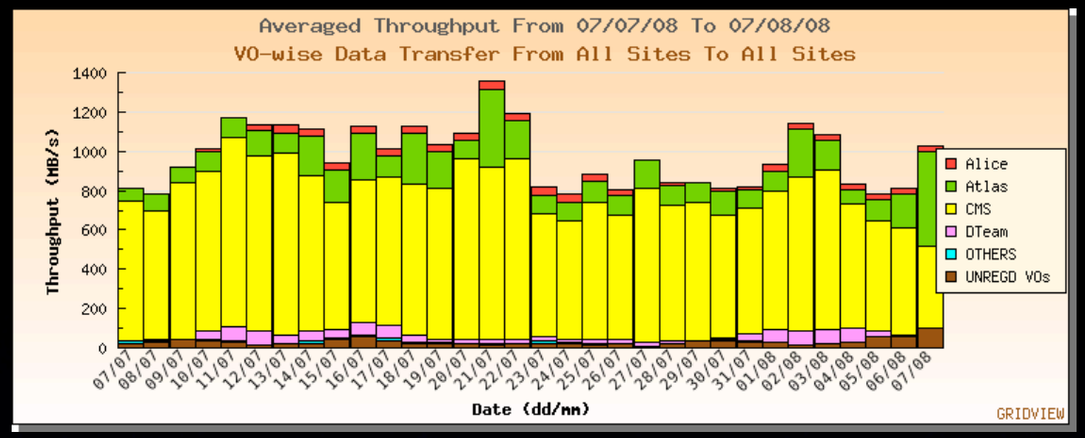


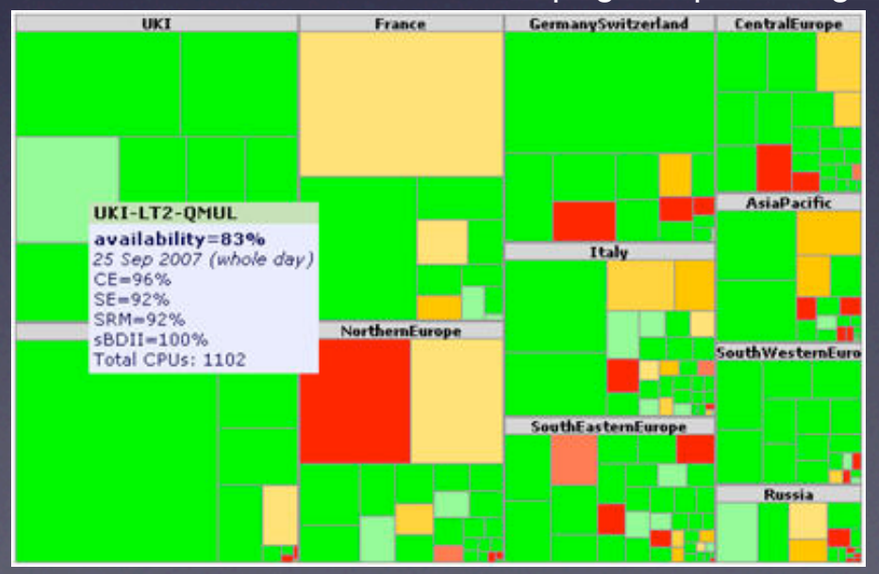
http://gridview.cern.ch/GRIDVIEW/

Computing Models for the LHC experiments

Patricia McBride
Fermilab/CMS

Hadron Collider Physics Summer School 2008

http://gridmap.cern.ch/gm/



https://twiki.grid.iu.edu/twiki/bin/view/Documentation/GlossaryOfTerms


Disclaimer

- This is not a tutorial. Data Analysis will be covered in another session.
- Focus is the development of the computing models and use ATLAS and CMS as examples.
- Tried to give a little history and to avoid using too many TLAs or experiment specific names - not always successfully.
- This is not a lecture on GRIDs. Think of the experiments' computing infrastructure as a layer on the GRID distributed computing facility.

- Many tutorials are available on the GRID. See http://csc.web.cern.ch/CSC/ for example.
- For news on operations of the WLCG GRID, see: https://twiki.cern.ch/twiki/bin/view/LCG/WLCGOperationsWeb
- More detailed ATLAS tutorial: http://indico.cern.ch/conferenceDisplay.py?confId=22616


Principles of LHC Computing

- > 2 PB RAW data expected each year from an LHC experiment.
  - These data must be processed promptly and made available to the users.
  - They must be stored safely for the lifetime of the experiments.
  - Reprocessing will be needed, particularly during the first year.
- Costs of single computing and storage elements have continued to decrease.
  - Unfortunately, we need many...
    - Reliability and availability are major factors => As consumers, we have come to expect computer systems and networks to be up all the time.
    - Procurements take time. We need to plan ahead.
    - Computing infrastructure is expensive and requires lots of power and cooling.
- Reliable networks permit movement of data around the globe (though not always easy).

Users are distributed around the world and expect access to the data.


Data Volume (CMS example - design)

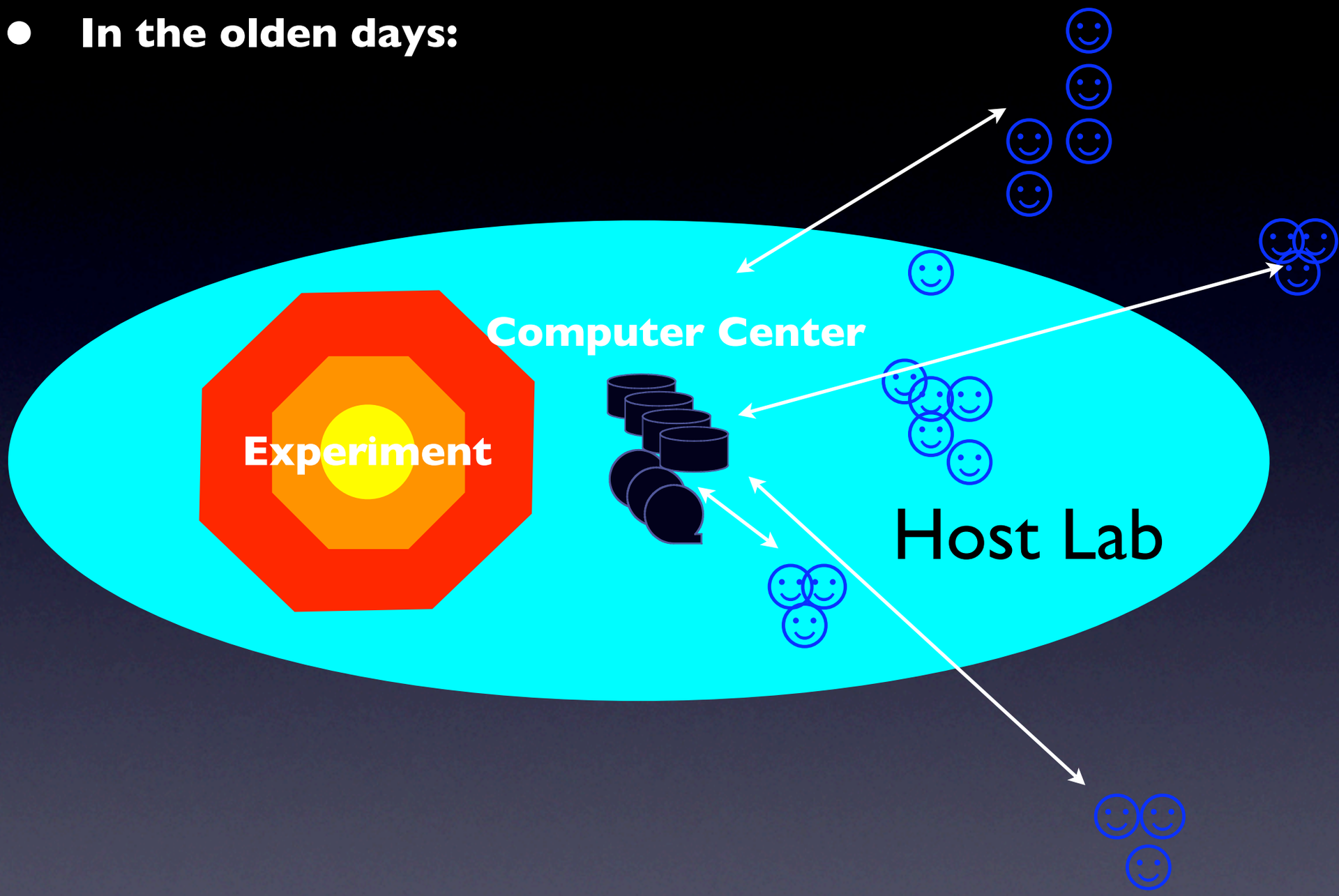
~O(10) PB/year

Event Format	Content	Purpose	Event size	Events/year	Data volume/year
RAW	Detector data, L1 and HLT result info	Input to Tier-0 and for data archive	1.5 MB	3.3×10^9 2 copies	5.0 PByte
RECO	Reconstructed Objects (tracks, vertices, jets) including hits/clusters	Output of Tier-0 reconstruction or re-reconstruction at Tier-1	250-500 kB	8.3×10^9 with reprocessing	2.1 PByte
AOD	Reconstructed Objects. small amount of hit info possible	Physics analysis	50-100 kB	53×10^9 with reprocessing copies at all T1s	2.6 PByte
SIM (RAW)	Generator info, simulated detector info	Simulation information for archive	2 MB	1:1	

assumes 150Hz trigger rate; 1 year = 10^7 sec

A little history

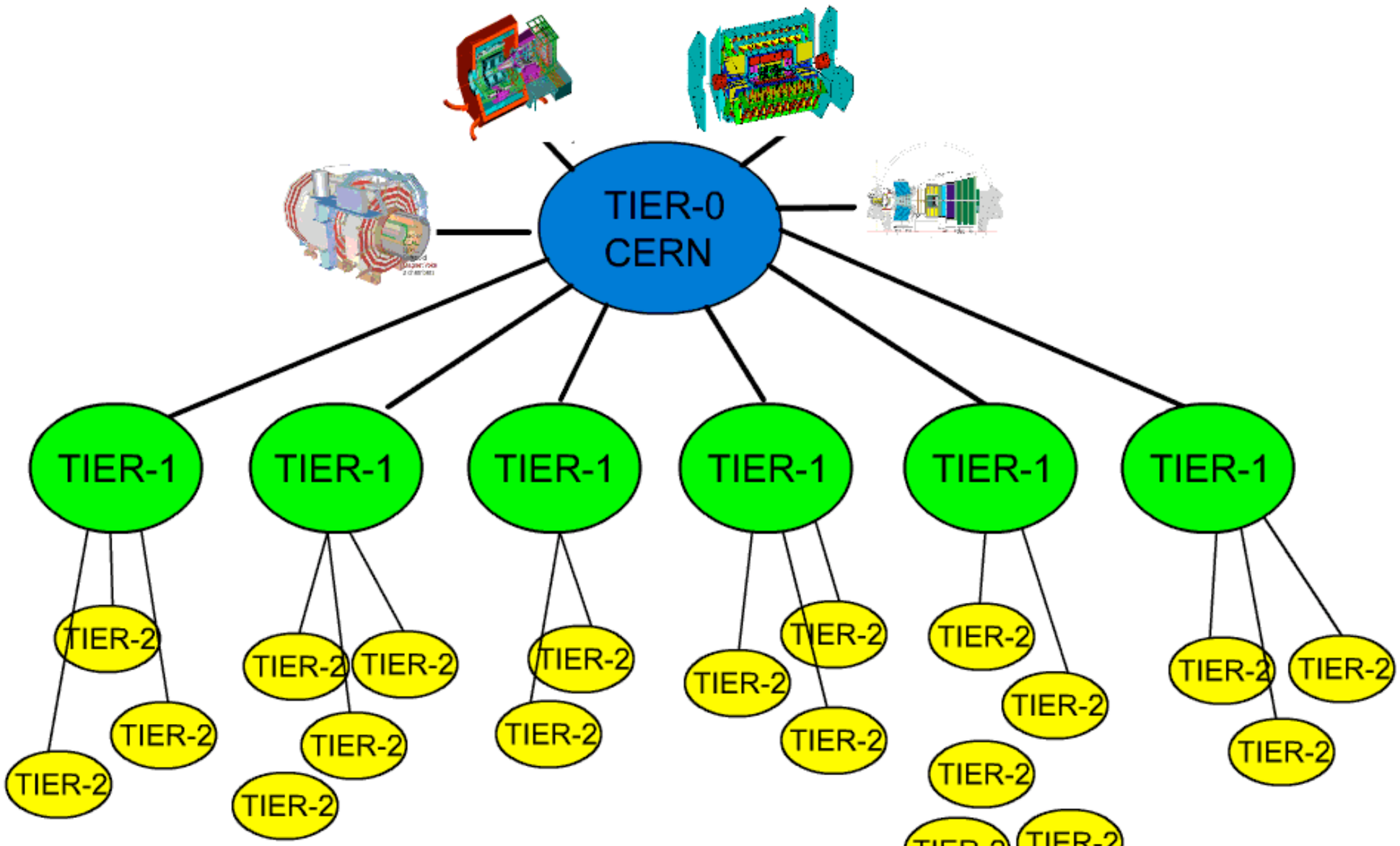
- **In the olden days:**




LHC Computing Models

- Not all computing can be done at the CERN.
- Hierarchical computing models try to optimize the use of resources and maximize user access to the data
- The majority of computing for analysis is NOT at CERN.
- Requirements:
  - 2 archived copies of the RAW data (one at CERN, the second copy elsewhere)
  - Prompt processing will (usually) be done at CERN
  - Simulated data will be produced continually and archived
  - Data will be distributed for reprocessing and data analysis
- Other observations:
  - Needs of the experiments and users will evolve with time
  - Technology may evolve quickly; facilities (buildings, power cooling) evolve slowly

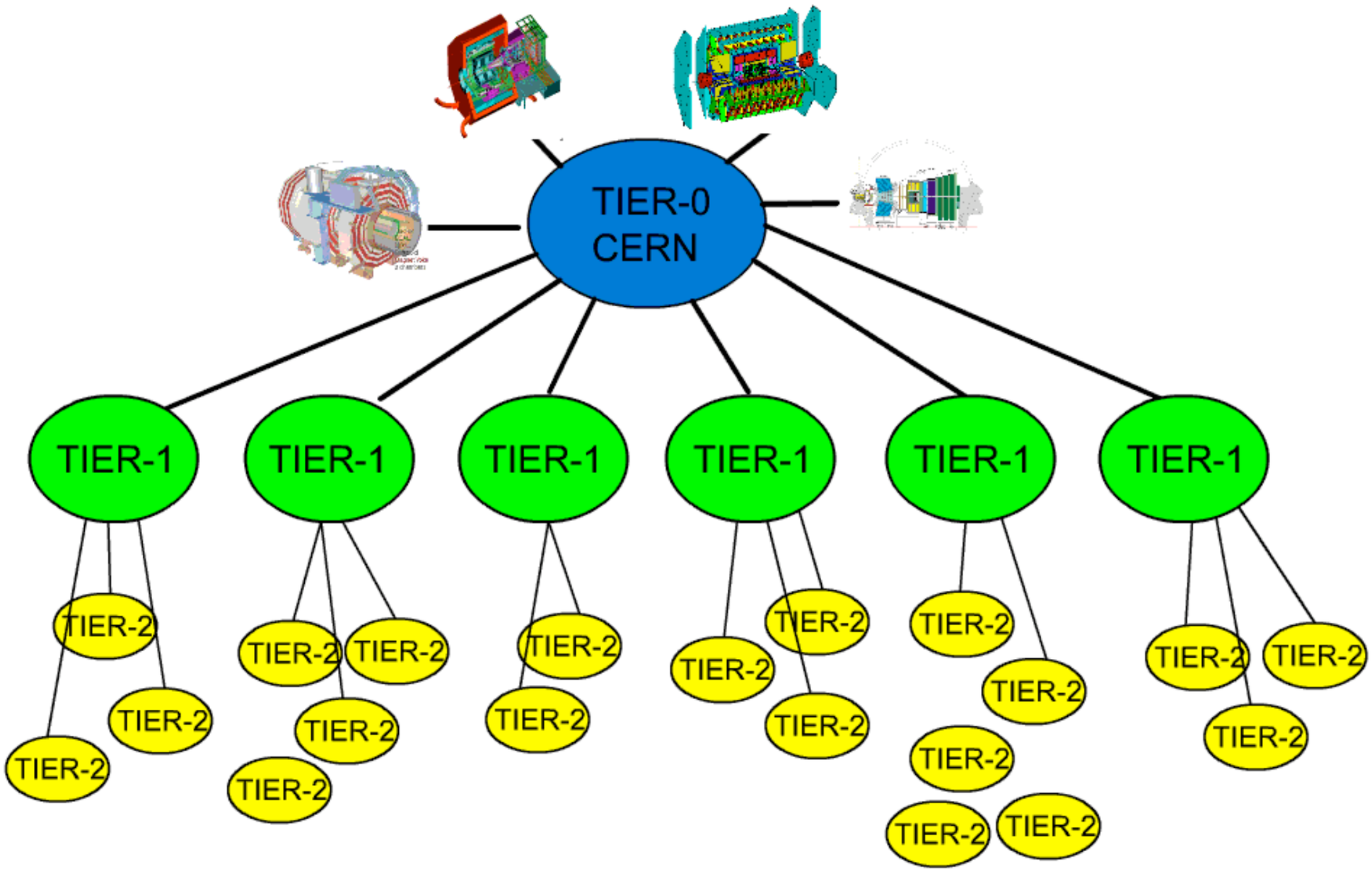
Computing Models



MONARC Tiered computing model came in the late 90's

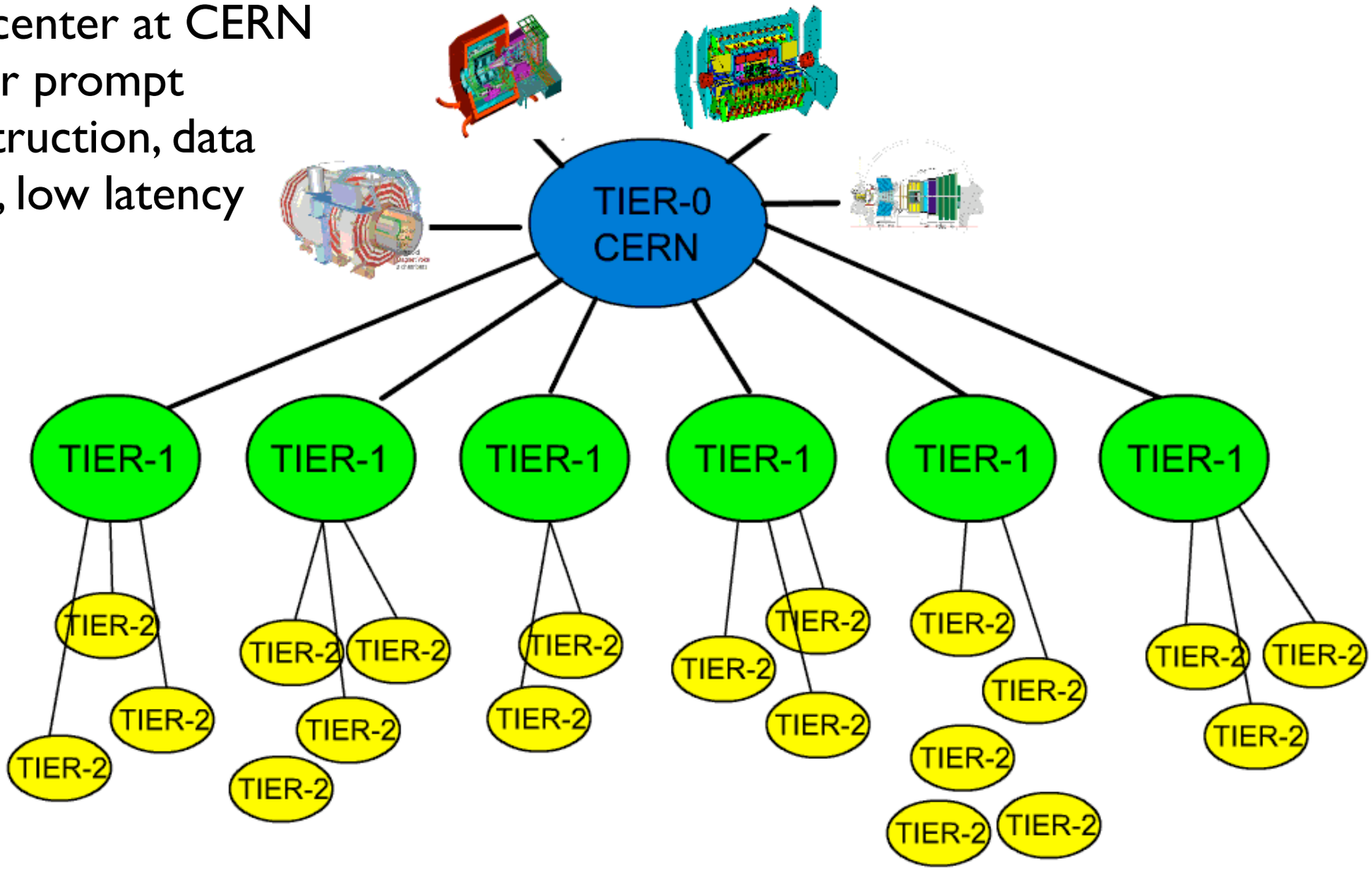
- Level of distribution motivated by the desire to empower and leverage resources and to share load, infrastructure, and funding

Computing Models

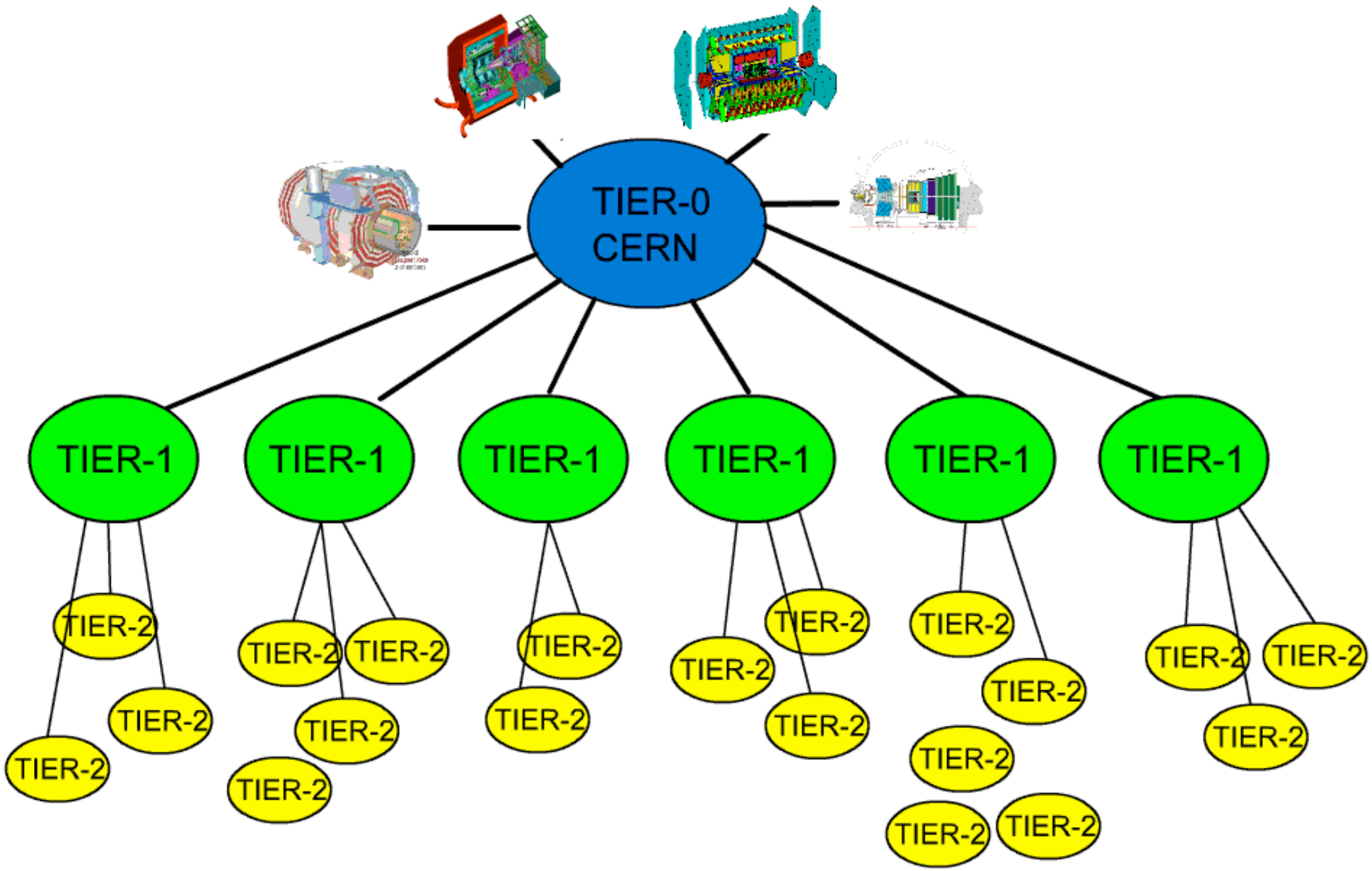


Computing Models

Tier-0 center at CERN used for prompt reconstruction, data archive, low latency work

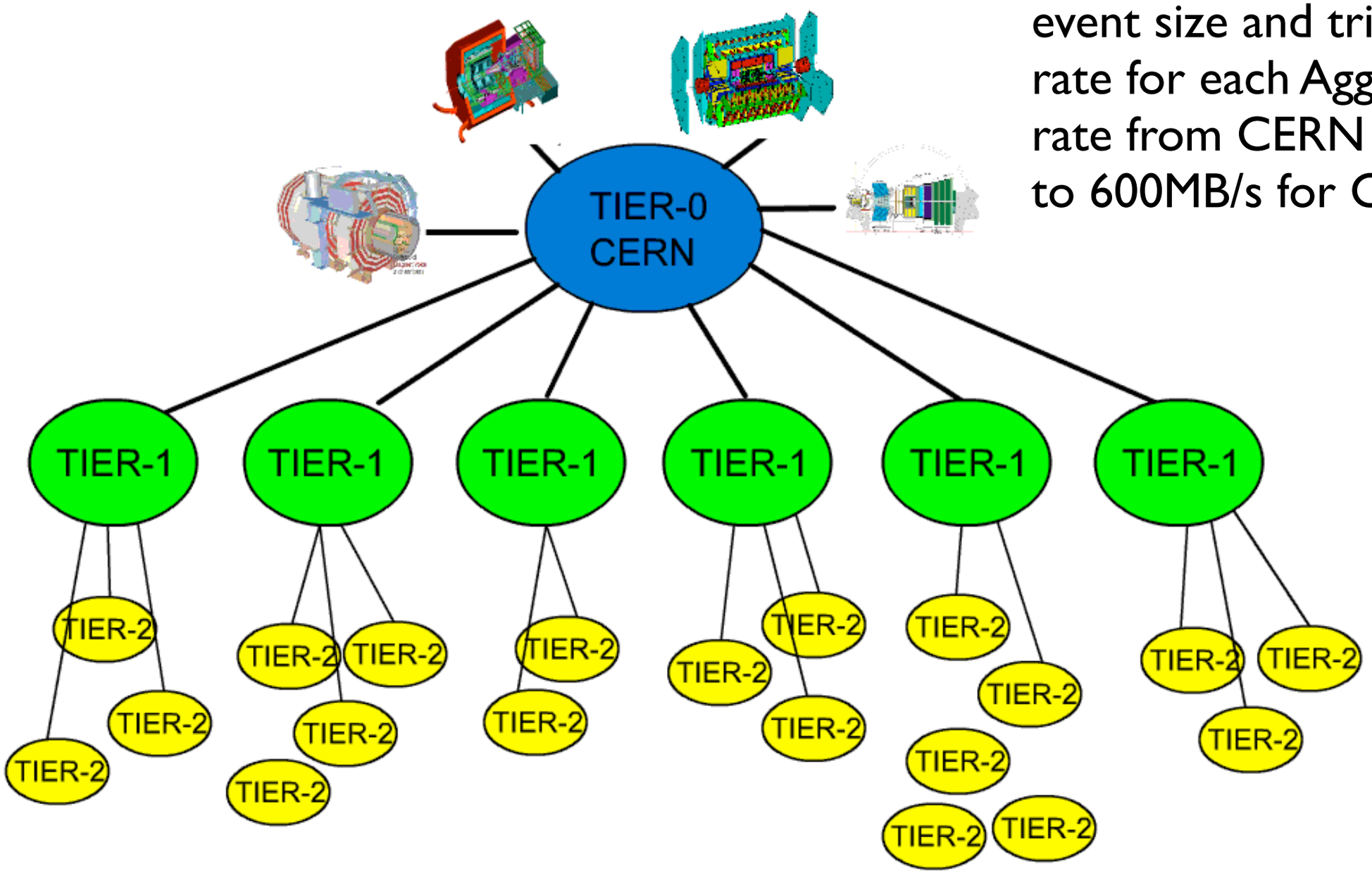


Computing Models

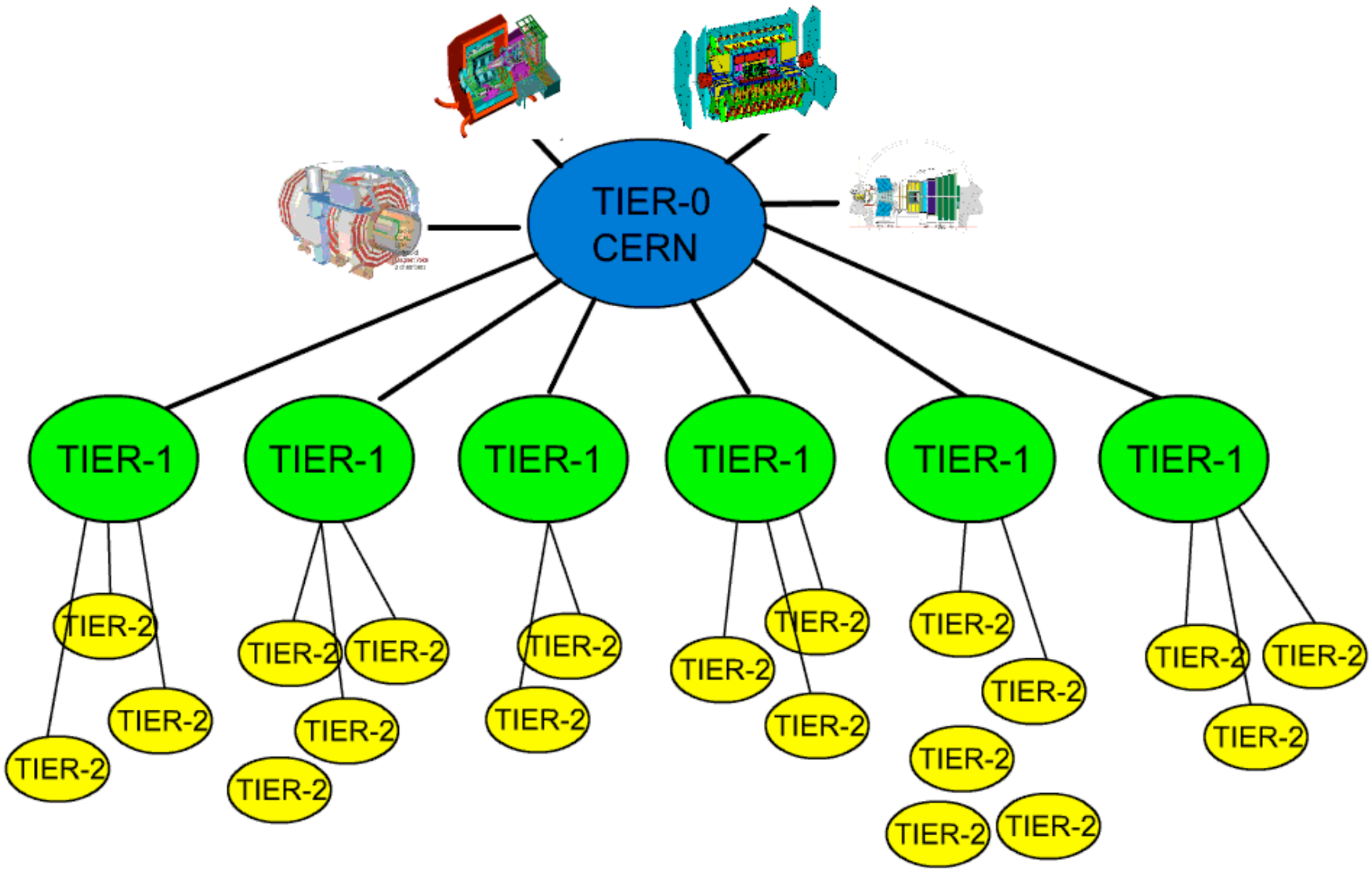


Computing Models

Rate to Tier-1 varies by event size and trigger rate for each Aggregate rate from CERN peaks to 600MB/s for CMS



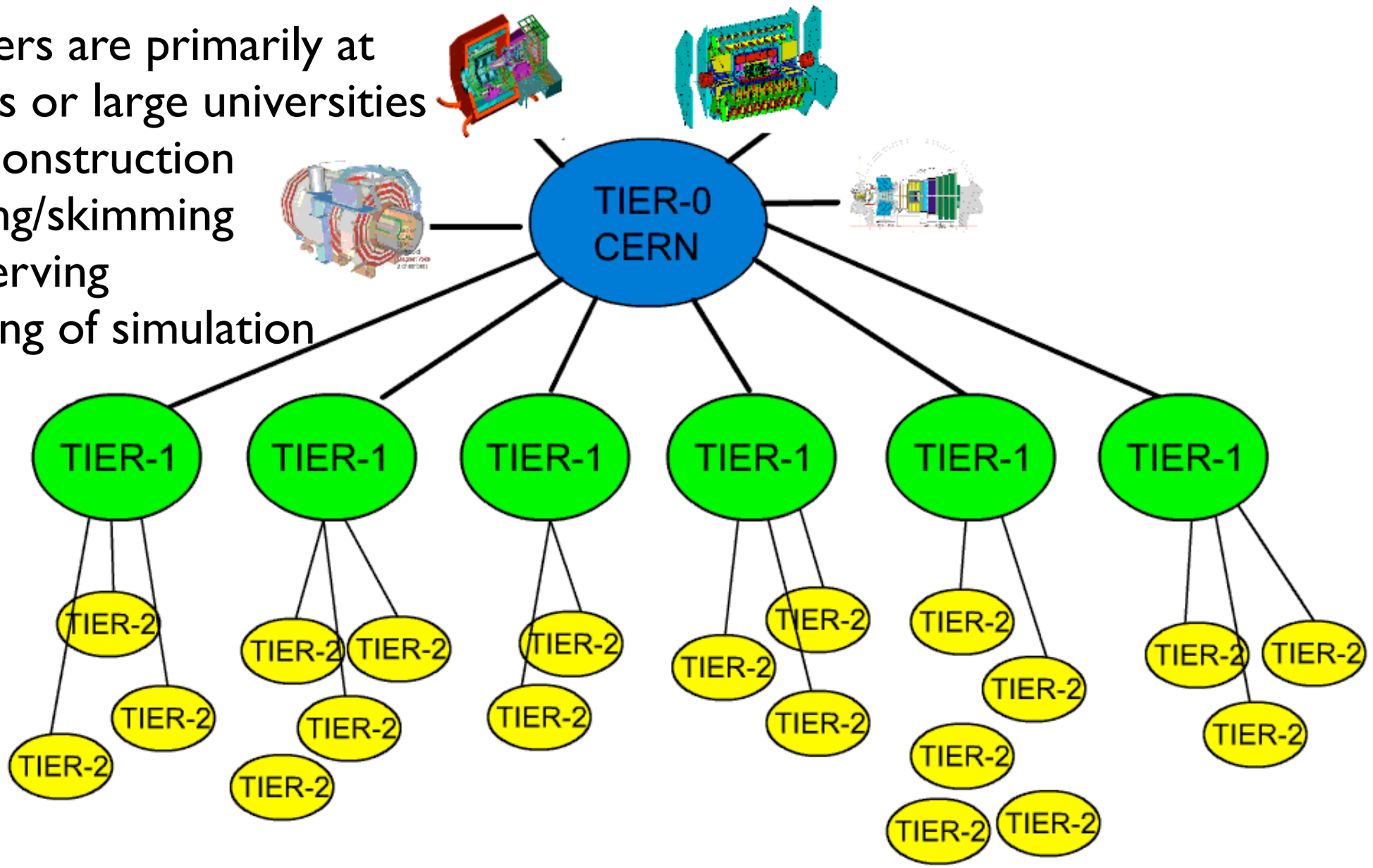
Computing Models



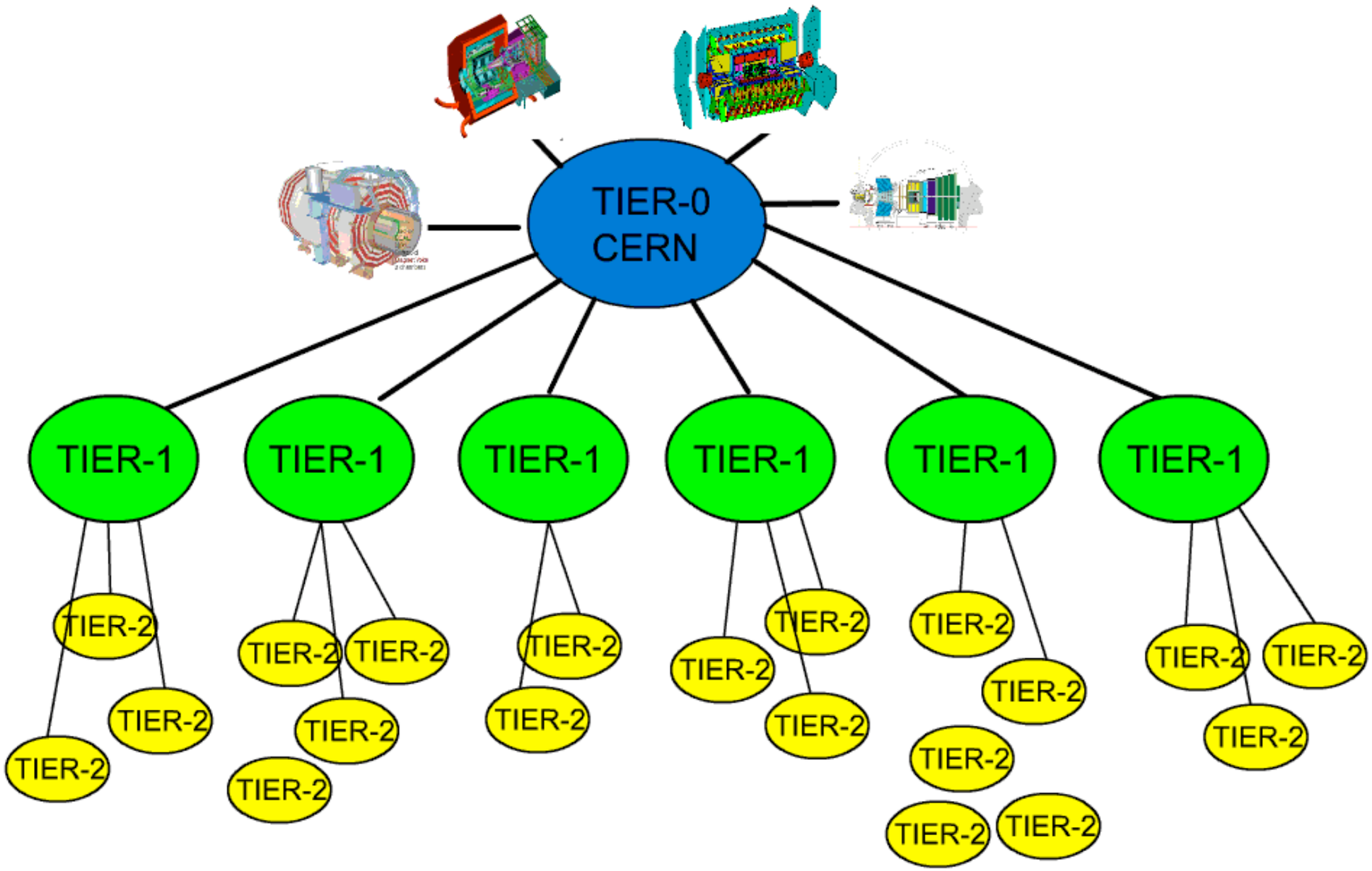
Computing Models

Tier-1 centers are primarily at national labs or large universities

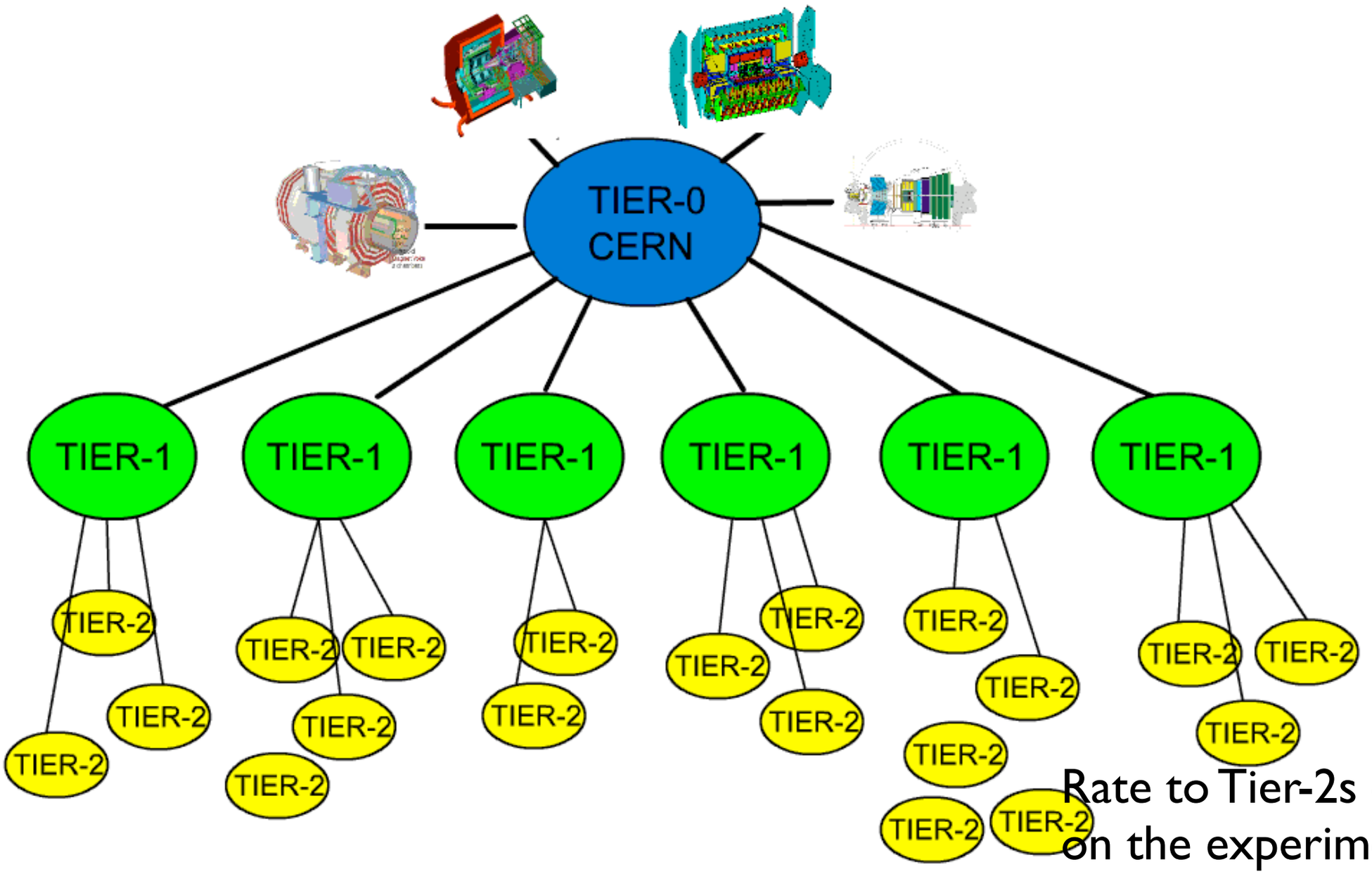
- Re-Reconstruction
- Stripping/skimming
- Data serving
- Archiving of simulation



Computing Models



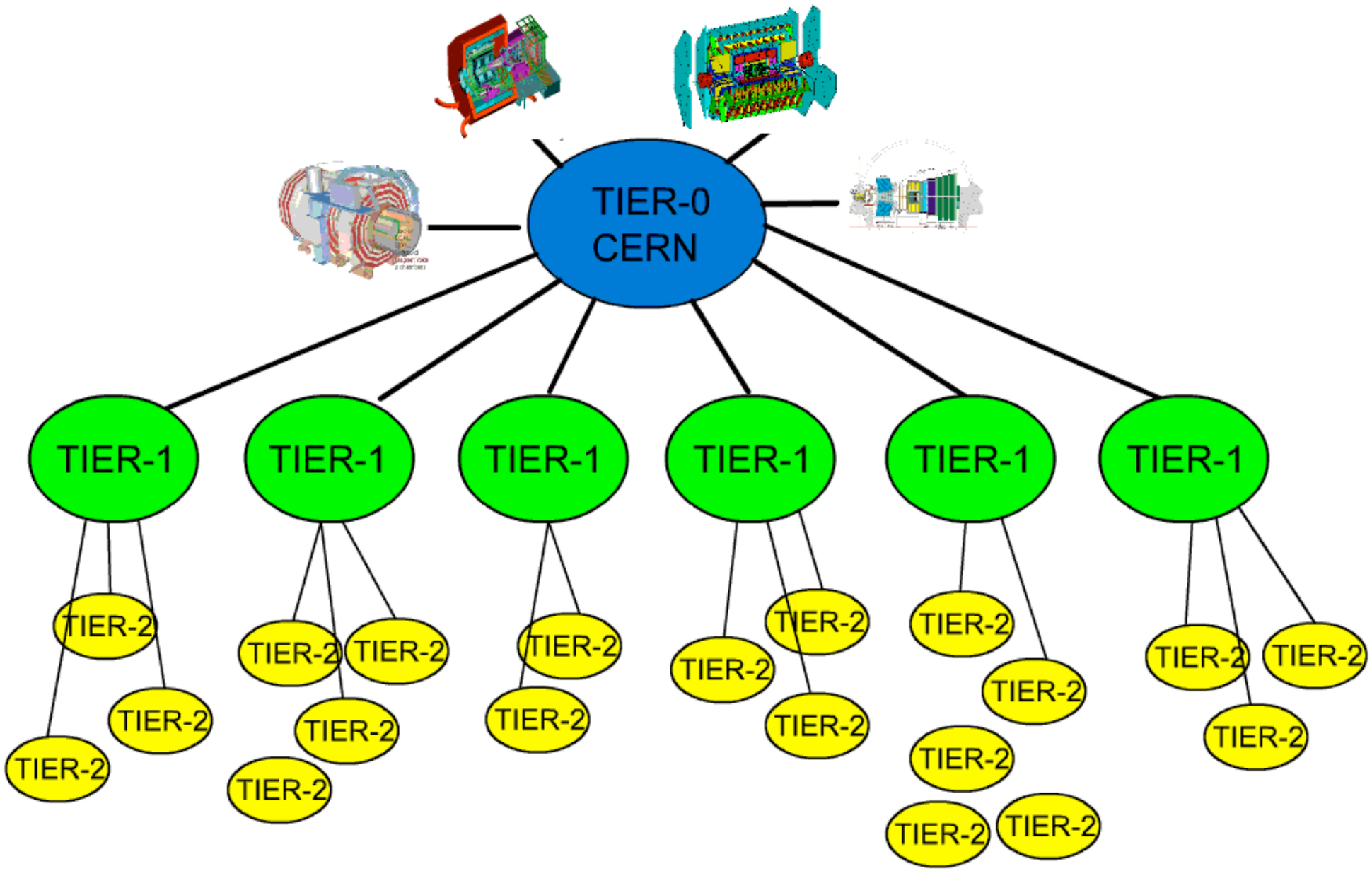
Computing Models



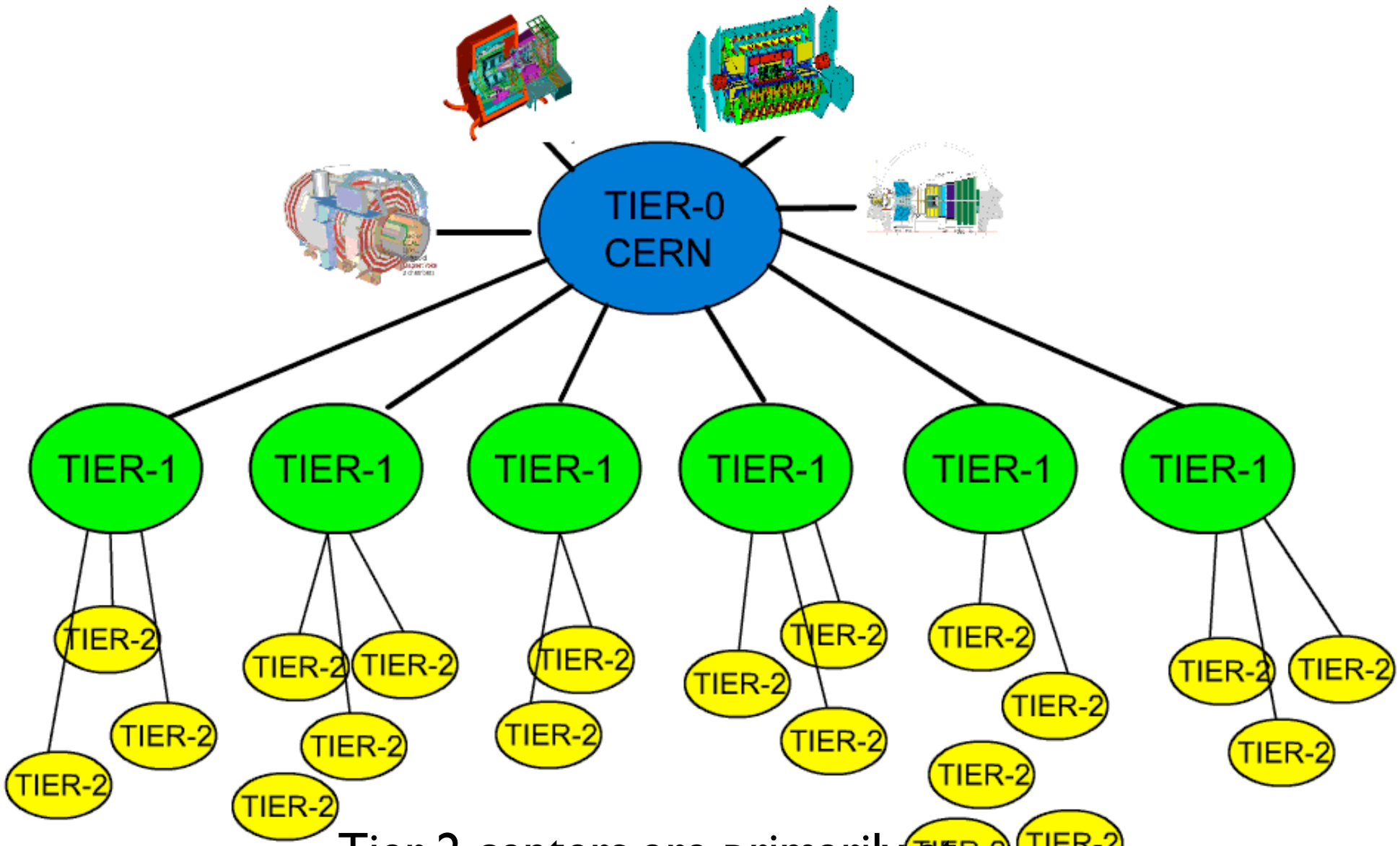
Rate to Tier-2s depends on the experiment and the expectations for updating storage

- Can burst with activity

Computing Models



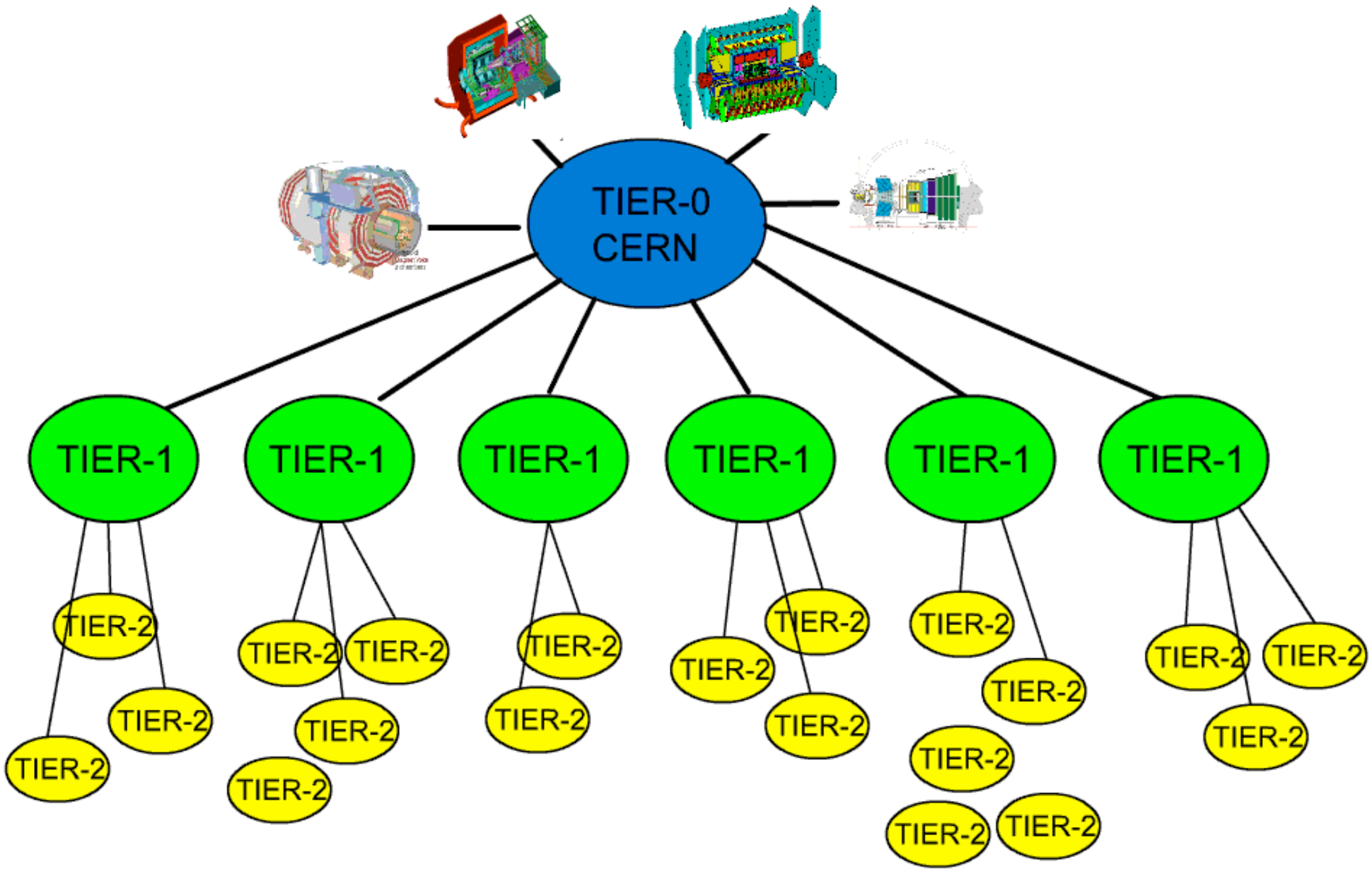
Computing Models



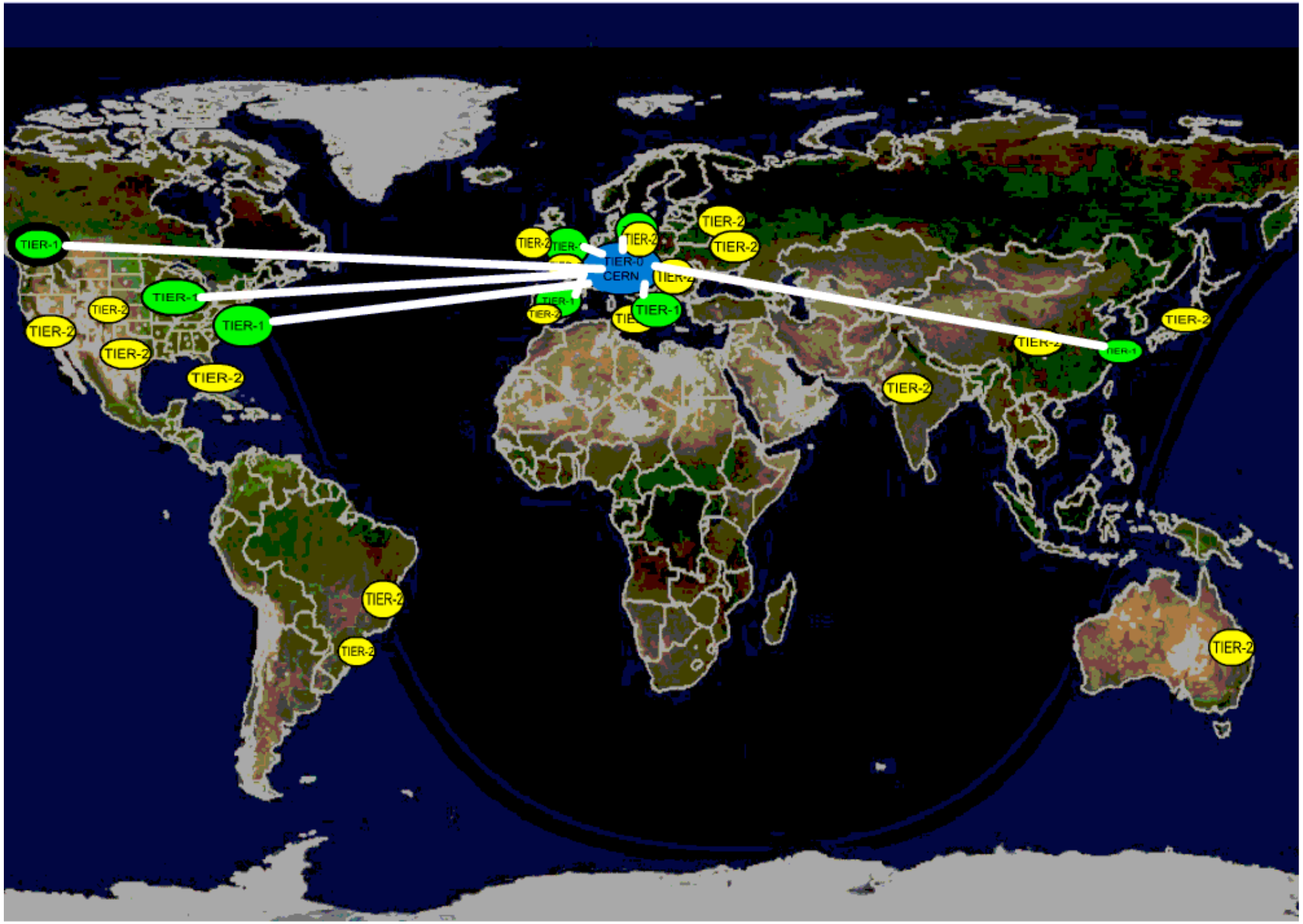
Tier-2 centers are primarily at universities

- Simulation
- User Analysis

Computing Models



Networking



Optical Private Network (OPN) connects CERN and Tier-1. Other connections handled by shared networks

WLCG

Memorandum of Understanding

for Collaboration in the Deployment and Exploitation of the Worldwide LHC Computing Grid

between

The EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH ("CERN"), an intergovernmental Organization having its seat at Geneva, Switzerland, as the Host Laboratory of the Worldwide LHC Computing Grid, the provider of the Tier0 Centre and the CERN Analysis Facility, and as the coordinator of the LCG project,

on the one hand,

and

all the Institutions participating in the provision of the Worldwide LHC Computing Grid with a Tier1 and/ or Tier2 Computing Centre (including federations of such Institutions with computer centres that together form a Tier1 or Tier2 Centre), as the case may be, represented by their Funding Agencies for the purposes of signature of this Memorandum of Understanding,

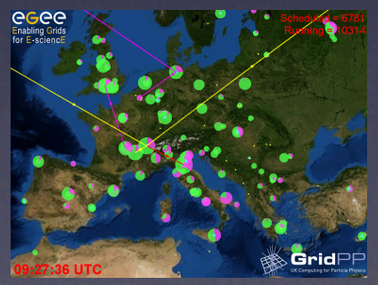
on the other hand,

WLCG MOU contains details of Tier-1, Tier-2 including pledges and commitments to the experiments:
http://lcg.web.cern.ch/LCG/C-RRB/MoU/WLCGMoU.pdf

"The purpose of the Worldwide LHC Computing Grid is to provide the computing resources needed to process and analyse the data gathered by the LHC Experiments. The LCG project, aided by the Experiments themselves, is assembling at multiple inter-networked computer centres the main offline data-storage and computing resources needed by the Experiments and operating these resources in a shared grid-like manner. One of the project's most important goals is to provide common software for this task and to implement a uniform means of accessing resources."


GRID computing - Living in the real world

- The reality is more complicated than imagined ten years ago.
- It was thought that a worldwide grid would be achieved, but technologies diverged and developed along different paths. The WLCG works to bridge the national and international operational GRID infrastructures by enabling interoperability. (EGEE, OSG, NorduGrid, GridPP and INFN Grid)
- In the computing models, it was imagined a nominal center for each Tier and that centers would be similar in size and in level of support. In reality, centers vary in size; a few support only one LHC experiment, some support all four.
- The GRID was a dream 10 years ago. The WLCG and other GRID organizations worked to make it a reality. The WLCG focused on GRID services needed for the experiments and brought the GRID into operation through a set of service challenges.

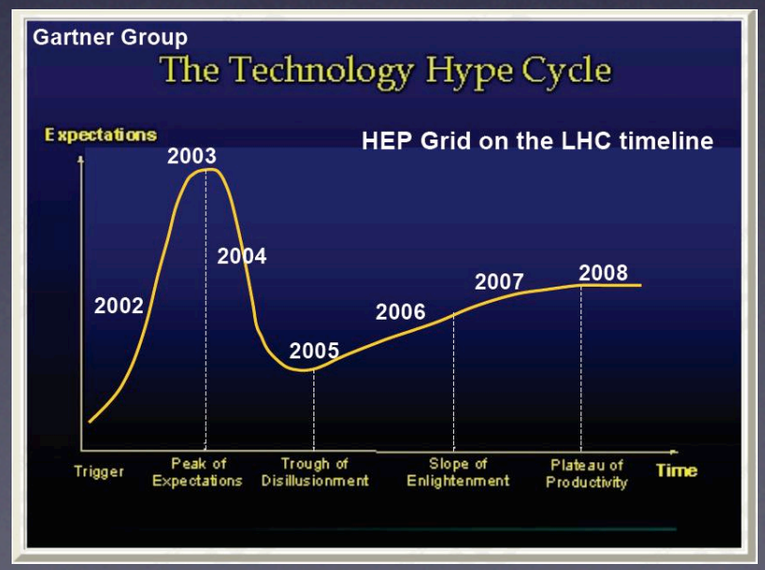


Open Science Grid

Priority WLCG baseline services:

- Transfer tools
- File transfer service
- VO management
- Workload Management
- Database services
- Job monitoring tools
- Information system
- Compute Element
- Storage Element

LCG TDR 2005




Computing resource requirements

- In 2005, the experiments defined their computing models and made projects of resource requirements in terms of capacity of cpu, disk and tape. Planning has been for 5 years period. (These requirements are updated periodically - already planning for 2012-13.)
- WLCG Requirements and Pledges for computing at CERN and T2s:

CPU
Disk
Tape

Split 2008	ALICE	ATLAS	CMS	LHCb	SUM 2008
Required	1800	3705	5300	360	11165
Offered	1800	3710	5300	360	11170
% of Req.	100%	100%	100%	100%	100%
Required	1600	152	400	270	2422
Offered	1600	153	400	270	2423
% of Req.	100%	101%	100%	100%	100%
Required	3300	2449	4400	630	10779
Offered	3300	2450	4400	630	10780
% of Req.	100%	100%	100%	100%	100%

Tier 0

CPU
Disk
Tape

Split 2008	ALICE	ATLAS	CMS	LHCb	SUM 2008
Required	500	2081	2100	0	4681
Offered	500	2080	2100	0	4680
% of Req.	100%	100%	100%	100%	100%
Required	100	1146	1800	80	3126
Offered	100	1146	1800	80	3126
% of Req.	100%	100%	100%	100%	100%
Required	0	370	900	0	1270
Offered	0	370	900	0	1270
% of Req.	100%	100%	100%	100%	100%

CAF

CERN Computing 2008

WLCG Tier 2 (20-40 per experiment)

CPU
Disk

Split 2008	ALICE	ATLAS	CMS	LHCb	SUM 2008
Offered	6572	17946	17467	4223	46208
Required	12500	17510	13400	4550	47960
Balance	-47%	2%	30%	-7%	-4%
Offered	1365	6466	4464	139	12434
Required	1700	7770	5100	9	14579
Balance	-20%	-17%	-12%	1443%	-15%

2008 T2

CPU
Disk

	SUM 2008	SUM 2009	SUM 2010	SUM 2011	SUM 2012
Offered	46208	69352	106749	138120	157151
Required	47960	80750	164540	218720	275010
Balance	-4%	-14%	-34%	-37%	-43%
Offered	12434	22238	34116	43115	49320
Required	14579	23023	34243	47903	61963
Balance	-15%	-3%	0%	-10%	-20%

T 2

Units: CPU: MSI2k, Disk :TB, Tape:TB
2MSI2k ~ 1000 batch slots in 2008


Level of Service at the Centers

Service Tier-1	*Maximum delay in responding to operational problems*			*Average availability*[2] *measured on an annual basis*	
	Service interruption	Degradation of the capacity of the service by more than 50%	Degradation of the capacity of the service by more than 20%	During accelerator operation	At all other times
Acceptance of data from the Tier-0 Centre during accelerator operation	12 hours	12 hours	24 hours	99%	n/ a
Networking service to the Tier-0 Centre during accelerator operation	12 hours	24 hours	48 hours	98%	n/ a
Data-intensive analysis services, including networking to Tier-0, Tier-1 Centres outwith accelerator operation	24 hours	48 hours	48 hours	n/ a	98%
All other services[3] – prime service hours[6]	2 hour	2 hour	4 hours	98%	98%
All other services[3] – outwith prime service hours[6]	24 hours	48 hours	48 hours	97%	97%

Service Tier-2	*Maximum delay in responding to operational problems*		*Average availability*[2] *measured on an annual basis*
	Prime time	*Other periods*	
End-user analysis facility	2 hours	72 hours	95%
Other services[3]	12 hours	72 hours	95%

Tier-0 and T1 Level of Service: availability of 97-99% over the year. This is not easy to achieve.

The WLCG has defined expected response times for services.

Tier-2 level of service does not demand weekend response to problems.

Prime time: 8-18 in the time zone of the site Mon-Friday expect on holidays.

WLCG Tier-1 Centers

Tier-1 Center	ALICE	ATLAS	CMS	LHCb
Canada, TRIUMF		x		
France, CC-IN2P3	x	x	x	x
Germany, FZK-GridKA	x	x	x	x
Italy, CNAF	x	x	x	x
Netherlands LHC/Tier1	x	x		x
Nordic Data Grid Facility	x	x		
Spain, PIC		x	x	x
Taipei, ASGC		x	x	
UK, RAL	x	x	x	x
USA, BNL		x		
USA, FNAL			x	

Planning for computing resources is done 3-5 years in advance.

Planning for facilities (buildings, cooling, power) takes even longer.

Purchasing can take 6-9 months!

Installation and commissioning also takes time.

Split 2008		ALICE	ATLAS	CMS	LHCb	SUM 2008
CPU	Offered	5541	19195	10291	2536	37563
	Required	10100	18120	9600	1770	39590
	Balance	-45%	6%	7%	43%	-5%
DISK	Offered	2395	10913	5546	1367	20221
	Required	4000	10730	7200	1025	22955
	Balance	-40%	2%	-23%	33%	-12%
TAPE	Offered	2983	7692	9429	1194	21298
	Required	5800	8070	9800	860	24530
	Balance	-49%	-5%	-4%	39%	-13%

2008 Tier-1 resources


LHC Computing Layers

- **Hardware and site fabric** - cpu, disk, robots, drives, batch systems, linux, storage systems
  - **GRID middleware and services** (from EGEE, OSG, NorduGrid) - uniform environment for job execution, interface for jobs and data, security, accounting
    - **Experiment services** - Data management and bookkeeping, calibration services, workload management, application software installation
      - **User applications** - physics and analysis codes, using the application development framework provided by the experiment - the purpose of the Framework and Event Data Model is to facilitate the development and deployment of reconstruction and analysis software.

- GOAL - A global environment where physicists can develop and share code and also easily find their dataset, access the data and run jobs where the datasets are located.


Role of the Tier-0

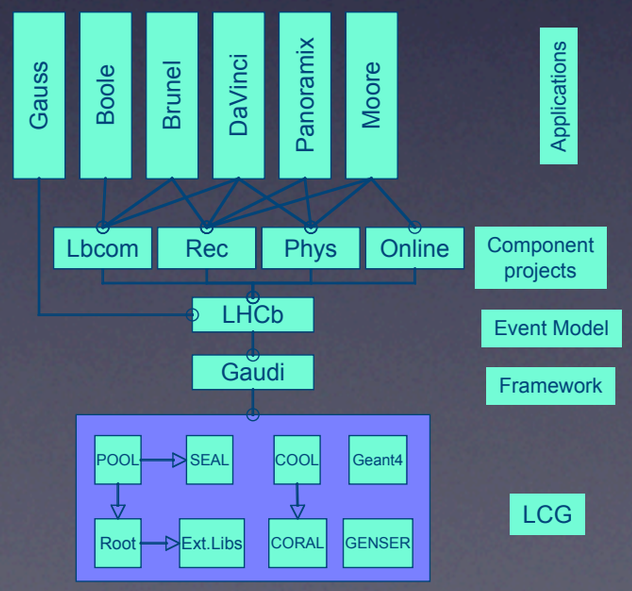
- Primary computing infrastructure at CERN (CPU + Disk + Mass Storage)
- State of the art networking - OPN
  - connectivity to the experiments
  - connectivity to the Tier-1 centers
- Requirements for custodial storage of datasets -
  - RAW data archive
  - Archive of reconstructed data and AOD
- Service level: 24x7 service
- Major roles in the experiments: **Prompt Reconstruction, data archiving, prompt alignment and calibration (T0/CAF) , transfer of RAW/RECO to T1s**

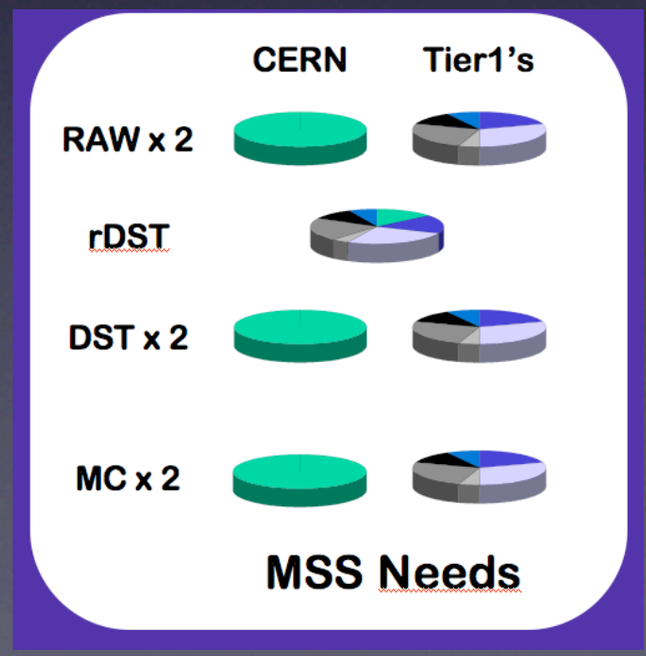

The Role of the Tier-1 (T1) Centers

- Computing infrastructure (CPU + Disk Storage + Mass Storage) on the Grid
- State of the art network -
  - 10Gb/s connectivity CERN on OPN, connectivity to other T1s
  - connectivity requirements to the T2s differ between CMS and ATLAS
- Requirements for custodial storage of datasets -
  - RAW data must remain accessible throughout the lifetime of the LHC
  - Long-term access to RECO/SIMU and analysis datasets
- National (regional) support role including training and user support
- Service level: 24x7 service
- Roles in the experiments: **(Re-)Reconstruction, analysis skims, archiving**
  - For ATLAS and CMS, the Tier-1s will be restricted to production and "organized" tasks


LHC-b Computing

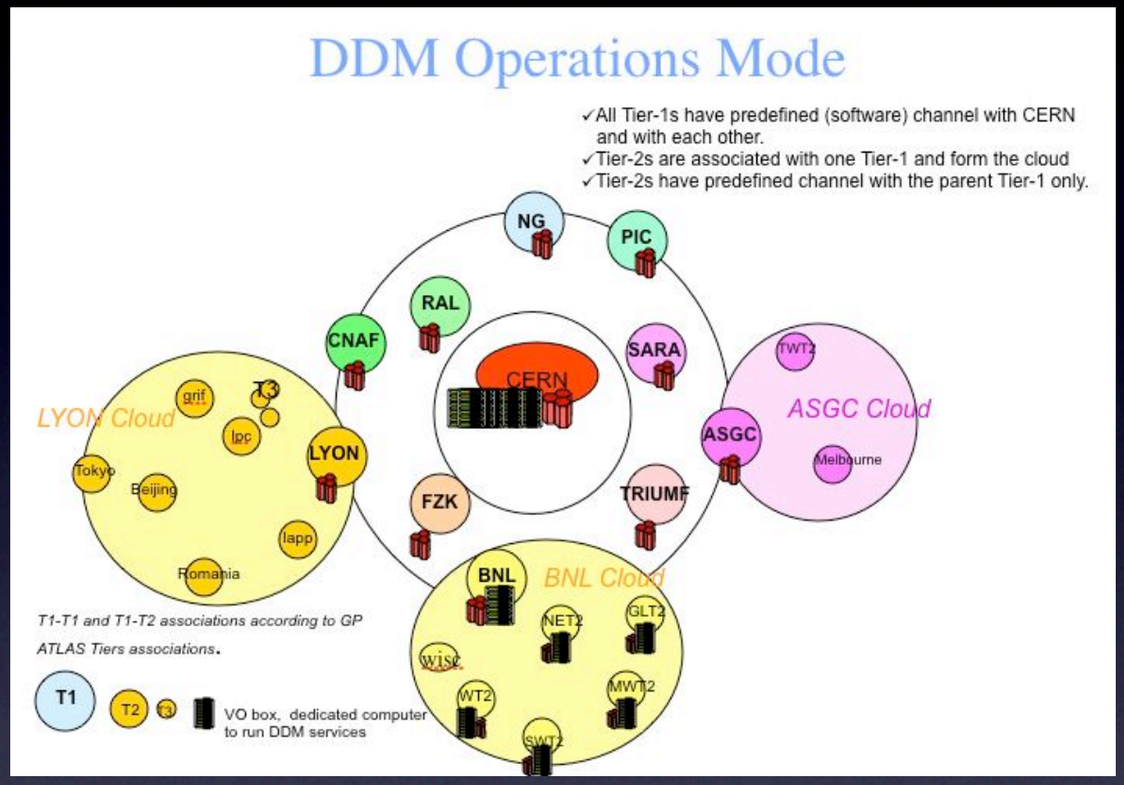
- LHC-b has a slightly different model (smaller event size, higher trigger rate); all data processing is distributed.
  - MC simulation dominates the CPU needs and will be done at sites with available resources (Tier-2,1,0)
  - reconstruction and stripping at will be done at the Tier-0 AND Tier-1s with the output of the stripping distributed to all Tier-1/Tier-0 centers
  - Analysis is planned to be at Tier-1 and Tier-0 but may move to Tier-2 centers. Jobs will run where the data are. Stripped DSTs will be on stored on disk at the T0/1.
- The analysis user interface is Ganga and DIRAC with integrated Workload and Data management is used for production jobs.






Atlas on the Grid

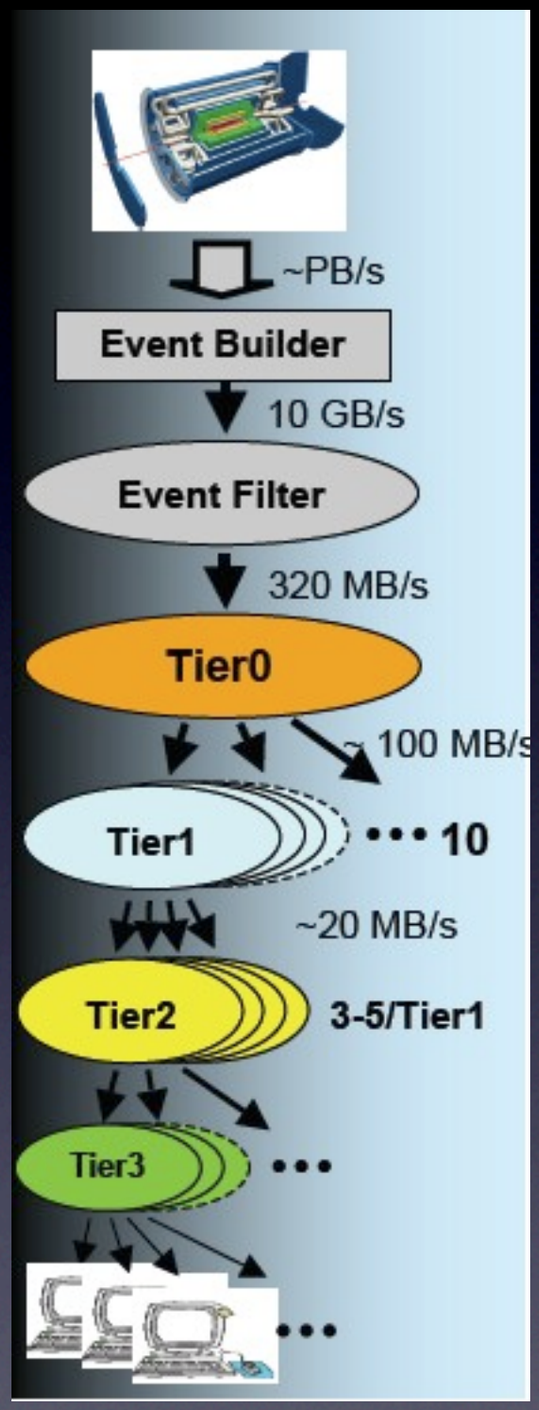
- ATLAS GRID Architecture has 4 Main Components
  - Distributed Data Management (DDM)
  - Distributed Production System (ProdSys)
  - Distributed Analysis (DA)
  - Monitoring and Accounting



- The Distributed Data Management (DDM) is the central link
  - A hierarchical definition of datasets
  - Central dataset catalogues
  - Data blocks as units of storage and replication
  - Distributed file catalogues
  - Automatic data transfer mechanisms (dataset subscription system)


Some definitions

- RECO or ESD: Primary output of the reconstruction from DATA
- SIMU: Simulated data that has been reconstructed; also contains generator information and MC truth
- RAW: data from the detector, input into the reconstruction at the Tier-0
- AOD: reduced event data for analysis
- TAGS

ATLAS planning for data distribution

RAW

- Original copy at Tier-0
- Complete replica distributed among all Tier-1 centers
  - randomized datasets to make reprocessing more efficient

ESD

- ESD produced by primary reconstruction reside at Tier-0 and are exported to 2 Tier-1s
- Subsequent versions of ESDs, produced at Tier-1s are stored locally and replicated to another Tier-1

AOD

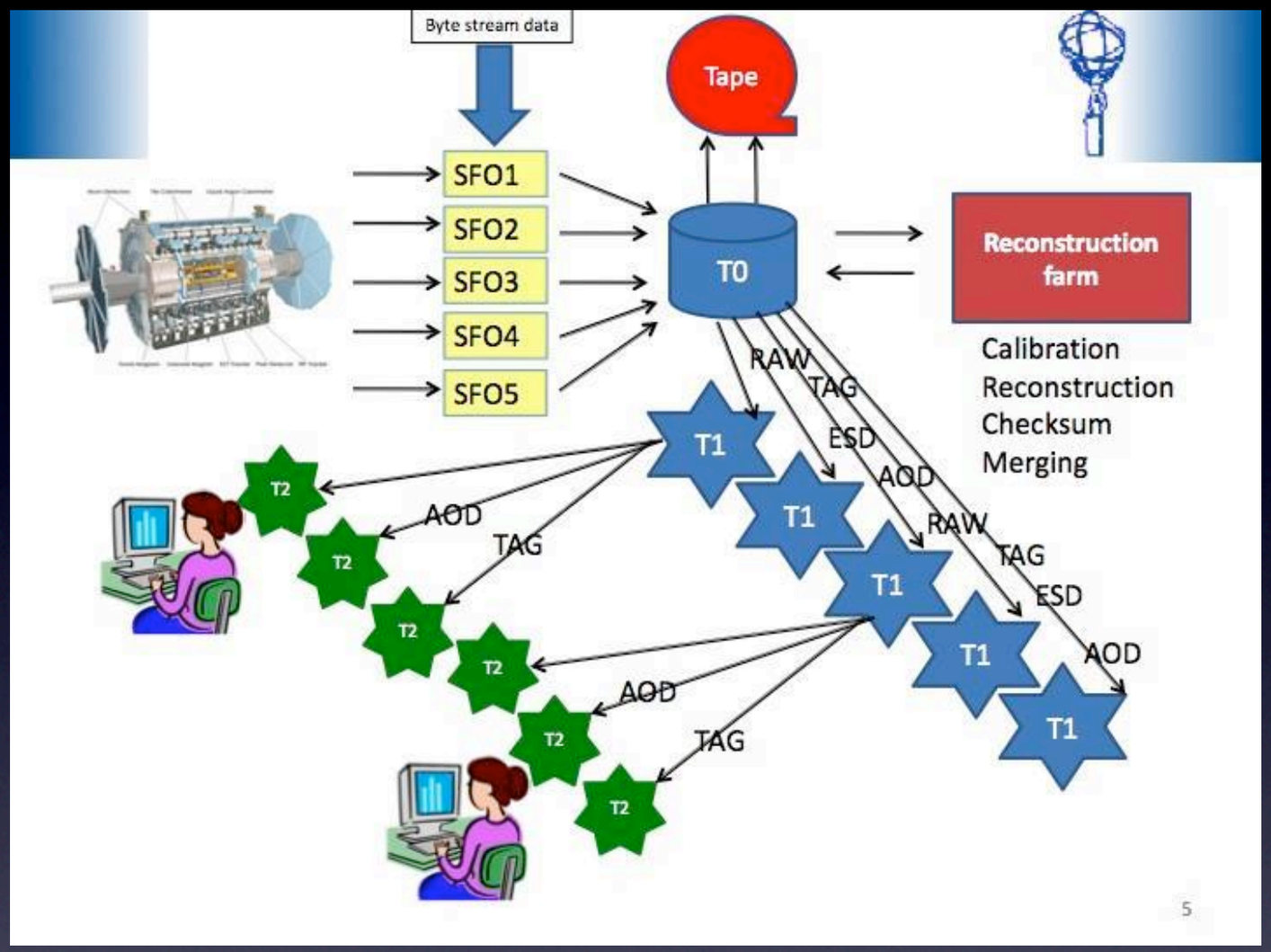
- Completely replicated at each Tier-1
- Partially replicated to Tier-2s so as to have at least a complete set in the Tier-2s associated to each Tier-1.

TAG

- TAG databases are replicated to all Tier-1s
- Partial replicas of the TAG will be distributed to Tier-2s as ROOT files.

Other data samples can be stored "anywhere" subject to availability.


Data Flow and Analysis Datasets



ATLAS
Derived Physics Datasets (DPD)

- Group-level DPDs will be produced in scheduled activity at Tier-1s with an overall coordinator and production people in each group
- User-level DPDs can be produced at Tier-2 and brought "home" to Tier-3s or desk/lap-tops if small enough
- Derived Physics Datasets (DPD) in ATLAS will consist (for most analyses) of skimmed/slimmed/ thinned AODs plus relevant blocks of computed quantities (such as invariant masses)

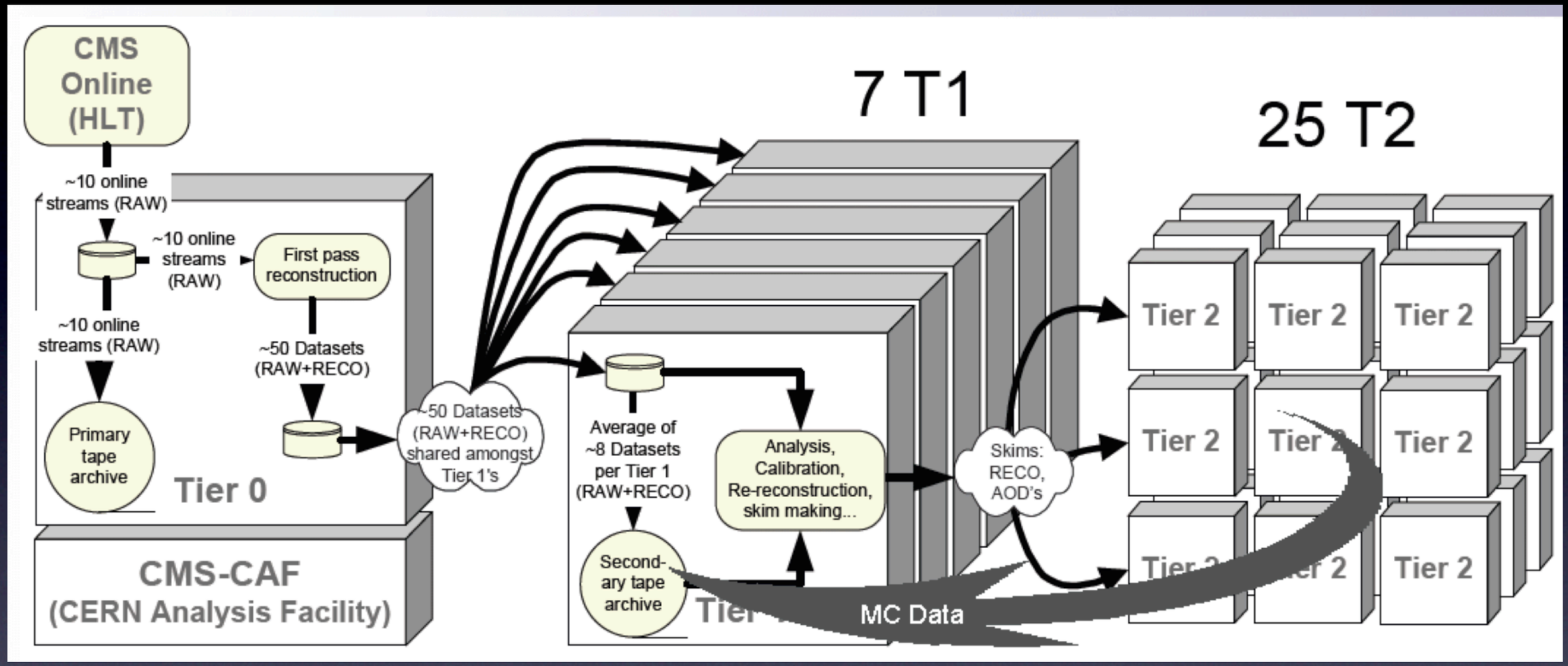
DPDs are stored in the same format as ESD and AOD so they will readable by Athena and ROOT.

CMS Data Model

The CMS Event Data Model (EDM) uses the concept of an *Event* as a C++ object container. During processing, data are passed from one module to the next via the Event, and are accessed only through the Event. All objects in the Event may be individually or collectively stored in ROOT files, and are thus directly browsable in ROOT.

Physics data sets for physics group analysis is planned to be made from organized skims of primary datasets (output usually in AOD format).
https://twiki.cern.ch/twiki/bin/view/CMS/WorkBookCMSSWFramework

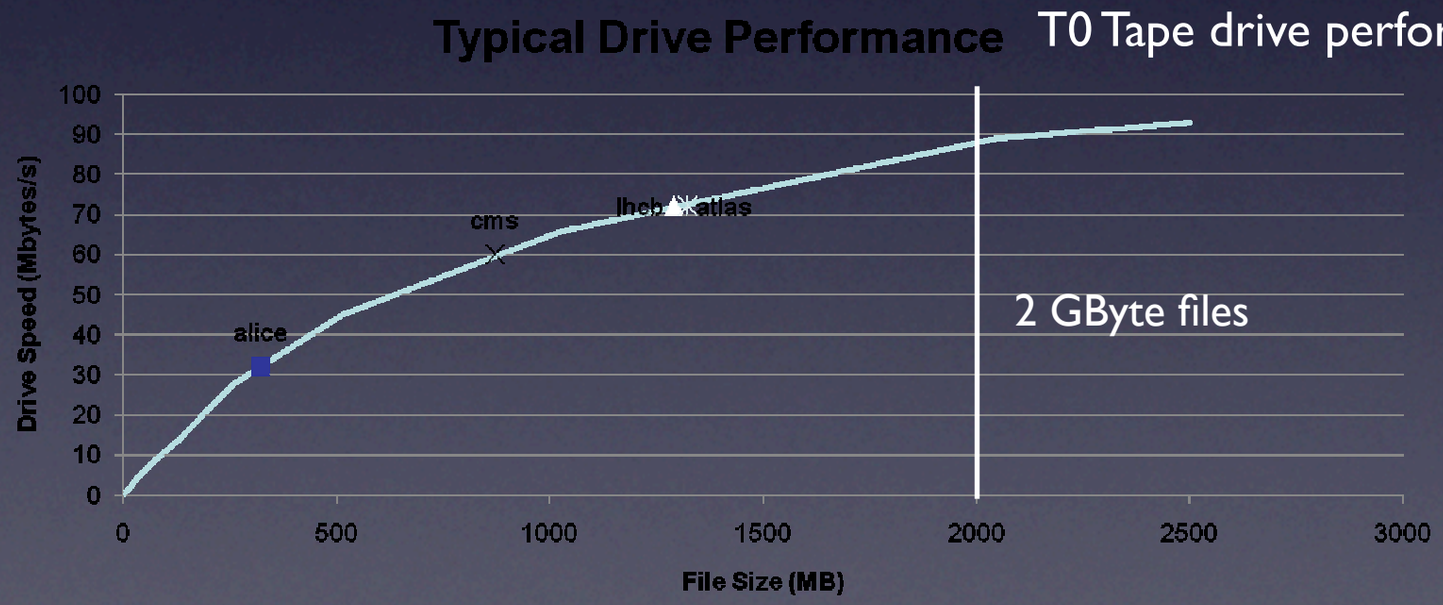

CMS Computing Model



- CMS Computing Model calls for a large number of Primary DataSets (originally 50) - determined by trigger information (muon, electron, jetMET...).
- The plan is to archive complete primary datasets at a single Tier-1.
- For 2008, two or more copies of RECO can be available at the Tier-1s.
- CAF is used for fast turn around.

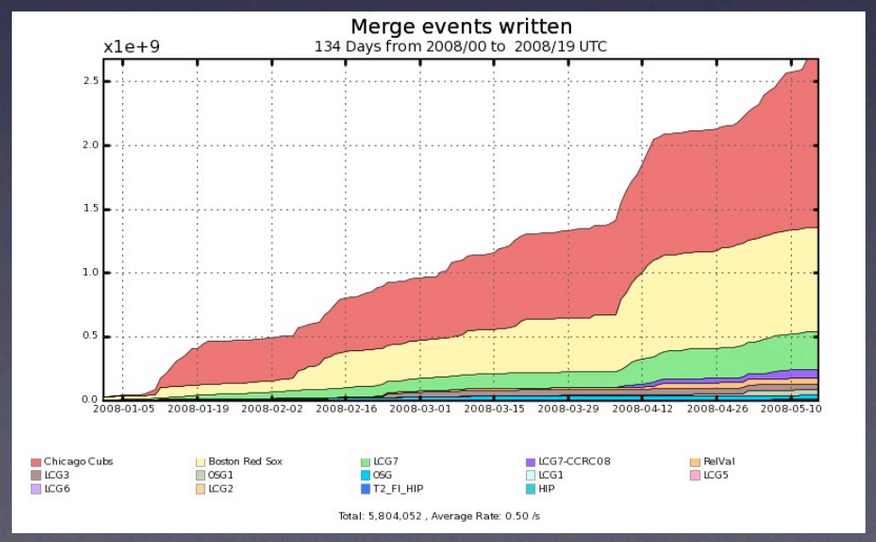

Resources at CERN

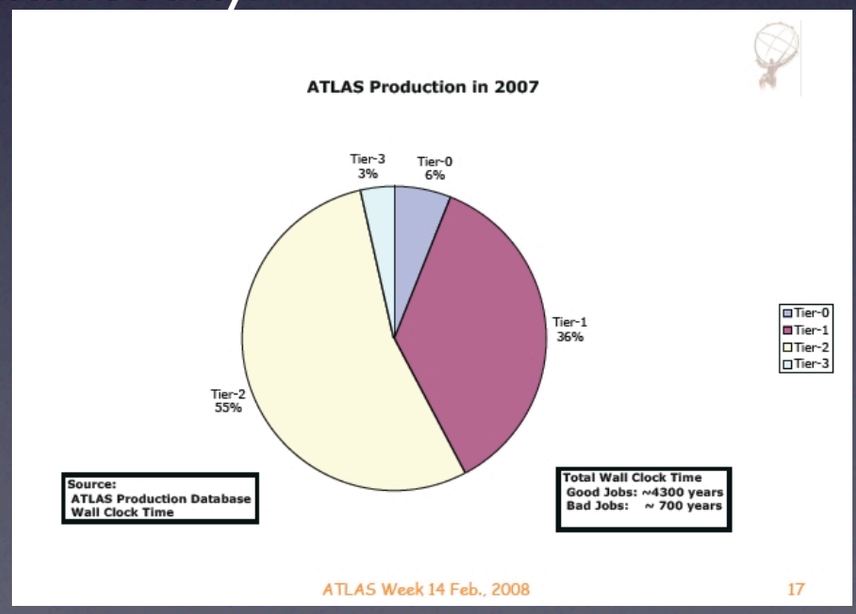
- Tier-0 - primarily for production use (reconstruction)
- CAF - for prompt calibration, alignment and priority activities
- **Users at CERN -** what will be available? There are CPU cycles available.
  - User disk space? - this should improve (?) WLCG working group is investigating.
  - Access to tape drives? - production activities and archiving RAW have priority!!!
    - Beware - small files degrade performance of tape systems (aim for ~2 GB files to tape)




Data Production

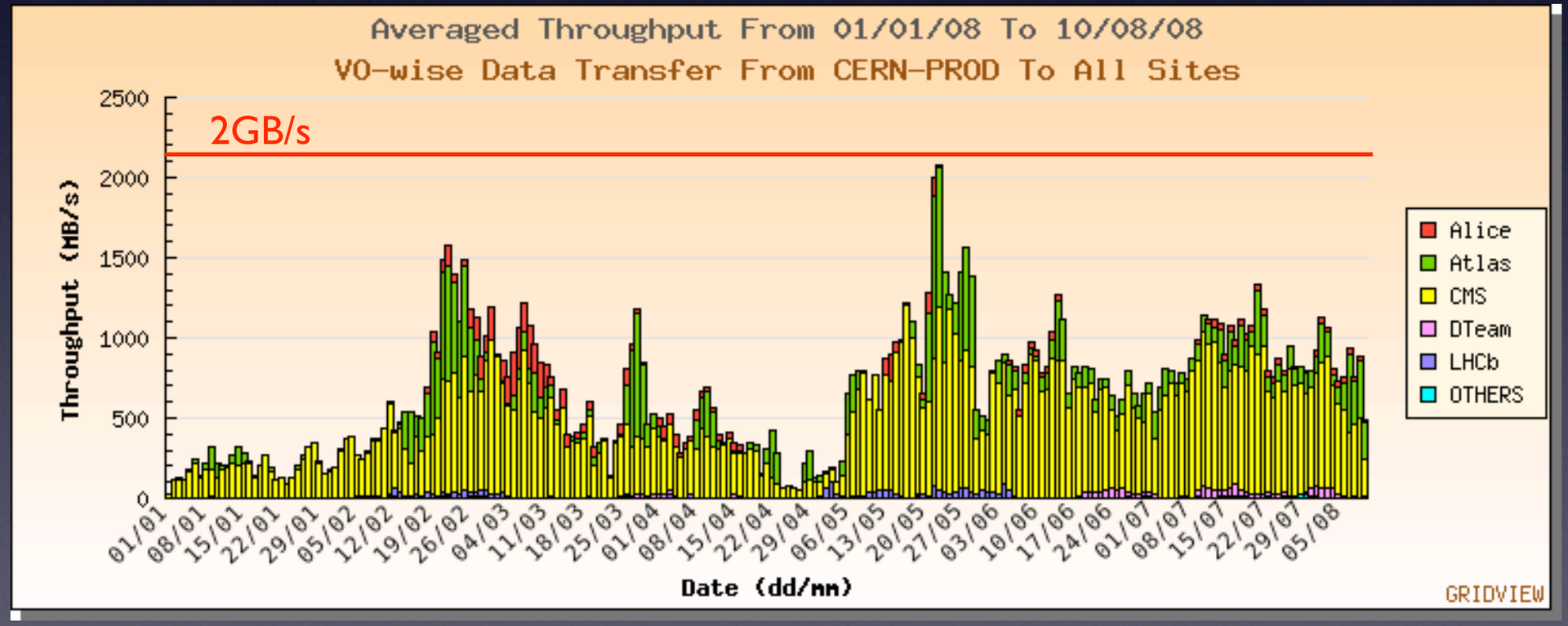
- MC Production has been on-going for some time. Operations teams at the experiments have tested the software and the sites (all Tiers) through MC production.
- Production with cosmic commissioning data has also been tested. Good check of software and the production and transfer systems. Lesson from challenges - if it hasn't been tested, it won't work.
- Workflows for calibration and alignment have been tested in recent challenges and dress rehearsals.
- Grid infrastructure has begun to stabilize. Sites have been upgrading continuously. Major downtimes impact the experiments and the users! Availability is what counts.
- Efficient access to data on tape will be key for re-reconstruction and for skimming. Still not fully tested at all sites by all experiments simultaneously.






Data Transfers

- Data is considered "safe" when it is archived to tape at the Tier-0 at CERN and at a Tier-1 site.
- Reliable data transfers from CERN to the Tier-1 sites are critical.
- On-going transfer tests and movement of production data have pushed the system to full scale.
- Transfers from CERN by VO since the beginning of 2008:




Organized Data Challenges

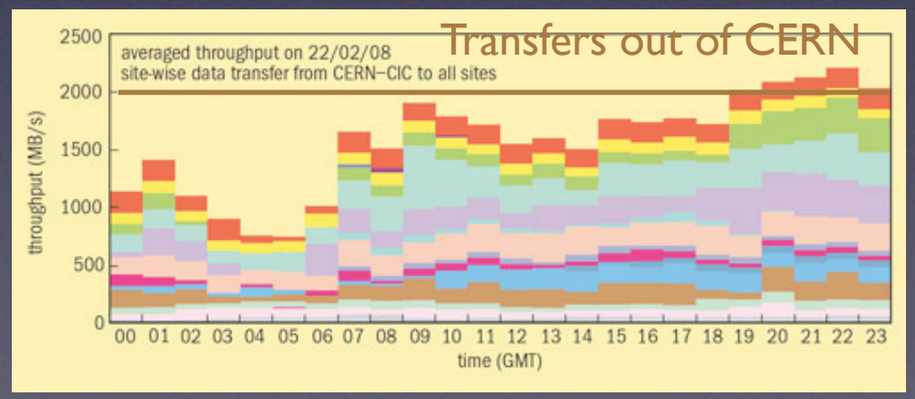
- Readiness and scale tests of infrastructure and sites have been important tools.
  - experiments have organized tests of their software and computing infrastructure
  - WLCG has organized tests of the grid tools and facility infrastructure at the site
  - Tests started many years ago. The GRID has evolved during that time. Realism has crept in... Some things are working much better, still some work to be done.

- Dress rehearsals for data production
  - major stress test of the computing systems and software
  - Important tests of alignment and calibration
  - Sometimes useful for physics studies - this has been a bonus

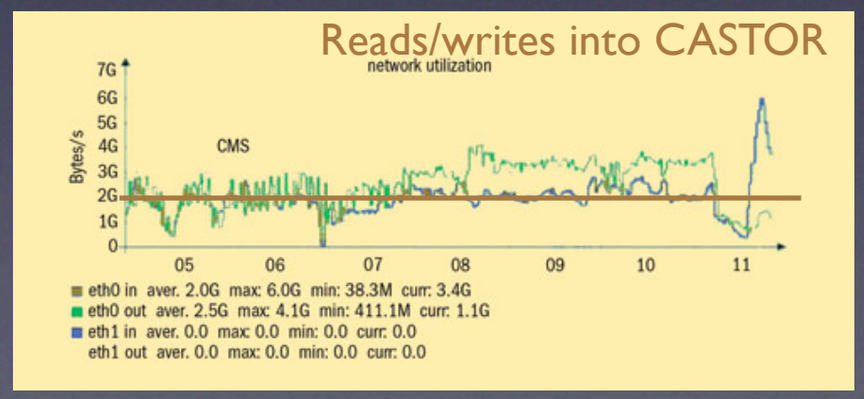
- An Analysis challenge at the Tier-2s... an elusive goal. Analysis has ramped up at Tier-2s in the last few months.

- In 2008 the WLCG orchestrated the CCRC tests: Combined Computing Readiness Challenge - with all 4 experiments contributing transfer/data access/CPU load at many sites.

Transfers out of CERN



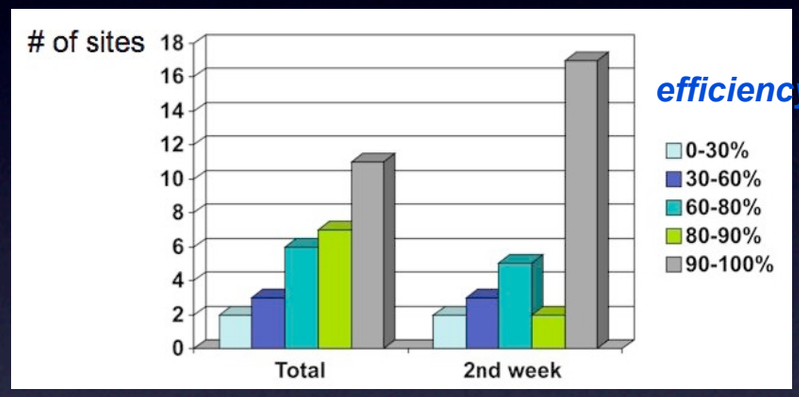
Reads/writes into CASTOR



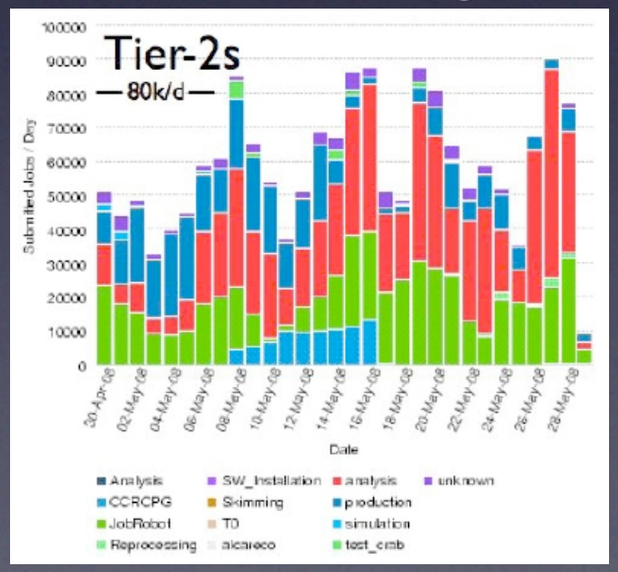

Role of the Tier-2s

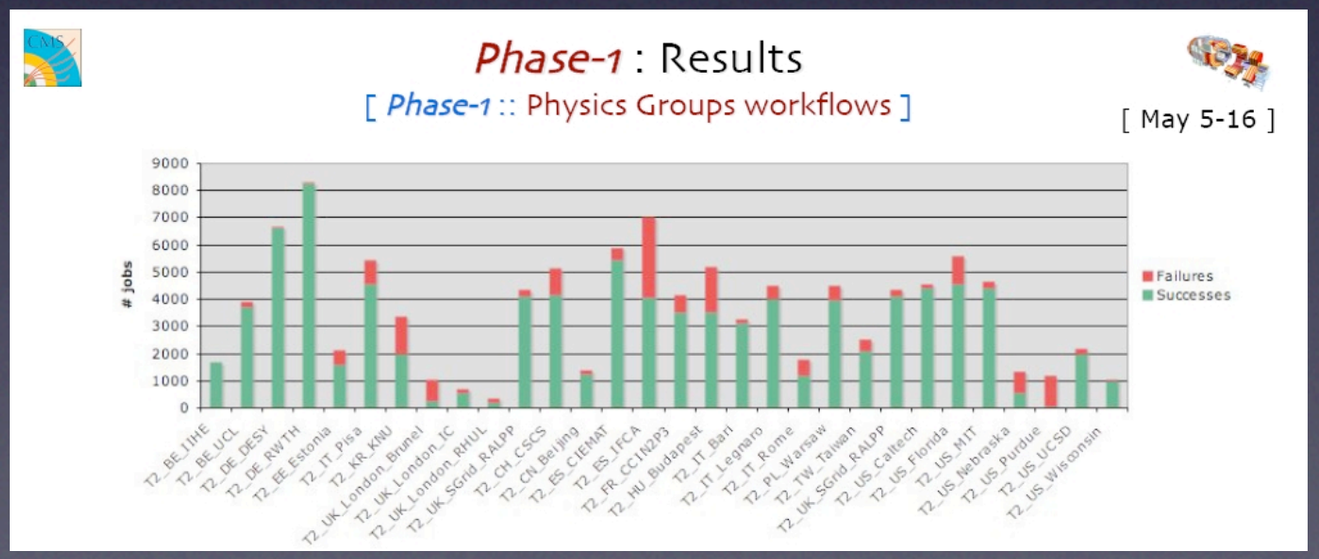
- **Monte Carlo production** - centrally organized; archived at regional Tier-1 - well tested by many computing and physics exercises

- ***User analysis and support for local community***
  - Users analysis at the Tier-2s is ramping up as more MC and commissioning data is available at the sites.
  - CMS tested analysis at the T2s during the May challenge.



- Tier-2s have agreements to provide support during local working hours. The expectation is that they will be up 24x7, but are not obliged to provide support on the weekends if something breaks.

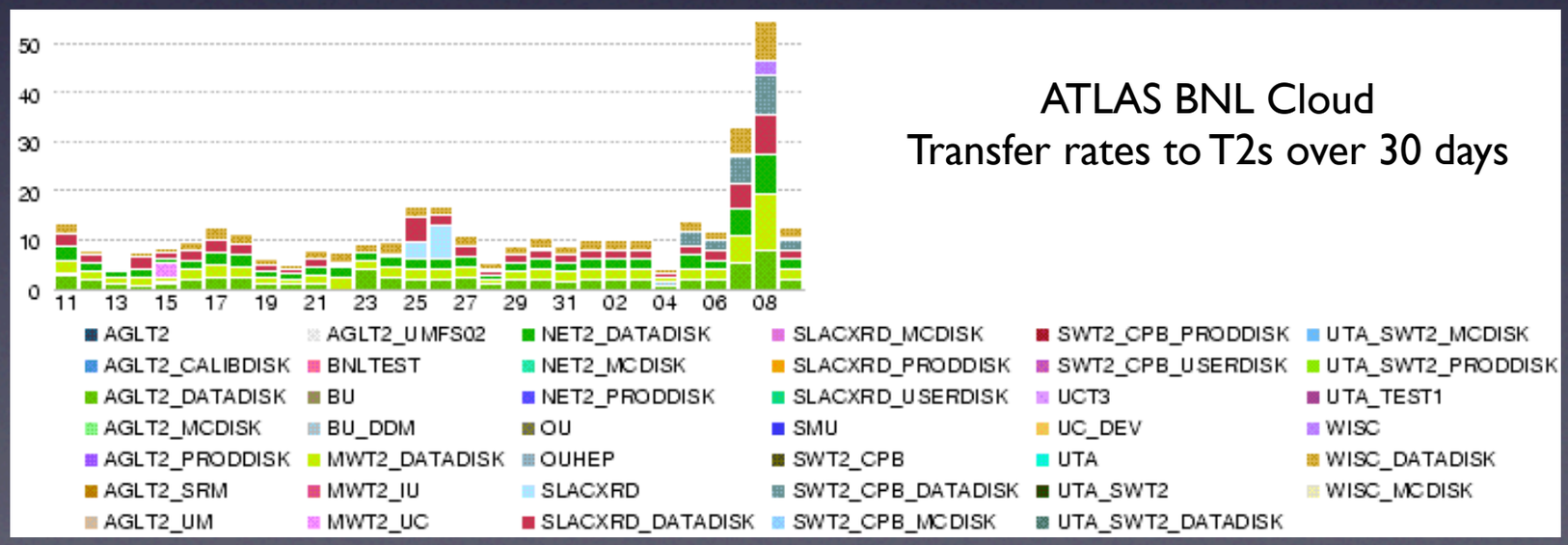




"controlled" (central) submission


Role of the Tier-2 centers

- ATLAS and CMS have slightly different plans for data distribution at the Tier-2s in their computing model.
  - For CMS the Tier-2s can pull data from any of the 7 Tier-1 sites; Tier-1 association is regional for production MC archiving.
  - ATLAS uses a cloud model with close association to the regional Tier-1 for dataset transfers.
  - Data distribution to the Tier-2 was tested in the recent Combined Computing Readiness Challenge (CCRC).



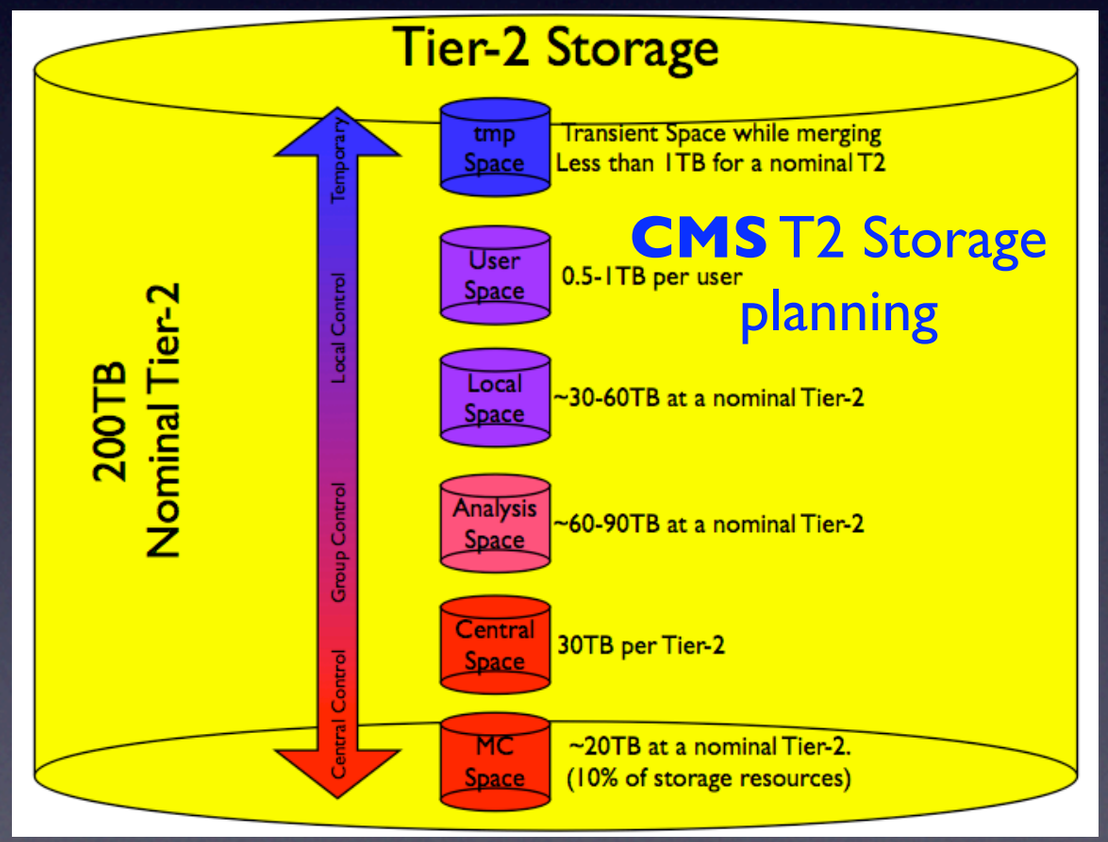
ATLAS BNL Cloud
Transfer rates to T2s over 30 days


Physics Analysis (CMS example)

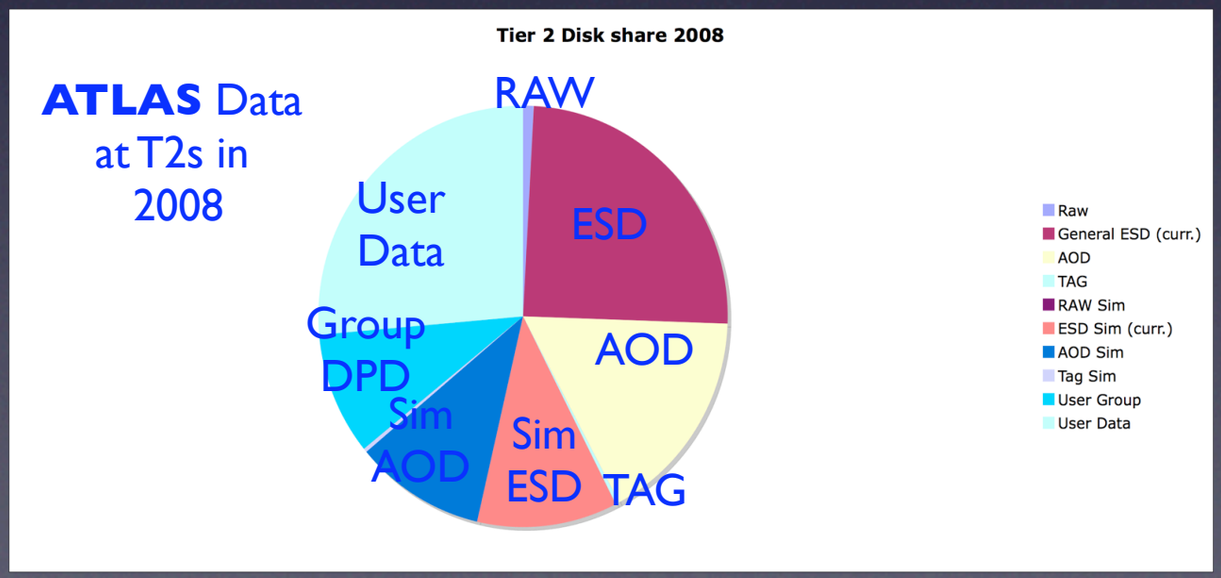
- In the computing model, CMS has two primary analysis resources: the CERN Analysis Facility (CAF) and a significant fraction of the combined Tier-2 resources.
- The CAF is intended for specific varieties of analysis with requirements of low latency access to the data.
  - The CAF is a very large resource in terms of absolute capacity, but is small wrt the total resources available for analysis in the global T2 system.
- The vast majority of the analysis capacity is located at the Tier-2 computing centers.
- The Tier-1 centers and CERN have provided much of the computing for analysis for CMS up to now.
- The Tier-2 sites storage were envisaged as dynamic caches whose contents are to be driven by the needs of users and physics groups.
- In reality, the storage at Tier-2 centers has been reasonably static and users have not felt able or empowered to control the storage.
- At least for 2008 - datasets will be placed at the Tier-2s.


Data Storage at the Tier-2 (2008 plans)

- Tier-2 centers must meet the needs of MC production and user analysis.
- This is where most of the analysis will be done. Need to use the storage wisely.
- Data storage at the Tier-2 will evolve. Current planning is focused on 2008.
- There will be storage space for user data at the Tier-2s. For CMS, expect space for 0.5-1TB per user with ~40 users/site.



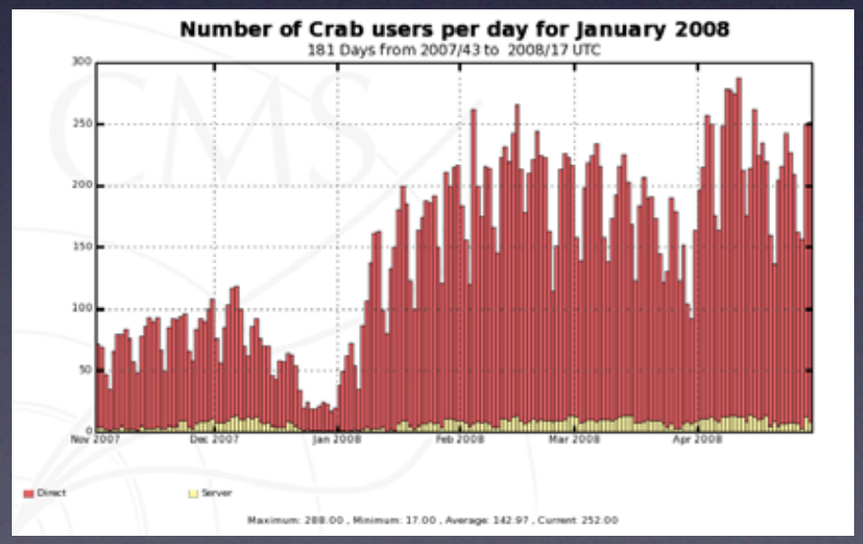
Some RAW and all ESD/RECO will be available at the T2 in 2008
Much of this will be "pre-placed" in early running.
AODs will be available as well as DPD(ATLAS) for physics groups.

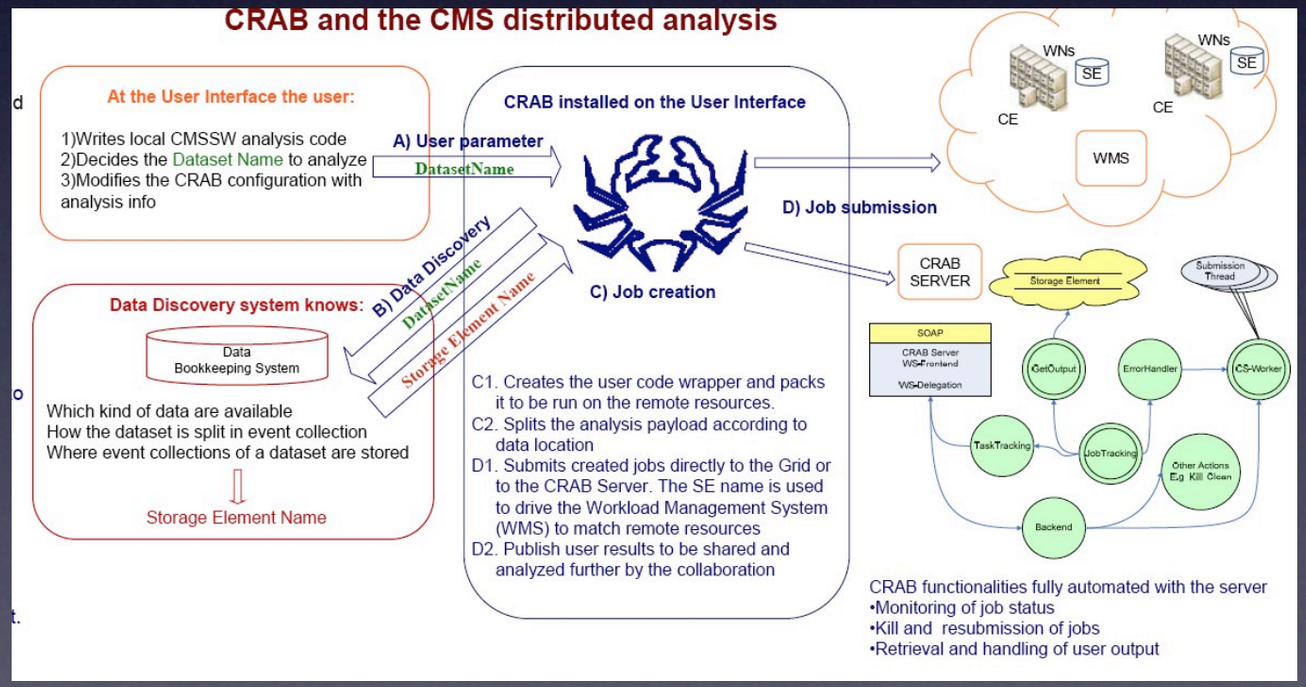



User Analysis on the GRID

- Concept:
  - Applications are developed by users on their laptop and sent out to the GRID
  - Data is placed at the GRID sites where official software is already available.
  - Anyone can request a data transfer to a site, but the site data manager must approve.
  - Site provides job output stage-out space
- The collaborations have developed tools to simplify the job of user analysis - locate the data, identify available resources, create jobs and submit, monitor the job and retrieve the output

In CMS this tool is CRAB:



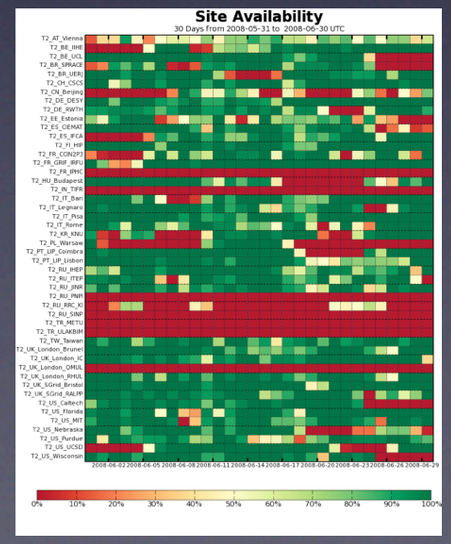


Tutorial for CRAB: https://twiki.cern.ch/twiki/bin/view/CMS/WorkBookRunningGrid

"The Grid distributed data analysis at CMS" - poster at CSFI08


Etiquette - User analysis at Tier-2s

- Users should access processing resources through the grid interface at Tier-2 centers.
- Users should expect the Tier-2 sites to pass availability tests and to respond to facility savannah tickets or GGUS tickets when facility grid interface problems are observed.
- Some sites provide direct login accounts for groups of users, but are not required to do so.
  - Sites providing direct access should ensure the CPU resources used by direct submission from interactive accounts are accounted to give proper credit to the site. Please use the tools that give credit to the site!
  - Sites should ensure that priority on the batch resources is defined by activity or association and not log in technique. A user performing the same activity should have the same priority whether he/she logged in to the cluster or submitted through the grid.
- Users must adhere to security policies. Sites will be strict!
- Think before you submit! What is your job doing?




Tier-3 computing

- Tier-3 centers do not have a well defined role in the experiments.
- Local computing at institutes for users. Resources are not scheduled by the collaborations. (T3s are not required to participate in MC production, for example)
  - Many T1/T2 have set aside resources for local T3 use
  - large/small clusters at universities
  - desktops/laptops
- Tier-3 access to the data - what are the limitations? Not a priority for production.
- An example - ATLAS "policy" for T3 data transfers
  - O(10GB/day/user) who cares
  - O(50GB/day/user) rate throttled
  - O(10TB/day/user) user throttled
  - everything is negotiable - large transfers are possible if negotiated


Computing Models for 2008

- Expect shorter run (~60 days) with lower operations efficiency (~40%?).
- Plan for calibration and alignment runs with cosmics.
- Expect high trigger rates and possibly larger event size.
- Assume more re-reconstruction passes.
- Physicists will want to access more information than is in the AOD.
- In short, the computing models will be different than would expected for steady operations - for example, in CMS the RECO will be available at the Tier-2s for 2008.


CMS Data distribution for startup (2008)

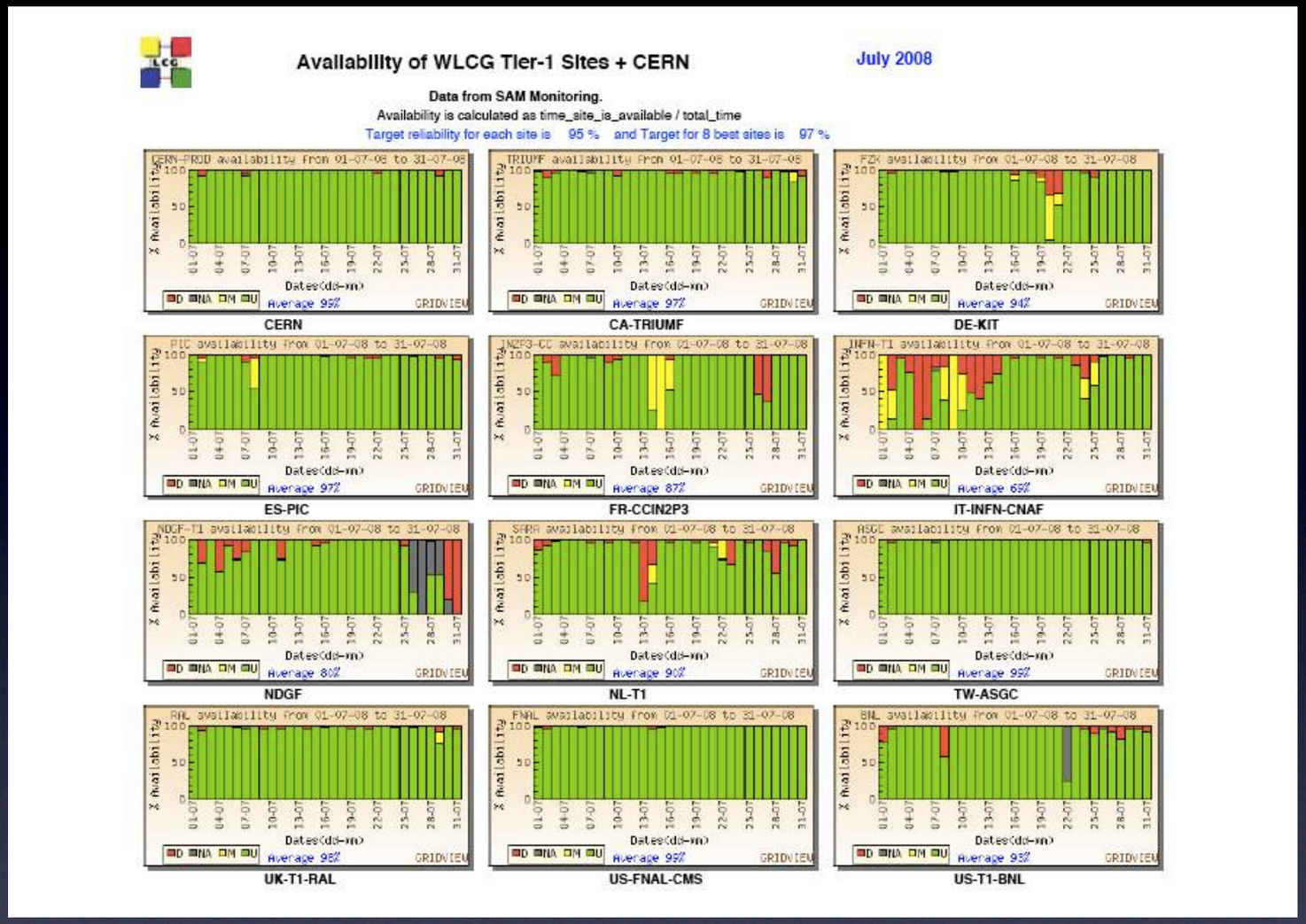
- T0 --> CAF
  - Initially, a full RECO will be copied to CAF at CERN for calibration, alignment and early physics studies
  - limitation: CAF disk space; one copy only, to be replaced whenever re-reconstruction takes place

- T0 --> T1
  - expect ~3 passes of reconstruction passes in 2008
    - simRECO and simAOD: at (one) custodial T1 only
    - realRECO and realAOD: as many copies as possible with at least one copy at external T1

- T1--> CAF and other T1s
  - in order to have full set of RECO at CAF, transfer re-reconstructed data to CERN
  - need to flush older versions quickly from CAF disk

- T1--> T2
  - use a portion of T2 storage to host RECO datasets by Primary Dataset
  - need at least to have ~2 copies among all T2s
  - matched to general interests of local groups
  - expect to assign 2-3 T2 centers per analysis group
  - effect on export bandwidth increases linearly

Monitoring the GRID

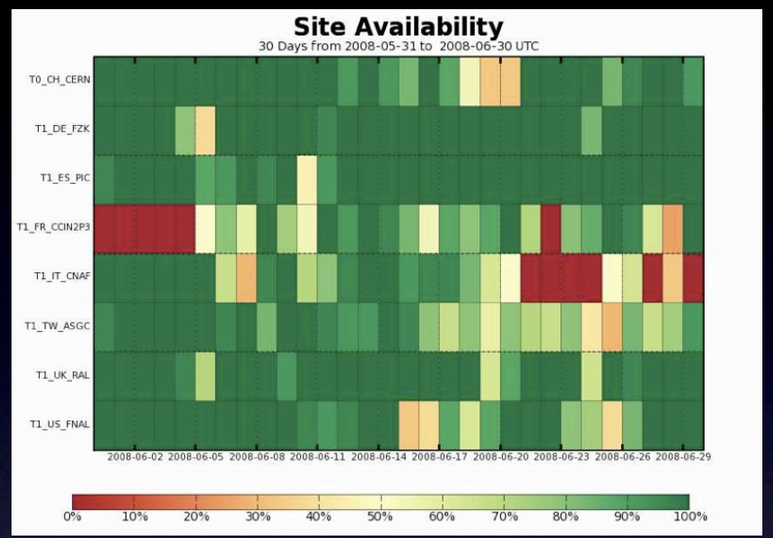


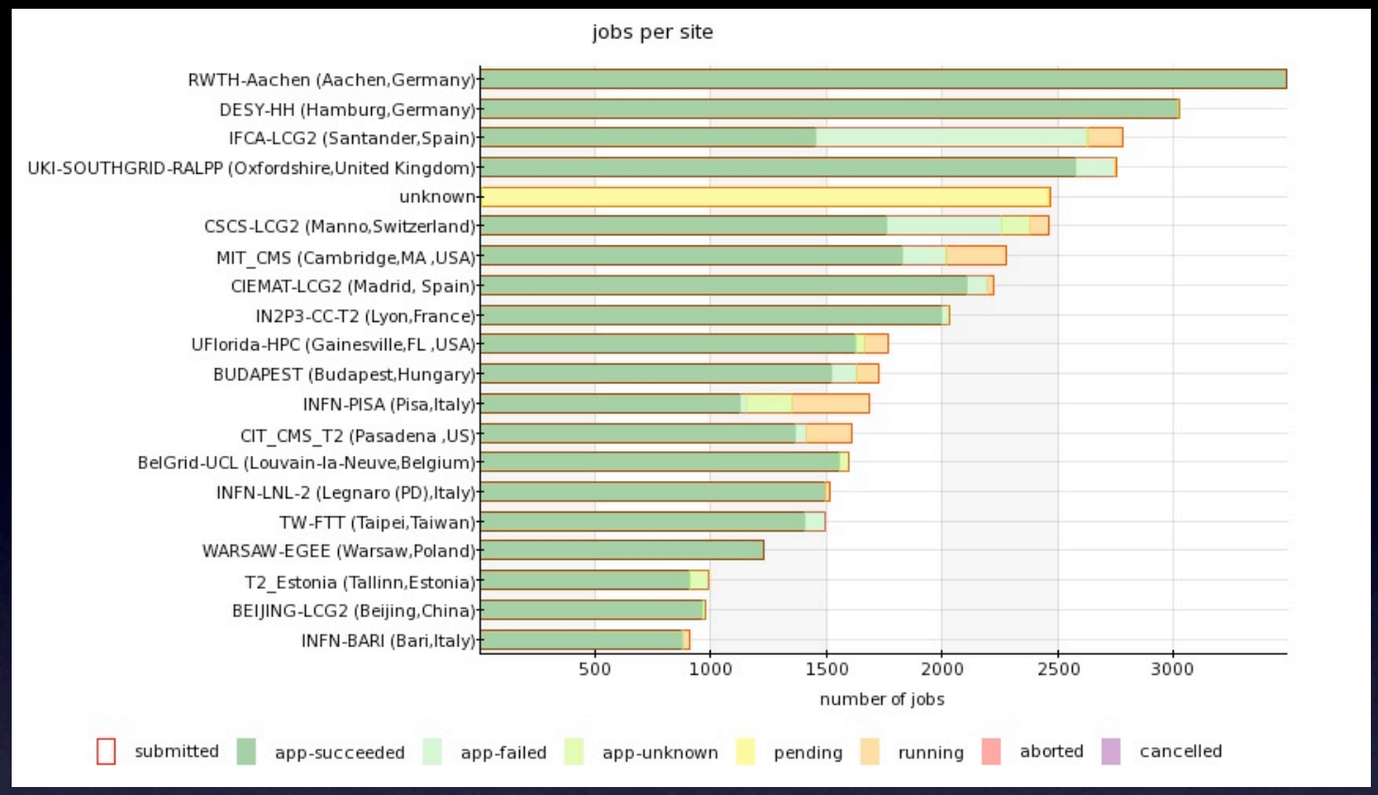
Tier-1 availability > 95%

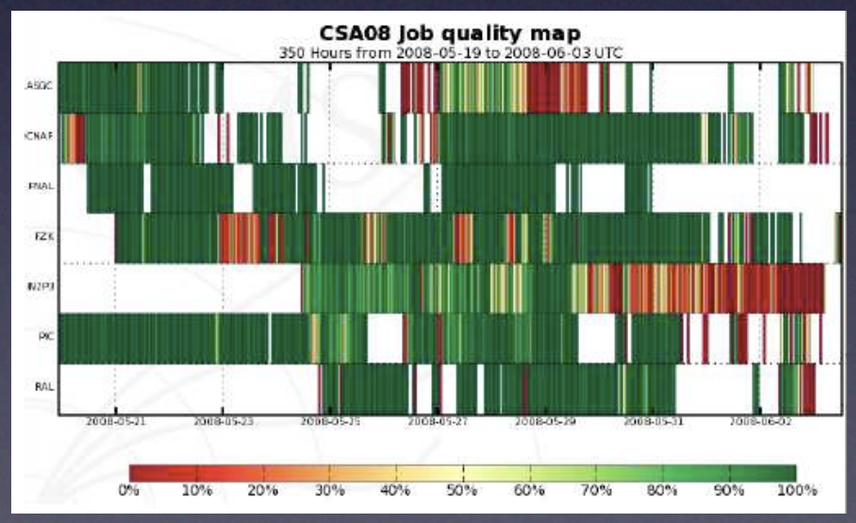
Availability = time site is available / total time

- Monitoring of GRID systems is essential for stability.
- Tests of availability "SAM" have been developed for WLCG services and for the experiments to answer "will my job run at this site".
- Availability is published every month and watched carefully by the experiments.
- The situation has improved over the past year, but there is still room for improvement.
- Scheduled downtimes impact availability and will cause difficulties if data is unavailable for analysis for long periods.

more on Monitoring







Monitoring computing in the experiments: http://arda-dashboard.cern.ch/

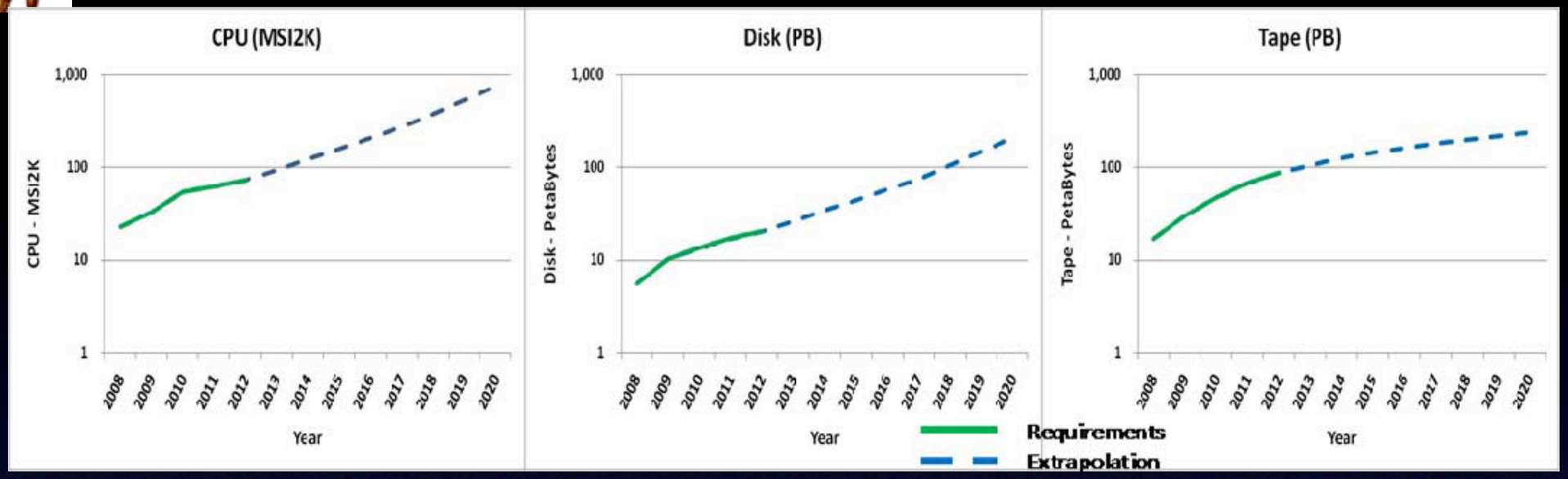
Ability to monitor jobs, transfers and sites.

Constantly improving. Need more "RT" monitoring - what is my job doing now?


Computer Security

- Certificates are issued by a CA (Certificate Authority) - for example by DOE or CERN who certifies the identity of the owner of the certificate.
- If CAs are trusted; then user certificates verified by the CA are trusted.
- Users (or their grid certificates) are registered with and become part of a Virtual Organization through a VOMS. (For an experiment VO, someone must verify that a user is a member of the collaboration.)
- Attributes are assigned to the user at the VOMS (Virtual Organization Management Service) to control the actions a VO member can perform. Default role - user can submit a grid job.
- Policy enforcement is local at the sites; privilege enforcement is decided by resource owners. (Experiments play a role.)
- ***Cyber security is taken very seriously by all sites and by all VOs. Please do your part to keep the systems secure.***

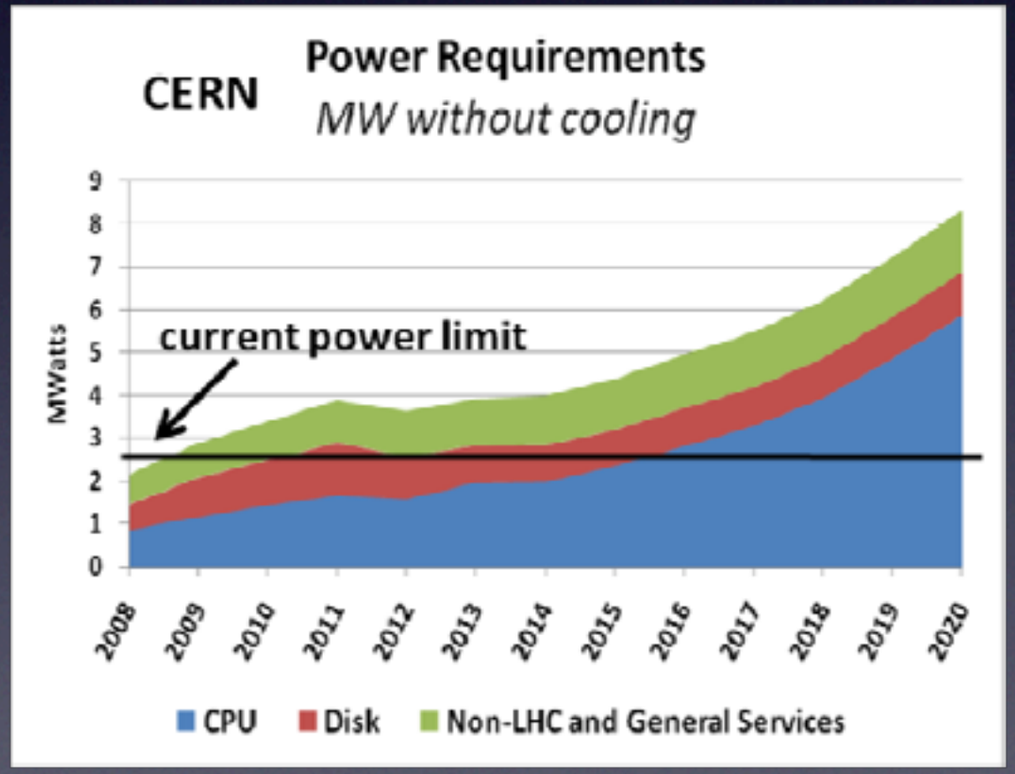
CERN computing: Looking towards the future



planned capacity at CERN

- The collaborations are projecting computing needs to 2012, but will need to revisit the requirements after some experience with data from the detectors.

- Power and cooling will be a limiting factor. Plans for new facility infrastructure at CERN are underway. Other centers will have similar issues as CPU capacity increases.



http://lcg.web.cern.ch/LCG/documents/Plan%20to%20meet%20LHC%20Experiment%20Requirements%20at%20CERN%20-%20Summary.pdf


Summary on Data Access

- Data management will be key to successful LHC physics analysis. Jobs will run where the data are found.
- Central management is expected at the Tier-0 and Tier-1. We have had to delete MC data this year since we have succeeded in using all the tape space at several centers. ($100/TB for tape media)
- Data management at the Tier-2s will be key for user analysis! Please use the resources wisely.
  - Understand what resources are available AND what you need.
  - CPU and disk are both valuable resources and eventually will be in short supply. Reflect on what you are doing. Am I using the resources effectively? Prepare for change.
- It has been our experience that it is difficult for physicists to delete data - even MC from years ago. This will be an on-going issue for the experiments.


Summary

- The computing models for the experiments were conceived about 10 years ago and refined in 2005. Now it is time to see how they work!
- I have skipped a lot of details and passed over a lot of systems. My apologies if I did not cover your work or your experiment in detail. Please go to tutorials in your collaboration for details.
- Expect these models to evolve over the next year or two. Adaptations to optimize use of resources and to take advantage of technology changes are expected.
- Informed users will help make the global computing systems work better.
  - Please inform yourselves about the software and the capability of the sites. Informed users will be happier users.
  - Please volunteer to teach others.
  - The global computing system for the LHC took many years and many people to realize. It is a work in progress. Suggestions and contributions are always welcome.