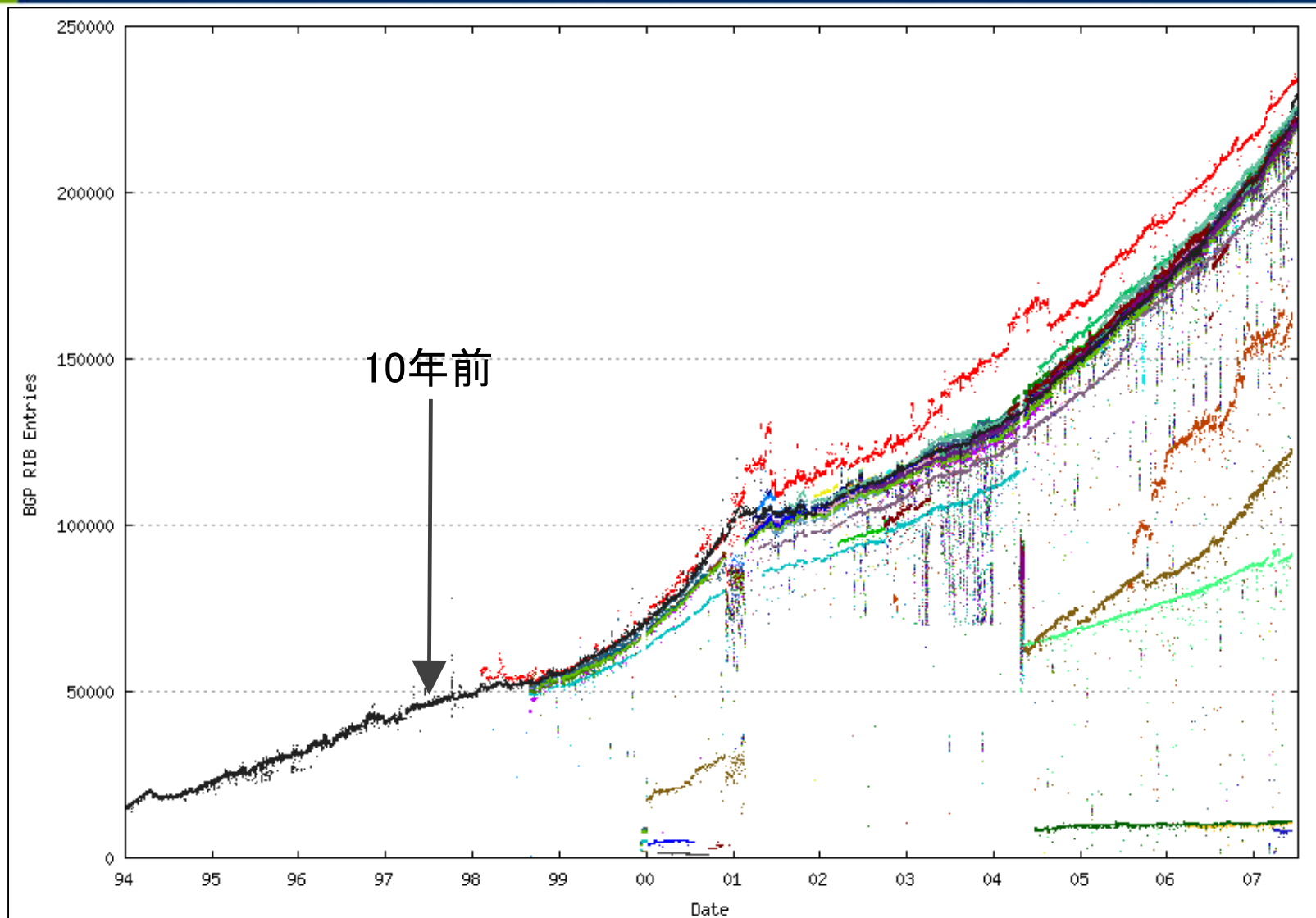


インターネット基盤技術はこれからの インターネットを支えきれるか？

石黒邦宏



BGP経路



JANOG10周年おめでとう！

- 10年前のインターネット経路は約5万経路
 - 今は20万経路以上
 - ここにいる皆さんがいなければ動かなかったでしょう
- 世界中のコンピューターを一つのネットワークにつなぐという夢は叶ったか？
 - 大きなチャレンジだったが、今や実現しつつあるのではないか
 - 携帯が接続され、電話を飲み込み、さらには放送へ
- インターネット・サービス・プロバイダーとは何だったか？
 - 電話会社主導の通信事業の中で果たしたISPの役割
 - そして我々インターネットオペレータとは？
- 社会のコミュニケーション基盤へ
 - IP技術への過度な期待と負荷
 - 自身の成功の犠牲者になる可能性もある

さて、10年前にタイムトリップすると

- Googleはまだありませんでした
 - 1998年9月設立
- Yahoo!はもうありました
 - 1995年3月設立
- Yahoo! Japanもありました
 - 1996年1月設立
- Amazonはもうありましたが
 - 1995年7月設立
- Amazon.co.jpはまだありません
 - 2000年11月設立
- 楽天はできたてです
 - 1997年2月7日設立
- ライブドアの前身のオン・ザ・エッジもありました
 - 1996年4月設立

ADSLはまだありません。ISDNを覚えていますか？

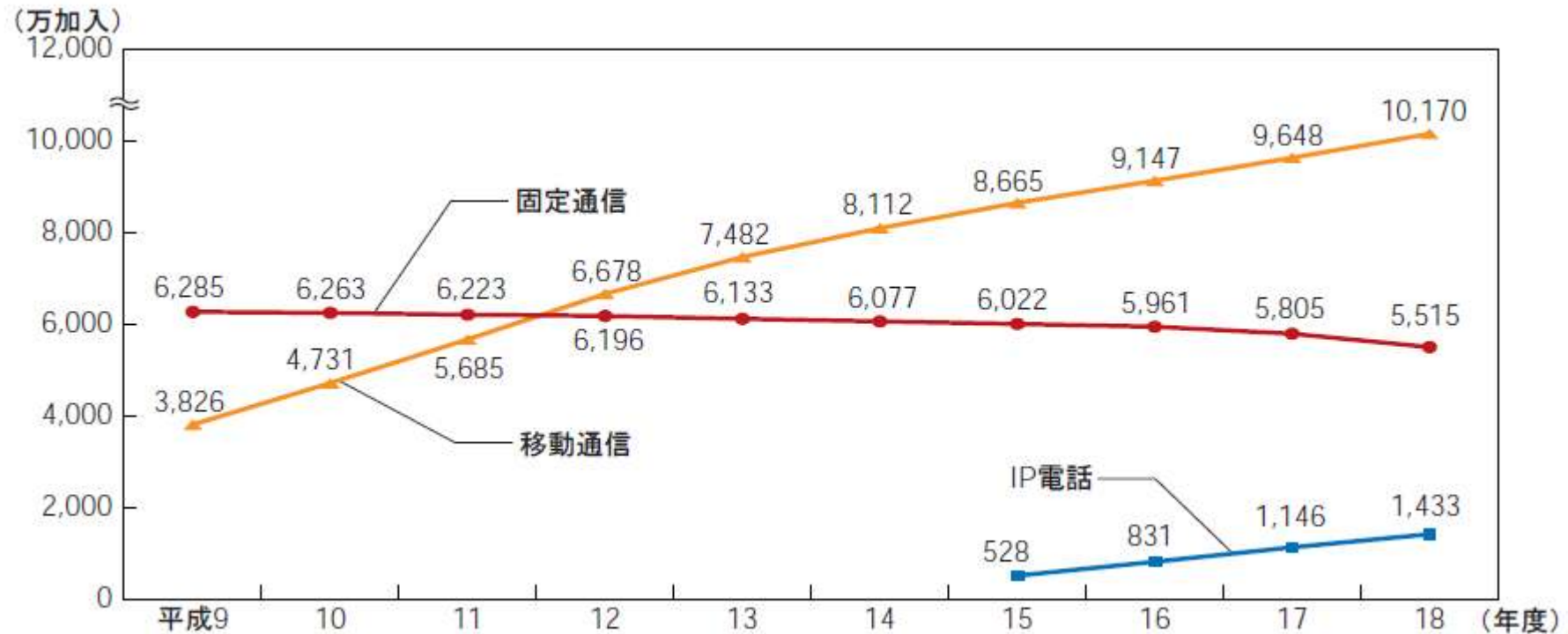
- Yahoo! BBはまだありません
 - 2001年サービス開始アナウンス
- 大丈夫！テレホーダイがあります
 - 深夜早朝の時間帯(23時～翌日8時)に限り、予め指定した2つまでの電話番号に対し、通話時間に関わらず料金が月極の一定料金となるもの(From Wikipedia)
 - 1995年より開始
- 大丈夫！rt100iがあります
 - rt100i-users メーリングリスト
- 大丈夫！MN128-SOHOがあります
 - ISDNターミナルアダプター＋ルーター機能
- Ascendという会社がアメリカにありました
 - pipeline 50, pipeline 75, MAX...

携帯からメールを送ろうと思っははいけません

○ i-mode はまだありません

- 1999年2月サービス開始

○ 携帯加入者数が固定電話加入者数を抜くのは2000年

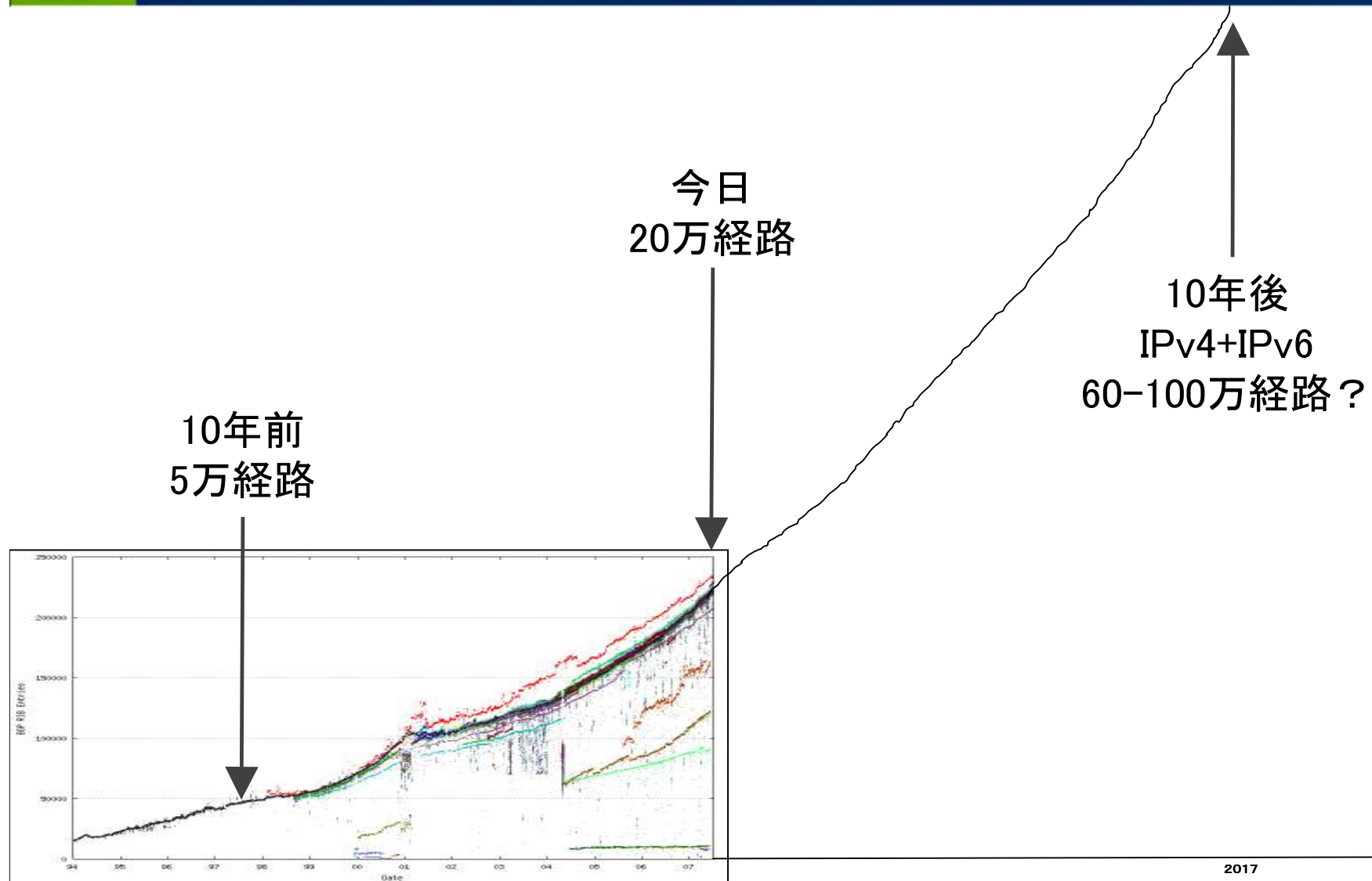


平成19年 情報通信白書より

次の10年

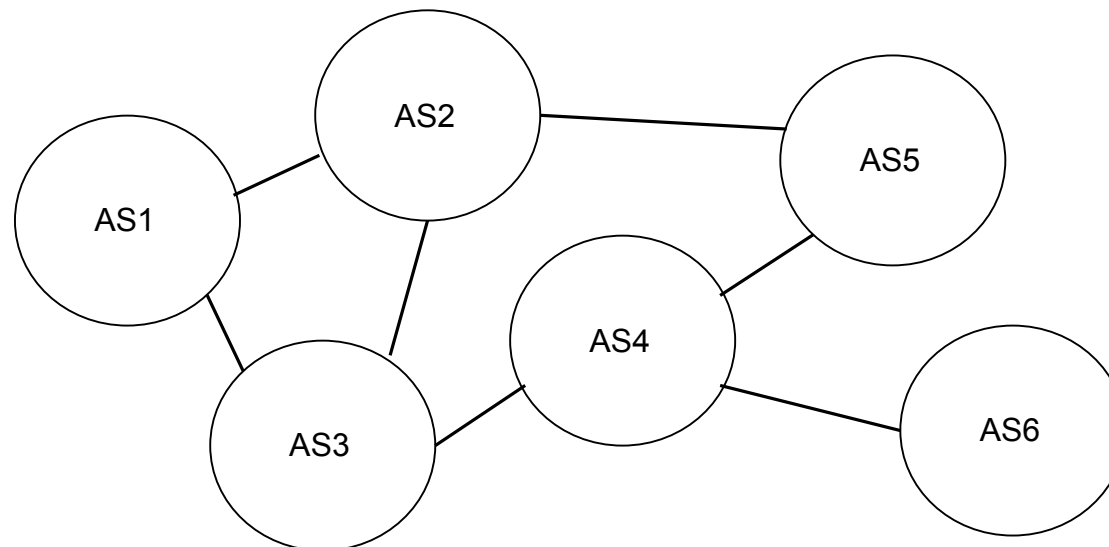
- インターネットのさらなる拡大
- 国内インターネット
- IP電話
- 専用線
- 網内経路制御問題
- 顧客トレンド
- 国内の電話の現状
- 非インターネットIP動画配信
- Ethernet高速化
- パケットネットワークによる電話網の構築

BGP経路表問題



インターネットの仕組み

- 自分の住所を人に伝え、聞いたひとはそれを次の人に伝えるという巨大なやぎさん伝言システム
- ひとまどりのネットワーク管理を指してASと呼ぶ
- AS間とAS内と2つのレベルの伝言システムがある



IAB - Routing and Addressing Workshop

○ IAB(IETFの上部団体)によるワークショップ

- 経路制御とアドレッシングに関するワークショップ
- 昨年2006 年10月18-19日にオランダのアムステルダムで開催
<http://www.iab.org/about/workshops/routingandaddressing/index.html>

○ ワークショップ報告書(draft-iab-raws-report-02.txt)

- はじめに
インターネットの経路制御およびアドレス割当てシステムが深刻なスケーラビリティ問題に直面していることは広く認識されている。これまでにないユーザー人口の急激な増加、同時にマルチ・ホーミング、トラフィック・エンジニアリングそしてポリシー・ルーティング等々の各種要因が、警告的な割合での経路表の増加をもたらしてきた。現行の経路制御アーキテクチャが深刻なスケーラビリティ問題を抱えてるかもしれないことは、長い間認識されていたにもかかわらず、効果的な解決策はいまだ見つけられていないし、開発もされておらず、普及もしていない。
- 緊急度(Section 2.4)
2.1(経路表増加)と、2.2(アドレス割当てシステム)問題については即座に行動を起こす必要がある。...中略...そして、数年以内に効果的な解決策が開発され、普及する必要があるだろう。

draft-yu-scaling-01-j.txt

- 1999年6月にJANOGで翻訳した文章
- 経路制御プロトコルの規模対応性
 - 今日、一般的に実装されている経路制御プロトコルは、内部経路制御(IGP)には IS-IS や OSPF、そして、外部経路制御(EGP)には BGPが使用されている。
 - 規模対応や他の観点において、これらのプロトコルは既にRIPやEGPのような前の世代のプロトコルから改善されている。しかしながら、規模対応性は、ネットワークが大規模な場合や、経路制御設計が規模対応に関する問題をあまり考慮していなかったり、あるいは、プロトコルの実装が非効率であるようなとき、重要な問題としていまだに存在している。
- ほとんど、1999年から問題は解決しておらず状況は悪化する一方か？

どれがいちばん問題なのか？AS間？AS内？

○ AS間経路制御

- Internet全体にインパクト
- 議論の中心はこれ

○ AS内経路制御

- もしかするとこちらのほうがより深刻かもしれない

○ ネットワーク規模の推測

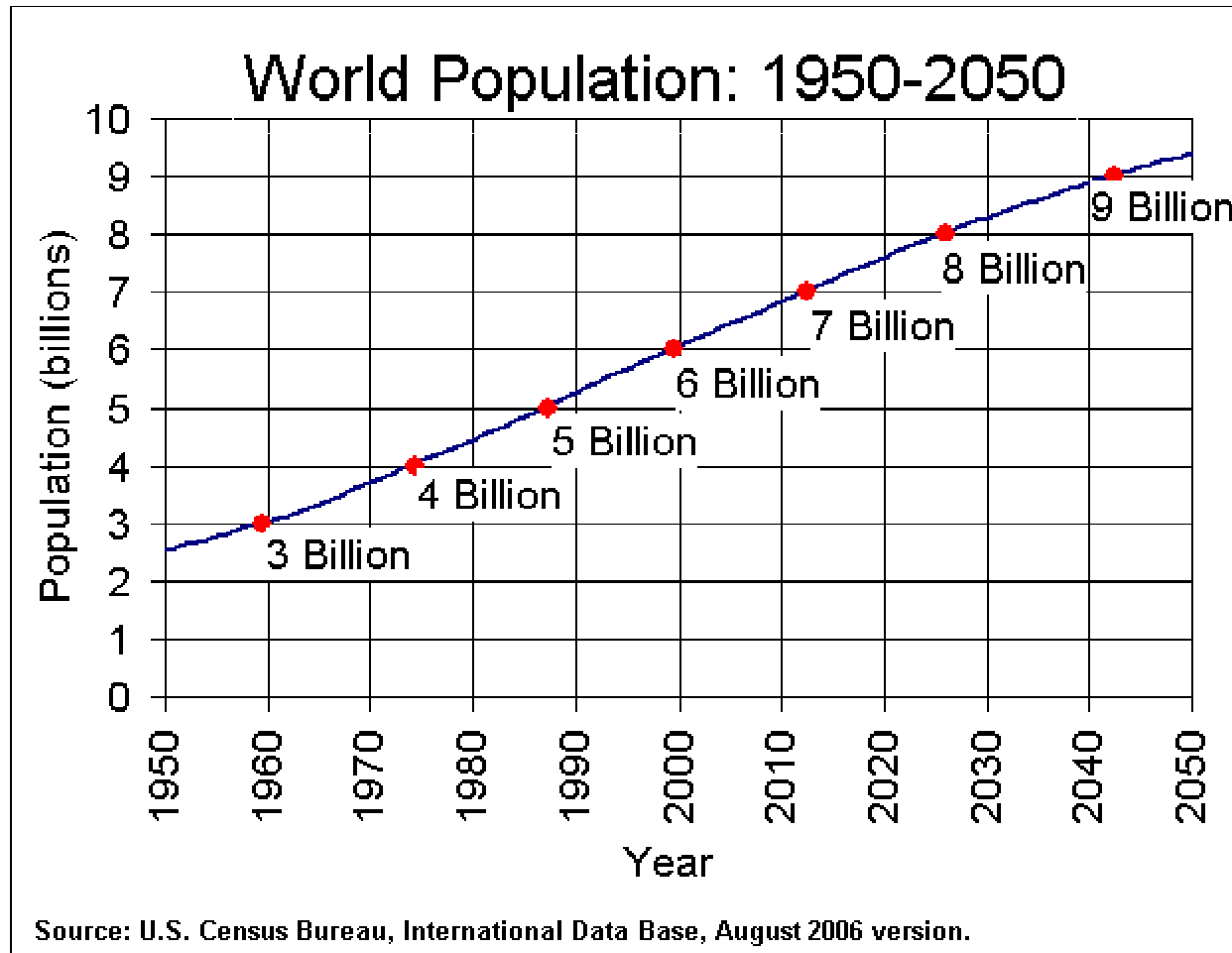
- IP網はインターネット網だけではない
- 固定電話網？携帯電話網？
- 放送網？

AS間経路制御 = EGP = BGP

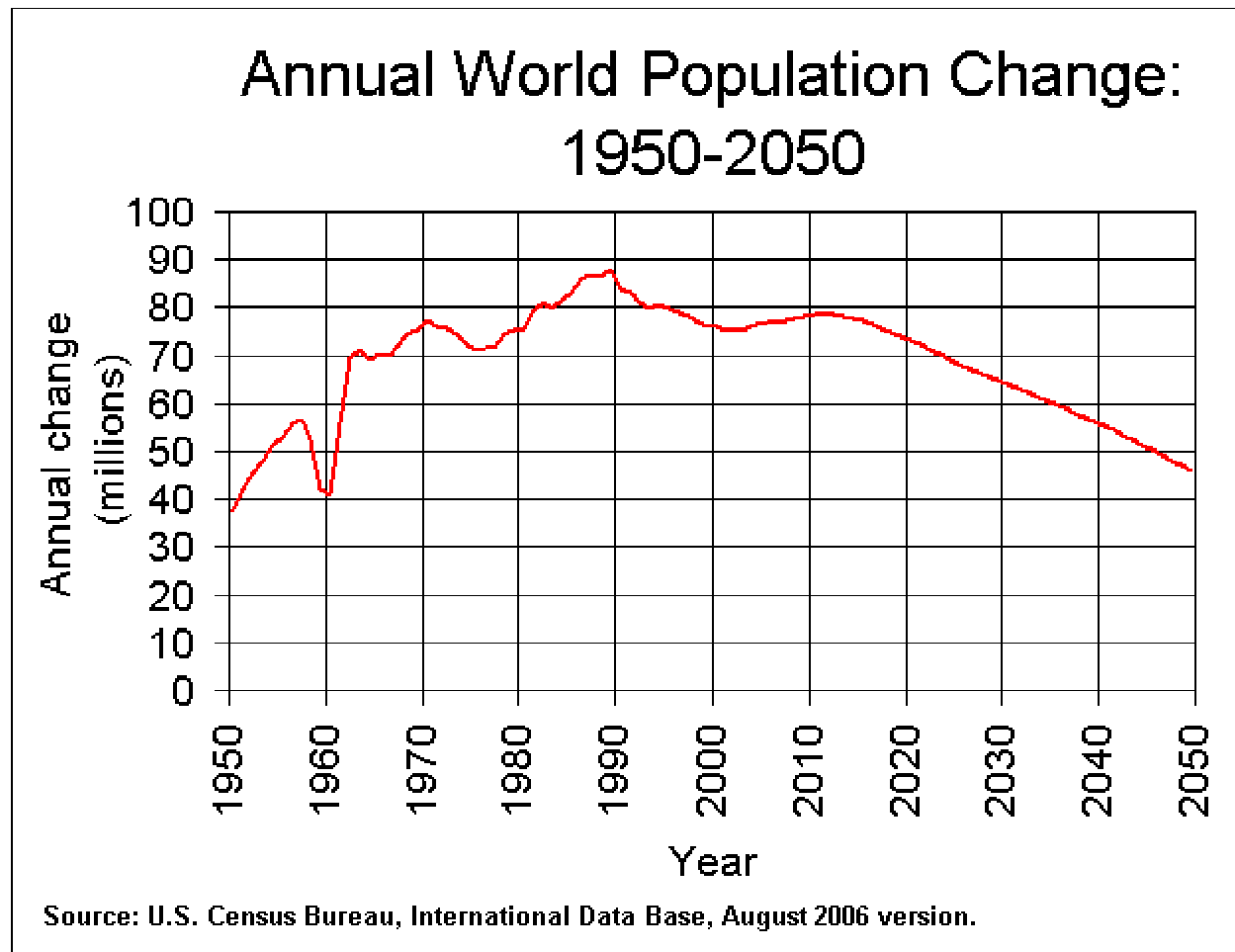
○ BGPの問題と各種提案

- BGPの経路数問題(20万+経路)
- Multihome Punching Hole -> LISPで減らす？
- BGPのコンバージェンス時間問題
 - MRAI Timer
 - Dampening Off
 - Path Length Dampening
 - Optimal Path Hysteresis
 - Delayed Best Path Selection
 - Interface Damening/no fast-external-fallover/BFD etc...
- BGPトランスポートパフォーマンス問題
 - BST(SCTP)
 - Peer-group
 - Large TCP Buffer Size/No Slow Startup
 - Path MTU Discovery

地球人口規模



人口の増加数



人口の増加速度



IPv4アドレス枯渇問題

- 経路制御の側からみるとちょうどいいマイルストーンかもしれない
- IPv4インターネットの規模が確定される
 - 半スタティック・ルーティングが可能になるかもしれない
 - ルーティングレジストリーからエントリー作成してルーターに注入？
 - なので、IPv4アドレスの売り買いとかはあまりしないで欲しい
- 最終的なIPv4 Default Free経路数は？
 - 5年後に新規割り当てが停止されたとして、50万経路？
- その後IPv4経路を減少させる方策
 - IPv6 only ネットワークに完全移行したらIPv4経路をwithdrawする？
 - LISP(Locator/ID Separation Protocol)？

世界のインターネット接続性

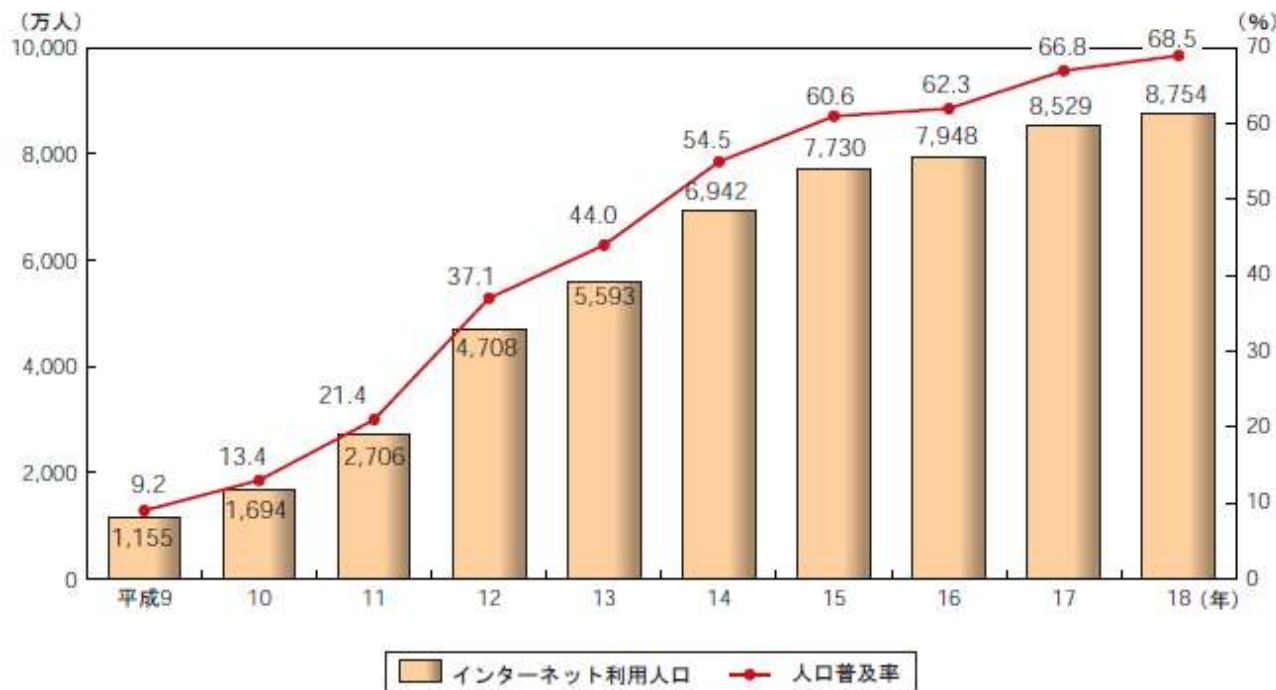
- 人口はこれまでリニアに近い速度で増加してきたが速度は徐々に減少
- これまでに接続されていない地域が接続するため、経路は増え続けるだろう
- 日本などのほぼ全人口にインターネット接続がいきわたった国の経路増加は鈍化する可能性がある
- マルチホームへの要求はさらに高まる
- IPv4+IPv6
- さまざまな要因のため今後の増加のカーブは非常に予測じづらい
- BGPはPath Vectorアルゴリズムなので、死に方は局地的かもしれない

次の10年

- インターネットのさらなる拡大
- 国内インターネット
- IP電話
- 専用線
- 網内経路制御問題
- 顧客トレンド
- 国内の電話の現状
- 非インターネットIP動画配信
- Ethernet高速化
- パケットネットワークによる電話網の構築

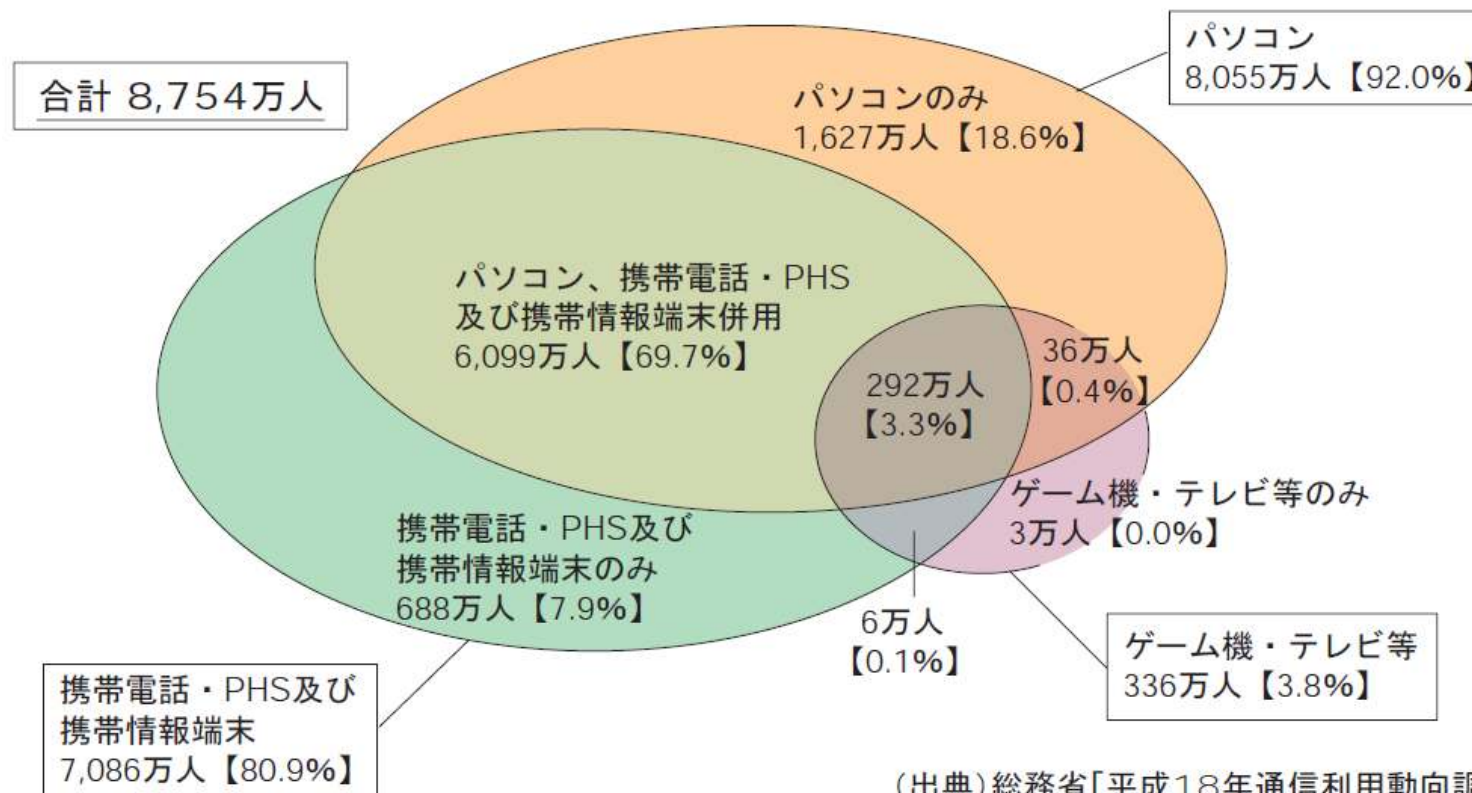
国内インターネット普及状況

- インターネット利用人口が9000千万人弱
- 総人口は1億2775万人
- 15歳から64歳人口は8355万71千人
- 人口普及率70%弱



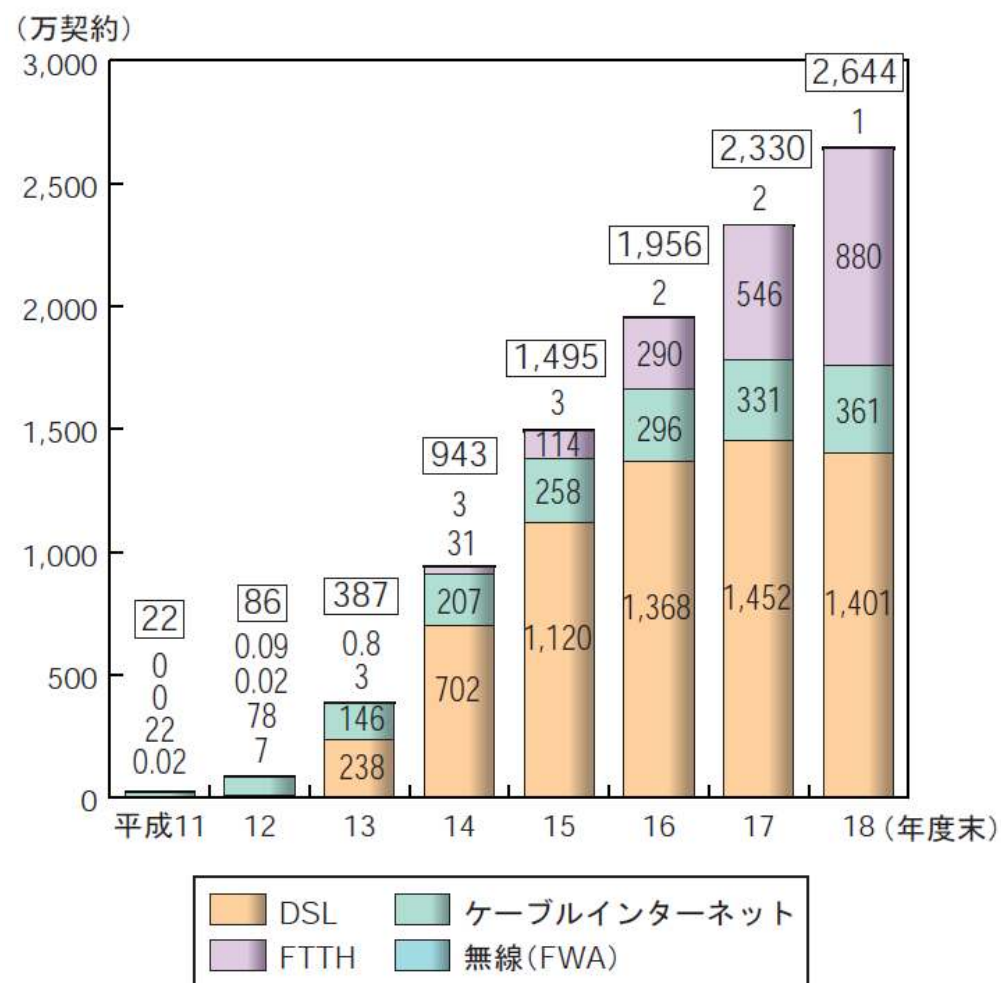
インターネット利用端末の種類

- 今後さらに携帯電話、ゲーム機、テレビからのアクセスが増えるだろう



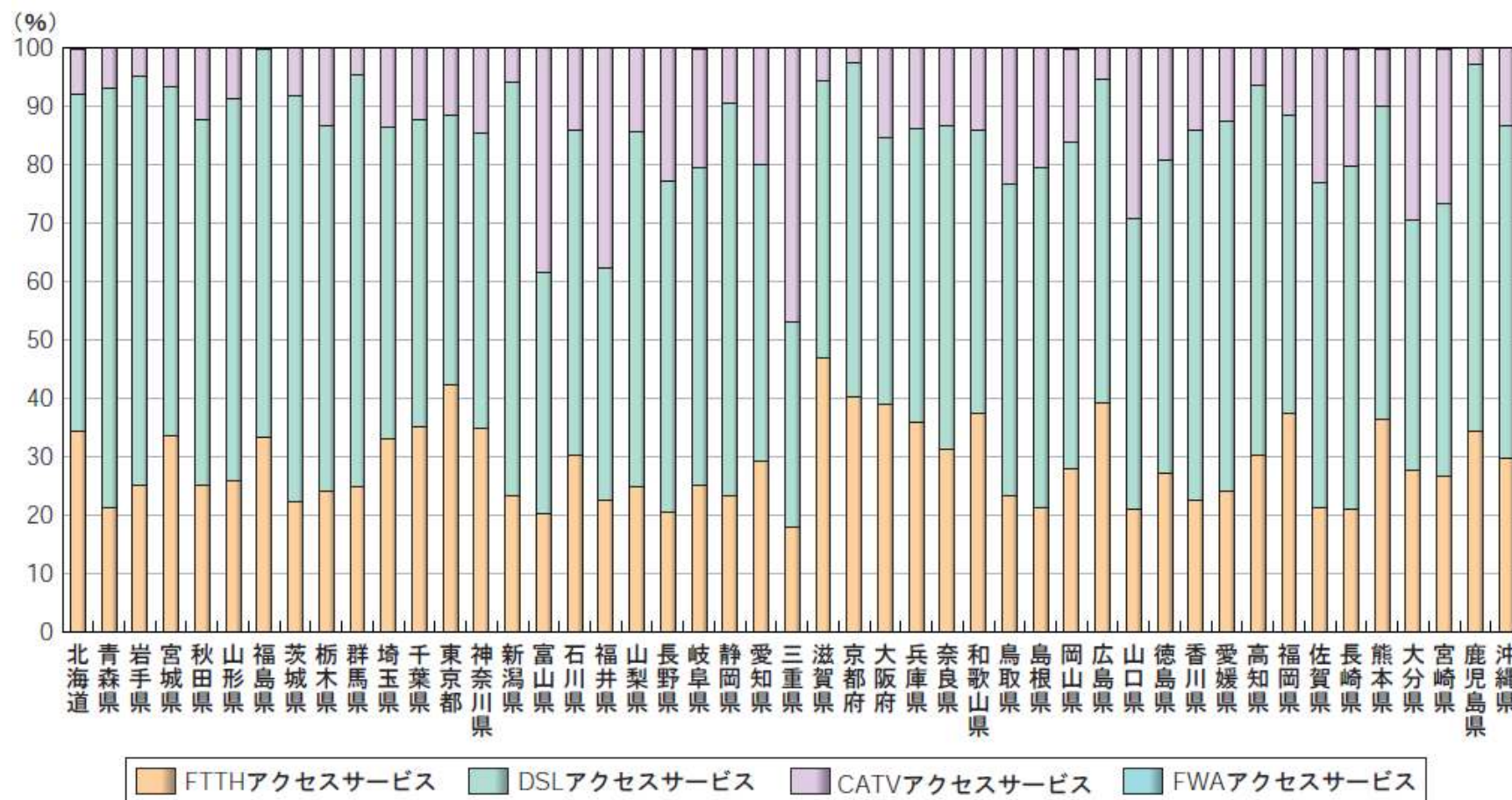
(出典)総務省「平成18年通信利用動向調査(世帯編)」

ブロードバンド契約者数の推移

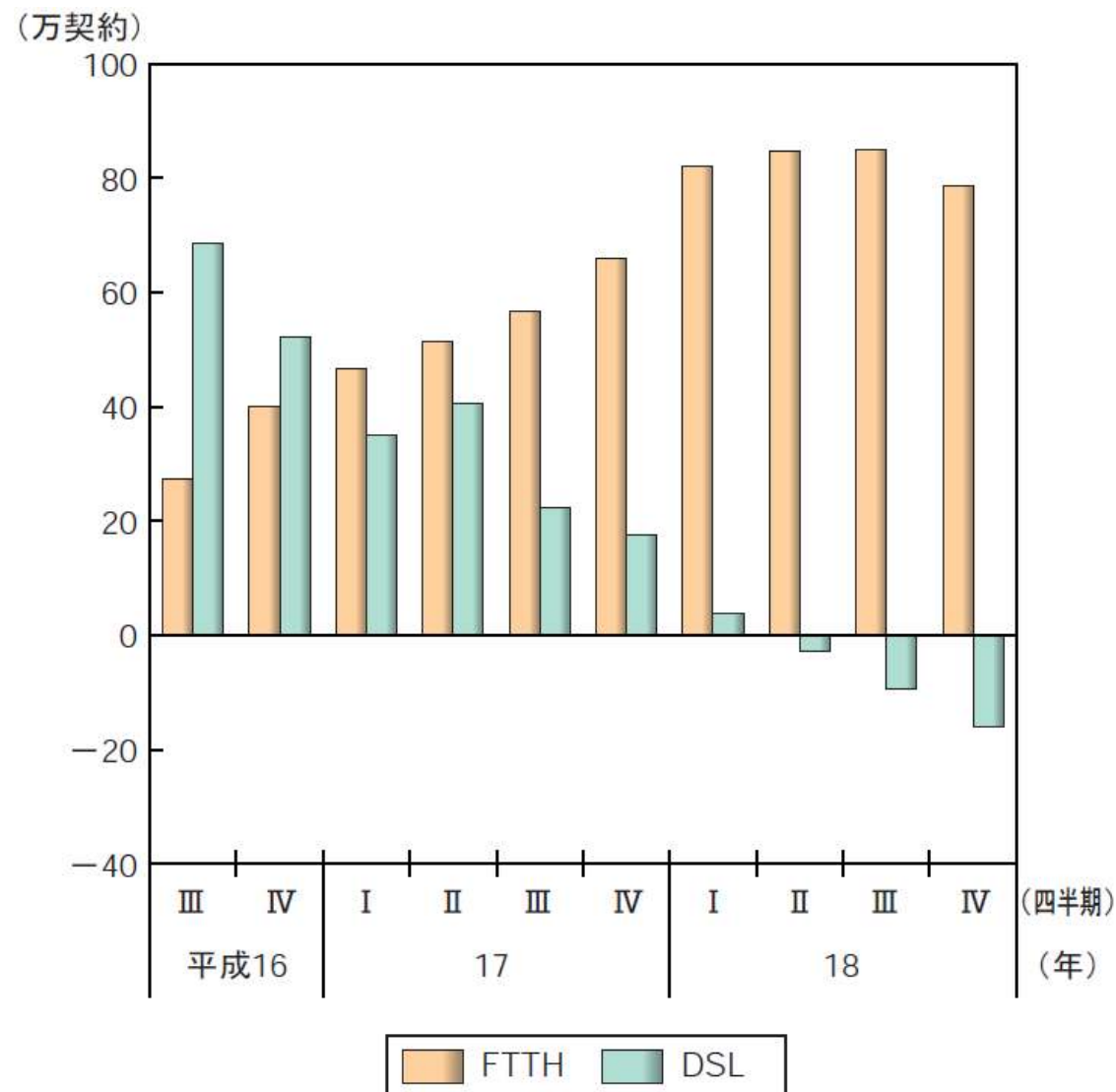


※ 平成16年度分以降は電気通信事業報告規則の規定により報告を受けた契約数を、それ以前は事業者から任意に報告を受けた契約数を集計

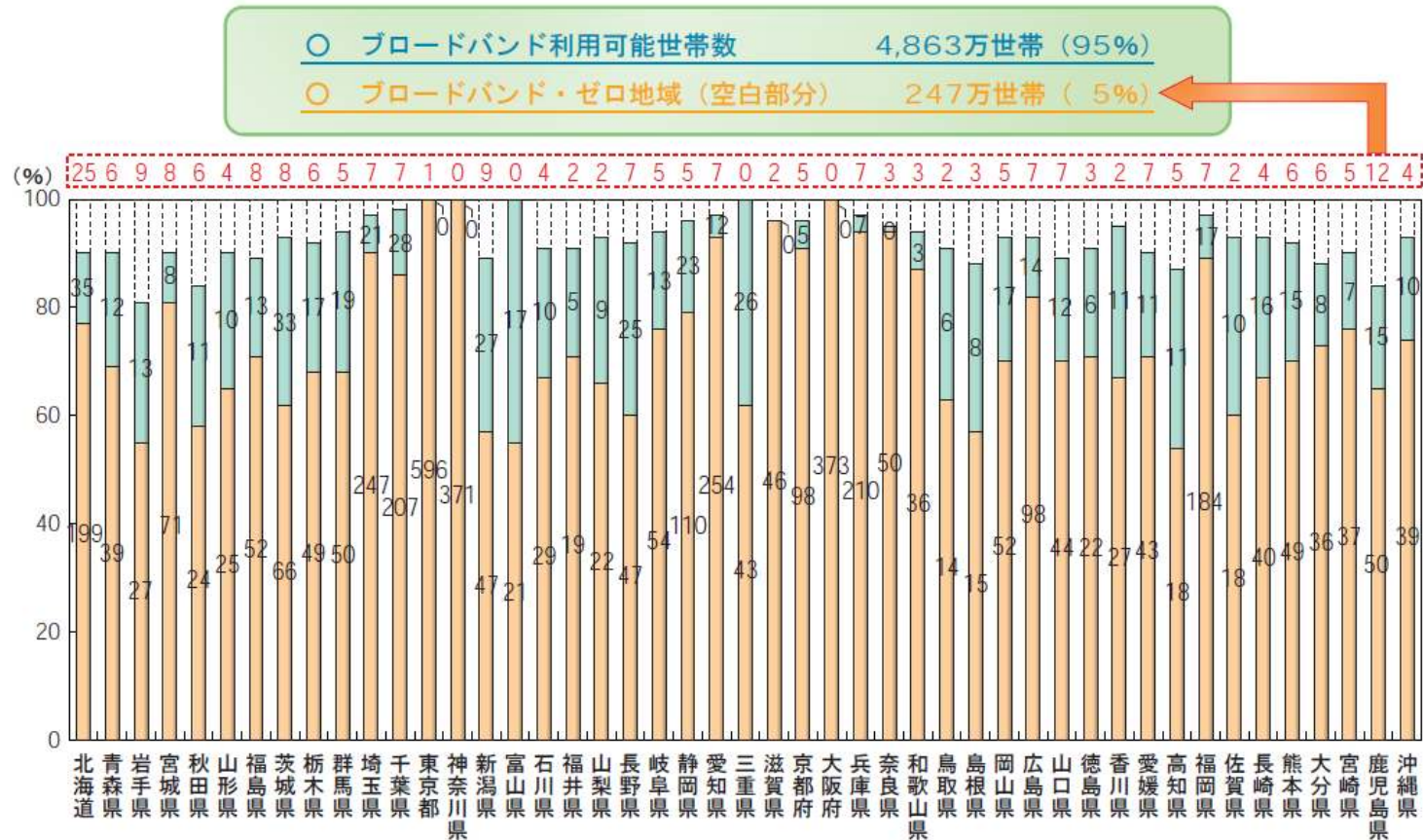
ブロードバンド契約の回線シェア



DSLとFTTHの契約純増数の推移



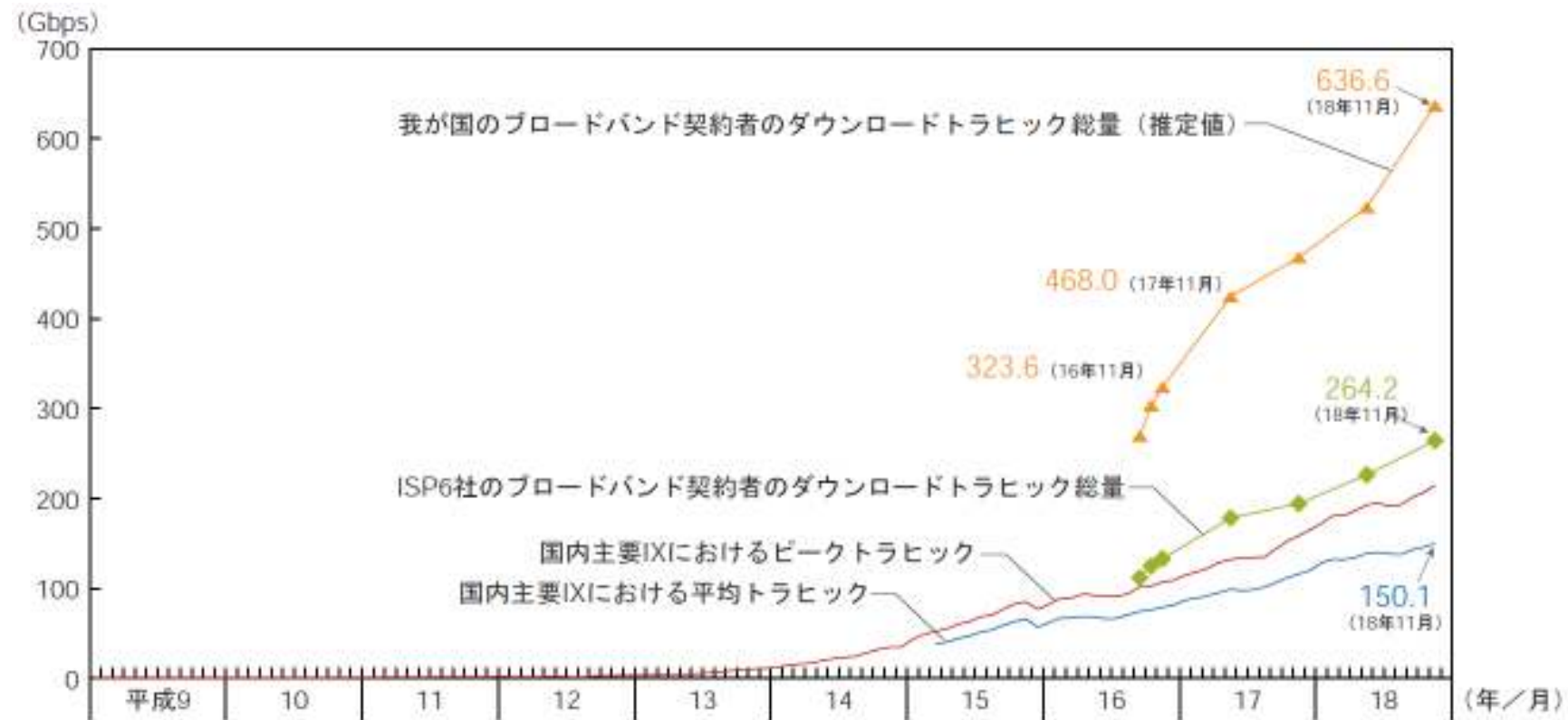
光ファイバーが全世帯をカバーする日



FTTHサービス (光ファイバ) が提供されている地域の世帯 (%)
 FTTHサービスは未提供だが、ADSL、ケーブルインターネット等の何らかのブロードバンドサービスが提供されている地域の世帯 (%)
 ブロードバンド・ゼロ地域の世帯 (%)

※ 事業者情報、国勢調査データ等に基づき推計
 なお、ADSLについては、サービスの提供地域内であっても、収容局からの距離が4 kmを超える世帯については信号の減衰が大きく実用に適しないことから、「未提供」に含めてある
 ※ グラフ中の数字の単位は万世帯

国内トラフィック推移



国内のインターネット接続性

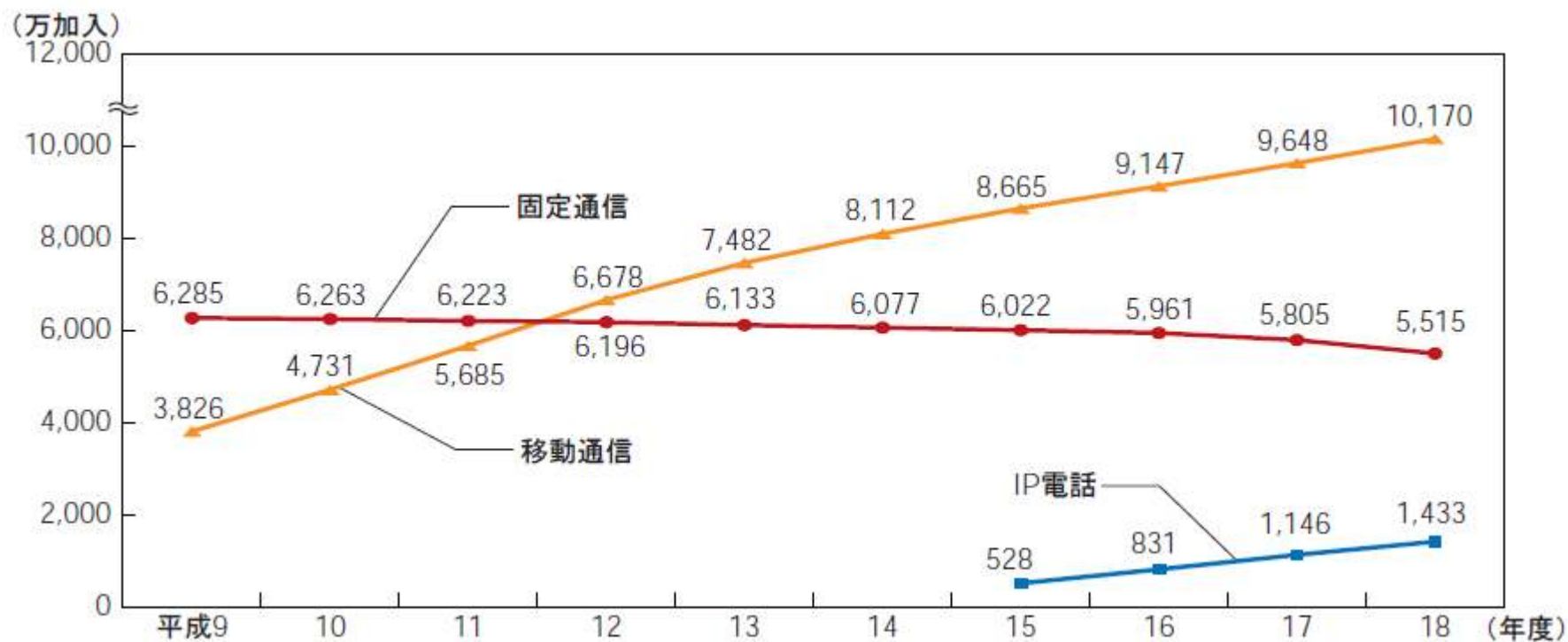
- 人口普及率70%
- ブロードバンド及びFTTHが全世帯へ
- 今後帯域は増えるだろうが、接続性の懸念は薄らいだ
- われわれは勝利しつつあるのか？
- 運用技術の積極的な公開

次の10年

- インターネットのさらなる拡大
- 国内インターネット
- IP電話
- 専用線
- 網内経路制御問題
- 顧客トレンド
- 国内の電話の現状
- 非インターネットIP動画配信
- Ethernet高速化
- パケットネットワークによる電話網の構築

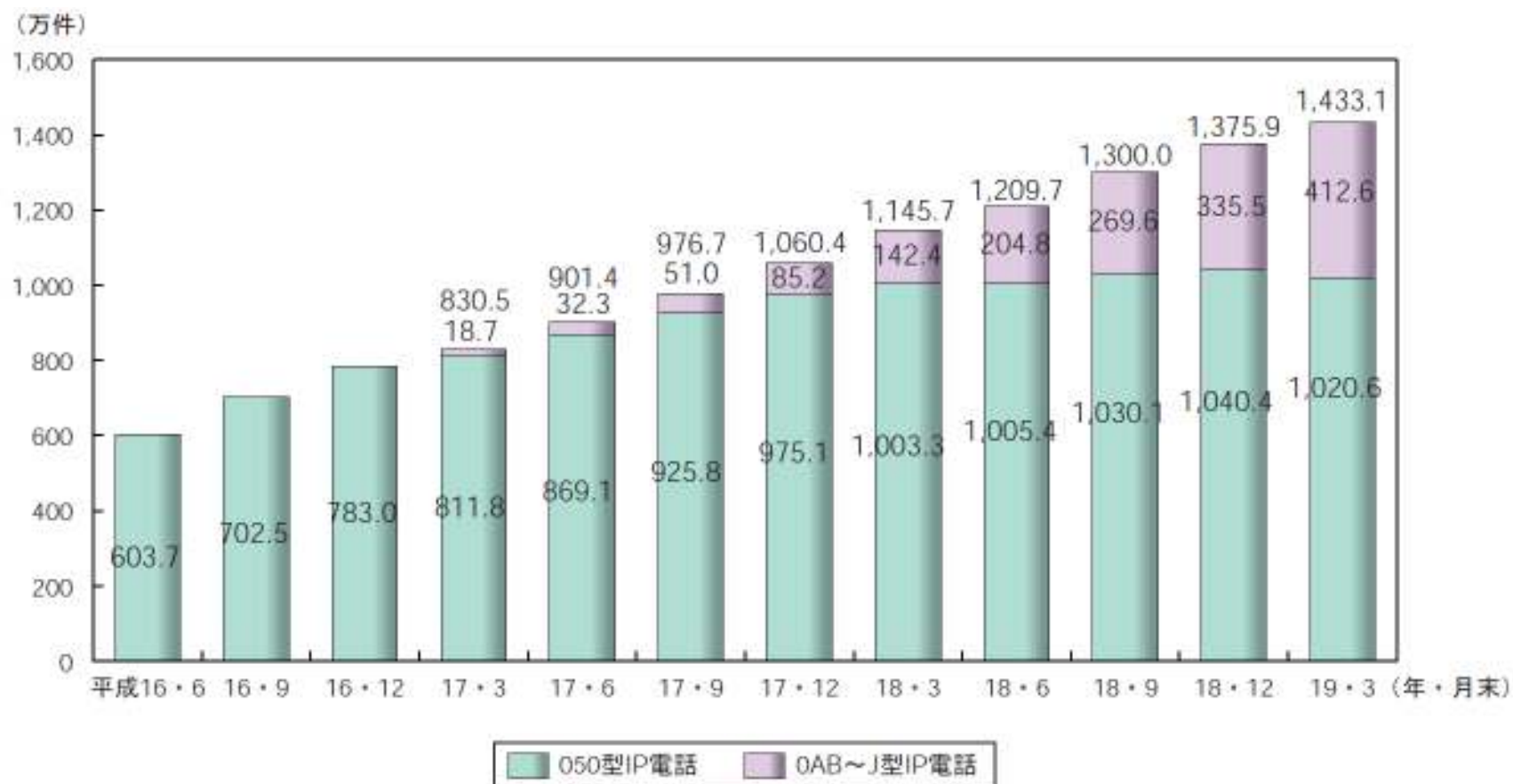
IP電話

- IP電話が急増(しかも固定電話の置き換え開始)
- インターネット網と同一である必要はないが。。。。。

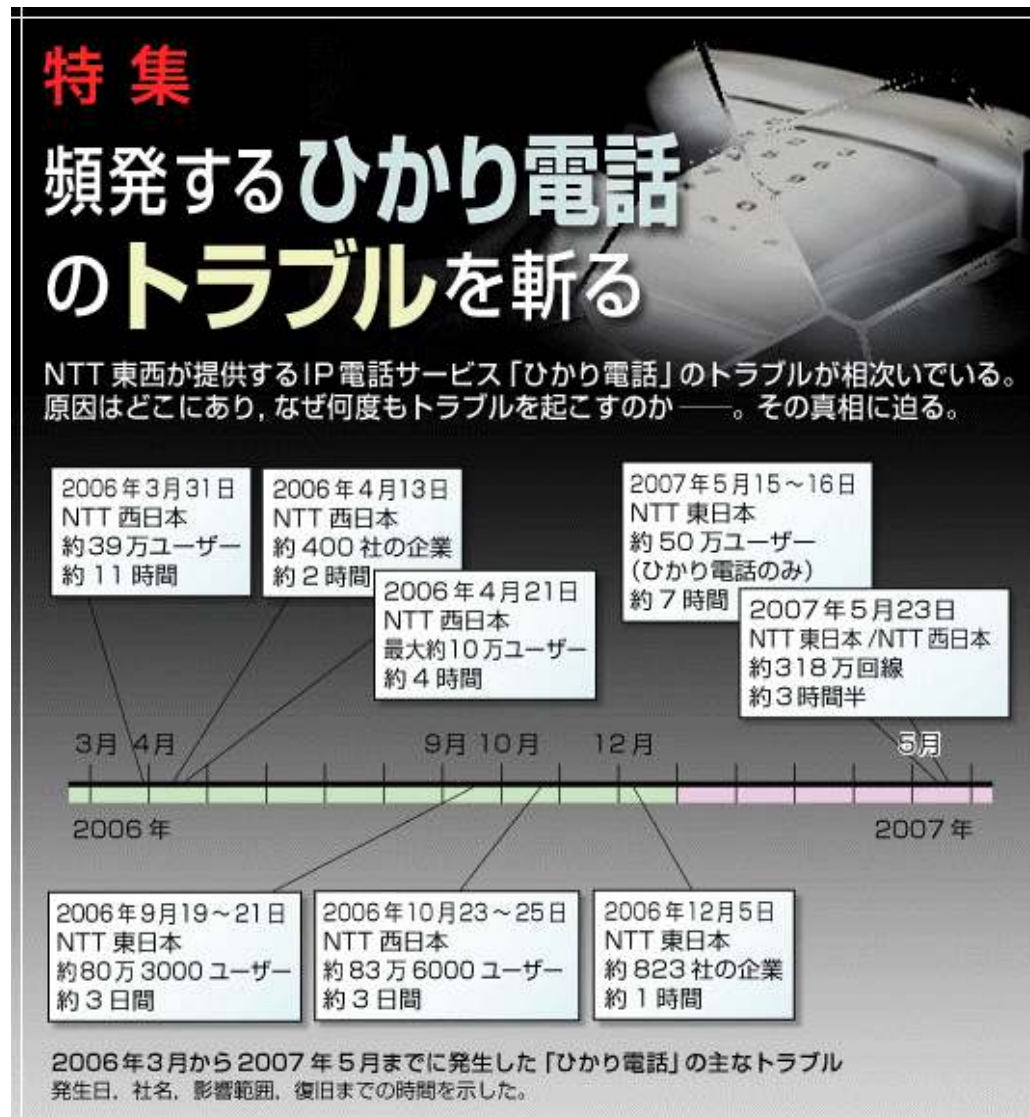


固定電話からIP電話へのシフト

○ IP電話の利用状況



ひかり電話の“電話”という言葉がよくないのでは？



サービスの質の違い？

- インターネットアプリ
 - つながる保証はない
 - TCP, UDP レベルでの再送の仕組み(切れても誤魔化せる)
- 電話
 - つながってあたりまえ
 - 緊急電話(110番119番)
- この間のギャップをどう埋めるか？
 - 現在のIP基盤技術で本当に電話を引き受けられるか？
- 安かろう悪かろう電話の導入？
 - 安い110番というのはあり得るか？

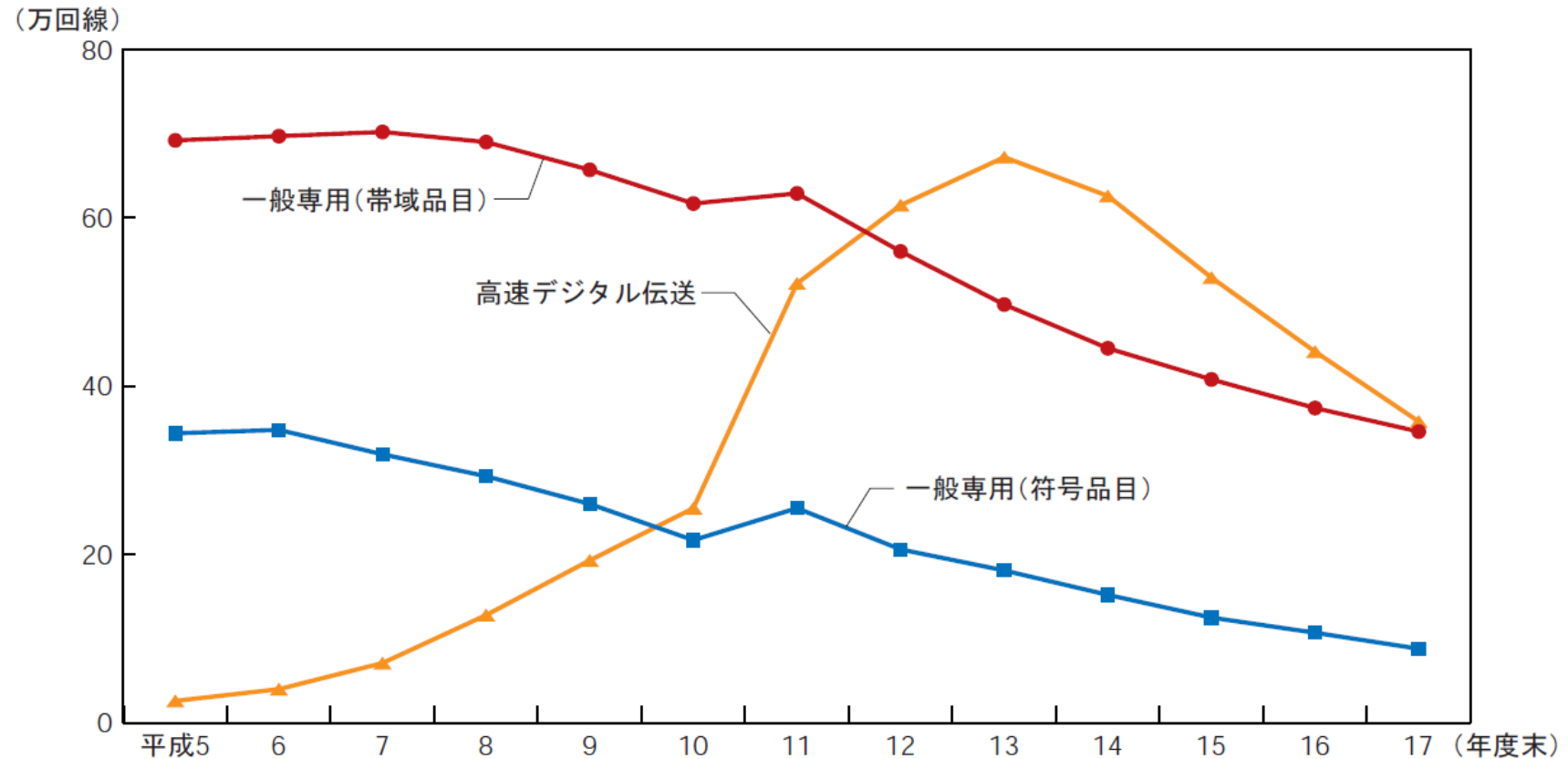
次の10年

- インターネットのさらなる拡大
- 国内インターネット
- IP電話
- 専用線
- 網内経路制御問題
- 顧客トレンド
- 国内の電話の現状
- 非インターネットIP動画配信
- Ethernet高速化
- パケットネットワークによる電話網の構築

専用線

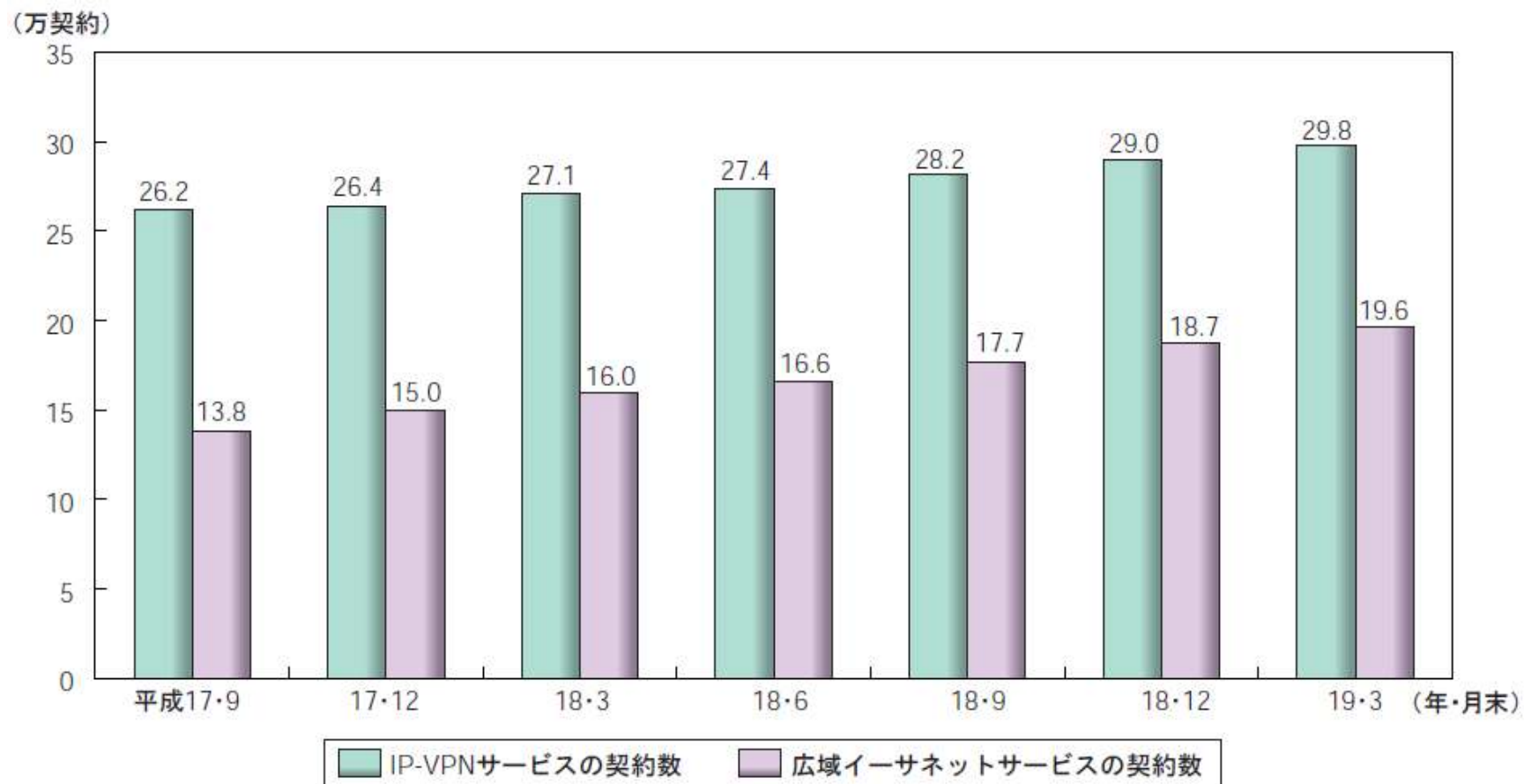
- 国内の専用線は79万2千回線
- 1秒以上の回線断が障害の本物の専用線
- 安い専用線から本物の専用線までさまざまなバリエーション
 - IP-VPN
 - 広域Ethernet
 - ATM

国内(本物)専用回線数の推移



IP-VPNおよび広域Ethernet

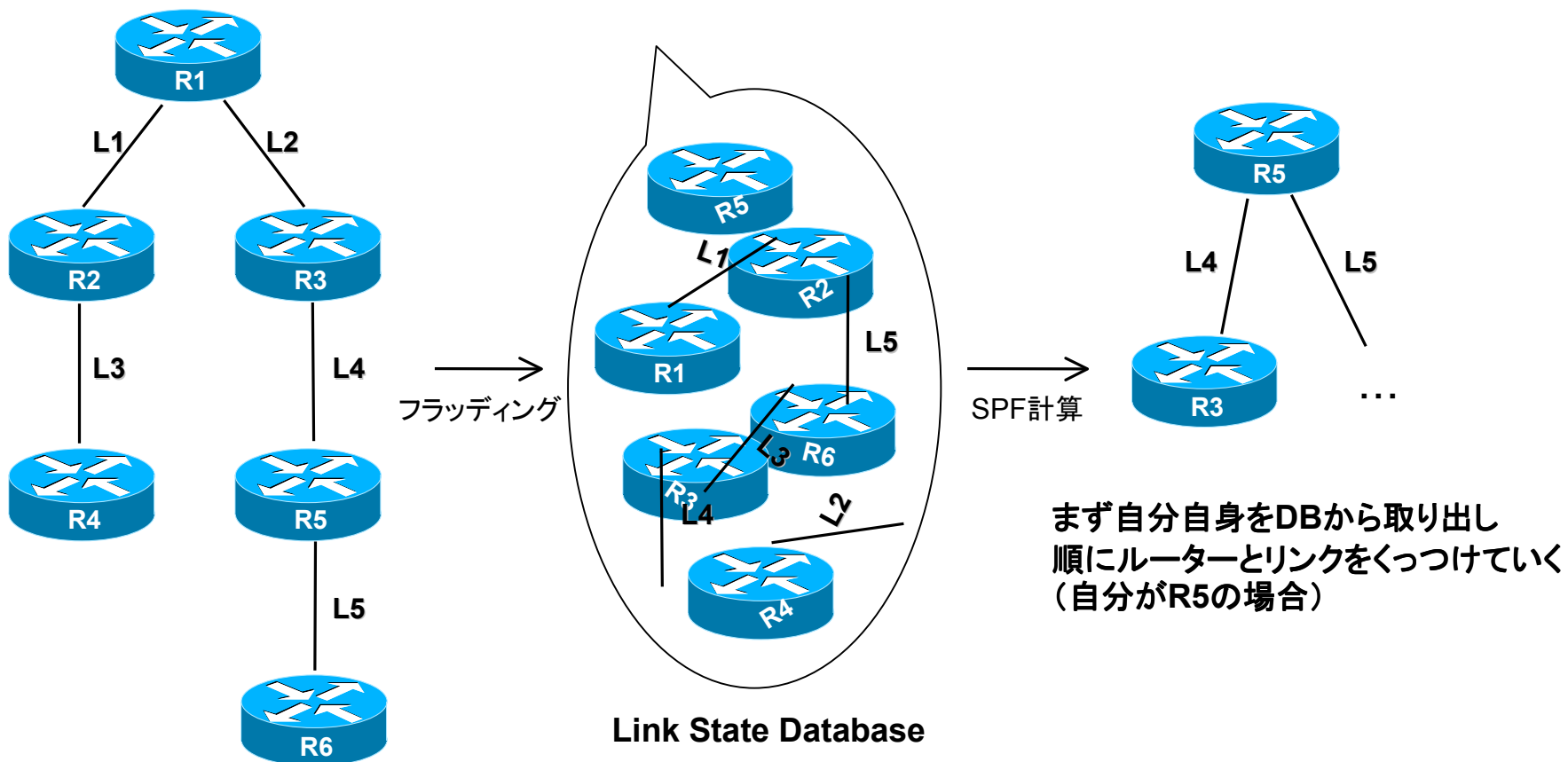
○ 安い専用線という概念



次の10年

- インターネットのさらなる拡大
- 国内インターネット
- IP電話
- 専用線
- 網内経路制御問題
- 顧客トレンド
- 国内の電話の現状
- 非インターネットIP動画配信
- Ethernet高速化
- パケットネットワークによる電話網の構築

AS内経路制御 = IGP = OSPF/IS-IS = SPF計算

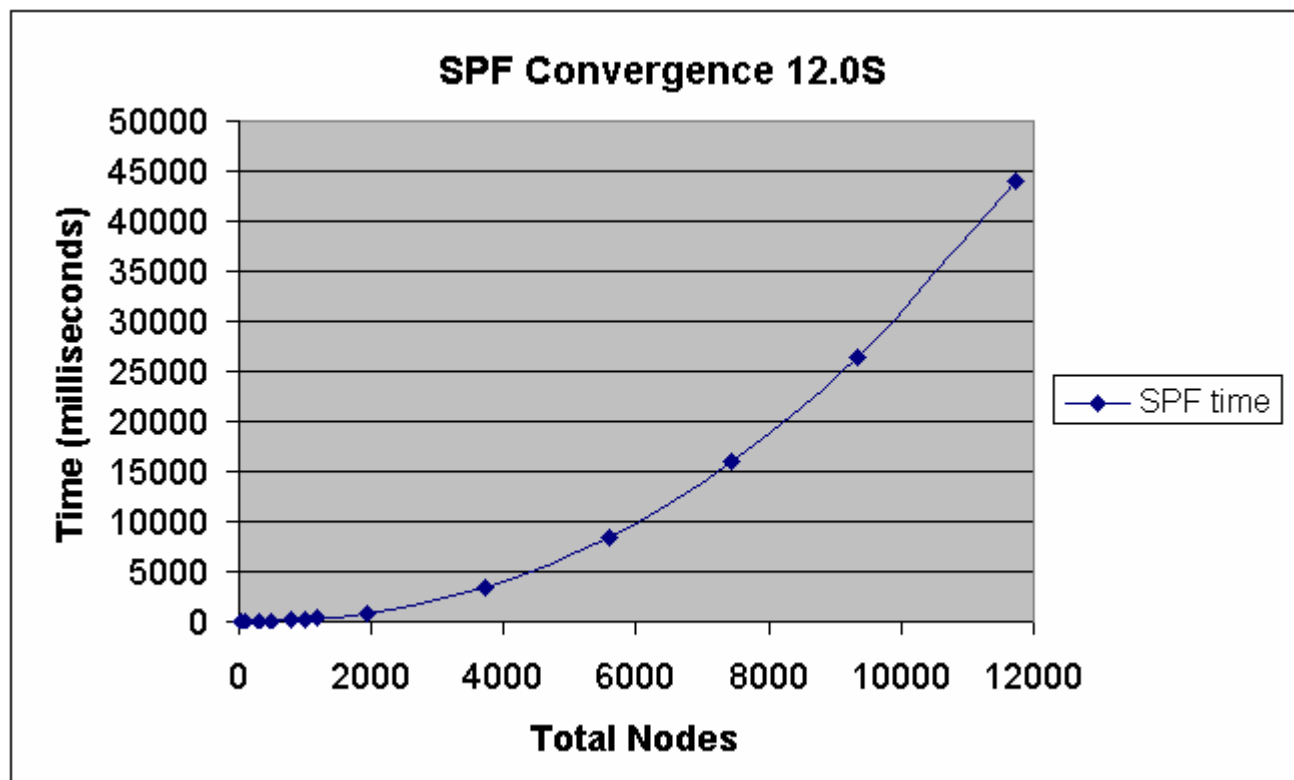


IGPの問題点

- あまりに便利のため、一度使うとやめられない
- 根拠はないが、何となくネットワークが大きくなってもそのまま動く気がする
- 破綻するぎりぎり寸前まで動くので死ぬ時は一気にいく
- SPF以外に接続性を担保する仕組みが無い場合がほとんど
- MPLSなど他レイヤーのプロトコルが依存している
 - SPFが死ぬとまきぞえで死亡
- 実際は数秒の断が継続的に発生していてもTCP,UDPレベルでカバーされるため傷が目立たない

IGPは無限にスケールするわけではない

- ネットワークが大きくなれば当前計算量も計算時間も大変
- ルーター数10000を超すとさすがにやばい
- CPU、メモリーの浪費と地球環境への影響



SPF全計算回避策

○ Incremental SPF Calculation

- ルーターおよびリンクのアップ&ダウン時に全部計算しない
- スタブへの変更→スタブ部分だけ再計算
- リンクのコストのみの変更(トポロジーに変化なし)

○ Ball&Chain

- 物理学の応用による手法
- リンクの切断時に、重力によって下にぶら下がるモデルを使用して計算量を減少させる

○ いずれにせよ、ラッキーな場合のみ適用可能

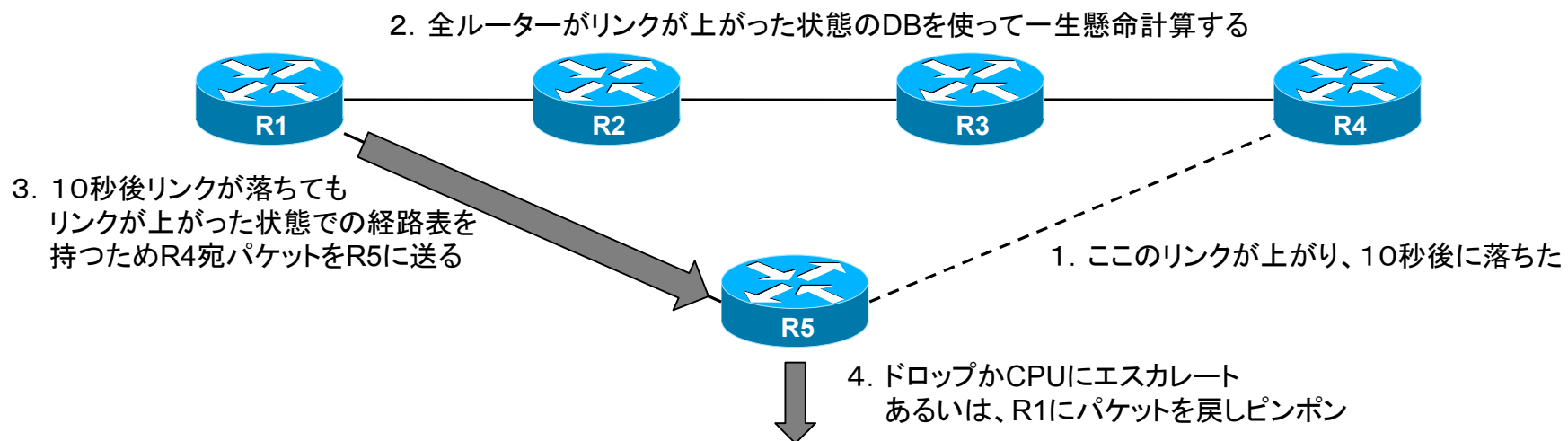
- SPF全計算が発生することを前提にすべき

SPF Exponential Backoff

- のべつまくなし計算していると、IGPメルトダウンの危険性が高まる
- 収束は遅くなる可能性があるが、SPFの開始のタイミングや SPF計算の間隔を、SPF計算の頻度により自動的に調節
- SPF Exponential Backoff の導入を考えた時点で、網の状況は危険だと考えるべき

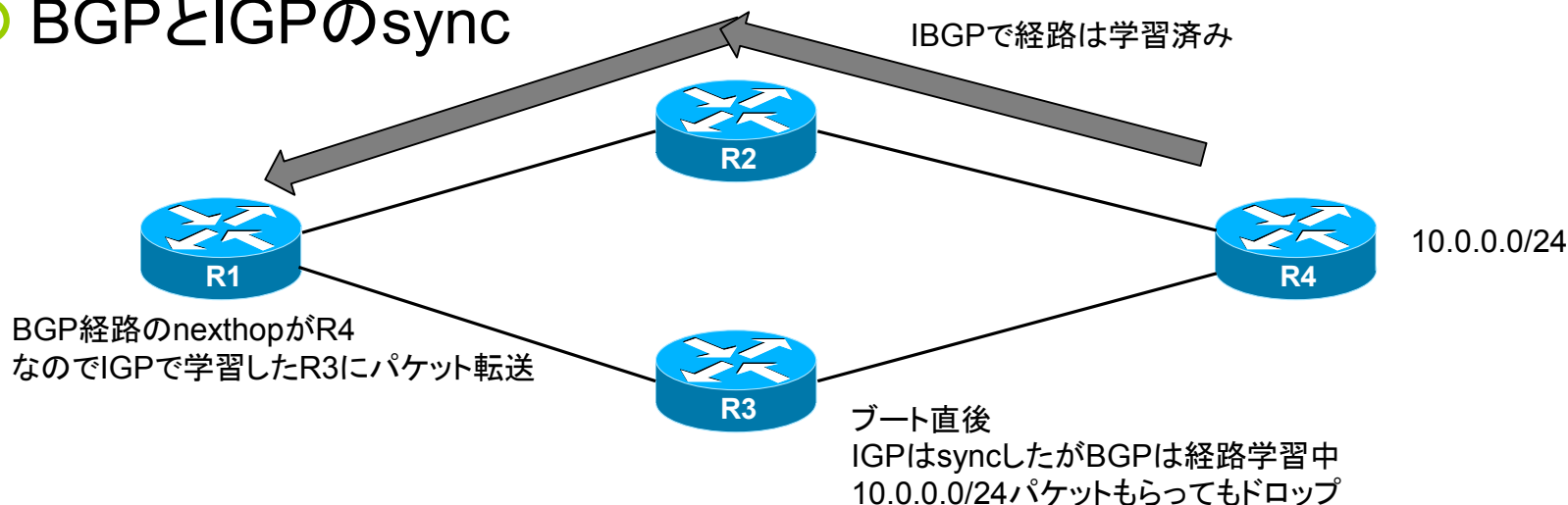
SPF Blackhole 問題

- SPFに時間がかかると、Blackholeの問題が発生する
 - SPF計算時間+SPFインターバル時間
- SPF Exponential Backoffが有効になっていると、さらに状況は悪くなる(最新のDBを使った計算までのタイムラグ)
- 存在しないリンクを使用したMPLS LSPが長時間存在する可能性も



プロトコル間の整合性問題(経路sync、タイムアウト)

○ BGPとIGPのsync



- 各プロトコル間でのタイムアウト時間の違い
- どんな場合でもCPUへのエスカレーションは止めたい
- 局所的には同一ベンダー同一バージョンで、
- 大局的にはマルチベンダ、マルチバージョンで

次の10年

- インターネットのさらなる拡大
- 国内インターネット
- IP電話
- 専用線
- 網内経路制御問題
- 顧客トレンド
- 国内の電話の現状
- 非インターネットIP動画配信
- Ethernet高速化
- パケットネットワークによる電話網の構築

最近のトレンド

○ Metro/Carrier Ethernetの引き合い急増

- Provider Backbone Bridge
- Provider Backbone Transport
- Ethernet OAM

○ Metro/Carrier Ethernet + MPLS

- PBT + PWE3
- PBT as IGP shortcut
- PBT as MPLS VPN BGP nexthop

○ Carrier Ethernet Transport

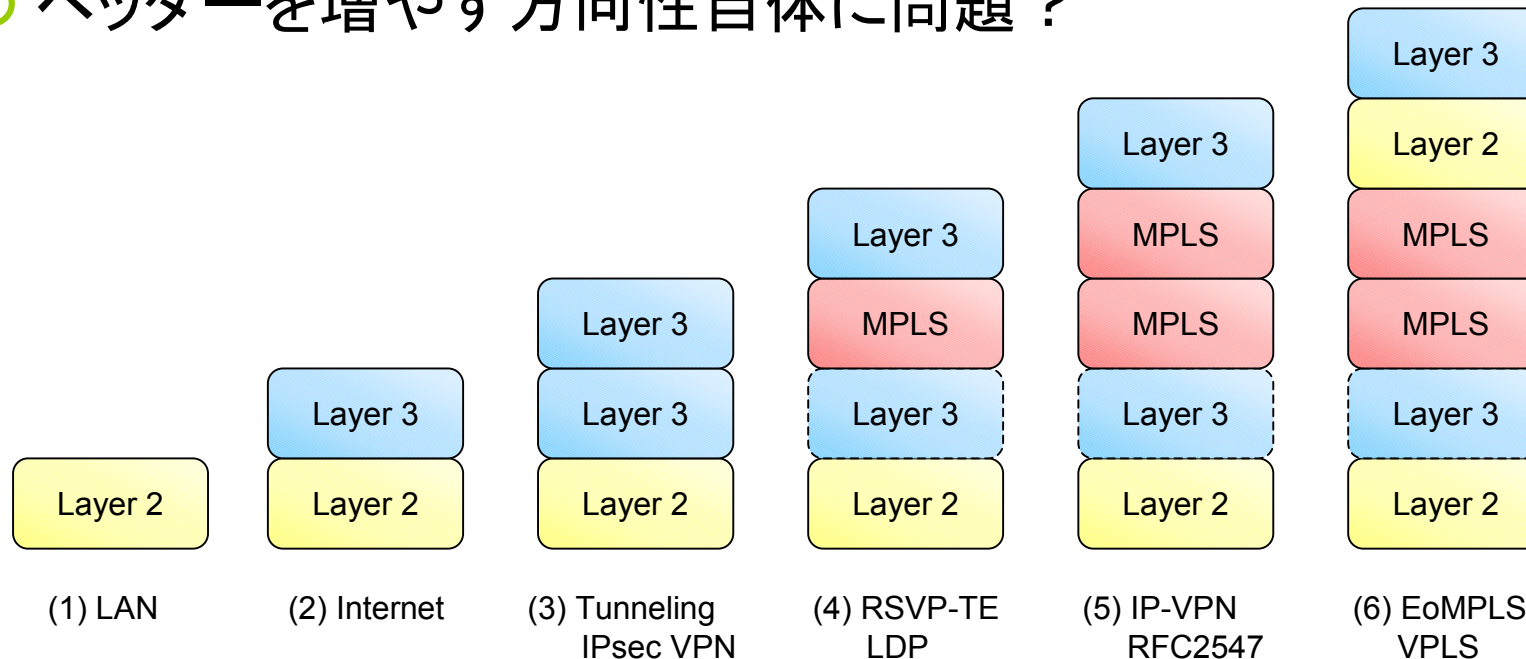
- T-MPLS
- Full Provisioning Network Management Systemへの要望

○ 電話網を引き受けるための技術整備

- NGN

History of Internet Routing

- これまでのInternet Routingの歴史
- ヘッダーが増えるごとに網の安定性は遡減する
- L3 IGP/EGPにスケーラビリティ問題
- ヘッダーを増やす方向性自体に問題？

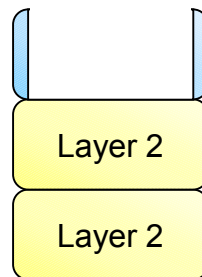


MPLSではパケット内にL3ヘッダーが現れないが、L3レイヤーでMPLSシグナリングを行うため、仮想的なヘッダとして表した

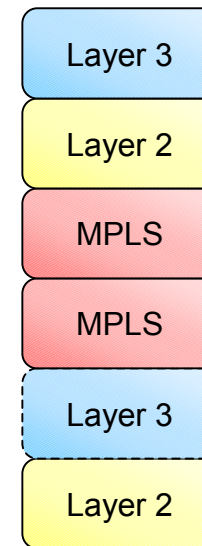
シグナリング/ルーティングを網から抜きたい

- すべては電話のために
- PBTあるいは、T-MPLSをつかって、固定設定で運用したい
- うちの商売あがったりか？

エッジでL3からL2へマッピング →



PBT



EoMPLS
VPLS

とにかくL3ノードを減らしたい

- すべては電話のために
- できるだけIGPのSPFドメインに参加するL3ノードの数を減らしたい
- 現状MPLSを動かすには、L3ノードであることが前提
- QoS(あるいはCoS)保障は絶対に必要
- すべてのノードが管理下にある時に自律分散が本当に必要か？

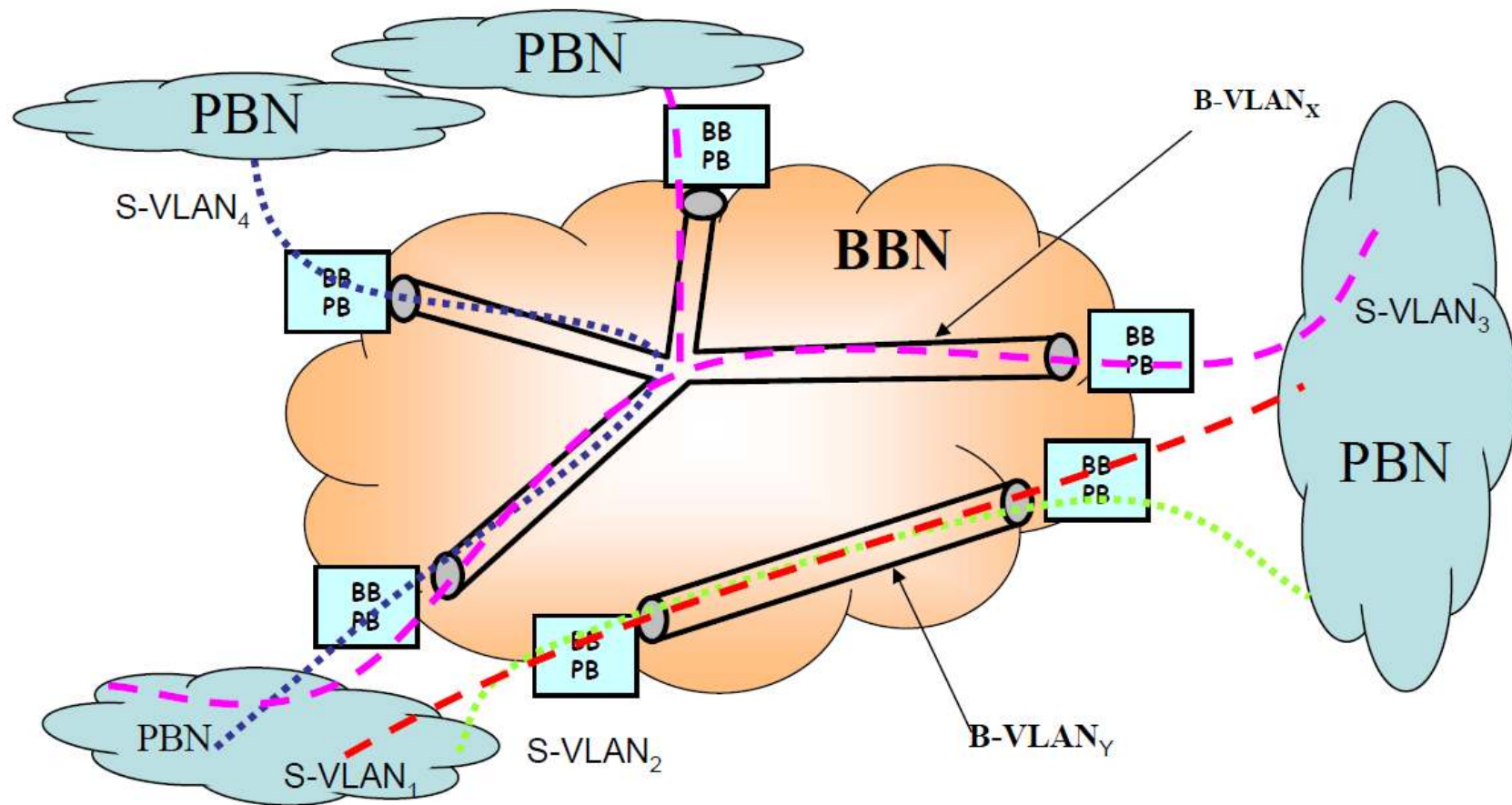
IP代替技術としてのEthernet技術の進化

- IPおよびMPLSの多様なサービスの受け皿がEthernetで可能になりつつある

	IP	Ethernet
VPN サービス	IP Tunneing IP-VPN	VLAN stacking(802.1ad) PBB(802.1ah)
Pseudo Wire	EoMPLS	PBT(802.1Qay)
OAM	Ping, Traceroute	Ethernet OAM ITU Y.1731

Provider Backbone Bridge

○ スケールするVLAN stacking

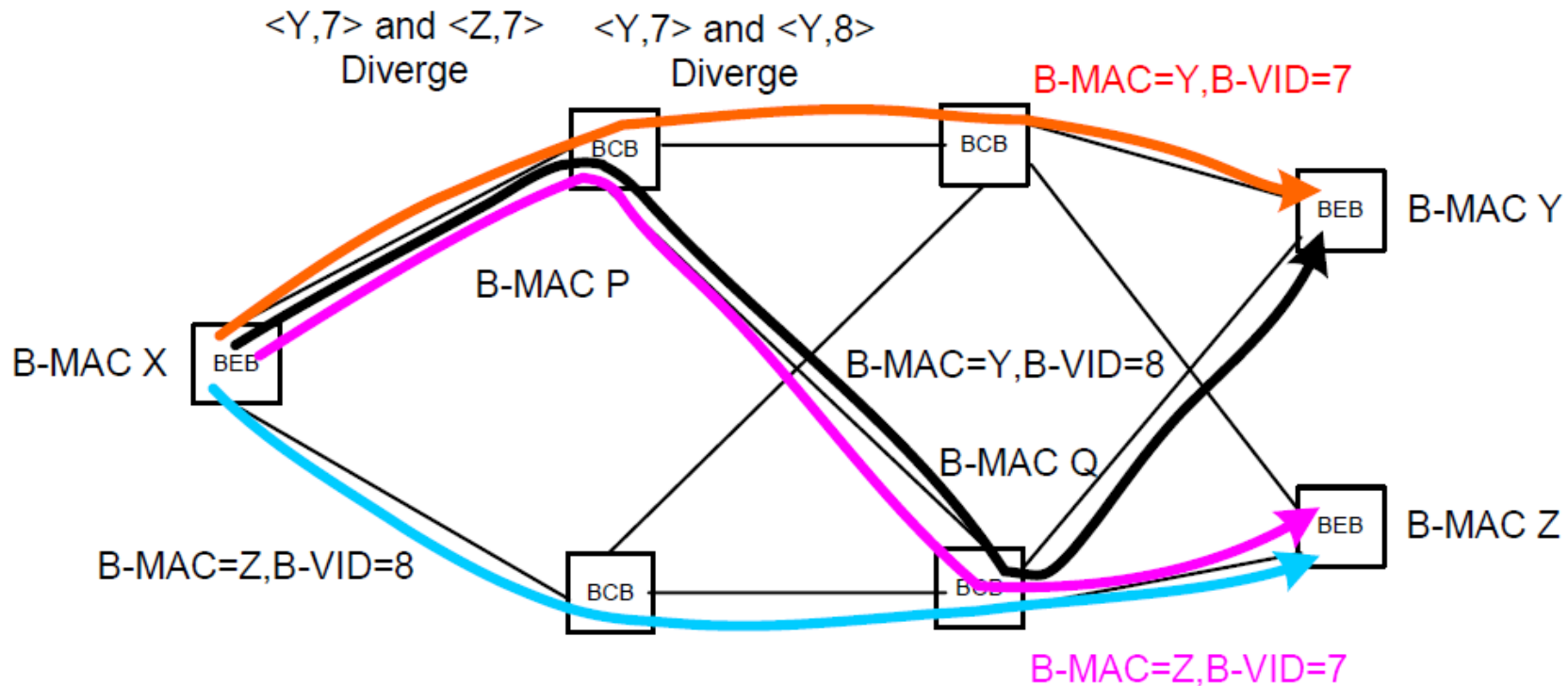


• **BB PB**: Backbone Provider Bridge Edge

Provider Backbone Transport

○ PBT(Provider Backbone Transport)

- PBB上にQoS保障をした、Point-to-Pointの経路を構築する技術
- EthernetによるPWE3技術



Carrier Ethernet Transport

○ CET(Carrier Ethernet Transport)

- L3のスケラビリティ問題と、Ethernet技術の進化を背景に、キャリアのIPバックボーンをL2 Ethernetで構築しようとする動きの総称
- Wikipediaにも載ったので一般用語化しつつあるのか？
- シグナリングは以下の2つのどちらか
 - PBT
 - T-MPLS

○ British Telecomの動きによって加速した

○ シグナリングがないため、手による設定で網を構築

○ すぐれたマネージメントソフトウェアが必要

○ すぐれたOAMシステムが必要

○ 広域Ethernetの運用経験資産を生かす

○ BPTとPWE3との類似性

BT NGN のケース

○ BT NGN案件、L3+MPLS(Cisco)とL2 PBT(Nortel)の対決

- 以下LightReading(<http://www.lightreading.com/>)の記事タイトル
- 2006年5月 BT Likes Nortel's New Ethernet Flavor
- 2006年9月 BT Rethinks 21CN Core Strategy
- 2007年1月 Nortel, Siemens Win PBT Deals at BT
- 2007年4月 BT Pressures Vendors Over PBT
- BT NGNはL2 PBTでの構築が決定

○ China Telecom

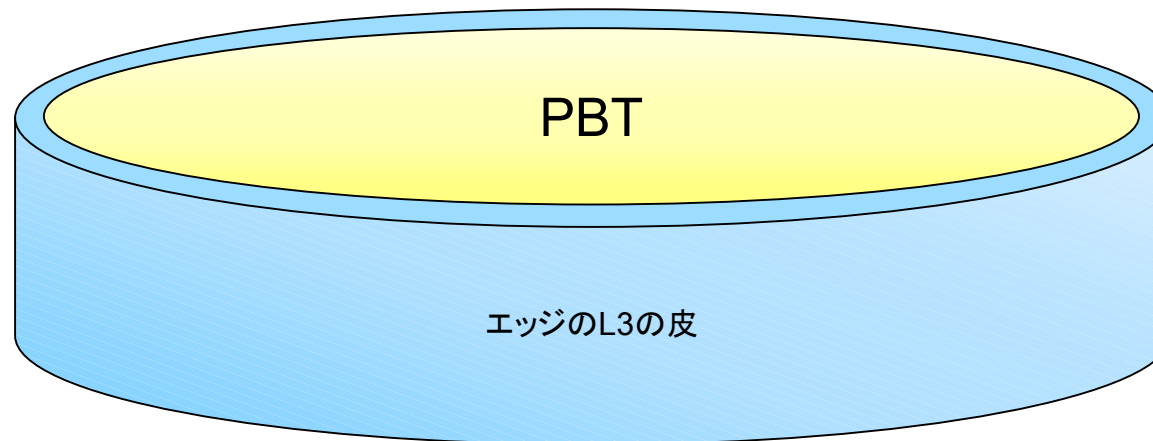
- 2006年10月 Nortel's PBT Debuts in China

○ 他ベンダーの動き

- 2006年12月 Cisco Tracks PBT Standards Process
- 2007年3月 Huawei Supports PBT
- 2007年4月 Avici Abandons Routing, Targets PBT

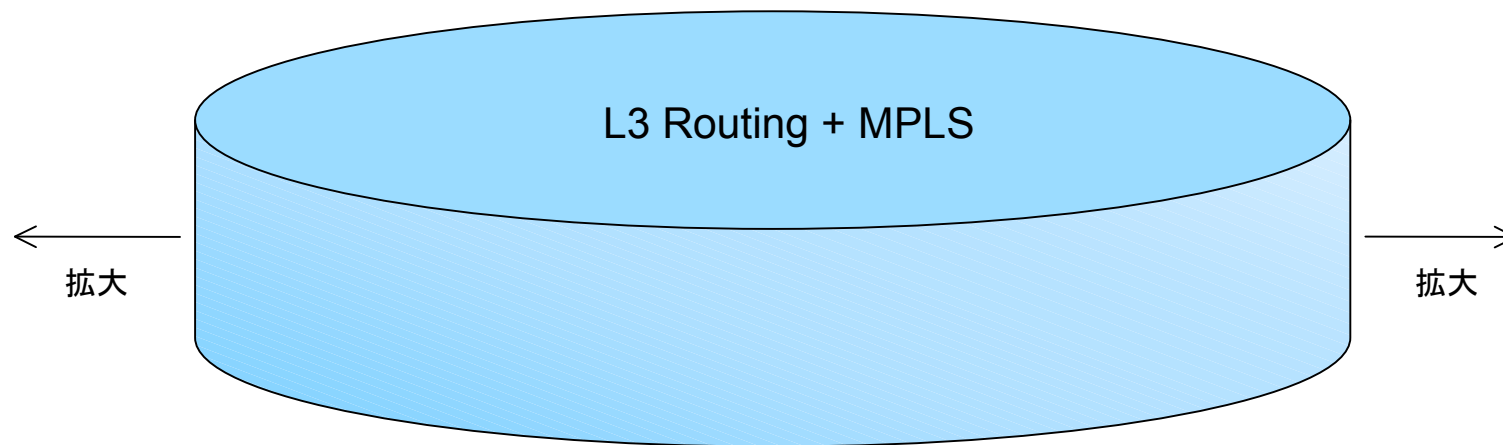
BT 21CNネットワークアーキテクチャ

- コアはPBTで構築（一部MPLS網も併用）
- IPサービスとのインターフェース部分のみL3で構築
- 新規網のため思い切ったアーキテクチャが可能だった



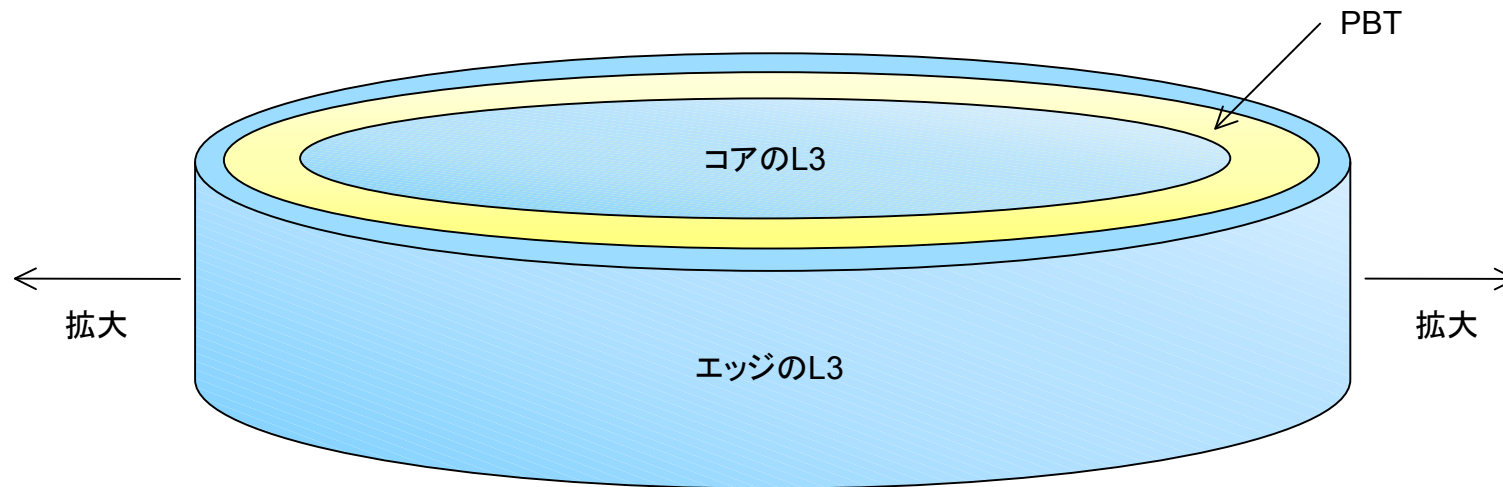
既存ISP/Carrier IP網アーキテクチャ

- 基本的にすべてのルーター間でIPルーティングを行っている
- 網の拡大と共にL3ネットワークも拡大せざるを得ない
- 深刻なスケーラビリティ問題に直面している



Carrier Scaling Ethernet

- 既存コア網を生かしつつ、L3ネットワークの拡張を防ぐ
- エッジで、L3->PBTへマップし、さらにPBTからコアL3ネットワークへ渡す
- ネットワークの拡大には、PBT網の拡張で対応する

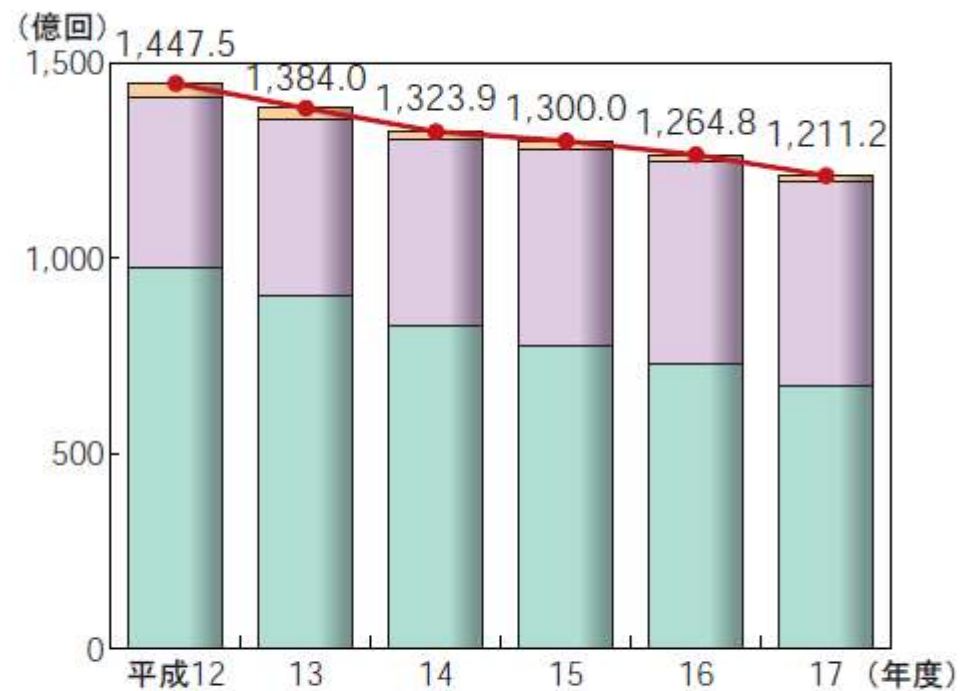


次の10年

- インターネットのさらなる拡大
- 国内インターネット
- IP電話
- 専用線
- 網内経路制御問題
- 顧客トレンド
- 国内の電話の現状
- 非インターネットIP動画配信
- Ethernet高速化
- パケットネットワークによる電話網の構築

電話を掛ける回数は減っている

図表2-1-18 通信回数の推移（発信端末別）

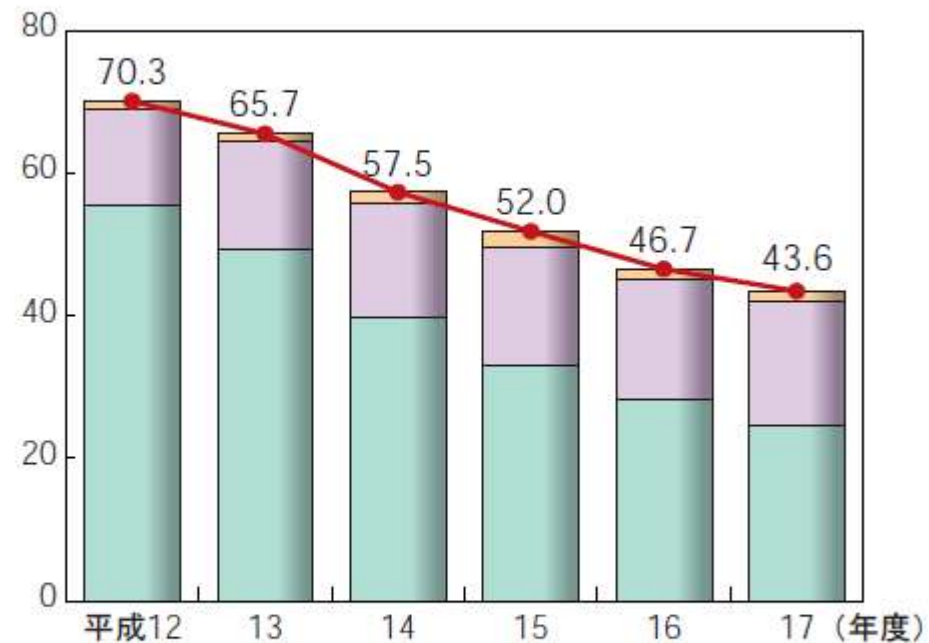


PHS	35.9	26.2	22.2	21.3	17.2	17.0
携帯電話	438.3	452.4	474.5	504.4	516.8	522.3
固定端末	973.3	905.4	827.2	774.4	730.7	671.9
総回数	1,447.5	1,384.0	1,323.9	1,300.0	1,264.8	1,211.2

電話で話す時間も減っている

図表2-1-19 通信時間の推移（発信端末別）

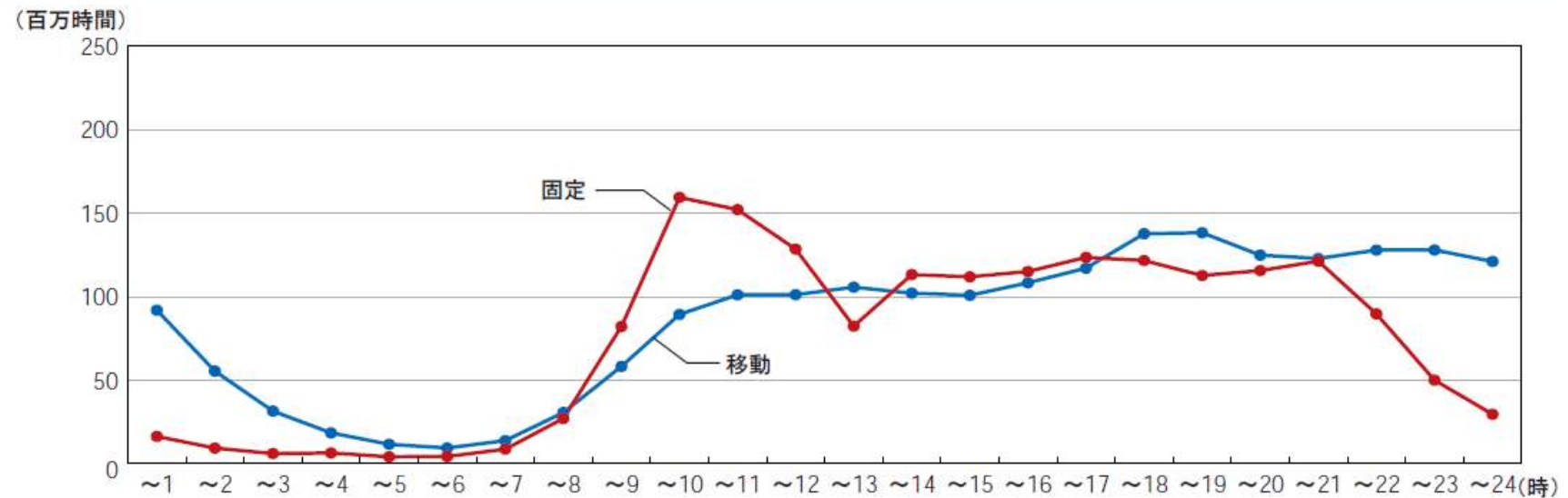
（億時間）



PHS	1.1	1.2	1.5	2.3	1.5	1.4
携帯電話	13.4	15.2	16.0	16.6	16.7	17.4
固定端末	55.7	49.4	40.0	33.2	28.5	24.8
総時間	70.3	65.7	57.5	52.0	46.7	43.6

固定は朝10時にピーク、移動は午後6時ごろ

図表2-1-24 固定通信と移動通信の時間帯別通信時間の比較



総務省「トラヒックからみた我が国の通信利用状況」により作成

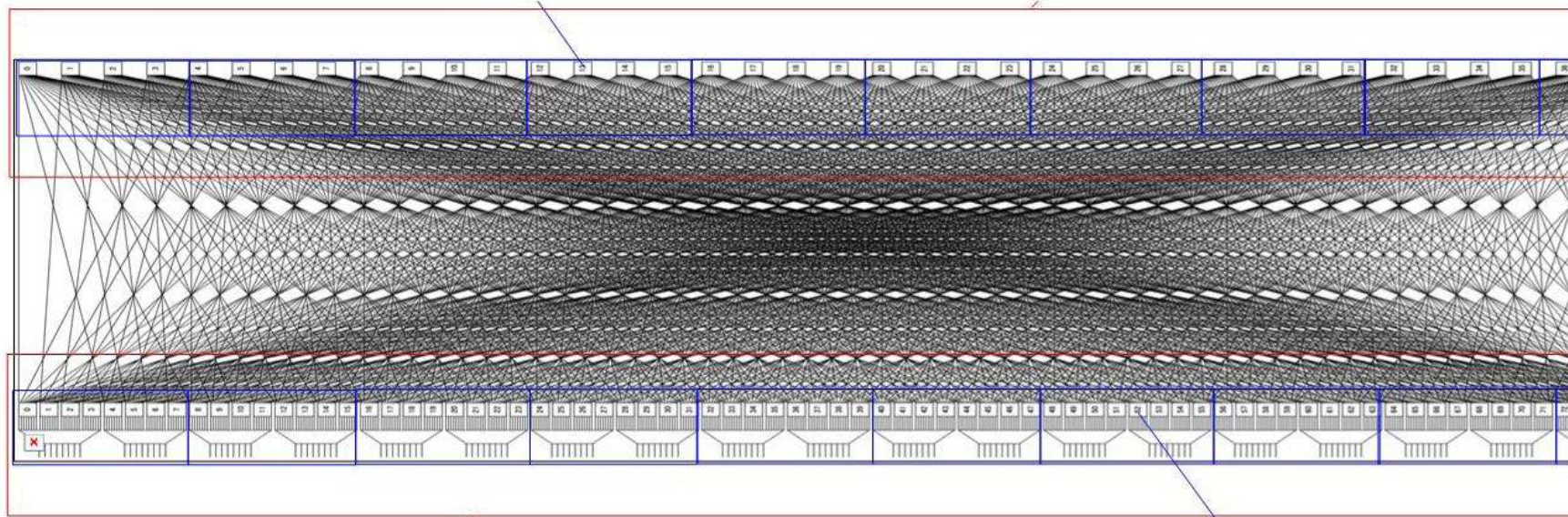
次の10年

- インターネットのさらなる拡大
- 国内インターネット
- IP電話
- 専用線
- 網内経路制御問題
- 顧客トレンド
- 国内の電話の現状
- 非インターネットIP動画配信
- Ethernet高速化
- パケットネットワークによる電話網の構築

非圧縮ハイビジョン配信 over IPネットワーク

○ 非圧縮ハイビジョン配信開始

- その筋の人によると一本2.5Gbps欲しいらしい
- そろそろ10G x 160portの放送用Ethernetスイッチが出そう
- そうすると、10G(4本) x 160 = 640本のハイビジョン非圧縮配信可能
- 日本程度の国土規模で、キー局から地方への配信には十分か
- 広域Ethernetで、スター型トポロジー



スーパーハイビジョン放送 over IPネットワーク

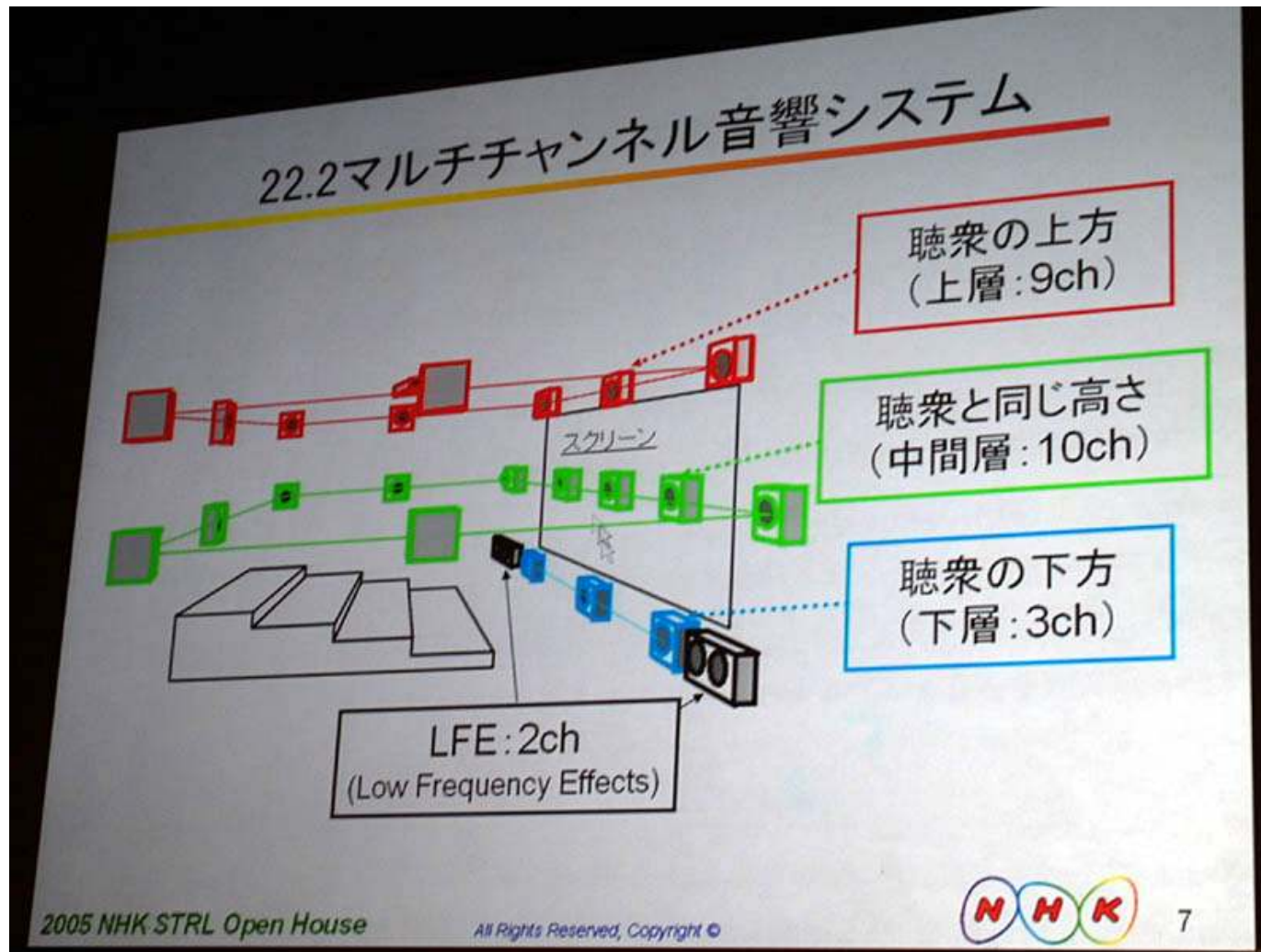
○ スーパーハイビジョン

- ハイビジョンの16倍の映像
- 7680x4320の解像度による画像
- 22.2チャンネルのサラウンド音声
- 非圧縮で50Gbps
- 3.5Tのディスクで18分録画できるそうです
- すいません、圧縮させてください
- 圧縮すると、200M-400Mbps(MPEG-4 AVC/H.264)

視野角100度！



22.2マルチチャンネル音声



衛星と言っているが、 数百MbpsならばIP網でいけるかもしれない

研究の背景と目的

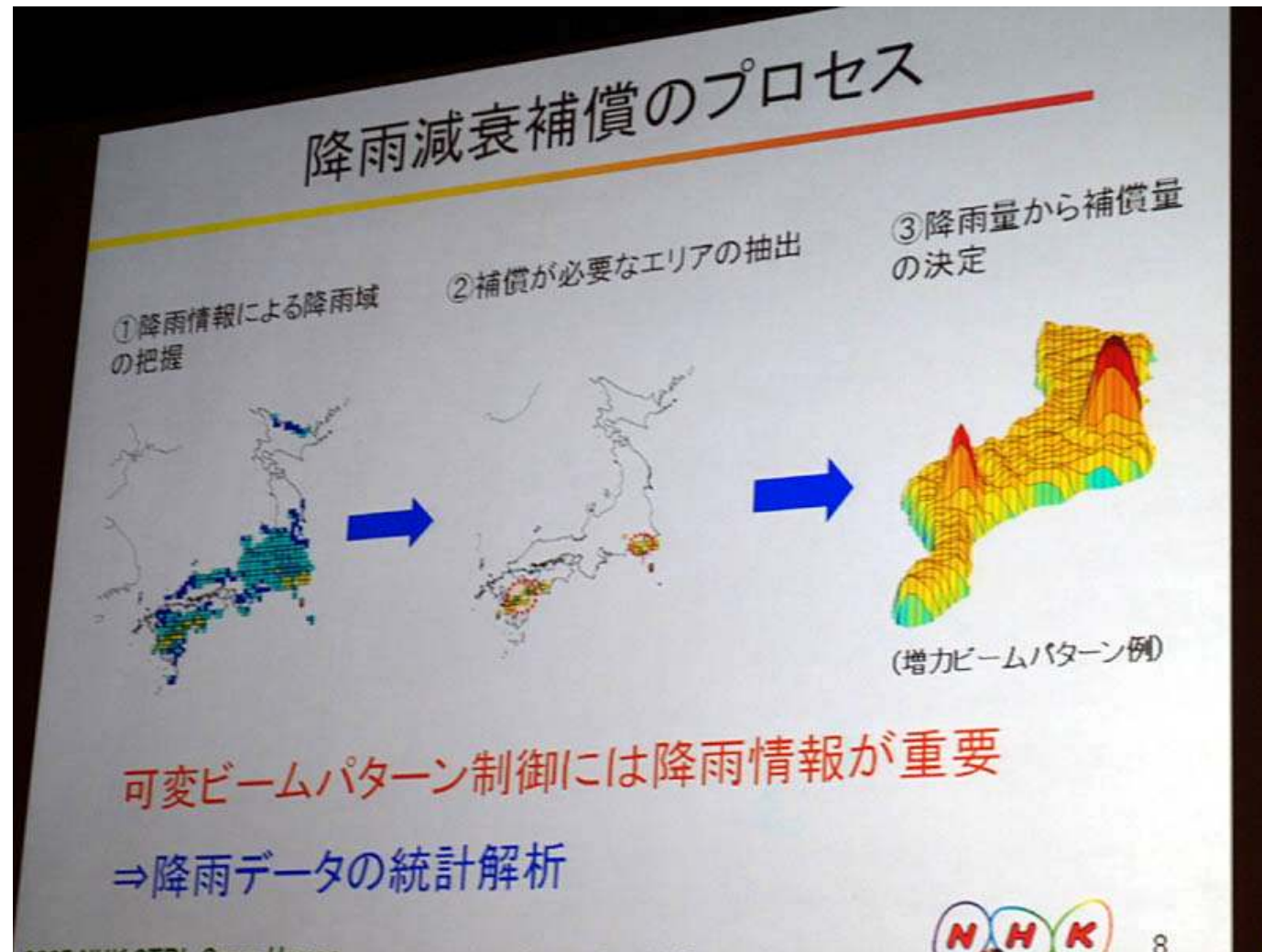
- 視聴者に圧倒的な高臨場感と没入感を与えるスーパーハイビジョン
- 走査線数は現行のハイビジョンの約4倍の4000本以上
- 非圧縮では40~50Gbpsの伝送速度が必要になるが、デジタル圧縮符号化技術により伝送速度は数百Mbps程度になると予想されている
- 現行の放送メディアでは家庭への伝送は困難



21GHz帯衛星放送バンド

- 広帯域放送チャンネルの確保できる
- スーパーハイビジョン放送の伝送路として期待できる

衛星の21Ghz帯衛星放送バンドは雨に弱い



次の10年

- インターネットのさらなる拡大
- 国内インターネット
- IP電話
- 専用線
- 網内経路制御問題
- 顧客トレンド
- 国内の電話の現状
- 非インターネットIP動画配信
- Ethernet高速化
- パケットネットワークによる電話網の構築

Ethernetのさらなる高速化

- 40Gは運用フェーズへ
- 100G Ethernet
 - [Road to 100G Alliance](#)
 - 100Gリンクアグリゲーション($x8 = 800G?$)
 - 2007年中に専用ASIC搭載機器
 - 2008年には出来合いのチップ出荷予定
 - PBB,PBT,OAMなど機能の入れ込み
- L3フォワーディング機能のチップ化
- MPLSフォワーディング機能のチップ化
- レイヤー間の連携
- この先は本当に光スイッチだろう。。。。。

次の10年

- インターネットのさらなる拡大
- 国内インターネット
- IP電話
- 専用線
- 網内経路制御問題
- 顧客トレンド
- 国内の電話の現状
- 非インターネットIP動画配信
- Ethernet高速化
- パケットネットワークによる電話網の構築

パケットネットワークによる電話網の構築

- 次の10年の一番の課題は「電話」だろう
- とにかく手に入る武器は全部使う必要がある
- 固定電話の回線交換からパケット交換へのシフト
 - 1877年グラハム・ベルAT&T設立以来の大変化
 - IP技術、Ethernet技術、光スイッチ技術をすべて動員
 - 非常に大変で難しい作業
- 携帯電話網のIP化？
 - 固定電話がちゃんと動くまで待つべきだろう(というか待ってほしい)
 - 必ずしもEnd-to-EndのIP化が必要ではないかもしれない
 - Ethernet技術の導入とその範囲

IPネットワークとは？

○ ベスト・エフォート

- しかしいまどきQoS/CoS保障なしでのオペレーションなど。。。

○ コネクション・レス

- バックボーンでは、BGP、MPLS、PWE3などコネクション・オリエンテッド技術満載、ネットワークプロトコルでTCPを使った時点で。。。

○ 自律分散

- 戦争があっても動くネットワーク、実は戦争が無くても落ちる？
- 非シグナリング志向ネットワーク

○ IPネットワーク≠IPルーティング

- ペイロードがIPパケットであれば、IPネットワークと言うべきか