IIJ INNOVATION INSTITUTE

IIJ Technical Week 2011

データセンターテストベッドの構築と展開

IIJ Innovation Institute Inc. IIJ技術研究所

宇夫陽次朗 <yuo@iijlab.net>

本セッションの概略

プログラムより抜粋

IIJグループで運用されている「外気冷却方式コンテナ型データセンター」の更なる機能および性能向上を目指した「研究開発用テストベッド」と今後の展望について紹介します。

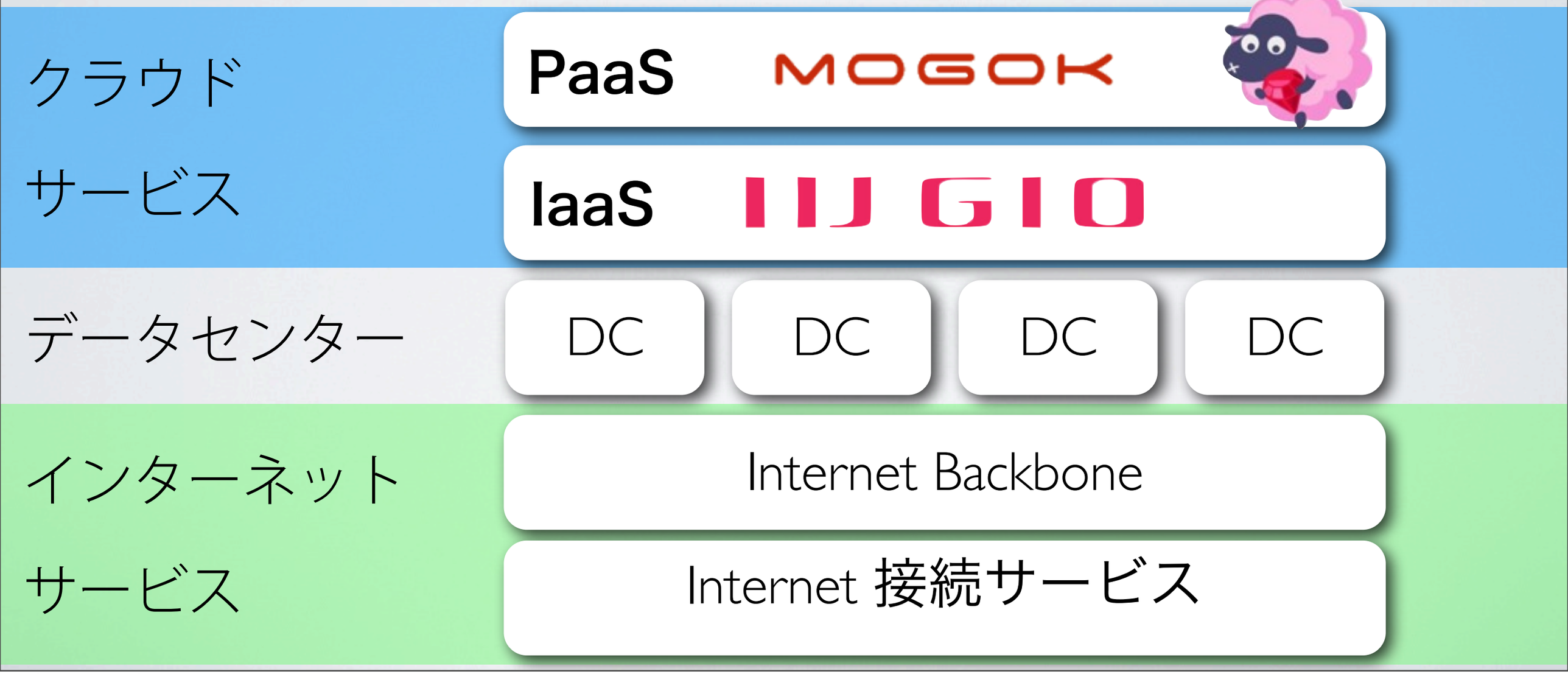
- IIJのデータセンター（特にクラウドプラットフォーム）関連研究開発を紹介します。

コンテンツ

- セッションの流れ
  - データセンター&クラウド関連研究開発の目的
  - コンテナ型データセンターテストベッド
  - 今後の展望とまとめ

Internet, Data Center, Cloud

- IIJグループの立ち位置
  - 上（Cloud)から下（インターネット接続）まで全部提供



IIJのデータセンター方向性

- 効率の良いデータセンターファシリティの実現へ

コンテナ型
データセンターモジュール

＋

外気導入冷却機構

→ 必要な分だけ設備を構築可能なモジュラリティ

→ 高いエネルギー効率のデータセンター冷却の実現

周辺環境と運用効率

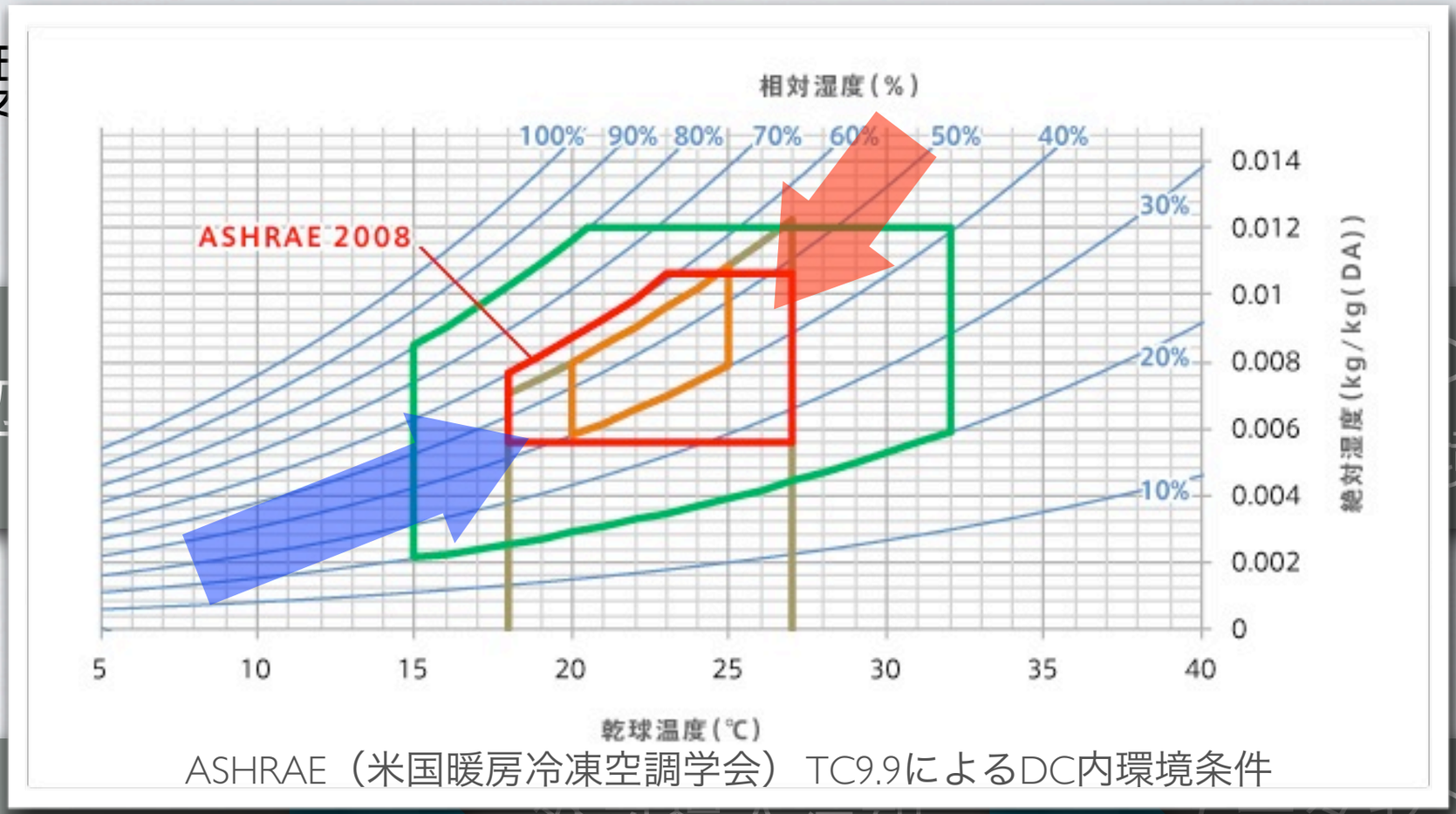
- 外気導入冷却はデータセンターの周辺環境の変動で必要なエネルギーが変動する仕組み

周りが暑いとき・寒いときはエネルギーを使って環境を改質

周りが適温なら全然エネルギーを使わないで良い

周辺環境と運用効率

- 外気導入冷却はデータセンターの周辺環境の変動で必要



ASHRAE（米国暖房冷凍空調学会）TC9.9によるDC内環境条件

周辺環境 → 外気導入冷却機構 → データセンター内環境

周りが適温なら全然エネルギーを使わないで良い

効率の良いデータセンターって？

- あたらしく導入されるパラメータがいっぱいある
  - 外気導入冷却　→　周辺の環境（空気）の変動
  - 外気導入冷却　→　設置環境のプロファイル
  - クラウド的な仕組み　→　システム負荷の分布
  - クラウド的な仕組み　→　システムサイズの変動
  - ・・・

研究開発のターゲット （1）

- データセンター≠単なるファシリティ
  - 「設備としての効率化」
    から
    「上位システムを含んだ最適化」へ

研究開発のターゲット（2）

- どうすれば、あたらしく導入されたパラメータをうまく扱って効率の良い仕組みを実現できるのか
- 商用施設と研究開発施設間でのノウハウの相互フィードバック

クンフー（経験）を積め！

- 「やってみないとわからないこと」がたくさん

知りたいこと／知るべきこと（1）

- 物理環境とデータセンターとサーバー
  - 日本の四季の変動
  - 日本の北と南、東と西、どこがサーバー運用に優しいか
  - 晴れと曇り、夕立と梅雨
  - 朝と昼と夜
  - 各地方の四季折々

知りたいこと／知るべきこと（2）

- 空気調和システム（空調）とデータセンター
  - 温度、湿度、気圧の「経路制御」
  - 空気の流れの「経路制御」
  - それぞれの操作と実現するまでにかかる時間
  - それぞれの操作と必要エネルギー

知りたいこと／知るべきこと（3）

- 単体ITサーバーの動きとデータセンター
  - 「仕様：XX度まで動きます」けれど何度で動かすのが電力効率最良か？
  - 各サーバーの冷却ファン動作プロファイルとDC内の空気の流れの関連性
  - サーバー機能の「一時停止」のコストと時間

知りたいこと／知るべきこと（4）

- データセンターネットワーク(DCN)と効率
  - ネットワークの規模とエネルギー効率の関連性
  - スイッチアーキテクチャとエネルギー効率
  - ネットワークトポロジとエネルギー効率
  - 通信／データ転送アルゴリズムとエネルギー効率
  - 通信クオリティ（レイテンシ／帯域）とエネルギー効率

知りたいこと／知るべきこと（5）

- システム全体の資源管理／単体と群体
  - 同じ{容量、品質}のサービスを提供するためのロードの分布とシステム全体のエネルギーの相関
  - ITサーバーの「エネルギー効率の良い」ロードパターン
  - システム全体の「エネルギー効率の良い」ロード配置

いろいろ知って実現したいことは？

- ある種の「オペレーティングシステム」for「Data Center as a Computer」
  - データセンター内の資源管理、資源配置
  - データセンター内のステートの管理
  - 様々な周辺環境や運用形態への適応性の提供
  - データセンターの抽象化とAPIの提供

どんな道具が必要か？

- 「実機」＝「研究開発用のデータセンター一式」
  - 研究対象としている領域全体を含んだ仕組みが必要
  - 冷却＋ファシリティ＋ITサーバー＋ネットワーク+計測系

クラウドデータセンターテストベッド

- データセンターテストベッド
  - 2010年度：日本での外気導入冷却の可能性の検証
  - 2011年度：データセンターファシリティの効率化
  - 2012年度：クラウドテストベッドの構築

テストベッドの構成

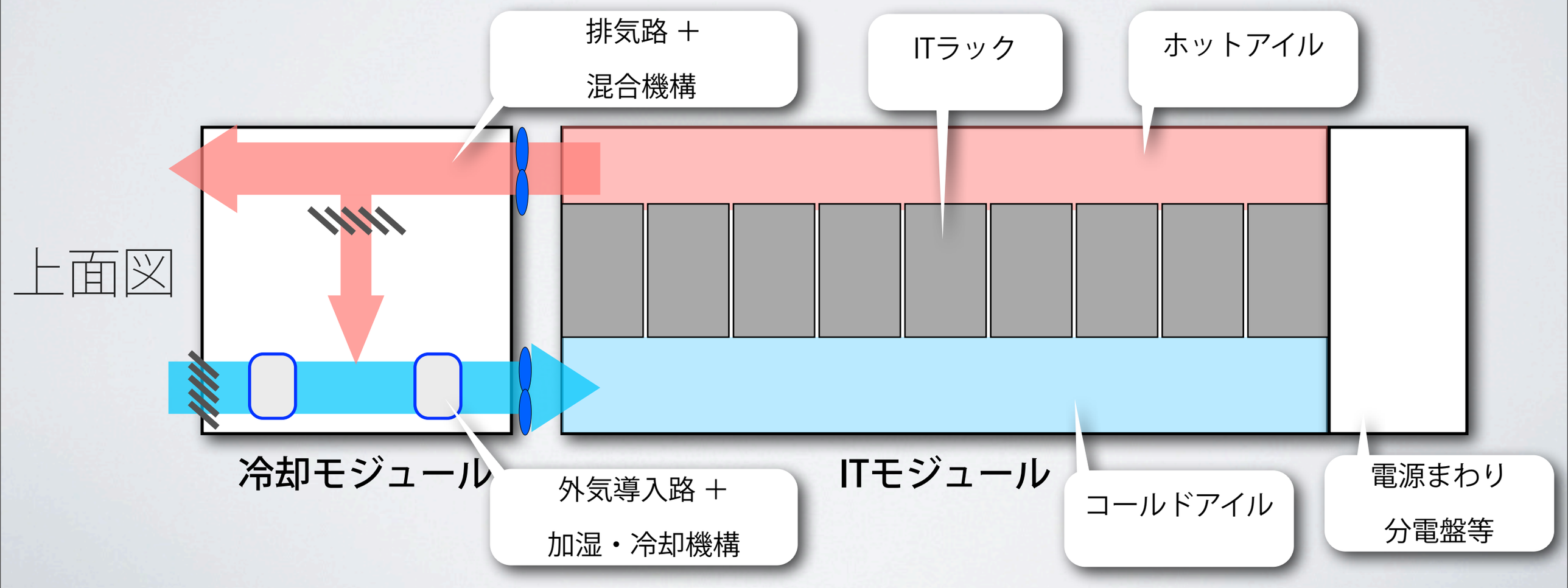
- 物理的ファシリティ
- センサー
- ITサーバー
- ネットワーク
- （ソフトウェア）

テストベッド | 物理ファシリティ(1)

- コンテナデータセンターのファシリティ

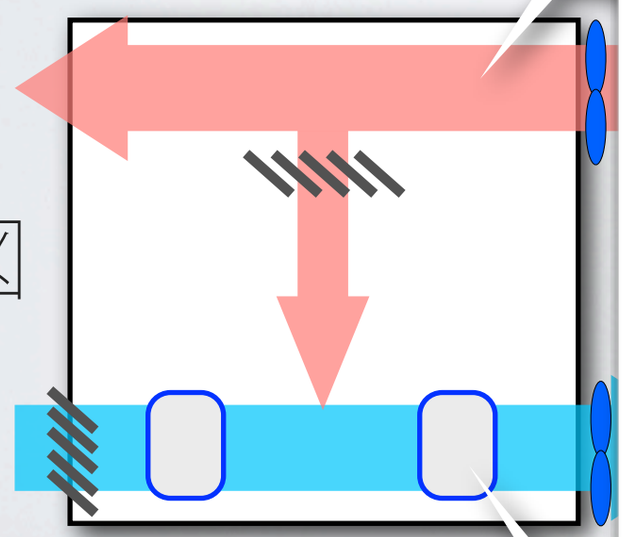
テストベッド|物理ファシリティ(2)

- 「外気導入冷却」＋「モジュール型コンテナデータセンタ」の占有実験設備として構築



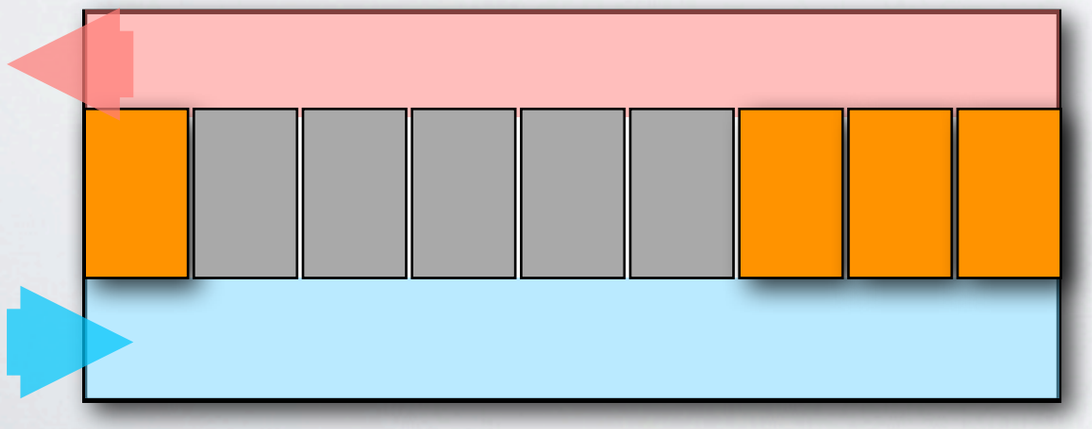
テストベッド|物理ファシリティ(2)

- 「外気導入冷却」＋「モジュール型コンテナデータセンタ」の占



テストベッド|物理ファシリティ(3)

- データセンター負荷のエミュレーション
  - 一部あえて電力効率の悪い古いサーバー設置して「熱源サーバ」として制御

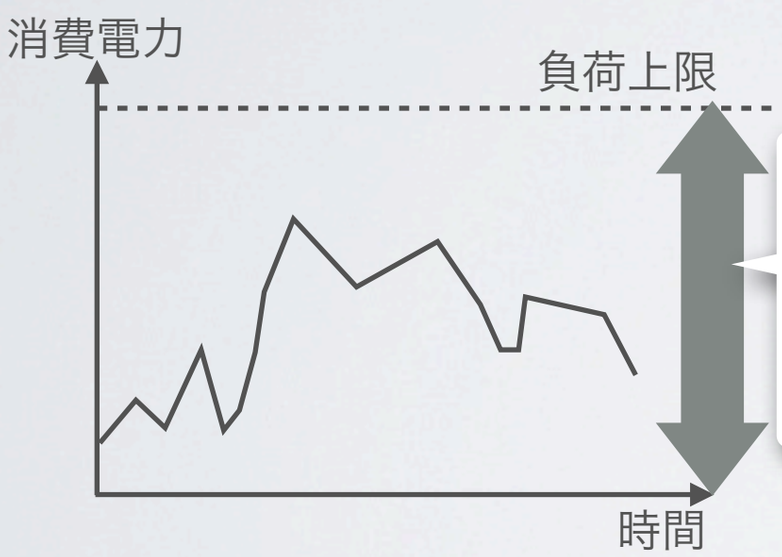


熱源サーバーラック

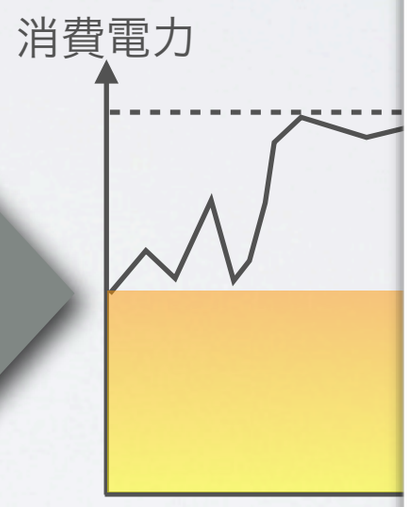
0-90KVA負荷

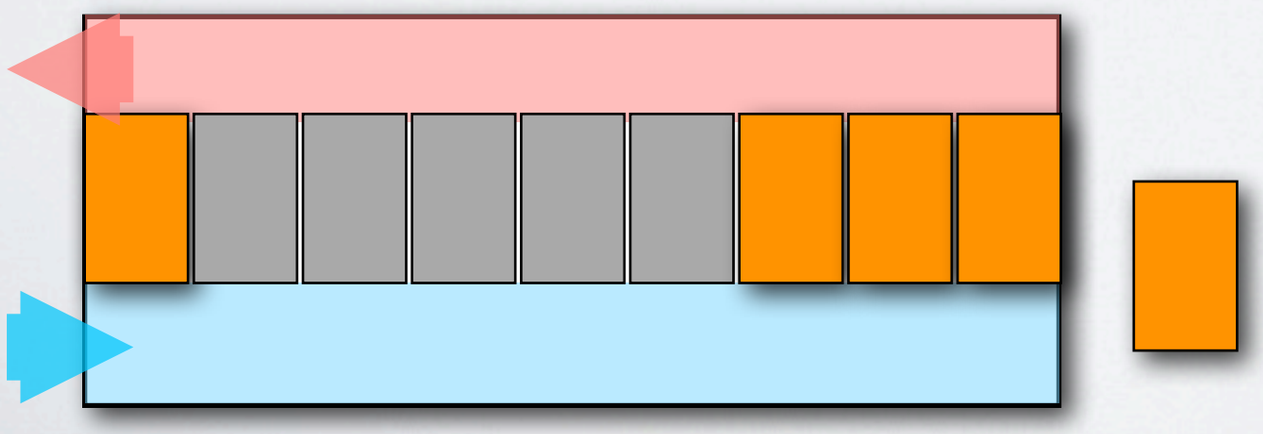
テストベッド | 物理ファシリティ(3)

- データセンター負荷のエミュレーション
  - 一部あえて電力効率の悪い古いサーバー設置して「熱源サーバ」として制御



負荷上限が大きい為、実システムだけで高負荷状態を作るの面倒





熱

0-9

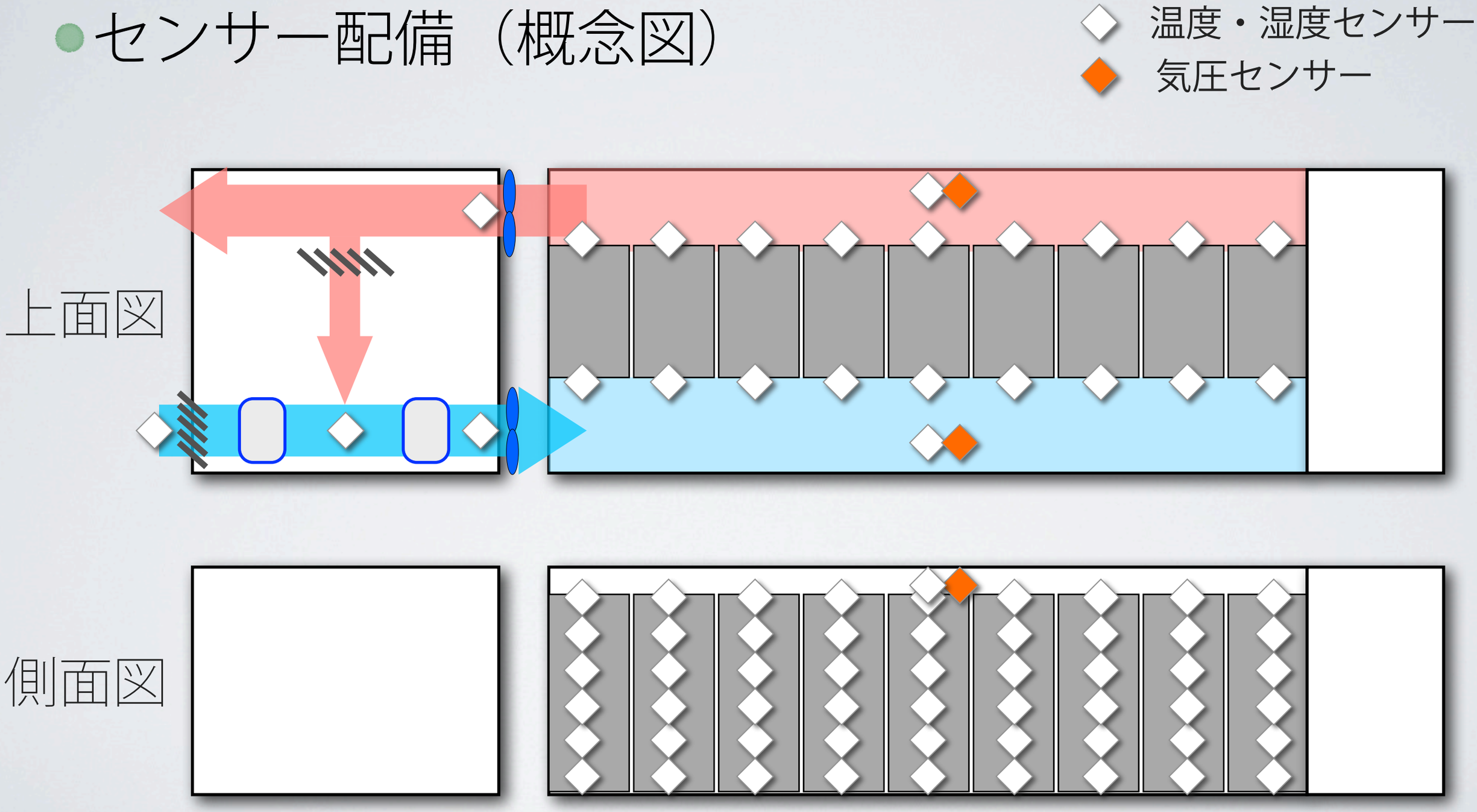
テストベッド|センシング（1）

- テストベッドに関するセンシング階層
  - データへのアクセスAPIを整備

センシングレイヤリング	収集情報	利用方法
IT機器(OS/アプリケーション)	今回は対象外（上位層で動くシステムが独自に収集、利用）。標準的なHaaS/IaaS statistics APIなどの検討は継続的に行う	サービス構築者が収集
IT機器(IPMI)	各CPU温度、ファン回転数、電源電圧および電流、全体消費電力、その他IPMI2.0で取得できる情報＋ベンダー拡張情報	機器所有者が収集
module DCファシリティ	全ラック:ラックあたり8系統毎の電力および前面3,後面3の温度湿度 部分的：コンセント単位の電力計測(分解能１W) ホットアイル・コールドアイルの気圧および環境温度湿度	コンテナDCの動作状況をまとめてアクセスAPIを提供予定
空調設備	空調制御装置経由から取得可能な、ファン風量、外気気温、湿度、動作モード（混合比、加湿機、コンプレッサの動作その他）	
外部環境センシング	複合気象センサ: 実験設備設置場所の温度、湿度、風向、風速、降雨量、気圧	

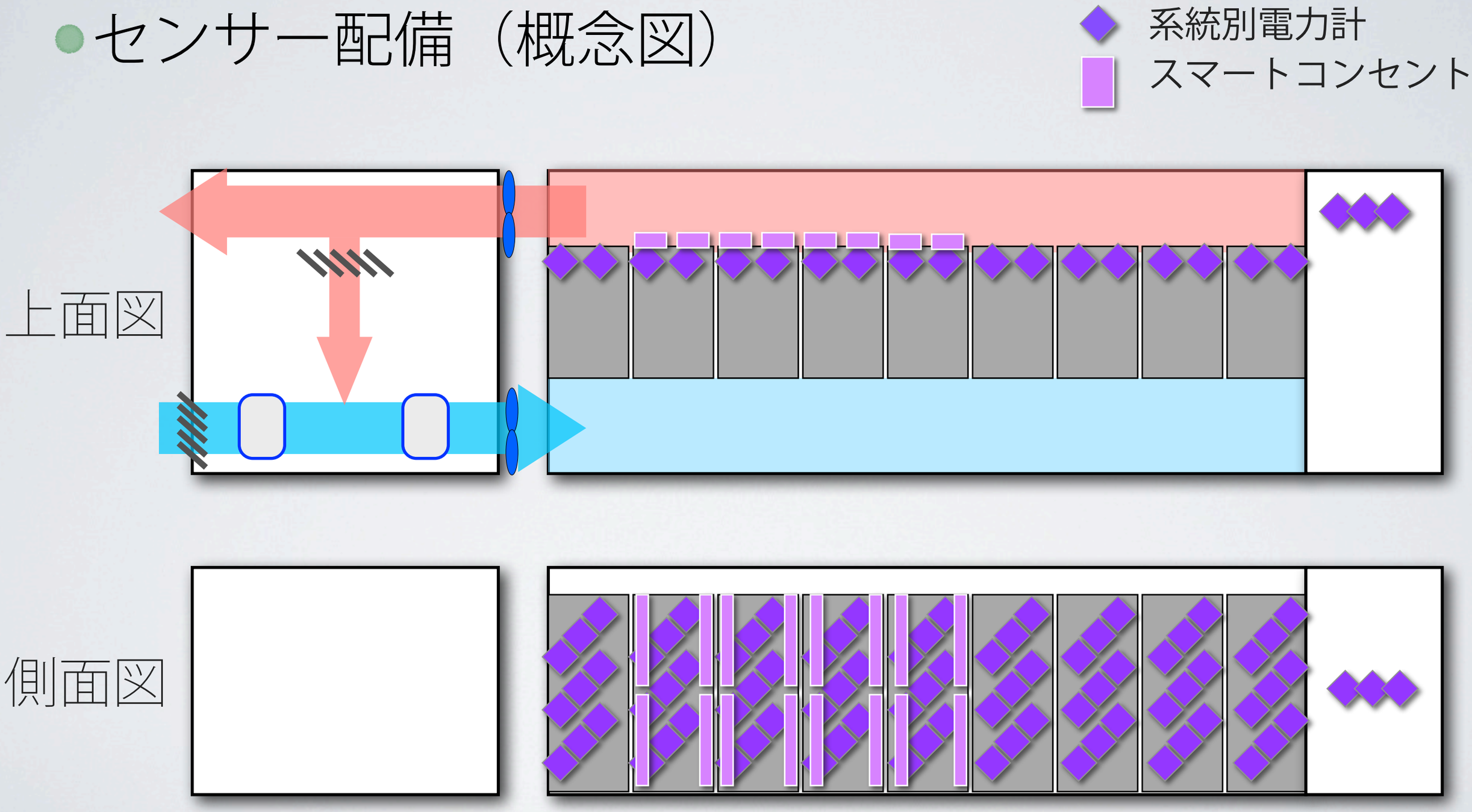
テストベッド|センシング（2）

- センサー配備（概念図）



テストベッド|センシング（3）

- センサー配備（概念図）

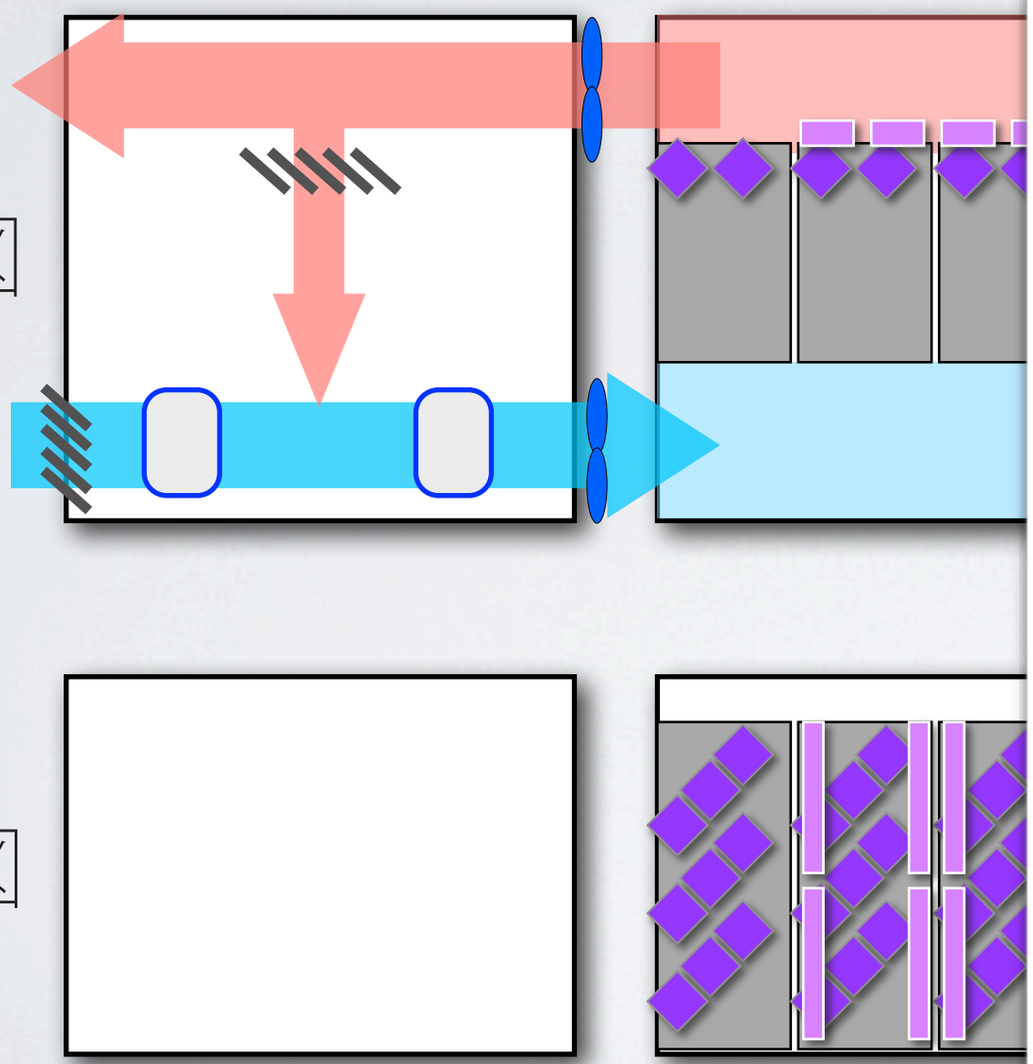


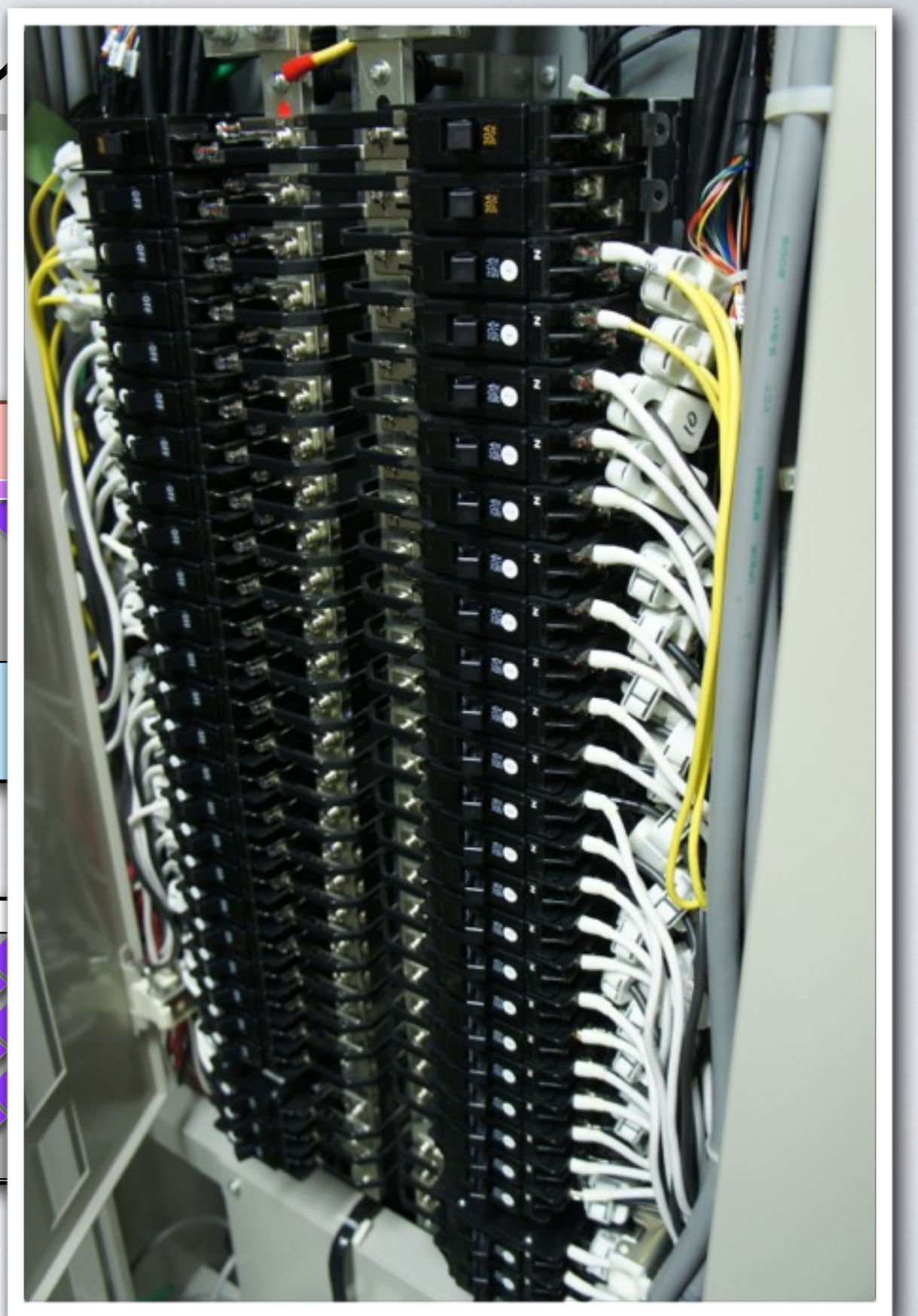
テストベッド | セン

- センサー配備（概念図）

上面図

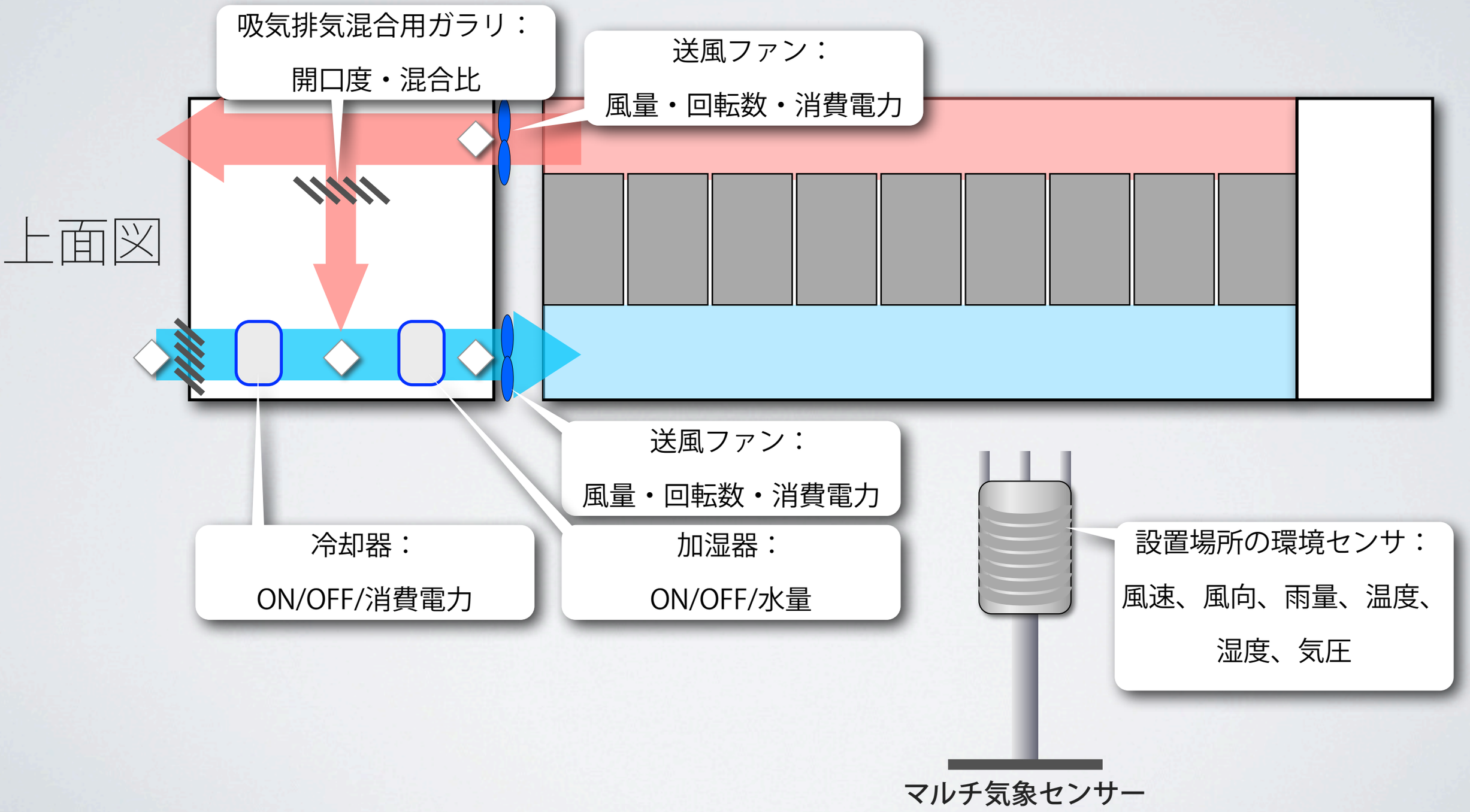
側面図





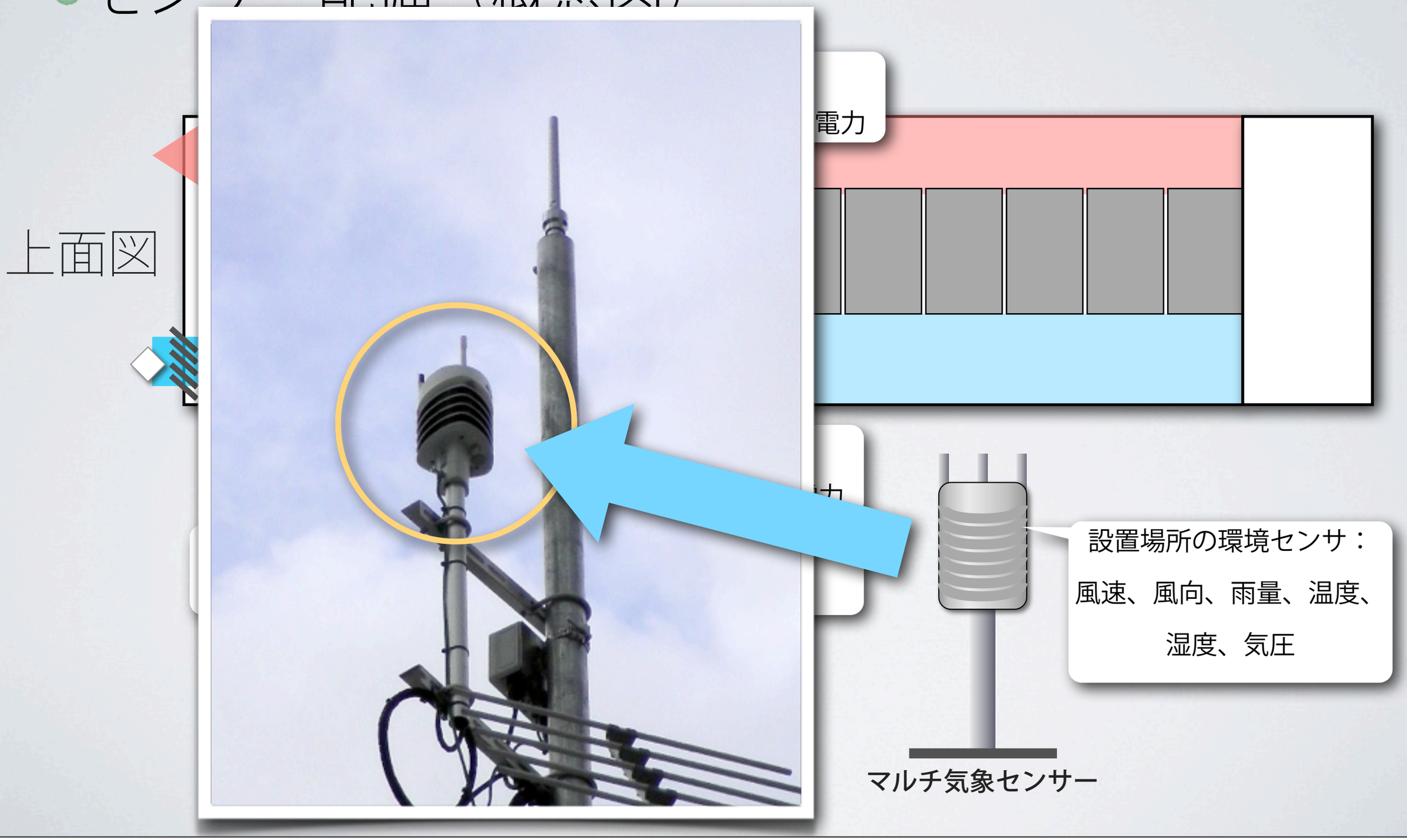
テストベッド|センシング（4）

- センサー配備（概念図）



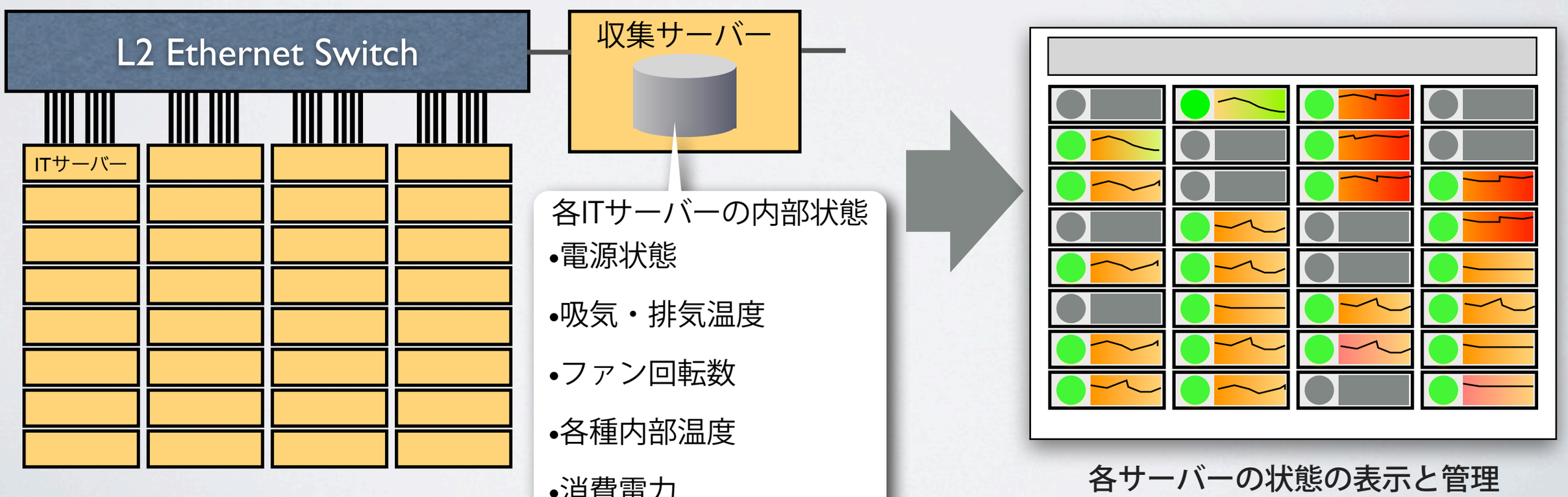
テストベッド｜センシング（4）

- センサー配備（概念図）



テストベッド | センシング（5）

- IT機器の内部センシング情報の収集と利用
  - IPMI(Intelligent Platform Management Interface) over LAN 経由で各ITサーバー内センサー情報を取得、集約、提示

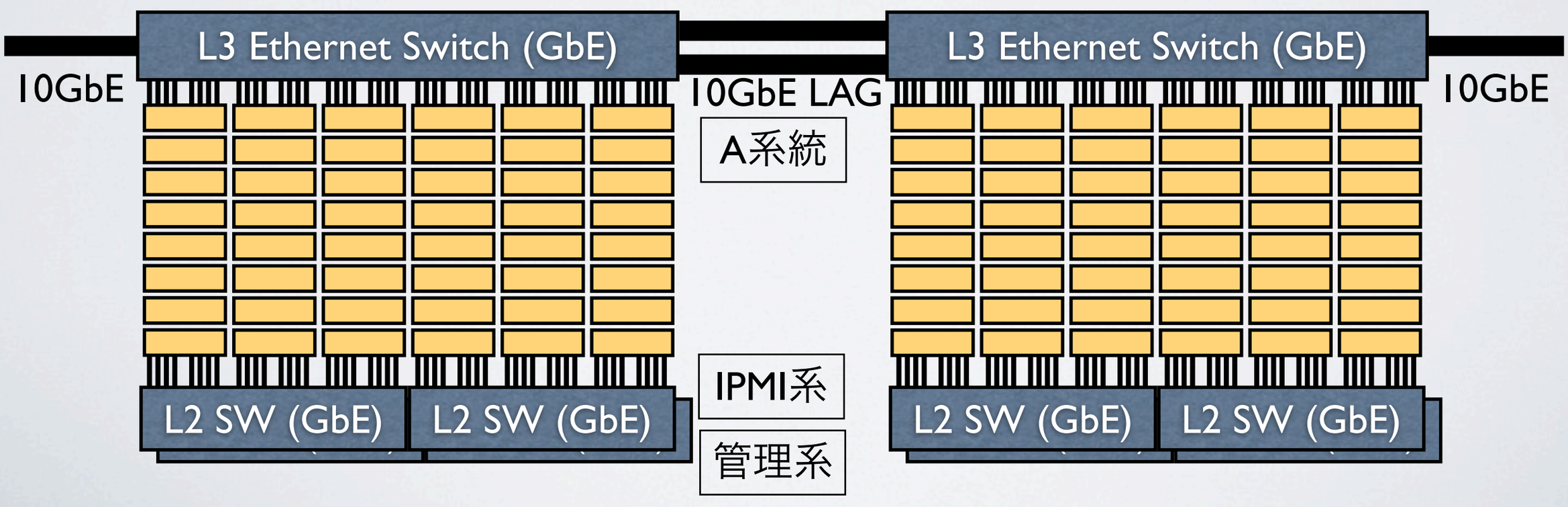


テストベッド|ソフトウェア環境

- 2011年度テストベッドは「中身」入り
  - 現行モデルのITサーバー：100台弱
  - 負荷生成用ITサーバー：50台
  - 熱源負荷用ITサーバー：100台
  - GbE - 10GbEクラスの内部ネットワーク

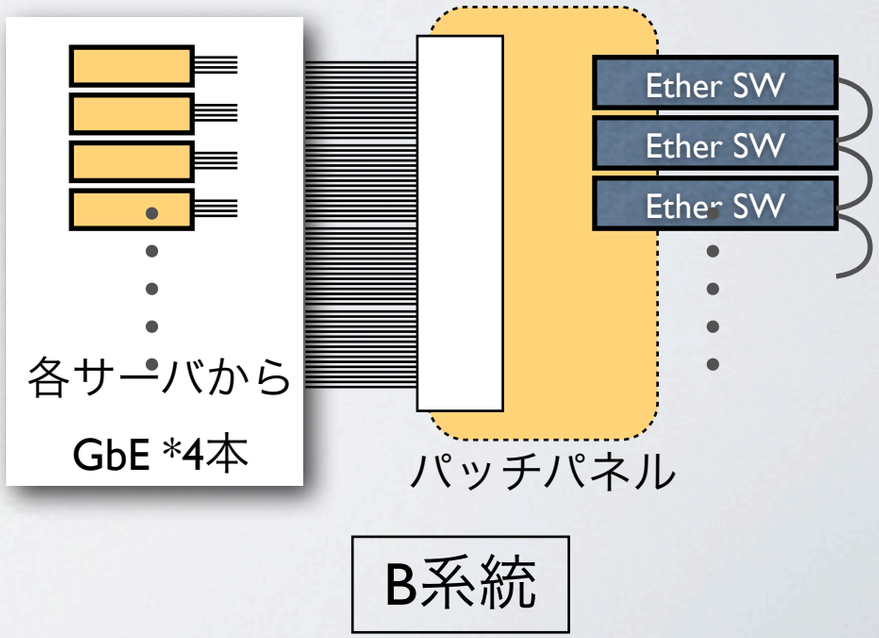
テストベッド | ITサーバー

- HaaS / IaaS / PaaSレイヤの構築
  - 普通のCloudサービス用機材と同等の設備を導入
  - 実験設備の規模：2.4GHz CPU 1000core, 主記憶 3TB, ストレージ 40TB, ネットワーク GbE * 6 + 10GbE * 2 + IPMI



テストベッド|ネットワーク

- 普通のCloudサービス用のネットワーク
  - 各サーバからGbEでスター型＋管理用＋IPMI
- 様々なネットワークトポロジを実現
  - 「データセンターネットワーク」研究・検証・計測用の物理的組み替え可能＋マルチインターフェイス環境
- マルチインターフェイス＋マルチパスネットワーク



こんな設備でできること(1)

- いろいろな要素間の相関の計測
- データセンター＆Cloudの新技術の検証
- CPUやメモリを用いた計算

できること | 相関分析→効率化

- システム全体の電力と各要素の相関
  - 周辺環境の変動
  - 上位システムのアルゴリズム・挙動・負荷変動
  - 内部ネットワークの構成・挙動・制御
  - 各ITサーバーの個別制御ロジック
  - 空調システムの制御ロジック
- 計測した相関に基づくパラメータ調整と効率化

できること｜新技術の検証

- 「新しいHaaS / IaaS / PaaS」のplayground
- 同一環境での検証による性能／挙動比較の容易化
- 「新しいデータセンターネットワーク技術」の検証・計測・運用可能性の検討
- 新しい「ハードウェア」と本システム親和性検証

できること|計算

- 「必要最低限のクラウド構築用環境」でも一人で使うには過剰な計算能力が提供できる
  - ログ解析などの大規模計算などに利用可能
  - 様々なアプリケーション利用時の稼働データを収集してシステム挙動の解析に利用

テストベッドの今後

- 規模の拡大による検証可能サイズの拡張
  - 多地点化（分散データセンターモデルの検証）
  - ネットワークの高容量化
  - サーバー数の拡張
- センシング密度の拡張による高精度化
  - センシングモデルの構築とリアルタイム計測の実現

本セッションのまとめ

- データセンター＆クラウドの研究開発に関するIIJの取り組みの一部を紹介しました
- 今後もデータセンター＆クラウドの効率化や高度化に向けた様々な研究開発に取り組んでいきます