





# だれも教えてくれない ODMサーバの導入ノウハウ

KDDI株式会社 プラットフォーム開発本部  
プラットフォーム技術部  
加藤 真人

# 自己紹介



KDDI株式会社 プラットフォーム開発本部  
クラウドサービス開発部

加藤 真人



KDDI入社以来インフラ系サービスの開発を担当。  
これまでに、お客様のシステムを監視するサービスや  
KCPSの前身であるバーチャルデータセンターなどを開発。  
その後、KCPS(KDDIクラウドプラットフォームサービス)の開発を初期メン  
バーとして参加し、  
現在は数千台のサーバを支えるインフラ設計業務を担当。ODM機器の採用や、  
Open Compute Projectへの参加など積極的にインフラのコモディティ化を推  
進してる。



## KCPSの概要

サーバインフラミッション

サーバ構築 短納期と自動化にチャレンジ

今後のOCPへのミッション

日本仕様のOCP

# KCPSの概要



# K C P S って何？

## K C P S (KDDI Cloud Platform Service)

『CloudStack専用ポータルから仮想サーバ環境をオンデマンド提供するサービス  
WVSイントラネット接続と占有型サーバを標準提供するIaaS』

- ・ 利用ユーザ：  
法人のお客様・KDDI事業用の設備（AU-Cloud）でも利用
- ・ データセンタ：4局、物理サーバ：数千台

### [インフラの特徴]

- ・ 稼働率99.9999997%（2016Q1実績）
- ・ サーバを台湾ODMから直接調達
- ・ KDDIイントラ網に直結
- ・ 物理機器の持ち込みができる！





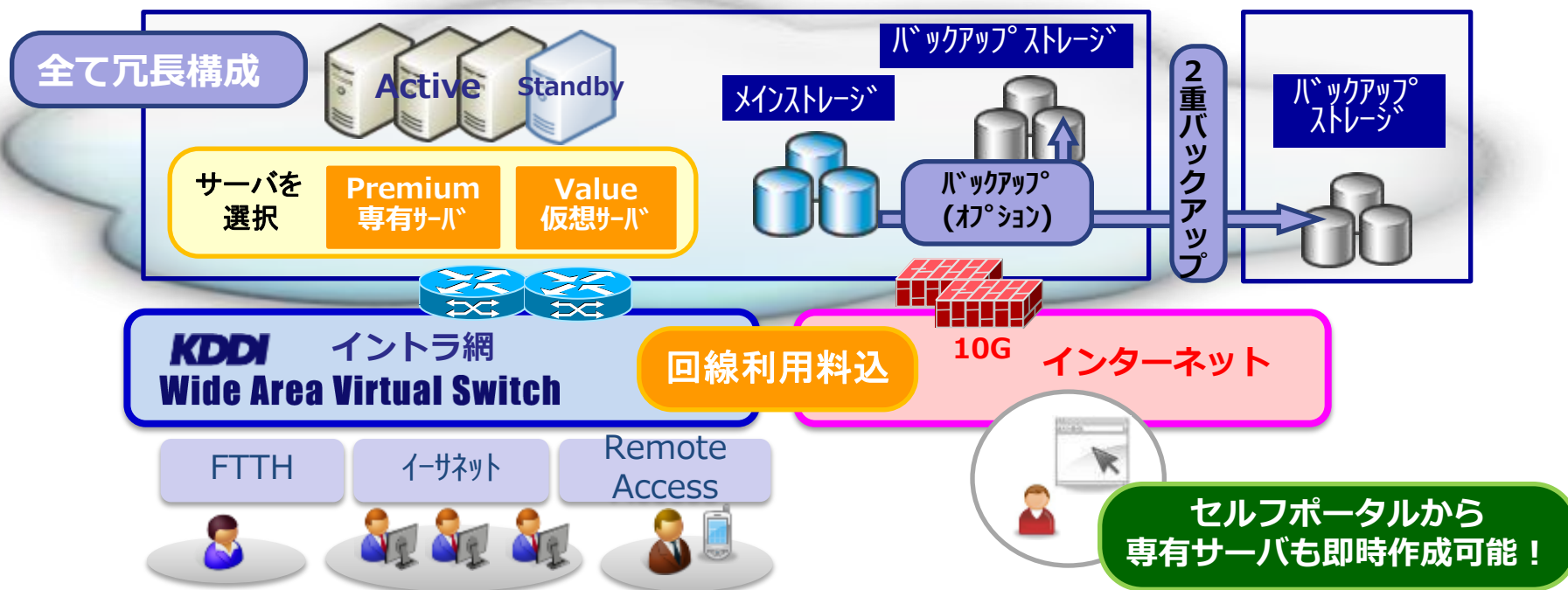
# KDDI Cloud Platform Service

① HA（フェイルオーバー）込みのQuality Cloud、SLA99.99%

② 社内システムのクラウド化に必須なイントラ回線接続料は無料

※お客様拠点のアクセス回線は別途費用がかかります

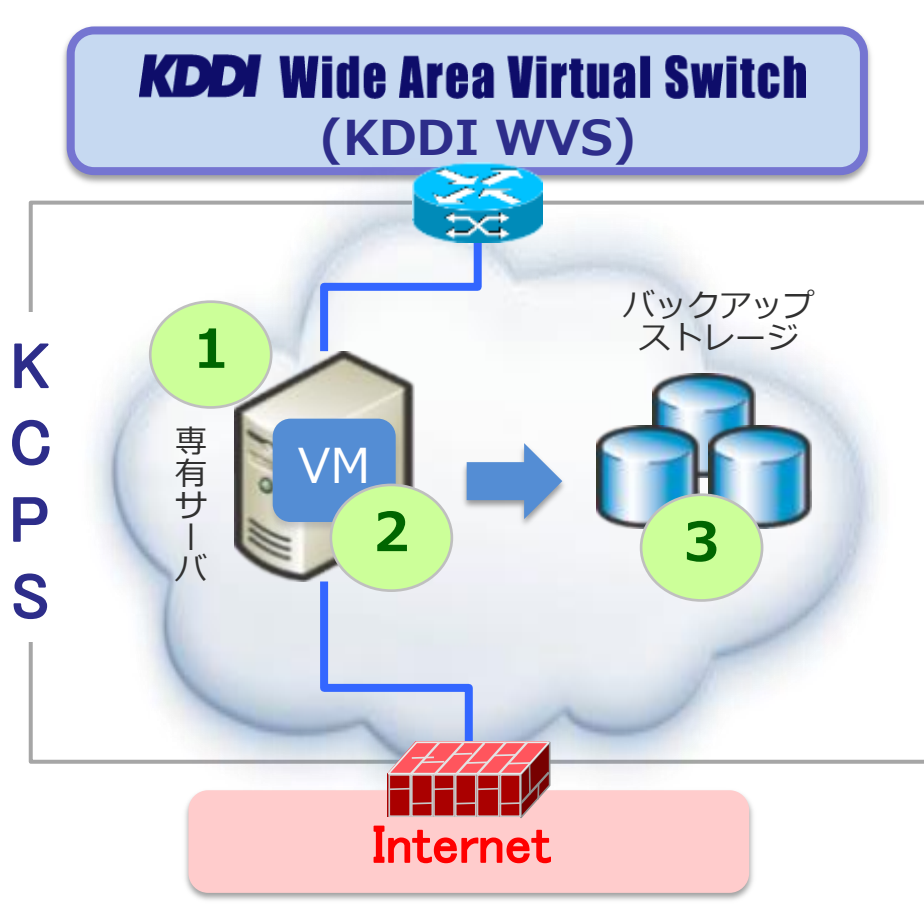
③ 専有サーバも提供。仮想サーバもいつでもリソースを増減可能



# KCPSの特徴① 専用サーバやイントラ網も利用できる

## 専用サーバをオンデマンドでタイムリーに構築可能

セルフポータル画面から仮想マシンを直ぐに作成できるから  
サイジングは気にせずオンプレと同等の環境をクラウド上で利用できる！



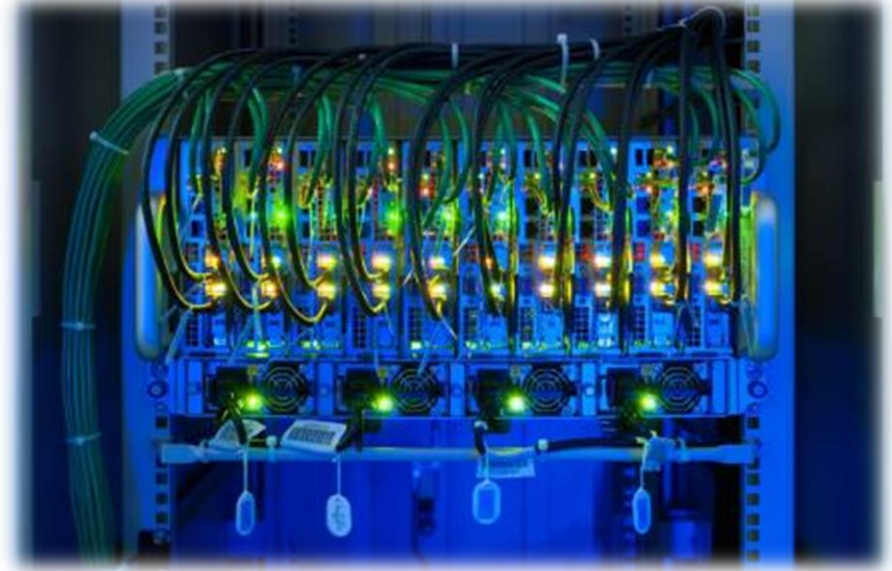
① 専用サーバ(Premiumメニュー)の追加

② 専用サーバ上に仮想マシンを作成(VM作成)

③ バックアップ設定



## KCPSの特徴② サーバの直接調達



- ✓ 大手サーバベンダーやFacebookにサーバを供給する台湾ODMベンダーから**直接サーバを購入**

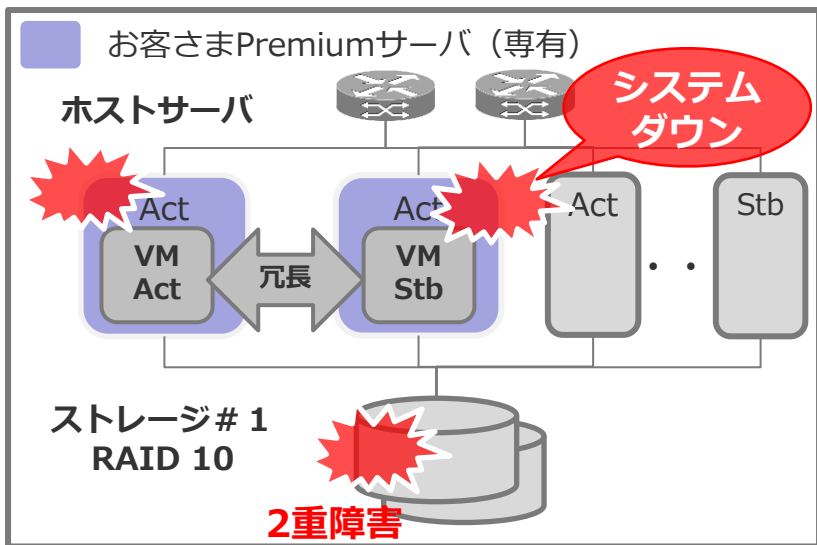
# KCPSの特徴③ 高信頼性機能の提供

ストレージハードウェアまで考慮した分散アーキテクチャ  
"エクストラ アベイラビリティ"を採用！  
これまでにない冗長構成をオンデマンドで利用でき  
お客さまシステムの可用性を大幅に向上！

国内クラウド  
事業者初！

一般的なクラウド

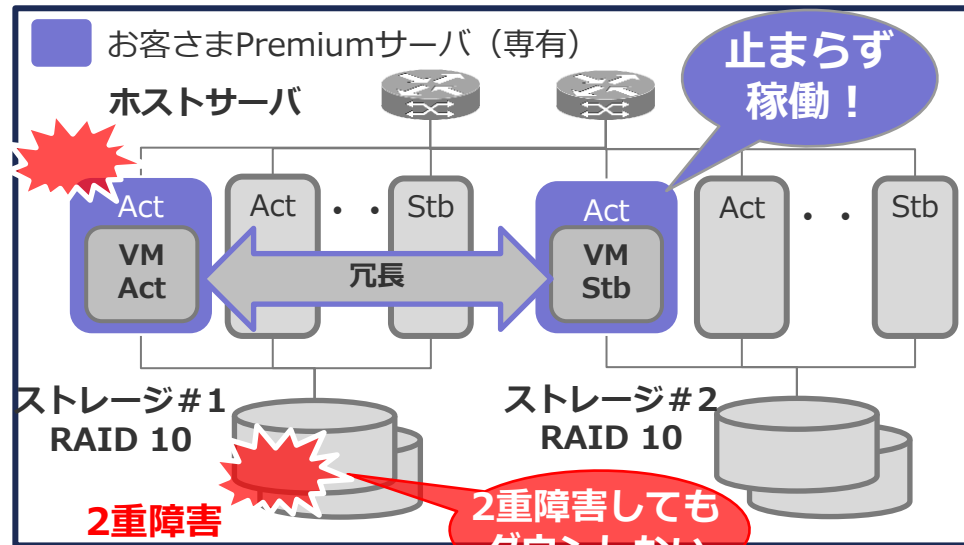
収容ストレージの指定できないため、  
仮想サーバレベルで冗長化しても  
ストレージ2重障害の影響を受ける



エクストラアベイラビリティ

2015年2月より提供開始

Premiumサーバ(専有サーバ)の収容  
ストレージを分散する事により、  
ストレージ2重障害の影響を受けない！



## KCPSの特徴④ オブジェクトストレージ

### ①高い堅牢性

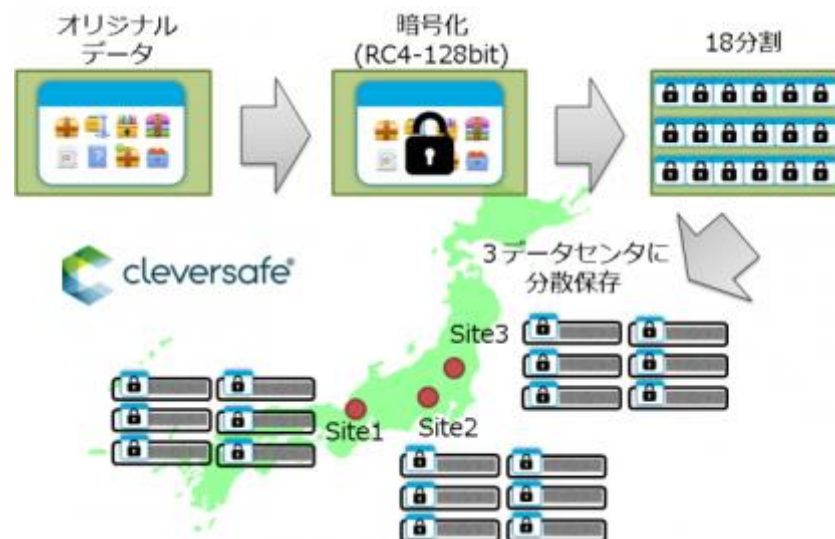
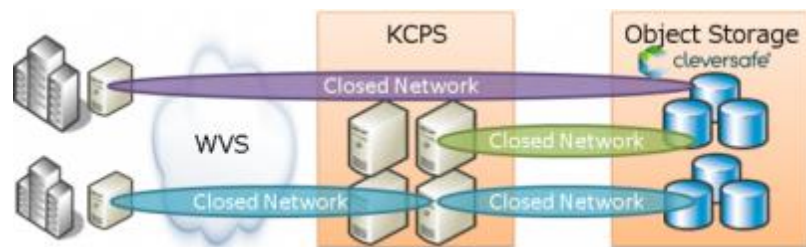
「99.99999999999999%の堅牢性」

### ②クローズドNWで利用

「ネットワーク接続料とデータ転送料が不要で安価」

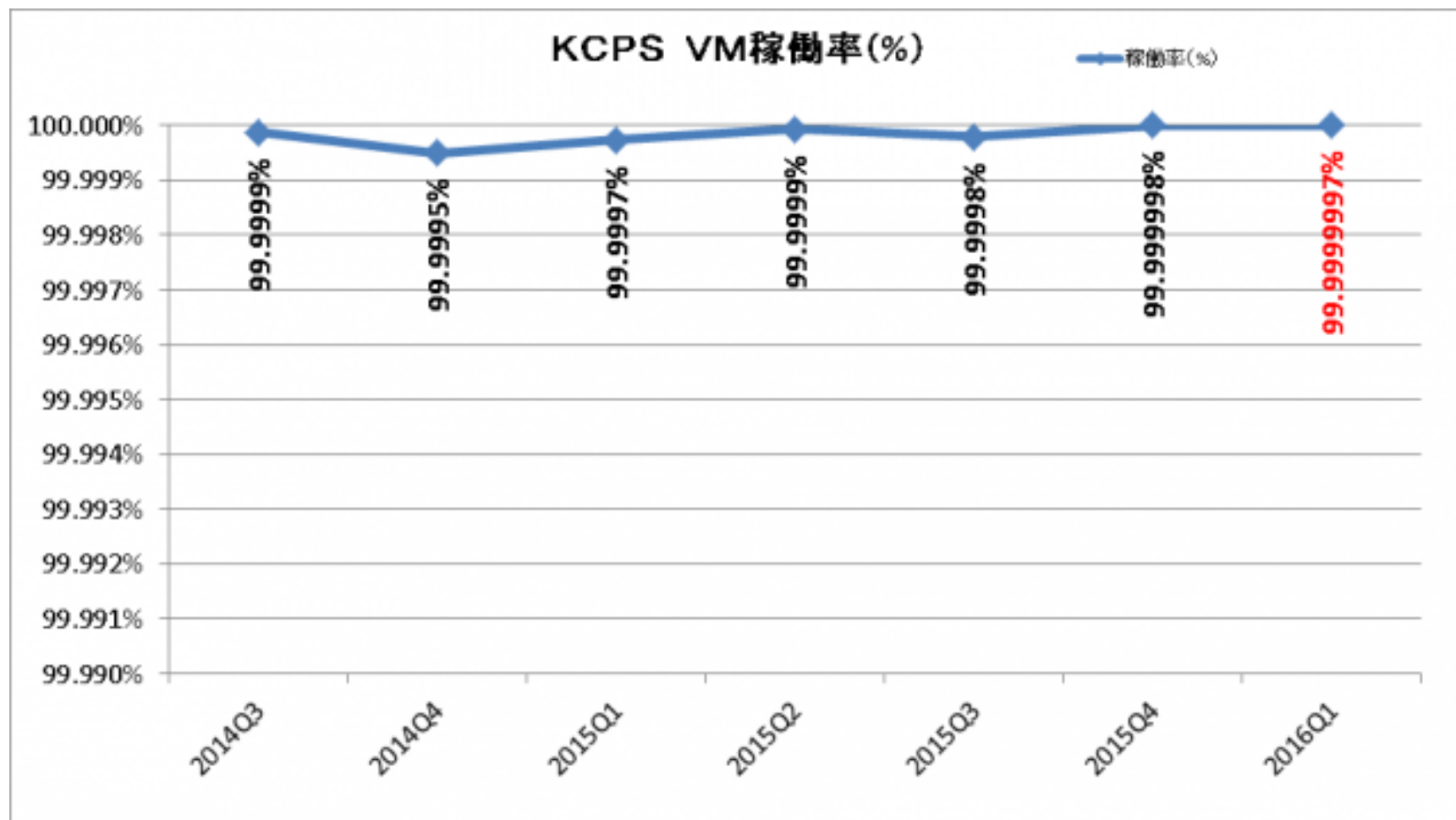
### ③高いセキュリティ

「auスマートパス基盤での実績を活かしたサービス」



## KCPSの特徴⑤ KCPSの高い稼働実績

KDDI Cloud Blogで稼働率と故障内訳を公開！  
2016Q1稼働率は**99.9999997%**





# ちょっと宣伝

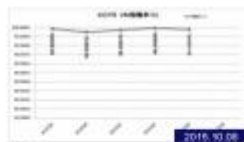
<http://cloudblog.kddi.com/?v=block>

**KDDI クラウドブログ で検索**

## KDDI Cloud Blog

HOME Category Archive Ranking Member Gallery Movie

Google Apps 3,000件以上プレゼント!



・KCPs  
KCPs 2015Q3の稼働率について

KDDIクラウドプラットフォームサービス以下、KCPsのサービス運用を担当している佐藤様です。KCPs(2015Q3) [一]



・Data Center  
大阪の人情味あふれる土地柄に馴染み、お客様に寄り添う運用を目指して～TELEHOUSE OSAKA 2 運用者インタビュー



・KCPs ・News  
国内初のイントラネット接続型モバイルアプリ/IoTデバイス開発基盤「KCPs mBaaS by Kii」提供開始



・Data Center  
厳しい要求を満たした堅牢な施設 TELEHOUSE OSAKA 2の建設現場とは？～大林組に聞く、建設時の苦勞とこだわり～



・IaaSPaaS ・KCPs  
KDDI×アマゾン データサービスジャパン株式会社の共催セミナーのご報告



# インフラ系の記事が注目されている

2016.04.19



• CloudStack • Data Center • IaaS/PaaS • Infrastructure • KCPS

## だれも教えてくれないODMサーバの導入ノウハウ

KDDIクラウドプラットフォームサービス(以下、KCPS)のインフラ担当の加藤真人です。KDDI Cloud Blogに [...]

加藤 真人



2015.05.15



• Infrastructure • KCPS

## 高品質なクラウドを支えるインフラ設計

KDDIクラウドプラットフォームサービス(以下、KCPS)のインフラ担当の加藤真人です。今回は、2月1日にリリースされた [...]

加藤 真人



2014.10.02



• Data Center • Infrastructure • KCPS

## クラウドを支える縁の下の力持ち (KCPSファシリティ紹介)

KDDIクラウドプラットフォームサービス(以下、KCPS)のインフラ担当の加藤真人です。今回は、インフラチームのファシリ [...]

加藤 真人



2014.08.05



• Data Center • IaaS/PaaS • Infrastructure • KCPS

## 実は、低コストと高品質なサーバは実現できる

初めまして、KDDIクラウドプラットフォームサービス(以下 KCPS) のインフラ担当 加藤真人です。プラットフォーム [...]

加藤 真人





# サーバインフラミッション

①低価格で高品質な都合のいいHWがほしい

②発注からリリースまで超短納期・自動化にしたい

③サーバの運用工数を削減したい

# ODMサーバという選択

## 低価格

- 不要な機能・パーツを排除
- ダイレクト購入だから納得価格
- 不足機能は追加実装可能

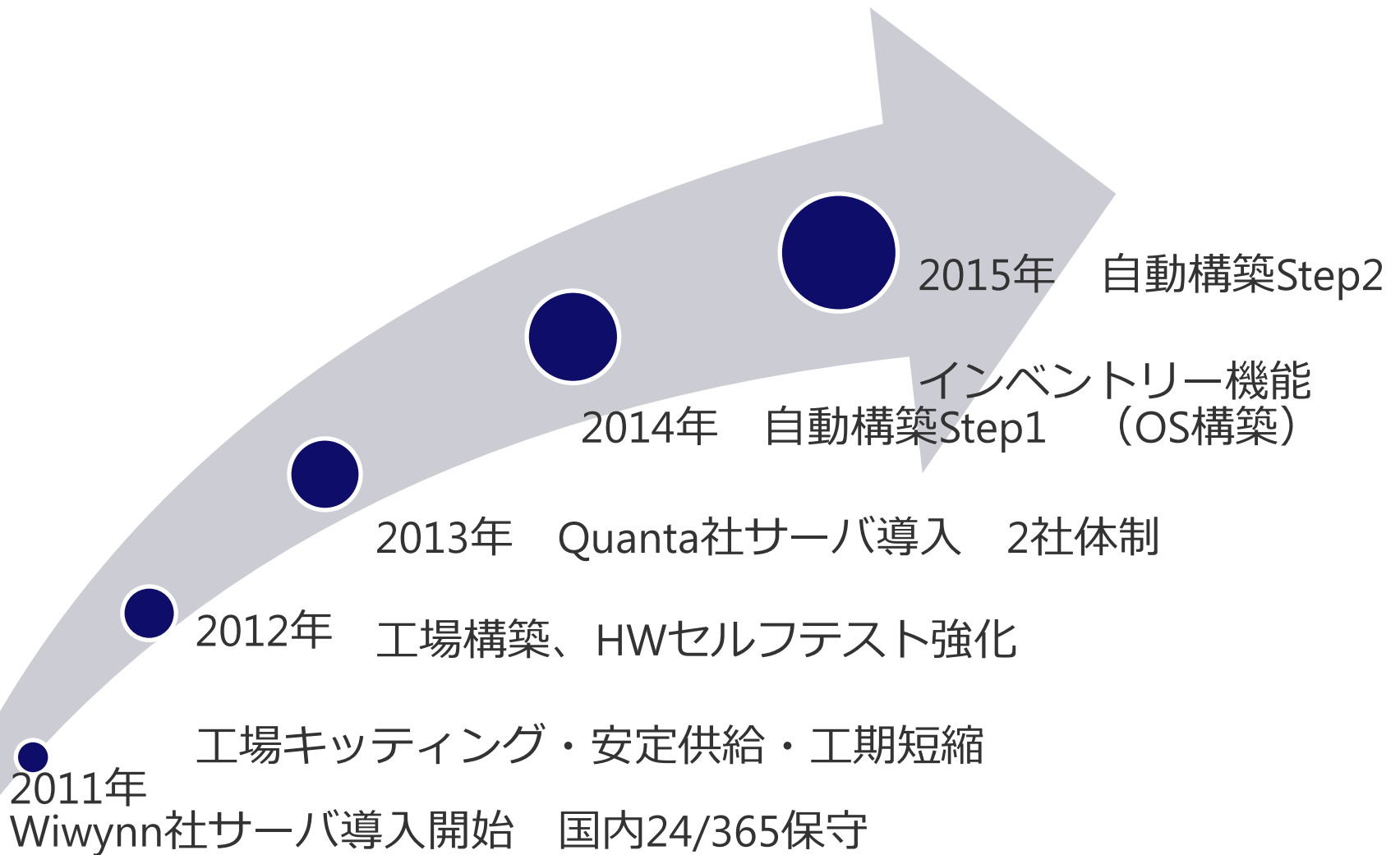
## 高品質

- メーカー製サーバと同一パーツ
- 品質の良いパーツの指定購入
- 不具合に対する特別対応

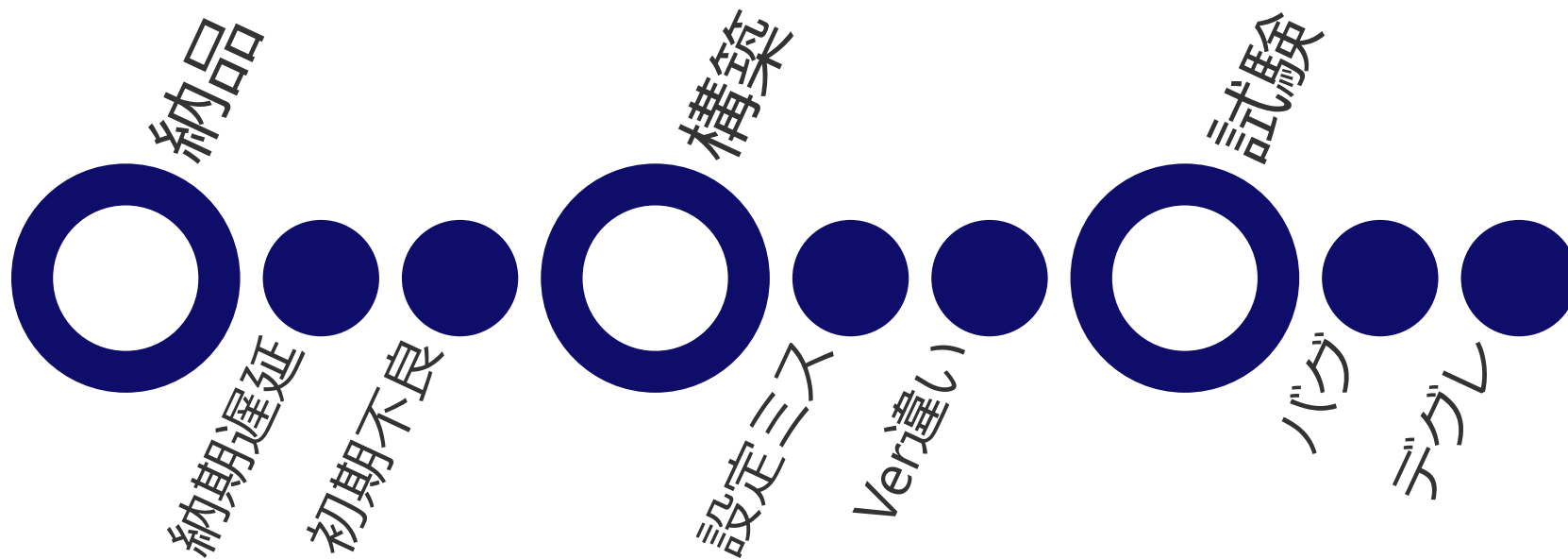
## スピード

- 工場での事前構築
- 新機能を逸早く実装可能
- サーバ仕様の変更が柔軟

# サーバ領域を活かしたSI領域へシフト



# ODMサーバでの課題



## 納期遅延 パーツ供給の遅れ

ODMメーカーは基本受注生産を行っており、大量のサーバパーツは確保していないため、外部要因によるパーツ遅延が発生し、サーバの出荷遅延となることがあるため、この対策が必要である。KDDIとしては、パーツの共通化や代替パーツの検討など遅延が発生した場合の別手段を確保し、納期遅延を未然に防いでいる。納品時期を複数回に分け、遅延時の影響を小さくする対策も有効である。

外部要因：

台風による洪水や地震による天災

パーツ工場の火災

パーツの製品変更 提供制限



海外ニュースやODMメーカーとの密な連携が重要



## 納期遅延 組み立て工程

組み立てラインから出荷までの工程で発生します。組み立て工程においてはラインのブッキングやストライキ、発送工程においては政府による輸出制限など想定外の理由で遅れることがあります。この遅れについて回避することが出来ないため、組み立て工場との工程確認を頻繁に行い、なにか不都合が発生した場合には情報を早くキャッチし、後工程で吸収できるよう調整を行うことが重要です。

工程要因：

組み立てラインのブッキング

中国政府の輸出制限

長期休暇による人員確保

パーツの不具合



ODMメーカーや工場との密な連携が重要

# 初期不良

数百台のサーバを一度に導入するため、初期不良は避けられない。工場でのエイジング内容を確認しKDDIオリジナルな試験工程を追加している。これにより、初期不良の発生率を一桁へ低減させることが出来た。輸送時や長時間稼働での初期不良はKDDIのDCに到着後実施する。後続工程へのスケジュール影響を出さないよう、高負荷長期運転、パーツの脱着試験（1000回）、DISK、メモリーの全面チェックなどを徹底的に行う。

## 試験ポイント

- ・ 負荷や温度による発生
- ・ 物理的な緩み、抜けなどの確認
- ・ パーツの正常性



この工程が一番重要！！

# 設定ミスやVer管理

BIOSやFMなどのHW設定は事前にODMメーカーの工場にて設定されてきますが、設定作業が人による手動設定なため、設定間違いが発生することは避けられません。対策としては、BIOSにおける初期設定をKDDI仕様に変更し、ROMに書き込むという対応を行いますが、システム開発の途中で変更が発生することがあり、どうしても手動設定箇所が残ります。各種設定を吸い上げ、チェックする仕組みを取り入れることが、ポイントです。

## ポイント

- ・ 機械的に設定ミスを見つける
- ・ チェック漏れを無くす
- ・ 設定間違いの経緯を調査する

低レイヤーをいかに機械的に制御するか！

# バグやデグレ

試験工程で発見された「バグ」や「不具合」は構築スケジュールに大きなインパクトを与えます。いかに早く、バグを解決し試験を再開させるのか？が運用工程においても重要なポイントになります。ODMメーカーのエンジニアと初期設計の段階から密な連携をとり、発生時にはパーツメーカーを含めた改修対応が重要です。また、改修されたプログラムにおいてデグレの発生率が非常に高い状況でした。これを改善するために、メーカーで実施する試験環境や試験方法を均一化しテストしなりを自動化することで、デグレ発生率を低減させています。

## ポイント

- ・ 試験環境をKDDIと同じにする
- ・ 試験工程を自動化、均一化する
- ・ エンジニアとの直接パスを作る

ここでの経験が運用工程で生かされる

# 自分達がサーバメーカーとなる

これまでの、サーバを購入して問題があればサーバメーカーへ問い合わせ、調査結果を確認して、試験を行うというような受け身の姿勢から。

パーツメーカー、ODMメーカーと一緒に改善、改修を行う体制が必要です。

これにより、利用しているサーバの部品一つ一つに責任をもち、くせや弱点を知ることにより、システム設計段階で弱点を補うことでシステム全体の安定を得ることができるのです。

独自のサーバを自ら設計する

①低価格で高品質な都合のいいHWがほしい

完了

②発注からリリースまで超短納期・自動化にしたい

③サーバの運用工数を削減したい



そこで

# サーバ構築 短納期と自動化

# KCPSでの構築工程



設計変更弱い

重要工程は人に頼る

## ① 設定・確認の自動化

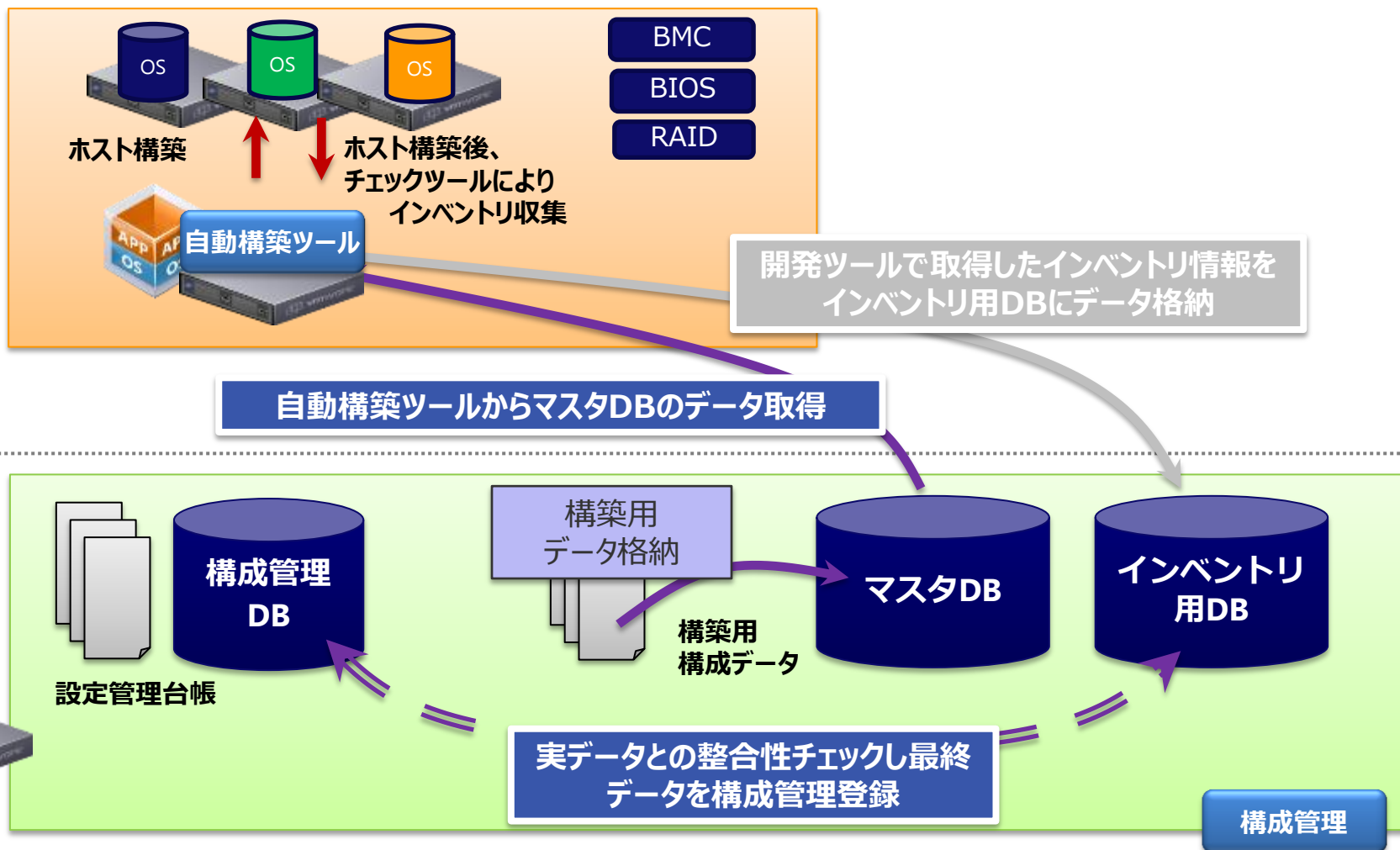
- BIOS, BMC, RAID, FWすべてをCLI・APIに対応

## ② OS導入自動化と高速化

- PXEを利用しないOS導入機能を搭載
- キックスタート廃止による設定変更の柔軟化
- 設定情報の自動取得・設定機能を実装

## ③ 試験・修復の自動化

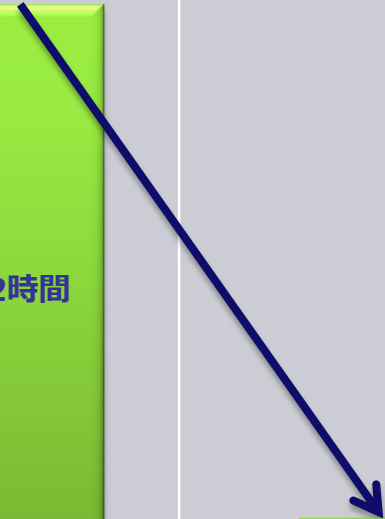
- HW試験から結合試験までを自動化
- 設定項目をリアルタイム取得確認変更を実装

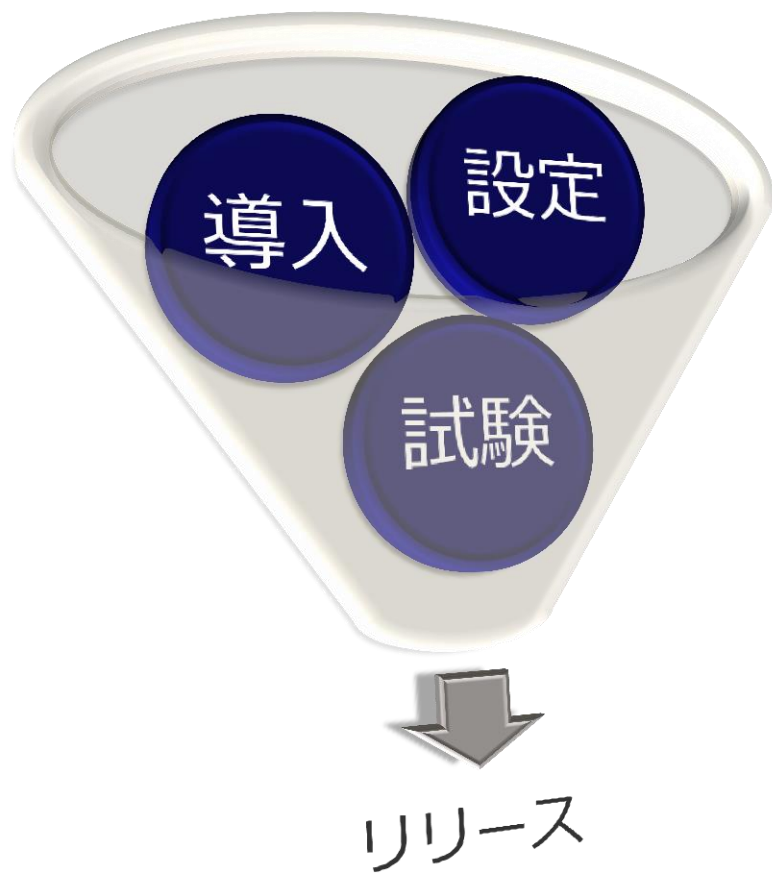


# 構築時間の比較

CONFIDENTIAL

ステップ	作業内容	手動	自動構築ツール
1	設定値自動取得 HWセルフテスト	20サーバの構築  約72時間  72時間	20サーバの構築  2時間未満  2時間
2	BIOS / BMC / RAID ファームウェアのアップデート		
3	BIOS / BMC / RAID の設定		
4	BIOS / BMC / RAID の設定値正当性確認 & 構成管理登録		
5	OSの展開 (各種OS対応)		
6	OSパラメータの設定変更		
7	OSの設定値正当性確認 & 構成管理DB保管		
8	OSセルフテスト 正常性確認		







①低価格で高品質な都合のいいHWがほしい

完了

②発注からリリースまで超短納期・自動化にしたい

完了

③サーバの運用工数を削減したい

道半ば



# 今後のOCPへのミッション

# Software-Defined Infrastructureを見据えて

今後、インフラ技術は抽象化されたハードウェアをソフトウェアによってリソースを様々な環境の変化に応じて動的にコントロールできる**SDI**に向っていく



ソフトウェアでコントロール出来ない超低レイヤを汎用的にコントロール可能にする

BIOS

RAID

BMC

etc

# HWレイヤと上位レイヤのつながりが必要



# HWに強い台湾だから「つなぎ」を全て実現出来る

ハードに必要な機能だけでなく  
システム構築に必要な機能を自由に実装することができる

これが 最大のメリットです。



自動構築ツール  
合同プロジェクト



### KCPSでのノウハウをOCPへ還元

- BMCの標準化
- BIOSの標準化
- サーバインフラ以外への拡大
- エンタープライズクラウド要素を追加  
→ODMサーバの標準化を推進

汎用的で自由な

インフラツール開発の共有を推進します。

# 日本仕様のOCP

## ①現行サーバとの違い

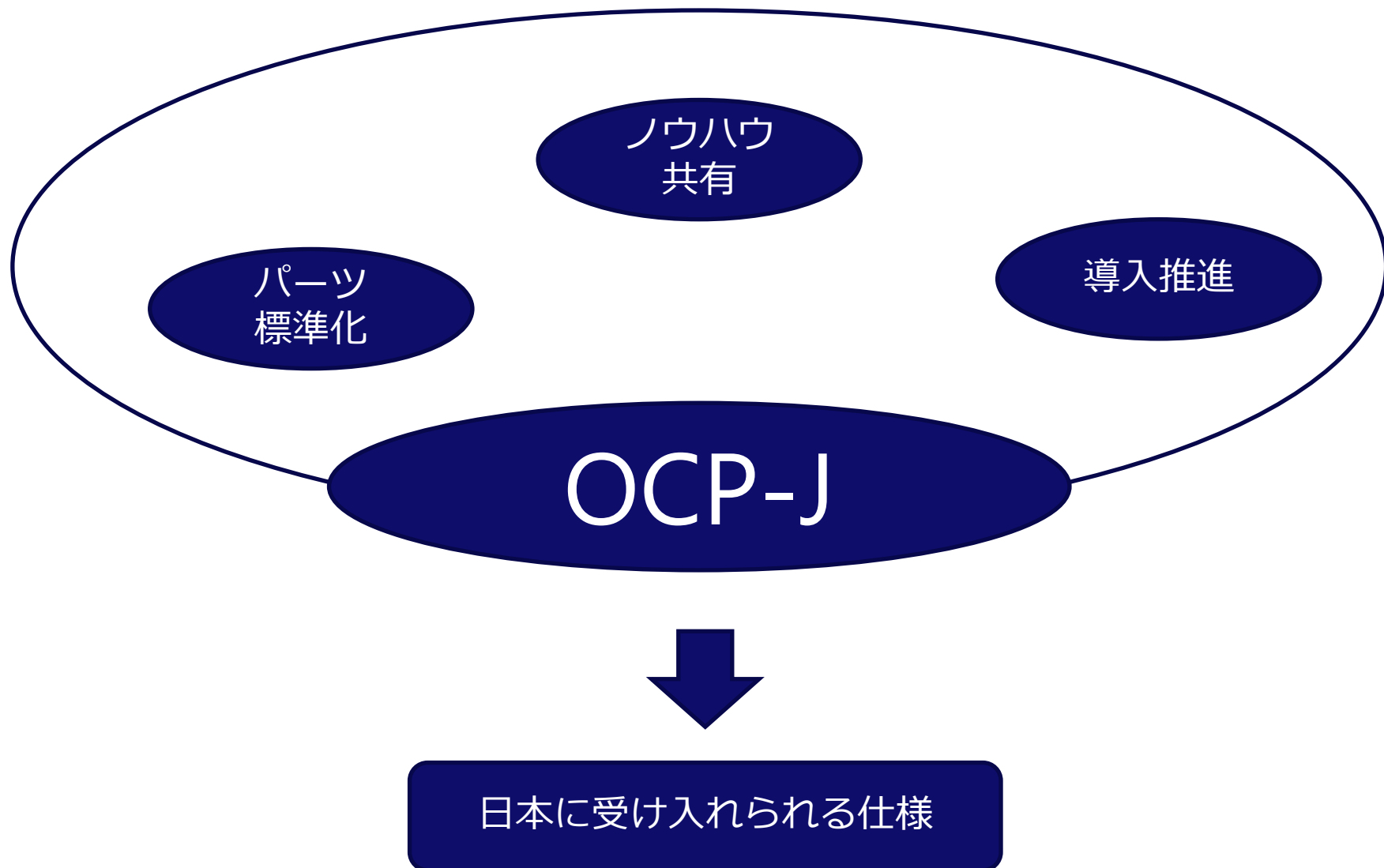
- ・ 運用時に利用する機能が少ない
- ・ 電源が特殊
- ・ ラック単位での導入が必要

## ②本当に日本にマッチしているのか？

- ・ 小規模構成においてコストメリットがあるか？
- ・ 扱える提案するSierがいるのか？
- ・ 日本のDCにマッチするのか？

結局 メガデータセンター向け





日本に受け入れられる仕様

# 一例として



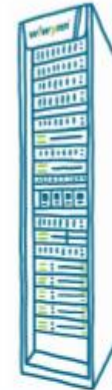
		SV324G2-2N	SV324G2-4N
<b>Processor</b>		Intel® Xeon E5-2600 V4	
<b>Node / CPU Socket</b>		2 / 4 (2 per node)	4 / 8 (2 per node)
<b>Node Specification</b>			
<b>Memory</b>		16 DIMM slots; Up to 512GB; DDR4 up to 2133 MT/s	
<b>Storage</b>	<b>3.5" Hot-plug Drive</b>	0	0
	<b>2.5" Hot-plug Drive</b>	6 SATA HDDs/SSDs	2 SATA HDDs/SSDs
	<b>M.2 SSD Module</b>	1	0
<b>Network</b>		One 1GbE LAN port	
<b>Expansion Slot</b>	<b>PCI-E x16</b>	0	1
	<b>PCI-E x8</b>	3	0
	<b>PCI-E x8 Mezzanine</b>	1	• 10/20/50 GbE ports via OCP mezz
<b>System Power Supply</b>		1600W (1+1)	
<b>System Dimensions (mm)</b>		2U; 87 (H) x 448 (W) x 800 (D)	

\* SV324G2 multi-node server will be available soon

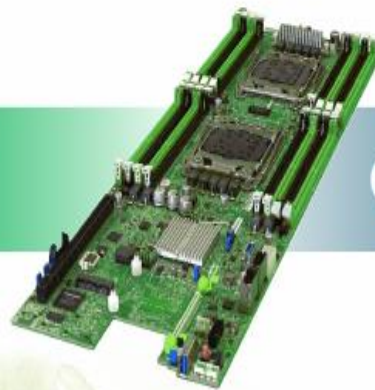
# From 21" to 19" (Leopard + Carmel)



21"



19"

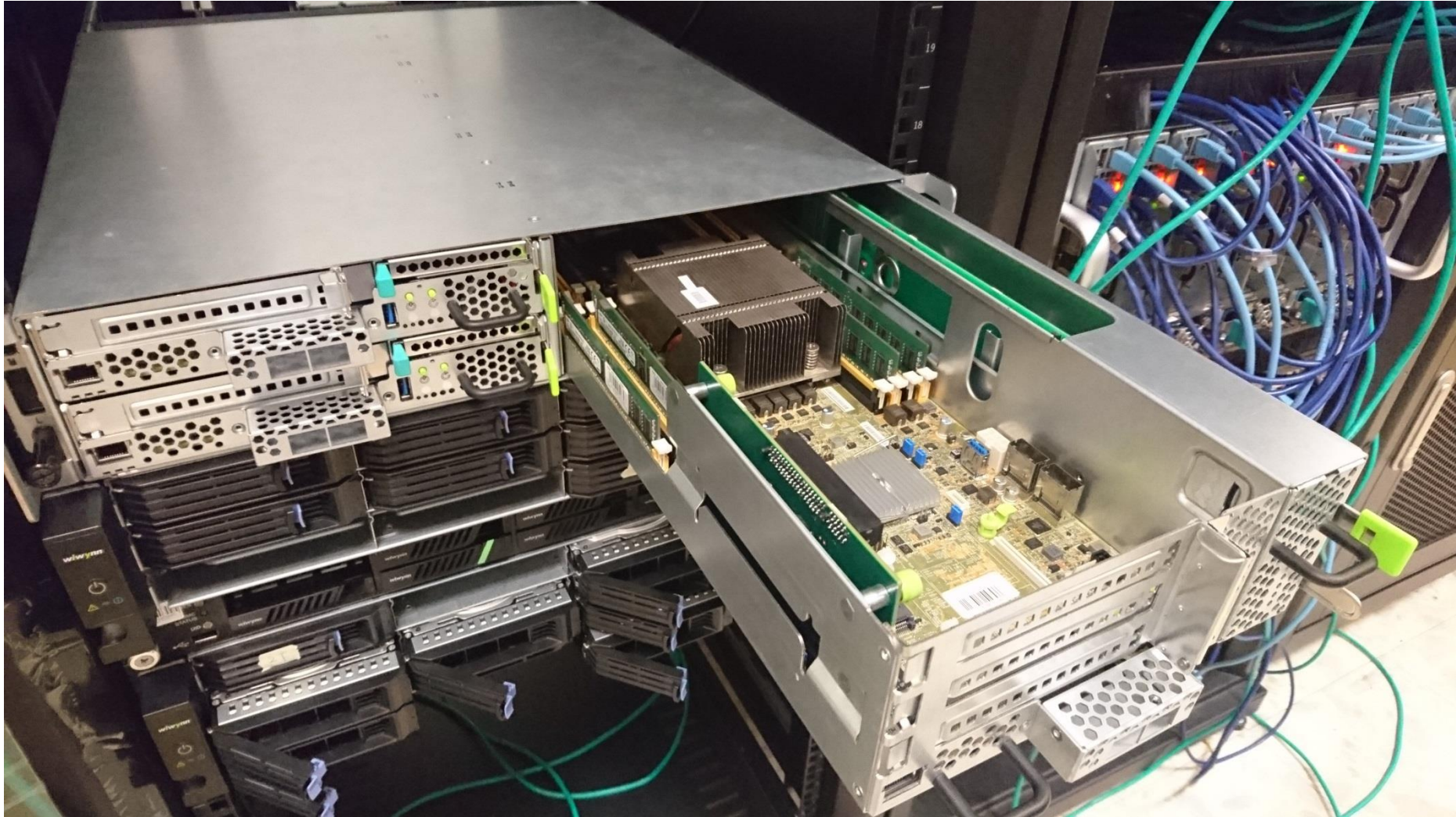


# 単独電源





# フロント



## Dashboard

Dashboard gives the overall information about the status of the device and remote server.

### Device Information

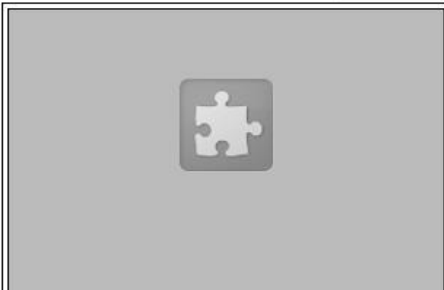
Firmware Revision: 2.16.65366  
 Firmware Build Time: Jul 23 2015 19:32:41 CST

### Network Information [\(Edit\)](#)

MAC Address: F8:0F:41:81:23:BE  
 V4 Network Mode: DHCP  
 IPv4 Address: 172.17.200.14  
 V6 Network Mode: DHCP  
 IPv6 Address: ::

### Remote Control

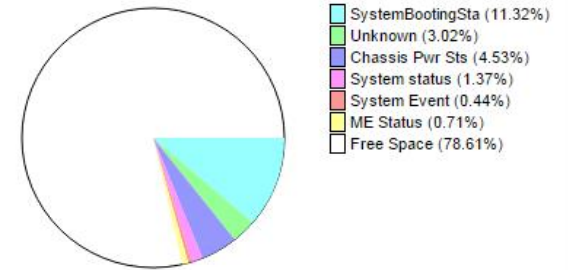
Launch



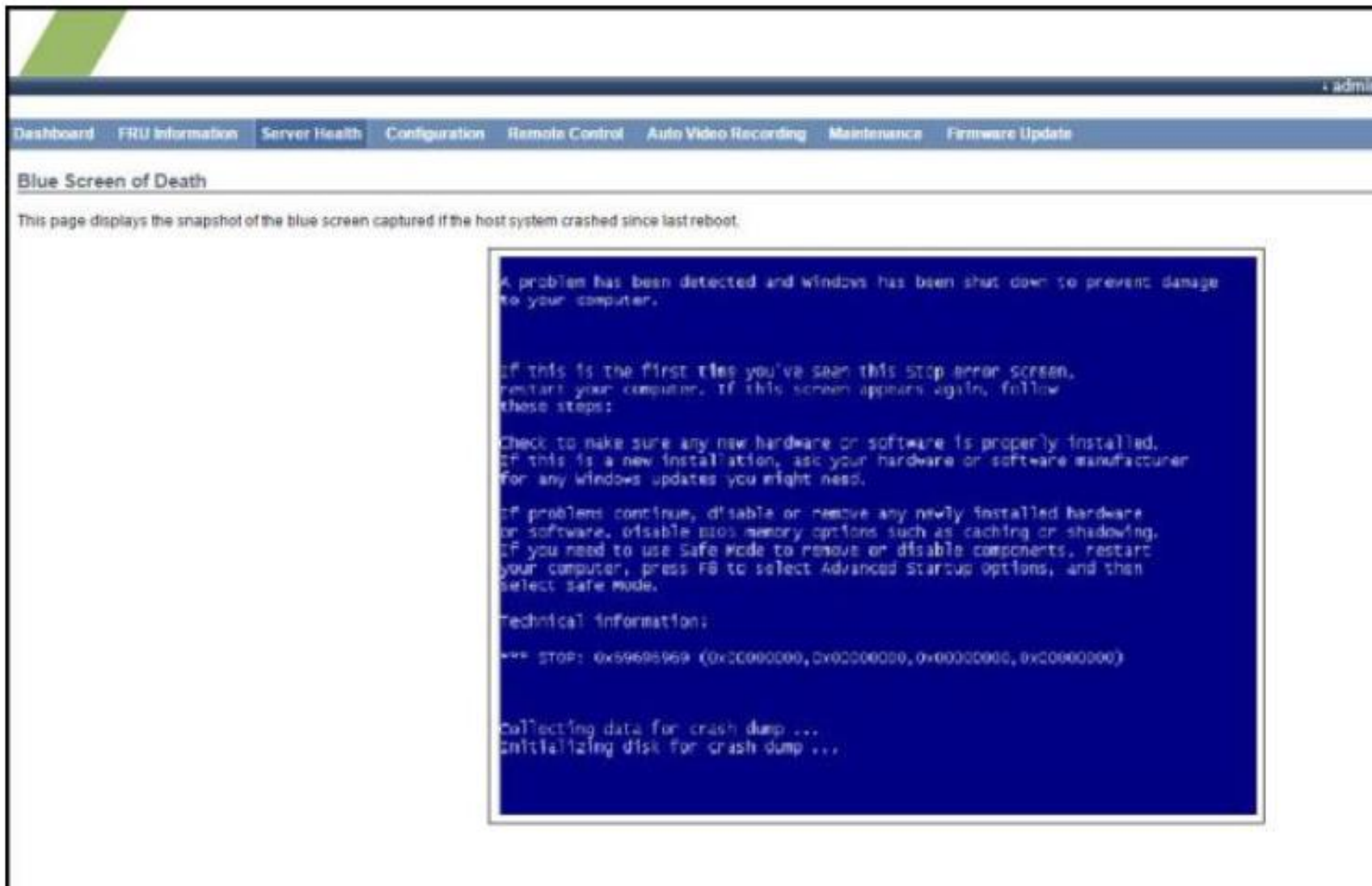
### Sensor Monitoring

Status	Sensor	Reading	
●	Outlet Temp	36 ° C	
●	P0 VR Temp	43 ° C	
●	P1 VR Temp	40 ° C	
●	Inlet Temp	30 ° C	
●	PCH Temp	46 ° C	
●	P0 Temp	42 ° C	
●	P1 Temp	40 ° C	
●	P0 DIMM VR0 Temp	41 ° C	
●	P0 DIMM VR1 Temp	45 ° C	
●	P1 DIMM VR0 Temp	41 ° C	
●	P1 DIMM VR1 Temp	40 ° C	
●	HSC Temp	33 ° C	
●	System status	0x8003	

### Event Logs



# ブルースクリーンをキャプチャー



The screenshot shows a web-based server management interface. At the top right, the user is logged in as 'admin'. The navigation menu includes: Dashboard, FRU Information, Server Health, Configuration, Remote Control, Auto Video Recording, Maintenance, and Firmware Update. The main content area is titled 'Blue Screen of Death' and contains the following text:

This page displays the snapshot of the blue screen captured if the host system crashed since last reboot.

```
A problem has been detected and windows has been shut down to prevent damage to your computer.

If this is the first time you've seen this STOP error screen, restart your computer. If this screen appears again, follow these steps:

Check to make sure any new hardware or software is properly installed. If this is a new installation, ask your hardware or software manufacturer for any windows updates you might need.

If problems continue, disable or remove any newly installed hardware or software. Disable BIOS memory options such as caching or shadowing. If you need to use Safe Mode to remove or disable components, restart your computer, press F8 to select Advanced Startup options, and then select safe mode.

Technical information:

*** STOP: 0x59695969 (0xC0000000, 0x00000000, 0x00000000, 0xC0000000)

collecting data for crash dump ...
initializing disk for crash dump ...
```



## List of Supported Host OS

- RHEL 5
- RHEL 5.3
- RHEL 5.4
- RHEL 6
- w2k3
- w2k8
- Windows 2008 R2
- Windows 2008 SP 2
- Win 2012 (64 bit)
- RHEL 4
- OpenSuse 11.2
- OpenSuse 10.x
- Ubuntu 8.10
- Ubuntu 9.10
- Ubuntu 11.04
- Ubuntu 11.10 Server
- Ubuntu Server 12.04 (64)
- SLES 11
- Debian 6
- CentOS 6.0





# Quality Cloud