Knowledge Transfer at CERN

Nick.Ziogas@cern.ch

Knowledge Transfer Group

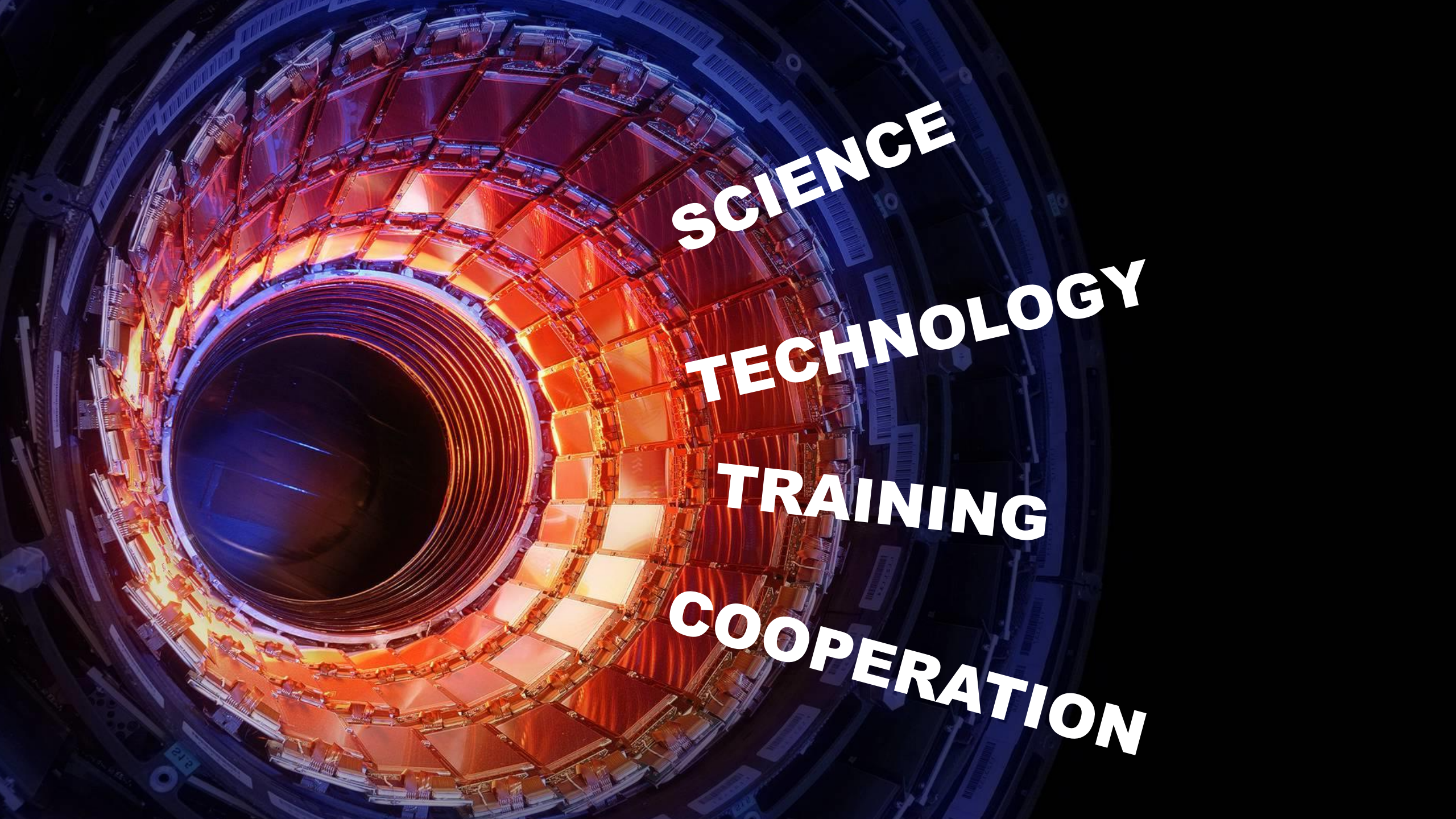
Central Macedonia Companies @ CERN – 10-11 Jan 2019

Why Knowledge Transfer

CERN's next challenge

Sample technologies

The KT Ecosystem


SCIENCE
TECHNOLOGY
TRAINING
COOPERATION

Knowledge Transfer: Mission

Maximize the technological and knowledge return to society, in particular through Member States industry;

Promote CERN's image as a centre of excellence for technology and innovation;

Demonstrate the importance and impact of fundamental research investments;

Key words: ***dissemination*** *and* ***impact***

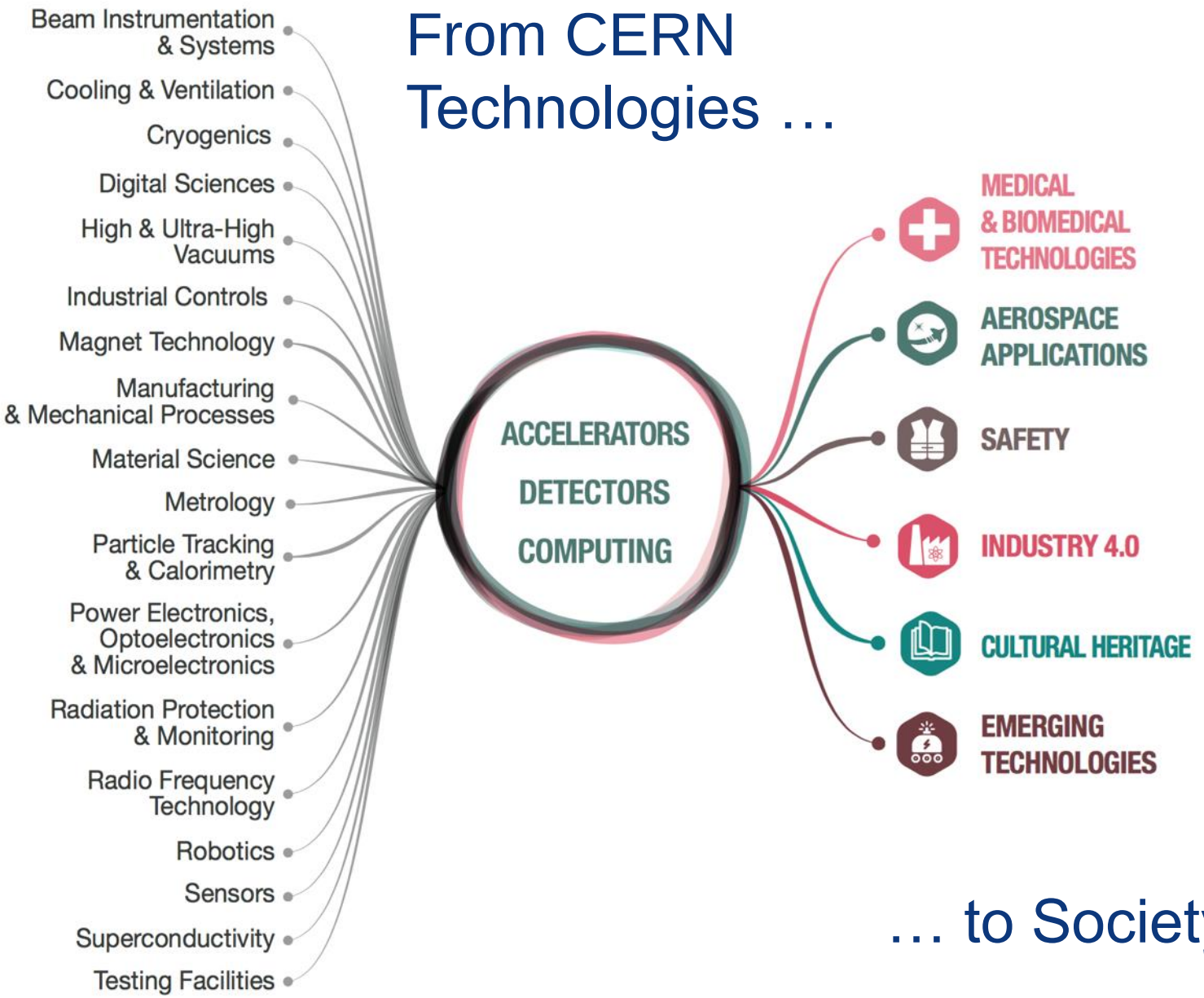
KT Strategy

Disseminate and maximize the impact of CERN's knowledge on society, in particular Member States industry;

Identify new opportunities and encourage CERN's researchers to engage in KT activities while providing a high quality service to them;

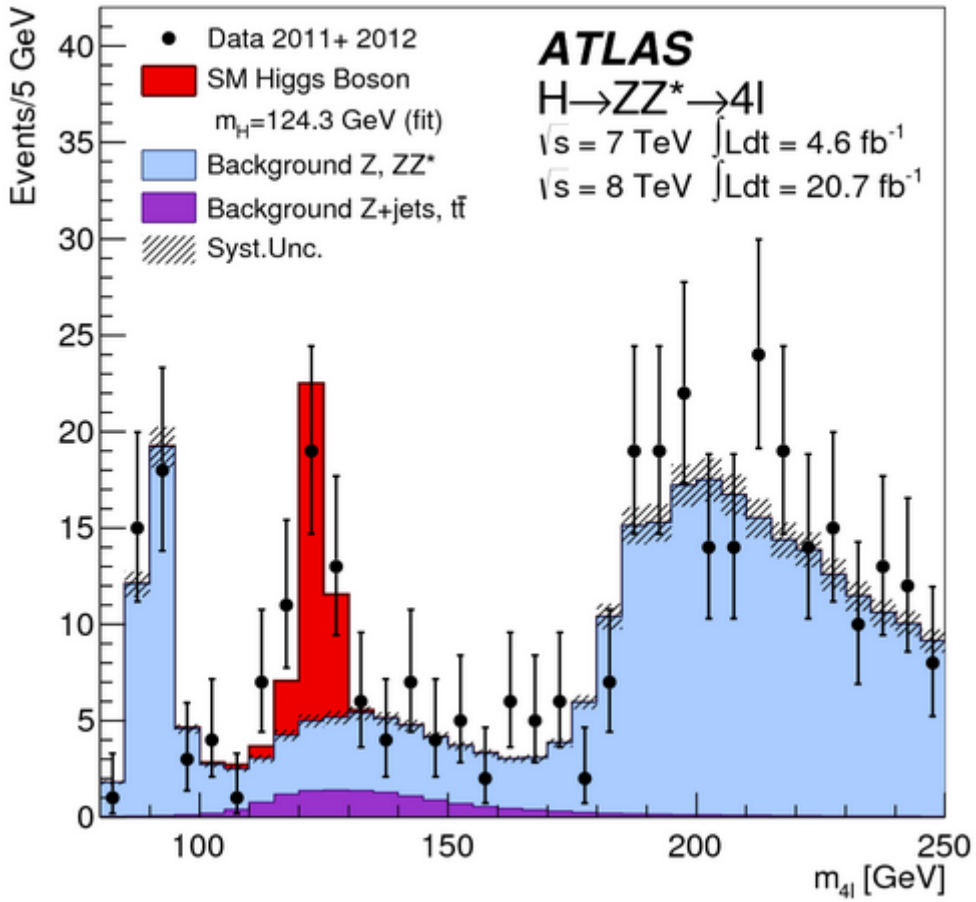
Communicate on CERN's KT activities to key stakeholders (in particular to decision makers in CERN's Member States).

From CERN Technologies …



… to Society

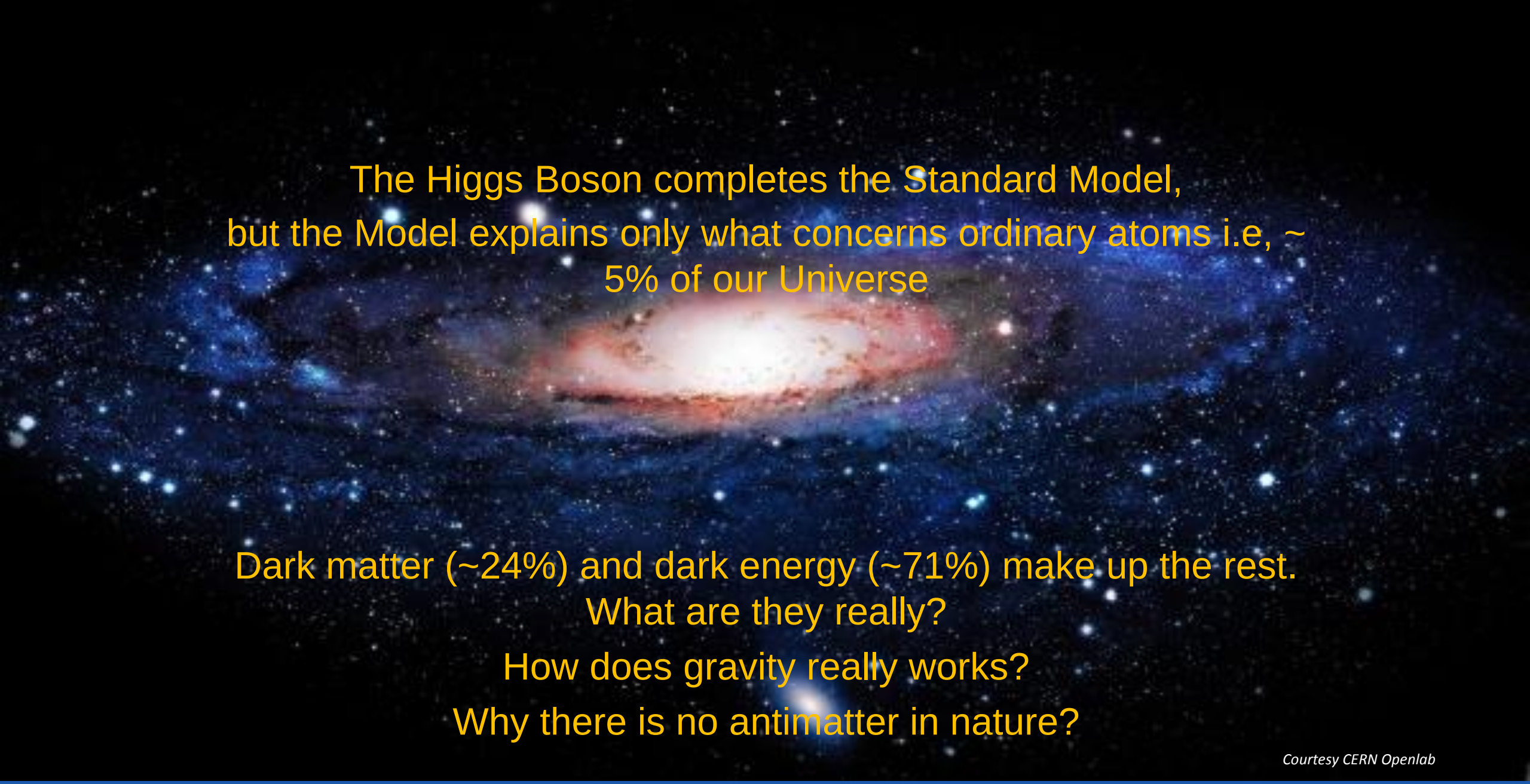
The Higgs Boson



Announced on the
4 July 2012

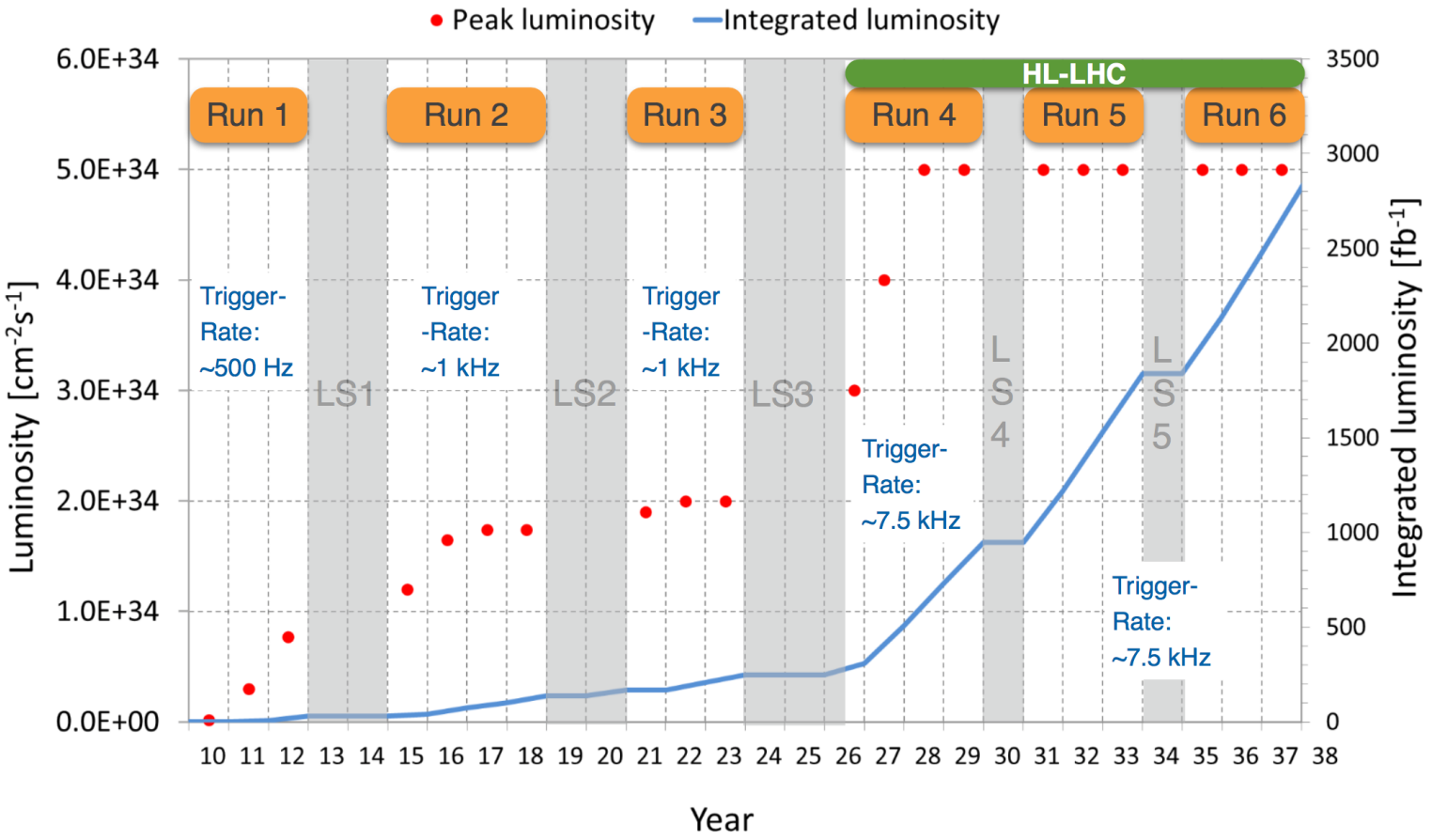
Nobel prize in Physics
8 October 2013

Courtesy CERN Openlab


The Higgs Boson completes the Standard Model,
but the Model explains only what concerns ordinary atoms i.e, ~ 5% of our Universe
Dark matter (~24%) and dark energy (~71%) make up the rest.
What are they really?
How does gravity really works?
Why there is no antimatter in nature?
Courtesy CERN Openlab

The NEXT Challenge

LHC Schedule



Courtesy CERN Openlab

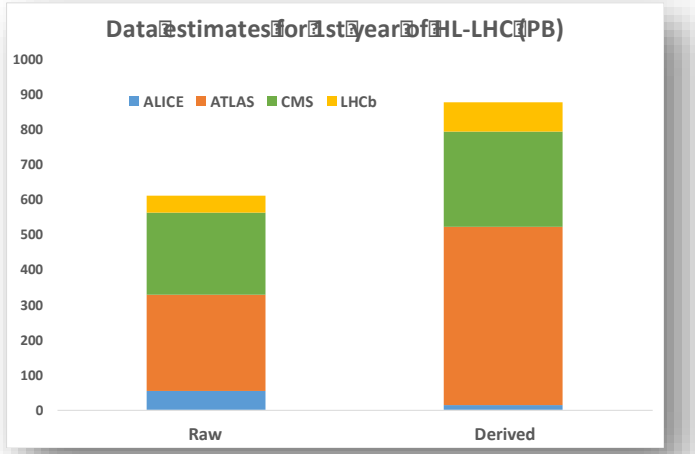
LHC Run3 and Run4

Scale and Challenges

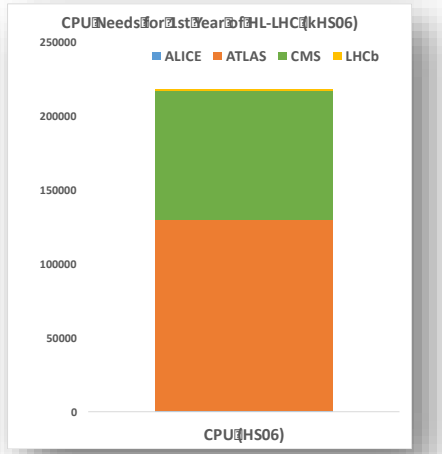
Raw data volume for LHC increases exponentially and with it processing and analysis load

Technology at ~20%/year will bring x6-10 in 10-11 years

Estimates of resource needs at HL-LHC x10 above what is realistic to expect from technology with reasonably constant cost



Data:
- Raw 2016: 50 PB → 2027: 600 PB
- Derived (1 copy): 2016: 80 PB → 2027: 900 PB



CPU:
- x60 from 2016

Technology revolutions are needed

Courtesy CERN Openlab

Deep Learning on FPGAs

High Luminosity LHC: Extreme data rates – 100 TB/sec (CMS)

Challenge: Filter events to reduce data rates to manageable levels → *Triggering*

- *Non filtered events are lost forever*
- Sophisticated techniques are necessary to preserve the physics in a increasingly complex collision environment

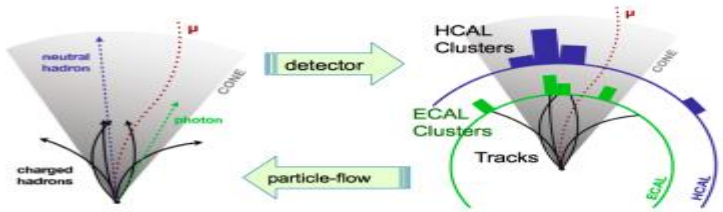
New machine learning trigger algorithms

The detector and its trigger system

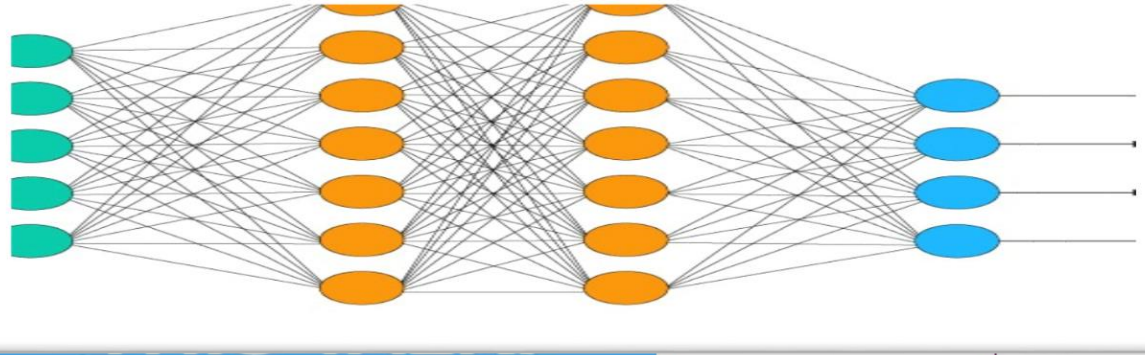
Output: trigger primitives
(calo energy clusters, muons, tracks)

Particle-flow algorithm @ L1

Output:
particle candidates



Machine Learning



Trigger decision

Courtesy Jennifer Ngadiuba

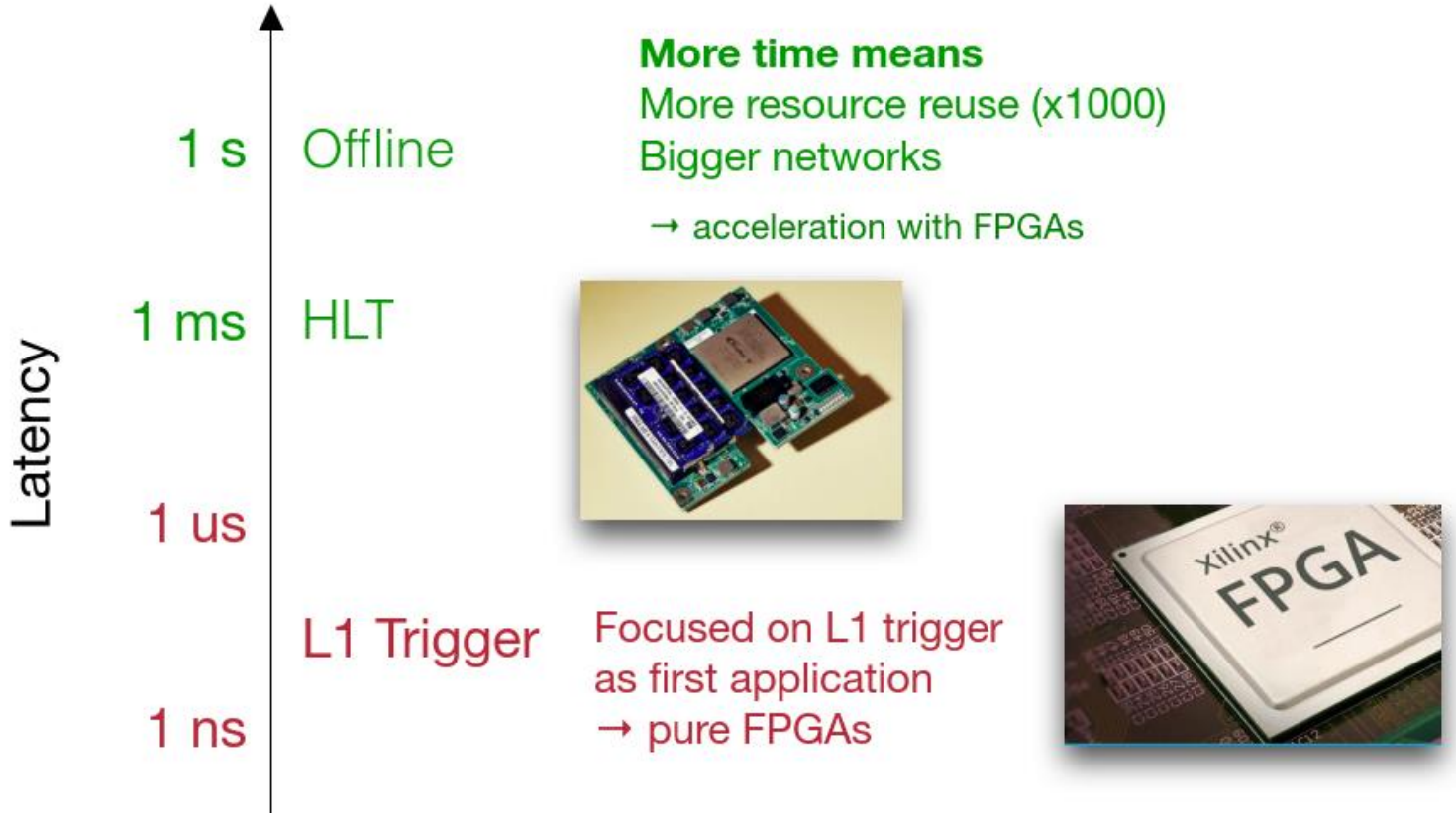
Why FPGAs & accelerator cards?

Offload computational heavy parts to FPGA "accelerator"

10x – 100x increased computational speed
10x reduced system size
10x-100x reduced power consumption

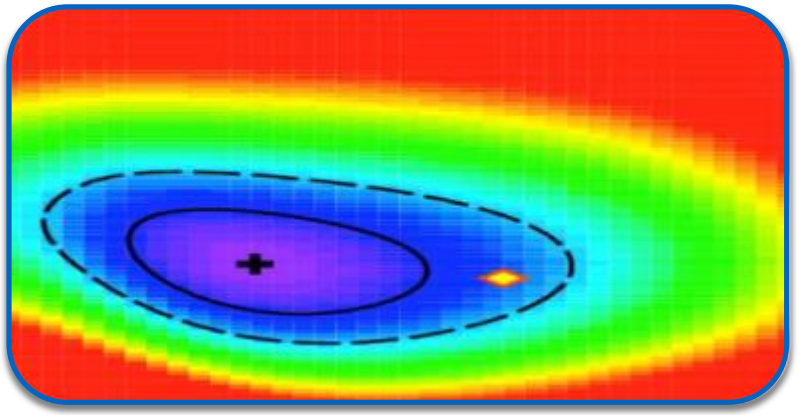
General use cases
Machine learning
Big data analytics
High performance computing (genomics)

@ LHC
Pure FPGAs at the first level trigger
FPGA accelerators for High Level Trigger
FPGA accelerators for offline computing resources



- What can we fit on an FPGA?
- The **hls4ml** package to produce firmware blocks fast! – OS Software

Big Data & Machine Learning Software



WHAT

ROOT / TMVA is a modular big data software framework, providing the functionalities needed to deal with big data processing, statistical analysis, visualisation and storage. It is mainly written in C++ but integrated with other languages such as Python and R. Integrated machine learning environment.

HOW

- Artificial neural networks
- Rectangular cut optimisation
- Projective likelihood estimation
- Multidimensional estimations
- Linear discriminant analysis
- Function discriminant analysis
- Boosted/bagged decision trees
- Predictive learning
- Support Vector Machine

WHY

- **Open source**
- **TMVA Toolkit for multivariate data analysis – machine learning environment integrated in ROOT**
- **Good for analysis of extreme large sets of homogeneous data**
- **Used in physics, biology, finance and insurance fraud analysis**
- **Possible application in genomic data**
-

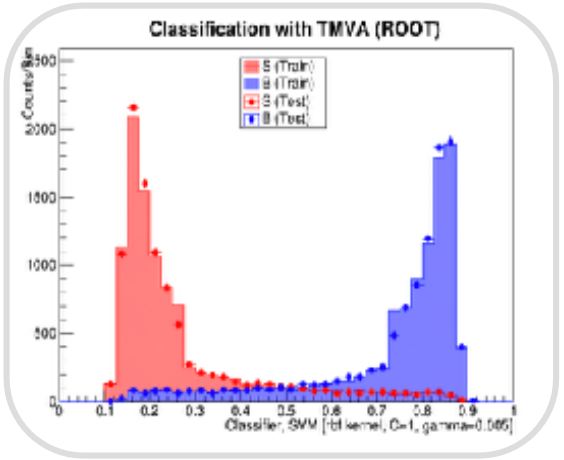
TMVA

- Toolkit for MultiVariate Analysis – ROOT machine learning tools
  - Set of algorithms, used in LHC experiments and in data analysis
  - Easy interface for beginners, powerful for experts
  - TMVA is evolving, several active contributors

- TMVA is also a common interface for classification and regression
- Provides for easy training of different methods on the same data set
- Embedded in ROOT

Big Data & Machine Learning Software

TMVA



USES outside HEP

- **Used in: Finance, Insurance fraud analysis, Auto crash testing data analysis;**
- **Astrophysics, biology, computational neuroscience..**
- **Processing and analysis of large medical datasets, for example**
  - Genomics data
  - EEG/ECG data
  - Biosensor data

WHO

- Over 12’000 Academic and Research users in various fields
- The Sanofi – Pasteur
- Renaissance Technologies financial research
- EU JRC – Institute for reference materials and measurements - IRMM
- Open Source community
- General Aviation startup
- Medical Applications startup looking for biomarkers

Key technology: EOS – Big Data Storage and Management

what

Very low latency, flexible, highly scalable, reliable disk storage platform.

apps

- Data centers with wide range of quality of service requirements.
- Cloud service providers.
- Research infrastructures generating large data quantities.

tech specs

- Low latency. Very fast simple namespace with no database dependency.
- Unlimited scalability. Tunable file layout per directory.
- High distribution of data across nodes makes EOS highly resilient to hardware failures
- Multi protocol support. Main access via xroot, but also http, fuse, gridftp.
- Open Source Software with long term support at CERN.

added value

- On demand quality of service.
  → **Allows to quickly adapt and meet changing performance and reliability requirements.**
- Highly scalable, multi protocol.
  → **From 2 Pb to 256 Pb in 6 years. No inherent limit.**
- Cost effective, reliable data storage.
  → **Based on commodity hard disks. Just a Bunch Of Disks layout.**

Key technology: File Transfer Service (FTS)

what

Global fast and reliable bulk data mover. Optimization of existing network and storage resources for best transfer outcome.

apps

- Where large distributed data processing is necessary.
- Cloud applications
- Where there is critical need for sustained, reliable big data transfers.

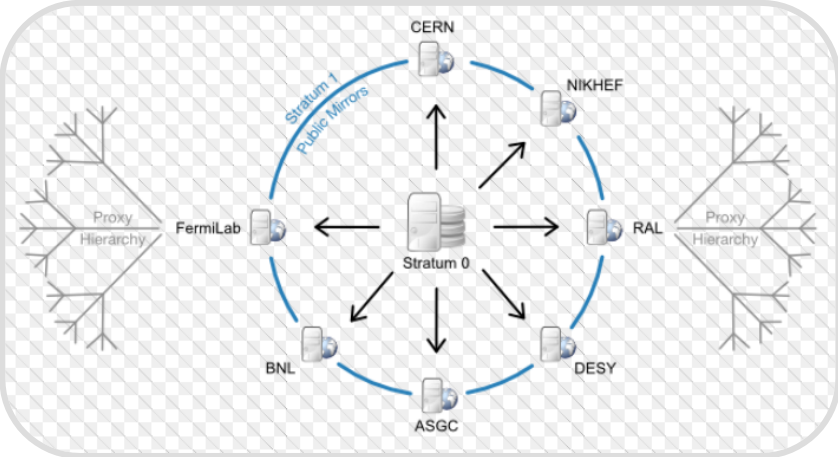
tech specs

- Multiprotocol, supports: http, xrootd, gridftp,srm..
- Endpoint centric. Automatically adapts the load to end point capabilities.
- Flexible multidimensional scheduler allows to customize resource sharing for specific jobs.
- User friendly real time web monitoring and dashboards.
- No configuration necessary for most cases. Easy web configuration if necessary.
- Open Source Software with long term support at CERN.

added value

- Experience running reliable transfers or the order of 20Pb/week for years.
- Intelligent, parallel transfers optimizing network and storage resources.
- Integrity & Reliability
  → **Checksums and retries ensure safe transfers**
- Simple to use.

Large scale reliable software distribution



WHAT

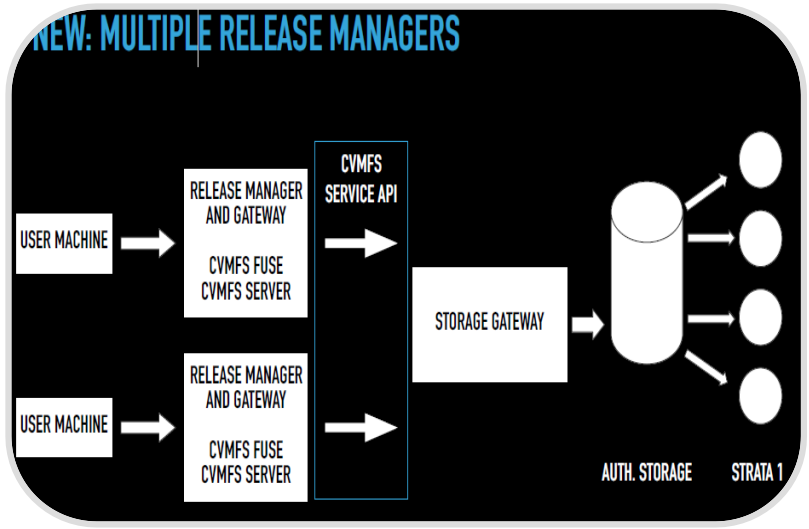
CernVM-FS is a Posix Read-only network file system based on HTTP.
It is designed an optimized for reliable software distribution of a large number of files to a large number of nodes.
It is used to deploy software on the worldwide-distributed computing infrastructure used to run data processing applications.

- ☑ Scalable, reliable and low- maintenance software distribution service;
- ☑ Transfers data and meta-data on demand and verifies data integrity by cryptographic hash;
- ☑ Aggressive caching and reduction of latency. Particularly adapted for software distribution;
- ☑ A Docker plugin allows serving the contents of image layers from CernVM FS repositories;
- ☑ Open source.

WHERE @ CERN

- ➢ All experiment software distribution uses CernVM-FS. Including the Grid nodes (T1 & T2) software distribution.

Large scale reliable software distribution



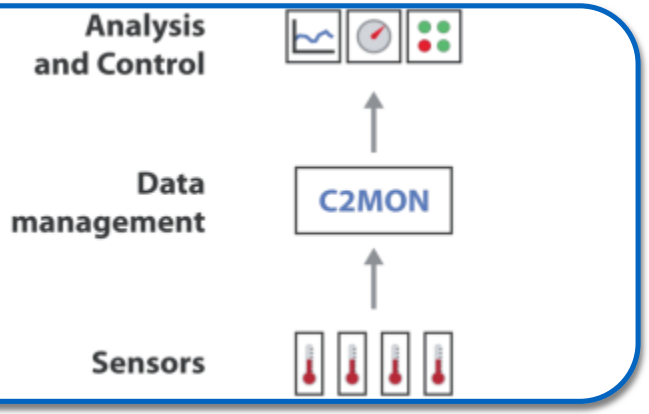
USES Outside HEP

- Multinational company with critical reliance on software
- Large data center applications
- Volunteer computing services for companies, monetiztion of sleeping resources
- Containerized, Data-Intensive Applications

WHO

- Widely used by the research community for software distribution (e.g. HEP, Biology)
- Big Pharma: Hoffman – La Roche.
- MesoSphere: Data Center Operating System (DCOS) (?)
- Aim to include the CernVM-FS plugin as part of the Docker distribution and widen use
- Open Source community, other similar potential users

Control and monitoring platform



WHAT

A modular Java framework called C2MON for large-scale industrial monitoring and control solutions. It has been developed for CERN's demanding infrastructure monitoring needs and is based on more than 10 years of experience. All core functionalities of a monitoring system are available and adaptable to a wide variety of monitoring systems.

HOW

- A modular and three-tier architecture: Data Acquisition, Server and Client API
- Decouples functionality and allows modular development
- Made to handle sudden and unforeseen machine breakdowns
- Integrated history browsing for industrial dashboards

WHY

- **Designed to use in large and complex control & monitoring environments with diverse infrastructure**
- **Robust, reliable and scalable open source architecture for many applications, like for example**
  - Can be used to realise IoT scenarios
  - Electrical grid operators
  - Oil & gas industry
  - Chemical industry
  - Patient monitoring
- **Modern HTML5 web interface for easy navigation**

A lot of systems to control

Controls
Computers

Electricity

Cryogenics

Magnets

85’000 Devices
> 2 Million I/O
Endpoints

***Much more** when*
including subsystems!

Safety

Cooling

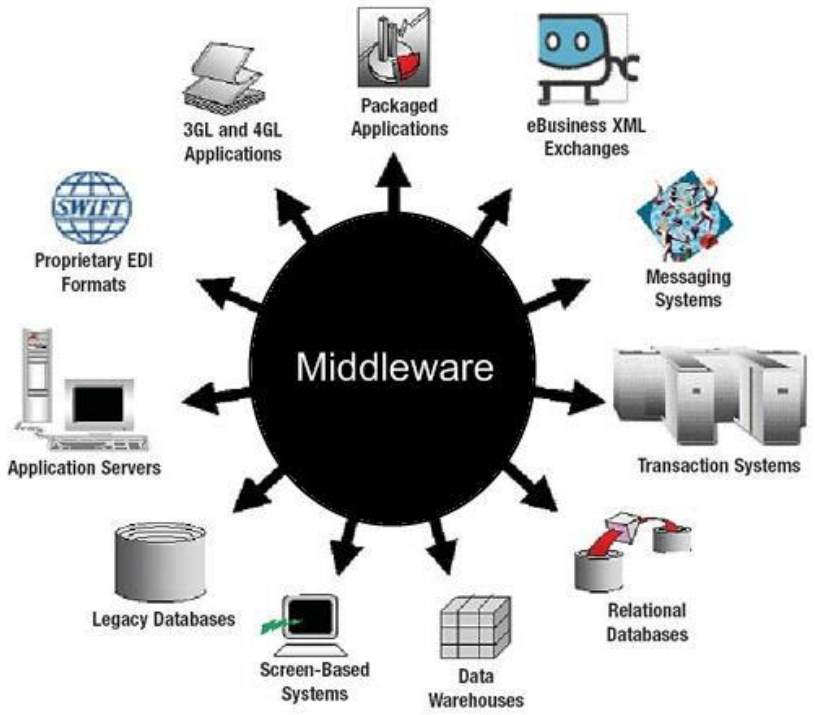
Ventilation

Vacuum

elasticsearch. as timeseries data storage

kibana ← **C2MON** → Grafana

CMW - Controls Middleware



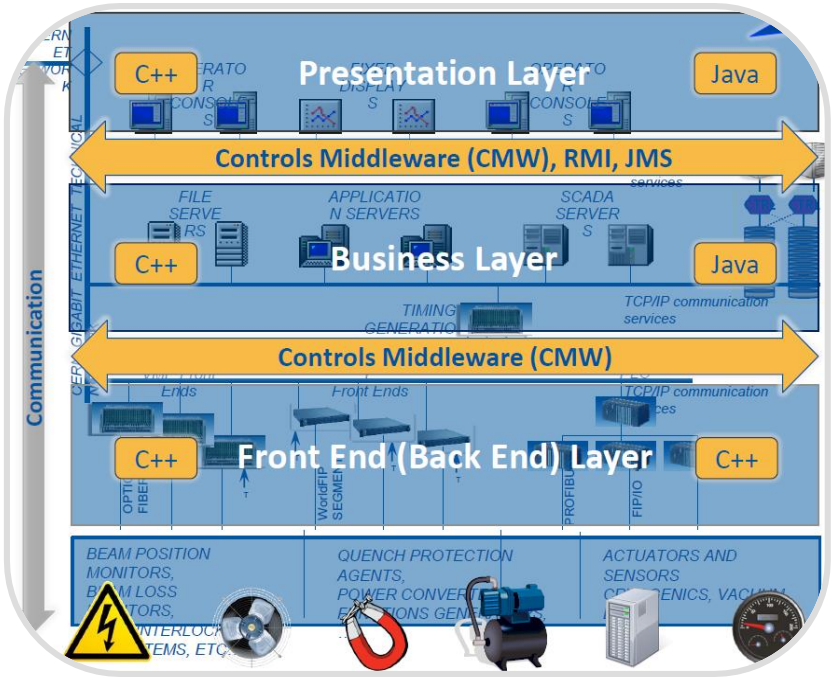
WHAT

Definition: Software*, which allows an application to interoperate with other software, without requiring the user to understand and to code the low-level operations required to achieve interoperability.*
CMW provides centrally managed middleware services (Directory/Naming, RBAC, Proxies,..) for all the whole accelerator complex.

WHERE @ CERN

- All CERN accelerator complex controls and interoperability.

CMW - Controls Middleware



WHY CMW

- ☑ Reliable and scalable transport of data between controls processes
- ☑ Critical core communication layer. Reliable communication in distributed environment; Decentralized – no brokers
- ☑ Centrally managed middleware services
- ☑ Real time handling of very large data sets: 4’000 CMW servers, 85’000 devices, 2’000’000 IO-points
- ☑ Uses ZeroMQ & RDA3 – Standards based
- ☑ Public API for developing clients & servers
- ☑ Provides comprehensive diagnostics
- ☑ CERN licence

WHERE @ CERN

- ➢ All CERN accelerator complex controls.

CMW - Controls Middleware

WHO

➢ LG-Display of Korea licensed CMW in 2017. To be used in all LG-D plants as part of the new MES system

Fast & easy machine-to-machine communication

WHAT

Open Platform Communications Unified Architecture (OPC-UA) is a machine to machine communication protocol (middleware) for industrial automation developed by the OPC Foundation based on a client/server model. The Quasar framework developed by CERN helps speed up and reduce cost of developing and maintaining OPC-UA components.

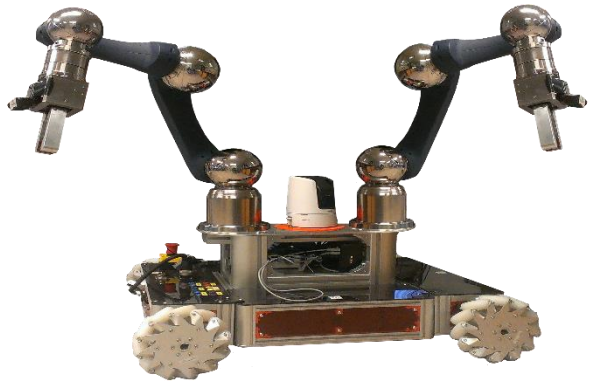
HOW

- Vertical (device/user oriented) and horizontal (peer to peer) system integration
- Model driven code generation
- Generation an executable OPC-UA server
- Development of a specific device OPC-UA server
- Re-use of common OPC-UA related code

WHY

- **Facilitate communication between machines.**
- **Virtual 'Plug & Play' for industrial components.**
- **Speeds up development of industrial components for companies and simplifies the maintenance process.**
- **Applications in building automation (HVAC, BAS), discrete manufacturing, process control solutions etc**
- **IoT devices applications**

Novel Robotic Platform - CERNBot



WHAT

CERNbot is a modular and flexible robotic platform developed at CERN for complex interventions in harsh environments. It is a highly customizable platform that can receive up to two commercial robotic arms working together in a coordinated manner.

HOW

- Autonomous interventions in hostile environments in particular in presence of ionizing radiation.
- Up to two coordinated commercial robotic arms installed providing for complex intervention capabilities.
- Real time operating system
- Future: Combined with a trained neural network to recognize scenes with the help of specific alpha-numeric markings

WHY

- **Highly versatile robotic platform.**
- **High payload for a platform this size.**
- **Very competitive cost base.**
- **Based on standard industrial components with predictable upgrade path.**
- **Intelligent inspection and intervention reducing employee exposure.**

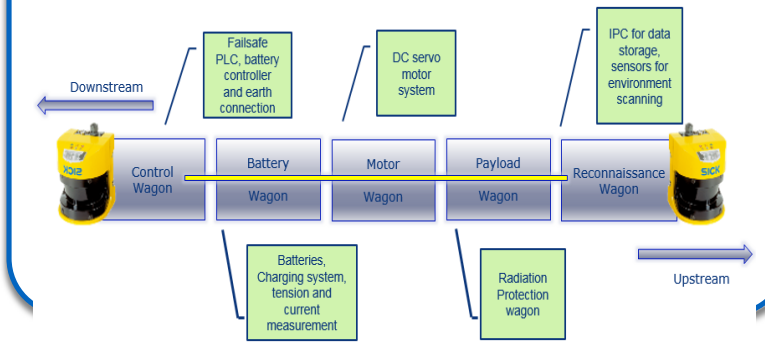
Train Inspection Monorail

WHAT

The Train Inspection Monorail is autonomous and versatile vehicle monitoring the 27-km long LHC tunnel and moving along a track suspended from the tunnel's ceiling. Packed with sensors for visual inspection, the robot can be programmed to perform real-time inspection intervention missions.

HOW

- Autonomous vehicle control
- Modular design
- Automated visual inspection
- Different sensors packages
- Handling robotics on board



WHY

- More safe as it does not expose humans to potential dangers
- Automated inspection missions provide wealth of data

USES

- Tele-manipulation & maintenance without human intervention
- Automated inspection missions provide wealth of data.
- Analytics being developed
- Photogrammetry applications for vital infrastructure
- Hazard monitoring & inspection (radiation, chemical, fire, etc)
- Safety – Fire brigade wagon
- Extendible to any other specific application as required by the tunnel operator

Extreme high resolution photon sensor

WHAT

Medipix is a family of read-out chips for particle imaging and detection. The original concept of Medipix is that it works like a camera, detecting and counting each individual particle hitting the pixels when its electronic shutter is open. This enables high-resolution, high-contrast, noise hit free images – making it unique for imaging applications.

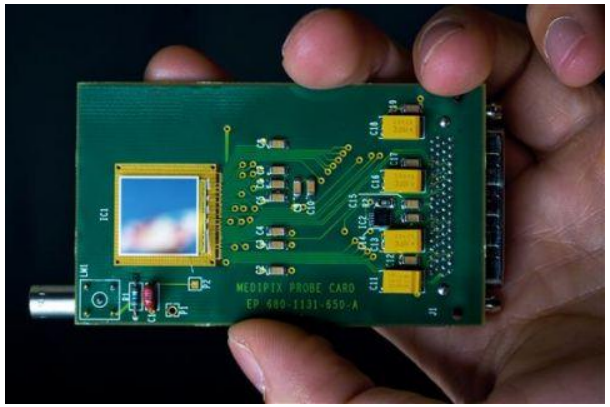
HOW

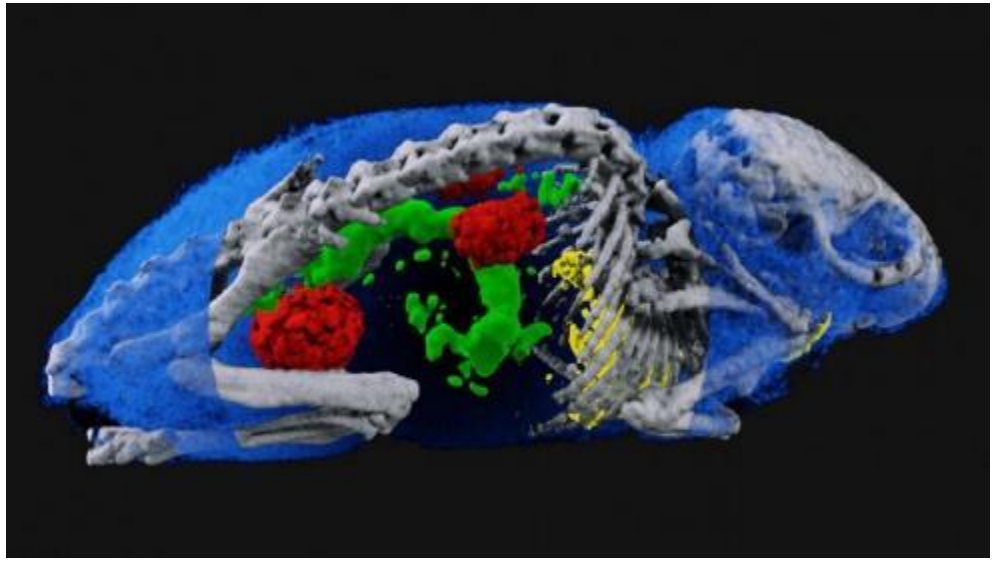
- "Clever" pixel electronics capable of processing every detected photon.
- Capable to record continuous stream of data, not just one frame
- Colours to indicate different energy levels of the photons

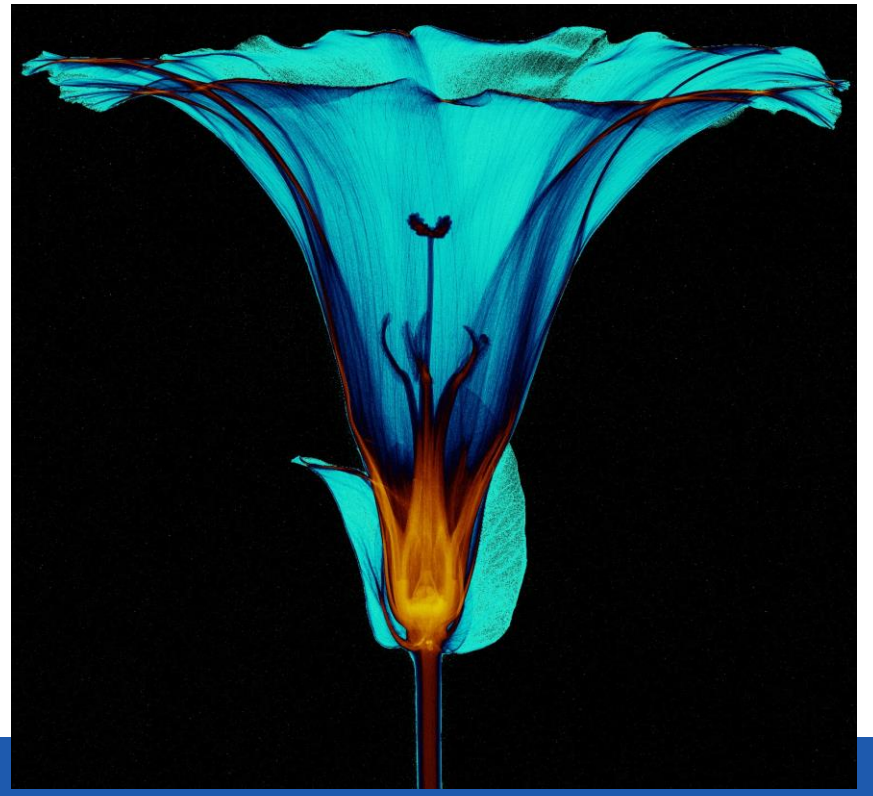
WHY

- **High resolution adds value in for example medical imaging, space dosimetry and material analysis**
- **Able to better visualize differences in material / tissue types**
- **Can also be used in non-destructive testing**
  - Detect various components
  - Detect cracks, voids
  - Detect contamination

Medipix

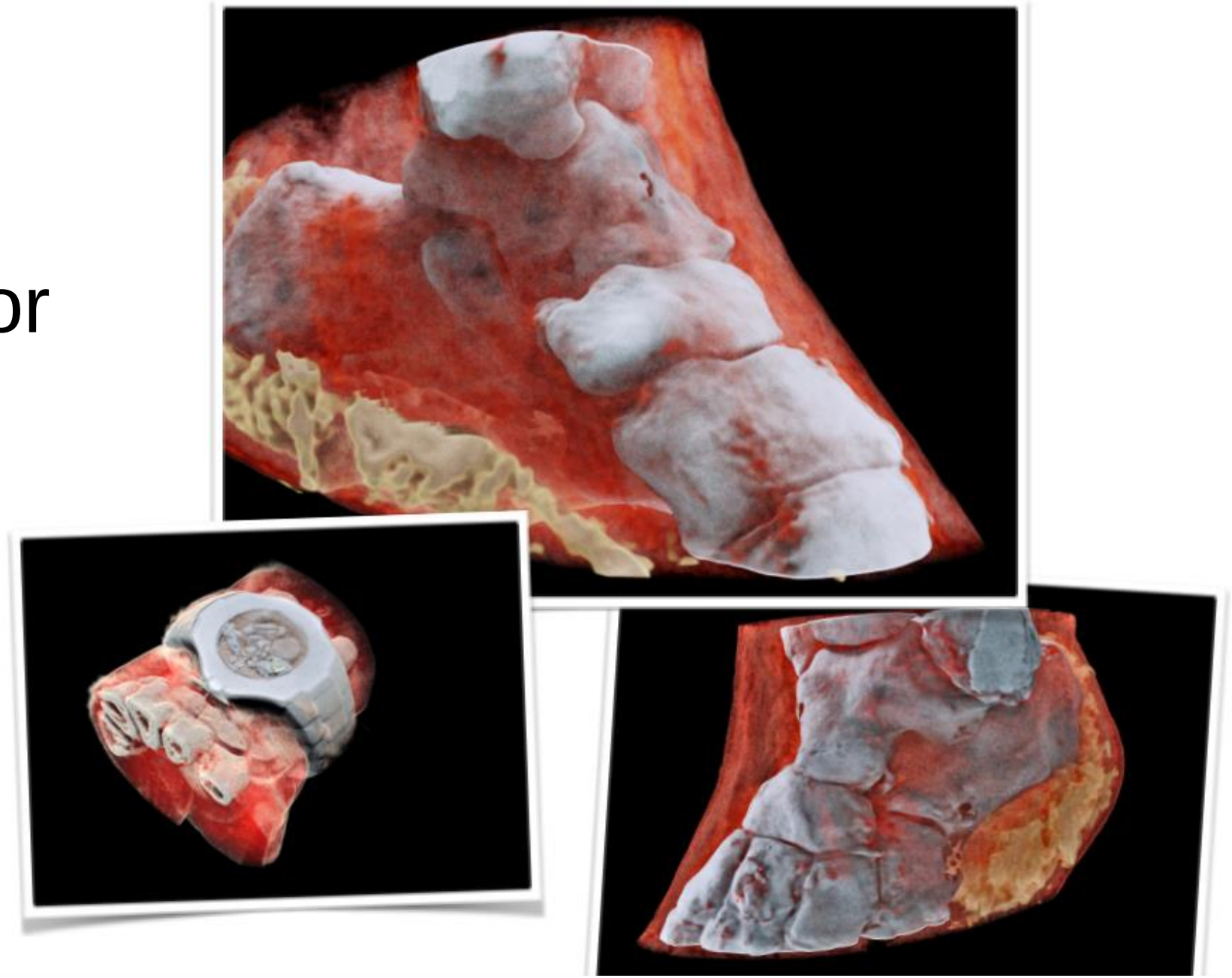






medipix collaboration

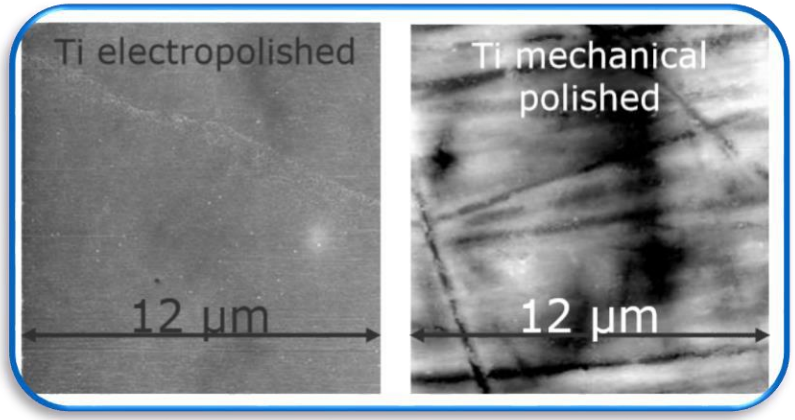
First in living human 3D Color X-ray



mars

Non Evaporative Getter thin film coatings

Titanium polishing



WHAT

Titanium polishing is a process to reduce the roughness, and thereby increase the brightness, of the metal surface. CERN's patented electropolishing method can produce polished titanium pieces with surface smoothness down to the nanometre level, with the additional benefit of lower power consumption compared to other methods.

HOW

- The metal can be polished down to the nanometer level
- Enables efficient detection of flaws in the surface
- Low power consumption
- Creates a shining, mirror like appearance
- Provides metalic purity and chemical passivity

WHY

- **Polish electrodes, which require an ultra-smooth surface to avoid sparks during operation**
- **There is practically no size limit on the sample to be treated, making the method appropriate for a wide range of applications, including:**
  - **Vacuum technology**
  - **Medical industry (e.g. implants, tools)**
  - **Jewellery, spectacle frames, watches**
  - **Aerospace (e.g. turbine blades)**
  - **Electronics, storage discs**

Compact Universal Cutter

WHAT

This orbital cutting machine has been designed to cut a broad range of pipes of different diameters and materials located in places which are particularly difficult to access. Once mounted on a pipe, the cutter operates autonomously without manual assistance, making it suitable to cut pipes which present health hazards.

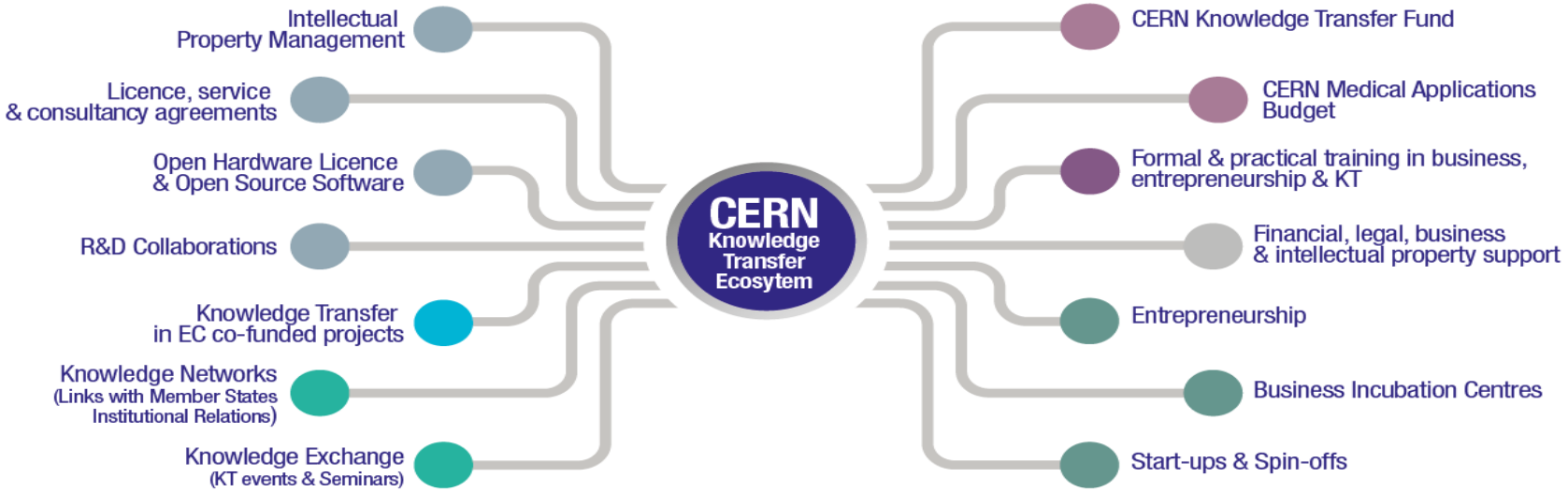
HOW

- Autonomous cutting
- Driven by hydraulic motor
- From 100 to 1200 mm
- Adaptable circular saw
- For different thicknesses
- For different materials

WHY

- **More safe as it does not expose humans to potential dangers**
- **One tool for many pipes**
- **Easy access**

Building an Ecosystem



Conclusions

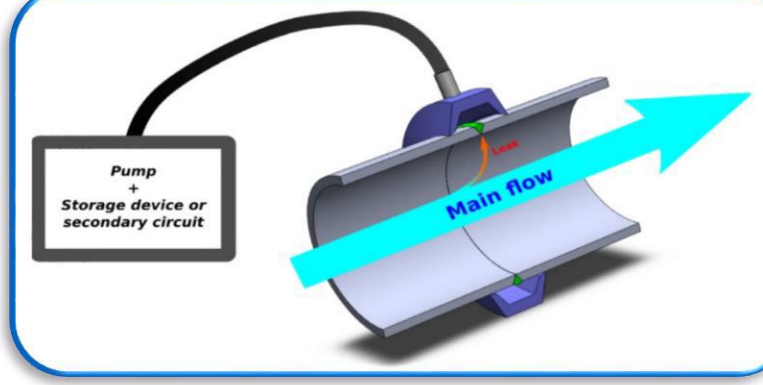
- The sample of CERN technologies shows the very wide spectrum of applications;
- From techs used in medical applications to autonomous robotic trains to pipe cutting;
- Where there is interest we will do our best to make collaboration happen!

Thank you for your attention !

www.cern.ch

www.cern.ch

Hood Clamshell Tool



WHAT

To guarantee the sealing of a joint, junction, pipe or tube, a very precise non-destructive technique with helium is used, which allows detection and measurement of small leaks. This technique provides a low cost option and is made possible through the use of the hood clamshell tool.

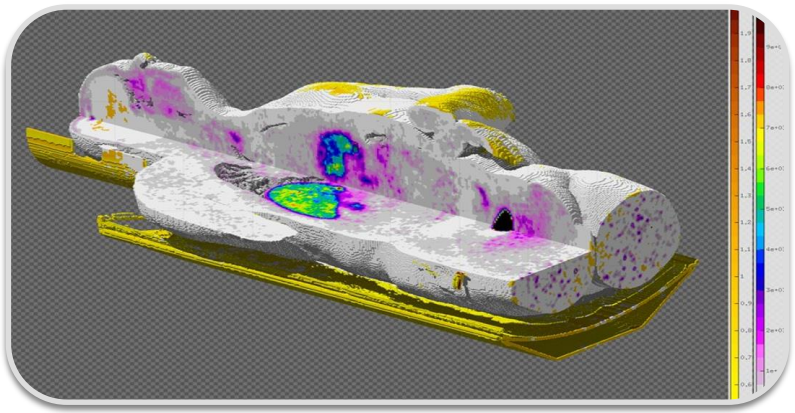
HOW

- Tool can fit any kind of seal, from 15mm to 500mm in diameter.
- Leak detector sensitivity: 10-10mbar.
- Equipped with a simple/innovative locking mechanism.
- Leak detector (helium sensor) is connected directly to the capsule.

WHY

- **Tool can be used in complex or restricted places.**
- **Quickness of installation and measurement.**
- **Easy to use.**
- **Adaptability to different diameters of tubes.**
- **Single compact unit tool.**
- **Self sealing system.**

Monte Carlo Simulation for particle interactions with matter



WHAT

FLUKA (Fluctuating Cascade) is a general purpose software tool for calculations of particle transport and interactions with matter. FLUKA can simulate the interaction and propagation in matter of about 60 different particles with high accuracy. FLUKA can handle very complex geometries and yields very accurate simulations.

WHERE @ CERN

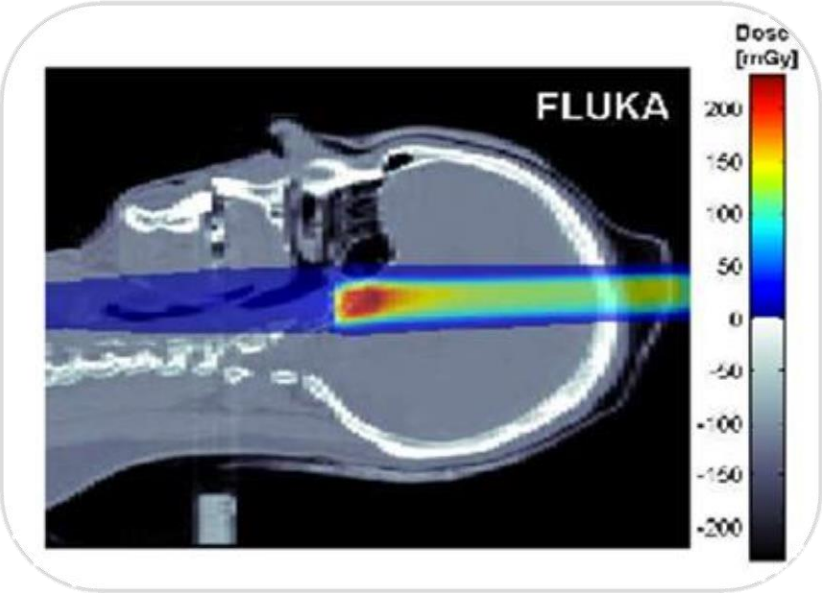
- Detector design
- Accelerator & components design (beam dumps etc)
- Target design
- Radiation protection & shielding
- Activated material management

Monte Carlo Simulation for particle interactions with matter

WHY FLUKA

- ☑ Sound and modern Physics models. High level of reliability. Predictability where no experimental data are directly available;
- ☑ Can simulate photons and electrons from 1 keV to >1000 TeV, neutrinos, muons, of any energy hadrons up to 20 TeV;
- ☑ Can track charged particles even in the presence of magnetic or electric fields;
- ☑ Can handle very complex geometries in terms of "voxels“;
- ☑ Rigorously controlled Physics models described in international scientific journals;
- ☑ A friendly user interface, FLAIR, is available, as well as a 3D visualisation too.

Monte Carlo Simulation for particle interactions with matter



USES outside HEP

- Hadron therapy Treatment planning systems (TPS)
- Optimization of medical particle therapy (including hw design)
- X-Ray simulations
- Shielding design (including space applications)
- Safety related to radiation protection
- Dose calculation
- Nuclear waste management

WHO

- Over 10’000 Academic and Research users
- Siemens AG
- Raysearch Laboratories
- Ansaldo Nucleare
- Dectris AG
- GMTH China
- Innocryst
- TUV SUD