

Sistema di backup e ripristino dati al CNAF

Enrico Fattibene, Alessandro Cavalli

INFN-CNAF

Workshop della CCR – Paestum 26 maggio 2022

Introduzione

- Il servizio di backup e ripristino dati mitiga il rischio di perdita di dati ritenuti critici
- I dati devono poter essere ripristinati nel minor tempo possibile
- Nel momento del bisogno si deve ripristinare tutto quello che serve
- I dati da salvare sono relativi a servizi amministrati da reparti diversi del CNAF o appartengono a progetti diversi (amministratori interni ed esterni)
- Bisogna sempre limitare i costi
- Il ripristino può non essere effettuato per mesi, ma nel momento del bisogno non può fallire



Dati sotto backup

- Servizi nazionali
 - Mail server (CNAF e Presidenza), mailing list, documentale, web server, repository Git e SVN, etc.
- Sistema informativo
- Servizi generali CNAF
 - Log server, dhcp, monitoring, immagini virtuali, domain controller
- Tier1
 - Home utenti, database Oracle, wiki, elog
- Sistemi distribuiti
 - Portali, siti, agenda, vari database, cvmfs, VOMS, configurazioni di servizi di produzione
- EPIC Cloud
 - Sistema certificato e gestito secondo le norme ISO per la sicurezza delle informazioni (27001, 27017, 27018)
 - Dati della infrastruttura gestita da personale CNAF
 - Dati delle applicazioni degli utenti

Dati <i>ACTIVE</i>	63 TB	172 M file
Dati <i>INACTIVE</i>	109 TB	85 M file

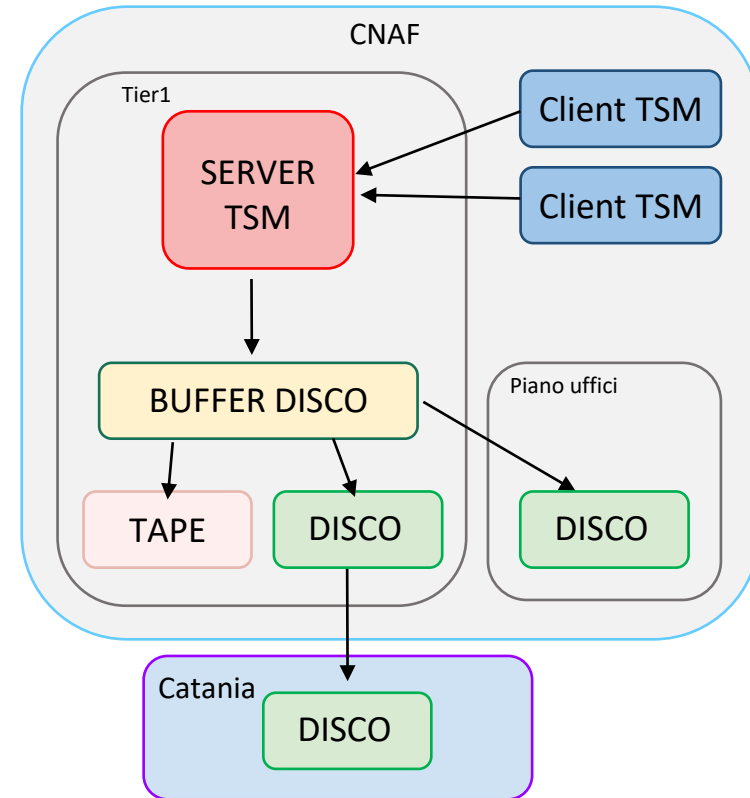
Infrastruttura

- 1 server IBM Spectrum Protect (TSM)
- 1 libreria tape (Oracle SL8500)
 - Principalmente usata per i dati degli esperimenti scientifici in modalità HSM
 - 6 tape drive T10000C
 - ~100 tape T2 di capacità 5,4TB l'una
- Un'area disco da 170TB
- Per i dati di EPIC e per il backup del DB di TSM
 - 1 server disco in stanza 45 (piano 1 CNAF)
 - Disco a INFN-Catania



Backup e archiviazione

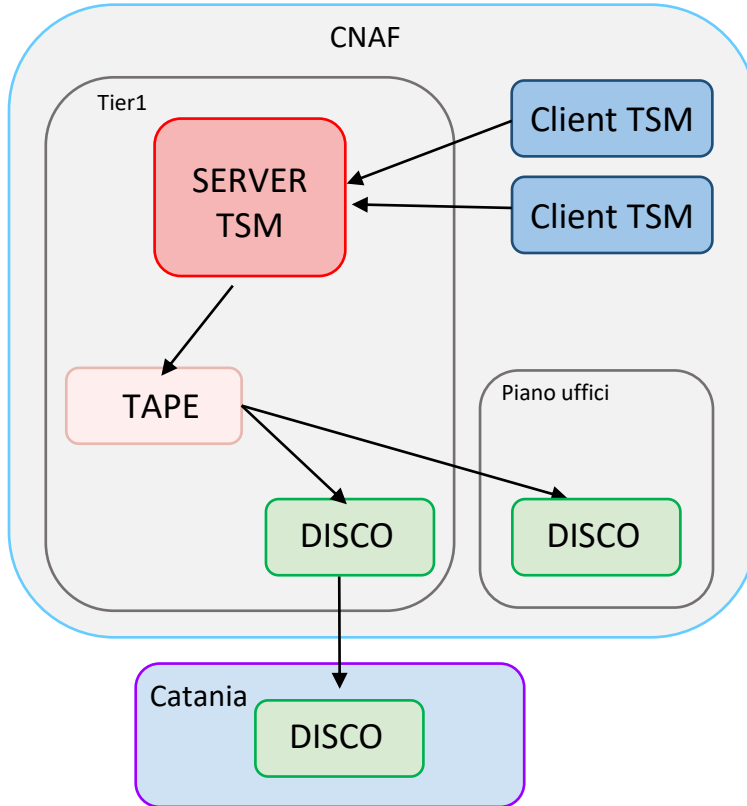
- Backup incrementale
 - Processo automatico (tipicamente giornaliero) su ciascuno dei client registrati
 - Per servizi interni CNAF:
 - Ultima versione dei file (*Active*) salvata sia su disco che su tape
 - Vecchie versioni di file modificati o cancellati (*Inactive*) salvati su tape
 - 2 diverse “Retention” (tempo per cui gli *Inactive* sono mantenuti nel sistema), a seconda del tipo di dato: 1 mese e 1 anno
 - Per dati di EPIC Cloud:
 - Tutti i dati salvati in 2 copie tape e 1 disco al CNAF e 1 copia disco a INFN-CT
 - Retention in accordo con gli utenti
- Archiviazione
 - Per dati che non verranno modificati nel tempo e che vanno conservati su nastro per un certo periodo



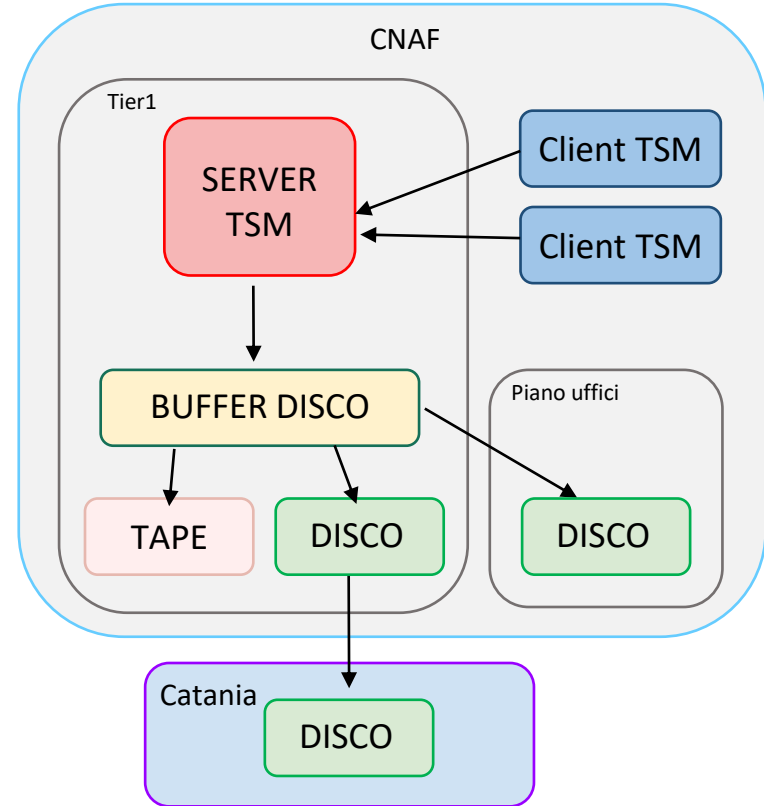
Client e licenze TSM

- Lavoro di razionalizzazione dei client
 - Per limitare i costi sono stati creati client collettori di dati di vari servizi
- Software IBM ILMT
 - Per poter usufruire del regime di sub-licensing
- 28 nodi client TSM su host fisici o virtuali
 - 24 client su Linux (80% CentOS 7)
 - 2 client standard su Windows Server 2012/2016
 - 2 client Oracle database (Tivoli Data Protection for Oracle Database - TDPO) su RedHat Enterprise Linux 6 e Oracle Linux 7
- Gestione mista
 - Personale CNAF: 19 client
 - Utenti esterni (dati dei progetti EPIC): 9 client

Flusso dei dati



Prima: scrittura su tape e copie successive



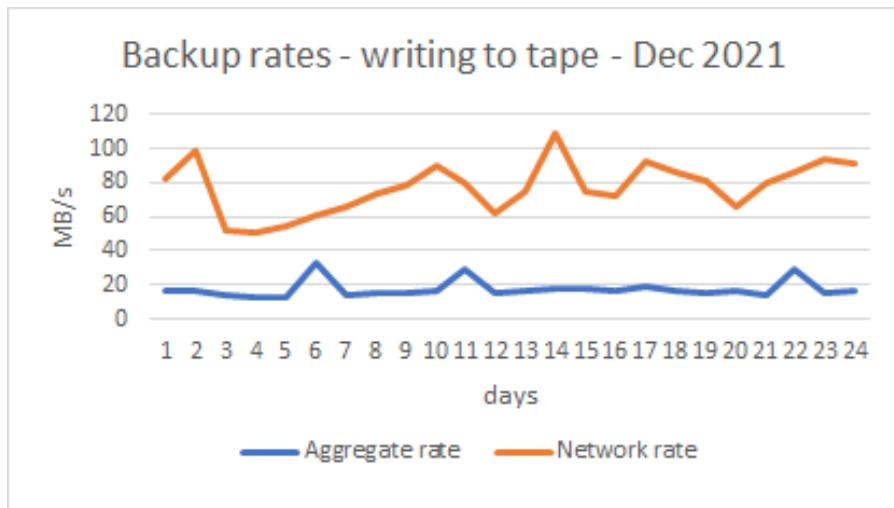
Ora: scrittura su buffer disco e migrazioni successive

Buffer disco

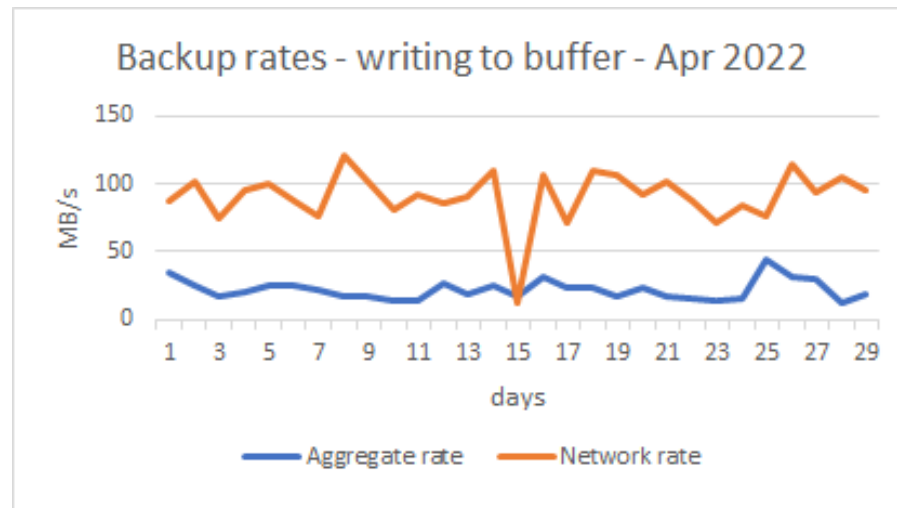
- Dati scritti su un disco che funziona da buffer davanti al nastro
 - Migrati su nastro e su disco permanente qualche ora dopo
 - In precedenza i dati provenienti dai client venivano scritti direttamente su tape
- Complica la configurazione di TSM
 - Divisione logica del buffer in diversi storage pool, uno per ogni storage pool di nastri
- Permette di minimizzare le risorse e quindi i costi
 - Drive utilizzati in modo ottimale, solo per il tempo necessario alle scritture dal buffer

Metodi di scrittura

Esempi di backup da uno stesso client con i due metodi



Rate effettivo medio (network): 77.2 MB/s
Rate aggregato medio: 17.5 MB/s
Efficienza aggregato/effettivo: 23%

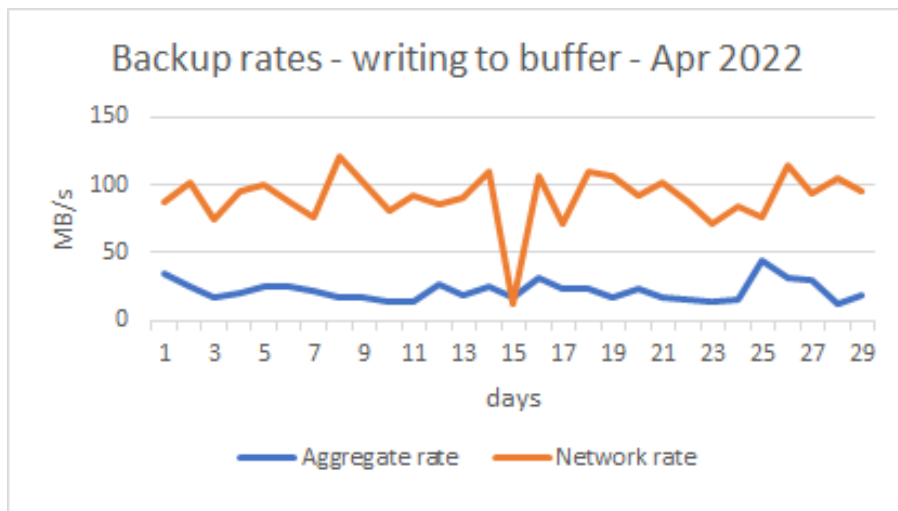


Rate effettivo medio (network): 90.9 MB/s
Rate aggregato medio: 22 MB/s
Efficienza aggregato/effettivo: 28%

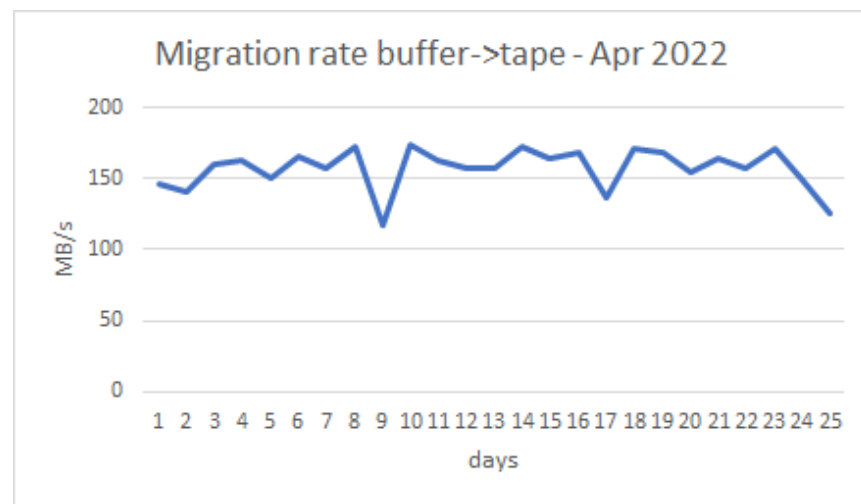
Con scrittura diretta su nastro, inefficienza nell'uso dei drive del 75%

Utilizzo drive ottimizzato

Scrittura su buffer e migrazione successiva
Rate nominale tape drive: 250 MB/S



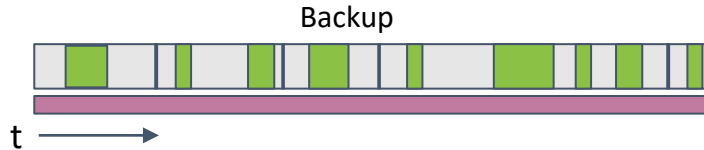
Rate aggregato medio: 22 MB/s





Rate aggregato medio: 157 MB/s


Scrittura via buffer vs diretta su tape

Scrittura diretta su tape

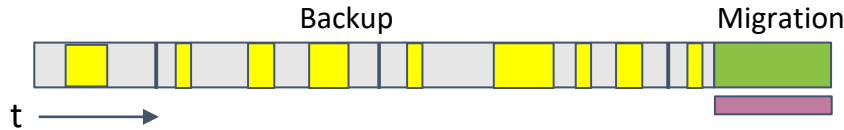


 Scrittura su buffer

 Scrittura su nastro

 Utilizzo tape drive

Scrittura via buffer



- Con scrittura via buffer
 - Drive utilizzati solo per il tempo della scrittura su nastro
 - Migliora il rate di scrittura su nastro (differenza destinata ad aumentare con tecnologie future)
 - È possibile scrivere in parallelo da più client
- La migrazione da buffer a nastro richiede al massimo 2 drive
 - Utilizzati in parallelo per le 2 copie su nastro
 - Obiettivo di ridurre il numero di drive per il backup (6 attuali)

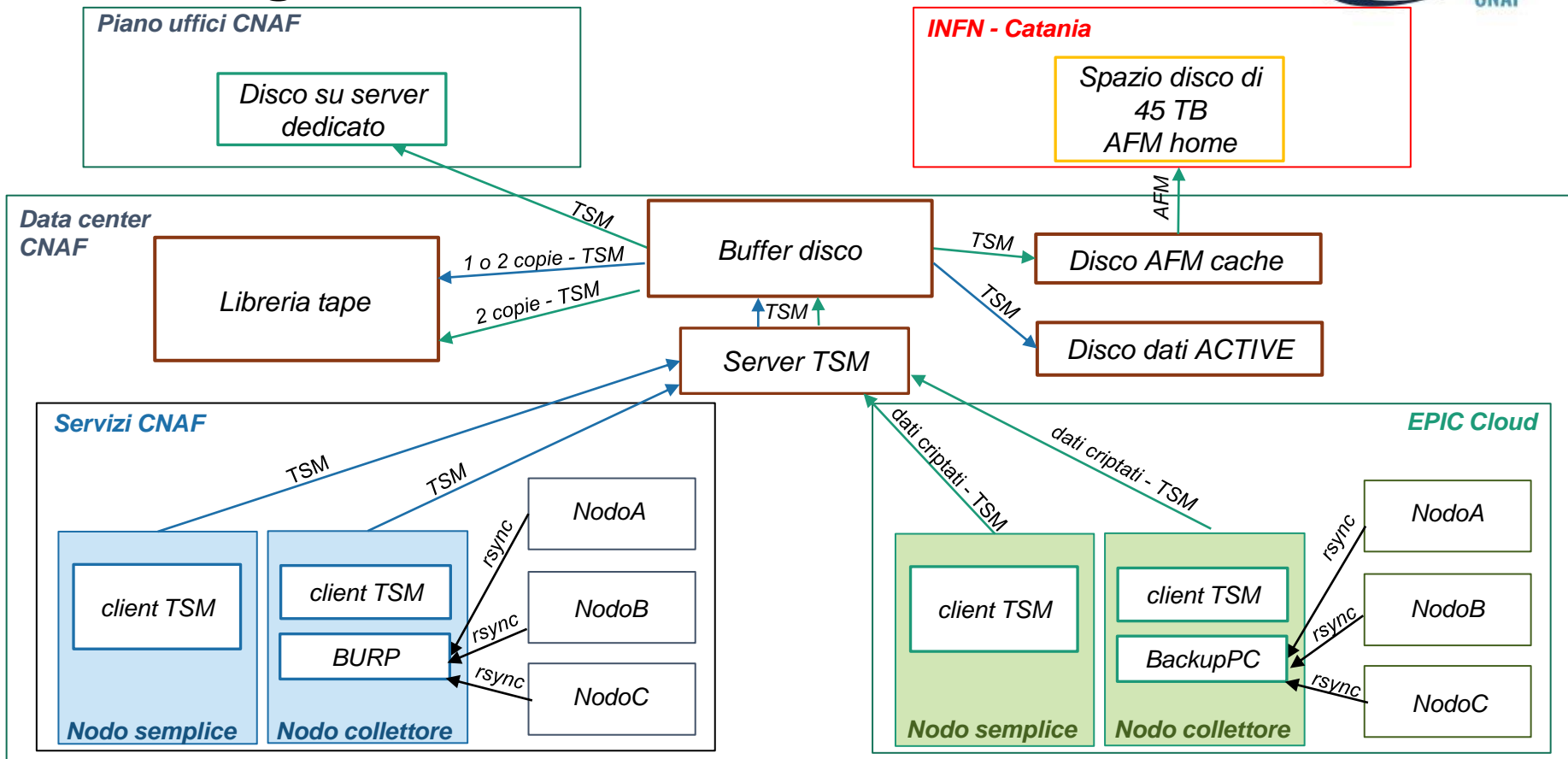
Backup in EPIC Cloud



- Usa il sistema di backup del CNAF tenendo conto delle norme per la sicurezza delle informazioni
 - Dati dell'infrastruttura e delle applicazioni degli utenti
 - Configurazione client fatta dagli utenti secondo nostra procedura
 - Documentazione condivisa con gli utenti
 - Procedura di backup
 - Configurazione lato client
 - Configurazione lato server di retention policy, directory da salvare, tempistiche, etc.
 - Precisa distinzione di responsabilità del CNAF e degli utenti
 - Comunicazioni tracciate sempre via ticket
 - Test di restore e DR con cadenza annual
 - Dati criptati sul client con AES (Advanced Encryption Standard) 256 bit
 - Distribuzione delle copie per minimizzare rischio di perdita dati
 - 15 TB con previsione di importante crescita



Schema generale



Backup su INFN-Catania



- Precedente metodo: rsync
 - Volumi TSM copiati da disco CNAF a INFN-Catania
 - Errori nel test di restore da dati a Catania
 - TSM riconosceva solo una parte dei file presenti nei volumi ricopiati da Catania al CNAF
 - In caso di DR, necessità di ricopiare tutti i volumi TSM sul disco a disposizione del nuovo server
 - Probabile inefficienza nel restore
- Nuovo metodo: AFM (Active File Management)
 - Integrato in GPFS
 - Cache al CNAF vista da TSM server, home su disco di Catania
 - Sincronizzazione cache-home in background
 - Scritture dirette via TSM
 - In caso di DR, lettura (trasparente) solo dei blocchi che contengono i file di cui fare restore
 - Risparmio in termini di tempo di attesa per avere i dati a disposizione

Test di restore EPIC Cloud



- I test di restore devono essere effettuati dagli amministratori dei client
- Una volta all'anno
- Per ogni client TSM
 - test di ripristino completo dell'ultima versione disponibile dei dati dalle copie di backup disponibili al CNAF (2 copie nastro e 1 copia disco)
 - test di ripristino completo della versione dei dati ad un momento antecedente alla data del test da una delle copie di backup

Test di disaster recovery EPIC Cloud

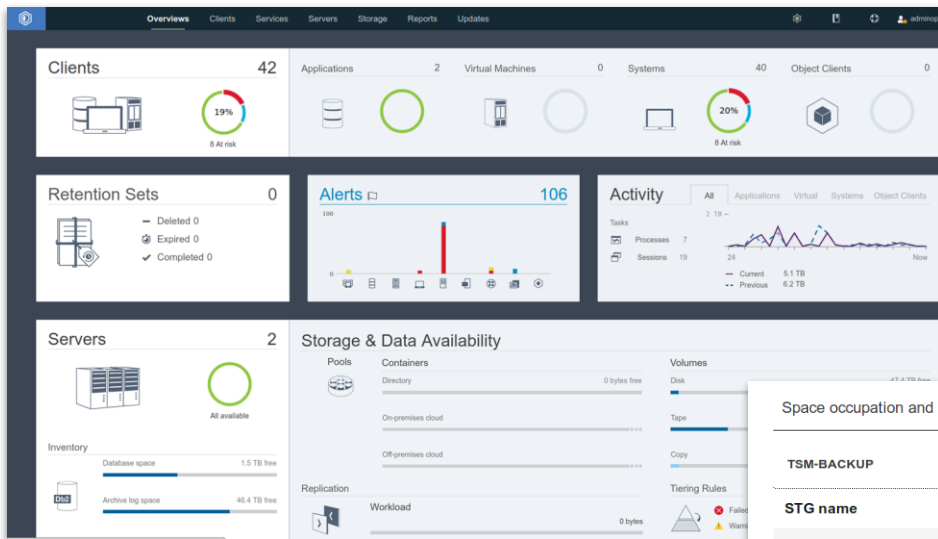


- Una volta all'anno
- Simuliamo la perdita del server TSM e di tutte le copie di backup al CNAF
 - Installazione nuovo nodo con server TSM
 - Configurazione di cache AFM su nodo che ospita TSM server
 - Recupero DB TSM da Catania
 - Restore del DB a partire dalla copia di Catania
 - Restore dei dati di un client
- Ultimo test effettuato con successo a marzo 2022

Gestione operativa

- Report giornaliero via mail per ogni reparto/progetto con info sui propri client e in particolare:
 - Backup programmati nelle ultime 24 ore
 - orario di schedula, di effettivo inizio e di fine backup, stato finale del backup - Completed, Failed, Future, Missed, Started, Restarted - ed output che informa sulla presenza o meno di errori/warning
 - Spazio occupato e numero di file salvati per client su ognuno dei pool (disco e nastro)
 - Spazio occupato e numero di file salvati per filespace complessivamente sui pool
- Ogni reparto è responsabile del controllo dell'esito dei processi di backup sulle proprie risorse
 - I problemi sui client EPIC sono sempre tracciati via ticket
- Ripristino
 - Sono previste prove di ripristino con cadenze varie per ciascun client

Operations Center - Mail Reports



Report personalizzato
20 maggio 2022 alle 9:15
Questo report comprende 3 query SQL

Scheduled backups in the last 24 hours

Node	Schedule time	Start time	End time	Status
DISKSERV-SAN-94	2022-05-20 00:01	2022-05-20 02:53	2022-05-20 04:46	Completed

Space occupation and number of files per client

Node name	Space in GB	Number of files	Location
DISKSERV-SAN-94	1590.45	10233961	Disk (active)
DISKSERV-SAN-94	2022.21	10254643	Tape 1st copy (active + inactive)
DISKSERV-SAN-94	2035.25	10254779	Tape 2nd copy (active + inactive)

Space occupation and number of files per storage pool

STG name	Space in GB	Number of files	Location
ACTIVE_STG	51132.76	178643855	Disk (active)
GENERAL-T1C	102071.37	122023037	Tape 1st copy (active + inactive)
HARMONY_PRIMARY	4776.52	4990556	Tape 1st copy (active + inactive)
ORACLE-T1C	3424.48	572	Tape 1st copy (active + inactive)
PRIMARY_C	47233.08	92487068	Tape 1st copy (active + inactive)
COPY_C	50374.37	92487776	Tape 2nd copy (active + inactive + archive)
HARMONY_COPY	4755.76	4949912	Tape 2nd copy (active + inactive)
ARCHIVE_C	31.71	136	Tape archive 1st copy

IBM Operations Center permette di monitorare, effettuare azioni ed inviare mail report ad amministratori o utenti del servizio. Al CNAF è utilizzato principalmente per i mail report e per accedere alle statistiche rapidamente.

Quanti dati abbiamo

- 63 TB / 172 milioni di oggetti *Active*
- 109 TB / 85 milioni di oggetti *Inactive*

- Spazio occupato su disco CNAF: 78 TB (escluso Buffer)
- Spazio occupato su disco Catania: 15 TB
- Spazio occupato su tape: 225 TB

Sviluppi futuri

- Nuovi progetti EPIC in arrivo
 - Da verificare se i casi d'uso siano compatibili con il servizio di backup attuale
- Possibili modifiche all'infrastruttura
 - 1 server per servizi CNAF interni e 1 server per EPIC Cloud
 - Pro: divisione nella gestione, database più snelli
 - Contro: risorse hardware separate -> bisogno di più risorse
 - Seconda copia tape su nuova libreria IBM

Conclusioni

- Il sistema di backup e ripristino dati del CNAF
 - E' a disposizione dei vari reparti del CNAF per tenere copie di dati critici
 - Preserva i dati di EPIC seguendo le norme ISO per la sicurezza delle informazioni
 - Copie su media diversi e geograficamente distribuite
 - E' stato ottimizzato per minimizzare uso di licenze software e risorse hardware
 - Dovrà far fronte alla futura crescita dei dati di EPIC