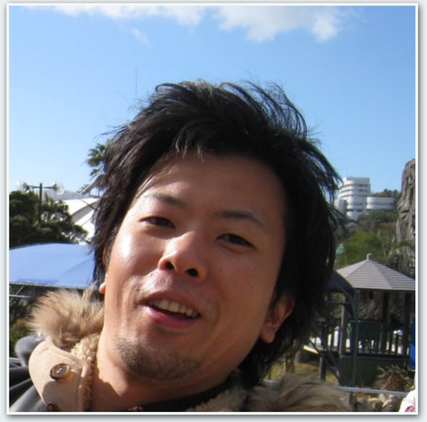
IX・ISP ネットワーク品質の守りかた



小島 慎太郎

株式会社コーダンス
代表取締役・ネットワークエンジニア

codeout

http://about.me/codeout

2004 ISP ntt.net

2009 IX JPNAP

2014 独立
ISP・IX・コンテンツ事業者
の技術支援

本日のゴール

- IX・ISP 事業者が どのようにしてネットワーク品質を維持しているか、事例について知ってもらう

- 事例から、何が学べるかをまとめる

IX・ISP の話を聞いてきました 👂

IX

- その国では最大級
- 国に閉じる国内事業者
- エンジニア 約20人
  (NOC 業務除く)

ISP

- 世界で最大級 (Tier1)
- 各国に拠点を持つ
  国際事業者
- エンジニア 約15人
  (+ NOC 35人)

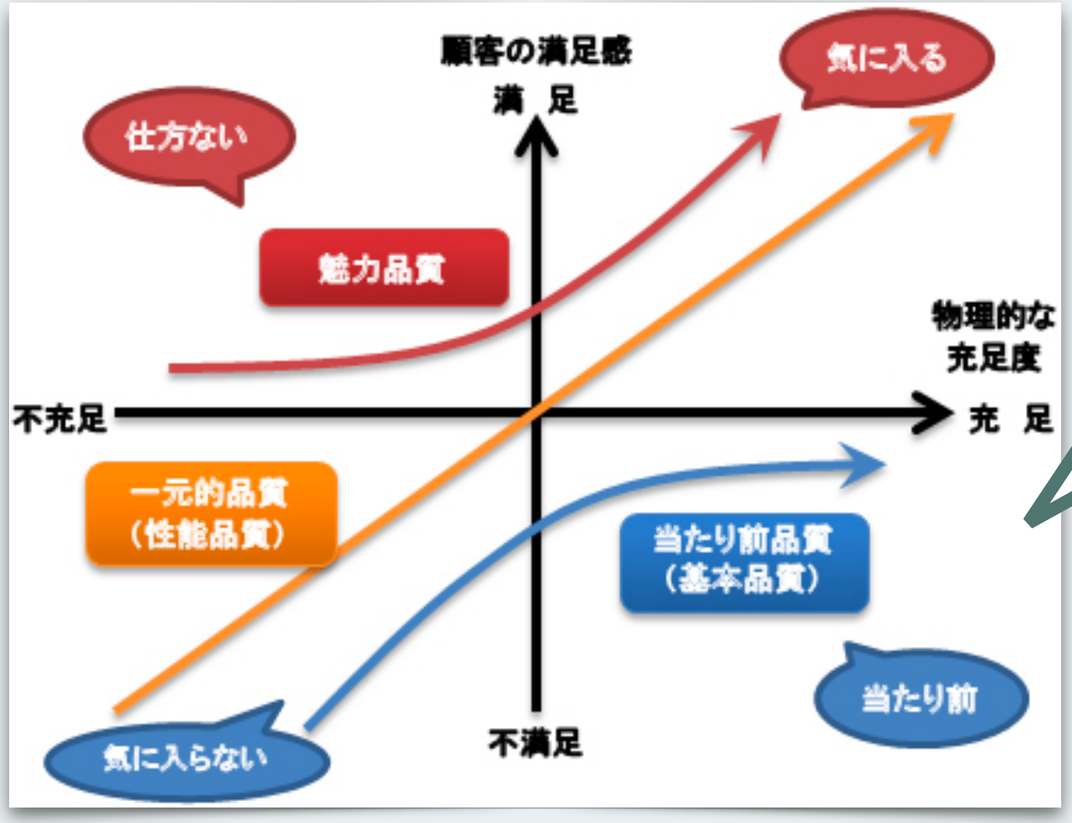
あるIX事業者の例

冗長設計と事前検証にコストをかけ、
監視・計測 をがんばらないパターン

IXのビジネス背景

1. **ユーザーにとって「ピアしたい事業者が集まっているか」が最大のポイント**
   - ピアしたい人気事業者 = だいたい大手 = 彼らが最も気にするのは価格
2. **価格に対する圧力 (ユーザーの期待) が強い**
   - ピアリングのためのインフラ → ユーザーにとってのコストが
     PNI フルメッシュ > IX でないと成立しない
   - ネットワーク機能がEthernet！ シンプルすぎる。機能で勝負しにくい
3. **価格以外での勝負ポイント**
   - プラットフォームとしての売り方 (IXユーザーが、IX上で自社サービス販売)
   - 付加機能 (Route Server、Traffic 可視化、GPRS Roaming など)
   - サポート (メンテウィンドウを融通してくれるとか)

IXビジネス背景における品質



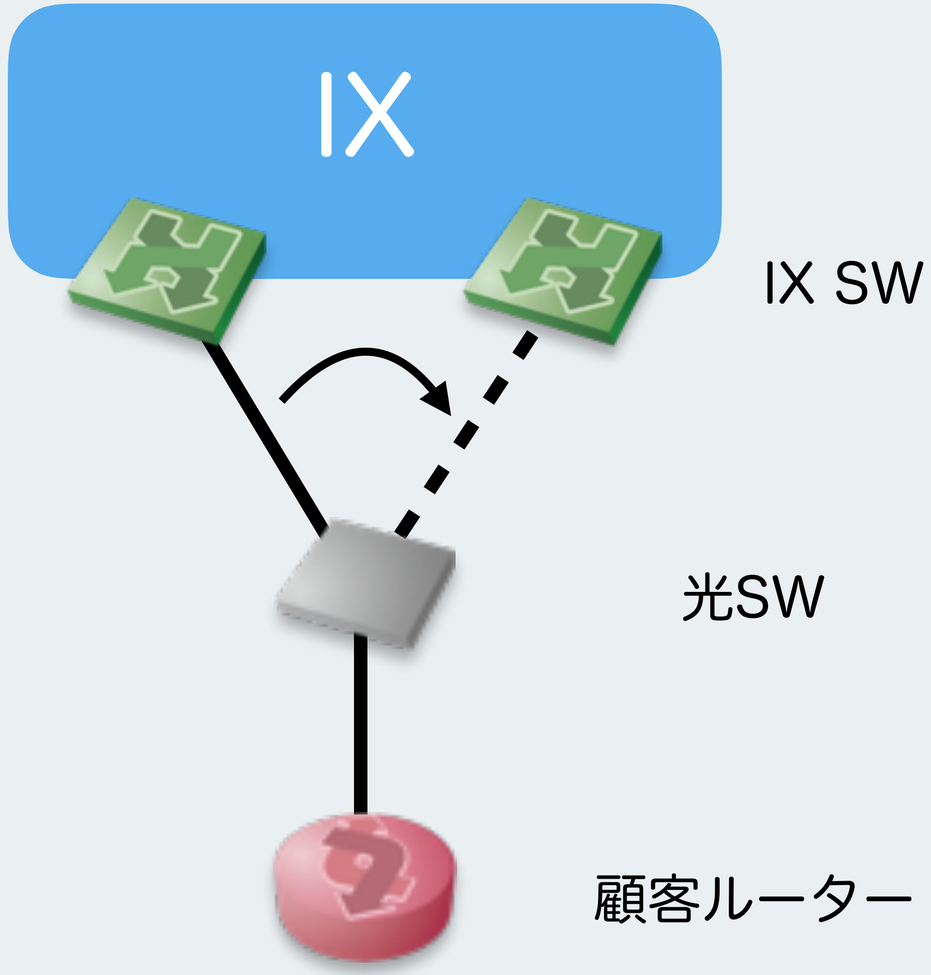
多くのユーザーにとっての
ネットワーク品質

コストをかけずに、
ネットワーク品質を下げ止める
(上げても効果は薄い)

狩野モデル
https://www.juse.or.jp/departmental/point02/08.html

ある IX 事業者の選択

- **価格競争力を上げるために**
  - シングルベンダーによるOPEX 削減
  - 自動化によるOPEX 削減
- **シングルベンダーリスクの低減**
  - 完全な冗長化 (光SW)
  - 事前検証
  - リッチなLAB



Q: 新スイッチ・OS の検証、なにやってます？

- スネークによるBERT
- LACP、VLAN、QinQ などEthernet feature 確認
- STP 系のプロトコル動作確認
- MAC制限機能確認
- Hardware の冗長確認。ブチ切りしてみるとか
- オーバーサブ時のふるまい確認・buffer 深さ確認
- L1 (DWDM・光SW) との相性チェック
- CLI、API、syslog、SNMP、sflow などの管理系
- バグのregression テスト

- 👈 のようなこと、おそらく経験がある人が想像できることは大体やっている
- 細かなテストケースはない。項目リスト spreadsheet があるくらい。毎回網羅するわけではなく、状況に応じて改変して対応
  - 例: 不具合が見つかった場合、そこを掘り下げる
- Regressionテストは、ベンダーに依頼

Q: 新スイッチ・OSのデプロイ
気にしてること、工夫していることは？

- Optical SW によって完全に冗長がとれている = **設計による工夫**。
  - HW / SW デプロイがやりやすい
  - = 問題があればすぐ戻せる / SPOF をなくす

Q: 通常運用時、なに監視してます？

- Device Perspective
  - Traffic
  - Optical Power
  - CPU Utilization / Temperature
- Network Perspective
  - 全スイッチにserver 配置、
    そこから全IPアドレスにping
  - IXP-Watch / arpsponge など
    一般的なIXPツール

- **たいしたことをしていない**
- **逆に、それでどうにかなっている**
- SLA はない

Q: 検証項目・デプロイ方法・監視項目について、なぜそこに至りましたか

- 明確な基準があるわけではない。そのときの担当者が基準を決める
  = 人に依存する
- 「**設計上、多少の見落としは許容できる** (品質を大きく落とす事故が起こらない) ようになっている」というのが我々のポリシー
  - Optical SW、management traffic のoutbound 化
  - 冗長化を常に意識した収容設計
    - 事故の局所化

Q: 検証項目・デプロイ方法・監視項目について、見直していますか

- **定量評価できていないが、heuristic なアプローチでうまくいっている**
  - 実運用に基づく経験を蓄積し、共有 (伝承) されている
  - 育成をしっかりやることで、維持できていると言える
    - しかし、育成を体系的にできているわけではない

Q: ほか、何か参加者に
有益そうなことがあればコメントください🤦

- LAB環境は重要
  - 事前にできる検証の質が圧倒的に違う
  - 場合によってはスケールダウン / VM版もOK

この事例から学べること

- **品質管理上の弱点にフォーカスしよう**
  - この事例でいえば、シングルベンダーであること
- **設計重要**
  - 「何かあってもユーザーに影響しない」の安心感
  - SLA / SLO のような数値目標がなくても、意外とうまくいく
- **残った問題は、しっかり検証する**
  - 許容できることは受け入れ、必要な分をちゃんとやる
  - "しっかり" = よく言えば柔軟、悪く言えば曖昧であるが、そこも設計の安全性でカバー可能

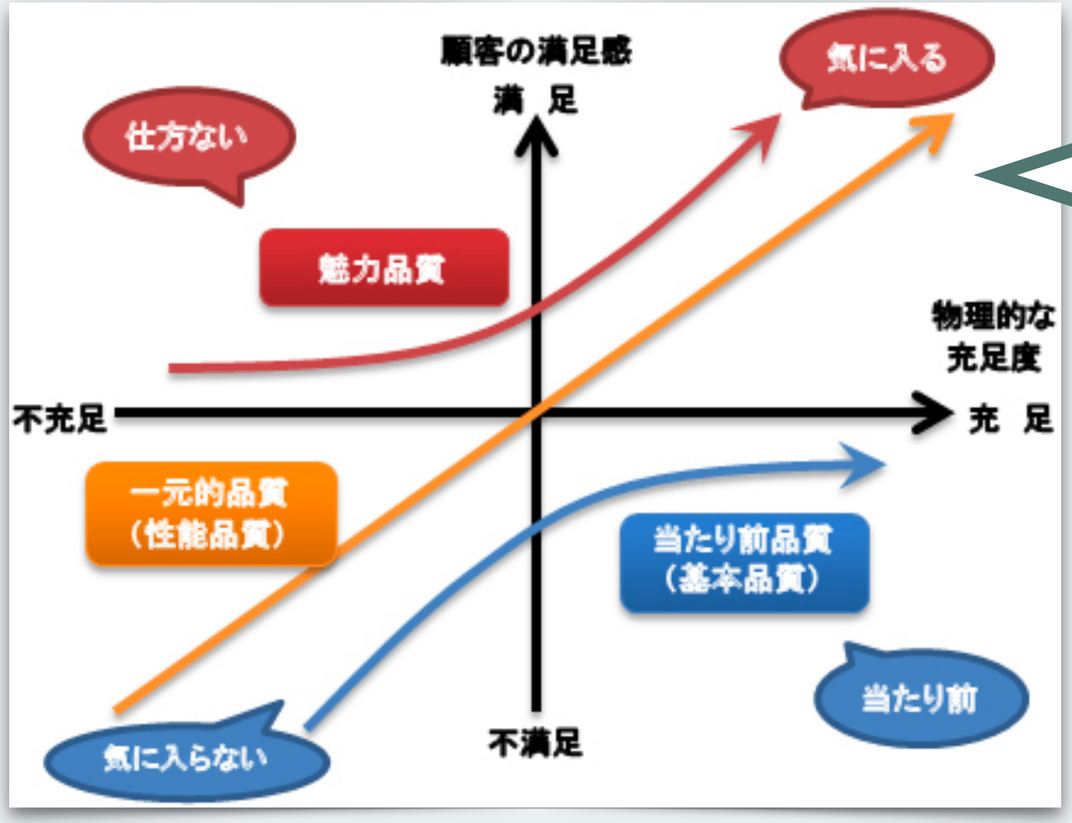
あるISP事業者の例

事前検証はがんばらず、
デプロイ・監視・計測 にコストをかける
パターン

Tier1 ISPのビジネス背景

1. **世界中のASに直接・間接的にTransitを売るか、Peerしないといけない**
   - いろんなユーザー・ピアリングパートナーがいる
2. **価格・ネットワーク品質・付加機能・API、優先度を決めにくいが、やはり価格になりそう**
   - どれを重視するかはユーザーによるが、Tier1 ISP を使うような事業者はISP の使い方を知っていて、高いと買ってくれない

ISPビジネス背景における品質



狩野モデル
https://www.juse.or.jp/departmental/point02/08.html

多くのユーザーにとっての
ネットワーク品質

- "日本のようなサービス品質ではないが、海外での顧客満足度は高い"
- コストをかけずに、
  ネットワーク品質を下げ止める。
  あわよくば上げる。
  (ビジネス上、上げるモチベーションがある)

ある ISP 事業者の選択

- **価格競争力を上げるために**
  - マルチベンダー、適材適所をつきつめたCAPEX 削減
  - 自動化によるOPEX 削減
    - NOC 業務を除き、IP ネットワークエンジニア15人(!) で世界をカバー
- **品質を上げるために**
  - デプロイ・監視・計測にコストをかける

Q: 新Hardware・OS の検証、なにやってます？

Hardware

- 基本 = 既存設備との互換性
- DDoS時のふるまい確認
  - Buffer が溢れたときのパケットDropについて。設計上、それが許容できるか
- Full Route 増加に対する、RE/CPM/RPのScalability

Software

- 基本 = 既存運用機能メインでテスト
  - BGP Scale、MPLS-TE、L2VPN、GRE、VRFなど、
- 将来サービス導入のための評価
  - BFD、QPPB、OAM、RR など
- バックエンドシステムとの互換性
- OSのバグ修正確認

Q: 新Hardware・OSのデプロイ
気にしてること、工夫していることは？

1. LABで基本機能のテスト
2. 商用ネットワークに顧客・peer 収容なしで導入
3. 一部顧客・peer でのfield trail
4. 導入判断
5. まずUSにデプロイ → 徐々に全世界に広げる
   - 「初物を入れてくれるな」「Bug Fix 早く」のような各国の思惑を調整する

Q: 通常運用時、なに監視してます？

- Chassisアラーム、リンク断、メモリ使用量
- ISIS/BGP State
- 商用網内のpacket loss (内部向け)
  - 観測された場合、発生POP間で回線新設
- “サイレント故障” に関連するlog
- SLA (packet loss、jitter、latency)
  - ユーザーの問い合わせ時、「SLA を満たせているか」を見る
- LSPのふるまい
  - LSP 設定変更
  - Primarily LSPの切り替わり
  - LSPあたりのBandwidth
  - 国際回線使用率
- 項目ではないが、Root Cause Analysis をがんばっている
  - 大元のアラームと付随アラームなど関連するアラームをまとめる
  - NOC の反応速度を上げるため

Q: 検証項目・デプロイ方法・監視項目について、なぜそこに至りましたか

- 既存検証項目テンプレート、商用網に必要な基礎機能（日本より少ない）しか検証しない
  - L2VPN、BGP、ISIS、GRE、VRF、VLAN、MPLS-TE、RFC2544など
  - フィールドトライアル時のトラブル・クレームを追加で検証することがある
- **いくら事前に検証しても、商用に入れると問題が起こる**
  - 「基本検証に時間をかけなくていいのでは？」という意見も

Q: ネットワーク品質維持の観点で気にしていることありますか？

- 取れるlog を全部取る
  - MIBで取得できないlogは、定期login + コマンド収集
  - 将来使えそうなものは、手当たり次第
  - トラブル時の原因特定スピードで、ベンダーを超えられる
- NOCのトラブルシュート能力を上げる

この事例から学べること

- **品質管理上の弱点にフォーカスしよう**
  - この事例でいえば、商用ネットワークでしか発生しない問題が多いこと
- **計測重要**
  - 監視・計測・logging にコストをかける
  - Field trial、Canary Deployment が可能なのは、監視・計測基盤があるから

まとめ

やってることが違う→リスクの違い

IX 事業者

- **設計 (冗長性確保) にコストをかける**
- **事前検証 をしっかりやる**
  - 検証項目は柔軟に
    許容できるものはあらかじめ捨てる
  - 抜け漏れさえも許容する
- 監視・計測 はがんばらない

ISP 事業者

- 事前検証 はがんばらない
- **デプロイ・監視・計測 にコストをかける**

この違いは、
- **クリティカルな問題でもコストをかけることで**
  予見可能 / 問題があっても設計で隔離可能 (IX)
- **コストをかけてもクリティカルな問題は**
  予見不可能 (ISP)

という、リスクの質の違い

ネットワーク品質を守るためのヒント

- **どんなリスクに対応するか**
  - トラブルチケットを遡るのがオススメ
- **そのリスクは、制御可能か**
  - 人的・金銭的・時間的制約に見合うか
  - 「何かあってもだいたい大丈夫」と思えるか
    - 「設備が不安」は投資が足りないし、「何が起こるか不安」は検証が足りない
    - 頭よすぎる設計は 検証・保守・情報共有・教育などの運用コストが高く、人の入れ替えに弱い
- **予見できないリスクに備えられているか**
  - 計測し、検知を早くしましょう

Questions ?