

第2章: ICP Kubernetes環境を構築するための 技術検討

章内目次

第2章: ICP Kubernetes環境を構築するための技術検討とその内容

第1節: コンピューティング基盤技術検討

第2節: ネットワーク基盤技術検討

第3節: ストレージ基盤技術検討

第4節: ハイパーバイザー層技術検討

第5節: OS層技術検討

第6節: Dockerエンジン構成技術検討

第7節: ICPクラスター構成技術検討

第8節: 関連システム連携技術検討

第9節: 構成上のNGパターン集

第10節: 検討結果の実装への反映

ICP k8s環境のシステム構成検討のために目指すべきゴール（その1）

■ ICP K8s環境のシステム構成の検討に取り組む際での目指すべきゴール（技術検討にあたっての主要なマイルストーン）を下表に列挙する

–ICP k8s環境のシステム構成検討について、下表のカテゴリ毎での解説を以降の各節で行う。

ソリューション・スタック内で 技術検討が属するレイヤー	レイヤー内での検討対象	技術検討により目指すべきゴール (技術検討にあたっての主要なマイルストーン)
物理基盤層	コンピューティング基盤	1. プラットフォーム（Linux x86_64 /Linux on Power (ppc64le)/Linux on IBM® Z）を選定すること 2. 「ノード」の構成方針（ノード役割の配置方針、およびノード冗長化方針）を明確にすること 3. 物理サーバー（もしくは仮想サーバー）の構成（スペック、および台数）を導き出すこと
	ネットワーク基盤	1. ICP k8s環境が利用するネットワークの一覧を明らかにすること 2. ICP k8s環境が利用するネットワークのICP k8s環境を用いる組織が運用するネットワーク群における位置づけ（サービスLANに属するのか？管理LANに属するのか？孤立しているのか？互いに疎通するのか？…etc.）を明らかにすること 3. ICP k8s環境の各ノードでの複数NIC構成方針を明らかにすること 4. ICP k8s環境でのノード間通信での実装方針（どのネットワークを使用するのか？通信は暗号化するのか？…etc.）を明らかにすること
	ストレージ基盤	1. ICP k8s環境上で稼動するアプリケーションが（データ保管のために）用いる永続ストレージの実装方式を明らかにすること 2. ICPの各ノードが（ICP k8s環境の運用のために）必要とするストレージ領域に関わる要件の内容を明らかにすること
Hypervisor層	-	1. ハイパーバイザー層の利用の有無を決定すること 2. 使用するハイパーバイザー製品を選択すること

ICP k8s環境のシステム構成検討のために目指すべきゴール（その2）

■ ICP K8s環境のシステム構成の検討に取り組む際での目指すべきゴール（技術検討にあたっての主要なマイルストーン）を下表に列挙する（続き）

–ICP k8s環境のシステム構成検討について、下表のカテゴリ毎での解説を以降の各節で行う。

ソリューション・スタック内で 技術検討が属するレイヤー	レイヤー内での検討対象	技術検討により目指すべきゴール (技術検討にあたっての主要なマイルストーン)
OS層	-	1. 使用するOS（Linuxディストリビューション）を選定すること 2. ICP製品が定める、OS構成に関わる各種要件を把握すること
コンテナ・ランタイム層	Dockerエンジン構成	1. Dockerエンジンの各ノードへの展開方式を明らかにすること 2. Dockerエンジンのバージョンを選定すること 3. Dockerエンジンを各ノードに自動展開する場合には、Docker関連の各種設定を各ノードに反映するための方式を理解すること
コンテナ・オーケストレーション層	ICPクラスター構成	1. ICPクラスターという概念を理解すること 2. ICPクラスターに対して適用出来る設定項目の全体像を把握すること
関連システム連携	-	1. ICP k8s環境が連携し得る関連システムの存在を把握すること 2. ICP k8s環境と関連システムとの間での連携方式を明らかにすること

第2章 第1節: コンピューティング基盤技術検討

コンピューティング基盤における検討事項: 検討のゴール

■ コンピューティング基盤での検討を通じて目指すべきゴールは以下の通り

1. プラットフォーム（Linux x86_64 /Linux on Power (ppc64le)/Linux on IBM® Z）を選定すること
2. 「ノード」の構成方針（ノード役割の配置方針、およびノード冗長化方針）を明確にすること
3. 物理サーバー（もしくは仮想サーバー）の構成（スペック、および台数）を導き出すこと

コンピューティング基盤構成における検討事項: プラットフォームの選択

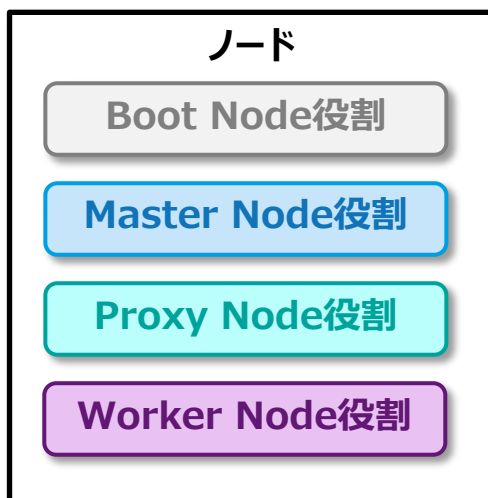
- ICP k8s環境がサポートするプラットフォームは以下の3種類である
 - Linux x86_64
 - Linux on Power (ppc64le) POWER8® or later
 - Linux on IBM® Z

- ICP製品が提供する機能へのサポート有無はプラットフォームにより異なる
 - 「Linux x86_64」プラットフォームが機能網羅性に優れる
 - 「Linux on IBM® Z」は「Linux x86_64」「Linux on Power (ppc64le)」と比較しサポートされる機能が少ないため、利用の際には以下の「Supported features by platform」を参照すること。
 - 参照: [IBM Knowledge Center - サポートされるオペレーティング・システムとプラットフォーム](#)

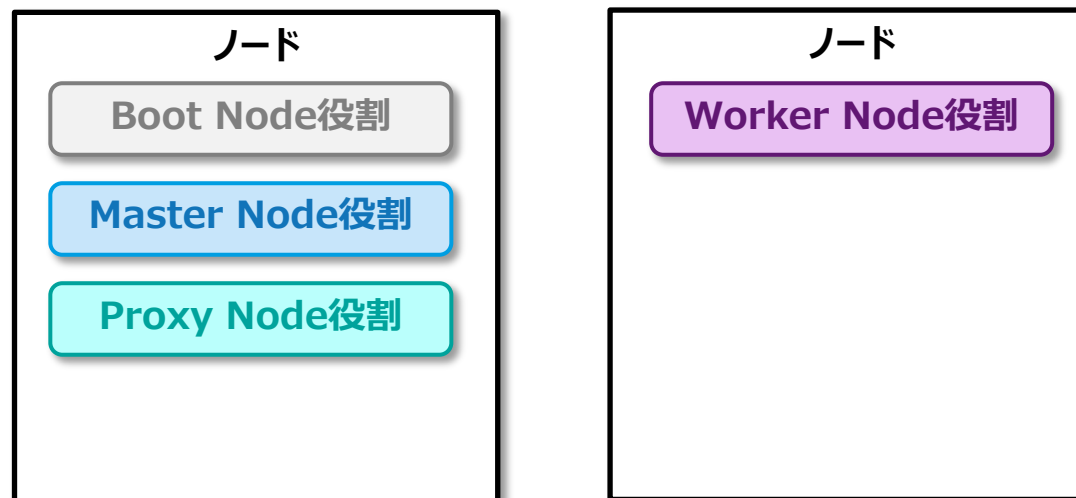
コンピューティング基盤における検討事項: ノード役割の配置 (その1)

- ICP k8s環境ではノードに対して複数の「ノード役割」を割り当てることが出来る
 - ノード役割のノードへの割り当て例を以下に示す

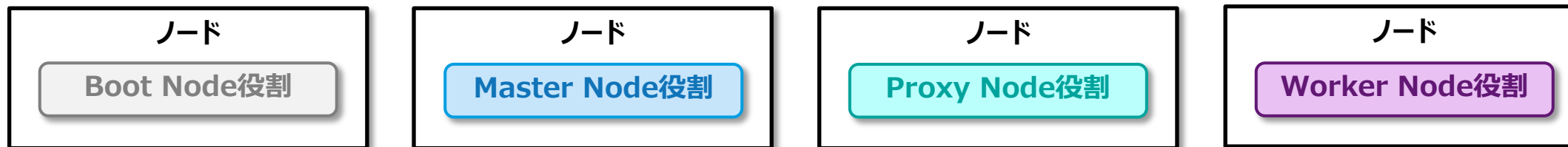
配置例: 単一ノードに全ノード役割を配置



配置例: Worker Node役割を他の役割から分離



配置例: 役割毎にノードを分割



コンピューティング基盤における検討事項: ノード役割の配置 (その2)

■ ICP k8s環境での「ノード役割」は以下の7種類である

- Boot Node
 - 役割: ICP k8s環境のセットアップを制御する
- Master Node
 - 役割: ICP k8s環境全体を制御する
- Management Node
 - 役割: ICP k8s環境を対象とした管理サービスを提供する
 - このノード役割は環境に存在しなくても良い（その場合、Master Nodeがその役割を代行する）
- Proxy Node
 - 役割: ICP k8s環境上で稼動する業務サービスに対する通信を中継する
- Worker Node
 - 役割: ICP k8s環境上で稼動する業務サービスを実行する
- VA Node
 - 役割: ICP k8s環境のPrivate registry上にあるContainer Imagesのセキュリティーチェックを行う
- etcd Node
 - 役割: etcdの分散key value storeを稼働させる
 - etcdノードを用意しない場合、etcdはMasterノードに配置される

コンピューティング基盤における検討事項: ノード役割の配置 (その3)

■ ICP k8s環境でノードに対して役割を割り当てる際の制約は以下の通り

– 役割が割り当てられるノードは、役割毎に定められた性能要件を満たすこと

• 参照: [IBM Knowledge Center - ハードウェア要件および推奨](#)

– 「IBM® Z」プラットフォームを用いるノードにはVAノード役割は割り当てられない

• 参照: [IBM Knowledge Center - サポートされるオペレーティング・システムとプラットフォーム](#)

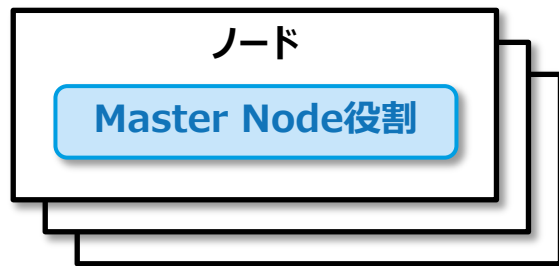
Table 2. Supported node type by platforms

Node type	Linux x86_64	Linux on Power (ppc64le)	Linux on IBM® Z (s390x)
Boot	Y	Y	Y
Master	Y	Y	Y
Management	Y	Y	Y
Proxy	Y	Y	Y
Worker	Y	Y	Y
VA	Y	Y	N

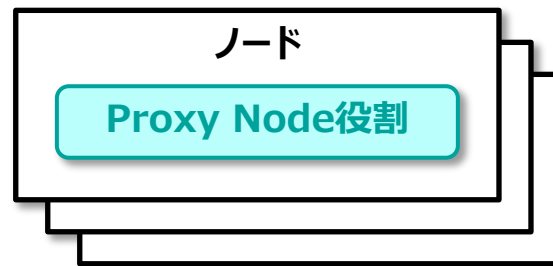
コンピューティング基盤における検討事項: ノード役割の配置 (その4)

- 可用性要件、性能要件への対応のため、ノードを冗長化することが出来る
 - Master Node、およびProxy NodeでのHA構成が可能
 - (Master NodeでのHA構成においては冗長度=3 or 5が推奨)
 - 参照:[IBM Knowledge Center - 高可用性 IBM® Cloud Private クラスター](#)
 - Worker Nodeでは任意のスケールアウト構成が可能
 - ICP k8s環境で稼動するコンテナを複数のWorker Node間に分散配置することにより、システムの可用性、およびスケーラビリティが向上する

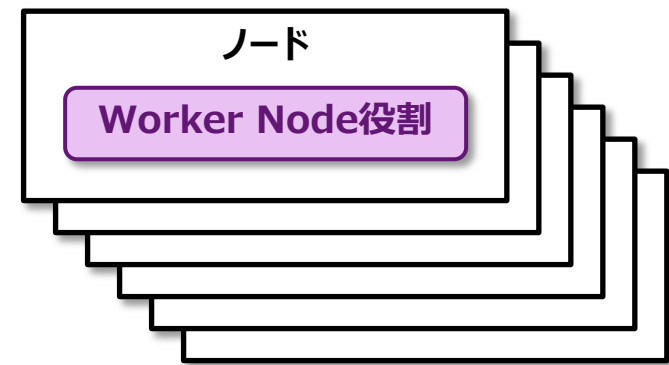
Master Node (3多重 or 5多重)



Proxy Node



Worker Node (多重度は任意)



コンピューティング基盤構成における検討事項: ノードサイジング

■ ノードを稼動させるための最低性能要件が課されている

- (ICP製品のバージョンアップに伴い要件は頻繁に変更されるため、常に最新の要件を確認すること)

• 参照: [IBM knowledge Center - ハードウェア要件および推奨](#)

Table 2. Minimum hardware requirements for a multi-node cluster

Requirement	Boot node	Master node	Proxy node	Worker node	Management node	VA node	etcd node
Number of hosts	1	1, 3, or 5	1 or more	1 or more	1 or more	1, 3, or 5	1 or more odd number of nodes
Cores	1 or more	8 or more	2 or more	2 or more	8 or more	8 or more	1 or more
CPU	>= 2.4 GHz	>= 2.4 GHz	>= 2.4 GHz	>= 2.4 GHz	>= 2.4 GHz	>= 2.4 GHz	>= 2.4 GHz
RAM	>=4 GB	>=16 GB	>=4 GB	>=4 GB	>=16 GB	>=16 GB	>=4 GB
Free disk space to install	>=100 GB	>=300 GB	>=150 GB	>=150 GB	>=300 GB	>=300 GB >=800 GB (production environment)	>=100 GB

■ Cluster構成を考える際の参考として、サンプルデプロイメントサイズに関するガイドが提供されている

- Small, Medium, Large(worker node数500の場合および1000の場合)に分けてサンプル構成を提供

• 参照: [IBM knowledge center - IBM® Cloud Private クラスターのサイジング](#)

コンピューティング基盤構成における検討事項: メモリー容量の確保

■ メモリー・スワップを利用しない前提の元で各ノードのメモリーを確保すべき

- コンテナによるメモリー・スワップ利用を一元的に制御するための機能をKubernetes製品が提供していないことに起因して、メモリー・スワップが生じ得る環境においてコンテナの性能を適切に維持・管理することには困難が伴う
- (こうした背景の元で) Kubernetes製品のバージョン1.8以降では、メモリー・スワップの利用が無効化されていないノードでのコンテナの稼働を行わない設定が既定値とされている

• 参照: [GitHub kubernetes/CHANGELOG-1.8.md](https://github.com/kubernetes/kubernetes/blob/master/CHANGELOG-1.8.md)

- 考慮事項①: ICP製品に組み込まれたKubernetes実装では、(オープンソース・ソフトウェアとしてのKubernetes実装とは異なり) OSによるメモリー・スワップの利用が有効化されているノードでもコンテナの稼働を認める設定が既定値となっている(そのため、意図せぬメモリー・スワップの利用に伴う性能上の不具合への配慮が必要)
- 考慮事項②: ICP製品が用いているコンテナ・ランタイム製品であるDocker製品自体は、個々のコンテナによるメモリー・スワップの利用を制御するための機能を提供している

• 参照: [Specify a container's resources | Docker Documentation](https://docs.docker.com/config/containers/container-networking/#memory)

第2章 第2節: ネットワーク基盤技術検討

ネットワーク基盤における検討事項: 検討のゴール

■ ネットワーク基盤での検討を通じて目指すべきゴールは以下の通り

1. ICP k8s環境が利用するネットワークの一覧を明らかにすること
2. ICP k8s環境の各ノードでの複数NIC構成方針を明らかにすること
3. ICP k8s環境でのノード間コンテナ通信での実装方針（どのネットワークを使用するのか？通信は暗号化するのか？…etc.）を明らかにすること

ネットワーク基盤構成における検討事項: 複数NIC構成

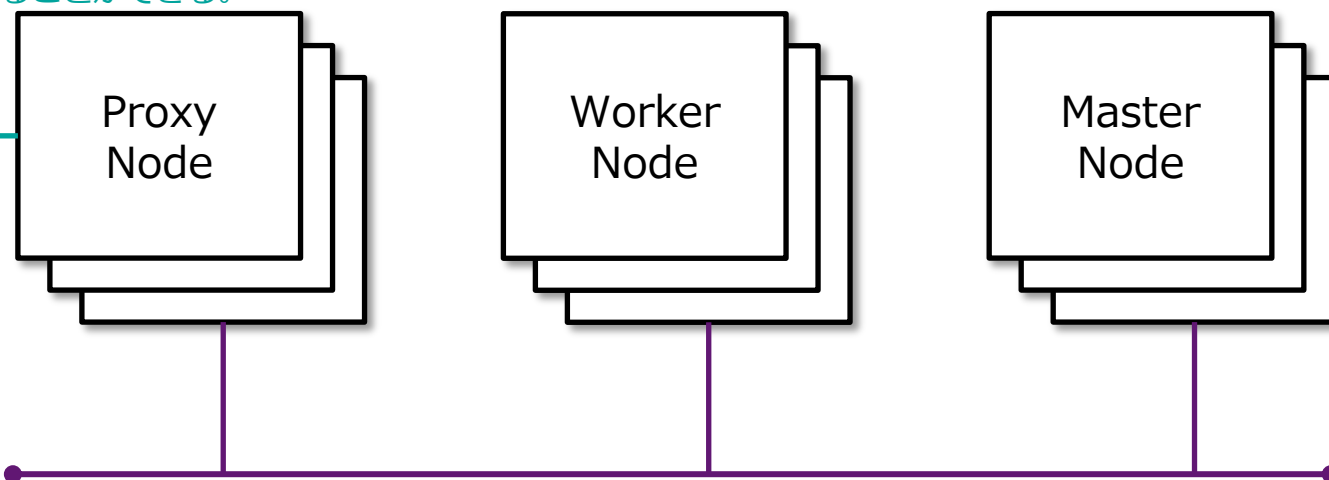
■ 各ノードでの複数NIC構成に関わるICPの設定項目を下図に示す

- ・参照: [IBM Knowledge Center - config.yaml ファイルを使用したクラスターのカスタマイズ](#)

フロントエンド・ネットワーク

NICが複数ある環境では、「[config.yaml](#)」設定ファイルで「`proxy_lb_address`」設定値を用いることにより、フロントエンド・ネットワークからNodePortリソースへのアクセスのためのIPアドレスを指定する。

Proxy NodeがHA構成を採る場合、「`proxy_vip`」設定値に仮装IPアドレスを設定する。また、「`proxy_vip_iface`」設定値には、「`proxy_vip`」が使用する仮想IPインターフェースを指定することができる。



クラスター管理ネットワーク

「[config.yaml](#)」設定ファイルでの「`cluster_vip`」設定値には、HA環境でのVIPを設定する。また、「`vip_iface`」設定値を用いることで、Master NodeがHA構成を採る場合での「`cluster_vip`」が使用する仮想IPインターフェースを指定することができる。また、NICが複数ある環境では、「[config.yaml](#)」設定ファイルで「`cluster_lb_address`」設定値に、クラスター内の管理サービス用のパブリックIPアドレスまたは外部IPアドレスを指定する。HA環境では、「`cluster_lb_address`」に設定したIPアドレスによって「`cluster_vip`」に設定したIPアドレスが置き換えられる。

ノード間コンテナ通信ネットワーク

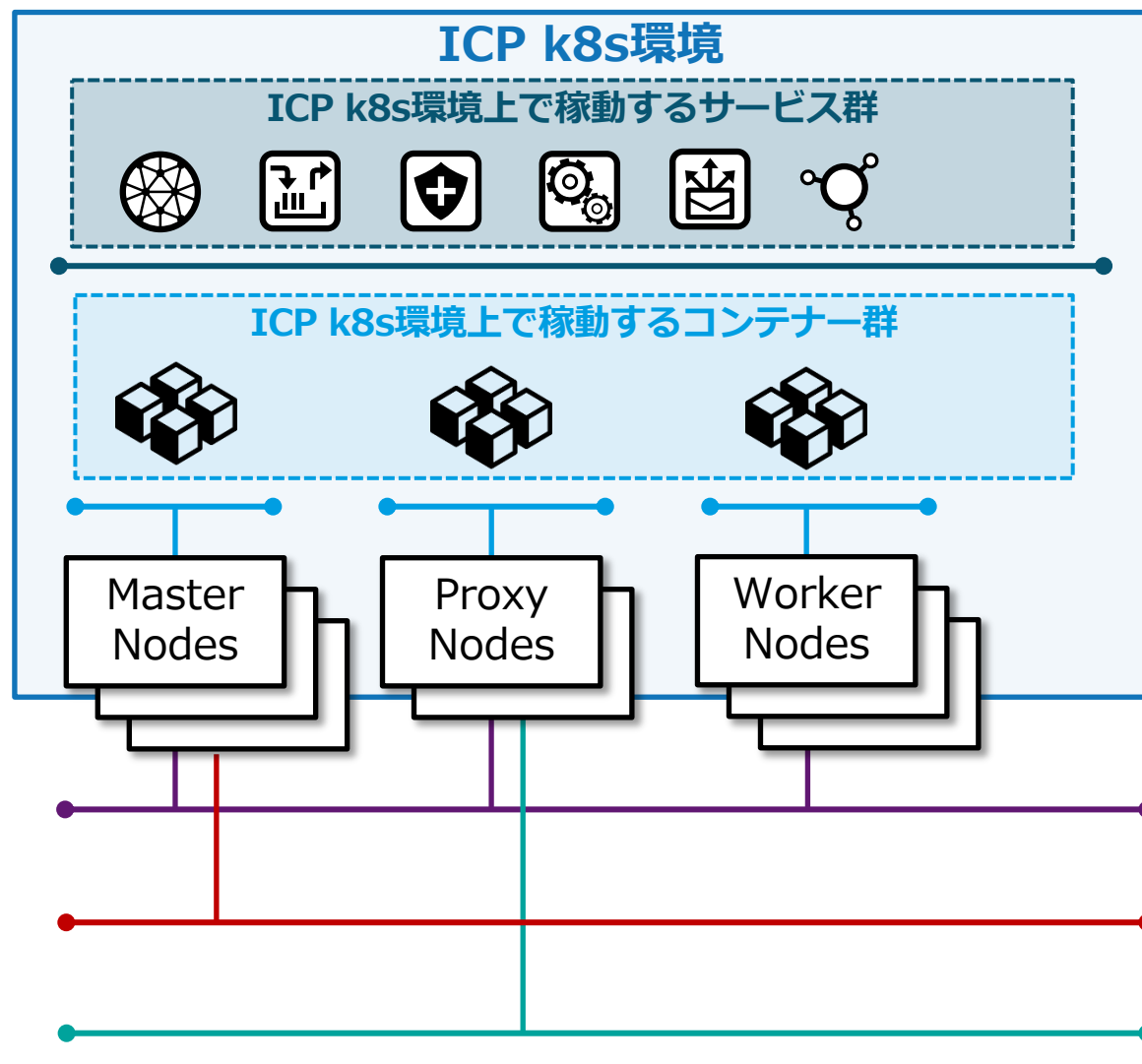
NICを選択する際、デフォルトでは「[config.yaml](#)」設定ファイルで「`calico_ip_autodetection_method: can-reach={{ groups['master'][0] }}`」となっており、`/cluster//hosts` ファイルに記載したMasterノードのIPアドレスと通信できるNICが選択される。

明示的にNICを変更したい場合は、「`calico_ip_autodetection_method: interface`」等の設定値を用いることにより設定が可能である。

参照: [Calico - Configuring calico/node](#)

ネットワーク基盤構成における検討事項: ネットワーク体系

- 下図に示す5種類のネットワークを明示的に意識したシステム構成が必要である。



クラスターIP所属ネットワーク

ICP k8s環境での「Service」オブジェクトに付与されたIPアドレスが所属する。
「[config.yaml](#)」設定ファイルでの「service_cluster_ip_range」設定値に相当する。

コンテナIP所属ネットワーク

ICP k8s環境のコンテナ（Pod）に付与されるIPアドレスが所属する。
「[config.yaml](#)」設定ファイルでの「network_cidr」設定値に相当する。

ノード間コンテナ通信ネットワーク

ノードを越えたコンテナ（Pod）間通信の転送経路となる

クラスター管理ネットワーク

ICPクラスターへの管理アクセスに用いる

フロントエンド・ネットワーク

ICP k8s環境上で稼動するサービスへのProxy Node経由でのアクセスを受け付ける

ネットワーク基盤構成における検討事項: Calico関連の検討事項

- ICP k8s環境でのノード間通信を制御/運用するCalicoソフトウェアに関連する以下の設定値、および、関連する実装に留意すること（下表）

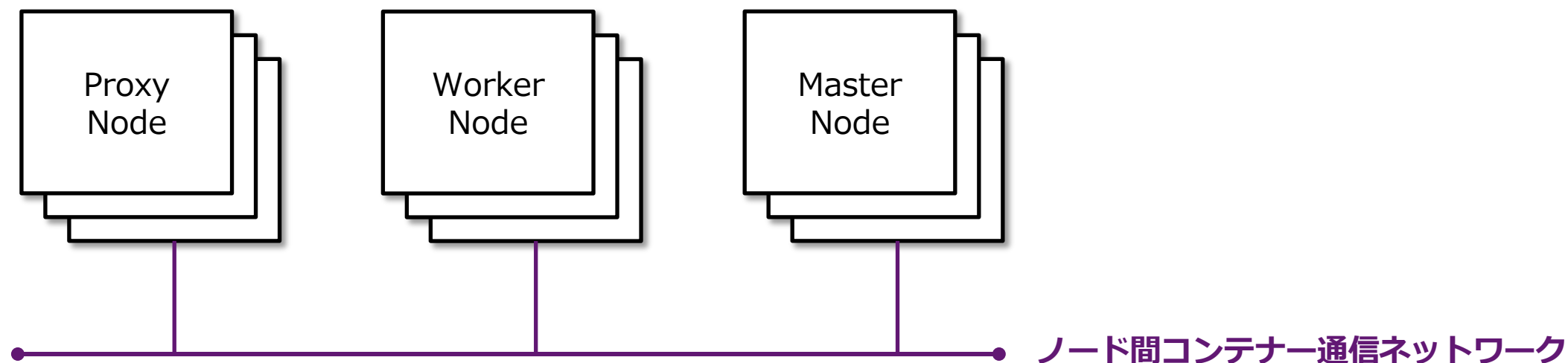
・参照: [IBM Knowledge Center - config.yaml ファイルを使用したクラスターのカスタマイズ](#)

設定箇所	設定値	既定値	解説
ICPクラスター・セットアップのための「config.yaml」ファイル	network_type	calico	「Calico」および「nsx-t」を選択可能。 「nsx-t」を使用する場合は、別途いくつかのパラメーターを設定する必要がある。 参照: IBM Knowledge Center - VMware NSX-T 2.3 と IBM Cloud Private の統合
	calico_ipip_mode	Always	デフォルトでは「Always」となっており、pod間の通信としてクラスターノードをまたいでIP in IPトンネルメッシュが作成される。異なるサブネットにあるノード間通信のためのIP in IPトンネルを作成したい場合はパラメータ値を「CrossSubnet」に設定する。 以下に示す状況においては、設定値は「Always」あるいは「CrossSubnet」としてIP in IP modeを有効化しなければならない。 ・ ノードが複数サブネットに分散しており、かつBGPが利用出来ない場合 ・ ノードがL3スイッチとして稼動することが許されない場合（例:OpenStack環境） 一方で、IP over IP modeを利用した場合、IP over IP処理の負荷が高いことに起因してネットワーク・スループットが落ちる可能性がある点に留意のこと。
	calico_tunnel_mtu	1430	CalicoでIP over IP modeを用いる場合のMTUサイズを定義する。 また、ノードのNICでのMTUを、ここで指定したMTUサイズより20以上大きくしない場合には通信スループットに悪影響が生じるとされている（そのため、IP over IP modeを処理するネットワークのMTUサイズでも同様の配慮を行うべき）。
	calico_ip_auto_detection_method:	can-reach={{ groups ['master'][0] }}	Calicoによる、ノード間通信のために用いるNICの選択方式を指定する。 デフォルトの設定では、/etc/hostsに記載したMasterノードのIPアドレスと通信できるNICが選択される。その他の設定を用いる場合は、「first-found」あるいは「interface」パラメータを設定することができる。 参照: Calico - Configuring calico/node

ネットワーク基盤構成における検討事項: 暗号化による通信の保護

■ ノード間コンテナ通信ネットワーク上のトラフィックを暗号化出来る

- ・参照: [IBM Knowledge Center - IPsec を使用したクラスター・データ・ネットワーク・トラフィックの暗号化](#)



「[config.yaml](#)」設定ファイルに「ipsec_mesh」セクションを追加することによりIPsecによる通信暗号化を有効とすることが出来る。
その際に、以下の設定値の指定が求められる。

- IPsecによる暗号化通信のために用いるノードでのネットワーク・インターフェース（専用のNICが必要）
- IPsecによる暗号化通信を有効とするネットワーク・サブネット
- IPsecによる暗号化通信の対象から除外するIPアドレス

また、ICPクラスターインストール後にもIPsecによる暗号化通信の有効/無効を切り替えることができる。

第2章 第3節: ストレージ基盤技術検討

ストレージ基盤における検討事項: 検討のゴール

■ ストレージ基盤での検討を通じて目指すべきゴールは以下の通り

1. ICP k8s環境上で稼動するアプリケーションが（データ保管のために）用いる永続ストレージの実装方式を明らかにすること
2. ICPの各ノードが（ICP k8s環境の運用のために）必要とするストレージ領域に関わる要件の内容を明らかにすること

ストレージ基盤構成における検討事項: 永続ストレージ (その1)

■ ICPのGUIコンソールから構成出来る永続ストレージの種別は以下の4種である

– NFS / Gluster Fs / vSphereボリューム※ / hostPath

※ vSphere Volumeの表記について、本章内では「vSphereボリューム」という表記に統一する

■ 加えて、ICPの「Create resource」画面にYAMLファイルを投入することによりKubernetesからの利用がサポートされる任意の永続ストレージを構成出来る

– Kubernetes v1.12から利用出来る永続ストレージ種別は以下の通り

- GCEPersistentDisk / AWSElasticBlockStore / AzureFile / AzureDisk / CSI / FC (Fibre Channel) / FlexVolume / Flocker / NFS / iSCSI / RBD (Ceph Block Device) / CephFS / Cinder (OpenStack block storage) / Glusterfs / vSphereボリューム / Quobyte Volumes / HostPath / Portworx Volumes / ScaleIO Volumes / StorageOS / Local / photonPersistentDisk
- 参照: [IBM Knowledge Center - PersistentVolume の作成](#)
- 参照: [IBM Knowledge Center - IBM Cloud Private のストレージ・オプション](#)
- 参照: [PersistentVolumeSpec - Kubernetes API Reference Docs](#)

ストレージ基盤構成における検討事項: 永続ストレージ (その2)

■ 永続ストレージへの同時アクセスの可否は永続ストレージ種別に依存する

– 以下、参照先情報の「Access Modes」の項を確認のこと

• 参照情報: [Persistent Volumes – Kubernetes](#)

– 利用したいストレージのサポート条件や制限を適宜確認すること

• 参照: [IBM Knowledge Center - ストレージ・ガイド](#)

• 参照: [IBM Knowledge Center - IBM Cloud Private のストレージ・オプション](#)

ストレージ基盤構成における検討事項: 永続ストレージ (その3)

- 「hostPath」 永続ストレージ種別は単一ノード(single-node cluster)構成のみで認められる。また、PersistentVolumeとしてのhostPathは、テスト環境としてのみ利用可能。

- ・参照: [IBM Knowledge Center - hostPath PersistentVolume の作成](#)
- ・参照: [PersistentVolumeSpec - Kubernetes API Reference Docs](#)

- 「vSphereボリューム」 永続ストレージ種別を用いる場合には、「vSphere クラウド・プロバイダー」利用のための前提条件を満たす必要が生じる

- ホスト名、およびVM名の命名にあたっての比較的厳しい制約が課されているため要注意
- 「vSphere クラウド・プロバイダー」を構成することはICPセットアップ時に明示的に指定すること

- ・参照: [IBM Knowledge Center - vSphere クラウド・プロバイダー](#)
- ・参照: [IBM knowledge Center - 前提条件および制限](#)

ストレージ基盤構成における検討事項: 永続ストレージ (その4)

- 「GlusterFS」 永続ストレージをWorker Node上に構成する際の前提条件、および手順が示されている
 - ・ 参照: [IBM Knowledge Center - GlusterFS PersistentVolume の作成](#)

- 「NFS」 永続ストレージを構成する際の前提条件、および手順が示されている
 - ・ 参照: [IBM Knowledge Ceneter - NFS PersistentVolume の作成](#)

ストレージ基盤構成における検討事項: ファイルシステム容量確保

- 各ノードのファイルシステム向けにストレージ容量を確保しなければならない
 - 容量が肥大化し易いファイルシステムのマウントポイントを分離することも検討すべき
 - 2019年6月時点でのファイルシステム容量の最小要件は以下の通り（ICP製品のバージョンアップに伴い要件は頻繁に変更されるため、常に最新の要件を確認すること）
 - 参照: [IBM Knowledge Center - ハードウェア要件および推奨](#)

Table 4 Minimum storage requirements for runtime

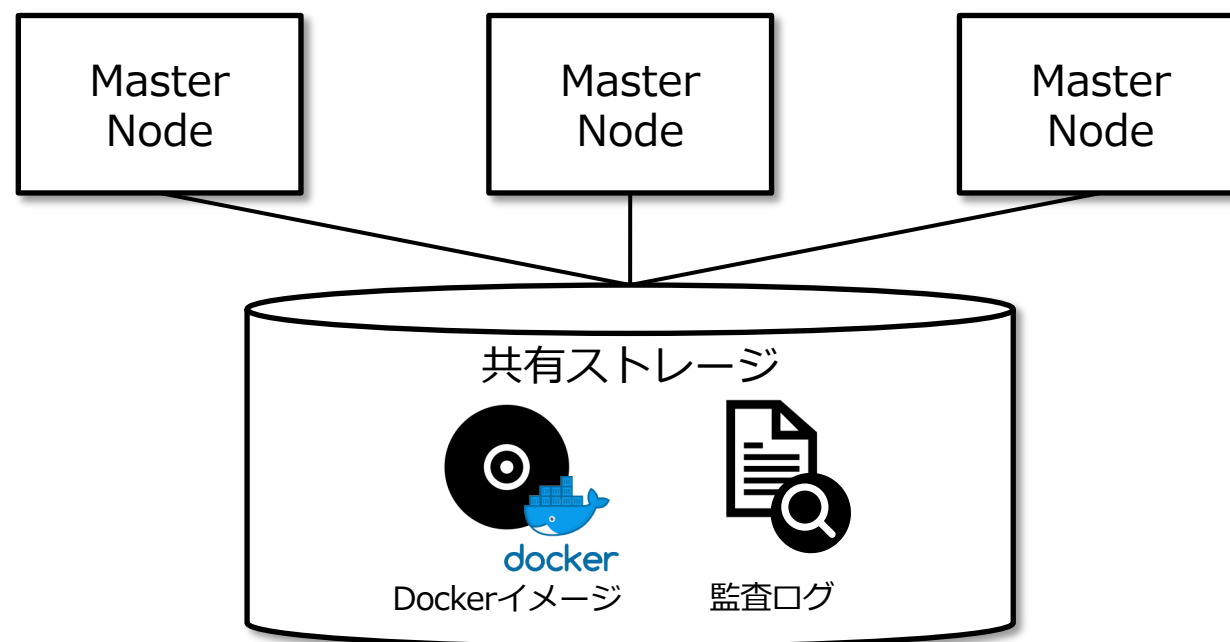
Location	Minimum disk space	Ideal disk space	Node
-/	300 GB	>=800 GB	Master and management
	200 GB	>=600 GB	Worker, proxy, and etcd
	300 GB	>=1000 GB	VA
---/tmp/	50 GB	>=50 GB	All nodes
---/var/	200 GB	>=700 GB	Master and management
	150 GB	>=550 GB	Worker, proxy, and etcd
	200 GB	>=900 GB	VA
---/lib/docker	100 GB	>=400 GB	All nodes
---/lib/etcd	10 GB	>=20 GB	Master or etcd
---/lib/etcd-wal	2 GB	>=4 GB	Master or etcd
---/lib/icp	50 GB	>=150 GB	Master and management
---/lib/icp/va	100 GB	>=350 GB	VA
---/lib/kubelet	30 GB	>=150 GB	All nodes

ストレージ基盤構成における検討事項: ICP HA構成への対応

- HA構成の元でICPをセットアップする場合には、MasterノードのOSファイルシステムでの特定パスを共有ストレージにマウントしなければならない

–対象のパスは以下の通り(ファイルのパーミッションは0755に設定)

- /var/lib/registry – ICP k8s環境のprivate image registry (Dockerイメージを保管する) の配置先
- 参照: [IBM Knowledge Center - IBM Cloud Private Cloud Native、Enterprise、および Community Edition のインストール](#)
- 参照: [IBM Knowledge Center - 高可用性 IBM® Cloud Private クラスター](#)



第2章 第4節: ハイパーバイザー層技術検討

ハイパーバイザーの検討事項: 検討のゴール

■ ハイパーバイザー層についての検討を通じて目指すべきゴールは以下の通り

1. ハイパーバイザー層の利用の有無を決定すること
2. 使用するハイパーバイザー製品を選択すること

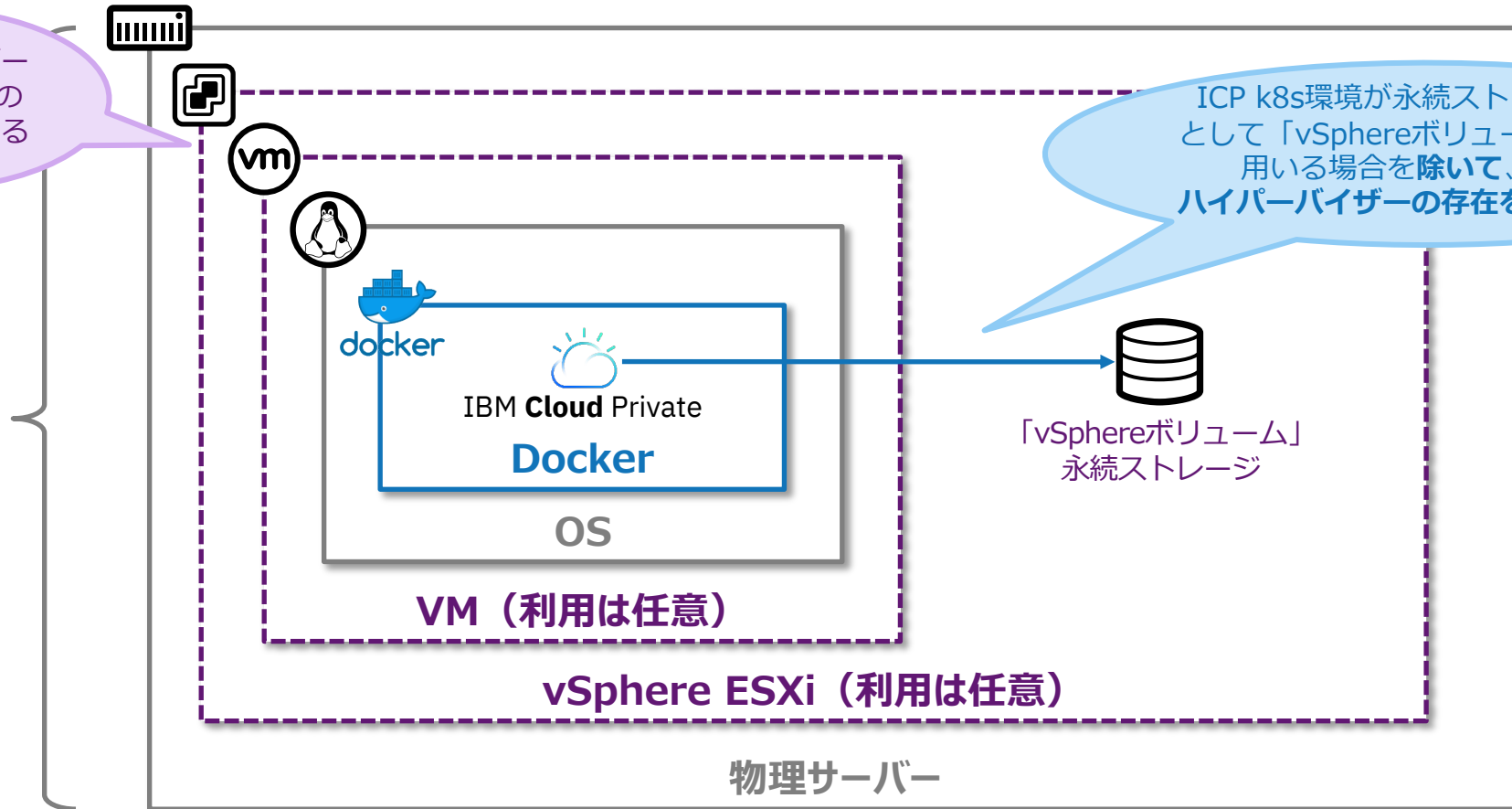
ハイパーバイザーの検討事項: ハイパーバイザーの利用有無（その1）

■ ICP k8s環境は物理マシン/仮想マシンのいずれの上にも実装出来る

- ハイパーバイザーを用いても、用いなくても、どちらでも良い
- 永続ストレージ種別として「vSphereボリューム」を用いる場合を除けば、ハイパーバイザーはICP k8s環境から透過的である

ハイパーバイザー
(およびVM) の
利用は任意である

ICP k8s環境の
ソリューション・スタック



ICP k8s環境が永続ストレージ領域
として「vSphereボリューム」種別を
用いる場合を除いて、ICPは
ハイパーバイザーの存在を意識しない

ハイパーバイザーの検討事項: ハイパーバイザーの利用有無（その2）

■ ICP k8s環境でハイパーバイザーを利用する利点/欠点を下表にまとめる

ハイパーバイザーを利用する利点	ハイパーバイザーを利用する欠点	利点/欠点に対する解説
ノード構成の柔軟性が高い		ハイパーバイザーが提供する仮想マシンを用いることにより、物理マシンを用いた場合と比較してノード構成（例：ノードスペック、ノード台数）の柔軟性が増す。
永続ストレージ種別として、「vSphereボリューム」を利用出来る		ハイパーバイザーとしてVMware vSphereを用いる場合には、永続ストレージ種別として、「vSphereボリューム」（その実体はVMware vSphereの仮想ディスク）を用いることが出来る。
ICP k8s環境の可用性が高まる		ハイパーバイザーが提供する可用性確保のための機能（例：vSphere vMotion、vSphere HA…etc.）を用いることにより、ハイパーバイザー上で稼動するICP k8s環境の可用性が向上する。
ICP k8s環境の可搬性が高まる		ハイパーバイザーを用いることによりICP k8s環境と物理サーバーが互いに疎結合となるため、ICP k8s環境の可搬性が高まる。
システム・バックアップが取得しやすい		ハイパーバイザーを用いることにより、物理マシンを用いた場合と比較してシステム・バックアップの取得が容易となる。
	ハイパーバイザーを運用するために必要なソフトウェア・ライセンス費用、および、サービス費用が増加する	ハイパーバイザーを運用するために必要なソフトウェア・ライセンス費用、および、サービス費用を要することが、ICP k8s環境の金銭的なコストを押し上げる。
	ハイパーバイザー層での製品不具合や設定不備がシステムの安定性に影響を与える。	ハイパーバイザー層での製品不具合や設定の不備に起因して、ICP k8s環境の安定性が損なわれる可能性がある。

ハイパーバイザーの検討事項: ハイパーバイザーのサポート

- IBMは、ICP製品の稼動環境として、いくつかのハイパーバイザー製品（もしくはIaaSプロバイダー）を「認定」している
 - 2019年6月時点での「認定」済みの製品を以下に示す（ICP製品のバージョンアップに伴い要件は頻繁に変更されるため、常に最新の状況を確認すること）
 - 参照: [IBM Knowledge Center - サポートされる IaaS、ハイパーバイザー、および環境](#)

IBM Cloud Private is certified on several Infrastructure as a Service (IaaS) providers.

Table 1. Supported IaaS versions

IaaS	Versions
Azure	
VMware vCenter	6.0, 6.5, 6.7
OpenStack	Mitaka and later
Hyper-V	10.0.15063.0 and later
Nutanix AHV(x86)	20170830.94 (AOS 5.5.2)
Nutanix AHV(IBM PowerPC)	20170331.78 (AOS 5.2.1.1, 5.10.0.7) and later
IBM PowerVC	1.4.1 and later

- ただし、それ以外のハイパーバイザーが非サポートであることを表明するものではない
 - 「contact your sales representative」として、営業担当者への照会を指示している

第2章 第5節: OS層技術検討

OS構成における検討事項: 検討のゴール

■ OS層についての検討を通じて目指すべきゴールは以下の通り

1. 使用するOS（Linuxディストリビューション）を選定すること
2. ICP製品が定める、OS構成に関わる各種要件を把握すること

OS構成における検討事項: サポートOS

- ICP k8s環境の各ノードではICP製品がサポートするOSを用いるべきである
 - 2019年6月時点でサポートされるOSを下表に示す（ICP製品のバージョンアップに伴い要件は頻繁に変更されるため、常に最新の状況を確認すること）
 - 参照: [IBM Knowledge Center - サポートされるオペレーティング・システムとプラットフォーム](#)

Table 1. Supported operating systems

Platform	Operating system
Linux x86_64	Red Hat Enterprise Linux (RHEL) 7.4, 7.5, and 7.6
	Ubuntu 18.04 LTS and 16.04 LTS
	SUSE Linux Enterprise Server (SLES) 12 SP3
Linux on Power (ppc64le) POWER8® or later	Red Hat Enterprise Linux 7.4, 7.5*, 7.5-Alt, and 7.6
	Ubuntu 18.04 LTS (KVM and Bare Metal only) and Ubuntu 16.04 LTS
	SUSE Linux Enterprise Server 12 SP3
Linux on IBM® Z	Red Hat Enterprise Linux 7.4, 7.5, and 7.6
	Ubuntu 18.04 LTS and 16.04 LTS
	SUSE Linux Enterprise Server 12 SP3

* RHEL 7.5 users on POWER9™ must be sure to install the latest kernel-alt kernel that is provided by RHEL.

- VAノードではSLESをサポートしていないため注意が必要。
 - 参照: [IBM Knowledge Center - サポートされるオペレーティング・システムとプラットフォーム](#)

OS構成における検討事項: ネットワーク・ポート構成

- ICP k8s環境がノード上で用いるネットワーク・ポートと、ノードに対応するLinux OS、およびOS上のソフトウェアが用いるネットワーク・ポートとの競合は避けなければならない
 - ICP k8s環境が用いる既定のネットワーク・ポート番号は以下から参照のこと
 - 参照: [IBM Knowledge Center - 必要なポート](#)
 - ICP k8s環境が用いるネットワーク・ポート番号のうち、一部ネットワーク・ポート番号は、既定値から変更することが出来る（以下のCustomizing the cluster with the config.yaml fileの内容を参照のこと）
 - 参照: [IBM Knowledge Center - config.yaml ファイルを使用したクラスターのカスタマイズ](#)
 - OS FirewallではICP k8s環境が用いるネットワーク・ポートを開放しなければならない

OS構成における検討事項: ファイアーウォール構成

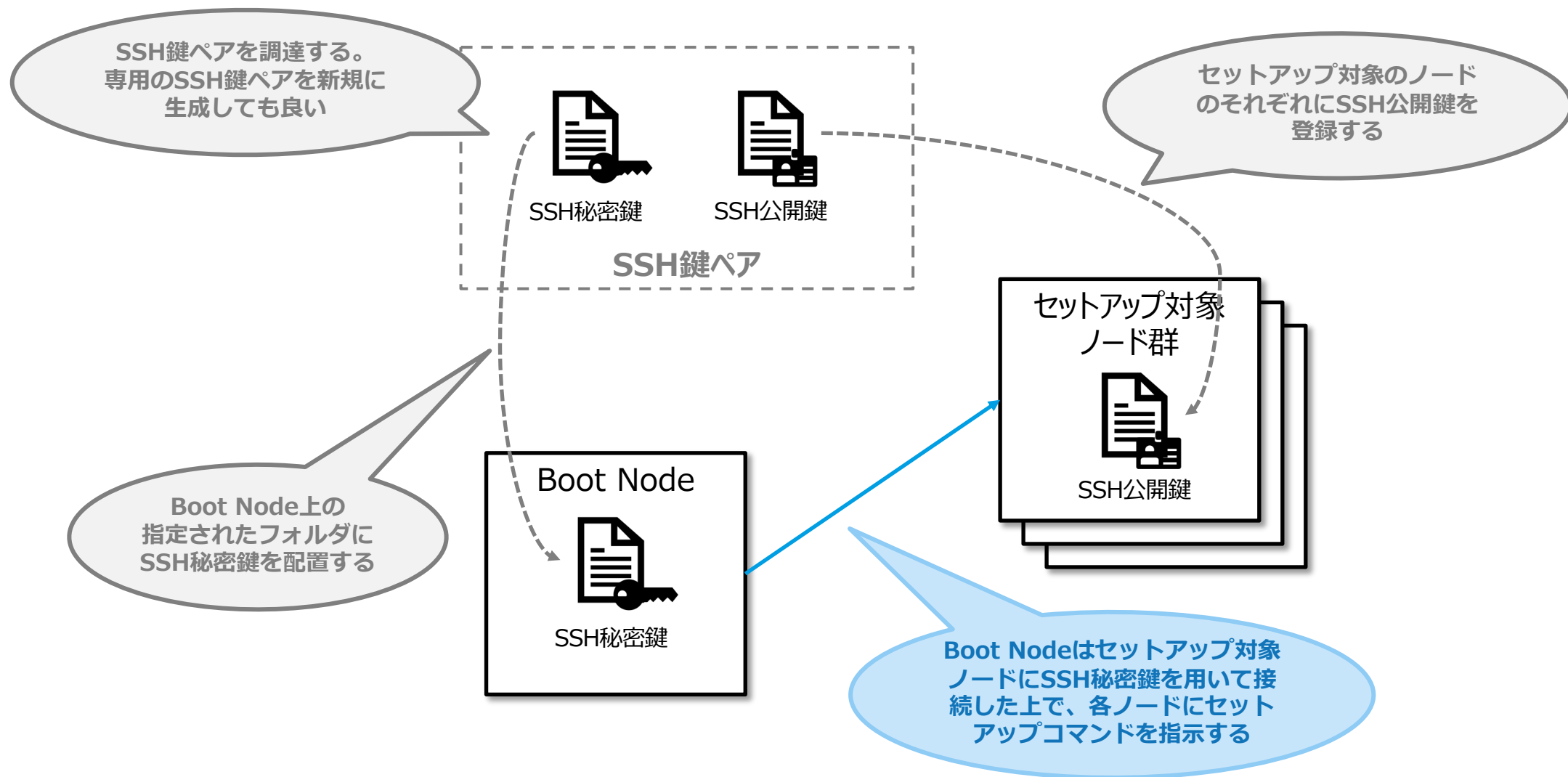
- ICP k8s環境をインストールするにあたり、ファイアーウォール設定について以下の点を考慮する必要がある
 - config.yaml上で「firewall_enabled: true」に設定すると、ローカルOSのfirewallが起動している状態を保持し、ICPインストールに必要なfirewall上のポートを解放する
 - 参照: [IBM Knowledge Center - config.yaml ファイルを使用したクラスターのカスタマイズ](#)
 - Master あるいは proxyノードに SLESを用いている場合、インストール中はクラスタ内の全てのFirewallを無効にしなければならない
 - 参照: [IBM Knowledge Center - IBM Cloud Private Cloud Native、Enterprise、および Community Edition のインストール](#)

OS構成における検討事項: BootノードからのSSH接続 (その1)

- インストール時にBootノードから他ノードを制御するために、Bootノードから他ノードへのSSH接続を行うことが必要となる
- Boot NodeからのSSH接続構成は以下の2通りのいずれかにより実装される
 - SSH鍵ペアを用いた認証
 - パスワードを用いた認証
- SSH鍵ペアを用いたSSH接続のための構成方法は以下にガイドされている
 - 参照: [IBM Knowledge Center - クラスター・ノードでの SSH 鍵の共有](#)
- パスワードを用いたSSH接続のための構成方法は以下にガイドされている
 - 参照: [IBM Knowledge Center - クラスター・ノードのパスワード認証の構成](#)

OS構成における検討事項: BootノードからのSSH接続 (その2)

■ SSH鍵ペアを用いた認証構成の概念図を以下に示す



OS構成における検討事項: 名前解決構成

■ ICP k8s環境の各ノードは互いに名前解決を行う必要がある

- /etc/hostsファイルを編集することにより名前解決を実装することが指示されている
- /etc/hostsファイルを編集する際は、以下の点に注意する必要がある
 - ローカル・ループバック・アドレスにホスト名を対応させるエントリは認められていない
 - /etc/hosts ファイル内の**ホスト名に大文字を使用することはできない**
 - クラスタに含まれるノードが1つの場合、そのノードのIPアドレスおよびホスト名をリストする必要がある
 - **127.0.1.1 および ::1 localhost で始まる行をコメントアウトする必要がある**
- 参照: [IBM Knowledge Center - ノードの準備](#)

```
>cat /etc/hosts
127.0.0.1    localhost
#127.0.1.1   nodehostname ← OSがこのようなエントリを既定で設けていた場合にはコメントアウトしなければならない
# The following lines are desirable for IPv6 capable hosts
#::1 localhost ip6-localhost ip6-loopback ← OSがこのようなエントリを既定で設けていた場合にはコメントアウトしなければならない
...
```

OS構成における検討事項: その他のセットアップ前提

■ 各ノードでのLinux OS構成に関わる各種の前提が指定されている

- 参照: [IBM Knowledge Center - ノードの準備](#)

- それらの各種前提の内容は以下に示す通り（2019年6月時点）

– クラスターの準備

- System requirementを満たしていること
- クラスターのアーキテクチャを決定し、各ノードにIPアドレスが割り当てられていること
- 環境内に異なるネットワークデバイス名を持つノードが含まれることがある。Calicoを使用する場合は、ネットワーク設定で必ず[interface] パラメーターを設定すること

– 各ノードの準備

- ノード上の全てのディレクトリが空であり、インストールに十分な容量が確保されていること
- 各クラスターノードで「service_cluster_ip_range」パラメーターにデフォルトゲートウェイあるいはルートが設定されていること
- 各ノードが、クラスター内の他の全てのノードに接続されていること
- ICP k8s環境の全てのノードにサポートされているバージョンのPythonが導入されていること
- ICP k8s環境の各ノードは互いに時刻同期していること
- 各ノードにSSHクライアントが導入されていること

第2章 第6節: Dockerエンジン構成技術検討

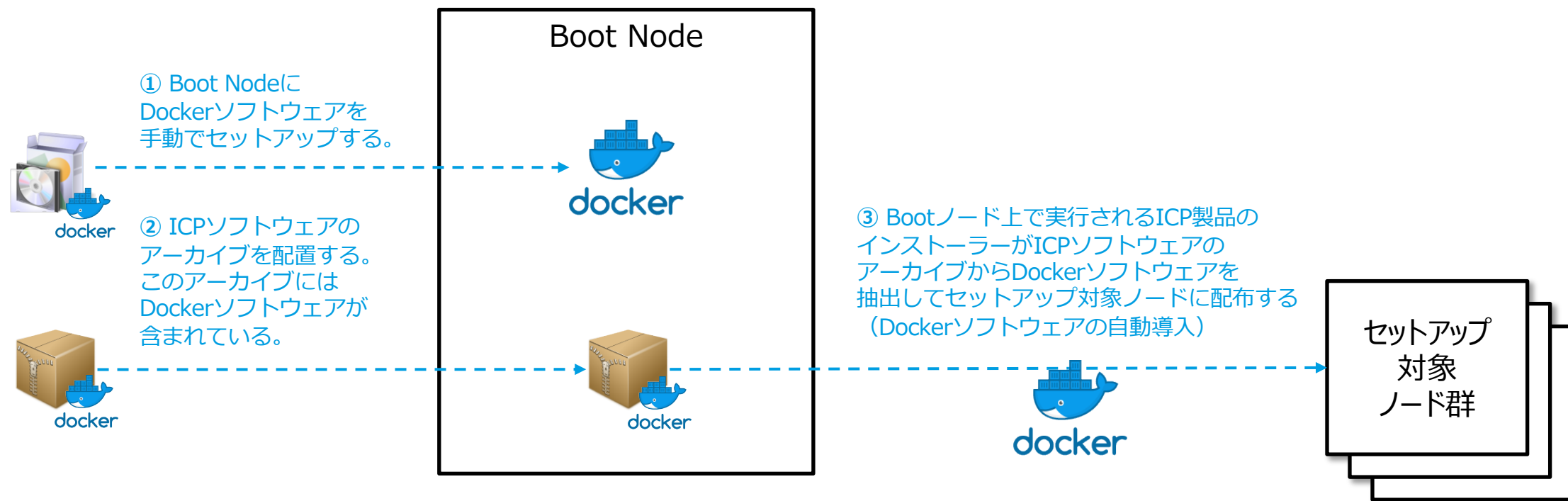
Dockerエンジン構成における検討事項: 検討のゴール

■ Dockerエンジンについての検討を通じて目指すべきゴールは以下の通り

1. Dockerエンジンの各ノードへの展開方式を明らかにすること
2. Dockerエンジンのバージョンを選定すること
3. Dockerエンジンを各ノードに自動展開する場合には、Docker関連の各種設定を各ノードに反映するための方式を理解すること
4. Dockerのストレージ・ドライバーを選定すること

Dockerエンジン構成における検討事項: Dockerの自動導入

- ICP k8s環境のセットアップ時に、各ノード（Boot Nodeを除く）にDockerソフトウェアを自動導入することが出来る（自動導入が必須という訳ではない）
 - Boot NodeではDockerソフトウェアを手動で導入しなければならない



- OSとしてSLESを利用するノードでは、手動でDockerをインストールする必要がある
 - 参照: [IBM Knowledge Center - Docker 自動インストールのためのクラスター・ノードの構成](#)

Dockerエンジン構成における検討事項: Dockerバージョン

■ ICP k8s環境がサポートするDockerのバージョンが定められている

- ICPパッケージには、あらかじめDockerエンジンが同梱されている。同梱されている以外のDockerエンジンを用いる場合は、ICPがサポートするDockerのバージョンを確認する必要がある (ICP製品のバージョンアップに伴い要件は頻繁に変更されるため、常に最新の状況を確認すること)

• 参照: [IBM Knowledge Center - サポートされる Docker バージョン](#)

プラットフォーム	オペレーティング・システム	パッケージ名	バイナリー名
Linux®	Red Hat Enterprise Linux (RHEL)、Ubuntu、SUSE Linux Enterprise Server (SLES)	IBM Cloud Private 3.1.2 Docker for Linux®	icp-docker-18.03.1_x86_64.bin
Linux® on Power® (ppc64le)	Red Hat Enterprise Linux (RHEL)、Ubuntu、SUSE Linux Enterprise Server (SLES)	IBM Cloud Private 3.1.2 Docker for Linux® on Power® (ppc64le)	icp-docker-18.03.1_ppc64le.bin
IBM® Z	Red Hat Enterprise Linux (RHEL)、Ubuntu、SUSE Linux Enterprise Server (SLES)	IBM Cloud Private 3.1.2 Docker for IBM® Z	icp-docker-18.03.1_s390x.bin

注: SLES オペレーティング・システムの場合のみ、icp-docker-18.03.1 bin ファイルは Docker のバージョン 17.09.1 をデプロイします。

Table 1. Versions of Docker that are supported by IBM Cloud Private

Platform	Operating system	Docker CE	Docker EE
Linux® x86_64	Red Hat Enterprise Linux (RHEL) 7.4, 7.5, and 7.6	See notes	17.03 to 18.03.1, 18.06.2
	Ubuntu 18.04 LTS	18.03.1, 18.06.2	18.03.1, 18.06.2
	Ubuntu 16.04 LTS	1.12 to 18.03.1, 18.06.2	17.03 to 18.03.1, 18.06.2
	SUSE Linux Enterprise Server (SLES) 12 SP3	See notes	See notes
Linux on POWER 64-bit Little Endian (LE)	Red Hat Enterprise Linux 7.4, 7.5, and 7.6	See notes	17.03 to 18.03.1, 18.06.2
	Ubuntu 18.04 LTS	18.03.1, 18.06.2	18.03.1, 18.06.2
	Ubuntu 16.04 LTS	1.12 to 18.03.1, 18.06.2	17.03 to 18.03.1, 18.06.2
	SUSE Linux Enterprise Server 12 SP3	See notes	See notes
Linux on IBM® Z	Red Hat Enterprise Linux 7.4, 7.5, and 7.6	See notes	17.03 to 18.03.1, 18.06.2
	Ubuntu 18.04 LTS	18.03.1, 18.06.2	18.03.1, 18.06.2
	Ubuntu 16.04 LTS	1.12 to 18.03.1, 18.06.2	17.03 to 18.03.1, 18.06.2
	SUSE Linux Enterprise Server 12 SP3	See notes	See notes

Dockerエンジン構成における検討事項: Dockerエンジンパラメーター

- 各ノード（Boot Nodeを除く）にDockerソフトウェアを自動導入する際には、Dockerエンジンのパラメーターを指定することが出来る
 - 「docker_env」設定値、および「docker_version」「install_docker」設定値を用いる
 - 参照: [IBM Knowledge Center - config.yaml ファイルを使用したクラスターのカスタマイズ](#)
 - Dockerソフトウェアの自動導入時にバージョンを指定する場合は、「config.yaml」内の「docker_version」パラメーターを用いる。（指定するバージョンのパッケージがbootノードの /<installation_directory>/cluster/runtime-engineディレクトリー内に配置されている必要がある）
 - 参照: [IBM Knowledge Center - config.yaml ファイルを使用したクラスターのカスタマイズ](#)
- ICP製品のインストール・ガイドで、特定のパラメーター値を設定に反映することが指示されている場合がある
 - 「docker.service」ファイルの「MountFlags」パラメーターの値を「share」とすること、もしくは、「MountFlags」パラメーターのエントリ自体を削除すること
 - 参照: [IBM Knowledge Center - IBM Cloud Private の Docker のセットアップ](#)

Dockerエンジン構成における検討事項: ストレージ・ドライバー

■ 必要に応じて、Dockerのストレージ・ドライバーの選定を検討する

- 参照: [IBM Knowledge Center - IBM Cloud Private の Docker のセットアップ](#)
- 参照: [Select a storage driver - docker docs](#)

– 推奨されるストレージ・ドライバーの種類は、OSの種類やバージョン、Dockerのバージョンにより異なるため、最新の情報を確認する必要がある

- 参照: [Compatibility Matrix – docker success center](#)

– 例えば、OSとしてRHELを用いる場合、ストレージ・ドライバーとしてoverlay2が推奨される
また、devicemapperを本番環境として利用する場合、direct-lvmモードを利用しなければならないなどの制限がある

- 参照: [Get Docker EE for Red Hat Enterprise Linux – docker docs](#)

第2章 第7節: ICPクラスター構成技術検討

ICPクラスター構成における検討事項: 検討のゴール

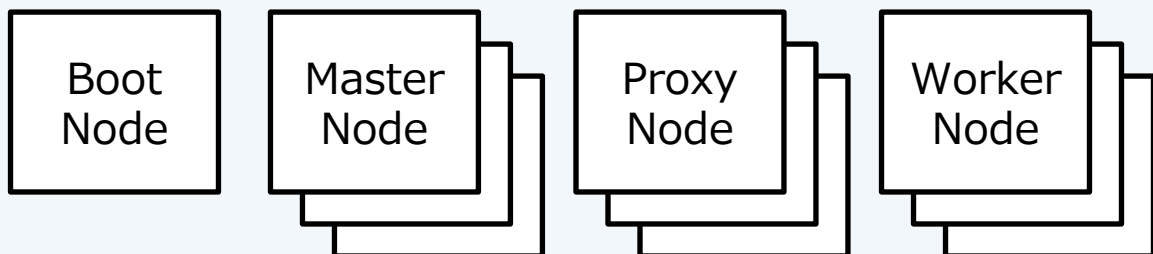
■ ICPクラスター構成についての検討を通じて目指すべきゴールは以下の通り

1. ICPクラスターという概念を理解すること
2. ICPクラスターに対して適用出来る設定項目の全体像を把握すること

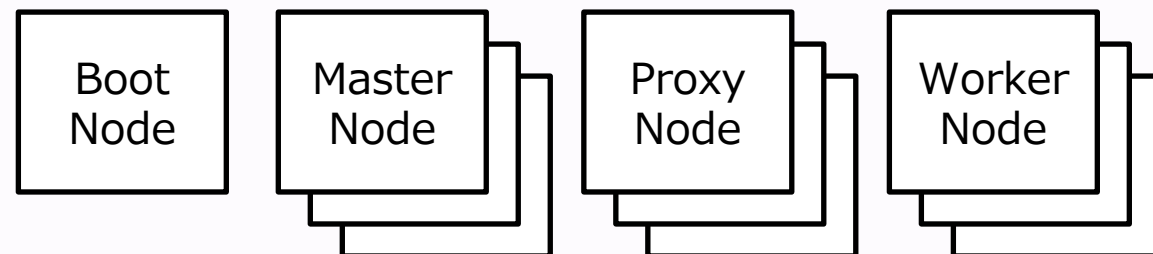
ICPクラスタ構成における検討事項: ICPクラスタという概念

- ICP k8s環境での「cluster（クラスター）」という用語は、ICPの実装に用いられているKubernetes製品での「cluster（クラスター）」という概念に対応している
 - 「cluster（クラスター）」という概念は「ICP k8s環境の自立したフルセット」と捉えられる
 - 何らかの要件（例えば、可用性要件）に基づき、ICP k8s環境の複数のクラスターを運用することも設計上の一つの選択肢である
 - ICP k8s環境の動作を制御するための設定の多くが、ICPクラスタを対象に行われる

ICP Kubernetesクラスタ A



ICP Kubernetesクラスタ B



要件に応じて、複数のクラスターを並行稼動させるケースも生じうる

ICPクラスター構成における検討事項: クラスターへの設定（その1）

- ICPクラスター設定値の指定はICPクラスターのセットアップ時に参照されるパラメーターファイルである「config.yaml」ファイルへの記述を通じて行う

– 「config.yaml」ファイルに記述する設定内容の分類を下表に示す

・参照: [IBM Knowledge Center - config.yaml ファイルを使用したクラスターのカスタマイズ](#)

設定項目の分類	解説	分類に含まれる設定項目例
汎用設定	他の分類に属さないような雑多な設定項目が属する。	offline_pkg_copy_path: （設定内容）オフラインインストール時に一時的にインストールファイルを保持しておくディレクトリを指定することができる。
Kubernetes 設定	ICP製品に組み込まれているKubernetes製品の動作を制御するための設定項目が属する。Kubernetes製品を構成する各サービスに引き渡す引数を定義する形で設定を行う。利用可能な引数は Kubernetes製品のガイド から確認する必要がある	kubelet_extra_args: （設定内容）kubeletサービスに引き渡す引数を指定する。例えば、ノード上で稼働させるPod数の上限を指定することができる。
ログ設定	ログ管理運用に関する設定項目が属する。Kibanaコンソールの導入を指定する設定項目もこの分類に属する。	kibana_install: （設定内容）Kibanaの導入有無を指定する。v3.1.0よりデフォルト値が「true」に変更。
ネットワーク設定	ネットワーク機能の動作に関わる設定項目が属する。それら項目の内容は本資料の「ネットワーク基盤構成における検討事項」の項で解説した通り。	network_cidr: （設定内容）ICP k8s環境上で稼働するコンテナに割り当てるIPアドレスのレンジを指定する。
クラスター・アクセス設定	ICPクラスターへのアクセスに関する設定項目が属する。	cluster_name: （設定内容）ICPクラスターの名前を指定する。

ICPクラスター構成における検討事項: クラスターへの設定 (その2)

– 「cluster.yaml」 ファイルに記述する設定内容の分類を下表に示す (続き)

・参照: [IBM Knowledge Center - config.yaml ファイルを使用したクラスターのカスタマイズ](#)

設定項目の分類	解説	分類に含まれる設定項目例
Docker 設定	Dockerエンジンの動作に関わる設定項目が属する。これら設定項目は、Dockerエンジンをクラスターに自動展開する際に反映される。	docker_env: (設定内容) Dockerエンジンが用いる環境変数を指定する。
プロキシ HA 設定 マスター HA 設定	ICPのHA構成時におけるクラスターIPアドレスの指定に関わる設定項目が属する。	proxy_vip: (設定内容) Proxy NodeのクラスターIPアドレスを指定する。
ユーザー設定	ICPクラスタインストール時、およびインストール後のユーザーに関わる設定項目が属する。	default_admin_password: (設定内容) ICPクラスターにアクセスする際のデフォルト管理者パスワードを指定する。
GlusterFS 設定	GlusterFSを利用するための設定項目が属する。具体的な設定内容の詳細が IBM Knowledge Centerでガイドされている 。	glusterfs: (設定内容) GlusterFSの利用有無を指定する。
クラウド・プロバイダー設定	vSphere クラウド・プロバイダーを利用するための設定項目が属する。クラウドプロバイダとして、“vSphere” “AWS” “Azure”について、それぞれガイドされている。	cloud_provider: (設定内容) cloud providerの種別を指定する。
IPsecを使用したクラスター・データのネットワーク・トラフィックの暗号化	ノード間通信のためにIPsec暗号化を利用するための設定項目が属する。項目の内容は本資料の「ネットワーク基盤構成における検討事項」の項で解説した通り。	(IBM Knowledge Center記述 を参照のこと)
管理サービス設定	Logging, Monitoring などマネジメントサービスの有効無効に関わる設定項目が属する。	(IBM Knowledge Center記述 を参照のこと)
VMware NSX-T 2.3 と IBM Cloud Private の統合	ICP上でVMware NSX-Tによるネットワーク仮想化の機能を利用する上での設定項目が属する。	(IBM Knowledge Center記述 を参照のこと)

ICPクラスター構成における検討事項: クラスターへの設定（その3）

- 「cluster.yaml」 ファイルに記述する設定内容の分類を下表に示す（続き）
 - 参照: [IBM Knowledge Center - config.yaml ファイルを使用したクラスターのカスタマイズ](#)

設定項目の分類	解説	分類に含まれる設定項目例
環境分離	ICPクラスターの環境分離に関わる設定項目が属する。	isolated_namespaces: （設定内容） namespaceおよびそのnamespace専用のworkerノードが属するホストグループを指定する。
etcd 設定	etcdノード役割に関わる設定項目が属する。	etcd_extra_args: （設定内容） etcdの追加構成情報を指定する。
Istio アドオン・セキュリティ設定	Istioアドオンに関わる設定項目が属する。	istio_addon: （設定内容） istio add-onのセキュリティ設定を行う。
NGINX ingress コントローラー	Nginx ingress controllerに関わる設定項目が属する。NGINXのカスタマイズに利用可能なパラメーターについては、 Github上のConfiguration options から確認する必要がある。	disable-access-log: （設定内容） アクセスログの有効無効設定ができる。
Minio 設定	minioストレージに関わる設定項目が属する。	（ IBM Knowledge Center記述 を参照のこと）

ICPクラスター構成における検討事項: クラスターへの設定（その4）

■ 優先的な検討の対象とすべきICPクラスター設定項目を下表に列挙する

・参照: [IBM Knowledge Center - config.yaml ファイルを使用したクラスターのカスタマイズ](#)

設定項目の分類	ICPクラスター設定項目	既定値	設定項目の趣旨	優先的な検討を必要とする背景
汎用設定	firewall_enabled	false	ICPクラスターのセットアップ対象とする環境でのファイアウォールの動作有無を指定する。	ファイアウォールが存在するネットワーク環境では既定値からの変更が必要。
ネットワーク設定	calico_ipip_mode	Always	ノード間通信のモードを指定する。「Always」あるいは「CrossSubnet」とした場合にはノード間通信がIP over IP modeとなる。	以下に示す状況においては、設定値は「Always」あるいは「CrossSubnet」としてIP over IP modeを有効化しなければならない。 - ノードが複数サブネットに分散しており、かつBGPが利用出来ない場合 - ノードがL3スイッチとして稼動することが許されない場合（例:OpenStack環境） 一方で、IP over IP modeを利用した場合、IP over IP処理の負荷が高いことに起因してネットワーク・スループットが落ちる可能性があるため、トレードオフが生じる。
	calico_tunnel_mtu	1430	CalicoでIP over IP modeを用いる場合のMTUサイズを定義する。	ICP k8s環境のノード間通信を運用するネットワークでのMTUサイズを確認した上で、当該のMTUサイズを20以上下回る値を設定すべき。
	network_cidr	10.1.0.0/16	ICP k8s環境の各コンテナが利用するIPアドレスのレンジを指定する。	ここで指定するアドレス・レンジが、組織内で使用しているプライベート・ネットワークのアドレス・レンジと競合することは、ネットワーク・ルーティング制御の観点より好ましくない。

ICPクラスター構成における検討事項: クラスターへの設定 (その5)

■ 優先的な検討の対象とすべきICPクラスター設定項目を下表に列挙する (続き)

・参照: [IBM Knowledge Center - config.yaml ファイルを使用したクラスターのカスタマイズ](#)

設定項目の分類	ICPクラスター設定項目	既定値	設定項目の趣旨	優先的な検討を必要とする背景
ネットワーク設定	calico_ip_autodetection_method	can-reach={{ groups['master'][0] }}	ノード間通信のために用いるNICの選択方式を指定する。	ノードが複数NIC構成を採る場合には、ノード間通信に用いるネットワークを明示的に指定すべきである。
	service_cluster_ip_range	10.0.0.0/16	ICP k8s環境上で稼動するサービスに付与する仮想IPアドレスのレンジを指定する。	ここで指定するアドレス・レンジが、組織内で使用しているプライベート・ネットワークのアドレス・レンジと競合することは、ネットワーク・ルーティング制御の観点より好ましくない。
クラスター・アクセス設定	cluster_name	mycluster	複数クラスター環境において、各クラスター名を指定する。	この情報は管理および認証の際に利用するため、ICPインストール後に変更してはならない項目である。
	cluster_lb_address	(None)	ノードに複数NICが構成されている環境でのICP管理通信に用いるIPアドレスを指定する。	Masterノードが複数NIC構成を採る場合には、ICP管理通信に用いるネットワークを明示的に指定すべきである。
	proxy_lb_address	(None)	ノードに複数NICが構成されている環境での外部からのサービス通信に用いるIPアドレスを指定する。	Proxyノードが複数NIC構成を採る場合には、外部からのサービス通信に用いるネットワークを明示的に指定すべきである。

ICPクラスター構成における検討事項: クラスターへの設定 (その6)

■ 優先的な検討の対象とすべきICPクラスター設定項目を下表に列挙する (続き)

・参照: [IBM Knowledge Center - config.yaml ファイルを使用したクラスターのカスタマイズ](#)

設定項目の分類	ICPクラスター設定項目	既定値	設定項目の趣旨	優先的な検討を必要とする背景
プロキシ HA 設定 および マスター HA 設定	proxy_vip_iface	(None)	Proxy NodeがHA構成を採る場合でのHAクラスターIPアドレスが使用するネットワーク・インターフェースを指定する。	Proxy Nodeが複数NIC構成を採る場合には、外部からのサービス通信に用いるネットワークを明示的に指定すべきである。
	vip_iface	(None)	Master NodeがHA構成を採る場合でのHAクラスターIPアドレスが使用するネットワーク・インターフェースを指定する。	Master Nodeが複数NIC構成を採る場合には、外部からのサービス通信に用いるネットワークを明示的に指定すべきである。

ICPクラスター構成における検討事項: クラスターへの設定 (その7)

■ 優先的な検討の対象とすべきICPクラスター設定項目を下表に列挙する (続き)

• 参照: [IBM Knowledge Center - config.yaml ファイルを使用したクラスターのカスタマイズ](#)

設定項目の分類	ICPクラスター設定項目	既定値	設定項目の趣旨	優先的な検討を必要とする背景
ユーザー設定	default_admin_user	admin	ICPクラスターにアクセスする際のカスタム管理者ユーザー名を指定する。	LDAPが有効化されている、かつ、ユーザー名がadminのLDAPユーザーが存在する場合、デフォルトで設定されているクラスター管理者名と競合するため他の値に変更する必要がある。
	default_admin_password	(None)	ICPクラスターにアクセスする際のデフォルト管理者パスワードを指定する。	ICPインストールの際に必ず行うべき設定である。
	password_rules	- '^[a-zA-Z0-9¥-]{32,}\$'	default_admin_passwordが満たすべきパスワードルールを指定する。	passwordルールに則る適切なパスワードを設定する必要がある。v3.1.2よりパスワードルールが強化され、デフォルトで 32文字以上 という非常に強固なパスワードルールとなっているので注意が必要。

第2章 第8節: 関連システム連携技術検討

関連システム連携構成に関する検討事項: 検討のゴール

- 関連システム連携構成についての検討を通じて目指すべきゴールは以下の通り
 1. ICP k8s環境が連携し得る関連システムの存在を把握すること
 2. ICP k8s環境と関連システムとの間での連携方式を明らかにすること

関係システム連携構成における検討事項: ICP k8s環境の関連システム

■ ICP k8s環境は関連システムと連携することによりその機能を発揮する

– ICP k8s環境と関連し得るシステムを下表に例示する（コンピュート基盤、ネットワーク基盤などは、ICP k8s環境に包含されているとの考えの下で関連システムとしては捉えていない）

関連システム	ICP k8s環境との関係/連携	関連システム連携のための検討負荷
DevOps系システム	ICP k8s環境外部のDevOps系システムが生成したアプリケーションの実行パッケージをICP k8s環境にデプロイして稼動させる。	大きい: (詳細について後続スライドで解説する)
LDAPサービス	ICP k8s環境での認証/認可に関連する運用のために、ICP k8s環境外部のLDAPサービスを利用する。	中程度: (認証・認可の詳細について7.2節で解説する)
監視基盤	ICP k8s環境で確認されたイベントを、ICP k8s環境外部の監視基盤に送信する。	大きい: (詳細について後続スライドで解説する)
名前解決基盤	ICP k8s環境上で稼動するサービスによるICP k8s環境外部への名前解決のための連携を行う。また、ICP k8s環境外部からのICP k8s環境上のサービスへの名前解決のための連携を行う。	中程度: (詳細について後続スライドで解説する)
バックアップ基盤	ICP k8s環境上で稼動する業務システムが永続ストレージ上に保管するデータをバックアップする。	中程度: (バックアップ運用の詳細について4.2節で解説する)
外部Helm Chart Repository、外部Docker Registry	ICP k8s環境上でのアプリケーションの展開に用いるHelm Chartを、ICP k8s環境外部の外部Helm Chartレポジトリから取得する。また、ICP k8s環境上でのアプリケーションの展開に用いるDockerコンテナ・イメージを、ICP k8s環境外部の外部Docker Registryから取得する	小さい: ICP側から参照するHelm Repository、もしくはDocker Registryの所在を指定すれば良い。 (詳細について5.2節で解説する)
NTPサービス	ICP k8s環境のノード間での名前解決を、ICP k8s環境外部のNTPサービスによりサポートする。	小さい: Linux環境での標準的な時刻同期実装を行えば良い。

関係システム連携構成における検討事項: DevOps連携

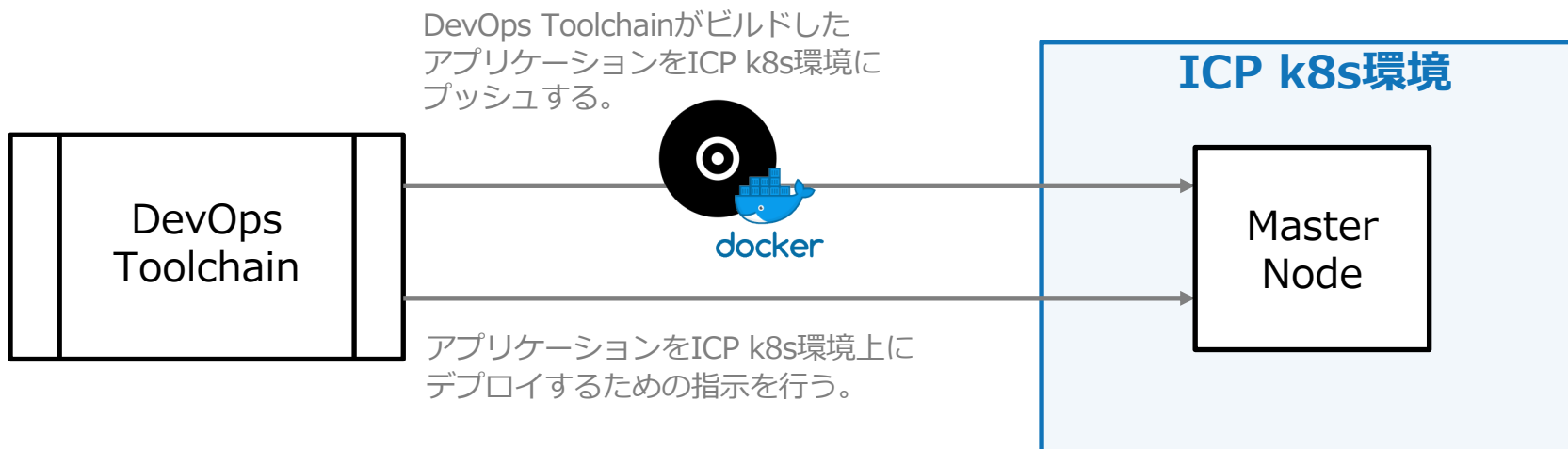
- ICP k8s環境外部のDevOps系システムが生成したアプリケーションの実行パッケージをICP k8s環境にデプロイして稼動させる
 - 連携のイメージ、および、重点的に検討すべきポイントを以下に図示する

アプリケーションのデプロイに関わる検討事項

- DevOps Toolchainの側がKubernetes上でのアプリケーションの展開操作をサポートしているか？（kubectlによる操作 / APIによる操作…etc.）
- DevOps Toolchainの側がKubernetes上でのアプリケーションの展開操作に関連する設定情報や認証情報を適切に保管、管理するための機能を保持しているか？

アプリケーションのプッシュに関わる検討事項

- DevOps Toolchainの側がKubernetes向けアプリケーションパッケージ形式（Dockerイメージ + Helm Chart）に対応しているか？
- DevOps Toolchainの側がプライベート・イメージ・レジストリーへのDockerイメージのPushに対応しているか？

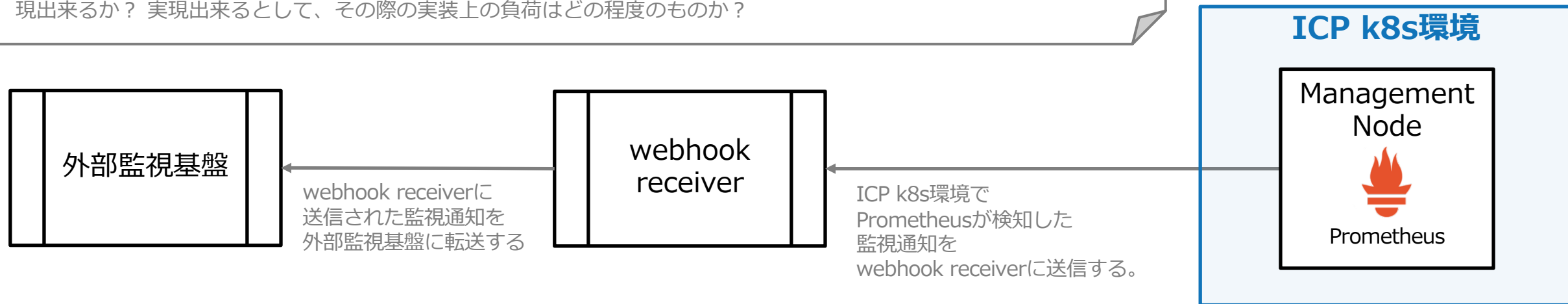


関係システム連携構成における検討事項: 監視基盤連携

- ICP k8s環境で確認されたイベントを、ICP k8s環境外部の監視基盤に送信する
 - 実装のイメージ、および、重点的に検討すべきポイントを以下に図示する

監視基盤連携に関わる検討事項

- ICP k8s環境で外部監視基盤への通知機能を提供するPrometheusサービスからの通知の受信を外部監視基盤がサポートするか？ Prometheusからの通知実装については「[Configuration | Prometheus](#)」の記述を参照のこと。Prometheus側で既定で提供されている通知実装（Email、Slack…etc.）を用いることが出来ないならば、webhookサービスをPrometheusと外部監視基盤との間に置いた上で、同サービスによるメッセージ変換などを行わなければならない（外部監視基盤側の実装に応じた通知手段のカスタマイズはwebhookサービスで吸収する製品思想となっている）。
- 外部監視基盤が求める通知ポリシー（例:同一障害に起因する重複メッセージの抑止）をPrometheusサービスを用いて実現出来るか？ 実現出来るとして、その際の実装上の負荷はどの程度のものか？



関係システム連携構成における検討事項: 名前解決基盤連携

- ICP k8s環境上で稼動するサービスによるICP k8s環境外部の名前解決のための連携を行う。また、ICP k8s環境外部からのICP k8s環境上で稼動するサービスに対する名前解決のための連携を行う

–実装のイメージ、および、重点的に検討すべきポイントを以下に図示する

名前解決基盤連携に関わる検討事項

- ICP k8s環境上で稼動するDockerコンテナ内部のプロセスがICP k8s環境外部に向けて名前解決を行う際の実装はどのようなものになるか？ この実装のための方式を大まかに分類すると以下に示す2通りとなる。
 - ① ICP k8s環境がホストするDNSサービス (kube dns) を介して名前解決を行う方式
 - ② ICP k8s環境の各ノードによるDNS参照を介して名前解決を行う方式。詳細については「[Configure private DNS zones and upstream nameservers in Kubernetes](#)」を参照のこと
- ICP k8s環境上で稼動するサービスを外部DNSサービスに登録する運用をどのように行うか？

ICP k8s環境上で稼動するサービスを
外部DNSサービスに登録する



ICP k8s環境で稼動するサービスが
環境外部のサービスの名前解決を行う

ICP k8s環境



第2章 第9節: 構成上のNGパターン集

ICP k8s環境構成 NGパターン集: はじめに

- 以降のスライドで、ICP k8s環境で構成を行うにあたってのNGパターン（技術的に許容されない/推奨されない構成パターン）を例示する

ICP k8s環境構成 NGパターン集: (その1)

■ ICP k8s環境構成にあたってのNGパターンを以下に例示する

ICP機能をサポートしない プラットフォームの選択

※利用したいICP機能をサポートしない
プラットフォームは使用しないこと

Storage GlusterFS機能は
IBM Z 環境を
サポートしないためNG



Storage GlusterFS機能

×



IBM Zがサポートしない ノード役割の割り当て

※サポートしないノード役割を
IBM Zノードに割り当ててはならない

IBM ZはVAノード役割はサポートしないためNG



VA Node役割

×



ICP k8s環境構成 NGパターン集: (その2)

■ ICP k8s環境構成にあたってのNGパターンを以下に例示する (続き)

HA構成で非推奨とされる ノード冗長度を使用

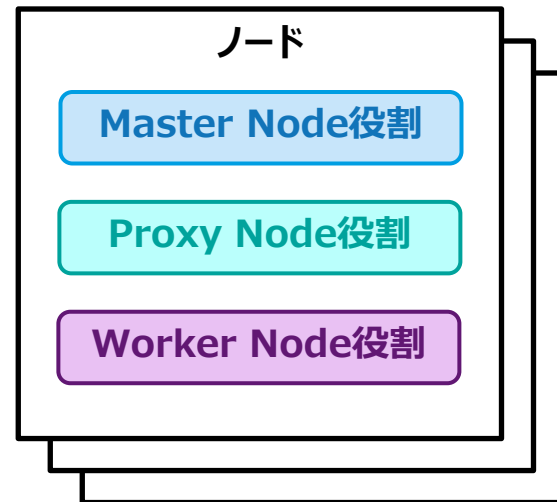
※Master NodeのHA構成では
冗長度=3 or 5以外は非推奨

HA構成での冗長度=4は非推奨であるためNG



Worker Node役割の他役割との 同一ノード配置 (中規模～環境)

※Worker Node役割が配置される全ノードがICP課金対象となることにより、
ノード役割を適切に配置した場合と比較して、ICP製品ライセンス費用の観点で不利になってしまう



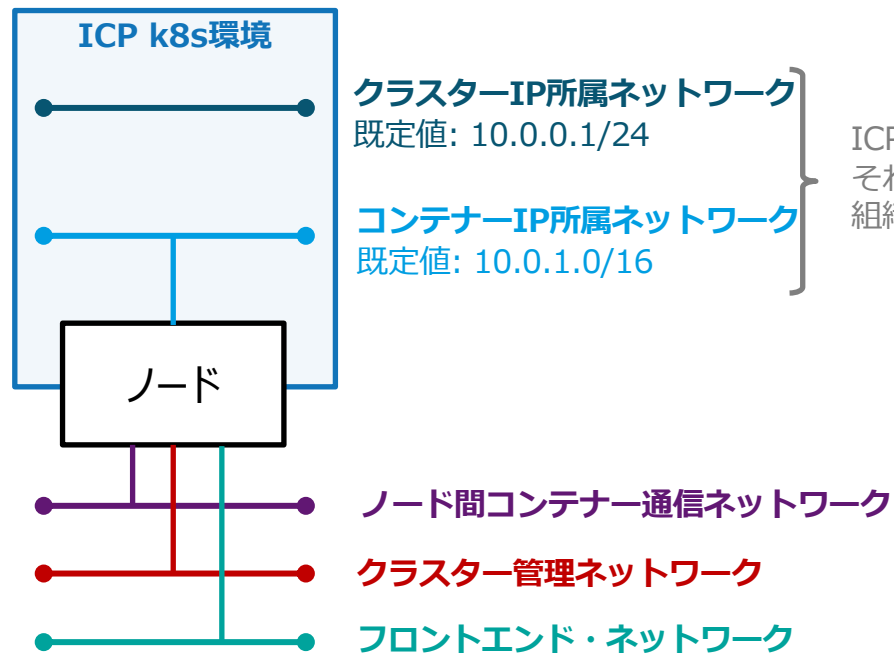
仮想プロセッサ・コア課金の対象となるノード上で
Master Node役割やProxy Node役割を稼動させることは
課金対象であるCPU資源の無駄遣いである。
Master Node役割、Proxy Node役割は別ノードに移すべき

ICP k8s環境構成 NGパターン集: (その3)

■ ICP k8s環境構成にあたってのNGパターンを以下に例示する (続き)

ネットワークアドレス・レンジの競合

※ICP k8s環境内部のアドレス・レンジと、組織内部のアドレス・レンジとの競合を避けること



ICP k8s環境のセットアップ時には、これらネットワークのアドレス・レンジを指定する。それらのアドレス・レンジが、ICP k8s環境を利用する組織内部のアドレス・レンジと競合する場合には、組織ネットワークの全体で、**ルーティング制御面での整合が保たれなくなるおそれ**が生じる。

ICP k8s環境構成 NGパターン集: (その4)

■ ICP k8s環境構成にあたってのNGパターンを以下に例示する (続き)

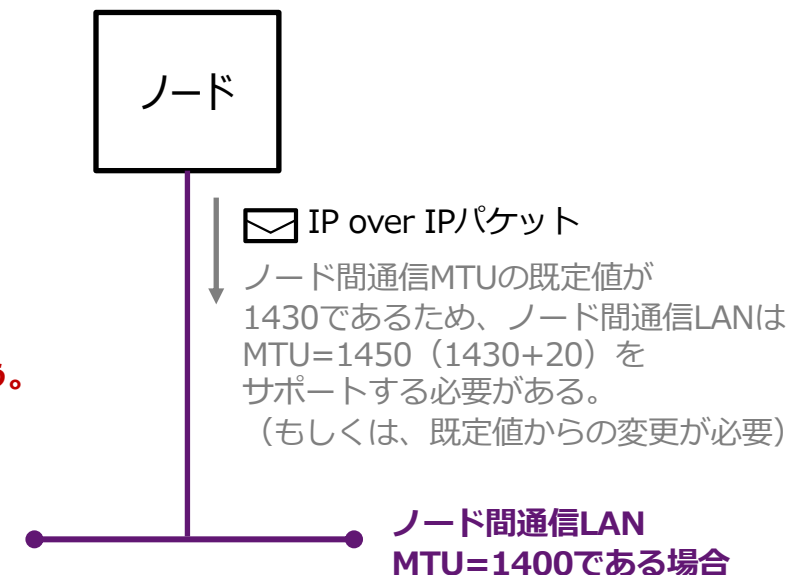
不適切な複数NIC構成

※複数NIC構成を設定として明示的に行わない場合、
各ノードのNICを用いた通信が意図せぬ形で行われ得る



MTU値の不整合

※ノード間コンテナ通信のためにCalicoの
IP over IP modeを用いる場合に、
ネットワーク基盤でサポートするMTU値が
ノード間通信でのMTU値を下回ると、
ネットワーク・スループットが低下し得る

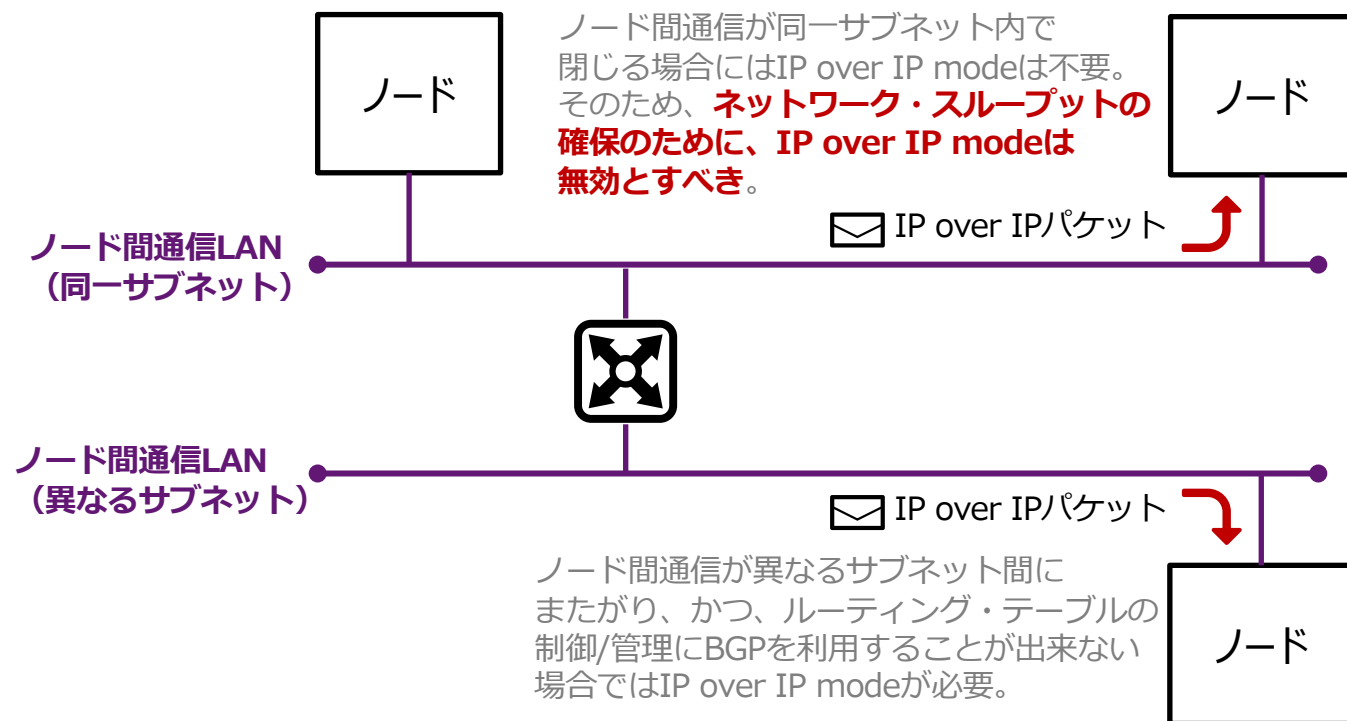


ICP k8s環境構成 NGパターン集: (その5)

■ ICP k8s環境構成にあたってのNGパターンを以下に例示する (続き)

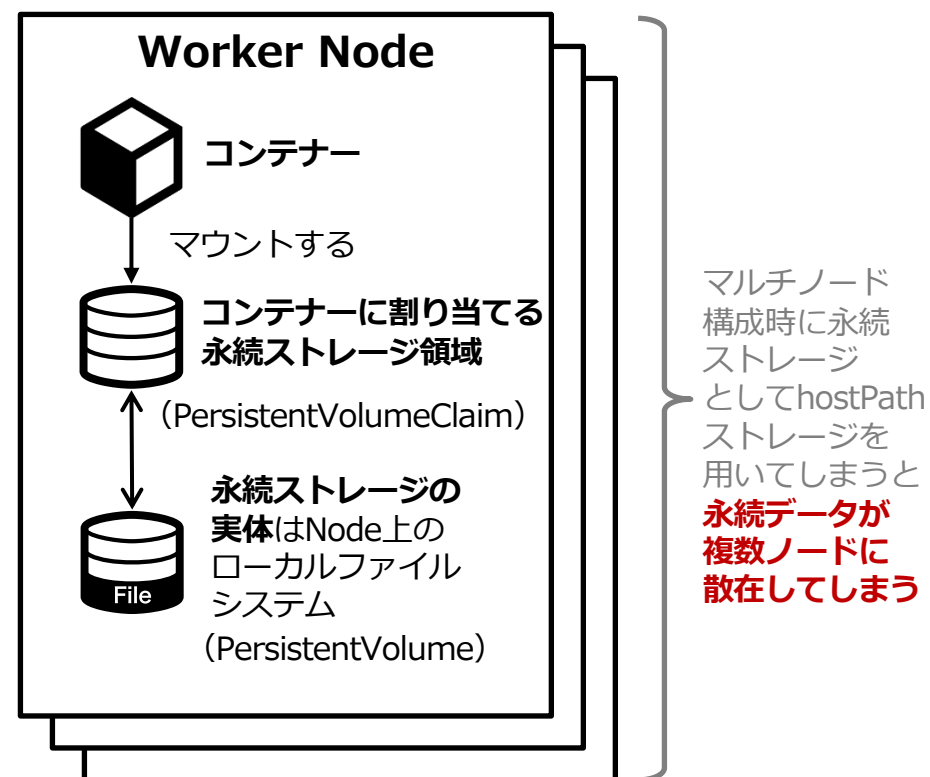
IP over IP modeの利用に伴う不必要な性能低下

※IP over IP modeはネットワーク・スループットの観点で不利であるため、ネットワーク性能が重視される環境では、IP over IP modeの使用は避けるべき



マルチノード構成でのhostPath永続ストレージ利用

※マルチノード構成下でhostPath永続ストレージを利用することは非推奨

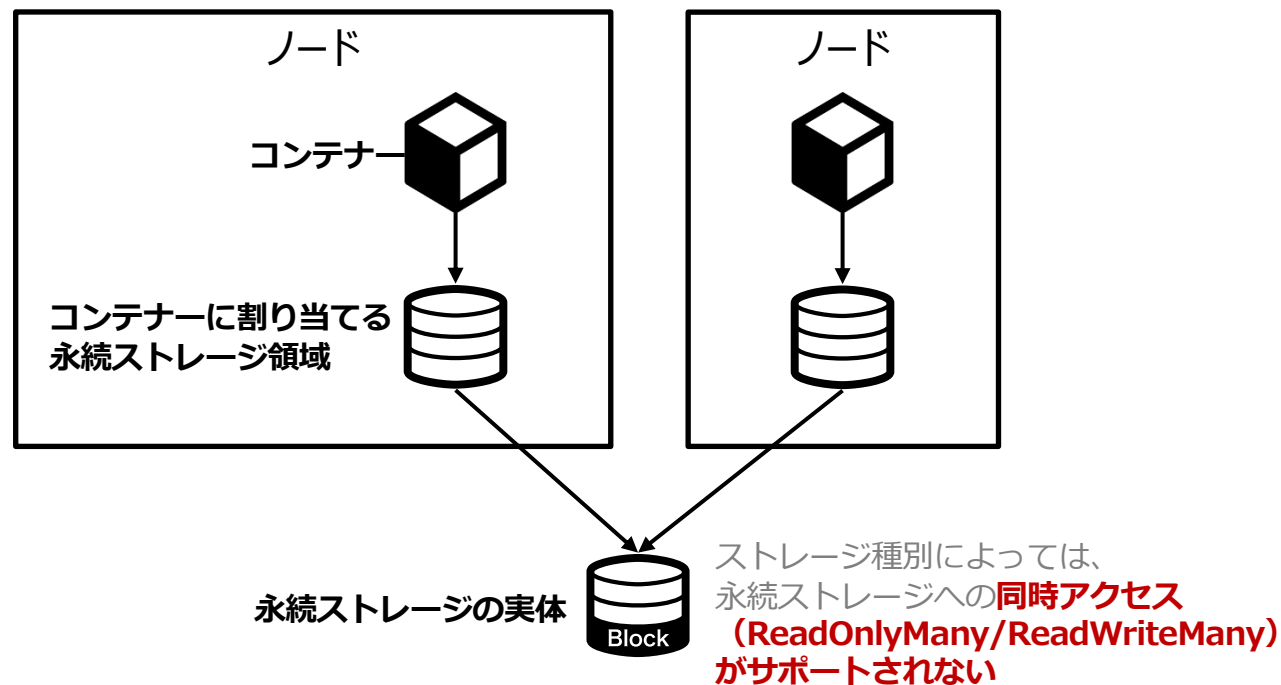


ICP k8s環境構成 NGパターン集: (その6)

■ ICP k8s環境構成にあたってのNGパターンを以下に例示する (続き)

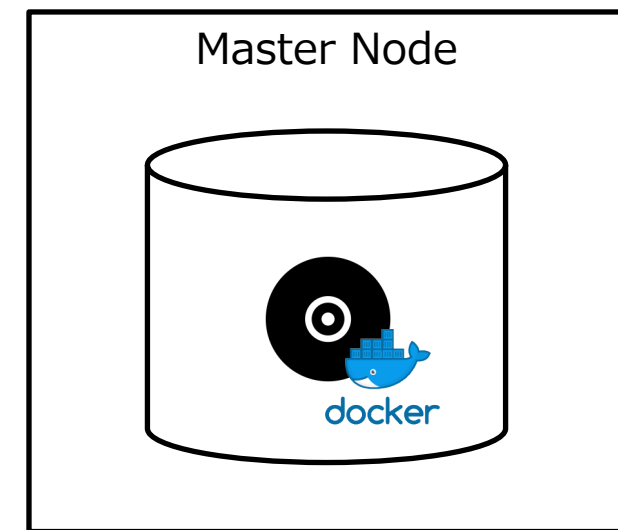
同時アクセスをサポートしない永続ストレージ種別の使用

※同時アクセスが要件とされる場合には、
同時アクセスをサポートする永続ストレージを用いなければならない



ノードでのファイルシステム容量不足

※各ノードのファイルシステムが肥大化する可能性がある



特に、Dockerイメージのプライベートレポジトリが
肥大化しがちであるため、サイジング時に留意のこと

ICP k8s環境構成 NGパターン集: (その7)

■ ICP k8s環境構成にあたってのNGパターンを以下に例示する (続き)

サポート対象外OSの使用

※各ノードではICPがサポートするOSを利用しなければならない

ネットワーク・ポートの競合

※ICPがノード上で使用するネットワーク・ポートと各ノード上の非ICPサービスが利用するネットワーク・ポートが競合してはならない

各ノード

サポートしないOSであるためNG



VAノード

VAノードはSLESをサポートしないためNG



ノード

ネットワーク・ポートが競合しているためNG



TCP 80で待ち受け

TCP 80で待ち受け

ICP k8s環境構成 NGパターン集: (その8)

■ ICP k8s環境構成にあたってのNGパターンを以下に例示する (続き)

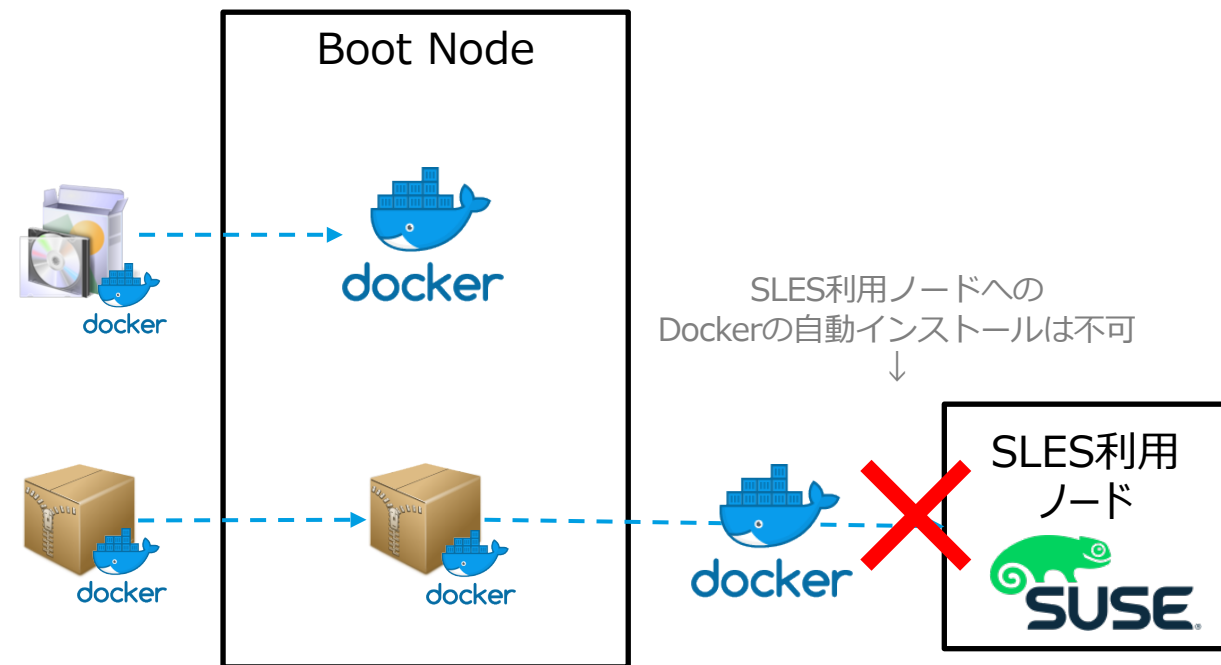
ICPインストール中における 許可されないFirewallの有効化

※MasterあるいはProxyにSLESを使用している場合、ICPインストール中はクラスター内のFirewallを全て無効化すること

Master あるいは Proxyノードに
SLESを利用している場合に
クラスター内のFirewallが有効化されているのはNG

SLESを用いたノードでの Dockerの自動インストール

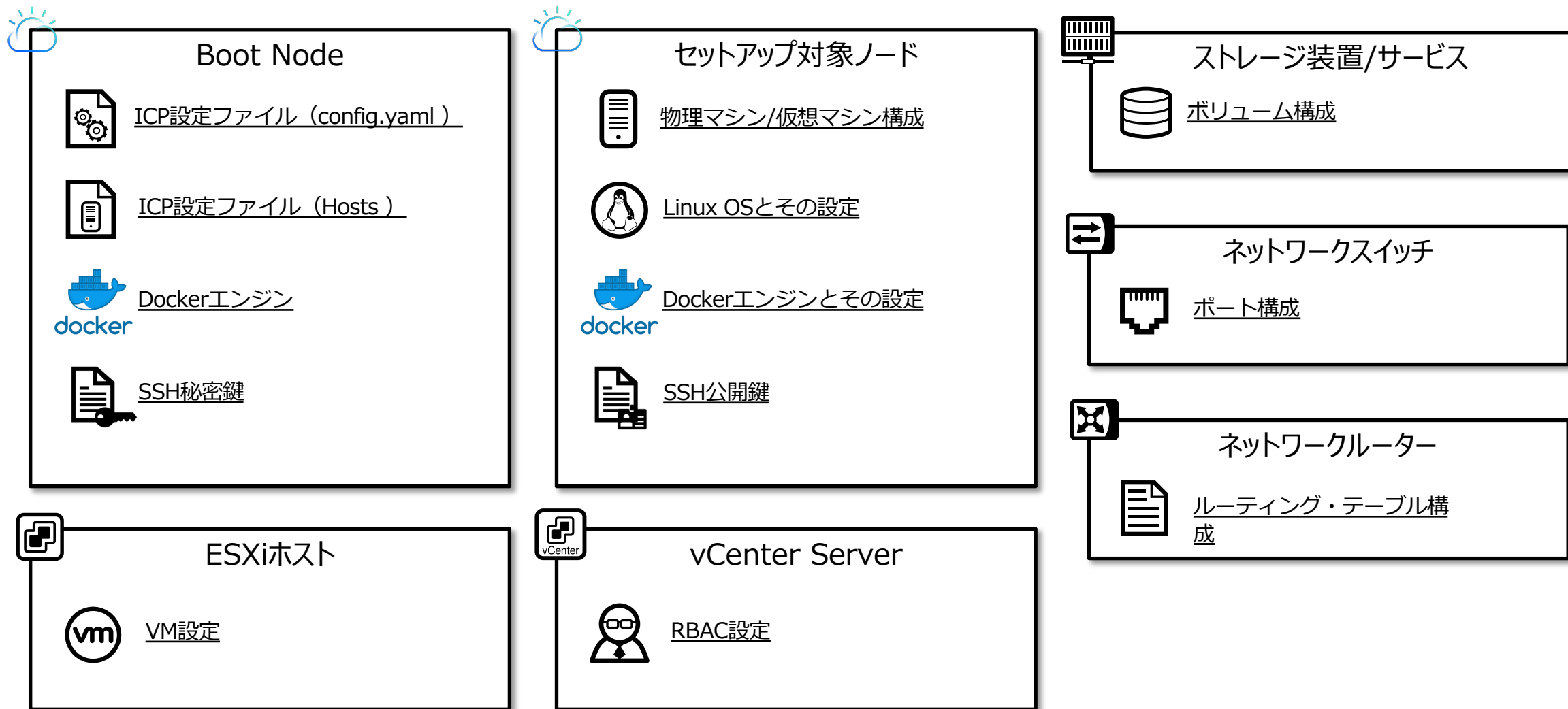
※SLESを用いたノードでは
Dockerを手動でインストールする必要がある



第2章 第10節: 検討結果の実装への反映

技術検討の結果は「どこの」実装に反映されるのか？（その1）

■ 技術検討の結果が反映されるシステム構成要素を以下に図示する



技術検討の結果は「どこの」実装に反映されるのか？（その2）

■ 技術検討の結果のシステム実装への反映について下表にまとめる

設定対象 (ノード、機器)	設定対象 (設定ファイル等)	技術検討結果の設定対象への反映方針	関連する IBM Knowledge Center情報
Boot Node	ICP設定ファイル (config.yaml)	[ネットワーク基盤技術検討の結果反映]: ICP k8s環境で使用するネットワーク、および、各ノードが使用するNICを設定として反映する。ノード間通信の動作に関わる各種の設定を投入する。 [ストレージ基盤技術検討の結果反映]: 永続ストレージとして「GlusterFS」、「vSphereボリューム」もしくは「NFS」を用いる場合には、そのための設定を反映する。 [OS構成技術検討の結果反映]: ノード間認証のためにパスワードを用いる場合には、そのための設定を反映する。また、ICP製品とOSとの間でのネットワーク・ポートの競合を避けるための設定を反映する場合も生じうる。 [Dockerエンジン技術検討の結果反映]: Dockerエンジンを各ノードに配布するにあたって、各ノードに引き渡す引数、およびDocker環境変数を反映する。 [ICPクラスター構成技術検討の結果反映]: ICP k8s環境での管理機能を提供している、Kubernetesに代表される各種ソフトウェアの動作を制御するための設定を反映する。	• IBM Knowledge Center - config.yaml ファイルを使用したクラスターのカスタマイズ • IBM Knowledge Center - IPsecを使用したクラスター・データ・ネットワーク・トラフィックの暗号化 • IBM Knowledge Center - GlusterFS PersistentVolume の作成 • IBM Knowledge Center - vSphere クラウド・プロバイダー • IBM Knowledge Center - NFS PersistentVolume の作成 • IBM Knowledge Center - クラスター・ノードのパスワード認証の構成 • IBM Knowledge Center - IBM Cloud Private Cloud Native、Enterprise、および Community Edition のインストール

技術検討の結果は「どこの」実装に反映されるのか？（その3）

■ 技術検討の結果のシステム実装への反映について下表にまとめる（続き）

設定対象 (ノード、機器)	設定対象 (設定ファイル等)	技術検討結果の設定対象への反映方針	関連する IBM Knowledge Center情報
Boot Node	ICP設定ファイル (Hosts) ※: 「/etc/hosts」 とは別物	[コンピューティング基盤技術検討の結果反映]: ICP k8s環境の各ノード役割を担うLinuxサーバーそれぞれのIPアドレスを反映する [OS構成技術検討の結果反映]: ノード間認証のためにパスワードを用いる場合で、かつ、ノード毎にパスワードが異なるならば、ノード毎のパスワードを反映する。 [ICPクラスター構成技術検討の結果反映]: ICP k8s環境の各ノードに引き渡す、Kubernetesなどの各種ソフトウェアの動作を制御するための設定を反映する。	• IBM Knowledge Center - hosts ファイルでのノード役割の設定
	Dockerエンジン	[Dockerエンジン技術検討の結果反映]: ICP製品がサポートするバージョンのDockerソフトウェアをBoot Nodeに導入する。	• IBM Knowledge Center - サポートされる Docker バージョン • IBM Knowledge Center - IBM Cloud Private の Docker のセットアップ
	SSH暗号鍵	[OS構成技術検討の結果反映]: ノード間認証のためにSSH鍵ペアを用いる場合には、SSH暗号鍵をBoot Nodeに配置する。	• IBM Knowledge Center - クラスター・ノードでの SSH 鍵の共有

技術検討の結果は「どこの」実装に反映されるのか？（その4）

■ 技術検討の結果のシステム実装への反映について下表にまとめる（続き）

設定対象 (ノード、機器)	設定対象 (設定ファイル等)	技術検討結果の設定対象への反映方針	関連する IBM Knowledge Center情報
セットアップ対象 ノード	物理マシン/仮想マ シン構成	[コンピューティング基盤技術検討の結果反映]: ICP k8s環境でのノード構成を満足する物理マシン、もしくは仮想マシンを手配する。 [ネットワーク基盤技術検討の結果反映]: ICP k8s環境の各ノードに対応する仮想マシンのそれぞれが必要なネットワークに接続するよう構成する。 [ストレージ基盤技術検討の結果反映]: 永続ストレージとして「GlusterFS」を用いる場合、GlusterFSの構成要件を満足するようにノードを構成する。	• IBM Knowledge Center - ハードウェア要件および推奨 • IBM Knowledge Center - サポートされるオペレーティング・システムとプラットフォーム • IBM Knowledge Center - システム要件
	OS構成全般	[ネットワーク基盤技術検討の結果反映]: ノードが所属するネットワークへの接続のための設定（例:ネットワーク・インターフェースへのIPアドレス付与）を行う。 [ストレージ基盤技術検討の結果反映]: 必要に応じて、永続ストレージとして用いるストレージ領域、およびICP k8s環境の運用データを保管するためのストレージ領域を各ノードからマウントする。また、Master NodeがHA構成を採る場合には、複数のMaster Node間でデータを共有するための共用ストレージ領域をマウントする。 [OS構成技術検討の結果反映]: ICP製品のセットアップのための一連の前提項目をOS構成に反映する。	• IBM Knowledge Center - ノードの準備 • IBM Knowledge Center - ハードウェア要件および推奨 • IBM Knowledge Center - IBM Cloud Private Cloud Native、Enterprise、および Community Edition のインストール

技術検討の結果は「どこの」実装に反映されるのか？（その5）

■ 技術検討の結果のシステム実装への反映について下表にまとめる（続き）

設定対象 (ノード、機器)	設定対象 (設定ファイル等)	技術検討結果の設定対象への反映方針	関連する IBM Knowledge Center情報
セットアップ対象 ノード	Dockerエンジン	[Dockerエンジン技術検討の結果反映]: ICP機能を用いたDockerソフトウェアの各ノードへの展開を用いない場合には、ICP製品がサポートするバージョンのDockerソフトウェアを各ノードに明示的に導入する。	• IBM Knowledge Center - サポートされる Docker バージョン • IBM Knowledge Center - IBM Cloud Private の Docker のセットアップ
	SSH公開鍵	[OS構成技術検討の結果反映]: ノード間認証のためにSSH鍵ペアを用いる場合には、SSH公開鍵を各ノードに登録する。	• IBM Knowledge Center - クラスター・ノードでの SSH 鍵の共有
ESXiホスト	仮想マシン	[ストレージ基盤技術検討の結果反映]: 永続ストレージとして「vSphereボリューム」を用いる場合には、永続ストレージ利用のためのvSphere クラウド・プロバイダーを実装するための各種の制約（例:VM名の制約）を反映する。	• IBM Knowledge Center - vSphere クラウド・プロバイダー/前提条件および制限
vCenter Server	RBAC設定	[ストレージ基盤技術検討の結果反映]: 永続ストレージとして「vSphereボリューム」を用いる場合には、永続ストレージ利用のためのvSphere クラウド・プロバイダーを実装するための設定を反映する。	• IBM Knowledge Center - vSphere クラウド・プロバイダー

技術検討の結果は「どこの」実装に反映されるのか？（その6）

■ 技術検討の結果のシステム実装への反映について下表にまとめる（続き）

設定対象 (ノード、機器)	設定対象 (設定ファイル等)	技術検討結果の設定対象への反映方針	関連する IBM Knowledge Center情報
ストレージ装置/ サービス	ストレージボリューム構成	[ストレージ基盤技術検討の結果反映]: ICP k8s環境の永続ストレージをサポートするためのボリューム構成を行う。	IBM Knowledge Center - ストレージ・ガイド
ノード対向のネットワークスイッチ	構成全般	[ネットワーク基盤技術検討の結果反映]: ICP k8s環境で発生する通信をサポートするための構成を行う。	IBM Knowledge Center - ネットワーク・ガイド
組織内のネットワークルーター	構成全般	[ネットワーク基盤技術検討の結果反映]: ICP k8s環境が占有するネットワークを疎通要件に応じてルーティング・テーブルに反映する。	IBM Knowledge Center - ネットワーク・ガイド