

2005年 JANOG16

株式会社アイ・ティ・フロンティア

メインフレームサーバーの凄さについて (モデルはzSeriesです)

小田木 寿俊



IT Frontier

©Copyright IT Frontier Corporation 2004

初めにお断り

- IBM zSeriesをメインフレームの代表として取り上げております。当然富士通、日立のメインフレームも存在し、似た特徴を持っていると思っています。メインフレームへの継続的な投資、サポートを正式に表明しているのはIBMという現状とメインフレームの元祖という視点からzSeriesを取り上げています。
- メインフレームとは、従来のアーキテクチャの流れを持つ物を指しています。

メインフレームの特徴

とりあえず一般的な特徴は・・・

信頼性が桁違いにすごい！（説明あり）

拡張性が桁違いにすごい！

仮想化技術が桁違いにすごい！（説明あり）

ネットワークもそれなりに凄い！（説明あり）

柔軟性が桁違いにすごい！（説明あり）

お値段が桁違いにすごい！（と見えるが）

心ないコンサルタントによく虐められるところがすごい！

メインフレームの信頼性と安定性

zSeries



信頼性

全てのH/W要素を冗長化
二つの演算回路による結果の検査・再実行
MTBF 30年とも40年以上とも
その稼働率は99.999%

安定性

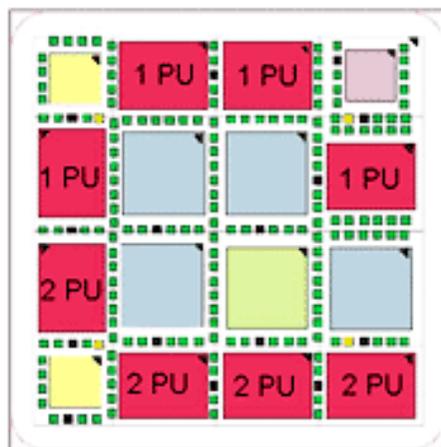
持てる能力を100%引き出す設計
CPUは本業に専念
IO処理はチャネルサブシステム、SAP
ハードウェアレベルのメモリー保護

メインフレームの信頼性と安定性

例えばCPUモジュール

MCM上に搭載する12CPUのうち同時に稼働するのは最大8個。残り4個のチップの内2つがSAPという入出力用に使われ、残りは予備となり、障害が発生した場合、予備のチップに自動的に切り替わる。システムの稼働には全く影響がでないような設計になっている。

z990の心臓部であるMCM(Multi Chip Module)



PU: Processing Unit

z990は最大4枚のMCM
を搭載可能

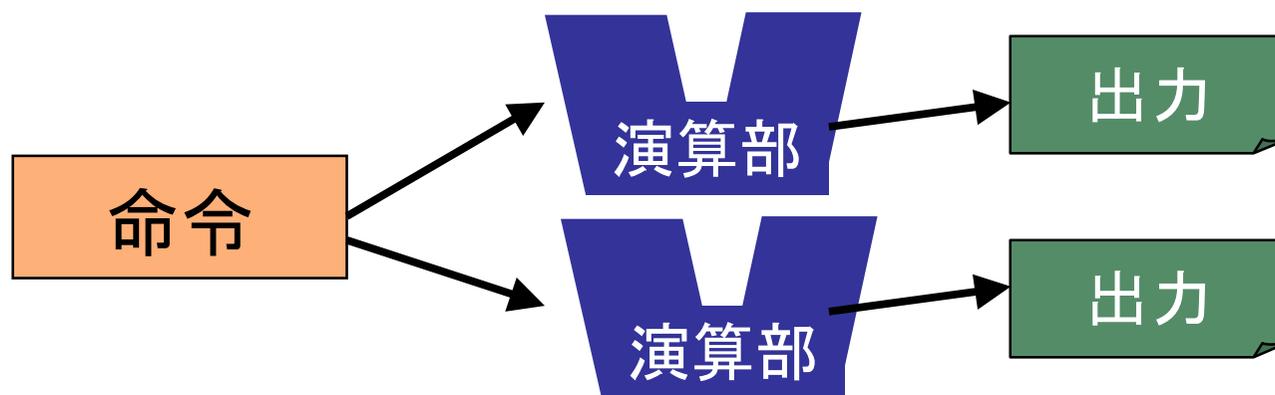
メインフレームの信頼性と安定性

例えばプロセッサエラー検査

複数の命令実行ユニットで同じ命令を実行
出力結果を検証

エラーがあれば再実行

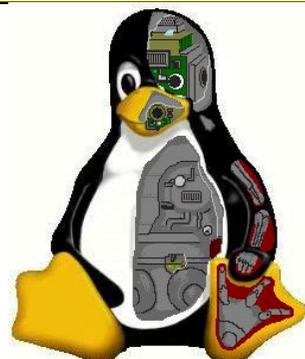
‘スーパースカラー’の技術を処理速度向上の為ではなく、信頼性の向上に利用



一致
しなければ
再実行

メインフレームの仮想化技術

zSeries



LPAR

筐体をロジカル分割
CPUを各LPARで共用可能
CPUパワーを細かく配分可能
DISKなどのIO装置も共用可能
動的に再配分も可能

VM

VMでソフトウェア的に分割
VM内では無数のゲストOSが稼動
一部物理的な制約を越えた指定が可能

メインフレームの仮想化技術

例えばLPAR

PR/SMと呼ばれる機能を用いる。

一つの筐体を複数に論理分割。

分割された論理区画(LPAR)が一つのサーバーとなる。

最新のzSeries990、zSeries890では30区画

LPAR間でCPU資源、殆どのIO資源を共用する事が可能

メインフレームの仮想化技術

例えばVM

ソフトウェア的に仮想マシンを構築する技術

前述のPR/SMはVMの機能の一部をH/W化したもの。

VMはH/Wテストの環境であり、OSの開発環境でもある。

VM上で稼動するOSをゲストOSという。

物理的な制約を越えて仮想的にゲストOSに資源を供給可能。

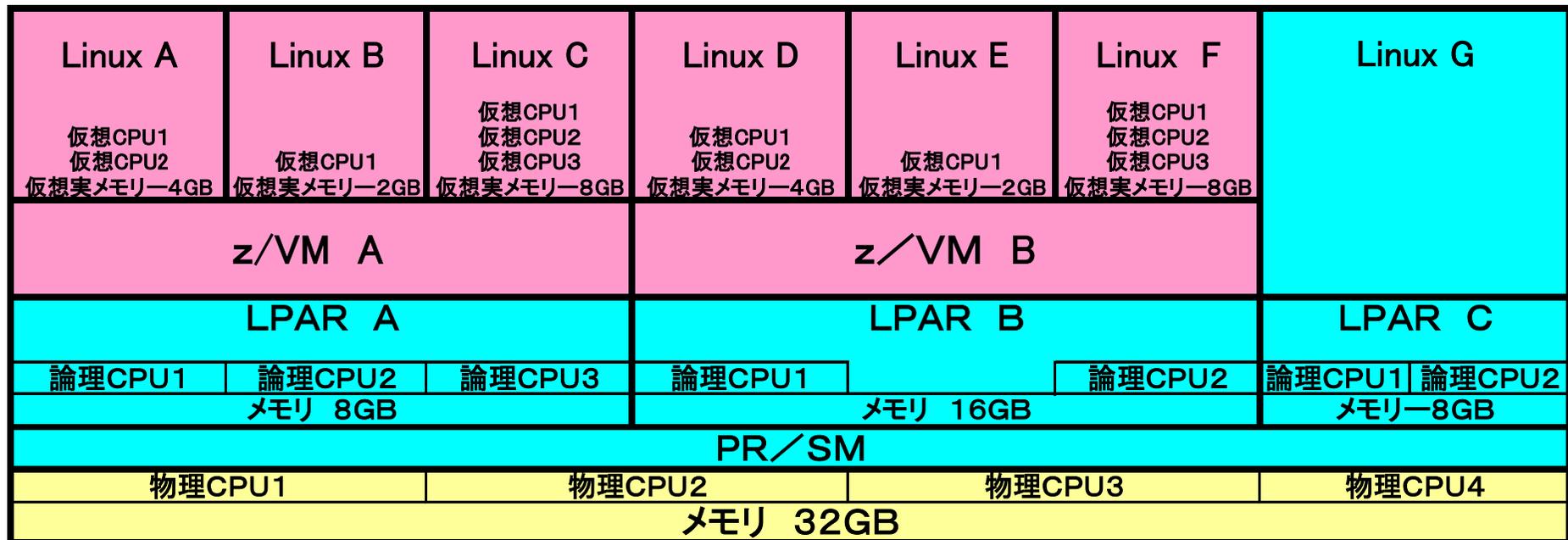
理論的には一つのVMの下で無数のゲストOSを稼動可能。

前述のLPARとVMを組み合わせた環境下で
Linuxサーバーを
統合したソリューションが今ブーム

メインフレームの仮想化技術

物理筐体とLPARとVMの関係

zSeriesで稼動するz/Linuxの概念図



zSeries Server

ゲストOSには、VMが割り当てた、仮想のCPU、仮想の実メモリが、実在するCPU、メモリに見える。

VMは、実際のCPU数、メモリ容量の壁を超えて、ゲストOSに仮想のCPU数、仮想実メモリを割り当てる可能。

各LPAR間で、設定によりCPUパワーの自動的な再配分(自由な貸し借り)が可能。

メインフレームのネットワーク技術

zSeries



OSA

外部LANへ接続する為のアダプター
複数のLPARやVMゲストOSで共用可能

HiperSockets

LPAR間を繋ぐ高速仮想LAN
OSのCPUコストを削減する大きなMTU

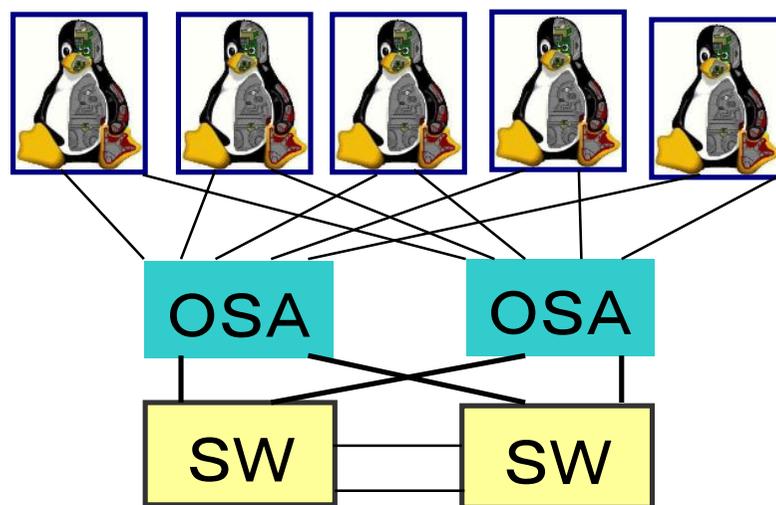
VM Guest LAN

VM内のゲストOS間を結ぶ仮想LAN
VM内でHiperSocketsをエミュレーション

メインフレームのネットワーク技術

例えばOSA

トークンリング・ギガビットイーサ、1000Base-Tへの接続が可能
OSAの物理ポートの共用により、スイッチなどの導入コストを削減
OSA二枚で複数の仮想Linuxサーバーの冗長化が可能
z990では最大物理ポート数48、ポート辺り2048のIPアドレス
MTU値 1492／8992



メインフレームのネットワーク技術

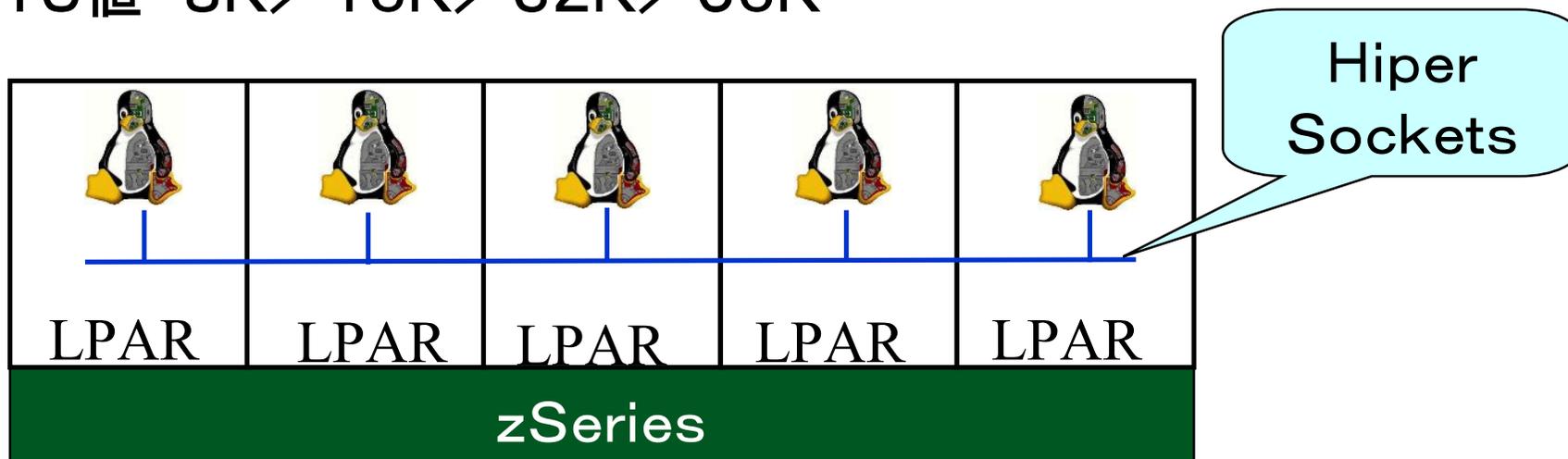
例えばHiperSockets

LPAR間の内部通信を実現する仮想ネットワーク
ネットワーク機器の導入・管理費用を削減

数Gbpsレベルのスループット？

外部からの改ざん、盗聴は不可能

MTU値 8K／16K／32K／56K



メインフレームのネットワーク技術

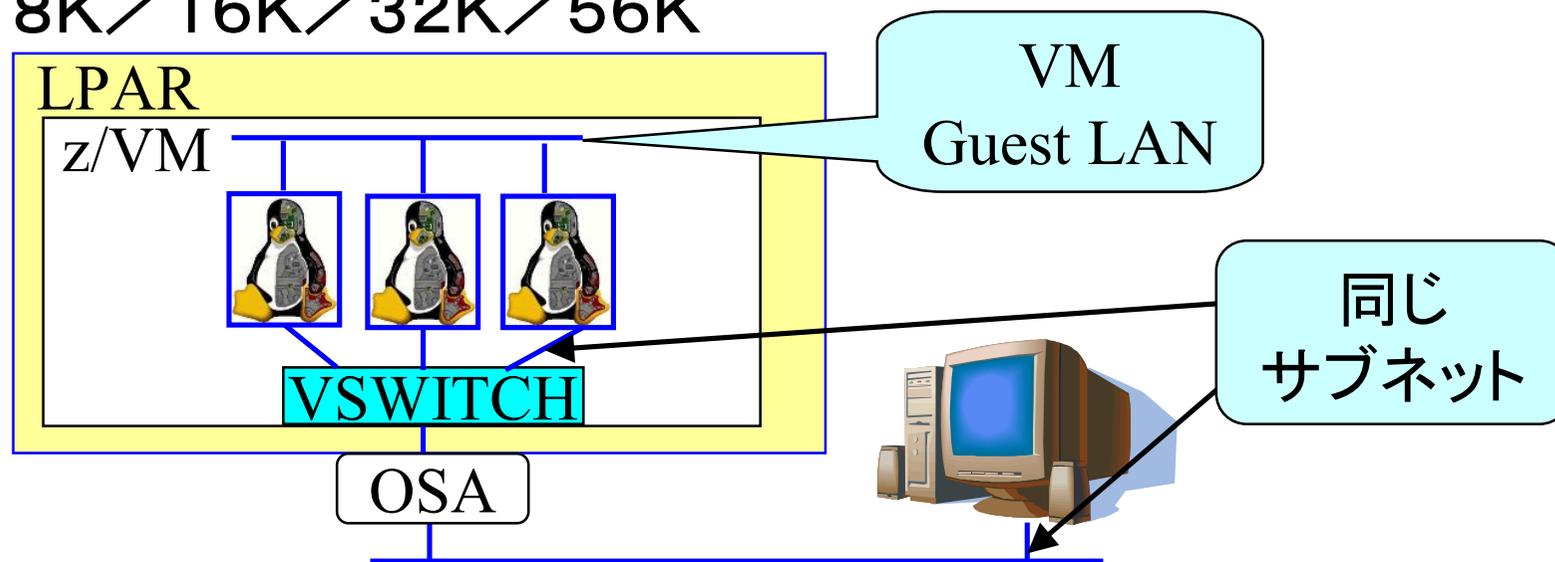
例えばVM Guest LAN

z/VMがHiperSocketsをエミュレーション

Guest OS間の内部通信を実現する仮想ネットワーク
その応用としてVSWITCHがある。

z/VMがレイヤー2スイッチやブリッジをエミュレーション

MTU値 8K/16K/32K/56K



メインフレームの柔軟性

zSeries



CUoD

稼働中の動的な型式変更

CoD

稼働中の一時的な処理能力の追加

リソース共有機能

仮想化してリソースを共有
システムリソースのダイナミックな再配分

メインフレームの柔軟性

例えばCUoD

同一物理モデル内で型式の変更が可能(プロセッサ追加起動)
メモリーについてもモデル内での稼働中の増設が可能
チャンネル・メモリーについても稼働中の追加が可能
(ゲージに余裕がある必要があります)

メインフレームの柔軟性

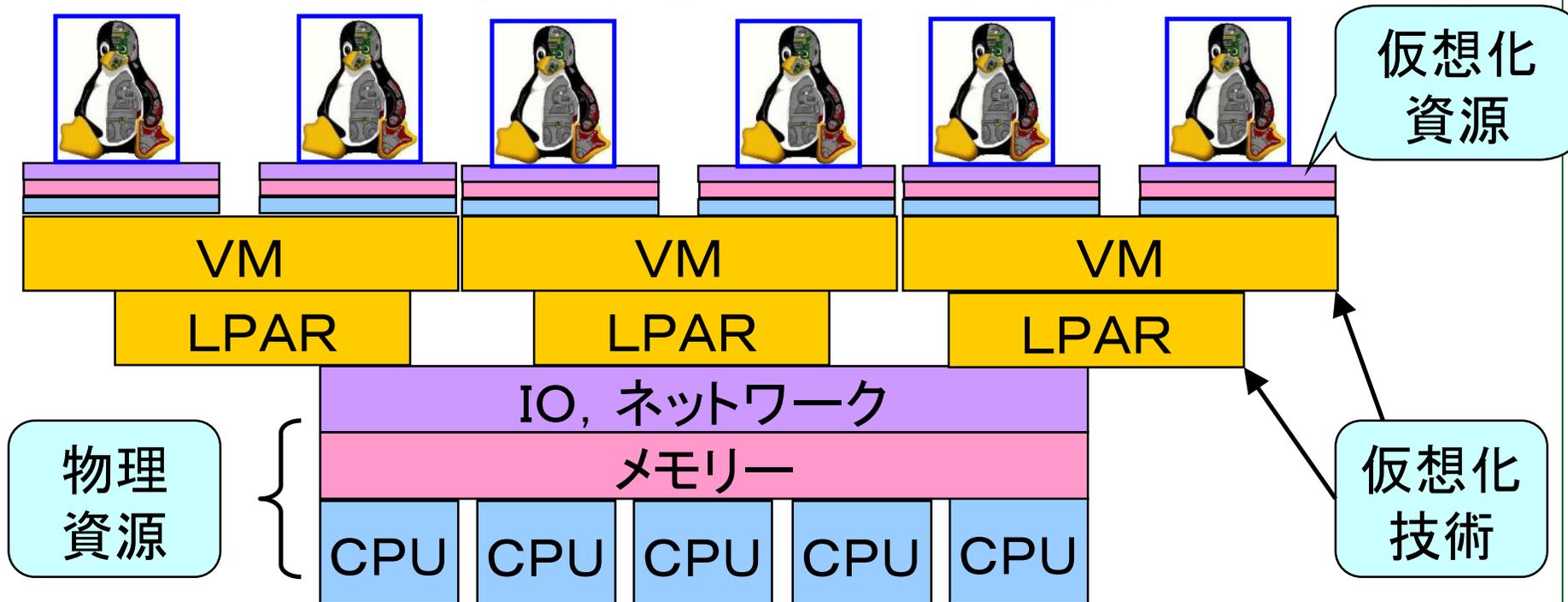
例えばCOD

システムを停止せずにCPUを動的に追加・削除
z/VMは再起動なしでCPUの追加・削除に対応可能
z/VM上のGuest OSであるLinuxもその恩恵を受ける事ができる
一定期間のみ一時的に筐体としてCPU能力の追加が可能

メインフレームの柔軟性

例えばリソース共用1

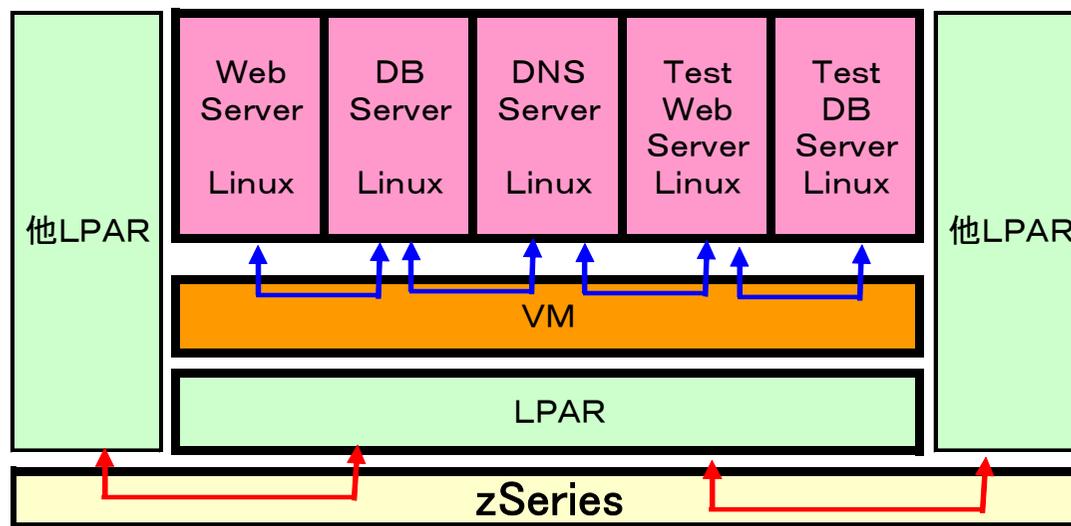
LPARやz/VM等のエミュレーション機能が背景
CPU、メモリー、IO装置、ネットワーク機能等の占有使用環境をエミュレーション
システムリソースを動的に仮想サーバー間で割振って無駄なく使用



メインフレームの柔軟性

例えばリソース共用2

LPAR内と外でシステムが自動的にワークロードの調整
急激なアクセス増に対しても、一時的に対象サーバーの処理能力
を高めることが可能



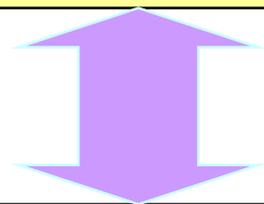
- VMがLPARに割当てられたCPU資源を動的に再配置
- LPAR間でのCPU資源の動的再配置

おまけ 故障かな? と思ったら

故障かな? と思ったら

メインフレームは自分でメーカーに電話をします

CEから連絡がきて保守日の相談を人間とします



この間も殆どの機構が冗長化されている為、
停止することはありません。
(そうありがたい)