

Apache e Mysql cluster

Affrontare il clustering da 4 a 400 nodi con strumenti open-source di livello enterprise (o meglio :>).

Deploy di ambienti clusterizzati (Load Balancing e High Availability).

Sommario:

1-3) Introduzione

4-7) Soluzioni da pover'uomo

8-10) Apache cluster (Load Balancing e High Availability)

11-12) Mysql cluster (Load Balancing e High Availability)

13) NAS/SAN (High Availability)

14) Domande?

Ragioni per clusterizzare

Load Balancing aka Bilanciamento del Carico:

- Un singolo server non può gestire il carico
- Macchine più veloci non sempre risolvono
- Differenti colli di bottiglia: Disk IO, CPU o Network
- SMP ha dei problemi intrinseci

High Availability aka Alta Disponibilità:

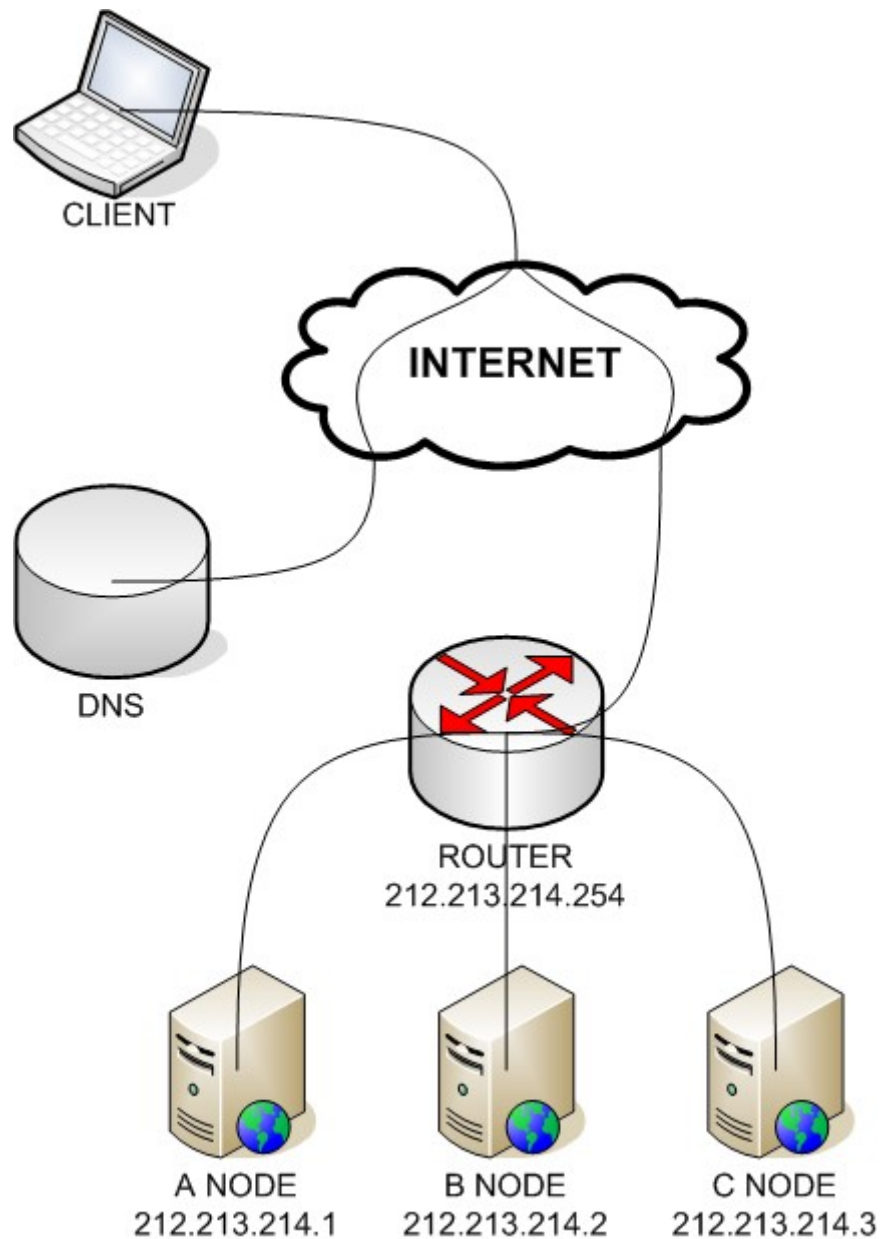
- Capacità di Failover
- High Availability (the 9s)

Scalability

Scalability: La facilità con cui un sistema o un componente può essere modificato per risolvere una categoria di problematiche.

- Riferita alla maggiore potenza di calcolo
- Un'applicazione scalabile permette di continuare ad aggiungere server senza aggiungere (molto) overhead
- Ovvero: Le performance dell'applicazione non soffrono per maggior carico o data size

Round Robin DNS



Utilizza il DNS per fornire IP differenti in maniera circolare

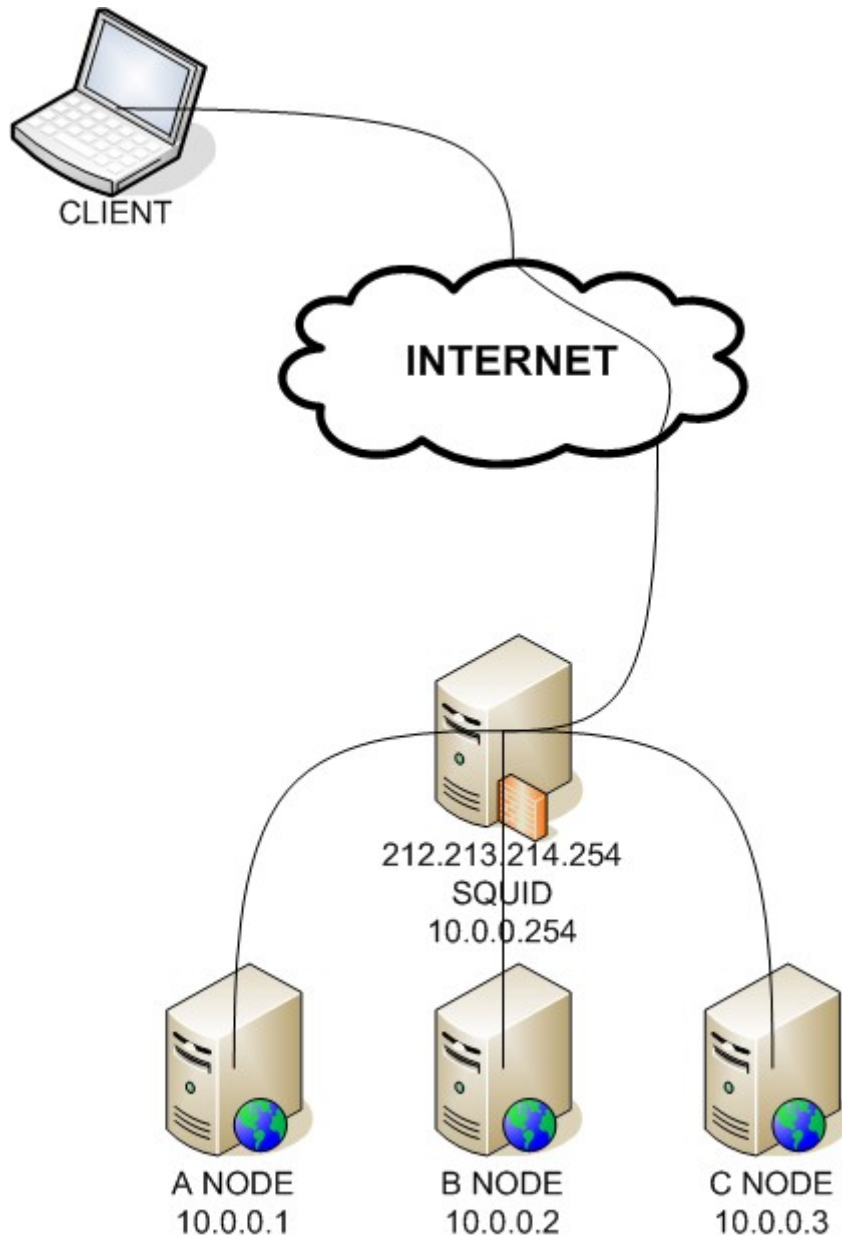
- Semplice da configurare
- Controllo limitato su quale server riceva la richiesta
- Il caching emula un meccanismo di conn-track

Esempio con BIND DNS:

```
www.ush.it. 60 IN A 212.213.214.1  
www.ush.it. 60 IN A 212.213.214.2  
www.ush.it. 60 IN A 212.213.214.3
```

- Un TTL breve è importante per evitare il caching

Squid Redirector



Squid come Front Proxy

- Squid è un HTTP Caching Proxy
- Può anche essere utilizzato come Front End Proxy
- Tramite uno script si può implementare un meccanismo di Load Balancing

Esempio con Squid Redirector:

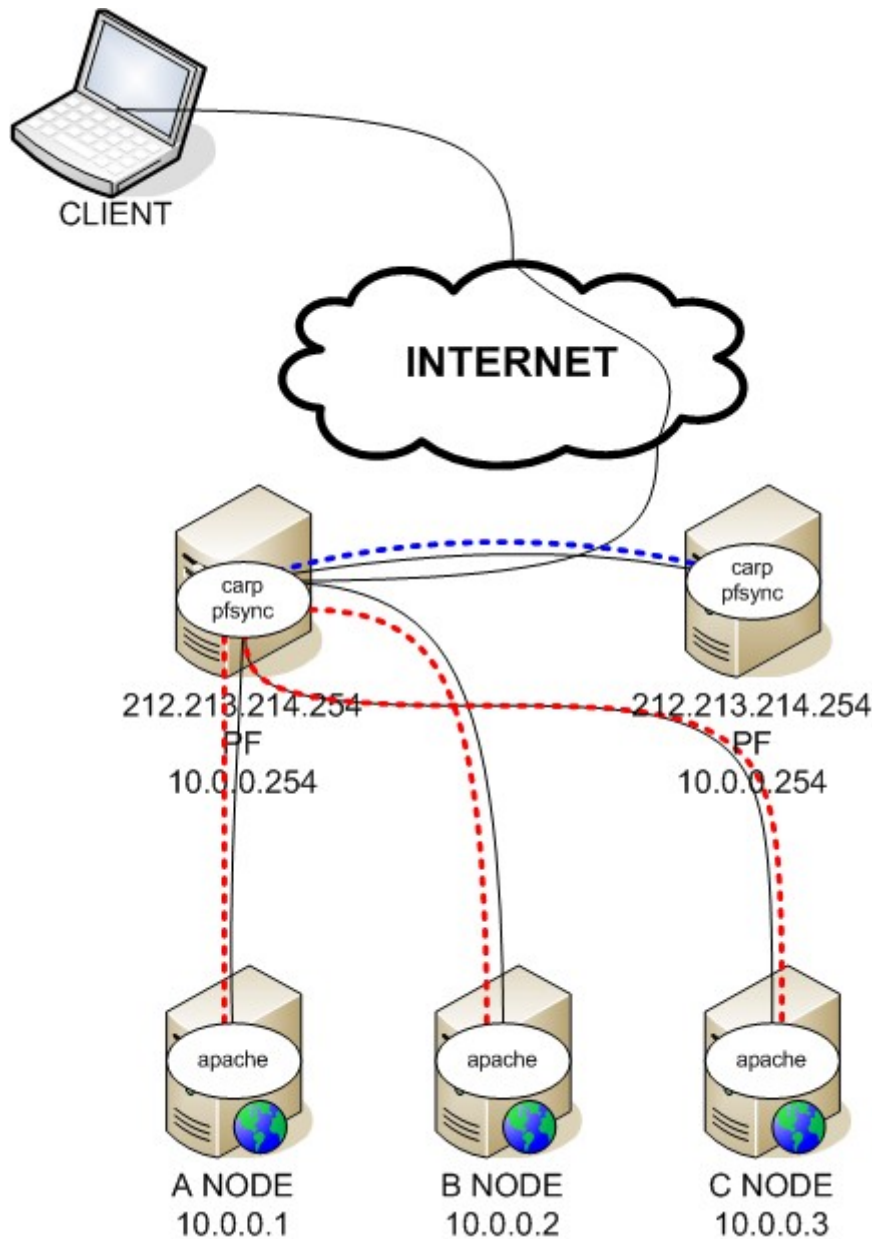
```
http_port 80
httpd_accel_host 127.0.0.1
httpd_accel_port 9090
```

```
httpd_accel_single_host on
httpd_accel_uses_host_header on
```

```
redirect_program /home/cl/redirect.sh
redirect_children 20
```

- Riscrive gli indirizzi (url rewrite)
- Distribuisce le richieste tra i nodi
- Monitora lo stato dei nodi

PF Address Pool



PF Address Pool

- PF è il firewall di OpenBSD
- Dispone di un meccanismo simile al round-robin chiamato IP Pool
- Simile al conn-track (“sticky connection”)

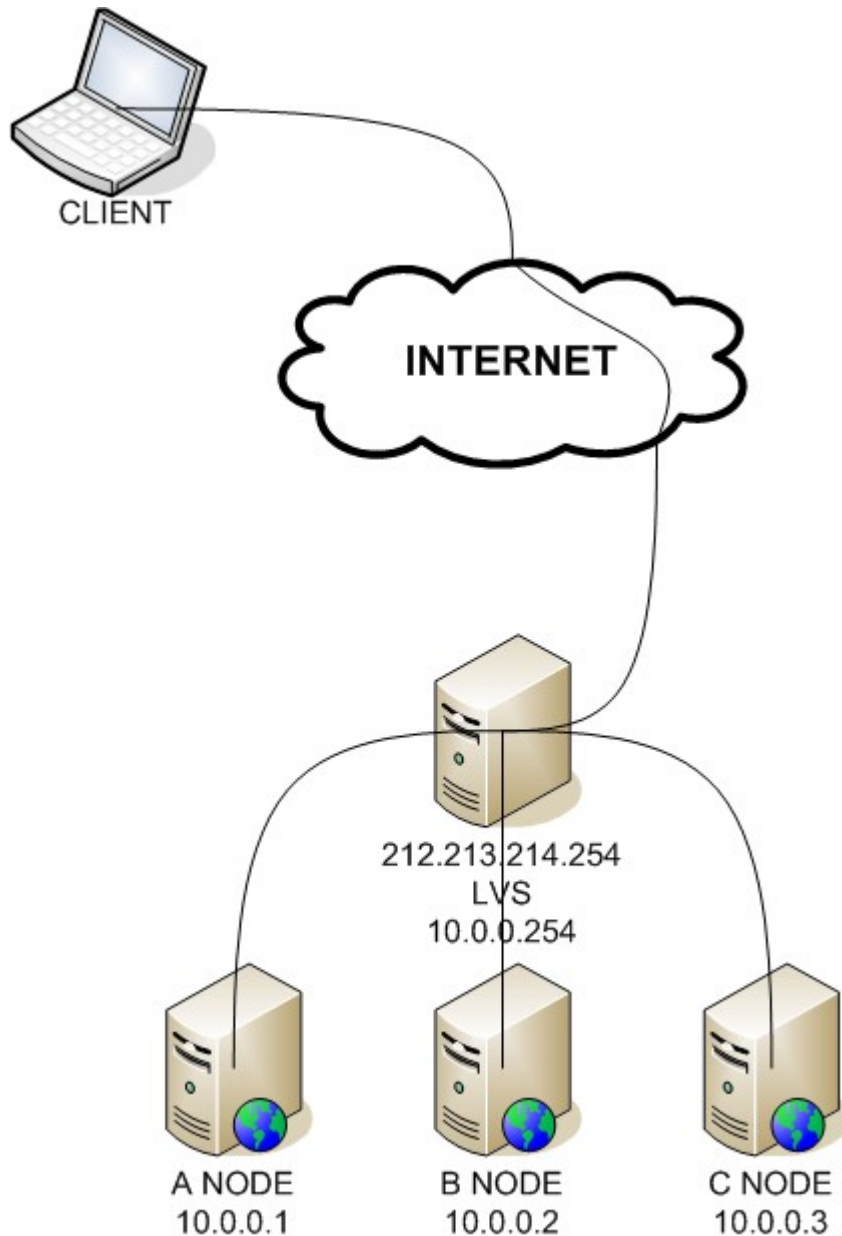
Esempio di IP Pool:

```
web_servers =  
"{ 10.0.0.1, 10.0.0.2, 10.0.0.3 }"
```

```
rdr on $ext_if proto tcp \  
    from any to any \  
    port 80 -> $web_servers \  
    round-robin sticky-address
```

- Tramite CARP e pfsync possiamo creare un backup redirector (fail-over)
- Distribuisce le richieste tra i nodi
- Non monitora lo stato dei nodi

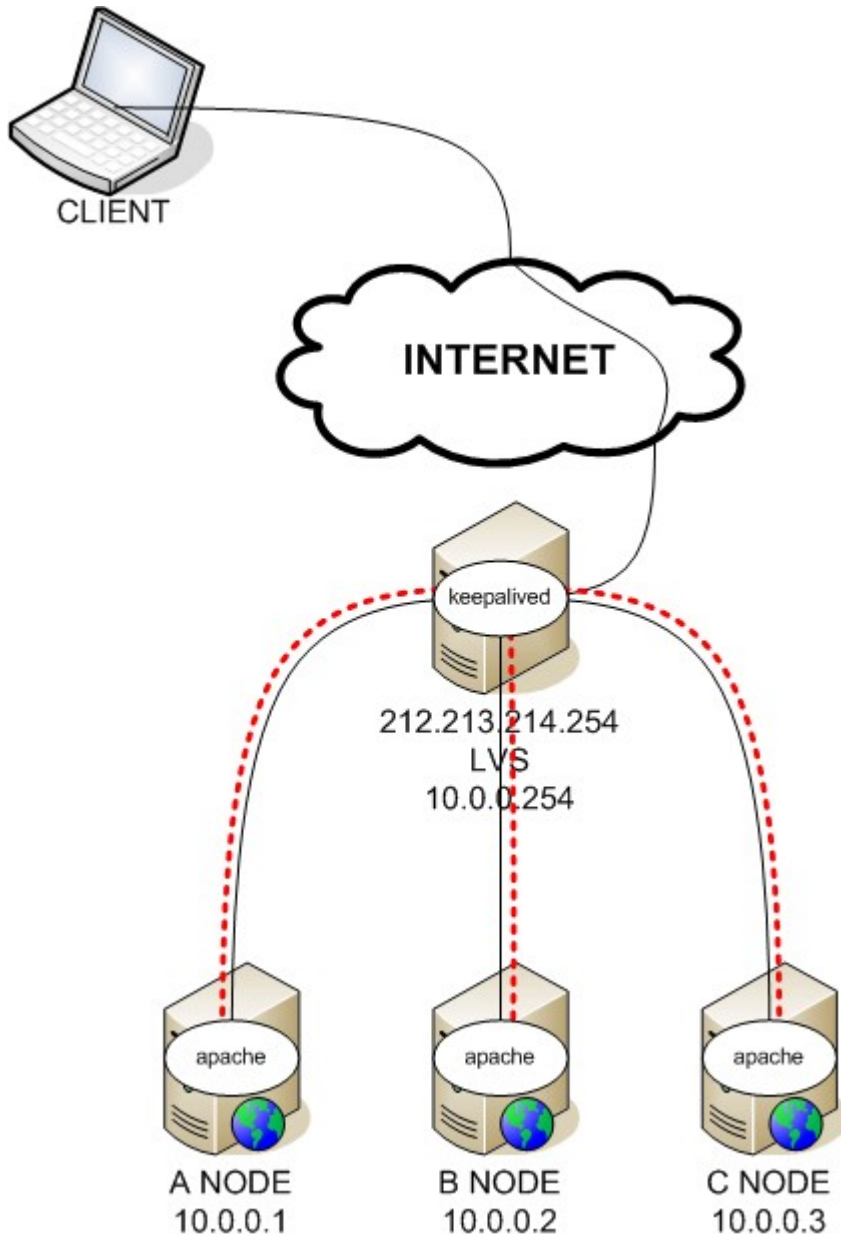
LVS Linux Virtual Server



Linux Virtual Server

- Il Front End Server (il Redirector) è il Load Balancer
- E' possibile in base al carico dei nodi determinare a chi indirizzare la richiesta
- Utilizza un modulo del kernel che implementa il load balancing a livello di trasporto all'interno del kernel (Layer-4 switching) chiamato IPVS
- Si possono usare diversi algoritmi di scheduling
- Non è fail-safe

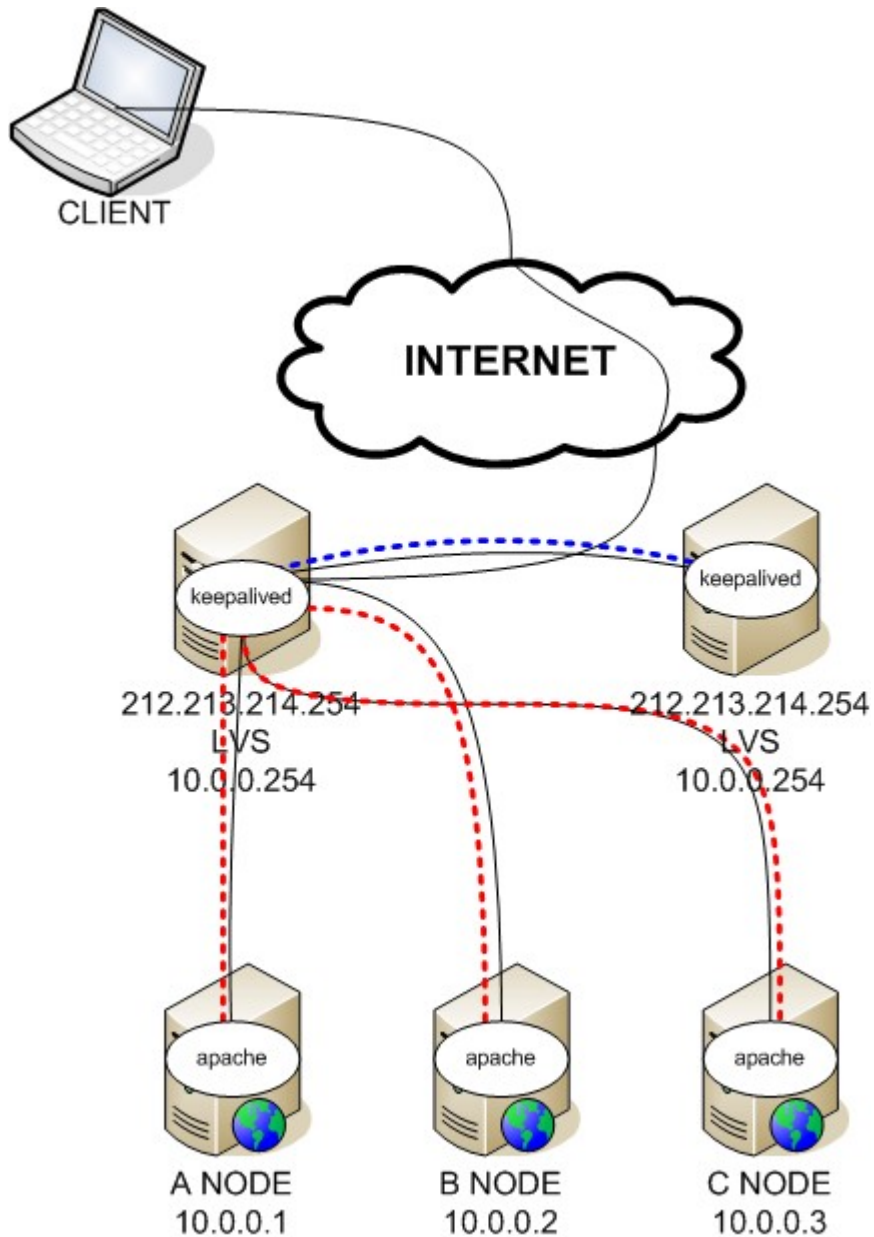
LVS + Keepalived



Keepalived + Linux Virtual Server

- Si innesta (hooking) in LVS
- Ha funzionalità di failover
- Controlla a livelli differenti: TCP, HTTP GET e script personalizzati

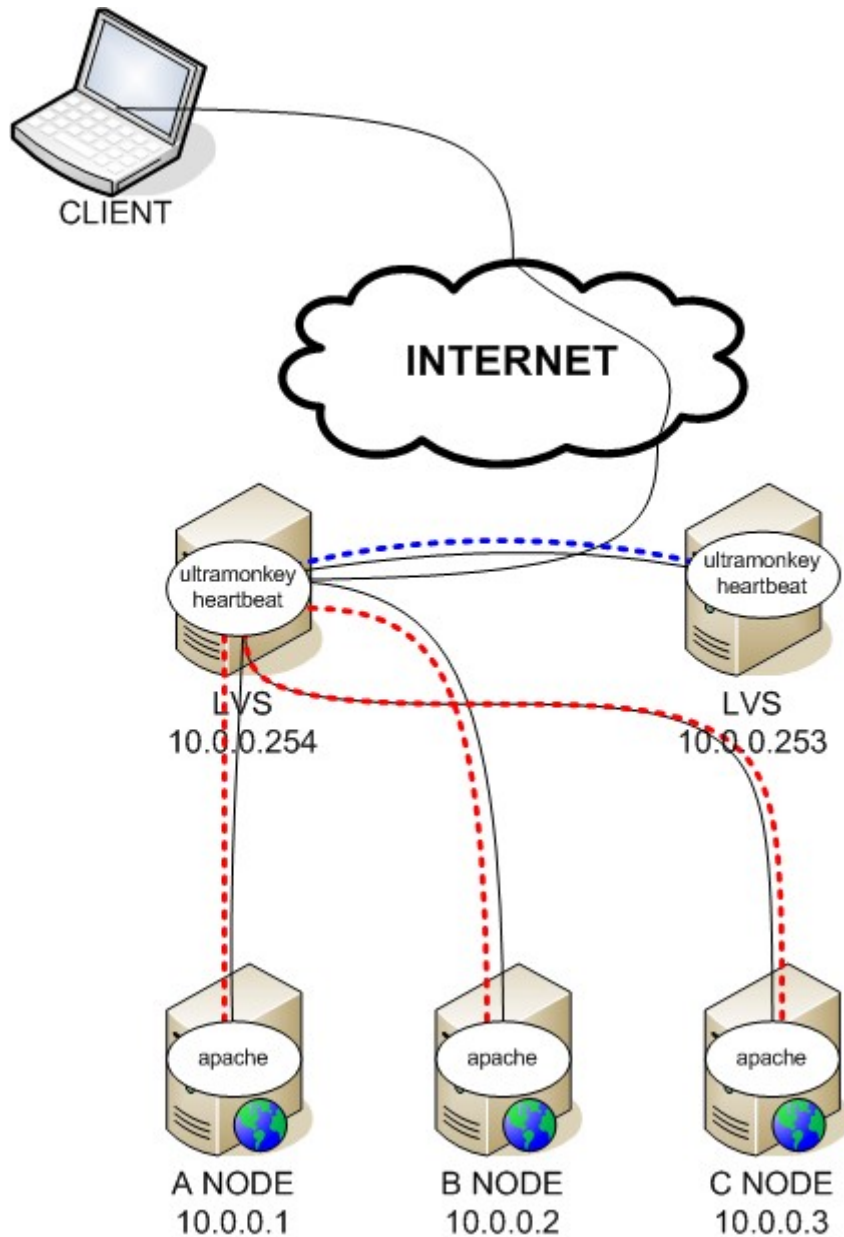
LVS Backup



Backup redirector

- Keepalived implementa il failover tramite VRPP
- Questo permette di avere un redirector di backup nel caso il primario non funzioni

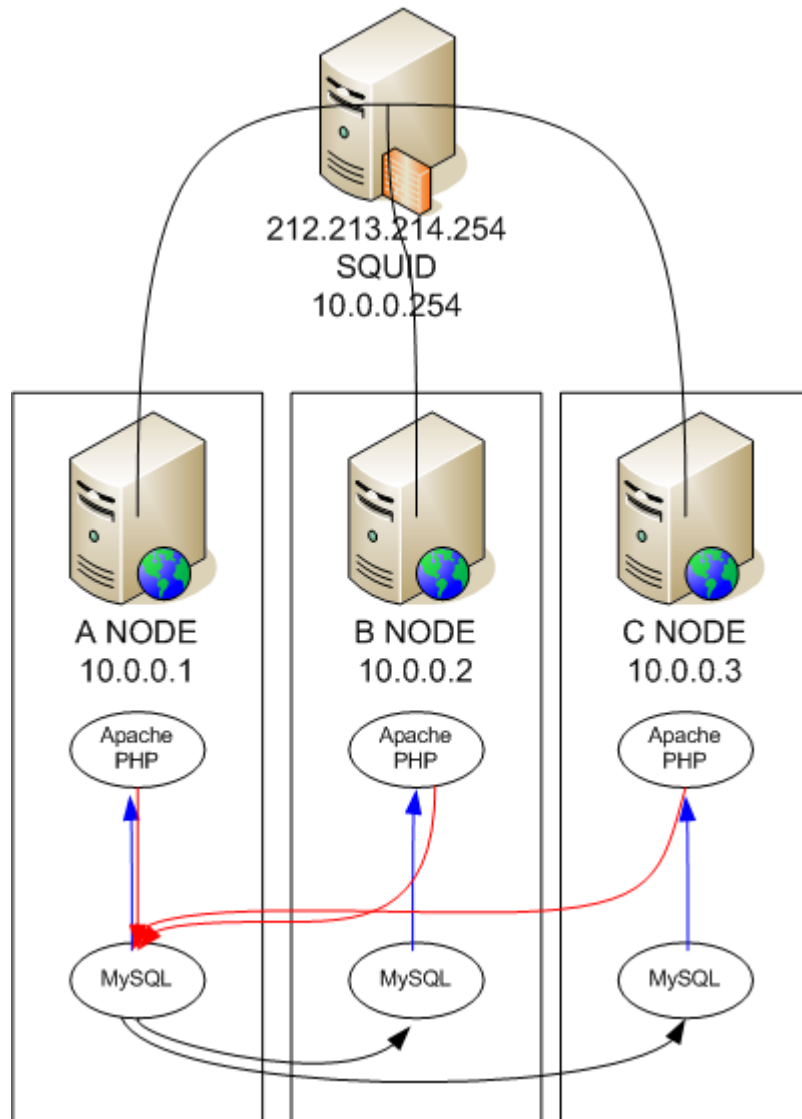
LVS + Ultramonkey



LVS + Ultramonkey

- Heartbeat monitora i bilanciatori e i nodi
- Ha funzionalità di fail-over
- Controlla a livelli differenti: TCP, HTTP GET e script personalizzati
- IP virtuale (mentre keepalived usa VRPP)

Database replication



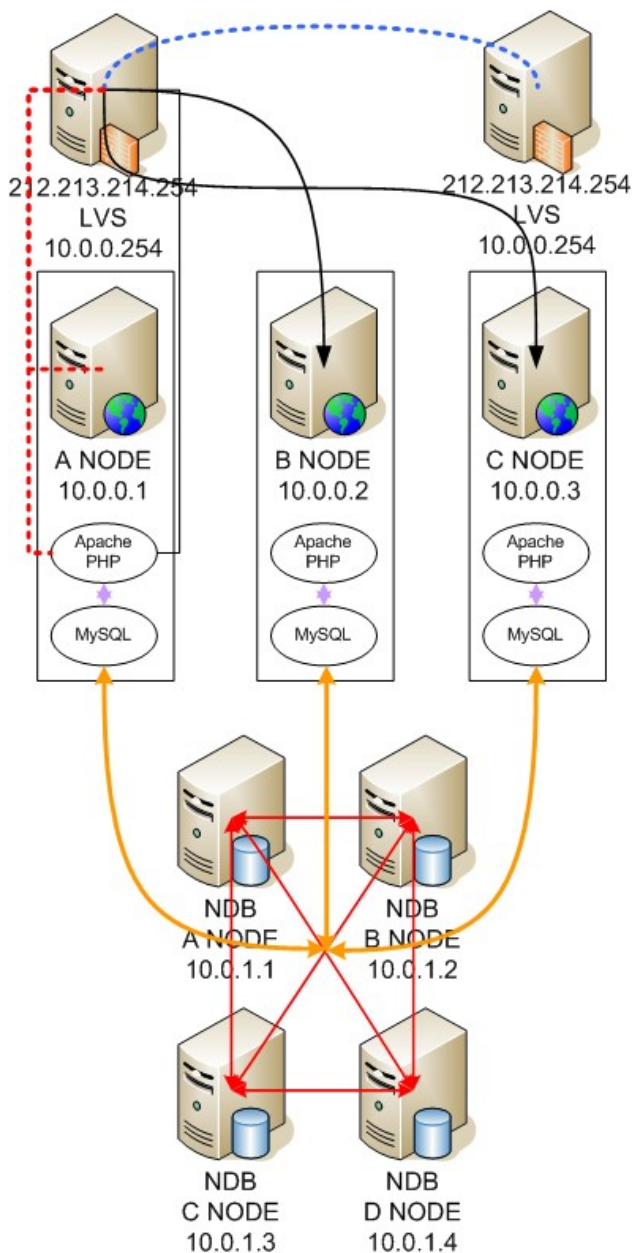
Database replication

- Master con multipli slave
- Tutte le WRITE vanno al master
- Le READ vanno agli slave

Problematiche:

- In realtà non è fail-safe
- Le transazioni non sono atomiche su tutti i nodi

MySQL Cluster



Il cluster NDB con MySQL 5 integrato nella rete HA

- Supporto NDB
- Repliche multiple per il fail-safe
- Le transazioni sono atomiche su tutto il cluster

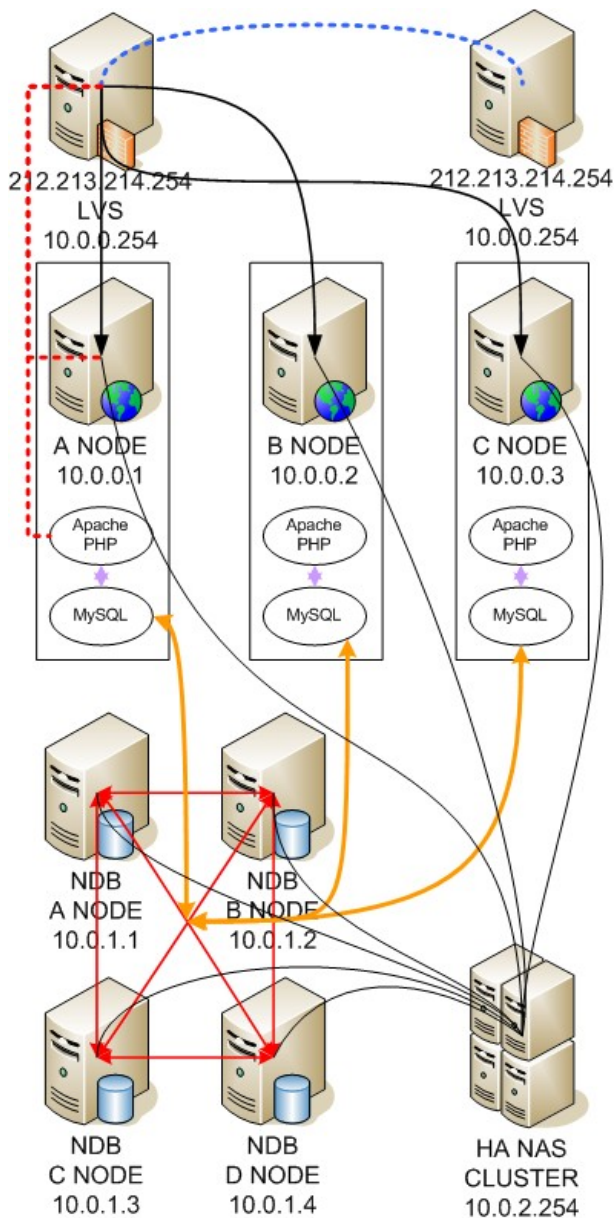
Più tipologie di nodi:

- Nodi di gestione (management)
- Engine SQL
- Nodi di memorizzazione (storage)

Problematiche:

- Tutto è nella memoria
- Ama usare molta memoria

Data Storage



Il NAS/SAN all'interno della rete HA

- Network File Systems: NFS, CODA
- Dedicated Hardware: NAS/SAN (Network Attached Storage)
- NAS: High Availability RAID
- SAN: Cluster

Domande?

Fine della storiella (-: Domande?

Risorse

Ilia Alshanetsky	http://ilia.ws/
Squid redirecting	http://www.squid-cache.org/Doc/FAQ/FAQ-15.html
Linux Virtual Server	http://www.linuxvirtualserver.org/
Keepalived	http://www.keepalived.org/
Ultramonkey	http://www.ultramonkey.org/
MySQL Cluster	http://dev.mysql.com/doc/refman/5.0/en/ndbcluster.html
PHP	http://www.php.net
SRM	http://www.vl-srm.net/

Contatti

Francesco 'ascii' Ongaro - ascii@ush.it - www.ush.it