



Η Πληροφορική @CERN

Florentia Protopsalti

IT-TD-SM

ΚΑΛΩΣΟΡΙΣΜΑ & ΕΙΣΑΓΩΓΗ



Computing @CERN

- Το CERN είναι ένα από τα πιο απαιτητικά περιβάλλοντα πληροφορικής στον ερευνητικό κόσμο.
- Η τεράστια ποσότητα δεδομένων που παράγεται από το LHC (μέχρι και petabyte ανά δευτερόλεπτο).
- Πώς η πληροφορική επιτρέπει την αποθήκευση, επεξεργασία, ανάλυση και προσομοίωση αυτών των δεδομένων.



Computing @CERN

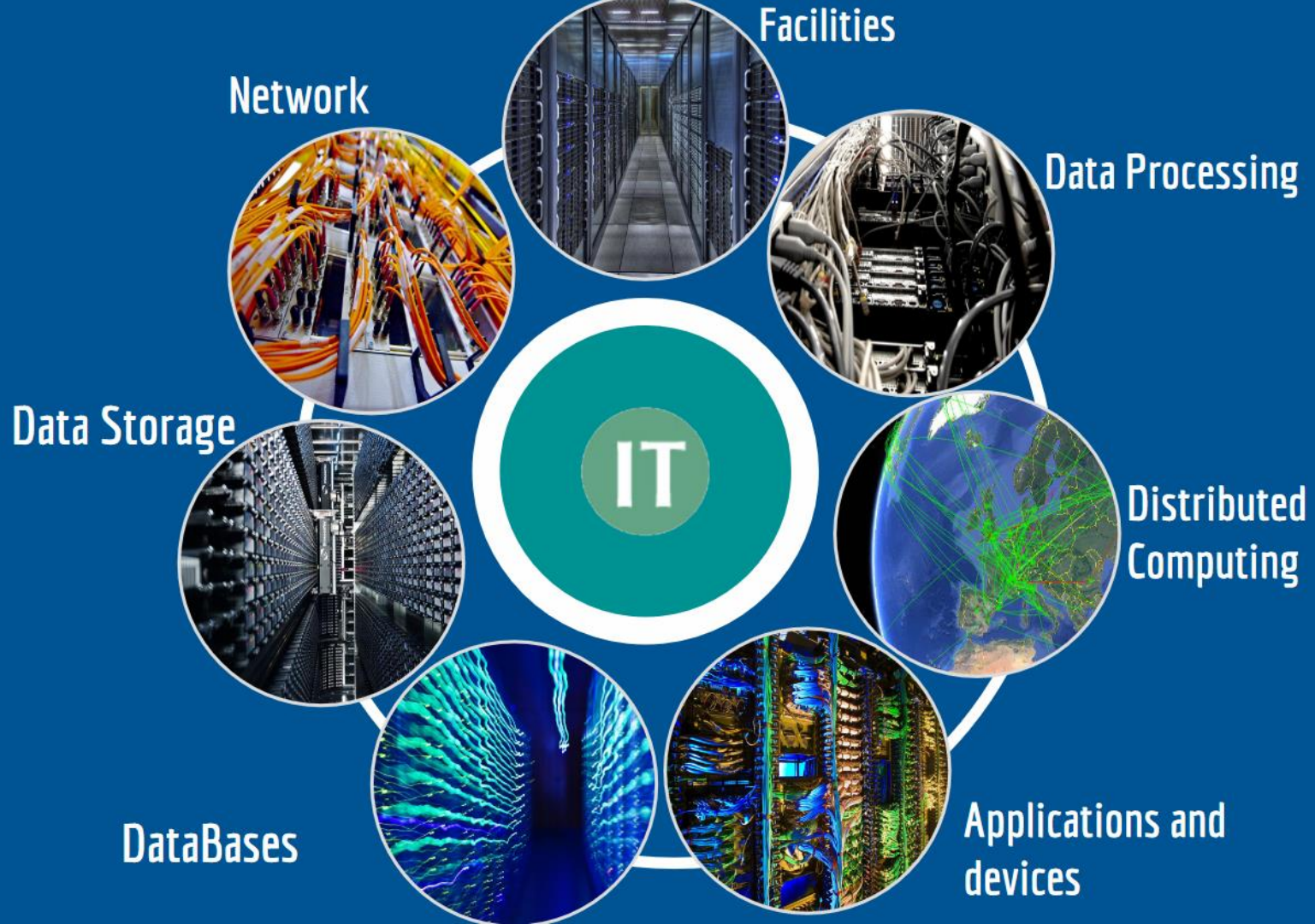
- Σύνδεση με πειράματα όπως ATLAS, CMS, ALICE, and LHCb.
- Η σημασία της επεξεργασίας δεδομένων σε πραγματικό χρόνο για την ανίχνευση σπάνιων γεγονότων και την ανάλυση συγκρούσεων.
- Ο Παγκόσμιος Ιστός δημιουργήθηκε και αναπτύχθηκε αρχικά στο CERN το 1991 από τον Tim Berners-Lee.

Ο Ρόλος του IT @CERN

ΚΥΡΙΟΣ ΡΟΛΟΣ ΣΤΗΝ ΕΡΕΥΝΑ
ΤΟ ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΥΠΟΣΤΗΡΙΖΕΙ ΤΗΝ ΕΡΕΥΝΑ
ΠΑΡΕΧΟΝΤΑΣ ΑΠΑΡΑΙΤΗΤΗ
ΥΠΟΛΟΓΙΣΤΙΚΗ ΙΣΧΥ ΚΑΙ
ΔΙΧΕΙΡΙΖΕΤΑΙ ΤΑ ΔΕΔΟΜΕΝΑ
ΑΠΟΤΕΛΕΣΜΑΤΙΚΑ

ΛΕΙΤΟΥΡΓΙΚΗ
ΑΠΟΔΟΤΙΚΟΤΗΤΑ
ΚΕΝΤΡΙΚΟ ΓΙΑ ΤΗ ΔΙΑΤΗΡΗΣΗ
ΤΗΣ ΛΕΙΤΟΥΡΓΙΚΗΣ
ΑΠΟΔΟΤΙΚΟΤΗΤΑΣ, ΤΟ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΔΙΑΣΦΑΛΙΖΕΙ ΙΣΧΥΡΑ
ΣΥΣΤΗΜΑΤΑ ΓΙΑ ΤΗΝ
ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ
ΣΕ ΠΡΑΓΜΑΤΙΚΟ ΧΡΟΝΟ

ΚΑΙΝΟΤΟΜΙΑ ΚΑΙ ΑΝΑΠΤΥΞΗ
ΕΣΤΙΑΣΜΕΝΟ ΣΤΗΝ
ΚΑΙΝΟΤΟΜΙΑ, ΑΝΑΠΤΥΣΣΕΙ
ΤΕΧΝΟΛΟΓΙΕΣ ΔΙΧΜΗΣ,
ΑΠΑΡΑΙΤΗΤΕΣ ΓΙΑ
ΠΡΩΤΟΠΟΡΙΑΚΕΣ
ΕΠΙΣΤΗΜΝΙΚΕΣ ΑΝΑΚΑΛΥΨΕΙΣ



Κύρια Συστατικά του Computing @CERN

- Data Centers



Data Centers

Meyrin Data Center (Switzerland)

- Δημιουργήθηκε το 1972
- 450 000 processor cores και 10 000 servers run 24/7
- Το 90% των πόρων για την υπολογιστική ισχύ στο Data Centre παρέχεται μέσω ενός ιδιωτικού cloud βασισμένου στο OpenStack

Data Centers

Prevessin Data Center (France)

- Εγκαινιάστηκε το Φεβρουάριο του 2024
- Ανάπτυξη δεδομένων (45 petabytes ανά εβδομάδα)
- 6000 τετραγωνικά μέτρα – 6 δωμάτια για εξοπλισμό πληροφορικής
 - Ενεργειακά αποδοτική δομή
 - Σύστημα ανακύκλωσης νερού



Κύρια Συστατικά του Computing @CERN

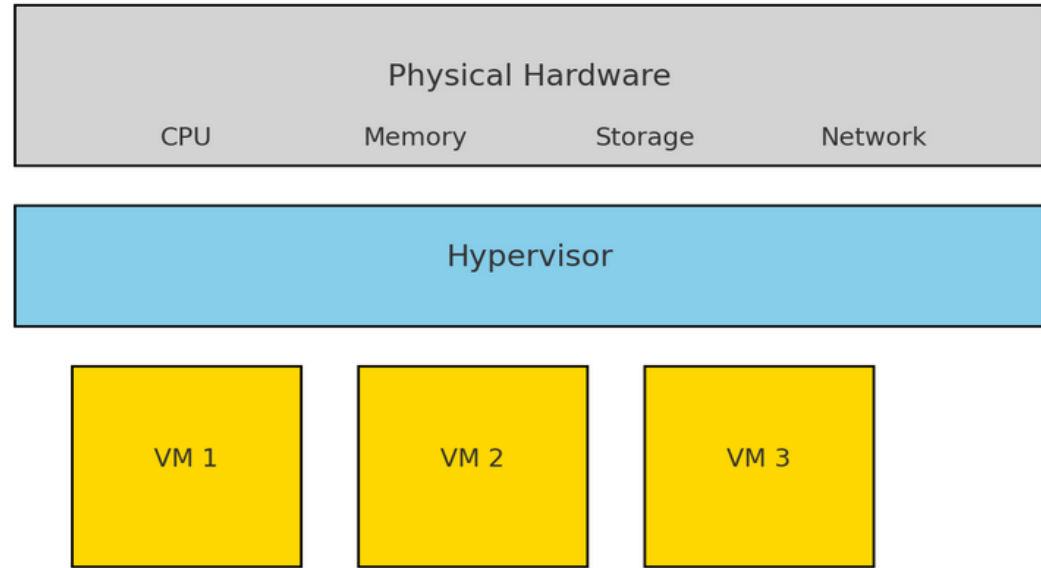
- Data Centers
- CERN Openstack Private Cloud





CERN OpenStack Private Cloud

- Ξεκίνησε η διερεύνηση το 2009
- Άρχισε τη χρήση της Εικονικοποίησης (Virtualization) το 2013
 - 300K cores and ~15k εικονικές μηχανές (VMs)
- Private: αποκλειστική χρήση, ελεγχόμενο περιβάλλον, προσαρμογές ασφαλείας



1. **Physical HW:** στοιχεία όπως CPU, η μνήμη, η αποθήκευση και το δίκτυο
2. **Hypervisor:** Το επίπεδο λογισμικού που διαχειρίζεται και κατανέμει πόρους από το υπολογιστικό υλικό προς τις Εικονικές Μηχανές (VMs).
3. **VMs:** Κάθε Εικονική Μηχανή εκτελεί το δικό της λειτουργικό σύστημα και εφαρμογές, απομονωμένα από άλλες Εικονικές Μηχανές.

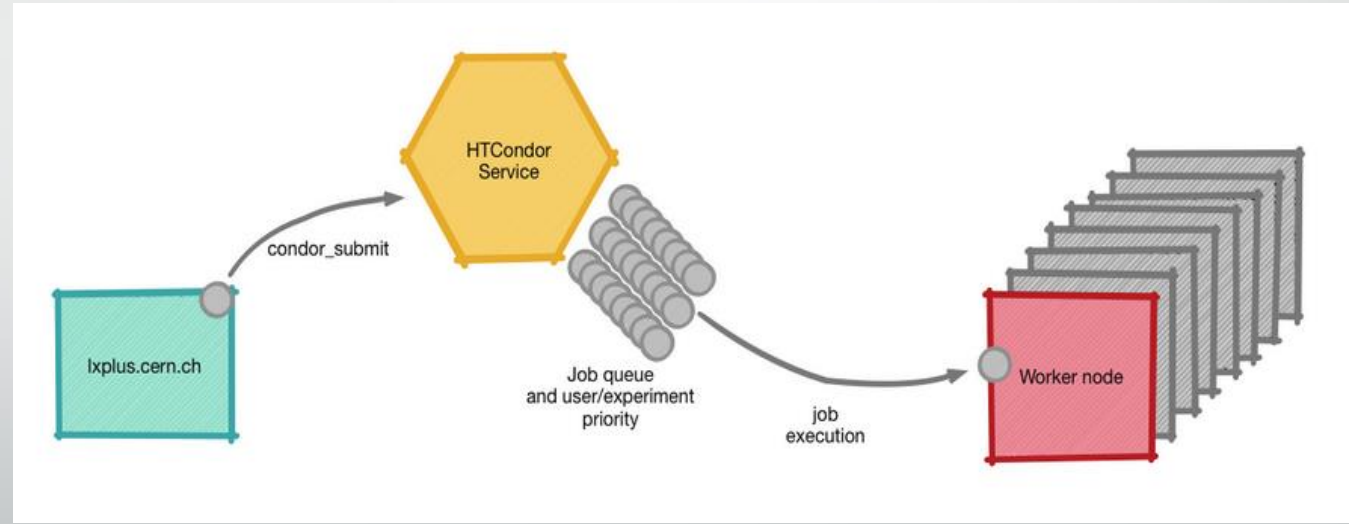


Γιατί Εικονικές Μηχανές?

- Αποτελεσματικότητα Κόστους
- Καλύτερη Χρησιμοποίηση Πόρων
- Εύκολη δημιουργία/διαγραφή τους χωρίς ανθρώπινη παρέμβαση
 - Εύκολες μεταφορές/αναβαθμίσεις
- **Παράδειγμα:** Batch Service (HTCondor) -> Μεγάλος καταναλωτής πόρων στο IT

Τι είναι το HTCondor?

[HTCondor](#) είναι ένα open-source λογισμικό για καλύτερη διαχείριση των πόρων



Κύρια Συστατικά του Computing @CERN

- Data Centers
- CERN Openstack Private Cloud
- High Performance Computing (HPC) Υποδομή



High-Performance Computing (HPC)

- Επιτρέπει την επεξεργασία και ανάλυση τεράστιων ποσοτήτων δεδομένων από πειράματα όπως ο Μεγάλος Επιταχυντής Αδρονίων (LHC).
- Παρέχει την απαραίτητη υπολογιστική ισχύ για πολύπλοκους υπολογισμούς.
- Το HPC επιτρέπει στους ερευνητές να εκτελούν προσομοιώσεις που αναπαράγουν τις συνθήκες και τα αποτελέσματα συγκρούσεων σωματιδίων.
- Αυτές οι προσομοιώσεις βοηθούν στην πρόβλεψη και επιβεβαίωση θεωρητικών μοντέλων
- Απαιτεί υπολογιστική ισχύ, συχνά περιλαμβάνοντας δισεκατομμύρια υπολογισμούς ανά δευτερόλεπτο.



High-Performance Computing (HPC)

Παράλληλη Υπολογιστική

- **Ανάπτυξη Αλγορίθμων:** Ανάπτυξη και βελτιστοποίηση αλγορίθμων ειδικά σχεδιασμένων για παράλληλους υπολογισμούς, διασφαλίζοντας ότι τα συστήματα HPC μπορούν να ανταποκριθούν στις μοναδικές απαιτήσεις της έρευνας.
- **Βελτιστοποίηση Απόδοσης:**
 - Το HPC βασίζεται σε τεχνικές παράλληλων υπολογισμών.
 - Οι εργασίες χωρίζονται σε μικρότερες, ανεξάρτητες εργασίες που μπορούν να επεξεργαστούν ταυτόχρονα.
 - Μειώνει τον χρόνο επεξεργασίας και αυξάνει την απόδοση.
 - **Παράδειγμα:** Monte Carlo προσομοιώσεις
 - Οι προσομοιώσεις χρησιμοποιούν τυχαία δειγματοληψία για να εκτιμήσουν τις πιθανότητες των διαφορετικών αποτελεσμάτων.

Κύρια Συστατικά του Computing @CERN

- Data Centers
- CERN Openstack Private Cloud
- High Performance Computing (HPC) Υποδομή
- Worldwide LHC Computing Grid (WLCG)



Worldwide LHC Computing Grid (WLCG)



Running jobs: 365644
Active CPU cores: 807139
Transfer rate: 21.54 GiB/sec

Worldwide LHC Computing Grid (WLCG)

- Ένα παγκόσμιο δίκτυο κατανεμημένων υπολογιστών.
- Συντονίζεται από το τμήμα πληροφορικής του CERN.
- Εκτείνεται σε 170 κέντρα δεδομένων σε περισσότερες από 40 χώρες.
- Παρέχει τους πόρους για την αποθήκευση, διανομή και ανάλυση PB δεδομένων κάθε χρόνο.

Πώς να το χρησιμοποιήσετε?

- Συμβάλλετε με πόρους υπολογιστών από το σπίτι(LHC@home)
 - Κατεβάστε και εγκαταστήστε μια εφαρμογή
- WLCG προγράμματα: Επικοινωνήστε με το τοπικό τμήμα φυσικής του πανεπιστημίου
 - Πρόσβαση στο LHC computing grid
 - Εγγραφείτε σε ένα Virtual organisation



Το WLCG οργανώνεται σε επίπεδα:

Επίπεδο-0: Το Κέντρο Δεδομένων CERN είναι η κύρια εγκατάσταση για την αρχική επεξεργασία των δεδομένων.

- Λαμβάνει τα ακατέργαστα δεδομένα απευθείας από τους ανιχνευτές LHC.
- Δημιουργεί αντίγραφα των δεδομένων για περαιτέρω διανομή.
- Αποθηκεύει ένα αντίγραφο των ακατέργαστων δεδομένων και των επεξεργασμένων δεδομένων για μακροχρόνια αρχειοθέτηση.
- Το Επίπεδο-0 διαθέτει μερικά από τα πιο προηγμένα συστήματα HPC στον κόσμο.



Επίπεδο -1 : Περιφερειακά Κέντρα Δεδομένων

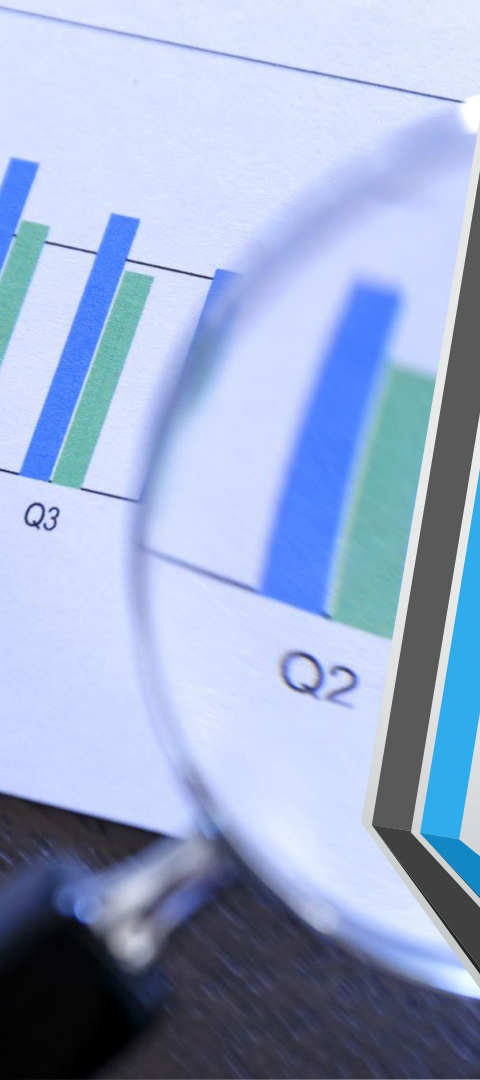
- Λαμβάνει δεδομένα από το Επίπεδο-0
- Εκτελεί εργασίες όπως αναγνώριση και ανάλυση.
- Κέντρα Επίπεδου-1 σε χώρες όπως η Γερμανία, η Ιταλία, η Γαλλία

Επίπεδα-2 και 3: Πανεπιστήμια και ερευνητικά ιδρύματα

Κύρια Συστατικά του Computing @CERN

- Data Centers
- CERN Openstack Private Cloud
- High Performance Computing (HPC) Υποδομή
- Worldwide LHC Computing Grid (WLCG)
- Διαχείριση και Αποθήκευση Δεδομένων





Διαχείριση και Αποθήκευση Δεδομένων

- **Μαζική Δημιουργία Δεδομένων:** Οι αρχικοί ρυθμοί δεδομένων του LHC είναι κλίμακας PB/s. Η διαχείριση του φόρτου δεδομένων είναι μια βασική ευθύνη του τμήματος IT.
- **Φιλτράρισμα και Μείωση Δεδομένων:**
 - **Φιλτράρισμα Δεδομένων σε Πραγματικό Χρόνο:** Χρήση προηγμένων αλγορίθμων για το φιλτράρισμα και μείωση των δεδομένων σε πραγματικό χρόνο (μειωμένα σε GB/s μετά το φιλτράρισμα δεδομένων).
 - **Συλλογή Δεδομένων Υψηλής Απόδοσης:** Διασφάλιση ότι τα συστήματα συλλογής δεδομένων μπορούν να διαχειριστούν τους εξαιρετικά υψηλούς ρυθμούς δεδομένων, ελαχιστοποιώντας την απώλεια δεδομένων και εξασφαλίζοντας την ακρίβεια.

Διαχείριση και Αποθήκευση Δεδομένων

Λύσεις Αποθήκευσης Δεδομένων

- Δεδομένα που χρειάζεται να προσπελαστούν γρήγορα αποθηκεύονται σε δίσκους υψηλής απόδοσης (disk arrays).
- Δεδομένα που προσπελάζονται λιγότερο συχνά αρχειοθετούνται σε μαγνητικές ταινίες (magnetic tapes).
- EOS - κατακεμημένο σύστημα αποθήκευσης
- CERNBox - Cloud storage service
- Ceph - Κατακεμημένο object storage system
- **Cloud Integration:**
 - CERN Openstack Cloud
 - Εμπορικοί Πάροχοι Cloud (e.g Amazon Web Services, Microsoft Azure)



Αρχειοθέτηση και Διατήρηση Δεδομένων

- **Μακροπρόθεσμη Διατήρηση Δεδομένων:** Διασφάλιση της μακροχρόνιας διατήρησης των επιστημονικών δεδομένων, διατηρώντας τα προσβάσιμα και χρήσιμα για μελλοντική έρευνα και επαναανάληψη.

Προσβασιμότητα και Κοινοποίηση Δεδομένων

- **Open Data Initiatives:** Δέσμευση του CERN για ανοικτή επιστήμη διευκολύνοντας την κοινοποίηση δεδομένων με την παγκόσμια ερευνητική κοινότητα (π.χ. πλατφόρμες).
 - **Παράδειγμα:** [CERN Open Data Portal](#), μια πλατφόρμα αφιερωμένη στην κοινοποίηση δεδομένων από διάφορα πειράματα, με την παγκόσμια ερευνητική κοινότητα.

Κύρια Συστατικά του Computing @CERN

- Data Centers
- CERN Openstack Private Cloud
- High Performance Computing (HPC) Υποδομή
- Worldwide LHC Computing Grid (WLCG)
- Διαχείριση και Αποθήκευση Δεδομένων
- Δικτύωση



Δικτύωση (Network)

- **Παγκόσμια Δικτύωση:** Διατήρηση μιας ισχυρής, υψηλής ταχύτητας δικτυακής υποδομής που συνδέει το CERN με ερευνητικά ιδρύματα παγκοσμίως, υποστηρίζοντας τη συνεργασία σε παγκόσμιο επίπεδο.
- **Υψηλή Ταχύτητα Συνδεσιμότητας**
- **Προετοιμασία για Μελλοντικά Πειράματα:**
 - Το CERN σχεδιάζει μελλοντικά πειράματα και αναβαθμίσεις του LHC.
 - Το τμήμα IT εργάζεται για την κλιμάκωση της δικτυακής υποδομής για να καλύψει τις αυξημένες απαιτήσεις.
 - Επέκταση του εύρους ζώνης για την επεξεργασία μεγαλύτερων όγκων δεδομένων.
 - Ενίσχυση της εφεδρείας για την εξασφάλιση της αξιοπιστίας του δικτύου.
 - Υιοθέτηση νέων τεχνολογιών.



Κύρια Συστατικά του Computing @CERN

- Data Centers
- CERN Openstack Private Cloud
- High Performance Computing (HPC) Υποδομή
- Worldwide LHC Computing Grid (WLCG)
- Διαχείριση και Αποθήκευση Δεδομένων
- Δικτύωση
- Λογισμικό και Ενδιάμεσο Λογισμικό



Λογισμικό και Ενδιάμεσο Λογισμικό (Software and Middleware)

- **Ανάπτυξη Λογισμικού:**
 - Ανάπτυξη και συντήρηση προσαρμοσμένων λύσεων λογισμικού προσαρμοσμένων στις ανάγκες του CERN.
 - Παραδείγματα:
 - ROOT: Ένα προσαρμοσμένο data analysis framework
 - EOS: Ένα σύστημα αποθήκευσης υψηλής απόδοσης και για τη διαχείριση δεδομένων.
 - Invenio: Μια ψηφιακή βιβλιοθήκη για τη διαχείριση και διατήρηση επιστημονικών δημοσιεύσεων και δεδομένων.
- **HTCondor**
 - Είναι ένα σύστημα λογισμικού που δημιουργεί ένα περιβάλλον Υπολογισμού Υψηλής Απόδοσης.

Κύρια Συστατικά του Computing @CERN

- Data Centers
- CERN Openstack Private Cloud
- High Performance Computing (HPC) Υποδομή
- Worldwide LHC Computing Grid (WLCG)
- Διαχείριση και Αποθήκευση Δεδομένων
- Δικτύωση
- Λογισμικό και Ενδιάμεσο Λογισμικό
- Ακεραιότητα και Ασφάλεια Δεδομένων



Ακεραιότητα και Ασφάλεια Δεδομένων

- Εξασφάλιση της Ακεραιότητας των Δεδομένων καθ' όλη τη διάρκεια του κύκλου ζωής τους.
- **Μέτρα Ασφαλείας:** Δεδομένης της ευαισθησίας και της σημασίας των δεδομένων, υπάρχουν προηγμένα πρωτόκολλα ασφαλείας για την προστασία από μη εξουσιοδοτημένη πρόσβαση, παραβιάσεις δεδομένων και άλλες κυβερνοαπειλές.
- Βασικά μέτρα ασφαλείας περιλαμβάνουν:
 - Τείχη προστασίας (Firewalls)
 - Παρακολούθηση σε πραγματικό χρόνο για προστασία από κυβερνοαπειλές.

Κύρια Συστατικά του Computing @CERN

- Data Centers
- CERN Openstack Private Cloud
- High Performance Computing (HPC) Υποδομή
- Worldwide LHC Computing Grid (WLCG)
- Διαχείριση και Αποθήκευση Δεδομένων
- Δικτύωση
- Λογισμικό και Ενδιάμεσο Λογισμικό
- Ακεραιότητα και Ασφάλεια Δεδομένων
- Αποδοτικότητα Ενέργειας



Αποδοτικότητα Ενέργειας

- Virtualization and Cloud Computing
- Βελτιστοποίηση Κώδικα για αποδοτικότητα ώστε να μειωθεί το υπολογιστικό φορτίο για την επεξεργασία δεδομένων.
- Μείωση εκπομπών άνθρακα σε όλες τις λειτουργίες.
- Πρωτοβουλίες για Αποδοτικότητας Ενέργειας:
 - Οι προσπάθειες περιλαμβάνουν ψύξη (cooling) ενεργειακής απόδοσης, χρήση ανανεώσιμων πηγών ενέργειας και βελτιστοποίηση λογισμικού για τη μείωση των αποβλήτων.



Κύρια Συστατικά του Computing @CERN

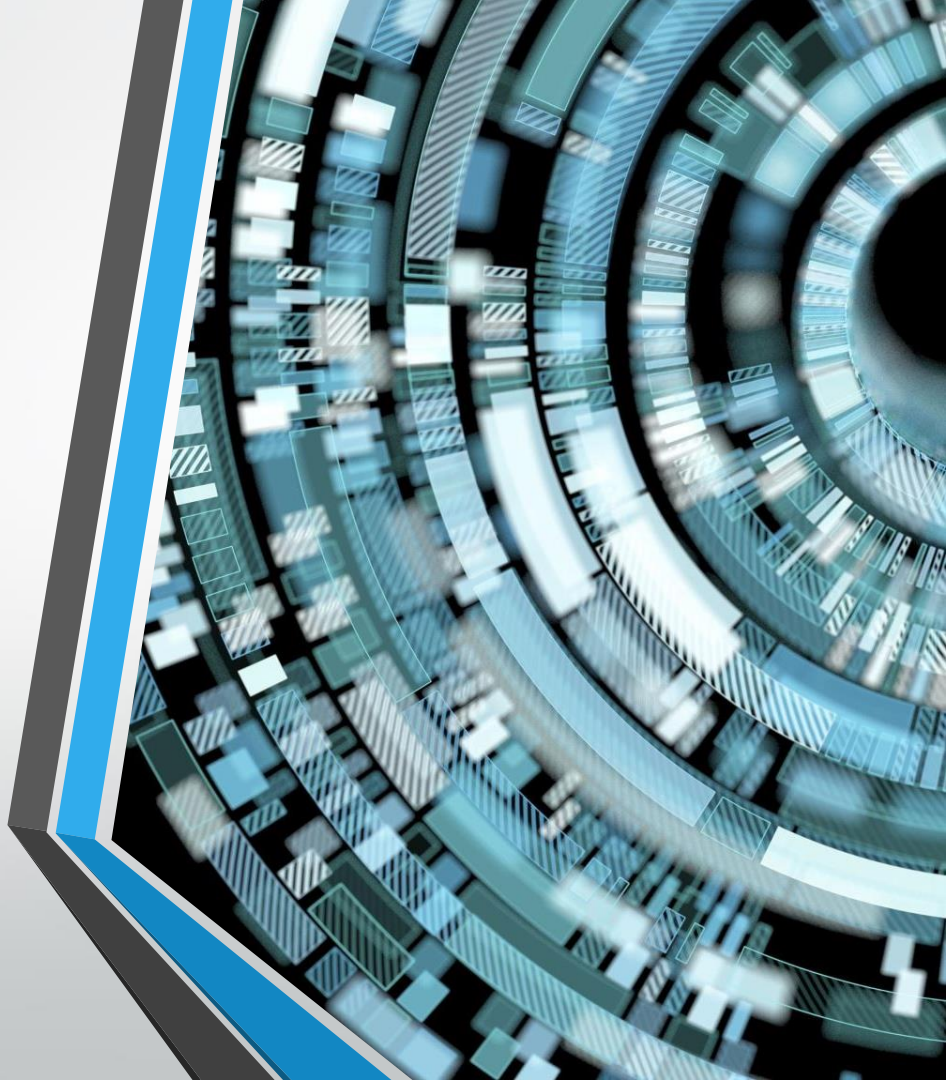
- Data Centers
- CERN Openstack Private Cloud
- High Performance Computing (HPC) Υποδομή
- Worldwide LHC Computing Grid (WLCG)
- Διαχείριση και Αποθήκευση Δεδομένων
- Δικτύωση
- Λογισμικό και Ενδιάμεσο Λογισμικό
- Ακεραιότητα και Ασφάλεια Δεδομένων
- Αποδοτικότητα Ενέργειας
- Υποστήριξη Συνεργασίας



Υποστήριξη Συνεργασίας

Διευκόλυνση Συνεργασίας στην έρευνα

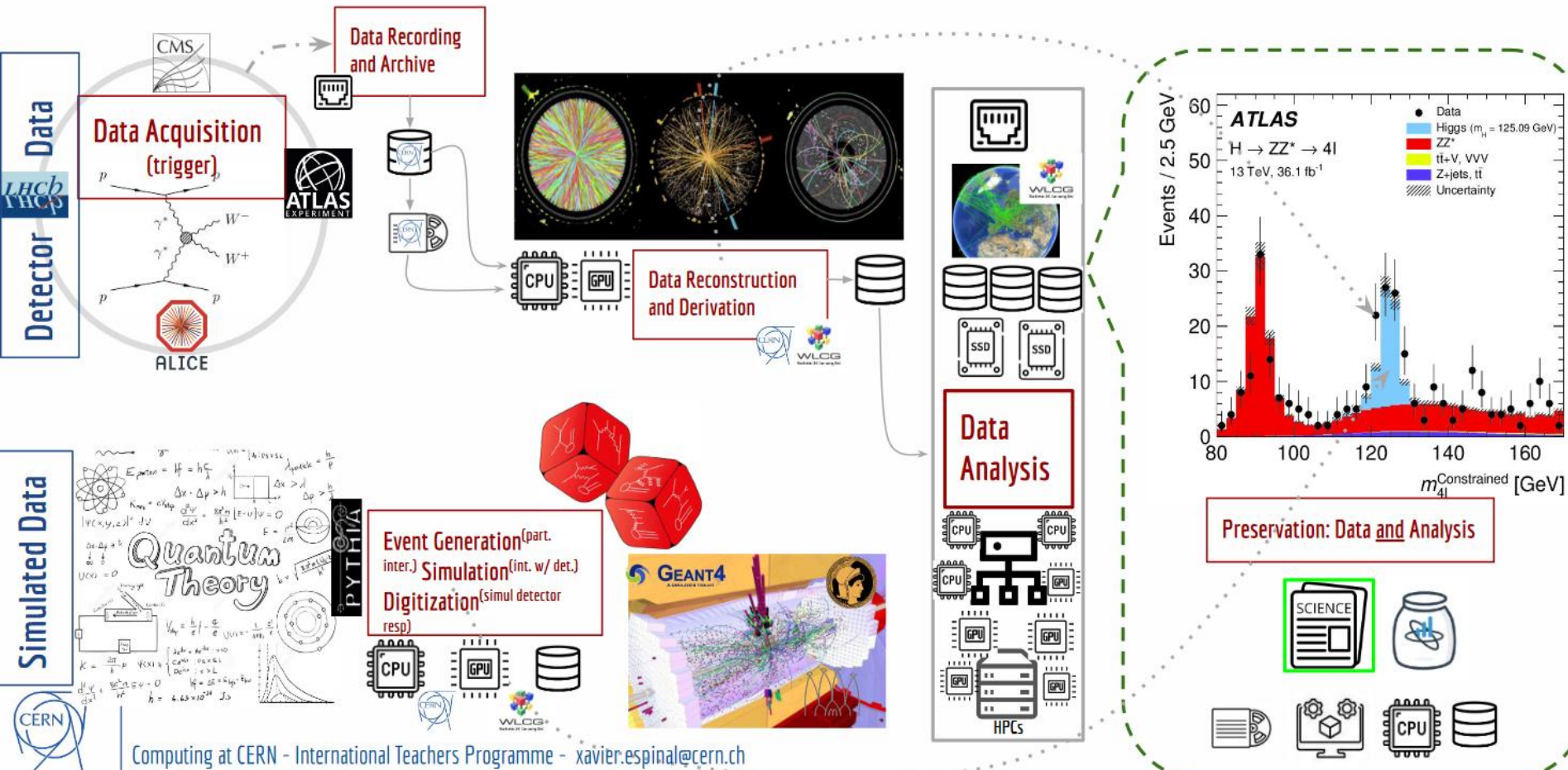
- Το τμήμα IT παρέχει εργαλεία και πλατφόρμες για αποτελεσματική παγκόσμια συνεργασία.
- **Παραδείγματα:**
 - **Indico:** Διαχείριση εκδηλώσεων και προγραμματισμός συνεδρίων.
 - **Zoom:** Βιντεοδιάσκεψη για εικονικές συναντήσεις και συζητήσεις.
 - **CERNBox:** Ασφαλής αποθήκευση στο cloud και κοινή χρήση αρχείων.
 - **GitLab:** Συνεργατική ανάπτυξη λογισμικού και έλεγχος εκδόσεων.
 - **ServiceNow:** Σύστημα υποβολής αιτημάτων (ticketing system).



Παράδειγμα: Ανάλυση Δεδομένων για το Πείραμα ATLAS

- Κατανομή Δεδομένων
- Κατανεμημένη Επεξεργασία
- Διαχείριση Εργασιών
- Πρόσβαση σε Δεδομένα
- Συνεργασία και Κοινοποίηση Αποτελεσμάτων

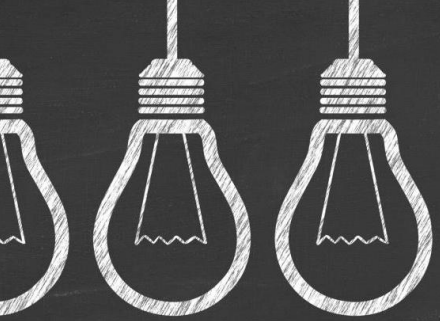
Computing at CERN: the big picture



Πρόοδοι στην Πληροφορική στο CERN

High-Luminosity LHC (HL-LHC) προετοιμασίες

- **Αυξημένες Απαιτήσεις Δεδομένων:** Το HL-LHC, το οποίο προγραμματίζεται να ολοκληρωθεί μέχρι το 2029, θα επιτρέψει στους φυσικούς να μελετήσουν τους μηχανισμούς με μεγαλύτερη λεπτομέρεια. Ο στόχος είναι να αυξηθεί η ολοκληρωμένη λαμπρότητα κατά έναν παράγοντα 10 πέρα από την τιμή σχεδίασης του LHC.
- Το LHC Υψηλής Λαμπρότητας θα παράγει τουλάχιστον 15 εκατομμύρια Higgs bosons ετησίως αντί για 3 εκατομμύρια.



CERN openlab

- **Δημόσια – Ιδιωτική Συνεργασία (~20 χρόνια)**
- **Συνεργασία μεταξύ της επιστημονικής κοινότητας και της βιομηχανίας**
- **Καινοτόμες Τεχνολογίες:** Οι καινοτόμες τεχνολογίες στο CERN περιλαμβάνουν εξελιγμένα εργαλεία οπτικοποίησης δεδομένων, αλγόριθμους μηχανικής μάθησης και προηγμένα μέτρα κυβερνοασφάλειας.
- **Επίδραση στην Έρευνα:** Οι καινοτομίες του τμήματος IT ενισχύουν σημαντικά την πειραματική ακρίβεια, οδηγώντας σε πρωτοποριακές ανακαλύψεις στη θεμελιώδη φυσική.
- **Τομείς:** HPC, Big Data Management, AI , Machine Learning
- **CERN Quantum Technology Initiative:** Εξετάζεται πώς η κβαντική υπολογιστική μπορεί να ωφελήσει τη φυσική υψηλής ενέργειας. Το CERN συνεργάζεται με παγκόσμιους ηγέτες στην κβαντική υπολογιστική για να διερευνήσει πιθανές εφαρμογές

Δέσμευση για ανοικτή επιστήμη (Open Science): Προώθηση της διαφάνειας στην έρευνα μέσω συνεργατικών έργων.

Πρακτικές κοινοποίησης Δεδομένων: Διευκολύνουν την ομαλή παγκόσμια συνεργασία και ενισχύουν την ακεραιότητα επιστημονικών ευρημάτων.

Προώθηση Παγκόσμιας συνεργασίας: Αξιοποίηση της τεχνολογίας για την ένωση ερευνητών και την ενίσχυση διασυνοριακών συνεργασιών.

Συνεργασία
και Open
Science

Επίδραση του Computing στην Κοινωνία

- Οι καινοτομίες της πληροφορικής στο CERN έχουν αλλάξει τομείς όπως η υγειονομική περίθαλψη, τα οικονομικά και η εκπαίδευση, χρησιμοποιώντας προηγμένη ανάλυση δεδομένων.
- Οι τεχνολογίες που αναπτύχθηκαν στο CERN βελτιώνουν την ιατρική απεικόνιση και την τηλεϊατρική, διευκολύνοντας τη διάγνωση και τη φροντίδα των ασθενών σε όλο τον κόσμο (π.χ. απεικόνιση PET, Θεραπεία με Πρωτόνια και Βαρέα Ιόντα, Ασφάλεια Ακτινοβολίας).

Χρήσιμοι Σύνδεσμοι

- Computing at CERN:
 - <https://home.cern/science/computing>
- The Worldwide LHC Computing Grid:
 - <https://wlcg-public.web.cern.ch/>
- The World Wide Web:
 - <https://home.cern/science/computing/birth-web>
- The Web at 30:
 - <https://web30.web.cern.ch/>
- Quantum Technologies:
 - <https://quantum.cern/>
- CERN openlab:
 - <https://openlab.cern/>





[Photo by Pietro Battistoni on Pexels](#)

Συμπεράσματα & Ερωτήσεις