



ATLAS Grid Computing - Introduction

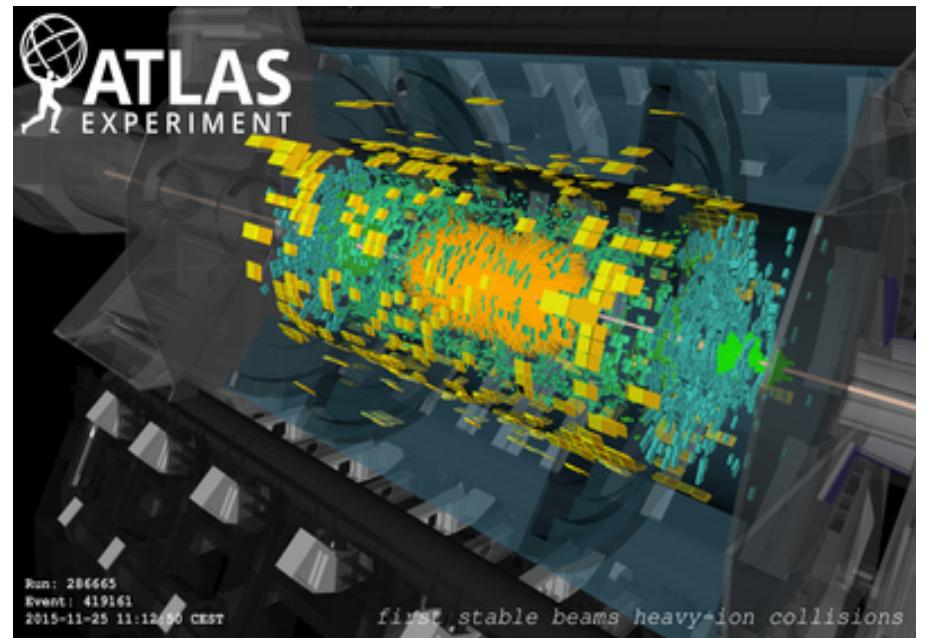
Practical Course on Parallel Computing – Sose18, 04.2018

Gen Kawamura, II. Physics Institute

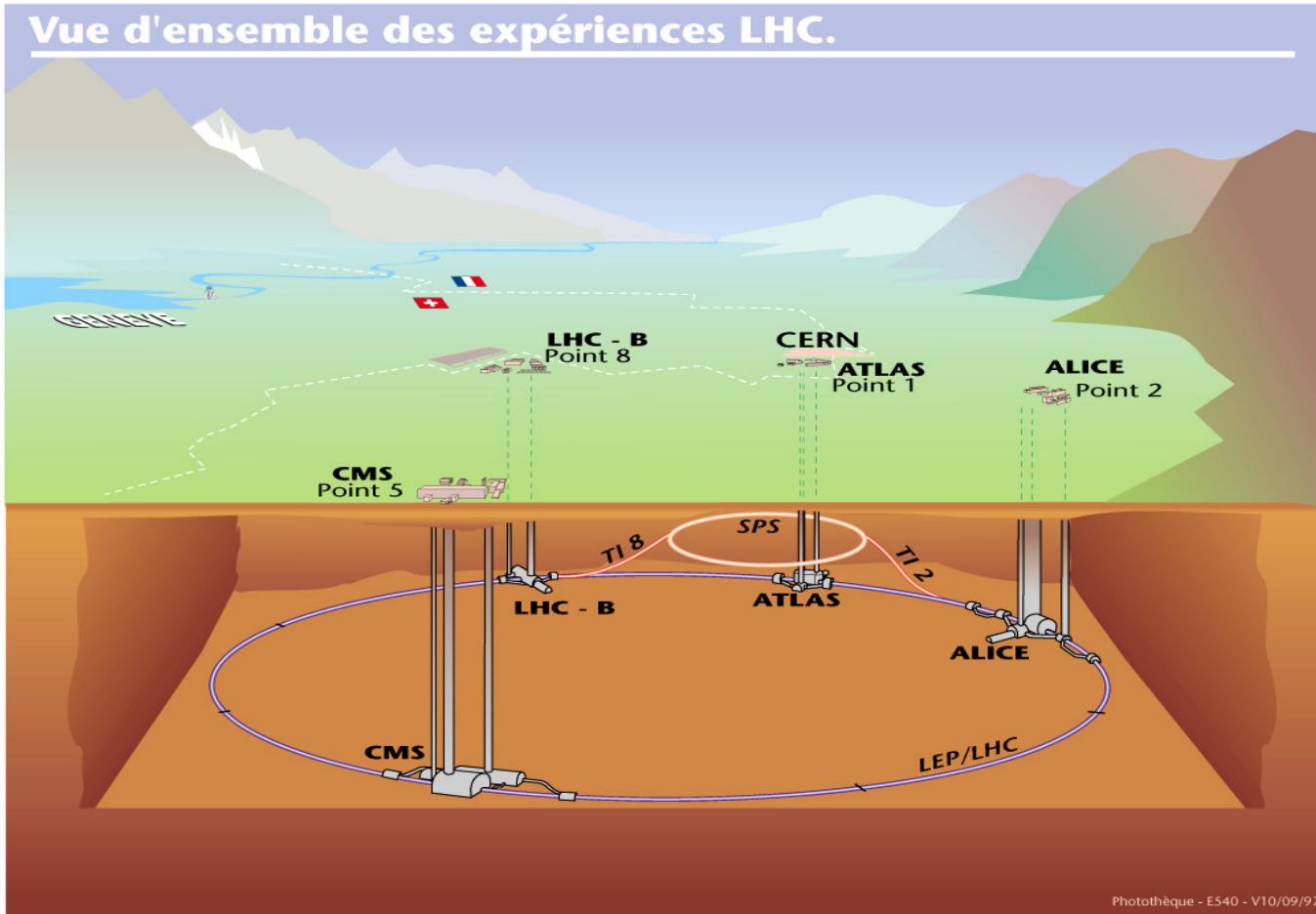
Overview of Introduction

- **Introduction to ATLAS data analysis model**
 - What are the LHC experiments?
 - The ATLAS detector
 - Basic Physics
 - ATLAS Grid computing and WLCG resources
 - Data flow in ATLAS
 - Basis of data analysis in high energy physics
 - Event Data Model (EDM) in ATLAS
 - Basis of data analysis
 - Basic tools: ATLAS CLI and CVMFS
- **ATLAS Metadata Interface (AMI)**
- **Rucio (ATLAS Data Management System)**
- **Introduction to World LHC Computing Grid (WLCG)**
- **PanDA (ATLAS Job Management System)**

Introduction to ATLAS data analysis model



What are the LHC experiments? - 1



The biggest hadron collier on this planet

- Hadron-Hadron collisions in TeV scale (14TeV in 2018)
- 100m below ground
- 27km circular ring
- The world largest distributed computing platform across 36 countries
- 4 main experiments (ATLAS, CMS, Alice and LHCb)
- **Already not super easy!**

Expectations

- New physics
- Understanding of early universe and fundamental laws
 - Validation of many theoretical models

What are the LHC experiments? - 2

- New discovery was announced by ATLAS and CMS in 2012
 - *Higgs boson* (was expected by Standard Model)
 - Nobel prize 2013



What are the LHC experiments? - 3

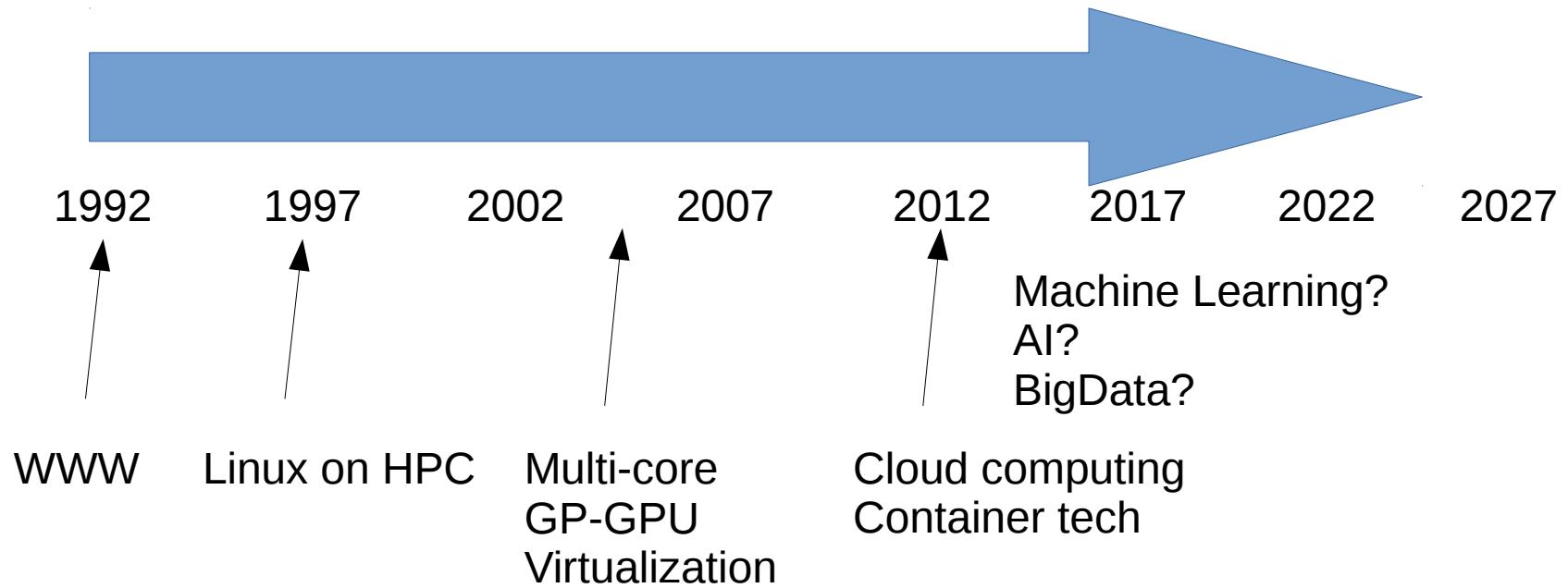
Standard Model of Elementary Particles

three generations of matter (fermions)					
	I	II	III		
mass	$\approx 2.4 \text{ MeV}/c^2$	$\approx 1.275 \text{ GeV}/c^2$	$\approx 172.44 \text{ GeV}/c^2$	0	
charge	2/3	2/3	2/3	0	
spin	1/2	1/2	1/2	1	
QUARKS	u up	c charm	t top	g gluon	Higgs
	$\approx 4.8 \text{ MeV}/c^2$	$\approx 95 \text{ MeV}/c^2$	$\approx 4.18 \text{ GeV}/c^2$	0	
	d down	s strange	b bottom	0	
LEPTONS	e electron	μ muon	τ tau	γ photon	Z boson
	$\approx 0.511 \text{ MeV}/c^2$	$\approx 105.67 \text{ MeV}/c^2$	$\approx 1.7768 \text{ GeV}/c^2$	0	
	-1	-1	-1	1	
	1/2	1/2	1/2		
	ν_e electron neutrino	ν_μ muon neutrino	ν_τ tau neutrino		± 1
	$< 2.2 \text{ eV}/c^2$	$< 1.7 \text{ MeV}/c^2$	$< 15.5 \text{ MeV}/c^2$	1	
	0	0	0		
	1/2	1/2	1/2		
SCALAR BOSONS					
GAUGE BOSONS					
GAUGE BOSONS					

This

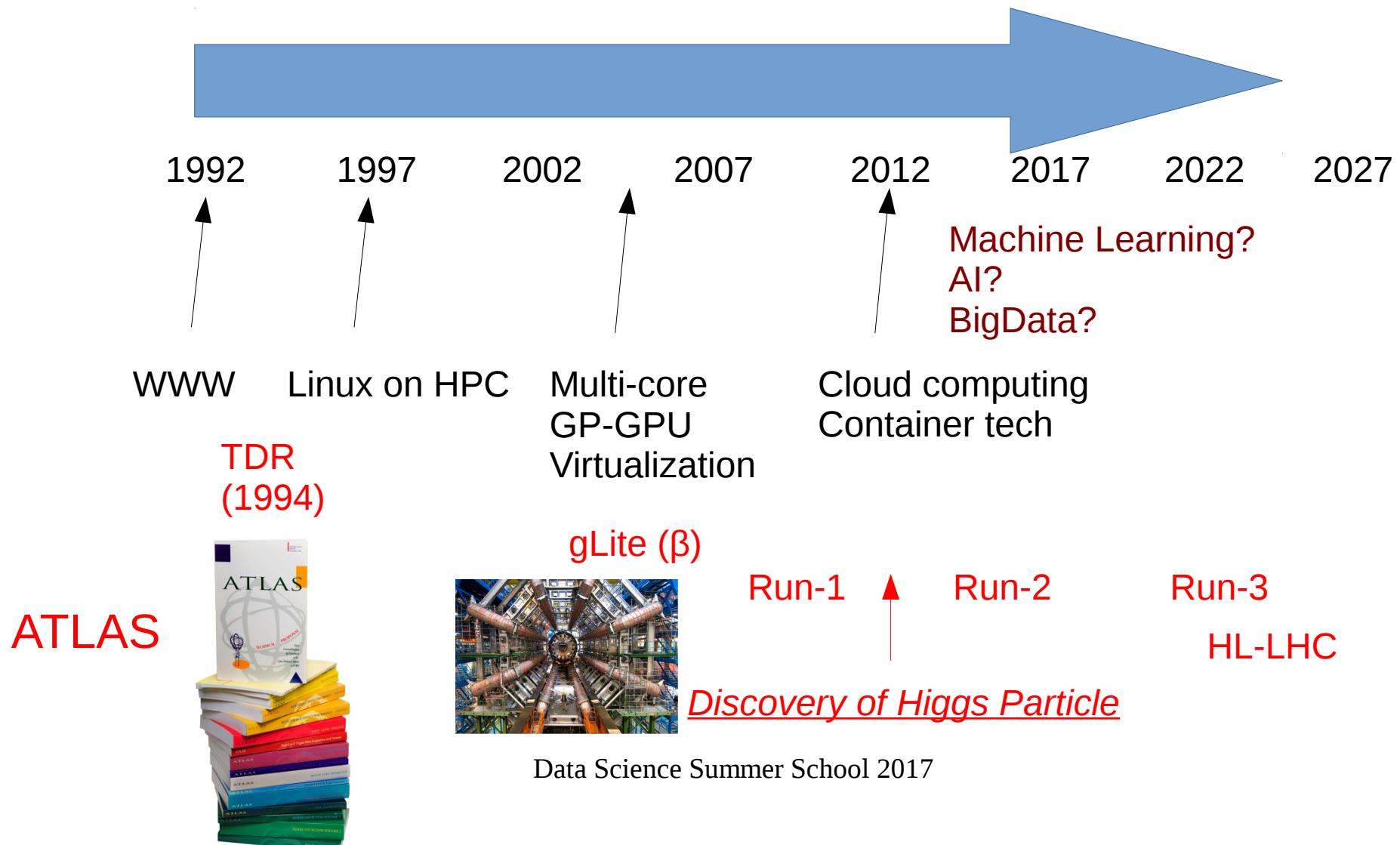
What are the LHC experiments? - 4

- Has continued since 1992



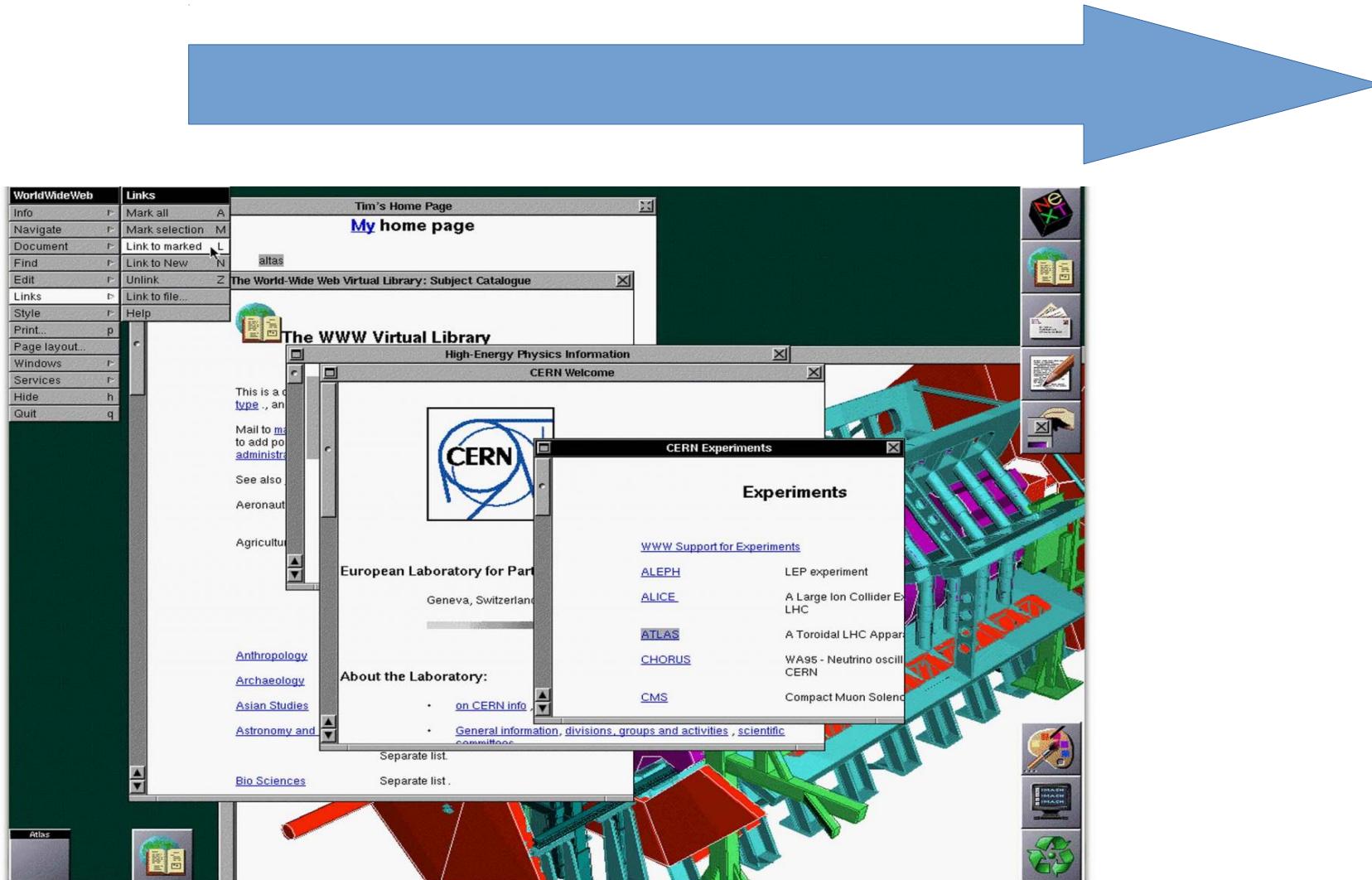
What are the LHC experiments? - 4

- Has continued since 1992



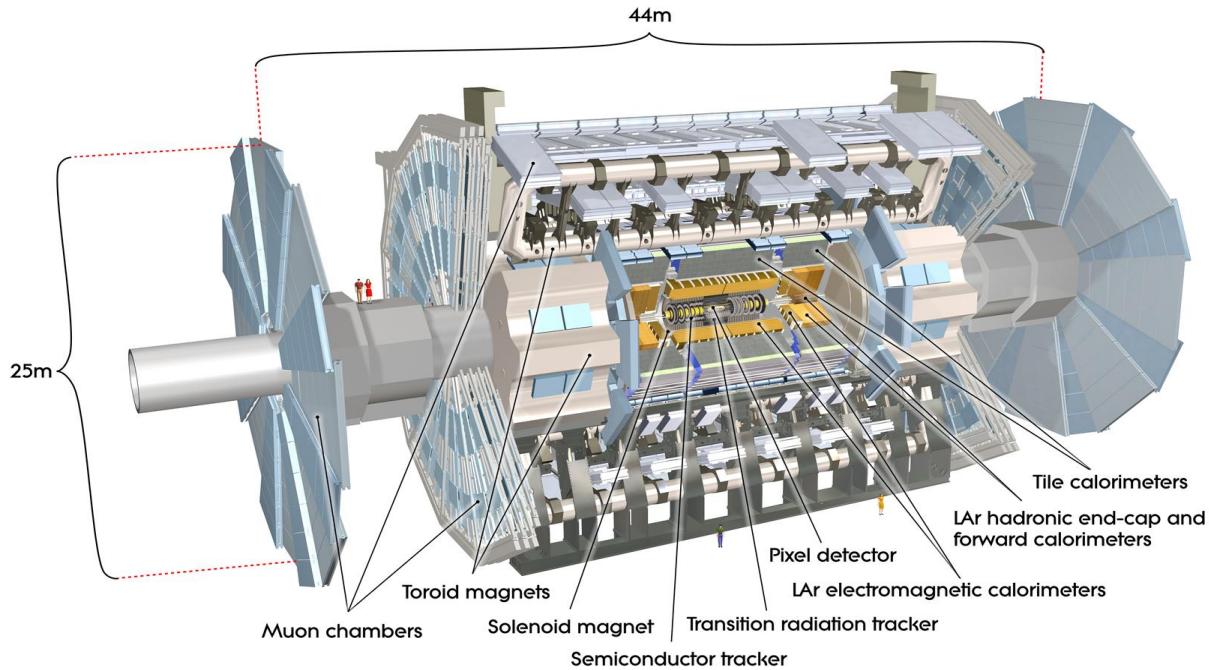
What are the LHC experiments? - 4

- Has continued since 1992



Screenshot of the original NeXT web browser in 1993 (Image: Berners-Lee/CERN)

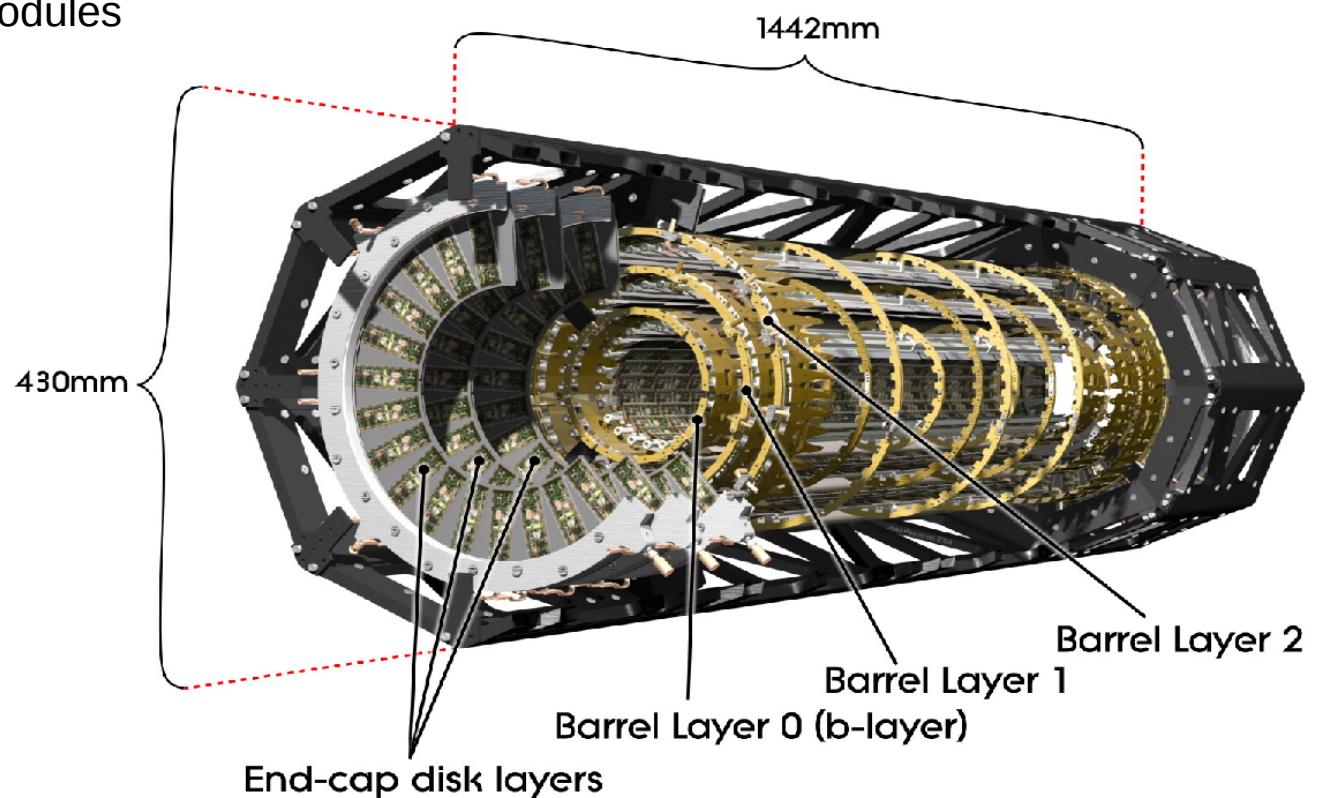
The ATLAS detector



- General-purpose HEP experiment
 - Reconstruction of outgoing particles from proton-proton collisions at 13TeV
 - pp-rate: 40MHz
 - Length: 46m, Height 25m, Weight 7kt, B-field 4T

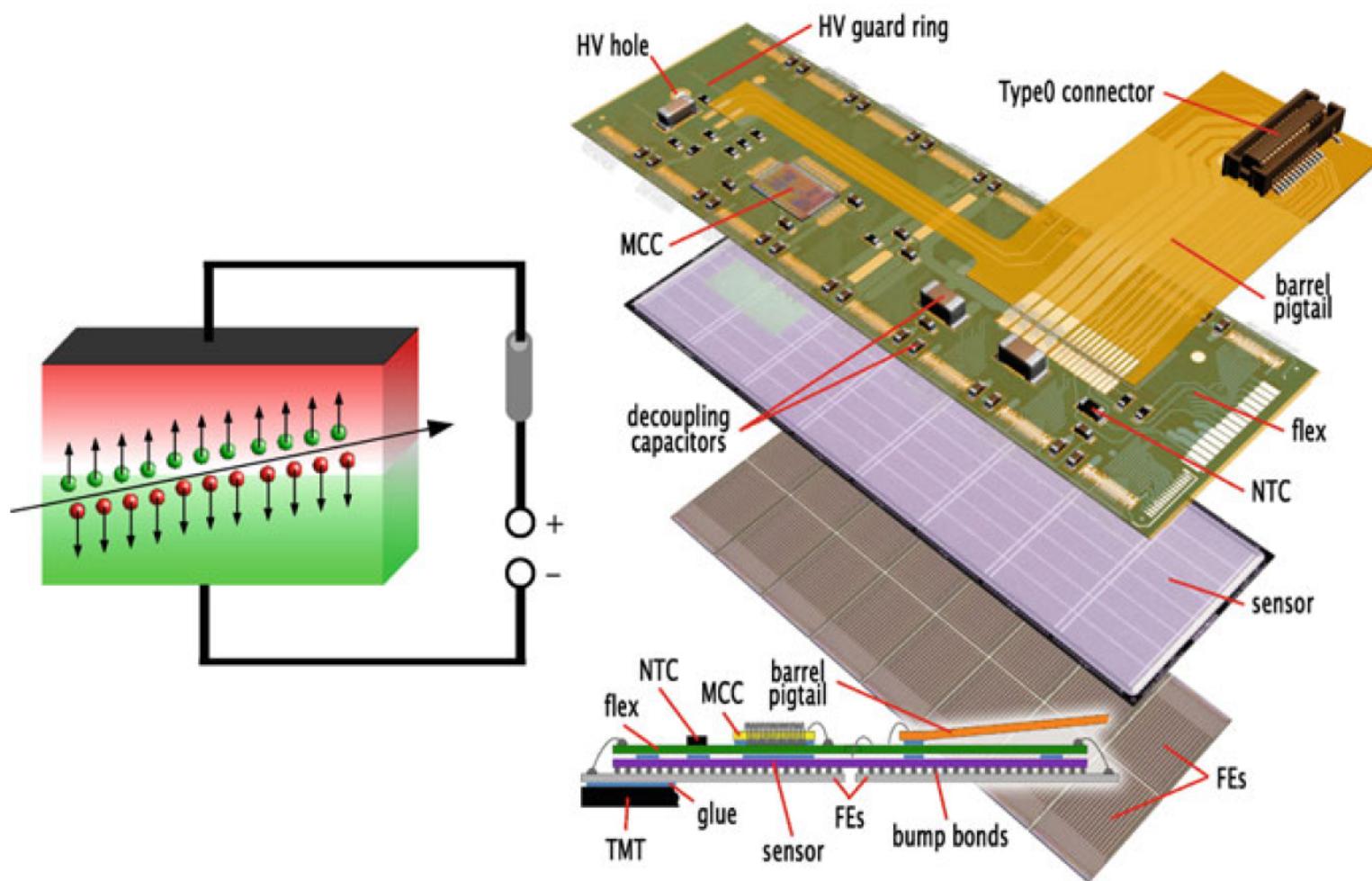
Inner Pixel Detector - 1

- 80 million pixels (80 million channels). Area 1.7m^2 . 15 kW power consumption
- Barrel has 1,744 modules (10cm^2) with 46,080 readout channels per module
- Pixel size $50 \times 400\mu\text{m}^2$. Resolution $14 \times 115\mu\text{m}^2$
- Three Pixel disks (in each endcap) have 6.6 million channels
- 3 barrel layers: 1,456 modules
- 3 disks in each end-cap: 288 modules

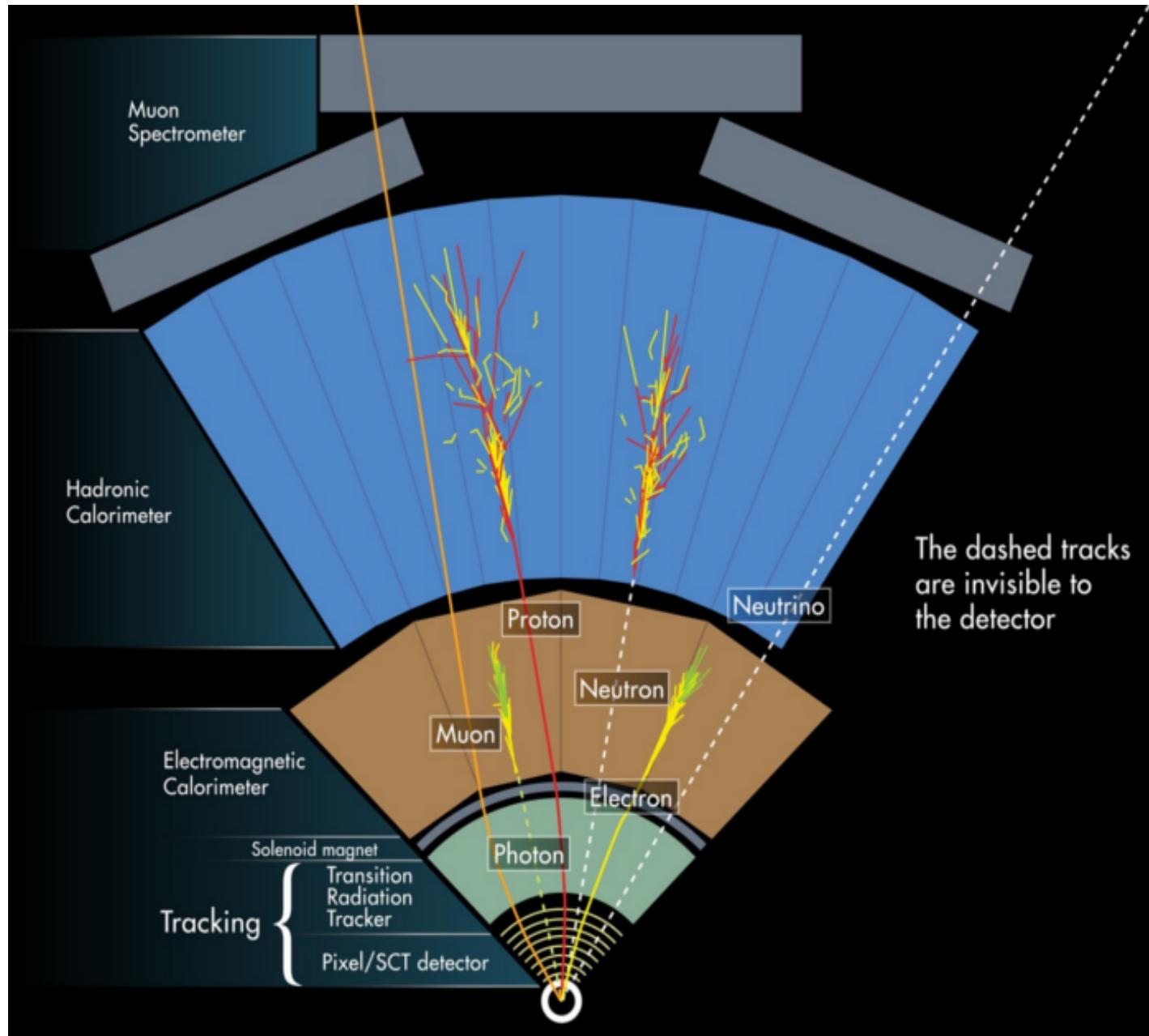


Inner pixel detector - 2

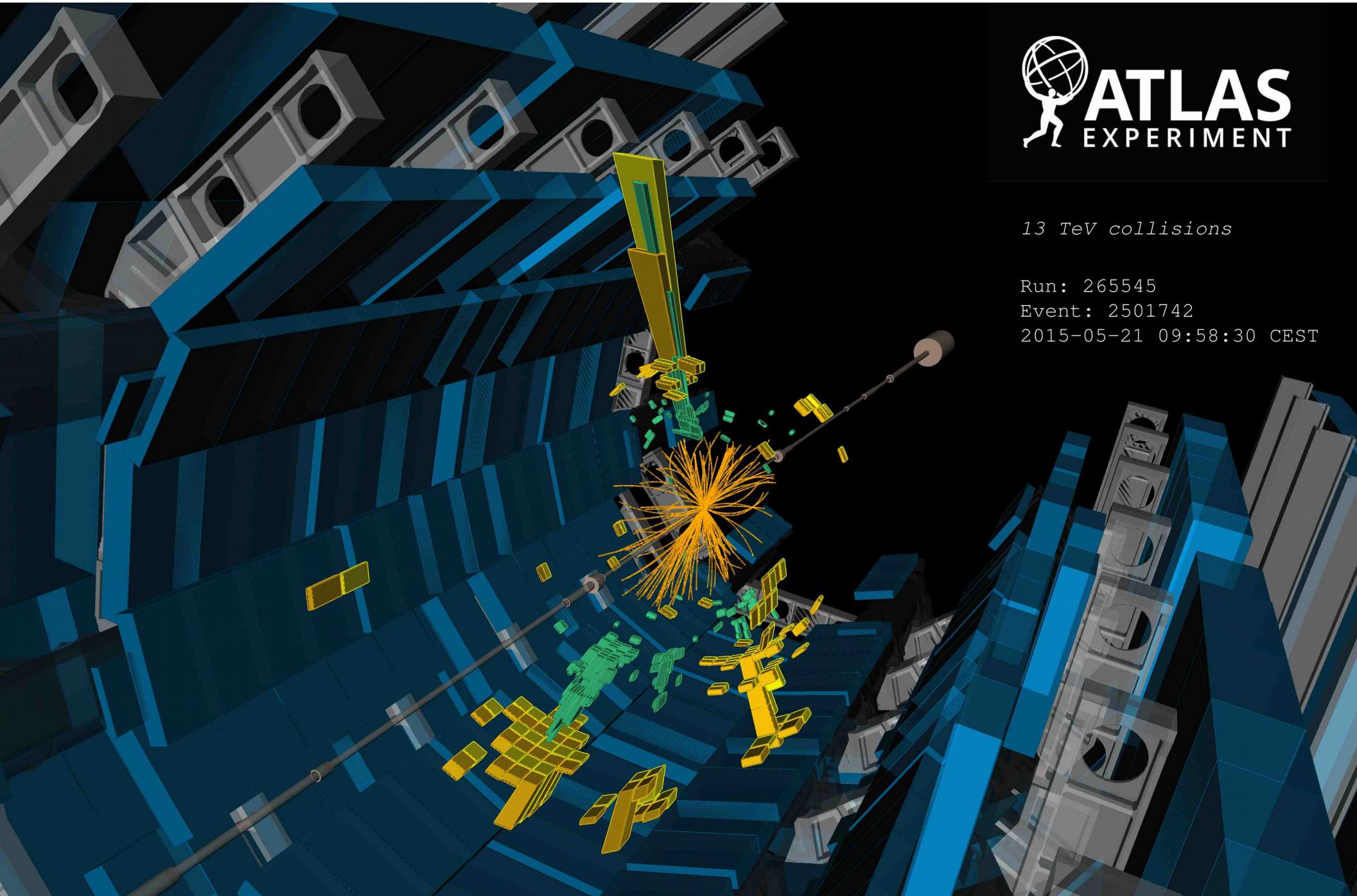
- Works like (very complicated) CCD camera



Silicon Tracker & Trigger - 2



One event



13 TeV collisions

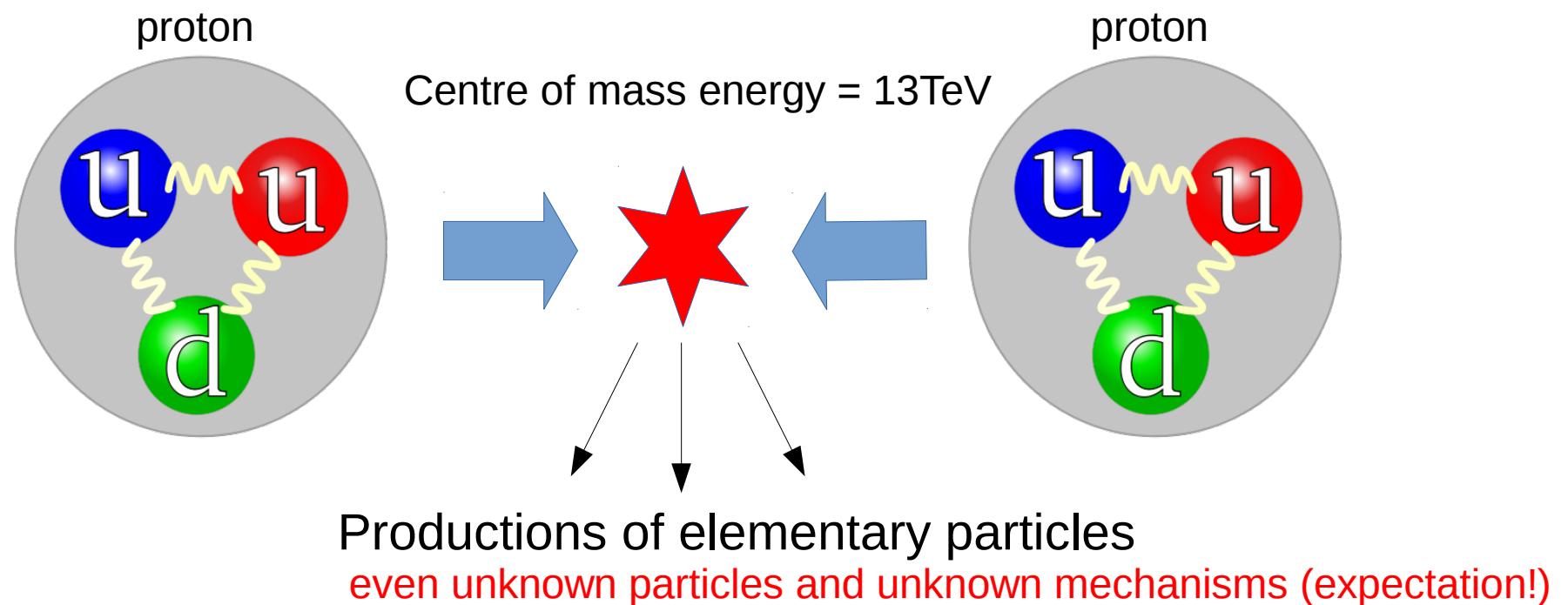
Run: 265545

Event: 2501742

2015-05-21 09:58:30 CEST

Basic Physics - 1

- Proton-Proton (Hadron-Hadron) Collisions
 - Proton = 2 Up quarks, 1 Down quark bound by Gluons (strong force)



Basic Physics - 2

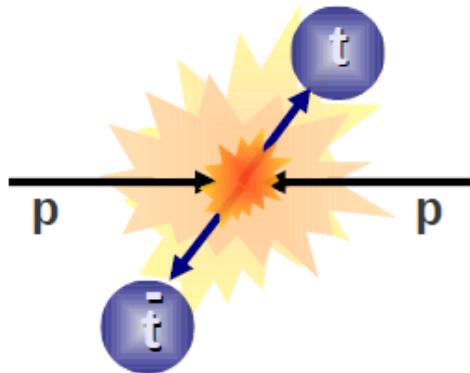
- Production of Teilchenzoo (particle zoo)



Basic Physics - 3

- For example, Top-Quark production

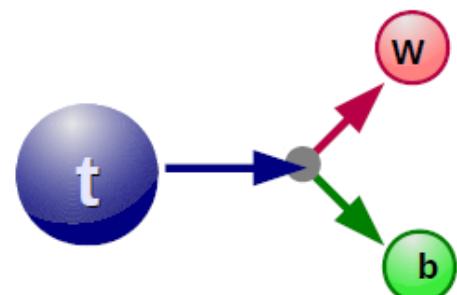
Production:



Intrinsic properties:



Decay:



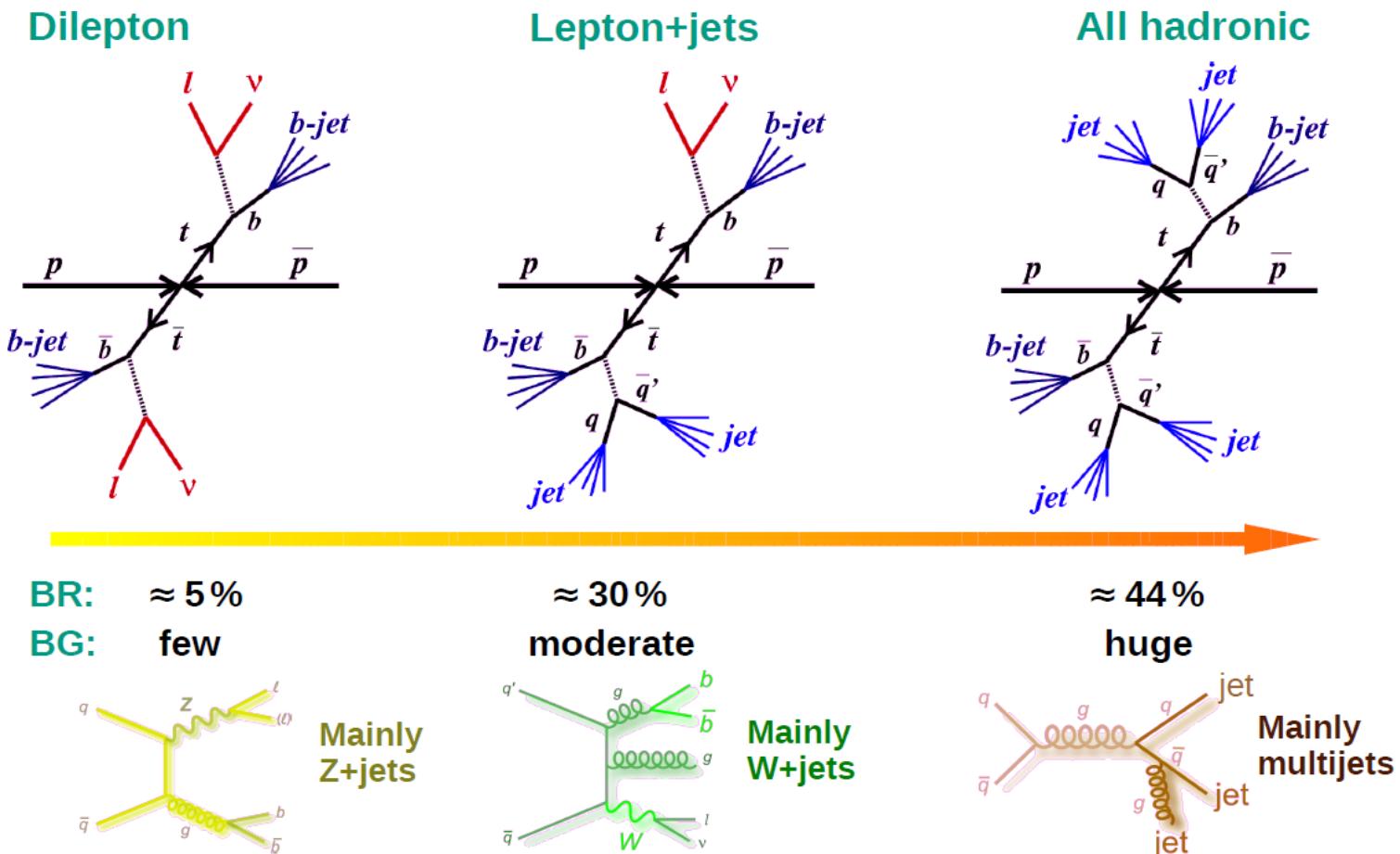
- Production rate of different processes (*strong and EW*)
- Differential distributions
- New production mechanisms

- Top quark mass \rightarrow constraints on Higgs mass
- Charge
- Lifetime

- Decay channels (*SM and new*)
- Couplings

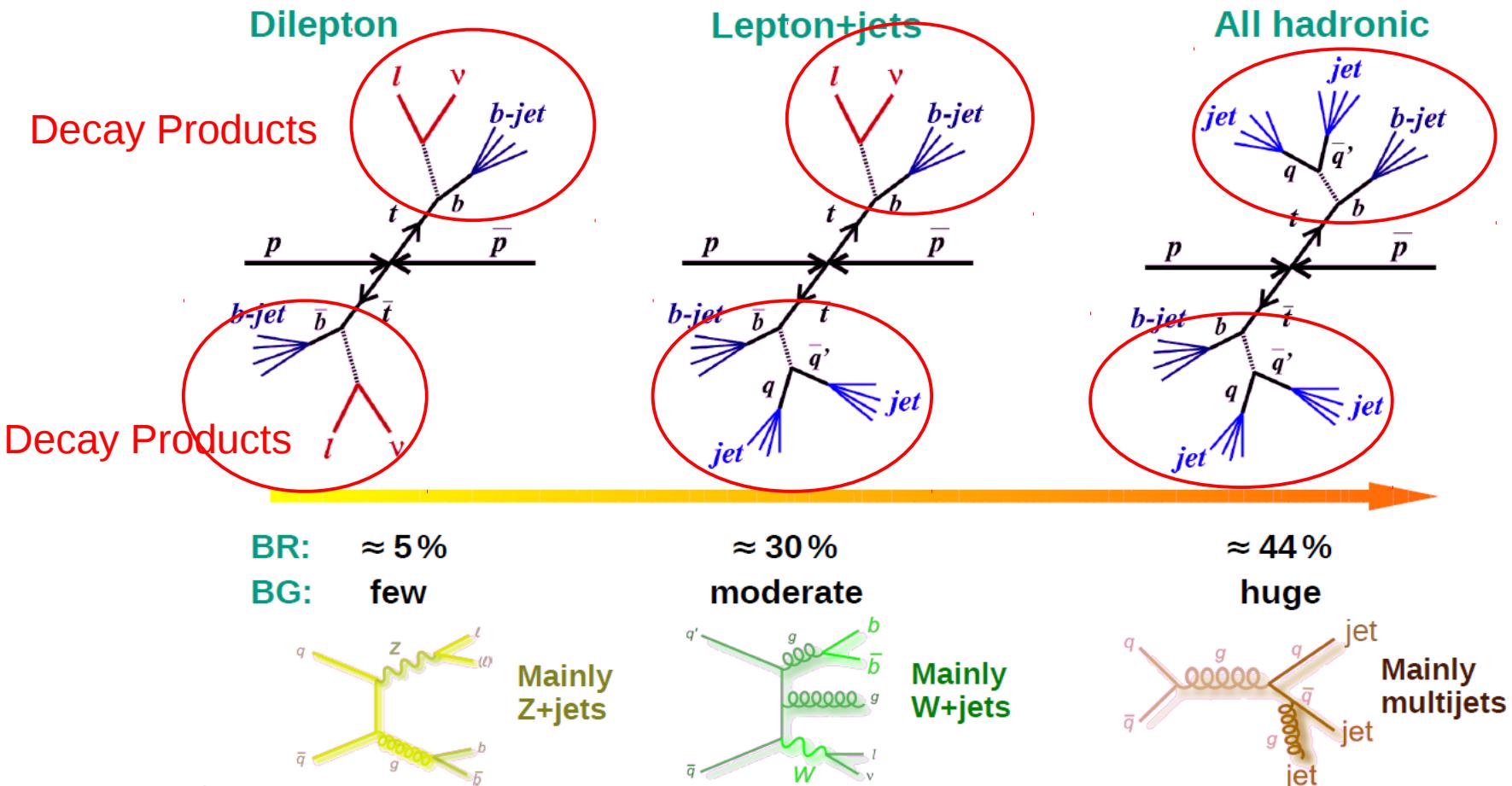
Basic Physics - 4

- For example, Top-Quark production
 - Event signatures



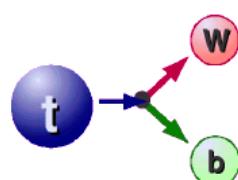
Basic Physics - 5

- For example, Top-Quark production
 - Event signatures



Basic Physics - 6

- For example, Top-Quark production
 - Decay products

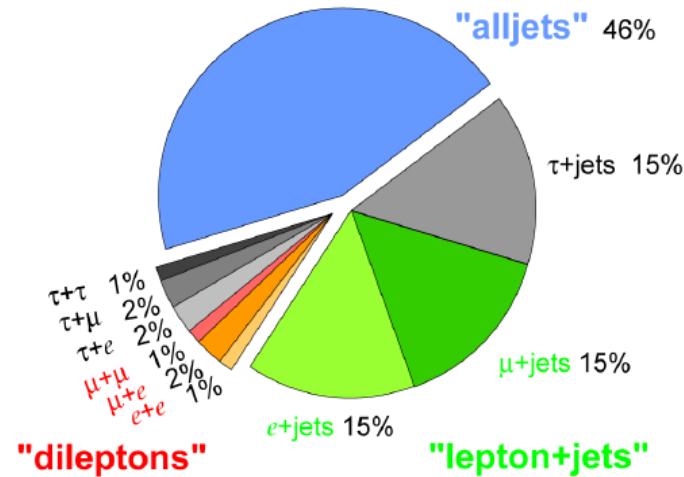


SM:

$$t \rightarrow b \ W \approx 100 \%$$

Decay of top quark pairs

$\bar{c}s$	electron+jets	muon+jets	tau+jets	all-hadronic
$\bar{u}d$				
$-\tau$	$e\tau$	$\mu\tau$	$\tau\tau$	tau+jets
$-\mu$	$e\mu$	$\mu\tau$	$\mu\tau$	muon+jets
$-e$	$e\mu$	$e\mu$	$e\tau$	electron+jets
W decay	e^+	μ^+	τ^+	$u\bar{d}$
				$c\bar{s}$



"lepton+jets"-channel:
($e+jets$ and $\mu+jets$)

- Moderate background
- Relative high branching ratio (~30%)

ATLAS Grid computing and WLCG resources - 1

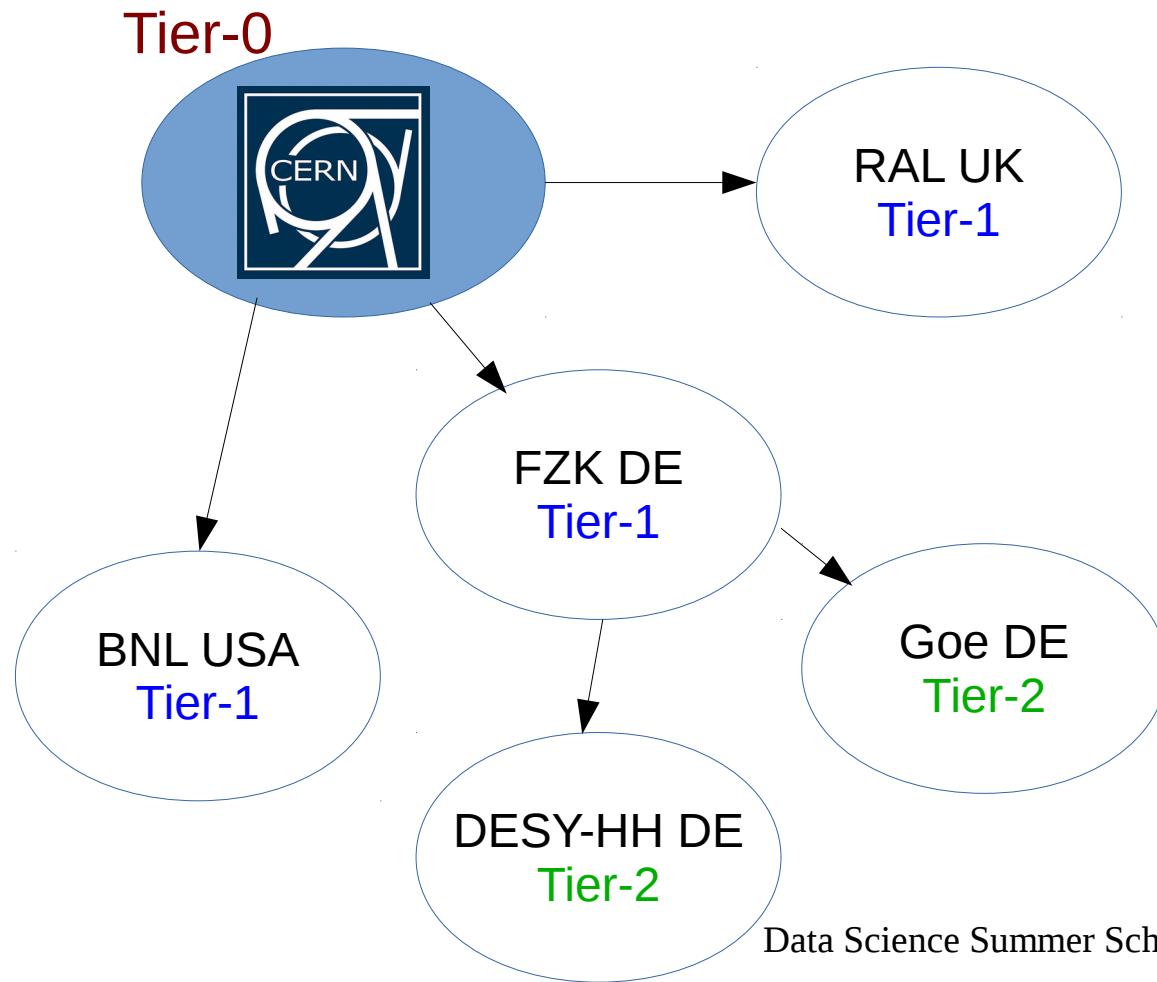
- World LHC Computing Grid (WLCG)
 - The most sophisticated data-taking & analysis system ever built for science, providing near real-time access to LHC data
 - 2 million computing jobs per day
 - 42 countries
 - 170 computing centres



ATLAS Grid computing and WLCG resources - 2

- LHC multi-tier structure

- WLCG = Worldwide LHC Computing Grid



Tier-0: Raw data, Data store (in tape), Pre-processing, Reconstruction

Tier-1: National LCG-Centre, faster network connections and larger storage spaces (e.g. Tape), MC production, user analysis, etc.

Tier-2: University or Facility level computing sites. MC production, user analysis, etc.

THE LHC AND EXPERIMENTS

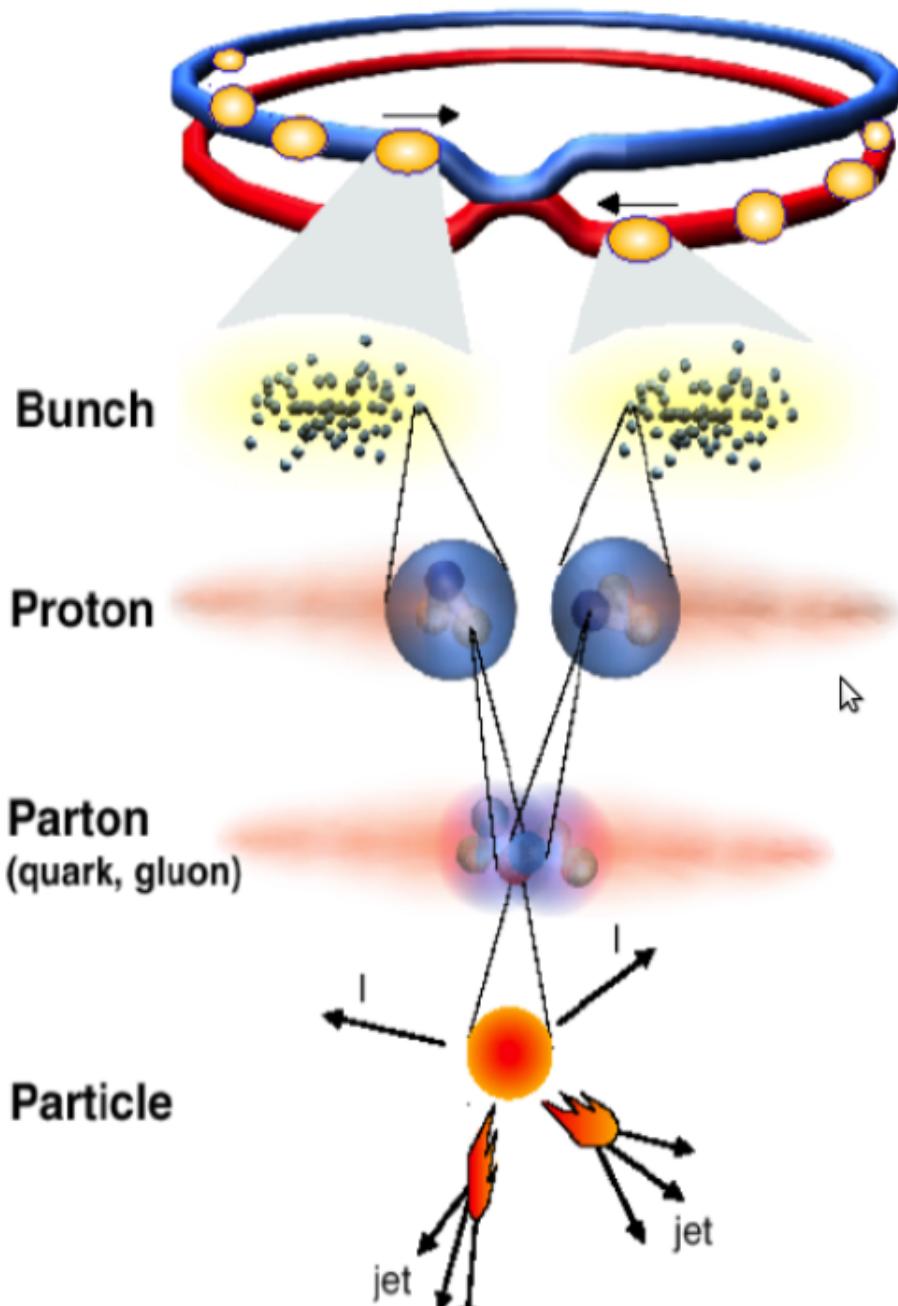


DETECTORS BUILT AND OPERATED BY A LARGE TEAM



Worldwide Collaboration of over 3000 physicists and engineers in ATLAS and CMS each + similar in LHCb and ALICE

COLLISIONS AT THE LHC



Proton-Proton-Kollisionen
2835 Teilchenbündel (Bunch)

10^{11} Protonen / Bunch
Kollisionsrate 40 MHz (25 ns)

Schwerpunktsenergie 14 TeV
(= $7400 \times$ Ruheenergie der kollidierenden Teilchen)

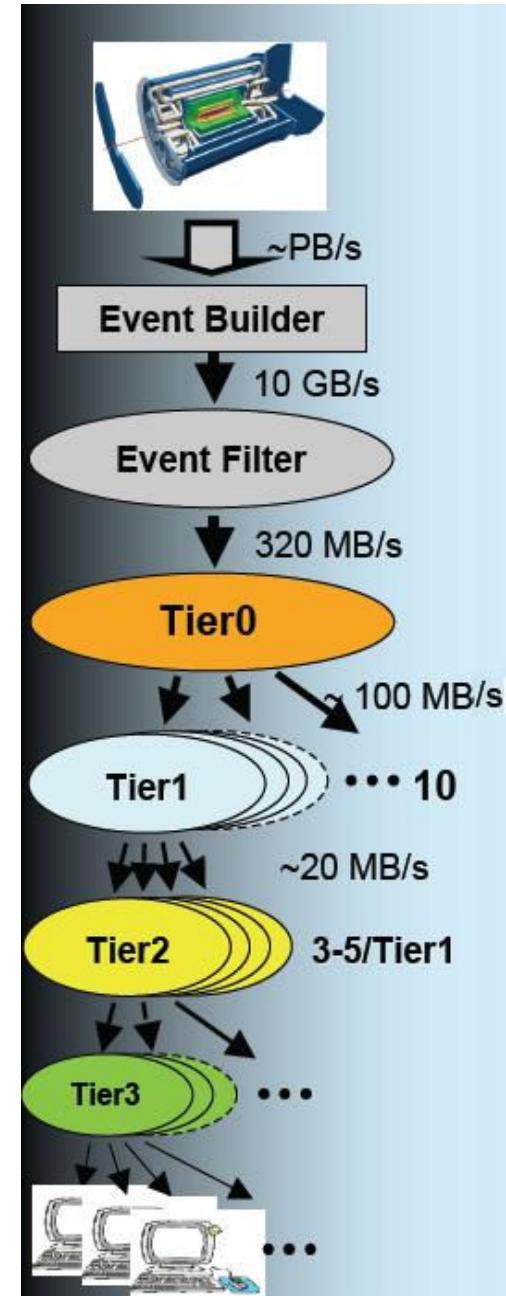
Schwerpunktsenergie der kollidierenden Quarks und Gluonen bis einige TeV

~25 pp-Kollisionen pro Bunch-Kollision

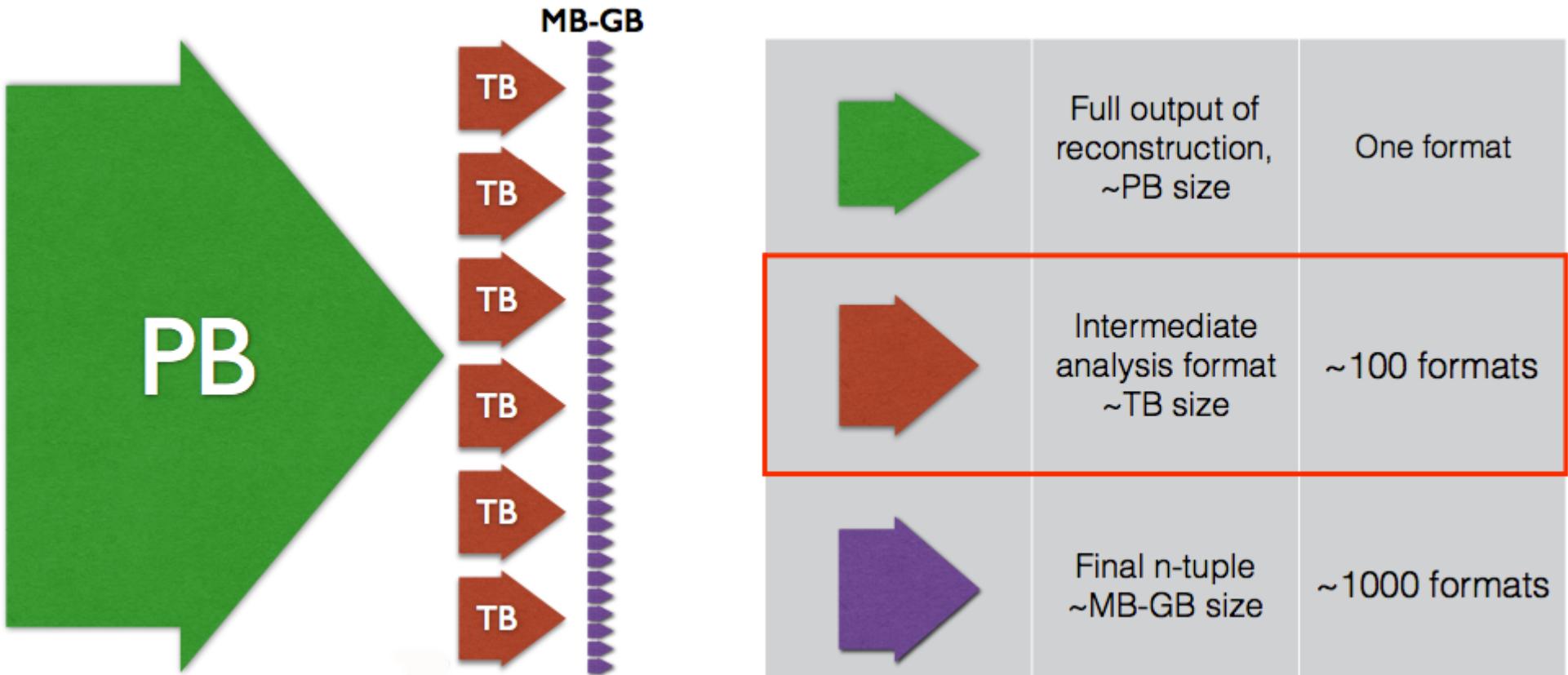
Interessante Ereignisse: $10^{-9} - 10^{-11}$ unterdrückt!

Data flow in ATLAS

- **RAW** = about 2 MB/event
 - Original data at **Tier-0** (ATLAS detector)
 - Complete replica distributed among all **Tier-1s**
- **ESD (Event Summary Data)** = 1 MB/event
 - ESDs produced by primary at **Tier-1s**
 - Globally 2 copies on disk
- **AOD (Analysis Object Data)** = 0.2 MB/event
 - Completely replicated at each **Tier-1**
 - Partially replicated to **Tier-2s** so as to have a complete set in the **Tier-2s** associated to each national **Tier-1**
 - Every **Tier-2** specifies which datasets are most interesting for their reference community; the rest are distributed according to capacity
- **TAG (Event Level Metadata Tags)** = 1 kB/event
 - TAG files or databases are replicated to all **Tier-1s** (Root/Oracle)
 - Partial replicas of the TAG are distributed to **Tier-2s** as Root files
 - Each **Tier-2** have all TAGs that correspond to the AODs stored



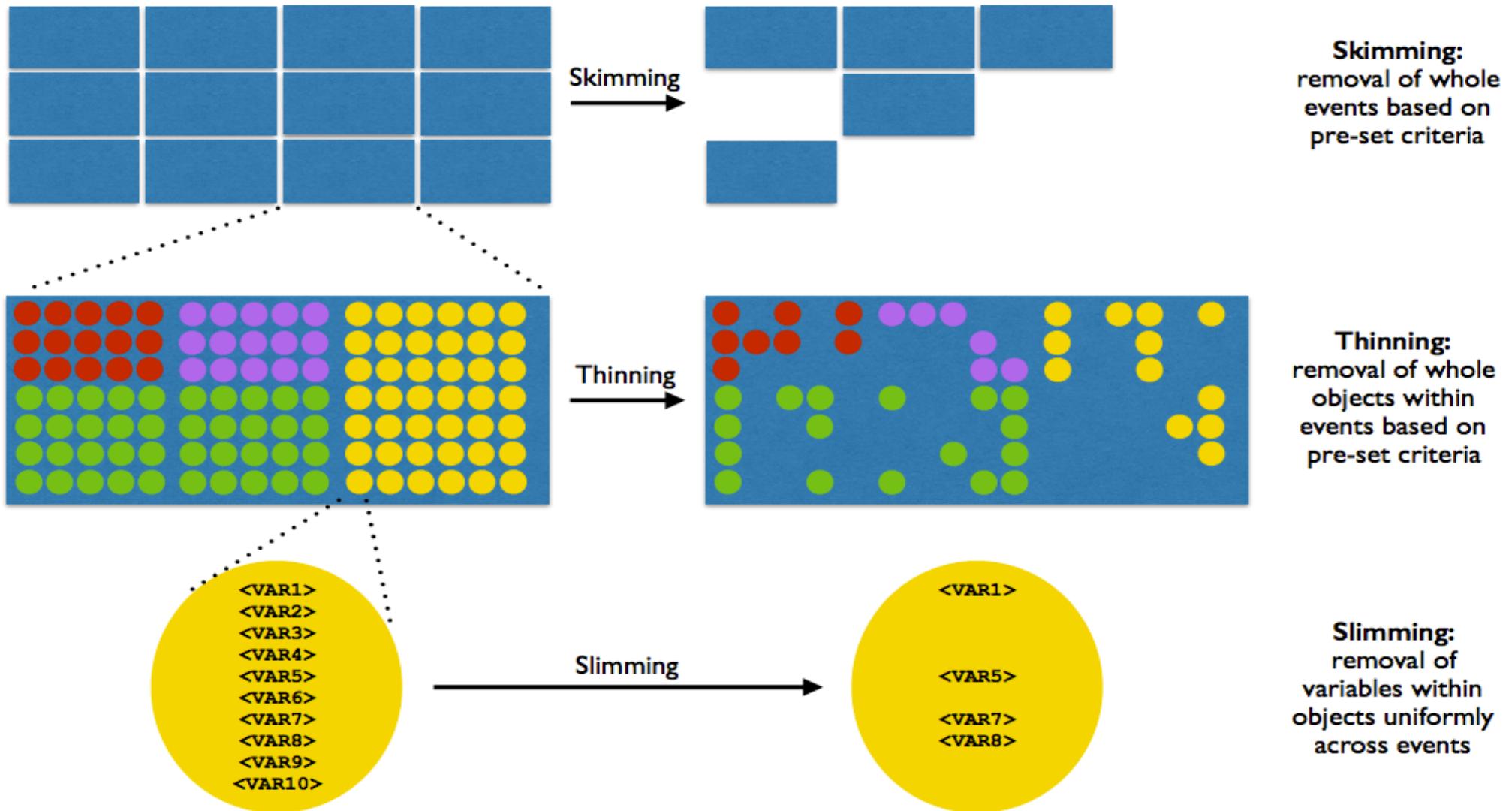
Event Data Model (EDM) in ATLAS



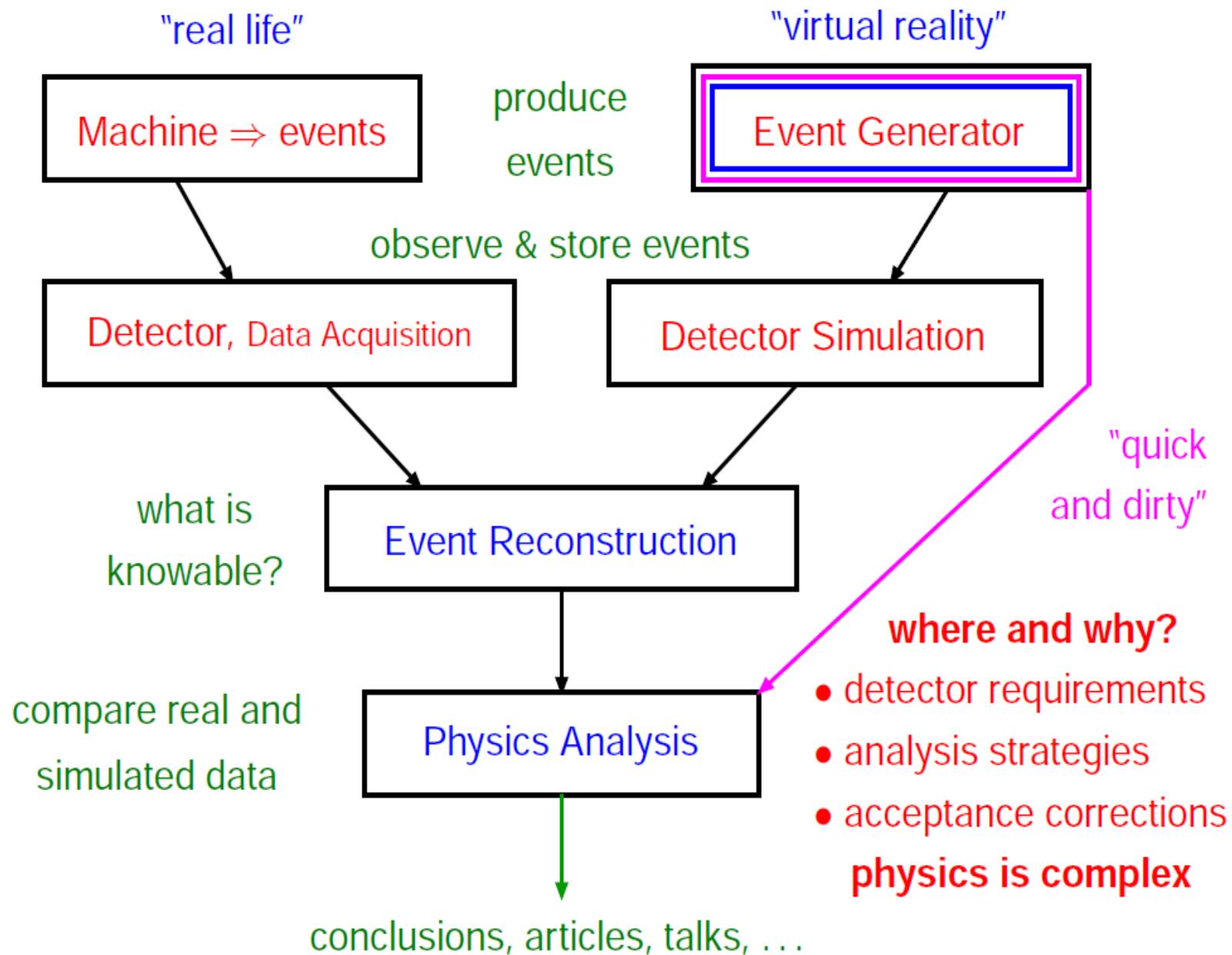
- AODs are still huge, so need to have manageable size of data → Derived AOD (DAOD)
- Derivation framework in Athena (ATLAS software framework)
- Each analysis group use own derivation framework

Event Data Model (EDM) in ATLAS

- Size reduction (= selection of events, particles and parameters)



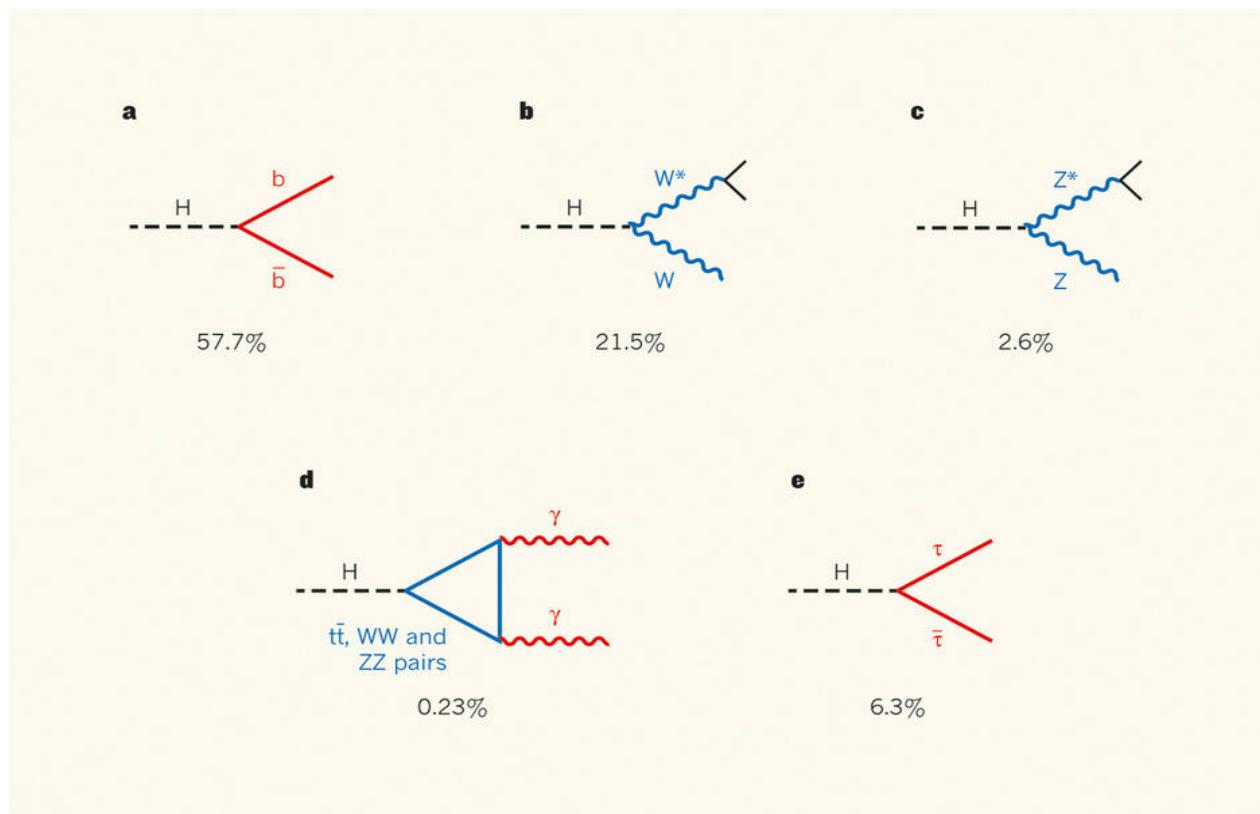
Basis of data analysis - 1



Basis of data analysis - 2

- Higgs decay channel

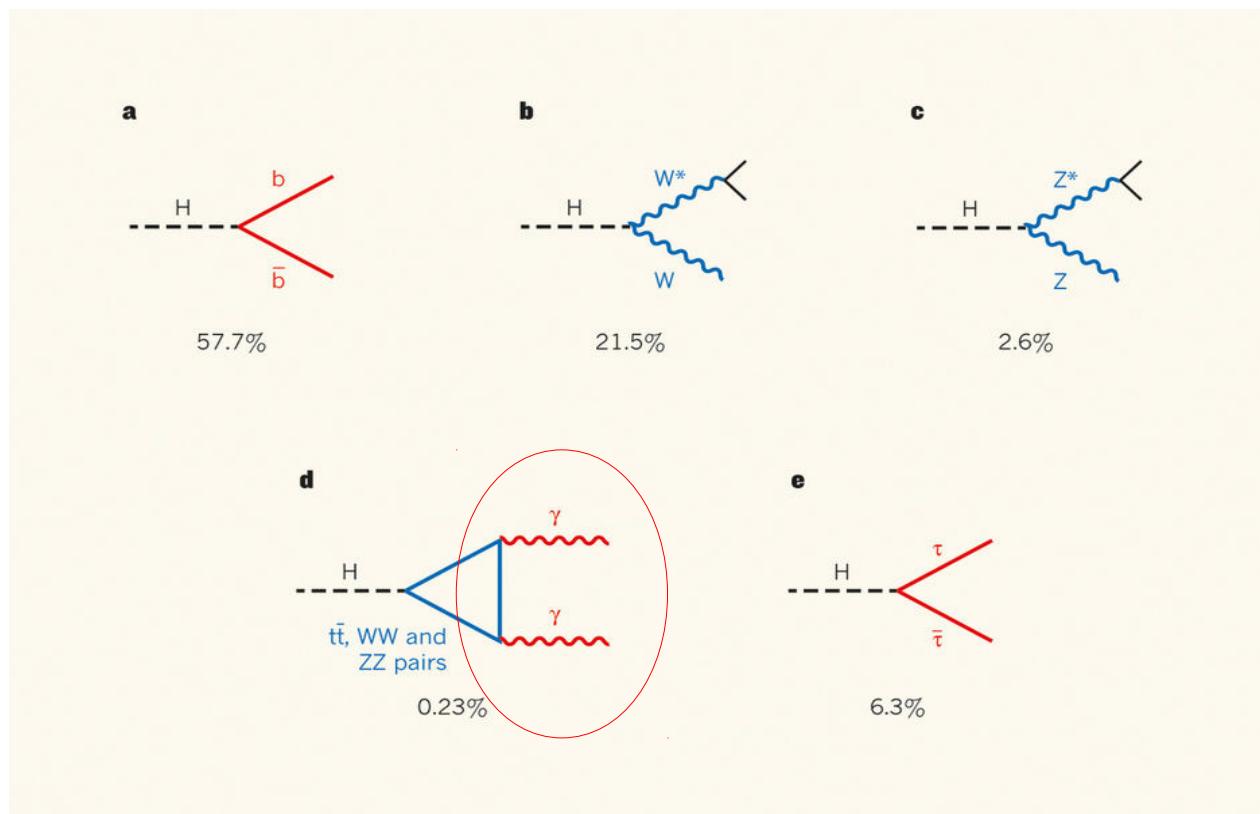
- 1 EUR \rightarrow 50 + 50 cent
- 1 EUR \rightarrow 50 + 20 + 20 + 10 cent
- 1 EUR \rightarrow 20 + 20 + 20 + 20 + 20 cent
- 1 EUR \rightarrow 50 + 20 + 10 + 10 + 10 cent



Basis of data analysis - 2

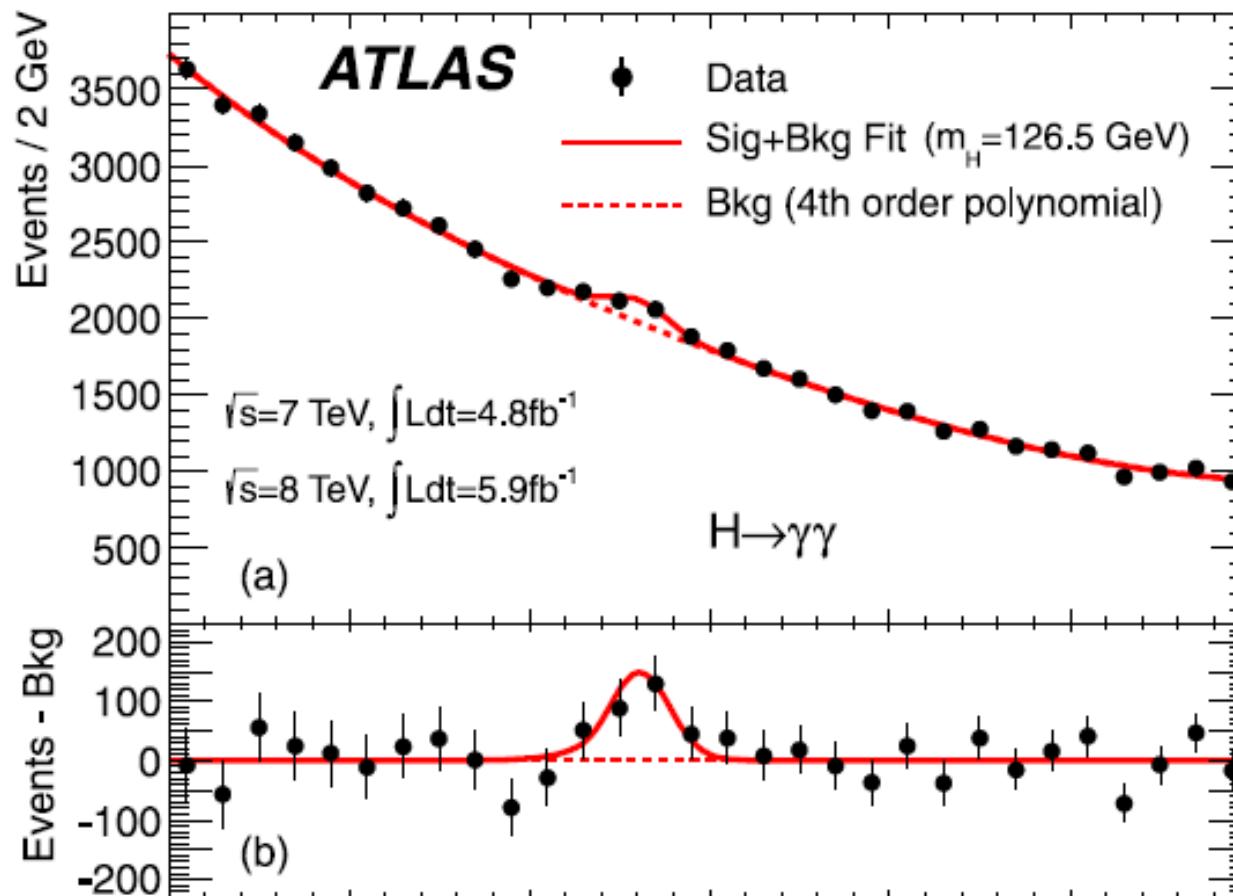
- Higgs decay channel

- 1 EUR \rightarrow 50 + 50 cent
- 1 EUR \rightarrow 50 + 20 + 20 + 10 cent
- 1 EUR \rightarrow 20 + 20 + 20 + 20 + 20 cent
- 1 EUR \rightarrow 50 + 20 + 10 + 10 + 10 cent



Basis of data analysis - 3

- Atlas Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC." arXiv preprint arXiv:1207.7214 (2012).



Basis of data analysis - 4

ATLAS Collaboration / Physics Letters B 716 (2012) 1–29

13

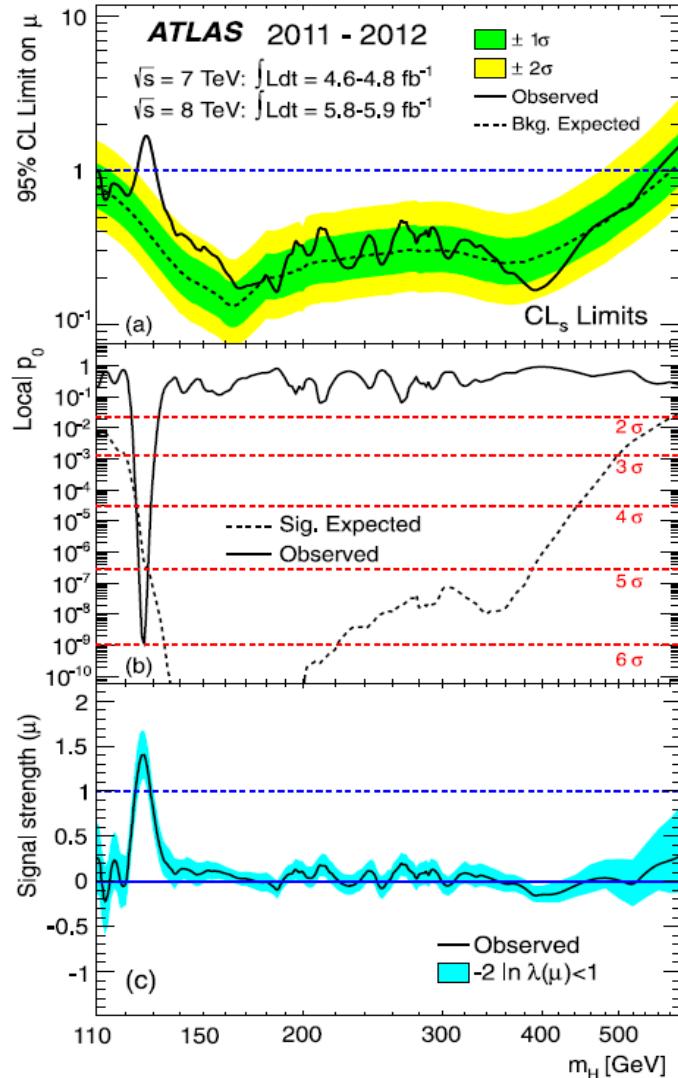


Fig. 7. Combined search results: (a) The observed (solid) 95% CL limits on the signal strength as a function of m_H and the expectation (dashed) under the background-only hypothesis. The dark and light shaded bands show the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties on the background-only expectation. (b) The observed (solid) local p_0 as a

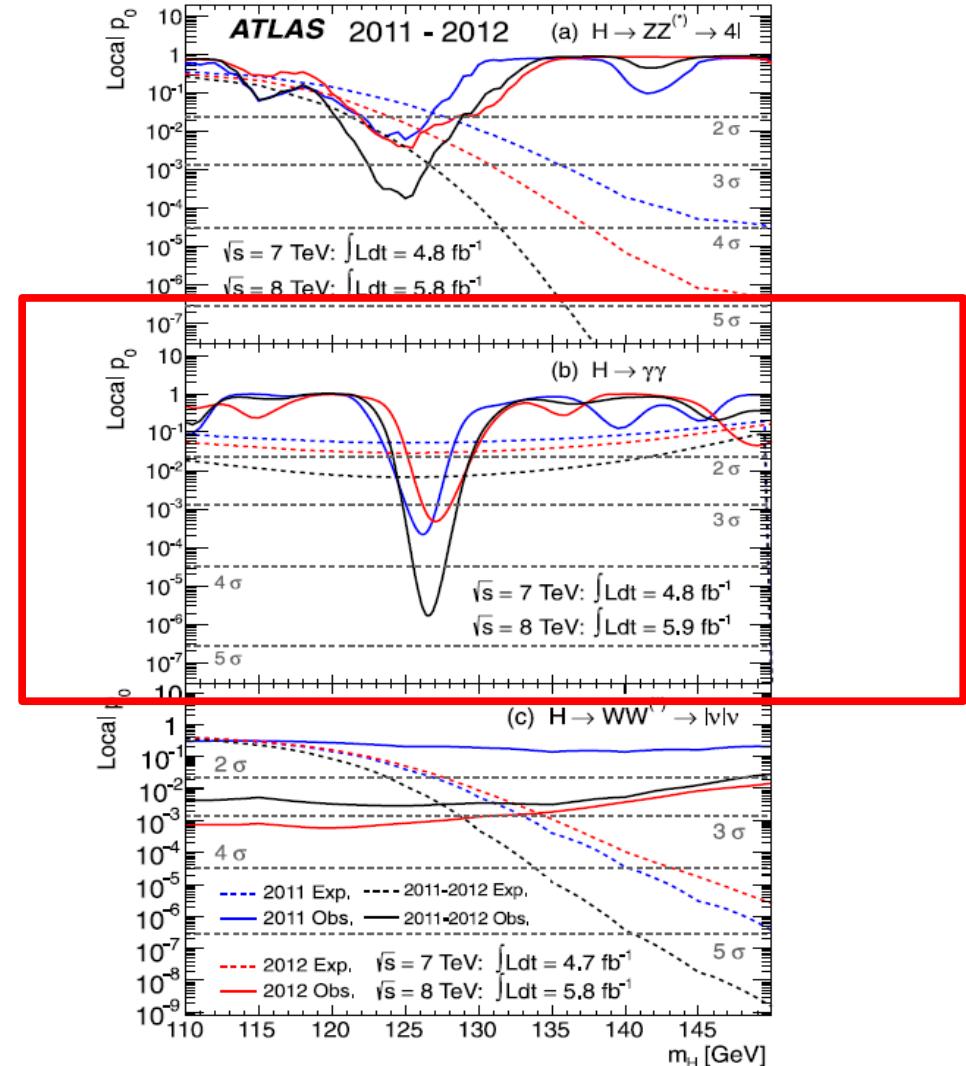


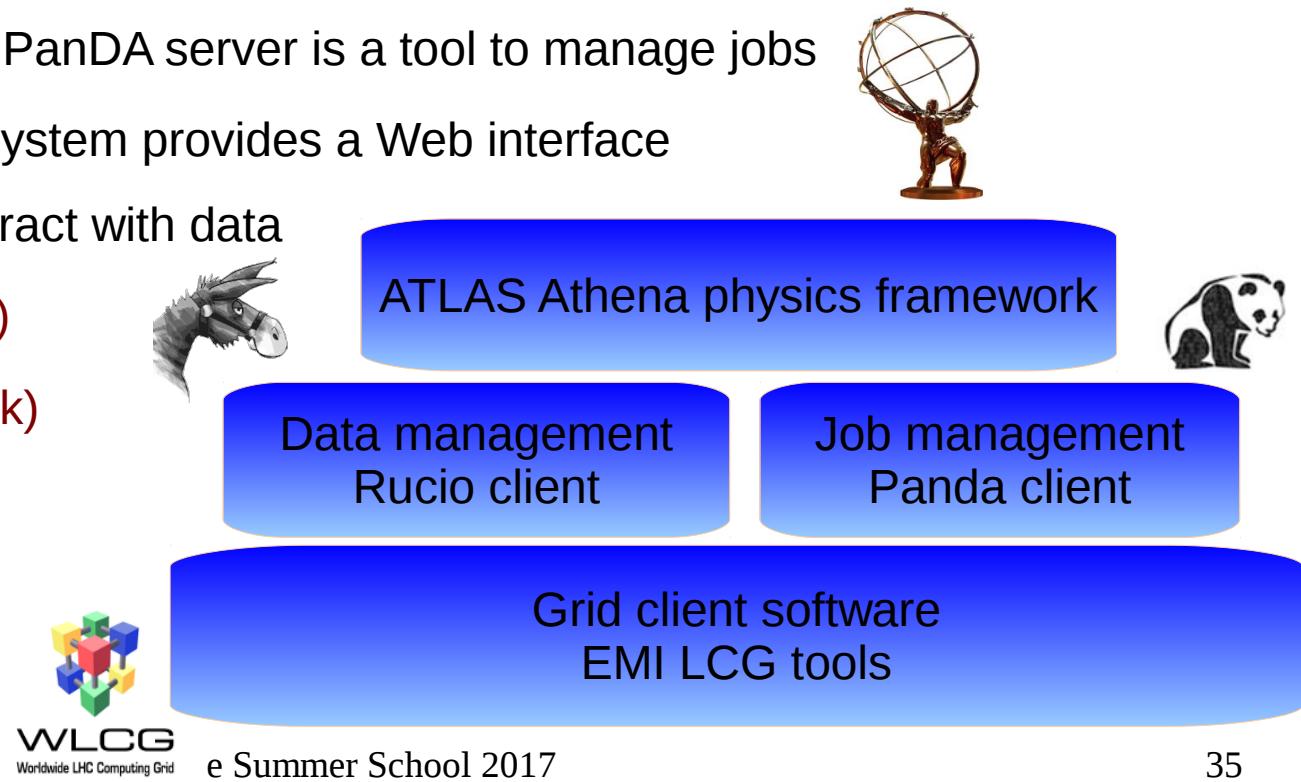
Fig. 8. The observed local p_0 as a function of the hypothesised Higgs boson mass for the (a) $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$, (b) $H \rightarrow \gamma\gamma$ and (c) $H \rightarrow WW^{(*)} \rightarrow l\bar{l}v\bar{v}$ channels. The dashed curves show the expected local p_0 under the hypothesis of a SM Higgs boson signal at that mass. Results are shown separately for the $\sqrt{s} = 7 \text{ TeV}$ data (dark blue in the web version), the $\sqrt{s} = 8 \text{ TeV}$ data (light red in the web version)

As a result: the computing model

- In principle, no need of MPI
 - The probability of each event is independent
 - The pseudo-random number generator: affects the quality of Monte-Carlo generator, event generator
- Easily paralleled and even widely distributed among data locations
- Many end-users from different institutes
- Budgets, manpower, computing facilities, networks
 - Also distributed

Basic Tools: ATLAS CLI and CVMFS

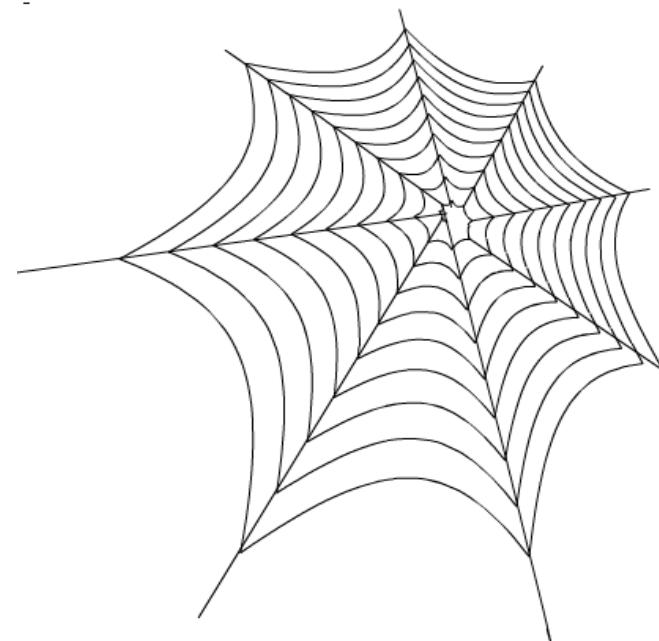
- Grid client software originally supported by *European Middleware Initiative (EMI)*
- CVMFS is a remote repository using FUSE file system
 - e.g. /cvmfs/atlas.cern.ch → A repository of all client software we need
 - Internally using HTTP → Need of network access
- PanDA client using ATLAS PanDA server is a tool to manage jobs
 - BigPanda monitoring system provides a Web interface
- Rucio client is a tool to interact with data
- ATLAS framework (Athena)
- ROOT (c++ HEP framework)



CVMFS (CERN VM File-System)

- Before CVMFS
 - Problems of distribution of common High Energy Physics software environments
 - Needed to build every software release in every university/facility
- After CVMFS
 - Everything is centrally managed/updated
 - Nightly build is also available
 - Nothing to do (e.g. build of every update) in local PCs / Clusters
 - Using general http proxy system like Squid → less network access and robust

ATLAS Metadata Interface (AMI)

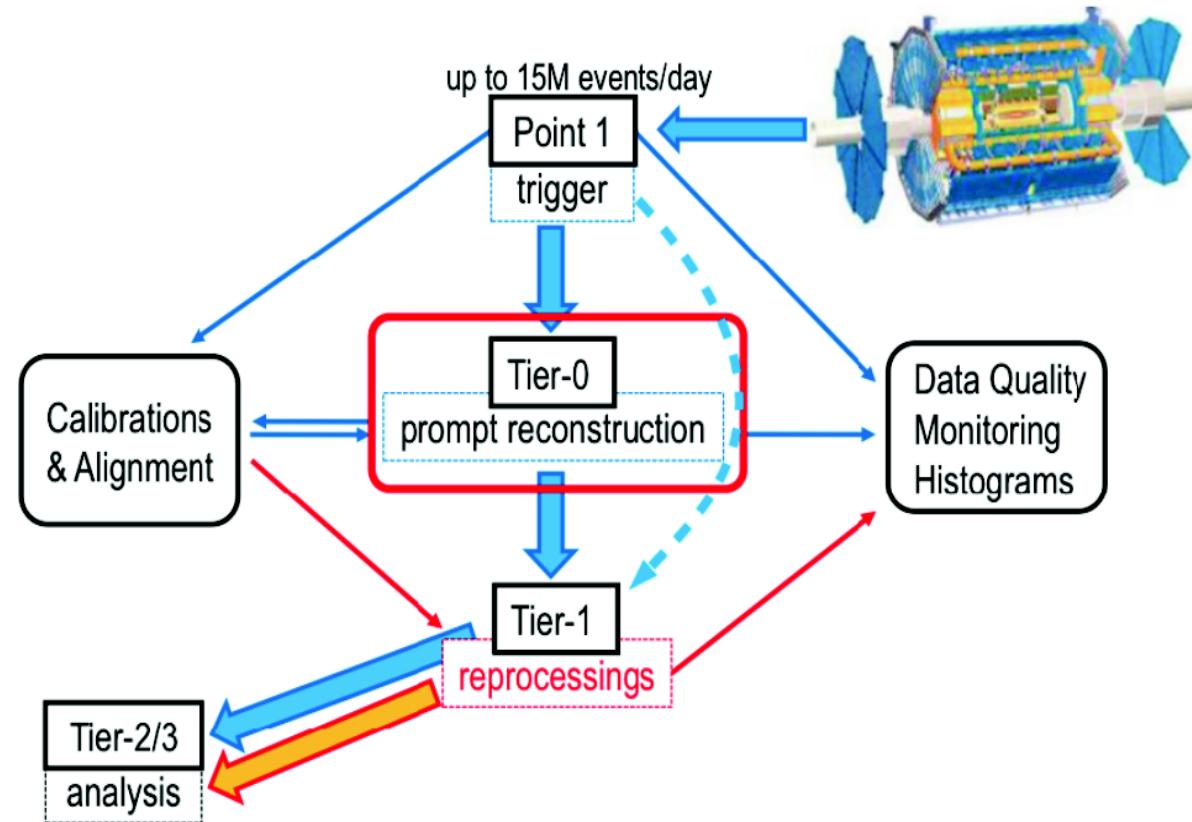


AMI (ATLAS Metadata Interface)

- **Metadata = Data of Data (Description of data)**
- Key of ATLAS Data Life Cycle Management
- On each step of data reconstruction in ATLAS, AMI Tags are generated
 - ATLAS data set, metadata link
 - Size and origin of dataset
 - File, number of events
 - Software parameter (AMI Tags)
 - MC parameter (PDF, generator, cross section, etc)
 - Lost files and Lumi blocks
 - Link to other applications (COMA, Rucio)
 - Data period
 - Luminosity, Trigger, Data of RAW data creation etc
 - Special Interface
 - AMI-Tags
 - Software Configuration Parameters
 - AMI-Glance
 - Data + Publication
 - Definition of data periods
 - Definition of physics containers
 - Reprocessing campaigns
 - Event count comparator

ATLAS Dataset - 2

- Tier-0
 - RAW is produced
 - Prompt data reconstruction
- ProdSys2
 - MC + Reprocessing
 - PanDA, JEDI, DEFT
- Distribution of dataset
 - Rucio



ATLAS Dataset - 3

- Data reconstruction → Example : x353 = AMI tag
 - `Reco_tf.py —AMI=x353 --inputBSFile=tier0_RawData.data`
- Reprocessing Campaign
 - After improvement of ATLAS software or framework, often huge reconstruction jobs can run
 - The campaign require many computing resources
 - As a result, many different versions of AOD (Analysis Object Data)
 - Derivation framework: AOD → Derived AOD
 - About 1% size of original AOD
 - Selection of particular event, parameter etc
 - Different configurations, software and conditions

ATLAS Dataset - 4

- Dataset = Collection of files
 - Collision data (data) and Monte Carlo (mc)

Data:

project tag:
2012 pp data
8 TeV Run number stream merged files Data type:
data12_8TeV.00209980.physics_Egamma.merge.AOD.f476_m1223
AMI tag describes configuration of
each step (Tier-0 bulk reconstruction **f**,
file merging **m**)

Simulation:

project tag: MC DSID
"mc12" setup unique #
8 TeV for process "human-readable" description of MC sample merged files Data type:
mc12_8TeV.119353.MadGraphPythia_AUET2BCTEQ6L1_ttbarW.merge.NTUP_SMWZ.
e1352_s1499_s1504_r3658_r3549_p1328/
AMI tag describes configuration of
each step (evt generation **e**, full simulation **s**,
reconstruction **r**, D3PD creation **p**)
/: is a "container" (points to other datasets)

AMI WebUI - 1

Datasets / Dataset Browser

Search Form 1: data12_001-real_data

1 dataset 1075 records

Query :dataset.amiStatus='VALID' AND (dataset.dataType like 'AOD') AND (dataset.streamName like 'physics_MinBias')

more fields +	logicalDatasetName ▾ Q	nFiles ▾ Q	totalEvents ▾ Q	totalSize ▾ Q	runNumber ▾ Q	period ▾ Q
details	data12_8TeV.00200804.physics_MinBias.merge.AOD.r4644_p1517 DQ2 - Provenance - GANGA export	18 182	585918 11645642	68.699 GB	200804 COMA Report - Periods - Run_Summary - Run_Query - DAQ_Config	A1 more Info - COMA - All Runs
details	data12_8TeV.00200805.physics_MinBias.merge.AOD.r4644_p1517 DQ2 - Provenance - GANGA export	2	54277	1.881 GB	200805 COMA Report - Periods - Run_Summary - Run_Query - DAQ_Config	A2 more Info - COMA - All Runs
details	data12_8TeV.00200841.physics_MinBias.merge.AOD.r4644_p1517 DQ2 - Provenance - GANGA export	4	186265	20.177 GB	200841 COMA Report - Periods - Run_Summary - Run_Query - DAQ_Config	A3 more Info - COMA - All Runs
details	data12_8TeV.00200842.physics_MinBias.merge.AOD.r4644_p1517 DQ2 - Provenance - GANGA export	4	194266	17.213 GB	200842 COMA Report - Periods - Run_Summary - Run_Query - DAQ_Config	A3 more Info - COMA - All Runs
details	data12_8TeV.00200863.physics_MinBias.merge.AOD.r4644_p1517 DQ2 - Provenance - GANGA export	4	99639	15.288 GB	200863 COMA Report - Periods - Run_Summary - Run_Query - DAQ_Config	A3 more Info - COMA - All Runs
details	data12_8TeV.00200913.physics_MinBias.merge.AOD.r4644_p1517 DQ2 - Provenance - GANGA export	4	120898	16.403 GB	200913 COMA Report - Periods - Run_Summary - Run_Query - DAQ_Config	A3 more Info - COMA - All Runs
details	data12_8TeV.00200928.physics_MinBias.merge.AOD.r4644_p1517 DQ2 - Provenance - GANGA export	2	47551	7.145 GB	200928 COMA Report - Periods - Run_Summary - Run_Query - DAQ_Config	A4 more Info - COMA - All Runs
details	data12_8TeV.00200965.physics_MinBias.merge.AOD.r4644_p1517 DQ2 - Provenance - GANGA export	8	290279	36.652 GB	200965 COMA Report - Periods - Run_Summary - Run_Query - DAQ_Config	A4 more Info - COMA - All Runs
details	data12_8TeV.00200982.physics_MinBias.merge.AOD.r4644_p1517 DQ2 - Provenance - GANGA export	2	40071	7.092 GB	200982	A4

1: number of results
2: default order, more recent first

3: query clauses
4: +/- fields
5,6: filter,calculator

7: conversion of units
8: group by, order by tools

AMI WebUI - 2

Element's information		Children elements	
logicalDatasetName	mc14_8TeV.129173.Pythia8_AU2CTEQ6L1_gammajet_DP140.merge.AOD.e1146_s1896_s1912_r5591_r5625 RucioInfo Provenance - Campaigns - GANGA export - Series	dataset_extra	4 Records
physicistResponsible	c.gwenlan1@physics.ox.ac.uk	dataset_keywords	5 Records
nFiles	200	dataset_comment	No records found
totalEvents	999500	files	3 200 Records
totalSize	672.801 GB	jobOptions	No records found
dataType	AOD	prodsys_task	1 Records
prodsysStatus	ALL EVENTS AVAILABLE	field	approx_crossSection
ECMEnergy	8000	value	5 1.2217E+02
physicsComment		field	approx_GenFitEff
PDF	CTEQ6L1 - LO with LO alpha_s	value	9.6932E-04
version	e1146_s1896_s1912_r5591_r5625 Datasets - Config_Tag	field	autoConfiguration
AtlasRelease	19.0.3	value	['everything']
crossSection	122.170 nb 2 Report an error - Jira issues	field	postInclude
Trans	1: provenance & rucio 2: JIRA link for X section pbs	value	[RecJobTransforms/UseFrontier.py]
data			

1: provenance & rucio
2: JIRA link for X section pbs

3: click for list of files
4: detail of prodsys task
5: cross section

Rucio (ATLAS Data Management System)

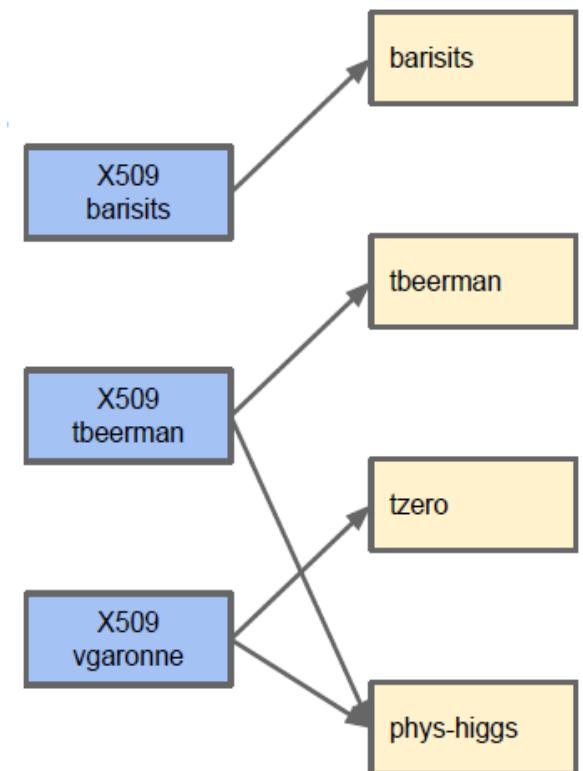


Basic concepts - 1

- Data Management System for Run-2 in ATLAS distributed computing system
 - Used to download outputs of Grid jobs, moving data and searching for them
- Rucio CLI tools from CVMFS
 - Web interface provides similar functionality

Basic concepts - 2

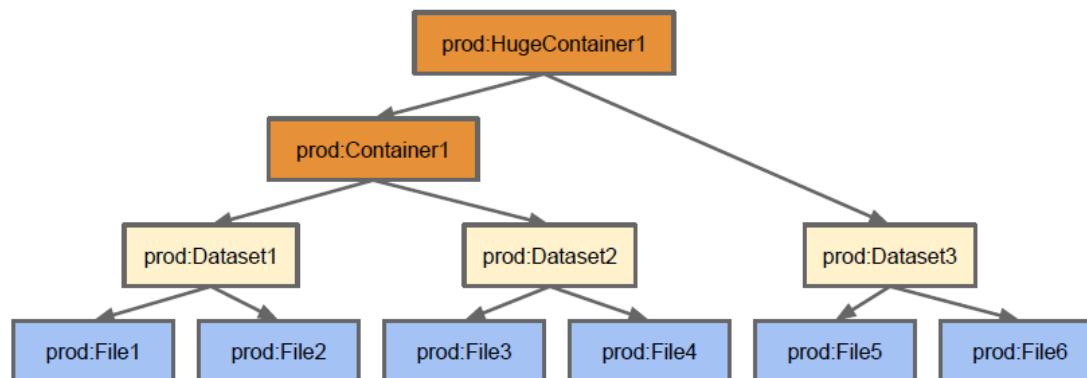
- Accounts
 - A Rucio Account can represent users, groups (phys-susy) or activities (panda, tzero)
 - Quota and permissions are associated to an account
 - One can connect to a Rucio account using X509 certificates/proxies, Kerberos
 - One credential can be used to map to different accounts.



Basic concepts - 3

- Rucio namespace

- 3 types of Data Identifiers (DIDs): File, Datasets, Containers.
 - Dataset: Collection of Files
 - Containers: Collection of Datasets or Containers
- The namespace is divided using scopes. A name is unique within a scope but can be used in other scopes. A DID is identified by a scope and a name



List DIDs

- A data Identifier is found by name and scope

```
## To list all DIDs within a scope
```

```
rucio list-dids user.gkawamur:*
```

SCOPE:NAME	[DID TYPE]
user.gkawamur:user.gkawamur.pruntest_7168.log	CONTAINER
user.gkawamur:user.gkawamur.tut.helloworldathena.test.log	CONTAINER
user.gkawamur:user.gkawamur.20160323144306.1.log	CONTAINER
user.gkawamur:user.gkawamur.tutorial.grid.deriv.test1.log	CONTAINER
user.gkawamur:user.gkawamur.pruntest_14222.log	CONTAINER

```
## Using filters (search)
```

```
rucio list-dids data16_13TeV:* --filter type=DATASET,datatype=AOD
```

SCOPE:NAME	[DID TYPE]	
data16_13TeV:data16_13TeV.00293572.physics_CosmicCalo.merge.AOD.x387_m1554		DATASET
data16_13TeV:data16_13TeV.00297447.physics_Standby.merge.AOD.x416_m1583		DATASET
data16_13TeV:data16_13TeV.00297041.physics_Main.merge.AOD.f686_m1583		DATASET
data16_13TeV:data16_13TeV.00297041.physics_CosmicCalo.merge.AOD.f686_m1583		DATASET
data16_13TeV:data16_13TeV.00297041.physics_ZeroBias.merge.AOD.f686_m1583		DATASET

Show metadata

```
## Showing metadata of a dataset
```

```
rucio get-metadata data16_13TeV:data16_13TeV.00303819.physics_Main.merge.AOD.f716_m1620
```

```
purge_replicas: None
campaign: None
is_new: None
is_open: False
closed_at: 2016-07-18 07:39:44
deleted_at: None
availability: AVAILABLE
eol_at: None
guid: None
panda_id: None
provenance: None
accessed_at: 2016-09-19 22:00:22
version: f716_m1620
scope: data16_13TeV
hidden: False
md5: None
events: 25475721
adler32: None
complete: None
lumiblocknr: None
monotonic: False
updated_at: 2016-09-19 22:00:42
obsolete: False
transient: None
did_type: DATASET
suppressed: True
expired_at: None
stream_name: physics_Main
account: tzero
run_number: 303819
name: data16_13TeV.00303819.physics_Main.merge.AOD.f716_m1620
task_id: None
datatype: AOD
created_at: 2016-07-17 04:17:53
bytes: 6462688464808
project: data16_13TeV
length: 2595
prod_step: merge
phys_group: None
```

RucioUI

- Move data between RSEs, use Rucio replication rules
- Such request can be generated by CLI and UI
- The UI tool is *the Rucio Rule Definition Droid (R2D2)*
<https://rucio-ui.cern.ch/r2d2>
- Basic need: X509 certificate in your browser



ATLAS Rucio UI Monitoring ▾ Data Transfers (R2D2) ▾ Reports ▾ pattern OR name OR rule id Search Using account: gkawamur ▾ Other Monitoring ▾ Help ▾

You are here: Rucio Rule Definition Droid - List Rules Rucio Version (WebUI / Server): 1.8.0 / 1.8.0

Rules

New request

Account	RSE	State	Activity	Interval
gkawamur	RSE		User Subscriptions	14 days

Apply

Show 10 entries

Search:

Name	Account	RSE Expression	Creation Date	State	Locks OK	Locks Replicating	Locks Stuck
No data available in table							

Name	Account	RSE Expression	Creation Date	State	Locks OK	Locks Replicating	Locks Stuck

Showing 0 to 0 of 0 entries

Previous Next

[delete rule](#)

[download as JSON](#)

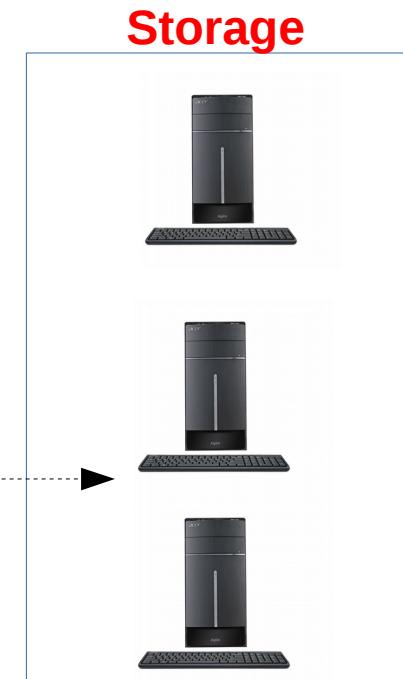
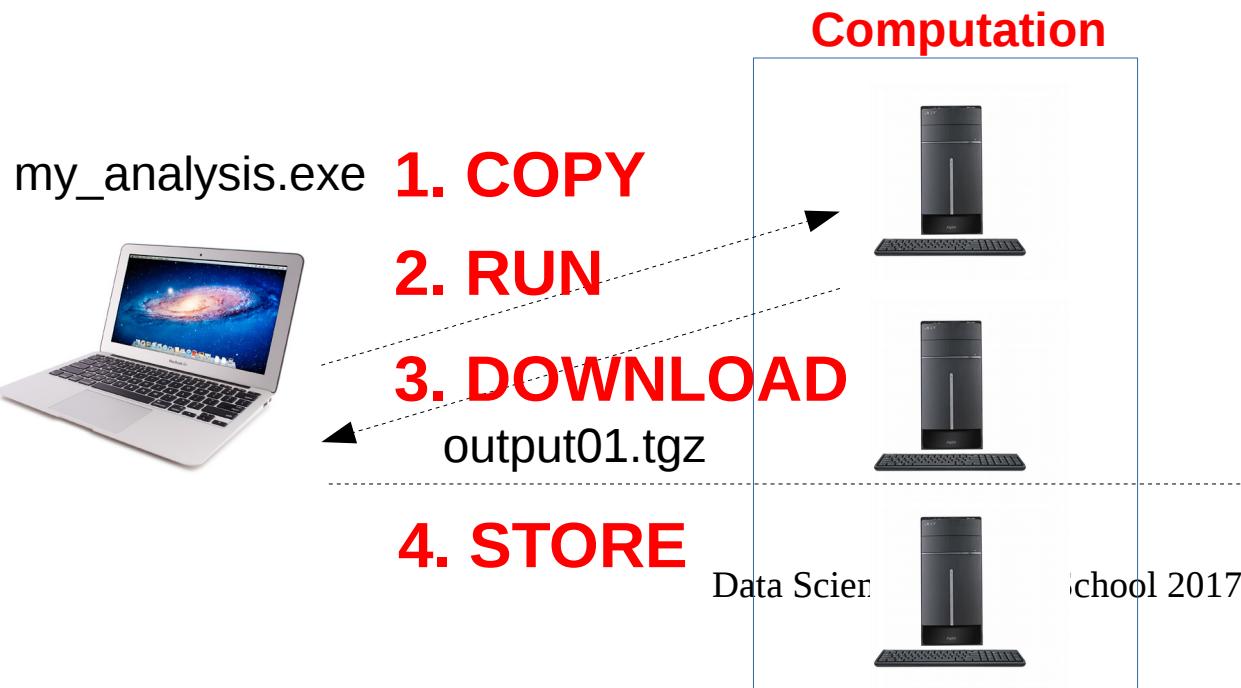
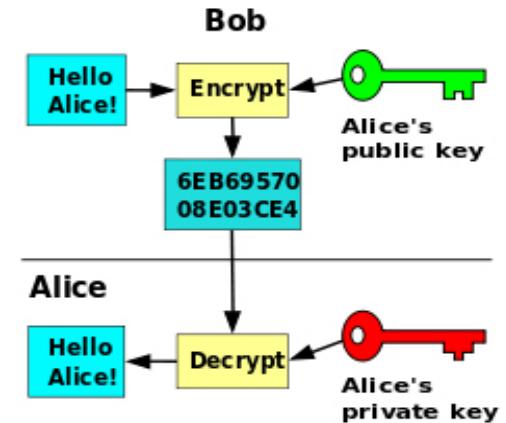
Introduction to World LHC Computing Grid



Concepts - 1

- How it works (*without Grid*)
 - Only PKI by openssl -> a simplest way

```
$ scp my_analysis.exe computer01:~/  
$ ssh computer01 “~/myprogram.exe run 01”  
$ scp computer01:output01.tgz ./output01.tgz  
$ scp output01.tgz storage02:output01.tgz
```



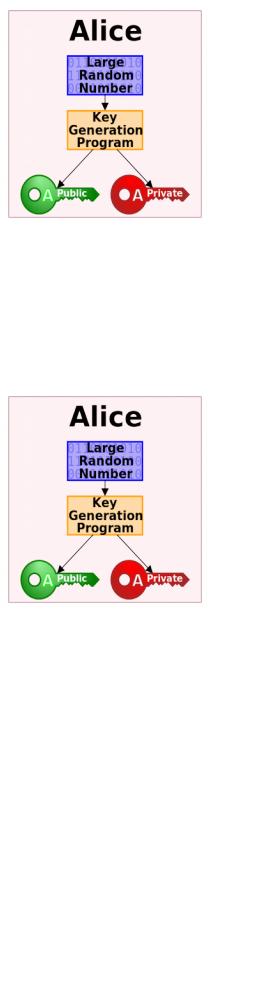
Concepts - 2

- How it works (*with Grid*)
 - Need of basic stuffs
 - Security policy based on PKI
 - Working with many users (physicists)
 - Recognition of users in affiliation, organization and experiment
 - Control of Job allocation
 - Control of storage space allocation and data replication
 - Control among geographically distributed computers
 - Interface of local computing infrastructures
 - Tree-like information search for computing clusters
 - E.g. like Domain Name Server
 - End user tools to make life easy



Concepts - 3

- Then, hierarchical signatures

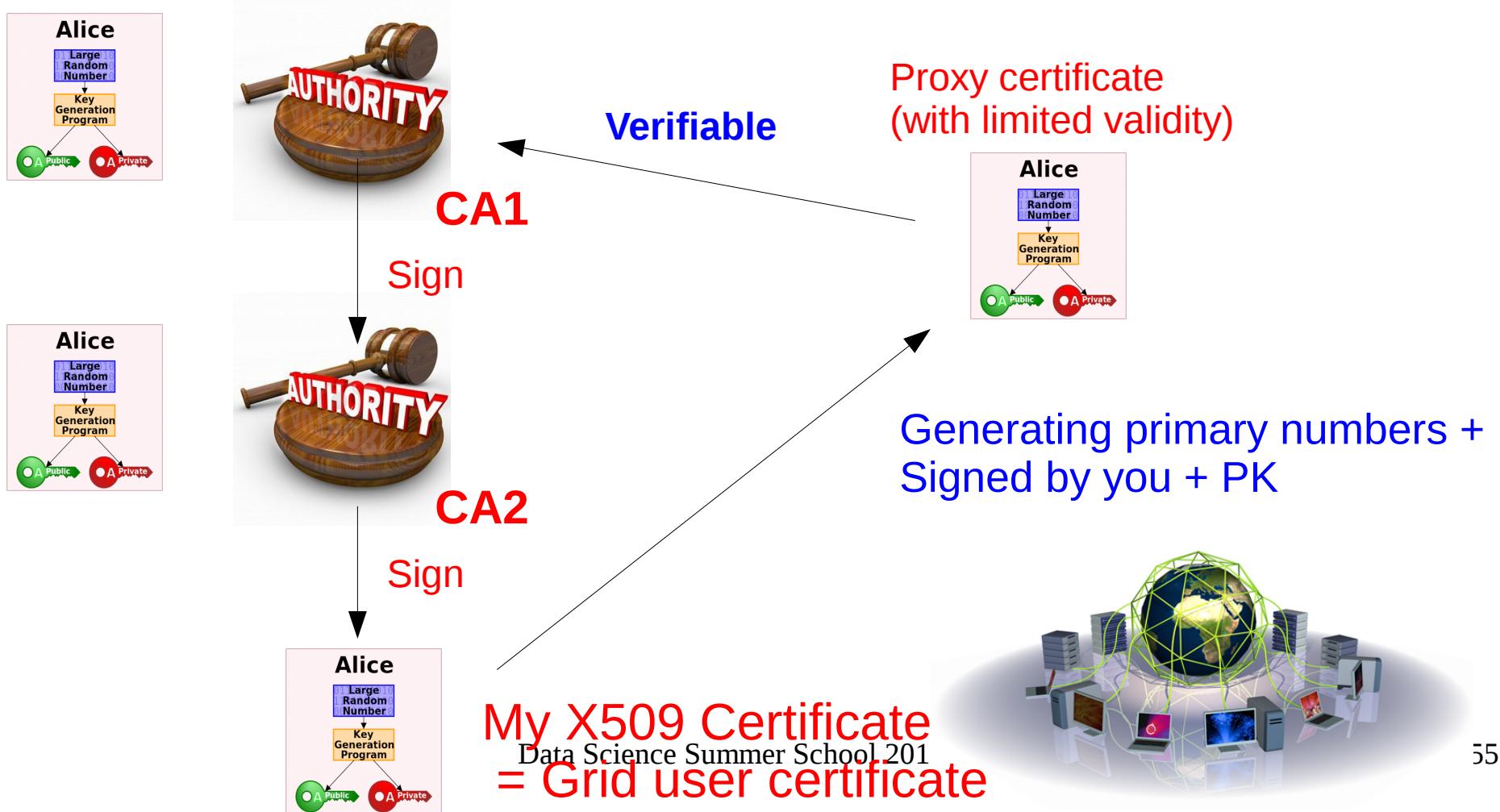


My X509 Certificate
Data Science Summer School 2011
= Grid user certificate



Concepts - 4

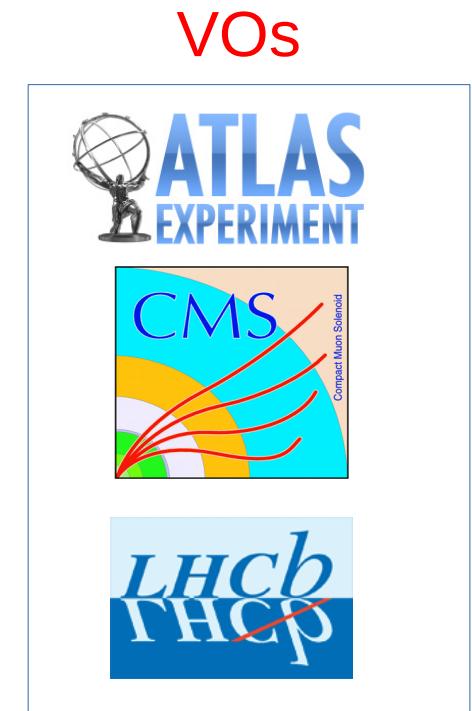
- Generating a new certificate = proxy certificate



Concepts - 5

- Virtual Organization (VO), VOMS
 - VOMS extends a certificate with a VO

Certificate Authority



Concepts - 6

- A similar analogy
 - Authenticated and authorized for your tasks

Certificate Authority



↓
Sign



User certificate

↓
Sign



Proxy certificate
= a limited copy of your certificate



Controlled



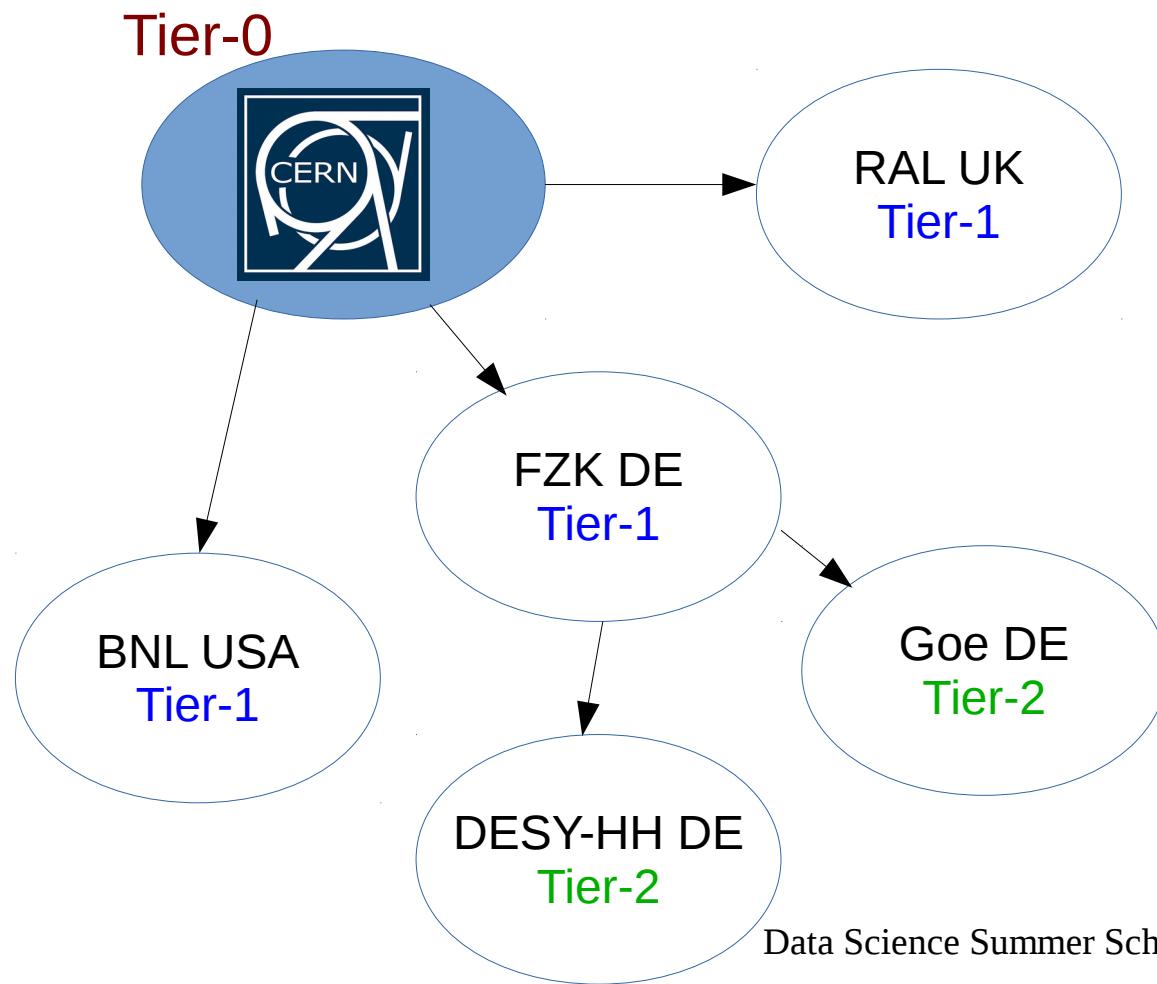
Authorization
= can enter a new land (Grid)



Authentication
= who you are?

ATLAS Grid job - 1

- LHC multi-tier structure
 - WLCG = Worldwide LHC Computing Grid



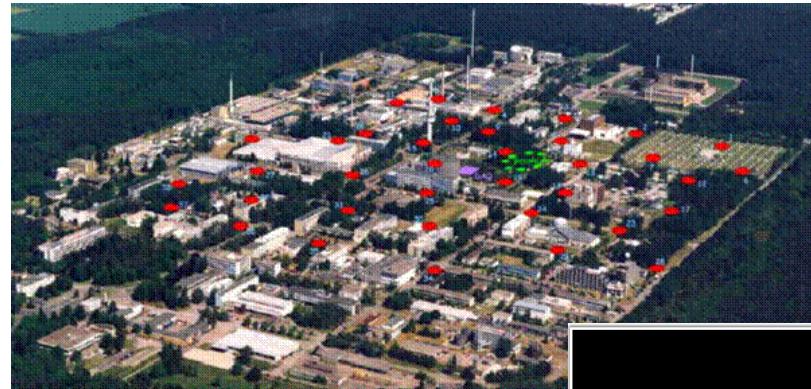
Tier-0: Raw data, Data store (in tape), Pre-processing, Reconstruction

Tier-1: National LCG-Centre, faster network connections and larger storage spaces (e.g. Tape), MC production, user analysis, etc.

Tier-2: University or Facility level computing sites. MC production, user analysis, etc.

ATLAS Grid job - 1

- FZK Tier-1



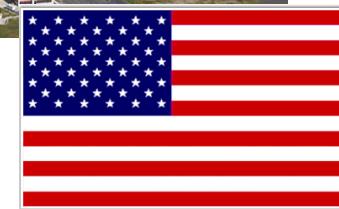
Forschungszentrum Karlsruhe
in der Helmholtz-Gemeinschaft



- BNL Tier-1



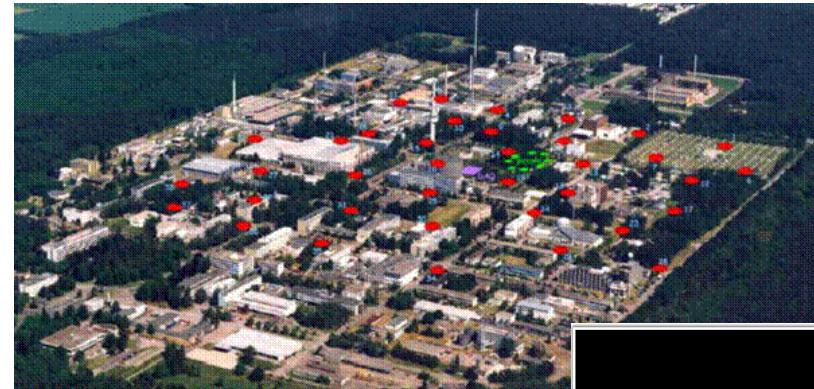
017



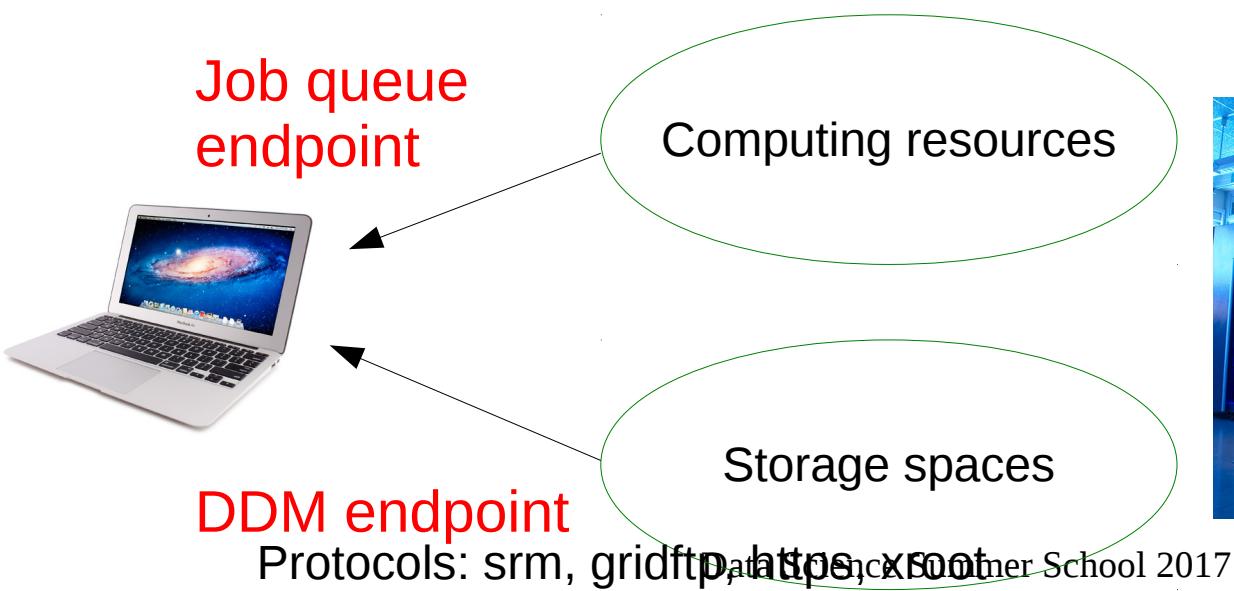
59

ATLAS Grid job - 1

- FZK Tier-1



Forschungszentrum Karlsruhe
in der Helmholtz-Gemeinschaft



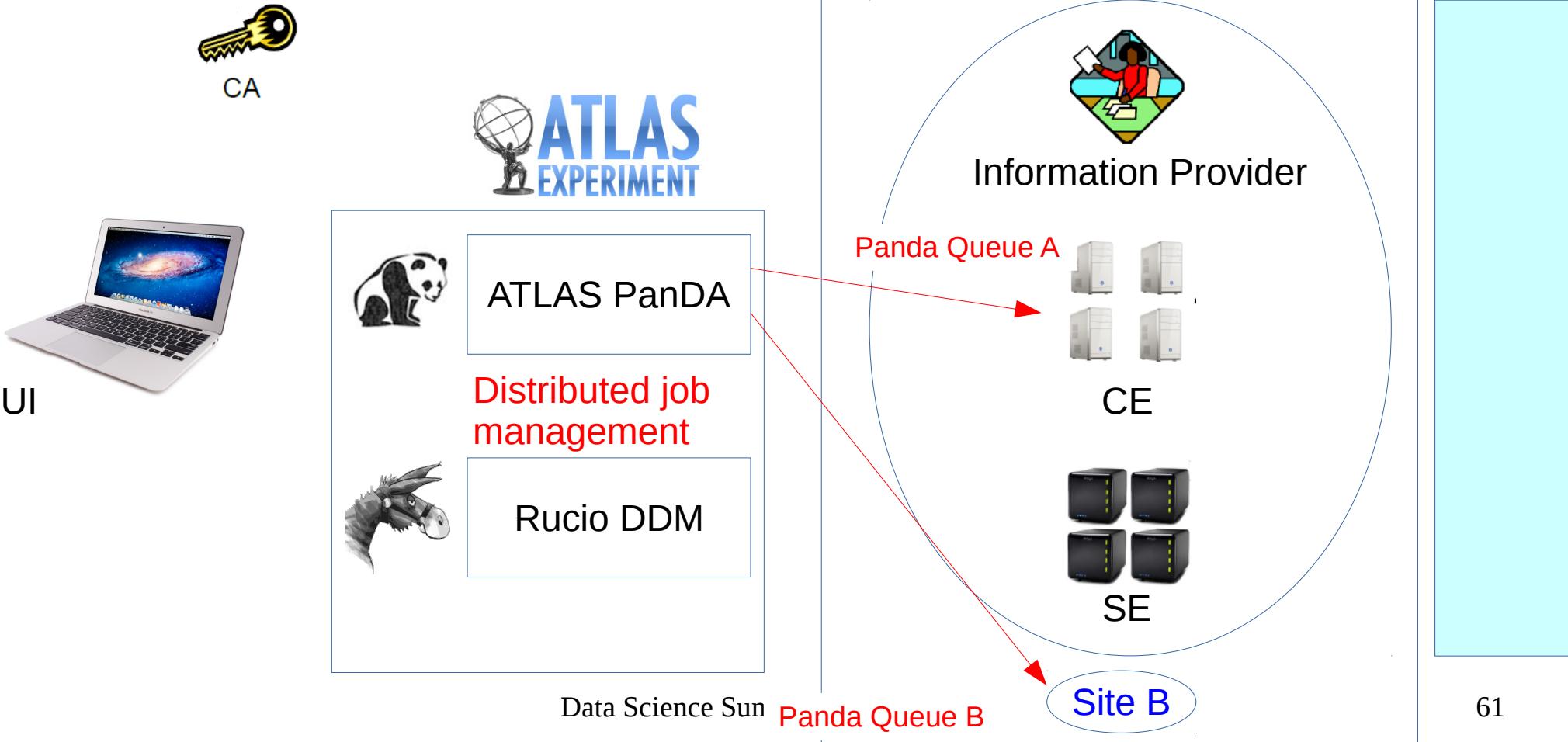
ATLAS Grid job - 2



WLCG
Worldwide LHC Computing Grid

Job
Status

- How do they work?



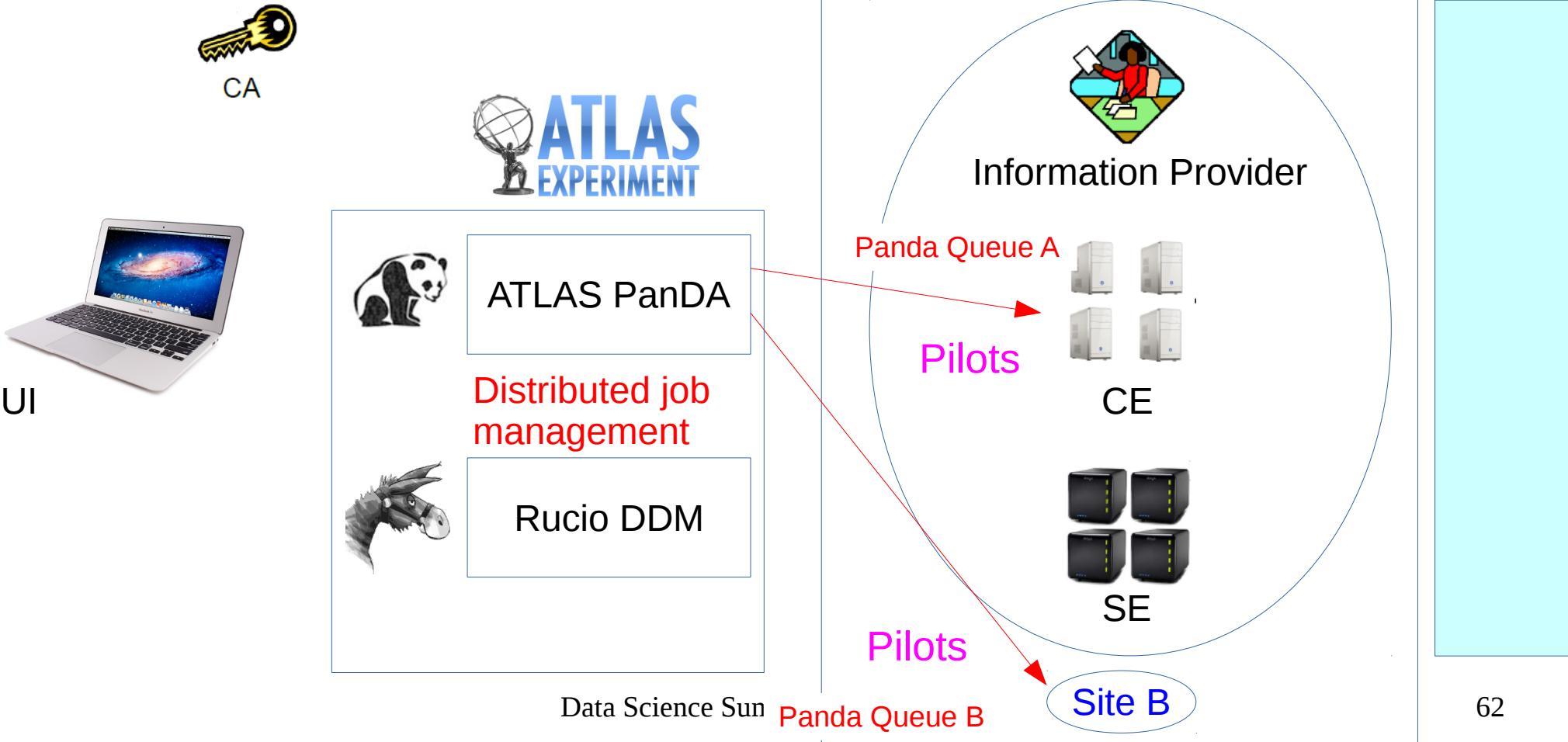
ATLAS Grid job - 3



WLCG
Worldwide LHC Computing Grid

Job
Status

- How do they work?

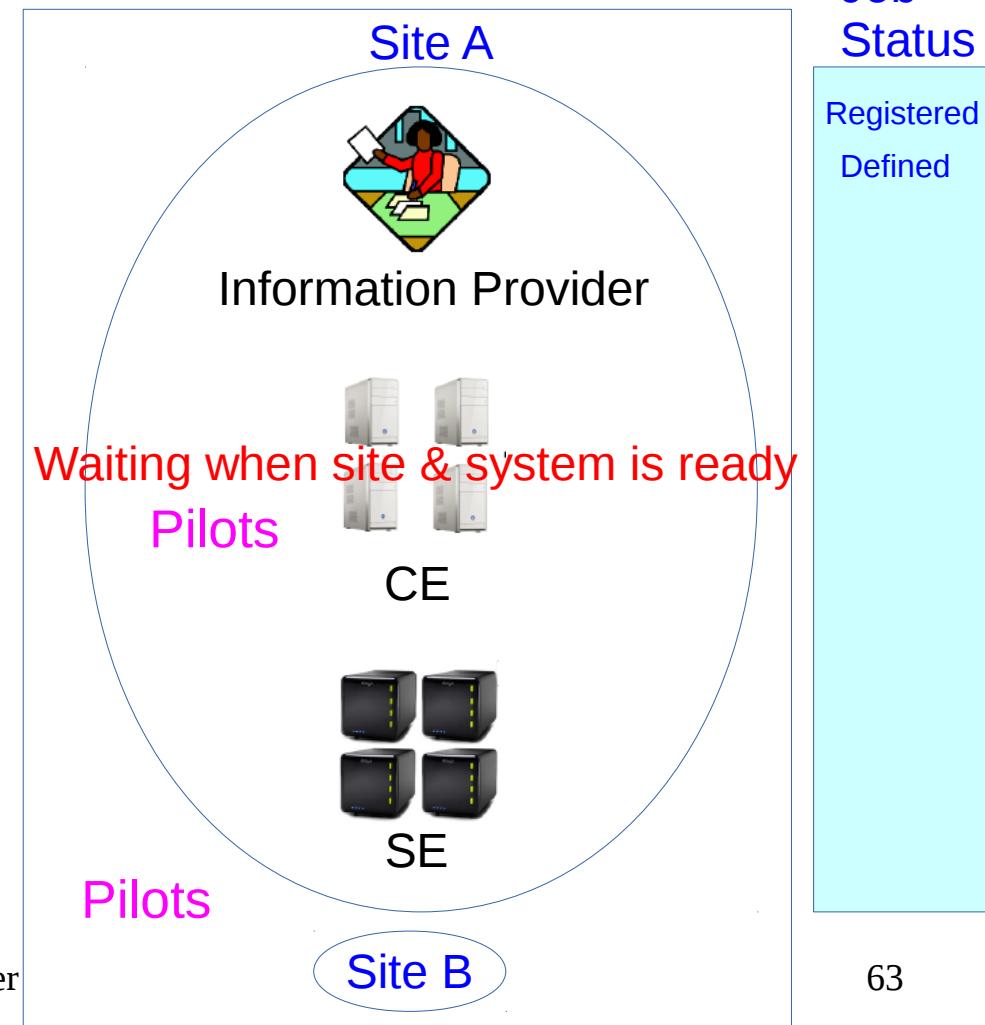
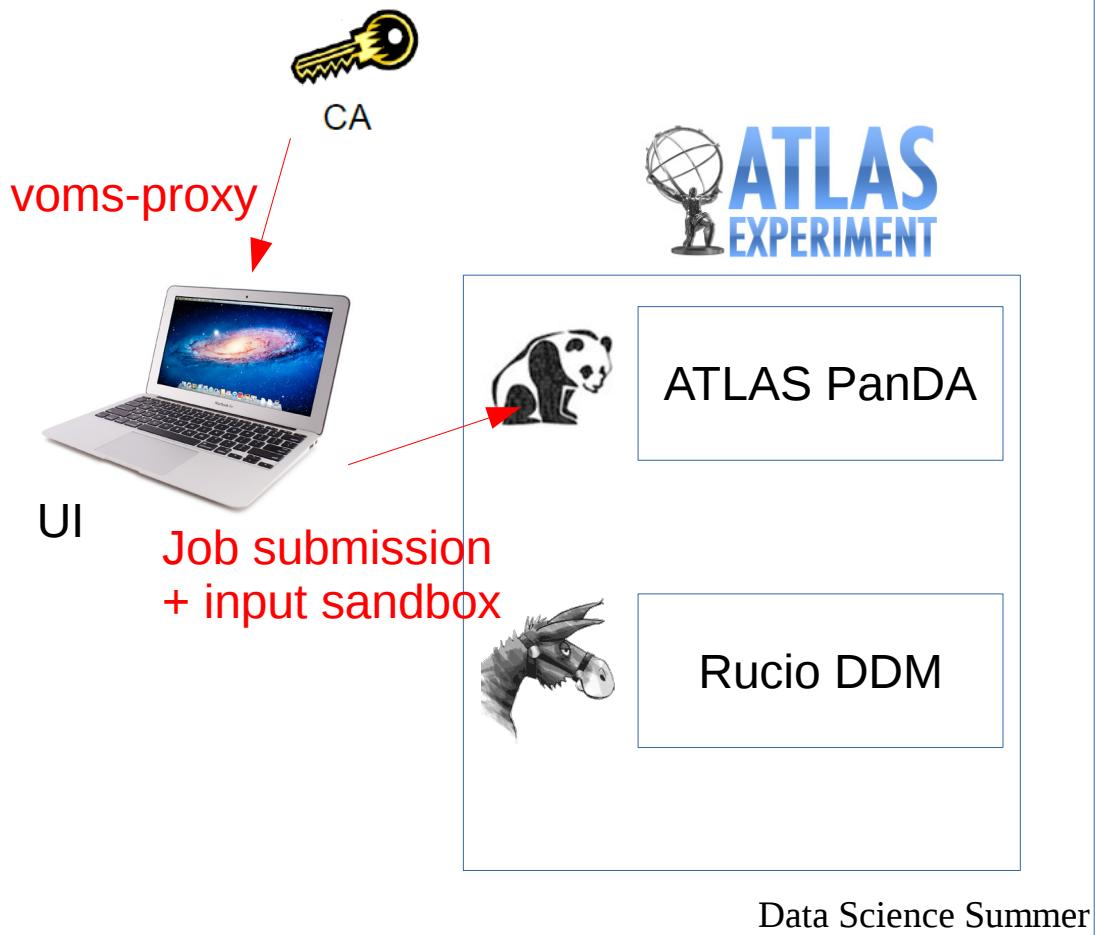


ATLAS Grid job - 4



WLCG
Worldwide LHC Computing Grid

- How do they work?

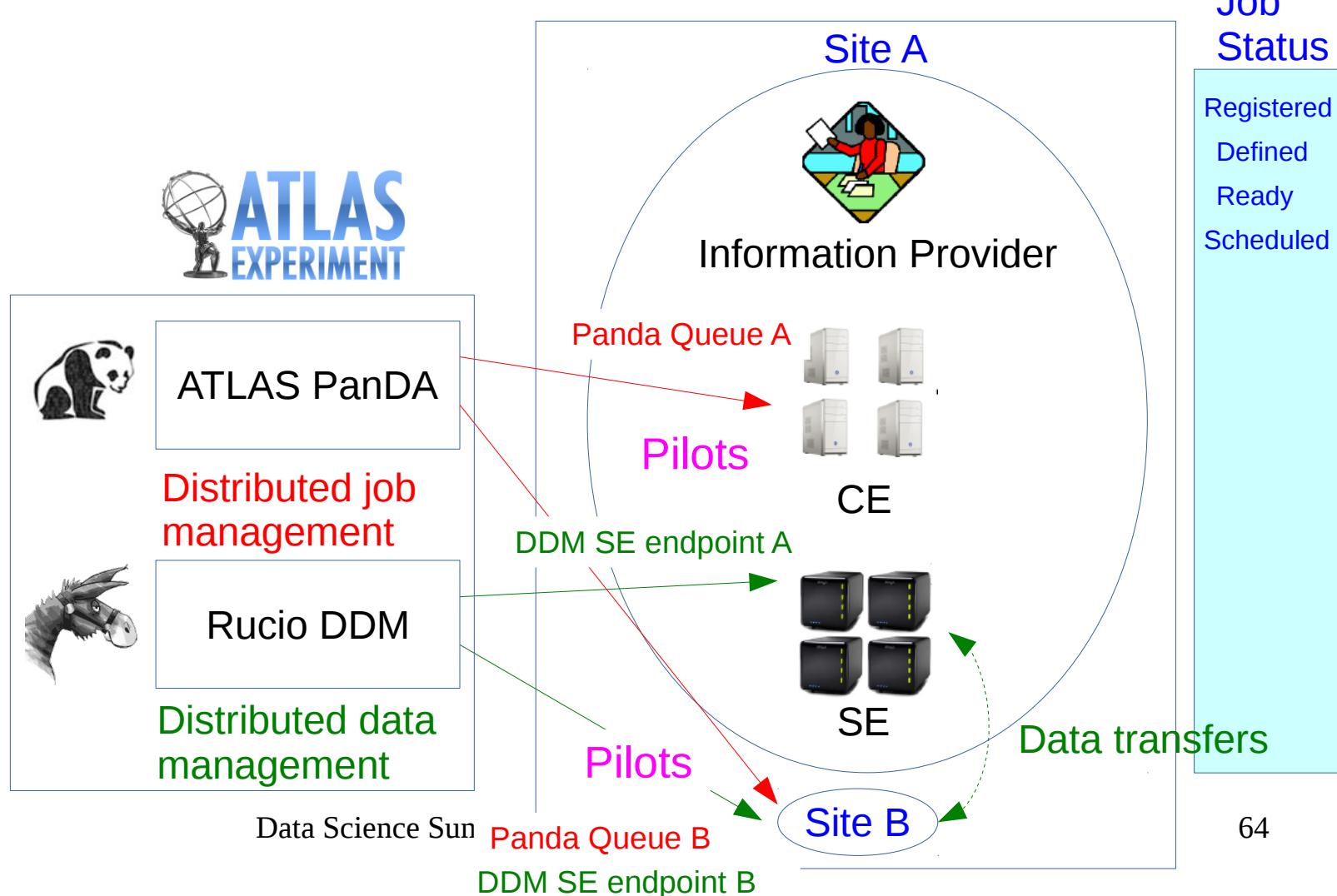


ATLAS Grid job - 5



WLCG
Worldwide LHC Computing Grid

- How do they work?

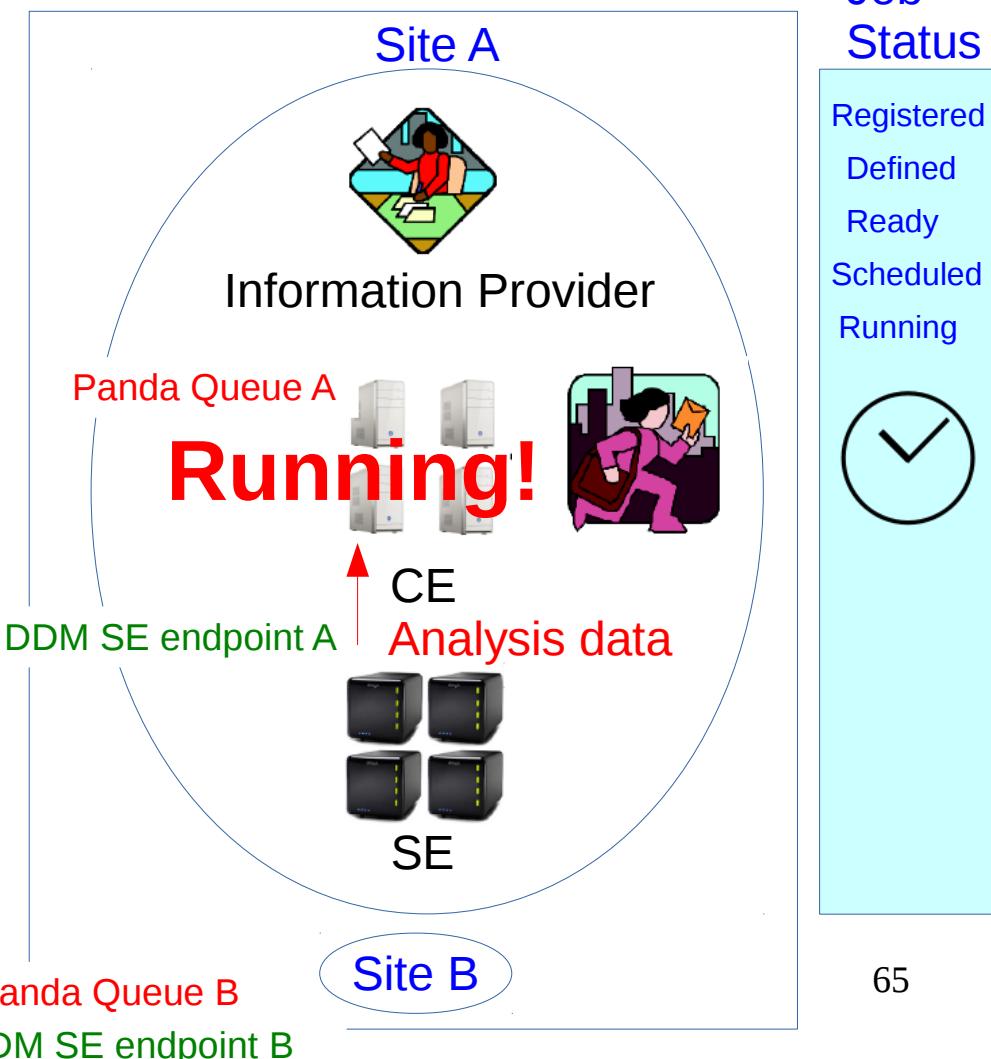
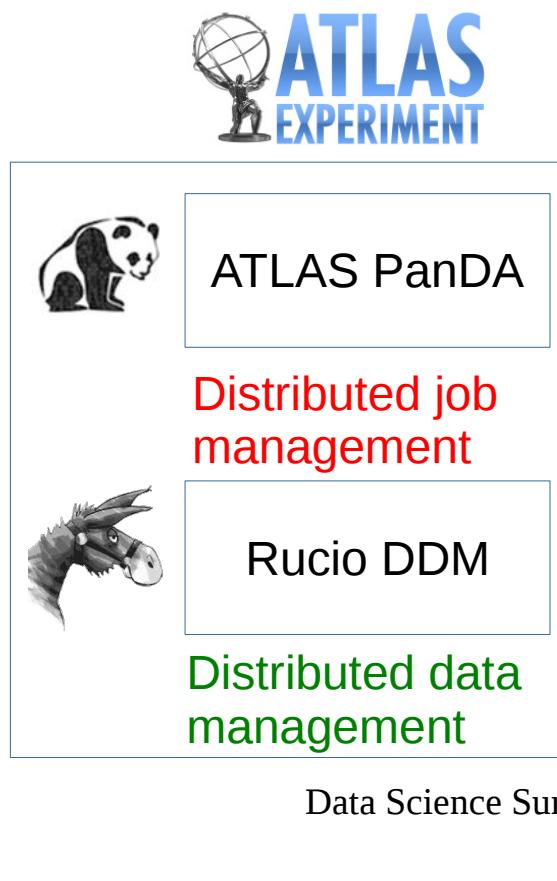


ATLAS Grid job - 6



WLCG
Worldwide LHC Computing Grid

- How do they work?

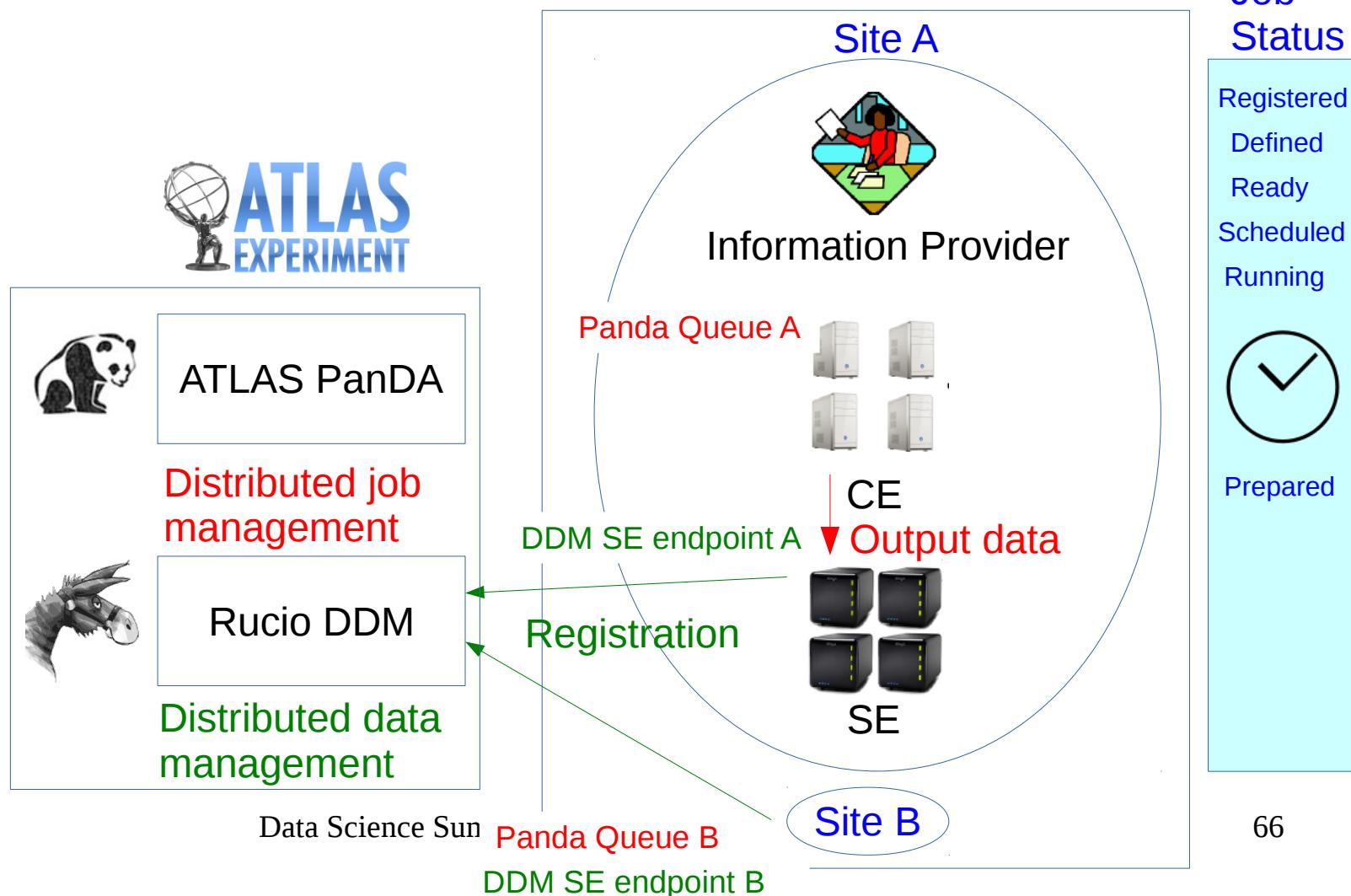


ATLAS Grid job - 7



WLCG
Worldwide LHC Computing Grid

- How do they work?

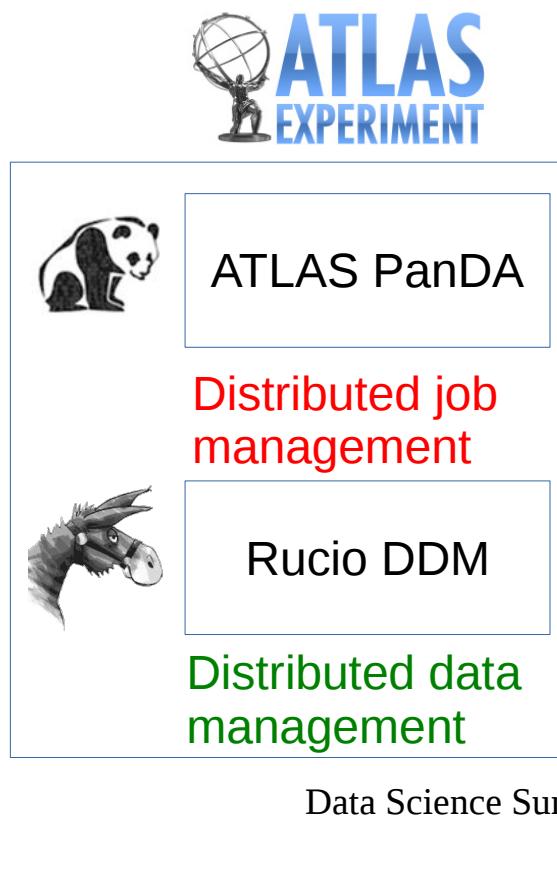


ATLAS Grid job - 8

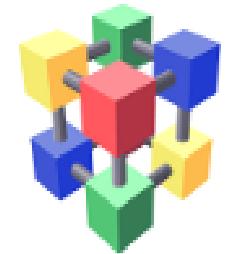


WLCG
Worldwide LHC Computing Grid

- How do they work?

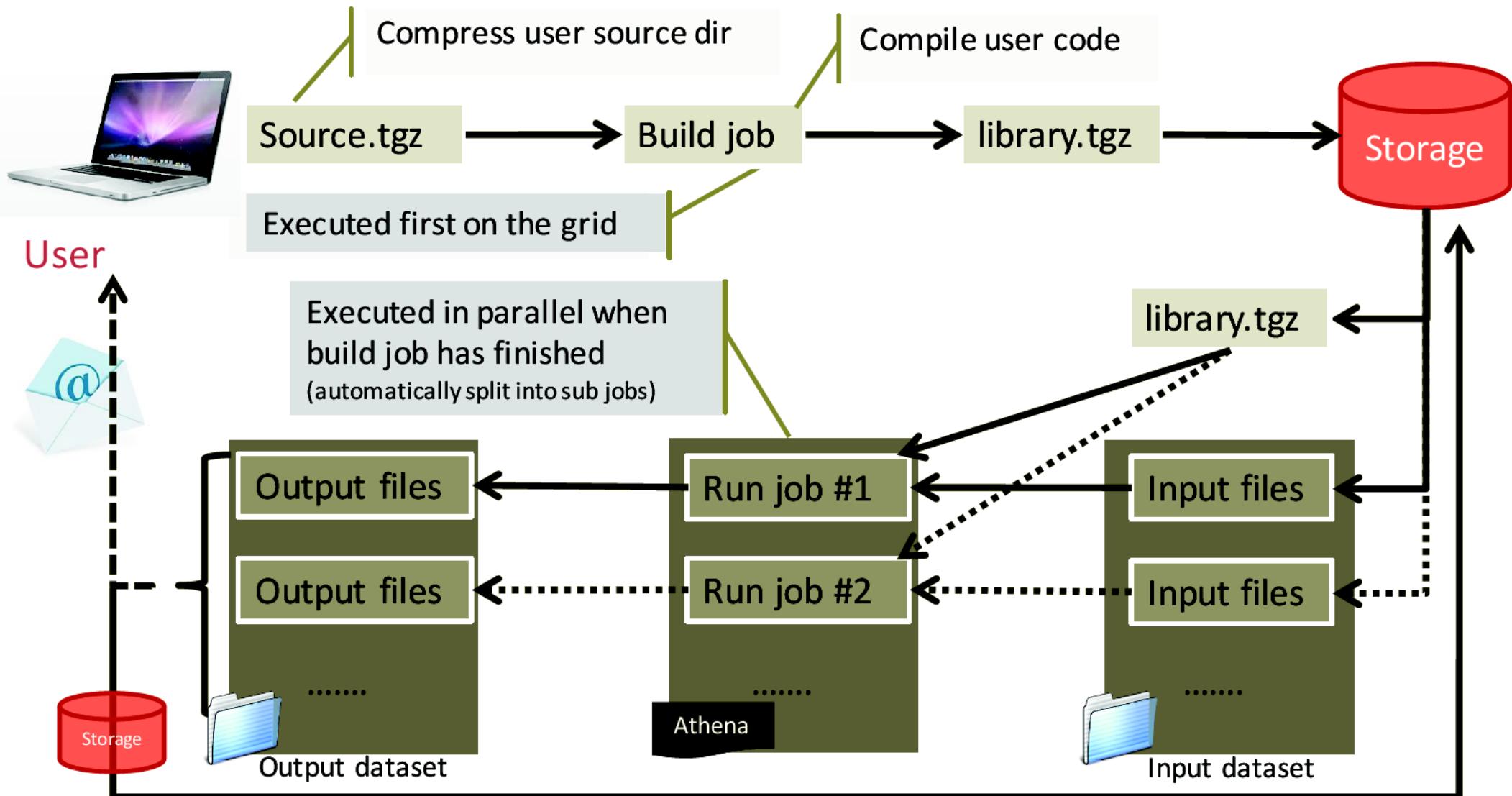


PanDA (ATLAS Job Management System)



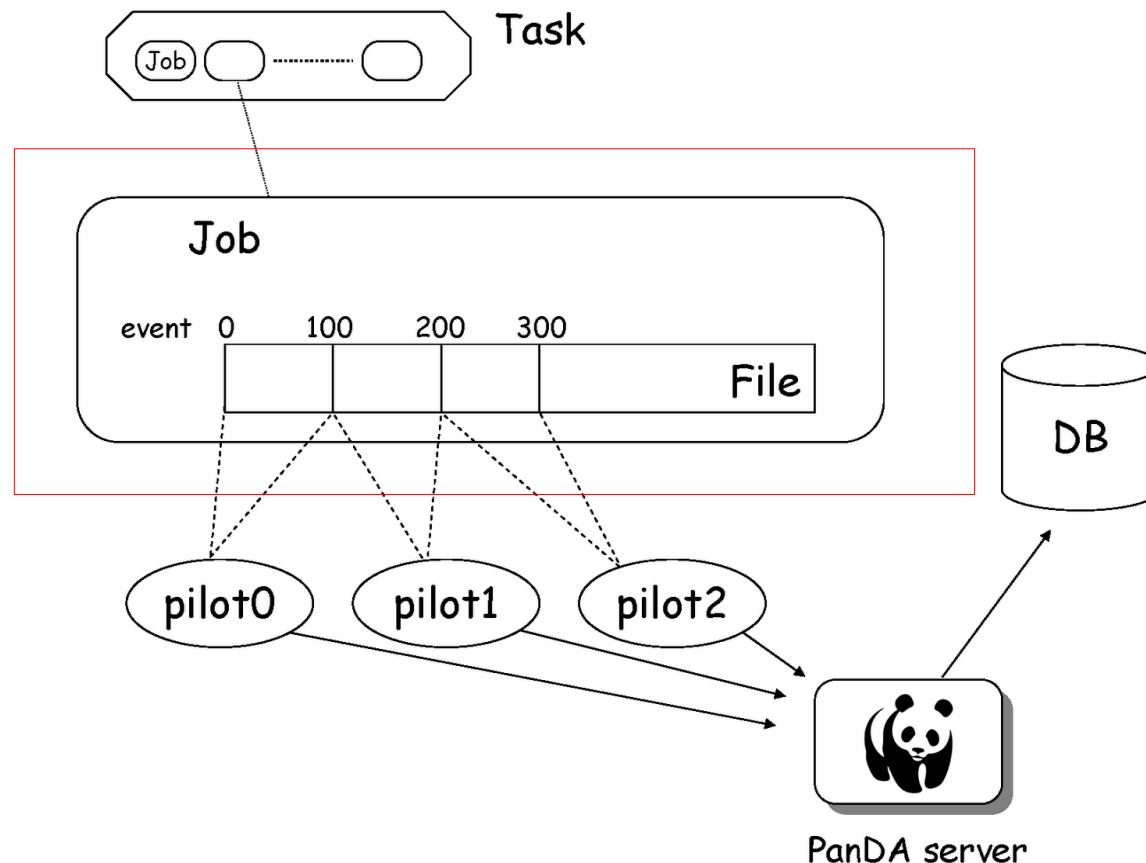
WLCG
Worldwide LHC Computing Grid

Build job & Run job



Event level processing

- Task (calculation) can be split by evenl level handler
 - Event Level Handler can generate basic pilot jobs according to events in given file(s)



Monitoring PanDA jobs (1/5)

Go to: <http://bigpanda.cern.ch> and enter your jediTTaskID

here →

The dashboard displays three stacked bar charts showing the number of slots of running jobs. The left chart covers 24 hours from 2014-10-06 09:00 to 2014-10-07 09:11 UTC. The middle chart covers 7 days from 2014-09-30 to 2014-10-07. The right chart covers 29 days from 2014-09-07 to 2014-10-07. The charts are color-coded by location: CERN (blue), NL (orange), EG (green), EU (red), and T0 (yellow). A legend at the bottom of each chart shows these colors.

Search

PanDA job ID or name Submit

Batch ID Submit

Task ID Submit

Task name Submit

News

- 20140923: Job pages directly link to object store based log tarballs
- 20140913: Link DEFT request page from task list and detail pages
- 20140912: On task page highlight input containers, hide datasets by default
- 20140911: Task name search supports wildcarding
- 20140911: Show wait time, duration for jobs not yet running, completed
- 20140818: Job attempt# off for user page, JEDI jobs. Not meaningful in JEDI.
- 20140818: Task attribute summary added to user page
- 20140817: Output container list added to task detail page
- 20140817: Support clarified. Use DAST list, as ever, for dist analysis support

Monitoring PanDA jobs (2/5)

Task is running

bigpanda.cern.ch/task/?jeditaskid=4212437&display_limit=200

ATLAS PanDA monitor Dashboards Tasks Jobs Errors Users Sites Incidents Search Prodsys Services VO Help Generated 2014-10-07 09:34 UTC

Task 4212437: user.nozturk.pruntest/

Task ID	Jobset	Type	WorkingGroup	User	Task status	Ninputfiles finished failed	Created	Modified	Cores	Priority	Parent
4212437	15168	anal		Nurcan Ozturk	running	1 0 (0%) 0 (0%)	2014-10-07 09:33	10-07 09:33	1	1000	

States of jobs in this task [Show jobs](#)

defined	waiting	pending	assigned	throttled	activated	sent	starting	running	holding	transferring	finished	failed	cancelled	merging
1								1						

Jump to [job parameters](#), [task parameters](#)

View: [job list \(access to job details and logs\)](#) [child tasks](#) [prodsys task page](#) [brokerage logger](#) [JEDI action logger](#) [error summary](#)

Output containers

[user.nozturk.pruntest.log/](#)

4 datasets, show/hide by type: [all](#) [lib\(1\)](#) [log\(1\)](#) [pseudo_input\(1\)](#) [tmp_log\(1\)](#)

Dataset, container name	Type	Stream	State	Status	Nfiles	Created	Modified

Job parameters

"

Monitoring PanDA jobs (3/5)

Task is done

Screenshot of the ATLAS PanDA monitor interface showing a completed task.

Task ID: 4212437 | **User:** Nurcan Ozturk | **Status:** done | **Ninputfiles:** 1 | **finished:** 1 (100%) | **failed:** 0 (0%)

States of jobs in this task: defined, waiting, pending, assigned, throttled, activated, sent, starting, running, holding, transferring, finished, failed, cancelled, merging. (The 'finished' state has a value of 2, indicated by a red arrow.)

Jump to job parameters, task parameters

View: job list (access to job details and logs) | child tasks | prodsys task page | brokerage logger | JEDI action logger | error summary

Output containers: user.nozturk.pruntest.log/ (indicated by a red arrow)

4 datasets, show/hide by type: all lib(1) log(1) pseudo_input(1) tmpl_log(1)

Dataset, container name	Type	Stream	State	Status	Nfiles	Created	Modified
-------------------------	------	--------	-------	--------	--------	---------	----------

Monitoring PanDA jobs (4/5)

Task details, list of jobs. Click on “PanDA ID” to go to the job details, job log files

click 

Job list Sort by Pandaid, ascending mod time, descending mod time, priority, attemptnr											
PanDA ID Attempt#	Owner Group	Task ID	Transformation	Status	Created	Time to start d:h:m:s	Duration d:h:m:s	Mod	Cloud Site	Priority	Job info
2280470784 Attempt 1	Nurcan Ozturk	4212437	runGen-00-00-02	finished	2014-10-07 09:33	0:06:28	0:0:01:04	10-07 09:49	ES ANALY_IFAE	1000	
	Job name: user.nozturk.pruntest/ #1 Datasets:										
2280470783 Attempt 0	Nurcan Ozturk	4212437	buildGen-00-00-01	finished	2014-10-07 09:33	0:00:16	0:0:01:04	10-07 09:40	ES ANALY_IFAE	2000	
	Job name: user.nozturk.pruntest/ #0 Datasets: Out: panda.1007093351.25596.lib._4212437										

 run job: runs the job at the grid site

 build job: recreates the athena environment at the grid site

Monitoring PanDA jobs (5/5)

How to find job log files

Screenshot of the ATLAS PanDA monitor interface showing job details for PanDA job 2280470784.

Job details for PanDA job 2280470784 (Generated 2014-10-07 10:52 UTC)

PandaID	Owner	TaskID	Status	Created	Time to start d:h:m:s	Duration d:h:m:s	Modified	Cloud Site	Priority
2280470784	Nurcan Ozturk	4212437	finished	2014-10-07 09:33	0:0:06:28	0:0:01:04	10-07 09:49	ES ANALY_IFAE	1000

Job name: [user.nozturk.pruntest/](#) type: panda-client-0.5.30-jedi-run transformation: runGen-00-00-02

Datasets:

Job information logfiles pilot job stdout, stderr, batch log logger records child jobs

Jobset 15168

check athena_stdout.txt (for pathena)/prun_stdout.txt (for prun) and pilotlog.txt (all commands issued by the pilot) from this link

5 job files						
Filename (Type)	Scope	Size (MB)	Status	Attempt (max)	Dataset	
panda.1007093351.25596.lib._4212437.241508507.lib.tgz (input)	panda	0	ready		panda.1007093351.25596.lib._4212437 (dispatch block: panda.1007093351.25596.lib._4212437)	
user.nozturk.pruntest.log.4212437.000001.log.tgz (log)	user.nozturk	0	finished		user.nozturk.pruntest.log.6779086	
user.nozturk.pruntest.log.4212437.000001.log.tgz (log)	user.nozturk	0	ready		user.nozturk.pruntest.log/ (destination block: sub0186192010)	
pseudo_lfn (pseudo_input)		0	finished		pseudo_dataset	
pseudo_lfn (pseudo_input)		0	unknown		pseudo_dataset	

Conclusion

- Successful deployments for huge collaborations and experiments in the real world
 - Infrastructure demonstrated to be able to support LHC data taking and analysis
- Distributed parallel computing system works among different institutes, batch systems and countries
- Significant operational infrastructure behind it
- Service Oriented Architecture: A reliable and robust service of many components need

Hands-on

Open the hands-on slide.

Use a grid proxy (Note: the certificate is valid
during the time slots of the hands-on)