

LPAR - logiczne partycjonowanie systemów

Bartek Franaszek

Paweł Duńczewski

Jakub Wojciechowski



		7	14	28
1	8	15	22	29
2	9	16	23	30
3	10	17	24	
4	11		25	
5	12	19	26	
6	13	20	27	

Agenda

1. Co to jest (D)LPAR?
2. Cel stosowania
3. Jak to wygląda w IBM – LPAR
4. Konsole HMC i IVM
5. Jak stworzyć własnego LPAR-a
6. Limity i ograniczenia
7. Jak to wygląda w SUN – LDOM`s
8. Jak to wygląda w HP – VSE
9. Hypervisor
10. Architektury wirtualne
11. Fujitsu-Siemens
12. Pytania?

Scenariusz

Serwer Studium Wychowania i Sportu UW

Obciążenie: zazwyczaj niskie,
dwa razy w roku bardzo wysokie



Serwer students Wydziału MIM UW

Obciążenie: zazwyczaj średnie,
większe do południa,
kilka razy w roku bardzo wysokie

Serwer BUW

Obciążenie: zazwyczaj średnie,
większe po południu



Czy można jakoś zoptymalizować wykorzystanie serwerów ?

Co to jest Logiczne Partycjonowanie?

- Logiczne partycjonowanie (LPAR) polega na dzieleniu zasobów komputera pomiędzy różne środowiska w taki sposób, by każde środowisko mogło pracować niezależnie.
- Dwuznaczność: LPAR = logiczna partycja
- Liczba logicznych partycji którą można stworzyć zależy od systemu
- Na ogół kolejne partycje są wykorzystywane do oddzielnych zadań
- Partycje komunikują się jak oddzielne maszyny.

Podstawowe Parametry określające Logiczną Partycję

Procesory

- Ilość procesorów widziana przez system.
- Moc procesorów.
- Możliwość automatycznego dobierania mocy w miarę zapotrzebowania
- Granice w jakich można dynamicznie zmieniać parametry procesorów dla danego LPAR-a.

Pamięć

- Ilość pamięci przydzielonej LPAR-owi.
- Granice w jakich można dynamicznie przydzielać/odbierać pamięć.

Fizyczne urządzenia przydzielone LPAR-owi (wszelkie karty rozszerzeń).

Wirtualne urządzenia przydzielone LPAR-owi (Virtual SCSI/Network).

Co to jest DLPAR?

Dynamic logical partitioning (DLPAR) rozszerza możliwości LPAR o logiczne przydzielanie i zarządzanie zasobami maszyny bez restartowania jej, np.:

Processor

- Dołożenie/odebranie mocy procesora
- Dołożenie/odebranie wirtualnych procesorów
- Zmiana trybu pracy procesora

Pamięć

- Dołożenie/odebranie pamięci

Urządzenia fizyczne

- Dołożenie/odebranie dowolnych adapterów fizycznych

Urządzenia wirtualne

- Dołożenie/odebranie wirtualnych kontrolerów SCSI
- Dołożenie/odebranie przestrzeni dyskowej udostępnianej przez VIOS.
- Dołożenie/odebranie wirtualnego adaptera ethernet.

Po co to wszystko?

- Konsolidacja serwerów
- Środowiska produkcyjne i testowe: możliwość stworzenia środowiska do testowania i przenoszenia oprogramowania, które działa na dokładnie takiej samej platformie, jak środowisko produkcyjne
- Podział danych i operacji: możliwość obsługi zestawu różnorodnych aplikacji i danych z tego samego fizycznego serwera przy utrzymaniu ścisłego rozdziału dostępu do zasobów i danych.
- Równoważenie skalowania
- Elastyczność konfiguracji, szczególnie wygodnie obsługiwana przez mechanizm DLPAR.

Jak to wygląda to w IBM?

- Technologia dla mainframe'ów
- Platformy IBM:
 - iSeries, eSeries, pSeries, zSeries
- Procesory Power4, Power5
- Konsole HMC i IVM

Hardware Management Console:

- Jedna, czasami dwie konsole.
- Dostęp do konsoli bezpośredni, lub zdalny:
http://adres_konsoli/remote_client.html
- Dane o konfiguracji LPAR-ów przechowywane są na serwerze a nie konsoli.
- HMC umożliwia robienie backup-ów danych o konfiguracji LPARów
- Awaria konsoli nie ma wpływu na działanie serwera.

Integrated Virtualization Manager (IVM)

- Alternatywa dla HMC do stosowania w „mniejszych” serwerach pSeries.
- Posiada interface WWW (https) + wiersz poleceń.

Ograniczona funkcjonalność w stosunku do konsoli HMC, brak:

- Możliwości tworzenie wielu profili na LPAR.
- Możliwości zarządzania wieloma serwerami z jednego punktu.
- Możliwości przydziału fizycznych urządzeń do LPAR-a

Nie można stosować jednocześnie konsoli HMC i IVM

Definiowanie LPAR-a

Create Logical Partition Wizard

This wizard helps you create a new logical partition and a default profile for it. You can use the partition properties or profile properties to make changes after you complete this wizard.

Ensure you have your logical partition planning information before you use this wizard. You may also find it helpful to be familiar with logical partition concepts. Click Help for more information.

To create a partition, complete the following information:

System name : Server-9406-595-SN65CFC0F

Partition ID :

Partition name :

Partition environment

- AIX or Linux
- i5/OS
- Virtual I/O server

Create Logical Partition -- Workload Groups

if you plan to use a workload management application on your server, you can include this partition in a partition workload group. You can include the partition in a partition workload group by specifying the following :

Will this partition be included in a workload group?

- No
- Yes, this partition is in a workload group.

Partition workload group :

Definiowanie LPAR-a – pamięć

Create Logical Partition Profile

A profile specifies how many processors, how much memory, and which I/O devices and slots are to be allocated to the partition.

Every partition needs a default profile. To create the default profile, specify the following information :

System name: Server-9406-595-SN65CFC0F

Partition name: PREZENTACJA SO

Partition ID: 25

Profile name: default

This profile can assign specific resources to the partition or all resources to the partition. Click Next if you want to specify the resources used in the partition. Select the option below and then click Next if you want the partition to have all the resources in the system.

Use all the resources in the system.

Help ? < Back Next > Finish

Create Logical Partition Profile - Memory

Specify minimum, desired and maximum amounts of memory for this profile using a combination of the gigabyte and megabyte fields below.

Installed memory (MB): 131072

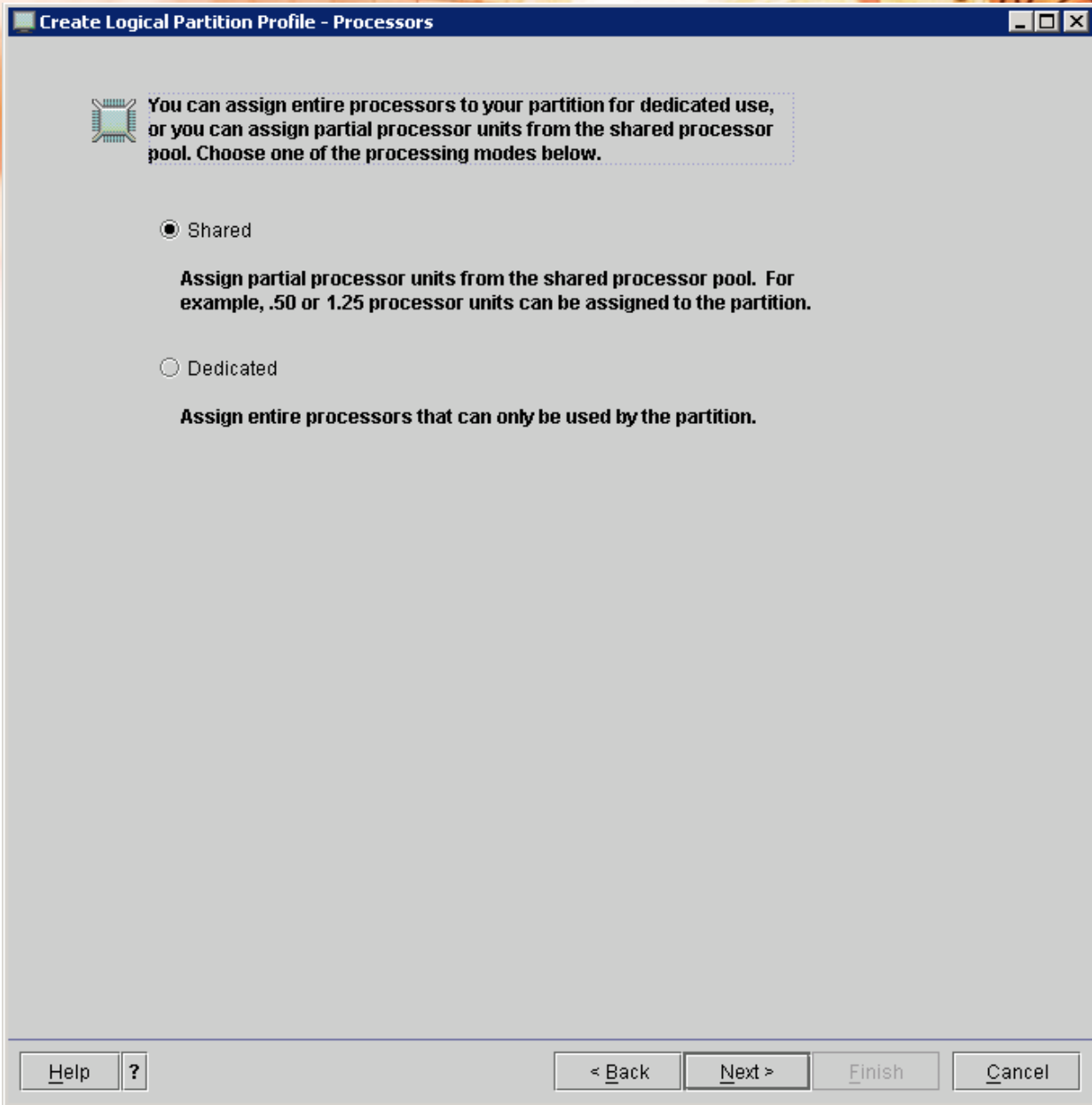
Current memory available for partition usage (MB) : 58112

Minimum memory	Desired memory	Maximum memory
0 GB	2 GB	4 GB
256 MB	0 MB	0 MB

Advanced Options

Show details


Help ? < Back Next > Finish Cancel



Definiowanie LPAR-a – tryb pracy procesora

Definiowanie LPAR-a – procesory

Create Logical Partition Profile - Processing Settings

 Specify the minimum, desired and maximum processing settings in the fields below.

Total usable processing units: 8.00

Minimum processing units:

Desired processing units:

Maximum processing units:

Advanced Processing Settings

Sharing modes

You must specify a processing sharing mode for this partition profile.

Capped

The processor usage never exceeds the assigned processing capacity.

Uncapped

Weight:

Processing capacity may be exceeded when the shared processor pool has spare processing power.

Virtual processors

The default virtual processor settings have been filled in for you. You may change the default settings below.

Minimum processing units required for each virtual processor:

Minimum number of virtual processors:

Desired number of virtual processors:

Maximum number of virtual processors:

Definiowanie LPAR-a – urządzenia fizyczne.

Create Logical Partition Profile - I/O

Select desired and required I/O components for this partition profile from the managed system I/O below. You can change the required attribute and specify the pool ID, if applicable, by double clicking the I/O Pool column.

Managed system I/O

Location Code	Description	Location Code
U5094.001.6	Unit U5094.001.6559789	
U5094.001.6	Bus 13	
U5094.001.6	Slot C11 PCI I/O Processor	
U5094.001.6	- Slot C12 PCI 10/100/1000Mbps Ethernet UT...	
U5094.001.6	- Slot C13 Empty slot	
U5094.001.6	- Slot C14 PCI Ultra Magnetic Media Controller	
U5094.001.6	- Slot C15 Empty slot	
U5094.001.6	Bus 14	
U5094.001.6	Slot C01 PCI I/O Processor	

I/O devices in the profile

Required	I/O Pool	Description	Loca
----------	----------	-------------	------

Create Logical Partition Profile - I/O Pools

You may edit, if applicable, the partition's I/O pool participation by adding and removing pool IDs from the partition I/O pools list.

You may add existing pool IDs and you may create pools to add by entering in the pool IDs in the I/O pools list.

I/O Pools

Number of I/O pools allowed : 1024

I/O pools :

Partition I/O Pools

Definiowanie LPAR-a – urządzenia wirtualne.

Create Logical Partition Profile - Virtual Adapters

SCSI Ethernet Serial

-- Virtual SCSI adapters

Type	Slot	Required	Remote Partition	Remote Slot
Client SCSI	2	Yes	Virtual IO Server 5(3)	23

Create server adapter... Create client adapter... P

-- Virtual Adapters

Slot	Type
0	Server Serial
1	Server Serial
2	Client SCSI

Maximum virtual adapt
Number of virtual adap

Help ? < Back Next > Finish Cancel

Create Logical Partition Profile - Virtual Adapters

SCSI Ethernet Serial

-- Virtual Ethernet adapters

Slot	Required	Bridged	VLAN
3	Yes	No	340

Create adapter... Properties... Delete

-- Virtual Adapters

Slot	Type
0	Server Serial
1	Server Serial
2	Client SCSI
3	Ethernet

Maximum virtual adapters : 10
Number of virtual adapters : 4

Help ? < Back Next > Finish Cancel

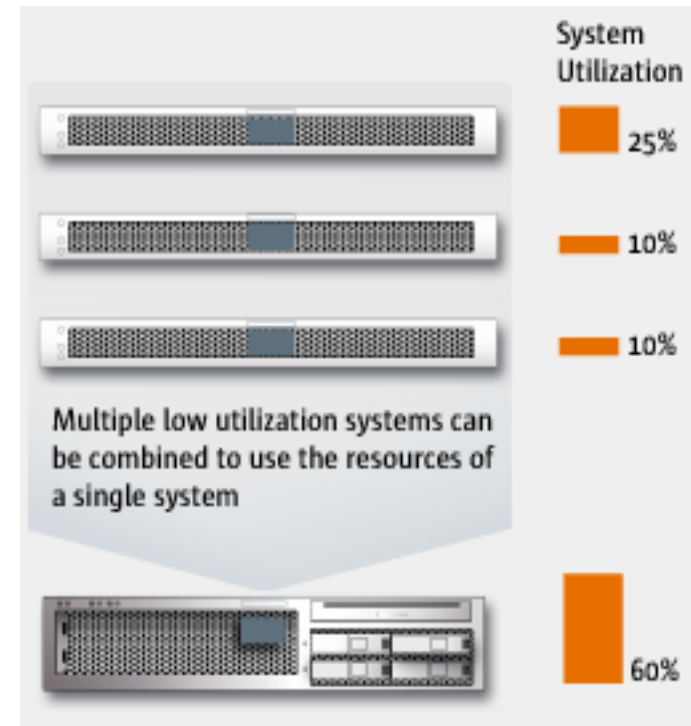


Limity i ograniczenia

- Maksymalna ilość LPAR-ów na serwer = **254**
- Maksymalna ilość wirtualnych procesorów na LPAR = **64**
- Minimalna ilość procesora na LPAR = **0,1**
- Maksymalna ilość mocy procesora na wirtualny procesor = **1**

Logical Domains (LDoms)

- Darmowa i otwarta technologia
- Łączy wiele serwerów fizycznych w jeden logiczny
- Zwiększa wykorzystanie mocy obliczeniowej
- Prostsza administracja i niższe rachunki za prąd
- ... razem do 10 tys. \$ na serwer



Wymagania dla Ldom 1.0.1

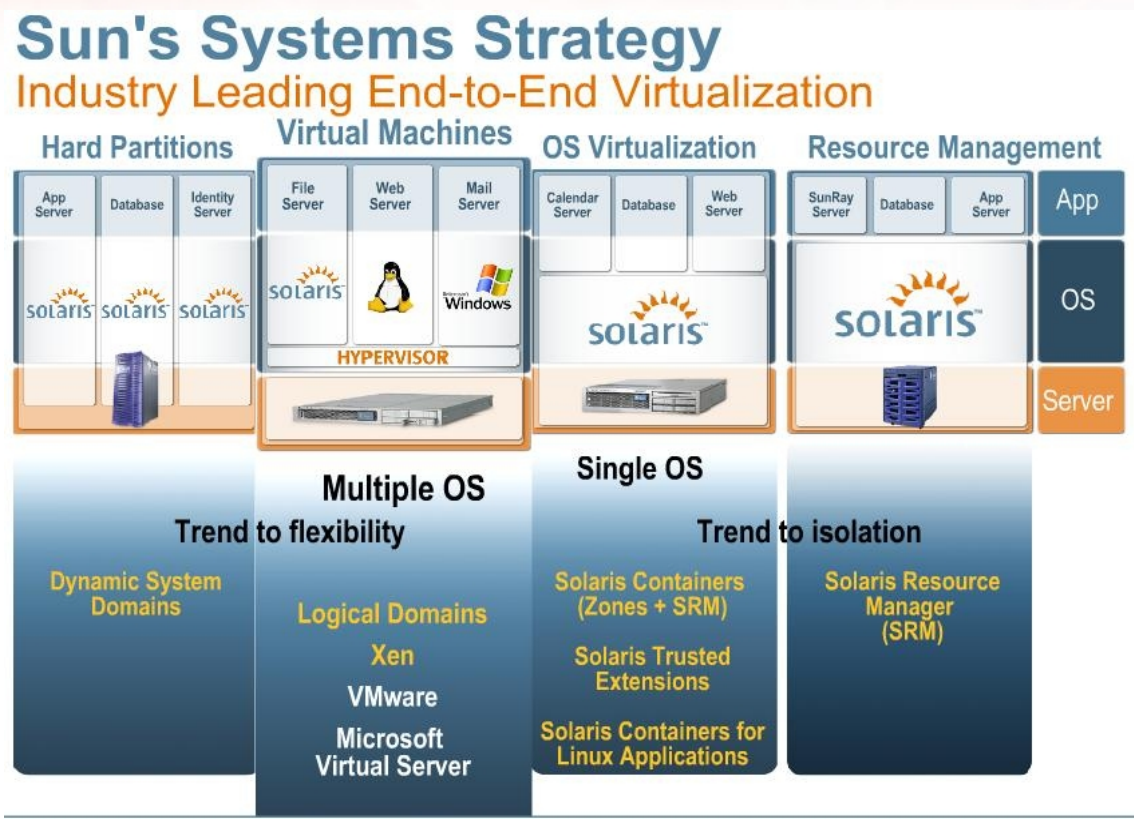
- Serwer wyposażony w procesor
 - **UltraSPARC T1**
 - **UltraSPARC T2**
- System **Solaris 10 11/ 06**
- ..lub inne systemy wzbogacone w specjalne moduły umożliwiające obsługę hypervisora
- Aktualne oprogramowanie firmware
 - 127580-01 at a minimum: **Sun SPARC Enterprise T5120 and T5220**
 - 127577-01 at a minimum: **Sun Fire and SPARC Enterprise T1000 Servers**
 - 127576-01 at a minimum: **Sun Fire and SPARC Enterprise T2000 Servers**
 - 127579-01 at a minimum: **Sun Blade T6300 Server Module**
 - 127578-01 at a minimum: **Netra T2000 Server**
 - 127575-01 at a minimum: **Netra CP3060 Blade**

Procesor UltraSPARC T2

- 8 SPARC V9 cores 1.4 Ghz
- 8 wątków na jądro (64)
- 4 mb L2 cache
- 64 obsługiwane domeny
- 1.5 watta na wątek
konkurencja ponad 30W



Ldom, a inne technologie



Ldom`s, a inne technologie

LDom	Dynamic Devices
Może używać części procesora (wątku – procesor=8 wątków)	Używa bloków złożonych z 4 procesorów
LDom oferuje większą elastyczność	

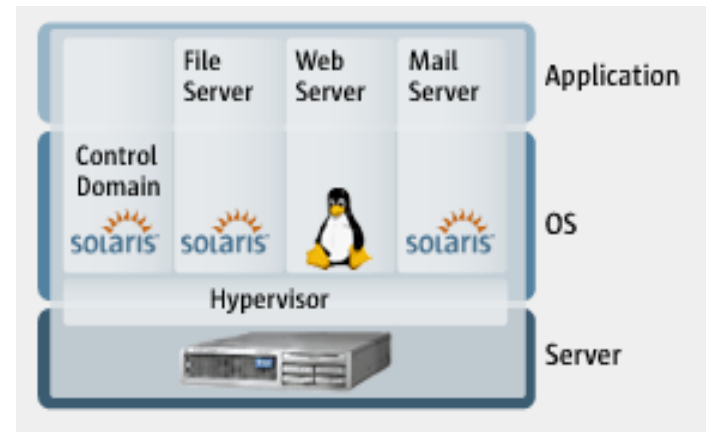
LDom	Solaris Containers
Oddzielne i niezależne systemy operacyjne (osobno bootowane)	Wirtualizacja w obrębie jednego Systemu Operacyjnego
LDom oferuje większą niezawodność	

LDom	VmWare, Xen
Wirtualizowany system wie o tym, (posiada specjalne moduły do kontaktu z hypervisorem)	Wirtualizowany system zachowuje się standardowo i wymaga zewnętrznego zarządzania zasobami
LDom oferuje większą wydajność	

Czym jest Logical Domain

Każda domena :

- Dla użytkowników sprawia wrażenie bycia odrębnym serwerem
- Posiada odrębny w pełni konfigurowalny system operacyjny
- Posiada własny zbiór zasobów tj. CPU, dysk, pamięć, urządzenia I/O
- Może być wyłączona i boot'owana niezależnie od innych domen





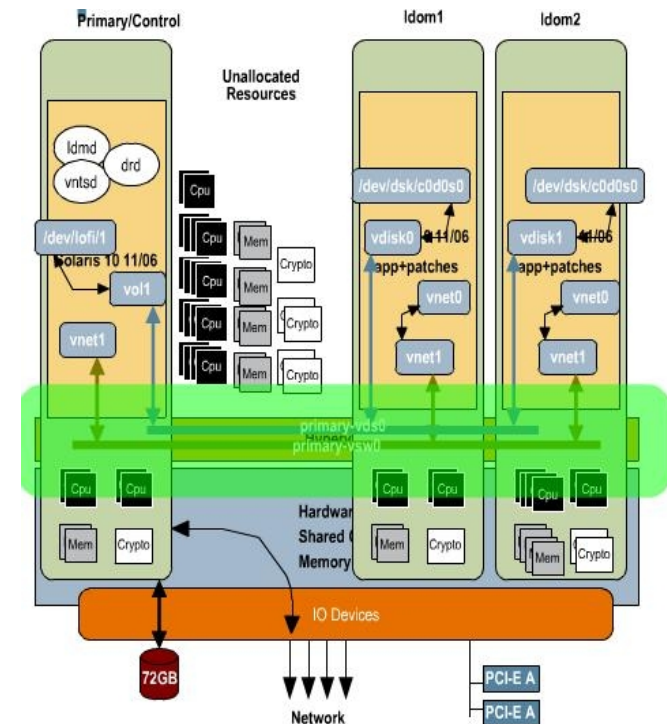
Kluczowe składniki

- Hypervisor
- Control Domain
- Service Domain
- I/O Domain
- Guest Domains
- Virtualized Devices

LDOM Hypervisor

Zapewnia niezależną i bezpieczną komunikację między użytkownikiem, a dostępnymi urządzeniami

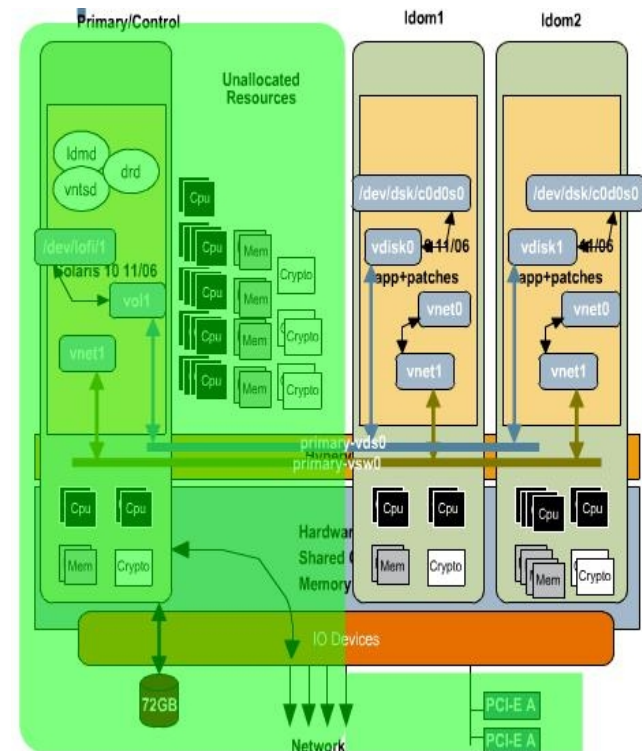
- Niskopoziomowy komponent będący oprogramowaniem typu firmware
- Udostępnia wirtualną sieć lokalną do komunikacji między domenami (LDC)
- Udostępnia fizyczny i wirtualny dostęp do urządzeń I/O
- Zawiera interfejs mapowania dla FMA (zarządzanie danymi o błędach)
- Zarządzanie konfiguracją z poziomu Control Domain



Control Domain

Główna domena umożliwiająca tworzenie i konfigurację pozostałych domen

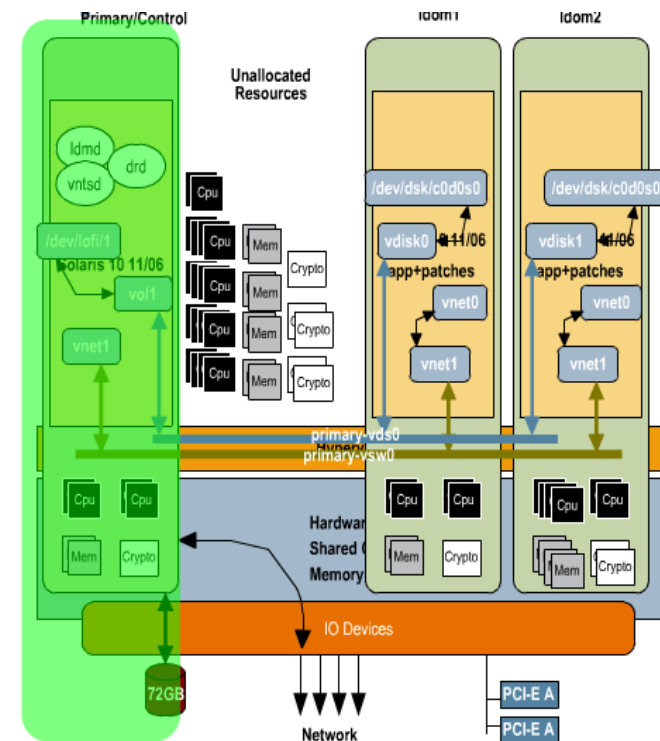
- Pierwsza domena aktywna po instalacji dająca możliwość tworzenia innych domen
- Działają na niej binaria aplikacji LDOM (dostęp przez komendę 'ldm')
- Zapewnia monitorowanie zasobów
- Poprzez hypervisora reguluje dostęp do zasobów
- Działa na niej system Solaris w minimalnej konfiguracji



Service Domain

Zarządza dostępem do urządzeń sprzętowych dla Guest Domains

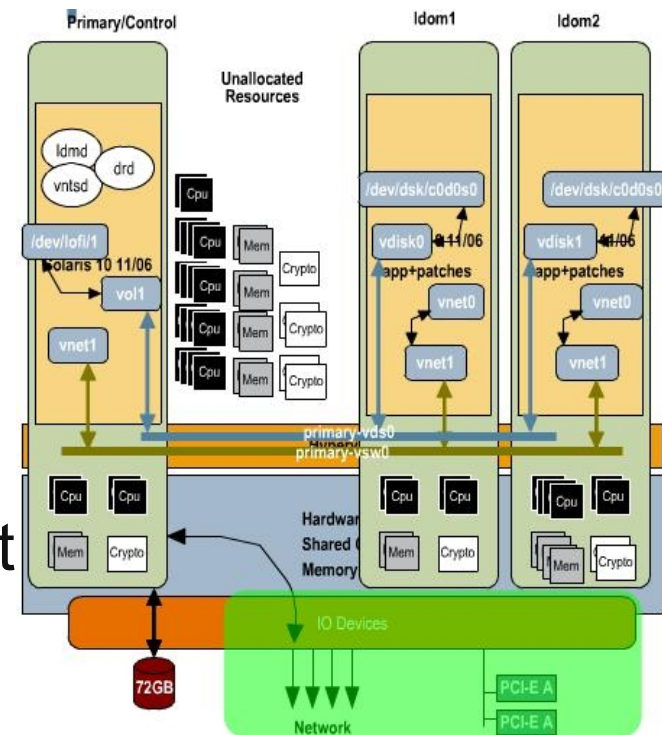
- Tworzy wirtualne usługi (services), np. dyski, terminale sieciowe
- Zapewnia dzielony lub wyłączny dostęp do urządzeń
- Po raz pierwszy utworzona Control Domain jest jednocześnie SD
- Może funkcjonować jako Guest Domain
- Posiada priorytetowy dostęp do jednej lub wielu szyn PCI-E



I/O Domain

Funkcjonuje podobnie do Service Domain, lecz zarządza tylko urządzeniami I/O

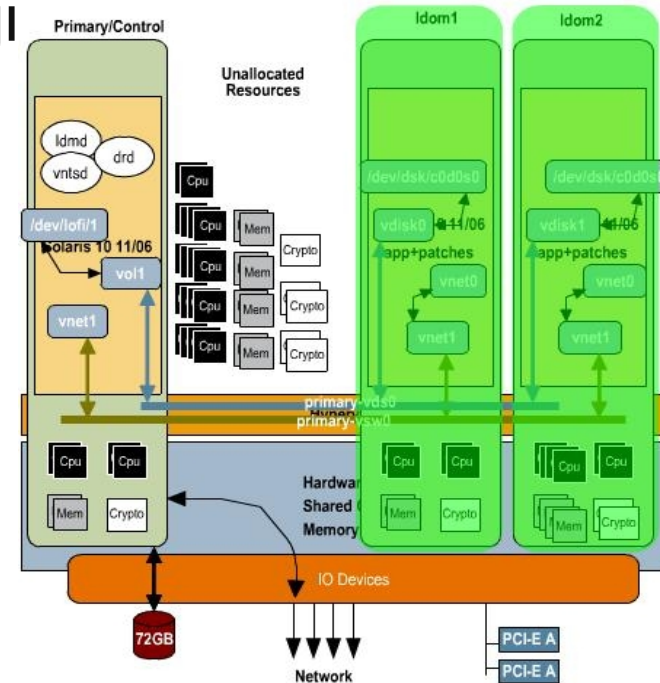
- Stwierdzono, że dostęp do urządzeń I/O jest na tyle ważny, że zasługuje na odrębną domenę
- Umożliwia bezpośredni dostęp do urządzeń I/O (sieci, dyski, PCI-E)
- Mogą jednocześnie działać maksymalnie dwie domeny I/O
- Nie udostępnia swoich zasobów w wirtualnej postaci domenom typu Guest
- Zastosowanie -> np. bazy danych



Guest Domains

Zapewnia środowisko działania dla uruchomionych aplikacji użytkownika

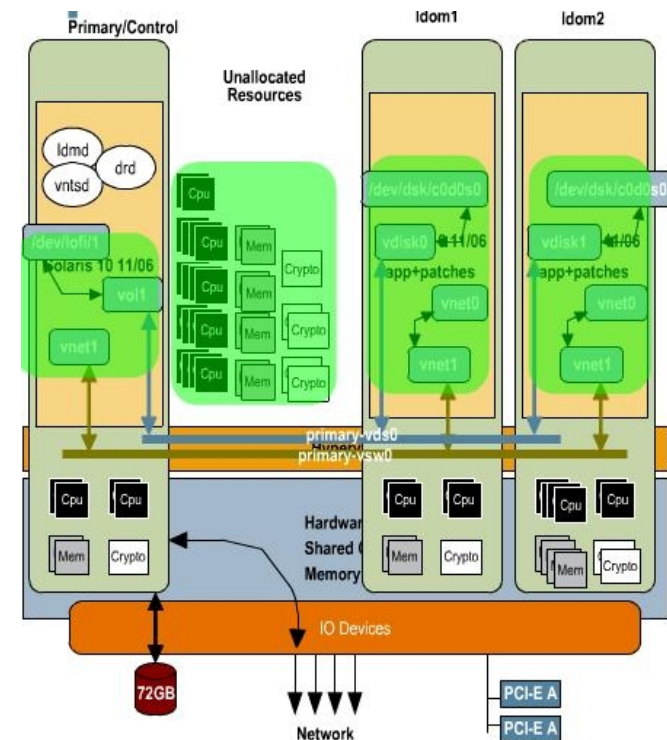
- Dla zwykłych użytkowników to „prawdziwe” domeny (pozostałe jedynie do konfiguracji)
- Zazwyczaj działa wiele niezależnych Guest Domain
- Mają dostęp do jednej bądź kilku usług (services)
- Mogą być dynamicznie konfigurowane z poziomu Control Domain
- Może na nich działać pełen system Solaris udostępniający kolejne narzędzia do wirtualizacji (Containers, SRM)



Virtualizes Devices

Wirtualna abstrakcja fizycznych urządzeń

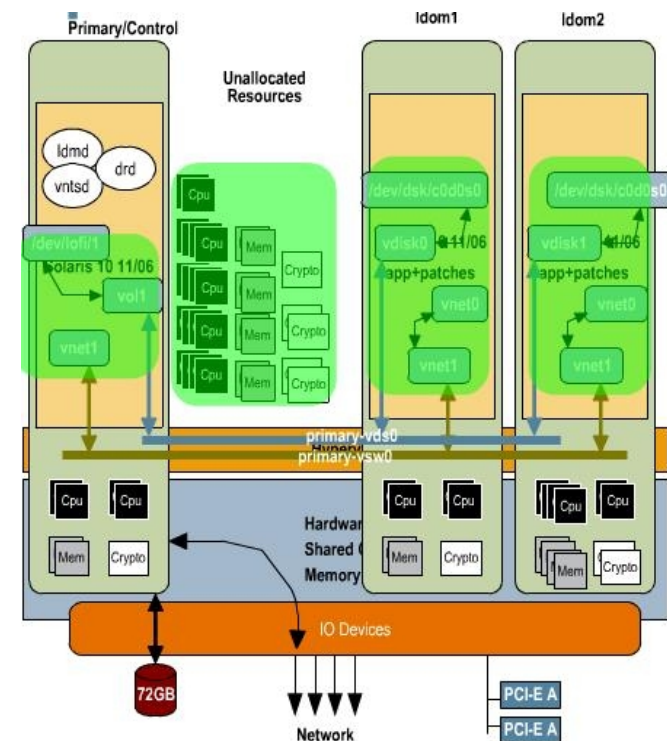
- Procesory (dynamiczna alokacja)
- Pamięć (**planowana** dynamiczna alokacja)
- Jądra kodujące (crypto cores)
- Przełączniki sieciowe
- Dyski
- Konsole
- NIC (Network Interface Controller)
- Terminal (Virtual Terminal Server)



Virtualizes Devices

Urządzenia są udostępniane w modelu klient serwer

- Do komunikacji służą Logical Domain Communication Channels zarządzane przez hypervisor (konfiguracja w CD)
- Wirtualny serwer abstrahuje zasoby: np. serwer dysku może być – kilkoma dyskami, pojedynczym dyskiem lub obrazem pliku
- Domeny funkcjonują na zasadzie klienta łączącego się z serwerem



Porady (Best Practices)

- Dokładnie zaplanuj konfigurację
- Sugestywnie nazywaj domeny i usługi
- Zabezpiecz oraz ogranicz do minimum instalację systemu na domenach Control i Services
- Dla lepszej wydajności zaalokuj wszystkie urządzenia za pomocą odpowiednio dopasowanych Service Domain
- Współpracuj z ZFS (umożliwia boot'owanie oraz robienie snapshot'ów)
- W trakcie tworzenia architektury pomyśl o podziale dostępu do zewnętrznych zasobów (np. internet)

Przykładowa konfiguracja

• Control Domain

- > 4 x vCPU's
- > 1 x Crypto units
- > 1 GB memory
- > 1 x e1000g0 NIC
- > 1 x Physical disk0
- > VNTSd ports 5000+
- > S10/JASS applied

• Service Domain

- > 8 x vCPU's
- > 1 x Crypto units
- > 1 GB memory
- > 2 x vsw0 vSwitch attached to e1000g1
- > no e1000gx plumbed
- > 1 x vds0 using physical disk1 under ZFS
- > S10/JASS applied

• Guest Domain

- > 8 x vCPU's
- > 2 x Crypto units
- > 4 GB memory
- > 2 x vnet0, vnet1 over vsw0
- > 1 x vdisk0 (lofi image via vds0)
- > Installed/config'd application
- > S10/ Patches applied

Jak to się robi?

Tworzenie domeny ldom2 z poziomu linii poleceń (CLI)

```
> ldm create ldom2
> ldm set-vcpu 6 ldom2
> ldm set-mem 8g ldom2
> ldm set-crypto 2 ldom2
> ldm add-vnet vnet1 primary-vsw0 ldom2
> ldm add-vdiskserverdevice /dev/dsk/c0t1d0 vol2@primary-vds0
> ldm add-vdisk vdisk0 vol2@primary-vds0 ldom2
> ldm bind ldom2
```

Uproszczenie BUI

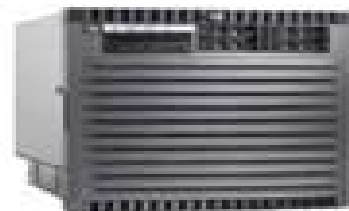
ang. Browser User Interface – ograniczone opcje w stosunku do linii poleceń; dostęp po wpisaniu <http://localhost:6789>

The image displays three overlapping screenshots of the Sun Logical Domain Management (LDM) web interface. The top-left screenshot shows the 'Common Administrative Tasks' menu, which includes options for domain management (View-Manage Domain, Create Domain, Edit Domain, Start Domain, Stop Domain, Delete Domains) and service management (View-Manage Network Service, View-Manage Disk Service, View-Manage Logical Service). The top-right screenshot shows the 'Virtual Switch Services (v)' table, which lists services and their associated network devices, MAC addresses, and nodes. The bottom-center screenshot shows a task list with 'View-Manage Domain' selected.

Service	Network Device	MAC Address	Node
primary.vsw0	lge2	14:05:322	-
primary.vsw1	lge1	14:05:221	-
LDM.vsw0	lge2	14:05:145	-
Full-Ext0	lrc0042	14:05:7ca	-

Technologia HP - VSE

- Virtual Server Environment – automatyczna infrastruktura umożliwiająca dynamiczną optymalizację wykorzystania zasobów serwera poprzez wirtualizację
- Przeznaczony do serwerów rodziny HP Integrity i HP 9000



Technologia HP - VSE

Komponenty :

- HP Capacity Advisor – zintegrowane narzędzie symulowania obciążenia aplikacji służące do lepszego planowania architektury.
- HP Virtualization Manager – zarządzanie wszystkimi zasobami oraz partycjami (fizyczne i wirtualne) za pomocą wygodnego interfejsu
- HP Global Workload Manager – monitoruje zużycie zasobów systemowych oraz dokonuje optymalizacji przenosząc zasoby z bezczynnych partycji do tych charakteryzujących się dużym obciążeniem

Technologia HP - VSE

Rodzaje partycji :

- HP Hard Partitions – partycjonowanie sprzętowe – system nPartitions (każda ma odrębne komórki z CPU i pamięcią) pełna izolacja sprzętowa
- HP Virtual Partions – partycjonowanie programowe, współistnienie kilku partycji na jednej nPartycji, każda ma przydzielone pewne zasoby, możliwa dynamiczna migracja zarówno procesorów jak i pamięci
- HP Integrity Virtual Machines – partycjonowanie programowe, oferuje izolację błędów, narzędzia do automatycznej alokacji zasobów i migracji pamięci przy awariach

parawirtualizacja

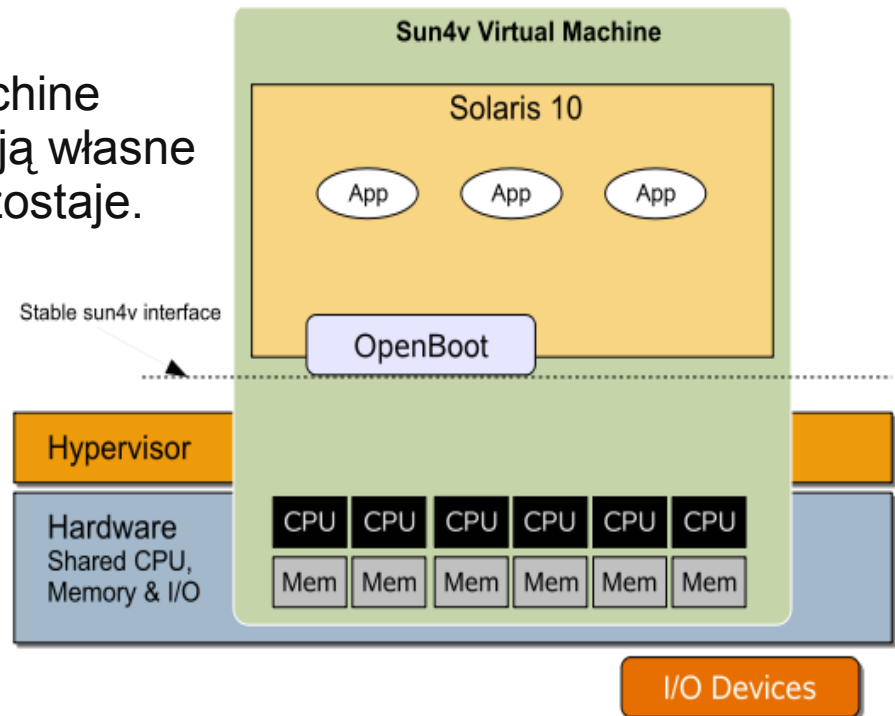
- udostępnienie rozszerzenia zestawu wywołań systemowych (*hypercalls*), które mogą być wykorzystane jedynie przez odpowiednio przygotowany system.
 - z tą metodą spotkamy się również w zastosowaniach dostępnych powszechniej niż firmy LPAR, np Xen.

Hypervisor

- Oprogramowanie (małe, zwarte!) stanowiące warstwę pośrednią pomiędzy sprzętem a partycjami, dostarczane jako firmware (np w pamięci flash).
 - abstrakcja sprzętu
 - dostarczenie jednolitego interfejsu architektury wirtualnej
 - we współpracy z systemami na partycjach dynamicznie zarządza przydziałem zasobów
 - dodatkowy poziom mapowania pamięci

jak to się je?

- (na podst. opisu API hypervisor do UltraSPARC T1, czyli dla architektury sun4v)
- hypervisor to inaczej Virtual Machine Monitor, różne firmy wprowadzają własne nazewnictwo, ale koncepcja pozostaje.



opis maszyny wirtualnej

- Machine description (MD) - udostępniane systemowi gościa
 - udostępniana jako jednolita struktura
 - unikalna dla gościa (w razie aktualizacji jest dostarczana nowa)
 - składa się z bloków czterech bloków
 - Header (opis pozostałych bloków),
 - Node block (węzły - podzielone wg zależności na kategorie, dane o dostępnych zasobach - cache, cpu/cpus, mblock, platform, etc.; w tym szczegóły - częstotliwość zegara, wspierane zestawy instrukcji, zakresy adresacji pamięci) ,
 - Name block (unikalne nazwy dla nazewnictwa wpisów węzła, ,
 - Data block (surowe dane do których może odnosić się Node block)

Domain Services

- instrukcje jakie może wywołać uprzywilejowane oprogramowanie:
 - `mach_exit`
 - `mach_desc` - pobranie opisu maszyny
 - `mach_sir` - software initiated reset (pamięć jest zachowywana, wybrany CPU jest przestawiany do punktu wejściowego SIR)
 - `mach_set_soft_state` - informowanie hypervisor o statusie maszyny (normalny/przejściowy), z krótkimi opisami.
 - ...
 - `mach_watchdog` - pierwsze wywołanie aktywuje watchdoga, w razie braku kolejnych wywołań hypervisor podejmie odpowiednie działania
 - warto pamiętać o wyłączaniu watchdoga przy debugowaniu :-)

CPU services

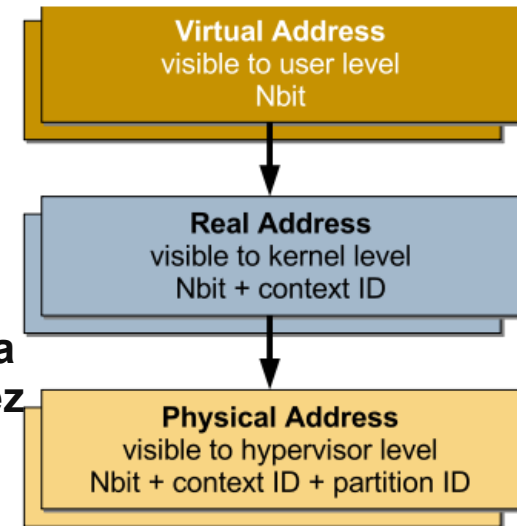
- CPU są zawsze w jednym z trzech stanów: stopped, running, error
- identyfikator CPU to 16bitowa unikalna liczba. operacje które mają być wykonane na wielu procesorach posiadają określoną listę procesorów.
 - **cpu_start**
 - **cpu_stop** (oczywiście nie wolno wyłączać bieżącego CPU, ani ostatniego dostępnego CPU - należy w tym wypadku użyć mach_exit)
 - **cpu_yield** - zawieszenie wykonania na bieżącym CPU, np do oszczędzania energii. przerwanie może zakończyć ten stan.
 - **cpu_modno_send** - wysłanie przerwania typu CPU do kolejek wskazanych CPU
 - **cpu_myid**
 - **cpu_state**

model obsługi przerwania

- Typy przerwania:
 - **CPU mondo** **CPU to CPU interrupt message.**
 - **Device mondo** **interrupt sent by an I/O device.**
- są 2 kolejki przerwania: dla CPU i dla urządzeń. nowe przerwania są dopisywane na końcach kolejek (w postaci interrupt reports). dopisanie nowego przerwania do kolejki generuje przerwanie rodzaju `cpu_mondo` lub `dev_mondo`. ich obsługa może zostać zablokowana.

pamięć

- zastosowanie LPAR sprawia, że jest dołożony dodatkowy poziom mapowania pamięci (fiz. - rzeczywiste)
- w architekturze sun4v:
 - domyślnie wyłączona translacja adresu wirtualnego, odwołania wykonywane przy użyciu adresów rzeczywistych.
 - gdy uruchomić translację, może być wykonywana na jeden z kilku dostępnych sposobów (nie-wprost przez TSB; wprost przez TLB przy pomocy specjalnych funkcji, mapowanie stałe). udostępniane funkcje do sterowania MMU.



a gdy coś nie gra...

- przy debugowaniu hypervisor przysłania sprzęt, więc utrudnia diagnozowanie błędów. żeby nie dopisywać do hypervisora funkcjonalności rzucania *core dump*, połączono tę akcję ze zrzutem z gościa - gość udostępnia hypervisorowi część swojej pamięci, gdzie ma on zapisać informacje przydatne przy debugowaniu (wymagany rozmiar pamięci jest znany z Machine Description).
 - **API hypervisora udostępnia funkcję `dump_buf_update`.**

komunikacja - Logical Domain Channel services

- **brak adresacji, przesyłanie krótkich pakietów z danymi**
- **brak kontroli transmisji, musi być dołożony dodatkowy protokół.**
- ze względu na ograniczenia możliwości bezpośredniego przypisania urządzeń do domen (np. PCI-E mogą być przypisane co najwyżej dwóm różnym domenom), lepiej jest wirtualizować zasoby - dostęp do nich uzyskując przez LDC.
- odebrane dane trafiają do kolejki w pamięci gościa.

komunikacja - sieci

- urządzenie vnet (wirtualna karta ethernet)
- urządzenie vsw (switch wirtualny)
- możliwość tworzenia sieci wydzielonych (również skomplikowanych konfiguracji, z licznymi switchami wirtualnymi)
- dostarczana jest pula adresów MAC, z których można korzystać adresując partycje

rekonfiguracja partycji

- może być:
 - dynamiczna
 - opóźniona

inne funkcje hypervisora, inne

- obsługa pamięci (mem_scrub, mem_sync), obsługa przerwań urządzeń, zegar, konsola, dostęp do urządzeń (PCI, PCI-X, PCI_Express, etc).
- OpenBoot PROM – warto wspomnieć

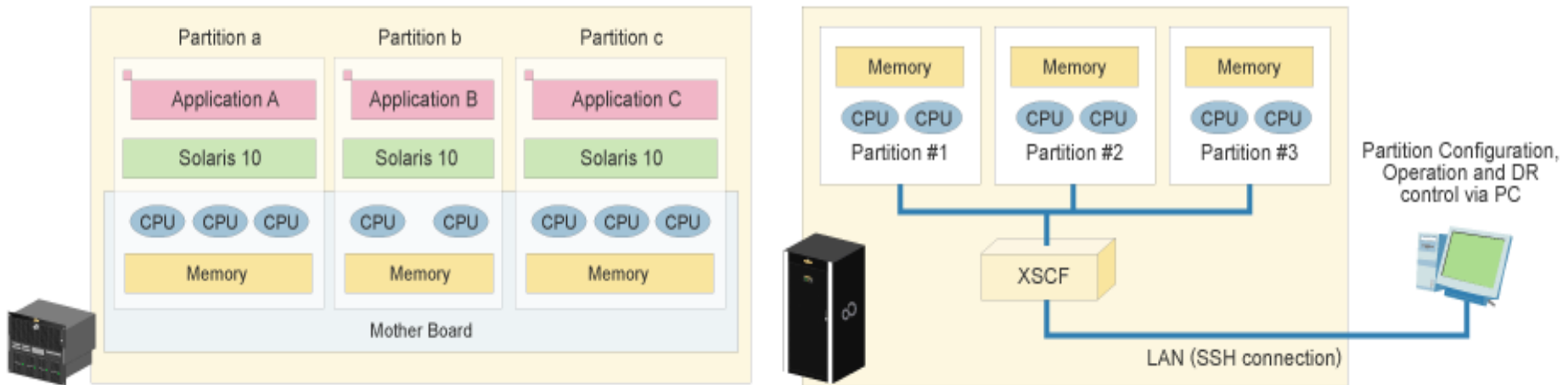


architektura wirtualna

- systemy mające pracować w LPAR powinny umieć korzystać z tych funkcji (przynajmniej z części). w praktyce taka specyfikacja przyjmuję nazwę osobnej architektury, np:
 - **POWER (IBM; wsparcie m.in. Linux)**
 - **sun4v (SUN; Solaris, Ubuntu, FreeBSD)**

Fujitsu

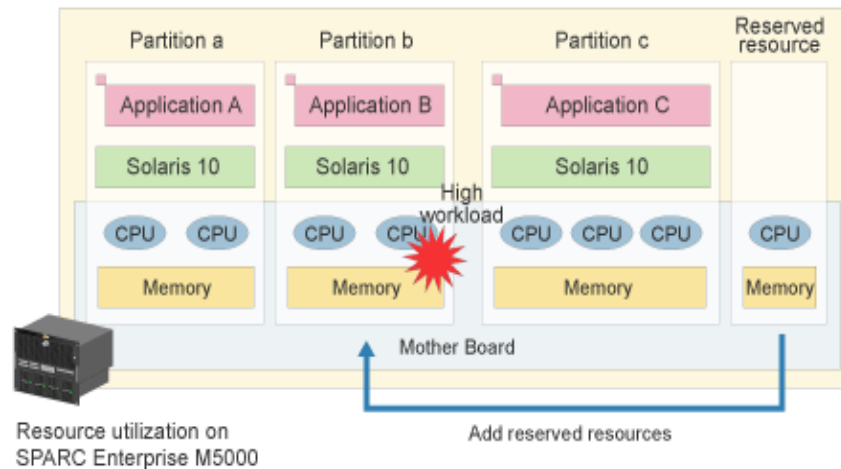
- rozwiązania Fujitsu dla serwerów SPARC Enterprise, partycje logiczne zarządzane są dynamicznie przez XSCF (extended system control facility)



Partitioning on SPARC Enterprise M5000

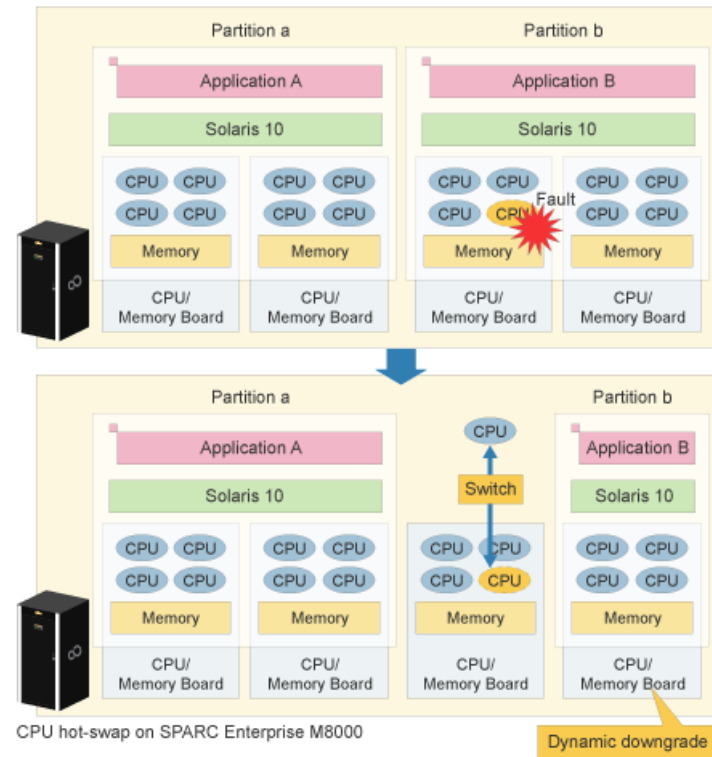
Fujitsu

- w przypadkach dużego obciążenia zarezerwowane zasoby mogą być łatwo dołączone do partycji



Fujitsu

- dla partycji obejmujących więcej niż jeden moduł CPU-pamięć, można przeprowadzić dynamiczny downgrade. w przypadku awarii – system działa dalej.



Źródła:

- Wikipedia
- <http://unixdays.pl/i/unixdays-prezentacje/2007/biedron.pdf>
- <http://www-03.ibm.com/systems/pl/p/lpar/>
- <http://opensparc-t1.sunsource.net/specs/Hypervisor-api-current-draft.pdf>
- http://ideasint.blogs.com/ideasinsights/2006/10/sun_logical_dom.html
- <http://www.sun.com/servers/coolthreads/ldoms/index.xml>
- <http://www.idg.pl/artykuly/54550.html>
- <http://opensolaris.org/os/community/ldoms/>

Dziękujemy za uwagę



Czy są pytania?