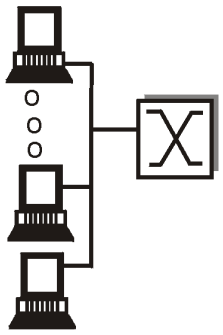
データリンク層の技術 (3)

Suguru Yamaguchi
Nara Institute of Science and Technology
Department of Information Science

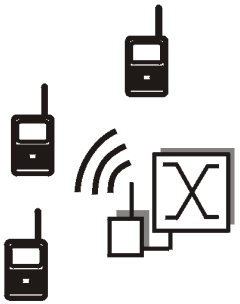
Medium Access Control (MAC)

Access Channel

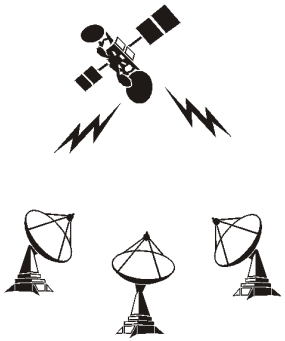
- 二種類の “リンク”:
  - Point-to-point
    - PPP :ダイヤルアップ接続
    - イーサネットスイッチとホストの間のPoint-to-point リンク
  - ブロードキャスト (LAN内など)
    - 伝統的なイーサネット
    - 802.11 wireless LAN



shared wire
(e.g. Ethernet)



shared wireless
(e.g. Wavelan)



satellite

cocktail party

MAC Protocols (1)

- 共有ブロードキャストチャネル
  - 別々のノードからの複数同時通信：干渉の発生
  - 複数のシグナルを同時に受信する：衝突の発生
- 理想的な Medium Access Protocol
  - 1つのノードが通信する時、伝送率*R*で伝送
  - *M*ノードが通信する時、平均伝送率*R/M*で伝送
  - 完全分散:
    - 通信を調停する特別なノードが存在しない
    - 時間同期、スロット同期を使用しない
  - 単純

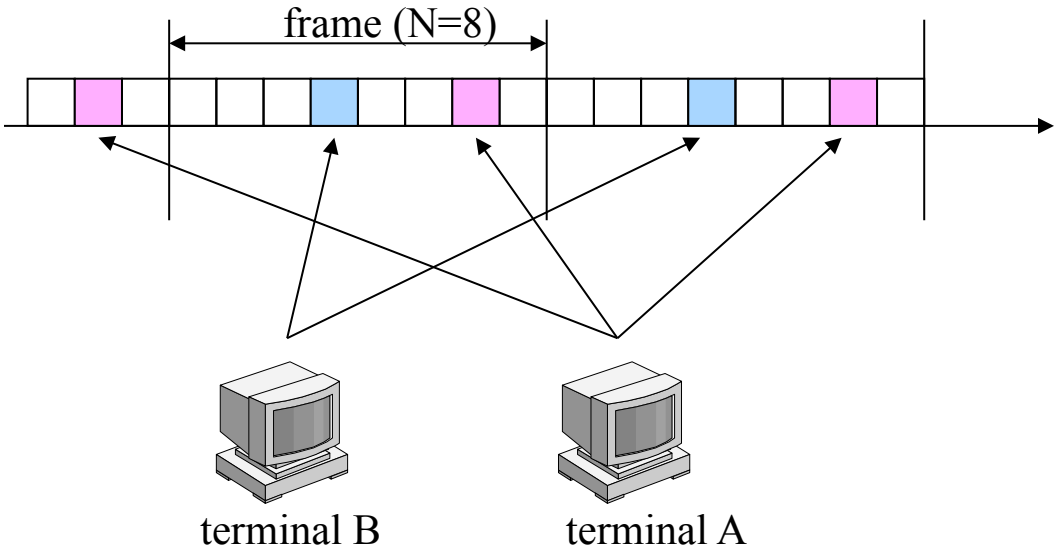
MAC Protocols (2)

3つの技術:

- Channel Partitioning
  - チャネルを小さな「ピース」(タイムスロット、頻度、コード等)に分割
  - 排他的利用のためにピースを各ノードに割り当てる
- Random Access
  - チャネルが分割されてなければ衝突が起きてしまう
  - 衝突からの回復
- Taking turns
  - 各ノードが自分の順番を持つ
  - 多くのデータを送るノードがより長い順番を持つ

TDMA (1)

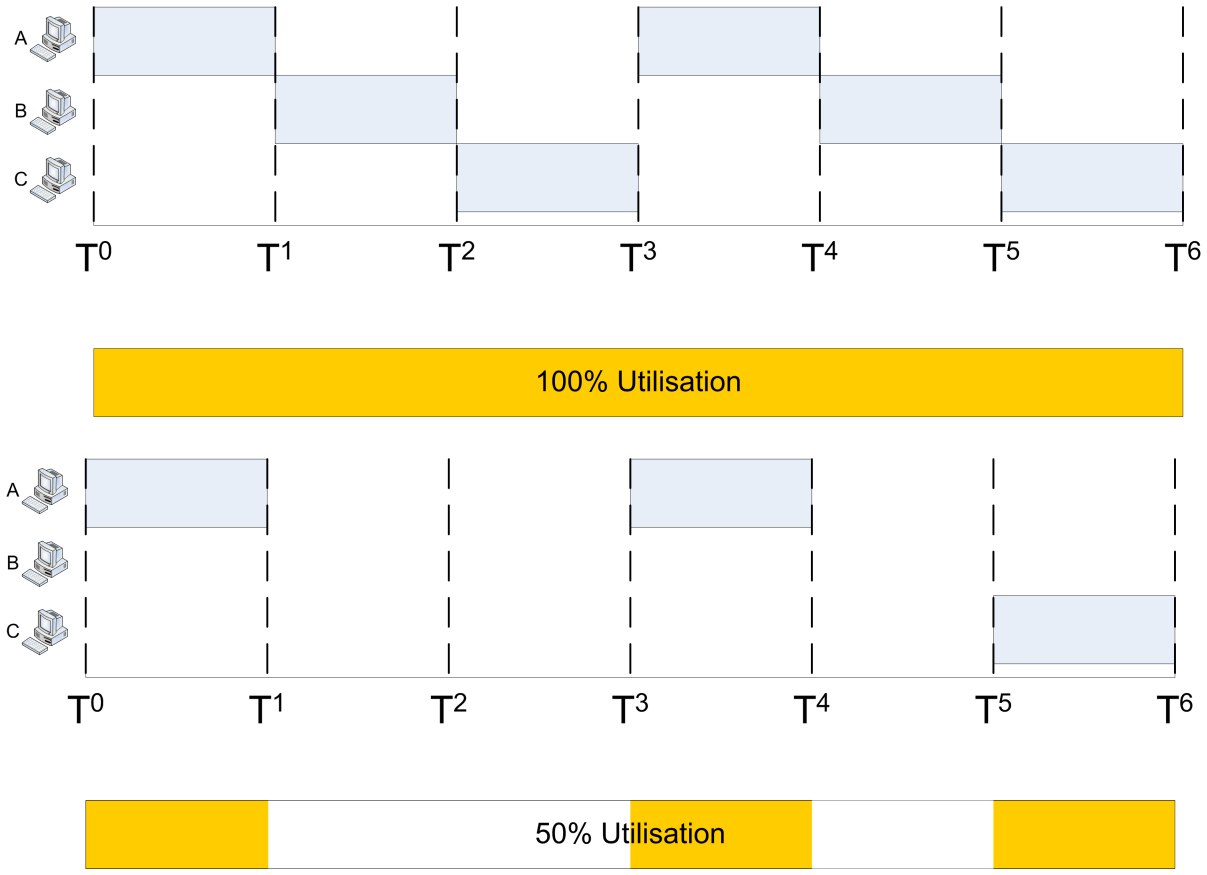
- Time Division Multiple Access
  - チャネルに順番にアクセス
  - 媒体メディアをタイムスロットに分割
  - 各ステーションは固定長のスロットを得る
    - Length = 各ラウンドの通信時間
  - 複数のスロットで1フレームを構成
  - 各端末は特定のスロットを使用する
  - 使わないスロットは
    アイドル状態にしておく



TDMA (2)

- スロットの確保と解放
  - 到底的多重化
    - 前もって管理者が割り当てておく
  - 動的割り当ては複雑
    - 各端末に特別な機能が必要
- 時間同期
  - 時間を基準にしてスロットを定義
  - 各端末とスイッチングノード間での時間同期が必要

TDMA (3)



不安定な通信では時間同期はうまくいかない

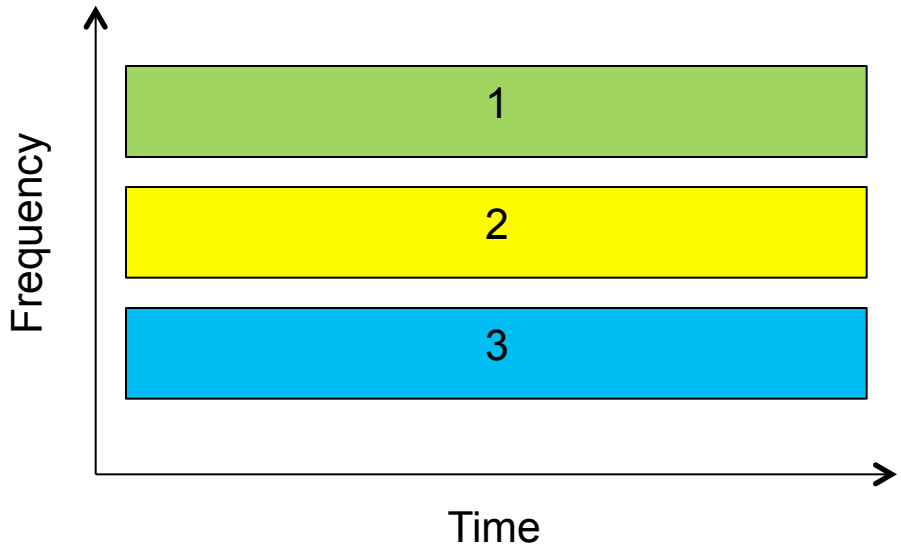
TDMA から ATMへ (余談)

- TDMAは媒体アクセス技術として広く利用されている
  - 多重化技術と同様に単純である
  - 性能を保証できる
  - TDMA は多くの製品で利用されている
    - Based on STM (Synchronous transfer mode)
    - 同時通信ネットワークは現在も利用されている
- TDMAを LANでの通信で使用するのは困難
  - 各端末の時間同期を保証できない
  - 端末が使用する帯域量を事前に予測することが出来ない

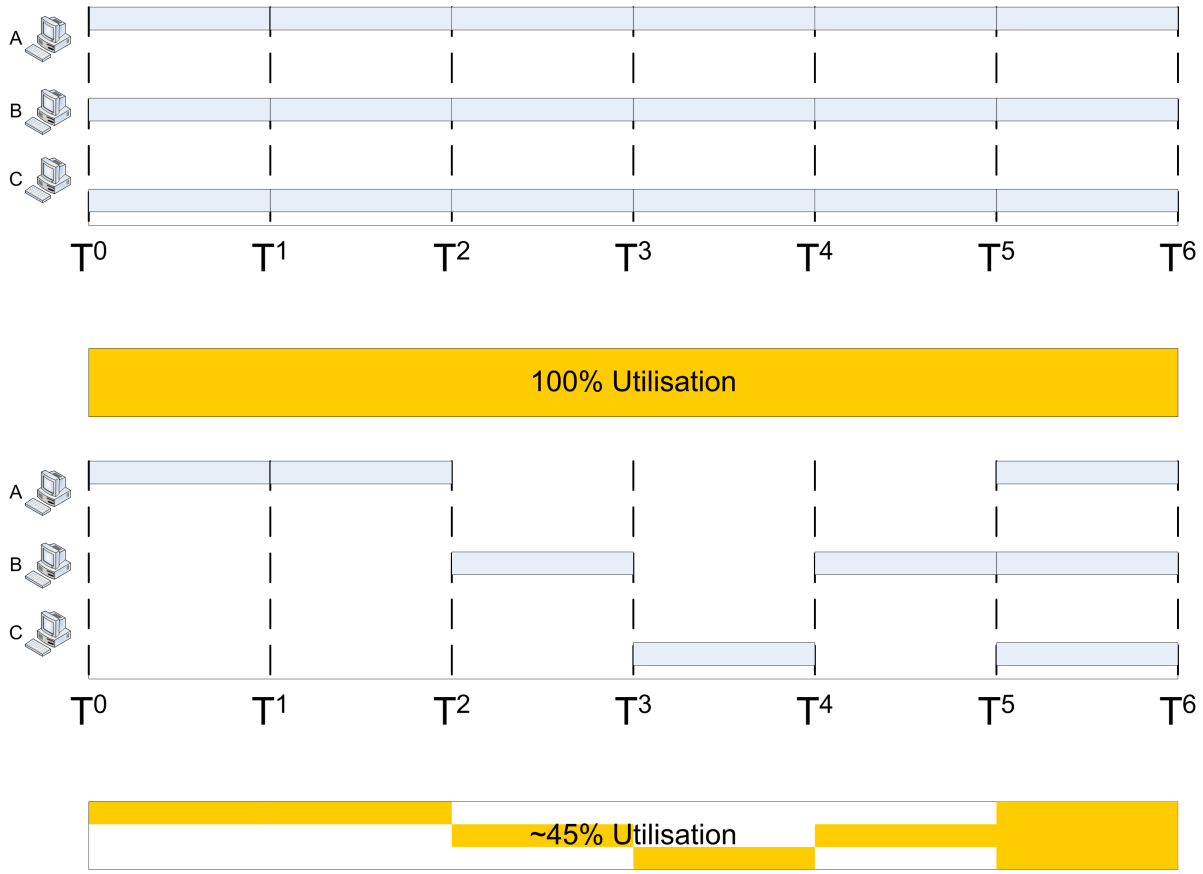
- ATMの出現
  - LANでの通信に適している
  - TDMAのように多重化を用いてパフォーマンスに焦点をあてている
  - 非同期通信モード

FDMA

- Frequency Division Multiple Access
  - チャネルスペクトラムをfrequency bandに分割
  - 各ステーションは固定frequency bandを持つ
  - frequency bandの使われていない時間はアイドル状態へ



FDMA (2)



通信の急増(バースト)のときはうまくいかない

OFDM

- Orthogonal Frequency Division Multiplexing
- 最新の高帯域デジタル通信はOFDMを利用している
  - 地上デジタルテレビ放送(日本、EU)
  - ADSL 高速モデム
  - 無線LAN (IEEE 802.11a/g/n)
  - WiMAX (IEEE 802.16d/e)
  - 4G

デジタル変調の原理

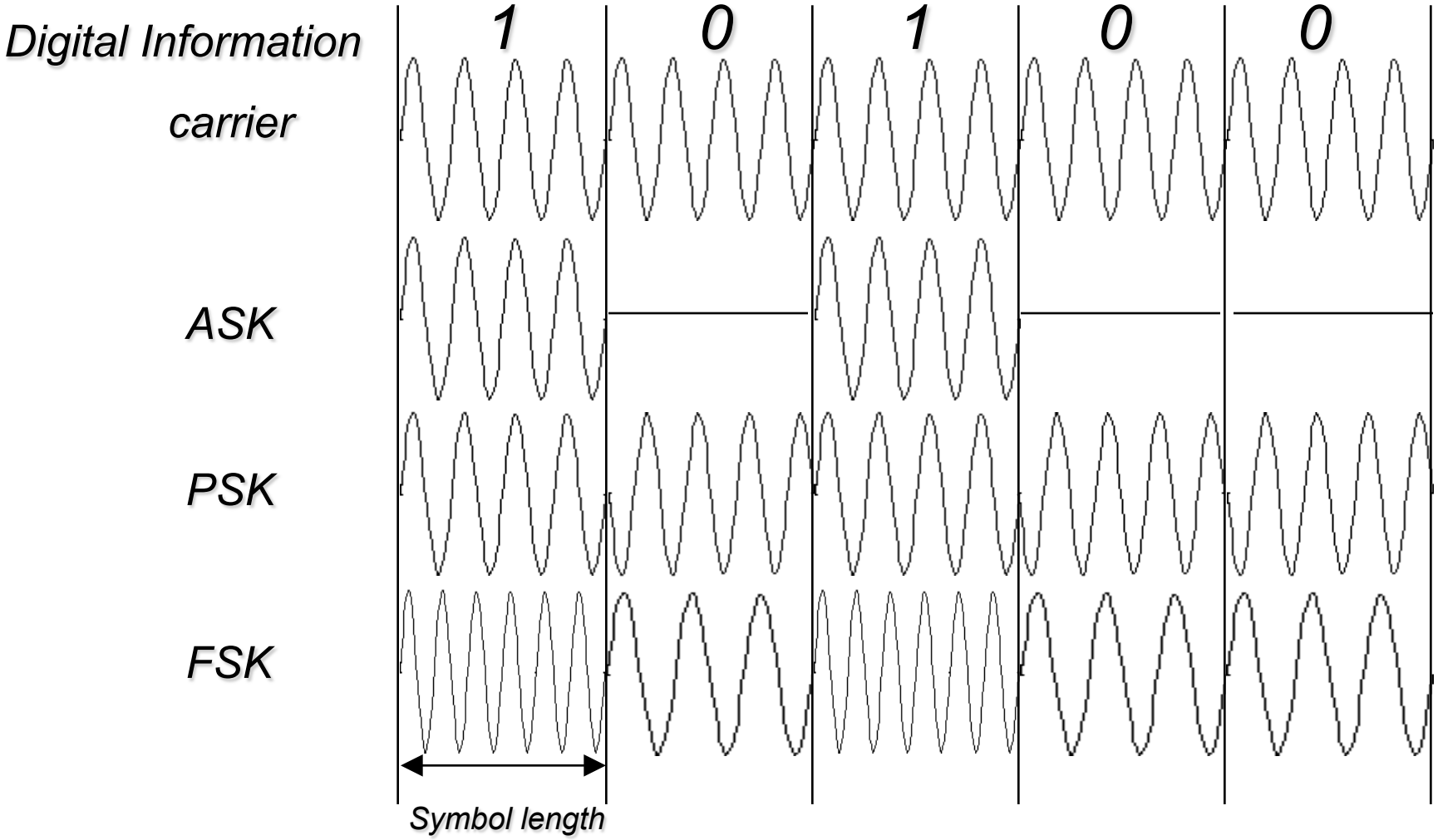
- デジタル変調 : 正弦波信号の3つのパラメータを調節
  - A, θ_k, f_c,

$$s(t) = A \cdot \cos(2\pi \cdot f_c \cdot t + \theta_k)$$

- 3種類のデジタル変調:
  - ASK : Amplitude Shift Keying
  - PSK : Phase Shift Keying
  - FSK : Frequency Shift Keying

- OFDM はASKとPSKを使用
  - Quadrature Amplitude Modulation (QAM)

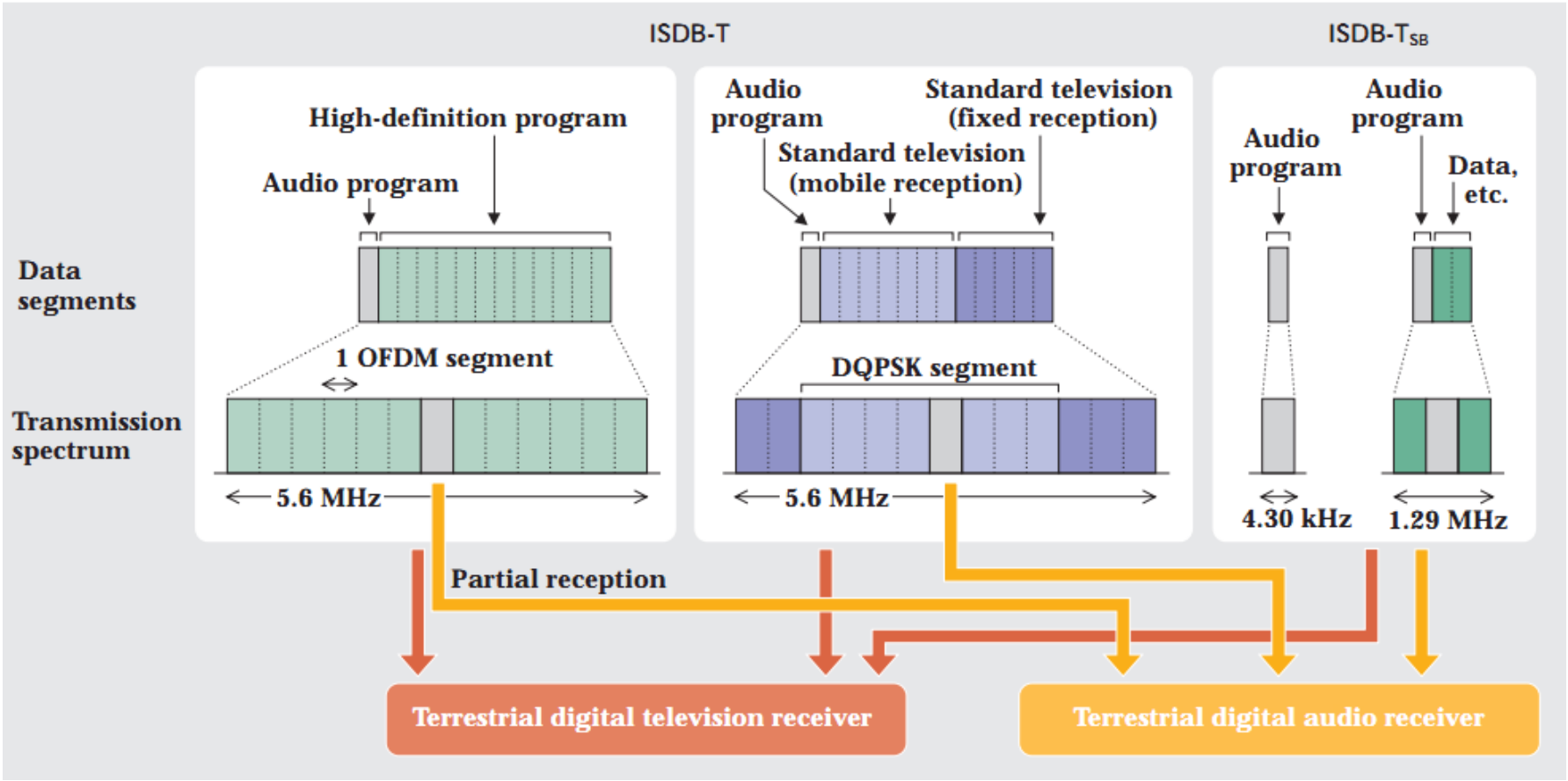
Symbol Waveform



日本の地上波デジタル放送 (1)

- Integrated Services Digital Broadcasting for Terrestrial Television Broadcasting (ISDB-T)
- 2006年に国内全体でサービス提供
  - 2011年:アナログ放送終了
- 5.6MHzの帯域を13のOFDMセグメントへ分割(~430 kHz 帯域)
  - HDTV or SDTV: 12 セグメント
  - Mobile TV : 1 セグメント

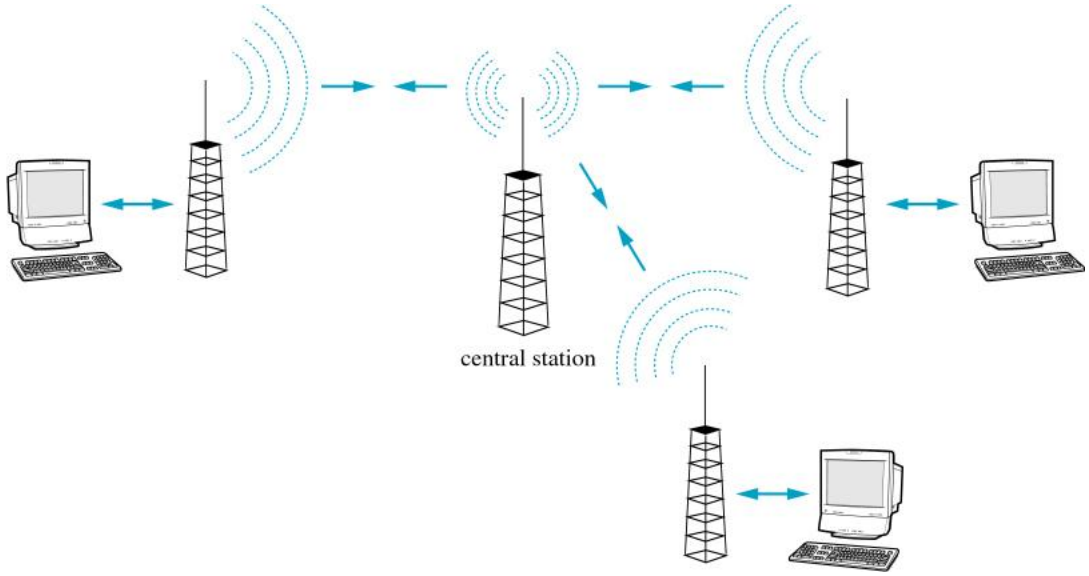
日本の地上波デジタル放送 (2)



NHK public paper: http://www.nhk.or.jp/strl/publica/bt/en/le0020.pdf

Random Access Protocols

- ノードが送信パケットを持つ時
  - 全チャネルで送信したときの伝送率R
  - 各ノードそれぞれへの優先割り当てはしない
  - 古くからの媒体共有問題：ラジオアンテナ
- 2つ以上の通信 ➜ “衝突”
- Random access MAC プロトコルの特徴:
  - 衝突検知の方法
  - 衝突回復の方法
- プロトコルの例:
  - ALOHA
  - Slotted ALOHA
  - CSMA
    - CSMA/CD
    - CSMA/CA

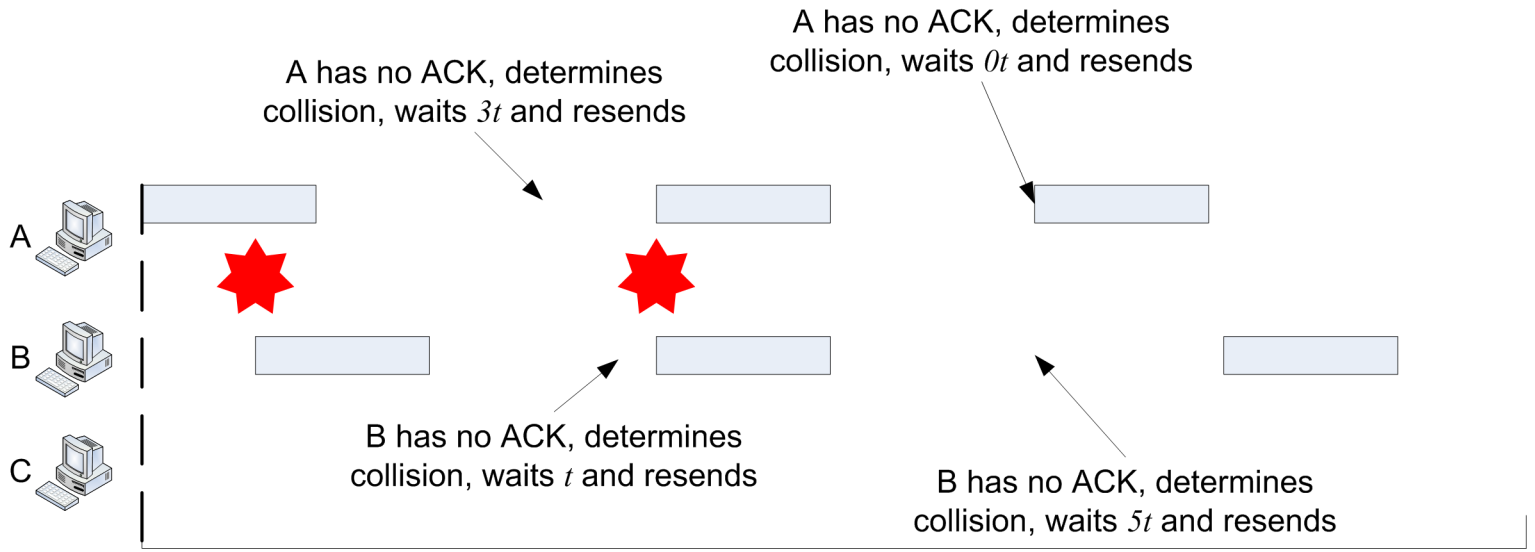


Pure ALOHA (1)

- 歴史
  - ハワイ大学の ALOHA systemで開発
  - 2つの無線回線
    - 端末-> ホスト (uplink)
    - ホスト-> 端末(downlink)
  - 2つの無線回線の共有方式を開発
- 方式:
  - 各端末は独立に動作
  - 端末は送信データがあればすぐに送信
  - 受信側はACKを発信
    - ACK自体は衝突することはない
    - ACKがRTT (Round Trip Time)時間内に帰ってこなければ、パケット衝突の可能性がある
  - バックオフ処理
    - [0,K] で乱数 r を発生
    - r 秒後に再送を行う

Pure ALOHA (2)

- 衝突
  - ランダム時間の待機後、再送
  - ランダム時間をどのように割り当てるかが性能の上で重要



Pure ALOHA (3) – 弱点

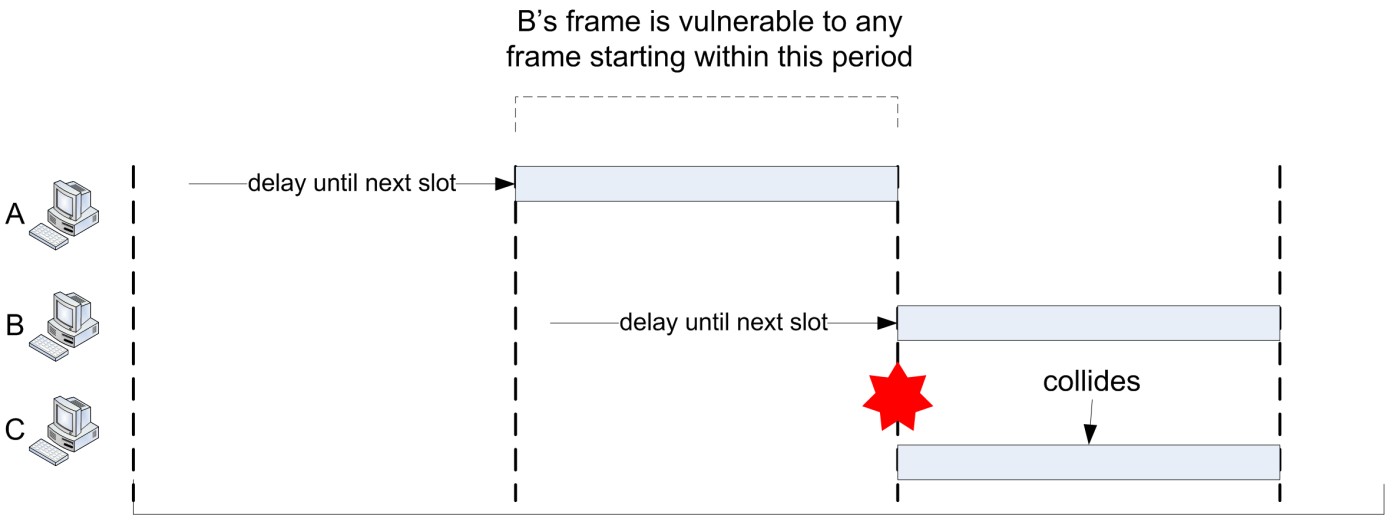
- 各端末が自立的にメディアにアクセス
  - 無秩序なアクセス方式のため、性能は悪い
  - 秩序はシステムの複雑化に繋がる
- 衝突の多発は性能劣化を引き起こす
  - 端末数の増加
  - 通信量の増加
- バックオフ時間
  - K
  - 性能に影響を与えるパラメータ

Slotted ALOHA (1)

- Slotted ALOHA
  - pure ALOHAの改良
  - 通信時間をタイムスロットに分割
  - 端末からの通信はタイムスロットに従わせる
  - “systematic order”
  - 回線アクセスの衝突をスロットの競合に
  - 競合回数を少なくする
  - 性能改善 (throughput 18.4%→36.8%)

Slotted ALOHA (2)

- Basic idea:
  - 全フレームを同じサイズに
  - 時間を等サイズのスロットに分割
  - 各ノードは各スロットの開始時刻にのみフレーム転送を開始できる
  - 各ノードは同期している
  - 2つ以上のノードが同じスロットで転送を開始した時に衝突を検知
  - 次のスロット開始までフレームを遅らせる



Pure ALOHA vs Slotted ALOHA

- どちらも、衝突後はランダム機能を使う
- Slotted ALOHAは衝突確率を下げる
  - フレーム送信準備が整った時:
    - Pure ALOHAではすぐに送信され、別のノードがその時に送信していたフレームを破壊するかもしれない
    - Slotted ALOHAでは、他の待機フレームを破壊するかもしれない

CSMA (1)

- Carrier Sense Multiple Access
- 送信する前にメディアに対し現在使用中かどうか尋ねる(キャリアセンス)
  - 使われていなければ直ちに送信
  - 伝送遅延のため、衝突が発生するかもしれない
  - RTT/2時間の間に他の端末が送信すると衝突が起こる
    - 伝送媒体全体に送信状態が電波する時間

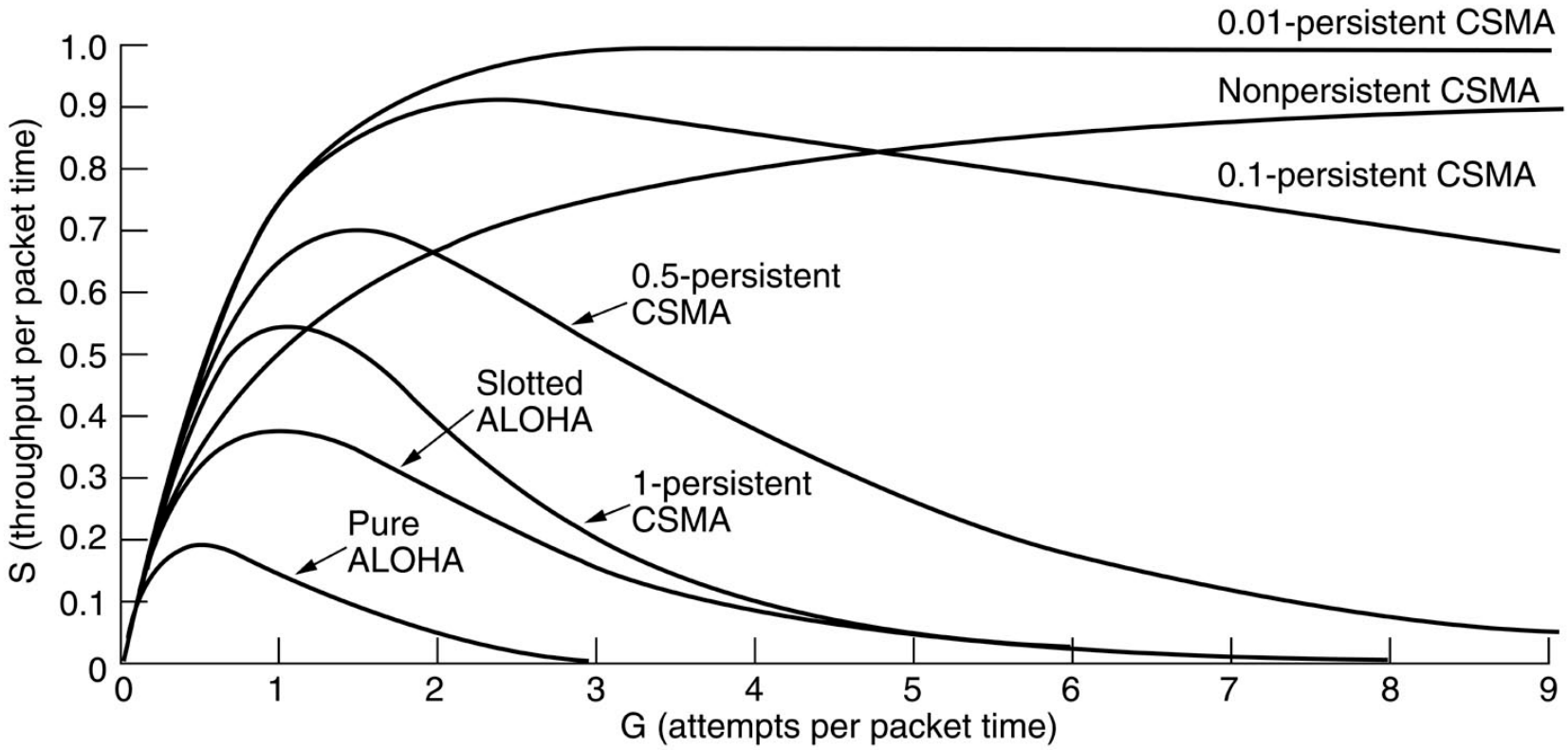
CSMA (2)

- キャリア(搬送電波)を検出した場合
  - 1-persistent
    - channel free: ただちに送信
    - channel busy: 1回は送信
  - non-persistent
    - channel free: 直ちに送信
    - channel busy: ランダム時間後に送信
  - (slotted) p-persistent CSMA
    - channel free: 確率pで送信
    - channel busy:次のスロットでアルゴリズムを繰り返す
- 衝突発生時
  - ランダム時間待機後、再びメディアに尋ねる

CSMA (3)

- CSMAは2つの特色を持つ
  - Collision Detection (/CD)
    - データフレームが衝突したことを確認できる
    - Ethernet 802.3プロトコルファミリーで使用
  - Collision Avoidance (/CA)
    - 特別なフレームを利用してチャネルを予約
    - データフレームは衝突しない
    - 無線LAN 802.11 プロトコルファミリーで使用

MAC Throughput Performance



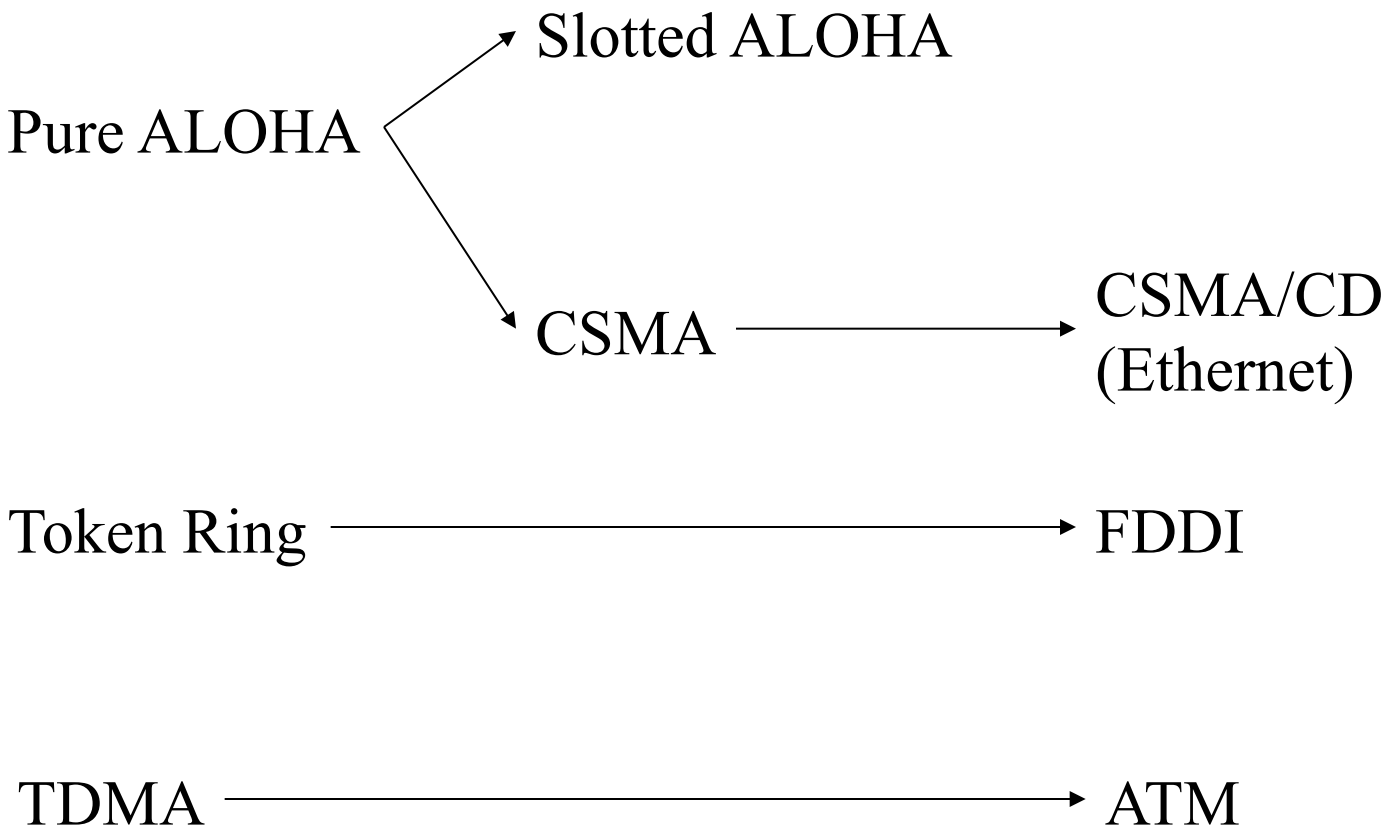
Taking Turn MAC Protocols

- MACプロトコルのチャネル分割:
  - 高負荷でのチャネル効率性や平等性を共有
  - 低負荷では非効率:
    - チャネルアクセス遅延
    - 1つのアクティブノードに対し1/N の帯域を割当
- Random access MAC protocols
  - 低負荷では効率的:
    - 1つのノードが最大限にチャネルを使用できる
  - 高負荷:
    - 衝突時のオーバヘッド
- “Taking turns” プロトコル
  - 2つの方法:Polling & Token

Token Passing

- 802.4 (Token Bus), 802.5 (Token Ring)で利用
- FDDIもToken Passingの一つ (実際はもっと複雑)
- 方法
  - 一定周期でノード間をトークンが移動
  - トークンを所持しているノードは通信可能
  - トークン保持タイマ
- 特徴
  - 他のデータリンク方式と違い、高負荷であっても性能が低下しない

Historical transition



Ethernet (IEEE 802.3 Family)

Ethernet

- CSMA/CDを利用したLAN技術
- IEEE802.3で標準化
  - 10Mbps, base-band transmission
  - 1500octet MTU
  - LANのバスタイプとして設計された
    - 同軸ケーブルとトランシーバ
    - 10Base5, 10Base2
- UTPケーブルを利用した10BaseT技術からの脱却
  - スター型構造
  - ハブやメッシュの誕生

Fast Ethernet

- 100Mbps Ethernet
  - 媒体はUTP/CAT5
  - 10 Mbps Ethernet 環境からの直接切り替え
  - 双方向通信の100BaseT
    - 双方向通信のための広い帯域
  - 自動センシング / 自動ネゴシエーション
    - コンテンションタイプは自動決定
  - 10BaseT/100BaseT, Unidirectional / Bidirectional
    - 相互運用性

Gigabit Ethernet (IEEE802.3z)

- 1Gbps
- 双方向データチャネル技術
  - Half-channel is 1Gbps
  - Full duplex
- 通信媒体は光ファイバー
  - 1000Base-SX (短波), 1000Base-LX (長波)
- 今日の殆どの製品が IEEE802.3z 標準に従っている
  - 相互運用性は着実に進歩している
  - スイッチ機能付きNICが普及している

Gigabit Ethernet (IEEE802.3ab)

- New Gigabit Ethernet standard
  - 1000Base-T
  - UTP
  - 1Gbps Ethernet
  - Maximum 100m

Ethernet Family Strengths

- Fast Ethernetの高速な普及
  - 10BaseTからの簡単な切り替え
  - 安価な製品
  - Fast Ethernetで10BaseT networkを代用

- 特別なネットワーク管理が必要ない
  - Plug-and-play
  - ATMとは大きく違う部分

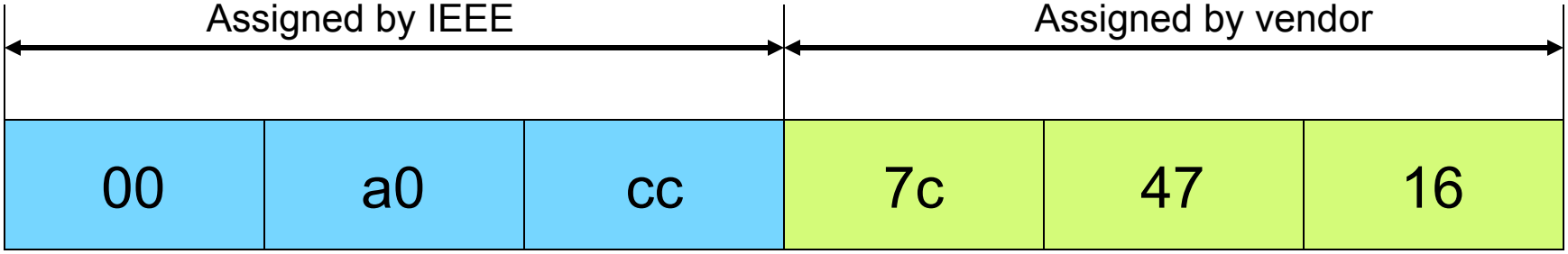
Ethernet Technologyの比較

	Ethernet	Fast Ethernet	Gigabit Ethernet
Data Rate	10Mbps	100Mbps	1Gbps
CAT 5 UTP	100m	100m	100m (CAT5+/6)
STP/Coax	500m	100m	N/A
Multi-mode Fiber	2Km	412m (hd*) 2Km (fd*)	550m
Single-mode Fiber	25Km	20Km	> 5Km

*IEEE half duplex/full duplex
note: IEEE 802.3z

Ethernet's address

- 6 オクテット (48 bit)
- 最初の3オクテットはベンダを識別(IEEEが定義)
- 次の3オクテットはベンダに割り当てられている
- IEEE MAC address
- “00:a0:cc:7c:47:16”



10Gbps Ethernet

- IEEE802.3ae
  - 2002年10月にIEEE が標準化
- WAN側とLAN側を分離
  - WAN PHY
    - SONET frameを使用
      - SONET/SDHの相互運用性を考慮
    - ～300m, wavelength: 850nm
    - 10km, wavelength: 1310nm
    - 40km, wavelength: 1550nm
  - LAN PHY
    - Ethernet Frameを使用
    - ～300m, wavelength: 850nm
    - 10km, wavelength: 1310nm
    - 40km, wavelength: 1550nm
    - 10km, wavelength: 1310nm
  - 2002年に市場に提供
    - 10GBASE-LR serial (1310nm)
    - 10GBASE-EX4(1550nm WWDM)
    - 通信距離e: 10Km – 50Km

40Gbps/100Gbps Ethernet

- IEEE802.3ba
  - 2009年7月にIEEEが定義

- 目的: 10Gbps以上の高速ネットワークをサポート
  - 全2重(full-duplex)通信のみをサポート
  - 現在の標準的なフレームサイズを保持
  - BER (Bit Error Rate) によって10^{-12}乗までの伝送品質をサポート
  - OTN（Optical Transport Network)をサポート

  - 40Gbps以上をサポートする物理層の仕様を提供:
    - SMF >10km
    - OM3 MMF >100m
    - Copper cable assembly >7m
    - Backplane >1m

  - 100Gbps以上をサポートする物理層の仕様を提供:
    - SMF >40km or >10km
    - OM3 MMF >100m
    - Copper cable assembly >7m

40Gbps Ethernet Single-mode Fibre PMD

- IEEE802.3bg
  - 2011年3月に標準化
  - PMD: Physical Medium Dependent
- 改善点:
  - MAC/PLSサービスインタフェースによってBERの10^{-12}以上をサポート
  - MACデータレートを40Gbpsをサポート
  - 40GBASE-R PCS と PMAを使用
  - IEEE 802.3baを更新したIEEE 802.3asの物理、論理インタフェースを使用
  - SMFにおいて最小2kmからの40Gbpsをサポートする物理層の仕様を提供
  - 現在の40Gbpsクライアントインタフェースにおける光学的互換性を提供 (OTU3/ STM-256/OC-768/40G POS)

Future of Ethernet

- Terabit Ethernet
  - P802.3bj が議論中
  - 100mまでの4x25Gマルチモードファーバーリンク
  - 2km以上のシングルモードファイバーリンクも構築中
- ネットワーク分割
  - “Switch agnostic” NIC Partitioning (NPAR)
  - 1つの物理ポートに4つの仮想ポート
    - Virtual ports: NIC, FCoE, or iSCI
- OpenFlow
  - research control planeを作成
  - 仮想ネットワークの準備
  - 高速ネットワークの導入

Requirements for the Data Link

Today's data link requirements

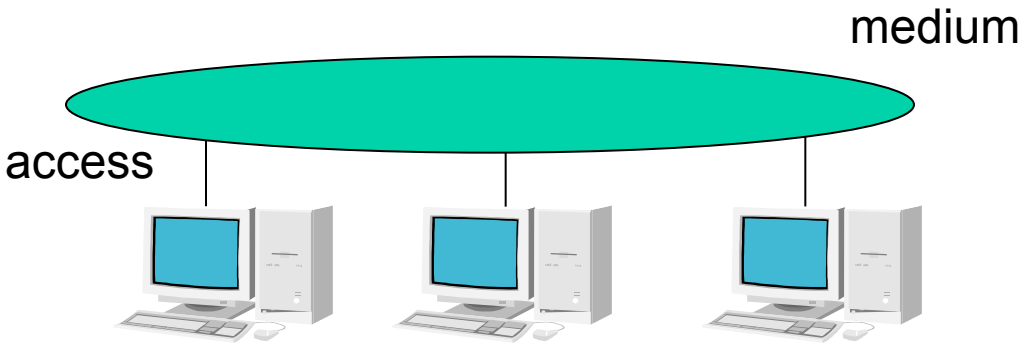
- Wide bandwidth
- Large scale
- Virtualization
- Coverage expansion

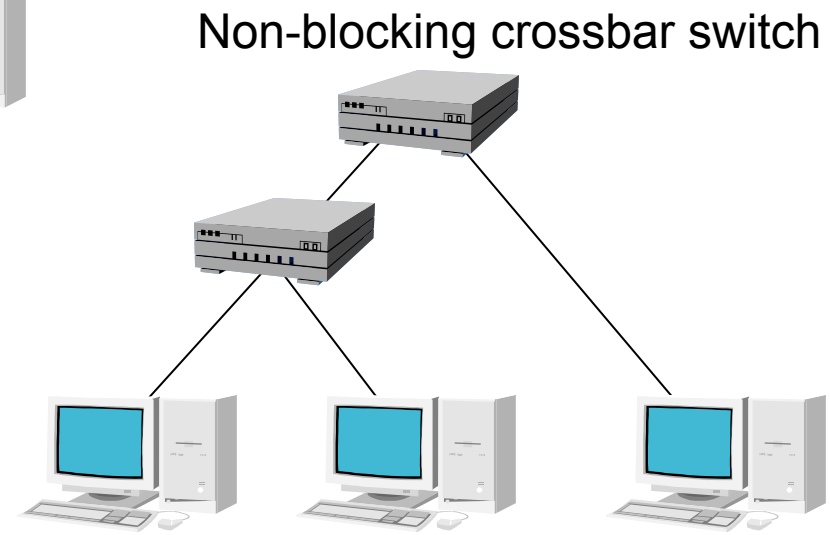
Development of data link technologies

- Wide bandwidth
- Large scale
- Virtualization
- Coverage expansion
- Switched media
- Bridges
- VLAN
- Broadband wireless, residential access, etc.

Wide bandwidth：shared media → switched media

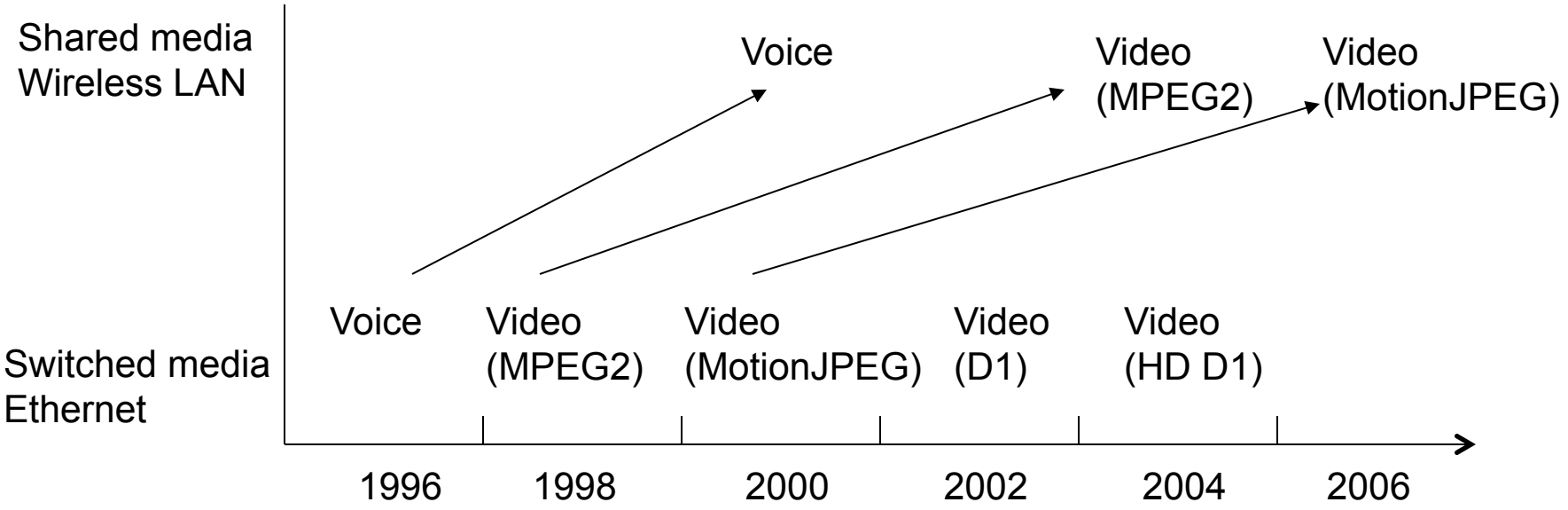
- High-bandwidthと commodity LAN
- Processor Interconnect Technologyの普及





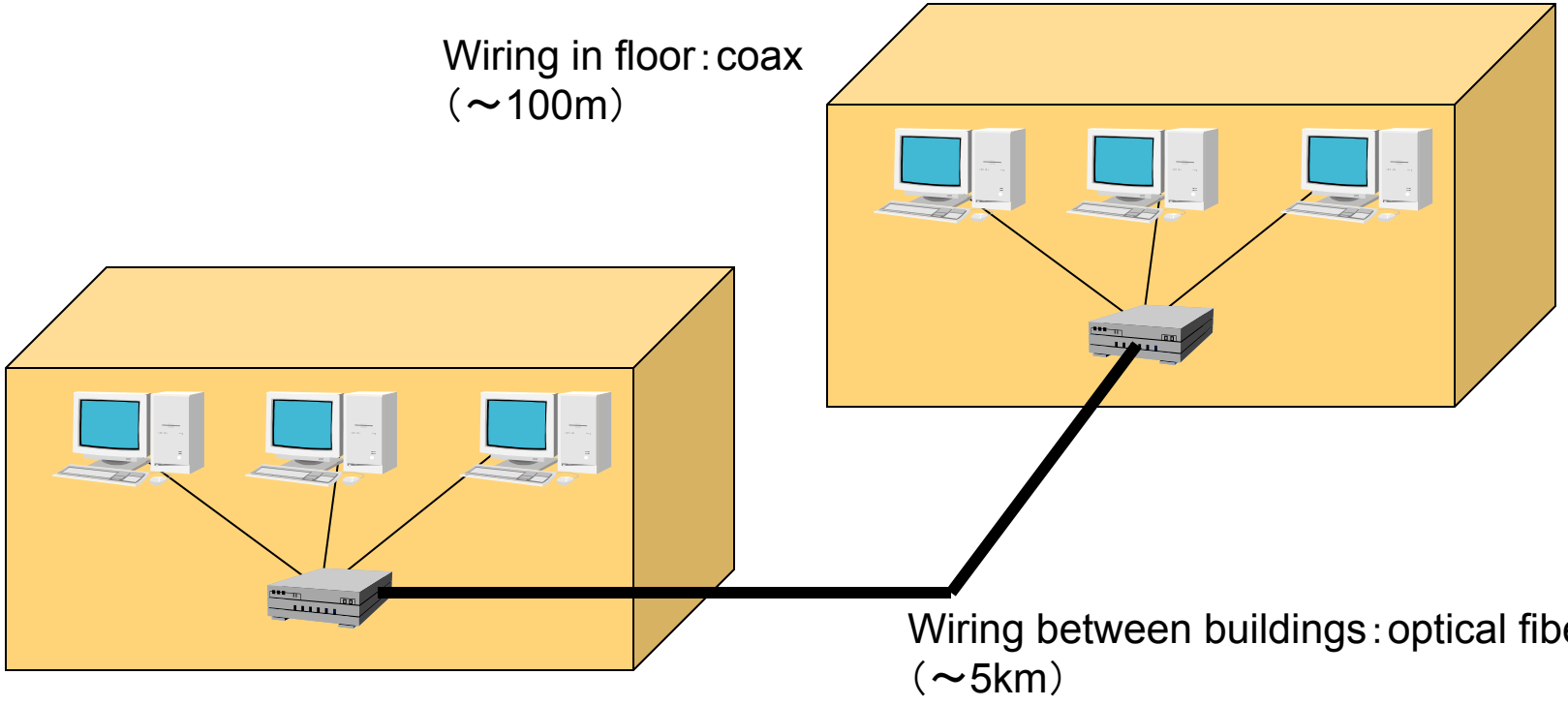
switched mediaへの移行

- switched media と shared mediaの混同は重大な問題
- デモの失敗



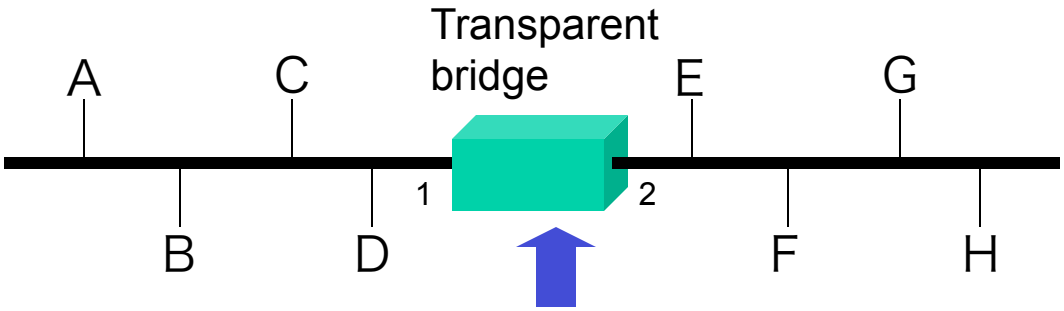
Large scale: Bridges

- 物理的な制約とLANの快適さとの互換性
  - 範囲,キャパシティ



Bridge basics: Transparent bridge

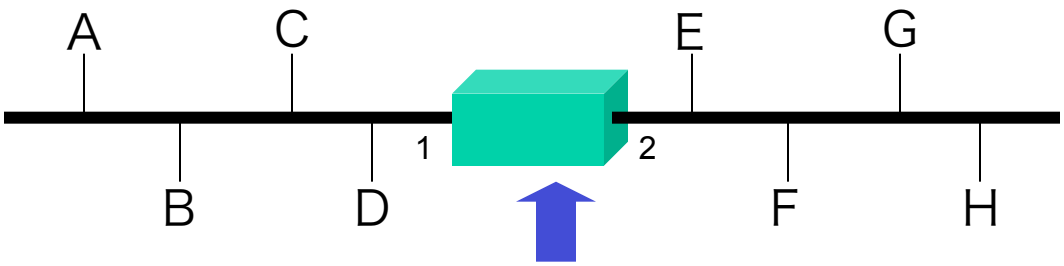
- ホストはブリッジを知らない
- Transparent bridge
  - MAC フレームは変更しない
  - Promiscuous: 全てのパケットをキャプチャ
  - 管理者はブリッジフォワーディングテーブルを設計



Fwd to 1	A, B, C, D
Fwd to 2	E, F, G, H

Learning bridge

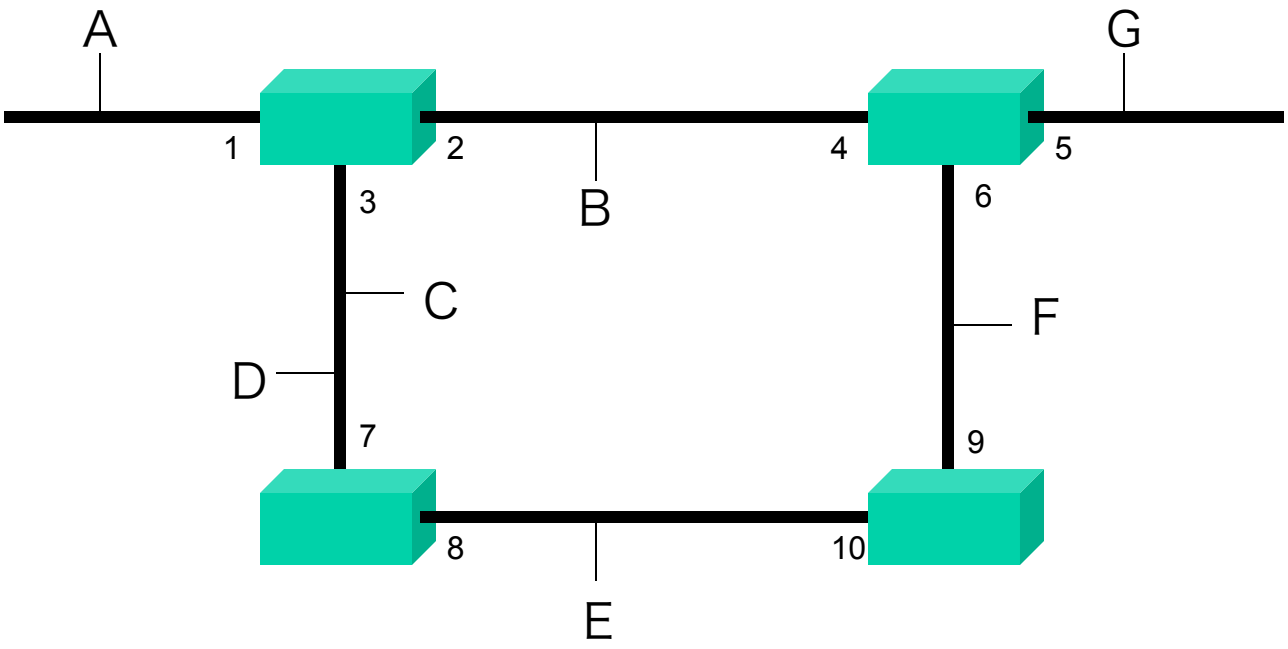
- 未知のMACアドレスが宛先のフレームを受信した時
  - 他のポートへ転送 (flooding)
  - 発信元ホストを記録 →フォワーディングテーブルへ記録
- 既知のMACアドレスが宛先のフレームを受信した時
  - 適切なポートに送信
- ブロードキャスト送信
  - 同じ媒体内の全アクティブホストが受信



Construct a table with learning

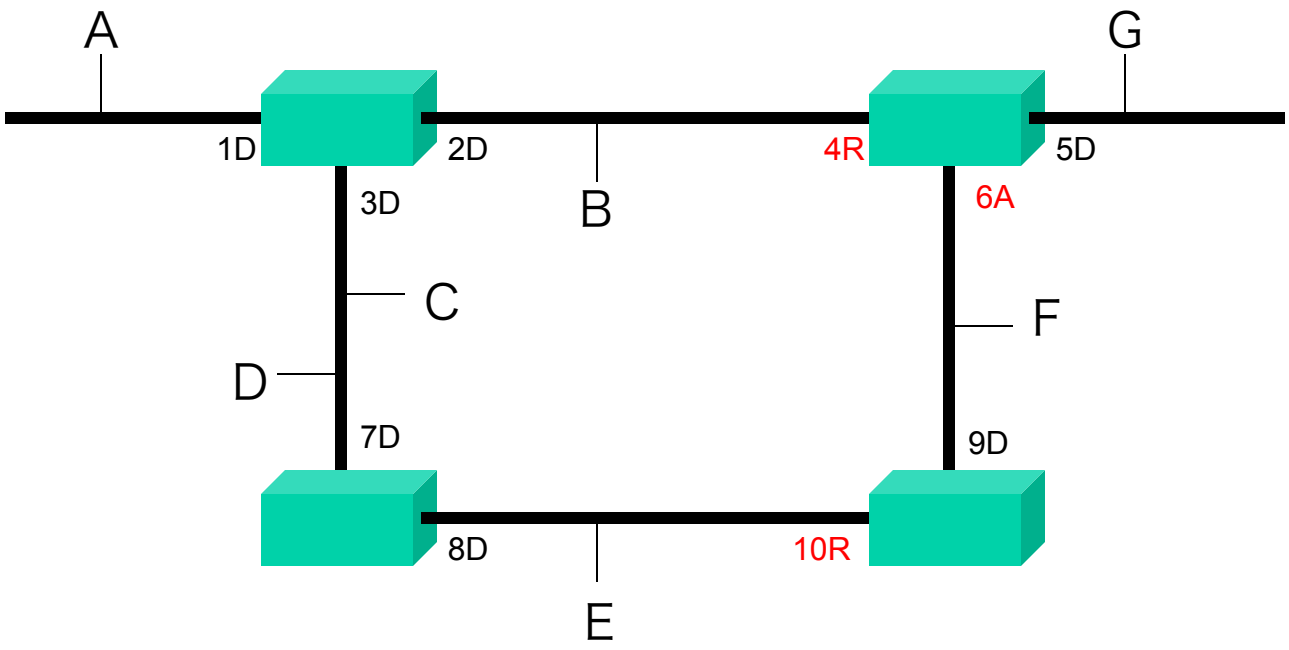
Loops in learning bridges

- 以前のアルゴリズムだと、AがG宛にフレームを送信するとどうなる？



Spanning tree bridges

- Spanning tree



R: root port
D: designated port
A: alternative port

More various technologies

- VLAN
  - ネットワーク仮想化
  - 今日、VLANはよく利用されている
  - 1つのメディアに複数のネットワークを配置
- WDM
- MPLS
- Broadband wireless, residential access, WiMAX, …

- 興味があったら自分で調べてみよう