

SPARC M12/M10 Oracle VM Server for SPARC 設計のポイント ～ゲストドメイン編～

2018年4月（第2.0版）

富士通株式会社



■ 目的

- 本書は、SPARC M12/M10で仮想化を実現するOracle VM Server for SPARCの豊富な機能を用いて、柔軟性の高いシステムを設計するためのポイントを説明しています。

■ 対象読者

- SPARC M12/M10とOracle VM Server for SPARCの導入を検討されている方
- Oracle Solaris, Oracle VM Server for SPARCの基礎知識を有している方

■ 留意事項

- 本書の内容は、Oracle VM Server for SPARC 3.5 / Oracle Solaris 11.3に基づいています。Oracle VM Server for SPARCのバージョンが3.4以前、またはOracle SolarisのバージョンがOracle Solaris 10の環境では、利用できない機能がある場合や操作方法が異なる場合があります。ご了承ください。
- サーバは、SPARC M12/M10（SPARC M12-1/M12-2/M12-2S/M10-1/M10-4/M10-4S）を対象としています。

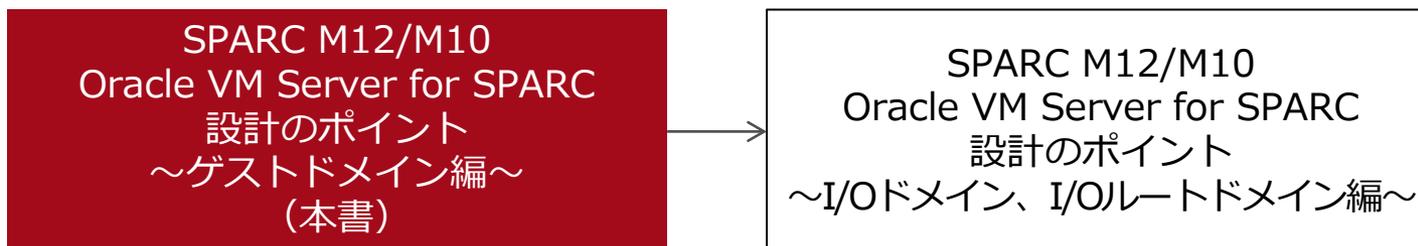
■ 本書での表記

- コマンドのセクション番号は省略しています。
例：
 - ls(1) ⇒ lsコマンド
 - shutdown(1M) ⇒ shutdownコマンド
- 以下の用語は略称を用いて表記する場合があります。

略称	正式名称
Solaris	Oracle Solaris
Oracle VM	Oracle VM Server for SPARC

■ ドキュメントの位置づけ

- Oracle VM Server for SPARCによる仮想化の柔軟性を生かしつつ、より高性能なシステムや可用性の高いシステムを設計する場合には、本書の内容を理解した上で『SPARC M12/M10 Oracle VM Server for SPARC 設計のポイント～I/Oドメイン、I/Oルートドメイン編～』を参照してください。



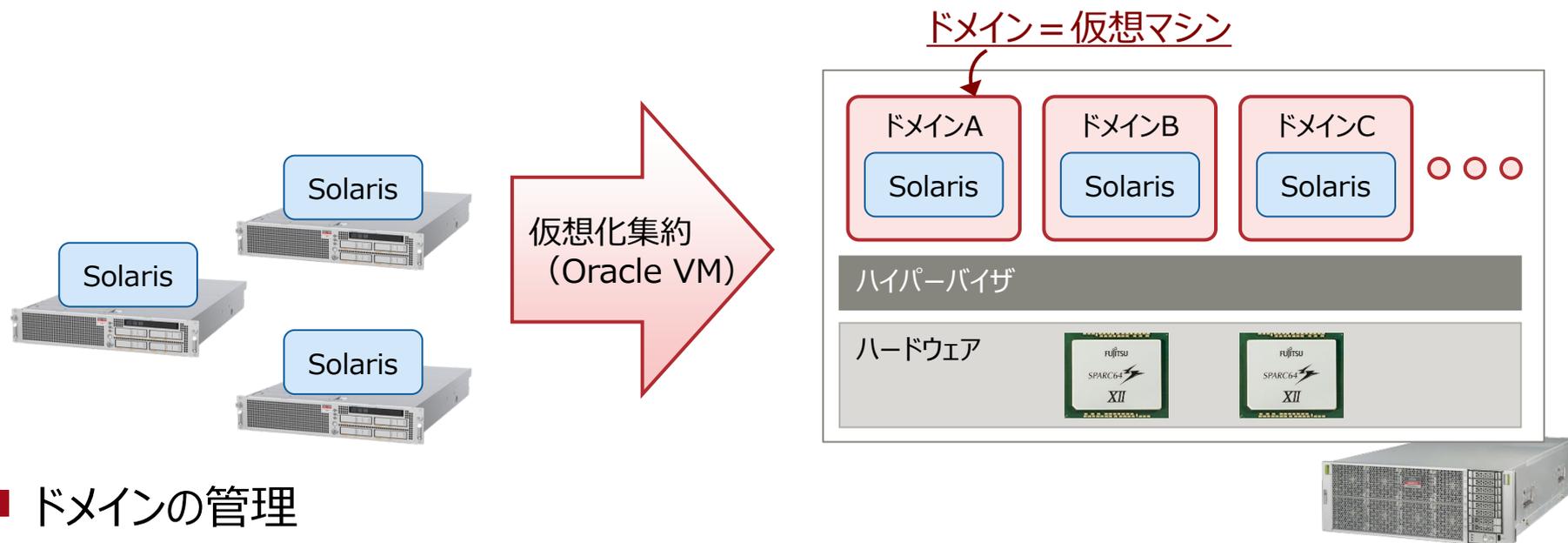
1. Oracle VM Server for SPARCの基礎知識
 2. ゲストドメインの概要
 3. CPU・メモリリソース設計のポイント
 4. I/Oデバイス設計のポイント
 5. 設計サンプル
 6. ライブマイグレーション
- 付録

1. Oracle VM Server for SPARCの 基礎知識

本章では、Oracle VMの概要や基本的な用語について解説します。

■ Oracle VM Server for SPARCの構造

- Oracle VMは、ハイパーバイザ層でサーバを論理的に分割して仮想化を実現します。サーバ内に「ドメイン」と呼ばれる仮想マシンを構築し、各ドメインで独立したSolaris（Solaris 10 or Solaris 11）環境を稼動できます。



■ ドメインの管理

- 各ドメインには、CPU、メモリ、I/Oデバイスを自由に割り当てることができます。また、ドメインは個々に起動、停止、再起動することができます。

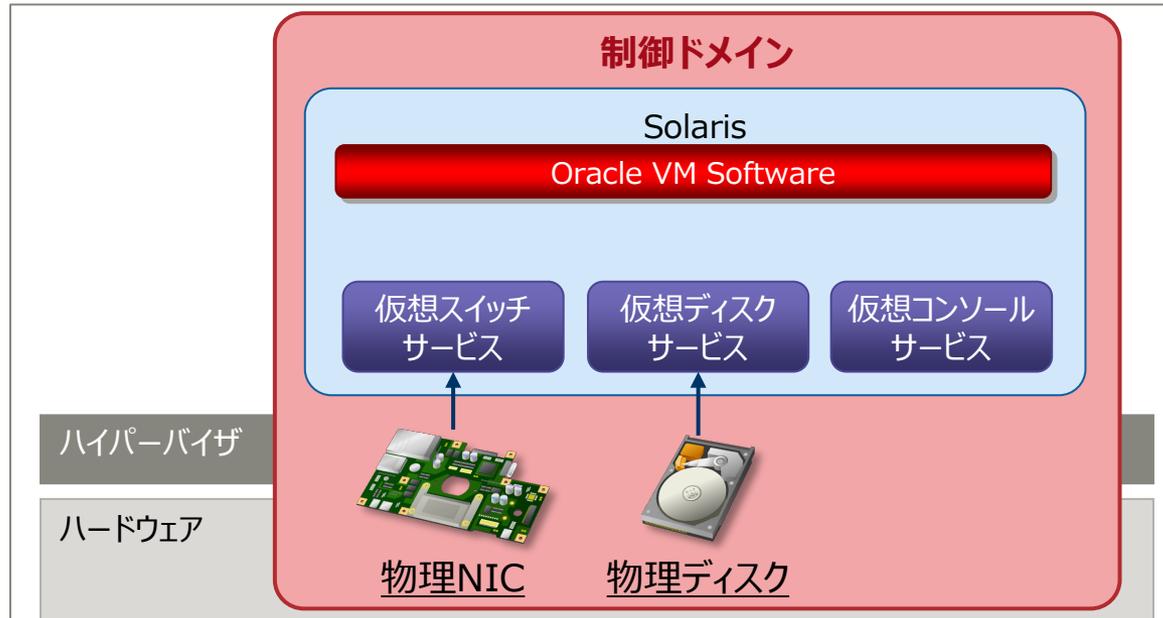
■ ドメインの操作性

- ドメイン（仮想環境）のOSの設定や操作方法は、従来（物理環境）とほとんど変わりません。既存の物理環境を仮想環境に移行しても、同等の運用および管理が可能です。

Oracle VMを構成し利用可能なドメインは、その役割や機能によって5つに分かれます。

■ 制御ドメイン

- 制御ドメインは、ほかのドメインを作成して管理し、仮想リソースをほかのドメインに割り当てることができるドメインです。制御ドメインは、サーバごとに1つだけ存在します。

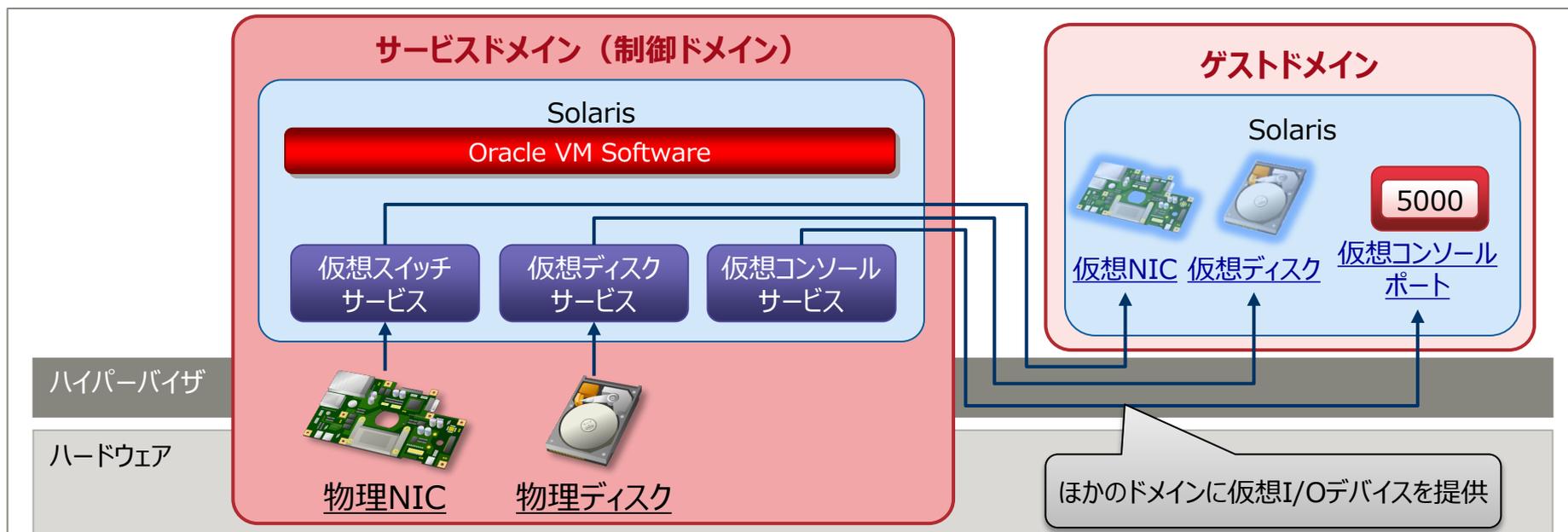


■ サービスドメイン

- サービスドメインは、仮想スイッチ、仮想ディスク、仮想コンソールなどの仮想サービスをほかのドメインに提供するドメインです。通常、制御ドメインまたはI/Oルートドメインがサービスドメインの役割を兼ねます。

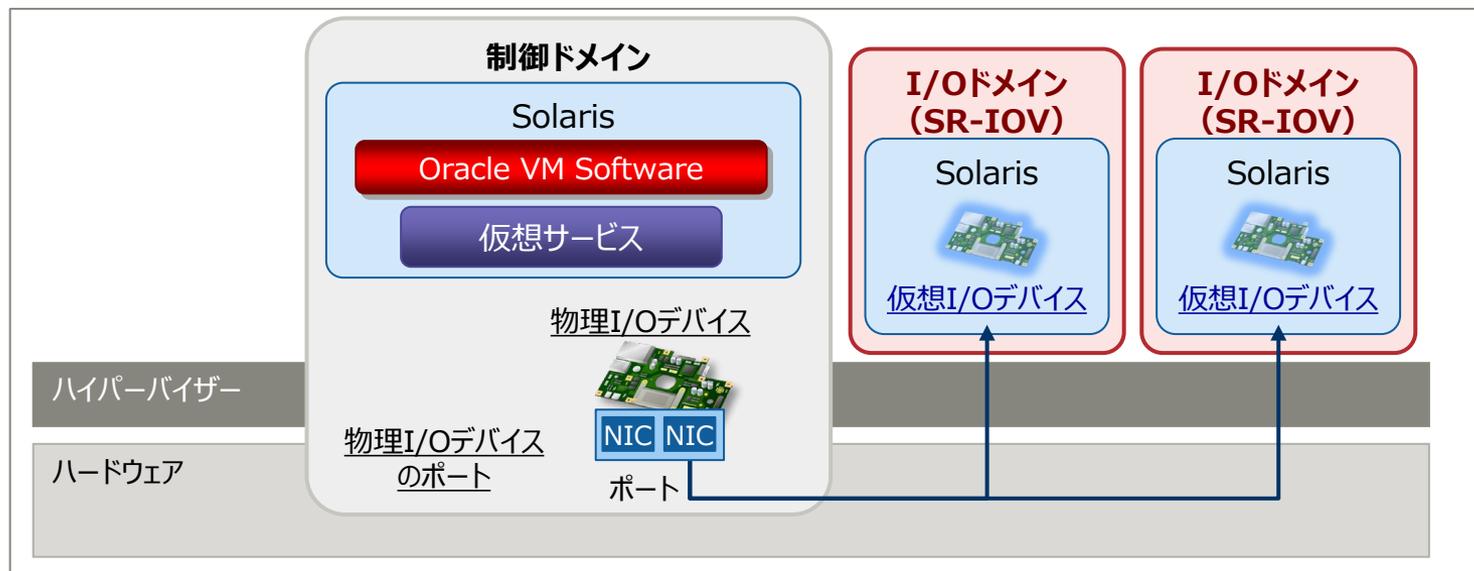
■ ゲストドメイン

- ゲストドメインは、物理I/Oデバイスの代わりに、制御ドメインやサービスドメインから提供される仮想デバイスや仮想サービスを使用するドメインです。通常、Oracle VMでは、ゲストドメインを多数構築することで、システムの集約性を高めます。



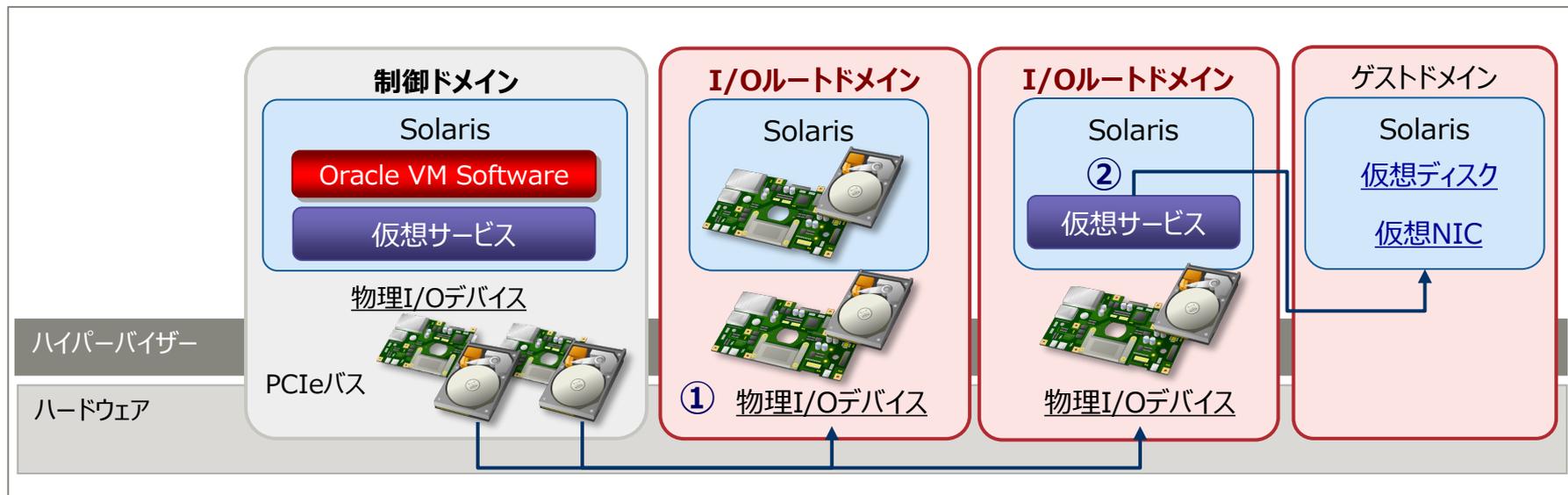
■ I/Oドメイン (SR-IOV)

- I/Oドメインは、Single Root I/O Virtualization (以降、SR-IOV) 機能によって、物理I/Oデバイスのポート (LANポートやFCポートなど) が仮想的なポートに分割されて割り当てられたドメインです。複数のドメインで物理I/Oデバイスを共有できます。
- I/Oドメインは物理I/Oデバイスに直接アクセスできるため、仮想I/Oに関連する性能のオーバーヘッドを避け、物理I/Oデバイスと同等のパフォーマンスを得ることができます。



■ I/Oルートドメイン

- I/Oルートドメインは、物理I/Oデバイス（PCIカードやディスクなど）が割り当てられたドメインです。物理I/Oデバイスを占有し、それらに直接アクセスすることができます。（①）
- I/Oルートドメインをサービスドメインとして使用することで、割り当てられている物理I/Oデバイスを仮想デバイスとしてほかのドメインと共有できます。（②）
- 1つまたは複数の物理I/Oデバイスに直接アクセスできるため、仮想I/Oに関連するパフォーマンスのオーバーヘッドを避け、物理I/Oデバイスと同等のパフォーマンスを得ることができます。



2. ゲストドメインの概要

本章では、Oracle VMのさまざまなドメインのうち、柔軟性の高い構成（ゲストドメイン）について解説します。

Oracle VMによる仮想環境は、大きく以下の3つに分かれます。

ゲストドメイン

構成図



制御ドメイン
(サービスドメイン)



構築概要

- 制御ドメインとゲストドメインで構成します。
- 制御ドメインはサービスドメインとして、ゲストドメインに**仮想I/Oデバイス**を提供します。

特長

- 物理I/Oデバイスを複数のドメインで共有可能な、柔軟性の高い構成です。

I/Oドメイン

構成図



制御ドメイン



構築概要

- 制御ドメインとI/Oドメインで構成します。
- 物理的なポートを**仮想的なポート**に分割して、I/Oドメインに割り当てます。

特長

- 柔軟性と高性能を実現できる、バランスのよい構成です。

I/Oルートドメイン

構成図



構築概要

- 制御ドメインとI/Oルートドメインで構成します。
- **PCIeバスごと**、物理I/OデバイスをI/Oルートドメインに割り当てます。

特長

- 物理I/Oデバイスを占有させるため、性能が高く、ハードウェアの障害からシステムを分離して可用性を高められる構成です。

本書の解説範囲

各構築パターンの比較

		ゲストドメイン	I/Oドメイン	I/Oルートドメイン
対象サーバ		SPARC M12/M10 全モデル	SPARC M12/M10 全モデル	SPARC M12-2/2S SPARC M10-4/4S
耐障害性 (障害隔離性)		△ 制御ドメイン障害時、 全業務システム停止	△ 制御ドメイン障害時、 全業務システム停止	◎ ドメイン障害時、ほかの ドメインに影響なし
性能		○ I/Oの性能オーバーヘッド が生じる場合がある (※1)	◎ CPU・メモリ・I/Oの オーバーヘッドなし	◎ CPU・メモリ・I/Oの オーバーヘッドなし
コスト		◎ I/Oデバイス (PCIカード類)を 共有可能	△ I/Oデバイスを 共有可能 ※NP-IV対応の FCスイッチが必要	△ ドメインごとに I/Oデバイスが必要
柔軟性	ドメインの数 (※2)	◎ ほぼ制限なし (2~256 ドメイン) ※CPU数に依存	◎ ほぼ制限なし (2~256 ドメイン) ※CPU数・物理I/O デバイス数・ポート数に依存	△ 多数の集約は不可 (2~6 ドメイン程度) ※PCIeバスの数に依存
	業務システムの ライブマイグレーション	◎ 可能	× 不可	× 不可

本書の解説範囲

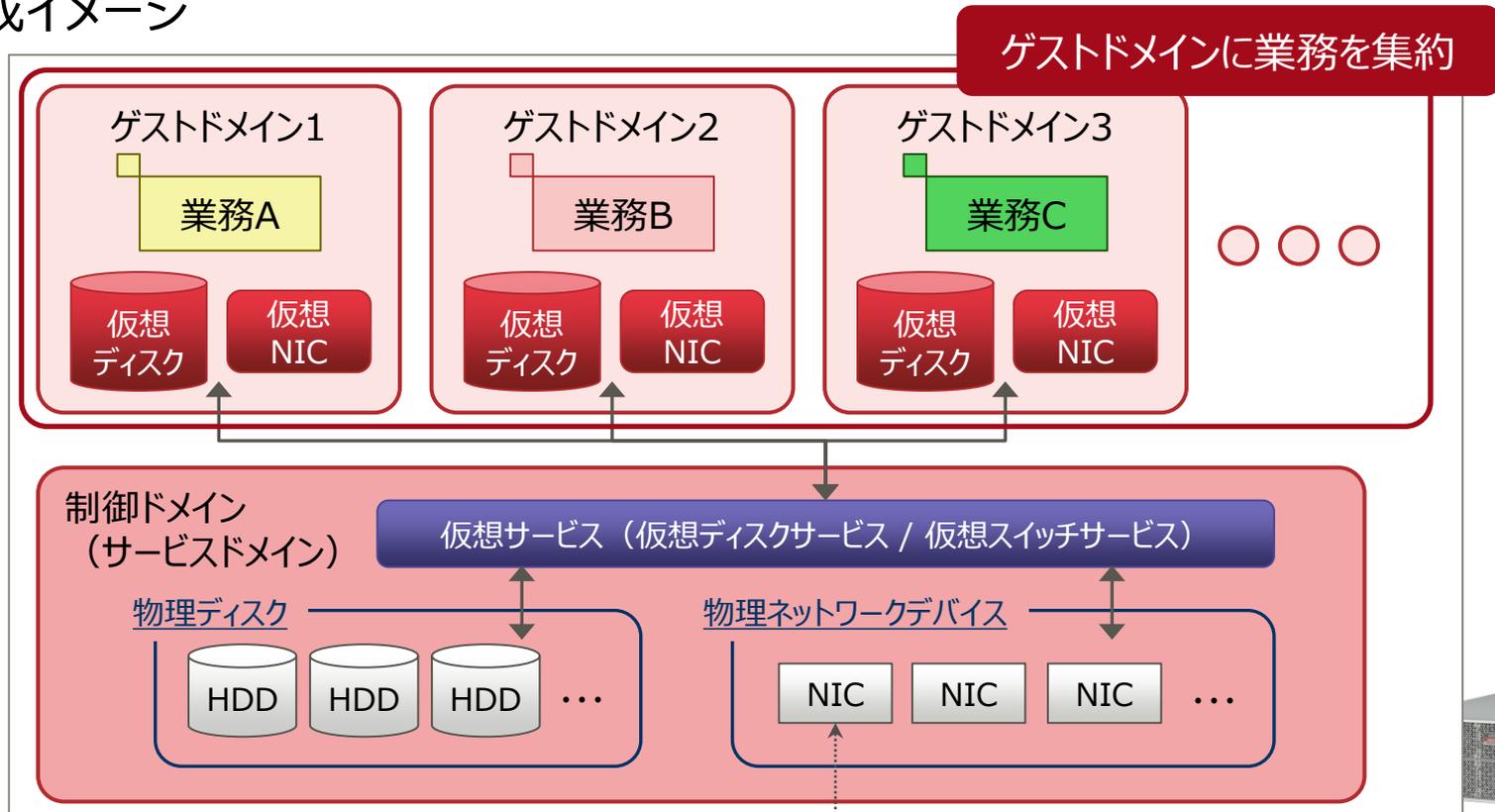
※1: CPU・メモリの性能オーバーヘッドはありません。

※2: 制御ドメインなどのサービスドメインを含めた、構築可能なドメインの数を記述しています。

■ 概要

- ゲストドメインで業務システムを稼働させる、Oracle VMの基本構成です。
- すべての物理I/Oデバイスを制御ドメインに割り当てます。
- 制御ドメインがサービスドメインとなり、すべての仮想I/Oデバイスを提供します。

■ 構成イメージ



本書では、物理LANポートを「NIC」と表記します。

■ 特長

- SPARC M12/M10のすべてのモデルで構成できます。
- 柔軟性が高く、業務システムの稼動中に以下の環境変更が可能です。
 - CPU、メモリ、仮想I/Oデバイスの追加、削除
 - ライブマイグレーション（他サーバへの業務システムの移動）

■ 留意事項

- 制御ドメイン（サービドメイン）に障害が発生した場合は、仮想I/Oデバイスが使用不可となり、すべての業務システムが停止します。

本事象の対策として、以下のような方法があります。

- I/Oルートドメイン上で業務システムを稼動
- ゲストドメインを含めたクラスタ環境を構築（PRIMECLUSTER, Oracle Solaris Clusterなど）



- 物理サーバと比較して、以下の理由によりゲストドメインのI/O（ネットワーク、ディスク）の性能が低下する場合があります。

- I/Oのアクセスがサービドメインを経由するため、オーバーヘッドが発生する。
- I/Oを複数のゲストドメインで共有するため、あるゲストドメインのI/Oの負荷がほかのゲストドメインのI/O性能に影響を与える。

3. CPU・メモリリソース 設計のポイント

各ドメインには、CPU・メモリのリソースを割り当てます。本章では、リソース割り当てに関する基本的なルールと推奨する設計指針を解説します。

■ 特長

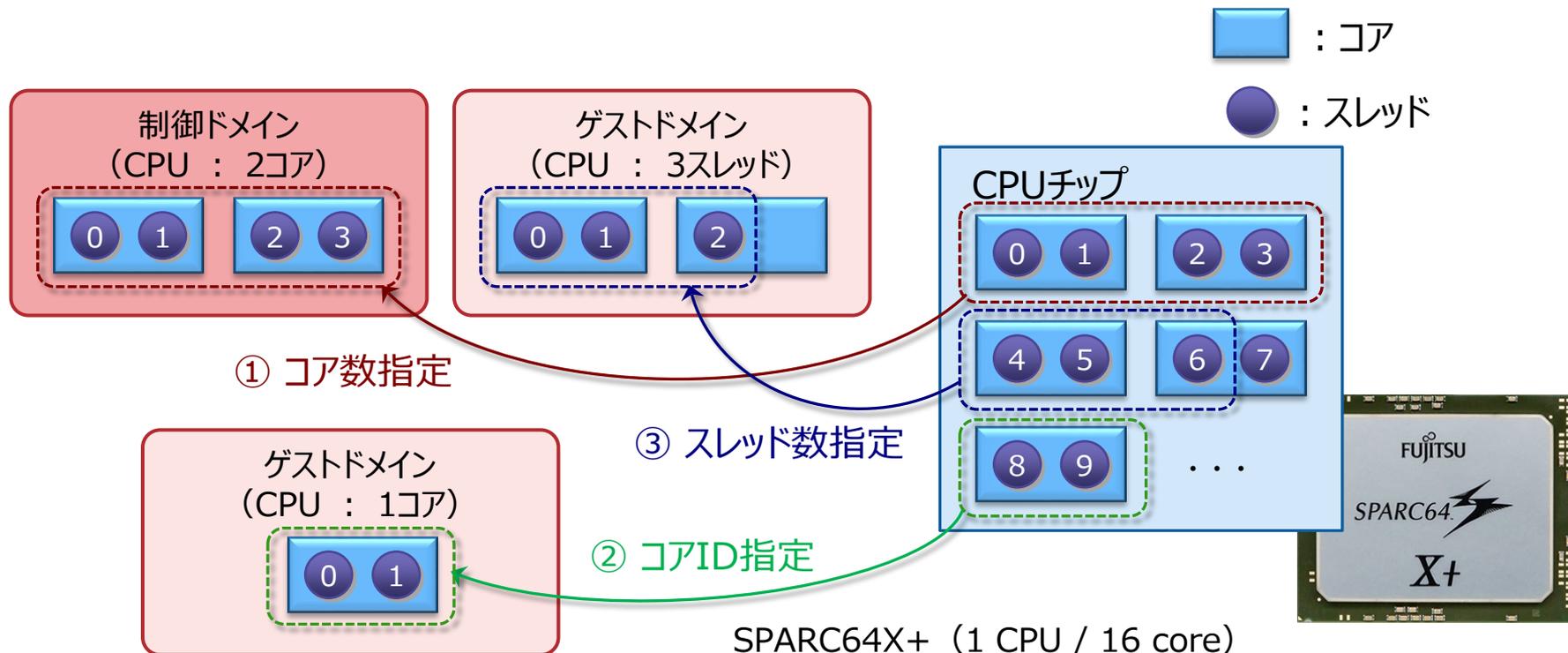
- 各ドメインに割り当てたCPU・メモリリソースは、そのドメインが占有します。
 - リソースプール化およびドメイン間でCPU・メモリリソースは共用しません（例：Oracle Solarisゾーン）。
- 割り当ての変更は、動的（ドメインが起動した状態）に実行可能です。
 - リソースの割り当て方式によっては、動的に変更できない場合があります。
（[「CPUの割り当て（2/2）」](#)、[「メモリの割り当て（1/2）」](#)参照）
- サーバ仮想化による性能のオーバーヘッドはありません。

■ 注意事項

- 業務システムが稼動するドメインに割り当てるCPU・メモリリソースの量は、従来の物理サーバと同様に見積もり・決定します。
 - サービスドメインのCPU・メモリリソースは別途確保しておく必要があります。
- 各ドメインに割り当てるメモリ容量の合計は、ハイパーバイザの動作に必要な分を除くため、物理的に搭載しているメモリ容量よりも小さくなります。
 - SPARC M12の場合2.5 GB、SPARC M10の場合2 GB分、小さくなります。
- オーバーコミット機能（物理搭載以上のCPUやメモリを擬似的に割り当てる機能）はありません。
 - ドメインに必要となるリソース、拡張予定のリソースを満たす、CPU・メモリを搭載したサーバを導入します。

■ 割り当て方法

- 各ドメインに割り当てるCPUの量は「コア単位 or スレッド単位」で指定します。
- 3種類の割り当て方式があり、ドメインごとに選択することができます。
 - ① コア数指定 (コア単位) : CPUのコア数を指定
 - ② コアID指定 (コア単位) : CPUのコアID (物理的な位置) を指定
 - ③ スレッド数指定 (スレッド単位) : CPUのスレッド数を指定



■ メリットとデメリット

■ CPUの割り当ては、「**① コア数指定**」を推奨します。

- 以下の表のとおり、「コア数指定」は、性能・信頼性・柔軟性に優れた割り当て方式です。

	① コア数指定	② コアID指定	③ スレッド数指定
性能	○ 問題なし	○ ・問題なし ・最適割り当て可	△ 性能が低下する 可能性あり
CPUコアの自動交替機能 (信頼性)	○ 対応	✕ 非対応	○ 対応
動的構成変更 (柔軟性)	○ 可能	✕ 不可	○ 可能
割り当ての粒度 (リソースの有効活用度)	△ CPUコア単位	△ CPUコア単位	○ CPUスレッド単位

• より細かい粒度でCPUを割り当てる場合は、「③ スレッド数指定」を使用します。

- SPARC M12の場合、4スレッド単位でドメインにCPUを割り当ててください（SPARC M10は1スレッド単位）。
- 同一コア内のCPUスレッドを別々のドメインに割り当てた場合、片方のドメインの負荷が高くなると、もう片方のドメインの性能に影響を与えることがあります。

• 「② コアID指定」は高度な設計のため、本書では詳細を記述していません。

■ 未使用のCPUコア（ライセンス未購入）を1コア以上確保します。

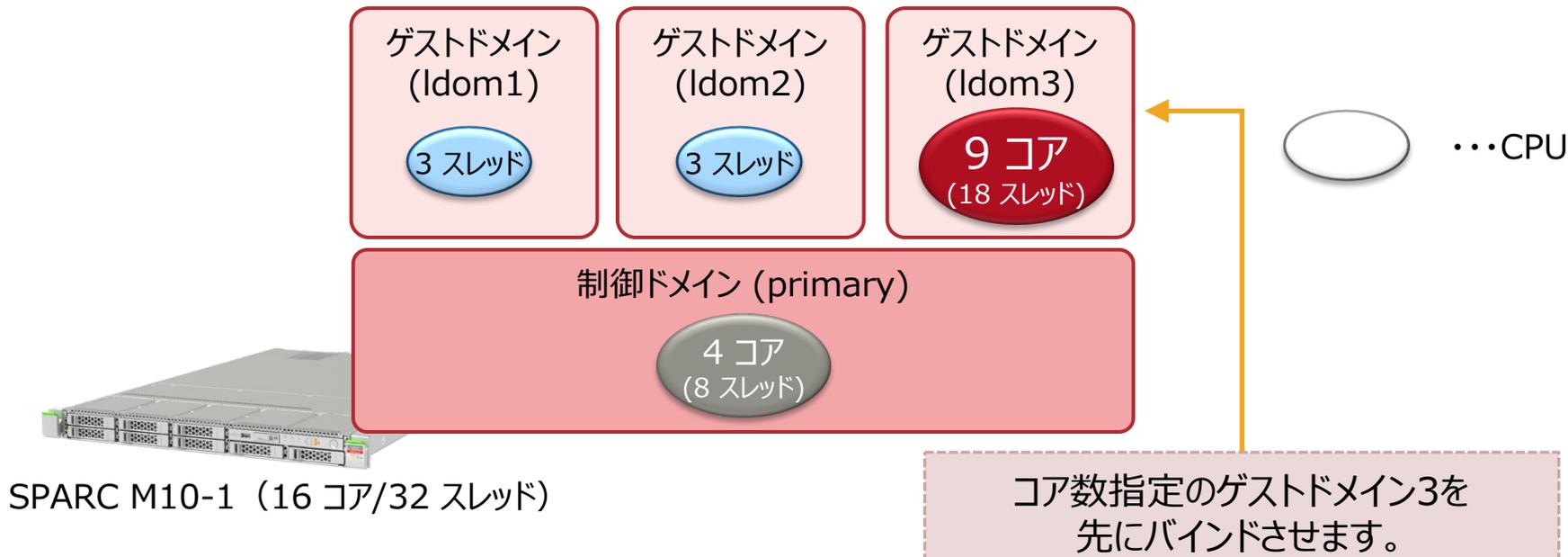
- 「CPUコアの自動交替機能」を有効にするためです。

《参考》ドメインのバインド順序 (1/2)

CPU割り当て方式が「コア数指定」のドメインと「スレッド数指定」のドメインの両方が存在する場合、「コア数指定」のドメインを先にバインドしてください。

- 「スレッド数指定」のドメインを「コア数指定」より先にバインドした場合、一部のドメインがバインドできないことがあります。
- ドメインを起動する順序に制約はありません。

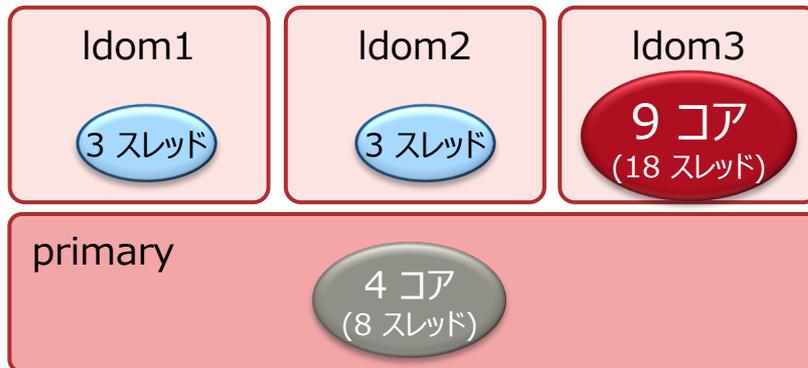
例：ゲストドメイン1とゲストドメイン2が「スレッド数指定」のドメイン、
ゲストドメイン3が「コア数指定」のドメインの場合



《参考》ドメインのバインド順序 (2/2)

■ ドメインをバインドしたときのCPUリソースの割り当てイメージ

構成例 :



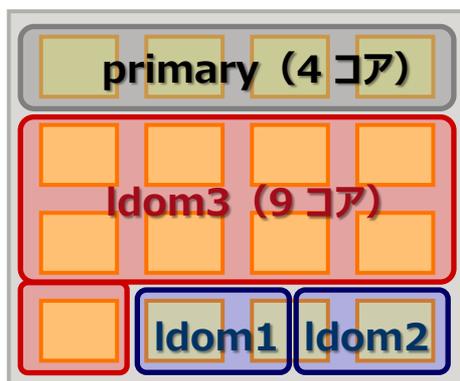
SPARC M10-1
(16 コア/32 スレッド)



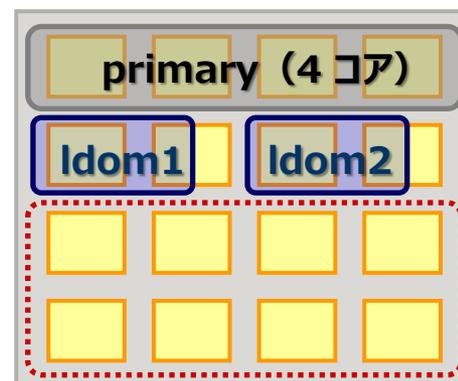
コア数指定 (Idom3) のドメインを先にバインド

スレッド数指定 (Idom1, Idom2) のドメインを先にバインド

□ : CPUコア



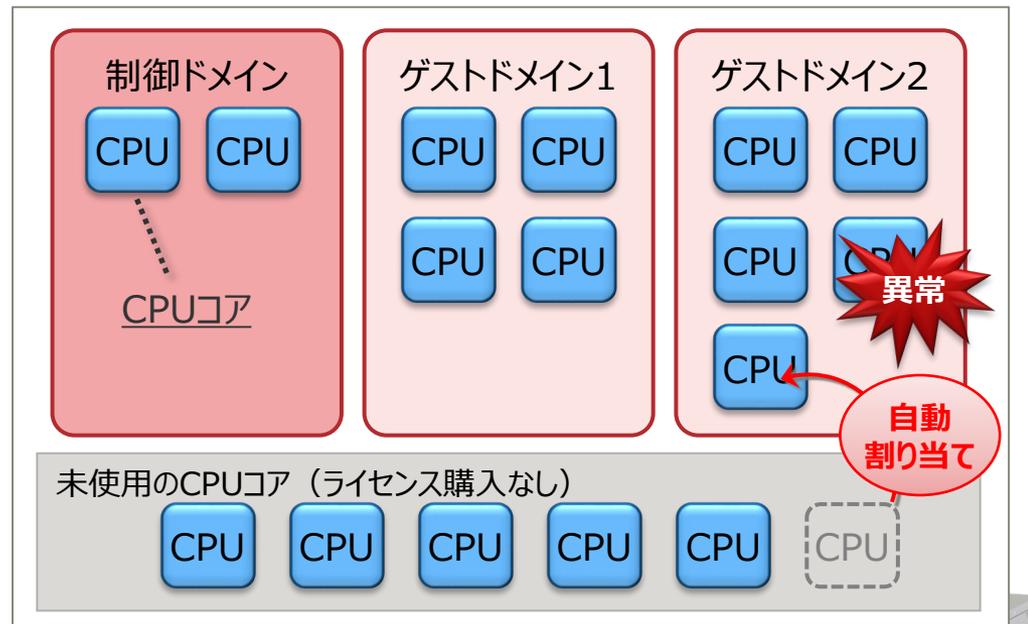
→ 各ドメインに正常にCPUリソースをバインド



→ CPU「コア」のリソース不足が発生して、一部ドメイン (Idom3 : 9 コア) のバインドが失敗

SPARC M12/M10では、CPUコアに異常が生じても、CPUコアの自動交替機能によりドメインの性能を維持したまま業務システムの稼動を継続できます。

ドメインのCPUコアに異常が発生
↓
当該のCPUが動的に縮退
↓
未使用のCPUコアを自動的に割り当て
↓
CPUリソースを維持して、業務継続



※SPARC M12/M10で有効な機能です。

SPARC M10-1 (16 CPUコア搭載)

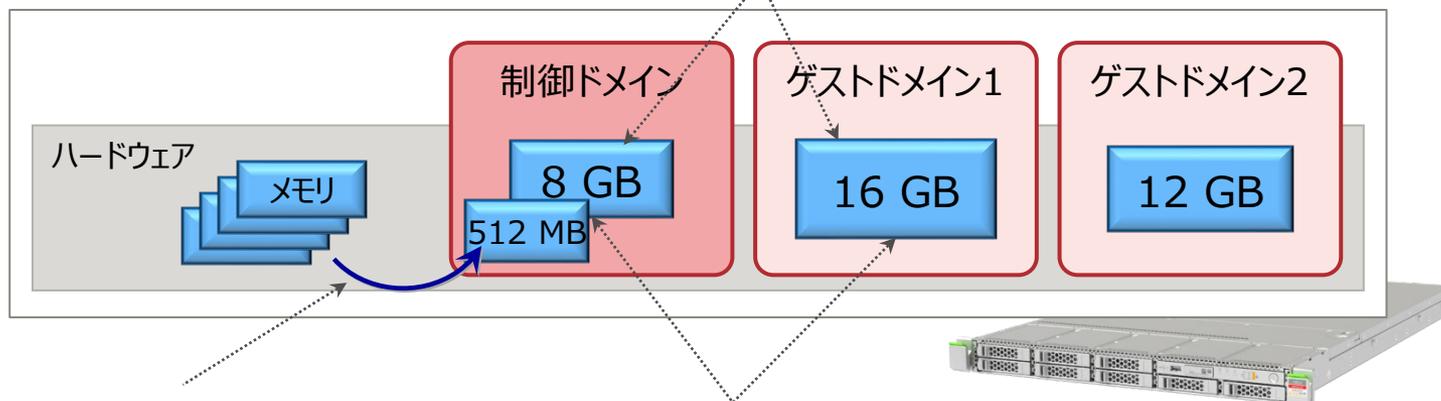


- 自動で割り当てられる分のCPUコアアクティベーション (CPUのライセンス) を購入しておく必要はありません。
※上図の場合、10 CPUコア分の購入でOKです。
- 「未使用のCPUコア (CPUコア アクティベーションで有効化していないコア)」を用意しておく必要があります。詳細は、以下をご参照ください。
 - 『初期投資を抑制しつつ、業務拡張や異常時にも柔軟に対応するCPUコア アクティベーション』
<http://www.fujitsu.com/jp/products/computing/servers/unix/sparc/featurestories/technology/scalability/cod/>

■ 割り当て方法

- 各ドメインに割り当てるメモリの容量は「MB単位 or GB単位」で指定します。…①
 - 割り当てるメモリの容量は、「4 MB」の倍数で指定します。
- ドメインを停止せずに割り当て容量を変更（追加 or 削減）する場合は、メモリの変動サイズが「256 MB単位」となるように指定します。…②
 - ドメインを停止させる場合は、256 MB単位である必要はありません。
- 各ドメインには、「4 GB」以上のメモリを割り当てます。…③
また、SPARC M12では、ゲストドメイン以外のドメインに対して「6 GB」以上のメモリを割り当てます。

① 各ドメインに、割り当てるメモリ容量のみを指定



② 動的に追加・削減する場合は、
256 MB単位で容量を指定

③ 4 GB以上のメモリを割り当て

■ 制限事項

- ハイパーバイザが必要とする「2 GB」以上のメモリを残します。
 - 例えば「64 GB」のメモリが搭載されたサーバの場合、ドメインに割り当てるメモリの合計は「62 GB」以下になるように設計します。
 - SPARC M12-1 / M12-2 / M12-2Sでは、同様の理由から「2.5 GB」以上のメモリを残します。
- 業務システム稼動中のメモリ追加を想定する場合は、さらに空きのメモリを確保します。
- 以下のような場合、業務システム稼動中のメモリ削減が実行できないことがあります。
 - ドメインのOSやミドルウェアなどがメモリを大量に使用している場合、動的な削減ができない、もしくは少ない容量のメモリしか削減できないことがあります。
 - このような場合も、ドメインを一度停止させれば、自由にメモリを削減させることができます。

4. I/Oデバイス設計のポイント

各ドメインには、I/Oデバイスを割り当てます。本章では、I/Oデバイスの割り当てに関する基本的なルールと推奨する設計指針を解説します。

■ 割り当ての種類

■ I/Oデバイスの割り当てには、以下の3種類があります。

本書の解説範囲

① 「仮想I/O」の割り当て（ゲストドメイン）

- 物理I/Oデバイスを仮想化し、複数のドメインに「占有または共有」させます。
- 物理I/Oデバイスを多数のドメインで共有可能なため、**柔軟性を重視**する場合に選択します。
- 動的な構成変更が可能。

② 「VF」の割り当て（I/Oドメイン）

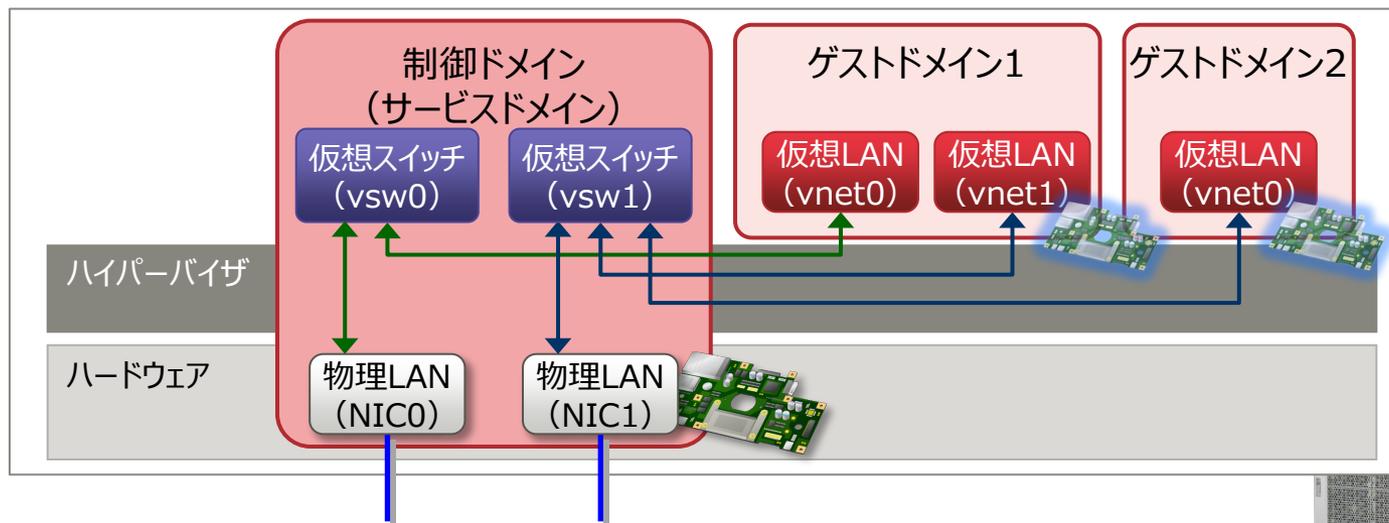
- 物理I/Oデバイスのポートを仮想的なポート(以下、VF)に分割し、複数のドメインに「占有または共有」させます。
- 仮想化処理はハードウェアで実行されるため、仮想化によるオーバーヘッドが発生しません。
- そのため、**性能を重視**する場合に選択します。
- 基本的に、動的な構成変更が可能（ただし、静的な変更が必要になる場合あり）。

③ 「PCIeバス」の割り当て（I/Oルートドメイン）

- ドメインにPCIeバス配下の物理I/Oデバイスをすべて「占有」させます。
- そのため、**性能と障害隔離性を重視**する場合に選択します。
- 基本的に、動的な構成変更が可能（ただし、静的な変更が必要になる場合あり）。

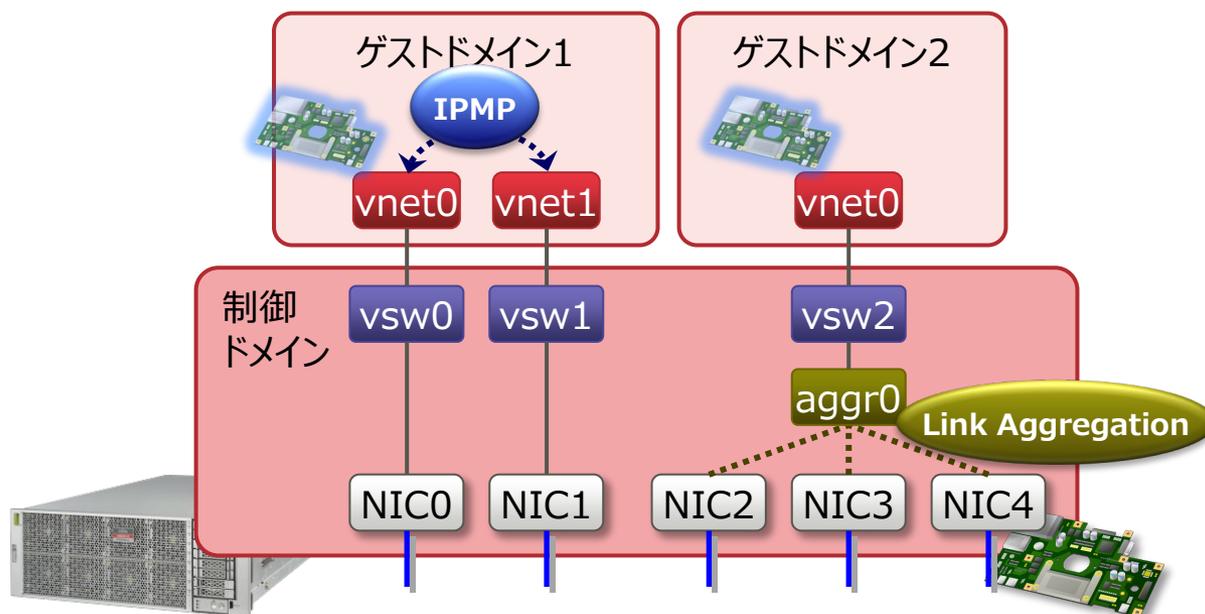
■ 基本的な考え方

- サービスドメインである制御ドメインまたはI/Oルートドメインの物理LANポートから仮想スイッチ（vsw）を構築します。
- 仮想スイッチを元に仮想LANポート（vnet）を作成し、ゲストドメインに割り当てます。
 - 通常の物理LANポートと同様に、ゲストドメインのOS上で以下を設定できます。
 - デバイスの有効化、IPアドレスの設定
 - ネットワークの冗長化設定（IPMP, PRIMECLUSTER GL）
- 1つのvswから複数のvnetを作成して、別々のドメインに割り当てることができます（物理LANポートの共用）。



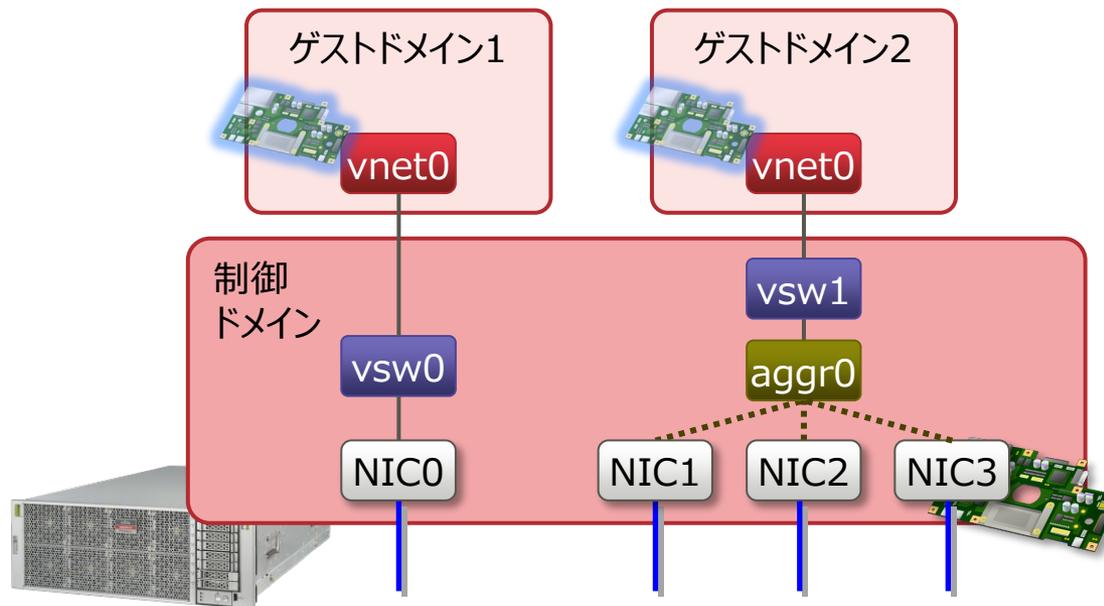
■ 仮想ネットワークの冗長化

- vnetでは、物理LANポートと同様に、以下の2つの方法による冗長化構成が可能です。
 - ゲストドメインでの冗長化
 - 2つの物理LANポートより、それぞれでvswとvnetを構築して冗長化
 - Solaris標準のIPMPやPRIMECLUSTER GLなどを使用
 - 制御ドメインでの冗長化
 - 複数の物理LANポートをLink Aggregationで集約してvswを構築
 - Solaris標準のLink Aggregationを使用



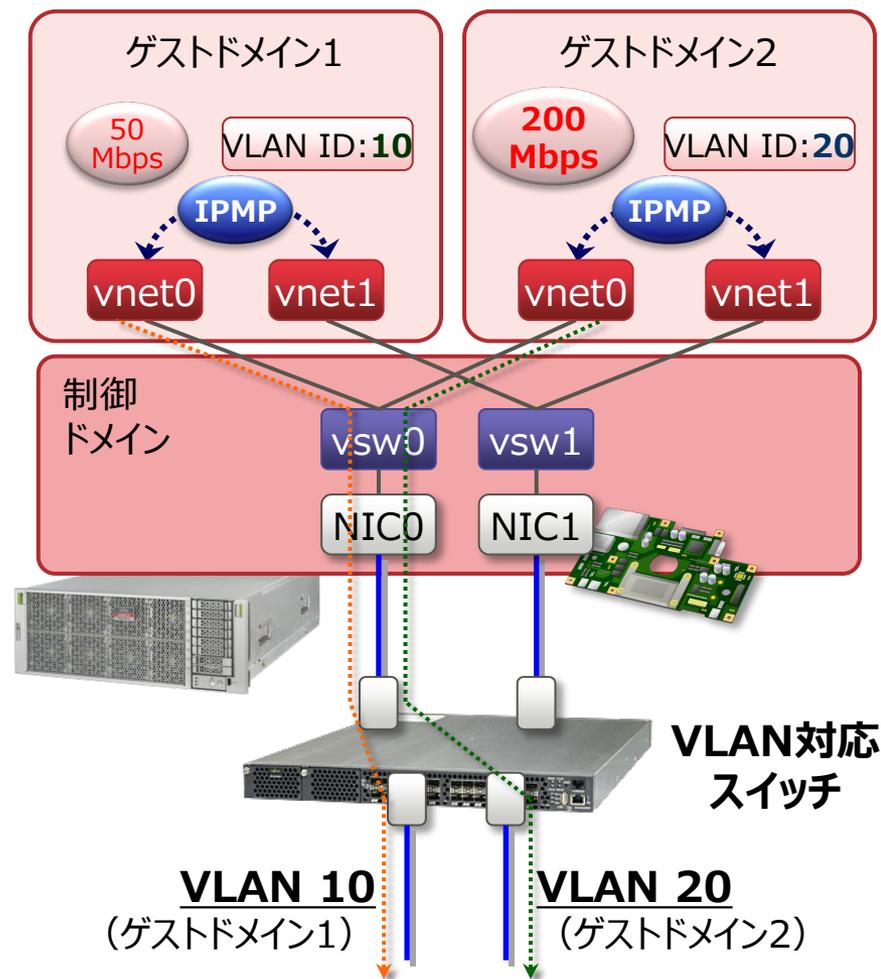
■ ネットワークの占有構成

- ゲストドメインの**ネットワーク性能を優先**する場合は、以下のいずれかの方法で物理ネットワークデバイスを1つのゲストドメインで占有させます。
 - 物理LANポートを占有（ゲストドメイン1）
 - Link Aggregationで作成したデバイスを占有（ゲストドメイン2）



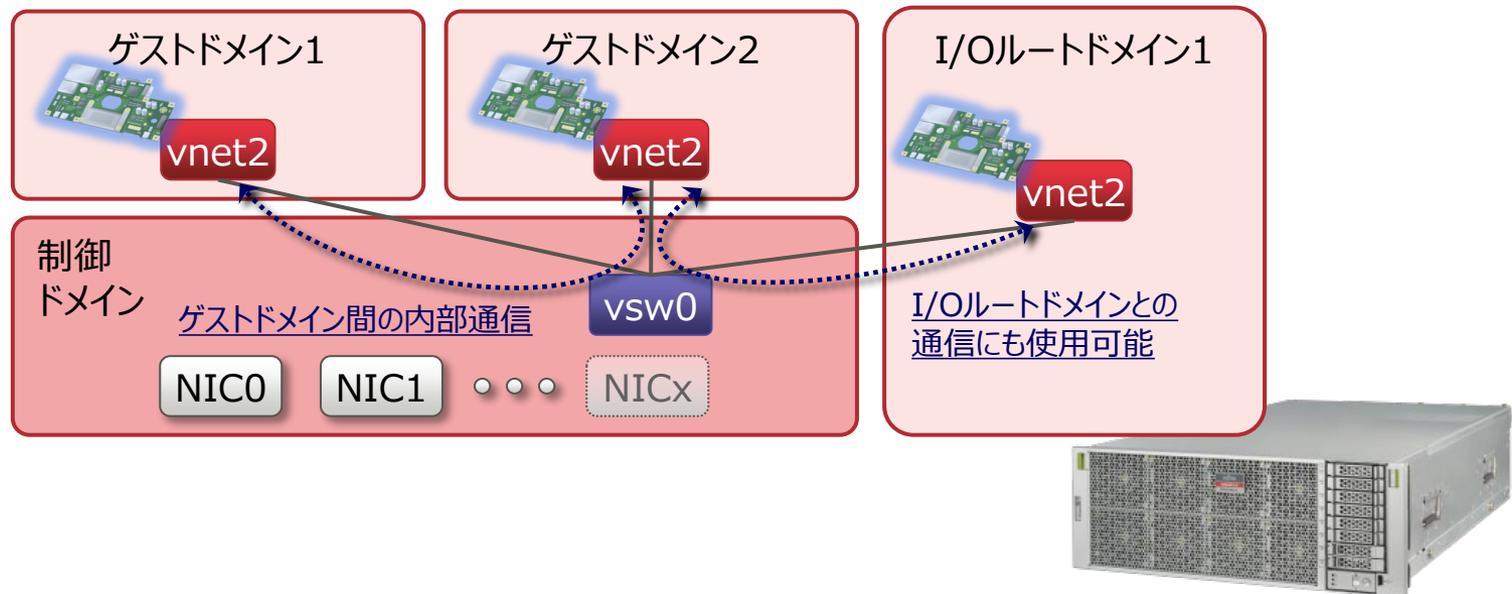
■ ネットワークの共有構成

- 使用する物理LANポートの数を抑え、**柔軟な仮想ネットワーク環境**を構築する場合は、複数のゲストドメインでネットワークデバイスを共有します。
- 「帯域制御」と「VLAN」の機能の併用が有効です。
 - 帯域制御機能により、各vnetの帯域をMbps, Gbps単位で指定（物理NICの帯域を超えない範囲の値を指定）できます。
 - VLAN機能により、サブネットの異なるネットワーク集約が可能です。ポートVLAN（タグなし）とタグVLANの片方または両方の設定が可能です。
 - タグLANとは、VLANの接続方法の1つです。データに識別子をつけることでネットワークを判別する方式です。



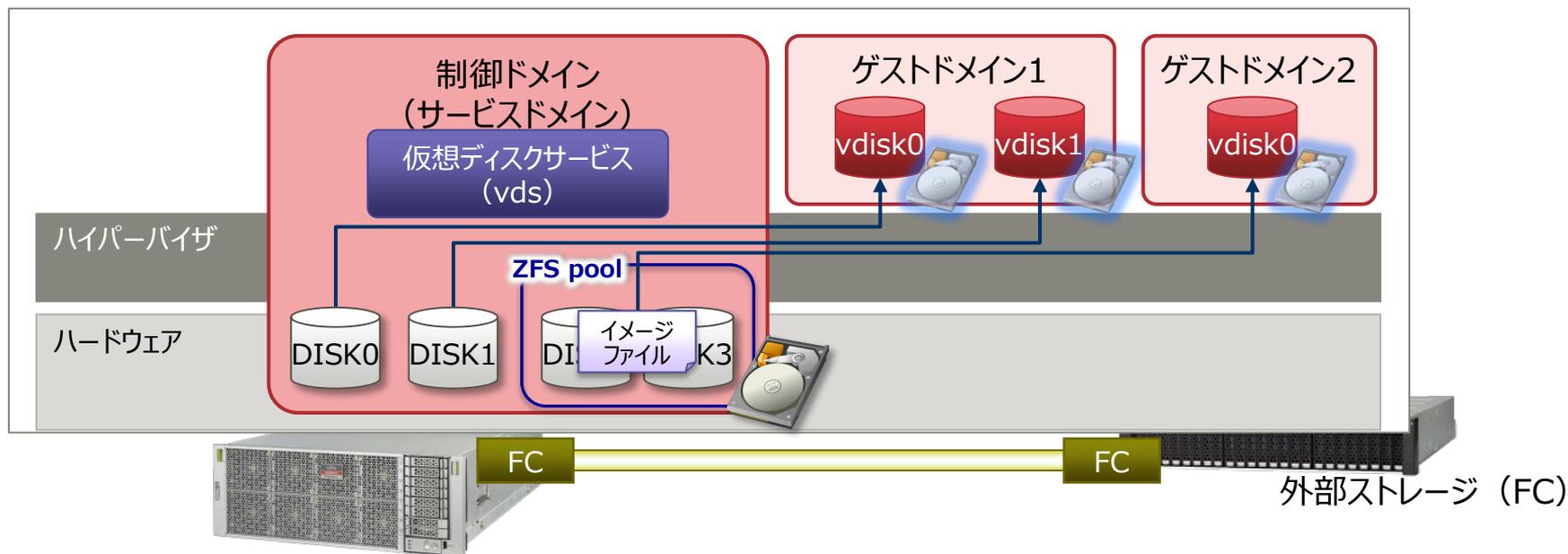
■ 内部ネットワーク機能

- ドメイン間の通信には、仮想スイッチの内部ネットワーク機能が有効です。
- サーバ外部を経由しない通信のため、**性能とセキュリティ**の両面で優れています。
- 内部ネットワーク機能では、物理NICを割り当てずに仮想スイッチを作成します。
 - 物理LANポートなしで通信を確立することができます。
 - 本機能はゲストドメインだけでなく、制御ドメインやI/Oルートドメインなどのすべてのドメイン間で適用可能です。



■ 基本的な考え方

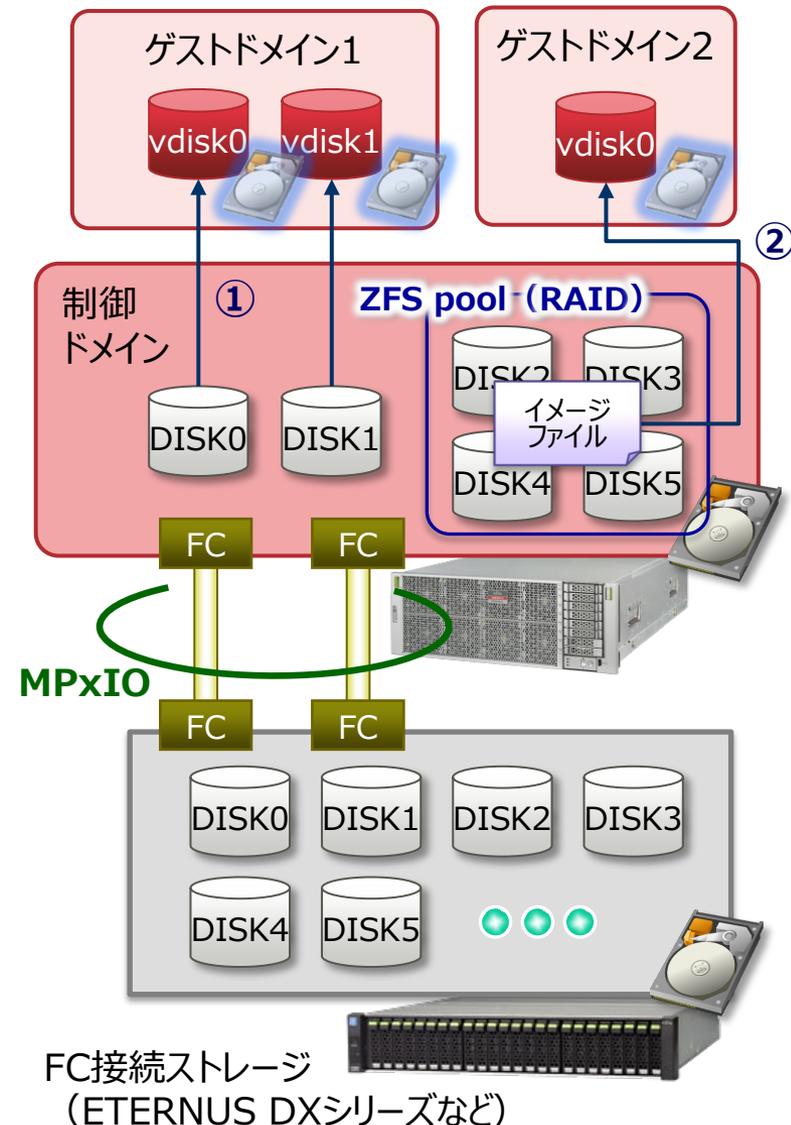
- 仮想ディスクの元となるデバイス（バックエンドデバイス）をサービスドメインで用意します。
 - 物理ディスク・外部ストレージのLUN（/dev/dsk/cXtXdXs2）
 - イメージファイル（mkfileコマンドで作成した領域）
- バックエンドデバイスを仮想ディスクサービス（vds）に登録し、ゲストドメインに割り当てます。



- ゲストドメインでは、通常の物理ディスクと同様に仮想ディスクを使用できます。
 - 仮想ディスクにはターゲットIDがありません（cXtXdXsX ⇒ cXdXsX）。
- バックエンドデバイスのマルチパスはゲストドメインではなく、サービスドメインで構築します。

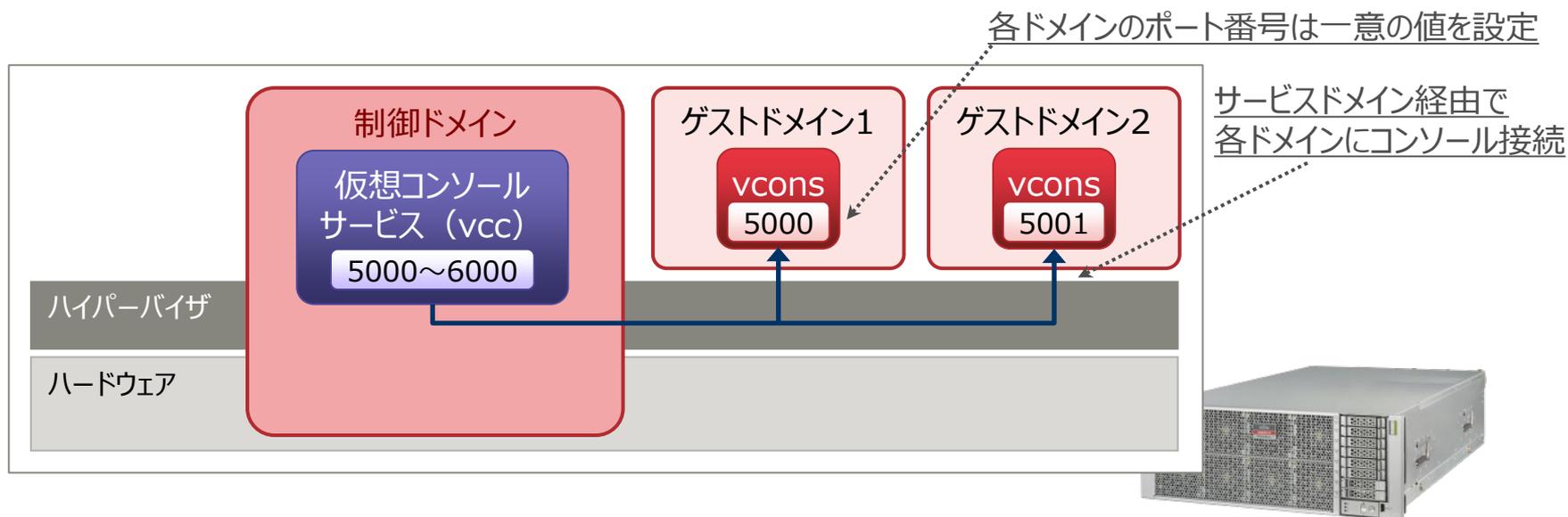
■ 障害対策

- サービスドメインで、ディスクのミラーリングやFC (ファイバーチャネル) のパスの冗長化を行います。
- ミラーリングは2種類の方法があります。
 - ディスク装置のハードRAID機能を使用する場合は、ハードRAIDを適用したディスク・LUNをそのままゲストドメインに割り当てます (①)。
 - ソフトRAID機能 (OS標準のZFSなど) を使用する場合は、ソフトRAIDによって生成されたデバイスやイメージファイルをゲストドメインに割り当てます (②)。
- FCのパスの冗長化ではSolaris標準のMPxIOやETERNUSマルチパスドライバを使用します。



■ 基本的な考え方

- サービスドメインで、仮想コンソールサービス (vcc) を構築します。
 - vccには仮想コンソール接続に使用するポート番号の範囲を指定できます (5000~6000など)。
- vccより、ほかのドメインに仮想コンソールポート (vcons) を割り当てます。
- 仮想コンソールには、telnetコマンド (プロトコル) で接続します。
 - 各ドメインのポート番号を指定することで、物理サーバと同様にtelnetによるコンソール接続ができます。
 - SSHによる仮想コンソール接続はできません。
- 本機能は、コンソール接続するすべてのドメインに設定することができます。



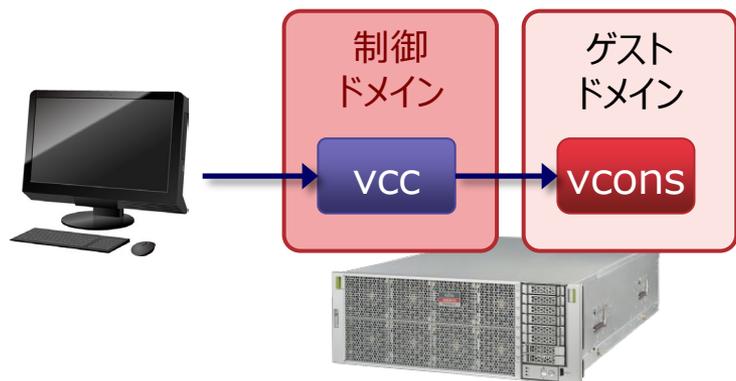
■ 仮想コンソールへの接続方法

■ 仮想コンソール接続方法は、以下の2種類を設定できます。

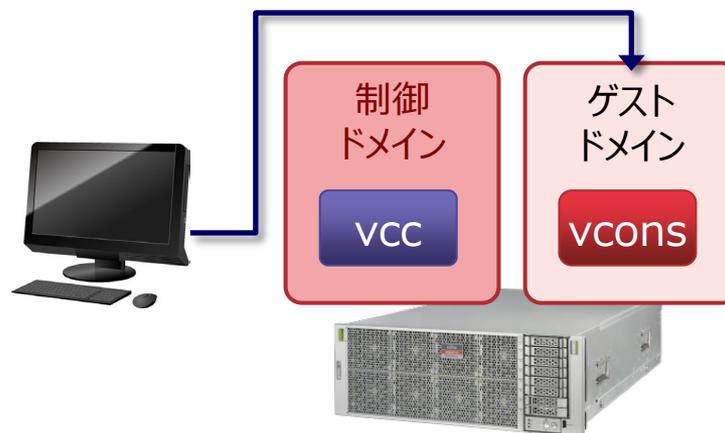
- 制御ドメインにログインした後に接続 (デフォルト)
- 外部ネットワークから直接接続

■ 仮想コンソール接続方法の変更は、サービスドメインのSMF (Service Management Facility) のプロパティを設定します。

- SMFサービス : `svc:/ldoms/vntsd`
- プロパティ : `vntsd/listen_addr`



制御ドメインからコンソール接続 (デフォルト)
[`vntsd/listen_addr = localhost`]



外部ネットワークから直接コンソール接続
[`vntsd/listen_addr = "制御ドメインのIPアドレス"`]

5. 設計サンプル

ここまでは、柔軟なスケールアップやスケールアウトを実現するためのOracle VMでの構築パターンとリソース設計について解説しました。本章では、それらのノウハウを活用した設計例を紹介します。

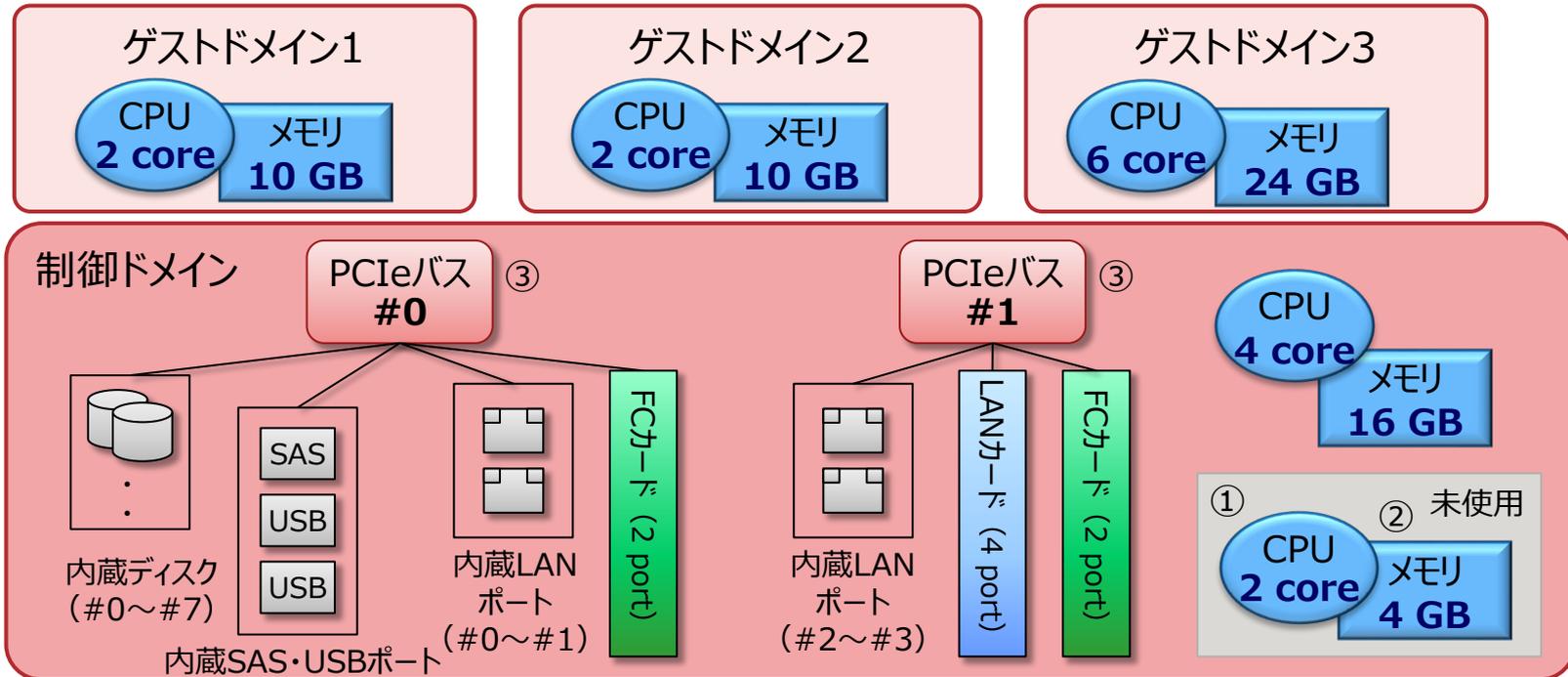
ゲストドメインの構成例 (1/3)

以下は、SPARC M10-1 (CPU: 16 core/メモリ: 64 GB) での構成例です。

■ CPU・メモリ・物理I/Oデバイス

■ 3つのゲストドメインを構築

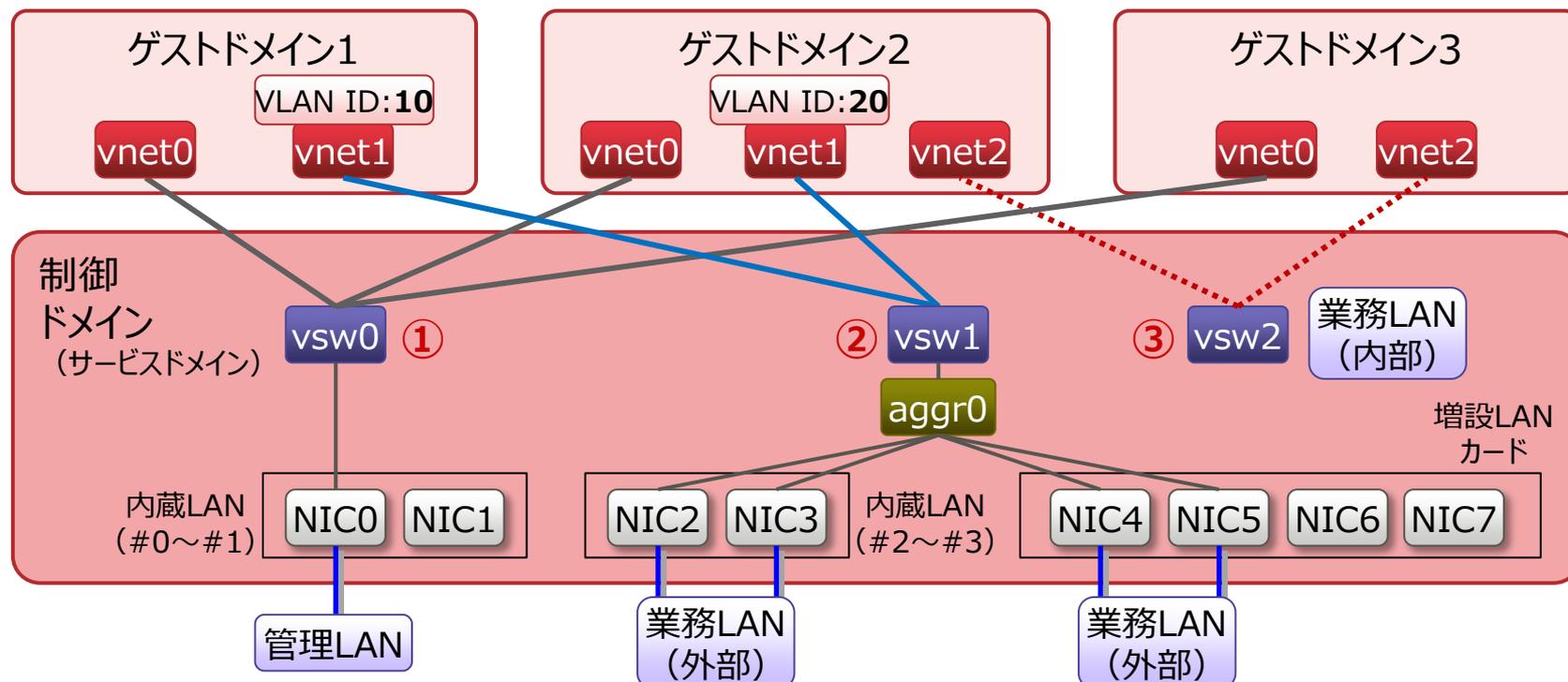
- 自動交替機能用として未使用のCPUを2コア確保…①
- ハイパーバイザが必要とするメモリ容量も含めて空きメモリを4 GB確保…②
- 仮想サービスとして提供する物理I/Oデバイスは、PCIeバスも含めてすべて制御ドメインに割り当て…③



ゲストドメインの構成例 (2/3)

■ 仮想ネットワーク

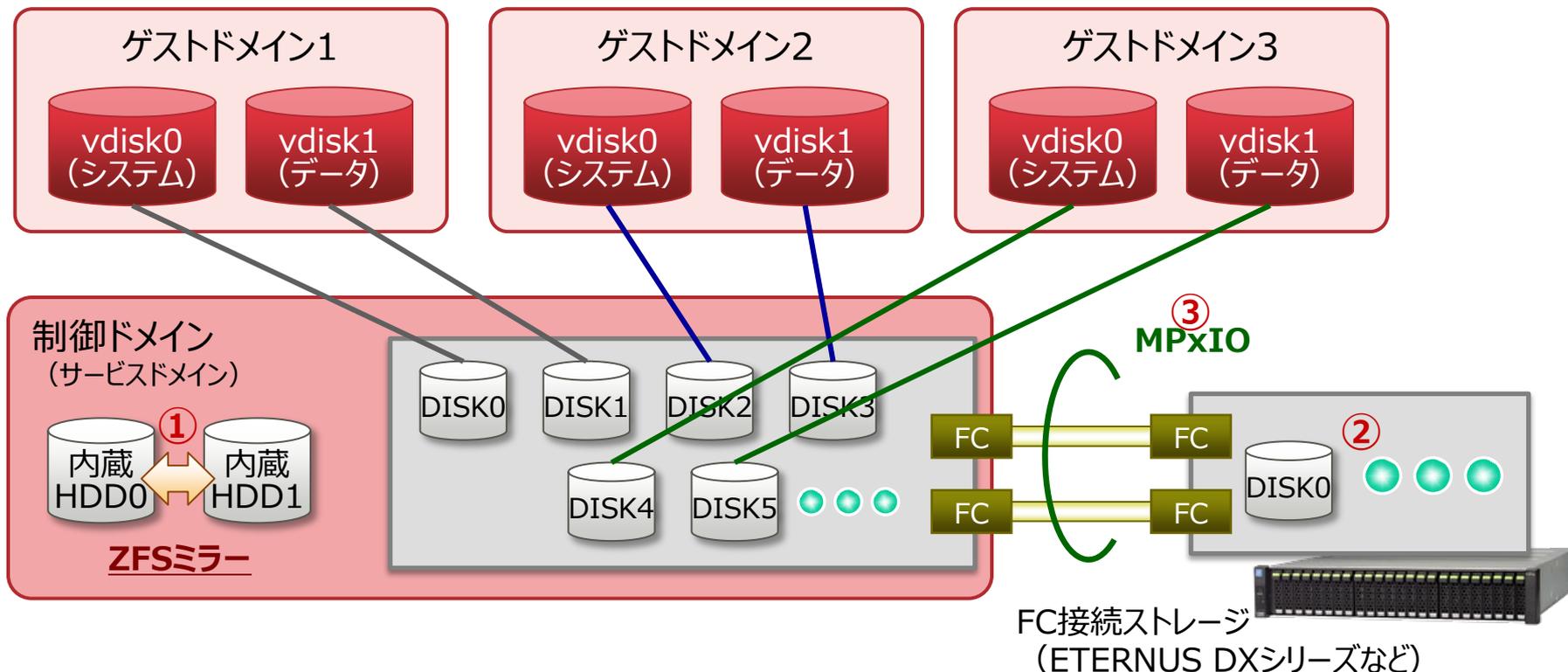
- 管理LAN (vnet0) は各ゲストドメインで1つの物理LANポートを共有 ……①
- 業務用の外部ネットワーク (vnet1) はLink Aggregationで作成したネットワークインタフェースを使用 ……②
- VLAN機能を併用して、異なるサブネットのネットワーク環境を集約
- 外部ネットワークが不要なゲストドメインは、ゲストドメイン間の通信に内部ネットワーク機能を使用 (vnet2) ……③



ゲストドメインの構成例 (3/3)

■ 仮想ディスク

- 制御ドメインのシステムディスクは内蔵HDDを使用 …①
 - OS標準のZFSによるミラーリングを構築
- ゲストドメインの仮想ディスクには、ストレージ装置のディスク (LUN) を使用 …②
 - ゲストドメイン上の冗長化設定は不要 (ストレージ装置のハードRAID機能を使用)
- FCのパスの冗長化はOS標準のMPxIOを使用 …③

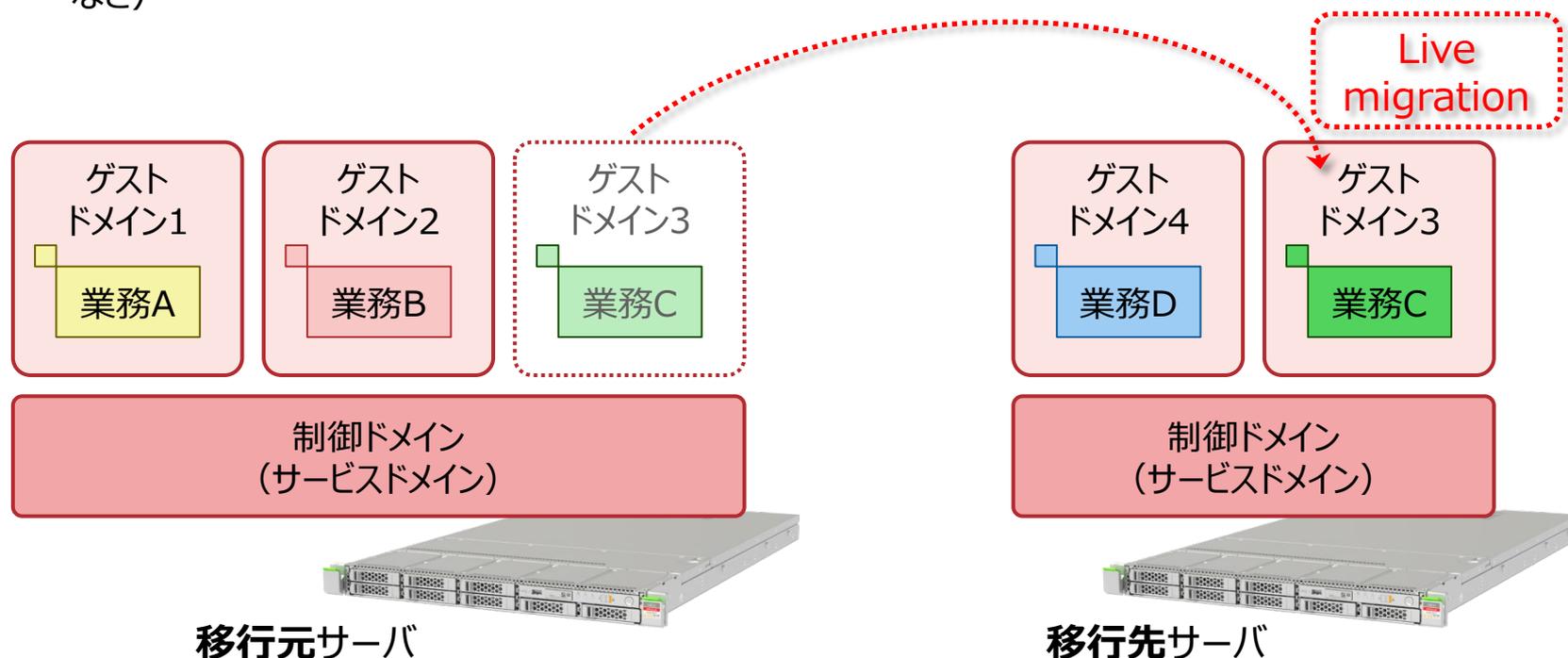


6. ライブマイグレーション

Oracle VMは起動中の仮想環境をほかのサーバへ移動させる、ライブマイグレーション機能を備えています。本章では、ライブマイグレーションを使用するための設計指針と留意事項を解説します。

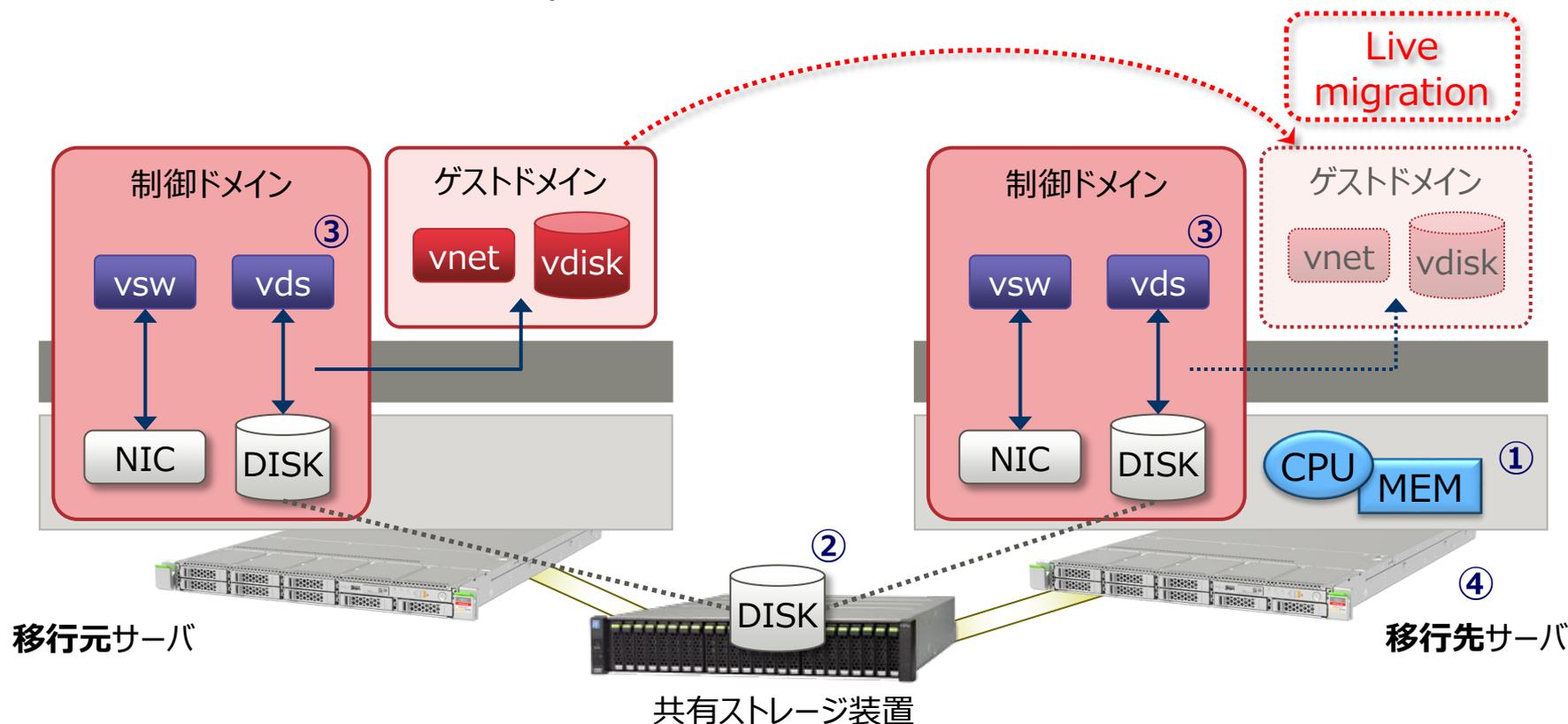
■ 概要

- ライブマイグレーションにより、ゲストドメインの業務・アプリケーションを動作させたまま、ほかのサーバへ移動させることができます。
- ゲストドメインのリソース（CPU・メモリ・仮想I/Oデバイス）を変更することなく移動できます。
- ライブマイグレーションは主に以下の用途で使用することができます。
 - 業務システム稼動中のサーバの保守・メンテナンス
 - ハードウェアリソースの最適化（CPU・メモリリソースが余っているサーバへゲストドメインを移動させる、など）



■ 設計のポイント

- 移行先サーバでCPU・メモリリソースを確保 … ①
- 仮想ディスクのバックエンドデバイスを共有ストレージに配置 … ②
- 同一の仮想サービスを定義 … ③
- 移行先サーバに対応した「cpu-arch」プロパティを設定 … ④



① 移行先サーバでCPU・メモリリソースを確保

■ 移行先サーバでゲストドメインのリソースを確保しておく必要があります。

- CPUはゲストドメインに割り当てたリソースと同数のCPUコア（スレッド単位の割り当ての場合はCPUスレッド）を確保しておきます。
- メモリはゲストドメインと同容量以上のメモリを確保しておきます。さらに、移行先サーバと同じ構成のメモリブロックを作成できるように十分な容量を確保しておきます。詳細は、以下をご参照ください。

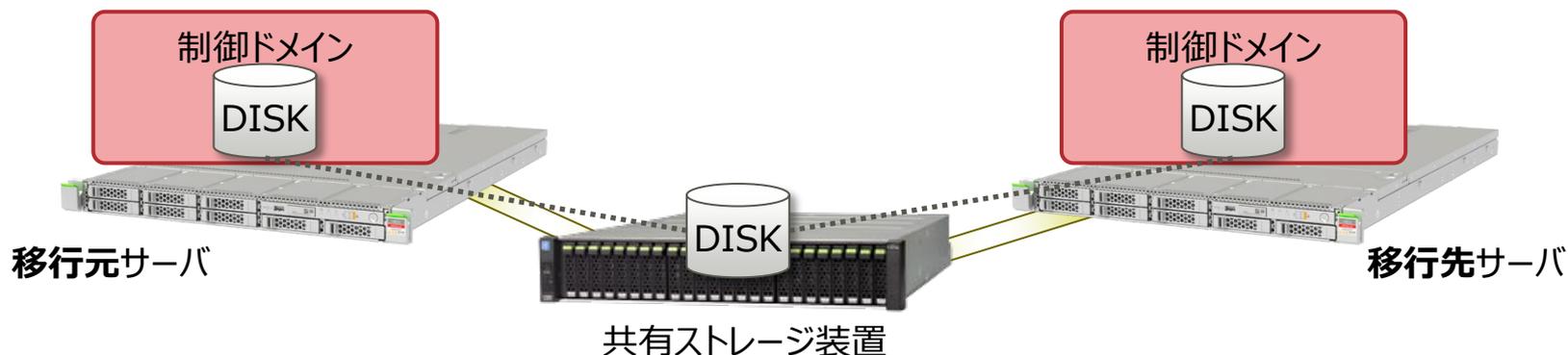
- 『ライブマイグレーション導入ガイド・運用手順書』

<http://www.fujitsu.com/jp/sparc-technical/document/solaris/#ovm>

② 仮想ディスクのバックエンドデバイスを共有ストレージに配置

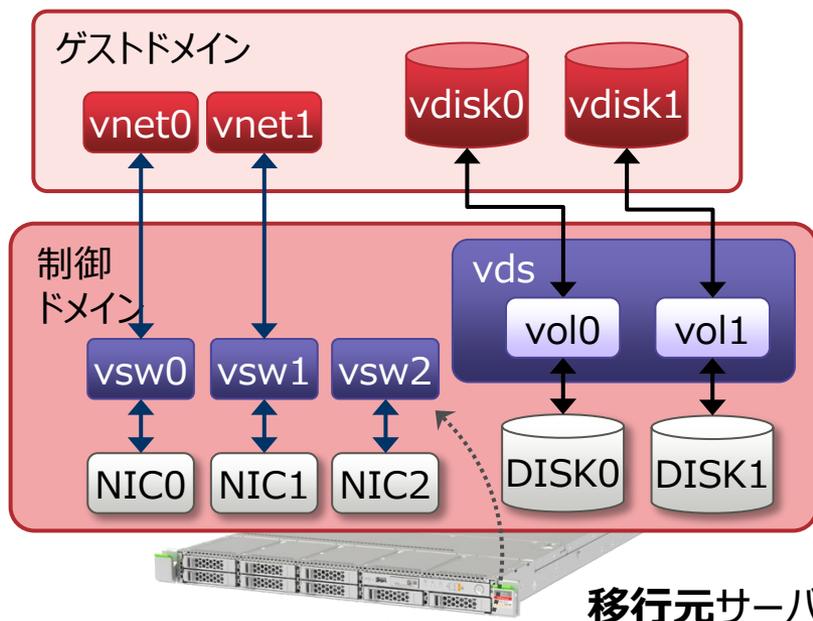
■ 移行元サーバと移行先サーバの両方の制御ドメイン（サービスドメイン）に、共有ストレージのディスクを認識させておきます。

- 通常、FC接続のストレージ装置を使用します。
- バックエンドデバイスにはディスク（LUN）を指定します。



③ 同一の仮想サービスを定義

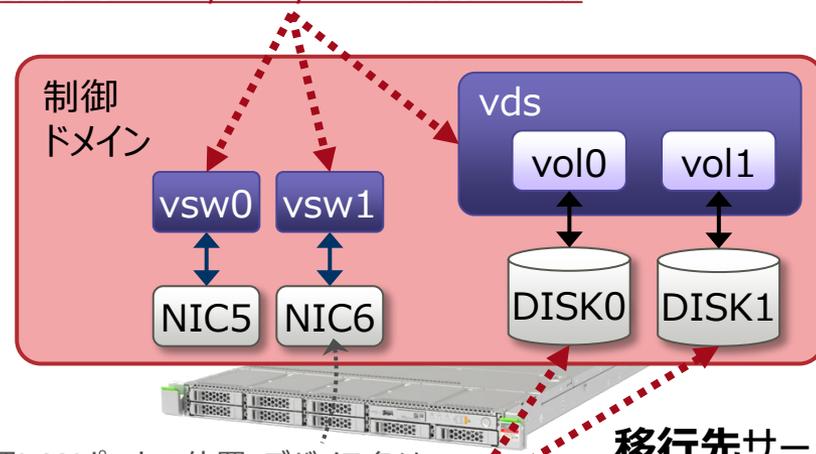
- 移行元サーバと移行先サーバのサービスドメインで、以下の名称を一致させておきます。
 - 仮想スイッチ (vsw)
 - 仮想ディスクサービス (vds)
 - 仮想ディスクボリューム (vol)
(仮想ディスクボリュームは、仮想ディスクのバックエンドデバイスをvdsに登録する際の定義名です)
- 仮想ディスクは、同一のバックエンドデバイスをvds, volに登録します。



ゲストドメインが使用していない仮想サービスは移行先に用意する必要はありません

移行元サーバ

移行元とvsw, vds, volの名称を一致



物理LANポートの位置・デバイス名は、同一にする必要はありません

移行先サーバ

移行元と同一のバックエンドデバイスを登録

- ④ 移行先サーバに対応した「cpu-arch」プロパティを設定
- 通常、ライブマイグレーションは「同じCPUタイプのサーバ間」で実行します。
 - ただし、ゲストドメインに「cpu-arch」パラメータを設定することで、CPUタイプが異なるサーバ間でのライブマイグレーションが可能となります。

「cpu-arch」プロパティの種類



※1: SPARC M10の場合、SPARC64 XとSPARC64 X+のいずれかのCPUが搭載されています。
同じCPUが搭載されたモデル同士であれば、SPARC M10-1とSPARC M10-4などの異なる機種間でも、cpu-archを「native」に設定してライブマイグレーションができます。

※2: SPARC M12の場合、Oracle VM 3.5以降かつ制御ドメインのOSがOracle Solaris 11以降で設定が可能です。

- ライブマイグレーションは「ゲストドメイン」に対する機能です。
 - ドメインに割り当てるI/Oデバイスは、すべて「仮想I/Oデバイス」にする必要があります。
 - I/OドメインとI/Oルートドメインをライブマイグレーションさせることはできません。
 - cpu-archの設定を変更する場合は、ゲストドメインを再起動させる必要があります。
 - 仮想コンソールのポート番号が変わることがあります。
 - ライブマイグレーションしたゲストドメインには自動で仮想コンソールの再設定が実行されます。そのため、仮想コンソールサービス・ポート番号が変更されることがあります。
 - 一時的にゲストドメインのCPUリソースを減らすことで、CPUコア アクティベーションキーも移行先のサーバへ移動させることができます。
 - SPARC M12/M10のCPUコア アクティベーション機能との併用は、以下の資料をご参照ください。
 - 参考 : 技術情報 Technical Park 『ビジネスの変化にスピーディに対応するSPARC M10』
<http://www.fujitsu.com/jp/sparc-technical/document/proposal/>
- ライブマイグレーションは「クラスタ」に該当する機能ではありません。
 - ライブマイグレーションは「移行元サーバの制御ドメイン」で実行させる機能です。
 - サーバ全体もしくは制御ドメインに障害が発生した場合は、ライブマイグレーションが実行できません。
- あらかじめ各ミドルウェアのライセンス形態を確認してください。
 - 移行元サーバと移行先サーバの両方のライセンス費用が必要となるミドルウェアもあります。

付録

『Oracle VM Server for SPARC Documentation』（Oracle社）

<http://www.oracle.com/technetwork/documentation/vm-sparc-194287.html>

『Oracle VM Server for SPARCを使ってみよう』

<http://www.fujitsu.com/jp/sparc-technical/document/solaris/#ovm>

『SR-IOV（Single Root I/O Virtualization）ドメイン環境 構築ガイド』

<http://www.fujitsu.com/jp/sparc-technical/document/solaris/#ovm>

『SPARC M12/M10 ドメイン構築ガイド』

<http://www.fujitsu.com/jp/sparc/lineup/m12-1/documents/>

■ SPARC/Solarisの構築に役立つドキュメントが満載

- ハイパーバイザベースの仮想化：

Oracle VM Server for SPARC

- Solarisベースの仮想化：

Oracle Solarisゾーン

- 最新ファイルシステム：

ZFS (Zettabyte File System)

- Solaris 8/9環境をそのままSolaris 10へ：

Oracle Solaris Legacy Containers

など

今すぐクリック!! 

<http://www.fujitsu.com/jp/sparc-technical/>

版数	更新日時	更新内容
初版	2014年12月	新規作成
第1.1版	2016年6月	仮想ディスクのバックエンドデバイスの設計方針を更新
第2.0版	2018年4月	SPARC M12に対応 Oracle VM 3.5に対応 構成見直し 分冊化

■ 使用条件

■ 著作権・商標権・その他の知的財産権について

- コンテンツ（文書・画像・音声等）は、著作権・商標権・その他の知的財産権で保護されています。本コンテンツは、個人的に使用する範囲でプリントアウトまたはダウンロードできます。ただし、これ以外の利用（ご自分のページへの再利用や他のサーバへのアップロード等）については、当社または権利者の許諾が必要となります。

■ 保証の制限

- 本コンテンツについて、当社は、その正確性、商品性、ご利用目的への適合性等に関して保証するものではなく、そのご利用により生じた損害について、当社は法律上のいかなる責任も負いかねます。本コンテンツは、予告なく変更・廃止されることがあります。

■ 輸出または提供

- 本製品を輸出又は提供する場合は、外国為替及び外国貿易法及び米国輸出管理関連法規等の規制をご確認の上、必要な手続きをおとりください。

■ 商標

- UNIXは、米国およびその他の国におけるオープン・グループの登録商標です。
- SPARC Enterprise、SPARC64、SPARC64ロゴおよびすべてのSPARC商標は、米国SPARC International, Inc.のライセンスを受けて使用している、同社の米国およびその他の国における商標または登録商標です。
- OracleとJavaは、Oracle Corporationおよびその子会社、関連会社の米国およびその他の国における登録商標です。
- その他各種製品名は、各社の製品名称、商標または登録商標です。

