

# クラウド基盤を 支える要素技術 2012

Internet Initiative Japan Inc.  
齊藤 秀喜 <saitou@iij.ad.jp>

- IJ Technical Week 2012 -

# 本セッションの概略

IIJが持つ研究開発用データセンターのインフラを例に、現在のクラウド(IaaS)基盤を支える要素技術と、今後の課題について解説する20分のセッションです。

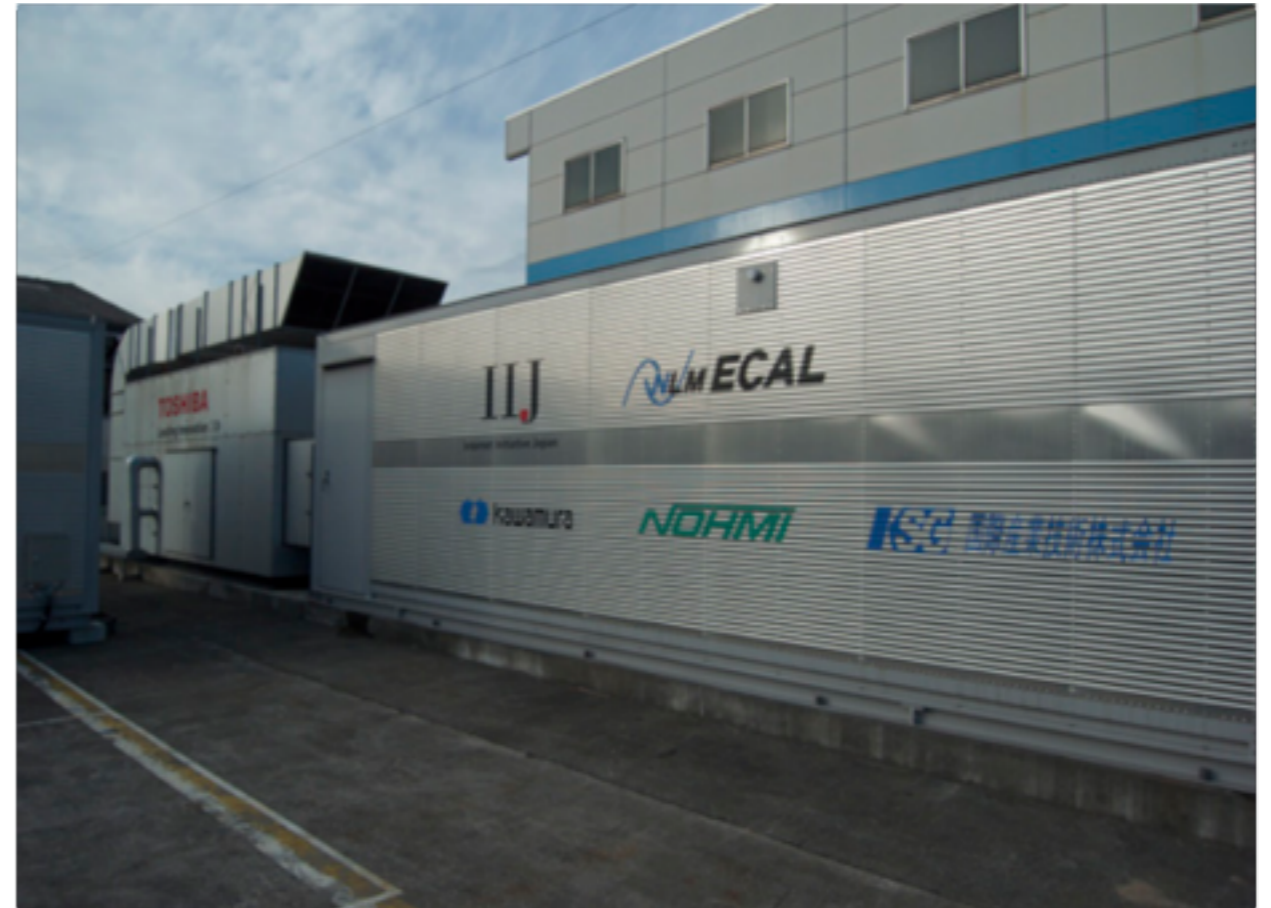
# 目次

- 研究開発用データセンターについて
- 実証実験用インフラを構築・維持管理する手法
- クラウド(IaaS)基盤を支える技術
- 今後のクラウド(IaaS)基盤に求められるもの

# 研究開発用DC

- コンテナ型データセンターを利用
- サーバ収容台数: 約200台
- 各種計測用センサー多数

コンテナDC内に収容しているサーバやネットワーク機器等の実験用設備を利用して、将来、標準化や普及が期待される技術についての実証実験を行う目的で稼働中。サーバ、ネットワーク機器共に実験内容に応じた様々なベンダ・スペックのものを収容している。



# 実証実験用インフラを 構築・維持・管理する 手法

# ITインフラ (クラウド時代以前)



# ITインフラ (クラウド時代)

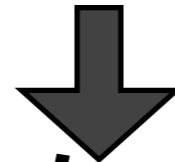
The screenshot shows the IIJ GIO website homepage. The browser address bar displays 'www.ij.ad.jp/GIO/'. The page features a navigation menu with 'トップ', 'IIJ GIOとは', 'サービス一覧', '導入事例', and 'パートナープログラム'. A large banner at the top left reads '日本のクラウド、IIJ GIO' and '1,000通り以上の豊富なシステム構成。' with a globe graphic. Below the banner are three promotional boxes: 'VMware Readyな仮想化プラットフォーム', '大ヒットゲームも続々稼働中! IIJ GIOソーシャルアプリ支援ソリューション', and 'IIJ GIO アプリケーションサービス SaaS' (listing 'サイボウズ ガルーン SaaS' and 'Smart POSサービス'). The middle section lists 'IIJ GIO プラットフォームサービス PaaS' (listing 'ホスティングパッケージサービス', 'ストレージサービス', and '仮想デスクトップサービス') and 'IIJ GIO コンポーネントサービス HaaS/IaaS' (listing 'コンポーネントサービス' and 'Cloud Migration Factory for Windows'). On the right side, there is a contact form with 'お問い合わせ' and '資料請求、見積依頼もお気軽に' buttons, a phone number '03-5205-4466', and social media links for Facebook and Twitter. A 'キャンペーン実施中' banner is at the bottom right.

研究開発用DCのインフラ  
にもGIOが使えれば幸せ...  
しかし、今回は残念ながら  
マッチせず



# 研究開発用DCCの 難しさ

- 仮想環境だけでなく物理環境も必要とされる
- 収容する機器は、時期・用途によりさまざま。設定・管理方法もさまざま
- 高頻度で(論理的に)破壊される
- これまでにない監視・性能計測の仕組みが必要
- 短期間で準備しなければならない



GIOではない専用の  
クラウド基盤を用意する

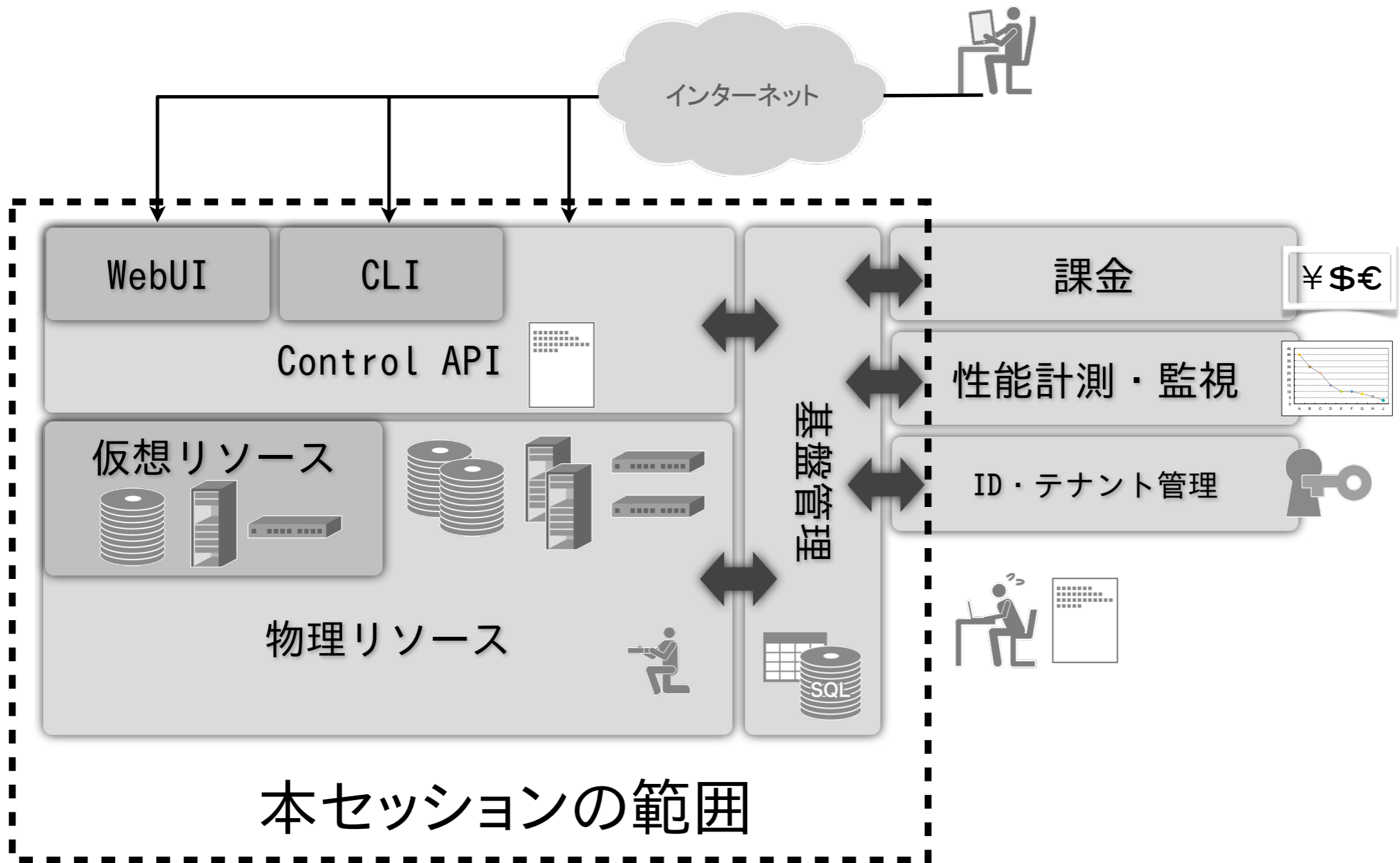
# クラウド(IaaS)基盤を 支える技術

- 物理基盤管理
- 仮想化基盤管理

本セッションの範囲

- ID・テナント管理
- システムの監視・性能計測
- 課金
- ユーザーインターフェイス

# 各要素の相関関係

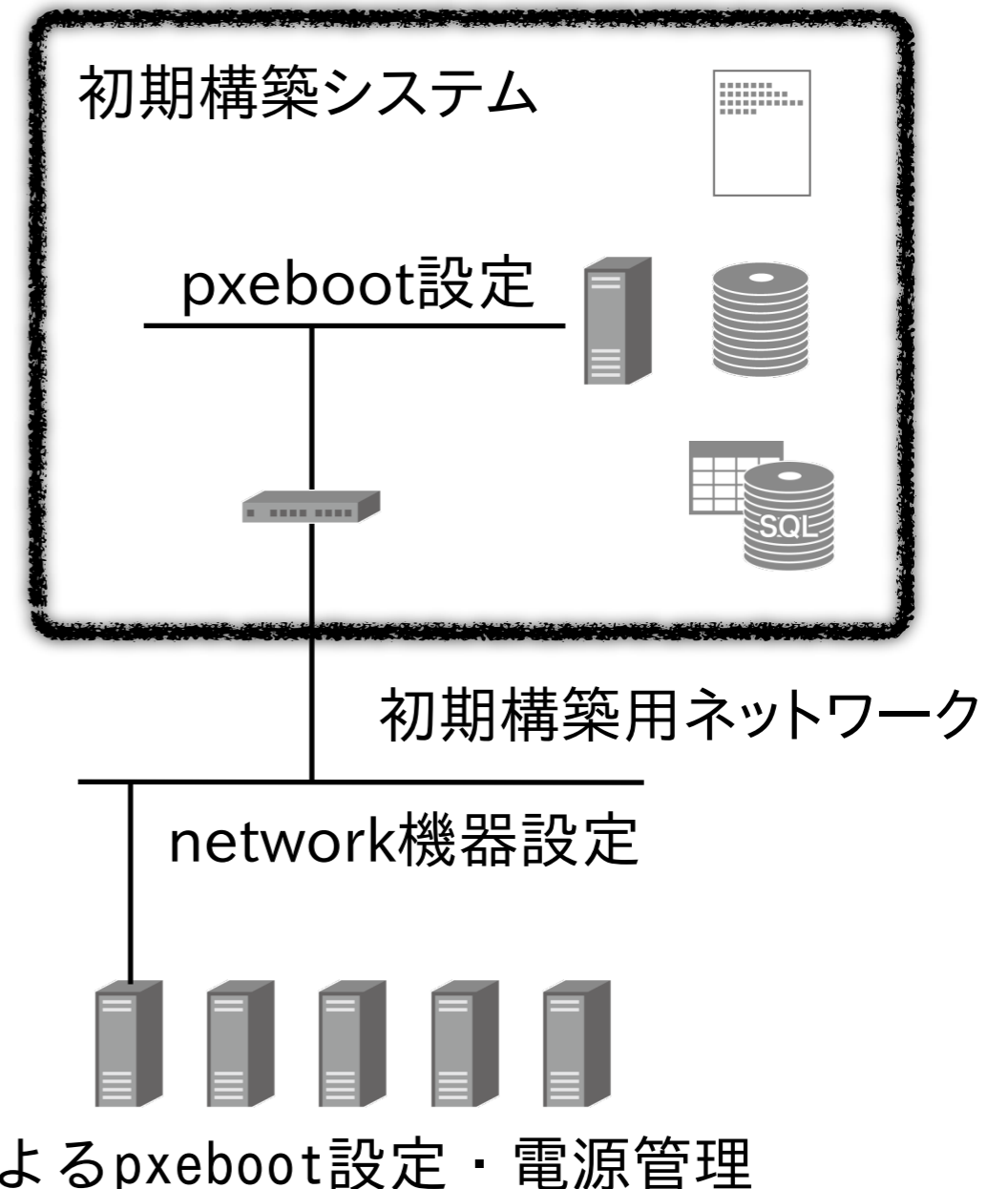


# 物理基盤管理(1)

仮想リソースを利用するにしても、結局は物理リソースの上に仮想環境を構築することとなる。

物理リソース管理するための仕組みは、特に新しい技術ではなく、これまで培ってきた技術・知見を活かすことが可能。

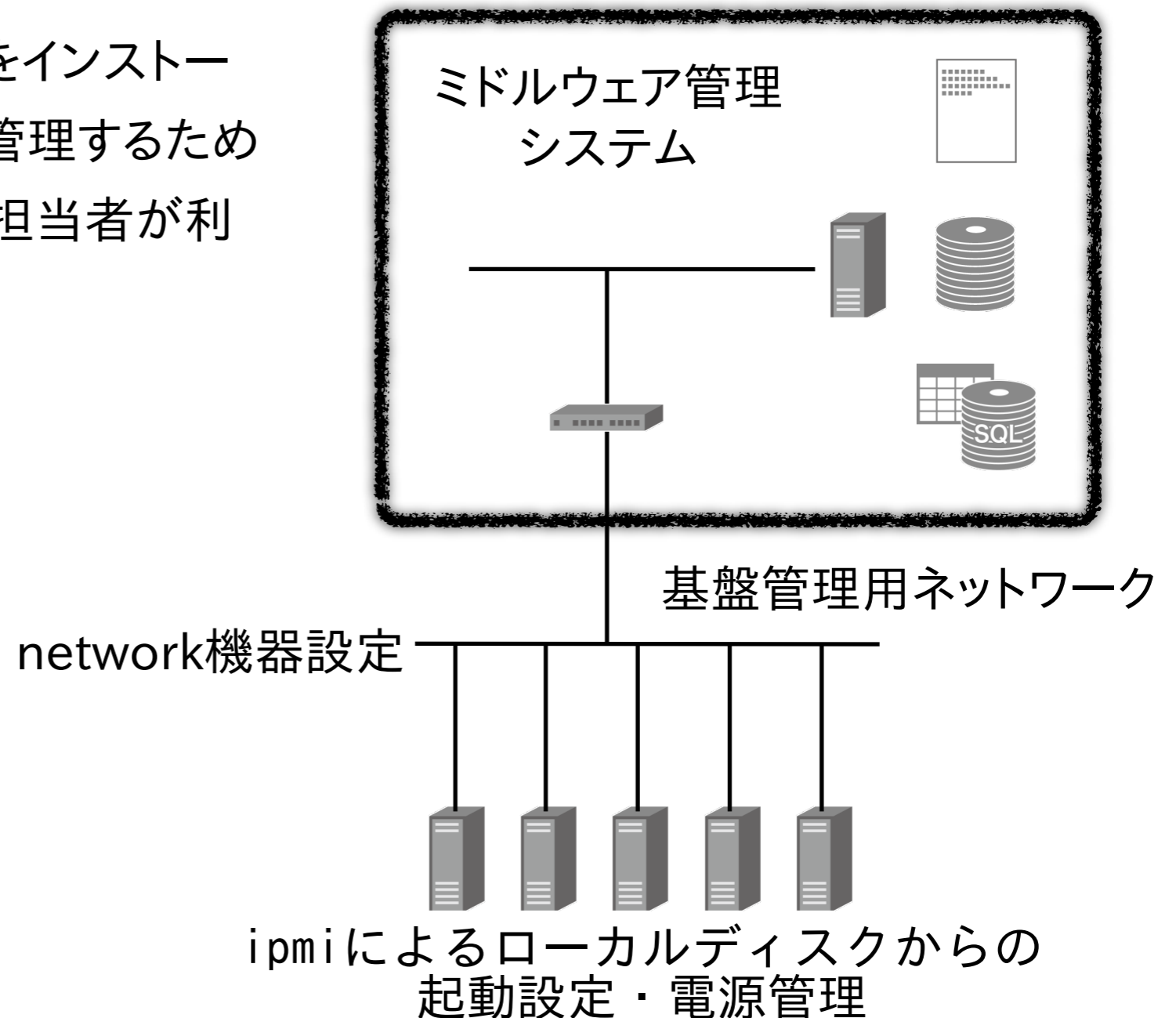
- netconf
- ipmi
- pxeboot/tftp/dhcp/syslinux
- syslinux
- ftp
- nfs



# 物理基盤管理(2)

OSインストール後にミドルウェアをインストール・設定する等、サーバの構成を管理するための手段にはいくつかあるが、作業担当者が利用しやすいものを提供。

- capistrano
- fabric
- expect
- pssh
- puppet
- chef
- Juju



# 仮想化基盤管理(1)

仮想マシン・ストレージ・ネットワーク管理などの仮想化基盤を管理するための技術については、ここ2年ほどで急速に進歩しており、商用・オープンソースともに多くのプロダクトがリリースされているが、オープンソースプロダクトとしては、以下の3つが主要なものとしてあげられる。

- OpenStack
- Apache CloudStack
- Eucalyptus

これらの中で、IJJの研究開発用データセンターでは、OpenStackを採用している。



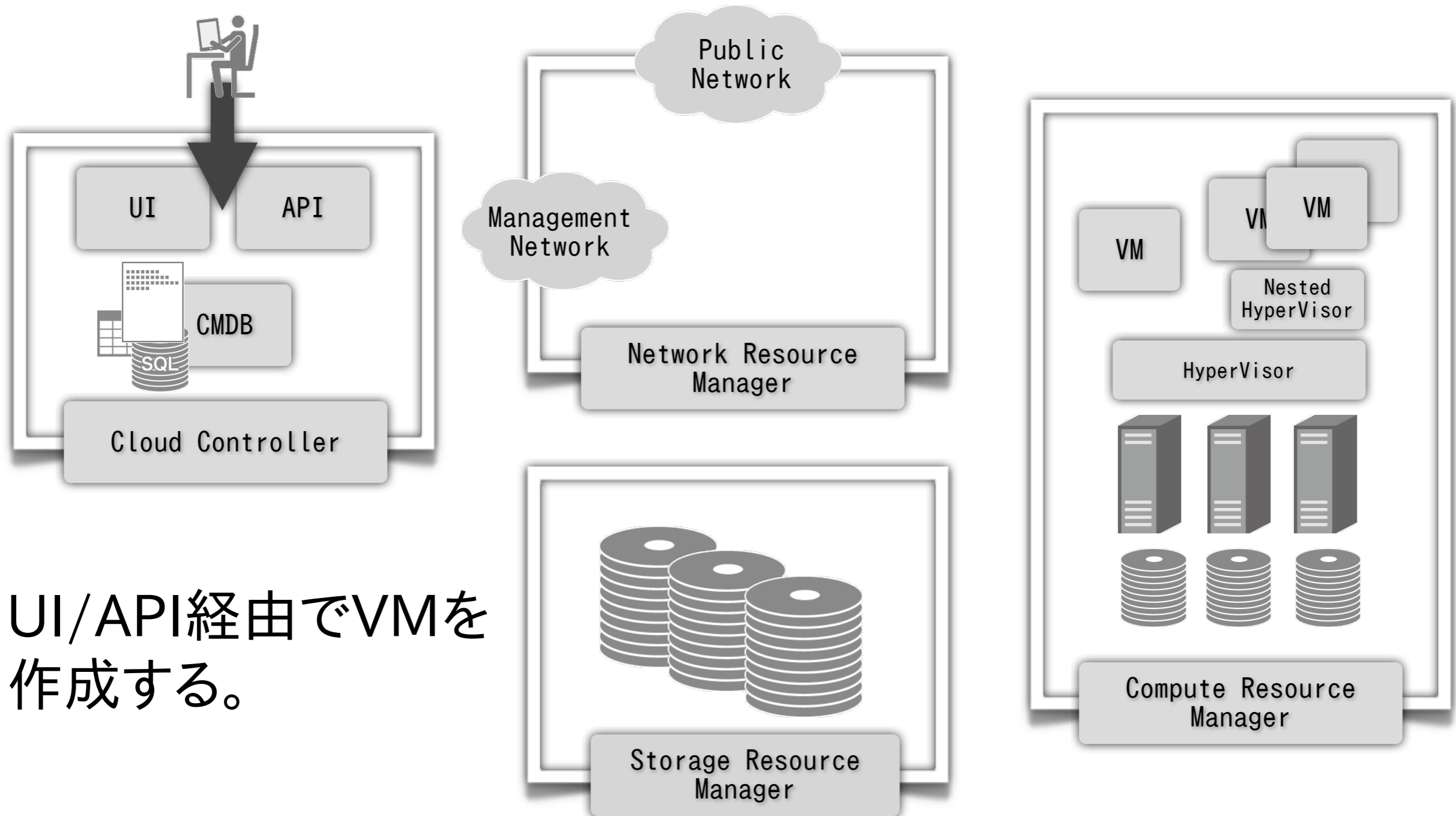
# 仮想化基盤管理(2)

OpenStackのバージョンはEssexを採用。  
最新リリースはfolsomで、Essexからの移行に向けて現在  
検証中。



- austin[2010.1] (2010-10-21)
  - bexar[2011.1] (2011-02-03)
    - cactus[2011.2] (2011-04-15)
      - diablo[2011.3] (2011-09-22)
        - **essex[2012.1] (2012-04-05)**
          - folsom[2012.2] (2012-09-27)

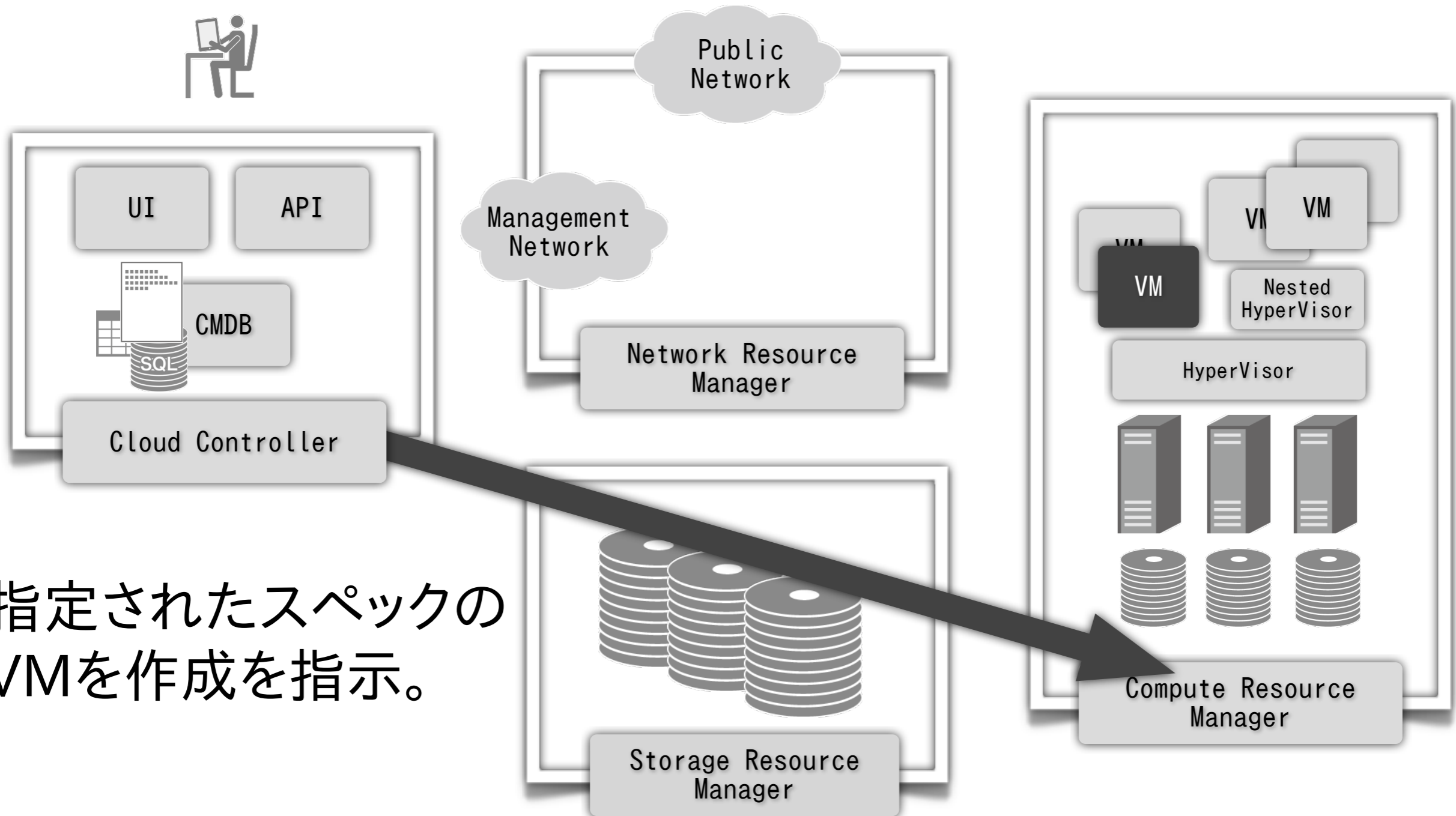
# 仮想化基盤管理(3)



UI/API経由でVMを作成する。

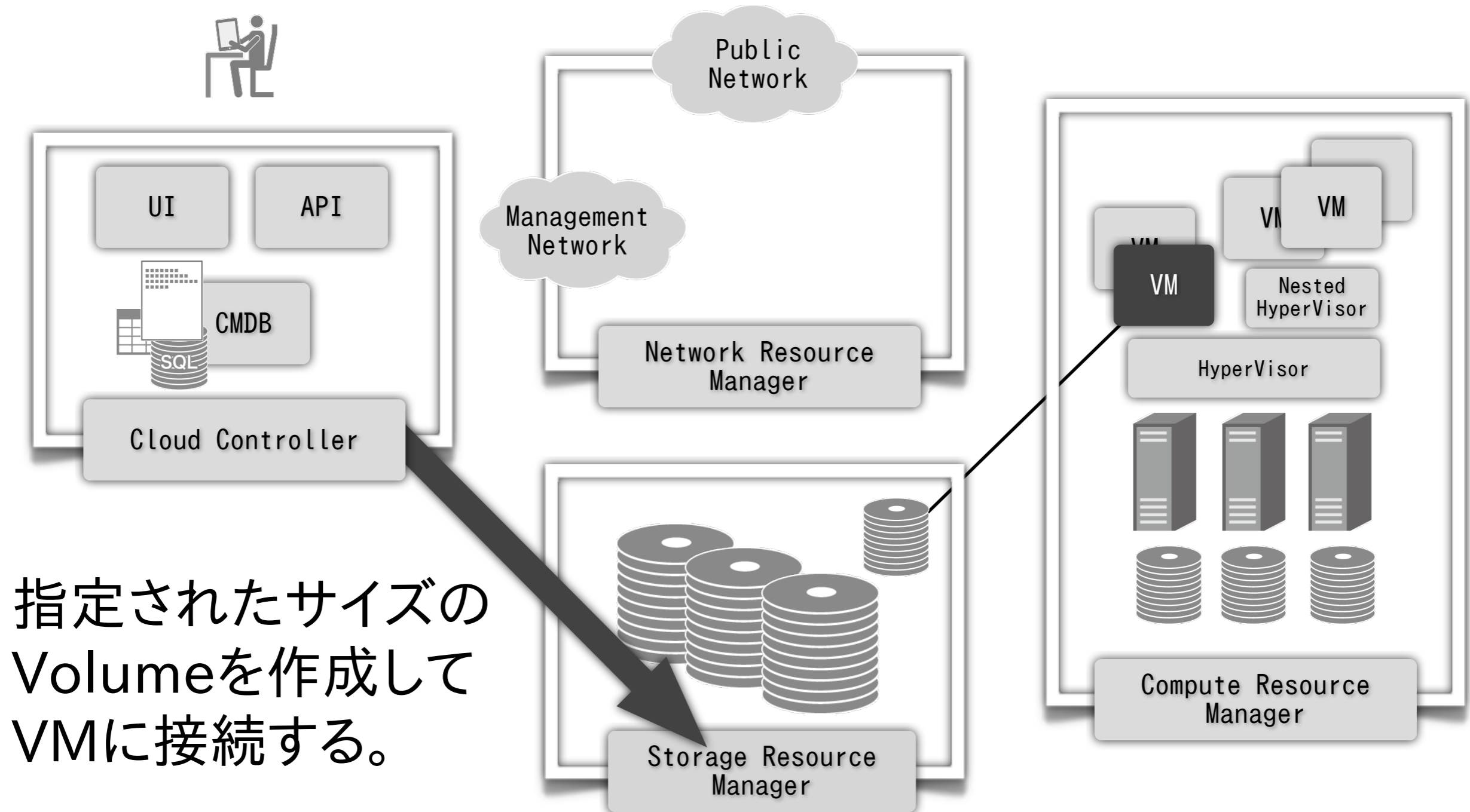


# 仮想化基盤管理(4)



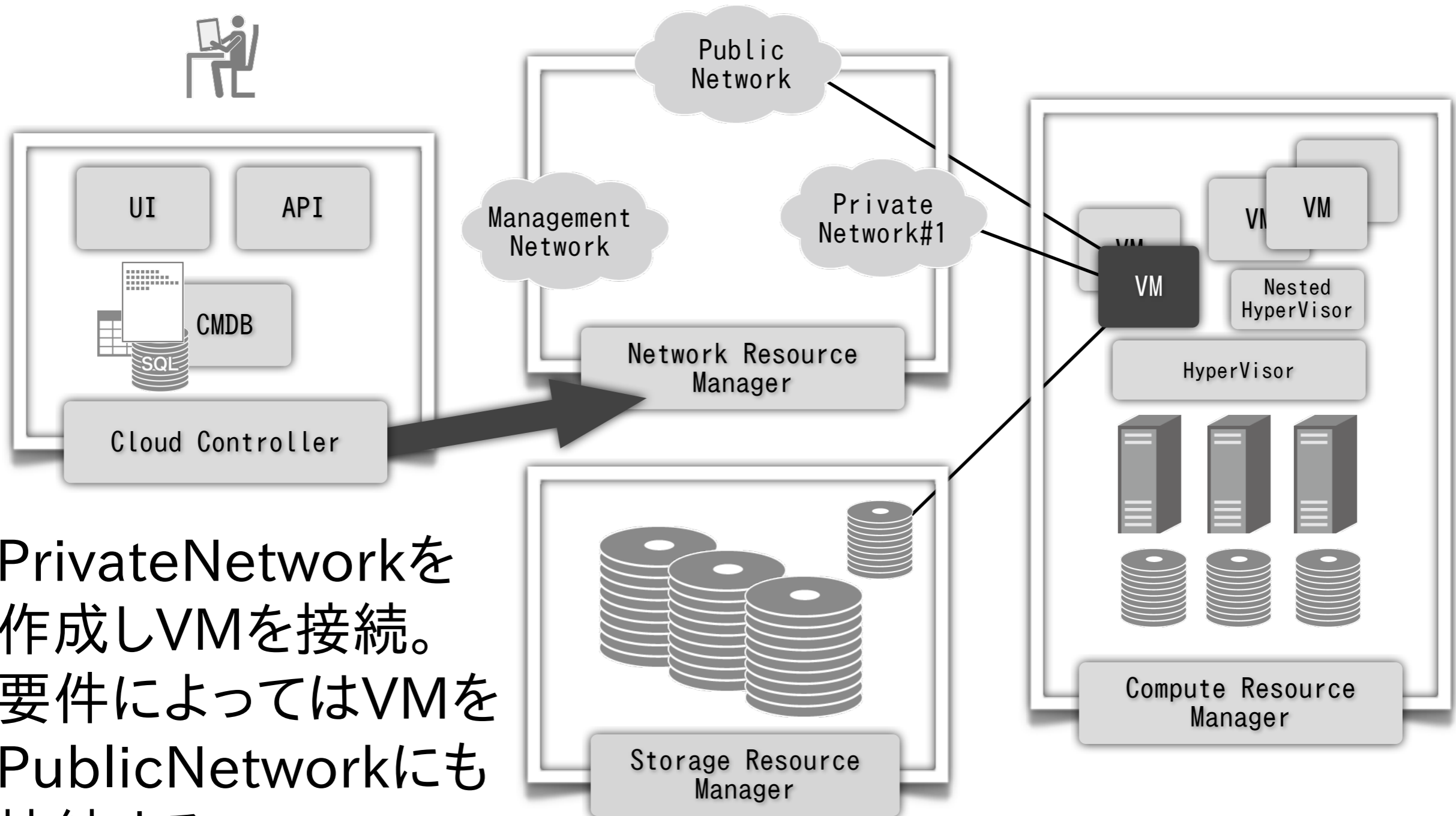
指定されたスペックの  
VMを作成を指示。

# 仮想化基盤管理(5)



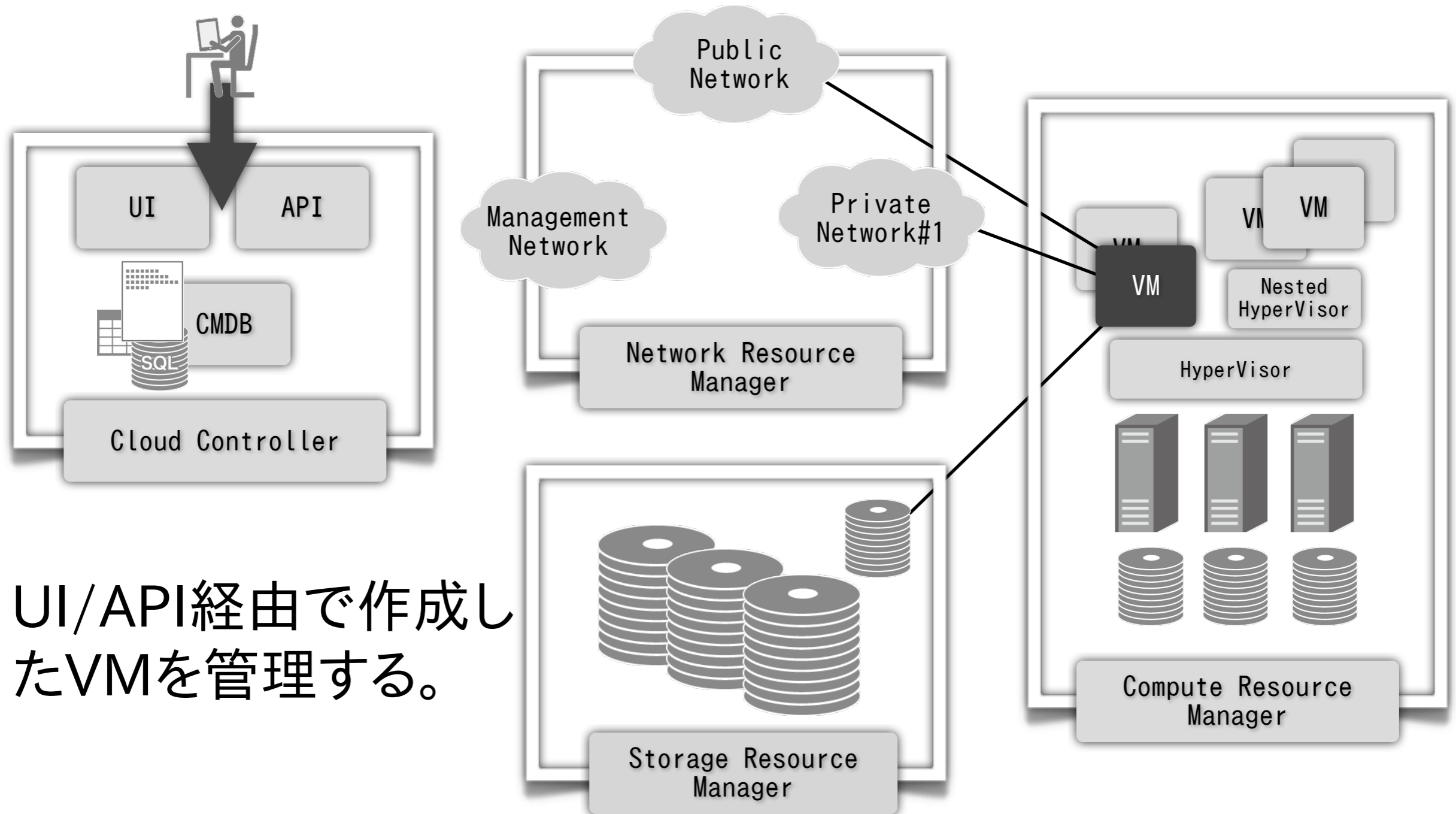
指定されたサイズの  
Volumeを作成して  
VMに接続する。

# 仮想化基盤管理(6)



PrivateNetworkを  
作成しVMを接続。  
要件によってはVMを  
PublicNetworkにも  
接続する。

# 仮想化基盤管理(7)



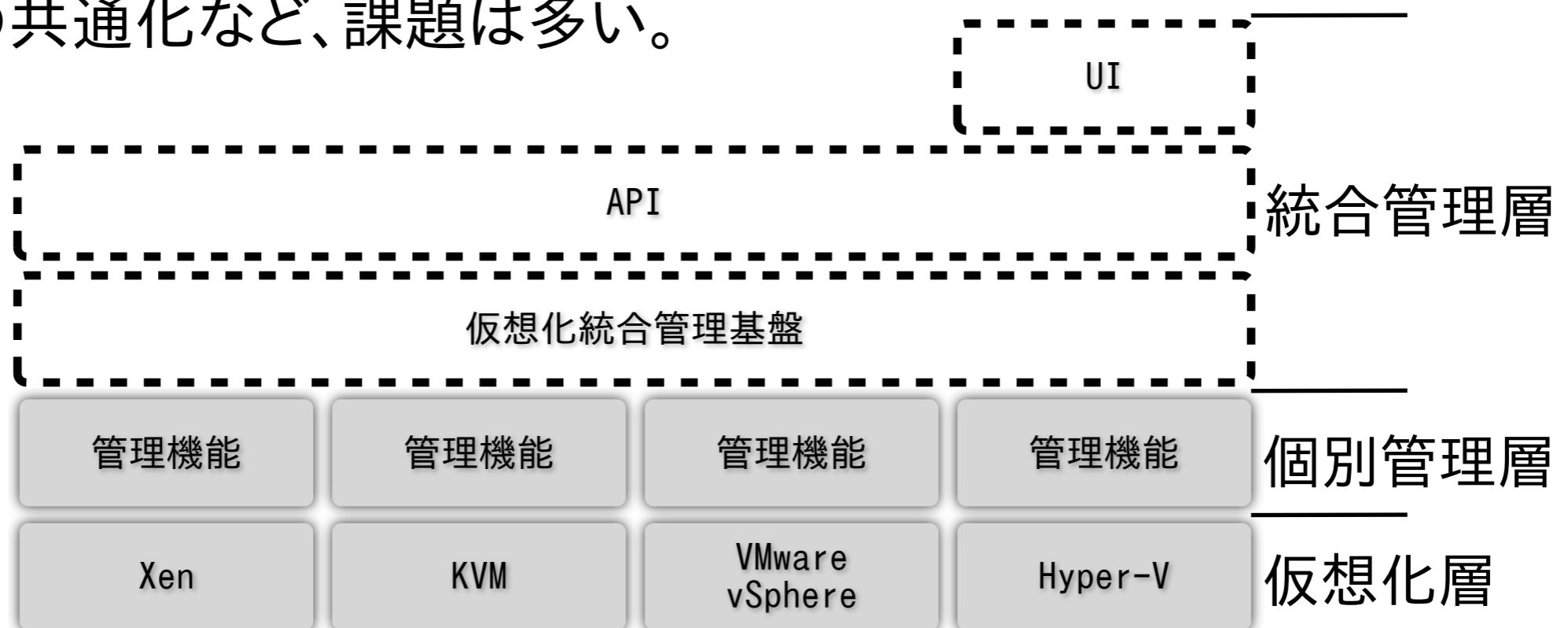
UI/API経由で作成したVMを管理する。

今後のクラウド(IaaS)基盤  
に求められるもの

# 仮想化基盤の 統合管理

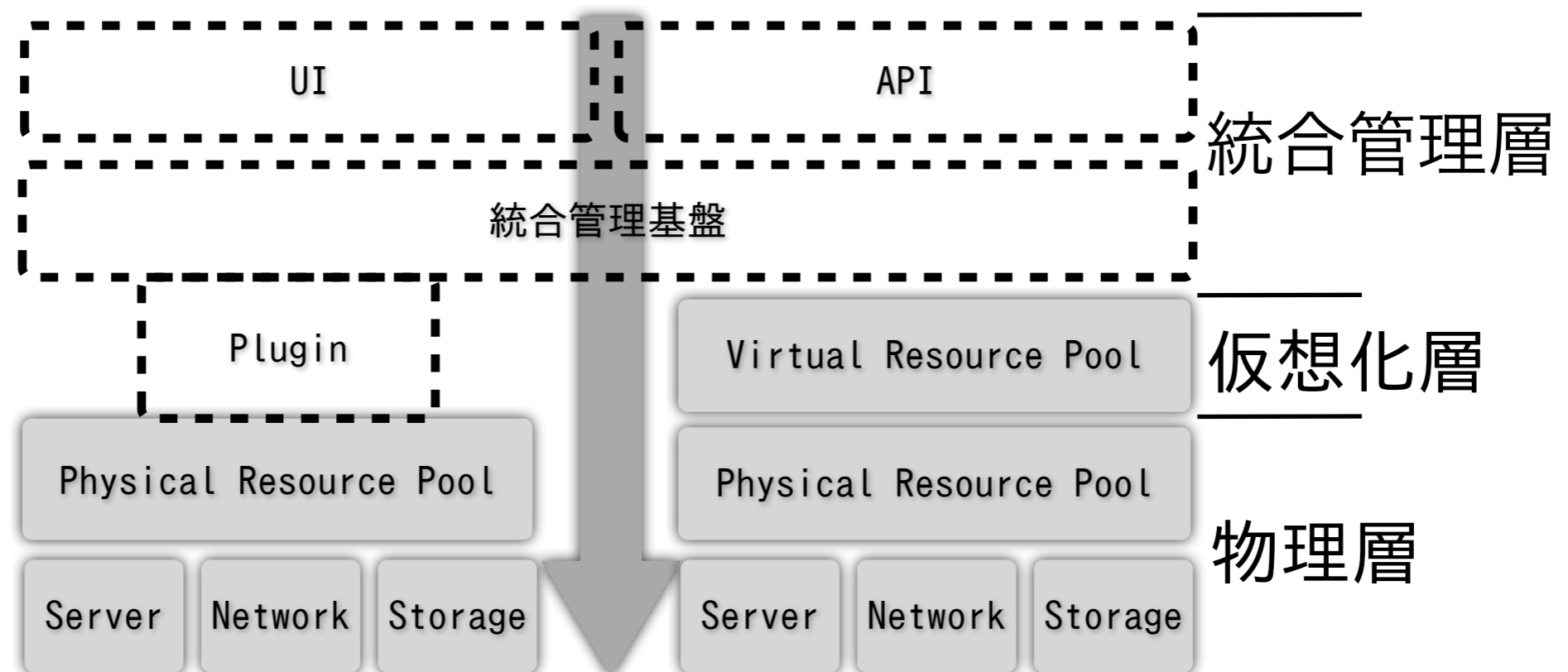
異なる仮想化基盤をフロントエンドで統合管理するための手法が必要となる。

仮想化基盤管理ソフトウェア (OpenStack、CloudStack、Eucalyptus など) は多様な仮想化基盤に対応する機能を持っているが、仮想化基盤それぞれが独自に用意している管理機能との親和性やID・テナント管理の共通化など、課題は多い。



# 物理 & 仮想化基盤の 統合管理

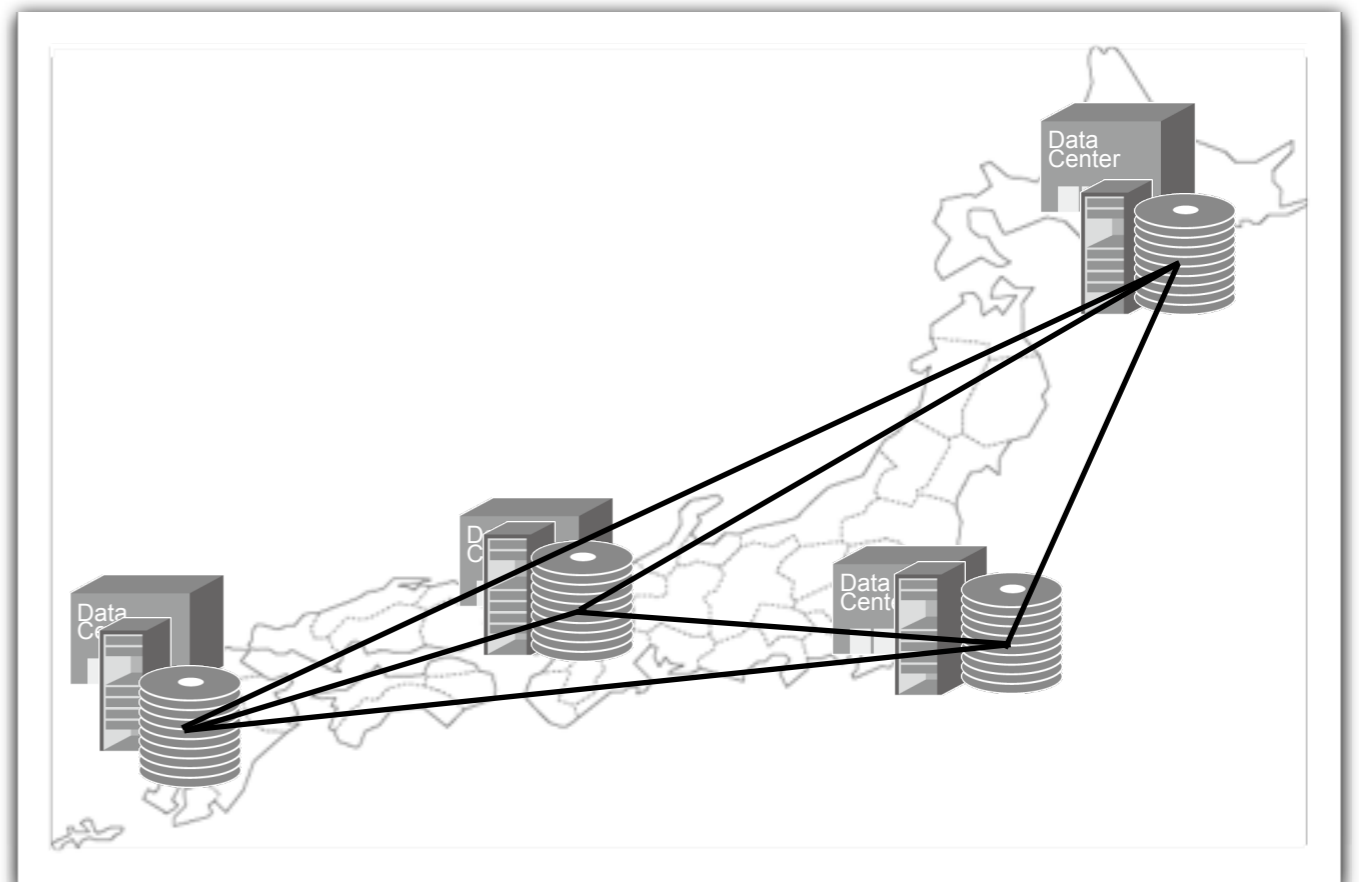
物理基盤を管理する機能については、前出の仮想化基盤管理ソフトウェアのプラグインとして既に提供されているものもあるが対応範囲は狭く、結局は利用するハードウェア毎にプラグインを独自開発する必要がある。



# 分散配置された クラウド基盤の連携

分散配置されたクラウド基盤を一元的に管理する手法は、物理・仮想化基盤を統合管理する技術の延長線上にあるため、統合管理基盤の構築は必須。

現状、サイトをまたがるライブマイグレーションの実現性や、物理サーバの配置問題など課題は多いが、ここ1年で、SDNによる柔軟なサイト間接続など機能限定的ながら実現性が高そうな要素技術も出てきている。





# まとめ

- クラウド基盤管理には物理・仮想どちらの技術も必要。
- 物理基盤管理は従来の技術と工夫で十分に対応可能。
- 仮想化基盤管理技術は、ここ2年ほどで急速に進化している。
- 今後は物理・仮想リソース管理のための個々の要素技術や、異なるサイトに分散した基盤を、いかに統合して一元管理できるかが鍵となる。

ご清聴ありがとうございました

# クラウド基盤を支える要素技術

## 研究開発用DCにおけるデータ収集



株式会社インターネットイニシアティブ  
プラットフォームサービス本部  
プラットフォームサービス部 プラットフォーム開発課  
阿部 博

Ongoing Innovation

20<sup>th</sup>  
Anniversary

# 目次

- 研究開発用データセンター
- データ収集の目的
- 収集するデータ
- データの収集手法／蓄積手法

# 研究開発用DC設備

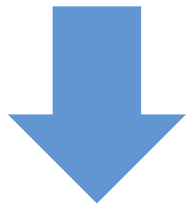
- サーバ約200台
- スイッチ数10台
- UPS
- 気象センサー
- 温室度センサー
- 空調制御装置
- 電力計



# データ収集の目的

---

- 空調制御効率化
- 実験時のデータ観測
- 実験機器の監視

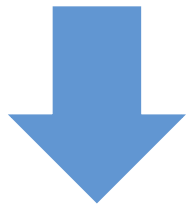


- 研究開発用DCの見える化

# データ収集の目的

---

- 空調制御効率化 ← 今回はこちらに注力!!
- 実験時のデータ観測
- 実験機器の監視



- 研究開発用DCの見える化

# 空調制御効率化に必要なデータ

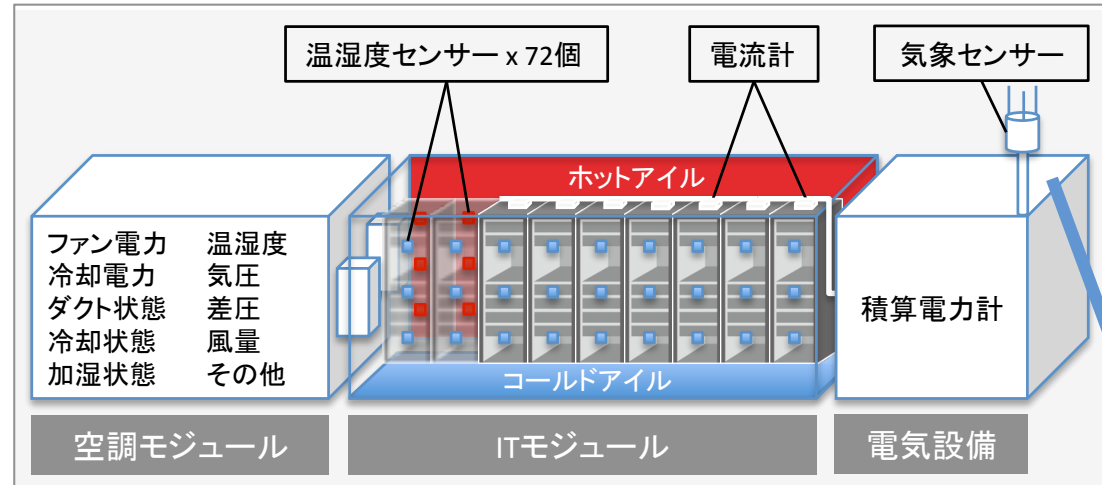
- 外気導入冷却では周辺環境の変動で必要なエネルギーが変動する仕組み



- 空調制御装置のデータ
- 外気温度／湿度などの気象データ
- 電気設備のデータ
- ITモジュール(コンテナ)内のデータ

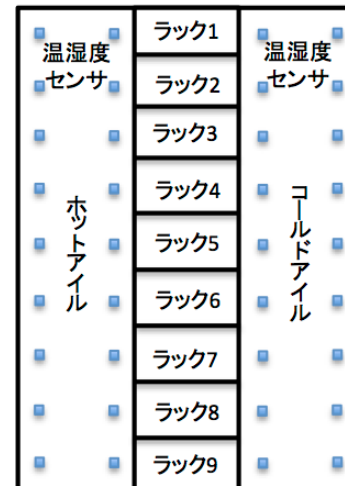
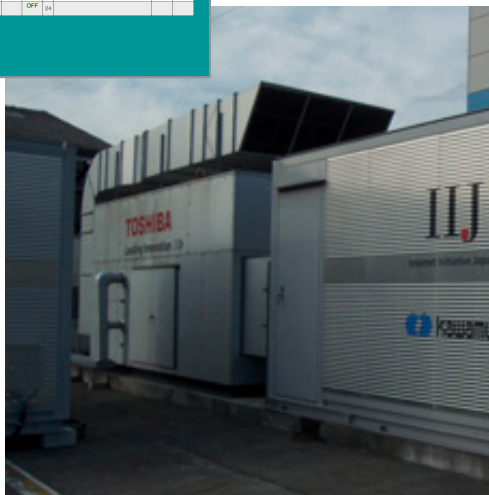


# 収集するデータ



コンテナサーバ 空調機

設備名	状態	設定	動作	故障
1. 空調機	ON	10	稼働	OFF
2. 空調機	OFF	11	停止	ON
3. 空調機	ON	12	稼働	OFF
4. 空調機	ON	13	稼働	OFF
5. 空調機	OFF	14	停止	OFF
6. 空調機	OFF	15	停止	OFF
7. 空調機	ON	16	稼働	OFF
8. 空調機	ON	17	稼働	OFF
9. 空調機	ON	18	稼働	OFF
10. 空調機	ON	19	稼働	OFF
11. 空調機	ON	20	稼働	OFF
12. 空調機	ON	21	稼働	OFF
13. 空調機	ON	22	稼働	OFF
14. 空調機	ON	23	稼働	OFF
15. 空調機	ON	24	稼働	OFF



# 商用のDC用センサーネットワーク

- もちろん売り物もたくさんあります
- でも高い
  - 某社さんでDC用は2000万とか...
- データは早めに取得したい
  - アドホックに作りたい
    - 作って壊される設備なので柔軟に
  - 自作！！

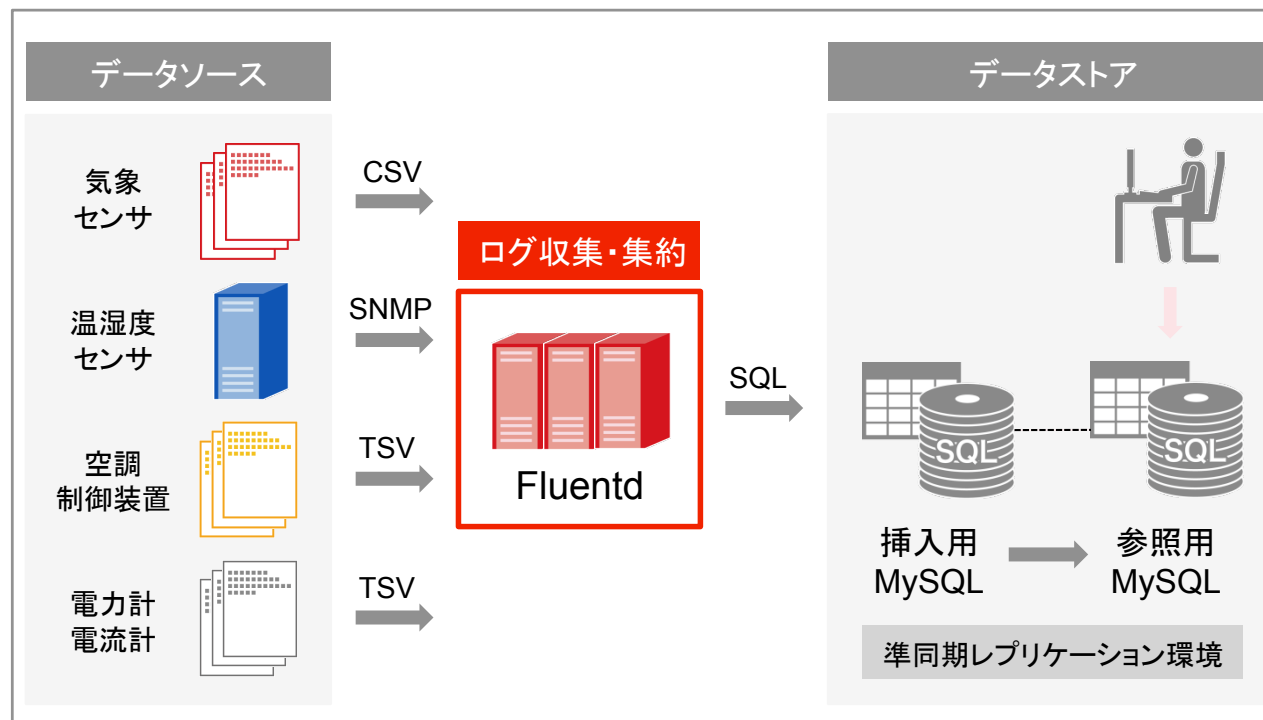


# 取得する情報

収集対象	取得時間	1日の情報量	収集間隔
気象センサー	外気温度、外気湿度、降雨量、 降雹量、風速、風圧、気圧	約4万件	30秒
温湿度センサー	温度、湿度(全72カ所)	約62万件	10秒
空調制御装置	温度、湿度、ダクト状態、ファン 回転数、ファン電力、気圧、差圧 風力、加湿状態、冷却状態、冷 却電力	約88万件	5秒
電力計	基幹積載電力計及び内部の電 力計毎の電流値	約6千件	60秒

# データの収集方法／蓄積方法

- Fluentd
- データ収集プログラム
- データ蓄積 (MySQL)

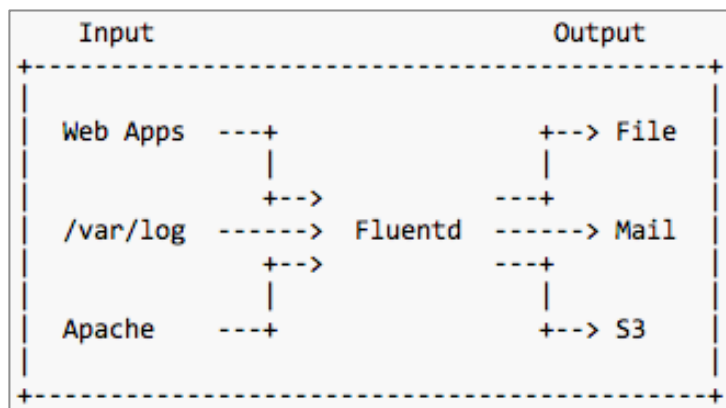


# Fluentd

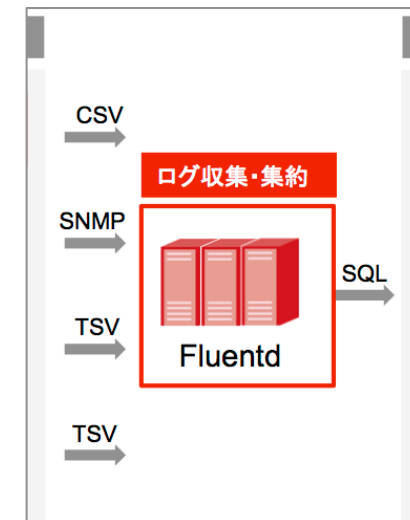
- <http://fluentd.org/>
  - Rubyで記述されたデーモンプログラム
  - メッセージはJSON形式で扱われる
  - 大規模なログ扱う処理系として国内でメジャー
  - 各種プラグインが開発されている
    - input, outputの両方のプラグインが充実



Fluentdアーキテクチャ

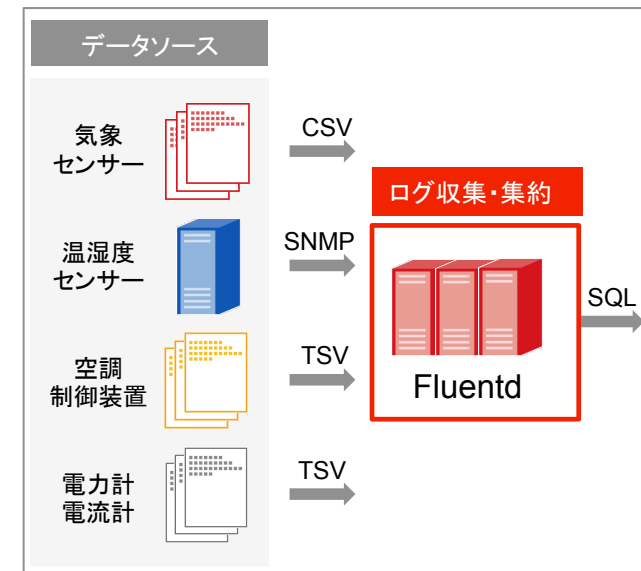


本件での使い方



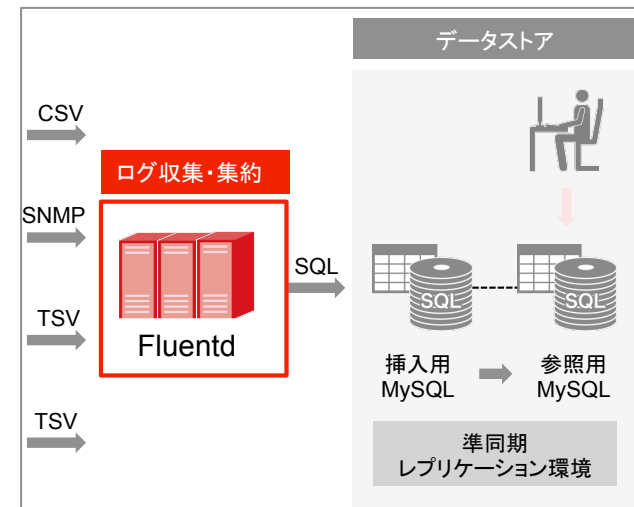
# データ収集プログラム

- 気象センサーデータ
  - 別システムでCSVファイルへ出力
    - tailプラグインを利用
- 温湿度センサーデータ
  - SNMPでポーリング
    - SNMPポーリング用プラグインを自作
- 空調制御装置、電力計データ
  - 別システムでTSVファイルに出力
    - tailプラグインを利用



# データ蓄積

- 蓄積対象はMySQL
  - FluentdからSQLのインサートへ変換
  - 問題
    - 各データはスキーマが違う
    - 1ヶ月単位で違うテーブルにデータを蓄積
    - これから収集対象が増えて行く



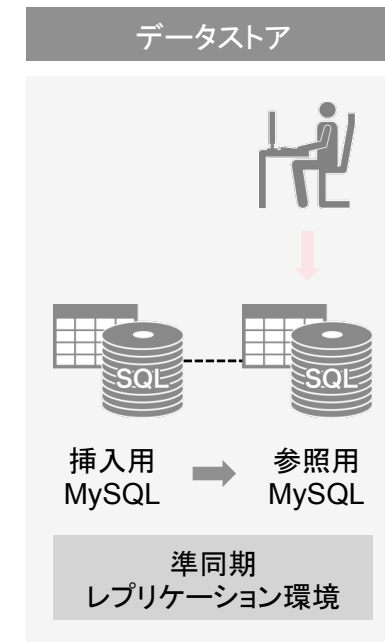
# データの参照と解析

- 参照用MySQL

- 解析用に重いSQLを走らせたい
- Insert処理に邪魔されたくない

- データの解析

- 相関を取るのはいくらから...
- 電気通信大学の市川研究室、川喜田研究室と共同研究を行っています
  - <http://www.ichikawa.hc.uec.ac.jp/pukiwiki/>
  - <http://www.kwkt.inf.uec.ac.jp/>





# 空調制御効率化以外のデータ収集

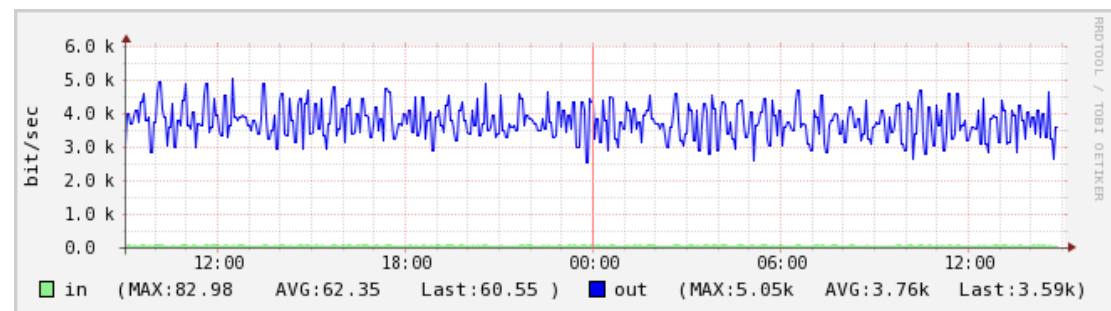
- サーバ
  - SNMP、IPMI
- ネットワーク機器
  - SNMP
- 気象センサー
  - 温度、湿度
- UPS
  - バッテリー残量

# 空調制御効率化以外のデータ収集

- サーバ
  - SNMP、IPMI
- ネットワーク機器
  - SNMP
- 気象センサー
  - 温度、湿度
- UPS
  - バッテリー残量

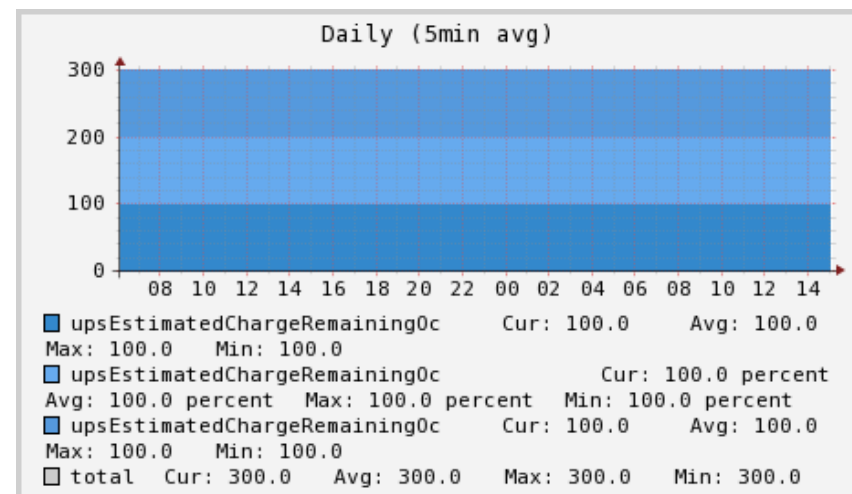
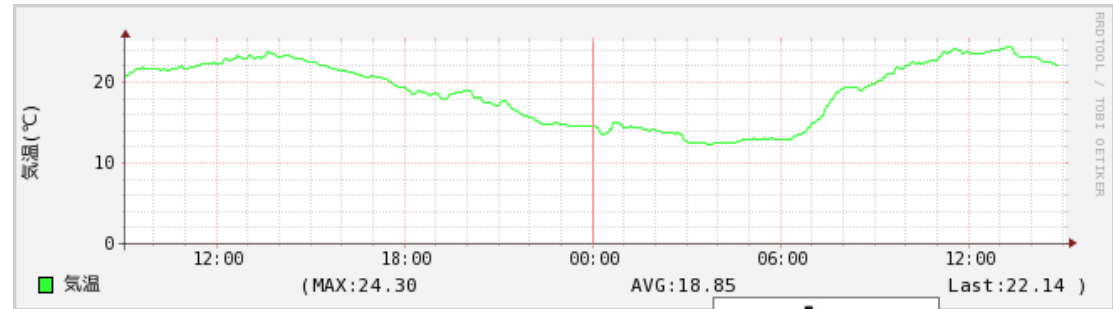


Status	Interface	Description	IP Address
	<a href="#">ALL</a>		
	<a href="#">eth0</a>		192.168.0.30/24
	<a href="#">eth0.100</a>		192.168.100.25/24
AdminDown	eth1		
	<a href="#">lo</a>		127.0.0.1/8
<b>Storage Used</b>			
	<a href="#">Physical memory</a>		1.79839G / 1.84822G
	<a href="#">Virtual memory</a>		1.85936G / 3.84822G
	<a href="#">Memory buffers</a>		165.891M / 1.84822G
	<a href="#">Cached memory</a>		1.03503G / 1.03503G
	<a href="#">Swap space</a>		62.4336M / 1.99999G
	<a href="#">/</a>		17.4695G / 29.5293G
	<a href="#">/var/service/asama</a>		32.3237G / 115.171G
<b>Users</b>			
	<a href="#">users</a>		0
<b>Load Average</b>			
	<a href="#">Load-1</a>		2.12
	<a href="#">Load-5</a>		2.13
	<a href="#">Load-15</a>		2.15
<b>Memory Pool</b>			
	<a href="#">memTotalFree</a>		1.98885G
	<a href="#">memAvailReal</a>		51.0273M



# 空調制御効率化以外のデータ収集

- サーバ
  - SNMP、IPMI
- ネットワーク機器
  - SNMP
- 気象センサー
  - 温度、湿度
- UPS
  - バッテリー残量



# 監視

- Nagios
  - サービスホストを監視
    - 無茶をする実験も一応監視
  - 設定プロビジョニングとの連携は開発中

**Nagios**

**Current Network Status**  
 Last Updated: Wed Oct 17 15:07:13 JST 2012  
 Updated every 30 seconds  
 Nagios® Core™ 3.2.3 - www.nagios.org  
 Logged in as nagiosadmin

**Host Status Totals**

Up	Down	Unreachable	Pending
42	40	0	0

**Service Status Totals**

Ok	Warning	Unknown	Critical	Pending
6	0	0	0	0

**Host Status Details For All Host Groups**

Host	Status	Last Check	Duration	Status Information
8101000	UP	10-17-2012 15:04:22	55d 18h 1m 55s	PING OK - Packet loss = 0%, RTA = 0.19 ms
8101001	UP	10-17-2012 15:03:02	55d 18h 18m 45s	PING OK - Packet loss = 0%, RTA = 0.20 ms
8101002	UP	10-17-2012 15:02:52	55d 17h 47m 34s	PING OK - Packet loss = 0%, RTA = 0.19 ms
8101003	UP	10-17-2012 15:02:42	55d 17h 42m 24s	PING OK - Packet loss = 0%, RTA = 0.18 ms
8101004	UP	10-17-2012 15:02:52	55d 17h 37m 14s	PING OK - Packet loss = 0%, RTA = 0.22 ms
8101005	UP	10-17-2012 15:02:52	55d 17h 36m 44s	PING OK - Packet loss = 0%, RTA = 0.21 ms
8101006	DOWN	10-17-2012 15:02:52	112d 13h 2m 41s	CRITICAL - Host Unreachable (172.16.0.7)
8101007	UP	10-17-2012 15:02:52	40d 3h 49m 50s	PING OK - Packet loss = 0%, RTA = 0.76 ms
8101008	UP	10-17-2012 15:02:42	60d 1h 15m 26s	PING OK - Packet loss = 0%, RTA = 0.42 ms
8101009	UP	10-17-2012 15:02:52	60d 1h 15m 26s	PING OK - Packet loss = 0%, RTA = 0.44 ms
8101011	UP	10-17-2012 15:04:02	60d 1h 15m 26s	PING OK - Packet loss = 0%, RTA = 0.18 ms
8101020	DOWN	10-17-2012 15:02:02	138d 7h 23m 27s	CRITICAL - Host Unreachable (172.16.0.100)



# まとめ

---

- 研究開発用DCのデータ収集を行っています
- データ収集の為にFluentdを使ってみました
- 研究開発用DCの見える化を進めます
- Webに記事を書いています
  - [http://www.ij.ad.jp/company/development/tech/activities/dc\\_data/index.html](http://www.ij.ad.jp/company/development/tech/activities/dc_data/index.html)

## ご清聴ありがとうございました

お問い合わせ先 IIJインフォメーションセンター  
TEL: 03-5205-4466 (9:30~17:30 土/日/祝日除く)  
info@ij.ad.jp  
<http://www.ij.ad.jp/>

Ongoing Innovation

本書には、株式会社インターネットイニシアティブに権利の帰属する秘密情報が含まれています。本書の著作権は、当社に帰属し、日本の著作権法及び国際条約により保護されており、著作権者の事前の書面による許諾がなければ、複製・翻案・公衆送信等できません。IIJ、Internet Initiative Japanは、株式会社インターネットイニシアティブの商標または登録商標です。その他、本書に掲載されている商品名、会社名等は各会社の商号、商標または登録商標です。本文中では™、®マークは表示していません。©2012 Internet Initiative Japan Inc. All rights reserved. 本サービスの仕様、及び本書に記載されている事柄は、将来予告なしに変更することがあります。



## インターネットの先にいます。

IIJはこれまで、日本のインターネットはどうあるべきかを考え、  
つねに先駆者として、インターネットの可能性を切り拓いてきました。  
インターネットの未来を想い、イノベーションに挑戦し続けることで、世界を塗り変えていく。  
それは、これからも変わることのない姿勢です。  
IIJの真ん中のIIはイニシアティブ ————— IIJはいつもはじまりであり、未来です。

Ongoing Innovation

お問い合わせ先 IIJインフォメーションセンター  
TEL: 03-5205-4466 (9:30~17:30 土/日/祝日除く)  
info@ij.ad.jp  
<http://www.ij.ad.jp/>

本書には、株式会社インターネットイニシアティブに権利の帰属する秘密情報が含まれています。本書の著作権は、当社に帰属し、日本の著作権法及び国際条約により保護されており、著作権者の事前の書面による許諾がなければ、複製・翻案・公衆送信等できません。IIJ、Internet Initiative Japanは、株式会社インターネットイニシアティブの商標または登録商標です。その他、本書に掲載されている商品名、会社名等は各会社の商号、商標または登録商標です。本文中では™、@マークは表示していません。

©2012 Internet Initiative Japan Inc. All rights reserved. 本サービスの仕様、及び本書に記載されている事柄は、将来予告なしに変更することがあります。