

# **”Change” our private cloud infrastructures from single-AZ to Multi-AZs**

2022/7/13 JANOG50

# Agenda

1. Multi-AZsインフラ基盤の設計概要  
鈴木 雄一郎 – NW室サービスネットワーク1チーム
2. Multi-AZsインフラ基盤を支えるバックボーンネットワーク  
向井 脩 – NW室サービスネットワーク1チーム
3. Multi-AZs環境化での各種NWファンクションのサポートについて  
川上 けんと – Verda platform室 ネットワーク開発チーム
4. Multi-AZs環境へのPlatform機能の追従、サーバリソースコントロール  
山田 英樹 – Verda platform室 Verda Reliability Engineering チーム
5. ディスカッション

# About me

- **プロフィール**

名前：鈴木 雄一郎 - Yuichiro Suzuki

所属：ITサービスセンターNW室サービスネットワーク1チーム

拠点：茨城県水戸市

趣味：テニス、子育て (3女,1男)

- **職歴**

- SB Telecom (-2012) : 法人向けIP-VPNサービス(solteria)の設計/構築

- NTT docomo(2012-2018): LTE/3G用のIPバックボーンの構築、EPC開発

- KDDI (2018-2019): 5G用NW設計/構築、ホワイトボックススイッチ用OS開発

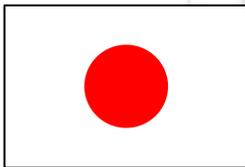
- LINE (2019～) : DCNW、DCI MPLS NW、Fintech NW設計/構築、AS38631

# About LINE

# LINEの月間アクティブユーザー（2022年3月時点）



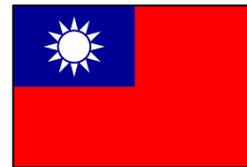
TOP 4 MAU:  
約1億7,600万人



日本  
約9,200万人



タイ  
約5,300万人



台湾  
約2,200万人



インドネシア  
約900万人

# LINE Services



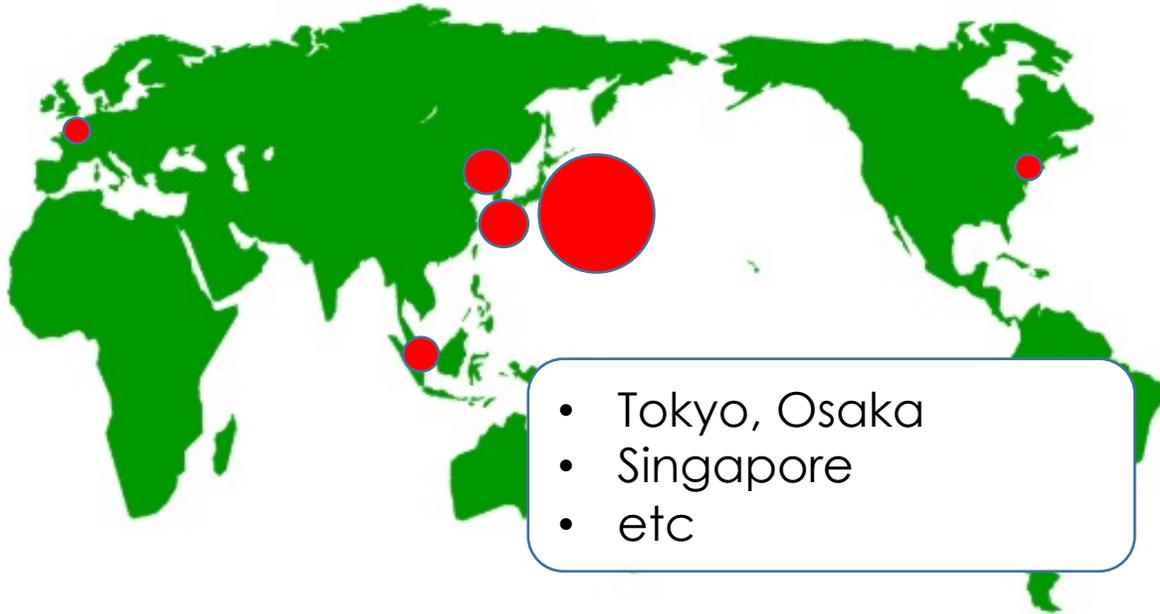
**LINE NEWS**

**LINE 証券**



**LINE GAME**

# LINEのインフラ規模



Global Office network 30+

Internet 800Gbps

CDN 2.6Tbps

VM 94,000+

Verda  
(private cloud)

PM 68,000+

NW Device 10,000+

Rack 3,700

# LINEのインフラ組織

## IT Service Center

### Verda Platform室

managed service開発  
Platform開発

### Network 室

datacenter設計・構築・運用  
network 設計・構築・運用

### DB 室

MySQL, HBase, MongoDB  
Redis ,etc

### System 室

Server, hardware , OS  
system integration

# 本題

# Multi-AZsインフラ基盤への移行のきっかけ

## 2020年11月9日に発生した「LINE」アプリ不具合について

2020.11.10 | LINE



平素はLINEをご利用いただきありがとうございます。

「LINE」アプリにおきまして、2020年11月9日 14時32分から14時52分の約20分間、メッセージを送受信できない等の不具合が発生しました。ご迷惑をおかけして大変申し訳ございませんでした。不具合の詳細と対応状況は以下の通りです。

### 1. 不具合の詳細

ネットワーク機器の保守作業において通信障害が発生し、通信ができない状態が数分続きました。

これに伴い「LINE」にログインして利用する機能・サービス全般（一部を除く）が使えない他、アカウントを作成できない状態となり、日本国内に加え一部海外のユーザーも影響を受けました。

LINEは14時32分に障害が発生した後、事象を認識してただちに状況確認や対応策を講じ、14時52分にはメッセージ送受信をはじめとする不具合は概ね解消いたしました。

22時08分にLINE公式アカウント等、LINEを利用する各サービスも含め正常通りご利用いただけるようになりました。

#### 【不具合の影響範囲】

- ・対象：2020年11月9日 14時32分から14時52分の間に「LINE」アプリを利用したユーザーの皆さま（国内／海外）
- ・障害影響時間：2020年11月9日14時32分から14時52分（その他LINE公式アカウントなど一部機能においては21時50分まで）
- ・障害発生原因：ネットワーク機器の計画的な保守作業において、迂回ルートに切り替えできない状態になったため
- ・障害範囲：「LINE」にログインして利用する機能・サービス全般（一部を除く）、アカウントの作成

### ●発生事象

- ・ネットワーク機器障害：数分間
- ・メッセージ送受信不可：約20分間
- ・全サービス完全復旧：約8時間

# Multi-AZsインフラ基盤への移行のきっかけ

## 2021年4月12日に発生した「LINE」アプリ障害について

2021.04.13 | ALL



平素はLINEをご利用いただきありがとうございます。

「LINE」アプリにおきまして、2021年4月12日 15時9分から16時1分の約50分間、メッセージを送受信できない等の障害が発生しました。ユーザーの皆様にご迷惑やご不便をおかけしましたことをお詫び申し上げます。障害の詳細と対応状況は以下の通りです。

### 1. 障害の詳細

当社が利用している国内データセンター提供事業者の誤作業に起因し、「LINE」のメッセージを送受信できない等の障害が発生しました。同外部事業者において、本来電源停止を伴わずに電源設備のメンテナンス作業を行うべきところ、同社の誤作業により複数のネットワークやサーバー機器の電源断が発生し、通信ができない状態が続きました。メッセージの送受信や通話ができない等の状態となり、日本国内に加え海外のユーザーも影響を受けました。LINEは15時9分に障害が発生した後、事象を認識してただちに状況確認や対応策を講じ、16時1分にはトーク機能や通話機能をはじめとする障害は概ね解消いたしました。17時21分には、プロフィール更新や既読表示、グループ参加等の障害も復旧いたしました。現在も引き続き、障害影響範囲の確認と一部復旧作業を続けております。なお、誠に申し訳ございませんが、15時9分から16時1分の期間にメッセージを送信された方は不達の可能性がございますので、お手数ですが、再度、送信をお願いいたします。すべての障害復旧が完了しましたら、こちらのサイトにてお知らせいたします。

#### 【障害の影響範囲】

- ・対象：2021年4月12日 15時9分から16時1分の間に「LINE」アプリを利用したユーザーの皆さま（国内及び海外）
- ・障害影響時間：2020年4月12日 15時9分から16時1分（障害影響範囲の確認と一部復旧作業は継続中 完全復旧済み）
- ・障害発生原因：当社が利用している国内データセンター提供事業者が、本来電源停止を伴わずに電源設備のメンテナンス作業を行うべきところ、同社の誤作業により複数のネットワークやサーバー機器の電源断が発生したため
- ・障害範囲：「LINE」にログインして利用する機能・サービス全般（一部を除く）、アカウントの作成

### ●発生事象

- ・複数のNW/サーバ電源断発生
- ・メッセージ送受信不可：約50分間

# サーバールーム障害への対応策

## 案1.集約NW機器の サーバールーム分散配置

## 案2.アプリの分散配置が可能な インフラ基盤の構築

### Pros

- ネットワークアーキテクチャの大幅な変更が不要

- ネットワークやサーバが落ちてもアプリケーションは止まらない。

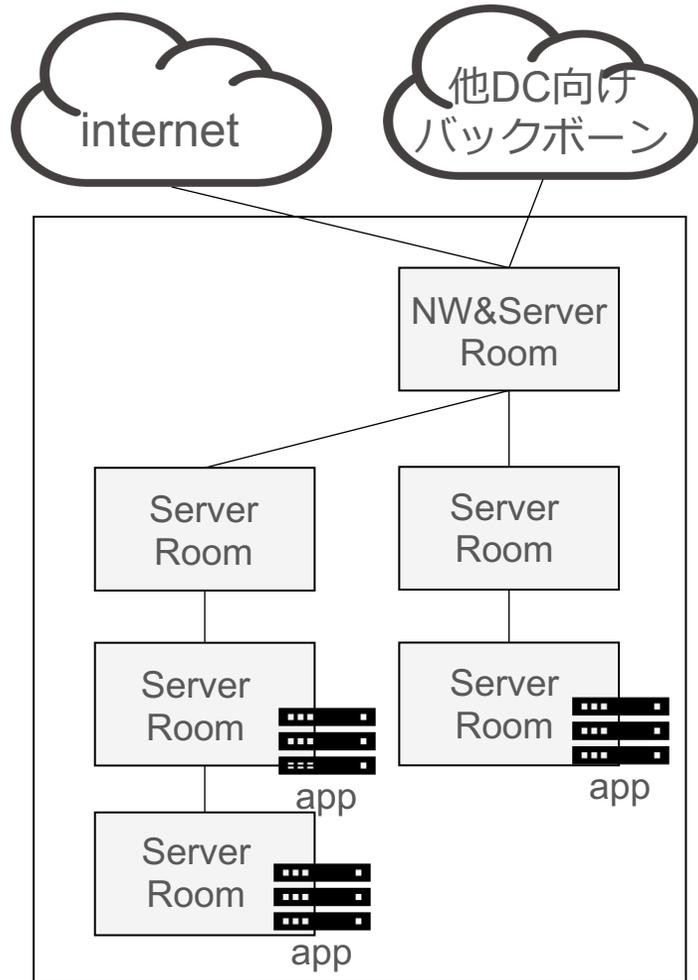
### Cons

- NW機器のオペミスの内容によってはサーバールーム単位でNWが全断し、アプリケーションは止まる。
- ラック電源断によりサーバが落ちる場合はアプリケーションは止まる。

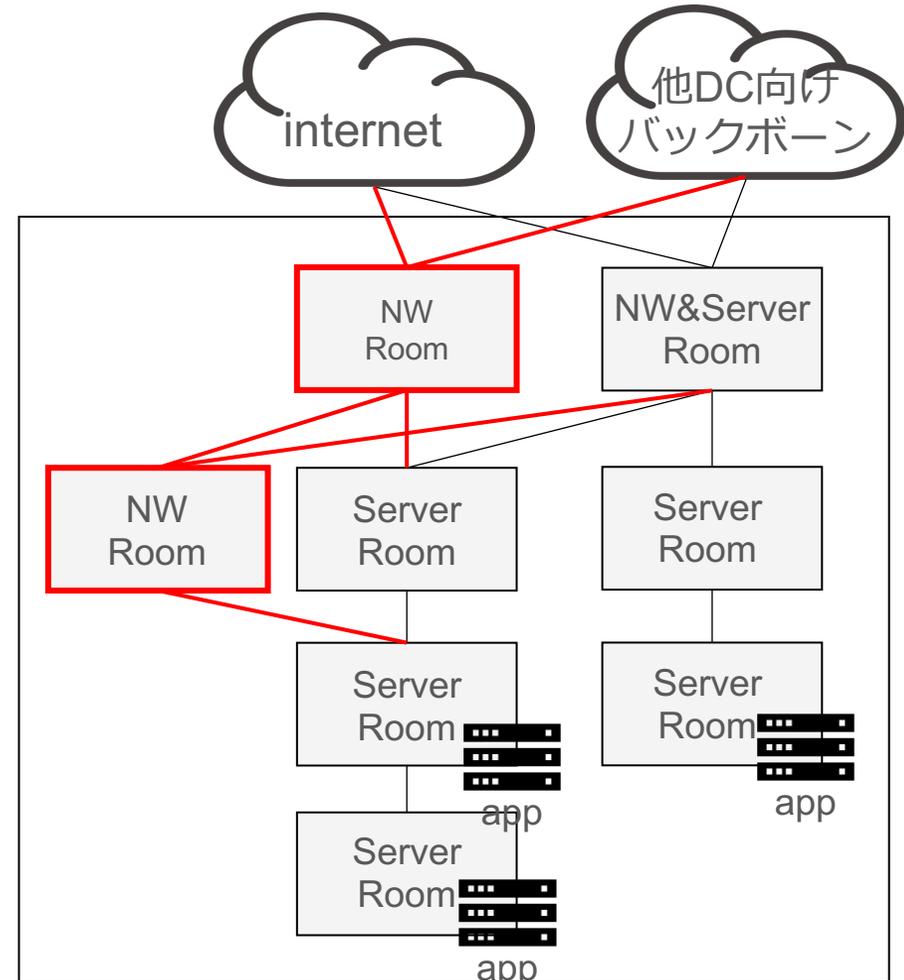
- ネットワークアーキテクチャの大幅な変更が必要
- ネットワークアーキテクチャ変更に合わせてNFV/Platform機能の追従も必要
- アプリケーションの分散配置が必要

# 案1.集約NW機器のサーバルーム分散配置

変更前

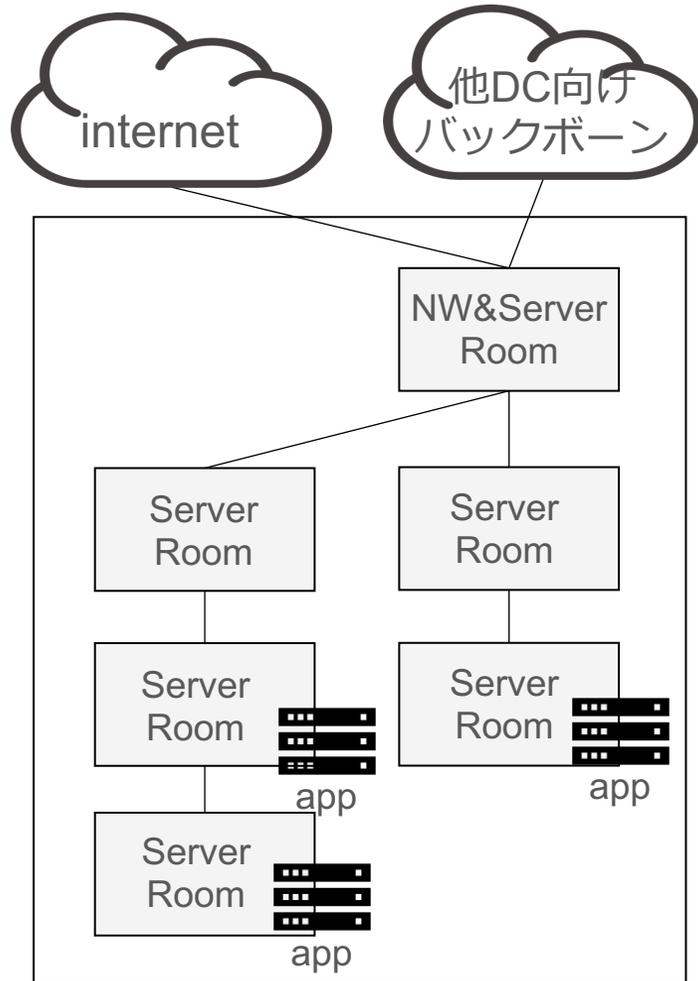


変更後

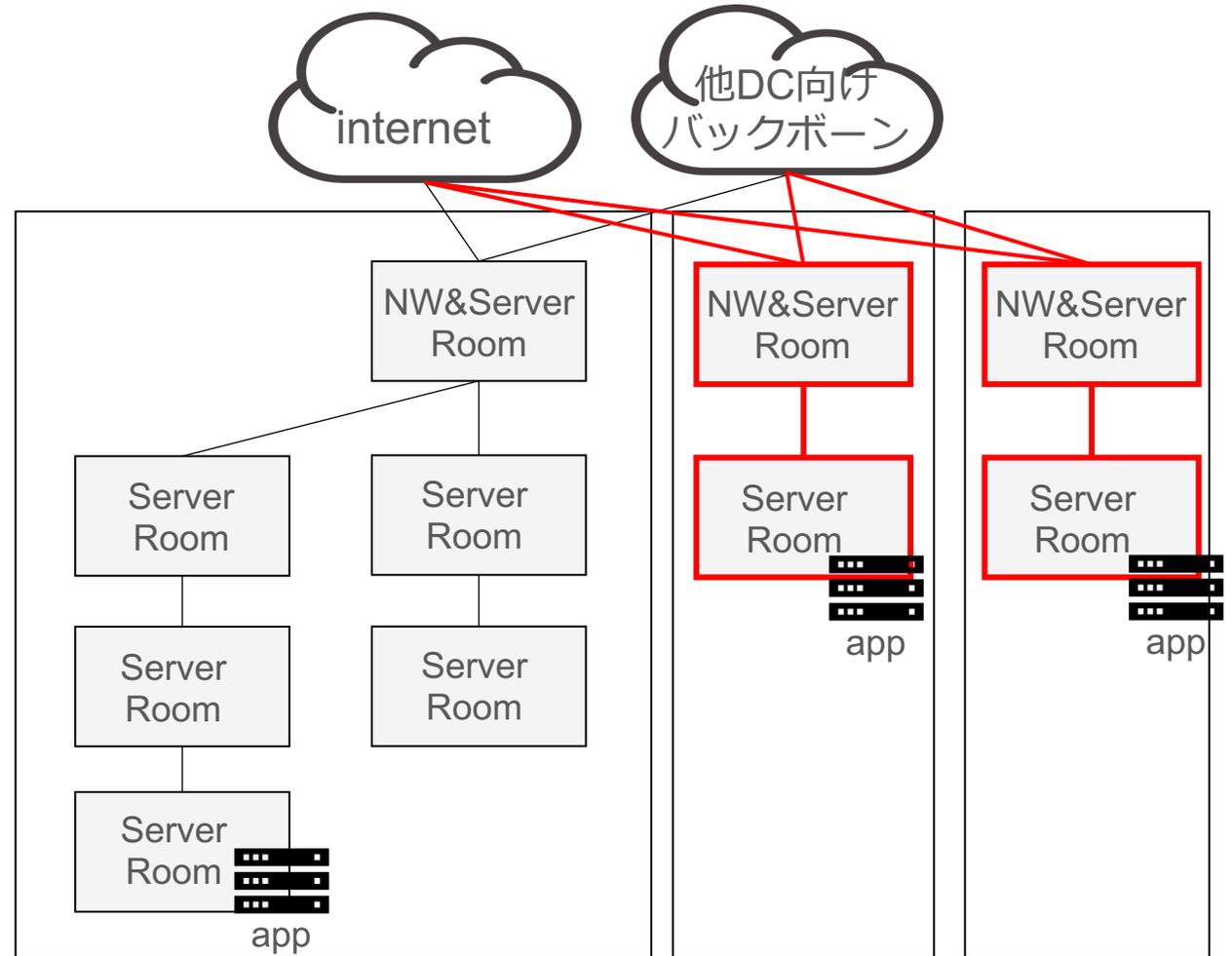


# 案2. アプリの分散配置が可能なインフラ基盤の構築

変更前



変更後



# サーバールーム障害への対応策

## 案1.集約NW機器の サーバールーム分散配置

### Pros

- ネットワークアーキテクチャの大幅な変更が不要

### Cons

- NW機器のオペミスの内容によってはサーバールーム単位でNWが全断し、アプリケーションは止まる。
- ラック電源断によりサーバが落ちる場合はアプリケーションは止まる。

## 案2.アプリの分散配置が可能な インフラ基盤の構築

- ネットワークやサーバが落ちてもアプリケーションは止まらない。

- ネットワークアーキテクチャの大幅な変更が必要
- ネットワークアーキテクチャ変更に合わせてNFV/Platform機能の追従も必要
- アプリケーションの分散配置が必要

# アプリの分散配置が可能なインフラ基盤構築の準備



## Datacenter/Network

- 互いに障害の影響をおよぼさないインフラの単位(Availability Zone)を3つ※以上構築  
※典型的なクラスタソフトウェアを意識
- Backbone networkの構成変更

## Network Function (DNS/LB/NAT)

- 各種アプリケーションが利用するNetwork Functionを各AZに配置する。
- AZ障害時は他AZのNetwork Functionに自動的に切り替える。

## Platform

- 社内アプリケーション開発者が複数のAZに容易にアプリケーションをdeployできるようにする。
- 各AZに必要なサーバリソースを確保する。

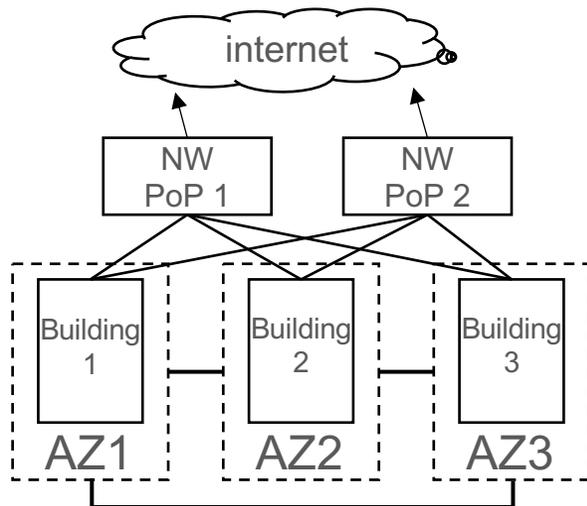
# AZの定義 および Multi-AZsモデル

- AZ(Availability zone)

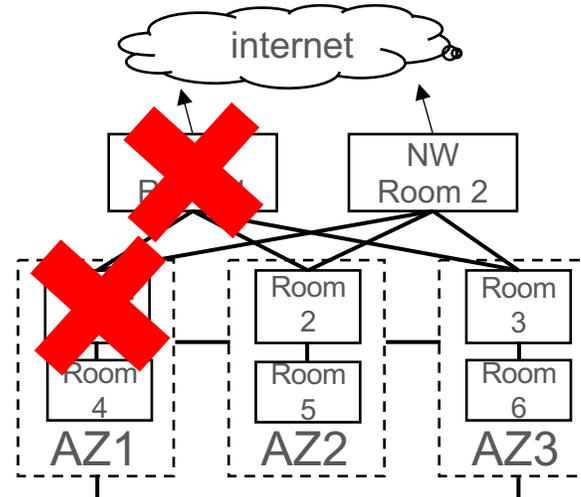
- 次の要素を他の単一のAZに依存していない、物理インフラの単位
  - 電源
  - 空調
  - internetへの接続性
  - 他のAZへの接続性

- Multi-AZs infrastructure model

Building-Aware AZ model



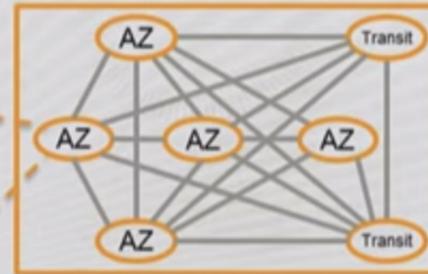
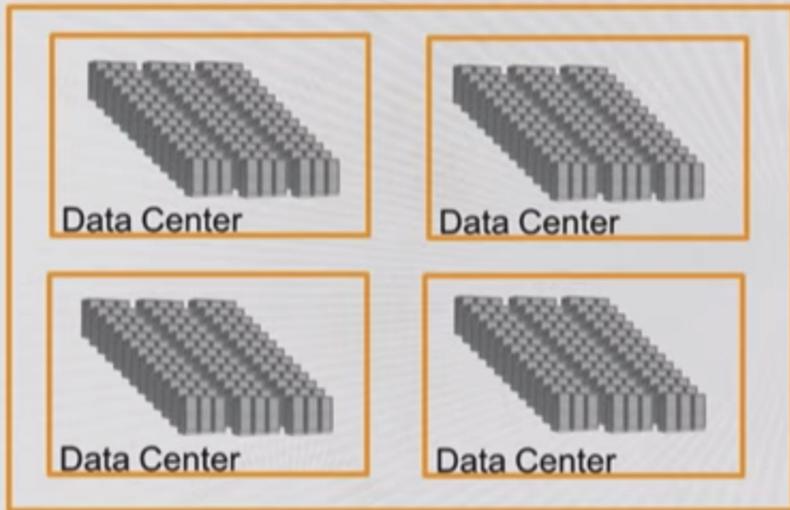
Room-Aware AZ model



アプリケーションは3つのAZに分散配置する前提

# (参考)AWSのAZ

## Example AWS Availability Zone



- 1 of 28 AZs world-wide
- All regions have 2 or more AZs
- Each AZ is 1 or more DC
  - No data center is in two AZs
  - Some AZs have as many as 6 DCs
- DCs in AZ less than  $\frac{1}{4}$  ms apart
  - Don't need inter-AZ independence
  - Do require low latency & full B/W

各AZは1つ以上のDCで構成

# AZ infrastructure model choice

Regionごとのインフラ規模に応じて  
どちらのデザインを採用するかを選択

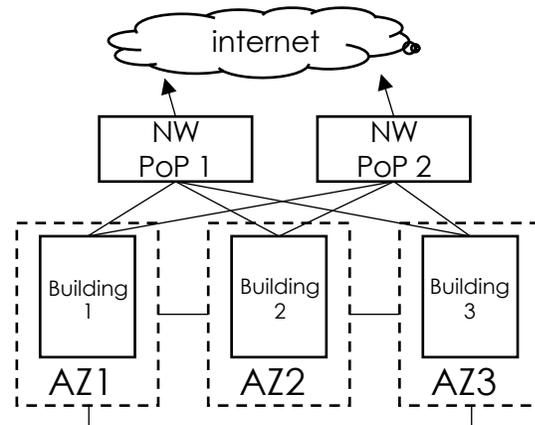
## Tokyo Region

- Building-Aware AZ modelを採用

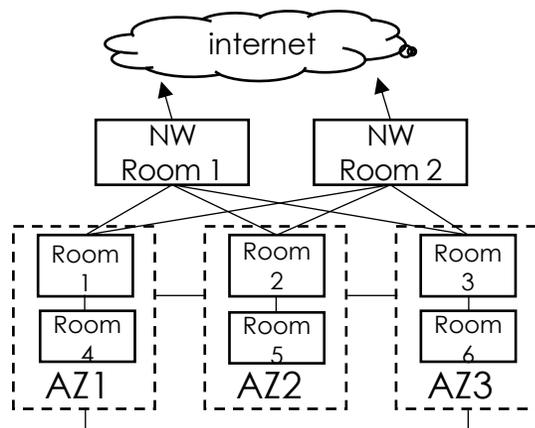
## その他のRegion

- Room-Aware AZ modelを採用予定

Building-Aware AZ model



Room-Aware AZ model



# Conclusion

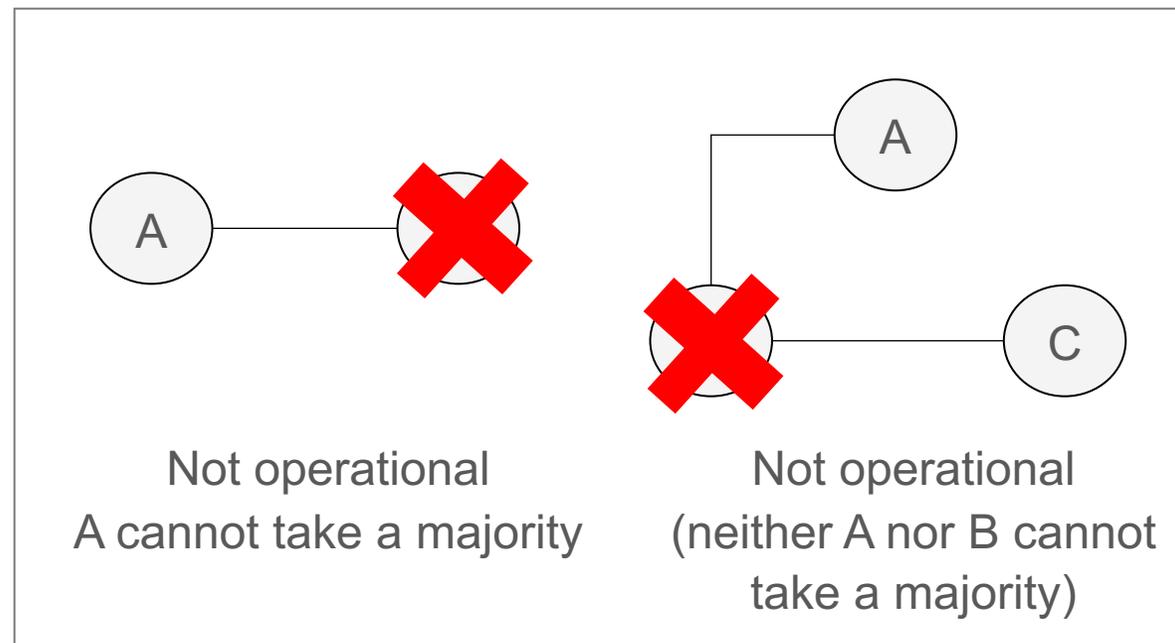
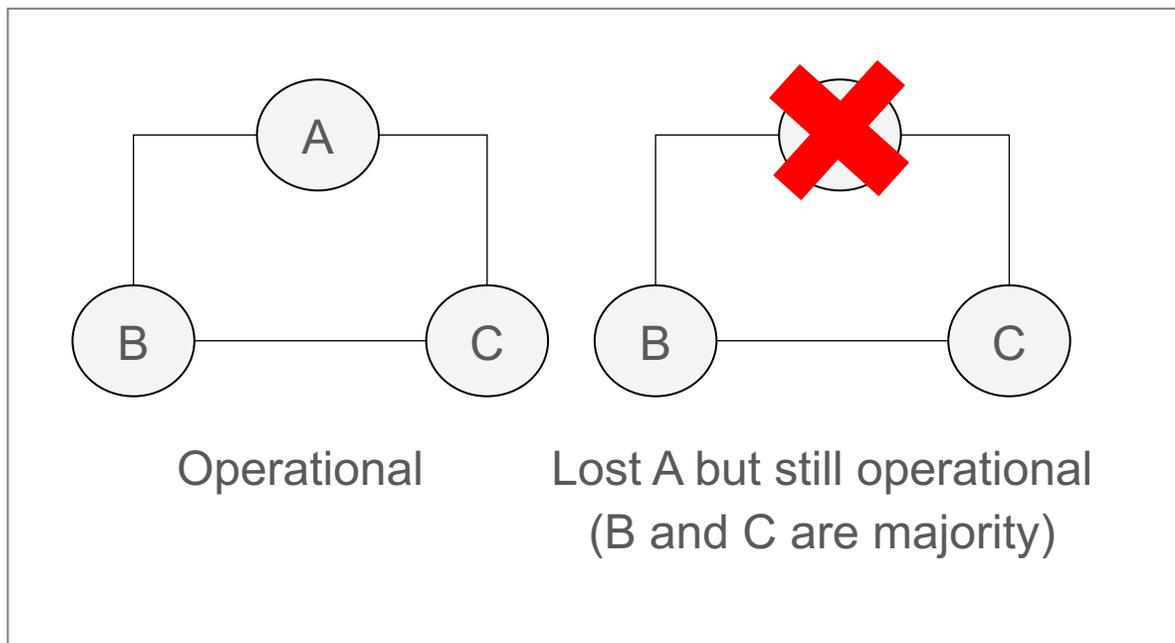
- LINEアプリでメッセージが送受信できなくなる大規模な障害を経験
- 対策としてアプリケーション分散配置可能なインフラ基盤への移行を決断
- LINEの規模にあったMult-AZsインフラ基盤のモデルを決定し構築を進行

**THANK YOU**

# appendix

# (参考)3AZを準備する理由

- Region内に3AZを準備する理由
  - 典型的なクラスタソフトウェアは、3つの互いに依存しないAZが存在することを期待している。
  - AZの数が2つであったり、3つ以上であっても互いに依存している場合、障害時にsplit brainが発生する。



# (参考) Region , DR

- リージョン

- 1つ以上のアベイラビリティゾーンの集合
- リージョン内の任意の2点間のレイテンシはおおむね1ms以下
- 物理的にはおおむね100km圏内

- DR

- 激甚災害を想定し、東京から十分離隔がある拠点をDRサイトとして準備済。

ディスカッション

- DC/NW観点
  - 設備構築する際にどのような障害が発生することを前提としているか？
    - 1ノード障害
    - 2重障害
    - ラック単位障害
    - ルーム単位障害
    - ビル単位障害
  - internet向けのNW PoPはどのように冗長しているか？
    - 地域内冗長
    - 東阪冗長
    - etc
- NFV観点
  - 新規のサービスインやマイグレーション時のサービスインをどう対処しているか？
- Platform観点
  - サービリソースのキャパシティプランニング方法