

ディペンダブルな自律連合型クラウドコン ピューティング基盤の研究開発

加藤和彦

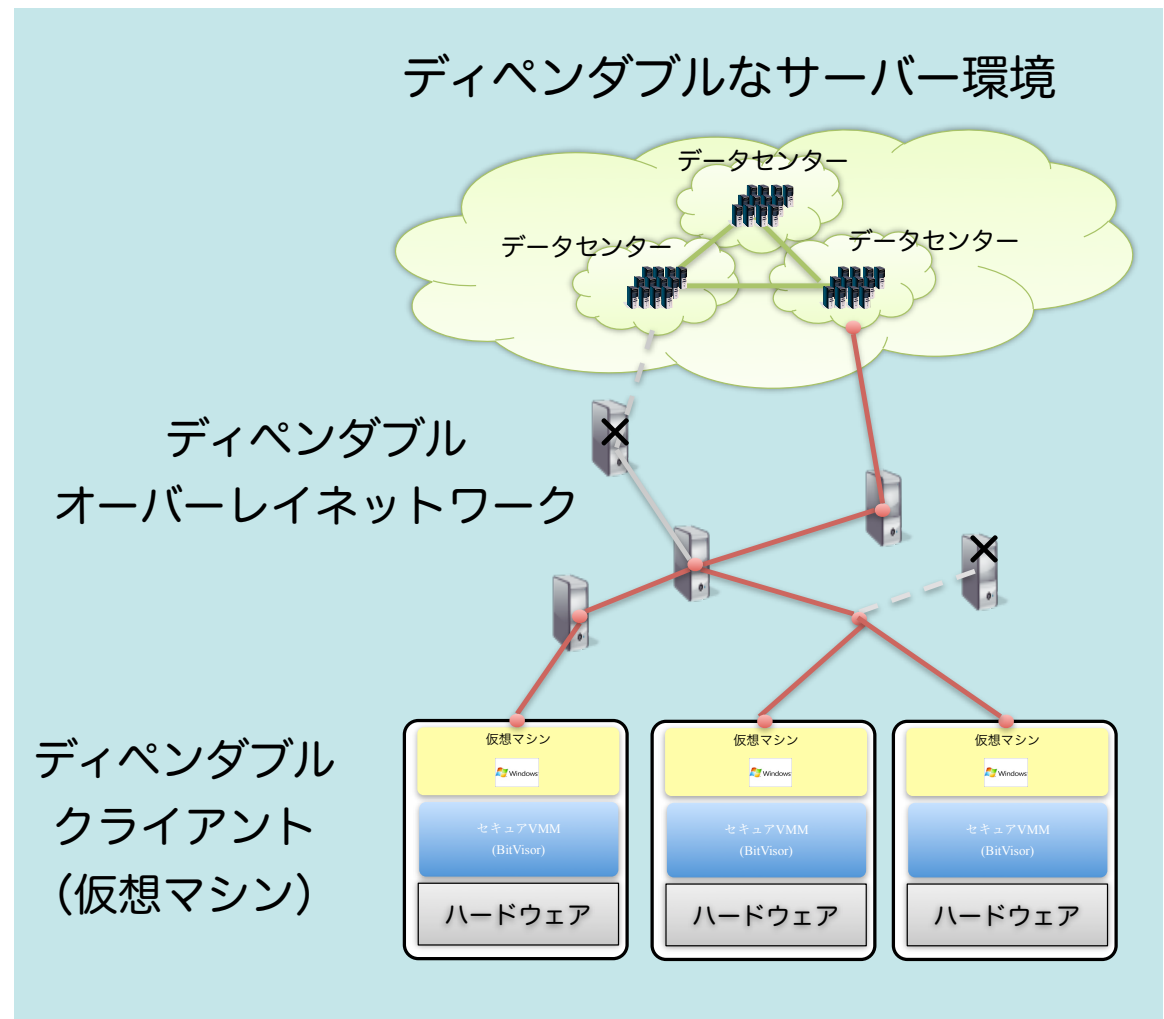
筑波大学大学院 システム情報工学研究科

ディペンダブルなクラウドコンピューティング基盤

クライアント環境，ネットワーク環境，サーバ環境における，ディペンダビリティを有するクラウドコンピューティング基盤技術の開発

ディペンダビリティ：

- ・ 信頼性
- ・ 可用性
- ・ 応答性
- ・ スループット
- ・ セキュリティ
- ・ プライバシー



主要メンバー

筑波大学

加藤和彦 教授, 杉木章義 助教, 長谷部浩二 助教
新城靖 准教授

東京大学

品川高廣 准教授 (2011/3まで: 筑波大学 講師,
2012/4~現在: 東京大学 准教授)

電気通信大学

大山恵弘 准教授

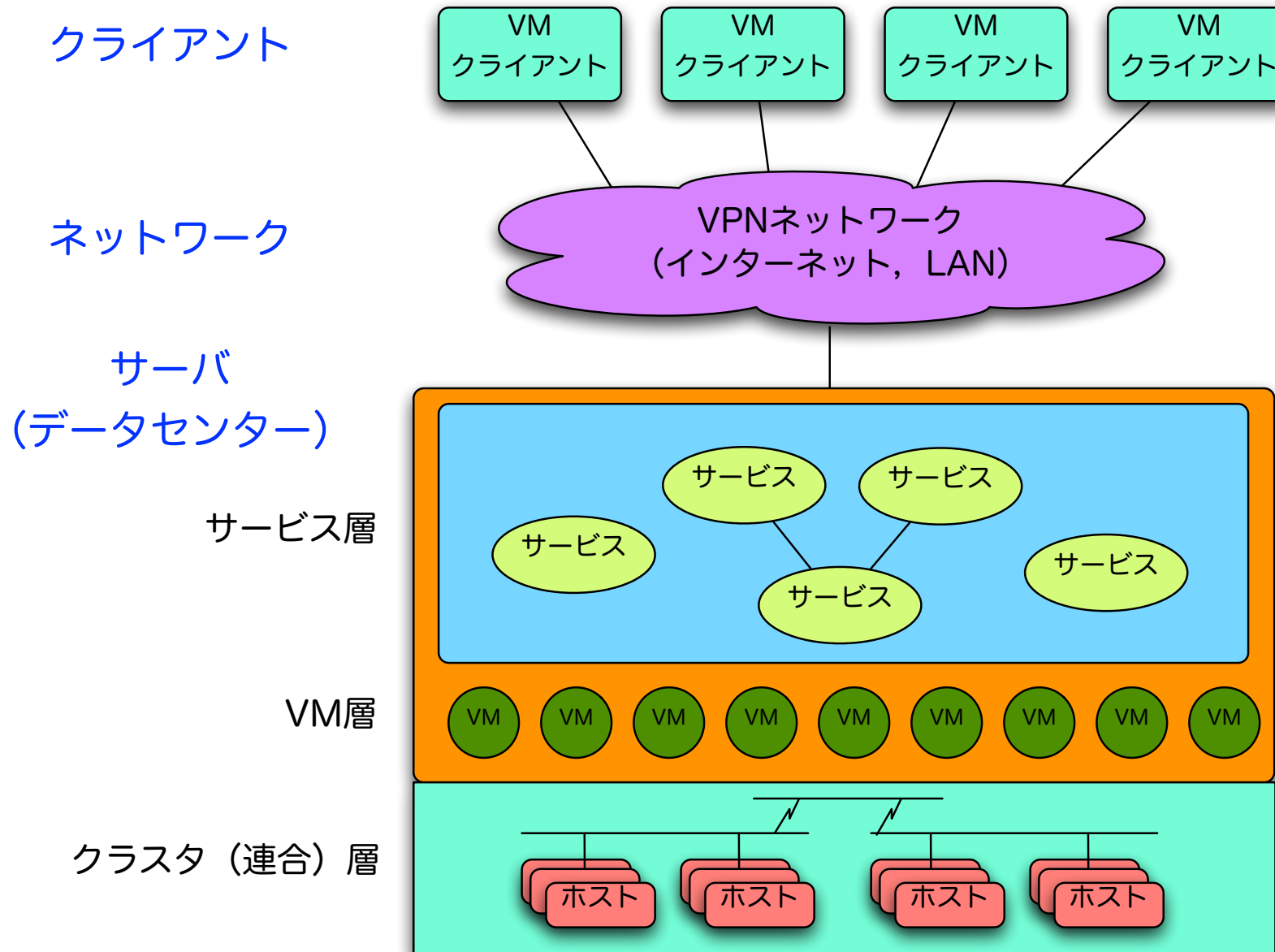
富士ソフト (株)

石井 嘉明, 矢野 恭平, 廣岡 誠之

ソフトイーサ (株)

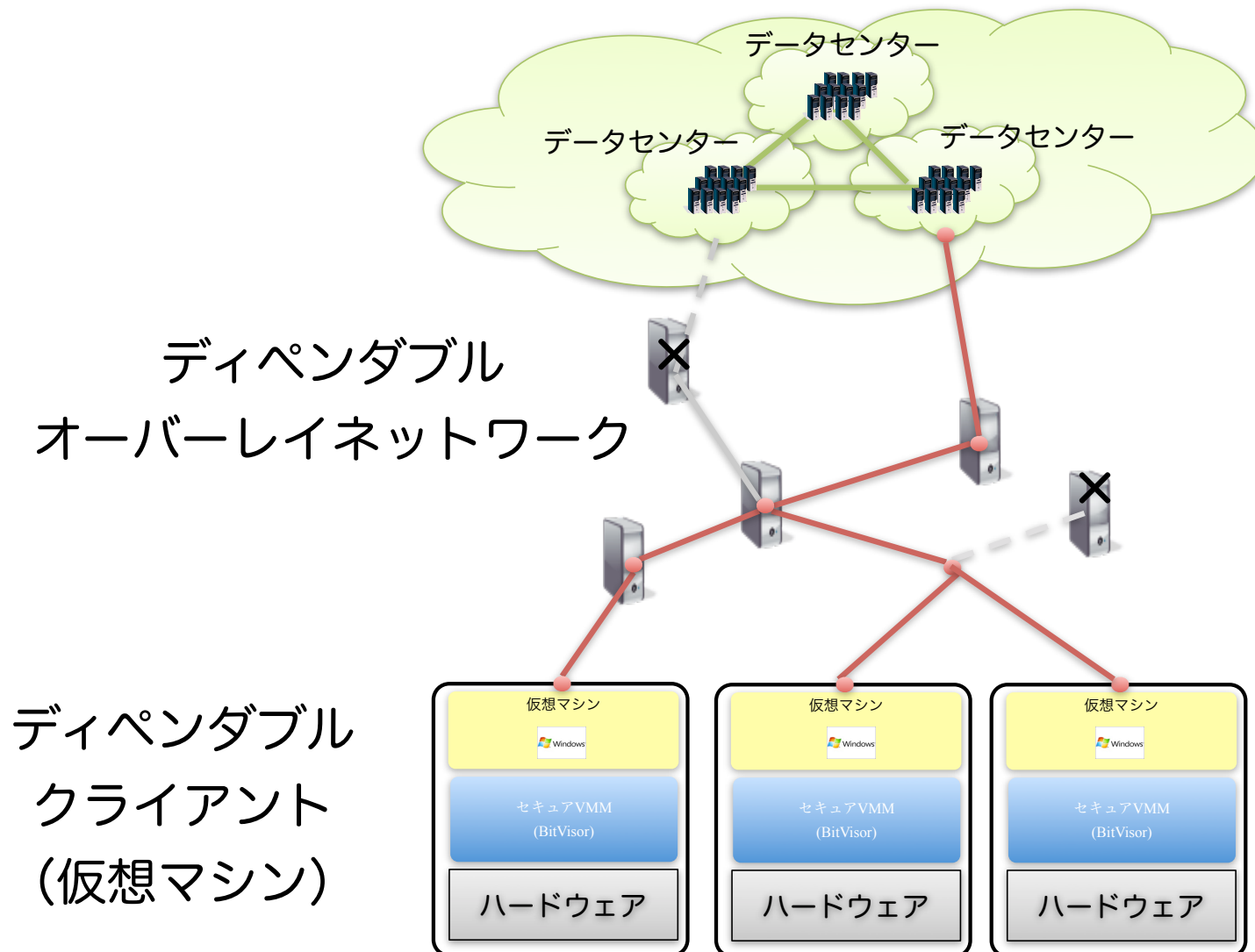
登 大遊

仮想マシンとクラウドコンピューティング

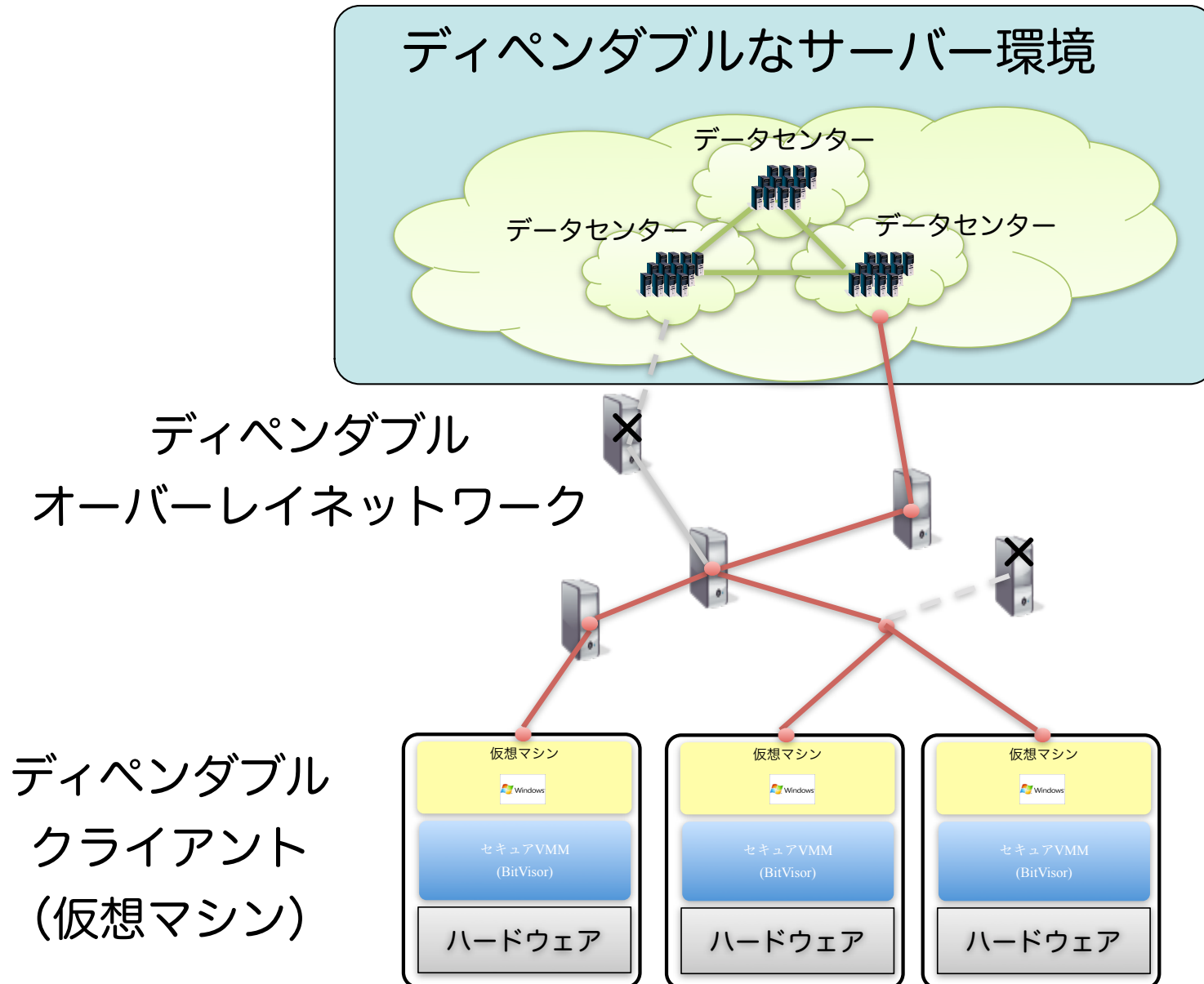


システムの全容

ディペンダブルなサーバー環境



(I) ディペンダブルなサーバー環境



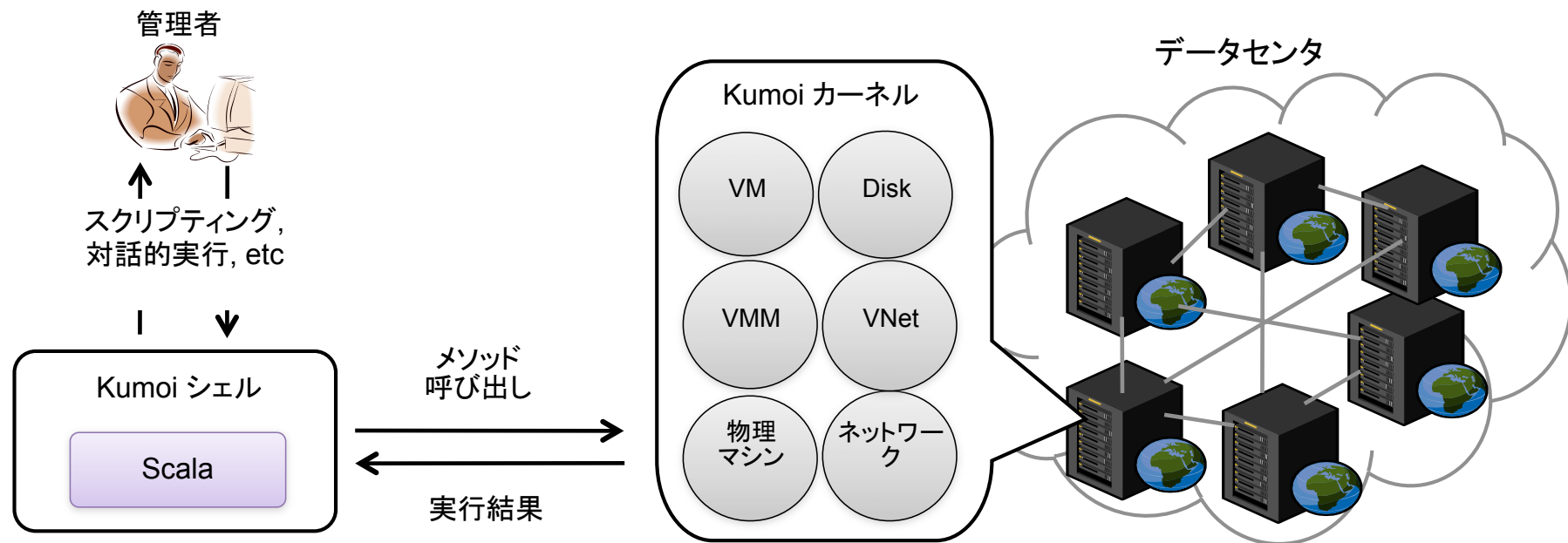
クラウド基盤ソフトウェア

Kumoi (雲居)

- Scala言語による記述：オブジェクト指向＋関数型
 - ✓ Java～Rubyレベルのスケラブルな記述，静的型検査
- オブジェクト／リソース マッピング
 - ✓ ハードウェア環境，仮想マシン，仮想ネットワーク等のシステムリソースを，ソフトウェアオブジェクトしてプログラミング可能とした。
- 分散並列資源管理アルゴリズムの内蔵
 - ✓ ゴシップ・アルゴリズム等。

オープンソースソフトウェアとして公開

Kumoiシステムの概要



ソースコード 現在3.4万行 (Scala言語)

Kumoiシエルの使用例

```
scala> pms.fliter(_.cpuRatio > 0.9).map(_.name)
```

pms: 利用可能な物理マシンのリスト

filter: フィルタ関数

map: マップ関数

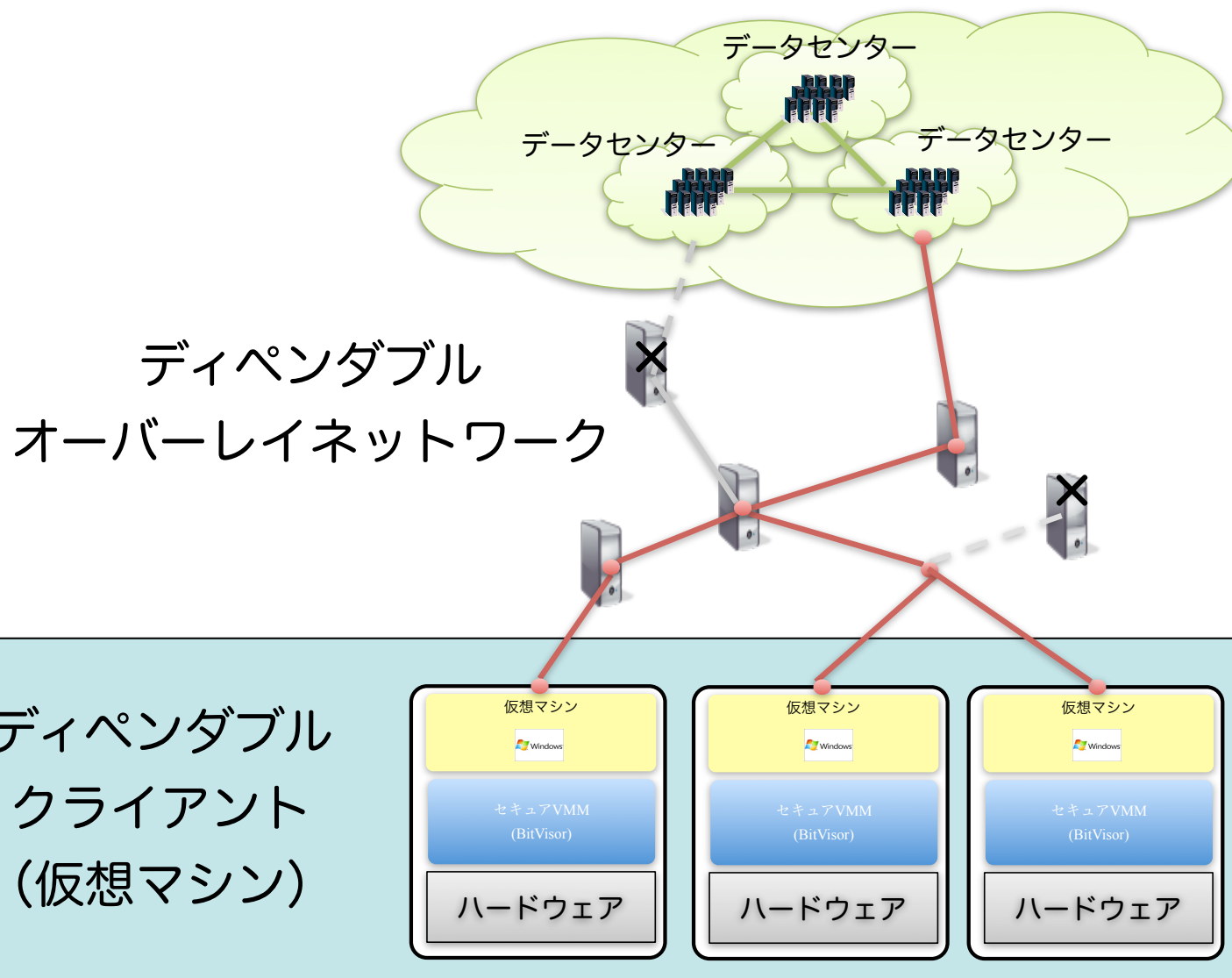
_: 高階関数の仮引数

Kumoiシステムプログラミング： VMコンパクション

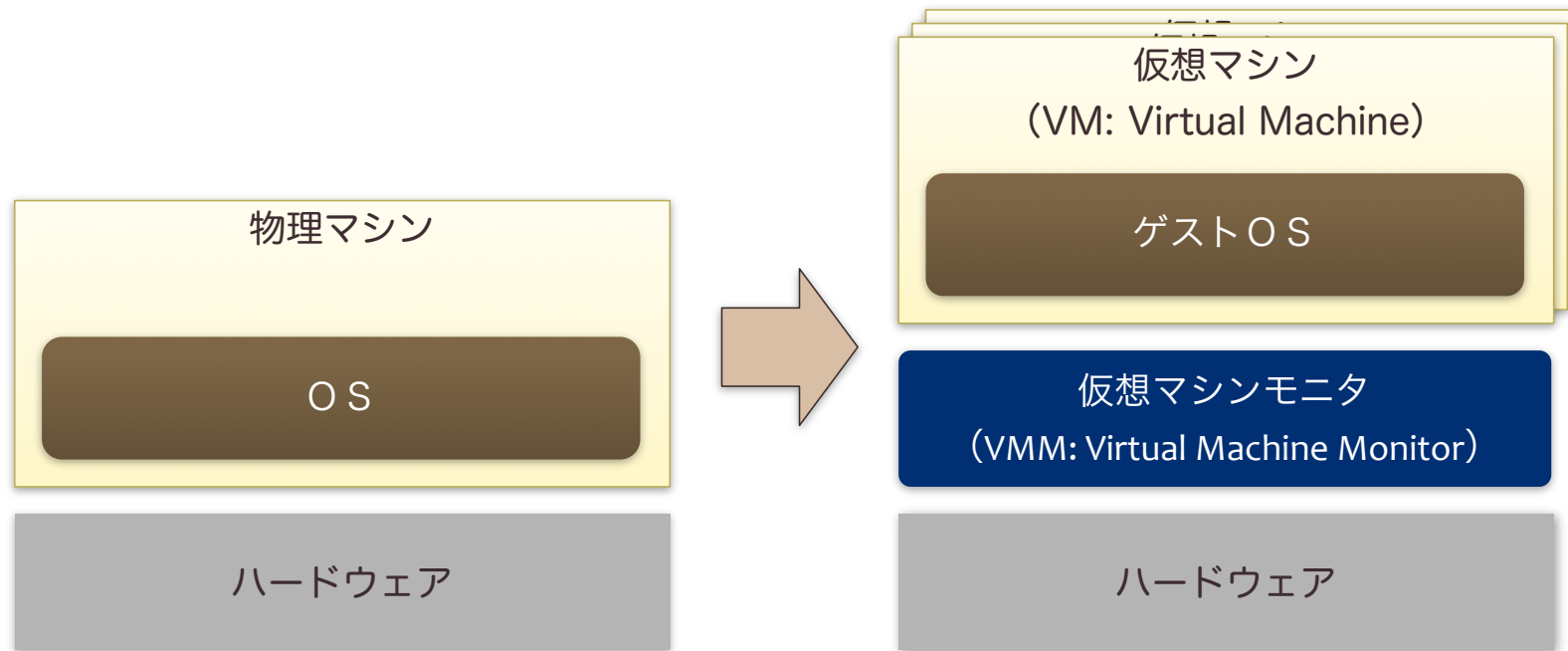
```
def compact(pms: List[VM]) {  
  def firstFit(v: VM, rest: List[VM]) {  
    rest match {  
      case h :: rs if h.cpuAvailable > v.cpuRatio => v.migrateTo(h)  
      case h :: rs => firstFit(v, rs)  
      case List() =>  
    }  
  }  
  def compacti(pms: List[VM]) {  
    pms match {  
      case h :: rest =>  
        h.vms.foreach(v => firstFit(v, rest.reverse))  
        compacti(rest)  
      case List() =>  
    }  
  }  
  compacti(pms.reverse)  
}
```

(II) ディペンダブル・クライアント

ディペンダブルなサーバー環境

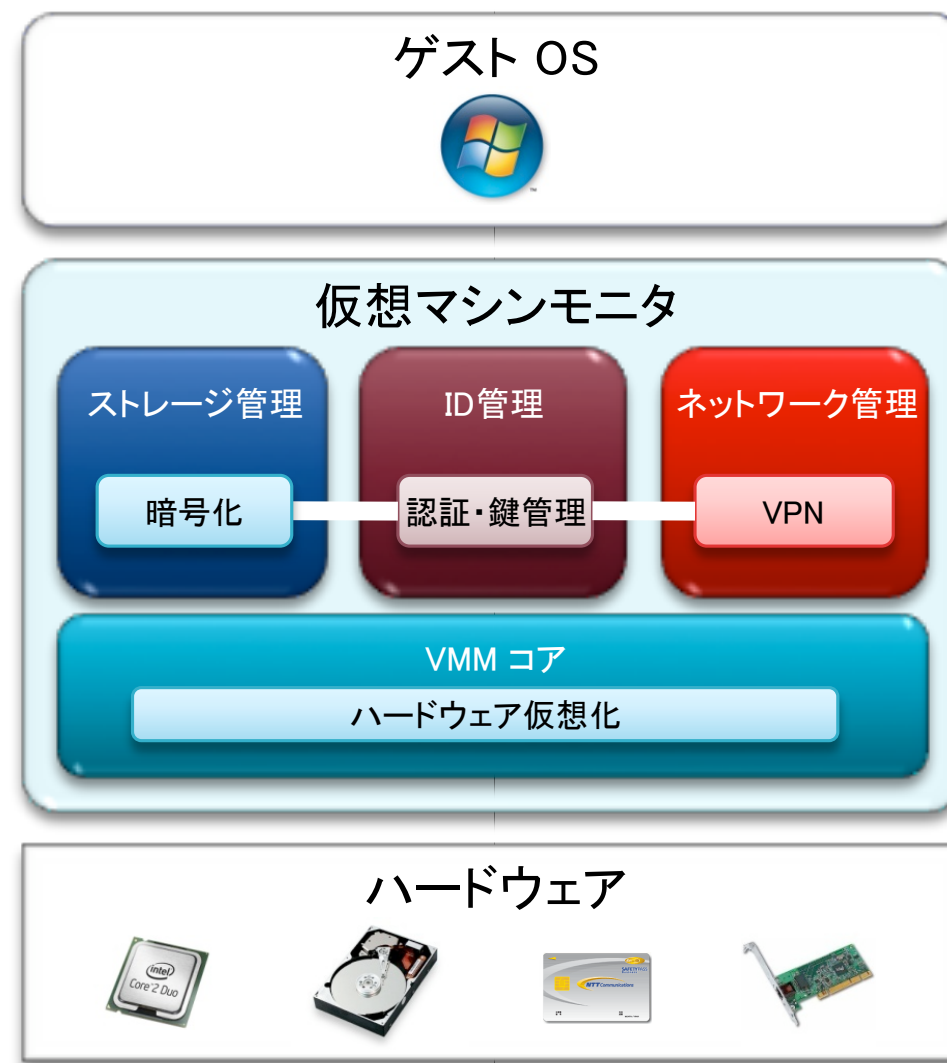


セキュア仮想マシンBitVisor



セキュア仮想マシンBitVisor

- ・ ストレージ管理
 - ✓ HDD及びUSBメモリの暗号化
- ・ ネットワーク管理
 - ✓ IPsecによるVPN接続
- ・ ID管理
 - ✓ ICカードによる鍵管理・認証
- ・ VMMコア
 - ✓ CPU及びメモリの仮想化

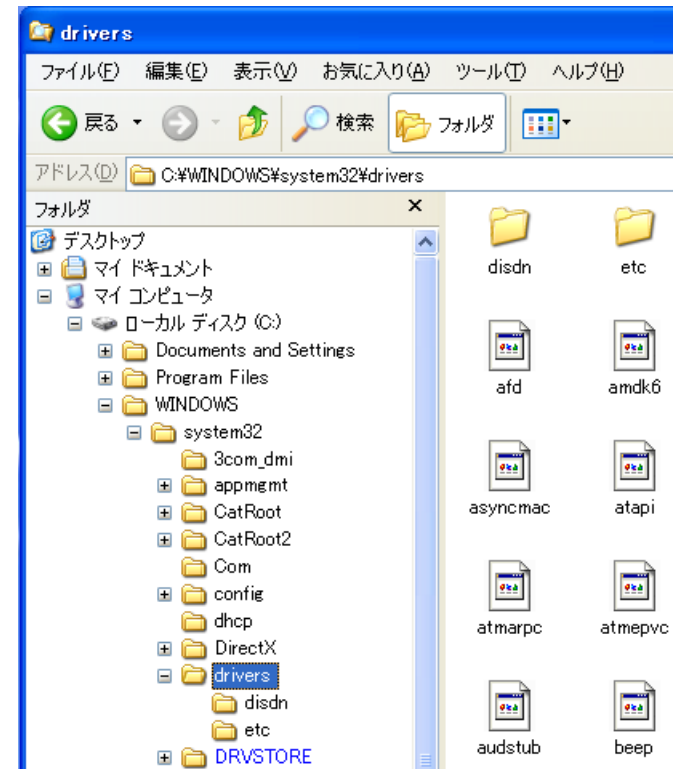


BitVisorの活用

- ゲストOSのシステムファイル保護
- マルウェア検出
 - ✓ VMM内でIDS
- 透過的VPN切り替え（次節）

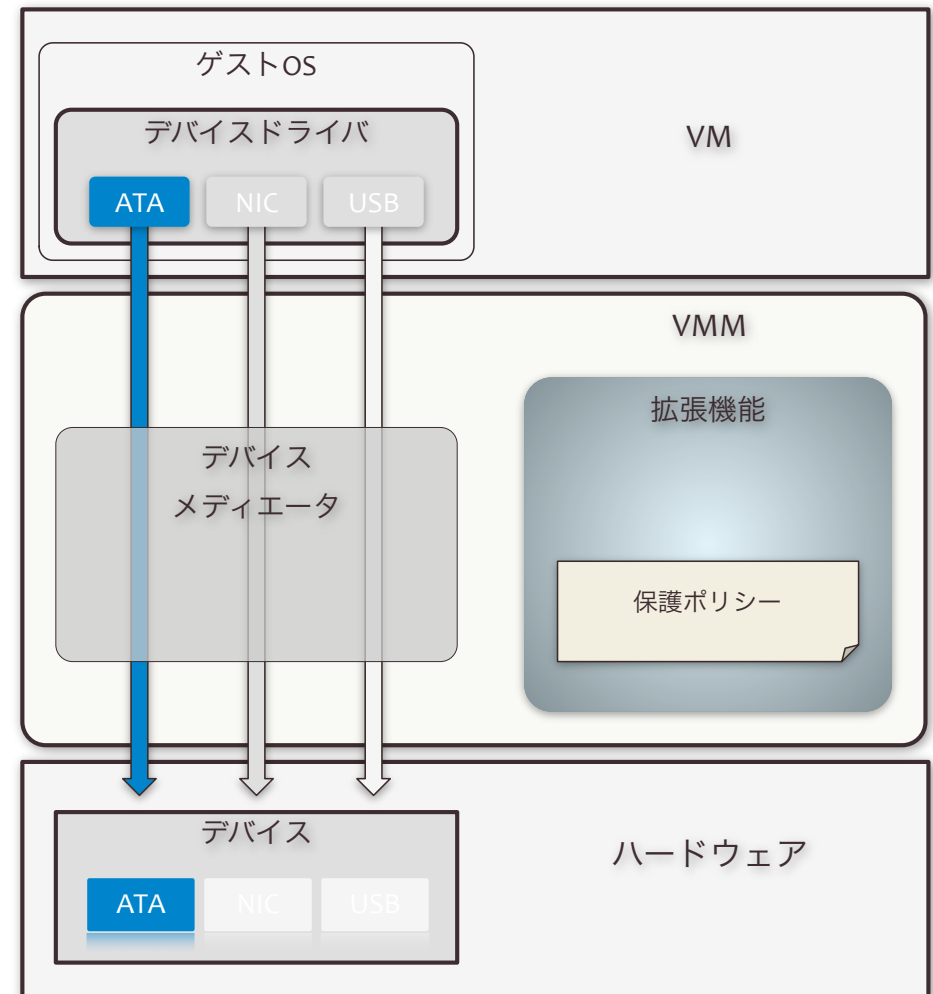
BitVisorを用いたゲストOSの システムファイル保護

- システムファイルを完全に保護する
 - ✓ カーネル・イメージ
 - ✓ デバイスドライバ等
- セキュリティソフトウェア
 - ✓ アンチウィルス等



システムファイル保護の実現法

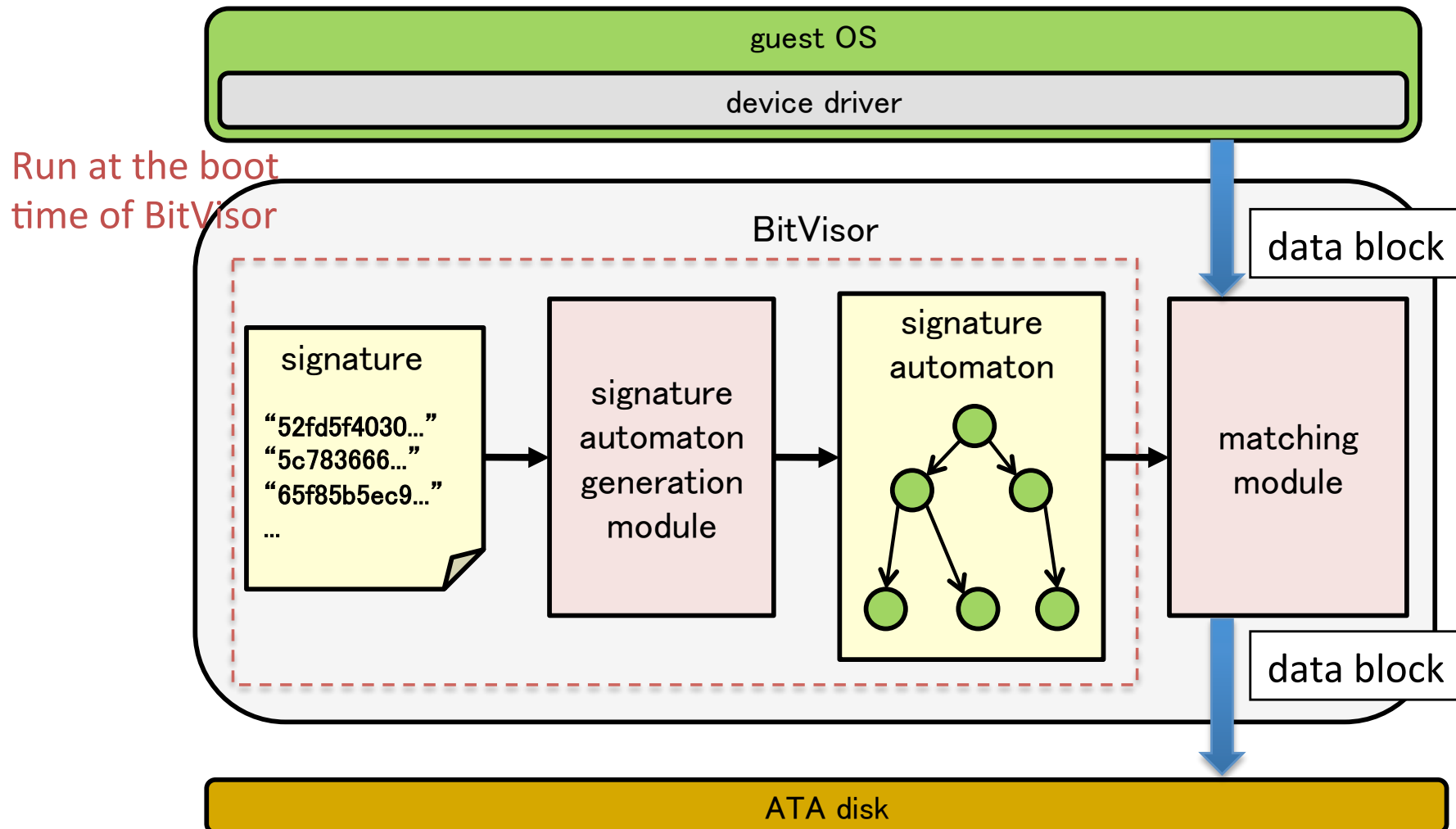
- BitVisorでストレージアクセスを監視
 - ✓ システムファイルへの書き換えを検出
 - ✓ ファイルのデータ領域
 - ✓ 管理データ領域
 - ✓ ディレクトリ構造
- ファイルとセクタのマッピングを管理
 - ✓ セマンティックギャップを克服



BitVisorでマルウェア検知：動機

- 既存のアンチ・マルウェアの問題点
 - マルウェアがカーネル乗っ取りに成功されると、稼働停止とされる可能性がある。
 - ユーザがアンチ・マルウェアを停止させる可能性がある。（故意に／ミス操作）

BitVisorでマルウェア検知：実現



BitVisorでマルウェア検知：実験

- 実験に用いたMalware
 - Linux: Bliss.A, Bliss.B, Linux.Caveat, Linux.Cassini.1618, Linux.Dido-478, Linux.Diesel.969.elf, Linux.Gildo, Linux.Godog.A, Linux.Godog.C, ...
 - Windows: CERE1482, MIRC.Bee, NT.Chatter, Trivial.64.C, Trojan.IWorm.MP.Virus, Trojan.SdBot-459, VGEN.32562.275, ...
- 結果：大部分を検出
 - 42 out of 44 Linux malware
 - 166 out of 169 Windows malware
 - BVMD could not detect malware for which ClamAV provides complex signatures (regular expression and file hash)
 - Linux: Linux.Caveat.a, Linux.Nel.a
 - Windows: W2K.Team, W32.Fugo, Worm.Zaka.A

BitVisorでマルウェア検知： オーバーヘッド

	physical machine	BitVisor	BVMD (30K sigs)
OS boot (s)	43	55	80
kernel build (s)	296	398	473



+34%



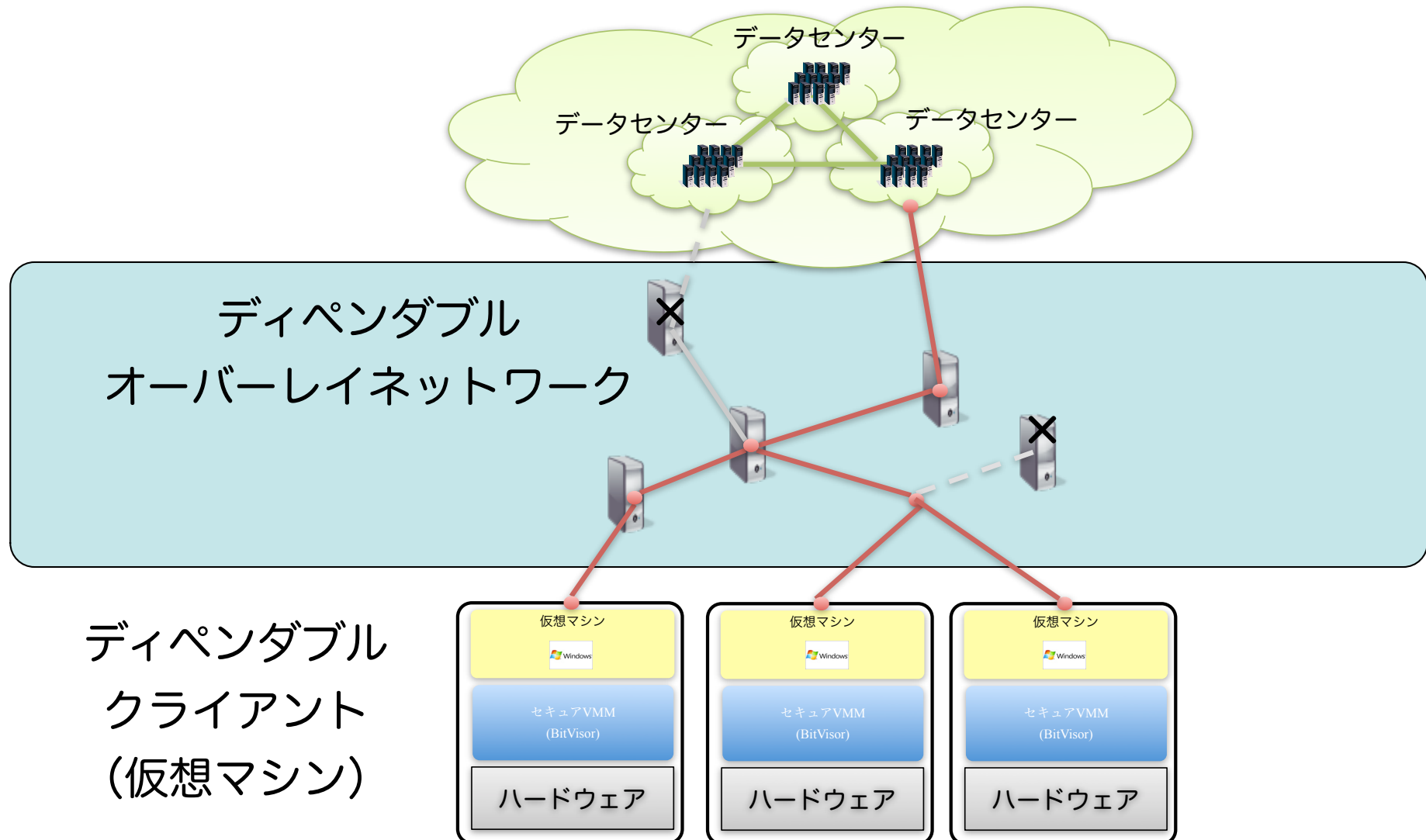
+19%

Platform:

- Intel Core 2 Duo P8600 2.4 GHz
- 2GB memory (1GB for BitVisor)

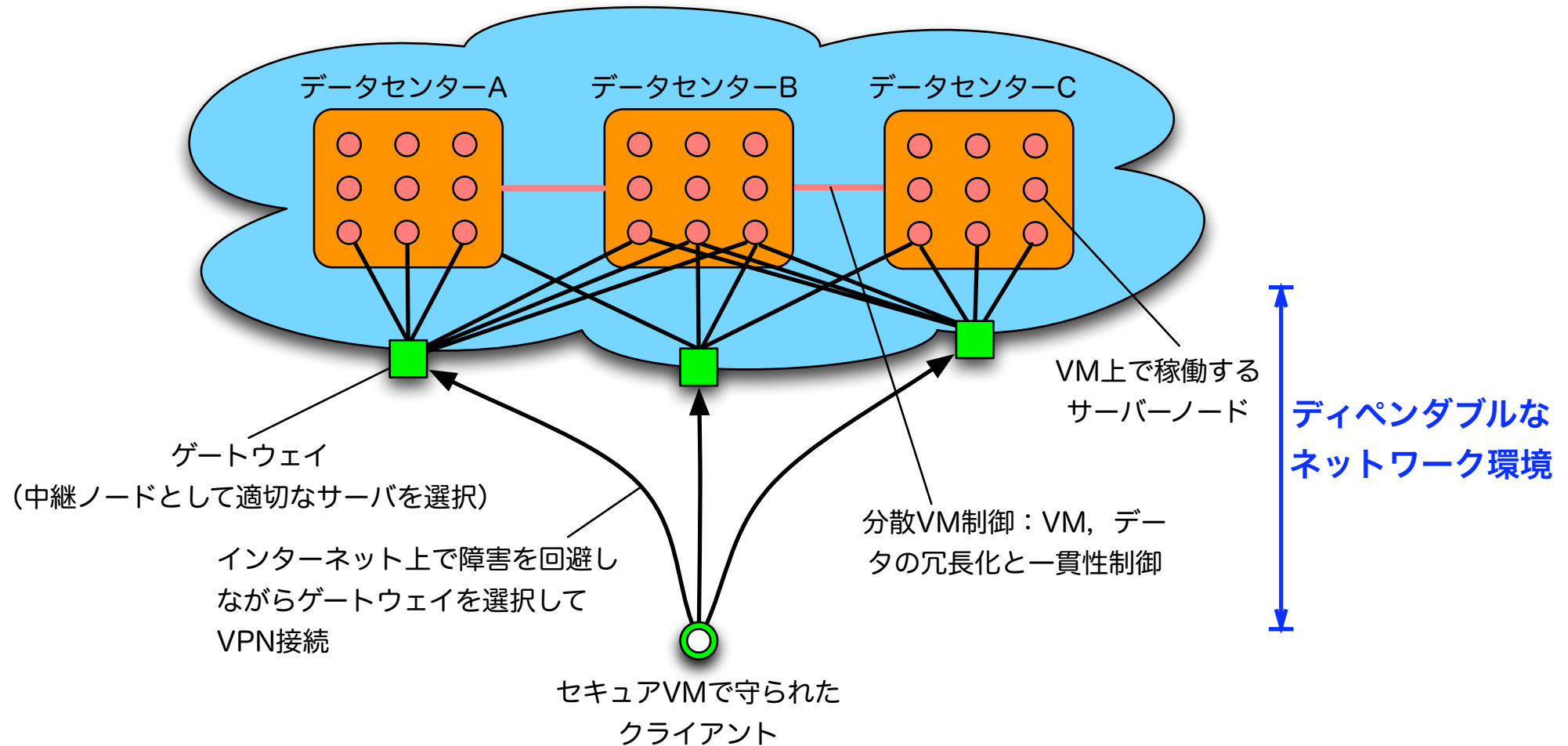
(III) ディペンダブル・オーバレイネットワーク

ディペンダブルなサーバー環境



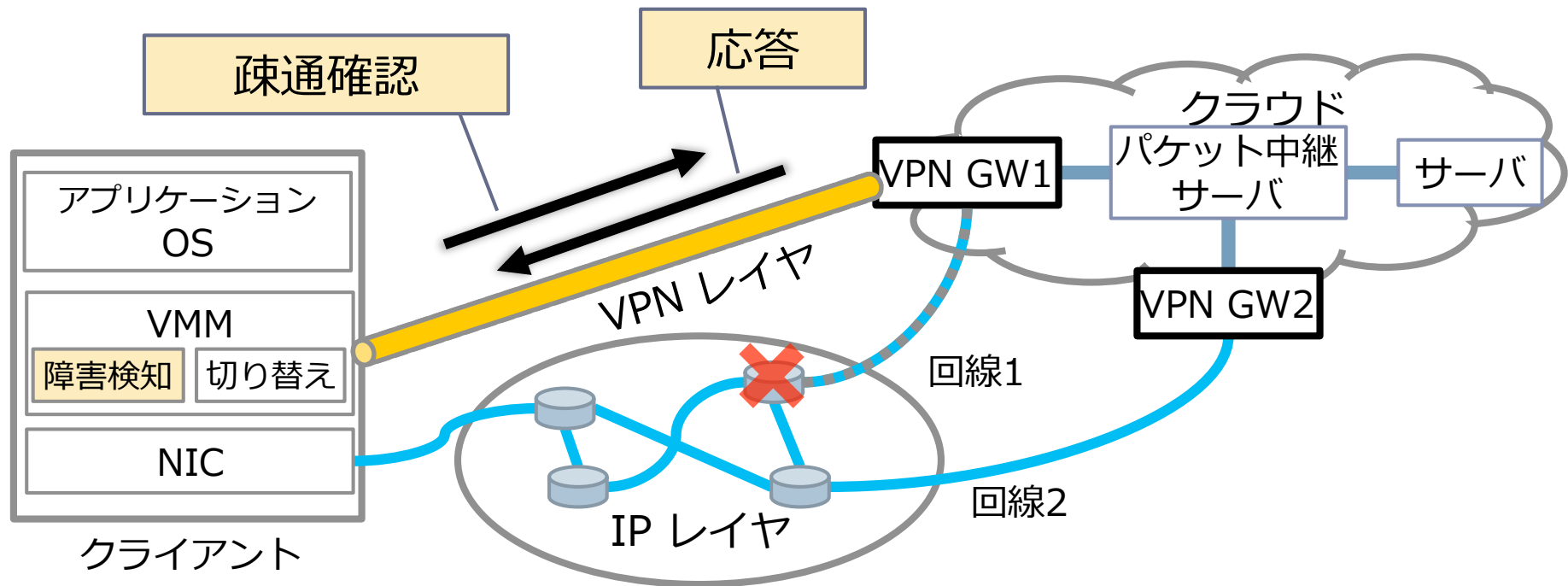
ディペンダブルなネットワーク環境

ディペンダブルなクラウドコンピューティング基盤



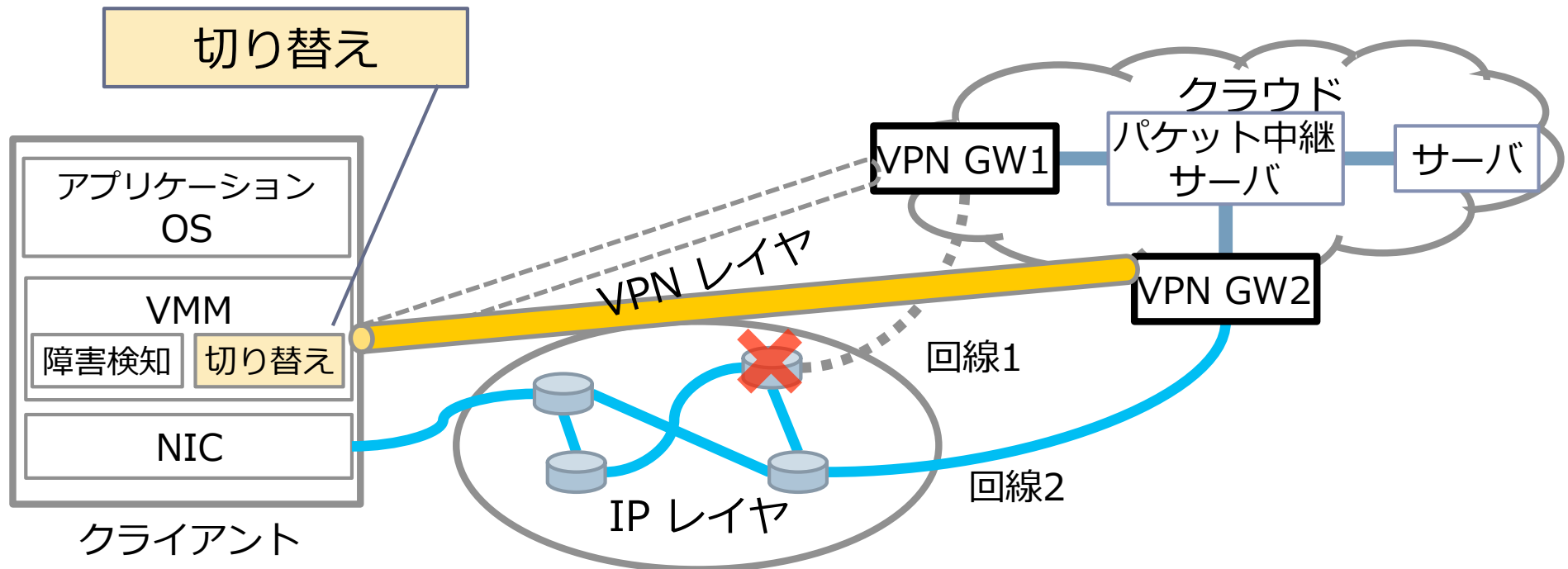
仮想マシンによるVPN自動切り替え1

通信と障害検知



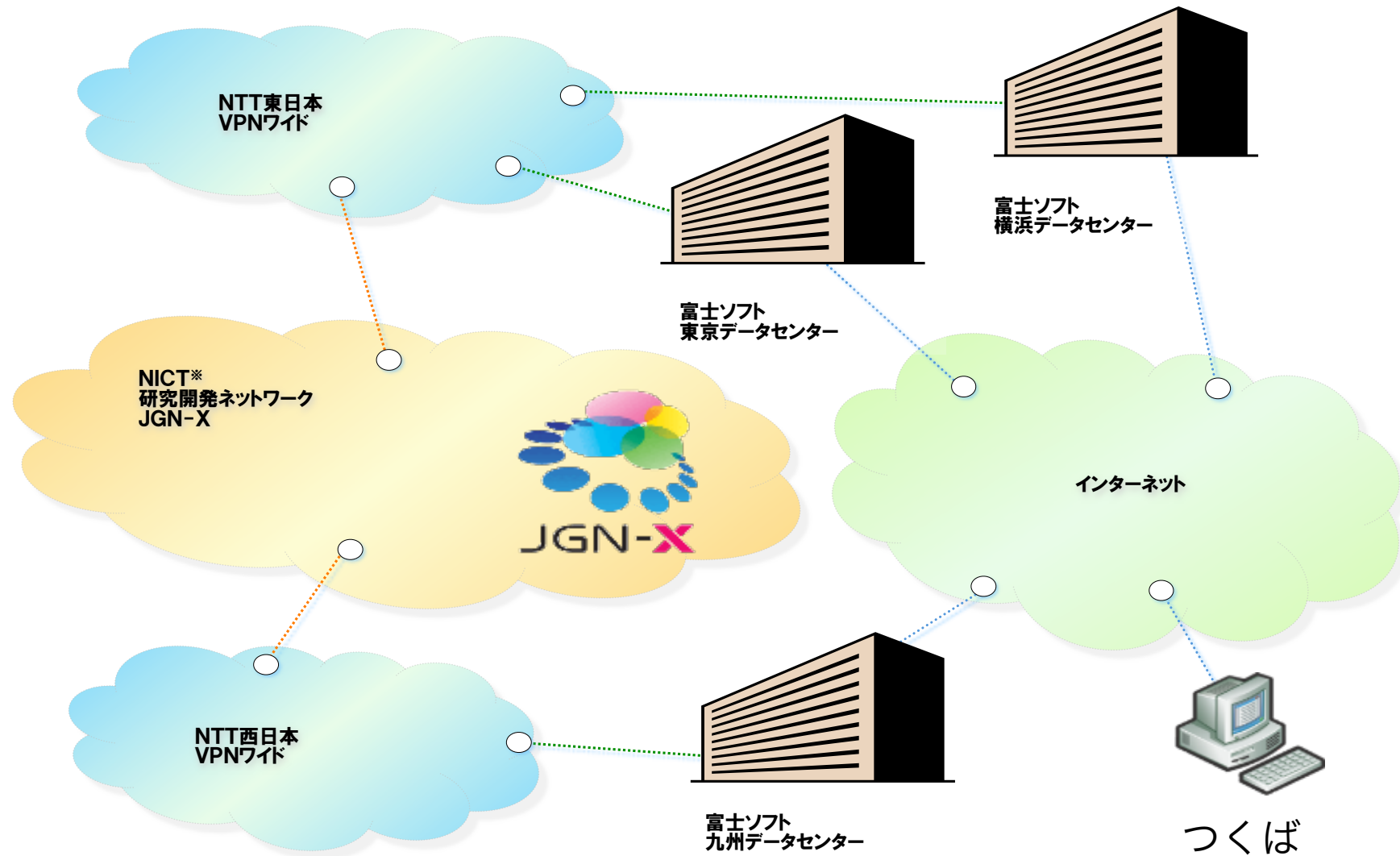
仮想マシンによるVPN自動切り替え2

VPN切り替え



JGN-X上でOpenFlowの利用を準備中

広域データセンターを用いた実証実験



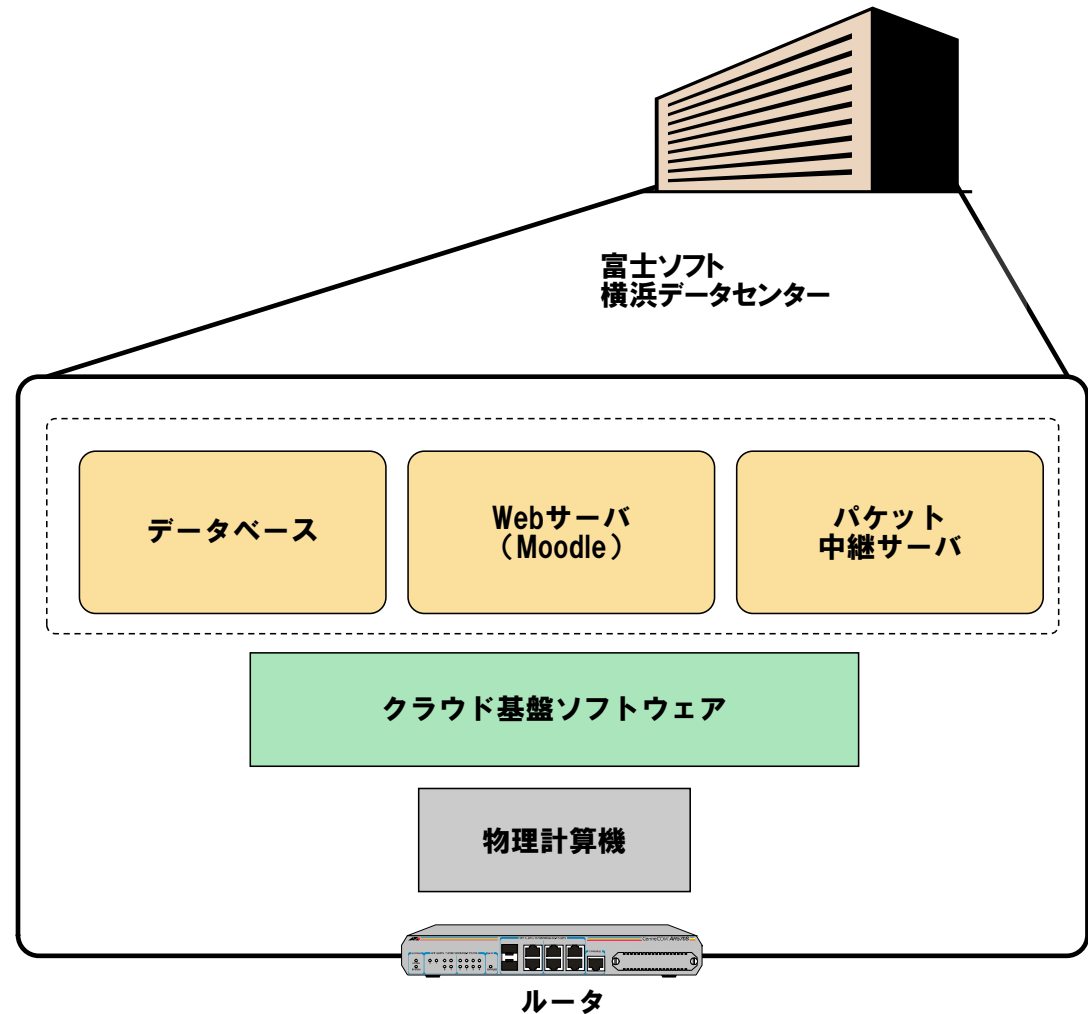
各データセンターのシステム構成

□ハードウェア

- * 物理計算機： HPブレード 6 台
- * ルータ： YAMAHA RTX1200

□ソフトウェア

- * OS： CentOS5.5
- * 仮想マシンモニタ： KVM
- * クラウド基盤ソフトウェア： Kumoi



VPN切換えに要する時間

「つくばー東京」 (56Km)から (直線距離)

「つくばー横浜」 (84Km)に切換

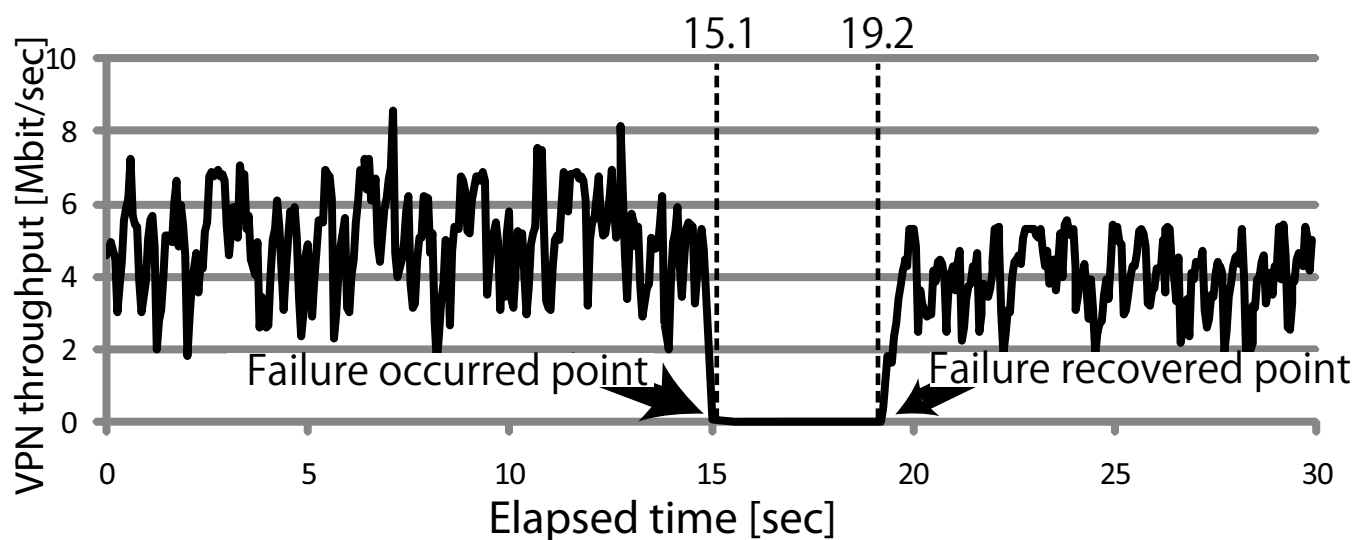


Figure 8. Throughput Transition over Failure

VPN切換えシステムの ネットワーク遅延とスループット

つくば—東京：56Km
つくば—横浜：84Km
つくば—福岡：926Km

(直線距離)

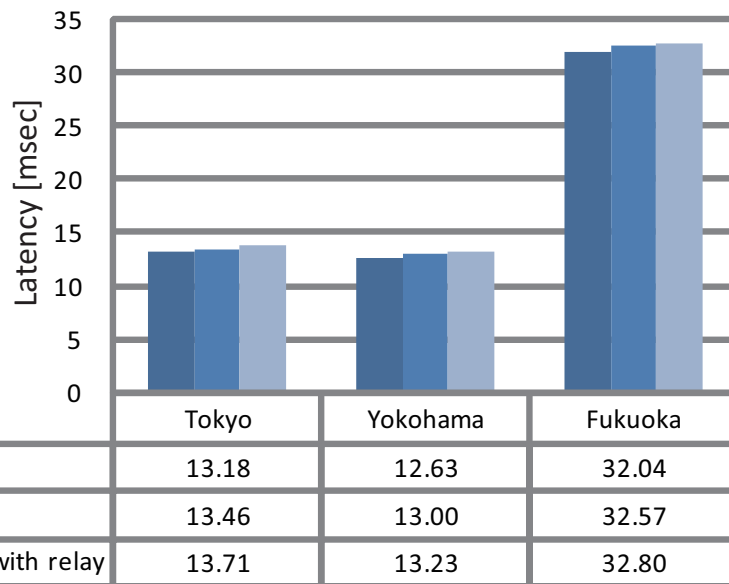


Figure 9. Latency

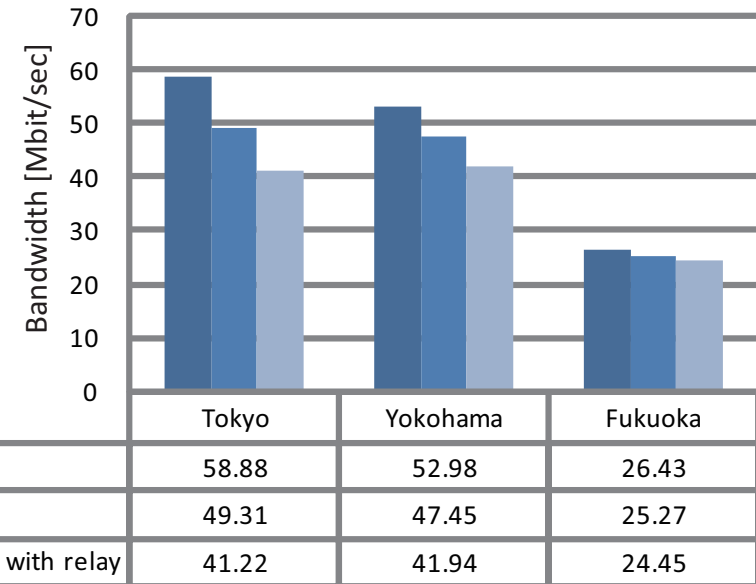


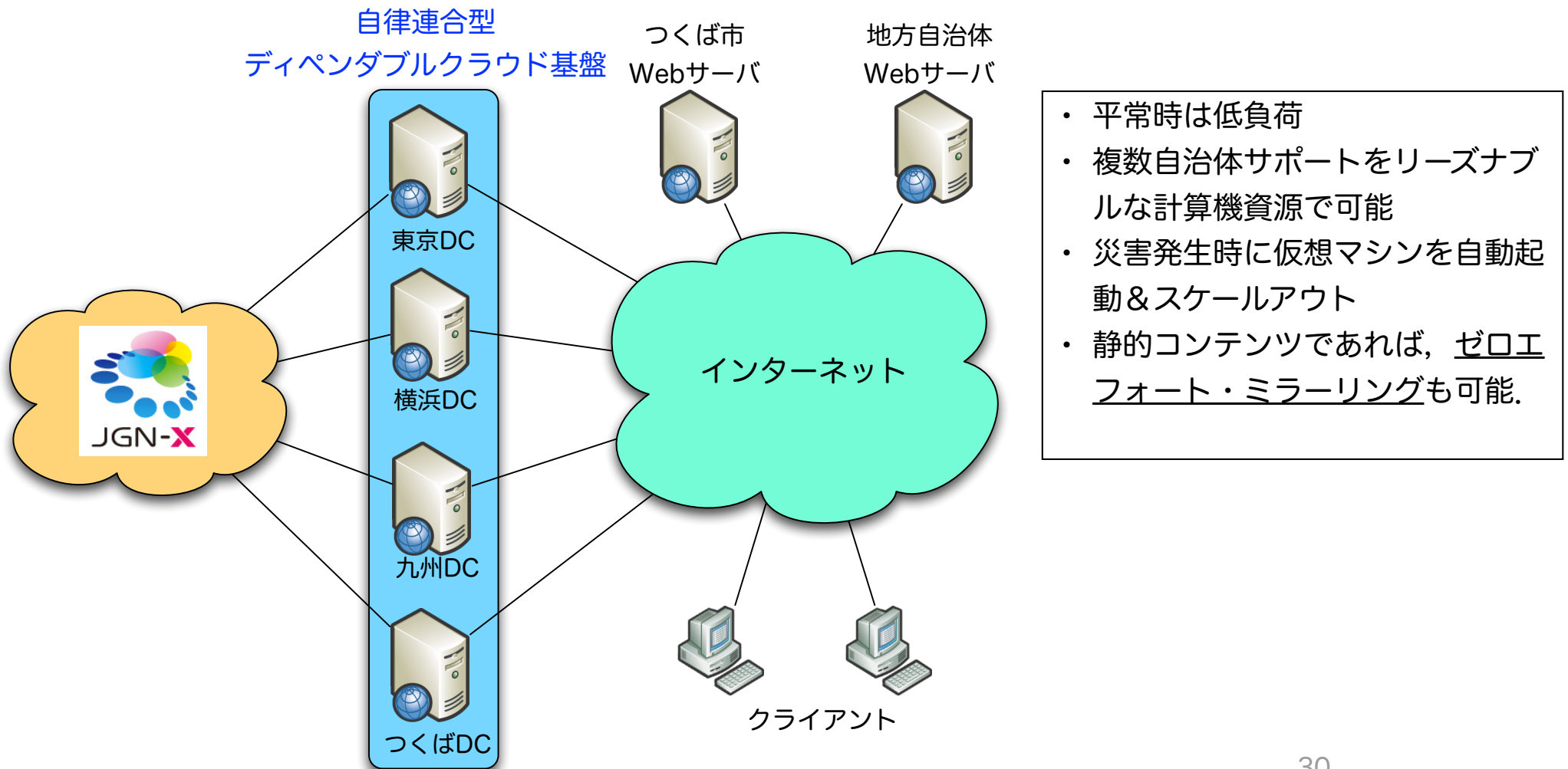
Figure 10. Throughput

活用

- ネットワークサービスのHA (highly-availability), オートスケール, 予見化
 - ✓ Moodle (オープンソースeLearningシステム)
- 災害時対応Webサーバ・ミラーリングシステム

災害時対応Webサーバ ミラーリングシステム

筑波大学加藤研究室 + (独) 情報通信研究機構(NICT) + つくば市

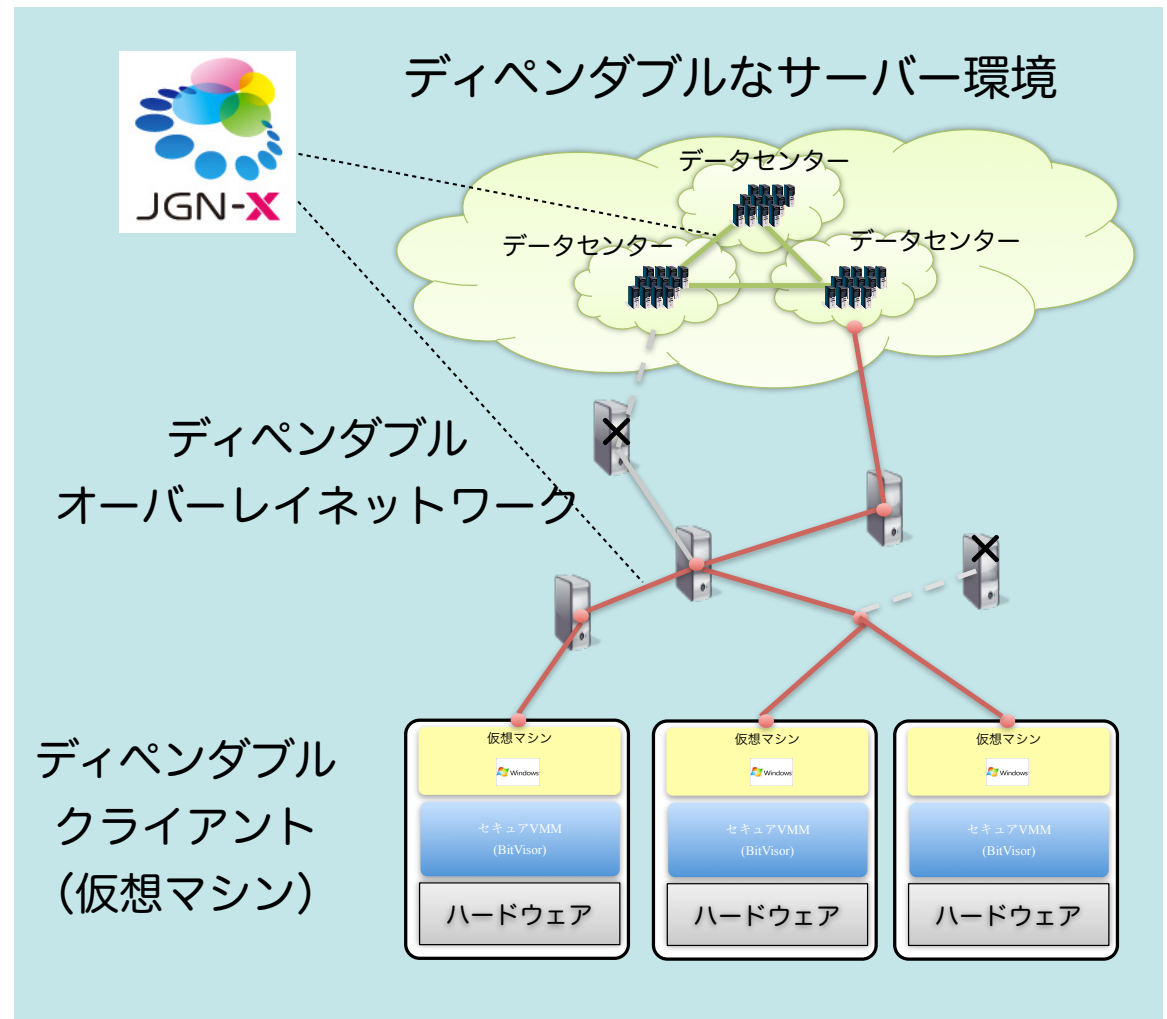


まとめ

クライアント，ネットワーク，サーバ環境を仮想化技術で
統合したコンピューティング環境

ディペンダビリティ：

- 信頼性
- 可用性
- 応答性
- スループット
- セキュリティ
- プライバシー



ディペンダビリティ

	クライアント環境	ネットワーク環境	サーバ環境
可用性	リモート管理機能 ネットワークブート機能	透過的VPN切り替え	VMのハイアベイラビリティ P2P型のメンバシップ管理
信頼性	プロテクションドメイン	広域データセンター環境	高信頼性言語
完全性	システムファイル保護 マルウェア検出機能 VMMコア	VPN (IPsec) 接続	VMの利用 LDAP認証&サンドボックス 仮想ネットワーク
保守性	リモートコントロール	VPN (IPsec) 接続	資源のオブジェクト化 スクリプティング環境
機密性	ID管理 ストレージ管理 VMMコア	VPN (IPsec) 接続	VMの利用 LDAP認証&サンドボックス 仮想ネットワーク

取り組んでいる課題

- Kumoiの拡張・応用
 - ✓ 仮想ネットワーク制御（OpenFlow等）
 - ✓ Failure-oblivious computing（失敗忘却計算）の導入
 - ✓ 通信キャッシング技術による高速化
 - ✓ 応用：並列分散パラメータ・チューニング
- BitVisor応用
 - ✓ 透明なネットワークブートシステム
 - ✓ ゲストOS起動高速化
 - ✓ VMM内空き時間利用計算
- クラウド用省電力分散ストレージシステム