



Enabling Grids for  
E-science in Europe

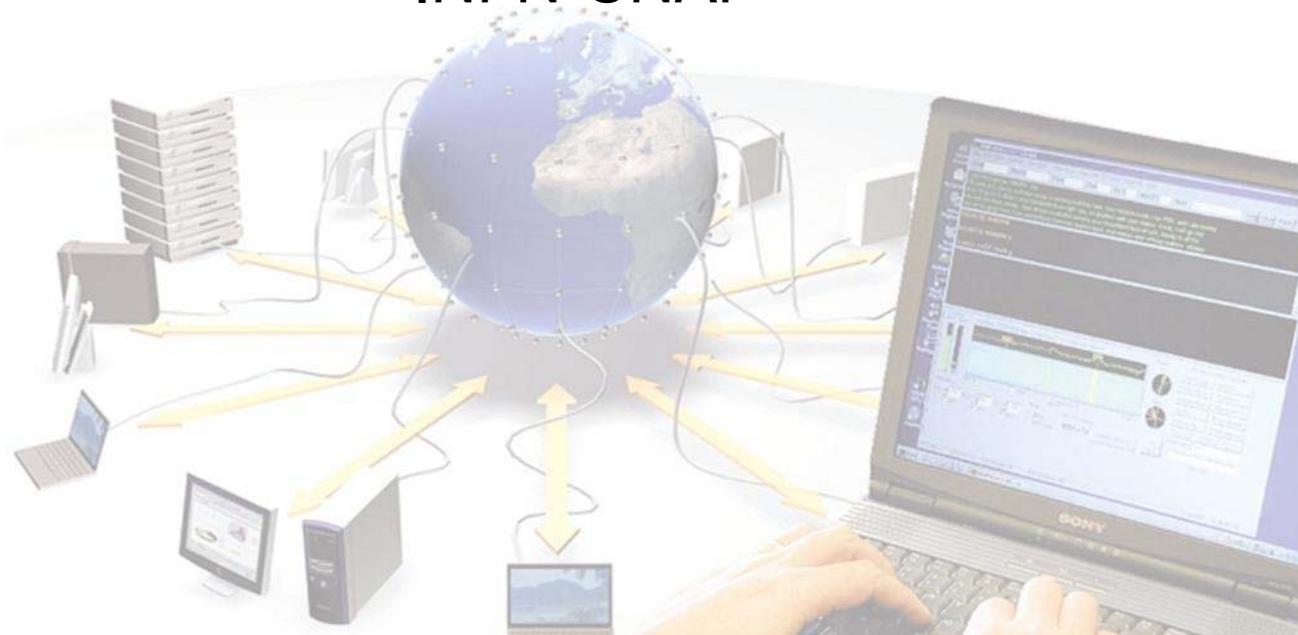
[www.eu-egee.org](http://www.eu-egee.org)



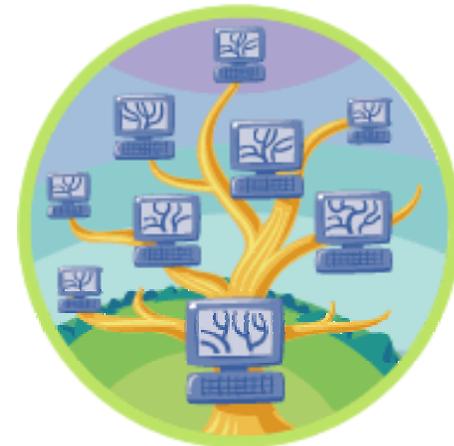
<http://grid.infn.it>

# Data transfer

Paolo Veronesi  
INFN-CNAF

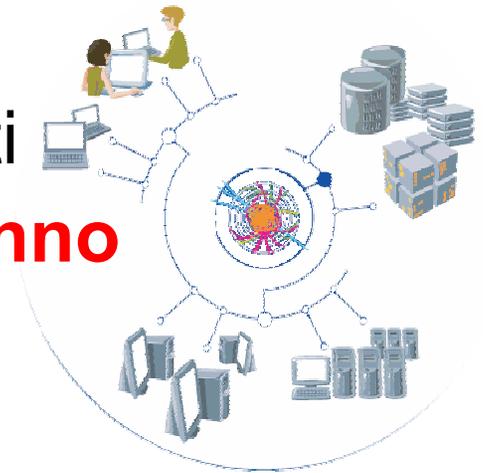


- **La sfida di WLCG**
- **Modello di calcolo degli esperimenti LHC**
  - Esempio: il modello LHCb
  - Esempio: il modello CMS
- **FTS: File Transfer Service**
  - Architettura e concetti;
  - Utilizzo e monitoring;
- **Note su firewall**



# La sfida LHC

- Volume di dati
  - Alto rate x tanti canali x 4 esperimenti
  - ➔ **15 PetaBytes di nuovi dati ogni anno**
- Potenza di calcolo
  - Analisi di eventi x migliaia di utenti
  - ➔ **100k delle piu' veloci CPUs di oggi**
- Analisi e fondi distribuiti Worldwide
  - Fondi per le risorse di calcolo e di storage distribuiti a livello regionale e nazionale
  - Analisi efficiente ovunque
  - ➔ **GRID technology**



# La distribuzione dei progetti EGEE e OSG

- I progetti **EGEE** e **OSG** sono la base del progetto Worldwide LHC Computing Grid Project **WLCG**



**INTEROPERABILITA'**

# La collaborazione WLCG

- La collaborazione
  - 4 LHC esperimenti
  - ~200 centri di calcolo
  - 12 (Tier-0, Tier-1)
  - 38 “Tier-2”
  - 40 Paesi coinvolti
  - Grids: EGEE, OSG

## TIER2:

- Università’;
- Sezioni INFN;
- Laboratori;



LCG-TDR-001  
CERN-LHCC-2005-024



[www.cern.ch/lcg](http://www.cern.ch/lcg)

**LHC Computing Grid**  
Technical Design Report

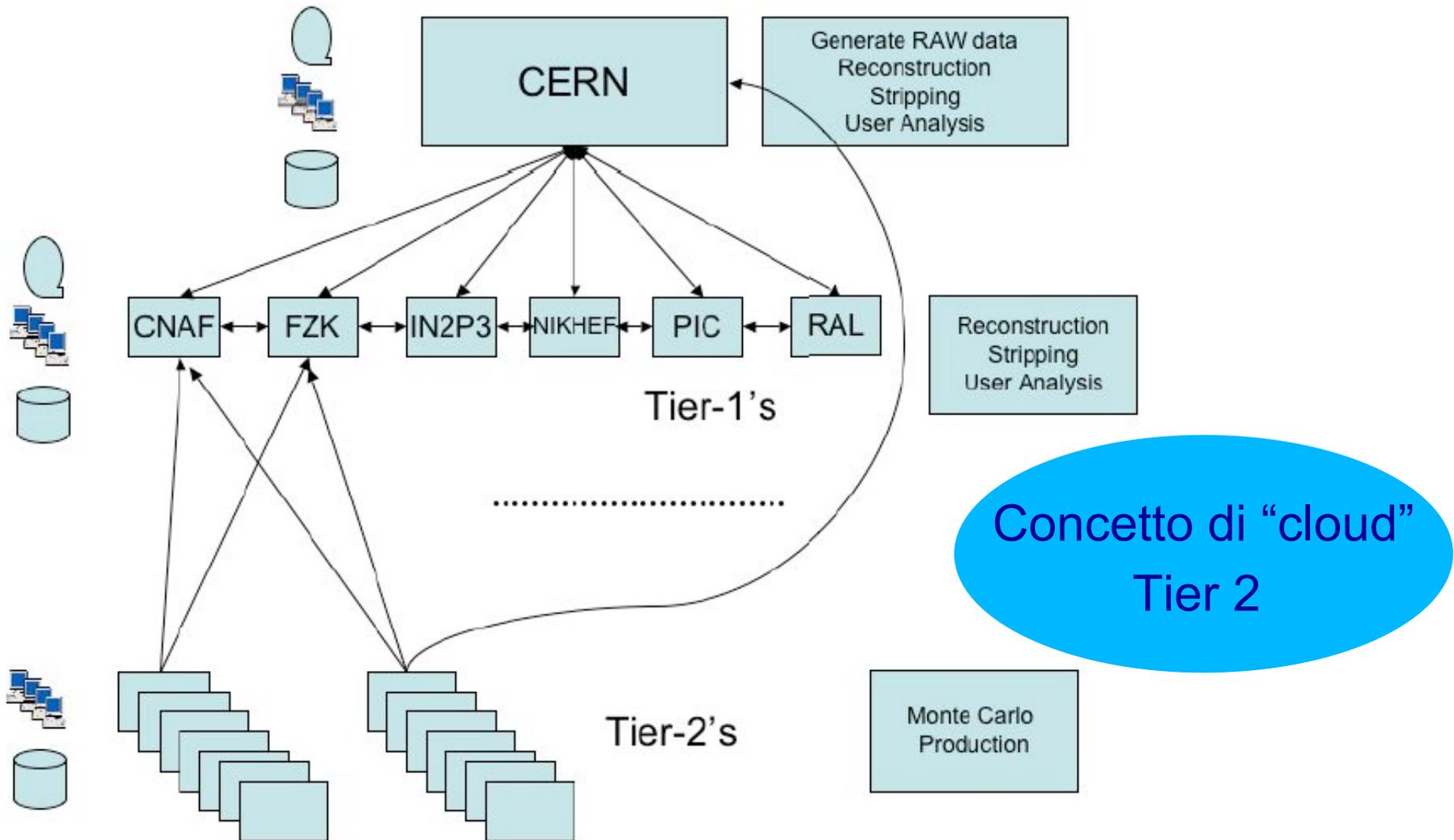
Editor: Jürgen Knobloch



# LHCb Computing Model 1/2

- **Modello gerarchico;**
- **TIER 0: CERN**
  - Centro di produzione centrale, responsabile per la distribuzione dei RAW data in quasi real-time a tutti i Tier1;
- **TIER 0 + 6 TIER 1**
  - Responsabili del processo di produzione associato ai dati reali (dati salvati al CERN e una copia distribuita sui tier1).
  - La maggioranza delle analisi fisiche saranno svolte al CERN e nei tier1.
  - Una certa porzione di dati (DST current year) verra' poi distribuita a tutti i tier2;
- **12 TIER2**
  - Usati quasi esclusivamente per **produzione di eventi simulati** (MonteCarlo), con CERN e TIER1 che faranno da repository centrali
  - Per analisi occorre ricevere gli ultimi DST in semi real-time (50MB/s)

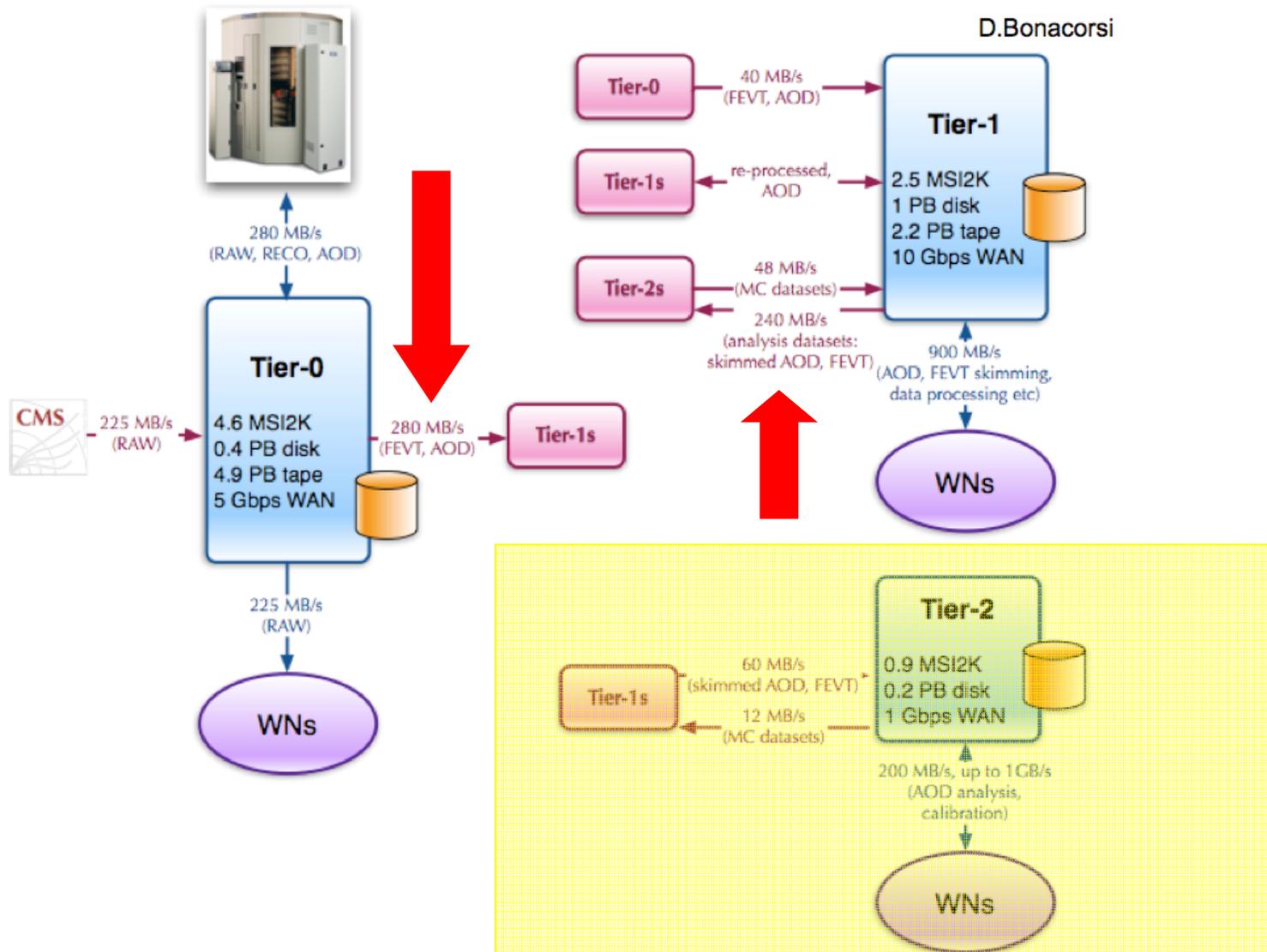
# LHCb Computing Model 2/2



# LHCb computing resource Stime 2006-2010

CPU(MSI2k.yr)	2006	2007	2008	2009	2010
CERN	0.27	0.54	0.90	1.25	1.88
Tier-1's	1.33	2.65	4.42	5.55	8.35
Tier-2's	2.29	4.59	7.65	7.65	7.65
<b>Total</b>	<b>3.89</b>	<b>7.78</b>	<b>12.97</b>	<b>14.45</b>	<b>17.88</b>
Disk(TB)					
CERN	248	496	826	1095	1363
Tier-1's	730	1459	2432	2897	3363
Tier-2's	7	14	23	23	23
<b>Total</b>	<b>984</b>	<b>1969</b>	<b>3281</b>	<b>4015</b>	<b>4749</b>
MSS (TB)					
CERN	408	825	1359	2857	4566
Tier-1's	622	1244	2074	4285	7066
<b>Total</b>	<b>1030</b>	<b>2069</b>	<b>3433</b>	<b>7144</b>	<b>11632</b>

# Il modello CMS



# INFN Tier 2 sites

- **ALICE**
  - **INFN-CATANIA**
  - **INFN-TORINO**
  - INFN-BARI (TBD)
  - INFN-LEGNARO (TBD)
- **ATLAS**
  - **INFN-ROMA1**
  - **INFN-NAPOLI**
  - INFN-MILANO (sj)
  - INFN-FRASCATI (in)
- **CMS**
  - **INFN-ROMA1**
  - **INFN-LEGNARO**
  - INFN-BARI (in)
  - INFN-PISA (sj)
- **LHCb**
  - **INFN-T1 (CNAF)**

- **File Transfer Service (FTS)** e' un servizio per il trasferimento di dati
  - E' un servizio **multi esperimento**, utilizzato per bilanciare l'uso delle risorse di un sito in accordo con le policy del sito stesso e delle **Virtual Organization (VO)** che supporta;
- **Perche' serve?**
  - Lato **utenti**, il servizio permette il trasferimento dei file da una sorgente ad una destinazione in modo affidabile;
  - Lato **site manager**, fornisce una soluzione affidabile ed elastica per soddisfare le richieste di trasferimento dati provenienti dagli esperimenti supportati;
  - Lato **production manager**, permette di controllare le richieste dei loro utenti (priorita', re-schedule transfer, ....)
  - **Il focus e' nel "servizio":**
    - Deve permettere queste operazioni in maniera semplice e trasparente

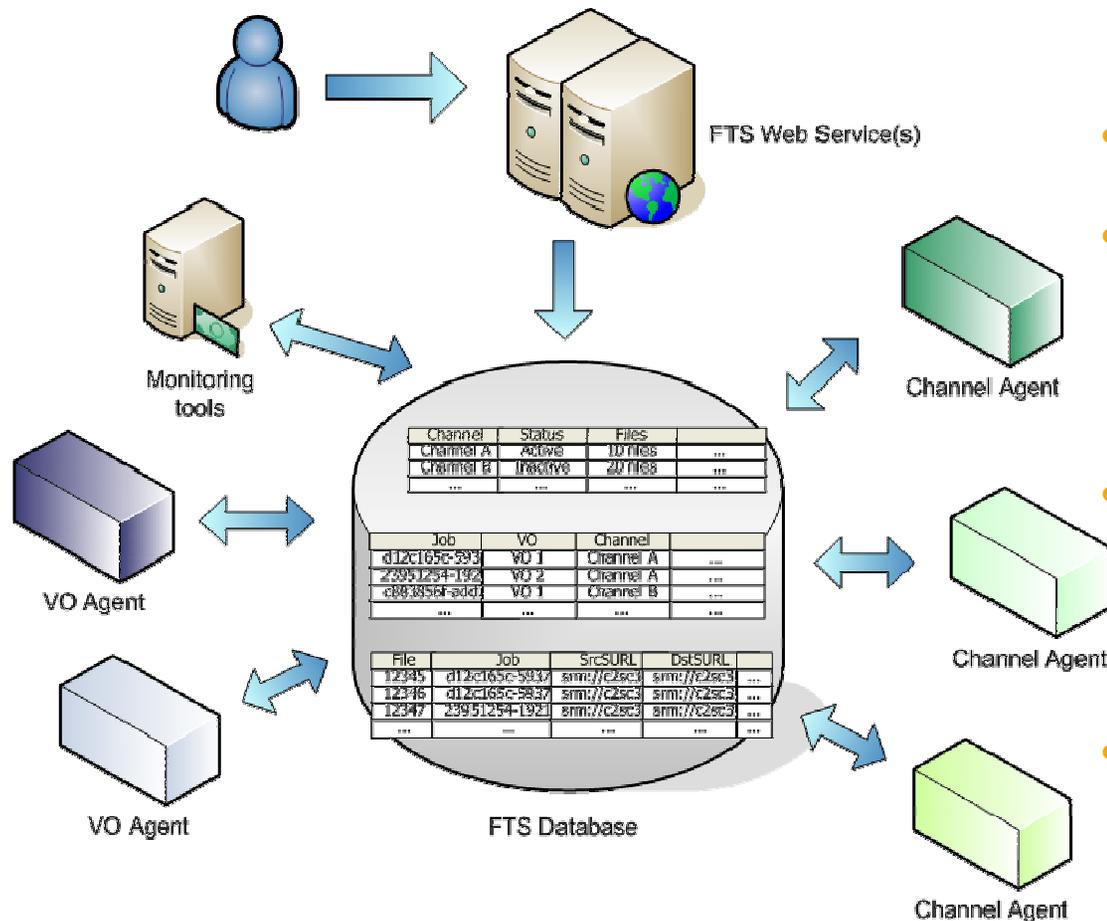
- **“CANALE”**: non e' un nome corretto 😊
  - Non si intende un percorso di rete ;
  - E' solo definito in maniera concettuale (nulla vieta di avere canali che si mappano su link fisici);
  - **“Coda”** sarebbe comunque meglio 😊
- **Un canale puo' essere:**
  - Una connessione punto a punto (sorgente -> destinazione)
  - Canali “catch-all” (e.g. Mondo->me, me->Mondo)
  - Una definizione piu' flessibile dei canali e' in via di sviluppo.
- **I canali sono uni direzionali**
- **I job di trasferimento vengono serviti secondo diverse prioritá':**
  - Priorita' Intra-VO per canale (Atlas 75%, CMS 25%)
- **Ogni canale ha il proprio set di parametri da configurare:**
  - Numero di trasferimenti concorrenti, numero di streams, buffer TCP, numero di risottomissioni in caso di fallimenti, etc.
- **Il servizio puo' dividere un job su piu' canali;**

- **FTS e' utilizzato dai frameworks degli esperimenti**
  - Di solito un utente finale non utilizza direttamente il servizio, ma attraverso il framework del proprio esperimento;
  - Production managers possono fare query dirette per debug e test
- **Esempio:**
  1. Il framework degli esperimenti e' responsabile dello staging-in dei dati;
  2. Vengono definiti dei set di sorgenti/destinazioni per un determinato gruppo di dati e sottomessi job di trasferimento a FTS;
  3. Lo stato di ciascun job viene tracciato durante le varie fasi del trasferimento;
  4. Il framework dell'esperimento puo' registrare lo stato dei job in ogni momento

## CMS

- TUTTI i transfers T0-T1 sono fatti con PhEDEx/FTS;
- La maggioranza dei trasferimenti T1-T2 sono fatti con PhEDEx/FTS

# Architettura di FTS



- **Web-service** come front-end
- **VO agents** si occupano di operazioni VO-specifiche (1 per VO)
- **Channel agents** si occupano di operazioni specifiche del canale (1 per canale)
- **Oracle** come backend DB.

# Configurazione FTS a INFN-T1

- Non solo LHC VOs
  - PAMELA
  - VIRGO
- 2 server dedicati in produzione (uno idle) + 1 server in pre-produzione
  - Problemi di cpu load con il DB sullo stesso;
  - FTS DB installato su un server dedicato (verrà installato su un cluster oracle)
- Matrice completa di canali T1<->T1 + T1<->T2 + STAR channels (39 canali in tutto);
- (Prototipo) Monitoring tool disponibile:
  - Web interface: <http://argus.cnaf.infn.it/fts/index-FTS.php>

# FTS monitoring tool

- DN utente recuperato dai log del web service;
- Tutte le altre informazioni recuperate dai log degli agenti dei canali e delle VO;
- Auto discovery di nuove VO e nuovi canali;
- UP & RUNNING da 5 mesi
  - ~ 70'000 job collezionati
- **Informazioni disponibile per job:**
  - – Canale
  - – VO
  - – DN Utente
  - – SRM Sorgente e destinazione
  - – FILE Sorgente e destinazione
  - – stato del job
  - – Inizio trasferimento
  - – Termine trasferimento
  - – Byte trasferiti
  - – Durata del job
  - – Rate
  - – Error code & Error Message
  - FAIR-SHARE
  - # job concorrenti
  - # di streams

Job details

Choose your preferred channel/VO:

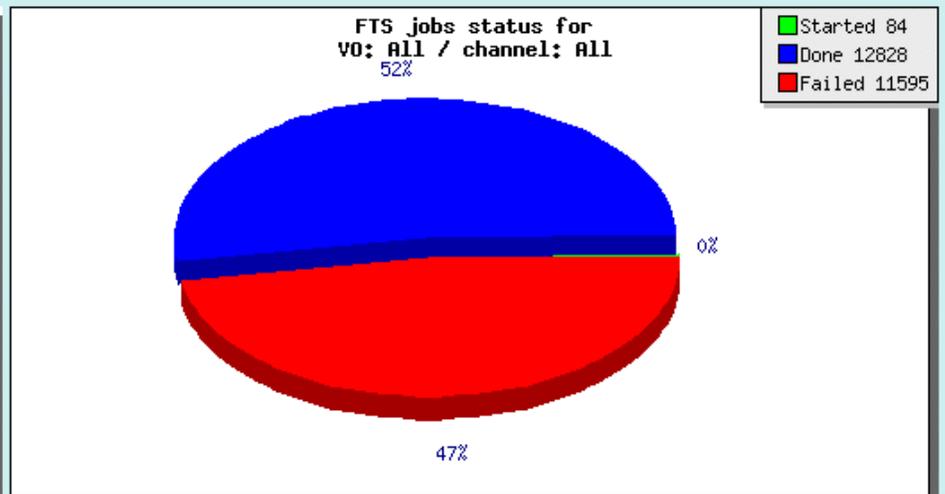
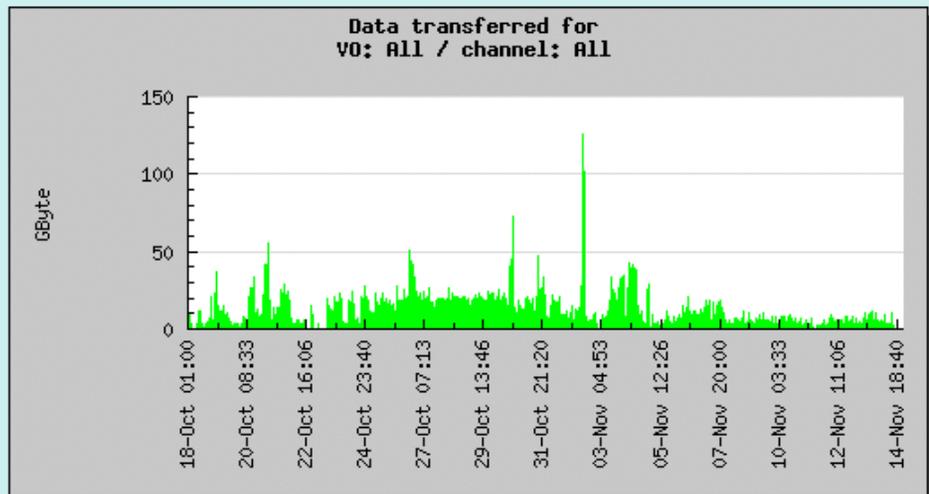
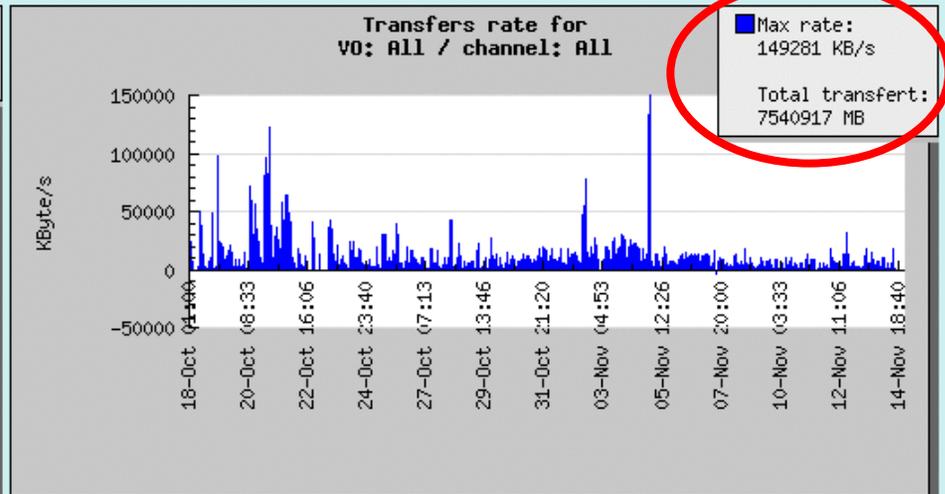
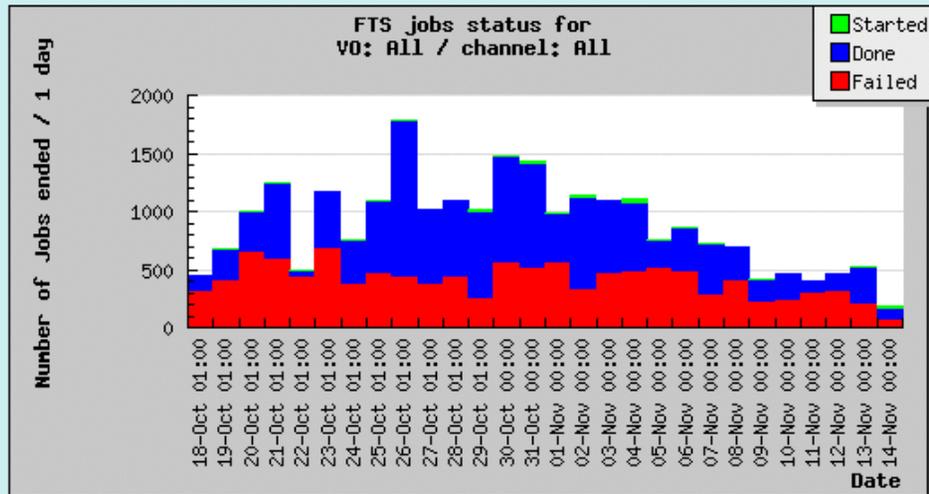
Channel: All VO: All

Choose your preferred end date:

November 14, 2006 for: 28 day Draw

**Failed = Aborted +  
Cancelled by users +  
Successfully Failed**

VO: All / channel: All



# FTS Monitoring tool

## dettagli del job

### Statistics graph

Choose your preferred channel/VO:

Channel:  VO:  Status:  Job user:

Choose your preferred end date:

for:  day

JOB ID	START TIME	END TIME	JOB BYTE (KB)	CHANNEL	JOB USER	STATUS	SOURCE_SRM
47fc6b26-7359-11db-b7ca-df6a401812b0	14 Nov 12:07	14 Nov 01:07	0	CMS		FAILED	CMSSRM.FNAL.GOV / 11 / STORE / CSA06 / CSA06-105-OS-ZMUMU-0 / RECO / CMSSW_1_0_5-RECO-H746EE88EDDAA52306CD016B2 1021 / 5215D12A-1D61-DB11-AC0C-00304885AE8C.ROOT
5ab50b7c-736f-11db-b7ca-df6a401812b0	14 Nov 12:33	14 Nov 12:33	0	DTEAM	Piotr Nyczyk 9654	FAILED	CASTORSC.GRID.SINICA.EDU.TW / CASTOR / GRID.SINICA.EDU.TW / SC / DTEAM / GENERATED / 2006_09_22 / FILE7C2A5905-0425-439B-A2DE-F15E339DF729 SRM_SOURCE: Failed on SRM get: Failed SRM get on http://castor.sc.grid.sinica.edu.tw:8443/srm/manager1 ; call. Error is specified file(s) does not exist

- Non e' proposta nessuna configurazione firewall automatica dagli script di installazione e/o configurazione;
- Se un sito e' protetto da firewall, occorre comunicare agli amministratori di rete di aprire alcuni "buchi" in entrata e/o uscita a seconda dei servizi.
- Una mappa completa delle porte utilizzate, suddivise per servizi, e' disponibile nel seguente URL:  
<http://j.home.cern.ch/j/jwhite/www/lcg-port-table.xls>

# Esempio di porte utilizzate

Node	Service	From		To	
		src	port (tcp)	dest	port (tcp)
SE					
	GridFTP Control	*{UI,SE,CE,WN}	C	localhost	2811
	GridFTP data (single channel)	*{UI,SE,CE,WN}	C	localhost	C
	GridFTP data (multiple channel)	*{UI,SE,CE,WN}	C	localhost	C
	SRM - httpd(apache)	*	*	localhost	80
C	<p>Controlable Ephemeral range (e.g. 20000-25000). Note: In practice, although this port-range is locally configurable using the GLOBUS_TCP_PORT_RANGE environment variable, <b>the values applying at a remote service cannot be predicted.</b> Consequently <b>reliable connection can only be established if all ports &gt;1023 are left open for outbound connections.</b></p>				

# References

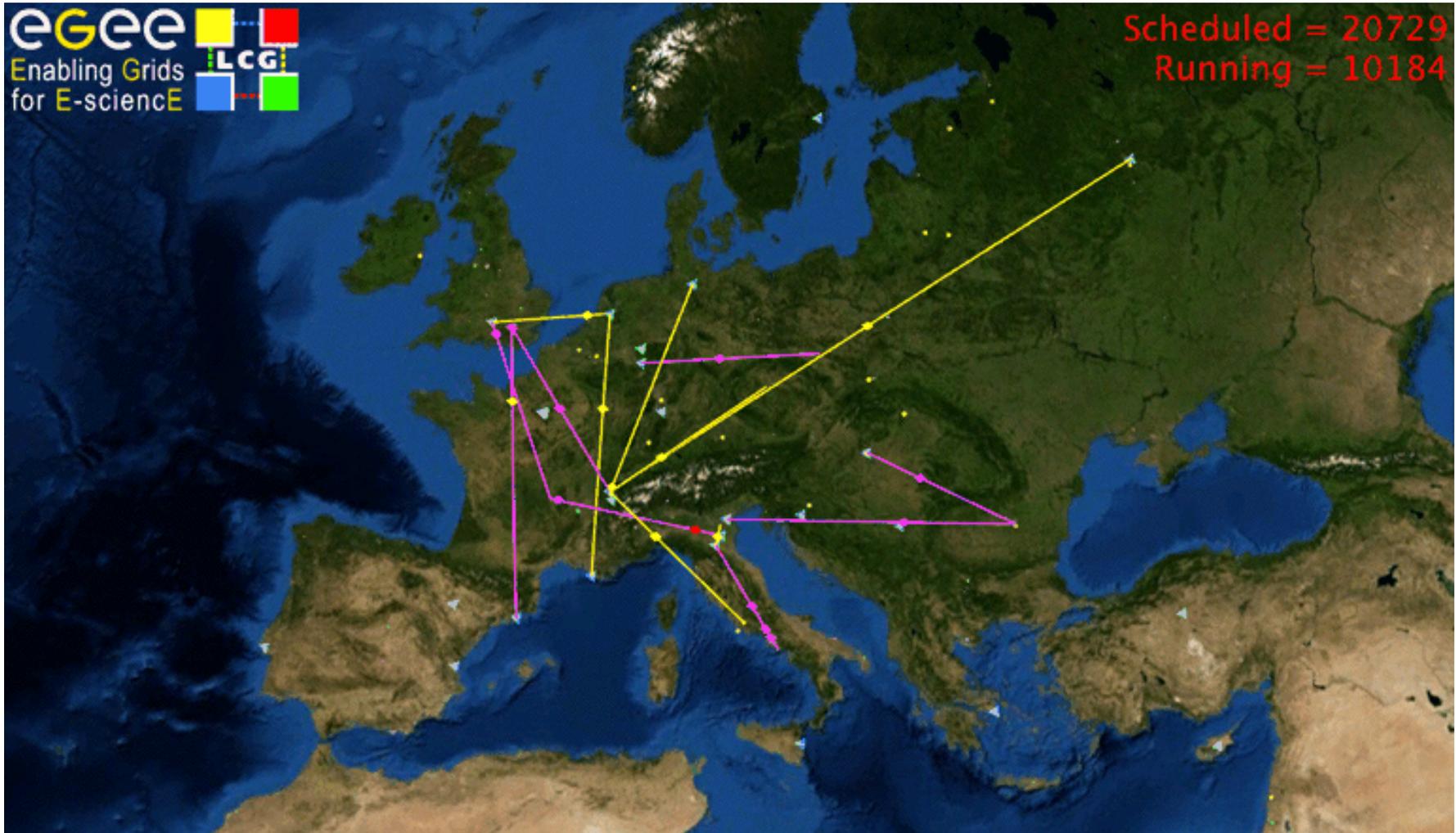
- Grid.IT – Italian Grid Infrastructure: <http://grid.infn.it/>
- EGEE Project: <http://www.eu-egee.org/>
- LCG - LHC Computing Grid Project <http://www.cern.ch/lcg>
- LHCb Computing Model (LHCC Review Report, 18/01/2005),  
CERN/LHCC-2004-036 LHCC-G-084
- CMS Computing Model, CERN-LHCC-2004-035 and CERN-LHCC-  
2005-023
- FTS Wiki page: <https://uimon.cern.ch/twiki/bin/view/LCG/FtsRelease15>

# Grid in action

**eGEE**  
Enabling Grids for  
E-science in Europe



Scheduled = 20729  
Running = 10184



**Tier-1 Centers: TRIUMF (Canada); GridKA (Germany); IN2P3 (France); CNAF (Italy); SARA/NIKHEF (NL); Nordic Data Grid Facility (NDGF); ASCC (Taipei); RAL (UK); BNL (US); FNAL (US); PIC (Spain)**