



# Disaster Recovery INFN

Galli Claudio 26-28/02/2014

Workshop CCR Laboratori Nazionali G.Sasso

# Gruppo

*Coordinatore :* Stefano Zani

*Componenti :* Sandro Angius  
Massimo Donatelli  
Claudio Galli  
Guido Guizzunti  
Dael Maselli  
Massimo Pistoni  
Claudio Soprano  
Riccardo Veraldi

*+ Collaborazione :*

Nunzio Amanzi  
Alessandro De Salvo  
(indicazioni su distributed File System)

...

# Aree di lavoro

DNS {distribuito + geo-replica}

MAILING {distribuito + mail relay }

SISTEMA INFORMATIVO :

Contabilità (CNAF) (→R12)

Portale Utente (CNAF)

Gestione Presenze (CNAF)

Documentale Alfresco (CNAF) [new]

Business Intelligence BI (CNAF)

Protocollo (CNAF) [new]

AAI + GODIVA (LNF)

Stipendiale + Sxgest2 (LNF)

Stipendiale + Cezanne (LNF) [new]

Protocollo (LNF)

Documentale (LNF)

Portale Unico (LNF) [new]

...



# Le Fasi di Intervento



## Step 0: Replica dei «Dati Grezzi», ovvero dei Backup di database e applicativi.

*(Permette in caso di disastro di recuperare i dati grezzi necessari per un successivo ripristino manuale dei servizi , nella sede originaria o di backup a seconda della gravità del disastro .)*

**[99.9%]**

# Step 0 : Replica Geografica dei Backup



- Apparati di Storage eterogenei tra LNF e CANF.
- Scartate le soluzioni di mirroring tra gli storage delle due sedi.
- E' stato implementato un meccanismo di sincronizzazione bidirezionale (CNAF-LNF) dei backup giornalieri che si basa sul **rsync+SSL**. Monitoring via **mail**.
- **Conf. vista dal CNAF (la situazione ai morsetti è idealmente speculare se vista da LNF) :**
  - 1 Macchina Fisica RH Invia (CNAF → LNF): [ → su Tape]**
    - Backup Freddo della contabilità (Db + App circa 500Gb dati, *procedura rvisionata dopo la messa in produzione del nuovo **Cluster R12** della Contabilità*).
    - Hot Backup del database della Contabilità R12 (2/day).
    - Backup delle conf. dell VM del SI. presenti al CNAF (Portale Utente, Sistema Presenze, Business Intelligence, Dashboard CNAF, ecc...)
  - 1 Macchina Virtuale RH Riceve (LNF → CNAF ): [ → su Tape]**
    - Backup notturno di Sxgest2 (Vecchio HR+Stipendiale-Documentale-Protocollo)
    - Hot Backup Godiva PreProd, Prod (2 full/settimana)
    - Hot Backup CEZANNE (area HR per il nuovo sistema Stipendiale)
    - **@work: l'invio dei backup delle VM OVirt legate all' area Cezanne-Sipert per il nuovo Stipendiale presso LNF. (Ultimata a brevissimo)**

# Le Fasi di Intervento



## Step 0: Replica dei «Dati Grezzi», ovvero dei Backup di database e applicativi.

*(Permette in caso di disastro di recuperare i dati grezzi necessari per un successivo ripristino manuale dei servizi , nella sede originaria o di backup a seconda della gravità del disastro .)*

**[99.9%]**

## Step 1: Replica geografica con allineamento in tempo reale delle banche dati critiche per l'ente [Oracle Data Guard 11g]. *[primi test 2012-13. Ripartirà **Marzo 2014** .]*



*(Garantisce il corretto livello di **data replication** e getta le basi per le evoluzioni successive.)*

# Step 1 : Replica Geografica dei Database 1/2



- Vista la quantità di installazioni di Oracle Database Ent. Ed. si è scelto di studiare e testare le repliche geografiche con **Oracle Data Guard 11gR2** (Soluzione consigliata in più occasioni e contesti )
- Sono stati fatti 3 test finora.
  - 1° test 2012 - Replica locale (tra due macchine interne al CNAF) di un database di test Oracle Enterprise **10gR2** con Data Guard 10g.  
(di difficile gestione/monitoraggio su Data Guard 10gR2)
  - 2° e 3° test - Replica Geografica LNF → CNAF del database di Pre-Prod di Godiva (11gR2). Utilizzato con successo **Oracle Data Guard 11g** e l'allineamento ha funzionato correttamente per una settimana.
- Si è deciso di uniformare il livello tecnologico alla versione **11gR2** dei Database per poter sfruttare alcune facility di **Oracle Data Guard 11gR2** .

# Step 1 : Replica Geografica dei Database 2/2



- **Database Godiva Prod e Pre-Prod** : già compatibili con le scelte effettuate.
- **Contabilità** : il Passaggio alla R12 mette a disposizione un Database Oracle enterprise Ed. 11gR2 compatibile con le scelte effettuate.
- **DB HR + Stipendiale** :
  - per il vecchio impianto (**SXGEST2**) verranno mantenuti in sync i soli backup, fintanto che ci saranno parti del vecchio impianto (hr+Stip+**Protocollo**) in uso.
  - Il nuovo impianto (**Cezanne**) sfrutta un Database Oracle Enterprise 11gR2 su ODA (LNF) per lo strato di persistenza dei dati, quindi è compatibile con le scelte effettuate.
- **Data Warehouse BI** : già compatibile con le scelte effettuate.  
Necessario comunque effettuare dei test e delle verifiche sulle procedure ETL di importazione dei dati
- **Sistema Presenze**: attualmente su DB Oracle Enterprise **10gR2**. E' necessario un upgrade del database alla versione 11gR2 prima di poter replicare efficacemente il database tramite Data Guard 11g.

# Le Fasi di Intervento

## Step 0: Replica dei «Dati Grezzi», ovvero dei Backup di database e applicativi.

*(Permette in caso di disastro di recuperare i dati grezzi necessari per un successivo ripristino manuale dei servizi , nella sede originaria o di backup a seconda della gravità del disastro .)*

**[99.9%]**

## Step 1: Replica geografica con allineamento in tempo reale delle banche dati critiche per l'ente [Oracle Data Guard 11g]. [primi test 2012-13. Ripartirà **Marzo 2014** .]

*(Garantisce il corretto livello di **data replication** e getta le basi per le evoluzioni successive.)*

## Step 2: Replica geografica in sync degli applicativi running.

*(La sincronizzazione degli strati applicativi sovrastanti le basi di dati consente di ripristinare i servizi in caso di disastro.)*

**[Technology tracking]**



# Step 2 : Replica Geografica degli Applicativi



## Fattori Abilitanti

### **DNS HA: (READY)**

La struttura DNS HA implementata dal CNAF in collaborazione con ROMA1 e altre sedi che ospiteranno le ulteriori istanze di DNS in ridondanza è un elemento fondamentale per la costruzione di servizi in alta affidabilità a livello geografico e consente di referenziare dinamicamente le istanze multiple e distribuite sul territorio (e su reti differenti).

### **Utilizzo di applicativi per la installazione automatica degli application server:**

L'utilizzo nelle sedi che ospitano servizi in DR di pacchetti per la installazione e gestione automatica degli application server (Per esempio PUPPET) renderebbe più veloce la reinstallazione di un servizio con contestualizzazione a livello di rete dei server.

### **Estensione di una rete locale a livello geografico fra le sedi: (DA SPERIMENTATE)**

Se riuscissimo ad estendere una rete locale a livello geografico, potremmo migrare macchine virtuali da una sede all'altra senza necessità di modificare la numerazione IP e senza dovere intervenire sul DNS-HA → **Recovery piu rapido** e possibile realizzazione di "Cluster distribuiti in alta affidabilità"

# Step 2 : Replica Geografica degli Applicativi

## technology tracking

### GlusterFS (Replica remota a livello di FileSystem):

Localmente al CNAF sono stati fatti alcuni test preliminari con la versione Comunity della Suite di Distributed Storage, ormai ufficialmente di Casa Red Hat (*Red Hat Storage Server*) . I primi test con Macchine Virtuali KVM hanno messo in evidenza alcune criticità nella gestione della consistenza delle VM in caso di disservizi (Interazioni con il Gruppo Coud hanno gettato dubbi sulla efficacia di questo FS a livello geografico)

Esistono molte alternative per File system distribuiti, ma è necessario identificare quali siano le più appropriate in base alle esigenze degli applicativi sovrastanti (GPFS?).



# Step 2 : Replica Geografica degli Applicativi

technology tracking

## Replica remota dei volumi (a livello di DEVICE)

Per tutto ciò che non è DB per i quali usiamo DataGuard , si potrebbero utilizzare specifiche features di replica remota fornite direttamente dai sistemi di storage.

Esempi: **Equallogic Auto Replication**

**EMC Mirror View**

**Netapp SnapMirror**

### PRO:

Una volta configurato a livello di volume, la replica è automatizzata

### CONTRO:

Questi sistemi pongono vincoli sul tipo di storage utilizzato nelle sedi (che dovrà essere omogeneo). Questi oggetti possono avere un costo rilevante

DUBBI: Come gestisco la consistenza delle immagini delle macchine virtuali che utilizzo?



# Step 2 : Replica Geografica degli Applicativi

technology tracking



## Hypervisor-based Replication

**Esempi:** VMware vCenter Site Recovery Manager (Commerciale)

**ZERTO** (<http://www.zerto.com/>) (Commerciale)

Lavora a livello di Hypervisor e replica anche le singole VM.

### PRO:

Sono soluzioni specifiche per il DR

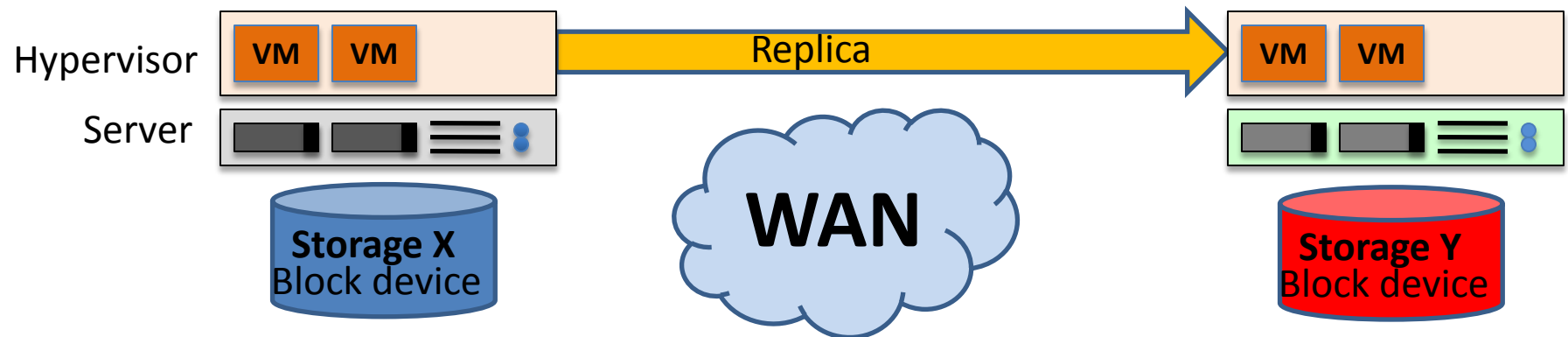
Sono indipendenti dall'HW utilizzato

Si occupano di mantenere consistenti i file delle immagini virtuali nelle repliche (da testare).

### CONTRO:

Supporta solo VMware (No KVM)

Non è FREE!!



# Le Fasi di Intervento



## Step 0: Replica dei «Dati Grezzi», ovvero dei Backup di database e applicativi.

*(Permette in caso di disastro di recuperare i dati grezzi necessari per un successivo ripristino manuale dei servizi, nella sede originaria o di backup a seconda della gravità del disastro.)*

**[99.9%]**

## Step 1: Replica geografica con allineamento in tempo reale delle banche dati critiche per l'ente [Oracle Data Guard 11g]. [primi test 2012-13. Ripartirà **Marzo 2014**.]

*(Garantisce il corretto livello di **data replication** e getta le basi per le evoluzioni successive.)*

## Step 2: Replica geografica in synch degli applicativi running.

*(La sincronizzazione degli strati applicativi sovrastanti le basi di dati consente di **ripristinare i servizi** in caso di disastro.)*

**[Technology tracking]**

## Step 3: Test Plan periodici e definizione delle Procedure di Ripristino.

*Le prove periodiche di ripristino dei servizi servono a verificare che le procedure di recovery funzionino correttamente.*

# Prossimi appuntamenti



**Marzo 2014:**

**10 - 14 : Ripresa lavori per la replica con Data Guard.**

**31 : Riunione di allineamento del gruppo.**

**Fine**