



SOFTWARE SCIENTIFICO DISPONIBILE SUL CLUSTER DI CALCOLO DI IASF-BOLOGNA

INAF/IASF Internal Report n. 693/2015

Eleonora Torresi, Adriano De Rosa, Giuseppe Malaguti

1. INTRODUZIONE

IASF-Bologna è dotato di un centro di calcolo all'avanguardia necessario per le molteplici attività di ricerca che vengono svolte all'interno dell'Istituto. Tale cluster è accessibile dagli utenti dell'IASF facendo il login sui server 'front-end' *tonno* o *bitonno* con il proprio username e password. Su tale cluster sono installate librerie pubbliche come GCC, librerie MPI, librerie di calcolo scientifico o di accesso ai dati scientifici (come CFitsIO e HEALPix). Molti dei software pubblici per l'analisi dei dati scientifici sono correttamente installati sul cluster e pronti ad essere utilizzati dagli utenti. In questo modo ogni utente può ridurre e analizzare i dati provenienti da differenti satelliti senza dover necessariamente installare sulla propria macchina il software appropriato, ma semplicemente inizializzando le variabili d'ambiente per una specifica analisi. Questo permette di risparmiare tempo e garantisce l'utilizzo di un software, e dei relativi file di calibrazione, sempre aggiornati.

Il presente documento è strutturato come segue:

- in Sezione 2 viene presentato lo stato dell'arte del cluster di calcolo (architettura generale, caratteristiche tecniche e sistemi operativi dei nodi di login, modalità di accesso ai nodi);
- la Sezione 3 è dedicata al software scientifico presente sul cluster di calcolo. In particolare, i pacchetti di riduzione e analisi dati installati sono elencati e descritti insieme con l'ambiente di sviluppo del cluster.

2. CLUSTER DI CALCOLO

Di seguito verrà presentato lo stato dell'arte dell'infrastruttura hardware e software del centro di calcolo di IASF-Bologna. Esso è dotato di un cluster di calcolo dedicato all'analisi scientifica, allo *storage* dei dati e allo sviluppo software, e di un cluster di macchine virtuali dedicato all'implementazione dei servizi come la posta elettronica, il servizio di stampa (cups), un server di licenze, etc. Tali macchine virtuali sono relative a progetti e/o task specifici, in gestione al centro di calcolo.

2.1 Architettura generale del cluster di calcolo

Il cluster di calcolo è costituito da:

- 4 nodi di login:
 1. bitonno.iasfbo.inaf.it
 2. tonno.iasfbo.inaf.it

3. login01.iasfbo.inaf.it

4. login02.iasfbo.inaf.it

- 4 nodi di calcolo con 64 core ciascuno (gestiti da uno *scheduler* “IBM loadlever”);

- 4 nodi di *storage*.

I nodi di login si differenziano tra loro in base all’utilizzo che l’utente ne vuole fare, in modo da poter distribuire il carico di lavoro il più equamente possibile tra i nodi. Per esempio: **bitonno** è dedicato alla lettura delle mail, al desktop remoto, all’accesso ai filesystem; **tonno** oltre che alle attività di bitonno è dedicato alla compilazione e sottomissione dei job sui nodi di calcolo. **login01** e **login02** sono invece interamente dedicati alle sessioni grafiche (X su ssh), all’accesso ai filesystem, alla compilazione e sottomissione dei job sui nodi di calcolo.

2.2 Caratteristiche tecniche e sistemi operativi dei nodi di login

Nodo	Caratteristiche tecniche	Sistema operativo Linux (RedHat like)
tonno	4 core intel xeon, 16GB ram	CentOS 6.4
bitonno	4 core intel xeon, 16GB ram	CentOS 5.5
login01	8 core amd, 32GB ram	CentOS 6.4
login02	8 core amd, 32GB ram	CentOS 6.4

2.3 Modalità di accesso ai nodi di login

E’ possibile effettuare l’accesso ai nodi di login in tre modalità:

1) SSH senza X forwarding

ssh username@tonno.iasfbo.inaf.it

ssh username@bitonno.iasfbo.inaf.it

ssh username@login01.iasfbo.inaf.it

ssh username@login02.iasfbo.inaf.it

in questa modalità l’utente non può utilizzare la finestra grafica.

2) SSH con X forwarding

```
ssh -X username@tonno.iasfbo.inaf.it  
ssh -X username@bitonno.iasfbo.inaf.it  
ssh -X username@login01.iasfbo.inaf.it  
ssh -X username@login02.iasfbo.inaf.it
```

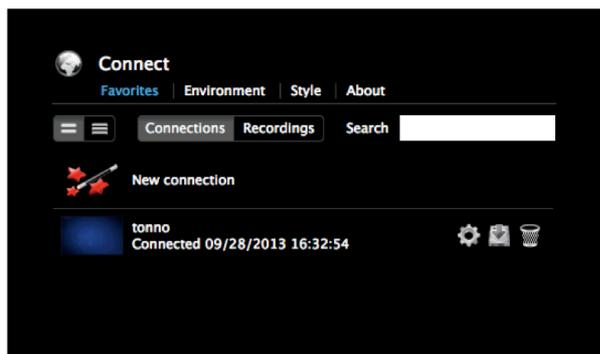
in questa modalità l'utente può utilizzare la finestra grafica.

3) Accesso da desktop remoto (solo su tonno e bitonno)

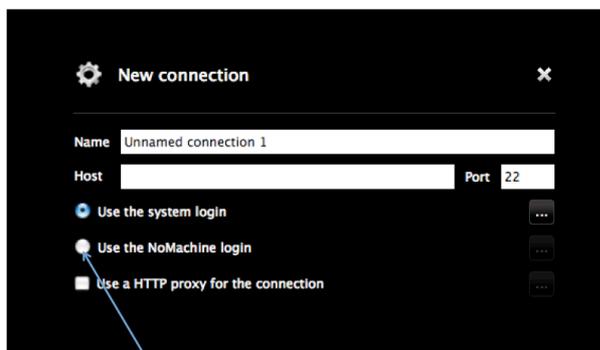
Questa tipologia di accesso è basata sul software open source FreeNX attraverso una connessione alla porta 22 (ssh).

Al fine di poter utilizzare questa modalità di accesso l'utente deve:

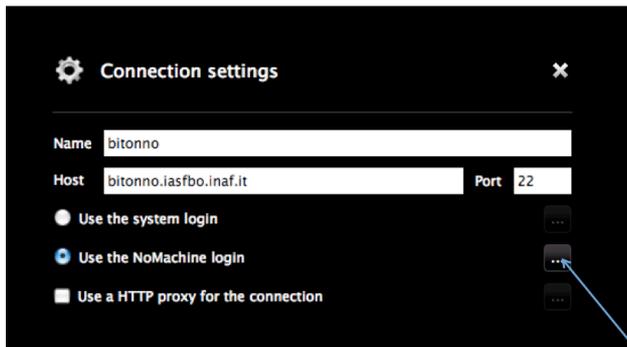
- scaricare il Player relativo al proprio sistema operativo dal sito <http://www.nomachine.com/download>
- installare il Player sul proprio pc
- eseguire NoMachine Player



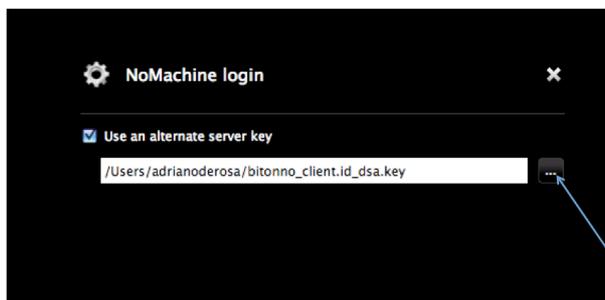
STEP 1



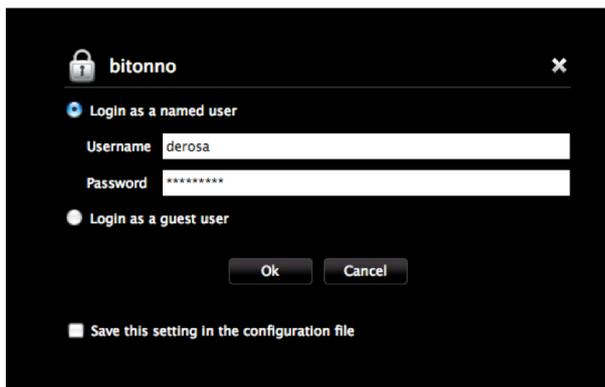
STEP 2



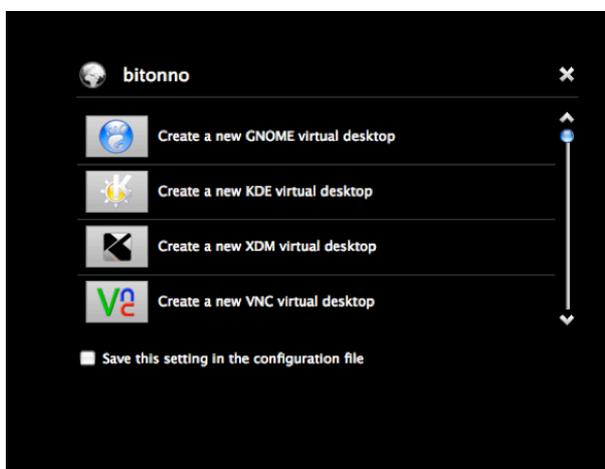
STEP 3



STEP 4 (è necessario collegarsi al server per scaricare la chiave in locale)



STEP 5



STEP 6

(è preferibile usare un desktop virtuale GNOME perché più leggero di KDE)

3. SOFTWARE SCIENTIFICO

IASF-Bologna è da sempre coinvolto in numerose missioni spaziali che operano nelle diverse bande dello spettro elettromagnetico. In generale, ogni missione ha il proprio pacchetto di riduzione dati e le proprie calibrazioni che devono essere aggiornate periodicamente per poter effettuare una corretta analisi dei dati scientifici.

L'implementazione del software scientifico installato sul cluster di calcolo, l'aggiornamento del portale web di Istituto relativamente a questo aspetto e il supporto agli utenti sono alcune tra le attività svolte dalla sottoscritta nell'ambito dell'Assegno di Ricerca *“Studio e realizzazione di strumenti informatici e di supporto generale per l'ottimizzazione dell'attività scientifica di Istituto, anche nell'ambito delle iniziative di didattica, divulgazione e public outreach”*.

Relativamente a tale attività di supporto, la sottoscritta ha contribuito all'implementazione del software scientifico in particolare, da maggio 2013, gli utenti di IASF-Bologna possono trovare installati, e correttamente funzionanti, molti dei pacchetti di riduzione dati di missioni spaziali per cui il software sia pubblico e scaricabile (insieme con i relativi file di calibrazione). Infine sul cluster possono essere presenti anche codici di calcolo e dati 'proprietary' a cui non tutti gli utenti hanno accesso. In questo modo viene garantita la riservatezza dei dati e del software.

L'utilizzo del software scientifico direttamente dal cluster ha il vantaggio per l'utente di non dover installare ogni pacchetto sul proprio pc, avendo però la certezza di lavorare con versioni sempre aggiornate e disporre di una macchina in generale più potente del proprio pc. Inoltre, ciascun utente può richiedere in qualunque momento l'installazione di un nuovo pacchetto scientifico.

Un'altra attività svolta dalla sottoscritta in questo ambito è stato il continuo aggiornamento della pagina web di istituto dedicata al 'Software Scientifico'

(<http://www.iasfbo.inaf.it/activities/software/>). In questo modo ciascun utente può facilmente recuperare la lista completa dei software installati sul cluster e le versioni corrispondenti.

Software scientifico

INTRODUZIONE | [ELENCO PACCHETTI PRESENTI SUL CLUSTER](#) | [ISTRUZIONI PER L'USO](#)

IASF-Bologna è attivamente coinvolto in numerosi progetti astrofisici che operano in diverse bande dello spettro elettromagnetico e permettono di studiare l'Universo dalle microonde ai raggi gamma.

La riduzione dei dati grezzi acquisiti da un satellite nello spazio (o da un telescopio a Terra) e la successiva analisi richiedono l'utilizzo di appropriati software scientifici e relative librerie che l'utente dovrebbe avere installato sul proprio computer di lavoro.

Tuttavia, gli utenti che volessero analizzare dati provenienti da diversi satelliti senza dover installare ogni volta il software appropriato, possono farlo attraverso il cluster di Istituto, a cui possono accedere semplicemente facendo il login sui server front-end tonno o bitonno con il proprio username e password.

Sul cluster sono installati librerie e software scientifici pubblici ([lista completa dei pacchetti installati](#)).

Oltre ai compilatori come GCC, librerie MPI, di calcolo scientifico o di accesso ai dati scientifici (Cfitsio e HEALPIX) sul cluster sono installati i software scientifici per la riduzione dei dati relativi a diverse missioni spaziali. In particolare:

- * SAS (Scientific Analysis System) è il pacchetto di riduzione per i dati raccolti dal satellite *XMM-Newton* nella banda X 2-10 keV;
- * un altro satellite che lavora nella stessa banda è *Chandra* il cui pacchetto di riduzione dati è CIAO (Chandra Interactive Analysis of Observations);
- * il pacchetto HEASoft fornisce invece una serie di tools per l'analisi dati, in particolare gli tools necessari, tra l'altro, per la riduzione dati dei satelliti *Suzaku* e *Swift*.

Altri software scientifici installati sono:

- * OSA (Integral Off-Line Scientific Analysis) per il satellite soft-gamma *INTEGRAL*;
- * Fermi Science Tools per l'analisi dati del satellite gamma *Fermi*.

Sempre nel pacchetto HEASoft l'utente troverà il software XSPEC con cui potrà proseguire con l'analisi spettrale dei dati appena ridotti. In alternativa a XSPEC sono disponibili: SHERPA all'interno del pacchetto CIAO e SPEX, che potrebbe essere molto utile per chi lavora con dati X ad alta risoluzione.

Inoltre, è stata recentemente installata una versione di IRAF (Image Reduction and Analysis Facility) e PyIRAF per l'analisi dei dati ottici e infrarossi.

Ogni pacchetto scientifico di riduzione dati è corredato dai relativi file di calibrazione che vengono aggiornati periodicamente.

Infine, tra gli altri pacchetti di cui potrà usufruire l'utente ci sono: i software di visualizzazione dati come ds9, gnuplot, smongo, etc., e il software per il calcolo e le simulazioni scientifiche come IDL, MATHEMATICA e Geant4.

Sul cluster d'Istituto possono essere presenti anche codici di calcolo e dati 'proprietary' delle missioni spaziali, come ad esempio viene fatto per le missioni *Planck* e *Agile*, garantendo la riservatezza dei dati e dei software.

Per segnalazioni riguardanti problemi con i pacchetti disponibili, software da aggiornare o da installare si prega di contattare torresi@iasfbo.inaf.it o derosa@iasfbo.inaf.it.



Di seguito viene riportata una lista dei compilatori, delle librerie e dei software attualmente presenti sul cluster. Per un elenco completo e le relative versioni si prega di visitare la pagina web <http://www.iasfbo.inaf.it/activities/software/elenco-software/>

Compilatori e librerie MPI	GCC
	Scalasca
	Intel
	openMPI

Librerie di calcolo scientifico	Scalapck
	Lapack
	MKL
	ACML
	NAG
	GSL

Librerie per accesso ai dati astrofisici (I/O)	HEALPIX
	Cfitsio
	HDF5

Software di visualizzazione dati scientifici	Paraview
	Paraview MPI
	Ds9
	pyDs9
	Gnuplot
	Grace, Xmgr
	smongo

Software di calcolo e simulazioni scientifiche	IDL
	MPIDL
	FASTIDL
	Mathematica
	Geant

Infine viene riportato l'elenco dei pacchetti per la riduzione e analisi dei dati scientifici, in particolare ottici/infrarossi, X e gamma (aggiornato a Dicembre 2015):

Software per analisi dati	HEAssoft v. 6.17	Set completo di software distribuito da HEASARC (è una release unica dei pacchetti FTOOLS e XANADU)
	Fermi Science Tools v9r32p5	Versione pubblica dei tools per l'analisi dei dati del satellite gamma <i>Fermi</i>
	CIAO v. 4.7 con CALDB 4.6.9	Chandra Interactive Analysis of Observations è il pacchetto di riduzione dati del satellite X <i>Chandra</i>
	SAS v. 14	Scientific Analysis System è il pacchetto per la riduzione e analisi dei dati del satellite X <i>XMM-Newton</i>
	SPEX v. 2.01.05	Software ottimizzato per l'analisi dei dati X ad alta risoluzione
	OSA v. 9.0	Integral Off-Line Scientific Analysis è il pacchetto di riduzione dati degli strumenti a bordo del satellite gamma <i>INTEGRAL</i>
	Scisoft v. 7.7.0	Insieme di software astronomici forniti da ESO
	IRAF v. 2.16	Image Reduction and Analysis Facility software generico per la riduzione e analisi dei dati astronomici in particolare ottici e infrarossi
	pyRAF v. 2.1.6	IRAF in linguaggio python
	Starlink Hikianalia version 2013-04-15	Progetto anglosassone di lunga data che supporta il riprocessamento dei dati astronomici
	SExtractor v. 2.19.5	Programma che costruisce un catalogo di oggetti partendo da un'immagine astronomica
	SWarp v. 2.38.0	Programma che ricampiona e unisce immagini in formato FITS
Le PHARE v. 2.2	Set di comandi fortran per calcolare i redshift fotometrici e performare i fit delle SED	

3.1 Ambiente di sviluppo del cluster di calcolo

Il cluster di calcolo ha un ambiente di sviluppo *modulare*, occorre cioè caricare ogni volta l'ambiente opportuno per svolgere la propria analisi. Questo al fine di evitare possibili conflitti tra i programmi installati e un non corretto funzionamento dei comandi per la riduzione e analisi dei dati.

La gestione di tale ambiente viene effettuata mediante il comando **module**. Questo comando permette di impostare le variabili d'ambiente necessarie all'utilizzo di un software.

Alcuni comandi che possono essere utili quando si caricano le variabili d'ambiente sono:

- **module av** (per avere una lista dei moduli disponibili sul sistema)
- **module load nome-modulo** (per caricare un modulo, es. `module load IDL-8.2`)
- **module unload nome-modulo** (per togliere un modulo già caricato)
- **module list** (per conoscere i moduli caricati)
- **module purge** (per ripulire l'ambiente)

Nel caso di software per la riduzione dati come HEASOFT, CIAO e SAS, dopo aver definito l'ambiente occorre anche inicializzarlo (equivale a lanciare il comando Linux 'source'):

- `module load HEASOFT-6.17`
 - `heainit`
- `module load SAS-14.0`
 - `sasinit`
- `module load CIAO-4.7`
 - `ciaoinit`

In questo caso il comando 'module purge' non è sufficiente per ripristinare l'ambiente di default ma è necessario aprire una nuova finestra di lavoro.

Per maggiori informazioni sui nodi del cluster e sui moduli attualmente presenti si rimanda alla seguente documentazione:

<http://www.iasfbo.inaf.it/download/4559/>
<http://www.iasfbo.inaf.it/download/4562/>
<http://www.iasfbo.inaf.it/download/4565/>