

Capítulo 2

Dentro del capítulo dos, se definen y especifican los recursos de hardware y software necesarios para la virtualización de servidores mediante la técnica de Solaris Containers, al igual que se explican los principales elementos que se utilizan en esta técnica.

2. RECURSOS PARA LA VIRTUALIZACIÓN DE SERVIDORES

Haciendo mención al capítulo anterior, existen diversas técnicas de virtualización de servidores, cada una con sus ventajas, desventajas y métodos de aplicación. El presente trabajo de Tesis está enfocado en la técnica de virtualización de servidores a nivel de Sistema Operativo mediante Solaris Containers, la cual elegimos basándonos en diferentes características, entre las que destacan:

- Es de distribución libre
- Aplica para plataformas SPARC y x86/x64
- Puede albergar versiones anteriores a Solaris 10
- Tiene la capacidad de virtualizar ambientes Linux
- Alto nivel de granularidad de uso del procesador

Cuando una organización se encuentre en la necesidad de implementar una solución que involucre la virtualización de servidores, deben de tomarse en cuenta diferentes aspectos para elegir la mejor técnica de virtualización. Dependiendo de las aplicaciones que se piensen virtualizar debe de elegirse la herramienta que mejor satisfaga los requerimientos de las aplicaciones, por ejemplo, algunas aplicaciones funcionan basándose en una versión específica de kernel del Sistema Operativo, para lo cual la mejor herramienta de virtualización sería aquella que cuente con el nivel de kernel aislado entre los diferentes ambientes virtualizados para evitar que alguna actualización del nivel de kernel afecte a los demás servidores virtualizados y que proporcione más independencia a cada ambiente. Algunas de las herramientas que aplicarían para éste tipo de solución son Oracle VM Server for SPARC (antes llamados Sun Logical Domains) y VMWare ESX, entre otros.

La Tabla Comparativa de las diferentes Técnicas de Virtualización presentada en el tema 1.3.4, puede servir como base y ayuda a las áreas de Tecnología de la Información al momento de tomar una decisión para elegir la mejor técnica de virtualización de servidores para sus ambientes.

Así como debe de elegirse la mejor herramienta de software para la virtualización de servidores, también debe de elegirse el mejor equipo de hardware para llevar a cabo la implementación de una solución de virtualización.

El tema de la virtualización de servidores se ha vuelto popular y más organizaciones son atraídas por esta nueva solución, lo cual ha generado que los diferentes proveedores de hardware inviertan en el desarrollo de nuevas tecnologías enfocadas a la virtualización. Tanto AMD como Intel han agregado como características a sus últimos procesadores métodos especiales para que el hardware asista y mejore varios de los procesos que se ven involucrados en la virtualización.

AMD tiene la extensión AMD Virtualization (AMD-V™) mientras que Intel utiliza su denominado Intel® Virtualization Technology (Intel® VT).

Dichas tecnologías permiten crear de manera más eficiente las máquinas virtuales, ayudan a la migración en vivo de máquinas virtuales y conceden el acceso directo de los dispositivos de entrada y salida de una máquina virtual, sin tener la necesidad de pasar por el hypervisor, entre otros beneficios.

En la elección del equipo de hardware deben de tomarse ciertas consideraciones, entre las que destacan:

- La arquitectura del sistema
- El procesador
- Cantidad de bahías para tarjetas PCI-E
- Número de bahías para discos duros
- Soporte a diferentes Sistemas Operativos

Por ejemplo, cuando se decide por una solución de virtualización con la herramienta Oracle VM Server for SPARC (antes llamada Sun Logical Domains) debe de considerarse implementar esta solución en un equipo de hardware que cuente con un procesador **SPARC®** de la tecnología Ultra SPARC CMT® (Chip Multi Threading), ya que éste procesador cuenta con la capacidad de generar varios hilos de ejecución simultáneos para atender varias tareas y la herramienta Oracle VM Server explota estas características al máximo, ya que el diseño de la herramienta está basado directamente en el procesador **Ultra SPARC CMT®**.

Así como se muestra en la Tabla Comparativa del capítulo 1.3.4, existen algunas herramientas de virtualización que sólo pueden implementarse en ciertas arquitecturas, como por ejemplo Windows Server 2008 Hyper-V, la cual sólo puede usarse en plataformas x86/x64 para virtualizar ambientes de la misma plataforma; por el contrario, la herramienta Solaris Containers puede implementarse en plataformas x86/x64 y plataforma SPARC® y puede virtualizar ambientes UNIX y Linux.

Otro punto importante que debe tomarse en cuenta al momento de decidir la adquisición de un equipo de hardware son los costos de software y hardware. Estos son cuantificados usando métodos tales como el Retorno de Inversión (ROI) o el Costo Total de Propiedad (TCO), los cuales son dos mecanismos de negocio importantes que se utilizan para determinar el valor económico de una decisión de compra.

Se pueden encontrar herramientas en línea como “Return of Investment (ROI) Calculator” de Microsoft y el “VMware ROI Calculator” de VMware que permiten simular el escenario y muestra como se obtiene un ROI favorable en la virtualización de servidores, ya que se han hecho varios estudios que indican esto. Para el análisis de TCO, también

existen herramientas que tienen un costo de licencia, por ejemplo “Gartner’s TCO Manager” y “Gartner’s TCO Analyst”, ambos son propiedad intelectual de **Gartner**.

2.1 Herramientas de hardware

En este capítulo, se especifican a detalle los recursos de hardware utilizados en la presente Tesis, tal como lo es el servidor Sun Fire T2000 que pertenece a la familia de servidores con procesadores UltraSPARC®, también llamados CMT (Chip MultiThreading).

El procesador CMT se caracteriza por tener la capacidad de correr varios hilos de ejecución al mismo tiempo, lo cual evita el “tiempo muerto” cuando se ejecuta una operación del procesador y pasa a memoria, es decir, el procesador ejecuta la siguiente tarea encolada sin quedarse ocioso.

Como herramienta de hardware elegimos el servidor Sun Fire T2000, del fabricante Sun Microsystems (recientemente adquirido por Oracle). Este equipo de hardware se eligió debido a que cuenta con tecnología CMT especial para la virtualización, además de que es producto del mismo fabricante de Solaris 10 y a su vez, de Solaris Containers, lo cual representa una ventaja ya que dichas soluciones están adecuadas al hardware del fabricante. Además, la herramienta Solaris Containers es capaz de aprovechar la capacidad de hilos simultáneos de ejecución que posee el procesador Ultra SPARC del servidor Sun Fire T2000.

2.1.1 Especificaciones del servidor

El equipo Sun Fire T2000 cuenta con procesador UltraSPARC® T1, el cual entrega hasta un máximo de 32 hilos simultáneos de ejecución por procesador, lo cual genera el procesamiento ideal no sólo para aplicaciones Web, también para aplicaciones de misión crítica, bases de datos y aplicaciones de Planificación de Recursos Empresariales (Enterprise Resource Planning ERP) y Administración Basada en la Relación con los Clientes (Customer Relationship Management CRM).

Tal como se describe en el documento de especificaciones del equipo Sun Fire T2000 (Ver Anexo 1), el equipo cuenta con capacidades de alta disponibilidad, tales como discos

duros instalables en caliente (“hot plug” en inglés, lo cual significa, que el servidor no tiene que apagarse para instalar un disco duro al chasis), fuentes de poder redundantes y reemplazables en caliente (“hot swappable”, es decir, si una fuente de poder falla, la otra sigue funcionando y la dañada se desmonta del chasis sin apagar el servidor y se instala la nueva sin reiniciar el equipo), monitoreo del ambiente del servidor, manejo de niveles de RAID, entre otras.

El servidor es un equipo de 2 unidades de Rack, el cual entrega un alto desempeño y capacidades de RAS con un mínimo costo de energía en un pequeño espacio, tiene una capacidad máxima de 64GB en RAM, 1 procesador por chassis y Sistema Operativo Solaris 10.

El equipo que usaremos durante el desarrollo de la Tesis, es un equipo Sun Fire T2000 con 1 procesador UltraSPARC® T1 Quad Core a una velocidad de 1.0 GHz, 8 GB en RAM y 2 Discos Duros de 146GB.

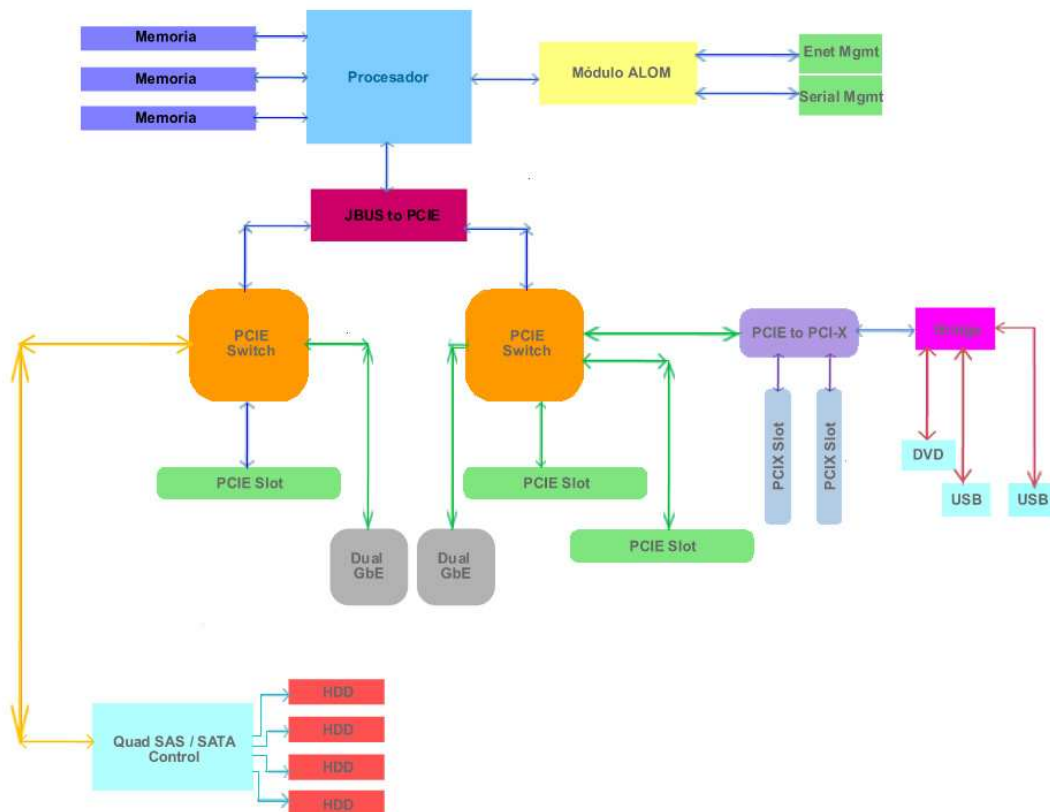


Figura 2.1 Diagrama de Servidor

En la figura 2.1 se muestra la conectividad entre los diferentes componentes que forman la arquitectura del sistema del servidor Sun Fire T2000. Se puede apreciar que los bancos de memoria están conectados directamente al procesador a través de un bus de datos.

También conectado al procesador se encuentra un módulo llamado “ALOM” el cual es un pequeño firmware que permite la administración del servidor a través de los puertos de administración marcados en los recuadros verdes conectados al módulo.

Para la comunicación entre el procesador y las tarjetas de conectividad a la red de datos, se encuentra un bus de datos especial, llamado “JBUS a PCIE”. Este bus es el canal de comunicación entre los switches PCIE y las tarjetas de datos conectadas a ellos. Como se puede ver en el diagrama, existen 5 bahías PCIE, de las cuales 2 bahías están configuradas para soportar tarjetas tipo PCIX; pero para que los switches PCIE reconozcan el microcódigo de éstas tarjetas PCIX, existe un pequeño bus que realiza una conversión de PCIX a PCIE para que la tarjeta sea reconocida por los switches PCIE y a la vez, por el procesador.

Dentro de la arquitectura del procesador existe también un “bridge”, el cual tiene la tarea de comunicar los puertos USB y DVD del chasis hacia el JBUS del procesador.

En cuanto a los discos duros, éstos se comunican al procesador a través de un “Control Quad SAS ó SATA” que parte de los dos Switches PCIE que se conectan al procesador. Este control Quad SAS ó SATA cuenta con 4 bahías de disco duro conectadas para instalar 4 discos duros al servidor.

2.1.2. Recursos del Servidor

El Servidor Sun SPARC T2000 con el que estamos trabajando, cuenta con los siguientes recursos de hardware:

- 2 Tarjetas HBA FC Dual Channel QLogic
- 4 puertos Gigabit Ethernet
- 8 DIMMs DDR2 de 1 GB
- 1 procesador Ultra SPARC T1 con 6 cores q 1.0 GHz
- 2 Discos Duros de 146GB

2.2. Herramientas de Software

Las herramientas de software que se utilizan para la implementación y diseño en este trabajo son descritas a continuación, explicando sus principales características y funcionamiento.

2.2.1. Solaris 10

Solaris 10 es un sistema operativo UNIX de código abierto lanzado por Sun Microsystems a principios del año 2005, es gratuito, sin importar la cantidad o tipo de sistemas en que se requiera utilizar y compatible con muchas de las arquitecturas de hardware de fabricantes como Dell, Fujitsu, Siemens Computers, IBM y HP. Funciona principalmente en arquitecturas SPARC de 32 y 64 bits, x86, AMD64 y EM64T, al ser un sistema operativo que emplea la técnica de multiprocesamiento con hilos de ejecución, se obtiene un mejor rendimiento de los procesadores.

El sistema operativo Solaris 10 es el software que maneja los recursos y gestiona las tareas que se ejecutan en el sistema, se encarga de los datos almacenados y de la comunicación con los dispositivos de entrada y salida, como monitores, discos duros, impresoras, entre otros. Solaris 10 cuenta con el software ONC+ (*Open Network Computing Remote Procedure Call*), protocolo con el que se proveen varios servicios de red, tales como NFS, NIS. El ambiente de escritorio de Solaris 10, si se cuenta con interfaz gráfica, es un ambiente de ventanas o GUI, con el que se accede a todas las funciones del sistema (Figura 2.2).

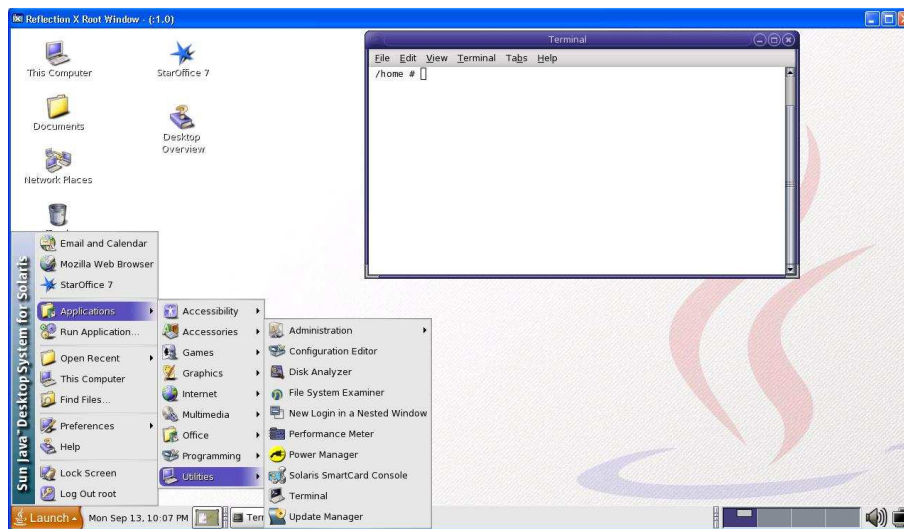


Figura 2.2 Interfaz gráfica de Solaris 10

Los tres principales componentes del sistema operativo Solaris 10 son:

- El **kernel** es el núcleo del sistema operativo, que maneja todos los recursos físicos, entre los que se incluyen: los sistemas de archivos y estructuras de datos, los dispositivos de almacenamiento, las llamadas al CPU, la administración de procesos y el manejo de memoria.
- El **Shell** es la interfaz entre el usuario y el kernel, acepta los comandos introducidos por el usuario, los interpreta y los pasa al kernel para que este los ejecute. Los shells que Solaris 10 principalmente contiene son: Bourne Shell, C Shell y Korn Shell.
- Los **directorios jerárquicos** son la forma en que están organizados el conjunto de archivos del sistema, la estructura tiene forma de árbol con raíz, por lo que la revisión y búsqueda de información se realiza de forma más sencilla, ya que el usuario debe indicar el directorio para localizar el archivo deseado, lo cual se conoce como ruta de acceso del archivo.

Entre las principales características que resaltan de Solaris desde versiones anteriores se encuentran:

- **Seguridad**, proporciona funciones avanzadas que permiten proteger mejor el sistema y consolidar servicios de forma segura, utiliza manejo de permisos de procesos y de usuarios, contenedores y sistemas de archivos. Las aplicaciones y los usuarios sólo son capaces de tener acceso de lectura a los datos, o de escritura, siempre y cuando tenga la autorización apropiada.
- **Rendimiento**, ofrece diferentes ventajas de rendimiento para servicios de base de datos, aplicaciones web y Java. Además de que tiene la capacidad para ejecutarse bajo diferentes cargas de trabajo en sistemas que van desde computadoras de escritorio con un solo procesador hasta sistemas más complejos en los que se requiere el multiprocesamiento.
- **Disponibilidad**, está diseñado para mantener el sistema funcionando en caso de producirse fallas de hardware y durante rutinas de mantenimiento.
- **Facilidad de monitoreo**, las características del kernel combinadas con software desarrollado por usuarios permite monitorear y analizar el comportamiento y rendimiento del sistema.
- **Manejo efectivo** de los recursos de hardware disponibles en el sistema hace posible soportar grandes cargas de trabajo.

La versión Solaris 10 incluye nuevas tecnologías que le permiten distinguirse de otros sistemas operativos existentes, entre las que podemos destacar las siguientes:

- **DTrace** es un conjunto de herramientas de diagnóstico que ayuda a visualizar la actividad que se está ejecutando en el kernel del sistema y las aplicaciones. Los cuellos de botella pueden identificarse y corregirse en el momento que suceden, sin necesidad de hacer pruebas por separado.
DTrace se ejecuta cuando el administrador necesita recopilar la información necesaria, lo que lo hace una herramienta dinámica ya que al terminar de ejecutarse automáticamente se deshabilita sin afectar el rendimiento del sistema.
- **SMF (Service Management Framework)** es un modelo unificado que provee un conjunto de comandos, utilidades y documentación que facilita el arranque y detención de los servicios que ofrece el sistema operativo.
- **Solaris Containers**, la tecnología de particionamiento de software que habilita la creación y administración de múltiples ambientes virtuales de ejecución sobre la misma instancia de Solaris 10.
- Las tecnologías de **Solaris Fault Manager** y **Solaris Service Manager** incorporan al sistema operativo herramientas para la detección de eventos y proporciona facilidades para la desactivación de componentes con fallas de forma dinámica reduciendo los tiempos de caída y de no disponibilidad.
- **ZFS (Zettabyte File System)** un nuevo sistema de archivos que proporciona una administración más sencilla e integridad en los datos, además de gran escalabilidad al terminar con los problemas asociados al crecimiento de los sistemas de archivos, puesto que cada sistema de archivos consume sólo el espacio que en realidad necesita.

2.2.2 Generadores de carga

Los generadores de carga son herramientas de software diseñadas para simular conexiones de usuarios, cargas de trabajo, accesos a dispositivos de almacenamiento y uso de memoria en un determinado momento en el servidor como si fueran reales, con el fin de realizar pruebas de rendimiento para determinar puntos de falla en las aplicaciones, en el hardware y errores de configuración en el sistema operativo.

Simular diferentes cargas de trabajo ayuda a predecir como la aplicación o servicio que brinda el servidor se va a comportar en diferentes niveles de concurrencia.

La **herramienta Stress** permite generar cargas de trabajo o imponer cierto nivel de estrés al procesador, a la memoria y los dispositivos de entrada y salida. La herramienta está desarrollada en el lenguaje C y es de distribución libre, generalmente se utiliza para realizar pruebas de escalabilidad y para exponer errores que solamente se manifiestan cuando el sistema está sometido a una carga de trabajo muy grande. Con la ayuda de esta herramienta se consigue un patrón de utilización de los recursos bajo diferentes escenarios.

La **herramienta JMeter** simula peticiones de diferentes clientes a las aplicaciones de red como HTTP, HTTPS, FTP y realiza pruebas sobre conexiones de bases de datos. Consigue realizar pruebas de funcionalidad y rendimiento, permitiendo la evaluación de las aplicaciones en diferentes cargas de trabajo y bajo distintos niveles de tráfico en la red. Esta herramienta forma parte del proyecto Apache Jakarta.

Existen diferentes herramientas para generar diferentes cargas de trabajo sobre recursos de hardware o aplicaciones, sin embargo, las presentadas en esta Tesis fueron escogidas debido a ser gratuitas y de código abierto. Existen herramientas similares pero tienen un costo de licencia y son recomendadas para pruebas con aplicaciones propietarias, por ejemplo *LoadGen* está diseñada para producir una carga de trabajo en una implementación de Exchange, *LoadRunner* para servidores HP, etc.

2.3. Containers

Un Container (Contenedor) en Solaris 10 se define como el ambiente de ejecución que utiliza las tecnologías de Zonas y de Resource Manager. Ambas tecnologías permiten aprovechar distintas ventajas de la virtualización, complementándose entre sí, mientras las Zonas proveen de un ambiente virtual aislado, Resource Manager permite controlar y administrar los recursos de cómputo destinados a cada Zona, dando como resultado un mejor aprovechamiento de los recursos físicos sin comprometer el nivel de servicio.

2.3.1. Zonas

Como se ha mencionado hasta ahora, la virtualización consiste en ejecutar dos o más instancias de sistema operativo en un mismo equipo físico, en Solaris 10 la tecnología de particionamiento que permite hacer esto es llamada Solaris Zones, o simplemente Zonas.

Dichas zonas consisten en un conjunto de procesos que permiten crear ambientes de ejecución aislados lógicamente. De esta forma cada ambiente se administra de forma independiente, sin afectar aplicaciones ejecutándose en otras Zonas.

Las Zonas separan las aplicaciones y los datos contenidos en ellas del Hardware en donde residen, facilitando la movilidad entre Hardware. Cada Zona posee sus propios recursos lógicos de sistema operativo, como si fuera una instancia instalada en un equipo físico, tales como passwords de administrador, usuarios, puertos de red, sistemas de archivos, procesos y servicios, interactuando únicamente con sus recursos locales (Figura 2.3).

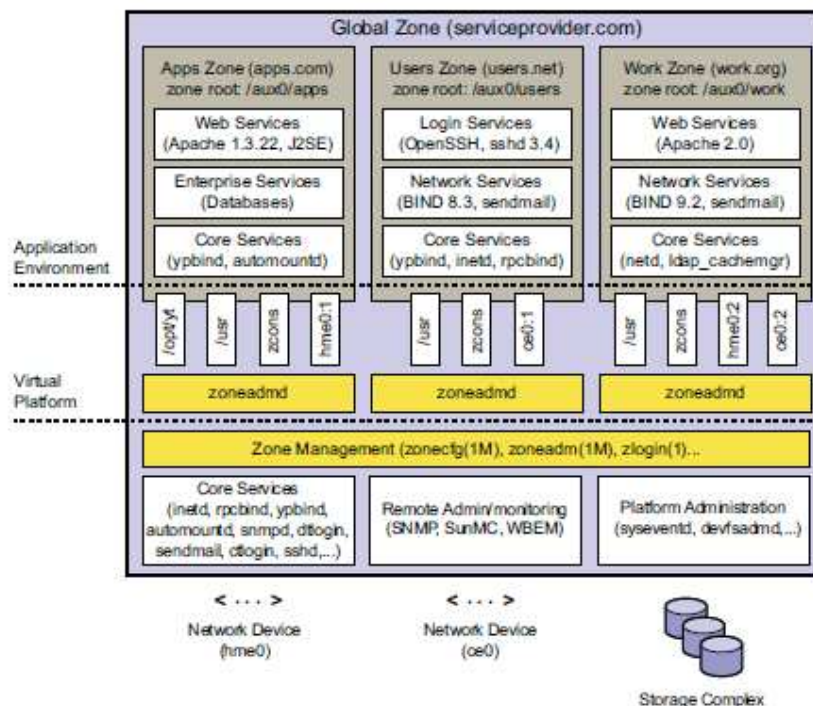


Figura 2.3 Sistema con Zonas Instaladas

Los recursos físicos son compartidos por todos los ambientes virtuales, conocidos como Zonas Locales, mientras que el sistema operativo nativo, llamado Zona Global, se encarga de administrar el uso y acceso de esos recursos, también es la que permite crear, instalar y administrar las Zonas Locales.

Las Zonas Locales contienen un subconjunto de los paquetes del sistema operativo Solaris, además de software adicional o compartido con la Zona Global, así como archivos, datos y configuraciones, sin embargo comparte la operación del Kernel con la Zona Global.

A diferencia de otras técnicas de virtualización de servidores, Solaris Zones tiene dos modelos para instalar el ambiente virtual, el llamado Spares Root Zone, que consiste en compartir la mayoría de archivos binarios con la Zona Global y las demás Zonas Locales, y el Whole Root Zone, semejante a las soluciones de otros fabricantes, en donde los archivos binarios son independientes entre un ambiente virtualizado y otro.

Así como existen dos modelos para instalar las Zonas Locales, existen dos tipos de configuraciones para la red, Shared-IP (IP compartida) y Exclusive-IP (IP exclusiva). En ambos casos existe una virtualización de las interfaces de red, sin embargo, para la configuración de IP compartida la interfaz física de red se comparte con la Zona Global y eventualmente con las otras Zonas Locales, mientras que la configuración de IP exclusiva utiliza una interfaz de red de manera que ni la Zona Global ni otras Zonas Locales puedan hacer uso de dicha interfaz.

Estas dos últimas características, el modelo de instalación del ambiente virtual y la configuración de la red, son elementos clave que se deben considerar al momento de planificar la virtualización, ya que por un lado pueden ofrecer flexibilidad y consolidación de recursos, mientras que por otro lado, de manera excluyente, pueden permitir una mayor independencia entre recursos de los ambientes virtualizados.

2.3.2. Administración de Recursos

Resource Management (administración de recursos) en pocas palabras se refiere a la habilidad de controlar los recursos disponibles que las aplicaciones pueden usar en un sistema, donde básicamente estos recursos son procesamiento y memoria. El Solaris Resource Manager es una tecnología integrada de manera nativa en Solaris 10, lo que significa una mejor integración con otras herramientas. Resource Manager permite reservar recursos de cómputo para una aplicación o proceso en particular, saber cómo están siendo usados los recursos y ajustarlos cuando sea necesario.

Gracias al uso de Resource Manager es posible modificar el comportamiento del sistema operativo respecto a diferentes cargas de trabajo, optimizando así el uso de los recursos, mejorando el rendimiento que se refleja en un mejor tiempo de respuesta y en un mejor servicio.

Existen tres mecanismos para el control de la administración de recursos en el sistema operativo Solaris: mecanismos de restricción, mecanismos de planificación y mecanismos de particionamiento.

Los **mecanismos de restricción** permiten definir los límites de consumo de recursos específicos para una carga de trabajo. Este mecanismo es usado para controlar cargas cuyo comportamiento incorrecto comprometa el nivel de servicio de otras aplicaciones.

Por otro lado, los **mecanismos de planificación** hacen una toma de decisión sobre la reserva y la asignación de los recursos basados en intervalos, así si un recurso está reservado para una aplicación, pero en un instante no está siendo usado, otra aplicación puede hacer uso de él. Solaris 10 Resource Manager hace uso de Fair Share Scheduler (Planificador de Reparto Justo) o FSS para administrar los recursos de procesamiento bajo este mecanismo.

Finalmente, el **mecanismo de particionamiento** describe el uso de subconjuntos de recursos destinados a una aplicación o carga de trabajo, de esta forma cada aplicación asignada a un subconjunto de recursos, dispondrá de un mínimo asegurado, sin embargo, no permite el uso de los recursos ociosos que pertenezcan a otro subconjunto.