High Performance Computing and Networks (HPCN)

Lidia Leoni, direttore
Piattaforma High Performance Computing and Networks (HPCN),
Centro Ricerche, Sviluppo e Studi Superiori in Sardegna
lidia.leoni@crs4.it

CRS4
Polaris Technological Park, 09010 Pula CA, Italy
http://www.crs4.it

Il CRS4

Il CRS4 è un centro di ricerca multidisciplinare che promuove lo studio, lo sviluppo e l'applicazione di soluzioni innovative a problemi provenienti da ambienti naturali, sociali e industriali.

Tali sviluppi e soluzioni si basano sulla Scienza e Tecnologia dell'Informazione e sul Calcolo Digitale ad alte prestazioni.

L'obiettivo principale è l'Innovazione.

Il CRS4, fondato dal Prof. Rubbia e dal Prof. Zanella nel 1990, è parte del Parco Scientifico e Tecnologico di Pula (POLARIS)

Informazioni generali

Fondato il **30 novembre 1990.**

Entrato nel suo **ventiquattresimo anno di attività** divenendo un punto di riferimento importante per la Regione

- 160 ricercatori/tecnologi
- budget ~ 15M€ di cui ~50% da fondi esterni
  - EU/National research project
  - Industrial contracts

Informazioni generali

Ricerca e sviluppo nei settori abilitanti

L'esperienza diretta nel contesto applicativo con focus primario su: Energia & Scienze ambientali; Società dell'informazione; scienze biomediche

Forte Rete di collaborazioni

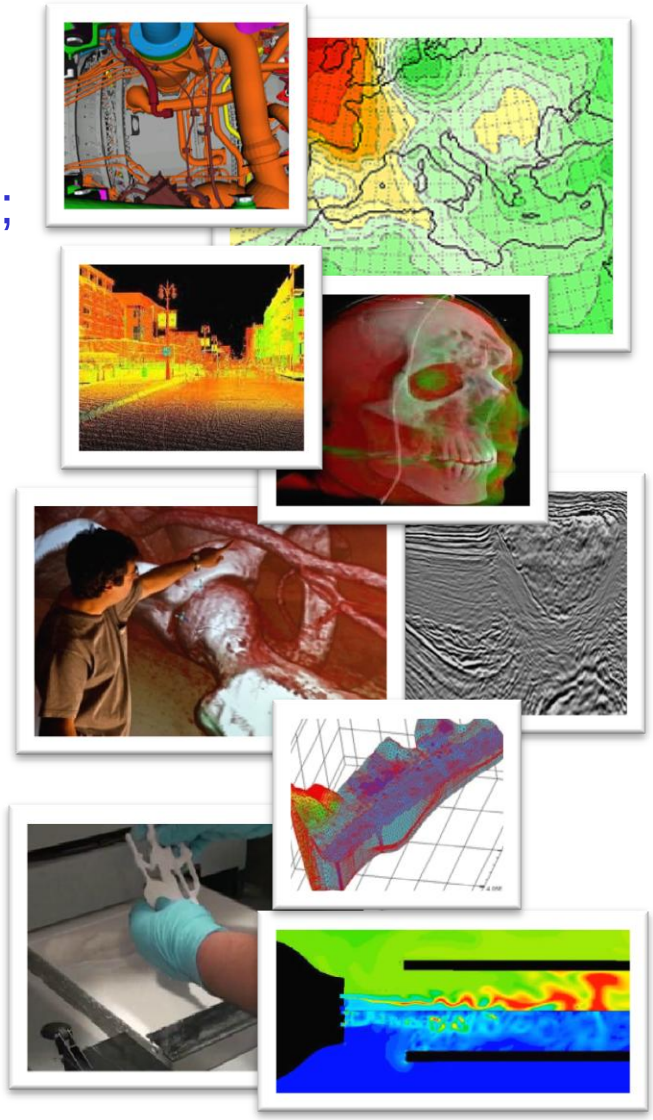
PON, EU FP7, Wellcome Trust, NIH, APL, ...

Trasferimento tecnologico alle industrie

ENI, INPECO, IBM, NICE, GEXCEL, ...

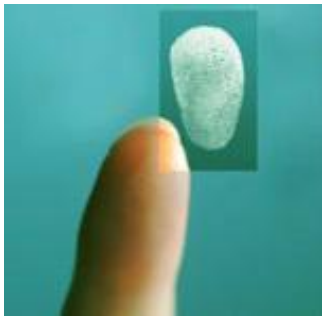
Trasferimento tecnologico alle istituzioni pubbliche

Sanità, Ambiente, Beni Culturali, ...



Partner Industriali

inpeco
automation in healthcare



eni

LUNA ROSSA

MAXELER
Technologies

NICE

GEXCEL

BASE 4

Alta Formazione e training on the job

- **Corso di Analisi Dati Next Generation Sequencing con Galaxy** (8 - 11 ottobre 2013)
- **Corso di amministrazione di cluster di calcolo** alle pubbliche amministrazioni
- **SWAT International Workshop** in Sardinia (September 30 - October 4 2013)
- **Summer School 2011-2012-2013** NGS and GWAS for Complex and Monogenic Disorders September (Sept. 9-13 2013)
- Master on **Bioinformatics** (2 editions) w University of Cagliari (2005-2006, 2012-2013)
- **Sardinian Summer School -** Genomic Analysis of Complex and Monogenic Disorders
- Master on **Renewable Energy** w industrial partners and University of Cagliari

  • up to 10 stagers per year from Universities

Settori di ricerca

Fedele alla sua missione fondativa, il Centro ha promosso programmi di ricerca e sviluppo tecnologico a **carattere multidisciplinare**;

Le Aree di Ricerca si sono evolute e specializzate, mantenendo come **orizzonte di riferimento quello internazionale**, sola dimensione realmente capace di contribuire allo sviluppo del territorio.

All'interno dei **programmi strategici**, non soltanto le **competenze scientifiche e tecnologiche** ma anche la capacità di **parlare a mondi diversi**, vitale per un centro di ricerca, si sono sviluppate e si svilupperanno, per posizionare il Centro su **filoni di frontiera**, cogliendo e anticipando le **opportunità** che si presenteranno.

- ***Energia e Ambiente***
- ***Biomedicina***
- ***Società dell'informazione***
- ***Data Fusion***

Piattaforme Tecnologiche

Il CRS4 possiede un centro di calcolo, la prima piattaforma in Italia dedicata alla genotipizzazione e al sequenziamento massivo del DNA e di un laboratorio di Visual Computing allo stato dell'arte.

Le nostre Piattaforme Tecnologiche sono:

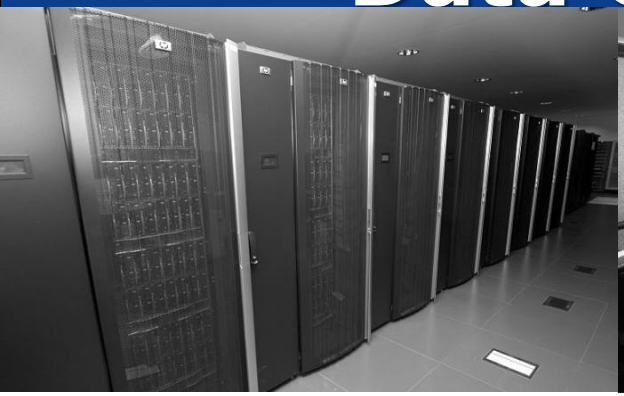
- ***High Performance Computing and Network***
- ***Genotyping and Massive DNA Sequencing***
- ***Visual computing***

La piattaforma HPCN

La piattaforma HPCN mette a disposizione degli utenti un'infrastruttura per il calcolo ad alte prestazioni ed i relativi servizi di supporto alla ricerca. In collaborazione con la sua comunità, valuta, implementa e supporta le nuove tecnologie emergenti. Il calcolo ad alte prestazioni è uno dei principali servizi forniti dal settore ai propri utenti.

- **HPC**
- **Networking & security**
- **System management**

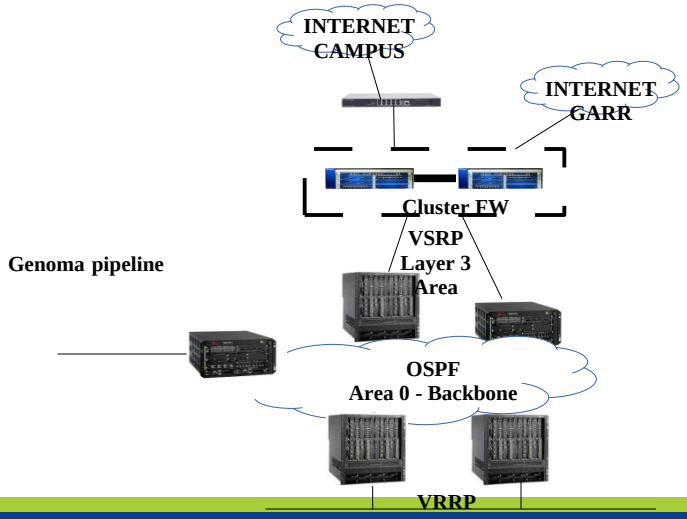
Data Center Overview

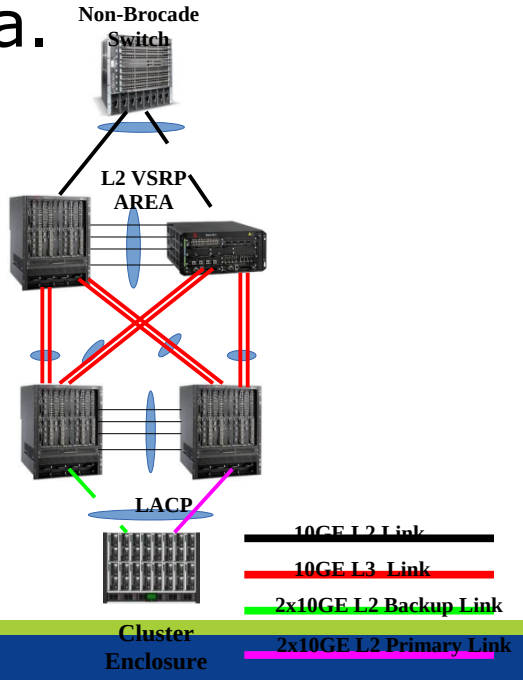


- 170+ TeraFlop di potenza di calcolo su 600+ compute node
- Cluster Standard e ibridi (GPU, IBM Cell, FPGA…)
- 5 Petabyte di storage
- 250 macchine virtuali per servizi, sviluppo e progetti specifici
- 1 Gbps link GARR, link a 100 Mbps con un provider
- 350 porte IB (DDR, QDR, FDR)
- 300+ porte 10GE e circa 1200 porte 1GE
- Rete Ottica DWDM (ROADM) – sino a 80 Gbps PtP nella rete CyberSar
- Rete Ottica CWDM

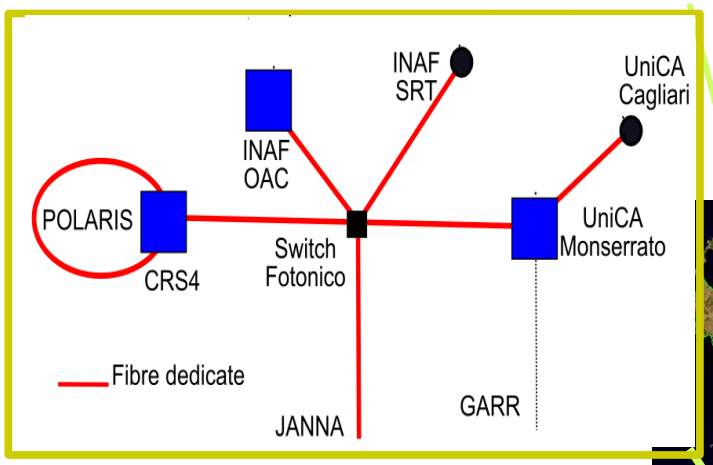
Rete e Sicurezza

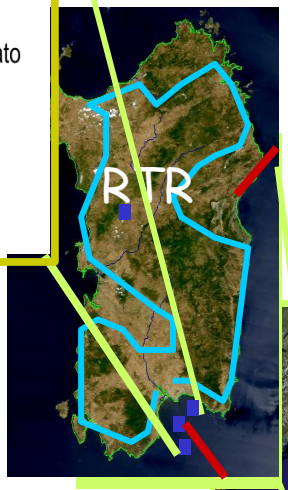
L'infrastruttura di rete è un componente essenziale e trasversale a tutti i progetti e le attività di ricerca del CRS4. Per rispondere adeguatamente alle esigenze dei ricercatori, negli anni, l'infrastruttura di rete è stata modificata sia dal punto di vista architetturale che tecnologico, utilizzando sempre soluzioni di ultima generazione. L'architettura corrente è stata disegnata per offrire elevate caratteristiche di robustezza, prestazioni e resilienza.



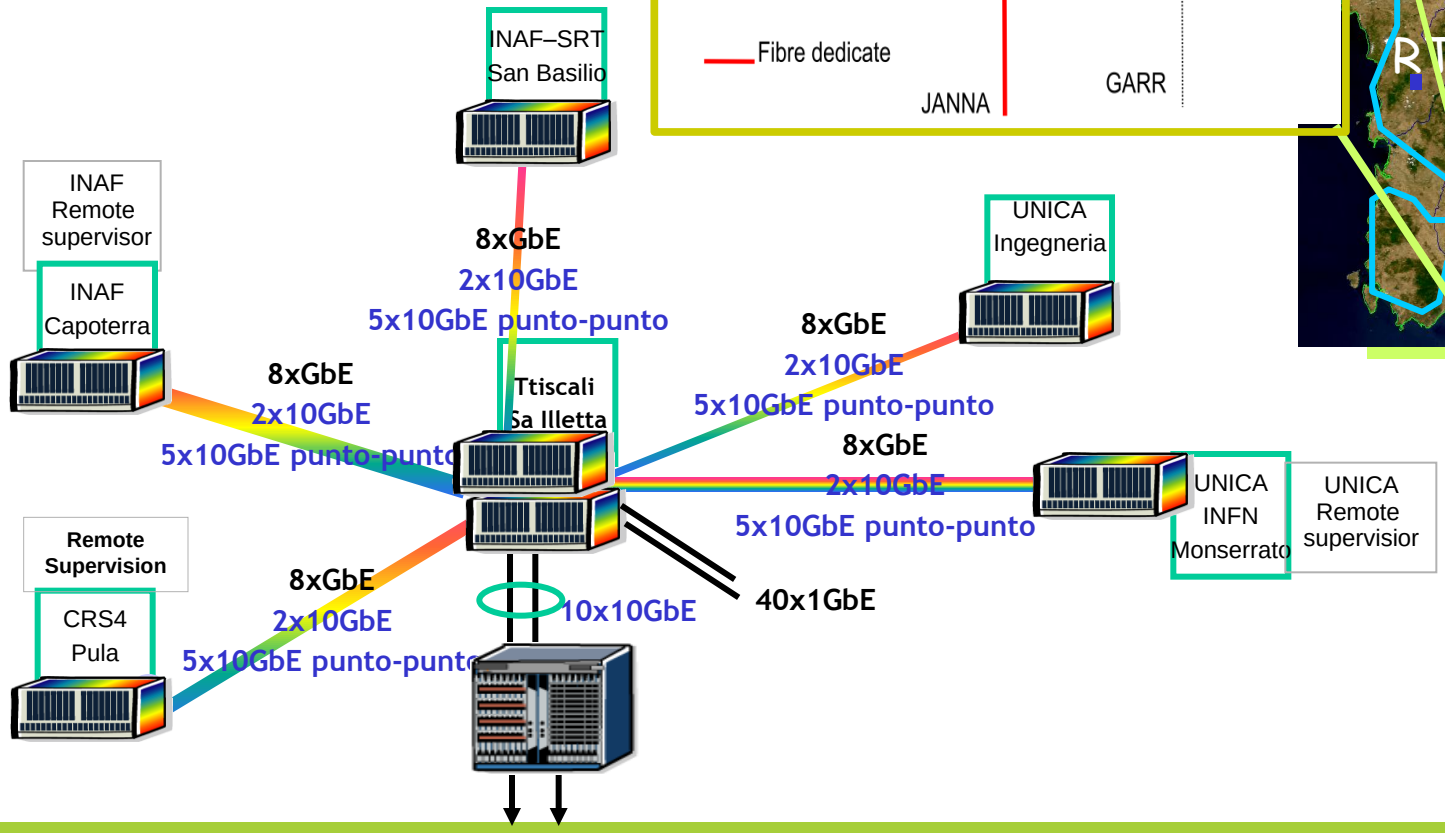


Infrastruttura della rete CyberSar su RTR-R

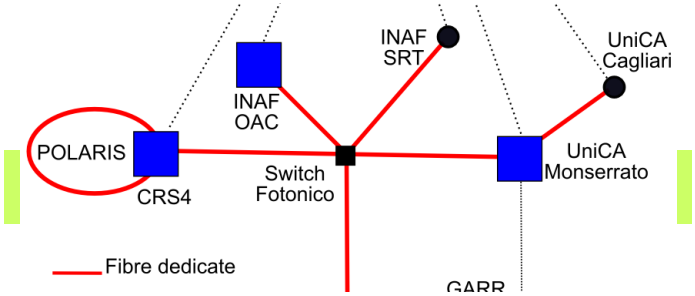




Janna dark fibe

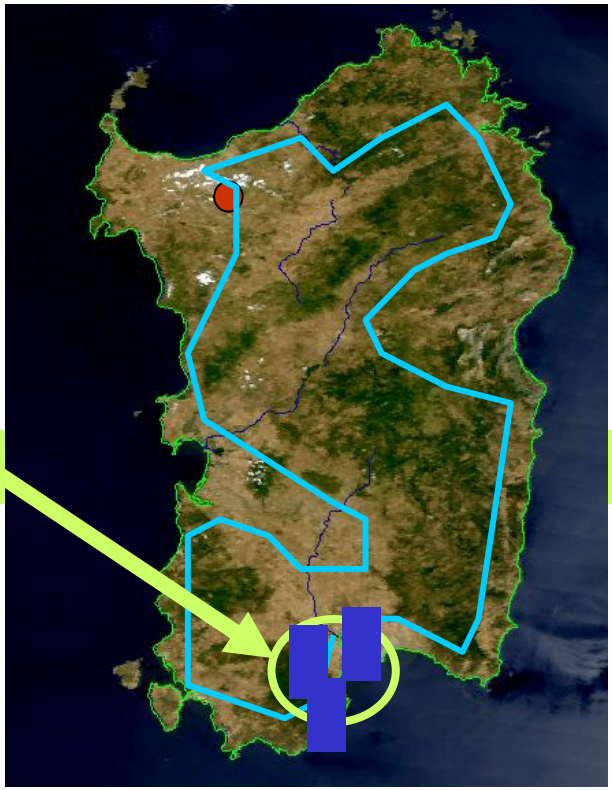


Rete Cybersar



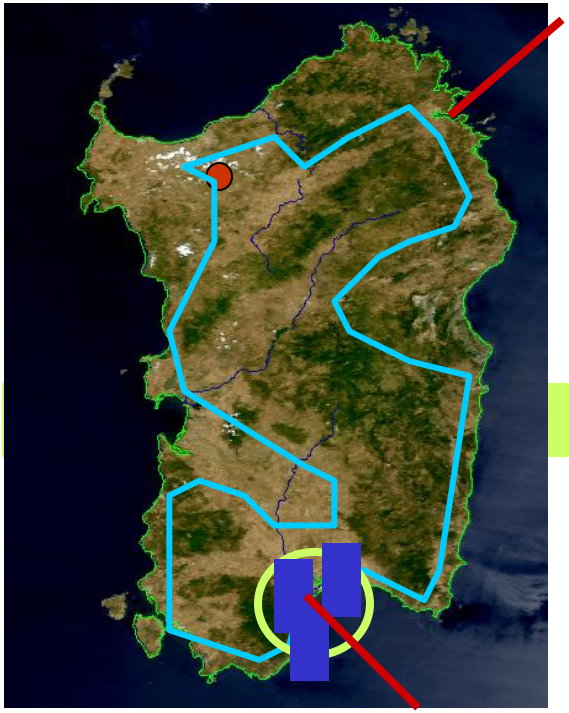
Dark fiber core

Rete Regionale Ricerca



RTR DWDM network
10Gb/s lambdas

RTR-R Rete Telematica Regionale - Ricerca



DWDM network
10 Gb/s lambdas

Janna dark fibers

HPC (High Performance computing)

Il CRS4 offre supporto alle attività di ricerca della comunità scientifica tramite il **Calcolo ad alte prestazioni** e le sue applicazioni, grazie a un centro di calcolo allo stato dell'arte. Il suo personale specializzato è altamente qualificato e affianca i ricercatori nell'utilizzo dell'infrastruttura tecnologica. La potenza di calcolo del CRS4 è di circa 170 TFlops, così suddivisi:

- ***GPU Nvidia Kepler K40 34 Tflops***
- ***GPU NVidia Kepler K10 90 Tflops***
- ***HP Cluster 34,6 Tflop (bassa e media latenza)***
- ***Sun Cluster 3 TFlops***
- ***Tesla cluster***
- ***GPU NVidia cluster***
- ***FPGA Maxeler***
- ***IBM Cluster and other resources 3 Tflops***

Formazione di personale da adibire ad attività sistemistiche e di ricerca

Settori di ricerca sul cluster CRS4

Fisica dei Materiali: CNR-IOM/Unica - interazioni tra molecole organiche e materiali (zinc oxide), proprietà elettriche e conduttive di materiali e particelle, problemi legati all'ambiente e all'ossidazione dei materiali come il piombo.

Meteorologia: previsioni multiple al fine di stimare la probabilità che un dato evento si verifichi

Bioinformatica e genomica: molecole di interesse farmaceutico e geni o proteine

Geofisica: studio e sviluppo di metodi numerici, basati sulle onde acustiche, per ricostruzione del sottosuolo

Chimica: celle a combustibile

Fluidodinamica: LunaRossaChallenge - Simulazioni su 1024 cores per ricerca CFD, Karalit, fluidodinamica dei materiali

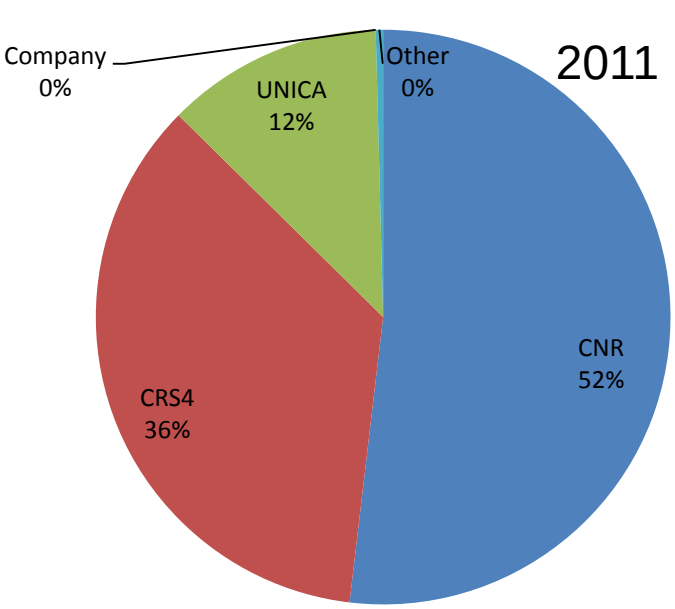
Chimica-bioinformatica: dinamica molecolare

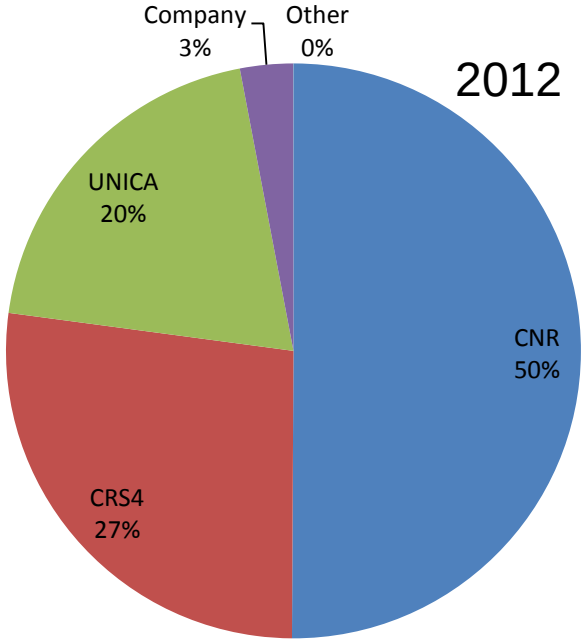
Termodiffusione: generazione di gradienti di concentrazione causati da gradienti di temperatura

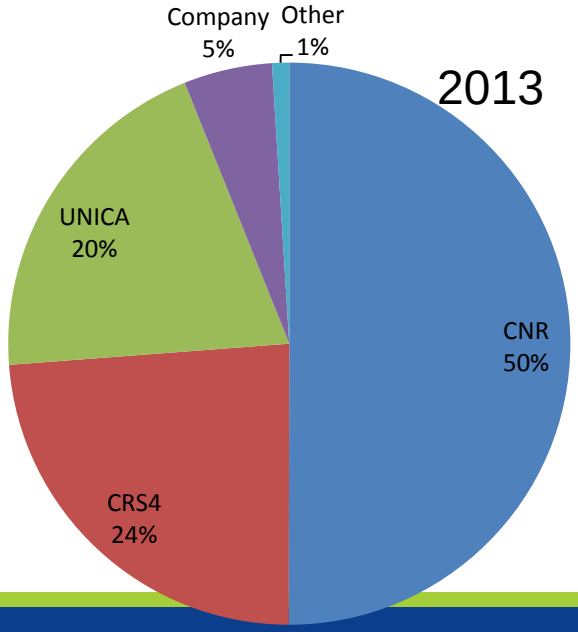
Bioinformatica: simulazione in silicio dell'attività elettrica del cuore

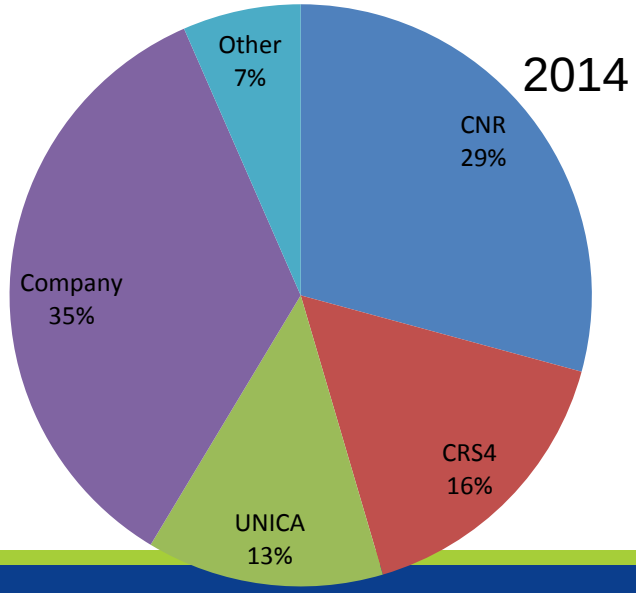
Dinamica molecolare: Atomi e molecole nello studio di patologie

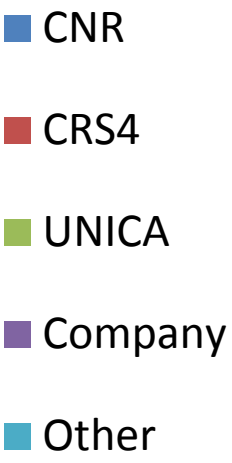
Utilizzo Risorse di calcolo CRS4 2011-2014











Utilizzo delle risorse HPC e software DRMS

Il principale software di DRMS (Distributed Resources Manager System) e' **OpenGridScheduler**. Abbiamo anche LSF di Platform (Load Share Facility)

I software DRMS controllano e sovrintendono l'utilizzo delle risorse di calcolo come:

- CPU
- Memoria
- spazio disco
- licenze
- rete.

Gli utenti richiedono le risorse sottomettendo, tramite il DRMS, i loro i lavori che possono essere sequenziali o paralleli.

Abbiamo piu' di 100 utente che attivamente utilizzano il cluster. Altri utenti sono raggruppati in progetti o utilizzano le risorse, anche dall'esterno del CRS4, tramite portali e servizi vari.

Software e supporto

Compilatori: **Intel, PGI (Portland Group), Compilatori Gnu, CUDA**

MPI: **MVAPICH, MPI, Openmpi**

Software: **StarCD/StarCcm+, Ansys Fluent, Paraview, NAMD, Gromacs, ACEMD, AMBER, DL_POLY, Grace, MayaVi, VMD, Quantum-Espresso, Annovar, OpenFoam, Schrodinger**

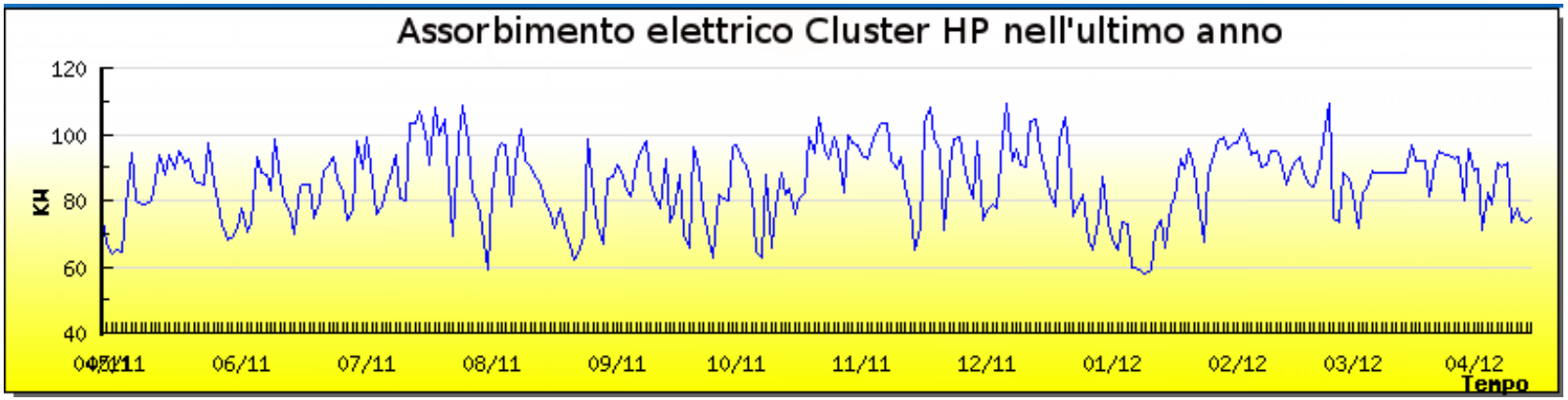
Tools: **Perl, Python, QT, R, Java, tcl/tk, Valgrind**

Librerie: **ATLAS, Lapack e Parpack, Blas, FFTW, ACML, MKL, Metis, Aztec, ….**

Il supporto agli utenti è fornito anche fornendo supporto personalizzato sulle risorse o sul software. Questa politica si è rivelata molto utile in diversi settori della ricerca scientifica.

Software di management MUCCA

Impatto del modulo GREEN di MUCCA nel consumo del cluster HP



Tutto il cluster acceso, senza jobs utente, assorbe circa 85KW/h!!!!

Risparmio sino a 30KW/h per brevi periodi

Forte integrazione tra l'hardware e lo scheduler

Policy e granularita' adattabili per gruppi di macchine

Cluster MEDUSA per CFD

Cluster con 640 core composto da

- 32 nodi dual CPU Intel X86_64
  - 2 CPU con 10 core, 2.8GHz
  - 128 GB RAM (circa 5GB per core)
  - dischi SSD
  - IB FDR (56 Gbps)
  - 1 GE

LUNA ROSSA

Infrastruttura di Storage

Spazio disco 5 PB disponibili

- Cluster storage DDN 3,8 PB
- Sun StorageTek 500 TB
- Cluster space 200 TB
- DAS e NFS File server 500 TB

Spazio di Backup approx 800TB

Infrastruttura DDN

2 Controller SFA 10K

- 1800 dischi da 3TB - **5,4 PB RAW**

12 Server con connessioni **IB QDR and 10 GbE**

1 Storage IBM Storwize V7000 con 24 dischi SAS da 600 GB 10000 RPM per la gestione dei metadata catalog, 2 Server con connessioni 10 GbE collegati allo Storwize

Filesystem utilizzato: **GPFS 3.4**

L'installazione supporta il collegamento istantaneo di oltre 1000 client

Lo spazio disco utilizzabile ammonta a 3.8 PByte ed è accessibile da un unico mountpoint

Il metadata catalog è ospitato su infrastruttura separata con dischi SAS da 10K rpm, configurati in raid 5

L'accesso al filesystem è garantito da 10 server IBM collegati allo storage con tecnologia IB

Architettura priva di single point of failure

DataDirect™ NETWORKS

Gestione criticità sulle infrastrutture IT

- Classificazione delle applicazioni e servizi in classi di criticità
  - Livello Alto
  - Livello Medio
  - Livello Basso

- Trattamento differenziato in corrispondenza di criticità alle varie classi di servizi con tempi di ripristino differenziati (dipende dal budget associato a servizi e progetti)

- Pubblicazione dellla procedura di spegnimento (protezione) e ripristino delle infrastrutture

Accesso utenti CRS4

4 differenti tipologie di utenti:

- *Personale CRS4* **-** Tutte le risorse
- *Utenti esterni* **-** Solo le risorse di progetto
- *Studenti* - Dipende dal loro tutor e dal tipo di stage
- *Guest* - Solo accesso alla rete pubblica

Tipi di accesso alle risorse

- *VPN*
- *SSH*
- *FTP (project area)*
- *http/s (wiki, SVN, webmail, cloud….)*

Infrastruttura tecnologica

Il Centro di calcolo del CRS4 è classificato Tier 2:

- Ridondanza su alcuni componenti (G.E., UPS, Chiller) che supportano il Data Center
- Nessuna ridondanza a livello dell' alimentazione (power center) o della distribuzione o per effettuare alcune manutenzioni strutturali senza fermate
- Collegamento su più fibre ma su unico cavedio
- Fermata del sistema per manutenzione annuale

Attivita' HPCN in ambito ambiente di Calcolo

Potenziamento degli strumenti di **management** creati al CRS4 **MUCCA 2.1)**

- moduli di gestione cluster e monitoring: hardware, software, e jobs utente
- Modulo GREEN: risparmio energetico gestione intelligente dei nodi non utilizzati)

Cloud Computing

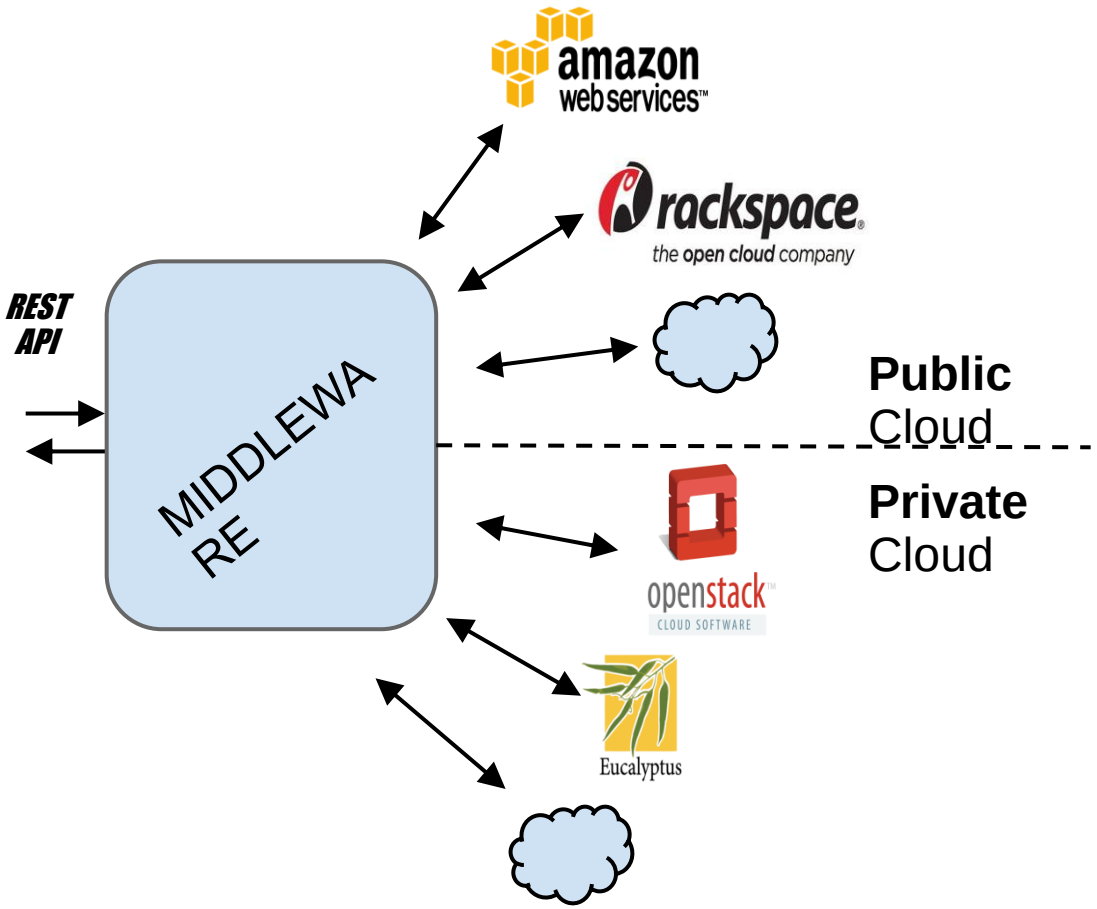
- Progetto interno basato su Openstack
  - Opensource, modulare e ben supportato (ampia community)
  - Orientata all'infrastruttura (IaaS, ottimizzazione risorse di calcolo e risparmio energetico)
  - Flessibile (raggruppamento ed isolamento)
  - Nuovo modello di fruizione dell'infrastruttura
- Collaborazioni esterne Bando PIA

Studio e sviluppo di **metodologie** per l'analisi e la gestione di **grandi quantità di dati**

- Analisi Python
- Copia parallela dei dati (5TB con 10 client paralleli in poche ore)

Esperienza sul Cloud Computing Open Source

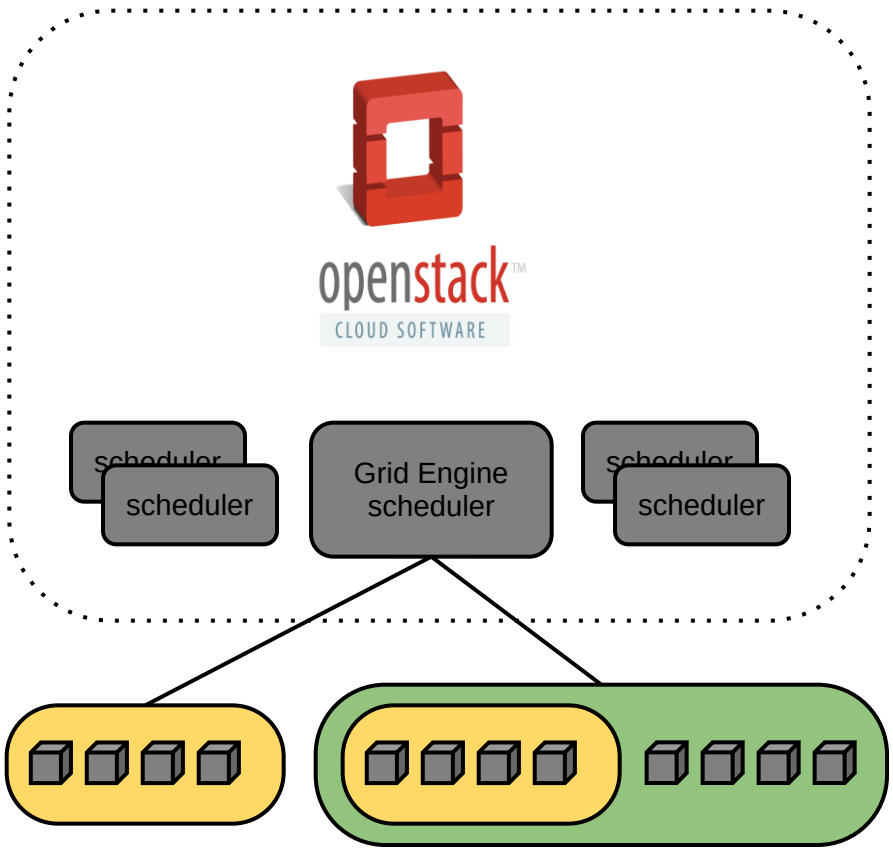
- 2009 - 2011 Eucalyptus
- 2011 - oggi **OpenStack**



E' stata sviluppato un Middleware che consente di interagire con cloud pubblici e privati usando un unico set di chiamate REST attraverso le API messe a disposizione.

In questo modo si centralizza il controllo e si semplifica l'accesso ai servizi.

Le API consentono di creare dei servizi aggiuntivi scollegati dal dialetto del cloud scelto e dalla localizzazione.

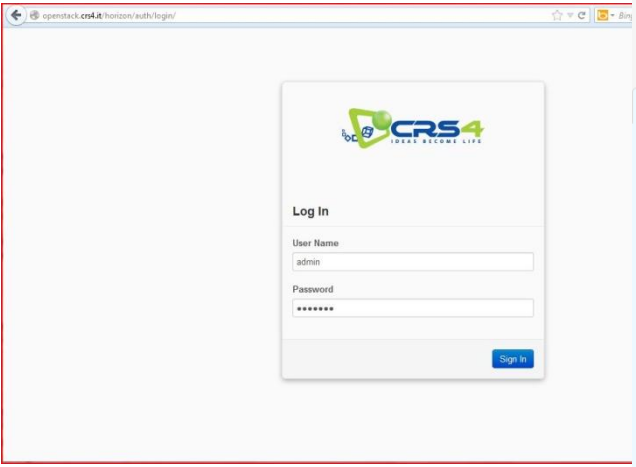


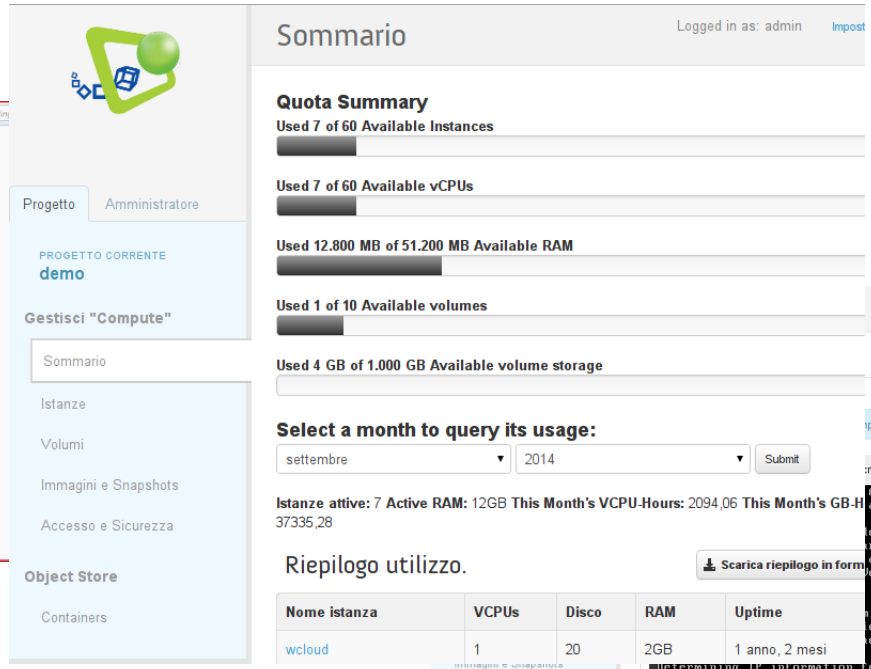
OpenStack farm

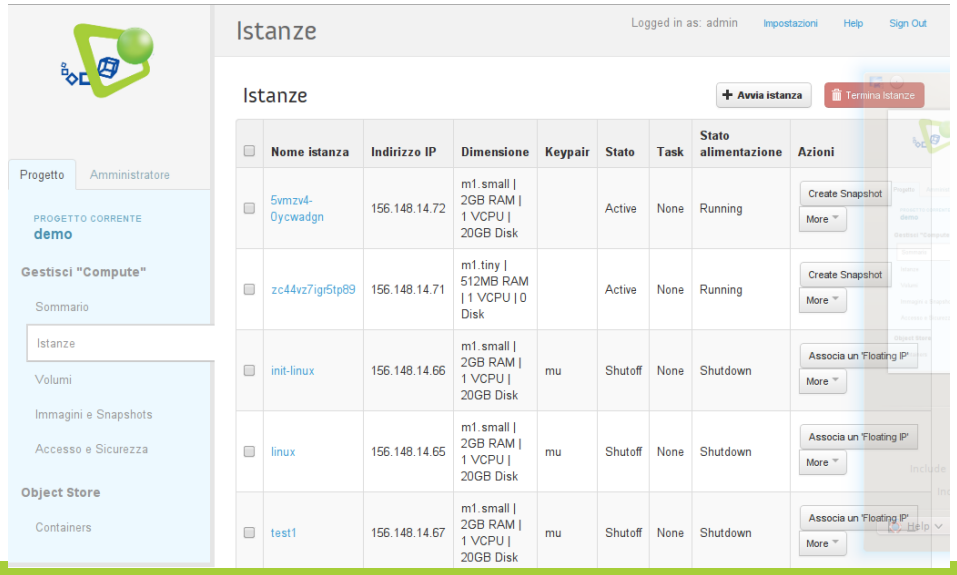
Grid Engine farm

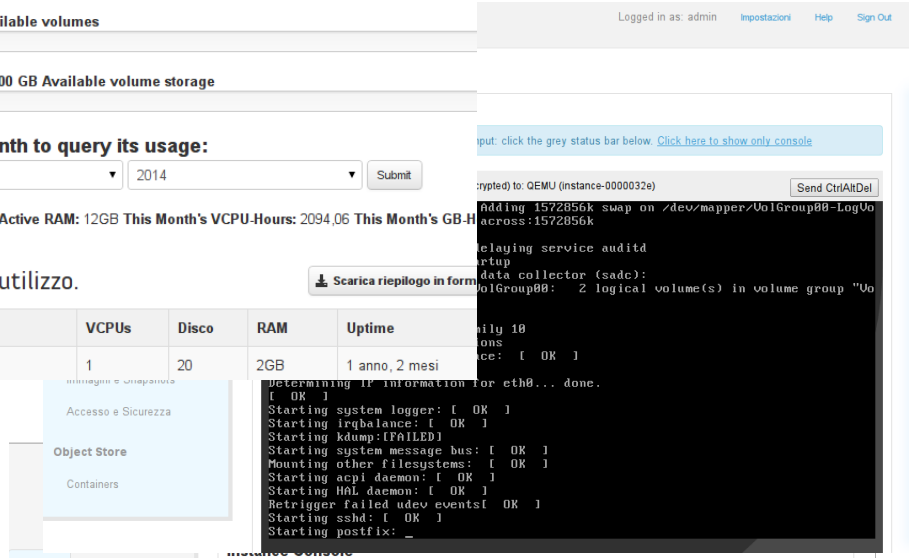
Grid Engine Scheduler e' un componente che abbiamo sviluppato che ci consente di far convivere le risorse tra OpenStack e Grid Engine.
In questo modo e' possibile gestire un pool di macchine condivise tra OpenStack e Grid Engine tanto facilmente quanto uno interamente dedicato.
Dal punto di vista dell'utente questo aspetto e' trasparente, dal lato gestione questo consente di dirottare la destinazione d'uso delle macchine fisiche verso le esigenze del centro di calcolo.

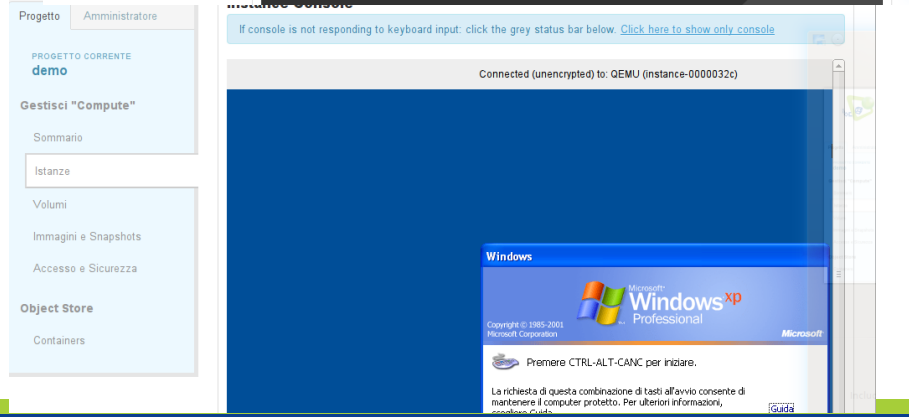
Portale Cloud CRS4











Sardinia Radio Telescope archive system

Il Sardinia Radio Telescope è capace di produrre flussi di dati di varie dimensioni, che vanno dai 5 KB/s (quasi 20 MB all'ora) del *total power* sino ai 512 MB/s (quasi 2 GB all'ora) delle ROACH.

Facendo una stima approssimativa dei tempi di utilizzo annui dei vari backend, possiamo dedurre che SRT attualmente è capace di produrre centinaia di TB di dati ogni anno, che dovranno essere archiviati da qualche parte.

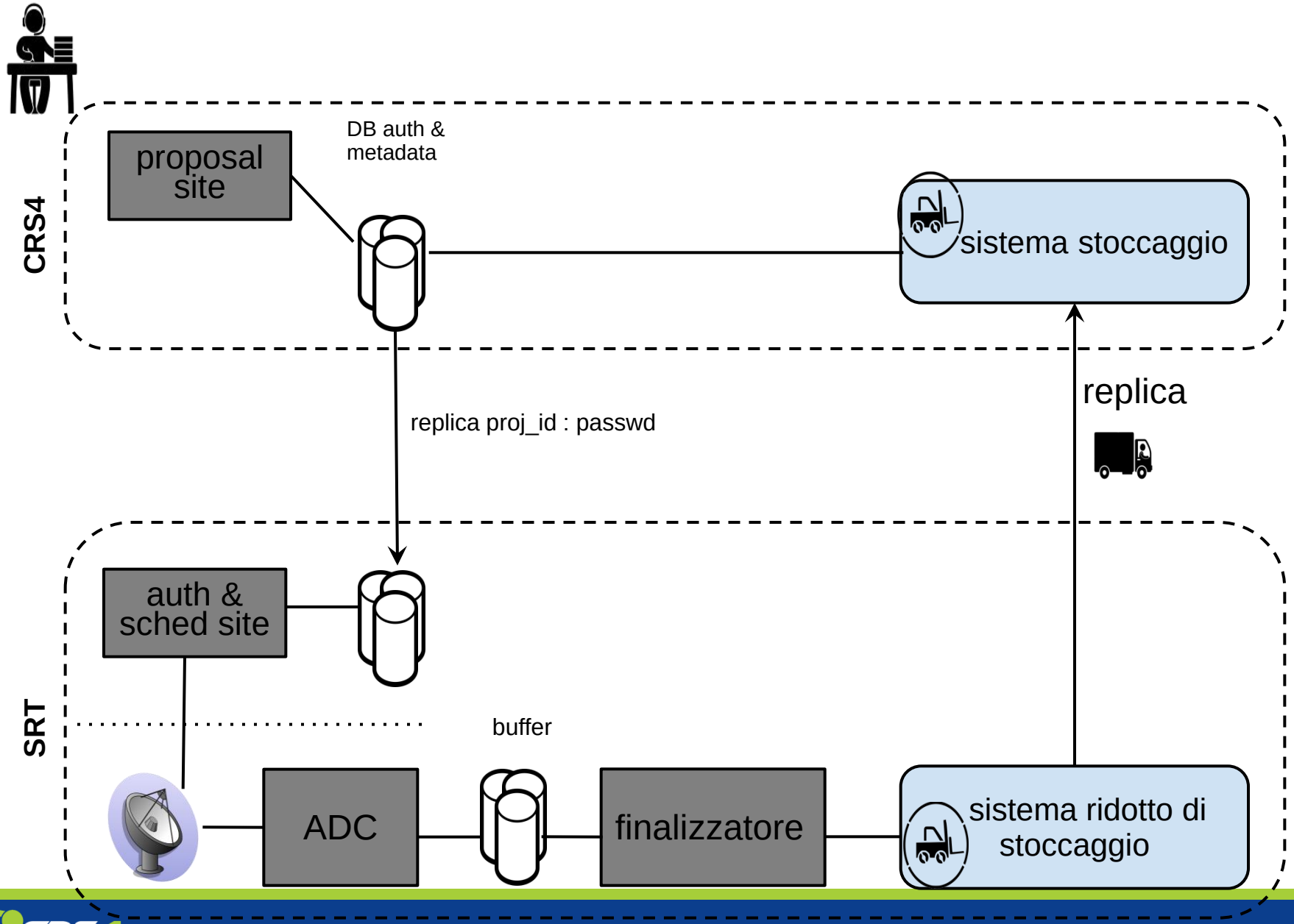
Nei prossimi anni saranno disponibili nuovi backend, come ad esempio la ROACH 2, capace di produrre un flusso di dati di 40 GB/s (quasi 150 TB all'ora).

Il trend è quindi abbastanza chiaro, e indica che la quantità di dati che annualmente dovremo archiviare crescerà con il passare del tempo.

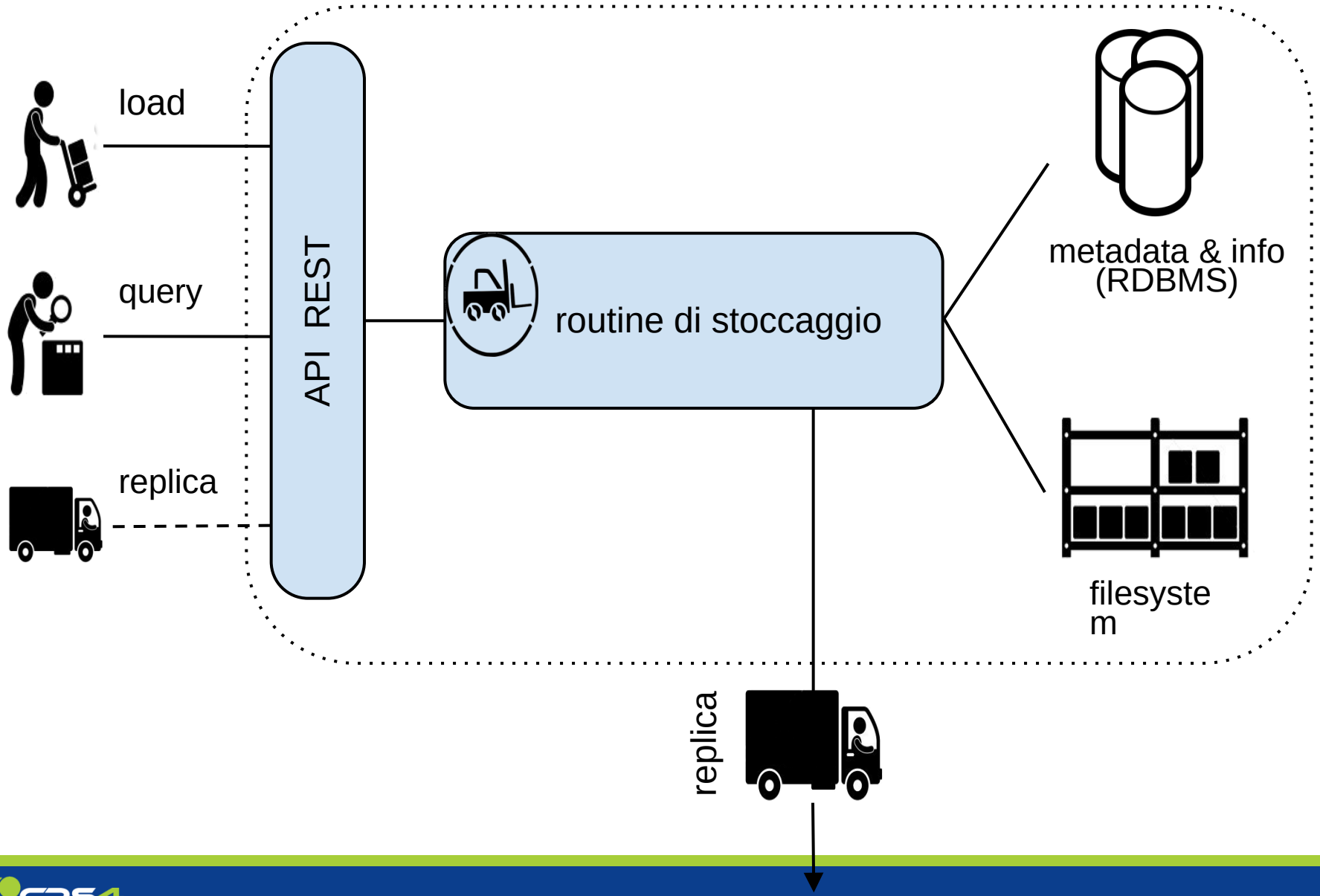
Il progetto si propone di realizzare un sistema di archiviazione

- **scalabile**
- **costi complessivi che decrescano con l'aumentare dei dati**
- **risorse computazionali.**

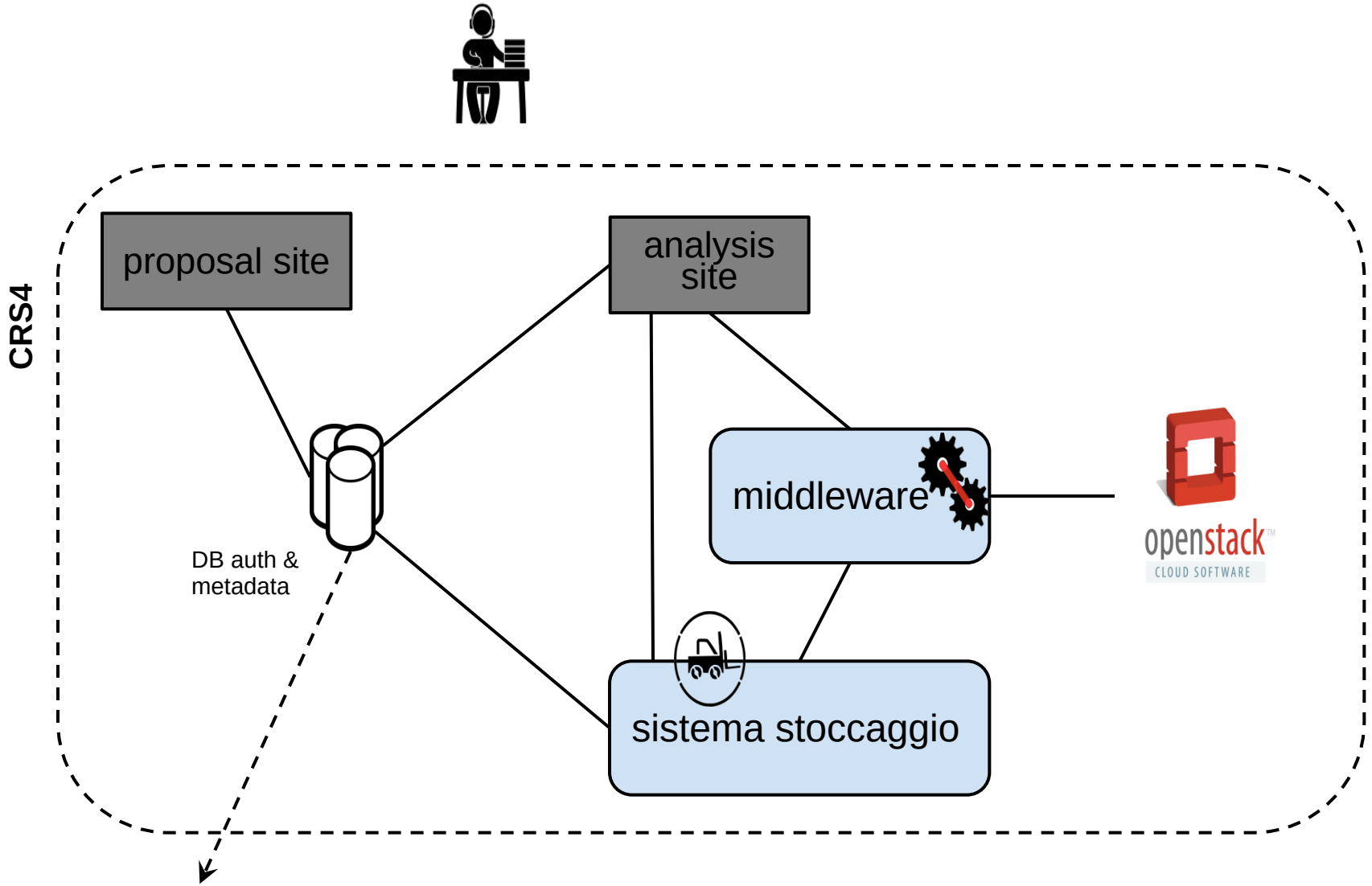
Sardinia Radio Telescope archive system: schema



Progetto Storage SRT: sistema di stoccaggio



Progetto Storage SRT: evoluzioni del flusso



Acquisizioni future

Il CRS4 ha previsto un piano di upgrade delle infrastrutture di calcolo che consentiranno l'acquisizione di:

- Cluster
- Storage
- Rete
- UPS, GE, impianto elettrico ….

Future project vision

REGIONE AUTÒNOMA DE SARDIGNA
REGIONE AUTONOMA DELLA SARDEGNA