

EMC²
where information lives

La Deduplicación en Entornos Virtualizados

powered by EMC Iberia TSSO

Bilbao, 9 de octubre de 2008



Agenda

EMC²
where information lives

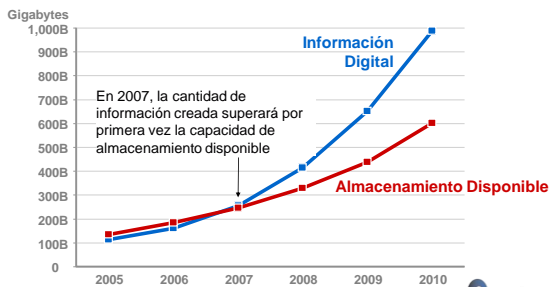
- El Problema
- La Tecnología
- Las Soluciones
- Conclusiones



2007/08 – En la Encrucijada

El mundo va a generar más Información de la que es capaz de almacenar

EMC²
where information lives



Source: IDC, "The Expanding Digital Universe," sponsored by EMC, March 2007

¿Cuántas veces se hace backup del mismo fichero?

llamado Ventas_Final_1998.xls
guardado por Juan Pérez el 24/01/1999...

que se jubiló en el 2001

¿Cuántas copias de ese fichero existen?

¿Cuántas veces se verifica si ha de ser salvado?

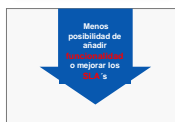
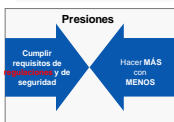
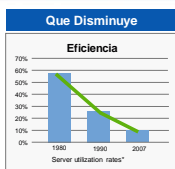
¿Cuántas veces se refresca o recoloca en cinta ?

¿Porqué debería ser restaurado en caso de una contingencia?

¿Cuántos millones de ficheros tienes?



La Realidad TI

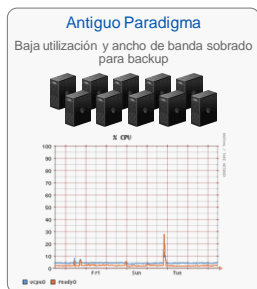


*Source: VMware, November 2007



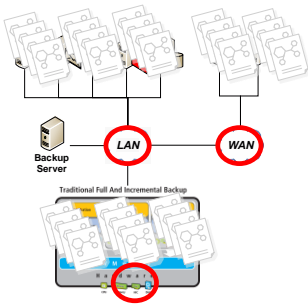
Arquitecturas Virtuales y Backup

La virtualización de servidores no elimina todas las ineficiencias, que son barreras potenciales para una mayor Consolidación y Valor



Problemática Entornos Distribuidos o Virtualizados

EMC²
where information lives



Problemática

Recursos físicos insuficientes

Entornos dimensionados para producción

No existe ventana de backup

Ineficiente y lento

En el Backup el 200% de los datos van por la LAN, WAN o recursos compartidos



La avalancha

EMC²
where information lives



Agenda

EMC²
where information lives

El Problema

La Tecnología

Las Soluciones

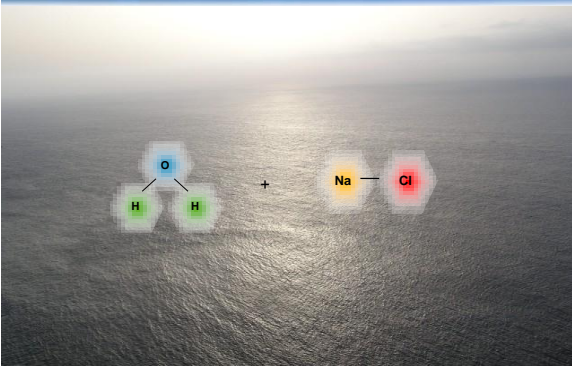
Conclusiones



¿ Como podríamos almacenar esto ?



Deduplicándolo



Los objetivos de la Deduplicación



Tipo de Dato	Datos Respaldados	Datos Movidos	Ratio de Deduplicación
Windows file systems	3,573 GB	6.1 GB	586:1
Mezcla Windows, Linux, y UNIX file systems	5,097 GB	11.7 GB	436:1
Ficheros eng. en NAS (NDMP backups)	3,265 GB	24.2 GB	135:1
Mezcla 20% databases, 80% file systems (Windows y UNIX)	9,583 GB	80.0 GB	120:1
Mezcla de Linux file systems y databases	7,831 GB	104.2 GB	75:1



¿Como funciona la deduplicación?



DeDuplicación Global de Datos

- El agente escanea los datos buscando cambios
- Segmentación Variable de los archivos modificados
- Sólo los segmentos únicos son protegidos



Busca en una cache local de archivos para identificar los datos



Comprime y aplica un hash a cada objeto

Busca en una cache local de hashes y en el Servidor central de Avamar para determinar si el objeto es único



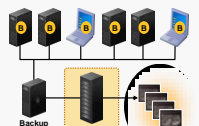
Avamar Data Store o Avamar Server

De-duplicación en Origen o Destino



Alternativas de Implementación en el Mercado

DEDUPLICACIÓN en Destino



Reducción de hasta 50x en espacio
Mueve el 200% de los datos semanalmente
 Los backups se recuperan desde Full + Incrementales
 Los datos son vistos como file system o librerías virtuales para el entorno de backup

DEDUPLICACIÓN en Origen



Reducción de hasta 50x en espacio
Mueve el ~2% de los datos semanalmente
 Hasta 500x de reducción en líneas de comunicaciones
 Backup más rápidos - Hasta 10x en backups full diarios
 Todos los backups son Full – Recuperación en un solo paso

Todas las Alternativas



EMC Avamar

Software de backup con deduplicación en origen
 Complementa o sustituye al software de Backup
 Visión estratégica de la deduplicación a largo plazo

- Ahorro de espacio y comunicaciones
- Accelera backups hasta 10X or more
- Regiones de deduplicated data across WAN for DR

Recuperación en un solo paso (No más incrementales)
 Oportunidad de introducir un elemento software en clientes competencia de Networker

- Funcionalidad no existente
- Los ahorros pueden justificar ambos productos o incluso la sustitución

Disk Library Family – con DL3D

Solución de VTL, con deduplicación en destino
 Se integra con el software de backup existente
 Visión resultadista de la deduplicación a corto plazo

- Se integra fácilmente
- Ahorro de comunicaciones en DR

Solución Flexible para todo tipo de clientes
 Soluciones de alto rendimiento y disponibilidad
 Oportunidad de introducir un elemento Software con mínimo impacto en el cliente desde el punto de vista operacional



Provides NetWorker installed base up-sell opportunity with integrated de-dupe solutions



Avamar Server based de-duplication

Agenda



El Problema

La Tecnología

Las Soluciones

Conclusiones

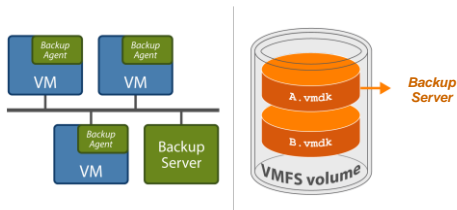


Alternativas de Protección



Guest Level Backup (Aprox. Tradicional)

Image Level Backup (Nueva Aproximación)



Soluciones Avamar



Alternativas de Implementación del Software

Solución Software

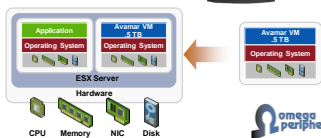
- Avamar de-duplication software
- Servidor Compatible
- Almacenamiento interno o CLARiON

Solución Completa Preconfigurada – Data Store

- Avamar de-duplication software + EMC-certified hardware
- Simplifica la solución
 - Adquisición (Un fabricante, Un Hardware certificado)
 - Implementación (Reduce tareas onsite)
 - Soporte (Punto único de contacto)

Solución Virtualizable – AVE (Avamar Virtual Edition)

- Avamar SW + VMWare
- Único en el mercado
- Acelerar el ROI



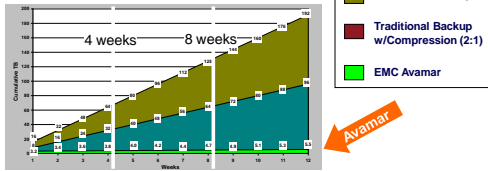
Beneficios de la De-duplicación



Optimización de Espacio

Traditional Backup v. EMC Avamar

Cumulative Media Required



Avamar rentabiliza el backup a disco

Total Cumulative Storage - 10TB MS Office file data
(Traditional 2 weekly fulls + daily incrementals, EMC Avamar 1 daily full)



25

Adiós a las cintas??



Eficiente en Coste

- Deduplicación Global → Ahorro de espacio y comunicaciones
 - Escalabilidad lineal, sin perder la Deduplicación Global
- Algoritmo flexible de bloque variable → Ahorro de espacio y comunicaciones
- Deduplicación en Origen → Ahorro de espacio y comunicaciones
- Múltiples topologías para ajustar el coste-rendimiento

Plataforma Segura

- Arquitectura redundante para DR
- Redundancia de nodos (RAIN)
- Redundancia en disco (RAID)
- Cifrado de datos en línea y en destino

Eficiencia de Gestión

- Administración centralizada
- Todos los backups son "full"
- Capacidad de búsqueda dentro cada fichero
- Posibilidad de acceso a cada backup mediante protocolos NAS



26

Avamar: Optimized Backup for VMware Infrastructure



Avamar De-duplicates at the Optimal Location and Level of Granularity

Optimal location: Reduces data at the source

- Duplicate data never traverses congested shared resources
- Data backed up is reduced from ~200 percent to ~2 percent of primary weekly
- Significantly reduces contention for shared resources
- De-duplicates within and across virtual machines and/or VMDK (virtual machine disk format) files

Optimal granularity: Sub-file, variable-length segments

- VMDK is one large file; any changes cause an incremental/full traditional backup
- Fixed-length segment de-duplication fails due to frame offset
- Sub-file, variable-length segment de-duplication finds changes anywhere in VMDK
- Dramatically cuts the amount of data backed up daily
- Significantly reduces daily backup time

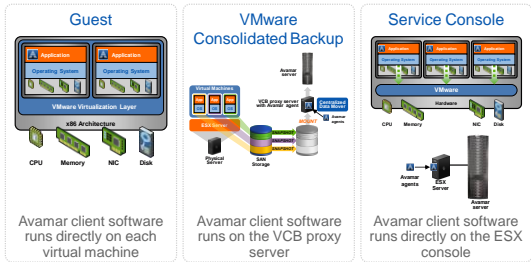
Ideal for the protection of VMware environments

- Enables fast, secure backups over existing virtual infrastructure
- Permits greater server consolidation and maximum value from VMware



27

Avamar Client Backup Solutions



Avamar Backup for VMware Guests

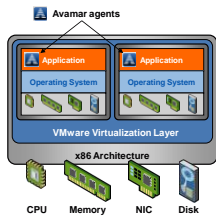


Avamar agent resides inside each virtual machine

De-duplicates data within the virtual machine, as if they were physical servers

Moves minimal backup data
 - Reduces resource contention and accelerates backups

Provides file-level restore for Windows, Linux, and Solaris



Optimized Backup for VMware Infrastructure



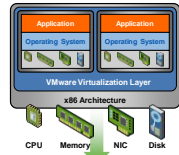
Avamar Efficiently Protects Virtual Machines

- Up to 95 percent reduction in data moved
- Up to 90 percent reduction in backup times
- Up to 50 percent reduction in disk impact
- Up to 95 percent reduction in NIC usage
- Up to 80 percent reduction in CPU usage
- Up to 50 percent reduction in memory usage

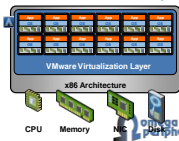
All backups are stored as "virtual full backups," ready for immediate restore

Maintains effective consolidation ratios without overtaxing CPU utilization

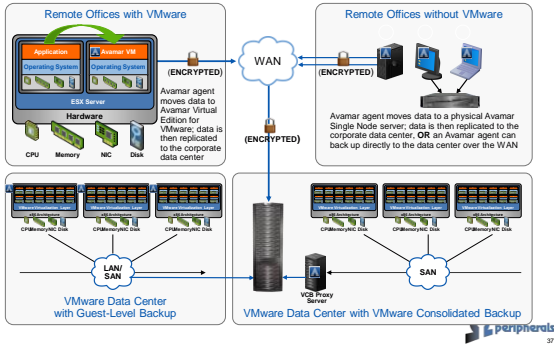
Traditional moves ~200% weekly



Avamar moves ~2% weekly



Avamar Products for VMware Infrastructure



Avamar Supported Operating Systems and Applications



Client Operating Systems Supported

- Microsoft Windows Server 2003 Standard and Enterprise
- Microsoft Windows 2000 Server and Advanced Server
- Microsoft Windows NT 4.0 Server
- Microsoft Windows XP, XP Professional, Vista
- Red Hat Linux 8.0, 9.0
- Red Hat Enterprise Linux (RHEL) 2.1, 3.0, 4.0, 5.0
- Solaris 2.6, 7, 8, 9, 10
- SUSE Linux Enterprise Server 8, 9, 10
- IBM AIX 5.1, 5.2, 5.3, 6.1
- HP-UX 11.0, 11V1, 11V2, 11V3
- Mac OS X 10.4x, 10.5x
- NetWare 6.5
- Novell Storage Services (NSS) OES 2

New Avamar software expands client operating system, application support

Application Modules

- Microsoft Exchange 5.5, 2000, 2003
- Microsoft SQL Server 7.0, 2000, 2005
- Oracle 9i, 10g, 10gR2
- IBM DB2 8.2.x, 9.5
- NDMP (EMC Celerra) DART 5.5
- NDMP (NetApp) Data ONTAP 6.3, 6.4, 6.5, 7.0.4, 7.0.5, 7.0.6, 7.1x

VMware Infrastructure

- VMware ESX Server versions 2.5.3, 3.0.x, 3.5, 3i
- ESX client OS-certified using RHEL 4 client (support within virtual machines depends on VMware support of client operating system)