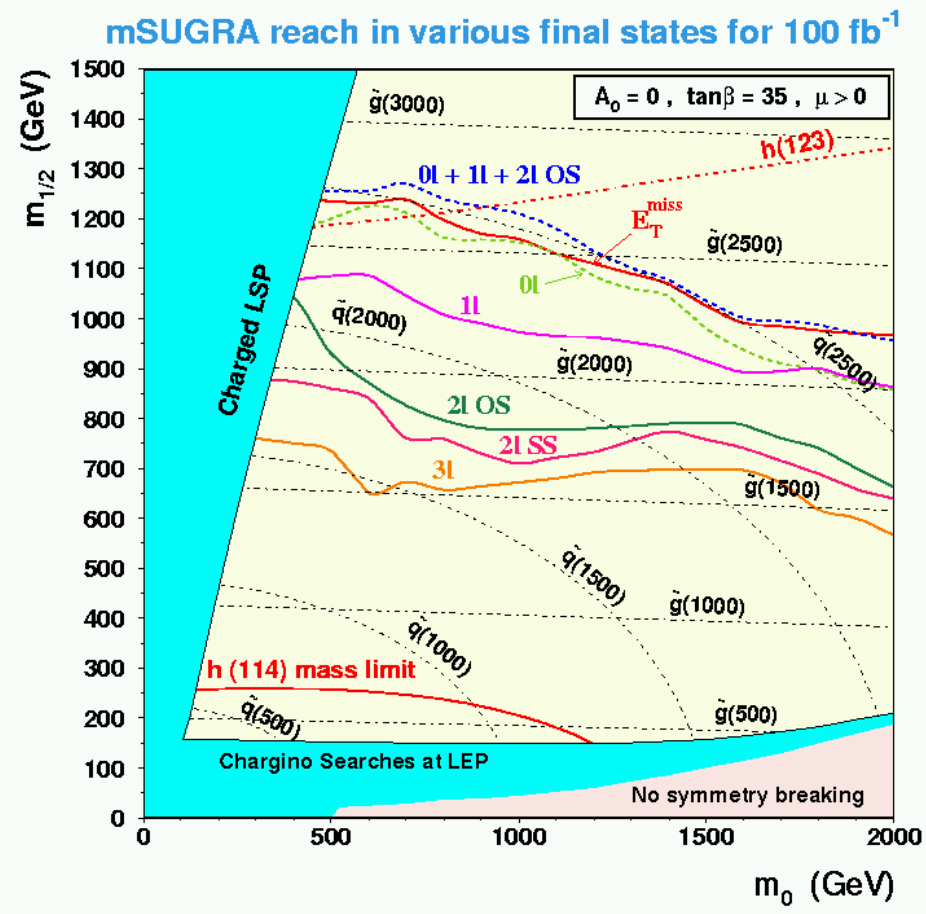
edge 78.9 +/- 2.1 GeV (generated 78.2)

repeat for non-b squarks.

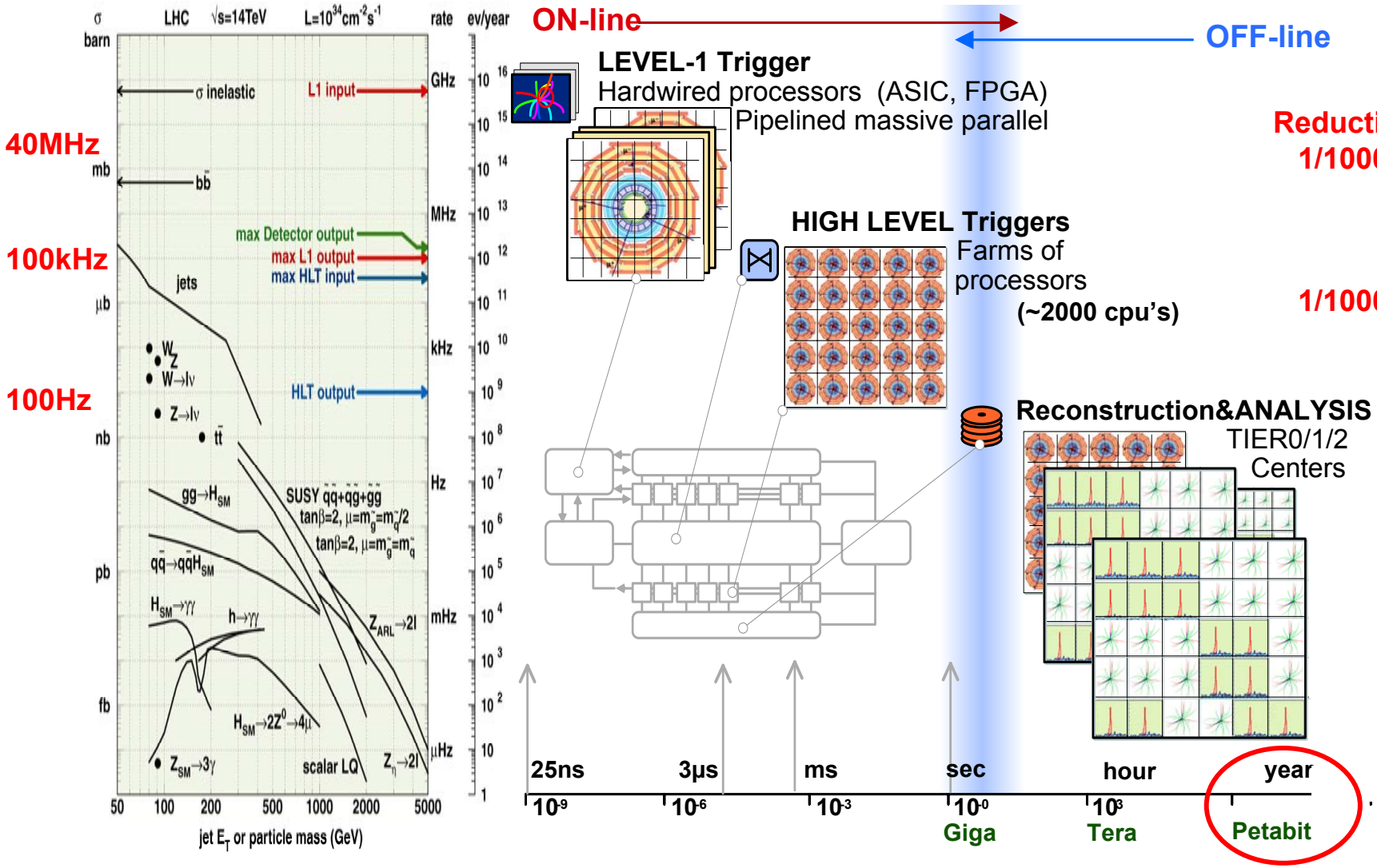
SUSY Search

TeV scale SUSY should be found at LHC. rather quickly… If it exists.





Physics Selection



Offline Computing

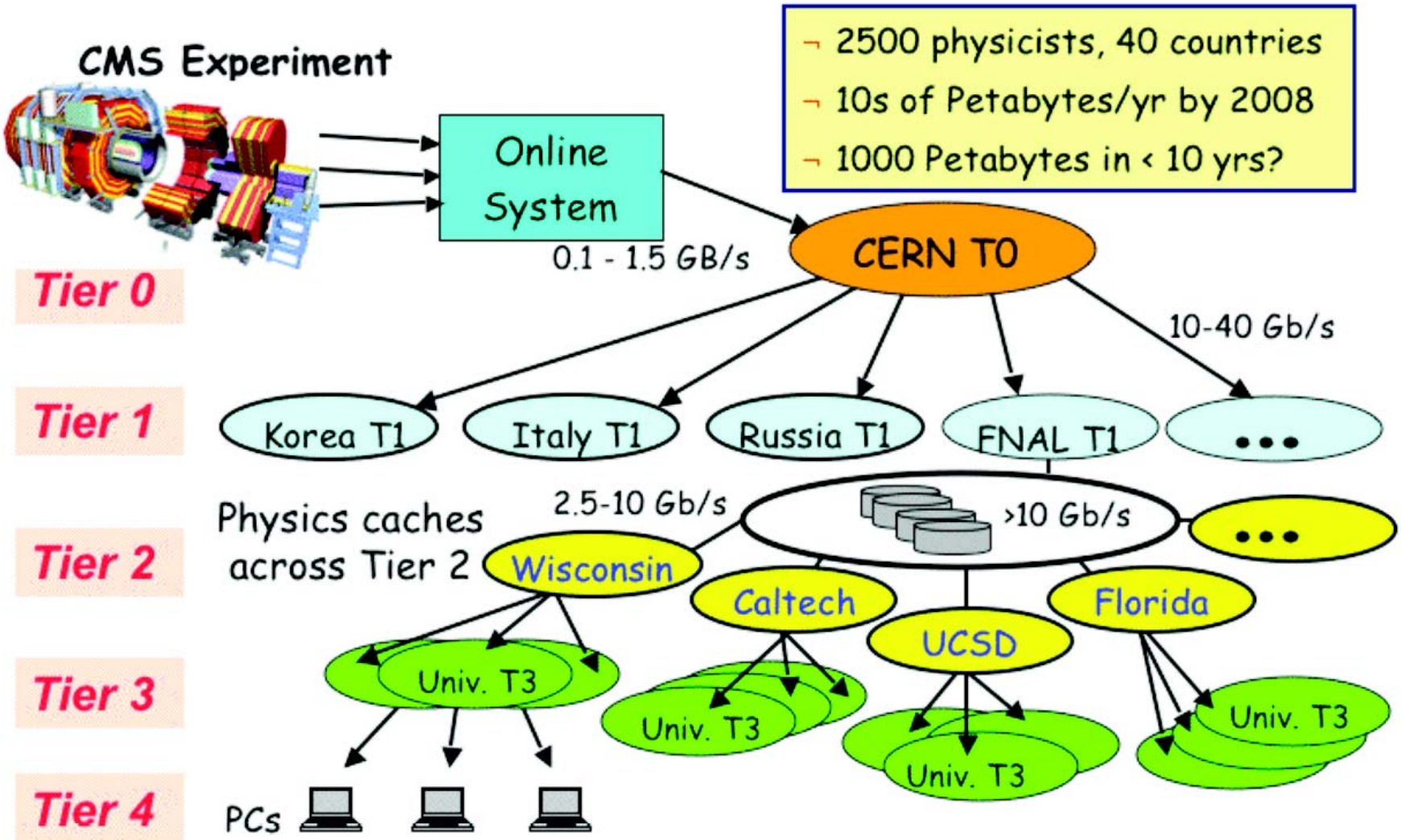
Data Processing

- **Full Detector Simulation**
  - Input – none. Output - O(100M) events per year.
  - CPU intensive (a few minutes per event) & 1PB data.
    - **Use many CPU's O(10K) in distributed world.**
    - **Transfer events to CMS data centers to store.**
- **Event Reconstruction**
  - Input & Output - O(1G) events/year
  - CPU intensive & O(1PB) of input/output data.
    - **Process events at CMS Tier-0 center and store. – a few K CPU's.**
    - **Reprocess a few times at Tier-1 centers.**

Data Access

- **Physics Analysis – find signal events from huge back ground.**
  - Interactive or batch analysis by ~2000 Physicists at T2, T3, …
  - O(100M) events or less per analysis. Repeated access.
  - I/O intensive
    - **Reduce event size and condense signal data sample – traditional way.**
    - **Grid**

CMS Data Grid



Projects and Funding in the U.S.

- Funding for Physics related Grid Projects, Middleware, LHC Computing, Networking comes from two different funding agencies in the U.S. – DOE and NSF
- Both agencies have provided significant funding to lab's, universities, experiment-led projects and to innovative groups of researchers. They have funded
  - GriPhyN (NSF) – an early Grid project - Physics + CS
  - PPDG (DOE) – also an early Grid project – Physics + CS
  - iVDGL (NSF) – to deploy a "lab" for Grid work – spawned VDT, iGOC
  - Condor and Globus and other middleware
  - SRM
  - Ultralight (NSF) and UltraScienceNet(DOE) and LamdaStation(DOE)
  - U.S. CMS and U.S. ATLAS Software and Computing efforts, which also supply funding to use/enhance/work together on some of the above
  - Experiments such as CDF, D0, Babar, SDSS, Ligo, STAR
  - TeraGrid(NSF)
  - LHCnet transatlantic link (DOE)
  - DOE Science Grid – Certificate authority
  - And much much more.....

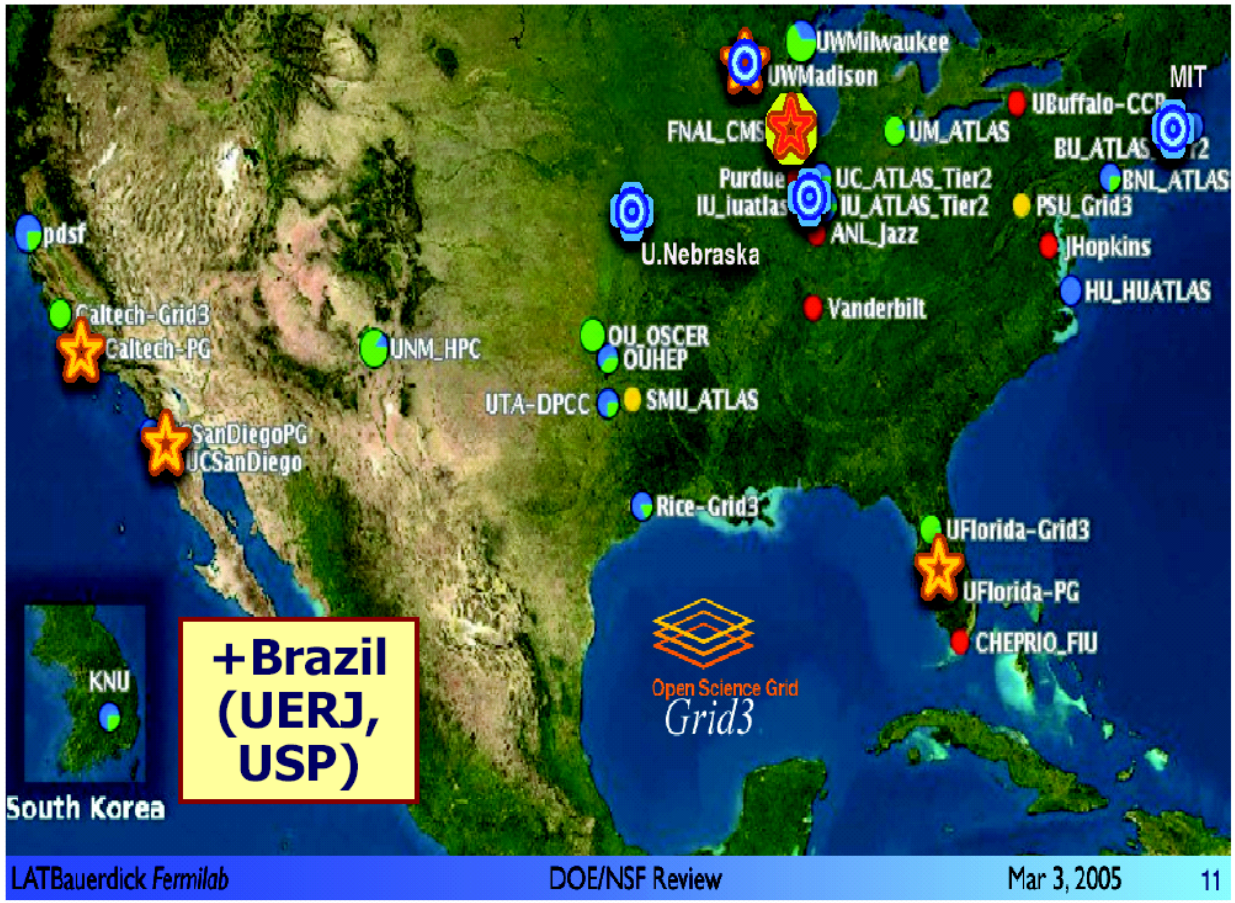
Next – move from R&D to integrated production system.

March 16, 2005 GDB - Open Science Grid in the US

Vicky White
Fermilab

GIRD3 and US-CMS Tier 1/2

Coordinated effort toward production system – started at SC2003



+Brazil (UERJ, USP)

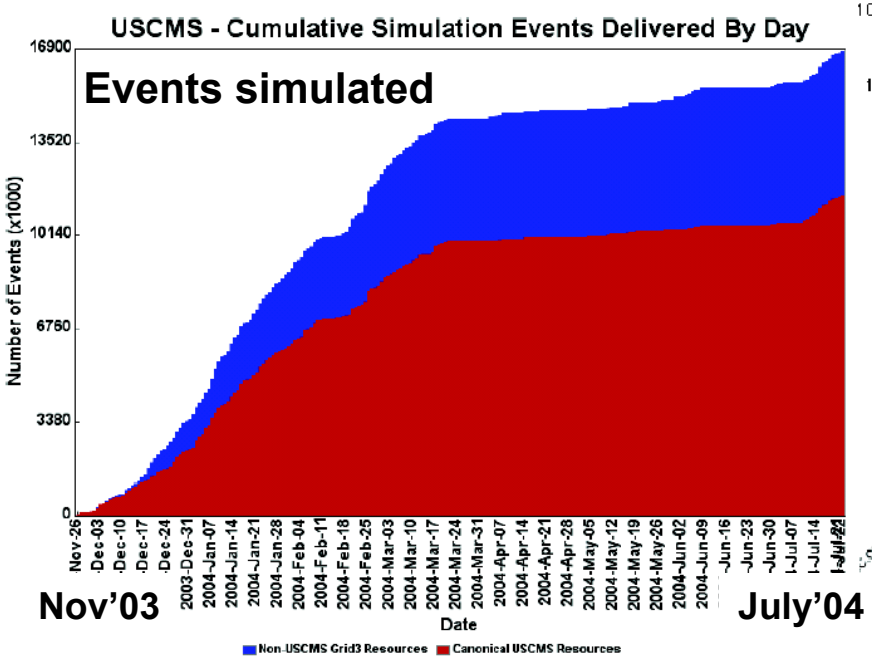
Grid3: A Nat'l Grid Infrastructure

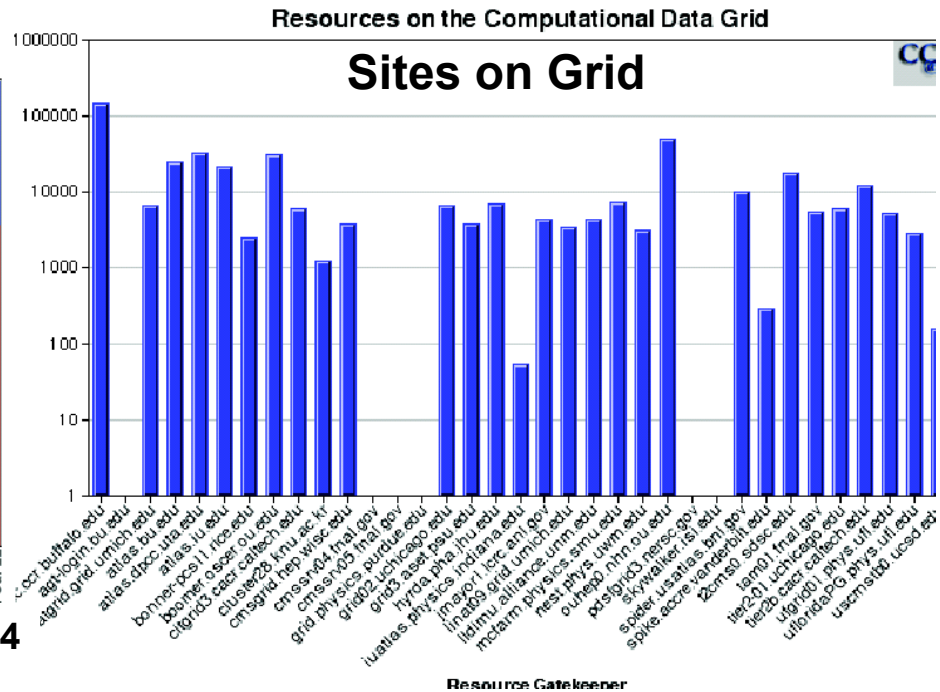
- 35 sites, 3500 CPUs: Univ. + 4 Nat'l labs
- Part of LHC Grid
- Running since October 2003
- HEP, LIGO, SDSS, Biology, CS

Transition to Open Science Grid: http://www.openscience.org

CMS Data Challenge '04 using Grid3
(50M events GEANT4 simulation worldwide)

- During 2004 33% of US CMS simulation production was done through opportunistic use of non-CMS resources.
- The jobs ran on 17 sites with an overall efficiency of ~65%. **~1200 CPU's**
- The failures were mainly due to hardware problems and filled disk partitions. The overall infrastructure was stable.
- The support load was <1 FTE to run production.





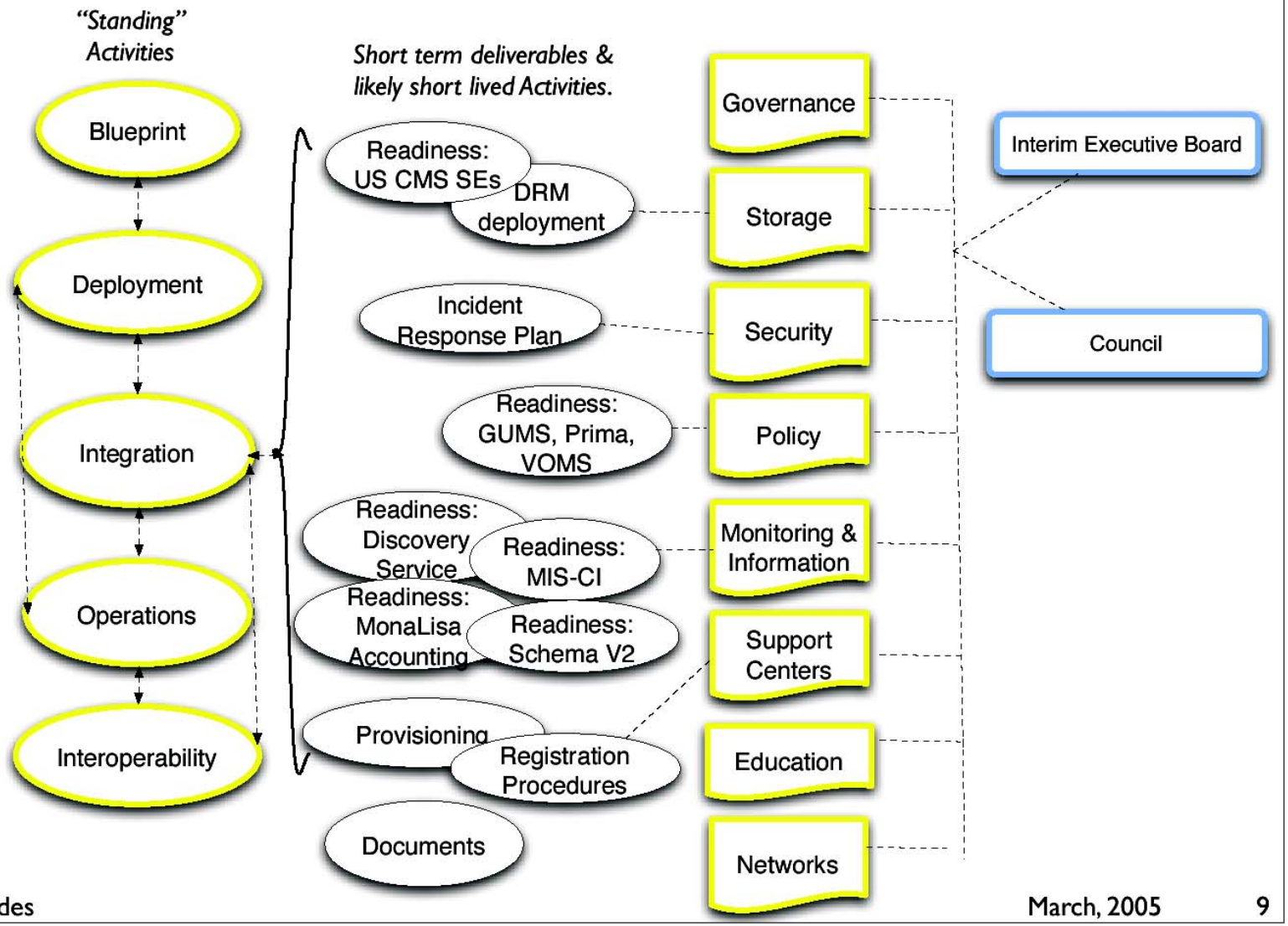
Open Science Grid Consortium

- The Open Science Grid (OSG) Consortium was formed in 2004 by teams from U.S. universities and national laboratories in order to build and support a production quality peta-scale Grid infrastructure for large scale science. The Open Science Grid will ensure that the U.S. plays a leading role in defining and operating the global grid infrastructure needed for large-scale collaborative and international scientific research
- The Open Science Grid Consortium is an organization, not a managed project. It sponsors collaborative activities and technical groups to achieve the shared goals of the members
  - You will hear much more about how this works in later talks
- http://www.opensciencegrid.org

Vicky White
Fermilab

Open Science Grid

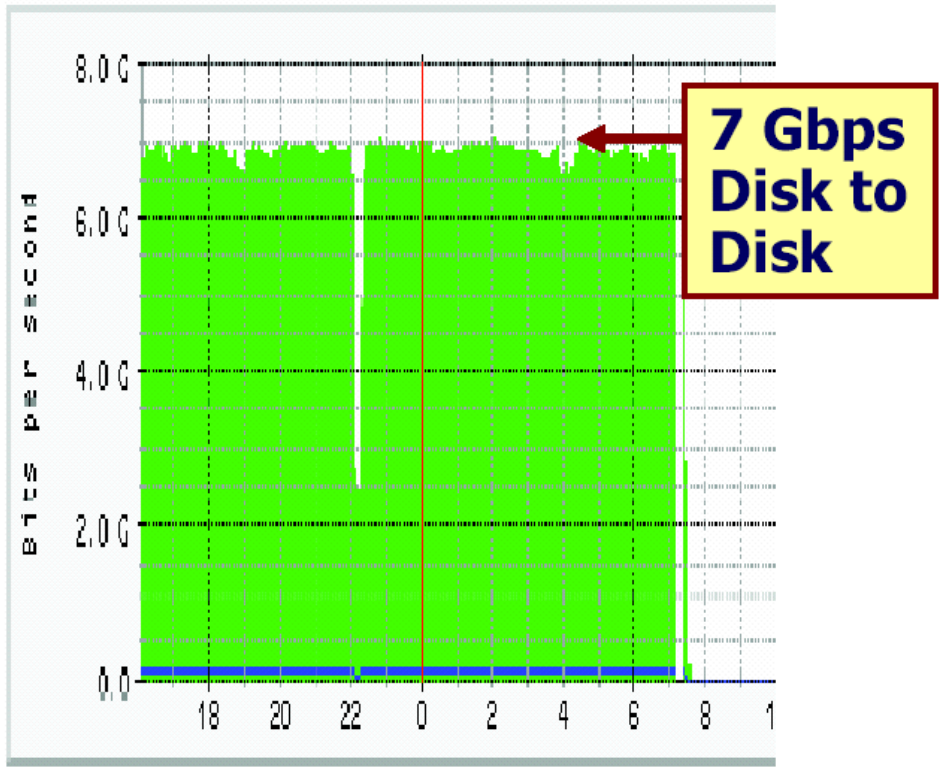
Activities



Ruth Pordes March, 2005 9

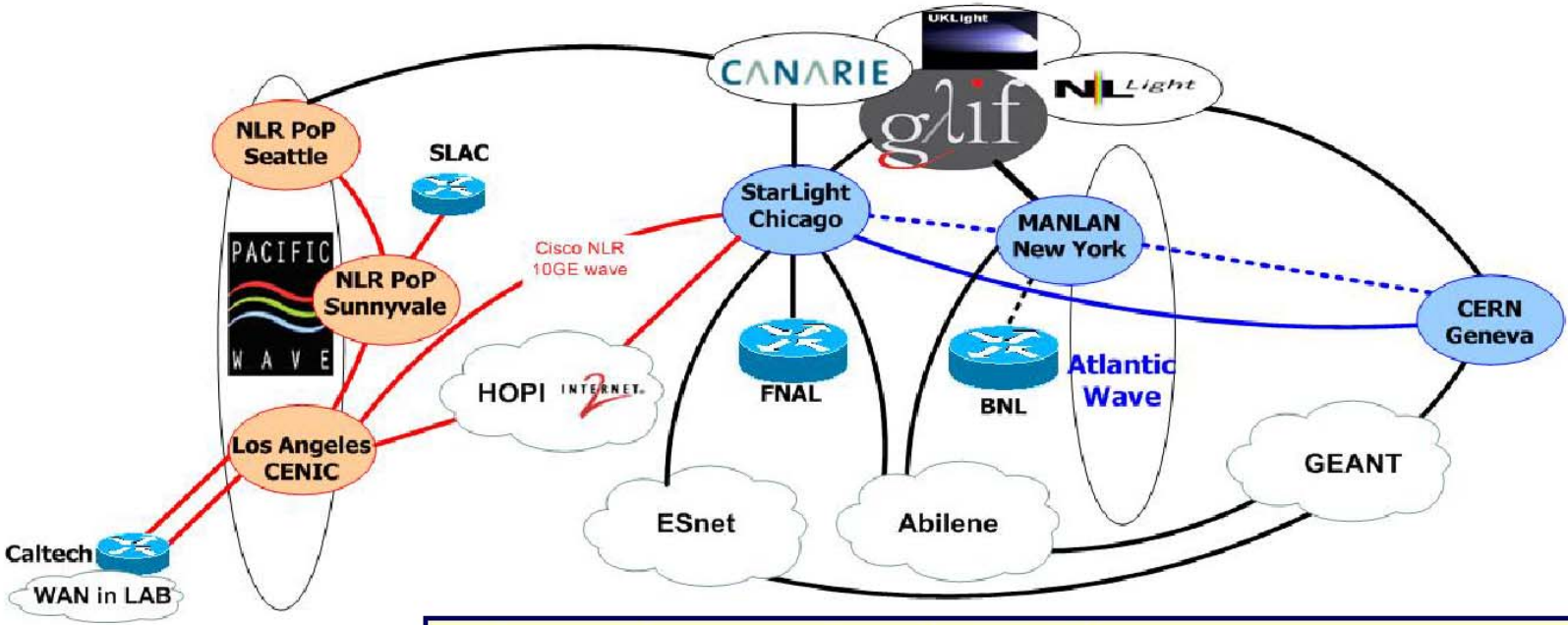
FNAL-CERN Service Challenge
Transfer over LHCNet (Starlight – CERN)

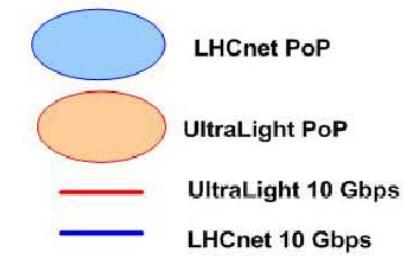
SC2
4/5/2005



Hour

CERN-US: Fall 2005 Plan





- Second transatlantic 10 Gbps wave
- New PoP in New York (MANLAN)
- FNAL: Dark fiber to Starlight => a natural growth path to N X 10G in the future
- BNL: Direct fiber to MANLAN being studied (2006)
- GLIF: Global Lambda Integrated Facility
- CERN RFP: 3rd Wave in 2006; 4th in 2007

DOE Science Network Roadmap

Science Areas	Today *End2End* Throughput	5 years End2End Throughput	5-10 Years End2End Throughput	Remarks
High Energy Physics	0.5 Gb/s	100 Gb/s	1000 Gb/s	High bulk throughput
Climate (Data & Computation)	0.5 Gb/s	160-200 Gb/s	N x 1000 Gb/s	High bulk throughput
SNS NanoScience	Not yet started	1 Gb/s	1000 Gb/s + QoS for Control Channel	Remote control and time critical throughput
Fusion Energy	0.066 Gb/s (500 MB/s burst)	0.198 Gb/s (500MB/ 20 sec. burst)	N x 1000 Gb/s	Time critical throughput
Astrophysics	0.013 Gb/s (1 TByte/week)	N*N multicast	1000 Gb/s	Computational steering and collaborations
Genomics Data & Computation	0.091 Gb/s (1 TBy/day)	100s of users	1000 Gb/s + QoS for Control	High throughput and steering

- See www.es.net/#research

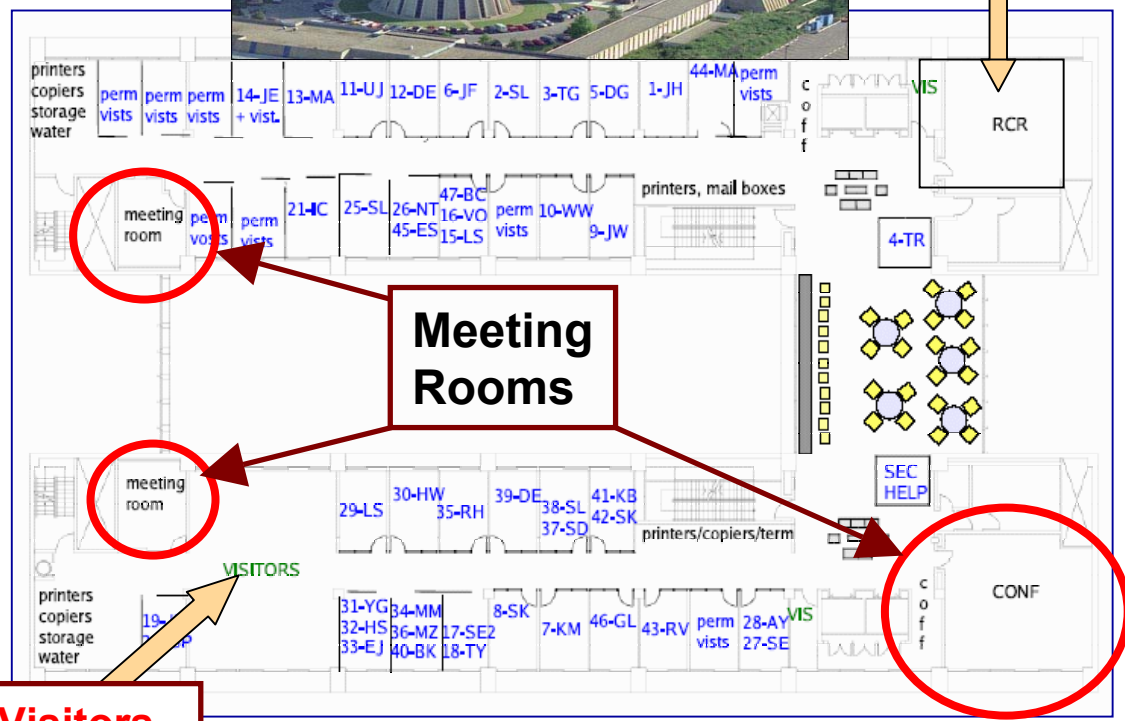
LPC and ROC at Fermilab

LPC (LHC Physics Center)

- A gathering place and a focus of expertise on all aspects of analysis, software and event processing, convenient to US physicists
- A center of excellence for LHC physics
- A place to train students/postdocs
- A place for workshops/conferences
- A development center for software and physics analysis development
- A tool to encourage a graceful Tevatron ➡LHC transition

ROC (Remote Operation Center)

- Test Beam (HCAL, MUON) started in May 2004
- Active shift operation, often-daily conclave to monitor/analyze test beam data from FNAL



Conclusion

LHC and CMS start in 2007. First physics run will be in 2008.

We expect rich physics at the new energy frontier.

We are pushing Grid for global physics analysis.

I hope to come back here in 2008 to report "Discovery in proton-proton collisions at the LHC" .

Additional Slides

Grid3 – a huge success

- Grid3 started out as a demonstration for SuperComputing2003 of a multi-site working Grid
- It ended up a sustained simple Grid infrastructure that could be (and was) used for real work
  - http://www.ivdgl.org/grid3
- Built with leadership from iVDGL and effort from many of the other Grid projects and experiment efforts
- The next step was clearly to create an organization to help us build a sustained Grid infrastructure in the U.S.
  - To provide computing for LHC
  - To support other experiments and other sciences than physics
  - To provide opportunities for CS research and education

Computing Center Specifications

Tier-0 Center

			Eff Factors
CPU scheduled	4588	kSI2K	85.00%
Disk	407	Tbytes	70.00%
Active tape	3775	Tbytes	100.00%
Tape I/O	600	MB/s	50%

Tier-1 Centers

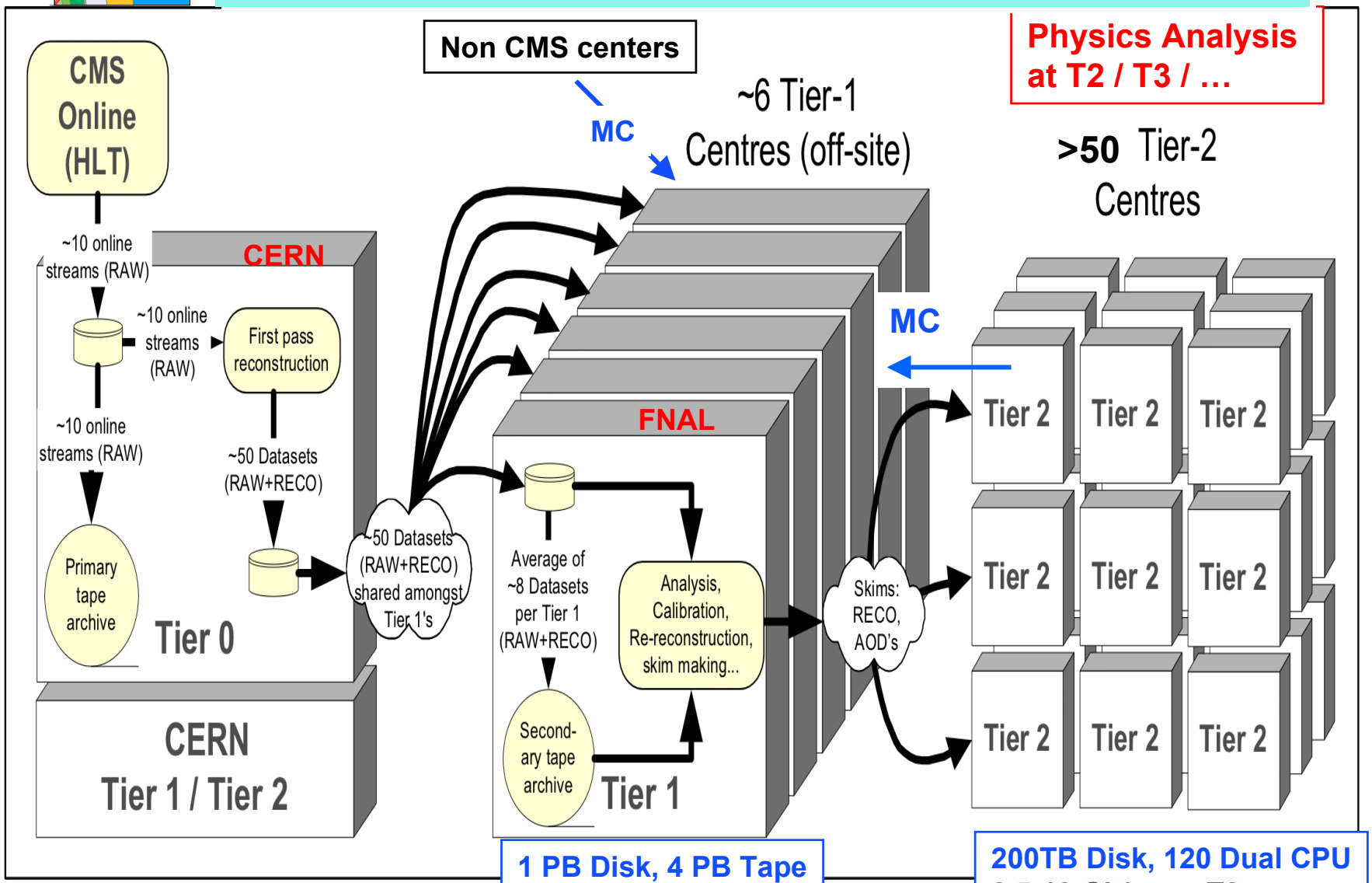
- 1/6
- US-CMS is larger

			Eff Factors
CPU scheduled	1199	kSI2K	85.00%
CPU analysis	929	kSI2K	75.00%
Disk	1121	Tbytes	70.00%
Active tape	1837	Tbytes	100.00%
Data Serving I/O Rate	800	MB/s	

Tier-2 Centers

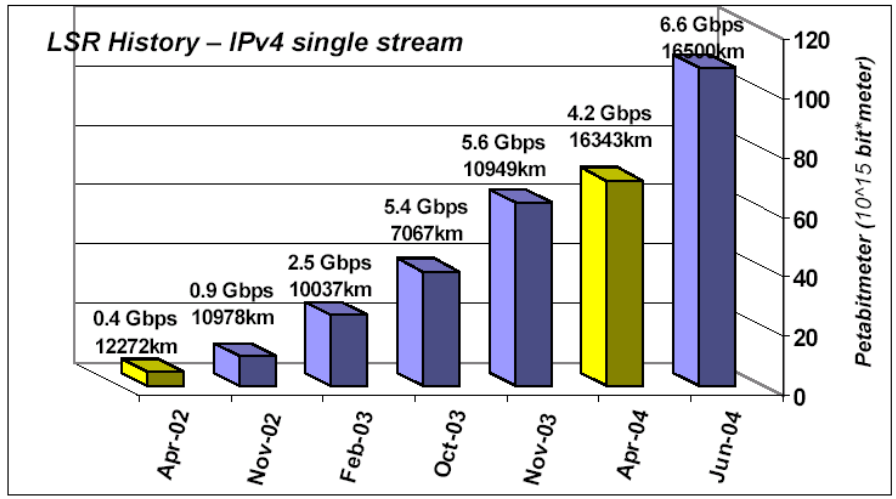
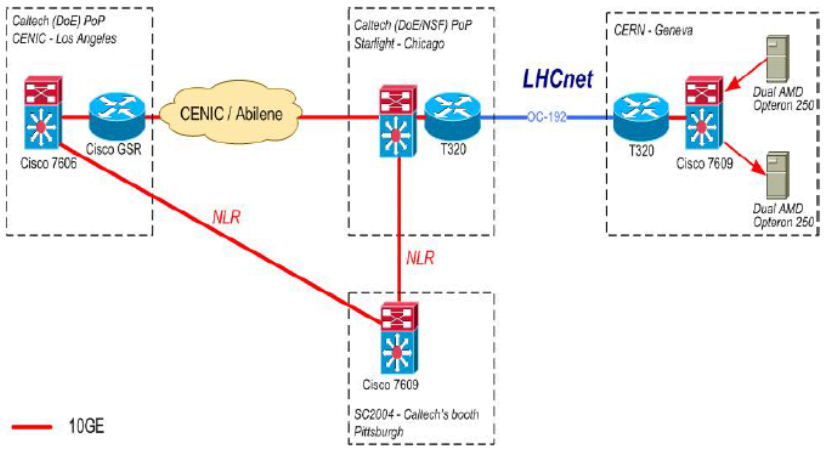
			Eff Factors
CPU scheduled	250	kSI2K	85.00%
CPU analysis	579	kSI2K	75.00%
Disk	218	Tbytes	70.00%

Data Distribution



Internet2: Land Speed Record

- Judged on product of transfer speed and distance end-to-end, using standard (TCP/IP) protocols, Across Production Nets: e.g. Abilene
- IPv4: 6.9 Gbps Caltech-CERN (27 kkm) At SC2004: Grand Tour of Abilene, NLR, US-CERN Forward & Back
- IPv6: 5.1 Gbps Geneva-Phoenix (SC2003)
- 7.48 Gbps X 16 kkm (Linux, 1 Stream) Achieved in July
- 11 Gbps (802.3ad) Over LAN in Sept.
- *Concentrate now on reliable Terabyte-scale file transfers*
  - *Note System Issues: CPU, PCI-X Bus, NIC, I/O Controllers, Drivers*



Nov. 2004 Record Network