

Dal progetto DHTCS-IT a un servizio per INAF ?

Fabio Pasian – INAF-OATs

Workshop Calcolo INAF, Monte Mario, 21 giugno 2016

Il progetto DHTCS-IT

Distributed High Throughput Computing and Storage in Italy

« Il progetto ha l'obiettivo di consolidare e far evolvere la attività che è stata realizzata nell'ambito di IGI in un quadro di portata internazionale. Il progetto intende, inoltre, riconfigurare sul panorama nazionale le attività nel campo del calcolo distribuito e agire da "collante" tra le varie iniziative esistenti (o pianificate) previste per specifici ambiti territoriali o per comunità di utilizzatori definiti, come i Progetti PON, i Progetti POR, i progetti Smart Cities e Cluster Tecnologici e nel campo internazionale i progetti ESFRI. » (dalla proposta, Nov 2012)

Obiettivi iniziali di DHTCS-IT

- Potenziare e far evolvere l'infrastruttura IGI
- Fornire servizi Grid e Cloud ai gruppi di ricerca dei soggetti partecipanti
- Supportare in modo adeguato, dal punto di vista infrastrutturale, i progetti ESFRI dei soggetti partecipanti
- Coordinamento e integrazione con altri progetti (es. GARR-X Progress)

DHTCS-IT: nuovi servizi

- DHTCS-IT come evoluzione/estensione dei servizi offerti da IGI (Grid → Cloud)
- Evoluzione verso servizi cloud open integrati in Grid o gestiti in modo indipendente/complementare
- Rafforzamento della vocazione multidisciplinare dell'infrastruttura

Finanziamento di DHTCS-IT

- Gestito da INFN, con partecipazione CNR, INAF, INGV e GARR (tramite INFN)
- Finanziamento MIUR 2013-14-15 come attività di ricerca a valenza internazionale (es. al pari di SKA, CTA, E-ELT e SRT per INAF)
- Totale su 3 anni: ~ 5.600.000 €
 - per INAF: ~ 560.000 €

DHTCS-IT @ INAF

Acquisito un cluster (hotCat)

- 800 core (HP DL560 G9)
- Infiniband 40 Gbps
- RAM 16 GB/core
- 300 TB HDD raw
- 1 CTER TD per il supporto
- logistica adeguata (freecooling,...)
- in corso di completamento e sviluppo
- da integrare con l'infrastruttura Italiana secondo regole TBD



Da DHTCS-IT a un servizio per INAF?



Use case per T-2/T-3

- progettazione/test di software di analisi dati e prototipi per grandi progetti;
- sviluppo e test di codici;
- studi di fase A di progetti;
- progettazione di sistemi innovativi;
- esigenze della comunità scientifica con minori necessità di calcolo;
- data processing di livello intermedio;
- preparazione di software per simulazioni HPC;
- piccoli run di tipo HPC;
- etc.

Un sistema di classe T-2 per INAF

- Nel piano triennale INAF 2015-17 viene esplicitamente menzionata la possibilità per INAF di dotarsi di un sistema di classe Tier-2
- Il costo di un tale sistema è attualmente dell'ordine di 1.5 M€, con una spesa di mantenimento (staff compreso) di circa 300 K€/anno.

Procurement/Sostenibilità/Know-how

- **Procurement** – vanno perseguite possibilità di finanziamento esterno per acquisire il nucleo del sistema, che INAF può integrare.
- **Sostenibilità** – è necessario elaborare un modello di costo che distribuisca le spese in modo opportuno sulla base delle statistiche di utilizzo.
- **Know-how/Networking** – manutenzione, sviluppo e supporto richiedono personale dedicato e una conoscenza distribuita a livello INAF.

Uno o più Centri di computing

- Differenti modalità di reperimento fondi → è possibile il procurement di più sistemi → centri
 - e-infrastructure distribuita ma sinergica, con la stessa tipologia di servizi e sw, politica di servizio e di sviluppo comune
 - di concerto con gli Uffici ICT e Science Data Management della Direzione Scientifica,
 - coordinata con il Centro Archivi dell'INAF - IA2,
 - coerente con le politiche nazionali e internazionali sul calcolo, e in particolare con i progetti di sviluppo tecnologico nel campo finanziati a livello nazionale e/o internazionale.
 - coordinati, tramite INAF DS, con la componente italiana dei progetti per le grandi infrastrutture osservative ESFRI ed ESA Cosmic Vision.
 - centri di spesa di terzo livello (?)

Proposta per la fase transitoria

- I fondi DHTCS-IT sono $> 1/3$ del costo di un T-2
- hotCat (800 \rightarrow 1000 core) come seme di T-2
- Proposta: hotCat + altri 1 o 2 sistemi (da > 200 core) come **primo centro INAF** per il calcolo, che rispetti i principi espressi precedentemente
- Scopo: **servizio** ad alta disponibilità per l'utenza
- **Committment** per mantenimento/sviluppo c/o **INAF** sulla base di un meccanismo che garantisca sostenibilità
- Sinergia con INFN e le infrastrutture di cloud italiana (DHTCS-IT) ed europea (EOSC), e con i progetti H2020 EGI-Engage, INDIGO-DataCloud e ASTERICS

Grazie per l'attenzione !
Domande ?