

# RECaS BARI



UNIVERSITÀ  
DEGLI STUDI DI BARI  
ALDO MORO



Istituto Nazionale di Fisica Nucleare

Bari, .....

## **Relazione sull'attività del Centro di Bari dell'Infrastruttura ReCaS dal 1 novembre 2017 al 31 ottobre 2018**

Giorgio Pietro Maggi

ReCaS-Bari Via Orabona, 4 – 70125 BARI

Telefono: +39 080 5442357; +39 080 5443134

Sito Web: [www.recas-bari.it](http://www.recas-bari.it) e-mail: [info@recas-bari.it](mailto:info@recas-bari.it) Pec: [recas-bari@pec.uniba.it](mailto:recas-bari@pec.uniba.it)

Supporto: [support@recas-bari.it](mailto:support@recas-bari.it)

<b>Introduzione</b>	<b>3</b>
<b>L'assetto organizzativo</b>	<b>3</b>
<b>I servizi e le risorse</b>	<b>5</b>
Farm HTC	6
Cloud@ReCaS-Bari	7
Tier1 @ReCaS-Bari	11
Cluster HPC	12
Disk storage	13
Tape storage	13
Housing of IT resources	13
<b>Le risorse umane</b>	<b>15</b>
Skill ed expertise	15
<b>Le attività</b>	<b>16</b>
Attività a carattere amministrativo	16
Stipula convenzioni e/o accordi	16
Emissione Ordinativi	16
Riscossioni contributi su convenzione e/o accordi.	17
Gestione tecnica del centro	17
Ricerca e Sviluppo	18
INFN Corporate Cloud (INFN-CC)	20
<b>Gli utenti</b>	<b>22</b>
<b>Pubblicazioni</b>	<b>23</b>
Papers:	23
Conference Papers:	24
Preprint:	24
Presentations:	25
Posters:	25

## Introduzione

La presente relazione, che copre il periodo dal 1 Novembre 2017 al 31 ottobre 2018, periodo che nel seguito viene anche indicato come “periodo di riferimento”, è redatta ai sensi del comma 6.3 c) dell’“Accordo per la gestione del Centro di Bari dell’infrastruttura “ReCaS”” (in seguito indicato come Accordo) tra l’Università degli Studi di Bari “Aldo Moro” (in seguito anche UNIBA) e l’Istituto Nazionale di Fisica Nucleare (in seguito anche INFN) firmato in data 26/7/2017.

## L’assetto organizzativo

L’assetto organizzativo del Centro di Bari dell’Infrastruttura ReCaS (nel seguito anche semplicemente Centro oppure ReCaS-Bari) è definito dal già citato “Accordo per la gestione del Centro di Bari dell’infrastruttura ReCaS”.

In base all’Accordo:

- La gestione del Centro è supervisionata da un **Coordinatore** nominato di comune accordo tra le Parti, tramite scambio di corrispondenza tra il Presidente dell’INFN e il Rettore di UNIBA, su proposta del Direttore della Sezione INFN e del Direttore del Dipartimento Interateneo di Fisica, sentito il Comitato Tecnico Scientifico.
- Le decisioni inerenti l’utilizzo delle risorse del Centro vengono adottate dal **Comitato Paritetico** a cui spetta anche il compito di approvare
  - a. la relazione annuale sulle attività svolte, predisposta dal Coordinatore;
  - b. e il regolamento di funzionamento del Comitato Tecnico Scientifico

Il Comitato Paritetico nell’anno di riferimento era così composto:

- il Coordinatore, che lo presiede (Prof. Giorgio Pietro Maggi)
- il Direttore della Sezione di Bari dell’INFN (Prof. Mauro de Palma)
- il Direttore del Dipartimento Interateneo di Fisica (Prof. Salvatore Vitale Nuzzo)
- un rappresentante di UNIBA, nominato dal Rettore (Prof. Roberto Bellotti)
- un rappresentante della Sezione di Bari dell’INFN, nominato dal Presidente (Dott. Domenico Elia)

L’accordo prevede anche le seguenti figure:

- Il **Vice Coordinatore**, nominato dal Coordinatore tra i membri del Comitato Paritetico, (prof. Roberto Bellotti).
- Il **Responsabile Tecnico**, nominato tra il personale INFN e UNIBA che concorre al funzionamento del Centro, coordina il personale tecnico ed opera le scelte tecniche per assicurare il funzionamento del Centro in accordo con gli obiettivi fissati dal Coordinatore e dal Comitato Paritetico (Dott. Giacinto Donvito).
- Il **Comitato Tecnico Scientifico** che è composto da sette membri designati dal Comitato Paritetico. Ciascun membro resta in carica tre anni.
  1. Prof. Filippo Lanubile, Presidente (UNIBA)
  2. Dott. Claudio Grandi (INFN)
  3. Prof. Tommaso Di Noia (POLIBA)
  4. Dott. Federico Ruggieri (GARR)
  5. Prof. Graziano Pesole (IBIOM-CNR)
  6. Dott. Antonio Zilli (DTA)
  7. Dott. Davide Salomoni (INFN)

ReCaS-Bari ha cominciato ad operare con il nuovo assetto organizzativo con la nomina del Coordinatore (Prof. Giorgio Maggi) in data 20 ottobre 2017 e del rappresentante dell'INFN nel Comitato Paritetico, il dott. Domenico Elia, in data 13 Novembre 2017.

Nel corso dell'anno oggetto della relazione si è riunito 4 volte:

- il 28 novembre 2017
- l'8 gennaio 2018
- il 4 aprile 2018
- il 19 luglio 2018

## I servizi e le risorse

La Fig. 1 mostra schematicamente i servizi disponibili su ReCaS-Bari e come le risorse sono ripartite tra i vari servizi.

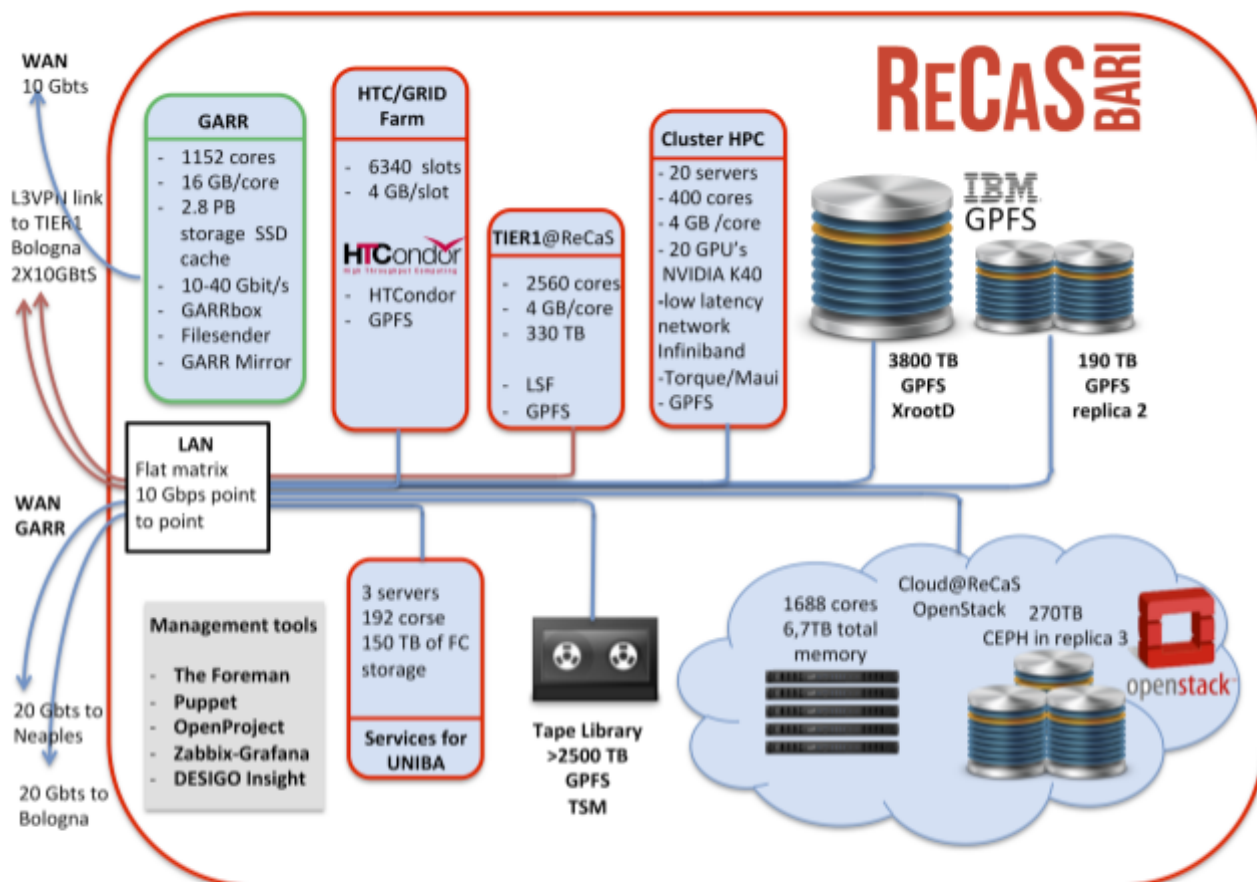


Fig. 1 Schema di massima dei servizi disponibili su ReCaS-Bari e ripartizione delle risorse sui servizi

Le risorse gestite da ReCaS-Bari sono quelle elencate nell'Allegato Tecnico all'Accordo a cui si sono aggiunte le acquisizioni 2017:

- Alice: 4 nodi di calcolo da 64 core (core=256, KHS06=3.07), 270 TB Netti di storage.
- CMS: 4 nodi di calcolo da 64 core (core=256, KHS06=3.07), 950 TB Netti di storage.

A causa del protrarsi della gara, purtroppo le risorse di storage non sono ancora state fisicamente acquisite, installate e messe in operazione. Le CPU invece sono già operative.

Per quanto riguarda invece le acquisizioni 2018, nel periodo di riferimento della presente relazione sono state avviate le gare (attualmente in fase di espletamento) per acquisire il seguente materiale:

- Alice: 3 nodi di calcolo da 40 core (core=120, KHS06=1.8), 280 TB raw di storage.
- CMS: 720 TB raw di storage.

In particolare, gli acquisti dello storage sono stati fatti con il CUP del Progetto IBiSCo, al momento ancora in fase di approvazione.

## Farm HTC

La farm HTC è il più importante dei servizi disponibili presso ReCaS-Bari, attraverso cui vengono svolte le funzioni di TIER2 (centro di calcolo di livello intermedio) per gli esperimenti ALICE (A Large Ion Collider Experiment) e CMS (Compact Muon Solenoid) in svolgimento presso il Large Hadron Collider (LHC) al CERN, a cui partecipano ricercatori dell'INFN, dell'Università di Bari e del Politecnico di Bari.

Per svolgere tali funzioni la farm batch di ReCaS-Bari è integrata nel Worldwide LHC Computing Grid (WLCG) e nella European Grid Infrastructure (EGI).

La farm HTC, nel periodo di riferimento, ha potuto contare su un numero di slot di calcolo variabile nel tempo e compreso tra 6000 e 9000 job slot, ciascuno dei quali equipaggiato da 4 GB di memoria RAM.

Ciascun server di calcolo, che contiene fino ad un massimo di 64 slot, può accedere, alla velocità di 10 Gbps, a tutto lo spazio disco di ReCaS-Bari, dell'ordine di 3800 TB in singola replica a cui si aggiungono 220 TB in doppia replica (per home degli utenti). Per la gestione dello storage è stato utilizzato il file system distribuito GPFS.

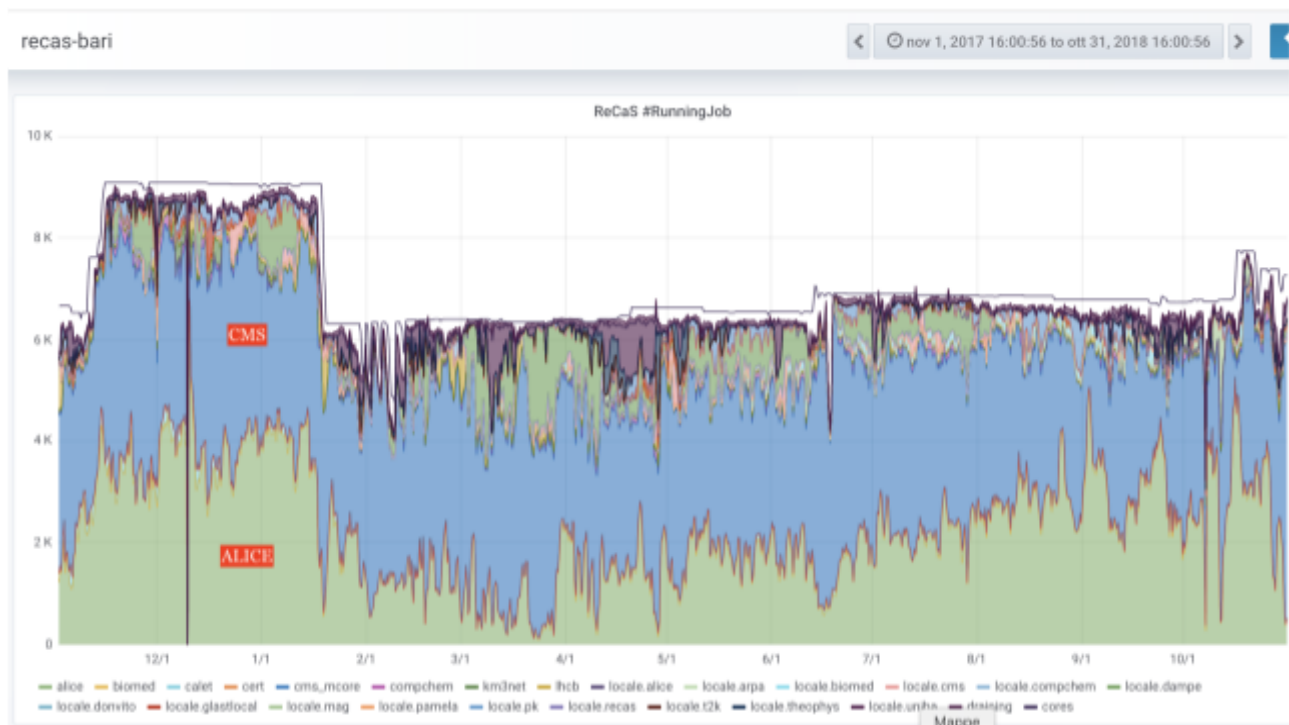


Fig. 2 - Occupazione della farm batch nel periodo 1 nov 2017- 31 ott 2018

Il sistema operativo utilizzato è stato Scientific Linux 6, il cui aggiornamento a CentOS 7 è iniziato verso la fine del periodo di riferimento. Le code sul sistema batch sono state gestite in maniera molto soddisfacente da HTCondor sia per job mono- che multi-core.

Il massimo tempo di esecuzione consentito per ciascun job è di 50 ore, fatto salvo per alcune eccezioni che richiedono un tempo di esecuzione più lungo dovuto sostanzialmente ad esigenze applicative.

I job possono essere sottomessi

- via GRID utilizzando un qualsiasi tool di sottomissione di job, come per esempio CRAB, Dirac, JST, etc.
- dai server di front end a cui si accede con le proprie credenziali.
  - per Planetek è stata realizzata una macchina di sottomissione ad-hoc in ambiente cloud per permettere una più facile integrazione nel loro workflow di analisi e pubblicazione dei dati.

Per eseguire job che richiedono un ambiente diverso da Scientific Linux 6 (o più recentemente CentOS 7), è possibile utilizzare container Docker sulle risorse batch tramite il software uDocker.

Dalla Fig. 2 si evince che la farm HTC ha sempre operato in condizioni di saturazione per tutto il periodo di riferimento.

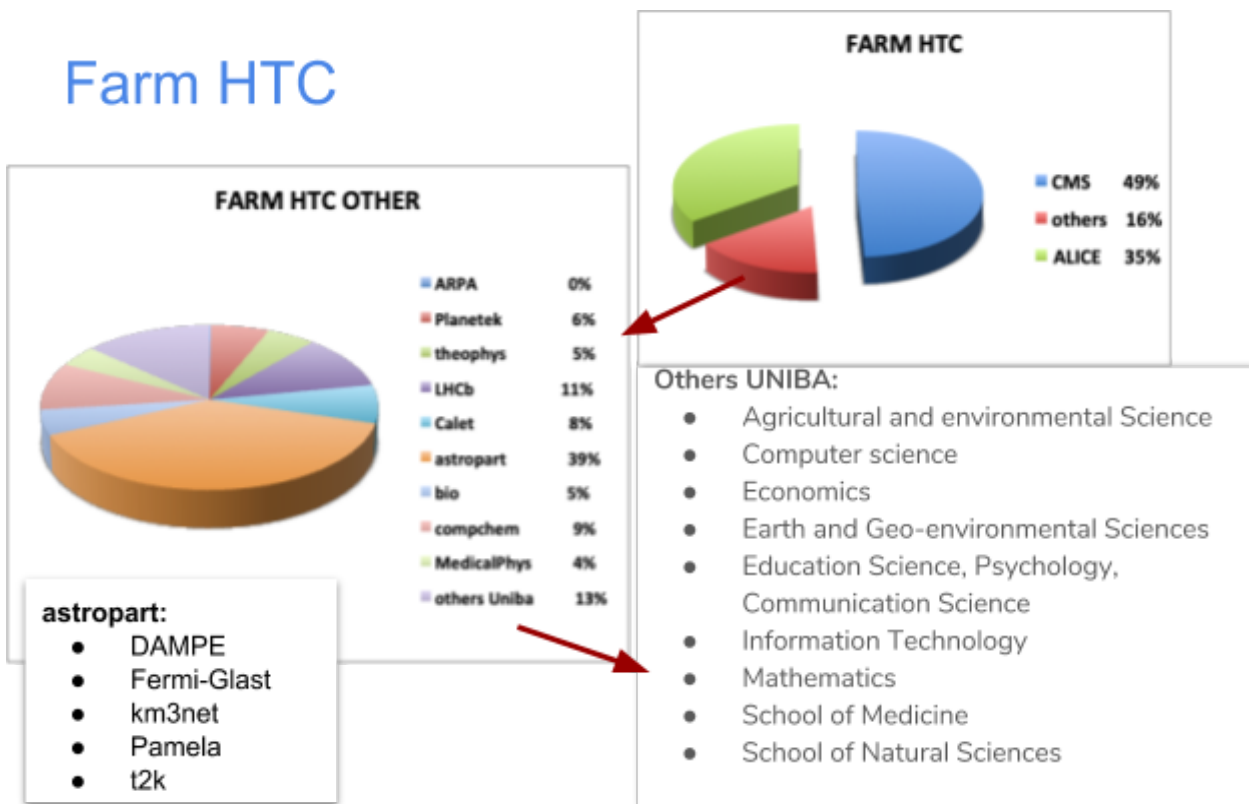


Fig. 3 - Gli utenti e il loro uso della Farm HTC

La Fig. 3 mostra che l'84% della risorse della farm HTC sono usate dagli esperimenti ad LHC ALICE e CMS, mentre il restante 16% è utilizzato per esperimenti di fisica astroparticellare e per attività locali da utenti accademici, dalla Pubblica Amministrazione Locale e da PMI attive sul territorio.

### Cloud@ReCaS-Bari

Il servizio Cloud@ReCaS-Bari è fortemente cresciuto con il passare degli anni, diventando di fatto il secondo tra i servizi disponibili sia in termini di risorse assegnate sia come richieste da parte degli

utenti. L'infrastruttura cloud di ReCaS-Bari, basata su OpenStack, conta su circa 1700 CPU core, 6.7 TB di RAM e 270 TB di spazio disco in replica 3 gestito da CEPH.

Il servizio viene utilizzato

- come ambiente di sviluppo e/o dimostratore per progetti (europei, PON, PRIN e progetti regionali) in cui sono coinvolti ricercatori dell'INFN e/o dell'Università di Bari,
- per fornire risorse virtuali (unità di calcolo e/o storage) per ricercatori, dottorandi e in qualche caso laureandi dell'Università di Bari o dell'INFN,
- per esporre servizi quali siti web istituzionali (Wordpress) e/o di e-learning (Moodle), nonché per fornire storage personale.

<ul style="list-style-type: none"> <li>• 2 controller / network nodes           <ul style="list-style-type: none"> <li>◦ Expose service api endpoints</li> <li>◦ HAproxy Load-Balancer</li> <li>◦ Provide a single entry-point to the cloud services</li> </ul> </li> <li>• 43 compute nodes, 64 cores AMD Opteron, 250GB RAM</li> <li>• 10 Storage Servers</li> <li>• RabbitMQ cluster (3 nodes)</li> <li>• MySQL database</li> <li>• Puppet is used for automatic installation/configuration</li> </ul>
---

Tabella 1 - Le risorse di Cloud@ReCaS-Bari

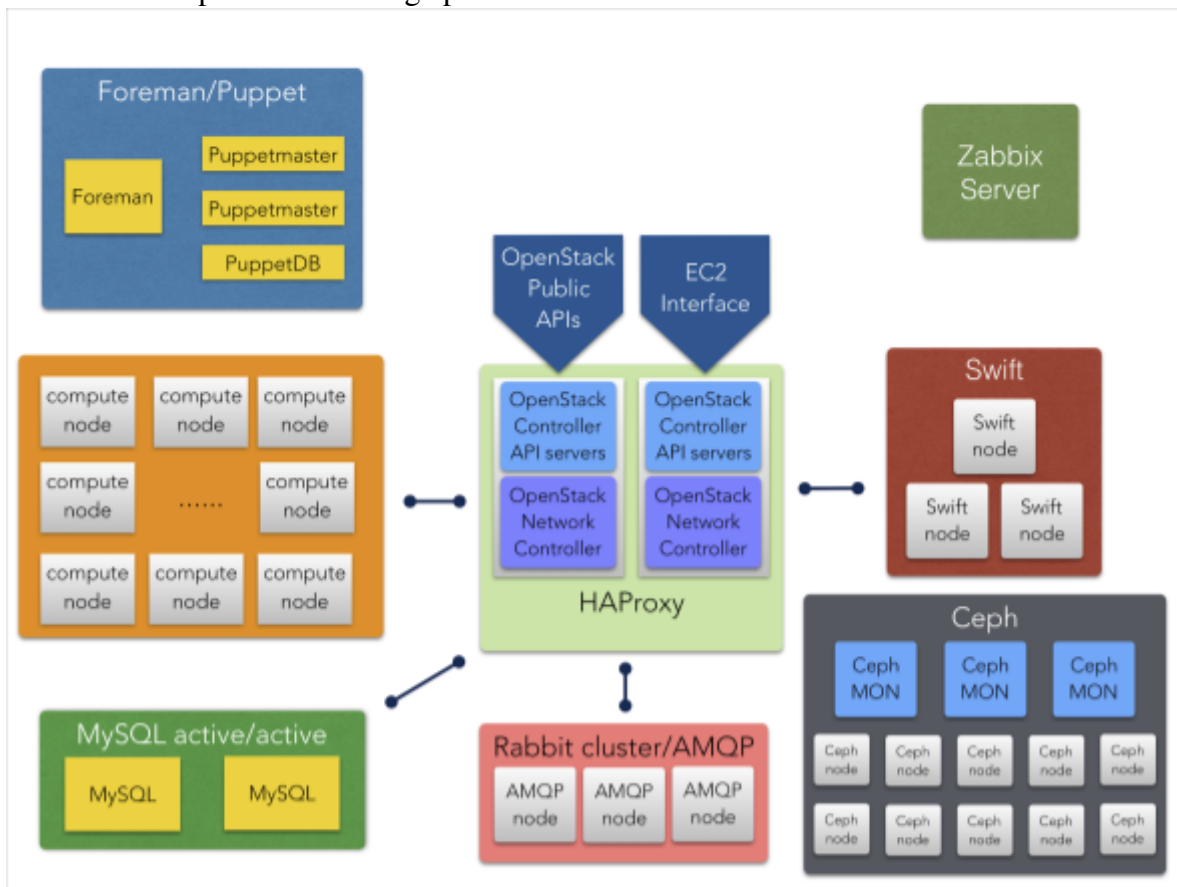


Fig. 4 - Schema a blocchi dell'infrastruttura Cloud di ReCaS-Bari

Sull'infrastruttura cloud è possibile istanziare macchine virtuali (VM) o cluster di macchine virtuali con diverse scelte del sistema operativo Linux: Ubuntu 16.04/18.04 LTS, Debian 8/9, CentOS 6/7, Fedora (su richiesta) oppure con un sistema operativo Windows Server R 2012 (versione trial,



attivabile dall'utente che possieda una licenza). L'utente può anche creare istanze a partire da immagini importate autonomamente.

Di norma le VM vengono istanziate con IP privato, in modo da non essere raggiungibili dall'esterno e risultare meno vulnerabili ad attacchi informatici: in questo caso l'utente potrà accedere alle proprie risorse attraverso una VPN. Laddove invece l'utente avesse la necessità di esporre servizi all'esterno (come un web server o una interfaccia), può fare richiesta di utilizzo di indirizzi IP pubblici.

Per l'accesso alle macchine virtuali è necessario utilizzare una coppia di chiavi pubblica/privata (iniettata nell'istanza al momento della creazione).

La Fig. 4 mostra lo schema a blocchi dell'infrastruttura Cloud@ReCaS-Bari i cui numeri principali sono riassunti nella Tabella 1.

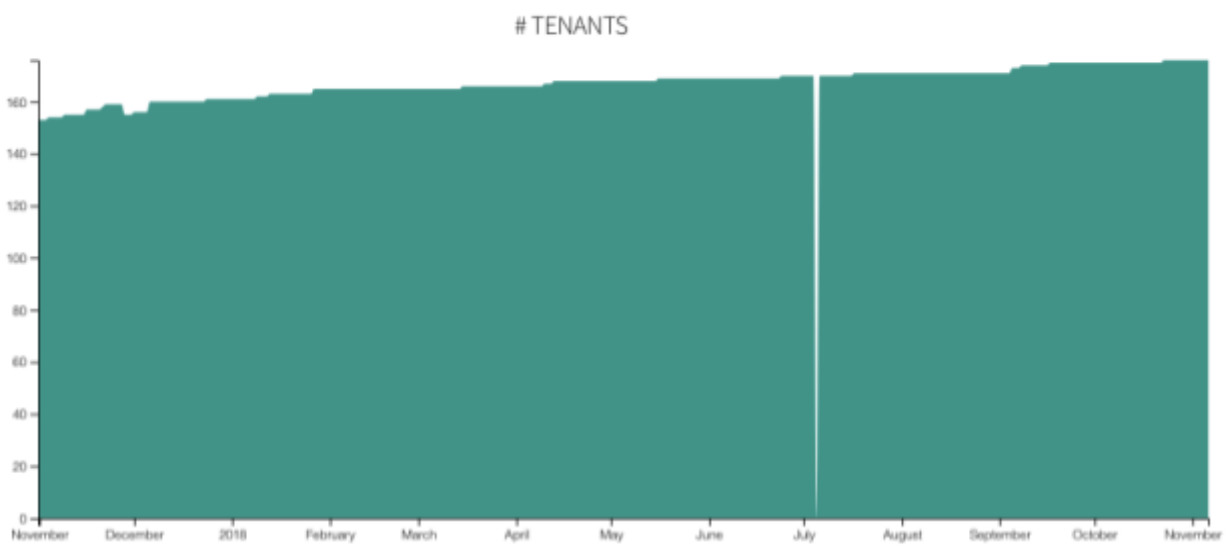


Fig. 5 - L'andamento del numero di "tenants" sulla Cloud@ReCaS-Bar nel periodo di riferimento

La Fig. 5 mostra l'andamento del numero di "tenant" che corrispondono a gruppi di ricerca differenti che utilizzano l'infrastruttura Cloud@ReCas-Bari in maniera indipendente. Ogni gruppo di ricerca può gestire diversi progetti, a ciascuno dei quali afferiscono diversi utenti, e fare uso di diverse macchine virtuali.

Dalla figura si nota come nell'ultimo anno il numero di tenant sia passato da 155 a poco meno di 180 con un incremento del 16%. Sono state comunque effettuate ripetute campagne di individuazione e rimozione delle VM non più utilizzate, lasciando sull'infrastruttura cloud solo effettivamente ancora necessarie agli utenti.

Il numero di progetti e di utenti serviti sono riassunti nella Tabella 2.

L'infrastruttura Cloud@ReCaS-Bari è integrata nella Federated Cloud di EGI a cui offre servizi anche in modalità pay-per-use.

<ul style="list-style-type: none"> <li>● 800 virtual machines</li> <li>● 580 virtual volumes (~ 180 TB)</li> <li>● 244 active projects</li> <li>● 620 users</li> </ul>
--

Tabella 2 - I numeri principali di Cloud@ReCaS-Bari

NGI\_IT > RECaS-BARI    DETAILS    STATUS

Timestamp	Availability	Reliability
2017-11	99.57	100.00
2017-12	90.51	92.14
2018-01	98.40	98.40
2018-02	99.84	99.84
2018-03	98.06	98.06
2018-04	99.77	99.77
2018-05	98.85	98.85
2018-06	99.83	99.83
2018-07	97.55	97.55
2018-08	96.83	96.83
2018-09	98.38	98.38
2018-10	100.00	100.00

Fig. 6 - L'Availability e la Reliability della Infrastruttura Cloud@ReCaS-Bari nel periodo di riferimento

L'integrazione nella EGI Federated Cloud porta come benefit il fatto che l'infrastruttura cloud è costantemente monitorata e confrontata con simili implementazioni a livello europeo. La figura seguente mostra la situazione della "Availability" e della "Reliability" come misurate da EGI (European Grid Infrastructure) della infrastruttura cloud del data center ReCaS-Bari per il periodo di riferimento della presente relazione da cui appare che i due parametri sono stati sempre molto elevati e sicuramente al di sopra della soglia di attenzione di EGI.

La fig. 7 mostra invece le percentuali di uso della infrastruttura Cloud@ReCaS-Bari da parte di progetti/esperimenti (BioVeL<sup>1</sup>, DEEP<sup>1</sup>, EOSC-Hub<sup>1</sup>, CMS-DODAS, GEMACA<sup>2</sup>, INDIGO<sup>1</sup>, MUSICA<sup>3</sup>, OCP<sup>4</sup>, PERSON<sup>2</sup>, PRISMA<sup>4</sup>, SDDS-CNAF, XDC<sup>1</sup>), progetti ESFRI (ELIXIR, Lifewatch), istituzioni accademiche/scientifiche (CNR-IREA, INGV, CNR-IIA, POLIBA e UNIBA), organizzazioni europee (EGI Fed Cloud), e organizzazioni private (Planetek, Reply, Servizi Locali, Sud Sistemi)

<sup>1</sup> Progetto europeo

<sup>2</sup> Progetto regionale

<sup>3</sup> Progetto con il Comune di Bari

<sup>4</sup> Progetto PON

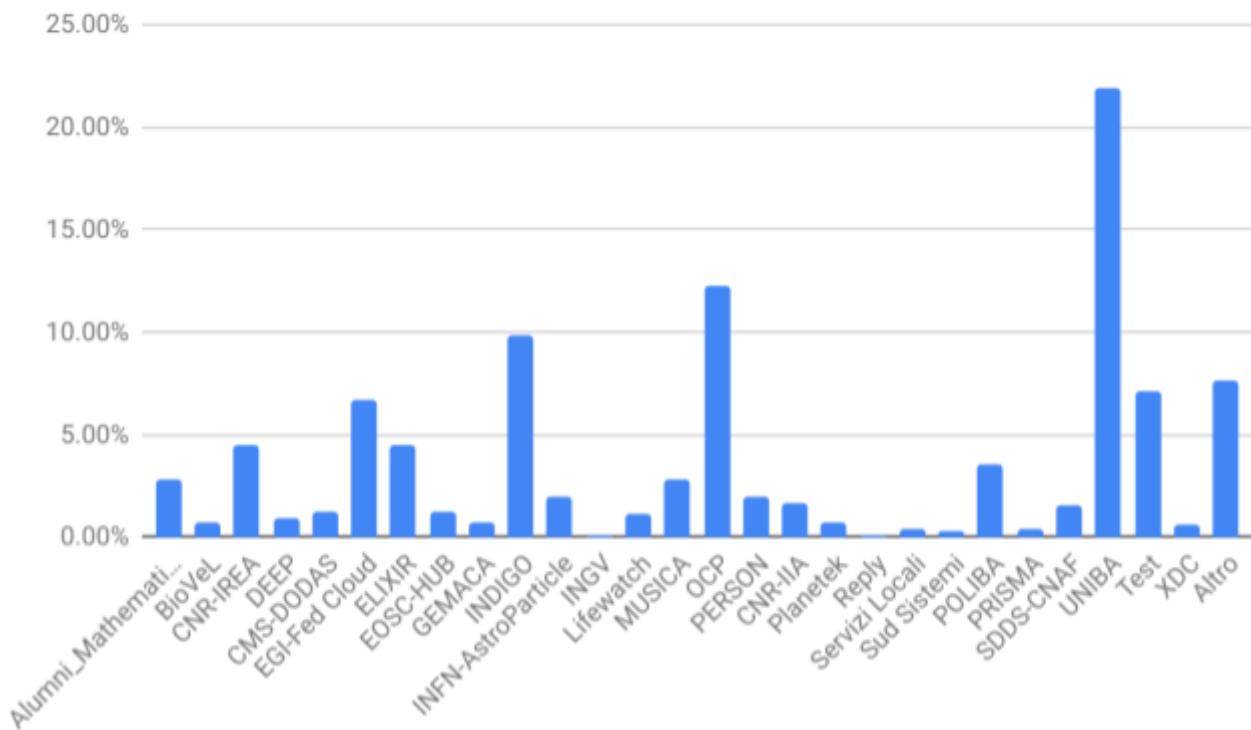


Fig. 7 - Le percentuali di utilizzo della Cloud@ReCaS-Bari da parte di progetti, istituzioni, PMI.

La figura 8, infine, riassume i diversi servizi forniti attraverso l'infrastruttura Cloud@ReCaS-Bari

- Mesos cluster on-demand
  - Heat/Tosca template
  - Ansible roles and playbook
- Galaxy cluster on-demand
- Desktop as a Service (web based)
- RStudio on-demand
- Jupyterhub on-demand
- ShareLaTeX on-demand
- Dropbox-like service based on ownCloud
- GitLab
- Moodle
- Wordpress

Fig. 8 - I principali servizi forniti da Cloud@ReCaS - Bari

### Tier1@ReCaS-Bari

Si tratta di un servizio fornito al TIER1 dell'INFN.

Il TIER1 è un centro di calcolo con un numero maggiore di risorse rispetto ai TIER2 e dispone di una tape library per la memorizzazione di lungo periodo dei dati. Il TIER1 dell'INFN è localizzato

a Bologna e gestito dal CNAF (Centro Nazionale dell'INFN “per la ricerca e lo sviluppo nelle tecnologie informatiche e telematiche”).

Il servizio *Tier1@ReCaS-Bari* si occupa di connettere logicamente alcune delle risorse informatiche di ReCaS-Bari (dell'ordine di 2500 core equivalenti ad una potenza di calcolo di 20 kHSE06) in maniera che appaiano come facenti parte del TIER1.

La sperimentazione condotta tra il CNAF e Bari è riuscita a mettere a punto una tecnica per l'utilizzo da parte del TIER1 di risorse di calcolo ovunque presenti in Italia, come per esempio quelle fornite dalle cloud pubbliche.

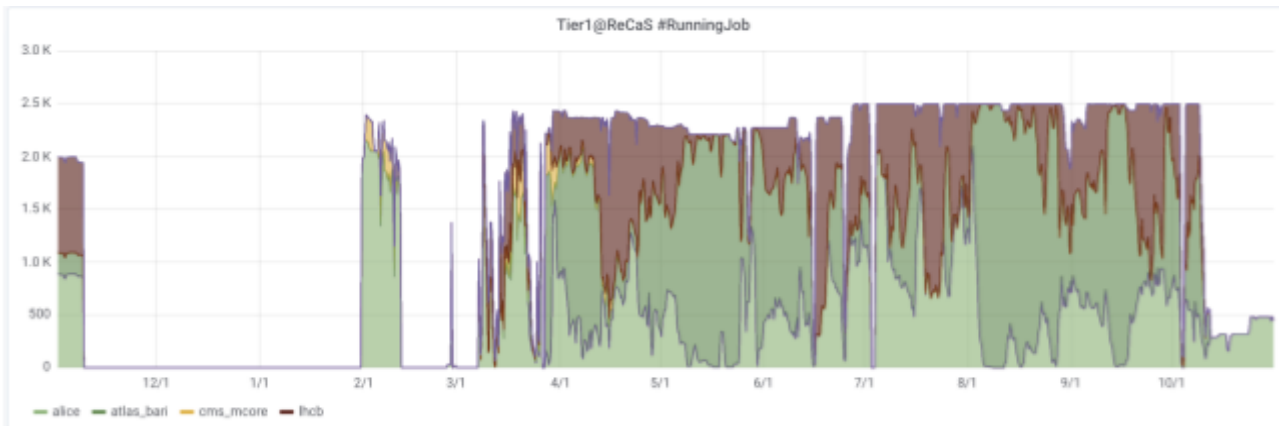


Fig. 9 - Andamento della occupazione delle risorse da parte del TIER1 nel periodo di riferimento. Si osservi che nel periodo dal 10 nov 2017 al 31 gen 2018 il TIER1 è stato in manutenzione e le sue risorse sono state aggiunte alla farm HTC

Alla fine del periodo di riferimento, il 10 ottobre 2018, il servizio *Tier1@ReCaS-Bari*, avendo raggiunto i suoi obiettivi, è stato in parte smantellato per recuperare le risorse informatiche necessarie per lo studio di fattibilità del progetto “Analisi satellitare delle deformazioni degli edifici scolastici italiani”.

### Cluster HPC

Il cluster dispone di fino a 400 core fisici (800 considerando l'hyperthreading). È composto da 20 server ciascuno con 256GB di RAM, connessi tra loro con connessioni a bassa latenza (InfiniBand), ciascuno dei quali è dotato di una scheda acceleratrice grafica (GPU) NVIDIA K40. Ciascun server di calcolo, può accedere, alla velocità di 10 Gbps, a tutto lo spazio disco di ReCaS-Bari (circa a

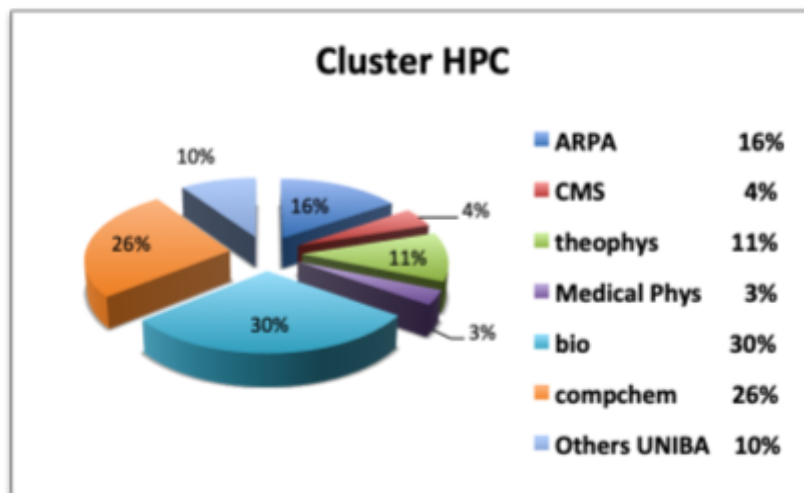


Fig. 10 - La ripartizione dell'uso delle risorse del cluster HPC tra le diverse categorie di utenti

3800 TB in singola replica più 220 TB in replica due) gestito attraverso il file system distribuito GPFS.

Il sistema operativo è stato migrato da CentOS 6 a CentOS 7 durante il periodo di riferimento. Per la gestione delle code viene utilizzato Torque/MAUI (PBS).

La Fig. 10 mostra come sono ripartite le risorse tra i diversi utenti del cluster HPC. Significativo è l'uso che ne

viene fatto dalla VO “bio” (essenzialmente attività legata ad ELIXIR, Chimica computazionale ed ARPA-Puglia).

L’occupazione delle risorse è mostrata in Fig. 11. Al contrario della farm HTC, il cluster HPC non è quasi mai saturato. Questo è in parte legato al gestore delle code che non è in grado di assicurare l’esecuzione di applicazioni critiche come quelle che giornalmente devono essere fatte girare da ARPA-PUGLIA. Per soddisfare questa criticità si preferisce tenere alcune risorse sempre disponibili, e quindi non occupate.



Fig. 11 - Occupazione del Cluster HPC nel periodo di riferimento

### ***Disk storage***

ReCaS-Bari pur disponendo di 3,8 PB per immagazzinamento di dati su disco in singola copia e 220 TB in doppia copia (per conservare dati difficilmente riproducibili o recuperabili da altre parti) e nonostante le ripetute campagne presso gli utenti per far cancellare i dati non più utili o di archiviare quelli che vengono acceduti molto di rado, ha ormai raggiunto una condizione di saturazione. I principali utilizzatori di spazio disco sono naturalmente gli esperimenti ALICE e CMS. Tutto lo storage è gestito mediante il file system GPFS.

### ***Tape storage***

La tape library IBM, System Storage TS3500, in dotazione a ReCaS-Bari è in grado di archiviare su nastro nominalmente 2500 TB (3900 TB di fatto) di dati.

La tape library viene utilizzata per due funzioni:

- replica su nastro dei file presenti anche su disco: in questo modo si dispone di una copia di backup che può essere utilizzata in caso di perdita di dati per malfunzionamenti del disco;
- archiviazione di lungo periodo di alcuni file acceduti molto di rado, che, una volta archiviati, possono essere cancellati dal disco, liberando spazio prezioso.

### ***Housing of IT resources***

Il data center ReCaS-Bari può ospitare al suo interno risorse computazionali di terzi.

In questo modo l'utente può sfruttare l'infrastruttura del data center, come il sistema di UPS e il generatore ausiliario, i sistemi di condizionamento, supervisione e allarmistica, i collegamenti di rete e, volendo, il controllo da parte degli operatori sulle risorse ospitate.

IN questo modo l’utente può evitare di realizzare l'infrastruttura necessaria per mantenere in esercizio le proprie macchine e di sostenerne i costi di gestione. Infine, localizzando le proprie

risorse all'interno di ReCaS-Bari, è possibile sfruttare altri servizi messi a disposizione dal data center, come quello di backup e archiviazione dati.

Usufruiscono di questo servizio

- il GARR, con 4 rack della Cloud Federata, che ha scelto una configurazione in cui non c'è alcuna connessione fisica e/o logica con il data center ReCaS-Bari. I due oggetti, pur essendo fisicamente nella stessa stanza si comportano come due entità completamente separate ed indipendenti: la loro unica connessione è attraverso la rete ricerca.
- il CNR-IBIOM, con un rack, che invece ha scelto una maggiore integrazione tra le sue risorse e il Centro.

## Le risorse umane

Nella tabella 3 è elencato il personale impegnato a vario titolo sulle attività che si svolgono nel centro ReCaS-Bari

<b>1. Personale con contratto di lavoro dipendente</b>		
1.	Marica Antonacci	TI tecnologo INFN
2.	Domenico Diacono	TI tecnologo INFN
3.	Giacinto Donvito	TI tecnologo INFN
4.	Riccardo Gervasoni	TI CTER INFN
5.	Alessandro Italiano	TI CTER INFN
6.	Ferdinando Lardo	TI Cat C UNIBA
7.	Alfonso Monaco	TI tecnologo INFN
8.	Stefano Nicotri	TI tecnologo INFN
9.	Vincenzo Spinoso	TI tecnologo INFN
10.	Roberto Valentini	TI CTER INFN
<b>2. Assegnisti di ricerca</b>		
1.	Due selezioni (una UNIBA e una INFN) in corso di espletamento	
<b>3. Management</b>		
1.	Roberto Bellotti	TI PO UNIBA
2.	Domenico Elia	TI Ricercatore INFN
3.	Giorgio Pietro Maggi	TI PO POLIBA

Tabella 3 - Elenco del personale Impegnato a vario titolo sulle attività del Centro

### *Skill ed expertise*

- Gestione di una infrastruttura di calcolo complessa con strumenti come the Foreman, Puppet, GitLab, OpenProject, Zabbix, Grafana, DESIGO Insight.
- Gestione di sistemi di storage e file system distribuiti quali GPFS, Ceph, Swift, GlusterFS, Lustre
- uso di molteplici batch system, fra cui, HTCondor, SLURM, PBS, utilizzati per la gestione di sistemi di calcolo parallelo e distribuito di grandi dimensioni.
- gestione e supporto nell'utilizzo di tutti i più comuni database management system, di tipo SQL e NoSQL/NewSQL.
- uso di sistemi HPC per il calcolo parallelo E/O di GPU per l'accelerazione di specifici calcoli scientifici
- virtualizzazione e uso di container,
- dispiegamento e amministrazione di infrastrutture cloud IaaS complesse con servizi in alta disponibilità,
- attivazione di servizi PaaS e SaaS.
- gestione e configurazione di framework e applicativi all'avanguardia quali OpenStack, VMWare, Mesos (Marathon/Chronos), Kubernetes, ...
- utilizzo in produzione di strumenti e piattaforme per l'analisi di big data quali Spark e Hadoop



## Le attività

Le attività di ReCaS-Bari si possono classificare nelle seguenti linee:

1. attività a carattere amministrativo
2. gestione tecnica del centro
3. attività di R&D (Ricerca e Sviluppo)

### *Attività a carattere amministrativo*

#### *Stipula convenzioni e/o accordi*

Nel corso del periodo di riferimento sono stati stipulati i seguenti accordi e/o convenzioni riguardanti il Centro ReCaS-Bari

- UNIBA (Dipartimento Interateneo di Fisica)
  - EGI-VO-OLA ReCaS-Bari\_Pilot-Vito (EGI pay-per-use)
  - EGI-VO-OLA\_ReCaS-Bari\_Pilot3-WUR-ALTERRA (EGI pay-per-use)
  - CONVENZIONE OPERATIVA TRA Dipartimento Interateneo di Fisica “M. Merlin” - Università degli Studi di Bari e Istituto per Rilevamento Elettromagnetico dell’Ambiente - Consiglio Nazionale delle Ricerche
- INFN
  - ACCORDO DI RICERCA COLLABORATIVA (rif. cod. TTB\_18BA\_019) tra INFN e INGV

E’ stato inoltre avviato l’iter procedurale per il rinnovo/la stipula

- della convenzione tra Planetek Italia srl e UNIBA
- dell’accordo di collaborazione tra INFN, UNIBA e ARPA-Puglia
- dell’accordo di collaborazione tra il Dipartimento Interateneo di Fisica di Bari e il Dipartimento di Scienze Fisiche della Terra e dell’Ambiente dell’Università degli Studi di Siena
- dell’accordo di collaborazione tra INFN, UNIBA e CNR-IIA

#### *Emissione Ordinativi*

Spese per manutenzione.

	IMPORTO	Istituzione che ha sostenuto la spesa
Rinnovo del contratto di manutenzione per il 2018 e il 2019 per il sistema di UPS	<b>€29,280.00</b>	INFN
Rinnovo del contratto di manutenzione della tape library per l’anno 2018 (comprende manutenzione di xx server SuperMicro e Switch Huawei)	<b>€24,252.38</b>	INFN
Rinnovo del contratto di manutenzione della tape library per l’anno 2019 (comprende manutenzione di xx server SuperMicro e Switch Huawei)	<b>€24,252.38</b>	INFN



Rinnovo del contratto di manutenzione biennale del sistema di supervisione	<b>€7,930.00</b>	INFN
Acquisto nuova versione delle Licenze DESIGO	<b>€3,050.00</b>	UNIBA
Aggiornamento del software di supervisione alla nuova release DESIGO	<b>€4,880.00</b>	INFN
Manutenzione cabina elettrica e gruppo elettrogeno	<b>€790.85</b>	UNIBA
Manutenzione Ascensore	<b>€1,029.49</b>	UNIBA
Pulizia uffici	<b>€3,000.00</b>	UNIBA
Manuetenziome bombole gas spegnimento incendi sala CED	<b>€2,000.00</b>	UNIBA
Sostituzione batterie centraline sistema di supervisione	<b>€488.00</b>	INFN
Sostituzione batteria UPS	<b>€962.00</b>	INFN
Sostituzione switch di rete	<b>€427.00</b>	INFN
<b>Totale</b>	<b>€102,342.10</b>	

*Riscossioni contributi su convenzione e/o accordi.*

È stato incassato il contributo di CNR-IBBE sull'ACCORDO DI RICERCA COLLABORATIVA N° TTB\_15BA\_044 per un importo di 45.000,00 euro.

E' stata emessa nota di debito (non ancora incassata) ad INGV sull'ACCORDO DI RICERCA COLLABORATIVA (rif. cod. TTB\_18BA\_019) tra INFN e INGV per un importo di 25.000,00 euro

Sono state infine formulate due offerte per attività di ricerca collaborativa

1. Preventivo per la fornitura di ricerca dell'Istituto Nazionale di Fisica Nucleare – Sezione di Bari, relativamente al progetto OMICS4FOOD, cod. 1JLZKD0, POR Puglia FESR-FSE 2014-2020 Azione 1.6., Avviso "Innonetwork", A.D. n.124 del 16/10/2017, CUP B37H1700480000 per un importo di 45.000,00 euro
2. Offerta per la fornitura di ricerca collaborativa dell'Istituto Nazionale di Fisica Nucleare – Sezione di Bari relativamente al progetto “Sistema Integrato di monitoraggio, allerta e prevenzione dello stato di CONservazione di Habitat ed ECOsistemi in aree interne e costiere protette e da proteggere (COHECO)” - a valere sul bando Innonetwork – – Aiuti a sostegno alle Attività di R&S, Regione Puglia per un importo di 28.000,00 euro.

### ***Gestione tecnica del centro***

L'attività di gestione tecnica del data center assorbe gran parte della manodopera disponibile. Un'idea della complessità dell'attività può essere desunta dai seguenti numeri:

1. diciassette riunioni generali di gruppo.
2. 235 ticket sottomessi e 105 risolti

3. 67 nuovi account creati, sia per l'utilizzo di risorse cloud o della Farm HTC.

A parte l'attività ordinaria, nel periodo di riferimento sono state svolte le seguenti attività:

1. spostamento delle risorse TIER1@ReCaS-Bari nella Farm HTC durante il periodo di down del TIER1 tra la fine del 2017 e l'inizio del 2018 ed la loro restituzione al TIER1 alla fine del periodo.
2. Fornito una soluzione alternativa per l'acquisizione dei dati per gli esperimenti DAMPE ed EEE e per il backup dei dati del sistema informativo durante il periodo di down del TIER1.
3. Applicazioni delle patch per superare diverse vulnerabilità tra cui la CVE-2015-7547.
4. Passaggio al sistema operativo CentOS 7 da Scientific Linux 6
5. Configurazione di un cluster ad-hoc per il Proof of Concept per l'attività di monitoring degli edifici pubblici ad uso di CNR/ASI

### **Ricerca e Sviluppo**

L'attività di ReCaS-Bari non si esaurisce con la gestione del data center. Accanto a questa si inserisce, in un modo del tutto naturale, una attività di R&D su problematiche di calcolo distribuito, le cui soluzioni vengono immediatamente implementate nella gestione del data center realizzando un corto circuito virtuoso che permette di mettere immediatamente a disposizione degli utenti le innovazioni individuate nel programma di R&D.

L'attività di R&D viene sviluppata all'interno di progetti di ricerca europei, nazionali o regionali, Nel periodo di riferimento l'attività di R&D è stata svolta nell'ambito di tre progetti di ricerca europei e di due progetti regionali:

1. **DEEP-Hybrid DataCloud (DEEP-H)** è un progetto H2020 coordinato da CSIC(Spagna) che ha come obiettivo lo sviluppo e l'implementazione di tecnologie cloud che consentano un accesso facile e trasparente a risorse eterogenee in ambito cloud e HPC. Il progetto sta lavorando per sviluppare strumenti in grado di realizzare in maniera automatica ambienti operativi per il deep e machine learning in diversi contesti applicativi. ReCaS-Bari coordina a livello INFN le attività tecniche, partecipa alla implementazione del layer PaaS del progetto coordinando il relativo work package e fornisce uno degli ambienti di sviluppo e test (testbed).
2. **eXtreme-DataCloud, (XDC)** è un progetto H2020 coordinato dall'INFN, in cui la Sezione di Bari dell'INFN è coinvolta in modo coordinato con il CNAF e con la Sezione di Padova. Il progetto ha l'obiettivo di implementare un set di servizi per federare risorse di storage, in modo da poter supportare scalabilità fino all'ordine delle centinaia di PB. Contestualmente, si vogliono supportare casi d'uso molto diversi fra di loro, che vanno dalla fisica delle alte energie fino agli studi clinici, passando per l'astrofisica e le ricerche basate sull'uso di laser a raggi X. ReCaS-Bari si occupa del coordinamento tecnico del progetto e del supporto agli use-case di WLCG e della bioinformatica.
3. **EOSC-hub** è un progetto H2020 che è iniziato il 1 gennaio 2010 e durerà tre anni. Il progetto intende realizzare la prima implementazione concreta di una infrastruttura aperta di cloud computing per la scienza (European Open Science Cloud: EOSC). ReCaS-Bari ha un ruolo rilevante nel progetto, infatti è coinvolto nelle attività di coordinamento tecnico del progetto, di supporto e mantenimento della PaaS di INDIGO, dell'implementazione di tool di analisi per la comunità di fisica delle alte energie e nelle operation dell'infrastruttura.
4. **PERSON** - "PERvasive game for perSONalized treatment of cognitive and functional deficits associated with chronic and Neurodegenerative diseases" è un progetto finanziato dalla Regione Puglia che si è concluso a ottobre 2017 ed ha realizzato un ambiente ICT diagnostico e terapeutico, basato su un sistema pervasivo e innovativo, con il quale

effettuare diagnosi precoce di malattie neurodegenerative e supportare processi di terapia personalizzata nel caso in cui siano rilevate situazioni patologiche.

5. DECISION - Sistema informativo in grado di analizzare i big data e di utilizzare il machine learning per fornire informazioni analitiche e previsionali di supporto alle decisioni a partire da immagini telerilevate.

Progetto	stato appr	INFN	UNIBA	POLIBA
<b>Progetti attivi</b>				
DEEP	approvato	X		
XDC	approvato	X		
EOSC-hub	aprovato	X		
PERSON	approvato	X	X	
DECISON	approvato	X	X	

La continuità del programma di R&D richiede che sempre nuovi progetti subentrino a quelli appena conclusi: ecco perchè particolare attenzione è stata rivolta alla preparazione e sottomissione di nuovi progetti.

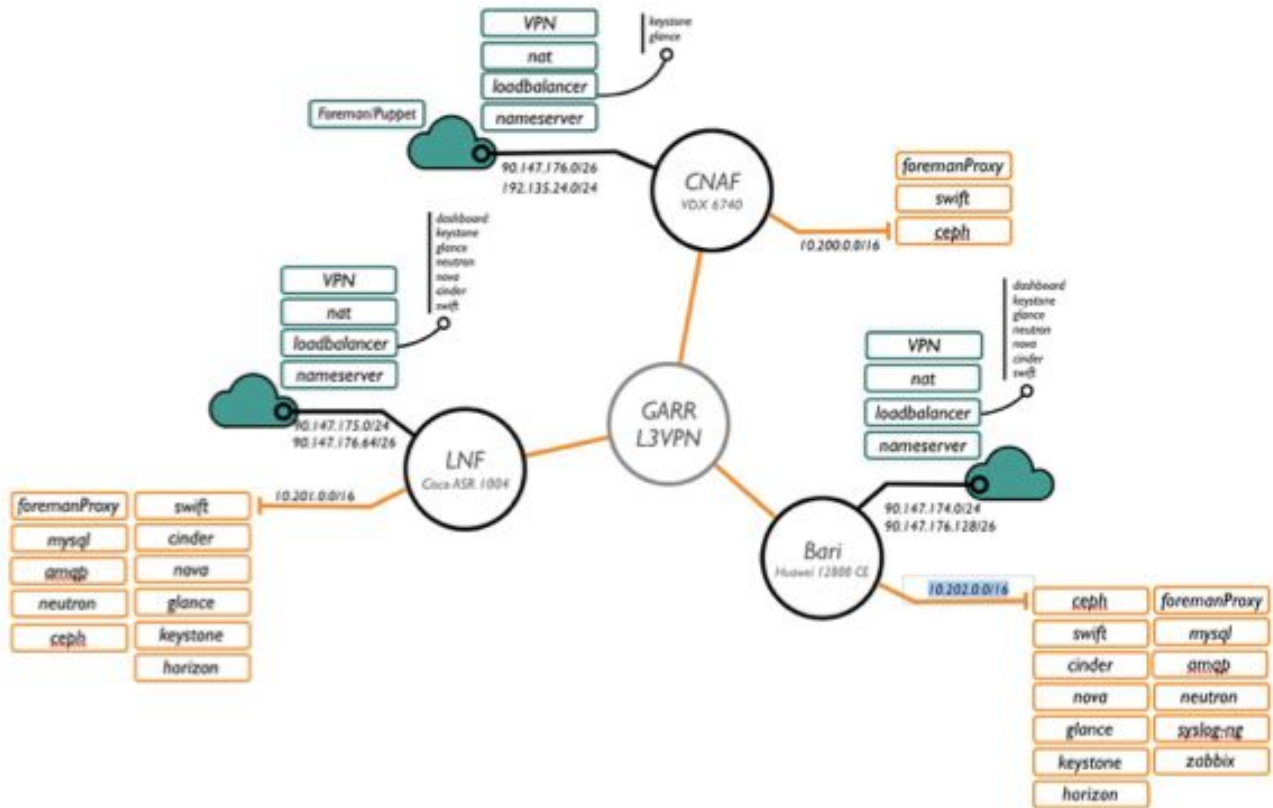
La tabella seguente mostra i progetti, preparati con il contributo di ReCaS-Bari, nel periodo di riferimento.

Progetto	stato appr	INFN	UNIBA	POLIBA
<b>Progetti PON avviso 1735 (Chiusura bando il 9 nov 2017) NB cofinanziamenti 50%</b>				
RPASInAIR - Integrazione dei Sistemi Aeromobili a Pilotaggio Remoto nello spazio aereo non segregato per servizi	approvato	€149,903.06	€134,000.00	€51,500.00
CLOSE - Close to the Earth	approvato		€ 237.500,00	
3D-CISMA - 3D City Smart MANager	non appr	X	X	
TEBAKA - Sistema per acquisizione conoscenze di base Del terroto	non finanziato		X	
BIO-D - SVILUPPO DI BIOMARCATORI DIAGNOSTICI PER LA MEDICINA DIPRECISIONE E LA TERAPIA PERSONALIZZATA	non finanziato	X	X	X
<b>Progetti PRIN 2017 (scadenza del bando 29 marzo 2018)</b>				
Studio e modellazione di sistemi fisici complessi. (PI Roberto Bellotti)	in valutazione	X	X	
Earth observation attraverso studio di immagini da satellite (PI.Riccardo Lanari)	in valutazione	X	X	
<b>Progetti PON avviso 424 (chiusura bando 15 giugno 2018)</b>				

<b>IBiSCo</b> - Infrastruttura per Big data e Scientific Computing (infrastruttura DHTCS (ora IPCEI-HPC-BDA) - Distributed High Throughput Computing and Storage)	in valutazione	X	X	
<b>CNRBiOmics</b> - Centro Nazionale di Ricerca in Bioinformatica per le scienze "Omiche" (infrastruttura ELIXIR)	in valutazione	X	X	
<b>LifeWatchPLUS</b> - LifeWatch, infrastruttura di eScience per la ricerca su biodiversità ed ecosistemi – Progetto di rafforzamento infrastrutturale (infrastruttura LIFEWATCH-ERIC)	in valutazione	X		
<b>Bando AIFA 2017</b>				
Reducing Antibiotic Misuse in Neonates - A Big Data-driven, machine learning approach to predict neonatal sepsis avoiding antibiotic misuse	non ammesso	X	X	
<b>Call europee</b>				
ICT-11-2018-2019 Call TOPIC : “HPC and Big Data enabled Large-scale Test-beds and Applications”	non appr	X		
SC1-DTH-07-2018 Call TOPIC “Exploiting the full potential of in-silico medicine research for personalised diagnostics and therapies in cloud-based environments”	non appr	X		
INFRAEOSC-05-2018-2019-b:	in valutazione	X		
ICT-11-2018-2019 (IoT/Streaming):	in valutazione	X		

### ***INFN Corporate Cloud (INFN-CC)***

ReCaS-Bari gioca un ruolo importante nel progetto INFN-CC. Si tratta di un progetto della Commissione Calcolo e Reti dell’INFN che ha come obiettivo quello di dotare l’INFN di una infrastruttura distribuita e solida, in grado di fornire in maniera nativa HA e disaster-recovery, indispensabili per tutta una classe di casi d'uso critici e di grande importanza per l'INFN.



INFN-CC accoppia strettamente alcune installazioni omogenee di OpenStack. Si presenta come un dominio amministrativo singolo, benché distribuito e punta a realizzare la replica delle risorse, la distribuzione e l'alta disponibilità, entrambe per i servizi di rete e per le applicazioni utente. Alla realizzazione di INFN-CC, che è ormai operativa da alcuni mesi, ReCaS-Bari, che ospita uno dei tre siti, ha contribuito nella sia nella fase di progettazione che nella sua realizzazione.

## Gli utenti

ReCaS-Bari supporta:

- gli esperimenti di fisica e in parte le attività di Fisica Teorica della Sezione INFN e del Dipartimento Interateneo di Fisica
- numerosi progetti sia europei, che nazionali e regionali (alcuni ormai conclusi)
- attività di altri istituti di ricerca: Dipartimenti di UNIBA, istituti del CNR, Istituto di Geofisica e Vulcanologia, Politecnico di Bari...
- progetti di scienze della vita: bioinformatica, fisica medica, biodiversità (progetti ESFRI)
- le analisi di dati satellitari (con aziende e spin off)
- le previsioni sulla qualità dell'aria nella Regione Puglia (ARPA)
- attività di R&D di PMI locali.
- il progetto MUSICA ( Sistema innovativo di Monitoraggio Urbano) del Comune di Bari-

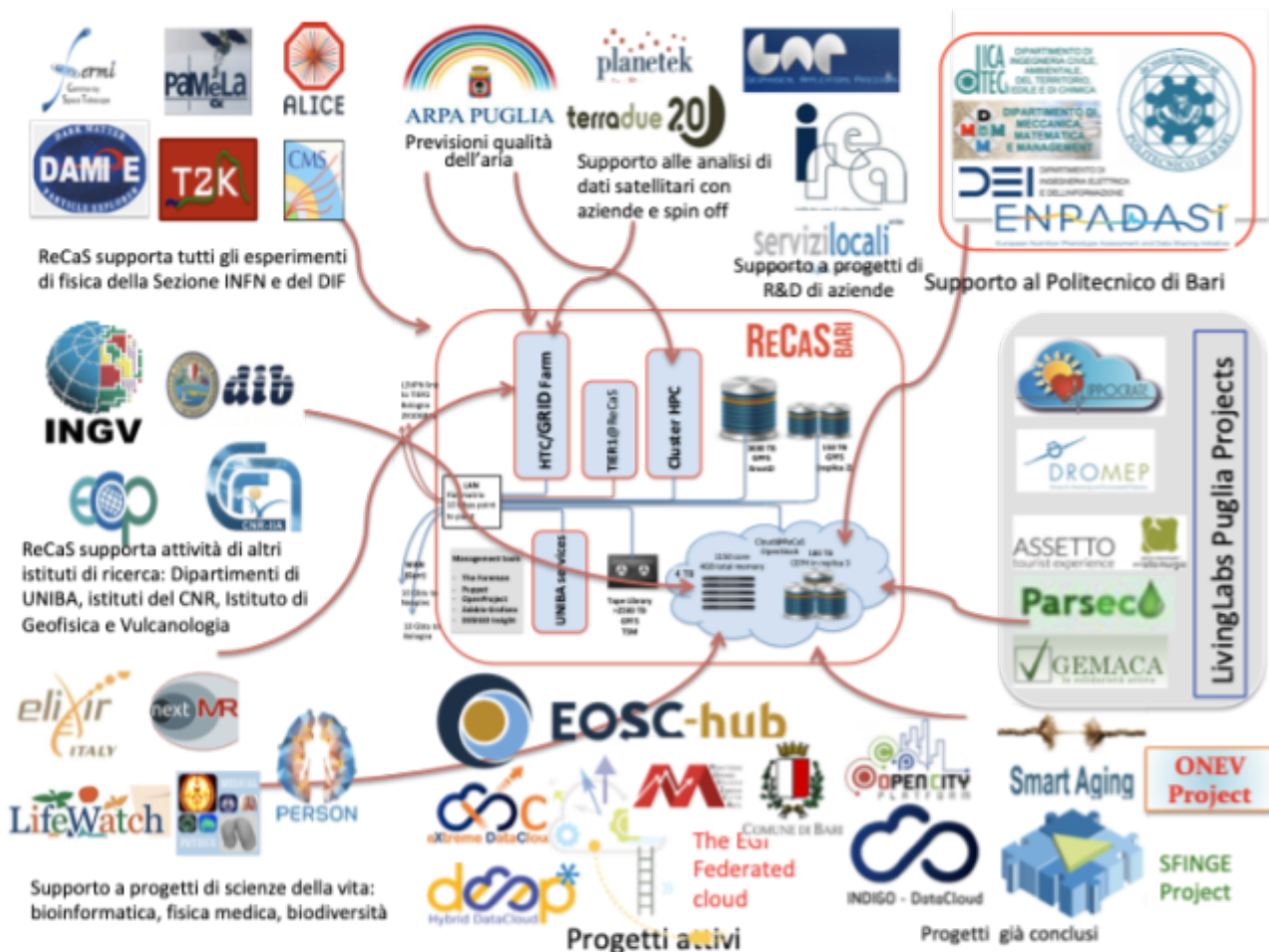


Fig 12 - Alcuni degli utenti di ReCaS-Bari

In un sondaggio parziale condotto all'inizio del 2018 è emerso che ReCaS-Bari era stato citato nei ringraziamenti di almeno 30 lavori scientifici di comunità che non comprendono gli esperimenti di riferimento (ALICE e CMS) o in cui un membro di ReCaS-Bari è firmatario del lavoro.



## Pubblicazioni

### *Papers:*

Alfonso Monaco, Gianluca Sforza, Nicola Amoroso, Marica Antonacci, Roberto Bellotti, Marina de Tommaso, Pierpaolo Di Bitonto, Eugenio Di Sciascio, Domenico Diacono, Eleonora Gentile, Anna Montemurno, Michele Ruta, Antonio Ulloa, Sabina Tangaro - **The PERSON project: a serious brain-computer interface game for treatment in cognitive impairment**

Health and Technology (September 2018)

DOI: <https://doi.org/10.1007/s12553-018-0258-y>

D. Salomoni, I. Campos, L. Gaido, J. Marco de Lucas, P. Solagna, J. Gomes, L. Matyska, P. Fuhrman M. Hardt, G. Donvito, L. Dutka, M. Plociennik, R. Barbera, I. Blanquer, A. Ceccanti, E. Cetinic, M. David, C. Duma, A. López-García, G. Moltó, P. Orviz, Z. Sustr, M. Viljoen, F. Aguilar, L. Alves, M. Antonacci, L. A. Antonelli, S. Bagnasco, A. M. J. J. Bonvin, R. Bruno, Y. Chen, A. Costa, D. Davidovic, B. Ertl, M. Fargetta, S. Fiore, S. Gallozzi, Z. Kurkcuoglu, L. Lloret, J. Martins, A. Nuzzo, P. Nassisi, C. Palazzo, J. Pina, E. Sciacca, D. Spiga, M. Tangaro, M. Urbaniak, S. Vallero, B. Wegh, V. Zaccolo, F. Zambelli **INDIGO-DataCloud: a Platform to Facilitate Seamless Access to E-Infrastructures**

Journal of Grid Computing (August 2018)

DOI: <https://doi.org/10.1007/s10723-018-9453-3>

Alfonso Monaco , Anna Monda , Nicola Amoroso, Alessandro Bertolino, Giuseppe Blasi, Pasquale Di Carlo, Marco Papalino, Giulio Pergola, Sabina Tangaro , Roberto Bellotti - **A complex network approach reveals a pivotal substructure of genes linked to schizophrenia.** PLoS ONE 13(1): e0190110. <https://doi.org/10.1371/journal.pone.0190110>.

Nicola Amoroso, Marianna La Rocca, Stefania Bruno, Tommaso Maggipinto, Alfonso Monaco, Roberto Bellotti and Sabina Tangaro - **Complex networks reveal early MRI markers of Parkinson's disease.** Med Image Anal. 2018 Aug;48:12-24. doi: 10.1016/j.media.2018.05.004. Epub 2018 May 17

Amoroso N, Diacono D, Fanizzi A, La Rocca M, Monaco A, Lombardi A, Guaragnella C, Bellotti R, Tangaro S; **Alzheimer's Disease Neuroimaging Initiative. Deep learning reveals Alzheimer's disease onset in MCI subjects: Results from an international challenge.** J Neurosci Methods. 2018 May 15;302:3-9. doi: 10.1016/j.jneumeth.2017.12.011. Epub 2017 Dec 26.

Balech B and Monaco A and Perniola M and Santamaria M and Donvito G and Vicario S and Maggi G and Pesole G - **DNA Multiple Sequence Alignment Guided by Protein Domains: The MSA-PAD 2.0 Method.,**

Methods in molecular biology (Clifton, N.J.) (2018)

doi = {10.1007/978-1-4939-7683-6\_13}

DAMPE Collaboration- **Direct detection of a break in the teraelectronvolt cosmic-ray spectrum of electrons and positrons.**},

Nature n. 12 (2017)

doi = {10.1038/nature24475}

Maggipinto, T. and Bellotti, R. and Amoroso, N. and Diacono, D. and Donvito, G. and Lella, E. and Monaco, A. and Antonella Scelsi, M. and Tangaro, S.- **DTI measurements for Alzheimer's classification**

Physics in Medicine and Biology {2017}, volume = {62}, number = {6}, pages = {2361-2375}

Aiftimiei, D.C. and Antonacci, M. and Bagnasco, S. and Boccali, T. and Bucchi, R. and Caballer, M. and Costantini, A. and Donvito, G. and Gaido, L. and Italiano, A. and Michelotto, D. and Panella, M. and Salomoni, D. and Vallero, S.- **Geographically distributed Batch System as a Service: The INDIGO-DataCloud approach exploiting HTCondor**

Journal of Physics: Conference Series (2017), volume = {898}, number = {5}

Lopez Garcia, A. and Zangrando, L. and Sgaravatto, M. and Llorens, V. and Vallero, S. and Zacco, V. and Bagnasco, S. and Taneja, S. and Dal Pra, S. and Salomoni, D. and Donvito, G.- **Improved Cloud resource allocation: How INDIGO-DataCloud is overcoming the current limitations in Cloud schedulers**

Journal of Physics: Conference Series {2017}, volume = {898}, number = {9}

### *Conference Papers:*

Marica Antonacci, Alberto Brigandì, Miguel Caballer, Giacinto Donvito, Germán Moltó, Davide Salomoni. **Piattaforme per l'analisi di Big Data istanziate on-demand tramite la PaaS di INDIGO-DataCloud**

Garr Conference 2017, THE DATA WAY TO SCIENCE, 15-17 November 2017, CA' FOSCARI UNIVERSITY OF VENICE

DOI: 10.26314/GARR-Conf17-proceedings-02

D. Spiga, M. Antonacci, T. Boccali, A. Ceccanti, D. Ciangottini, G. Donvito, C. Duma, D. Salomoni. **A container-based solution to generate HTCondor Batch Systems on demand exploiting heterogeneous Clouds for data analysis**

2017 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)

DOI: 10.1109/NSSMIC.2017.8533143

D. Elia, G. Vito, S. Bagnasco, A. Crescente, G. Donvito, A. Franco, S. Lusso, D. Mura, S. Piano, G. Platania on behalf of the ALICE Collaboration - **A Dashboard for the Italian Computing in ALICE**

CHEP 2016 San Francisco - Proceeding - Published online: 23/11/2017 -

<http://iopscience.iop.org/volume/1742-6596/898?pageTitle=IOPscience>

<http://iopscience.iop.org/article/10.1088/1742-6596/898/9/092054>

### *Preprint:*

Marco Antonio Tangaro, Giacinto Donvito, Marica Antonacci, Matteo Chiara, Pietro Mandreoli, Graziano Pesole, Federico Zambelli. **Laniakea: an open solution to provide Galaxy "on-demand" instances over heterogeneous cloud infrastructures.**

DOI: <https://doi.org/10.1101/472464>



***Presentations:***

Marica Antonacci-Oral Presentation: **Digital archive as a service: automatic deployment of an Invenio-based repository using TOSCA orchestration and Apache Mesos**

CHEP2018 23rd International Conference on Computing in High Energy and Nuclear Physics, 9-13 July, Sofia, Bulgaria

Marica Antonacci-Oral Presentation: **Piattaforma di sviluppo per AAI basata su Apache Mesos**  
CCR Workshop, 11-15 June 2018, Rimini

Marica Antonacci

1) **Lightning Talk: DEEP - Brief project overview**

2) **INDIGO PaaS Orchestrator Demo**

EOSC-HUB week, 16-20 April 2018, Malaga, Spain

Marica Antonacci-Oral Presentation: **On-demand large-scale data analytics platform with INDIGO-DataCloud PaaS**

Garr Conference 2017, THE DATA WAY TO SCIENCE, 15-17 November 2017, CA' FOSCARI UNIVERSITY OF VENICE

G. Vino, D. Elia e G. Donvito - Oral Presentation - **Developing a monitoring system for geographically distributed datacenters based on OpenStack**

30/05/2018 - Workshop GARR 2018

<https://www.eventi.garr.it/it/ws18/programma/workshop/speaker/266-gioacchino-vino>

***Posters:***

G. Vino , M. Antonacci, G. Donvito, D. Elia. - **Developing a monitoring system for Cloud-based distributed datacenters**

CHEP2018 23rd International Conference on Computing in High Energy and Nuclear Physics, 9-13 July, Sofia, Bulgaria

M. Antonacci , G. Donvito, D.C. Duma, S. Stalio. **Distributed and Integrated IaaS and PaaS Cloud for Research - The INFN Corporate Cloud solution**

CHEP2018 23rd International Conference on Computing in High Energy and Nuclear Physics, 9-13 July, Sofia, Bulgaria

A Monaco et al - **Il progetto Person: Uno strumento riabilitativo per malattie neurodegenerative fornito come servizio SaaS**

AIFM - 10° Congresso Nazionale Bari 2018