JPRS トピックス＆コラム

■「512の壁」を越える ～EDNS0の概要と運用上の注意～

DNSプロトコルに存在する制限の一部を緩和し、DNSSECやIPv6などへの対応を可能にするためのDNSの機能拡張技術である「EDNS0」について解説します。

■DNS における UDP の採用と 512 バイト制限

DNS の通信ではドメイン名や IP アドレスなど、比較的短いデータを含む **DNS メッセージ**が頻繁にやりとりされることが一般的です。こうした通信形態では、相手との接続の事前確立を必要としない **UDP** を用いる方が、接続の事前確立が必要な TCP よりも通信にかかるコスト・時間を抑制でき、システム全体の効率や処理能力を向上させることができます。

このような背景から、DNS では主な通信プロトコルとして UDP が採用され、TCP は権威 DNS サーバー間のゾーン転送やある程度の長さを超える DNS メッセージを取り扱う際にのみ使用されるように設計されました。そして、UDP において取り扱い可能な DNS メッセージの最大長は当初、**512 バイトまで**と定められました。

■512 バイト制限の理由

なぜ、UDP における DNS メッセージの最大長は 512 バイトまでと定められたのでしょうか。

インターネットで使われている IP（IPv4）の仕様では一度に受信可能なデータグラム（ヘッダーを含むパケット）として、576 バイトを保証しなければならないと定められています。この値は、64 バイトのヘッダーと 512 バイトのデータブロックを格納可能な大きさとして選択されたものです（RFC 791 3.1. Internet Header Format）。

このため、UDP における DNS メッセージサイズの最大値を 512 バイトまでとすることで、通常の DNS 通信は IPv4 ネットワークにおいて**必ず 1 パケットで送受信可能**になります。

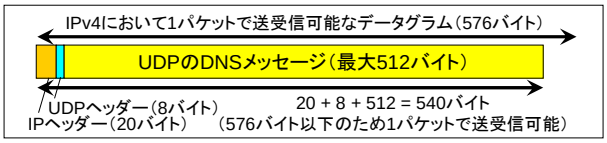

図 1 UDP における DNS メッセージサイズ

このことは通信の信頼性が高くなかった当時のインターネットにおいて、DNS を実用的に使用可能にするための重要な要素となりました。

■TCP フォールバックと「512 の壁」

そして、512 バイトを超える DNS メッセージを応答するための方法として、「**TCP フォールバック**」と呼ばれる方式が定められました。

DNS サーバーが返そうとする応答が 512 バイトを越えた場合、512 バイト以下に切り詰めた上で「データが切り詰められた」というビット（TC）をセットした応答を送信元に返します。TC がセットされた応答を受け取った場合、送信元は同じ問い合わせを TCP で再送します[1]。この一連のやりとりを TCP フォールバックと呼びます。

TCP フォールバックは必ず UDP による通信の後で発生します。そのため、名前解決にかかるコスト・時間が最初から TCP を使う場合よりも更に大きくなります。この UDP における DNS メッセージサイズの制限は後に、DNS における「**512 の壁**」と呼ばれるようになりました。

■DNSSEC の標準化と「512 の壁」の顕在化

その後、1990 年代に DNS キャッシュポイズニングが関係者の間で問題視され[2]、**DNSSEC** の標準化作業が開始されました。そして、1997 年に DNSSEC の最初の仕様（RFC 2065）が標準化されました。

DNSSEC を導入した場合、DNS 応答を検証するための鍵や署名により DNS メッセージのサイズが増大します。この、メッセージの増大に伴う UDP メッセージのオーバーフローと、オーバーフローした場合の TCP フォールバックの高いオーバーヘッド（512 の壁）が、DNSSEC 導入における障害として顕在化してきました。

---

[1] TCP では 65,535 バイトまでの DNS メッセージサイズを処理できます。
[2] 1990 年に書かれた S. Bellovin 氏の論文がきっかけとなりました。

※掲載内容は2013年11月現在のものです。

■EDNS0 の概要と特徴

IPv4 における 576 バイトの制限は、1980 年代当時の通信網の信頼性や制限を考慮したうえで設定されました。その後、ローカルエリアネットワークにおいて広く普及した Ethernet では 1 回の伝送で送信可能な最大値（MTU）は 1500 バイトとなっており、かつ広域ネットワークの信頼性も当時と比べ、飛躍的に向上しています。

そのような背景から、通信に従来の UDP を使いながら DNS メッセージサイズの制限を緩和するための拡張機能である「Extension Mechanisms for DNS（**EDNS0**[3]）」が、1999 年に標準化されました（RFC 2671）。EDNS0 には DNS メッセージサイズの制限緩和以外のプロトコル拡張も含まれており、DNS を DNSSEC や IPv6 に対応させる場合、EDNS0 への対応が必須とされています（RFC 3226）。EDNS0 はその後、細部の仕様や推奨値などが 2013 年に変更されています（RFC 6891）。

EDNS0 では Ethernet における DNS メッセージサイズの推奨値を、**1280 バイトから 1410 バイトまでの間**としています。

■EDNS0 による通信のしくみ

EDNS0 による通信では **OPT** というリソースレコードが、DNS メッセージの additional セクションに付加されます。OPT レコードは A レコードや NS レコードなどとは異なり、DNS における通信（トランザクション）の際にのみ現れることから、「疑似 RR」（pseudo-RR）と呼ばれています。

EDNS0 に対応したキャッシュ DNS サーバーが権威 DNS サーバーと通信する場合、まず OPT レコードを用いた DNS メッセージを通信相手の権威 DNS サーバーに送信します。相手が EDNS0 をサポートしている場合には正しい応答が返るため、その後のその相手との通信では EDNS0 が有効になります。

相手が EDNS0 をサポートしていなかった場合、OPT レコードを用いた DNS メッセージに対しエラー応答が返るか、OPT レコードが無視されることになります。その場合、その相手とは従来の DNS プロトコルにより通信を継続します。

[3] EDNS0 の「0」は、バージョン番号に由来しています。

相手が EDNS0 に対応しているかどうかは応答を受け取ったキャッシュ DNS サーバーにより、一定時間キャッシュされます。

■EDNS0 環境における注意

EDNS0 を用いた環境では、**512 バイトよりも大きな UDP による DNS 応答**が送受信されます。また、MTU を超える大きな DNS 応答では、**IP フラグメンテーション**[4]が発生する可能性があります。

一部のルーターやファイアーウォールなどのネットワーク機器はこのような DNS メッセージに対応しておらず、**誤動作**や**データの喪失**などが発生する可能性があります。その場合、**DNS の誤動作**や**名前解決エラー**などにつながる危険性があります。

そうしたトラブルの発生を防ぐため、自組織で運用しているネットワーク機器が EDNS0 に対応した DNS メッセージを正しく取り扱えるかを事前確認し、必要に応じた機器の更新やファームウェア・ソフトウェアのバージョンアップなどの対応を実施しておく必要があります。

また、EDNS0 をサポートしていない相手との通信や、EDNS0 により設定される DNS メッセージサイズ（ペイロードサイズ）よりも大きな DNS メッセージを応答する場合、従来からの TCP フォールバックによる再問い合わせが実行されます。そのため、各組織において設定するファイアーウォールなどでは、DNS のための設定として UDP の 53 番ポートだけではなく、**TCP の 53 番ポートへのアクセスも許可しておく**必要があります。

■今回のまとめ

① DNS メッセージサイズにおける 512 バイト制限は、IPv4 の仕様に由来している
② 512 バイトを超える DNS メッセージを応答する方法として、TCP フォールバックという方式が定められた
③ DNSSEC の標準化が「512 の壁」を顕在化させた
④ 「512 の壁」を超える仕組みとして EDNS0 が開発された
⑤ DNSSEC や IPv6 対応では EDNS0 対応が必須
⑥ 使用中のネットワーク機器が EDNS0 に対応しているか、確認しておく必要がある
⑦ UDP/53 に加え、TCP/53 も忘れずに許可しておくこと

[4] IP の内部処理において、MTU よりも大きなデータを小さなデータに断片化することをいいます。

