NII市民講座
ネットワークのセキュリティ

国立情報学研究所
アーキテクチャ科学研究系
福田 健介 (kensuke@nii.ac.jp)

http://www.flab.nii.ac.jp/~kensuke/index-j.html

今日の話題

- はじめに (インターネットのおさらい)
- ネットワークスキャンと脅威
- スキャン検出手法
- IPv4 vs IPv6
- 最近の研究の紹介

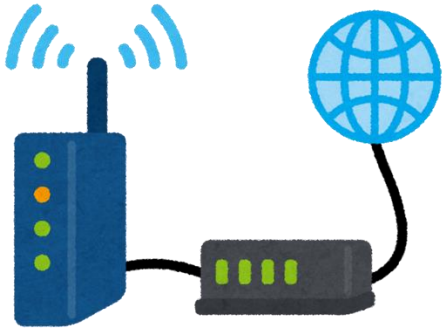
背景

- インターネットは重要なインフラの一つ
  - 何億もの機器 (パソコン, スマートフォン,...)が接続
- インターネットにおけるさまざまな脅威
  - DDoS (分散サービス不能攻撃)
  - ランサムウェア
  - フィッシング

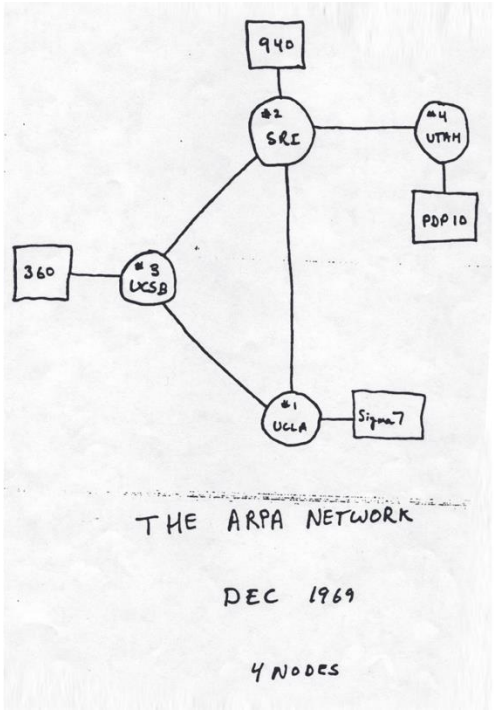
- 今日の発表ではネットワーク上のセキュリティを考えます

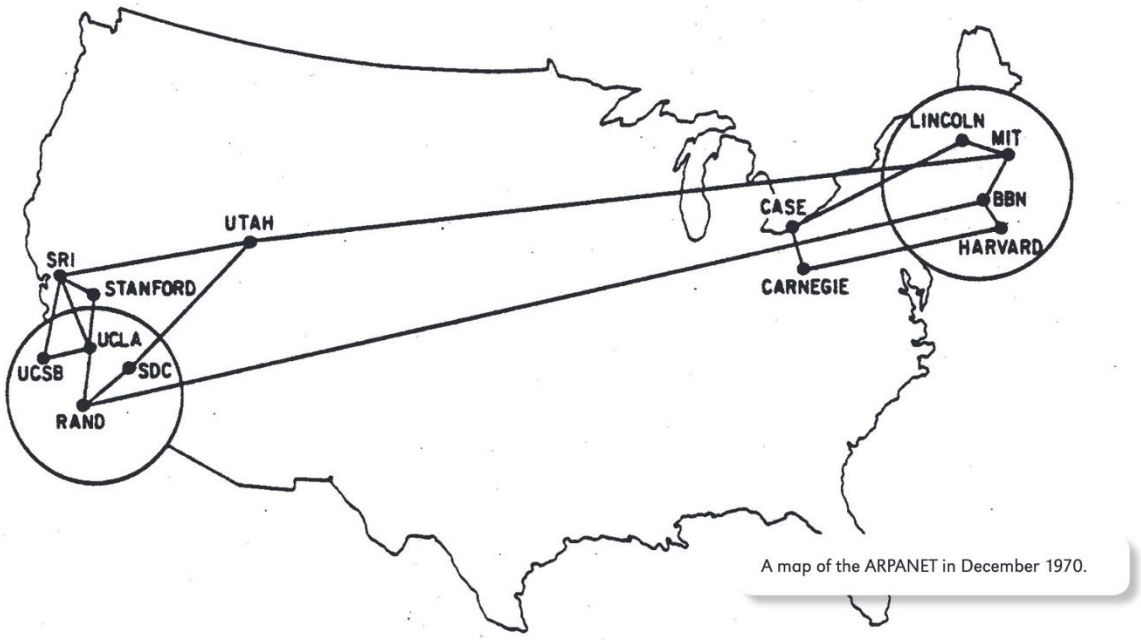
おさらい：インターネットとは

- インターネット＝ネットワークのネットワーク
- ネットワークはノード (ホスト, ルータ)とエッジから構成
- 中央集権ではなく自律分散
  - 世界の全容はわからない
- プロトコル (ルール)にしたがって通信が行われる
- データはパケットとして転送



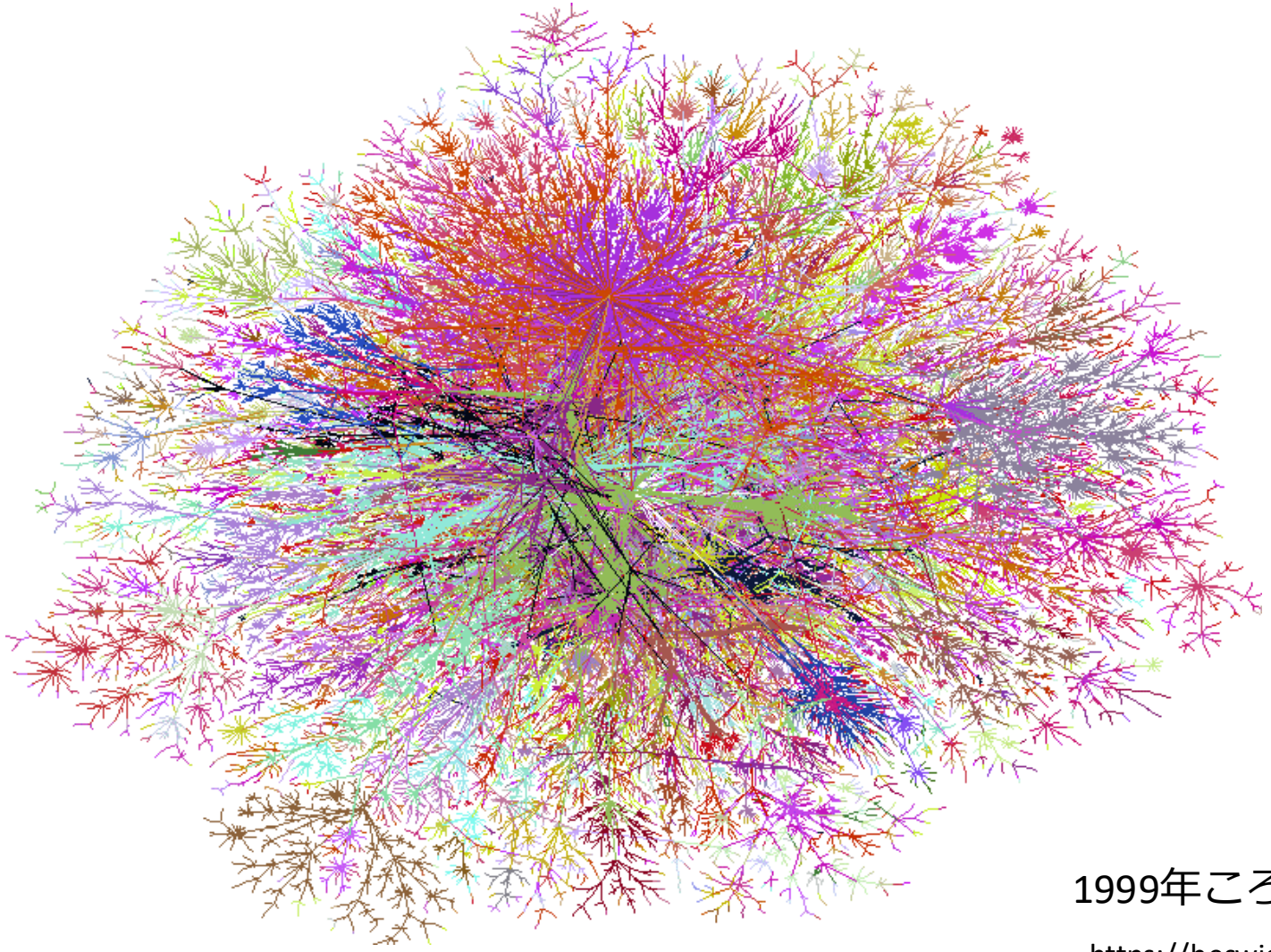
インターネットの発展





https://www.darpa.mil/attachments/(2O15)%20Global%20Nav%20-%20About%20Us%20-%20History%20-%20Resources%20-%2050th%20-%20Internet%20(Approved).pdf

インターネットの発展

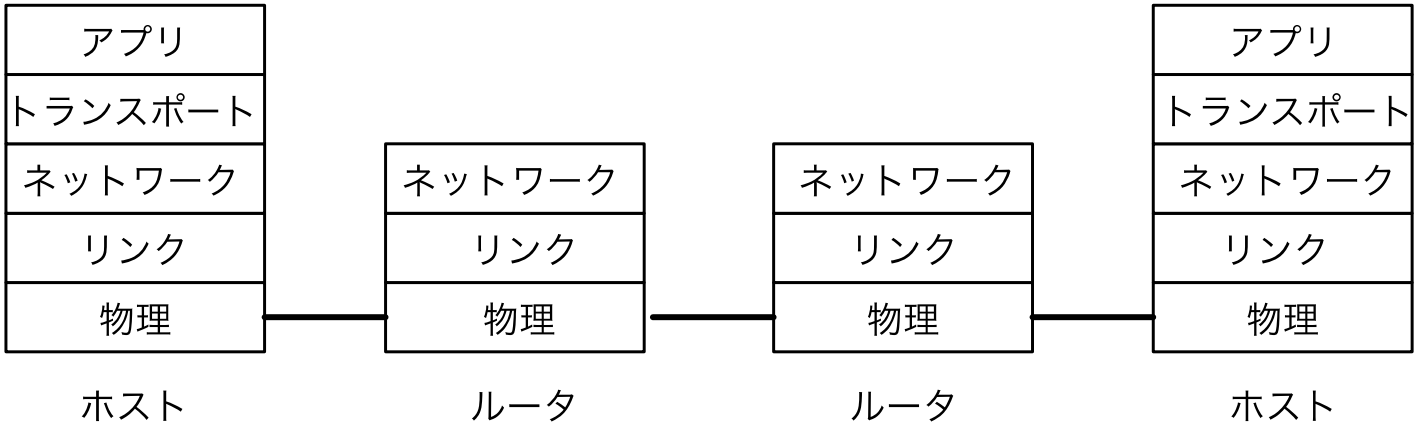


1999年ころ(一部)

https://heswick.com/ches/map/

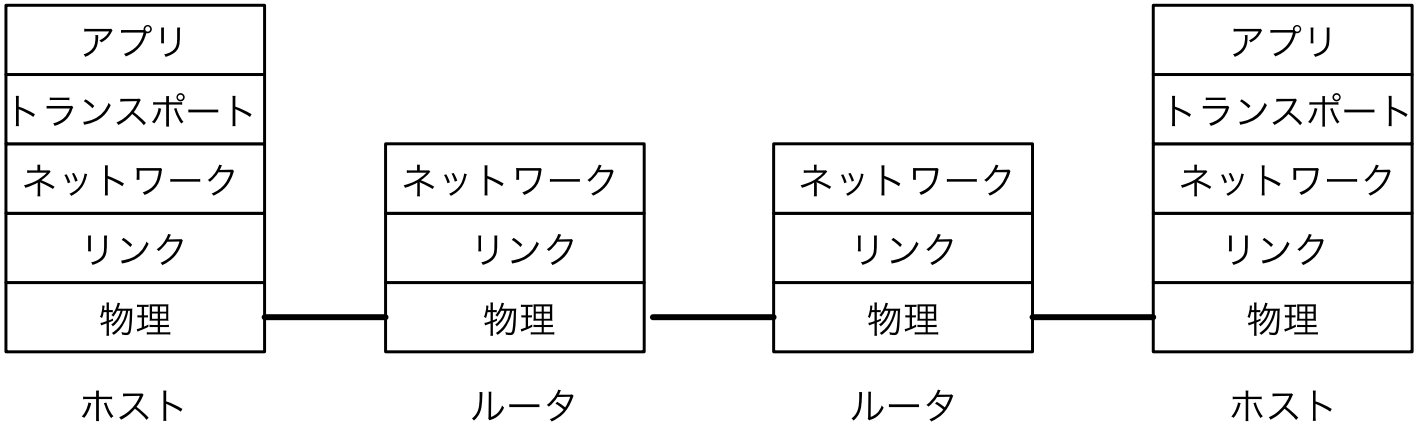
プロトコルとは

- インターネットで通信するための手順 (プロトコル)
  - アプリケーション層: HTTP(S), DNS, …
  - トランスポート層: TCP, UDPプロトコル (エンド間)
  - ネットワーク層: IPプロトコル (ノード間)
  - リンク層: Ethernet (デバイス間)



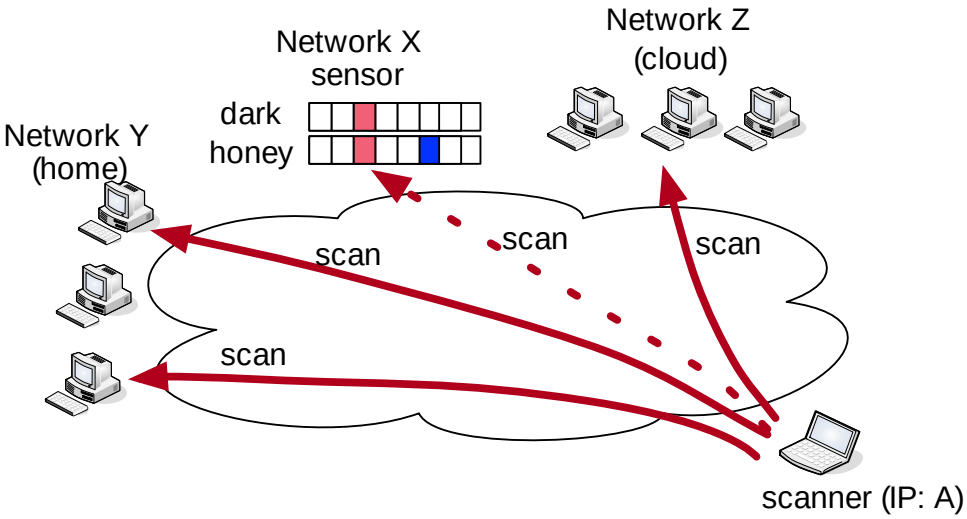
プロトコルとは

- IPアドレス
  - ネットワーク層(IP層)で使われる識別子
  - インターネット上で一意に定まる
  - IPv4: 32bit (e.g., 192.168.2.3)
  - IPv6: 128bit (e.g., 2001:db8::10)



ネットワークスキャン

- ネットワークスキャン: ネットワーク上で動作しているIPアドレス(デバイス(ホストやルータ))を探す試み
- スキャンの目的
  - デバイスへの到達性を調べる
  - デバイスの脆弱性（ぜいじゃくせい)を調べる



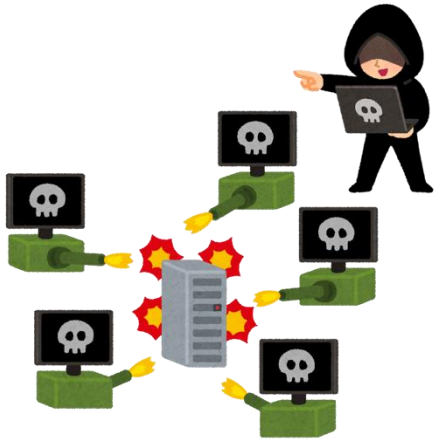
なぜスキャンが重要・問題か?

良くあるシナリオ:

- 新しい脆弱性が発見・情報公開
- 攻撃者がスキャンを開始
- 脆弱性のあるデバイスをリストアップ
- リストにあるデバイスにさらなる調査 (攻撃)
- デバイスの乗っ取り

デバイスを乗っ取られると

- Botnetの一部として操られる
  - 数十万台のデバイスが他のホストを攻撃
  - 知らない間に攻撃に加担
- デバイス内の情報の漏洩



自分の管理するデバイスのセキュリティ対策が重要です

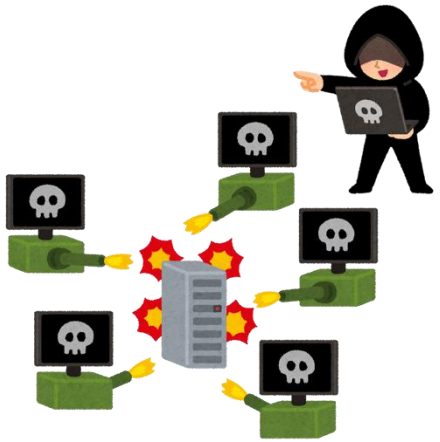
情報セキュリティ10大脅威2024 [組織]

1. ランサムウェアによる被害
2. サプライチェーンの弱点を悪用した攻撃
3. 内部不正による情報漏洩
4. 標的型攻撃による情報漏洩
5. ゼロデイ攻撃
6. 不注意による情報漏洩
7. 脆弱性対策情報の公開に伴う悪用
8. ビジネスメール詐欺
9. テレワーク等のニューノーマルな働き方を狙った攻撃
10. 犯罪のビジネス化

https://www.ipa.go.jp/security/10threats/10threats2024.html

デバイスを乗っ取られると

- Botnetの一部として操られる
  - 数十万台のデバイスが他のホストを攻撃
  - 知らない間に攻撃に加担
- デバイス内の情報の漏洩



自分の管理するデバイスのセキュリティ対策が重要です

でも自分のデバイスには来ないのでは?

全IPv4アドレスをスキャンするには?

- IPv4 アドレス空間 32bit = 43億アドレス (10^9)
  - 2進数で32桁: 11000000000000000000001000000001
  - 192.0.2.1 のように8bit単位で区切って10進数で表現
- 43億アドレスをスキャンするために必要な時間?
  - 数分
  - 数時間
  - 数日
  - 数週間

全IPv4アドレスをスキャンするには?

- IPv4 アドレス空間 32bit = 43億アドレス (10^9)
  - 2進数で32桁: 11000000000000000000001000000001
  - 192.0.2.1 のように8bit単位で区切って10進数で表現
- 43億アドレスをスキャンするために必要な時間?
  - 数分
  - 数時間
  - 数日
  - 数週間

- 答え：数時間
  - 専用ソフトウェア
  - 普通のパソコンと少し速めのネットワーク

スキャンと脅威との関係

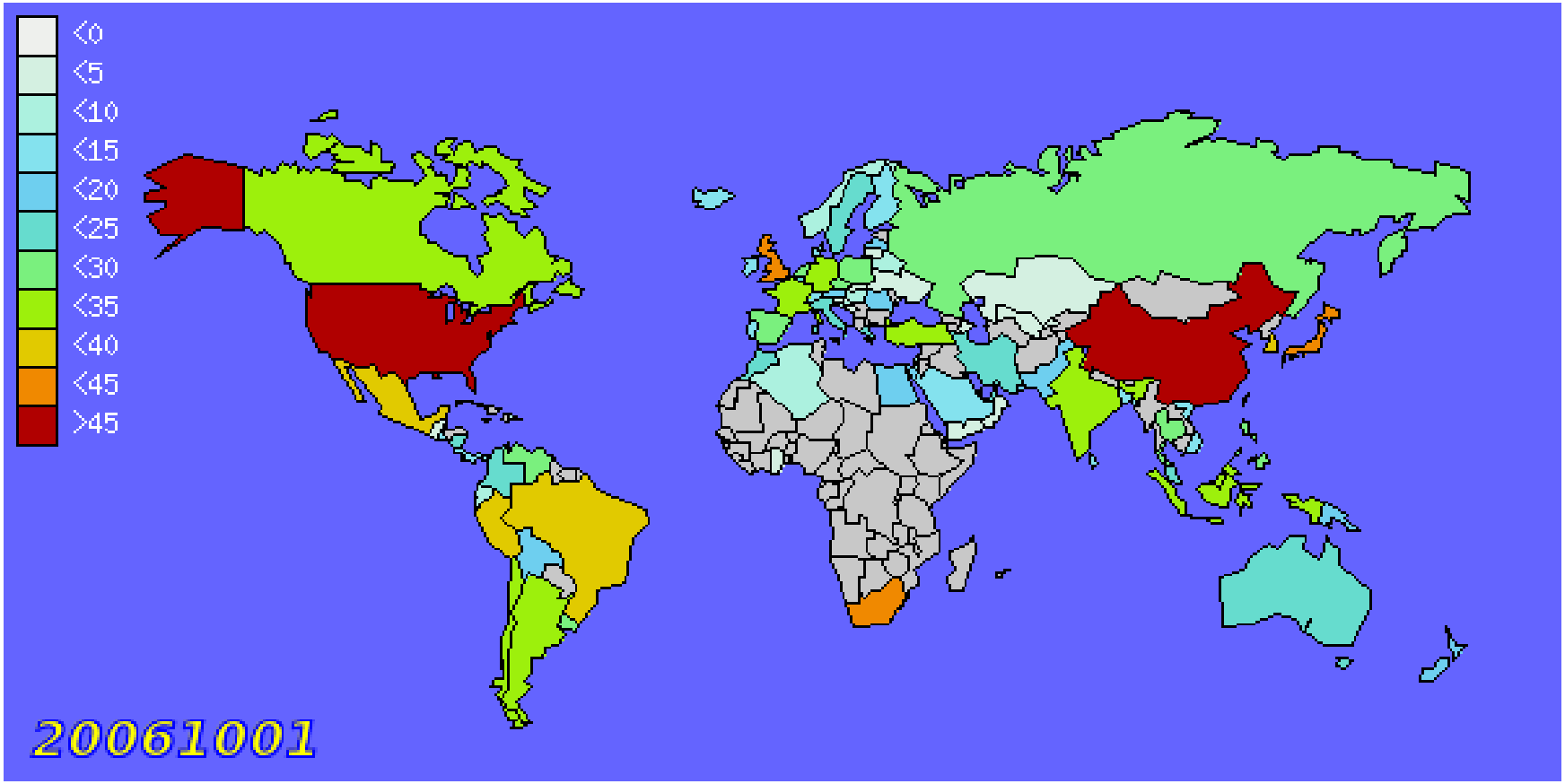
- 新しい脆弱性が発見・情報公開
- 攻撃者がスキャンを開始
- 脆弱性のあるデバイスをリストアップ
- リストにあるデバイスにさらなる調査 (攻撃)
- デバイスの乗っ取り

- 自分は安心と思っていても、、、
- たくさんのスキャンが来ている!

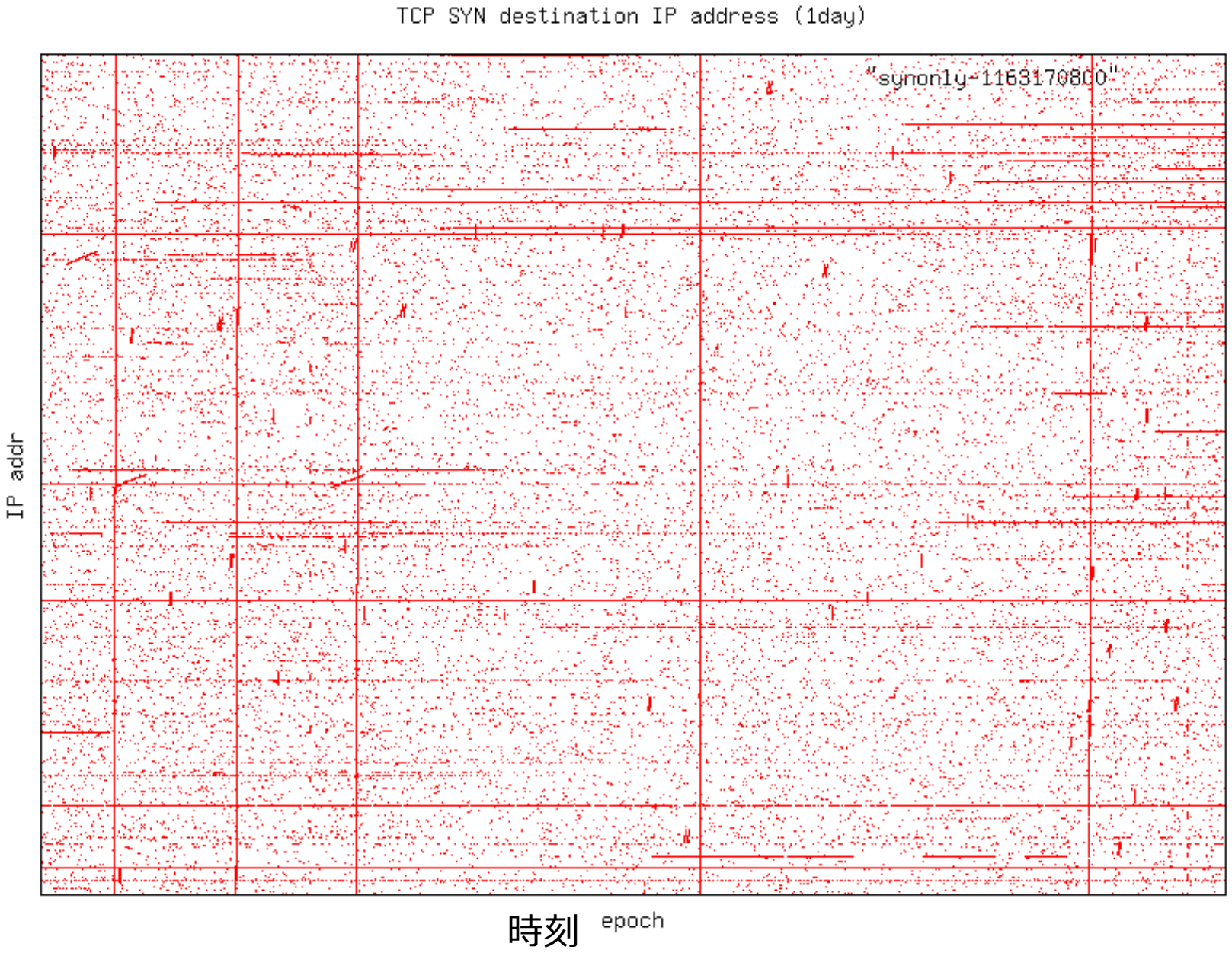
実際のスキャンデータ

- 世界地図
- スキャッタープロット

センサーへのスキャン



- 送り元のIPアドレスを国にマッピング



スキャンの送り先IPアドレス

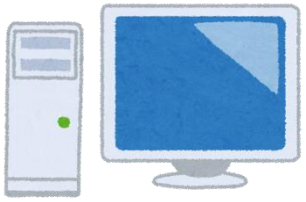
- 各点がスキャンパケット

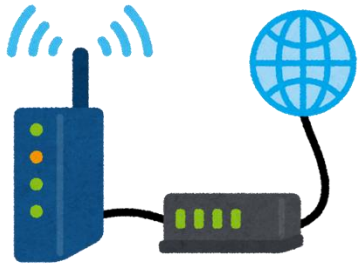
水平方向：同一ホストへスキャン

垂直方向：多数のホストへのスキャン

本日の大事なメッセージ

- 自分の接続している機器の設定見直し
  - ホームルータやパソコンの設定・アップデート
  - 古いホームルータはサポート切れの可能性も
- 境界防御だけでなくネットワーク内のホストにも
  - 多層防御
  - ホストベースのFW (e.g., Windows defender)
- さらにゼロトラストの考え方が導入されつつある
  - 誰も何も信じないモデル
  - リソースのアクセスには必ずチェック





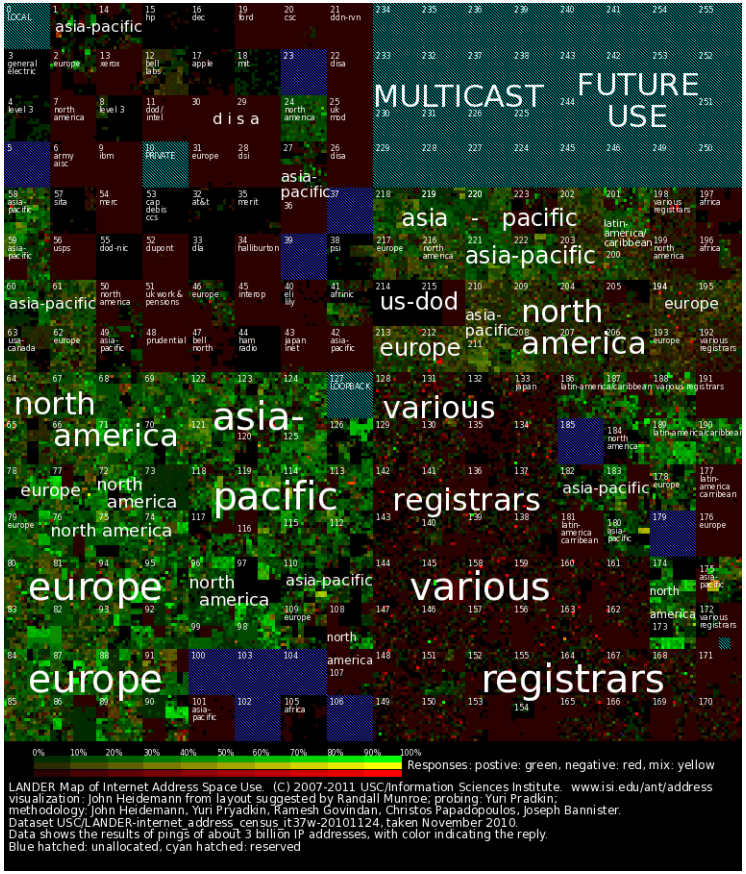
ここからは少し専門的な話

- スキャンの応用
- スキャン検出手法
- IPv6でのスキャン検出

スキャンの応用

定期的なIPアドレス空間へのスキャン

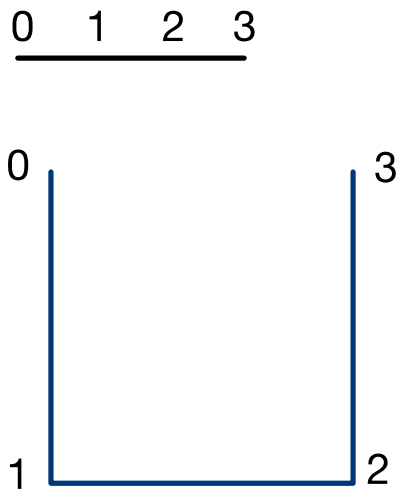
- 脆弱性把握
- アドレス利用状況把握
- ネットワーク接続性確認
  - 自然災害による停電



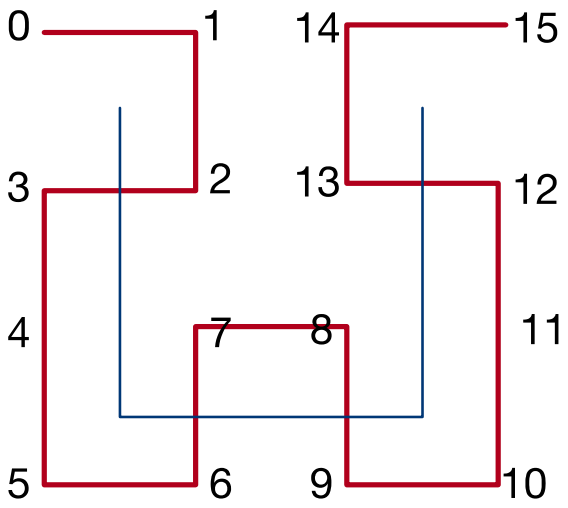
https://ant.isi.edu/address/index.html

参考: ヒルベルト曲線

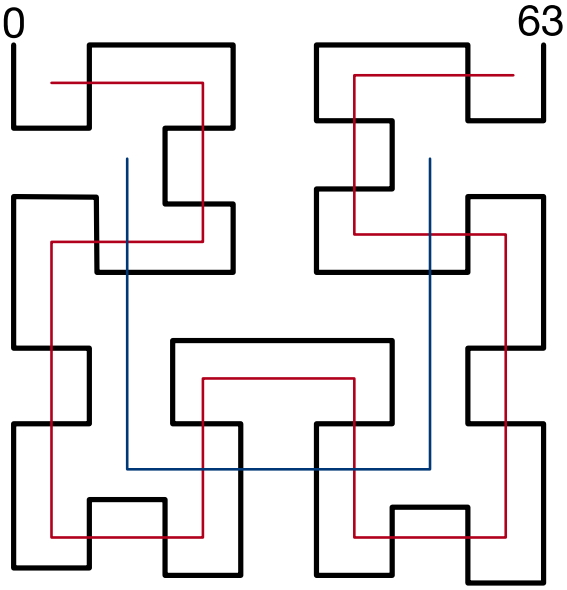
- IPアドレスは32bitの整数 (1次元)
- 1次元データを2次元で表現 (フラクタル)



2bit



4bit



6bit

スキャンの検出

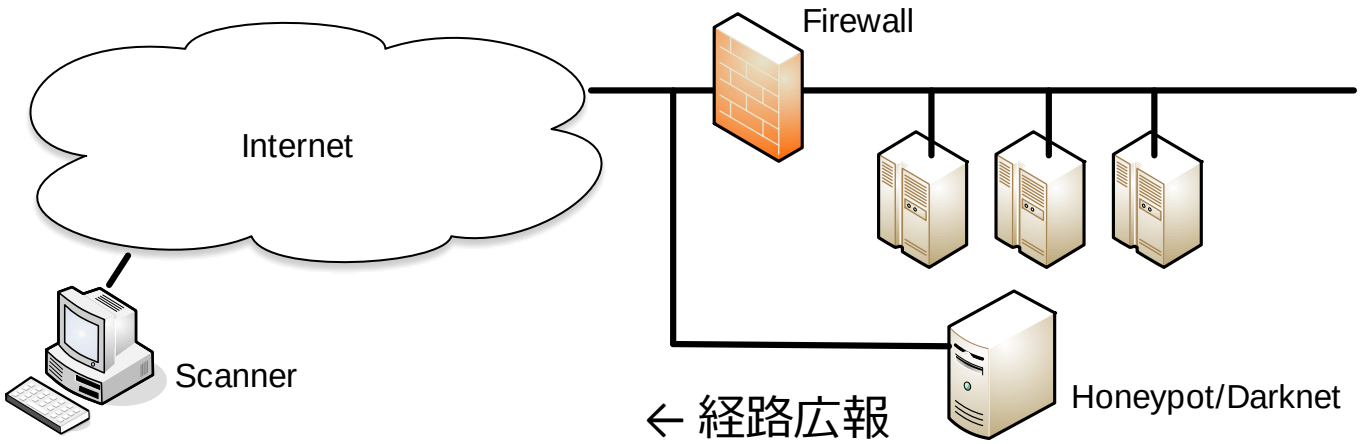
目的：新たな大規模スキャンを検出し早期対処

- どこで?
  - 自分のコンピュータで見張る
  - ネットワークの真ん中(IXP, BB)で見張る
  - 分散されたサーバ(FW)で見張る
- どうやって?
  - 正常な通信以外を見張る
  - 正常と異常の定義とは?

- 専用のセンサー(ネットワーク)
  - ハニーポット、ダークネット

センサーネットワーク: ハニーポット・ダークネット

- ハニーポット: おとりのホスト (応答あり)
- ダークネット: ホストのいないネットワーク (応答なし)



IPv6とは?

- IPv4アドレス43億個は割り当て済み
  - マーケットで売買 (1アドレス 30USD程度)
- IPv6が1990年代後半に標準化
  - アドレスサイズ128bit (2001:db8::1のように表記)
  - IPv4との互換性なし
  - YoutubeやGoogle等のサービスは既にIPv6対応
  - モバイル・ブロードバンドアクセスで普及
    - 日本での普及率: 51.2% (2024年)
- 多くのデバイスはIPv4およびIPv6の両方を利用可 (デュアルスタック)

Google IPv6

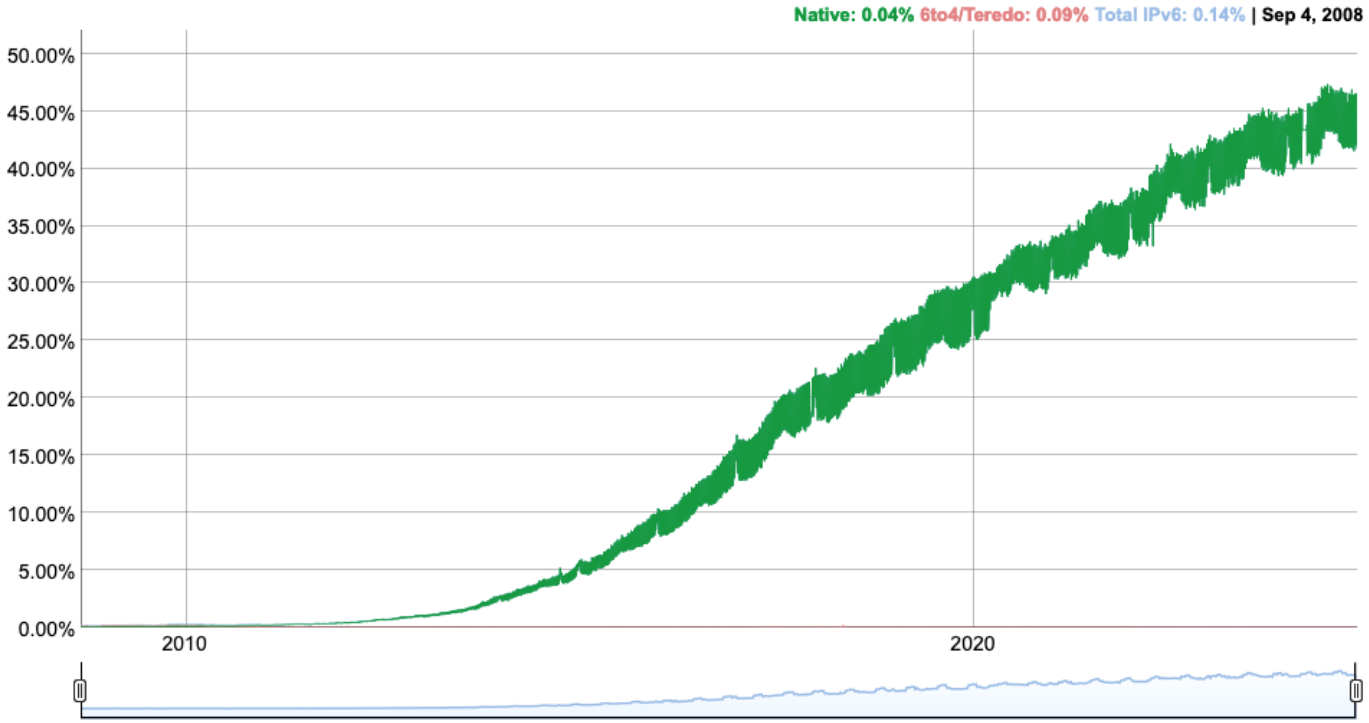
Statistics

Google collects statistics about IPv6 adoption in the Internet on an ongoing basis. We hope that publishing this information will help Internet providers, website owners, and policy makers as the industry rolls out IPv6.

IPv6 Adoption | Per-Country IPv6 adoption

IPv6 Adoption

We are continuously measuring the availability of IPv6 connectivity among Google users. The graph shows the percentage of users that access Google over IPv6.



https://www.google.com/intl/en/ipv6/statistics.html

全IPv6アドレスをスキャンするには?

- IPv6 アドレス空間 128bit = 340澗 (43億 x 43億 x 43億 x 43億; 10^36)
- 340澗 アドレスにスキャンするために必要な時間?
  - 数ヶ月
  - 数年
  - 数十年
  - 数百年

かん

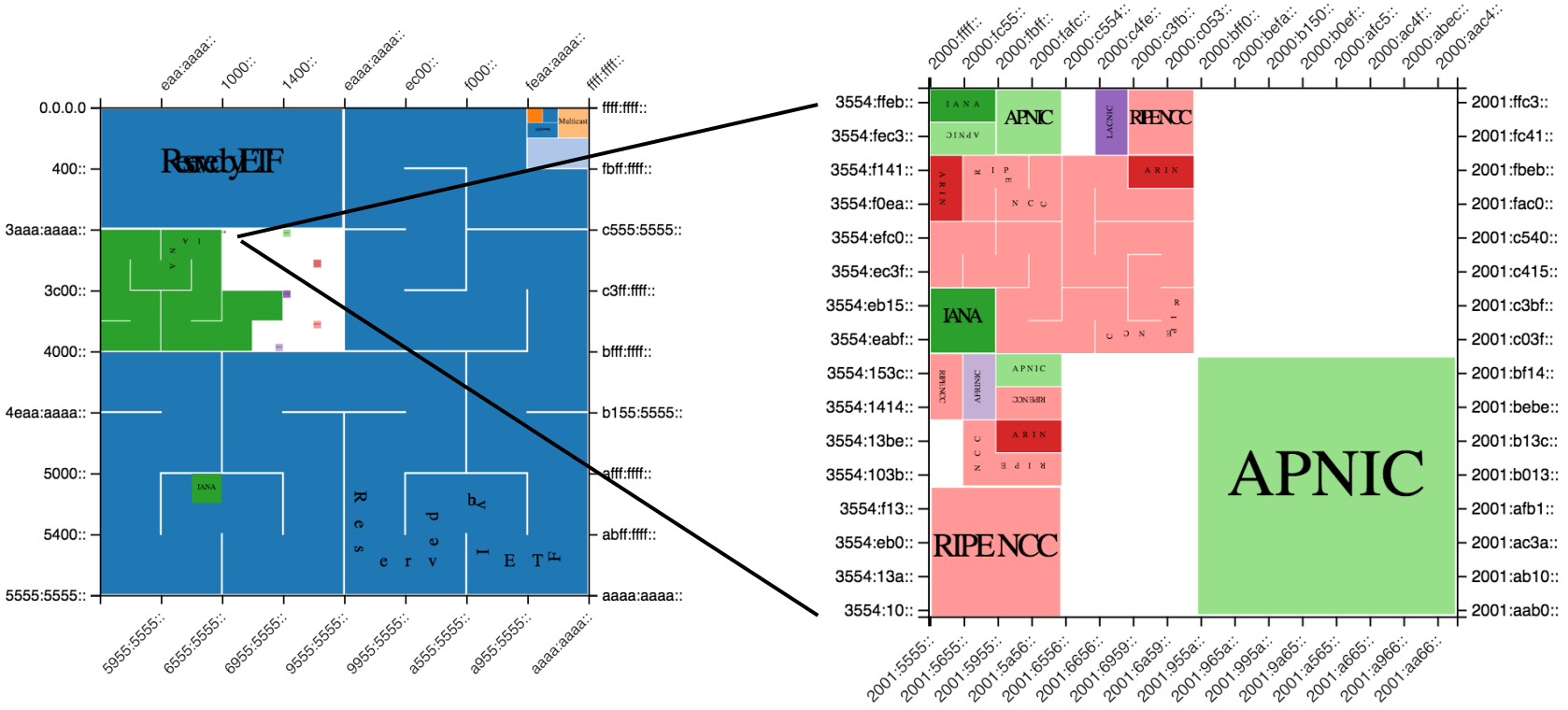
一、十、百、千、万、億、兆、京、垓、秭、穣、溝、澗、正、載、、、

全IPv6アドレスをスキャンするには?

- IPv6 アドレス空間 128bit = 340澗 (43億 x 43億 x 43億 x 43億; 10^36)
- 340澗 アドレスにスキャンするために必要な時間?
  - 数ヶ月
  - 数年
  - 数十年
  - 数百年

- 答え：50億年 (地球の歴史!)たっても終わらない
  - 専用ソフトウェア
  - 普通のパソコンと少し速めのネットワーク

IPv6アドレス空間



https://bl.ocks.org/vasturiano/ba0bdc27ddc2b13012e85b2a0b7a1c43

では、何もしなくても安全?

- ランダムスキャンでホストを検出することは困難
  - アドレスが知られなければ大丈夫(?)
- 到達可能なIPv6アドレスのリスト (ヒットリスト)を用いたスキャン
  - IPv6アドレス探索の各種アルゴリズムが提案
  - 公開ヒットリストも存在

大規模IPv6スキャンは起きてるのか?

- 2018年ころより、複数の手法で大規模スキャン観測を実現
  - 分散ファイアウォール
  - CDN
  - バックボーン/IXP
  - DNSからの推定
  - センサーネットワーク (honeypot/darknet)
- IPv6セキュリティは今後ますます重要
  - 今のところ多くはセキュリティベンダや研究機関
  - IPv4に比べてIPv6ネットワークの設定は無関心

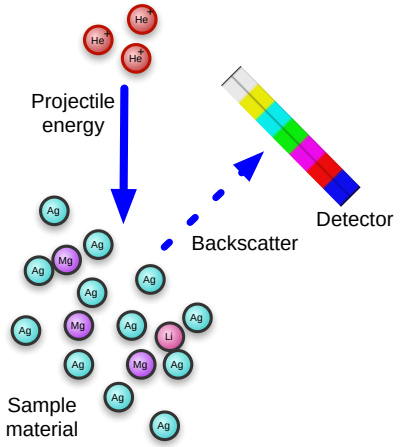
最後に、最近の取り組みについて紹介します

DNSベースのIPv6スキャン検出

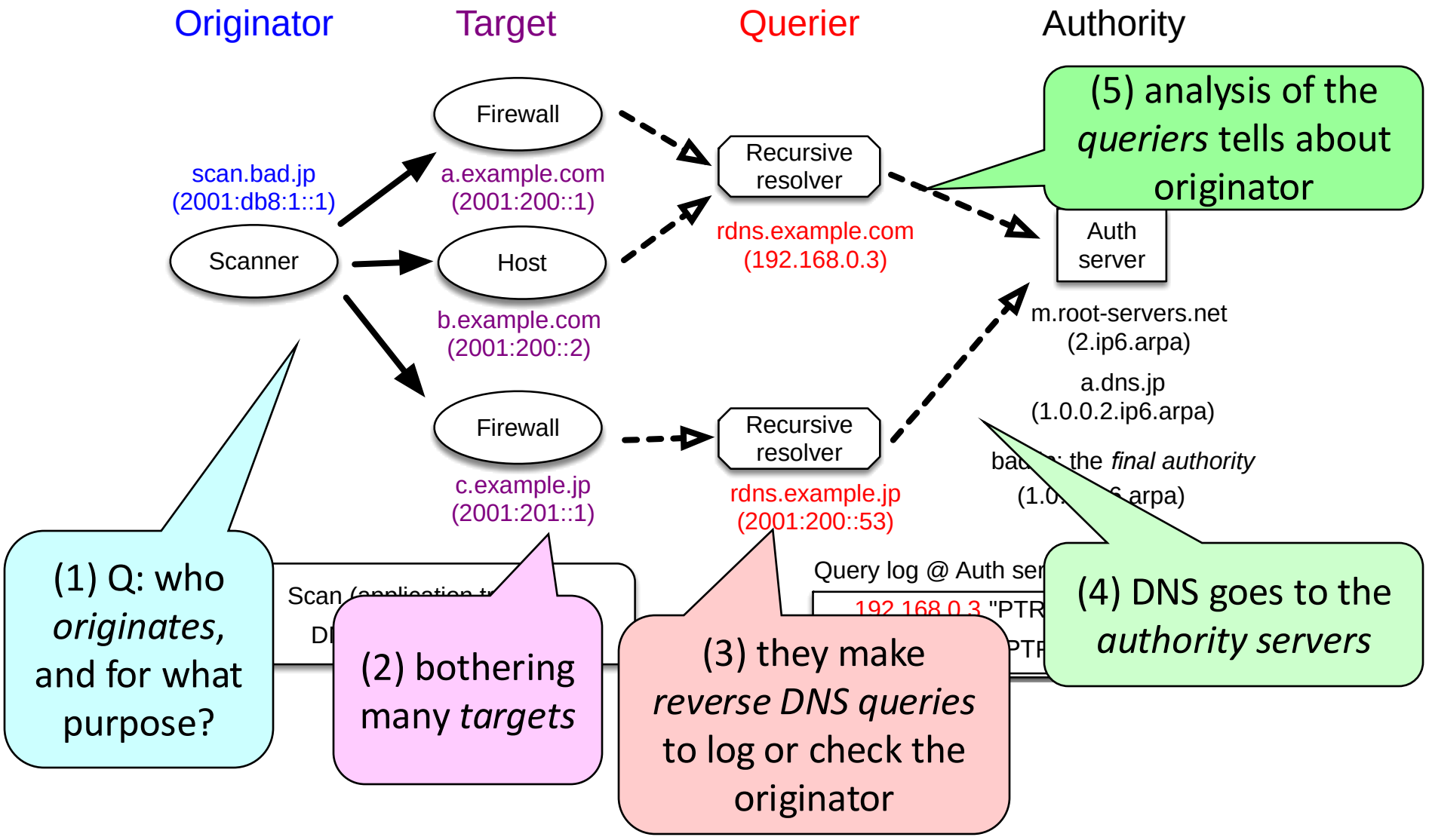
K.Fukuda, J.Heidemann, “Who Knocks at the IPv6 Door? Detecting IPv6 Scanning”, in ACM IMC’18

DNSベースのIPv6スキャン検出

- 問題：世界中で起きているスキャンを検出したい
- キーアイディア：世界中のFWが検出したスキャンをDNSデータから推定
  - FWではスキャンのIPアドレスのホスト名を検索
    - FQDN of 2001:db8:::10?
  - 権威DNSでの集合的な検索クエリからスキャンを推定



スキャンが検出できる仕組み



スキャン検出結果

	IP	MAWI			Backscatter	Dark	ASN	info
		#days	port	scan type	#weeks	#weeks		
(a)	2001:48e0:205:2::/64	6	TCP80	Gen	1 (5)	1	40498	New Mexico Lambda Rail
(b)	2a02:418:6a04:178::/64	2	ICMP	rand IID	2 (4)	0	29691	Nine, CH
(c)	2a02:c207:3001:8709::/64	2	TCP80	rand IID	2 (2)	0	51167	Contabo, DE
(d)	2a03:f80:40:46::/64	2	ICMP	rDNS	2 (3)	0	5541	ADNET-Telecom, RO
(e)	2405:4800:103:2::/64	2	ICMP	rDNS	0 (4)	0	18403	FPT-AS-AP, VN
(f)	2a03:4000:6:e12f::/64	1	ICMP	rDNS	0 (0)	0	197540	NETCUP-GmbH, DE
(g)	2800:a4:c1f:6f01::/64	1	ICMP	rDNS	0 (0)	0	6057	ANTEL, UY

スキャン検出結果

	IP	MAWI			Backscatter	Dark	ASN	info
		#days	port	scan type	#weeks	#weeks		
(a)	2001:48e0:205:2::/64	6	TCP80	Gen	1 (5)	1	40498	New Mexico Lambda Rail
(b)	2a02:418:6a04:178::/64	2	ICMP	rand IID	2 (4)	0	29691	Nine, CH
(c)	2a02:c207:3001:8709::/64	2	TCP80	rand IID	2 (2)	0	51167	Contabo, DE
(d)	2a03:f80:40:46::/64	2	ICMP	rDNS	2 (3)	0	5541	ADNET-Telecom, RO
(e)	2405:4800:103:2::/64	2	ICMP	rDNS	0 (4)	0	18403	FPT-AS-AP, VN
(f)	2a03:4000:6:e12f::/64	1	ICMP	rDNS	0 (0)	0	197540	NETCUP-GmbH, DE
(g)	2800:a4:c1f:6f01::/64	1	ICMP	rDNS	0 (0)	0	6057	ANTEL, UY

DNSセンサーで検出可能

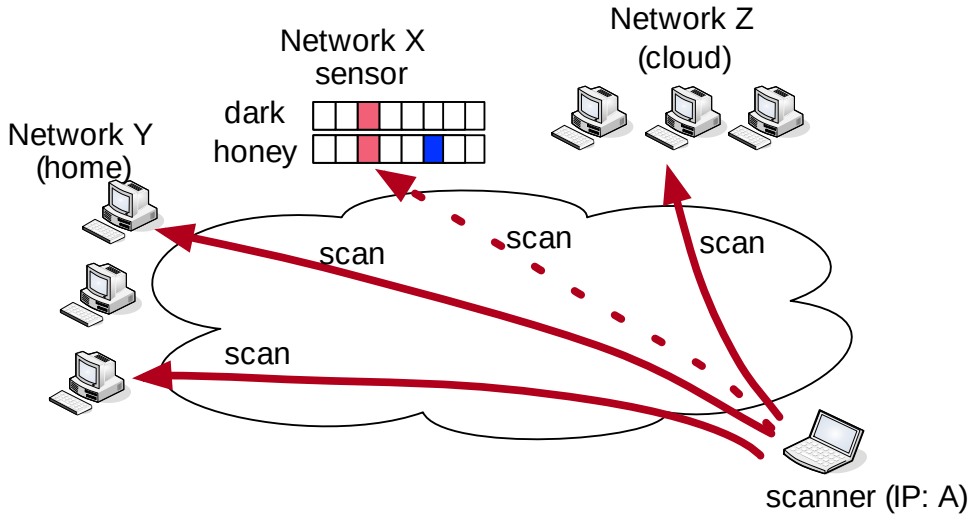
Darknetにはスキャンが来ない

センサーネットワークの改善

L.Zhao, S.Kobayashi, K.Fukuda, “Exploring the Discovery Process of Fresh IPv6 Prefixes: An Analysis of Scanning Behavior in Darknet and Honeynet”, PAM’24

センサーネットワークの改善

- 問題：スキャンがセンサーに到着しない
- キーアイディア：スキャナーのヒットリストにセンサーアドレスを載せる
  - センサーアドレスを積極的にインターネット上に公開
  - どのような方法が有効か?



アドレス公開手法

- IPv4 reverse: 192.168.1.1 -> foo.bar.com -> 2001:db8::10
  - 対応するDNSエントリを追加
- IPv6 enumeration: DNSツリーの探索
- IPv6 special: 2001:db8::1など人が推定しやすいものをDNS登録
- IPv6 popular name: www.bar.com -> 2001:db8::80
- No exposing: 何もしない

どの手法が効率的か?

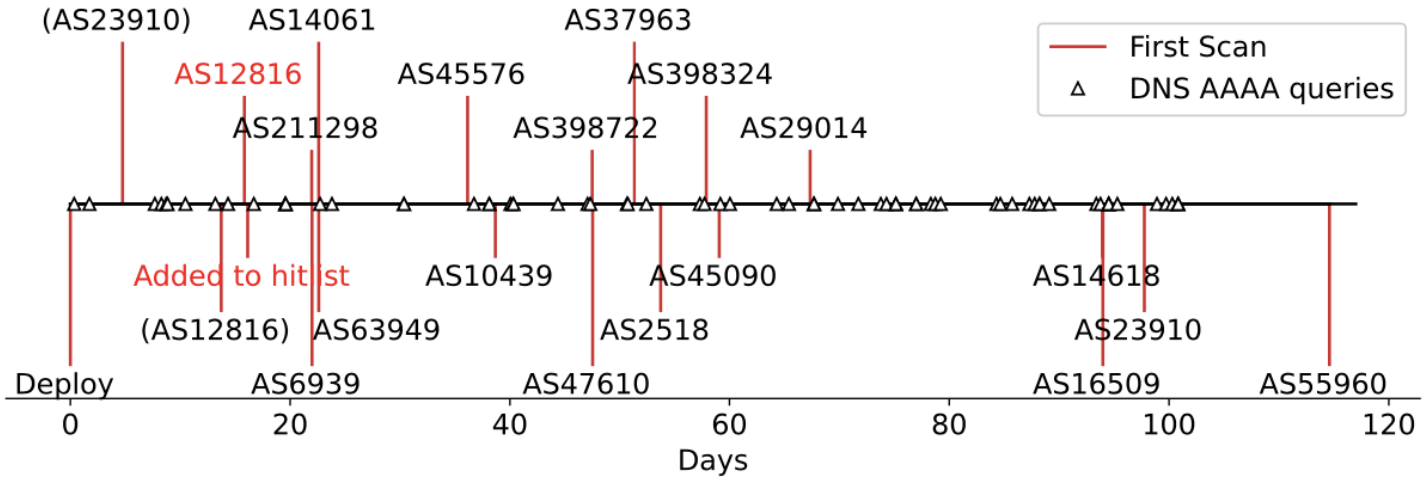
アドレス公開手法

- IPv4 reverse: 192.168.1.1 -> foo.bar.com -> 2001:db8::10
  - 対応するDNSエントリを追加
- IPv6 enumeration: DNSツリーの探索
- IPv6 special: 2001:db8::1などよく推定しやすいものをDNS登録
- IPv6 popular name: [illegible]
- No exposing: 何もし[illegible]

-	Darknet	Honeynet
No exposing	5	1
IPv4 reverse	1.7K	2.0K
IPv6 enumeration	2	2
IPv6 special	1	1
IPv6 popular name	4	2

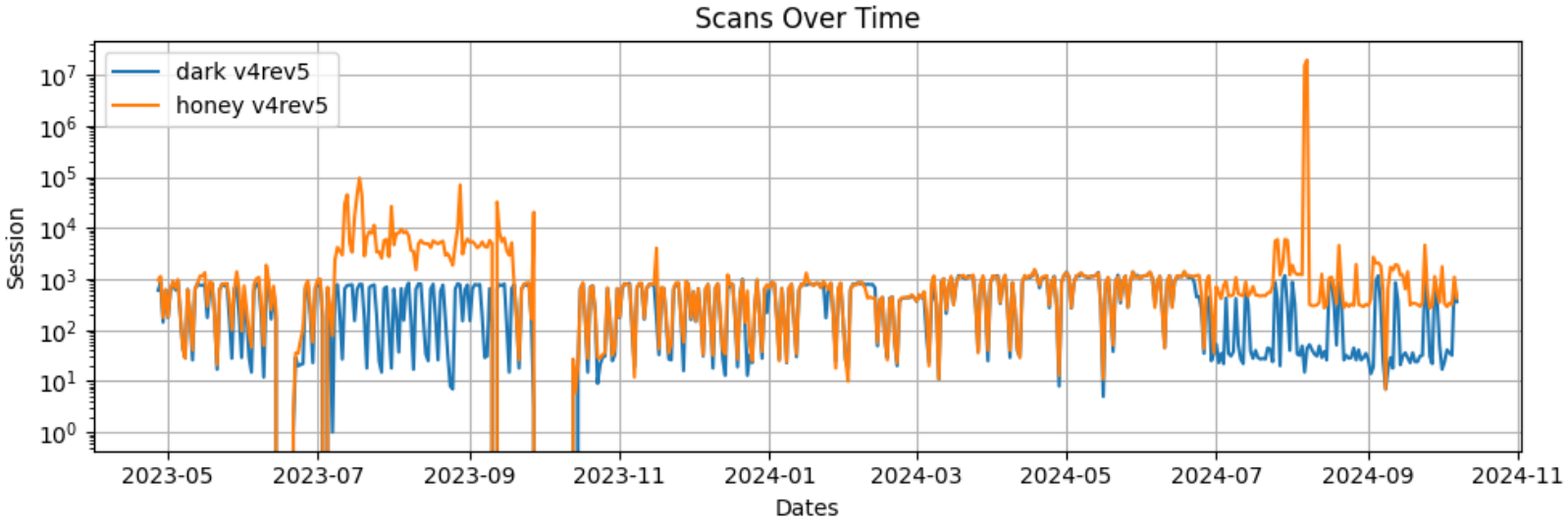
どの手法が効率的か?

スキャンの到来間隔



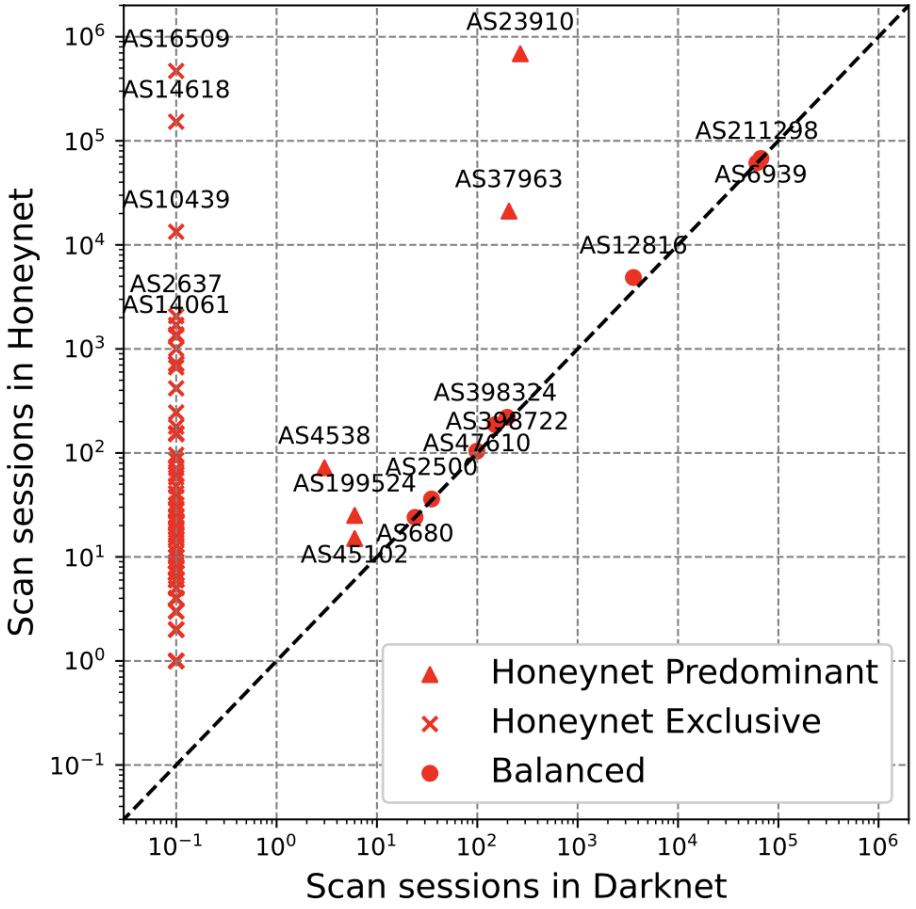
- アドレスを公開してから数日でスキャンが到来

スキャンの到来間隔



- 突発的なスキャンイベントも観測

公開・自作ヒットリストの利用

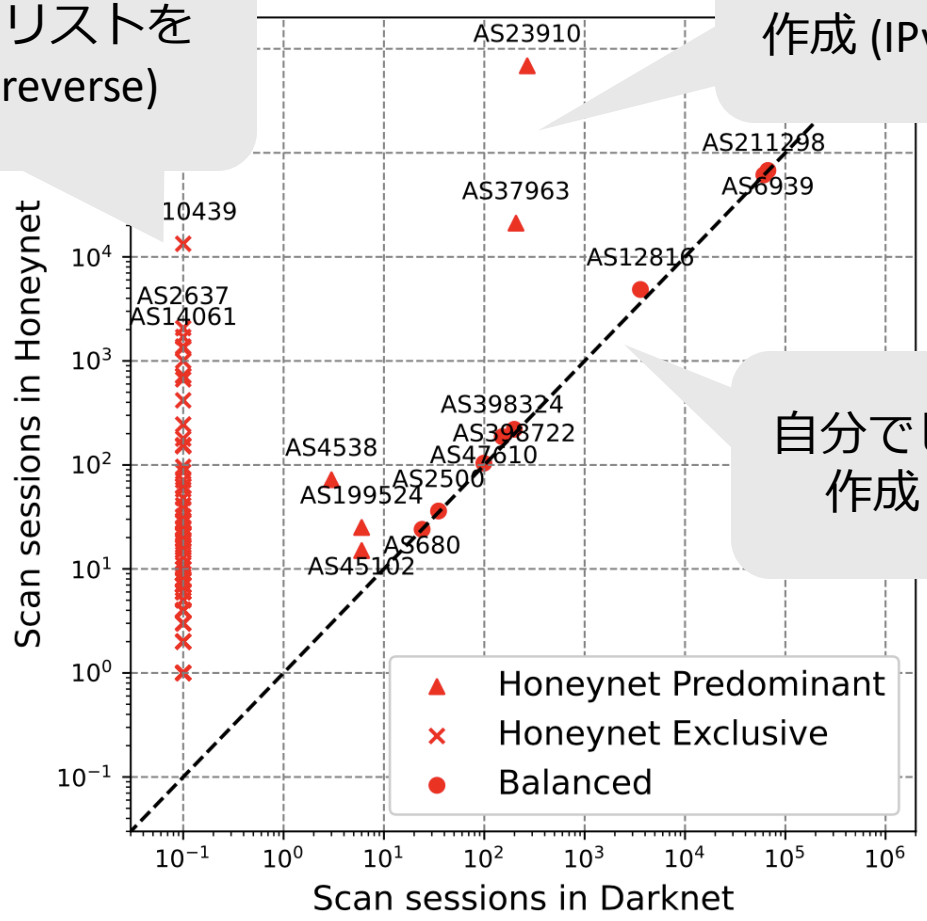


公開・自作ヒットリストの利用

公開ヒットリストを
利用(IPv4 reverse)

自分でヒットリストを
作成 (IPv4 reverse + random)

自分でヒットリストを
作成 (IPv4 reverse)



まとめ

ネットワーク上のスキャンとその脅威

- IPv4ではスキャンを回避することは困難
  - 機器のセキュリティが重要
  - 境界防御だけでなく多層防御
- IPv6でも大規模スキャンが現実的に
  - 高精度・広範囲の観測基盤の必要性
  - 将来のセキュリティ課題