

経路制御

(関連WGの動向)

古河電気工業株式会社
ファイテルネットワーク研究所
武藤 大
dai@inf.furukawa.co.jp

2002年4月24日

内容

- ⌘ Routing Area の動向
- ⌘ 経路制御の問題点
 - ⊡ 数 (アドレス枯渇, フィールド拡張) の問題
 - ⊡ セキュリティの問題
- ⌘ 経路制御関連の WG の動向
 - ⊡ Rpsec BOF
 - ⊡ Forces WG
 - ⊡ その他 WG
- ⌘ まとめ

Routing Area

⌘ Routing area

- ☑ 経路制御全般を扱う Area

⌘ Area Director

- ☑ 交代が激しい, 最近 1 年だけでも,
David Oran, Rob Coltun, Abha Ahuja
- ☑ 現在は以下の2人
Randy Bush <randy@psg.com>
Bill Fenner <fenner@research.att.com>

⌘ WG 数

- ☑ WG の整理が進んでいる, 現在の WG 数は 13

Routing area の WG

- ⌘ Bgmp (Border Gateway Multicast Protocol)
- ⌘ Forces (Forwarding and Control Element Separation)
- ⌘ Idmr (Inter-Domain Multicast Routing)
- ⌘ Idr (Inter-Domain Routing)
- ⌘ Isis (IS-IS for IP Internets)
- ⌘ Manet (Mobile Ad-hoc Networks)
- ⌘ Msdp (Multicast Source Discovery Protocol)
- ⌘ Ospf (Open Shortest Path First IGP)
- ⌘ Pim (Protocol Independent Multicast)
- ⌘ Rip (Routing Information Protocol)
- ⌘ Ssm (Source-Specific Multicast)
- ⌘ Udldr (UniDirectional Link Routing)
- ⌘ Vrrp (Virtual Router Redundancy Protocol)

IGP, EGP, Multicast, others

Routing Area の WG の開催動向

⌘ 第 53 回 IETF

☑ FORCES, IDR, MANET, MSDP, PIM, UDLR, VRRP, RPSEC(Bof)

⌘ 第 52 回 IETF

☑ FORCES, IDR, PIM, UDLR, VRRP

⌘ 第 51 回 IETF

☑ FORCES, MOBILEIP, ISIS, IDR, MANET, MSDP, PIM, UDLR

⌘ RIP, OSPF はしばらく開催されていない

Routing Area の動向

- ⌘ 流行の MPLS, Traffic engineering 等は Sub-IP area に移動, 整理される
 - ⌘ 地味, 話題が少ない
 - ☑ 現状の問題点を解決するための地道な作業が進んでいる
 - ☑ が, それほど進んでいないようにも見える
- “Overview of the 1998 IAB Routing Workshop”
(RFC2902)

経路制御の問題点

⌘ 数の問題

- ⊡ IPv4 アドレス空間
- ⊡ AS 番号数
- ⊡ その他フィールド拡張

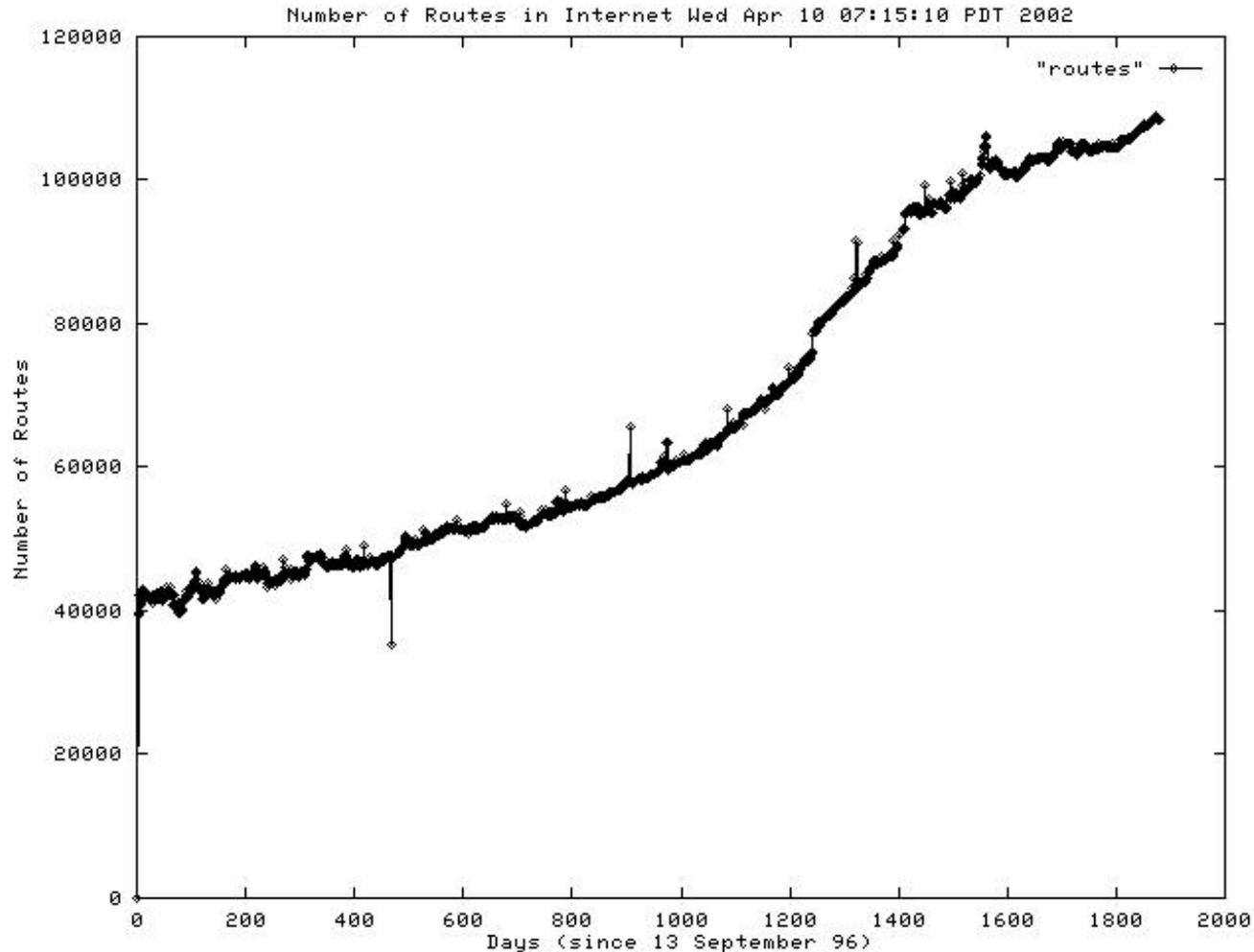
⌘ 用途の変化(インターネットが生活基盤になり得るのか?)

- ⊡ セキュリティ
- ⊡ モバイル(ユビキタス)
- ⊡ 安定度

IPv4 アドレス空間

- ⌘ インターネットの普及によるIPv4 アドレス空間の枯渇
 - ☑ IPv4 のアドレス数は世界の人口にも満たないおよそ43 億
 - ☑ ユーザ数の増加, 利用形態(モバイル)の変化により, 必要とされるアドレス空間は増加している
- ⌘ 経路表の増加によるルータの負荷の増加
 - ☑ 現在, 10 万を超える経路が交換されている
 - ☑ 最適経路の検索
 - ☑ フラップ等による間接的負荷

IPv4 経路数の遷移



<http://www.employees.org/~tbates/cidr-report.html>

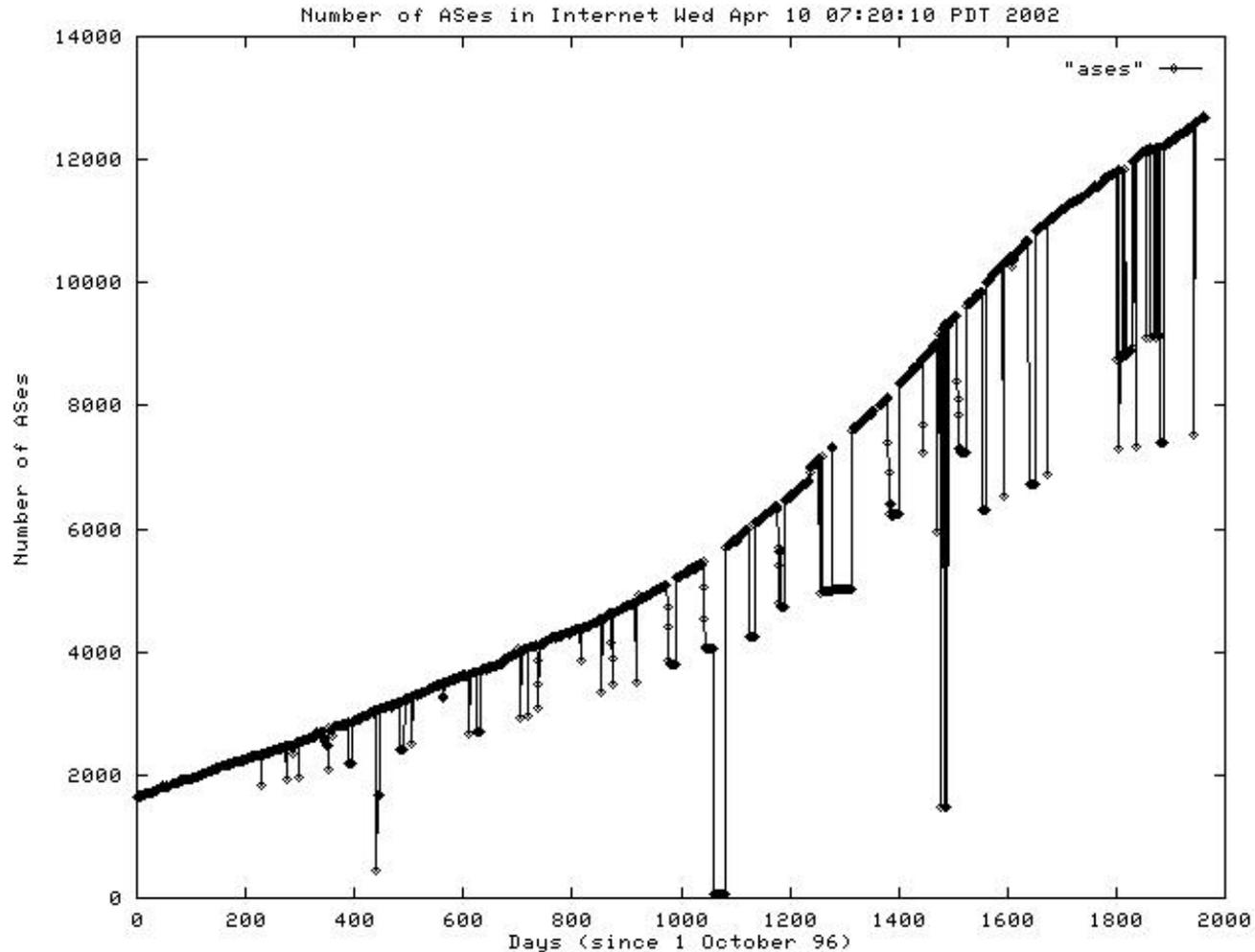
Classless Inter-Domain Routing (CIDR)

- ⌘ “Classless Inter-Domain Routing (CIDR):
an Address Assignment and Aggregation
strategy” (RFC1519)
- ⌘ 経路の効率の良い集約
- ⌘ Cidr 以外の効果として, NAT の効果も無視できない
“Traditional IP Network Address Translator (Traditional
NAT)” (RFC3022)

AS (Autonomous System)

- ⌘ AS とは, 共通のポリシーや同じ管理下で運用されているルータやネットワークの集合を意味する (例: ISP, 会社)
- ⌘ AS という仮想単位毎に経路を Cidr 等を利用して, 出来るだけ集約して交換
- ⌘ AS 割り当ての条件はマルチホーミング
“Guidelines for creation, selection, and registration of an Autonomous System (AS)” (RFC1930)

AS 番号数の遷移



<http://www.employees.org/~tbates/cidr-report.html>

AS 番号の枯渇

- ⌘ グラフは利用されている AS の数。割り当て済みの数は、さらに多い。
- ⌘ AS番号の拡張
 - ⊡ 現在, 2バイト長
 - ⊗ 1 ~ 65535 (ただし, 64512 ~ 65535 は, プライベート AS 番号)
 - ⊡ 4バイトへの拡張が, IDR WG で議論されている
 - ⊗ “BGP support for four-octet AS number space” (draft-ietf-idr-as4bytes-04.txt)
 - ⊗ 影響範囲が大きく, 移行は進んでいない
 - ⊗ WG のステータスは, 実装が進むのを待っている状態

IPv6

- ⌘ IPv4 アドレス空間の枯渇に対する最終的解答
- ⌘ インターネットへの要求の変化ユビキタス, モバイル, マルチメディア (Everything over IP)
- ⌘ ネットワークトポロジーに基づいたアドレス配布
経路集約: 現在, 数百のオーダ

IPv6 の問題点 1 (/48 の割り当て)

- ⌘ 小さなサイトも “ /48” (IPv4 での class B 相当) のネットワークアドレスが配布される
- ⌘ どこまで設定が簡略化できるか？
 - ☑ サイト内の設定, ネットワークアドレス(プレフィックス)配布方法
 - IPv6 WG : Router renumbering, PPP 拡張, DHCPv6, DHCPv6 subset, 標準 RA, RA 拡張, ICMPv6 拡張
 - ☑ Default router 等の設定

IPv6 の問題点 2 (マルチホーム)

- ⌘ IPv6: アドレス割り当ても、最初から再出発
 - ☑ ネットワークトポロジーに基づいたアドレス配布による経路情報の効率的集約
- ⌘ マルチホームにより、現在の IPv4 のように経路表の爆発的増加もあり得る
 - ☑ ルータ、ホストの負荷が小さい、シンプルな機構を考える必要がある
 - ☑ Site Multihoming in IPv6 (multi6) WG
 - 第 51 回 IETF 報告会 「マルチホームと IPv6」(太田)
 - (<http://www.net.intap.or.jp/INTAP/ietf/51st/4.files/frame.htm>)

Extended Community 属性の拡張

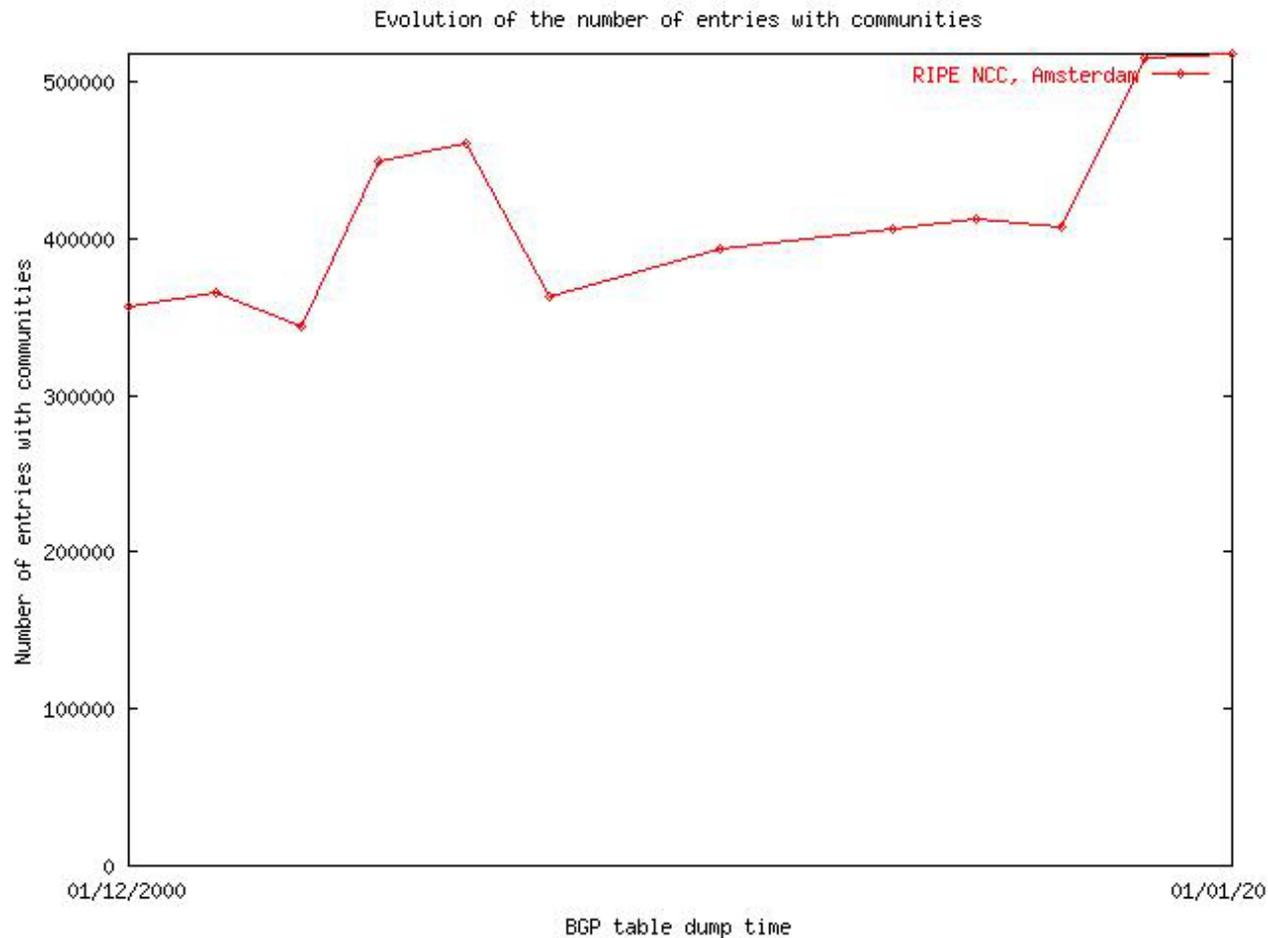
⌘ Community 属性

- ☒ "BGP Communities Attribute" (RFC1997)
- ☒ タイプコード値は "8"

⌘ Extended Community 属性

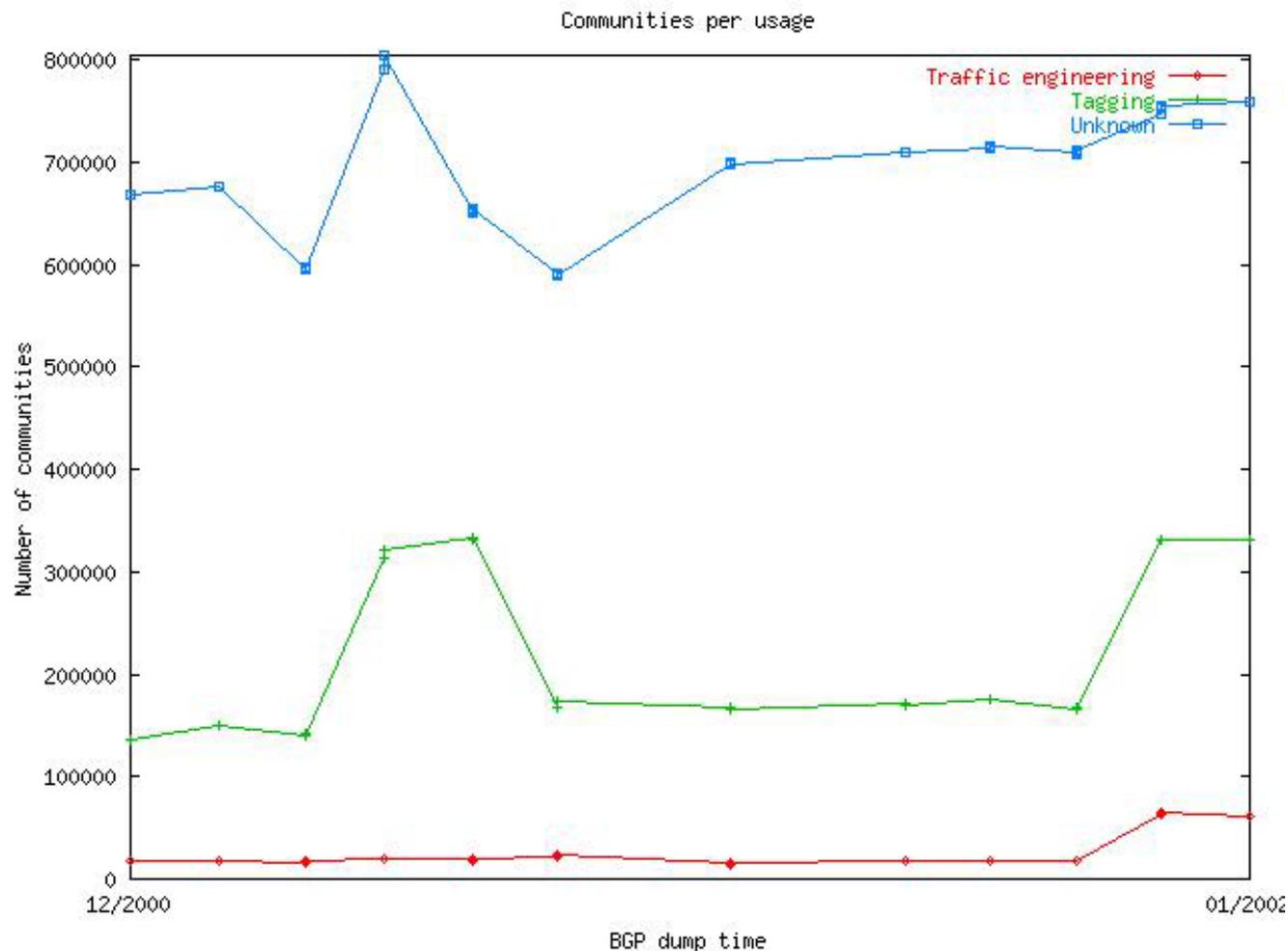
- ☒ "BGP Extended Communities Attribute"
(draft-ietf-idr-bgp-ext-communities-03.txt)
- ☒ タイプコード値は "16"
- ☒ 重複の心配なく community 属性を使える

Community 属性の利用状況 1



<http://alpha.infonet.fundp.ac.be/anabgp/>

Community 属性の利用状況 2



<http://alpha.infonet.fundp.ac.be/anabgp/>

Community 属性の利用状況 3

- ⌘ グラフは <http://alpha.infonet.fundp.ac.be/anabgp/>
- ⌘ RIPE NCC, Amsterdam で観測されたデータ
- ⌘ 全体 (12 万弱) の経路のうち, 約 5 割弱の 5 万強の経路で community 属性が利用されている
- ⌘ “A survey of the utilization of the BGP community attribute”
(draft-quoitin-bgp-comm-survey-00.txt)
- ⌘ 上記 web には, LINX (London), SFINX (Paris), JPIX (Japan) 等のデータもある。community 属性の利用状況はまちまち

Community 属性の利用状況 4

AS	TE	Tag	Others
RIPE NCC, Amsterdam	60345	331316	758089
LINX, London	14371	16283	13315
SFINX, Paris	31	8	261
AMS-IX, Amsterdam	462	356	1868
CIXP, Geneva	11879	5473	3270
VIX, Vienna	39626	42056	14006
JPIX, Otemachi	0	0	0
University of Oregon	314841	388406	2125204

“A survey of the utilization of the BGP community attribute”
(draft-quoitin-bgp-comm-survey-00.txt)

Rpsec 1 (経路制御とセキュリティ)

- ⌘ 多くの WG でセキュリティに関する拡張が議論されている
- ⌘ Routing Protocol Security Requirements (rpsec) Bof
 - ⊞ Chirs:
 - Russ White <riw@cisco.com> ,
 - Danny McPherson <danny@tcb.net>
 - ⊞ 目的
 - ルーティングプロトコルへのセキュリティに関する要求事項の整理とそのモデル作り
 - ⊞ <http://www.ietf.org/ietf/02mar/rpsec.txt>

Rpsec 2 (ビザンチン問題 1)

Rpsec Bof での Radia Perlman による現在の経路障害の問題点の報告

ネットワークの障害は、フェイルストップ、つまり構成要素が突然完全に動かなくなるような場合の他に、ビザンチン問題 (Byzantine failure) 型障害がある。ビザンチン障害では、ネットワーク端末は動作を続け、不正な動作を続ける。フェイルストップ型の障害に比べると、障害の原因の発見が遅れる場合が多く、障害発見後も障害の発生原の特定、障害の解明が困難な場合が多い。

Rpsec 3 (ビザンチン問題 2)

ビザンチン障害は、以下に示すように多くの原因がある。

- ☒ ハードウェアの障害
- ☒ ソフトウェアのバグ, 仕様の問題
- ☒ 設定の誤り
- ☒ 意図的な不正経路の広告等の故意の妨害

今後のインターネットの経路のセキュリティ, スタビリティを高めるためには、ビザンチン障害に耐性のある、ソフトウェア、ハードウェア、ユーザインタフェースの設計が必要である。

Rpsec 4 (経路制御のセキュリティに関する I-D 1)

- ⌘ 1. Flaws in packet's authentication of OSPFv2, Jerome Etienne
<http://www.ietf.org/internet-drafts/draft-etienne-ospfv2-auth-flaws-00.txt>
- ⌘ 2. Flaws in RIPv2 packet's authentication, Jerome Etienne
<http://www.ietf.org/internet-drafts/draft-etienne-ripv2-auth-flaws-00.txt>
- ⌘ 3. BGP Security Vulnerabilities Analysis, S. Murphy
<http://www.ietf.org/internet-drafts/draft-murphy-bgp-vuln-00.txt>
- ⌘ 4. BGP Security Protections, S. Murphy
<http://www.ietf.org/internet-drafts/draft-murphy-bgp-protect-00.txt>
- ⌘ 5. OSPF with digital signature against an insider, Jerome Etienne
<http://www.ietf.org/internet-drafts/draft-etienne-rfc2154-flaws-00.txt>

Rpsec 5 (経路制御のセキュリティに関する I-D 2)

⌘ 6. Secure BGP (S-BGP)

<http://www.net-tech.bbn.com/sbgp/draft-clynn-s-bgp-protocol-00.txt>

⌘ 7. OSPF Version 2 (RFC2328)

<http://www.ietf.org/rfc/rfc2328.txt>

⌘ 8. RIP-2 MD5 Authentication (RFC2082)

<http://www.ietf.org/rfc/rfc2082.txt>

⌘ 9. IS-IS Cryptographic Authentication

<http://www.ietf.org/internet-drafts/draft-ietf-isis-hmac-03.txt>

⌘ 10. Protection of BGP Sessions via the TCP MD5 Signature Option

<http://www.ietf.org/rfc/rfc2385.txt>

<http://www.ietf.org/internet-drafts/draft-ietf-idr-rfc2385bis-00.txt>

Forces 1

⌘ Forces: Forwarding and Control Element Separation WG

Chirs:

Patrick Droz <dro@zurich.ibm.com> ,

David Putzolu <David.Putzolu@intel.com>

⌘ 目的:

☑ ネットワークデバイスのコントロールプレーンとフォワーディングプレーン間のインターフェースプロトコルの標準化。

⌘ 現状

☑ Requirement, Framework, Applicability に関する I-D が出揃ったところ。まだ RFC はない。

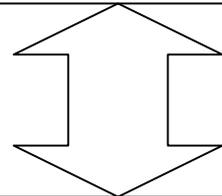
Forces 2 (Forcesに関する I-D)

- ⌘ draft-ietf-forces-netlink-02.txt
 - ☞ Linux の CE-FE 間インタフェースとしての netlink
- ⌘ draft-ietf-forces-requirements-02.txt
 - ☞ Forces 要求仕様
- ⌘ draft-crouch-forces-applicability-01.txt
 - ☞ Forces の適用領域
- ⌘ draft-anderson-forces-arch-00.txt
 - ☞ Forces アーキテクチャ
- ⌘ draft-gopal-forces-fact-00.txt
 - ☞ Forces 準拠プロトコル FACT
- ⌘ draft-gopal-forces-femodel-00.txt
 - ☞ FE の論理モデル

Forces 3 (Forces プロトコル)

- コントロールエレメント(CE)

- 汎用CPU
- シグナリング(RIP, OSPF, LDP, RSVP),
マネジメントプロトコル(SNMP, COPS)

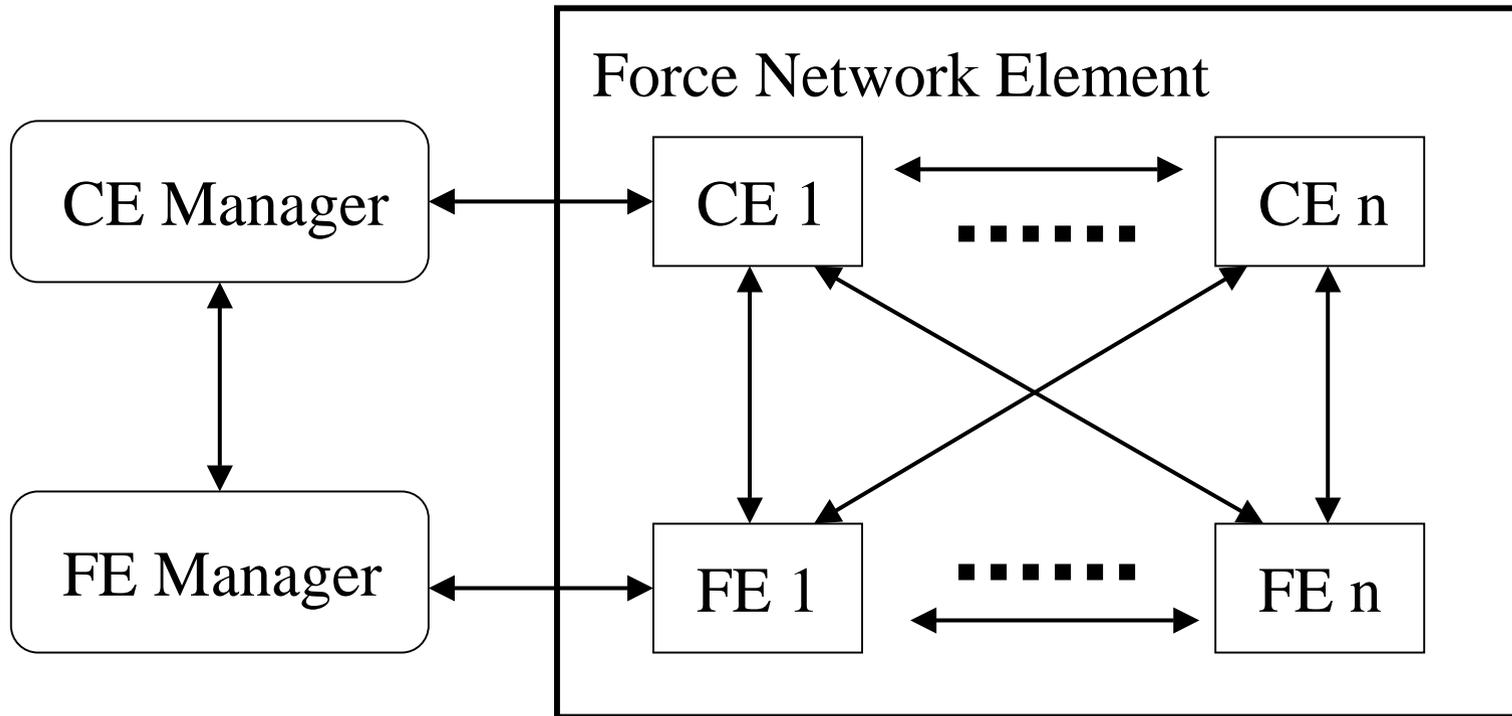


FORCESプロトコル

- フォワーディングエレメント(FE)

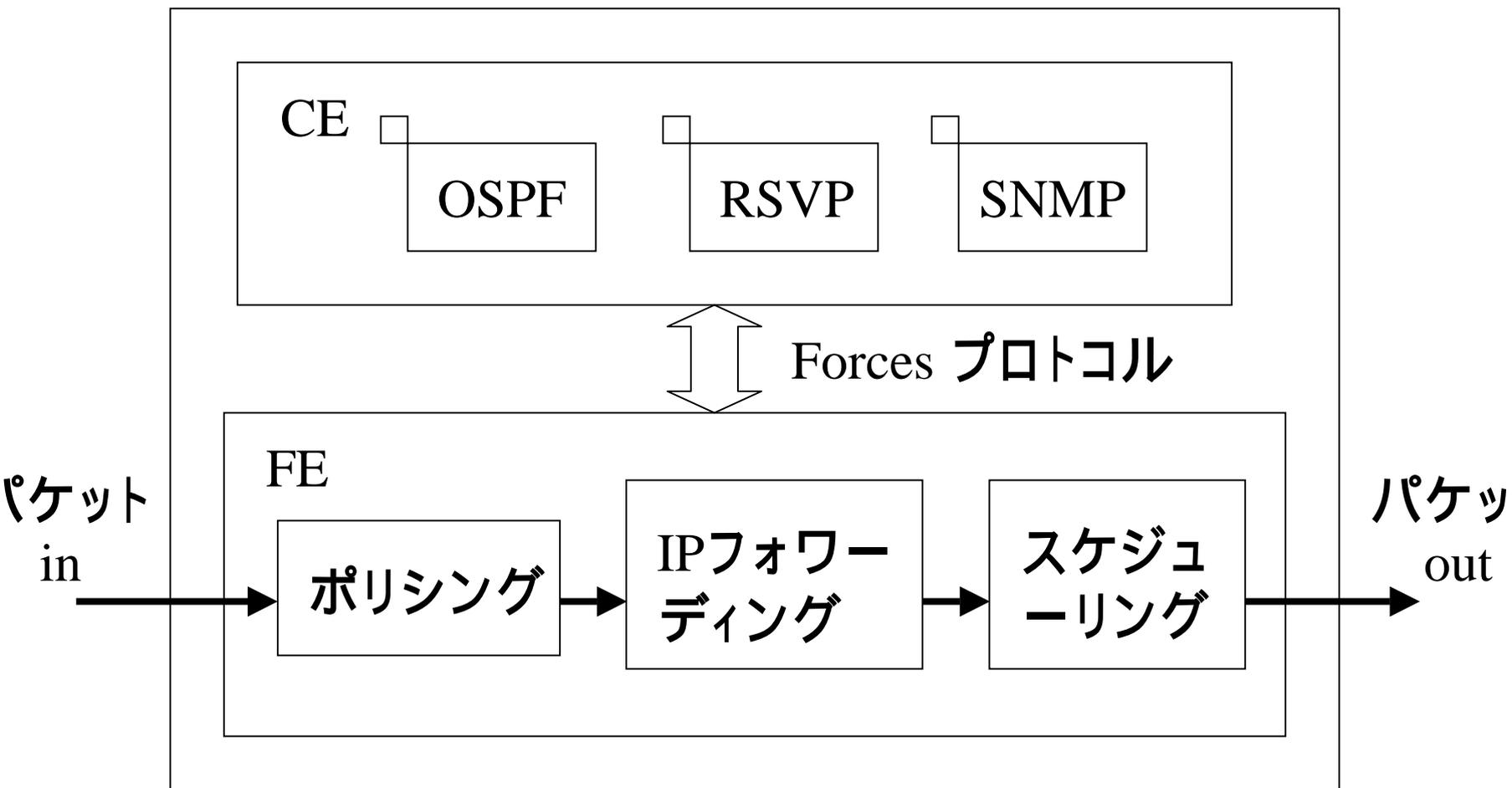
- ネットワークプロセッサ, ASIC
- フォワーディング, ファイアウォール,
NAT, スケジューリング

Forces 4 (Forces アーキテクチャ)



CE (FE) Manager: 各 CE (FE) の通信相手となる FE (CE) を決定

Forces 5 (ルータアーキテクチャ例)



Forces 6 (主な プロトコルRequirements)

⌘ スケーラビリティ

- ⊡ 少なくとも数百 FE と数万ポートのサポート

⌘ セキュリティ

- ⊡ メッセージの認証, 完全性の保証

⌘ 動的アソシエーション

- ⊡ CE および FE の動的な追加, 削除のサポート

⌘ 非同期通知機能

- ⊡ FE上のイベントを非同期でCEに通知する機能

Forces 7 (Netlink)

- ⌘ Linux 上に実装されている Forces ライクな ユーザ-カーネル間 (カーネル-カーネル間含) インターフェース
- ⌘ ソケットを用いた CE (ユーザ空間)-FE (カーネル空間) 通信を実現
- ⌘ ルーティングテーブル操作, ファイアウォール設定, ARPテーブル操作等を実装

Forces 8 (適用領域)

⌘ Applicable

- ☑ Discovery, Capability Information Exchange
- ☑ Topology Information Exchange
- ☑ Configuration
- ☑ Routing, QoS, Security, Filtering Exchange
- ☑ Encapsulation/Tunneling Exchange
- ☑ NAT, Application-level Gateway, Measurement, Accounting
- ☑ Diagnostics, Redundancy

⌘ Not Applicable

- ☑ Label Switching, Multimedia Gateway Control (MEGACO)

性能測定, 性能指標を扱う WG

ネットワークの輻輳による, データの配送遅延等, 現在インターネットで発生している種々の問題, 経路制御も含めたネットワークの性能を正しく表現するパラメータ, 測定方法の確立が重要となってきている

- ⌘ bmwg (Benchmarking Methodology WG)
- ⌘ ippm (IP Performance Metrics WG)
- ⌘ ipfix (IP Flow Information Export WG)

まとめ (問題はセキュリティとスケーラビリティ)

⌘ <http://www.ietf.org/proceedings/02mar/index.html>

⌘ セキュリティ

- ⊞ どの WG もセキュリティ関係の議論が多い, I-D 乱立状態
- ⊞ Rpssec bof の登場により, プロトコルに共通した統一モデルへの動き

⌘ スケーラビリティ

- ⊞ IPv4 アドレス空間の枯渇
- ⊞ 経路表の増加
- ⊞ AS という仮想集合の有効性と限界
- ⊞ 回線の高速化, トラフィック量の増加によるルータの高速化

⌘ マルチキャスト (PIM WG, MSDP WG)

- ⊞ SSM (Source specific multicast) 登場後, あまり動きはないよう
- ⊞ UDLR WGから動きがあるかもしれない