

壊れながらも使えるネットワーク

Matsuzaki 'maz' Yoshinobu

<maz@ij.ad.jp>

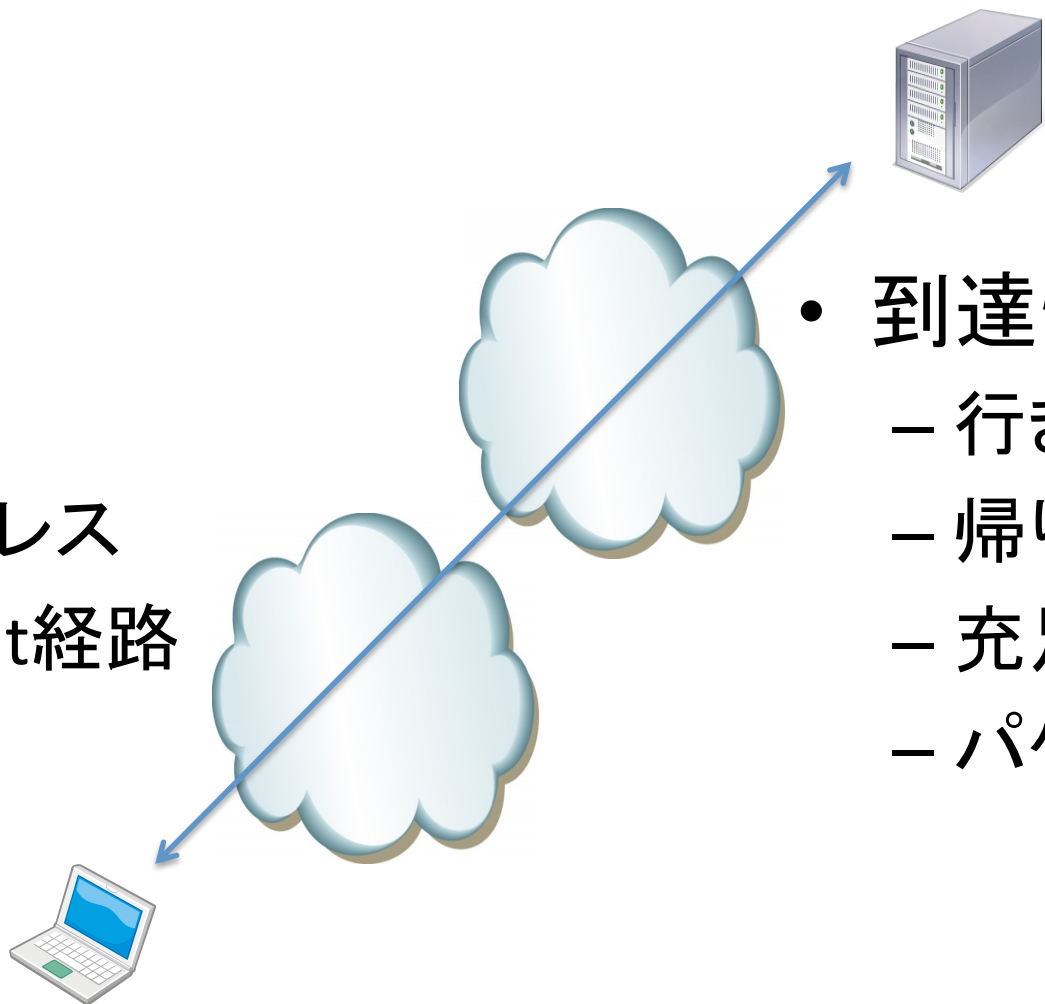
使えるネットワーク

- 接続性

- 認証
- IPアドレス
- default経路
- DNS

- 到達性

- 行きの経路
- 帰りの経路
- 充足した帯域
- パケットフィルタ



壊れる

- 回線切断
- 機器故障
- 電源障害

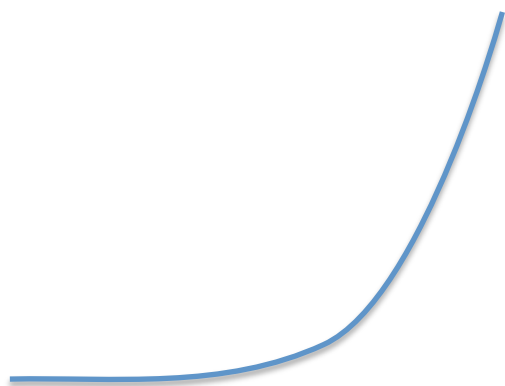


何とかして**対障害性**を高める

耐障害性

高信頼化

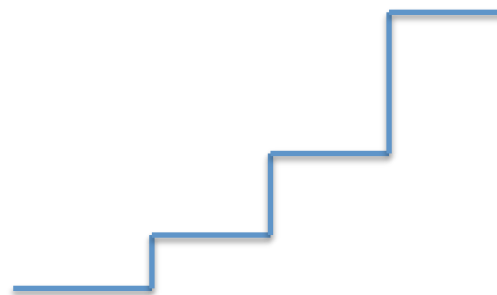
- 壊れない機器
- バグ無しソフトウェア
- 落ちない回線



個で頑張る

冗長化

- 機器の多重化
- 異経路の回線
- 余剰帯域の確保



全体で頑張る

冗長設計

- 障害を想定
 - どこ&何が壊れるか
- 設計
 - 冗長化する箇所
 - 迂回に関わるプロトコル/制御
 - コスト
- 冗長設計は知見の宝庫
 - どんな冗長設計してますか？

冗長化のコスト

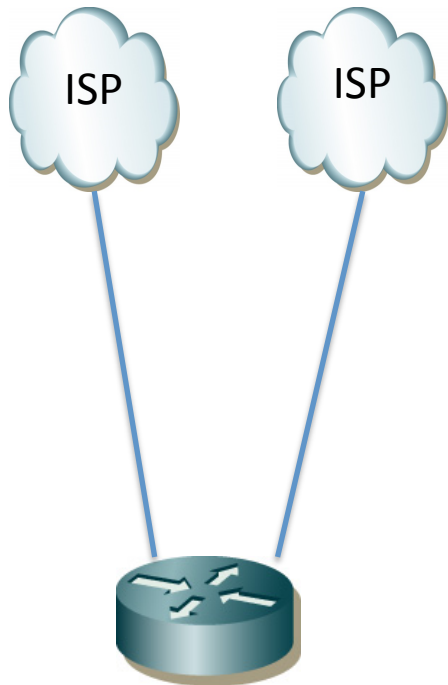
- 費用 対 効果
 - 余剰帯域や余剰設備 → コスト
 - 冗長による耐障害性 → 効果
 - 主要なコストは時代とともに変化
 - 回線費用
 - 機器費用
 - 電力や場所代の費用
- 時々見直した方が良い

冗長構成

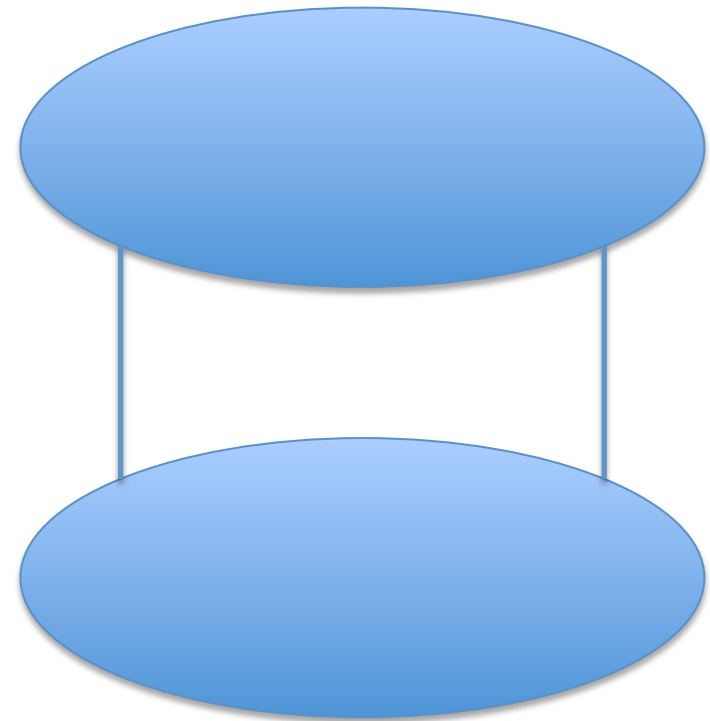
- 大きな迂回は影響が大きい
 - 他のネットワークへの影響
 - 網内での迂回経路の見積もり
- 必要十分な迂回を行えば良い
 - 全体としてコストも小さくなる
 - 設計の見通しも良くなる

大きな迂回が発生しうる所

複数ISPからのトランジット



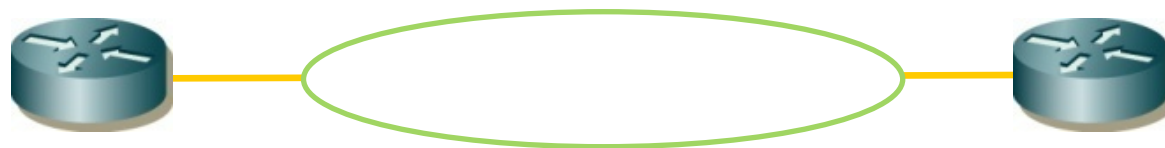
AS間の相互接続



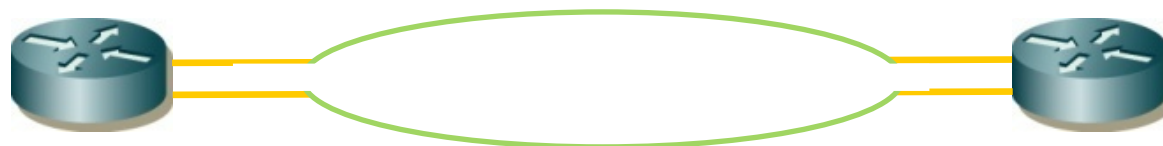
個の信頼度と全体のバランス

- リングプロテクション 対 IP迂回

リングプロテクションで1回線



ハンプロテクションだけど2回線



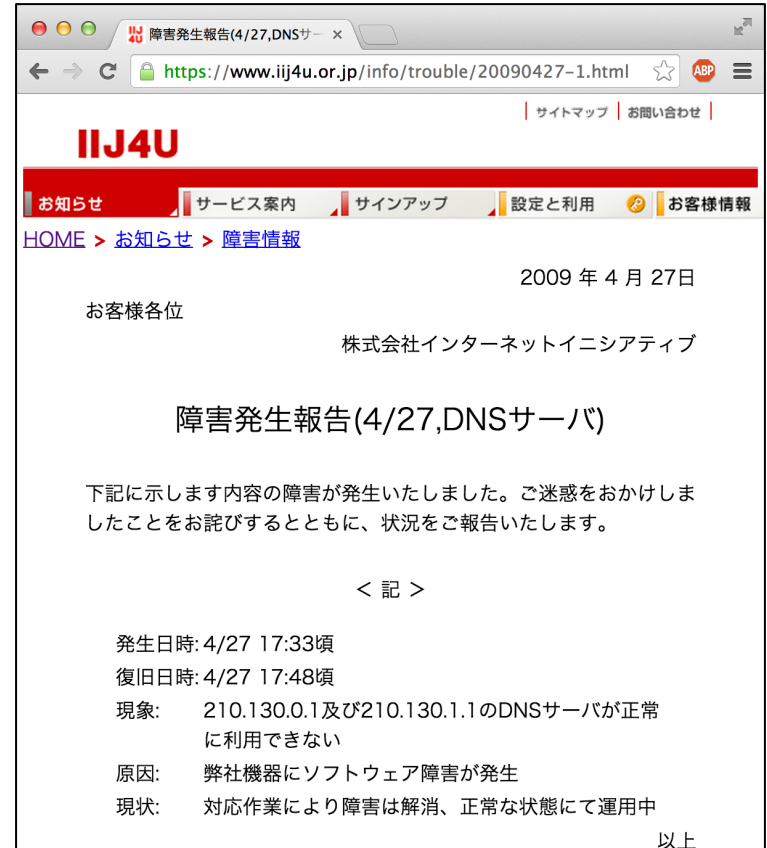
- 太平洋など長距離区間では、プロテクションのコストが高い

冗長化の弱点

- コストがかかる
 - 余剰の設備を維持しないといけない
- 複雑になる
 - 構成や切り替わりの状態
- 想定した障害にしか対応できない
 - 想定外の障害が起こると困る

障害事例

- 個人向けサービスのキャッシュDNS障害
 - 2009年4月
- 全システムで障害発生
 - 最初に片系で障害発生
 - 次いでもう片系でも発生
 - 両システムが15分程度障害

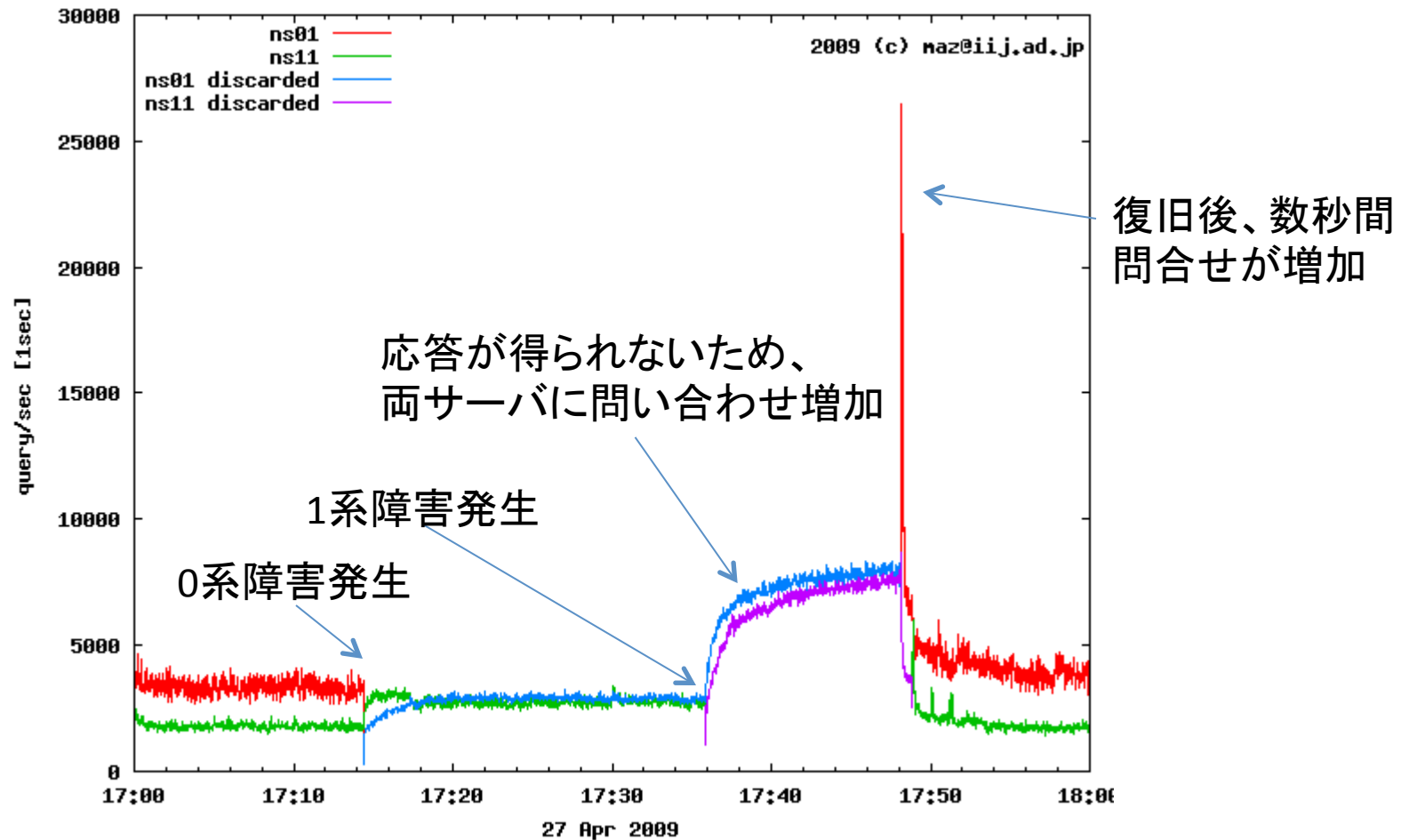


The screenshot shows a web browser window with the URL <https://www.ij4u.or.jp/info/trouble/20090427-1.html>. The page header features the IJ4U logo and navigation links for 'お知らせ' (Notice), 'サービス案内' (Service Guide), 'サインアップ' (Sign Up), '設定と利用' (Settings and Usage), and 'お客様情報' (Customer Information). The main content area is titled '障害発生報告(4/27,DNSサーバ)' (Outage Report (4/27, DNS Server)) and is dated '2009年4月27日'. The text addresses customers, stating that a DNS outage occurred on the specified date and time, and that the service is now restored. The details of the outage are as follows:

項目	内容
発生日時	4/27 17:33頃
復旧日時	4/27 17:48頃
現象	210.130.0.1及び210.130.1.1のDNSサーバが正常に利用できない
原因	弊社機器にソフトウェア障害が発生
現状	対応作業により障害は解消、正常な状態にて運用中

The page concludes with '以上' (End).

キャッシュDNSと障害



知見

- 片系統の障害時はサービス継続できていた
 - 異常なクエリ傾向無し → 想定通り
- 両系統の障害時ではクエリの増加を確認
 - 再問い合わせが急増していた
- 復旧時には瞬間的に問い合わせが増加
 - 通常時の7倍程度

インターネットと障害

- インターネットのほとんどは他者による運用
 - 自分の運用範囲はごく一部
- 誰がどこでどんな障害起こすか分からない
 - 何が起こるかわからない
- しかも変動が続いている
 - 相互接続、ポリシ

困った

動的経路制御バンザイ

- 想定外の障害でもまだ動くかも
 - 全障害を想定して設計している訳じゃない
 - 使える経路があれば、勝手に使ってくれる
- インターネットで僕たちが取りうる手段
 - 充実した相互接続
 - 異経路による余剰帯域

(仮)復旧を目指す

- 接続性の提供
 - 認証、IPアドレス、DNS
- 到達性の確保
 - 疎通の担保
 - 輻輳の改善
 - 冗長構成の回復
- どこから手をつける？
 - ユーザの利用方法に依存

流量

prefix	実トラヒックの占有率
ふがふが.135.0/24	5.46%
もふもふ.0.0/16	4.99%
ふよつと.0.0/16	3.87%
もよつと.128.0/17	2.29%
もがもが.20.0/22	1.92%
ふがもが.0.0/16	1.72%

- ブロードバンドユーザ向けトラヒックを24時間観測し、経路毎にトラヒックを累積
- ちなみに全517528経路中の63176経路からトラヒックを観測していた

DNS

- サーバにアクセスする際にDNSで名前解決
 - いっぱい引かれている名前はよく使われている
- DNSで応答されるAレコードを調べてみよう
 - IPアドレスを経路情報にmap
 - みんながよく使うprefixが分かるはず！

みてみた

- コンシューマ向けキャッシュDNS
- 24時間でのべ 464,750,044 のIPアドレスを応答
- 経路情報にmapすると 44328経路
 - full routeは517528経路
 - ざっくりfull routeの9%弱ぐらい

利用状況

prefix	DNSでのヒット率	実トラヒックの占有率
ほげほげ.0.0/16	28%	1.3%
ふがふが.135.0/24	8%	9.1%
ほにより.148.0/22	2.4%	0.1%
ほげっと.0.0/14	2.3%	0.002%
ふがっと.0.0./13	2.2%	0.001%
むごむご.108.0/22	1.7%	0.1%
むがっと.128.0/17	1.3%	0.6%

トラヒック占有率との相関は無さそげ

復旧の優先順位

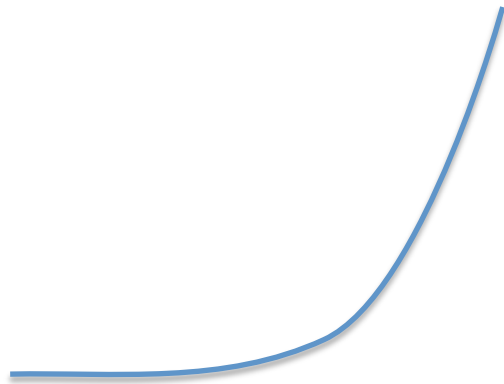
- 帯域の大きな回線から直していくのが良い
 - 主要な設備
 - 多くの人が利用していると考えられる
- と思ってるんですが、何か知見ありますか？

聞いてみたい事

- 冗長設計
 - 考え方とか、構成とか
 - 困ってる事とか
 - みんなが忘れてそうな障害想定とか
- 復旧の優先順位
 - 何か知見ありますか？

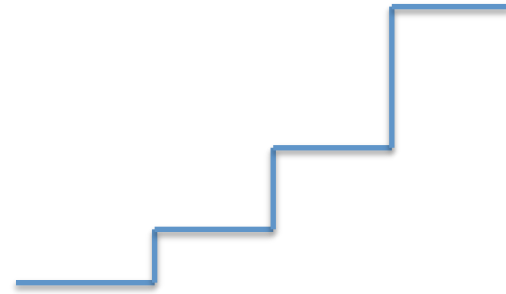
個の信頼度と全体 その2

- 高信頼化



個で頑張る

- 冗長化



全体で頑張る

僕たちが、「全体」だと思っているサービスや機能は他者とか利用者から見ると「個」かもしれない

データの多さと情報量

国民へ発信する重要情報のフ x

Japan Agency for Local Authority Information Systems [JP] <https://www.j-lis.go.jp/lasdec-archive/cms/12,22060,8...>

通知日 平成23年3月18日 20時0分

自治体CEPTOARとしての業務

件名

(参考) 国民へ発信する重要情報のファイル形式について

内容要旨

震災発生から時間経過とともに、全国民が注目・閲覧を要する重要コンテンツ、政府、地方公共団体から提供される情報の多くが、PDFやExcelファイル形式で情報配信されております。インターネットを通じて多数の国民に円滑に閲覧していただきたいところですが、アクセスが集中し、PDFやExcelファイル形式の場合、容量が大きく、サーバー・回線リソースを圧迫し、重要情報が閲覧できない事象が頻出しています。

また、被災してPCが故障していたり、PCを持っておらず、携帯電話で情報を確認している方もおり、携帯電話の細い回線で容量の大きなファイルを開覧することが困難なケースもあります。

このため、現在各団体がアップロードしておられる情報は、1分1秒を争うものも多々あるかと存じますが、より多くの方に簡易に情報を受け取っていただけるよう配慮した情報提供も心がけるべきと思われまます。

是非アップロードするファイルのデータ形式についてご一考いただき、国民への円滑な情報提供をご検討いただけますと幸いです。

ご参考まで、具体的なファイルの公開方法（混雑時にそれらの閲覧でしのげるようにする予防的措置）は以下の通りです。

1. アップロードするファイルはPDFだけではなく、誰もが見ることができ、小容量な「HTML形式（テキスト）」のファイルも公開する。
2. 紙資料のスキャンファイルはPDFではなく、誰もが見ることができて、再利用が容易な「JPEG形式（静止画像データ）」とする。
3. 表形式データのファイルは、容量が大きくなるExcel形式ではなく、どのソフトでも開くことができ、再利用が容易な「CSV形式（表計算ソフトなどで開くことを前提に、データをカンマ(,)や、スペース、tab等で区切って並べたテキスト形式）」とする。

【具体的な方法例】

ワード・エクセル文書をHTML形式で保存する方法

「ファイル」→「名前を付けて保存」→ファイルの種類にて『Webページ(*.htm,*html)』で保存

みんなに優しい

- 端末に優しい
 - 簡素なファイルフォーマット
 - 再利用可能な形式
- ネットワークに優しい
 - 必要最小限で通信が完了する

いつもの連絡手段を試すよね

- “普段”使っているアプリケーション
 - 電話
 - SNS
 - マイクロブログ
 - チャットツール
- 時代とともに変わる
- それぞれのコミュニティによって手段は異なる

非常時の通信モード

- 輻輳の発生
- 有限のバッテリー
- こんな環境でも、“いつもの”アプリケーションが使えると嬉しい
 - 多少の機能縮退を我慢するとして

例えばこんなこと

- 輻輳対策
 - アクセスの集中に耐える
 - 最小限の packets 数で通信が完了する
 - 過度のデコレーションを諦める
 - ストア & フォワード強いね
- バッテリ対策
 - CPUに優しいファイル形式を使う
 - 人間が迷わない様な情報メニュー

非常時の通信ってどう考えますか？

- サービス提供者に提言することある？
 - 非常時モードいかがでしょう
- 利用者にお問い合わせすることある？
 - どんな利用の仕方が困るか
 - どんな利用方法をして欲しいか
 - 非常時には非常時のアプリケーションに誘導？
- 僕たちで出来る事がありますか？
 - 公平制御？

まとめ

- 壊れそうなところは冗長化
 - 依存関係を見落とさない様に
 - 個別の信頼度が高いと忘れがち
- 実はもっとできる事があるかもしれない
 - しかも日常の利用方法のままに
 - 情報提供は再利用や弱い端末に配慮を
 - アプリケーションは非常時モードがあると嬉しい