DNS Privacy

Internet Week 2017 DNS Day

東京大学総合文化研究科

石原知洋

本日の内容

- DNS のプライバシーとその背景
- IETF での動き
- 標準化された機能
- 現在の議論

DNS のプライバシーとは？

- 端的に言えば、「DNS クエリの内容を他人に知られない」こと

- さまざまな実現手法が提案されている
  - クエリに含まれる情報の最小化
  - 暗号化

DNSの情報から何がわかる？

- DNSから分かる情報は、クライアントが「どこと通信しようとしているか」
  - どの web ページを見ようとしているか
  - どこにメールを送ろうとしているか
  - その他、クライアント上のアプリケーションがどこと通信しているか

- 通信そのものの情報ではない
  - 実際の通信は暗号化通信で守られる
  - しばらくの間「そこまでの機密性はない」と手付かず

DNS プライバシーの背景

- ・・・・が、実際に DNS を広域で解析している某組織があることが判明
  - 2013年の事件
- IETF でも Pervasive surveillance, Monitoring が話題に（IETF88）
  - RFC7258: Pervasive Monitoring Is an Attack が発行
- 2014 年に dprive wg 開始
  - RFC7626: “DNS Privacy Considerations” が発行

Morecowbell / QuantumDNS

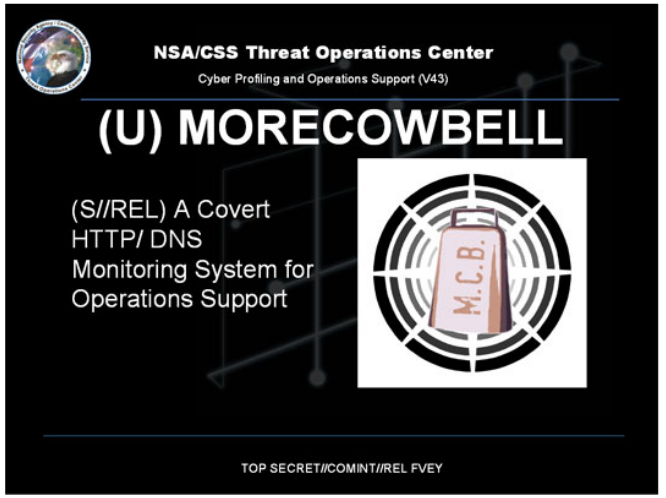
- 受動・能動での DNS 情報収集システム
- RFC7626 にも「Actual Attack」の項目に名指し

dward Snowden, on croyait
et par l'agence de
Security Agency (NSA). Or

u consulter un nouveau lot
s'attaque de façon massive
, qui gère les répertoires de

requête vers un nom de
e » indispensables, reçoivent
es formulées en langage
monde.fr »), puis ils trouvent
les machines

es possèdent leurs propres
es noms sont toujours
vec les grands « serveurs
ntralisent les répertoires pour
es de serveurs racine. Ils

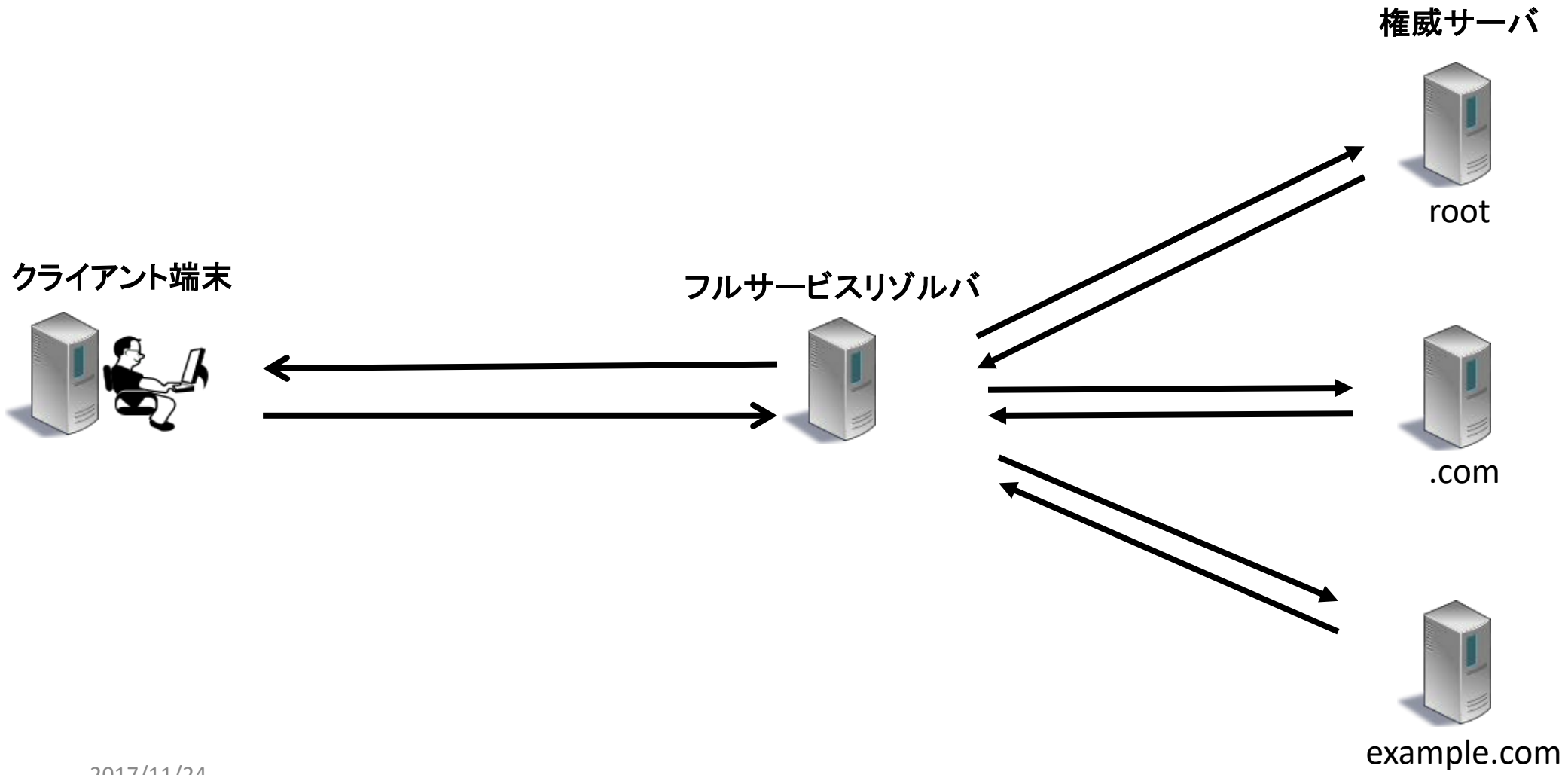


http://www.lemonde.fr/economie/visuel/2015/01/24/cowbells-nouvelles-revelations-sur-les-pratiques-de-la-nsa_4561547_3234.html

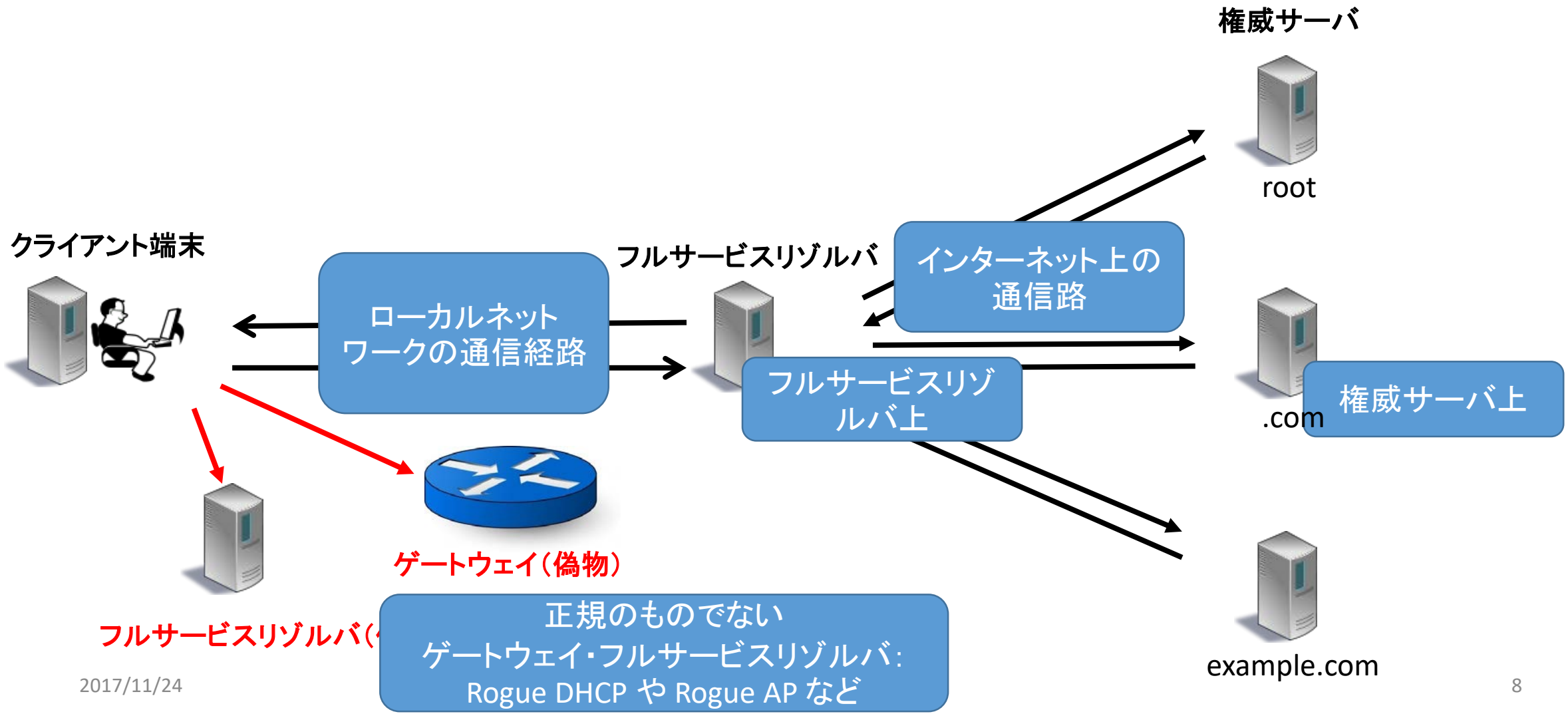
QUANTUMDNS	• DNS injection/redirection based off of A Record queries. • Targets single hosts or caching name servers.	Dec 2008	Operational	**Successful** (High priority CCI target exploited)

https://snowdenarchive.cjfe.org/greenstone/collect/snowden1/index/assoc/HASH2b8f.dir/doc.pdf

DNSの情報はどこで傍受・取得されるのか



DNSの情報はどこで傍受・取得されるのか



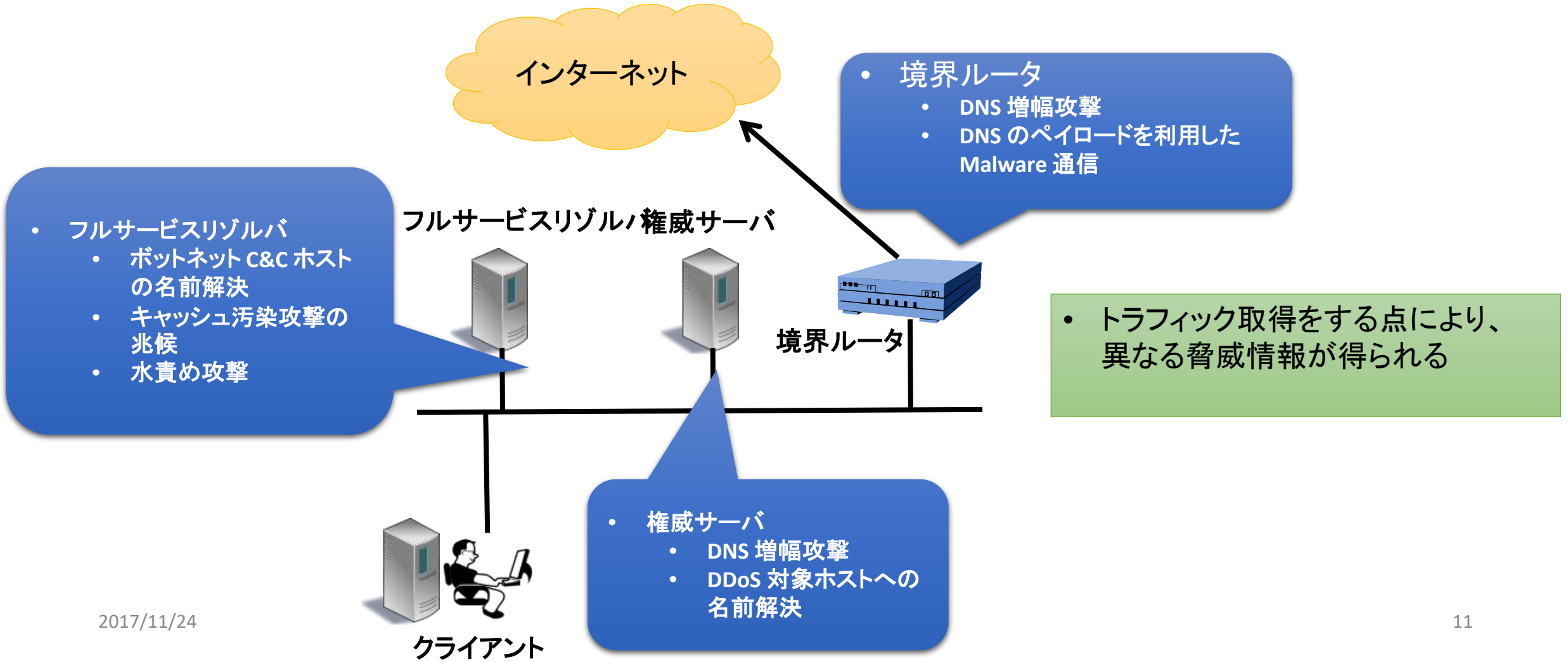
誰が DNS の情報を欲しがっているのか

- 悪意を持つ攻撃者
- xSP
  - 運用上の理由
  - マーケティング上の理由
- 国家機関
  - 情報収集
  - 検閲

収集点により異なる情報レベル

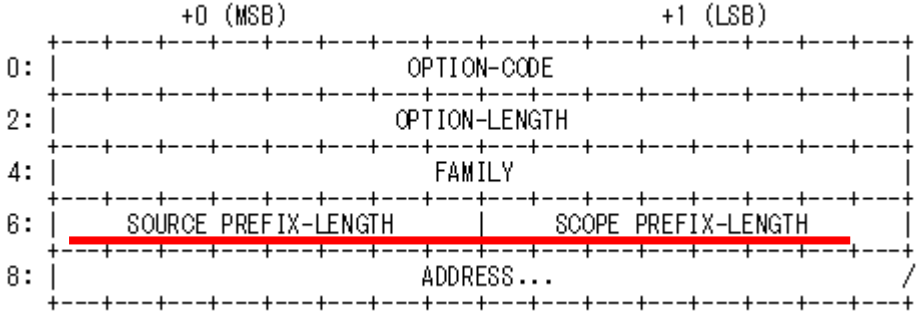
- ローカルネット・フルサービスリゾルバ上での収集（偽物含む）
  - エンドユーザの問い合わせ内容
  - エンドユーザの IP アドレスと対応できるので、最も多くプライバシー情報を含む
- インターネット上の通信路・権威サーバ上での収集
  - フルサービスリゾルバからの問い合わせ内容
  - フルサービスリゾルバで集約され、またキャッシュもあるため、エンドユーザと問い合わせが一対一で対応される場合は少ない
- フルサービスリゾルバのユーザが多ければ多いほど「外から見た時の」匿名性が大きくなる
  - しかし、そのフルサービスリゾルバ自体には大量のプライバシー情報が蓄積される

DNSトラフィック解析点により得られる脅威情報



EDNS(0) Client Subnet(ECS)

- RFC7871
- フルサービスリゾルバから権威サーバへの DNS の問い合わせに、そのクエリの元となる再帰問い合わせ要求を出したクライアントの IP サブネットを入れ込む
  - CDN事業者が、Public DNSを利用しているエンドユーザーのある程度の所在地を把握し、配信の最適化を行うため



- サブネットとはいえ、ある程度クライアント側の情報が権威サーバ、およびインターネット網に出ていく

EIL:
Dealing with the Privacy Problem of ECS

- NDSS 2017 DNS Privacy Workshop での発表
  - CNNIC
  - draft-pan-dnsop-edns-isp-location-03 として ID も出ている

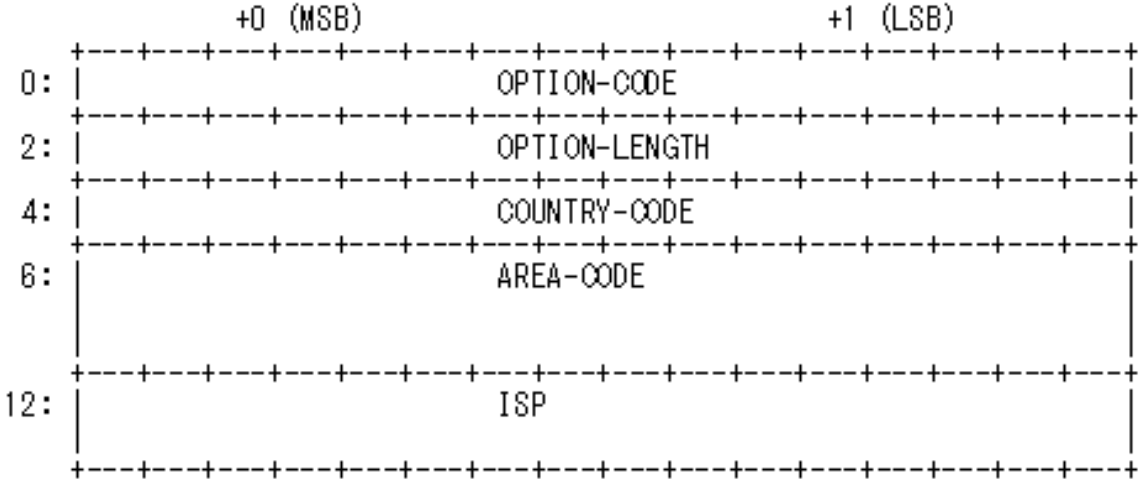
- ECS で、サブネットはわかるが、実はそれほど有用な情報にならないのでは？という問題提起
  - 多重 Forwarding
  - NAT 配下

- サブネットではなく、フルサービスリゾルバがクライアントの IP アドレスを GeoIP で地名データにして、その情報を埋め込む

EIL: Format

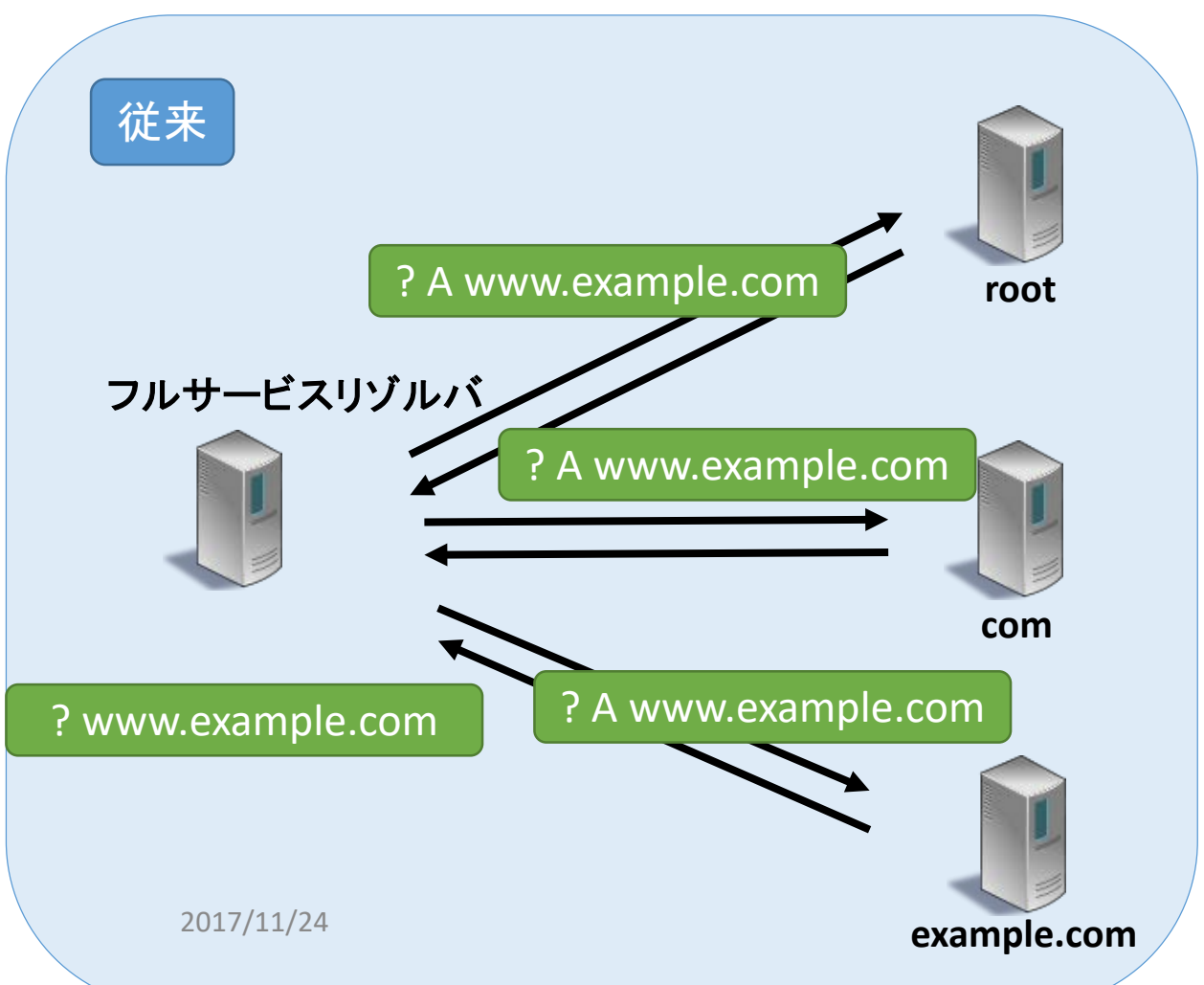
- COUNTRY-CODE
  - ISO3166 の country code
- AREA-CODE
  - 同ISO3166 の country subdivision code
  - 中国:省レベル　日本:県レベル
- ISP
  - 文字列
  - Country-Code 内でユニークであればいい

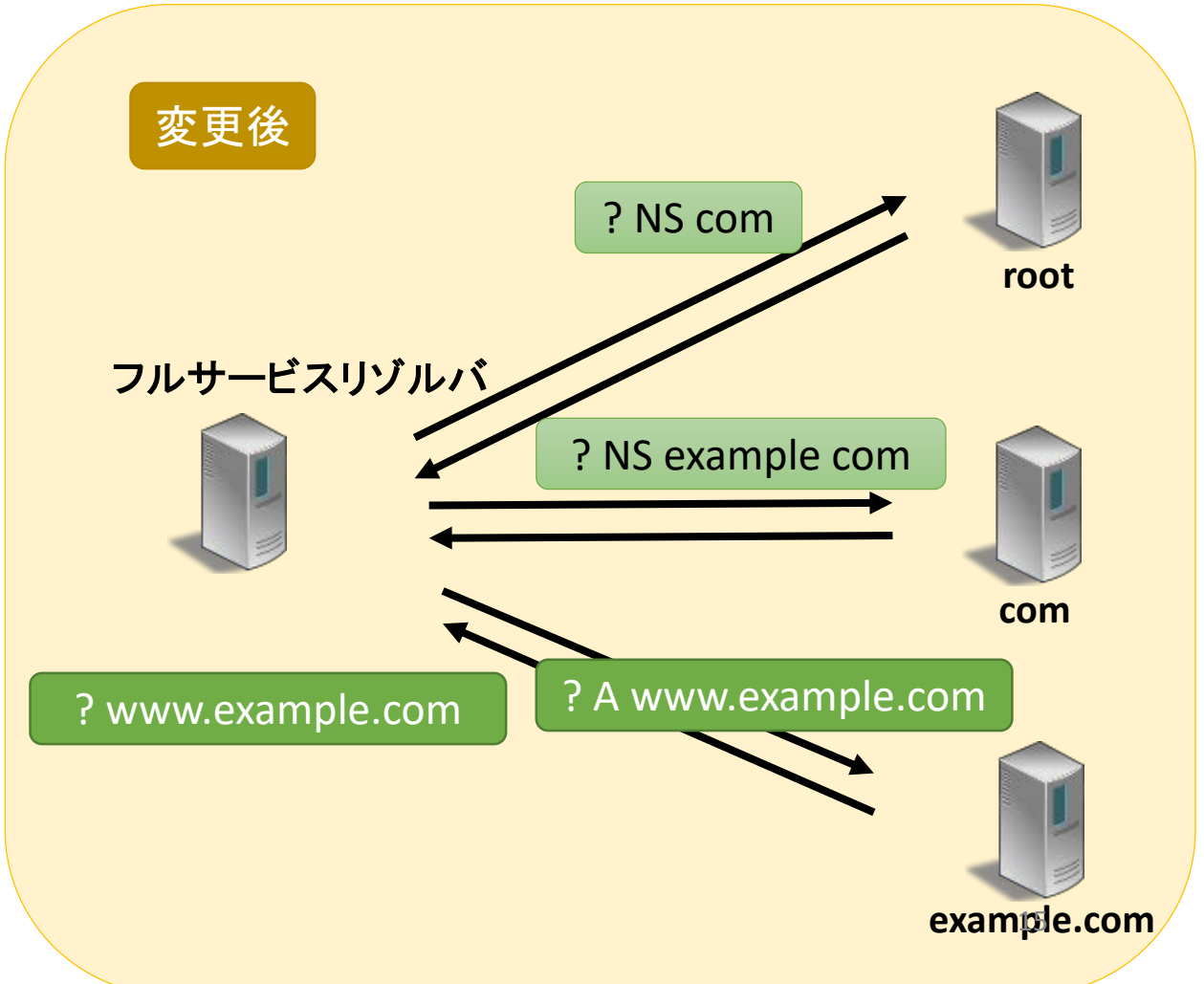
- EIL の方がサブネットより実質的には機微情報が少ない、という議論も
  - わかって ISP と県レベル



Qname Minimization

- 権威サーバに送られる情報を最小化





IETF dprive WG

- 2014/10/17 に設立
- RFC7258 Pervasive Monitoring Is an Attack の内容を受ける形
- DNS に機密保持の機能を追加することがミッション
  - DNS の機密とはなんぞや、というのが RFC7626 : DNS Privacy Considerations で定義

- DNS のプライバシー機能についての RFC を発行
  - EDNS(0) padding Option（RFC7830)
  - DNS-TLS（RFC 7858）
  - DNS-DTLS（RFC8094）

DNS 通信の暗号化

DNS over TLS

- RFC7858: Specification for DNS over Transport Layer Security(TLS)
- DNS通信の暗号化をおこなうプロトコル
- スタブリゾルバ（エンドホスト）とフルサービスリゾルバの間の通信を暗号化
- 通常DNSとはポートを共有しない（TCP853）、プロトコルレベルでのfallback もしない（強制的に unsecure なものを使わされる攻撃などを防ぐ）

- ブートストラップ問題：フルサービスリゾルバをどのように信用するか
  - Opportunistic Privacy Profile
    - DHCP などで配られた情報と、リゾルバから来た公開鍵をとりあえず信じる
  - Out-of-Band Key-Pinned Privacy Profile
    - 別の方法で配られた鍵を信用し、それ以外は信じない
    - 鍵の Rollover などはまた別途考える必要がある（RFC7858 のスコープ外）

DNS over TLS の使い方

- フルサービスリゾルバ実装
  - Unbound
  - Knot resolver
- クライアント実装
  - getdns(API)
  - stubby(ローカルフルリゾルバ)

getdns
Unbound security

DNS Curve

- DNSSEC とは異なる DNS のセキュリティ拡張
- D. J. Bernstein により提案
- DNSSEC と異なる点
  - リゾルバと権威サーバの間の通信を暗号化する
  - レコードレベルでの正当性検証ではなく、サーバレベルでの正当性検証をおこなう
  - スタブの存在は想定しない（＝エンドホストが再帰問い合わせ・検証を行う）
- 暗号化されるため、引いた名前と対応するレコードは隠される

DNS Curve：問題点

- サーバ側での計算資源が必要
- 権威サーバと外部ネットワークにエンドホストの IP アドレスがわかってしまう
  - DNS 的には、権威サーバの IP アドレスがわかれば、どのドメインの名前を引いているかというのはわかってしまう
  - 少なくとも dmm.com の権威サーバに行ってるからFXかソシャゲか動画見てるか英会話教室に行ってる
  - 権威サーバが多数のドメインをサービスしてれば情報量は薄まるかもしれない。。。
- DNS over TLS でリゾルバが複数ユーザを集約するのと、どちらがプライバシー的に許容できるのか

DNS Privacy through ~~Mixnets~~ Onions and Micropayments (DNS over Tor)

- DNS のトラフィックを Tor を通すという大胆な提案
  - NDSS 2017 DNS Privacy Workshop での発表

- DNS のプライバシー機能により、ある程度の改善した
  - DNS over TLS によりスタブリゾルバとフルサービスリゾルバ間の通信は守られる
  - Qname minimization により、通信路で漏洩する情報は減る
  - DNS Curve に類する物を使えば、フルサービスリゾルバと権威サーバ間も暗号化できそう
- しかし、フルサービスリゾルバが single point of failure となりえる
  - クライアントの IP アドレス、問い合わせ履歴、結果、全ての情報を持っている

- だったらいっそ Tor を使ってみては・・・・という提案

ゾーン情報の秘匿
Zone Confidentiality

ゾーン内容の機密性

- DNS はもともとパブリックなシステム
- しかし・・・・
  - ホスト名・IPアドレス
  - 攻撃の足がかりに使われる可能性
- あまりゾーンの内容をおおっぴらにしたくはない
- 通常の DNS であればドメイン名を知らなければ情報は取れないが・・

NSEC：DNSSECでの不存在証明

- 「あるレコードが NXDOMAIN」であることを証明するための仕組み
- ゾーンをアルファベット順に並べ、それぞれのレコードに「アルファベット順の次のレコード」を示すレコードを付与

aa.example.com IN NSEC ac.example.com .....

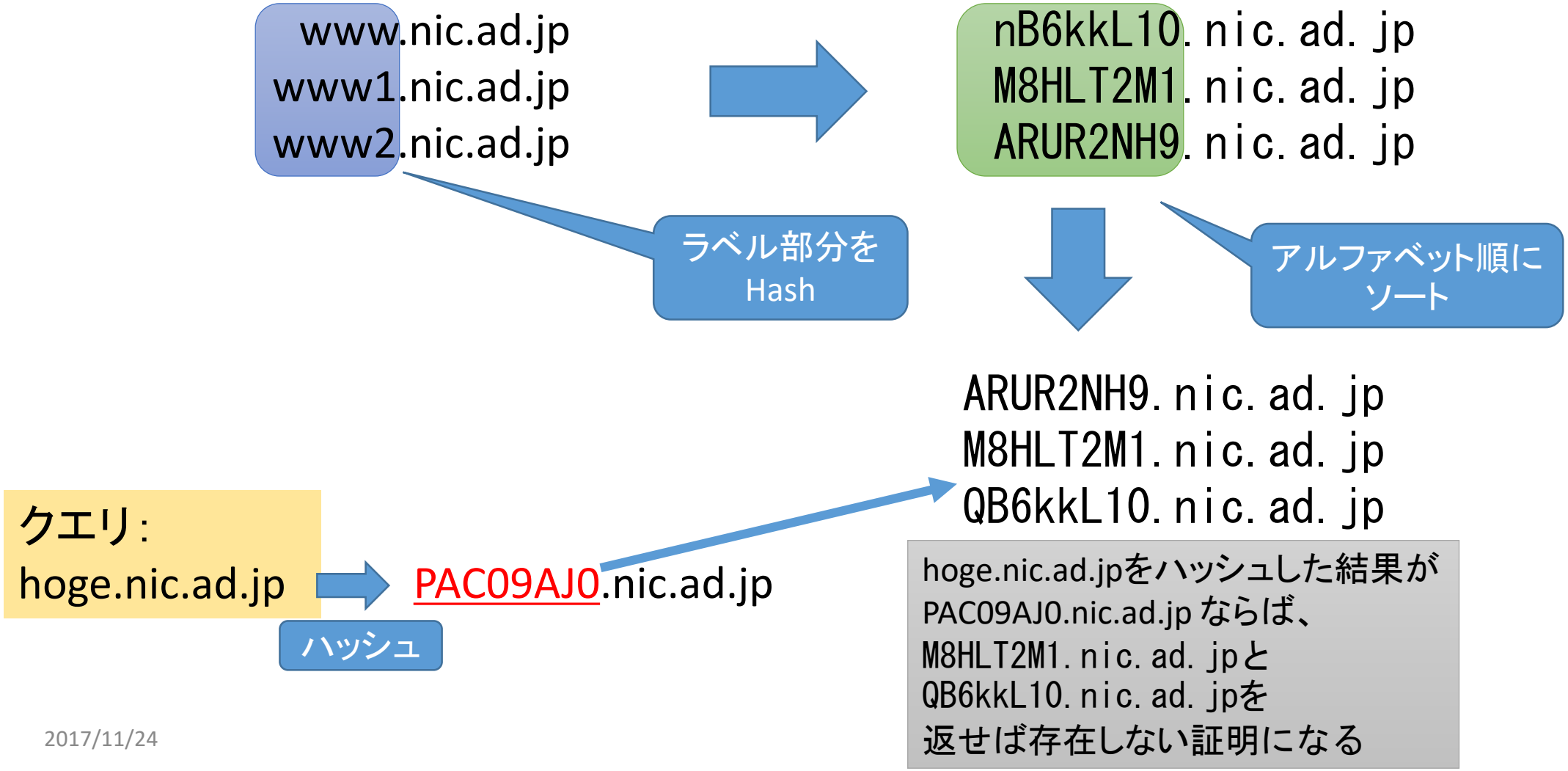
  - この場合、ab.example.com というレコードは存在しない証明になる
  - レコード自体の正当性は DNSSEC の署名で担保

- このレコードを辿っていけば、ゾーン内の名前はすべて列挙できてしまう

NSEC3

- NSEC を名前列挙されないように改良した提案
- 実際の名前を利用するのではなく、一方向ハッシュを使用

NSEC3 の仕組み


www.nic.ad.jp
www1.nic.ad.jp
www2.nic.ad.jp
ラベル部分を
Hash
nB6kkL10.nic.ad.jp
M8HLT2M1.nic.ad.jp
ARUR2NH9.nic.ad.jp
アルファベット順に
ソート
ARUR2NH9.nic.ad.jp
M8HLT2M1.nic.ad.jp
QB6kkL10.nic.ad.jp
クエリ:
hoge.nic.ad.jp
ハッシュ
PAC09AJ0.nic.ad.jp
hoge.nic.ad.jpをハッシュした結果が
PAC09AJ0.nic.ad.jp ならば、
M8HLT2M1.nic.ad.jpと
QB6kkL10.nic.ad.jpを
返せば存在しない証明になる

NSEC3とドメイン名列挙攻撃

- NSEC3 は辞書攻撃されうる
  - Breaking DNSSEC(D. J. Bernstein)
  - NSEC3 でも「ハッシュは」列挙できる
  - ハッシュのリストがあれば「一般的にありえそうなラベル」の辞書をハッシュにかけて推測可能
    - その辞書の名前を直接引いても同じことができるが、列挙に必要なクエリ数は劇的に減る

- 列挙されない仕組みとして NSEC5 が提案

NSEC5

- DNS 列挙攻撃を防ぐために提案
- DNS 列挙攻撃を防ぎつつ、セカンダリが勝手に嘘の NXDOMAIN を返せないようにする

- プライマリ、セカンダリがそれぞれ別の NSEC5 専用の公開鍵、秘密鍵を持つ
- 権威サーバはその場で QNAME を署名し NXDOMAIN を返す
  - NSEC/NSEC3のような、有限の署名で列挙攻撃を防ぐことは数学的に不可能（というテクニカルレポートがある）
  - 当然ながらより多い計算資源が必要

検閲への対応
Censorship Resistance

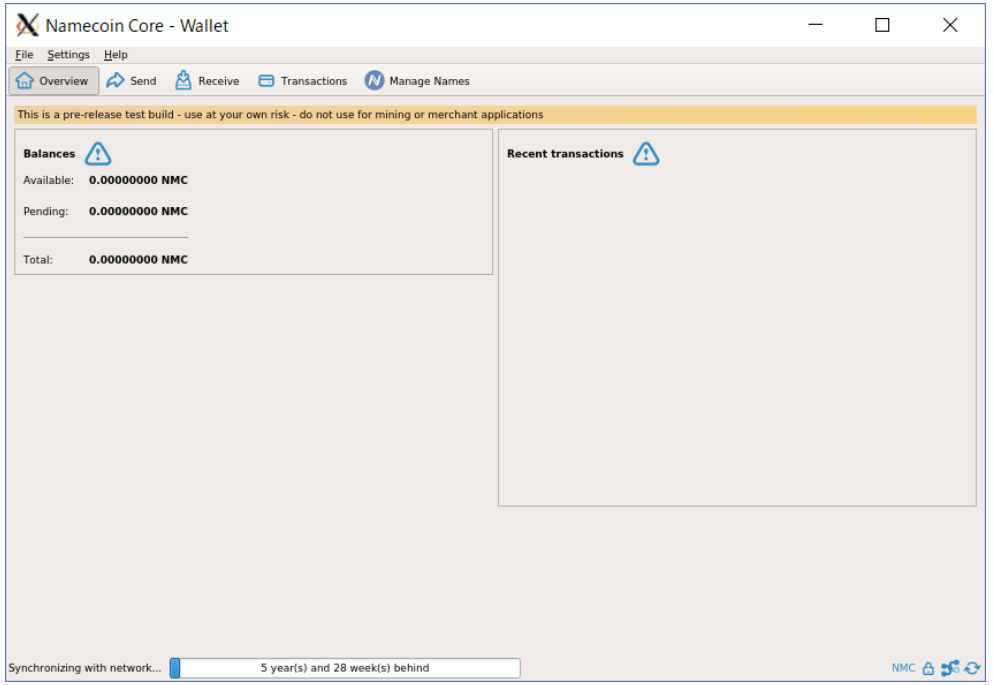
DNS Namecoin

- Bitcoin の技術を使った Key / Value 登録システム
- Non-root・耐検閲（Censorship resistance）
- 登録・変更がトランザクションとなる
  - 手数料として Namecoin(altcoin) を使用
- 名前空間は .bit ドメインを使用

- 名前については First come, First served（早い者勝ち）モデル
  - 登録すると、36000ブロック（約半年ぐらい）の間有効
  - Censorship resistance を謳っており、どこかにピンポイントで介入しても名前のデータを書き換えることは困難

DNS Namecoin の使い方

- 実装
  - 登録・変更
    - Namecoin クライアント
  - 名前解決
    - NMControl（ローカルネームサーバ）
    - パブリック DNS ゲートウェイ
    - ブラウザ拡張
- Namecoin クライアントを実行すると、ブロックチェーンの同期がされ、登録できるようになる
  - 同期には数日単位が必要。。。



ブロックチェーン暗号による実現

- システム全体の構造としては P2P のネットワーク
- 各フルノードが同じブロックチェーンの情報を持つ
  - Proof of Work を示したノードだけが新しいブロックの追加権を持つ

- すなわち、全部のフルノードは namecoin の名前空間に対して完全なアクセスを持っている
  - 列挙はとても簡単、zone confidentiality はない
  - 全部のツリーを保持しているので、フルノードが手元にあれば問い合わせが外に飛んでいかない。問い合わせの秘匿性は高い

- ある意味で先祖返り
  - 全員が同じ Hosts ファイルを同期するという黎明期モデル、大丈夫？
  - 「でも bitcoin は現に今動いているじゃないか」という説得力

DNSプライバシーの今後
研究・運用の変化

DNS プライバシーの技術がもたらす影響

- 暗号化により、DNS トラフィックから得られる情報が減少
  - DNS トラフィック解析による IDS の効力は減る
  - フルサービスリゾルバでログを取る or フルサービスリゾルバ自身に IDS 機能
  - 「DNS はプライバシーが守られて当然」の世界になれば、フルサービスリゾルバのログの取り扱いにも注意が必要（保存期間・匿名化）

- QNAME Minimization により、ツリー上位権威サーバでの DITL などで得られる情報量の低下
  - 悪性ドメインを対象とするクエリの時間・空間分布の解析など、既存の解析手法の一部は不可能に

普及の問題

- DNS-TLS は実装はそろっており、小規模であれば導入は容易
  - しかし、DNS 通信が TLS か平文かユーザは（おそらく）気にしない
  - 今日ではユーザのネットワーク環境はコロコロ変わる
  - 現在守られているか、いないのか？それをどうやってユーザに伝えるか？
  - 「守られてなかったら通信しない」まで行けるのか
  - Opportunistic Privacy Profile って本当に安全？

- DNSCurve / DNSCryptoは DNSSEC で苦労してきた分と
  似たような苦労がもう一度必要

まとめ

- DNS のプライバシーとその背景
- IETF での動き
- 標準化された機能
- 現在の議論