

A Large Ion Collider Experiment



AliEn in ALICE

Predrag Buncic

Alice Environment @ Grid

User Interface		
VTD/OSG stack	AliEn stack	EDG stack



Nice! Now I do not have to worry about ever changing GRID environment...



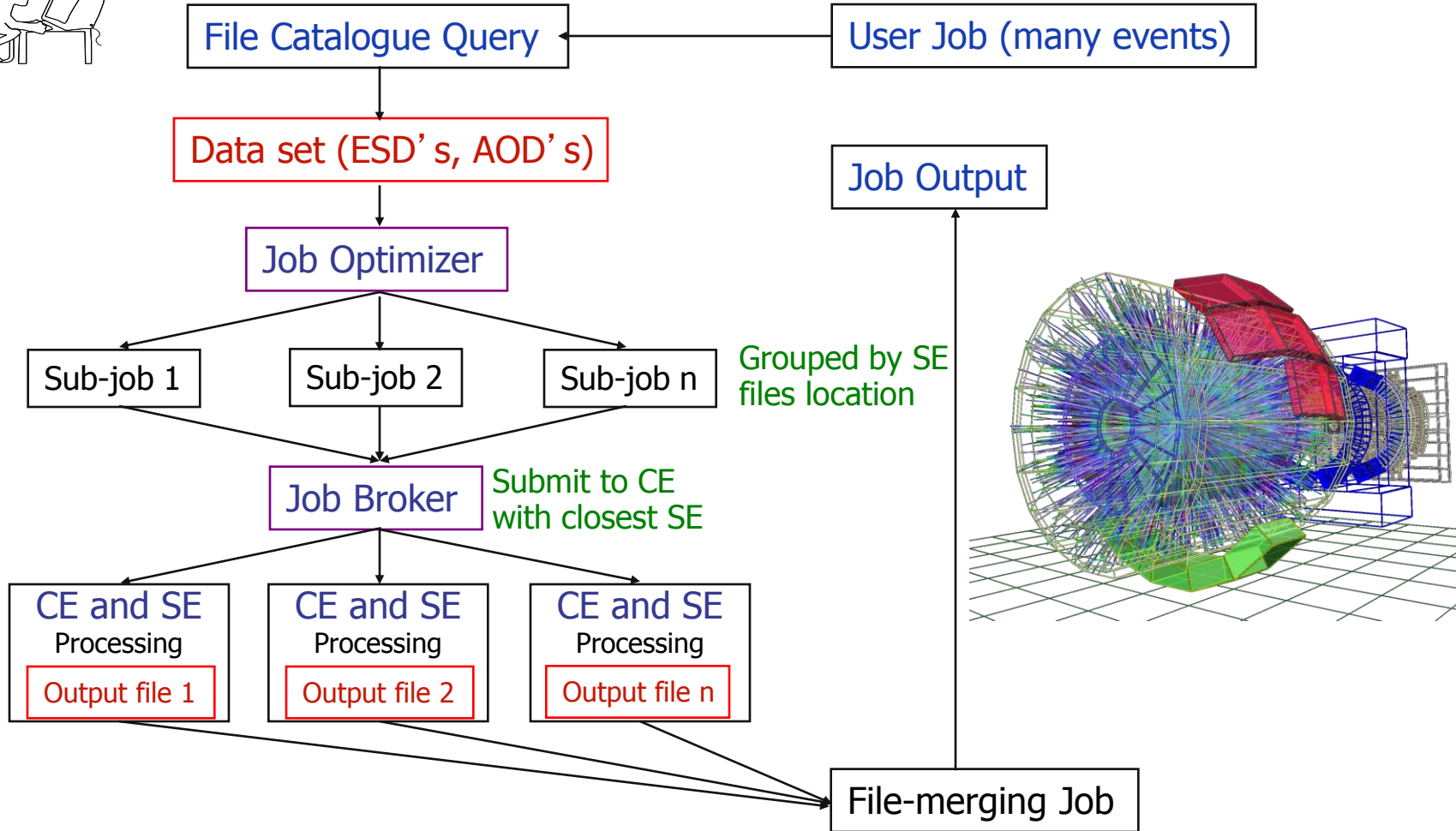
AliEn v1.0 (2001)

- New approach
 - Using standard protocols and widely used Open Source components
 - Interface to many Grids
- End-to-end solution
 - SOA (Service Oriented Architecture)
 - SOAP/Web Services (18)
 - Core Services (Brokers, Optimizers, etc)
 - Site Services
 - Package Manager
 - Other (non Web) Services (ldap, database proxy, posix I/O)
 - Distributed file and metadata catalogue
 - API and a set of user interfaces
- Used as production system for ALICE since the end of 2000

AliEn v2.0 (2007)

- New API Service and ROOT API
 - Shell, C++, perl, java bindings
- Analysis support
 - Batch and interactive
 - ROOT/PROOF interfaces
 - Complex XML datasets and tag files for event level metadata
 - Handling of complex workflows
- New (tactical) SE and POSIX I/O
 - Using xrootd protocol in place of aiod (glite I/O)
- Job Agent model
 - Improved job execution efficiency (late binding)

Distributed Analysis



ROOT / AliEn UI

```

X alientest@pcarda02:~
[pcarda02] /home/alientest > alien/api/bin/aliensh
[ aliensh 2.0.4 (C) ARDA/Alice: Andreas.Joachim.Peters@cern.ch/Derek.Feichtinger@cern.ch]
*****
* Welcome to the ALICE VO at alien://pcapiserv01.cern.ch:10000
* Running with Server V2.0.5
*****

*****
AliEn v.2-10 has been released.
*****
aliensh:[alice] [1] /alice/cern.ch/user/p/peters/macros/ >ls
.esdTree.C
.esdTree.h
.MyBatchAnalysis.C
esdAna.C
esdAna.h
esdTree.C
esdTree.h
MyBatchAnalysis.C
aliensh:[alice] [2] /alice/cern.ch/user/p/peters/macros/ >

```

```

X apiclient@pcapiserv01:~/root
root [12] TGrid::Connect("alien://");
=> Trying to connect to Server [0] http://pcapiserv01.cern.ch:9000 as User peters
*****
* Welcome to the ALICE VO at alien://pcapiserv01.cern.ch:9000
* API Service written by Derek Feichtinger/Andreas-J.Peters
* Running with Server V2.0.0
*****

root [13] TAlienCollection* collection = new TAlienCollection("/tmp/example1.xml");
root [14]

```

Workflow engine

Run 197692 processing details

<p>261753606</p> <p>LHC period LHC13g - CPass0 (reconstruction) (OCDB)</p> <p>7.933 KB</p>	<p>261802598</p> <p>LHC period LHC13g - CPass0 (reconstruction)</p> <p>168.3 MB</p> <p>48127 ev.</p>	<p>262047448</p> <p>LHC period LHC13g - CPass0 (merging+OCDB)</p> <p>2.055 MB</p>	<p>262055695</p> <p>LHC period LHC13g - CPass1 (reconstruction) (OCDB)</p> <p>7.944 KB</p>	<p>262063914</p> <p>LHC period LHC13g - CPass1 (reconstruction)</p> <p>6.685 GB</p> <p>49043 ev.</p>	<p>262102834</p> <p>QA_LHC13g_Stage1: Intermediate merging stage</p> <p>706.4 MB</p>	<p>262108278</p> <p>QA_LHC13g_Stage2: Intermediate merging stage</p> <p>110.5 MB</p>	<p>262114662</p> <p>QA_LHC13g_Stage5: Final merging stage</p> <p>45.48 MB</p> <p>49043 ev.</p>
<p>262029844</p> <p>LHC period LHC13g - reco for Muon + Calorimeters (OCDB)</p> <p>7.366 KB</p>	<p>262033399</p> <p>LHC period LHC13g - reco for Muon + Calorimeters</p> <p>9.664 GB</p> <p>380499 ev.</p>	<p>262066247</p> <p>QA_LHC13g_Stage1: Intermediate merging stage</p> <p>439.4 MB</p>	<p>262084300</p> <p>QA_LHC13g_Stage2: Intermediate merging stage</p> <p>71.41 MB</p>	<p>262094107</p> <p>QA_LHC13g_Stage5: Final merging stage</p> <p>28.53 MB</p> <p>353859 ev.</p>			
<p>301912337</p> <p>LHC period LHC13g - VPass1 (OCDB)</p> <p>8.294 KB</p> <p>56603 ev.</p>	<p>301914351</p> <p>LHC period LHC13g - VPass1</p> <p>2.745 GB</p> <p>62414 ev.</p>	<p>301936110</p> <p>QA_LHC13g_Stage1: Intermediate merging stage</p> <p>166.4 MB</p>	<p>301938227</p> <p>QA_LHC13g_Stage5: Final merging stage</p> <p>26.48 MB</p> <p>56603 ev.</p>				
		<p>262098145</p> <p>AODmerge_LHC13g: AOD merging</p> <p>2.341 GB</p> <p>380499 ev.</p>					
		<p>262098146</p> <p>Merged tags creation</p> <p>19.84 KB</p>					
		<p>301956354</p> <p>AODmerge_LHC13g: AOD merging</p> <p>1.125 GB</p> <p>62414 ev.</p>					
		<p>303143829</p> <p>Merged tags creation</p> <p>1.105 KB</p>					

Analysis trains

The LEGO framework



Usage statistics

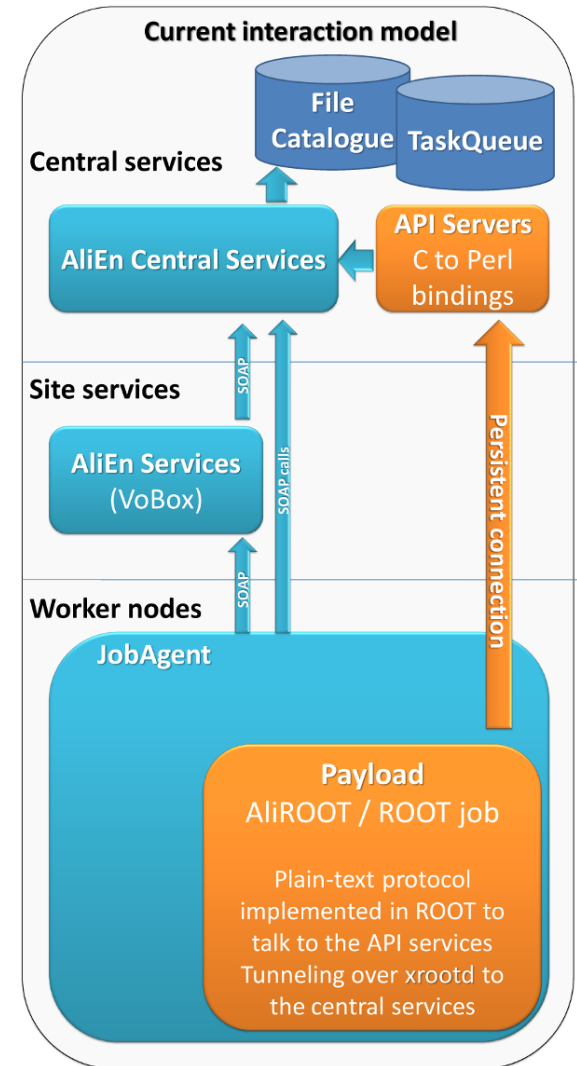
ALICE Analysis Trains

Welcome pbuncic - Help

PWG	Train name	I'm in	Last run	Description	Train operator(s)
CF	CF_PbPb		18 Oct 2013	Train for data PbPb running	jpgrosseo, miweber, sjena
CF	CF_PbPb_MC		30 Sep 2013		jpgrosseo, miweber, sjena
CF	CF_PbPb_MC_AOD		18 Oct 2013		jpgrosseo, miweber, sjena
CF	CF_pp		18 Oct 2013	Train for AOD pp correlation analyses	jpgrosseo, miweber, rodrigua
CF	CF_pPb		18 Oct 2013	pPb AOD express analysis train (Data)	jpgrosseo, miweber, sjena
CF	CF_pPb_MC		08 Oct 2013	pPb AOD express analysis train (MC)	jpgrosseo, miweber, sjena
CF	CF_pp_ESD		07 Aug 2013		jpgrosseo
CF	CF_pp_MC		09 Oct 2013	Train for AOD MC pp correlation analyses	jpgrosseo, miweber, rodrigua
CF	CF_pp_MC_ESD		12 Oct 2013		jpgrosseo
DQ	DQ_PbPb		20 Sep 2013	ESD	cbaumann, jbook, tgunji
DQ	DQ_PbPb_AOD		16 Sep 2013	AOD train	cbaumann, jbook, tgunji
DQ	DQ_PbPb_MC		28 Sep 2013	MC train for Pb-Pb ESD/AOD	cbaumann, jbook, tgunji
DQ	DQ_pp		15 Jul 2013	Data pp: AOD	cbaumann, jbook, tgunji
DQ	DQ_pPb		04 Oct 2013		cbaumann, jbook, tgunji
DQ	DQ_pPb_MC		27 Jun 2013		cbaumann, jbook, tgunji
DQ	DQ_pp_ESD		21 Aug 2013		cbaumann, jbook, tgunji
DQ	DQ_pp_MC		01 May 2013	MC: ESD	cbaumann, jbook, tgunji
GA	GA_PbPb		19 Oct 2013	Data train for PWGGA on ESDs of PbPb.	fbock, hqvigsta, mcosenti
GA	GA_PbPb_AOD		20 Sep 2013	Data train for PWGGA on AODs of PbPb.	fbock, hqvigsta, mcosenti
GA	GA_PbPb_MC		19 Oct 2013	MC train for PWGGA on ESDs of PbPb.	fbock, hqvigsta, mcosenti
GA	GA_PbPb_MC_AOD		29 Sep 2013	MC train for PWGGA on AODs of PbPb.	fbock, hqvigsta, mcosenti
GA	GA_pp		12 Jul 2013	Data train for PWGGA on ESDs of pp.	fbock, hqvigsta, mcosenti
GA	GA_pp_AOD		01 Aug 2013	Data train for PWGGA on AODs of pp.	fbock, hqvigsta, mcosenti

AliEn Summary

- 3-layer system that leverages the deployed resources of the underlying WLCG infrastructures and services
 - including the local variations, such as EGI, NDGF and OSG
- Interfaces to AliRoot via ROOT plugin(TAliEn) that implements AliEn API
- Complex workflows including distributed analysis built on top of AliEn API
- Used by Panda (at GSI)





MonALISA - Monitoring



MonALISA Repository for ALICE

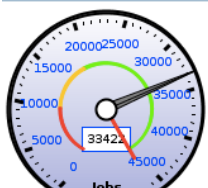


My jobs | My home dir | Catalogue browser | LEGO Trains | Administration Section | ALICE Reports | Alert XML Feed | Firefox Toolbar | MonALISA GUI

- ALICE Repository
 - ALICE Repository
 - Google Map
 - Shifter's dashboard
 - Run Condition Table
 - Production Overview
 - Production info
 - Job Information
 - SE Information
 - Services
 - Network Traffic
 - FTD Transfers
 - CAF Monitoring
 - SHUTTLE
 - Build system
 - HepSpec
 - Dynamic charts

This page: bookmark, URL

Running Jobs trend





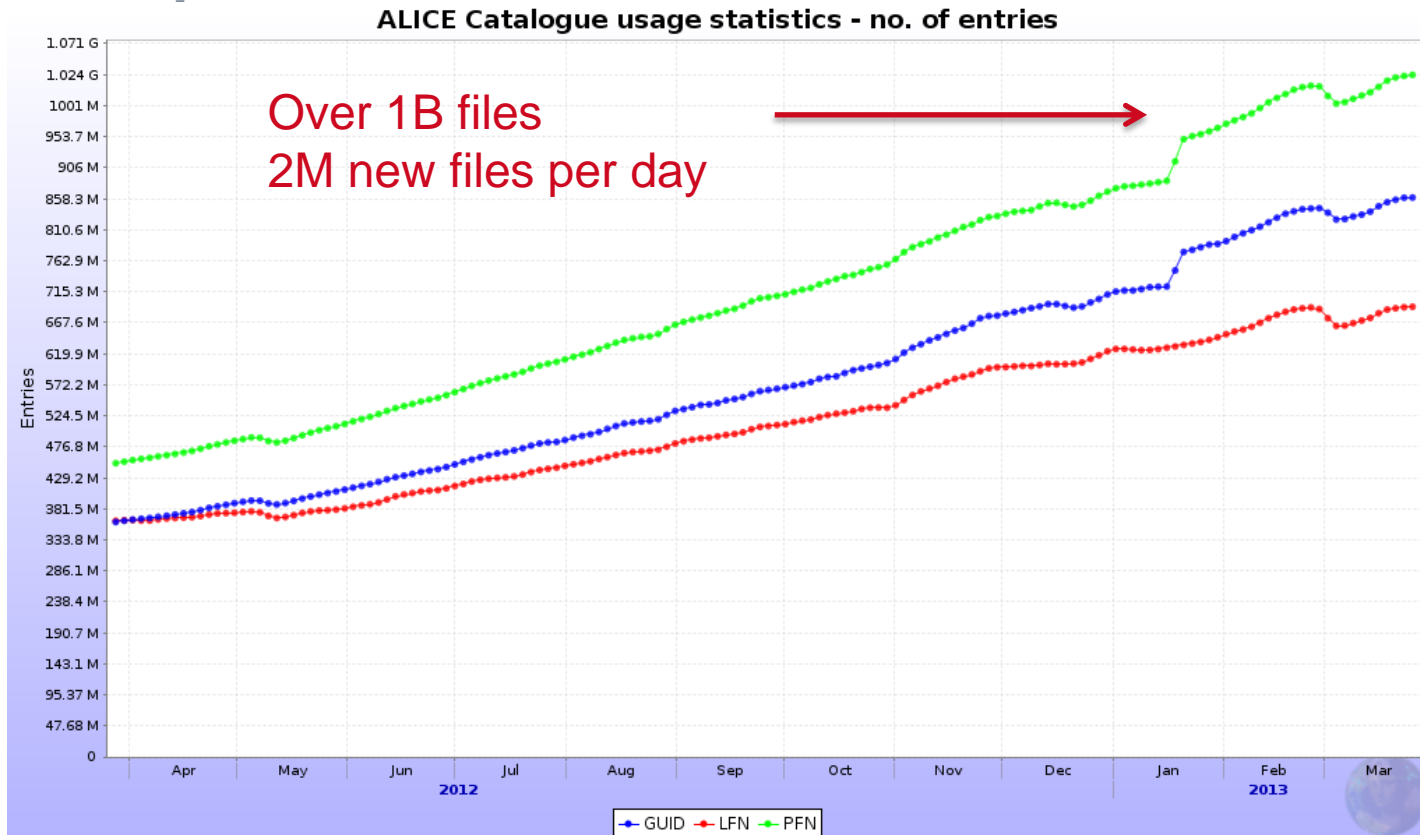
ALICE Upgrade plans

- ALICE plans some serious detector upgrades
- Run2 (2015-2017)
 - 4 fold increase in instant luminosity for Pb-Pb collisions
 - consolidation of the readout electronics of TPC and TRD (readout rate x 2)
- Run3 (2019-2021)
 - Continuous readout TPC, ITS upgrade
 - 50kHz Pb-Pb interaction rate (current rate x 100)
 - 1.1 TB/s detector readout
 - Needs online reconstruction in order to better compress data for storage

Why changes in AliEn?

- While the system currently fulfills all the needs of ALICE users for reconstruction, simulation and analysis there are concerns about scalability of the file catalog beyond Run2
- Need to address the use for emerging cloud, volunteer as well as the opportunistic resources for ALICE
- In general, no manpower for maintenance and continuous development of the current system

AliEn File Catalog



Data management is built in AliEn, data access using xrood

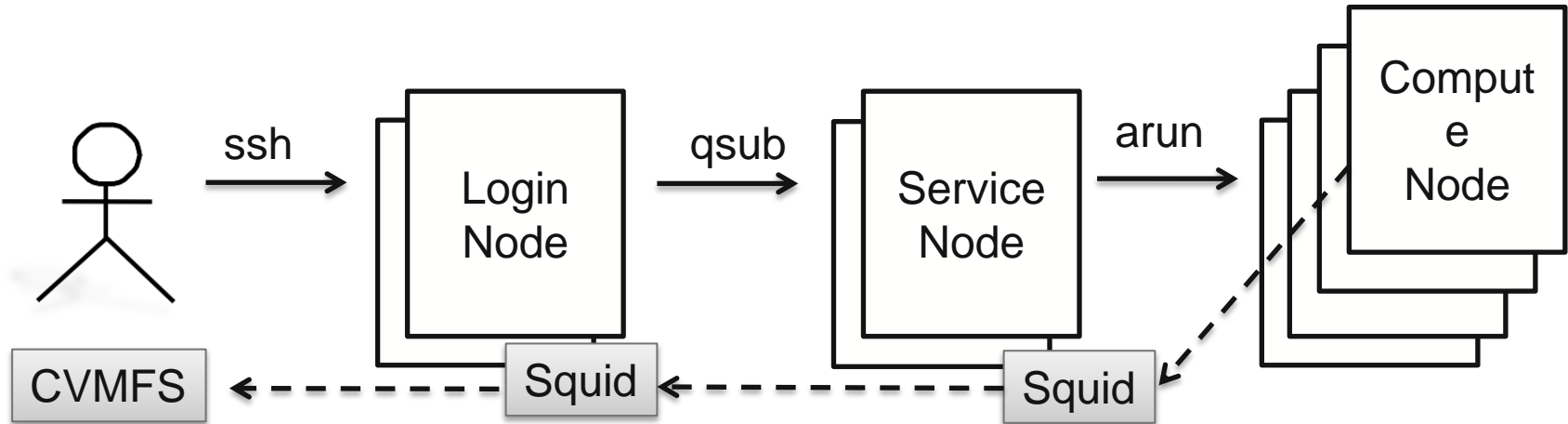
Opportunistic use of SC resources for simulation



- 18,688 nodes with total of 299,008 processor cores and one GPU per node. And, there are spare CPU cycles available...
- Ideal for event generators/simulation type workloads (60-70% of all CPU cycles used by the experiments).
- Simulation frameworks must be adapted to efficiently use such resources where available

Not an easy job...

No internet, minimal OS



```
$ [titan-batch6][12:12:09][/tmp/work/atj/cvmfs_install/bin]$ aprun ./parrot_run -t/tmp/scratch /cvmfs/alice.cern.ch/bin/alienv setenv AliRoot -c aliroot -b
```

```
*****
* WELCOME to ROOT *
* Version 5.34/08 31 May 2013 *
* You are welcome to visit our Web site
* http://root.cern.ch *
* *
*****
```

Thanks to CMS parrot (from cctools) extended to allow access to CVMFS (some restrictions apply)

```
ROOT 5.34/08 (v5-34-08@v5-34-08, Jun 11 2013, 10:26:13 on linuxx8664gcc)
```

```
CINT/ROOT C/C++ Interpreter version 5.18.00, July 2, 2010
Type ? for help. Commands must be C++ statements.
Enclose multiple statements between { }.
2+2
(const int)4
```

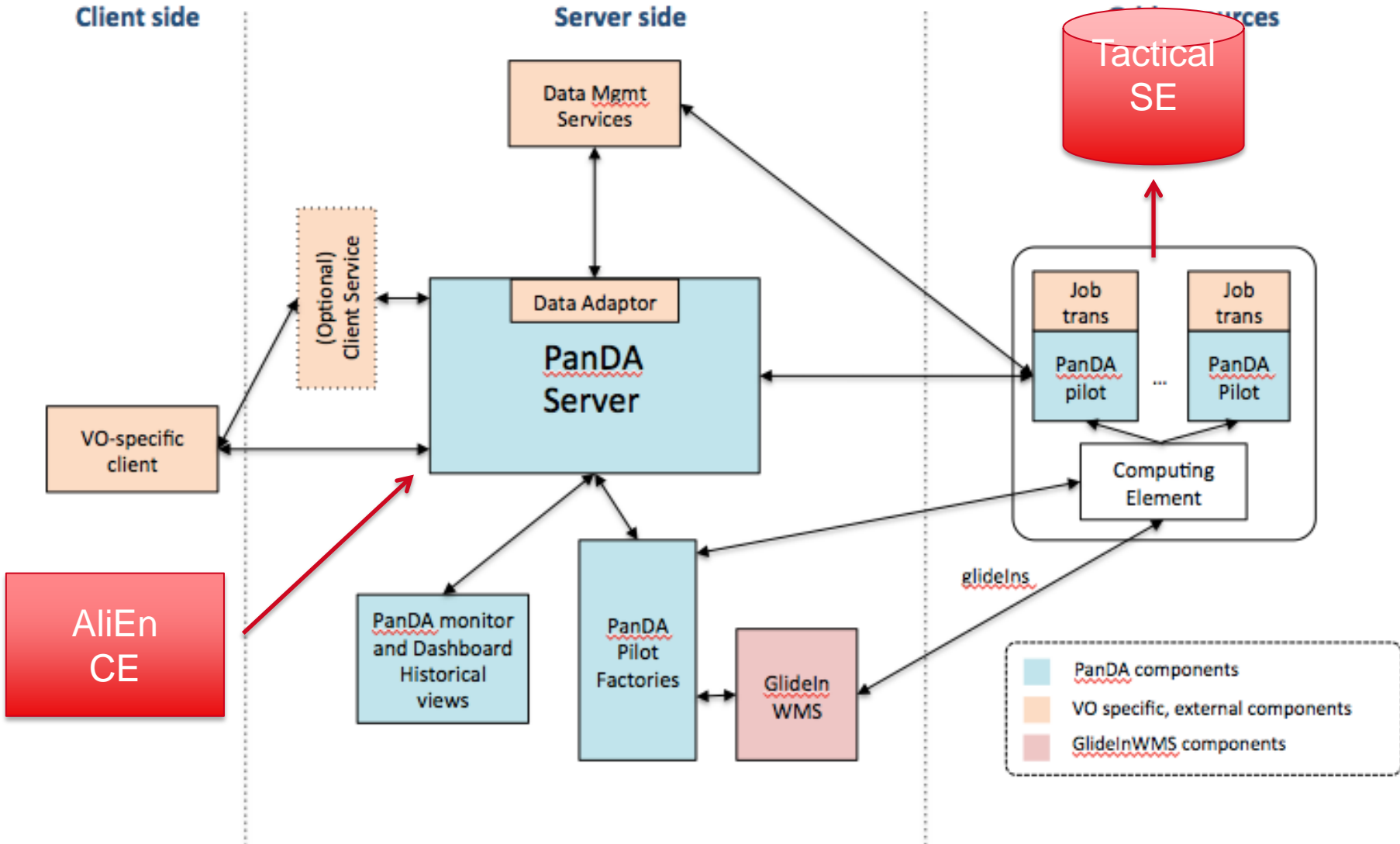
Adam Simpson, ORNL

Interfacing with PanDA (Plan A)

Client side

Server side

Resources

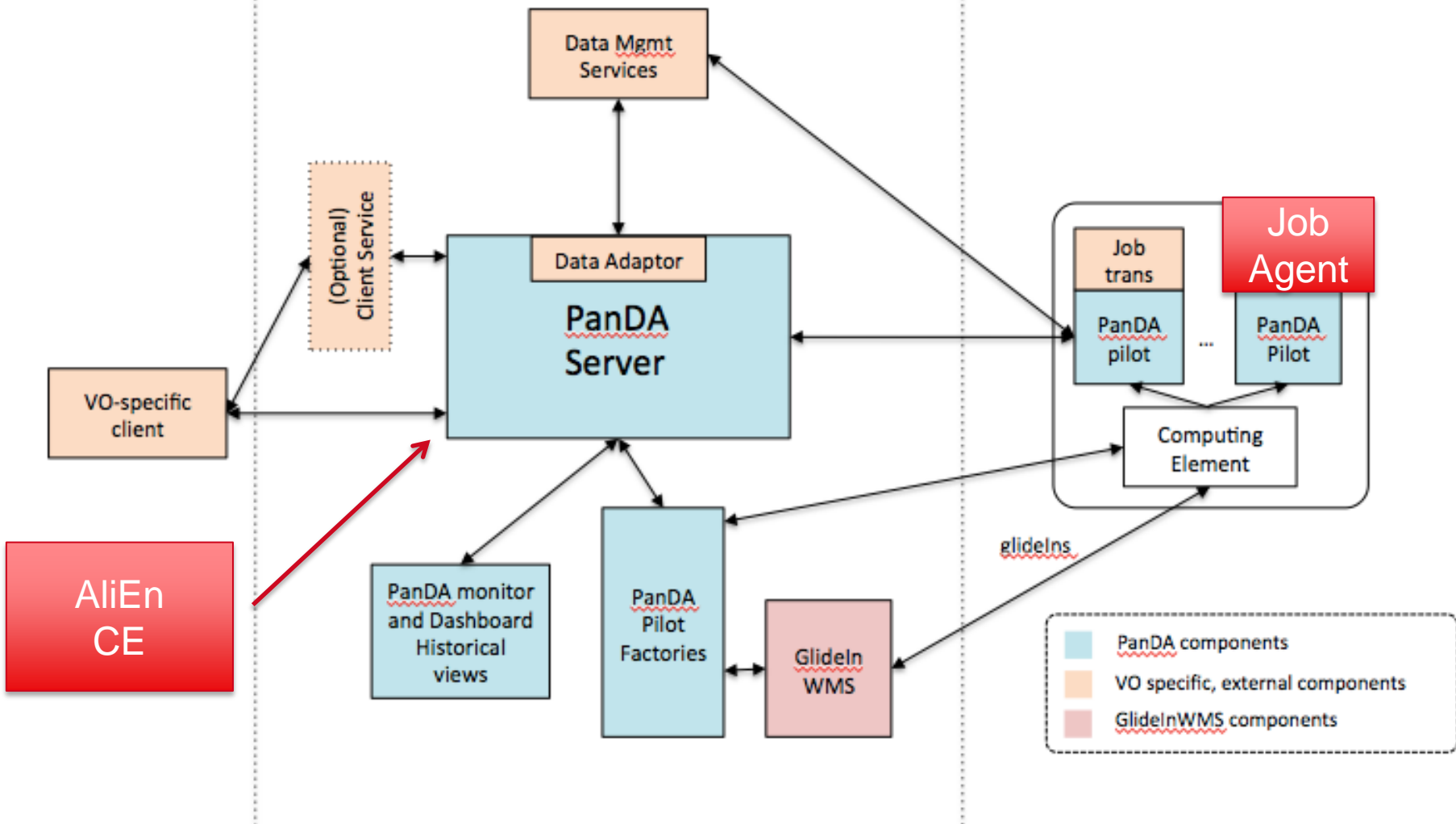


Interfacing with PanDA (Plan B)

Client side

Server side

Grid resources

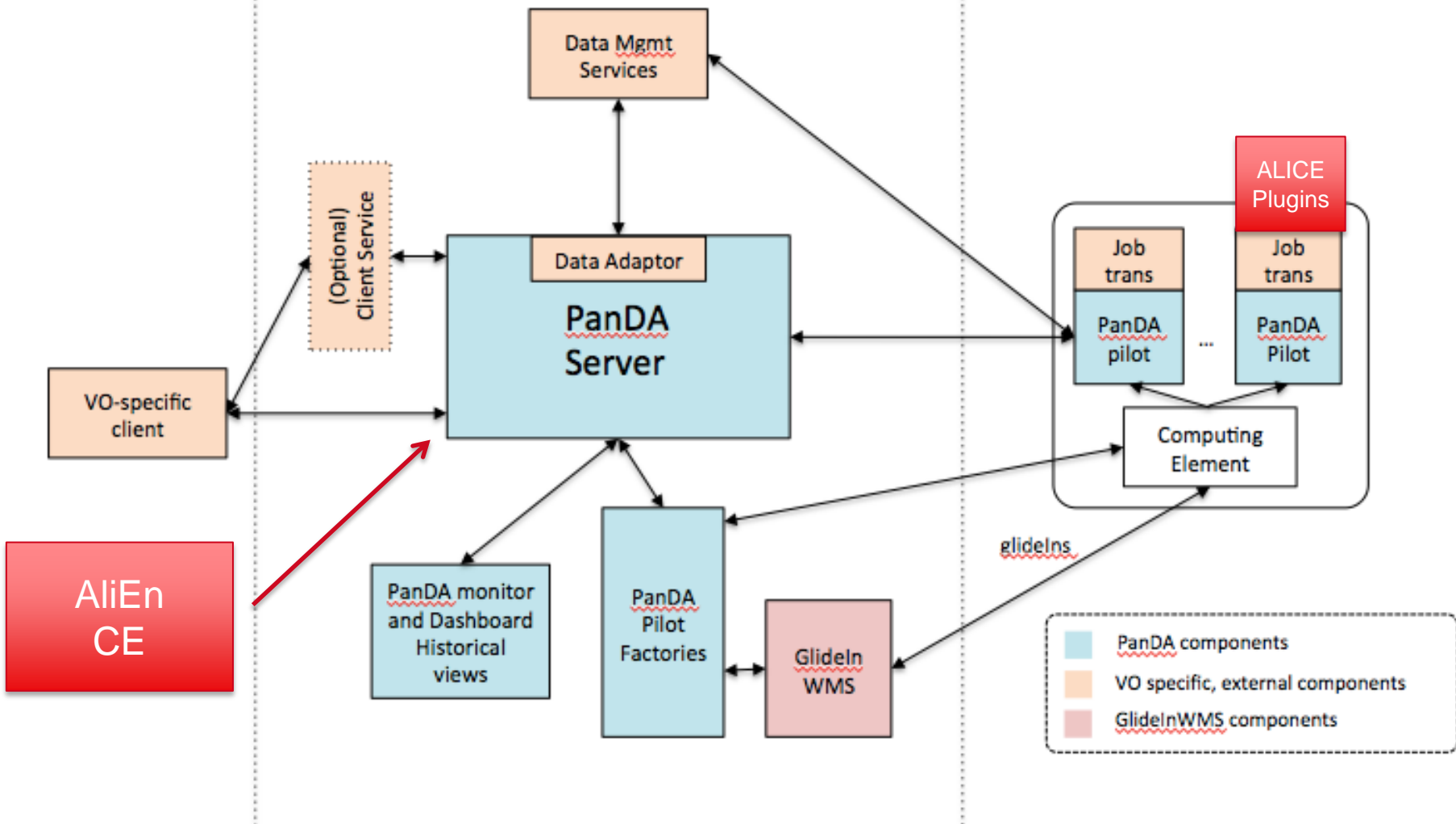


Interfacing with PanDA (Plan C)

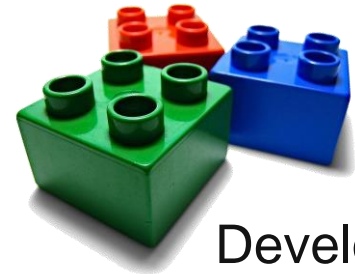
Client side

Server side

Grid resources



Conclusions



Development is fun, maintenance is boring

- We have to concentrate on upgrade and facing manpower issues
- We expect to carry on with current AliEn implementation throughout Run2
 - Room for improvement (catalog, data management, clouds)
- Run3 will impose much higher requirements compared to Run2
 - 100x more events to handle
 - Huge simulation needs
- We need to prepare ourselves to use opportunistic resources to complement our simulation needs
 - PanDA is our chance to open a door of potentially large supercomputing facilities
- Several scenarios of integration exist, need to find the best one
- We have to preserve AliEn APIs to preserve investment in tools built on top of AliEn