



Enabling Grids for
E-science in Europe

Εισαγωγή στις Τεχνολογίες Πλέγματος *Introduction to Grid Technologies*

Αθανασία Ασίκη
aassiki@cslab.ece.ntua.gr
ΕΔΕΤ



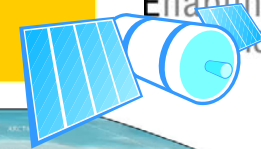
Τεχνολογίες Πλέγματος (Grids)



1. Τι είναι το Grid ?
2. Ιστορική αναδρομή
3. Κατηγορίες Grid συστημάτων
4. Βασικές αρχές του Grid
5. Δυνατότητες του Grid
6. Χρήστες του Grid
7. Αρχιτεκτονική του Grid

1. Τι είναι το Grid ?
2. Ιστορική αναδρομή
3. Κατηγορίες Grid συστημάτων
4. Βασικές αρχές του Grid
5. Δυνατότητες του Grid
6. Χρήστες του Grid
7. Αρχιτεκτονική του Grid

Πριν το Grid...



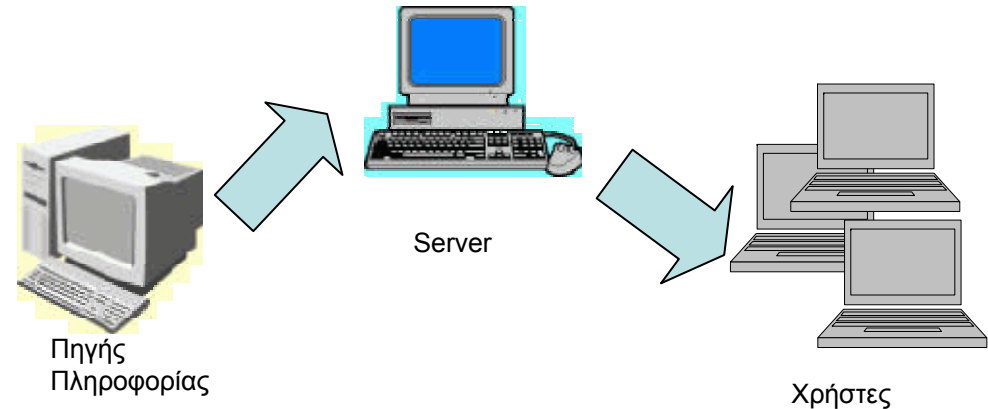
Μετά το Grid...



Grid Computing equates to the world's largest computer

Τι είναι το Grid?

- Το **World Wide Web** παρέχει πρόσβαση σε πληροφορίες αποθηκευμένες σε εκατομμύρια διαφορετικές τοποθεσίες ανά τον κόσμο.



- Κατ' αναλογία το **Grid** είναι μία αναπτυσσόμενη υποδομή που παρέχει πρόσβαση σε υπολογιστική ισχύ και αποθηκευτικό χώρο κατανεμημένα σε όλον τον κόσμο.

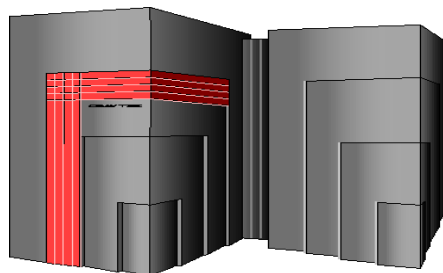


Τι είναι το Grid ?

- Συλλογή γεωγραφικά κατανεμημένων ετερογενών υπολογιστικών πόρων
“Most generalized, globalized form of distributed computing”
- **“An infrastructure that enables flexible, secure, coordinated resource sharing among dynamic collections of individuals, institutions and resources”**
Ian Foster and Carl Kesselman
- Προσφέρει πρόσβαση σε ένα ενιαίο ισχυρό εικονικό υπολογιστή
- Υποβολή από τους χρήστες μεγάλου πλήθους εργασιών χωρίς να ενδιαφέρονται που θα εκτελεστούν

- Μία οντότητα που πρόκειται να μοιραστεί

π.χ. Μονάδες Επεξεργασίας
Αποθηκευτικές Μονάδες
Λογισμικό

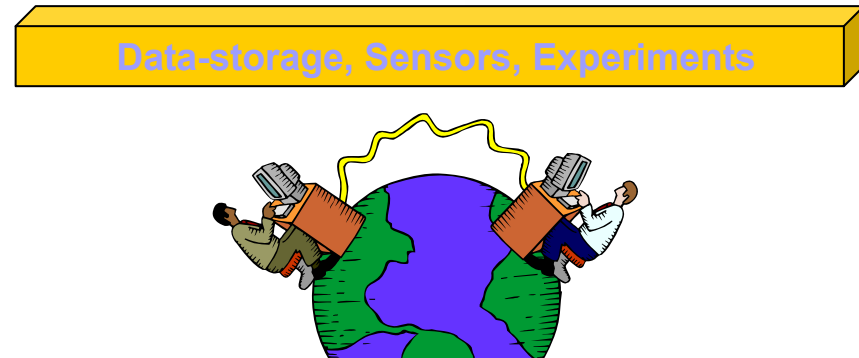
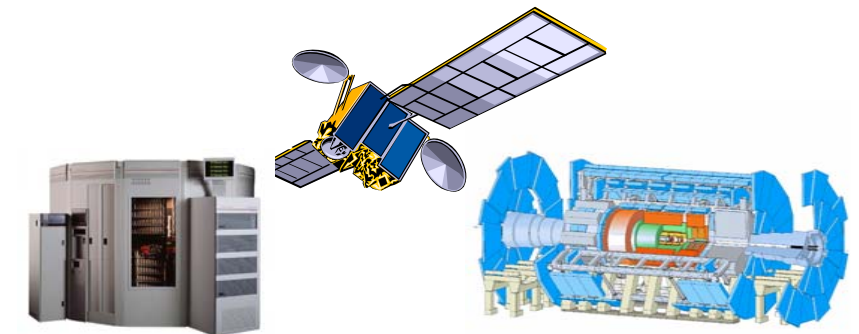
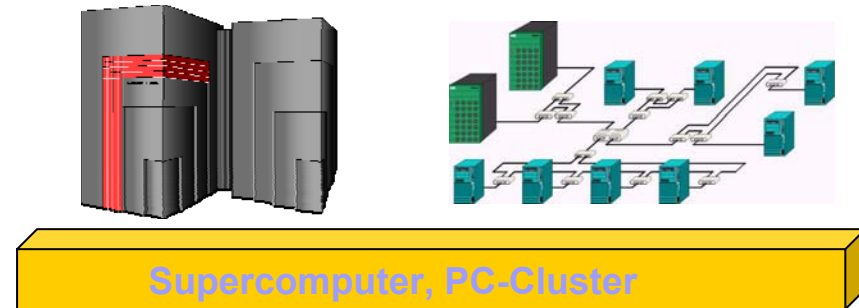
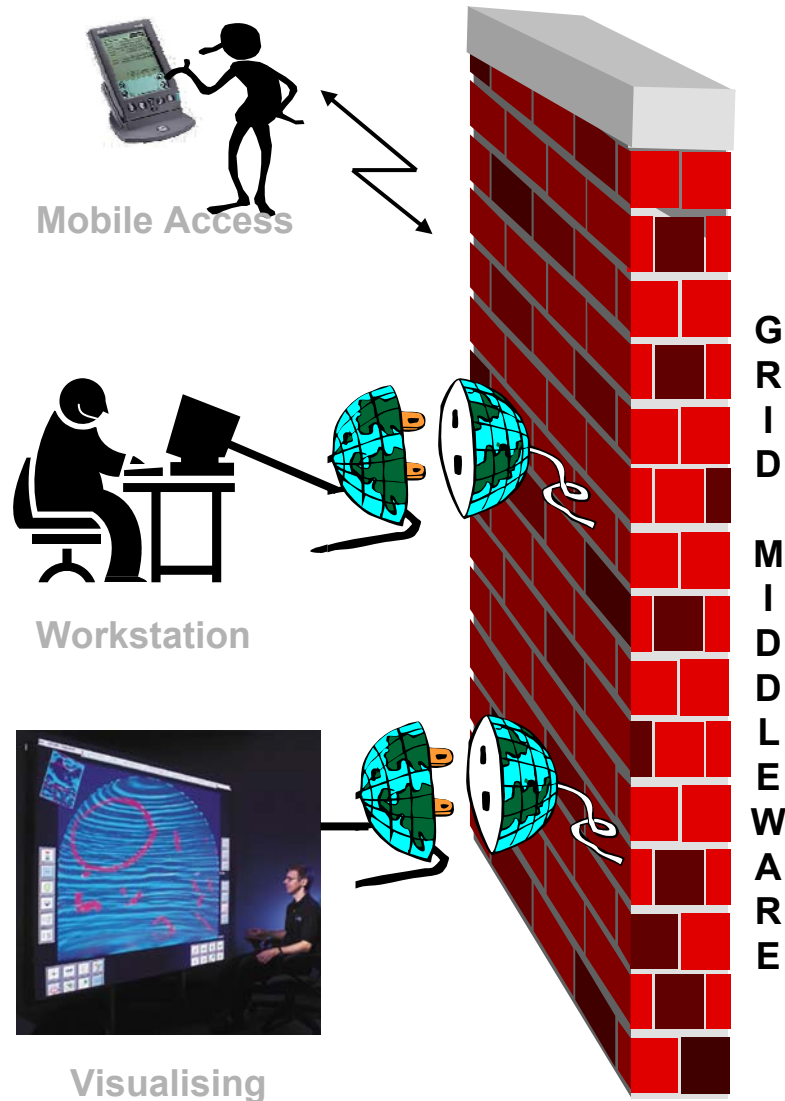


1. Τι είναι το Grid ?
2. **Ιστορική αναδρομή**
3. Βασικές αρχές του Grid
4. Κατηγορίες Grid συστημάτων
5. Δυνατότητες του Grid
6. Χρήστες του Grid
7. Αρχιτεκτονική του Grid

Grid ηλεκτρικής ενέργειας vs Grid

Grid ηλεκτρικής ενέργειας	To Grid
Χρησιμοποιείς την ηλεκτρική ενέργεια χωρίς να σε ενδιαφέρει πού και με ποιον τρόπο παράγεται. Είναι διαθέσιμη, προσιτή και αξιόπιστη.	Έχεις πρόσβαση σε υπολογιστική ισχύ και αποθηκευτικό χώρο που βρίσκονται σε όλον τον κόσμο
Τεράστια υποδομή που περιλαμβάνει ετερογενή συστήματα (σπίτια, σταθμούς παραγωγής, σταθμούς μεταφοράς, καλώδια).	Τεράστια υποδομή που περιλαμβάνει ετερογενή υπολογιστικά συστήματα και υπηρεσίες (επεξεργαστές, αποθηκευτικός χώρος, αισθητήρες).
Μπορείς να χρησιμοποιήσεις ηλεκτρικές συσκευές παντού, απλά βάζοντας αυτές στην πρίζα.	Θα μπορείς να έχεις πρόσβαση σε απομακρυσμένους υπολογιστικούς πόρους μέσω οποιασδήποτε πλατφόρμας, απλά χρησιμοποιώντας το Web.
Ζητάς ηλεκτρικό ρεύμα και σου δίνεται. Μοναδική δέσμευση ότι θα πρέπει να πληρώσεις το λογαριασμό.	Ζητάς υπολογιστικούς πόρους και πρόσβαση σε δεδομένα και σου δίνονται. Προς το παρόν είναι “δωρεάν”!

To Grid μεταφορικά ...



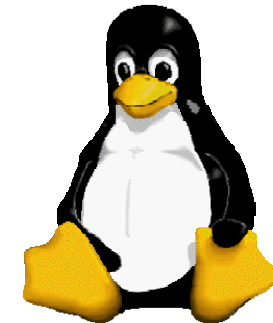
Σταθμοί προς το Grid

1990 : World Wide Web (CERN)



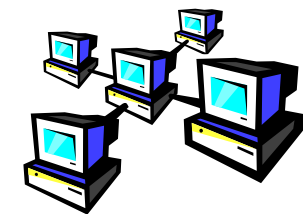
1991: Το λειτουργικό σύστημα
Linux (University of Helsinki)

✓ λειτουργικό σύστημα ανοικτού κώδικα



1994: Συστοιχίες υπολογιστών
(cluster) (NASA)

✓ χρήση καρτών Ethernet για διασύνδεση
υπολογιστών με υψηλή ταχύτητα



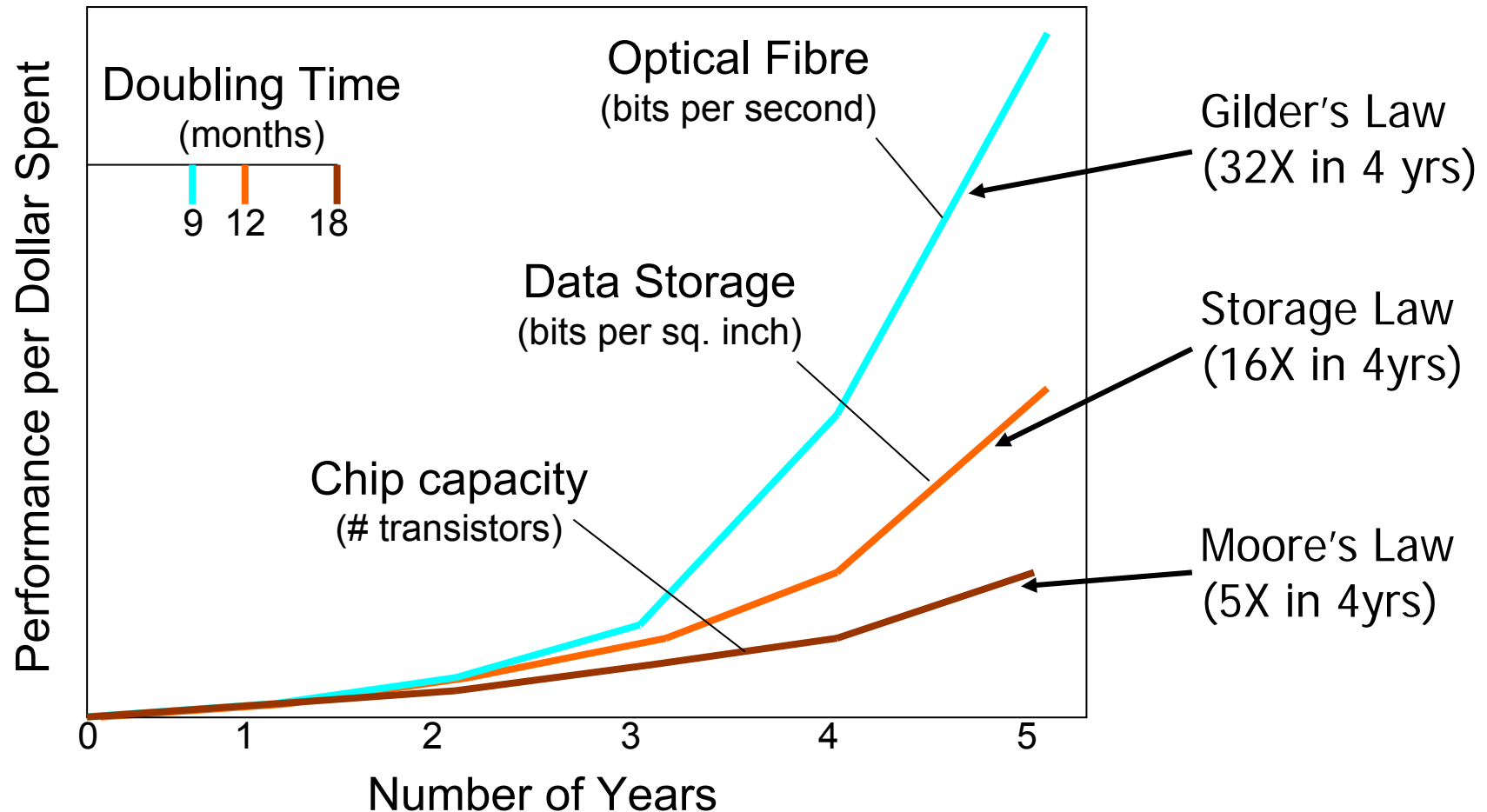
⇒ Το Grid αποτελείται από διασυνδεδεμένα cluster μέσω του Internet και στηρίζεται στο λειτουργικό σύστημα Linux

- Επιστήμη που έγινε δυνατή με το διαμοιρασμό πόρων (δεδομένων, επιστημονικών οργάνων, υπολογιστικών πόρων, εξειδικευμένων ατόμων) μέσω του Internet
 - ✓ Εφαρμογές που απαιτούν υπολογιστική ισχύ
 - ✓ Εφαρμογές που διαχειρίζονται πολλά δεδομένα (είτε παράγουν σαν αποτέλεσμα μεγάλο όγκο δεδομένων εξόδου είτε απαιτούν πρόσβαση σε συλλογές δεδομένων)
 - ✓ Καταργούνται τα όρια μεταξύ οργανισμών και διαχειριστικών τομέων

Γιατί τώρα ?

- Αύξηση της ταχύτητας των δικτύων υπολογιστών και εύκολη πρόσβαση σε αυτά
 - ✓ οπτικές ίνες
 - ✓ ασύρματες ζεύξεις
 - ✓ νέες τεχνικές που χρησιμοποιούνται στο Διαδίκτυο (ADSL, WiMax)
- Ο νόμος του Moore είναι παντού ...
 - ✓ Ραγδαία ανάπτυξη των υπολογιστικών μονάδων, των οργάνων, των αισθητήρων
 - ⇒ Πρόκληση να συνδυαστούν για το καλύτερο αποτέλεσμα!
- Οι εργασίες που εκτελούνται σε υπολογιστικά συστήματα απαιτούν μεγάλο αριθμό υπολογισμών και την επικοινωνία μεταξύ των ατόμων που τις εκτελούν

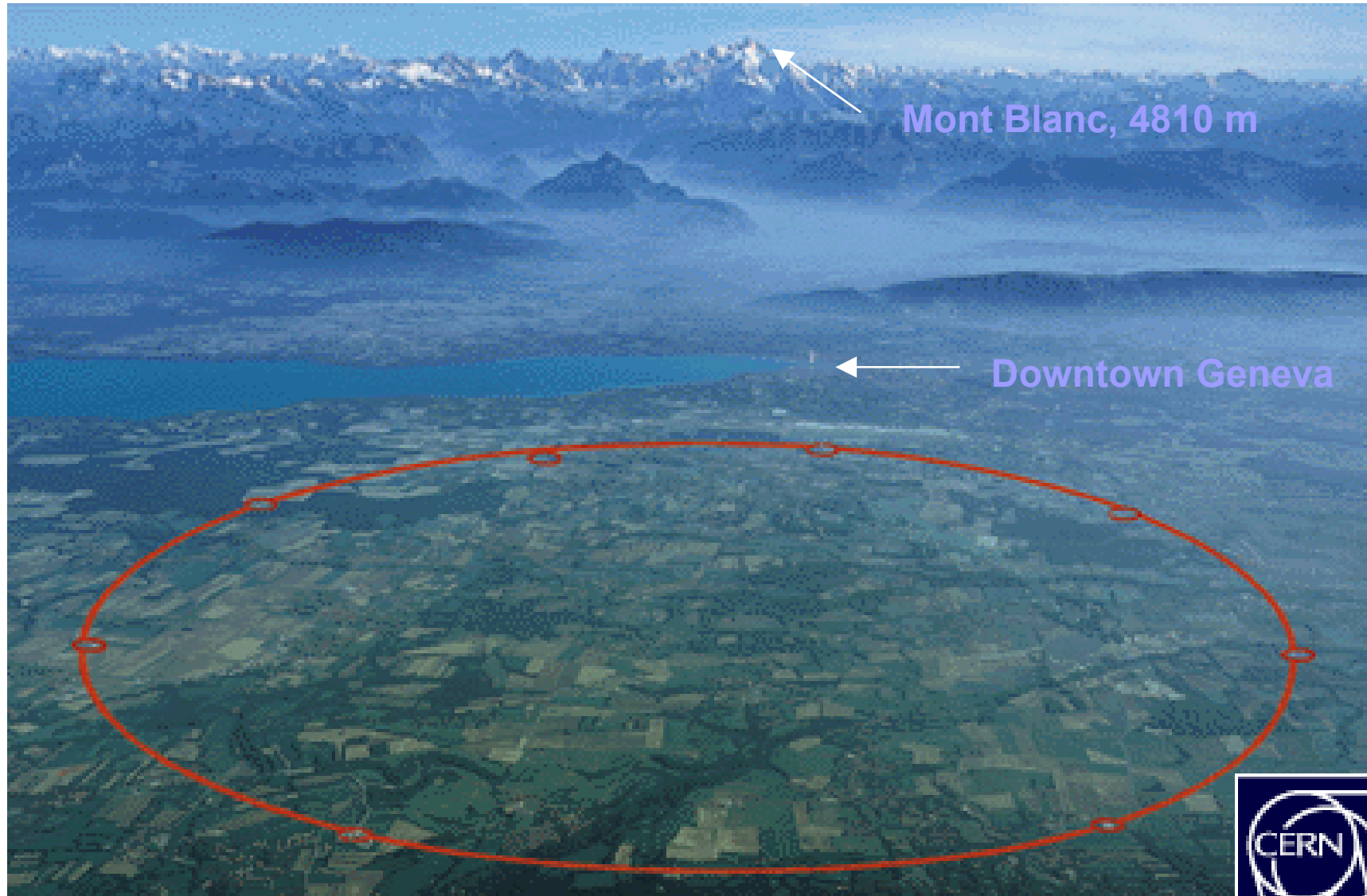
Εκθετική αύξηση



Triumph of Light – *Scientific American*. George Stix, January 2001

CERN

eGEE
Enabling Grids for
E-science in Europe



Πειραιάς, 30/03/1



LHC (*Large Hadron Collider*)

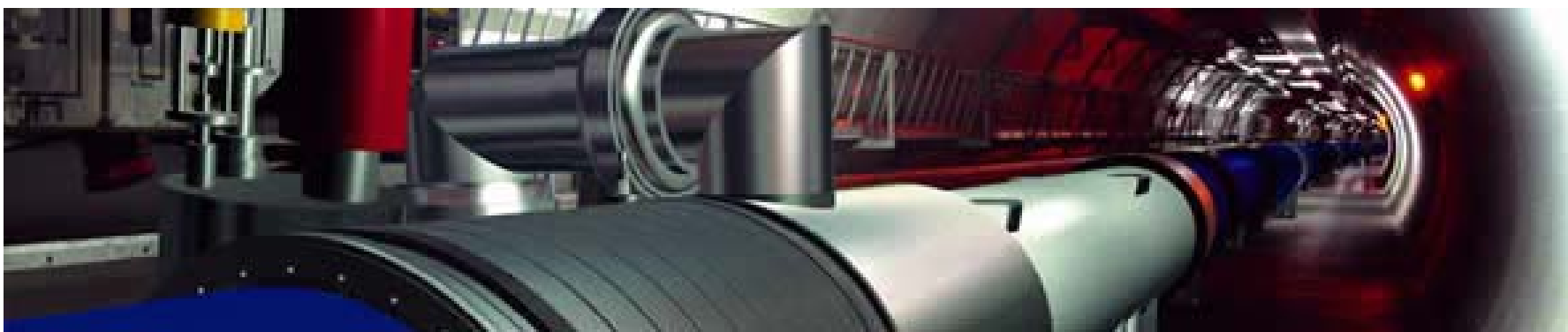
- Το LHC θα φέρει σε σύγκρουση δύο ακτίνες πρωτονίων ενέργειας 14TeV
- Εάν το Higgs μποζόνιο υπάρχει, τότε σίγουρα το LHC θα το βρει!
- Διεξαγωγή 4 πειραμάτων με ανιχνευτές

ALICE

ATLAS

CMS

LHCB

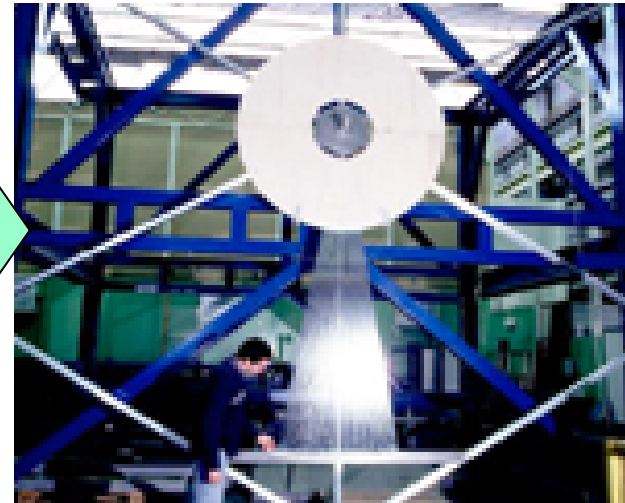


Ανιχνευτές του LHC

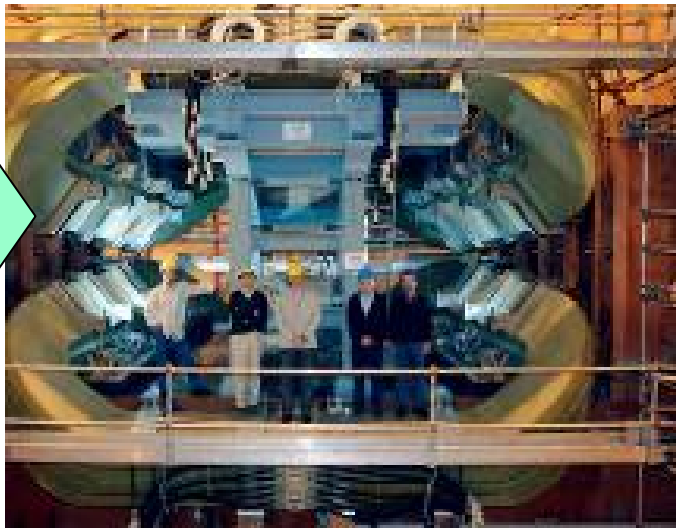
ATLAS



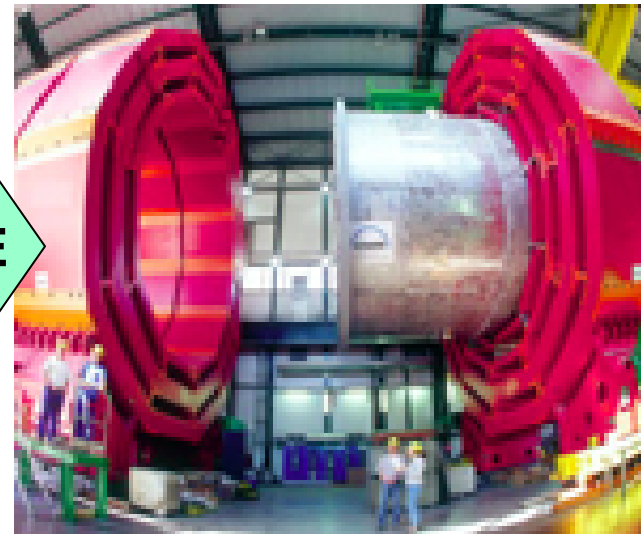
CMS



LHCb

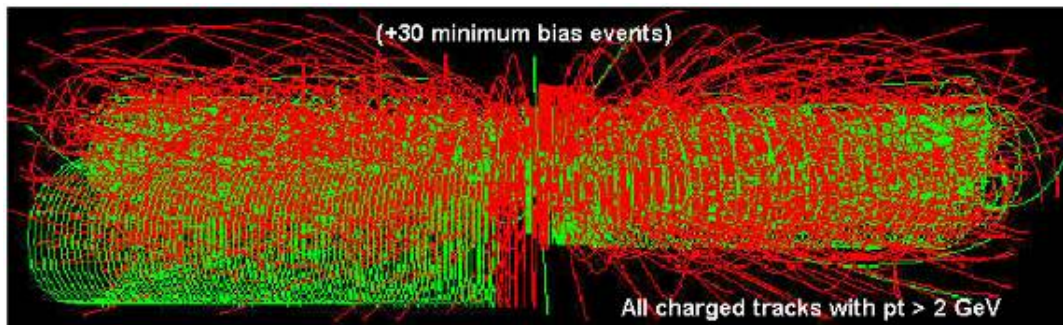


ALICE

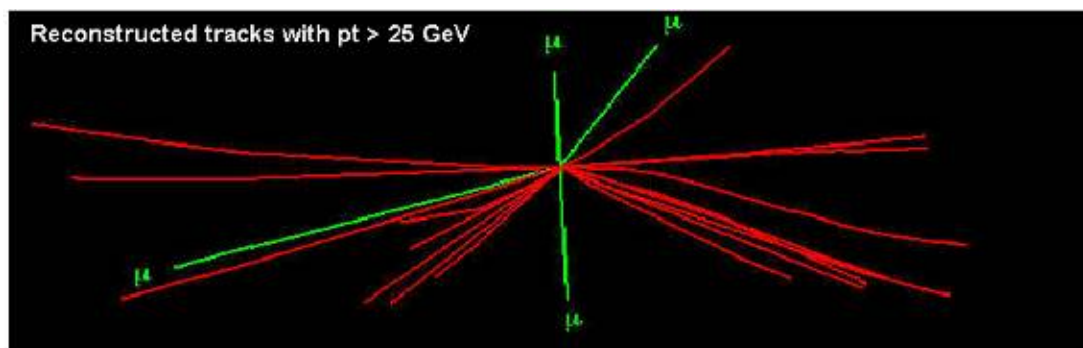


Η λειτουργία του LHC

Ξεκινώντας από αυτό το γεγονός ...



Αναζήτηση αυτού του αποτυπώματος ...



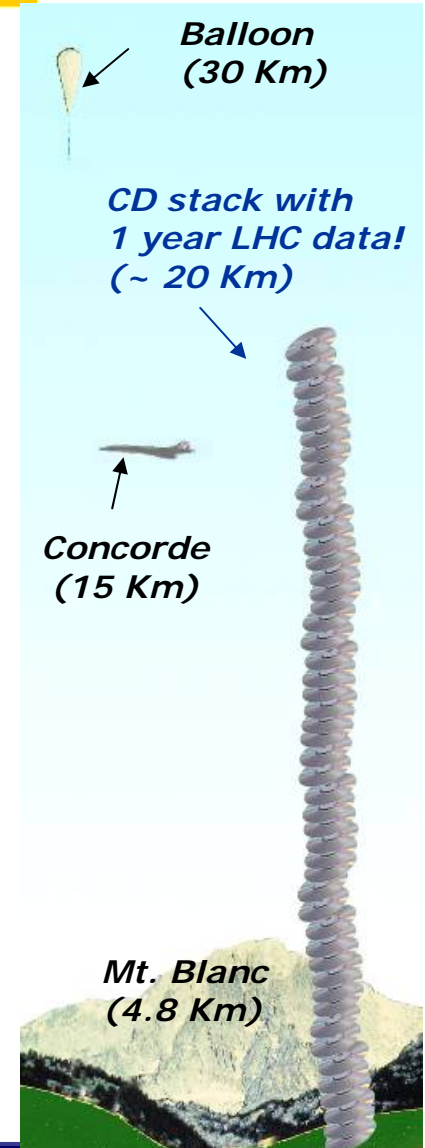
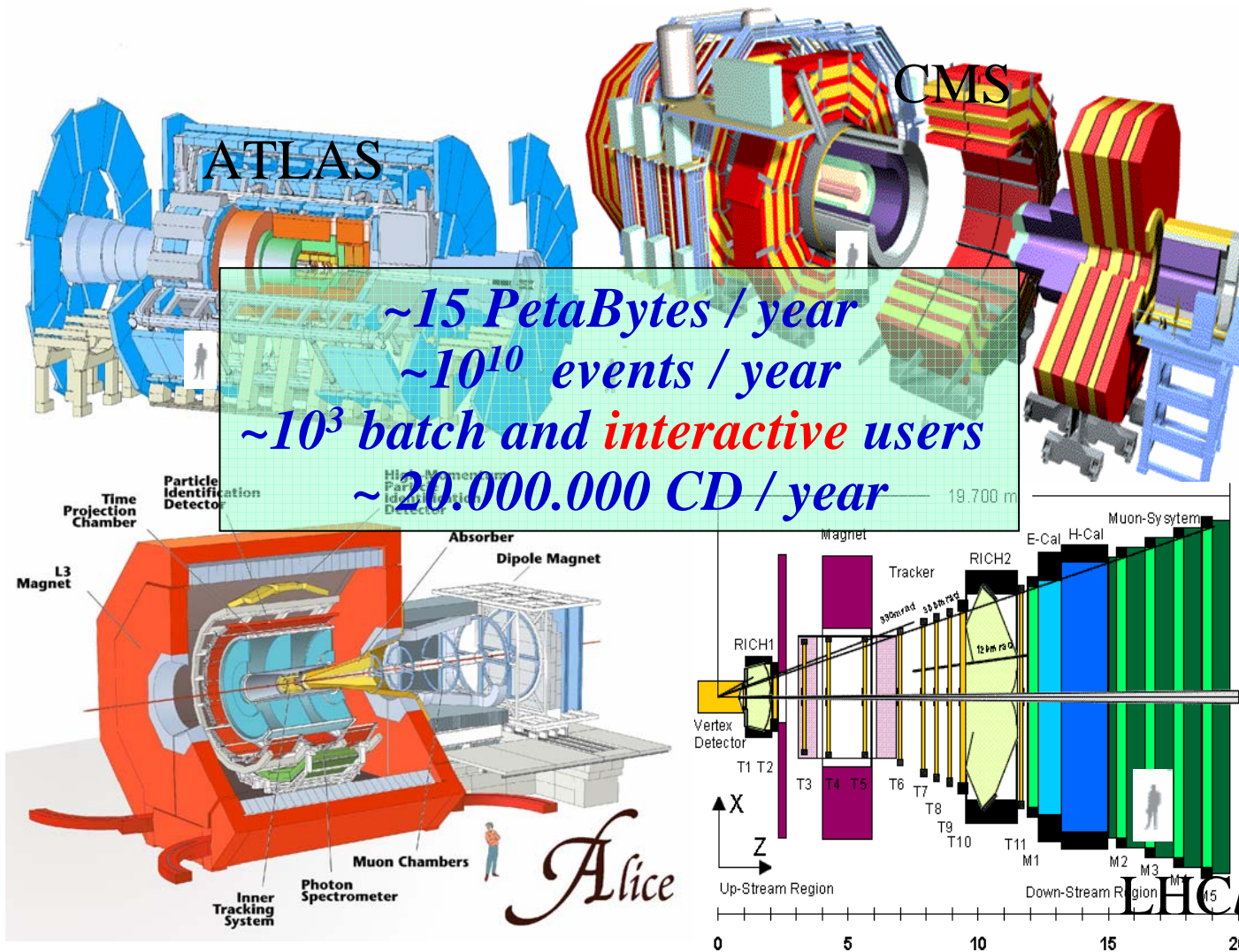
- ✓ Καταγραφή γεγονότων
- ✓ Αποθήκευση γεγονότων
- ✓ Επεξεργασία γεγονότων

Επιλεκτικότητα: 1 in 10^{13}

Σαν να αναζητούμε ένα άτομο σε χιλιάδες άτομα ανά τον κόσμο!

Η μία βελόνα σε 20 εκατομμύρια άχυρα!

Παραγωγή δεδομένων από τους Ανιχνευτές του LHC



Τι θέλουν οι επιστήμονες?

- Απεριόριστη υπολογιστική ισχύ
- Απεριόριστο αποθηκευτικό χώρο
- Αξιόπιστη πρόσβαση σε κοινούς πόρους, δεδομένα, εφαρμογές και αποτελέσματα που προέρχονται από την επεξεργασία των δεδομένων
- Δυνατότητα να συνεργάζονται με συναδέλφους απλά, αποτελεσματικά

Είδη υπολογιστικών προβλημάτων (1)

- *“Computer – centric” προβλήματα*
 - ✓ Ανάγκη υπολογιστικής ισχύος
 - ⇒ το Grid συνδυάζει μεγάλης κλίμακας υπολογιστικούς πόρους
- *“Data-centric” προβλήματα (“Data-intensive” προβλήματα)*
 - ✓ Τεράστια ποσά δεδομένων
 - ⇒ Το Grid αποτελεί μία συλλογή από γεωγραφικά κατανεμημένα repositories, βάσεις δεδομένων και ψηφιακές βιβλιοθήκες
 - συγκεντρώνει δεδομένα
 - αποθηκεύει δεδομένα
 - αναλύει δεδομένα

Είδη υπολογιστικών προβλημάτων (2)

- *“Community-centric” problems (collaborative εφαρμογές)*
 - ✓ Αλληλεπίδραση άνθρωπου προς άνθρωπο
 - ✓ Συνεργασία μεταξύ ατόμων ή κοινωνικών ομάδων
 - ✓ “Virtual shared space”
 - ⇒ κοινή χρήση αρχείων δεδομένων, εξομοιώσεων
 - ✓ Απαιτήσεις Πραγματικού Χρόνου

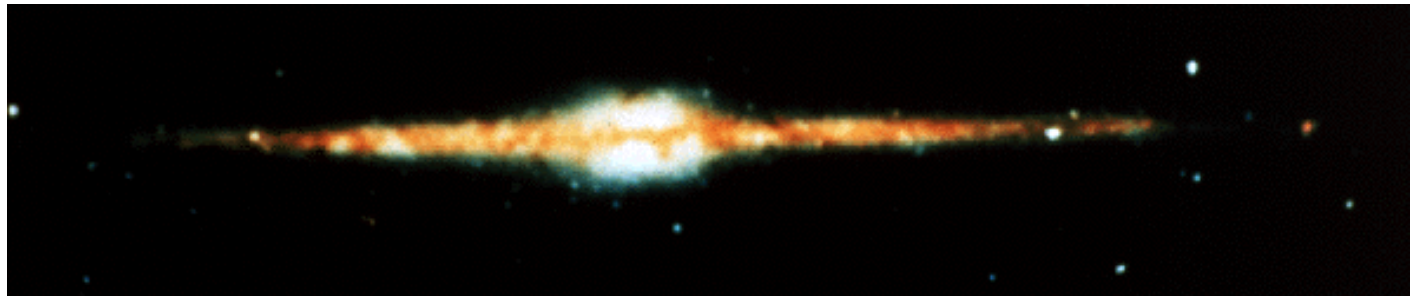
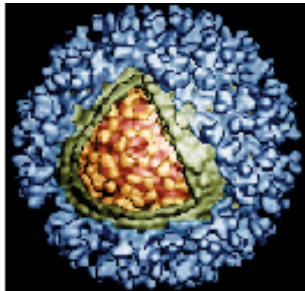
High-throughput εφαρμογές

- Χωρισμός του προβλήματος σε πολλά διαφορετικά tasks ανεξάρτητα μεταξύ τους
- ⇒ Προγραμματισμός των tasks για εκτέλεση στο Grid σε ανενεργούς υπολογιστικούς πόρους
- ✓ @home εφαρμογές



High-performance εφαρμογές

- **Supercomputing**
- “Computer – centric” προβλήματα
- Επιστημονικές εφαρμογές
 - Αστροφυσική
 - Αεροδιαστημική βιομηχανία
 - Αυτοκινητοβιομηχανία
 - Οικονομικά μοντέλα
 - Μετεωρολογικά μοντέλα
 - Κατανεμημένες Εξομοιώσεις



1. Τι είναι το Grid ?
2. Ιστορική αναδρομή
3. **Κατηγορίες Grid συστημάτων**
4. Βασικές αρχές του Grid
5. Δυνατότητες του Grid
6. Χρήστες του Grid
7. Αρχιτεκτονική του Grid

Κατηγορίες Grid Συστημάτων (1)

- **Υπολογιστικά Grids (Computational Grids)**

- ⇒ Συλλογή κατανεμημένων υπολογιστικών υποδομών οι οποίες λειτουργούν ως ενιαίος επεξεργαστής
- ⇒ Πραγματοποίηση επεξεργασίας δεδομένων με μεγάλες υπολογιστικές απαιτήσεις
 - ✓ ταχύτερα
 - ✓ αποτελεσματικότερα
 - ✓ με μικρό κόστος
 - ✓ χρησιμοποιώντας υπάρχουσες υποδομές
- ⇒ Εφαρμογές:
 - ✓ Επιστημονικός χώρος
 - ✓ Έρευνα
 - ✓ Βιομηχανία

Κατηγορίες Grid Συστημάτων (2)

- **Grids Δεδομένων (Data Grids)**

⇒ Οι χρήστες και οι εφαρμογές διαχειρίζονται πληροφορίες από βάσεις δεδομένων που βρίσκονται σε κατανεμημένες πλατφόρμες:

- ✓ εύκολα
- ✓ αποτελεσματικά

⇒ Μειωμένο κόστος γιατί δεν υπάρχει ανάγκη για μεταφορά, αντιγραφή και συγκέντρωση δεδομένων σε ένα κεντρικό σημείο

⇒ Αυξημένη αξιοπιστία κατά την πρόσβαση στα δεδομένα

Κατηγορίες Grid Συστημάτων (3)

- **Grids Υπηρεσιών (Service Grids)**

- ↳ Πραγματοποίηση επεξεργασίας Πραγματικού Χρόνου

- ↳ Προϋποθέσεις:

- ✓ η συλλογή δεδομένων από φυσικά κατανεμημένα εργαστήρια
- ✓ η ανάλυση των δεδομένων
- ✓ η διαχείριση των δεδομένων

1. Τι είναι το Grid ?
2. Ιστορική αναδρομή
3. Κατηγορίες Grid συστημάτων
4. **Βασικές αρχές του Grid**
5. Δυνατότητες του Grid
6. Χρήστες του Grid
7. Αρχιτεκτονική του Grid

Βασικές αρχές του Grid (1)

- **Διαμοιρασμός των υπολογιστικών πόρων**



Χρήστης που αποκτάει πρόσβαση στο Grid

⇒ χρησιμοποιεί απομακρυσμένους πόρους, που θα του επιτρέψουν να εκτελέσει εργασίες που δεν έχει τη δυνατότητα να εκτελέσει σε ένα μεμονωμένο υπολογιστή ή σε μία συστοιχία υπολογιστών

⇒ αποκτάει πρόσβαση σε απομακρυσμένο λογισμικό, υπολογιστικούς πόρους, δεδομένα, απομακρυσμένους αισθητήρες, τηλεσκόπια, επιστημονικά όργανα που ανήκουν σε άλλα ερευνητικά κέντρα



Προβλήματα:

- ✓ Οι προσφερόμενοι πόροι ανήκουν σε διαφορετικούς ανθρώπους και η χρήση τους υπόκειται σε διαφορετικές πολιτικές και περιορισμούς
- ✓ Ετερογενείς πόροι (διαφορετικά λογισμικά, διαφορετικοί κανόνες πρόσβασης και ασφάλειας)

Βασικές αρχές του Grid (2)

- **Ασφαλής πρόσβαση**

- ↳ **Πολιτική Πρόσβασης (Access policy)**

Σαφής καθορισμός της πρόσβασης των προμηθευτών των πόρων και των χρηστών και υπό ποιες προϋποθέσεις

- ↳ **Ταυτοποίηση (Authentication)**

Μηχανισμός που εξακριβώνει ποια είναι η ταυτότητα των χρηστών και των πόρων

- ↳ **Εξουσιοδότηση (Authorization)**

Μηχανισμός που καθορίζει τις εργασίες που επιτρέπονται με βάση τους κανόνες που ισχύουν

- **Προβλήματα:**

- Ασφάλειας στους υπολογιστές των Grid υποδομών
- Στον τρόπο μεταφοράς των δεδομένων μέσω του διαδικτύου

Βασικές αρχές του Grid (3)

- **Αποτελεσματική χρήση των πόρων**
- Αύξηση του αριθμού των χρηστών
 - ⇒ Ανεπάρκεια των υπολογιστικών πόρων
 - ⇒ Δημιουργία ουρών αναμονής πριν την εκτέλεση μίας εργασίας
- ✚ Ανάπτυξη αλγορίθμων για την βέλτιστη ανάθεση των εργασιών στους πόρους που διατίθενται
- ✚ Βέλτιστη ανάθεση:
 - ✓ Αριθμός εργασιών σε μία ουρά αναμονής
 - ✓ Υπολογιζόμενος χρόνος για την εκτέλεση των εργασιών που προηγούνται
 - ✓ Υπολογιστική ισχύς των πόρων

Βασικές αρχές του Grid (4)

- **Εξάλειψη της γεωγραφικής απόστασης**
 - ⇒ Απαιτούνται δίκτυα υψηλών ταχυτήτων για την διασύνδεση των πόρων (οπτικές ίνες, ανάπτυξη των διαδικτυακών τεχνολογιών)
 - ⇒ Χρόνος μεταφοράς των δεδομένων μεγάλος
 - ⇒ Μη αποτελεσματική η επεξεργασία δεδομένων σε ισχυρότερους αλλά πολύ μακρινούς υπολογιστικούς πόρους
 - ⇒ Μικρή καθυστέρηση στην επικοινωνία των μονάδων έτσι ώστε η συνεργασία να γίνεται σε πραγματικό χρόνο

Βασικές αρχές του Grid (5)

- **Δημιουργία κοινών προτύπων (Open Standards)**

- ✓ Μηχανικοί που αναπτύσσουν τις τεχνολογίες Grid
- ✓ Εταιρείες που αναπτύσσουν τεχνολογίες διαδικτύου
- ⇒ Μία εργασία που εκτελείται σε μία υποδομή Grid θα εκτελείται με τον ίδιο τρόπο και σε μία άλλη αν υπάρχουν κοινά πρότυπα

- ⇒ **Global Grid Forum**

- Καθιέρωση κοινών προτύπων (π.χ. OGSA)
- Σχεδιασμός μελλοντικών εφαρμογών και υπηρεσιών του Grid

- ⇒ **Globus Toolkit**

- Αναπτύσσεται από το Globus Alliance
- Λογισμικά εργαλεία για τη δημιουργία υπολογιστικών πλεγμάτων

1. Τι είναι το Grid ?
2. Ιστορική αναδρομή
3. Κατηγορίες Grid συστημάτων
4. Βασικές αρχές του Grid
5. **Δυνατότητες του Grid**
6. Χρήστες του Grid
7. Αρχιτεκτονική του Grid

Δυνατότητες του Grid (1)

- **Καλύτερη εκμετάλλευση πόρων**
- ⇒ Το μηχάνημα του χρήστη μπορεί να χρησιμοποιεί το μεγαλύτερο μέρος της ισχύς του
- ⇒ Χρήση ενός grid για την εκτέλεση μίας εφαρμογής (*Grid application*) σε ένα απομακρυσμένο μηχάνημα
- ⇒ Εκτέλεση της εφαρμογής σε κάποιο άλλο αδρανές (*idle*) μηχάνημα στο Grid χωρίς απόκτηση καινούργιων πόρων
- ⇒ Ισοροπία στο φορτίο των υπάρχοντων πόρων (*load balancing*)
- ⇒ Προϋποθέσεις :
 - ✓ Σχεδιασμός εφαρμογής με βάση τους μηχανισμούς του Grid
 - ✓ Το απομακρυσμένο μηχάνημα να διαθέτει κατάλληλο υλικό και λογισμικό για την εκτέλεση της εφαρμογής

Δυνατότητες του Grid (2)

- **Παράλληλη υπολογιστική επεξεργασία (Parallel CPU Capacity)**
 - ⇒ Εφαρμογές υψηλών απαιτήσεων σε υπολογιστική ισχύ
 - ⇒ Χωρισμός αλγορίθμων σε τμήματα που μπορούν να εκτελεστούν παράλληλα

Προβλήματα

- ⇒ Δύσκολο να χωριστεί μια εφαρμογή σε τελείως ανεξάρτητα κομμάτια
- ⇒ Συγκρούσεις όταν υπάρχουν κοινά αρχεία ή κοινές βάσεις δεδομένων για ανάγνωση και αποθήκευση δεδομένων

Δυνατότητες του Grid (2)

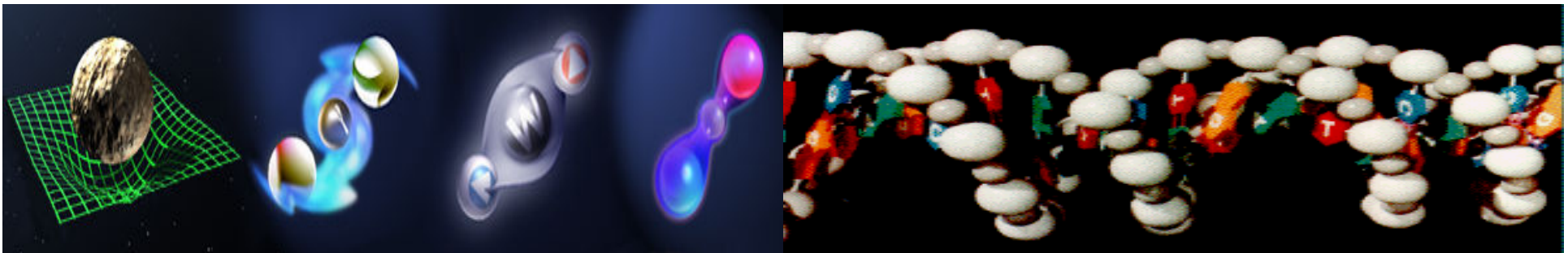
- **Παράλληλη υπολογιστική επεξεργασία (Parallel CPU Capacity)**

↗ Τεχνικά προβλήματα στην επικοινωνία παράλληλων εργασιών:

- ✓ η περιορισμένη χωρητικότητα δικτύου
- ✓ τα πρωτόκολλα συγχρονισμού
- ✓ το εύρος ζώνης προς συσκευές αποθήκευσης

↗ Πολλές επιστημονικές εφαρμογές όπως:

- ✓ Σωματιδιακή Φυσική
- ✓ Βιοιατρική
- ✓ Οικονομικά Μοντέλα
- ✓ Animation
- ✓ Επεξεργασία video
- ✓ Μετεωρολογικά Μοντέλα



Δυνατότητες του Grid (3)

- **Συνεργασία μέσω εικονικών οργανισμών**
- “A set of individuals and / or institutions defined by highly controlled sharing rules, with resource providers and consumers defining clearly and carefully just what is shared, who is allowed to share and the conditions under which sharing occurs”
Ian Foster
- Μεμονωμένα άτομα ή οργανισμοί που μοιράζονται υπολογιστικούς πόρους με έναν ευέλικτο, ασφαλή και οργανωμένο τρόπο
- ↳ Άλλα κοινά “αγαθά” μεταξύ των εικονικών οργανισμών
 - ✓ πόρους
 - ✓ λογισμικό
 - ✓ ειδικό εξοπλισμό
 - ✓ άδειες
 - ✓ υπηρεσίες (services)
 - ✓ εύρος ζώνης σύνδεσης στο internet

Εικονικοί Οργανισμοί (VOs)

- Ενταγμένοι στο πείραμα LHC
 - ALICE ⇒ Πείραμα ALICE
 - ATLAS ⇒ Πείραμα ATLAS
 - CMS ⇒ Πείραμα CMS
 - DTEAM ⇒ Grid (LCG) Deployment Group
 - LHCB ⇒ Πείραμα LHCB
 - SixTrack ⇒ Single Particle Tracking Code

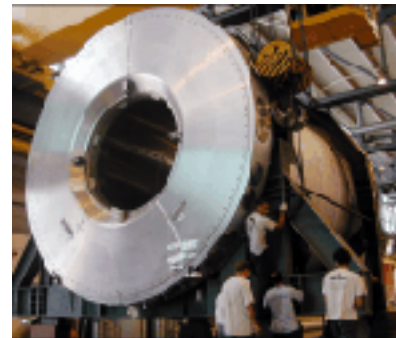
CMS



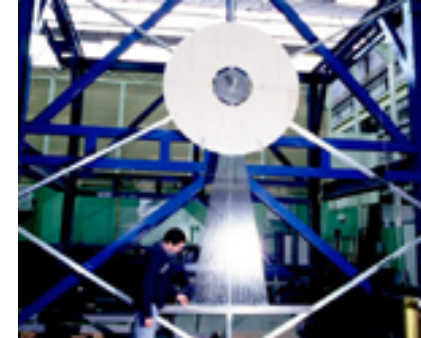
LHCb



ATLAS



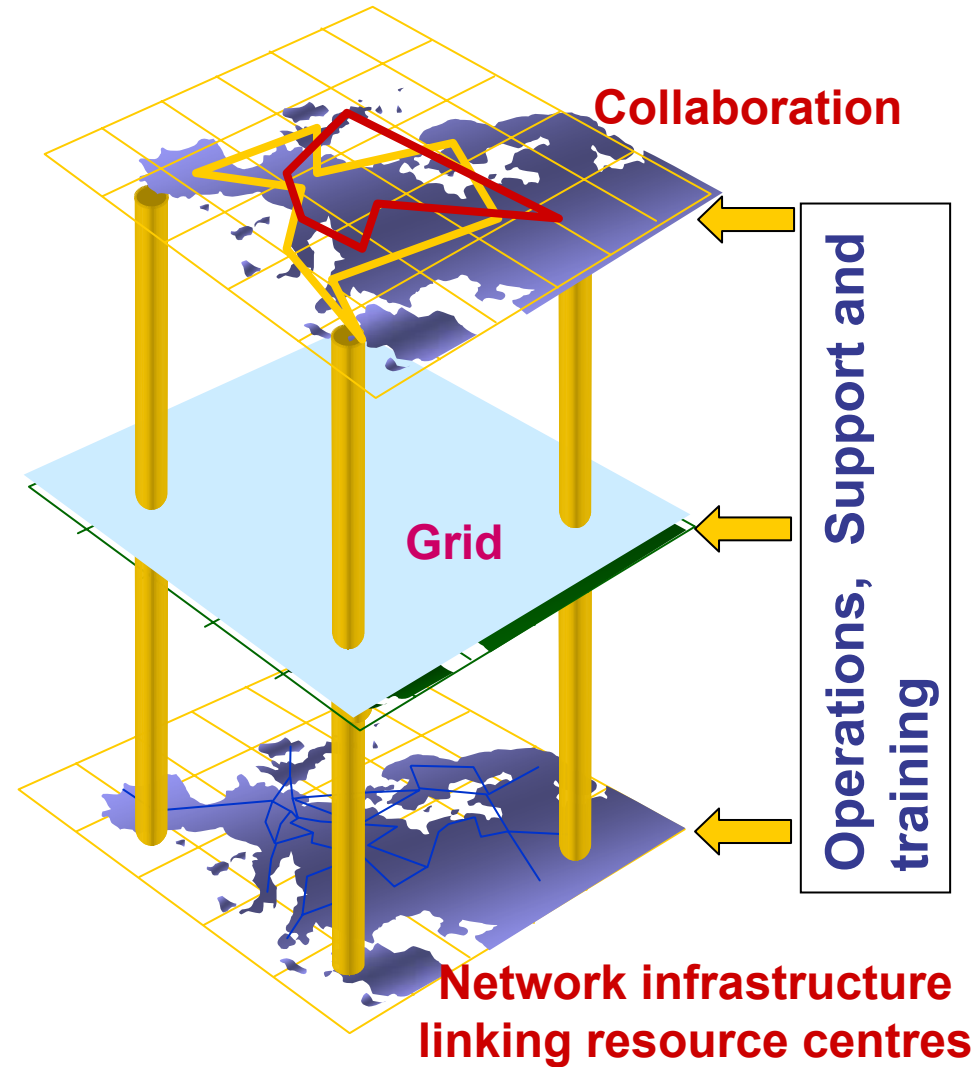
ALICE



Εικονικοί Οργανισμοί (VOs)

- Μη ενταγμένοι στο πείραμα LHC
 - Babar ⇒ Πείραμα Babar
 - D0 ⇒ Πείραμα D0
 - H1 ⇒ Πείραμα H1
 - Zeus ⇒ Πείραμα Zeus
 - ILC ⇒ ILC Community
 - Biomed ⇒ EGEE Biomedical Activity
 - ESR ⇒ Earth Science Research
 - EGEODE ⇒ Expanding GEOsciences on Demand
 - PhenoGrid ⇒ Particle Physics Phenomenology
 - CompChem ⇒ Computational Chemistry
 - SEE-VO ⇒ South Eastern Europe VO
- Περισσότερα στοιχεία για τους εικονικούς οργανισμούς σε επόμενη διάλεξη

To Grid



1. Τι είναι το Grid ?
2. Ιστορική αναδρομή
3. Κατηγορίες Grid συστημάτων
4. Βασικές αρχές του Grid
5. Δυνατότητες του Grid
6. **Χρήστες του Grid**
7. Αρχιτεκτονική του Grid

↳ **Χρήστες εφαρμογών (End Users):**

Χρησιμοποίηση των εφαρμογών

Αξιοποίηση της υπολογιστικής ισχύς και του αποθηκευτικού χώρου που τους παρέχεται

↳ **Σχεδιαστές εφαρμογών (Application Developers):**

Σχεδίαση και ανάπτυξη εφαρμογών που μπορούν να εκτελεστούν σε Grid υποδομές

- ✓ Σχεδιαστές εφαρμογών των εικονικών οργανισμών
 - ⇒ ανάπτυξη εφαρμογών για τις προσομοιώσεις πειραμάτων

✚ Διαχειριστές συστημάτων Grid (**Grid Administrators**):

Διαχείριση των επιμέρους Grid υποδομών και εξασφάλιση της σωστής λειτουργίας τους.

(διαχειριστές δικτύου, διαχειριστές cluster, διαχειριστές cluster ευρύτερης γεωγραφικής περιοχής, διαχειριστές των προγραμμάτων που χρησιμοποιούνται από τους εικονικούς οργανισμούς)

✚ Σχεδιαστές συστημάτων Grid (**Grid Developers**):

Ανάπτυξη, σχεδιασμός και εφαρμογή των νέων υπηρεσιών

✚ Σχεδιαστές εργαλείων (**Tool Developers**):

Ανάπτυξη εργαλείων, compilers, βιβλιοθηκών

⇒ Παροχή αποδοτικών εφαρμογών των προγραμματιστικών μοντέλων που θα χρησιμοποιηθούν από τους προγραμματιστές εφαρμογών

Αξιοποίηση του Grid (1)

- **Η κυβέρνηση και Διεθνείς οργανισμοί**

- ↗ Χρήση των ισχυρότερων υπολογιστικών πόρων σε καταστάσεις έκτακτης ανάγκης
- ↗ Έμμεση αντίδραση σε καταστροφές (πλημμύρες, πυρκαγιές)
- ↗ Ανταλλαγή αρχείων δεδομένων μεταξύ ιδρυμάτων και οργανισμών πιο απλά και πιο αποτελεσματικά
- ↗ Ανάπτυξη οικονομικών μοντέλων



- **Εκπαίδευση**

- ↗ Οι ηλεκτρονικές βιβλιοθήκες μπορούν να επωφεληθούν από Grid-based εργαλεία για πρόσβαση σε κατανεμημένα δεδομένα
- ↗ Η ηλεκτρονική μάθηση μπορεί να επιτευχθεί με τη δημιουργία εικονικών τάξεων με μαθητές, πόρους και εκπαιδευτές που βρίσκονται σε διαφορετικά μέρη

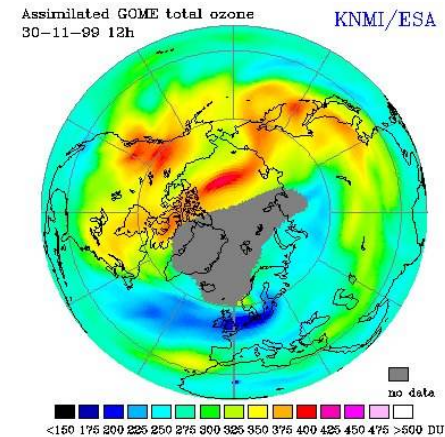


Αξιοποίηση του Grid (2)

- **Η Επιστημονική κοινότητα και η Τεχνολογία**
 - ⇒ Εξομοίωση των εφαρμογών σε πραγματικό χρόνο
 - ⇒ Εκτέλεση διάφορων υπολογισμών
 - ⇒ Σύνδεση απομακρυσμένων μηχανημάτων σε φάρμες υπολογιστών έτσι ώστε να μπορούν να παρακολουθούν και να διαχειρίζονται απομακρυσμένα όργανα και αισθητήρες
 - ⇒ Απαιτούνται:
 - ✓ Πολλές μονάδες επεξεργασίας
 - ✓ Αρκετός αποθηκευτικός χώρος
 - ✓ Πρόσβαση σε κοινές βάσεις δεδομένων.
 - ⇒ Φυσική Υψηλών Ενεργειών, Αστρονομία, Μετεωρολογία, Βιοϊατρικές Επιστήμες

Αξιοποίηση του Grid (3)

- **Το περιβάλλον**
 - ⇒ Εξομίωση διάφορων περιβαλλοντολογικών προβλημάτων, όπως
 - ✓ Τρύπα του όζοντος
 - ✓ Φαινόμενο του θερμοκηπίου
 - ✓ Μόλυνση του νερού και του αέρα
 - Μοντελοποίηση και πρόβλεψη σεισμών
 - Πρόβλεψη καιρού, τυφώνων
- **Ιδιωτικός τομέας**
 - ⇒ Παρέχουν λογισμικό εφαρμογών και υπηρεσίες ειδικού ενδιαφέροντος
 - ⇒ Χρήστες των τεχνολογιών Grid για τις δραστηριότητες τους
- ✓ Η γεωγραφική διασπορά των τμημάτων των εταιρειών
 - ⇒ ανάπτυξη τοπικών Grids (**intra – grids**) κατά αναλογία με τα intranets



Αναπτυξιακά έργα Grid ανά τον Κόσμο

Access Grid
DISCOM
DOE Science Grid
Condor
ESG (Earth System Grid)
Fusion Collaboratory
Globus
GrADSoft (Grid Application
Development Software)
Grid Canada
GRIDS (Grid Research
Integration Development &
Support Center)
GriPhyN (Grid Physics
Network)
iVDGL (International Virtual
Data Grid Laboratory)
Music Grid
NASA Information Power Grid
NCSA Alliance Access Grid

AstroGrid
AVO (Astrophysical Virtual
Observatory)
Comb-e-chem
CrossGrid
DAME (Distributed Aircraft
Maintenance Environment)
DAMIEN (Distributed Applications and
Middleware for Industrial Networks)
DataTAG
Discovery Net
DutchGrid
EDG (European DataGrid)
EGSO (European Grid of Solar
Observations)
GEODISE (Grid Enabled Optimisation
& Design Search for Engineering)

GRIA (Grid Resources for
Industrial Applications)
Grid-Ireland
GridLab (Grid Application
Toolkit and Testbed)
GridPP
LCG (LHC Computing Grid)
MyGrid
NGIL (National Grid for
Learning Scotland)
Nordugrid (Nordic Testbed for Wide
Area Computing and Data Handling)
PIONIER Grid
Reality Grid
ScotGrid

ApGrid
ApBioNet
Grid Forum Korea
PRAGMA (Rim Applications and Grid Middleware Assembly)
Grid Datafarm for Petascale Data Intensive Computing
Gridbus Project

- ✚ Αναπτυξιακό έργο της Ευρωπαϊκής Ένωσης με σκοπό την δημιουργία πλέγματος (Grid) στην Ευρώπη
- ✚ Η υποδομή αυτή θα είναι διαθέσιμη στην επιστημονική κοινότητα 24 ώρες την ημέρα και 7 ημέρες την εβδομάδα
- **Αυτό το έργο αρχικά επικεντρώνεται σε τρεις βασικούς τομείς:**
 - ✓ Ο πρώτος τομέας είναι η ανάπτυξη υποδομής πλέγματος που παρέχει όλες τις απαραίτητες υπηρεσίες για την ενσωμάτωση νέων πόρων και η υποστήριξη των χρηστών Φυσικής Υψηλών Ενεργειών και Βιο-Πληροφορικής
 - ✓ Ο δεύτερος τομέας είναι η συνεχής βελτίωση και διατήρηση του μερισμικού έτσι ώστε να παρέχονται αξιόπιστες υπηρεσίες στους χρήστες
 - ✓ Ο τρίτος τομέας είναι η προσέλκυση καινούργιων χρηστών τόσο από τη βιομηχανία όσο και από τον επιστημονικό χώρο και εξασφάλιση ότι οι χρήστες αυτοί λαμβάνουν τη καλύτερη δυνατή εκπαίδευση και υποστήριξη που χρειάζονται

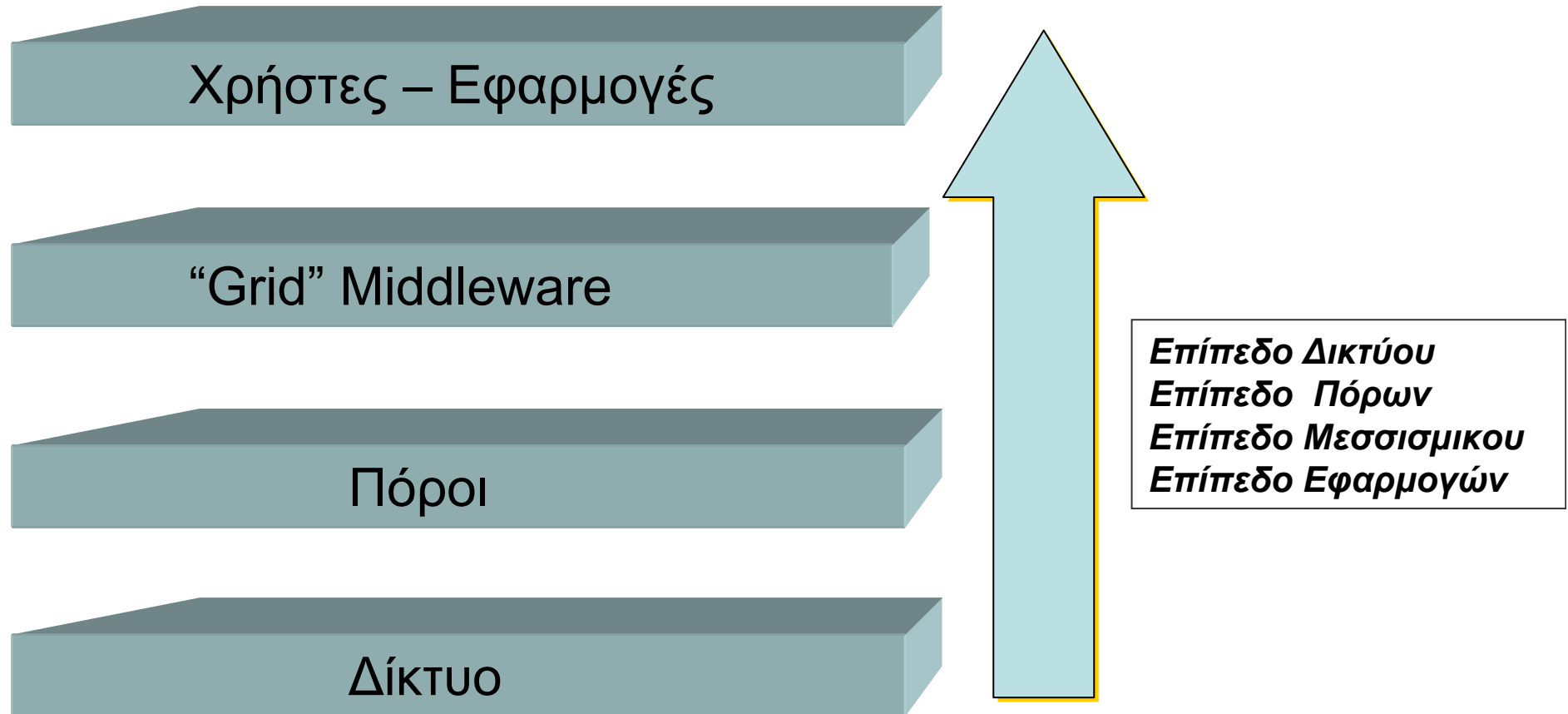
LHC Computing Grid

- Σκοπός
 - Το LCG project στοχεύει στην εγκατάσταση και την εξασφάλιση της λειτουργίας ενός Grid για τη συγκέντρωση και την ανάλυση των δεδομένων που θα προκύψουν από τους ανιχνευτές του LHC.
- ⇒ Η επεξεργασία και η ανάλυση των δεδομένων αυτών θα απαιτήσει τεράστια υπολογιστική ισχύ και τεράστιο αποθηκευτικό χώρο
- Προετοιμασία κοινής υποδομής:
 - ✓ βιβλιοθηκών
 - ✓ των εργαλείων
 - ✓ των frameworksπου απαιτούνται για να υποστηρίξουν τα προγράμματα εφαρμογών φυσικής

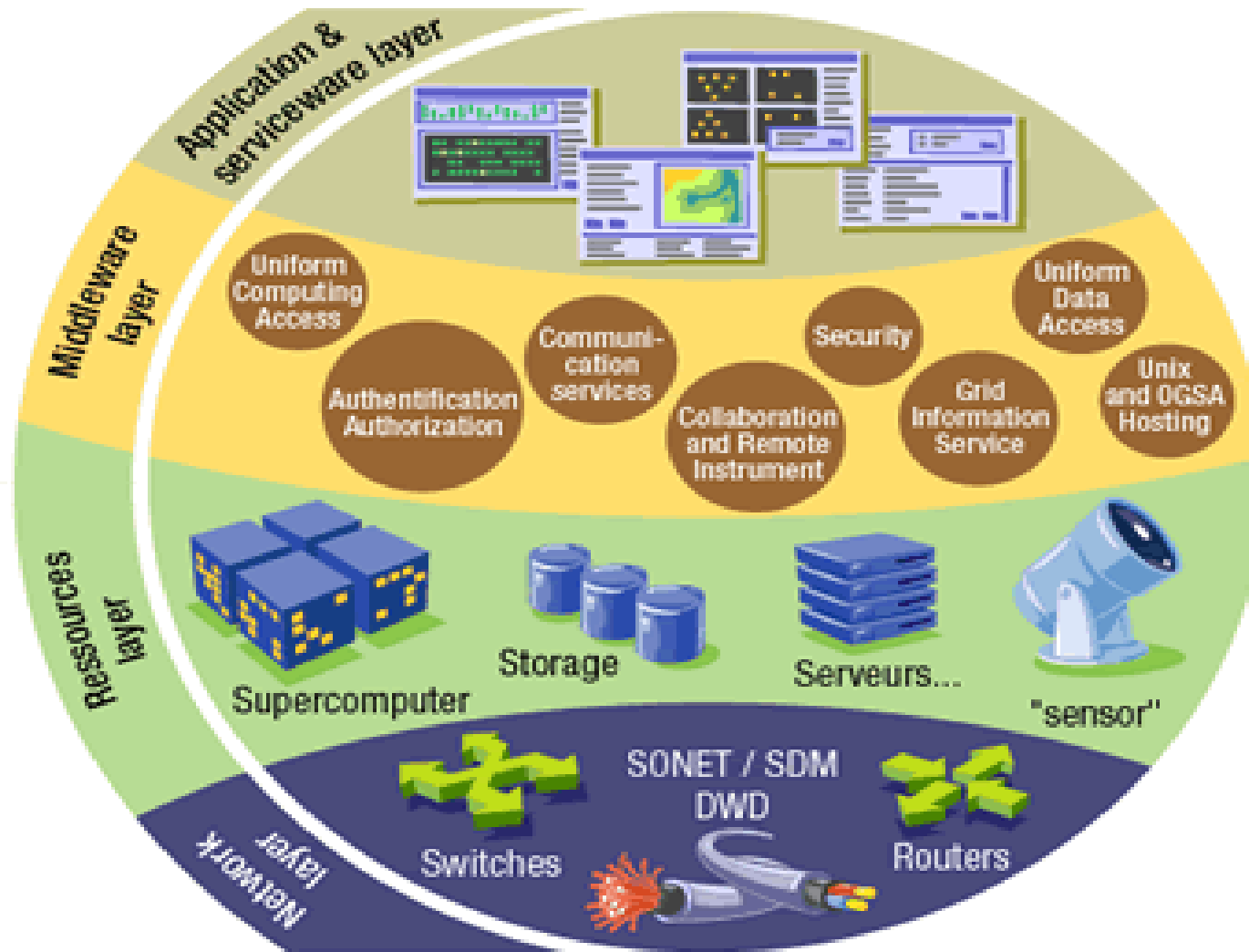


1. Τι είναι το Grid ?
2. Ιστορική αναδρομή
3. Κατηγορίες Grid συστημάτων
4. Βασικές αρχές του Grid
5. Δυνατότητες του Grid
6. Χρήστες του Grid
7. **Αρχιτεκτονική του Grid**

Αρχιτεκτονική του Grid (1)

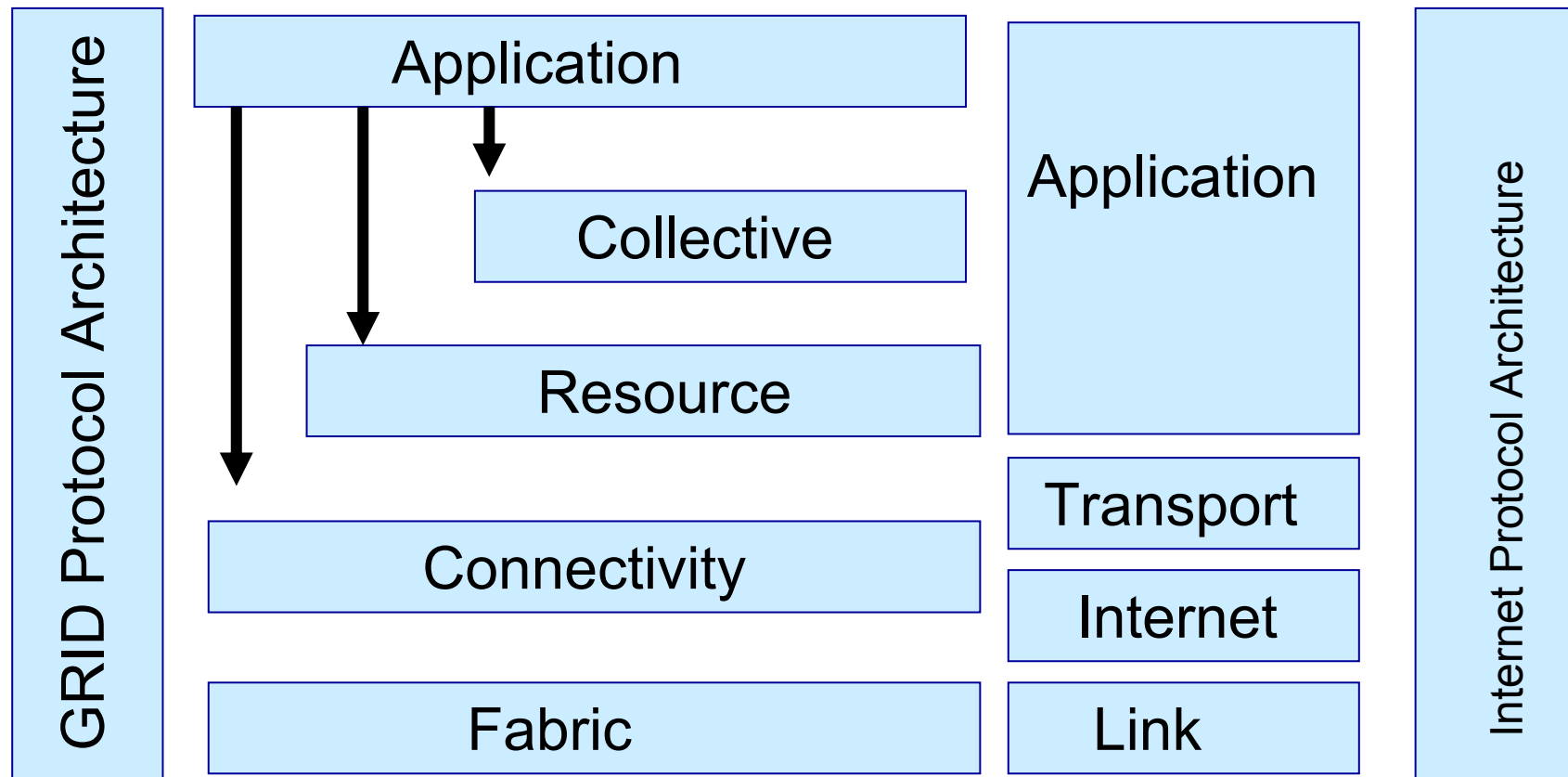


Αρχιτεκτονική του Grid (2)



Αρχιτεκτονική του Grid (3)

- **A layered grid architecture and its relationship to the Internet protocol architecture**
(Foster, Kesselman, & Tuecke)



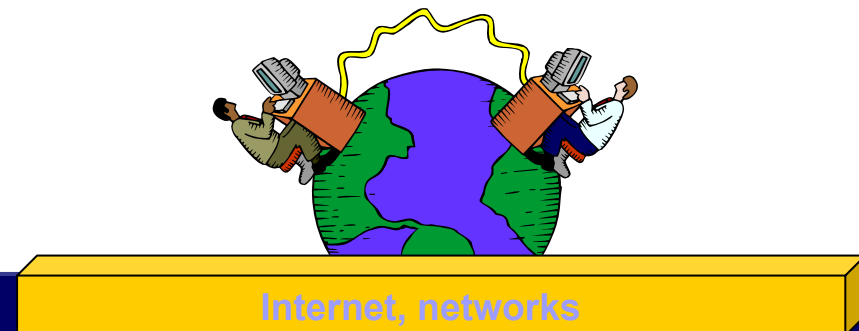
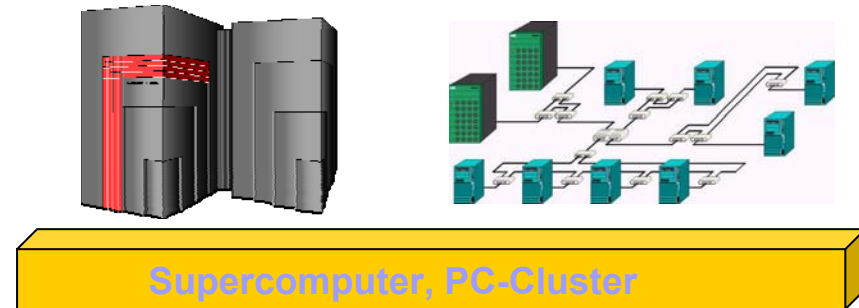
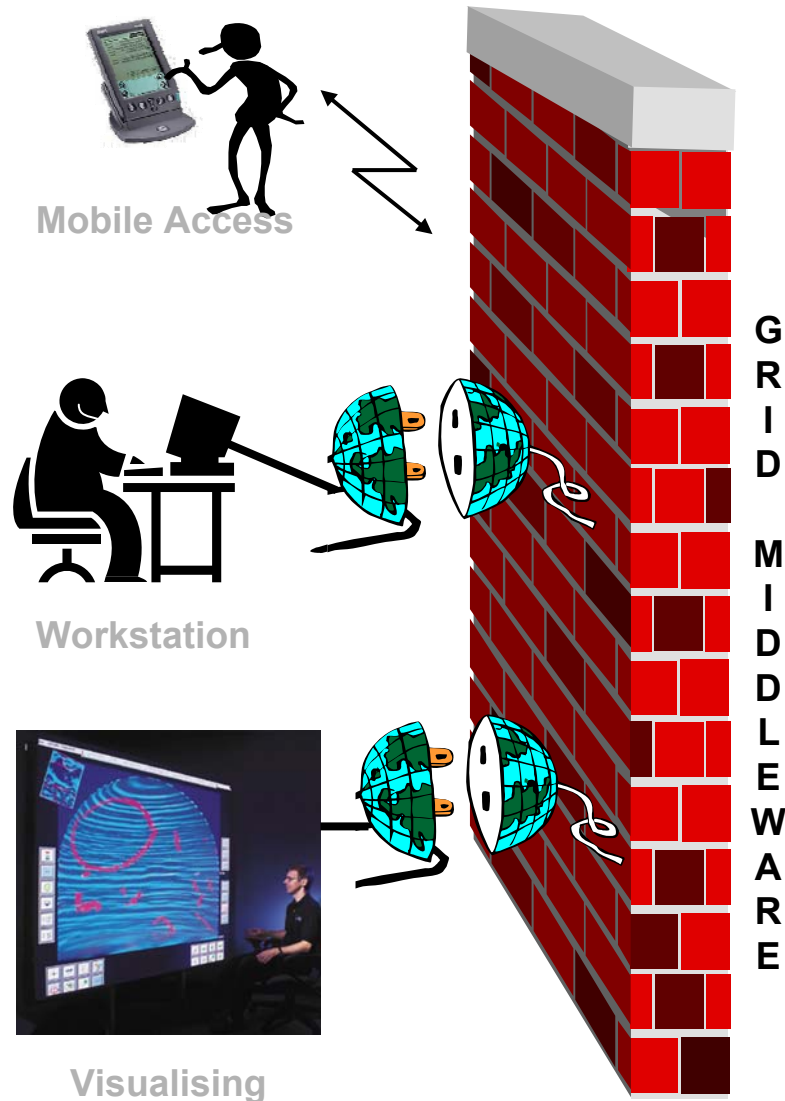
Μεσισµικό (Middleware)

- Λογισµικό που οργανώνει και ενοποιεί διαφορετικούς υπολογιστικούς πόρους που ανήκουν σε ένα Grid
- Το µεσισµικό του Grid κρύβει µεγάλο µέρος από την πολυπλοκότητα του Grid περιβάλλοντος από το χρήστη και του δίνει την εντύπωση ότι όλοι οι πόροι είναι διαθέσιµοι σε αυτόν σε ένα ενιαίο εικονικό κέντρο
- LCG, Globus, Condor



Condor
High Throughput Computing

To Grid μεταφορικά ...



Μεσισμικό (Middleware)

- Βελτιστοποιεί τη χρήση των γεωγραφικά κατανεμημένων πόρων
 - Εξασφαλίζει αποτελεσματική πρόσβαση σε επιστημονικά δεδομένα
 - Είναι υπεύθυνο για την ταυτοποίηση των χρηστών όταν υποβάλλουν μία εργασία σε ένα site
 - Βρίσκει το “κατάλληλο μέρος” για να εκτελεστεί μία εργασία
 - Υπεύθυνο για εκτέλεση των εργασιών
 - Καταγράφει την πορεία εκτέλεσης μίας εργασίας
 - Επανακάμπτει από προβλήματα
- ⇒ Ενημερώνει το χρήστη όταν μία εργασία εκτελεστεί και επιστρέφει το αποτέλεσμα

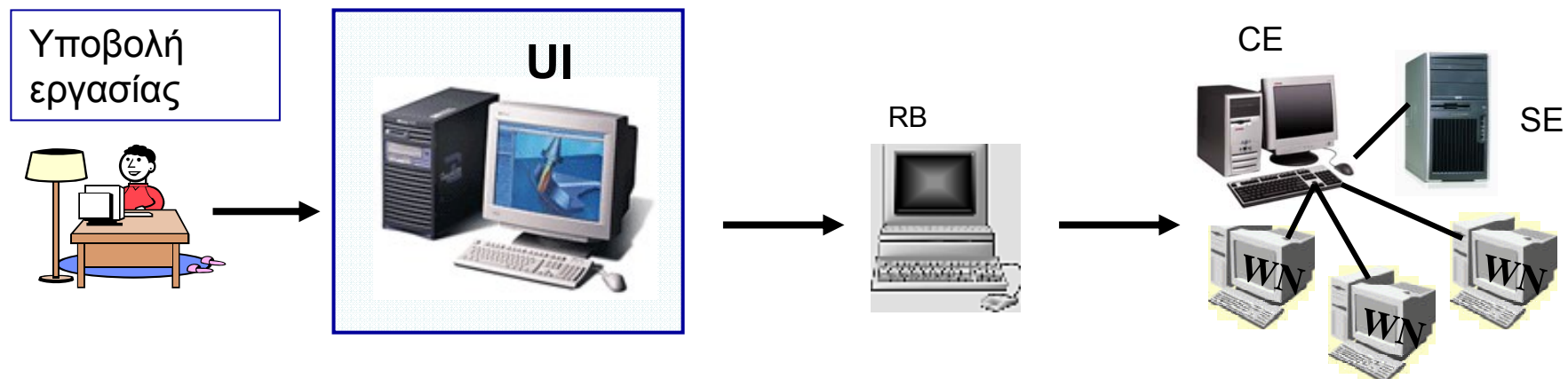
Globus Toolkit

- Grid project
 - Πρωτόκολλα
 - Υπηρεσίες
- Αναπτύσσεται από το Globus Alliance
- Λογισμικά εργαλεία για τη δημιουργία υπολογιστικών πλεγμάτων
- ✚ Υποδομή “ανοιχτού κώδικα” που περιλαμβάνει πολλές υπηρεσίες που χρησιμοποιούνται για την ανάπτυξη εφαρμογών Grid που σχετίζονται με την ασφάλεια, την ανακάλυψη πόρων, την διαχείριση πόρων και την πρόσβαση σε δεδομένα
- Επιλογή υπηρεσιών ανάλογα με τις ανάγκες των σχεδιαστών των εφαρμογών
- GRAM, GSI, MDS, GRIS, GIIS, GridFTP, Replica Catalog, Replica Management System



User Interface (1)

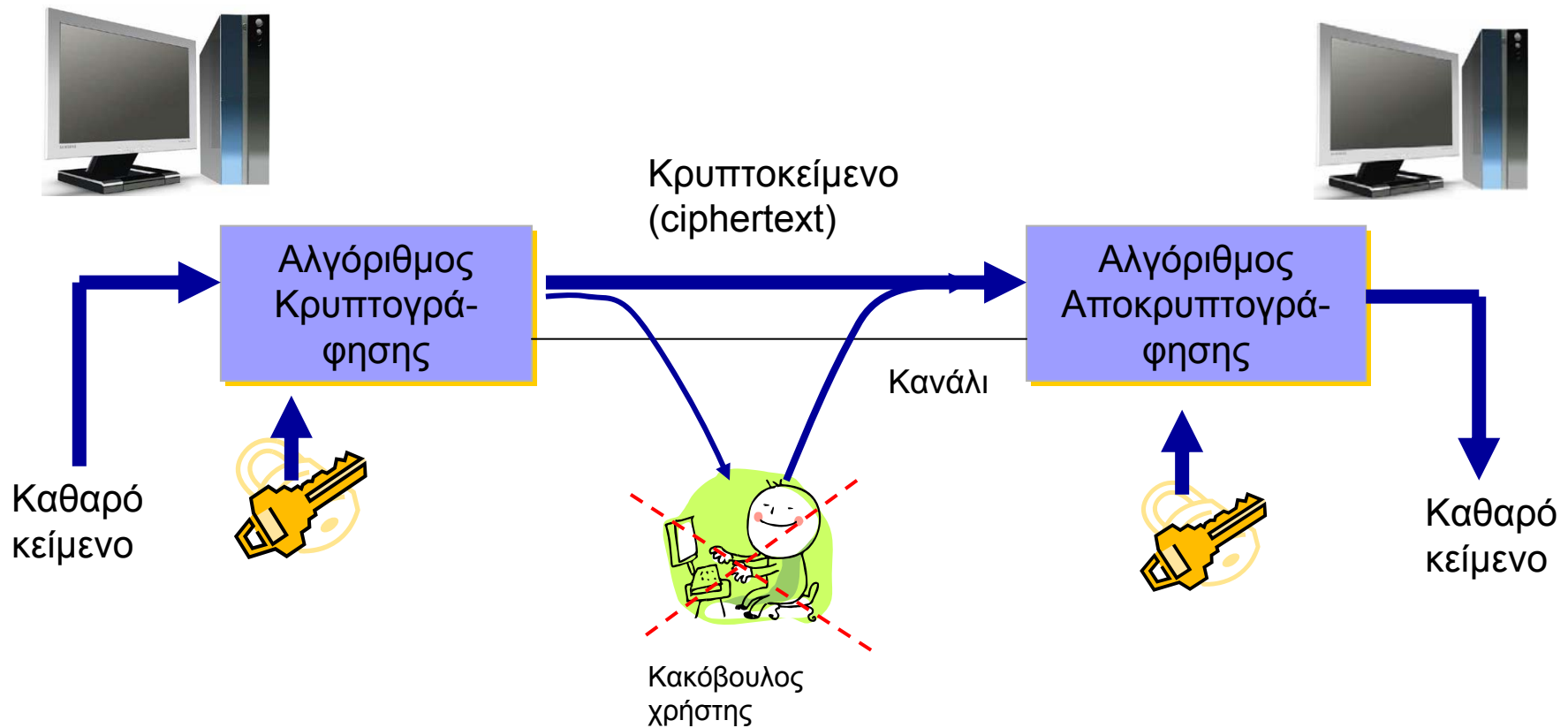
- Επιτρέπει στους χρήστες να έχουν πρόσβαση στις λειτουργίες του Grid
- Οι χρήστες αποκτούν σε αυτό προσωπικό λογαριασμό και εγκαθιστούν το προσωπικό τους πιστοποιητικό
- Πύλη για τα Grid Services



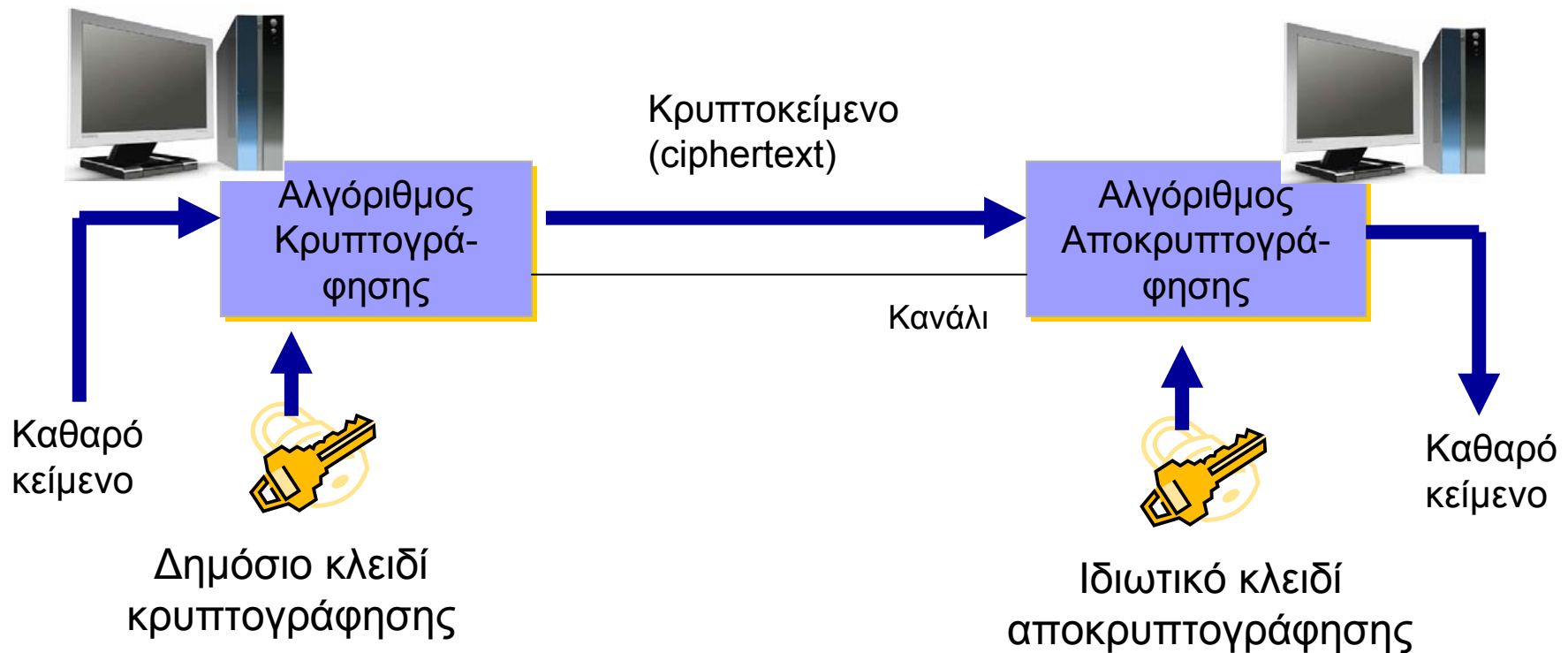
User Interface (2)

- Παρέχει στο χρήστη ένα *Command Line Interface* για την υλοποίηση κάποιων βασικών λειτουργιών, όπως :
 - ↪ Εύρεση όλων των υπολογιστικών πόρων που είναι συμβατοί με τις απαιτήσεις μιας υποβαλλόμενης εργασίας
 - ↪ Υποβολή (submit) μίας εργασίας
 - ↪ Παρακολούθηση της πορείας εκτέλεσης της εργασίας
 - ↪ Ακύρωση μίας ή περισσότερων εργασιών
 - ↪ Ανάκτηση των πληροφοριών υποβολής μίας εργασίας
 - ↪ Λήψη της εξόδου μίας ή περισσότερων εργασιών που έχουν ολοκληρωθεί
 - ↪ Ανάκτηση των δεδομένων εξόδων από τις εργασίες που εκτελέστηκαν

Συστατικά κρυπτογράφησης



Κρυπτογραφία δημοσίου κλειδιού



Ψηφιακό πιστοποιητικό X.509

- Κάθε οντότητα (χρήστης, υπολογιστικό μηχάνημα, άλλος πόρος) πρέπει να αποκτήσει ένα πιστοποιητικό
- Το πιστοποιητικό πρέπει να περιλαμβάνει πληροφορίες, όπως την ημερομηνία λήξης του, την Αρχή Πιστοποίησης από την οποία υπογράφηκε, το δημόσιο κλειδί του κατόχου και ένα θέμα
- Το θέμα του πιστοποιητικού (DN) προσδιορίζει μοναδικά τον κάτοχο του στο Grid και αποτελείται από τα εξής πεδία:

C = Χώρα του κατόχου

O = Οργανισμός του κατόχου

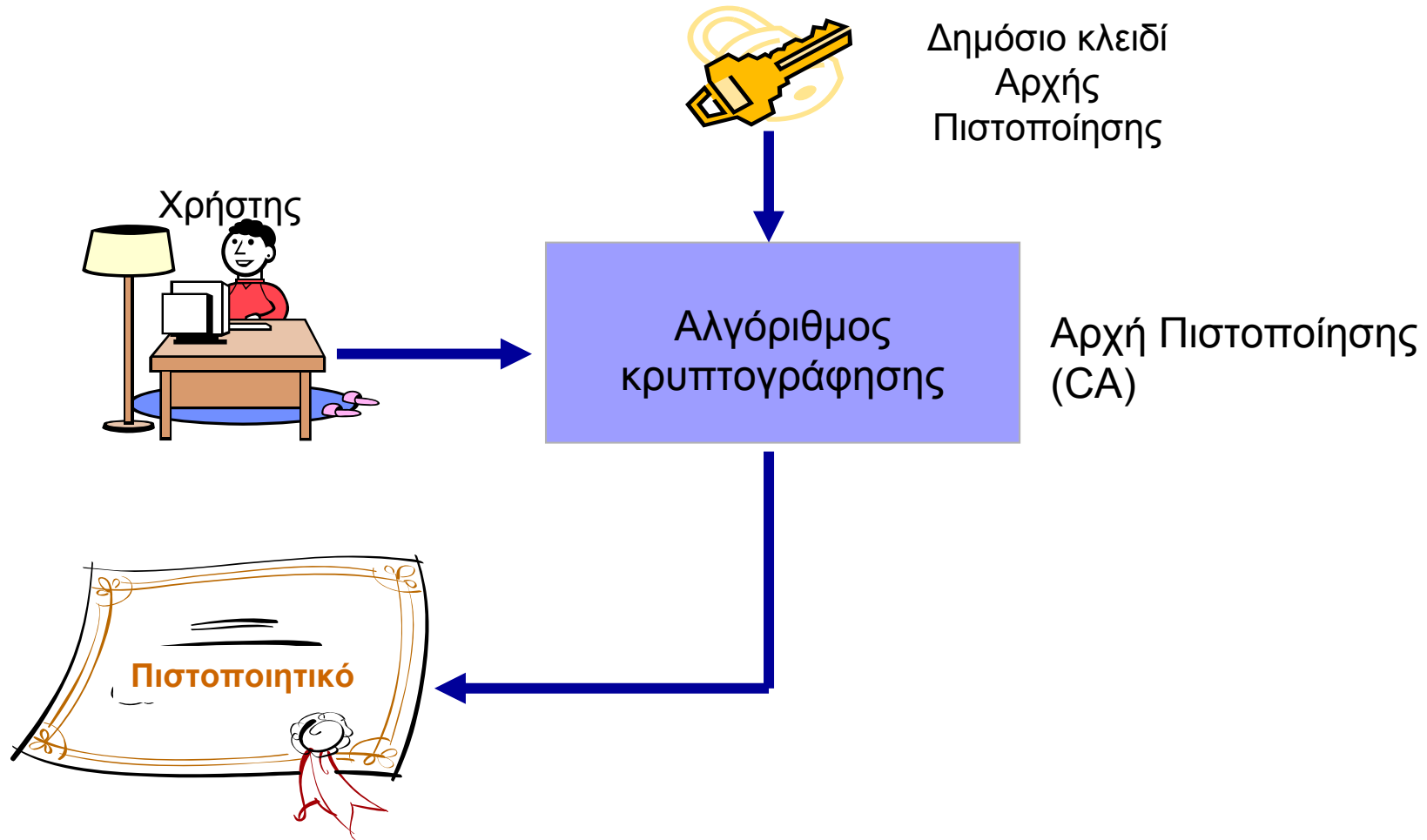
OU = Τμήμα του οργανισμού στο οποίο ανήκει ο κάτοχος

CN = Όνομα του χρήστη ή του πόρου

Αρχή Πιστοποίησης (Certification Authority)

- Είναι ένα εμπιστευόμενο τρίτο μέλος, το οποίο:
 - ✓ εκδίδει τα δημόσια κλειδιά
 - ✓ πιστοποιεί τις πληροφορίες που περιέχονται σε ένα πιστοποιητικό υπογράφοντας το με το δικό της ιδιωτικό κλειδί
- Η υπογραφή της Αρχής Πιστοποίησης γίνεται μέρος του πιστοποιητικού, το οποίο μπορεί να επαληθευτεί με τη χρήση του δημοσίου κλειδιού της Αρχής Πιστοποίησης από οποιονδήποτε την εμπιστεύεται

Πιστοποιητικό από Αρχή Πιστοποίησης

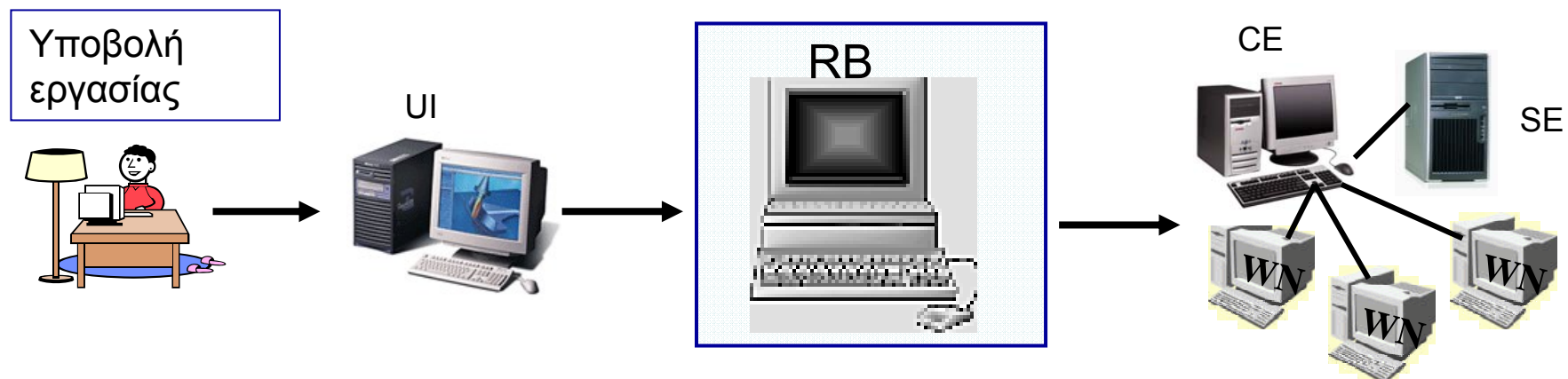


Proxy πιστοποιητικά

- Δημιουργία νέου προσωρινού πιστοποιητικού με τη χρήση του κανονικού πιστοποιητικού
 - ⇒ δημιουργείται ένα καινούργιο ζεύγος δημοσίου και ιδιωτικού κλειδιού για χρήση όσο το προσωρινό πιστοποιητικό είναι ενεργό
- Το καινούργιο ιδιωτικό κλειδί δεν ασφαλίζεται από κωδικό του χρήστη
- Οι κίνδυνοι από τη μη προφύλαξη του ιδιωτικού κλειδιού ελαττώνονται γιατί
 - ✓ το *proxy* πιστοποιητικό μπορεί να ρυθμιστεί να είναι έγκυρο για μικρό χρονικό διάστημα
 - ✓ χρησιμοποιεί ένα διαφορετικό ιδιωτικό κλειδί από αυτό του κανονικού πιστοποιητικού

Resource Broker

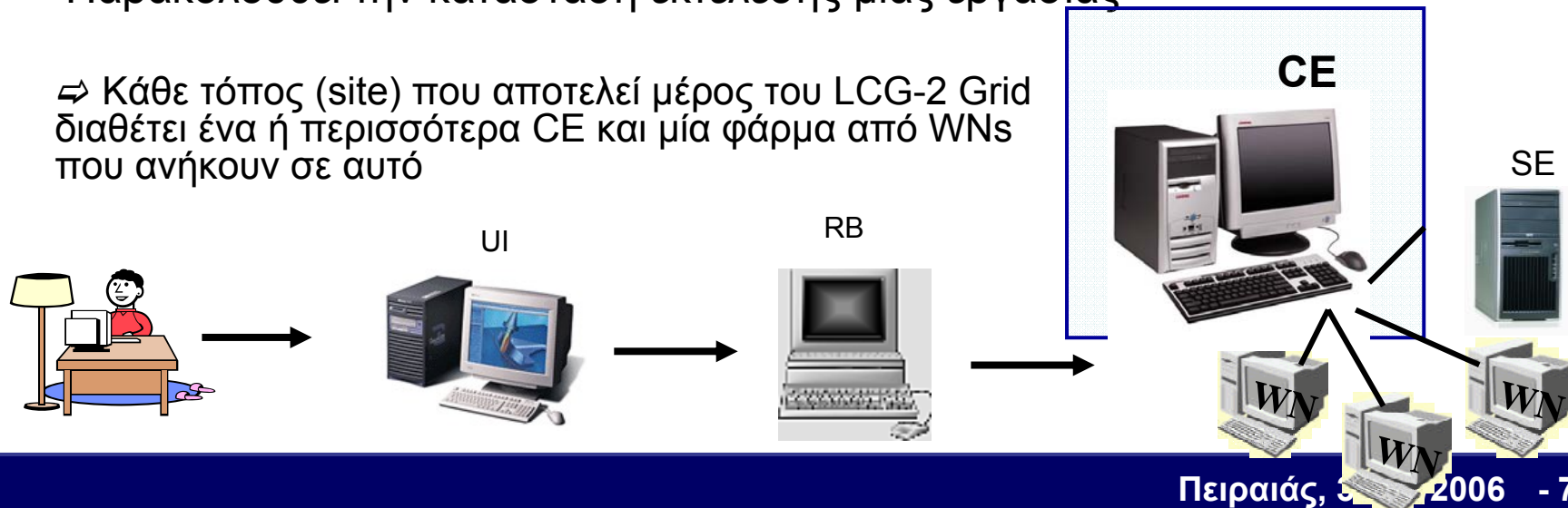
- Λαμβάνει τις εντολές χρηστών για την υποβολή μίας εργασίας
- Εξετάζει τους καταλόγους πληροφοριών για να βρει τους κατάλληλους υπολογιστικούς πόρους για την εκτέλεση της εργασίας



Computing Element

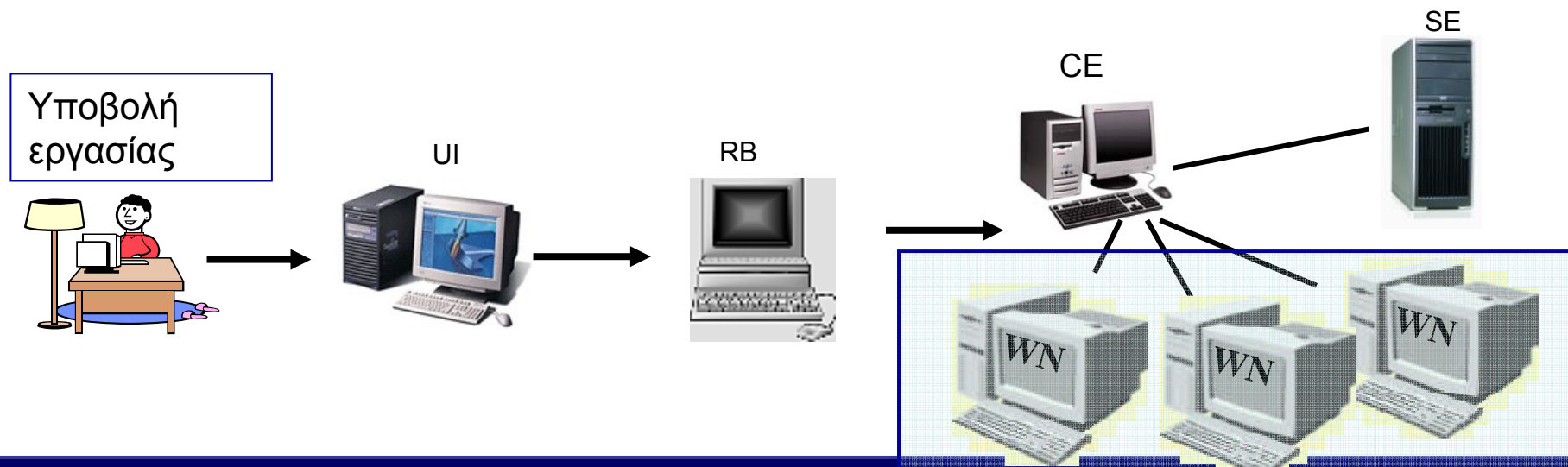
- “Grid interface”
- Διαχειρίζεται μία φάρμα ομογενών υπολογιστικών κόμβων τα οποία ονομάζονται Worker Nodes
- Εκτελεί τις βασικές συναρτήσεις των ουρών αναμονής
- Χρησιμοποιείται για την υποβολή των εργασιών για εκτέλεση στα Worker nodes
- Παρακολουθεί την κατάσταση εκτέλεσης μίας εργασίας

⇒ Κάθε τόπος (site) που αποτελεί μέρος του LCG-2 Grid διαθέτει ένα ή περισσότερα CE και μία φάρμα από WNs που ανήκουν σε αυτό



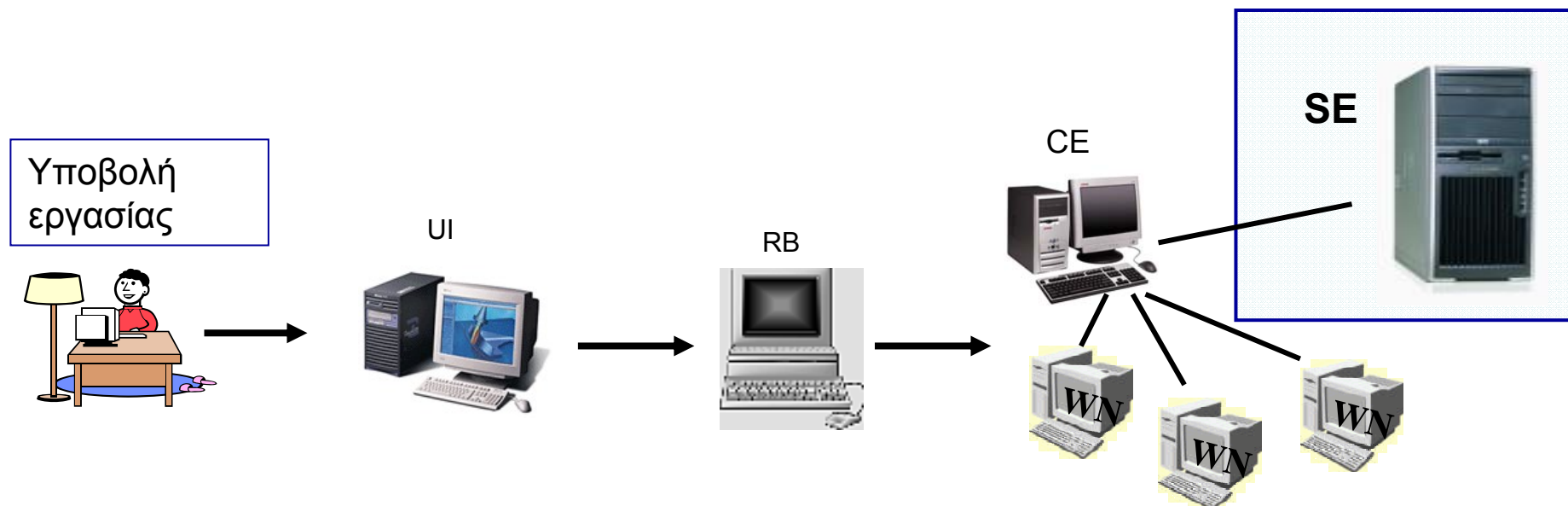
Worker Node

- Κόμβοι για να εκτελούν τις εργασίες
- Απαιτείται μόνο μικρό μέρος του middleware για να είναι συμβατά με την τεχνολογία grid
π.χ βιβλιοθήκες εφαρμογών
Application Programming Interfaces (API)
εντολές



Storage Element

- Πρόσβαση και υπηρεσίες σε αποθηκευτικούς χώρους (μπορεί να ελέγχει απλούς εξυπηρετητές δίσκων, συστάδες δίσκων ή συστήματα μαζικής αποθήκευσης (*Mass Storage Systems - MSS*))
- Κάθε τόπος (site) που αποτελεί μέρος του LCG-2 Grid έχει ένα ή περισσότερα Storage Element διαθέσιμα



Information Service (IS)

- Παρέχει πληροφορίες σχετικά με τους υπολογιστικούς πόρους και την κατάσταση στην οποία βρίσκονται
 - ⇒ Αυτές οι πληροφορίες είναι απαραίτητες για τη λειτουργία μίας υποδομής πλέγματος!
- Οι πληροφορίες αυτές γνωστοποιούνται από υπηρεσίες που εκτελούνται στους ίδιους τους κόμβους και αποθηκεύονται σε βάσεις δεδομένων
- Χρησιμοποιούνται για την αντιστοίχιση των εργασιών με τους κόμβους που ικανοποιούν τις απαιτήσεις τους και για τη δρομολόγηση τους σε αυτούς
- Χρησιμοποιούνται για την επιλογή αποθηκευτικών πόρων και την εύρεση αντιγράφων
- Οι πληροφορίες που δημοσιεύονται χρησιμοποιούνται ακόμα για
 - ✓ “monitoring” σκοπούς ⇒ για την ανάλυση της χρησιμοποίησης της απόδοσης του Grid και χρησιμεύει για την ανίχνευση λαθών
 - ✓ “accounting” σκοπούς ⇒ για τη δημιουργία στατιστικών για τις εφαρμογές που εκτελούνται από τους χρήστες σε διάφορους πόρους

Monitoring services

- **GridICE**

- ✓ Αποτελεί το πιο ολοκληρωμένο monitoring system

- συλλογή
- αποθήκευση
- απεικόνιση

πληροφοριών για την κατάσταση των υπολογιστικών πόρων

- ⇒ Πληροφορίες χαμηλού επιπέδου όπως το φορτίο των υπολογιστικών μονάδων, διαθέσιμη μνήμη, χρήση αποθηκευτικών μέσων, κτλ

- ⇒ Πληροφορίες για την κατάσταση των υπηρεσιών

- ⇒ Πληροφορίες για το Grid, όπως τον αριθμό των υπολογιστικών μονάδων που χρησιμοποιούνται, τον αριθμό των εργασιών που εκτελούνται και τον αριθμό των εργασιών που αναμένουν να εκτελεστούν, τις ελεύθερες υπολογιστικές μονάδες και τον διαθέσιμο αποθηκευτικό χώρο, κτλ.

INFN - GridICE - Grid Monitoring Service - Microsoft Internet Explorer

Αρχείο Επεξεργασία Προβολή Αγαπημένα Εργαλεία Βοήθεια

Norton AntiVirus

Πίσω Αναζήτηση Αγαπημένα

Διεύθυνση <http://gridice2.cnaf.infn.it:50080/gridice/site/site.php> Μετάβαση

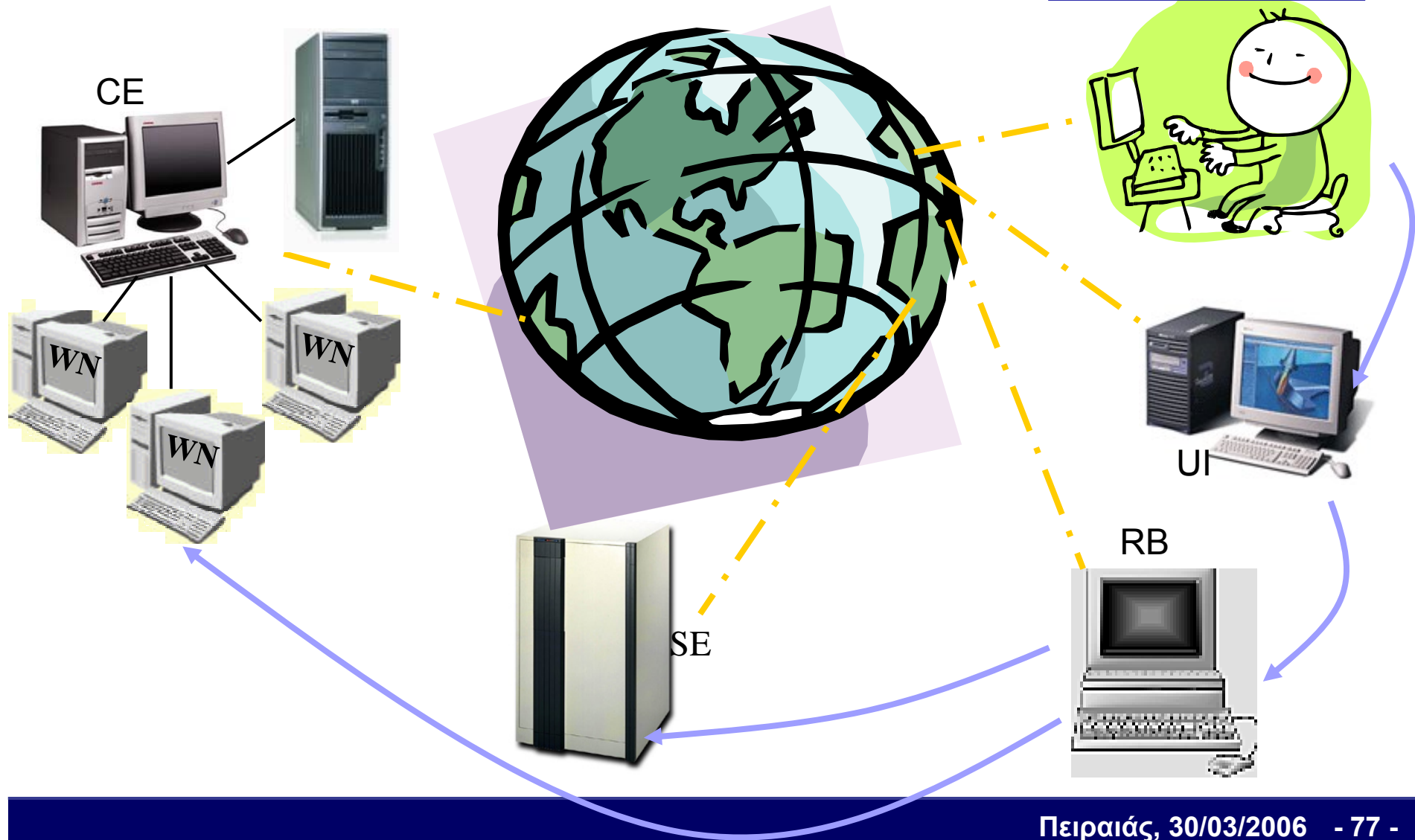
Google Search 1 blocked Check AutoLink AutoFill Options

Site ▼	Computing Resources											Storage Resources		
	GK#	Q#	RunJob	WaitJob	SlotLoad	MH#	Power	WN#	CPU#	CPUload	Available	Total	%	
CIEMAT-LCG2	1	5	2	0	3%	40	613K	24	96	0%	2.4 TB	4.1 TB	0%	
CINES	1	2	0	0	0%	2	-	-	-	-	4.5 GB	7.5 GB	0%	
CNB-LCG2	1	4	8	0	5%	11	126K	8	32	25%	63.5 GB	67.7 GB	6%	
CNR-ILC-PISA	1	6	2	1	5%	4	18K	2	4	25%	381.3 GB	381.6 GB	0%	
CSCS-LCG2	1	5	0	0	0%	3	-	-	-	-	5.8 TB	6.8 TB	14%	
CY01-LCG2	1	7	2	1	8%	5	-	-	-	-	629 GB	655.2 GB	4%	
CYFRONET-IA64	1	6	0	0	6%	-	-	-	-	-	-	-	100%	
CYFRONET-LCG2	1	10	2	0	2%	1	-	-	-	-	103.5 GB	2 TB	95%	
DESY-HH	4	41	190	17	87%	12	-	-	-	-	4.1 TB	6.9 TB	0%	
Durham	1	8	4	1	5%	-	-	-	-	-	1.2 TB	1.8 TB	29%	
ENEA-INFO	1	3	0	0	7%	17	67K	14	14	0%	14.8 GB	18.2 GB	19%	
ESA-ESRIN	1	3	0	0	0%	8	12K	6	6	0%	389.1 GB	845.1 GB	54%	
FMPHI-UNIBA	1	3	0	2	0%	1	-	-	-	-	2.6 TB	2.7 TB	6%	
FZK-LCG2	2	29	518	4	97%	-	-	-	-	-	2.3 TB	10.3 TB	78%	
GOG-Singapore	1	6	0	0	14%	2	-	-	-	-	36.2 GB	59.6 GB	39%	
GR-01-AUTH	-	-	-	-	-	9	-	-	-	-	465.2 GB	478.7 GB	3%	
GR-02-UoM	1	8	3	5	38%	3	-	-	-	-	856.1 GB	873.6 GB	2%	
GR-03-HEPNTUA	2	8	3	0	23%	2	-	-	-	-	846.5 GB	873.6 GB	3%	
GR-04-FORTH-ICS	1	9	3	2	60%	6	41K	3	8	28%	0.9 GB	7.6 GB	89%	
GR-05-DEMOKRITOS	1	2	4	1	10%	4	-	-	-	-	1.1 TB	1.1 TB	2%	
GRIF	2	5	4	0	22%	10	-	-	-	-	995.4 GB	1.1 TB	11%	
GSI-LCG2	4	6	0	0	0%	-	-	-	-	-	65.3 GB	67.7 GB	4%	
HEPHY-UIBK	1	2	1	0	4%	2	-	-	-	-	1.1 TB	1.6 TB	32%	
HG-01-GRNET	1	15	14	0	22%	28	512K	23	92	14%	34 TB	35.4 TB	4%	
HPC2N	1	1	33	864	34%	2	-	-	-	-	-	-	-	

Internet

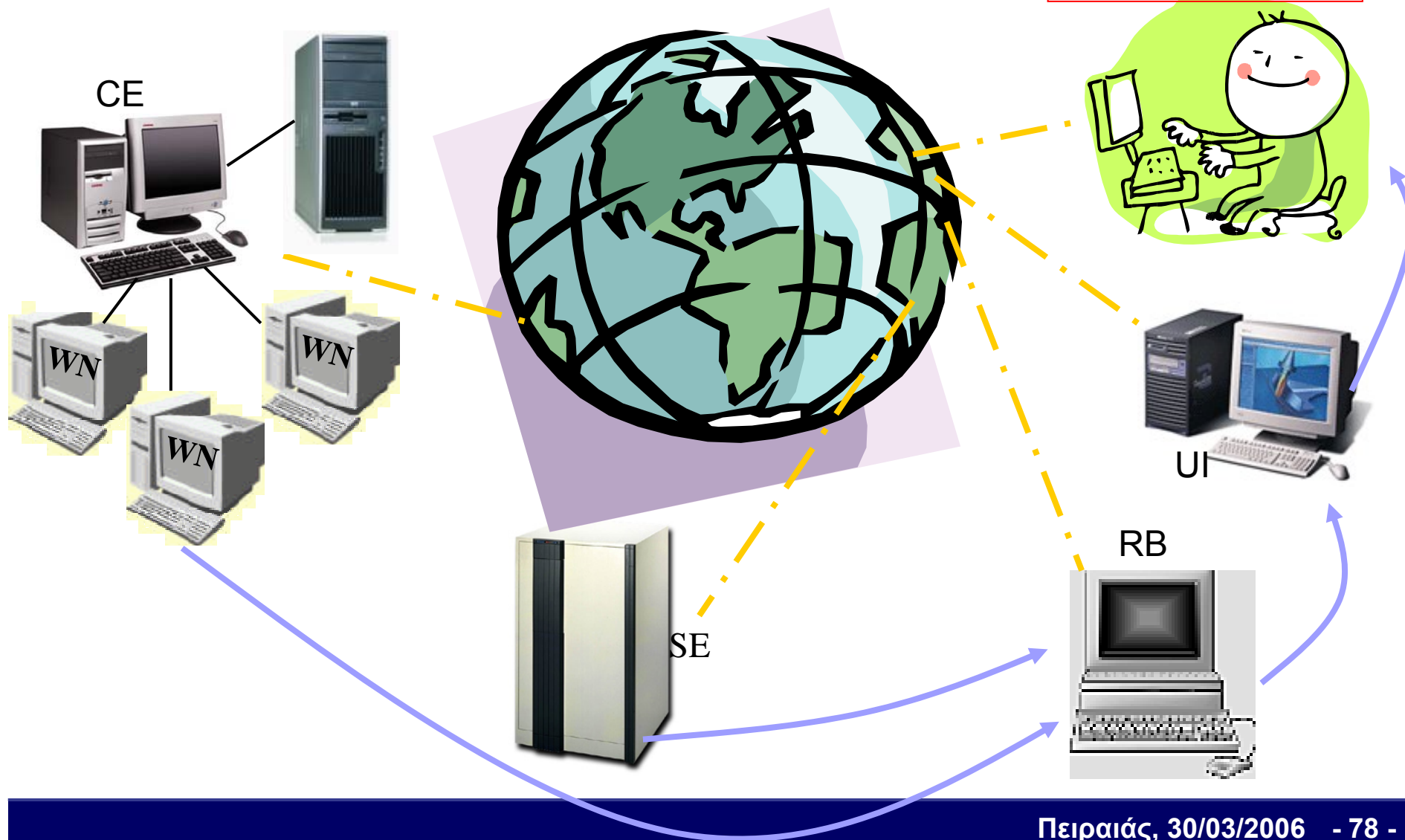
Υποβολή εργασίας στο Grid

Υποβολή εργασίας



Υποβολή εργασίας στο Grid

Εργασία εκτελέστηκε



Q&A



Αναφορές

- ✓ Grid café:
<http://gridcafe.web.cern.ch/gridcafe>
- ✓ Global Grid Forum:
<http://www.gridforum.org/>
- ✓ Gridtoday:
<http://www.gridtoday.com/gridtoday.html>
- ✓ Grid Computing
http://en.wikipedia.org/wiki/Grid_computing
- ✓ Distributed Computing
http://en.wikipedia.org/wiki/Distributed_systems
- ✓ Supercomputing
<http://en.wikipedia.org/wiki/Supercomputing>
- ✓ LCG-2 User Guide
http://egee.itep.ru/User_Guide.html
- ✓ EGEE (Enabling Grids for E-science)
<http://public.eu-egee.org/intro/>