Virtualizzazione con KVM su sistemi operativi Ubuntu

Francesco Stablum

Studente di Informatica, sviluppatore software, talvolta sistemista

Vantaggi della virtualizzazione

- Scalabilità
- Migrabilità
- OS sandbox
- QoS Isolation
- Consolidamento server

Vantaggi: scalabilità

- Possono essere in esecuzione sulla stessa macchina fisica un numero arbitrario di macchine virtuali con una variegata fauna di sistemi operativi differenti
- L'unico limite è la potenza di calcolo, la quantità di memoria disponibile e lo spazio su disco

Vantaggi: migrabilità

- Possiamo spostare una macchina virtuale da un sistema all'altro con estrema facilità.
- È sufficiente usare `scp`, invece di un furgone

Vantaggi: OS Sandbox

- Come in wikipedia, la Sandbox consente di fare pasticci devastanti senza ansia
  - p.es. installazione di un modulo per Apache sul web server
- Effettuare esperimenti sulle macchine di produzione dopo averle ”clonate”
- Filosofia Modifica/Annulla
- Testare virus
- HoneyPot/HoneyNet
- ...

Vantaggi: QoS Isolation

- Quality-Of-Service Isolation:
  - Evitare interferenze tra vari componenti software
  - Ad esempio, due sistemi virtuali con installate due istanze di Apache con configurazioni diverse (per esempio, una con mod_php e l'altra con mod_python).

Vantaggi: consolidamento server

- In alcuni contesti ci può essere una proliferazione di macchine fisiche sotto-utilizzate.
- Si virtualizzano i sistemi che necessitano di una macchina dedicata e si posizionano su un numero ridotto di macchine fisiche.
- Risparmio sull'hardware
- Risparmio energetico

Precisazioni

- Tutte le prove effettuate sono state fatte con Ubuntu 8.04

Struttura di un sistema di virtualizzazione

- Ipervisore
  - Chiamato anche VMM (Monitor di Macchine Virtuali)
- Macchine virtuali
  - Interfacce di rete virtuali
  - Dischi virtuali (sotto forma di un file memorizzato sulla macchina fisica)
  - Periferiche CD/DVD virtuali
    - E' possibile far leggere delle immagini ISO come se fossero dei CD/DVD reali sulla macchina virtuale
- Segmenti di rete virtuali
- Sistemi operativi sulle macchine virtuali

Accenni alla teoria

- Requisiti introdotti da Popek e Goldberg nel 1974
- Tre proprietà:
  - Equivalenza: un sistema che gira su una macchina virtuale deve mostrare un comportamento identitico a quello che mostrerebbe su una macchina fisica.
  - Controllo delle risorse: le risorse virtualizzate devono essere controllate completamente dall'ipervisore.
  - Efficienza: una porzione significativa del lavoro effettuato da una macchina virtuale deve essere effettuato senza l'intervento dell'ipervisore.

KVM

- Acronimo per Kernel-based Virtual Machine
- Utilizza funzionalità di virtualizzazione assistita dall'hardware.
- E' un modulo del kernel: `kvm.ko` e un modulo specifico per l'architettura della macchina ospitante: `kvm-intel.ko` e `kvm-amd.ko`
- Utilizza una versione modificata di QEMU
- Giovane ma promettente
  - Comunità di sviluppo molto attiva

Immagini disco per KVM

- La nostra macchina KVM ha bisogno di un disco (virtuale ovviamente).
- Il disco si può creare facilmente come file tramite il comando `qemu-img`:
- `qemu-img create dsl1.img -fqcow2 512M`

Formato qcow2 per immagini disco KVM

- qcow2: funzionalità salienti:
  - ottimizzazione: solo le parti di disco scritte vengono effettivamente allocate.
  - snapshot
  - crittografia AES
  - compressione dati
  - può essere convertito in formato raw
    e successivamente montato nel file system

Interazione con KVM: VNC

- VNC: e' possibile collegarsi via rete accedendo all'I/O della macchina virtuale
- `-vnc :0`
- Kvm ora diventa un server VNC e ci si può collegare semplicemente con il comando:
  - `vncviewer localhost:0`

Interazione con KVM: il Monitor

- E' una console che permette di inviare comandi a KVM
- Parametri di configurazione all'avvio di KVM
  - `-monitor stdio`
  - `-monitor tcp::8888`

Interazione con KVM: il Monitor

- comandi di controllo:
  - `stop` : arresto dell'emulazione
  - `system_reset` : reset della macchina
  - `sendkeys` : invio di input da tastiera
  - `savevm` : salvataggio di uno snapshot del sistema
  - `wavcapture` : salvataggio dello stream audio
  - Debugging
  - ...

TUN/TAP

- Costruzione di interfacce di rete virtuali
- TAP lavora a layer 2 (Ethernet)
- TUN lavora a layer 3 (IP)
- Scopo: permettere a un programma in user-space di ”attaccarsi” all'interfaccia virtuale come se fosse un terminale in LAN.
  - Proprio quello che ci serve per KVM!!

TUN/TAP: Mini-Tutorial (1)

- Installiamo il pacchetto contentente i tool necessari
  - `apt-get install uml-utilities`
- Creiamo il bridge virtuale
  - `brctl addbr br0`
- Creiamo l'interfaccia di rete virtuale
  - `tunctl -t tap0`
- Colleghiamo l'interfaccia virtuale al bridge
  - `brctl addif br0 tap0`

TUN/TAP: Mini-Tutorial (2)

- Lanciamo kvm con distro DamnSmallLinux:
- ```
  kvm
  -hda ./dsl2_qcow2.img
  -cdrom ./dsl-4.2.5-initrd.iso
  -net nic,macaddr="12:34:56:78:9a:bc"
  -net tap,ifname=tap0,script=no
  ```
- Il comando è rappresentato su più linee ma è da considerarsi su una sola
- I due comandi `-net` sono utilizzati per collegare l'interfaccia di rete del sistema virtualizzato all'interfaccia tap
  ```

TUN/TAP: Mini-Tutorial (3)

- Nella macchina ospitante bisogna impostare correttamente il routing verso le macchine virtuali:
  - `route add -net 10.77.77.0/24 dev br0`
- Attiviamo anche il forwarding:
  - `echo 1 > /proc/sys/net/ipv4/ip_forward`

TUN/TAP: Mini-Tutorial (4)

- Impostiamo l'indirizzo IP del bridge
  - `ifconfig br0 10.77.77.1 up`
- Nella macchina virtualizzata, l'interfaccia di rete è disponibile come eth0
- Impostiamo l'indirizzo IP di eth0 all'interno della macchina virtualizzata e il routing
  - `ifconfig eth0 10.77.77.2 up`
  - `route add default gw 10.77.77.1`

HeartBeat

- HeartBeat è un componente di un progetto più ampio, Linux-HA (High Availability)
- È in grado di:
  - stabilire se un nodo ha cessato di erogare un servizio (ad esempio, è esploso) tramite continuo pinging
  - effettuare il failover: un'altra macchina prende il controllo per mantenere una disponibilità continua

HeartBeat: un caso pratico

- Configurazione Attivo/Passivo
  - Cluster con due nodi:
    - Primary (192.168.1.2): Attivo
    - Secondary (192.168.1.3): Passivo
  - Il cluster ha un'indirizzo IP (192.168.1.1) che deve essere sempre disponibile anche in caso di fallimento di Primary
  - Pacchetti HeartBeat mandati in broadcast attraverso un interfaccia e un segmento di rete dedicato.

HeartBeat: un caso pratico

- Quando Primary non reagisce ai ping:
  - Secondary assume, oltre al proprio, anche l'indirizzo IP ”logico” 192.168.1.1 utilizzato per erogare i servizi all'esterno.

DRBD

- È un sistema di ridondanza in grado di creare una partizione in mirroring
- Utilizza un sistema di ”proiezione” delle scritture di un nodo ”master” su un nodo ”slave”
- Necessità di una canale di comunicazione ad alta velocità

DRBD

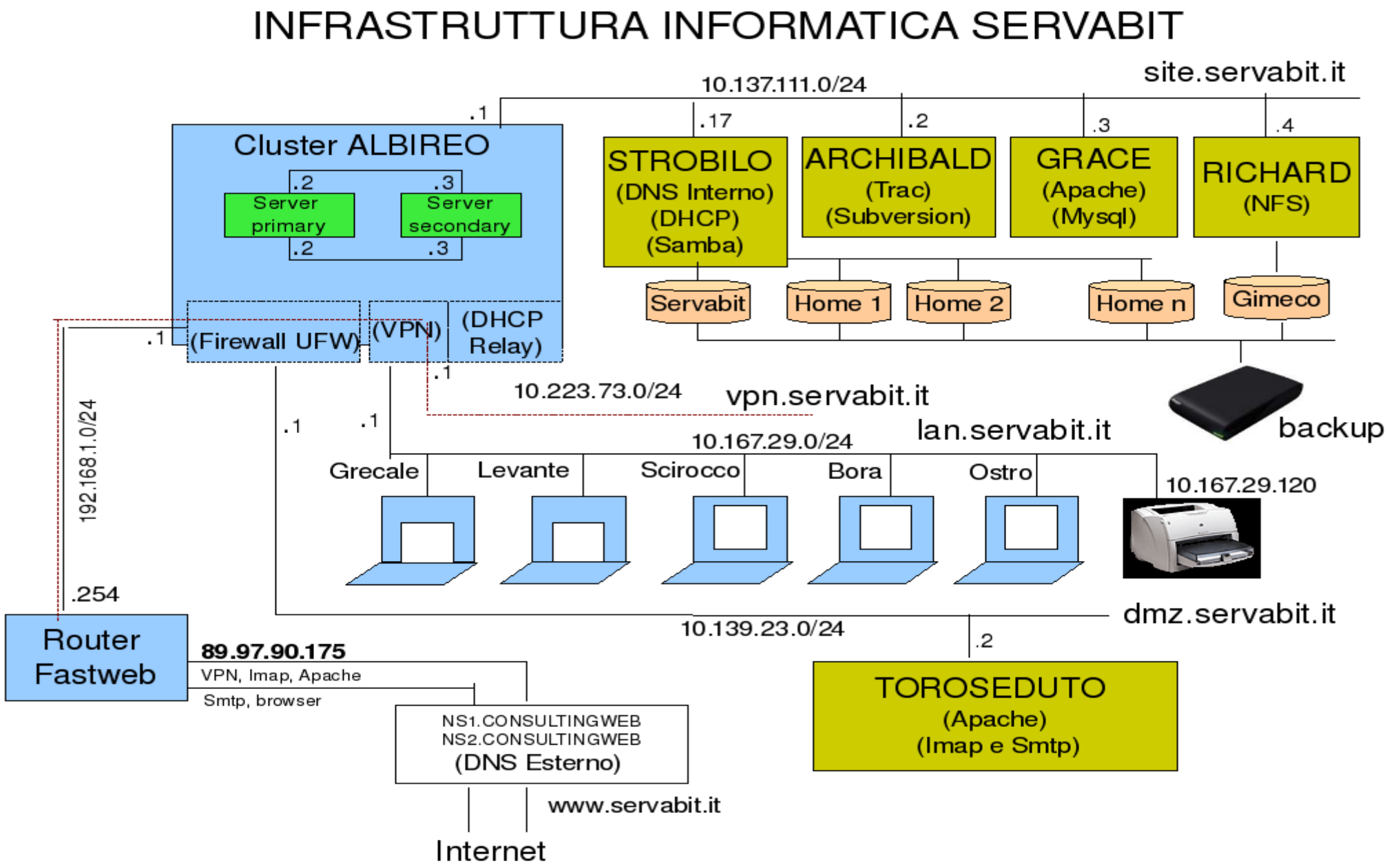
- Il sistema di invio delle scritture può essere:
  - Sincrono
    - LAN
  - Asincrono
    - Grandi distanze, grande ritardo

DRBD: Come lo usiamo?

- Nella partizione in mirroring ci mettiamo tutte le immagini dei sistemi operativi in esecuzione.
- Se Primary fallisce:
  - HeartBeat su Secondary rileva il problema
  - Secondary riavvia tutte le macchine virtuali
  - Le macchine virtuali al momento del riavvio si trovano in uno stato inconsistente, ma i sistemi virtualizzati se ne rendono conto e fanno partire `fsck`

Ma funziona?

Ma funziona? Certamente!



Strutturazione della rete

- *.dmz.servabit.it
  - Adibito alla zona de-militarizzata
  - Virtuale: macchine kvm su tap*
- *.site.servabit.it
  - Adibito ai servizi accessibili dall'interno
  - Virtuale: macchine kvm su tap*
- *.lan.servabit.it
  - Segmento di rete fisico.
- *.vpn.servabit.it
  - Adibito alla Virtual Private Network

Funzionalità offerte

- Funzionalità offerte da *.site.servabit.it
  - DHCP
  - DNS interno
  - Subversion
  - Samba
  - NFS
  - Trac
  - Macchine di sviluppo applicativi Web (Apache)
    - mod_python
    - mod_php
  - MySQL
  - PostgreSQL
  - OpenERP

Funzionalità offerte

- Funzionalità offerte da *.dmz.servabit.it
  - Imap con SSL
  - SMTP
  - Sito web accessibile da www.servabit.it (Apache)
- Risulta difficile da credere, ma tutti i servizi erogati da *.dmz e da *.site sono effettivamente in esecuzione su UNA SOLA macchina fisica, grazie a KVM
- Ogni macchina virtuale è un sistema Ubuntu Server con un proprio indirizzo IP

Ringraziamenti

L'azienda per la quale lavoro
per la straordinaria esperienza formativa

Tutti i colleghi per la pazienza
in particolar modo Davide Bolcioni che mi ha guidato
con la sua esperienza e i suoi preziosi consigli

Approfondimenti e Linkografia

- http://bellard.org/qemu/user-doc.html
- http://kvm.qumranet.com
- http://calamari.reverse-dns.net:980 /cgi-bin/moin.cgi/QemuAndTuntap
- http://www.linux-ha.org/Heartbeat