JANOG56 NOCツアー
BBチーム

BB | 目次

- チーム紹介
- BBチームの概要
- 構成説明
  - 全体構成
  - 会場～堂島構成
- ピアリングの告知と紹介
- 検証環境
- 監視

BB | チーム紹介

- 社会人 4名
- 学生　3名

BBチームの概要

- 基本構築
  - IIJさんとピアを行い、トランジットを受信しフルルート頂く
  - 3IX（JPNAP・BBIX・JPIX）で接続し各事業者とピアリング
  - 会場 ～ 堂島間でインターネット越しにVPN接続

- アドバンスド構築
  - 迂回経路切り替えにBFDを導入
  - RPKIを使用した経路制御 (不正経路の遮断など)

オペレーションのやり方やどういう風に活動したか

- オペレーション
  - チームリーダがタスクを細分化
  - 1タスクに社会人1名、学生1名を割り当てる
  - 基本的に学生が主担当、社会人が副担当

- 取り組み方
  - タスクに関する手順書を作業当日までに作成
    - 学生が手順書を作成し、社会人の方にフィードバックもらいながら完成させる
  - 作業当日は、主担当が手順書を見ながら設定を入れ、副担当はサポートと確認
    - オンラインや現地で画面を共有しながらの作業

検証環境と検証構成

- containerlabを用いて本番と同じ検証環境を構築
  - 検証環境VMのスペック
    - CPU: 12vCPU (virtual CPU)
    - メモリ 50GiB
    - ストレージ: 256GB
  - 直面した課題
    - フルルート読み込ませるとメモリがいっぱいで経路学習ができなかった
      - 仮想マシンのリソースはそこそこあるのに
    - 仮想マシンと実機でコマンドの差異があった
    - 定期的に仮想ルータに入れなかった
      - 共有環境であるため

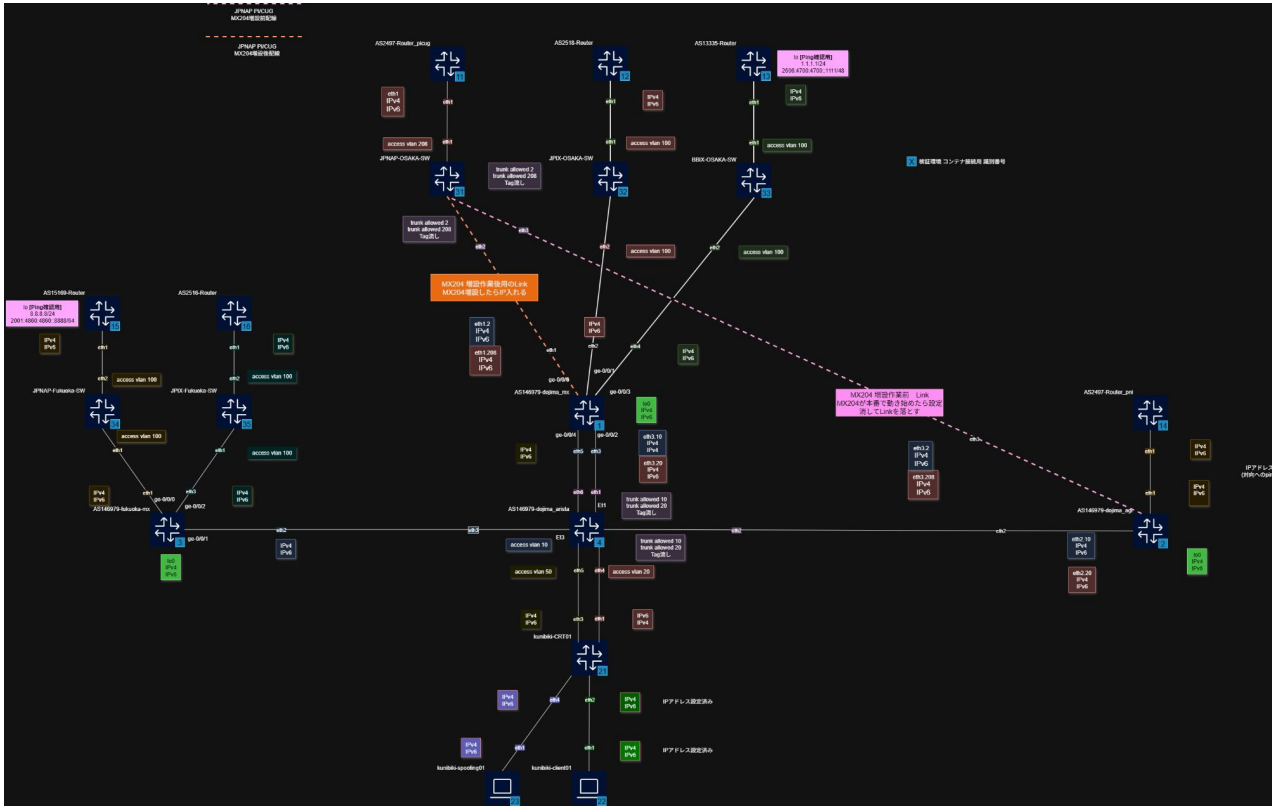
検証環境 構成図

・IXの収容変更などもあったので、理想構成で検証環境は構築

・切り替え作業などはContainerlabで事前にリンクを設定して実際の作業シミュレーションが出来る。

・IPv4/IPv6と記載と記載している箇所は実際は本番環境で使用するIPをそのまま記載してます。

・IX, Transitなどは全てGoBGPで構築。IXは指定経路の学習・Transitはroute-viewsから経路ファイスを入手しLinuxコンテナに注入



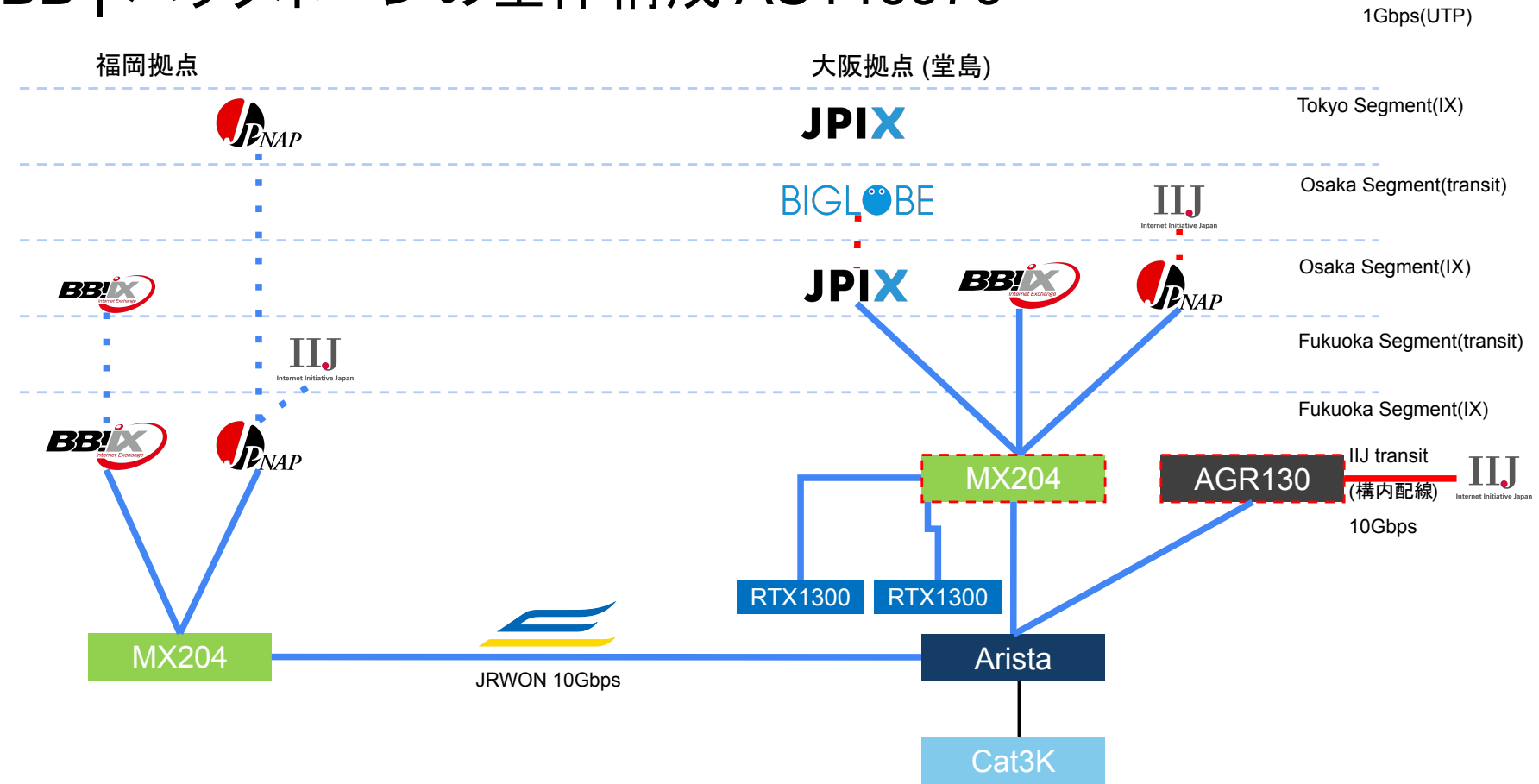
チームに入る前と入った後の心境の変化(1/2)

- チームへ入る前
  - ネットワーク機器もあまり触れたことがないため、構築や運用ができるか不安でした。
    チームの方へ迷惑をかけてしまうなども思ったりしていました。 (高田)
  - 今まで AS146979のサーバ関連を担当していましたが、この度 BBとして参加させていただき、BBの知識がなくリーダを務めることができないのではないかと感じていました (津田)
- チームへ入った後
  - 手厚いサポートを頂きながら、質問できる環境があり、さらに意欲があれば挑戦させてもらえるおかげで、大きな成長を実感できてとても楽しかったです。 JANOG NOCをやり切ったことで、できることが確実に増えましたし、何より大きな自信につながりました。 (高田)
  - AS運用経験を通じて、スキルアップすることができました。また、 BBの大変さを知ることができました。BBがとても楽しくなったので、AS146979のサーバだけでなく BBの運用にも参加してさらに勉強したいと感じました。 (津田)

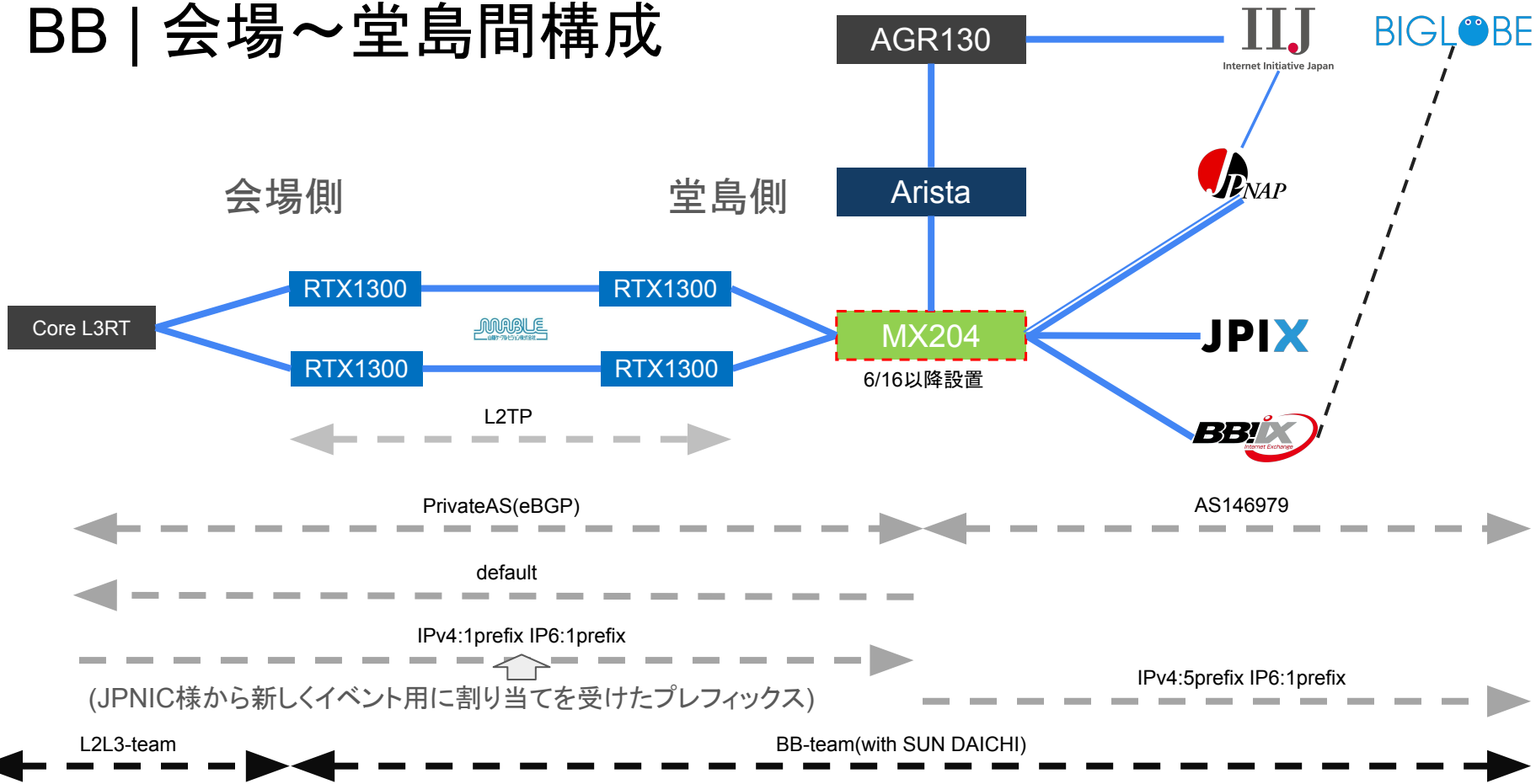
チームに入る前と入った後の心境の変化(2/2)

- チームへ入る前
  - ネットワークの中でも特に未知だった「 BGP」について一から学ぶ必要があり、知識がないことへの不安、技術力がないことへの不安、チームメンバーに迷惑をかけないかの不安など、様々な不安を抱えた状態でのスタートでした (高橋)
- チームへ入った後
  - 分からないことが悪いことではないという雰囲気と、非常に質問しやすい環境により、迷惑をかける不安が「自分はここに居ても大丈夫」であることの安心感に変わりました。活動全体を通してエンジニアとしての技術力や人間性など様々な面での成長を実感できて嬉しく思っています (高橋)

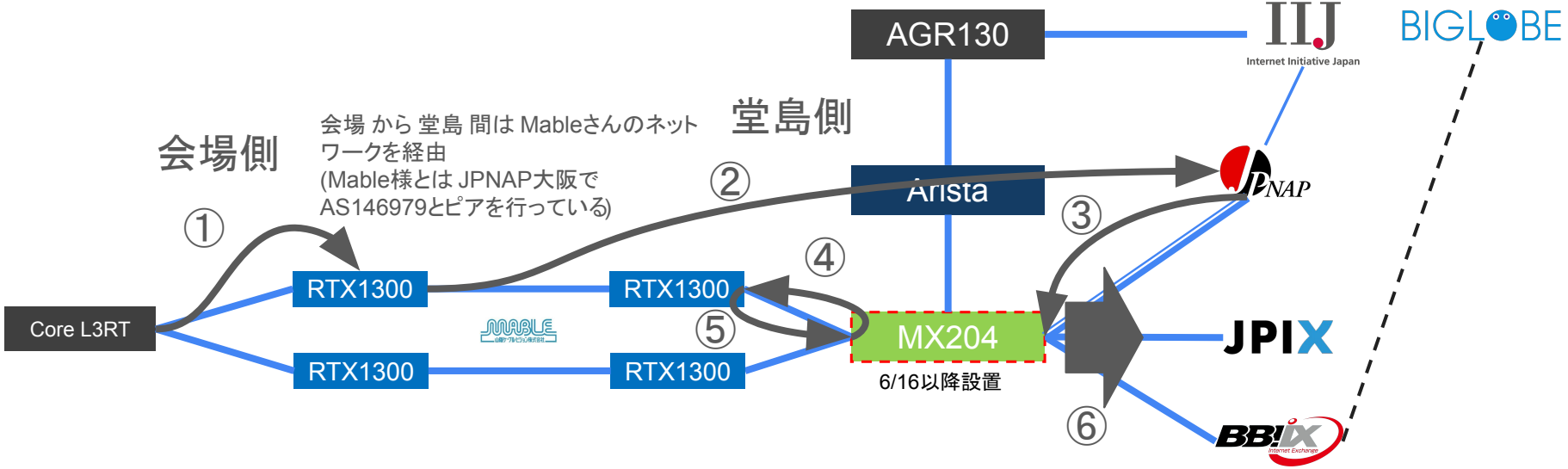
BB | バックボーンの全体構成 AS146979



BB | 会場～堂島間構成


AGR130
IIJ
Internet Initiative Japan
BIGLOBE
会場側
堂島側
Arista
JPNAP
RTX1300
RTX1300
Core L3RT
MX204
JPIX
RTX1300
RTX1300
6/16以降設置
BBIX
L2TP
PrivateAS(eBGP)
AS146979
default
IPv4:1prefix IP6:1prefix
IPv4:5prefix IP6:1prefix
(JPNIC様から新しくイベント用に割り当てを受けたプレフィックス)
L2L3-team
BB-team(with SUN DAICHI)

BB | 会場～堂島間構成 Traffic Flow

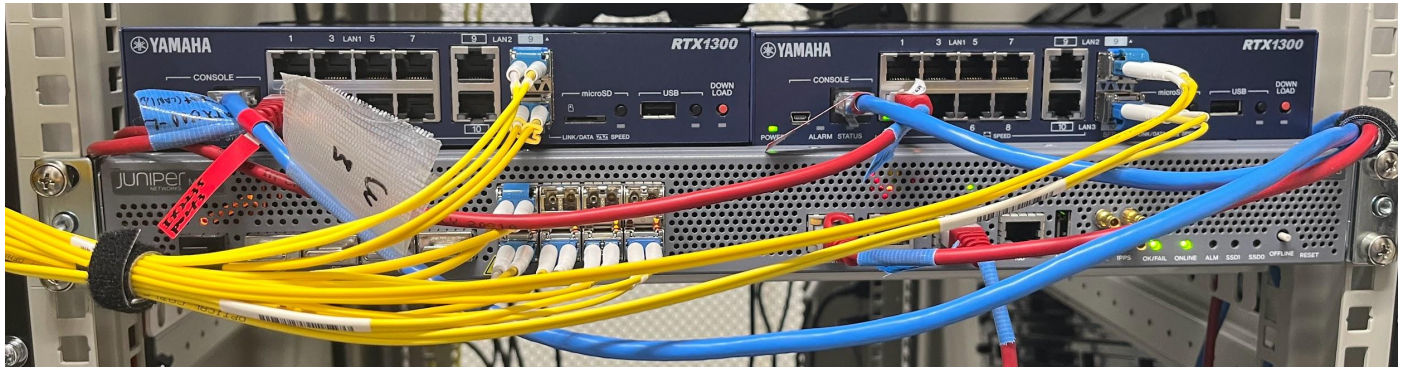


BB | 堂島側 機材

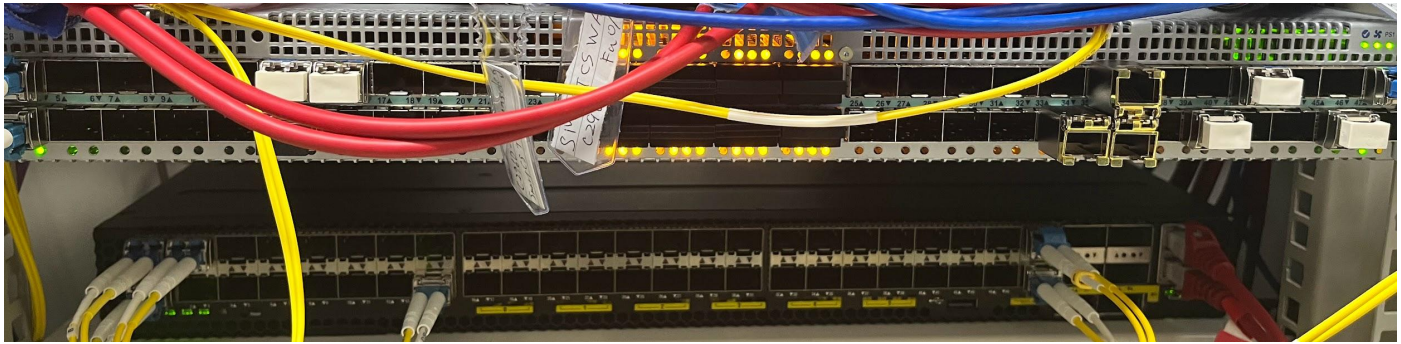
DC ラック

RTX1300

MX204



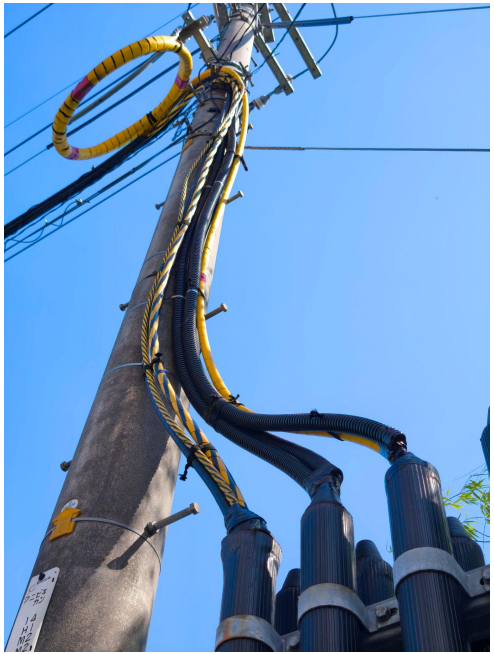
※写真は 3IXを MX204に移設する前
(この時は AGR130に接続)



Arista

AGR130

BB | 会場上流回線



MABLE様局舎近くから共同溝に降りる地点

共同溝の管路が川を渡る様子

BB | 会場上流回線

MABLE様局舎から約1.7km (会場裏手)

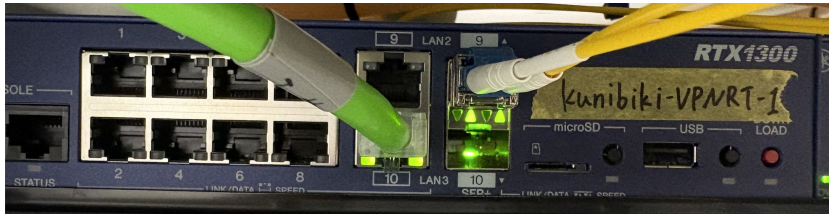
1F 多目的ホール ↑

BB | 会場上流回線

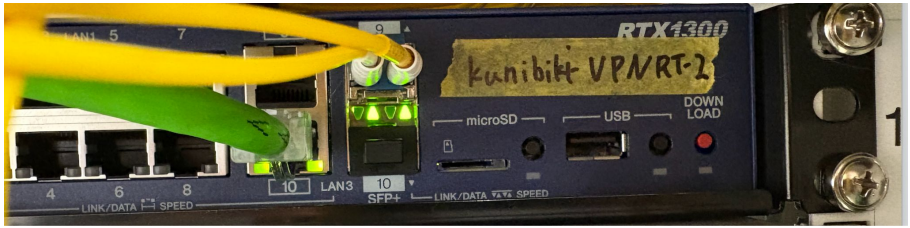
1F NOC部屋入って左手にONU
ONUからRTX1300までは
C6A UTPケーブル 75m
(ケーブルチーム作)

BB | 会場側

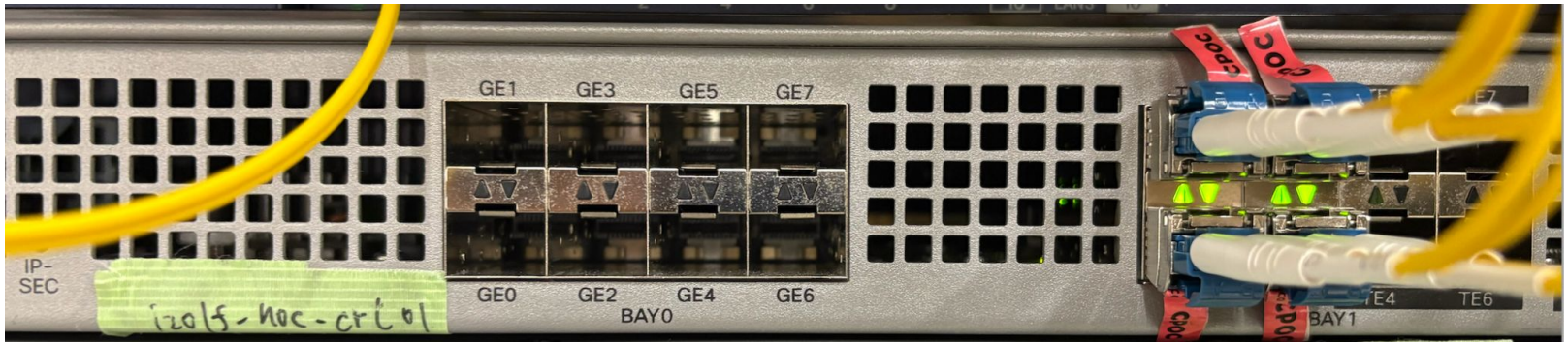
RTX1300(Active)



RTX1300(Standby)



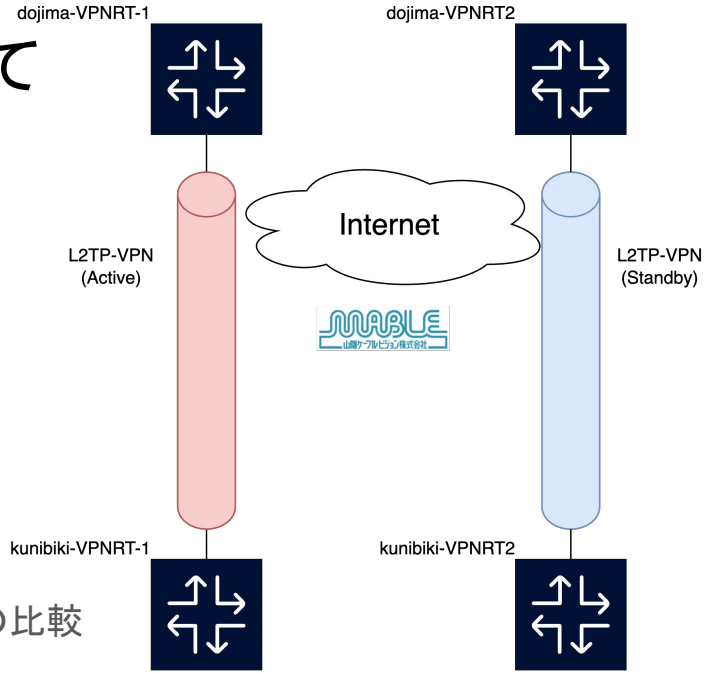
会場コアルータ



BB | 会場側上流　冗長構成について

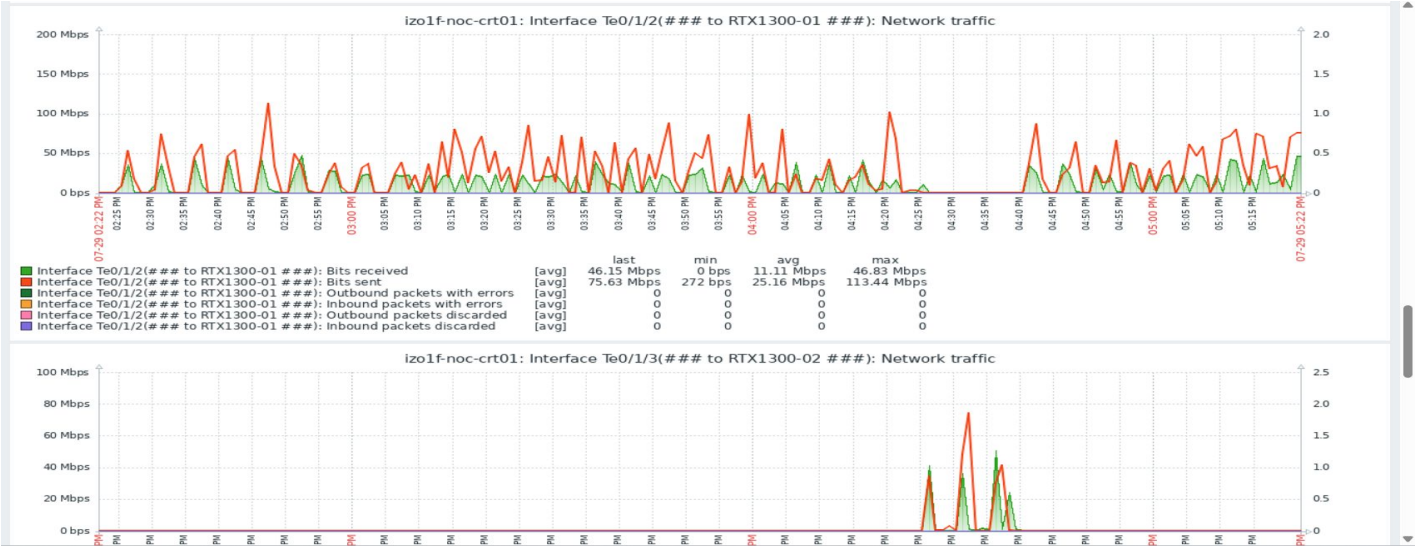
- 会場側コアルータ ～ 堂島MX204
  - Active: 通常時の経路
  - Standby: Activeが故障時の経路

- 経路迂回試験の確認項目
  - 経路迂回の可視化
  - 経路の迂回にかかる時間
    - BFD導入前後での経路迂回の切り替え時間の比較
      - 初期値は3000msec(3s)
      - 会期中は調整して300msec(0.3s)



BB | 経路迂回

- 障害発生時に迂回経路を通るか (mtrコマンドで観察)
  - 経路の迂回にかかる時間 : 約60s → BFD導入後での経路の迂回切り替え時間: 約1s
- 迂回の可視化 (4:25 PM付近でActive側をダウン, 上側: Active, 下側: Standby)
  - Traffic生成はiperf3で検証



BB | 障害試験

- 試験内容
  - 物理線を抜いてAct/Stbの切り替わりを確認
  - Act側のルータの電源を抜いてStbへの切り替わり確認
  - ActとStbを10秒ごとに切り替えて負荷をかけ、通信が保たれるかを確認
- 確認項目
  - トラフィック量, エラー率
- 結果 (障害前/障害後)
  - 約1分でact/stbの切り替わり
  - この後BFDを導入し、切り替えがさらに高速に！

	切り替え時間	パケロス率
BFD無し	約1分	10.3%(検証時の参考値)
BFDあり	1秒未満(0.3*3秒)	0.9%(検証時の参考値)

BB | RPKIによる経路制御

- RPKI(Resource Public Key Infrastructure)を使用
- 受信したBGP経路情報をROAと検証する
  合っていればvalid, 合っていなかったらinvalid, ROAがなければ unknown

Peer ＼ ROV	valid (LP)	unknown (LP)	invalid
バイラテラル	300	200	reject
各IX Route Server	250	150	reject
IIJ (PI/CUG)様	100	75	reject
Biglobe様	50	25	reject

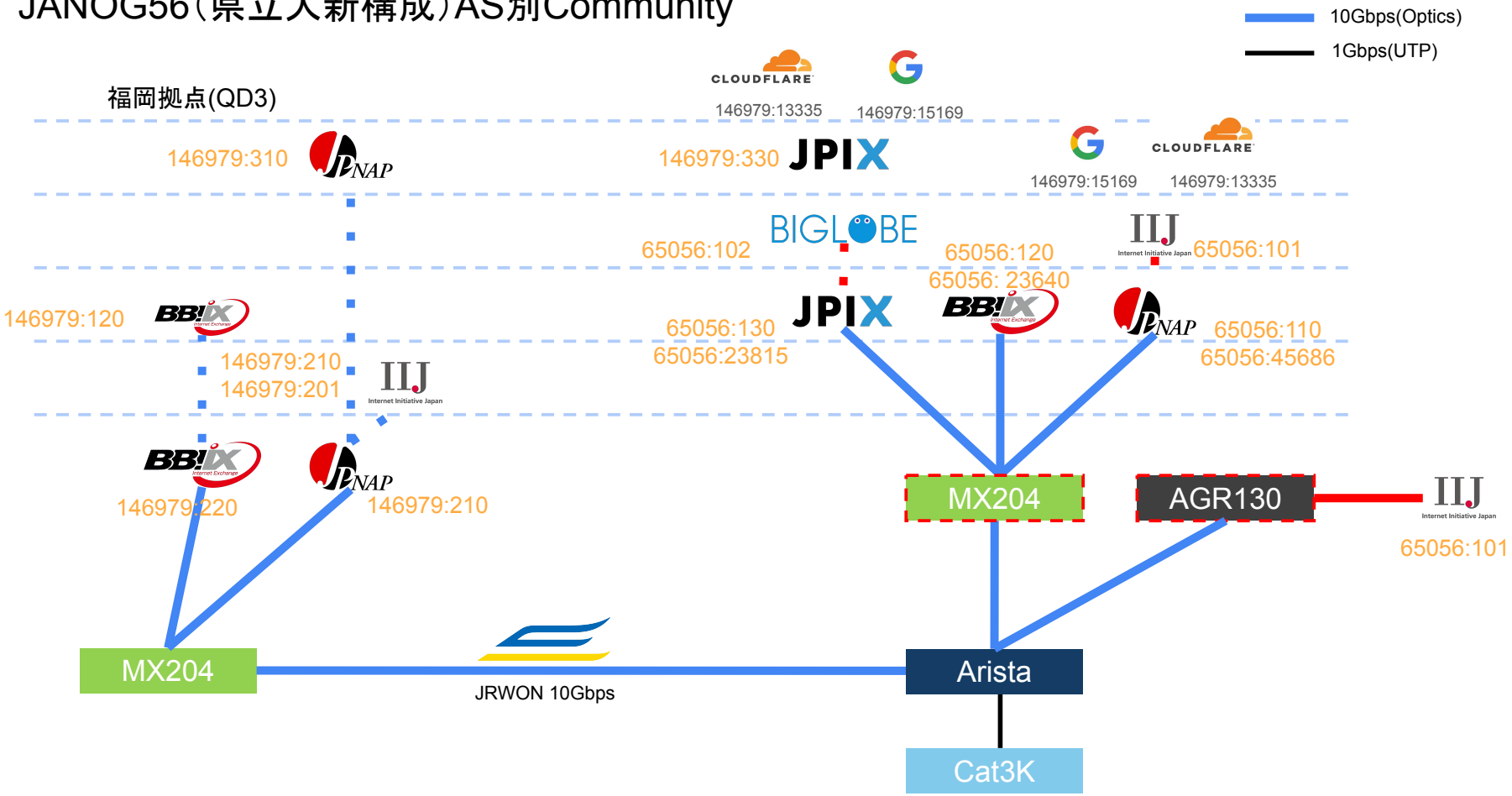
参考Config: https://www.mfeed.ad.jp/rpki/routersettings.html

BB | Route-Policy

- AS-PATH-Prepend
  - AS-PATHを長くすることで優先順位を下げる

Peer	MED	AS-PATH Prepend
バイラテラル	100	
各IX Route Server	110	
IIJ (PI/CUG)様	150	146979
Biglobe様	160	146979 146979

JANOG56（県立大新構成）AS別Community



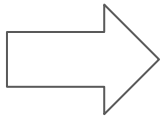
BB | community 一覧

community	members
BBIX-OSAKA	65056: 120
BBIX-OSAKA-RS	65056: 23640
JPIX-OSAKA	65056:130
JPIX-OSAKA-RS	65056:23815
JPIX-OSAKA-Transit-BIGLOBE	65056:102
JPNAP-OSAKA	65056:110
JPNAP-OSAKA-RS	65056:45686
JPNAP-OSAKA-Transit-IIJ	65056:101
RPKI-Invalid	65056:9999
RPKI-Unknown	65056:999
RPKI-Valid	65056:99

BB | RPKIによるルータ負荷

RPKI実施 前

CPU Idle割合 (5sec)	99%
メモリ使用率	13%



RPKI実施 後

CPU Idle割合 (5sec)	75%
メモリ使用率	19%

IXの ROA Cacheサーバを用いた RPKIを実施した際のルータの負荷

- CPU Idle (5sec)の割合が 24% 減少
- メモリ使用率が 6% (約 1.9GB) 上昇

BB | RPKI導入で確認できた Invalid経路数

Invalid経路数

IPv4	1516
IPv6	1051

IPv4, IPv6 合計 2567経路が Invalidとなった

ルーティングポリシーによる経路制御を行い、これら経路を rejectし、
会場から不正な経路へパケットが流れることを防止

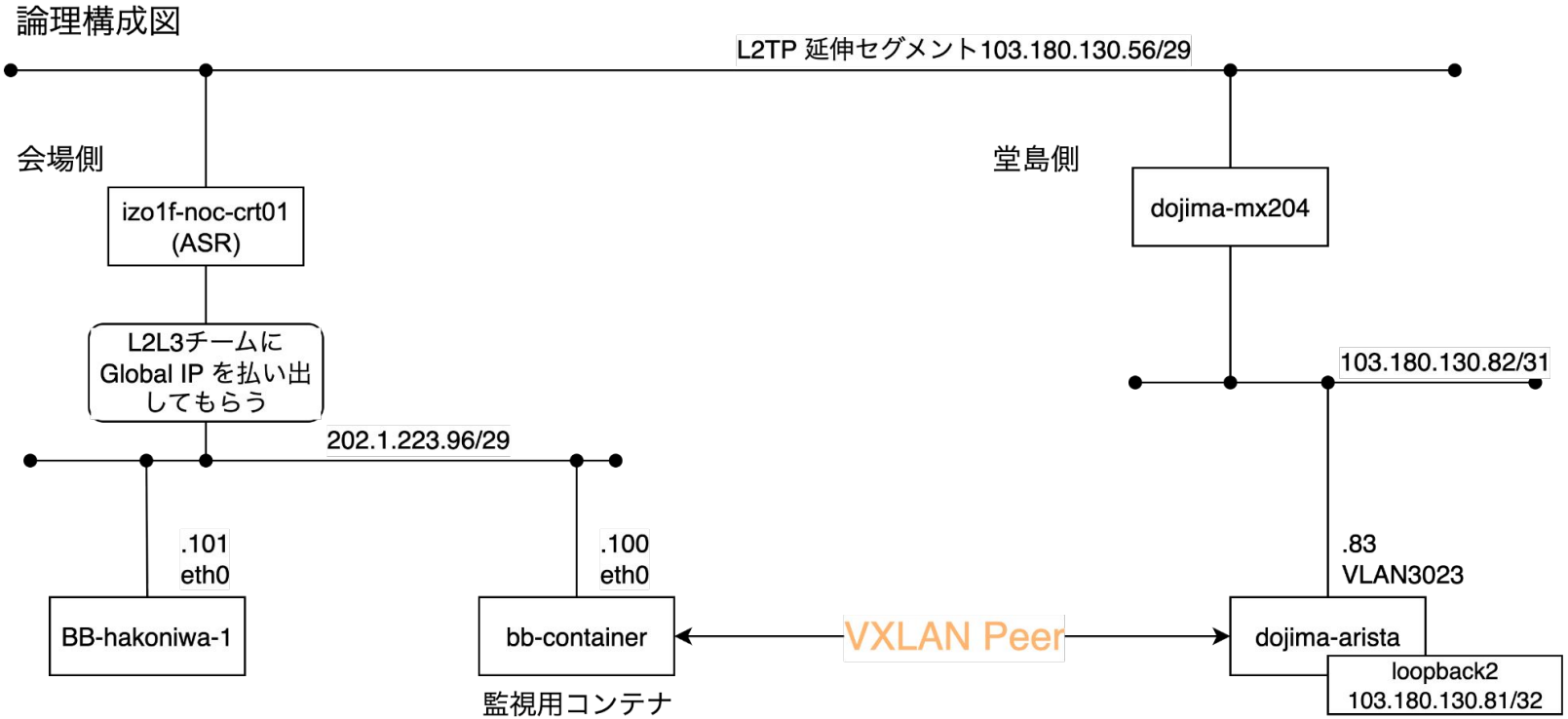
BB | 監視設計

- DCS 7280からBB ContainerへVXLANを張る
  - VXLAN : L3ネットワーク上にL2ネットワークを構築するトンネリングプロトコル

- 長県大の VLANをカプセル化して会場側へ送る
  - 全体の範囲を会場側でマネジメントポートから監視可能

BB | 監視設計

VXLANの論理構成



BB | 監視基盤、項目

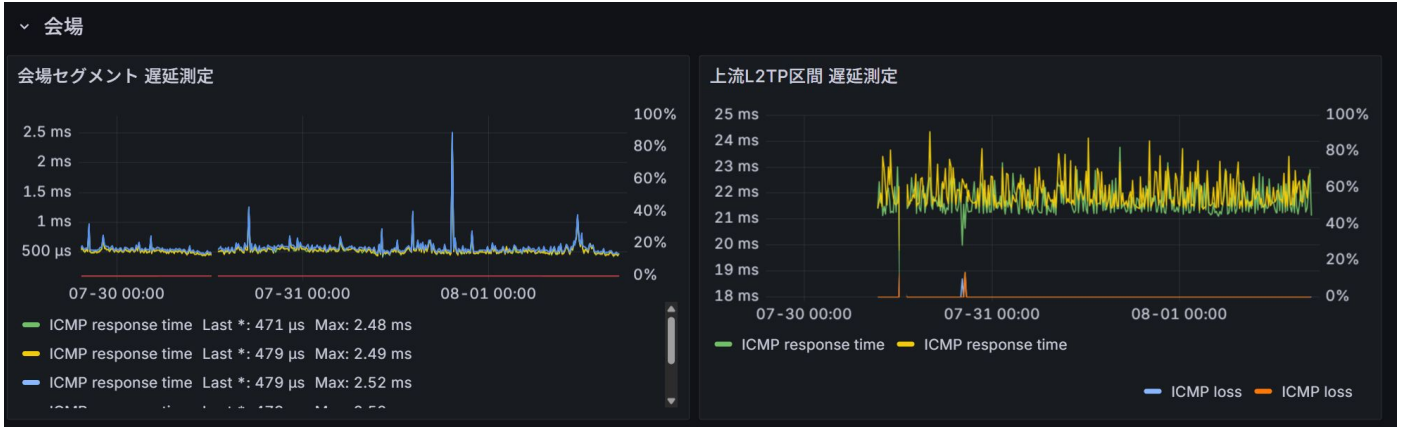
- 監視基盤
  - Zabbix 7.0 LTS
  - Grafana
  - Prometheus Blackbox Exporter (外形監視)
- 監視項目 (取得したデータ)
  - 各機器の syslog
  - 堂島、会場 RTX1300の死活
  - 会場関連の機器での遅延測定
  - 対外 ICMP Loss, Response
  - 対外(3IX、IIJ様の PNI, PI/CUG)トラフィック量
  - 特定Webサーバへのアクセスレイテンシ
  - AS146979 福岡 堂島 間 トラフィック
  - 堂島 会場 間 RTT(Round-Trip Time)

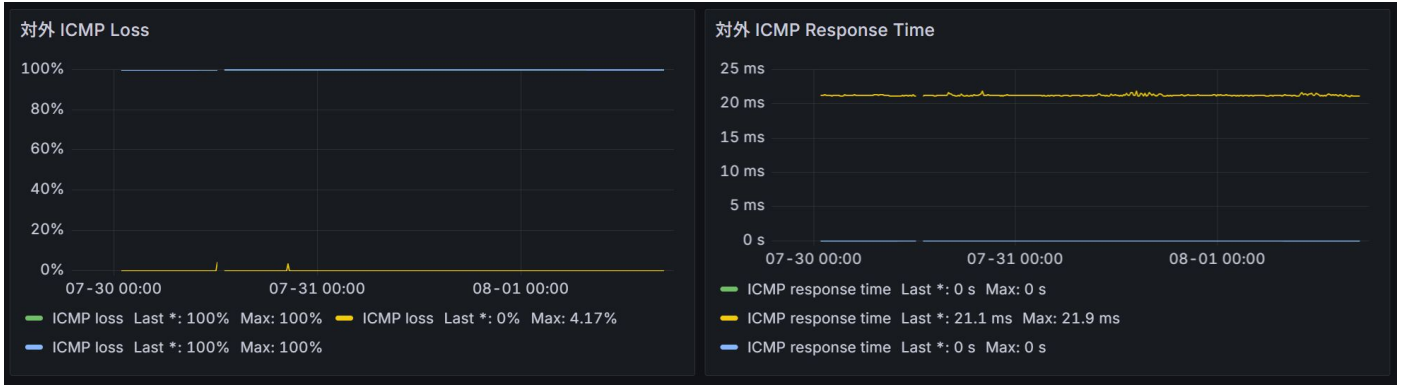
BB | 監視データ

Grafanaでの可視化 [syslog, 堂島、会場 RTX1300の死活]

BB | 監視データ

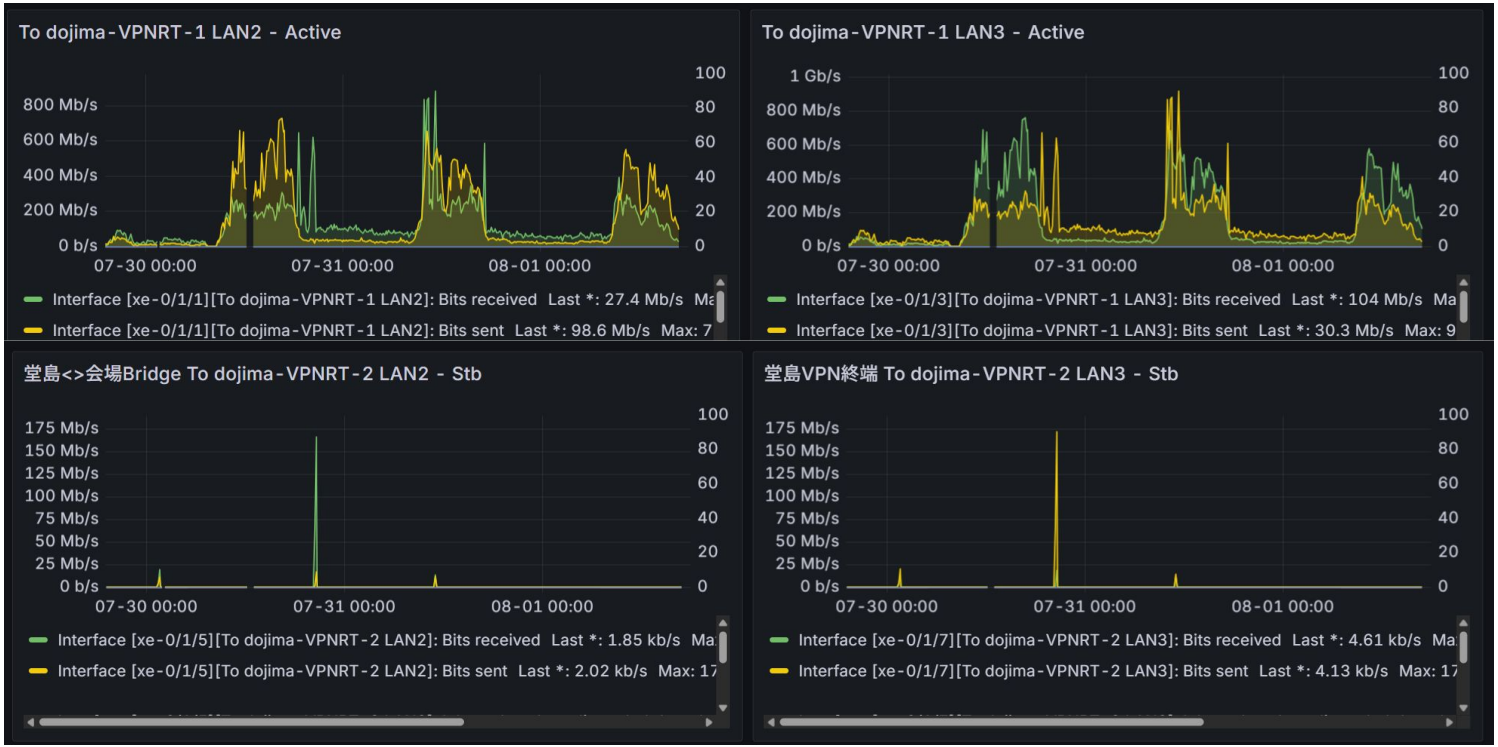
Grafanaでの可視化 [会場関連の機器での遅延測定、対外ICMP]





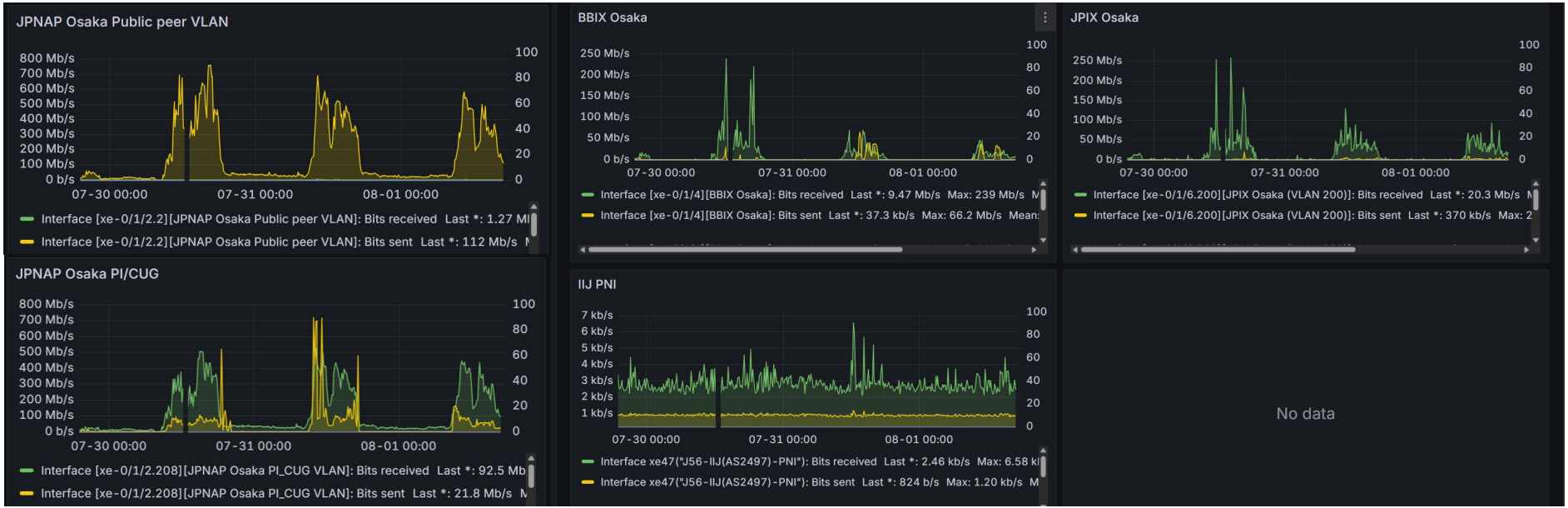
BB | 監視データ

Grafanaでの可視化 [会場設置 RTX1300(act, stb)のトラフィック量]



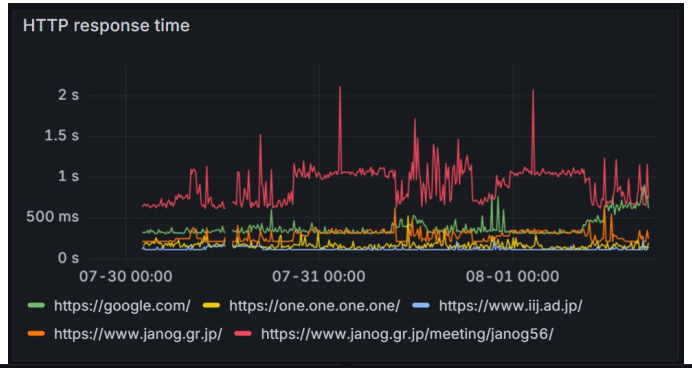
BB | 監視データ

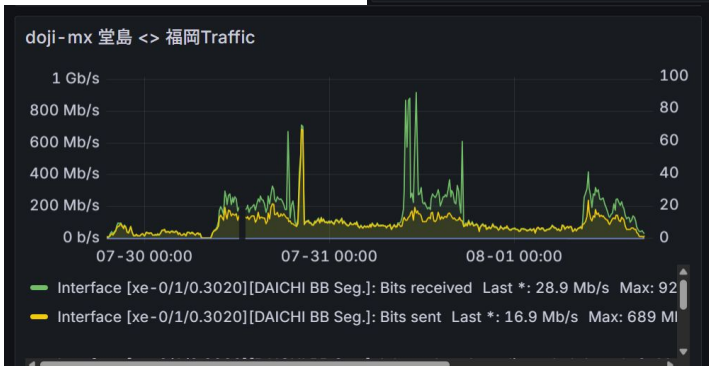
Grafanaでの可視化 [対外(3IX、IIJ様の PNI, PI/CUG)トラフィック量]

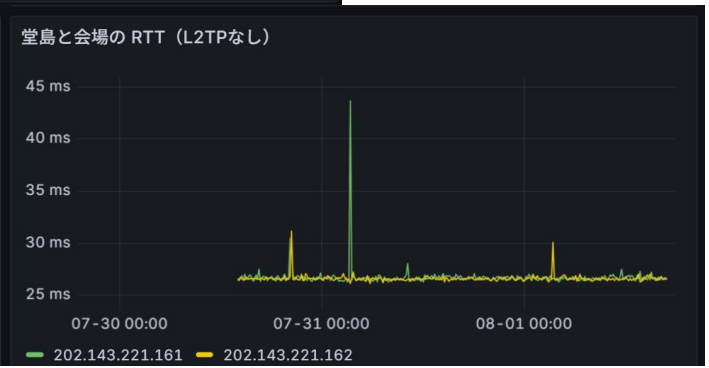


BB | 監視データ

Grafanaでの可視化 [特定Webサーバへのアクセスレイテンシ、
AS146979 福岡 堂島 間 トラフィック、堂島 会場 間 RTT(Round-Trip Time)]







BB | ピアリングの紹介

- 下記の企業様にピアリクエストを頂きました
  - 株式会社シナプス さま
  - 株式会社STNet さま
  - GMOネットアイアールディー さま

ピア申請してくださった組織の皆様
本当にありがとうございます！！