

Prezentacja LPAR

- Przygotowana przez:
 - Konrada Błachnio
 - Pawła Koniecznego
 - Michała Lenarta

Plan prezentacji:

- Co to jest LPAR i DLPAR
- Zastosowania
- Trochę historii
- Opis konkretnych rozwiązań
- Podsumowanie

Co to jest LPAR

Partycja logiczna – podzbiór zasobów maszyny, wirtualizowany jako jeden system.

Technologia LPAR umożliwia:

- Uruchomienie wielu systemów operacyjnych na jednej maszynie
- Odizolowanie od siebie poszczególnych systemów
- Przydzielanie zasobów dla poszczególnych partycji logicznych

Co to jest DLPAR

DLPAR – możliwość dynamicznej rekonfiguracji zasobów.

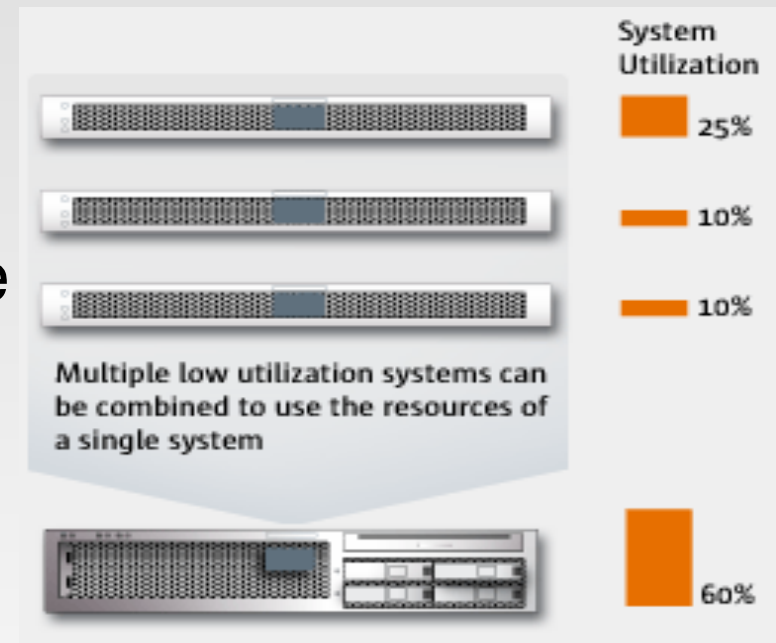
- Umożliwia zmianę przydziału CPU, pamięci, urządzeń I/O bez restartu systemu operacyjnego
- Wymaga wsparcia ze strony sprzętu i systemu operacyjnego

DLPAR - jak to działa

- **Hypervisor** – zarządca partycji. Umożliwia przydział zasobów sprzętowych poszczególnym systemom.
- Może działać w warstwie firmware'u (np. Sun Logical Domains, IBM LPAR)
- Może występować także jako część oprogramowania uruchamiana przed właściwym systemem operacyjnym (np. Xen)

Do czego to jest przydatne

- **Konsolidacja** - mając kilka maszyn marnujemy zasoby – każdy system wykorzystuje tylko niewielką ich część
- Możliwość reagowania na zmienne obciążenie systemów umożliwia lepsze wykorzystanie zasobów
- Oszczędność (sprzęt, energia, administracja)
- Używane do konfiguracji maszyn serwerowych



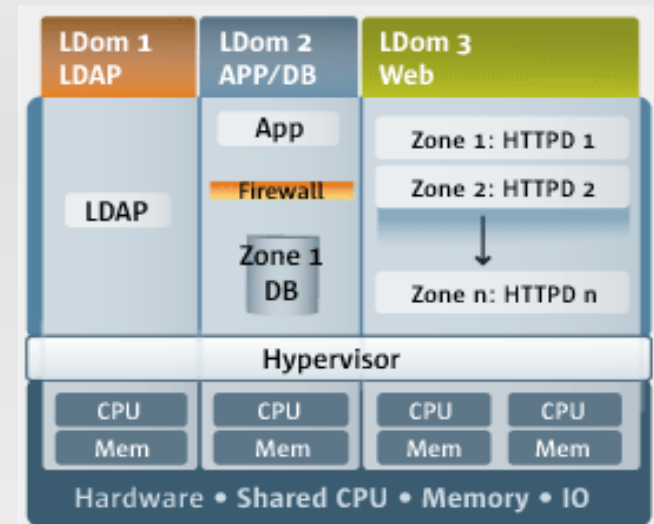
Do czego to jest przydatne

Środowiska produkcyjne i testowe:

Możliwość stworzenia środowiska do testowania i przenoszenia oprogramowania, które działa na dokładnie takiej samej platformie, jak środowisko produkcyjne (co zapewnia pełną zgodność), ale jest od niego całkowicie odizolowane.

Bezpieczeństwo

Przykład – firma używa serwerów www oraz bazy danych i serwera plików z poufnymi danymi.



- Aby zapewnić odpowiedni poziom bezpieczeństwa, należy odizolować serwery www od bazy danych
- Logiczne partycjonowanie systemów zapewnia taką izolację – każdy system ma niezależne zasoby

Skąd to się wzięło

- Technologia stworzona w okolicach roku 1990
- Pierwsze implementacje wykonane przez IBM
- Początkowe zastosowania dla serwerów klasy mainframe (architektury ESA/390, zSeries, System 9)
- Następnie idea rozszerzona również na maszyny nie będące mainframe'ami (pSeries, iSeries)
- Systemy wspierające LPAR to w chwili obecnej z/OS, z/VM, z/VSE, z/TPF, AIX, Linux, Windows oraz i5/OS

Wprowadzenie do IBM LPAR

- zakłada izolację systemów umieszczonych na poszczególnych partycjach logicznych (LPAR)
- umożliwia wirtualizację na poziomie firmware oraz sprzętu (specjalne procesory)
- daje możliwość dzielenia zasobów maszyny między różne systemy działające na niej równolegle i zupełnie niezależnie
- ogromne znaczenie biznesowe
- mnogość rozwiązań ułatwiających administrację

Sprzęt IBM obsługujący LPAR

- pSeries (Linux, AIX)
- eSeries i5
- iSeries (i5/OS, AIX, Linux, Windows Server)
- OpenPower
- BladeCenter

- zSeries (z/OS, z/VM, z/VSE, VSE/ESA, TPF, Linux)

Niektóre elementy Advanced POWER

Virtualization

- POWER Hypervisor
- Micro-Partitioning (AIX lub Linux)
- Virtual I/O Server
- Partition Load Manager
- Live Partition Mobility (nowość, dla POWER6)

POWER Hypervisor

- serce wirtualizacji, zarządza partycjami logicznymi
- tworzy wirtualną sieć LAN między partycjami
- zarządza pulą zasobów przydzielanych poszczególnym partycjom
- zarządza przerwaniem sprzętowymi
- zapewnia spójność partycji (izolację i niezależność)
- sterowany np przez HMC (Hardware Management Console)

Sterowanie Hypervisorem

Integrated Virtualization Manager

Welcome admin [Edit my profile](#) | [Help](#) | [Log out](#)

Partition Management

- [View/Modify Partitions](#)
- [Create Partitions](#)
- [View/Modify System Properties](#)

Virtual Ethernet Management

- [View/Modify Virtual Ethernet](#)

Storage Management

- [View/Modify Devices](#)
- [Advanced View/Modify Devices](#)
- [Create Devices](#)

User Management

- [View/Modify User Accounts](#)
- [Create User Accounts](#)

Service Management

- [Service Focal Point](#)
 - [Manage Serviceable Events](#)
 - [Service Utilities](#)
 - [Create Serviceable Event](#)
 - [Manage Dumps](#)
 - [Collect VPD Information](#)
- [Updates](#)
- [Backup/Restores](#)
- [Application Logs](#)
- [Hardware Inventory](#)

View/Modify Partitions

To perform an action on a partition, first select the partition or partitions, and then select the task.

System Overview

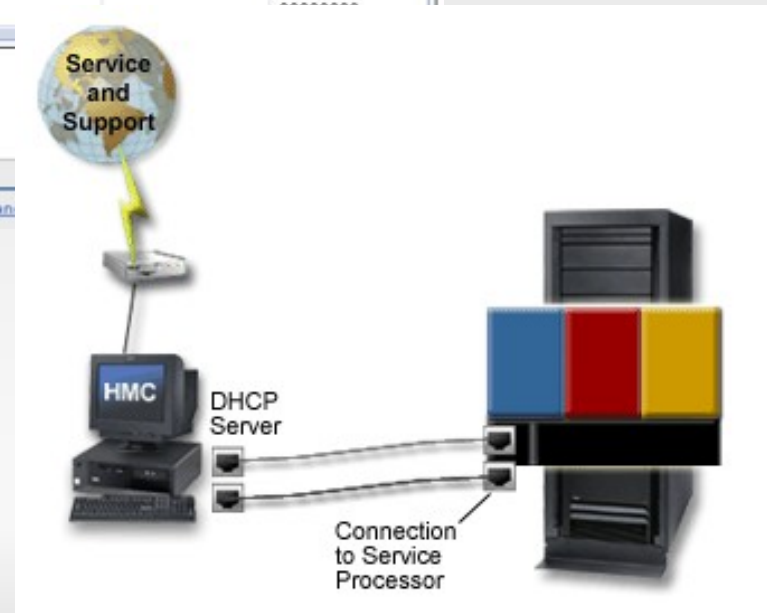
Total system memory:	8 GB	Total processing units:	2
Memory available:	1.78 GB	Processing units available:	1.1
Reserved firmware memory:	224 MB	Processor pool utilization:	7.7%
System attention LED:	Inactive		

Partition Details

Select	ID	Name	State	Uptime	Memory	Processors	Processing Units	Processor Utilization	Reference Code
<input type="checkbox"/>	1	eServer-p505	Running	7.26 Days	1 GB	2	0.2	22.2%	
<input type="checkbox"/>	2	AIX 5.3	Running	7.25 Days	1 GB	2	0.2	1.7%	
<input type="checkbox"/>	3	SuSE SLES 9sp2	Running	7.25 Days	1 GB	2	0.2	21.7%	SuSE Linux ppc64
<input type="checkbox"/>	4	RedHat AS EL4 u1	Running	7.25 Days	1 GB	2	0.2	19.7%	Linux ppc64
<input type="checkbox"/>	5	SUSE 9 SP2	Not Activated		2 GB	1	0.1	

Tasks

[Properties](#) | [Activate](#) | [Shutdown](#) | [Delete](#) | [Operator panel service functions](#) | [Referen](#)



Partition Load Manager

- automatyzuje proces rozdzielania zasobów pojedynczej maszyny między partycje logiczne
- zarządza CPU i pamięcią na podstawie konfiguracji sporządzonej przez administratora
- optymalizuje obciążenie poszczególnych partycji

Live Partition Mobility

- umożliwia przenoszenie działających systemów z jednej maszyny na drugą z zapamiętanym stanem procesora, pamięci, połączonymi użytkownikami itd.
- znacznie zwiększa dostępność i niezawodność systemu
- daje możliwość szybkiej reakcji na zwiększone obciążenie poprzez automatyczne przeniesienie systemu na mniej obciążoną maszynę
- wykorzystuje IBM High Availability Cluster Multiprocessing

Prezentacja idei DLPAR na podstawie rozwiązań IBM

- Co chcielibyśmy aby DLPAR oferowało
 - lista propozycji
- Co dostajemy do ręki
 - lista dostępnych mechanizmów
- Do czego to jest przydatne
 - większa przydatność DLAPR niż LPAR

Czego byśmy sobie życzyli

•Skoro dynamicznie to:

- Dodanie procesora
- Odjęcie procesora (zabranie ostatniego?)
- Przydzielenie dodatkowej pamięci
- Zabranie pamięci
- Dodanie/usunięcie urządzeń I/O
- Co jeszcze jesteśmy w stanie sobie wymyślić

Co dostajemy do ręki

• IBM dostarcza mechanizmy:

- Virtual I/O Serwer (VIOS)
- Virtual Ethernet (VIOS)
- CoD - (Capacity on Demand)
- Dynamic CPU Guard
- Micro-Partitioning

Virtual I/O Server

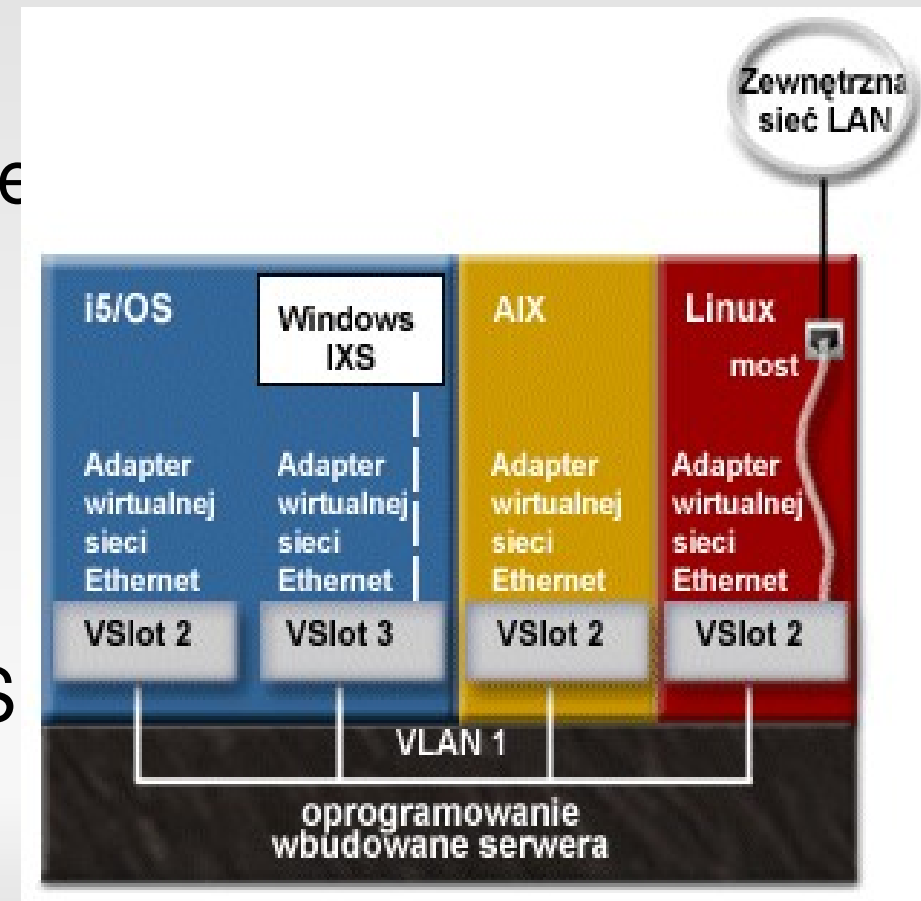
- VIOS to specjalna partycja udostępniająca fizyczne urządzenia pozostałym partycjom poprzez wirtualne interfejsy
- zarządzanie zasobami przydzielonymi partycjom
- zarządzanie gniazdami Ethernet, SCSI oraz urządzeniami storage
- współdzielenie jednego urządzenia

Virtual I/O Server

- wirtualizowanie wewnętrznych dysków SCSI
- wirtualizowanie zewnętrznych dysków SCSI
- dynamiczne dodawanie i usuwanie urządzeń dyskowych
- mały spadek szybkości obsługi urządzeń dyskowych - parę procent
- małe zaangażowanie procesora

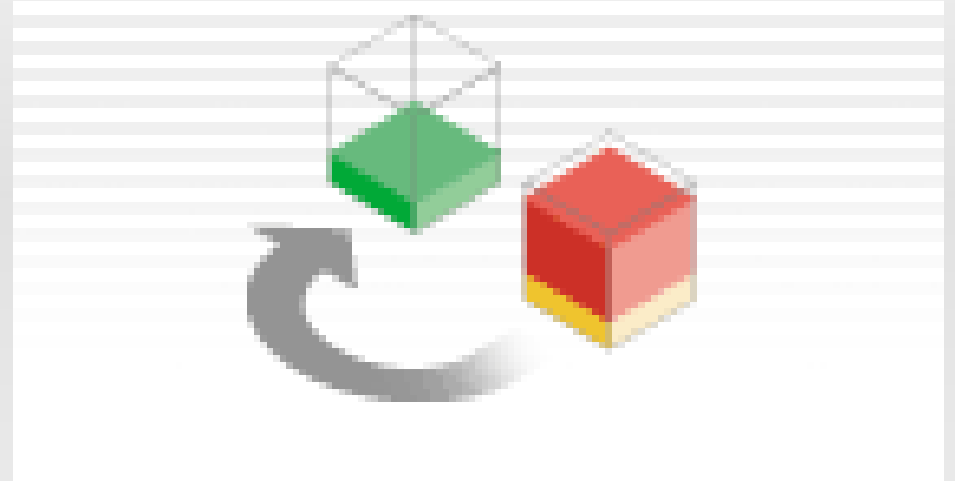
Virtual Ethernet

- korzysta z partycji VIOS
- do 4094 osobnych wirtualnych sieci LAN
- do 32 767 adapterów wirtualnej sieci Ethernet
- duże zużycie procesora
- w obrębie jednego LPAR porozumiewanie się bez VIOS



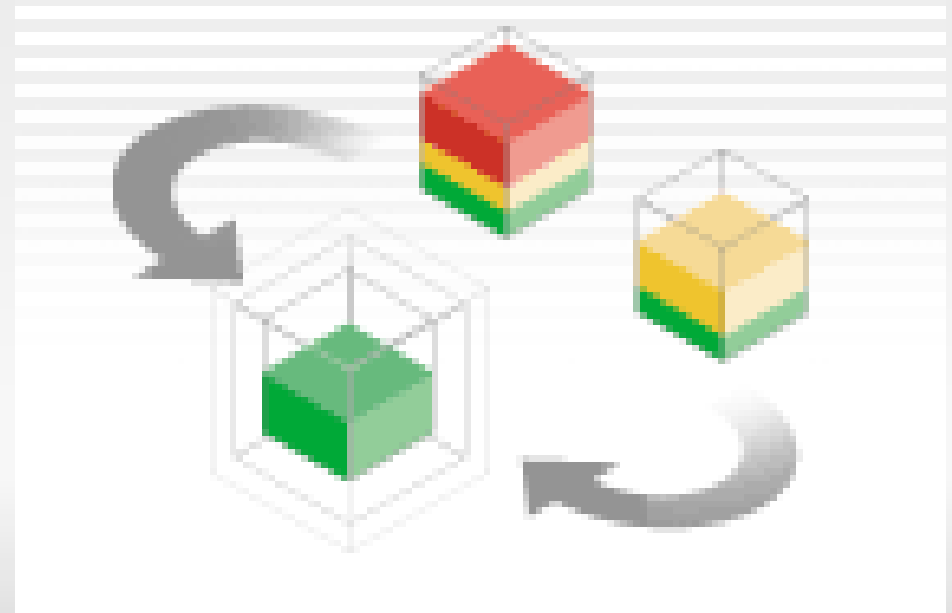
Capacity on Demand

- przenoszenie procesorów i pamięci pomiędzy partycjami
- zmiany nie wymagają restartowania partycji
- nie można zabrać ostatniego procesora
- nie można odebrać minimalnej jednostki pamięci



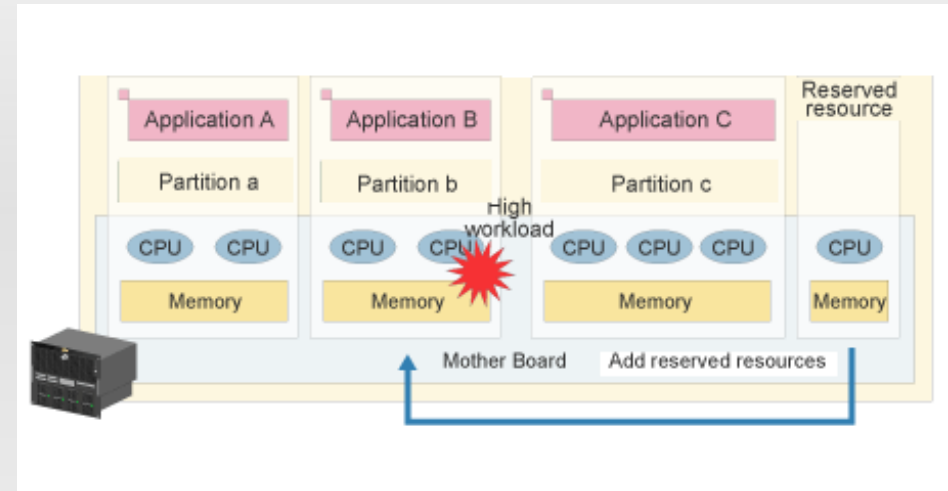
Capacity on Demand

- można odebrać przydzielony wcześniej procesor
- w razie awarii możliwość wymiany źle działającego procesora (Dynamic CPU Guard)



Dynamic CPU Guard

- monitoruje działanie procesorów
- w przypadku zaobserwowania nieprawidłowości odłącza procesor zanim ten zawiedzie
- podłącza procesor z puli wolnych procesorów
- z punktu widzenia aplikacji nie ma zauważalnych zmian
- cały proces nie wymaga podejmowania żadnych akcji od administratora



Micro-Partitioning

Mikropartycjonowanie jest to zdolność dzielenia procesorów fizycznych oraz współużytkowanie ich przez wiele logicznych partycji

Micro-Partitioning IBM

- Tryb użytkowania partycji limitowany (capped) czyli zużycie mocy procesora nie może przekroczyć jej przypisanej mocy przetwarzania
- Tryb użytkowania partycji nielimitowany (uncapped) czyli zużycie mocy procesora może przekroczyć przypisaną jej początkowo moc przetwarzania

Micro-Partitioning IBM

- Procesor dedykowany jest przypisany do danej partycji nie może być wykorzystywany przez inną partycję
- Procesor współużytkowany może być dynamicznie przydzielany do innej partycji jeśli obciążenie się zmieniło

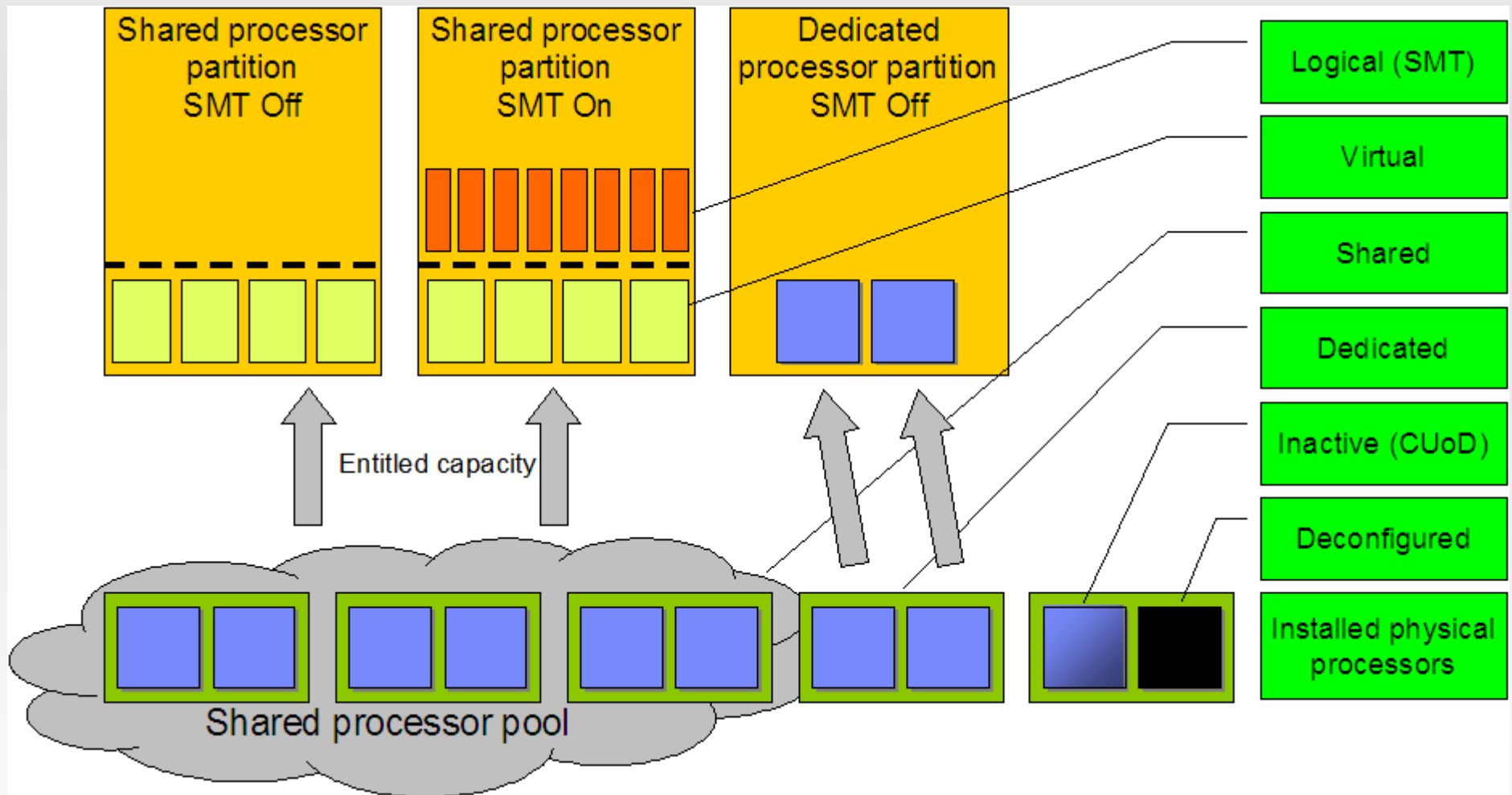
Micro-Partitioning IBM

- partycja nie może korzystać jednocześnie z procesorów dedykowanych i współużytkowanych
- partycja może mieć minimalnie 0.1 mocy jednego wirtualnego procesora
- maksymalnie 64 wirtualne procesory
- jeden procesor wirtualny to maksymalnie jeden procesor fizyczny

Micro-Partitioning

- partycje nielimitowane mogą mieć różne wagi jeśli chodzi o kolejność przyznawania mocy procesorów wirtualnych (0-255)
- procesory wolne mogą przełączać się pomiędzy trybami dedykowanym i współdzielonym
- na serwerze może być co najwyżej 256 partycji
- jedna partycja może mieć przydzielone nie więcej niż 64 wirtualne procesory

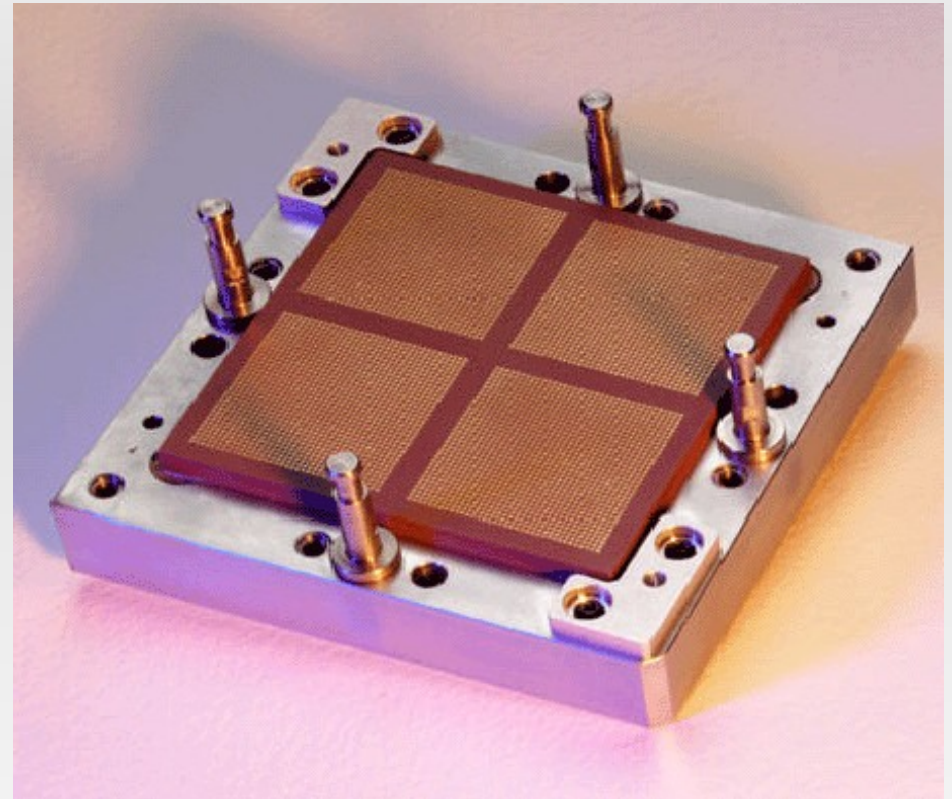
Micro-Partitioning



Co dostajemy do ręki

• Procesor POWER4

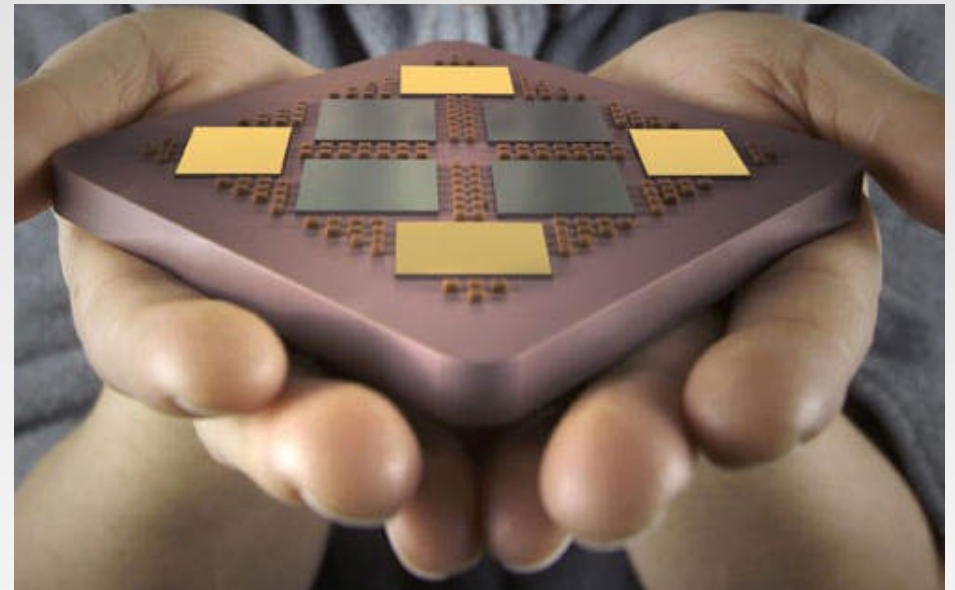
- dwa rdzenie
- dwie partycje
- mechanizm CoD
- Dynamic CPU Guard
- już od 439 euro



Co dostajemy do ręki

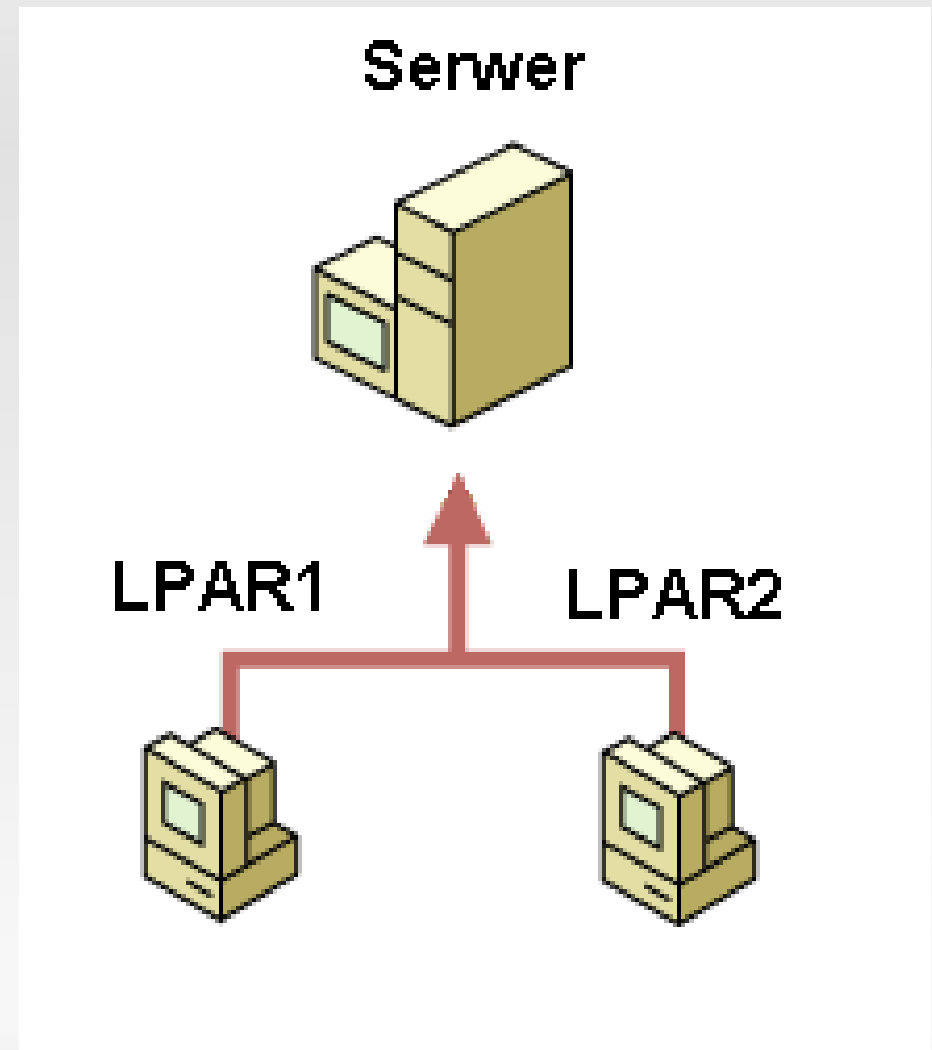
• Procesor POWER5

- dwa podwójne rdzenie
- dziesięć partycji
- mechanizm CoD
- Dynamic CPU Guard
- micro-partitioning
- już od 4.341 dolarów



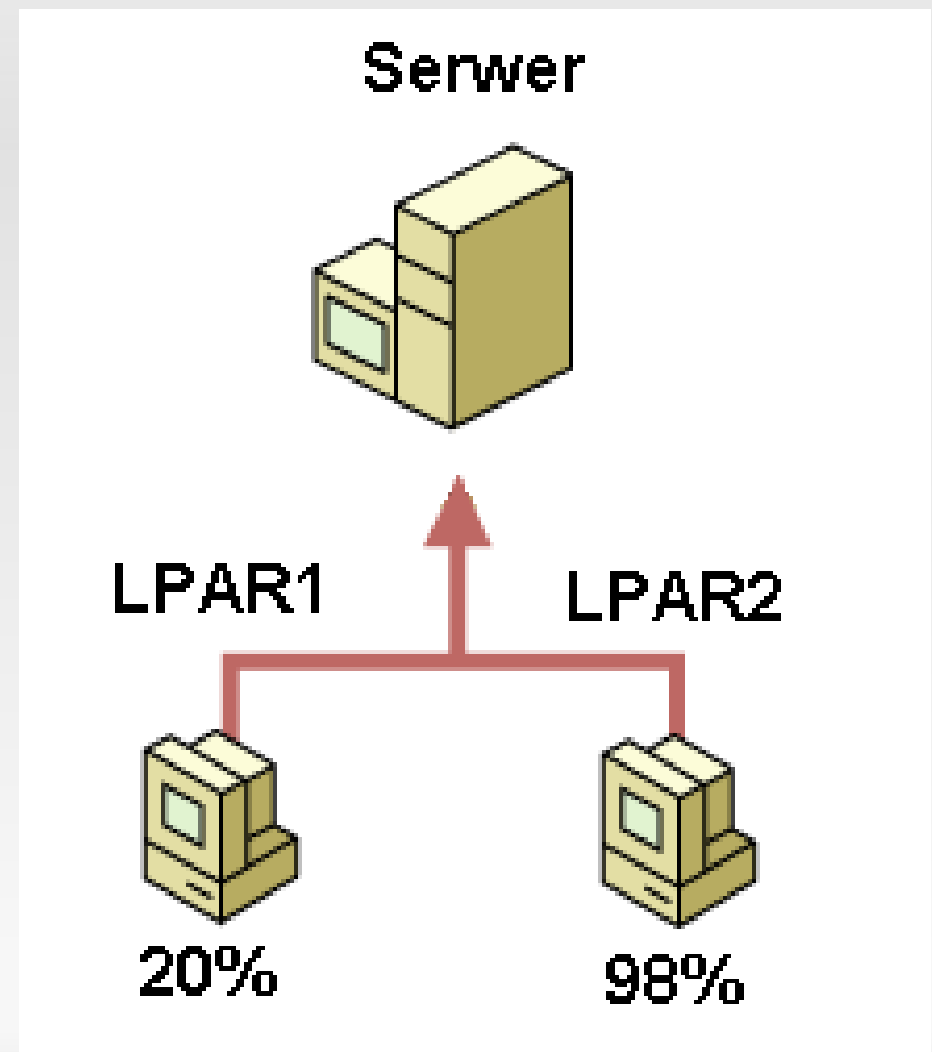
Do czego to jest przydatne

- jedna fizyczna maszyna
- dwóch klientów
- dwa logiczne serwery
- dzięki LPAR
- dzięki DLPAR
- mikroparycjonowanie



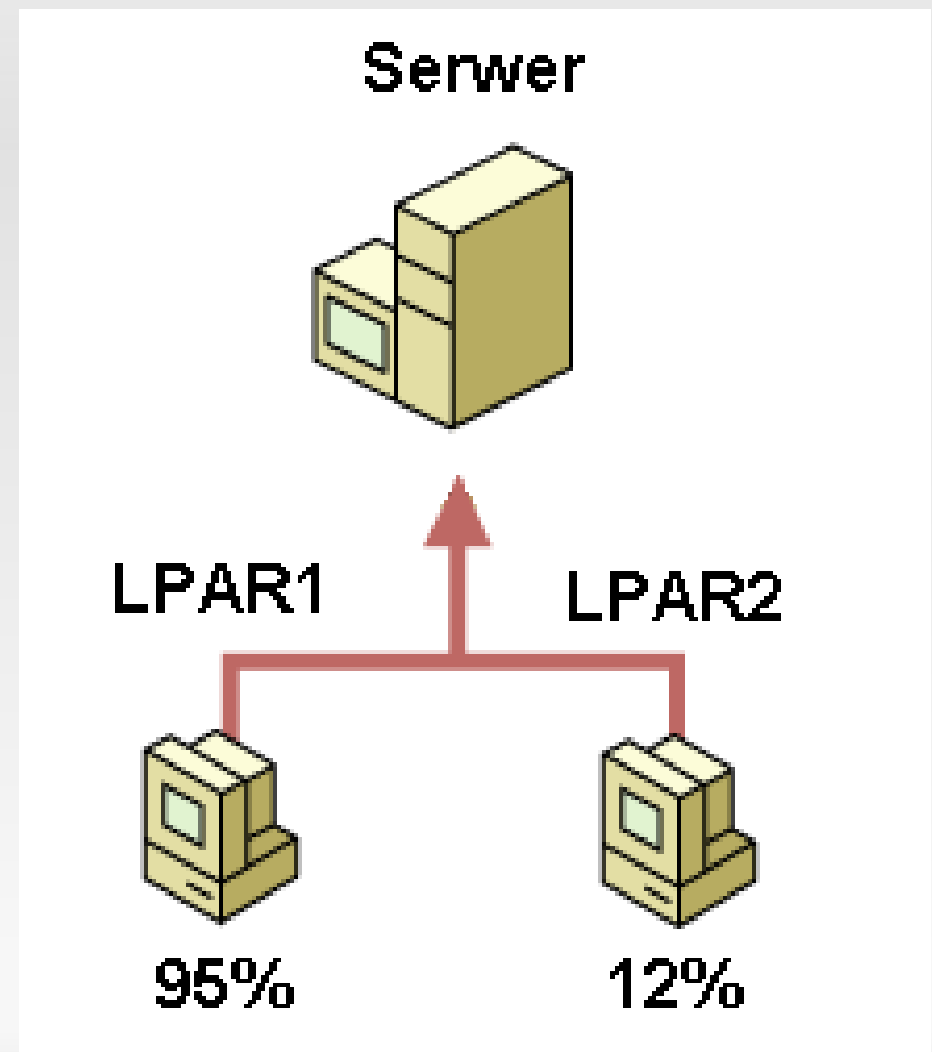
Do czego to jest przydatne

- nierówne obciążenie
- możliwości LPAR
 - konieczność restartu
- możliwości DLPAR
 - bez restartu
- mikropartycjonowanie
 - automatyczna



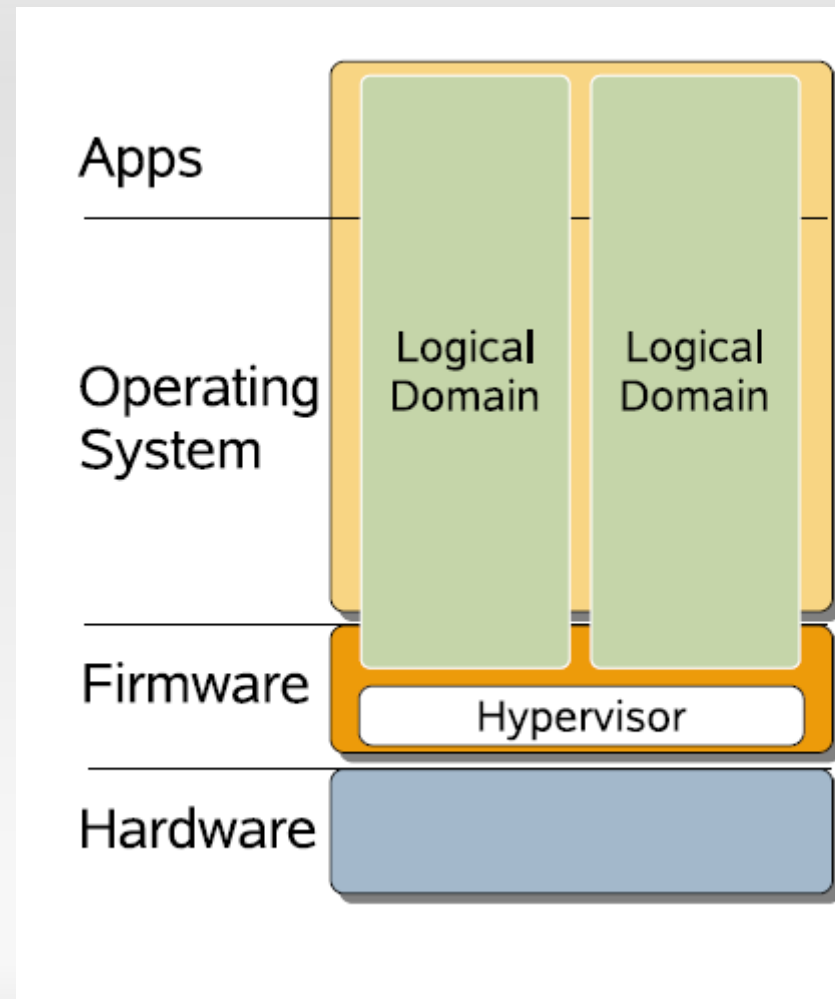
Do czego to jest przydatne

- nagła zmiana obciążenia
- możliwości LPAR
 - konieczność restartu
- możliwości DLPAR
 - ingerencja administratora
- mikropartycjonowanie
 - automatycznie



Sun Logical Domains

- Hypervisor na poziomie firmware'u - znajduje się na pamięci flash na płycie głównej
- Wymagana architektura - sun4v
- obecnie procesory UltraSPARC T1 oraz UltraSPARC T2



Sun Logical Domains

Domena logiczna (ang. Logical Domain) – oddzielna maszyna wirtualna, odizolowana od pozostałych, korzystająca z zasobów przydzielonych przez nadzorcę (hypervisor).

Dynamiczny przydział zasobów dla domeny – w trakcie działania systemu gościa możliwa jedynie zmiana przydziału zasobów CPU. Musi to obsługiwać system operacyjny.

Logical Domains - architektura

- Control Domain – kontroluje tworzenie i usuwanie pozostałych domen oraz przydział zasobów.
- I/O Domain – posiada bezpośredni dostęp do fizycznych urządzeń we/wy – np. kart sieciowych
- Service Domain – dostarcza (wirtualne) zasoby – dyski, sieć, konsole - najczęściej korzystając z I/O Domain
- Guest Domain – ma dostęp jedynie do wirtualnych zasobów (dysków, sieci, CPU) – na niej uruchamiamy np. Linuksa z serwerem www.

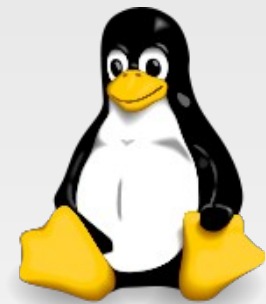
Logical Domains – systemy operacyjne

System operacyjny musi obsługiwać polecenia sun4v oraz wirtualne urządzenia tworzone przez Logical Domains.



Wsparcie dla sun4v posiadają:

- Solaris 10
- Linux – Ubuntu i Gentoo
- FreeBSD
- OpenVZ



Logical Domains – platforma sprzętowa

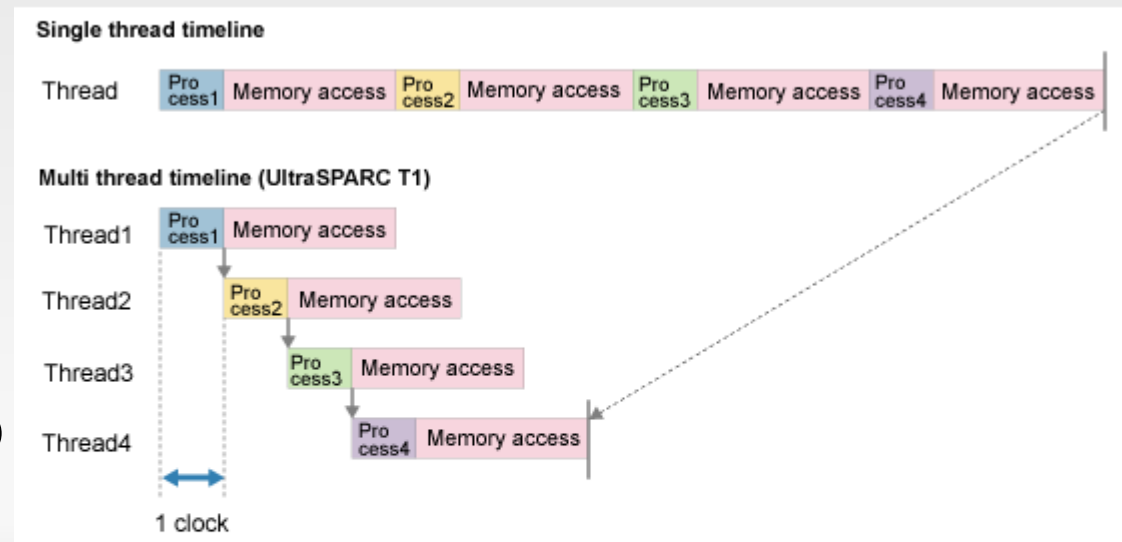
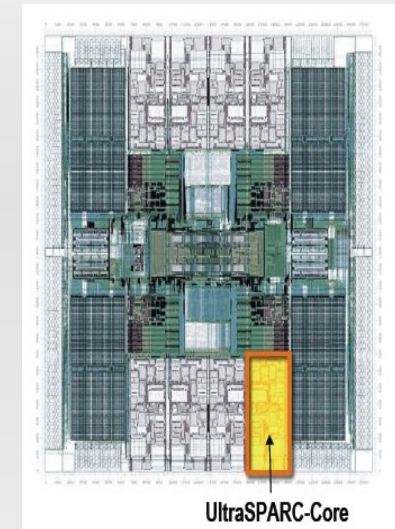
Architektura sun4v – tryb pracy hyper-privileged – korzysta z niego Hypervisor

- Procesory UltraSPARC T1 i T2

- Tylko jeden, 8-rdzeniowy procesor

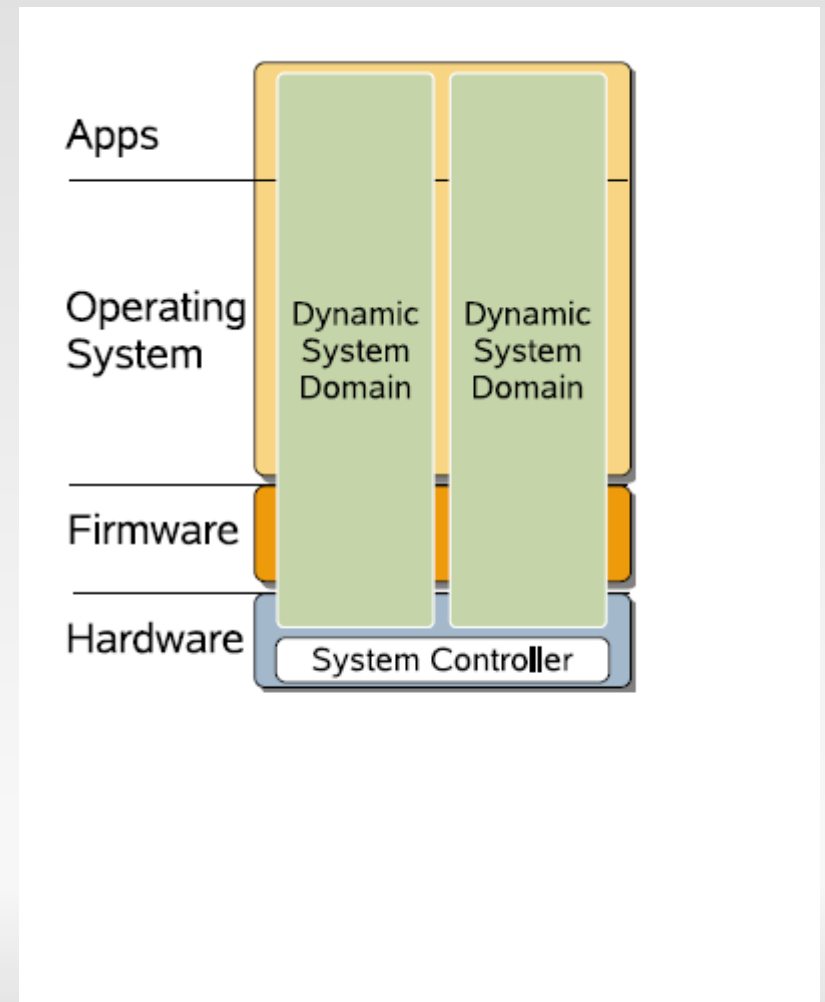
- Każdy rdzeń ma po 4 (T1) lub 8 (T2) wątków

- Może więc być max. 32 lub 64 domen logicznych



SPARC Enterprise Server

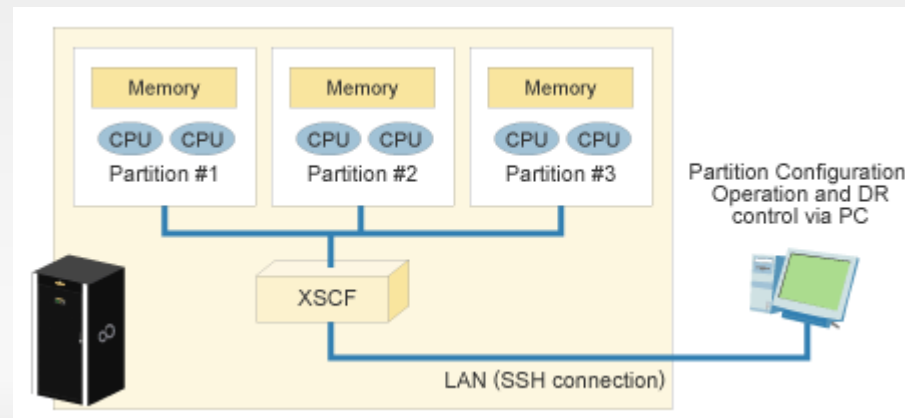
- Opracowane przez firmy Sun i Fujitsu
- Możliwe na serwerach SPARC Enterprise M4000, M5000, M8000, M9000
- Hardware Partitioning – systemy uruchamiane na oddzielnych (fizycznie) płytach zawierających niezależne CPU i pamięć



SPARC Enterprise Server

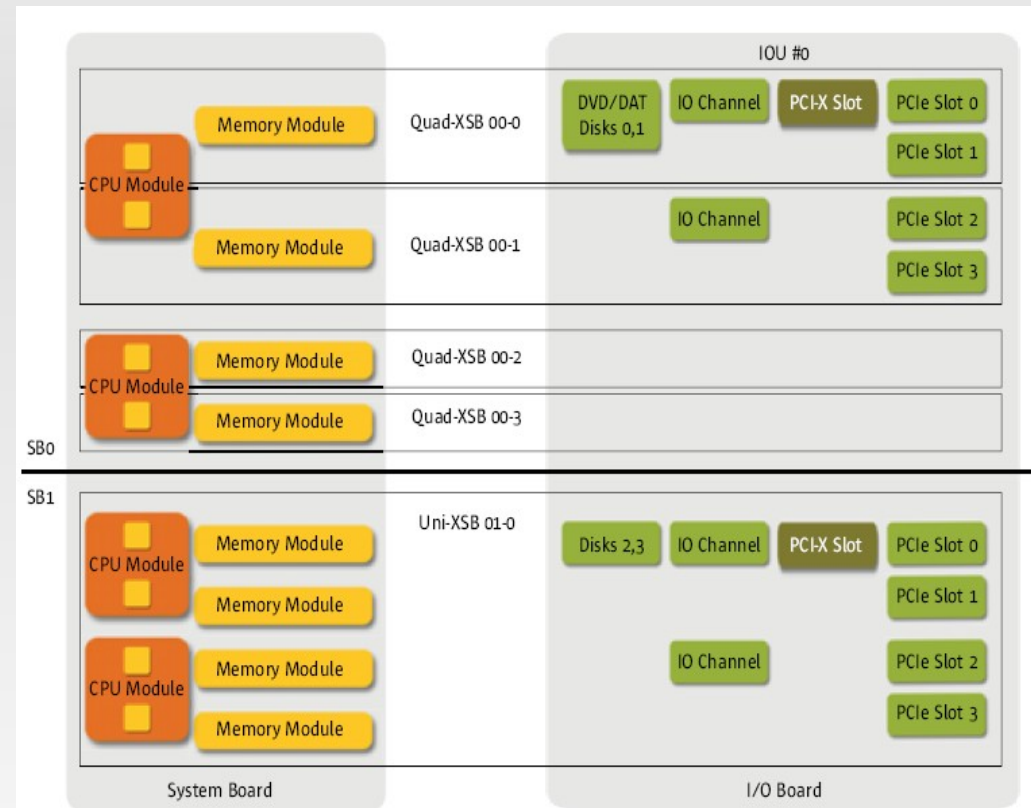
eXtended System Controller Facility

- eXtended System Controller Facility (XSCF) – firmware odpowiadający za konfigurację całego serwera
- umożliwia dynamiczną rekonfigurację domen, monitoruje błędy w systemie, zapewnia bezpieczeństwo..
- komunikacja między poszczególnymi domenami zabroniona



SPARC Enterprise Server

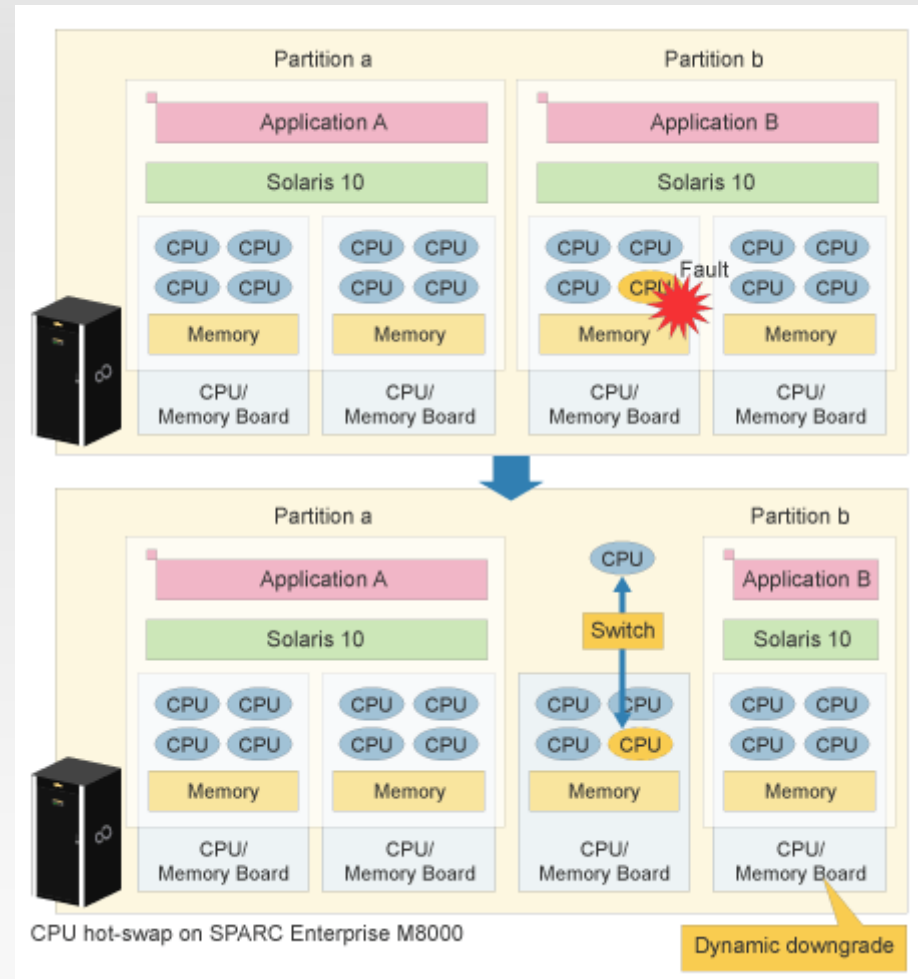
- Physical System Board (PSB)- płyta zawierająca max. 4 CPU
- eXtended System Board (XSB) = PSB lub „1/4” PSB (1 CPU)
- System używa jednej lub więcej XSB
- Możliwość dynamicznej zmiany zasobów (przydzielanie i usuwanie XSB) bez restartu systemu gościa



SPARC Enterprise Server

odporność na awarie

- Jeden z procesorów jest wadliwy
- Z Partycji b usuwamy PSB z wadliwym CPU
- **Nie przerywając** działania systemu B wymieniamy wadliwy CPU
- Duży plus - awaria jednego elementu nie zakłóca działania całego systemu



Kto w ogóle korzysta z (D)LPAR (wg IBM)?

- Royal Bank of Canada
- Audi
- DaimlerChrysler
- Whirlpool (pralka Whirlpool może pełnić rolę serwera systemu operacyjnego)
- Mayo Clinic

Korzyści z wprowadzenia LPAR (wg IBM)

- obniżenie kosztów energetycznych o 65% (dane dla pServer)
- obniżenie kosztów infrastruktury o 72% (dane dla pServer)
- duża elastyczność zmiany konfiguracji sprzętu i rozwoju technologii
- scentralizowane zarządzanie systemami
- zwiększenie niezawodności systemów
- zwiększenie bezpieczeństwa systemów

Praktyczne wyniki testów

Zwycięzcy testów w swoich klasach:

TPC	Dużo złożonych losowych zapytań do ogromnej bazy danych	IBM eServer p5 595 -> 100,512 zapytań na godzinę (przy 3TB) IBM pServer 575 -> 180,108 zapytań na godzinę (przy 10TB)
SPECweb99	Sprawdza, ile jednoczesnych połączeń serwer jest w stanie obsłużyć	IBM eServer p5 570 -> 25,000 połączeń
LINPACK	Rozwiązywanie dużych układów równań ze zmiennymi zmiennoprzecinkowymi	IBM pServer p5 570 -> 235,1 Gigaflopów na sekundę IBM pServer p5 595 -> 217,1 Gigaflopów na sekundę IBM eServer p5 595 -> 418 Gigaflopów na sekundę
SPECjbb2005	Symuluje zapytania do serwisu udającego sklep internetowy	IBM pServer p5 570 -> 691,975 operacji biznesowych na sekundę

We'd take a victory lap, but these things are heavy.



Literatura

- <http://www.sun.com/servers/coolthreads/ldoms/index.xml> - Sun Logical Domains
- <http://www.sun.com/blueprints/0207/820-0832.html> - Beginners Guide to LDoms
- <http://www.opensparc.net/operating-systems-supporting.html>
- <http://www.fujitsu.com/global/services/computing/server/sparcenterprise/tech>
- <http://www-941.ibm.com/collaboration/wiki/display/virtualization/Home>
- <http://en.wikipedia.org/wiki/DLPAR>