

INAF



HPC

La nuova convenzione con il Cineca

U. Becciani

**ICT Workshop
14-18 Novembre, 2016
Trieste**

INAF – Big Science

INAF ha raggiunto l'eccellenza scientifica in diversi settori: dalla Astrofisica Computazionale (modellizzazione teorica in cosmologia e astrofisica) ai grandi esperimenti e progetti.

Le *nuove frontiere* in Astrofisica richiedono e richiederanno l'uso massivo di infrastrutture di tipo HPC.

Le piattaforme HPC rappresentano oggi e nell'immediato futuro facilities di *fondamentale rilevanza* e sono necessarie oggi per mantenere l'eccellenza raggiunta.

INAF – HPC nel recente passato

INAF - CINECA convenzioni per l'HPC

Computational Astrophysics, Visualization, Archiving and Post-Processing Data Analysis

- 1997 - 2001
- 2001 - 2005
- 2005 - 2007
- 2008 - 2010

Nell'ultimo periodo importanti esperimenti hanno richiesto risorse di calcolo specifiche di tipo Tier-0 e Tier-1 per l'analisi dei dati sperimentali e per il confronto con i dati simulati (missione Planck), per cui sono stati presi specifici accordi con il Cineca

- 2010-2012
- 2012-2014

INAF – POR+PON

- POR 2005-2007 TRIGRID VL @OACT
- Cometa Consortium (PI2S2 project 2005-2009) (OACT+OAPA)
- Cosmolab Consortium (Cybersar project 2005-2009) (OACA)

INAF – POR+PON

- ➔ POR 2005-2007 TRIGRID VL @OACT + Cometa Consortium (PI2S2 project 2005-2009) (OACT+OAPA)
- ➔ Cosmolab Consortium (Cybersar project 2005-2009) (OACA)



The TriGrid Project & Cometa Consortium

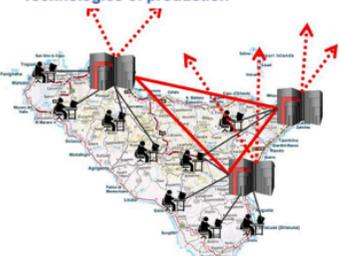
<http://www.pi2s2.it/applications>

(Some of) The Applications of TriGrid VL

- **Astrophysics**
 - Virtual Observatory
 - Visualizzazione 3D Visualization
 - The GAIA Mission
- **Biomedicine**
 - Analysis and classification of medical images (mammograms, etc.)
- **High Energy Physics**
 - Data analysis of CERN's Large Hadron Collider (LHC) Experiments
 - Theoretical Physics
- **Engineering**
- **Cultural heritage**
- **Analysis and monitoring of environmental risks**
- **Technologies of production**



APPLICATION	DOMAIN	Details
COMP_FILTER	Astrophysics	Details
AstroDyn	Astrophysics	Details
VisVOserver	Astrophysics	Details
P-Gadget2	Astrophysics	Details
MultIDIRECT	Astrophysics	Details
ADAM	Astrophysics	Details
ZEUSMP	Astrophysics	Details
PLUTO	Astrophysics	Details
lc_par	Astrophysics	Details
GFLASH	Astrophysics	Details
HEASOFT	Astrophysics	Details
GEX	Astrophysics	Details
FLY	Astrophysics	Details
JPLASH	Astrophysics	Details
CORSIKA	Astrophysics	Details
ClustalW	Bioinformatics	Details
SpitzTree	Bioinformatics	Details
GeneTrinder	Bioinformatics	Details
Paratree	Paratree	Details



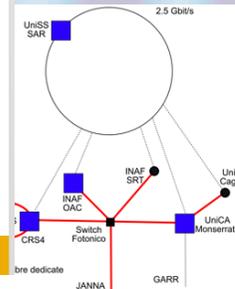
1. 1500+ cores AMD Opteron IBM Blade
2. 250 TB of memory
3. LSF 6.1 HPC everywhere
4. Infiniband-1X for HPC apps.

~15.000.000 € in 3 years!
~300 FTE's ! (2/3 new hired staff)

Cybersar Project

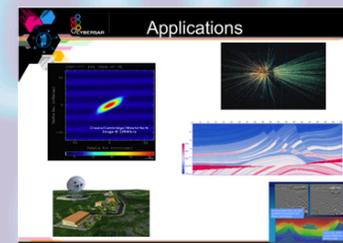


- 4 main sites
- Optical link (cold)- λ-Grid
- Optical Switching
- Bandwidth Unlimited Computing
- Application Programmable Computing Resources



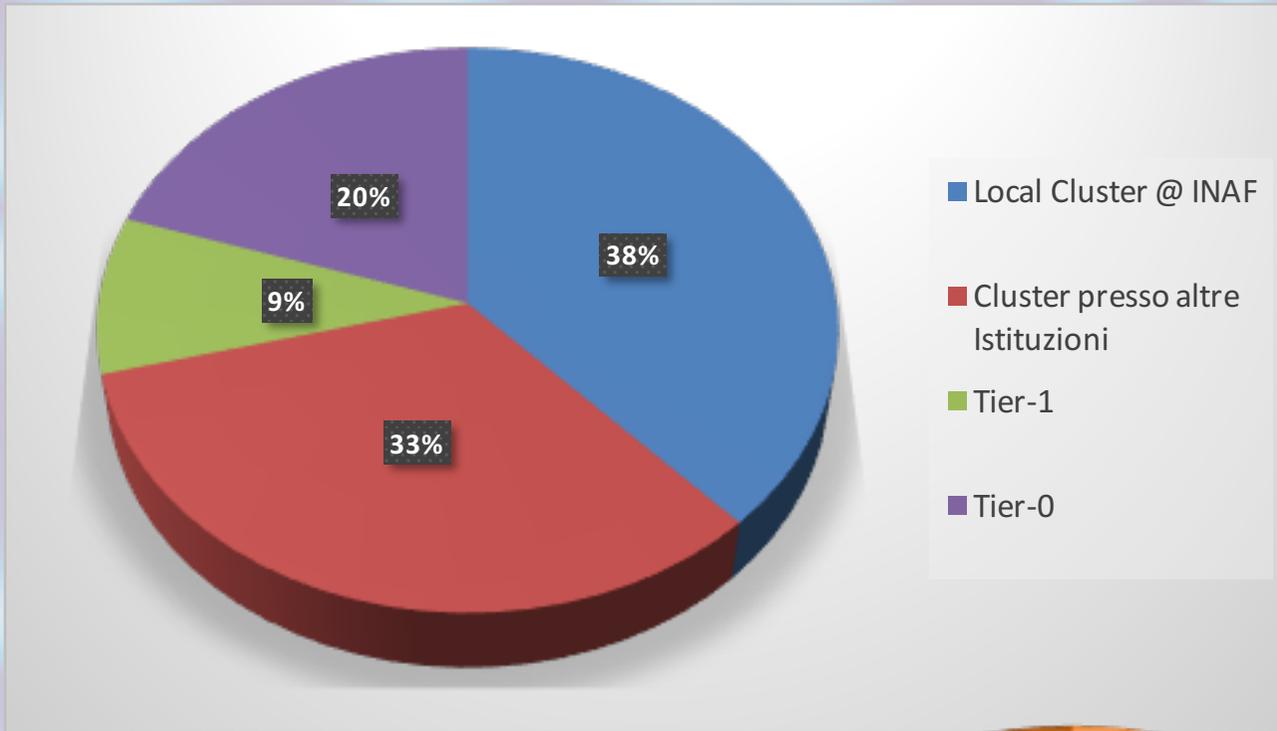
	U-CA	CRS4	OAC	Tot
Cpu cores†	408	312	288	1,008
RAM (TB)	0.8	0.8	0.6	2.2
Disk (TB)	70	80	45	195

+ AMD dual opteron 28x cpus
 •InfiniBand 4xDDR

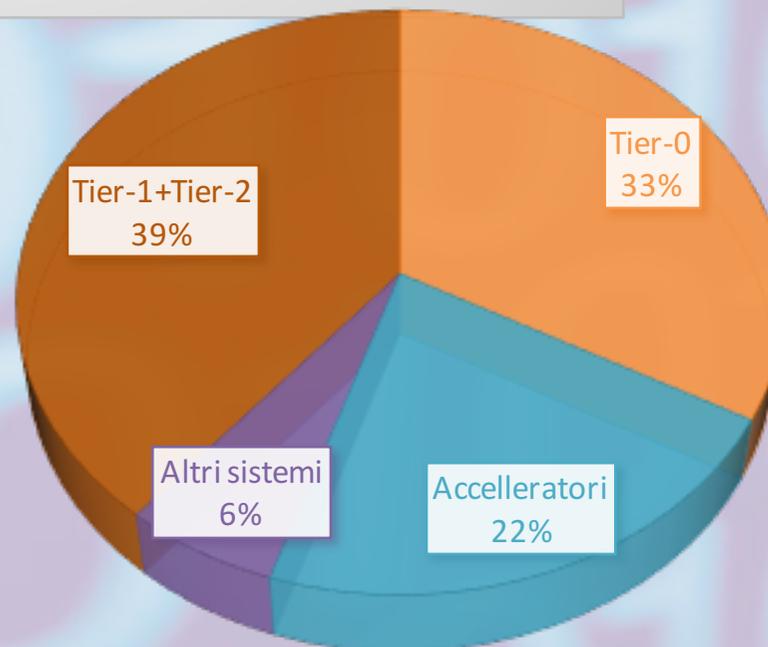


INAF HPC Survey

Stato Attuale



Necessità per il triennio 2016-2019



INAF HPC Survey

Ricercatori coinvolti: → **70 (staff & collaboratori)**

Osservatori/Istituti coinvolti: → **9 su 17**

Pubblicazioni peer reviewed (2012-2014): → **143**

Conference reports (2012-2014): → **129**

Risorse utilizzate in termini di CPU/core hours (2012-2014)

- → **255 Milioni, 68% Cluster (Tier1-Cineca) e IBM Fermi Cineca**
 - **MPI → 55%**
 - **MPI+OMP → 45%**
- CPU/core hours e storage richiesto
→ **550 Milioni e 710 TB disk space**

Supporto Richiesto

65% porting activity per l'uso dei sistemi HPC di nuova generazione

Necessità confermata e discussa

- **ICT Workshop: What about computing @ INAF ? INAF Headquarter
Giugno 20-21, 2016**
- **Meeting delle Macroaree – Giugno/Settembre 2016**

- In 2015 the computing resources in Cineca were:
 - Tier-0: **FERMI** (acquired in summer 2012)
 - Tier-1: **GALILEO** (acquired in Jan 2015)
 - Front-end, Viz, BigData: **PICO** (acquired in Nov 2014)
- FERMI arrived at the expected end of its activity.
- The Cineca governing bodies, aimed at supporting scientific research, approved a development plan, with an investment of Euro 50 million in two phases, from 2016 to 2020:
 - 2 x 5 → 10 Pflops in 2016-2017
 - 10 x 5 → 50 Pflops in 2019-2020

 **Marconi**

A tender was issued in 2015 and assigned Jan 2016 to **lenovo**

The system will be delivered in three phases:

- A1: April 2016 (BRD 2 PFs)
- A2: Sept 2016 (KNL 11 PFs)
- A3: July 2017 (SKL 5 PFs)

In total:

- 18 PFs peak performance
- 17PB row storage

Peak Perf.	Comp. Nodes	Socket	RAM/CN	Interconnect	Rack #	Service & Mgmt nodes
A1 (half reserved to EUROfusion)						
2PFs	1512	2x Intel Broadwell 18cores @2.3GHz	128 GB	Intel OmniPath 2:1 100Gb/s	21	10 Front End Nodes (2xBDW 18c +128GB RAM)+ 2 MGMT nodes

Core tot: 54.432
Core-h/anno=476.824.320

Peak Perf.	Comp. Nodes	Socket	RAM/CN	Interconnect	Rack #	Service & Mgmt nodes
A2						
Core tot: 244.800 Core-h/anno=2.144.448.000						
11 PFs	3600	Intel KnightsLanding 68cores @1.4 GHz	96 GB	Intel OmniPath 2:1 100Gb/s	50	4 Front End Nodes (2xBDW 18c +128GB RAM)+ 48 I/O nodes

Peak Perf.	Comp. Nodes	Socket	RAM/CN	Interconnect	Rack #	Service & Mgmt nodes
A3 (great part reserved to EUROfusion)						
Core tot: 60.480 Core-h/anno=529.804.800						
5PFs	1512	2x Intel SkyLake 20cores @2.3GHz	192 GB	Intel OmniPath 2:1 100Gb/s	21	2 Front End Nodes + 2 MGMT nodes (2xSKL 20c +192GB RAM) (2xSKL 20cc +

MARCONI IL NUOVO SISTEMA TIER-0 DEL CINECA

Architecture: IBM BlueGene/Q
Model: 10 racks
Processor Type: IBM PowerA2, 1.6 GHz
Computing Cores: 163840
Computing Nodes: 10240, 16 core each
RAM: 16 GB/node, 1GB/core
Internal Network: custom with 11 links -> 5D Torus
Disk Space: 2.6 PB of scratch space
Peak Performance: 2PFlop/s

Fermi....
a fine Luglio è stato spento



GAIA: MoU INAF - Cineca

Phase 1

Aprile 2013 - Dicembre 2013

- Test preliminari del codice parallelo Solver Module, re-engineering e test phase.
- Definizione e test delle procedure di trasferimento file dal DPCT al Cineca
- Risorse disponibili: **5 Milioni CPU core hours** (senza scadenza)

Phase 2

Gennaio 2014 - Dicembre 2015

- Test phase (*cont.*)
- Inizio della fase di test/produzione su dati reali e nuove implementazioni **GSR-Cycles 1-4**
- Risorse Disponibili: **25-30 Milioni CPU core hours** (senza scadenza)

Phase 3

Gennaio 2016 - Febbraio 2021

- Nuovo sistema **Marconi KNL**
- Soluzioni per **GSR-Cycles 5-10**
- Risorse Disponibili Iniziali: **15 Milioni CPU core hours/anno per 3 anni** (senza scadenza)
- Oltre **50 TB Disk Space permanente**

GAIA: MoU INAF - Cineca

Others

- Annesso tecnico al MoU rinnovato
- **Technical support** dello staff Cineca
- Comitato di Gestione dell'accordo
- *High priority queues*

A Luglio 2016 abbiamo rinnovato il MoU per il triennio 2016-2018

Stiamo usando inizialmente il Marconi BRD, poi transitiamo sul KNL e faremo il re-engineering del codice per sfruttare tutte le potenzialità del KNL+AVX

Il Cineca ha privilegiato il rapporto con la missione Gaia (come Flagship Project) rinegoziando il monte ore assegnato ***da 5 Milioni a 15 Milioni*** di ore Marconi KNL/anno.

II NUOVO Mou/Framework INAF - Cineca. Punti Qualificanti

- **Necessità di adeguamento dei programmi esistenti e di nuovo sviluppo con l'utilizzo di sistemi Tier-0 di ultima generazione basati su sistemi di accelerazione per non perdere l'eccellenza raggiunta.**
- **Accordo quadro valido per 3 anni ma economicamente valido per un anno da rinnovare esplicitamente e a pari condizioni economiche per un massimo di 3 anni (ulteriori 2 rinnovi)**
- **Creazione di Framework per Flagship Project. I Flagship Project, ai fini di questo accordo quadro, sono indicati dalla DS**
- **Risorse di calcolo per totali 50 Milioni di ore cpu/core Marconi KNL ma utilizzabili su tutti i sistemi del Cineca (35Mh inizialmente previsti! → + 40%)**
- **Utilizzo di code a più alto privilegio per l'esecuzione di job di elevata rilevanza**
- **Ogni Flagship Project avrà riservato il 5% delle risorse di calcolo e di supporto del monte ore di cui sopra. Il Cineca si impegna a riconoscerne ulteriore 5% in modo aggiuntivo alle risorse già concordate. Massimo allocabile 20% delle risorse per INAF di cui sopra per i progetti Flagship.**

La proposta di MoU/Framework INAF - Cineca. Punti Qualificanti

- Comitato di gestione paritario di 6 persone: 4 Cineca e 2 INAF presieduto da INAF. Il **comitato** garantisce la corretta applicazione dell'accordo e l'utilizzo delle risorse in modo efficiente (anche con la richiesta di **High Priority** dei job in coda) ed adeguato in base alle necessità dei ricercatori.
 - **INAF People: U. Becciani (OACT / ICT), G. Murante (OATS), P. Rossi (OATO), G. Mulas (OACA)**
- **Assegno di ricerca di tipo PostDoc (bando INAF) annuale rinnovabile: posizione di software engineer con esperienza in HPC new generation. A disposizioni di tutta INAF per aiutare lo sviluppo e il porting dei nostri codici e far crescere le nostre competenze**
- Tale unità di personale **lavorerà presso il Cineca** che **il Cineca si impegna per il supporto logistico di tale unità di personale per tutta la durata dell'accordo quadro.**
- Il coordinamento di tale unità è dato dall'INAF nella figura del responsabile del Comitato di Gestione
- **Effettiva data di inizio dell'accordo dopo la presa di servizio del PostDoc**

Criteri di Valutazione

Valutazioni Scientifiche (Calcolo Teorico, Sperimentale e Flagship Project)

→ *Obiettivi Raggiunti vs Obiettivi Previsti*

➤ *Produttività*

- Quanti programmi supportiamo, risultati scientifici raggiunti e qualità e rilevanza

➤ *Outcome: Risultati che creano valore aggiunto*

- Collaborazioni, Formazione, Mobilità, Divulgazione Scientifica

→ **Supporto ai progetti Flagship: Sfide del futuro**

- Quali e quanti progetti Flagship hanno trovato beneficio
- Attività sperimentali avviate e loro rilevanza
- Porting di applicativi

Criteri di Valutazione

→ Valutazioni Tecnologiche

➤ **Percentuale di utilizzo**

→ Quote di distribuzione ai vari gruppi e loro effettivo utilizzo

→ Utilizzo “mensile”

➤ **Ricercatori coinvolti**

➤ **Codici supportati**

➤ **Adeguatezza delle risorse: cosa c'è e cosa manca**

→ Disponibilità effettiva delle risorse

→ Post Processing

→ Sistemi di calcolo e storage: performance e disponibilità effettive

REMARK: Probabilmente 1 solo anno di convenzione è troppo poco per una valutazione obiettiva

Opportunità

➤ **Rapporti con il Cineca**

- Coinvolgimento in attività di Ricerca con canali di Finanziamento
- Supporto per specifici progetti (per esigenze particolari/speciali, vedi Gaia)
- Agevolazioni nelle richieste di supporto sistemistico

➤ **H2020 e altri canali**

- Coinvolgimento nei progetti come utenti di che fanno uso sistematico con centri di HPC
- Specifici progetti di sviluppo di nuove architetture HPC

➤ **POR + PON**

- Utilizzo delle competenze acquisite
- Cineca come partner di supporto

➤ **Crescita di personale ICT con competenze tecnologiche nel campo HPC**

Comitati e comunità per la valutazione

- **Commissione Calcolo** (*Calcolo Teorico, Flgaship Project, MoU Gaia*)
- **ICT – Gruppo di coordinamento**
- **Comitato di Gestione della Convenzione**

SUMMARY

- **La convenzione nasce per porting/adeguamento dei nostri codici in un momento storico di cambiamento di paradigma in ambito HPC (acceleratori), ma ci sarà anche spazio per sostenere alcuni progetti con CPU/core hours.**
- **Aprire opportunità ai Flagship projects: avviare o raggiungere livelli di competenza necessarie per le sfide dei grandi progetti**
- **Ci offre la possibilità di poter interloquire con soggetti diversi con maggiore possibilità nelle sfide del prossimo futuro (programmi H2020, Nazionali e Internazionali in ambito ICT/HPC)**
- **Pubblicità: Commissione Calcolo, MA, Media INAF, Indico/ICT, conferenze ecc.**
- **TEMPISTICA**
 - **Cineca è pronto da subito alla firma**
 - **Firma attesa entro Novembre**
 - **Pubblicazione Bando per AdR per inizio Dicembre**
 - **Selezioni a Gennaio**
 - **Febbraio effettivo inizio**