

第 15 回輪講資料

「コンピュータ・ネットワーク」

第 4.3 節 イーサネット

岡崎雅大

2019 年 5 月 20 日

目次

1	はじめに	1
2	イーサネット	1
2.1	クラシック・イーサネットの物理層	1
2.2	クラシック・イーサネットの MAC 副層プロトコル	2
2.2.1	べき乗バックオフを用いた CSMA/CD	4
2.3	イーサネットの性能	4
2.4	スイッチ式イーサネット	5

1 はじめに

本稿では、教科書「コンピュータネットワーク」より、第 4 章 4.3 節 4.3.1 項から 4.3.4 項までの内容をまとめたものである。これまでに述べたチャネル割当プロトコルが、実際のシステムへどのように適用されているかについて述べていく。とくに本稿では世界でもっとも広く普及しているコンピュータネットワークであろう、イーサネットについて取り扱う。

2 イーサネット

イーサネットには クラシック・イーサネット (classic Ethernet) と スwitch式イーサネット (switched Ethernet) の 2 種類がある。クラシック・イーサネットはこの章で触れる技術を用いて、マルチ・アクセスにより生じる問題を解決している。スイッチ式イーサネットはコンピューター間をつなぐために「スイッチ」と呼ばれるデバイスを用いる。これらの両方が「イーサネット」と呼ばれるが、両者は大きく異なっている。クラシック・イーサネットはイーサネットのもともとの形態である。スイッチ式イーサネットはイーサネットの現在の形である。ここからは、イーサネットの歴史的な変遷をそれらがどのように開発されたと共に時系列に沿って論じていく。

2.1 クラシック・イーサネットの物理層

イーサネットの歴史は ALOHA の歴史と同時に始まった。Bob Metcalf という学生が M.I.T で学士号を取得し、博士号のコースをハーバード大学で歩み始めたときである。研究を進める中で、彼は Abramson の研究に出会った。その後、彼はハーバードを卒業し Xerox PARC (Palo Alto 研究所) で働き始めた。そこで彼が目にしたものは、後に「パーソナル・コンピューター」と呼ばれるもの、Xerox Alto であった。ただし、個々の機械は互いに独立していた。彼は、Abramson の研究に関する知識を用いて同僚の David Boggs とともに、はじめての LAN を設計・実装し

た。伝送媒体として同軸ケーブルが用いられ、速度は 3Mbps であった。彼らはかつて電磁放射を伝えると考えられていた発光性エーテルにちなんで、このシステムをイーサネット (Ethernet) と呼んだ。

Xerox のイーサネットが大成功したので、DEC と Intel、そして Xerox は **DIX 規格 (DIX Standard)** と呼ばれる 10Mbps の規格を制定した。この DIX 規格に小さな変更を加えることで、1983 年に IEEE802.3 という規格が作られた。Xerox は、すでにパーソナル・コンピュータのような独創的発明をしていたにもかかわらず、この規格の商業化に失敗した。Xerox が規格の標準化以外にほとんど興味を示さなかったため、Metcalfe はイーサネット・アダプターを販売する会社、3Com を設立し、アダプターを何百万個も販売した。

クラシック・イーサネットでは、一本の長いケーブルが建物中を這い回り、すべてのコンピュータが接続された。このアーキテクチャを図 4-13(p274) に示す。

- シック・イーサネット (thick Ethernet)
 - ケーブルは黄色い庭のホースに似ている (規格として色が指定されているわけではないが、黄色が推奨されている)
 - 2.5m ごとにコンピュータを接続する場所を示す印が示されている
 - 1 つのセグメントの長さは 500m 以内
 - 接続台数は 100 台以内
- シン・イーサネット (thin Ethernet)
 - より曲げやすく、工業規格である BNC コネクタを用いる
 - 安価かつ導入が安易である
 - 1 つのセグメントの長さは 185m 以内
 - 接続台数は 30 台以内

どちらのイーサネットを用いても信号伝搬距離からセグメントあたりの長さの上限が決まっているため、より大規模なネットワークを構築するためには、複数のケーブルをリピーター (repeater) と呼ばれる装置で接続する必要がある。リピーターは、信号を受信・増幅し、その信号を機器の両方向に返す物理層のデバイスである。ソフトウェアから見ると、リピーターで接続した複数のケーブルは一本のケーブルと同じように見える。

これらのケーブルにより、情報はマンチェスタ符号化を用いて送信される (マンチェスタ符号化については 2.5 節を参照)。イーサネットは複数のケーブルで構成することができるが、どのトランシーバー間も 2.5km 以上離すことはできないほか、2 つのトランシーバー間に 4 つ以上のリピータを挟むこともできない。これらの制限は MAC プロトコルを正常に動作させるためのものである。

2.2 クラシック・イーサネットの MAC 副層プロトコル

送信フレームのフォーマットを図 4-14(p274) に示す。

- プリアンブル (8 バイト)

1~7 バイトはいずれも 10101010 というビットパターンを持つ。最終バイトだけは最後の 2 ビットが 11 となる。この最終バイトは IEEE802.3 でフレーム開始区切り (Start of Frame delimiter) と呼ばれている。このパターンを符号化すると長さ 6.4 μ s の 10 MHz の矩形波となり、これによって送受信者間のクロックを同期させる。また、最後の 2 ビットによって残りのフレームを送信することを受信者に伝える。
- あて先アドレス、送信元アドレス (6 バイト)

あて先アドレスで最初に送られるアドレスは、通常のアドレスの場合 0、グループ・アドレスの場合 1 である。グループ・アドレスにより、1 つの場所から送られたものを複数の端末で受信できるようになる。フレームがグループ・アドレス宛てに送られると、グループ内のすべての局がそのフレームを受信する。

 - マルチキャスト (multicasting)

グループ・アドレスに送信することをマルチキャストという。ブロードキャストに比べて選択的であるが、各局とグループの対応の管理が必要となる。

- ブロードキャスト (broadcasting)

特別なアドレスとして、あて先アドレスの全ビットがすべて 1 のものはブロードキャスト用として予約されている。このフレームは、ネットワーク上のすべての局に送信される。局同士を区別する必要がないため、グループ分けの管理も不要である。

局が持つ送信元アドレスは、IEEE によってどの局間でも重複が起こらないように割り振られており、全世界で一意である。アドレスの最初の 3 バイトは **OUI(Organizationally Unique Identifier)** に用いられている。この値は IEEE により割り振られ、メーカーを示す。メーカーは残りの 3 バイトを自ら割り振り、NIC に登録する。

- タイプ、長さ (2 バイト)

フレームがイーサネットか IEEE802.3 であるかに依存して内容が異なる。

- イーサネット：タイプ・フィールド

受信者にこのフレームをどのプロセスに渡すべきかを明示する。しかし、これではデータの中身を見なければフレームの長さがわからないため、階層化の決まりに反してしまう。

- IEEE802.3：長さフィールド

フレーム長を保持している。受信者にとってフレームをどう処理すべきか判断する方法は、**LLC (Logical Link Control：論理リンク制御)** プロトコル用の別のヘッダーをデータ内に追加することで提供される。

- データ (0-1500 バイト)、パッド (0-46 バイト)

データは 1500 バイト以下である。これは、イーサネット標準が制定されたとき、トランシーバーがフレームを格納するのに十分な RAM を必要とし、かつ、当時 RAM は高価であったことに基づいてある程度恣意的に選ばれたものである。最大フレーム長に加え、最小フレーム長も存在する。イーサネットでは、正しいフレームとそうでないフレームの区別を容易にするため、正しいフレームはあて先アドレスからチェックサムまで少なくとも 64 バイトの長さを持つ必要があると定めている。よって、フレームのデータ部分が 46 バイトより短い場合、フレームの長さを広げるためにパッドフィールドが用いられる。

- 最小フレーム長

- * 衝突を検知したとき、送信中のフレームの送信は打ち切られるため、バラバラとなったビットとフレームのかけらがいくつもケーブル上に存在してしまう。もしデータが極端に短いフレームが送信されると、これらと見分けがつきにくい。
- * 送信した最初のビットがケーブルの終端へ到着する前にフレームの送信が終了してしまうと、送信側が衝突を検知する前に送信が正常に終了したと誤認する恐れがある。そのため、衝突を検知したときにフレームを送信しておくためには、送信時間を終端までの遅延時間の 2 倍以上とする必要がある。

- パッド

フレームが終端に到着する遅延時間を τ とする。上記より、すべてのフレームは送信に 2τ 以上の時間をかけなければならない。2.5 km の最長距離と 4 つのリピータを持つ約 10M ビット/秒の LAN では、リピータの遅延時間を含めラウンド・トリップ時間は最悪の場合で約 50 μ s と算出される。1 ビットの転送に 100 ns かかるため、最小のフレーム長は 500 ビットとなる。いくらかの余裕を付加して、512 ビット (64 バイト) に丸められる。データフィールドまでのフィールドの長さを合計すると、22 バイトであるため、差し引き 46 バイト必要である。よって、データフィールドが 46 バイト未満であるとき、不足分をパッドフィールドで埋め合わせる。

- チェックサム (4 バイト)

フレームのビットが正しく受信されたか判断するための誤り検出コードとして用いられている。誤りが検出された場合、そのフレームを破棄する。32 ビット CRC であり、生成多項式を用いて定義されている (詳細は 3.2 節を参照)。

2.2.1 べき乗バックオフを用いた CSMA/CD

クラシック・イーサネットは 1-永続 CSMA アルゴリズムを用いている。送信すべきフレームを持っている局が媒体を監視し、媒体が使用されなくなるとすぐにフレームを送信する。局は送信時にチャネル上の衝突を監視する。衝突が発生した場合、局は短いジャム信号を発して送信を中止し、ランダムな時間間隔のあとに再送信を行う。衝突が発生したとき、ランダムな時間間隔をどのように決定するかを考える。

- 環境

図 4-5(p262) のモデルを用いる。衝突が発生すると、時間は離散的なスロットに分割される。スロットの長さは、イーサネットの最長往復伝搬時間 (2τ) である。よって、スロット時間は 512 ビット時間すなわち $51.2\mu\text{s}$ である。

- 流れ

1. 衝突が発生すると、局はフレームの送信を中止する。
2. それぞれの局は 0 または 1 スロット時間の中から待ち時間をランダムに選択し、待機したあと再送信する。
3. 双方が同じ待ち時間を選択した場合、再び衝突する。
4. それぞれの局は 0,1,2,3 の中から数値をランダムに選択し、そのスロット時間分待機する。
5. 3 回目の衝突が発生した場合、0 から $2^3 - 1$ の中から数値を選択し、そのスロット時間分待機する。
6. i 回目の衝突が発生した場合、0 から $2^i - 1$ の中から数値を選択し、そのスロット時間分待機する。しかし、10 回以上の衝突が発生した場合、数値の最大値は 1023 に固定される。
7. 16 回連続で衝突が発生した場合、送信を中断し、失敗を通知する。その後の回復は上位層に任される。

このアルゴリズムを べき乗バックオフ (binary exponential backoff) と呼ぶ。送信を行おうとしている局に対して、待ち時間を動的に適応させる事ができる。連続した衝突が起こるにつれて、ランダムに選ぶ数の範囲を指数関数的に増加させることで、衝突を起こしている局の数が少ないときには短い遅延で処理され、多くの局が衝突しているときには衝突が現実的な時間で解決されることを保証している。また、バックオフを 1023 で打ち切ることで、待ち時間が大きくなりすぎることを防いでいる。

衝突が発生しなかった場合、送信者はフレームが相手へ伝送されたものとみなす。すなわち、CSMA/CD もイーサネットも受信確認を行わない。これは、誤り率の低い有線チャネルや光ファイバチャネルに適している。よって、いかなる誤りも CRC によって検出し、上位層で回復しなければならない。

2.3 イーサネットの性能

クラシック・イーサネットの性能について考える。べき乗バックオフ・アルゴリズムの完全な解析は非常に複雑であるため、以下の仮定を行う。

- k 個の局が常に送信可能な状態にあり、一定の高い負荷が与えられている。
- 各スロットにおいて再送確立は一定である

各局が確立 p で競合しているスロットに転送すると、いずれかの局がそのスロットにチャネルを獲得する確率 A は次のようになる。

$$A = kp(1-p)^{k-1} \quad (1)$$

A は $p = 1/k$ のとき最大になり、 $k \rightarrow \infty$ で $A \rightarrow 1/e$ (e は自然対数の底) となる。競合期間がちょうど j スロットである確率は $A(1-A)^{j-1}$ であるため、1つの競合あたりの平均スロット数は次の式で与えられる。

$$\sum_{j=0}^{\infty} jA(1-A)^{j-1} = \frac{1}{A} \quad (2)$$

各スロットは、 2τ 時間待つので、平均競合期間 ω は、 $2\tau/A$ である。最適な再送確率 p を設定すると、競合スロットの平均数は e を超えない。したがって ω はせいぜい $2\tau e = 5.4\tau$ である。

平均的なフレームを転送するのに P 秒かかるとすると、多くの局がフレームを送信したい場合

$$\text{チャネル効率} = \frac{P}{P + 2\tau/A} \quad (3)$$

となる。ここで、2つの局間の最大ケーブル距離がどのように性能に影響するかを考える。ケーブルが長くなるにつれて、競合期間も長くなることから、なぜイーサネットの標準が最大ケーブル長を仕様として定めているかが推察できる。

式をフレームあたりの最適な競合スロット数 e のときについて、フレーム長 F 、ネットワークの帯域幅 B 、ケーブルの長さ L 、信号の伝達速度 c を用いて書き換えると、 $P = F/B$ であるから、

$$\text{チャネル効率} = \frac{1}{1 + 2BLE/cF} \quad (4)$$

となる。分母の第2項が大きくなると、ネットワークの効率は下がる。具体的には、ネットワークの帯域幅かケーブルの長さ (B と L の積) が増すと、与えられたフレーム長に対するネットワーク効率が下がってしまう。クラシック・イーサネットは広い帯域幅を持った長距離の通信に最適であるとはいえない。

2.4 スイッチ式イーサネット

イーサネットは1本の長いケーブルによって構成されるクラシック・イーサネットとは異なる進化をした。断線やゆるい接続を探すことの煩わしさを解消するため、それぞれの局を専用のケーブルで中央のハブ (**Hub**) へ接続するという配線パターンが考え出された。ハブは接続されたすべての通信線を電氣的に接続する。この構成を図 4-17(a)(p280) に示す。

- ハブ
 - 配線：電話会社のツイスト・ペア・ケーブル
 - * 多くの建物がすでにこのケーブルで配線されていた。
 - * 予備のケーブルも豊富であった。
 - ケーブルの最大長はハブから 100m (高品質なカテゴリ 5 を使用した場合 200m)。
 - 局の追加や撤去が簡単となり、ケーブルの断線も容易に発見できるようになった。
 - 論理的には1本の長いケーブルで接続する従来のクラシック・イーサネットと等価であるため、通信容量を増加させることはできない。局が増えるにつれ1台あたりの容量は減少し、最終的には LAN が飽和してしまう。

増大する負荷に対応する別の解決法は、スイッチ式イーサネット (switched Ethernet) である。このシステムの中心は、図 4-17(b)(p280) のように、すべてのポートをつなぐ高速なバックプレーンを備えた **switch** (スイッチ) である。

- スイッチ
 - 配線の抜き差しによって簡単に局の追加や撤去ができる。
 - 個々のケーブルやポートは、たいてい1つの局にのみ影響を及ぼすため、故障箇所を用意に発見できる。
 - ハブとは異なり、フレームのあて先になっているポートにしか出力しない。
 - * ポートがフレームを受け取ると、スイッチはフレームのイーサネット・アドレスを調べ、あて先ホストがどのポートに接続されているかを判断する。そのため、スイッチはポートとアドレスの対応を把握しておく必要がある。
 - * その他のポートにはフレームが存在することすらわからない。

2つ以上の局、あるいはポートが同時にフレームを送信しようとした場合、どのようなことが起こるだろうか。ハブではすべての局が同一の衝突ドメイン (**collision domain**) を持ち、転送のスケジューリングに CSMA/CD を用いる。スイッチでは各ポートは各自の衝突ドメインを持つ。一般に全二重通信では衝突が起らず、CSMA/CD も必要ないため、局とポートの両方が他の影響を受けずに通信線へ同時にフレームを送信できる。ただし、半二重通信の場合は局とポートが CSMA/CD を用いて転送権を争う。

- ハブに対するスイッチのメリット

- 衝突が起らないため、回線をより効率的に使用することができる。
- 異なる局から複数のフレームを同時に送信することができる。
- 以上よりハブに比べて桁違いの性能を得ることができる。

スイッチはセキュリティ面においてもメリットをもたらした。大抵の LAN インタフェースには プロミスキューモード (**promiscuous mode**) がある。これは、送信先に指定した端末だけでなく、すべての端末がフレームを受け取ることができる。ハブを用いると、接続されたすべての端末から、他のすべての端末のトラフィックを監視することができる。一方、スイッチを用いるとトラフィックは目的のポートにしか転送されない。この制限のおかげで通信は独立し、トラフィックが漏れたり、誤った相手に届くといったことが起こりにくくなる。しかし、本当にセキュリティが必要な場合はトラフィックを暗号化することが望ましい。

スイッチは各入力ポート上でイーサネット標準のフレームを待っているだけであるので、いくつかのポートをコンセントレータ（集約器）として使うことができる。図 4-18(p281) で上部右隅のポートは1つの局に接続しているのではなく、12ポートのハブに接続している。フレームがハブに到着すると、それらは衝突やベキ乗バックオフを含む通常の方法でイーサネット・フレームとして競合する。競合に勝ったフレームはハブを経由してスイッチに送られ、他のフレームと同様に処理される。一度スイッチに入ると、それらは正しい出力線に転送される。なお、送信先がハブで接続された回線上にいてもよく、その場合フレームはポートにすでに届いているのでスイッチはそれを渡すだけでよい。ハブはスイッチより単純でかつ安価である。