



Rapporti Tecnici INAF INAF Technical Reports

Number	219
Publication Year	2023
Acceptance in OA@INAF	2023-01-09T17:10:22Z
Title	Centro di calcolo oas plesso cnr struttura e servizi offerti
Authors	TACCHINI, ALESSANDRO
Affiliation of first author	OAS Bologna
Handle	http://hdl.handle.net/20.500.12386/32846 ; https://doi.org/10.20371/INAF/TechRep/219

CENTRO DI CALCOLO OAS
PLESSO CNR
STRUTTURA E SERVIZI OFFERTI

Autore: Alessandro Tacchini

INDICE

1. Introduzione
2. Networking
3. Cluster Virtual Machine
4. Cluster di Calcolo Parallelo
5. Storage
6. Server di Progetto
7. Impianti

1. Introduzione

Questo documento fornisce una sintetica descrizione dei servizi presenti nel Centro di Calcolo di OAS situato nel Plesso CNR.

Il documento è stato preparato nel 2019 e risulta molto importante per capire l'evoluzione del Centro di Calcolo negli ultimi anni.

Il Centro di Calcolo di OAS non è un semplice locale che contiene dei server ma una entità complessa ed in continua evoluzione.

I server che vi sono contenuti definiscono l'infrastruttura informatica per il Plesso CNR di OAS, sia come servizi generali che come servizi dedicati ai progetti ed ai ricercatori, e contribuiscono in modo significativo a tutto OAS oltre che all'infrastruttura informatica di INAF¹.

Il funzionamento dei server richiede anche una continua messa a punto degli impianti elettrici e di climatizzazione.

¹ Ad esempio vengono ospitati i server di MediaINAF e di EduINAF.

Aggiungere supporto ad un nuovo progetto significa modificare l'intera infrastruttura in modo che il servizio sia adeguato alla qualità richiesta.

2. Networking

Nel Centro di Calcolo è situato il cuore dell'infrastruttura di rete, qui sono situati tutti gli apparati che forniscono la connettività.

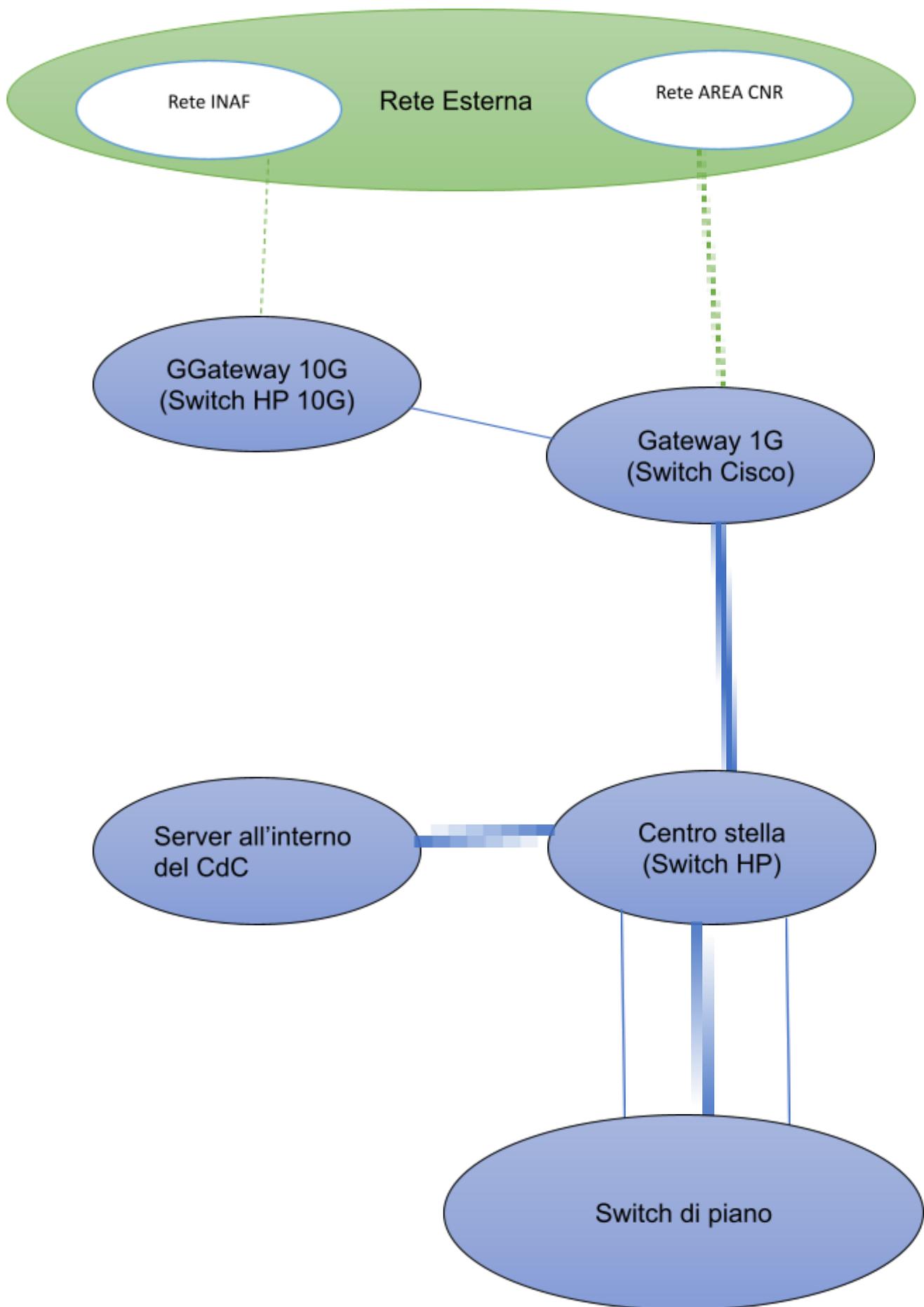
Sono presenti: un pannello (patch panel) da cui partono i cavi ethernet che collegano tutte le prese di rete in tutte le stanze dell'edificio; un pannello da cui partono i cavi in fibra ottica che realizzano il collegamento ad 1 Gb/s verso il router di Area CNR; un pannello da cui partono i cavi in fibra ottica che realizzano il collegamento a 10Gb/s verso il router INAF (sullo stesso pannello è presente il collegamento per il CTA Project Office).

I collegamenti ai vari piani sono ripartiti tramite tre switch dedicati, uno per piano, che a loro volta sono collegati in modo ridondato ai due switch che fungono da "Centro Stella".

Sempre al "Centro Stella" sono collegati una gran parte dei server del Centro di Calcolo e due switch Cisco (il vecchio centro stella) che fungono da portale verso l'esterno mediante il collegamento in fibra ottica.

Oltre ai vari switch che distribuiscono la connettività all'interno del centro di calcolo e che sono dedicati sia ai servizi di OAS Bologna che di INAF in generale, sono presenti anche apparati di rete di proprietà di CTA Project Office che permettono il collegamento verso l'esterno² a tutti i loro locali.

² Attraverso il POP Garr BO03-Gobetti



3. Cluster Virtual Machine

È un cluster formato da sette macchine, più una di controllo, non omogenee su cui è installato Oracle VM Server.

Un gruppo di macchine, per la precisione tre, ha 8 processori AMD e 16GB di RAM, due macchine hanno 8 processori Intel e 32 GB di RAM ed infine due macchine gemelle hanno 32 processori AMD e 128GB di RAM.

Le risorse Hardware quindi sono di dimensioni finite ma adeguate alle attuali esigenze dell'Osservatorio.

Il software di Oracle realizza un'infrastruttura robusta che offre alta affidabilità ed alta disponibilità.

Le macchine virtuali ospitate sul sistema sono di due tipi: macchine di un progetto/ricercatore e macchine per servizi di struttura.

Le macchine relative ad un progetto/ricercatore sono state richieste dai singoli ricercatori che necessitavano di avere a disposizione un server in poco tempo e senza caratteristiche particolari; sono macchine che realizzano servizi che vanno dal semplice web server dedicato al più impegnativo server per la pipeline di Agile.

Le macchine virtuali che danno servizi spaziano dal DNS/DHCP ad LDAP, dal Cloud (Owncloud) a RADIUS (per il servizio Eduroam) ed anche al server di posta elettronica (Zimbra).

4. Cluster di Calcolo Parallelo

Il cluster di calcolo parallelo è costituito da 5 nodi di calcolo per un totale di 296 Core e 768GB di RAM, 5 nodi di login, per un totale di 40 CPU, 3 server di Storage (da portare al più presto a 4 per avere la giusta ridondanza) che offrono più di 100TB di spazio disco.

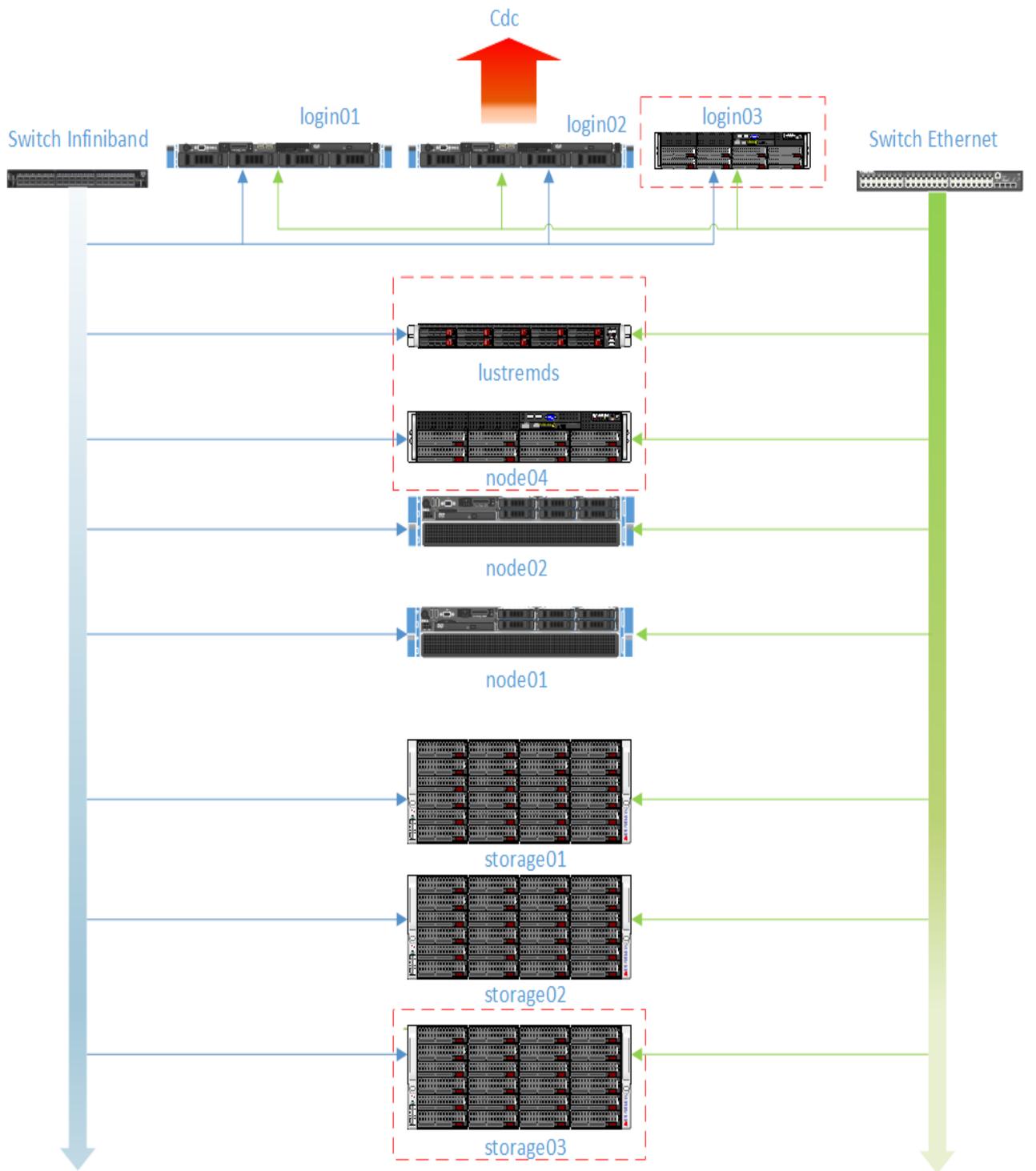
I nodi sono tutti collegati fra di loro con rete Infiniband 40Gb/s a bassa latenza. Il cluster è gestito utilizzando prodotti enterprise (offerti da IBM gratuitamente agli enti che fanno alta formazione), ovvero da uno scheduler (loadleveler) e da un filesystem distribuito (GPFS).

Il vantaggio di un cluster interno risiede nell'assenza di una policy per la prenotazione del cluster, a parte la policy di gestione dei jobs da parte dello scheduler.

In altre parole gli utenti abilitati possono sottomettere i propri job ad una coda del cluster e sarà poi lo scheduler che si occuperà di gestire l'allocazione delle risorse e l'esecuzione temporale.

Questa possibilità è data da due fattori: l'esiguo numero degli utenti e la non eccessiva richiesta di risorse di calcolo.

Nonostante la dimensione sia più piccola di quelle dei cluster del CINECA o di altre strutture INAF il nostro cluster risulta ideale per fare pratica e per ottimizzare il software, inoltre risulta comunque un buono strumento per il calcolo in generale.



5. Storage

Uno dei servizi più importanti ed anche meno conosciuti è lo Storage. Al giorno d'oggi lo spazio di archiviazione è diventato una risorsa indispensabile, la tendenza del futuro è un forte incremento della richiesta di spazio a tutti i livelli (singola persona, singolo progetto, intera infrastruttura).

L'organizzazione dello Storage all'interno del Centro di Calcolo ha seguito le richieste che si sono via via presentate nel corso del tempo e le macchine che lo realizzano sono dei sistemi SAN o NAS collegati ad una rete locale, in questo modo si garantisce accessibilità ma si mantiene un certo grado di protezione dalle intrusioni.

Tra i sistemi che formano l'insieme dello Storage si annoverano:

- Un Nas dedicato alla posta elettronica. In questo modo i dati risiedono su un supporto separato dal server di posta permettendone la portabilità in caso di incidente.
- Un NAS dedicato ai dati degli utenti. Ogni utente ha a disposizione una cartella Home in cui può tenere i propri dati (la dimensione di questo spazio è limitata).
- Un NAS dedicato ai dati di alcuni progetti di lungo corso il cui finanziamento è terminato da tempo (es. Integral).
- Un NAS dedicato al Backup. Il sistema di Backup, che impiega l'applicazione Tivoli di IBM, integra delle macchine NAS su disco con una unità a nastro, sempre di IBM.
- Una SAN dedicata al Cluster di Calcolo Parallelo. Con una dimensione di 110TB fornisce lo spazio necessario ai dati per questo tipo di calcolo.
- Una SAN che fornisce spazio disco alle macchine virtuali.

6. Server di Progetto

Oltre ai servizi comuni viene offerta ospitalità ed assistenza di base anche a quei progetti che necessitano della presenza di una o più macchine dedicate alla propria attività di ricerca.

- **ASTRI/CTA:** Per questo progetto il nostro istituto è responsabile del servizio GIT per tutto il SW, il sistema di collaborazione REDMINE, il test bed per lo sviluppo e il test del SW e il sistema di sviluppo e test del sistema di monitoring dell'ICT. Questi servizi sono realizzati utilizzando: 14 Server, con 188 core, 572 GByte di RAM e 120TB di spazio disco;
- **CTA RTA (Real Time Analysis)** per sviluppare la Pipeline di RTA abbiamo 2 Server molto potenti un OpenPower con 192 core, 128 GByte di RAM e 4 TB di spazio disco e un server Intel compatibile con 56 core, 128GB di RAM e 16TB di spazio disco;
- **Planck Main Backup Repository for Mission Data:** cluster di 6 server, con uno dedicato al solo accesso esterno, uno di backup, e gli altri fondamentalmente configurati come NAS, asserviti sia a detto portale che al cluster di Istituto, con 20 core totali e storage complessivo di circa 30 TB (in configurazioni RAID miste 1/10/5/6), in parte organizzato utilizzando un file system distribuito;
- **AGILE:** Per la missione AGILE abbiamo la responsabilità della pipeline di analisi dei dati del satellite in tempo reale per la generazione delle allerte scientifiche e la gestione delle allerte dall'esperimento LIGO. La Pipeline e l'analisi dati in tempo reale sono gestite utilizzando 12 Server, con 116 core, 280GB di RAM e 76TB di spazio disco;

- REM/GAIA/GRAPPA: nell'ambito di questi di progetti nazionali e internazionali curiamo sia la riduzione dati che il server web e archivio dati serviti da DBMS. In particolare ospitiamo tutte le immagini del telescopio REM (circa 3 milioni di immagini), il mirror di più di 50 cataloghi astronomici usati nell'ambito del progetto GAIA (24 TB), un archivio di milioni di modelli di spettri di galassie usate per il progetto GRAPPA. Altri progetti che prevedono l'uso di web-tool con accesso ad archivi locali sono in fase di sviluppo. Tutte queste risorse sono accessibili remotamente sia attraverso web-form che direttamente da programmi degli utenti. Per tutto questo sono usati 6 server con 40 Core, 196Gbyte di RAM e 82TB di spazio disco.

Oltre al supporto a progetti specifici viene fornito supporto anche ad alcuni servizi INAF, MediaINAF e EduINAF. Entrambe le macchine sono gestite dai rispettivi responsabili.

7. Impianti

IMPIANTO DI CONDIZIONAMENTO

L'impianto di condizionamento, con una potenza frigorifera complessiva di circa 100K BTU, è costituito da:

- Due condizionatori di precisione per “centro calcolo” Emerson con sistema di freecooling, ovvero con un sistema che rileva la temperatura esterna e nel caso in cui essa sia più bassa di quella del locale impiega l'aria esterna (opportunamente filtrata) per il raffreddamento del locale. I due condizionatori sono in grado di scambiarsi comunicazioni sul loro stato per minimizzare i consumi.

- Un condizionatore ad acqua fredda che utilizza l'impianto dell'istituto per la generazione dell'acqua per il condizionamento.

Abbiamo quindi un sistema completamente ridondato per il condizionamento.

IMPIANTO ELETTRICO

L'impianto elettrico del centro di calcolo è interamente collegato al gruppo di continuità che offre un sistema trifase con neutro da 60kVA. È quindi possibile ottenere 3 fasi distinte a 220Volt. Ogni server ha una doppia alimentazione che si collega a due fasi distinte, in tal modo è possibile effettuare interventi su di una fase senza che il centro di calcolo ne risenta. Allo stato attuale manca il commutatore tra la rete Enel ed il gruppo elettrogeno che dovrebbe entrare in funzione automaticamente entro pochi minuti dall'interruzione dell'erogazione della corrente. Tale assenza impedisce un corretto funzionamento del sistema che esaurita l'energia immagazzinata nel pacco batterie si spegne brutalmente. Tenuto conto dei consumi totali delle macchine installate la durata del pacco batterie è mediamente di 20 minuti. Si richiede quindi di ripristinare quanto prima tale commutatore per avere le giuste misure di sicurezza nell'erogazione della potenza elettrica.

SISTEMA ANTINCENDIO

Il sistema antincendio del Centro di Calcolo è affidato a due estintori a CO2 posti nel locale.