Sicurezza dei sistemi

Vincenzo Calabrò

Roma, 1/12/2008

Il supporto ai sistemi

Guardando da molto in alto, una infrastruttura tecnologica per la gestione dell'informazione, piccola o grande che sia, si può pensare come un insieme di sistemi amministrato autonomamente e collegato in vari modi ad altre strutture o realtà

E' un Sistema di sistemi

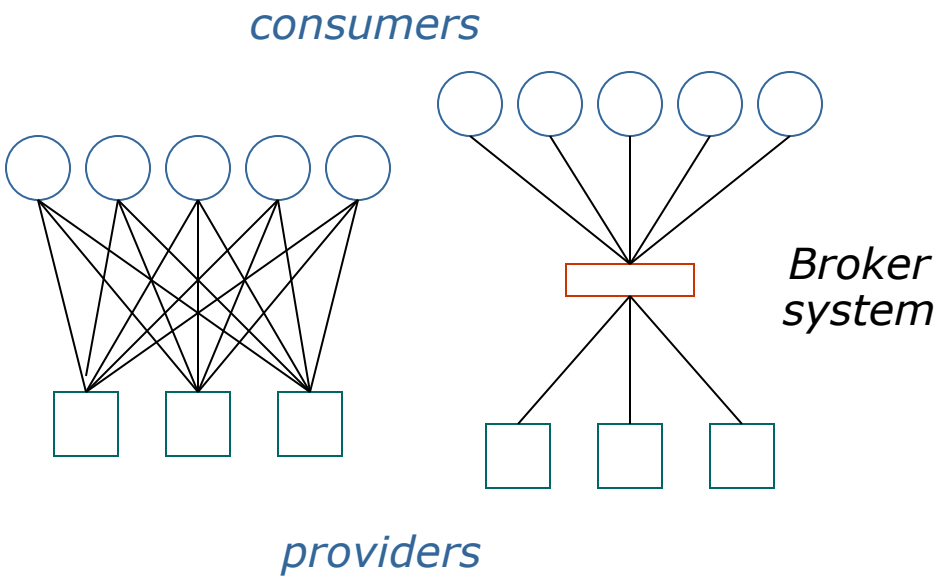
Garantire la sicurezza dell'informazione in esso contenuta e da esso gestita e trasmessa implica l'esercizio di una visione globale e unitaria, e l'effettuazione di uno sforzo di sintesi e coordinamento di molte pratiche specifiche dei sistemi componenti (dati, applicazioni, sistemi operativi, rete, ambiente, persone, skill, …), al fine di gestire e dominare la complessità d'insieme

Brokering

Il sistema di sistemi per vari motivi è in continua evoluzione e quindi cresce di complessità continuamente.

Per gestire nel tempo una complessità crescente controllando che i costi di questa gestione non esplodano occorro introdurre il *brokering*, che disaccoppia i fornitori delle soluzioni dai fruitori delle stesse

- monitoring
- configuration
- support
- …



Brokering

Il brokering si applica a tutti i livelli della gestione dei sistemi, dal più basso e tecnologico fino a quelli più alti applicativi e organizzativi:

NETWORK MGMT

SNMP, NetView, OpenView, …

SYSTEM & CONFIGURATION MGMT

Tivoli, LanDesk, SMS

APPLICATION MGMT

Candle

SUPPORT MGMT

Remote control, trouble ticketing

USER SUPPORT

Systems Help Desk, Application Help Desk

Baseline

In questo scenario si inserisce uno degli approcci fondamentali all'ottenimento della sicurezza e della disponibilità dei sistemi, ossia il raggiungimento di un punto di equilibrio, un livello di base di sicurezza a partire dal quale sia più semplice riconoscere proattivamente le anomalie, per poi evitarle o rimuoverle con specifiche azioni correttive

- policy di sicurezza e configurazione
- sistemi di monitoraggio e automazione
- gestione delle patch e dei cambiamento
- gestione delle anomalie e degli incidenti
- verifiche periodiche

Logging

La raccolta delle informazioni è il primo anello della catena di *brokering*

- Agenti SNMP ecc.
- Agenti per il controllo del sistema
- Agenti per il controllo di applicazioni
- Dirottamento e concentrazione dei LOG

Queste informazioni vengono inviate a server centrali di analisi e monitoraggio

Monitoring

Uno o più sistemi centrali ricevono, elaborano e organizzano le informazioni raccolte dagli agenti installati nei sistemi

- SNMP traps
- Alerts
- Actions

I sistemi centrali tengono una storia degli eventi e permettono di avere l'evidenza dei problemi anche successivamente

Big Brother, Hobbit Monitor

Big Brother

Abbastanza diffuso e molto efficace, perché semplice ed estensibile; ottimo strumento per sostenere una linea di base di servizio e di sicurezza. E' costituito da un server centrale e da servizi opzionali installati sui sistemi da monitorare

- BBNetwork, BBClient, BBDisplay

Hobbit Monitor

Evoluzione di BB: maggiori performances (C anziché shell script), configurazioni gestite centralmente, più modulare (permette la suddivisione gerarchica delle informazioni in pagine separate) ed estensibile

Nagios

Nagios è un “host and service monitor” con le seguenti funzionalità (dalla homepage):

Monitoring of network services (SMTP, POP3, HTTP, NNTP, PING, etc.)

Monitoring of host resources (CPU load, disk and mem usage, processes, logs, etc.)

Monitoring of environmental factors such as temperature

Simple plugin design (allows to easily develop host and service checks)

Ability to define network host hierarchy (avoiding multiple alerts)

Contact notifications (via email, pager, or other user-defined method)

Support for implementing redundant and distributed monitoring servers

Scheduled downtime management (suppressing notifications during planned outages)

Ability to acknowledge problems via the web interface

Web interface for viewing current network status, notification and problem history, log file, etc.

Simple authorization scheme that allows you restrict what users can see and do from the web interface

Hyperic HQ

Hyperic HQ è un “systems and application monitor”.

Conosce approfonditamente sistemi operativi, piattaforme applicative, application server, mail server, RDBMS, middleware di messaging, tecnologie di Microsoft, tecnologie di virtualizzazione e si interfaccia con altri sistemi di monitoraggio.

Architettura client server; è nato originariamente come systems monitor, sviluppato in ambiente JEE, l’agent (client) conosce e capisce i dettagli dei sistemi in cui si installa e ne permette di acquisirne la configurazione di base automaticamente

Se nel sistema monitorato è installato un RDBMS o un application server conosciuto questo viene “esplorato” e la sua configurazione diventa disponibile per il monitoraggio; per esempio, riconosce il deployment delle applicazioni JEE e le EAR diventano risorse gestite

Rende disponibili all’amministratore una quantità di metriche e un cockpit centralizzato con cui effettuare

- *application status monitoring and performance management*
- *cross platform SLA monitoring*
- *application resource visibility across virtual and physical environments*

Monitoring goal

Oltre alla funzione fondamentale di monitoraggio, l'efficacia di un monitor ben configurato consiste nell'ottenimento di uno stato di equilibrio nel controllo del Sistema di sistemi, ossia nella conoscenza della baseline dell'intera struttura informatica sotto controllo.

Man mano che si risolvono i problemi evidenziati dal monitor, ci si riavvicina al punto di equilibrio e si riduce l'instabilità dell'intero Sistema

La configurazione del sistema di monitor stesso (server centrali e agenti sui sistemi da monitorare) deve comunque essere sempre aggiornata e consistente per consentire al sistema di monitoraggio stesso di governare l'intera struttura informatica

Remote control

Quando il fornitore del supporto interviene a fronte di problemi, ha bisogno di farlo nel modo più efficiente possibile, cioè facendo uso di mezzi per la gestione remota dei sistemi e delle applicazioni in essi contenute.

- Raggiungibilità

Normalmente i sistemi stessi forniscono questi mezzi: per esempio i Terminal Services dei sistemi operativi server Microsoft hanno una modalità di funzionamento a scopo di amministrazione che consente due connessioni “Remote Desktop” al server, tramite il protocollo RDP.

Remote control

Un sistema spento, malfunzionante o in fase di boot, quando la rete non è disponibile o è in una sede remota possono essere necessarie altre soluzioni, eventualmente complementari

- Datacenter & remote branches
- KVM (Keyboard, Video & Mouse) management
- Serial, console & power management

Le soluzioni utilizzabili nei datacenter locali sono più efficienti ma più semplici di quelle utili per gestire sedi remote, dove servono diversi algoritmi di compressione, uso di canali di comunicazione alternativi alla rete, gestione dei dispositivi periferici, gestione della potenza elettrica, ecc

Remote control

I dispositivi più diffusi per gestire un'infrastruttura locale (datacenter) sono gli Switch KVM over IP

Sono dispositivi interfacciati con le console fisiche dei server tramite connettori VGA o DVI e connettori mini-DIN per mouse e tastiera (i dispositivi recenti utilizzano cablaggio CAT 5 con breakout boxes per minimizzare l'ingombro del cablaggio KVM)

Questi dispositivi sono interfacciati con un server centrale di gestione (per il dispatching delle connessioni e le autorizzazioni); i gestori usano una console software (anche web based) che li collega ai dispositivi stessi e al server prescelto tramite un protocollo di rete proprietario ed efficiente.

Configuration management

Si parla di configurazione di un sistema riferendosi a tutte le attività di installazione e manutenzione hardware e software dello stesso

La sistematica identificazione, proceduralizzazione e documentazione di queste attività si chiama configuration management e consente di far evolvere nel tempo i sistemi nel loro insieme, in modo organico e coerente

Dotarsi di buone procedure di configuration management dei sistemi è un obiettivo necessario per governare non solo la funzionalità degli stessi, ma anche la loro sicurezza e la loro disponibilità

Software Distribution

Tra gli oggetti che implicano configuration management c'è la distribuzione del software.

I sistemi di configuration management solitamente hanno un componente che consente la preparazione di pacchetti software e la loro successiva distribuzione sui sistemi controllati a seconda delle esigenze dell'utenza.

- Imaging
  Preparazione dei pacchetti mediante installazione su un sistema di riferimento
- Provisioning
  Distriuzione delle immagini secondo varie politiche e con vari meccanismi

Software Distribution

La maggioranza delle implementazioni di sistemi di software distribution funzionano in modalità client-server, ossia l'agent a bordo dei computer gestiti viene istruito sulla disponibilità di un pacchetto e comandato di scaricarlo e installarlo, ritornando al server il suo nuovo stato. Questa architettura è assolutamente adeguata all'uso in rete locale.

In una rete distribuita geograficamente, soprattutto nel caso di sedi piccole e dotate di connettività limitata, avere un unico punto di distribuzione centrale è una grossa limitazione

Software Distribution

Recentemente si è sempre più affermata un'architettura a due livelli, un primo client-server e un secondo peer-to-peer, funzionante in streaming o anche come le note reti di file sharing. Questa architettura più moderna è in generale più efficiente e anche più adatta nel caso di reti distribuite

Queste soluzioni prevedono che, oltre ai server di distribuzione centrali, anche i singoli nodi della rete possano contribuire alla diffusione dei pacchetti software o degli altri dati oggetti di distribuzione (ad es. gli aggiornamenti del prodotto stesso)

Software Distribution

Lo sviluppo e della diffusione delle tecnologie di virtualizzazione ha investito anche la software distribution, semplificando il packaging in forma di *sandbox* che vengono distribuiti in streaming sui sistemi finali e consentono alle applicazioni distribuite di eseguire in un ambiente isolato dal sistema destinatario ed eliminando le interazioni con le altre applicazioni oggetto della distribuzione

In questo caso, sfruttando le potenzialità dell'application virtualization, i package delle applicazioni sono molto più semplici da realizzare e gli oggetti da distribuire sono i *sandbox* stessi

Patch management

E' un altro sottoproblema del configuration management di particolare impatto sulle metodologie per la sicurezza dei sistemi

To patch or not to patch?

Questa domanda ha una risposta risolvendo il problema seguente: *il rischio di applicare al sistema la patch è superiore al rischio della vulnerabilità che la patch corregge?*

Non è un calcolo facile e comunque non va effettuato caso per caso e al di fuori di una metodologia chiara e ben pianificata.

Patch management

Una corretta metodologia di *patch management* può essere espressa in varie fasi

1. Baseline definition
2. Test Environments
3. Backout Plans
4. Patch collection and evaluation
5. Consolidation
6. Deployment
7. Reporting

Patch management

1) Occorre gestire un inventario dettagliato dei sistemi (nome, location, indirizzo, o.s. & software type & levels ecc). La security baseline è ottenuta con installazioni controllate o gestite automaticamente derivanti da prototipi sottoposti a vari test di funzionamento e sicurezza (es. hardening, scan, probing ecc)

2) Occorre installare un ambiente per il test, che riproduce i sistemi in produzione e su cui bisogna effettuare una simulazione del patching. Del test environment fanno parte anche gli eventuali utenti pilota che provano le patch

3) Prima di applicare le patch occorre effettuare il backup dei dati e delle applicazioni, ove possibile creandosi copie-immagine dei sistemi per realizzare il “bare-metal restore” o comunque una rapida ricostruzione del sistema in caso di necessità

Patch management

4) Il sistema usato per la gestione delle configurazioni deve essere in grado di classificare i server e permettere anche un'adeguata classificazione delle patch disponibili, per una successiva approvazione, test e applicazione

5) Dopo che le patch sono state approvate, poi applicate e testate sul test environment, deve essere documentato e divulgato il risultato (gli utenti informati, l'help-desk allertato, ecc) prima del deployment sui sistemi di produzione

6) Le patch vengono applicate ai sistemi interessati, quando stabilito e dichiarato nell’apposita comunicazione di servizio

7) Si produce infine la documentazione, corredata di eventuali report per il management e per la consultazione a posteriori delle attività effettuate

Upgrade management

Apparentemente analogo al patch management, tuttavia squisitamente diverso in quanto è un problema di insieme anziché del singolo sistema

To upgrade or not to upgrade?

Questa domanda ha invece come risposta solo un'altra domanda: When?

Con l'inarrestabile evoluzione dell'Information Technology non è semplicemente pensabile mantenere troppo disallineati i diversi sistemi nel tempo, altrimenti in pochi anni smettono di funzionare, non sono più supportati e si corre il rischio ben più grave di doverli dismettere e sostituire rapidamente, cosa non sempre possibile

Upgrade management

I sistemi (server, sistemi operativi e applicazioni) molto vecchi, anche se apparentemente rispondono ancora ai requisiti per cui sono stati installati, nel tempo soffrono nell'essere ospitati all'interno di un ambiente in continua evoluzione: tutti gli altri sistemi nel frattempo evolvono, il carico aumenta, i protocolli di comunicazione cambiano, compaiono in rete nuovi pacchetti, cambiano i dispositivi di rete cui è collegato, cambiano le applicazioni con cui è stato integrato ecc.

In pratica, anche se un sistema "funziona" in realtà il suo utilizzo reale si allontana sempre più dall'ambito originale in cui è stato usato e testato, per cui il rischio di suoi malfunzionamenti aumenta continuamente

Systems management

Gli strumenti di systems management hanno l'ambizione di offrire una soluzione integrata a tutti questi aspetti della gestione dei sistemi

Sono sistemi client-server modulari, dotati di una console di gestione che consente di governare le informazioni che i moduli di gestione (agenti) installati su tutti i sistemi (server e PC) mandano e ricevono dal server centrale di system management, dove risiede il database di inventario hardware e software, le immagini e gli aggiornamenti dei sistemi operativi e del software da installare, le regole di gestione.

La console offre poi funzionalità di monitoraggio, controllo remoto, assistenza all'help desk, ecc.

Hardening

E' una tecnica di configuration management dei sistemi, che permette di analizzare e affinare la configurazione degli stessi con l'obiettivo di conseguire la massima sicurezza di base per uno specifico sistema

Ciascun tipo di server richiede tecniche di hardening proprie, che ovviamente variano anche in funzione della sua visibilità e dell'informazione in esso contenuta. L'utilizzo di tool specifici permette di ottenere con ripetibilità l'applicazione sistematica delle policy di sicurezza scelte

Questi tipo di hardening si può definire statico, in quanto è indipendente dall'attività presente sul sistema e dai processi in esso utilizzati

Hardening

Si può parlare anche di hardening dinamico, nel caso si realizzino meccanismi di protezione che funzionano a ogni specifica azione effettuata sui dati dai processi ospitati dal sistema operativo

Alcune estensioni dei sistemi operativi permettono infatti di realizzare politiche di sicurezza molto più mirate ai dati e alle attività di accesso agli stessi

Hardening Windows

- Ci si affida alla Casa Madre
- Esempio di hardening per un web server
  - Pathes & updates (MBSA, latest patches, security notifications…)
  - IISLockdown + URLScan
  - Services (unnecessary, least-privilege, …)
  - Protocols (WebDAV, TCP/IP, NetBIOS, SMB)
  - Accounts (unused, guest, admin, IUSR_machine, …)
  - Files & Directories (NTFS, restrictions, SDK & samples, …)
  - Shares, Ports, Registry, Auditing, Logging, Scripts, Sites, Virtual Directories, ISAPI Filters, IIS Metabase, …

Hardening Unix

Strumenti come Bastille Linux permettono all'amministratore di automatizzare vari cambiamenti nella configurazione di un sistema (Linux, HP/UX, Mac)

- Patch checking
- File permissions
- Account and Boot security
- Inetd e altri daemon (servizi inutili, ecc.)
- User tools (compilatori, ecc.)
- Logging, printing
- Sendmail, Apache, DNS, FTP
- Directory /tmp
- IPTables, …

Hardening

Anche senza effettuare azioni esplicite di hardening, è comunque opportuno implementare il patching del fornitore. Le patch fornite includono sia fix per problemi funzionali dei prodotti, sia fix per problemi di sicurezza dell'intero sistema operativo

SELinux

Security-Enhanced Linux, disponibile anche in distribuzioni commerciali come RHEL, è una estensione del kernel di Linux che realizza un framework di sicurezza basato sull'implementazione a di meccanismi di autorizzazioni di tipo MAC (Mandatory Access Control).

MAC

Il sistema viene rappresentato in termini di *subjects* (processi) e *objects* (devices, files, sockets, …). SELinux consente di definire esplicite *policy* su tutti gli accessi, ossia gli usi che i subjects fanno degli objects del sistema.

DAC

Il meccanismo standard, tipico degli Unix, di Windows, dei file servers ecc, si chiama DAC (Discretionary Access Control).

DAC

In regime di **Discretionary Access Control**, ogni utente è considerato proprietario di alcuni oggetti, cui può assegnare diritti di accesso ad altri. Il concetto di *ownership* di un oggetto costituisce un potenziale pericolo per l'intero sistema:

- un utente potrebbe per errore consentire l'accesso indebito a informazioni riservate cambiando i permessi di accesso di un file
- un processo lanciato da un utente può effettuare qualunque operazione sugli oggetti di proprietà dell'utente
- un HTTP server compromesso ha accesso a tutti i file del gruppo web

Eccetera…

MAC

Al contrario, in regime di **Mandatory Access Control**, ovvero non-discretionary access control, non esiste il concetto di ownership: solo l'amministratore della sicurezza può (e di fatto deve) definire come ogni *subject* interagisce con gli oggetti del sistema.

Il kernel controlla dunque non solo l'identità dell'utente, ma anche tutte le altre informazioni in suo possesso, prima di consentire a un processo di effettuare una qualunque operazione su un oggetto. In questo modo a ciascun processo può essere dato puntualmente l'insieme dei soli permessi di cui ha bisogno per funzionare.

Si può cioè instaurare la regola del *minore privilegio*. Ad esempio, se una policy stabilisce che gli oggetti nella home di un utente sono accessibili solo dall'utente stesso, questo varrebbe anche se questi lanciasse ***`chmod -R a+rw ~`***

Storage management

A parte le piccole realtà i cui sistemi si riducono a qualche server, i dispositivi di memorizzazione in una infrastruttura informatica richiedono una specifica comprensione, gestione e pianificazione

Caratterizzazione: costo specifico, performance, disponibilità, accessibilità, modularità

Topologia e gestione: distribuito, centralizzato

Tecnologie: ESDI, SCSI, SAS, {P,S}ATA, SSA, F.C., iSCSI, DLT, LTO, AIT

Anche senza contare le tecnologie a nastro, i costi specifici variano di 3 ordini di grandezza: da 50€ a 50k€ per TB e più

Storage management

I costi variano per innumerevoli motivi, dalle quantità di dispositivi prodotti alle metodologie di test e progettazione, dalle necessità di collegamento alle funzionalità diagnostiche, dalla facilità di gestione alle caratteristiche di ridondanza e tolleranza ai guasti

Ad esempio, si consideri uno storage device che possa ospitare sia dischi F.C. che SATA: se è vero che il costo per TB di questi ultimi è molto più basso (anche di un fattore 5), è pure vero che anche l'MTBF è inferiore di uno o due ordini di grandezza; questo porta a concludere che anche l'utilizzo di questi dispositivi dovrà essere diverso

Storage management

I criteri per l'adozione di soluzioni storage sono dunque molteplici, ad esempio:

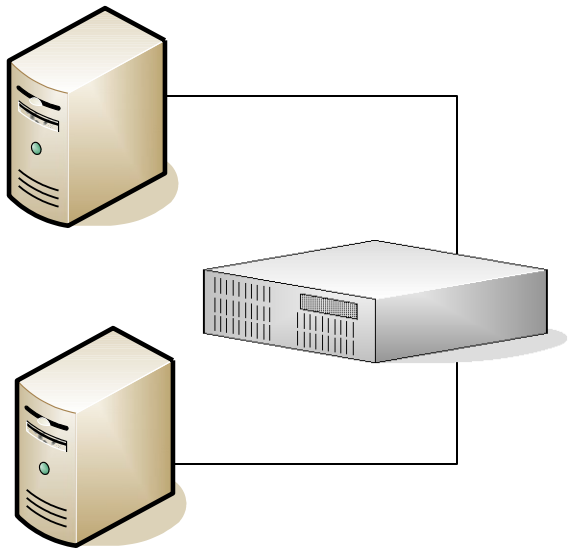
- Capacità e accesso
- Performances desiderate
- Scalabilità
- Disponibilità e affidabilità
- Protezione dei dati
- Disponibilità di skill per la gestione
- Disponibilità finanziarie

Storage management

Si parla di Direct Attached Storage per i dispositivi disco collegati a sistemi che hanno la diretta gestione dello spazio disco e della loro organizzazione: i dischi interni ai server, array JBOD o RAID interni o esterni con collegamenti SCSI o SAS sono tutti esempi di sistemi DAS

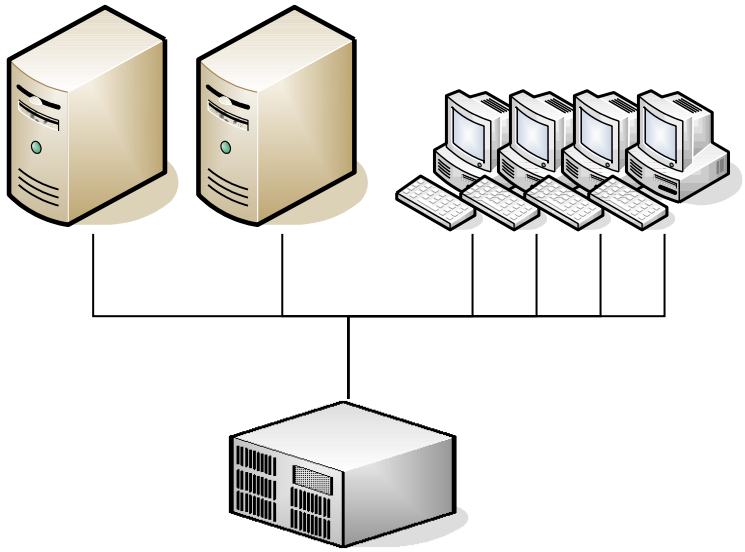
Per grosse quantità di dati, dove il numero di sistemi sia molto ridotto o non vi siano grossi problemi di scalabilità, l'installazione di uno o più dispositivi DAS consente di razionalizzare la gestione degli array disco

Altri eventuali componenti di sistema operativo (es. Logical Volume Manager) consentono poi la gestione dello spazio disco disaccoppiando ulteriormente partizioni e filesystem dai dispositivi disco collegati, fisici o logici che siano



Storage management

Si parla di Network Attached Storage in presenza di uno o più dispositivi di memorizzazione che di fatto sostituiscono file server e il tipo di accesso ai dati non è classificabile come "block I/O"



Supportano quindi protocolli di file sharing come NFS, CIFS/SMB, IPX, ecc. e sono integrati o integrabili con directory (es. AD) per la definizione degli utenti e l'assegnazione dei permessi sui dati

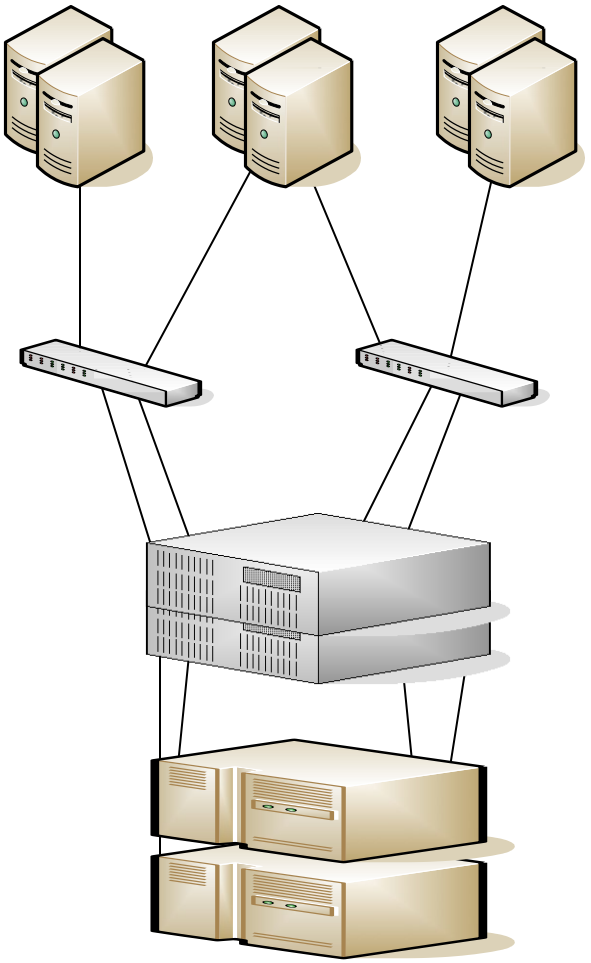
I NAS proteggono i dati mediante opportune ridondanze fisiche e logiche sui dischi e consentono di organizzare lo spazio disco complessivo suddividendolo in volumi logici che sono resi disponibili in rete locale ai sistemi (server o PC) che necessitano di accedere ai dati

Per quanto siano estremamente semplici deployment e manutenzione, con i NAS può essere difficoltosa e costosa l'integrazione con sistemi di sicurezza come antivirus e backup

Storage management

Una Storage Area Network è una soluzione modulare che consente di centralizzare lo storage in una rete specializzata e pensata per trasportare un protocollo dati per l'accesso al disco

I componenti di una SAN sono switch, HBAs, Controllers, disk enclosures, dischi e alimentatori, le cui configurazioni e collegamenti sono pensati e organizzati per massimizzare le ridondanze della memorizzazione e dell'accesso ai dati: doppio controller, doppi percorsi verso ciascun disco (FCAL) e verso gli switch (almeno due); ciascun server a sua volta può arrivare allo storage attraverso percorsi singoli o doppi (con gestione del Multipathing da parte del sistema operativo)

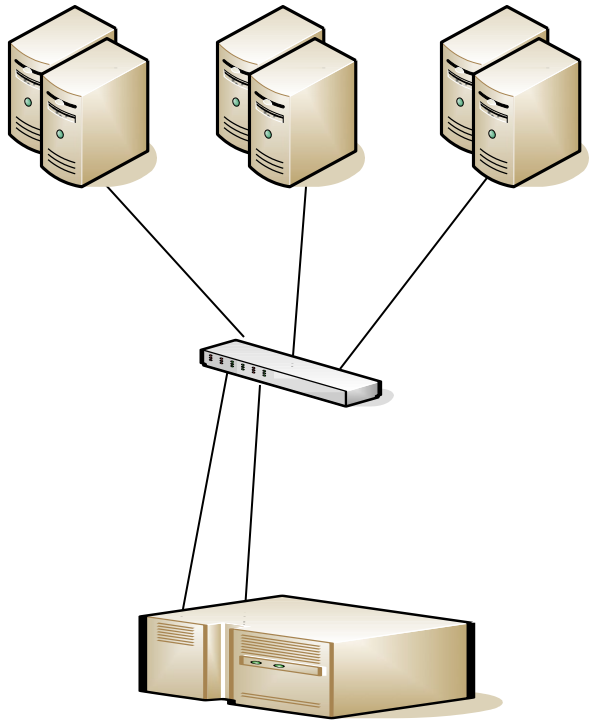


Storage management

Una iSCSI SAN è una SAN semplificata, che sfrutta l'infrastruttura GbE per consentire il collegamento dei dispositivi disco e dei server

Le prestazioni non sono confrontabili con quelle di una SAN in tecnologia Fiber Channel (anche se le velocità di trasmissione nominali lo sarebbero: 1, 10 Gbps per iSCSI e 2, 4 Gbps per F.C.)

In compenso, la semplicità infrastrutturale è indiscutibile e richiede pochi o nessun investimento, anche in termini di skill; soprattutto il collegamento con i server è grandemente semplificato perché non richiede necessariamente l'installazione di HBAs, in quanto può essere usata una delle due porte GbE che tutti i server hanno integrate sulla scheda madre utilizzando un initiator software



Storage management

Le SAN hanno un software di gestione che consente di configurare i dischi, definendone la configurazione in termini di ridondanza fisica e logica, creando raggruppamenti (disk groups) in differenti configurazioni RAID e associandovi dischi hot spare; un disk group opportunamente configurato diventa un unico spazio disco virtuale dotato opportune caratteristiche di ridondanza e prestazioni e tolleranza ai guasti dei singoli dischi fisici

Lo spazio disco virtuale di ciascun disk group viene poi segmentato in LUN, unità disco virtuali che sono associate ai server collegati alla SAN (presentazione)

Ciascun server vede e tratta le LUN ad esso presentate come se fossero dischi fisici

Storage management

Concentrando tutta la gestione fisica dello storage e implementando in un unico sistema ridondanze, scalabilità e gestione, si può aumentare la disponibilità dei sistemi e semplificare la gestione dello spazio disco; se i dispositivi lo supportano, diventa possibile anche la differenziazione dello stesso in tier che forniscono diversi livelli di servizio, a diversi costi/TB

- Tier 1: databases
  (RAID 0 o 10 + dischi F.C. piccoli 15kRPM)
- Tier 2: general use & High Performance file servers
  (RAID 5 o 6 + dischi F.C. grandi 10kRPM)
- Tier 3: archiving and other file servers
  (RAID 5 + dischi SATA/F.C. molto grandi)

Storage management

Anche i dispositivi a nastro sono molto importanti nella progettazione dello storage

Oltre alla tecnologia scelta per le cartucce, sono importanti anche la capacità, la scalabilità, il tipo di collegamenti, il parallelismo del dispositivo, le funzionalità diagnostiche, ecc

Le librerie automatiche possono ospitare dalle diecine alle migliaia di cartucce e hanno diversi drive anche di diverso tipo (es. LTO 2 e 3); nei modelli più sofisticati la robotica può essere duplicata e la libreria stessa può crescere modularmente

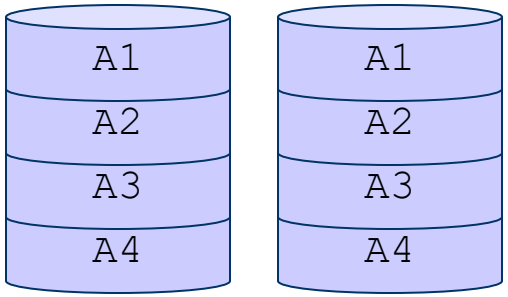
Collegamenti alle librerie in fibra ottica sono spesso usati per tenere i nastri o una loro copia in una sede separata senza “trasportarli”

Storage redundancies

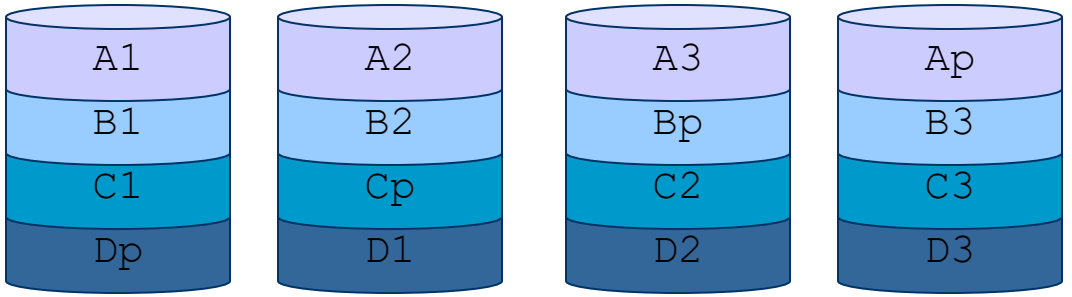
Anche se i dischi moderni hanno la capacità di far fronte a un determinato numero di problemi o guasti al materiale magnetico, riallocando automaticamente un numero finito di settori problematici su tracce spare (es. la tecnologia S.M.A.R.T. dei dischi ATA), per tollerare i guasti dei dischi si fa uso di configurazioni RAID (Redundant Array of Inexpensive Disks), che consentono ai dati di essere scritti su più di un singolo dispositivo fisico

RAID levels più usati: 1 (mirroring), 5 (distributed parity)

RAID 1



RAID 5



Storage redundancies

Nei dispositivi da una certa classe in su esistono inoltre meccanismi per la replicazione dei dati all'interno del dispositivo stesso (es. flash copy) o da un dispositivo a un altro (remote copy) tramite collegamenti fra switch della SAN

La funzionalità di flash copy consiste nella rapida duplicazione delle LUN; con la copia si possono poi effettuare operazioni di backup, sia file che image, senza fermare i sistemi

Con la remote copy è invece possibile mantenere una copia sempre allineata di LUN importanti su un dispositivo gemello situato in un'altra sede, eventualmente collegato a un server dello stesso tipo; è più costosa perché duplica tutti i blocchi disco

Tra queste operazioni la flash copy è svolta duplicando puntatori di allocazione a blocchi disco, per cui è estremamente efficiente e praticamente istantanea: dopo che i puntatori sono stati copiati, la copia viene lasciata intatta e nell' originale si duplicano solamente i puntatori ai blocchi disco che vengono modificati durante il periodo di esistenza della flash copy (copy on write)

Consolidations

Sono metodologie di rinnovamento tecnologico che riscuotono da tempo notevole successo commerciale per i risvolti economici particolarmente favorevoli

Storage consolidation

Server consolidation

Entrambe *non risolvono* specifici problemi di sicurezza e disponibilità dell'informazione, ma hanno importanti implicazioni che le riguardano direttamente e profondamente

Storage consolidation

DAS

Gestione diversificata e complessa, configurazioni errate, logging & monitoring mancanti o incompleti, scarso utilizzo delle risorse disponibili

NAS

Scarso controllo sui firmware e gli o.s., supporto antivirus, supporto a directory esistenti

SAN

Complessità di progettazione e zoning, vincoli sulle matrici di compatibilità hw/sw, costi di ingresso elevati

Server consolidation

Software-centrica

Raccolta di servizi multipli operanti su sistemi diversi, ottimizzazione e realizzazione con uno o pochi sistemi più grossi

Hardware-centrica

Virtualizzazione a livello di sistema operativo di più sistemi operanti su uno o pochi sistemi più grossi

“Problem consolidation”

Problemi

Le operazioni di consolidation ottimizzano le risorse e i costi del rinnovamento tecnologico, ma possono essere in contrasto con le politiche di sicurezza già in essere

- In-depth security: aggregare più servizi su un unico sistema comporta più benefici di sicurezza di quelli che si hanno se i servizi sono su sistemi separati?
- Security efforts: proteggere un singolo sistema complesso è più facile che proteggerne diversi semplici?
- Compatibility: servizi come il real-time scanning di un antivirus sono disponibili nella soluzione consolidata?
- IP storage: “nooo, ancora IPSEC”?

Enterprise backup

Sottoproblema dello storage management

Progettato per eseguire con flessibilità e automaticamente il backup di tutte le risorse di una rete

- Backup e archive
- Modello client-server
- Modello LAN-free
- File level & system level

Enterprise backup

Il vero problema del backup è in realtà il restore

Se il ripristino ordinario di file cambiati o cancellati accidentalmente non è critico, lo è invece il ripristino di interi sistemi a fronte di gravi malfunzionamenti hardware o perdite di dati in dispositivi di storage centralizzati

La criticità consiste nel fatto che il tempo impiegato per il ripristino dei dati deve essere compatibile con il fermo delle attività che sono coinvolte dall'operazione

Il ripristino di un sistema può avvenire con tecniche tradizionali (reinstallazione, configurazione e restore dei dati) o più avanzate (bare metal restore); in entrambi i casi occorre comunque partire da un hardware nuovo o ripristinato, il cui approntamento (se non già disponibile) richiede peraltro un ulteriore tempo che si va a sommare a quello necessario per il ripristino del sistema

Enterprise backup

Anche indirizzando adeguatamente il restore, la distribuzione delle risorse applicative fa porre l'attenzione non tanto sul funzionamento dell'intero processo di backup e della verifica di restorabilità dei dati, quanto alla

subset consistency

degli stessi. Più un'applicazione è distribuita, più è difficile effettuarne il restore dei dati in caso di problemi. Per ogni applicazione infatti occorre analizzare quali dati usa e dove sono, in modo che sia possibile effettuarne il restore in modo consistente. Molte applicazioni condividono poi risorse (come i DBMS per esempio) ed è complesso ottenere un ripristino dei dati consistente per ogni applicazione ospitata.

Disaster Recovery

Quando si pensa a guasti o indisponibilità di molti o tutti i sistemi, dovuti a incendi, allagamenti ecc., diviene naturale ipotizzare metodologie di Disaster Recovery

Il tradizionale modello di inizio anni '90 era focalizzato solamente sulle risorse IT; consiste per esempio nel programmare, a intervalli di tempo dichiarati, l'esportazione di una seconda copia dei nastri di backup da depositare in un sito diverso e nel pianificare la disponibilità di hardware necessario ad effettuare il ripristino dei dati

Questo problema va comunque visto in discipline più complete, come la Business Continuity, e che studiano come consentire a un'intera attività aziendale (nelle sue componenti di ricerca, di servizi, commerciale, distributiva e produttiva) di ripartire nel più breve tempo possibile a seguito di catastrofi o danni di grave entità a una o più sedi della stessa

Business Continuity

La Business Continuity *non* è un problema IT

A che cosa serve infatti poter ripristinare tutte le risorse informatiche di supporto alla produzione e alla vendita se non si riescono parimenti a ripristinare le risorse primarie atte a svolgere queste funzioni (es. telefonia, magazzino, linea di produzione)?

Enterprise scheduling

Sottoproblema dell'application management

Il *cron* di Unix e le *operazioni pianificate* di Windows sono esempi di semplice *scheduling*. L'*enterprise scheduling* è invece più complessa: è l'unione di sistemi e metodologie che permettono l'automazione di task distribuiti su diversi sistemi; possono essere coordinati nel tempo e correlati fra loro da una logica di controllo dotata di semantica applicativa.

I prodotti che implementano queste metodologie sono di derivazione mainframe e sono diffusi solo nelle grandi realtà a causa dei costi non proprio moderati.

Il grande valore di questi prodotti nell'ambito della sicurezza dell'informazione consiste nella specifica capacità di razionalizzare e centralizzare la "logica" di scheduling che altrimenti verrebbe di fatto distribuita fra i sistemi senza alcuna possibilità di creare correlazioni e automatismi, perdendone così nel tempo il controllo e la conoscenza.

Enterprise scheduling

Uno scheduler enterprise è un sistema client-server, con uno o più motori di scheduling, e con agenti installati nei diversi sistemi su cui devono girare i task. E' dotato di una console di amministrazione che serve per il governo e la programmazione delle regole che attivano e correlano i task stessi.

Esempio di schedulazione distribuita.

Far partire il backup di un server alle 04:00 è un task facilmente automatizzabile con cron. Non si riesce invece facilmente a realizzare un automatismo che permetta di fare la stampa dei tabulati del magazzino alle 07:00 se e solo se

- *l'elaborazione X sul sistema Y è terminata con successo*
- *la replica delle informazioni provenienti dal sistema Z è completa*
- *è lunedì*

Il primo punto nasconde una complessità: una qualunque elaborazione può andare a buon fine innanzitutto se tutti i suoi prerequisiti hanno terminato con successo; la prima condizione in realtà ne nasconde e implica una quantità di altre condizioni o task terminati con successo

System redundancies

Per massimizzare l'integrità e la disponibilità dell'informazione è necessario ricorrere a ridondanze nei sistemi hardware e software

- Dischi in RAID con hot spare multipli
- Alimentatori *n* + *1* e *n* + *n*
- ECC-RAM, Chipkill-RAM, “RAID”-RAM
- Processori multipli SMP e NUMA
- Partizionamento fisico e logico multiplo
- Components hot swappability
- Network card bonding
- Multiple path to storage
- …

System redundancies

La tendenza recente è di estrarre dai server alcune componenti comuni quali

- Alimentatori
- Porte di rete
- Porte SAN
- KVM
- Alcune periferiche (CD, ecc)

Nascono così i blade systems, sistemi rack 19” contenenti tutti i componenti comuni e ospitanti i server in formato ultracompatto; i componenti comuni sono ridondanti e controllati da un sistema di gestione integrato unico per tutti i blade server montati nell’enclosure

System redundancies

Supercomputers

Vecchi criteri: ridondare tutto, anche il BUS ecc

Forte dipendenza dal sistema operativo: solo un certo S.O. popteva gestire HW ridondante →Z sistemi proprietari come SUN (ora Oracle)

E le comunicazioni con l'esterno?

Anche le applicaizoni standard più di tanto non gradiscono l'esecuzione su questi sistemi (es riallocazione dell'I/O su bus alternativo)

COSTI ^ ^ ^

System redundancies

Autonomic Computing

Recenti criteri di progettazione dei sistemi suggeriscono di evitare le ridondanze dello hardware e piuttosto dotare lo stesso di funzionalità autodiagnostiche predittive; questo approccio consente un notevole risparmio complessivo:

- la complessità totale a livello hardware è poco diversa da un progetto tradizionale e non vi è completa o anche parziale duplicazione (e spesso inutilizzo) di componenti funzionali
- non è necessario modificare i sistemi operativi con nuove funzionalità e semantiche per utilizzare le ridondanze
- consente di spostare le attività di manutenzione dal post-incident al pre-incident e quindi di pianificare gli interventi

Server clusters

Fault-tolerant
Tempi di recovery trascurabili, nessuna perdita transazionale

High Availability
Tempi di recovery elevati (secondi), perdita di transazioni uncommitted

Disaster Recovery
Perdita di transazioni committed

Computationally intensive
HPC e MPP

Load Balancing
Solo distribuzione del carico

Cluster architectures

Single image

Single-init, *single-root*, *single process space* e *process migration*, *single IPC, networking, device space*

Shared disk

Diverse immagini di OS e applicazioni condividono uno stesso spazio disco

Shared nothing

Diverse immagini di OS e applicazioni hanno ciascuno un proprio spazio disco

Cluster architectures

Shared disk - MS e RedHat

Microsoft e RedHat come esempi di tradizionale cluster shared disk

- spazio disco condiviso e accessibile dalle immagini in modalità esclusiva
- *quorum*: spazio disco condiviso che mantiene lo stato del cluster insieme con quello delle immagini, accessibile contemporaneamente da entrambe in modalità raw
- le applicazioni sono suddivise in gruppi, affidati in gestione a un nodo o dell'altro
- le applicazioni sono presentate all’esterno mediante indirizzi virtuali, sostenuti da un nodo o dall’altro
- i nodi si controllano a vicenda mediante meccanismi di heartbeat
- quando un nodo fallisce l'altro fa partire le applicazioni dei gruppi che non possiede e assume il corrispondente indirizzo

Cluster architectures

Shared disk - Oracle RAC

Oracle propone invece un cluster FT e shared-disk denominato RAC

- *cluster filesystem* ad accesso multiplo, per coordinare l'accesso delle istanze allo spazio disco condiviso
- *cache fusion*: ciascun nodo contribuisce a una parte della “cluster cache” complessiva dei blocchi disco
- *block shipping* vs *function shipping*
- quando un nodo fallisce un altro prende possesso dei blocchi disco appartenuti alla cache di quel nodo

Cluster architectures

Shared disk e nothing – IBM DB2

IBM ha sempre evidenziato il vantaggio di avere un'architettura di RDBMS che può essere installata come un cluster tradizionale e come un cluster shared nothing contemporaneamente, consentendo una adeguata tolleranza ai guasti e funzionalità avanzate di partizionamento

Per potersi confrontare con le più avanzate funzionalità di clustering di Oracle (RAC), IBM ha recentemente arricchito il suo prodotto con HADR (High Availability & Disaster Recovery), una configurazione con due sistemi attivi (primario e secondario; il primario manda continuamente i log delle transazioni al secondario che così può effettuare le stesse modifiche al database;

Cluster architectures

Shared nothing – Lotus Domino

E’ basato sulla *core-technology* di Domino: la funzionalità di *replica* dei database di documenti (i DB sono gli elementi applicativi di Domino e la tecnologia che ne ha decretato il successo commerciale è replica)

- i nodi del cluster replicano continuamente i DB definiti come clustered
- la logica di failover è nel client
- ciascun nodo ha il suo storage dedicato
- nessun vincolo particolare

I nodi possono anche essere remoti, diversi fra loro come versione di software, ecc; è sufficiente che ciascuno abbia storage sufficiente a contenere la propria copia di ciascun DB replicato

Virtualization

Hardware level
Partitioning, VMWare, Microsoft Hyper-V, RHEV

Operating System level
FreeBSD jails, XEN, UML, KVM

High Level Language level
p-code, SmallTalk, Java, CLR

Compatibility
VM di vendor diversi potrebbero essere interoperabili, e comunque lo sono all'interno di una soluzione single vendor

Isolation
dalle altre VM e dal sistema operativo ospite

Encapsulation
il VM monitor consente di controllare l'esecuzione delle VM per ottimizzare memoria e uso della CPU e per effettuare validazioni a runtime (es. type checking)

Benefits/Performance
la virtualizzazione deve risultare complessivamente conveniente

Virtualization

Paravirtualization

Si parla di paravirtualization quando il supervisor o host offre specifiche interfacce all'ambiente guest volte a ottimizzare le performance; XEN ad esempio consente al guest di "vedere" alcuni dispositivi di I/O (dischi, schede di rete) non come dispositivi virtualizzati, ma con interfacce simili a quelle del sistema host:

il guest cioè può installare driver disco e di rete specifici e che lavorano in modo coordinato con quelli dell'host, in modo che l'accesso all'I/O delle VM non sia virtualizzato ma reale

Virtualization

Application virtualization

E’ una forma di virtualizzazione parziale, in cui l’hypervisor consente di creare ambienti isolati (*sandbox*) in cui installare software in un sistema ospite, principalmente MS Windows

Softricity (ora Microsoft App-V), Altiris (ora Symantec), VMWare ThinApp, Citrix XenApp…

Gli oggetti che vengono virtualizzati in questo caso sono, insieme con l’applicazione, quelle parti del sistema che vengono modificate dall’installazione e dall’esecuzione dell’applicazione stessa: filesystem e registry, incluse dll e driver specifici.

Le applicazioni vengono in realtà installate una tantum e fornite tramite tecniche di streaming

Virtualization

Vmware

Non è un'idea nuova: IBM VM (1960)

Attualmente è il più completo ambiente di virtualizzazione hardware level disponibile sul mercato

Il successo di prodotti come questo, che realizza la virtualizzazione hardware level per le architetture Intel è la conseguenza dell'enorme potenza elaborativa disponibile nei calcolatori odierni e dall'esigenza di razionalizzarne l'uso per vari motivi, soprattutto organizzativi ed economici

Versioni: Player, Workstation, Server, Enterprise

Virtualization

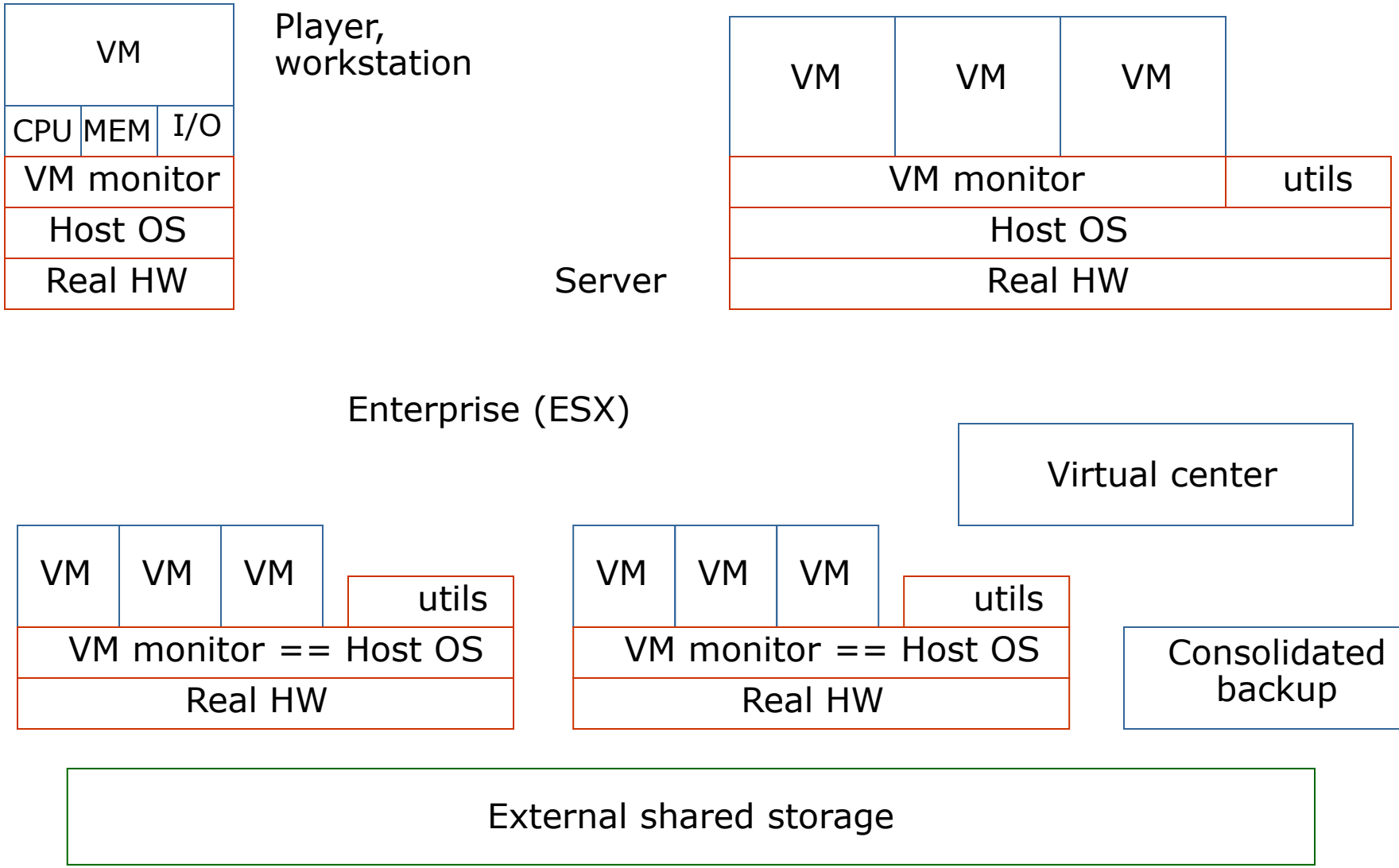
Altri sistemi di virtualizzazione emergenti:

Microsoft Hyper-V

...

RHEV
Red Hat Enterprise Virtualization

Virtualization



Virtual infrastructure

Concetti base

Hardware virtuale (CPU, memoria dispositivi di I/O), dischi virtuali (persistenti o undoable), switch di rete virtuali

Vmware Enterprise (ESX)

Al contrario di tutte le altre versioni, in cui il VM monitor è un processo del Sistema Operativo ospite, in ESX il VM monitor è il Sistema Operativo stesso (custom Linux) o una versione "stripped down" (ESXi)

Questo consente un controllo molto fine dell'assegnazione delle risorse fisiche a quelle virtuali (CPU, memoria e I/O) e quindi performance e funzionalità enormemente superiori

Virtual Infrastructure

E' l'insieme di almeno due nodi ESX con storage condiviso e coordinati da un gestore di funzionalità comune a tutti i server (Virtual Center) e un gestore del backup (VMWare Consolidated Backup)

Virtual infrastructure

VMFS
filesystem ad accesso multiplo ottimizzato per l'I/O su file di grandi dimensioni e da parte di hardware virtuale; più server fisici (nodi) ESX accedono contemporaneamente alle partizioni VMFS realizzate su storage esterni (SAN)

Snapshots e dischi virtuali
I dischi virtuali possono essere configurati come "undoable": vengono "fotografati" e tutte le modifiche apportate sono scritte in un REDO LOG, che può essere cancellato (si ritorna al momento dello snapshot) oppure applicato (ridiventa disco normale e lo snapshot sparisce)

Virtual SMP
Capacità di virtualizzazione di hardware multiprocessor, eventualmente mappato sull'architettura del processore fisico (hyperthreading, dual-core, ecc.)

Virtual infrastructure

Ballooning

Ottimizza l'utilizzo della memoria fisica da parte delle VM; le pagine di memoria inutilizzate sono identificate e il monitor può assegnarle a macchine che ne hanno bisogno; permette di effettuare l'*overbooking* della memoria

Swapping

Le aree di swap delle VM vengono ricavate prioritariamente nella memoria fisica del monitor, poi in aree disco dedicate, infine nella memoria virtuale (swap) del monitor stesso

Converter

Consente di copiare e convertire sistemi fisici in virtuali o a convertire VM da server VMWare di tipo diverso

Virtual infrastructure

VMotion
Le VM possono migrare da un nodo a un altro anche a caldo, senza interruzione del servizio

Stesso storage :: diversi host

Storage VMotion
Le VM possono migrare da uno storage a un altro, mentre sono in esecuzione su un determinato nodo

Stesso host :: diversi storage

Distributed Resource Manager
Gli host fisici vengono raggruppati in *server pool* caratterizzati da regole di esercizio comuni; il monitor gestisce la ripartenza automatica delle VM in caso di crash e consente anche l'allocazione automatica delle VM all'interno di ciascun pool (tramite vmotion) sui server fisici più opportuni, onorando le regole stabilite per ciascun pool (ad esempio sull'occupazione di CPU, sul carico di I/O, l'occupazione di memoria eccetera)

Virtual infrastructure

Consolidated Backup (VCB)

Un server separato con accesso ai filesystem VMFS condivisi consente di effettuare il backup immagine di ciascuna VM senza fermarla; l'operazione consiste nell'esecuzione della sequenza: *snapshot start, copy disk, snapshot release, backup disk, delete disk*

Vsphere 4: RAID 1 VM

Nella versione 4 di ESX (denominata ora Vsphere 4) è stata introdotta una funzionalità che somiglia al mirroring dei dischi e potrebbe essere chiamata "Clusterless High Availability": su due nodi fisici diversi due macchine virtuali sono in realtà una la copia dell'altra, mantenuta continuamente aggiornata con la stessa tecnologia di replica della memoria e del disco usata per le feature "vmotion" e "storage vmotion"; quando il nodo che ospita la VM principale fallisce, subentra quella secondaria in tempi molto brevi (anzi comunque inferiori a quelli di un cluster tradizionale)

Virtual infrastructure

Usi e motivazioni

Gli usi più comuni delle infrastrutture virtuali e della virtualizzazione in genere sono molteplici, tutti aventi come caratteristiche la razionalizzazione delle risorse e la maggiore economia e semplicità di utilizzo rispetto ai sistemi fisici

- *fast provisioning mediante macchine virtuali "template"*
- *copie di sistemi di sviluppo e test del software*
- *sistemi ad elevata disponibilità senza fare uso di cluster*
- *cluster in-a-box, across-boxes e physical to virtual*
- *failover rapido dei sistemi*
- *copie di sistemi per disaster recovery*
- *test environment per patching e upgrades*
- *production-cloning-test provisioning*

Virtual infrastructure

Non si può virtualizzare tutto

Esempi lampanti sono: il Virtual Center, il VCB, il server che gestisce gli switch KVM

Il Virtual Center non va virtualizzato perché governa l'intera infrastruttura, il VCB perché richiede l'accesso fisico alla SAN, eventuali server necessari al funzionamento di dispositivi come gli switch KVM non possono essere virtualizzati perché altrimenti quando si fa manutenzione dell'infrastruttura virtuale le console dei server fisici diventerebbero inaccessibili

Non si deve virtualizzare tutto

La virtualizzazione di servizi essenziali come il DNS della rete o il Domain Controller di un AD Microsoft va ben valutata e progettata per evitare che tali servizi non vengano a mancare quando si fa manutenzione dell'infrastruttura virtuale, per esempio prevedendo dei secondari fisici (o viceversa)

Alcuni sistemi (es. Oracle) non vengono supportati dal fornitore se utilizzati in produzione su ambienti virtuali

Virtual infrastructure trends

Questi ultimi anni hanno visto il consolidarsi delle tecnologie di virtualizzazione e la loro diffusione soprattutto nel consolidamento delle sale server. C'è poca differenza fra le tecnologie e le caratteristiche funzionali su cui le varie soluzioni si confrontano fra loro sono soprattutto gestionali ed economiche. Aspetti da considerare, per scegliere o cambiare una soluzione, sono comunque:

- funzionalità di base
- tecniche di gestione dell'infrastruttura
- tool di migrazione
- licensing e costi

E' inoltre in corso di diffusione la proposta di virtualizzare anche i client: l'idea di base è che i PC possano essere virtualizzati e risiedano nel datacenter sotto forma di immagine accessibile via rete e anche da remoto

Domande?

Vincenzo Calabrò
info@vincenzocalabro.it