分散システム特論レポート

RFC792 Internet Control Message Protocol (ICMP)
RFC2463 Internet Control Message Protocol (ICMPv6)
for the Internet Protocol Version 6(IPv6) Specification

情報工学専攻 修士 1 年 荒木研究室 柴田 賢介

1 はじめに

ICMP は、IP(Internet Protocol) をサポートするためのプロトコルである。データが送り先に届かないといったエラーの通知や、ルータがデータを転送するためのバッファが不足しているといったことの通知というように、エラーに対するフィードバックを提供することを目的として定められている。ICMP は IP の上位に位置するプロトコルと考えられることもあるが、実際には IP の一部であり、欠くことのできないプロトコルである。

次世代インターネットプロトコルとして IPv6(Internet Protocol Version 6) の研究開発が進められているが、これに伴い、IPv6 に対応する ICMP も定められている。これが ICMPv6 である。以下に、ICMP、ICMPv6 のメッセージフォーマット、特徴を述べ、ICMPv6 を利用したソフトウェアの例として traceroute6 を挙げ、動作の概要とソースコードの簡単な説明を行う。

2 ICMP

2.1 概要

IP は信頼性のない通信を行う。パケットが送り先に届くことは保証されておらず、たとえ送り先に届かない場合でも、パケットが再送されることはない。ICMP は先程述べたように、エラーに対するフィードバックを提供するためのプロトコルである。しかし、エラーに対するフィードバックが返される保証はなく、ICMP のレポートが行われないままデータの配信に失敗する場合もある。つまり、ICMP は IP を信頼性のあるものにするためのプロトコルではなく、信頼性はより高い層のプロトコルで保証されることになる。

ICMP は次節以降で述べるように、多くの種類のメッセージをやりとりするが、ICMP メッセージが届かないことに対する ICMP メッセージは送信されない。また、大きなサイズのパケットが送信される際にパケットを分割して送ることがあるが (IP フラグメント)、フラグメント化されたデータについてはオフセットが 0、つまり、分割されたもののうち、最初のパケットのみを扱うように定められている。

次節以降では、ICMP で用いられるメッセージについて述べることにする。

2.2 ICMP メッセージ

ICMP には、表 1 で示しているように、多くの種類のメッセージタイプがある。表中左の数字は、メッセージタイプを識別するための番号である。

表 1: ICMP メッセージ

タイプ	メッセージ
0	エコー応答
3	終点到達不可
4	始点抑制
5	リダイレクト
8	エコー要求
11	時間超過
12	パラメータ不良
13	タイムスタンプ要求
14	タイムスタンプ応答
15	情報要求
16	情報応答

2.2.1 終点到達不可 (タイプ=3)

まず最初に、メッセージタイプ 3 の終点到達不可メッセージについて説明する。終点到達不可メッセージのメッセージフォーマットは図 1 のようになる。

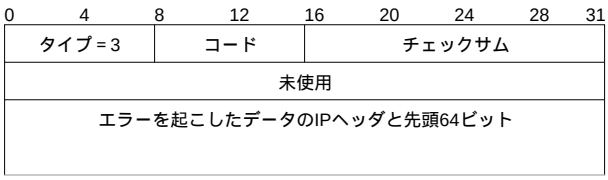

図 1: 終点到達不可メッセージのメッセージフォーマット

メッセージは、ヘッダ部とデータ部に分かれており、ヘッダ部にはタイプ、コード、チェックサムなどが含まれる。データ部には、エラーの対象となったパケットの IP ヘッダと、エラーの対象となったパケットのデータ部 64 ビット分が含まれる。データ部 64 ビットを含めているのは、この部分にポート番号が格納され

ているため、エラーパケットを送信したホストでプロセスの照合に使用できるからである。

ヘッダ部のメッセージタイプを示すビットの次に、「コード」という部分があるが、このコードは各メッセージの中でさらに種類を分けるために用いられる識別子である。終点到達不可メッセージは 6 種類のコードを持つ。これを示したのが表 2 である。

表 2: 終点到達不可メッセージ

コード	種類
0	ネットワーク到達不可
1	ホスト到達不可
2	プロトコル到達不可
3	ポート到達不可
4	要フラグメント化だが DF ビットがセット
5	始点経路指定失敗

このうちコード 0、1、4、5 のメッセージはルータから送られ、コード 2、3 のメッセージはホストから送られるものである。コード 4 のメッセージは、フラグメント化が必要であるにも関わらず、DF(Don't Fragment) ビットがセットされており、フラグメント化できない場合に送られるものである。コード 5 は、IP においてあらかじめデータグラムが通る経路を指定できる IP ソースルーティングオプションを使用した場合に、指定された経路の通りにデータグラムの転送ができない場合、メッセージを返すというものである。

2.2.2 時間超過 (タイプ=11)

次に、メッセージタイプ 11 の時間超過メッセージについて説明する。時間超過メッセージは表 3 のような 2 種類のコードを持つ。

表 3: 時間超過メッセージ

コード	種類
0	転送中のゼロ TTL による時間超過
1	フラグメントを再構成中の時間超過

コード 0 はルータがデータを転送する際に TTL(Time To Live) がゼロとなり、これ以上転送できないという場合に送信される。コード 1 はフラグメント化されたデータを再構成する際に、再構成のためのデータが揃っていないといった場合に送信される。

2.2.3 パラメータ不良 (タイプ=12)

このメッセージは、データのヘッダに含まれるパラメータに問題があった場合に送信される。ICMP メッセージのヘッダ部には、エラーが発見された箇所を示すポインタが含まれる。

2.2.4 始点抑制 (タイプ=4)

始点抑制メッセージは、ルータがキューに十分なバッファスペースを持たない場合、もしくは送信先ホストが届いたパケットを処理しきれなかった場合に、送信者に送出レートを下げるように促すために送信される。

2.2.5 リダイレクト (タイプ=5)

リダイレクトメッセージは、送信元ホストがデフォルトルートとして最適でない経路を使用しているのを検出した場合に送信される。メッセージには最適経路となるルータのアドレスを含む。但し、このメッセージが送信される対象は送信ホストのみであり、ルータには送られない。つまり、ルータの経路を変更するものではないということである。

2.2.6 エコー応答 (タイプ=0)、エコー要求 (タイプ=8)

このメッセージは、エコー要求が届いた場合に、エコー応答で返すというものである。ネットワークの到達性を調べるコマンドとして ping というものがあるが、これは、調べたいホストに対してエコー要求を出し、この要求に対して返るエコー応答から、相手のホストに到達できるかどうかを調べている。エコー応答では、エコー要求を出した送信元が決めた識別子とシーケンス番号をそのまま返す。

2.2.7 タイムスタンプ応答 (タイプ=14)、タイムスタンプ要求 (タイプ=13)

これらはエコー要求、エコー応答と似ているが、要求、応答にタイムスタンプが含まれる。タイムスタンプはグリニッジ標準時の午前 0 時からの時間を 64 ビット (単位 ms) で表したものである。メッセージには要求が送信された時のタイムスタンプと、要求を受信した時のタイムスタンプ、そして応答を送信するときのタイムスタンプが含まれる。また、エコー応答と同様に、応答メッセージは送信元が決めた識別子とシーケンス番号をそのまま返す。

2.2.8 情報要求 (タイプ=15)、情報応答 (タイプ=16)

これらのメッセージはホストが自分の IP アドレスが分からない場合に使用されるものである。送信元が IP ヘッダの送信元アドレスと送信先アドレスを 0 にして情報要求を行い、要求が届いたホストは IP アドレスを格納して応答を返す。しかし、現在はこのメッセージは廃止されており、使用されない。

3 ICMPv6

3.1 概要

次世代インターネットプロトコルとしてIPv6(Internet Protocol Version 6) の研究開発が進められている。IPv6 は、128 ビットの広大なアドレス空間を持ち、拡張ヘッダとして認証ヘッダやカプセル化セキュリティペイロードなどを含めることが可能なため、セキュリティへの対応が容易となっている。IPv6 の仕様は文献 [1] に定められている。これに伴い、IPv6 に対応する ICMP も定められている。これが ICMPv6 である。ICMPv6 は、ノードによってエラー通知のために使用される。IPv4 の場合と同様に、ICMPv6 は IPv6 には欠くことのできないプロトコルである。また、ICMPv6 は全ての IPV6 ノードに実装されなければならないと定められている。ICMPv6 メッセージのメッセージフォーマットは図 2 のようになる。

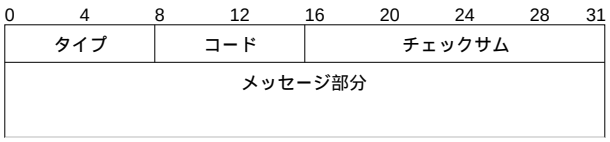

図 2: ICMPv6 メッセージのメッセージフォーマット

ここでは、主に ICMP と ICMPv6 とで異なる部分について説明する。

3.2 ICMPv6 メッセージ

ICMPv6 は、表 4 で示しているようなメッセージタイプを持つ。ICMPv6 メッセージは、大きく分けて 2 種類あり、1 つは ICMPv6 エラーメッセージ、もう 1 つが ICMPv6 情報メッセージである。表中左の数字は、メッセージタイプを識別するための番号である。

表 4: ICMPv6 メッセージ

タイプ	ICMPv6 エラーメッセージ
1	終点到達不可
2	パケット過大
3	時間超過
4	パラメータ不良
タイプ	ICMPv6 情報メッセージ
128	エコー要求
129	エコー応答

3.2.1 パケット過大 (タイプ=2)

ICMPv6 のメッセージには、ICMP にはなかったパケット過大メッセージがある。これは、パケットサイズが大きく、次ホップに転送できない場合に送信される。メッセージには次ホップの MTU(Maximum Transfer Unit：最大転送単位) を含める。このメッセージを受け取った送信元ホストでは、このメッセージを上位層に渡さなければならないことになっている。

3.3 メッセージ送信元アドレスの決定に関する記述

ICMPv6 が ICMP と異なる点として、ICMPv6 のメッセージを送信しようとしているノードが 2 つ以上のユニキャストアドレスを持っていた場合に、メッセージに含める送信元アドレスをどのようにして決定するかについて明確に述べているという点が挙げられる。以下に、このアドレスの決定方法について述べる。

ユニキャストアドレスに対して送信されたパケットに対するメッセージ　この場合は、送信元アドレスはそのユニキャストアドレスとする。

マルチキャスト、エニキャストアドレスに対して送信されたパケットに対するメッセージ　この場合、パケットを受信したインタフェースのユニキャストアドレスを送信元アドレスとしてメッセージを送信する。

ノードに属さないユニキャストアドレスに送信されたパケットに対するメッセージ　この場合は、エラーを診断するのに最も効果的なアドレスを送信元としてメッセージを送信する。例えば、パケットの転送を完了できなかった場合のエラーメッセージには、送信元アドレスとして、パケットの転送に失敗したインタフェースについているユニキャストアドレスを用いる。

それ以外の場合　上記の全てに属さない場合については、ルーティングテーブルをチェックし、そのパケットを転送するのに使用されたインタフェースを調べる。そして、そのインタフェースについたユニキャストアドレスを送信元アドレスとして、メッセージを送信する。

3.4 セキュリティに関する考慮

IPv6 でセキュリティへの対応が容易となっていることから、ICMPv6 はこれを利用した仕組みが追加されている。ICMPv6 では、メッセージの認証が可能となっており、ICMP パケットを交換する際には、IP 認証ヘッダでの認証が可能である。ICMP パケットを受信したホストは、認証ヘッダにおいて正当性を照合し、不正なパケットの場合には破棄する。

しかし、ICMP はさまざまな攻撃を受けやすいため、防護策が必要である。この防護策については、文献 [2] に定められている。

4 ソフトウェアの実装例

4.1 traceroute6 の動作概要

ここでは、ICMP を使用したソフトウェアの実装例として、traceroute6 を挙げる。traceroute6 は、始点から終点へ IP データグラムがたどるパスを明らかにする traceroute というソフトウェアを IPv6 対応にしたものである。

traceroute6 は、始点から終点に対して、probe パケットというパケットを UDP を使用して送信する。送信する際に、始めは TTL を 1 とする。もし、probe パケットが TTL が足りずに終点に到達できなかった場合には、始点から終点までのパスの途中に位置するルータが ICMP 転送中時間超過エラーを返す。これにより、始点には途中のルータからエラーが送られて来るため、パスに位置するルータの IP アドレスが分かる。始点からは、TTL を 1 ずつ増やして、probe パケットを送信し続け、終点に到達した場合には、終点から ICMP ポート到達不可エラーを返させることにより、probe パケットが終点に到達したことを知る。

図 3 に traceroute6 の動作の概要を示す。

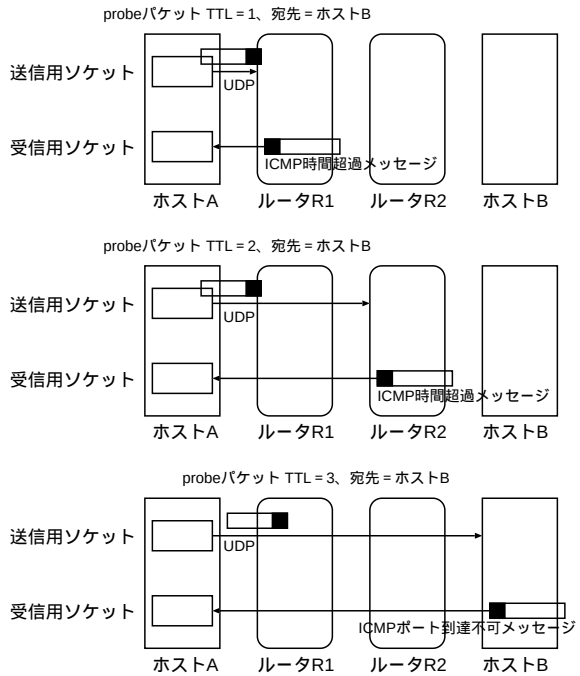

図 3: traceroute6 の動作の概要

この図に示しているように、ホスト A からホスト B に対して、probe パケットが TTL=1 ~ 3 で送信され、これに対し、ICMP の転送中時間超過エラー、ポート到達不可エラーがそれぞれルータ R1、R2、ホスト B から返る。このようにして、始点から終点へのパスはホスト A → R1 → R2 →ホスト B であることが分かる。

4.2 traceroute6 のソースコード

付録 1 として、traceroute6 のソースコード (一部) を掲載する。このプログラムでは、送信用と受信用の 2 つのソケットが必要となるため、まず最初に 2 つのソケットを作成している。受信用ソケットは、ICMP のメッセージパケットを受信できるようにするために、raw ソケットとして準備する。socket 関数の第一引数として IPv6 のソケットであることを示す AF_INET6 を、第二引数として raw ソケットであることを示す SOCK_RAW を渡している。送信用ソケットの方は UDP を用いるため、socket 関数の第二引数は SOCK_DGRAM となる。

次にプログラムの核となる部分であるが、まず一番外側のループで TTL を 1 ずつ増やし、終点に到達するまで probe パケットの送信を繰り返す (ア)。次に probe パケットを定められた回数だけ送信する (デフォルトでは 3 回同じ TTL で probe パケットを送信する。)(イ)。そして probe パケットを送信し (ウ)、ICMP のメッセージが返って来るまで待つ (エ)。ICMP メッセージが返って来ると、メッセージの種類によって分岐し、ポート到達不可メッセージだった場合には、終点に到達しているので got_there フラグを 1 にしてその後ループから抜けることになる (オ)。転送中時間超過エラーの場合には、何もせずにループを続行する。

4.3 traceroute6 の実行結果

付録 2 として、traceroute6 の実行結果を示す。示している結果は、九州大学内のホスト (3ffe:501:2c24:700:290:27ff:feaf:85f3) から、www.kame.net に対して実行したものである。結果は見やすくするため、改行を加えてある。この結果を見ると、始点から終点までの間に、8 つのルータを介していることが分かる。traceroute を使用することにより、例えばネットワークが不通になった場合に、始点からどこのルータまでは到達することができるかを調べることができるため、問題が発生しているルータを特定するといった用途に使用することができるであろう。

参考文献

[1] S. Deering and R. Hinden. Internet Protocol, Version 6 (IPv6) Specification, RFC 2460. December 1998.

[2] S. Kent and R. Atkinson. Security Architecture for the Internet Protocol, RFC 2401. November 1998.

＜付録 1：traceroute6 のソース（一部抜粋）＞

```
/*受信用ソケットの作成（raw ソケット）*/
icmp_sock = socket(AF_INET6, SOCK_RAW, IPPROTO_ICMPV6);

/*送信用（UDP）データグラムソケットの作成*/
if ((sndsock = socket(AF_INET6, SOCK_DGRAM, 0)) < 0) {
  perror("traceroute: UDP socket");
  exit(5);
}

/*省略...*/

for (ttl = 1; ttl <= max_ttl; ++ttl) {---(ア)
    ...
    for (probe = 0; probe < nprobes; ++probe) {---(イ)
        ...
        send_probe(++seq, ttl);---(ウ)
        ...
        while ((cc = wait_for_reply(icmp_sock, &from, reset_timer)) != 0) {---(エ)
           ...
           if ((i = packet_ok(packet, cc, &from, seq, &t1))) {
              ...
              switch(i - 1) {
              case ICMPV6_PORT_UNREACH:---(オ)
                   ++got_there;
                   break;
              ...
```

< 付録 2：traceroute6 の実行結果 >

```
[shiba@pleasure iputils]# ./traceroute6 www.kame.net
traceroute to kame220.kame.net (3ffe:501:4819:2000:210:f3ff:fe03:4d0)
from 3ffe:501:2c24:700:290:27ff:feaf:85f3, 30 hops max, 16 byte packets
 1  janus.v6.cc.kyushu-u.ac.jp (3ffe:501:2c24:700:210:4bff:fe11:6c8)
0.435 ms  0.409 ms  0.363 ms
 2  3ffe:501:2c24:a00:200:e2ff:fe12:5393 (3ffe:501:2c24:a00:200:e2ff:fe12:5393)
0.581 ms  0.554 ms  0.546 ms
 3  3ffe:501:2c00::1 (3ffe:501:2c00::1)
0.912 ms  0.926 ms  0.87 ms
 4  swift.sakyo.wide.ad.jp (2001:200:0:5201:2d0:b7ff:fe1e:8834)
10.902 ms  10.908 ms  10.883 ms
 5  june.nara.wide.ad.jp (3ffe:501:800::1)
17.499 ms  17.552 ms  17.479 ms
 6  2001:200:0:6001:2a0:ccff:fe73:386c (2001:200:0:6001:2a0:ccff:fe73:386c)
25.474 ms  25.55 ms  25.5 ms
 7  pc3.yagami.wide.ad.jp (2001:200:0:1c04::1000:2000)
26.41 ms  26.447 ms  26.566 ms
 8  gr2000.k2c.wide.ad.jp (3ffe:501:4819:2000::2000:1)
27.265 ms  26.83 ms  26.999 ms
 9  apple.kame.net (3ffe:501:4819:2000:210:f3ff:fe03:4d0)
26.814 ms  26.777 ms  26.72 ms
[shiba@pleasure iputils]#
```