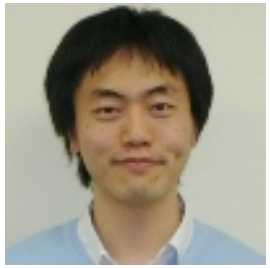
ＩＥＴＦ８７報告会
～ＣＧＮ関連技術～
behave/sunset4/opsawg

２０１３年０９月０５日
NTTコミュニケーションズ
Kaname, NISHIZUKA

自己紹介

西塚要



- 2006年 NTTコミュニケーションズ入社
  - 2006～ OCN(AS4713) エッジNW設計・開発
  - 2008～ 顧客ISP向け 提案・運用
  - 2012～ IAC(先端IPアーキテクチャセンタ)にて、CGN関連技術のIETF標準化活動等
- 社外活動
  - JANOG28 実行委員長
  - JANOG31 会場運営委員長など

【IETF86~87】RFC6888が発行されました。

※最初のID提出が2008/7

[RFC6888:BCP127]2013/4/30 BEHAVE WG

Common Requirements for Carrier-Grade NATs(CGNs)

通常のNATと異なるキャリア網NATとしての要求事項

- セッション上限機能
- Endpoint Independent Mapping
- Endpoint Independent Filtering
- Port再利用のルール
- セッション保持時間
- NAT Logの記録

各WGとCGN関連ドラフト

WG	タイトル	バージョン	IETF87発表
BEHAVE	Common Requirements for Carrier-Grade NATs (CGNs)	RFC6888	
BEHAVE	draft-ietf-behave-ipfix-nat-logging	01(WG)	○
BEHAVE	draft-ietf-behave-syslog-nat-logging	02(WG)	○
BEHAVE	draft-ietf-behave-requirements-update	00(WG)	○
BEHAVE	draft-ietf-behave-nat-mib	07(WG)	○
BEHAVE	draft-nishizuka-cgn-deployment-considerations	00	○
BEHAVE	draft-tsou-behave-natx4-log-reduction	04	
BEHAVE	draft-donley-behave-deterministic-cgn	06	
SUNSET4	draft-chen-sunset4-cgn-port-allocation	02	○
OPSAWG	draft-ietf-opsawg-lsn-deployment	03	○

CGNの実際の運用・デプロイに関連する draft が behave 以外のWGにも分散してしまっている現状。

BEHAVE WG item: NAT logging

draft-ietf-behave-ipfix-nat-logging
draft-ietf-behave-syslog-nat-logging

Both IPFIX and SYSLOG drafts were adopted as WG documents in IETF86
両方とも同じ内容・表記であるべきで、syslogとipfixの違いを無くすことが目的

議論の余地
- destination logging をするかどうか
- pre-NAT addresses をどのように表現するか(NATx4)
- port block allocation をどのように表現するか

CGN ～実NWへの展開へ向けたドラフト提出

Carrier-Grade-NAT (CGN) Deployment Considerations

draft-nishizuka-cgn-deployment-considerations-00

CGNを展開する上で考慮すべきこと

中～大規模ISPと同等のNWをStarBED(@北陸)にて構築して行った実機検証に基づいた結果を記載

[POINT]

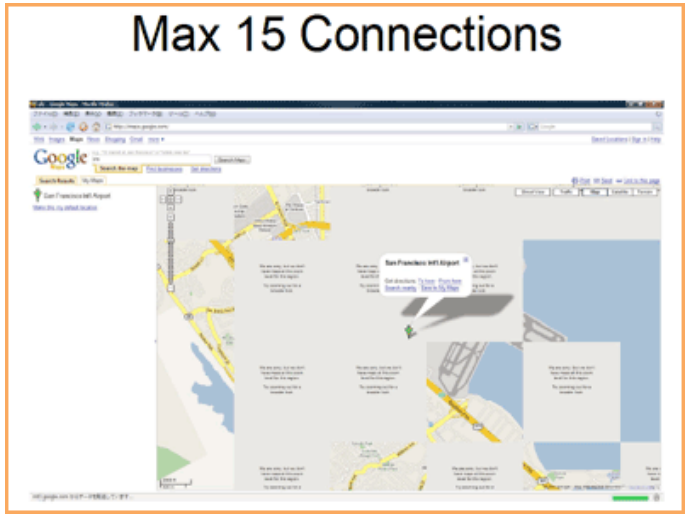
1. CGNの収容上限(キャパシティ)の見積もり

2. ポート割当手法によるIPv4アドレス節約効率とNATログ見積もり

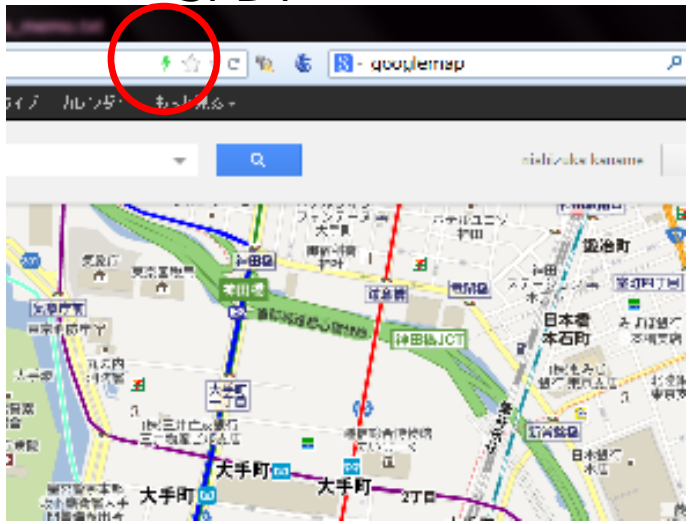
3. CGNの最適配置

1．CGNの収容上限(キャパシティ)の見積もり

SPDY/WebSocketなどの次世代Web技術によってセッションが重畳されるため、CGN(NATx4)のようにStateを持つNW機器への影響が変わってきています。



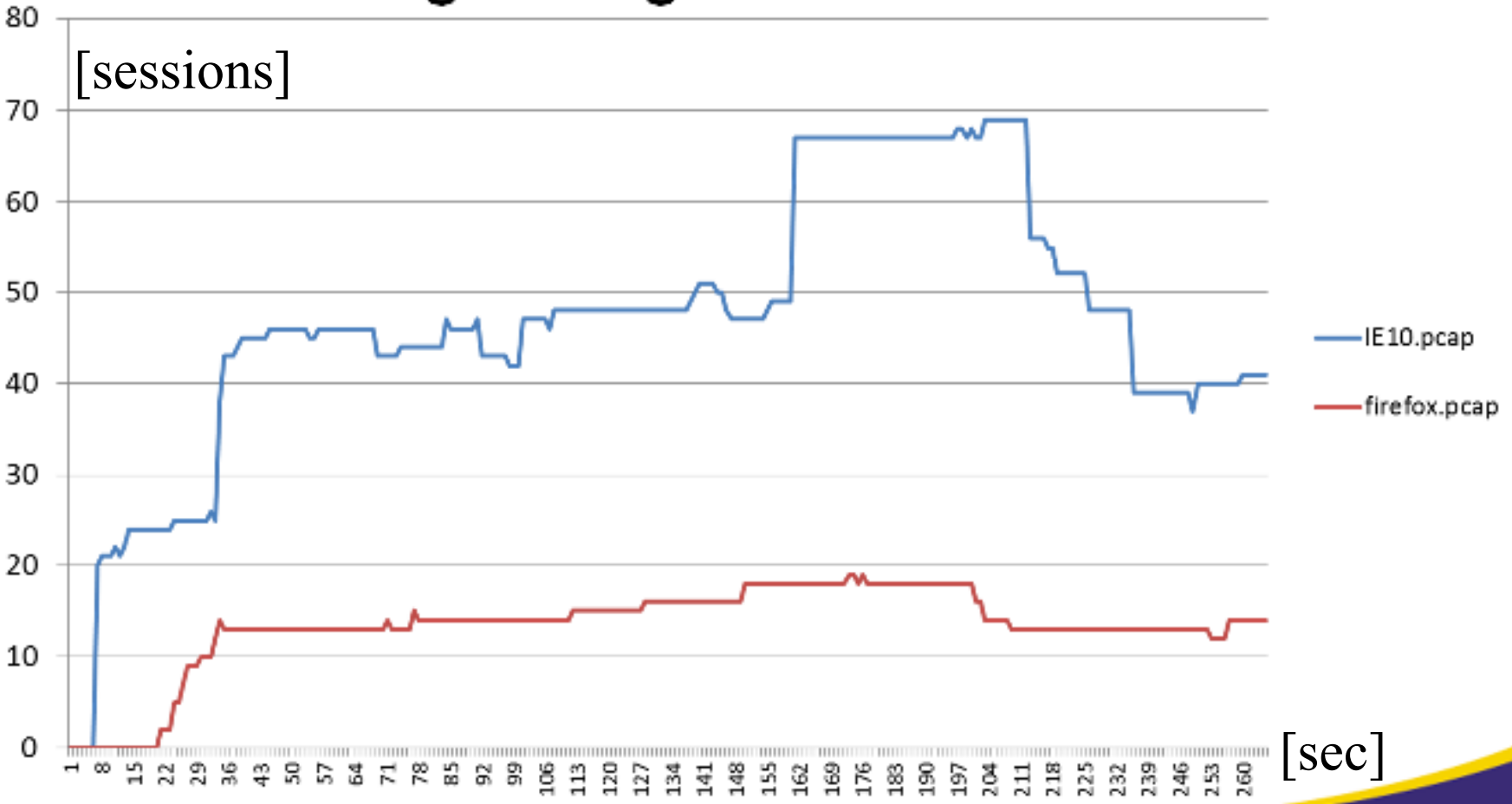
SPDY



Google Mapは複数ポートを利用するアプリケーションの象徴として使われてきましたが、現在はほとんどポートを消費しません(～10)

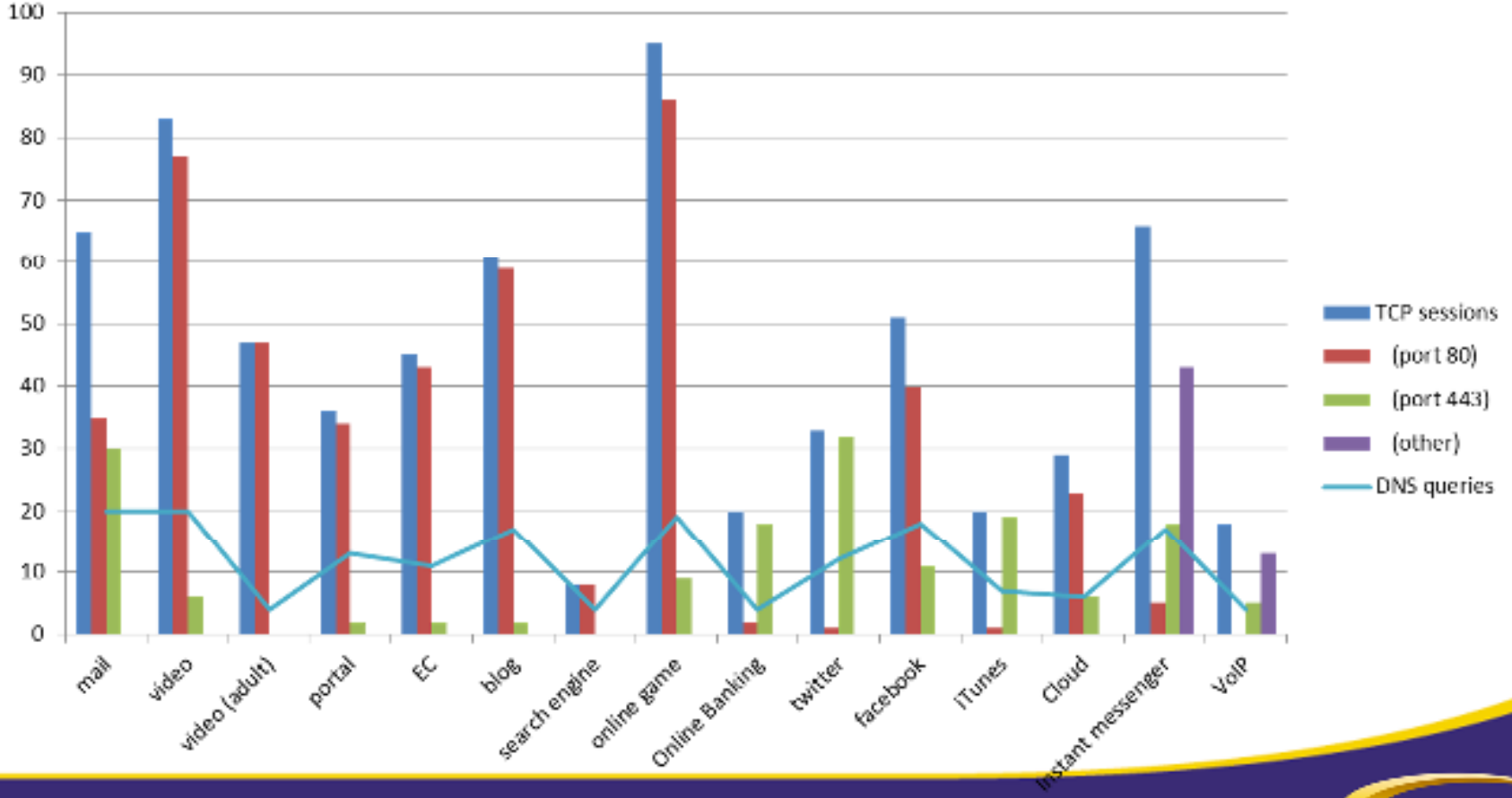
The impact of SPDY(or HTTP/2.0)

- **IE10 (not SPDY) vs FireFox21.0 (SPDY)**
  - **Browsing "Google MAP"**



The number of sessions

- **The average number of sessions has become less than in 2009.**



1．CGNの収容上限(キャパシティ)の見積もり

CGNの性能は以下の3点で測る。

指標	1ユーザあたり
トラフィック転送性能	50～100kbps
同時セッション数	100
セッション到着率(CPS)	0.05～0.2

CGN製品に、現実に近いトラフィックを印加したところ、16万ユーザを収容できる性能が認められた。また、HA構成による切替も可能であった。

CGNは当初(2009年当時)と比較して、
- 高集約可能な機器となった
- SPOF(Single Point of Failure)ではない
- 高価なソリューションではない

2．ポート割当手法によるIPv4アドレス節約効率とNATログ見積もり

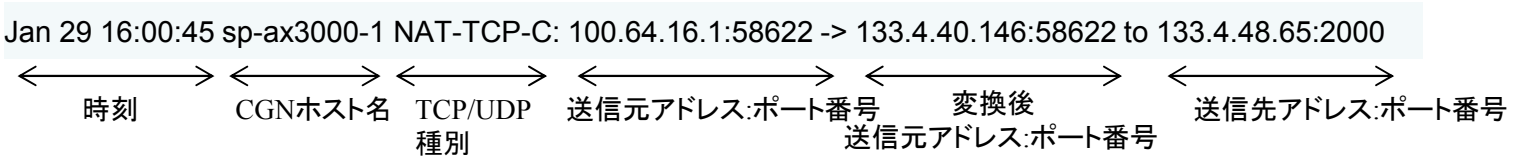
どのポートの割当手法を用いるか。

- Dynamic Assignment
  - 1ユーザあたりの平均ポート利用数(＝100)がkey factor
  - 約600ユーザで1グローバルアドレス
  - **# of pool address (P) =**
    # of Subscriber (S) * a * N / (65536 - R)
    a=25%, N=400.
- Static Assignment
  - 1ユーザの最大ポート利用数(＝1000)がkey factor
  - 約60ユーザで1グローバルアドレス
  - **# of pool address (P) =**
    # of Subscriber (S) * M / (65536 - R)
    M=1000

Dynamic Assignmentは、Static Assignmentと比較して、約10倍のアドレスシェア効率をもつ。

2．ポート割当手法によるIPv4アドレス節約効率とNATログ見積もり

NATログの見積もり



ASCII format: 120byte/record
Binary format: 26byte/record (20～25%)

information	byte
timestamp	8
CGN hostname(ID)	2～16
Transport protocol	1
Add/Delete flag	1
untranslated source address/port	6
translated source address/port	6

for 1,000,000 users, the size of log is piled up to 6.4TB per day.

NATログは非現実的なレベルではない
Port Block Assignなど、ログの量を軽減する手法も発達

３．CGNの最適配置

◆既存のNWにどのようにCGNを導入するか

CGNの性能向上によってコアに近い区間に配置することが可能になった。
しかし、NATはルーティングドメインを分割するため、従来のルーティング設計に影響を与える。
⇒インパクトを極小にするためには
- 既存のルーティングドメインと分割
- CGNで動的ルーティングが可能
- CGNのHA機能とルーティングプロトコルが連動できることが重要

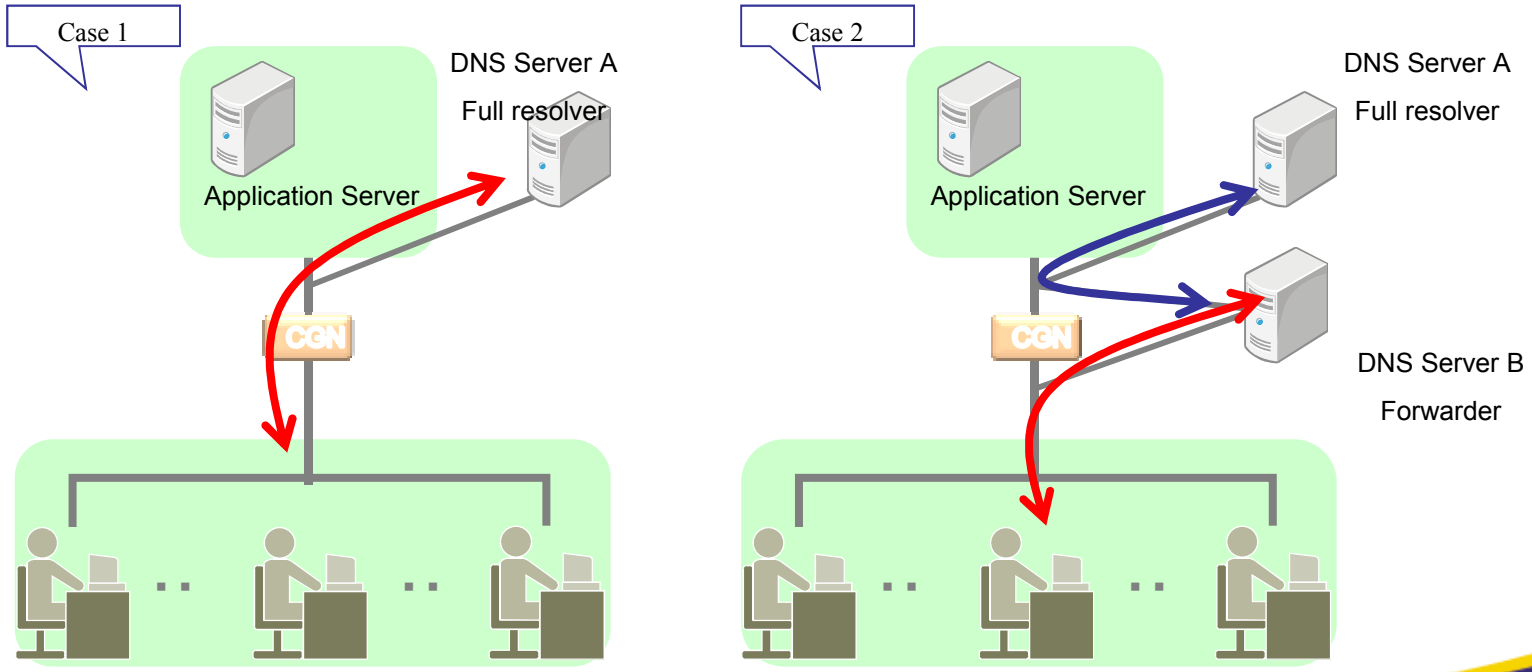
後述のOPSAWGのdraftと同様の問題意識

Where to place the CGN

<DNS considerations>

Should we bypass DNS queries to avoid intensive port consumption?

⇒No, NAT table of DNS timeouts in 3 seconds, so the consumption of NAT tables is not so much. As the result, it did not affect the performance of CGN.



draft feedback

draft 投稿後[behave/opsawg WG ML]
- DNSの配置についての議論が興味深い(Dan Wing)
- ウェブ/アプリケーションのIPv6対応調査やブラウザごとの挙動の違いを記述することで、より強力なレファレンスとなりうる(Shishio Tsuchiya)
- ポートの動的割り当て/静的割り当ての見積もりは良い指標となる(多数)

IETF87会場質疑
- オペレータにとって貴重なレポートである(Dave Thaler)
- NAT64での我々の調査結果と近い内容である(China Mobile)

IETF87後
- このドキュメントが非常に参考になった。GCM (Google Cloud Messaging)でタイムアウトの問題に直面している(Yutaka Ishizaki)

CGN Deployment with BGP/MPLS IP VPNs

draft-ietf-opsawg-lsn-deployment

V. Kuarsingh(Rogers Communications)

NAT対象 / NAT対象外のトラフィックを分離するアーキテクチャの提案

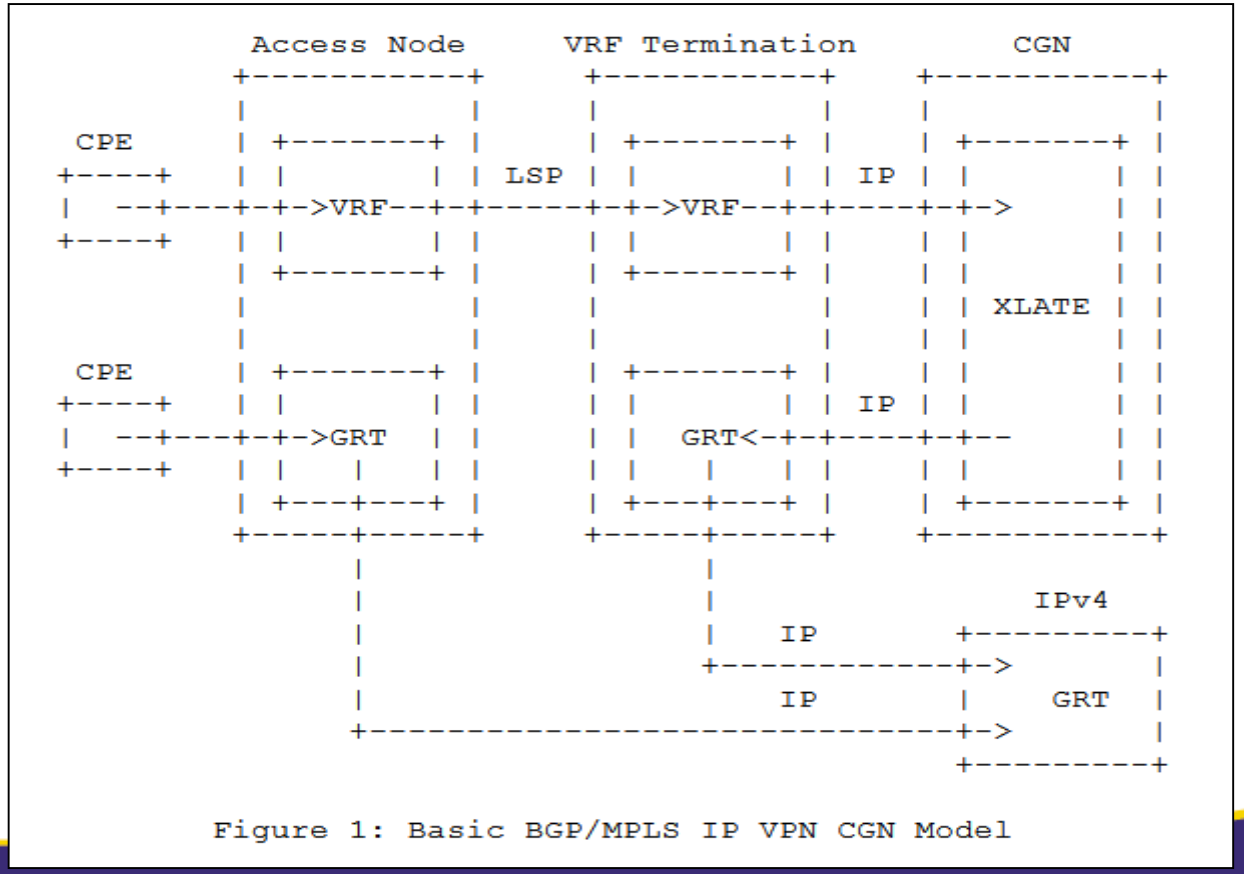


Figure 1: Basic BGP/MPLS IP VPN CGN Model

[OPSAWG]

[POINT]

1. NAT444環境の肯定
2. Routing Plane Separation
   - Policy Based Routingでも可能であるが、VRFによる分離はオペレータにとってより優しい選択である。
3. Flexible Deployment Options
   - CGNの集約効率が高く、またCGNへの依存度が次第に減っていっても分割損が発生しにくい

[質疑について]

何人かのオペレータからオフラインで多くの肯定的なコメントがあり、実際に環境を構築して、動作することを確かめたとのこと。
しかし、IDを読んだ人間は少なく、活発な議論は起こらなかった。

Analysis of NAT64 Port Allocation Method

draft-chen-sunset4-cgn-port-allocation

G. Chen(China mobile)

[POINT]

1．NAT64推進の理由づけ

- IPv6を同時提供することによって、対応済のアプリやサイトはIPv6通信となるため、CGNのセッションを消費しない。

2．NAT64環境のポート消費調査

[質疑について]

特になし。

結論

◆CGN(NAT444)が実際に使われ始めている
⇒CGN設計指針となるDraftが必要だった。

◆CGN関連のドラフトがBEHAVE以外のWGに分散しており、横断的に追うのが難しい。pcp-WGの議論もウォッチする必要があるが、pcpは実際のデプロイには至っていない。

◆「NATx4について複数のドラフトが提出されており、マージした方がいいものもあるだろう。もしより完璧で統一的なドラフトがあれば、それをWG itemにするのがよいと思われる。」(Sunset4 Chair: Wesley George)