

---

12

*Análisis diseño e implementación de un sistema de respaldo de  
datos y restauración de la información basado en Software  
AVAMAR.*

---

*Chango Wilson, Sayago Jaime, Orozco Fausto, Jácome Ligia.*

*Recibido: octubre 2017  
Aprobado: diciembre 2017*

---

# **Análisis diseño e implementación de un sistema de respaldo de datos y restauración de la información basado en Software AVAMAR.**

## **Analysis design and implementation of a data backup system and restoration of information based on AVAMAR Software.**

Chango Wilson<sup>1</sup>, Sayago Jaime<sup>2</sup>, Orozco Fausto Ernesto<sup>3</sup>, Jácome Ligia<sup>4</sup>;

PUCE-SE, Ecuador, wilson.chango@pucese.edu.ec<sup>1</sup>,

PUCE-SE, Ecuador, Jaime.sayago@pucese.edu.ec<sup>2</sup>,

PUCES-SD, Ecuador, oife@pucesd.edu.ec<sup>3</sup>,

UTI, Ecuador, ligiajacome@uti.edu.ec<sup>4</sup>,

### **Resumen**

*El objetivo de esta investigación está en realizar un sistema para el respaldo de la información de ambientes virtuales y de computadores de usuarios finales a través de la red LAN, se cubrió, para este efecto, el análisis, diseño e implementación a nivel de laboratorio, de un sistema de respaldos especializado para dichos ambientes, de esta manera se comprobó así sus prestaciones y su factibilidad de ejecución, desde un punto de vista técnico. AVAMAR, surge como una alternativa de solución para enfrentar los retos de aquellos ambientes de TI en donde los respaldos y recuperación de información demandan ventanas de tiempo menores, respuestas más ágiles de recuperación, respaldo a usuarios finales y máquinas virtuales. La arquitectura y tecnología de esta solución de respaldo y recuperación hace ideal para un estudio referente en el que se muestra todo lo que se puede lograr en el complejo y variado mundo del respaldo de la información. El escenario que se planteó trata de una infraestructura típica que consta de servidores físicos, virtuales, una red LAN y una red SAN, usuarios de computadores de escritorio y laptops; los usuarios finales tenían conectividad hacia los servidores a través de la red LAN, estos equipos de cómputo tenían Windows como sistema operativo, se estableció un ambiente virtual que se ejecutó sobre los servidores físicos, El ambiente planteado fue totalmente operativo y funcional, es decir todos los servicios estuvieron activos y los usuarios podían acceder a los mismos desde la red LAN. A partir de este escenario, se analizó las prestaciones del sistema de respaldos de la información de los usuarios finales y máquinas virtuales, optimizando los recursos como son hardware, software y tiempo de ejecución de las tareas de respaldo y recuperación.*

**Palabras clave:** Resguardo de la información, Recuperación de información, Ambiente virtual, AVAMAR.

### **Abstract**

*The objective of this research is to make a system for the backup of the information of virtual environments and computers of end users through the LAN network, covering, for this purpose, the analysis, design and implementation at the laboratory level, of a system of specialized backups for these environments in this way was thus checked for its performance and its feasibility of execution, from a technical point of view. AVAMAR, emerges as a solution alternative to face the challenges of those IT environments where backups and retrieval of information demand shorter time windows, more agile recovery responses, end user support and virtual machines. The architecture and technology of this backup and recovery solution makes it ideal for a reference study that shows everything that can be achieved in the complex and varied world of information backup. The scenario that was raised concerns a typical infrastructure consisting of physical, virtual servers, a LAN and a SAN network, users of desktop computers and laptops, end users had connectivity to the servers through the LAN, these computer equipment had Windows as an operating system, a virtual environment was established that was executed on the physical servers, the environment was fully operational and functional, all services were active and users could access them from the LAN network. From this scenario, we analyzed the benefits of the system of backups of the information of end users and virtual machines, optimizing resources such as hardware, software and execution time of backup and recovery tasks.*

**Keywords:** Information protection, Information retrieval, Virtual environment. AVAMAR

## 1. INTRODUCCIÓN

El ambiente de TI de las empresas está cambiando y evolucionando cada día. El resguardo de la información que genera una empresa siempre ha sido una tarea importante para el equipo de TI. En sus inicios esta tarea, era muy rústica y tenía muy poco nivel de automatización y confiabilidad. Copias de información a una unidad de disco externo, memoria USB o cinta magnética han sido las soluciones típicas implementadas ante la imperiosa necesidad de respaldar la información. Sin embargo, este tipo de soluciones se convierten por sí mismas en un problema para las empresas modernas por sus limitaciones y escasas garantías al momento de la recuperación, y no son aplicables para empresas grandes las cuales incluyen cientos de máquinas virtuales y máquinas de usuarios finales que necesitan ser respaldadas.

Una posible solución es el de disponer de una red completamente independiente de la de producción para el ambiente de respaldo y restauración de la información. Sin embargo, no todas las empresas pueden darse este lujo ya que implicaría destinar recursos de hardware como tarjetas NICs, concentradores, etc., dedicados a la red de respaldo; resulta entonces necesario estudiar esta problemática de los ambientes actuales y encontrar una alternativa de solución que contemple todos los factores críticos indicados y provea mecanismos que permitan realizar las tareas de respaldo y recuperación de una manera eficiente y optimizando el consumo de recursos de la infraestructura productiva de tal manera que cause el menor impacto posible.

AVAMAR facilita la integración con ambientes virtuales como VMware e Hyperv, la incorporación de módulos especializados para respaldo de máquinas virtuales y de usuarios finales, entre otras características, hace la herramienta ideal para el estudio de los sistemas de respaldos en estos escenarios particulares.

El escenario de análisis planteado, corresponde a un escenario típico de una empresa moderna, el cual cuenta con servidores ejecutándose sobre la plataforma virtual VMware[1] y varios usuarios de esos servicios, trabajando desde la red LAN. El estudio consistió en analizar las soluciones de respaldos más apropiada para este ambiente típico, determinando primero las necesidades del ambiente virtual y de usuarios finales, diseñando y modelando la herramienta y finalmente implementándola en un ambiente de laboratorio[2].

Se debe anotar que los ambientes virtuales VMware proveen características inherentes que deben ser aprovechadas por terceras herramientas para la optimización del respaldo, por ejemplo la capacidad conocida como CBT (Changed Block Tracking), de VMware, cuya función es la de reconocer únicamente los bloques de información que cambian de una máquina virtual, la compresión de la información durante las tareas de respaldo es crítica para optimizar el consumo de ancho de banda que una herramienta de respaldos consumirá en su operación[3].

Ya Chango[4] presentó un análisis para el respaldo de datos de máquinas virtuales y usuario final a través de una red LAN y SAN

AVAMAR, surge como una alternativa de solución para enfrentar los retos de aquellos ambientes de TI en donde los respaldos y recuperación de información demandan ventanas de tiempo menores[5], respuestas más ágiles de recuperación, consistencia de backups, respaldo a usuarios finales y máquinas virtuales, además, de que esta herramienta aprovecha esos puntos de integración de los hipervisores de los ambientes virtuales y ofrece módulos especializados para el respaldo en línea de máquinas de usuario final [6].

## 2. SISTEMA DE RESPALDOS

La cantidad de información que maneja una empresa está creciendo a pasos inmensos y la necesidad de resguardar la misma también. Existen diversas fuentes de información como servidores de aplicaciones, los usuarios, máquinas virtuales, etc. Los respaldos, o backups son utilizados básicamente por tres motivos: recuperación de desastres, restauraciones operacionales y almacenamiento de información a largo tiempo. La recuperación de desastres está pensada para recuperar toda, o gran parte de la información, de la infraestructura de TI en el caso de un desastre natural o un evento de daño mayor como un incendio, inundación, etc. La recuperación normalmente se la realiza en un sitio alternativo o de contingencia que tiene una infraestructura igual o similar al sitio principal y que está lista para levantar el ambiente de producción a partir de los respaldos.

El respaldo operacional es una copia de la información en un punto de tiempo determinado con el propósito de recuperar la data ante un evento de corrupción lógica de la misma que puede ocurrir durante la operación normal de una aplicación, servidor u otra fuente de información. La mayor incidencia de pérdida de datos en una organización corresponde a este tipo. La operación de respaldo en

una empresa es útil para tener una copia de la información disponible para una eventual recuperación en el caso de daños físicos de los almacenamientos, servidores, computadores, etc.

### 3. TIPOS DE RECUPERACIÓN

Existen recuperaciones a nivel de archivo o carpeta (es la más común), otras recuperaciones que tienen que ver con aplicaciones de correo electrónico, base de datos, etc. Y restauraciones de sistemas completos, la cual es aplicada en muy pocas ocasiones.

La recuperación de archivos individuales o directorios, es la más común, para este tipo de recuperación es importante saber las características de la aplicación a respaldar y la naturaleza de la información. Es clave conocer cuál será la periodicidad con la que se realizará el respaldo, por cuánto tiempo deberá estar disponible el mismo, etc.

De ello dependerá el dimensionamiento de la solución de respaldos. Otro tipo de restauración es aquella que se implica la recuperación en caso de desastres. Este tipo de recuperación, implica restaurar toda o parte de la información de la infraestructura de TI del sitio principal en un sitio alternativo. Para esto es necesario que la información respaldada esté físicamente en un sitio alternativo, y para este propósito existen algunas alternativas como son el traslado de medio físicos (cintas, discos externos, etc.) hacia el sitio alternativo o alternativas más sofisticadas como replicación a nivel de respaldo a través de la red WAN por el protocolo IP.

Por último, un método de recuperación denominado Bare Metal Recovery (BMR) el cual contempla la recuperación total de un servidor de producción en otro hardware. Este tipo de recuperación implica que toda la metadata, las configuraciones del sistema, de las aplicaciones, etc., serán restauradas a su estado original en un hardware de similares características que el servidor inicial.

Esta investigación se concentró en tareas de recuperación de archivos, carpetas, máquinas virtuales, etc., corruptos. Es decir, se contempló el escenario de recuperación más común que es el de usuarios finales y ambientes virtuales.

### 4. RESPALDO DE LOS AMBIENTES VIRTUALES Y USUARIO FINAL

El mercado de las telecomunicaciones ha cambiado mucho en los últimos años. La mayoría de empresas han adoptado o están en el proceso de adoptar la política de consolidación de recursos de tecnología en su Data Center. Los centros de datos

tradicionales ocupaban un servidor dedicado para cada una de las aplicaciones del negocio, lo que normalmente daba como resultado la sub-utilización de recursos de memoria, CPU, operaciones de I/O, etc.; y un incremento en costos a medida que la empresa escalaba.

La principal ventaja que brinda los ambientes virtuales permite optimizar el uso de recursos físicos al distribuirlos de manera óptima entre varios servidores virtuales (máquinas virtuales o VM). Es decir, sobre un mismo servidor físico pueden existir, ejecutándose decenas o cientos de máquinas virtuales, las cuales comparten los mismos recursos. Si bien es cierto, este esquema de consolidación tiene muchas ventajas a nivel técnico y económico, como consolidación de los ambientes heterogéneos, único punto de control y visibilidad de recursos, optimización del uso del hardware, reducción de costos, escalabilidad, etc., el momento de hablar de respaldo y recuperación de esa información el escenario es un poco más complicado.

Los esquemas de respaldo y recuperación en ambientes virtuales tienen el problema que para su operación deben utilizar los mismos recursos físicos que el ambiente productivo. Es más grave aún el problema, ya que esos recursos físicos son compartidos por varias máquinas virtuales. El impacto por consumo excesivo de recursos para una operación de respaldo para un solo cliente implica la degradación del performance de todas las máquinas virtuales que se ejecutan en el mismo servidor físico.

En el mundo “físico” no existe tal problema, ya que los recursos pertenecen a una sola aplicación y normalmente están sub-utilizados. Es por esto que una solución de respaldos especializada en un ambiente de servidores físicos normalmente no es la más apropiada para ambientes virtuales ya que no toma en cuenta las variables anteriormente indicadas.

De este particular, las empresas están muy conscientes, y es por ello que invierten cada vez más esfuerzos para mejorar sus esquemas de respaldo y recuperación del ambiente virtual[7].

### 5. ARQUITECTURA DE RESPALDO Y RECUPERACIÓN AVAMAR.

AVAMAR es una solución de respaldos y recuperación cliente servidor diseñada para trabajar en ambientes diversos, como son los ambientes virtuales, usuarios finales, servidores físicos, etc.

Esta arquitectura incluye tanto hardware como software, y puede escalar de manera vertical para acoplarse a las necesidades de empresas pequeñas, medianas y grandes. La clave de su arquitectura es el proceso de compresión de la información a través del análisis en línea y en el origen de la data antes de ser respaldada.

Este análisis le permite identificar información redundante y solo guardar aquella información nueva, ahorrando así recursos de almacenamiento, CPU en el host y ancho de banda. Este proceso de compresión, puede explicarse si se pone atención a la siguiente figura [8].

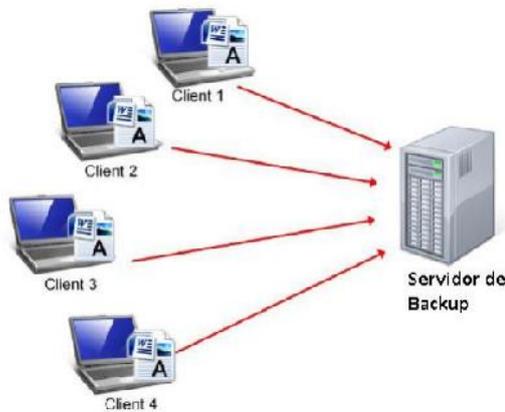


Figura 1. Proceso de respaldo en un ambiente tradicional

Fuente: Elaboración Propia basada en el caso de estudio

Fig1. Se muestra el proceso de respaldo en un ambiente tradicional cliente servidor. Todos los clientes tienen exactamente el mismo documento para respaldar, un sistema de respaldo tradicional guardaría 4 copias exactas del documento, una por cada cliente, sin notar que se trata de la misma información, consumiendo espacio innecesariamente.

## Materiales y métodos

La metodología utilizada en el análisis del caso de estudio fue basada en un enfoque cuantitativo con el cual primero se llegó a la elaboración de una matriz en la que se establece los requerimientos básicos para la implementación de la solución y los pasos posteriores para realizar las pruebas.

**Tabla 1**

*Información de los Usuarios Finales y sus requerimientos de espacio para respaldar*

Tipo usuario	Cantidad usuarios	Información Usuario	Crecimiento anual	Espacio total	Espacio año
		GB	%	GB	GB
Oro	10	50	5	500	525
Plata	10	25	5	250	263
Bronce	20	10	2	200	204
Total	40	85	12	950	992

**Nota.** El resultado anual de capacidad para almacenamiento de información para los usuarios tipo Oro, Plata y Bronce es de 992 GB. Fuente: Elaboración propia basada en la configuración de los tipos de usuarios para respaldar la información.

Según la tabla 1, en el ambiente estudiado, existen tres tipos de usuarios:

Oro con 10 usuarios con mayor jerarquía que tiene más información que respaldar, en promedio 50GB por usuario, con un crecimiento anual de esa data del 5%, y un total de 500GB de espacio actual para respaldar esa información. La cantidad de información generada proyectada a 1 año según la tabla sería de 525GB.

Plata: son 10 usuarios con jerarquía intermedia que tiene información que respaldar, en promedio 25GB por usuario, con un crecimiento anual de esa data del 5%, y un total de 250GB de espacio actual para respaldar esa información. La cantidad de información generada proyectada a 1 año según la tabla sería de 263GB.

Usuarios Bronce: son 40 usuarios con jerarquía baja que tiene información que respaldar, en promedio 10GB por usuario, con un crecimiento anual de esa data del 2%, y un total de 200GB de espacio actual para respaldar esa información. La cantidad de información generada proyectada a 1 año según la tabla sería de 204GB.

En cuanto al espacio requerido para el Backup de los servidores virtuales se establece la siguiente tabla.

**Tabla 2**

*Información de las Máquinas Virtuales y sus requerimientos de espacio para respaldar*

Tipo	Sistema	Información	Crecimiento	Espacio	Espacio
------	---------	-------------	-------------	---------	---------

Servidor	Operativo	actual [GB]	Anual [%]	Total [GB]	año [GB]
Proxy Server	Linux	60	5	60	63
Directorio Activo	Windows	100	8	100	108
Base de Datos	SQL 2008	120	10	120	132
Total		280	23	280	303

**Nota.** El resultado anual de capacidad de almacenamiento para los servidores virtuales en de 303 GB al año. Fuente: Elaboración propia basada en la configuración de los servidores virtuales.

Según la tabla 2, Proxy Server: se cuenta con un servidor Linux con 60GB de espacio actual, se prevé un crecimiento a un año del 5%, lo que implica al final de un año un requerimiento de 63GB para este servicio Directorio Activo: se cuenta con un servidor Windows 2012 con 100GB de espacio actual, se prevé un crecimiento a un año del 8%, lo que implica al final de un año un requerimiento de 108GB para este servicio Base de Datos: se cuenta con un servidor SQL 2008 con 120GB de espacio actual, se prevé un crecimiento a un año del 10%, lo que implica al final de un año un requerimiento de 132GB para este servicio, para el ambiente virtual entonces, la solución de respaldos debe contemplar un requerimiento total actual de 280GB y futuro a 1 año de 303GB.

Existen otros factores a tomar en cuenta que afectan el dimensionamiento de la solución final, tales como las políticas de respaldo definidas, períodos de retención, periodicidad de los respaldos.

### Investigación Bibliográfica.

Se utilizó para profundizar teorías sobre respaldos de información, se revisó artículos técnicos sobre visualización de servidores, además de la obtención de información a través de internet para las configuraciones de los servidores virtuales y de las máquinas de escritorio[9]

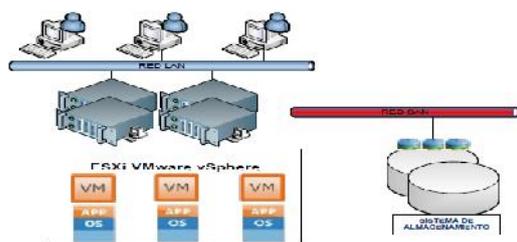


Figura 2. Diagrama del Escenario Actual del Caso de Estudio Fuente: Elaboración Propia basada en el caso de estudio

Fig2. Se puede apreciar los componentes físicos y lógicos que constituyen el escenario de estudio del presente trabajo. Se trata de una infraestructura típica que consta de los siguientes elementos: dos servidores de físicos Dell PowerEdge R720 con 4 CPUs Intel Xeon E5530 x 2.40 GHz y 12GB en RAM.

El ambiente virtual de las empresas actuales cada vez se vuelve más crítico ya que la tendencia es la migración hacia los ambientes virtuales de aquellos servicios que normalmente tenían destinado un host físico con recursos dedicados[10].

Incluso aplicaciones altamente transaccionales como las bases de datos han demostrado tener un performance adecuado en los ambientes virtuales, por lo que la tendencia está marcada y será más fuerte en los años venideros[11].

De igual manera, los usuarios de una empresa generan grandes volúmenes de información a diario[12]. Es común encontrar en los computadores de los usuarios finales información crítica como documentos de ofimática, correos electrónicos, archivos de multimedia, etc., que forman parte de ese activo intangible de la empresa tan importante[13].

El escenario planteado está considerado a nivel de laboratorio, cuenta con todos los componentes de un escenario real pero en un ambiente controlado, la infraestructura planteada es completa, los usuarios finales cuentan con conectividad hacia los servidores a través de la red LAN, éstos cuentan con Windows como su sistema operativo, existe un ambiente virtual ejecutándose sobre los servidores físicos con VMware vSphere 5, este permite tener, para el escenario actual, tres máquinas virtuales que actúan como servidores para los usuarios finales, éstas son: servidor con Linux para servicio de proxy, servidor con Windows 2012 para servicio de Directorio Activo, servidor con SQL para servicio de Base de Datos[14].

Los espacios de almacenamiento son presentados desde el storage VNX 5200 hacia los servidores a través de la red SAN implementada como LUNs (Logical Unit Number)[15]. VMware vSphere se encarga de convertir esas LUNs en Data Stores para los servidores ESXi, desde los cuales se aprovisionan los espacios en disco requeridos por las tres máquinas virtuales, el ambiente es totalmente operativo y funcional, es decir todos los

servicios están activos y los usuarios pueden acceder a los mismos desde la LAN[16].

Punto relevante es que todos los usuarios manejan el sistema operativo Windows 7 de 64 bits. Esto es importante tomarlo en cuenta por compatibilidad con la herramienta de respaldos a ser considerada[17].

## 6. PRUEBAS REALIZADAS

### Primera Prueba

El primer backups del dominio ORO, tuvo una duración de 1 minuto con 8 segundos para respaldar un total de 14.5MB. Podemos destacar que la herramienta casi no encontró información duplicada, esto se debe a que fue el primer respaldo y AVAMAR no tenía información previa contra la cual comparar la nueva y tratar de encontrar duplicidad[18].

El primer backup del dominio ORO, tuvo una duración de 1 minuto con 8 segundos para respaldar un total de 14.5MB. Podemos destacar que la herramienta casi no encontró información duplicada, esto se debe a que fue el primer respaldo y AVAMAR no tenía información previa contra la cual comparar la nueva y tratar de encontrar duplicidad.

### Segunda Prueba

Ahora bien, se realizó un segundo full Backup del mismo grupo sin alterar la información de la carpeta C:\Mis Backups, esta operación tardo apenas 10 segundos a pesar de que la carpeta a respaldar tiene los mismos 14.5MB de información, ésta indica 0%, lo que quiere decir que AVAMAR no identificó información nueva y no pasó un solo byte de información a través de la red

**Tabla 3**

*Resumen de Resultados de los dos primeros Backups*

DOMINI	TIEMPO BACKUP		B. RESPALDAR		B. NO DUPLICADOS	
	1 Backups	Backu p	1 Backup	2 Backup	1 Backu p	2 Backu p
ORO	1m:8s	0m:10s	14.5M B	14.5M B	99.7%	0%
PLATA	0m:57s	0m:10s	39.4M B	39.4M B	71.7%	0%
BRONCE	1m:45s	0m:10s	21MB	21MB	99.1%	0%

**Nota.** El resultado de Bytes reales no duplicados de cada grupo de usuario es 0%. Fuente: Elaboración propia

basada en los resultados de los backups de los grupos de usuarios.

Se puede notar la tendencia de este tipo de herramienta de backups: a medida que más información se tenga respaldada más posibilidad de encontrar data duplicada existe. En el segundo backup la herramienta, al tener ya la información del primer respaldo, fue capaz de identificar duplicidad en toda la información por lo tanto los Bytes Reales guardados en el segundo backup son 0. De igual manera cabe destacar que para el caso del primer respaldo del dominio PLATA se encuentra una importante tasa de compresión, ya que la herramienta solo guardo 71% de la información leída. Esto es porque ya se tenía los respaldos previos de los dominios ORO y BRONCE.

### Tercera Prueba

En esta tercera prueba se añadió mucha más información a la carpeta C:\Mis Backups de uno de los usuarios en el dominio ORO. Exactamente esa misma información (archivos, carpetas) fue añadida a otro usuario del dominio PLATA.

**Tabla 4**

*Resumen de Resultados de los dos primeros Backups*

Sistema Almacenamiento	I. Respalda	I. Respalda	T. Respalda
	[MB]	[MB]	[MB]
Avamar	820	2,46	23
Tradicional	820	820	66

**Nota.** Resultados de los dos primeros backups.

En una solución tradicional, la información total respaldada sería, para este ejercicio, de 820MB sobre 2.46MB que respaldó AVAMAR. Es decir que el sistema fue aproximadamente 300 veces más eficiente un sistema tradicional.

### Pruebas de Restauraciones de Información de Usuario Final

Ponemos a prueba las operaciones de restauración de la información del usuario final a partir de los respaldos obtenidos se prueba la capacidad de recuperación granular de la herramienta, es decir recuperar un archivo o carpeta particular sin necesidad de restaurar todo el Backups.

#### Prueba de Restauración 1.

En esta prueba se simula la pérdida de un archivo en la carpeta C:\Mis Backups, de un usuario final en el dominio ORO. Se borra intencionalmente el

archivo “Velocity\_provider\_playbook.pdf” cuyo tamaño es de 27MB.

El proceso de restauración, en esta ocasión se la realizó desde la herramienta de administración de AVAMAR, pero también la puede generar el usuario final siempre que éste tenga los accesos habilitados [19].



Figura 4. Recuperación de un único Archivo

Fuente: Elaboración Propia basada en la recuperación de un único archivo.

Se puede apreciar que la recuperación puede ser tan granular que incluso llega a nivel de archivo. Desde la consola de administración se puede explorar un backup específico y navegar por las carpetas respaldadas hasta encontrar el archivo específico a respaldar.

### Prueba de Restauración 2.

Para esta prueba se simulará haber perdido todo el contenido de la carpeta C:\Mis Backups. Para este ejercicio la carpeta tiene un tamaño de 800MB aproximadamente. Se medirá la rapidez del respaldo en este escenario.

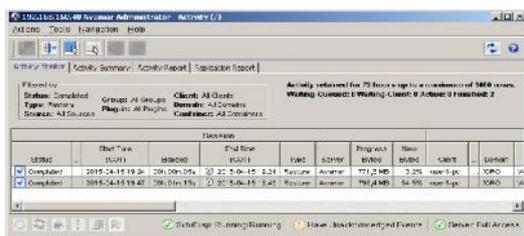


Figura 5. Tarea de Recuperación de toda la carpeta de respaldos  
Fuente: Elaboración Propia basada en la recuperación de toda la carpeta.

La recuperación de toda la carpeta de backups tardó 1 minuto con 13 segundos. Considerando que son 800MB es un tiempo bastante bueno, implica que el consumo de ancho de banda para esta operación fue de 88Mbps.

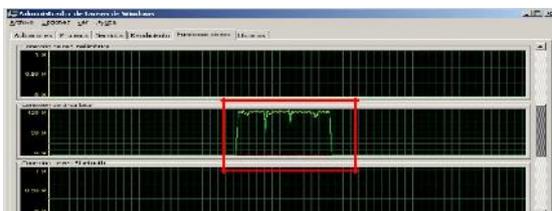


Figura 6. Consumo de Ancho de Banda durante la Operación de Recuperación

Fuente: Elaboración Propia basada en el consumo de ancho de banda durante la recuperación de la información.

Fig8, para esta prueba se midió también el comportamiento de la tarjeta de red NIC. Comprobando que esta tarea de restauración llegó a consumir casi el 100% del ancho de banda disponible en la interfaz.

Para casos de recuperación éste puede ser un comportamiento admisible. Pero para los backups a veces es necesario restringir el ancho de banda que una operación de respaldo puede utilizar. Esto es para resguardar recursos para las demás aplicaciones críticas del cliente que usan la red como medio de comunicación. Esta prueba se la realizará más adelante.

Se observa también, que, para los procesos de recuperación, la compresión y la capacidad de hallar data duplicada no es un factor que entra en juego ya que la información debe llegar al cliente con cada uno de sus bytes, por lo que no se espera obtener los tiempos tan bajos que se obtuvieron durante los procesos de respaldo.

### 3. 2 Pruebas sobre el Modulo de Respaldo de Máquinas Virtuales

Para realizar las pruebas de las Herramienta de respaldo de máquinas virtuales se considera un ambiente heterogéneo que se compone de máquinas con Sistemas Operativos Windows y Linux. En estos equipos se ejecuta deliberadamente acciones que degradaron el funcionamiento de los equipos dejándolos inoperativos. A partir de ahí se realiza la recuperación de las máquinas virtuales mediante la integración de la herramienta de respaldo al Hypervisor de VMware para la recuperación de las mismas. En la figura siguiente se explica el escenario de pruebas que se ha utilizado.

#### Escenario 1.

En el Escenario 1 se ejecutó un respaldo inicial de las máquinas virtuales e inmediatamente se ejecuta un segundo respaldo para observar los tiempos que le toma a la herramienta completar la tarea y la optimización del tiempo de respaldo mediante la tecnología de duplicación.

#### Escenario 2.

En el Escenario 1 se trabajó con una máquina virtual basada en Linux que sufrió un fallo que la

dejó inoperativa y con daños a nivel de los bloques de archivos del Sistema operativo. Posterior al fallo del equipo se ejecutó una recuperación de la máquina a un estado previo con el Backup existente. Se realizó pruebas de consistencia y funcionamiento del equipo virtual.

#### **Escenario 3.**

En el Escenario 2 se ha trabajado sobre un servidor Windows que ha sufrido un daño irreversible y es necesaria su restauración desde un respaldo de la Herramienta de RespalDOS. Para este ejercicio se realizará una eliminación completa de la máquina en VCenter y posterior restauración.

#### **Escenario 4.**

En el Escenario 4 se tiene una base de Datos SQL que tiene su información albergada en la unidad D, la cual ha sufrido un fallo y se ha perdido información de la misma. En este escenario se busca recuperar la unidad D y realizar pruebas sobre la Base de Datos para verificar que la información se encuentra sin cambios al estado previo a la pérdida de datos.

#### **Resultados Pruebas sobre el Escenario 1.**

Posterior a la configuración de la herramienta se realiza un respaldo inicial de las máquinas virtuales en las cuales se buscó determinar la eficiencia de la Herramienta de Backup con su tecnología de compresión una vez que se realizan respaldos posteriores. En las siguientes figuras se puede observar como en el segundo respaldo la nueva información que ingresa al equipo es bastante menor a la que se ingresó en el primer respaldo

#### **Resultados Pruebas sobre el Escenario 2**

La primera acción que se ejecutó sobre la máquina virtual consistió en ejecutar el comando indicado en la figura que dejó inoperativa la máquina. Ante este escenario se procedió a forzar el apagado de la máquina y se procedió a ejecutar una restauración de un Respaldo previo de la máquina virtual.

#### **Resultados Pruebas sobre el Escenario 3**

En este escenario se ha simulado la pérdida total de la máquina virtual desde el VCenter, para esto se borró la máquina virtual del DataStore de VMware y se verificó que no existe conectividad ni posibilidad de acceso a la misma. Posterior a esto se ejecuta la restauración completa de la máquina virtual y se verifica los tiempos de recuperación completa de la máquina virtual.

#### **Resultados Pruebas sobre el Escenario 4**

Para la ejecución de esta prueba se realizó una verificación inicial de la Base de Datos SQL para verificar el estado inicial de la misma, posterior a esto se procedió a borrar la base de datos

## **7. CONCLUSIONES.**

- ) A lo largo de la elaboración del presente estudio se pudo constatar que la consolidación de servidores en ambientes virtuales es una tendencia mundial que seguramente se mantendrá gracias a las ventajas que esta tecnología ofrece, tales como la utilización eficiente de recursos de CPU, memoria, tarjetas de red, etc.
- ) Un sistema de respaldo y recuperación apropiado para ambientes virtuales debe tomar en cuenta lo expuesto en el punto anterior, por lo tanto, deberá utilizar mecanismos que permitan la no contención al momento de lanzar una operación de backup. La prioridad estos sistemas es la optimización en la utilización de recursos del ambiente virtual.
- ) Mediante el uso del mecanismo de respaldo a través de snapshots o imágenes de las máquinas virtuales, se utiliza todas las facilidades que ofrece el Hypervisor, VMware, por ejemplo, para optimizar los backups. Por ejemplo, Avamar se integra con VMware a través de VADP para poder respaldar las VM de una manera no disruptiva sin tiempos de downtime del ambiente productivo.
- ) La mejor estrategia para el respaldo del ambiente de usuarios finales, es la de colocar un agente o software en el computador del cliente final. Este agente puede manejar de manera automática las operaciones de respaldo y restauración y además aplicar mecanismos de optimización del respaldo, por ejemplo, compresión de información para encontrar duplicada en los datos y enviar por la red solo información nueva.
- ) La capacidad de compresión que logra la herramienta de respaldos permite ahorrar en recursos del host respaldado (CPU, RAM, etc.) y en ancho de banda. El cliente de Avamar solo envía información nueva por la red, por lo tanto, el ahorro de ancho de banda, tiempo, y recursos físicos es sustancial. Esto se pudo corroborar en las pruebas realizadas a la herramienta.
- ) Se pudo probar la funcionalidad de la herramienta, mediante la cual se puede limitar el ancho de banda al agente de respaldo. Durante las pruebas de respaldos se limitó el ancho de banda del agente a 20Mbps y fue posible constatar que durante la operación de backups el consumo de ancho de banda se mantuvo alrededor de este valor. Esta funcionalidad es muy importante en ambientes muy pesados donde el ancho

de banda es vital para la operación de ciertas aplicaciones críticas del negocio.

) La recuperación de máquinas virtuales que han sufrido daños irreversibles desde un Respaldo de la Herramienta requieren de una Máquina Virtual de similares características a la máquina dañada. Este procedimiento se asemeja bastante a las recuperaciones Bare Metal que consisten en recuperar una máquina física respalda en un Hardware de similares características del quedó inoperativo.

) Los respaldos granulares, como en el Escenario de Respaldo y Recuperación de la Base de Datos SQL, permiten recuperar información de discos locales de la Máquina Virtual sin afectar Discos específicos como el de Sistema Operativo que pudieron tener cambios y actualizaciones que pueden requerirse no regresar a puntos anteriores en el tiempo.

## 8. RECOMENDACIONES.

) Para el respaldo de ambientes virtuales especialmente, se recomienda tener las precauciones debidas respecto a la herramienta de backup seleccionada para este propósito. La solución de backup debería al menos manejar dos sistemas de respaldo (por agente y por imagen de VM), integrarse con la plataforma virtual y utilizar los APIs disponibles del Hypervisor.

) Es muy importante pensar en los respaldos como en un sistema integral. Muchas de las empresas tienen sistemas de respaldo independientes para cada ambiente. Se recomienda analizar las opciones del mercado y buscar una herramienta que integre todos estos escenarios en una sola solución centralizada desde la cual se pueda controlar todo el ambiente de backups, esto es usuarios finales, máquinas virtuales y servidores físicos.

) Para tener un ambiente de respaldos y recuperación confiable se recomienda en primer lugar tener una política sólida de backups en la empresa. Deben definirse las prioridades de las aplicaciones correctamente, la periodicidad de los respaldos, períodos de retención, etc. Estas políticas dependerán del RTO y RPO particular del negocio.

) Para el respaldo de máquinas virtuales críticas se recomienda realizarlo a través de la utilización del método de imágenes o snapshots usando un Proxy. El proxy permitirá el respaldo no disruptivo de las máquinas virtuales que un agente sí

provocaría. Además, la utilización de este esquema libera los recursos del host ya que la carga del proceso del backup la asumirá el proxy y el VCenter en el caso de VMware.

) Para el ambiente de respaldos de usuarios finales se recomienda tener presente el consumo de ancho de banda, ya que necesariamente este respaldo debe realizarse en horas laborables mientras el computador del cliente esté prendido. Por lo tanto, se debería considerar una herramienta que permita por una parte el ahorro de ancho de banda con compresión, y por otro el control del ancho de banda que el agente pueda utilizar para limitarlo y que no lo monopolice.

## 7. ACCESO DE REFERENCIAS

- [1] «A la venta el libro Virtualización Corporativa con VMware - El blog de Josep Ros». [En línea]. Disponible en: <http://www.josepro.com/2009/05/1a-venta-el-libro-virtualizacion.html>. [Accedido: 08-sep-2017].
- [2] «Avamar. Respaldo y recuperación con deduplicación de datos. EMC». [En línea]. Disponible en: <https://spain.emc.com/data-protection/avamar.htm>. [Accedido: 07-sep-2017].
- [3] «h10584-avamar-technical-considerations-sp-wp.pdf | Virtual Private Network | Cloud Computing», *Scribd*. [En línea]. Disponible en: <https://es.scribd.com/document/326014247/h10584-avamar-technical-considerations-sp-wp-pdf>. [Accedido: 07-sep-2017].
- [4] C. Sailema y W. Gistavo, «Análisis, consideraciones de diseño e implementación en laboratorio de un sistema de respaldo de datos de máquinas virtuales y usuario final a través de la red LAN. Caso de estudio AVAMAR», Quito/PUCE/2015, 2015.
- [5] «App Volumes Database Best Practices». [En línea]. Disponible en: <https://www.vmware.com/techpapers/2017/vmware-app-volumes-database-best-practices.html>. [Accedido: 08-sep-2017].
- [6] J. Hamilton y E. W. Olsen, «Design and implementation of a storage repository using commonality factoring», en *20th IEEE/11th NASA Goddard Conference on Mass Storage Systems and Technologies, 2003. (MSST 2003). Proceedings.*, 2003, pp. 178-182.
- [7] J. Hamilton y E. W. Olsen, «Design and implementation of a storage repository using

- commonality factoring», en *20th IEEE/11th NASA Goddard Conference on Mass Storage Systems and Technologies, 2003. (MSST 2003). Proceedings.*, 2003, pp. 178-182.
- [8] K. Dimosthenis, *Data Intensive Storage Services for Cloud Environments*. IGI Global, 2013.
- [9] «Backup: Jim Butcher, Mike Mignola: 9781596061828: Amazon.com: Books». [En línea]. Disponible en: <https://www.amazon.com/Backup-Jim-Butcher/dp/1596061820>. [Accedido: 08-sep-2017].
- [10] «Cloud Computing Changes Data Integration Forever: What's Needed Right Now - IEEE Journals & Magazine». [En línea]. Disponible en: <http://ieeexplore.ieee.org/document/7962124/>. [Accedido: 08-sep-2017].
- [11] «Cloud Security Engineering: Theory, Practice and Future Research - IEEE Journals & Magazine». [En línea]. Disponible en: <http://ieeexplore.ieee.org/document/8024044/>. [Accedido: 08-sep-2017].
- [12] «Big Data analytics - IEEE Conference Publication». [En línea]. Disponible en: <http://ieeexplore.ieee.org/document/6398180/>. [Accedido: 08-sep-2017].
- [13] «SGSI support throught malware's classification using a pattern analysis - IEEE Conference Publication». [En línea]. Disponible en: <http://ieeexplore.ieee.org/document/7778516/>. [Accedido: 08-sep-2017].
- [14] J. Ros, «EMC Avamar, Backup de usuarios y VMs», *El blog de Josep Ros.* .
- [15] «Administrador de Servidores (1)». [En línea]. Disponible en: <https://es.scribd.com/doc/207673220/Administrador-de-Servidores-1>. [Accedido: 08-sep-2017].
- [16] «Administración de Storage y Backups de Maximiliano Dante Cantone: Editorial Ra-Ma 9788499640648 Paperback - Irish Booksellers». [En línea]. Disponible en: <https://www.iberlibro.com/Administraci%C3%B3n-Storage-Backups-Maximiliano-Dante-Cantone/22513281707/bd>. [Accedido: 07-sep-2017].
- [17] «Descubre y domina VMware vSphere™ 5 by José María González (Paperback) - Lulu». [En línea]. Disponible en: <http://www.lulu.com/shop/jos%C3%A9-mar%C3%ADa-gonz%C3%A1lez/descubre-y-domina-vmware-vsphere-5/paperback/product-21150606.html>. [Accedido: 08-sep-2017].
- [18] «Avamar Backup Recovery VMware Environments Wp», *Scribd*. [En línea]. Disponible en: <https://es.scribd.com/document/200981015/Avamar-Backup-Recovery-Vmware-Environments-Wp>. [Accedido: 07-sep-2017].
- [19] «Hacking Web Technologies». [En línea]. Disponible en: <http://0xword.com/es/libros/81-hacking-web-technologies.html>. [Accedido: 08-sep-2017].