あなたのシステムを OSSで冗長化しよう

いつか来るシステム障害にそなえて

サイオステクノロジー株式会社
くろき　ひろし

講師紹介

名前: 黒木 博 (くろき ひろし)

担当: DRBDサポートを担当

資格 LINBIT認定 DRBDエンジニア
日本で一番DRBDに詳しい人

メールアドレス:h-kuroki@sios.com

2

この写真が私、黒木です。

システムを冗長化する理由

停止したくないシステムを動かし続けるため

私達のまわりの情報システムは日々止まることなく動いています。

現在はいろいろな情報システムが企業や学校や病院などなどで利用されています。
システムが複雑に絡み合い、止められないシステムが多くなってきました。

しかしシステムはコンピュータ機器なので、いつかは壊れる事があります。
ハード故障以外にいろいろな原因で止まってしまうかもしれません。
それは5年後かもしれませんし、10年後かもしれません。
もしかしたら明日壊れるかもしれません。

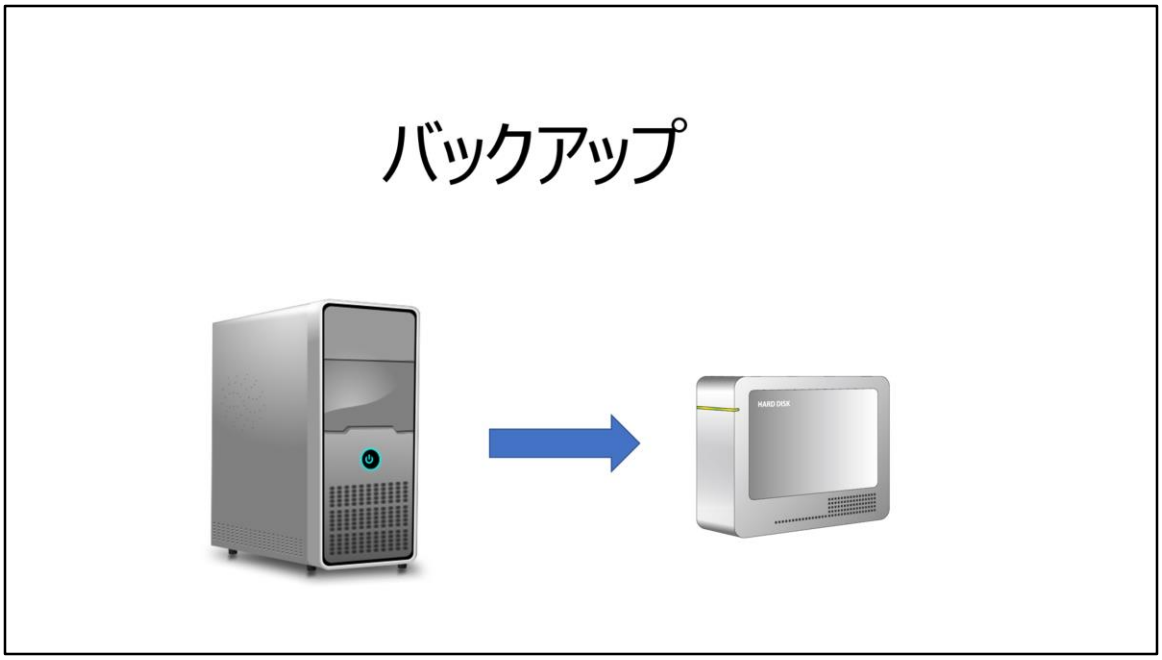
冗長化がシステムを救う

万が一システムが壊れても大丈夫なようにするいくつかの方法がありますが、
本日はシステムの冗長化についてお話します。

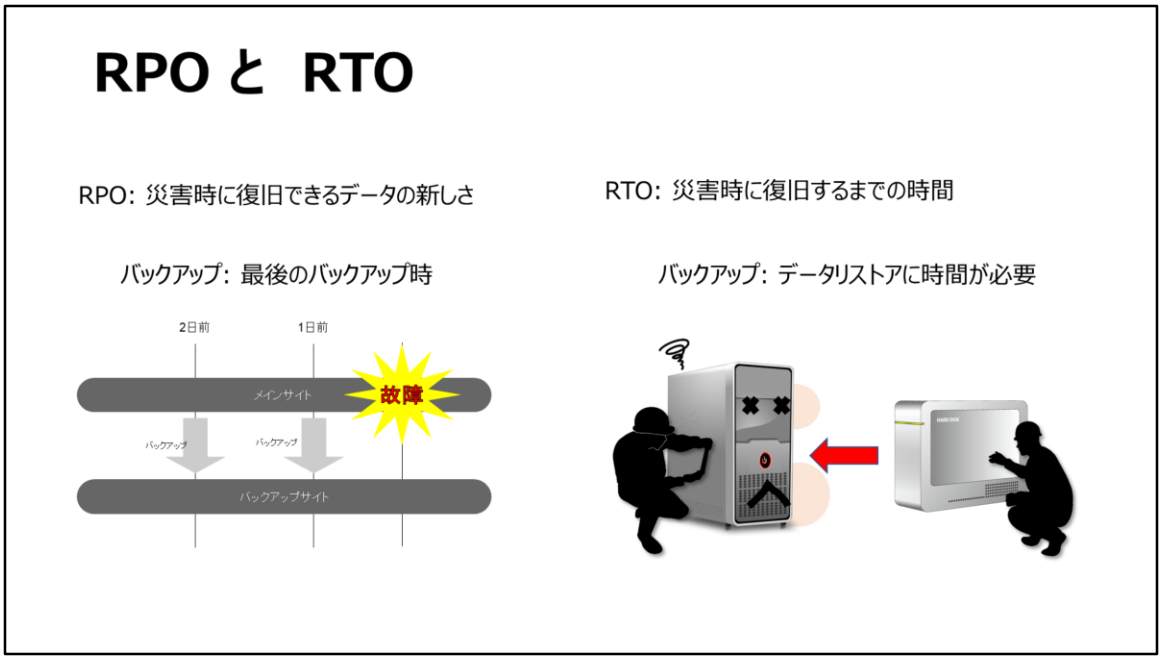
システム故障への対策

大事なシステムを守るため

それでは機器の故障からシステムを守るいろいろな方法を説明していきます。



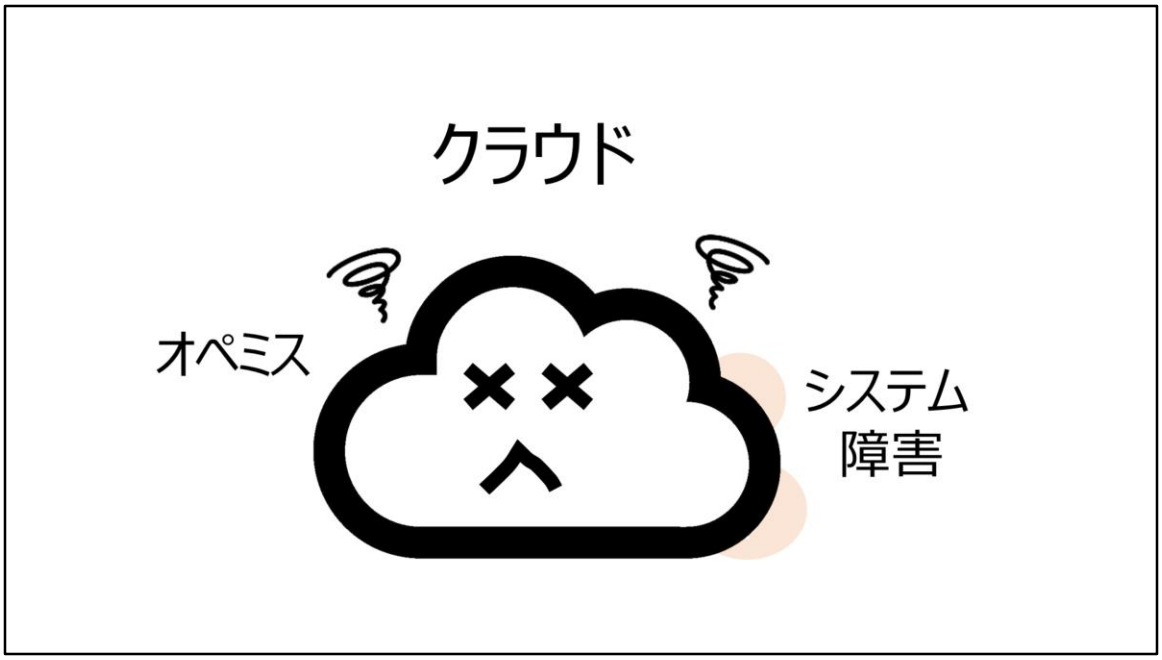
■機器故障への対策その１
まずはシステムのバックアップを取ることです。
定期的にバックアップを取っておけば、バックアップを元に壊れたシステムを復旧できます。



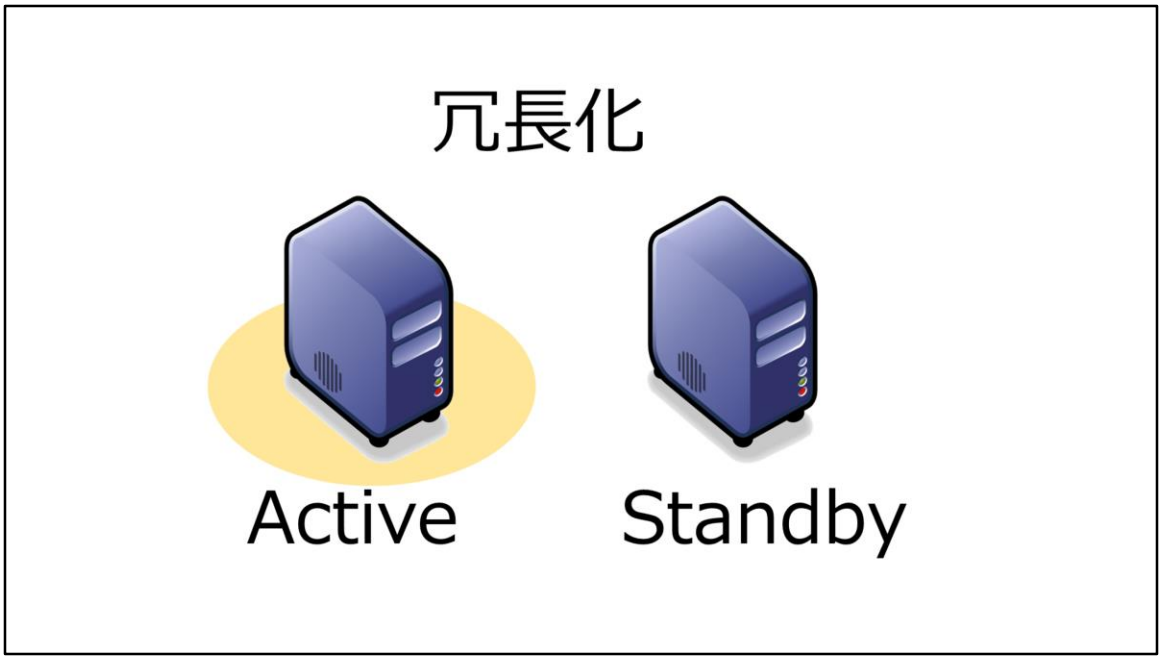
RPOとRTOというキーワードをご存知でしょうか。

バックアップされるデータは前回のバックアップ時点のデータです。
また、バックアップを使った復旧はシステムを構築しなおして、バックアップからデータをリストアするため、復旧まで時間がかかります。

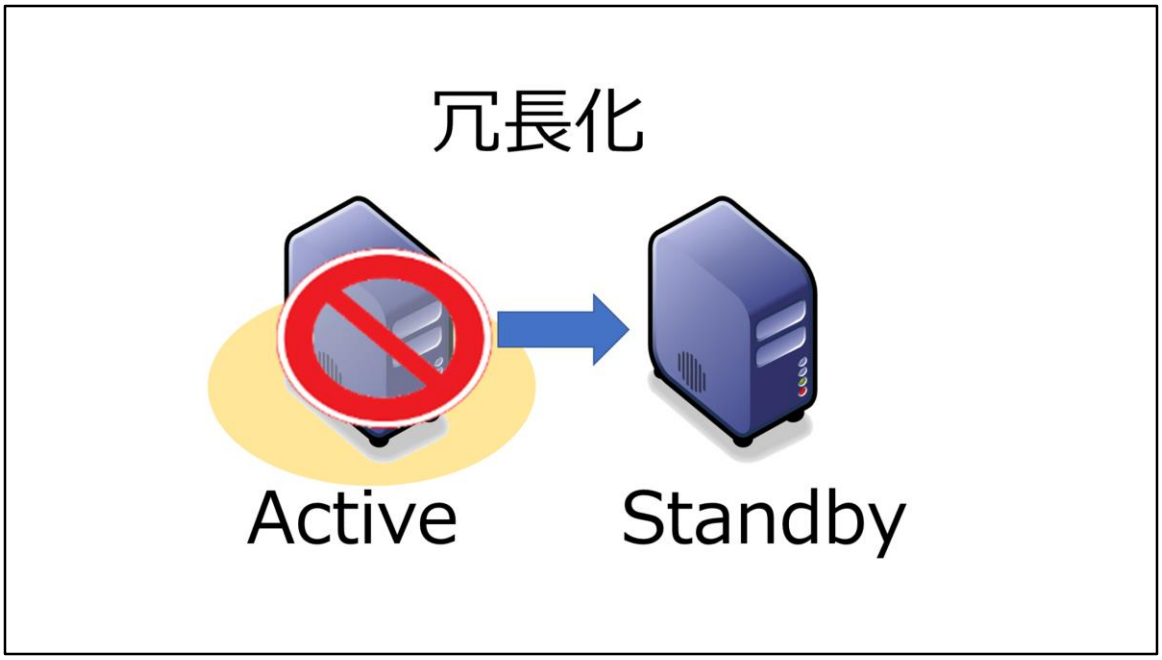
バックアップは簡便ですが、復旧には時間がかかる方法です



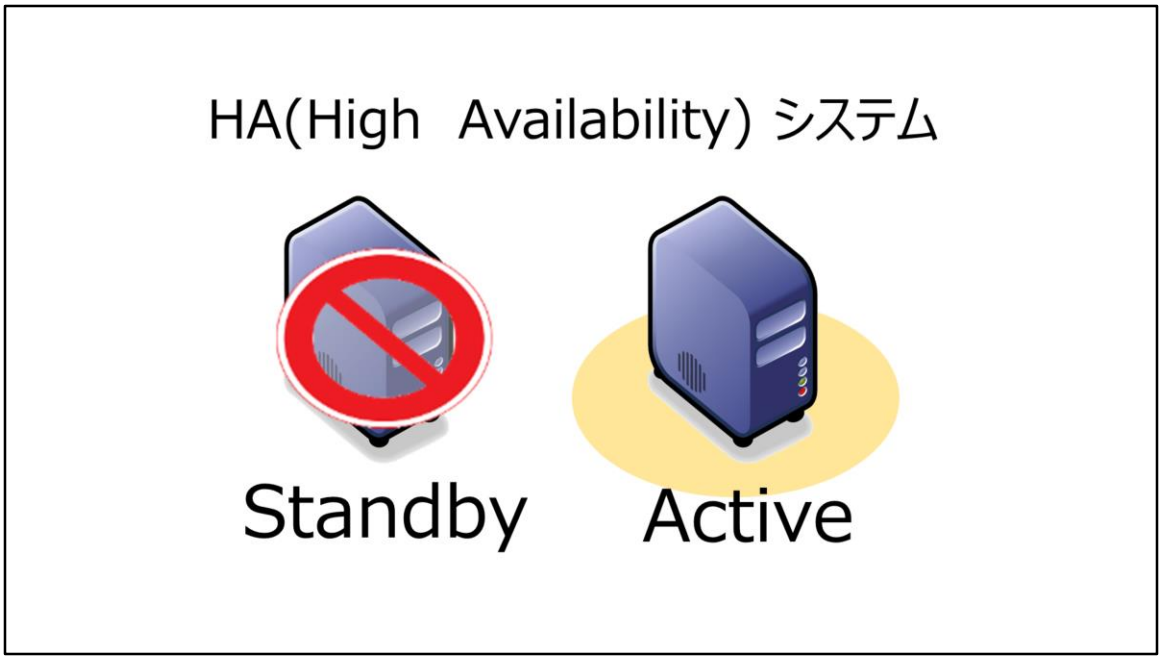
■機器故障への対策その２
クラウドに使えばどうでしょう？
クラウド環境ではクラウドの業者様がシステムの強靭化の対策をいろいろ
行っているので、致命的にシステムが壊れる心配はありません。
しかし、クラウドでシステム障害が発生する事もありますし、利用者がオペ
レーションミスでシステムを壊してしまう事もあります。
なにより大事なデータをクラウドに置けないという事もあります。



■機器故障への対策その３
機器の２重化を考えてみましょう。
サーバを２台構築して、２台のサーバに同じデータを置きます。１台のサーバが壊れても、もう１台サーバが残っているので、業務の継続ができます。



壊れたサーバでサービスが動いている時、予備機への切り替えが自動でできると、さらに素晴らしいシステムになり、サービスの管理者は安心して夜寝ることができます。



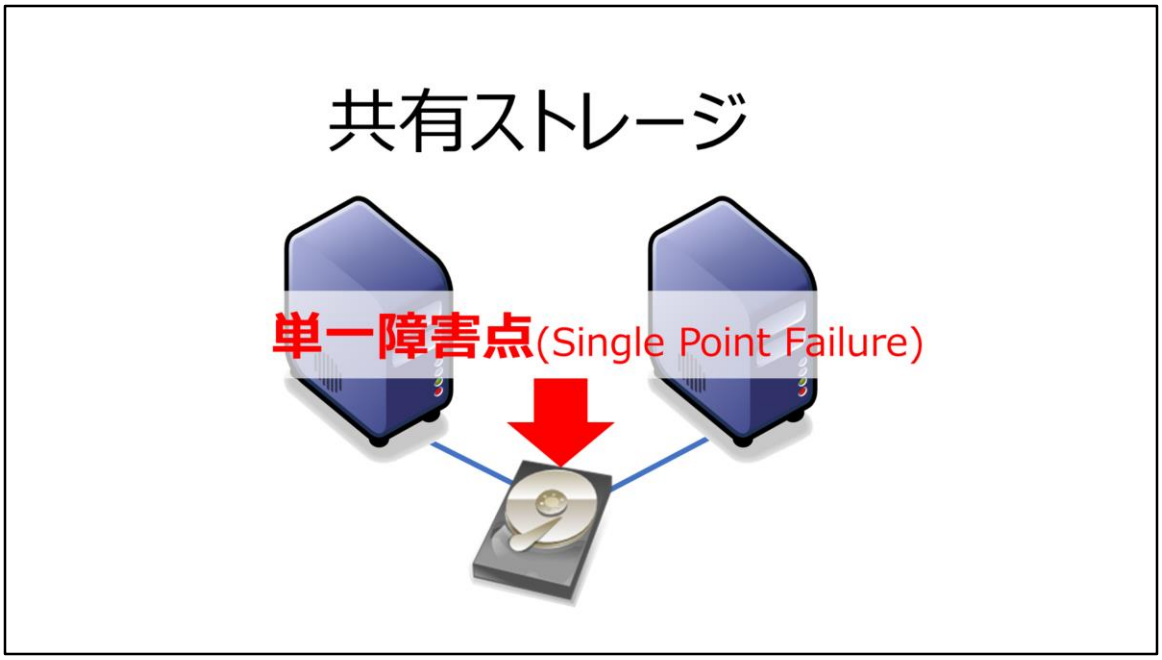
機器の２重化でシステムを冗長にすることを高可用性システム、通称HA（High　Availability）システムと呼びます。

データの冗長化

HAシステムでデータを共有する

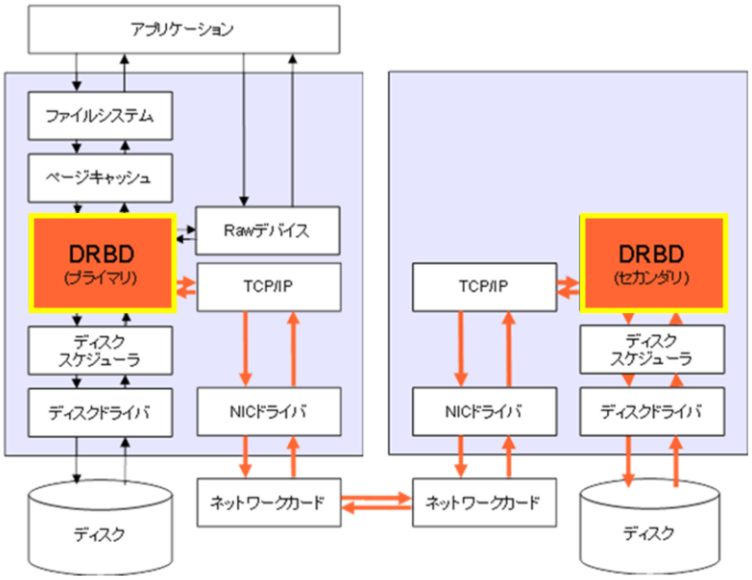
２台のサーバは同じデータが使える必要があります。
システムを構築した時のデータは日々の業務で更新され、スタンバイ側の
サーバのデータは古くなります。

もっとも簡単な方法はrsyncコマンドで定期的にバックアップを取る方法です。



NFSやSANなどの共有ストレージを使う方法もあります。
ただしこの構成はストレージが壊れてしますと、サーバーが無事に動いていてもサービスが停止してしまいます。
これを単一障害点と言います。

この単一障害点リスクを回避する優れた技術がこれから紹介する「DRBD」のレプリケーションです。


アプリケーション
ファイルシステム
ページキャッシュ
DRBD
(プライマリ)
Rawデバイス
TCP/IP
ディスク
スケジューラ
ディスクドライバ
NICドライバ
ディスク
ネットワークカード
ネットワークカード
TCP/IP
DRBD
(セカンダリ)
ディスク
スケジューラ
NICドライバ
ディスクドライバ
ディスク

DRBDのディスクミラーのメリット

RPO：災害時に復旧できるデータの新しさ

バックアップ：最後のバックアップ時

DRBD：クラッシュ直前

RTO：災害時に復旧するまでの時間

バックアップ：データリストアに時間が必要

DRBD：すぐに復旧できる

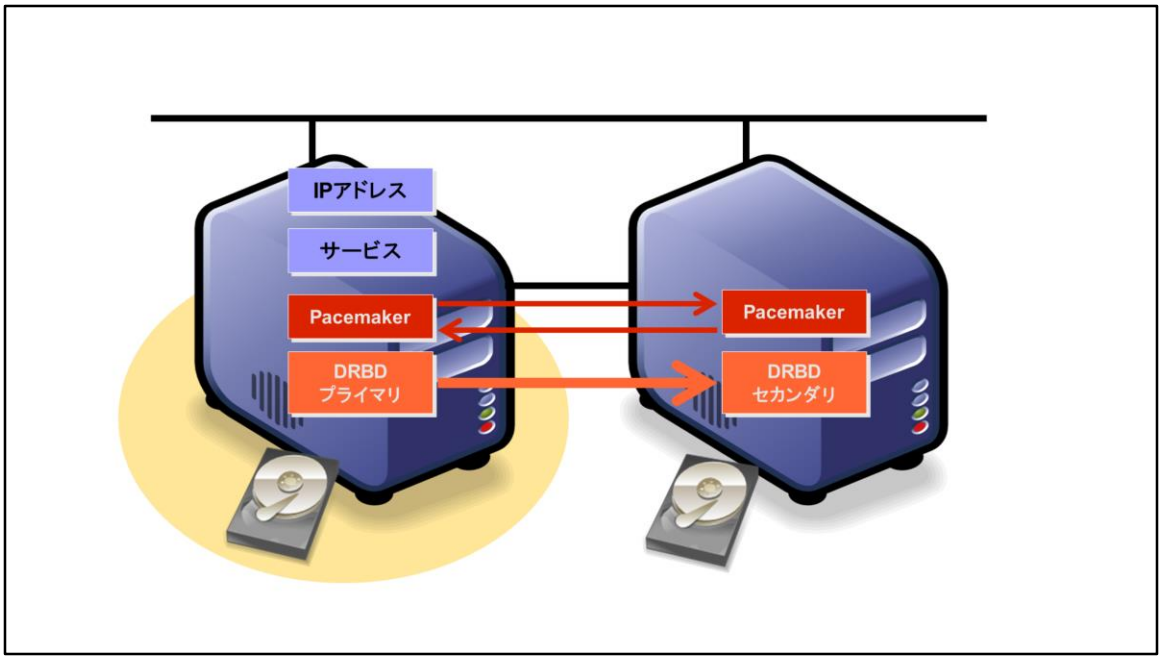
バックアップでは前回のバックアップの実行時のデータに戻ってしまいますが、DRBDの場合、障害が発生する直前のデータを使ってサービスを再開できます。

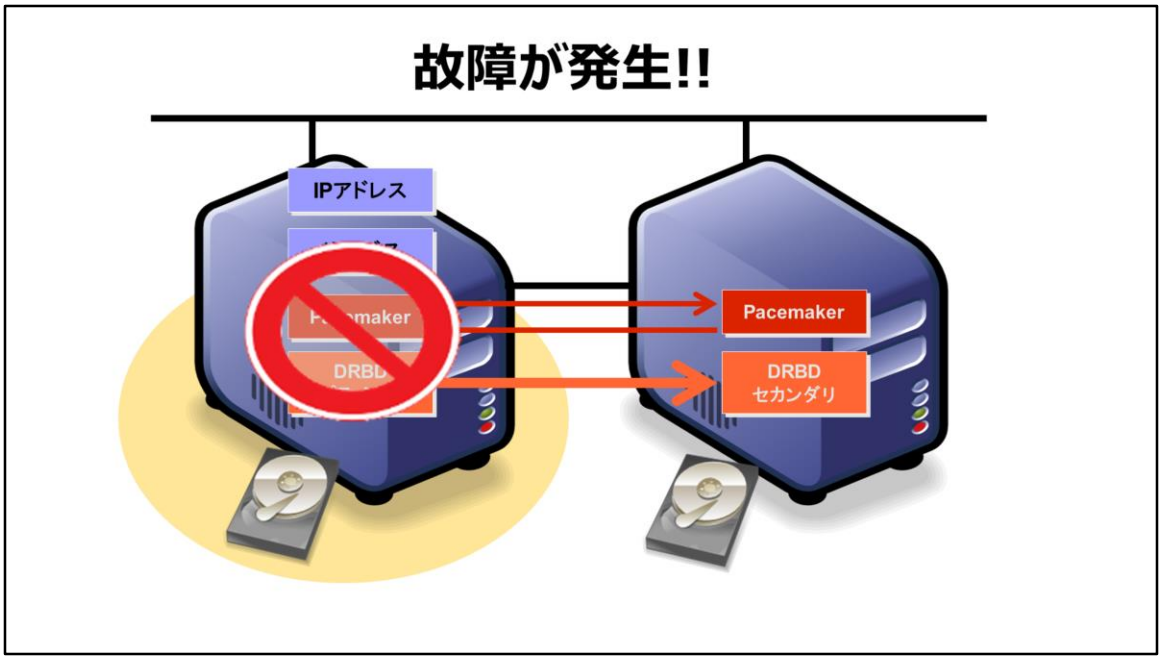
サービスの冗長化

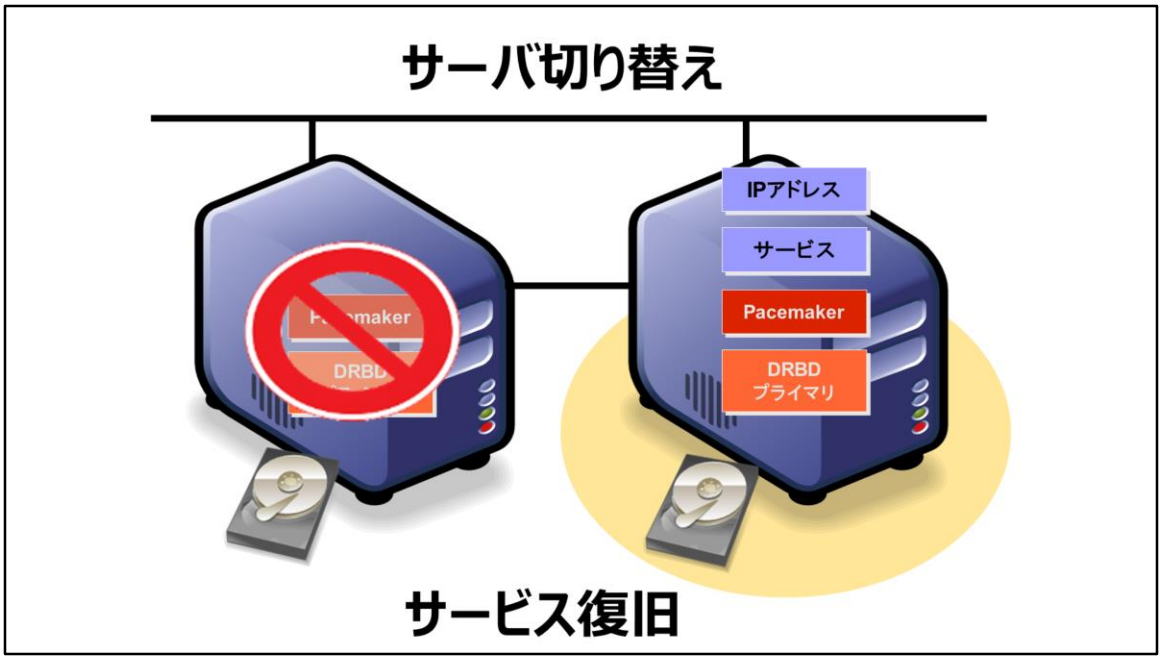
HAシステムでサービスを管理する

データの同期にはDRBDが使えることが判りました。
サービスを監視して、異常生じた時に自動的にサーバを切り替える仕組みがあると便利です。

これはPacemakerというソフトウェアを使うとサービスの管理ができます。


IPアドレス
サービス
Pacemaker
DRBD
プライマリ
Pacemaker
DRBD
セカンダリ


故障が発生!!
IPアドレス
Pacemaker
DRBD
Pacemaker
DRBD
セカンダリ


サーバ切り替え
IPアドレス
サービス
Pacemaker
DRBD
プライマリ
サービス復旧

DRBDとPacemaker、Corosyncを組み合わせると、
アクティブ-スタンバイなHAシステムが構築できます。

この設定例はオンプレミスなサーバでの構築例ですがAWSやAzureなどのクラウド環境でも利用できます。

Pacemaker と DRBD®による
HAクラスターのメリット

OSSなので無料で作れる

数多くの実績

構成がシンプル

LinuxでのHAシステムの構築は・・・

25