

PowerGres on Linux HA 技術資料

White Paper



2011 年 3 月

SRA OSS, Inc. 日本支社
〒171-0022 東京都豊島区南池袋 2-32-8 8F
Tel: 03-5979-2701 Fax: 03-5979-2702
<http://www.sraoss.co.jp/>

目次

1.PowerGres HA とは.....	3
1.1.PowerGres とは.....	3
1.1.1.PowerGres on Linux とは.....	3
1.2.クラスタとは.....	4
1.3.LifeKeeper とは.....	4
1.4.データレプリケーション方式の PowerGres on Linux HA とは.....	6
2.PowerGres HA の内部構成.....	7
2.1.PowerGres HA のシステム構成（共有ディスク方式）.....	7
2.2.PowerGres on Linux HA のシステム構成（データレプリケーション方式）.....	8
2.3.PowerGres リソースの階層構造.....	10
2.4.PowerGres の状態監視.....	11
2.5.障害発生時の動作例.....	12
2.6.クラスタ構成例.....	14
2.6.1.アクティブ・アクティブ構成.....	14
2.6.2.1 : n 構成.....	15
3.導入事例.....	16
3.1.PowerGres on Linux HA.....	16
4.サポートハードウェア.....	17
5.FAQ.....	23

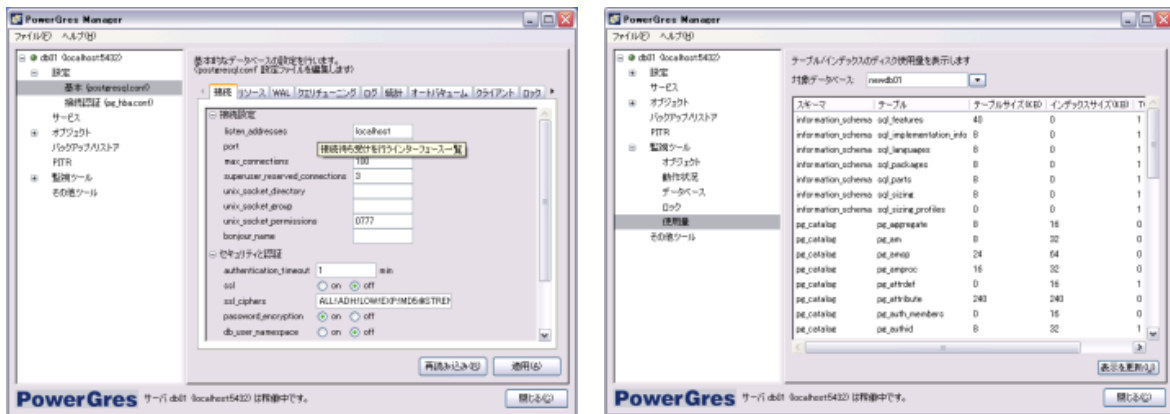
1. PowerGres HA とは

PowerGres HA は PowerGres の信頼性をより高めるためのソリューションです。PowerGres HA では、複数のサーバによって構成されたクラスタ環境内にデータベースを置き、障害発生時のシステムの停止時間を最小限に抑えることが可能です。PowerGres HA には、Linux 向けの PowerGres on Linux HA および PowerGres Plus HA と、Windows 向けの PowerGres on Windows HA があります。本資料では、PowerGres on Linux HA について紹介します。

1.1. PowerGres とは

PowerGres はオープンソースデータベースである PostgreSQL をベースとした製品です。これを使って HA クラスタ構成を組んだものが PowerGres HA にあたります。Linux で動く PowerGres には PowerGres on Linux と PowerGres Plus という 2 つの製品がありますが、PowerGres on Linux HA では PowerGres on Linux をベースとしています。

1.1.1. PowerGres on Linux とは



a. 設定ファイルの編集

b. 監視情報の閲覧

図 1 PowerGres on Linux の GUI 管理ツール

PowerGres on Linux は PostgreSQL をベースに GUI 管理ツールを付属した製品です。GUI 管理ツールではユーザの管理や設定ファイルの編集、データベースの起動や停止、監視情報の閲覧などを GUI から行うことができます。また、インストールも簡単で、インターフェイスには C、JDBC が含まれているので、すぐにアプリケーションの開発が始められます。

1.2. クラスタとは

PowerGres HA でクラスタ構成を組むことによって信頼性を高めていますが、そもそもクラスタとは何でしょうか。クラスタは複数のサーバを1つにまとめることによって信頼性や処理性能を高めることを目的としたシステムです。クラスタにはいくつかの種類がありますが、PowerGres HA は共有ディスク方式及びデータレプリケーション方式のフェイルオーバークラスタに分けることができます。

フェイルオーバークラスタとは、HA（High Availability：高可用性）クラスタの1つで、図2,3に示すように稼働系サーバに障害が発生した場合に待機系サーバにサービスやデータを切り替えることによってシステムの停止時間を最小限に抑えるというクラスタです。その中でも待機系サーバへのデータの引き継ぎに共有ディスクを使うものを共有ディスク方式のフェイルオーバークラスタ、ネットワーク経由でデータの複製を行うものをデータレプリケーション方式のフェイルオーバークラスタと呼びます。

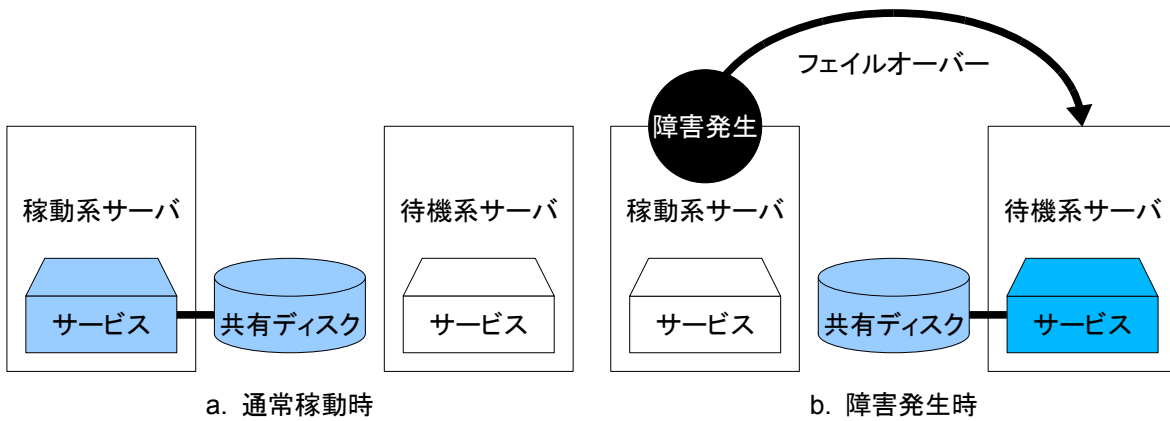


図2 フェイルオーバーの動作例(共有ディスク方式)

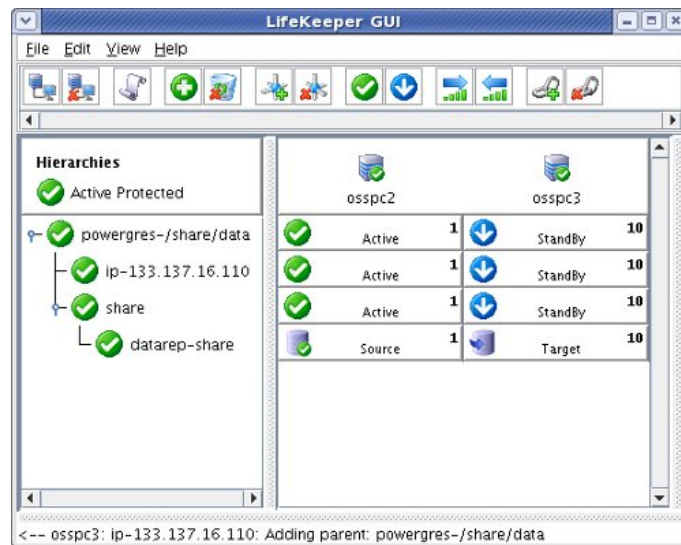


図 4 LifeKeeper の GUI 管理ツール

LifeKeeper では、アプリケーションのサービスをクラスタ環境内で保護するため、リカバリーキットを使ってリソースを定義する必要があります。リカバリーキットとは、アプリケーションごとのリソースの定義、リソース間の依存関係の定義、障害の検出、フェイルオーバーなどを提供するものです。PowerGres HA には PowerGres のサービスを LifeKeeper の保護対象として登録、管理するための PowerGres リカバリーキットが付属しています。なお、LifeKeeper には、標準で付属しているファイルシステムリカバリーキットや IP リカバリーキットの他にも、様々なアプリケーションを保護するためのリカバリーキットが提供されています（別途購入が必要）。

- | | |
|--------------------------------------|---|
| メール | – SDD Multipath ARK |
| – Postfix ARK | – PowerPath ARK |
| ウェブ | – Device Mapper Multipath (DMMP) ARK |
| – Apache ARK | – Hitachi Dynamic Link Manager Software (HDLM) ARK |
| データベース | – NEC iStorage StoragePathSavior (SPS) Recovery Kit |
| – DB2 ARK | その他 |
| – PostgreSQL ARK | – Generic ARK for JP1/AJS3 Manager |
| – Oracle ARK | – Generic ARK for JP1/AJS3 Agent |
| – Informix ARK | – Software RAID (md) ARK |
| – MySQL ARK | – WebSphere MQ ARK |
| ファイルシステム | |
| – NFS Server ARK | |
| – Logical Volume Manager (LVM) ARK | |
| – Network Attached Storage (NAS) ARK | |
| – Samba ARK | |
| マルチパス | |

※引用元 サポートアプリケーション一覧 http://www.sios.com/pdf/product/lifekeeper/l_support.pdf

1.4.データレプリケーション方式の PowerGres on Linux HA とは

PowerGres on Linux HA は、バージョン 6 からは共有ディスク構成の他にデータレプリケーション構成にも対応しています。データレプリケーション構成とは、稼動中のサーバのデータをネットワーク経由で他の待機サーバへレプリケーション(複製)する構成です。ローカルに接続されたディスクを使用するため、共有ディスクは不要になります。ただし、LifeKeeper のオプション Data Replication for Linux が必要になります。LifeKeeper の基本機能に Data Replication for Linux を組み合わせたものが Protection Suite for Linux として提供されています。

2. PowerGres on Linux HA の内部構成

2.1. PowerGres on Linux HA のシステム構成(共有ディスク方式)

共有ディスク方式の PowerGres on Linux HA の基本的なシステム構成は図 5 に示されるようになります。この構成では、稼働系サーバと待機系サーバ上の LifeKeeper が互いにサーバの状態、PowerGres や共有ディスク、ネットワークなどのリソースの状態を監視しています。もし、システムに障害が発生した場合は LifeKeeper によって稼働系サーバで起動していた PowerGres のサービスが待機系サーバにフェイルオーバーされます。また、このように稼働系サーバのみでサービスを提供する構成をアクティブ・スタンバイ構成と呼びます。

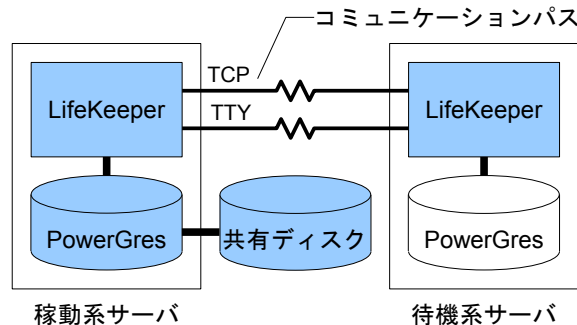


図 5 PowerGres on Linux HA の基本システム構成(共有ディスク方式)

なお、LifeKeeper ではサーバや PowerGres などのリソースの状態を監視するため、コミュニケーションパスを通してサーバ間での通信を行っています。コミュニケーションパスのインターフェイスには TCP または TTY が使えますが、LifeKeeper ではコミュニケーションパスが最低 2 本なければなりません。

このようなことから、共有ディスク方式の PowerGres HA を導入するにはハードウェアが最低限以下のことを満たしている必要があります。

- サーバが 2 台（稼働系サーバと待機系サーバ）以上あること
- 共有ディスクがあること
 - 共有ディスクとそのアダプタがサポートされていること（4 章を参照）
- それぞれのサーバにはネットワークインターフェイスが 2 本以上またはネットワークインターフェース 1 本以上と TTY 1 本あること（ネットワークインターフェース 1 本はサービス用とコミュニケーションパス用で兼用）

ネットワークインターフェースは、サービス用とは分けてコミュニケーションパス専用 2 本用意するのが推奨です。

また、ここで述べたシステム構成はあくまで基本です。LifeKeeper では他のリカバリーキットを組み合わせてることによって様々なクラスタ構成を組むことができます。

2.2. PowerGres on Linux HA のシステム構成(データレプリケーション方式)

PowerGres on Linux HA のデータレプリケーション方式の基本的なシステム構成は、図 6 に示されるようになります。共有ディスク方式との違いは、共有ディスクの代わりに稼働系サーバ・待機系サーバが各々データを保持し、ネットワーク経由で常にデータをレプリケーション（複製）している点です。

レプリケーションする領域はディスクパーティション単位で指定が可能です。データの転送元になる稼働系サーバのデータ領域を「Source」、データの転送先になる待機系サーバのデータ領域を「Target」と呼びます。もし、システムに障害が発生したときは、LifeKeeper によって稼働系サーバで動作している「Source」が待機系サーバにフェイルオーバーされ、待機系サーバがデータの転送元として働きます。

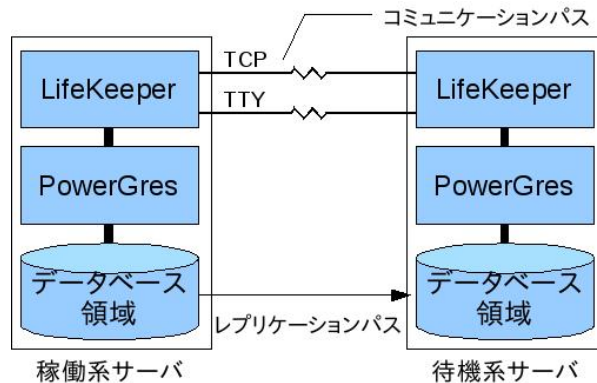


図 6. PowerGres on Linux HA (データレプリケーション方式)

データレプリケーションの動作を図 7 に示します。データをネットワーク経由でレプリケーションさせるために、LifeKeeper では RAID-Level1 及び NBD(Network Block Device)機能を利用しています。また、同期・非同期モードによるデータレプリケーションが可能です。非同期モードについては Linux kernel 2.6.16 以降を必要とします。

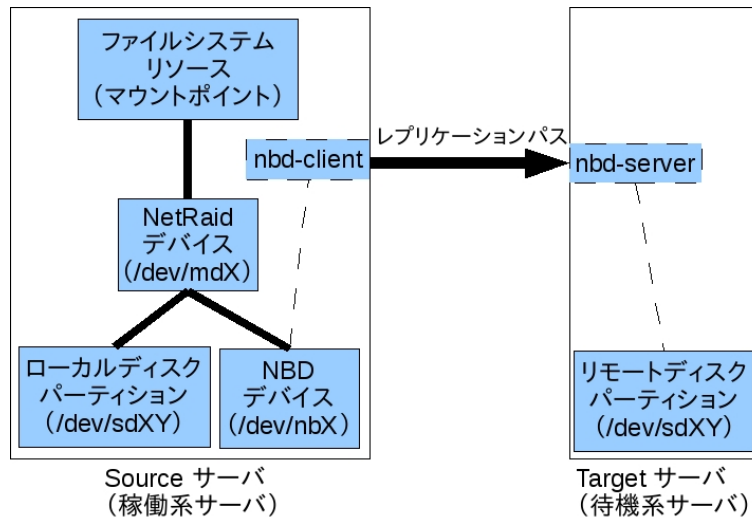


図 7.データレプリケーションの動作

また、データの複製をネットワーク経由で行う性質上、共有ディスク方式と比較して I/O 性能が低下します。

このようなことから、データレプリケーション方式の PowerGres on Linux HA を導入するにはハードウェアが最低限以下のことを満たしている必要があります。

- サーバが 2 台（稼働系サーバと待機系サーバ）以上あること
- それぞれのサーバにはネットワークインターフェイスが 2 本（コミュニケーションパスに 2 本、兼用でサービスの提供に 1 本、データレプリケーションパスに 1 本）以上あること
- データレプリケーション用ネットワークインターフェイスの通信速度規格が 1Gbps 以上あること

データレプリケーション用のパスはコミュニケーションパスのうちの 1 本を使用します。推奨は、コミュニケーションパスに 2 本、そのうちの 1 本をデータレプリケーション用とし、サービス用に別途 1 本の計 3 本以上のネットワークインターフェースを用意することです。

データレプリケーション方式の PowerGres on Linux HA で同期モードを利用した場合、単体サーバで動作させたときと比較して、どの程度パフォーマンスが低下するのか弊社で検証した結果を図 8 に示します。

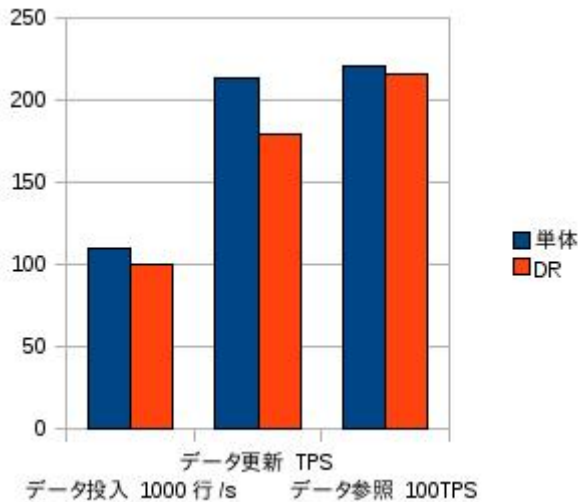


図 8 PowerGres on Linux HA の性能比較
単体サーバの動作 VS データレプリケーション動作（同期モード）

- 検証機器: HP DL380G4（2台）
- レプリケーションパス用 NIC: GigaBit Ethernet Adapter
- ディスク: Ultra320 SCSI (SmartArray)
 - Data 領域 / 72GB x 2 (RAID0)
 - OS 領域 / 72GB x 1
- テスト内容: pgbench Scale:10 Connect: 100 Transaction:100

検証結果より、データ投入・データ参照では性能差は僅かです。データ更新を含むトランザクション性能では、15%程度のダウンになります。なお、非同期モードの場合はデータ更新を含むトランザクション性能も差異がなくなります。

2.3. PowerGres リソースの階層構造

PowerGres HA では、LifeKeeper に対して PowerGres のサービスを登録して PowerGres リソースを作成すると、自動的にデータベースのデータを格納するディレクトリが存在するファイルシステムもリソースとして依存関係が作成されます。また、データベースへの接続は通常仮想 IP アドレスを通して行われるので、IP リソースを作成して PowerGres リソースとの依存関係を作成することが推奨されています。そのようにして作成された PowerGres のリソース階層は図 9 に示すようになります。

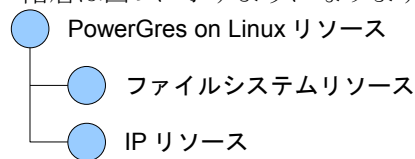


図 9 PowerGres リソースの階層構造

このリソース階層の親子関係はリソース間の依存関係を表しています。例えば、図 9 では PowerGres on Linux リソースはその子であるファイルシステムリソース、IP リソースに依存しています。そのため、PowerGres on Linux リソースに障害が発生したときはもちろん、それが依存するファイルシステムリソースや IP リソースに障害が発生したときも、PowerGres on Linux リソース全体としてフェイルオーバーが発生します。

2.4.PowerGres の状態監視

PowerGres on Linux HA では、PowerGres が正常に動作していることを確認するため、LifeKeeper によって PowerGres リソースの状態監視が定期的に行われています。LifeKeeper ではリカバリーキットを用いてそれぞれのサービスの状態監視を行っており、PowerGres の状態監視には PowerGres リカバリーキットが使われます。

PowerGres リカバリーキットでは以下のことについて上から順番に確認します。この確認に 1 つでも失敗した場合、PowerGres が正常に動作していないものと判断します。

1. データベースデーモン (postmaster) のプロセスが存在するか？
2. データベースに接続でき、SQL を実行できるか？

このようなリソースの状態監視は初期設定では 120 秒に 1 度行われます。なお、リソースに異常が発生した場合の動作については 25 節を参照してください。

2.5.障害発生時の動作例

アクティブ・スタンバイ構成において障害が発生した際の動作は図 10 に示すようになります。

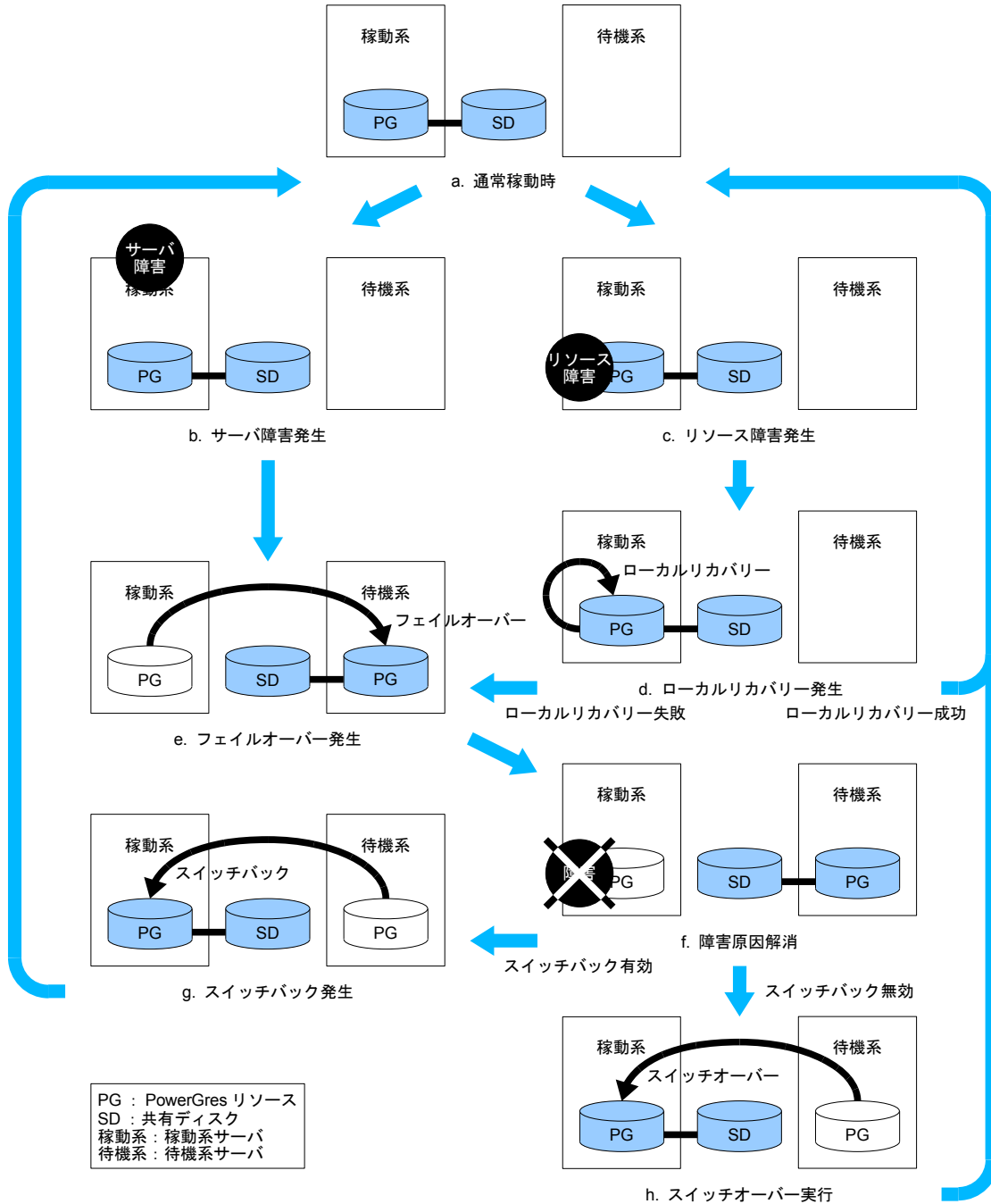


図 10 障害発生時の動作例(共有ディスク方式の場合)

a. 通常稼働時

アクティブ・スタンバイ構成における通常稼働とは稼働系サーバでサービスが提供されている状態を意味します。

b. サーバ障害発生

サーバに障害が発生すると LifeKeeper によってそれが検出されます。なお、初期設定では 5 秒ごとにコミュニケーションパスを通してサーバの状態が監視され、5 回続けて応答がない場合は相手のサーバに障害が発生していると考えられます。

c. リソース障害発生

LifeKeeper によって PowerGres や共有ディスク、仮想 IP アドレスなどといったリソースに障害が発生したことが検出されます。なお、リソースの状態は初期設定では 120 秒ごとに監視されています。リソース障害の原因としては PowerGres の停止や LAN ケーブルの切断など、様々な理由が考えられます。

d. ローカルリカバリー発生

LifeKeeper がリソース障害を検出した場合、まずは障害が発生したサーバでそのリソースのリカバリーを試みます。例えば、PowerGres が停止してしまった場合であれば PowerGres を起動しようとします。ローカルリカバリーに成功すると通常稼働時に戻ります。ローカルリカバリーに失敗した場合はフェイルオーバーが発生します。

e. フェイルオーバー発生

サーバに障害が発生した場合や、リソースに障害が発生してローカルリカバリーに失敗した場合はフェイルオーバーが発生します。リソース障害によるフェイルオーバーでは障害が発生したリソースと依存関係があるリソースもフェイルオーバーされます。例えば、PowerGres に障害が発生した場合であれば、それと依存関係がある共有ディスクや仮想 IP アドレスのリソースも稼働系サーバにフェイルオーバーされます。

f. 障害原因解消

フェイルオーバー発生後は待機系サーバでサービスが提供されます。ただし、稼働系サーバにはまだ障害が残っているので、障害の原因を調査して解消する必要があります。

g. スイッチバック発生

稼働系サーバで障害の原因を解消した際、スイッチバックオプションが有効に設定されていればスイッチバックが発生します。スイッチバックとはフェイルオーバー後に障害原因が解消された際にリソースを優先度の高いサーバに戻してサービスを再開することです。スイッチバックが発生すると稼働系サーバでサービスが提供されることになり、通常稼働時に戻ります。

h. スイッチオーバー実行

スイッチバックオプションが有効に設定されていない場合、稼働系サーバで障害の原因が解消されてもサービスは待機系サーバで提供され続けます。もし、稼働系サーバでサービスを提供したいときは手動でスイッチオーバーを実行する必要があります。スイッチオーバーとは任意のサーバにリソースを切り替えてそのサーバでサービスを再開することを意味します。スイッチオーバーを実行すると通常稼働時に戻ります。

2.6. クラスタ構成例

これまで PowerGres HA の基本的なシステム構成としてアクティブ・スタンバイ構成について述べましたが、PowerGres HA ではアクティブ・スタンバイ構成以外にも様々なクラスタ構成を組むことができます。

2.6.1. アクティブ・アクティブ構成

アクティブ・スタンバイ構成では稼働系サーバがサービスを提供し、待機系サーバでは稼働系サーバに障害が発生した場合のみしかサービスを提供しません。そのようにフェイルオーバーが発生しない限り、使われることのない待機系サーバを有効的に使うための構成がアクティブ・アクティブ構成になります。アクティブ・アクティブ構成では図 11 に示すようにそれぞれのサーバでサービスが提供され、互いのサーバがフェイルオーバー先となります。

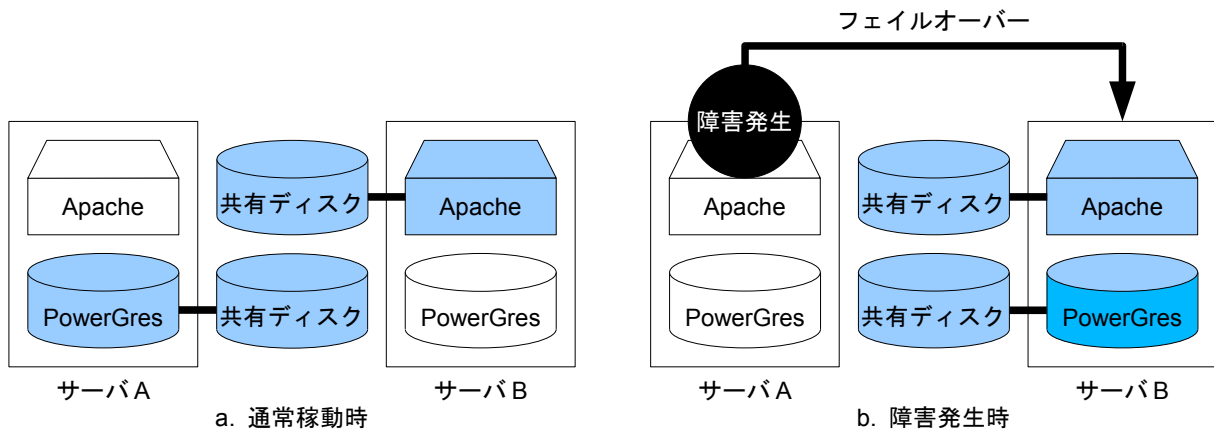


図 11 アクティブ・アクティブ構成におけるフェイルオーバー動作例 (共有ディスク方式の場合)

例えば、PowerGres HA に加えて Apache Web Server Recovery Kit (別途購入が必要) を使い、サーバAではPowerGres、サーバBではApacheのサービスを提供するようなアクティブ・アクティブ構成が可能です。もちろん、アクティブ・アクティブ構成であっても図 11 に示すようにサービスのフェイルオーバーには問題ありません。また、図 11 ではApacheによって提供されるコンテンツを共有ディスクに置くという構成ですが、コンテンツが静的なものであれば共有ディスクを使わずにそれぞれのサーバに置くという構成もできます。

2.6.2.1:n 構成

アクティブ・スタンバイ構成やアクティブ・アクティブ構成は1台の稼働系サーバに対して1台の待機系サーバという1:1構成ですが、PowerGres HA では1台の稼働系サーバに対して複数台の待機系サーバを置くという1:n構成も可能です。

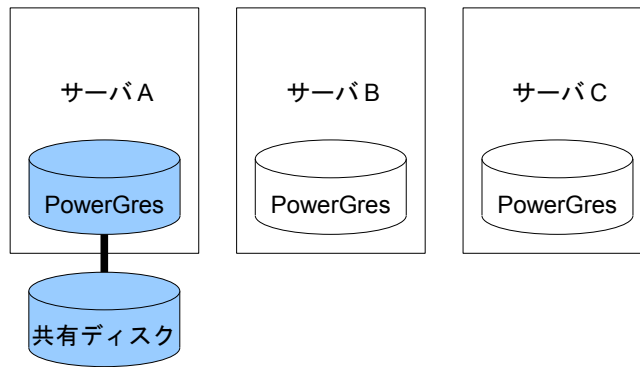


図 12 1:n 構成例(共有ディスクの場合)

例えば、図 12 に示される 1:n 構成の例では1台の稼働系サーバと2台の待機系サーバという構成によって PowerGres を LifeKeeper の保護下に置いています。この構成では、サーバ A で障害が発生するとサーバ B に、さらにサーバ B で障害が発生するとサーバ C にフェイルオーバーさせることが可能になります(どのサーバにフェイルオーバーするかはそれぞれのサーバの優先度によって決定します)。

このように待機系サーバを複数台置くことによってフェイルオーバー後における障害への対応が可能となり、可用性をさらに高めることができます。ちなみに 1:n 構成においてもアクティブ・アクティブ構成を組むことができます。

3.導入事例

3.1.PowerGres on Linux HA

某社では PowerGres on Linux HA を不動産関係のシステムのデータベースとして導入しています。このシステムでは図 13 に示すようにアプリケーションサーバ 2 台と PowerGres on Linux HA が導入されたデータベースサーバ 2 台から構成されています。

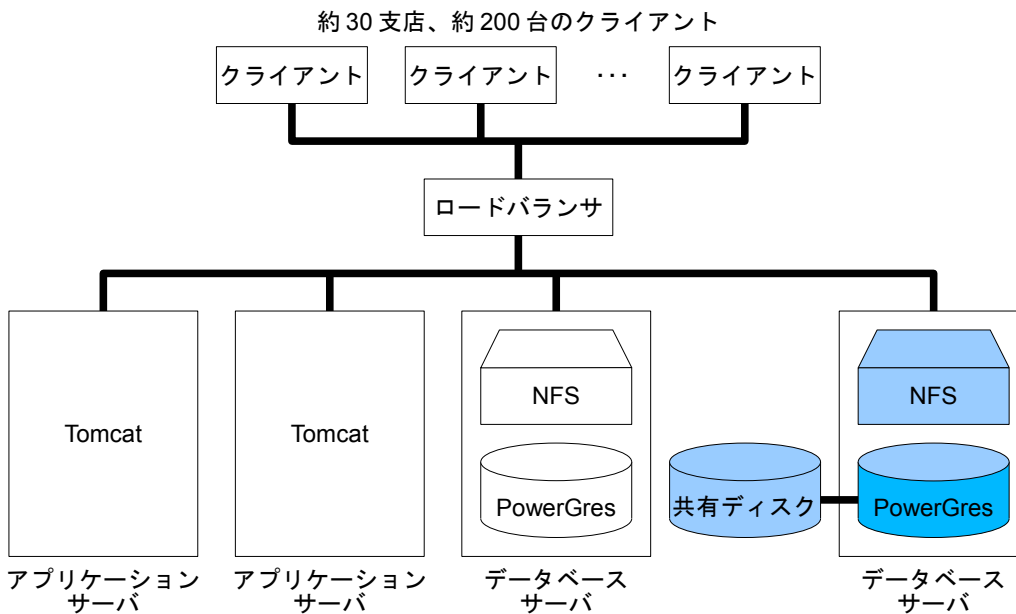


図 13 PowerGres on Linux HA を使ったシステム構成

このシステムには約 30 支店、約 200 台のクライアントからの要求があり、その要求はロードバランサによって 2 台のアプリケーションサーバに振り分けられ、負荷分散が行われています。アプリケーションサーバではその要求に応じてデータベースへの接続が行われます。データベースサーバでは PowerGres on Linux HA によってクラスタ構成が組み立てられており、全体として高い処理性能と信頼性を実現したシステム構成となっています。また、このシステムでは NFS も使用されていますが、それもまとめて LifeKeeper の保護下に置かれています。

4. サポートハードウェア

LifeKeeper でサポートされているストレージ（共有ディスク）およびそのアダプターは表 1、表 2 に示されるとおりです。なお、この情報はサイオステクノロジー株式会社のウェブサイト（<http://www.sios.com/>）にあるリリースノートから抜粋したものです。日本ではモデル名などが異なる場合があるので各ベンダーにお問い合わせください。また、最新のサポートハードウェアに関してはサイオステクノロジー株式会社のウェブサイトを確認するようにしてください。

表 1 サポートストレージ¹

ベンダー	ストレージモデル
ADTX	ArrayMasStor P ArrayMasStor L ArrayMasStor FC-II
Altix	TP9100
Baydel Storage Arrays	DAR3/5SE68C DAR3/C/5SE68C
Consan	CRD5440 CRD7220 (f/w 3.00)
DataCore	SANsymphony
Dell	650F (CLARiiON) Dell EMC CX3-10c, CX3-40c, CX3-20c, CX3-80, CX3-40(F), CX3-20(F) Dell EMC CX300, CX400, CX500, CX600, CX700 PowerVault (w/ Dell PERC, LSI Logic MegaRAID) Dell MD3000 Dell EqualLogic PS4000、PS5000、PS6000
EMC	Symmetrix 3000 Series Symmetrix 8000 Series Symmetrix DMX/DMX2 Symmetrix DMX3/DMX4 CLARiiON CX200, CX400, CX500, CX600, CX700 CLARiiON CX300 → PowerPath ARK を使用するとマルチパス構成が可能となり、 PowerPath ARK を使用しない場合は、シングルパス構成となります。

1 2010年12月現在

ベンダー	ストレージモデル
	CLARiX CX3-20 CLaRiiON CX3FC and combo 40290 CLaRiiON CX310c CLaRiiON AX4-5 CLaRiiON CX4-120, CX4-240, CX4-480, CX4-960
Fujitsu	ETERNUS3000 (w/ PG-FC105, PG-FC106, PG-FC107) →シングルパス構成のみサポート ETERNUS6000 (w/ PG-FC106) →シングルパス構成のみサポート ETERNUS4000 Model 80 , Model100(w/ PG-FC106, PG-FC107, PG-FC202) →シングルパス構成のみサポート FibreCAT S80 ETERNUS SX300 (w/ PG-FC106, PG-FC107) →マルチパス構成のみサポート ETERNUS2000 シリーズ: Model 50, Model 100, Model 200 (w/ PG-FC202) →シングルパス、マルチパスの両方の構成をサポート ETERNUS4000 シリーズ: Model 300, Model 500 (w/ PG-FC202) →シングルパス、マルチパスの両方の構成をサポート ETERNUS DX80 / DX60 (w/ PG-FC203L) →シングルパス、マルチパス(ETERNUS MPD)の両方の構成をサポート ※ マルチパスドライバを使用したシングルパス接続はサポートされません。
Hitachi Data Systems	HDS 7700 HDS 5800 HDS 9570V HDS 9970V HDS 9980V AMS 500 SANRISE USP/NSC (TagmaStore USP/NSC) ※LifeKeeper for Linux v5.1 にて検証を実施しました。 BR1600 AMS2100 AMS2300

ベンダー	ストレージモデル
HP/Compaq	<p>AMS2500</p> <p>RA 4100</p> <p>MA/RA 8000</p> <p>MSA1000 / MSA1500 (active/active and active/passive firmware configurations)</p> <p>HP MSA1000 Small Business SAN Kit</p> <p>EVA3000, 5000</p> <p>EVA4X00, 6X00, 8X00</p> <p>※EVA4400, 6400, 8400 は次の構成をサポートいたします。</p> <ul style="list-style-type: none"> ・ OS: Red Hat Enterprise Linux 5 Update 3 以降 ・ HBA ドライバー: 上記 OS に同梱の qllogic ドライバー ・ マルチパスドライバー: 上記 OS に同梱の Device Mapper Multipath ・ LifeKeeper: LifeKeeper for Linux v6 Update 4, v7 以降 ・ マルチパス ARK: DMMP ARK v6 Update 4, v7 以降 <p>(シングルパス、マルチパスどちらの構成でも DMMP ARK が必要です。)</p> <p>MSA500 (formerly Smart Array Cluster Storage)</p> <p>MSA500 G2</p> <p>MSA2000 Fiber Channel</p> <p>MSA2000 iSCSI</p> <p>MSA2000 SA</p> <p>MSA2300 Fiber Channel</p> <p>MSA2300i (DMMP Recovery Kit が必要)</p> <p>MSA2300sa (DMMP Recovery Kit が必要)</p>
IBM	<p>FAStT200</p> <p>FAStT500</p> <p>DS4100</p> <p>DS4200</p> <p>DS4300 (FAStT600)</p> <p>DS4400 (FAStT700)</p> <p>DS4500 (FAStT900)</p> <p>DS4700</p> <p>DS4800</p> <p>※シングルパス、マルチパス構成の詳細については、リリースノートをご覧ください。</p> <p>http://us.sios.com/Linux_95.htm</p>

ベンダー	ストレージモデル
	DS5000 DS6800 DS8100 DS400 (シングルパスのみのサポート) DS3400 DS3200 DS3300 EXP300 (w/ Server RAID) EXP400 (w/ ServerRAID) ESS Model 800 IBM eServer xSeries Storage Solution Server Type445-R for SANmelody IBM eServer xSeries Storage Solution Server Type445-FR for SANmelody IBM SAN Volume Controller
JetStor	JetStor II
MicroNet	GenesisONE
MTI	Gladiator 2550 Gladiator 3550 Gladiator 3600
NetApp	(NAS 構成) F8xx Series FAS2xx Series FAS9xx Series FAS3xxx Series FAS3050c (Data ONTAP7.1.1.1.) FAS6xxx Series (ONTAP 7.2.x) (SAN 構成) FAS3xxx Series (w/ QLogic QLE246x and DMMP) http://www.sios.com/product/lifekeeper/solution/solution-netapp.html
NEC	NEC iStorage S500 / S1500 / S2500 (シングルパスのみ) 以下、シングルパスおよび SPS Recovery Kit を使用してのマルチパス構成をサポート NEC iStorage S Series NEC iStorage D3-10 / D1-10 NEC iStorage D8-10 / D8-20 / D8-30

ベンダー	ストレージモデル
Newtech	SweeperStor SATA SweeperStor SAS
nStor	NexStor 4320F
ProCom	Reliant 1000
Radion Systems	Rack U2W Microdisk U2W
SGI	SGI InfiniteStorage 4600, Linux MPP driver
SILVERstor	Giant GT-3000 series
Sun	StorEdge 3310 StorEdge 3510 FC (w/ Sun StorEdge 2Gb PCI Single FC Network Adapter) StorEdge 6130 FC (w/ Sun StorEdge 2Gb PCI Single FC Network Adapter) StorageTek 2540 (w/ Sun StorageTek 4Gb PCI-E Dual FC Host Bus Adapter or Sun StorageTek 4Gb PCI Dual FC Network Adapter)
TID	MassCare RAID II
Winchester Systems	FlashDisk OpenRAID (SCSI) FlashDisk OpenRAID (FC,)
Xiotech	Magnitude 3D

※引用元 http://www.sios.com/pdf/product/lifekeeper/LKHW_Info.pdf

表 2 サポートアダプター

アダプタータイプ	アダプターモデル
Differential SCSI Adapter	Adaptec 2944 W, Adaptec 2944 UW, or Adaptec 2940 U2W Compaq 64bit PCI Dual Channel Wide Ultra2 SCSI Adapter Compaq SA 5i, 6i, 532, and 642 PCI Dual Channel Wide Ultra3 SCSI Adapters IBM ServeRAID-4Lx, ServeRAID-4Mx, ServeRAID-4H, and ServeRAID-6M SCSI Controllers Dell PERC 2/DC, PERC 4/DC LSI Logic MegaRAID Elite 1600 (Dell PERC 3/DC is the OEM version of this adapter) Adaptec 39160 Note that following adapters are Fujitsu tested in LifeKeeper configurations involving non-shared storage with IP failover only, or when using LifeKeeper Data Replication.

アダプタータイプ	アダプターモデル
	Adaptec ASR-2010S (Fujitsu PG-140C / CL) Adaptec ASR-3200S (Fujitsu PG-142B / C / D) LSI Logic MegaRAID SCSI 320-2 (Fujitsu PC-142E)
Fibre Channel	QLogic QLA 2100, QLogic QLA 2200, QLogic QLA 2340, QLogic QLA 200 (HP Q200) HP StorageWorks 2Gb 64-bit/133MHz PCI-X to Fibre Channel Host Bus Adapter (FCA2214) Compaq 64 bit/66MHz Fibre Channel Host Bus Adapter 120186-B21 Sun StorEdge 2Gb PCI Single FC Network Adapter (OEM'ed QLogic QLA 2310) Sun StorageTek 4Gb PCI-E Dual FC Host Bus Adapter Sun StorageTek 4Gb PCI Dual FC Network Adapter Emulex LP9002 (PG-FC105), Emulex LP1050, Emulex LP10000 (Emulex HBA には、2.4 カーネルの場合 lpfc ドライバ v7.1.13 以降が必要、2.6 カーネルの場合 v8.0.16 以降が必要となります) HP QLogic QMH2462 4Gb FC HBA Qlogic QLE2460 (4Gb HBA), Qlogic QLE2462 (4Gb HBA) FC1142SR 4Gb シングルチャネル PCI-Express Fibre Channel アダプタ FC1242SR 4Gb デュアルチャネル PCI-Express Fibre Channel アダプタ ※MSA1000 とのシングルパスおよびマルチパス構成で動作を確認しています。
Serial Attached SCSI (SAS)	DELL SAS 5/e adapters

※引用元 http://www.sios.com/pdf/product/lifekeeper/LKHW_Info.pdf

参考までに弊社での PowerGres HA におけるハードウェアの導入実績を図 14 に示します。もちろんこれに含まれていないベンダーのハードウェアもサポートされていればまったく問題はありません。

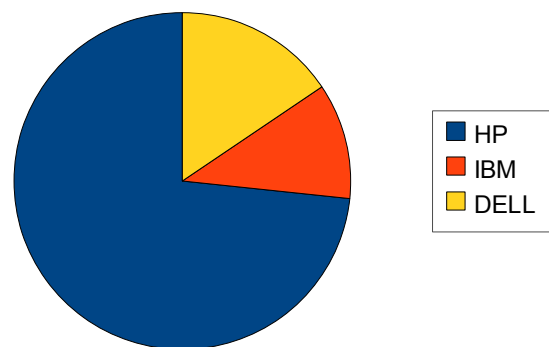


図 14 PowerGres HA におけるハードウェア導入実績

5.FAQ

質問 PowerGres on Linux HA と SteelEye Technology, Inc.から提供されている PostgreSQL Recovery Kit は何が違うのでしょうか？

回答 PowerGres on Linux HA は、PowerGres と弊社オリジナルのリカバリーキット、そしてサポートを含めた HA クラスターのソリューションです。従って、PowerGres on Linux HA では、PowerGres のバージョンアップ情報やパッチ情報、バグ情報を含む情報配信サービスや PowerGres に関する質問にお答えする QA サービス、障害対応という充実したサポートが含まれています。また、障害対応についてもお客様はそれが LifeKeeper の問題なのか、それとも PowerGres の問題なのかを切り分けする必要はまったくありません。

質問 フェイルオーバーはどのような原因で発生するのでしょうか？

回答 フェイルオーバーの原因はサーバの障害とリソースの障害という 2 つに分けられます。

まず、サーバ障害の原因として多いのはメモリやハードディスクなどのハードウェアに関する障害です。とくに 24 時間 365 日稼動するようなシステムではハードウェアの消耗は避けては通れません。もう 1 つの原因としてはリソースの障害が挙げられます。リソースの障害としては PowerGres がパニックで停止することや、LAN ケーブルが抜けてしまうことなどが原因として考えられます。

このような障害が発生したときでも、PowerGres HA では待機系にサービスをフェイルオーバーすることによってサービスを提供し続けることができます。

質問 PowerGres on Linux HA（共有ディスク方式）を導入する上での注意点などはありますか？

回答 PowerGres on Linux HA（共有ディスク方式）を導入するには、2.1 節でも述べましたがハードウェアが最低限以下のことを満たしている必要があります。

- サーバが 2 台（稼動系サーバと待機系サーバ）以上あること
- 共有ディスクがあること
 - 共有ディスクとそのアダプタがサポートされていること（3 章を参照）
- それぞれのサーバにはネットワークインターフェイスが 2 本またはネットワークインターフェース 1 本以上と TTY1 本あること

また、LifeKeeper のライセンスはネットワークカードの MAC アドレスに関連付けられているので、LifeKeeper を導入後にハードウェアの交換が行われた場合、ライセンスの再発行が必要になることがあります。

質問 PowerGres on Linux HA（データレプリケーション方式）を導入する上での注意点などはありますか？

か？

回答 PowerGres on Linux HA（データレプリケーション方式）を導入するには、2.2節でも述べましたがハードウェアが最低限以下のことを満たしている必要があります。

- サーバが2台（稼動系サーバと待機系サーバ）以上あること
- それぞれのサーバにはネットワークインターフェイスが2本以上あること
- データレプリケーション用ネットワークインターフェイスの通信速度規格が1Gbps以上あること