

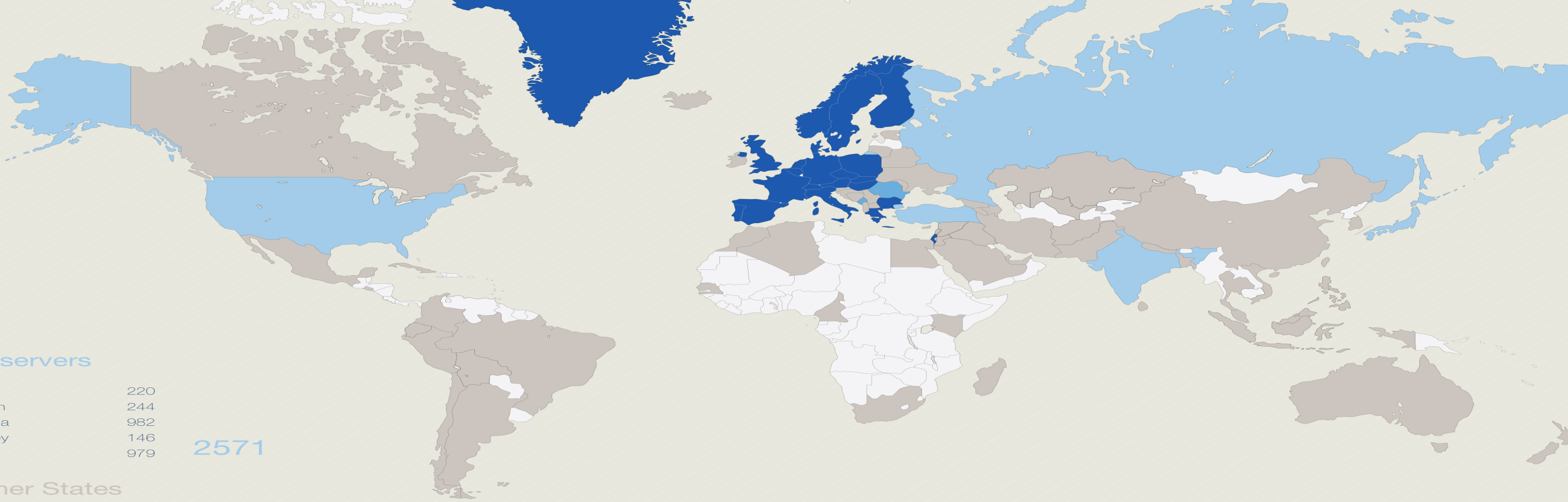
Oracle Big Data Discovery for CERN's Control Data

Antonio Romero Marin

What is CERN

- CERN - European Laboratory for Particle Physics
- Founded in 1954 by 12 countries for fundamental physics research in a post-war Europe
 - “Science for Peace”





2571

Observers

India	220
Japan	244
Russia	982
Turkey	146
USA	979

Other States

Afghanistan	1	El Salvador	1	Pakistan	41
Albania	2	Estonia	16	Palestine (O.T.),	4
Algeria	8	Georgia	36	Peru	8
Argentina	11	Gibraltar	1	Philippines	1
Armenia	25	Hong Kong	1	Saudi Arabia	3
Australia	25	Iceland	4	Senegal	1
Azerbaijan	8	Indonesia	1	Singapore	2
Bangladesh	4	Iran	28	Sint Maarten	2
Belarus	47	Ireland	22	Slovenia	27
Bolivia	3	Jordan	2	South Africa	16
Bosnia &		Kenya	1	Sri Lanka	5
Herzegovina	1	Korea, D.P.R.	1	Syria	2
Brazil	108	Korea Rep.	117	Thailand	12
Cameroon	1	Kuwait	1	T.F.Y.R.O.M.	1
Canada	134	Lebanon	12	Tunisia	6
Cape Verde	1	Lithuania	19	Ukraine	55
Chile	12	Luxembourg	4	Uzbekistan	4
China	280	Madagascar	4	Venezuela	9
China (Tapei)	45	Malaysia	15	Viet Nam	9
Colombia	30	Mauritius	1	Zimbabwe	2
Croatia	35	Mexico	64		
Cuba	7	Montenegro	3		
Cyprus	16	Morocco	12		
Ecuador	3	Nepal	5		
Egypt	19	New Zealand	7		

1415

Member States

Austria	99	Greece	152	Slovakia	88
Belgium	106	Hungary	68	Spain	337
Bulgaria	75	Israel	51	Sweden	75
Czech Republic	202	Italy	1686	Switzerland	180
Denmark	53	Netherlands	153	United Kingdom	640
Finland	87	Norway	61		
France	751	Poland	229		
Germany	1150	Portugal	109		

6352

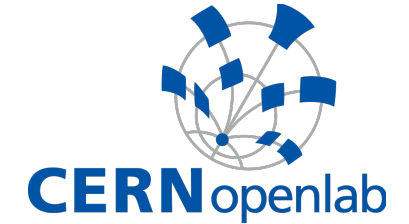
Candidate for Accession

Romania	118
---------	-----

Associate Members in the Pre-stage to Membership

Serbia	41
--------	----

CERN openlab



- Public-private partnership between CERN and leading ICT companies and research institutes
- Accelerate cutting-edge solutions for the worldwide LHC community and wider scientific research.
- Designed to create and disseminate knowledge
 - Publication of reports and articles
 - Workshops or seminars
 - CERN openlab Student Programme



CERN Mission



Discoveries

Seeking answers to questions about the Universe



International Cooperation

Uniting bringing nations together through science



Science & Education

Training tomorrow's scientists and engineers



Science & Technology

Advancing the frontiers of technology

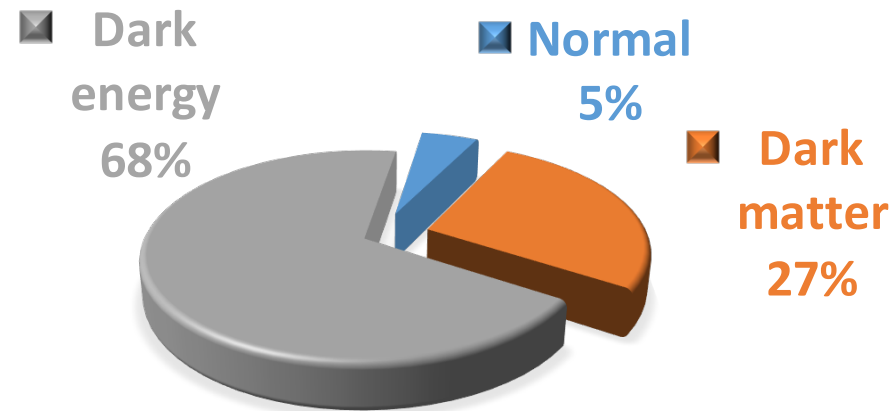


What is the Universe made of?

How does it work?

Fundamental Research

- Why do particles have mass?
- What is dark matter and dark energy?



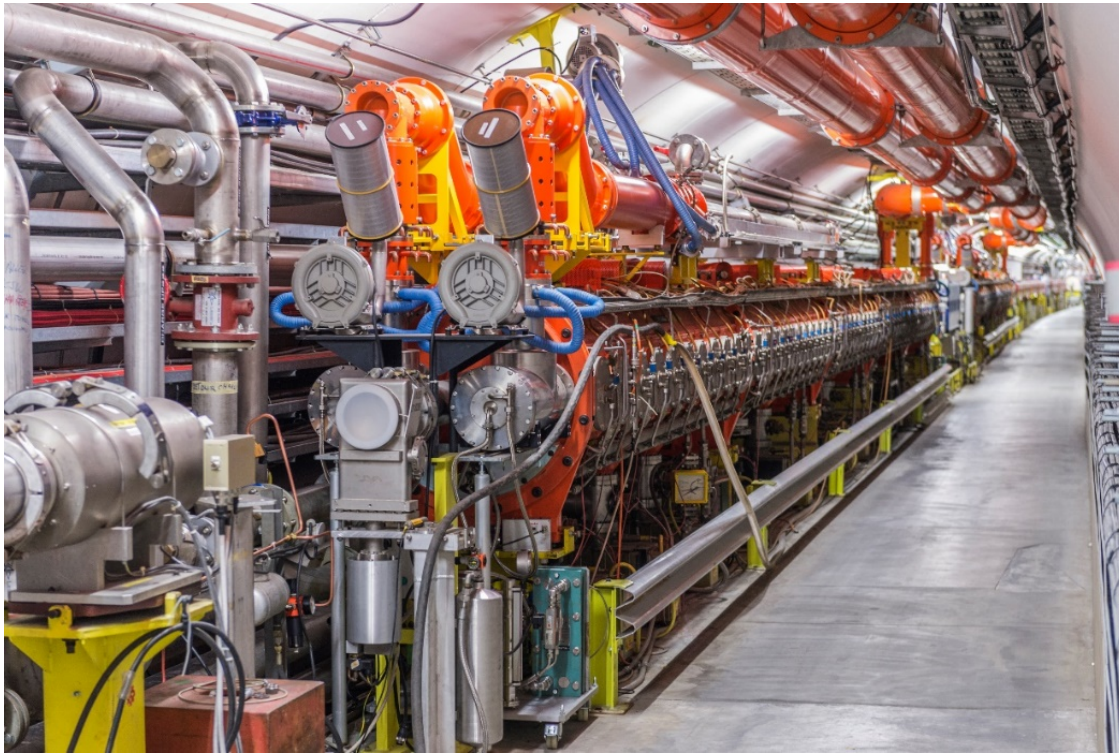
Composition of the Universe

- Why is there far more matter than antimatter in the universe?
 - Big Bang should have created equal amounts of matter and antimatter

CERN Instruments

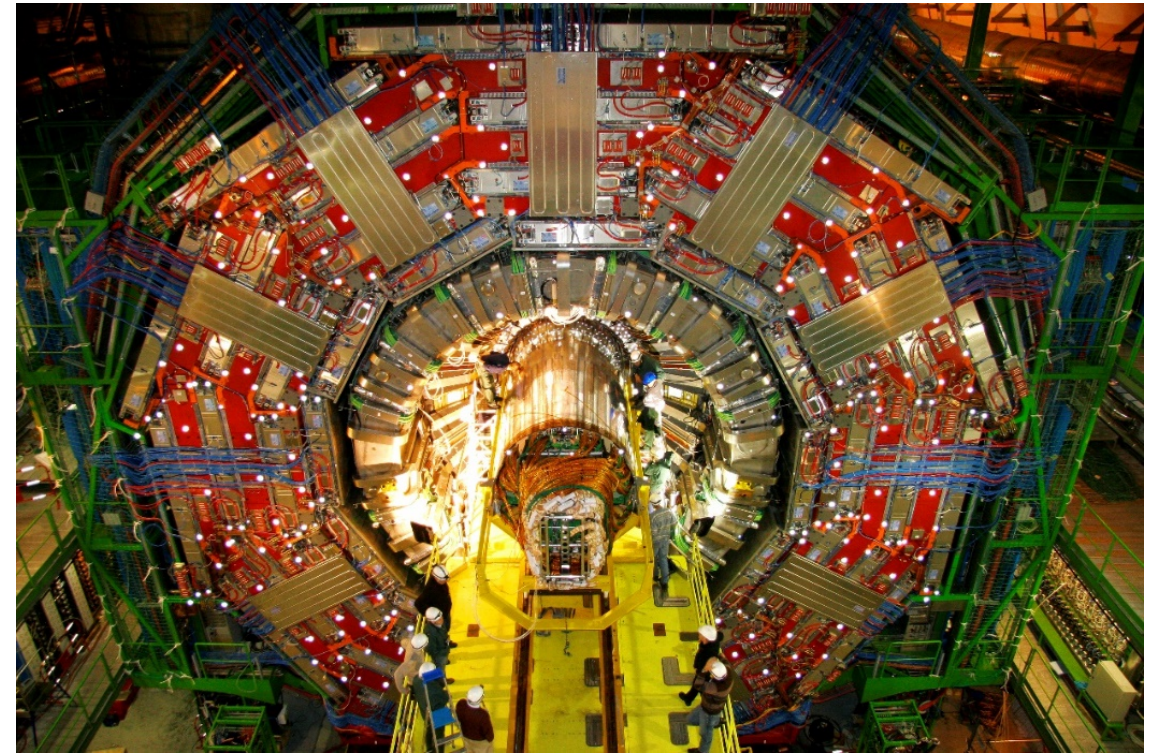
Accelerators

Boost particles to high energies and speed to collide



Detectors

Observe and record the results of these collisions



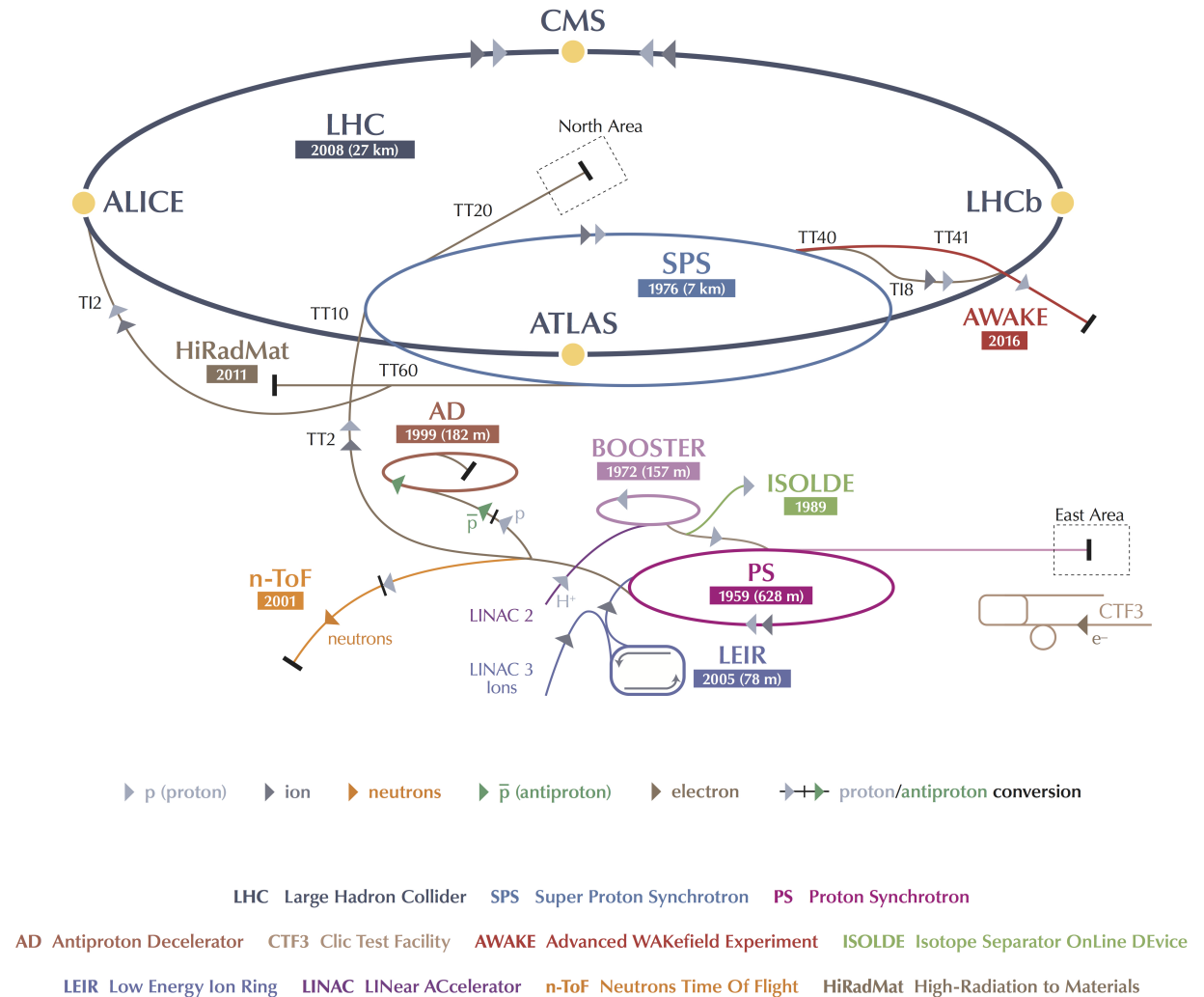
The Large Hadron Collider (LHC)



- **Largest machine** in the world
 - 27km, 6000+ superconducting magnets
- **600 million** collisions per second
 - Generating approximately one petabyte of data per second
- One of the **coldest places** on Earth
 - Main magnets operate at a temperature of 1.9 K (-271.3°C)
- **Hottest spot** in the galaxy
 - During Lead ion collisions create temperatures 100000x hotter than the heart of the sun

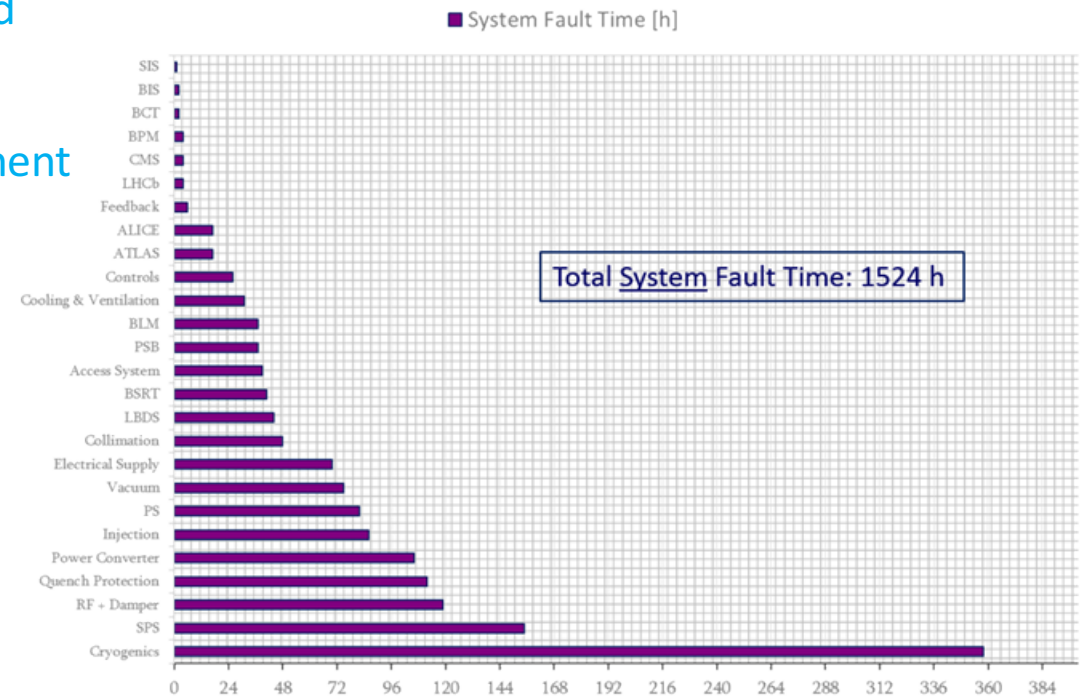
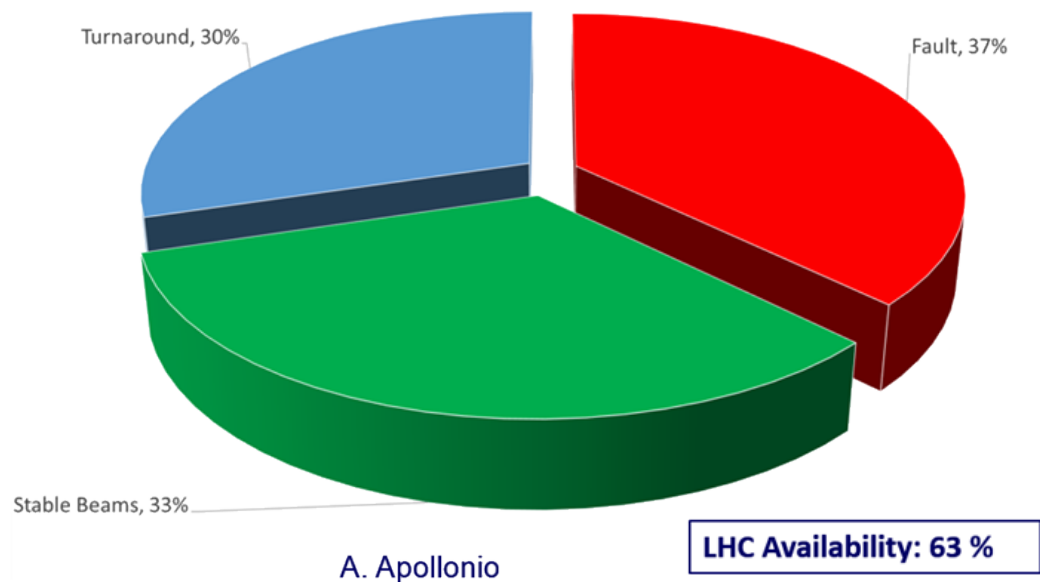
CERN Accelerator Complex

- Control and operations
 - Millions of sensors, signals
 - Large number of control devices
 - Equipment
- Monitoring and logging
- Supporting IT infrastructure
 - Databases
 - Network
 - Services



Data Analytics Challenges

- Some faults cannot be avoided
- Decrease the availability for running physics
- Corrective interventions needed
 - Fix equipment or components after failure has occurred
- Preventive maintenance is not enough
 - Does not take into account the condition of the equipment



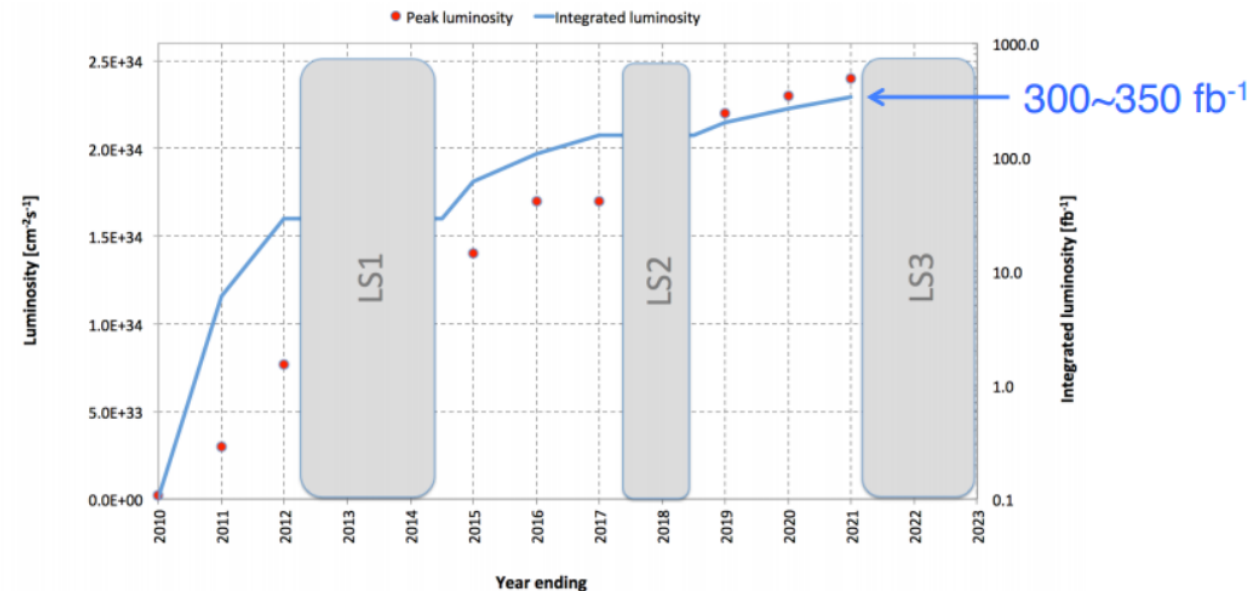
B. Todd, Evian 2012

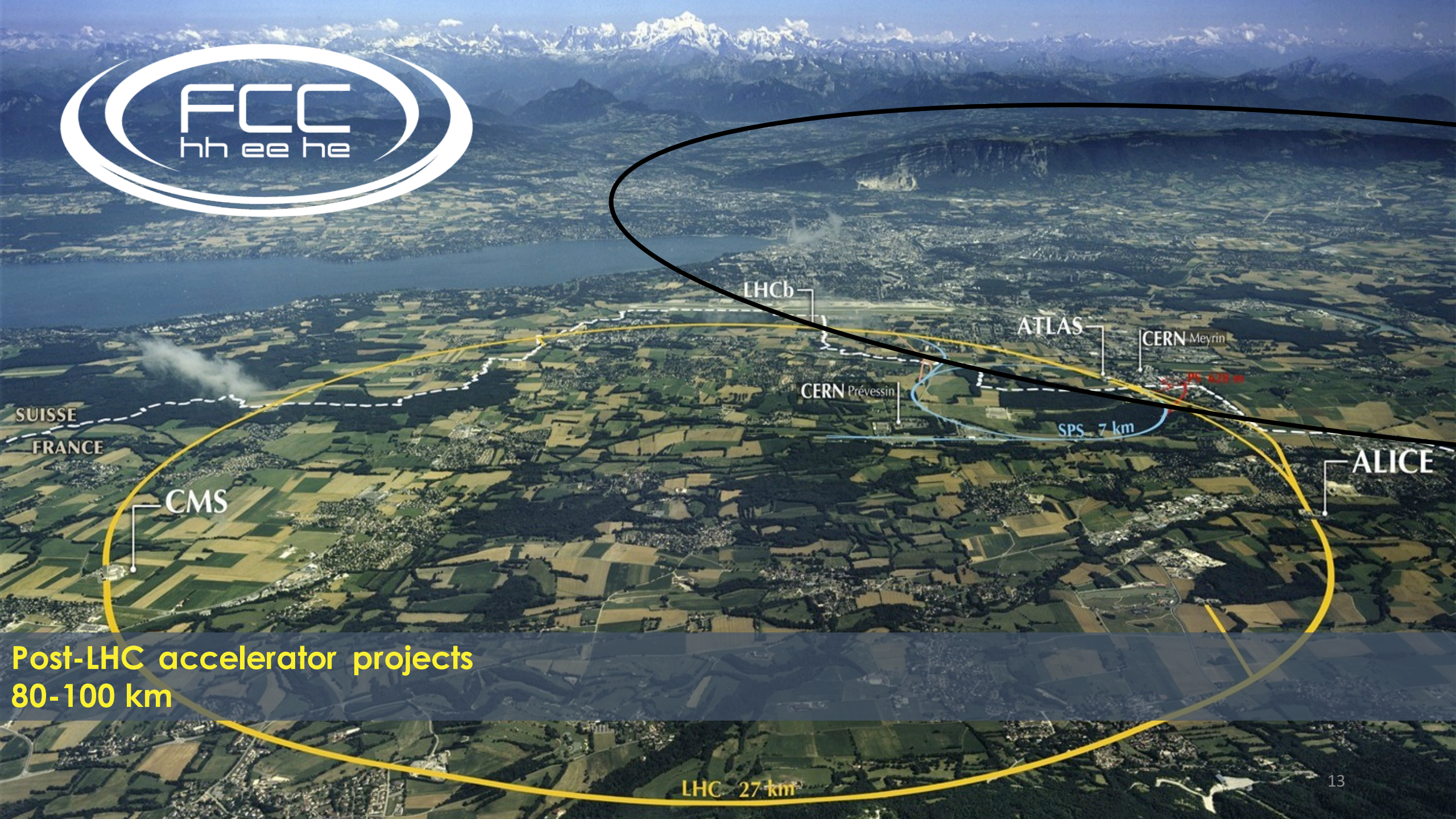
A look into the Future

- LHC upgrades will further increase luminosity
 - Computing resources needs will be higher
 - Data generated will increase drastically

Table 17: HLT Rate Evolution

Hz	ALICE	ATLAS	CMS	LHCb
2012	400 Hz 330 MB/s (p-p) 540 MB/s (p-Pb)	550 Hz 440 MB/s	460+360 Hz 328 MB/S	5000 Hz 300 MB/s
2015	500 Hz 525 MB/s (p-p) 810 MB/s (p-Pb) 3750 MB/s (Pb-Pb)	1000 Hz 800-1000 MB/s	1000 Hz 600 MB/S	10000 Hz 750 MB/s





LHCb

ATLAS

CERN Meyrin

CERN Prévessin

SPS 7 km

PS 6.18 km

SUISSE
FRANCE

CMS

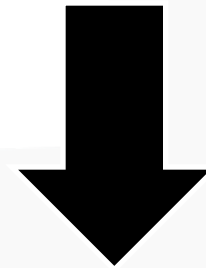
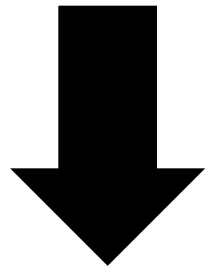
ALICE

Post-LHC accelerator projects
80-100 km

LHC 27 km

Data Analytics Objective

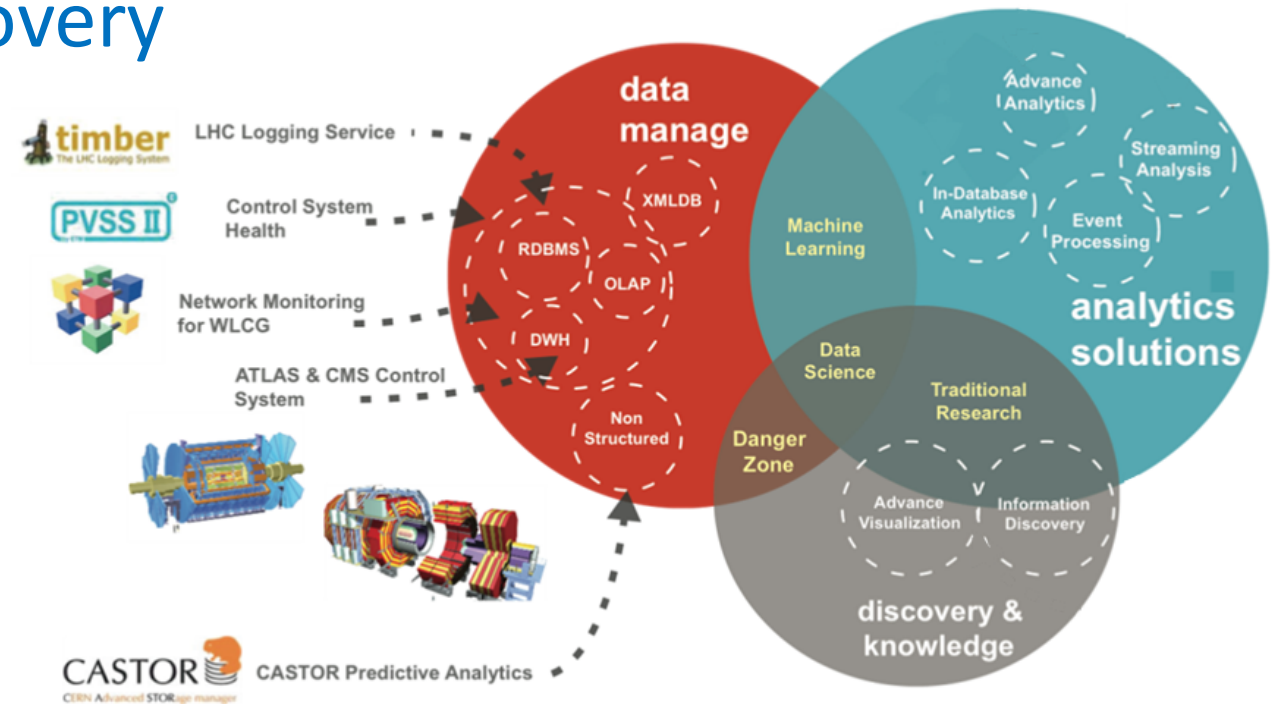
Control and Monitoring Systems



Intelligent, Predictive and Proactive Systems

Areas of investigation

- Predictive maintenance and system optimization
- Data extraction, transformation and loading (ETL)
- Data Visualization and Discovery



Use Case - FCC RAMS studies

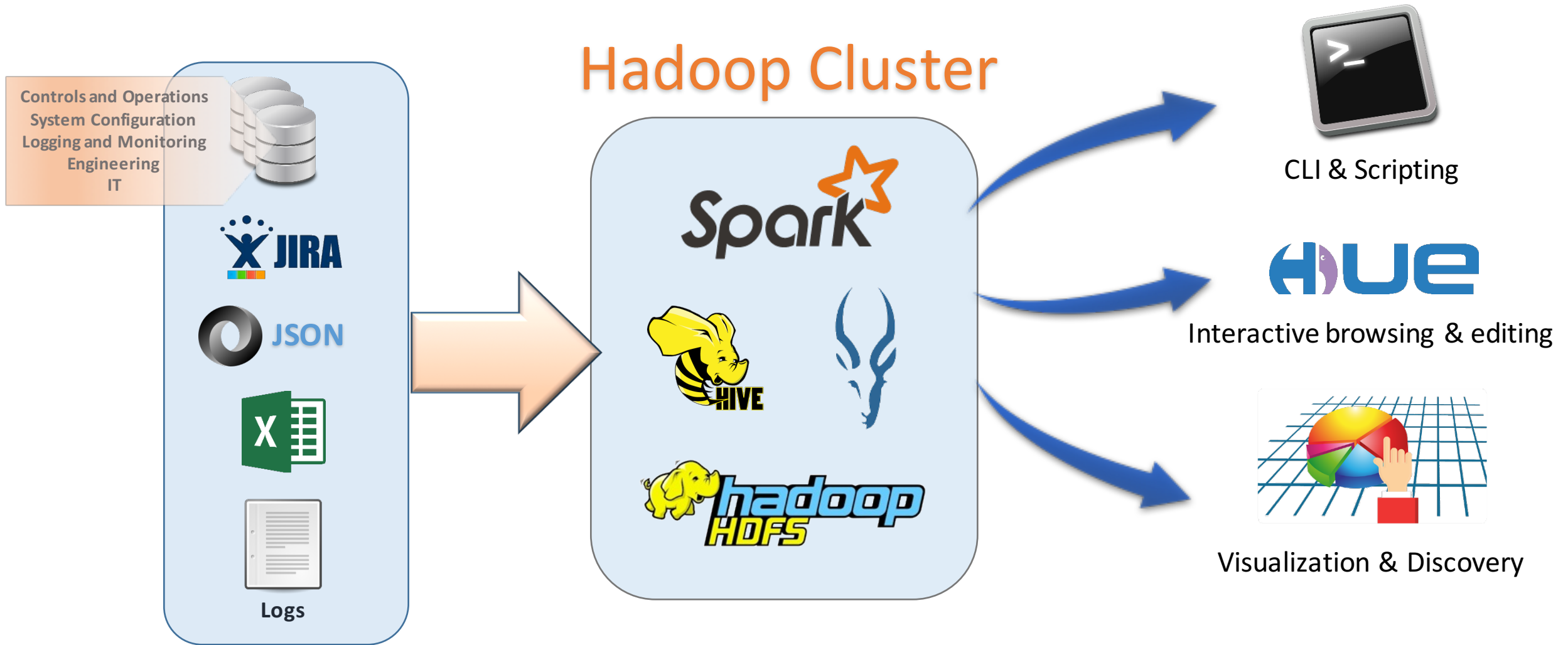
- Reliability, Availability, Maintainability and Safety (RAMS) studies for the Future Circular Collider (FCC)
- Study and increase the reliability and availability of the LHC
- Use RAMS findings to assess the feasibility of the needs of FCC
- Data distributed across multiple sources
 - Operations e-logbook
 - Accelerator Fault Tracking project
 - Accelerator logging service
 - Accelerator schedules
 - Cryogenics
 - Add more in the future
 - Vacuum, Power Converters, etc.



Requirements

- Flexible
 - Data => structured, semi-structured and non-structured, data editing
 - Use => Interactive, CLI & Scripting
- ETL functionalities
- Scalable
 - Data is foreseen to increase significantly (+datasets)
 - Processing
- Powerful
 - Browse data
 - Correlate Information
 - Visualization
 - Analytics

Hadoop based solution



HUE – Hadoop User Experience

- Hue is an open source suite of web-based applications for analyzing data with any Apache Hadoop
- It features:
 - SQL Editors for Hive, Impala, MySQL, PostGres, Sqlite and Oracle
 - Dynamic search dashboards for Solr
 - Spark Notebooks
 - Browsers for YARN, HDFS, Hive table Metastore, HBase, ZooKeeper
 - Pig Editor, Sqoop2, Oozie workflows Editors and Dashboards
 - Wizards to import data into Hadoop

HUE – Hadoop User Experience

The screenshot displays the HUE interface with two charts. The top chart is a bar chart showing the count of variables for various POSST identifiers. The bottom chart is a line chart showing the average value of variables over time for a specific POSST identifier.

Top Chart: Bar Chart

SQL Query:

```
select variable_name, count(*) as numrows
from lhlog_cryo
group by variable_name
order by numrows desc
limit 10
```

Chart Configuration:

- X-Axis: variable_name
- Y-Axis: numrows
- Sorting: Descending

variable_name	numrows
QRLAB_23L1_GT943.POSST	6,887,08
QRLHA_05R4_GT938.POSST	6,000,000
QRLAB_15L6_GT947.POSST	5,800,000
QRLHA_05R4_GT937.POSST	5,000,000
QRLFE_04L8_GT930.POSST	5,000,000
QRLCC_07L4_GT947.POSST	4,800,000
ATLAS:LUMI_TOT_INST	4,600,000
QRLDE_06R8_GT931.POSST	4,500,000
QRLHA_05L4_GT935.POSST	4,400,000
QRLEA_06L2_GT931.POSST	4,300,000

Bottom Chart: Line Chart

SQL Query:

```
select avg(value) as value, extract(utc_timestamp, "hour") as hour_
from lhlog_cryo
where utc_timestamp > "2015-06-12"
and utc_timestamp < "2015-06-13"
and variable_name = "QRLAB_23L1_GT943.POSST"
group by hour_
order by hour_
```

Chart Configuration:

- X-Axis: hour_
- Y-Axis: value
- Sorting: Ascending

hour_	value
0	12.33722
1	12.337
2	12.325
3	12.326
4	12.337
5	12.318
6	12.313
7	12.321
8	12.324
9	12.320
10	12.320
11	12.327
12	12.310
13	12.318
14	12.322
15	12.323
16	12.326
17	12.323
18	12.310
19	12.316
20	12.321
21	12.328
22	12.327
23	12.323

HUE – Hadoop User Experience

- One tool to use multiple Hadoop components
- Easy to use
- Compatible with multiple versions, open source
- Extensible
- But
 - Requires language knowledge to explore and transform the data
 - Limited for Data Discovery

Data Discovery

- Interactive and visual analytics
 - Find hidden patterns
 - Get **new insights**
- Intended to be used by the end users
 - Enabling them to use their intuition and knowledge of the data
- Powerful customization of dashboards and visualizations
 - Without intervention of IT
- Integrate multiple data sources
 - Analyze information of any type and any source

Oracle Big Data Discovery Overview

- **Data Exploration & Discovery**
 - Interactive catalog of all data
 - Assess attribute statistics, data quality and outliers
 - Quick data exploration or create dashboards and applications
- **Data Transformation with Spark in Hadoop**
 - Apply built-in transformations or write your own scripts
 - Data Enrichment
 - Text: Entity extraction, relevant terms, sentiment, language detection
 - Geographical information: address, IP, reverse
 - Preview results, undo, commit and replay transforms
- **Collaborative environment**
 - Share and bookmarks
 - Create and share transformed datasets

Components

- BDD Data Processing (Spark on YARN)
 - Hive Table Detector
 - Profiling and sampling
 - Transformations and enrichments
 - Refresh/incremental update datasets automatically or manually
- Dgraph (In-Memory Discovery Indexes)
 - In-memory, columnar, multi-core architecture
- Web Studio
 - Catalog, explore, transform and discover UI's

Architecture overview

cloudera®

CDH 5.5.1
16 nodes, 24 GB ram
Intel Xeon L5520 @ 2.27GHz
165 TB HDFS

Oracle Big Data Discovery
Libraries + Hive table detector



Resource Management (YARN)

Data Storage



Data Integration



C
o
o
r
d
i
n
a
t
i
o
n

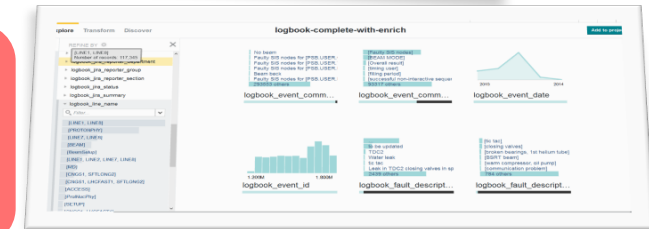
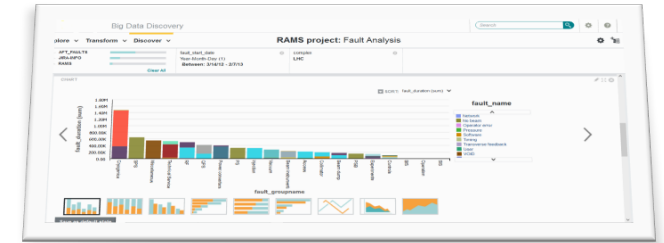


Dgraph
In-Memory Columnar
Database

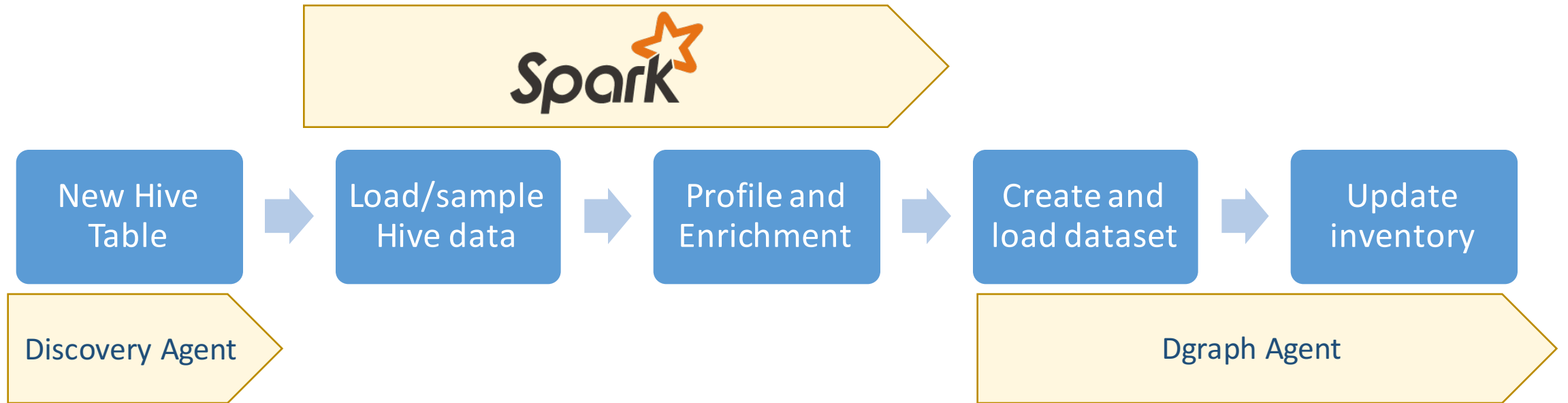
ORACLE®
EXALYTICS



4x Xeon E7-8895 v2 (15 cores each)
2 TB RAM
4.8 TB Flash + 6 x 1.2 TB 10K HDD



Data Processing Workflow

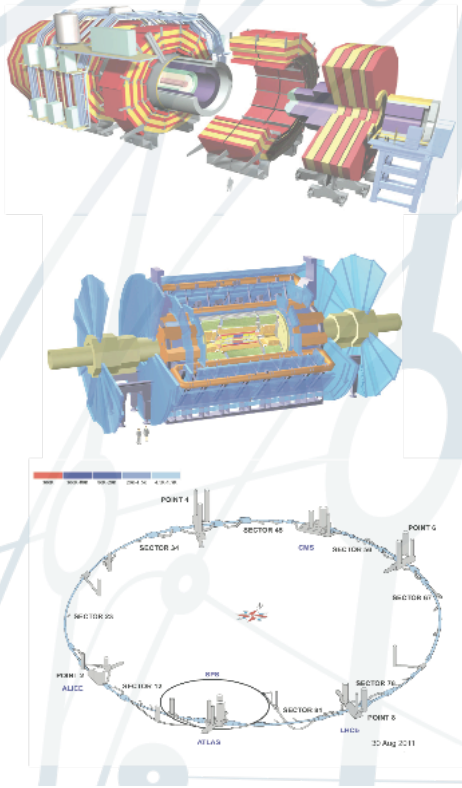


Important Technical Features

- Supports HDP and CDH Hadoop distributions
- Kerberos support
- Spark on YARN
- Data set and project level security in Studio
- Incremental updates and refresh datasets
- Applications and dashboards
- Ability to publish and share transformation scripts
- Custom visualization support (Javascript, D3, EQL)

Use Case: FCC RAMS

Scenario





- Power Converters
- Cryogenics
- Machine Protection
- Accelerator Major Events
- Accelerator Fault Tracking
- Accelerator Logging
- Operations logbook

•NoSQL
•XML
•JSON
•Text
•RDBMS



Datasets Catalog




 **7 Projects**
View all

 **11 Data Sets**
View all








 **Add Data Set**

- Refine By
- USAGE
 - Created By Me
 - CONTENT
 - Contains Dates
 - Contains Locations
 - METADATA
 - Project Author
 - Data Set Author
 - Project Tags
 - Data Set Tags
 - Last Modified
 - Number of Records
 - Number of Attributes
 - Data Source Type

Recently Viewed Data Sets

aft_faults Data Source: fcc_rams.aft_faults (1.8k records)  Preview	op_logbook Data Source: fcc_rams.op_logbook (1.1M records)  Preview	cms_lhcfills Data Source: cms_lhcfills.csv (621 records)  Preview
--	--	---

Most Popular Data Sets [View More](#)


op_logbook Data Source: fcc_rams.op_logbook (1.1M records)  Preview	apache_mwctl_prod_db... Data Source: apache-mwctl-test-db-a0... (183 records) Preview	apache_mwctl_prod_db... Data Source: apache-mwctl-test-db-a0... (745 records) Preview	aft_faults Data Source: fcc_rams.aft_faults (1.8k records)  Preview	cms_lhcfills Data Source: cms_lhcfills.csv (621 records)  Preview	aft_cardiogram Data Source: fcc_rams.aft_cardiogram (1.2M records)  Preview	cms_runs_for_fill Data Source: cms_runs_for_fill.csv (2.5k records) Preview
cms_tmb_rates Data Source: cms_tmb_rates.csv (820.6k records)  Preview	high_level_summary_sc... Data Source: fcc_rams.high_level_sum... (63 records)  Preview	lhclog_cryo Data Source: fcc_rams.lhclog_cryo (1.1B records)  Preview New	naive_ml_prediction Data Source: naive_ml_prediction.csv (1.2M records) Preview			

lhclog_cryo (1,078,271,323 records)

Data Set Info **Used in Projects (1)** Related Data sets (4)

 **Tags** To add tags, click the Tags button at left.

Attributes


Data source: fcc_rams.lhclog_cryo
Data source type: Hive
Hive Table name: fcc_rams.lhclog_cryo
Created on: 1/17/2016 4:52:14 PM (UTC)
Access: Public [edit](#)
Data set key: edp_cli_edp_bfb068c8-3280-4bcf-a107-87a3d4ef7e04

Actions

- [Explore](#)
- [Add to project](#)
- [Edit tags](#)
- [Reload data set](#)
- [Delete](#)

Summary

0 Views
Last Updated
1/17/2016 4:52:16 PM (UTC)

Newly Added Data Sets [View More](#)

lhclog_cryo Data Source: fcc_rams.lhclog_cryo (1.1B records)  Preview New	op_logbook Data Source: fcc_rams.op_logbook (1.1M records)  Preview	apache_mwctl_prod_db... Data Source: apache-mwctl-test-db-a0... (745 records) Preview	apache_mwctl_prod_db... Data Source: apache-mwctl-test-db-a0... (183 records) Preview	cms_tmb_rates Data Source: cms_tmb_rates.csv (820.6k records)  Preview	cms_runs_for_fill Data Source: cms_runs_for_fill.csv (621 records) Preview	cms_lhcfills Data Source: cms_lhcfills.csv (621 records)  Preview
---	--	---	---	---	--	--

Quick Data Exploration

OP_LOGBOOK

Clear All

REFINE BY

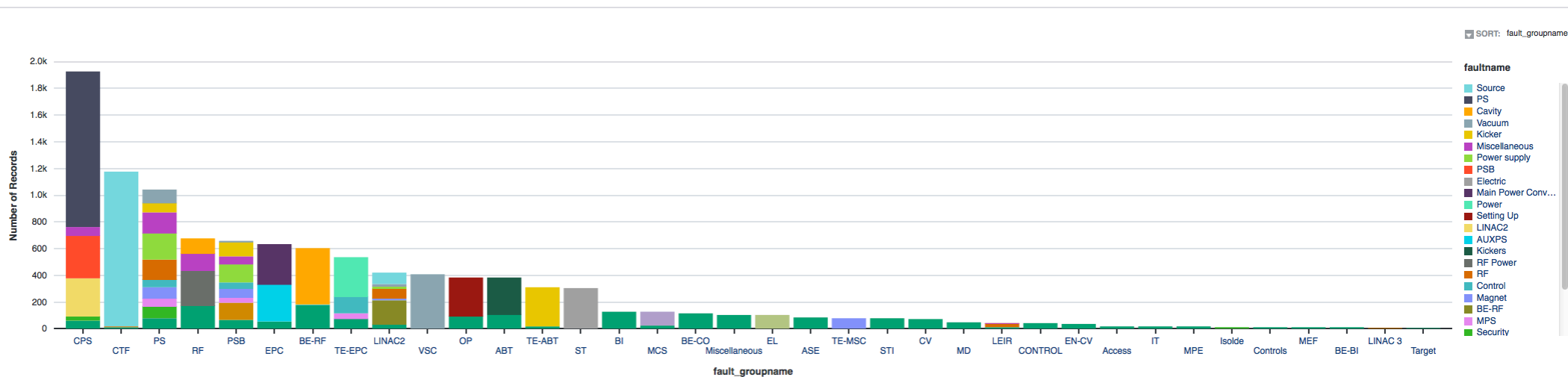
- element
- event_comment
- event_comment_terms
- event_date
- event_id
- fault_created
 - Between: 5/4/14 - 11/30/15
 - 10555 results
 - Update
- fault_description
- fault_groupname
- fault_id
- faultname
 - Filter...
 - Source
 - PS
 - Cavity
 - Vacuum
 - Kicker
 - Miscellaneous
 - Power supply
 - PSB
 - Electric
 - Main Power Converters
 - Power
 - Setting Up
 - LINAC2
 - AUXPS
 - Kickers
 - RF Power
 - RF
 - Control
 - Magnet
 - BE-RF
 - More
 - Select All
- linenames
- logbook_complex
 - Filter...
 - PS Complex

fault_created Year-Month-Day (1)
Between: 5/4/14 - 11/30/15

logbook_complex TESTS

faultname is Not Null VOID

fault_groupname faultname



1.1M Records, 10.6k Filtered Records, 15 Attributes

FAVORITES DATA TYPE NAME HIDDEN

All Attributes

Sort: By name

<p>#285 others</p> <p>Thermo-ionic Gun</p> <p>Heater control and HV breakdown... to be updated</p> <p>TBIU T6</p> <p>BLM</p> <p>element</p>	<p>No beam detected for SFTPRO1</p> <p>No beam detected for SFTPRO2</p> <p>Beam back for SFTPRO1</p> <p>No beam detected for SFTION2</p> <p>Start of the LHC filling period</p> <p>No beam detected for SFTION3</p> <p>6443 others</p> <p>event_comment</p>	<p>No beam</p> <p>SFTPRO1</p> <p>SFTPRO2</p> <p>Access</p> <p>Sequence</p> <p>Beam back</p> <p>11198 others</p> <p>event_comment_terms</p>	<p>2014</p> <p>2015</p> <p>event_date</p>	<p>0</p> <p>2292702</p> <p>2292684</p> <p>2440775</p> <p>2292679</p> <p>2440780</p> <p>9241 others</p> <p>event_id</p>	<p>2014</p> <p>2015</p> <p>fault_created</p>	<p>568 others</p> <p>to be updated</p> <p>Since Mon 3/8/15. Problem not yet to be updated</p> <p>No fault reported, no beam in ring.. Dipole 30330</p> <p>Power Cell - EMD114*9 blown-up</p> <p>7744478</p> <p>1044683</p> <p>1044424</p> <p>1039957</p> <p>1039372</p> <p>1039959</p> <p>3130 others</p> <p>fault_description</p>	<p>fault_id</p>
---	---	--	---	--	--	--	-----------------

Data Transformation UI - ETL

Transformation Editor

Use refinement state as a conditional statement

Enable automatic typeahead + Functions + Attribute

```
1 diffDates(op_end_time, creation_time_utc, SECONDS)
```

Configure Output Settings

- Apply to "op_duration"
- Create New Attribute

New Attribute Name
op_duration_seconds

Data Type
Double

Single Assign

Cancel Preview Add to Script

TRANSFORM SCRIPT

- elogbook_fault_id - Transform
- fault_id - Create
- fault_classification_id - Transform

Commit to Project

1.8k Records 1.8k Filtered Records 21 Attributes

FAVORITES DATA TYPE NAME HIDDEN

All Attributes

Sort: By preview order

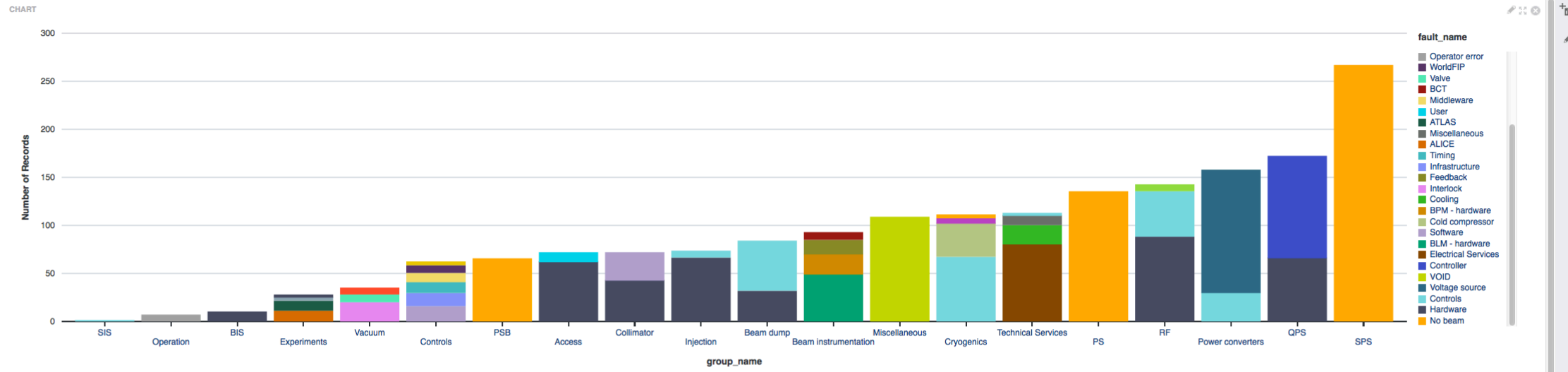
fault_id	fault_name	fault_state	group_name	is_root_cause	# op_duration	# op_duration_seconds	op_end_time	prevents_beam_op	prevents_injection	start_time
15160		OP_ENDED		N	4,266	10,932	2015-04-14 08:28:38 UTC	N	Y	2015-04-14 07:17:32 UTC
15223		CANCELLED		N				N	N	2015-04-15 06:30:00 UTC
secto... 3761		OP_ENDED		N	21,588	47,292	2014-11-24 17:30:38 UTC	N	N	2014-11-24 11:30:50 UTC
f 3765		OP_ENDED		N	99,735	66,405	2014-11-25 15:17:19 UTC	N	N	2014-11-24 11:35:04 UTC
roup.... 1909	Feedback	OP_ENDED	Beam instrumentation	N	66	56,455,621	2013-02-03 04:20:39 UTC	N	N	2013-02-03 04:19:33 UTC
condit... 3781		OP_ENDED		N	78,786	-9,846	2014-11-25 09:23:53 UTC	N	N	2014-11-24 11:30:47 UTC
roup.... 1908	No beam	OP_ENDED	SPS	N	20,870	56,457,295	2013-02-03 03:52:45 UTC	N	N	2013-02-02 22:04:55 UTC
roup.... 1924	No beam	OP_ENDED	SPS	N	147	55,550,285	2013-02-13 15:49:35 UTC	N	N	2013-02-13 15:47:08 UTC
roup.... 1907	Infrastructure	OP_ENDED	Controls	N	6,760	56,538,706	2013-02-02 05:15:54 UTC	N	N	2013-02-02 03:23:14 UTC
roup.... 1471	VOID	OP_ENDED	Miscellaneous	N	2,614	73,416,728	2012-07-21 20:55:32 UTC	N	N	2012-07-21 20:11:58 UTC
roup.... 1923	Hardware	OP_ENDED	Access	N	3,985	55,624,990	2013-02-12 19:04:30 UTC	N	N	2013-02-12 17:58:05 UTC
roup.... 1889	No beam	OP_ENDED	SPS	N	4,292	56,956,446	2013-01-28 09:13:34 UTC	N	N	2013-01-28 08:02:02 UTC
roup.... 1906	No beam	OP_ENDED	PSB	N	8,765	56,548,197	2013-02-02 02:37:43 UTC	N	N	2013-02-02 00:11:38 UTC
roup.... 117	Hardware	OP_ENDED	RF	N	3,750	137,030,838	2010-07-16 14:20:22 UTC	N	N	2010-07-16 13:17:52 UTC
roup.... 1470	No beam	OP_ENDED	PSB	N	1,322	73,419,672	2012-07-21 20:06:28 UTC	N	N	2012-07-21 19:44:26 UTC
roup.... 1922	Hardware	OP_ENDED	Injection	N	14,065	55,629,441	2013-02-12 17:50:19 UTC	N	N	2013-02-12 13:55:54 UTC
roup.... 1888	No beam	OP_ENDED	PSB	N	7,090	56,970,250	2013-01-28 05:23:30 UTC	N	N	2013-01-28 03:25:20 UTC
roup.... 1905	Controls	OP_ENDED	Beam dump	N	2,773	56,556,800	2013-02-02 00:14:20 UTC	N	N	2013-02-01 23:28:07 UTC
roup.... 1774	Cold compressor	OP_ENDED	Cryogenics	N	3,307	63,595,623	2012-11-12 13:00:37 UTC	N	N	2012-11-12 12:05:30 UTC
roup.... 116	User	OP_ENDED	Access	N	5,859	137,044,782	2010-07-16 10:27:58 UTC	N	N	2010-07-16 08:50:19 UTC
roup.... 1469	Controls	OP_ENDED	Beam dump	N	4,625	73,421,636	2012-07-21 19:33:44 UTC	N	N	2012-07-21 18:16:39 UTC
roup.... 1921	Hardware	OP_ENDED	RF	N	20,968	55,624,990	2013-02-12 19:04:30 UTC	N	N	2013-02-12 13:15:02 UTC
roup.... 1887	No beam	OP_ENDED	SPS	N	2,107	57,082,801	2013-01-26 22:07:39 UTC	N	N	2013-01-26 21:32:32 UTC
roup.... 1633	Hardware	OP_ENDED	RF	N	6,148	67,696,057	2012-09-26 02:00:03 UTC	N	N	2012-09-26 00:17:35 UTC
roup.... 1904	No beam	OP_ENDED	SPS	N	318	56,594,555	2013-02-01 13:45:05 UTC	N	N	2013-02-01 13:39:47 UTC
roup.... 1773	Hardware	OP_ENDED	RF	N	10,413	63,601,399	2012-11-12 11:24:21 UTC	N	N	2012-11-12 08:30:48 UTC
roup.... 115	Controls	OP_ENDED	Power converters	N	3,815	137,052,743	2010-07-16 08:15:17 UTC	N	N	2010-07-16 07:11:42 UTC
roup.... 1468	No beam	OP_ENDED	SPS	N	2,733	73,441,164	2012-07-21 14:08:16 UTC	N	N	2012-07-21 13:22:43 UTC

Discovery Applications

FCC RAMS v1: Faults

REFINE BY

- AFT_CARDIOGRAM
- AFT_FAULTS
 - creation_time_utc
 - description
 - element
 - elogbook_fault_id
 - fault_classification_id
 - fault_classification_name
 - fault_creation_source_name
 - fault_description
 - fault_id
 - fault_name
 - fault_state
 - group_name
 - is_root_cause
 - op_duration
 - op_end_time
 - prevents_beam_op
 - prevents_injection
 - start_time
 - system_name
- Other
 - expert_duration
 - expert_end_time



RESULTS TABLE

General ▾ 0 RECORDS SELECTED VIEW OPTIONS ▾ ACTIONS ▾

<input type="checkbox"/>	group_name	op_end_time (Year-Mo-...)	element	prevents_beam_op	fault_creation_source_...	elogbook_fault_id	fault_description	fault_name	creation_time_utc (Yea...	fault_classification_id	is_root_cause	fault_state	description
<input type="checkbox"/>		4/14/15		N	LHC Logbook	1041380	null		4/14/15	4	N	OP_ENDED	
<input type="checkbox"/>				N	AFT Web Application		test from Ben		4/15/15	8	N	CANCELLED	
<input type="checkbox"/>		11/24/14		N	AFT Web Application		PLC problem in Pt 4, all ...		11/25/14	5	N	OP_ENDED	
<input type="checkbox"/>		11/25/14		N	AFT Web Application		all patrols lost in point 4		11/26/14	5	N	OP_ENDED	
<input type="checkbox"/>	Beam instrumentation	2/3/13	OFSU crash	N	LHC Logbook	1039013	{\"fault_description\": {\"gro...	Feedback	11/18/14	1	N	OP_ENDED	
<input type="checkbox"/>		11/25/14		N	AFT Web Application		interlock access during c...		11/25/14	1	N	OP_ENDED	
<input type="checkbox"/>	SPS	2/3/13		N	LHC Logbook	1039002	{\"fault_description\": {\"gro...	No beam	11/18/14	1	N	OP_ENDED	
<input type="checkbox"/>	SPS	2/13/13		N	LHC Logbook	1039123	{\"fault_description\": {\"gro...	No beam	11/18/14	1	N	OP_ENDED	
<input type="checkbox"/>	Controls	2/2/13	logging	N	LHC Logbook	1038991	{\"fault_description\": {\"gro...	Infrastructure	11/18/14	1	N	OP_ENDED	Connection to LHC-OP-L
<input type="checkbox"/>	Miscellaneous	7/21/12	ofc down	N	LHC Logbook	1034471	{\"fault_description\": {\"gro...	VOID	11/18/14	1	N	OP_ENDED	
<input type="checkbox"/>	Access	2/12/13	patrollost in PM45	N	LHC Logbook	1039117	{\"fault_description\": {\"gro...	Hardware	11/18/14	1	N	OP_ENDED	
<input type="checkbox"/>	SPS	1/28/13		N	LHC Logbook	1038905	{\"fault_description\": {\"gro...	No beam	11/18/14	1	N	OP_ENDED	
<input type="checkbox"/>	PSB	2/2/13		N	LHC Logbook	1038986	{\"fault_description\": {\"gro...	No beam	11/18/14	1	N	OP_ENDED	
<input type="checkbox"/>	RF	7/16/10	Ch 8 on CIB.UA47.R4B1	N	LHC Logbook	1021242	{\"fault_description\": {\"gro...	Hardware	11/18/14	1	N	OP_ENDED	RF interlock on BIS
<input type="checkbox"/>	PSB	7/21/12		N	LHC Logbook	1034469	{\"fault_description\": {\"gro...	No beam	11/18/14	1	N	OP_ENDED	
<input type="checkbox"/>	Injection	2/12/13	tune kicker	N	LHC Logbook	1039116	{\"fault_description\": {\"gro...	Hardware	11/18/14	1	N	OP_ENDED	
<input type="checkbox"/>	PSB	1/28/13	cooling	N	LHC Logbook	1038904	{\"fault_description\": {\"gro...	No beam	11/18/14	1	N	OP_ENDED	
<input type="checkbox"/>	Beam dump	2/2/13	Trigger/Retrigger fault	N	LHC Logbook	1038981	{\"fault_description\": {\"gro...	Controls	11/18/14	1	N	OP_ENDED	
<input type="checkbox"/>	Cryogenics	11/12/12	Cold Compressor S34	N	LHC Logbook	1037943	{\"fault_description\": {\"gro...	Cold compressor	11/18/14	1	N	OP_ENDED	
<input type="checkbox"/>	Access	7/16/10		N	LHC Logbook	1021240	{\"fault_description\": {\"gro...	User	11/18/14	1	N	OP_ENDED	

Explore Multiple Datasets

Search



ORACLE Big Data Discovery

FCC RAMS v1: Faults and Logbook

Explore Transform Discover

AFT_FAULTS OP_LOGBOOK Clear All

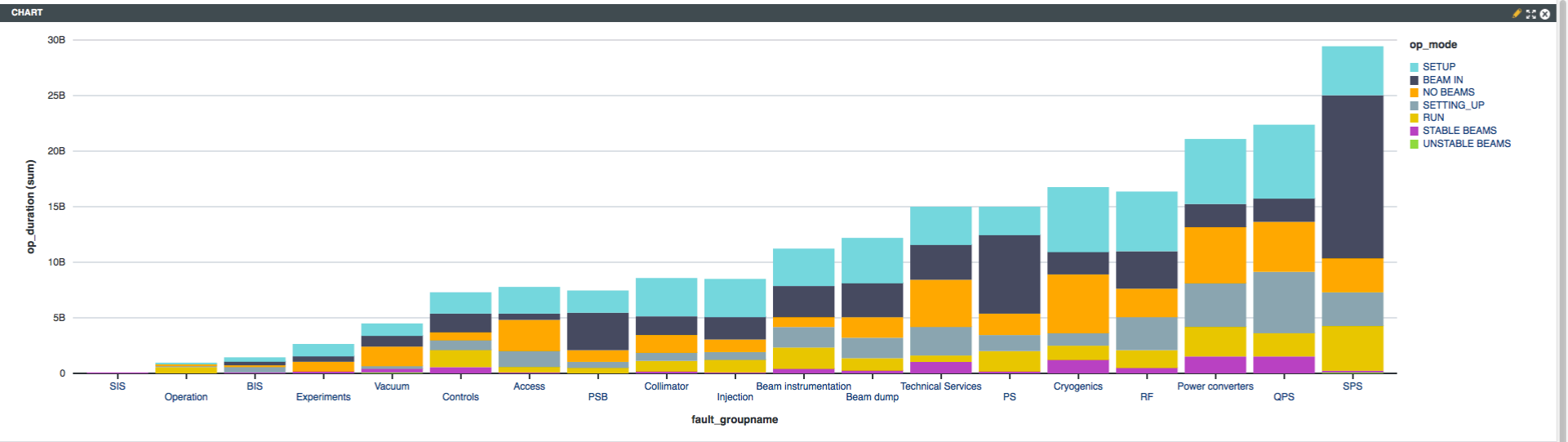


CHART SETTINGS

CHART OPTIONS

CHART TYPE

CHART SUB-TYPE

X-axis

fault_groupna... (by members)

Y-axis

op_duration (sum)

Color

op_mo... (by members)

Trellis

Type or select attribute

DISPLAY OPTIONS

Orientation

Chart height 450 pixels

Chart width Auto 500 pixels

Show chart title

DATA VIEW

Data View aft_faults - linked

Save as default state

RESULTS TABLE

General 0 RECORDS SELECTED

VIEW OPTIONS ACTIONS

	fault_description	fault_name	creation_time_utc (Year-Month-...)	fault_classification_id	is_root_cause	fault_state	description	prevents_injection	system_name	fault_classification_name	start_time (Year-Month-...)
<input type="checkbox"/>	("fault_description": ("gro...	Feedback	11/18/14	1	N	OP_ENDED		N	Feedback	External	2/3/13
<input type="checkbox"/>	("fault_description": ("gro...	No beam	11/18/14	1	N	OP_ENDED		N	No beam	External	2/2/13
<input type="checkbox"/>	("fault_description": ("gro...	No beam	11/18/14	1	N	OP_ENDED		N	No beam	External	2/13/13
<input type="checkbox"/>	("fault_description": ("gro...	Infrastructure	11/18/14	1	N	OP_ENDED	Connection to LHC-OP-L...	N	Infrastructure	External	2/2/13
<input type="checkbox"/>	("fault_description": ("gro...	Hardware	11/18/14	1	N	OP_ENDED		N	Hardware	External	2/12/13
<input type="checkbox"/>	("fault_description": ("gro...	No beam	11/18/14	1	N	OP_ENDED		N	No beam	External	1/28/13
<input type="checkbox"/>	("fault_description": ("gro...	No beam	11/18/14	1	N	OP_ENDED		N	No beam	External	2/2/13
<input type="checkbox"/>	("fault_description": ("gro...	Hardware	11/18/14	1	N	OP_ENDED	RF interlock on BIS	N	Hardware	External	7/16/10
<input type="checkbox"/>	("fault_description": ("gro...	No beam	11/18/14	1	N	OP_ENDED		N	No beam	External	7/21/12
<input type="checkbox"/>	("fault_description": ("gro...	Hardware	11/18/14	1	N	OP_ENDED		N	Hardware	External	2/12/13
<input type="checkbox"/>	("fault_description": ("gro...	No beam	11/18/14	1	N	OP_ENDED		N	No beam	External	1/28/13
<input type="checkbox"/>	("fault_description": ("gro...	Controls	11/18/14	1	N	OP_ENDED		N	Controls	External	2/1/13
<input type="checkbox"/>	("fault_description": ("gro...	Cold compressor	11/18/14	1	N	OP_ENDED		N	Cold compressor	External	11/12/12
<input type="checkbox"/>	("fault_description": ("gro...	User	11/18/14	1	N	OP_ENDED		N	Patrol lost	External	7/16/10
<input type="checkbox"/>	("fault_description": ("gro...	Controls	11/18/14	1	N	OP_ENDED		N	Controls	External	7/21/12
<input type="checkbox"/>	("fault_description": ("gro...	Hardware	11/18/14	1	N	OP_ENDED		N	Hardware	External	2/12/13
<input type="checkbox"/>	("fault_description": ("gro...	No beam	11/18/14	1	N	OP_ENDED		N	No beam	External	1/26/13
<input type="checkbox"/>	("fault_description": ("gro...	Hardware	11/18/14	1	N	OP_ENDED	M1B1 power converter n...	N	Hardware	External	9/26/12

REFINE BY

- AFT_CARDIOGRAM
- AFT_FAULTS
 - creation_time_utc
 - description
 - element
 - eologbook_fault_id
 - fault_classification_id
 - fault_classification_name
 - fault_creation_source_name
 - fault_description
 - fault_id
 - fault_name
 - Filter...
 - No beam
 - Hardware
 - Controls
 - Voltage source
 - Controller
 - Electrical Services
 - BLM - hardware
 - Software
 - Cold compressor
 - BPM - hardware
 - Interlock
 - Cooling
 - Feedback
 - Infrastructure
 - Timing
 - ALICE
 - User
 - ATLAS
 - Miscellaneous
 - Middleware
 - More
 - Select All
 - fault_state
 - group_name
 - is_root_cause
 - op_duration
 - op_end_time
 - prevents_beam_op
 - prevents_injection
 - HIGH_LEVEL_SUMMARY_SCHEDULE
 - LHCLOG_CRYO
 - OP_LOGBOOK

Data Visualization

FCC RAMS v1: Extended Charts

Explore Transform Discover

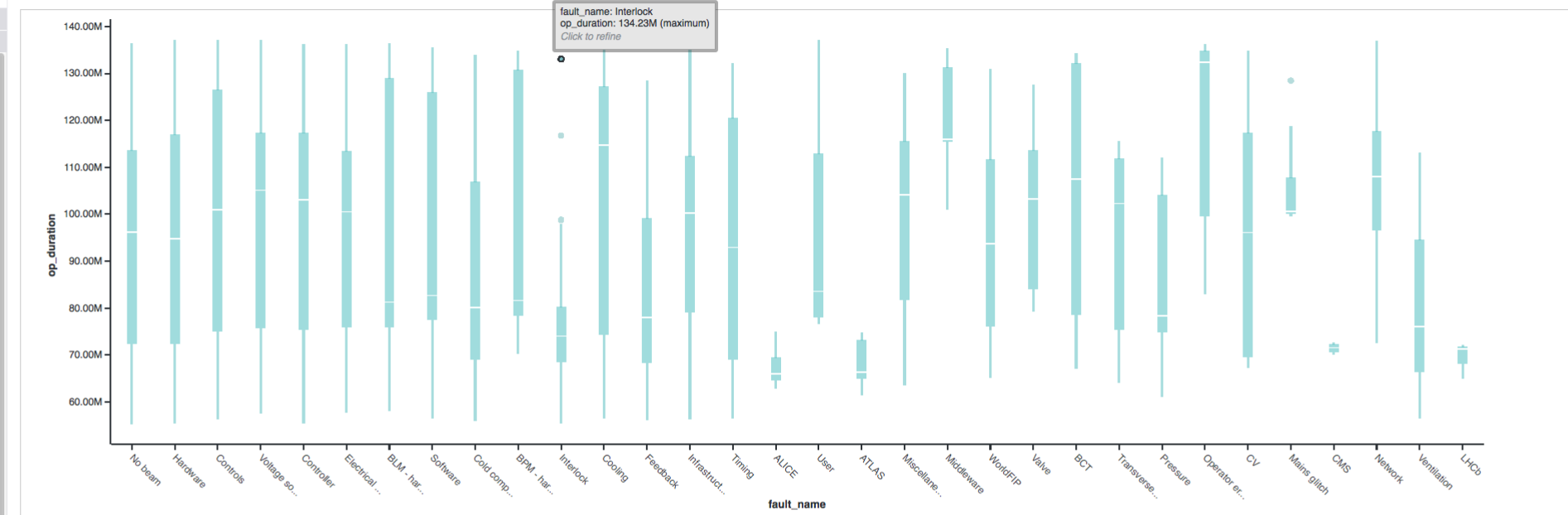
AFT_FAULTS
OP_LOGBOOK
[Clear All](#)

faultname
VOID

REFINE BY

- AFT_CARDIOGRAM
- AFT_FAULTS
 - creation_time_utc
 - description
 - element
 - eologbook_fault_id
 - fault_classification_id
 - fault_classification_name
 - fault_creation_source_name
 - fault_description
 - fault_id
 - fault_name
 -
 - No beam
 - Hardware
 - Controls
 - Voltage source
 - Controller
 - Electrical Services
 - BLM - hardware
 - Software
 - Cold compressor
 - BPM - hardware
 - Interlock
 - Cooling
 - Feedback
 - Infrastructure
 - Timing
 - ALICE
 - User
 - ATLAS
 - Miscellaneous
 - Middleware
 - [More](#)
 - [Select All](#)
 - fault_state
 - group_name
 - is_root_cause
 - op_duration
 - op_end_time
 - prevents_beam_op
 - prevents_injection
 - HIGH_LEVEL_SUMMARY_SCHEDULE
 - LHCLOG_CRYO
 - OP_LOGBOOK

BOX PLOT

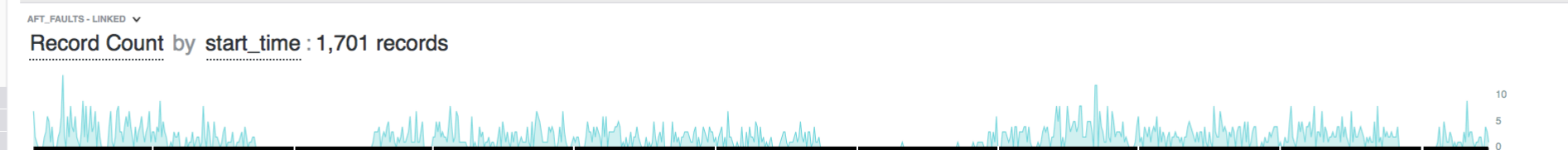
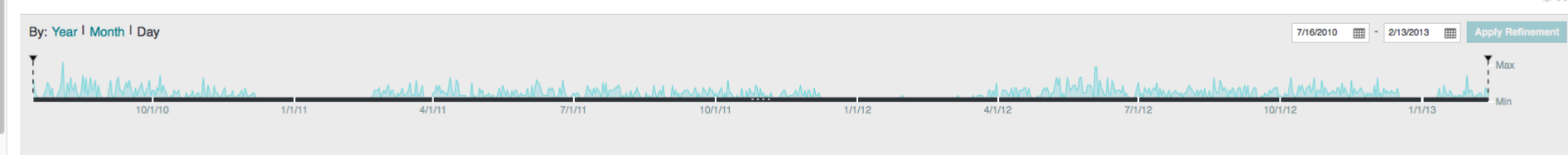


VIEW: aft_faults | METRIC: op_duration | DIMENSION: fault_name

ADD COMPONENT

-
- Box Plot
- Chart
- Column Container
- Histogram Plot
- IFrame
- Map
- Parallel Coordinates Plot
- Pivot Table
- Results List
- Results Table
- Summarization Bar
- Tabbed Container
- Tag Cloud
- Thematic Map
- Timeline
- Web Content Display
- Donut Pie (Unpublished)

TIMELINE



Data Visualization

CMS Fills: tmb_rates

Explore Transform Discover

CMS_RUNS_FOR...
 CMS_TMB_RATES Clear All

runnumber
 199608

REFINE BY

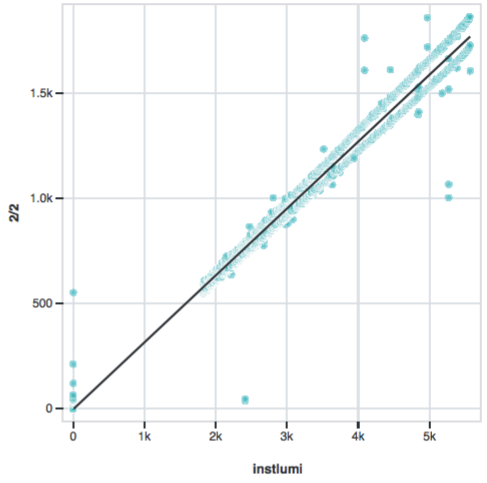
- ▶ CMS_LHCFILLS
- ▶ CMS_RUNS_FOR_FILL
- ▼ CMS_TMB_RATES
 - ▶ 1/1
 - ▶ 1/2
 - ▶ 1/3
 - ▶ 2/1
 - ▶ 2/2
 - ▶ 3/1
 - ▶ 3/2
 - ▶ 4/1
 - ▶ 4/2
 - ▶ fct_time
 - ▶ instlumi
 - ▶ instlumierror
 - ▶ me
 - ▶ pileup
 - ▶ pileuprms
 - ▼ runnumber

Filter...

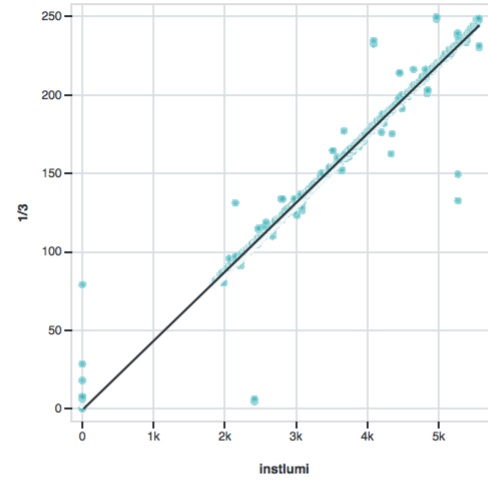
 - 252504
 - 255038
 - 173692
 - 177730
 - 258177
 - 207454
 - 200091
 - 206745
 - 166512
 - 194050
 - 195552
 - 166380
 - 195398
 - 172822
 - 258158
 - 165993
 - 193621
 - 210658
 - 202504
 - 176309
- ▶ starttime
- ▶ stoptime

COLUMN CONTAINER

CHART



CHART



RESULTS TABLE

0 RECORDS SELECTED

	runnumber	1/1	1/2	1/3	2/1	2/2	3/1	3/2	4/1	4/2	fct_time (Year-Month-D...	instlumi	instlumierror
<input type="checkbox"/>	199608	67,350.85	2,642.26	206.71	31,179.86	1,442.91	15,986.39	1,996.61	14,363.54	27,744.87	7/24/12	4,712.53	0.
<input type="checkbox"/>	199608	67,365.72	2,642.37	207.29	31,161.58	1,443.19	15,975.94	1,994.15	14,358.52	27,796.3	7/24/12	4,715.04	0.
<input type="checkbox"/>	199608	67,454.42	2,643.77	207.7	31,204.16	1,442.88	16,002.51	1,997.64	14,372.98	27,779.07	7/24/12	4,721.7	0.
<input type="checkbox"/>	199608	67,509.79	2,646.2	207.54	31,248.23	1,443.6	16,016.13	1,999.95	14,384.57	27,841.57	7/24/12	4,722.9	0.
<input type="checkbox"/>	199608	68,388.87	2,682.23	210.2	31,660.51	1,463.29	16,227.77	2,025.69	14,576.27	28,205.57	7/24/12	4,785.57	0.
<input type="checkbox"/>	199608	67,534.9	2,650.59	208.6	31,258.85	1,447.08	16,030.38	2,001.05	14,396.56	27,861.61	7/24/12	4,726.93	0.
<input type="checkbox"/>	199608	68,442.37	2,685.5	210.66	31,672.86	1,466.04	16,229.6	2,028.39	14,572.22	28,179.21	7/24/12	4,787.21	0.
<input type="checkbox"/>	199608	57,447.28	2,242.93	175.22	26,464.05	1,223.52	13,601.92	1,693.48	12,218.61	23,482.16	7/24/12	4,000.56	0.
<input type="checkbox"/>	199608	67,631.33	2,649.78	208.45	31,308.91	1,446.96	16,056.75	2,002.4	14,426.47	27,878.9	7/24/12	4,729.61	0.
<input type="checkbox"/>	199608	68,650.39	2,690.54	210.96	31,721.55	1,467.61	16,260.08	2,031.54	14,605.17	28,235.37	7/24/12	4,790.64	0.
<input type="checkbox"/>	199608	41,589.1	1,623.45	126.95	19,276.53	881.4	9,906.8	1,229.46	8,899.11	16,878.53	7/25/12	2,912.43	0.
<input type="checkbox"/>	199608	57,490.17	2,240.56	175.86	26,474.51	1,225.51	13,605.17	1,693.86	12,224.77	23,468.41	7/24/12	4,004.74	0.
<input type="checkbox"/>	199608	67,674.77	2,652.31	208.49	31,322.94	1,449.79	16,055.61	2,006.47	14,424.57	27,900.23	7/24/12	4,734.54	0.
<input type="checkbox"/>	199608	70,365.46	2,743.45	215.43	32,362.59	1,497.95	16,585.46	2,070.38	14,905.43	28,829.61	7/24/12	4,894.94	0.
<input type="checkbox"/>	199608	69,031.09	2,686.93	211.71	31,739.6	1,466.16	16,273.74	2,029.41	14,625.27	28,267.1	7/24/12	4,798.35	0.
<input type="checkbox"/>	199608	150.47	36.83	3.62	314.42	29.67	289.97	42.38	261.89	58.85	7/24/12	0.32	0.
<input type="checkbox"/>	199608	28,283	1,166.89	91.61	14,532.89	687.88	7,316.02	885.32	6,180.48		7/25/12	2,088.17	0.

Next steps

- Check new upcoming features presented during OOW 15
 - Scalability improvements
 - Streaming integration (Kafka)
 - Add more advanced charts and visualizations
 - Machine Learning integration
- Evaluate Big Data Discovery cloud
- Extend to more CERN use cases
 - Controls and Operations
 - Accelerator Fault Tracking
 - Diagnostics and Monitoring
 - IT Infrastructure Monitoring
 - Server logs analysis
 - Database latency
 - Human Resources



Conclusions

- Data visualization and discovery is an important area in data analytics
 - Facilitates users to visualize and explore their data
 - Find correlations, extract insight and useful information
- Important points
 - Flexible and user-friendly platform
 - Advanced data visualization and exploration
 - Collaborative
- Application to different domains
 - Controls and Operations
 - IT Infrastructure Monitoring
 - Human Resources



www.cern.ch