Big Data Analytics for the Exploitation of the CERN Accelerator Complex

Antonio Romero Marín





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"





India	220	
Japan	244	
Russia	982	
Turkey	146	OF
USA	979	25

Afghanistan	1	El Salvador	1	Pakistan	
Albania	2	Estonia	16	Palestine (O.T.).	
Algeria	8	Georgia	36	Peru	
Argentina	11	Gibraltar	1	Philippines	
Armenia	25	Hong Kong	1	Saudi Arabia	
Australia	25	Iceland	4	Senegal	
Azerbaijan	8	Indonesia	1	Singapore	
Bangladesh	4	Iran	28	Sint Maarten	
Belarus	47	Ireland	22	Slovenia	
Bolivia	З	Jordan	2	South Africa	
Bosnia &		Kenya	1	Sri Lanka	
Herzegovina	1	Korea, D.P.R.	1	Syria	
Brazil	108	Korea Rep.	117	Thailand	
Cameroon	1	Kuwait	1	T.F.Y.R.O.M.	
Canada	134	Lebanon	12	Tunisia	
Cape Verde	1	Lithuania	19	Ukraine	
Chile	12	Luxembourg	4	Uzbekistan	
China	280	Madagascar	4	Venezuela	
China (Tapei)	45	Malaysia	15	Viet Nam	
Colombia	30	Mauritius	1	Zimbabwe	
Croatia	35	Mexico	64		
Cuba	7	Montenegro	З		
Cyprus	16	Morocco	12		
Ecuador	З	Nepal	5		
Egypt	19	New Zealand	7		14

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	E	5352

Candidate for Accession

Romania

Associate Members in the Pre-stage to Membership

Serbia

Distribution of All CERN Users by Nationality on 14 January 2014

CERN Mission





Science & Education

Training tomorrow's scientists and engineers

Science & Technology

Advancing the frontiers of technology

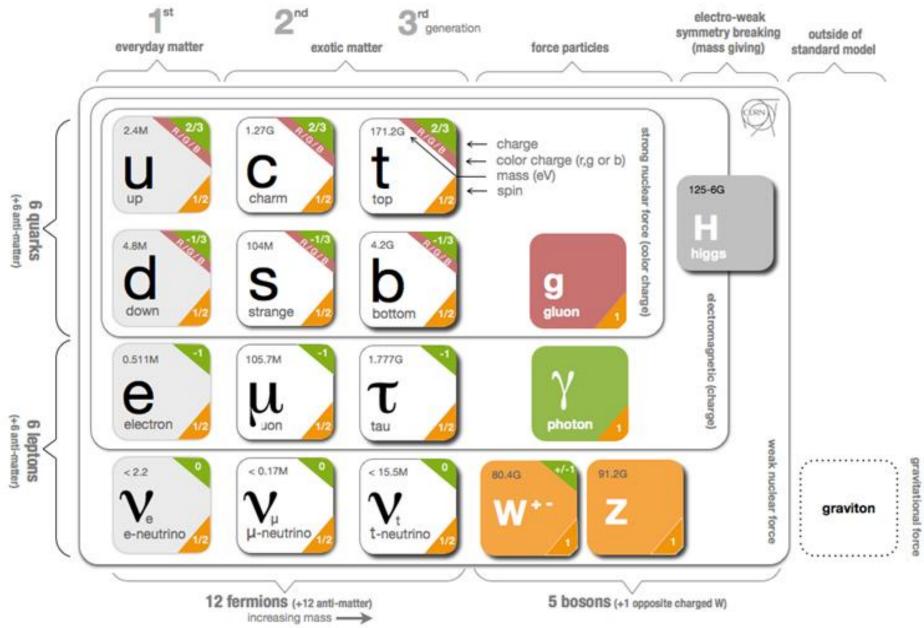


Milan – 11/03/2015

What is the Universe made of?

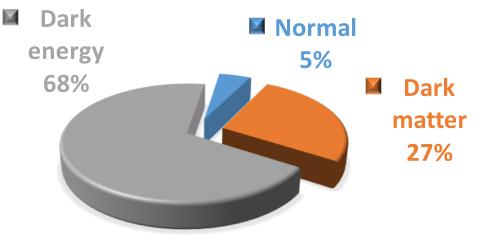
How does it work?

The Standard Model



Fundamental Research

• What is dark matter and dark energy?



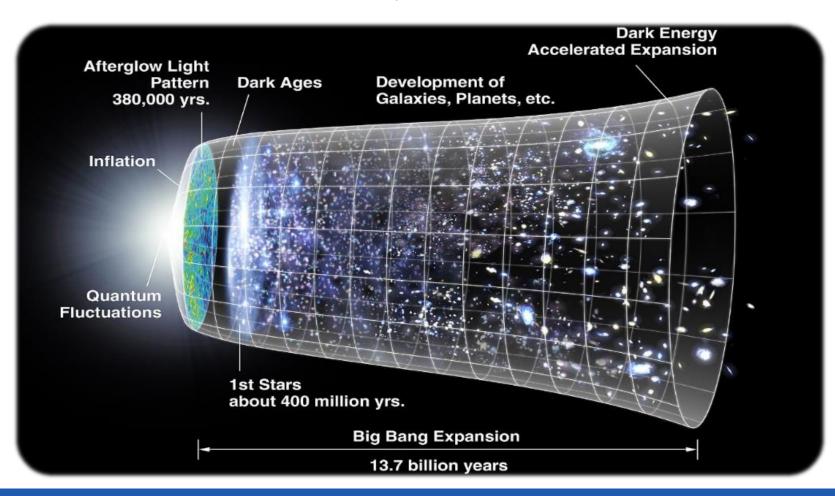
Composition of the Universe

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



Fundamental Research

• What was the state of the matter in the very first moments of the Universe?





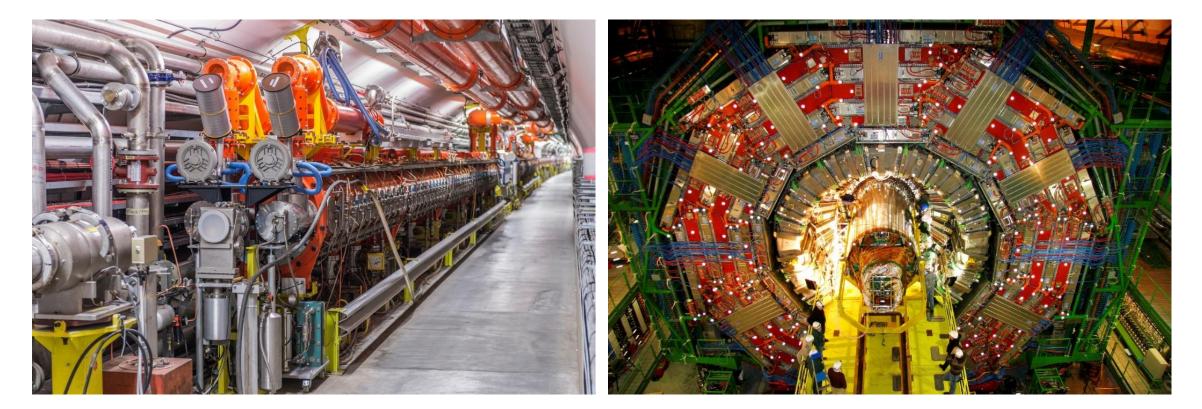
CERN Instruments

Accelerators

Detectors

Boost particles to high energies and speed to collide

Observe and record the results of these collisions





The Large Hadron Collider (LHC)

• Largest machine in the world

CMS

SUISSE

FRANCI

- 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

HCh

LING 4/ NIT

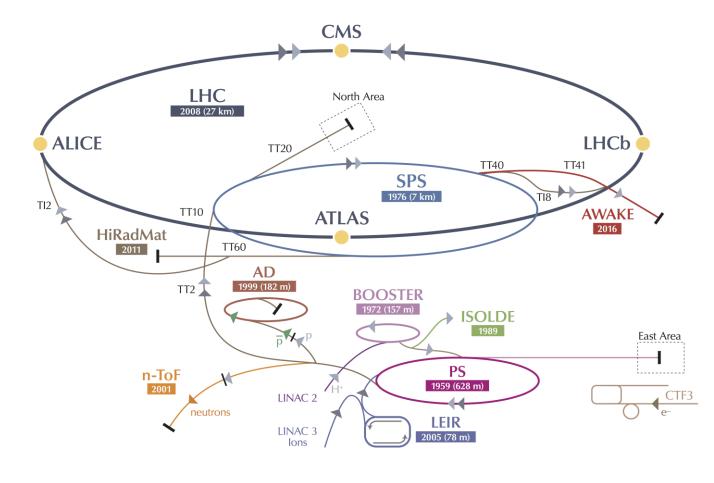
CERN Prévessin

ATLAS

CERN Meyrin

ALICE

The CERN Accelerator Complex



▶ p (proton) \blacktriangleright ion \blacktriangleright neutrons \blacktriangleright p (antiproton) \blacktriangleright electron \rightarrow + \vdash proton/antiproton conversion

LHC Large Hadron Collider SPS Super Proton Synchrotron PS Proton Synchrotron

AD Antiproton Decelerator CTF3 Clic Test Facility AWAKE Advanced WAKefield Experiment ISOLDE Isotope Separator OnLine DEvice LEIR Low Energy Ion Ring LINAC LINear ACcelerator n-ToF Neutrons Time Of Flight HiRadMat High-Radiation to Materials

CERN is an extreme data environment

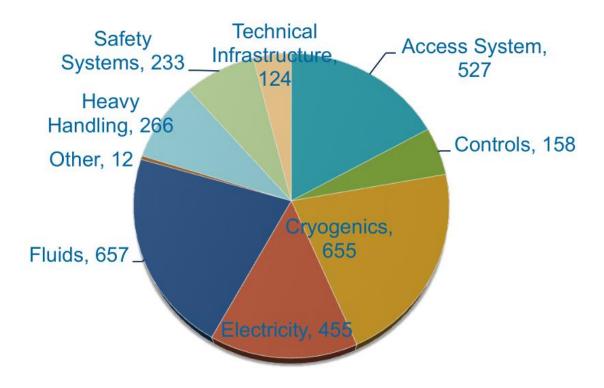
- Control and operations
 - Millions of sensors, signals
 - Large number of control devices
 - Equipment
- Monitoring and logging
- Supporting IT infrastructure
 - Databases
 - Network
 - Services
- CERN has great monitoring and logging systems
 - Large amount of data has been stored over years





Data Analytics Challenges

- Some faults cannot be avoid
- Decrease the availability for running physics
- Corrective interventions needed





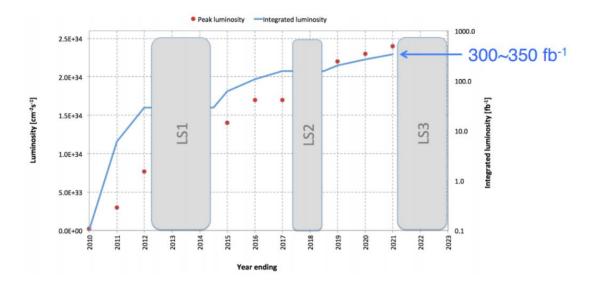
A look into the Future

• LHC upgrades will further increase luminosity

- Computing resources needs will be higher
- Data generated will increase drastically

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

Table 17: HLT Rate Evolution

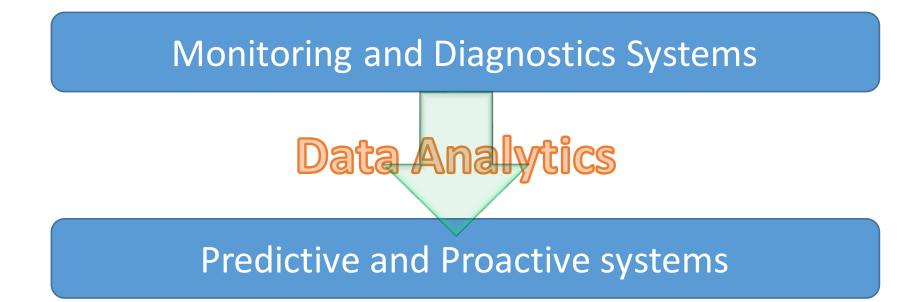


- Next accelerators
 - Future Circular Collider (80-100 km)





The objective – Improve our systems





Milan – 11/03/2015

openlab Data Analytics Project

• Optimize our systems

- Reducing and predicting faults and corrective interventions
- Increase the availability and operations efficiency

• Profit from CERN data investment by using data analytics

- Extract knowledge
- Discover useful information
- Suggest conclusions
- Support decision making

Control and Monitoring Systems

- Proactive
- Predictive
- Intelligent





CERN openlab



SIEMENS

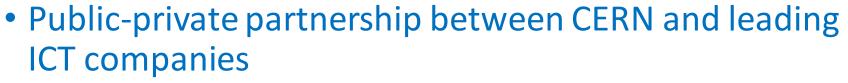
Contributors

C rackspace.

Associates

Yandex





- Accelerate cutting-edge solutions to be used by the worldwide LHC community
- Designed to create and disseminate knowledge
 - Publication of reports and articles
 - Workshops or seminars
 - CERN openlab Student Programme

Data Analytics Use Cases



Use Case: CERN Advance Storage Manager (CASTOR)

- Mass Storage Solution for managing physics data files
 - 12k disks, 30k tapes
 - 100 PB on tape, 50 PB on disk
 - +300 M files



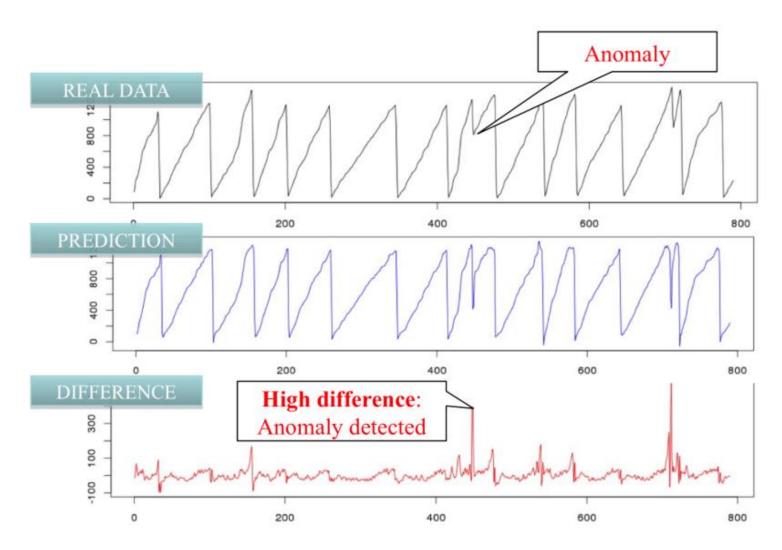




Use Case: CERN Advance Storage Manager (CASTOR)

- Optimization
 - Performance
 - Cause of errors

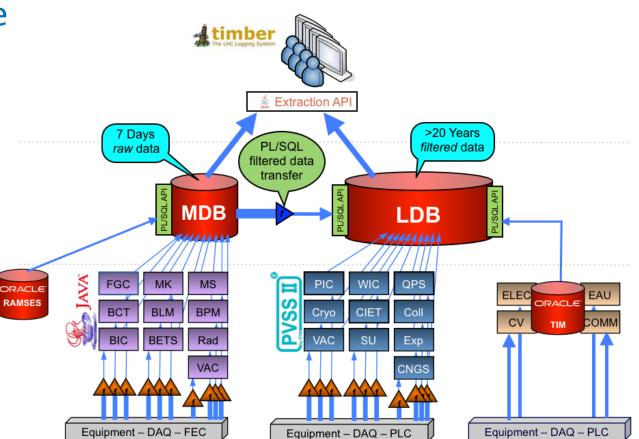
- Predictive analytics
 - Anomaly detection
 - Early warning systems





Use Case: Operation and Control Systems

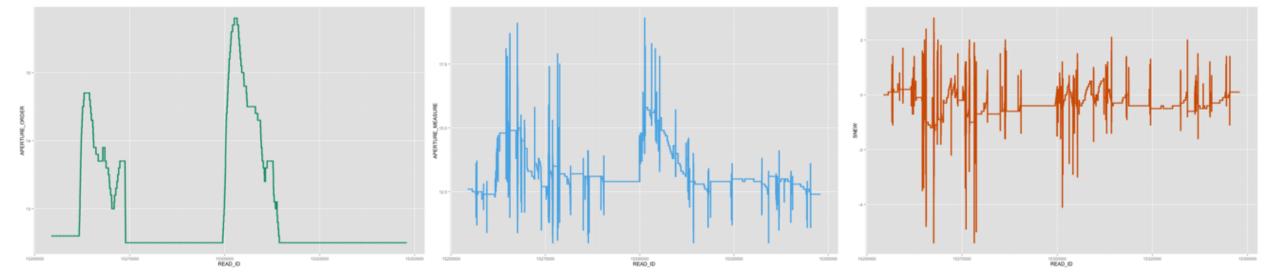
- LHC Accelerator Logging Service
 - Around 1 million signals
 - Temperatures, electrical currents
 - Magnetic field strengths
 - Vacuum pressures, etc.
- Control system Health
 - Gas Breakdown
 - Vacuum
 - Machine Protection
- Predictive maintenance





Use Case: Cryogenics Faulty Valves Detection

- What is the objective?
 - Predict faulty valves before they actually fail
- How?
 - Valve receive an aperture order value (aperture order)
 - Effective aperture realized by the valve (aperture measured)
 - Analyzing the difference between both (**S** = aperture order aperture measured)





Use Case: Cryogenics Faulty Valves Detection

- Signals used
 - S = aperture order aperture measured
- Features extractions based on S
 - Variance
 - Percentile 99.9
 - Rope distance
 - Noise Band
- Three different status
 - Faulty
 - Not faulty
 - Unknown
- Predictive model
 - SVM Support Vector Machine



- Interactive and visual analytics by end users
 - Flexible and easy to use but powerful
 - Analyze information of any type and any source
 - Get new insights from data
- Electronic Logbook
 - Log of events in the accelerator complex

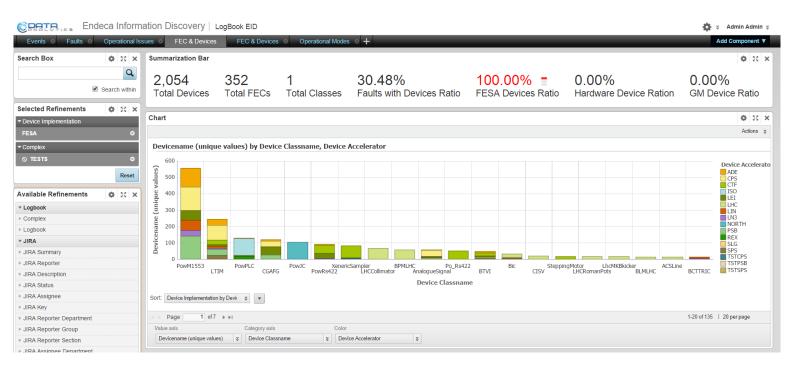
Stati	stics for the PS V From: 2012 eLogbook:	20815 Period: Morning To: 20	120817 🕮 Period: Morning 🔻
		Availabilities	
Lines	In Super Cycle	In Fault	Availabilities (%)
AD	55 [h]	7 [h] 57 [min] 37[s]	85%
EASTA	55 [h]	7 [h] 57 [min] 37[s]	85%
EASTB	55 [h]	7 [h] 57 [min] 37[s]	85%
SFTPRO	6 [h] 54 [min] 29[s]	5 [h] 45 [min] 18[s]	16%
EASTC	55 [h]	7 [h] 57 [min] 37[s]	85%
CNGS	55 [h]	10 [h] 22 [min] 58[s]	81%
LHCPROBE	55 [h]	7 [h] 57 [min] 37[s]	85%
I_LHC	31 [h] 31 [min] 33[s]	4 [h] 37 [min] 40[s]	85%
LHC	55 [h]	7 [h] 57 [min] 37[s]	85%
TOTAL	423 [h] 26 [min] 02[s]	68 [h] 31 [min] 38[s]	83%

Systems for AD							
GROUP NAME	DURATION						
PS	Power supply	3 [h] 47 [min] 34[s]					
PS	RF	3 [h] 13 [min] 33[s]					

	1		Global Post Mortem Event
4	23:42	SUP	Event Timestamp: 10/06/12 23:42:39.163 Fill Number: 2718 Accelerator / beam mode: PROTON PHYSICS / STABLE BEAMS Energy: 4000080 [MeV] Intensity Bl/B2: 15509 / 14217 [e^10 charges] Event Category / Classification: PROGRAMMED_DUMP / MULTIPLE_SYSTEM_DUMP First BIC input Triggered: First USR_PERMIT change: Ch 1-Programable Dump b1: A T -> F on CIB.CCR.LHC.B1
5	23:42	SUP	Global Post Mortem Event Confirmation Dump Classification: Programmed Dump Operator / Comment: papotti / End of physics fill, clean dump.
			BEAM MODE > BEAM DUMP
6	23:42	SUP	LHC RUN CTRL: BEAM MODE changed to BEAM DUMP
			BEAM MODE > BEAM DUMP
7	23:42	SUP	LHC RUN CTRL: BEAM MODE changed to BEAM DUMP
8	23:42	SUP	ELOGBOOK: STARTING B1 MKISS
9	23:43	SUP	ELOGBOOK: STARTING B2 MKISS
10	23:44	SUP	LHC SEQ: beam dump handshake closed; LHC=STANBY, EXP=VETO
11	23:44	SUP	LHC SEQ: MCS checks finished
12	23:45	SUP	LHC SEQ: SMP pre-operational checks finished
13	23:45	SUP	LHC SEQ: BIS pre-operational checks finished
			BEAM MODE > RAMP DOWN
14	23:48	SUP	LHC RUN CTRL: BEAM MODE changed to RAMP DOWN
15	23:48		LHC SEQ: BPMLHC calibration finished. Overall result: SUCCESS Chosen bunch spacing: (B1 & B2) BUNCH_50NSEC (manually chosen)
			(For more details see BI-LHC ELogBook)



- Integrating Data Sources
 - Electronic Logbook
 - Controls Configuration DB
 - JIRA (JSON)
- Text analysis
 - Extract most relevant terms
 - Entity extraction



Tag Cloud

4 20 ×

active filter beam dump BSRT beam cable shorted causing orbit fluctations circuit breaker cold box COLD COMPRESSOR COMMUNICATION problem communication stop compressor trip cryogenics problem current lead DFB valve control loop electrical glitch electrical interruption electrical problem Faulty temperature gauge headtall instability heater discharging injection septa interlocked positions. Left jaw Lost communication Lost S34 motor position oil pump Optical link power converter Power converter power cut power cycle power glitch power module power problem power supply power trigger power supply pressure interlocks provoking loss required access right jaw Settings mismatch stable filter stopped water temperature sensor vacuum problem Vacuum problem vacuum valves Warm Compressor water fault

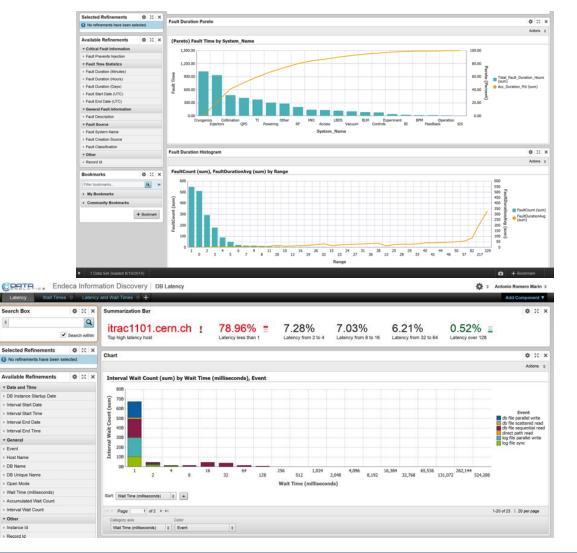


- Better LHC Operations
 - Events Analysis
 - Correlate Information
 - Fault Tracking System
 - Operational Modes
 - Operational Issues
 - Control Equipment

CDATA End	deca Inform	ation Discovery	eLogBook	EID						🗱 🗧 Admin Admin 🌣
Events Faults	Operational Is	ssues Availability	Logbooi	k ○ +						Add Component V
Search Box	• :: ×	Summarization Bar								0 11 ×
	Q	130,396	2,70	6 2.08%	0.02	69.99% =	0.00%	30.01% ≡		
	Search within	Total Events	Total Faul		TI Major Events	PS Complex		нс		
Selected Refinements	• x ×	Tag Cloud	_							ø x x
 Complex (2 Values) 	٥									1.4.1.4.1.4
PS Complex	•	BCT BLM - hardware	BPM - hardw	are Cold compressor	Control Controller Con	trois cooling cv Electri	C Electrical Services	External Feedback Hardy	vare in displacement	Infrastructure Interlock Kicker Magnet Main
цнс	۰	Maria		No. No. be	Dour	ar augalu Dr		1.7.250		
	Reset	Middleware MISCO	laneous	MPS Network NO DB	am Operator error FOW	er supply Pressure RF		rity Software Source Timing	g Transverse feedback L	Jser Vacuum Valve Ventilation VOIL
Available Refinements	0 % X					Voltage sour	CB Water WorldFIP			
* LogBook										
* Complex		Explore Fault Name		by Number of records						
Search for specific value	9	Contract Descention		1.4.00000000000000000000000000000000000						
SPS (66269) HISTORY (14800)		Component Contain	er							0 % ×
TESTS (9548)		Chart				O 30 3	Chart			o x ×
ISOLDE (8999)						Actions \$				Actions \$
BE-BI (2458)		Fault Duration (sum), Faul	L_Duration (average)	by Fault Groupname		Fault Id (unique	values) by Operational I	Mode	
TE-ABT (1179)		1.400		()	of the contraction			runce) of operational		
TE (862) UA9 (773)		1,400 1,200 1,000 1,000 1,000 1,000 1,000 1,000 1,000 1,000 1,000 1,000 1,000	TIT			70 Fault				
TE-VSC (641)		0 1,000 E 800								
RF (455)		600 mg	A			- 30 a Fault_Duration				Operational Mode
EN-ICE (123)		400 1 200				20 S (average)				BEAM IN
SHUTDOWN (53)		- 0	IIII as as			0 1				NO BEAMS
AT-MCS (1)	(Territoria)	2 5 8	RF Technic	CONTROL Vecuum External-Electric Beam dump MISCELLANEOUS CPS EXTERNAL	TIMING Beam Inst Isolde ECCOLING ECCOLING TS EIS EIS CTF	Cont				SETTING_UP
	Select All	penio	Nical D	CONTROL Accuum External-Electric Beam dump MISCELLANEOU DYTERNAL EXTERNAL	r suppi	<u>0</u>		A.		STABLE BEAMS
Eugbook			Servi	NEO NEO	3 June				1	UNSTABLE BEAMS
✓ Events			8	S n	tat					
Event Date					9					
Event Comment				Fault_Groupnan	e					
Operational Mode		Sort: Fault_Groupname	by Fault 💲	•			Sort: Operational Mode	• ¥ .		
Line Name		Page 1	of 1 + +:			1-46 of 46 50 per page	Page 1	of 1 + +1		1-8 of 8 Show all
* Faults		Category axis	011			reio or eo 1 ou per pege	Color			redite anow as
Fault Groupname Fault Name		Fault_Groupname	*				Operational Mode	*		
 Fault Description 										
Fault Start Date										
Fault End Date		Results Table	O X ×							
Fault Duration		Faulta	\$ 0 rea	ords selected						View Options \$ Actions
* Shifts		Recor	A DI D	Fault Groupname	Fault Name	Fault Description	Fault Start Date (Years	Fault End Date (Year-M	Fault Duration	
Shift Start Date						t man a start publi		5/27/11		
Shift End Date			6	Isolde	Security		5/27/11	5/27/11		5

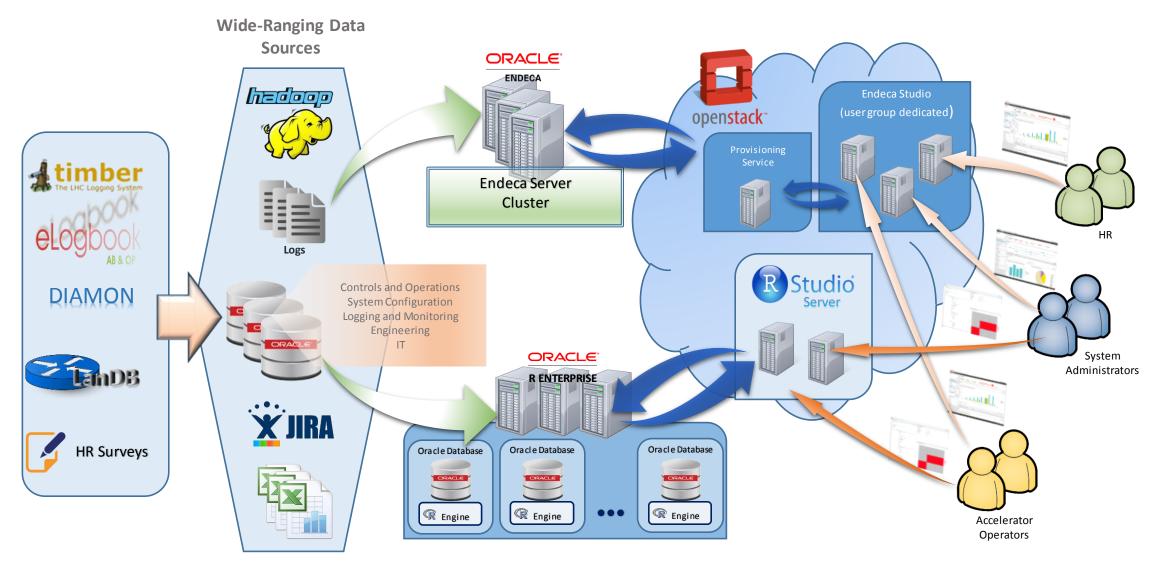


- Many different Applications
 - Controls and Operations
 - Accelerator Fault Tracking
 - Diagnostics and Monitoring
 - IT Infrastructure Monitoring
 - Server logs analysis
 - Database latency
 - Accelerator Logbook
 - Human Resources





Infrastructure







www.cern.ch