

Virtualization with SPARC & Solaris

This presentation focus on methods to achive a redundant and flexible infrastructure where SPARC based Solaris workloads are deployed and virtualized on T-series servers with binary compatibility with the previous generation of servers, achieving platform refresh and virtualization without large-scale software migration projects and cost.

Lars Tunkrans
Distinguished Engineer
Fujitsu Sweden

lars.tunkrans@se.fujitsu.com

Why virtualize with SPARC and Solaris ?

- Because its free ! WHAT ?
- There are NO licensing fees ! Really ?
- YES it is GRATIS , Free , NO Payment !!!
- The Oracle Virtual Machine for SPARC hypervisor is included with every T4 server, No extra charge.
- Virtualizing with Microsoft or Vmware or Citrix costs licenses per cpu or gigabyte. Virtualizing with Solaris costs nothing more than the server it runs on.
- Todays powerful hardware makes virtualization the only way to capitalize on the investment

Design Elements

This presentation will describe a server / storage/ network infrastructure designed to support virtualization. Virtualization requires redundancy of hardware to be able to support virtual machines that are able to migrate from one physical server to another. The design therefore needs to include :

- Multiple physical servers
- Multiple redundant network switches
- Multiple redundant storage fabric (switches)
- A storage array with multiple redundant controllers

Design goals

Active N+1 redundancy

- There shall be spare capacity equal to the power of one physical server in case of a server breakdown.
- All network links to be redundant and aggregated to multiple switches in case of switch failure.
- All storage volumes to be redundant (RAID 5 or 6) and accessible over multiple paths / controllers.

Components for redundancy

To build a redundant infrastructure it is necessary to have in place or invest in a SAN disk array with multiple storage controllers/processors and a storage network and High speed LAN built with multiple switches for FCP/ISCSI and IP traffic. The storage platform must support Solaris storage multipathing.



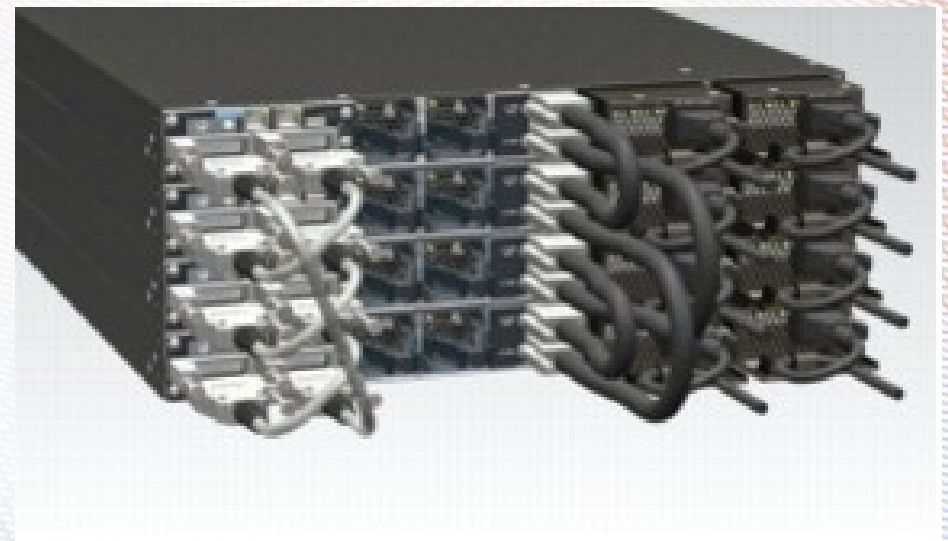
Networks where switches are single point of failure, makes the infrastructure unreliable, and limited by bandwidth of links between switches.

Components for redundancy

Switch infrastructure is extremely Important for a redundant virtualization Platform. The switches need to provide:

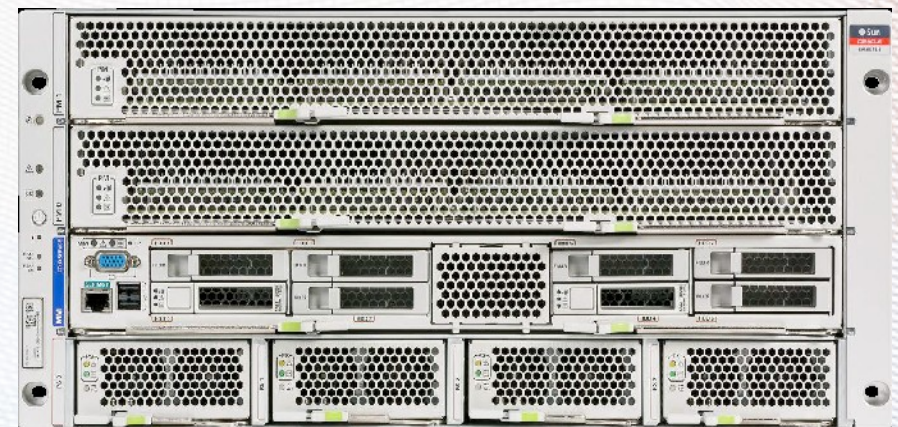
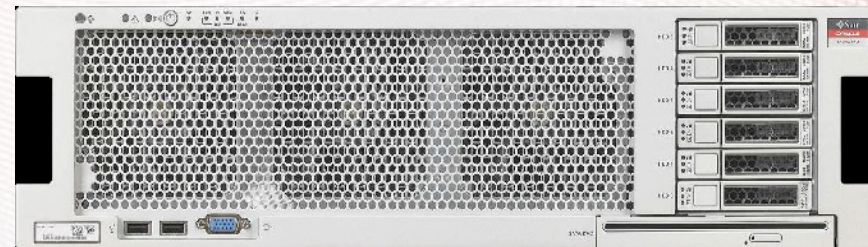
- Hardware redundancy
- Spare power supply capacity
- A fast switching backplane.
- Link aggregation
- LACP protocol
- VLAN (802.1Q) & VLAN routing

Cisco 3750x series of switches are one of the platforms that does this well. They work well with the networking stack in Solaris to deliver a VLAN directly to the desired virtual machine.



SPARC server components

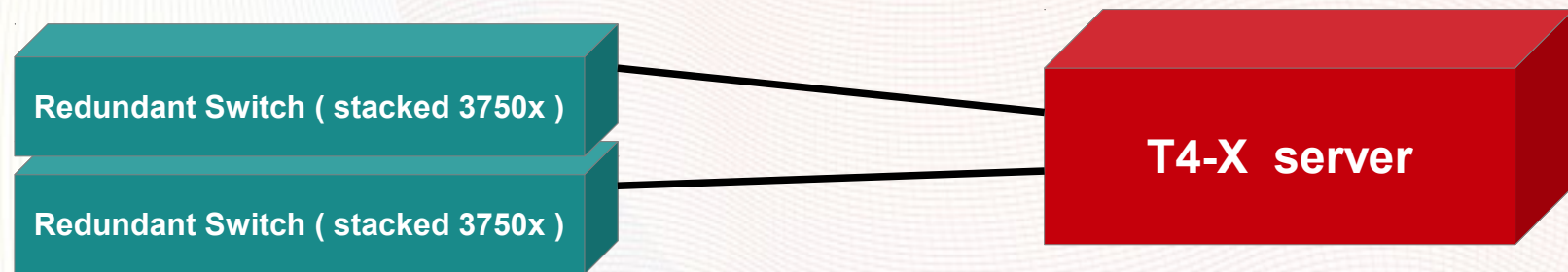
- The Oracle T-series Servers exists in 3 sizes :
- The 2 Rack units T4-1 with 8 Cpu cores (2.85 GHz) and 256 GB RAM
- The 3 Rack units T4-2 with 16 cpu cores and 512 GB RAM
- The 4 Rack units T4-4 with 32 cpu cores (3 GHz) and 1 TB RAM .
- The previous generations of T-Series servers, T2 and T3 servers can also be used.



Virtualization & Redundancy

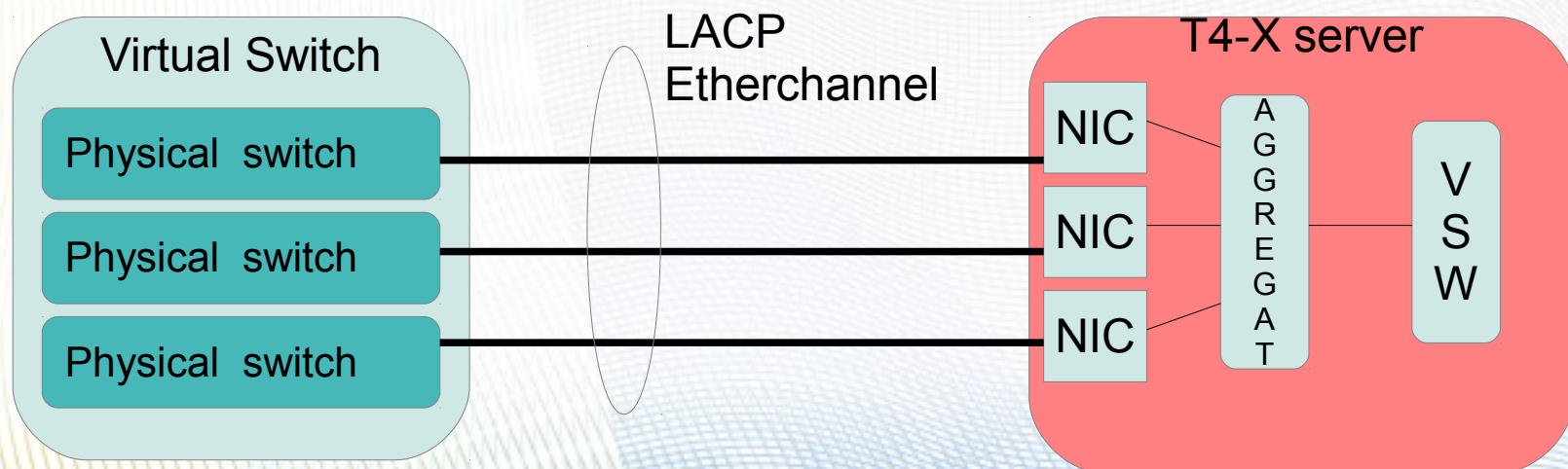
- In order to keep a virtual machine as portable as possible it is a good strategy to create the redundancy elements of the design – outside - the virtual machine.
- Two redundancy enhancing features that exists in the Solaris environment are "storage multipathing , mpxio" and " network interface, or link aggregation"
- Configuring the server base with aggregated network links and multipathed storage volumes allows the virtual solaris machines to have a simple & basic I/O design setup.

Network aggregated connections

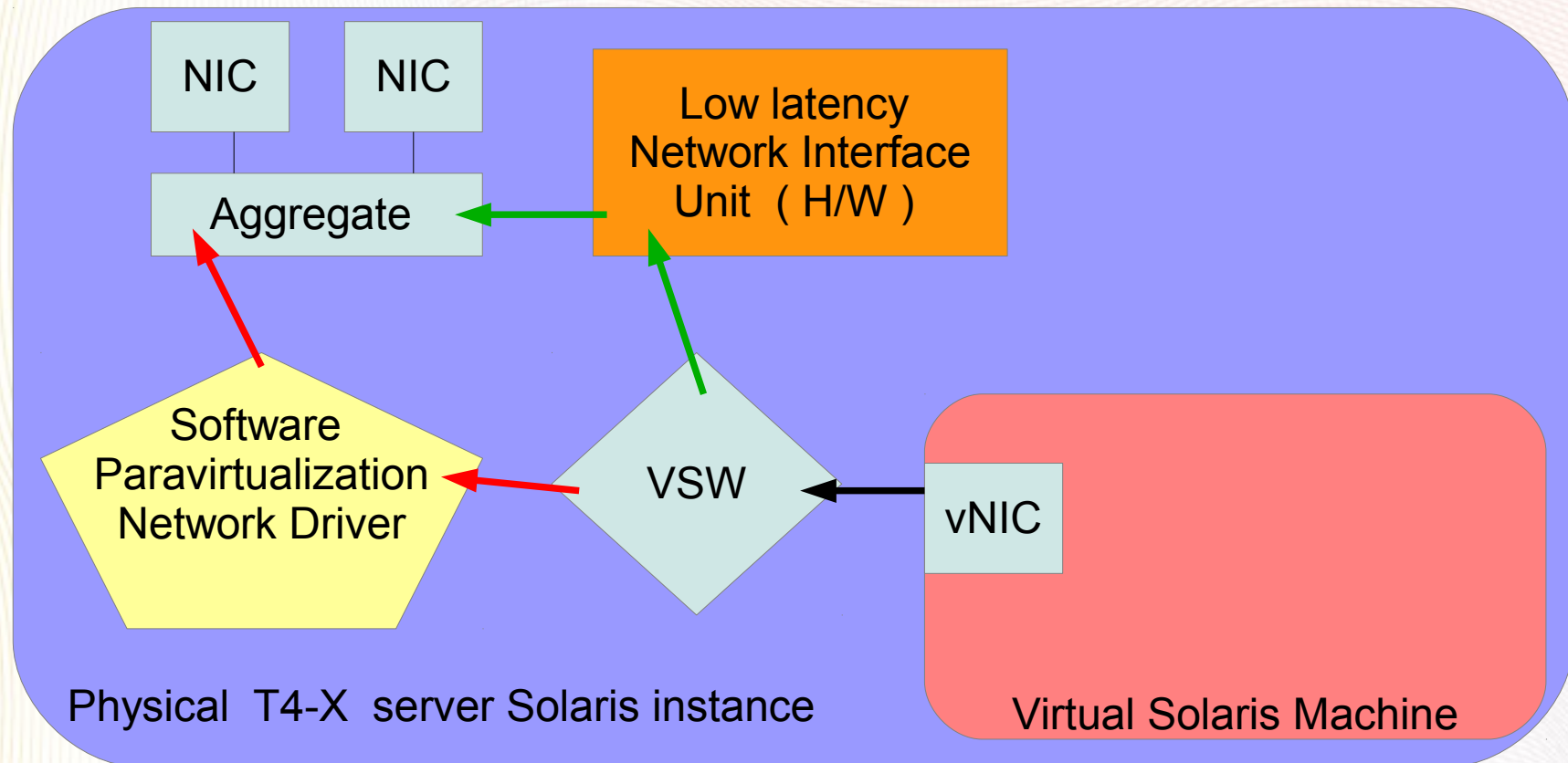


The key technologies in the switch stack is the Ether-channel or Port-channel that is compatible with the Solaris network aggregated link setup. Cisco switches and Solaris both support Link Aggregation Control Protocol and the link therefore is capable of surviving a broken cable or broken hardware interface.

The aggregated Link is then in Software connected to a Virtual SWitch (VSW) which provides network connections to virtual machines



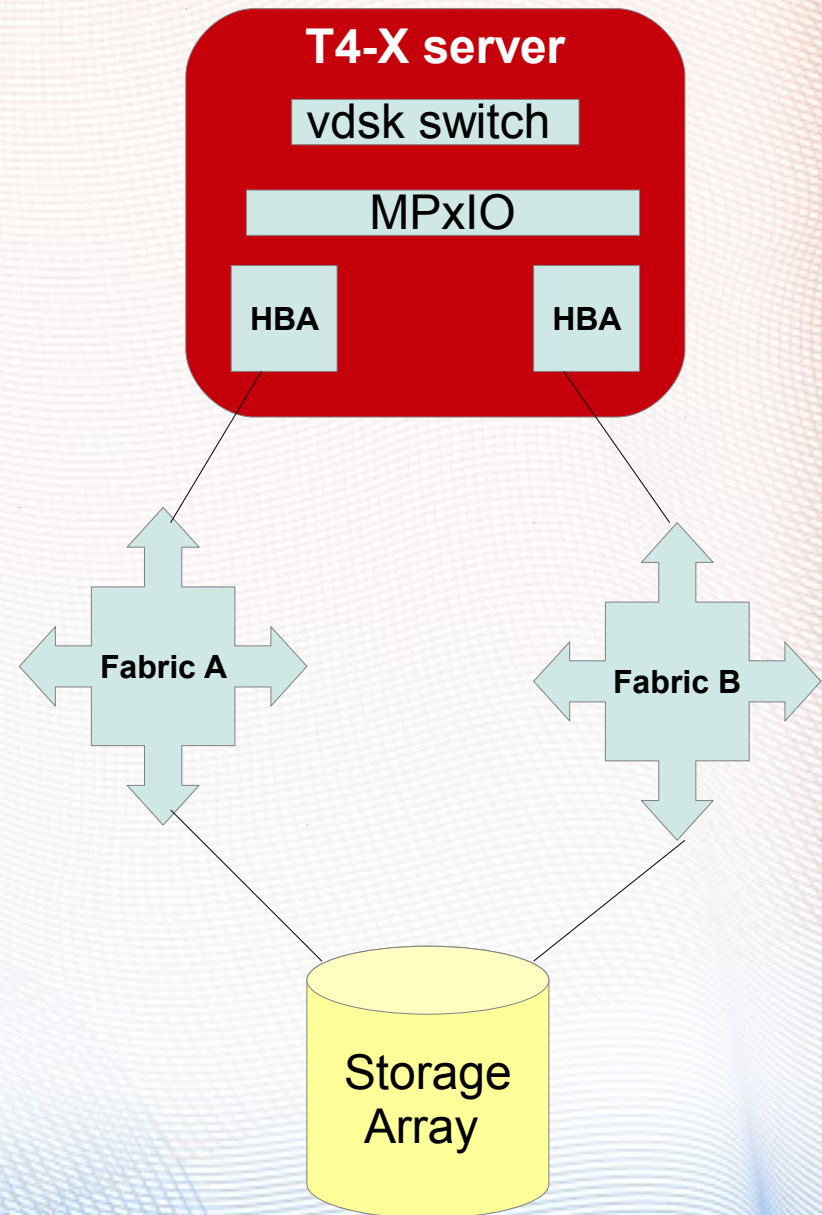
Virtual networking with NIU



Virtual networking can in the T4 server (and T2 and T3) be done in two ways. Either entirely In software with paravirtualization drivers in the HOST O/S, or there is limited resources to run hardware assisted DMA directly from the Virtual NIC in the VM to the physical NIC's on the physical server with the NIU circuitry built into the CPU, for six VM's for each T4 cpu-chip.

Storage multipathing

- Dual redundant storage fabrics makes the volumes in the storage array visible as two different volumes from the server.
- The MpxIO layer removes this doublesightedness and presents one logical diskvolume to the virtual disk switch
- The virtual disk switch serves up storage volumes to the virtual machines

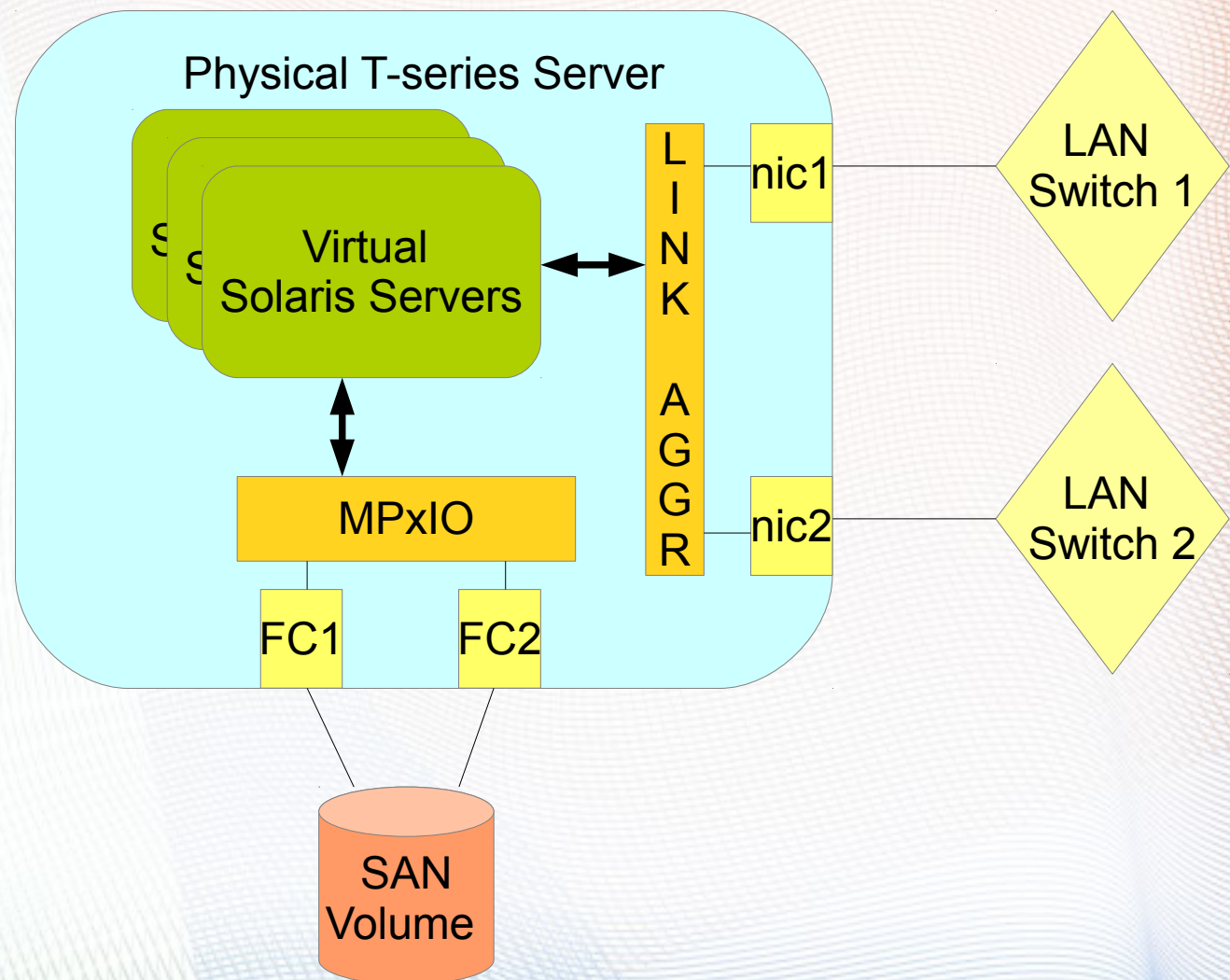


Virtualization & Redundancy

The virtual machines use redundancy features in Solaris on the Physical server.

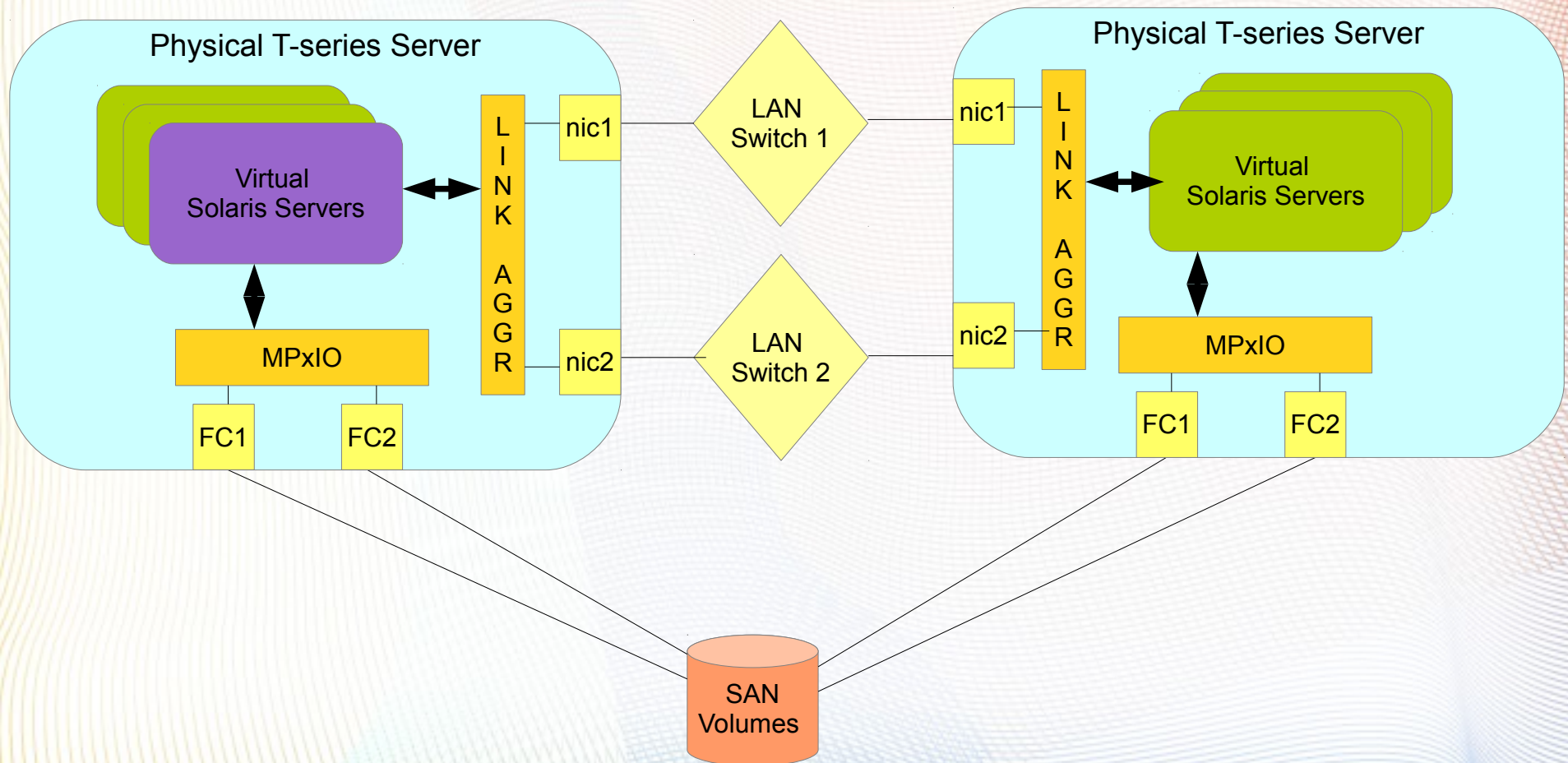
Link Aggregation Control Protocol or LACP can be used with competent switches to disconnect bad connections between switch and server and avoid network problems. Aggregation also hides physical NIC names for the VM, making it portable.

MpxIO hides the complexity of the FC fabric for the storage consumer, the VM, so that a unique disk volume has the same ID across multiple fabrics and multiple servers. The SAN volume becomes “portable” to the neighbour server host.



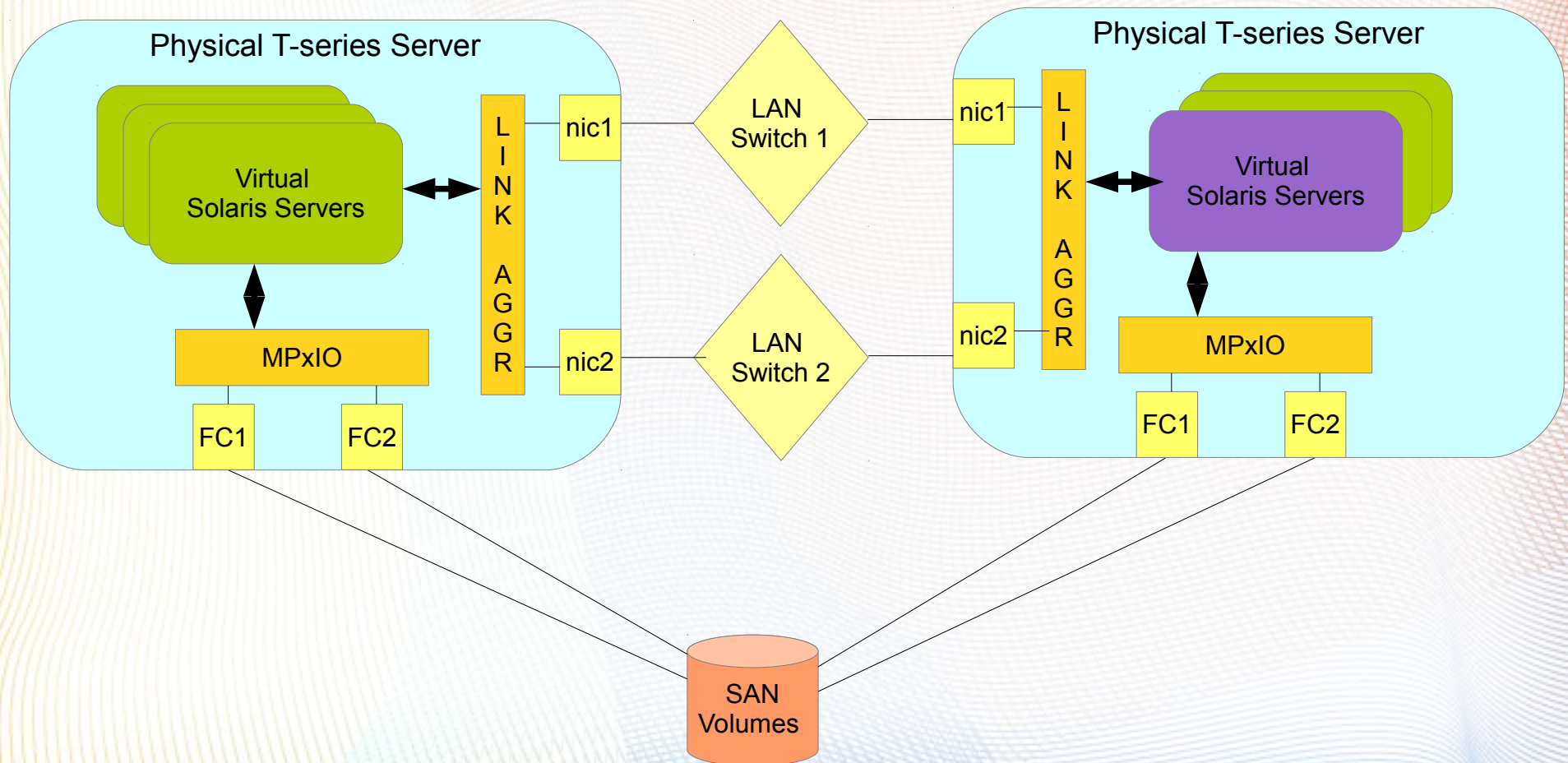
Virtualization & Redundancy

Multiple Servers configured the same way, have the ability to do *LIVE Migration* of virtual machines between physical servers. Load balancing and peak utilization periods for different virtual machines can be managed dynamically.

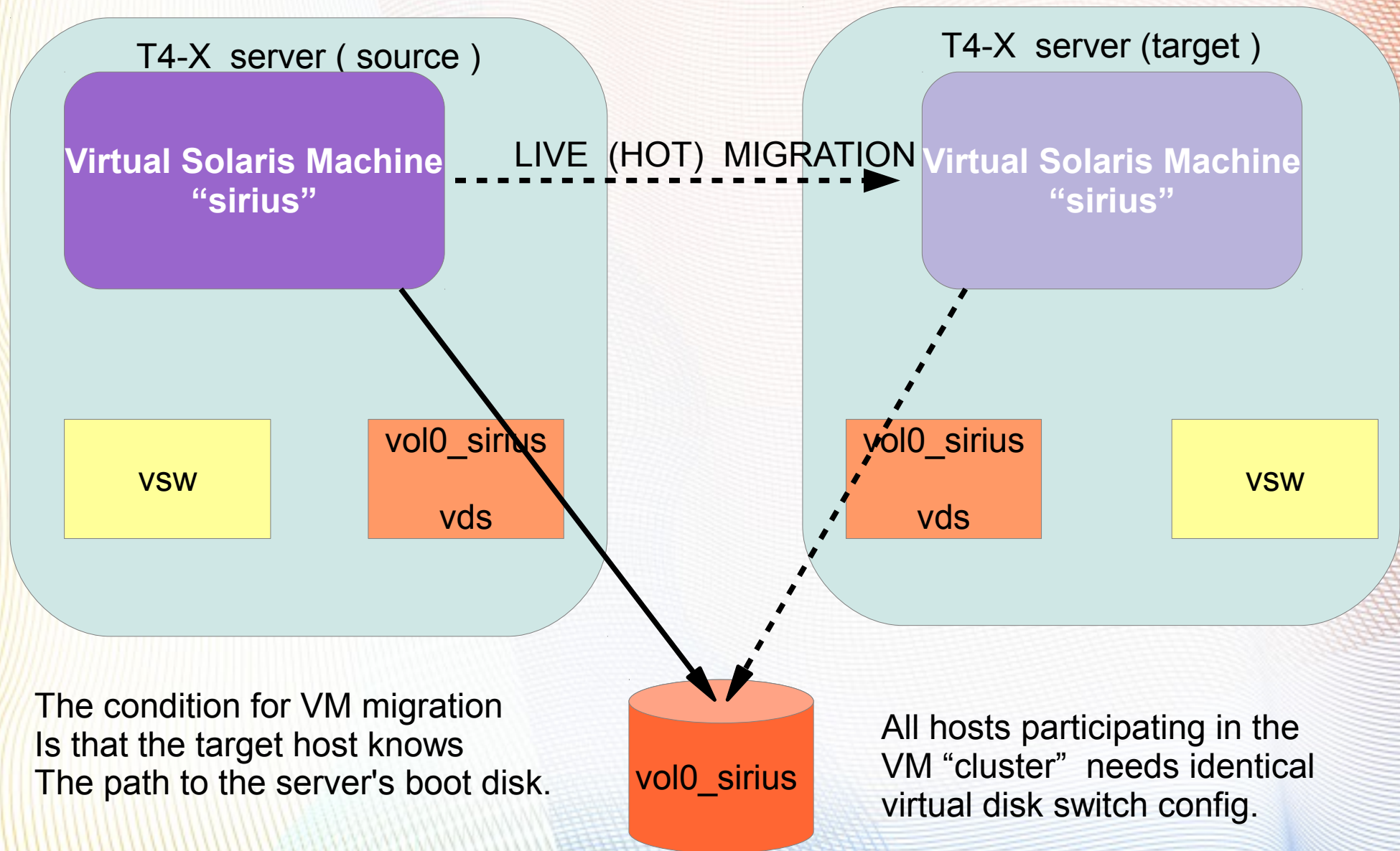


Virtualization & Redundancy

Multiple Servers configured the same way, have the ability to do *LIVE Migration* of virtual machines between physical servers. Load balancing and peak utilization periods for different virtual machines can be managed dynamically.

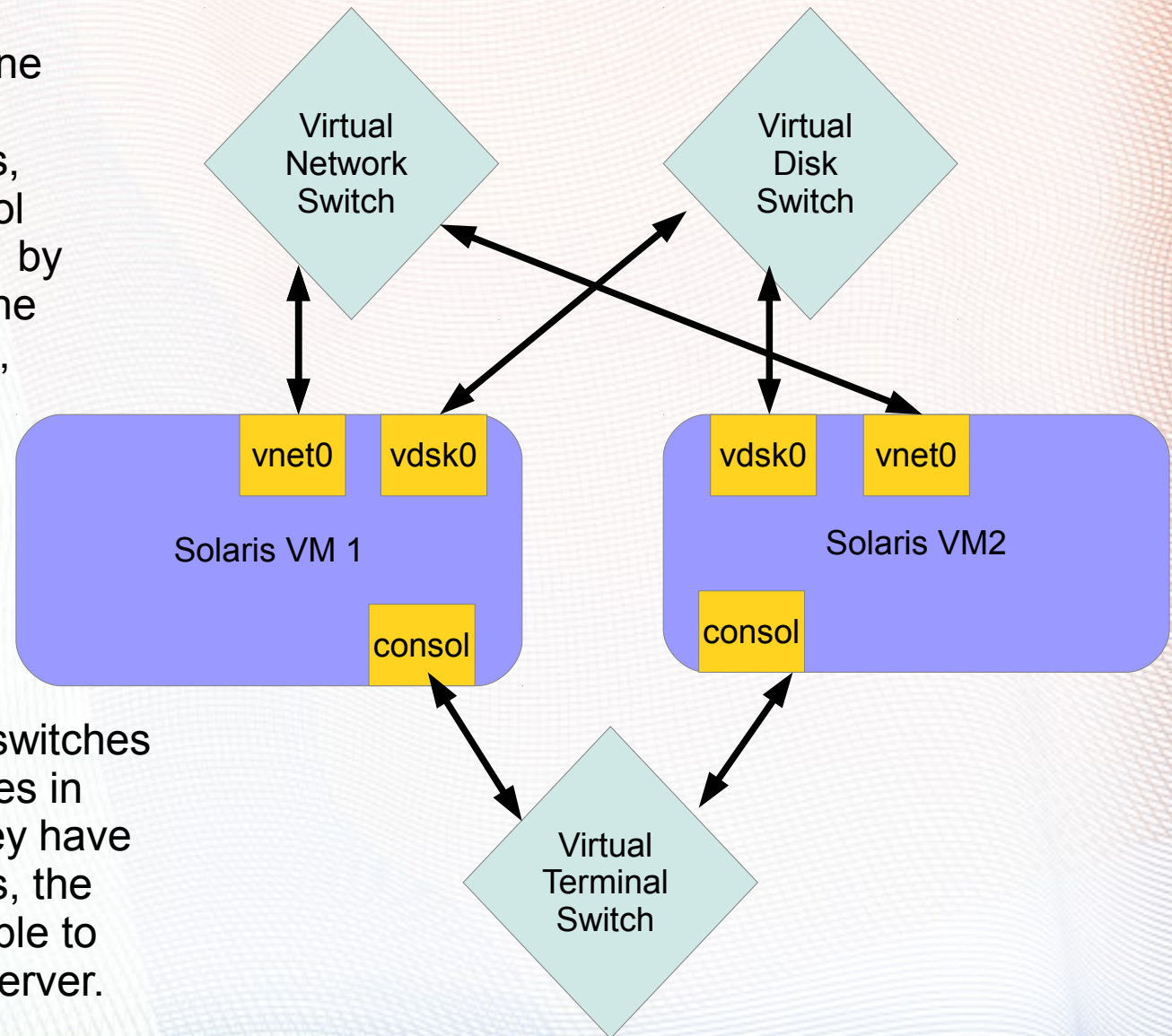


Virtual Machine Migration



The Virtual Machine environment

The virtual solaris machine uses at least 3 different types of virtual resources, Network, Disk and Consol switches that is supplied by the Oracle Virtual Machine for SPARC environment,



As long as these virtual switches are given the same names in different servers, and they have common storage devices, the Virtual machine will be able to migrate from server to server.

Sizing & consolidation

- The T4servers come in three sizes. Either 8, 16 or 32 CPU cores, with 256, 512 or 1024 GB ram.
- A typical Application server like JBOSS or Tomcat/Apache or Websphere works well with 4 to 6 GB ram and 0,5 to 1 CPU core (there is 8 vCPU's per Core so its possible to give a virtual machine 4 vCPU's)
- Its concievable to run 14 virtual machines on a T4-1 server with 8 cores and 128 GB ram,
- 28 virtual machines on a T4-2 and 60 virtual machines on a T4-4 server
- Therefore its possible to consolidate a large number of small servers like the Fujitsu Primepower 250 or the Sun Ultrasparc 3 V240 server onto just a few T4 series machines

Sizeing & consolidation

- Its necessary to understand that a **very large amount of RAM** is needed to support the number of virtual machines a T4 server is capable of running.
- If you plan to replace your 20 old SUN V210 servers that runs your website, each of which are populated with 8 GB RAM.
- You would need 3 of the T4-1 servers with enough RAM to run the 20 servers in 2 of the 3 T4-1's to build a solution where one of the T4-1 servers can be taken offline for service.
- $20 \text{ server} * 8 \text{ GB} / 2 = 80 \text{ GB Ram}$ For Each T4 for virtual machines + 8 GB for the Primary Solaris Server.
- This shows that 64 or 128 GB RAM is a good ballpark figure for the Memory size of the T4-1 server.
- Similarly you need to populate the T4-2 and T4-4 to about half its memory capacity to utilize its processing power.