

LPAR - logiczne partycjonowanie systemów

Anna Kulka, Edyta Luty, Michał Ziemba

MIMUW

19 grudnia 2007

O czym będzie mowa...

1 Ogólnie

- Co to tak naprawdę jest
- Do czego służy
- Gdzie jest przydatne
- Cele
- Technologie
- Podstawowe pojęcia

2 LPAR

- Wprowadzenie do LPAR
- Platforma sprzętowa
- Funkcjonalność i zastosowanie
- Rekonfiguracja
- Zmiany potrzebne w Linuksie na LPAR

O czym będzie mowa...

- 3 LDOM
 - Platforma sprzętowa
 - Zarządca partycji
 - Dynamiczna rekonfiguracja
 - Zmiany w systemach operacyjnych
- 4 PPAR, XPAR
 - PPAR, XPAR
- 5 Podsumowanie
 - Zalety
 - Przykłady użycia
 - Pytania
 - Zakończenie

LPAR - co to jest

- logiczne partycjonowanie systemów
- dwuznaczność nazwy - mechanizm i sposób jego realizacji przez IBM
- mechanizm pozwalający obsłużyć wymagania wielu systemów za pomocą jednej maszyny fizycznej
- koncepcja wirualizacji dzięki ktorej nabywcy serwerów średniej klasy mają do dyspozycji tak potężne technologie, zwiększające możliwości i wydajność.

Do czego służy

Do czego służy

- pozwala na uruchamianie wielu pojedynczych systemów w ramach jednej centralnie zarządzanej maszyny
- oprogramowanie zarządzające może automatycznie przesuwać zasoby między maszynami

Gdzie można tego używać

Gdzie można tego używać

- aplikacje które przewidują duże obciążenie sezonowo np. sklepy przed świątami

Cele

Cele

- konsolidacja serwerów
- tworzenie środowisk produkcyjnych i testowych
- podział danych i operacji
- równoważenie skalowania
- elastyczność konfiguracji

Technologie

Technologie

- LPAR - IBM
- LDOM - SUN
- PPAR, XPAR - Fujitsu
- nPar, vPar - HP

Podstawowe pojęcia

Hypervisor

Zarządca partycji, pośrednik między warstwą sprzętową i systemu operacyjnego:

- Dzieli sprzęt na logiczne partycje (LPARy)
- Zapewnia izolację między nimi
- Dzieli pracę procesorów fizycznych między potrzeby poszczególnych partycji
- Zapewnia komunikację między partycjami (Virtual SCSI i Virtual Ethernet)

Podstawowe pojęcia

Micropartitioning

Rozszerzenie idei LPAR, polegające na współdzieleniu procesorów w jeszcze większym stopniu.

LPAR w sensie pojedynczej partycji

Wycinek serwera stanowiący niezależną maszynę

Virtual I/O Server

Oprogramowanie umieszczone na jednym z dedykowanych LPARów umożliwiające połączenia Virtual SCSI i Virtual Ethernet

LPAR i IBM

LPAR

- technologia zaprojektowana dla mainframe'ów ok '90r przez IBM
- bazuje na parawirtualizacji z wykorzystaniem hypervisora (zarządcy sprzętowego)
- izolacja systemów na poszczególnych partycjach

DLPAR

- mechanizm dynamicznej rekonfiguracji (przydziału procesorów, RAMu itd)
- NIE (sic!) wymaga zamykania systemu działającego na LPARze

LPAR – Co nie jest technologią LPAR?...

...ale też jest obsługiwane przez eSeries i5

- Virtual I/O
- Virtual Ethernet
- Capacity on Demand

Platformy obsługujące LPAR:

średniej wielkości maszyny

- iSeries
- eSeries i5
- pSeries
- Blade Center
- OpenPower

mainframe'y

- System z9
- zSeries

Elementy wsparcia:

Procesor

Power4, Power5

Konsola zewnętrzna

HCM

Współdzielenie gniazd

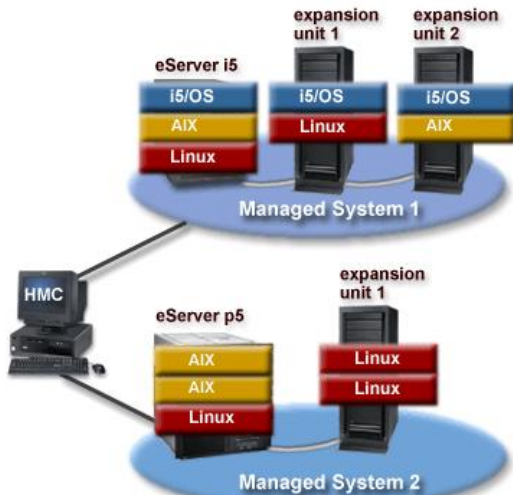
PCI, Ethernet

Funkcjonalność

Możliwość lepszego wykorzystania

- procesorów (mocy obliczeniowej-CuOD)
- pamięci
- gniazd PCI
- gniazd Ethernetowych

Hard Management Console



Integrated Virtualization Manager

- Alternatywa dla HMC do stosowania w mniejszych serwerach pSeries.
- Nie wymaga dedykowanej konsoli, instalowany wraz z oprogramowaniem Virtual I/O Serwera.
- Posiada interface WWW (https) + wiersz poleceń.
- Możliwości tworzenie wielu profili na LPAR.
- Pełnego wsparcia dla CoD.
- Możliwości zarządzania wieloma serwerami z jednego punktu.
- Możliwości przydziału fizycznych urządzeń do LPAR-a (wszystkie LPAR-y w pełni zwirtualizowane).

Nie można stosować jednocześnie konsoli HMC i IVM

CuOD-Dynamic Capacity on Demand

- pozwala zamawiać i instalować systemy z nadmiarowymi procesorami
- procesory trzymane w rezerwie do czasu nadmiernego obciążenia
- Administrator włącza dodatkowe zasoby bez potrzeby restartowania
- Dynamic CPU Guard - w razie awarii dynamiczne odłączenie i zastąpienie

Obsługiwane systemy operacyjne

Obsługiwane systemy

- z//OS
- z//VM
- z//VSE
- z//TPF
- AIX
- Linuksy (szczególnie Red Hat i SUSE)
- Linux on Power
- i5/OS

Wymagania dla Linuksów

Konieczne...

- jądro od 2.6.x
- Platform Enablement Librar librtas-1.3.1-0.ppc64.rpm
- Hardware Inventory lsvpd-0.15.1-1.ppc.rpm

... i dostateczne, aby wykorzystywać możliwości

- SRC – zarządzanie demonami
- RSCT – Reliable, Scalable, Cluster Technology
- CSM – autoryzacja z HMC
- ServiceRM – Service Resource Manager oraz....

Wymagania dla Linuksów

... i dostateczne, aby wykorzystywać możliwości

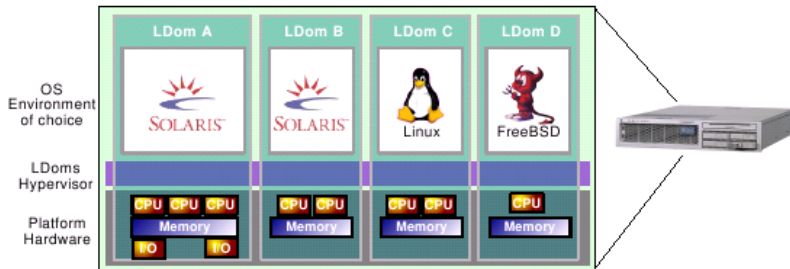
- DynamicRM –
 - Dynamicznie dodaje i usuwa procesory i gniazda we/wy z uruchomionych partycji
 - Optymalizuje wykorzystanie zasobów
 - Generuje wstrzymywanie, bądź zabijanie procesów na l-partycji
- Service log
- Error Log Analysis
- PCI Hotplug Tools
- Dynamic Reconfiguration Tools
- Inventory Scout

LDOM – Co to takiego?

LDOM

- technologia opracowana przez Sun Microsystems
- oferuje wirtualne środowisko obliczeń
- każda domena jest wirtualną maszyną z konfigurowalnymi zasobami sprzętowymi
- systemy na poszczególnych domenach pracują niezależnie

Technologia LDOM



Platforma sprzętowa

Platforma

- architektura sun4v
- procesory UltraSPARC T1 i UltraSPARC T2
- mechanizm Chip Multi Threading, wykorzystywany przez LDOM
- wypuszczone na rynek w kwietniu 2007

UltraSPARC T1

UltraSPARC T1

- jeden z pierwszych procesorów wielordzeniowych i wielowątkowych
- 4,6 lub 8 rdzeni
- do 32 wątków
- technologia CoolThreads
- niskie zużycie energii

UltraSPARC T2

UltraSPARC T2

- 8 rdzeni
- do 64 wątków
- przeznaczony głównie do zastosowań sieciowych
- pierwszy procesor, którego blueprints są ogólnodostępne
- w sprzedaży od października 2007 roku

Wsparcie sprzętowe

UltraSPARC T1

- do 32 domen logicznych
- wątki procesora przypisywane indywidualnie do LDOM
- dwie szyny PCI, które można przypisać do różnych LDOM

UltraSPARC T2

- do 64 domen logicznych
- Network Interface Unit (10 GbE) przypisywalny do dowolnej LDOM

Obsługiwane systemy operacyjne

Obsługiwane systemy

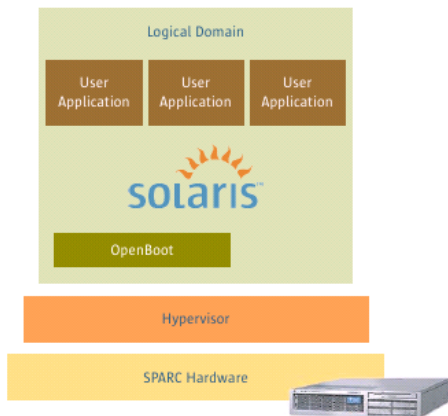
- Solaris 10
- OpenSolaris
- Ubuntu Linux Server Edition
- FreeBSD
- Wind River

Hypervisor

Funkcjonalność

- łączy środowisko operacyjne z platformą sprzętową, tworzy abstrakcję hardware'u (czyli to co w LPAR)
- używa rozszerzeń wbudowanych w procesor sun4v
- udostępnia Logical Domain Channels (LDCs), które umożliwiają komunikacje domen między sobą

Hypervisor



Hypervisor

LDOM Manager

- aplikacja umożliwiająca konfigurację domen
- tylko jedna domena może ją uruchamiać
- udostępnia interfejs Hypervisora przez linię komend

LDOM Manager

```
# ldm add-domain myldom1
# ldm add-vcpu 12 myldom1
# ldm add-memory 4G myldom1
# ldm add-vnet vnet1 primary-vsw0 myldom1
# ldm add-vdiskserverdevice /dev/dsk/c0t1d0s2 voll@primary-vds0
# ldm add-vdisk vdisk1 voll@primary-vds0 myldom1
# ldm set-variable auto-boot\?=false myldom1
# ldm set-variable boot-device=/virtual-devices@100/channel-devices@200/\
disk@0 myldom1
# ldm bind-domain myldom1
# ldm start-domain myldom1
# telnet localhost 5000
```


Hypervisor

Role Domen Logicznych

- Control Domain
- Service Domain
- I/O Domain
- Guest Domain

Dynamiczna rekonfiguracja

Dynamiczna rekonfiguracja

- wymaga co najmniej Solaris 10 Update 3
- możliwa dynamiczna rekonfiguracja CPU (bez restartowania systemu)
- rekonfiguracja pamięci wymaga restartu systemu
- tak samo rekonfiguracja urządzeń wejścia / wyjścia

Dynamiczna rekonfiguracja

```
primary# ldm set-vcpu 16 myldom1
```

```
myldom1# psrinfo -vp
```

```
The physical processor has 16 virtual processors (0-15)  
UltraSPARC-T1 (cpuid 0 clock 1000 MHz)
```

```
primary# ldm set-memory 2048m myldom1
```

```
Initiating delayed reconfigure operation on LDom myldom1. All configuration  
changes for other LDom's are disabled until the LDom reboots, at which time  
the new configuration for LDom myldom1 will also take effect.
```

Zmiany w systemach

Niezbędne patche – Solaris OS

- 118833-36 – nakładka na jądro
- 124921-02 – update'y dla sterowników i narzędzi do LDOM
- 125043-01 – update'y dla konsoli (qcn)

Ogólnie

PPAR, XPAR

Wirtualizacja stworzona przez Fujitsu-Siemens dla serwerd'ż'w PrimePower z serii 900,1500,2500

- obecnie dostępnA SUN w ramach APL na SPARC Enterprise M4000, M5000, M8000 i M9000.

Dwa rodzaje partycjonowania systemowego:

- PPAR - analogicznie do domen systemowych
- XPAR - rozszerzone partycjonowanie

Wymagania

PPAR

- wysokie ziarno podziału zasobów
- jedna płyta CPU/MEM (do 8CPU)

XPAR

- efektywniejsze ziarno podziału zasobów
- wymagane wykorzystanie specjalnie logicznie podzielonych płyt procesorowych

Zalety

Dlaczego właśnie tak?

- Podniesienie wydajności
- Redukcja kosztów
- Zwiększenie stopnia wykorzystania zasobów

BZWBK

147 LPAR'ów

- 7 z systemem i5 OS
- 136 z AIX 5.3
- 4 z SUSE Linux Enterprise Server

potrzebuje tylko:

- 79 procesorów POWER 5
- 320 GB RAMu
- 25 adapterów ethernetowych
- 20 adapterów FC

BZWBK

zapewniają także ...

- nie tylko wymaganą wydajność
- ale także pełną redundancję:)

SunFire T2000

Zadania

- 1 Uruchomienie zadania zabierającego 100% procesora,
- 2 Uruchomienie równoległe 16 zadań zabierających 100% procesora,
- 3 Pobranie 40MB pliku i zapisanie go na dysk
- 4 Pobranie równoległe 16 plików o wielkości 40MB i zapisanie na dysk.

SunFire T2000

Uzyskane czasy (domena kontrolna vs. logiczna

(kontrolna 4 wątki - 1 core, 1GB RAM vs. logiczna 16 wątków
- 4 core, 1 GB RAM)

- 1 2 minuty 30 sekund vs. 2 minuty 30 sekund
- 2 31 minut 13 sekund vs. 7 minut 48 sekund
- 3 4 sekundy vs 4 sekundy
- 4 1 minuta 1 sekunda vs 1 minuta 3 sekundy

Pytania

Serdecznie ...

... zachęcamy do zadawania pytan...

Zakończenie

Serdecznie ...

... dziękujemy za wysłuchanie prezentacji