The Grid Challenges in LHC Service Deployment

Patricia Méndez Lorenzo

CERN (IT-GD)

Linköping 18th October 2005

6th Annual Workshop on Linux Cluster for Super Computing & 2nd User Meeting for Swedish National Infrastructure for Computing

Information Society

Outlook

CERN and LHC

LCG: LHC Computing Model

Service Challenges: Phases

Current Phase: SC3

Support to the Services

Support to the Experiments

Experiences of the Experiments

Next Step: SC4

Summary

The European Organization for Nuclear Research

The European Laboratory for Particle Physics

Fundamental research in particle physics

Designs, builds & operates large accelerators

Financed by 20 European countries
ıber states) + others (US, Canada,
a, India, etc)

2000 staffs + 6000 users from all over the world

Next huge challenge: LHC (starts in 2007) experiment: 2000 physicists, 150 universities, with an operation life greater than 10 years

LHC- Physics Goals

Higgs particle

Key particle in the Standard Model that could explain the elementary particle masses

Search for super-symmetric particles and possible extra dimensions

Their discovery would be a serious push for Super Symmetric theories or "String Theories" aiming at the unification of the fundamental forces in the Universe

Anti-matter issues

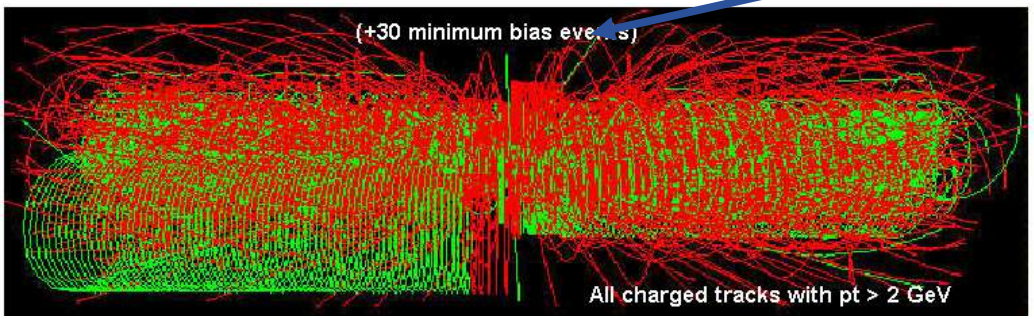
Why the Universe is made of matter instead of an equal quantity of matter and antimatter

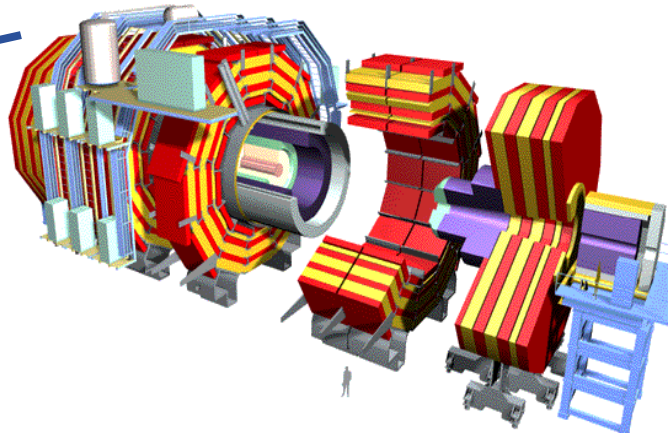
Understand the early Universe (10^{-20} – 10^{-6} seconds)

Soup of quarks and gluons stabilized into nucleons and then nuclei and atoms

The LHC Experiment

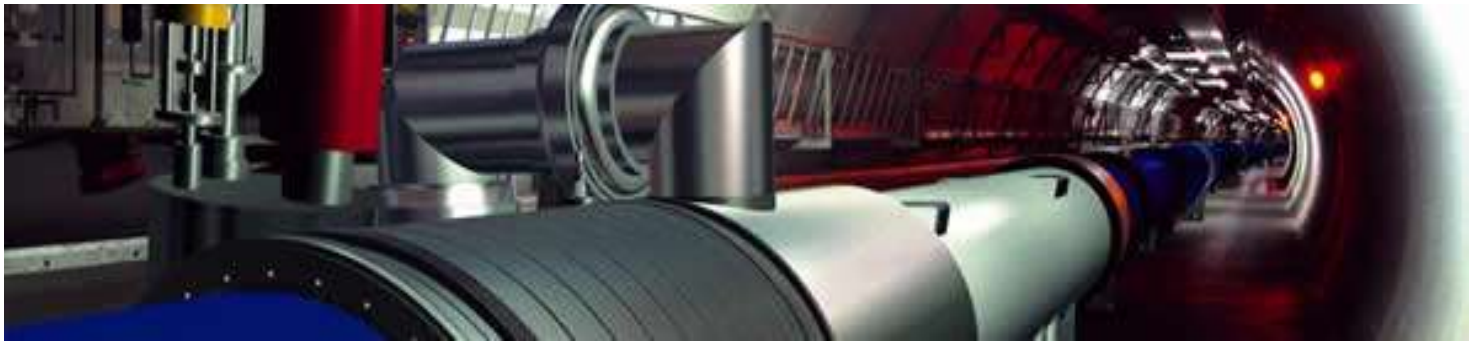
Reduce by online computers that filter out a few hundred good events per sec





The LHC: G... ...vents) per second

Recorded on disk and magnetic tape at 100-1000 MB/sec: 15 PB/year



The LHC Computing Environment

LCG (LHC computing Grid) has been developed to build and maintain a storage and analysis infrastructure for the entire high-energy physics community

LHC is beginning the data taking in summer 2007

- Enormous volume of data
  - Few PB/year at the beginning of the machine operation
  - Several hundred PB yearly produced for all experiments in 2012
- Large amount of processing power

As a solution a LCG world-wide Grid is proposed

- Established using a world-wide distributed federal Grid
- Many components, services, software, etc, to coordinate

Takes place at an unprecedented scale

- Many institutes, experiments and people working closely together

LCG must be ready at full production capacity, functionality and reliability in less than 1 year!

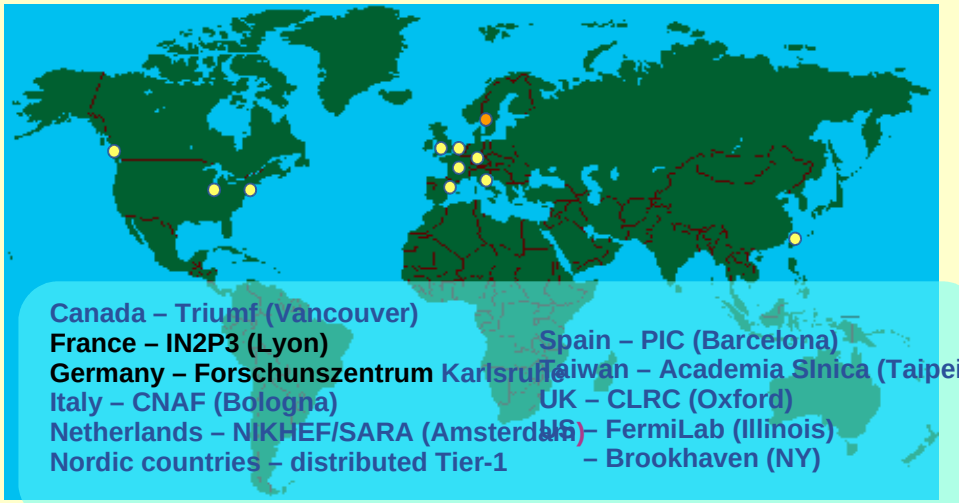
LCG is an essential part of the chain allowing the physicsts to perform their analyses

- It has to be a stable, reliable and easy to use service

LCG: The LHC Computing Grid

Tier-0 – the accelerator centre

- Data acquisition and initial Processing of raw data
- Distribution of data to the different Tier's



Canada – Triumf (Vancouver)
France – IN2P3 (Lyon)
Germany – Forschunszentrum Karlsruhe
Italy – CNAF (Bologna)
Netherlands – NIKHEF/SARA (Amsterdam)
Nordic countries – distributed Tier-1
Spain – PIC (Barcelona)
Taiwan – Academia Sinica (Taipei)
UK – CLRC (Oxford)
US – FermiLab (Illinois)
– Brookhaven (NY)

Tier-1 – "online" to the data acquisition process → high availability

- Managed Mass Storage – → grid-enabled data service
- Data-heavy analysis
- National, regional support

Tier-2 – ~100 centres in ~40 countries

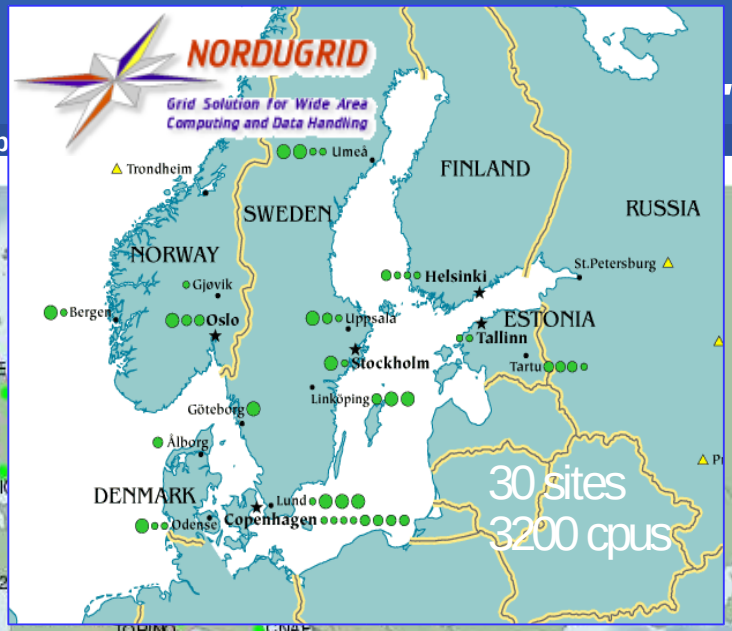
- Simulation
- End-user analysis – batch and interactive

Tier 1 centers

				ALICE	ATLAS	CMS	LHCb
1	GridKa	Karlsruhe	Germany	X	X	X	X
2	CCIN2P3	Lyon	France	X	X	X	X
3	CNAF	Bologna	Italy	X	X	X	X
4	NIKHEF/SARA	Amsterdam	Netherlands	X	X		X
5	NDGF	Distributed	Dk, No, Fi, Se	X	X		
6	PIC	Barcelona	Spain		X	X	X
7	RAL	Didcot	UK	X	X	X	X
8	TRIUMF	Vancouver	Canada		X		
9	BNL	Brookhaven	US		X		
10	FNAL	Batavia	US			X	
11	ASCC	Taipei	Taiwan		X	X	

A US Tier1 for ALICE is also expected

LCG in the World



May 2005

140 Grid sites

34 countries

12000 CPUs

8 PetaBytes

Current Activities

Experiments and Grid teams are ramping up

Service Challenges (SC): Brief introduction

Testing the required services to the necessary level of functionality, reliability and scale

Preparing, hardening and delivering the production of LCG environment

Running in an environment as realistic as possible

Data Challenges (DC): General aspects

Experiments test their LCG based production chains and the performance of the Grid fabric

- ➳ Processing data from simulated events
- ➳ Emulating step by step the scale they will have to face during real data taking

From 2005 experiments include SC testing as part of their DC

The Service Challenges

The Main Purposes

➳ Understand what it means to operate a real grid service – run for days/weeks outside the Data Challenges of the experiments

➳ Trigger/encourage the Tier1 & Tier2 planning – move towards real resource planning

➳ Get the essential grid services ramped up to target levels of reliability, availability, scalability, end-to-end performance

➳ Set out milestones needed to achieve goals during the service challenges

This is focused on Tier 0 – Tier1 and large Tier 2

Service Challenge Phases

SC1 (December 2004) and SC2 (March 2005)

- Focused on the basic infrastructure
- Neither experiments nor Tier-2 were involved
- Extremely useful feedback to build up the services and to understand the issues involved in offering stable services

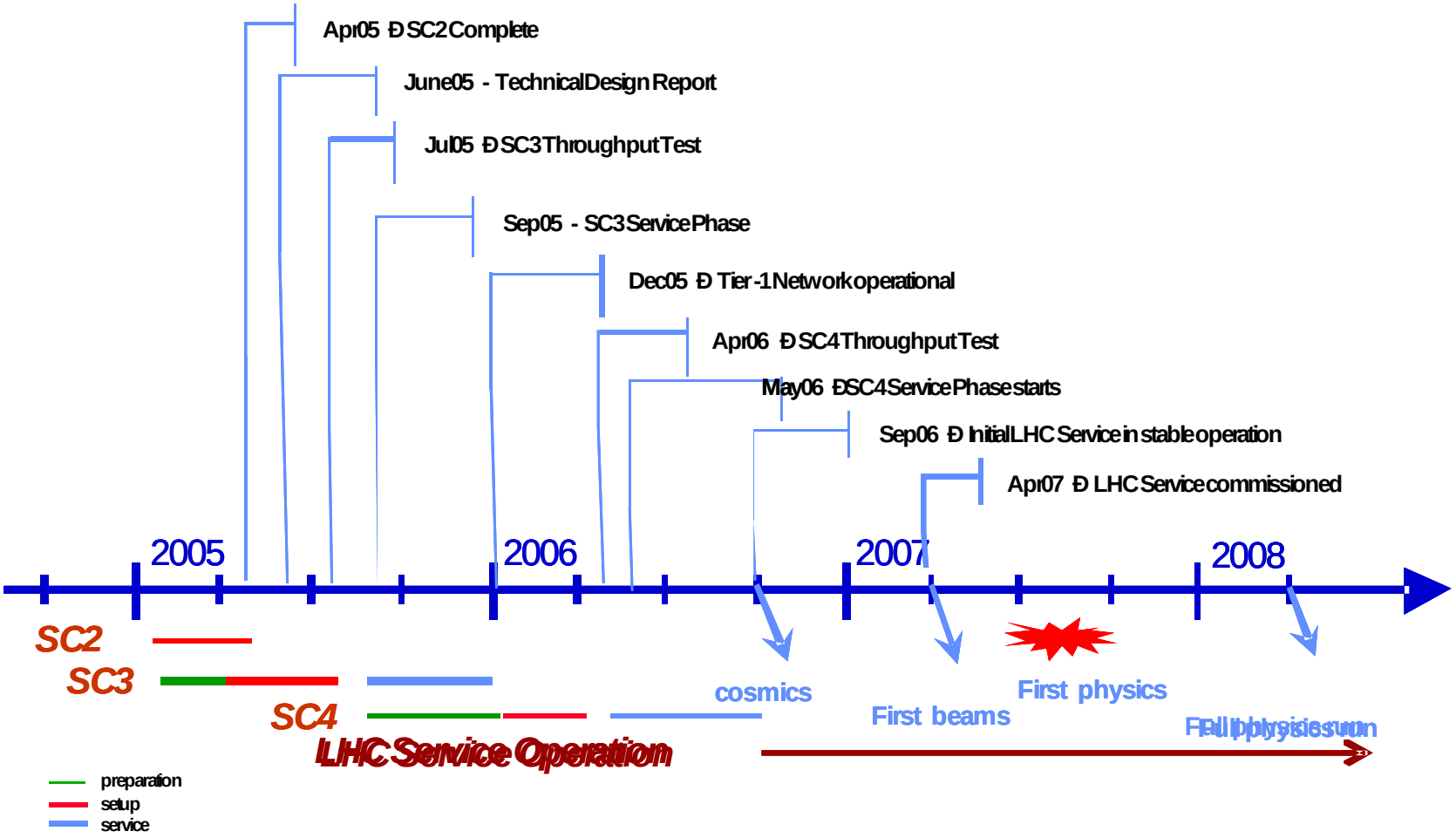
Until end 2005: SC3

- Include increasing levels of complexity and functionality
- All experiments and all Tier1s are now involved!
- Including experiment-specific solutions

Beyond 2005: SC4

- Designed to demonstrate that all of the offline use cases of LHC can be handled at an appropriate scale
- The service that results becomes the production service of LHC

LCG Deployment Schedule



The Current Challenge: SC3

Principal Challenge of this Phase: Prototype the data movement service needed by LHC

Generalities (1): In terms of Architecture

- The elements used in this service already existing but never interconnected
- Using the same architecture and Design than in previous challenges

The Current Challenge: SC3

Generalities (2): In terms of Setup and Service

- **Setup phase**
  - Starting 1st of July 2005
    - Includes throughput test
      - Maintenance for one week an average throughput of 400 MB/s from disk (CERN) to tape (Tier-1)
  - 150 MB/s disk (CERN) -> disk (Tier-1)
  - 60 MB/s disk (CERN) -> tape (Tier-1)

- **Service Phase**
  - Stating September until the end of 2005
  - Including real experiment data
    - CMS and Alice at the beginning
    - ATLAS and LHCb joining in October/November

Description of the Services provided in SC3

SRM (Storage Resource Manager v1.1)

- General interface to manage the different storage systems placed in the sites, in an homogeneous way
- V2.1 foreseen for SC4

FTS (File Transfer Service provided by gLite)

- It is a lowest level data movement service
- Responsible for moving sets of files from one site to another one in a reliable way and that's all
- Designed for point to point movement of physical files
- New versions (to be deployed end October) dealing with Logical File Names and Catalogs

LFC (LCG File Catalog)

- grid catalog capable to provide local and global cataloguing functionality including file replication across different sites.

Description of the Services provided by SC3

Summary of the services requested by each LHC Experiment during the SC3 phase

Service	VOs
SRM (1.1 in SC3 and 2.1 in SC4)	all VOs
LFC	ALICE, ATLAS
FTS	ALICE, ATLAS, LHCb
CE	Not required
Site BDII	Not required
R-GMA	Not required

Results of SC3 in terms of Transfers

The target data rates 50% higher than during SC2

All T1 have participated in this challenge

Important step to gain experience with the services before SC4

Site	Daily average (MB/s)
ASCC, Taiwan	10
BNL, US	107
FNAL, US	185
GridKa, Germany	42
CC-IN2P3, France	40
CNAF, Italy	50
NDGF, Nordic	129
PIC, Spain	54
RAL, UK	52
SARA/NIKHEF, NL	111
TRIUMF, Canada	34

Best daily throughputs sustained per site

Support for the Services

During the throughput phase the experiments had the possibility to test the services at CERN

- To get confident with their software
- To gain familiarity with the services

NOT intended to perform production

For the FTS

- Set of pilot servers set at CERN per each experiment
- Access restricted to a small set of persons
- Set of channels defined between T0-T1s
- FTS UI setup at CERN for the tests

For the LFC

- Each experiment had a pilot LFC server placed at CERN

VO BOXES

The experiments require a mechanism to run long-live agents per site

They perform activities on behalf of the experiment

- ➳ Job submissions, monitoring, software installation, etc

Provided for all the experiments which require it in SC3 participant sites

The Base Line Services Workgroup identified the need for the experiments to run specific services on the sites

Service Provided to the experiment

Dedicated nodes to run such agents

Delivered at each site for each experiment

- ➳ Direct gsissh loggin is allowed for certain set of persons per exp.
- ➳ Registration of a proxy for automatic renewal
- ➳ Delivered for the experiment software installation

Experiences with the Experiments

Main Goals of the experiments using the SC3

- Test of the data transfer and storage services to work for realistic use
- Evaluation of the services provided
- Data Challenges productions using SC3 services
  - ➳ Test of the experiment models
  - ➳ Analysis of reconstructed data
- Testing the Workload Management components

Experiences with experiments:ALICE

Alice is coordinating with SC3 to run their DC05 in the SC3 framework
Already beginning in September

Requirements and Status:

VO-managed node at each site

- ➳ Available in all sites: T1 and T2 participating in SC3
- ➳ Alice performs the job submissions from this node

LFC required as local catalog in all sites

- ➳ Central Catalog (Metadata Catalog) provided by Alice
- ➳ LFC APIs implemented in the Alice Framework (AliEn)
- ➳ More than 10000 entries (LFC as unique catalog)

FTS available in all sites

- ➳ Tests among T0-T1 performed this summer
- ➳ FTS APIS implemented in the Alice Framework

➳ First bunches of jobs already submitted at CERN

Full Production Mode

Experiences with Experiments: ATLAS

ATLAS is joining the Production in October

Requirements and Plans:

VO box deployed at all sites

- ➳ Required at all sites

Deployment of FTS (Throughput phase)

- ➳ Comparison of FTS performance with the ATLAS framework (Don Quijote)
- ➳ Following the throughput phase and investigating the integration of FTS in their space

For the moment RLS (previous catalog) entries are being migrated to global LFC (Throughput phase)

- ➳ Using their pilot as a global copy of the RLS to run some analysis
- ➳ Cleaning the entries in the catalog before splitting it in many local catalogs

The 2nd phase (from October) foresees tests to their new Data Management infrastructure

Experiences with the experiments: CMS

CMS is in production from middle September

Requirements and Status:

Vo BOX not required

Transfers: Done with PhEDEx

- FTS is not currently used
- Integration with PhEDEx likely to happen later
- Full production already going on

Catalog: Local file catalog is needed by PhEDEx

- Testing the different POOL backend
- Functionality and scalability tests

Experiences with the Experiments: LHCb

LHCb is already in production from October

Requirements and Status:

VO BOX not required right now

LFC central catalog

- Implementation of the different catalogs in the general Management System
- Catalogs centralized at CERN
- Very good experiences with LFC

Data Transfer: FTS

- Integration of the FTS client into the LHCb Data Management system already done
- LHCb is in full production mode

Service support activities

Service challenge Wiki Page

- Contact persons
- Experiment Requirements
- Configuration of the Services
- T0, T1 and T2 information
- Daily log files and tracking of the status

Multi-level support structure

- First Level Support
  - Box-level monitoring and interventions
- Second Level Support
  - Basic Knowledge of the applications
- Third Level Support
  - Experts

Daily contacts with experiments

One person full support for each experiment

Regular Meeting with the experiments and the sites

Regular visits and workshops at each site

Next Step: SC4

The aim of SC4 is to demonstrate that the data processing requirements of the experiments can be fully handled by Grid at the full nominal data rate of LHC

All T1 and most of T2 involved

Triple "challenge":

Services: Successfully completed 6 months before the data taking

They have to ready to a high stability level

Sites: Ramp up their capacity to twice the nominal data rates expected for the production phase

Experiments: Their own services will be included in the next phase → their feedback is fundamental

The involvement of the experiment should enable a sizeable fraction of users to migrate to the grid and use it for their daily activity

The service resulting from SC4 becomes the production model

The full communities of LHC will be now supported

Target data rates for CERN and T1 centres in SC4

Center	ALICE	ATLAS	CMS	LHCb	Target Data Rate MB/s
TRIUMF		X			50
CC-IN2P3	X	X	X	X	200
GridKa	X	X	X	X	200
CNAF	X	X	X	X	200
NIKHEF/SARA	X	X		X	150
NDGF	X	X	X		50
PIC		X	X	X	100
ASGC		X	X		100
UK,RAL	X	X	X	X	150
US,BNL		X			200
US,FNAL			X		200
At CERN					1600

Summary

LHC, the world largest collider will begin to take data in 2007

- Computing infrastructure challenges
- Data Management Challenge

The LHC Computing Model LCG has been developed to build and maintain a storage and analysis infrastructure for the entire high-energy physics community

- More than 140 sites distributed in 34 countries in a Tier-Model structure

The Service Challenges purpose is to understand what it means to operate a real grid service and to deal in advance all the situations the experiment will have to face during the real data taking

- It is not an exercise, it is the real thing
- The current phase (SC3) is being used and tested by the HEP experiments
- Next step: SC4 will include all Grid communities and its results will be deployed for the real data taking