

# ZFS オープンストレージ OS NexentaStor 及びアプライアンス製品のご紹介

= Contents =

1. Data & Topics
2. NexentaStor 概要
3. ZFS の特徴
4. NexentaStor プラグイン
5. 機能比較
6. アプライアンス製品
  - PrimeSTOR ZFS
  - PrimeGATE ZFS

コアマイクロシステムズ株式会社

平成 22年 10月 11日



## Data & Topics

---



# コアマイクロシステムズ次期戦略・新製品発表会



## コアマイクロシステムズ 次期戦略・新製品発表会 開催レポート

2009年11月13日開催 於:丸ビル ホール&コンファレンススクエア



### ■開催概要

名称 :コアマイクロシステムズ 次期戦略・新製品発表会

主催 :コアマイクロシステムズ株式会社

協賛 :Nexenta Systems, Inc.



日時 :2009年11月13日(金曜日) 14時~16時20分(開場 13時30分)

会場 :丸ビル ホール&コンファレンススクエア ルーム#5

対象 :ソリューションプロバイダー様

開催当日は100名近い方々にご参加を賜り、盛会にて「コアマイクロシステムズ 次期戦略・新製品発表会」を無事、終えることができました。これもひとえに、日頃ご支援を頂いているパートナー様お客様の御蔭であると深く感謝申し上げます。本当にありがとうございました。

開催日当日は、フラグシップ製品であるSolid STOR(業界最速水準30万IOPSを実現するDRAM型半導体ストレージ)の技術を高度に発展させた、B-TREEフル連想型大容量キャッシュ装置「Super CACHE」及び応用装置である「Super SSD」を発表しました。また、次世代データセンタにおける仮想化環境のトータルQoSを効果的にかつ効率的に実現する革新のストレージ自動階層化/仮想化技術ZFSをベースとしたNexentaStor オープンストレージ・ソリューション「Prime STOR ZFS」及び「Prime GATE ZFS」の解説を行いました。営業・マーケティング戦略としては「Solid STOR」、「Super CACHE」、「Super SSD」などの「スーパーストレージ」と「Primeシリーズ」、「Data SAFE」を重点製品と位置づけ、それぞれのターゲットとビジネスチャンス进行解説しました。

当日の資料ダウンロード> :<http://www.cmsinc.co.jp/seminar/20091113.html>



“ストレージ・ソリューションのリーディング・プロバイダ” コアマイクロシステムズ株式会社

Copyright © Core Micro Systems Inc., All rights reserved.

# IaaSトップベンダー「ITコア」様採用事例発表

## ITコアの全層仮想化IaaS「GrowServer2010-II」に仮想マシンの高速かつ高効率バックアップ用ZFSアプライアンスとして「PrimeSTOR ZFS」を提供

コアマイクロシステムズ株式会社(本社:東京都板橋区、代表取締役社長:高橋晶三、以下コアマイクロシステムズ)は業界初のプライマリストレージとしてインライン重複排除機能を実装した「PrimeSTOR ZFS」(NexentaStor 3.0を搭載したストレージ装置)が、IaaSのリーディングカンパニーである株式会社ITコア(本社:東京都中央区、代表取締役社長山田敏博、以下ITコア)全層仮想化クラウドサービス「GrowServer2010-II」に採用されたことを2010年7月27日に発表しました。

「GrowServer2010-II」は完全に仮想化された国内初の本格的IaaSです。「GrowServer2010-II」の仮想マシンを高速かつ高効率バックアップするためのストレージを高品質かつ低コストで構築するべく、ITコアはPrime STOR ZFSを採用しました。

株式会社ITコア 代表取締役社長 山田敏博氏は「コアマイクロシステムズのPrime STOR ZFSを採用したことにより、低価格で高品質なZFSアプライアンスを利用できました。」と述べています。



### 企業発表

#### コアマイクロシステムズ、ITコアがストレージ製品「PrimeSTOR ZFS」を採用

2010年7月27日

コアマイクロシステムズ株式会社

コアマイクロシステムズ、ITコアの全層仮想化IaaS「GrowServer2010-II」に仮想マシンの高速かつ高効率バックアップ用ZFSアプライアンスとして「PrimeSTOR ZFS」を提供

コアマイクロシステムズ株式会社(本社:東京都板橋区、代表取締役社長:高橋晶三、以下コアマイクロシステムズ)は業界初のプライマリストレージとしてインライン重複排除機能を実装した「PrimeSTOR ZFS」(NexentaStor 3.0を搭載したストレージ装置)が、IaaSのリーディングカンパニーである株式会社ITコア(本社:東京都中央区、代表取締役社長 山田敏博、以下ITコア)全層仮想化クラウドサービス「GrowServer2010-II」に採用されたことを発表しました。

「GrowServer2010-II」は完全に仮想化された国内初の本格的IaaSです。

「GrowServer2010-II」の仮想マシンを高速かつ高効率バックアップするためのストレージを高品質かつ低コストで構築するべく、ITコアはPrime STOR ZFSを採用しました。

**報道発表原文** > <http://www.cmsinc.co.jp/news/2010/news201007271.html>



「ストレージ・ソリューションのリーディング・プロバイダ」 コアマイクロシステムズ株式会社

Copyright © Core Micro Systems Inc., All rights reserved.



# Prime STOR ZFS ENTERPRISEの発表

強力なストレージ仮想化機能と半導体ストレージ「Super SSD」により、  
オンデマンドでの性能強化と柔軟な容量拡張を実現した  
「Prime STOR ZFS ENTERPRISE」を10月1日より発売開始

コアマイクロシステムズ株式会社(本社:東京都板橋区、代表取締役社長:高橋晶三、以下 コアマイクロシステムズ)は近年急速に進むサーバ仮想化とストレージ効率化のニーズに対応する革新のストレージシステムPrime STOR ZFS ENTERPRISEを2010年10月1日より発売開始することを本日発表いたしました。本ストレージシステムは、ZFSベースの強力なストレージ仮想化機能と自社独自開発の半導体ストレージ「Super SSD」によるオンデマンドでの性能強化と柔軟な容量拡張に対応し、仮想統合環境や各種のエンタープライズアプリケーションの将来的なストレージニーズに的確に応えます。コアマイクロシステムズはPrime STOR ZFS ENTERPRISEにより初年度2億円の売り上げ増を見込んでいます。

本ニュースは、キーマンズネット様、インターネットコム様、毎日新聞様、Yahoo!ニュース様、infoseekニュース様、livedoorニュース様、@Niftyニュース様、Exciteニュース様、msnニュース様など、9メジャーメディアに掲載されています。

報道発表原文>

<http://www.cmsinc.co.jp/news/2010/news201007271.html>



“ストレージ・ソリューションのリーディング・プロバイダ” コアマイクロシステムズ株式会社

Copyright © Core Micro Systems Inc., All rights reserved.

## ■ 特長

- ・SAN(FC、iSCSI)、NAS(CIFS、NFS)フルユニファイドストレージ
- ・シンプロビジョニング/ダイナミックストレージプール仮想ストレージ
- ・完全に自動化されたTiered Storage アーキテクチャ
- ・圧縮に加えてインラインブロックレベル重複排除機能
- ・オンラインでの無制限の容量拡張に対応
- ・サイレントデータコラプションの検出と自動データ修復
- ・HDD3台の同時障害にも耐えるRAID7に対応
- ・無制限スナップショット/クローニングによる自律CDP
- ・VMware/Xen/Hyper-V へのビジュアルプロビジョニング
- ・完全に冗長化されたクロスミラーHAクラスタ構造
- ・高速/高効率ブロックレベルのリモートレプリケーション

## ■ 仕様

- ・ホストプロトコル FCP、iSCSI、CIFS、NFS、FTP、WebDAV
- ・ホストインターフェース 8G FC、10GbE、GbE
- ・L1キャッシュ/ノード 24GB~(最大512GB)
- ・L2キャッシュ/ノード 400GB~(最大28TB)
- ・書き込み同期キャッシュ 50GB~(最大1TB)
- ・ストレージプール容量(物理) 96TB~(容量無制限)
- ・ハイブリッドRAID 10、50、60、70、11、51、61
- ・最大IOPS(アクティブ、アクティブ時) 100万回以上/4KBブロック
- ・最大スループット(アクティブ、アクティブ時) 5GB/S以上/128KBブロック
- ・サイズ 16U~

## ■ 価格

- ・96TBモデル~ オープンブライズ
- ※価格は、保守レベル、構成で異なります。

## ■ 保守形態

- ・スタンダード 平日/9時~17時 対応
- ・ミッションクリティカル 365日/24時間 対応
- ・データセンタ向け FLEX保守



# テクニカルコラム「かんたん♪ ZFS」

コアマイクロシステムズでは、弊社Prime STOR ZFSやPrime GATE ZFS、NexentaStorなどのZFS関連製品のご採用メリットをよりご理解いただくべく、ZFS関連のテクニカルコラムを連載しています。なお、当コラムはメールマガジン「CNL」に掲載しております。定期的にご愛読をされる場合は、CNLのご登録をお勧めいたします。

## 概要

対象:ZFSにご興味がある方々全般

執筆者:コアマイクロシステムズ株式会社 高田 浩生

掲載周期:不定期(月一回を目安としています)

URL: [http://www.cmsinc.co.jp/mm/mailmagazine/nexenta\\_index.html](http://www.cmsinc.co.jp/mm/mailmagazine/nexenta_index.html)

## バックナンバー

2010/10/01 第十回: DTrace

2010/07/29 第九回: システムモニタ

2010/06/30 第八回: 増分データの送信

2010/05/12 第七回: ZFS データの送受信

2010/04/15 第六回: スナップショット

2010/03/19 第五回: ZFS ファイルシステム

2010/02/25 第四回: ZFS のインターフェース

2010/01/01 第三回: OpenSolaris / Nexenta

2009/11/18 第二回: 根久善太

2009/10/22 第一回: ZFSの特徴





# NexentaStor

---



“ストレージ・ソリューションのリーディング・プロバイダ” コアマイクロシステムズ株式会社  
Copyright © Core Micro Systems Inc., All rights reserved.

# Nexenta Stor とは?

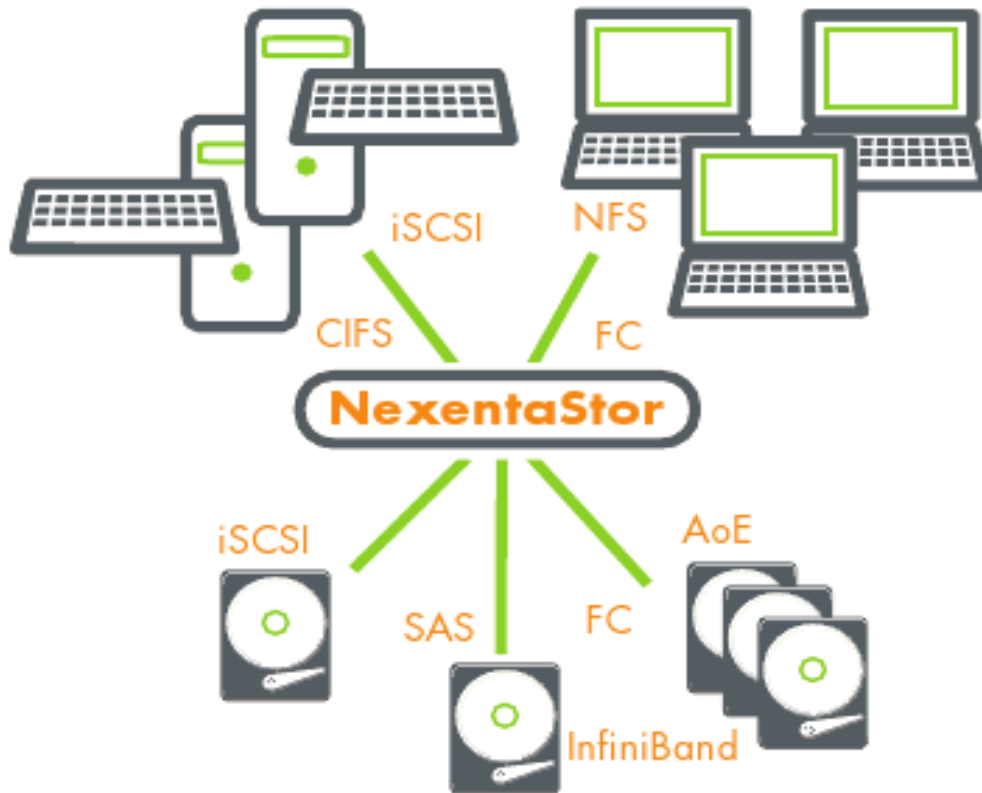
統合ストレージ環境 : block and file

オープンストレージOSの先鋭  
IA サーバハードウェアで動作

OpenSolaris / ZFS によりエンタープライズストレージシステムを従来の三割程度のコストで実現

- 強力なデータ保護機能
- ファイルサイズ無制限
- スナップショット無制限

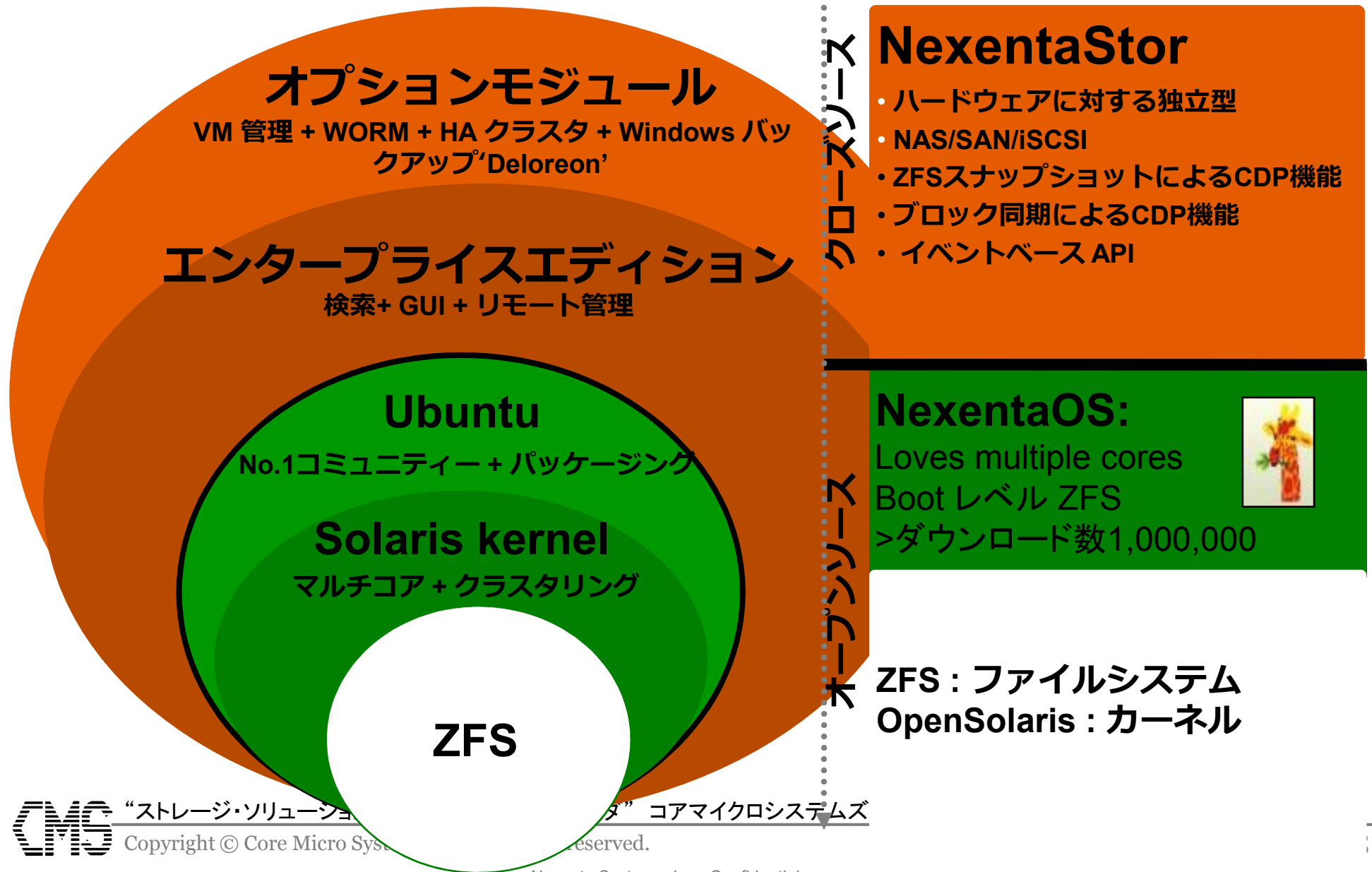
仮想環境対応ストレージ



“ストレージ・ソリューションのリーディング・プロバイダ” コアマイクロシステムズ株式会社

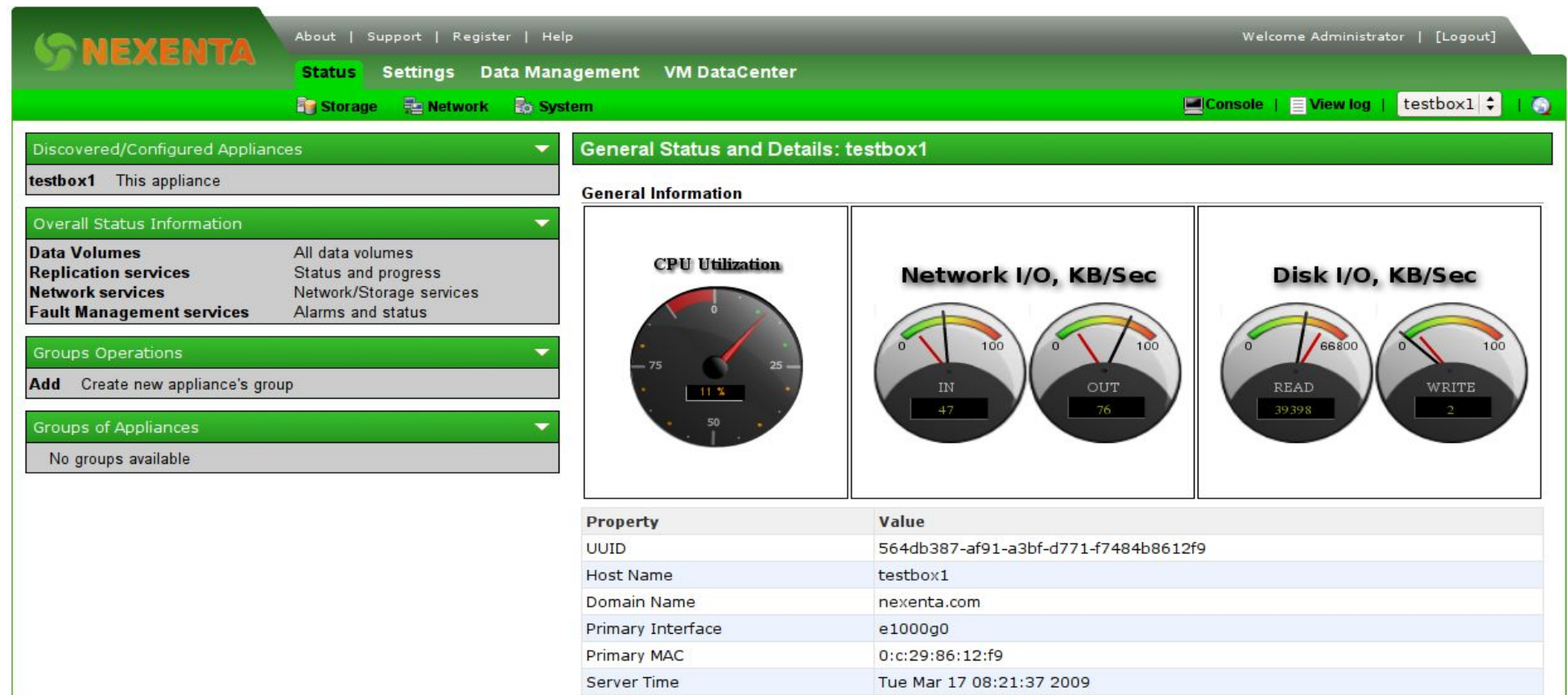
Copyright © Core Micro Systems Inc., All rights reserved.

# オープンストレージOS構造





# 管理画面イメージ



# 顧客とパートナー



## •Resellers:



## •Partners:



“ストレージ・ソリューションのリーディング・プロバイダ” コアマイクロシステムズ株式会社

Copyright © Core Micro Systems Inc., All rights reserved.



## ZFS の特徴

# 従来のファイルシステムとの違い

1. ストレージプール
  - 物理デバイス固定割り当てからの解放
  - ハードディスク、ストレージユニットを追加するだけで使用領域が増える
  - ストライプ幅自動調整 RAID
  - 管理の容易さ
2. トランザクションファイルシステム
  - 常にデータの一貫性が保たれる
  - 障害耐性が高い
3. 自己修復データ保護機構
  - チェックサム
4. スケーラビリティ
  - 128 bit ファイルシステム
5. ポータビリティ



Storage Pool



# ZFS の構造

## 1. Interface Layer

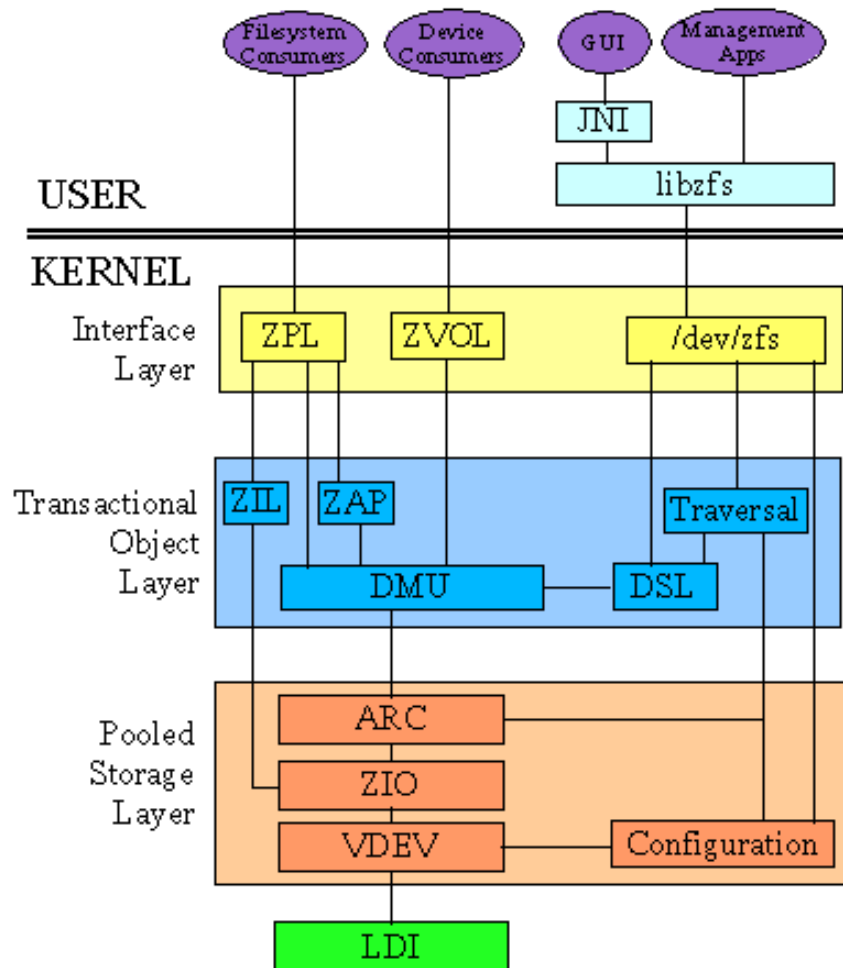
- POSIX インターフェース
- ボリュームインターフェース

## 2. Transaction Object Layer

- DMU オブジェクト
- トランザクション
- ZIL
- スナップショット

## 3. Pooled Storage Layer

- I/O パイプライン
- キャッシュ (ARC)
- ボリュームマネージャ





# Interface Layer

## 1. ZPL (ZFS POSIX Layer)

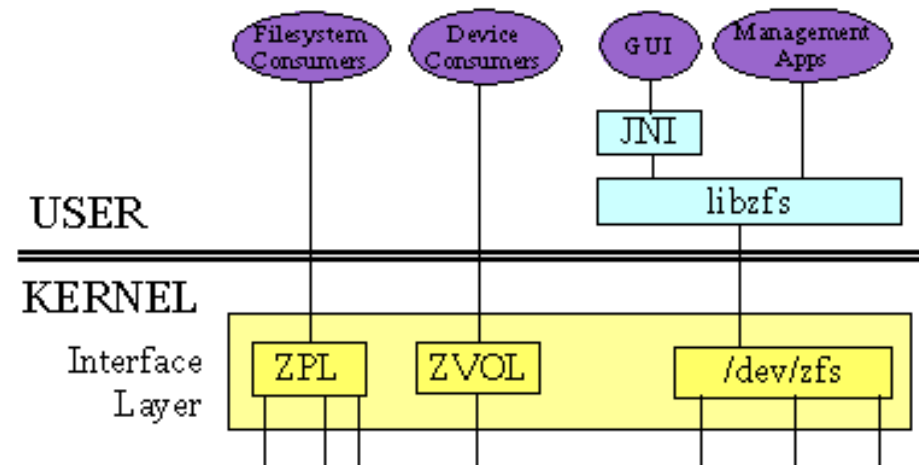
- VFS インターフェース ( mount、umount)
- vnode インターフェース ( open()、read()、write()、mmap() 、 fsync() )
- ACL

## 2. ZVOL (ZFS Volume)

- ボリュームインターフェース (/dev/zvol/[r]dsk/....)
- ZPL とは別に vnodeops を定義

## 3. /dev/zfs

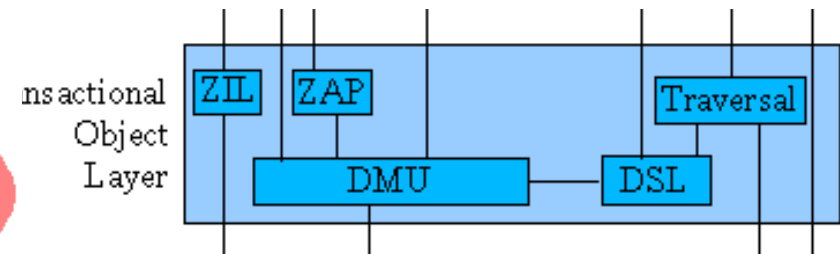
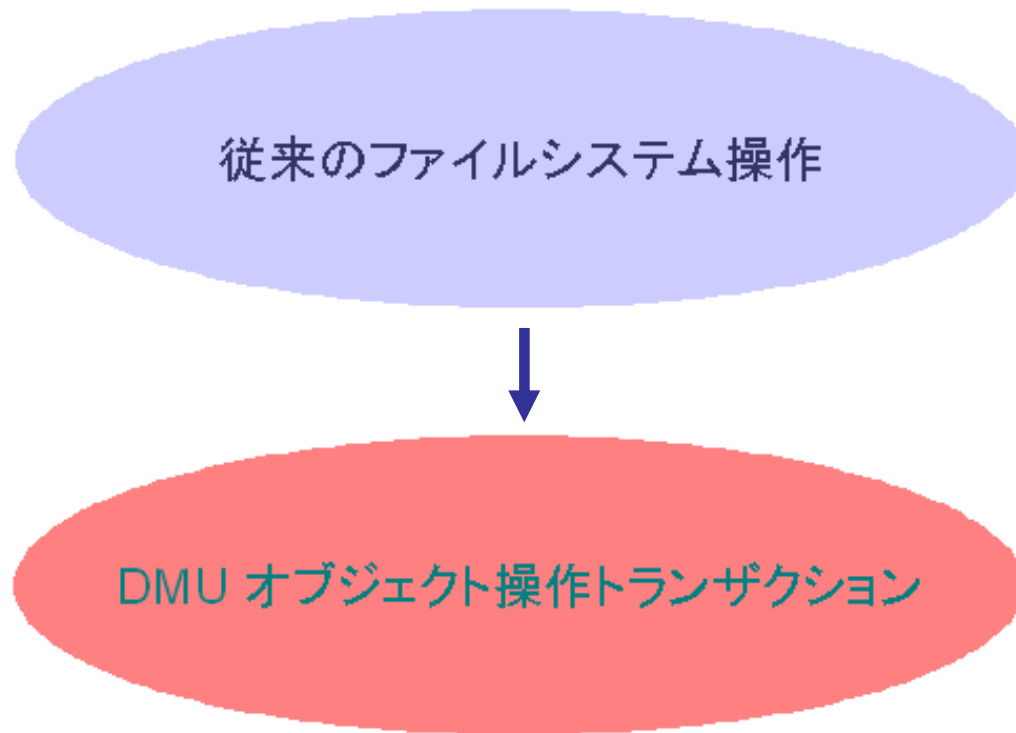
- ioctl() による ZFS の制御



# Transaction Object Layer

## 1. オブジェクトベースのファイルシステム

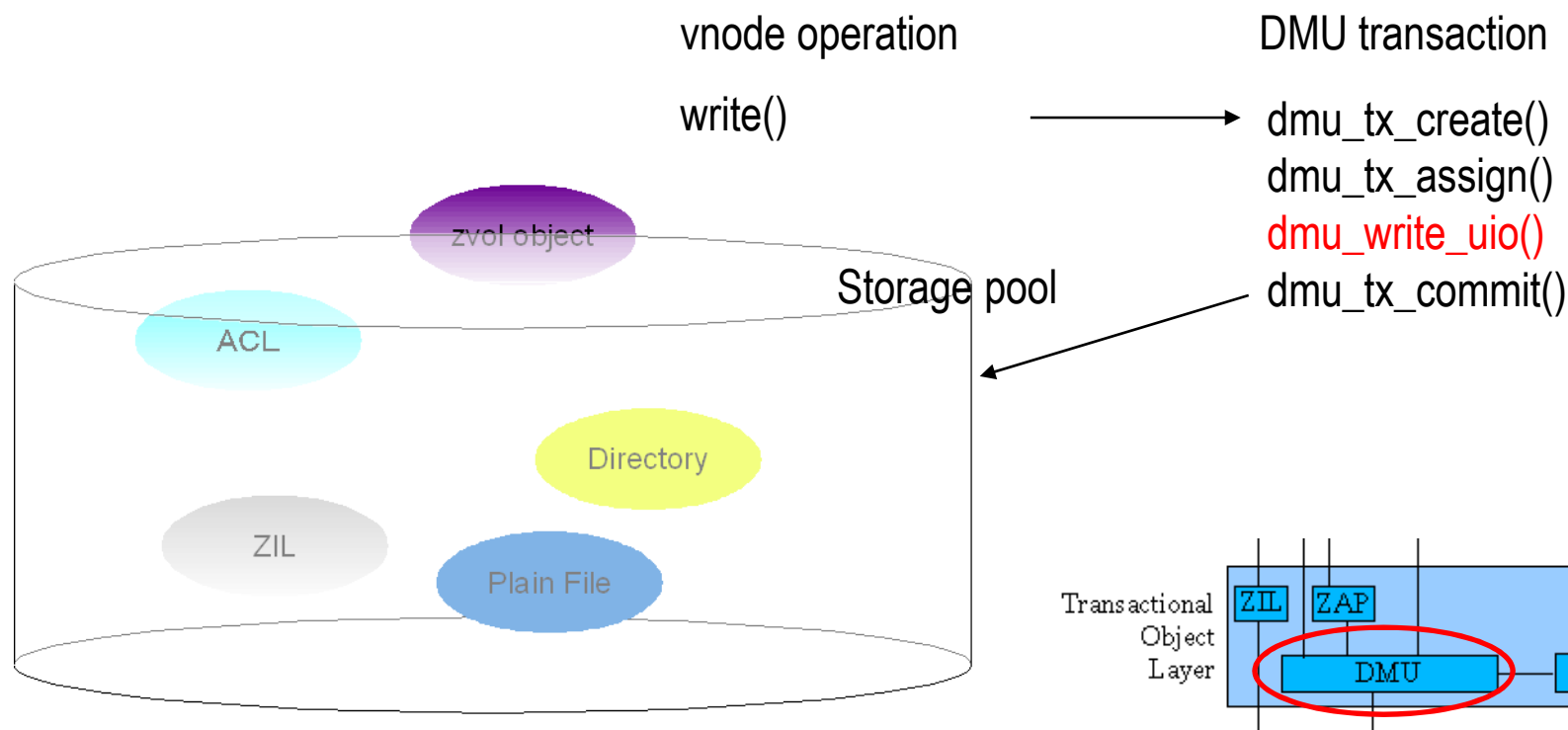
- メタデータとデータをオブジェクトとして扱う
  - ファイル、ディレクトリ、ACL、ZIL



# Transaction Object Layer

## 1. インターフェースレイヤに DMU オブジェクト操作を提供

- トランザクション管理
- DMU オブジェクト操作 (ZIO ヘマップ)



# Transaction Object Layer

## 1. DMU オブジェクト例

```
# zdb -v tank/test
Dataset tank/test [ZPL], ID 34, cr_txg 501, 1.56G, 7 objects
```

Object	lvl	iblk	dblk	lsize	asize	type
0	7	16K	16K	16K	15.0K	DMU dnode
1	1	16K	512	512	1K	ZFS master node
2	1	16K	512	512	1K	ZFS delete queue
3	1	16K	512	512	1K	ZFS directory
4	1	16K	512	512	1K	ZFS directory
5	3	16K	128K	800M	800M	ZFS plain file

```
# zdb -v tank/vol
Dataset tank/vol [ZVOL], ID 40, cr_txg 541, 802M, 3 objects
```

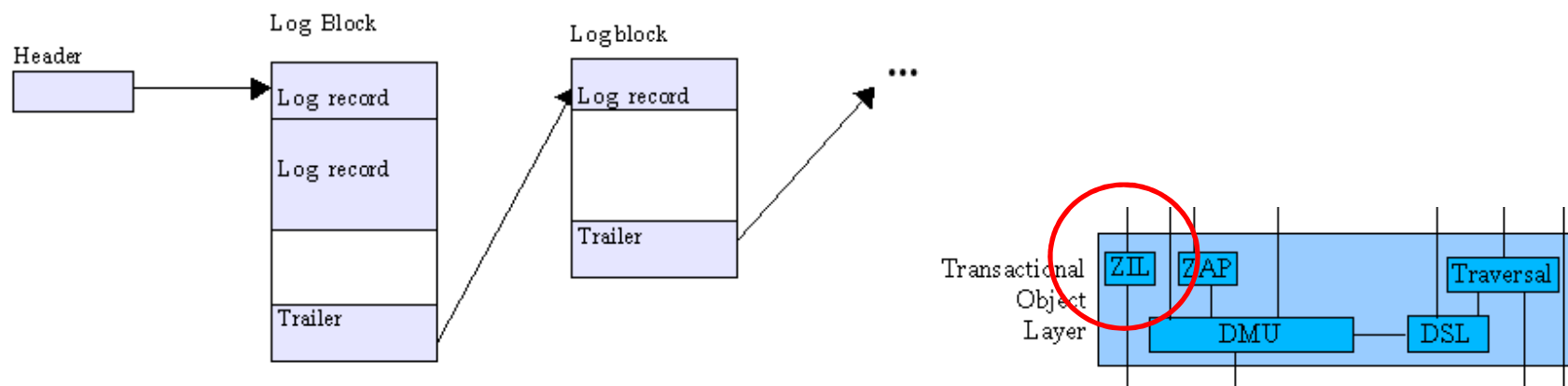
Object	lvl	iblk	dblk	lsize	asize	type
0	7	16K	16K	16K	14.0K	DMU dnode
1	4	16K	8K	800M	802M	zvol object
2	1	16K	512	512	1K	zvol prop



# Transaction Object Layer

## 1. ZIL (ZFS Intent Log)

- ログとしてシーケンシャルに書き込む (ジャーナルログ)
- 専用の高速ログデバイスを指定可能
  - フラッシュディスク
  - ミラー構成
- データベースなど同期書き込みが多いアプリケーション向け





# Transaction Object Layer

## 1. トランザクション

- txg\_time 以内に vdev へ反映
- 障害発生時のトランザクションは破棄される (チェックサム)
- システムコールと一対一に対応しない (大きな write() は分割される)
- メタ構造とデータの対応について常に一貫性が保たれている
  - アプリケーションのデータ一貫性が保護されるという意味ではない
  - fsck 不要？ → それでも整合性チェックはした方が良い - scrub



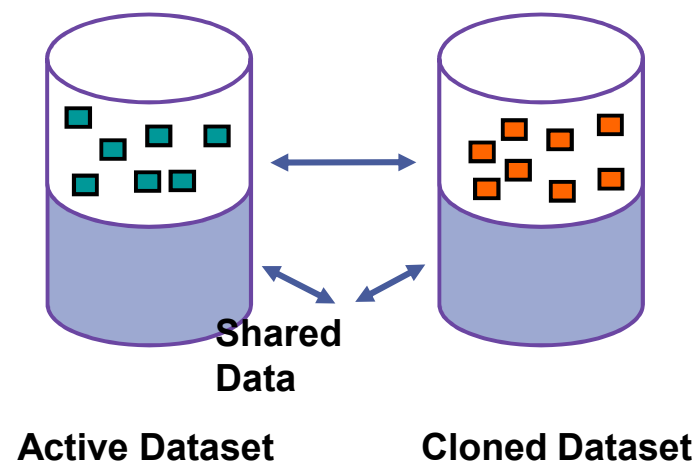
# Transaction Object Layer

## 1. スナップショット

- 静止点を記録するのみ
- パフォーマンスへの影響がない
- 無駄に容量が増えない (COW)

## 2. スナップショット操作

- ロールバック
- 書き込み可能なクローンとしてマウント



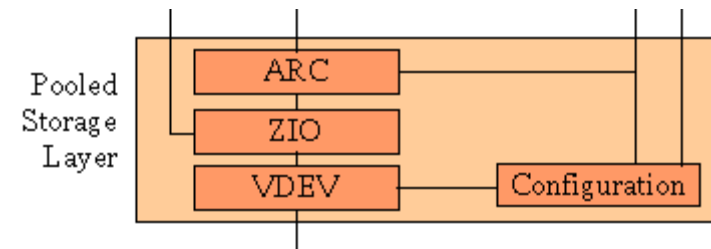
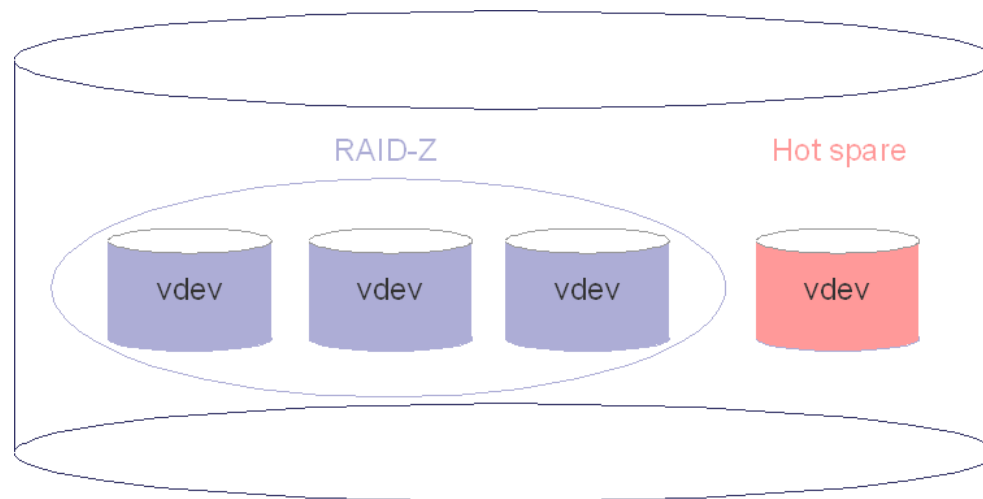
```
# zfs snapshot tank@snap1
```

```
# zfs list -t snapshot
```

NAME	USED	AVAIL	REFER	MOUNTPOINT
tank@snap1	0	-	19K	-

# Pooled Storage Layer

1. vdev に依存しないフラットアドレス空間を提供
2. vdev – ハードディスク、ファイル (通常は rdsk)
3. キャッシュ (ARC)
4. I/O パイプライン (ZIO)
5. ボリュームマネージャ (VDEV : ストライプ、ミラー、RAID 5/6)



# Pooled Storage Layer

## 1. プールの作成

```
# zpool create tank mirror c1d1 c2d0 spare c2d1
# zpool status tank
pool: tank
state: ONLINE
scrub: none requested
config:
  NAME      STATE  READ WRITE CKSUM
  tank      ONLINE  0   0   0
    mirror  ONLINE  0   0   0
      c1d1   ONLINE  0   0   0
      c2d0   ONLINE  0   0   0
  spares
    c2d1     AVAIL
errors: No known data errors
```



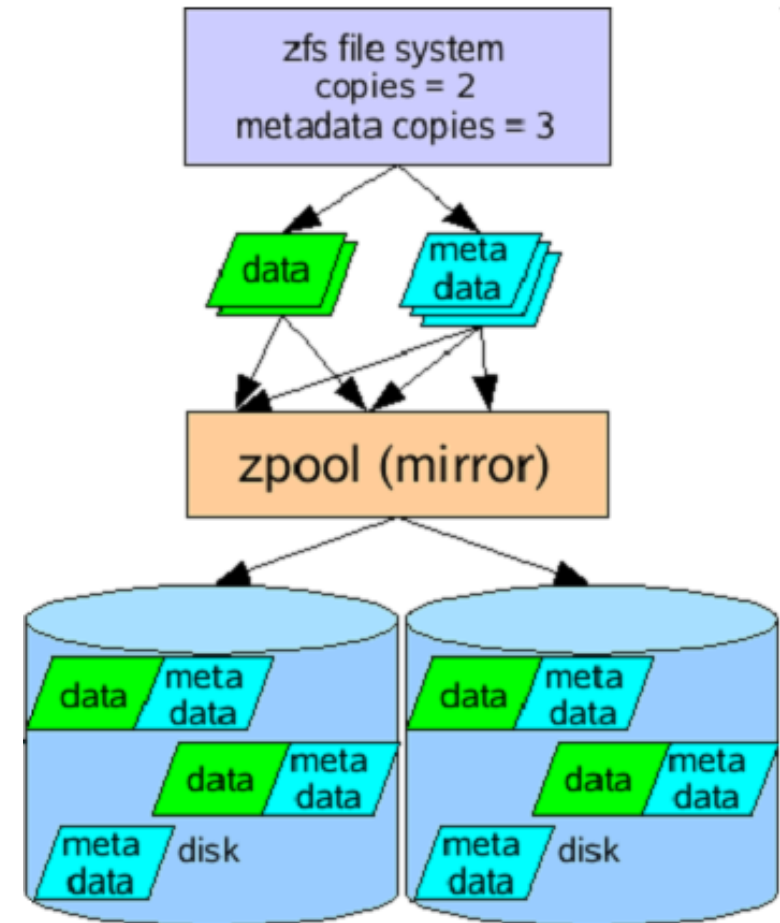
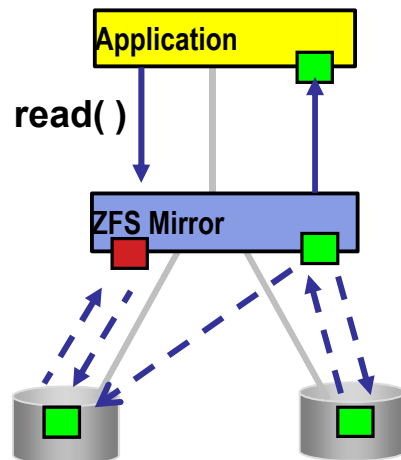
# Pooled Storage Layer

## 1. 冗長構成によるデータ保護

- メタデータは  $n + 1$
- データは指定数だけ冗長化
- $\text{copies} = n$  はディスク障害を救済しない
  - mirror、RAID-Z を使う

## 2. チェックサム

- fletcher2、fletcher4、SHA256

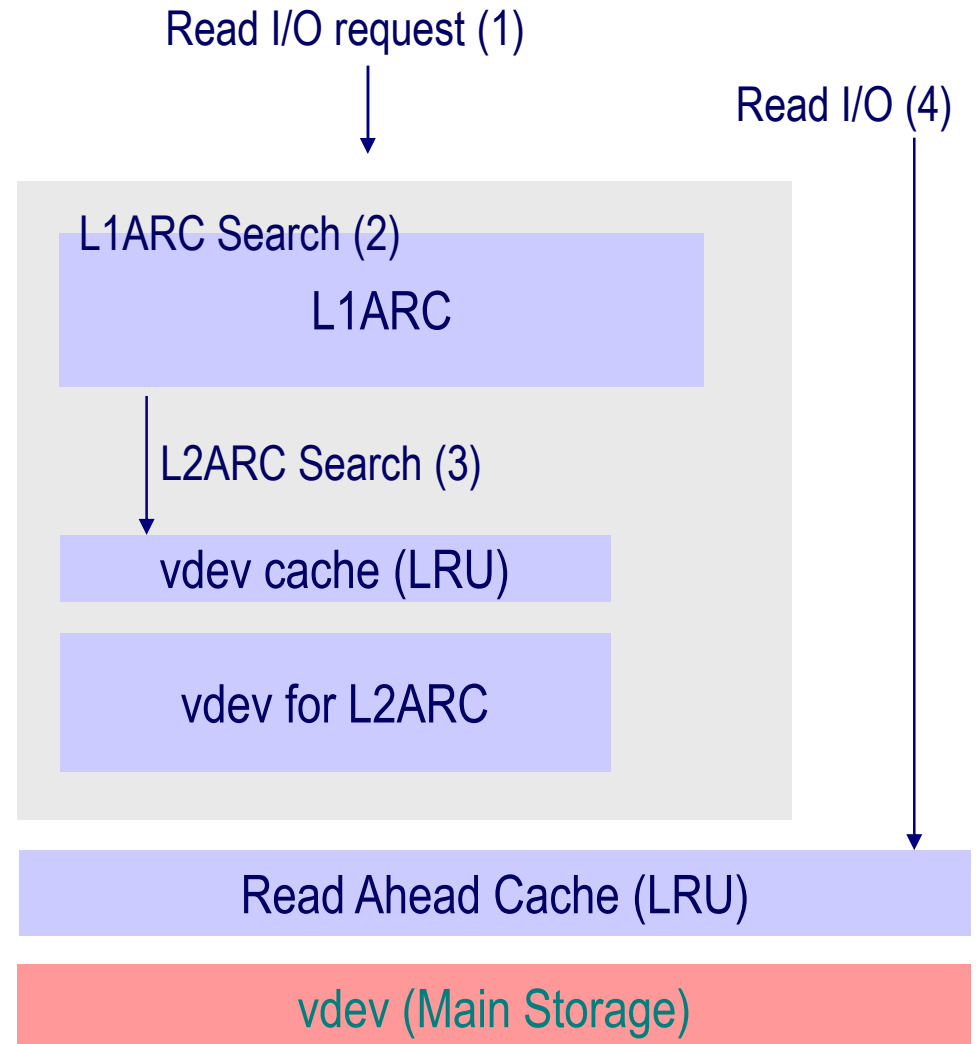




# Pooled Storage Layer

## 1. 読み込み キャッシュ

- メインメモリを可能な限り使用
  - Read ahead cache (10 MB)
  - Adaptive Replacement Cache
- L2ARC – 専用外部デバイス



# Pooled Storage Layer

## 1. サポートする RAID 構成

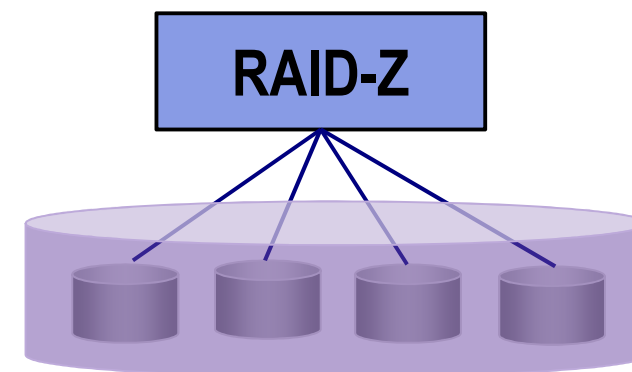
- RAID 0
- RAID 1 (mirror)
- RAID 5 (RAID-Z)
- RAID 6 (RAID-Z2)
- RAID 7 (RAID-Z3 – NexentaStor 次期バージョン対応予定)
- ホットスペア

## 2. RAID-Z

- 可変ストライプ幅

## 3. チェックサムによる自己データ修復

- 静かなデータ崩壊 (Silent Corruption) 抑制



# ポータビリティ (ベンダロックインからの解放)

1. オープン仕様
  - お客様のデータはお客様のもの
2. スナップショットの送受信 (send/receive)
  - 標準入出力
  - SSH
  - 遠隔バックアップ
3. 物理デバイスの移動
  - エクスポート・インポート
  - 適応型エンディアン制御
4. オープンソース
  - FreeBSD (10 日でポーティング)
  - Linux (FUSE)





## NexentaStor プラグイン

---

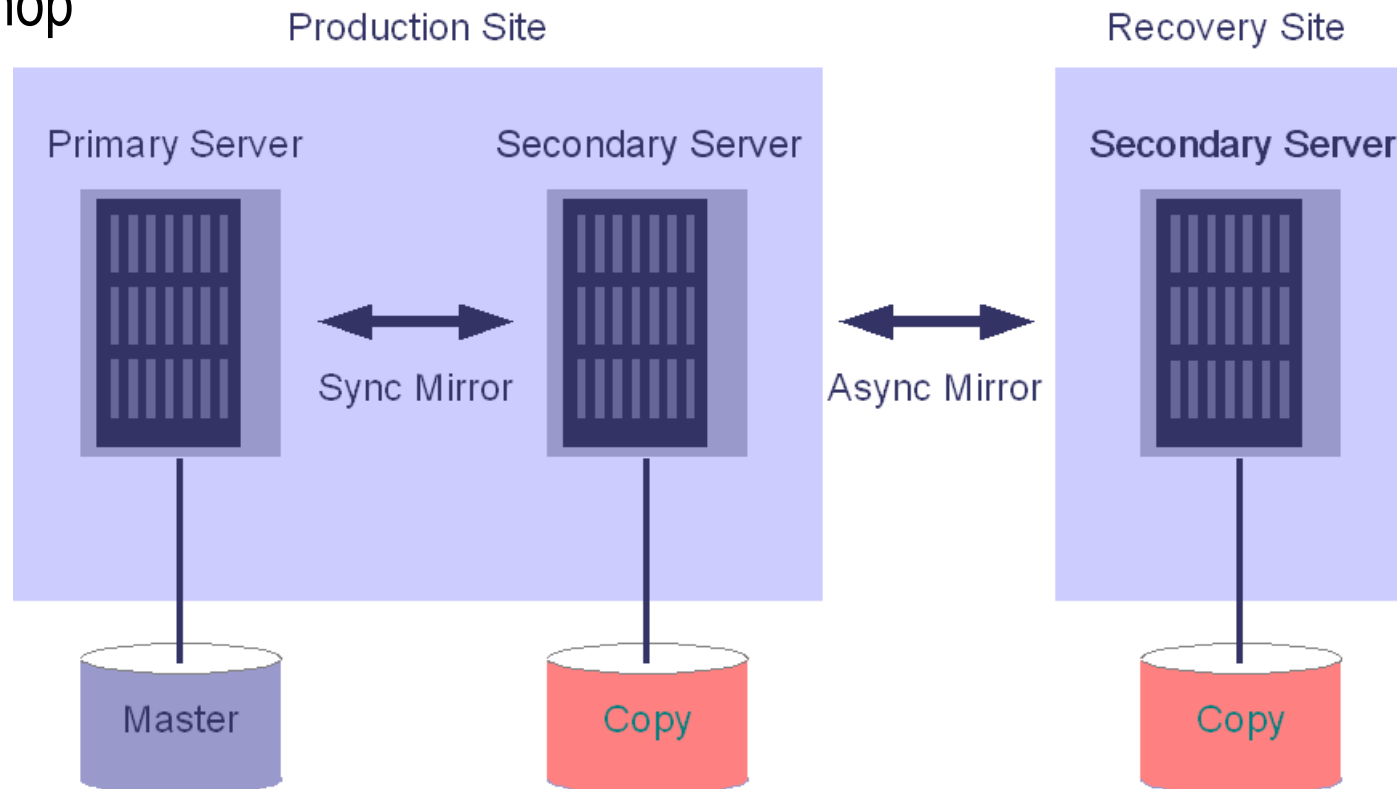


“ストレージ・ソリューションのリーディング・プロバイダ” コアマイクロシステムズ株式会社

Copyright © Core Micro Systems Inc., All rights reserved.

# Auto-CDP

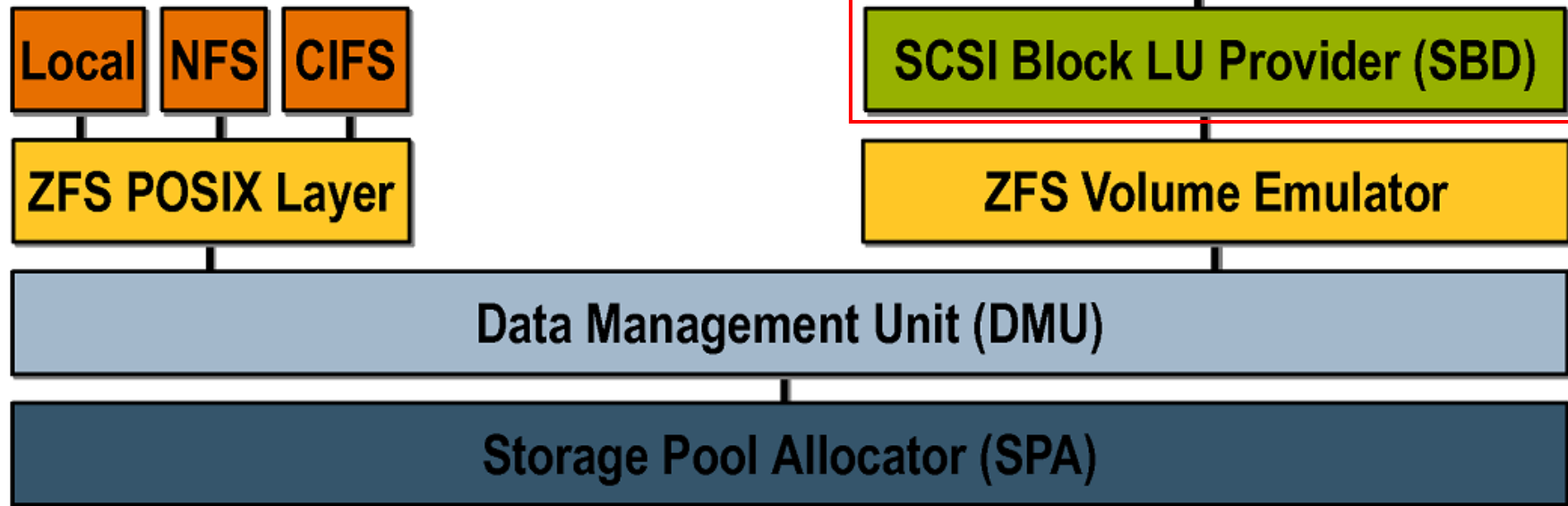
1. ブロックレベルレプリケーションソフトウェア
2. Sync / Async
3. Forward / Reverse
4. Multihop



# COMSTAR (Target 2.0)

## 1. SCSI ターゲット機能

- iSCSI
- FC
- SAS (Future)
- FCoE (Future)



# VM Data Center

## 1. 仮想環境向けストレージサービス統合管理機能

- VMWare
- Xen
- Hyper-V

**NEXENTA** About | Support | Register | Help Welcome Administrator | Logout

◆ Status ◆ Settings ◆ Data Management ◆ Analytics ◆ VM DataCenter

Dashboard VM Hosts Inventory Console View log

VM DataCenter Summary VM DataCenter Hosts Summary

192.168.100.9 (ESX, 2x2802 MHz, 2.00 GB)

192.168.100.9 ESX, 2x2802 MHz, 2.00 GB  
Virtual Machines vms on 192.168.100.9  
> winxp Running  
VDisk Storages vstorages on 192.168.100.9  
> Storage1 (2) (vmfs)  
> xxx (vmfs)

192.168.100.4 (ESX, 4x2800 MHz, 4.00 GB)

**VIRTUAL MACHINE SUMMARY: 192.168.100.9: WINXP**

**winxp**  
Microsoft Windows XP Professional (32-bit)

Host CPU Usage: 28.0 MHz  
Host Memory Usage: 30.0 MB  
Power State: Running

**Take Snapshot**  
Snapshot virtual machine "winxp". Snapshot will be taken at both VM Host and ZFS levels, to ensure coherency

**Move VM on Other Host**  
Use VMotion VHost capability to move migrate virtual machine on other host. Both VHosts origin and target one should have VMotion capability enabled.

**VStorage Info**

VStorage Name	Type	Local Backstore	Capacity	Free
xxx	vmfs	vpool/xxx	6.00 GB	3.60 GB

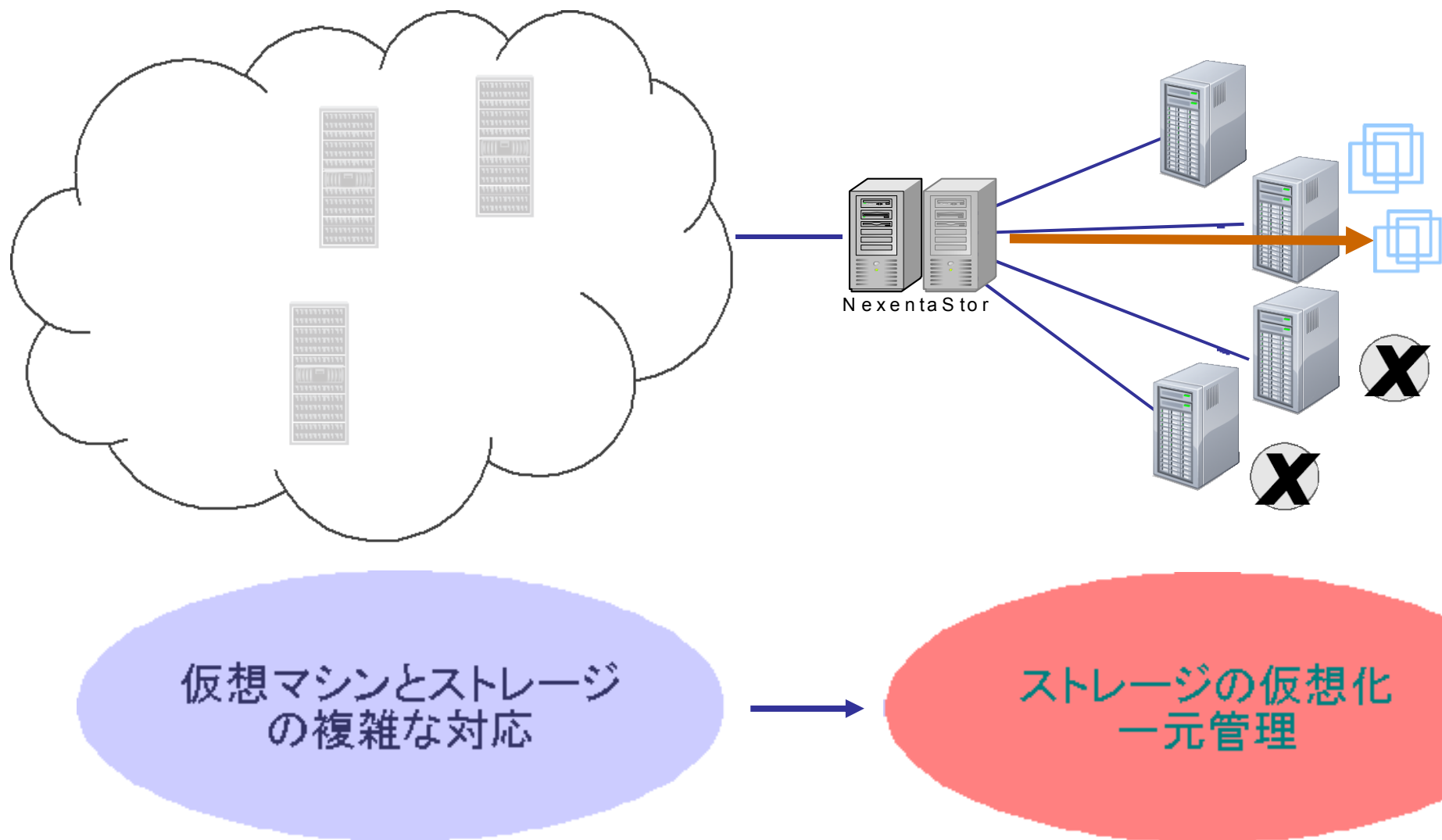
**VDisk Info**

VDisk Name	VStorage Name
winxp.config	xxx
winxp:3000	xxx

Found a bug? Feature request? Request Technical Support



# VM Data Center

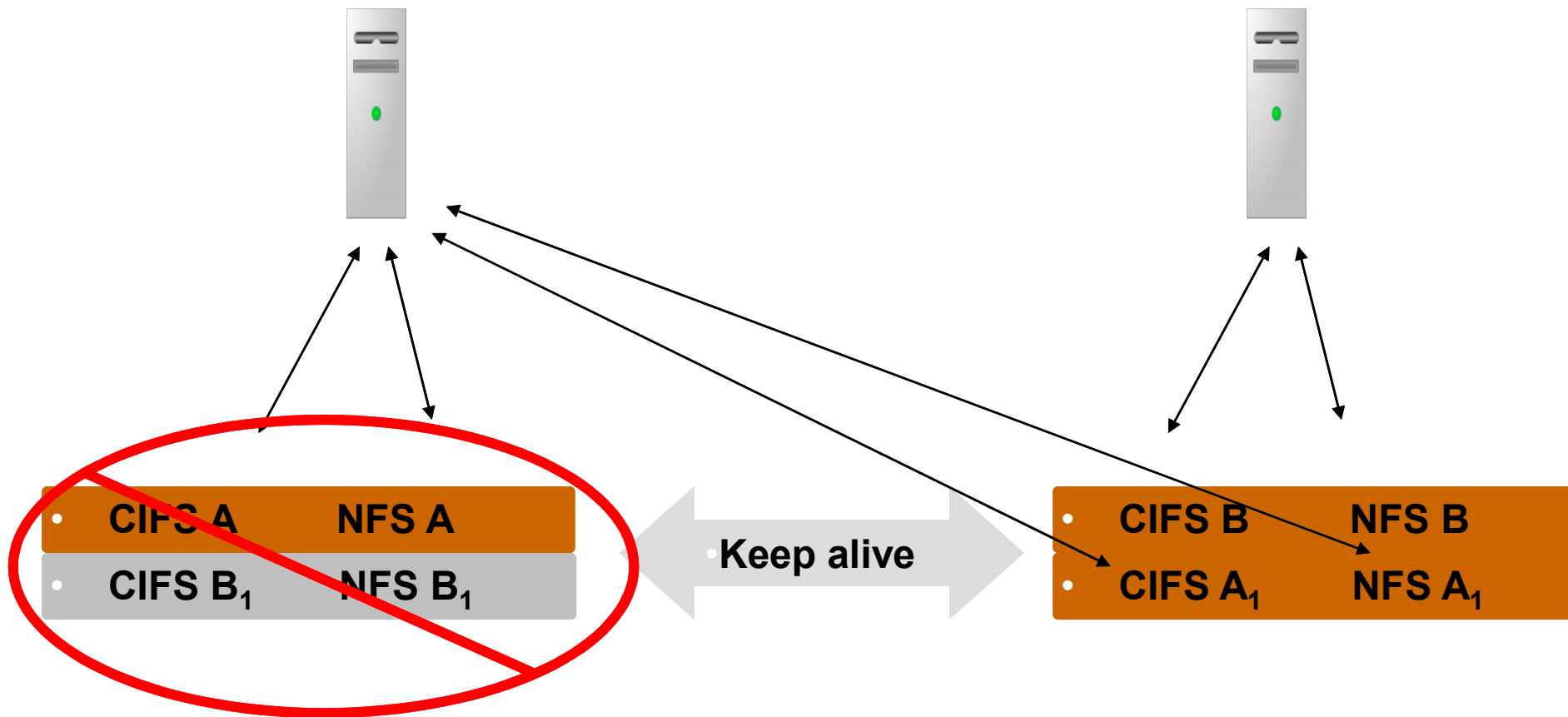




# HA Cluster 1.0

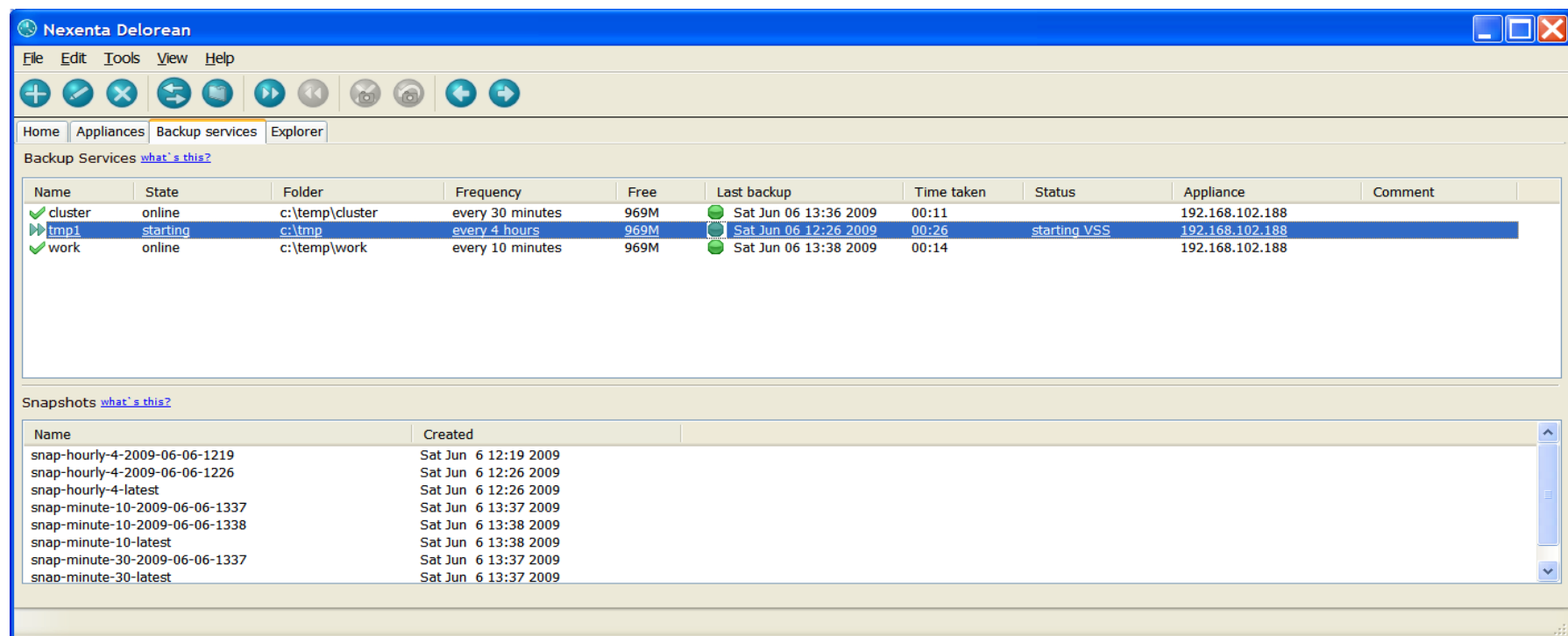
## 1. HA ソフトウェア

- High-Availability.Com 社 RSF-1 ベース



# Delorean

1. Windows ホストの CDP バックアップソフトウェア
2. VSS 対応
3. Windows エクスプローラ風操作 GUI





## 機能比較

---



“ストレージ・ソリューションのリーディング・プロバイダ” コアマイクロシステムズ株式会社  
Copyright © Core Micro Systems Inc., All rights reserved.

# 機能比較

Features	NexentaStor	N社同クラス製品	E社同クラス製品
HA Cluster	有り (ファイルサービスのみ)	有り	有り
Protocols	NFS v3/v4, CIFS, iSCSI, HTTP, FTP, NDMP, WebDAV		
Hybrid Storage Pool	有り	無し	無し
重複排除	実装される予定	有り	有り
データ圧縮	有り	無し	無し
スケーラビリティ	128-bit (Zettabyte)	64-bit	64-bit
RAID	有り	有り	有り
シンプロビジョニング	有り	有り	有り
スナップショット	有り	有り	有り
スナップショットリストア	有り	有り (オプション)	有り (オプション)
非同期ミラー	有り (オプション)	有り (オプション)	有り (オプション)
クローン	有り	有り (オプション)	無し
リモートアーカイビング	有り	有り (オプション)	有り (オプション)
電子メール通知	有り	有り	有り
データ検索	有り (MS Office、PDF etc..)	無し	無し



## アプライアンス製品

---



“ストレージ・ソリューションのリーディング・プロバイダ” コアマイクロシステムズ株式会社

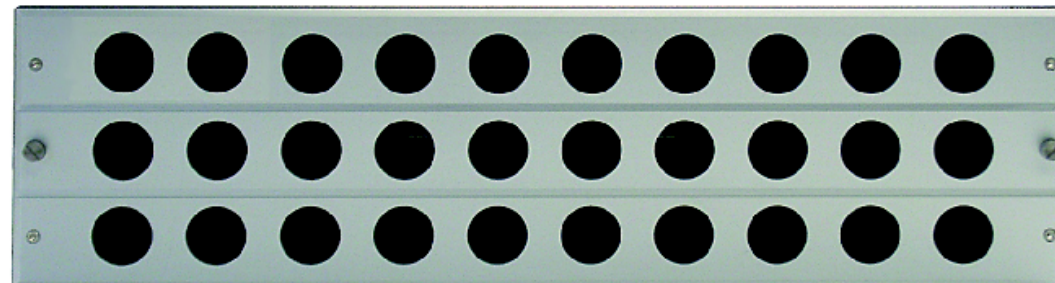
Copyright © Core Micro Systems Inc., All rights reserved.



# Prime STOR ZFS

大規模同時アクセスに応える先進のユニファイドストレージ

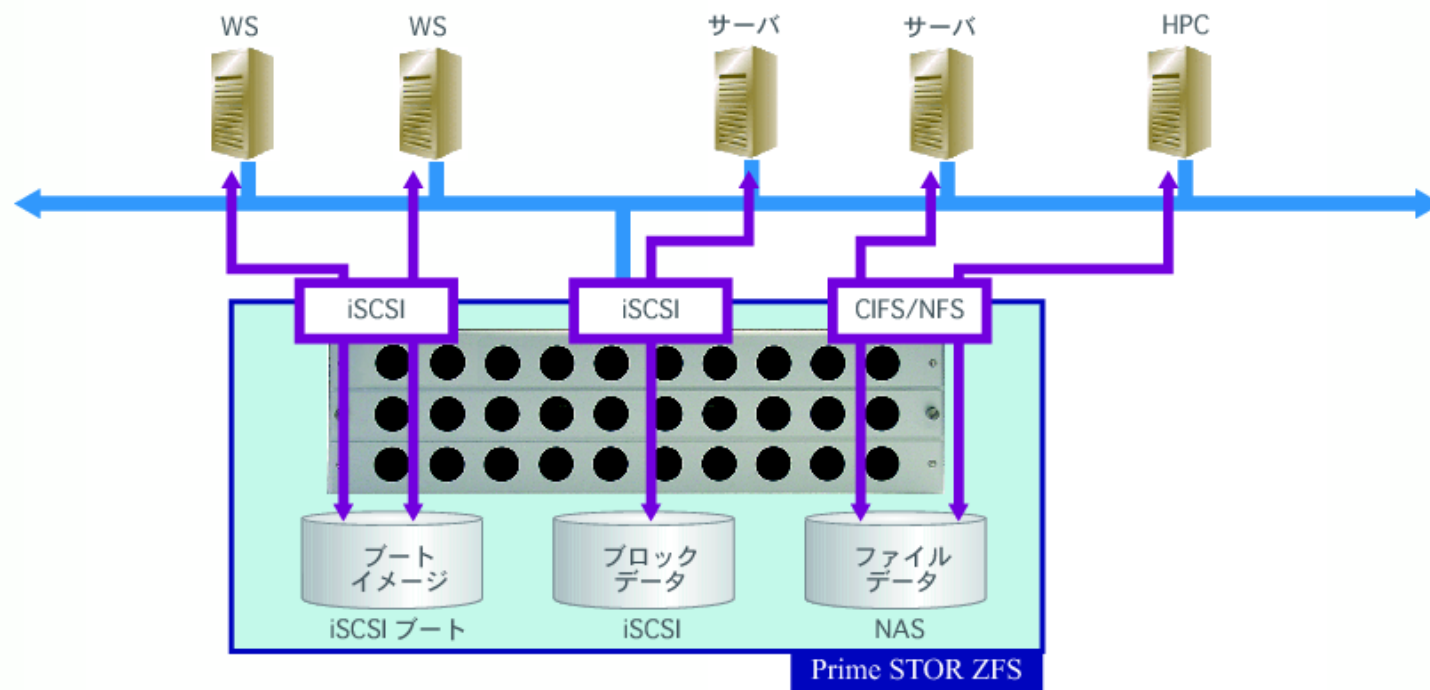
- トランザクションI/O を最適化する階層キャッシュ構造
- スケーラブルストレージプール
- シンプロビジョニング
- 次世代 10 GbE 対応 iSCSI / NAS ユニファイド I/O



# Prime STOR ZFS

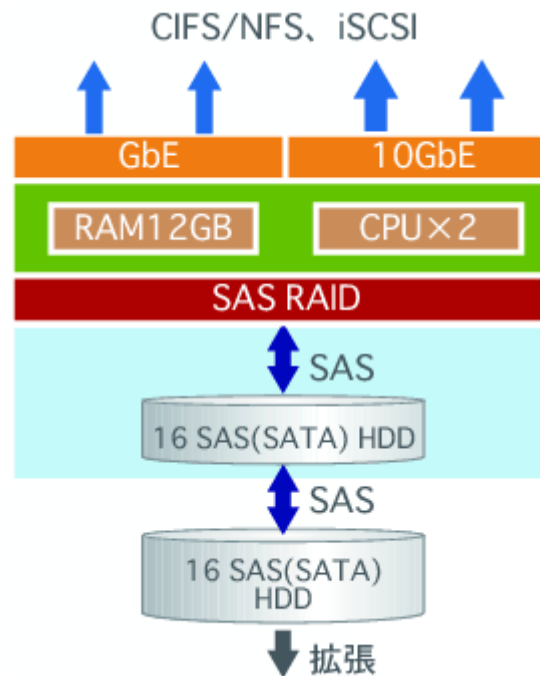
## 1. 用途

- 仮想サーバ / 仮想マシン用
- EDA / CAE 大規模ワークグループ
- メディカルイメージング / CG レンダリング



# Prime STOR ZFS

1. DRAM / SSD / HDD を階層化
2. 書き込みキャッシュ (DRAM / SSD オプション)
3. 読み出しキャッシュ (DRAM /SSD オプション)
4. レプリケーションソフトウェアオプション







# Prime GATE ZFS

大規模同時アクセスに応える先進のストレージ仮想化ゲートウェイ

- トランザクションI/Oを最適化する階層ストレージアーキテクチャ
- ダイナミック&シンプロビジョニング仮想化ストレージプール
- iSCSI 及び FC SAN 及び NAS ユニファイド I/O サービス



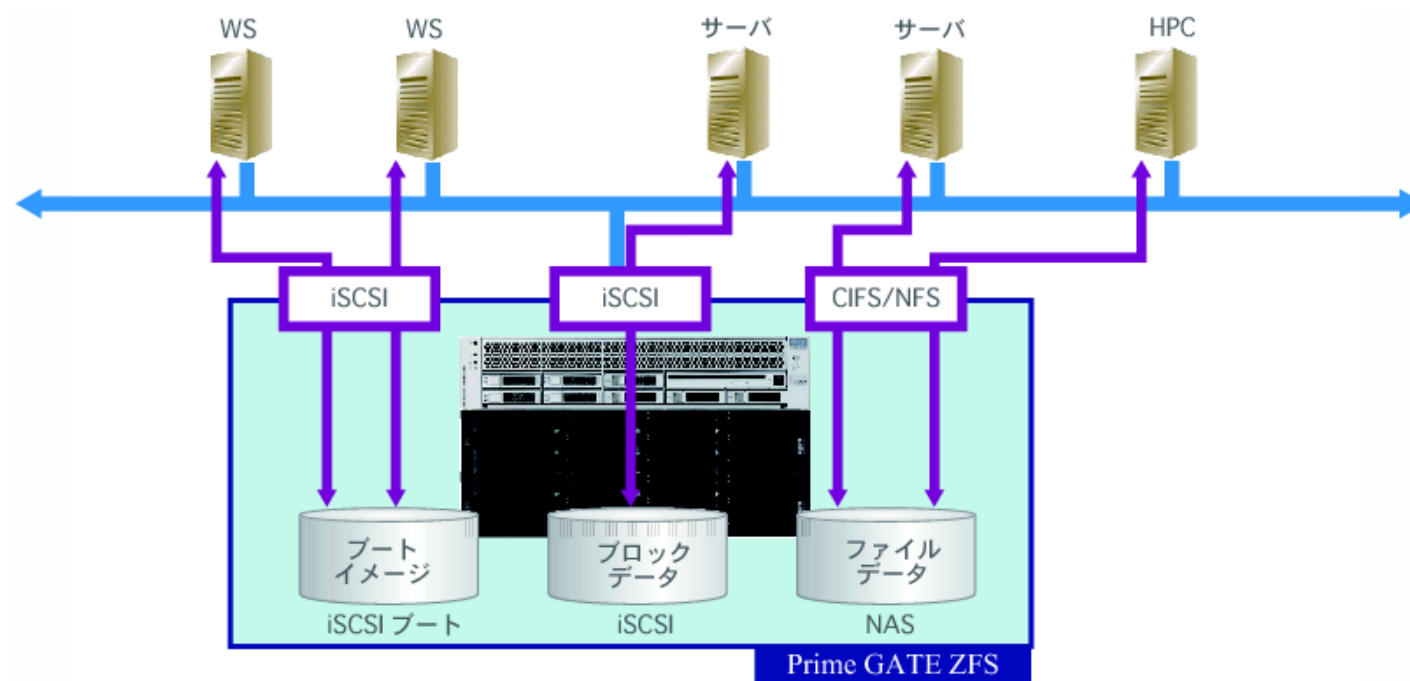
“ストレージ・ソリューションのリーディング・プロバイダ” コアマイクロシステムズ株式会社

Copyright © Core Micro Systems Inc., All rights reserved.

# Prime GATE ZFS

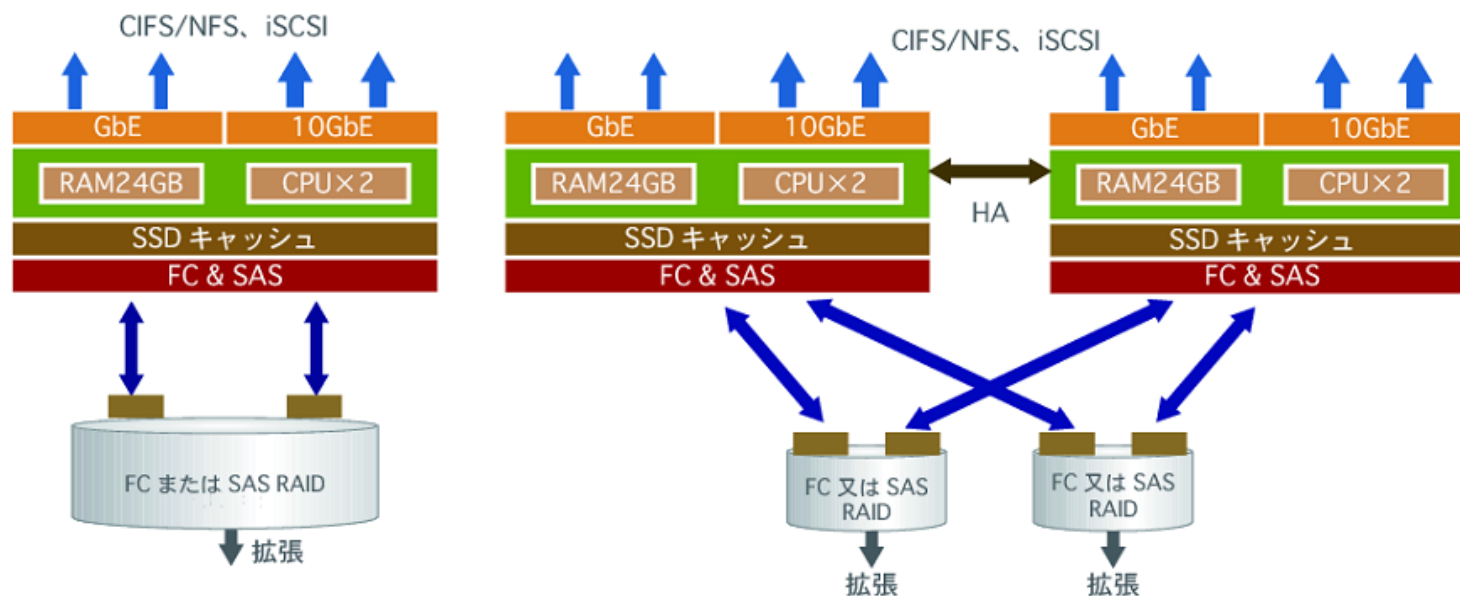
## 1. 用途

- エンタープライズ
- データセンター
- 仮想サーバ用バックストレージ



# Prime GATE ZFS

1. DRAM / SSD / HDD を階層化
2. 書き込みキャッシュ (DRAM / SSD)
3. 読み出しキャッシュ (DRAM / SSD)
4. HA オプション
5. レプリケーションソフトウェアオプション





## コアマイクロシステムズ株式会社

Core Micro Systems, Inc.

URL: <http://www.cmsinc.co.jp/> Mail: [sales@cmsinc.co.jp](mailto:sales@cmsinc.co.jp)

TEL: 03-5917-6451 IP Phone: 050-5558-5410 FAX 03-5917-6452

本社 〒173-0026 東京都板橋区中丸町11-2 ワコーレ要町ビル9F



Copyright © Core Micro Systems Inc., All rights reserved.