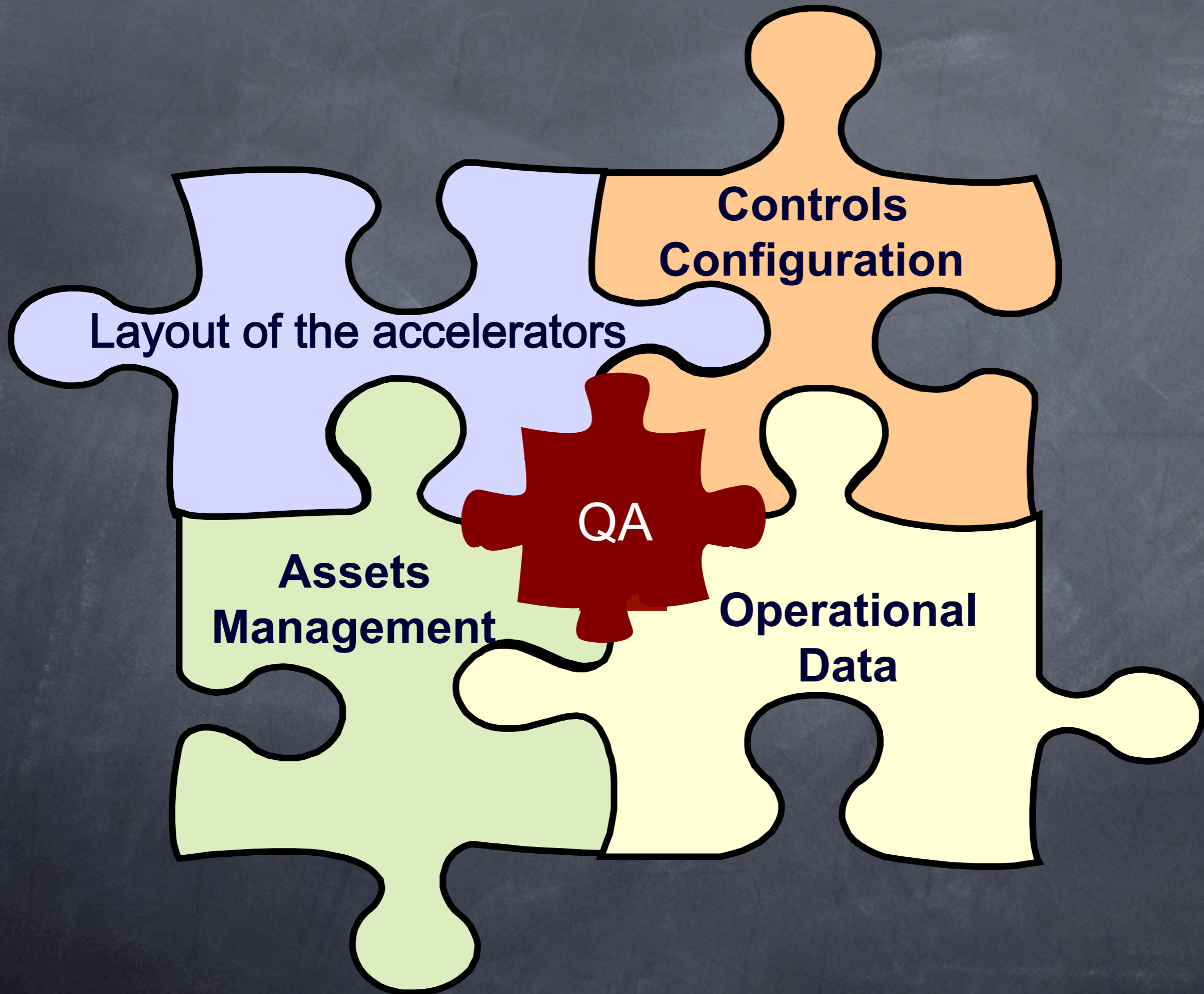


Data Management for Accelerator Control & Operation

Part II

Chris Roderick, Zory Zaharieva
CERN, BE-CO-DM

Database Futures Workshop, 2011-06-06



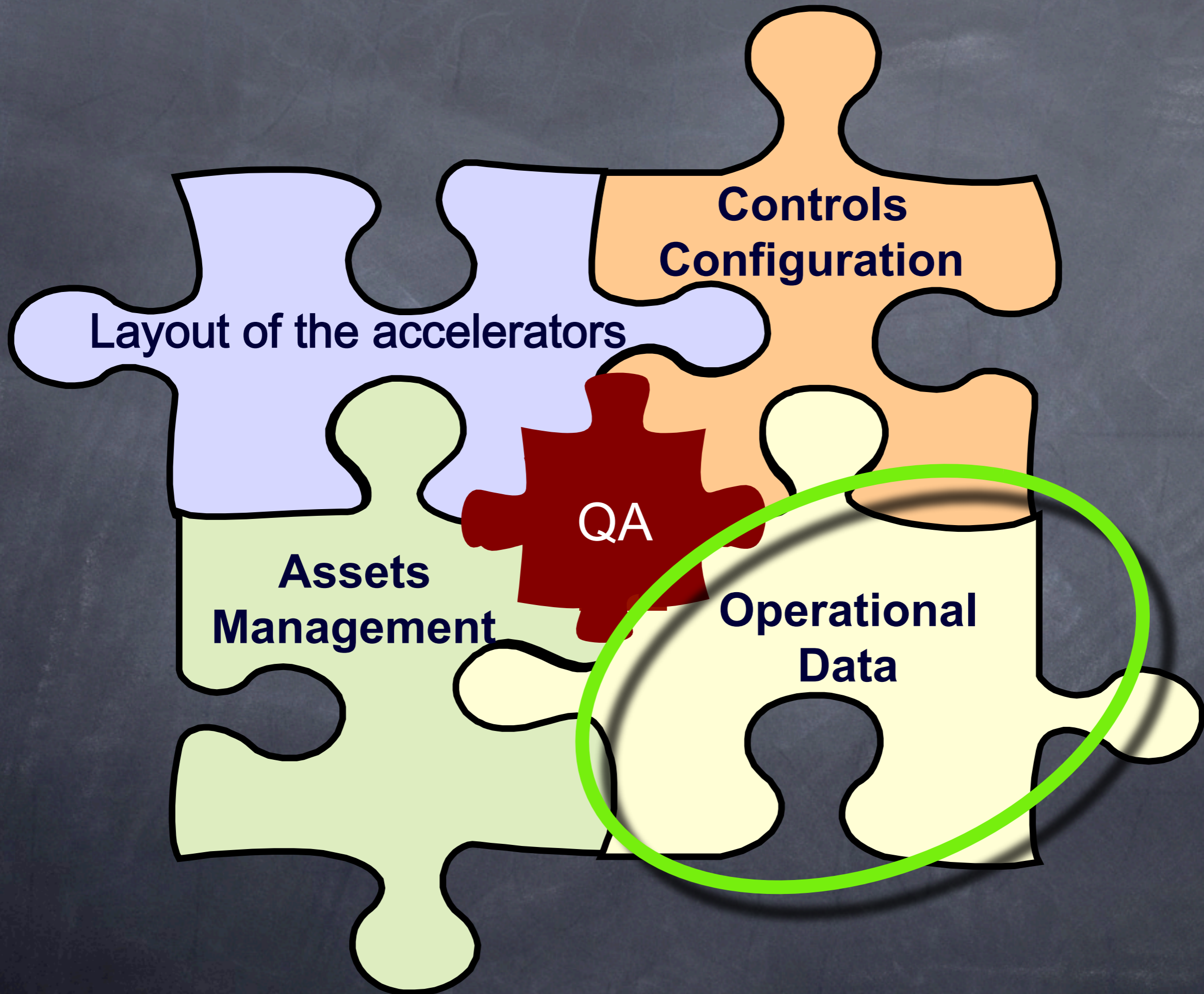
Layout of the accelerators

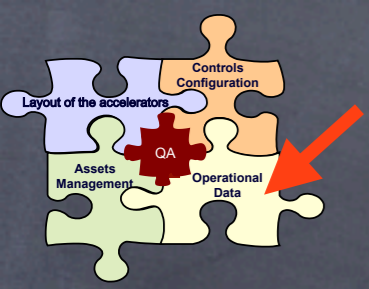
**Controls
Configuration**

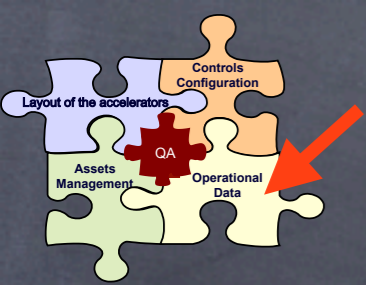
**Assets
Management**

**Operational
Data**

QA

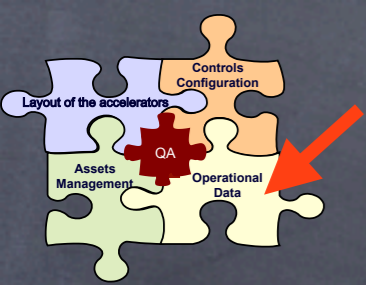






Accelerator Settings Management

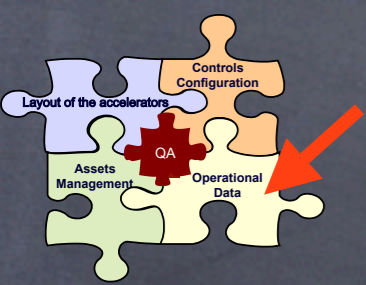
modify & drive settings to 10,000's of
accelerator devices to **control particle beams**



Accelerator Settings Management

modify & drive settings to 10,000's of
accelerator devices to **control particle beams**

maintain **full history of all changes**
revert to prior settings

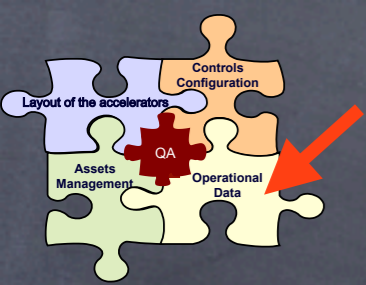


Accelerator Settings Management

modify & drive settings to 10,000's of
accelerator devices to **control particle beams**

maintain **full history of all changes**
revert to prior settings

exploitation impossible with erroneous settings



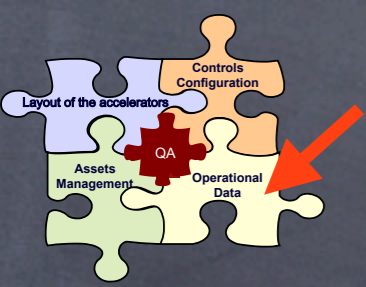
Accelerator Settings Management

modify & drive settings to 10,000's of
accelerator devices to **control particle beams**

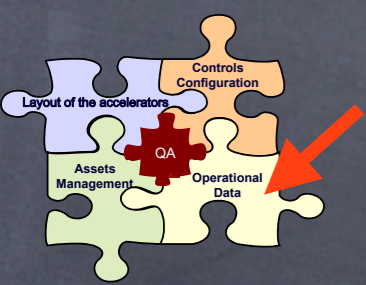
maintain **full history of all changes**
revert to prior settings

exploitation impossible with erroneous settings

MISSION CRITICAL



Accelerator Settings Management



Accelerator Settings Management

Very complex domain

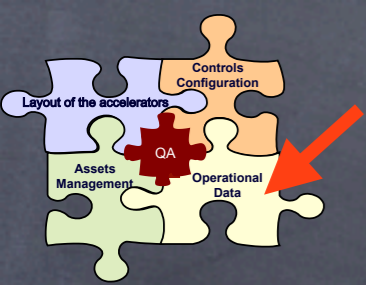
relational model using Oracle DBMS

317 tables

573 indexes

1,476 constraints

16k lines of PL/SQL



Accelerator Settings Management

Very complex domain

relational model using Oracle DBMS

317 tables

573 indexes

1,476 constraints

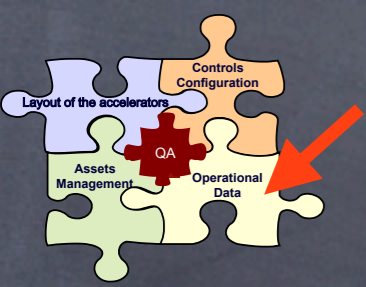
16k lines of PL/SQL

43 million setting headers

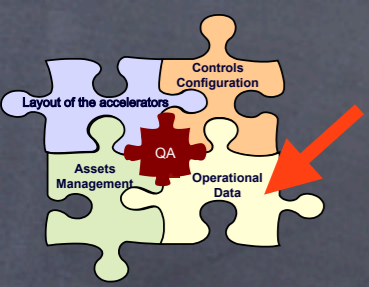
3.9 billion setting values (~40GB)

39k logical + physical devices

5 particle accelerators



Accelerator Settings Management

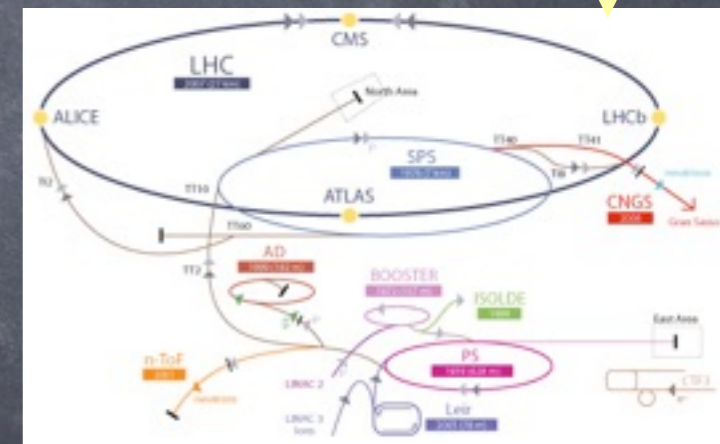
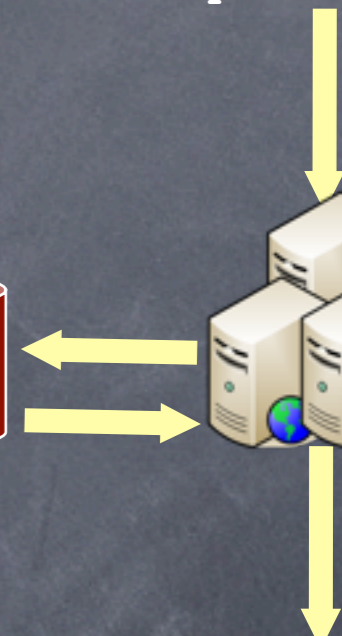


Accelerator Settings Management

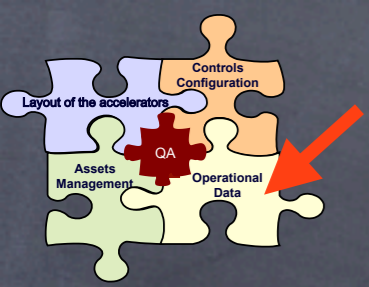


<Acc Operators>

Database is used **on-line**



<Accelerators>



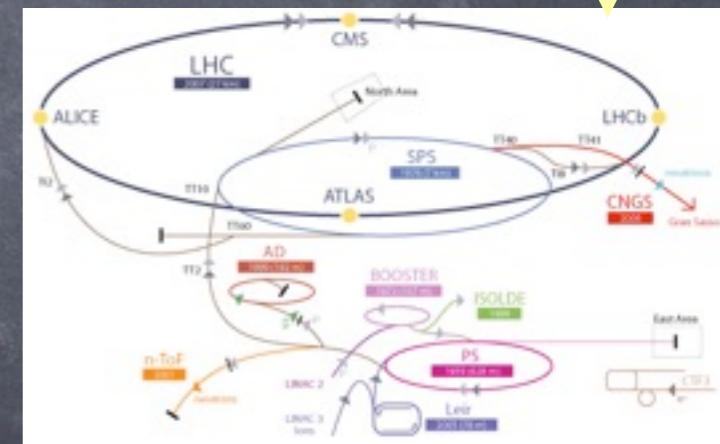
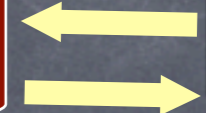
Accelerator Settings Management



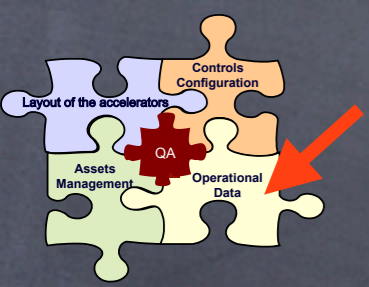
<Acc Operators>

Database is used **on-line**

Response times & High Availability → Critical



<Accelerators>



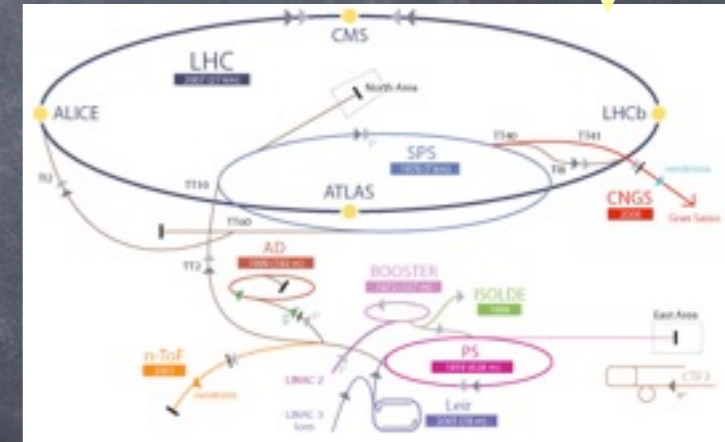
Accelerator Settings Management



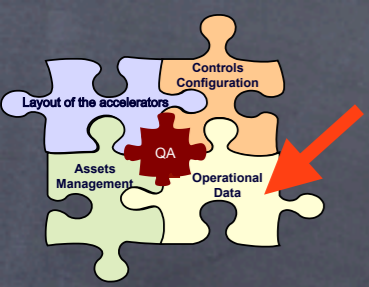
<Acc Operators>

Database is used **on-line**

Response times & High Availability → Critical



<Accelerators>



Accelerator Settings Management



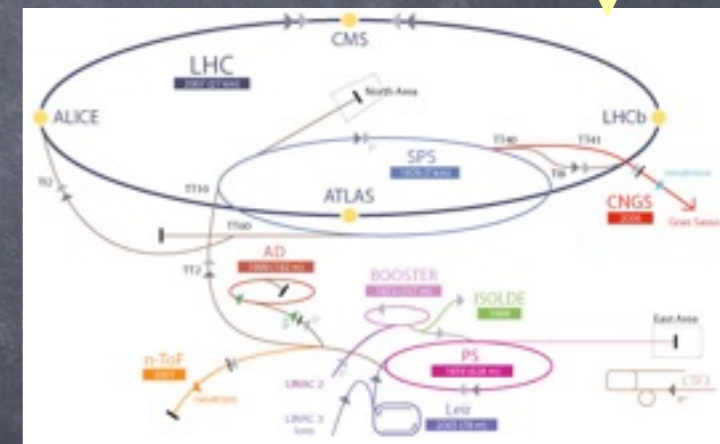
<Acc Operators>

Database is used **on-line**

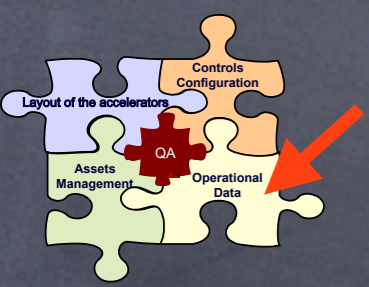
Response times & High Availability → Critical



LSA Database
ORACLE



<Accelerators>



Accelerator Settings Management



<Acc Operators>

Database is used **on-line**

Response times & High Availability → Critical

Many proprietary features



Hash / Range Partitioning

IOTs

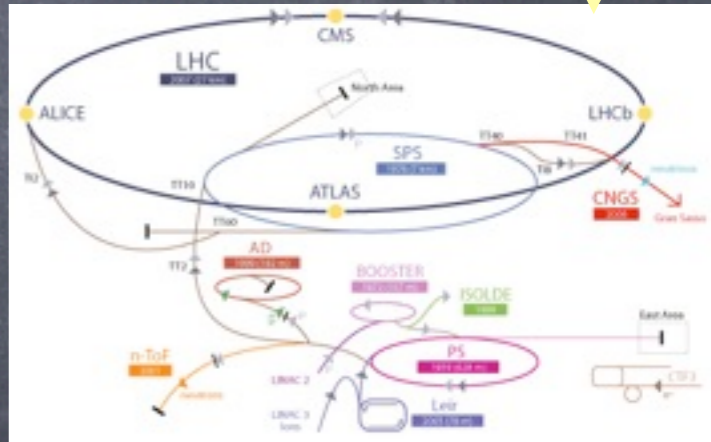
Function / Bitmap Indexes

PL/SQL

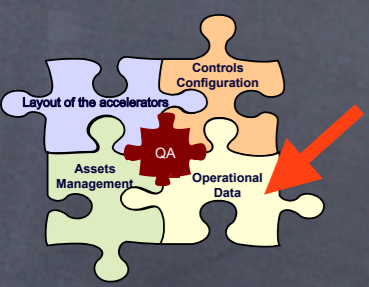
Analytic Functions

Compression

Bulk Operations



<Accelerators>



Accelerator Settings Management



<Acc Operators>

Database is used **on-line**

Response times & High Availability → Critical

Many proprietary features



Hash / Range Partitioning

IOTs

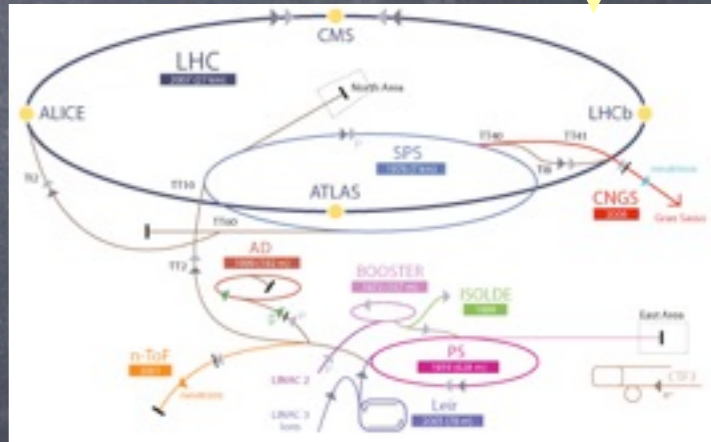
Function / Bitmap Indexes

PL/SQL

Analytic Functions

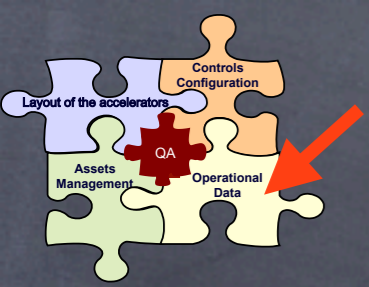
Bulk Operations

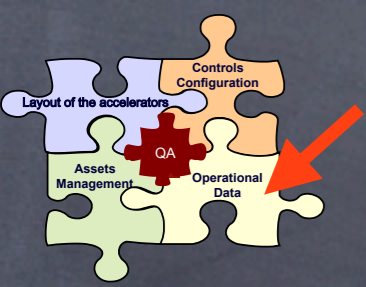
Compression



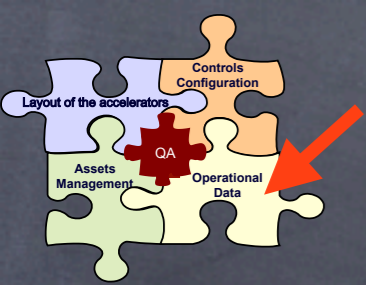
<Accelerators>

HA via 2-node RAC with NAS



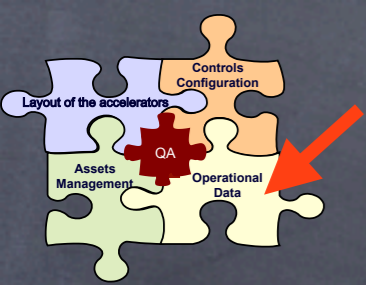


Accelerator Logging Service



Accelerator Logging Service

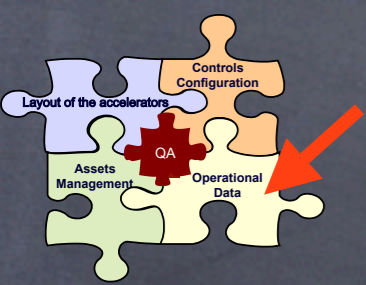
store accelerator beam & equipment data
on-line, beyond the lifetime of the LHC (>20y)



Accelerator Logging Service

store accelerator beam & equipment data
on-line, beyond the lifetime of the LHC (>20y)

analyse behaviour accelerators & sub-systems
over long periods of time

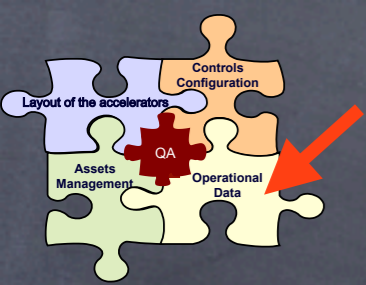


Accelerator Logging Service

store accelerator beam & equipment data
on-line, beyond the lifetime of the LHC (>20y)

analyse behaviour accelerators & sub-systems
over long periods of time

MISSION CRITICAL



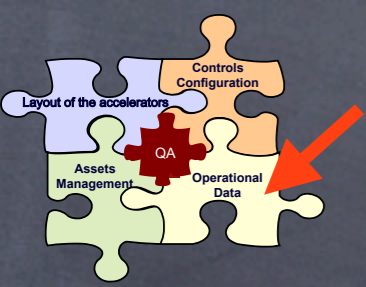
Accelerator Logging Service

store accelerator beam & equipment data
on-line, beyond the lifetime of the LHC (>20y)

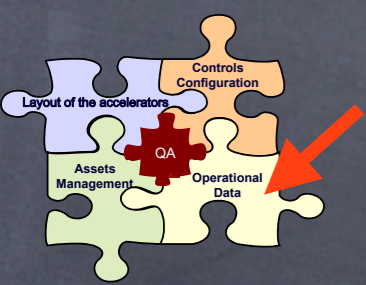
analyse behaviour accelerators & sub-systems
over long periods of time

MISSION CRITICAL

In 2001, based on LEP experience
estimated data rate: 1TB/year...



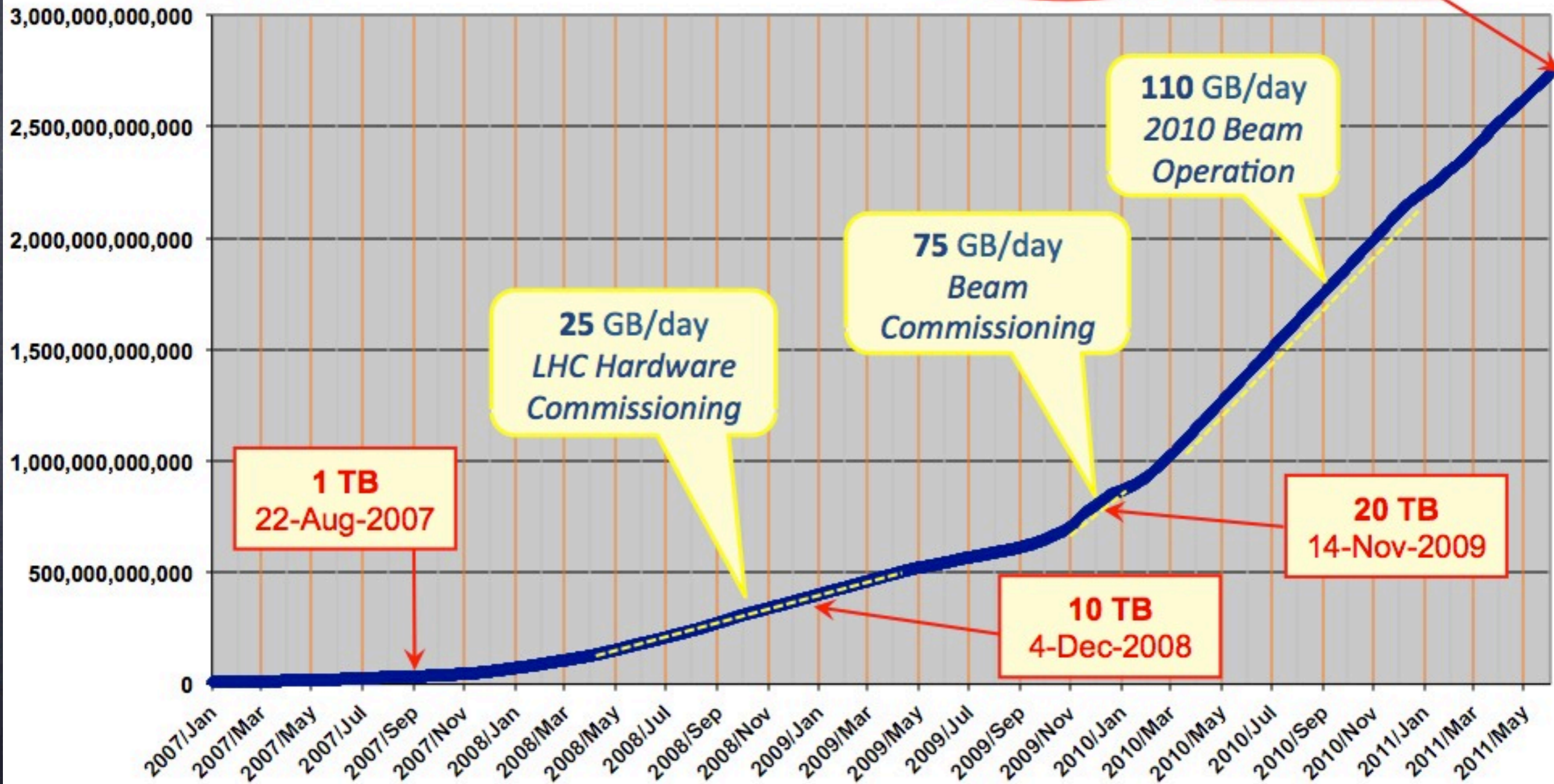
Accelerator Logging Service

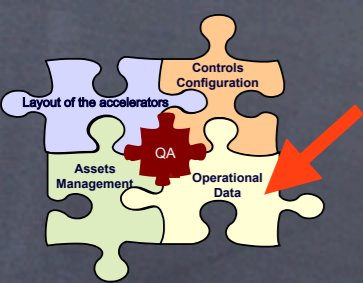


Accelerator Logging Service

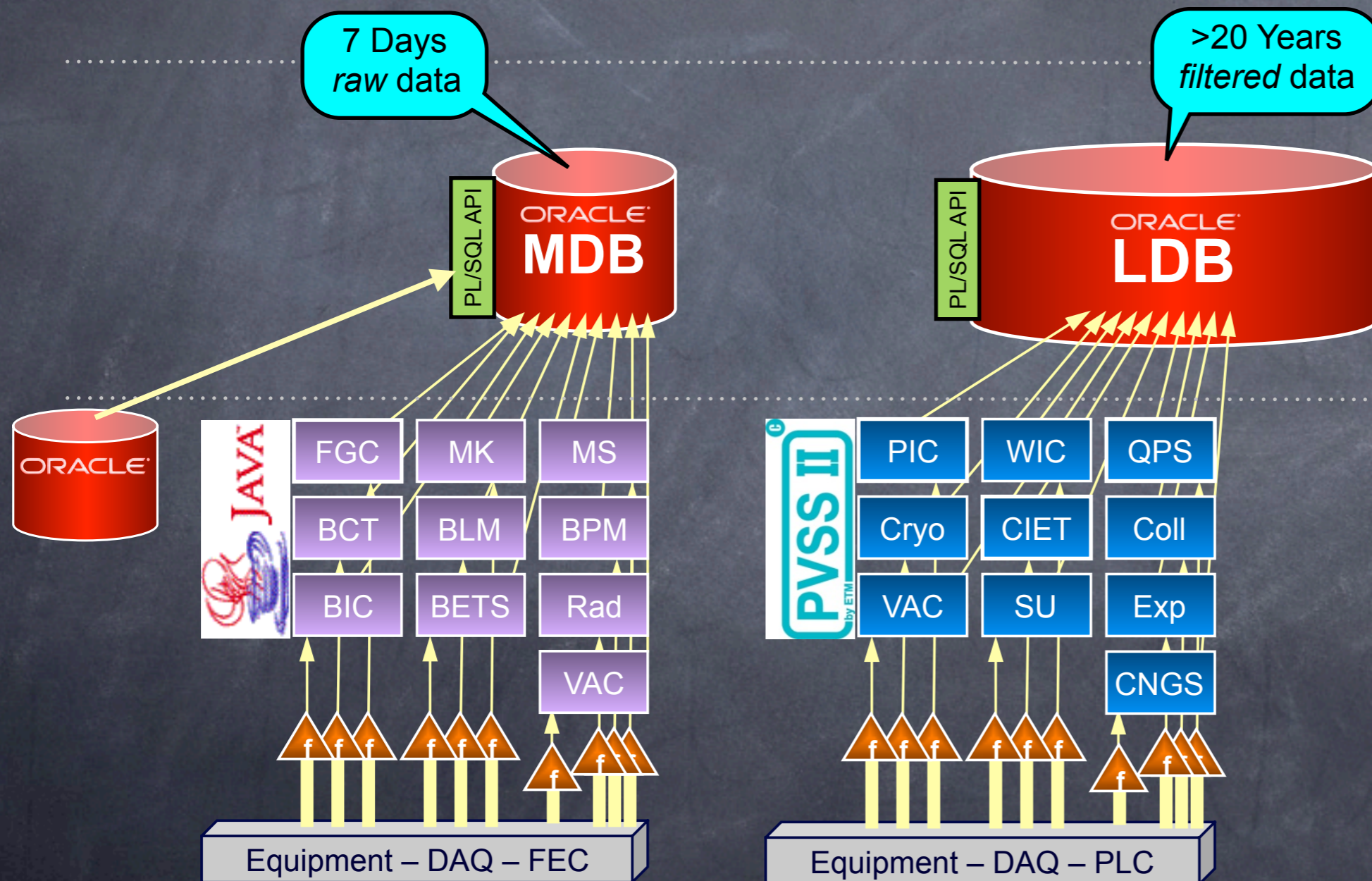
LHC Logging Service - Logged records

◆ logged records

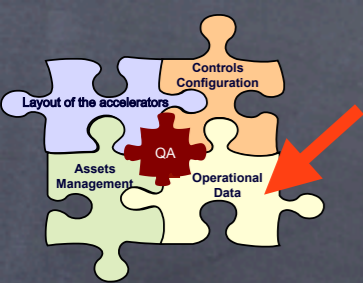




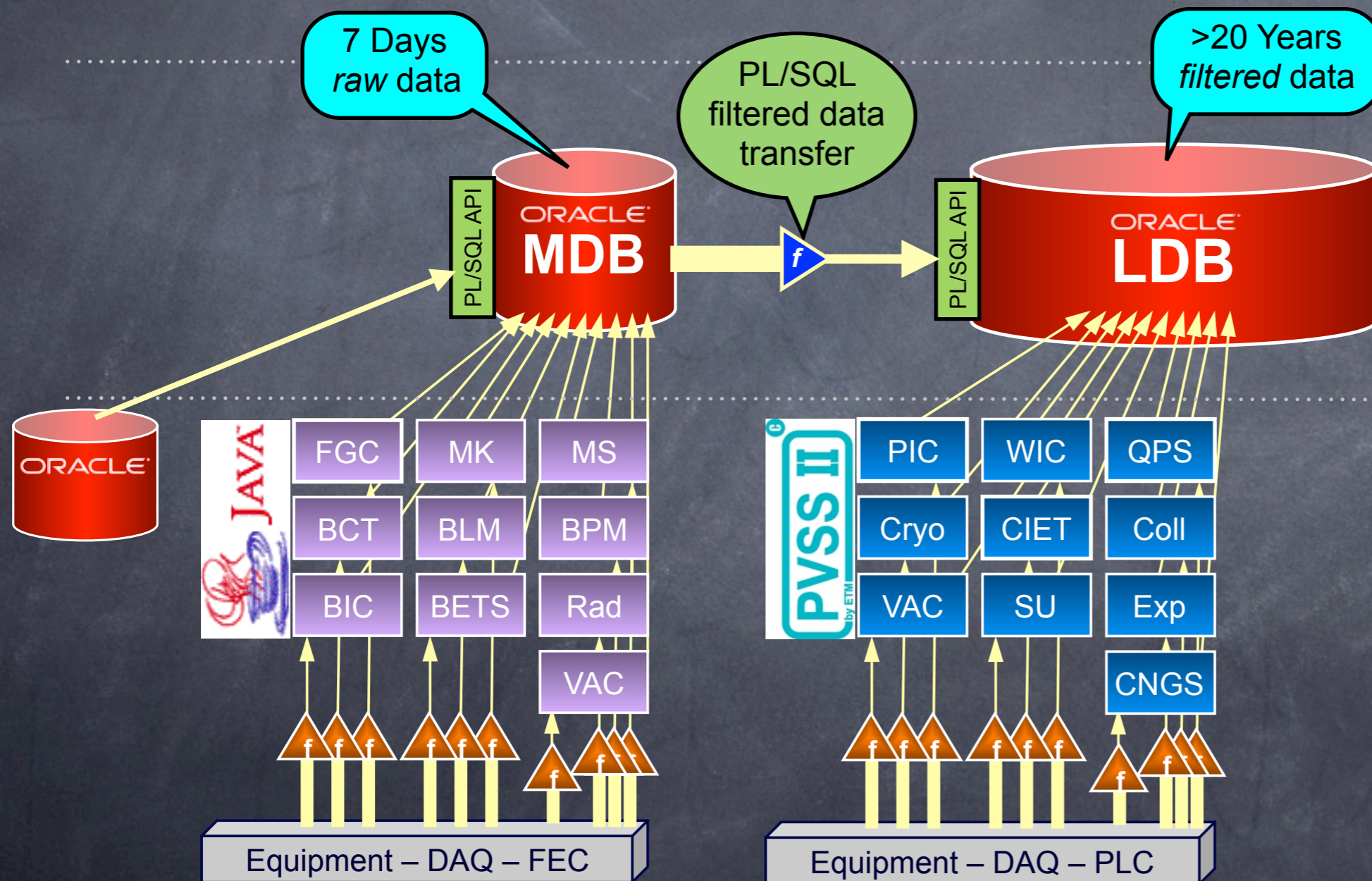
Accelerator Logging Service



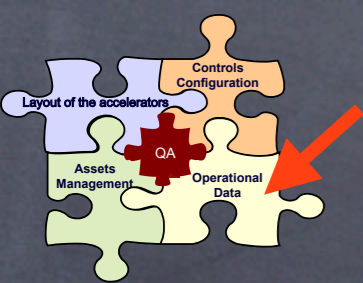
Data Consumers
Data Persistence
Data Providers



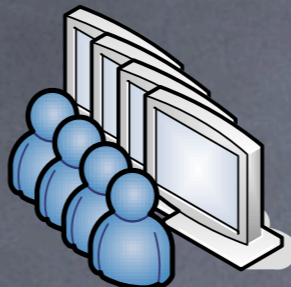
Accelerator Logging Service



Data Consumers
Data Persistence
Data Providers



Accelerator Logging Service

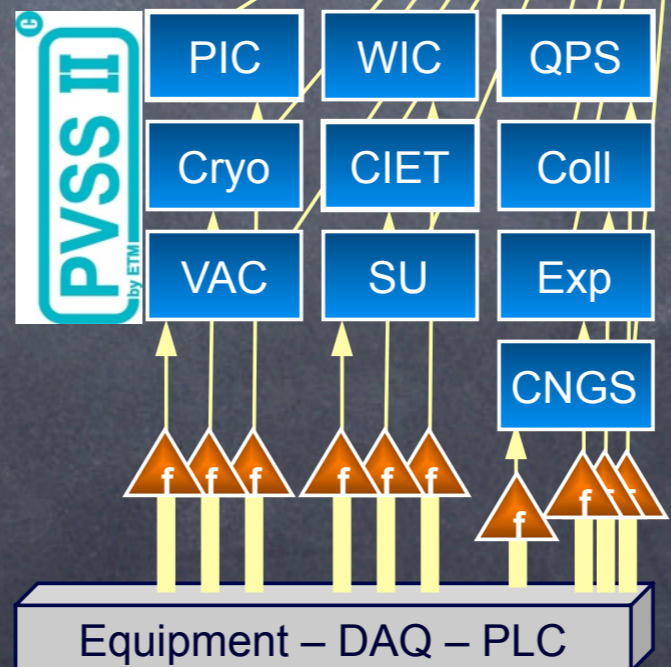
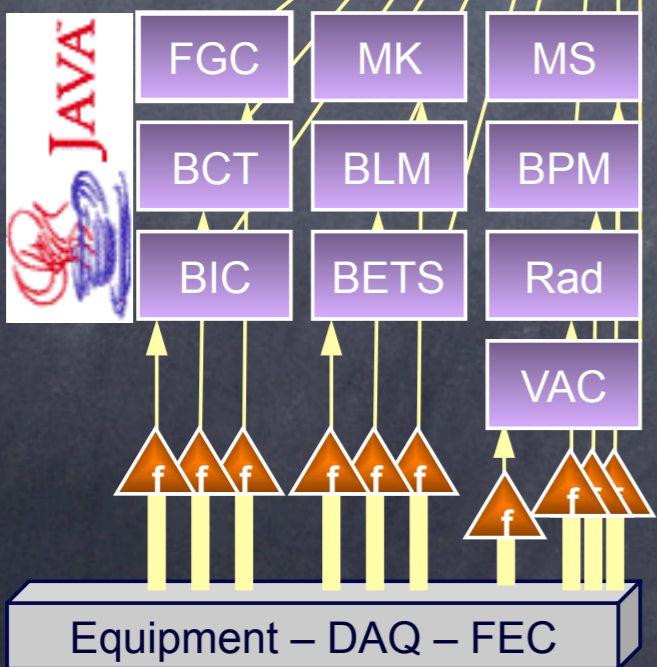
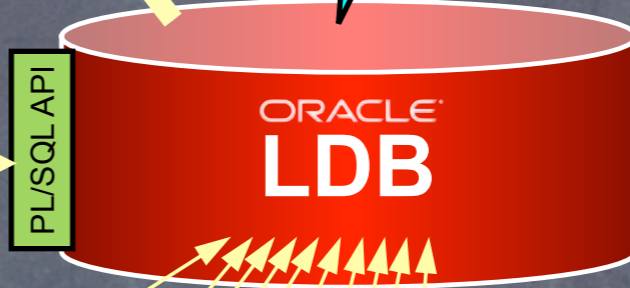
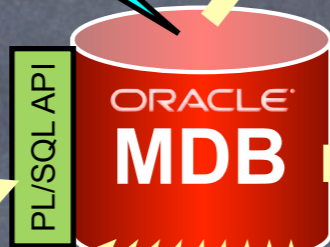


Extraction API

7 Days raw data

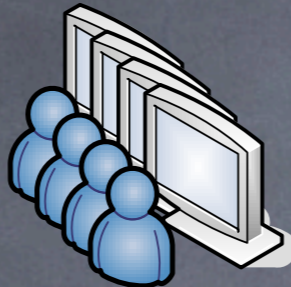
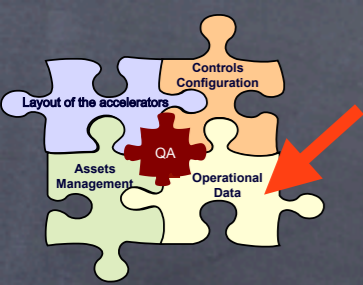
PL/SQL filtered data transfer

>20 Years filtered data



Data Consumers
Data Persistence
Data Providers

Accelerator Logging Service



Extraction API

7 Days raw data

>20 Years filtered data

~ 250'000 Signals
 ~ 14 data loading processes
 ~ 5.4 billion records / day
 ~ 250 GB / day
 → 90 TB / year throughput

PL/SQL filtered data transfer

PL/SQL API

PL/SQL API

ORACLE MDB

ORACLE LDB

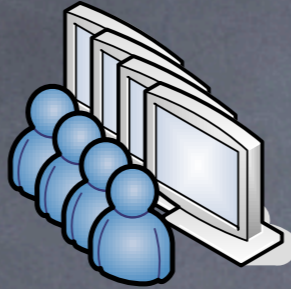
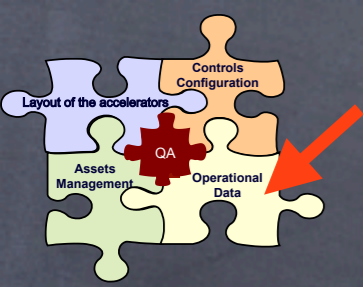


Equipment – DAQ – FEC

Equipment – DAQ – PLC

Data Consumers
 Data Persistence
 Data Providers

Accelerator Logging Service



Extraction API

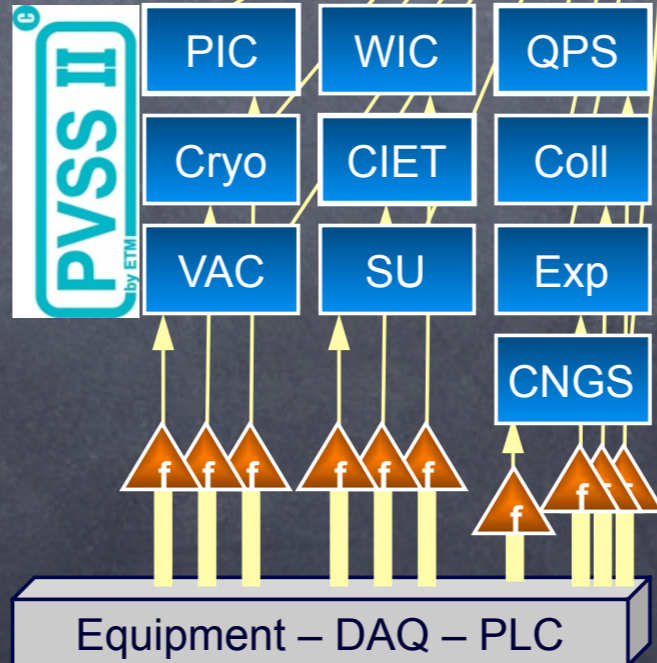
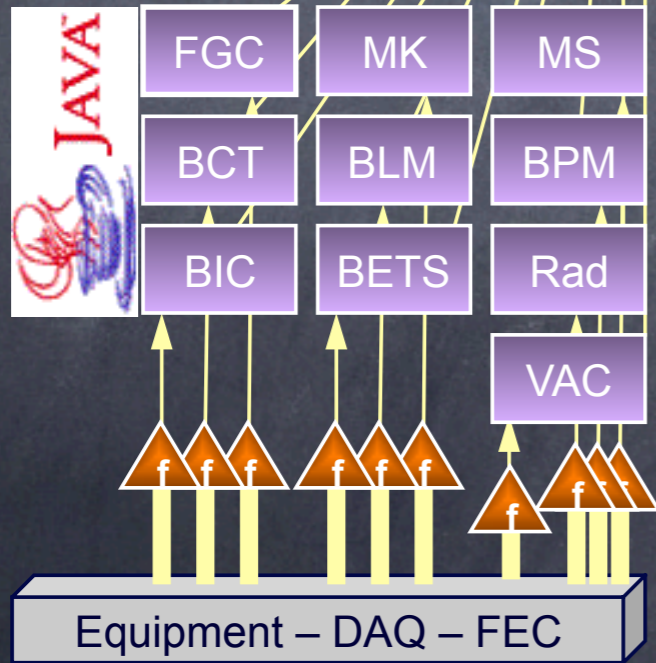
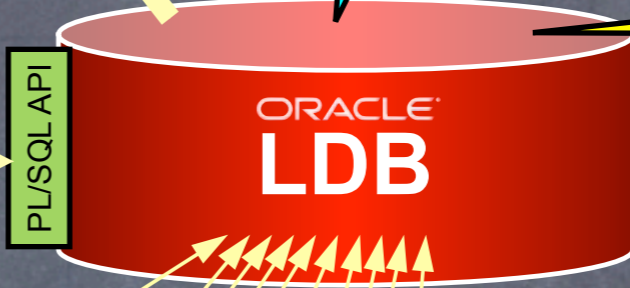
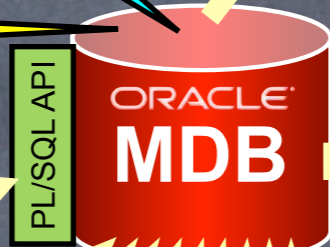
7 Days raw data

PL/SQL filtered data transfer

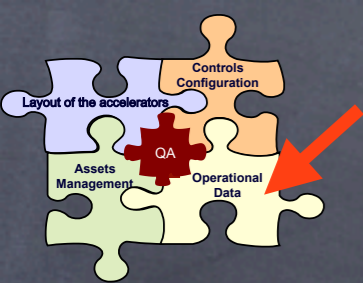
>20 Years filtered data

~ 250'000 Signals
 ~ 14 data loading processes
 ~ 5.4 billion records / day
 ~ 250 GB / day
 → 90 TB / year throughput

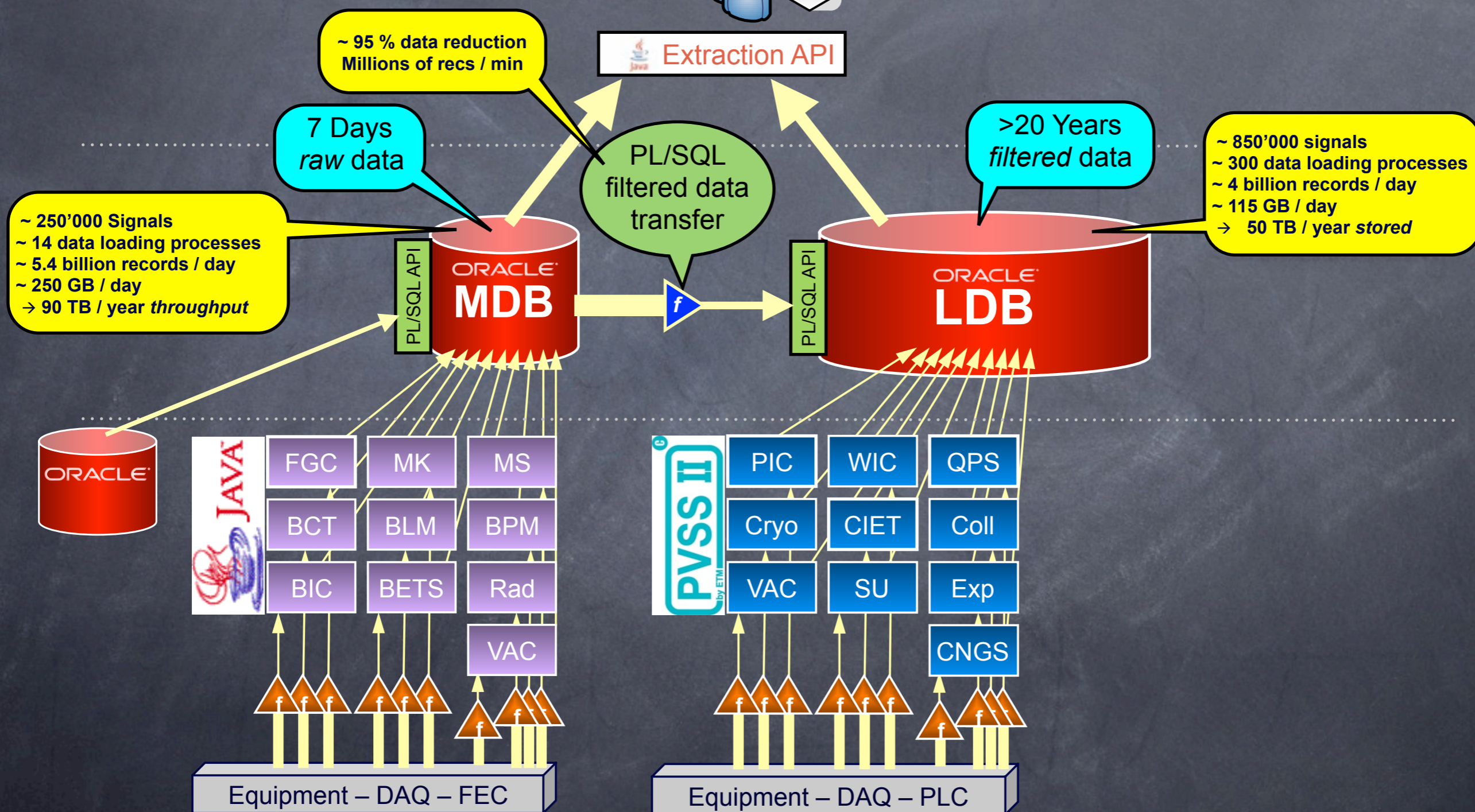
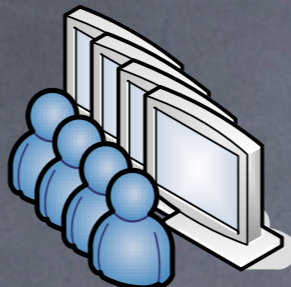
~ 850'000 signals
 ~ 300 data loading processes
 ~ 4 billion records / day
 ~ 115 GB / day
 → 50 TB / year stored



Data Consumers
 Data Persistence
 Data Providers



Accelerator Logging Service



~ 250'000 Signals
 ~ 14 data loading processes
 ~ 5.4 billion records / day
 ~ 250 GB / day
 → 90 TB / year throughput

7 Days raw data

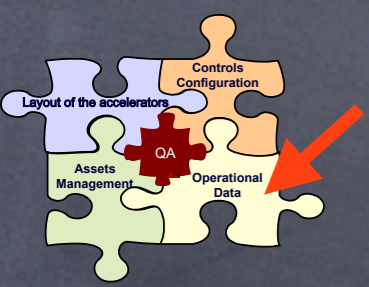
~ 95 % data reduction
 Millions of recs / min

PL/SQL filtered data transfer

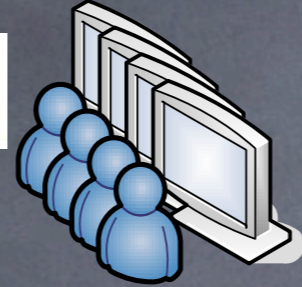
>20 Years filtered data

~ 850'000 signals
 ~ 300 data loading processes
 ~ 4 billion records / day
 ~ 115 GB / day
 → 50 TB / year stored

Data Consumers
Data Persistence
Data Providers



Accelerator Logging Service



~ 500 active extraction clients
 ~ 90 custom applications
 ~ 2 million extraction requests / day

Extraction API

~ 95 % data reduction
 Millions of recs / min

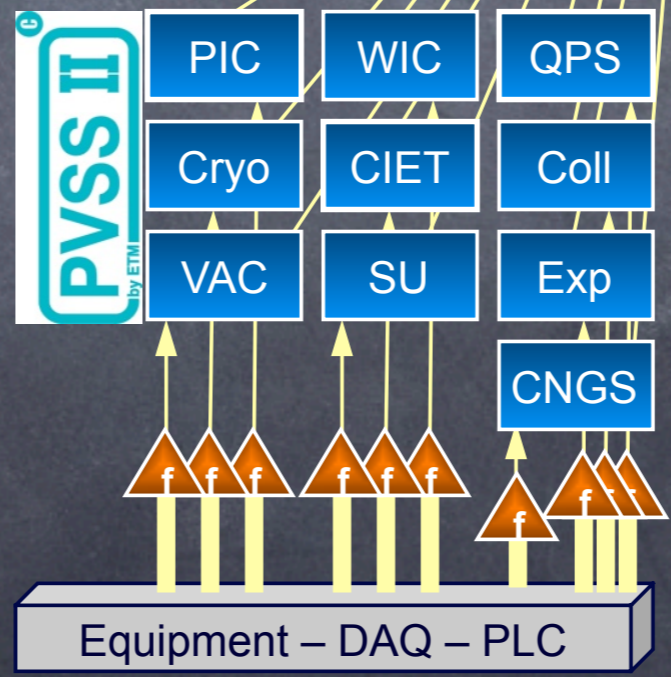
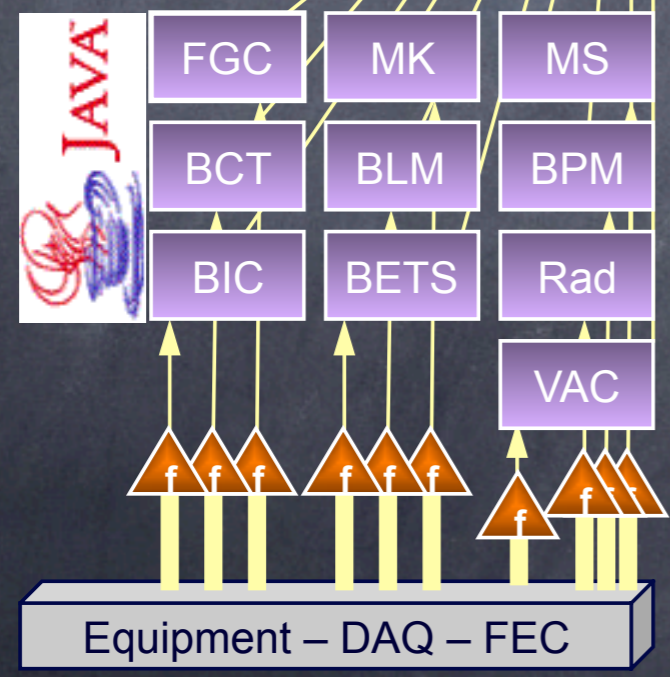
7 Days raw data

PL/SQL filtered data transfer

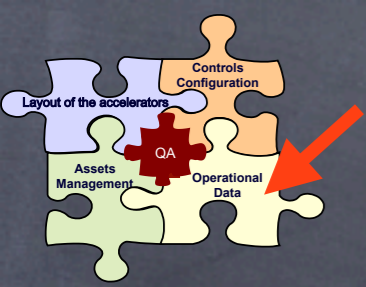
>20 Years filtered data

~ 850'000 signals
 ~ 300 data loading processes
 ~ 4 billion records / day
 ~ 115 GB / day
 → 50 TB / year stored

~ 250'000 Signals
 ~ 14 data loading processes
 ~ 5.4 billion records / day
 ~ 250 GB / day
 → 90 TB / year throughput

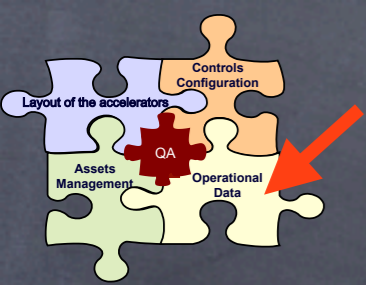


Data Consumers
Data Persistence
Data Providers



Accelerator Logging Service

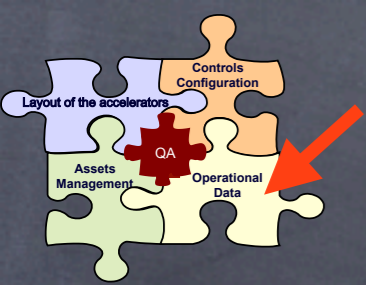
Simple schema, Many proprietary features



Accelerator Logging Service

Simple schema, Many proprietary features

```
create table data_numeric (  
  variable_id number,  
  utc_stamp   timestamp (9),  
  value       number,  
  constraint dn_pk primary key (variable_id, utc_stamp)  
)  
organization index compress 1  
partition by range (utc_stamp) (  
  partition dn_2008091900 values less than (  
    timestamp'2008-09-19 01:00:00'  
  ),  
  partition dn_2008091901 values less than (  
    timestamp'2008-09-19 02:00:00'  
  ),  
  .....  
);
```



Accelerator Logging Service

Simple schema, Many proprietary features

```
create table data_numeric (  
  variable_id number,  
  utc_stamp   timestamp (9),  
  value       number,  
  constraint dn_pk primary key (variable_id, utc_stamp)  
)  
organization index compress 1  
partition by range (utc_stamp) (  
  partition dn_2008091900 values less than (  
    timestamp '2008-09-19 01:00:00'  
  ),  
  partition dn_2008091901 values less than (  
    timestamp '2008-09-19 02:00:00'  
  ),  
  .....  
);
```

1,590 partitions
~2.7 trillion rows



Be wary of the users!



Big difference between how:

Big difference between how:

You think users will read/write data

Users tell you they intend to read/write data

Users actually read/write data

System misuse

can have severe

negative impact on stability

API-ONLY access

API-ONLY access

Instrument everything!



What is Instrumentation?

What is Instrumentation?

Capturing information about
system activity in real time
($\&$ over time)

What is Instrumentation?

Capturing information about
system activity in real time
($\&$ over time)

Who?

What?

Where?

How?

How Long?

We **know** what the system is doing

We **know** how the system is performing

Throughput & Response times

Instrumentation is everywhere
within logging service

Instrumentation is everywhere
within logging service

Every read/write capture & persist:

who, requested what (signals, time window, filters), from where (real client), using which products, how long it took, #records

Instrumentation is everywhere
within logging service

Every read/write capture & persist:

who, requested what (signals, time window, filters), from where (real client), using which products, how long it took, #records

Short-term details, long term aggregates

Instrumentation is everywhere within logging service

Every read/write capture & persist:

who, requested what (signals, time window, filters), from where (real client), using which products, how long it took, #records

Short-term details, long term aggregates

Performance analysis / tuning

Complex problem diagnosis

Capacity Planning

Internal Instrumentation

ORACLE Enterprise Manager
Grid Control 11g

Cluster: [acclug_cluster](#) > Cluster Database: [ACCLUG.com.ch](#) > Top Sessions > Database Instance: [ACCLUG.com.ch_ACCLUG2](#) > Top Activity >

Session Details: 383 (LHCLOG)

Collected From Target Feb 1, 2011 9:37:20 AM

View Data: Real Time: 15 Second Refresh | Refresh

Kill Session | Enable SQL Trace

General | **Activity** | Statistics | Open Cursors | Blocking Tree | Wait Event History | Parallel SQL

Drag the shaded box to change the time period for the detail section below.

Legend:

- SQL*Net more data from client
- log file sync
- db file sequential read
- CPU Used

Anchor shaded box to latest time | Go

Detail for Selected 5 Minute Interval

Start Time: Feb 1, 2011 9:29:28 AM | View: Show Aggregated Data | Run ASH Report

Sample Time	SQL ID	SQL Type	Plan Hash Value	Wait Event	P1 Value	P2 Value	P3 Value	Time Waited (mhu s)	Object	File	Block Number	Program	Module	Action	Client ID	QC Session ID	QC Instance ID
2/1/11 9:34:26 AM		UNKNOWN Q		log file sync	12404	301243609	0	18167	LHCLOG.DN_PK /ORA/dbs...ta_6n3kxs3d_dbf	265835	DataWriter	ENTER	loadNumericData	0	0	0	
2/1/11 9:33:16 AM	3pkrm4ccbknra	INSERT	0	db file sequential read	2169	71285	1	72556	LHCLOG.DN_PK /ORA/dbs...ta_6n3kxs3d_dbf	71285	DataWriter	ENTER	loadNumericData	0	0	0	
2/1/11 9:32:27 AM	3pkrm4ccbknra	INSERT	0	SQL*Net more data from client	1413697536	2	0	167140	LHCLOG.DN_PK /ORA/dbs...ta_6n3kxs3d_dbf	283997	DataWriter	ENTER	loadNumericData	0	0	0	
2/1/11 9:32:17 AM	3pkrm4ccbknra	INSERT	1575093030	CPU								DataWriter	ENTER	loadNumericData	0	0	0
2/1/11 9:32:16 AM	1f4h83gj20032	SELECT	4091179982	CPU								DataWriter	ENTER	checkVariables	0	0	0
2/1/11 9:31:01 AM	4hw28r7x7ub55	SELECT	4091179982	CPU								DataWriter	ENTER	checkVariables	0	0	0

General | **Activity** | Statistics | Open Cursors | Blocking Tree | Wait Event History | Parallel SQL

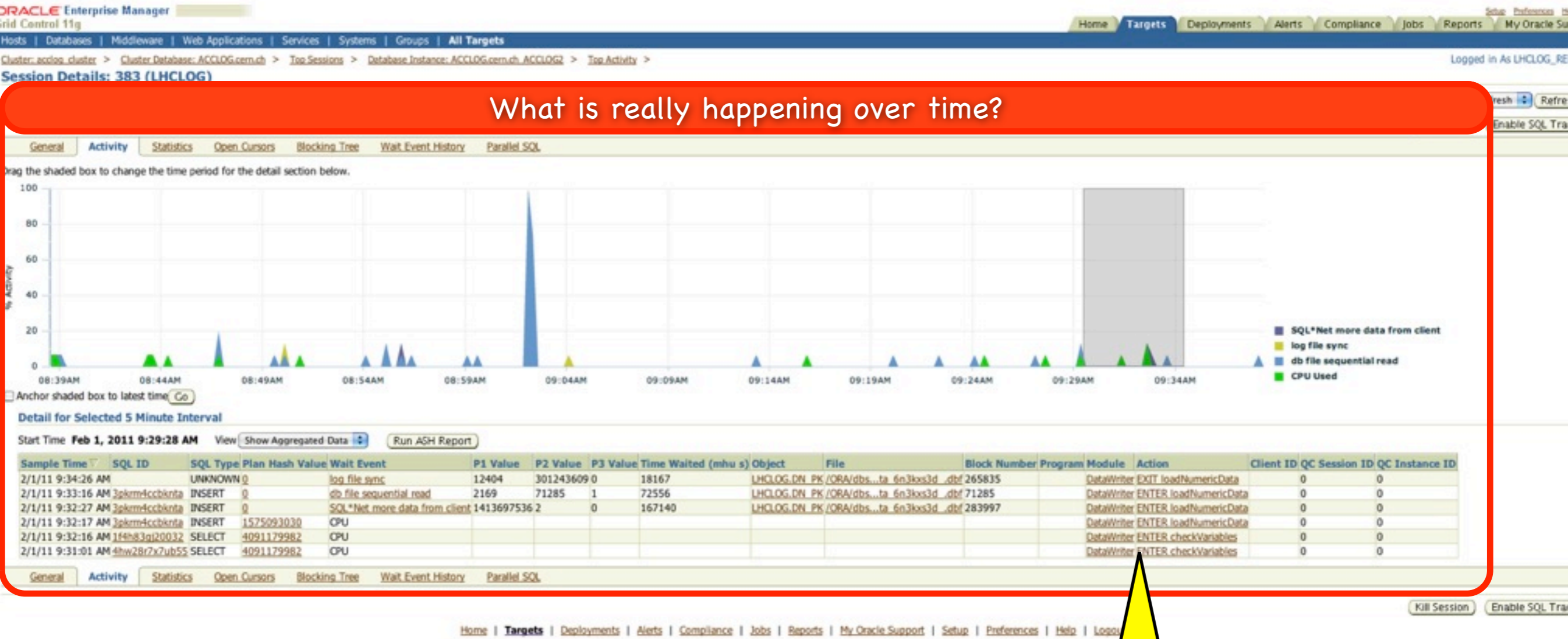
Kill Session | Enable SQL Trace

Home | **Targets** | Deployments | Alerts | Compliance | Jobs | Reports | My Oracle Support | Setup | Preferences | Help | Logout

Copyright © 1996, 2010, Oracle and/or its affiliates. All rights reserved.
Oracle is a registered trademark of Oracle Corporation and/or its affiliates.
Other names may be trademarks of their respective owners.
[About Oracle Enterprise Manager](#)

Internal Instrumentation

What is really happening over time?



Augment internal DB instrumentation with contextual information

Internal Instrumentation

ORACLE Enterprise Manager
Grid Control 11g


Hosts | Databases | Middleware | Web Applications | Services | Systems | Groups | All Targets

Cluster: [acclog_cluster](#) > Cluster Database: [ACCLOG.cern.ch](#) > Top Consumers > Database Instance: [ACCLOG.cern.ch ACCLOG2](#) >

Module: DataWriter

[Actions](#) | [Activity](#) | [Statistics](#)

Top Actions



Action	Percentage
ENTER loadNumericData	84.2%
ENTER checkVariables	9.9%
EXIT loadNumericData	5.4%
FINISHED	0.5%

Active Actions

View: [Active Actions](#)

Action	Activity (% for the last 5 minutes)	Aggregation Enabled	SQL Trace Enabled	D
ENTER loadNumericData	84.2	FALSE	FALSE	
ENTER checkVariables	9.9	FALSE	FALSE	
EXIT loadNumericData	5.4	FALSE	FALSE	
FINISHED	.5	FALSE	FALSE	

[Actions](#) | [Activity](#) | [Statistics](#)

[Home](#) | [Targets](#) | [Deplo](#)

Copyright © 1996, 2010, Oracle and/or its affiliates. All rights reserved.
Oracle is a registered trademark of Oracle Corporation and/or its affiliates.
Other names may be trademarks of their respective owners.
[About Oracle Enterprise Manager](#)

Internal Instrumentation

ORACLE Enterprise Manager
Grid Control 11g

Hosts | Databases | Middleware | Web Applications | Services | Systems | Groups | All Targets

How is time being spent within a Module?

Module: DataWriter

Actions | Activity | Statistics

Top Actions



Active Actions

View Active Actions

Action	Activity (% for the last 5 minutes)	Aggregation Enabled	SQL Trace Enabled	D
ENTER loadNumericData	84.2	FALSE	FALSE	
ENTER checkVariables	9.9	FALSE	FALSE	
EXIT loadNumericData	5.4	FALSE	FALSE	
FINISHED	.5	FALSE	FALSE	

Actions | Activity | Statistics

Home | [Targets](#) | [Deplo](#)

Copyright © 1996, 2010, Oracle and/or its affiliates. All rights reserved.
Oracle is a registered trademark of Oracle Corporation and/or its affiliates.
Other names may be trademarks of their respective owners.

[About Oracle Enterprise Manager](#)

To Summarize...

To Summarize...

Oracle DBMS relational features

- Accurately model Accelerator domain complexity
- Ensure data integrity

Well considered designs, &

significant usage of proprietary Oracle functionality

- deliver necessary performance
- satisfy other functional requirements

Questions?