

ITP09 – Alta disponibilità e site resilience in Exchange Server 2013



Michele Betelli

micheleb@icubed.es

<http://www.icubed.es>



Grazie a



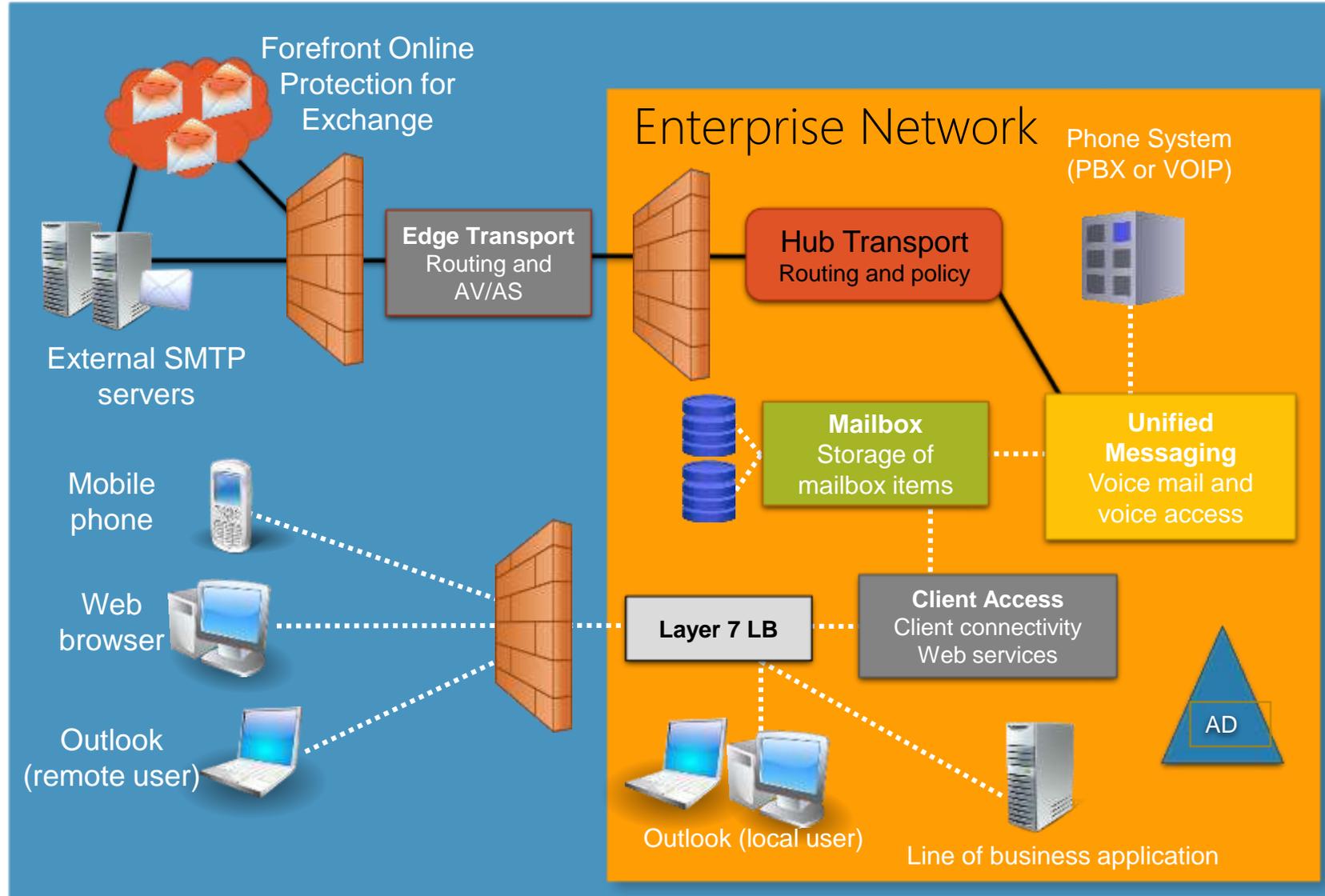
Sponsor



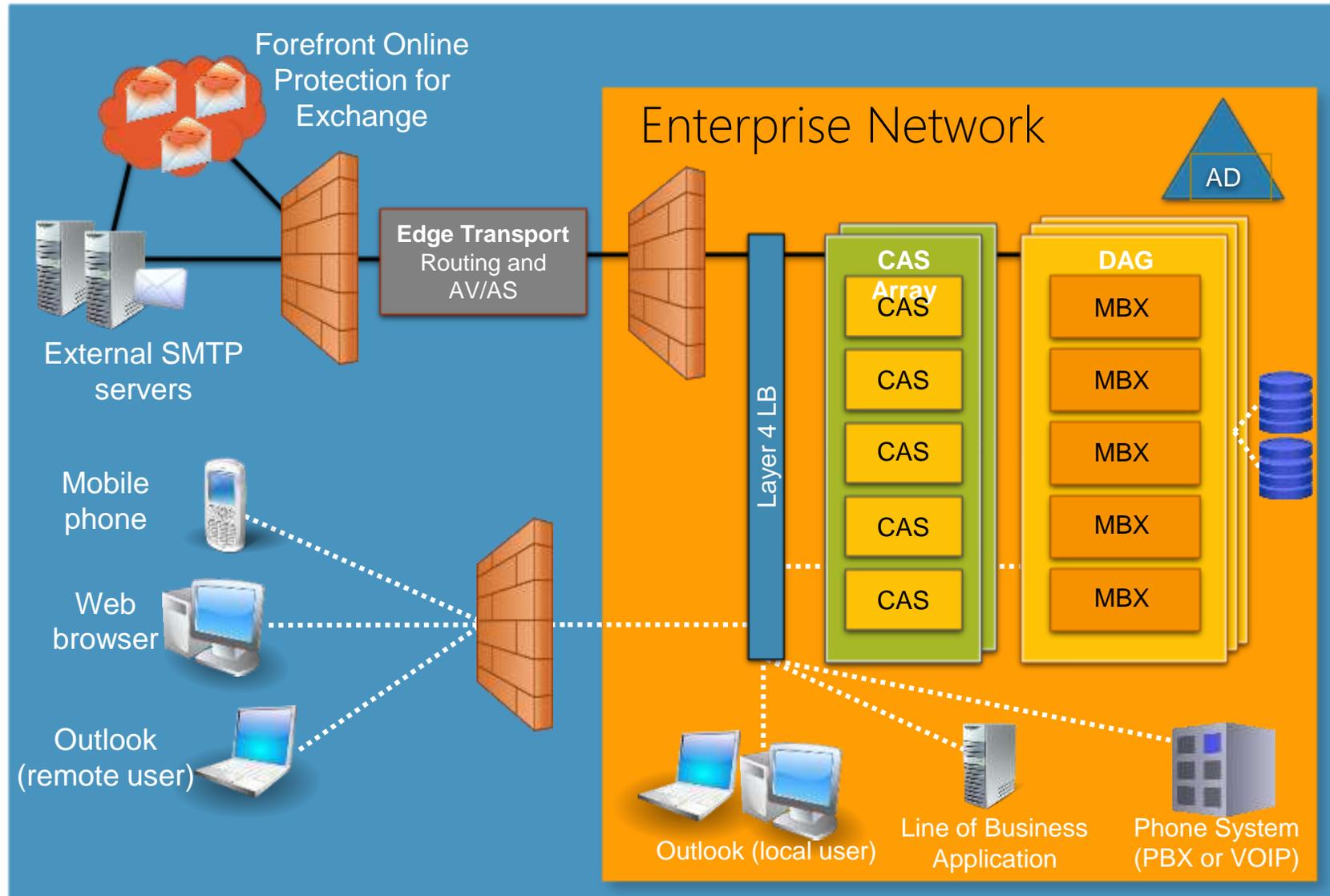
Agenda

- Ripasso
 - Architettura Exchange Server 2010
 - Architettura Exchange Server 2013
- Alta disponibilità
 - Database Availability Groups
 - Load Balancing
 - DNS Round Robin
 - Network Load Balancing
 - Hardware Load Balancing L4 e L7
- Site resilience

Architettura Exchange Server 2010



Architettura Exchange Server 2013



Alta disponibilità

- Abbiamo solamente 2 ruoli da proteggere (senza contare Edge Server)
- Alta disponibilità per il ruolo Mailbox usando i DAGs
- Alta disponibilità per il ruolo CAS usando load balancing

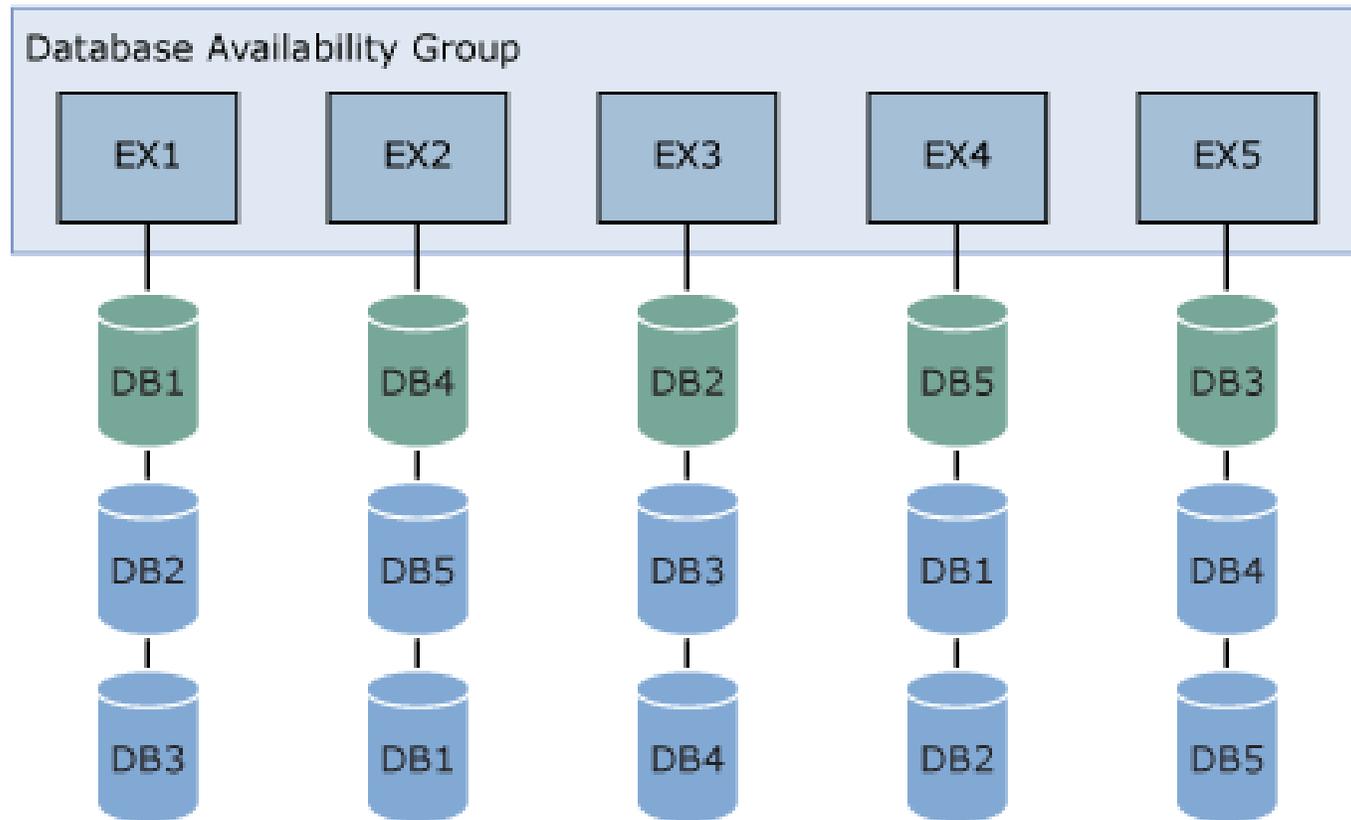
Database Availability Groups

- Gruppo di Mailbox Server (fino a 16 nodi)
- Replica a livello applicativo
- Ogni server ha il suo storage
- Il failover lo abbiamo a livello di DB e non di server
- Utilizza Failover Cluster di Windows (attenzione alla versione di 2008 server)
- Può coesistere con il ruolo di Client Access (*)
- Può essere esteso a site remoti

* Failover Cluster non è compatibile con NLB

Database Availability Groups

DAG with five members



Database Availability Groups (preprequisiti)

- I nodi devono far parte tutti dello stesso dominio AD
- I nodi devono avere lo stesso sistema operativo
- Il sistema operativo deve supportare Failover Cluster
- Il percorso dei DB e LOG dovrà essere uguale su tutti i nodi
- Witness server in caso di nodi pari
- Due schede di rete per ogni nodo su due subnet distinte isolate
- Massimo 500ms di latenza a livello network

Novità dei Database Availability Groups

- Managed Availability (restore del servizio automatico)
- Reseed automatico (copia dal DB attivo verso disco di spare)
- Configurazione automatica delle Network (default, ma posso farlo manualmente)
- Nuovi CmdLet per la gestione con PowerShell
- La Best Copy Selection tiene conto della disponibilità dei singoli servizi (OWA, POP, IMAP, etc...)
- Miglioramenti alle copie Lagged
 - Integrazione con Safety Net
 - Replay automatico in caso di necessità (spazio disco esaurito, etc.)
- Nuovi controlli per avviso allarmi sulla Single Copy

Load Balancing

- DNS Round Robin
- Network Load Balancing
- Hardware Load Balancing

DNS Round Robin

- Un solo nome DNS con molti indirizzi IP
- Non vale per i protocolli non HTTP (es. POP3 o IMAP4)
- Quasi tutto il traffico verso Exchange 2013 è HTTP (anche MAPI)
- I client salta alla seconda IP se nel primo risponde con un Hard TCP Failure
- Metodo utilizzato per Site Resiliency
 - Se perdetevi il VIP verso il bilanciatore del primo sito puntate al bilanciatore del secondo
- Il namespace non è più uno SPOF

Svantaggi di DNS Round Robin

- Non fa il controllo della salute del server
- Non fa il controllo del carico di lavoro
- Il server perso rimane nella lista degli IP validi se non rimosso a mano
- Il TTL del record DNS e applicativo client potrebbe impattare sul bilanciamento

Network Load Balancing

- E' disponibile in tutte le versione di Windows Server in forma gratuita
- Un solo nome e una sola IP condivisa fra i membri del cluster NLB
- Usa un suo algoritmo per bilanciare il traffico
- Garantisce alta affidabilità quando nota una failure e ridistribuendo il traffico tra i nodi restanti
- Fornisce affinity a livello di IP sorgente

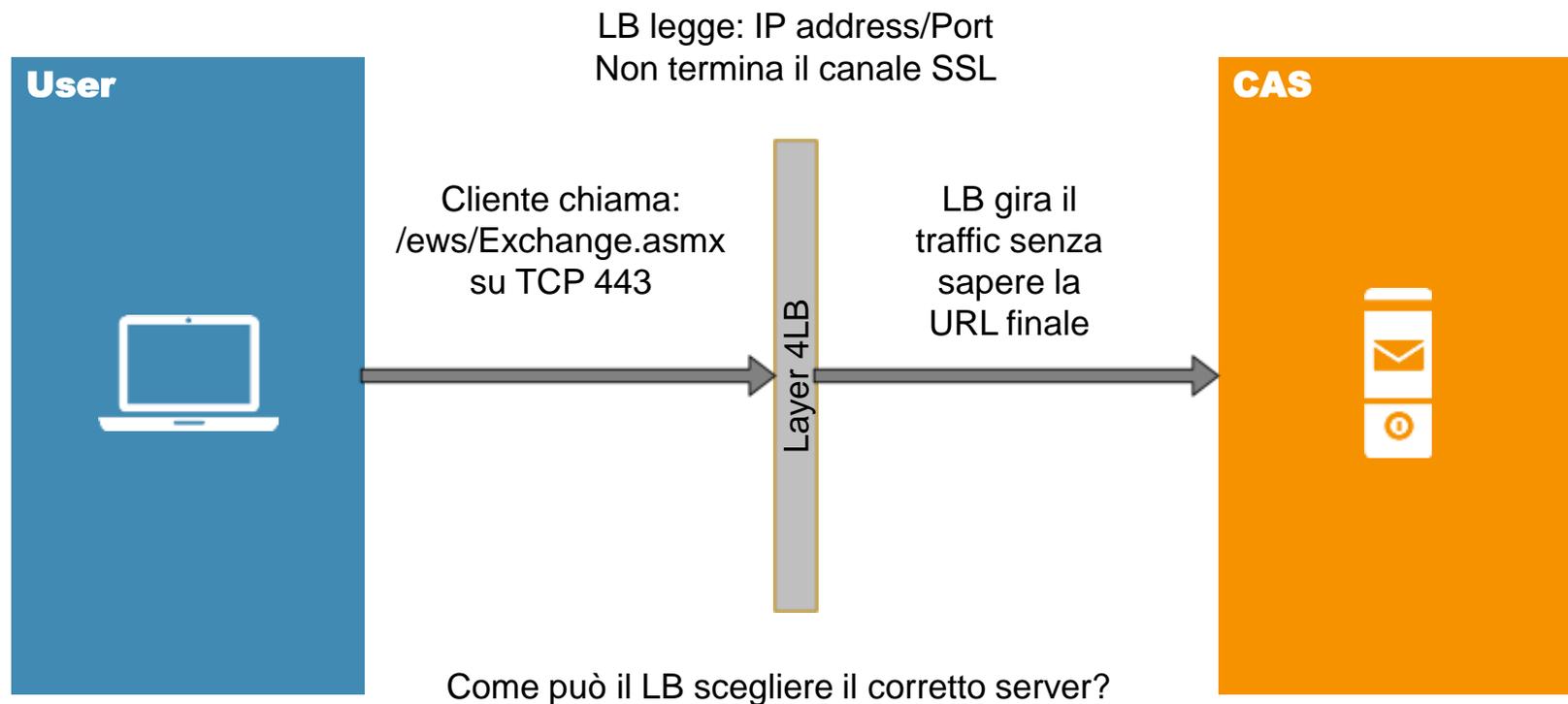
Svantaggi di NLB

- Non è raccomandato se abbiamo più di 8 server
- Non è in grado di riconoscere problemi a livello applicativo
- Gestisce l'affinity a livello IP quindi non molto utile se dietro NAT
- Non supportato su membri di DAG per la presenza di Failover Cluster
- Può produrre Port Flooding

Hardware Load Balancing

- Disponibile come hardware o come macchina virtuale (anche in Azure)
- Possono fornire bilanciamento a Livello 4 (TCP) o Livello 7 (applicativo)
- I bilanciatori a L4 sono più semplici da configurare rispetto a quelli L7
- Lavorando a L4 non sappiamo esattamente dove è destinata la richiesta, il bilanciatore la invierà ai server Exchange ma non saprà se è una richiesta OWA piuttosto che Autodiscover o altro..
- Per lo stesso motivo a L4 non potremo fare un probe al servizio richiesto, ma faremo un check generico al server di destino
- Lavorando a L7 posso terminare il canale SSL e leggere e gestire l'intero pacchetto

Bilanciando a L4...



Controllo di salute a L4...

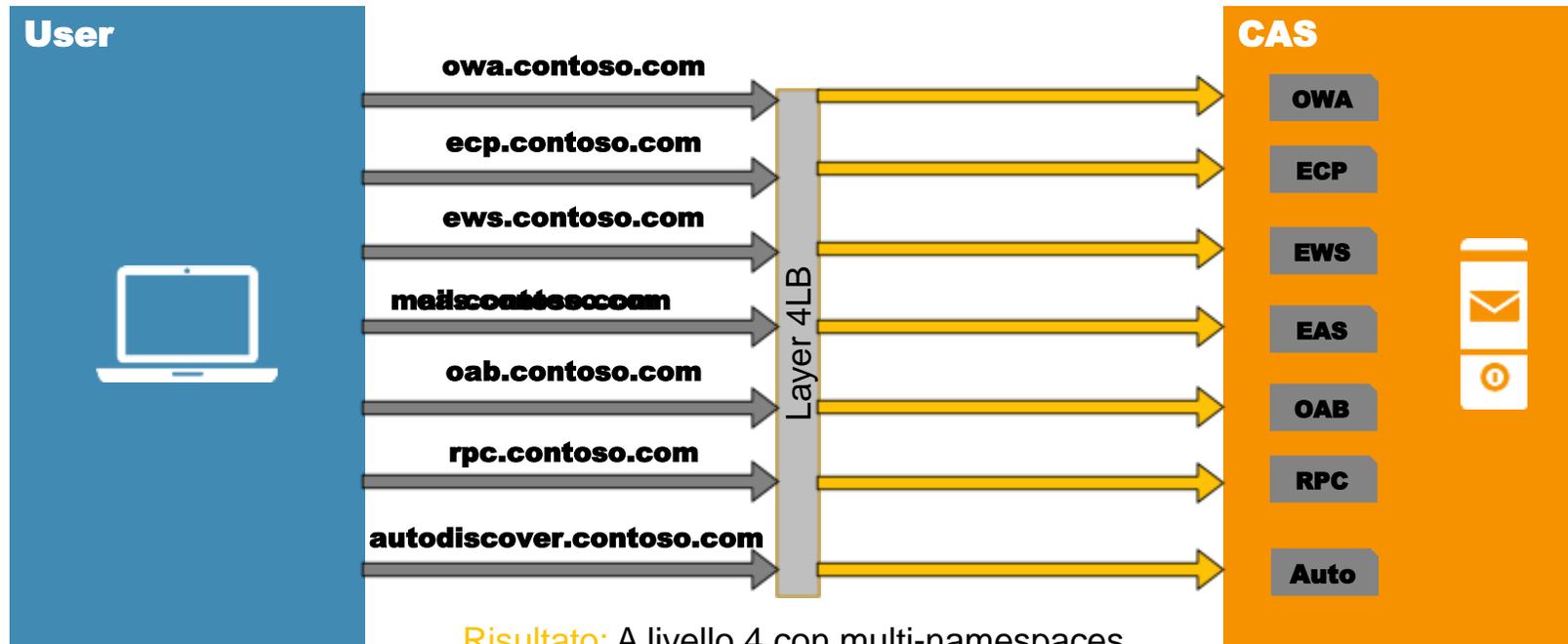
Se possiamo scegliere una sola Vdir per determinare lo stato di tutto il server, quale scegliamo?



Risultato: A livello 4 con un solo namespace il check è per server e NON per protocollo

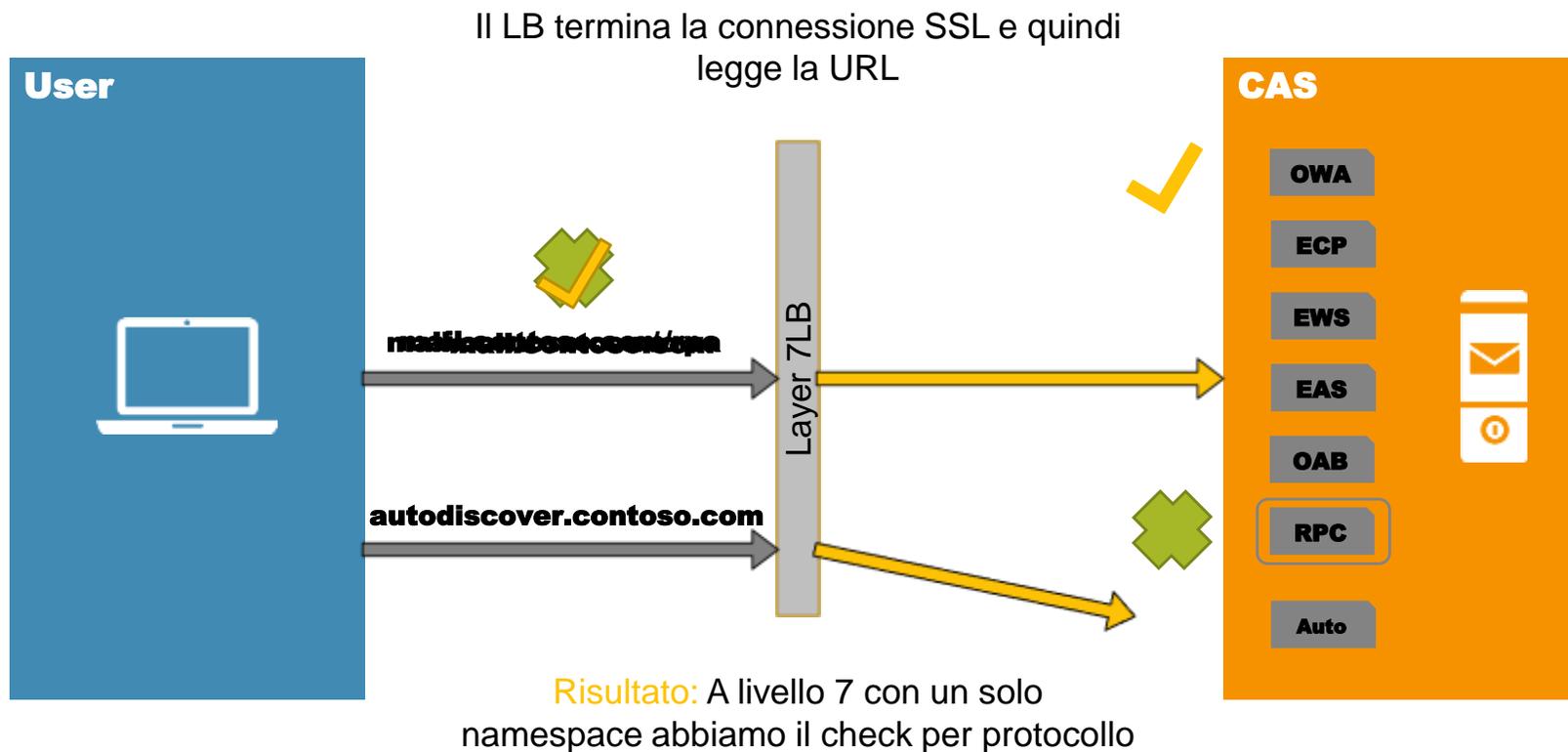
L4 con multi-namespaces

Un VIP per ogni namespace



Risultato: A livello 4 con multi-namespaces, il check è per protocollo.

Controllo di salute a L7



LB + Managed Availability

- Exchange 2013 include una pagina web comandata dal Managed Availability per sapere se un determinato servizio sta funzionando correttamente
 - <https://mail.contoso.com/ews/healthcheck.htm>
 - <https://mail.contoso.com/oab/healthcheck.htm>
- Se il servizio é OK allora la pagina risponderà un 200 OK
- Questo controllo può essere utilizzato dal LB per verificare il servizio richiesto

Riassunto

- DNS Round Robin
 - Non è un bilanciatore di carico
 - Buono per site resilience
 - I client HTTP fanno failover automaticamente
- Network Load Balancing
 - Bilanciatore di tipo software con limitazioni
- Hardware Load Balancing
 - Bilanciatore di tipo hardware o VM che fa controlli di salute del server
 - Lavorando a L7 posso avere un check esatto del servizio che viene richiesto

Site resilience in Exchange 2010

- Ambiente complesso
- Il restore dei ruoli dovevano essere collegati, CAS e MBX insieme
- Il namespace era un Single Point of Failure

Site resilience in Exchange 2013

- Le procedure di *-over sono state semplificate
- Adesso il restore di MBX e CAS è indipendente
- Il namespace garantisce la ridondanza e non è un SPOF
- Posso implementare failover automatico a livello di datacenter

Procedure semplificate

- In precedenza la perdita di una parte dell'infrastruttura poteva portare a uno switchover del datacenter
- In Exchange 2013, il ripristino del servizio è praticamente automatico
- L'amministratore si deve preoccupare di risolvere il problema, non di ripristinare il servizio

Come lavorano CAS e MAILBOX

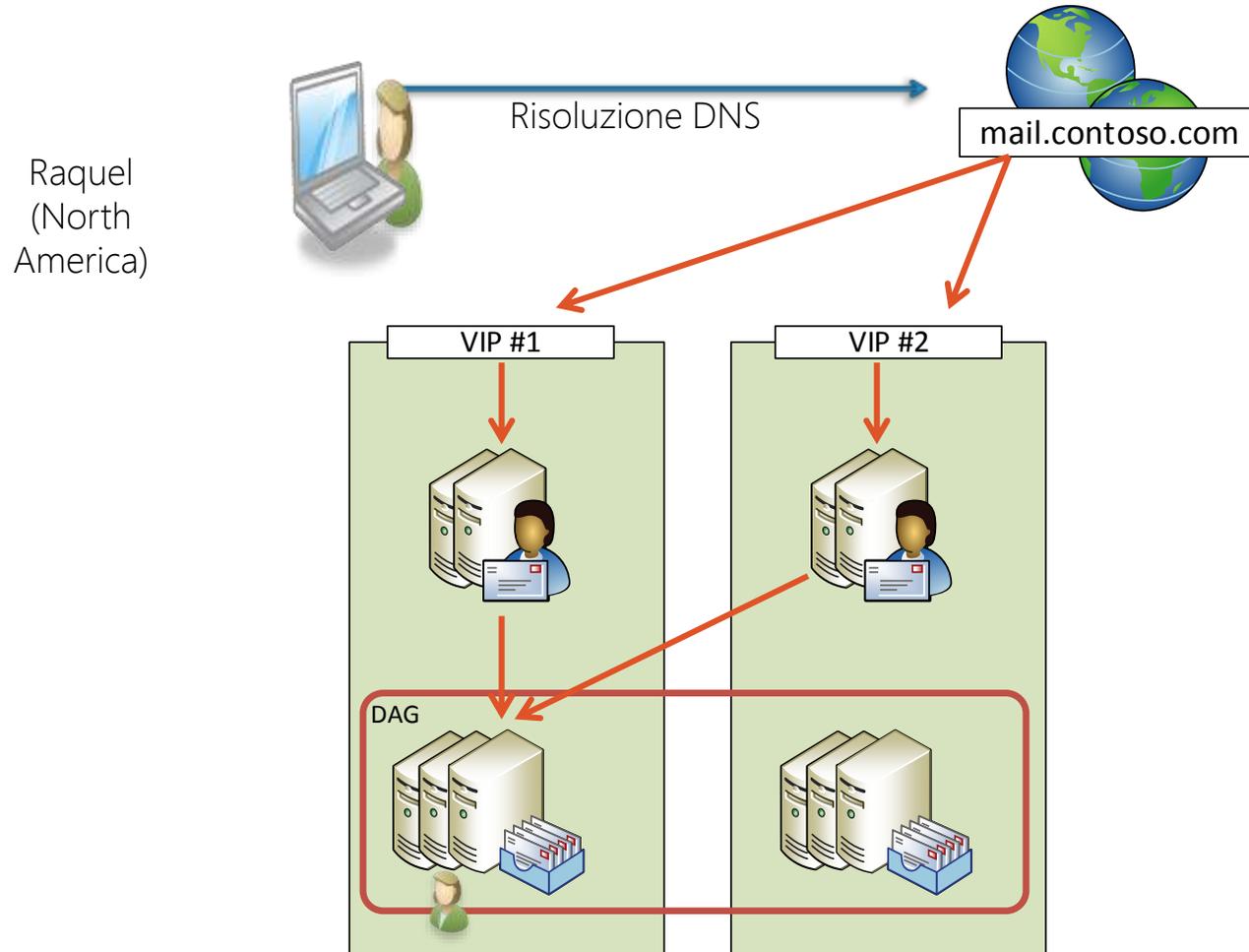
- In Exchange 2013 il CAS funziona solo da Proxy verso il ruolo MBX con il DB attivo
- Se passo da un CAS ad un altro, chi svolge il lavoro sarà sempre il MBX con il DB attivo
- In caso di failover del ruolo MBX il CAS reindirizzerà il traffico sul MBX dove il DB si è attivato, anche se si trova in un altro site

Site resilience + DNS Round Robin

- In Exchange 2013 quasi tutto si basa su HTTP
- I client HTTP hanno l'IP failover integrato
- Se un IP genera un Hard Failure, si passa al prossimo IP
- Si può fare uno switchover del site solo rimuovendo il VIP dal DNS
- Posso usare lo stesso namespace per datacenter differenti
- Non ci sono problemi di latenza per il cambio di record DNS (TTL)

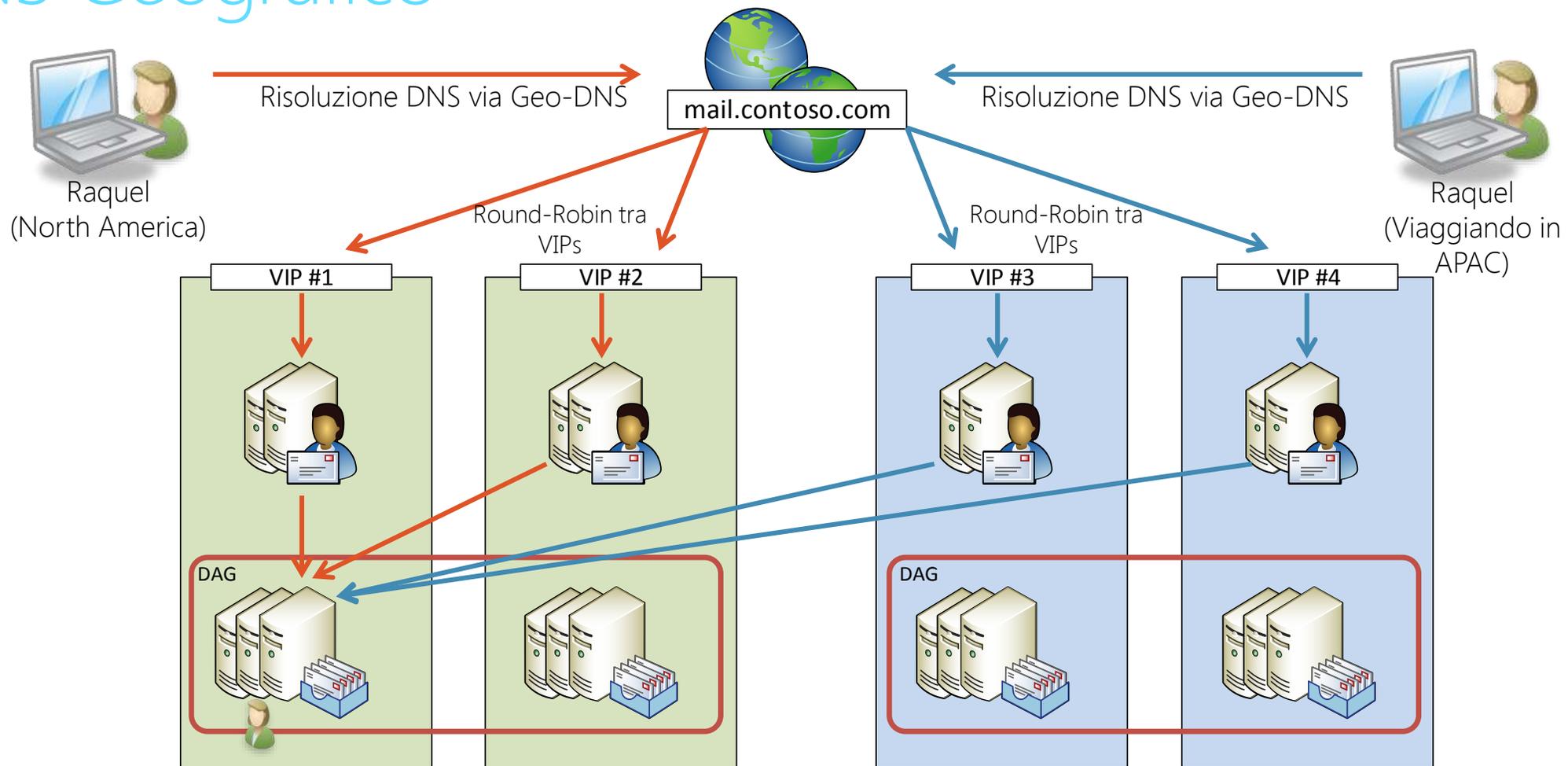
Esempio di un solo namespace

Round-Robin DNS



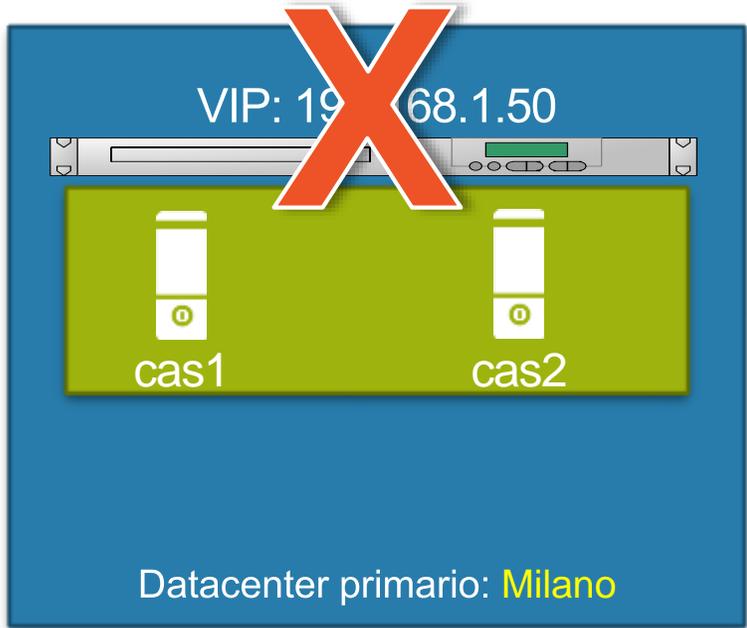
Esempio di un solo namespace

DNS Geografico

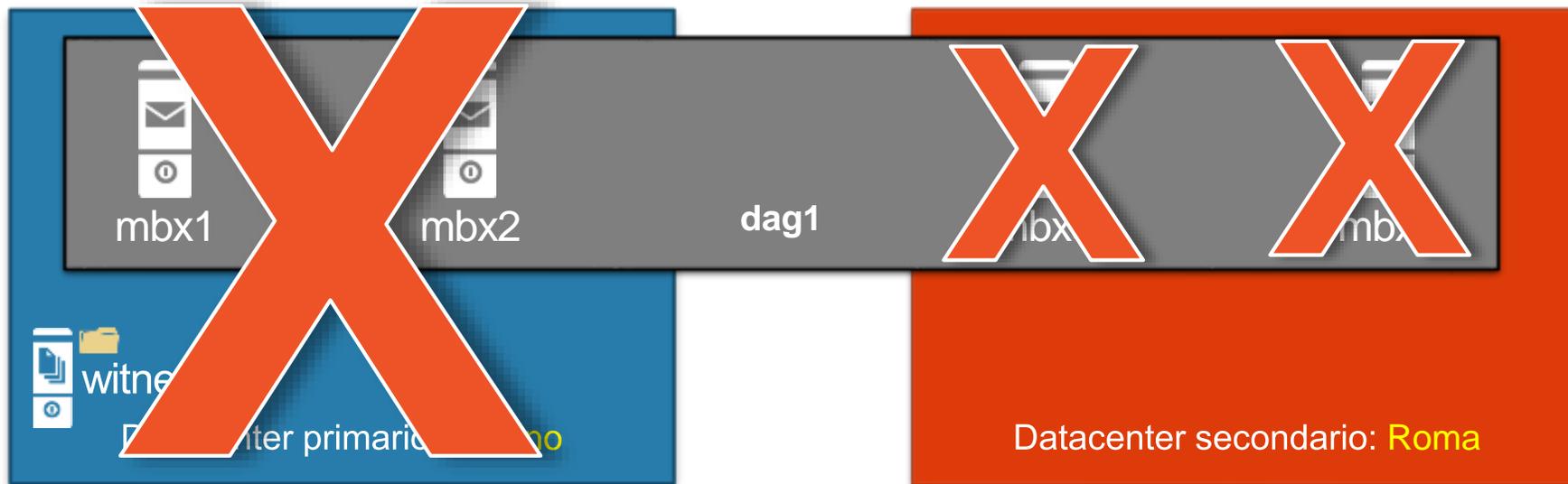


Site resilience con due sites (CAS)

Se si ha un'opzione IP del DNS che permette di controllare i servizi client si sposteranno automaticamente su un altro VIP
mail.contato.com 192.168.0.50, 10.0.1.50

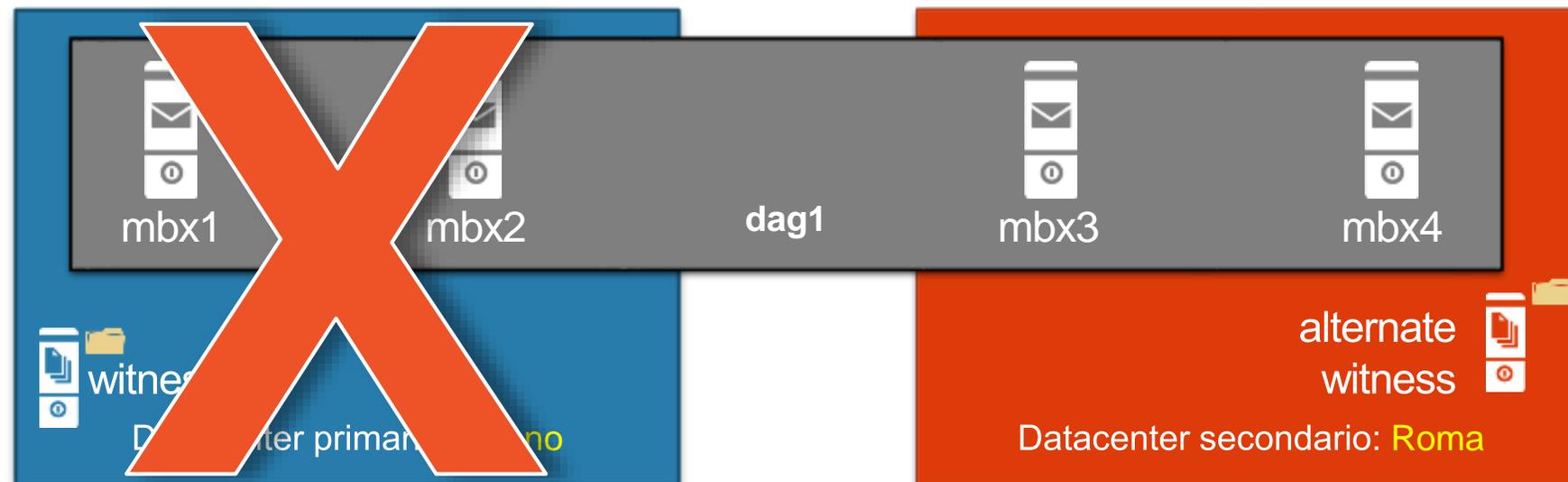


Site resilience con due sites (MBX)



Site resilience con due sites (MBX)

1. Configurare il site come down: `Stop-DatabaseAvailabilityGroup DAG1 –ActiveDirectorySite:Milano`
2. Stop del servizio cluster sugli altri nodi del DAG: `Stop-Clusvc`
3. Attivare il DAG nel secondo Datacenter: `Restore-DatabaseAvailabilityGroup DAG1 –ActiveDirectorySite:Roma`

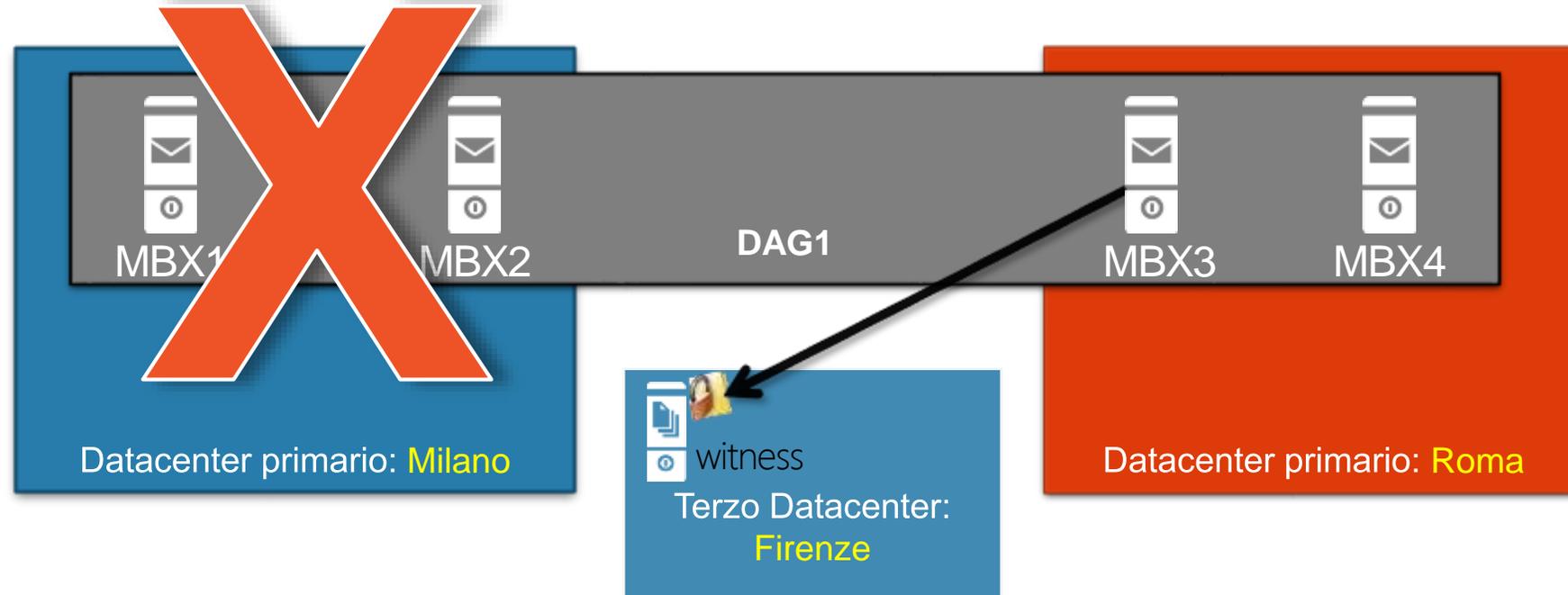


Site resilience e failover automatico

- Con tutte le modifiche alla nuova architettura, ora possiamo avere un failover automatico fra sites
- Requisiti:
 - 2 Datacenter per Exchange e 1 per File Share Witness (FSW) del DAG
 - La latenza tra i datacenter deve essere inferior a 250ms e la connessione affidabile
 - Ogni datacenter dovrà avere una connessione propria verso il site del FSW

Site resilience con tre sites

Assunto che MBX3 e MBX4 stiano lavorando e uno di loro possa connettersi alla FSW, avremo un failover automatico



Q&A

Tutto il materiale di questa sessione su

<http://www.communitydays.it/>

Lascia il feedback su questa sessione,
potrai essere estratto per i nostri premi!

Seguici su

Twitter @CommunityDaysIT

Facebook <http://facebook.com/cdaysit>

#CDays14

