

Le griglie computazionali nella fisica delle alte energie



Mario.Reale@cnafe.infn.it

EDG / INFN GRID

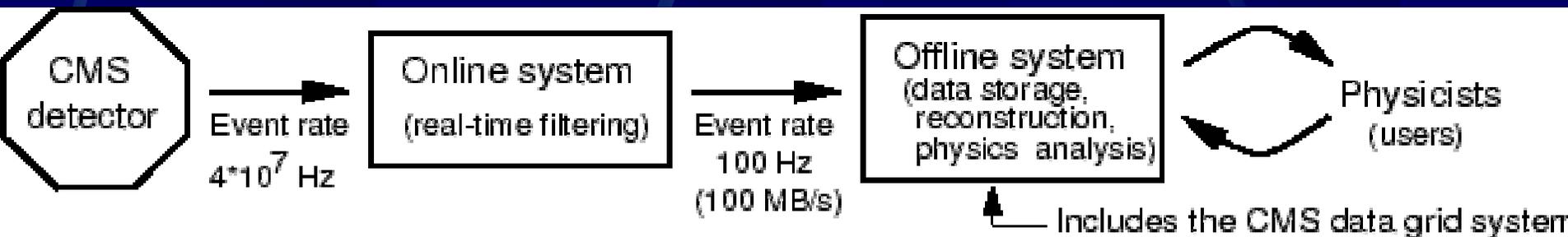


Contenuti

- Perché usare le griglie in Fisica delle Alte Energie
- Breve lista dei progetti esistenti
- European Data Grid : uno sguardo d'insieme e sul networking
- DEMO : Job Submission e GENIUS

Perche' utilizzare le Griglie in HEP

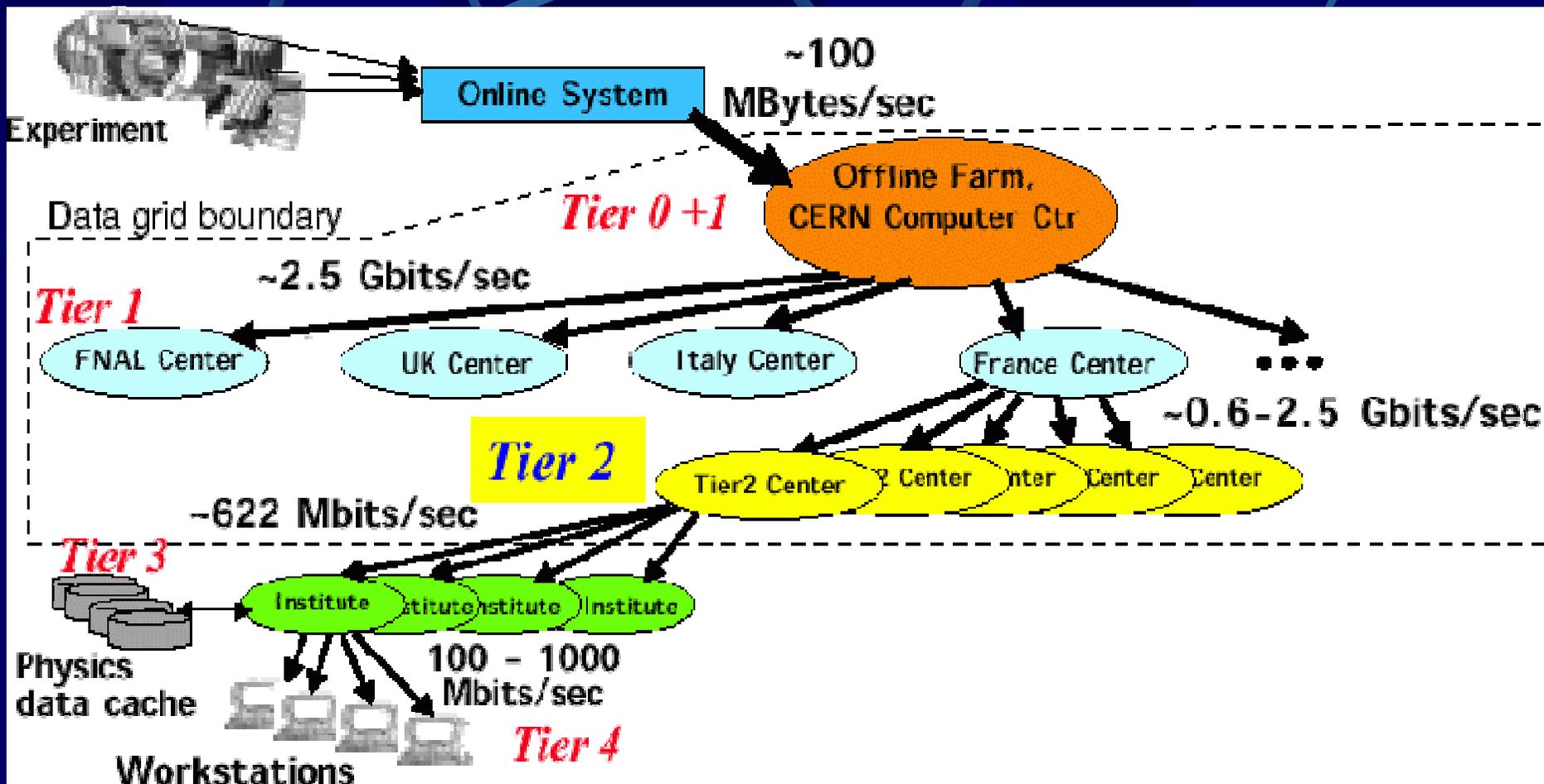
- Natura distribuita del problema: una comunita' mondiale di fisici analizza una mole mai sperimentata prima di dati
- Esempio: Un esperimento di LHC come CMS:
 - On line : ~ 1 PB /sec
 - Off-Line : input to Tier 0 : 100 - 200 MB/s



Qualche ordine di grandezza

- Mole di dati che LHC fornirà:
 - 10^9 evt raw/anno + altrettanti da simulazione
 - Dimensione 1 evento : 1 MB (CMS), 25 MB (Alice)
 - Eventi Simulati : 2 MB (CMS), 2 GB (Alice)
- Quante CPU ci serviranno per analizzarli:
 - Stime CMS: Tier0 : 455,000 SI95
 - Circa 3000 PC in Tier 0
- Quanti dati trasferiremo sulla rete:
 - Difficile da stimare : per trasferire l'equivalente di tutto quanto acquisito e simulato al Tier 0 in un anno da 1 esperimento come CMS (= 3 - 4 PB)
 - a 2.5 Gbps servono 420 ore ~ 18 giorni
 - a 155 Mbps servono 6770 ore ~ 290 giorni

Il modello Monarch - CMS



Griglie Computazionali

- Una Griglia e' un insieme coordinato di risorse distribuite e condivise in maniera dinamica da utenti appartenenti a Organizzazioni Virtuali
- Una Griglia ha continuo bisogno di informazioni aggiornate sullo stato effettivo attuale delle risorse che la compongono (CPU, spazio disco, files, ambiente software disponibile, collegamenti di rete.....)
- Usare le griglie in HEP ha il vantaggio di aumentare enormemente le capacita' di calcolo e la facilita' di accesso a tutte le risorse per ogni utente appartenente ad una VO – rendendolo utente potenziale di ogni risorsa del suo esperimento (non piu' solo del suo Istituto di Ricerca o Centro di Calcolo)

HEP e griglie computazionali: Progetti Esistenti

- EU
 - EU DataGrid (HEP, BioMed, EO Pioneering Project, Globus based)
 - CrossGrid (QoS, RealTime, Streaming)
 - Numerosi Progetti Nazionali (INFN GRID, eScience)
- USA
 - iVDGL
 - GriPhyN
 - PPDG
- DATA TAG

European Data Grid EDG

- Progetto Europeo (EU) , 3 anni.
- Kick Off : 1.1.2001 – Fine : 31.12.2003
- 6 Partner Principali : INFN, CERN, CNRS, ESA-ESRIN, NIKHEF, PPARC

EDG: Partner Secondarii

Industriali:

- Datamat (Italy)
- IBM-UK (UK)
- CS-SI (France)

Istituti Accademici e di Ricerca

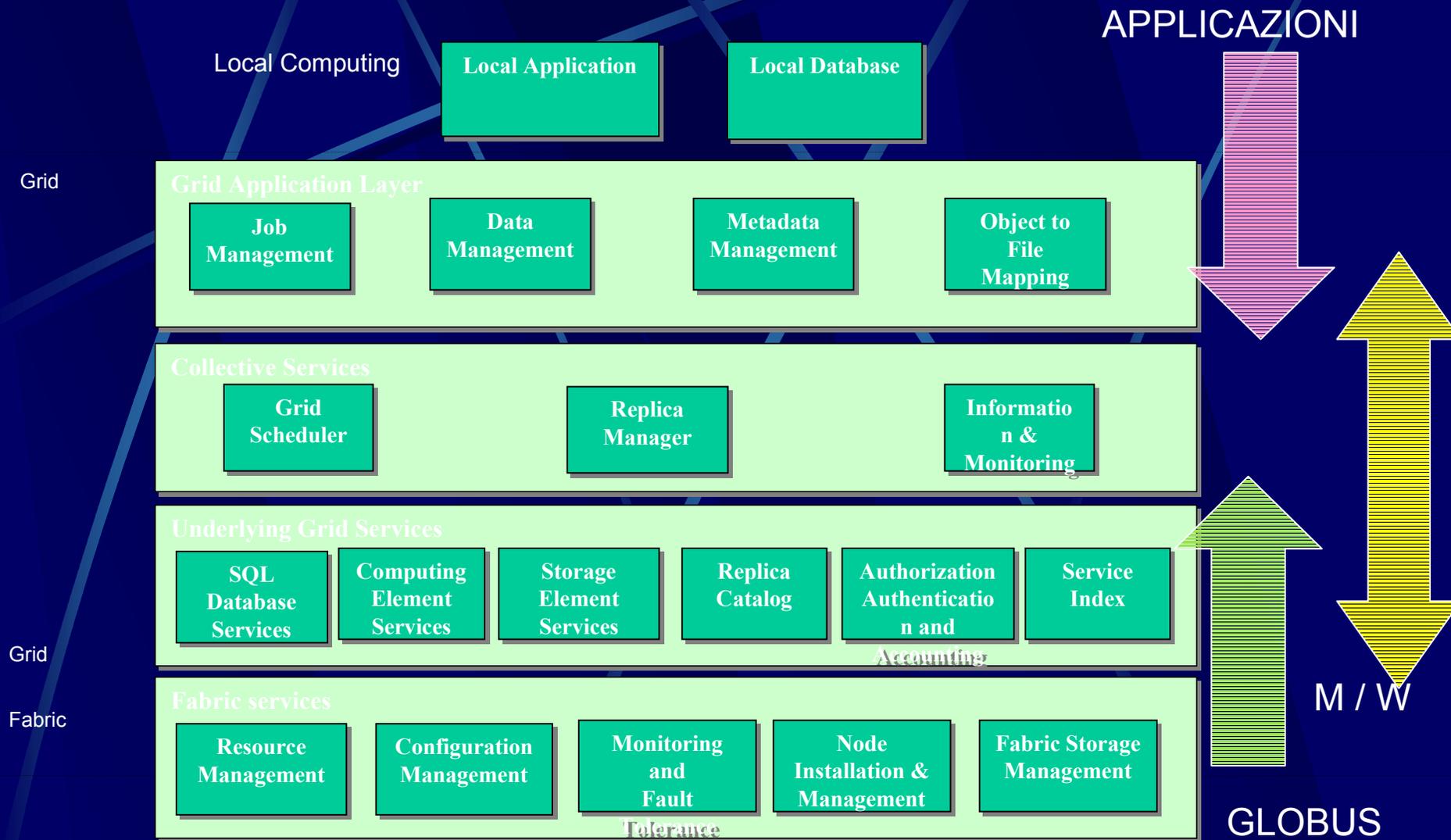
- CESNET (Czech Republic)
- Commissariat à l'énergie atomique (CEA) – France
- Computer and Automation Research Institute, Hungarian Academy of Sciences (MTA SZTAKI)
- Consiglio Nazionale delle Ricerche (Italy)
- Helsinki Institute of Physics – Finland
- Institut de Fisica d'Altes Energies (IFAE) - Spain
- Istituto Trentino di Cultura (IRST) – Italy
- Konrad-Zuse-Zentrum für Informationstechnik Berlin - Germany
- Royal Netherlands Meteorological Institute (KNMI)
- Ruprecht-Karls-Universität Heidelberg - Germany
- Stichting Academisch Rekencentrum Amsterdam (SARA) – Netherlands
- Swedish Research Council - Sweden



EDG : stato attuale del progetto

- EDG fornisce attualmente un insieme di servizi di middleware – che vanno dal Job Management, Data Management, GRID monitoring, Network monitoring, Security, Authentication, Authorization – da installarsi su una piattaforma Linux Red Hat 6.2 tramite tools di fabric management, install & config - e un insieme di servizi comuni centralizzati disponibili per i siti che desiderano entrare in EDG (Resource Broker, VO-LDAP servers per Autenticazione, VO-based Replica Catalogs, VO-management services, Logging and Bookkeeping)
- 9 siti EDG appartengono attualmente al testbed EDG tb 1 :
CERN(CH), RAL(UK), Manchester(UK), NIKHEF(NL), CNAF(I), Catania(I), Torino(I), Padova(I), Lyon(F)
- EDG Release 1.2 sta venendo distribuita e installata sui siti distribuiti dopo essere stata collaudata al CERN
- Intenso Sviluppo middleware in corso :
 - Nuove funzionalità per WMS (WP1) : split up jobs, billing e accounting
 - Nuove Funzionalità di Data Management : API di Replica Management
 - Integrazione delle informazioni di networking sulle politiche di Brokering

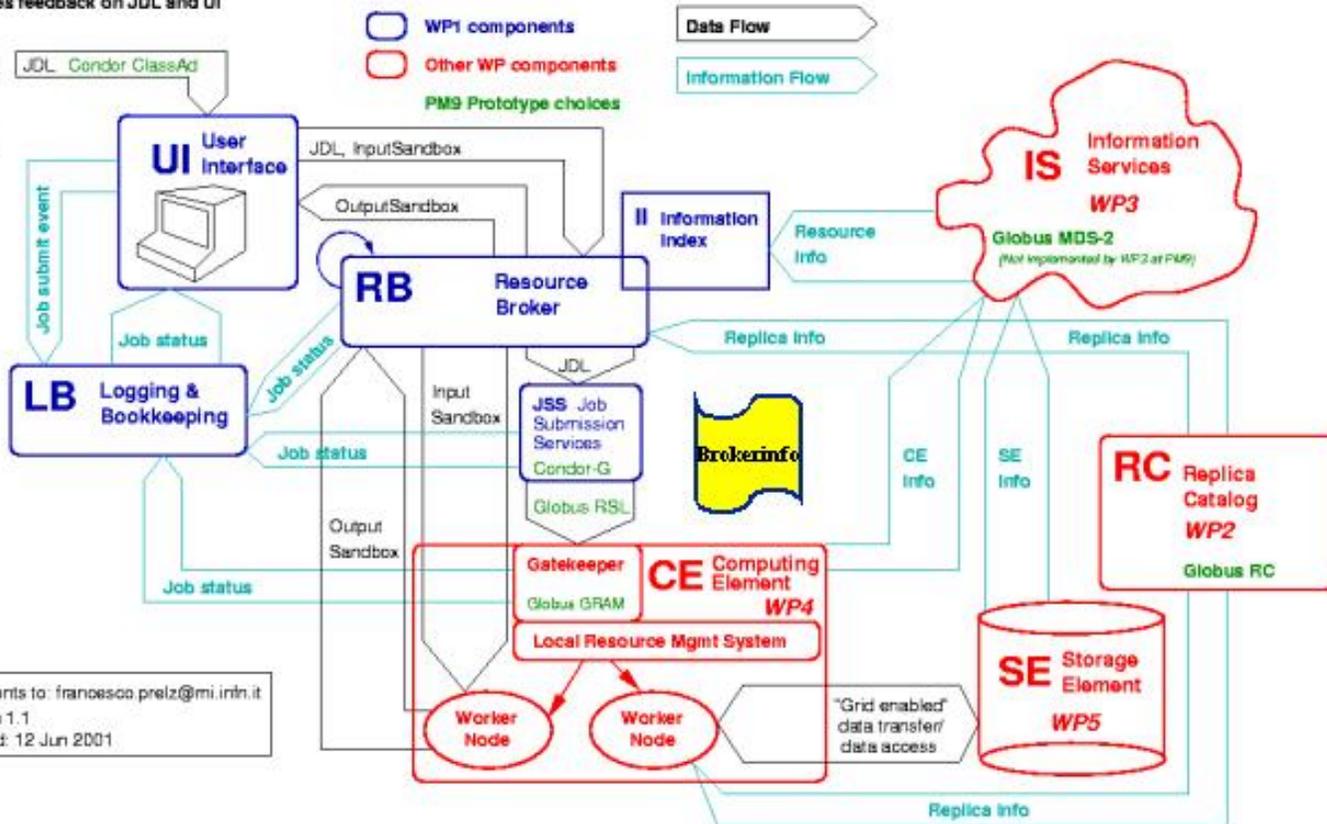
Architettura del MiddleWare EDG



Architettura del MiddleWare EDG

End User **WP8,9,10**

- Specifies job using JDL
- Submits job using UI
- Controls and monitors job(s)
- Provides feedback on JDL and UI

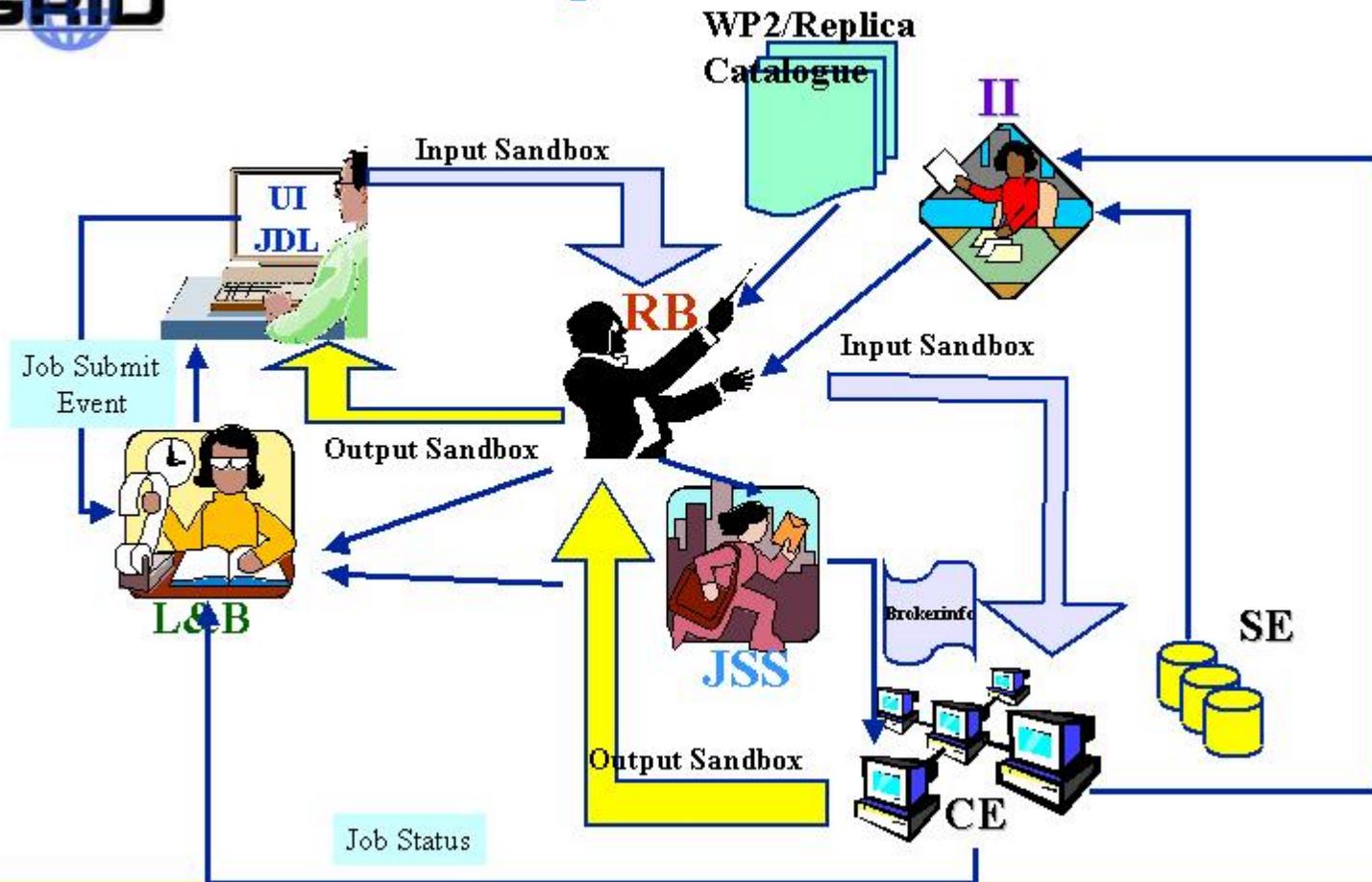


Flavia Donno The EDG Middleware User Interface ° 7

EDG : schema funzionale delle componenti principali



WP1 PM9 Integration



Flavia Donno: The EDG Middleware User Interface * 8

Networking Info all' interno di EDG

- Attualmente EDG (WP7) fornisce nel m/w tools di monitoring delle prestazioni della rete (tra siti appartenenti al testbed):
 - EDG-IPERF
 - Misura il throughput TCP/ IP
 - EDG-UDPMON
 - Misura UDP throughput
 - Misura lost packets rate
 - Misura inter packet jitter
 - EDG-PINGER
 - Misura il round-trip-time (min/max/ave)
 - Misura il packet-loss rate

Inclusione delle informazioni sulle performances della rete

- Attualmente disponibili tramite visualizzazione ad un osservatore umano (user, network/system administrators)
- Per identificare e risolvere problemi (bottlenecks, point of unreliability, quality of service needs, topology...)
- Per ottenere previsioni sulle performance della rete e ottimizzazione — Capacity planning

3 Campi principali di lavoro/ sviluppo per il Networking in EDG:

- Packet forwarding Differenziato

- **Supporto per la Quality of Service:**
 - classification
 - metering
 - policing
 - marking
 - scheduling
 - shaping

- Monitoring e Configurazione del protocollo di trasporto

- Ottimizzazione network-based della griglia

Packet Forwarding Differenziato

- Proposta l' introduzione di una architettura DiffServ (basata percio' su aggregati di traffico e non microflows individuali)

traffic Traffic aggregates and not individual micro flows

- Proposta l' introduzione di diversi livelli di servizio a seconda dell' applicazione di griglia. I Servizi Proposti sono basati su
 - Expedited Forwarding: garanzia sul ritardo (IP Premium)
 - Assured Forwarding (Garanzia su Banda, non su rit. & jitter)
 - Proposto utilizzo di LessThanBestEffort (LBE) per applicazione di testing intrusivo

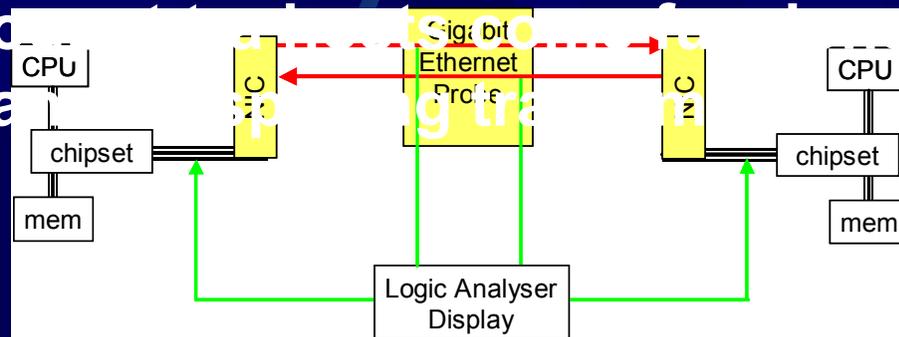
Monitoring e Configurazione del protocollo di trasporto

- EDG WP7 ha effettuato numerosi studi di ottimizzazione del protocollo TCP-IP
 - Proposto l' utilizzo di web100 in ogni sito EDG che necessita di alte performances di rete (per esempio frequently accessed Storage Elements) – Necessita di un kernel patch. Monitoring di variabili fondamentali per TCP IP come
 - numero e tipo di pacchetti trasmessi e ritrasmessi
 - ECN (RFC 3168) and SACK (RFC 2061) availability
 - timeouts e timeouts di ritrasmissione (RTO)
 - La soglia slow-start (`ssthresh` `ssthresh` `ssthresh`) e la dimensione della congestion window size (`wnd` `wnd`)

Studi Sperimentali sulla configurazione del protocollo di trasporto

- *Round trip latency as a function of Ethernet frame size using request-response message* Round trip delay come funzione dell' Ethernet Frame Size (protocol stuck performances , comportamento delle NICs)

- UDP throughput as a function of frame size e tra



Ottimizzazione Network-based della griglia

Inclusione nelle politiche di Brokering:

- Parametri di Network performance possono essere usati per ottimizzare l'allocazione delle risorse (replicazione files , MPI, Accesso a files remoti...)
- Metriche sulle performances della rete devono:
 - Essere pubblicate nel Grid Information System
 - Essere accessibili tramite funzioni aggregate chiamate dai servizi del GRID Resource Broker (computing e data storage).

Ottimizzazione Network-based della griglia

- Proposta l' introduzione di un NGO : Network Grid Optimizer node che, basandosi sulla definizione di Cost Functions – calcolate in base alle funzioni fornite dal Grid Monitoring System – fornisce informazioni al resource broker sul costo di accesso ai dati richiesti a partire da un dato Storage Element.

Conclusioni

- La disponibilità di reti ad alta velocità renderà il futuro delle griglie computazionali sempre più avvincente e consentirà di realizzare l'ambiente di accesso alle risorse distribuite sempre più efficiente e operativo
- In EDG molti sviluppi di networking sono in corso, in maniera da includere la rete come componente viva di griglia e non solo come supporto per i servizi forniti