JANOG53 発表資料

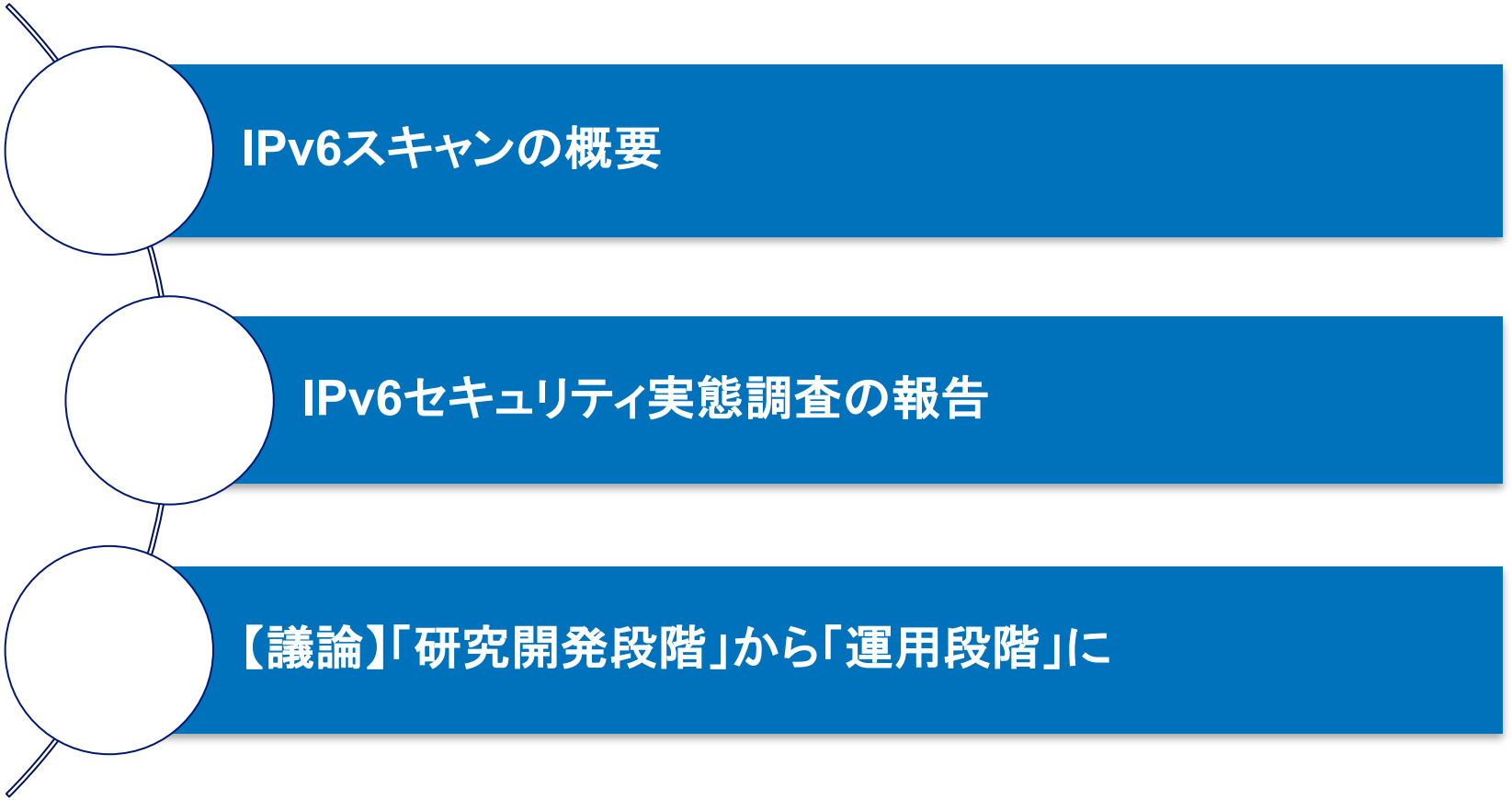
IPv6インターネットのセキュリティスキャン技術を運用する未来の作り方

NTT社会情報研究所　竹村 達也

自己紹介

- **名前：竹村 達也（Tatsuya Takemura）**
- 所属：NTT社会情報研究所（入社3年目）
  - サイバー攻撃対策技術の研究開発に従事
    - 攻撃発生前に脆弱ホストに対処
    - 特に、**IPv6アドレス空間のセキュリティスキャン**に着目
- 修士の研究テーマ：BGP経路ハイジャック対策関連
- JANOG：3回目の現地参加（45@札幌、50@函館、53@博多）

目次



目次

IPv6スキャンの概要

・世界的にどこまで研究されているかの紹介

・その中での我々の立ち位置

IPv6セキュリティ実態調査の報告

【議論】「研究開発段階」から「運用段階」に

ネットワークスキャンとIPv6

NTT

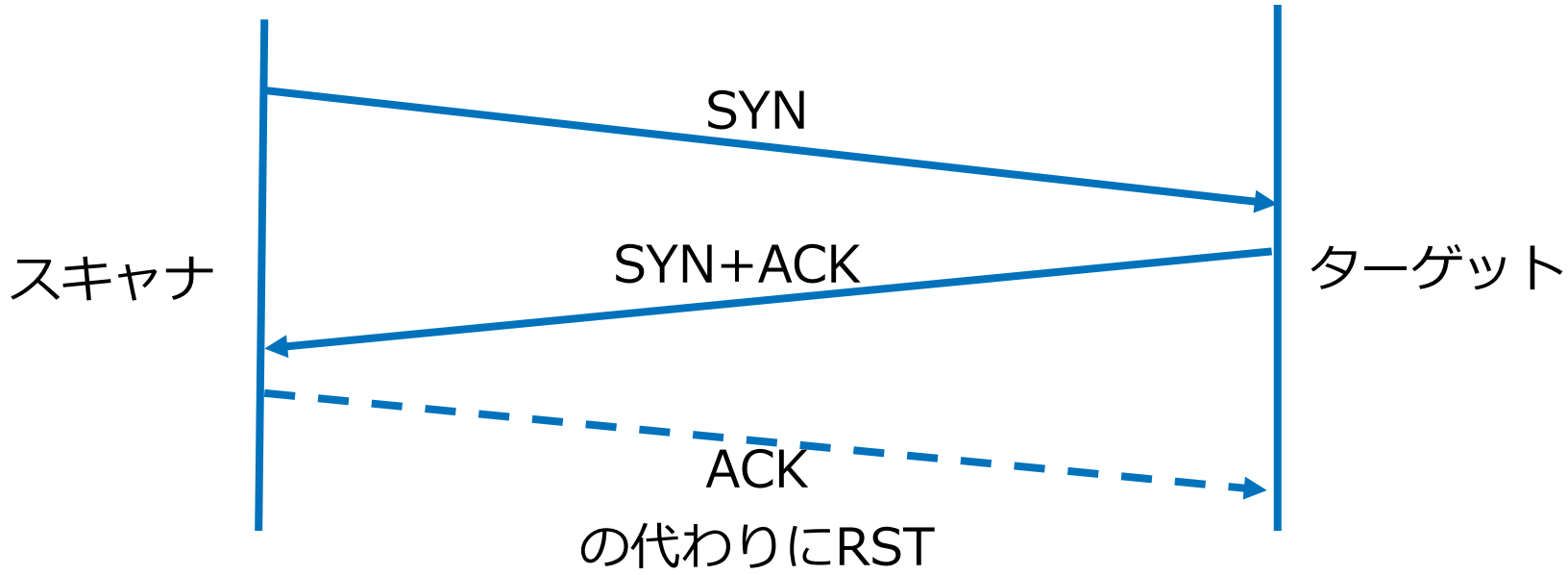
- スキャンは脆弱性への対応状況の調査など幅広い領域に貢献[1]
- 広大なIPv6アドレス空間では総当たりスキャンが不可能[2]
- 我々の目的：IPv6でもできるだけ効率的にアプリケーションレベルのスキャンを行い、攻撃発生前に脆弱ホストに対処したい

	IPv4	**IPv6**
ビット数	32 bit	128 bit
アドレス数	4.3×10^{9} （枯渇）	3.4×10^{38}
総当たりスキャン	可能 （ZMap[3]など）	不可能

[1] F. Li et al., “You’ve got vulnerability: Exploring effective vulnerability notifications”, USENIX Security’16.
[2] P. Foremski et al., “Entropy/IP: Uncovering structure in ipv6 addresses”, IMC’16.
[3] Z. Durumeric et al., “ZMap: Fast internet-wide scanning and its security applications”, USENIX Security’13.

ステートレススキャンの台頭

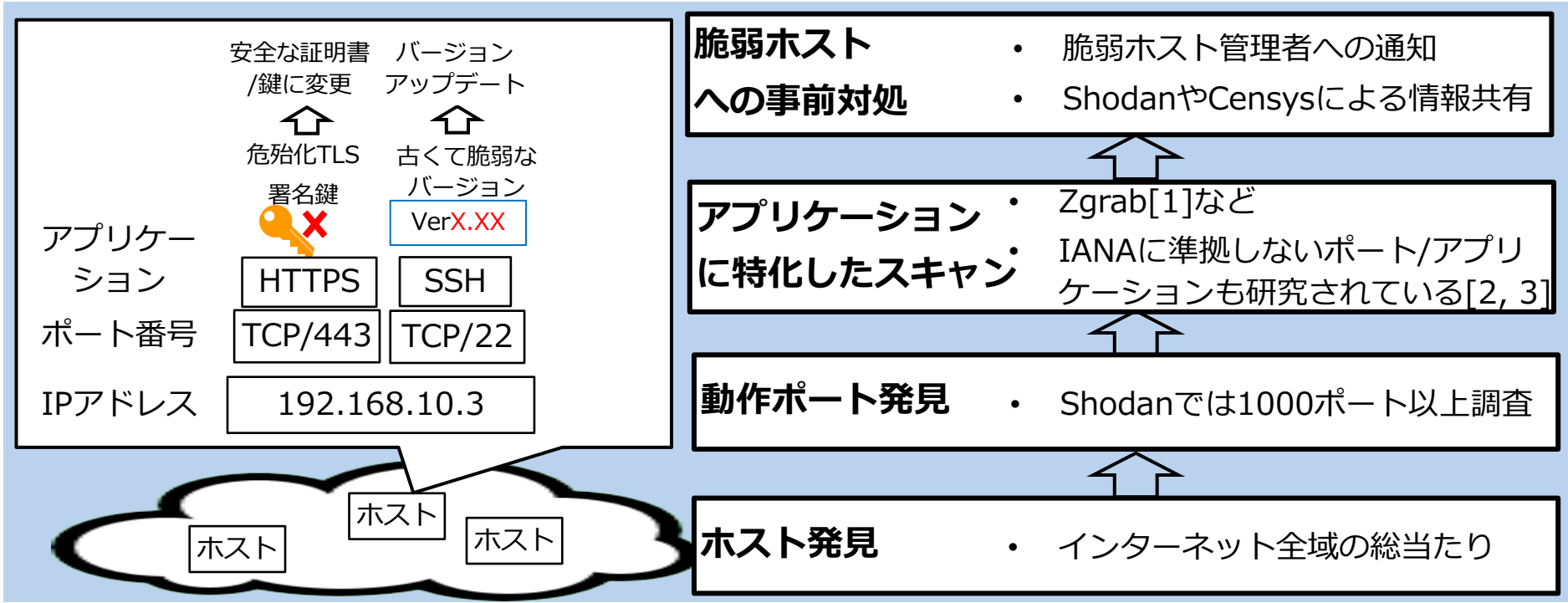
- ホストとコネクション確立せずに稼働ポートを調査
- ZMap[1]ではIPv4インターネット全体を1ポートあたり45分以下でスキャン可能



[1] Z. Durumeric et al., “ZMap: Fast internet-wide scanning and its security applications”, USENIX Security’13.

IPv4でのセキュリティスキャン

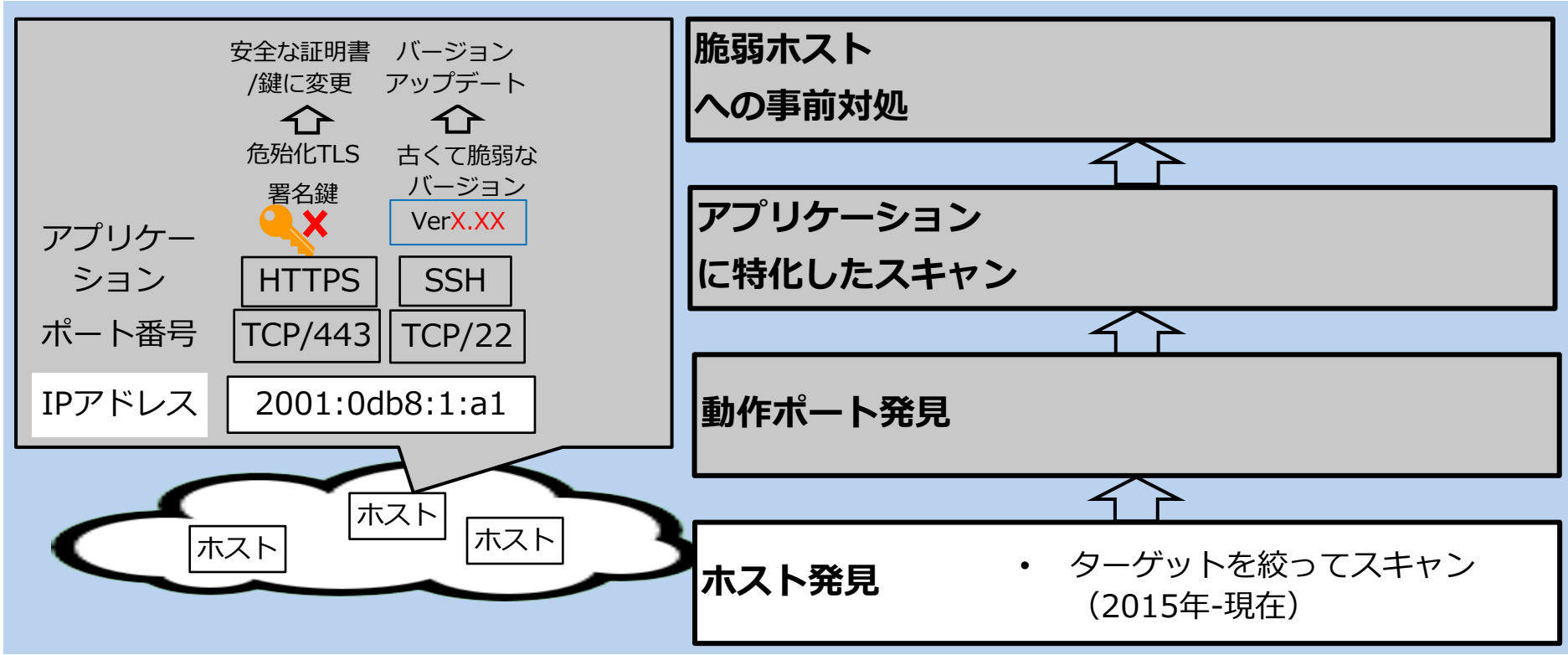
- 様々なポート/アプリケーションでセキュリティスキャンされている



[1] Z. Durumeric et al., “A Search Engine Backed by Internet-Wide Scanning”, CCS’15.
[2] Izhikevich et al., “LZR: Identifying unexpected internet services” USENIX Security’21.
[3] Izhikevich et al., “Predicting ipv4 services across all ports”, *SIGCOMM*’22.

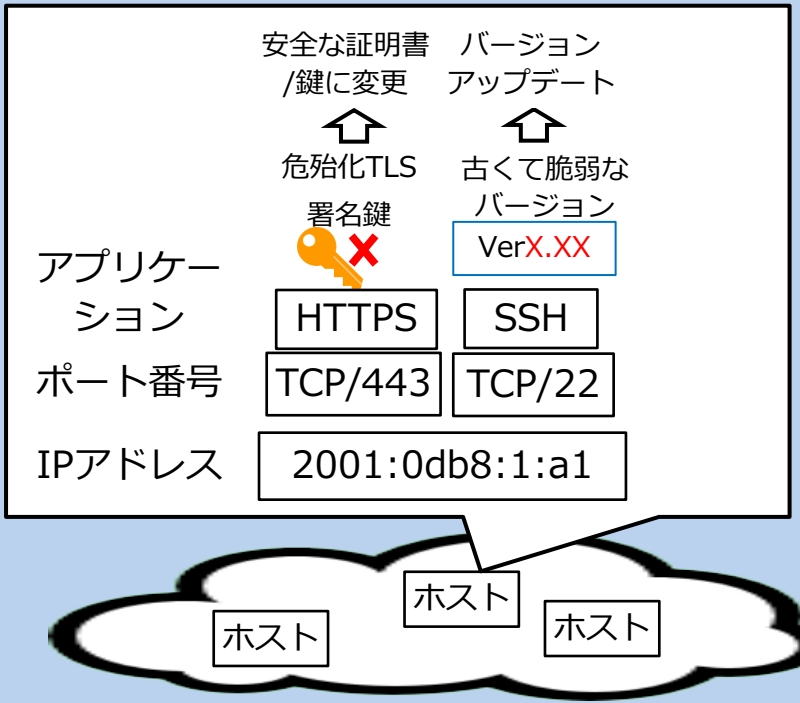
IPv6でのセキュリティスキャン

- 既存技術では、(一部の)ホストの存在レベルまでしか把握できていない



我々の取り組み

- 多様なポートでアプリケーションレベルのスキャンをできるようにした



脆弱ホストへの事前対処
- IPv6脆弱ホストが密集するASにメールで通知

アプリケーションに特化したスキャン
- ポートに該当するアプリケーションの脆弱性調査

動作ポート発見
- 多様なポートに対応

ホスト発見
- ターゲットを絞ってスキャン（2015年-現在）

IPv6ホスト発見技術

- IPv6スキャンの考え方：総当たりではなく、選択的にターゲットを生成
- 重要な3要素：ヒットリスト・TGA・ターゲットスキャン

ヒットリスト

- 存在が既知のIPv6アドレスのリスト

2001:0DB8::0001
2001:0DB8::0002
2001:0DB8::0004

→代表例：IPv6 Hitlist Service
- DNSデータやRIPE Atlasなどからの IPv6アドレスを収集
- 現在1.1Bアドレス (ICMPv6応答：9M, TCP/80応答：1.8M)

シードとして入力

TGA
(Target Generation Algorithm)

- アドレスを分析して新しくターゲットを生成

例：Entropy/IP, 6Gen, 6Tree, 6GCVAE, 6VecLM, 6GAN, 6Hit, 6Graph, AddrMiner, DET, 6Forest, 6Scan, …

→現在、13種類。それぞれが多様なターゲットを生成

ターゲットを出力

ターゲットスキャン

- 生成したターゲットをスキャン

2001:0DB8::0001
2001:0DB8::0002
2001:0DB8::0003
2001:0DB8::0004
2001:0DB8::0005
2001:0DB8::0006
…

→スキャン自体の本質はIPv4と変わらない

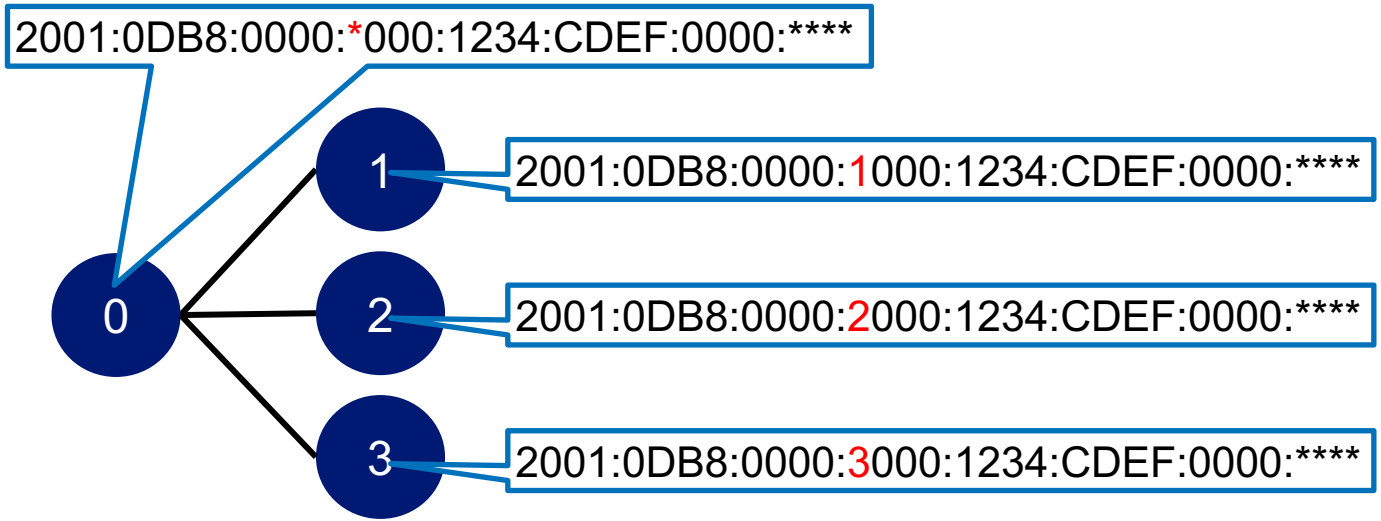
代表的なターゲット生成アルゴリズム

	アプローチ	Alias 処理	スキャンフィードバック
Entropy/IP [1]	エントロピー+ クラスタリング		
6Gen [2]	木	✅	
6Tree [3]	木	✅	✅
6Hit [4]	木+機械学習	✅	✅
6GAN [5]	機械学習	✅	
6Forest [6]	機械学習		

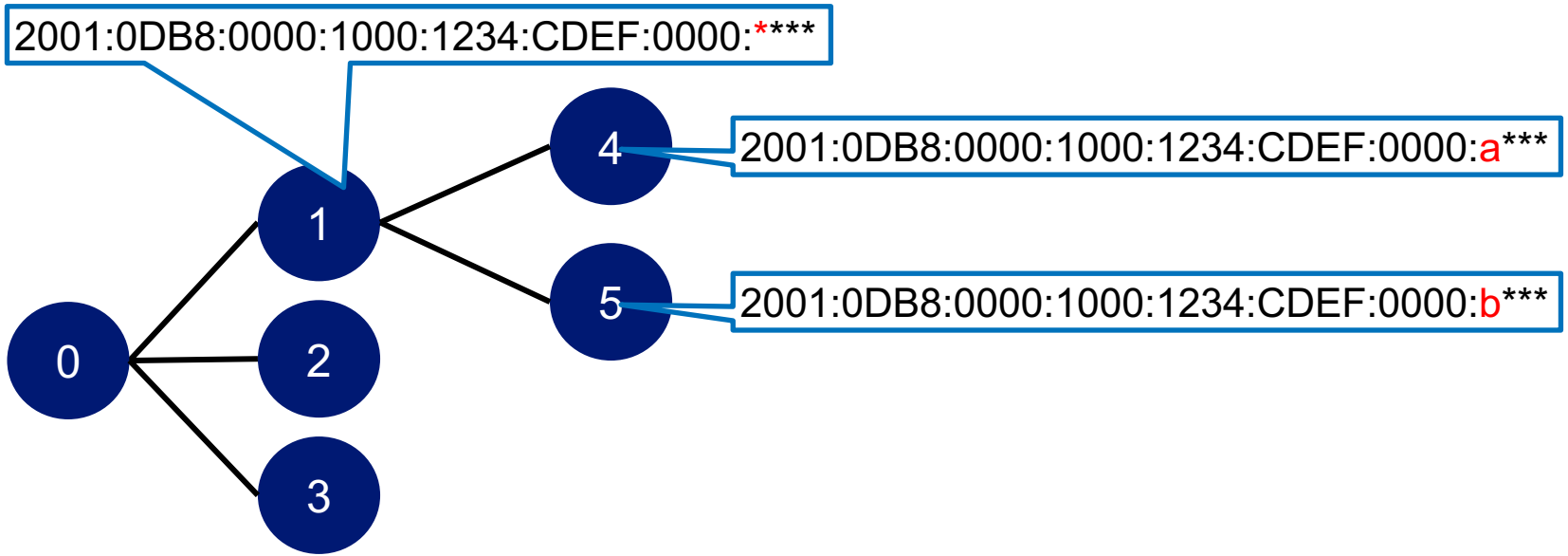
[1] P Foremski et al., Entropy/ip: Uncovering structure in ipv6 addresses, IMC'16.
[2] Z Liu et al., Target generation for internet-wide IPv6 scanning, IMC'17.
[3] B Hou et al., 6Tree: Efficient dynamic discovery of active addresses in the IPv6 address space, Computer Networks'19.
[4] B Hou et al., 6Hit: A Reinforcement Learning-based Approach to Target Generation for Internet-wide IPv6 Scanning, INFOCOM'21.
[5] T Sui., 6GAN: IPv6 Multi-Pattern Target Generation via Generative Adversarial Nets with Reinforcement Learning, INFOCOM'21.
[6] T Yang et al., 6Forest: an ensemble learning-based approach to target generation for internet-wide IPv6 scanning, INFOCOM'22.

例：6Tee（木をベースとした手法）

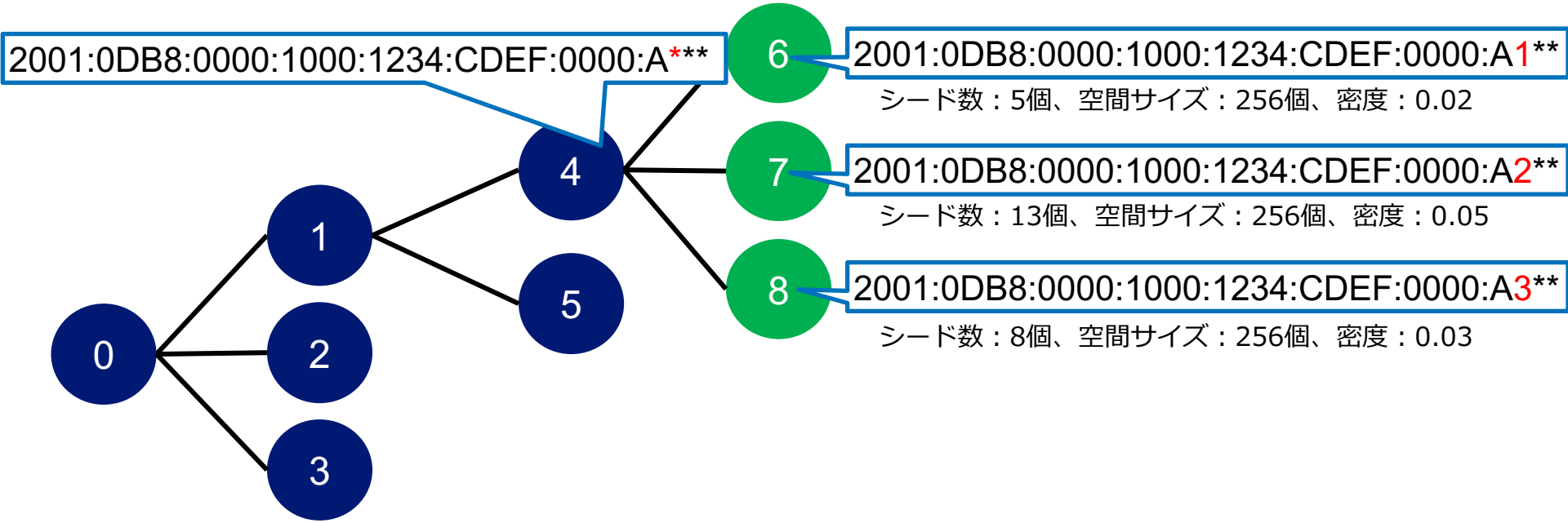
B Hou et al., 6Tree: Efficient dynamic discovery of active addresses in the IPv6 address space, Computer Networks'19.



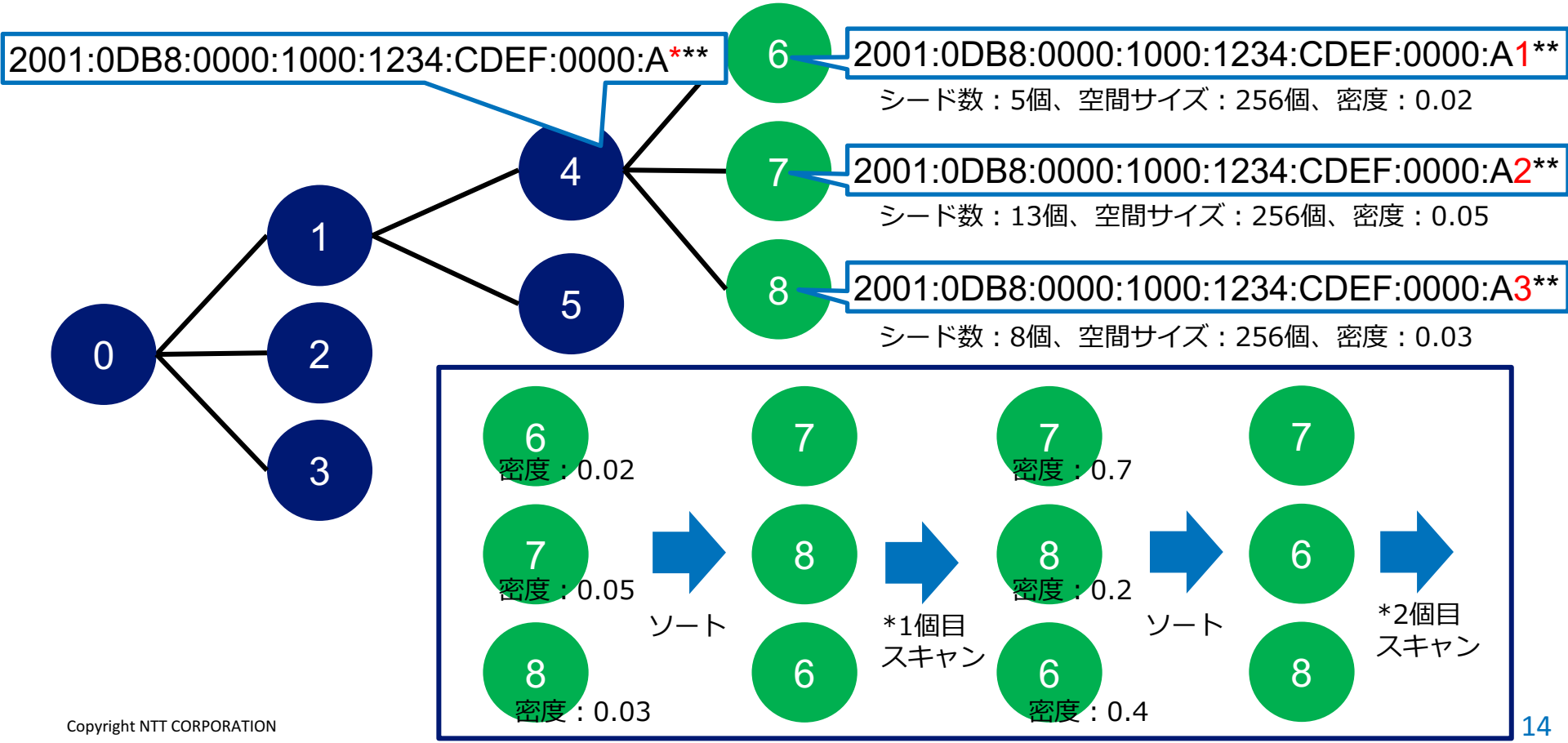
例：6Tee（木をベースとした手法）



例：6Tee（木をベースとした手法）

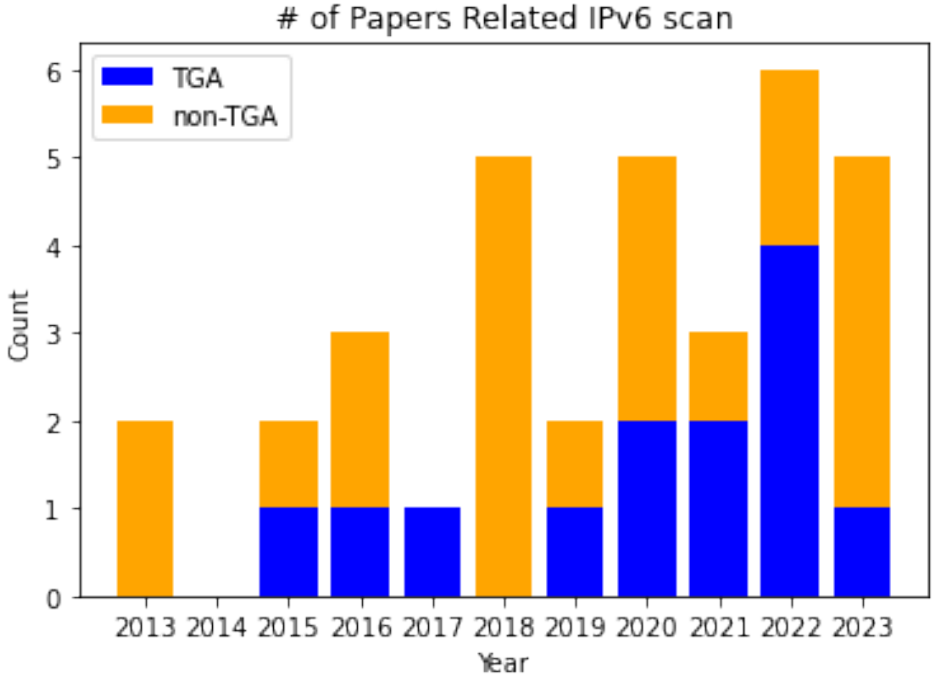


例：6Tee（木をベースとした手法）



IPv6スキャン関連研究の傾向

- 特に、TGA (Target Generation Algorithm) の研究が多く、今後も色々と提案される可能性
- 他には、ヒットリスト自体・EUI-64関連・ダークネット的な分析の研究など



※セキュリティ・通信分野の難関国際会議に採録された論文が中心

目次

IPv6スキャンの概要

IPv6セキュリティ実態調査の報告

・我々の独自技術による調査

【議論】「研究開発段階」から「運用段階」に

多種ポート対応スキャンフレームワーク（独自技術）

- ホストの存在に加えて、その上で動くポートまで効率的に発見するためのスキャンフレームワークを開発
- 例えば、TCP23/Telnet ではフレームワークを利用することで10倍以上効率化

ヒットリスト
（存在が既知のホスト）

```
2001:0DB8::0001
2001:0DB8::0002
2001:0DB8::0004
```

フレームワーク

Entropy/IP

6Gen

6Tree

- プレフィックスごとにポートの利用傾向を分析することで効率的にポートスキャン
- 任意のTGAに対応

稼働している(ホスト, ポート)
の組合せ

```
2001:0DB8::0001, tcp80, 443
2001:0DB8::0002, tcp80, 21, 22
2001:0DB8::0003, tcp23
2001:0DB8::0004, tcp110, 143
2001:0DB8::0005, tcp80, 8080
2001:0DB8::0006, tcp23, 179
…
```

※技術詳細は事後資料にて紹介

IPv6セキュリティの実態調査

- 我々の独自フレームワークが発見した稼働状態の (ホスト, ポート) のうち、TCP22/SSH、TCP3306/MySQL、TCP143/IMAP、TCP23/Telnetでのバナー調査を実施 [289AS]
- 各ASのIPv4アドレス空間も同様に調査
- クラウド環境のようなホストが密集している場所では発見されやすい

	IPv4のみ脆弱	IPv6のみ脆弱	IPv4/IPv6どちらも脆弱
SSH TCP/22	4 (1.38%)	5 (1.73%)	1 (0.35%)
MySQL TCP/3306	76 (26.3%)	3 (1.04%)	2 (0.69%)
IMAP TCP/143	34 (11.76%)	0 (0.0%)	1 (0.35%)
Telnet TCP/23	137 (47.4%)	8 (2.77%)	31 (10.73%)

IPv6が新しいからセキュリティ設定が適切に行われているとは限らない（例：IPv4でSSHを利用していて、意図せずIPv6でもSSHが動いている）

IPv4もIPv6も脆弱な運用がなされているASであり、IPv4のために作成された攻撃シナリオをIPv6でも利用できる可能性（例：IPv6アドレス空間への効率的なボットネットの拡大）

- IPv6アドレスがいつでも知られうる前提で機器などを適切に設定するのが重要
- IPv4で標的とされたり問題が発生したアプリケーションの情報を利用して、IPv6運用をよりセキュアに

目次

IPv6スキャンの概要

IPv6セキュリティ実態調査の報告

【議論】「研究開発段階」から「運用段階」に

・悪用されないようにしつつ、セキュリティ用途で活用できる未来

議論：IPv6スキャン技術運用の課題

- IPv4での歴史からわかるように、スキャン技術はセキュリティ用途だけでなくマルウェア感染拡大用途など悪用もされうる「諸刃の剣」[1]
- IPv6ではさらにスキャン技術の活用方法が複雑に

	IPv4	IPv6
NW管理者/運用者以外の第三者からの注意喚起	• インターネット全体で可能	• インターネット全体は不可能 • ヒットリスト次第

NWによっては、注意喚起なかったが攻撃用途のスキャンがきて悪用の側面が上回る可能性

→各NWを管理する人が個別に対策する必要

[1] Antonakakis et al., Understanding the mirai botnet, USENIX Security‘17.

議論点

1. そもそも、現時点では「IPv6は広大なのでスキャンされない想定」or「IPv6もそれなりにスキャンされるのを想定」のどちらで運用されていることが多いか？
2. IPv6の脆弱ホストについては、IPv4と比較して数が少ない今のうちに、できる限り対処しておいた方が良いのではないか？
3. IPv6で効率的なスキャン技術がOSSとして公開されたり、ShodanやCensysが当たり前になる（つまり攻撃者が手軽にスキャンしてくる）前に、各NWを管理する人が個別に準備する期間はどの程度必要か？