

Stato dell'infrastruttura dei SSNN con HA geografico

Stefano Longo
03 Giugno 2019

Recap [1 / 3]

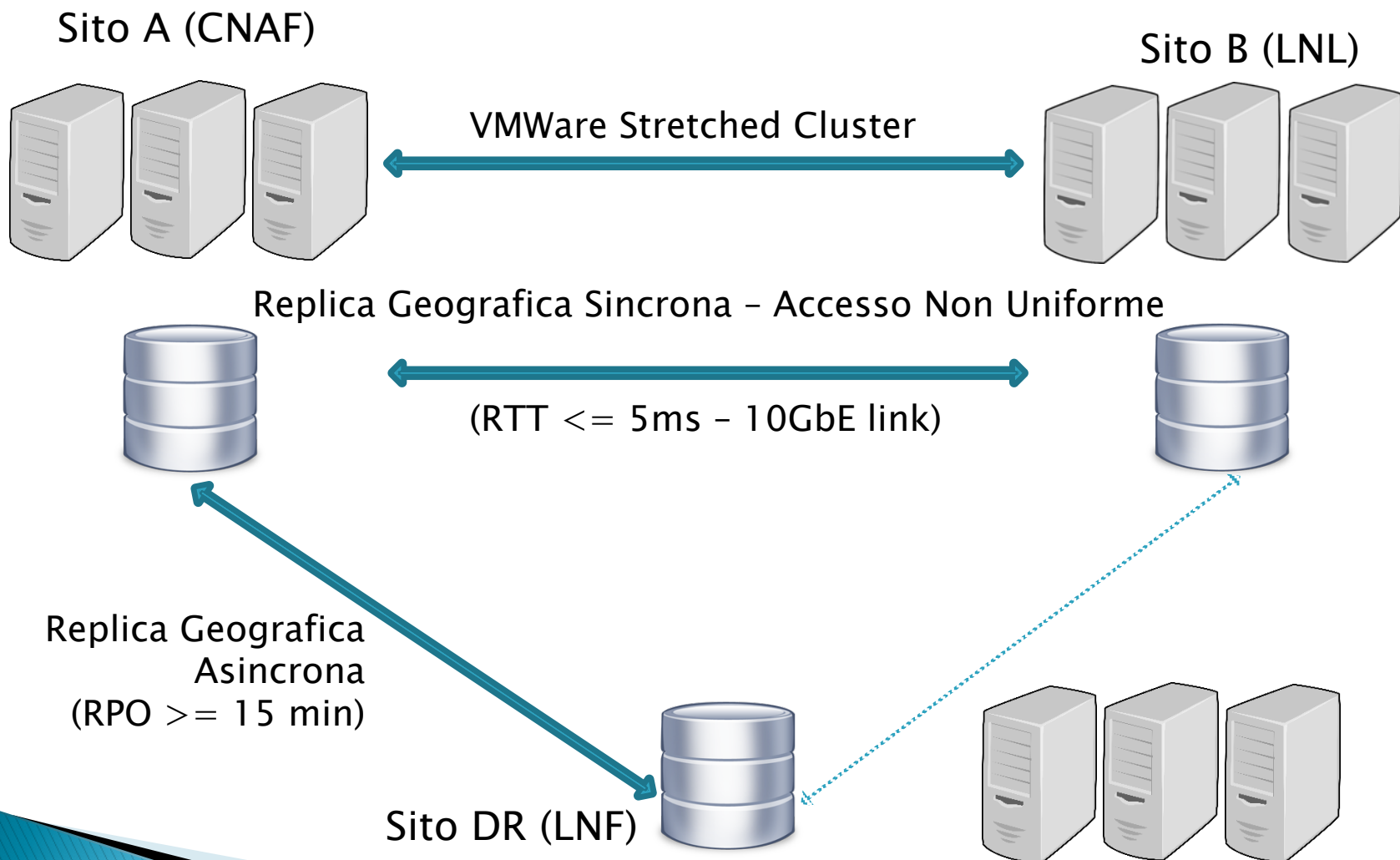
Richiesta per aumentare affidabilità e disponibilità di dati e applicazioni:

- ▶ Anche in caso di indisponibilità di uno dei centri che erogano Servizi
- ▶ Con un RPO ~ 0
- ▶ Con un RTO $\rightarrow 0$
- ▶ Senza introdurre particolari requirements a livello applicativo

Soluzione proposta a livello infrastrutturale

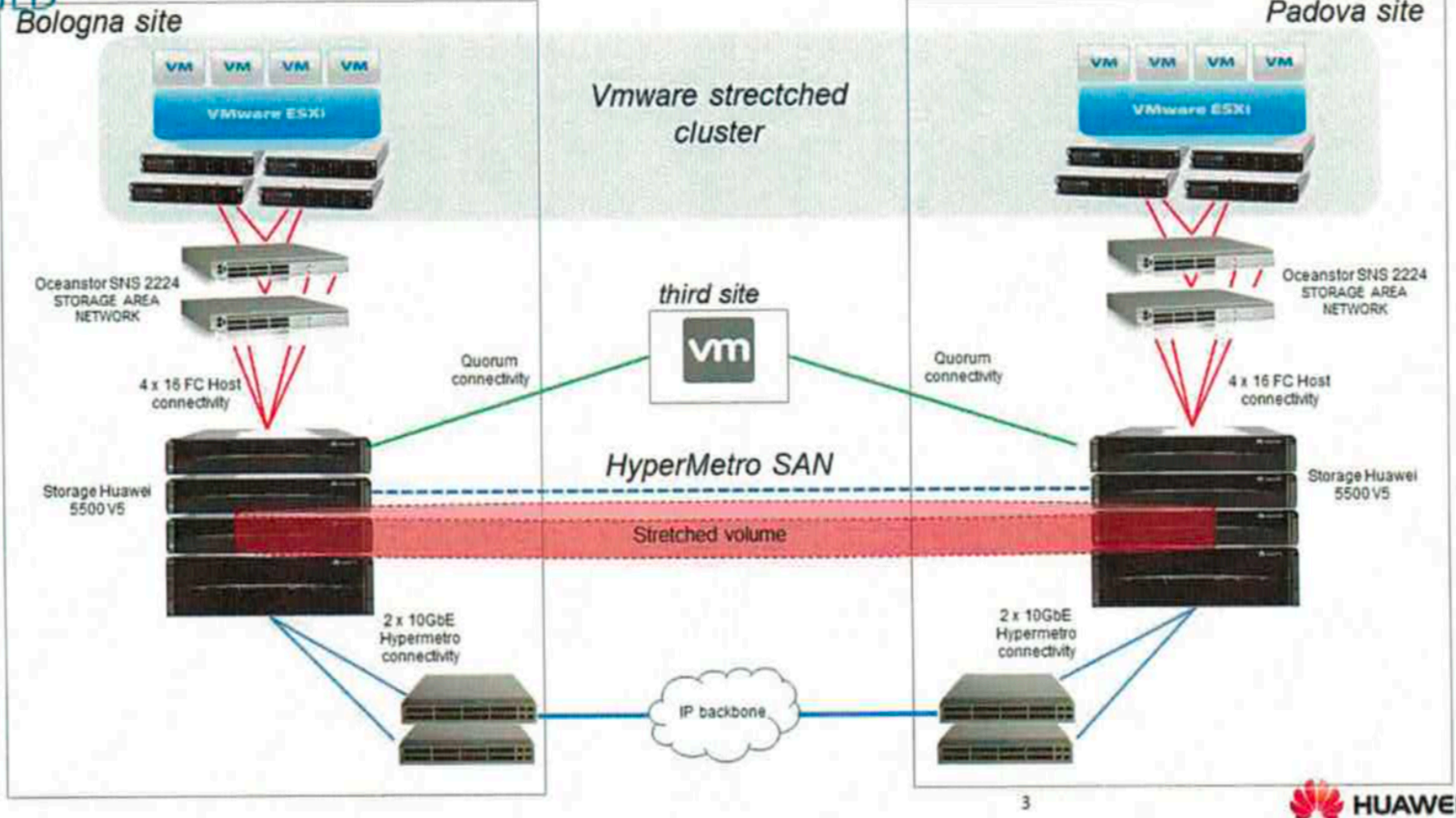
- ▶ Comprensiva di DR per adempiere alla corrente normativa e per facilitare recovery di eventuali disastri

Recap [2/3]



Recap [3/3]

HLD

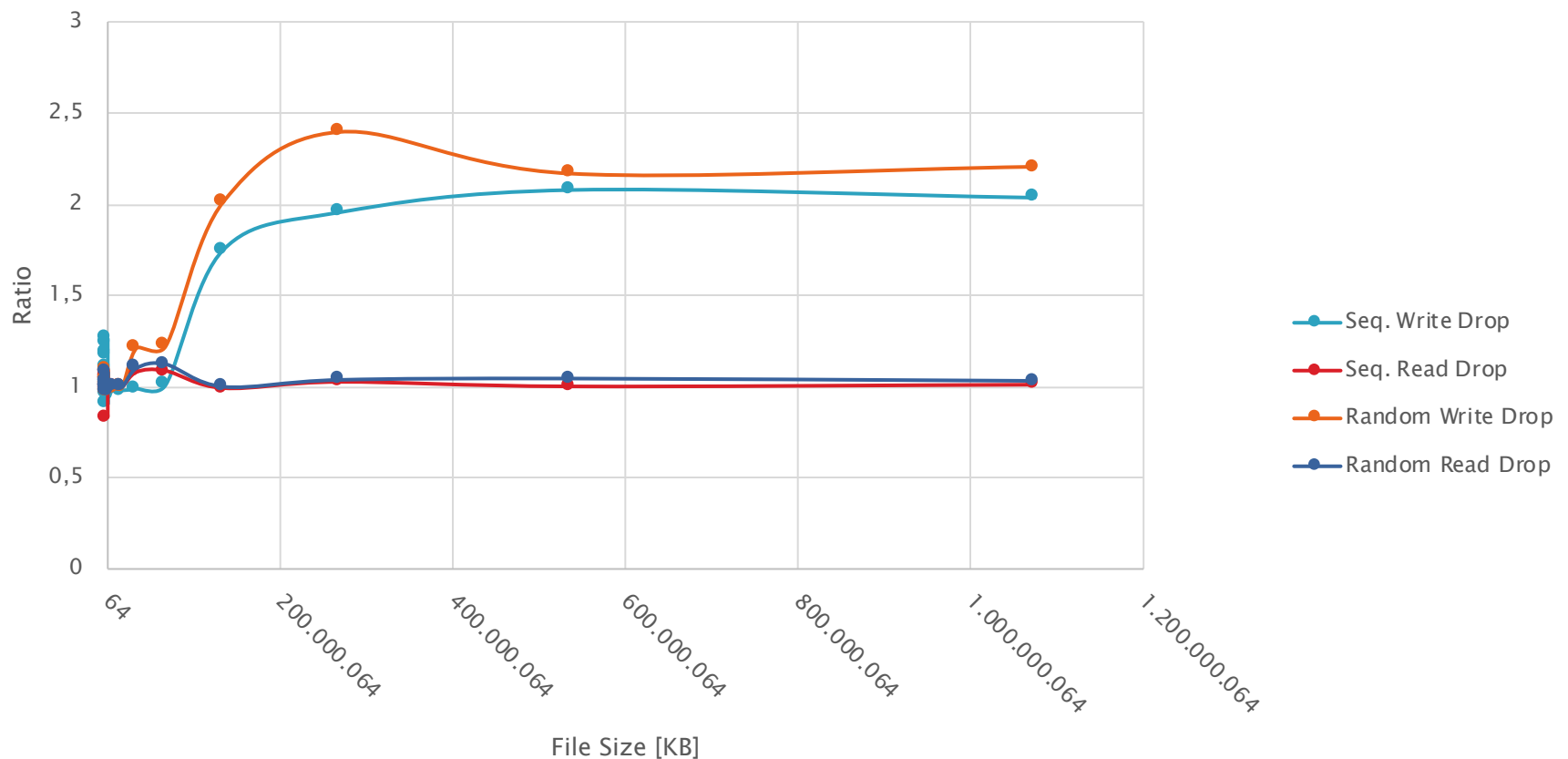


Misure SAN

- ▶ Setup hardware e prima serie di misure della SAN presentate alla scorsa riunione di Parma (https://agenda.infn.it/event/17960/contributions/87383/attachments/62087/74283/SSNN_03_04_2019.pdf)
- ▶ Misure throughput (Parma) effettuate con IOZone. Seconda trancia test di performance effettuate con Flexible I/O Tester, ottenendo risultati analoghi
- ▶ Effettuate misure di latenza in aggiunta al throughput
- ▶ Verificato comportamento in condizioni fallimento HW

Locale vs Hypermetro [1 / 3]

Performance Drop



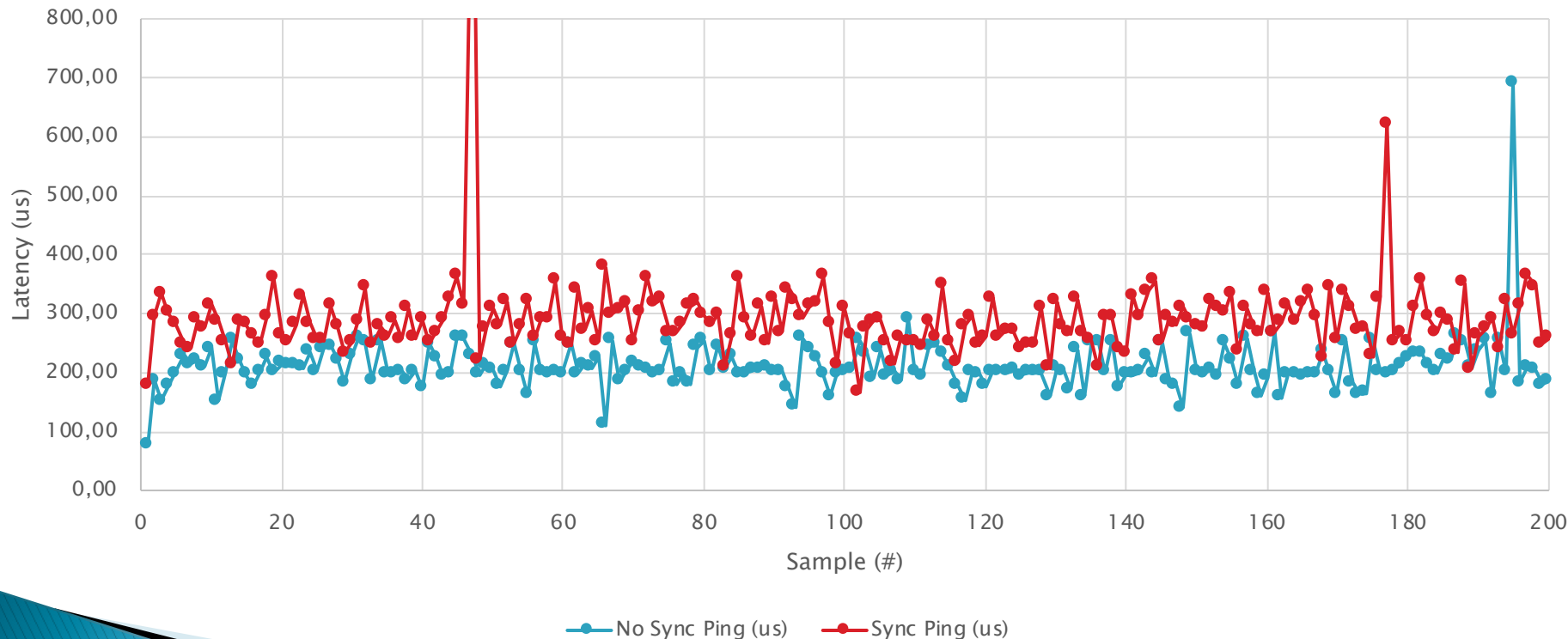
Locale vs Hypermetro [2 / 3]

Latenza LUN indipendente vs LUN in Hypermetro (ioping 4KiB packet)

LUN Indipendente ~ 210us

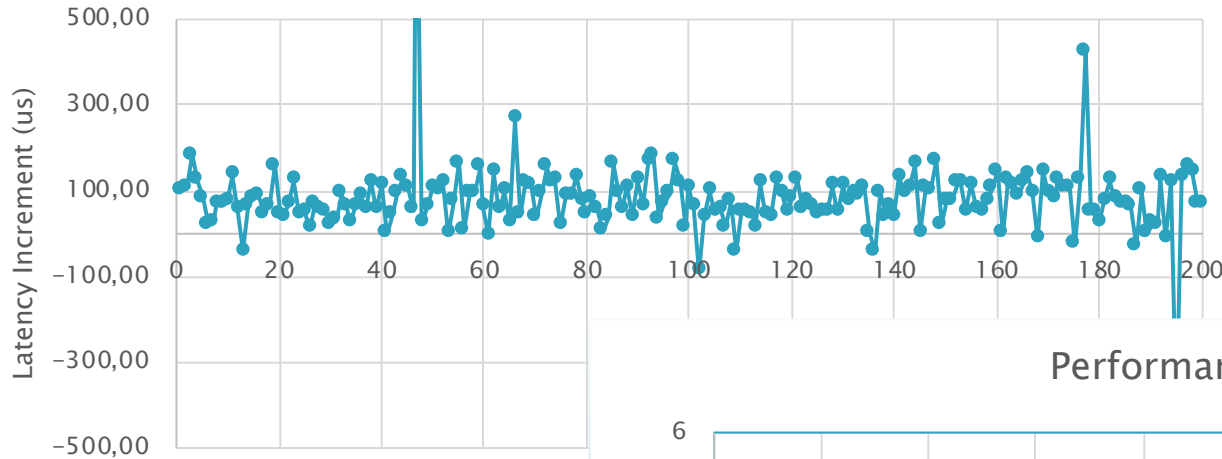
LUN Hypermetro ~ 290us

Latency



Locale vs Hypermetro [3 / 3]

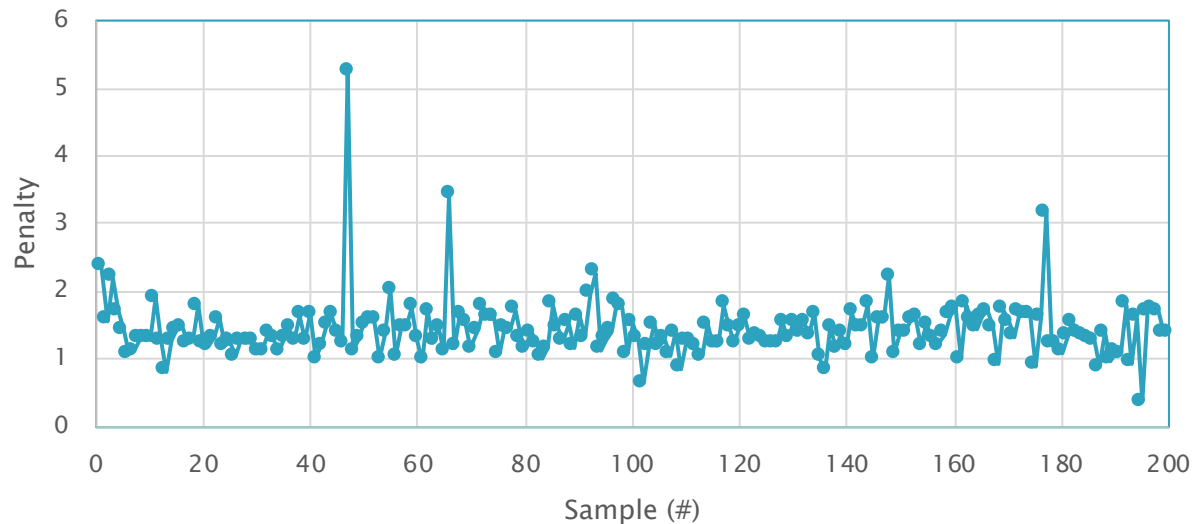
Latency Increment (us)



Incremento $\sim 80\mu s$

Latenza Sync/No
Sync $\sim 1,43$

Performance Drop



Infrastruttura [1 / 3]

Setup infrastruttura di virtualizzazione:

- ▶ Per i dettagli relativi all'Hardware/Connettività, rimando alla presentazione della scorsa riunione di Parma (https://agenda.infn.it/event/17960/contributions/87383/attachments/62087/74283/SSNN_03_04_2019.pdf)
- ▶ 6 Nodi M640 – VMware 6.5
- ▶ 2x10Gb dedicati al management
- ▶ 2x10Gb dedicati ai servizi
- ▶ 2 LUN da 10TB in Hypermetro (modalità Highest)
- ▶ Preferred site CNAF

Infrastruttura [2/3]

vmware vSphere Web Client

Navigator

- userv-test-20.inf.n.it
- vcneter-sn-cnaf.inf.n.it
- vdaprotection-hero-sn-cnaf.inf.n.it
- vdaprotection-sn-cnaf-02.inf.n.it
- vdaprotection-sn-cnaf-03.inf.n.it
- vdaprotection-sn-cnaf.inf.n.it
- vista.inf.n.it
- vrealize-log.inf.n.it
- vrealize-operations.inf.n.it
- vreplication-01-sn-cnaf.inf.n.it
- vreplication-02-sn-cnaf.inf.n.it
- web.inf.n.it
- www-presid.inf.n.it
- www-e-rhs.eu
- www-e-rhs.it
- www-extreme-datacloud.eu
- www-iperionch.eu
- www-iperionhs.eu
- xmpp.inf.n.it
- SSNN-GHA
 - ssnn-gha-cnaf-enc01-bl-01.cnaf.inf.n.it
 - ssnn-gha-cnaf-enc01-bl-02.cnaf.inf.n.it
 - ssnn-gha-cnaf-enc01-bl-04.cnaf.inf.n.it
 - ssnn-gha-inl-enc01-bl-01.cnaf.inf.n.it
 - ssnn-gha-inl-enc01-bl-02.cnaf.inf.n.it
 - ssnn-gha-inl-enc01-bl-04.cnaf.inf.n.it
 - ssnn-gha-cnaf-netmon.cnaf.inf.n.it
 - ssnn-gha-inl-netmon.cnaf.inf.n.it
 - test-mgmt
 - test1-www

M640

Getting Started | **Summary** | Monitor | Configure | Permissions | Hosts | VMs | Datastores | Networks | Update Manager

M640
Total Processors: 168
Total vMotion Migrations: 20

CPU: FREE: 434.93 GHz
USED: 696.00 MHz CAPACITY: 435.62 GHz
MEMORY: FREE: 2.21 TB
USED: 28.85 GB CAPACITY: 2.24 TB
STORAGE: FREE: 20.99 TB
USED: 53.10 GB CAPACITY: 21.04 TB

vSphere HA initiated a failover action on 1 virtual machines in cluster M640 in datacenter SSNN-GHA

Cluster Resources

Hosts	6 Hosts
Total virtual flash resources	0,00 B
EVC mode	Disabled

vSphere Availability

vSphere HA: Protected

CPU	0%	50%	100%
Memory	0%	50%	100%

CPU reserved for failover: 50 %
Memory reserved for failover: 50 %
Proactive HA: No provider applied
Host Monitoring: Enabled
VM Monitoring: Disabled

Tags

Assigned Tag	Category	Description
This list is empty		

vSphere DRS

Perfectly Balanced

Migration automation level: Fully Automated
Migration threshold: Apply priority 1, priority 2, and priority 3 recommendations
Power management automation level: Off
DRS recommendations: 0
DRS faults: 0

Cluster Consumers

Resource pools	0
vApps	0
Virtual machines	

Related Objects

Work in Progress

- Deploy OVF Template (4)
- Deploy OVF Template (3)
- Deploy OVF Template (2)
- Deploy OVF Template (1)
- userv-test-01.inf.n.it - Ed...
- Deploy OVF Template
- Add Host
- New Virtual Machine
- Configure Replication for x...

Alarms

All (2) | New (1) | Acknowl...

- vcneter-sn-cnaf.inf.n.it
- VMware vAPI Endpoint Service H...
- M630
- vSphere HA failover in progress

Recent Objects

Viewed	Created
M640	
SSNN-GHA	
ssnn-gha-cnaf-enc01-bl...	
ssnn-gha-cnaf-enc01-bl...	
ssnn-gha-cnaf-enc01-bl...	

Recent Tasks

Task Name	Target	Status	Initiator	Queued For	Start Time	Completion Time	Server
Check new notifications	vcneter-sn-cnaf.inf...	✓ Completed	VMware vSphere Up...	606 ms	27/05/2019 19:03:01	27/05/2019 19:03:02	vcneter-sn-cnaf.inf.n.it
Migrate virtual machine	userv-03.inf.n.it	✓ Completed	System	3 ms	27/05/2019 18:55:50	27/05/2019 18:56:27	vcneter-sn-cnaf.inf.n.it

Infrastruttura [3/3]

Configurazione:

- ▶ Singolo Cluster (vSphere HA)
- ▶ Soglia di ingresso 50% (CPU e RAM)
- ▶ DRS automatizzato sull'intero cluster
- ▶ 1 «Isolation address» (da aumentare per produzione)
- ▶ Permanent Device Loss + All Path Down (Conservative Power Off and Restart)
- ▶ Località definita attraverso gruppi di host
- ▶ Definite «must/should rules» per imporre o suggerire affinità VM/Host

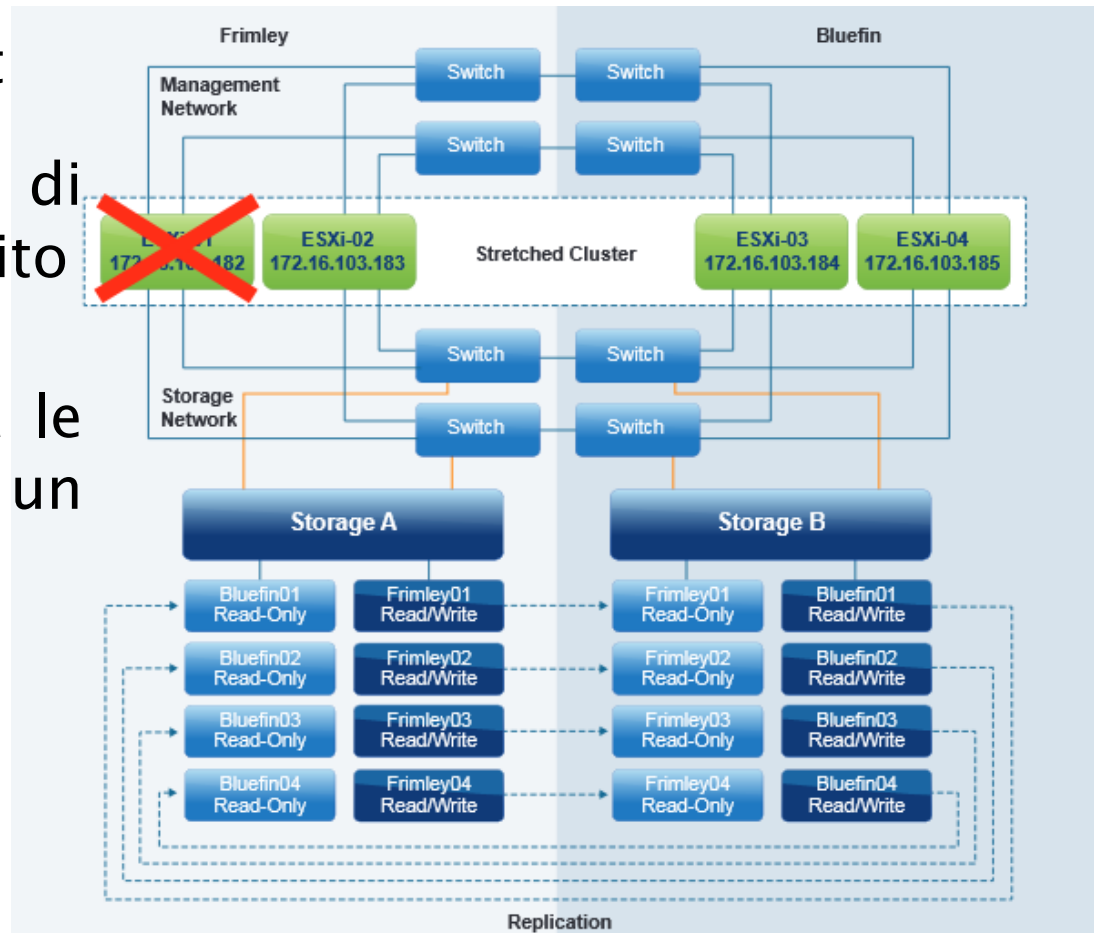
Scenari Testati [1 / 8]

Fallimento singolo host

- ▶ Simulazione rottura di un hypervisor in un sito

vSphere HA rischedula le VM in esecuzione su un altro host disponibile

Esito: OK



Scenari Testati [2 / 8]

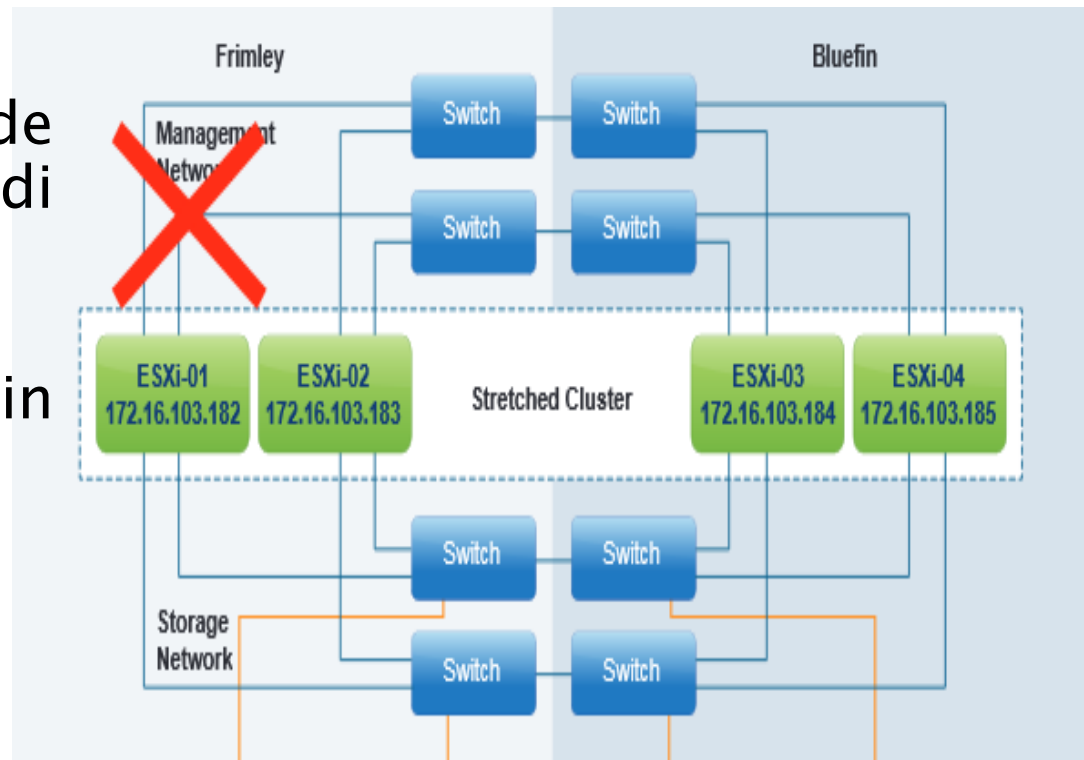
Isolamento singolo host

- ▶ Un hypervisor perde connettività alla rete di management

Le VM rimangono in esecuzione sull'hypervisor

Esito: OK

Da valutare diverso setup



Scenari Testati [3 / 8]

Partizionamento Storage

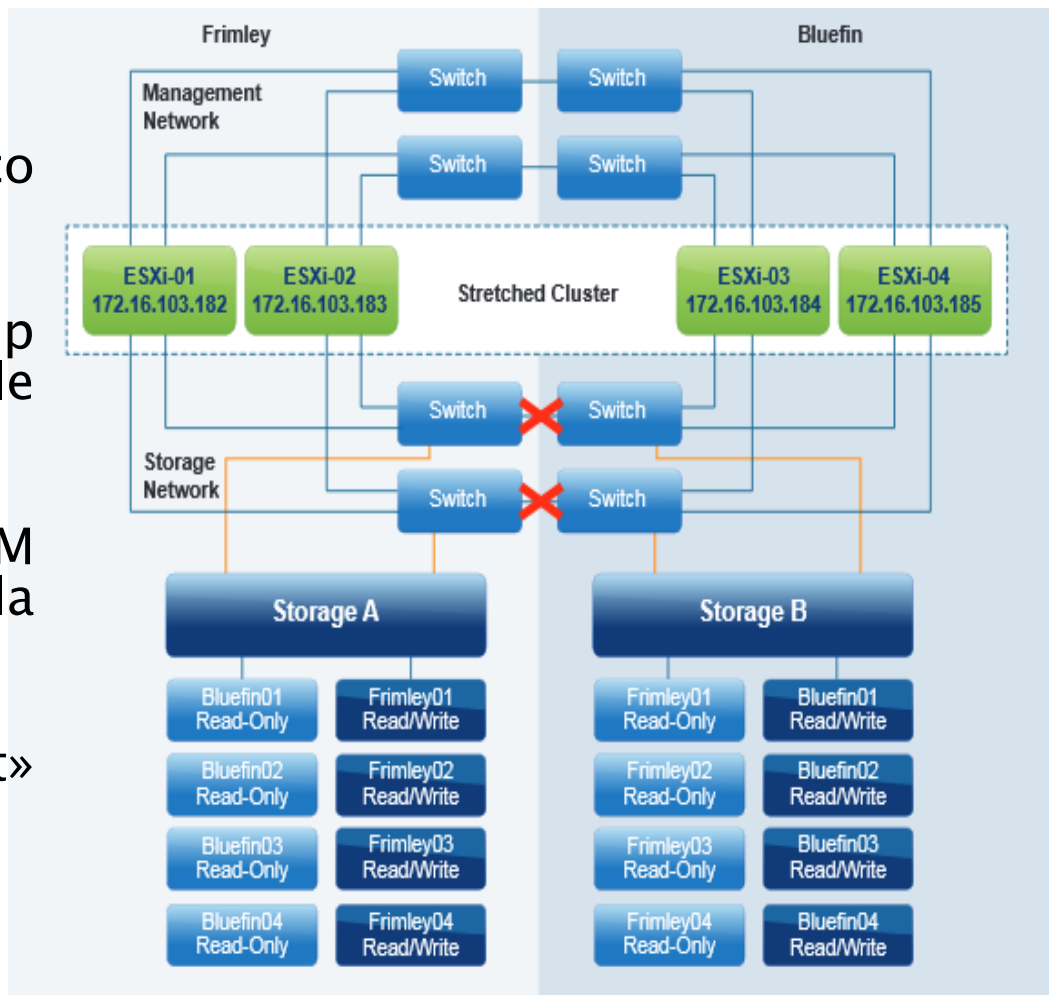
- ▶ Perdita del link geografico per la sincronizzazione

In funzione del setup «preferred site» delle singole LUN viene segnalato il PDL

vSphere HA rischedula le VM sugli host con accesso alla SAN

VM associate a regola «must» vengono messe offline

Esito: OK



Scenari Testati [4/8]

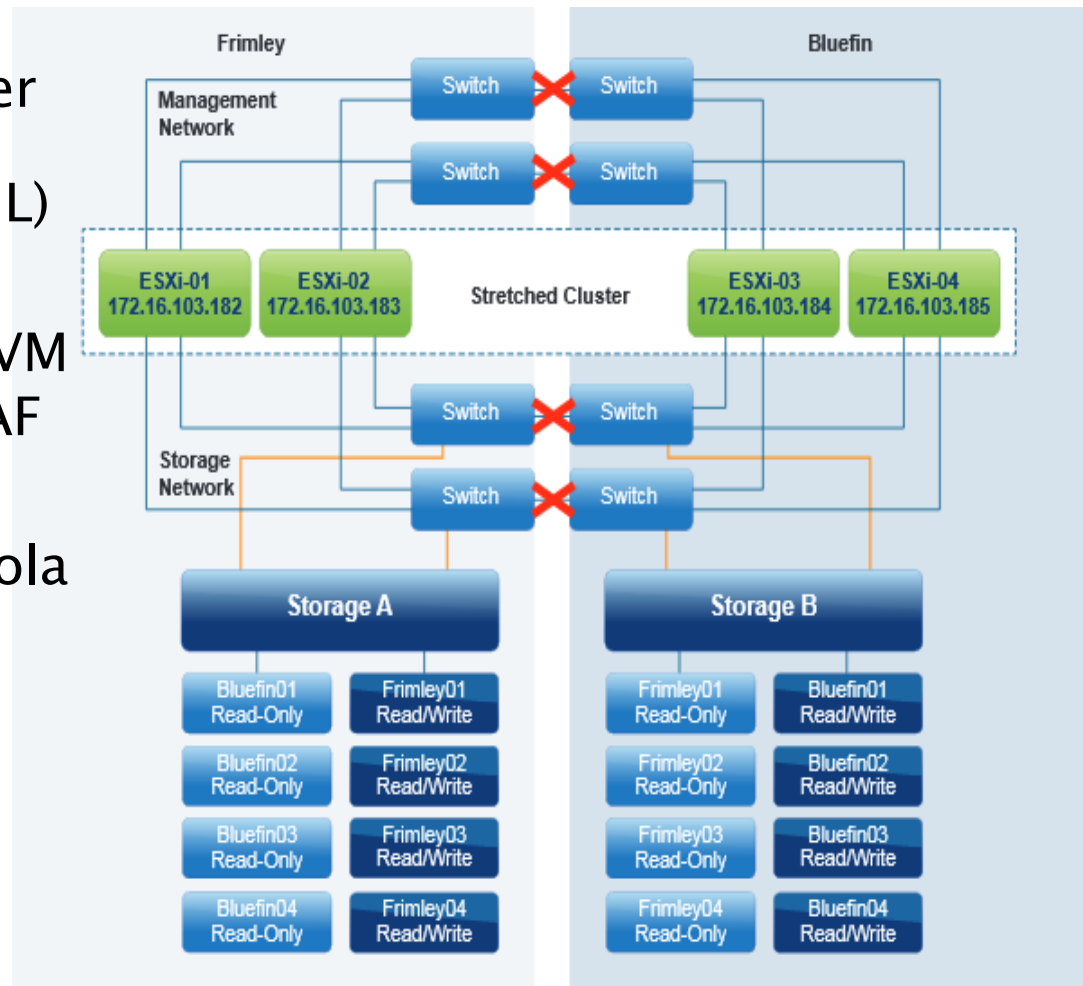
Partizionamento datacenter

- Un sito viene isolato (LNL)

vSphere HA rischedula le VM in esecuzione sul sito CNAF

VM associate a regola «must» rimangono offline

Esito: OK

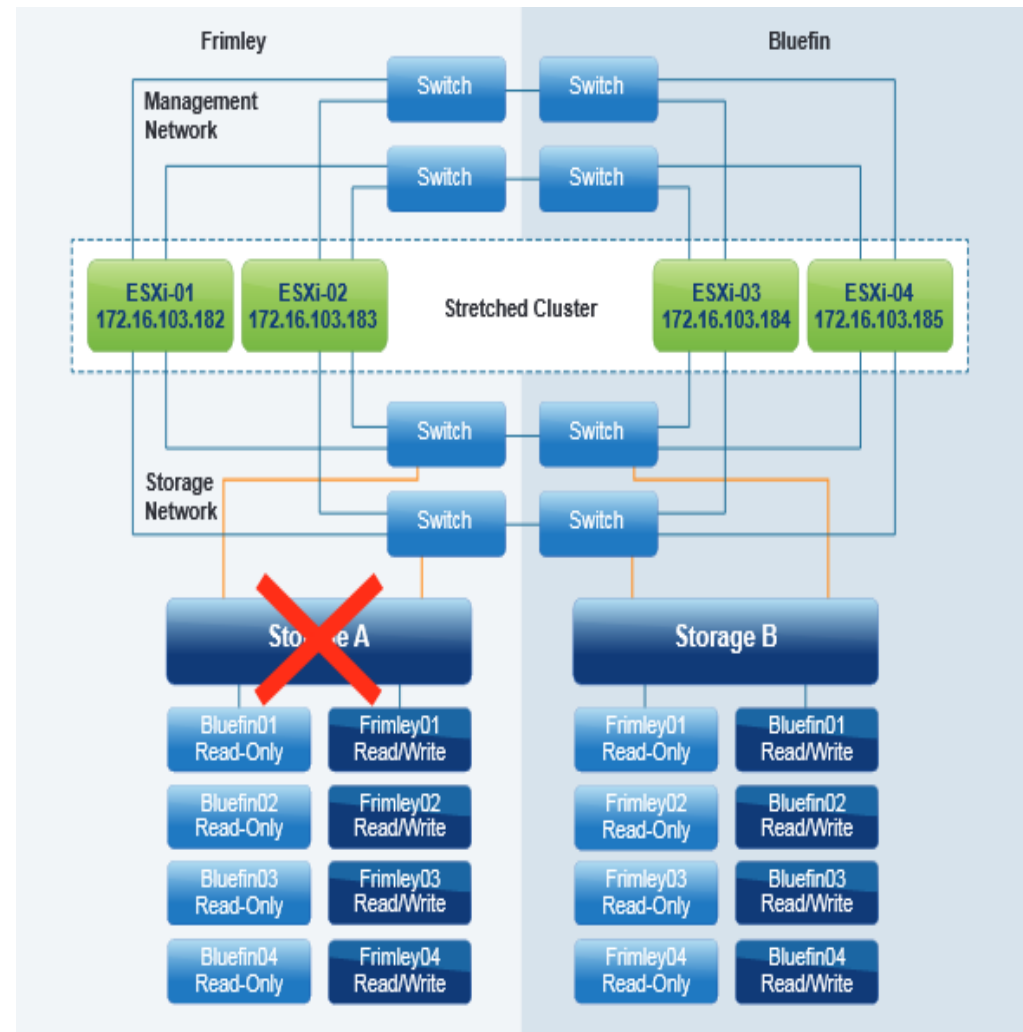


Scenari Testati [5 / 8]

Fallimento storage

- ▶ Simulazione rottura di un intero sistema di storage
- ▶ Il sistema risponde alla condizione di APD con i timeout previsti.
- ▶ Vengono messi offline gli hypervisor che hanno perso accesso allo storage
- ▶ Le VM vengono rischedulate nel datacenter con SAN superstite
- ▶ VM associate a regola «must» rimangono offline

Esito: OK

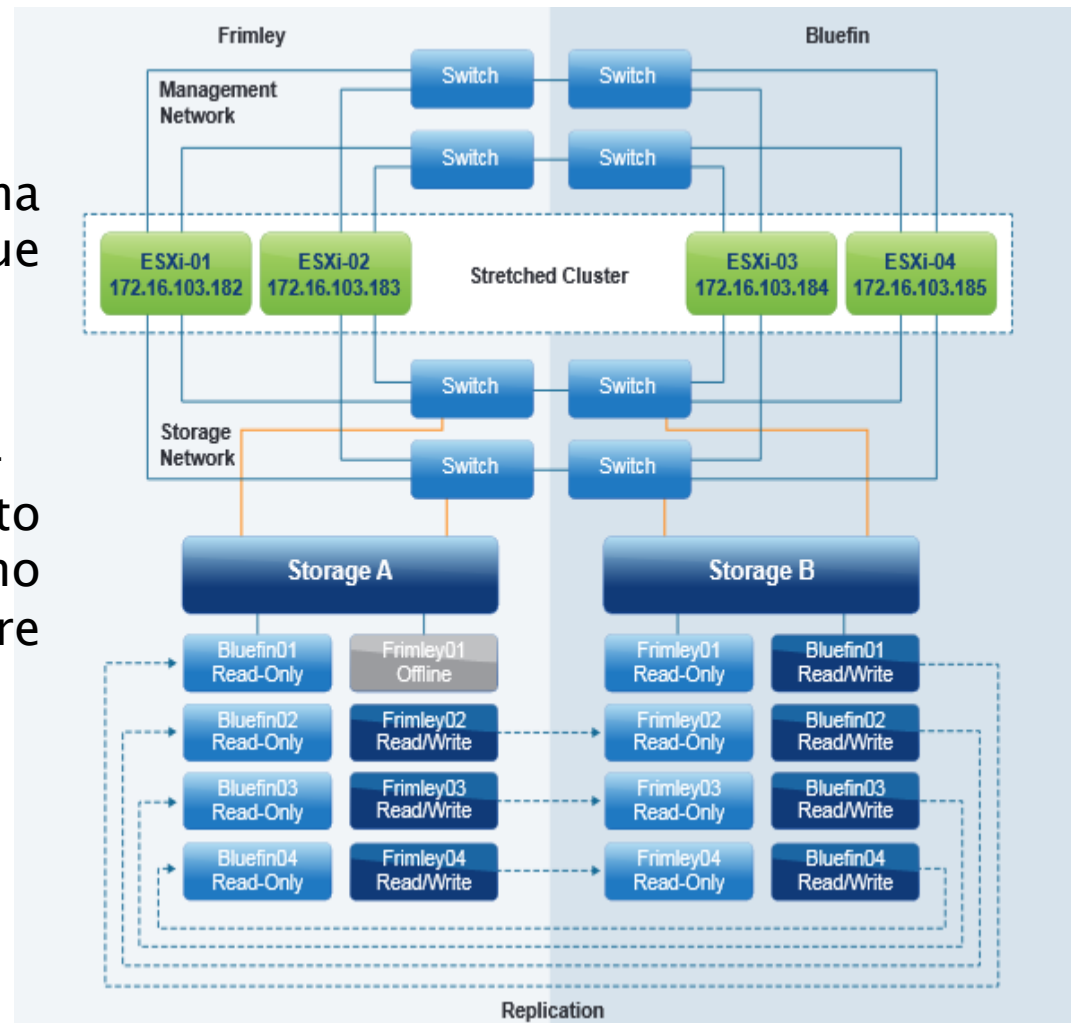


Scenari Testati [6/8]

Permanent Device Loss

- ▶ Viene perso l'accesso ad una LUN in uno dei due datacenter
- ▶ vSphere HA intercetta il PDL
- ▶ Le VM in esecuzione nel sito che genera il PDL vengono spente e fatte ripartire nell'altro datacenter

Esito: OK

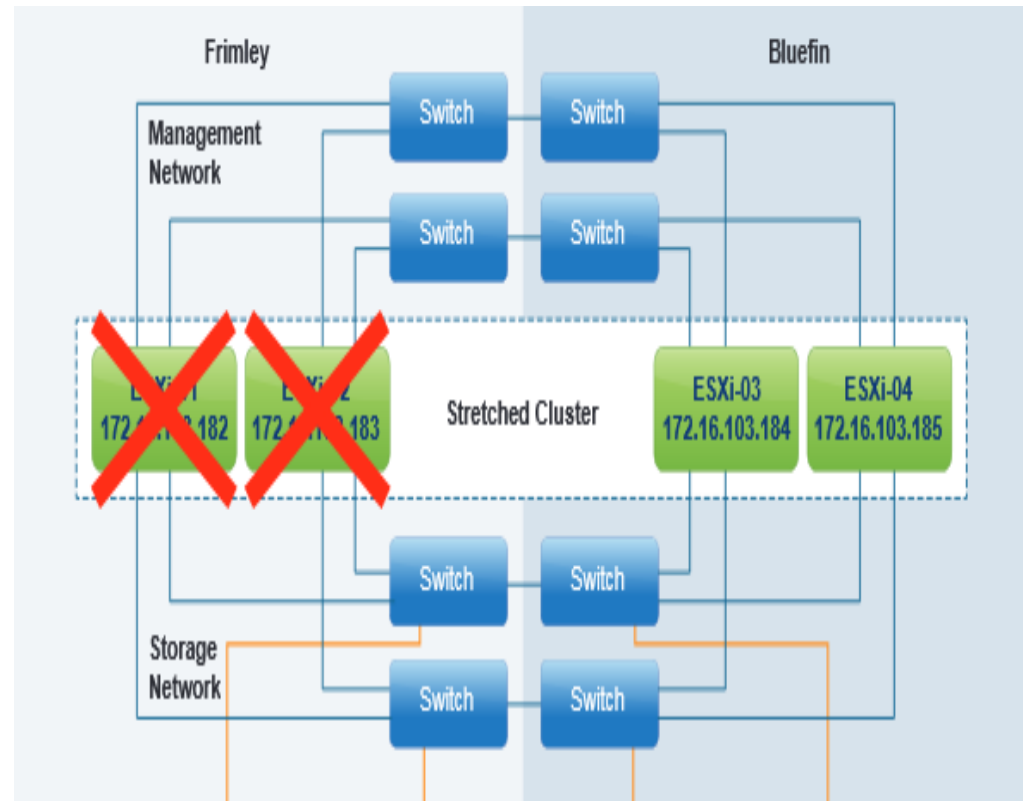


Scenari Testati [7/8]

Fallimento di tutti gli hypervisor di un datacenter

- ▶ Tutti gli hypervisor di un datacenter risultano non funzionanti
- ▶ vSphere HA rischedula autonomamente l'esecuzione delle VM nell'altro datacenter
- ▶ VM associate a regola «must» rimangono offline

Esito: OK

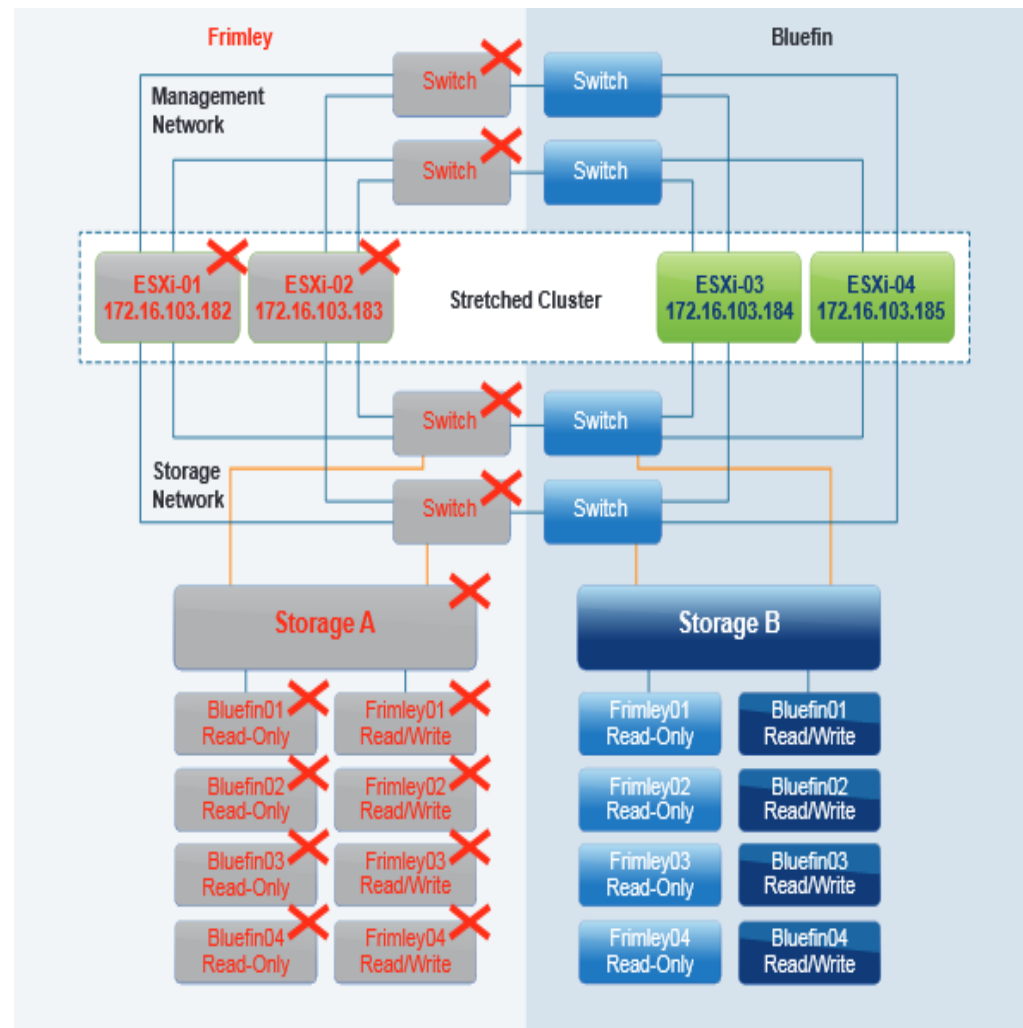


Scenari Testati [8/8]

Perdita di un datacenter

- ▶ Un datacenter (LNL) risulta completamente inaccessibile
- ▶ vSphere HA rischedula autonomamente l'esecuzione delle VM nell'altro datacenter
- ▶ VM associate a regola «must» rimangono offline

Esito: OK



WhatNext?

L'infrastruttura di virtualizzazione (Hypervisor + SAN) è operativa
Per arrivare al setup di design i prossimi step sono:

- ▶ «Virtualizzare» i gateway delle diverse VLAN
- ▶ Annunciare le sottoreti di servizio anche da LNL (GARR)
- ▶ Definire le regole di accesso alla VLAN dei servizi (ACL + Firewall)
- ▶ Posizionare il quorum server della SAN nel datacenter designato (LNF)
- ▶ Disporre dell'infrastruttura di virtualizzazione necessaria per il DR

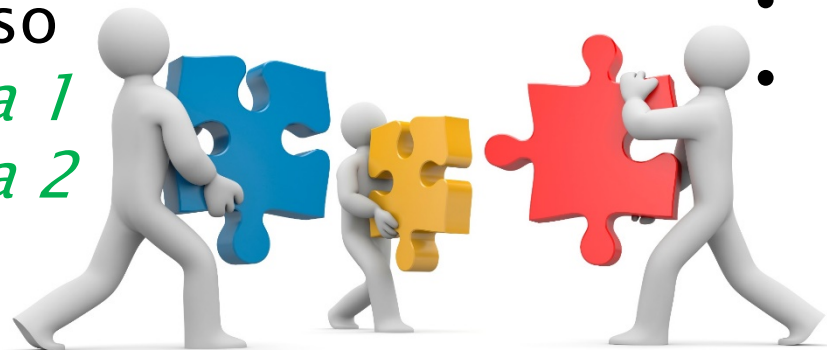
Parallelamente al setup della parte network, andranno condotti i test sulla sincronizzazione/HA dei DBMS Oracle (SI)

Nell'infrastruttura è stato installato Vmware 6.5 per consentire continuità al sistema di DR. Terminata la migrazione di quest'ultimo si procederà all'upgrade a Vmware vSphere 6.7

Il Team

Servizi Nazionali:

- Stefano Antonelli
- Stefano Longo
- Felice Rosso
- *Assegnista 1*
- *Assegnista 2*



Calcolo@LNL:

- Luciano Berti
- Massimo Biasotto
- Sergio Fantinel
- Michele Gulmini

NET@CNAF:

- Lorenzo Chiarelli
- Donato De Girolamo
- Stefano Zani

Q&A