

*The King of Cluster Solution*

Linuxによるミッションクリティカルシステムを実現する

***DNCWARE Cluster Perfect***

東芝ITソリューション株式会社

情報機器技術部オープンソフトウェア担当

石井 正悟

# LinuxにおけるHAクラスタの高まり

- サーバOSとしてのLinux市場の成長
- 企業ユーザがLinux採用へ
- 適用分野の広がり
  - 従来のインターネットサーバから徐々にアプリケーションサーバやDBサーバへと適用分野が広がりつつある
- 高信頼性を求めるユーザの変化
  - Φ1 (1996～): 信頼性の高いUNIXクラスタシステム
  - Φ2 (1998～): コスト面からNTクラスタが伸びる
  - Φ3 (2001～): 信頼性とコストの面からLinuxクラスタへ



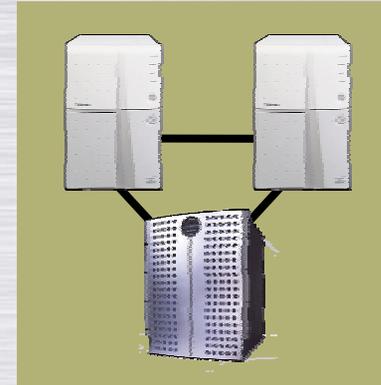
LinuxでのHAクラスタのニーズの増加

# LinuxでのHAクラスタ適用分野

## ■ DBサーバやアプリケーションサーバなどの基幹システム

特長

- 信頼性重視
- 共有ディスク構成が多い

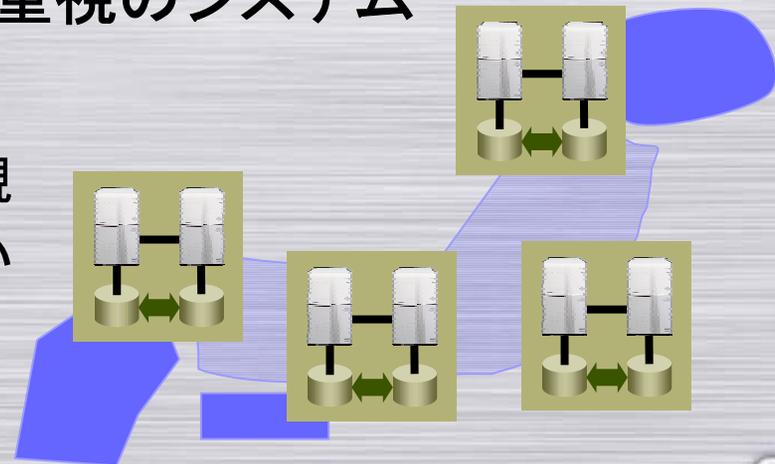


DBシステム

## ■ 大規模展開を行なうコスト重視のシステム

特長

- 大量導入のためコスト重視
- 非共有ディスク構成が多い



# ClusterPerfectの位置づけ

HA クラスタ

*ClusterPerfect*  
(Linux, Solaris, NT)

分散処理クラスタ

OPS, 分散OS, etc.

*ClusterPerfect for OPS*  
(NT)

データレプリケーション

ディスク間でのデータレプリケーション

ネットワーク負荷分散クラスタ

TurboClusterなど

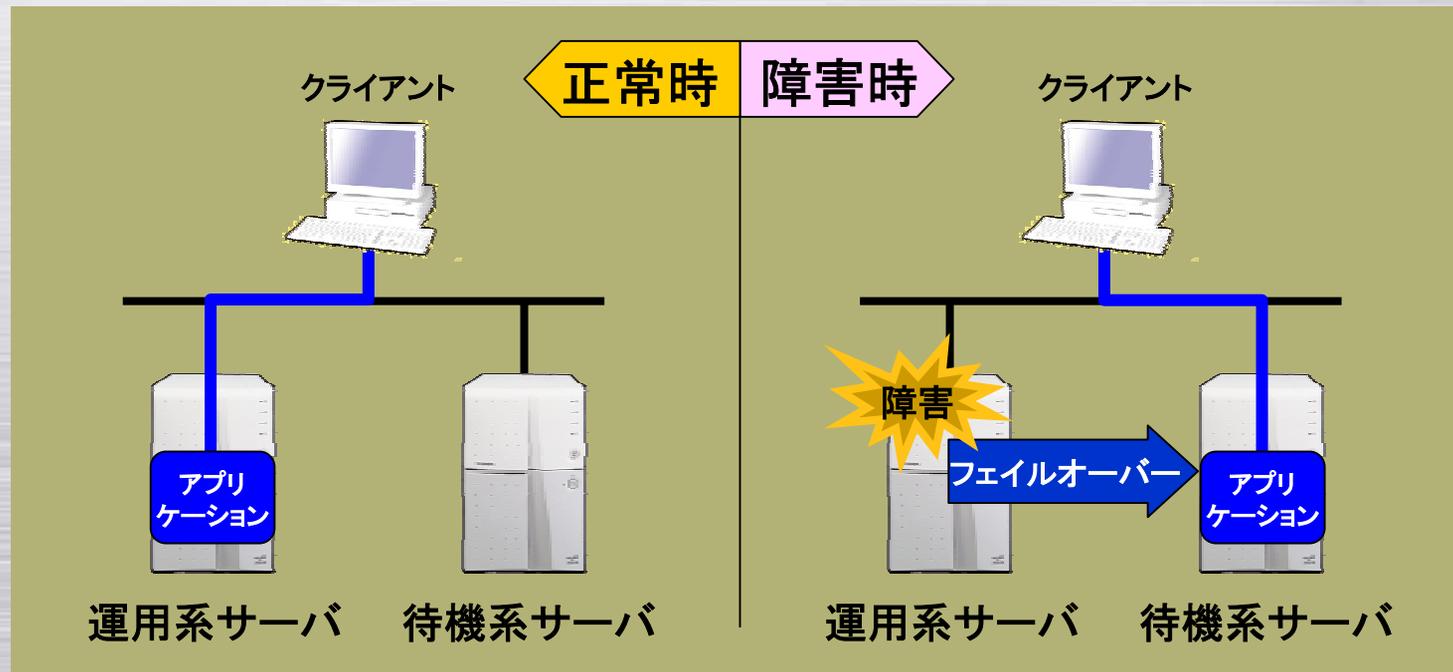
*ClusterPerfect for Web*  
(Linux, Solaris, NT<sub>予定</sub>)

数値計算型クラスタ

アプリケーションごとに設計

# HAクラスタシステムとは

- ネットワーク上で複数のサーバを連携させ、1つのシステムとして利用するシステムです。
- どれか1つのサーバで障害が発生しても、残りのサーバでバックアップして業務を継続することができます。



# クラスタシステムに対する要望

- 簡単、確実にクラスタを組みたい
- フレキシブルに構成を選びたい
- サーバやディスク選択の幅が欲しい
- 既存アプリケーションやミドルウェアをそのまま使いたい



統合クラスタソフトウェア

**DNCWARE** *ClusterPerfect*

# ClusterPerfectの特長

## Flexible

### ■ 様々なニーズに柔軟に対応

2台のHA構成から256台の多ノード構成、共有ディスクの有無、多彩なバックアップ方式など、ニーズに合わせて最適なシステムを自由に選択できます。

## Easy

### ■ 高度なクラスタシステムを手軽に実現

あらかじめシステム定義したテンプレートの利用により、システムを驚くほど簡単に実現できます。

## Open

### ■ マルチベンダ・マルチプラットフォーム対応

オープンアーキテクチャにより、幅広いサーバやディスク選択を可能にします。さらに、既存のアプリケーションやミドルウェアの変更は不要です。

## Scalable

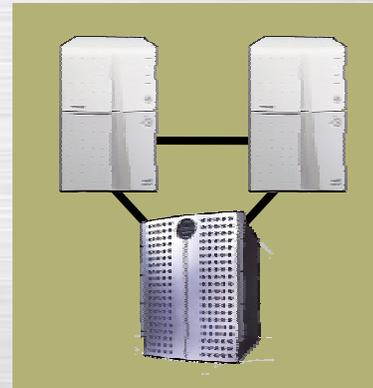
### ■ 最大256台のクラスタ構成

最大256台のサーバによるクラスタ構成が可能であるため、将来の業務拡張にも余裕をもって対応できます。

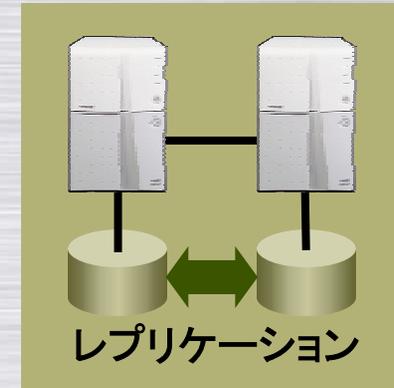
## 業務に合った最適なクラスタシステムを構築

クラスタを構成するノード数、共有ディスクの有無、バックアップ方式、ノード間接続方式など多くの選択肢を用意しています。業務に合わせて最適なクラスタシステムの構築が可能です。

## データ引き継ぎ方式

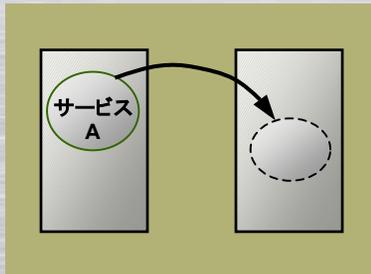


共有ディスク方式

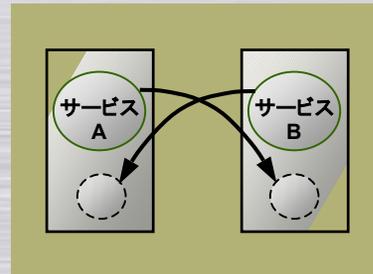


非共有ディスク方式

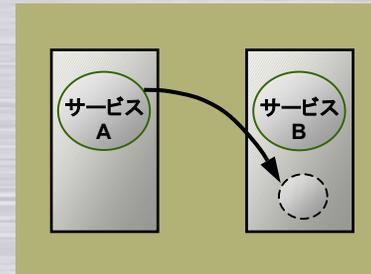
## バックアップ形態例



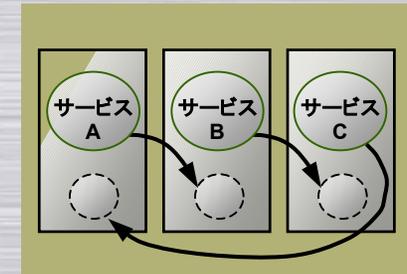
スタンバイ方式



相互バックアップ方式



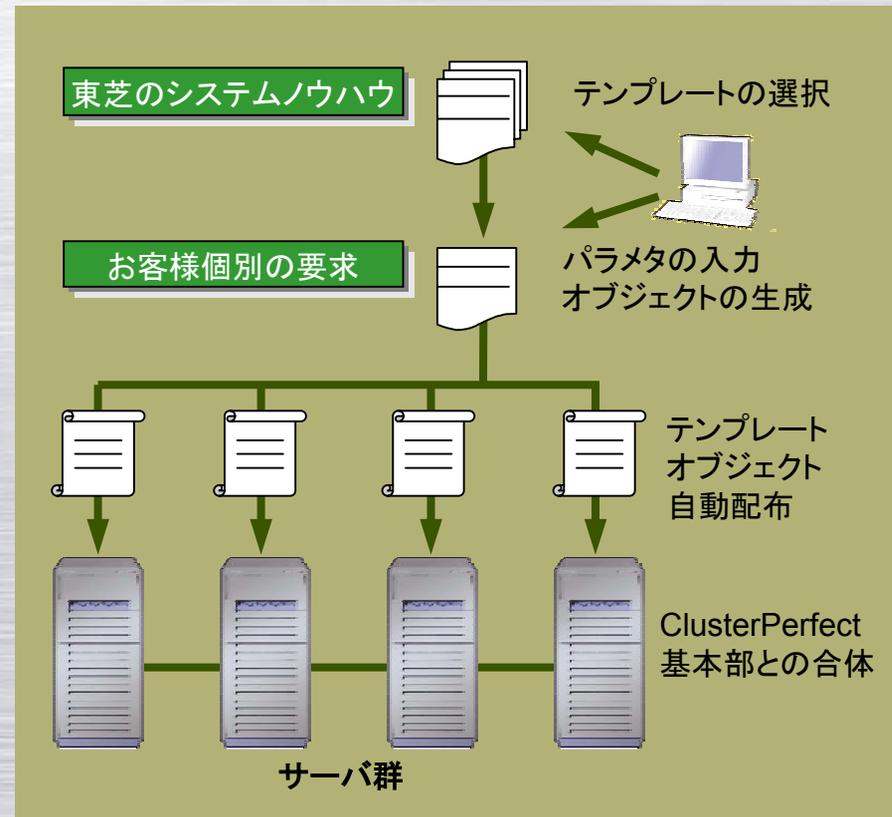
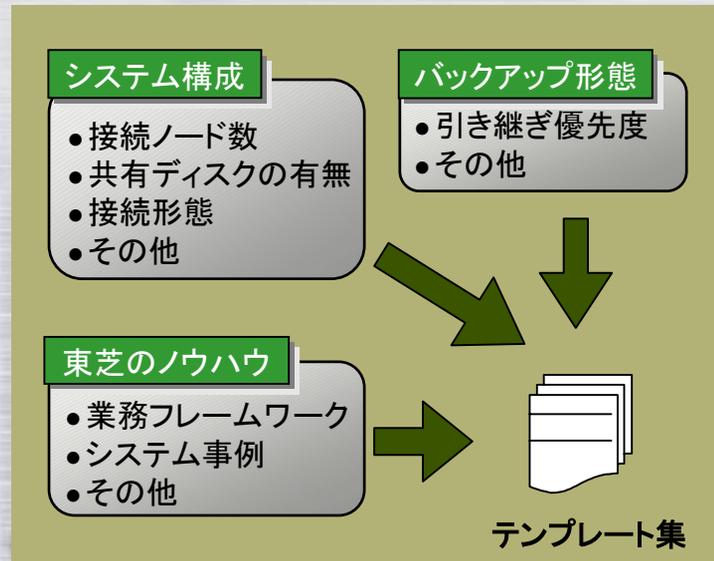
片系バックアップ方式



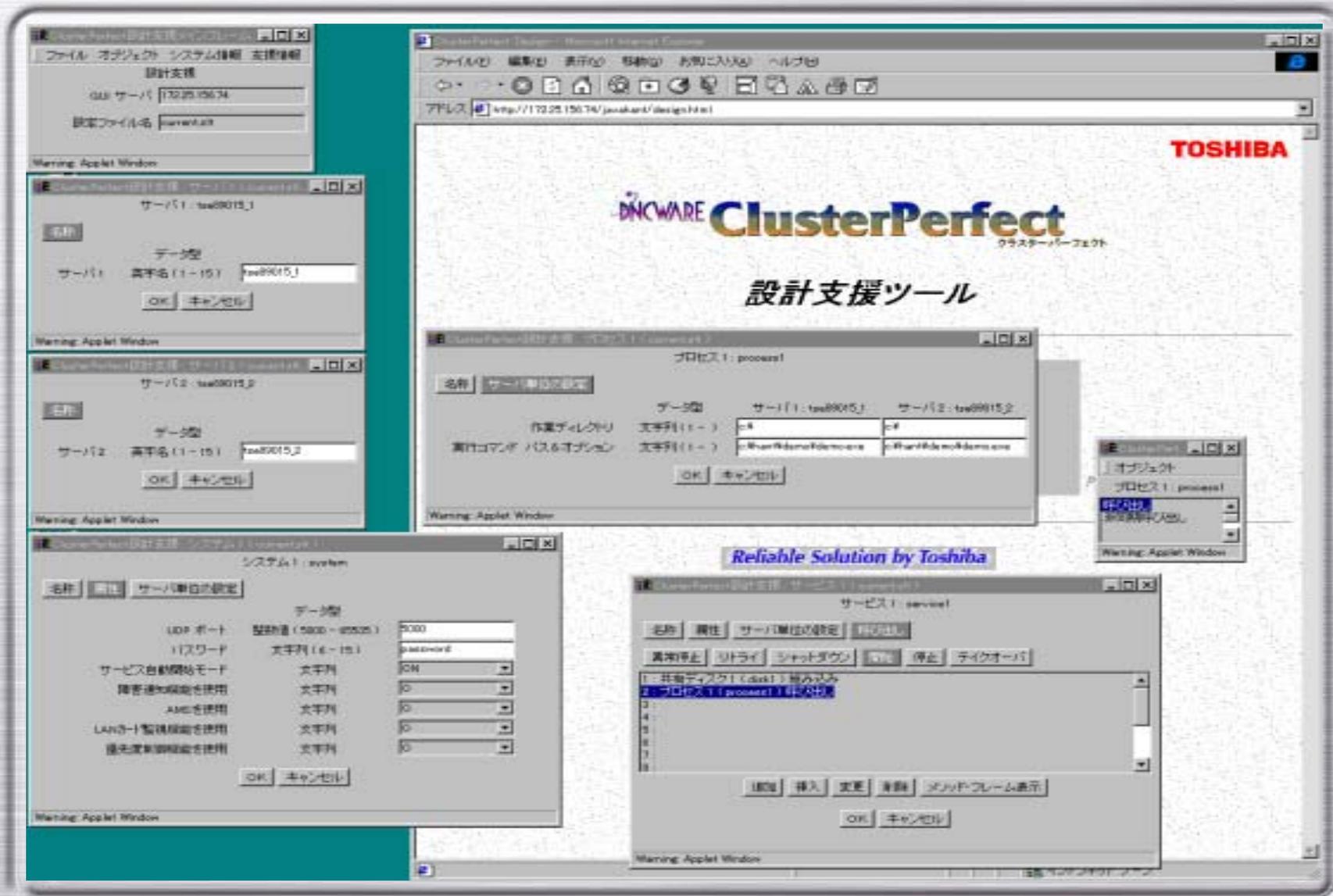
ローテート方式

# テンプレートの利用で簡単 &スピーディにクラスタシステム設計が可能 *Easy*

東芝が長年培ったクラスタのノウハウをテンプレート集として提供します。あらかじめシステム構成が定義されているテンプレートの中から業務にあったクラスタ構成を選択し、IPアドレスなどのいくつかの項目を埋めるだけで簡単に設定が行なえます。



## ブラウザとJavaによる設計支援ツール



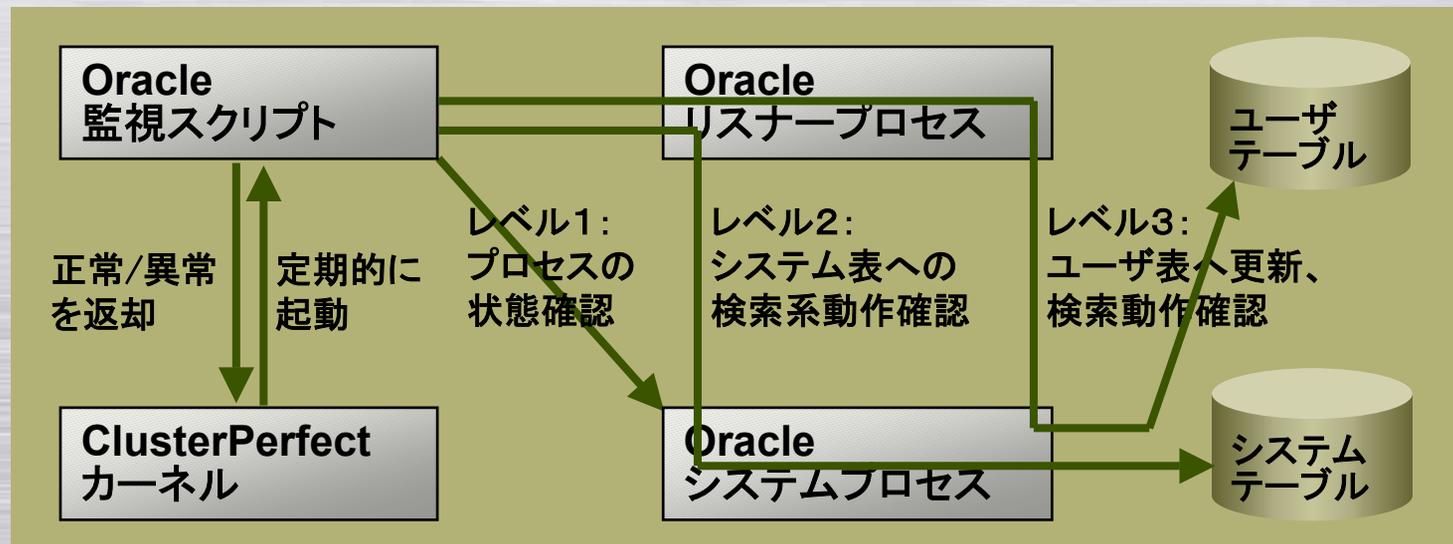
※本内容は変更する場合があります。  
詳細は、仕様書あるいは説明書をご覧ください。

## アプリケーション、ミドルウェアの変更不要

アプリケーション、ミドルウェアの状態監視を、プログラム外部の監視スクリプトで行ないます。

プログラム内部にクラスタ用APIを埋め込む必要がないので、クラスタシステム構築のためにプログラムを変更する必要は一切ありません。また、OracleやLANカードなどの障害を監視するための監視スクリプトをあらかじめ用意しています。

## 【Oracle監視スクリプトの例】



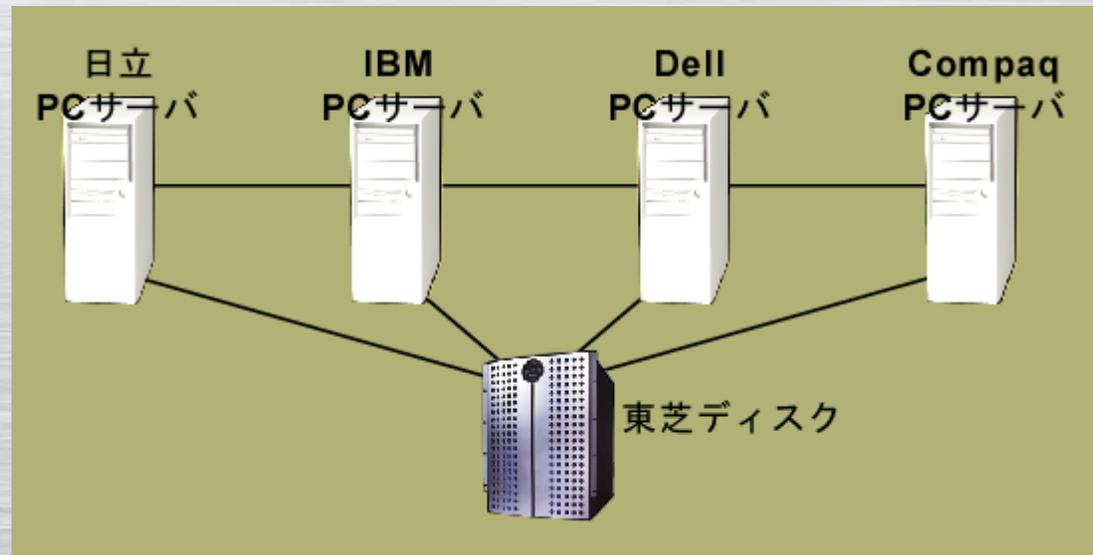
## マルチベンダ・マルチプラットフォーム対応

徹底したオープン性の追求により、複数メーカーのハードウェアによるクラスタ構成が可能です。またマルチプラットフォーム環境においても、統一したオペレーションによるクラスタ管理が可能です。

## 【マルチベンダによるクラスタシステム例】

DNCWARE ClusterPerfect  
対応プラットフォーム

Windows NT サーバ4.0  
Windows 2000  
Solaris  
Solaris Intelプラットフォーム  
Linux



# ClusterPerfectの最大の特徴 一経験と実績一

1970s 1980s 1990~ 1994 ~ 1996 1997 ~ 1999 2000~

密結合クラスタ(共有メモリ・ディスク)

4ノード

8ノード

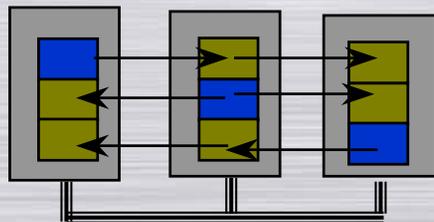
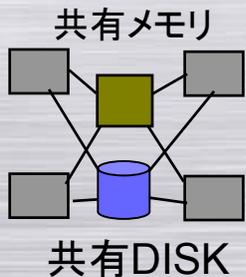
**T-7/70 G8000 VL2000**



疎結合クラスタ  
最大16ノード  
**G8000**

**UNIX/PCサーバ複合系**

UXシリーズ (Solaris)  
GSシリーズ (WinNT)  
最大2ノード  
共有ディスク



**DNCWARE for HA**  
**DNCWARE for OPS**

最大4ノード  
スタンバイ型 (HA)  
分散DB型 (OPS)

**ClusterPerfect**

最大256ノード  
高速切替  
レプリケーション  
負荷分散, etc

※本内容は変更する場合があります。  
詳細は、仕様書あるいは説明書をご覧ください。

# ClusterPerfectの安全性

ClusterPerfectでは、Split Brain※やデータ破壊などを防ぐための工夫がきちんと考慮されています。

## ① ハートビートパスの多重化

ハートビートを行なう通信LANは、単一故障を配慮して多重化を実施。  
ハートビートに使用するネットワークの障害で引き起こされるSplitBrainを防止。

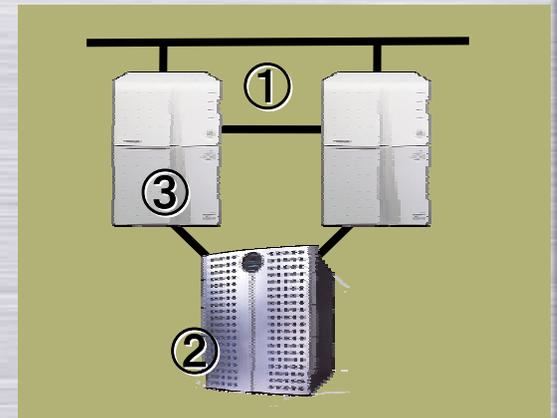
## ② 共有ディスクのH/Wレベルでの排他機構

SCSIのリザーブ/リリースコマンドにより、H/W的に排他を実施。

## ③ 強制パニック

万一、共有ディスクが他系にリザーブを奪われた場合は、  
自系を強制的にパニックさせて正常系の処理に悪影響を  
与えない工夫を実施。

※クラスタ構成で他系の状態がわからなくなった状態。  
両ノードが自分が稼働系であると判断して両ノードでサービスを起動してしまう。



# ClusterPerfect(Linux版)事例1

## 某ECサイト様向け

### ■使用目的

ECサイトの顧客情報管理システム

### ■ポイント

- Linuxを用いたハイエンドDBシステム
- 24H365D稼動

### ■システム構成

- IAサーバ  
他社製
- 共有ディスク  
東芝ArrayFort AF1500
- DB  
Oracle8i
- OS  
TurboLinux Server 6.1

インターネット



# ClusterPerfect(Linux版)事例2

## 某公共システム向け料金収集システム

### ■使用目的

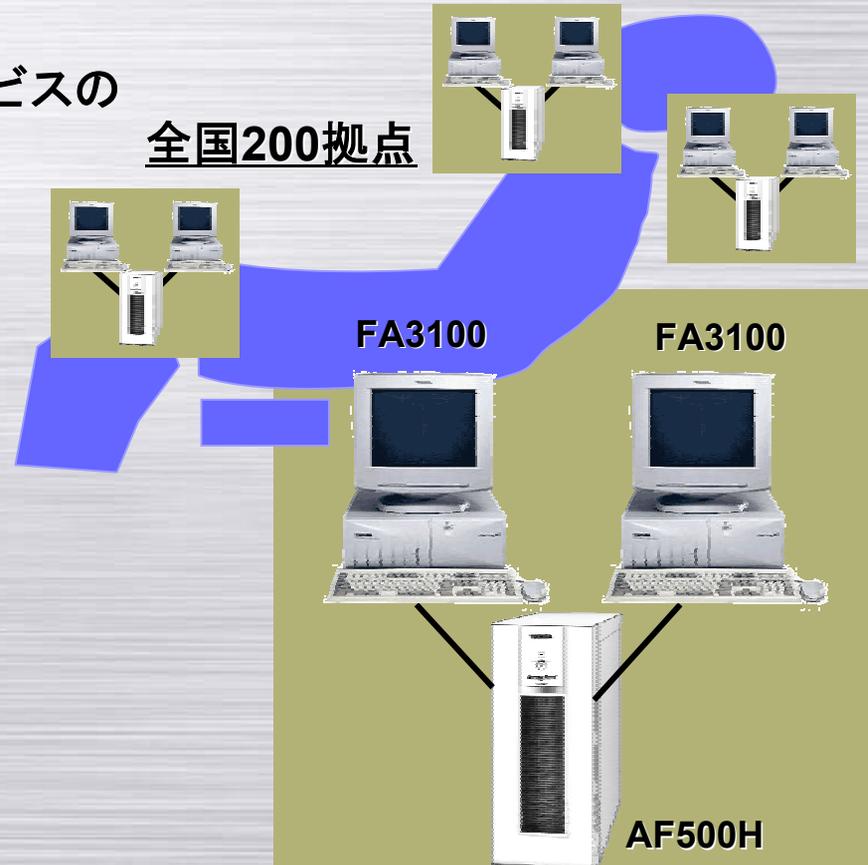
全国拠点を発生する公共サービスの  
利用料金収集システム

### ■ポイント

- 全国200拠点到展開
- ローコストなHAソリューション
- 24H365D稼働

### ■システム構成

- IAサーバ  
東芝産業用コンピュータFA3100
- 共有ディスク  
東芝ArrayFort AF500H
- OS  
TurboLinux Server 6.1



# ClusterPerfectシリーズ商品構成

<b>DNCWARE ClusterPerfect</b>		Linux	Solaris	WinNT /2000	クラスタ基本パッケージ
各監視 オプション	<b>LAN監視オプション</b>	Linux	Solaris	WinNT /2000	
	<b>Oracle監視オプション</b>	Linux	Solaris	WinNT /2000	
	<b>Notes監視オプション</b>			WinNT /2000	
	<b>SQL Server監視オプション</b>			WinNT /2000	
	<b>Apache監視オプション</b>	Linux			
特定ミドル ウェア用専用 パッケージ  2ノード分ClusterPerfectと 監視オプションをパッケージ	<b>DNCWARE ClusterPerfect for Oracle</b>			WinNT /2000	Oracle専用2ノードクラスタパッケージ
	<b>DNCWARE ClusterPerfect for Notes</b>			WinNT /2000	Lotus Notes専用2ノードクラスタパッケージ
	<b>DNCWARE ClusterPerfect for SQL Server</b>			WinNT /2000	SQL Server専用2ノードクラスタパッケージ
<b>DNCWARE ClusterPerfect for Oracle Quick Recovery</b>			Solaris	WinNT	Oracle専用2ノードクラスタパッケージ Oracle高速リカバリ機能付き
<b>DNCWARE ClusterPerfect for OPS</b>				WinNT	Oracleパラレルサーバ
<b>DNCWARE ClusterPerfect for Web</b>		Linux	Solaris		Web負荷分散クラスタ

※本内容は変更する場合があります。  
詳細は、仕様書あるいは説明書をご覧ください。

# LinuxでHAクラスタを開発することのメリット ー1

## ■ 確実な方法で障害検出可能

障害発生を早く、確実に検出する必要がある。

誤検出は全系停止に成り得る大問題。

オープン・ソースなのでカーネルの動作をソースコードから確認して確実な障害検出方法を採用可能。

## ■ 商用UNIXからの移植が容易

POSIX準拠なので、マルチプラットフォーム対応のソフトウェアならば殆ど再コンパイルでok。

商用UNIXでの開発knowhowが活かされる。

# LinuxでHAクラスタを開発することのメリット ー2

## OS種別によるトラブル・シューティングの違い

### ■プロプライエタリOS

ソース・コードからの解析が可能

迅速なパッチ対応が可能

### ■商用OS

ブラックボックスなので真の原因追求が困難

回避策を提示し, OSメーカーによるパッチ配布を待つ

開発元にエスカレーションしても却下の可能性あり

### ■Linux

ソース・コードからの解析が可能

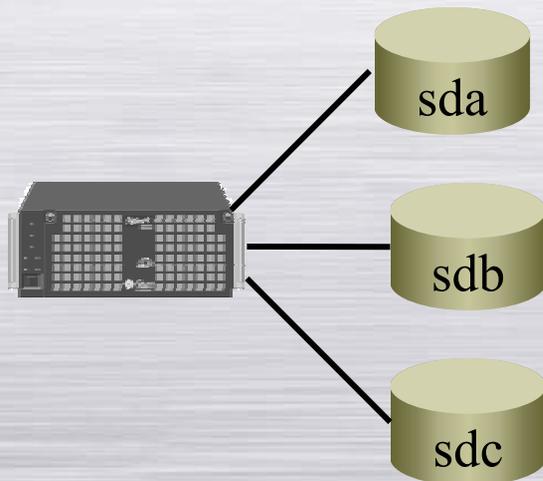
迅速なパッチ対応が可能

コミュニティからの迅速なパッチ公開

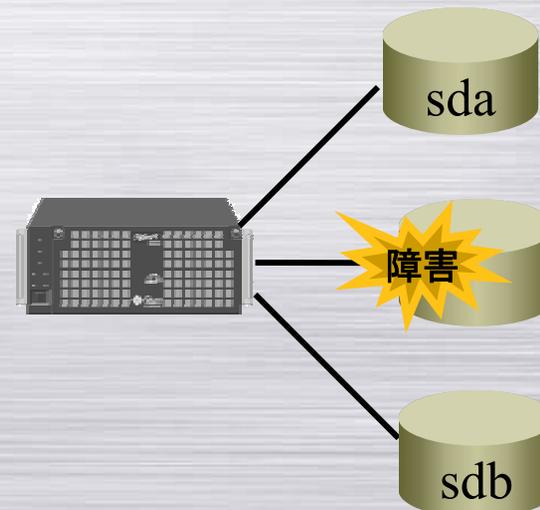
# LinuxでHAクラスタを開発する際の技術的課題

- デバイスを一意に認識する手段がない  
デバイス名がboot時の認識順番によって決定される。

正常時



/dev/sdbに障害発生



☆ devfs のディストリビューションへの取り込みに期待。

# Linuxは本当にコストパフォーマンスが高い？

## 問題提起

### ■導入時

安価なハードウェアでも高性能である。

開発コストの大半がハードとソフト・ライセンス料のシステムならば非常にコストパフォーマンスが高い。

大規模なアプリ開発を伴うシステムでは、OSの違いよりも開発手法・開発ツールの違いの方がコストへの影響は大きい。

### ■運用時

OSを代えることでどの部分のコストが下がる？

# Linuxに対する勘違い

- Linuxシステム構築費は安い  
システム構築費用に占めるOSのライセンス料は僅か
- 旧バージョンでも保守可能  
基幹システムではOSをバージョンアップしたくない。  
枝分かれバージョンのソース・コードを自分で管理する必要がある。  
ソース・コードを解析できる技術者が必要
- サポートも安い  
解析方法, 解析結果の信憑性, 解決までのリードタイムなどの違いはあっても, 解析, 対策の検証, ユーザへの報告など労力は変わらない。

# 信頼性の高さで基幹システムにLinuxを選択

## ■コストパフォーマンス

導入時コストがTCOの大半を占めるシステムではLinuxを選択することで高いコストパフォーマンスが得られる。  
例えば、クローンシステムを全国展開する場合。

## ■信頼性の高さ

24H365日の連続運用が可能。  
(週に一度の reboot は不要)

Linuxならば真の原因を追求し、確実に修正できる。  
コミュニティからの迅速なパッチ公開。

## まとめ

- ClusterPerfect は、多くの実績のある信頼性の高いクラスタウェアです。
- 操作性、運用性のすぐれたクラスタウェアです。

- 
- Linuxは、Linus Torvalds氏の米国及びその他の国における登録商標または商標です。
  - TurboLinuxは、TurboLinux,Inc.の商標です。
  - Solaris,Java、およびすべてのJava関連の商標及びロゴは、米国およびその他の国における米国Sun Microsystems,inc.の商標または登録商標です。
  - Oracleは、Oracle Corporationの登録商標です。
  - Intelは、Intel Corporationの商標です。
  - Windows,Windows NT,SQL Serverは、米国Microsoft Corporationの米国およびその他の国における登録商標または商標です。
  - Lotus Notesは、Lotus Development Corporationの登録商標です。
  - 本資料に掲載の商品の名称は、それぞれ各社が商標として使用している場合があります。