社団法人 情報処理学会 研究報告
IPSJ SIG Technical Report

2008－DSM－48 (4)
2008－QAI－26 (4)
2008／3／6

DNS キャッシングサーバにおける異常クエリ分析

豊野剛 石橋圭介 西田晴彦 三宅延久
日本電信電話株式会社 NTT 情報流通プラットフォーム研究所†
† 〒180-8585 東京都武蔵野市緑町 3-9-11

DNS はインターネットの基盤となるサービスである．本論文では DNS キャッシングサーバに注目し，実際に大規模ネットワークにおけるユーザクエリの挙動を分析した．キャッシングサーバにおける異常な挙動のクエリを定義し，分析した結果，ユーザクエリにも多量の異常クエリが含まれていることが確認された．また，ユーザ毎の単位時間当たりのクエリ発信数（クエリレート）で閾値を設けることで，ほぼ異常クエリだけを分離することが可能なことが分かった．

An Analysis of Anomalous Queries on DNS Caching Servers

Tsuyoshi Toyono, Keisuke Ishibashi, Haruhiko Nishida and Nobuhisa Miyake
Nippon Telegraph and Telephone Corporation
NTT Information Sharing Platform Laboratories‡
‡ Midori-cho 3-9-11, Musashino, Tokyo, 180-8585 Japan

The DNS (Domain Name System) domain names to be used in the Internet transactions instead of IP addresses. We analyzed user DNS queries on caching servers in a large scale network. These measurements show user queries have the large number of anomalous repeat queries, and we can separate these anomalous queries by using simple per-user query rate filtering.

1. まえがき

DNS（Domain Name System）はインターネット上においてドメイン名（Domain Name）と IP アドレスを相互変換する機能を有する分散データベースシステムである．インターネット上のノードは一意な IP アドレスを持ち，これを識別子として用いている．しかしユーザがインターネット上のサービスを利用する際には IP アドレスを用いてノードを指定することは少なく，実際にはドメイン名によってアクセス先のノードを指定している．今後数年で一般ユーザも IPv4 及び IPv6 の Dual Stack 環境を利用するようになると予測されている[1]が，その場合には DNS は IPv4 および IPv6 双方のアドレス空間とドメイン名空間の相互変換を行うことになる．更に，DNS は上記のドメイン名と IP アドレスの変換の他にもメール配送先決定や ENUM [2]など，各種サービス検索に利用されており，今日のインターネットのインフラともいえる重要な機能を担っている．この DNS の信頼性・耐障害性を確保し，性能劣化を回避することは極めて重要である．

DNS のサーバには大別するとネームサーバとキャッシングサーバという 2 種のサーバが存在する．DNS は単一のドメイン名空間を多くのサーバで分割して保持しており，一方向の木構造データベースとなっている．このドメイン名空間を保有する一連のサーバ群がネームサーバである．そのため，ドメイン名空間を維持する上でこれらのサーバが正常に動作していることは重要であり，ネームサーバの応答性能や信頼性・耐障害性に関する研究[3] [4] [5]が数多く行われている．また応答性能の劣化要因として，設定ミスや実装上のバグ，DoS 攻撃や異常クエリなどが挙げられ，この要因に関しても分析が行われている[6] [7] [8]．

一方，ユーザからのドメイン名と IP アドレスの変換の問い合わせ（クエリ）を受け，ネームサーバへの問い合わせを仲介するものがキャッシングサーバである．ユーザのクエリを受けるキャッシングサーバの応答性能は，ユーザのネットワーク利用の体感速度に直接影響を及ぼす．特に ISP や企業の社内網など，多くのユーザを抱える大規模ネットワークになるとキャッシングサーバが処理しなければならないユーザクエリ数は膨大なものとなり，負荷も高くなる可能性がある．

キャッシングサーバの挙動に関しては，主に応答性能や耐障害性に関する分析として，キャッシュのヒット効率[9]や Root ネームサーバ，TLD[i]ネームサーバといった基幹ネームサーバへの反復クエリ状況，ユーザへのレスポンス時間などの分析[10] [11]が行われてきた．しかし，キャッシングサーバにおいては，応答性能の劣化要因に関する分析はまだ十分であるとは言えない．

これらの事を踏まえ，本論文ではキャッシングサーバにおいて，応答性能に影響を与える要因を明らかにすることを目的とし，特にユーザとキャッシングサーバ間の DNS トラフィックに注目する．まずユーザからのクエリの中で，キャッシングサーバの応答性能に影響を与える負荷要因となり得る問い合わせを異常クエリとして定義し，分類する．そして実際のネットワーク上で観測したユーザクエリに対してこの異常分類に基づき行った分析の結果を紹介する．更に，分析結果を基にキャッシングサーバの性能劣化を回避し応答性能および耐障害性を確保するための考察を行う．

2. DNS トラフィックデータ

2.1 DNS 概要

DNS では，一部の例外を除いて主に UDP を用いたセッションレスな通信が行われる．DNS の挙動はプロトコルとして定められており[12] [13]，大きく分けて以下の (a) (b) (c) の 3 つのシステムから成り立っている（図 1）．

[i] Top Level Domain はドットで区切られた階層構造を持つドメイン名のうち，最後尾の項を指す．例として jp., com. など．TLD ネームサーバはこれを管理するネームサーバを指す．

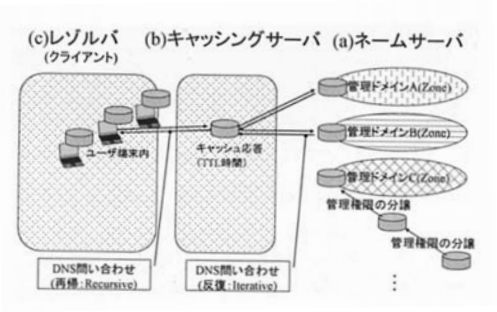

図 1 DNS の概要

(a) ネームサーバ (Name server)

ドメイン名 (Fully Qualified Domain Name : FQDN) [ii] と IP アドレスのマッピングデータベースを保有し，そのドメイン名に対して応答する権利を有するサーバをネームサーバと呼ぶ．データベースはゾーンと呼ばれるドメイン単位毎に分割され階層的に権限を委譲して管理される．通常は同じデータベースを保有する複数台が同期を取った上で分散して設置され，冗長化される．

(b) キャッシングサーバ (Caching server)

組織やネットワークドメインの境界毎に設置され，ユーザの問い合わせとネームサーバの応答を仲介する．キャッシングサーバはユーザからの問い合わせ（再帰クエリ）を受けると，代理としてネームサーバへの問い合わせ（反復クエリ）を順次行い，最終的に得られた回答をユーザへ応答する．同時に得られた回答内容をキャッシュとして TTL (Time To Live) で指定された期間だけ保持し，他のユーザから同じ問い合わせがあった場合にはキャッシュから応答することでネームサーバへのクエリの氾濫を抑制する．またユーザへの応答にも TTL (Time To Live) を付与する．

(c) レゾルバ (Stub resolver)

DNS ではクライアントを特にレゾルバと呼ぶ．レゾルバはキャッシングサーバに問い合わせを依頼することで間接的に名前解決を行う．レゾルバが出す問い合わせを再帰クエリと呼ぶ．通常，レゾルバは OS に組み込まれユーザ端末内で動作する．レゾルバもまた得られた回答内容をキャッシュとして保持することで一定時間は同じ問い合わせを抑制できる仕組みになっている．

2.2 DNS トラフィックデータの取得

本論文ではインターネット上での実際のユーザの問い合わせの振る舞いを正確に把握するために，多数のユーザに利用されている大規模ネットワークの DNS キャッシングサーバ (2.1 項 (b)) におけるトラフィックデータを取得した．取得するデータは DNS サーバを送信元あるいは宛先とする UDP/TCP port 53 の DNS トラフィックのみを対象とした．今回の分析に用いたデータは平日の 24 時間分である．

3. DNS での異常クエリ

3.1 DNS サーバ応答性能の劣化要因

多数のユーザに影響を与えるキャッシングサーバでも，多くのネームサーバと同様，サーバの信頼性・耐障害性を確保し，性能劣化を回避することは重要である．

一般的なサーバ-クライアント型通信と同様に，DNS サーバにおいてもサーバ性能に比して処理可能なトランザクション数に限界がある．特にキャッシングサーバの場合，ユーザから受けた問い合わせをネームサーバに反復的に問い合わせてから応答するという作業を行うため，1 つのユーザクエリがその何倍ものトランザクション処理の要因となり得る．このため，キャッシングサーバの応答性能に影響を与える要因として，ユーザからのクエリ状況を正確に把握することが重要となる．

1 項で挙げた Brownlee, N et al. [7]による Root ネームサーバの研究では，Root ネームサーバが受信している問い合わせのうち，応答性能に影響を与える要因となる異常 (Bogus) なクエリの分析を行っている．異常クエリの中には，プライベート空間 IP アドレスに対する逆引きクエリなどのプロトコル的な異常クエリのほか，特に，

(a) 存在しないドメイン名 (Root サーバの場合は存在しないトップレベルドメイン名) に対する問い合わせ．

(b) サーバの応答した TTL を無視し同じ内容を繰り返し続ける問い合わせ．

の 2 種に分けられるものが多く，更に (b) に関しては 1 分以内に繰り返される“リピートクエリ”が全クエリのうち 2 割から 5 割を定常的に占有していることが示されている．また，リピートクエリには多くの再帰クエリが含まれていることが示され，これらはネームサーバではなくローカルサーバ (すなわちキャッシングサーバ) で処理されるべきであると述べられている．DNS の問い合わせの階層構造を考慮すると，これらの異常クエリはユーザからキャッシングサーバに向けても同様に問い合わせられている可能性が高く，キャッシングサーバの負荷上昇，性能劣化要因となっている可能性があるため，今回の分析対象とする．

3.2 ユーザ異常クエリの定義

ユーザからキャッシングサーバへの問い合わせのうち，キャッシングサーバの負荷増大に影響し，応答性能の劣化を引き起こす可能性のあるクエリを，今回は“異常クエリ”と定義する．3.1 項で述べたように，Root ネームサーバでの分析で全クエリのうち多くの割合を占めていたのは存在しないドメインへの問い合わせ，およびリピートクエリである．そしてこれらのクエリはキャッシングサーバへ向けても問い合わせられている可能性が高いと推測できる．

一方，ユーザからのリピートクエリに関しては，Mass Mailing Worm の動作や DDoS Attack の際に DNS 上で観測されることが知られている[14] [15] [16]．ユーザからのリピートクエリの発生要因を調べるためには，リピートクエリの性質によってより詳細な場合分けが必要となる．

これらを踏まえ，今回は異常と思われるクエリを振る舞いに応じて以下の 7 つにパタン分けし，これを DNS キャッシングサーバでのユーザ異常クエリと定義した．リピートクエリに関しては，Mass Mailing Worm に見られる MX クエリ，及び DDoS Attack の際に見られるエラークエリについてを特に分けて分類した．

1. NxQtype (存在しないクエリタイプ)
   DNS では問い合わせ内容 (タイプ) を Qtype，問い合わせドメイン名を Qname として指定するが，Qtype が通常の DNS で定義されている 10 タイプ (A, NS, CNAME, SOA, PTR, MX, TXT, AAAA, SRV, ANY) 以外の問い合わせ．DNS クエリとして正しくない壊れたパケットもこの分類に含む．
2. NxTLD (存在しない Top Level Domain)
   存在しないトップレベルドメイン名に対する問い合わせ．“.localhost”や“.workgroup”など，ドメイン名

[ii] ホスト名と全てのドメイン名を完全表記したものだが，本論文では便宜上ドメイン名と称する．

で無いもののほか，IP アドレスに対する IP アドレス問い合わせ（A for A）などのクエリもこの分類に該当する．

3. RFC1918（プライベート空間 IP アドレス逆引きクエリ）
   RFC1918 [17]に定義されているプライベート IP アドレス空間の IP アドレスからドメイン名を牽く（逆引き）問い合わせ．これらはローカルネットワーク内で処理されるべきクエリであり，キャッシングサーバまで到達しているのは異常と判断する[iii]．
4. Ignore TTL（TTL 無視）
   キャッシングサーバからの応答に含まれる TTL を無視して同じ問い合わせ（同じ Qtype 及び Qname）を繰り返すもの．TTL には二種類あり，問い合わせレコードが存在する場合は当該レコードの TTL，問い合わせたレコードが存在しない場合（NXDomain または NODATA [iv]）は，応答した Default（SOA）TTL となる．後者は存在しないレコードに対する問い合わせを抑制するために用いられる．
5. Repeat MX（MX リピートクエリ）
   前回の問い合わせから短時間以内に，（問い合わせドメイン名は異なるが）問い合わせタイプ MX のクエリを繰り返すもの．Bot や Worm にこのような挙動を示すものがある．ここでは短時間を 0.01 秒とし，異なるドメイン名を問い合わせた場合も含むこととする．
6. Repeat Err（エラーリピートクエリ）
   サーバからの応答がエラー（ServFail，FromErr，Refused）だったにも関わらず，前回の問い合わせから短時間以内に同じ問い合わせを繰り返すもの．ここでは短時間を 1 秒とする．これらの応答はネームサーバ側の設定や Lame Delegation [v]などが要因であることが多いが，その場合タイムアウトを待たなければならず，キャッシングサーバのトランザクション処理に多大な負荷を及ぼす可能性がある．
7. Repeat（リピートクエリ）
   上記 1～6 に該当せず，前回の問い合わせから短時間以内に同じ問い合わせを繰り返すもの．DNS ではレゾルバの適切な再問い合わせ間隔は 4 秒以上とされている[18]．ここでは短時間を 1 秒と定める．また，分類 4 の Ignore TTL と同等のクエリでも，サーバが何らかの理由で TTL を含んだ応答を返していなかった場合，そのクエリはここに分類される．

上記の 7 つの異常分類に当てはまらないクエリは，今回は全て「正常クエリ」とみなす．これらの中には今回は定義していないが性質的には異常とするべきクエリも含まれている可能性があるが，これに関しては後の 5.1 項で述べる．

4. DNS トラフィックの分析

2.2 項で取得した DNS トラフィックデータについて，今回定義した異常クエリに関する分析を行った．

DNS は基本的には UDP を用いたセッションレスの通信で完結するが，DNS パケットに付与された送受信 IP アドレス，DNS の 16bit のメッセージ ID，問い合わせタイプ（Qtype），問い合わせドメイン名（Qname）によってクエリと応答のトランザクションは比較的容易に推測可能である．これらを用いて問い合わせと応答のトランザクションを構築し，一定期間保持することで，以前に同一のユーザ[vi]から問い合わせられたクエリ内容との比較を行う．扱うトランザクション数が膨大であったため，本分析ではトランザクションの保持期間を 5 分とした．

分析した 24 時間のクエリ数時間変動をグラフ化したものを図 2 に示す．未明に流量が落ち込む，日中 12 時に流量が増加するなど，いわゆる一般的に見られるトラフィック流量の日変動グラフと同様の傾向を示していることから，DNS トラフィックはユーザのネットワーク利用形態を反映していると考えられる．

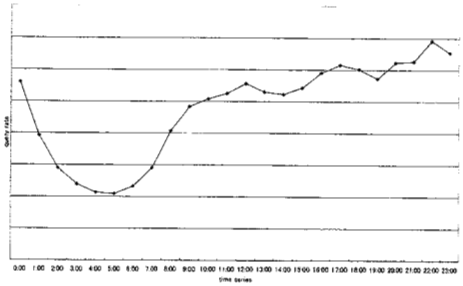

図 2 受信したクエリ数の推移（24 時間）

4.1 応答分析

まず，ユーザからのクエリに対して，キャッシングサーバがどのような応答を返しているかを分析した．DNS パケットには RCODE と呼ばれるフィールドが付されており，ユーザの問い合わせがエラーになった場合はこのフィールドで通知される．図 3 はキャッシングサーバからユーザへの応答を RCODE の割合でグラフ化したものである．このグラフから，ネームサーバへの問い合わせが失敗していることを示す“ServFail”や“Refused”は割合としては少ないことが分かる．そして問い合わせが正常に応答されたことを示す“No Error”が 78%を占め，また，問い合わせられたドメイン名が存在しなかったことを示す“NxDomain”が 17%を占めていることも分かる．この 2 点から（1）キャッシングサーバがユーザの代理にネームサーバに問い合わせを行う際に直接的に障害となるようなユーザクエリは少数であること，（2）95%のクエリに対しては正常もしくは回答無しの何らかの応答が返せていること，が分かる．

一方，同じデータを今回 3.2 項で定義した 7 つの異常分類で分析し，割合を円グラフにしたものを図 4 に示す．これを見ると，ユーザクエリで「正常」と分類できるものはわずか 15%にしか過ぎず[vii]，特に分類 4 の“TTL 無視”と分類 7 の“Repeat”が全クエリの約 80%を占めていることが分かる．

すなわち，実際のインターネット環境においてのユーザクエリは，ネームサーバへの問い合わせ自体は正しく行えており，プロトコル的には正常終了しているが（図 3），内容で分類するとそのうち 8 割はキャッシングサーバに負荷を与える可能性がある不適切なリピートクエリであることが判明した（図 4）．

[iii] http://as112.net/

[iv] No Error かつ Answer 0 の応答で，該当クエリタイプに対する回答が無いことを示す．

[v] http://www.apnic.net/services/rev-del/lame-del/

[vi] 本論文では“同一の送信元（Source）IP アドレス”からのクエリを便宜上“同一のユーザからクエリ”と称す．厳密には NAT 等により同一 IP アドレスから複数ノードのクエリが発せられている可能性や，分析期間中に IP アドレスが変わったノードが存在する可能性がある．

[vii] 図中では“Legitimate”と表記．

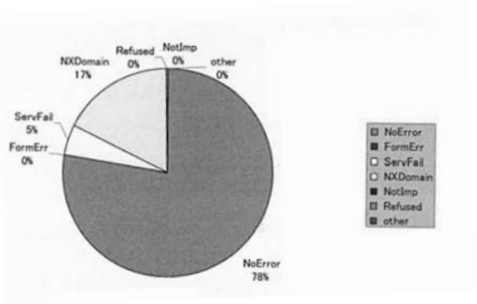

図 3 応答における RCODE 割合

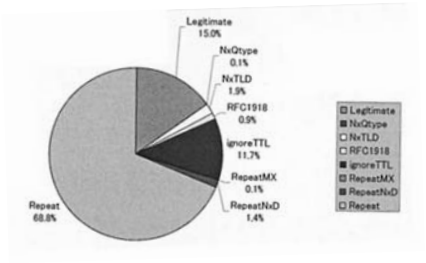

図 4 ユーザクエリにおける異常割合

4.2 クエリレート分析

前項では応答全体に占めるリピートクエリが約 8 割であることを示した．本項ではさらにキャッシングサーバで受信したユーザ毎のクエリ数および異常クエリの傾向について分析する．

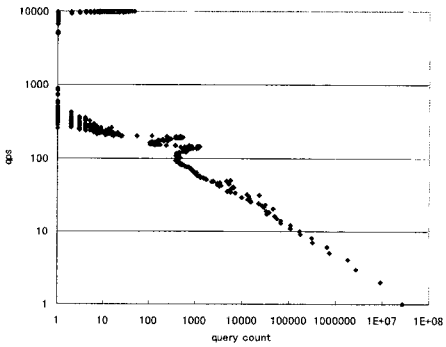

図 5 クエリレートの分布

図 5 はキャッシングサーバで観測されたユーザ毎の受信クエリ数を，単位時間当たりのクエリレート（qps: query per second）でプロットしたものである．この図では単位時間スロット毎にユーザ毎受信クエリ数をプロットしているため，単一のユーザが複数時間スロットに出現した場合は複数回プロットされている．ここで示されるように，単位時間当たりクエリ数とその受信回数については，いわゆる Zipf's law に従った分布が見られる．しかし外れ値として 10000qps 以上という膨大なクエリを投げ続けているユーザが複数観測された．

図 6 はクエリレートの受信回数の相補累積度数分布を示したものである．このグラフからは，ほぼ全てのユーザは，1 秒間に数回程度の問い合わせ（1pqs～10qps）を発しているに過ぎないことが分かる．これは Web やメールの利用といった一般的なネットワークアプリケーションの利用形態と合致していると考えられる．高いクエリレートを発出しているユーザは，例えばこの分析期間中に一度でも 100qps を越えるクエリを送出してきたユーザは全ユーザ数のうちわずか 0.075%に過ぎなかった．そしてそのクエリ量は全クエリ量の約 0.14%であった．

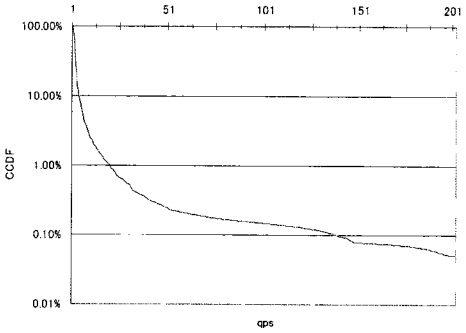

図 6 クエリレート相補累積度数分布

表 1 クエリレート毎にみた異常割合

異常分類	100qps 以上	200qps 以上	300qps 以上	400qps 以上	500qps 以上
Legitimate (正常)	0.09%	0.01%	0%	0%	0%
NxQtype	0%	0%	0%	0%	0%
NxTLD	0%	0%	0%	0%	0%
RFC1918	0.80%	0%	0%	0%	0%
IgnoreTTL	1.63%	0.05%	0.01%	0%	0%
RepeatMX	0.01%	0%	0%	0%	0%
RepeatErr	0.64%	0%	0%	0%	0%
Repeat	59.69%	59.69%	59.69%	59.69%	59.69%

表 1 はクエリレート毎に分析した異常クエリおよび正常クエリの割合である．なお，示した数値は 100qps 未満のクエリも含んだ総受信クエリ量に比する割合である．この表から，100qps を超えるクエリの中には，まだ 0.09%の正常なクエリが含まれていることが分かる．しかし 300qps を超えると正常な内容のものは含まれず，全てが何らかの異常クエリであるという結果になった．

また，高クエリレートの異常割合は，全体を分析した際の異常割合とほぼ同じ傾向を示し，分類 4 の“TTL 無視”および分類 7 の“Repeat”が大きな比率を占めた．ただし，特徴的な傾向として，全体を分析した際には TTL 無視と同じ程度の割合で存在していた正常クエリが大きく減少し，ほとんど観測されないことが分かった．

すなわち，正常なクエリは低いクエリレートを保って問い合わせられているのに対して，少数ユーザから高いクエリレートで同じ問い合わせ内容を繰り返し続ける異常リピートクエリが，サーバに処理負荷を与えていることが明らかになった．

5. 応答性能の確保に向けた検討

5.1 高クエリレートユーザの詳細分析

4.2 項に示したクエリレート分析を踏まえ，250qps 以上の異常リピートクエリを発するユーザ(IP アドレス)について，それぞれについて更に実際のクエリ内容を分析したところ，特徴的なものは大きく分けて 4 つに分類できることが分かった．なお付記したパーセンテージは高クエリレートユーザ数全体に占める割合を表す．

- NTP サーバ検索系（3.9%）

  単一の NTP サーバ名を解決しようとして，失敗している．time.stdtime.gov.tw など，世界的に用いられている単一の NTP サーバのドメイン名を問い合わせ続けている様子が観測された．特徴として 10000qps 以上の速度で問い合わせを行い続けており，単位時間当たりのクエリ数が膨大なため，ユーザ数的な比率では約 4%に過ぎないが，クエリ数比では全数の 50%以上を占める．4.2 項図 5 において，10000qps 付近の高クエリレートを発しているユーザの殆どはこれに該当している．ネットワーク機器の中にはこれら NTP サーバのリストをハードコーディングしてしまっているものもあるため，こういった機器が関係している可能性もある．これらのクエリは異常と推定できる．

- メールサーバ検索系（76.4%）

  メール配送を担う SMTP サーバと推測される“mail”，“mx”などの文字列が入ったサーバ群に対し，問い合わせタイプ A（IP アドレス），MX（メールサーバ名）を繰り返し問い合わせ続けている様子が観測されるもの．通常ユーザの利用形態からはこれらメールサーバの検索を繰り返し行うことはない．しかし高クエリレートユーザ数全体に占める割合は大きく，76%がこの分類に該当した．これに該当する挙動を示すユーザの端末は，ドメイン名が解決された後，更に Worm によって SPAM 配信などに利用されている可能性もある．これらのクエリは異常と推定できる．

- メッセンジャーサーバ検索系（7.8%）

  著名なメッセンジャーサービスのサーバのドメイン名を繰り返し問い合わせ続けている様子が観測されるもの．通常のメッセージングサービスの利用ではこのような多量の繰り返しクエリは発生しない．これに該当する挙動を示すユーザの端末は，メッセンジャーサービスを利用した SPAM 配信などに利用されている可能性もある．これらのクエリは異常と推定できる．

- 逆引き系（7.8%）

  IP アドレスからドメイン名を牽く（逆引き）を行う問い合わせタイプ PTR を大量に検索しているもの．Web サーバのアクセスログ解析アプリケーションなどは，定められた時刻にログ解析のためにこのように IP アドレスからドメイン名への変換を大量に問い合わせるという挙動を示すことがある．このクエリは，クエリレートと問い合わせタイプから異常と推定することは難しい．

上記に挙げたようにおよその傾向があるものの他にも，分類不可能な高いクエリレートのユーザが存在した．例えば，“pic”“img”“photo”などの文字列が含まれたドメイン名を検索し続けるユーザや，内容的には異常と思われる特徴が特に見られなかったユーザも見られた．上記の“逆引き系”を含め，これらはクエリレートからのみでは無用なクエリであるとは判断できない．実際には有用なクエリであるにも関わらず“巻き添え”になって異常クエリと判断されている可能性がある．

一方，これとは逆に今回の定義においては正常クエリと判断されているものの，実際にはサーバに高負荷を与えうる異常なクエリである“見逃し”も発生している可能性がある．例えばあるユーザが発信元 IP を詐称して高クエリレートで送信したりした場合，無用な高クエリレートの問い合わせであるにも関わらず，今回の分類では異常と判別されない．

今後分析を進めていく上で，これらに関しては更に考慮が必要であろう．

5.2 クエリタイプ分析

前項の詳細な分析を踏まえ，問い合わせタイプ（Qtype）によって異常クエリの傾向に有意な差があるならば，これによって分類することで更に高精度に高クエリレートユーザを正常か異常か判別することが可能になるのではないかと推測した．これが行えるならば，高クエリレートユーザが発しているクエリが Qtype　MX ならば異常クエリである傾向が高く，Qtype　PTR ならば異常クエリである傾向が低いといった判別が可能になる．

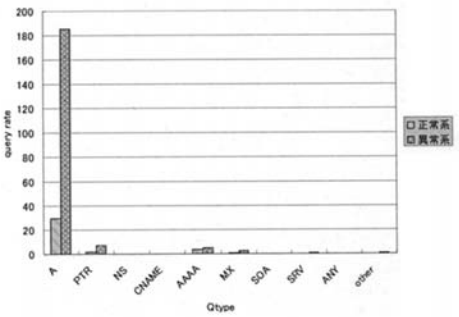

図 7 QTYPE 別 クエリ受信数

図 7 は問い合わせタイプ別にクエリの受信数を異常分類したものである．定義した 7 つの異常分類に合致したものを合わせて「異常系」これに該当しないものを「正常系」と記した．Qtype　A はドメイン名から IPv4 アドレスへの変換問い合わせ（正引き）を，Qtype　PTR は IP アドレスからドメイン名への問い合わせ（逆引き）を示す．図から，現状の利用ではこの A クエリに関する異常クエリが突出していることが分かる．

また，図 8 は Qtype 別に，異常クエリと正常クエリの割合を 100%を基準とした積み上げグラフに示したものである．

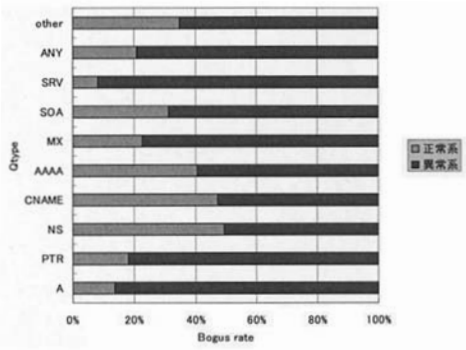

図 8 QTYPE 別 異常割合

このグラフにより，図 7 では総数で突出している Qtype A のみの異常クエリが目立つが，実際には他の Qtype の問い合

わせにおいても総受信回数のうち少なくとも半数は異常に分類されるクエリを受信していることが分かる．

このことから，実際には今回の異常クエリを更に Qtype 単独では判別することが困難であると言える．

5.3 キャッシングサーバの応答性能の確保に向けた検討

クエリレートに関する分析と異常クエリ分析の結果から，単位時間当たりのクエリレートの高いユーザからのクエリ内容は，そのほぼ全てが今回定義した異常クエリに合致していることが明らかになった．このことから，一定の閾値を設けて単位時間当たりのクエリ数の多いユーザ Top N を抽出することで，サーバの応答性能に影響を与える恐れのある異常クエリを分離できることが分かる．すなわち，ネットワークエッジでのアクセスコントロールリストの設定やサーバ側の応答制御においてクエリレートによる簡便な閾値設定を行うだけで，正常な問い合わせを行っているユーザへの影響を最小限に留めたまま効率的に異常クエリによる負荷上昇および応答性の劣化を抑止することが可能となる．ただし 5.1 項での分析の結果から，単位時間当たりのクエリレートが高いユーザの中にも，本来異常とするべきかどうか判断の難しいクエリがあり，巻き添えが発生する可能性があることには注意すべきである．

今後はクエリ内容を 5.1 項で示された特徴などで更に絞り込むことで，正常なユーザを巻き込むことなく高精度に異常クエリを振るい分けることが可能になっていくと期待できる．

6. むすび

本論文では DNS キャッシングサーバに着目し，特にユーザからのクエリに関する考察を行った．ユーザからのクエリのうち，キャッシングサーバの性能劣化に繋がるものを異常クエリとして定義し，今回は 7 つに分類した．実際に運用されている大規模キャッシングサーバの DNS トラフィックデータを分析し，一般ユーザのクエリの挙動を明らかにした．まず，応答内容を分析することによって，ユーザクエリのうち 8 割はキャッシングサーバに負荷を与える可能性がある不適切なリピートクエリであることが判明した．また，クエリレートを分析することによって，正常なクエリは低クエリレートで問い合わせられているのに対し，異常リピートクエリは少数のユーザから高クエリレートで問い合わせ続けられているということが分かった．これらの結果から，単位時間当たりのユーザ毎のクエリ数で一定の閾値を設けるのみで，正常な問い合わせを行っているユーザへの影響を最小限に抑えたまま，負荷上昇および応答性能の劣化の要因となる異常クエリを効率的に振るい分けすることができることが分かった．

今後はキャッシングサーバの信頼性・耐障害性を向上するために，応答性能に影響を与える要因となる異常クエリに関してネームサーバからの応答も含めて分析し，更に高精度に追跡可能な手法を検討していく予定である．

参考文献

[1]. G. Huston, "IPv4 Address Space Report," http://www.potaroo.net/tools/ipv4/ .

[2]. P. Faltstrom, and M. Mealling, "The E.164 to Uniform Resource Identifiers (URI) Dynamic Delegation Discovery System (DDDS) Application (ENUM)," RFC 3761, April 2004.

[3]. T. Lee, B. Huffaker, M. Fomenkov, and Kc claffy, "On the problem of optimization of DNS root servers' placement," in *Proc. Passive and Active Measurement Workshop (PAM) 2003*, April 2003.

[4]. N. Brownlee, Kc claffy, and E. Nemeth, "DNS root/gTLD performance measurements," in *Proc. USENIX LISA2001 Conference*, December 2001.

[5]. N. Brownlee, and I. Ziedins, "Response time distributions for global name servers," in *Proc. Passive and Active Measurement Workshop (PAM) 2002*, March 2002.

[6]. D. Wessels, "Is your caching resolver polluting the Internet?," SIGCOMM 2004 NetTS Wokrshop, September 2004.

[7]. N. Brownlee, Kc claffy, and E. Nemeth, "DNS measurements at a root server," in *Proc. IEEE GLOBECOM '01. IEEE*, Vol. 3, pp. 1672-1676, November 2001.

[8]. A. Broido, E. Nemeth, and Kc claffy, "Spectroscopy of private DNS update sources," in *Proc. the Third IEEE Workshop on Internet Applications. WIAPP 2003*, pp. 19-29, June 2003.

[9]. J. Jung, E. Sit, H. Balakrishnan, and R Morris, "DNS performance and the effectiveness of caching," *IEEE/ACM Transactions on Networking*, Vol. 10, No. 5, pp. 589-603, October 2002.

[10]. 加藤朗，関谷勇司，"ISP の DNS サーバの DNS トラヒックの解析，" *The transactions of the Institute of Electronics, Information and Communication Engineers. B*, Vol. J87-B, No. 3, pp. 327-335, March 2004.

[11]. T. Toyono, H. Nishida, and K. Ishibashi, "An analysis of the queries from caching servers to root servers," 2007 OARC DNS Operations Workshop, June 2007.

[12]. P. V. Mockapetris, "Domain names - concepts and facilities," STD 0013, RFC 1034, November 1987.

[13]. P. V. Mockapetris, "Domain names - implementation and specification," STD 0013, RFC 1035, November 1987.

[14]. K. Ishibashi, T. Toyono, H. Matsuoka, K. Toyama, M. Ishino, C. Yoshimura, T. Ozaki, Y. Sakamoto, and I. Mizukoshi, "Measurement of DNS traffic caused by DDoS attacks," in *Proc. IEEE SAINT 2005 Workshop on Network Security Threats and Countermeasures*, pp. 118-121, January 2005.

[15]. N Chatzis, "Mass mailing worm detection by means of situation aware DNS," in *Proc. the Eighth international Symposium on Autonomous Decentralized Systems*, pp. 279-286, March 2007.

[16]. Y. Musashi, and K. Rannenberg, "Detection of mass mailing worm-infected PC terminals by observing DNS query access," *IPSJ SIG Technical Reports, Computer Security 27th*, Vol. 2004, No. 129, pp. 39-44, December 2004.

[17]. Y. Rekhter, B. Moskowitz, D. Karrenberg, G. J. de Groot, and E. Lear. "Address allocation for private Internets," RFC 1918, February 1996.

[18]. A. Kumar, J. Postel, C. Neuman, P. Danzig, and S. Miller, "Common DNS implementation errors and suggested fixes," RFC 1536, October 1993.