

スケールアウト & 分散アーキテクチャを基盤 に持つ簡単プライベートクラウドの構築

Macnica Networks Corp.

CTO 鈴木 秀一

<http://www.macnica.net>

Version: 2014/1/21



OpenStack Summit

2013/11/5-8

OpenStack Summit in Hong Kong



VMware vForum2013

2013/11/6-7

vForum2013 in Tokyo

比較してみました。

	vForum 2013	OpenStack Summit
来場者数 ※主催者発表	9000	3000
セッション数	109	160以上
スポンサー数	66社	63社
同じもの	<ul style="list-style-type: none"> ・活気 ・充実したエコシステム(パートナー) ・Software Defined というメッセージ 	
違うもの	<ul style="list-style-type: none"> ・スーツ姿 ・エンタープライズ事例 ・セキュリティ&HA 	<ul style="list-style-type: none"> ・Tシャツ姿 ・サービス事業者事例 ・スケラビリティ&アジリティ

若干状況に変化が。。

◎ OpenStackの急速な成長

- 年2回のアップデート、モジュール化された主要コンポーネントの取り込み。(Neutron、Ceilometer、Heat、TripleO等々)
- 今後、セキュリティ等にも積極的に取り組んでいくというコミュニティの姿勢がKeynoteからもわかる。

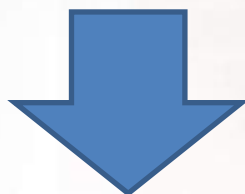
◎ 急成長するOpenStack関連ビジネス

- 2013年 \$600M → 2015年 \$1B(\$1000M)
- ディストリビューションやターンキーソリューションの売上は2014年に50%以上成長し、OpenStackビジネスの成長エンジンになる。

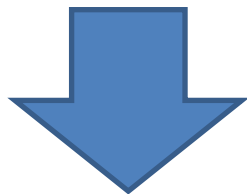
※ソース:451 Report

今後の動きを大胆に予測すると。。

2014年は、より簡単にかつセキュアにOpenStackを使う環境が充実してくる。



多くのエンジニアリソースを擁する大手のサービス事業者だけでなく、エンジニアリソースに制限がある比較的小規模な事業者やエンタープライズがOpenStackを使い始める。



2014年はエンタープライズにおけるOpenStack
元年になる！！

OpenStack元年に合わせて。。

- ◎ マクニカネットワークスでも社内の技術部門システムでOpenStack利用の検討を開始。
- ◎ どのように我々がOpenStackを使ったシステムを組もうと考えているかを、デモを交えてご紹介。

マクニカネットワークスの紹介

セキュリティ

APT対策、エンドポイント暗号
Firewall、IPS、IDP
Web/Email GW
知的財産保護



ネットワーク

ワイヤレス、LAN
リモートアクセス
可視化、高速化、負荷分散、FW
Web-Sec Gateway



プラットフォーム

BigData、クラウド基盤
IT・IDマネジメント
コンピューティング
ストレージ



キャリア/OEM

通信キャリア基盤
トラフィック最適化
VoIPゲートウェイ、暗号鍵管理
キャッシュ、SeGW



エンジニア部隊ではサポートや検証を結構がっつりやります。

パフォーマンス&セキュリティテスト環境

- 160Gbpsスループットテスト (アプリケーションレイヤ)
- L4-L7負荷分散
- 150以上のアプリケーションロードテスト
- 25,500以上のセキュリティ テストシナリオ (attack pattern)

2009年以前のサーバー検証環境

- ◎ 各個人、各プロダクトチームで勝手にいろんなサーバーを立てていた。
- ◎ ベアメタルサーバーあり、VMware あり、Xenサーバーありの乱立状態。

2010年に検証環境と運用サーバ環境の仮想化、統合の最初のフェーズをスタート

Phase 1:とにかく統合化 & 仮想化しよう

◎ 2010年

- すでに活用中のVmware Workstation, ESX、ESXiをVMware vSphereに統合することでスタート
- 取り扱い商材や自分たちが興味がある技術を極力使った構成(Altor, Vkernel、McAfee、Nexenta等々)
- 最初はホスト2台からのスモールスタート



Phase1: スタートから2年、、、

サーバ11台、ストレージ4台 (SSDストレージも含む)



Phase 1: 課題

1. 増大する運用負荷

- ◆ ハード増設、ネットワーク構成見直し、ストレージのメンテ
- ◆ 機器の増加(=障害ポイント増加)に合わせて続発するトラブル
→ 本来のサポート&検証業務ができず。。。。

2. 増大する使用スペースと電力

- ◆ 6U → 30U
- ◆ 増え続ける消費電力&電源工事
→ 本来検証すべき機器がラックに置けない。。。。

3. 劣化するパフォーマンス

- ◆ 劣化するサーバ処理により、検証が進まない。
- ◆ 検証対象機器が遅い?それとも検証設備が遅い?
→ 何を検証しているのかわからない状態に。。。。

Phase 2: 課題を解決するために。。



を導入してみました。

Before



After



スペース

30U → 4U

87% Off

消費電力

7.7kW → 2.1kW

73% Off

パフォーマンス

5時間 → 30分

10倍 !!

Phase 2: Nutanix導入効果と新たな課題

◎ 導入効果

- 運用負荷大幅軽減 → ほぼトラブルフリー & 追加が簡単
(担当エンジニア泣いて喜んでいきます。)
- 消費電力とラックスペースの大幅節減
- パフォーマンスの大幅改善

◎ 新たな課題

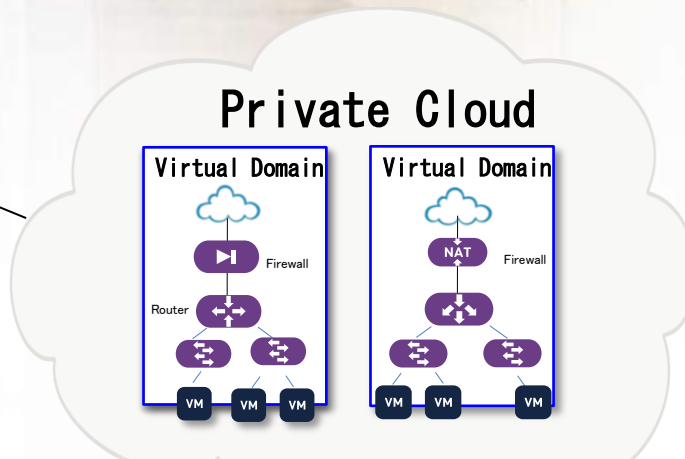
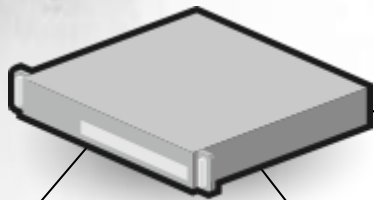
- スケールさせようとする、コスト高になる。(ライセンス費)
- リソース管理ができていない。管理しようとする、非常に煩雑。
- ネットワーク構成が非常に硬直的。ネットワーク機器との連携が困難。またアイソレーションできない。
- ロックインされてない? → 新しい技術、イノベーションへの追従ができない

Phase 3: OpenStackを使ったシステムのゴールイメージ

- ◎ 各ユーザがセルフ・プロビジョニングができ、かつリソース管理がしっかり行える。
- ◎ SDN機能を使い、各ユーザ毎の独自ネットワークをNeutronを使い、Horizon上あるいは独自Viewから簡単に構成できる。
- ◎ 物理アプライアンスをこの独自ネットワーク上に簡単に取り込み、検証ができる。
- ◎ パフォーマンス劣化が少なく、サーバリソースが足りない場合はHWを追加するだけで簡単に拡張できるシステム。
- ◎ 運用担当者に負担をかけないシステム。

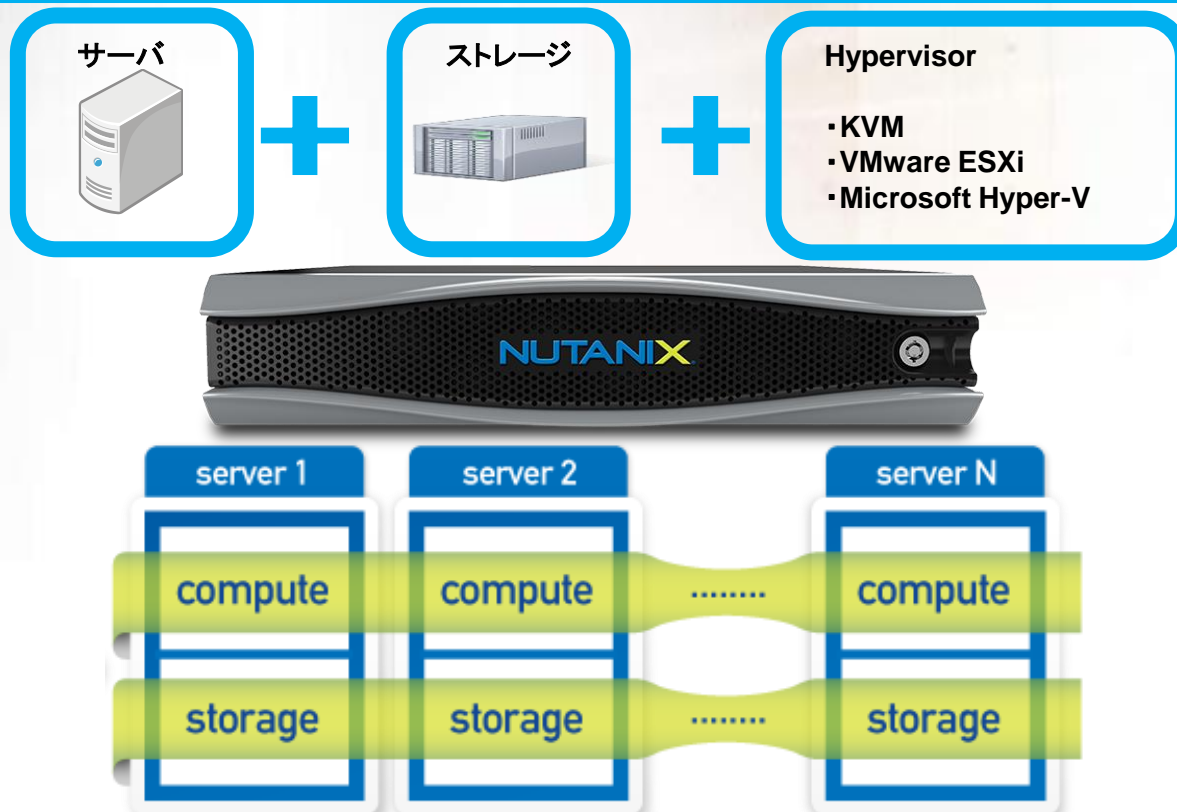


Phase 3: システム構成



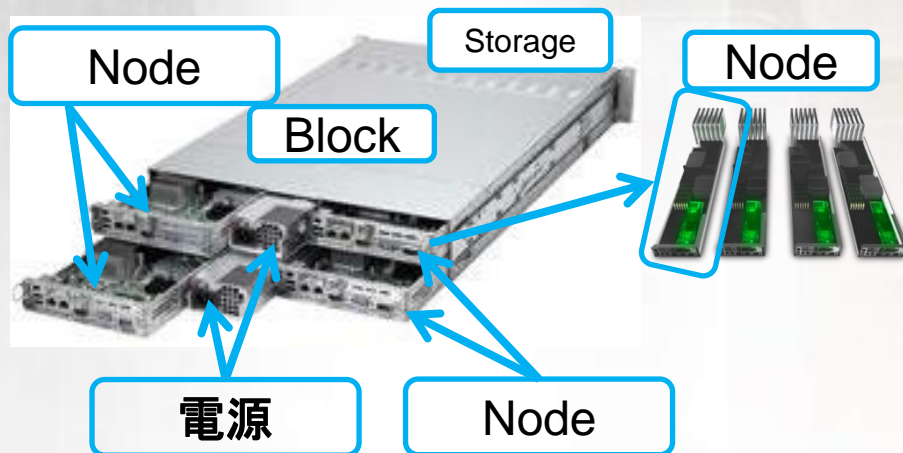
コンバートドシステム : Nutanix

サーバ群とストレージ階層を単一の統合されたアプライアンスに集約した
コンバートドインフラシステム



GoogleやFacebook、Amazon等と同様の設計原理や技術を活用した分散ファイルシステム(NDFS)により、企業の成長とワークロードの必要性に合わせて、新しいノードの自動検出をサポートし、既存のクラスターリソースを壊すことなく、1つずつノードを追加することができます。

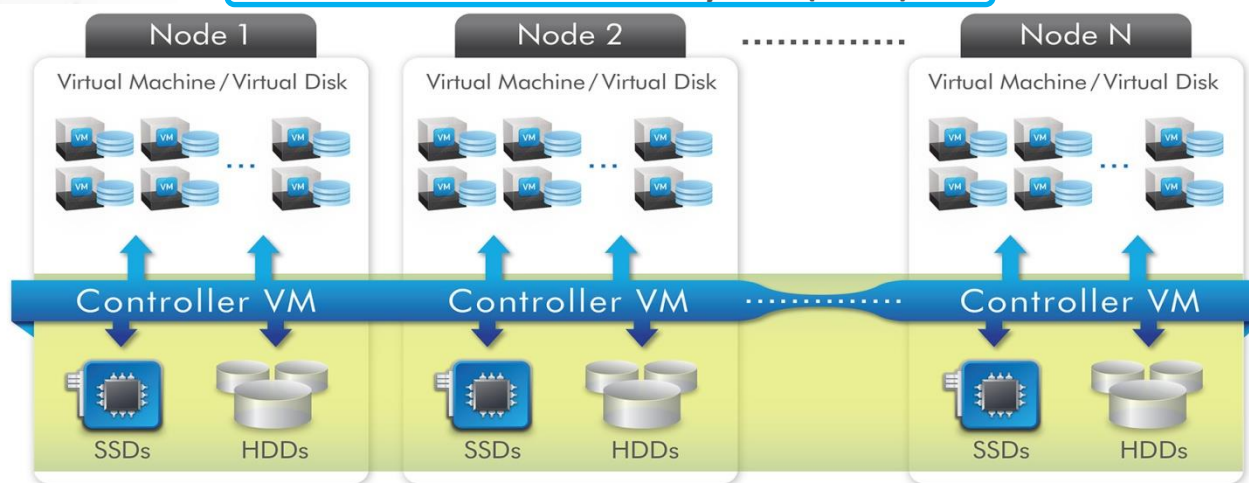
コンバージドシステム : Nutanix



Nutanixの特徴

1. Hypervisorがプリ・インストール済み
2. ストレージ専用ネットワークを排除してSANの構築不要
3. Google等が採用するデータ分散ストレージアーキテクチャ
4. パフォーマンスを維持しながら容易に拡張可能
5. 全冗長化構成で障害時のユニット交換はノンストップ
6. SSD→HDDのストレージ自動階層化(ティアリング)

Nutanix Distributed File System (NDFS)



- 最適なリソース配分をVMコントローラーにより自動制御する分散ファイルシステム

OpenStack on Nutanix



OVM
Horizon, Keystone,
Nova, Glance,
Cinder

CVM

Nutanix Node 1
(KVM/QEMU + nova-
compute/Cinder)

CVM

Nutanix Node 2
(KVM/QEMU + nova-
compute/Cinder)

...

CVM

Nutanix Node N
(KVM/QEMU + nova-
compute/Cinder)

◎ 容易な環境構築

- OpenStackの基盤となるKVM環境はプリインストール済
- サーバやストレージの細かいサイジングや設計を必要としない

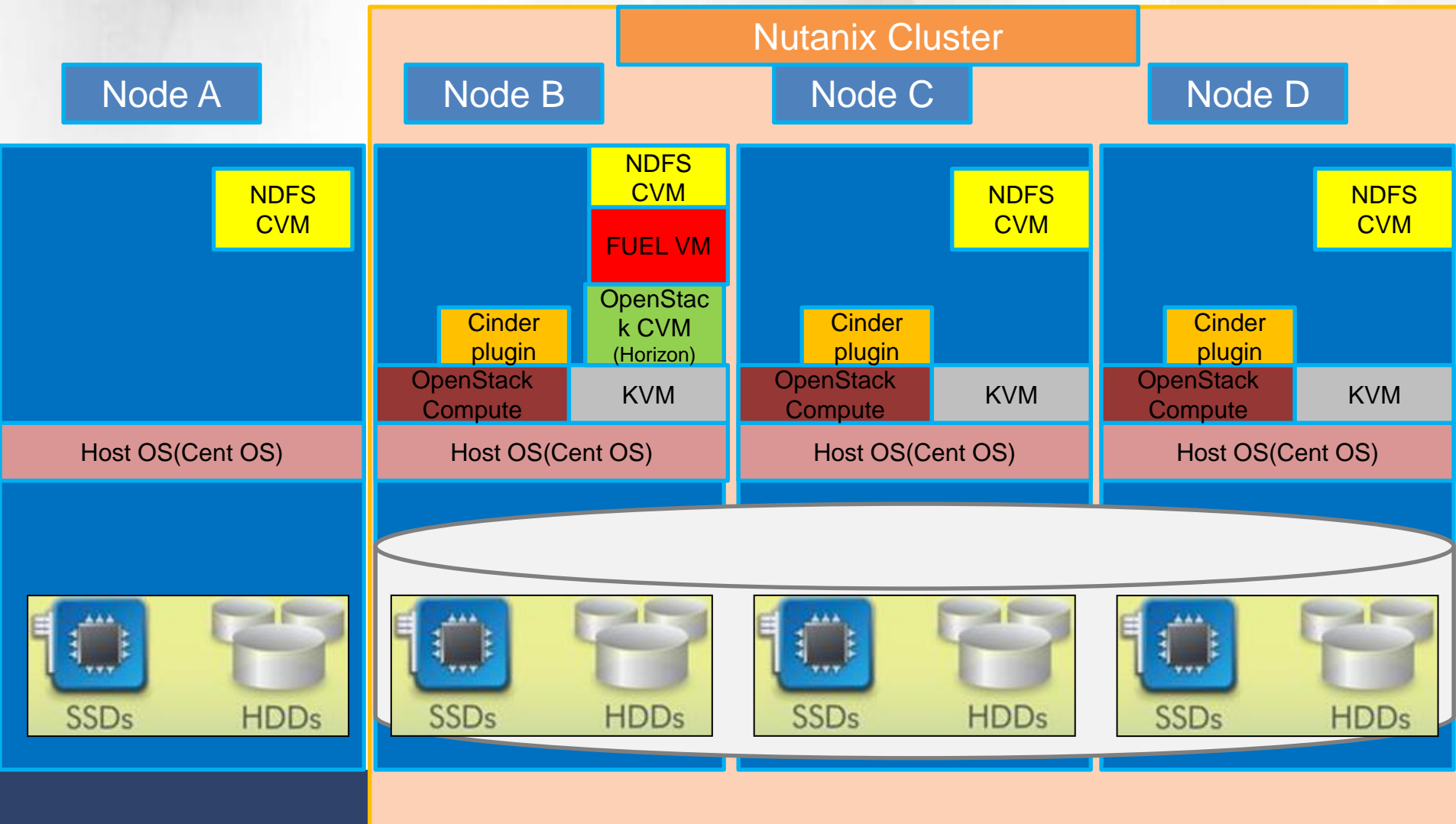
◎ 優れた拡張性と分散処理

- 特別な設定をする事なくノード追加で簡単拡張
- パフォーマンスを維持しながらスケールアウトを実現

【デモ】

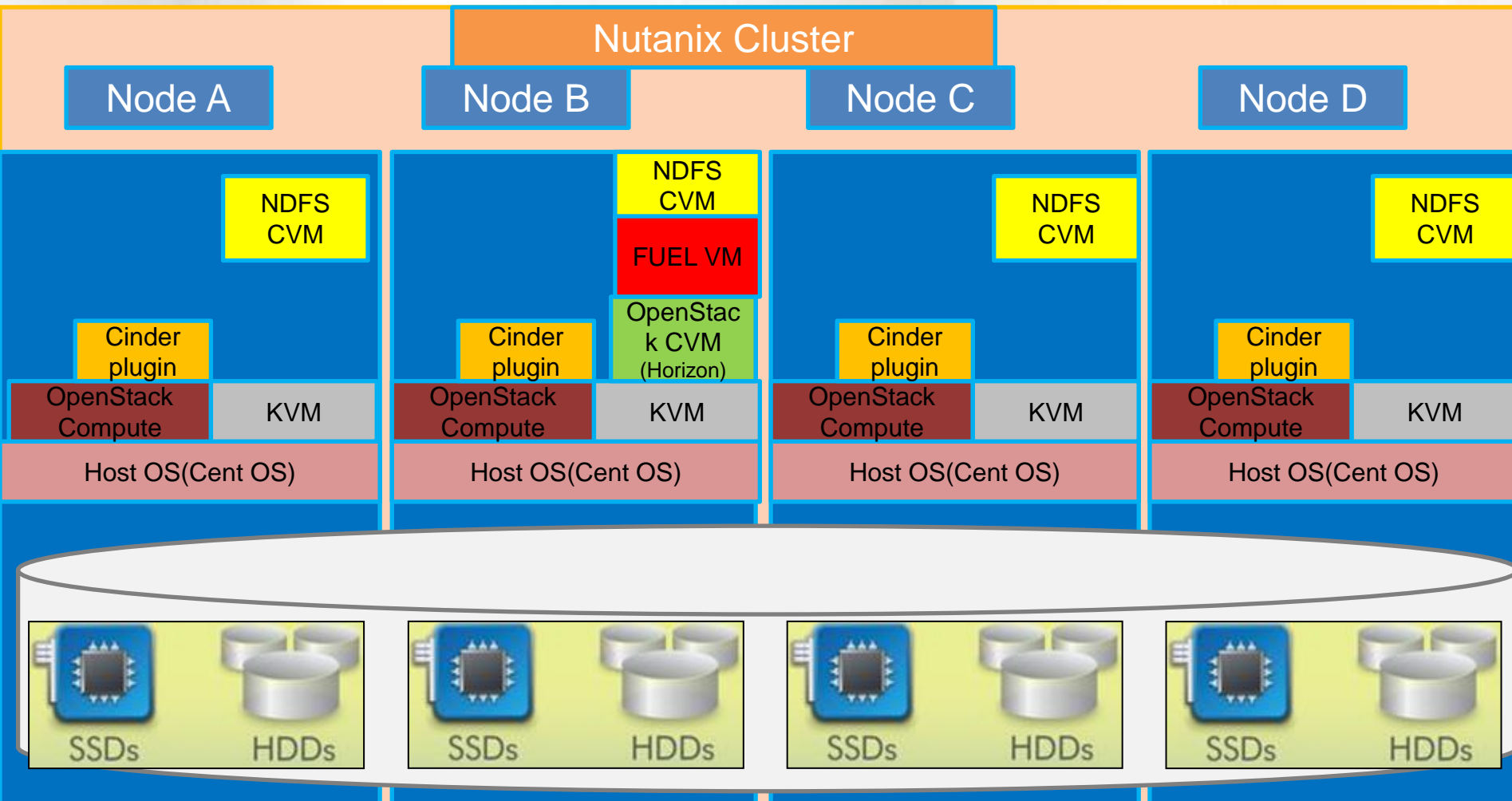
OpenStack on Nutanix

Nutanix: デモ環境



Nutanix: デモ環境

Nutanix管理コンソールから新規にノードを追加する事で、サーバ層とストレージ層のスケールアウトに合わせて、OpenStack層も同時に追加完了 ※OpenStack層自動スケールアウト機能は今後実装予定



Parse Lookup Update Modify grid

オーバーレイ+豊富なネットワーク機能をソフトウェア提供

仮想ドメイン

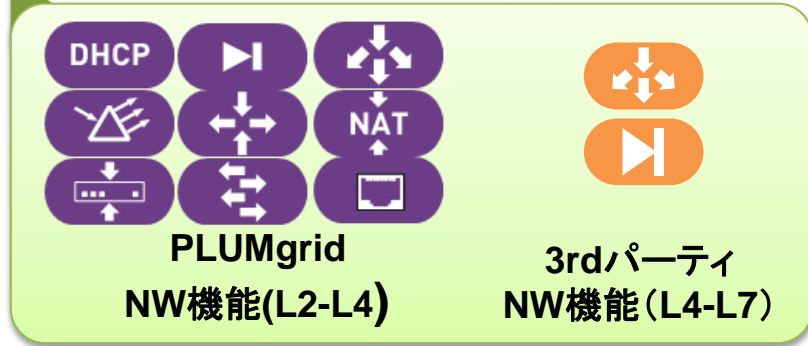
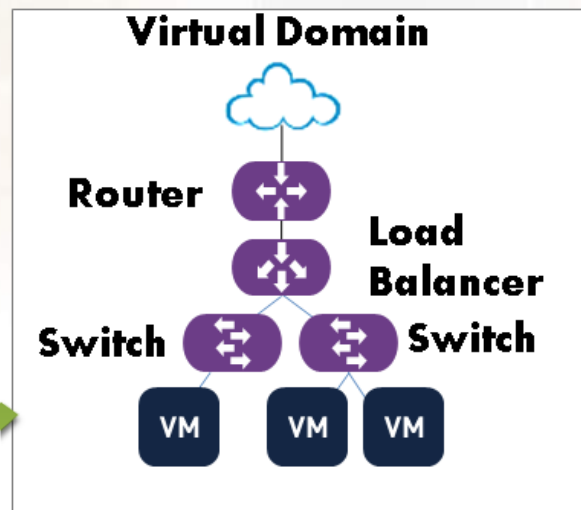
- マルチテナントソリューション
- ユーザはオンデマンドで、仮想ドメインを作成クローン、移行を行う。

独自のネットワーク機能の提供

- ルータ、DHCP、NAT機能を提供
(今後はFireWall、Load Balancerも)
- スケーラブルなコントロールプレーン

マルチハイパーバイザー対応

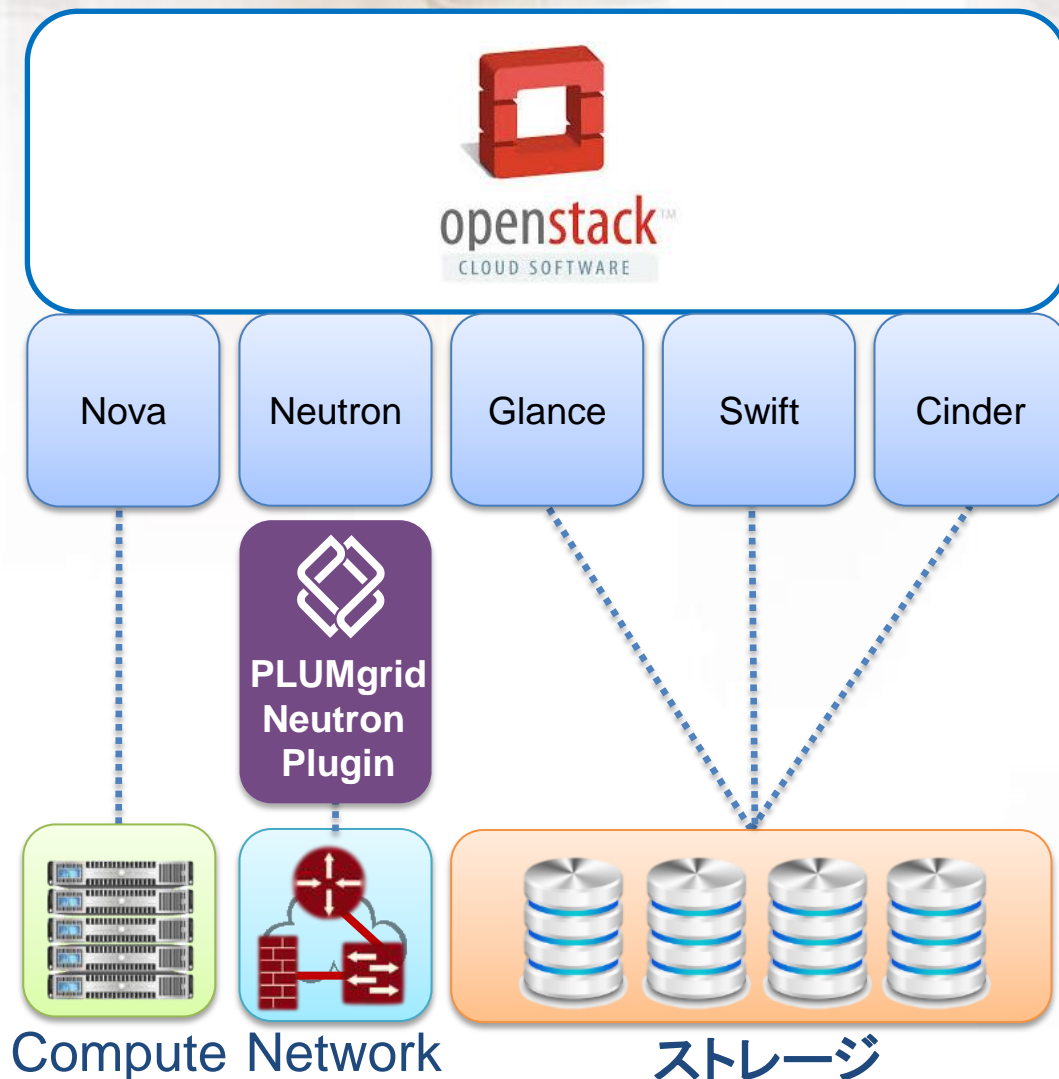
- KVM, ESXiに対応



OpenStack環境で必要なネットワーク機能をNeutronプラグインを介して提供

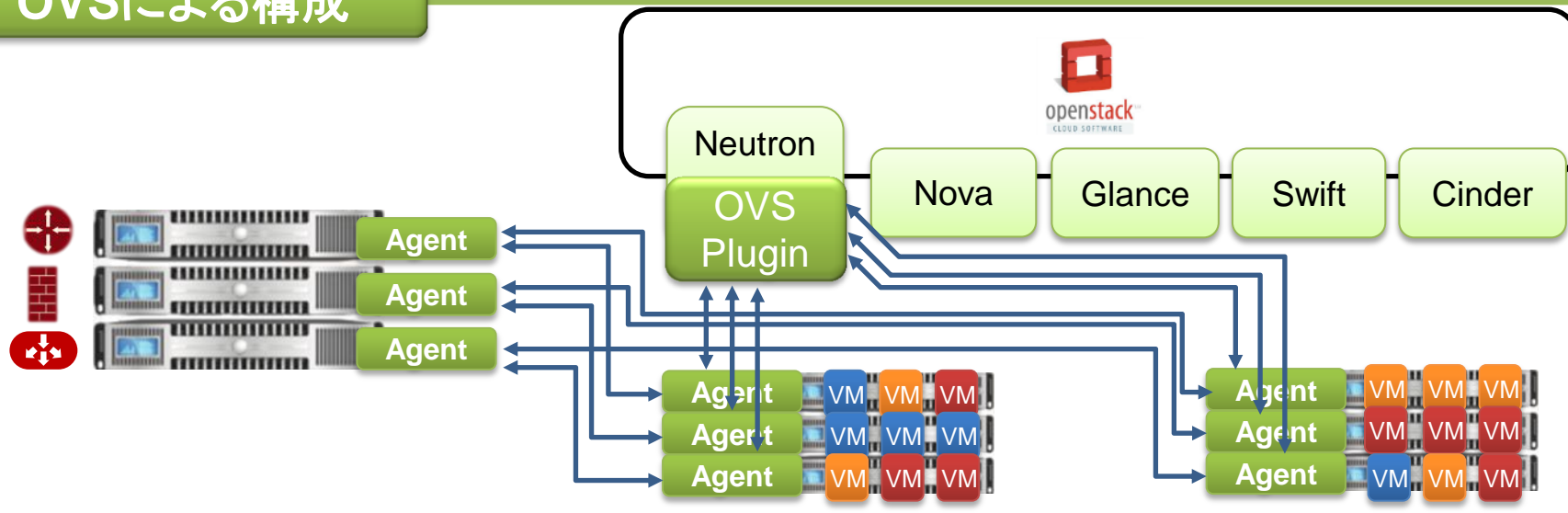
PLUMgrid OpenStack連携

- マルチテナント
 - Virtual Domain (仮想ドメイン)
 - アイソレーション (VxLAN, NVGRE)
- 拡張機能
 - 各種ネットワークサービスの提供
 - No OVS, No Flat-network
- スケーラビリティの向上
 - No VLAN, No Agents,
 - No Controller bottle neck
- オープンプラットフォーム
 - Neutron 連携
 - 3rd パーティ製ネットワーク機能連携
- ネットワーク可視化
 - PLUMgrid モニタリングと解析機能

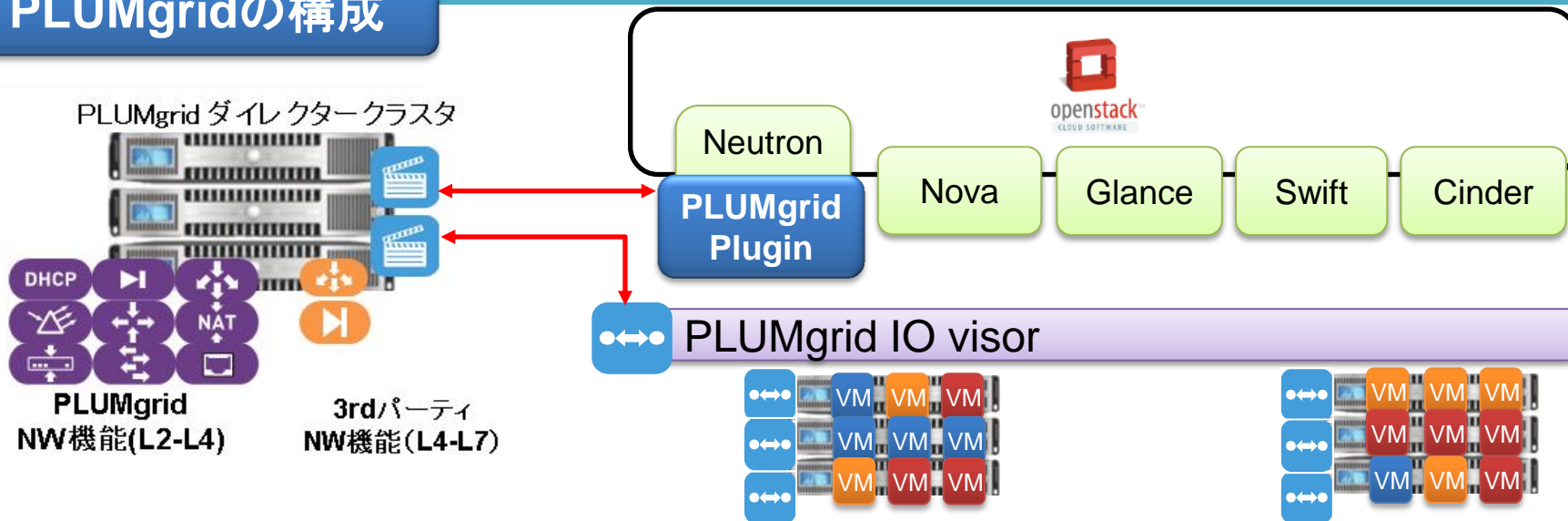


ネットワーク機能のシンプル化

OVSによる構成



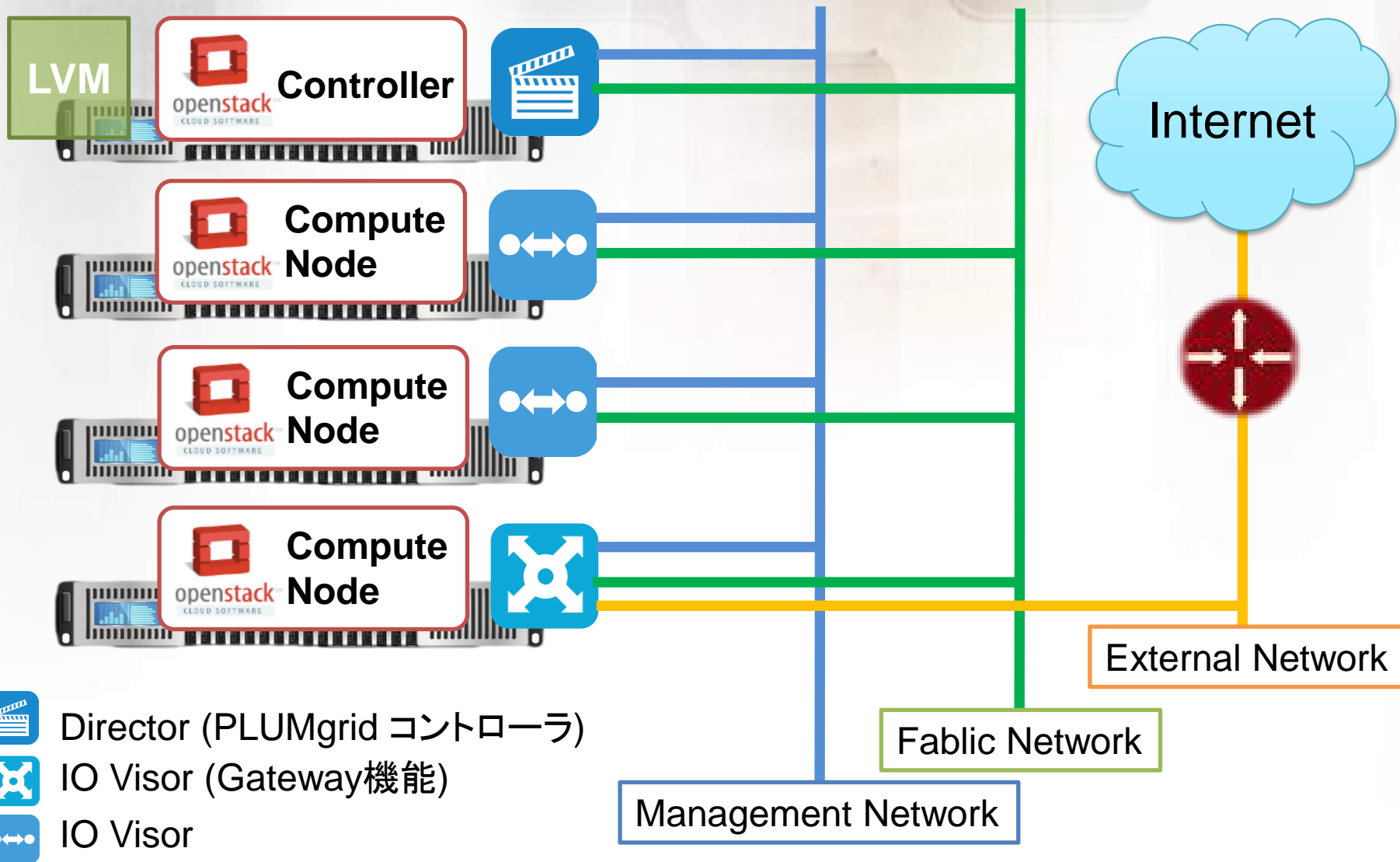
PLUMgridの構成



【デモ】

PLUMgridによる仮想ネット ワークドメイン構築

PLUMgrid デモ環境構成



Splunk とは？ あらゆるデータを一元化

どんなソースからのどんな形式のテキストデータでも
高速にインデックス化を行います



Ad hoc
search



Monitor
and alert



Report
and
analyze



Custom
dashboards



Developer
Platform



Not RDBMS



事前のスキーマ設計不要



データ別に収集用のコネクタ不要



データのフィルタ等不要

splunk>

Other Big Data
stores

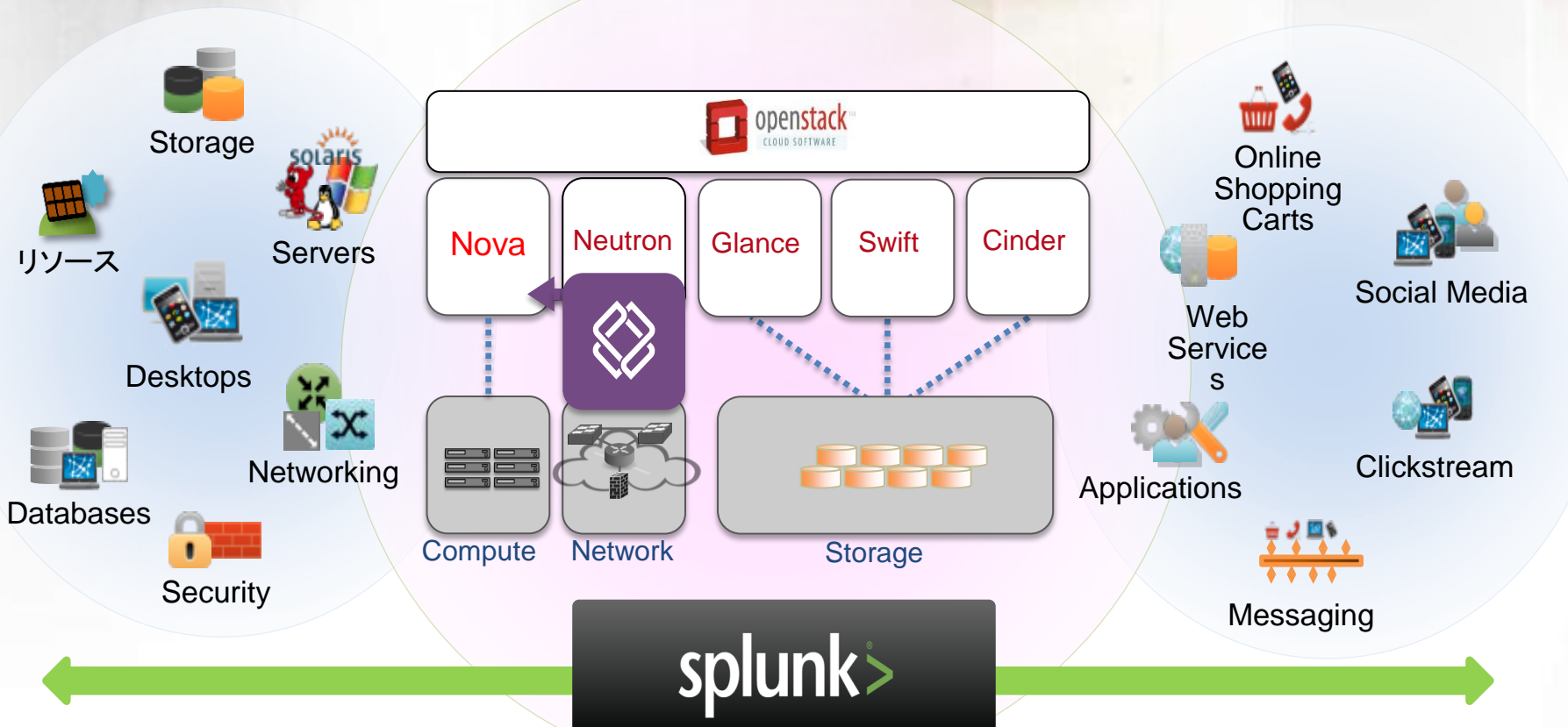


OpenStack周辺に拡大するデータソース

ハードウェア

OpenStack

アプリケーション



異なるログ形式の分散環境でのログをSplunkで集約し、横断的な
 サーチ、アラート、レポート

【デモ】

Splunkによる横断的なログ 解析手法

今後のインプリ予定

- ◎ OpenStack on NutanixのGA版に移行
- ◎ OpenStack on Nutanixの環境にPLUMgridを移植。(2014/4にFUELに対応予定。)
- ◎ Splunkをこの上で動作させ、まずはいろいろなログ (OpenStack、PLUMgrid) を食わせ、運用ルールを作成。
- ◎ 外部のフィジカル環境とつなぐためCumulus Linuxや他のSoftware Definedなスイッチとこのシステムを連携させる。