

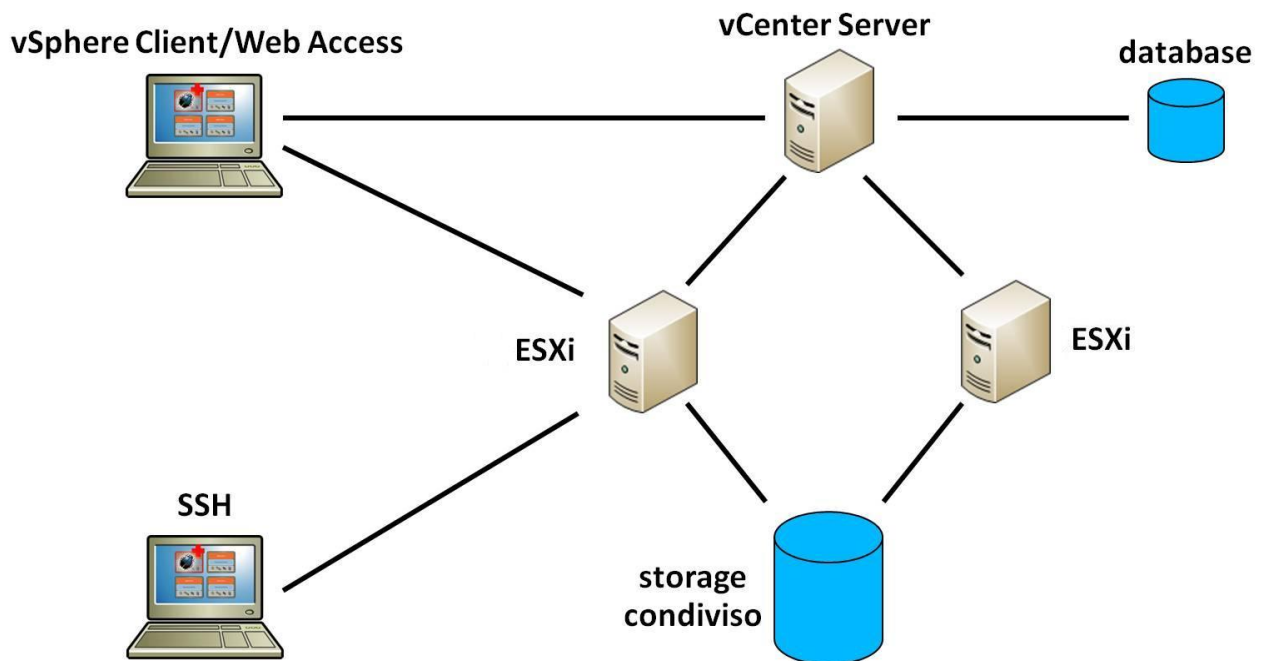


Sizing di un'infrastruttura server con VMware

v1.1

Matteo Cappelli

Vediamo una serie di best practices per progettare e dimensionare un'infrastruttura di server virtuali con VMware vSphere 5.0. Innanzitutto parlando di VMware, e quindi vSphere per la parte server, è bene ricordare l'architettura di base, visibile in figura.



I punti da tenere in considerazione durante la progettazione sono i seguenti (ordinati in base all'importanza):

1. tipo di server (entry level, midrange, enterprise, ...), al quale sono strettamente legati tutti i punti successivi, e l'espandibilità futura (scale-out e scale-up);
2. RAM, l'elemento più critico che può limitare fortemente le VM e le loro performance all'interno dell'infrastruttura;
3. CPU, in termini di numero di core, socket, e frequenza;
4. storage, quantità dello spazio necessario e tecnologia usata;
5. networking.

Di seguito vengono presi in esame questi 5 punti sia per la configurazione del server su cui girerà l'hypervisor (ESXi), sia per quello vCenter Server. Ricordo che ancor prima di progettare il server, così come per lo storage e ogni altro dispositivo (incluse le singole componenti, NIC, HBA, ...), è fondamentale controllare che VMware supporti il dispositivo in questione. Per verificare la compatibilità hardware è possibile visitare il seguente link ed inserire il modello di server scelto:

<http://www.vmware.com/resources/compatibility/search.php>

Server per ESXi

In questo paragrafo viene esaminato il procedimento da adottare per progettare adeguatamente il server che ospiterà ESXi, l'unico sistema utilizzabile con vSphere 5, mentre fino alla versione 4.1 era possibile optare anche per ESX. Per dimensionare il server è necessario prima di tutto sapere se verrà usata la versione installabile o quella embedded:

- nel caso in cui ESXi venga installato nei dischi interni, ovviamente questi devono essere predisposti nel server;
- se invece viene usata la versione embedded, con ESXi installato in una chiavetta USB, o se viene fatto boot da SAN, i dischi interni del server sono un plus e possono essere anche omessi.

Per quanto riguarda la definizione delle caratteristiche minime per questi server, come punto di partenza è possibile utilizzare il seguente link ufficiale:

http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1003661

per poi considerare le risorse necessarie alle VM. Lo step consigliato è quello di eseguire all'interno dell'infrastruttura fisica da virtualizzare un tool specifico, che raccolga informazioni sui server e sul loro carico di lavoro. Ad esempio VMware mette a disposizione il tool *VMware Capacity Planner*, che in modo automatico (richiede solo una fase di installazione e configurazione iniziale) raccoglie tutte le informazioni necessarie per il consolidamento. In alternativa è possibile dimensionare l'infrastruttura "manualmente", conoscendo le caratteristiche dei server fisici da consolidare, il SO, i tipi di servizi/applicazioni in esecuzione, e le caratteristiche della rete (switch che verranno collegati agli host ESXi, velocità, VLAN definite, ...). In ogni caso l'attività di pianificazione e dimensionamento dei nuovi server è piuttosto lunga e complessa, e soprattutto nel dimensionamento "manuale" bisogna studiare attentamente l'infrastruttura di partenza per evitare colli di bottiglia nella nuova infrastruttura. Provate ad immaginare cosa succederebbe se l'infrastruttura creata non avesse RAM sufficiente per eseguire le VM previste? E in aggiunta se il server avesse già raggiunto la massima espandibilità e non fosse possibile aggiungere ulteriori banchi di memoria? Sarebbe un disastro!!!

Partiamo dalla configurazione della RAM. La memoria necessaria ad un host sarà equivalente alla somma di più fattori:

- memoria necessaria all'hypervisor (2 GB, anche se in teoria potrebbe essere sufficiente 1/1,5 GB);
- memoria per le VM;
- memoria per l'overhead di ogni VM.

È bene infatti ricordare che ogni VM, oltre alla quantità di RAM per essa definita, utilizza una piccola quantità ulteriore per il suo funzionamento e gestione da parte dell'hypervisor. Nello

screenshot sottostante è possibile vedere la memoria definita per la VM (4 GB), e la memoria di overhead (183 GB circa).

General	
Guest OS:	Microsoft Windows Server 2008 R2 (64-bit)
VM Version:	7
CPU:	1 vCPU
Memory:	4096 MB
Memory Overhead:	183.05 MB
VMware Tools:	Not installed
IP Addresses:	
DNS Name:	
EVC Mode:	N/A
State:	Powered Off
Host:	192.168.126.42
Active Tasks:	

L'overhead generalmente è compreso tra 150 e 300 MB, in base alla quantità di RAM definita per la VM e al numero di vCPU, e come riferimento può essere utilizzata la tabella seguente (fonte VMware). Non sono riuscito a trovare riferimenti a casi con più di 8 VCPU, dato che vSphere 5.0 ha il limite di 32, anche se dubito che siano molto usate VM con VCPU maggiori di 8.

Overhead Memory on Virtual Machines

Memory (MB)	1 VCPU	2 VCPUs	3 VCPUs	4 VCPUs	5 VCPUs	6 VCPUs	7 VCPUs	8 VCPUs
256	113.17	159.43	200.53	241.62	293.15	334.27	375.38	416.50
512	116.68	164.96	206.07	247.17	302.75	343.88	385.02	426.15
1024	123.73	176.05	217.18	258.30	322.00	363.17	404.34	445.52
2048	137.81	198.20	239.37	280.53	360.46	401.70	442.94	484.18
4096	165.98	242.51	283.75	324.99	437.37	478.75	520.14	561.52
8192	222.30	331.12	372.52	413.91	591.20	632.86	674.53	716.19
16384	334.96	508.34	550.05	591.76	900.44	942.98	985.52	1028.07
32768	560.27	863.41	906.06	948.71	1515.75	1559.42	1603.09	1646.76
65536	1011.21	1572.29	1616.19	1660.09	2746.38	2792.30	2838.22	2884.14
131072	1912.48	2990.05	3036.46	3082.88	5220.24	5273.18	5326.11	5379.05
262144	3714.99	5830.60	5884.53	5938.46	10142.83	10204.79	10266.74	10328.69

Vediamo un esempio completo. Supponiamo di avere una infrastruttura fisica di 50 server, ognuno con una sola vCPU e 4 GB di RAM, e vogliamo virtualizzarli in 3 nodi ESXi. La memoria totale necessaria sarà di 200 GB, più un overhead di almeno 200 MB per VM, ossia altri 10 GB, più 2 GB per l'hypervisor in ognuno dei server: totale $200+10+(2 \times 3)=216$ GB di RAM. Poiché i nodi della nuova infrastruttura sono 3, il totale deve essere suddiviso per 3, dunque 72 GB di RAM per ogni host. Perfetto!? Non del tutto... Non abbiamo ancora considerato un eventuale crash di un host, e qualora la licenza di VMware acquistata preveda la funzionalità HA bisogna rivedere la quantità di RAM. Se si guasta uno dei 3 nodi, i due rimanenti infatti devono essere in grado di eseguire tutte o

quasi tutte le VM (a seconda della richiesta del cliente, e delle VM con priorità maggiore) che erano in esecuzione sull'host crashato. Un'infrastruttura di questo tipo viene tipicamente chiamata "2+1". Per continuare ad operare senza alcun problema, i due server rimanenti devono disporre di una quantità ulteriore di RAM per poter gestire il carico delle VM del host guasto. Poiché il server crashato ha 73 GB, entrambi i server funzionanti devono avere ulteriori 35 GB di RAM (non si considera la memoria per l'hypervisor) per entrambi questi nodi: totale di ognuno dei due nodi $73+35=98$ GB. Non sapendo però quale dei tre nodi si possa guastare, ognuno dovrà avere 98 GB di RAM. Fino ad ora non abbiamo considerato le tecniche di ottimizzazione della memoria di vSphere, dette memory overcommitment (transparent page sharing, memory compression, swapping, ballooning), alcune delle quali permettono di risparmiare qualche GB, soprattutto il tps. Tuttavia il mio consiglio è di **non risparmiare troppo sulla quantità di memoria, anche se il memory overcommitment potrebbe far risparmiare fino al 20% della RAM**. Queste tecniche sono difatti molto variabili, e la memoria liberata dipende da molti fattori che è difficile pianificare. Nella migliore delle ipotesi, in cui le VM finali grazie a queste tecniche utilizzano meno RAM di quella prevista, sarà possibile caricare un numero maggiore di VM o aumentare la loro quantità di memoria (se necessario). Una lettura interessante sul dimensionamento della RAM è la seguente:

<http://jpaul.me/?p=1757>

Per la CPU, che deve essere a 64 bit, l'elemento più importante è il numero di core e di socket, dai quali dipendono il numero di VM che possono essere create. In VMware ogni vCPU di una VM viene associata ad un core fisico, dunque una VM con 4 vCPU verrà eseguita da 4 core fisici diversi (chiamati HEC da VMware), così come 4 VM con una singola vCPU ciascuna. Il limite imposto da VMware è di 25 vCPU eseguibili da un core fisico (HEC), ovviamente per quanto riguarda CPU e server di fascia alta. A livello entry un core fisico può eseguire 4/5 VM con una vCPU, mentre a livello intermedio 10 o anche più. Nel design dell'infrastruttura bisogna considerare che la console di management in ESXi (come la vecchia Service Console di ESX) utilizzerà un core, e solitamente il core 0 del socket 0. Inoltre la frequenza e il modello della CPU di una VM invece corrisponderanno a quelle della CPU fisica, ad esempio se la CPU fisica ha una frequenza x GHz anche la frequenza della vCPU della VM sarà x GHz.

Per lo storage, quello locale ha scarsa importanza in presenza di uno storage condiviso esterno per la memorizzazione delle VM. Solitamente nello storage interno di un host viene installato ESXi, quando non viene usata la versione embedded o non è fatto boot da SAN. Per questi dischi non si ha bisogno di grandi performance, dunque possono essere sufficienti 2 dischi SAS 73GB a 10k giri. Diversa cosa è lo **storage esterno e condiviso, indispensabile in un ambiente con HA e vMotion**. (Non prendo al momento in considerazione la vSphere Storage Appliance, con la quale è possibile condividere tra i nodi i dischi interni ed avere una sorta di spazio condiviso per HA e vMotion. La funzionalità è interessante ma la reputo prematura.)

Come linee guida sulla tecnologia da usare si possono usare le seguenti tabelle.

Networked Storage that ESX Supports

Technology	Protocols	Transfers	Interface
Fibre Channel	FC/SCSI	Block access of data/LUN	FC HBA
iSCSI	IP/SCSI	Block access of data/LUN	<ul style="list-style-type: none"> ■ iSCSI HBA (hardware iSCSI) ■ NIC (software iSCSI)
NAS	IP/NFS	File (no direct LUN access)	NIC

vSphere Features Supported by Storage

Storage Type	Boot VM	vMotion	Datastore	RDM	VMware HA and DRS	VCB
Local Storage	Yes	No	VMFS	No	No	Yes
Fibre Channel	Yes	Yes	VMFS	Yes	Yes	Yes
iSCSI	Yes	Yes	VMFS	Yes	Yes	Yes
NAS over NFS	Yes	Yes	NFS	No	Yes	Yes

Altre linee guida per il design dello storage sono le seguenti:

Componente	Considerazioni
VMFS	<ul style="list-style-type: none"> > È possibile creare un solo volume VMFS per ciascun LUN. > È possibile utilizzare più VMFS per mantenere divisi gli ambienti di test e di produzione.
RDM	È possibile utilizzare i file RDM con le VM per: 1) i cluster da fisico a virtuale o cluster tra più unità e 2) per creare istantanee hardware dell'array di dischi.
iSCSI	Per ottimizzare le prestazioni e la sicurezza, è consigliabile installare lo storage iSCSI su una rete IP dedicata e isolata.
NAS/NFS	<ul style="list-style-type: none"> > Per ottimizzare le prestazioni e la sicurezza, è consigliabile installare lo storage NAS su una rete IP dedicata e isolata. > Per impostazione predefinita ciascun host ESX supporta il montaggio di 8 NFS fino a un massimo di 64.

Per maggiori informazioni sullo storage in ambiente VMware in generale potete vedere il seguente link che contiene ampia documentazione:

<http://www.vmware.com/technical-resources/virtual-storage/resources.html>

mentre solamente per le SAN è disponibile una guida ufficiale:

http://www.vmware.com/pdf/vsp_4_san_design_deploy.pdf

Per quanto riguarda le schede di rete, il loro numero dipende dalla velocità che supportano (1Gb o 10Gb) e in parte dalle funzionalità previste con la licenza acquistata (Essentials, Ess. Plus,

Standard, ...). Nel caso in cui si opti per schede ethernet a 1Gb, la configurazione ottimale è la seguente:

- 2 porte ethernet per il management del nodo;
- 2 porte per vMotion;
- 2 porte se è previsto l'utilizzo della Fault Tolerance;
- 2 porte se si prevede l'accesso a dischi con NFS o iSCSI;
- almeno 2 porte per il traffico delle VM.

Le porte possono essere suddivise in più schede di rete, ad esempio 1 scheda dual-port e 2 schede quad-port, ed in modo tale che ci sia sempre la ridondanza per ogni punto precedente. In ogni caso, anche per le soluzioni più piccole come ad esempio quelle che prevedono la licenza Essentials (senza vMotion e HA), dovrebbero essere previste sempre almeno 4 porte ethernet suddivise preferibilmente in 2 schede dual-port.

Se invece la rete esistente supporta la tecnologia 10Gb, è possibile anche usare schede ethernet a 10Gb, e in questo caso possono essere sufficienti 2 schede per tutto il traffico.

Infine vediamo in che modo scegliere la tipologia di server (entry level, midrange, enterprise, ...). Se vi state chiedendo come mai il punto più importante debba essere scelto alla fine, il motivo è che il tipo di server può essere scelto solo sapendo quanta RAM e quanta potenza di calcolo sono necessarie. Non ha senso ad esempio progettare subito un server di fascia entry perché non si ha bisogno di eccessiva potenza di calcolo, per poi calcolare la RAM necessaria e vedere che il server scelto ne supporta solo la metà.

Altri punti da tenere in considerazione sono:

- presenza di VM in fault tolerance, nel quale caso ci sarà una seconda VM identica eseguita su un host differente. Dunque devono essere calcolate anche le risorse necessarie a questa seconda VM in fatto di core usati e RAM;
- restrizioni particolari del cluster, ad esempio se alcune VM devono obbligatoriamente restare in esecuzione sempre sullo stesso host (con DRS attivo).

Server per vCenter

Il server che agisce da vCenter può essere implementato in 3 modi, come VM, come macchina fisica, oppure come appliance virtuale (introdotta con vSphere 5). Nei primi due casi vCenter deve essere installato obbligatoriamente su un sistema Windows (2003, 2008, 2008R2, ...) a 64 bit. **Il mio consiglio è di implementarlo come VM qualora la licenza di VMware sia almeno una Essentials Plus (con vMotion ed HA).** I vantaggi di avere vCenter virtuale sono:

- risparmio nell'acquisto di un ulteriore server fisico;
- ripristino in caso di guasto del server, tramite HA;
- backup semplice, in quando il backup può essere fatto parallelamente a quello delle altre VM;
- possibilità di spostare vCenter tra i nodi con vMotion, ad esempio in caso di manutenzione hardware.

Tali vantaggi non si hanno invece se vCenter è implementato in una macchina fisica, infatti si è costretti o a creare un cluster Microsoft (MSCS) per garantire continuità di servizio in caso di crash, o acquistare e configurare vCenter Server Heartbeat. Per ulteriori approfondimenti potete visitare i link seguenti:

- <http://communities.vmware.com/docs/DOC-11115>
- <http://communities.vmware.com/docs/DOC-11197>

Per dimensionare correttamente la macchina fisica che ospiterà vCenter Server può essere usato questo link:

http://kb.vmware.com/selfservice/microsites/search.do?cmd=displayKC&docType=ex&bbid=TSEB_B_1295981518194&url=&stateId=0%200%20184622712&dialogID=184624300&docTypeID=DT_K_B_1_1&externalId=1022101&sliceId=1&rfid=

Tipicamente il database necessario a vCenter viene installato sulla stessa macchina, almeno per gli ambienti piccoli/medi, in modo da fornire la migliore comunicazione possibile. Il database è infatti cruciale, perché basta una perdita di connessione tra esso e vCenter per un blocco di vCenter stesso. Per ambienti enterprise invece, si rende necessario l'utilizzo di un database opportuno situato su un server a parte, diverso da quello su cui si trova vCenter, e configurato in cluster. Nel caso in cui il DBMS venga installato sulla stessa macchina del vCenter è necessario tenere conto anche delle risorse utilizzate dal DBM. Con il database default (Microsoft SQL 2008 Express), i requisiti sono i seguenti:

<http://www.microsoft.com/sqlserver/2005/en/us/system-requirements.aspx>

In questo caso è sufficiente aggiungere 1 GB di RAM a quella prevista (quindi totale 4 GB), ricordando che Microsoft SQL 2008 Express può gestire al massimo 5 host e/o 50 VM. Diversa è la cosa se viene utilizzato un database diverso da quello default, poiché saranno necessarie quasi sicuramente risorse ben superiori.

Bisogna fare particolare attenzione durante il dimensionamento della macchina vCenter nel caso in cui siano installati altri applicativi e/o servizi al suo interno, come Update Manager (che a sua volta ha bisogno di un database diverso da quello di vCenter) o vCenter Converter. In questi casi la memoria minima richiesta può salire a 6 oppure 8 GB. Da ricordare che vCenter non può fare da domain controller nel caso in cui sia utilizzato un Active Directory.

Per quanto riguarda la CPU, gli unici requisiti è che sia a 64 bit, e abbia almeno due core a 2 GHz. Direi che anche in presenza di un database sulla stessa macchina questi valori sono sufficienti, almeno che l'ambiente da gestire non sia di grandi dimensioni. Come linee guida possono essere osservate le seguenti tabelle (fonte VMware).

Up to 50 Hosts and 500 Powered-On Virtual Machines

Product	Cores	Memory	Disk
vCenter Server	2	4GB	5GB
vSphere Client	1	200MB	1.5GB

Up to 300 Hosts and 3000 Powered-On Virtual Machines

Product	Cores	Memory	Disk
vCenter Server	4	8GB	10GB
vSphere Client	1	500MB	1.5GB

Up to 1000 Hosts and 10000 Powered-On Virtual Machines

Product	Cores	Memory	Disk
vCenter Server	8	16GB	10GB
vSphere Client	2	500MB	1.5GB

Infine la terza possibilità è quella di usare un'appliance virtuale che agisce da vCenter. Questa è liberamente distribuita da VMware, e in quanto appliance virtuale il suo funzionamento è simile a quello di una VM. Le sue principali caratteristiche sono le seguenti:

- basata su SUSE Linux Enterprise Server 11 x64;
- distribuito come OVF con relativi dischi VMDK;
- configurato con 2 vCPUs, 8 GB vRAM, LSI Logic Parallel, VMXNET 3, 15GB + 60GB VMDK e VMware Tools;
- include un database (embedded DB2 database) adatto per ambienti di test o per ambienti "piccoli" con meno di 5 nodi e 50 VM (equivalente a Windows vCenter Server + SQL Express);
- supporto di Oracle e IBM DB2 come database esterno, per ambienti medio/grossi.

Rispetto ad avere vCenter su una VM o su un server fisico dedicato, i pro di questa appliance sono:

- l'appliance è una sorta di black box che fornisce le funzionalità di management;
- deploy più rapido in pochi minuti contro gli almeno 30 minuti della versione Windows (spesso sono molti di più tra un aggiornamento e l'altro);
- la parte di WebClient è già inclusa e pronta all'uso (altro tempo risparmiato nel deploy);
- riduce i costi operativi in quanto è "semplice" da aggiornare;
- riduce il TCO perché non richiede licenze Windows aggiuntive (in realtà può non essere un problema se si hanno versioni Datacenter).

Mentre i contro sono:

- richiede più memoria della versione installabile (minimo 8 GB vRAM contro i 4 GB vRAM della versione installabile). Da notare che nelle nuove licenze vSphere 5 c'è un limite nella vRAM totale allocabile;
- Microsoft SQL Server al momento non è supportato (forse in versioni future);
- vCenter Server Linked Mode, Heartbeat, e Update Manager non funzionano;
- non è possibile usare vCenter con ambiente VDI/View;
- il VSA Manager non può essere utilizzato, visto che richiede anch'esso Windows;
- IPv6 non è al momento supportato;
- il modo con cui è licenziato il sistema operativo guest non è al momento chiaro. SuSE Enterprise non è OpenSuSE e in qualche modo andrà pure licenziato. No?!