DNS Day 2006

EDNS0関連の状況と設定について

2006/12/06
DNSOPS.JP **森 拓也** (tak@scs.co.jp)

EDNS0をサポートしているサーバ

- **サポート**
  - BIND 8.3.0**以降**
  - BIND 9
  - Nominum ANS/CNS
  - Windows 2003 server
  - NSD
- **未 (不) サポート**
  - djbdns
  - Windows 2000 server

MS Windowsサーバ

- Windows Server 2003で EDNS0が標準設定に
  - Windows Server 2000からのアップグレードで問題多発
  - さまざまなファイアウォールがEDNS0をサポートしていなくて問題に
    - PIX Firewall(2003/7/25対応)
    - Firewall-1 (2004/05/30?)
    - Etc.

ブロードバンドルータ

- フォワーディングサーバ機能を持つものが多い
- いくつかの実装
  - ヤマハ
    - EDNS0の問い合わせは無視→エラーを返すようになった
      - RT56v Rev.4.07.37、RT57i Rev.8.00.19 (2003/09)
  - ADSLモデム NVIII
    - EDNS0パケットは通すが、応答は512Byteで切られる
  - Alaxala
    - フォワーディング対象は、UDPのみ (EDNS0未サポート)
- 現実には、数回エラーするとフォールバックするので、通信自体ができないことはなさそう
  - EDNS0→(DNS)→TCP

いくつかのクライアントの振る舞い

	UDP	EDNS0	TCP
Windows XP SP2	✓	× (既定値)	✓
Vista RC1 JP	✓	× (既定値)	✓
MAC OS X 10.4.8	✓	× (resolv.confで、「option edns」を指定しても同様)	✓
RedHat EL 3など	✓	× (既定値) ✓ (最近のライブラリでは、resolv.confで、「option edns」を指定することで可能)	✓

SCS 住商情報システム株式会社 Sumisho Computer Systems Corporation

5

Windows Vista!

- ネットワーク関連の変化
  - IPv6
    - 既定でインストールされる
    - 有効である
    - 優先して使用される
  - DNS
    - 名前解決
      - A レコードと AAAA レコードの両方を問い合わせる
      - AAAA レコードが返れば、AAAA レコードが優先
    - DNS IPv6トランスポート
      - Windows Server 2003 では実装済み
      - Windows XP では AAAA レコードの問い合わせであっても、IPv4 使用
        - » ピュア IPv6 の環境では DNS への問い合わせができないという問題
  - DHCPv6 サポート
    - DHCPv6 をクライアントとサーバーともにサポート
    - DHCPv6 prefix delegation もサポート
- 既存ネットワークへの影響
  - 今すぐ大きな影響が出るとは思えない
  - IPv6を用いたメジャーなサービスが出てきたら、変化は急激?

SCS 住商情報システム株式会社 Sumisho Computer Systems Corporation

6

EDNS0とTCPフォールダウン

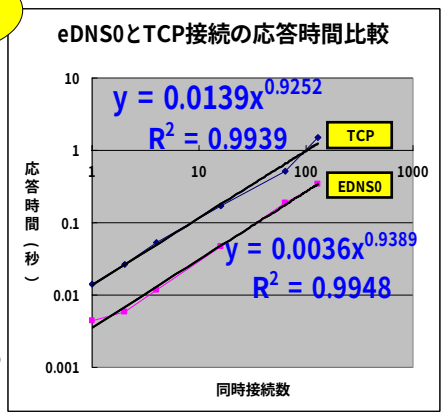
- UDPの場合と同じく、EDNS0でもadditionalセクションの返答で応答サイズを調整（BIND9.3.2）
  - 例：応答サイズ1011バイトの場合

Bufsize DiGの設定値	Protocol	Response size DiGの表示値	# of Add. Sec.
1011	EDNS0	1011	2
1010	EDNS0	995	1
979	EDNS0	979	0
978	TCP	1011	2

TCPフォールダウン時のレスポンス低下

- 環境
  - Pentium 166MHz
  - FreeBSD 5.4R
  - BIND 9.3.1
  - LAN環境
- 結果
  - 上がTCP、下がEDNS0
  - ベキ係数はほぼ等しい(約0.93乗)
  - 係数はTCPが約3.5倍
  - CPU負荷は高い
  - 外挿すると、同時セッション2000で、TCP15.7秒、EDNS0 4.5秒
- TCPの数値はさらに悪くなるかも?
  - 回線のレーテンシ
  - 同時接続数が増えた場合のオーバヘッド

最新CPUの1/50



実験結果を伊藤高一さんにいただきました。ありがとうございました。

EDNS0の問題点（セッション状態が悪い場合）

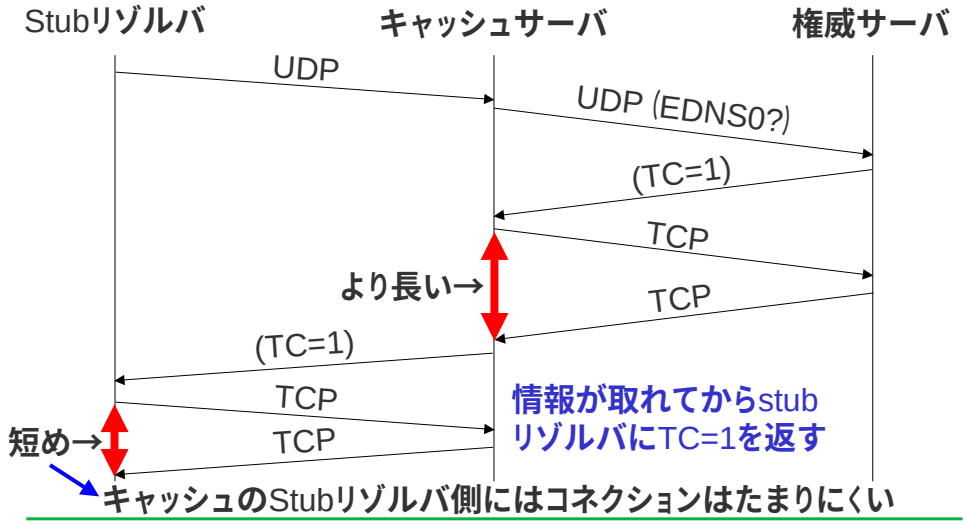
- EDNS0を使うとフラグメント化されたパケットが送られる
  - ネットワーク機器の問題
  - デフラグの負荷は?
  - ドロップに弱い
    - MTU=576だと、最大8パケットに分割（MTU=1500だと3パケット）
    - ドロップ率をαとすると、応答が正しく返る確率は、（1-α）^8
    - TCPなら再送機能があるので、影響を受けにくい
    - ドロップ率1%を超えるとしんどそう
    - end-2-endのドロップ率って、もっと高くなるかも
    - いつでもどこでもedns-udp-sizeを4096とかにするのは微妙かも
    - まあ、ISPのキャッシュDNSサーバならO.K.?

ドロップ率	3パケット	8パケット
5.00%	14.26%	33.66%
2.00%	5.88%	14.92%
1.00%	2.97%	7.73%
0.50%	1.49%	3.93%
0.20%	0.60%	1.59%

EDNS0 query First?

- BIND9では、EDNS0のクエリをまず出す
  - メリット
    - EDNS0が受け入れられれば、効率がよい
      - EDNS0
      - EDNS0（×TC=1）→TCP
  - デメリット
    - EDNS0が受け入れらないと、ペナルティーが大きい
      - EDNS0 (NOTIMPL, FORMERR, or SERVFAIL)→UDP
      - EDNS0 (×同上)→UDP(×TC=1)→TCP
    - EDNS0自体もオーバヘッドが若干はある
- Nominum社CNS：まずはnon-EDNS0（DNSSEC以外）
  - EDNS0の状況はサーバ単位でキャッシュされるので、ペナルティは大きくないとNominum社は主張

Stubリゾルバ･キャッシュサーバ･権威サーバ



EDNS0設定 (bind9)

- edns-udp-size
  - 自分がリクエストを送る場合にのみ使われる
    - リゾルバへの応答には使用されない
      - DNS Ampアタックを緩和しようと思って、edns-udp-sizeを指定しても何も起こらない
  - UDPパケットサイズを制限したい場合
    - BIND9.4.0b3では、max-udp-sizeが使用できる
    - それ以前では、EDNS0の上限 (4096:BIND9の場合) までは送られてしまう
- edns [yes|no]
  - 指定のサーバに対し、EDNS0を用いた問い合わせを行わない

全体的に

- EDNS0を使おう
  - TCPはやっぱり性能が出ない
  - できればUDPですませたい
  - EDNS0はネットワーク状況に影響受けやすいことに注意
- 適切なedns-udp-sizeの値は?

適切なedns-udp-sizeの値は?

- eDNS-udp-sizeを指定してもクエリに対する応答には影響しない
  - AMPアタックの増倍率に直接は関連しない
  - 外部一般公開サーバでは、max-udp-sizeを設定した方がよい?（でもまだβ）
- K,L rootでの調査結果
  - **EDNS0 deployment**
    - http://www.nlnetlabs.nl/downloads/edns0.pdf
  - 約7割が、2048byte。1280/4096はそれぞれ1割/2割程度
  - ちなみに2004年はEDNS0クエリ率は20%
    - 2003年からの増加は、1.4-2.0倍

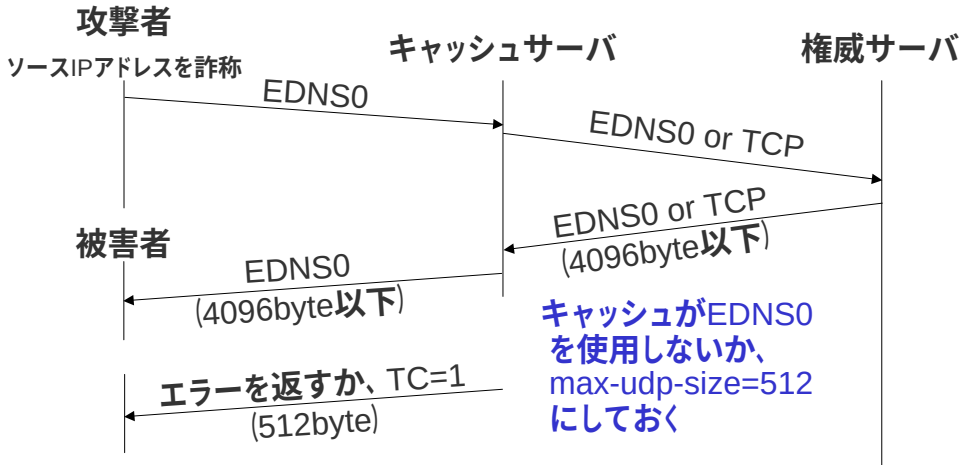
適切なedns-udp-sizeの値は?

- いくつかの候補値
  - 「DNS 512byteの壁」(2004年、Janog15、吉村さんほか) によると、TCPの応答平均値は744byte。ただし、NS RRに限れば平均1520byte。
    - http://www.janog.gr.jp/meeting/janog15/data/12-dns-yoshimura.pdf
    - とりあえず平均値の二倍くらい (1400byte) 以上あれば、ある程度はパフォーマンスあがる?
  - 特に考えず4096でよい?
  - 小さめにする?
    - IPv4のよく使われるMTU (PPPoEとかIPSecも考える必要あり) を考えると、1400byteくらい?
    - 1280byte-α:IPv6の最小パケット長を考慮?
    - EDNS0の応答状態はキャッシュされるので、届かなくても気にしない?

SCS 住金情報システム株式会社 Sumikin Computer Systems Corporation

15

Ampアタックとキャッシュ DNSサーバ



Ampアタックが成立しにくいか、増倍率を抑えられる

SCS 住金情報システム株式会社 Sumikin Computer Systems Corporation

16

EDNS0周りのサーバ設定

- キャッシュネームサーバ
  - 権威サーバ/forwarder への問い合わせはEDNS0を使う方がよい
    - 権威サーバへの負荷も軽くできる
  - クライアントへの応答および最大サイズに関してはちょっと微妙
    - 実際にはmax-udp-sizeしかいじれるところがない
- 権威ネームサーバ
  - BINDがまずEDNS0で送ってくる
    - 応答が512byteを超えなくとも、EDNS0は利用可能にしておいた方が楽
  - 逆に、EDNS0を止めておく必然性はなさそう
  - 設定しましょう！

(おまけ) 新パフォーマンステストツール

dnsperf/resperf/dhcperf

- Nominum社が開発したFreeの性能測定ツール
  - dnsperf：DNSサーバの性能評価ツール
  - resperf：キャッシュサーバ向き性能評価ツール
  - dhcperf：DHCPサーバ用性能評価ツール
    - 前の2つはソース提供。 dhcperfはバイナリで提供
  - URL
    - http://www.nominum.com/testing_tools.php

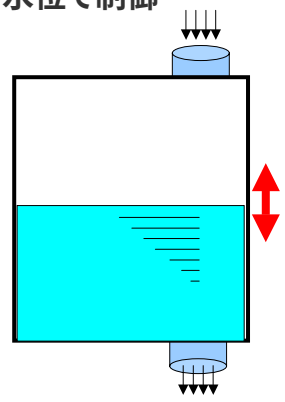
（おまけ）dnspefとresperfの相違点

- dnsperfは、Nominum版queryperfの後継
  - 連続したバースト的なクエリを送信
  - レスポンスを受信後、新規にクエリを送信
    - 権威サーバや、LAN環境でのキャッシュサーバのテストではO.K.
    - キャッシュのテストでWAN回線を使用した場合は結果が不十分になる可能性あり
      - レスポンス状況でdnsperf、queryperf の出す負荷(qps)が変化するため
- resperf
  - レスポンス状態にかかわらず負荷を上げていくことができる
    - したがって、テスト時にWAN環境の影響を受けにくい

（おまけ）相違点のイメージ

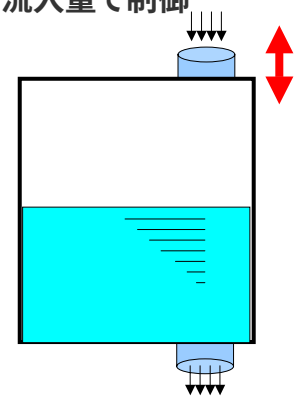
queryperf/dnsperf

水位で制御

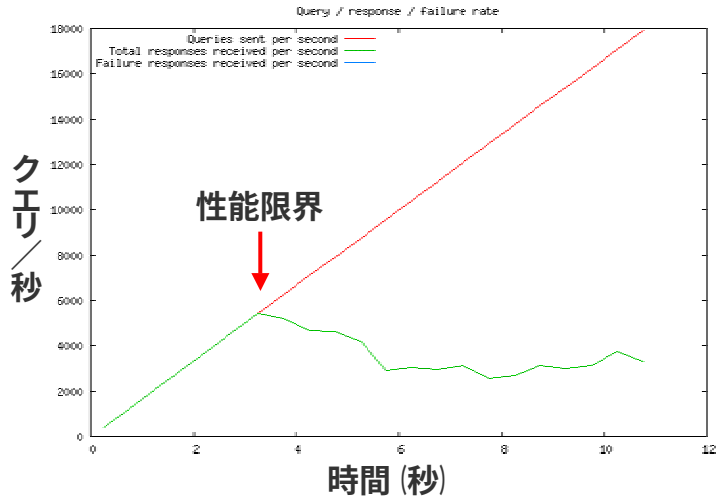


resperf

流入量で制御



(おまけ) resperf 結果グラフ その2 (Query/sec)



(おまけ) resperf 結果グラフ その3 (レイテンシ)