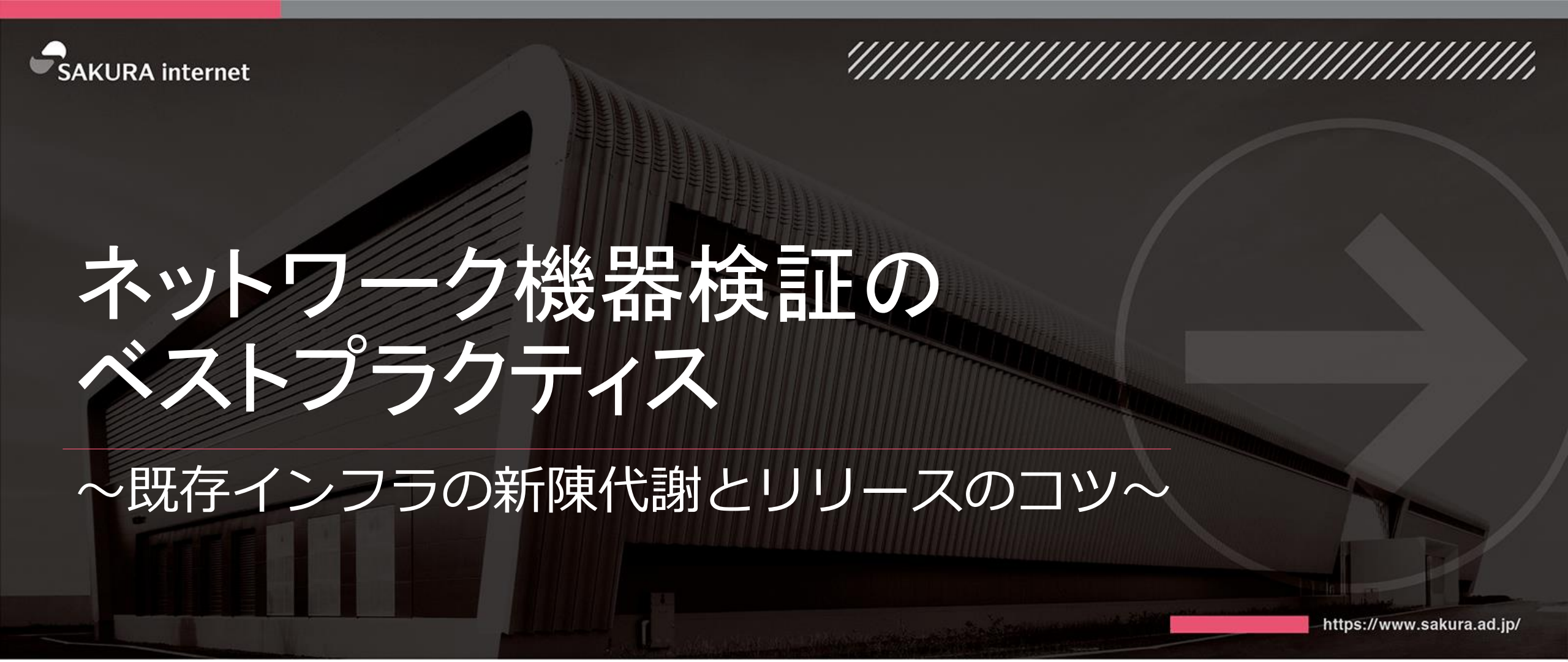


SAKURA internet

ネットワーク機器検証のベストプラクティス

～既存インフラの新陳代謝とリリースのコツ～

https://www.sakura.ad.jp/

DAY

2019/11/28 Internet Week 2019
S12 サービス設計とリリース

DEPARTMENT

さくらインターネット(株)

NAME

大久保 修一

自己紹介

SAKURA internet

さくらインターネット研究所 所属
大久保 修一　@jh1vxw

- 2003年入社、最初はバックボーンネットワークの運用担当
- 2009年よりさくらインターネット研究所
- 2011年よりIaaS「さくらのクラウド」の開発に参加
- コーディング、構築、企画、運用、お客様サポートなどなどを担当
- 当社のサービス間プライベート接続サービスも担当

IaaSの運用視点でお話させていただきます

サービス用機器導入～提供の流れ

場合によっては
設計を見直し

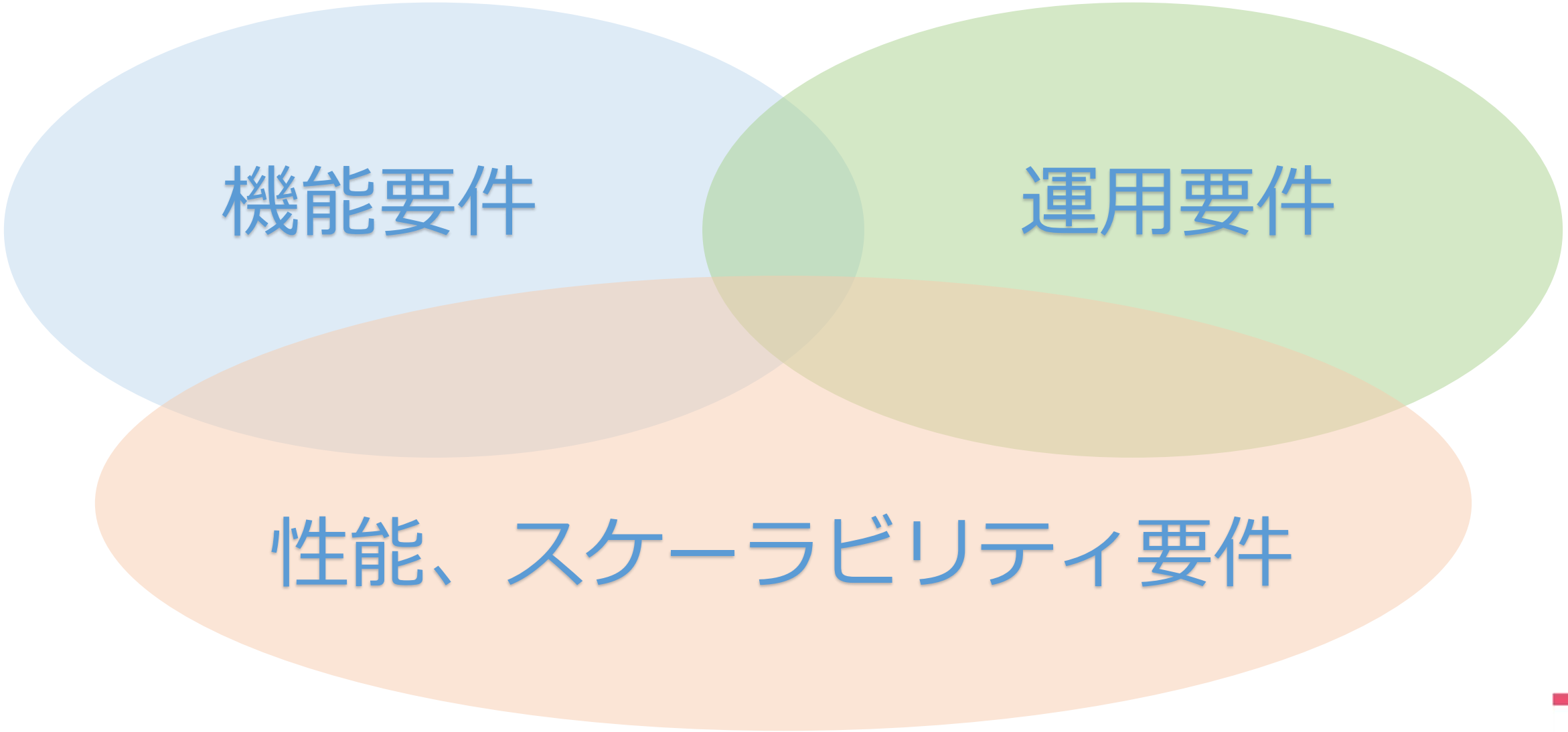
1. サービス仕様策定、システム全体設計
2. 各機器の想定機能要件、収容要件洗い出し
3. RFP
   - 要件を満たしそうで価格感がマッチする機種をいくつか選定
   - メーカさん、ベンダさんに相談したり
4. 機器検証
   - 検証項目洗い出し、検証構成検討
   - 検証機材をお借りしたり、ベンダさんのラボに訪問したり
5. 機器選定、導入判定、購入
   - バグ修正、追加インプリの条件をお願いすることも
6. システム構築、運用体制、総合テスト、サービスイン

このパートで
お話する部分

要件、検証項目のリストアップ

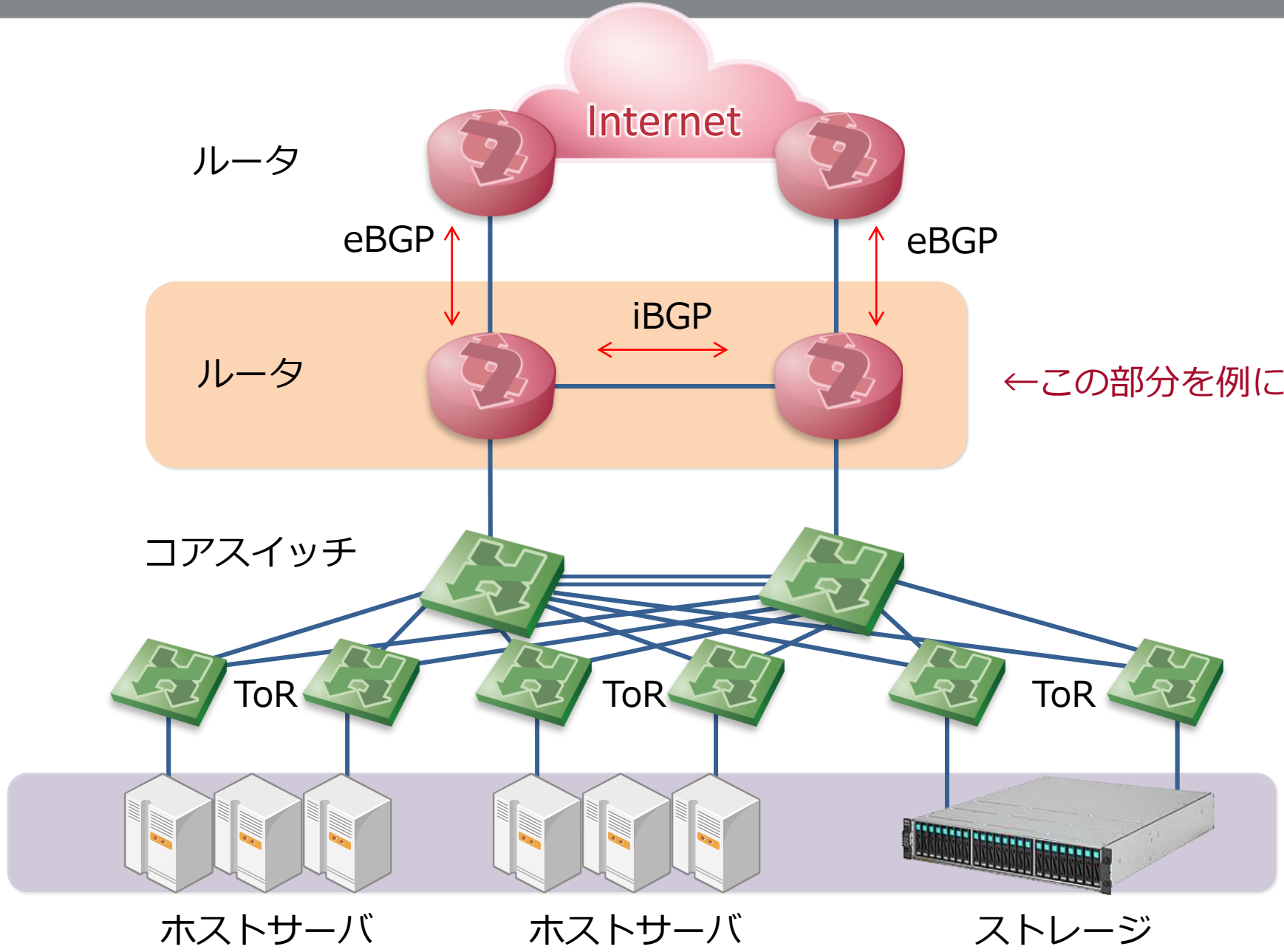
SAKURA internet

大きく3つの切り口で洗い出しを行う



ケーススタディ
～クラウド(IaaS)収容用途のルータ～

IaaSネットワーク構成例



上流とはプライベートASを用いたBGP接続

VRRPを用いた冗長化

3,500VLANくらい

1万VMくらい

要件抽出例(大項目)

SAKURA internet

- 機能要件
  - ルーティングプロトコル: OSPF, BGP IPv4/IPv6デュアル
  - 冗長化プロトコル: VRRP (もしくは同等のもの)
  - その他: IPv6 RA, VLAN単位での帯域制限, ACL
- 性能、スケーラビリティ要件
  - スループット: 10Gbps以上、そこそこMpps
  - SVI: 2,000以上
  - VRRP数: 2,000以上
  - ARP/NDPエントリ数: 12,800以上
- 運用要件
  - 管理、監視しやすいか？消費電力・サイズ・ファシリティ面
  - 大量のConfigをいれてもサクサク動くか？

要件抽出例(大項目)

- 機能要件
  - ルーティングプロトコル: OSPF, BGP IPv4/IPv6デュアル
  - 冗長化プロトコル: VRRP (もしくは同等のもの)
  - その他: IPv6 RA, VLAN単位での帯域制限, ACL
- 性能、スケーラビリティ要件

RFPはこれくらいの粒度で
検証項目はこれらをブレークダウンしてリストアップする

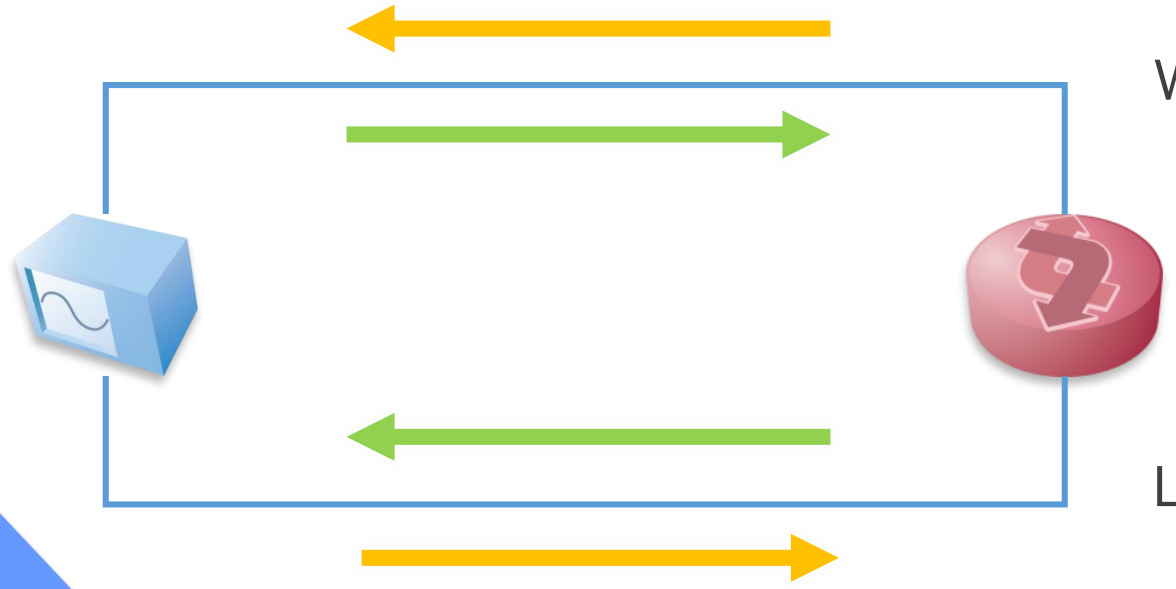
  - ARP/NDPエントリ数: 12,800以上
- 運用要件
  - 管理、監視しやすいか？消費電力・サイズ・ファシリティ面
  - 大量のConfigをいれてもサクサク動くか？

検証例

ルータ単体試験の例

機能要件 → 性能・スケール → 運用要件の順に
検証作業をすすめる

2,000SVI相当の
設定を投入



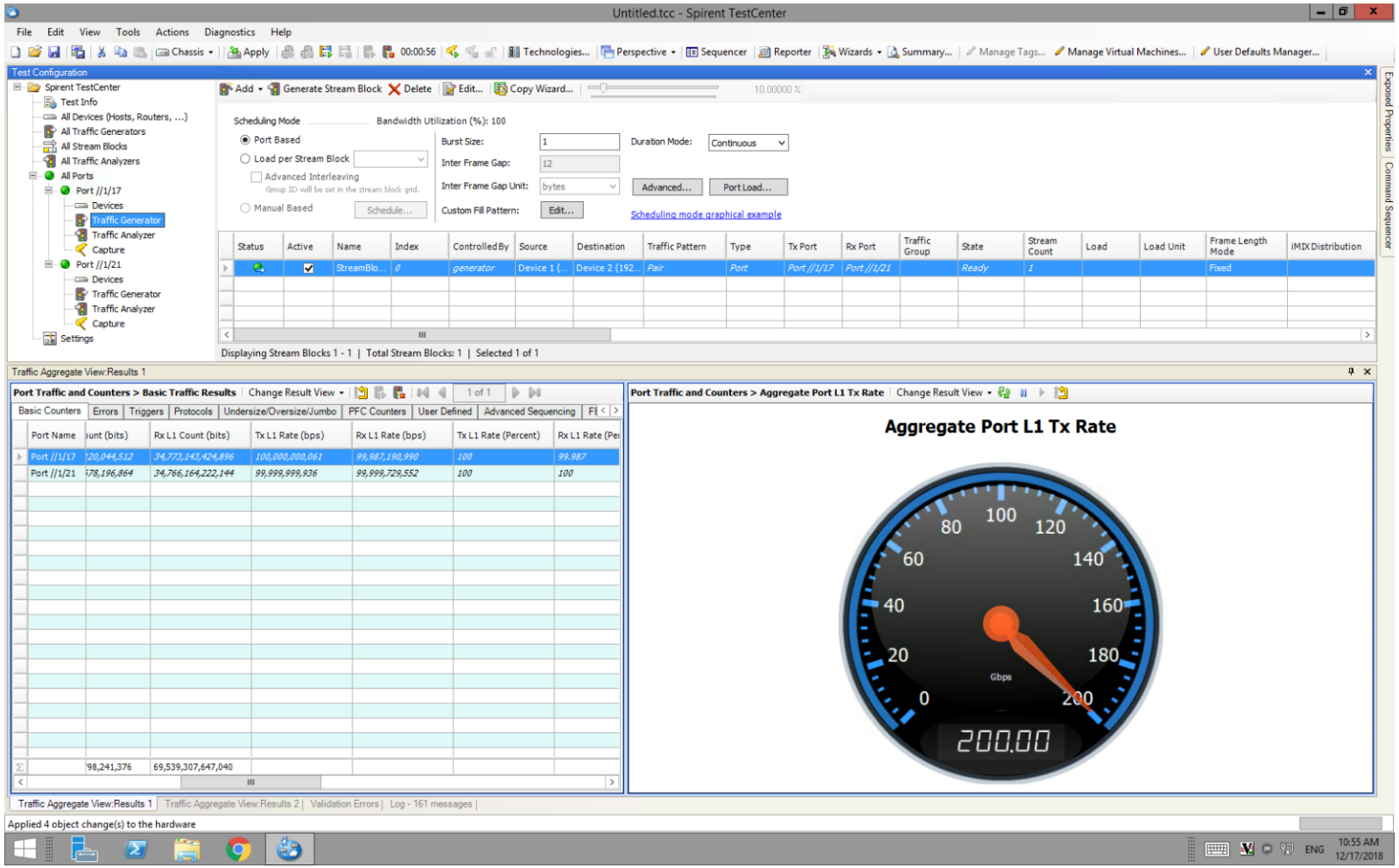
12,800VMを
エミュレーション

- ユーザ→インターネット向けを模擬
- インターネット→ユーザ向けを模擬

※ 上りと下りでは別世界

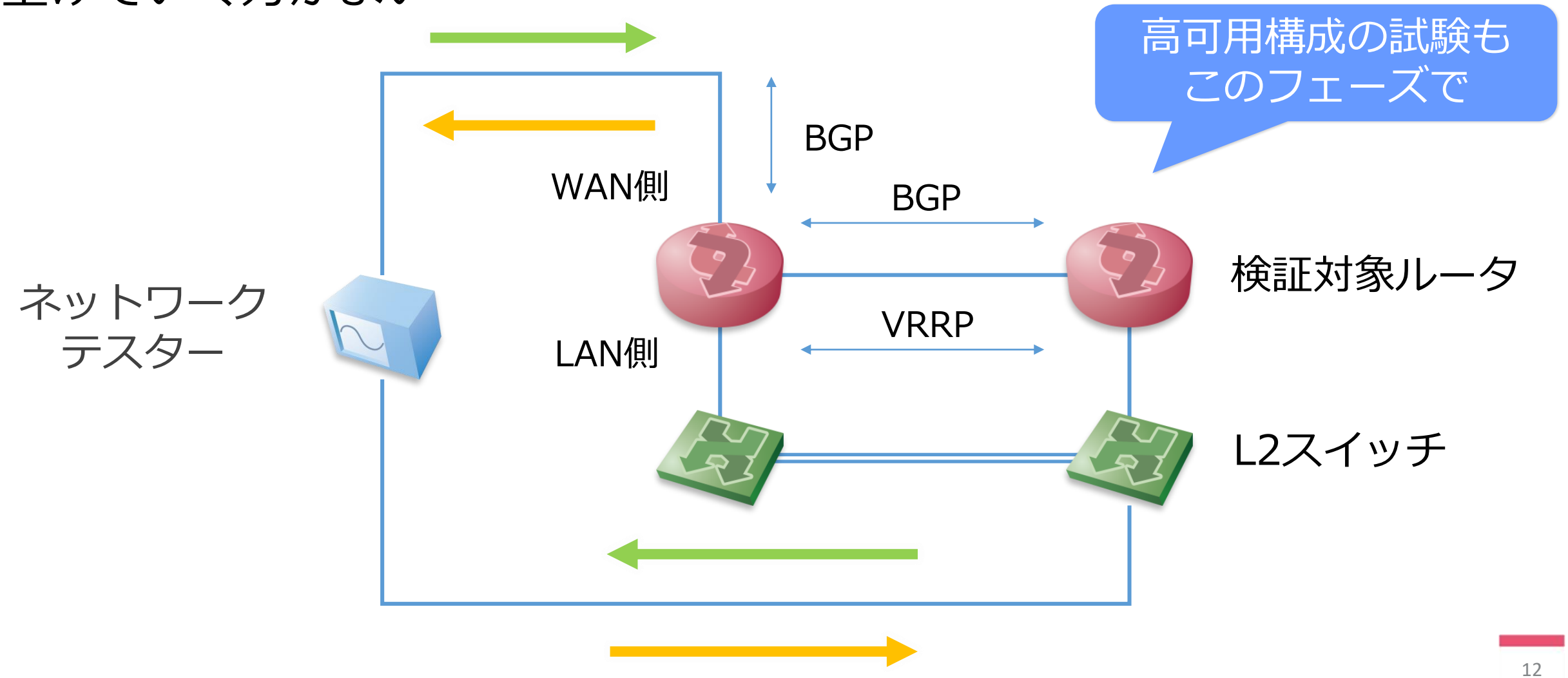
ネットワークテスターの活用

手元において置けると便利だが、結構お高いので・・
買えない場合はベンダさんのラボお借りしたりする



結合試験

いきなり組み上げずに、各機器を単体で試験しておき、徐々に組み上げていく方がよい

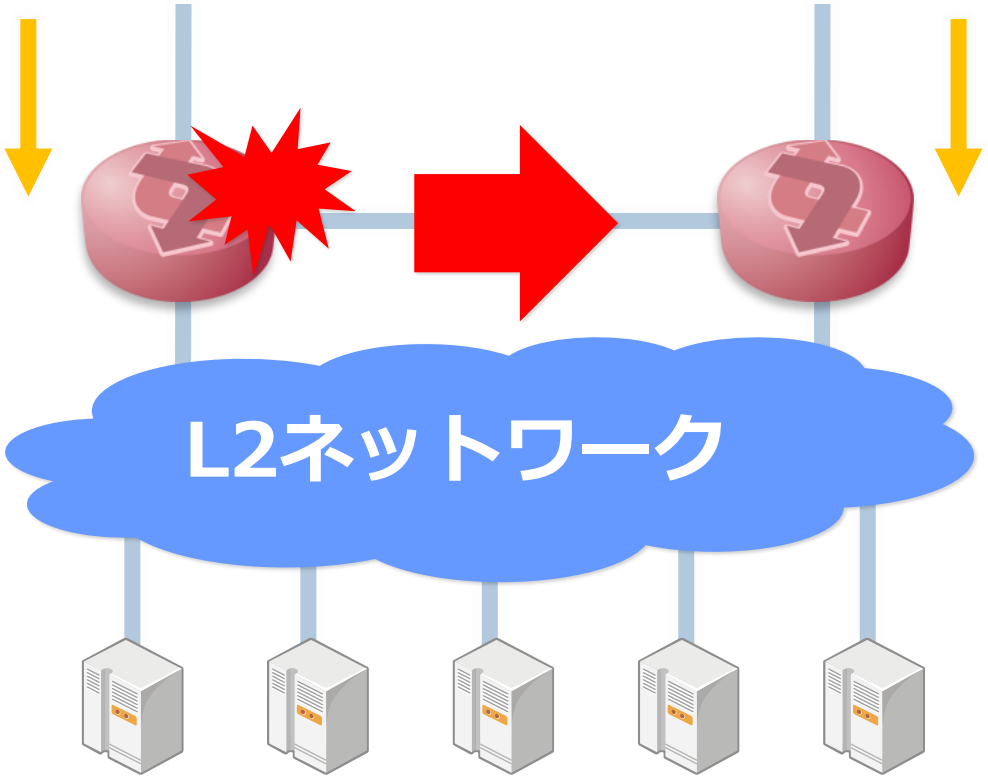


運用要件の確認

SAKURA internet

- 大量のConfigを入れたり抜いたり
  - クラウドだとお客様操作で自動でConfigされる
  - OSがクラッシュしないか確認
- ARP/NDPの学習スピード
  - 切り替え時間に影響

1万ノードの
ARP学習



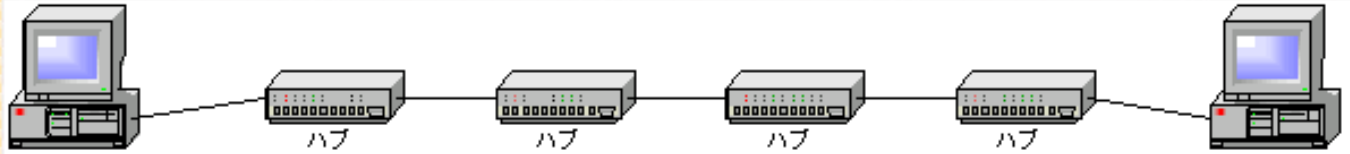
プロトコルレベルでは
すぐに切り替わっても
数分間通信が復旧しないことも

検証Tips

ショートパケットワイヤレート必要？

- 64バイトショートパケットワイヤレートの性能は本当に必要か？
  - 例えば10GbEだと14.8Mpps
- そもそもなぜイーサネットの最小フレーム長は64バイトなのか？

ご承知のとおり、**イーサネットの媒体アクセス制御方式は CSMA/CDです**。そのCSMA/CD で考えなければいけない重要なことは、**「最悪 の条件の下ですべての端末が衝突を検出」**できなければいけません。最悪な条件とは、一番はなれた2台の端末間での通信のときです。10BASE-Tイーサネットでは、4台のハブを経由する場合でしかも各 ケーブルが100mのときが最悪の条件です。



出典 http://www.n-study.com/network/minframe.htm

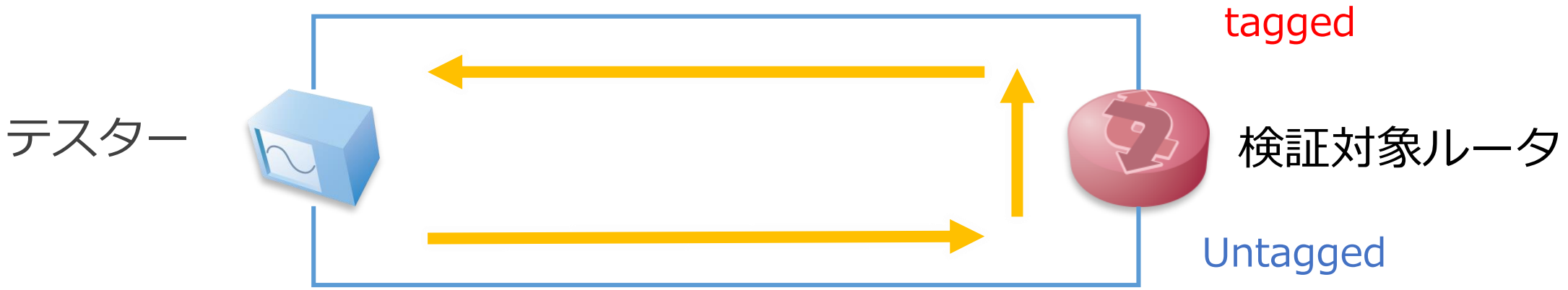
- たまたま100mのコリジョン検出のために64バイトに規定された
  - もし200m(128Bytes相当)だったら半分の性能(7.4Mpps)で済んでたかも

パケットバッファサイズの測定

SAKURA internet

バーストトラフィックをどれくらい吸収できるか？

タグVLANにすると1フレームあたり
4bytesずつバッファに貯まる



ロングパケットで99%、100%負荷時のレイテンシの差を求める
例: 99%:2.77us　100%:849.99us　差:847.22us

10Gbpsで847.22us分のパケットをバッファリング可能

10,000,000,000/8*0.00084722=1,059,025Bytes

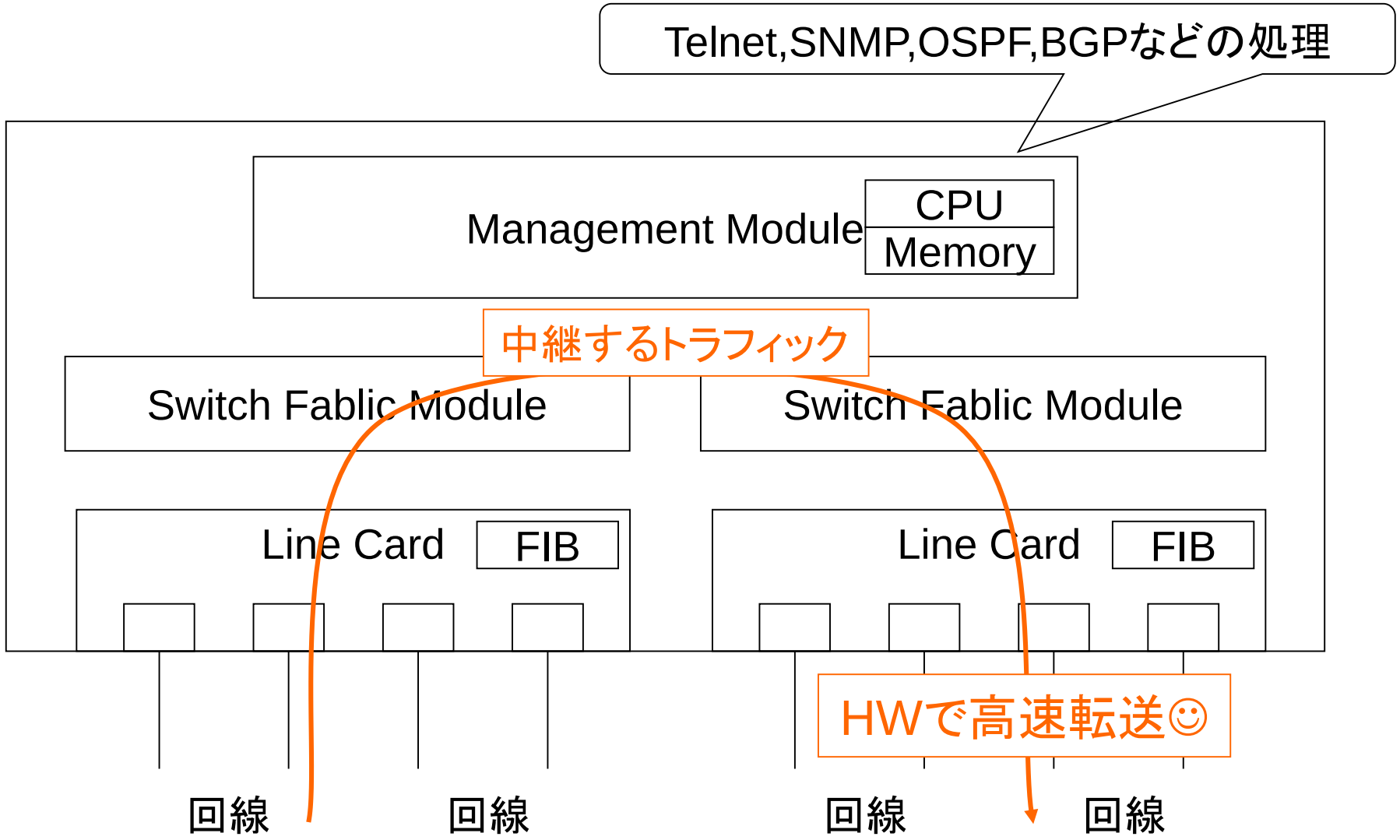
コントロールプレーンのロバスト性評価

SAKURA internet

- ルータ宛てのDoSアタック
  - UDPフラッド、ICMP echo request
  - SYNアタック(22番/23番/179番ポート)
- CPUエスカレーションされるパケット
  - TTL=1のパケット(TTL=0でICMP Time Exceeded生成)
  - ARP/NDP解決不能な宛先のパケット
  - トラフィックを流している状態でclear ip arp/clear ipv6 neighborsする
- 上記を10Gワイヤレートでぶつける

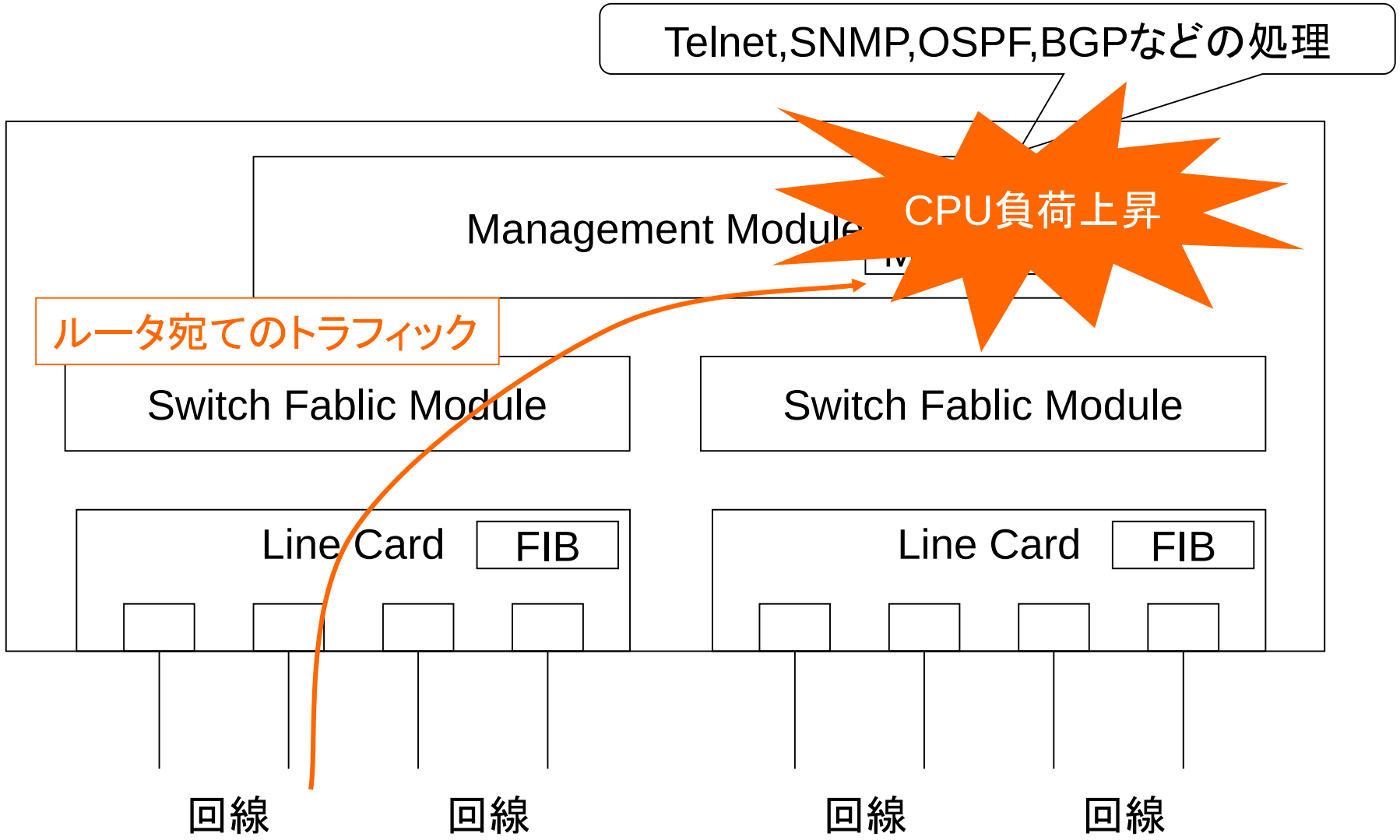
ルータの作り込まれ具合、叩かれ度合があからさまに

コントロールプレーンのロバスト性評価



http://irs.ietf.to/past/docs_20090521/ より

コントロールプレーンのロバスト性評価



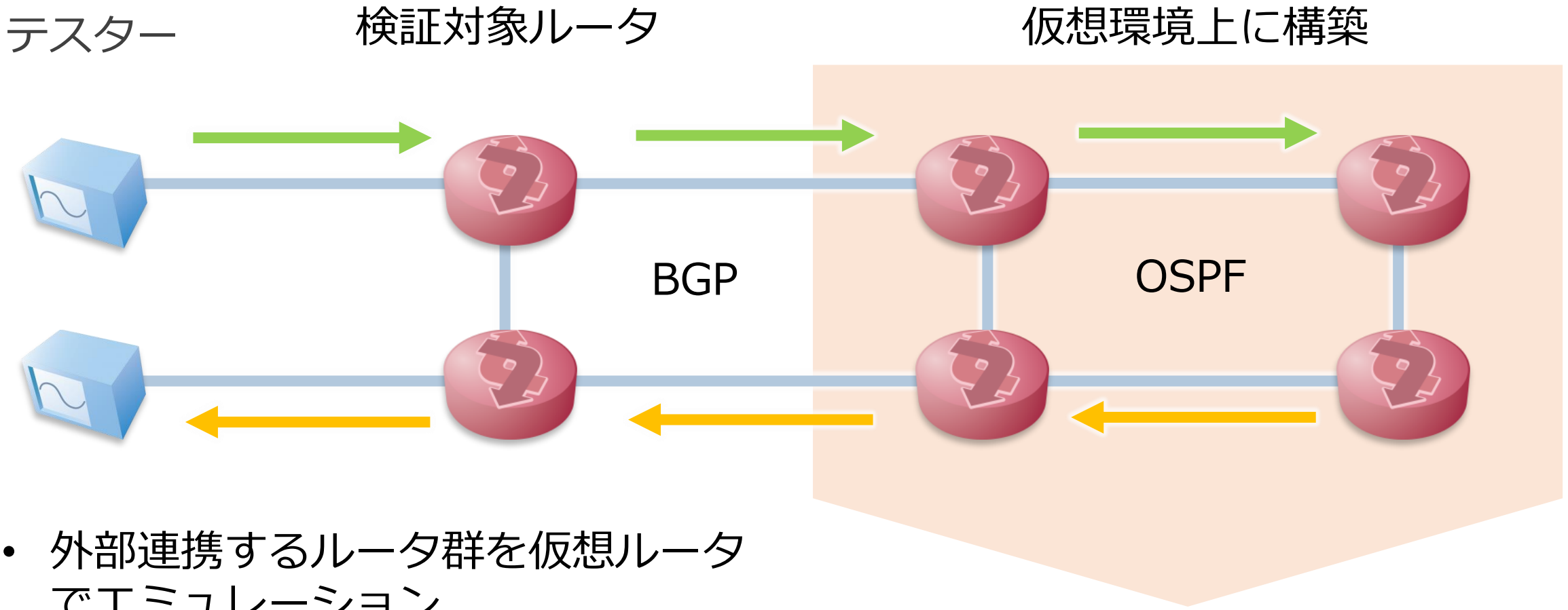
http://irs.ietf.to/past/docs_20090521/ より

発生した不安定事象の例

SAKURA internet

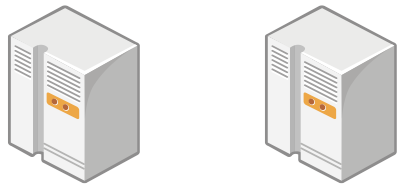
- VRRP Act/Stbフラップ
  - Standby側がVRRP Hello受信時に取りこぼし
- CLI操作が不能/重くなる
  - 23番/179番ポート宛てのDoSアタック
  - ARP/NDPエントリを大量に持った状態
- BGPピア断/OSPF Neighborダウン
  - clear ipv6 neighborsを実行
- LACPダウン

仮想ルータを組み合わせた検証環境



- 外部連携するルータ群を仮想ルータでエミュレーション
- 性能測定ではなく機能検証で活用
- 仮想環境なのでトポロジは柔軟にでき、検証環境の構築の手間も省ける

ホストサーバ



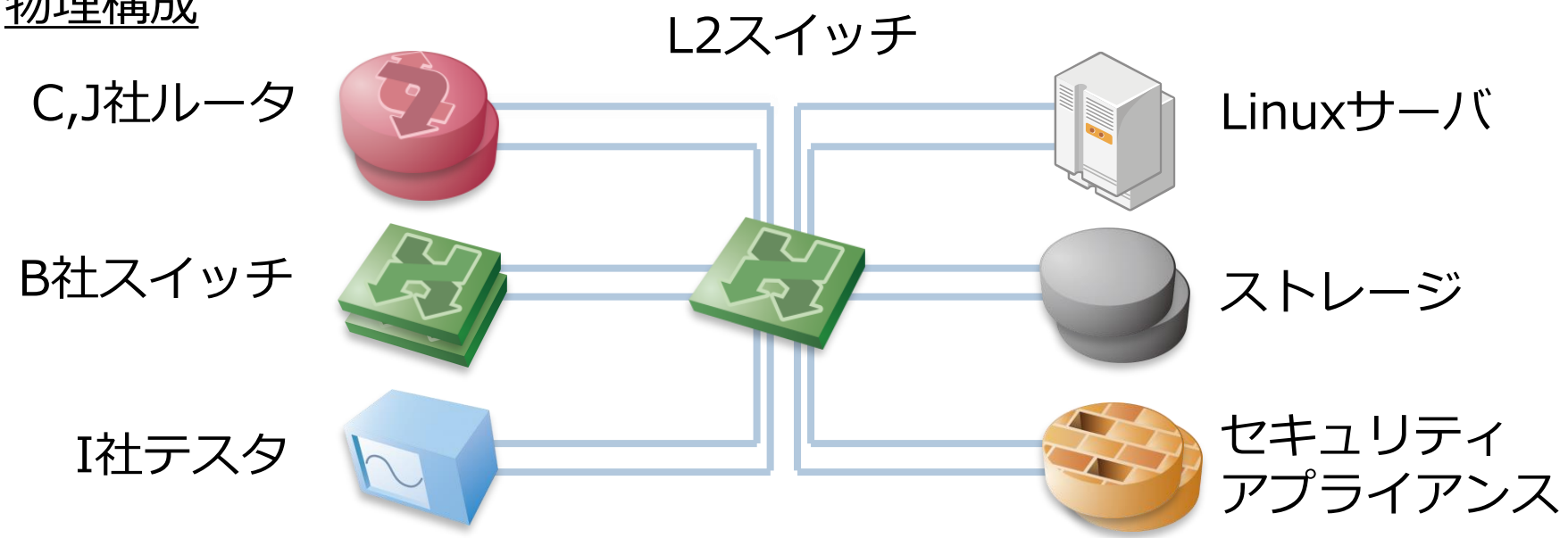
検証の手間を省くための工夫

SAKURA internet

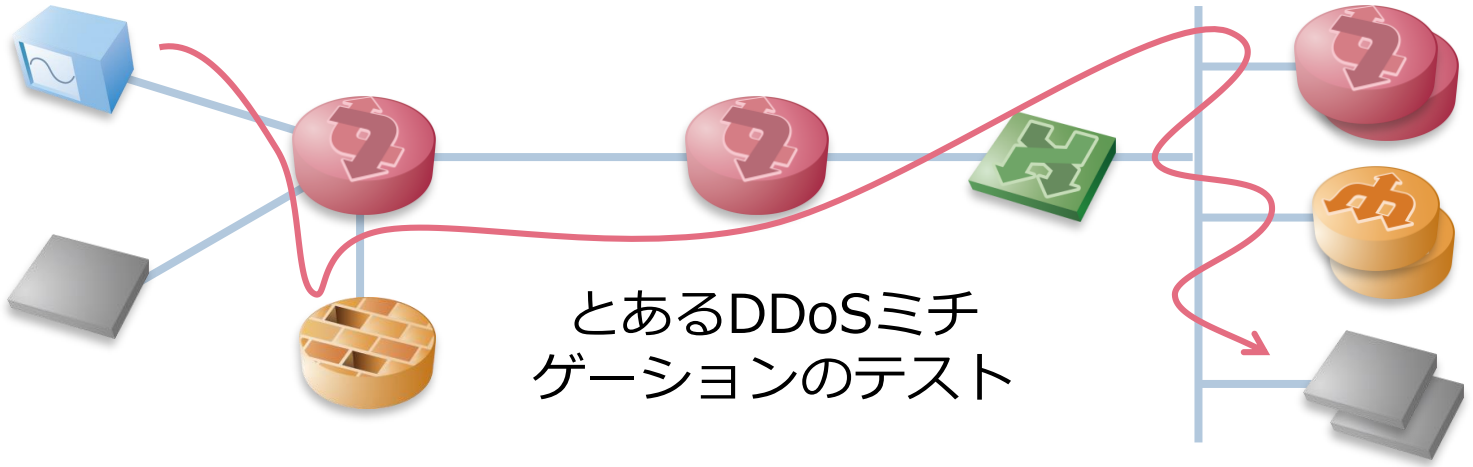
弊社のラボの構成例

写真（会場のみ）

物理構成



論理構成

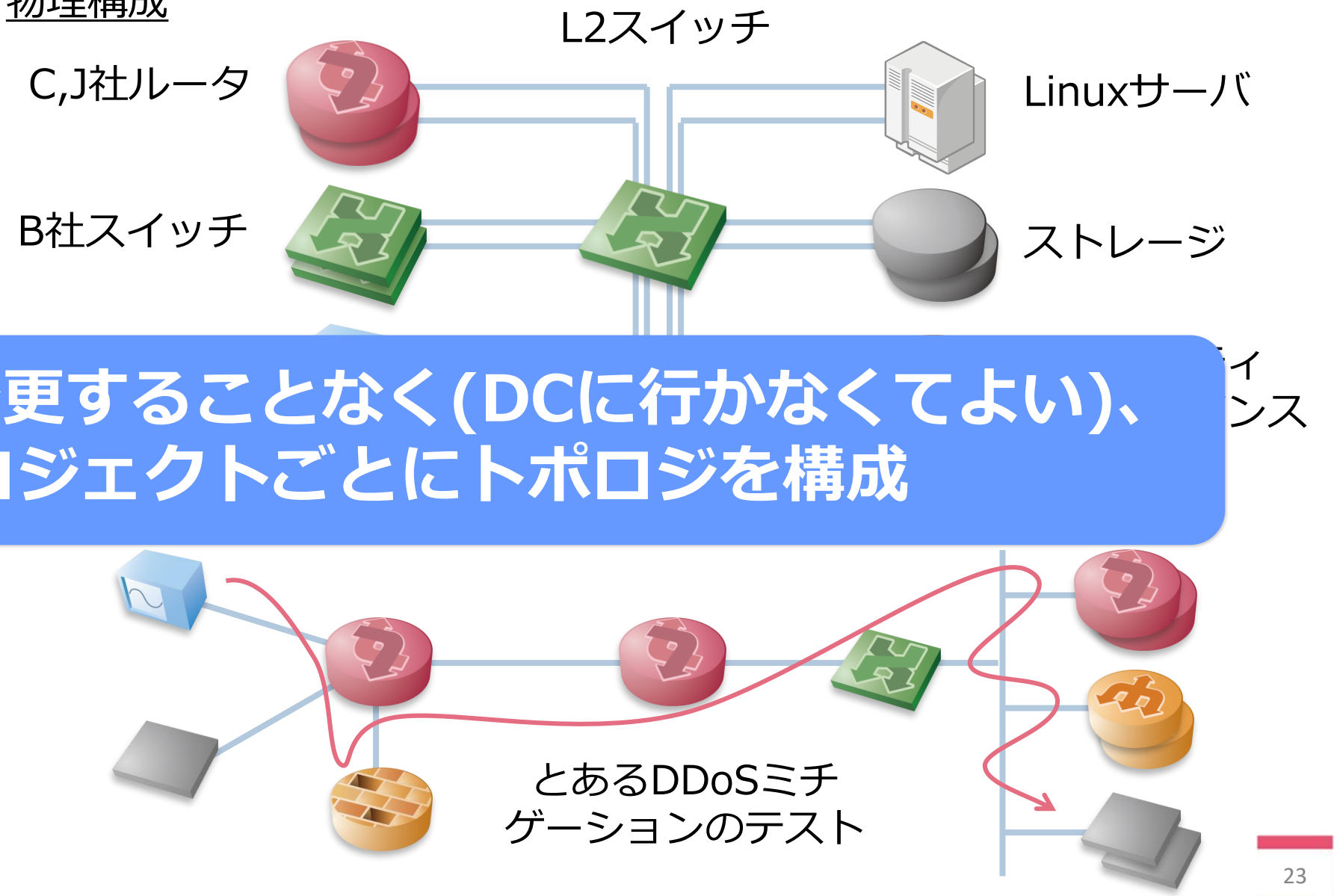


検証の手間を省くための工夫

SAKURA internet

弊社のラボの構成例

物理構成



物理構成を変更することなく(DCに行かなくてよい)、プロジェクトごとにトポロジを構成

検証結果のまとめ方

SAKURA internet

- 各検証項目ごとに、とりあえず試したことを随時メモしておく
  - 検証構成、config、印加試験トラフィックの内訳
  - 結果、ログ、グラフ、気づいた点
- 最初からきれいにまとめようとせずに
  - Slackとかに貼りまくっておくとあとでまとめるときに便利
- ある程度ネタが揃ったらドキュメントにまとめる
  - 検証レポートには、検証の目的、利用の想定、検証結果(マルバツ表)、サマリ(issueのリストと回避できるか？など)をまとめておくと良い
- マニュアル、機器の写真、ベンダさんへの問い合わせ内容もメモしておくと後で役に立つかも

検証作業の心得

SAKURA internet

- とりあえず溢れさせてみる
  - カタログスペックは信じない(言わずもがな)
  - 実運用で発生する多くの問題は、負荷や収容数に起因。
  - 収容上限を超えるConfigを突っ込んだり、ありえない負荷をかけてみたりする。
- 機器をダウンさせるあらゆる方法を考え、やってみる
  - 不具合が見つからないのは検証が足りてない証拠
  - Configに魂を込め、精根尽き果てるまで戦う
- メーカ/ベンダさんのレスポンスも要確認
  - 実運用に入って困ったときに助けてもらえそうか？
  - 何がおきてもなんとかなりそう、という感触を得られるか？

機器選定・導入判断の要素

SAKURA internet

- 検証結果
- 発見したissueは改修(回避 or 許容)できるものか？
- 運用性・保守性
- 価格
- スケジュール
- サービスの要求品質
- ベンダ・メーカさんの対応やカルチャー
  - サービスを一緒に作っていくパートナーとして
- いけると思ったかどうか？使いたいかどうか？
  - 完全に要求仕様を満たすものはない
  - 導入の最終決断は自身の覚悟をもって

ご清聴ありがとうございました