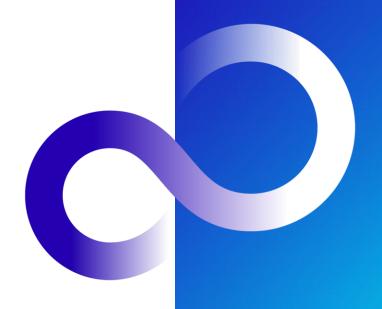


SR-IOV (Single Root I/O Virtualization) ドメイン環境構築ガイド

2016年 6月(第1.0版) 富士通株式会社



はじめに 1/2



■ 目的

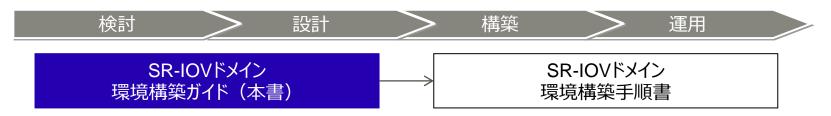
Oracle VM Server for SPARCにおけるI/Oリソースの割り当て方法の1つであるSingle Root I/O Virtualization (以降、SR-IOV) の概要や、SR-IOVを利用してI/Oドメインを構築する手順について紹介します。

■ 対象読者

- Oracle Solaris 11の基礎知識を有している方
- Oracle VM Server for SPARCの基礎知識を有している方

■ 留意事項

- 本資料は、Oracle VM Server for SPARC 3.3 / Oracle Solaris 11.3に基づいています。
 - ※Oracle VM Server for SPARCのバージョンが3.2以前、またはOracle SolarisのバージョンがOracle Solaris 10の場合、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。
- Oracle VM Server for SPARCの基本的な情報は、以下の資料をご参照ください。
 - Oracle VM Server for SPARC を使ってみよう
 http://www.fujitsu.com/jp/sparc-technical/document/#ovm-use-11
- Oracle VM Server for SPARCの最新情報は、Oracle社のマニュアルをご参照ください。
 - Oracle VM Server for SPARC Documentation
 http://www.oracle.com/technetwork/documentation/vm-sparc-194287.html
- ■ドキュメントの位置づけ



はじめに 2/2



■ 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	

目次



- 1. Oracle VM Server for SPARCにおけるI/Oリソースの割り当て
- 2. SR-IOV (Single Root I/O Virtualization) の概要
- 3. 注意事項
- 4. 構築手順
- 5. 参考情報
- ■付録



1. Oracle VM Server for SPARCにおけるI/Oリソースの割り当て

- I/Oリソースの割り当て方法の種類
- PCIeバス単位の割り当て -I/Oルートドメイン-
- SR-IOVによるI/O割り当て -I/Oドメイン-
- 仮想I/OによるI/O割り当て -ゲストドメイン-
- I/O割り当ての仕様比較

I/Oリソースの割り当て方法の種類



- Oracle VMには、主に3種類のI/O割り当て方法があります。
 - 割り当て方法により、ドメインの名称が変わります。

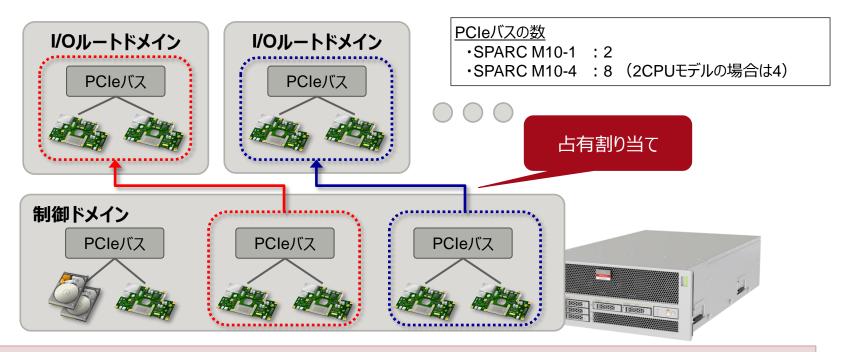
I/O割り当て方法	ドメインの名称	割り当て単位・仕組み
PCIe バス単位	I/Oルートドメイン	PCIeバス (PCIeバスに紐づくPCIeカードやその他I/Oデバイスを含む)を割り当て、占有させます。
SR-IOV	1/0ドメイン	I/Oデバイスの物理的なポート(LANポートやFCポートなど)を <u>仮想的なポートに分割</u> し、割り当てます。
仮想I/O	ゲストドメイン	I/Oデバイス(主にディスクとLANポート)を 仮想化して、仮想I/Oデバイスとして割り当 てます。

次ページ以降で、これらのI/O割り当ての仕組みについて解説します。

PCIeバス単位の割り当て -I/Oルートドメイン-



- 物理I/OデバイスをPCIeバス単位でドメインに割り当てます。
 - 1つのドメインでI/Oデバイスを占有します。
 - PCIeバスの数(作成可能なI/Oルートドメインの数)は、サーバのモデルにより異なります。



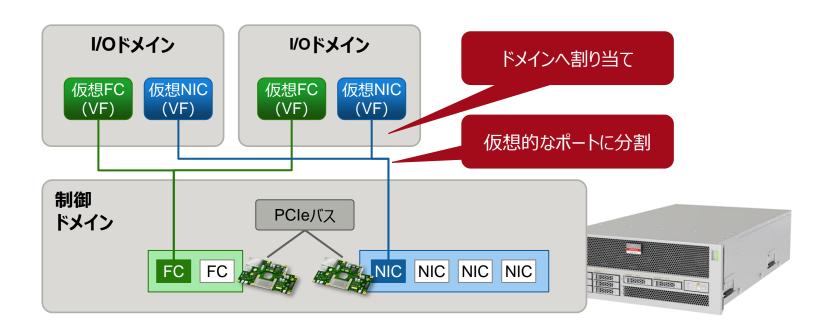
PCIeバスごとに、含まれるI/Oデバイスは異なります。詳細は、以下の資料をご参照ください。

- 『SPARC M10 システム システム運用・管理ガイド 』「付録A SPARC M10システムのデバイスパス一覧」 http://www.fujitsu.com/jp/sparc/lineup/m10-1/documents/
- ・『SPARC M10 / Oracle VM Server for SPARC 設計のポイント』 http://www.fujitsu.com/jp/sparc-technical/document/#ovm-design-11

SR-IOVによるI/O割り当て -I/Oドメイン-



- I/Oデバイスの物理的なポート(LANポートやFCポートなど)を 仮想的なポートに分割し、ドメインへ割り当てます。
 - 1つの物理ポートを1つのドメインで占有、または複数のドメインで共有させることができます。
 - 分割可能な数は、デバイスにより異なります。(「《参考》VFの作成可能数」参照)

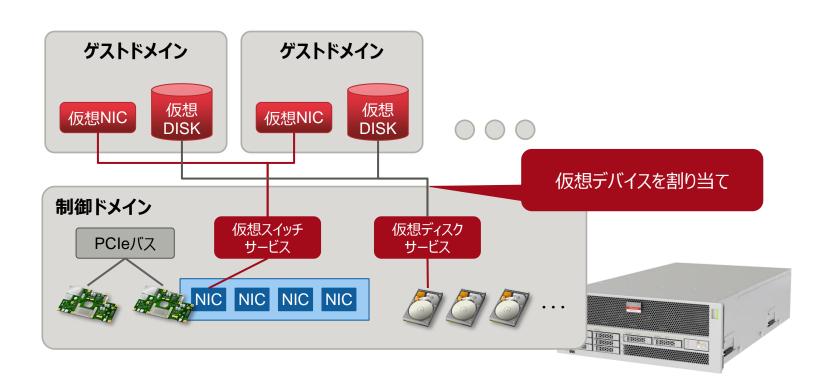


SR-IOVによるI/O仮想化処理は、ソフトウェアではなくハードウェア(PCIデバイス)上で実行されます。 そのため、仮想I/O割り当て(次ページ参照)よりも、高いI/O性能の仮想化環境を構築できます。

仮想I/OによるI/O割り当て -ゲストドメイン-



- I/Oデバイス(主にディスクとLANポート)を仮想化して割り当てます。
 - 仮想I/Oデバイスを1つのドメインで占有、または複数のドメインで共有させることができます。
 - 最も柔軟性が高く、構築が容易な割り当て方式です。



I/O割り当ての仕様比較



_	***************************************				
	項目	PCIeバス割り当て	SR-IOV	仮想I/O	
	I/O性能	◎ 物理環境と同等	◎ 物理環境と同等	○ 性能のオーバーヘッドが 生じる場合あり	
柔軟性	ドメインの 作成可能数 (※1)	△ 8 (PCIeバスの数に依存)	◎ 128 (CPU数、物理I/O数 (ポート数)に依存)	◎ 128 (CPU数に依存)	
	物理I/Oの 共有	X 共有不可	○ 複数のドメインで共有 (※2)	◎ 複数のドメインで共有	
	ライブ マイグレーション	× 不可	X 不可	◎ 可能	
運用	ハードウェア 障害時の影響度	◎ ハードウェア層で隔離	○ ハードウェアは共有	○ ハードウェアは共有	

本書では、SR-IOVによるI/O割り当ての機能や、I/Oドメインの構築手順を解説します。

- (※1) 制御ドメインを含めた、作成可能なドメインの数を表記しています。 SPARC M10-4 (4CPU) を最大構成として記述しています。
- (※2) サーバに搭載されているPCIカードおよびI/Oポートの数に依存します。 SPARC M10の場合、PCIeバスあたり24個のドメインまで共有可能です。



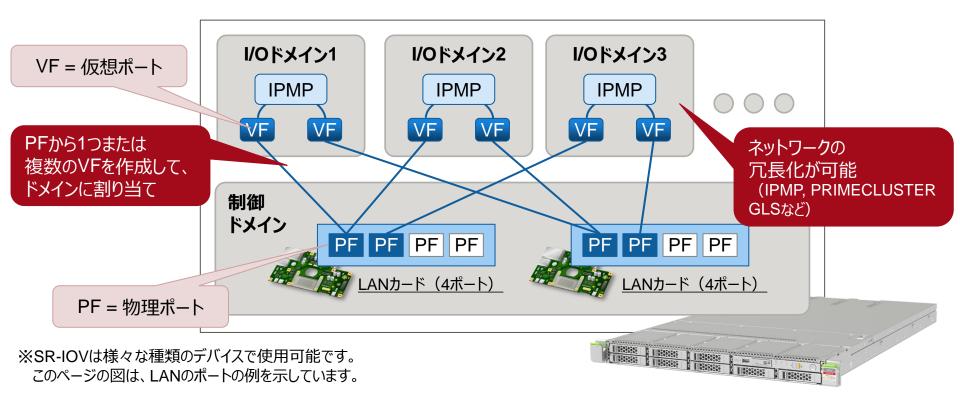
2. SR-IOV (Single Root I/O Virtualization) の概要

- SR-IOVを利用した物理ポートの分割と割り当て
- SR-IOVにおけるI/O冗長化
- 物理I/O資源の有効活用
- ■適用シーン

SR-IOVを利用した物理ポートの分割と割り当て



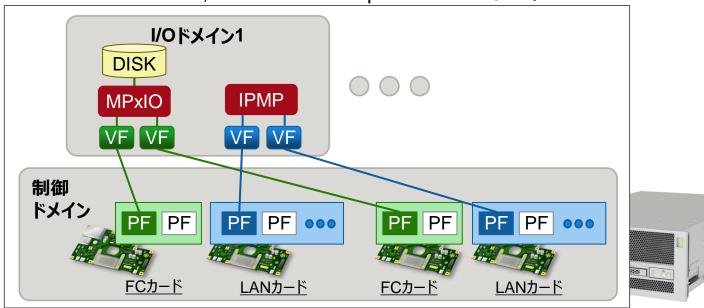
- I/Oデバイスの物理ポート(LANやFCなど)を仮想的に分割して、 ドメインに割り当てます。
 - 1つの物理ポートを1つのドメインで占有、または複数のドメインで共有させることができます。
 - I/Oのポートを以下のように呼びます。
 - PF (Physical Function): PCIカードやオンボードデバイスなどの、物理的なポートを指します。
 - VF (Virtual Function): PFをもとにして作成する仮想的なポートです。



SR-IOVにおけるI/O冗長化 1/2



- I/Oドメインの仮想ポート(VF)は冗長化させることを推奨します。
 - 複数の物理ポート (PF) からI/OドメインにVFを割り当て
 - I/Oドメイン上でマルチパスを設定(マルチパスソフトは主に以下を使用可)
 - ネットワークのマルチパス : IPMP, PRIMECLUSTER GLS (※1)
 - FCのマルチパス : MPxIO, ETERNUS Multipath Driver (※2)



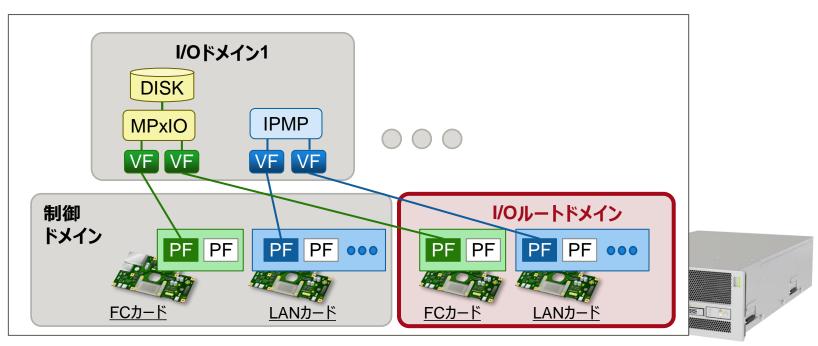
複数のデバイスからVFを作成して割り当てることで、 ハードウェアやネットワークの耐障害性を高めることができます。

- (※1) Link Aggregationは対応していません。
- (※2) ETERNUS Multipath Driverは、VF提供ドメインの二重化構成(次ページ参照)には対応していません。

SR-IOVにおけるI/O冗長化 2/2



- VFを提供するドメインを冗長化することもできます。
 - I/Oルートドメインを作成
 - 制御ドメインおよびI/Oルートドメインから、I/OドメインにVFを割り当て
 - I/Oドメイン上でマルチパスを設定

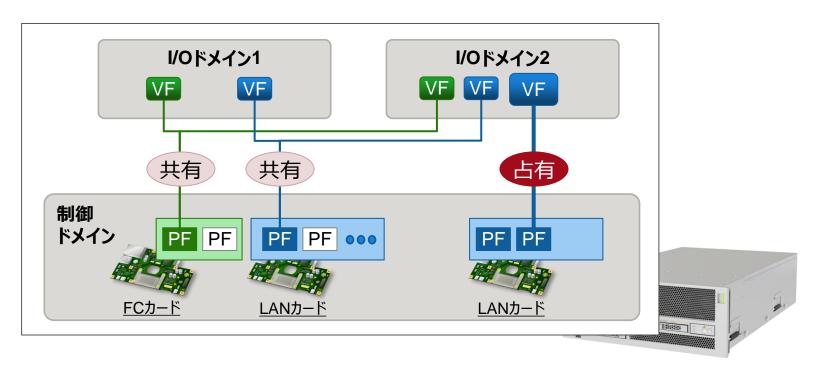


VF提供元ドメイン(制御ドメインまたはI/Oルートドメイン)のいずれかが停止した場合も、I/Oドメインを稼動させ続けることができます。

物理I/O資源の有効活用

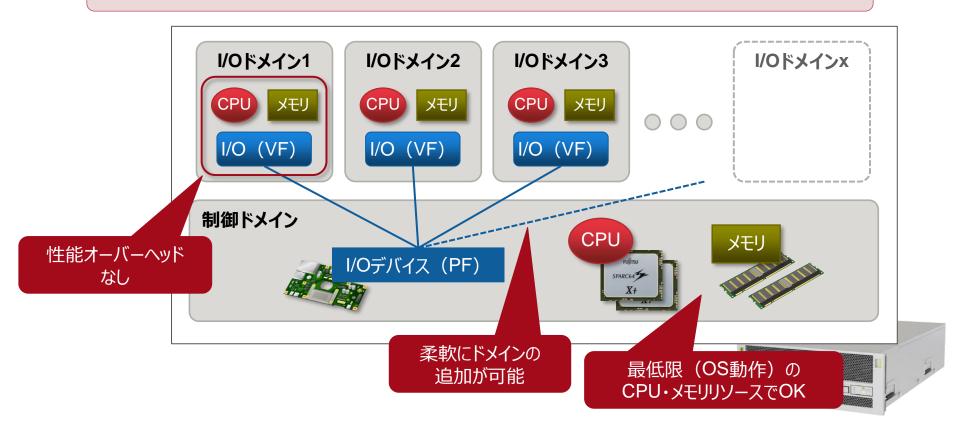


- 物理ポートの共有・占有のどちらも構成可能です。
 - 通常は、1つの物理ポートを複数のI/Oドメインで共有させます。
 - 仮想I/O (ゲストドメイン) 構築時と同様です。
 - 物理ポートの帯域を複数のI/Oドメイン(VF)で共有します。
 - 1つの物理ポートを1つのI/Oドメインで占有させることもできます。
 - 1つの物理ポートから1つだけVFを作成して、割り当てます。
 - 性能を重視するネットワークやドメインにおいて有効です。





SR-IOVは、多数の仮想環境を構築するのに適しています。



- CPU、メモリ、I/Oすべてにおいて、性能オーバーヘッドがない仮想環境を構築できます。
- SR-IOVの構成を柔軟に変更し、簡単に新規のI/Oドメインを追加できます。
- I/Oリソースの仮想化において、制御ドメインのCPU・メモリリソースを消費しないため、大規模の仮想化 基盤として有効です。



3. 注意事項

- 動作要件·推奨環境
- VF割り当て可能なドメインの数
- ネットワーク設計・設定時の注意事項
- FCストレージ/FCスイッチの接続
- 構成変更時の注意事項
- VFの追加作成とWWN設定変更時の注意事項
- VFの割り当て解除と削除の注意事項

動作要件·推奨環境



SPARC M10でSR-IOVを使用する場合は、以下の環境・要件を事前に確認します。

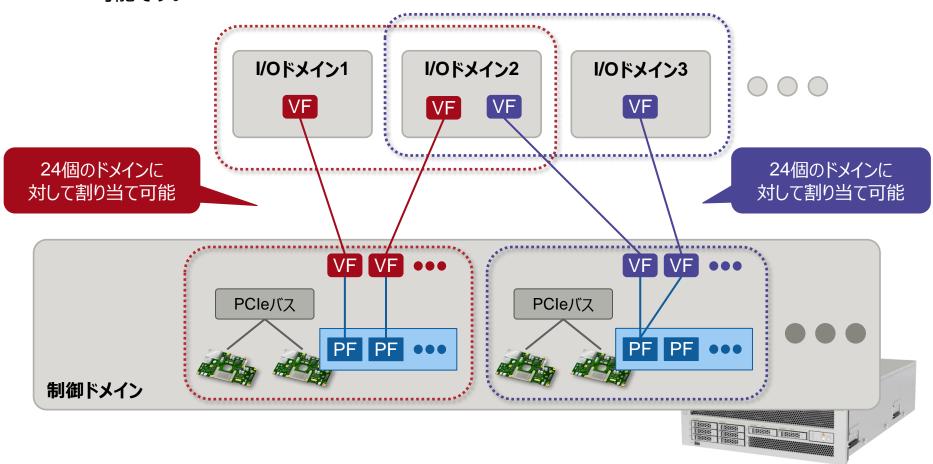
項目	動作要件・推奨環境
サーバのファームウェア	XCP2210以降
使用可能なI/Oデバイス	■ SPARC M10 内蔵LANポート ■ PCIカード • Ethernet(LAN)カード • InfiniBandカード • ファイバーチャネルカード 詳細は、以下のマニュアルをご参照ください。 • 『SPARC M10 システム PCI カード搭載ガイド』 「付録B SR-IOV対応カード」 http://www.fujitsu.com/jp/sparc/lineup/m10-1/documents/
OS ・ 制御ドメイン ・ I/Oルートドメイン ・ I/Oドメイン	Oracle Solaris 11.3 以降を推奨 (※) 最新のSRUを適用しておくことを推奨します。
Oracle VM Server for SPARC ・制御ドメイン	Oracle VM Server for SPARC 3.3 以降を推奨 (※)

[※] Oracle Solaris 11.2 以前または Oracle VM Server for SPARC 3.2 以前の環境の場合、一部使用できない機能があります。詳細は、Oracle社のマニュアルをご参照ください。

VF割り当て可能なドメインの数



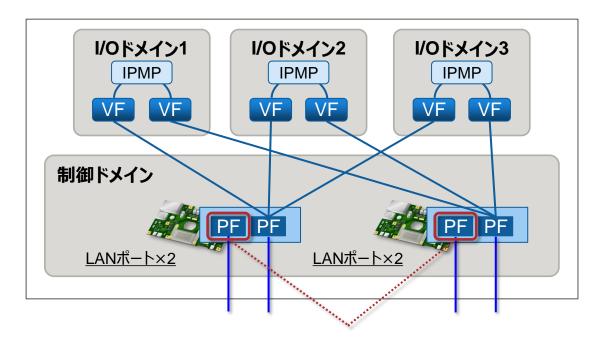
- 同一のPCIeバスからVFを割り当てられるドメインの数には、制限があります。
 - 制限の数は、サーバのモデルにより異なります。
 - SPARC M10の場合、PCIeバスあたり24個のドメインに対して、VFの割り当てが可能です。



ネットワーク設計・設定時の注意事項



■ VF提供元のドメイン(制御ドメイン/I/Oルートドメイン)が使用する LANポートとI/Oドメインが使用するLANポートは物理的に分けることを 推奨します。



制御ドメインが使用するLANポート(VF作成なし)

FCストレージ/FCスイッチの接続



■ SR-IOVを使用する場合は、基本的にI/OドメインをSAN Boot構成 (FC接続のストレージから起動)にします。 そのため、以下のハードウェアおよび設定が必要です。

■ FC接続のストレージ装置

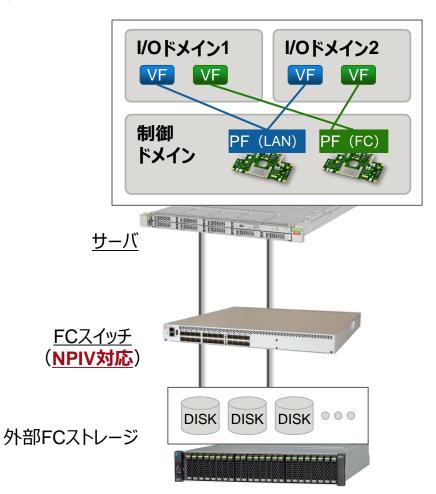
I/Oドメインのシステムディスクおよび データディスクとして使用します。

■ NPIV (※) 対応のFCスイッチ

- サーバと外部FCストレージをNPIV対応の FCスイッチで接続する必要があります。
 - 使用するFCスイッチが対応しているか、 事前に必ずご確認ください。
 - (※) NPIV: N-Port ID Virtualization

■ FCスイッチのゾーニング設定

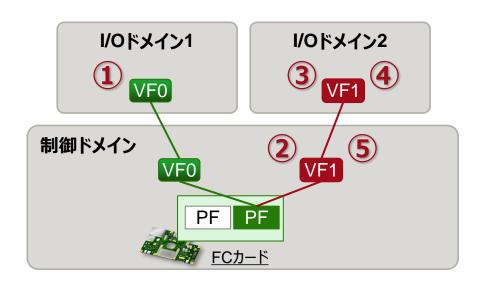
- 「WWN ゾーニング」を使用する ことを推奨します。
 - WWNゾーニングについては、 FCスイッチ・ストレージ装置の マニュアルをご参照ください。



構成変更時の注意事項



- SR-IOVの構成は随時変更できます。
 - ① VFの設定変更 (World Wide Nameの設定を変更する際など)
 - ② VFの追加作成 (I/Oドメインを追加作成する際など)
 - ③ VFの追加割り当て
 - 4 VFの割り当て解除(I/Oドメインの削除やPCIカードを交換する際など)
 - 5 VFの削除



- VF提供元のドメイン (制御ドメインなど) の再起動が必要な場合があります。
 - →詳細は、「VFの追加作成とWWN設定変更時の注意事項」参照
- その他の作業が必要な場合もあります。
 - →詳細は、「VFの割り当て解除と削除の注意事項」参照

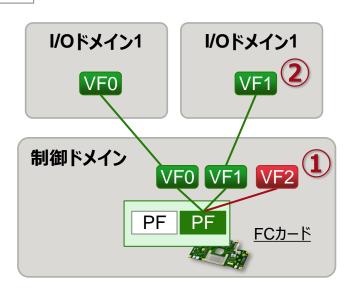
VFの追加作成とWWN設定変更時の注意事項



- VF提供元の再起動 (※1) が必要な構成変更
 - ① VFの追加作成

I/Oドメインを停止している場合、再起動は不要です。

- ※ 同じPFから作成したVFを割り当てているすべてのI/Oドメインの停止 (右図の場合は、I/Oドメイン1とI/Oドメイン2)
- ② VFのWWN (World Wide Name)の設定変更(※2)
- 再起動を必要としない構成変更(上記以外)
 - 作成済みのVFをI/Oドメインへ割り当て
 - VFをI/Oドメインから割り当て解除 (※3)
 - VFを削除(※3)
 - VFの設定(WWN以外)



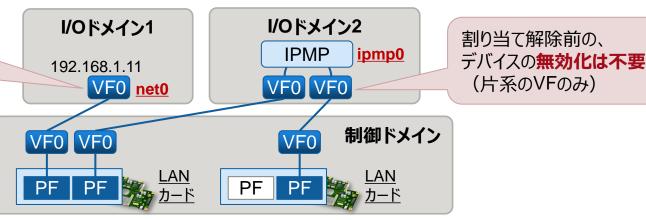
- (※1) VF提供元を遅延再構成状態にし、VFの追加作成またはWWNの設定変更をしたあとに再起動します。
- (※2) WWNを設定するVFが、I/Oドメインに割り当てられていない必要があります。
- (※3) 詳細は、「VFの割り当て解除と削除の注意事項」をご参照ください。

VFの割り当て解除と削除の注意事項



- VFの割り当て解除 ※LANポートの場合
 - 事前に、ipadmコマンドによりデバイスを無効にしてください。
 - デバイスを有効にしていない場合、またはマルチパス構成の場合は不要です。

割り当て解除前に、 デバイスの無効化が必要 # ipadm delete-ip net0



■ VFの削除

■ I/Oドメインに割り当てられていないVFを削除できます。

同一のPFから最後に作成したVFのみ、削除できます。

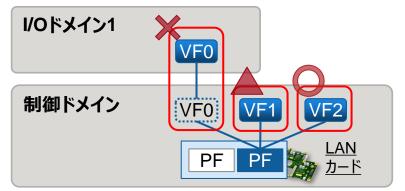
例: 「VF0 → VF1 → VF2 | の順にVFを作成

:削除可能

:削除不可 (VF2の削除後に削除可能)

:削除不可

(VFOの割り当てを解除し、 VF1, VF2の削除後に削除可能)



(片系のVFのみ)



4. 構築手順

- 本書で構築する仮想化構成
 - ディスク構成
 - ネットワーク構成
- ■構築手順

※『SR-IOVドメイン環境構築手順書』とあわせてご確認ください。

本書で構築する仮想化構成

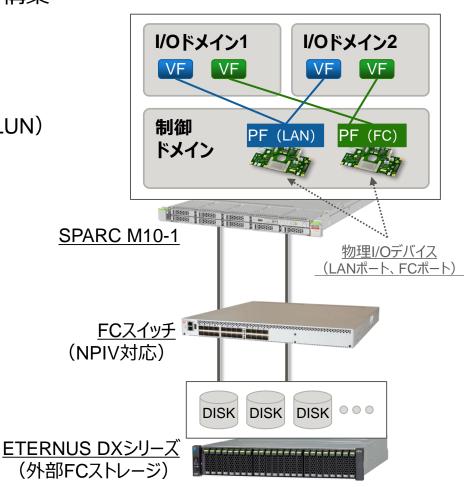


■ 構築概要

- SPARC M10-1において、I/Oドメインを2つ構築
 - 各I/Oドメインには、制御ドメインの LAN / FCポートより作成したVFを割り当て
- 各I/OドメインはSAN Boot構成
 - 外部ストレージ(ETERNUS)のディスク(LUN) を使用

■ 構築のポイント

- SR-IOVの基本的な構成
 - 制御ドメインにすべての物理I/Oデバイス (PCIeバス、PCIカード)を割り当て
 - I/Oドメインは制御ドメインから提供された SR-IOVによる仮想ポート(VF)を使用
- I/Oドメインの追加構築可能
 - VFを追加作成することで、物理I/Oデバイスの構成を変更せずに、I/Oドメインの追加が可能

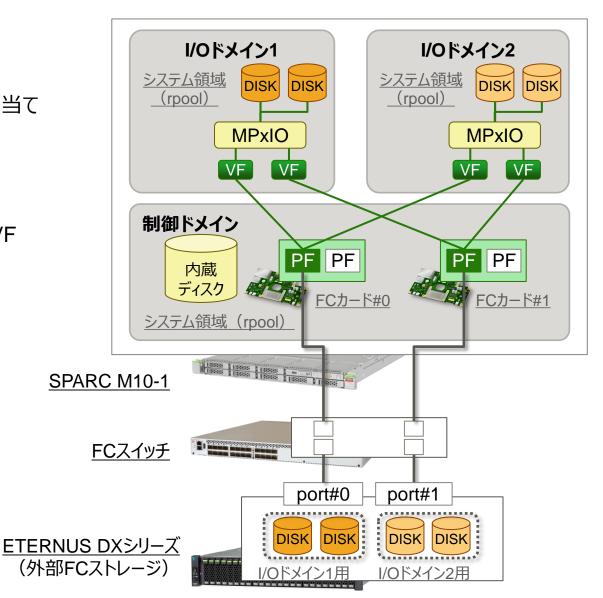


(※) SR-IOVの構成(PF, VF)は、簡略化して記述しています。 詳細は、次ページ以降をご確認ください。

ディスク構成



- ディスク/FCポートの構成
 - FCのディスクパスを冗長化
 - FCカード×2の各ポートからVFを作成し、I/Oドメインに割り当て
 - I/Oドメインでマルチパス設定 (Solaris標準のMPxIO)
 - FCスイッチでゾーニング設定
 - 外部FCストレージのディスクとVF のアクセスを制御 (各I/Oドメインは使用する ディスクのみ認識)

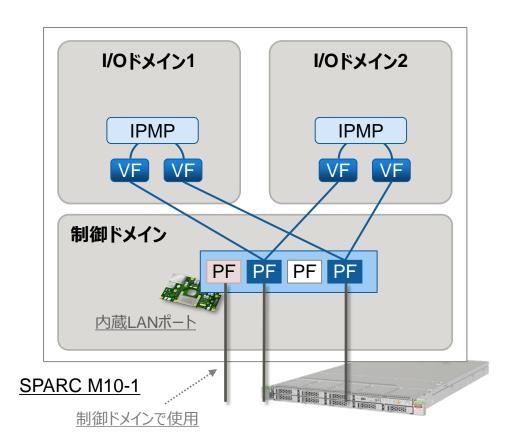


ネットワーク構成



- LANポートの構成
 - ネットワークを冗長化
 - 2つのLANポートからVFを作成し、 I/Oドメインに割り当て
 - I/Oドメインでマルチパス設定 (Solaris標準のIPMP)

- ※ 制御ドメインが使用しないLANポートから VFを作成し、I/Oドメインに割り当てます。
- ※ 制御ドメインのネットワーク(LAN)は、 別途冗長化することを推奨します。



構築手順





1. 制御ドメインの設定

2. 外部FCストレージへのアクセス設定

3. I/Oドメインの構築

4. I/Oドメインの設定

5. バックアップ・リストア

参考. VF提供元ドメインの二重化構成

参考. PCIカード交換時の設定復旧

制御ドメイン設定の概要



1

・サーバ・OSの環境確認

XCPファームウェア、OSとSRU、Oracle VMなどの版数を確認します。

2

・仮想サービスの作成

• 仮想コンソール、仮想ディスクサービスを作成します。

(**3**)

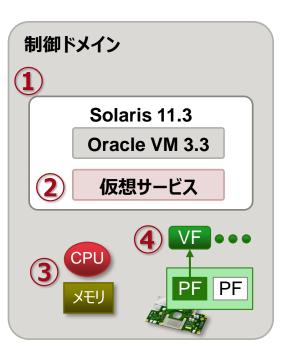
・リソースの設定

• 制御ドメインのCPU、メモリリソースを設定します。



VFの作成

• I/Oドメインに割り当てるVFを作成します。



ドメイン、リソース、仮想サービスの設定については、以下の資料もご参照ください。

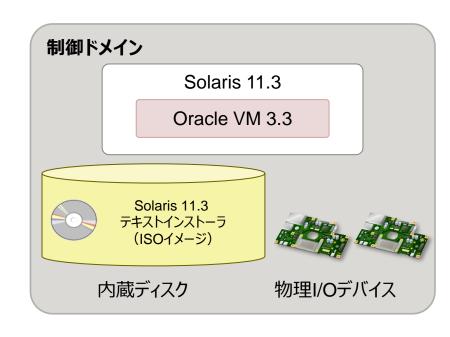
・『Oracle VM Server for SPARCを使ってみよう』

http://www.fujitsu.com/jp/sparc-technical/document/#ovm-use-11

① サーバ・OSの環境確認



- 制御ドメインの環境を確認します。
 - XCPファームウェア
 - 版数がXCP2210以降であることを確認します。
 - OS
 - OS版数およびSRU版数を確認します。
 - 本書は、Solaris 11.3を前提としております。
 - 最新のSRUを適用しておくことを推奨します。
 - Oracle VM Server for SPARC
 - 版数が3.3以降であることを確認します。
 - OSインストール媒体の格納 (※)
 - Solarisのテキストインストーラの イメージファイル (ISOファイル) を 用意します。
 - I/OドメインのOSインストールに使用します。

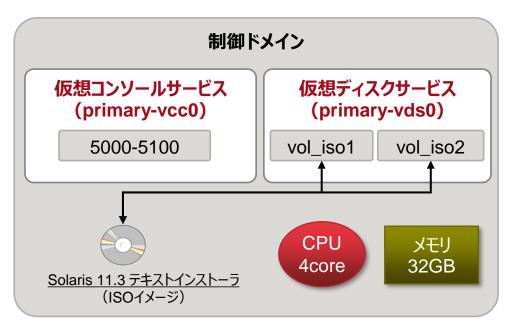


(※) インストールサーバを使用してI/OドメインのOSを インストールする場合は不要です。

② 仮想サービスの作成/③ リソースの設定



- I/Oドメインが使用する2つの仮想サービスを作成します。
 - 仮想コンソールサービス (vcc)
 - 仮想ディスクサービス (vds)
 - I/OドメインのOSインストールには、 Solaris 11.3のテキストインストーラのイメージファイル(ISO)を使用します。
 - イメージファイルを仮想ディスクサービスに登録します。
- 制御ドメインのCPUとメモリのリソースを設定します。
 - CPU
 - <u>コア単位の指定</u>を推奨します。
 - 割り当てるコアのID(物理位置)は 自動で選択されます。
 - メモリ
 - ・ 4GB以上の容量指定を推奨します。



④ VFの作成/割り当て手順の概要



制御ドメインの設定時にPF(物理ポート)からVF(仮想ポート)を作成します。 作成したVFはI/Oドメインの構築時に割り当てます。

1. PCIeバスのSR-IOVの有効化

SR-IOVが有効化されていることを確認します。

2. VFの作成

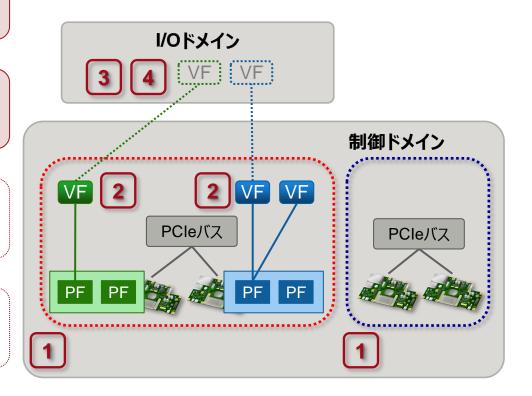
PFごとにVFを作成します。

3. VFの設定 (→「<u>3. I/Oドメインの構築</u>」)

VFのMACアドレスやWWNなどの 設定を行います。

4. VFの割り当て

(→「<u>3. I/Oドメインの構築</u>」) I/OドメインへVFを割り当てます。

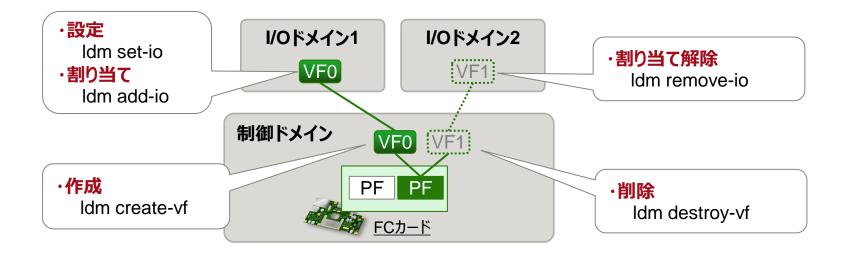


VFの作成/割り当てに使用するコマンド



■ VFの作成や割り当てには、以下のコマンドを使用します。

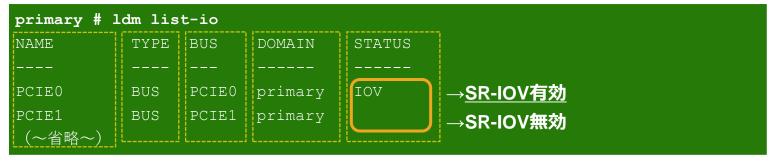
コマンド	解説
Idm create-vf	VFの作成
ldm set-io	VFの設定 (WWNやMACアドレスなど)
ldm add-io	VFをドメインへ割り当て
Idm remove-io	VFのドメインへの割り当てを解除
ldm destroy-vf	VFの削除



4-1 PCIeバスのSR-IOVの有効化



- PCIeバスごとに、SR-IOVの有効/無効を設定できます。
 - Idm list-ioコマンドで、SR-IOVの使用可否を確認します。
 - PCIeバスの[STATUS]欄が"IOV"と表示されている場合、 そのバスはSR-IOVが有効です。



デバイス名

PCIeバス 提供先 デバイス

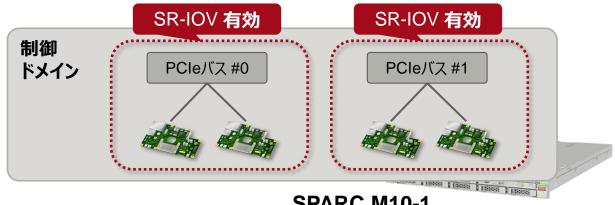
ステータス

の種類

デバイスの種類(TYPE)

•BUS : PCIeバス •PCIE : PCleカード

•PF : PF •VF : VF ■ デフォルトでは、すべてのPCIeバスにおいて、 SR-IOVが有効に設定されています。



《参考》 PCIeバスのSR-IOVの有効化手順



- PCIeバスごとにSR-IOVを有効にすることができます。
 - 1. SR-IOVを有効化するPCIeバスが割り当てられたドメインを、遅延再構成状態にします。

```
primary# ldm start-reconf primary
```

2. SR-IOVを有効にします。

【書式】Idm set-io iov=on <PCIeバス名>

```
primary# ldm set-io iov=on PCIE1
```

3. Oracle VMの構成情報をXSCF(サービスプロセッサ)に保存します。

```
primary# ldm remove-config config_initial
primary# ldm add-config config_initial
```

4. PCIeバスを所有するドメインを再起動します。

```
primary# shutdown -i6 -g0 -y
```

5. SR-IOVが有効になっていることを確認します。

```
primary# ldm list-io
 NAME
                                 TYPE
                                        BUS
                                                  DOMAIN
                                                            STATUS
 PCIE0
                                        PCIE0
                                                  primary
                                 BUS
                                                            IOV
                                                  primary
                                                            IOV
 PCTE1
                                 BUS
                                        PCTE1
                                                                   →SR-IOV有効化
 (~省略~)
```

4-2 VFの作成



■ 物理ポートのデバイス名(PF名)を指定して、VFを作成します。

【書式】Idm create-vf [-n 作成数 | max] <PF名>

primary# ldm create-vf -n 2 /SYS/MBU/NET0/IOVNET.PF1

Created new vf: /SYS/MBU/NET0/IOVNET.PF1.VF0
Created new vf: /SYS/MBU/NET0/IOVNET.PF1.VF1

- 上記コマンドでは、SPARC M10-1の内蔵LANポート#1から、VFを作成しています。
 - デバイス名(PF名)の確認方法は、次ページ以降をご参照ください。
- -n オプションで作成するVFの数を指定します。
 - "-n max"と指定した場合、最大数のVFを作成します。
 - 作成可能なVFの数は、使用するPCIeカードやポートの種類により異なります。 (詳細は、「《参考》VFの作成可能数 |参照)
 - -n オプションを指定しない場合、VFを1つ作成します。

VFの作成数

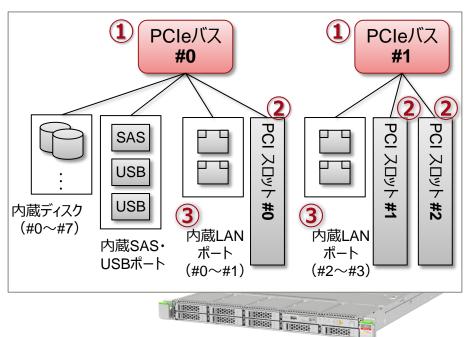
- 基本的に作成するVFの数は、物理ポート(PF)を共有するI/Oドメインの数を指定します。
- ただし、システム稼動後にドメインやネットワークの追加などで、さらに多くのVFを必要とする可能性がある場合は、**予備としてVFを多めに作成**しておくことを推奨します。
 (システム稼動後にVFを追加で作成(Idm create-vf)する場合は、制御ドメインまたはI/Oルートドメインの再起動が必要になります。)

《参考》物理デバイスの一覧表示



■ Idm list-io コマンドで、物理デバイスを確認できます。

SPARC M10-1 (1CPU / PCIeバス×2)



例:

primary# ldm list-io NAME	TYPE	BUS	DOMAIN	STATUS
PCIE0	BUS	PCIE0	primary	IOV
PCIE1	BUS	PCIE1	primary	IOV
/SYS/MBU/NETO	PCIE	PCIE0	primary	OCC
/SYS/PCIO	PCIE	PCIE0	primary	OCC
/SYS/MBU/SASHBA	PCIE	PCIE0	primary	OCC
/SYS/MBU/NET2	PCIE	PCIE1	primary	OCC
/SYS/PCI1	PCIE	PCIE1	primary	OCC
/SYS/PCI2	PCIE	PCIE1	primary	OCC
/SYS/MBU/NET0/IOVNET.PF0	PF	PCIE0	primary	
/SYS/MBU/NET0/IOVNET.PF	PF	PCIE0	primary	
/SYS/PCI0/IOVFC.PF0	PF	PCIE0	primary	
/SYS/PCI0/IOVFC.PF1	PF	PCIE0	primary	
/SYS/MBU/NET2/IOVNET.PF0	PF	PCIE1	primary	
/SYS/MBU/NET2/IOVNET.PF1	PF	PCIE1	primary	
/SYS/PCI1/IOVFC.PF0	PF	PCIE1	primary	
/SYS/PCI1/IOVFC.PF1	PF	PCIE1	primary	

デバイス名(NAME)

① PCIeバス ··· PCIEx

② PCIカード ··· /SYS/PCIx

③ 内蔵LANポート ・・・・ /SYS/MBU/... (SPARC M10-1)

••• /BB0/CMU/... (SPARC M10-4)

<u>デバイスの</u> 種類 :: PCIeバス 提供先 ステータス

デバイスの種類(TYPE)

・BUS : PCIeバス
・PCIE : PCIeカード

PCIE : PCIePF : PF

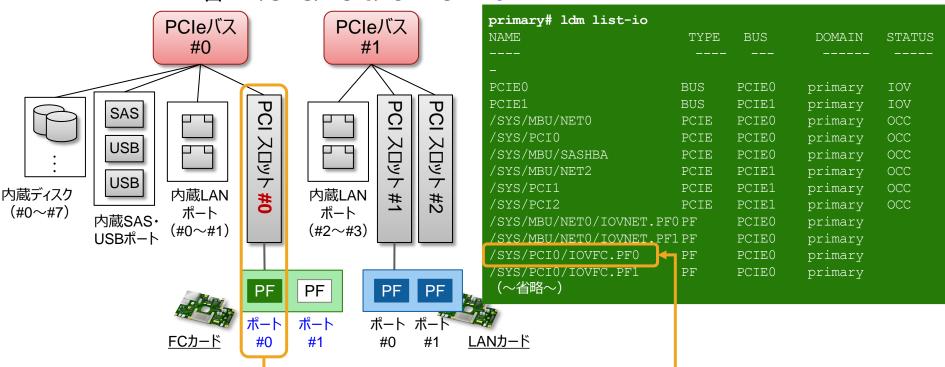
•VF : VF

デバイス名

《参考》物理ポートのデバイス名(PF名)の確認方法1/2



- PCIeカードの場合
 - Idm list-io コマンドで確認できます。
 - 「/SYS/PCIx/.... PFx」と表示されます。
 - ポート番号 (※) PCIスロット番号
 - 例: <u>SPARC M10-1: PCIスロット #0(FCカード搭載) ポート#0</u>
 - ⇒ PF名: /SYS/PCIO/IOVFC.PFO

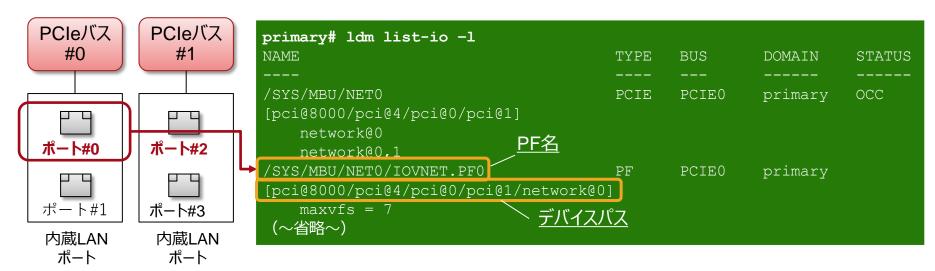


《参考》物理ポートのデバイス名(PF名)の確認方法2/2



- 内蔵LANポートの場合
 - デバイスパス (pci@...) をもとに確認します。
 - 例: SPARC M10-1: 内蔵LANポート#0の場合
 - 1. デバイスパスの確認
 - 以下のマニュアルからご確認ください。
 『SPARC M10 システム システム運用・管理ガイド』「付録A A.1 SPARC M10-1のデバイスパス」
 http://www.fujitsu.com/jp/products/computing/servers/unix/sparc/lineup/m10-1/documents/

 内蔵LANポート#0 のデバイスパス : /pci@8000/pci@4/pci@0/pci@1/network@0
 - 2. PF名の確認
 - Idm list-io -I コマンドで、デバイスパスに対応したPFを確認します。



※ SPARC M10-1の内蔵LANポートのみ表記

《参考》VFの作成可能数



- 物理ポート (PF) によって、作成可能なVFの数が異なります。
 - 例1: SPARC M10-1の内蔵LANポート: 最大7個のVFを作成可能

最大 7個

■ 例2:FCカードのFCポート: 最大16個のVFを作成可能

最大 16個

VFの作成可能な数は、SPARC M10のマニュアルでも確認できます。

・『SPARC M10 システム PCI カード搭載ガイド』

http://www.fujitsu.com/jp/sparc/lineup/m10-1/documents/

構築手順



1. 制御ドメインの設定



2. 外部FCストレージへのアクセス設定

3. I/Oドメインの構築

4. I/Oドメインの設定

5. バックアップ・リストア

参考. VF提供元ドメインの二重化構成

参考. PCIカード交換時の設定復旧

外部FCストレージへのアクセス設定の概要

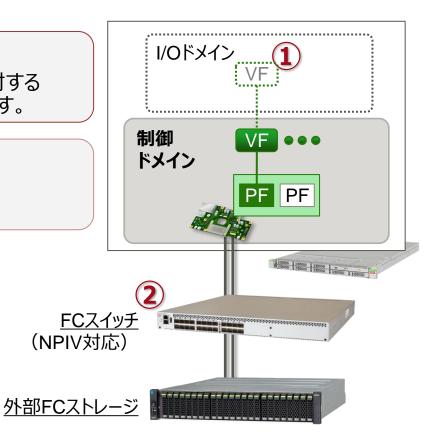


1

・VFに対するWWNの決定

• I/Oドメインに割り当てる、FCポートのVFに対する WWN (World Wide Name) を決定します。

- ・FCスイッチの設定
 - WWNゾーニングを設定します。

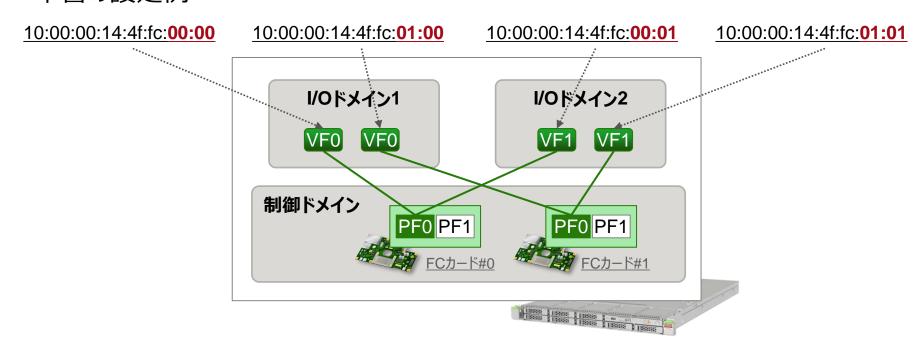


① VFに対するWWNの決定



FCポートのVFに、WWNを設定できます。 設定するWWNの値は事前に決定しておくことを推奨します。

■ 本書の設定例



VFのWWNは、以下の範囲の値を設定します。

XX:XX:00:14:4F:FC:00:00 ~ XX:XX:00:14:4F:FF:FF

《参考》WWNのデフォルト設定の注意事項

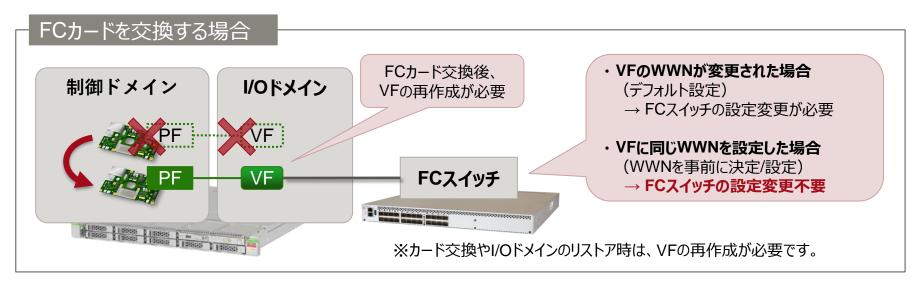


■ 運用上の注意事項

■ VFのWWNを設定しない(デフォルト)場合、WWNは自動で決定されます。

VFのWWNの設定	設定可能なWWNの範囲
自動(デフォルト)	XX:XX:00:14:4F:F8:00:00 ~ XX:XX:00:14:4F:FB:FF:FF
手動	XX:XX:00:14:4F:FC:00:00 ~ XX:XX:00:14:4F:FF:FF

- デフォルトでは、VFを作成する度に、そのVFのWWNが新しく自動で設定されます。
- FCカードの交換などにより、VFを再作成してWWNが変更された場合は、 FCスイッチやストレージの再設定が必要になることがあります。

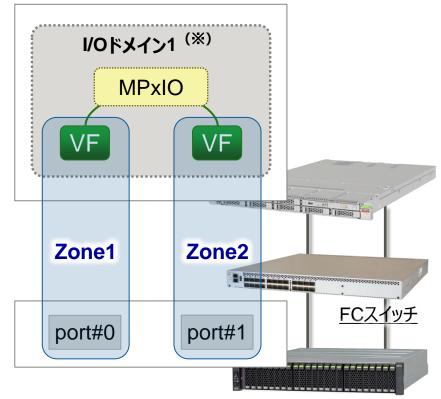


② FCスイッチの設定



I/Oドメインと外部ストレージのディスクのアクセス制御設定を行います。

- 設定のポイント
 - <u>I/OドメインのVFのWWN</u> (「① <u>VFに対するWWNの決定</u>」で指定) と<u>ストレージのFCポートのWWN</u>を 指定して、WWNゾーニングを 構成します。
 - FCのパスを確実に冗長化させるため、 I/Oドメイン内の複数のVFから、 別々のストレージのFCポートに アクセスするように設定してください。
- (※) ほかのI/OドメインのVFに対しても、同様にWWNゾーニングを設定してください。(本ページの図では、ほかのI/Oドメインの表記を省略しています。)



外部ストレージ

- ・WWNゾーニングは、I/Oドメインを起動させる前に設定する必要があります。
- ・設定方法は、FCスイッチのマニュアルなどをご参照ください。

構築手順



- 1. 制御ドメインの設定
- 2. 外部FCストレージへのアクセス設定



- 3. I/Oドメインの構築
- 4. I/Oドメインの設定
- 5. バックアップ・リストア

参考. VF提供元ドメインの二重化構成

参考. PCIカード交換時の設定復旧

I/Oドメイン構築の概要



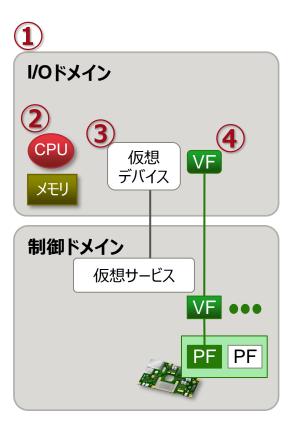
・ドメインの作成

ドメインの定義を作成します。

・リソースの設定

• CPU、メモリリソースを設定します。

- ・仮想デバイスの割り当て
 - 仮想コンソールを割り当てます。
 - 仮想ディスク(OSインストール用)を割り当てます。
- ・VFの設定/割り当て/確認
 - VFにWWNを設定し、I/Oドメインに割り当てます。



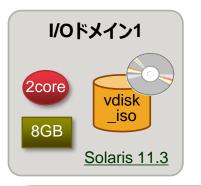
4

3

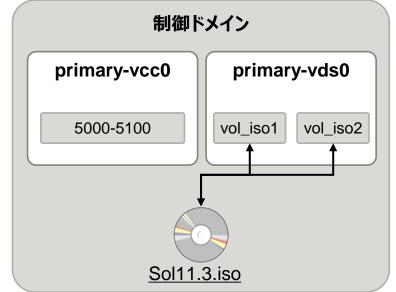
① ドメイン作成/② リソース設定/③ 仮想デバイス割り当て



- ドメインを作成します。
 - Idm add-domain コマンドによって、 ドメインの定義を作成します。
- ドメインにリソースを設定します。
 - I/OドメインのCPU、メモリリソースを 指定します。
- ドメインに仮想デバイスを割り当てます。
 - 仮想ディスクと仮想コンソールポートを I/Oドメインへ割り当てます。
 - 仮想ディスク
 - Solaris 11.3テキストインストーラのイメージを 仮想ディスクとして割り当て
 - 仮想コンソールポート
 - I/Oドメインのコンソール接続に使用する ポート番号を指定







4-1 VFの設定



■ 作成済みのVFを確認します。

【書式】Idm list-io

primary# ldm list-io			
NAME	TYPE	BUS	DOMAIN STATUS
PCIE0	BUS	PCIE0	primary IOV
(~省略~)			
/SYS/MBU/NET0/IOVNET.PF1.VF0	VF	PCIE1	
/SYS/MBU/NET0/IOVNET.PF1.VF1	VF	PCIE1	LANポートのVF
/SYS/MBU/NET2/IOVNET.PF1.VF0	VF	PCIE1	
/SYS/MBU/NET2/IOVNET.PF1.VF1	VF	PCIE1	
/SYS/PCI0/IOVFC.PF0.VF0	VF	PCIE0	
/SYS/PCI0/IOVFC.PF0.VF1	VF	PCIE0	FCポートのVF
/SYS/PCI1/IOVFC.PF0.VF0	VF	PCIE1	
/SYS/PCI1/IOVFC.PF0.VF1	VF	PCIE1	

デバイスの種類(TYPE)

·BUS : PCIeバス

・PCIE : PCIeカード

·PF : PF

•VF : VF

デバイス名

デバイスの種類

PCIeバス

提供先

ステータス

- VFの設定をします。
 - FCポートのVFに、「<u>VFに対するWWNの決定</u>」で決定したWWNを設定します。
 - VF提供元のドメイン (制御ドメインまたはI/Oルートドメイン) を遅延再構成にします。

primary# ldm start-reconf primary

※ VFにWWNを設定する場合のみ、遅延再構成にする必要があります。

• WWNを設定します。

【書式】 Idm set-io [設定するプロパティ=設定値] <VF名>

primary# ldm set-io port-wwn=10:00:00:14:4f:fc:01:00 node-wwn=20:00:00:14:4f:fc:01:00 /SYS/PCI1/IOVFC.PF0.VF0

※ LANポートのVFに、MACアドレスやVLANを設定する場合も同様です。

④-2 VFの割り当て/割り当ての確認



- VFをI/Oドメインへ割り当てます。
 - 作成したVF (FC, LAN) をI/Oドメインへ割り当てます。

【書式】 Idm add-io <VF名> <ドメイン名>

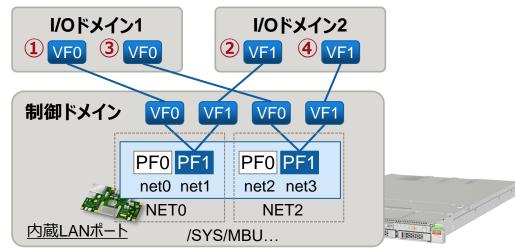
Primary# 1dm add-io /SYS/MBU/NET0/IOVNET.PF1.VF0 1dom1 ←VF0をldom1へ割り当て

VFがI/Oドメインに割り当てられていることを確認します。
【書式】 ldm list-io

VFがI/Oドメインへ
割り当てられています

```
primary# ldm list-io
 (~省略~)
/SYS/MBU/NET0/IOVNET.PF1.VF0
                                                            ldom1
                                                   PCIE1
                                            VF
/SYS/MBU/NET0/IOVNET.PF1.VF1
                                                            ldom2
                                                   PCIE1
                                            VF
/SYS/MBU/NET2/IOVNET.PF1.VF0
                                                            ldom1
                                            VF
                                                   PCIE1
/SYS/MBU/NET2/IOVNET.PF1.VF1
                                                            ldom2
                                            VF
                                                   PCTE1
 (~省略~)
```

※ PFを指定して、割り当て状態を確認することもできます。(【書式】 Idm list-io <PF名>)



《参考》仮想的なI/Oデバイスの一覧表示



■ VFと仮想デバイス(vdisk, vnet)の提供元と提供先ドメインを表示

【書式】Idm list-dependencies <-r> <-l> <ドメイン名>

primary# ld	m list-dependen	cies -r -l	
DOMAIN	DEPENDENT	TYPE	DEVICE
primary	ldom1	VDISK	primary-vds0/vdisk_iso
		IOV	/SYS/MBU/NET0/IOVNET.PF1.VF0
(~省略~)		IOV	/SYS/PCI0/IOVFC.PF0.VF0
		IOV	/SYS/MBU/NET2/IOVNET.PF1.VF0
		IOV	/SYS/PCI1/IOVFC.PF0.VF0
	ldom2	VDISK	primary-vds0/vdisk_iso

提供元ドメイン 提供先ドメイン デバイスの種類 デバイス名

Idm list-dependenciesコマンドのオプション

• -r : 提供元ドメイン、提供先ドメインの順に表示

・-I : 詳細(デバイス名)を表示

・ドメイン名:指定したドメインに関する情報のみ表示

仮想的なデバイスのみ、表示させることができます。物理的なデバイス (PCIeバス、PCIeカード) の割り当て先を確認する場合は、 前項のコマンド "Idm list-io"を使用してください。

4-3 VFの設定の確認



■ VFのWWNやMACアドレスを確認します。

- PFごとに確認します。
- <PF名>を省略することで、 すべてのPF, VFを表示できます。

LANの場合

- ・MACアドレス
- VLAN ID
- •MTU

本書では、LANのVFにデバイス 設定は実施していません。 そのため、mac-addr (Oracle VMで 自動で設定) とmtu (OSのデフォルト 値) のみ、表示されます。

FCの場合

WWN (port-WWN / node-WWN)

【書式】Idm list-io -I <PF名>

```
primary# ldm list-io -1 /SYS/MBU/NET2/IOVNET.PF0
NAME
                                          TYPE
                                                BUS
                                                        DOMAIN
                                                                  STATUS
/SYS/MBU/NET2/IOVNET.PF0
                                          PF
                                                PCIE1
                                                       primary
 (~省略~)
[pci@8100/pci@4/pci@0/pci@0/network@0]
    maxvfs = 7
/SYS/MBU/NET2/IOVNET.PF0.VF0
                                          VF
                                                PCIE1
                                                        1dom1
[pci@8100/pci@4/pci@0/pci@0/network@0,80]
    Class properties [NETWORK]
        mac-addr = 00:14:4f:f9:12:d1
        vlan IDs = 2, 3, 4
        mtu = 1500
 (~省略~)
```

```
primary# ldm list-io -1 /SYS/PCI1/IOVFC.PF0
 (~省略~)
/SYS/PCI1/IOVFC.PF0
                                                PCIE1 primary
                                          PF
[pci@8100/pci@4/pci@0/pci@1/SUNW,qlc@0]
   maxvfs = 16
/SYS/PCI1/IOVFC.PF0.VF0
                                         VF
                                                PCIE1
                                                      ldom1
[pci@8100/pci@4/pci@0/pci@1/SUNW,qlc@0,2]
    Class properties [FIBRECHANNEL]
        port-wwn = 10:00:00:14:4f:fc:00:00
        node-wwn = 20:00:00:14:4f:fc:00:00
        bw-percent = 0
```

《参考》VFの割り当て解除/削除



■ I/OドメインへのVFの割り当てを解除

【書式】 Idm remove-io <VF名> <ドメイン名>

primary# ldm remove-io /SYS/MBU/NET2/IOVNET.PF0.VF0 ldom1

■ VFの削除

【書式】Idm destroy-vf <VF名>

primary# ldm destroy-vf /SYS/MBU/NET0/IOVNET.PF1.VF0

←VFを1つ削除

primary# ldm destroy-vf -n 2 /SYS/MBU/NET0/IOVNET.PF1

←VFを2つ削除(※)

primary# ldm destroy-vf -n max /SYS/MBU/NET0/IOVNET.PF1

←VFをすべて削除

※ 最後に作成したVFからn個、削除します。

- 事前に、「VFの割り当て解除(Idm remove-io)」を実施しておく必要があります。
- 削除できるVFは同一PF内のVFのうち、最後に作成されたVFのみです。

注意事項の詳細は、「VFの割り当て解除と削除の注意事項」をご参照ください。

《参考》ネットワークデバイスの一覧表示



■ 各ドメインのネットワークデバイス(物理/仮想)を一覧で表示

【書式】Idm list-netdev <-l> <-b> <-o デバイス名> <ドメイン名>

orimary NAME	CLASS	MEDIA	STATE	SPEED	OVER	LOC
net0	PHYS	ETHER	up	1000	 igb0	/SYS/MBU/NETO
net1	PHYS	ETHER	unknown	100	igb1	/SYS/MBU/NET0
ldoms-net1.vf1	VNIC	ETHER	up	0	net1	
ldoms-net1.vf0	VNIC	ETHER	up	0	net1	
DOMETH						
DOMAIN ldom1						
NAME	CLASS	MEDIA	STATE	SPEED	OVER	LOC
net0	PHYS	ETHER	ир	100	igbvf0	/SYS/MBU/NET0/IOVNET.PF1.VF0
net1	PHYS	ETHER	up	100	igbvf1	/SYS/MBU/NET2/IOVNET.PF1.VF0
ipmp0	IPMP	IPMP	ok		net1,net()

Idm list-netdevコマンドのオプション

・-I : 詳細(MTU、IPアドレス、MACアドレスなど)を表示

・-b: 仮想サービスやIPMPの設定情報を表示しない

・-o デバイス名 : 指定したデバイスのみ表示(ドメイン名の指定が必須)

・ドメイン名: 指定したドメインに関する情報のみ表示

I/Oドメインに割り当てたLANのVFのデバイスパスを表示できます。 また、-I オプションにより、I/Oドメイン上で割り当てたIPアドレス(IPMP含む)を確認できます。

構築手順



- 1. 制御ドメインの設定
- 2. 外部FCストレージへのアクセス設定
- 3. I/Oドメインの構築



- 4. I/Oドメインの設定
- 5. バックアップ・リストア

参考. VF提供元ドメインの二重化構成

参考. PCIカード交換時の復旧手順

I/Oドメイン設定の概要



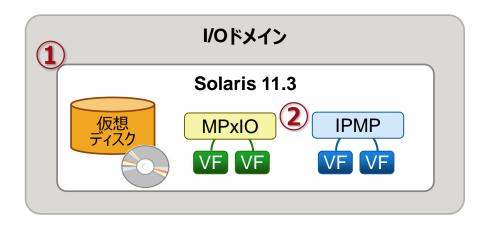
 $(\mathbf{1})$

・I/OドメインへOSインストール

仮想ディスクによってSolarisのローカルインストールを 実施します。

・マルチパスの設定(Solaris標準)

- MPxIOによって、FCのマルチパスを設定します。
- IPMPによって、ネットワークのマルチパスを設定します。



① // のドメインへのSインストール



ゲストドメインと同様の手順でOSをインストールします。

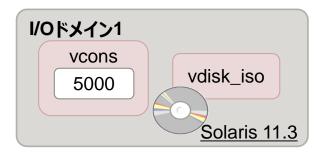
- ローカルブートによるOSのインストールを実施します。
 - 制御ドメインからI/Oドメインへコンソール接続します。

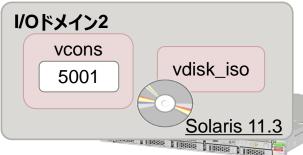
primary# telnet localhost 5000

- ※ I/Oドメインの仮想コンソールポートのポート番号を指定します。
- OBPからOSインストール用の仮想ディスクを指定して、インストールを実行します。

{0} ok boot vdisk iso

- ※ 仮想ディスクのエイリアス(本書・別冊の手順書では、vdisk iso)は、 仮想ディスク割り当て時に設定した値となります。 OSのインストール手順は、『Oracle Solaris 11を使ってみよう』も参考にしてください。
 - http://www.fujitsu.com/jp/sparc-technical/document/#solaris11







🧨 •OSインストール後は、物理サーバと同様に、SRUおよびESF(Enhanced Support Facility) などの必要なパッケージをインストールしてください。

②-1 FCのマルチパスの設定(MPxIO)



■ 物理サーバと同様の手順でマルチパスを設定します。

本書では、Solaris標準のMPxIOを使用します。

- ① 設定ファイル (scsi_vhci.conf) を編集します。
 - 外部ディスクの情報 (Vendor ID, Product ID) を追記します。
- MPxIOを有効化します。
 - stmsbootコマンドを実行します。

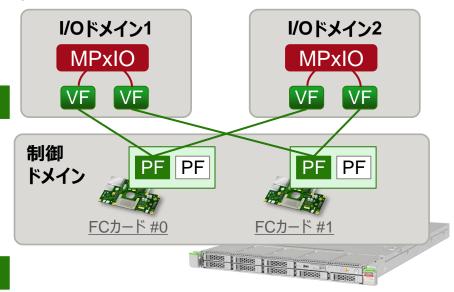
ldom1# stmsboot -D fp -e

※コマンド実行後、I/OドメインのOSが再起動します。

- ③ OBPパラメーターを設定します。
 - FCの冗長パスをboot-device パラメーターに設定します。

{0} ok setenv boot-device raid1 raid2

※「raid1」「raid2」はFCのデバイスパスのエイリアスを指しています。



MPxIOによるFCのマルチパスの設定は、以下の資料もご参照ください。

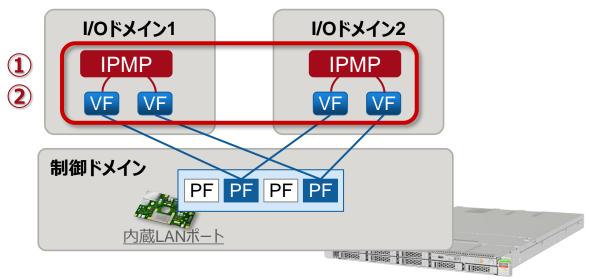
・『SAN Boot環境構築ガイド (Leadvilleドライバ編) for SPARC M10/SPARC Enterprise』

②-2 ネットワークのマルチパスの設定 (IPMP)



■ 物理サーバと同様の手順でマルチパスを設定します。

本書では、Solaris標準のIPMPを使用します。



2つの仮想NIC (VF) からIPMPグループを構築します。

【書式】ipadm create-ipmp <-i NICのデバイス名 (netX) > [-i NICのデバイス名 (netX) ···] <IPMPグループ名>

ldom1# ipadm create-ipmp -i net0 -i net1 ipmp0

② IPMPグループに対して、IPアドレスを設定します。

ldom1# ipadm create-addr -T static -a local=192.168.10.11/24 ipmp0/v4

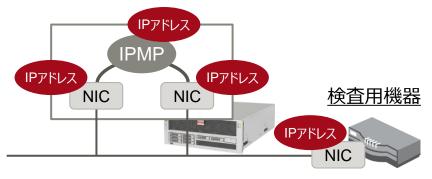
《参考》IPMPの動作



IPMPの監視(障害検出)と構成には、それぞれ2種類の設定があります。

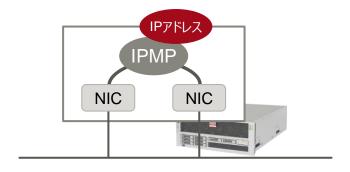
■ 障害検出

- プローブベース (検査信号ベース)
 - ルータなどの同一ホスト上の監視用機器 (ターゲット)を用意します。
 - 各NICからターゲットに対して、検査信号 (ICMP) を送信して、障害検出を行います。
 - 引き継ぎIPアドレスのほかに、 各NICにIPアドレスを設定します。



■ リンクベース

- 監視用機器の用意は不要です。
- NICのドライバがリンク状態を監視して、 障害の検出を行います。
- 設定するIPアドレスは引き継ぎIPアドレス のみです。



■構成

- 待機型構成
 - 通常時は1つのNICのみを通信に使用します。
 - ・ 残りのNICは、障害発生後に使用します。

■ 分散型構成

- すべてのNICを通信に使用します。
- 送信パケットは各NICに分散されます。

※ 本書では、リンクベース / 分散型構成 を使用しています。

構築手順



- 1. 制御ドメインの設定
- 2. 外部FCストレージへのアクセス設定
- 3. I/Oドメインの構築
- 4. I/Oドメインの設定



5. バックアップ・リストア

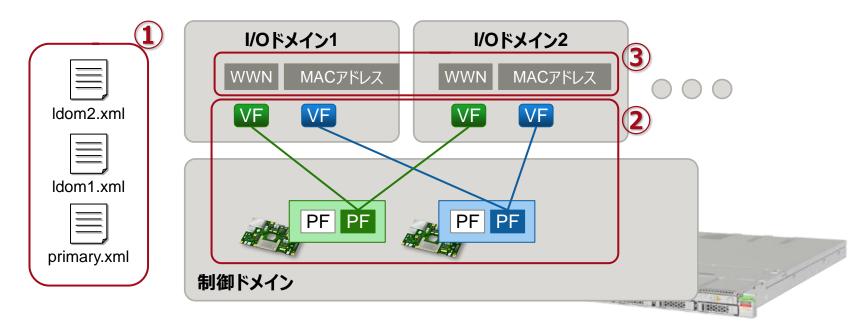
参考. VF提供元ドメインの二重化構成

参考. PCIカード交換時の設定復旧

バックアップの概要



- I/Oドメインのシステムボリュームとデータボリュームの バックアップ以外に、以下のバックアップが必要です。
 - Oracle VMの構成情報ファイル
 - PFとVFの構成
 - VFのデバイス設定



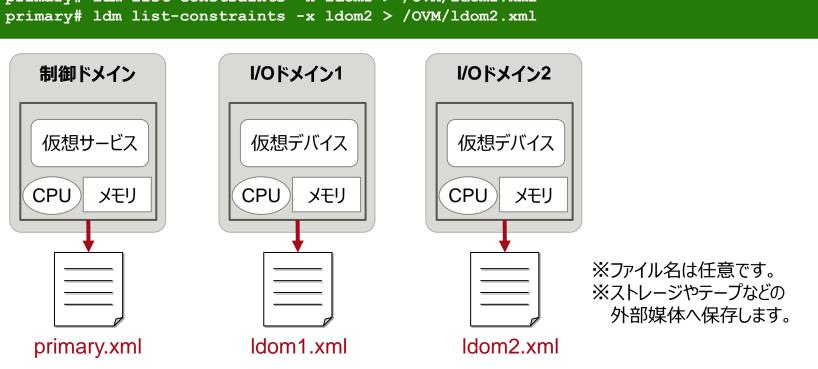
バックアップ (Oracle VMの構成情報)



- Oracle VMの構成情報は、xml形式で保存します。
 - ドメインごとに保存します。

【書式】 Idm list-constraints -x <保存するドメイン名> <保存先>

```
primary# ldm list-constraints -x primary > /OVM/primary.xml
primary# ldm list-constraints -x ldom1 > /OVM/ldom1.xml
primary# ldm list-constraints -x ldom2 > /OVM/ldom2.xml
```



XMLファイルには、SR-IOVに関する設定(作成したVFおよびVFの設定内容)は記録されません。 そのため、次ページ以降で解説する「<u>PFとVFの構成</u>」と「<u>VFのデバイス設定</u>」を記録しておきます。

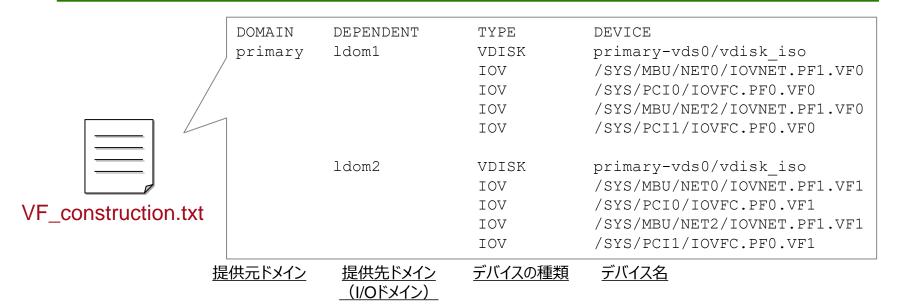
バックアップ(PFとVFの構成)



- VFの提供元ドメイン、提供先ドメイン、デバイス名を記録します。
 - テキスト形式で保存します。

【書式】Idm list-dependencies -r -l

primary# ldm list-dependencies -r -l > /OVM/VF_construction.txt



- ※ ファイル名は任意です。
- ※ ストレージやテープなどの外部媒体へ保存します。

リストア時はこの情報をもとに、手動でVFを作成する必要があります。

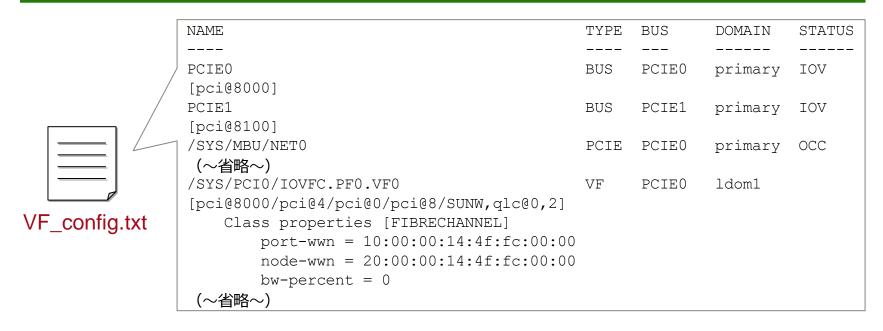
バックアップ(VFのデバイス設定)



- VFに設定したWWNなどのデバイスの設定を記録します。
 - テキスト形式で保存します。

【書式】Idm list-io -I <PF名>

primary# ldm list-io -1 > /OVM/VF_config.txt



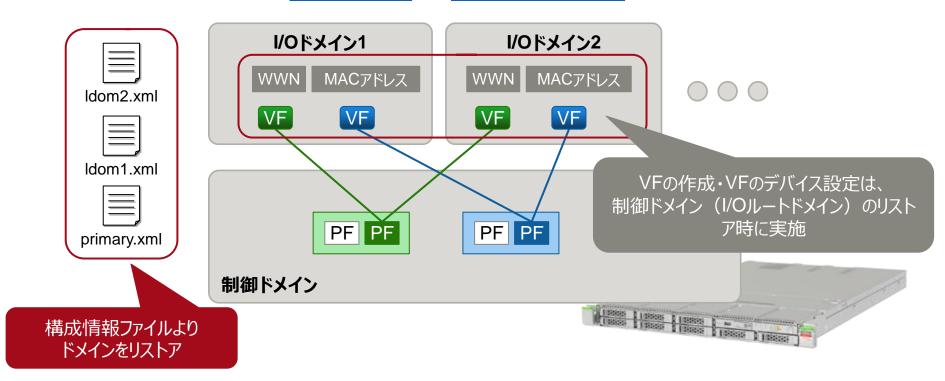
- ※ ファイル名は任意です。
- ※ ストレージやテープなどの外部媒体へ保存します。

リストア時はこの情報をもとに、手動でVFの設定を変更する必要があります。

ドメインのリストアの概要



- 全ドメイン共通
 - 構成情報ファイル(XMLファイル)をもとに、ドメインを復元させます。
- 制御ドメイン・I/Oルートドメイン(I/OドメインにVFを提供しているドメイン)
 - ドメインの復元に加え、「<u>VFの再作成</u>」と「<u>VFのデバイス設定</u>」を実施します。



各ドメインのリストア (Oracle VMの構成情報)



- I/Oルートドメイン・I/Oドメインのリストア
 - ■「<u>Oracle VMの構成情報(XMLファイル)</u>」によってドメインを復元します。

【書式】 ldm add-domain -i <構成情報ファイル>

```
primary# ldm add-domain -i /OVM/ldom1.xml
primary# ldm add-domain -i /OVM/ldom2.xml
```

VFを提供するI/Oルートドメインをリストアする場合は、 制御ドメインと同様に、VFの再作成・VFのデバイス設定を実施します。

- 制御ドメインのリストア
 - ■「 <u>Oracle VMの構成情報(XMLファイル)</u> 」によってドメインを復元します。

【書式】 Idm init-system -r -i <構成情報ファイル>

primary# ldm init-system -r -i /OVM/primary.xml



制御ドメインをリストアする場合、 事前にOracle VMの構成を 初期化させておきます。 (factory-default)

- VFを再作成します。
 - バックアップ時に取得した「PFとVFの構成」をもとに、VFを作成します。
- VFのデバイス設定をします。
 - バックアップ時に取得した「VFのデバイス設定」をもとに、 VFのMACアドレスやWWNなどを設定します。
 - ※ VFの作成および設定手順は、「VFの作成」「VFの設定」をご参照ください。
 - ※ 制御ドメインリストア後、I/Oルートドメイン・I/Oドメインのリストアが必要です。

構築手順



- 1. 制御ドメインの設定
- 2. 外部FCストレージへのアクセス設定
- 3. I/Oドメインの構築
- 4. I/Oドメインの設定
- 5. バックアップ・リストア



参考. VF提供元ドメインの二重化構成

参考. PCIカード交換時の設定復旧

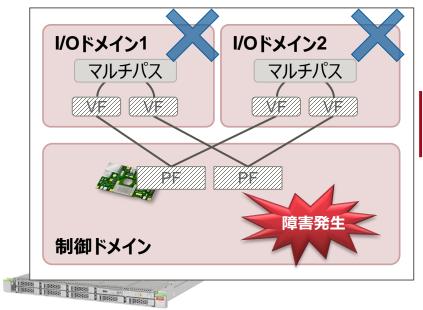
VF提供元ドメインの二重化



- VF提供元のドメイン障害に対応した、信頼性の高い仮想環境を構築できます。
 - 制御ドメインとI/Oルートドメインの両方から、VFを提供させます。

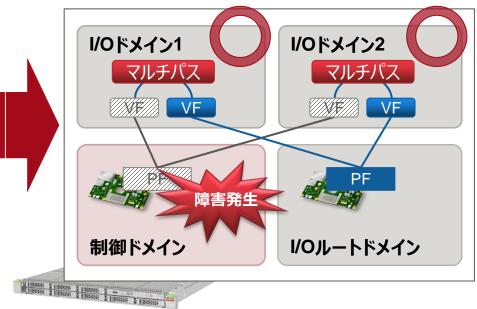
制御ドメイン障害時の動作の違い

SR-IOVの基本構成



制御ドメインからすべてのVFを提供✓ 制御ドメイン停止時は、すべてのI/Oドメインが停止

VF提供元ドメインの二重化構成



制御ドメインとI/Oルートドメインの両方からVFを提供✓ 制御ドメイン停止後も、I/Oドメインは継続して起動

本書におけるVF提供元ドメイン二重化の構成

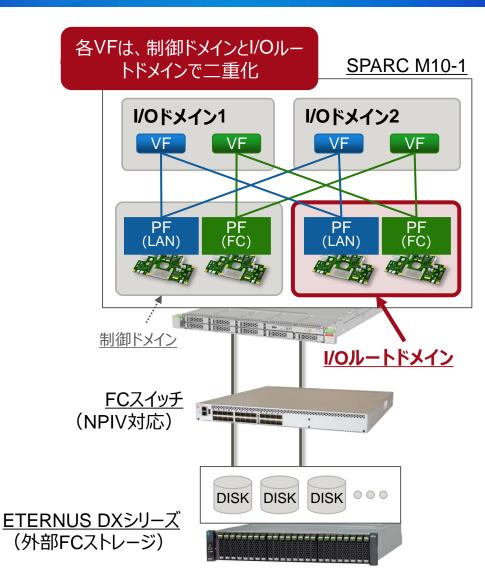


■ 構築のポイント

- I/O (VF) を制御ドメインとI/Oルート ドメインで二重化
 - 制御ドメインまたはI/Oルートドメインが 停止しても、I/Oドメインは継続して稼動可能
- I/Oドメインの追加構築可能
 - VFを追加作成することで、物理I/Oデバイスの 構成を変更せずに、I/Oドメインの追加が可能

■ 構築概要

- SPARC M10-1において、 <u>I/Oルートドメインを1つ</u>、 I/Oドメインを2つ構築
 - 各I/Oドメインには、制御ドメインと I/OルートドメインのそれぞれのLAN/ FCポートから作成したVFを割り当て
- I/Oルートドメインと各I/Oドメインは SAN Boot構成
 - 外部ストレージ(ETERNUS)のディスク (DISK)を使用

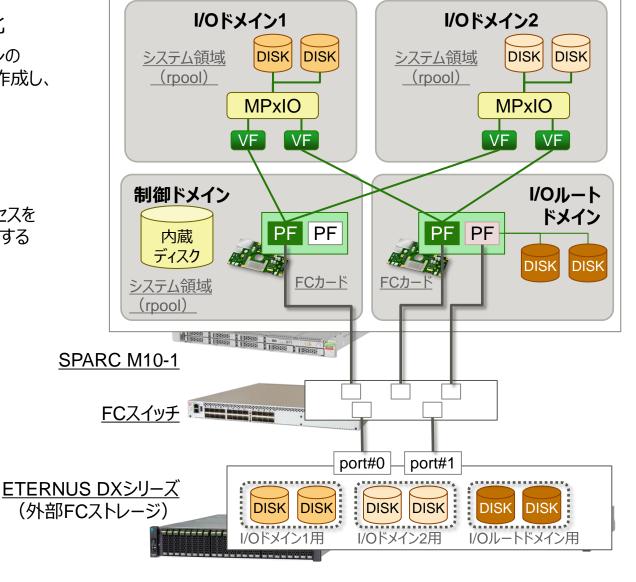


(※) SR-IOVの構成(PF, VF)は、簡略化して記述しています。 詳細は次ページ以降をご確認ください。

ディスク構成



- ディスク/FCポートの構成
 - I/Oドメインのディスクを冗長化
 - 制御ドメインとI/Oルートドメインの それぞれのFCポートからVFを作成し、 I/Oドメインに割り当て
 - I/Oドメインでマルチパス設定 (Solaris標準のMPxIO)
 - FCスイッチでゾーニング設定
 - ディスク (DISK) とVFのアクセスを 制御 (各I/Oドメインは、使用する ディスクのみ認識)

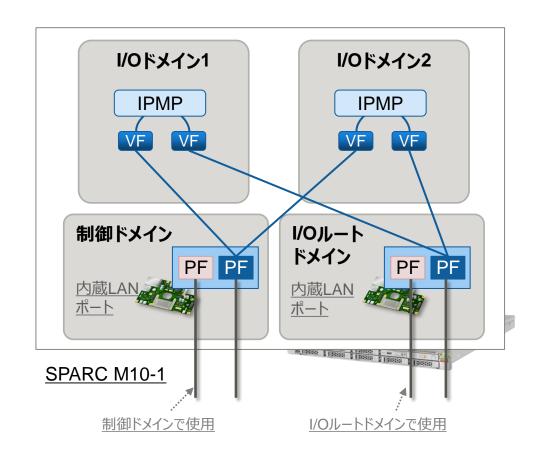


ネットワーク構成



■ LANポートの構成

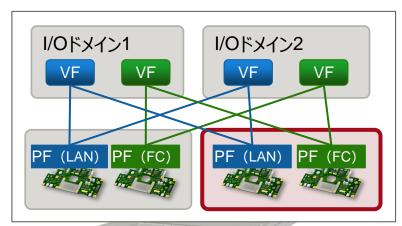
- ネットワークを冗長化
 - 制御ドメインとI/Oルートドメインの それぞれのLANポートからVFを作成し、 I/Oドメインに割り当て
 - I/Oドメインでマルチパス設定 (Solaris標準のIPMP)
 - ※ 制御ドメインおよびI/Oルートドメインが使用しないLANポートからVFを作成し、I/Oドメインに割り当てます。
 - ※ 制御ドメインのLAN(ネットワーク)も 別途冗長化することを推奨します。



VF提供元ドメイン二重化構成の構築手順概要



- 構築のながれは「4. 構築手順」とほぼ同じです。
 - 詳細は上記、および『SR-IOVドメイン 環境構築手順書』をご参照ください。
 - 1. 制御ドメインの設定
 - 2. 外部FCストレージへのアクセス設定
 - 3. I/Oルートドメインの作成
 - 4. I/Oドメインの構築
 - 5. I/Oドメインの依存関係設定
 - <u>6. I/Oドメインの</u>設定





I/Oルートドメインの作成



- 作成手順概要
 - ドメインの作成・リソースの設定
 - ゲストドメインやI/Oドメインと同様にドメインを作成して、 CPUやメモリなどのリソースを設定します。
 - 制御ドメインからPCIeバスを切り離します。

【書式】 Idm remove-io <PCIeバス名> <PCIeバスを切り離すドメイン>

primary# ldm remove-io PCIE1 primary

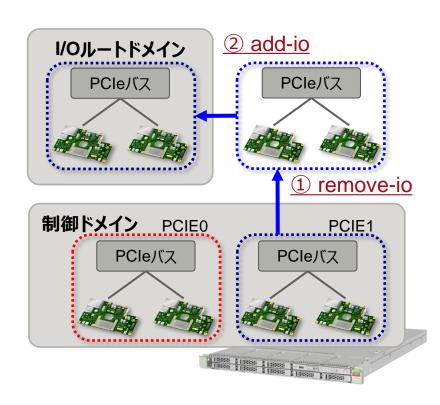
② ドメインにPCIeバスを割り当てます。

【書式】 ldm add-io <PCleバス名> <PCleバスを割り当てるドメイン>

primary# ldm add-io PCIE1 iodom

③ PCIeバスの割り当て状態を確認します。

primary# ldm list-io



I/Oドメインの依存関係設定



- VF提供元ドメインに障害(リブート、リセット、パニック)が発生した場合に、I/Oドメインを停止させないように依存関係を設定します。
 - ① I/Oドメインの依存先に、VF提供元ドメインを設定します。

【書式】 Idm set-domain master= < VF提供元のドメイン名 [, VF提供元のドメイン名][,...]> < I/Oドメイン名>

```
primary# ldm set-domain master=primary,iodom ldom1
primary# ldm set-domain master=primary,iodom ldom2
```

② VF提供元ドメイン障害時のI/Oドメインの動作を設定します。

【書式】Idm set-domain faulure-policy= <ポリシー名> <VF提供元ドメイン名>

```
primary# ldm set-domain failure-policy=ignore primary
primary# ldm set-domain failure-policy=ignore iodom
```

※ デフォルトで "ignore" が設定されています。

ポリシー (I/Oドメインの動作)

·ignore: I/Oドメインには影響を与えない

・panic : I/Oドメインにパニックを

発生させる

・reset : I/Oドメインを再起動させる

・stop : I/Oドメインを停止させる

I/Oドメイン1 ① master=primary,iodom







《参考》I/Oルートドメインの起動・停止



I/Oルートドメインの起動・停止方法は、ゲストドメインやI/Oドメインと同一です。

■ I/Oルートドメインの操作

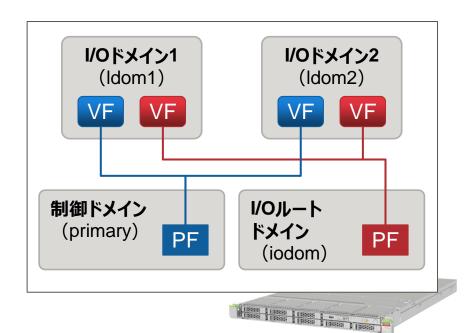
■ リソースのバインド・起動

primary# ldm bind-domain iodom primary# ldm start-domain iodom

■ ドメインの停止

primary# ldm stop-domain iodom

• I/Oルートドメイン上で、shutdownコマンドを 実行して停止することもできます。





/↑ I/Oルートドメインの停止について

・I/Oルートドメインを停止させる場合は、基本的に依存する I/Oドメイン(右図のIdom1, Idom2)を事前に停止させておきます。

構築手順



- 1. 制御ドメインの設定
- 2. 外部FCストレージへのアクセス設定
- 3. I/Oドメインの構築
- 4. I/Oドメインの設定
- 5. バックアップ・リストア

参考. VF提供元ドメインの二重化構成



参考. PCIカード交換時の設定復旧

PCIカード交換時の復旧手順の概要



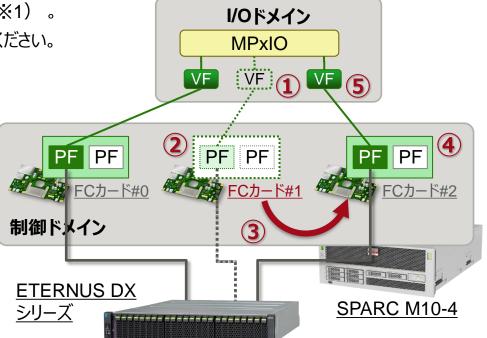
PCIカード交換時は、VFの再作成およびデバイス再設定などが必要です。

- 復旧手順(FCカード#1)
 - 制御ドメインおよびI/Oドメインの停止は不要です(※1)。
 - 各手順の詳細は、「4. 構築手順」の各章をご参照ください。

事前. <u>PFとVFの構成</u>採取 VFのデバイス設定採取

- ① VFの割り当て解除
- ② VFの削除
- ③ PCIカード交換 (※1)
- 4 VFの作成/デバイス設定 (※2)
- S VFの割り当て
- ※1 SPARC M10-1の場合、カード交換時にシステムの停止が必要です。 交換方法の詳細は、ご使用のサーバのサービスマニュアルをご参照ください。 http://www.fujitsu.com/jp/sparc/lineup/m10-1/documents/
- ※2 FCカードを交換する場合は、VFを作成(Idm create-vf)する際に同時にWWNを指定します。 (VF作成後にWWNを設定する場合、制御ドメインの再起動が必要となります。)

primary# ldm create-vf port-wwn=10:00:00:14:4f:fc:01:00
node-wwn=20:00:00:14:4f:fc:01:00 SYS/PCI1/IOVFC.PF0





5.参考情報

参考資料



■ Oracle VM Server for SPARCの技術資料

http://www.fujitsu.com/jp/sparc-technical/document/#ovm-use-11

- 『Oracle VM Server for SPARCを使ってみよう』
- 『Oracle VM Server for SPARC Idmコマンドクイックリファレンス』
- SPARC M10-1 サーバマニュアル

http://www.fujitsu.com/jp/products/computing/servers/unix/sparc/lineup/m10-1/documents/

- 『SPARC M10 システム 早わかりガイド』
- 『SPARC M10 システム ドメイン構築ガイド』
- 『SPARC M10 システム システム運用・管理ガイド』
- Oracle VM Server for SPARCの最新情報およびマニュアル (Oracle社)

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

技術情報 Technical Park



本製品のお問い合わせ

導入をご検討中のお客様

著作権とリンクについて

関連製品・サービス

• 総合索引 ・商標について

SPARC/Solaris

■ SPARC/Solarisの構築に役立つドキュメントが満載

技術情報 Technical Park

- Oracle VM Server for SPARC
- バ SPARC M10の技術情報を集めた、ポータルサイトです Solarisベースの仮想 Oracle Solarisソーン
- 最新ファイルシステム:

ZFS (Zettabyte File System)

ドキュメント

富士通のエンジニアが実際に検証し、ノウハウをまとめました。

Solaris 8/9環境をそのままSolaris 10へ

Oracle Solaris Legacy Containers

新 活用提案 ハードウェア Oracle Solaris 11 Oracle Solaris 10 ミドルウェア その他

今すぐクリック!!

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

など

改版履歴



版数	更新日時	更新内容
第1.0版	2016年6月	新規作成

使用条件·商標



■ 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下さい。

■ 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。



