

モンスターストライク® x IBM Cloud

株式会社ミクシィ

SRE 清水 勲

ネットワークエンジニア 吉野 純平



自己紹介

清水 勲 @isaoshimizu

- SIerで受託開発、自社プロダクト開発、運用を約8年
- 株式会社ミクシィ
 - 2011.8～ 運用部 アプリ運用グループ所属、SNSの運用
 - 2014.4～ モンスターストライクの運用にジョイン
 - 2015.8～ XFLAG スタジオ システム開発部
 - 2016.7～ XFLAG スタジオ ゲーム開発室 SRE グループ
 - 2018.2～ ヴァンテージスタジオ みてね事業部
- 楽器演奏、クラフトビールが大好き 🍺

このセッションでお伝えしたいこと

- マルチクラウドの活用事例
- クラウドを扱うためにやるべきこと、必要なこと
- IBMクラウド（ベアメタル）の特徴
- ネットワーク、サーバーの問題をどうやって解決していったか

アジェンダ

- マルチクラウドとIBMクラウド
- 2017-2018年末年始の負荷対策
 - ネットワーク編
 - サーバー編
- IBMクラウドを使ってみてどうだったか

マルチクラウドを活用したインフラ

日本版モンスターストライクのインフラは、ハイブリッド、マルチクラウドの構成

- ハイブリッド
 - オンプレとクラウド
- マルチクラウド
 - AWS、GCP、GMO、IBMの4クラウドを併用

マルチクラウドのメリット

- リスク分散
 - リソースのキャパシティ
 - AZ単位での障害リスク回避
- クラウド独自の機能が使いやすくなる
 - クラウドAからクラウドBのリソースを使う
 - クラウドBにしかないハードウェア（例: GPU、高速ストレージなど）、コストパフォーマンスが良いなど

IBMクラウドの良さとは

※ベアメタルの話に限定しています

- 新しいハードウェア
- 高パフォーマンス
- 弊社DCや利用しているクラウドとの相性
(低レイテンシ)
- ベアメタルもAPIで操作可能

2017-2018年末年始

2017-2018年末年始の負荷対策

- 年末年始は各種キャンペーンでアクセスが急激に増える
- アプリケーションサーバー向けのサーバーリソース確保が急務
- 弊社DCにあるデータベースやキャッシュサーバーと近い距離のクラウドが望まれていた

年末年始対策：ネットワーク編

自己紹介

吉野 純平 @junpei_y

- 新卒でミクシィに入社
 - 2008.4～開発部 運用グループ インフラチームでネットワークや広告システムを運用
 - 2010.4(?)～ネットワークを引退しアプリケーション運用へ
 - 2012.12(?)～全レイヤ戦を行う人に
 - 2013年 バックオフィス戦士として無双時代
 - 2017年 ネットワーク周りの開発を頑張っている

最初は何をしたかったのか

- ハイブリッドクラウドを運用している
- スケールアップをしたかった
- スケールアップ = 時間を買うこと
 - 負荷対策の優先順位を選べるようになる

クラウド接続技術 振り返り

弊社ケースの時系列

- **Cloud IX研究会** www.cloudix.jp
 - インターネット、クラウド間接続技術の検証や技術共有、実網での実装
 - 2013年ごろ
- **InternetWeek2015 T3セッション**
 - <https://www.nic.ad.jp/ja/materials/iw/2015/proceedings/t3/>

InternetWeek 2015のサマリー

- レイヤ3での接続にフォーカス
- 1つの面では運用が辛いことを学ぶ
- 汎用ASICの進化でMPLSをサポートする機器が増える流れが見えた
 - 2014年からMPLSに技術投資を始めた
 - キャリアさんのVPN網の技術なので枯れていると想像

レイテンシ の悩み

Latency は物理依存

tcpでの内部アクセスN回すると
RTT*N以上でのオーバーヘッド

Nが1,000等のアクセスは数秒オーダーの
処理時間を要する

レイテンシが小さいクラウドを探す

IBM接続

タイムライン（泣

10月12
日

11月

12月

1月

書類
仕事

LOA

敷設

IBM工期

X

X

X

技術
検討

事前
試験

機器
準備

本番
検証

商用利用

技術検証 フェーズ

技術検証

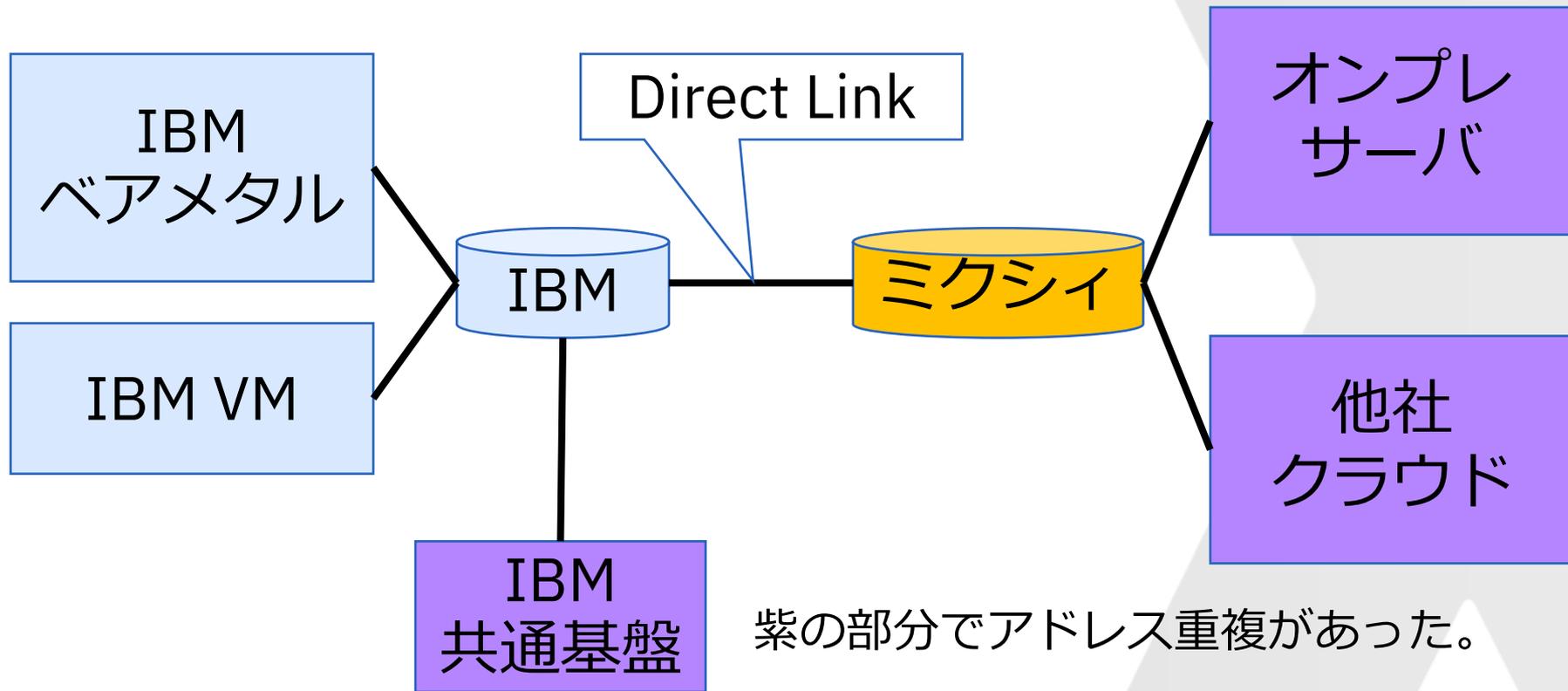
他のクラウドとデザインの方角性が違ふ

- まず仕様書が見つけづらふ
- よくあるVPCみたいなものがない
- 複数の事業を収容しやすすくない

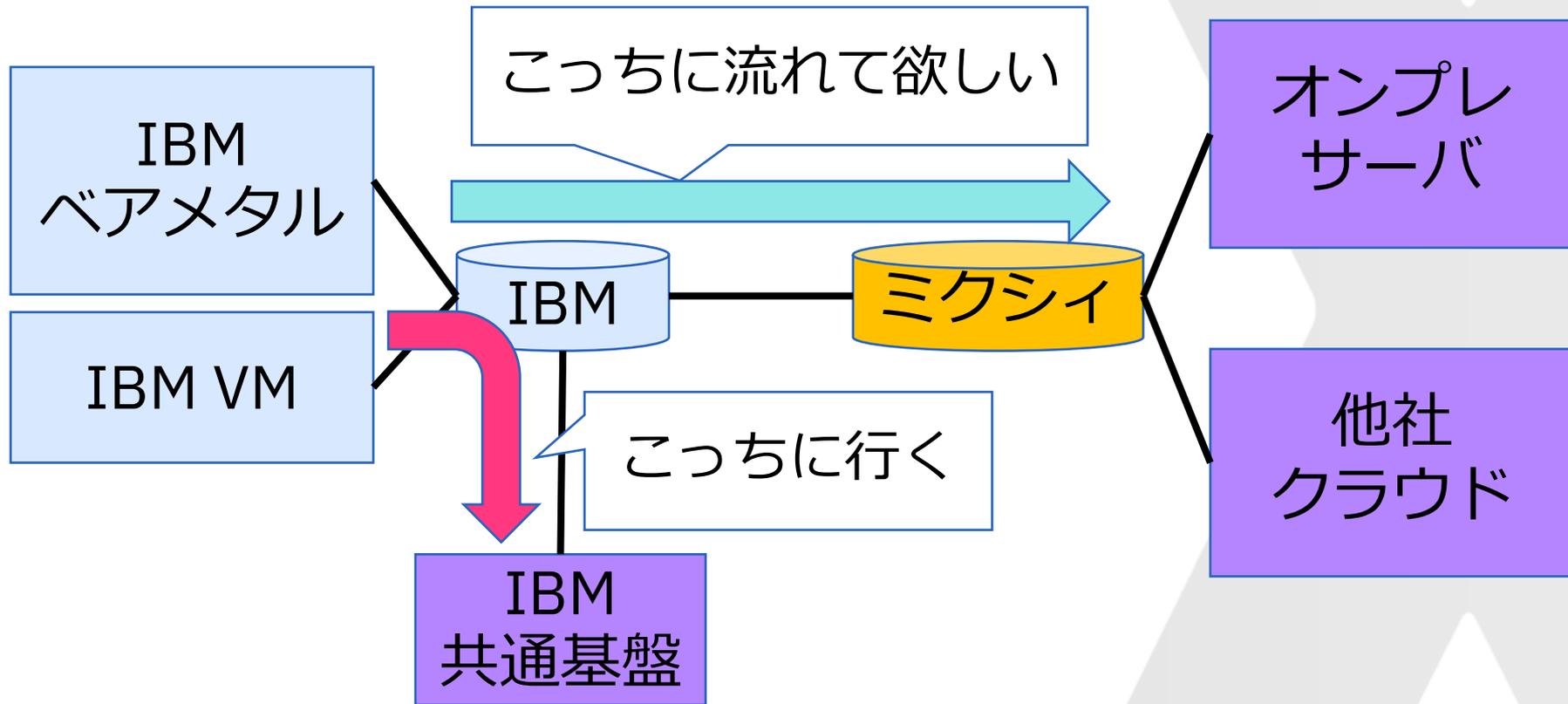
プライベートIPアドレスの制約

- **事業者側からアサインされる**
 - リージョンごとにレンジが決まっている
 - 世界で一意的なアドレスになる
 - 基盤で使われているアドレスはDirectLink経由で疎通不可能
- **Custom Private Addressingは？**
 - ご利用は相談の上、慎重に（とのこと）

弊社はバッティングしてた



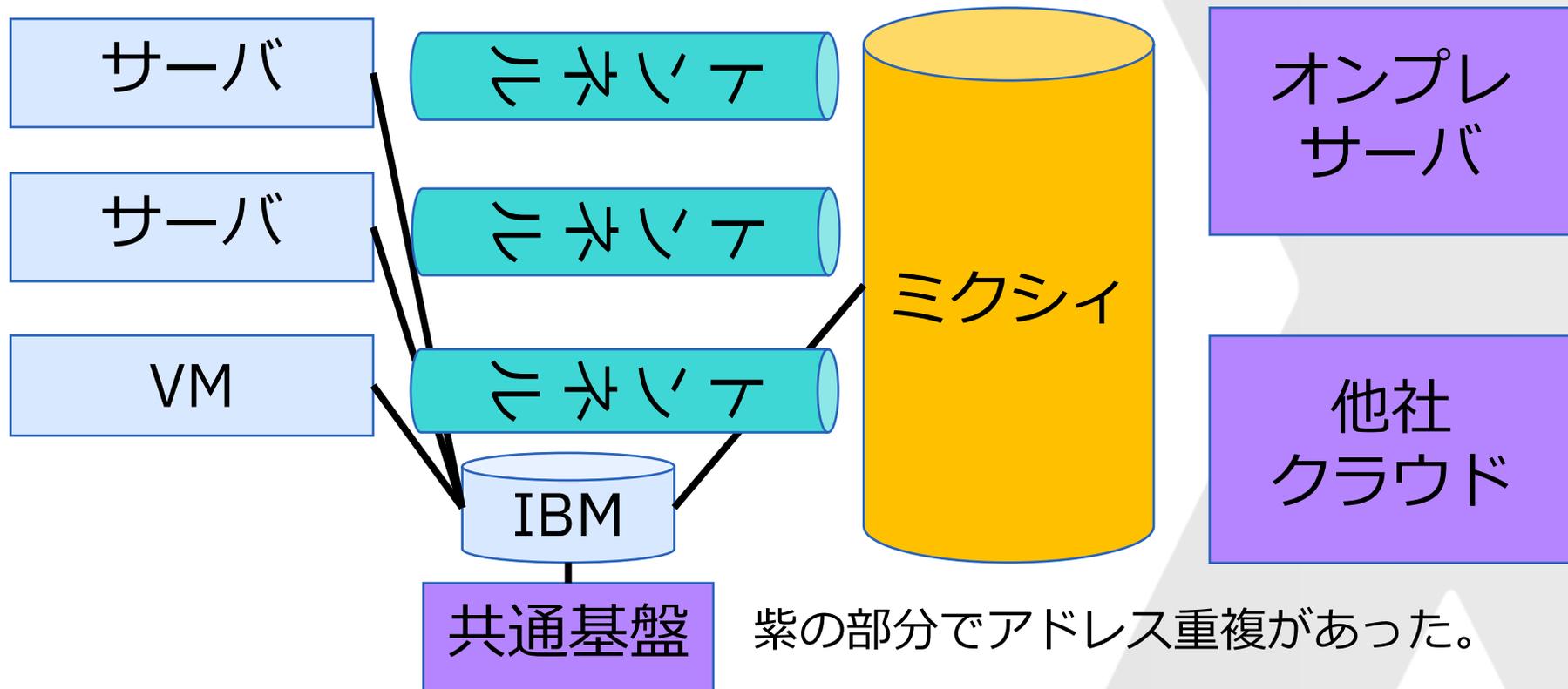
疎通ができない



解決策案

- **既存設備の該当範囲のリナンバ**
 - 工数で納期を考えると無理
- **既存設備との間でのNAT**
 - アプリケーションの改修が必要な箇所あり
- **IBM上のルータとトンネル接続**
 - パケットをCPUで処理するタイプであったため、10Gbpsをショートパケットは無理

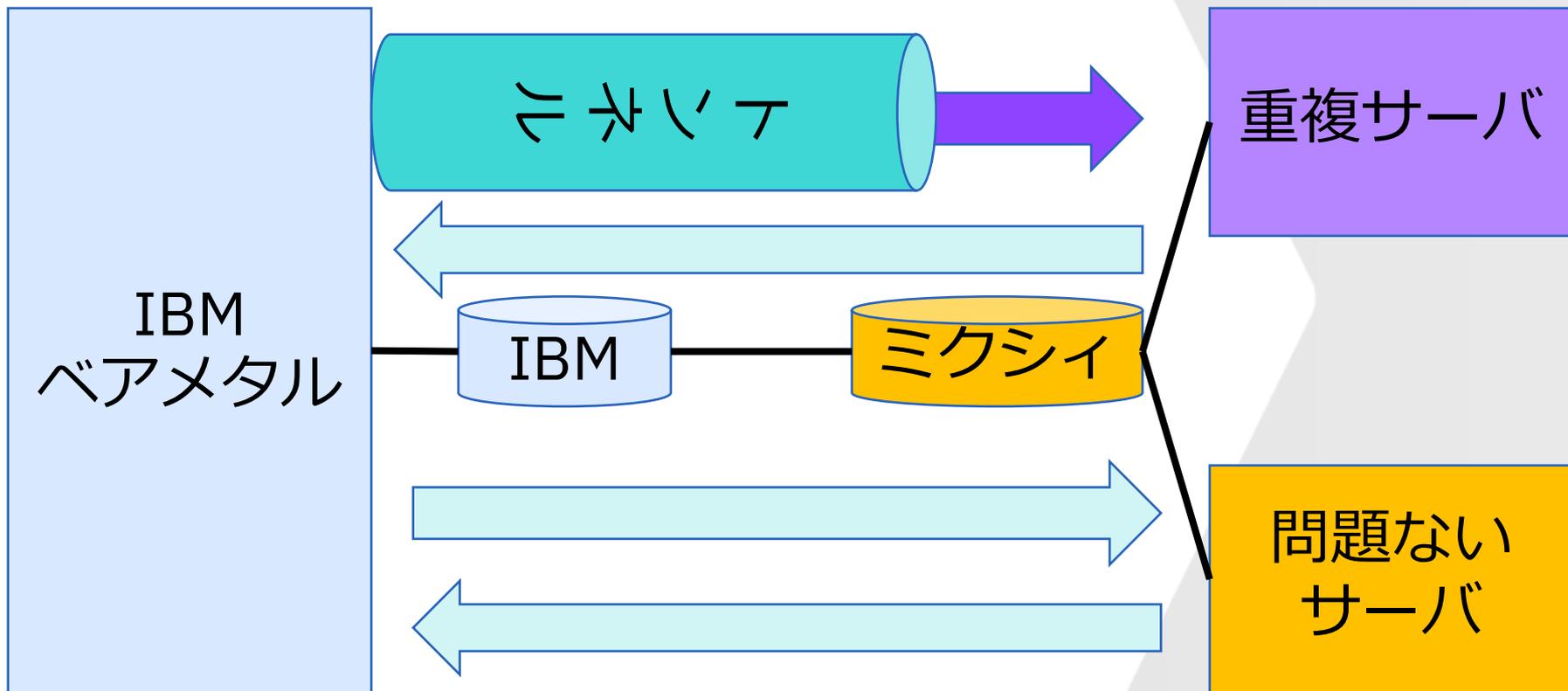
採用した案：サーバでのトンネル



トンネルの技術選択

- 色々あるがvxlanとGREを比較
- GREを採用した
 - トンネルヘッダーが短い（重要）
 - 当時使っていたKernelでモジュールロードだけで動く
 - 非対称で使いやすいかも
 - 利用NWハードウェアの技術制約が少なそう

特徴的な例



本番検証 フェーズ

予定外な問題が発生

レイテンシが思いの外大きかった

GREなしpingで200usec

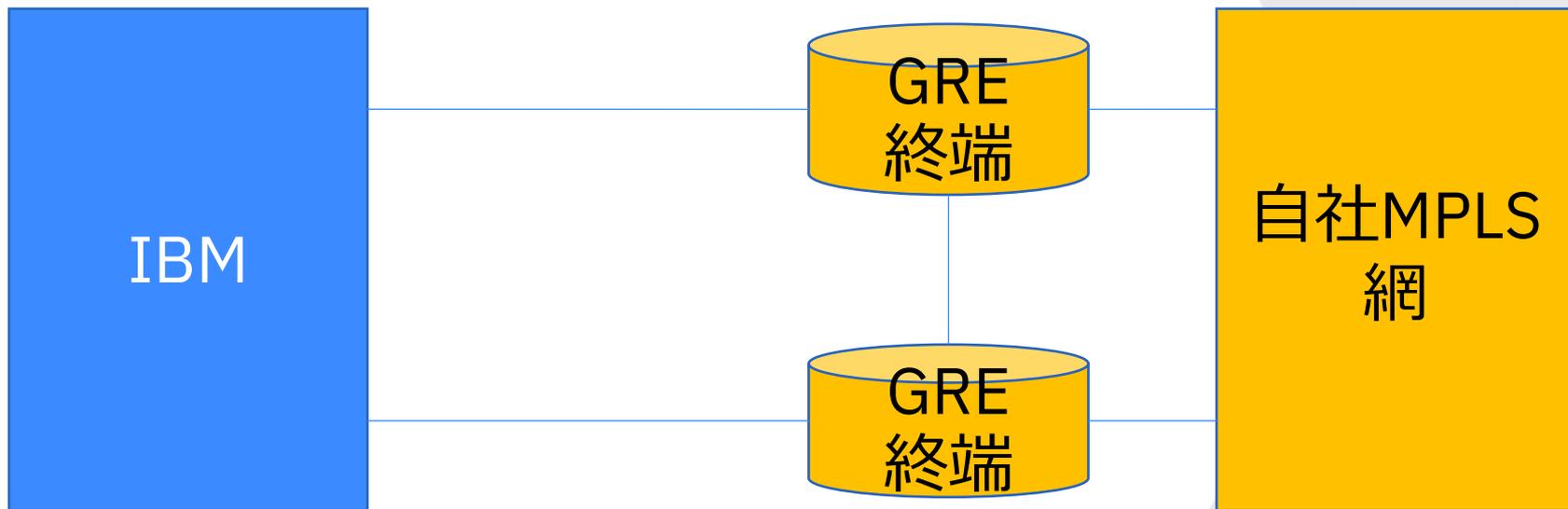
GREありpingで11,000usec（前後）

原因は、HWに処理を落とせてなかった
CPU処理になっていたため性能が出ない

対策

- 事前検証と結果が異なった
- 実網との比較から違いをパターン別に実装して比較データ収集
- 該当チップで処理されて運用しやすいようなコンフィグを選び実装した

最終構成



両系とも終端アドレスを1つとしてANYCAST
Active/Active構成を経路制御で実現

年末年始対策：サーバー編

2017-2018年末年始の負荷対策

- アプリケーションサーバーの増強
- 通常のキャンペーンは、約2,000-4,000vCPU増強程度
- 年末年始は約10,000vCPUをさらに追加する必要があった

利用したIBMクラウドの規模

追加が必要な約10,000vCPUのうち、
IBMクラウド（ベアメタル）の分は
どのくらいか見積もり

当初の計画では、
アプリケーションサーバー用 50台
E5-2690 v4(28core) 2,800vCPU分

50台オーダーした結果

オーダーしたマシン

マシンにインストールされたOS/ミドルウェア環境がLAMP環境だった 😬

不要なプロセスが多く起動していた

マシンをどうするか

- 各種サーバーのプロビジョニングはChefを使って適用して構築する運用
- 最小限の環境が欲しかった
- 50台分手作業でWebコンソールで作業はしたくない…
(作業漏れが出たり、ミスするリスク)

OS再インストールツールを作った

- Go言語でAPIクライアントツールを作った
- SDK、APIの対応によりWebコンソールに頼らず高速に作業できる
- ツール利用によって、誰がやっても同じ構成が簡単に作れる（超重要）
- APIの使い勝手を知る上でもちょうどよい規模感だった

なぜGoを選んだか

- 動作が高速
- ポータビリティ (one binary)
- 開発しやすい言語
- 技術情報が豊富にある
- Go製のソフトウェア (Docker、Kubernetes、Terraform、etc...) が豊富
- すでにGoでの開発実績、知見があった

50台の再インストール

- 自作のツールを利用して再インストール
- APIを介して50台一気に実行
- 作業漏れの心配がない
- コマンドさえあれば誰でも何度でも実行できる
- 無事再インストールできて本番投入

これですべて完了と思っていた

急遽追加のオーダー

- 想定外のリソース不足が発生（IBMクラウドの問題ではない）
- 不足分をIBMクラウドで補いたい
- ロンドンからCPU空輸対応（間に合わず）
- 東京の代替マシン47台（E5 2690-V3 24core 1,128vCPU分）のオーダー、セットアップ、本番投入

無事予定通りのリソース確保 🎉

IBMクラウドを使ってみて

よかったこと

- ベアメタルにも関わらずAPIで操作できる良さ（GoのライブラリがGitHubにあった）
- 低レイテンシ
- ハイパフォーマンス
- 営業、技術のサポート体制の手厚さ

IBMクラウドを使ってみて

期待したいこと

- Webコンソールの高速化、UIの統一
 - SoftLayer ? Bluemix ?、レスポンスが遅い、UIに古さを感じる
- ドキュメント、サンプルコードの充実、正確性
 - どのドキュメントを読むべきか、どのSDKを使うべきか。わかりづらい。

Thank you!



ワークショップ、セッション、および資料は、IBMまたはセッション発表者によって準備され、それぞれ独自の見解を反映したものです。それらは情報提供の目的のみで提供されており、いかなる参加者に対しても法律的またはその他の指導や助言を意図したのではなく、またそのような結果を生むものでもありません。本講演資料に含まれている情報については、完全性と正確性を期するよう努力しましたが、「現状のまま」提供され、明示または暗示にかかわらずいかなる保証も伴わないものとします。本講演資料またはその他の資料の使用によって、あるいはその他の関連によって、いかなる損害が生じた場合も、IBMは責任を負わないものとします。本講演資料に含まれている内容は、IBMまたはそのサプライヤーやライセンス交付者からいかなる保証または表明を引き出すことを意図したもので、IBMソフトウェアの使用を規定する適用ライセンス契約の条項を変更することを意図したものでなく、またそのような結果を生むものでもありません。

本講演資料でIBM製品、プログラム、またはサービスに言及していても、IBMが営業活動を行っているすべての国でそれらが使用可能であることを暗示するものではありません。本講演資料で言及している製品リリース日付や製品機能は、市場機会またはその他の要因に基づいてIBM独自の決定権をもっていつでも変更できるものとし、いかなる方法においても将来の製品または機能が使用可能になると確約することを意図したものではありません。本講演資料に含まれている内容は、参加者が開始する活動によって特定の販売、売上高の向上、またはその他の結果が生じると述べる、または暗示することを意図したもので、またそのような結果を生むものでもありません。パフォーマンスは、管理された環境において標準的なIBMベンチマークを使用した測定と予測に基づいています。ユーザーが経験する実際のスループットやパフォーマンスは、ユーザーのジョブ・ストリームにおけるマルチプログラミングの量、入出力構成、ストレージ構成、および処理されるワークロードなどの考慮事項を含む、数多くの要因に応じて変化します。したがって、個々のユーザーがここで述べられているものと同様の結果を得られると確約するものではありません。

記述されているすべてのお客様事例は、それらのお客様がどのようにIBM製品を使用したか、またそれらのお客様が達成した結果の実例として示されたものです。実際の環境コストおよびパフォーマンス特性は、お客様ごとに異なる場合があります。