

～NTTレゾナント、IcehouseからLibertyへ～

実録！大規模環境のOpenStack アップグレードの考え方と実施のコツ

2016/7/6

NTTソフトウェア株式会社

NTTレゾナント株式会社

 **NTTソフトウェア**

 **NTTレゾナント**

本日の内容

- OpenStackのアップグレードとは
- NTTレゾナントのOpenStack環境について
- アップグレード全体の流れ
- 検証内容についての紹介
- まとめ



自己紹介

三木 遼



2010年～(現職)

NTTソフトウェア

仮想化関連技術に関する設計・開発・検証等

2012年（Essexの頃）からOpenStack関連プロジェクトに在籍

橋本智哉



2001年～2012年

NTTレゾナント

gooブログ・教えてgooなど主要サービスの設計・構築・運用

2012年～(現職)

NTTレゾナント

サーバ基盤の設計・構築・運用 統括

大木 和也



2011年～2014年

テプコシステムズ

東京電力グループIT基盤の設計・構築

2014年～2015年

NTTデータ先端技術

ログ可視化パッケージ製品の販売・保守・技術サポート

2016年～(現職)

NTTレゾナント

サーバ基盤の設計・構築・運用



NTTレゾナントについて



辞書
Dictionaries

郵便番号
ZIP codes

ラボ
Laboratory

からだログ
Bodycloud

住宅・不動産
Housing and real estateSearch

検索
Search

ベビー
Baby-care

映画
Movies

自動車&バイク
Car and bike

地図
Maps

路線
Navigation

占い
Horoscopes

ランキング
Rankings

ブログ
Blogs

ニュース
News

天気
Weather

ヘルスケア
Healthcare

スマホ部
Smartphone applications

旅行
Travel

転職
Job search

恋愛&結婚
Love and marriage

NTTグループのオンラインストア
NTT-X Store a goo service
Online store

「goo」は、使えば使うほど「あなたにフィット」していく
NTTグループのポータルサイトです。Web検索やブログ、メール、
Q&Aサイト「教えて!goo」など、60を超える行動支援サービスを提
供しています。

Webポータルサイト goo

<http://www.goo.ne.jp/>



19周年!



NTTソフトウェアについて

- 高い技術力でシステムの設計・開発・運用を手がける
 - 近年の注力キーワードは「クラウド」と「セキュリティ」



<https://www.ntts.co.jp/>



OpenStack案件のお手伝いします!!



OpenStackのアップグレードとは



OpenStackのアップグレードについておさらい

- コミュニティとしてのリリースサイクル
 - 半年ごとにメジャーリリース
 - 約一年後にEOLとなる
 - 新機能は最新のバージョンにしか追加されない

Series	Status	Initial Release Date	EOL Date
<i>Ocata</i>	<i>Future</i>	TBD	TBD
<i>Newton</i>	<i>Under Development</i>	2016-10-06 (planned)	TBD
<i>Mitaka</i>	Current stable release, security-supported	2016-04-07	TBD
<i>Liberty</i>	Security-supported	2015-10-15	2016-11-17
<i>Kilo</i>	EOL	2015-04-30	2016-05-02
<i>Juno</i>	EOL	2014-10-16	2015-12-07
<i>Icehouse</i>	EOL	2014-04-17	2015-07-02
<i>Havana</i>	EOL	2013-10-17	2014-09-30

* <http://releases.openstack.org/>



OpenStackのアップグレードについておさらい

• 主なディストリビューションでのリリースサイクル

- RedHatはリリースから3年間のサポートを提供
- Ubuntuはバージョンによって異なるが最大で5年間のサポートを提供

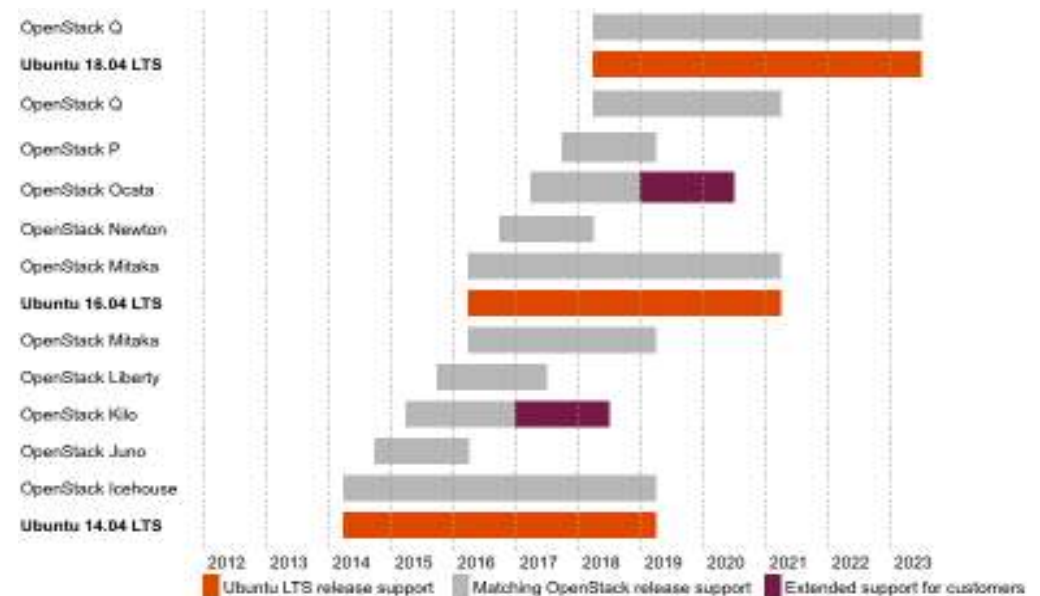
■ RedHat*

Red Hat OpenStack Platform Release	General Availability	End of Production, Phase 1	End of Production, Phase 2
3(Grizzly)	July 10, 2013	n/a	July 31, 2014
4(Havana)	December 19, 2013	June 19, 2015	June 19, 2015
5(Icehouse)	June 30, 2014	June 30, 2015	June 30, 2017
6(Juno)	Feb 9, 2015	Feb 9, 2016	Feb 17, 2018
7(Kilo)	August 5, 2015	August 5, 2016	August 5, 2018
8(Liberty)	April 20, 2016	April 20, 2017	April 20, 2019

* <https://access.redhat.com/support/policy/updates/openstack/platform>

** <https://wiki.ubuntu.com/OpenStack/CloudArchive>

■ Ubuntu**



OpenStackのアップグレード手法について

- アップグレードはHavanaの頃から考慮されている
- 無停止でのアップグレードもサポートされている

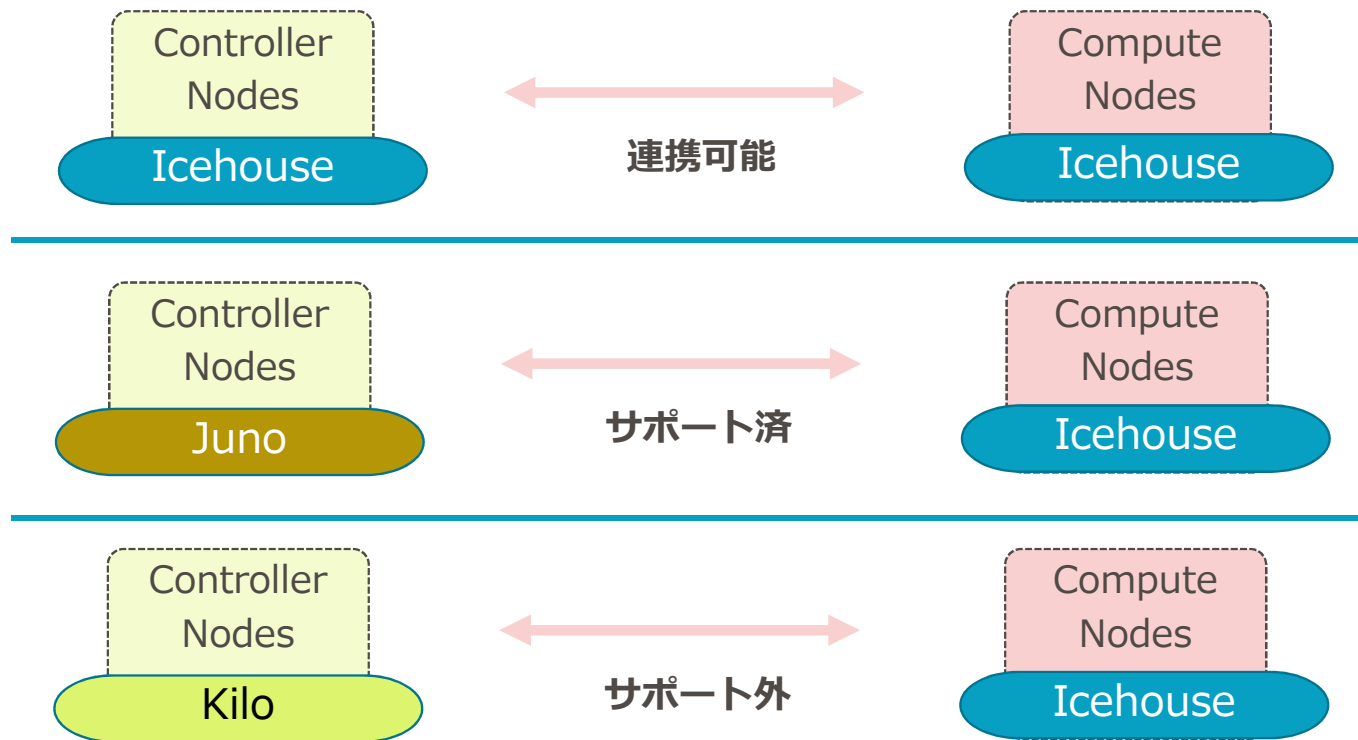
	コールドアップグレード	ローリングアップグレード
考え方	OpenStackを完全に停止して更新する※	OpenStack停止時間の最小化を目指す
主な実現項目	<ul style="list-style-type: none">• 1バージョン前とのConfigの互換性の維持• DBスキーマ更新方法の提供• 1バージョンアップグレード手順の提供• コミュニティCIによる検証が実施されている	<ul style="list-style-type: none">• コールドアップグレードと同様の項目• 複数のホストで異なるバージョンが混在する状態での動作
対応中のプロジェクト	Horizon, Keystone, Glance, Neutron, Nova, Swift, Ceilometer, Cinder, Fuel, Heat, Manila, Sahara	Nova, Swift, Neutron

※OpenStackが停止してもVMやボリュームは停止しません



OpenStackのアップグレード手法について

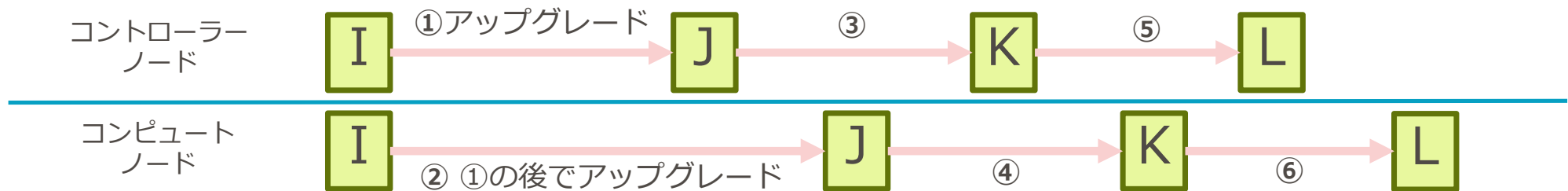
- ローリングアップデートでの制約事項
 - 基本的に一つ前のバージョンとの連携しかテストされていない



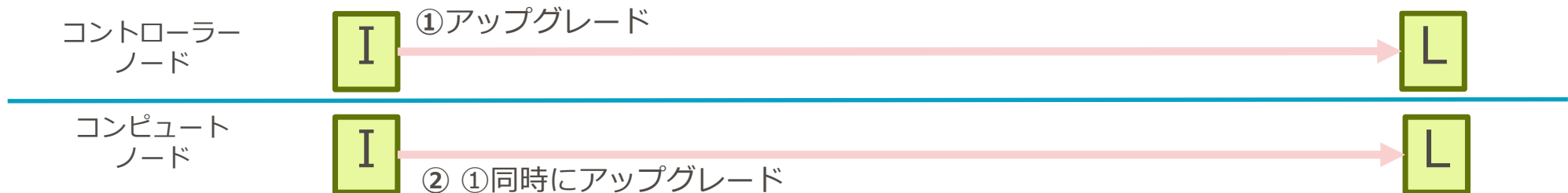
OpenStackのアップグレード手法について

- 基本的にはコールドアップデートがお勧め
 - IからLなど数段階のアップグレードをする場合、ローリングアップデートでは手順が煩雑

■ ローリングアップデート



■ コールドアップデート



アップグレードへの考え方

商用ディストリビューションで長期サポートを確保する

- システム廃棄期限までアップグレードしない
- 新機能の追加は行わない

コールドアップグレードできる環境作り

- OpenStackを停止することへの組織内での合意形成
- OpenStackを停止できればアップグレードは比較的容易



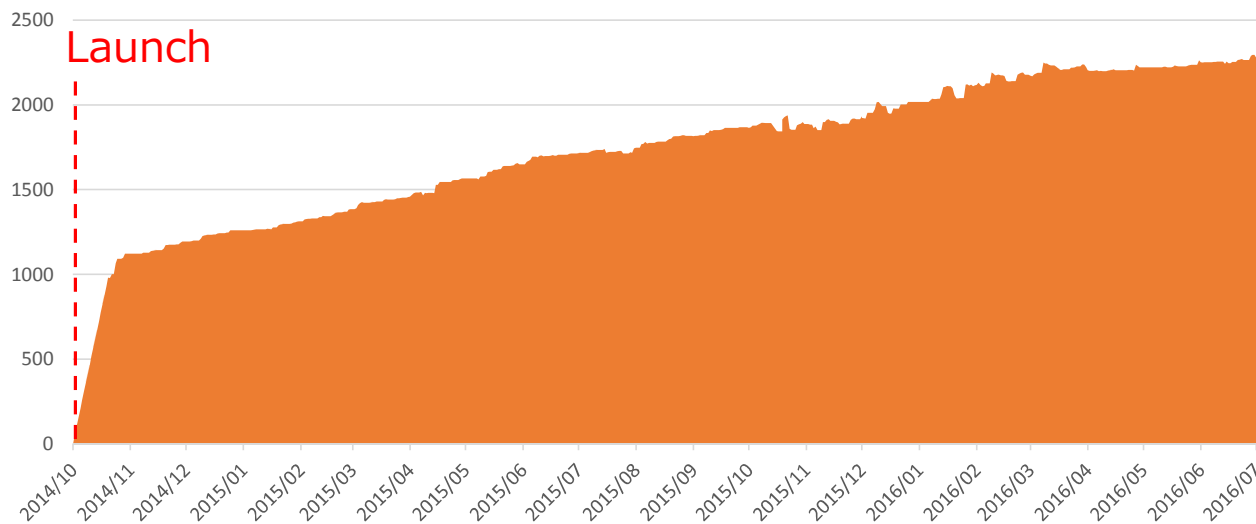
NTTレゾナントのOpenStack環境について



NTTレゾナントのOpenStackについて

NTTレゾナントのメインデータセンタにOpenStackを導入

- 2014年10月から運用中のプライベートクラウド
- 大小80種類以上のサービスを収容
- 月間 10億PV を支える環境
 - 400台の物理サーバ & 4800物理CPUコア



【VM起動数】
2016年7月現在
2200台以上



なぜアップグレードするのか？

- プライベートクラウド機能追加要望への対応
 - Kiloの新機能であるLBaaS(v2)等の利用を検討している
- その他の今後発生する新機能への対応を考慮
- アップグレード手順の早期確立

Series	Status	Initial Release Date	EOL Date
<i>Ocata</i>	<i>Future</i>	TBD	TBD
<i>Newton</i>	<i>Under Development</i>	2016-10-06 (planned)	TBD
<i>Mitaka</i>	Current stable release, security-supported	2016-04-07	TBD
<i>Liberty</i>	Security-supported	2015-10-15	2016-11-17
<i>Kilo</i>	EOL	2015-04-30	2016-05-02
<i>Juno</i>	EOL	2014-10-16	2015-12-07
<i>Icehouse</i>	EOL	2014-04-17	2015-07-02
<i>Havana</i>	EOL	2013-10-17	2014-09-30

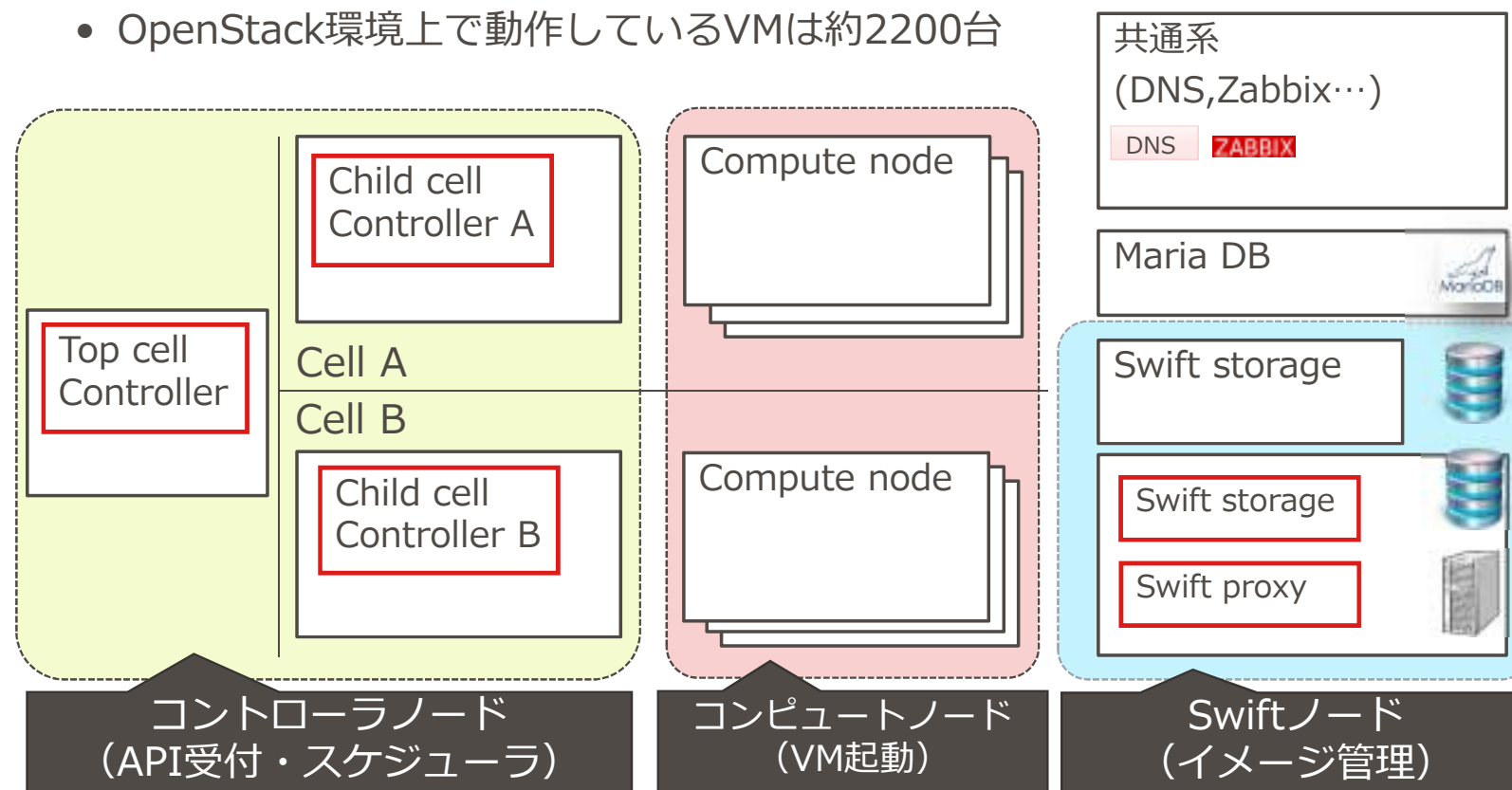
* <http://releases.openstack.org/>



OpenStack環境構成(1)

OpenStack環境の全体像

- 物理サーバは約400台で、すべてスペックは同一のものを使用
 - ※赤枠はKVM上で動作しているマシン
- OpenStack環境上で動作しているVMは約2200台



OpenStack環境構成(2)

- ・ 利用モジュール・バックエンド等
 - ・ VMのパフォーマンスを最重要視した構成としている

モジュール名	利用中	バックエンド	備考・特徴など
Keystone	Yes	認証基盤： →既存の共通認証システムと連携	特になし
Nova	Yes	仮想化： →Qemu/KVM ストレージ： →Qcow2をローカルディスクに保存	多量のComputeノードが存在する環境に対応するため、 Nova Cell(v1)を使用して、DBとMQを分割している
Neutron	Yes	テナント間NW隔離： →VLAN	IPとMacアドレスの管理に使用している。 仮想ルータや仮想FWなどのNFV機能は使用していない。
Glance	Yes	イメージ格納先： →Swift	イメージとスナップショットの管理に使用
Swift	Yes	オブジェクト格納先： →ローカルディスク	特になし
Horizon	Yes	-	運用ルールに基づいてカスタマイズして利用
Cinder	-	-	ブロックストレージサービスは提供しない



OpenStack環境構成(3)

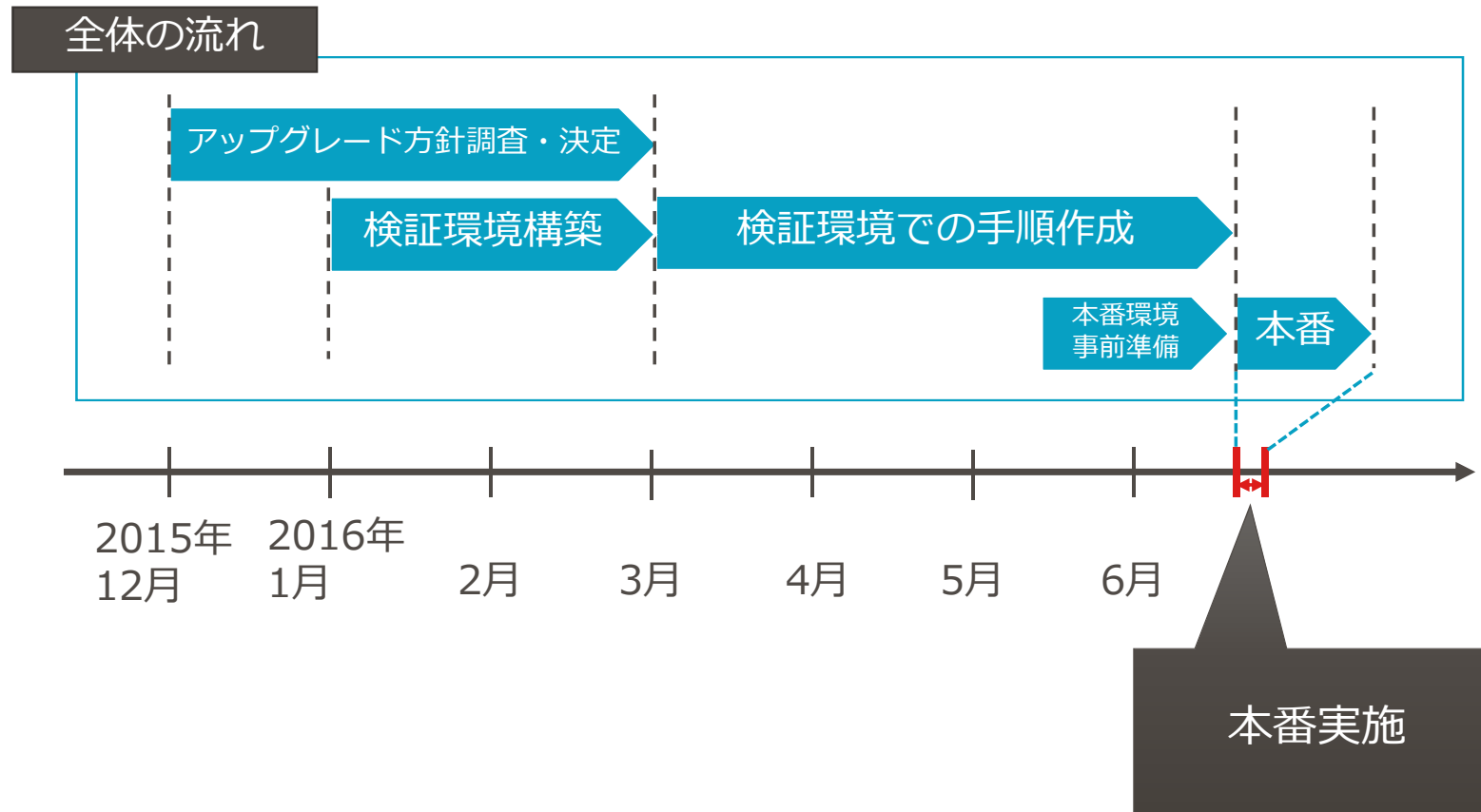
- 利用中のOS
 - CentOS 6.x 及び 7.xを使用
- OpenStackインストーラ
 - Icehouse版Packstack(Puppet)をベース
 - 前述の環境構成を実現するために独自に改造
- OpenStackパッケージ
 - RDOベース
 - Horizon修正と、致命的なバグは独自に適用



アップグレード全体の流れ

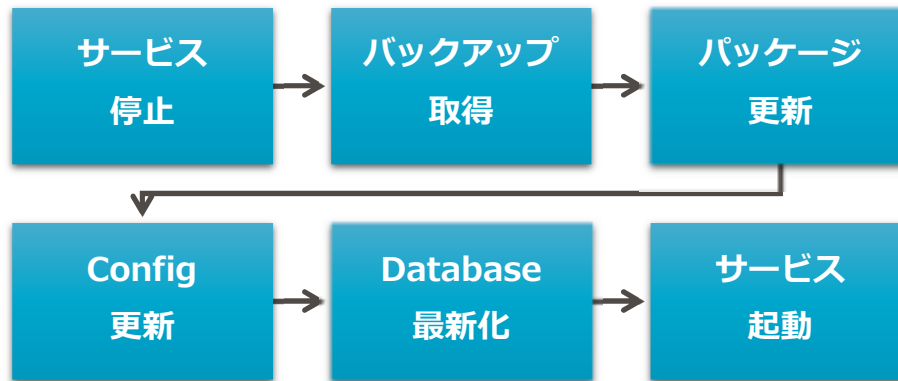


全体の工程・実施期間



アップグレード実施方針の調査

- 公式ドキュメントのガイドラインを調査*
 - アップグレードの基本的な流れが紹介されている



- 詳細な手順は環境に依存するため、独自に確立する
- 事前検証は絶対に必要
 - 「極力本番に近い環境を用意すること」と記されている
 - 今回は商用クラウドを用いて本番に近い環境を構築した

* http://docs.openstack.org/ops-guide/ops_upgrades.html



アップグレード実施方針の決定（1）



▼コールドアップグレードの方式を選択

- ・理想はユーザにアップグレード作業を意識させないレベルのローリングアップグレードだが、自動化や検証等の稼働増加を懸念
- ・APIサービスの停止は数日間可能という背景があるため（ユーザ資源（VM）の停止は当然NG）

▼Liberty搭載のサーバを新たに構築し、Icehouseのサーバから切り替えるとする方式とした

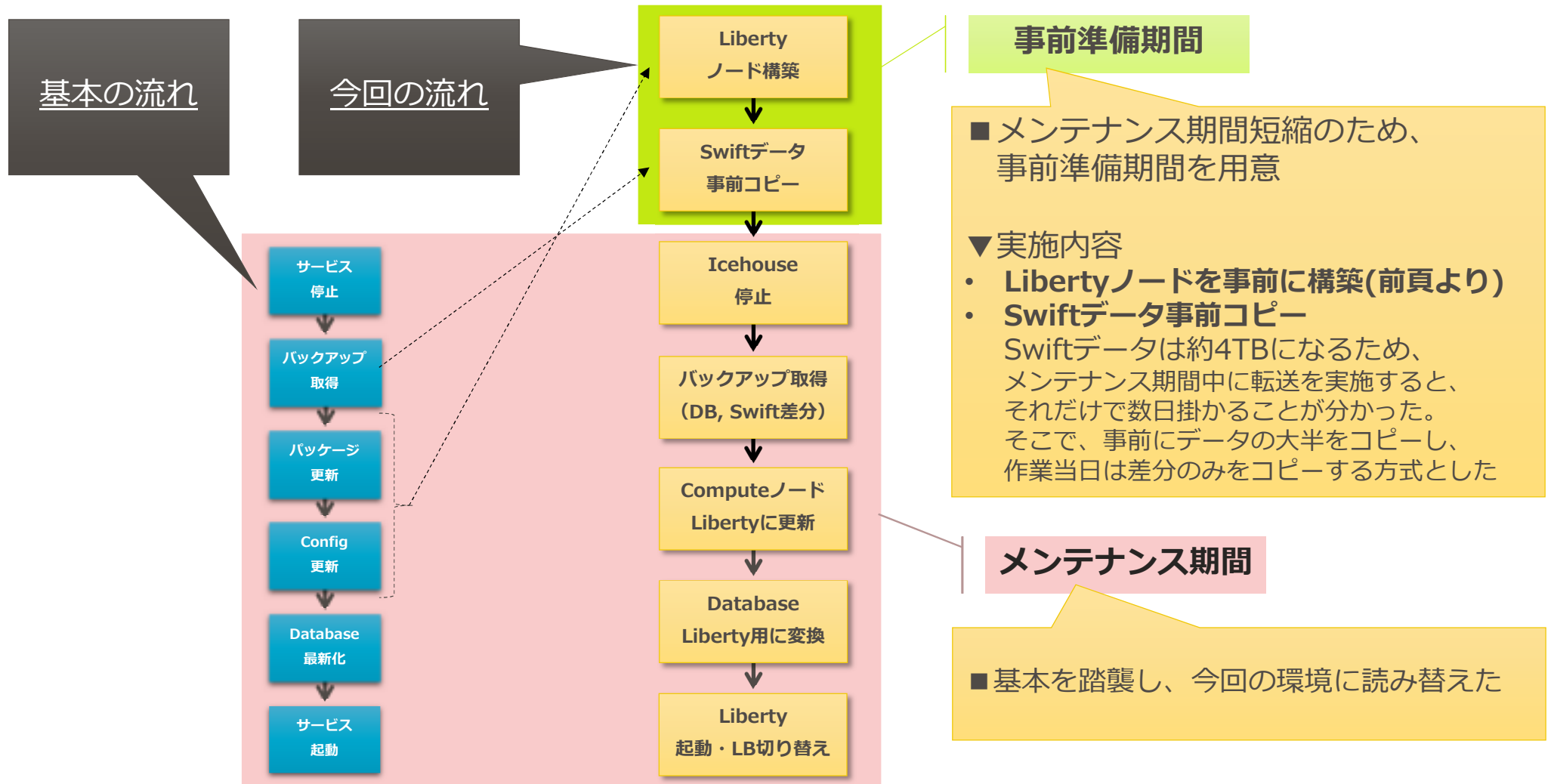
- ・具体的にはComputeノード以外を新たに用意する
- ・万が一に備え、ロールバックするための手順は必要と考えており、Icehouseに切り替えるだけで、簡単に元に戻せるようにするため

▼パッケージ更新とConfig更新は、構成管理ツール(Puppet)を利用して自動化する

▼DB更新はコミュニティが提供しているツールを利用して実施する

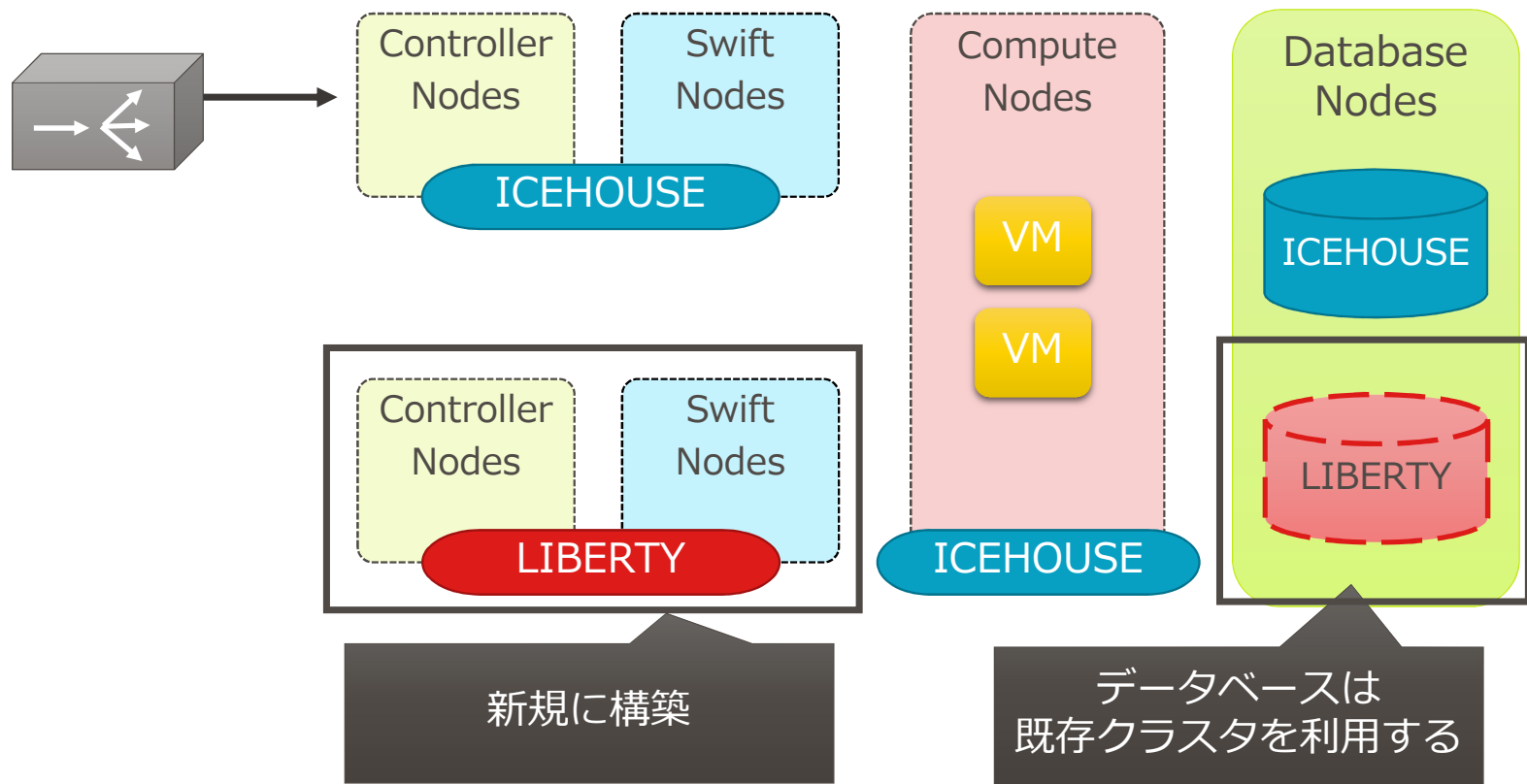
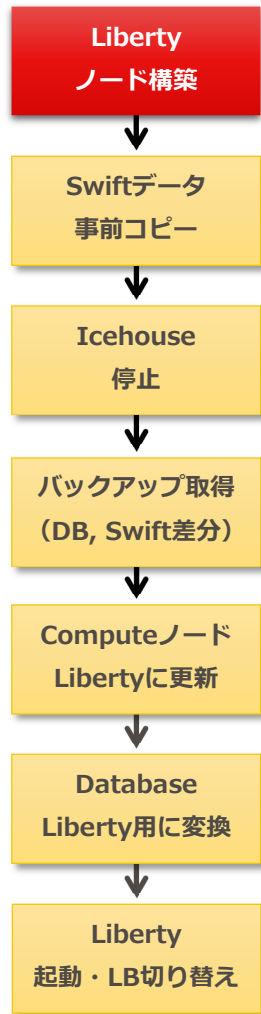


アップグレード実施方針の決定（2）



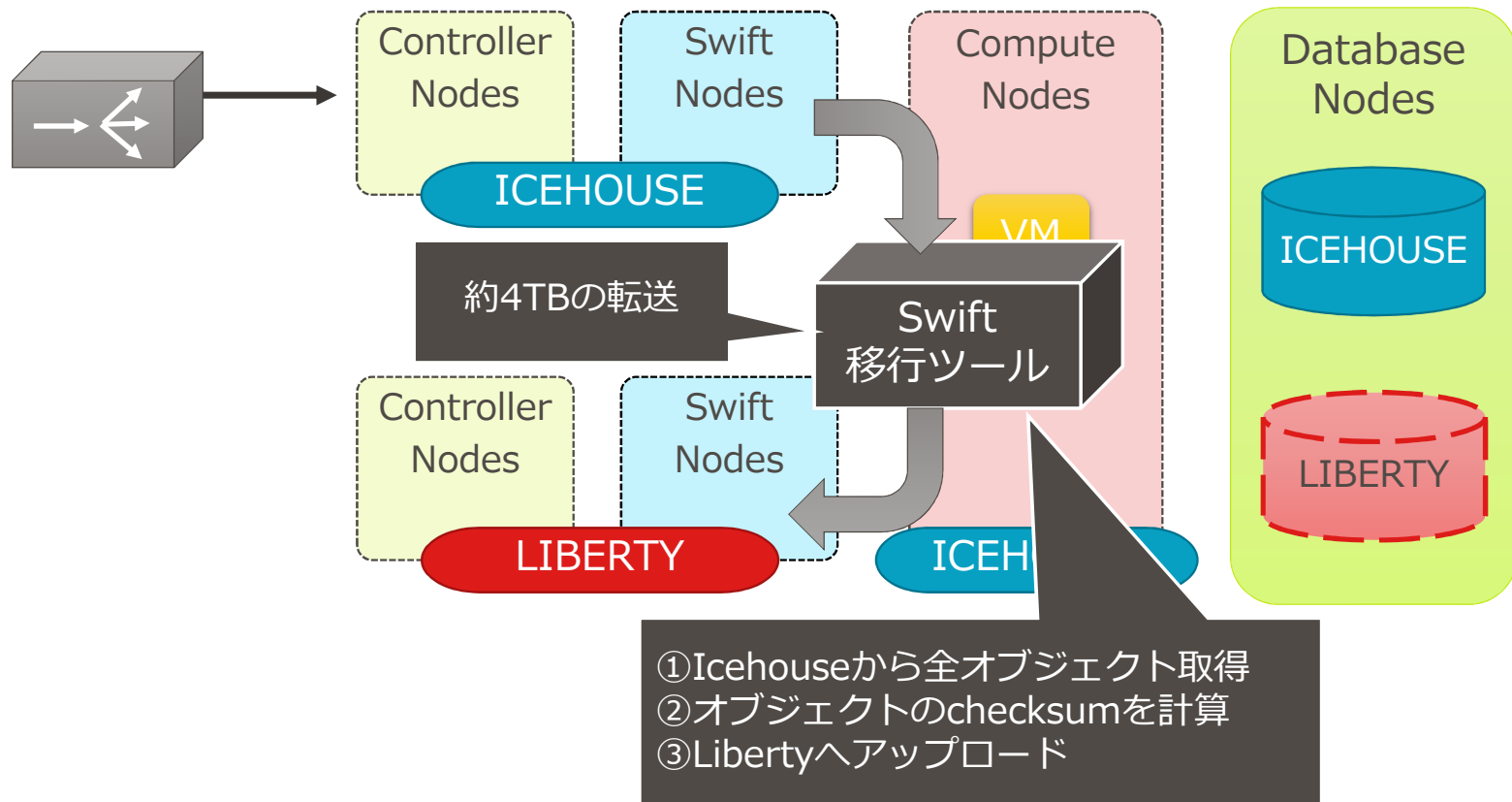
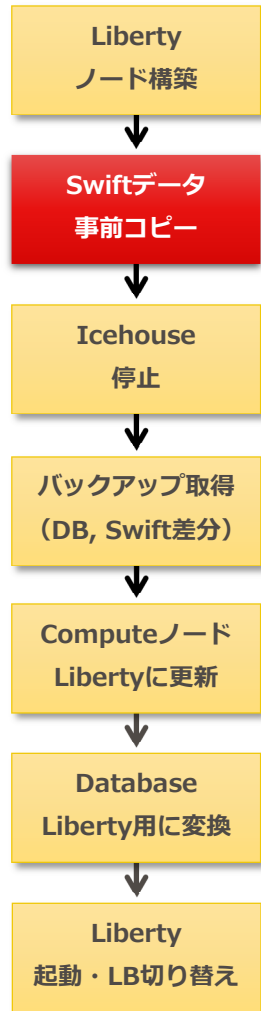
アップグレード実施手順の説明

- ControllerノードとSwiftノードを新たに構築する



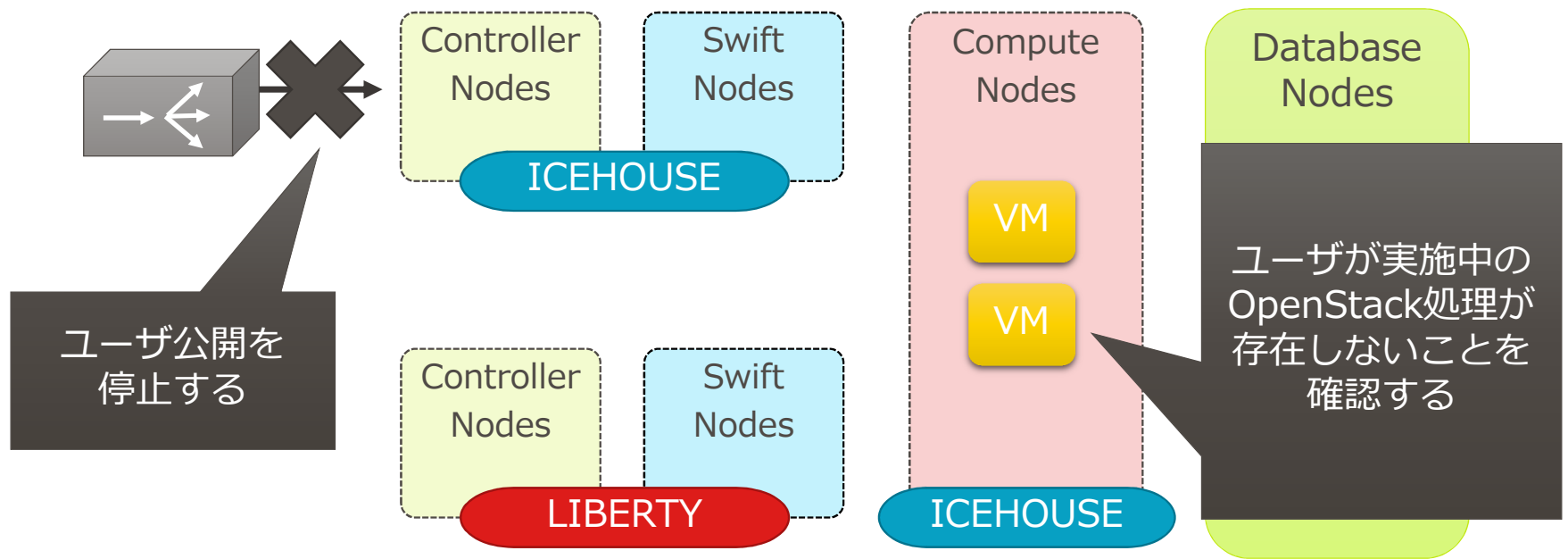
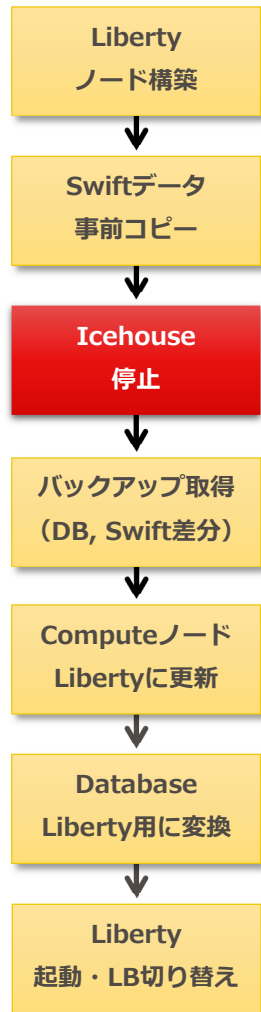
アップグレード実施手順の説明

- Icehouseサービス中にSwiftのデータをコピーする
 - ※サービス提供中の差分は公開停止後にコピーする



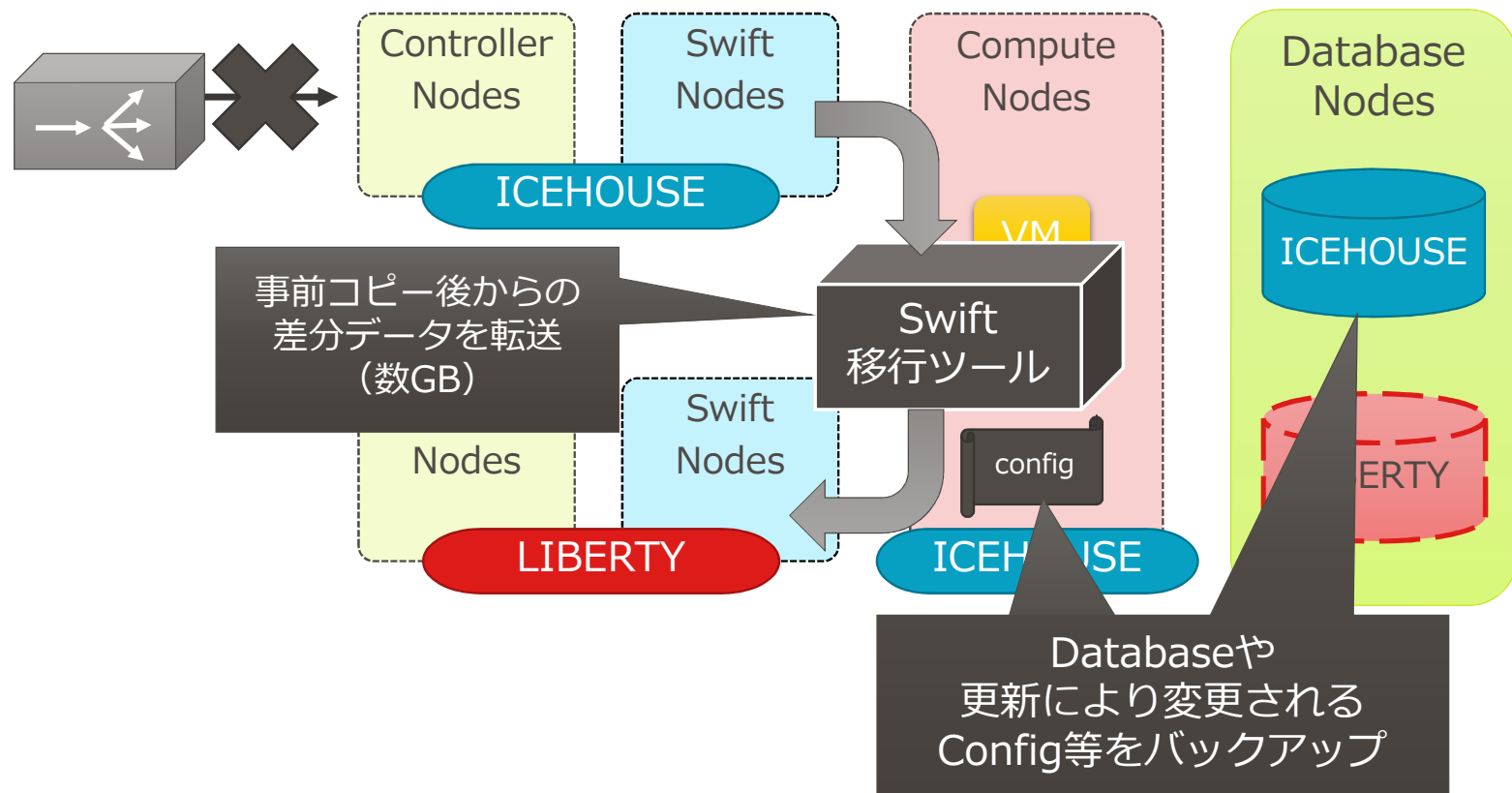
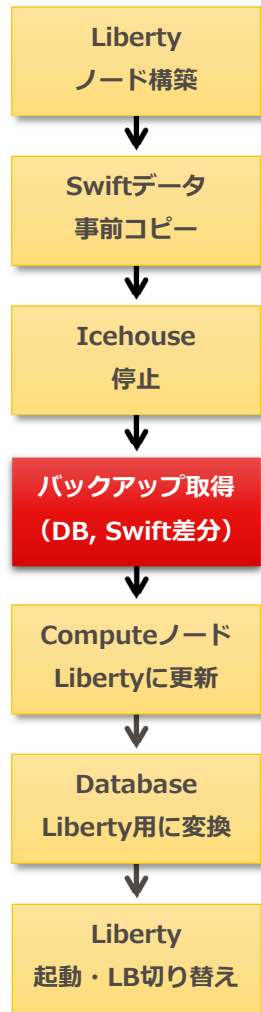
アップグレード実施手順の説明

- OpenStackのユーザ向けの公開を停止する
 - ※プロセス停止は、後述のSwiftデータ転送後



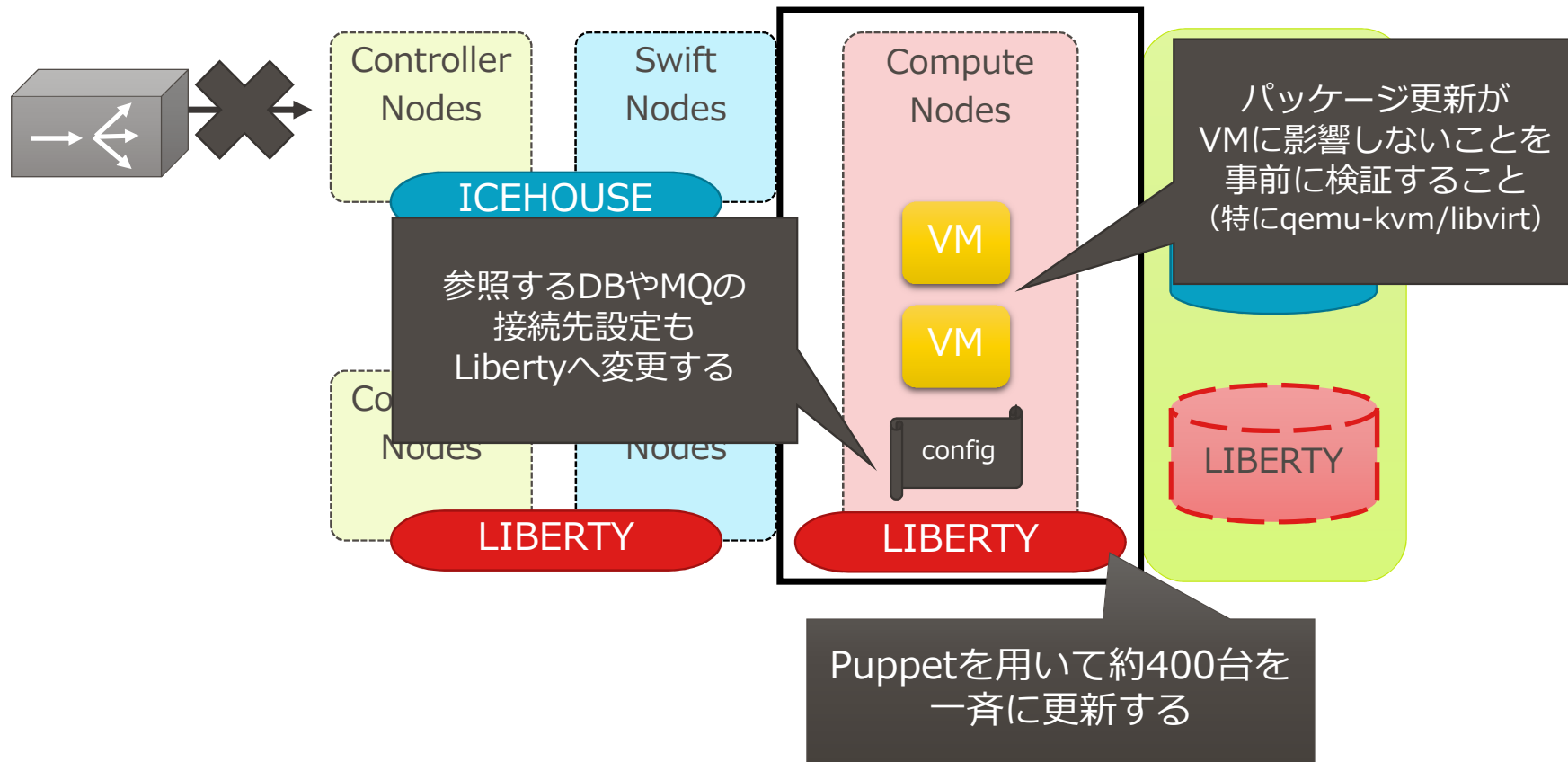
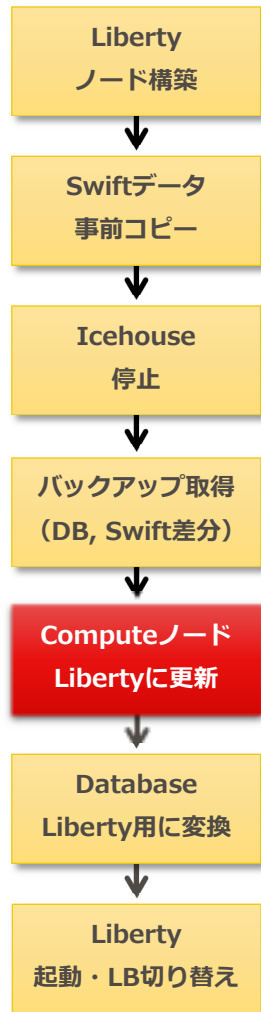
アップグレード実施手順の説明

- DBやConfig等、必要なバックアップを取得する
 - 差分データを転送後、OpenStack及び監視を完全停止する



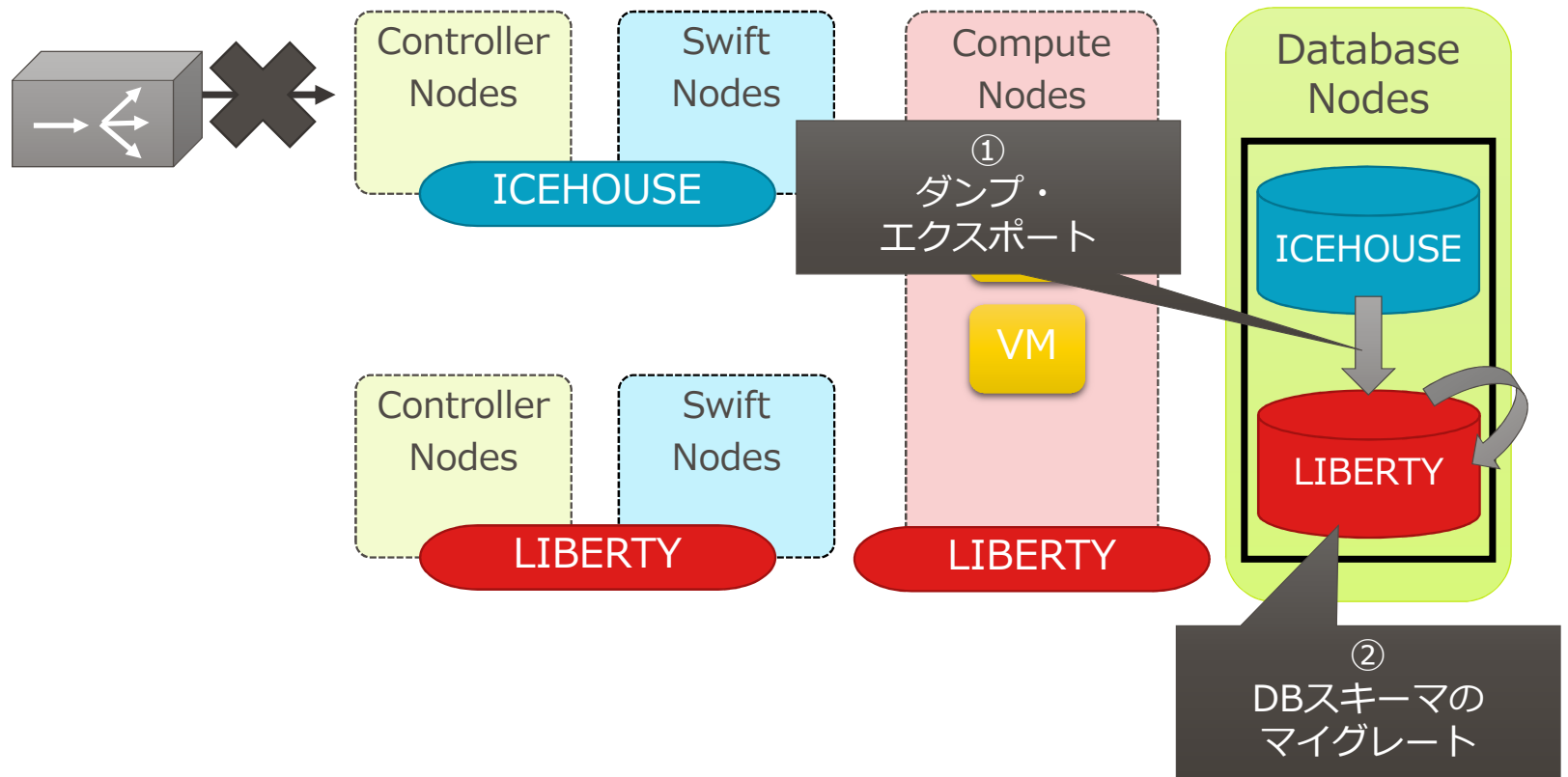
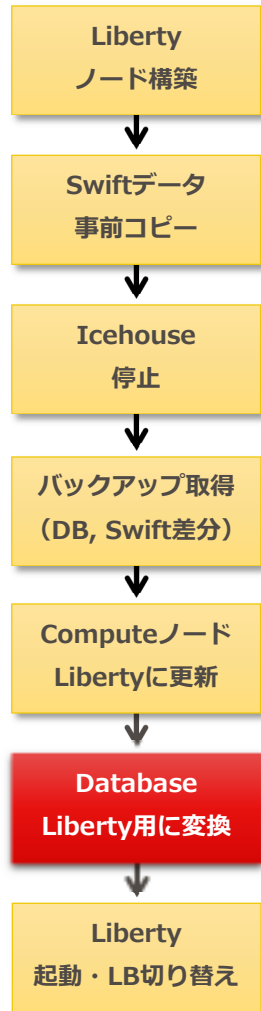
アップグレード実施手順の説明

- 「パッケージ」「Config」をLibertyのものへ更新する



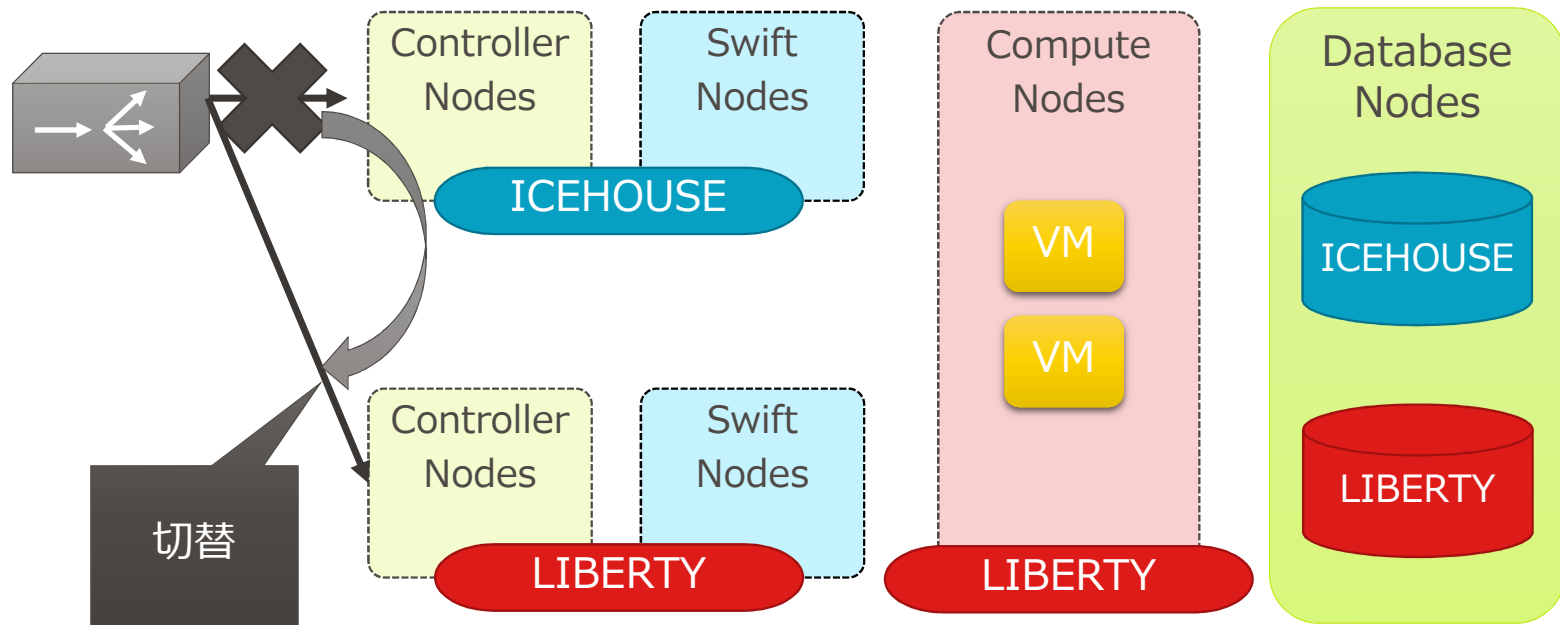
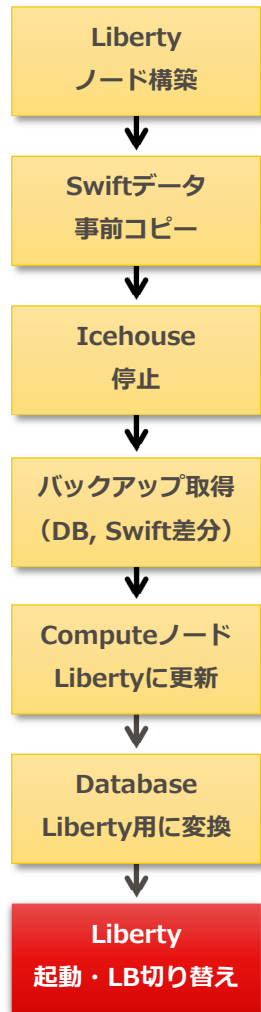
アップグレード実施手順の説明

- 作成済みのデータベースの中身を作成する
 - マイグレートにはCTノード(Liberty)の管理用ツールを利用



アップグレード実施手順の説明

- LBをlibertyノードに向け、サービスを再開する
 - 基本的な動作確認後に監視を再開する



検証内容についての紹介



事前検証の方針・観点

• 方針

- 可能な限り本番環境と同一構成の検証環境を構築し、アップグレード手順の作成およびアップグレード・切り戻しの実施を行う
- アップグレード後、正常性確認試験を実施する

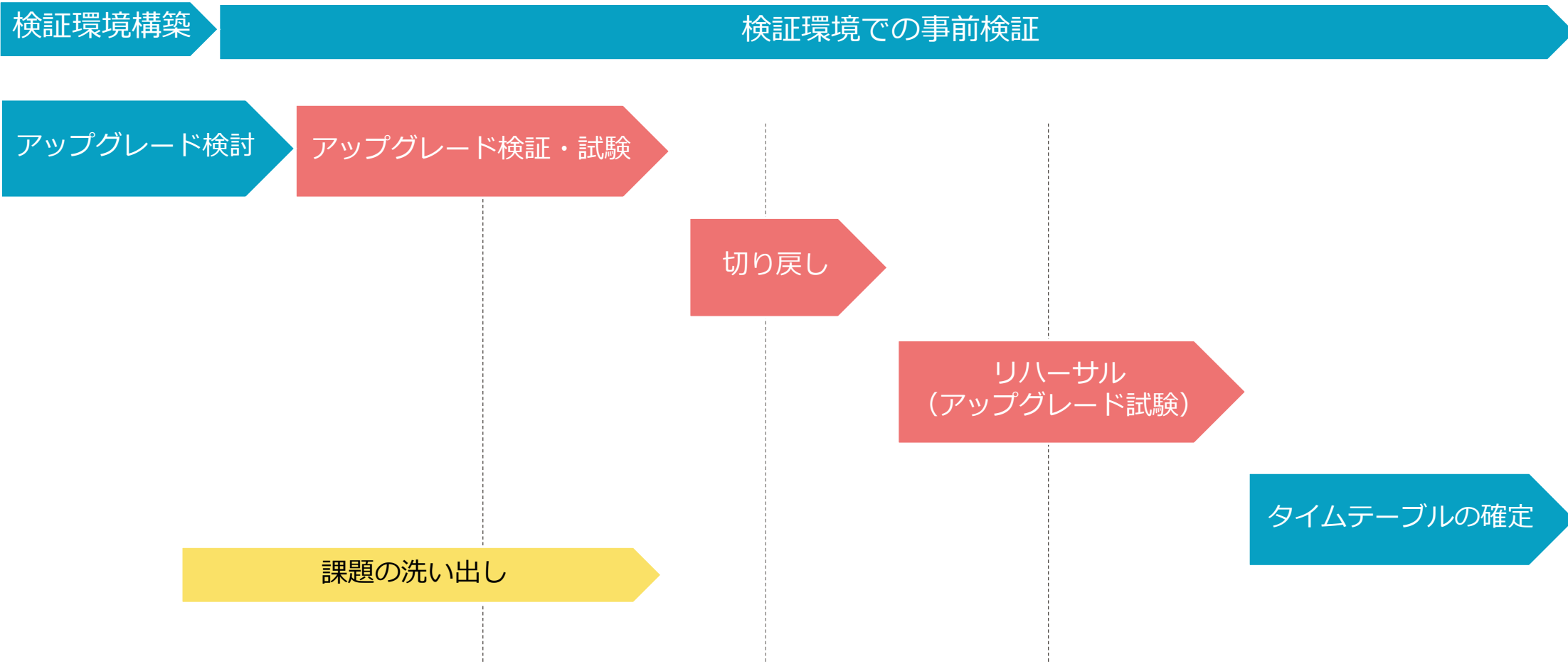
• 観点

- アップグレード手順に抜け漏れがないか
- アップグレード手順で既存VMに影響がないか
- アップグレード後、切り戻しが可能かどうか
- アップグレードおよび切り戻し手順にどの程度の時間を要するか
- 既存監視システムへの影響があるかどうか
- 本番データでDBマイグレートが実施できるかどうか



事前検証の流れ

検証環境での事前検証では、アップグレード2回、切り戻し1回を実施しました。



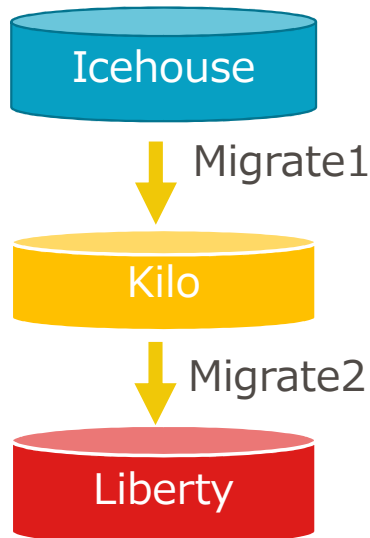
検証で発生した課題の紹介(DBマイグレート)

DBマイグレート試験にて課題を発見

Kilo環境(VM)

ダミー
RabbitMQ

DB



- 本番環境データを利用したマイグレート試験を行ったところ、LibertyへマイグレートするためにはKiloを経由する必要があることが発覚。
- Cell環境でマイグレートを実施する際はダミーのRabbitMQが必要になることが判明。

OpenStackの構成や本番環境データの状態によって、マイグレートが想定通りにいかないことがあります。

本番環境データおよび本番と同等構成の環境でDBマイグレート試験を早期実施することをオススメします。

検証で発生した課題の紹介(OpenStack)

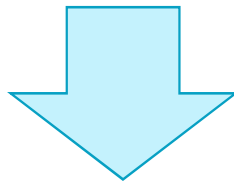
- 正常性確認試験にて、OpenStackの課題を発見
 - 致命的なバグが検出されることもあるため、事前検証は重要です

概要	対処
\$ nova listコマンドでVMのIPアドレスが表示されない	コミュニティからバックポート
Horizon上でプロジェクトの切り替えが出来ない	コミュニティからバックポート
Horizonの一部で2バイト文字の扱いに失敗する	独自対処 (コミュニティ報告中)
Shelveを実行すると複数のイメージが出来る	独自対処 (コミュニティ報告中)

コミュニティからのバックポートやコード修正、テストなどを確実に実施できる有識者、パートナーとの協力が重要となります。

検証で発生した課題の紹介(監視)

- 試験にて、旧バージョンで解決済みだった監視の問題が再燃
 - OpenStackの監視においては、ログの監視を行っており、無視しても良いログを正規表現にて育てていた。



- 正常性試験時に監視システムがエラーを検知し発報
 - ログのフォーマットが変更になったことにより、正規表現によって無視していたログが検知されるようになった。

検証環境にも本番環境と同様の監視環境を構築し、監視の検証も行う必要がある

環境によっては監視環境に過大な負荷が掛かり監視が停止する恐れも……



まとめ



まとめ

アップグレード手法は基本的にはコールドアップデートがお勧め

新バージョンのサーバを並行して構築しておくことで切り替え時間短縮・切り戻しも可能となる

実データを用いて初めて発覚する問題もあるので、早期に実データを用いた試験を実施しよう

コミュニティからのバックポートやコード修正などを実施できる有識者、パートナーとの協力が重要

本発表が皆様のOpenStack導入や、
アップグレードの実施など、何らかのお役に立てれば幸いです



質疑応答

- NTTソフトウェアは22番ブースに出展しています
 - 是非お越しください！！

