インターネットインフラ特論

Advanced Lecture on Internet Infrastructure

1. インターネットの原理

1. The Principle of the Internet

太田昌孝

mohta@necom830.hpcl.titech.ac.jp

講義の構成

- 前期
 - インターネットインフラ特論
 - 物理層、データリンク層、ネットワーク層
- 後期予定
 - インターネット応用特論
 - トランスポート層、アプリケーション層

Structure of Lecture

- 2nd Quater
 - Advanced Lecture on Internet Infrastructure
 - Physical, Datalink and Network Layers
- 4th Quater
 - Advanced Lecture on Internet Applications
 - Transport and Application Layers

通年の講義の目的

- インターネットの原理、インターネット流の プロトコル設計の「こつ」を理解する
 - エンドツーエンド原理(RFC1958)
 - グローバルコネクティビティ原理
 - スケーラビリティ原理
- ・ 今後はプロトコルの時代
 - 多用なアプリケーションに応じたプロトコル
 - APIは二の次

Purpose of the Lectures

- Understand the Pinciple of the Internet and Knowhow of Internet Style Protocol Design
 - The end to end principle (rfc1958)
 - principle of global connectivity
 - principle of scalability
- We are in Protocol Era
 - Protocols designed for various applications
 - APIs are of secondary importance

講義の参考となる資料

- RFC(一部和訳もあり)
- 「本当のインターネットをめざして」、情報処理学会誌、全36回(1999年4月号~2002年3月号)
- 「インターネットの真実」、週間東洋経済(2 001年1月より2002年4月まで連載)
- 昨年度の講義のOHP
 - ftp://chacha.hpcl.titech.ac.jp/2015/infra*.ppt

Reference Articles for Lecture

- RFC (Request for Comments)
- 「本当のインターネットをめざして」、情報処理学会誌、全36回(1999年4月号~2002年3月号)
- 「インターネットの真実」、週間東洋経済(2 001年1月より2002年4月まで連載)
- Slides used in the Previous Fiscal Year (in Japanese only)
 - ftp://chacha.hpcl.titech.ac.jp/2015/infra*.ppt

前期の構成(1)

- 1. インターネットの原理:エンドツーエンド原理、CATENETモデル
- 2. 物理層:アクセス網、幹線網、光ファイバー、ADSL、電波
- 3. データリンク層: イーサネット、ATM、PPP、IOG
- 4. インターネットワーキング層: IPv4、A RP

Topics of the 2nd Quater (1)

- 1. Principle of the Internet: End to end principle, CATENET model
- 2. Physical Layer: Access Network,
 Backbone Network, Optical Fiber, ADSL,
 Radio Wave
- 3. Datalink Layer: Ethernet, ATM, PPP, IOG
- 4. Internetworking Layer: IPv4, ARP

前期の構成(2)

- 5. インターネットワーキング層:IPv6、N
- 6. インターネットワーキング層:IPセキュリティ
- 7. インターネットワーキング層:NAT、D HCPとAUTOCONF
- 8. ルーティング: IGP、ポリシー、IX、マルチホーミング、モビリティ

Topics of the 2nd Quater (2)

- 5. Internetworking Layer: IPv6, ND
- 6. Internetworking Layer: IP security
- 7. Internetworking Layer: NAT, DHCP & AUTOCONF
- 8. Routing: IGP, Policy, IX, Multihoming, Mobility

前期の構成(3)

- 9. ルーティング:トラフィックエンジニアリング、ROLC、MPLS
- 10. ルーティング:マルチキャスト、経路 表の縮約の不可能性
- 11. 帯域•遅延保証:Guaranteed Servic、DiffServe
- 12. 超高速ルータ、光ルーティング
- 13~? おまけ?

Topics of the 2nd Quater (3)

- 9. Routing: Traffic Engineering, ROLC, MPLS
- 10. Routing: Multicast, Impossibility of Route Aggregation
- 11. Bandwidth & Delay Guarantee: Guaranteed Service, DiffServe.
- 12. Ultra High Speed Routers, Optical Routers
- 13~?. Auxiliary Topics

評価方法

- 中間レポートと、期末レポート
- 出席はとらないが、、、
- 質問やコメントを義務付ける
 - 期中、講義に関する技術的な内容の質問やコメントを最低2回、授業中に行うこと
 - よい質問やコメントは、成績の加点対象
 - 質問者は、講義終了後に名前と学籍番号を申告のこと

Evaluation Method

- Interim and Final Report
- Attendance is not Checked, but, ...
- Questions or Comments are Mandated
 - In the quater, questions or comments with technical content must be made at least twice during lecture
 - Good questions and comments will be awarded with points
 - Declare your name and student ID after each lecture, if you make questions or comments

プロトコルとは?

・ ネットワークで通信する手順

What is Protocol?

• Procedure to communicate over networks

ところで、インターネットとは?

- 電子メイルのことではない
 - 20年前には大真面目で主張されていた
- ウェブのことでもない
 - 現在は勘違いしている人が多い
- ・アプリケーションのことではない
- インターネットはIP(インターネットプロトコル)を用いて、インターネットの原理に基づいて、端末間が直接接続された網である

BTW, What is the Internet?

- Not e-mail
 - seriously thought so 20 years ago
- Not web, either
 - many still misunderstand so
- Is not applications
- The Internet is a network directly connecting terminals based on the principle of the Internet using IP (Internet Protocol)

エンドツーエンド原理 網の中抜き原理

- 端末(エンド)でできることは網側ではやらない
 - 網機器は単能(端末を結ぶだけ)、高速
- 直接関係する端末でできることは他の端末ではやらない
 - スケーラブル(負荷が集中しない)
 - 高信頼(端末が動いてなんらかの経路で通信 さえできればシステムは動作)

End to End Principle Disintermediated Networking

- Don't do things implemented by terminals (end) in the network
 - network equipment has only single function (to connect terminals) and is high speed
- Don't do things implemented by directly involved terminals by other terminals
 - scalable (no load consentration)
 - highly reliable (system works if only terminals are working and can communicate each other over some route)

インターネットじゃないもの(1) 電子メイル

- UUNET(JUNET)はインターネットではな かった
- パソコン通信もインターネットではなかった
- 電子メイルはインターネット上でも動くアプリケーションの一種
 - その他のネットワーク(電話網)上でも動く
- 昔の(海外)電子メイルは有料だった
 - 今でも携帯電話網上の電子メイルは一部有料

What is not the Internet (1) e-mail

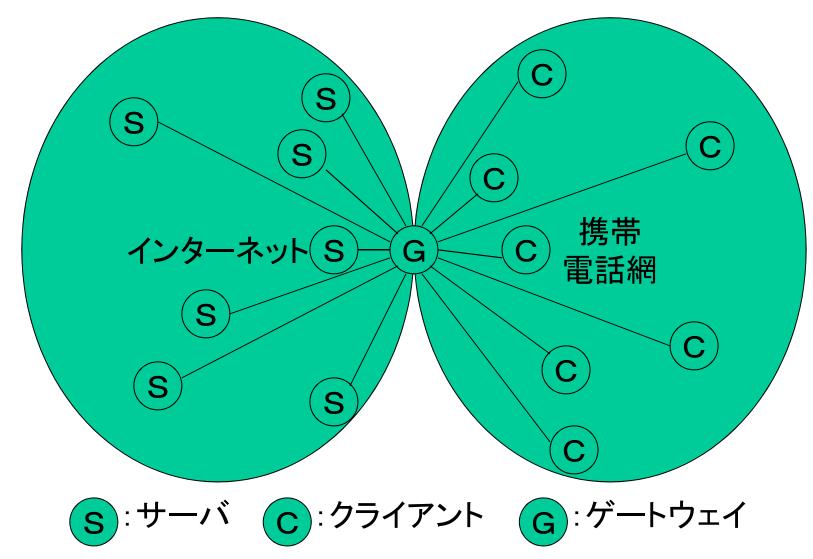
- UUNET (JUNET) was not the Internet
- Communication System for Personal Compters was not the Internet
- E-mail is an application works on the Internet
 - also works on other (phone) networks
- In the past, (oversea) e-mail was charged
 - some mail on mobile phone is still charged

インターネットじゃないもの(2) ウェブ

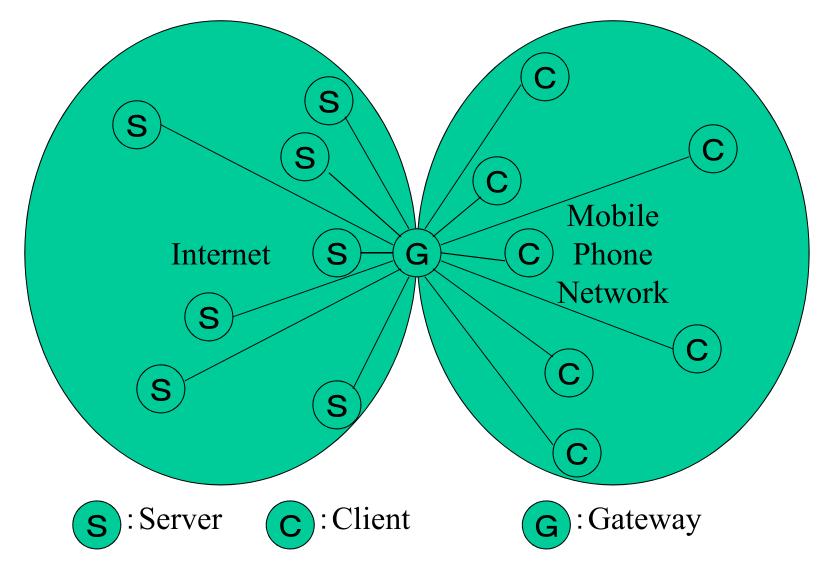
- ウェブはインターネットではない
 - マイクロソフトはそのへんを誤魔化しているが
- ウェブもインターネット上でも動くアプリケーションの一種
 - その他のネットワーク上でも動く
 - 携帯電話網からウェブを見るのは本来有料
 - もちろん、携帯電話網はインターネットではない

What is not the Internet (2) Web

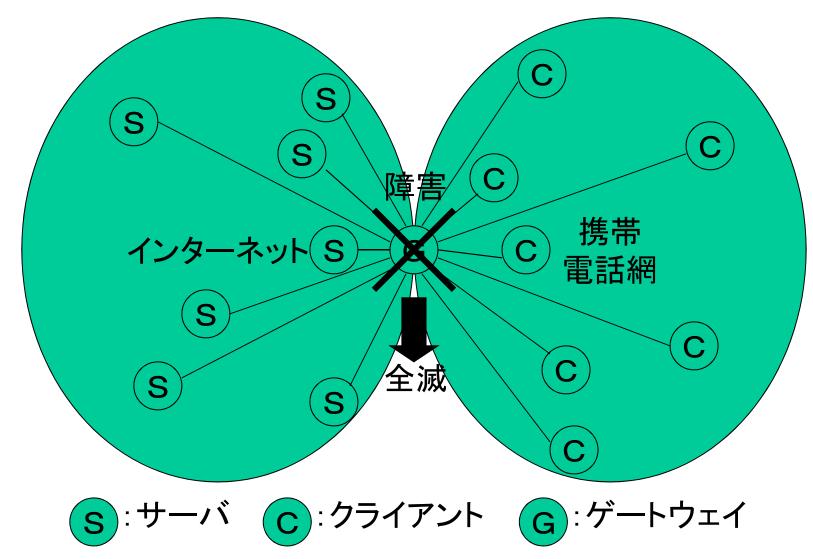
- Web is not the Internet
 - though Microsoft makes it obscure
- Web, too, is an application works on the Internet
 - also works on other networks
 - web browsing from mobile phone network is, in principle, charged
 - mobile phone network is not the Internet, of course



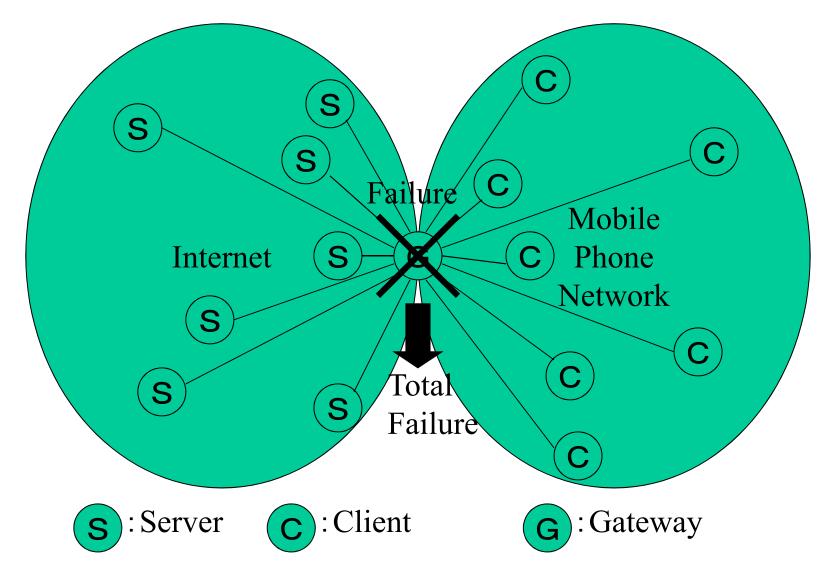
携帯電話網からのウェブの利用 サーバとクライアントが分離されゲートウェイ経由で通信



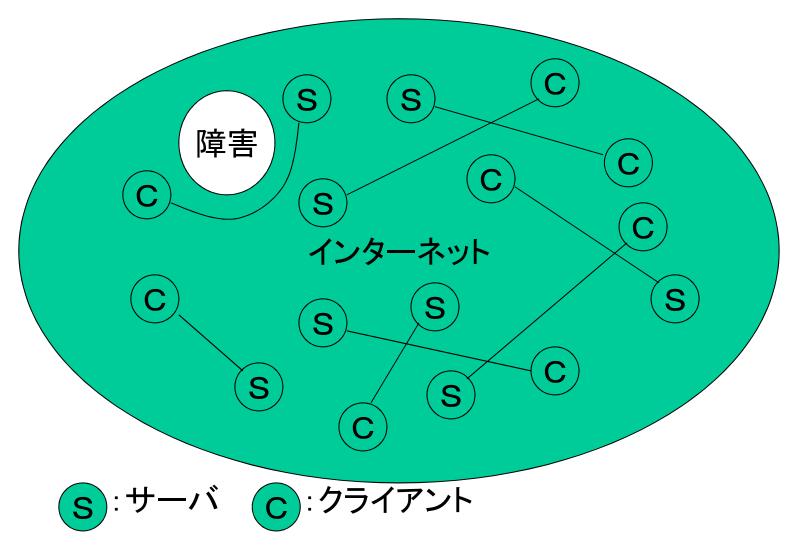
Web Browsing from Mobile Phone Network
Servers and clients are separated and communicate through the gateway



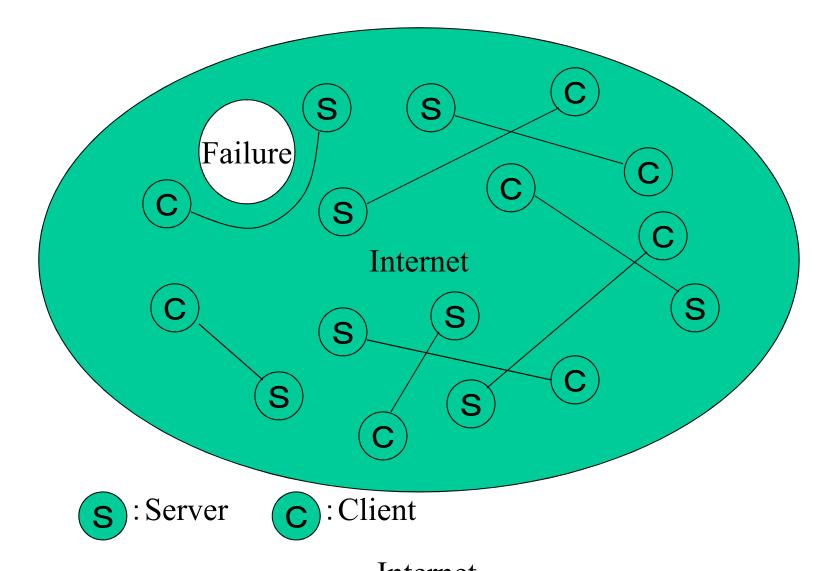
携帯電話網からのウェブの利用 サーバとクライアントが分離されゲートウェイ経由で通信



Web Browsing from Mobile Phone Network
Servers and clients are separated and communicate through the gateway



インターネット サーバとクライアントが混在して最短経路て通信



Internet
Servers and clients are mixed and communicate with shortest pathes

インターネットじゃないもの(3) 電話

- ・電話は電話網ではない
- 電話は電話網上でも動くアプリケーション
 - その他のネットワーク上でも動く
 - インターネット電話自体は無料
- もはや電話はインターネットの販促ツール
 - わざわざ電話網をサポートする価値なし
 - 携帯電話はモバイルインターネットの販促ツール

What is not the Internet (3) Phone

- Phone is not phone network
- Phone is an application works on phone networks
 - also works on other networks
 - phone over the Internet is free of charge
- Phone, today, is a promotion tool to sell Internet services
 - no need to support phone network
 - mobile phone is a promotion tool to sell mobile

インターネットじゃないもの(4) 電話網

- ・電話は電話網ではない
- 電話は電話網上でも動くアプリケーション
 - その他のネットワーク上でも動く
 - インターネット電話自体は無料
- ・もはや電話網は不要
 - 携帯電話網も同じ
 - 電話番号も不要
 - IPアドレスが基本(内線番号もあってもいい)

What is not the Internet (4) Phone Network

- Phone is not phone network
- Phone is an application works on phone networks
 - also works on other networks
 - phone over the Internet is free of charge
- We don't need phone networks
 - neither are mobile phone networks
 - phone numbers are not necessary, either
 - basically required is IP addresses (though alternative

ネットワーク

- 物流網
 - 郵便、宅配便、コンビニ
- 情報通信網
 - 出版網(書籍、新聞、レコード(CD)、映画)
 - 金融網
 - 電話網
 - 放送網
 - **インターネット**

Networks

- Physical Distribution Networks
 - Postal Service, Parcel Services, Convenience
 Stores
- Information Communication Networks
 - Publishing Network (Book, News Paper, CD, Movie)
 - Financial Network
 - Phone Network
 - Broadcast Network
 - _ the Internet

インターネットは情報通信網を 中抜きする

- インターネットは情報通信の価格破壊
 - 出版網、金融網、電話網、放送網は消える
 - 社会の情報通信コストの削減
 - ISPビジネス自体は儲からない
- ・ 出版、金融、電話、放送というサービスは
 - インターネット上に移行して残る
 - 社会の活力は増大する

Internet Disintermediate ICN

- Price Destruction of ICN
 - Publishing, financial, phone and broadcast networks will disappear
 - IC cost of the society decreased
 - ISP business itself is not profitable
- Publishing, financial, phone and broadcast services will:
 - remain, but, on the Internet
 - social activities increase

出版網

- 同じ情報を大量に配布
- 情報流通は遅くていい
- 著作権法による保護
- いまのインターネットの好餌
 - 壊滅寸前

Publishing Network

- Mass Distribution of Same Information
- Delay of the Distribution may be Tolerated
- Protected by Copyright Act
- The First Victim of the Internet
 - Collapsing

金融網

- お金のやりとりを管理
- 物流網でもあるが、今や、情報通信網としての面がはるかに大きい
- セキュリティー!!!
 - 事故を起こさないことではない
 - 事故の時、誰が損失をかぶるか

Financial Network

- Manage Transfer of Money
- Partly, Phisical Distribution Network, but, today, mostly ICN
- Security!!!
 - Not that there is no accident
 - Who will pay the loss on accidents

電話網

- ・ 音声を実時間で伝送する網
 - 音声伝送の帯域を確保
 - 音声伝送の遅延を最小化(保証)
- 専用線事業も
 - あくまで音声伝送事業が主
- 遅くて高い
- ・電電公社として保護、電気通信事業法で 開放

Phone Network

- Network for Realtime Voice Transfer
 - Allocate bandwidth for voice transfer
 - Minimize (guarantee) delay for voice transfer
- Dedicated Line Business may be Offerred
 - but, primary business is voice transfer
- Slow and Expensive
- Was Protected as Public Company
 - Leberated by Telecommunication Business Law

放送網

- 音声、画像を実時間で多数に伝送する網
 - 伝送帯域を確保
 - 遅延を最小化
- ・電波による広域一対多通信
 - ブロードキャスト/マルチキャスト
- ・放送法による保護

Broadcast Network

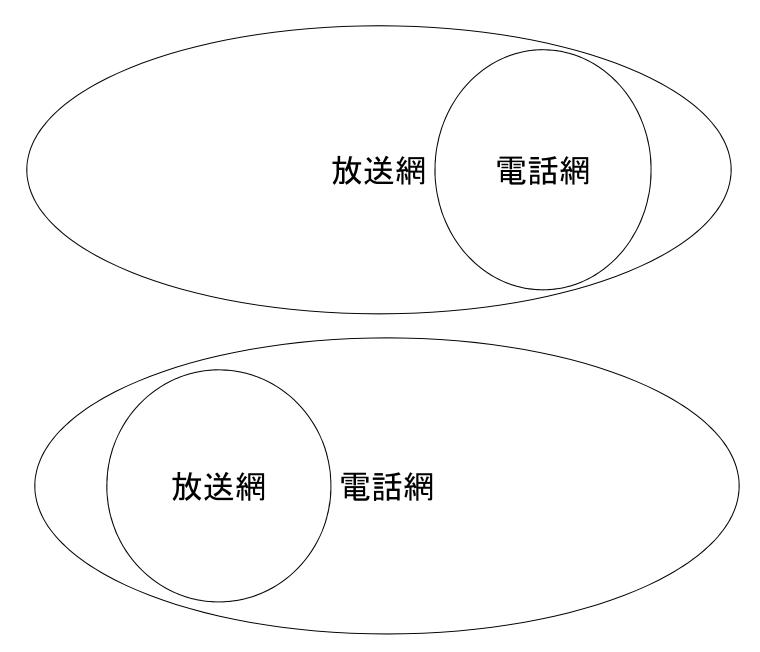
- Network to Transfer Voice/Image to Many in Realtime
 - Allocate bandwidth for the transfer
 - Minimize delay
- Wide Area One to Many Communication over Radio Waves
 - Broadcast/Multicast
- Protected by the Broadcast Act

放送と通信の融合

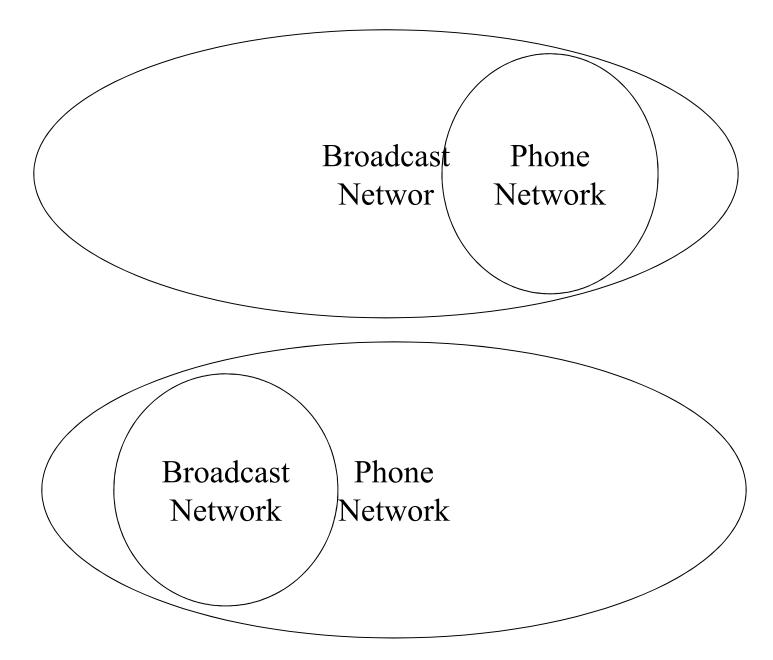
- ・電話網からみて
 - 電話網でも一対多通信は可能
 - ・ 電話網(BISDN)への放送網の統合
- ・放送網からみて
 - 放送のフィードバックを電話網から受ける
 - 電波でも1対1通信は可能
 - 放送網への電話網の統合?
- ・インターネットからみて
 - ぜんぶインターネットに統合

Integration of Broadcast and Telecommunication

- Viewed from Phone Network
 - one to many transfer over phone network
 - integration of broadcast network to phone network (BISDN)
- Viewed from Broadcast Network
 - receive feedback through phone network
 - one to one communication over radio wave possible?
 - integration of phone network to broadcast network
- Viewed from the Internet



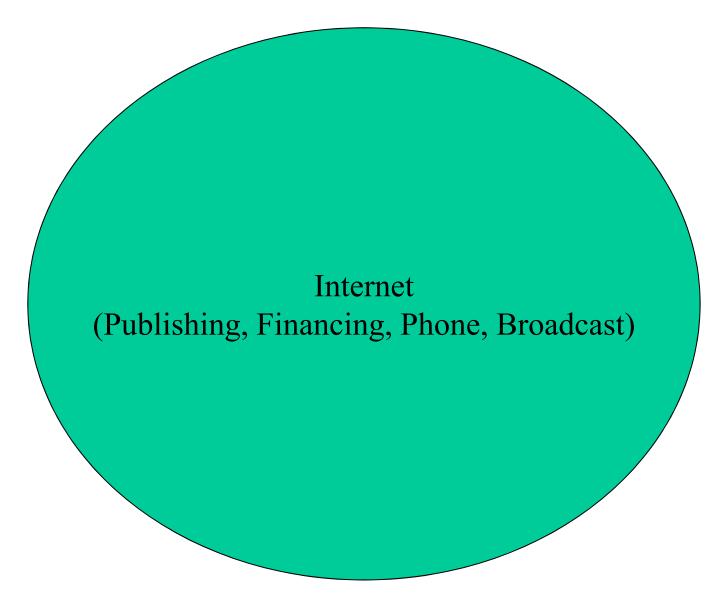
同床異夢の「放送と通信の融合」



Two Visions on "Integration of Broadcast and Telecommnication



インターネットによる情報通信サービスの統合



Integration of IC Services by the Internet

プロトコルのレイヤリング

- 扱う対象の抽象度に応じて層化
 - OSIの7層モデル
 - _ インターネットの5層モデル
- プログラムでいえば、サブルーチン化(構造化)に相当

Layering of Protocols

- Have Layers based on Level of Abstraction
 - 7 layer model of OSI (Open Systems Interconnection)
 - 5 layer model of the Internet
- Corresponds to Subroutines (Structuring) in Programming

アプリケーション層 7層 プレゼンテーション層 6層 セッション層 5層 トランスポート層 4層 ネットワーク層 3層 データリンク層 2層 物理層 1層

OSIのレイヤリング構造

Application Layer	Layer 7
Presentation Layer	Layer 6
Session Layer	Layer 5
Transport Layer	Layer 4
Network Layer	Layer 3
Datalink Layer	Layer 2
Physical Layer	Layer 1

Layering Structure of OSI

アプリケーション層
トランスポート層
ネットワーク層
データリンク層
物理層

インターネットのレイヤリング構造

Application Layer

Transport Layer

Network Layer

Datalink Layer

Physical Layer

Layering Structure of the Internet

物理層

- ・ 物理現象と情報を対応させる
 - 電圧<->0、1
 - 光のあるなしくー>0、1
 - 振幅と位相<一>O、...、63
- ・ 複数の同質の物理層の統合も可能(リピーター)
- プログラミングでいえば、機器制御のファームウェア

Physical Layer

- Map Physical Phenomena and Information
 - Voltage high/low <-> 0/1
 - Light on/off <-> 0/1
 - Amplitude and Phase <-> 0/1/.../63
- Multiple Similar Physical Layers may be Integrated (Repeaters)
- Corresponds to Control Firmware in Peripheral Devices in

データリンク層

- ・ ネットワーク層と物理層のつなぎ
- ・物理層が3個以上の機器からなる場合
 - 相互の区別が必要
- 複数の異質な物理層の統合も可能
 - 局所的な中継(ブリッジ)
- プログラミングでいえば、デバイスドライバ

Datalink Layer

- Joint between Network and Physical Layers
- If Physical Layer have more than 2 devices
 - distinction between them is necessary
- Multiple Dissimilar Physical Layers may be <u>Integrated</u>
 - Local Relays (Bridges)
- Corresponds to Device Drivers in Programming

ネットワーク層 インターネットワーキング層

- ・多数のデータリンク層を<u>統合</u>して大域的な ひとつの網として動作させる
- 大域的な中継(ルータ、ゲートウェイ)
- プログラミングでいえば、ファイル管理やプロセス間通信

Network Layer Internetworking Layer

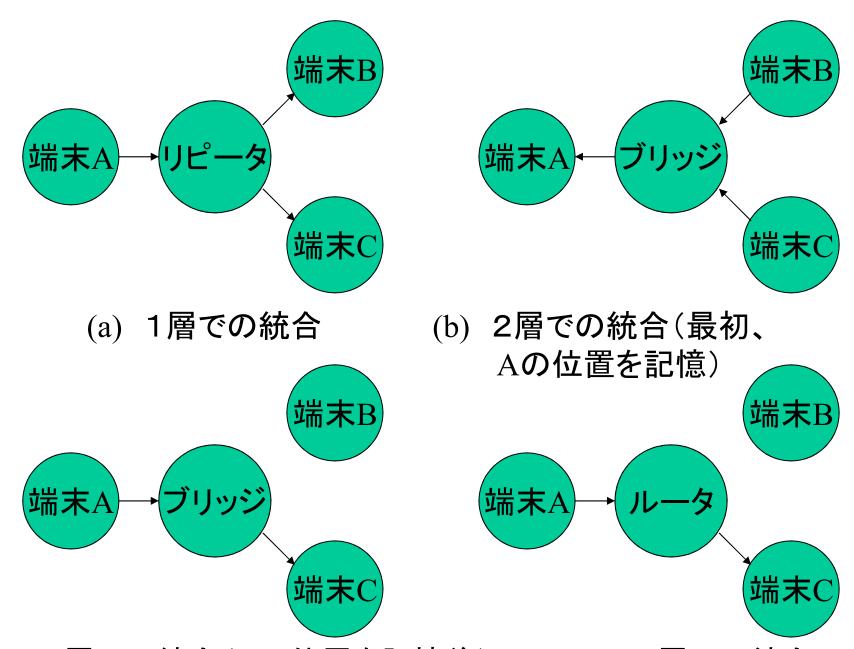
- Integrate Many Datalink Layers to form a Global Network
- Global Relaying (Routers, Gateways)
- Corresponds to File Manager and Interprocess Commnication in Programming

「統合」をどこでやる?

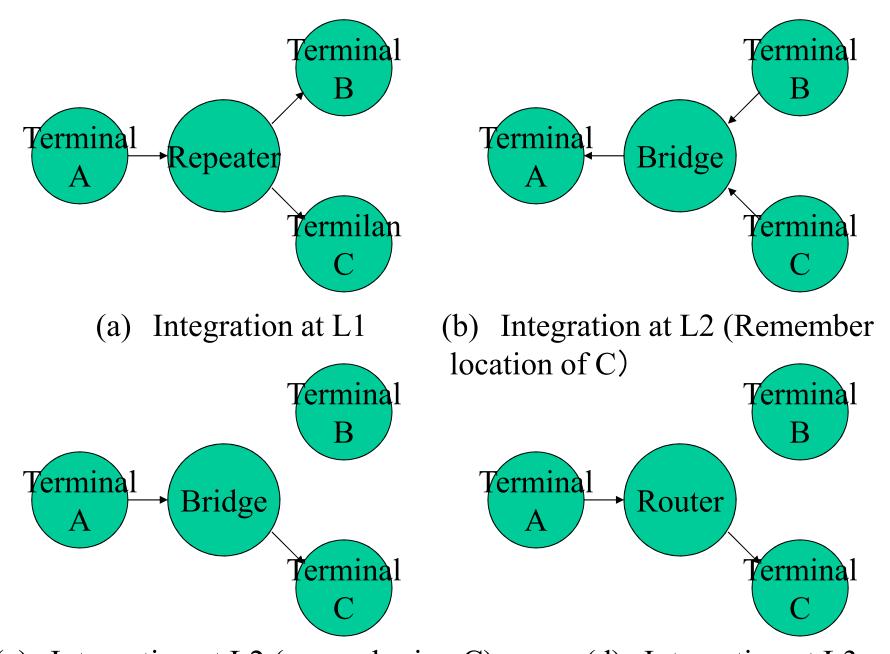
- どの層でやってもいい
 - だからといってあちこちでやるのは無駄
- どこかの層だけに任せるのがきれい
 - インターネットワーキング層では統合は必然
 - データリンク層でやる必要はない
 - ・データリンク層の単純化
 - 物理層でやると帯域の無駄
 - ・多くの端末が一つの物理現象を共有

Which Layer Should be Used for the "Integration"

- Any of the 3 Layers
 - but, is wasteful to do so everywhere
- Clean to do it only at one Layer
 - Internetworking layer MUST Integrate
 - Not necessary at datalink layer
 - simplify datalink layer
 - Waste bandwidth at physical layer
 - Many terminals share single physical phenomena



(c) 2層での統合(Cの位置を記憶後) (d) 3層での統合層ごとの中継機器の振る舞い(AからCへのパケット)



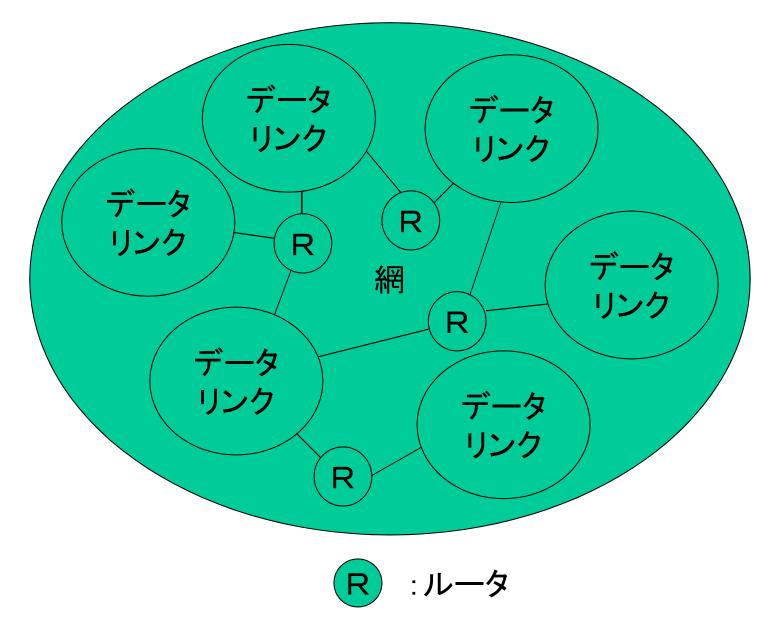
(c) Integration at L2 (remembering C) (d) Integration at L3 Relaying Behavior at each Layer (Packet from A to C)

網間接続

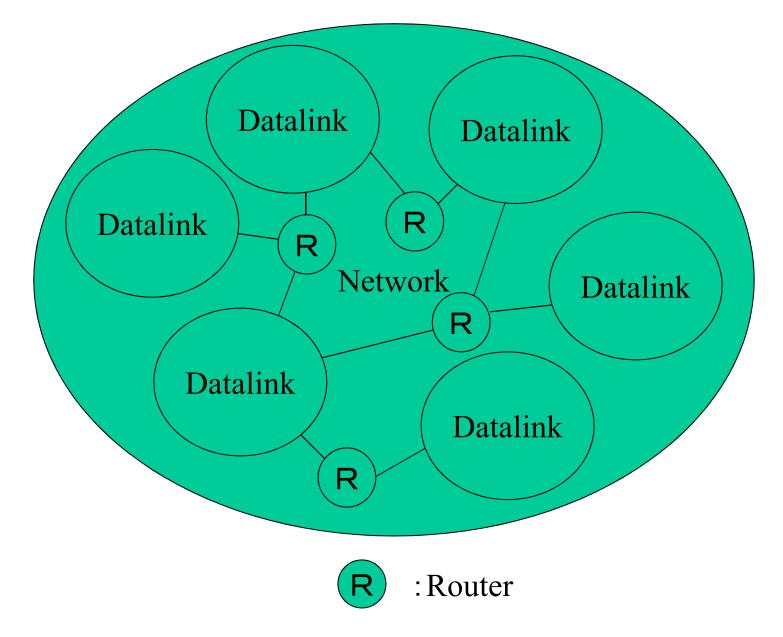
- 網どうしの「統合」
 - ネットワーク層以上の層での統合
- 矛盾!
- エンドツーエンド原理とは端末の直接通信
 - インターネットでは網間接続はありえない
 - 1つの網(インターネット)だけで全体を覆う
 - グローバルコネクティビティ原理
 - 網間接続の仕様をあれこれ考えるより楽

Interworking

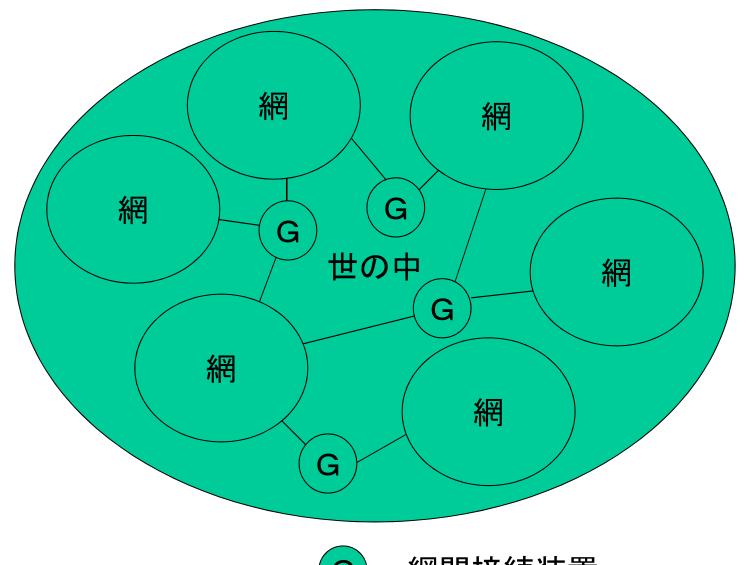
- Integration of Networks
 - above the Network Layer
- Contradiction
- The End to End Principle requires direct communication between Terminals
 - Interworking impossible on the Internet
 - A single network should covers the entire world
 - principle of global connectivity
 - Easier than specifing various interworking



ネットワークとデータリンク

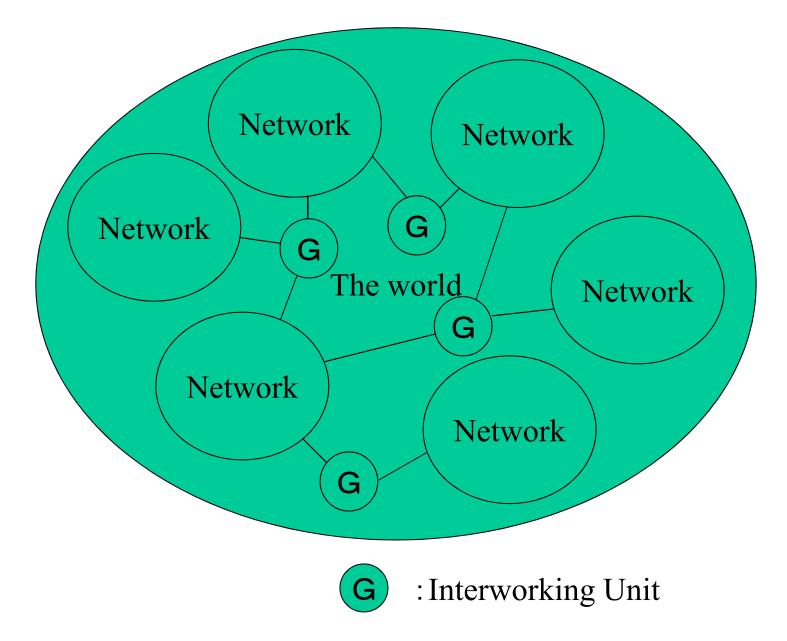


Network and Datalink

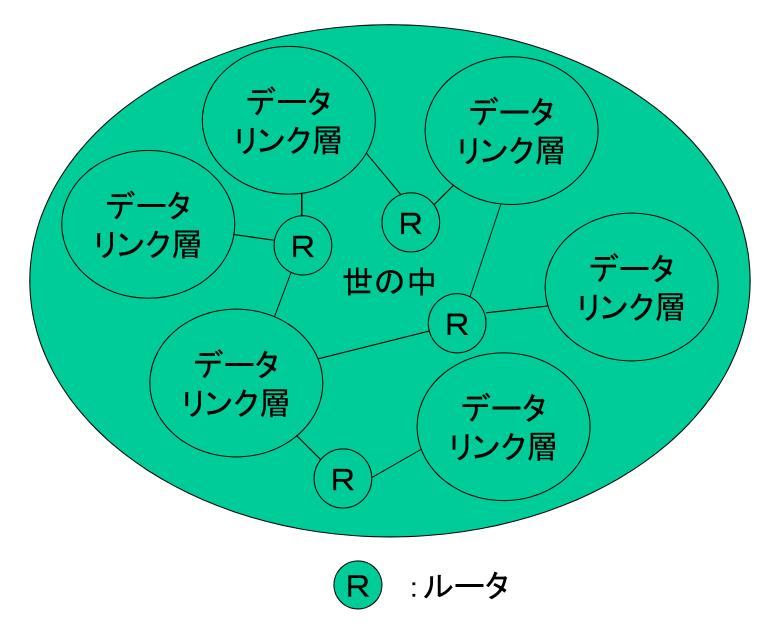


G:網間接続装置

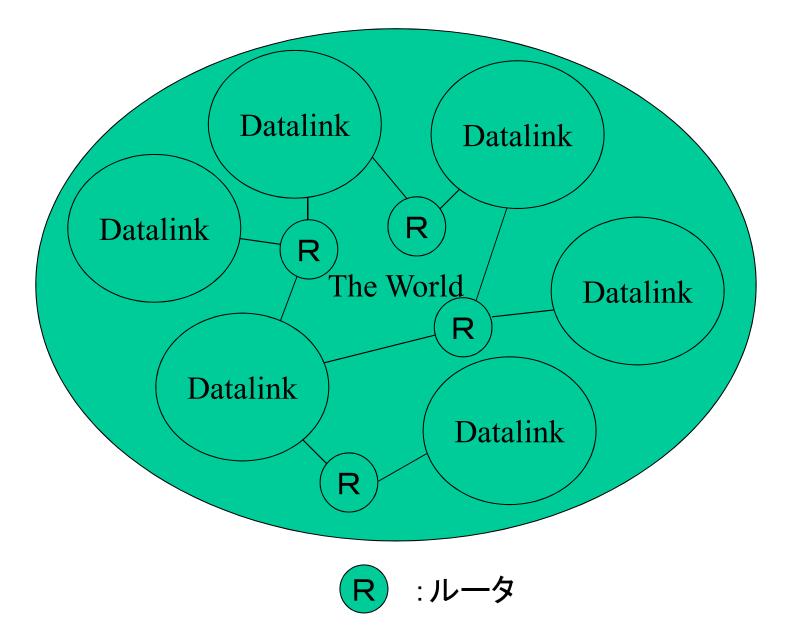
世の中とネットワーク層



The World and Networks



世の中とインターネット



The World and the Internet

トランスポート層

- ・ ネットワーク層では個々の端末を識別
- トランスポート層では個々の通信を識別
 - 同一端末間に複数の通信がありえる
 - ・ 別の通信は別のプロセスで処理
 - ・ネットワーク中ではそれらの要求する帯域などが 異なるかも、、、
 - ベストエフォートインターネットではどうでもいい
- プログラムでいえば、プロセス管理

Transport Layer

- Network Layer Identify Each Terminal
- Transport Layer Identify Each Commication
 - Multiple communications may exists between a pair of terminals
 - different communication is processed by different process
 - may require different bandwidth et.
 - unimportant in the current best effort Internet
- Corresponds to Process Manager in

セッション層、プレゼンテーション層、アプリケーション層

- プログラミングでいえば、個々のプロセス 内部の構造に相当
- ・ ネットワークでいっても、端末内の個々の プロセス内部の構造に相当
- 網間接続がなければ、区別に意味なし
 - インターネットではアプリケーション層のみ

Session, Presentation and Application Layers

- Corresponds to Internal Structure of each Process in Programming
- Corresponds to Internal Structure of each Process in Terminals
- Without Interworking, there is no Point to have 3 Layers
 - Layering of the Internet has the Application
 Layer only

トランスポート層とアプリケーショ ン層

- ベストエフォートでは、端末内部の区別で しかない
- 個別通信の個別プロセスへの振り分けだけはトランスポート層
- ・それ以上の区別には意味がない
 - 多くのアプリケーションで共通に使うプロトコル (TCP(信頼性確保と帯域管理のプロトコル) 等)は慣例的にトランスポート層に分類

Transport and Application Layers

- With Best Effort Network, Distinction is within Each Terminal
- Assigning Packets of Each Communication to Corresponding Process is Transport
- Further Distinctions is not Meaningful
 - Protocols shared by many applications (e.g.
 TCP (assure reliability and manage bandwidth))
 are, traditionally, classified as Transport

インターネットのレイヤリング

- 物理層、アプリケーション層は必須
- インターネットワーキング層はできる限りの ことをやる
- データリンク層、トランスポート層は極力なにもやらない

Layering of the Internet

- Physical and Application Layers are Essential
- The Internetworking Layer does as Much Things as Possible
- Datalink and Transport Layers should Avoid to do Thing

アプリケーション層

トランスポート層

インターネットワーキング層 こここそインターネット

データリンク層

物理層

インターネットのレイヤリング構造

Application Layer

Transport Layer

Internetworking Layer

Datalink Layer

Physical Layer

Here is the Essence of the Internet

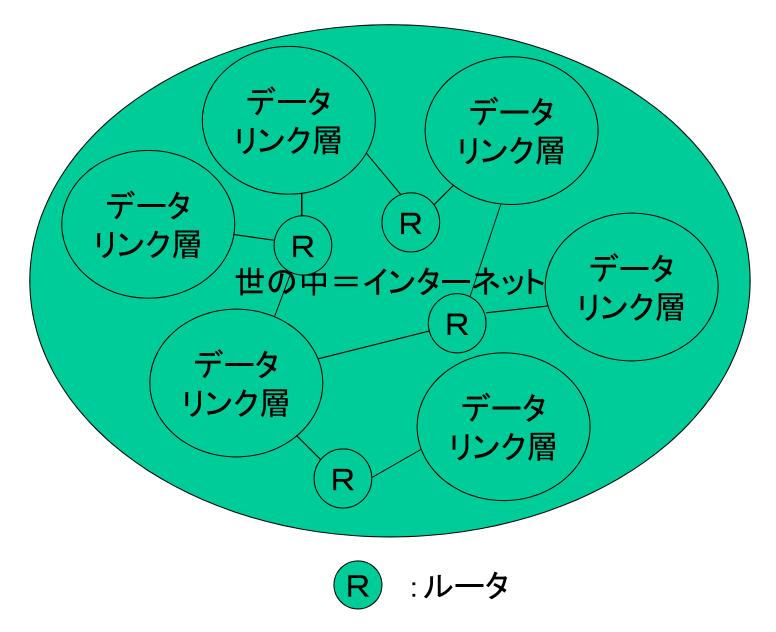
Layering Structure of the Internet

インターネットの構造

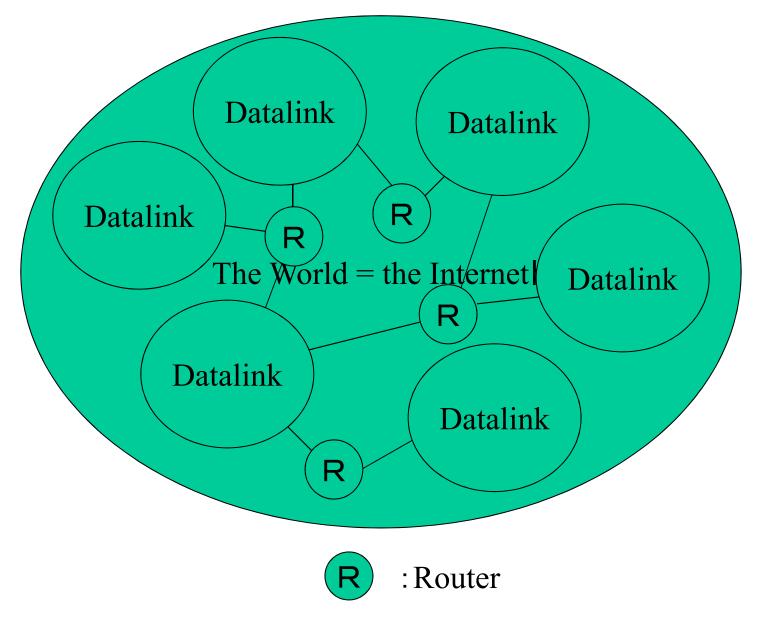
- CATENETモデル
 - 多数の小さな(機器の数が少ない)データリンク層をIP(Internet Protocol)ルータで相互接続したもの

Structure of the Internet

- CATENET Model
 - Many small (w.r.t. # of devices) datalinks
 interconnected by IP (Internet Protocol) routers



CATENETモデル



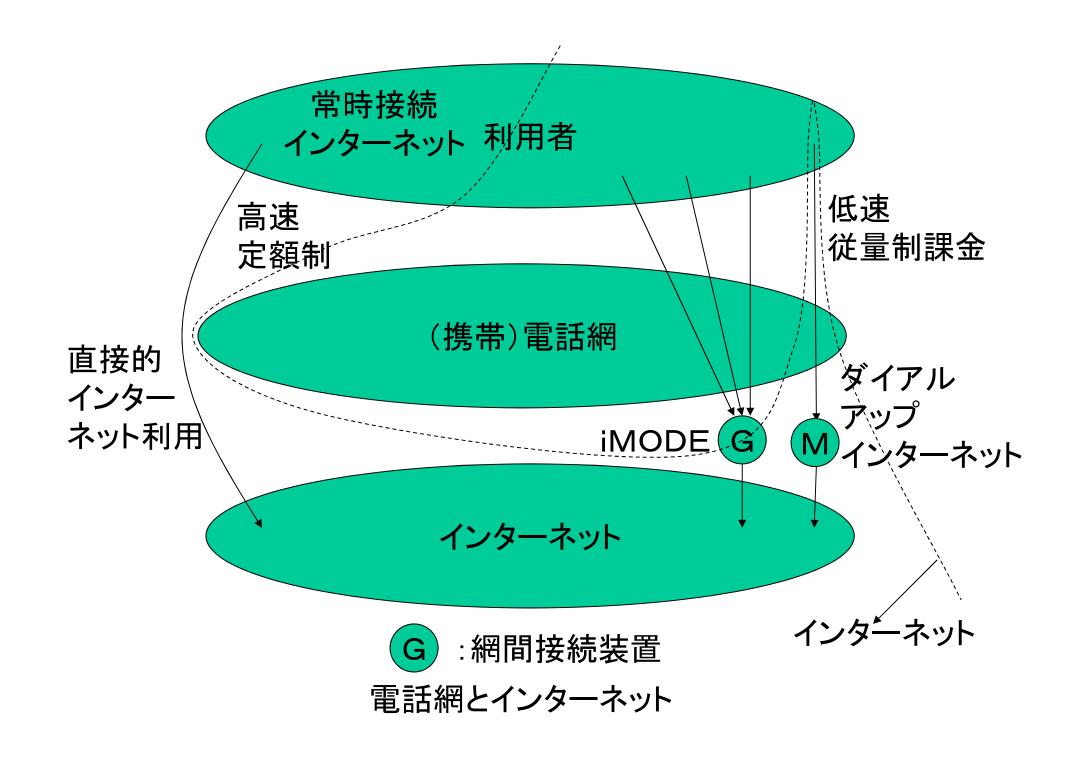
CATENET Model

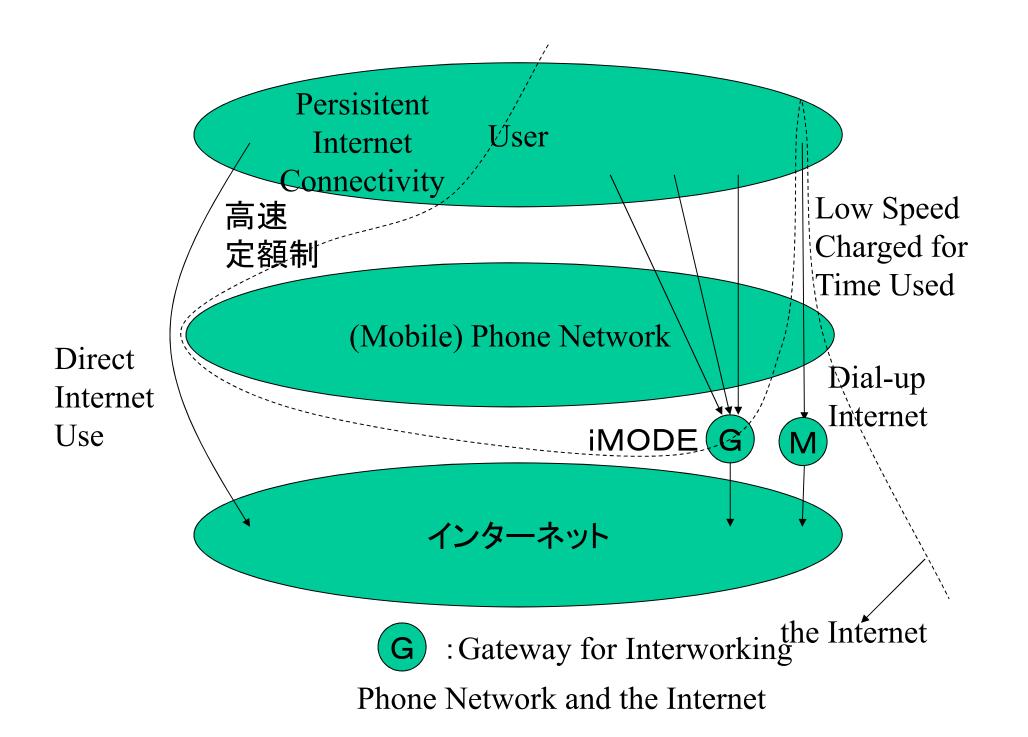
インターネットと網の構造

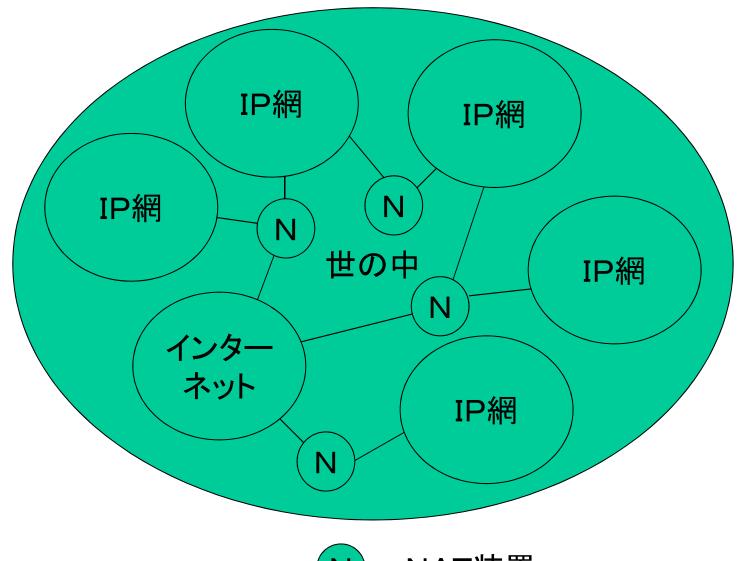
- インターネットの例
 - ダイアルアップインターネット
- インターネットでない例
 - iモード
 - IPだが、トランスポート層も中継
 - 従来のNAT
 - IPだが、アドレス等を書換え端末には隠蔽
 - ― 隠蔽の為の書換えで、トランスポート層以上で網間接続

The Internet and Structure of Networks

- Example of Internet
 - Dial-up Internet
- Example of non Internet
 - i-mode
 - IP, but, relaying at transport layer
 - Legacy NAT
 - IP, but, addresses etc. are modified, which is not visible to terminals
 - Interworking at the transport layer and above

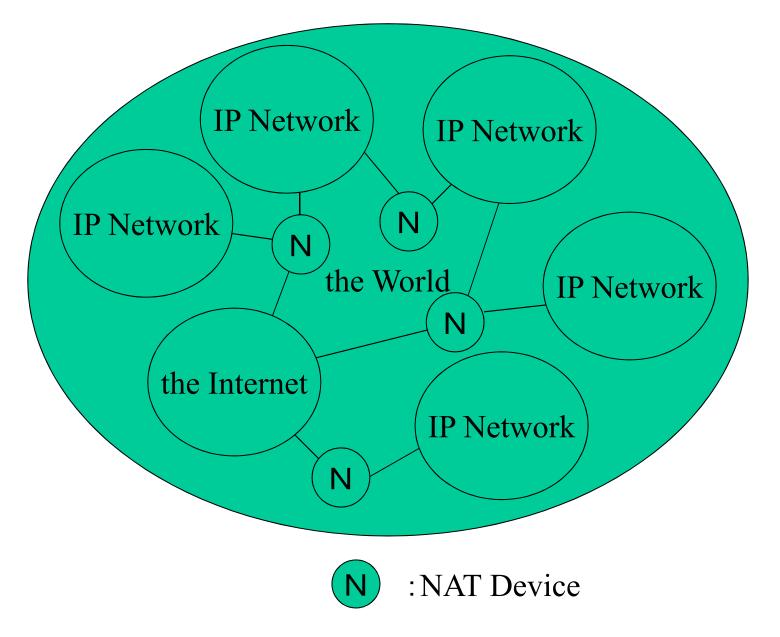






N :NAT装置

NATとインターネット



NAT and the Internet

インターネットのデータ形式

- データをパケットにまとめて転送
- パケットは個別に行き先をもつ
- IPv4では、20バイトのインターネットワーキング層へッダを付加
- トランスポート層へッダも付加される

Data Format of the Internet

- Data are Assembled to Form Packets
- Each Packet has its Own Destination
- With IPv4, 20B Internetworking Layer Header is Attached
- In Addition to a Transport Layer Header

4 ヘッダ長パケット長その他情報イ層プロトコルヘッダーチェックサム送信者アドレス受信者アドレスオプション(可変長、実際には使われない)送信者ポート番号送信者ポート番号受信者ポート番号

トランスポートヘッダの残りとペイロード

4 Header Length		Other Information Packet Length	
		L4 Protocol	Header Checksum
Source Address			
Destination Address			
Optional Header (Variable Length, not Actually Used)			
Source Port Number			Destination Port Number

4 Bytes

Remaining Transport Header and Payload

Format of IPv4 Packets

オプションIPヘッダを使わない (が使えない)わけ

- ルータで処理が必要なオプション
 - ルータの処理が複雑になる
 - ルータが遅くなる
 - ルータが落ちる
 - そもそも、ルータで処理が必要か?
 - ・エンドツーエンド原理によりオプションは有害無益
- ルータで処理が不要なオプション
 - トランスポート層以上のオプション

Reason Why Optional OP Header is not (can not be) Used

- Options Processed by Routers
 - Router Processing become Complex
 - routers become slower
 - routers may crash
 - In processing by routers necessary at all?
 - according to the end to end principle, options are harmful and useless
- Options not Processed by Routers
 - Options at or above the transport layer

今回のまとめ

- エンドツーエンド原理はインターネットの基本原理
- 網間接続はエンドツーエンド原理違反
 - グローバルコネクティビティ原理の導出
 - 7層モデルは無意味
- オプションIPヘッダはエンドツーエンド原理 違反

Wrap-up

- The End to End Principle is the Fundamental Principle of the Internet
- Interworking is against the End to End Principle
 - global connectivity principle derived
 - 7 layer model is meaningless
- Optional IP Headers are against the End to End Principle