

# IX最新動向

---

IRS14

11 Oct. 2007

Internet Multifeed Co. / JPNAP

<http://www.mfeed.co.jp/>

Toshinori Ishii

石井利教

All communication flows through here.

全てのコミュニケーションはここを通る



## 自己紹介

- インターネットマルチフィードのエンジニア
- JPNAPの運用もしています
- 日本のIXやブロードバンド事情について海外でプレゼンを試みたりもします
- 「最新動向」という名の下に取り留めのない話を...

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



## 初心に戻って...

- IX(IXP)とは

※JPNIC/APNICの定義より抜粋

- Internet eXchange Point の略
- 物理的なネットワーク
- 独立したISP間でのインターネットトラフィックの交換(=ASを持つISP同士をつなぐグローバルネットワーク)
- 接続ISP数は最低でも三つあるべき
- 参加するための明確でオープンなポリシーがなければならない

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



# 今日は扱いません

- Layer3 IX
  - 自分自身がASを持ち、IPトランジットを行うネットワーク
  - 技術的にはISPそのもの
- SIP-IX
  - SIP事業者間で互いにトラフィック交換を行うネットワーク
  - アプリケーション層での交換
  - 技術的にはNNTP配送系やIRCサーバと似ている

IXと名前はついているけれど、どのレイヤで情報を交換するかで技術が大きく異なる。

日本で通常IXと言うとLayer2 IXを指す場合がほとんど。

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



## Layer 2 IX

- ほとんどのL2IXはEthernetベース
  - みんなが慣れている(運用者、ユーザ)
  - パスの設定が不要
  - 製品も比較的安い
- (比較的)安定し、枯れている技術

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



# IXの現状

All communication flows through here.

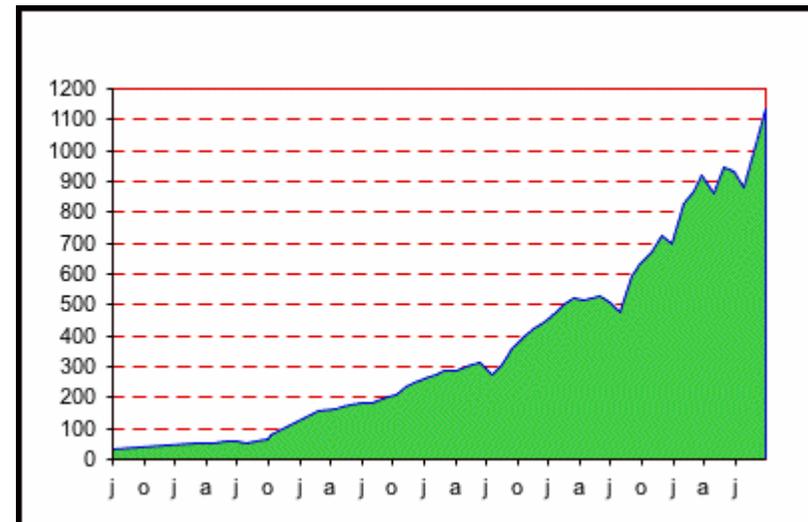
全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



## 世界のIXトラフィック動向

- 欧州はIXメイン
  - プライベートは少ない
  - 国が近いいため冗長化が図りやすい
- 米国は下火
  - 広大な国土
  - Tier 1 支配、トランジット
- アジアは国による
  - ISPが多く自由競争の行われている国ではIXトラフィックが多い



ヨーロッパのIXトラフィック動向 (Gbps)  
2007.09.27現在

j: January  
a: April  
j: July  
o: October

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



# Euro-IX

## • IXの連盟

- 元は欧州の連盟
- IXのみで構成される(RIPE EIX, APNIC IX-SIGとの違い)
- パトロンと呼ばれるスイッチベンダーによるスポンサー制度あり
- JPNAPと意気投合し欧州以外に準会員制度を作る

Member IXPs:	36
Associate member IXPs:	7
Countries represented:	25
Total number of IXP sites:	229
Listed IXP customers in DB:	3.094
Listed ASNs in DB:	2.872
Unique ASNs in DB:	1.932
Aggregated IXP peak traffic:	1.132Tbps
Average IXP peak traffic:	34.311Gbps
Total IXP switches in DB:	203
Cisco switches in DB:	88
Extreme switches in DB:	19
Foundry switches in DB:	67
Force10 switches in DB:	9
Glimmerglass switches in DB:	9
Other vendor switches in DB:	11

All communication flows through here.

全てのコミュニケーションはここを通る

2007.09.27現在

(c) INTERNET MULTIFEED CO.



## アジアのIX

- 通信の自由化が進んでいない国が多い
  - ISPの独占が強く、IXも支配している
  - IXが準国営
- ISPの数が少ない
- 通信の多くが海外向け
  - 自国内ピアリングをしようという気にならない
  - 自国内コンテンツが充実していない

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



# 日本のIXとトラフィック

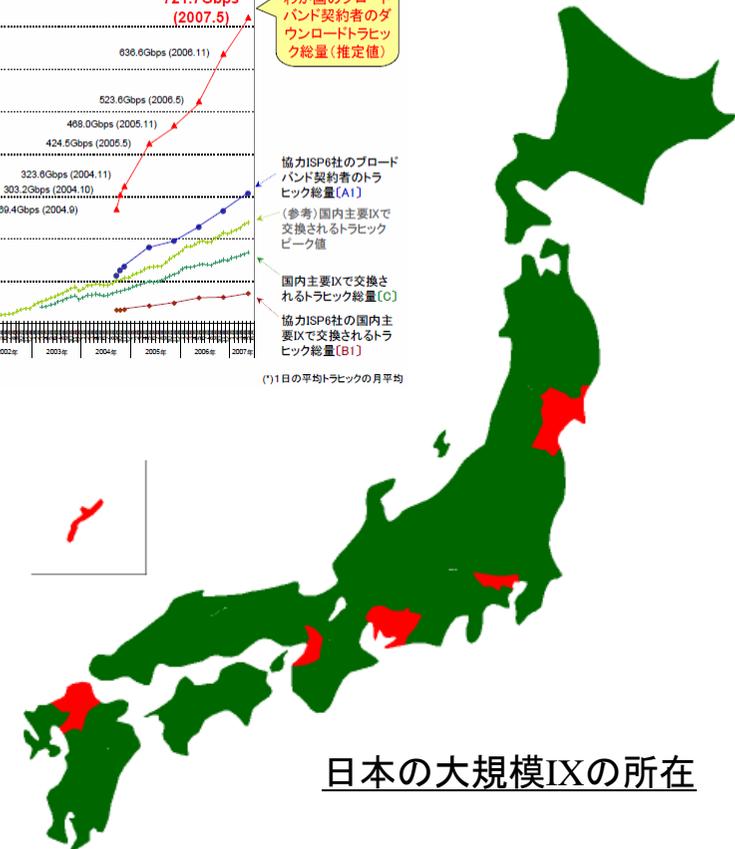
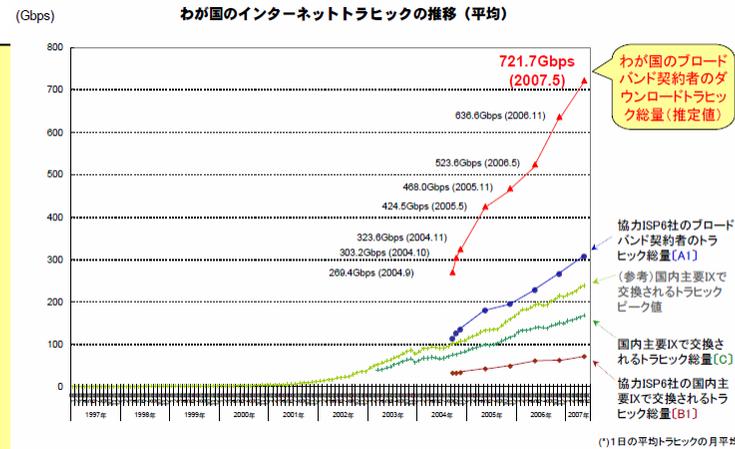
## ● 代表的なIX

- BBIX
- DIX-IE\*
- JPIX\*
- JPNAP\*

\* : 総務省統計に含まれるもの

## ● 日本のトラフィック

- 約1/4が代表的なIXを経由
- IX線のバックアップにはトランジット線を使うISPが多い模様

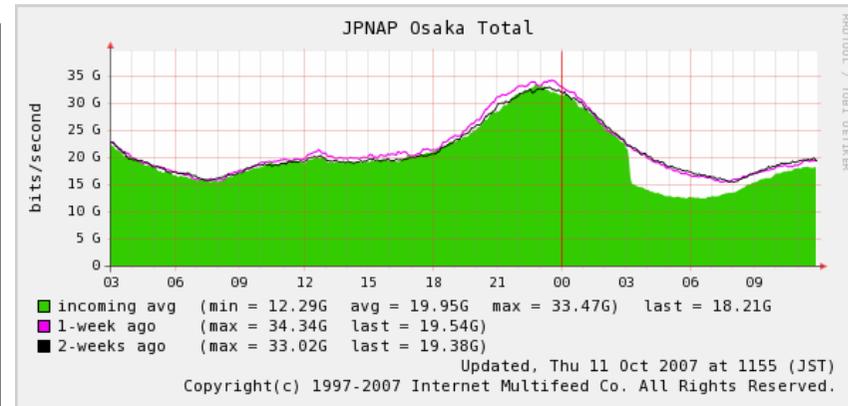


All communication flows through here.

全てのコミュニケーションはここを通る

# 東京と大阪

- JPNAPの例
  - JPNAP
  - JPNAP大阪
  - 500km離れている
  - 相互接続していない
- 大阪のトラフィック
  - 順調に伸びている
  - 35Gbps
  - 世界10位に入るIXにまで成長



All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



# IXの障害対策

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



## 耐障害性-拠点分散

- 場所を変えることで障害の波及を防ぐ
  - 災害による拠点障害
  - ループ等ネットワーク障害
- Netnod
  - スウェーデン国内5ヶ所
  - 相互接続していない



All communication flows through here.

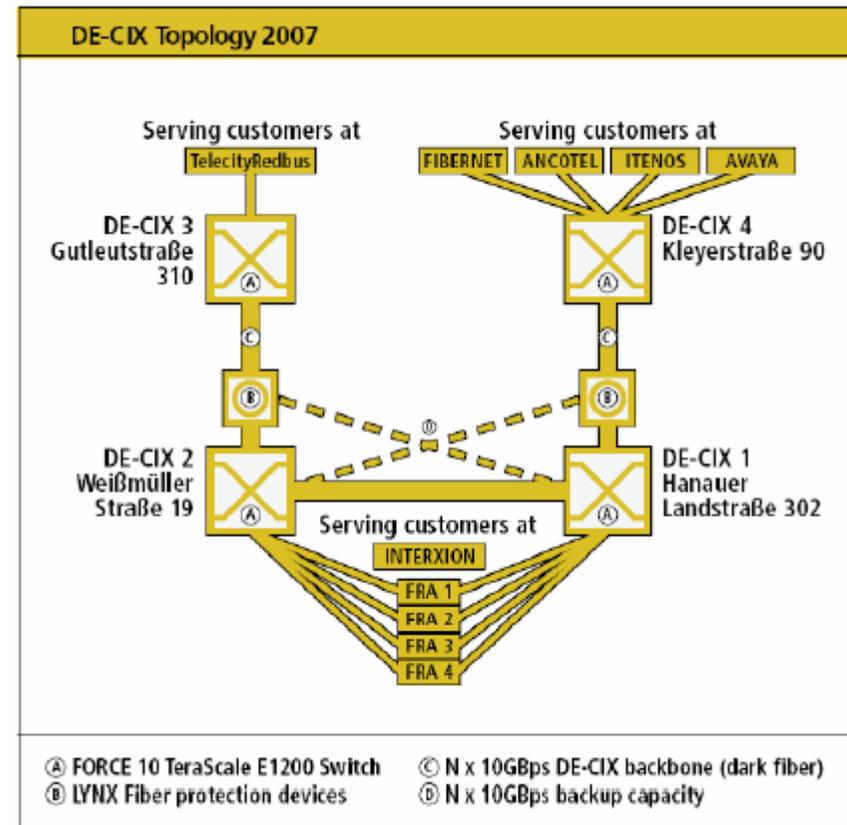
全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



# 耐障害性-局所化

- 光スイッチによる中継路の切り替え
  - コア拠点の障害回避
  - 切り離すことで障害を局所化する
  
- DE-CIX
  - フランクフルト近郊4ヶ所
  - コアサイトと相互接続

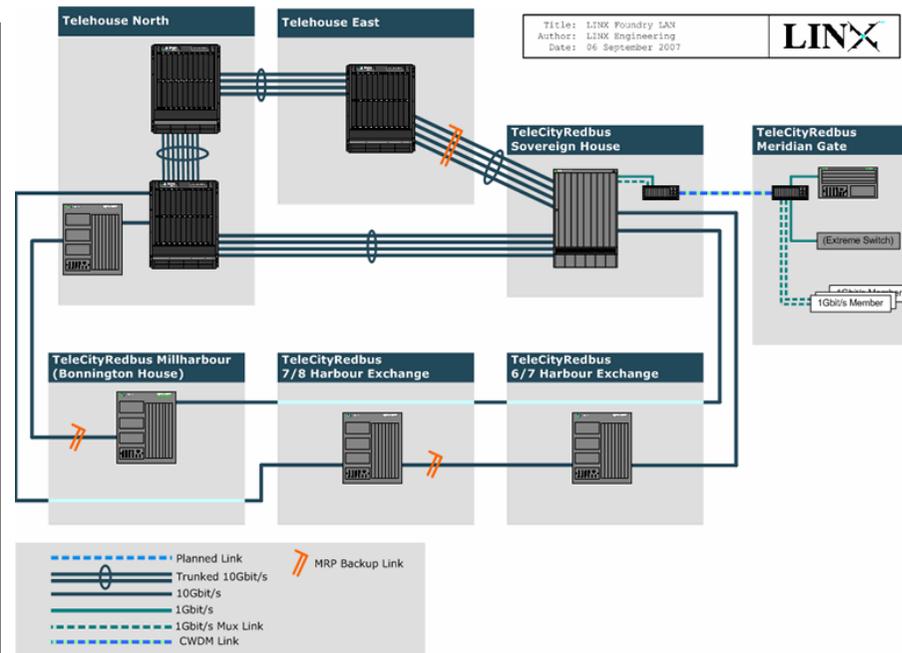


All communication flows through here.

全てのコミュニケーションはここを通る

# 耐障害性-冗長化

- リング構成による中継路の切り替え
  - 中継回線の障害回避
- LINX
  - MRP, EAPSによる冗長化
  - 図の他にもう一面あり

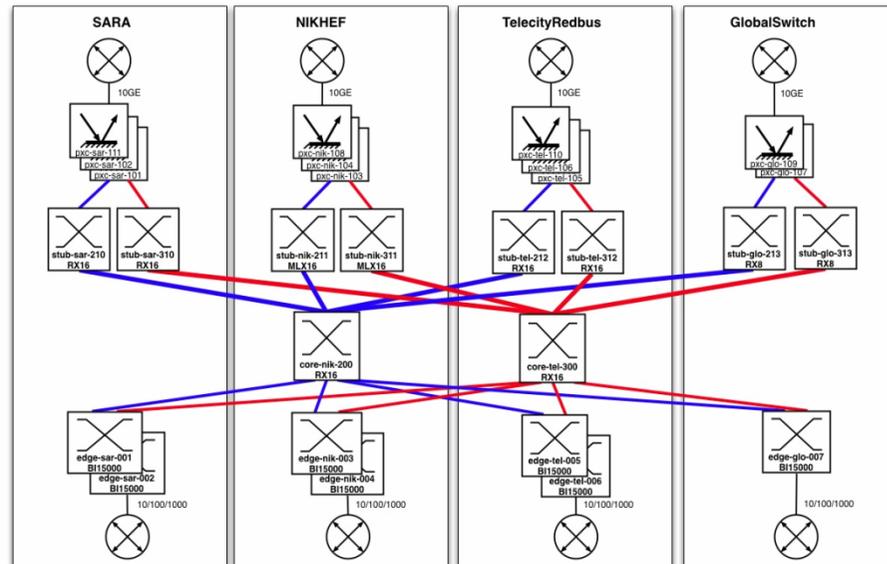


All communication flows through here.

全てのコミュニケーションはここを通る

# 耐障害性-バックアップ

- バックアップ系への切り替え
- AMS-IX
  - 光スイッチにより系全体をバックアップ系に切り替える
  - VSRPを用い、メイン系のFDBをバックアップ系に移す

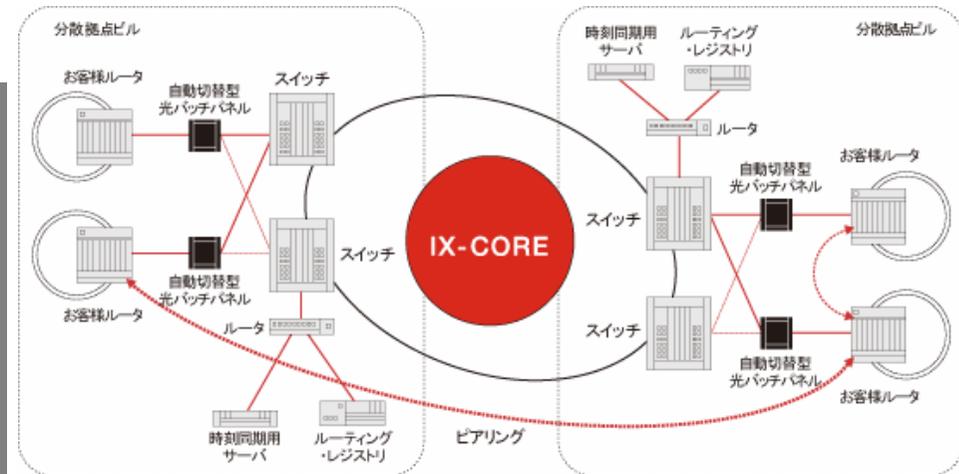


All communication flows through here.

全てのコミュニケーションはここを通る

# 自分たちを振り返る

- 拠点分散・局所化
  - JPNAP東京Ⅱ  
<http://www.mfeed.ad.jp/press/2007/20071002.html>
  - 現JPNAP(東京Ⅰに改称予定)と相互接続しない
  - 大手町エリアに大規模障害が発生しても生き残れるネットワーク構成を実現
  - 東京の利便性と障害回避の両立を目指す
  - 500km離れた大阪も有り



- 冗長化
  - リング構成(右図)
- バックアップ
  - 自動切替型光パッチパネルによるバックアップ系スイッチ収容

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



## その他障害対策関係

- リンク断検知、片線断検知
  - 片端だけDownしてパケロスやループすることを防ぐ
  - とくにメディコンが間に挟まっているときに有効
  - 例) Auto nego, LFS, UDLD, BFD, リンク断伝達機能 etc
  - 古い技術でも未だに相互接続性が...☹
- 冗長化プロトコル
  - リングトポロジーなどで使われる
  - 例) STP, RSTP, MRP, EAPS, etc
  - IX網内に閉じているので相互接続性の確認は少し楽
  - 外からの流入を防ぐのが重要

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



# IXの将来

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



## IXの付加価値

- もともとL2サービスのIXではサービスの差別化要因は少ない
- sFlowによるトラフィック解析
  - 現時点ではL2SWでフロー情報を得るにはsFlowしか選択肢がない
  - 例) IX加入者ルータ間のpeeringトラフィック表示など
  - JPNAPでは2005年からサービス提供開始 (PeerWatcher®)
  - 最近、各地のIXで流行中

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



# IXに流れるものの移り変わり

- IPv6
  - NSPIXP, JPIX, JPNAPともにIPv6専用IXを併設
  - 日本はIPv6への取り組みが早かったため、ISPから専用IXを作ることを求められた(ルータの安定性を不安視?)
  - 現在はデュアルスタックが主流
  - デュアルスタック化すると、sFlowでしかトラフィックを測定できない
- 4Byte AS
  - 4-2-4-2のようにつながるとヤバいルータがいる?
  - JPNAPに4バイトASが2ついるので勇気のある人は...☺

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



## 高トラフィック対応

- 100G Ethernet
  - IEEE802.3HSSGにて検討中
  - 100GEだけでなく40GEも標準化されることになった
  - 標準化のスケジュールはまだまだ(2010年春予定)
  - 2009年夏ごろ仕様がほぼ固まるので、それ以降で100GEを独自規格で出してくるメーカーがある模様
  - 日本メーカーもがんばれ
- それまでは既存の技術でなんとか対処するしか...

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



## 将来までのつなぎ

- LAG(Link Aggregation)
  - 複数本の回線を論理的に1本に見せる
  - 耐障害性の向上にも使える
  - 最近のスイッチは、L2設定であってもL3やL4情報をもとにロードバランスすることが可能
  - 帯域の効率利用ができる
  - ユーザ収容ポートでLAGした場合には、IPアドレスを追加する必要がないのでpeerの移設の必要がない(ISPにとってのメリット)
  - LACPは相互接続性に問題がある場合も☹

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



# IXP Switching “Wishlist” (1)

- MACアドレス学習の制限 ※RIPE EIXより抜粋
  - ループ防止
  - ルータ以外の機器の接続の制限
- RSTP BPDUの無視
  - 網の混乱の元
- ワイヤースピードでのL2/L3ヘッダによるACL
  - 不要なether typeを止める、など
- Broadcast, Multicast, Unknown unicastのコントロール
  - 該当パケットの監視、制限
  - OSPF、RIPの検知

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



## IXP Switching “Wishlist” (2)

- ARP snooping
  - 範囲外IPアドレスの使用禁止
- MARP?のサポート
  - MultiAccess Reachability Protocol?(存在しない)
  - L2多段スイッチ構成での故障を検知するプロトコル
- ロギングの柔軟化
  - syslogファシリティやレベルの柔軟な設定
  - syslog over SSL

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



## IXP Switching “Wishlist” (3)

- 管理インタフェース
  - 10G IF での management VLAN設定
  - telnet も snmp も sFlow も同じIFで使いたい
- (番外編)sFlow
  - ある特定のACLに引っかかったパケットだけ、別のサンプリングレートでsFlowパケットを出す

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



## おわりに

- IXトラフィックは順調に伸びています。
- 処理トラフィックが増えるにつれ重要度も増してきています。
- 今後も安心して使っていただけるような網を設計、運用していきます。
- スイッチメーカーの頑張りにも期待しています。

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.



# Thanks To

---

- ご清聴いただいた皆様
- データを提供してくれた方々
  - Arnold Nipper (DE-CIX)
  - Kurtis Lindqvist (Netnod)
  - John Souter (LINX)
  - Steven Bakker, Cara Mascini (AMS-IX)
  - Serge Radovcic (Euro-IX)
- 引用
  - [http://www.soumu.go.jp/s-news/2007/070822\\_2.html](http://www.soumu.go.jp/s-news/2007/070822_2.html)
  - <http://www.ripe.net/ripe/wg/eix/index.html>

All communication flows through here.

全てのコミュニケーションはここを通る

(c) INTERNET MULTIFEED CO.

