

Eucalyptus Hacks

～現場で上手にEucalyptusを活用するために～



2010年7月23日
株式会社NTTデータ
技術開発本部 伊藤雅典

INDEX

01 自己紹介

02 ねらい

03 Eucalyptus1.6.2の目玉

04 Eucalyptusの構造

05 注意しておくべき留意点と解決法

06 今後のEucalyptusへの期待

07 まとめ

01 自己紹介

変える力を、ともに生み出す。
NTT DATAグループ



氏名

伊藤 雅典 (いとう まさのり)

所属

株式会社NTTデータ 技術開発本部 ITアーキテクチャ&セキュリティ技術センタ

担当業務

NTTデータの総合クラウドサービス bizXaaS™ (<http://bizxaas.net/>) の、
「フルOSSクラウド構築ソリューション」の開発に従事

<http://www.nttdata.co.jp/release/2010/040801.html>

Eucalyptus 本体のデバッグを中心にトラブルシュートを担当

トラブルを起こすものは、Eucalyptusに限らずkernelでも何でも直します。😊

その他、活動領域

日本Eucalyptusユーザーズグループ (JEUG)、VIOPS InterCloud-SIG、GICTF (グローバルクラウド基盤連携技術フォーラム、<http://gictf.jp/>) 等でも活動中

02 ねらい

変える力を、ともに生み出す。
NTT DATAグループ



本日のプレゼンテーションでは、現場で上手にEucalyptusを活用するために、以下のトピックについてご紹介します。

1

Eucalyptus 1.6.2の 目玉

Eucalyptus 1.6.2は、なぜ注目に値するのか？

2

Eucalyptusの構造

Eucalyptus はどのような構造で実現されているのか？
(ネットワーク視点を中心に)

3

活用にあたっての留意点

Eucalyptus 1.6.2 を活用する上で、なお注意が必要な事項
とは何か？

4

今後の期待

今後、期待される改善とは何か？

03 Eucalyptus 1.6.2の目玉

変える力を、ともに生み出す。
NTT DATAグループ



1

EBS安定化

1.5系では、永続ストレージ (EBS) の管理ロジックに問題があり、EBSボリュームの中身が勝手に壊れることが多々あった

2

NW安定化

1.5系では、iptables によるNAT設定やフィルタ設定が壊れ、VMが通信不能に陥ることが多々あった

3

マルチクラスタサポート

ラック単位の増設ができるようになった

主要な未サポート機能

- Windows インスタンスサポート
- EBSからのインスタンス起動
- EBS インフラとしての iSCSI サポート等

04 Eucalyptusの構造

変える力を、ともに生み出す。
NTT DATAグループ



Eucalyptusの論理コンポーネント構成

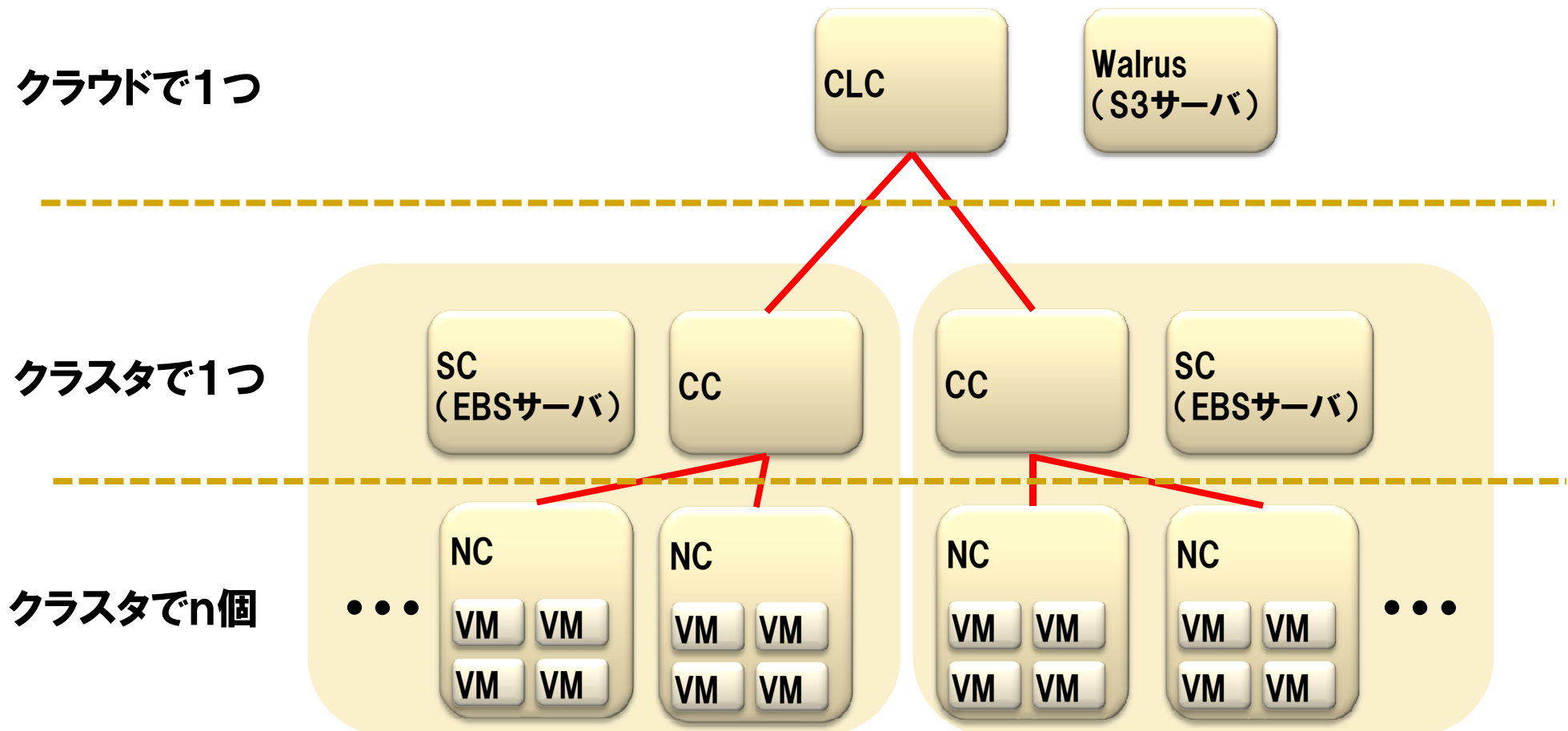
論理コンポーネント	略称	役割	実装単位	備考
Cloud Controller	CLC	クラウド全体の制御	eucalyptus-cloud プロセス	クラウドで1つ
Walrus	Walrus	S3サーバ(イメージ データ等の保持)	同上	クラウドで1つ
Storage Controller	SC	EBSサーバ	同上	クラスタで1つ
Cluster Controller	CC	NAT、VLAN制御、 DHCP	CCプロセス	クラスタで1つ
Node Controller	NC	VMの実行	NCプロセス	クラスタごとにN台

04 Eucalyptusの構造

変える力を、ともに生み出す。
NTT DATAグループ



Eucalyptusの論理コンポーネント構成

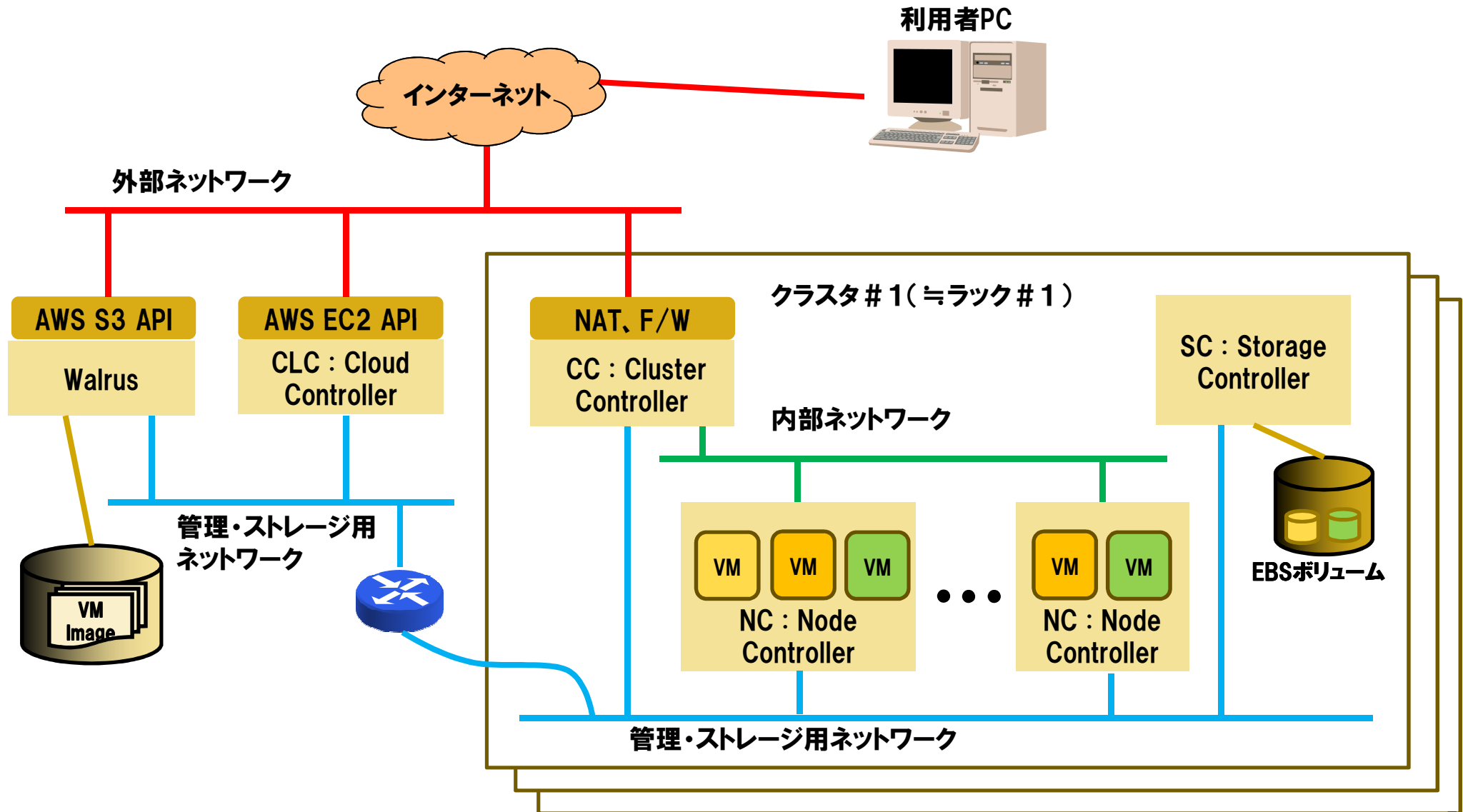


04 Eucalyptusの構造

変える力を、ともに生み出す。
NTT DATAグループ



Eucalyptusによって構築したIaaS環境の、NW構成例



04 Eucalyptusの構造

- **eucalyptus-cloud プロセス**
 - CLC、SC、Walrus の実体(分離してノード分散させて動作させるのも可能)
 - 起動処理はC言語で記述されており、libjvm.so を動的リンクして途中からJVMとして動き始める(！？)
 - 使っている主要な部品は、Mule ESB、hsqldb、hibernate、jetty、GWT等…
 - EBSサーバ(SC)のストレージ実体は、SCのローカルファイルシステム上のファイル。これをLVMで論理ボリューム化し、AoE(vblade)で NC にエクスポートする
 - ✓ I/Oネックになりがち
 - S3サーバ(Walrus)のストレージ実体は、普通のローカルファイルシステム。これを、S3のSOAPプロトコルで SC や NC と共有する。
 - ✓ I/Oネックになりがち

04 Eucalyptusの構造

- CCプロセス
 - Apacheのモジュールとして、axis2c を使って実装
 - NAT/FireWallの機能は iptables で実現
 - ✓ VMのパブリックIPは、すべてここで制御している
 - ユーザごとのNWアイソレーションは vconfig でTag VLANを切り出して実現
 - ✓ 内部ネットワークに用SWポートはトランクポートとして構成しておく必要がある
(MANAGEDモードで使う場合)
 - VMのprivate IP は、dhcpd を外部コマンドとして利用して割り当て
 - 複数クラスタ間の内部ネットワーク接続は Vtun で実現
 - ✓ Private IP 空間、VLAN Tag 空間がクラスタ間で一意になる

04 Eucalyptusの構造

- **NCプロセス**
 - Apacheのモジュールとして、axis2c を使って実装
 - libvirt APIを呼び出し、VMの起動・停止・構成変更等を制御
 - CCと協調してVLAN を構成(vconfigコマンド)、VMに仮想NICとして接続
 - VMのディスクイメージは、Walrus からローカルにダウンロードして、起動
 - ✓ キャッシュ機能があるため、毎回ダウンロードするわけではない
 - ✓ それでもキャッシュからのコピー処理が I/O ネットワークになりがち
 - SCがエクスポートするAoEボリュームに接続、EBSボリュームとしてVMに attach
 - ✓ iSCSIデバイスを扱うコードはまだ開発中

05 注意しておくべき留意点と解決法

Eucalyptus 1.6.2 でも発生するトラブルと解決方法の事例を紹介します
(特に、ネットワーク関連で顕在化する事例を中心に)

□ Hack1 : 23ユーザ以上収容しようとする通信できないVMが出る！？

□ 種別

□ Eucalyptusのバグ

□ 理由

□ CCが内部的に dhcpd を呼び出す際のコマンドバッファ長に不足があり、
必要なNIC (bridge) の一覧が dhcpd の引数として渡らず、DHCP の
lease に失敗するため。Uptime の長いVMから顕在化する。

□ 解決方法

□ パッチ適用・再ビルドする以外に回避方法はありません。

□ 参考

□ <https://bugs.launchpad.net/eucalyptus/+bug/603998>

05 注意しておくべき留意点と解決法

□ Hack2 : CLCを再起動したら、確保したはずのEIPが消えた！？

□ 種別

□ Eucalyptusのバグ

□ 理由

□ EIP (Elastic IP) は、クラウドで一意的な資源として、CLCで管理されているが、永続化のために使用している O/R mapper 使い方に問題があり、EIPの entity が正しく永続化されていないため。

□ 解決方法

□ パッチ適用・再ビルドする以外に回避方法はありません。

□ 参考

□ <https://bugs.launchpad.net/eucalyptus/+bug/581779>

05 注意しておくべき留意点と解決法

□ Hack3 : NC上のVMに対して scp すると、場合によっては通信が stall する！？

□ 種別

□ 設定ノウハウ

□ 理由

□ Eucalyptus CC/NCのプライベートネットワーク用NICに、Xenのデフォルトブリッジ(xenbr0)が設定されたNIC経由で通信すると、Linuxの不具合にひっかかる(正確には、NICの多段構成が問題)

□ 解決方法

□ Eucalyptus の NCでは xenbr0等 は使用しないため、Xenの設定で無効にするのが無難。具体的には、`/etc/xen/xend-config.sxp` の `(network-script network-bridge)` をコメントアウトする。

□ 関連事項

□ あわせて、libvirtがデフォルトで作成するNAT用の bridge の virbr0 が存在すると、CCが払い出すプライベートIPと衝突するため、virbr0 も無効にしておくのが無難。具体的には、以下のシンボリックリンクを削除する。
`/etc/libvirt/qemu/networks/autostart/default.xml`

05 注意しておくべき留意点と解決法

Hack1～Hack3 は、ほんの一部です。

Eucalyptus本体の改善や修正は、日本Eucalyptusユーザー会を通して発信していきます。

参考:その他の留意点の例

- VMどうしが Public IP で通信できない(パッチ)
- 1つのEBSボリュームが、2つのインスタンスに重複アタッチされることがある(パッチ)
- EBSトラフィックをVLANで分離できない(パッケージ入れ替え)
- CentOS 5.x等でEBSを使うと kernelが“BUG”と表示して Call Traceを出す(カーネルパッチ)
- インスタンスをリブートに失敗する(パターン1、パターン2)(パッチ)
- Web UI の管理画面に乗り込めなくなることがある(パッチ)
- JVM系のログ出力をする方法、GCオプションの指定方法が分からない(ノウハウ)

等々

05 注意しておくべき留意点と解決法

NTTデータからのコントリビューション(抜粋)

Eucalyptus BTS

- <https://bugs.launchpad.net/eucalyptus/+bug/424212>
- <https://bugs.launchpad.net/eucalyptus/+bug/481154>
- <https://bugs.launchpad.net/eucalyptus/+bug/581612>
- <https://bugs.launchpad.net/eucalyptus/+bug/581779>
- <https://bugs.launchpad.net/eucalyptus/+bug/582220>
- <https://bugs.launchpad.net/eucalyptus/+bug/582721>

Eucalyptus Forum

- <http://open.eucalyptus.com/forum/nc-and-cc-do-not-start-automatically>
- <http://open.eucalyptus.com/forum/ebs-does-not-work-8021q-vlan>
- <http://open.eucalyptus.com/forum/ebs-io-causes-kernel-call-traces-under-xen-environment>
- <http://open.eucalyptus.com/forum/q-why-modprobe-aoe-sc-needed>
- <http://open.eucalyptus.com/forum/improving-eucalyptus-cloud-start-files-rhelcentos>

RedHat社 Bugzilla (Eucalyptusの動作に影響のある kernel 修正等)

- https://bugzilla.redhat.com/show_bug.cgi?id=593902
- https://bugzilla.redhat.com/show_bug.cgi?id=593905

06 今後のEucalyptusへの期待

- ストレージインフラの強化
 - iSCSIの正式サポート
 - IaaS環境に適したクラスタストレージの適用
 - ・ 例: SheepDog (<http://www.osrg.net/sheepdog/>)
- マルチクラスタサポートのブラッシュアップ
 - 今のVtunの使い方では、十分なスケーラビリティが確保できないため。必要なら、openvswitch や UT-VPN の活用、L3化の検討等。
- 管理データの永続化(特にCC)、SPOFの除去(CLC, CC)
- LiveMigrationサポート
- IPv6 サポート
- 等々
- 余談
 - ところで、EEE (Eucalyptus Enterprise Edition)ってどうなのだろう？

07 まとめ

変える力を、ともに生み出す。
NTT DATAグループ



- Eucalyptusも1.6系からそれなりに使えるようになってきました
- でも、上手に使いこなしていく上では、まだまだ留意事項もあります
- Eucalyptus本体に対する改善やノウハウは、日本Eucalyptusユーザーズグループ Webサイトを通して発信していきます
 - JEUG Webサイト : <http://eucalyptus-users.jp/>
- みんなで使って良くしていきましょう！



日本 Eucalyptus ユーザーズグループ



ご清聴ありがとうございました

変える力を、ともに生み出す。

NTT DATAグループ



本文中に記載の会社名、商品名、製品名などは、一般に各社の商標または登録商標です
ただし本文中では、TMや®マークは明記してありません