【Solaris 活用ナビ ~Practical Tips for SPARC】

第 5 回 : SPARC Servers / Solaris 上で IoT 分析環境構築(1/5) 〜概要編〜

2018年3月

今回からサンプルシステムの構築を通して、SPARC Servers / Solaris の良さを解説していきたいと思っています。

はじめに~CPUの性能向上とシステム構成の変化~

昨今、CPU の性能は上がっていますが、それに付随する I/O は早くなっていません。従来のように単体のマシンで作業 を行うと、CPU を使い切る前に I/O その他がボトルネックとなり性能が上がらなくなります。例えば、シンプルな Web サ ーバなどでは 2~3 コアで 10Gbps の LAN の帯域を使い切ってしまいます。

そのため、さまざまなサーバを仮想化して統合するようになってきました。それだけでなく SDN などのように専用ハードウ ェアで行ってきたようなこともサーバに任せるようになってきました。

仮想集約では PC サーバが多くの場合使われますが、SPARC/Solaris には OS とハードウェアを統合した利点が数多くあります。

このシリーズではそれらの利点にスポットライトをあてていきたいと思います。

仮想統合基盤としての SPARC Servers / Solaris

ここでは仮想統合基盤として SPARC Servers / Solaris がなぜよいかについてお話します。

SPARC / Solaris で使用できる仮想化機能

SPARC / Solaris では3種類の仮想化機構が使用可能です。

ーつ目はハードウェアパーティショニングで物理的に分離するものです。一部の機種のみで使用可能ですので今回は詳 しくは説明しません。

二つ目は Oracle VM Server for SPARC です。Linux では Xen サーバに相当します。

三つ目は Solaris Zones です。Solaris OS を仮想的に分割することで、独立した仮想マシン環境であるゾーンを構築する ことができます。Linux では LXC(Docker)に相当します。

Oracle VM Server for SPARC の特徴

Oracle VM Server for SPARC (以下 OVM) では、物理サーバをファームウェア層のハイパーバイザーにより論理ドメイ ンと呼ばれる仮想マシンに分割することで、独立した Solaris OS が動作する仮想マシンを構成します。ファームウェア層 で各論理ドメインを制御しているため仮想化によるオーバーヘッドはほとんどありません。

各論理ドメインの OS は完全に独立しています。ある論理ドメインがクラッシュしても他の論理ドメインは問題なく動作す ることが可能です。

OVM では論理ドメインチャンネル(以下 LDC)を利用することができます。LDC はハイパーバイザーを介した仮想マシン 間の通信を行うための機構です。ファームウェアベースで動作することにより、CPU 負荷を抑えたネットワーク通信が可 能です。

OVM では上記のようにハードウェアと結びついた仮想化支援機構を利用することが可能であるため、仮想統合基盤とし て非常に有用であると思われます。

Solaris Zone の特徴

OVM ほど仮想マシンの独立性が必要でない場合は Solaris Zone が非常に有効です。 OVM とは違い Solaris Zone は同一のカーネル空間で動作するため、不要な処理が動作せず、リソースの使用量を抑え ることができます。

実際のシステムでは OVM と組み合わせることで、隔離性とリソース配分を最適化することが可能です。

サンプルシステムの概要

サンプルモデルとしては IoT 機器からインターネットを介してセンサーメッセージを受信して、Hadoop に蓄積して Big Data 分析を行うという処理を考えています。



図1 処理の流れ

このブログでのシステム構成は以下のようになります。



各機能は仮想マシンとして実装します。DMZ やイントラネットなど厳密に隔離する必要があるセクションは OVM で仮想 化します。アプリケーションサーバなど隔離する必要が低いものは各 OVM 内で Solaris Zone を使用して構成します。 ハードウェアは M12-2 を前提としています。使用するソフトウェアの冗長化の関係で3 台構成としています。

以下は各処理とその概要です。

ODMZ ドメイン

メッセージの受診を行う以下の仮想マシンを配置する論理ドメインです。このドメイン上の仮想マシンはインターネットからアクセスされることを前提としています。

• ロードバランサーゾーン

メッセージを受信して Web/AP サーバに転送する Solaris Zone です。ロードバランサーには Solaris 付属の Integrated Load Balancer を使用します。各ポートのアクセス制限は Solaris 付属の Packet FIlter を使用します。 • アプリケーションサーバゾーン

ロードバランサーから転送されたメッセージを受信しメッセージングサービスゾーンに送信する Solaris Zone です。 アプリケーションサーバとしてはオープンソースの Payara を使用します。各ポートのアクセス制限は Solaris 付属 の Packet FIlter を使用します。

Oファイアーウォールドメイン

DMZ ドメインとイントラネットドメインの間のファイアーウォールを行う論理ドメインです。各ポートのアクセス制限は Solaris 付属の Packet FIlter を使用します。セキュリティの観点から OVM の論理ドメインとし、その上でファイアーウォー ルの機能を実行する Solaris Zone を構成します。

Oイントラネットドメイン

メッセージを受信/蓄積し、分析を行いう論理ドメインです。

- メッセージングサービスゾーン アプリケーションサーバからメッセージを受け取りキューに溜めておき、メッセージを取り込まれるのを待ちます。オ ープンソースの Apache Kafka を使用します。
- データ分析ゾーン
 メッセージングサービスゾーンからメッセージを取り出し分析可能な形式で蓄積します。オープンソースの Apache
 Hadoop と Apache Spark を使用します。
- クラスター管理ゾーン

Apache Zookeeper を使用して、Hadoop と Kafka のクラスター制御を行います。

構成について

物理マシンと各論理ドメインは以下のようにします。3台のマシンともドメイン名以外は同一の構成にします。



稼働環境は以下を前提としています。OSS 関連は該当の章で記載します。

- サーバ SPARC M12-2
- OS Solaris11.3 SRU 25.3

ハードウェア設定と論理ドメインの構築

物理マシンの設定

M12 には独立した 3 つのネットワーク物理ポートが必要です。インターネット向けと DMZ 用とイントラネット用です。サン プルでは物理カードを増設して使用することを想定しています。

セキュリティのため各ドメインのグローバルゾーンはこれらのポートには接続しません。

オンボードの LAN ポートはシステム管理用です。パッチの適用などの各ドメインのメンテナンス時に使用します。各ドメインのグローバルゾーンのみ接続し、ノングローバルゾーンは接続しないようにします。

今回はネットワークの冗長化は行っておりません。実システムでは IPMP やリンクアグリゲーション等で冗長化してください。

制御ドメインの設定

制御ドメインの構成ですが、今回はコア数を2、メモリ量を14GBとします。ホスト名は任意でかまいません。

以下はゲストドメイン向けサービスの定義となります。制御ドメイン上で root にて作業してください。

まずディスク関係の設定を行います。

各ゲストドメインのインストール用の仮想ディスクを ZFS の ZVOL として作成します。容量の目安としては dmz ドメイン用で 30GB、ファイアーウォール用で 20GB、イントラネットドメイン用にはデータを蓄積する関係で可能な限り多い領域を割り当ててください。

zfs create -V 30G rpool/dmz0
zfs create -V 20G rpool/fw0
zfs create -V 100G rpool/intra0

仮想ディスクに各ゲストドメインからアクセスするため、仮想ディスクサーバを設定します。

Idm add-vdiskserver vds0 primary

仮想ディスクサーバにディスクを追加します。

Idm add-vdiskserverdevice /dev/zvol/dsk/rpool/dmz0 dmz0@vds0

Idm add-vdiskserverdevice /dev/zvol/dsk/rpool/fw0 fw0@vds0

Idm add-vdiskserverdevice /dev/zvol/dsk/rpool/intra0 intra0@vds0

ディスク関連の設定の次はネットワーク関連の設定を行います。

まず、物理ポートとリンク名の確認を行います。以下は構成例です。

# dladm show-	-phys				
LINK	MEDIA	STATE	SPEED	DUPLEX	DEVICE
net7	Ethernet	unknown	0	unknown	igb3
net6	Ethernet	up	1000	full	igb2
net5	Ethernet	up	1000	full	igb1
net4	Ethernet	up	1000	full	igb0
net2	Ethernet	unknown	0	unknown	ixgbe2
net1	Ethernet	unknown	0	unknown	ixgbe1
net0	Ethernet	up	1000	full	ixgbe0
net3	Ethernet	unknown	0	unknown	ixgbe3

各ネットワーク物理ポートに仮想スイッチを作成します。

#	dm	add-vsw	net-dev=net	0 vswO	primary

Idm add-vsw net-dev=net4 vsw1 primary

Idm add-vsw net-dev=net5 vsw2 primary

Idm add-vsw net-dev=net6 vsw3 primary

制御ドメインからゲストドメインにアクセスするための仮想コンソールを用意します。

Idm add-vconscon port-range=5000-5999 vcc0 primary

仮想コンソールサービスを起動します。

svcadm enable vntsd

DMZドメインの設定

ドメインを作成します。

Idm add-domain dmz

コア数の設定をします。GZに1コア、二つのNGZに各1コアの計3コアを割り当てます。

ldm set-core 3 dmz

メモリ量の設定をします。アプリケーションサーバが動くので 32GB を割り当てます。

Idm set-memory 32g dmz

仮想ディスクを追加します。

Idm add-vdisk dmz0 dmz0@vds0 dmz

仮想ネットワークデバイスを設定します。NGZ が使用する仮想ネットワークデバイスには vnic 用の mac address を追加します。vnet を作成した順番にドメイン上では vnet0, vnet1 と設定されていきます。

Idm add-vnet vnet0 vsw0 dmz

Idm add-vnet alt-mac-addrs=auto, auto vnet1 vsw1 dmz

Idm add-vnet alt-mac-addrs=auto, auto vnet2 vsw2 dmz

OS のインストール時にマシン名は dmz1、dmz2、dmz3 とします。net0(vnet0)が管理 LAN 用のインターフェースとなります。

ファイアーウォールドメインの設定

ドメインを作成します。

Idm add-domain fw

コア数の設定をします。GZに1コア、NGZに1コアの計2コアを割り当てます。

Idm set-core 2 fw

メモリ量の設定をします。ファイアーウォール向けだけですので、16GBを割り当てます。

Idm set-memory 16g fw

Idm add-vdisk fw0 fw0@vds0 fw

仮想ネットワークデバイスを設定します。NGZ が使用する仮想ネットワークデバイスには vnic 用の mac address を追加します。vnet を作成した順番にドメイン上では vnet0, vnet1 と設定されていきます。

Idm add-vnet vnet0 vsw0 fw
Idm add-vnet alt-mac-addrs=auto vnet1 vsw2 fw
Idm add-vnet alt-mac-addrs=auto vnet2 vsw3 fw

OS のインストール時にマシン名は fw1、fw2、fw3 とします。net0(vnet0)が管理 LAN 用のインターフェースとなります。

┃ イントラネットドメインの設定

ドメインを作成します。

Idm add-domain intra

コア数の設定をします。GZ に 1 コア、各 NGZ に 2 コアの計 7 コアを割り当てます。

Idm set-core 7 intra

メモリ量の設定をします。Hadoop や Spark 用に 512GB を割り当てます。

Idm set-memory 512g intra

仮想ディスクを追加します。

Idm add-vdisk intra0 intra0@vds0 intra

仮想ネットワークデバイスを設定します。NGZ が使用する仮想ネットワークデバイスには vnic 用の mac address を追加します。vnet を作成した順番にドメイン上では vnet0, vnet1 と設定されていきます。

Idm add-vnet vnet0 vsw0 intra

Idm add-vnet alt-mac-addrs=auto, auto, auto vnet1 vsw3 intra

OS のインストール時にマシン名は intra1、intra2、intra3 とします。net0(vnet0)が管理 LAN 用のインターフェースとなります。

今回は処理概要の説明と各仮想マシンの作成準備までを行いました。 次回は DMZ ドメインの NGZ 上でロードバランサーとアプリケーションサーバのインストールを行います。

【免責事項】

- 富士通(株)は、本コンテンツの内容について、妥当性や正確性について保証するものではなく、一切の責任を負い兼ねます。
- 本コンテンツや URL は、予告なしに変更または中止されることがあります。あらかじめご了承願います。
- 理由の如何に関わらず、情報の変更及び本コンテンツの掲載の中断または中止によって生じるいかなる損害についても責任を負うものではありません。

関連情報:http://www.fujitsu.com/jp/about/resources/terms/copyright/index.html