



Tier1, Tier2 e Tier3

CMS Italia



Modello di Calcolo per CMS Italia

Il Modello di CMS e' basato su una Gerarchia di Funzioni e Risorse (Tier-n). Tale Gerarchia e' "integrata" (e quindi resa piu' flessibile) dai Tools di GRID.

◆Le *Attivita'* che devono essere soddisfatte dal Modello sono catalogabili in:

- Attivita' schedulate (processing, re-processing, MC di larghe dimensioni, selezioni comuni di dati, etc.)
- Attivita' "caotiche" tipiche del processo di Analisi ("selezione / ricostruzione" dei dataset per Analysis Group, analisi degli individui, MC "mirati", etc)

◆L'analisi e' complessa e diversa da come oggi la si intende

- Il processo di analisi ad LHC passa attraverso la ripetitiva ricostruzione (DPD, AOD, ESD) dei dati di parti dell'Apparato.
- La ricostruzione richiede la specifica conoscenza sia del software di quelle parti di apparato che della fisica che si vuole analizzare, oltre ad una padronanza "tecnica" dei tools usati a vari livelli (Databases, OS, environment, etc.).
- Solo alla fine di ognuno di questi processi ripetitivi si ottengono grafici e plot di "tipo tradizionale"



Realtà del Modello

◆ Per rendere possibili le Attività citate (in particolare la seconda, l'Analisi) occorre disporre di Personale MOTIVATO:

- Dai risultati scientifici ottenibili da CMS
- Dalle attività di “Computing” Locali e Non, legate a CMS
- Se non è motivato dalla curiosità scientifica legata ai risultati di CMS, non acquista le necessarie competenze (Software, Detector, Tools)
 - La “curiosità” può essere più verso la “Fisica” o il “Computing”, ma comunque legata ai successi di CMS

◆ Competenze differenziate e complementari nelle Sedi, anche “alla MONARC”, ovvero diverse funzionalità (e quindi anche risorse)

➤ Flessibilità del Modello: Tier1, Tier2 e Tier3

◆ Il Modello (non solo questo di CMS) è realizzabile solo se si considerano le *funzioni ed le attività* dei Tier1, Tier2 e Tier3 insieme



Tier1 unico al CNAF per CMS Italia

- ◆ **Il Tier1 italiano (CMS o comune) fa parte del “Computing Detector” di CMS**
- ◆ **Il Tier1 di CMS Italia e’ la Base per le attivita’ “schedulate” dalla Collaborazione CMS**
 - Supporto per lo storage dei RAW data
 - Supporto per la produzione degli ESD da re-processing coordinato dalla Collaborazione e/o dai Physics Groups
 - Controllo dei Database distribuiti in tutta la Collaborazione
 - Supporto per le produzioni di grandi Datasamples (anche AOD e/o simulazioni)
 - Parziale supporto sistemistico per gli Utenti CMS
 - Coordinazione “tecnica” verso gli altri Centri di CMS (ovviamente insieme ai responsabili di CMS Italia)
 - Supporto e sviluppo di “alcuni” tools di Middleware (GRID) di interesse comune anche di CMS
 - Etc.
- ◆ **Il Tier1 unico al CNAF si puo’ inserire nel Modello di CMS con queste attivita’ se complementate da quelle dei Tier2 e Tier3**



Tier2 e Tier3 per CMS Italia

◆ **Il personale dei Tier2 e Tier3 e' direttamente interessato ai risultati di Fisica di CMS e "controlla" le risorse direttamente**

◆ **I Tier2 e Tier3 italiani sono pertanto la sede primaria delle attivita' di ANALISI ("caotiche")**

- Funzioni specifiche per il Computing di CMS (Core software, supporto del software specifico, database, middleware, tools di Analisi, etc.)
- Supporto agli Utenti distribuiti e Specifiche responsabilita' di "packages"
- Storage degli ESD necessari alle attivita' di Analisi supportate in loco
- Selezione e (ri)produzione degli AOD, DPD ed n-tuple per l'Analisi
- Responsabilita' di coordinamento scientifico verso la Collaborazione CMS
- Capacita' di gestione dei Database in modo consistente con la Collaborazione
- Capacita' di supporto tecnico, sistemistico e di analisi fisica per le attivita' collegate al sito
- Capacita' di coordinamento con gli altri Tier2 coinvolti nelle stesse analisi (italiani e non), nonche' con il Tier1 di riferimento

◆ **I Tier3 sono sostanzialmente dedicati ad Analisi "monotematiche"**

- Storage degli AOD necessari e produzione / rifelezione degli DPD ed n-tuple.



Il Modello di CMS Funziona

◆ **Un prototipo del Modello di CMS e' funzionante, e produce i risultati aspettati per gli studi di HLT e di Fisica: unico sistema "distribuito" in tutta CMS (CERN, Caltech, FNAL)**

▪ **Da almeno due anni CMS Italia ha responsabilita' ed ottiene successi in:**

- Produzioni, Simulazioni ed Analisi
- Tools di GRID
- Sviluppo di Software specifico di Detector
- Sviluppo di Software "Core"
- Programmazione del Computing di CMS

◆ **La time schedule dei prototipi deve tenere conto delle attivita' "reali" di "produzione di CMS (es. HLT e Physics TDR)**

◆ **Le Sedi di CMS Italia hanno gia' personale coinvolto e sono committed ad almeno mantenerlo:**

	LNL	Ba	Bo	Pd	Pi	Rm1	Ct	Fi	Pg	Pv	To
FTEs	~3	~2.5	~2.5	~3.5	~2.5	~2.0	~1.0	~0.5	~1.0	~0.5	~1.5



Dimensioni dei Tier-n italiani di CMS

- ◆ **Le dimensioni delle risorse dei Tier-n di CMS in Italia sono quelle espresse nei vari documenti di Review (per i quali vale la regola 1/3, 2/3).**
- ◆ **Le risorse necessarie (senza EFFICIENZA d'uso, Network, Infrastruttura e Personale)**
 - Tier1 (per CMS) di ~100kSI95 e ~1000 TB storage (dischi e nastri)
 - Tier2 da ~30 kSI95 a ~80 kSI95, e da ~400 TB a ~120 TB storage (dischi, possibilmente senza nastri)
 - Tier3 ~5 kSI95 e ~20 TB storage (dischi)
- ◆ **Commitment preliminare di Tier1, Tier2 e Tier3:**
 - Tier1 (comune): CNAF
 - Tier2: LNL+Pd, Ba, Bo, Pi, Rm1
 - Tier3: Ct, Fi, Pg, Pv, To
- ◆ **Lo sharing delle risorse pertanto:**
 - **Storage:**
 - ➔ 50% nel Tier1, 50% distribuito nei Tier2/3
 - **CPU:**
 - ➔ 35% nel Tier1, 65% distribuito nei Tier2/3



Sedi multi-Tier

◆ **Possibili sinergie di Infrastrutture e Personale sono possibili, ma programmabili solo se i siti sono definiti al piu' presto (long term commitments delle Sedi)**

- **CNAF: Tier1 CMS, Tier1 Alice, Tier1 Atlas, Tier1 LHCb, ...?**
- **Legnaro: Tier2 CMS, Tier? Alice**
- **Bari: Tier2 CMS, Tier2 Alice**
- **Bologna: Tier2 CMS, Tier2 Alice, Tier3 LHCb**
- **Padova: Tier2 CMS, Tier2 Alice**
- **Pisa: Tier2 CMS, Tier3 Atlas**
- **Roma1: Tier2 CMS, Tier2 Atlas, Tier? Alice, Tier? LHCb**
- **Catania: Tier3 CMS, Tier2 Alice**
- **Firenze: Tier3 CMS, Tier? LHCb**
- **Pavia: Tier3 CMS, Tier3 Atlas**
- **Perugia: Tier3 CMS, ?**
- **Torino: Tier3 CMS, Tier2 Alice**



Gestione del Tier1 unico

◆ Il Tier1 unico non sara' Operativo come prototipo prima di un paio di anni (Risorse, Infrastruttura e Personale)

- Occorre pero' un Centro che assuma il ruolo di "BOOT" per il Tier1: CNAF
- Occorre definire attraverso quale processo vengono definiti gli impegni nel (i)MoU del Computing per il Tier1 dell'INFN (e quindi anche per CMS),

◆ E' necessario definire da subito:

- Milestones del Progetto
- Comitato di "Referee" degli Esperimenti
- Comitato di gestione che risolva conflitti d'uso: del Prototipo oggi, del Centro domani
- Programma di validazione con Applicazioni di Esperimento (legato alle milestones)
- Quali competenze specifiche di Esperimento sono necessarie nel Tier1 per le attivita' e funzionalita' devolute al Centro

◆ La prototipizzazione del Tier1 deve essere integrata il piu' presto possibile con la prototipizzazione dei Tier2/3 che stanno realizzando le Produzioni degli Esperimenti e con lo sviluppo di GRID.



Conclusioni

Nello scenario dove il Tier1 non e' piu' solo per CMS e quindi direttamente di responsabilita'/controllo di CMS:

◆Cambiano poche cose, considerata la Flessibilita' e dinamicita' del Modello di CMS

- Il Tier1 e' la sede delle attivita' (servizi) "Collaboration scheduled" e "Collaboration coordinated" (e.g. Re-processing dai RAW data, MC di larghe dimensioni e di comune interesse, repository dei dati, interfaccia tecnica verso gli altri Tier1 e Tier0, etc.)
- I Tier2 sono la sede primaria delle attivita' di Analisi e processamento dei dati per ottenere i Risultati di Fisica e quindi la naturale interfaccia di CMS Italia verso il resto della Collaborazione
- I Tier3 hanno un ruolo un po' piu' importante, dovendosi assumere alcune funzionalita' e responsabilita' anche a livello della intera Collaborazione

◆Occorrono tuttavia garanzie:

- Sulla immediata e continuativa realizzazione dei Tier2
- Sulla immediata definizione di un Centro "boot" del Tier1, che definisca Milestones precise del progetto
- Sulla nomina di un Comitato "degli Esperimenti" che segua la realizzazione del prototipo di Tier1