DDoSアワー
Worm拡散の早期検知について

(OCNでの不正宛先向けトラフィックの解析)

2005年1月20日

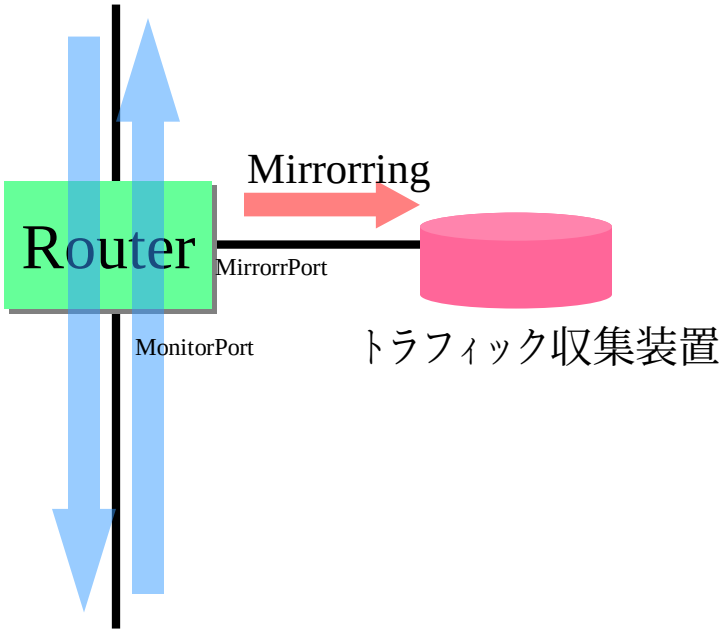
須藤　年章＜sudo@ocn.ad.jp＞
安田　歩＜yasuda@ocn.ad.jp＞
NTTコミュニケーションズ株式会社

Agenda

- モニタリング手法
- トラフィックの眺め方
- モニタリング、解析の実例
  不正宛先向けトラフィックの解析によるWorm拡散の早期検知

トラフィックモニタリング手法

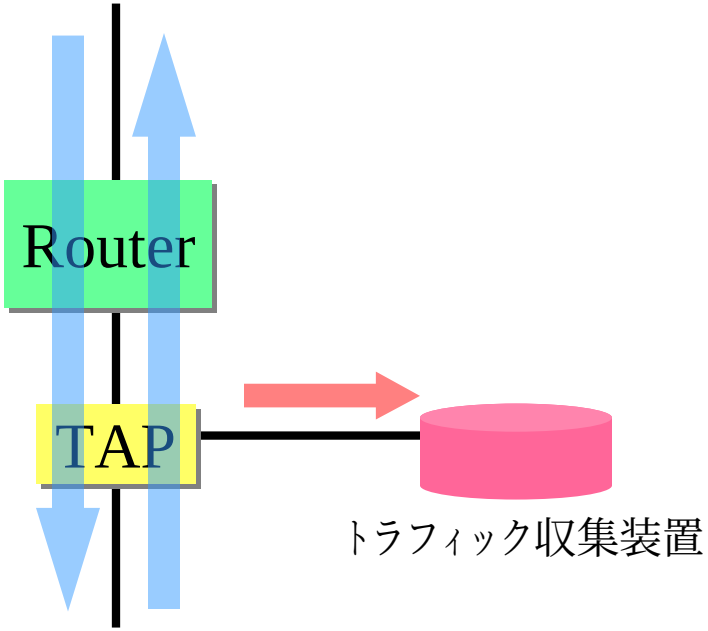
Port Mirrorring



■ルータ/SWのPort mirror機能を使用する。
■解析にはsnifferやids等を使用。

・全パケットを収集可能
・パケット全体を収集可能
・ルータ負荷については装置それぞれ。
・Mirroring機能的に制限がある場合がある。
　内部BUSまたぎのmirrorの高負荷

Tapping

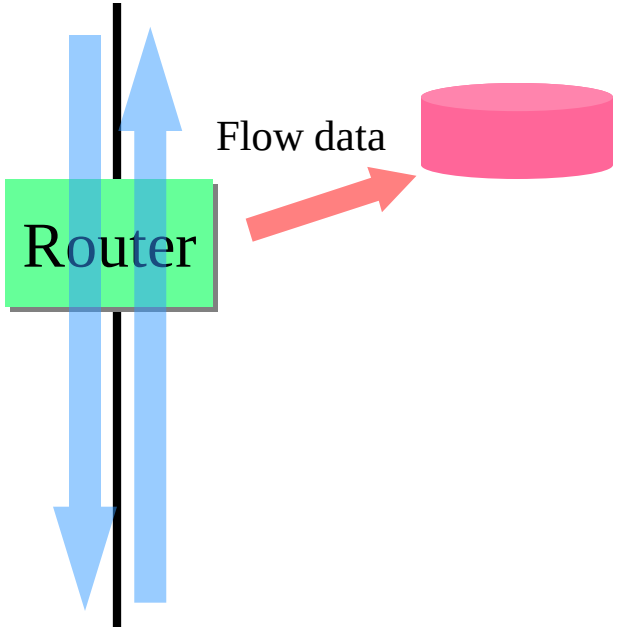


■UTP TAP/Optical Tap等で物理的に分岐する。
■解析にはsnifferやids等を使用。

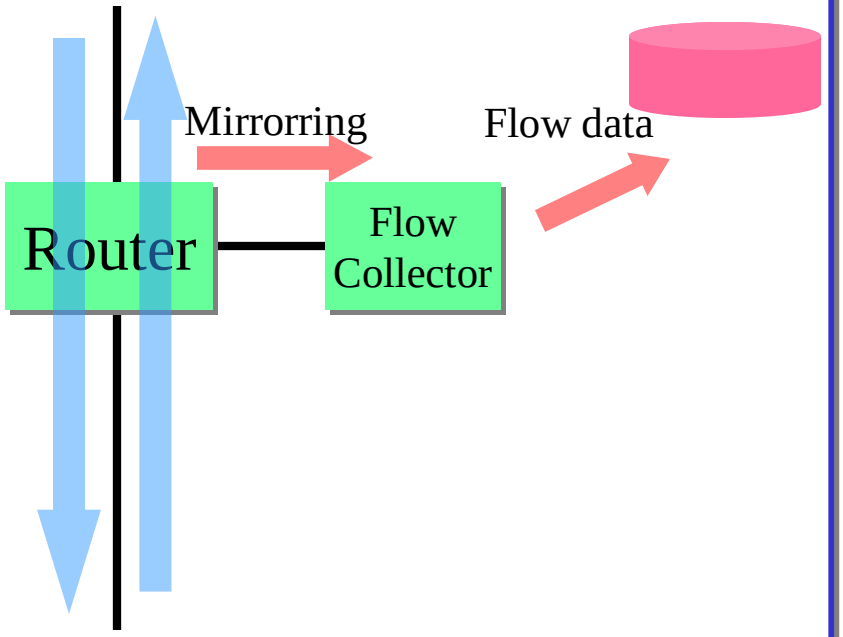
・全パケットを収集可能
・パケット全体を収集可能
・ルータ負荷はない。

トラフィックモニタリング手法

Flow collector



Monitoring/Tapping+Flow collector



■ルータのflow collector機能を使用する。
■解析にはflow解析用のソフト/アプライアンスを使用
■ルータの負荷を軽減するためにPortMirroringと組み合わせることも可能。

・CPU負荷が高い。
・基本的サンプリング。すべてのパケットを収集できない。
・IPパケットのヘッダー情報のみを収集。

トラフィックモニタリング手法

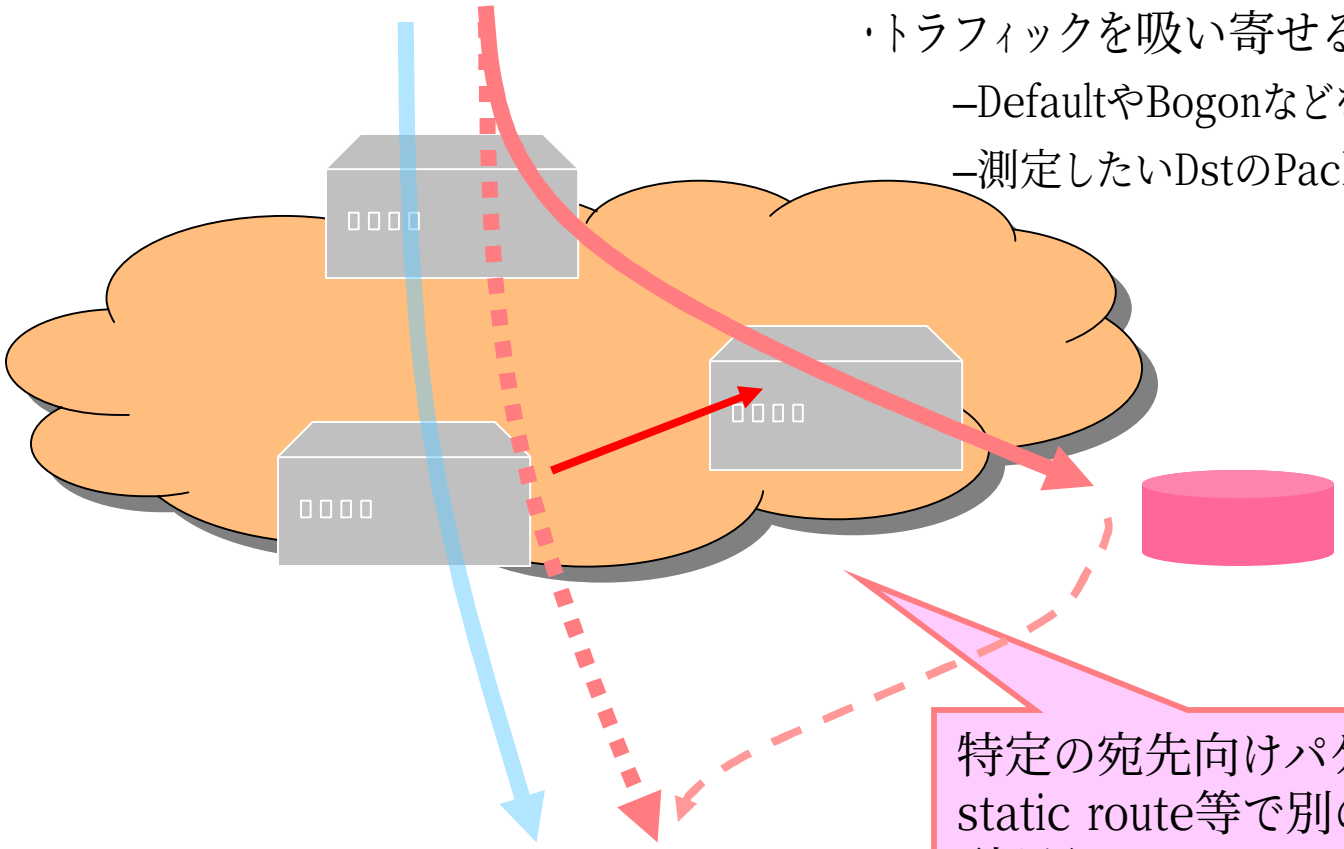
Traffic引き込み

・Honey Pot

・トラフィックを吸い寄せる

–DefaultやBogonなどをルーティング

–測定したいDstのPacketをルーティング



特定の宛先向けパケットを
static route等で別の方向へ
引き込む

トラフィックモニタリング手法

	Netflow collector	Port mirroring/ Tapping	Honey Pot/Traffic 引き込み
導入	容易（設定のみ）	配線やTAP等の機器が必要になる。	容易（ルーティング）
拡張性	高い	低い（リンク毎に必要）	高い
正確性	低い（Sampling rate次第）	高い（解析方法によるが）	高い（解析方法によるが）
トラフィック制限	あり（Samplingで制御）	なし（基本的にインタフェース速度だが最近の装置ではそういった機能もある）	一部のトラフィックのみが対象とするので調整可能。
設置箇所	ルータ	対処のルータ/リンク	ネットワークのどこか
収集対象	対象のルータ/リンクがあつかうトラフィック	対象のルータ/リンクがあつかうトラフィック	一部のネットワーク
その他	Vender毎にFlowデータの仕様が違う		

トラフィックの眺め方

- 全体的な量の変化
  - ・急に増えた⇒何かがおこっている？
- 特定の種類のパケットの増減
  - ・特定の宛先ポート向け
- 特定の送信元、宛先に関するパケットの増減
  - ・一つの送信元から一つの宛先　⇒DoS？アタック？
  - ・一つの送信元から不特定多数の宛先　⇒何かアタック？Worm？
  - ・不特定多数の送信元から一つの宛先　⇒DDoS？
- 特定の時間帯のトラフィックの増減

- 使用されるツール
  - ・MRTG
  - ・NMS
  - ・Netflow / sflow等のデータ解析アプライアンス
  - ・プロトコルアナライザ、sniffer
  - ・IDS　など

} 総量を眺めるには十分

} より高位レイヤを眺める

モニタリング/解析の実例

実例としてユーザーNWから送信される不正宛先向けの不要トラフィックを解析することで、Wormの蔓延度、Worm拡散トラフィックの状態を観測する。

- 不要トラフィックとは
  - 正当な宛先を持たず、廃棄されるパケット
- 対象となる宛先IPアドレス
  - プライベートアドレス
  - その他予約されたアドレス
  - full route に存在しないアドレス

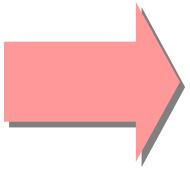
RFC 3330 Summary

Address Block	Present Use	Reference
0.0.0.0/8	"This" Network	[RFC1700, page 4]
10.0.0.0/8	Private-Use Networks	[RFC1918]
14.0.0.0/8	Public-Data Networks	[RFC1700, page 181]
24.0.0.0/8	Cable Television Networks	--
39.0.0.0/8	Reserved but subject to allocation	[RFC1797]
127.0.0.0/8	Loopback	[RFC1700, page 5]
128.0.0.0/16	Reserved but subject to allocation	--
169.254.0.0/16	Link Local	--
172.16.0.0/12	Private-Use Networks	[RFC1918]
191.255.0.0/16	Reserved but subject to allocation	--
192.0.0.0/24	Reserved but subject to allocation	--
192.0.2.0/24	Test-Net	
192.88.99.0/24	6to4 Relay Anycast	[RFC3068]
192.168.0.0/16	Private-Use Networks	[RFC1918]
198.18.0.0/15	Network Interconnect Device Benchmark Testing	[RFC2544]
223.255.255.0/24	Reserved but subject to allocation	--
224.0.0.0/4	Multicast	[RFC3171]
240.0.0.0/4	Reserved for Future Use	[RFC1700, page 4]

Worm拡散活動トラフィックを対象とした訳

Worm拡散時のトラフィックがNWへ与える影響は非常に大きい

- 無駄な帯域を消費する。
- 装置負荷の上昇による正規トラフィック処理への影響
  - ・Forwarding用のCacheの消費
  - ・セッションテーブル溢れ
    NATやFirewallなど。



CPUを圧迫

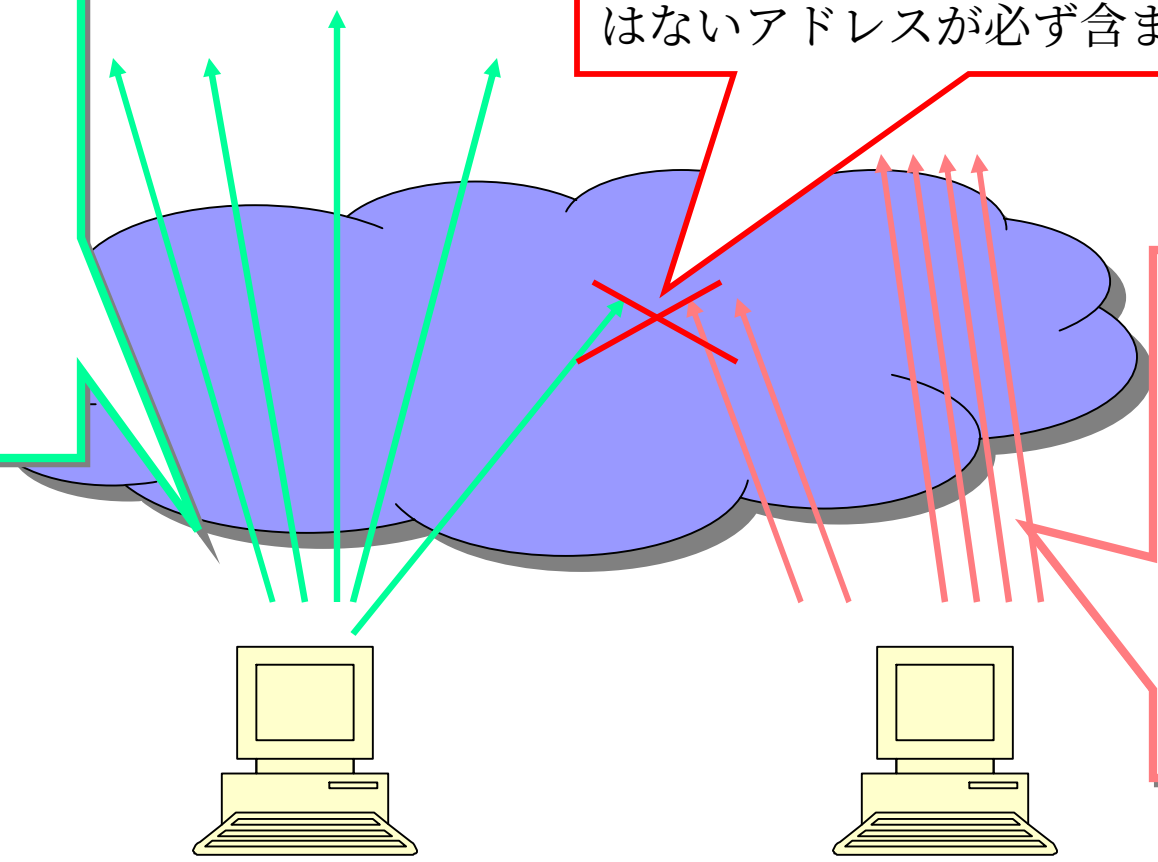
Memoryの消費

- ユーザー間の不要トラフィックの増加。
- 外部NWへの影響
- 閉域網で発生すると目も当てられない。

ワームの感染活動時の挙動

ワームの感染パケットの宛先には通常は到達性のないアドレス（Full route）にはないアドレスが必ず含まれる。

Destination address
1.A.A.253
100.C.A.2
192.168.A.B
201.D.A.5
D.1.0.100
10.0.D.C
⋮

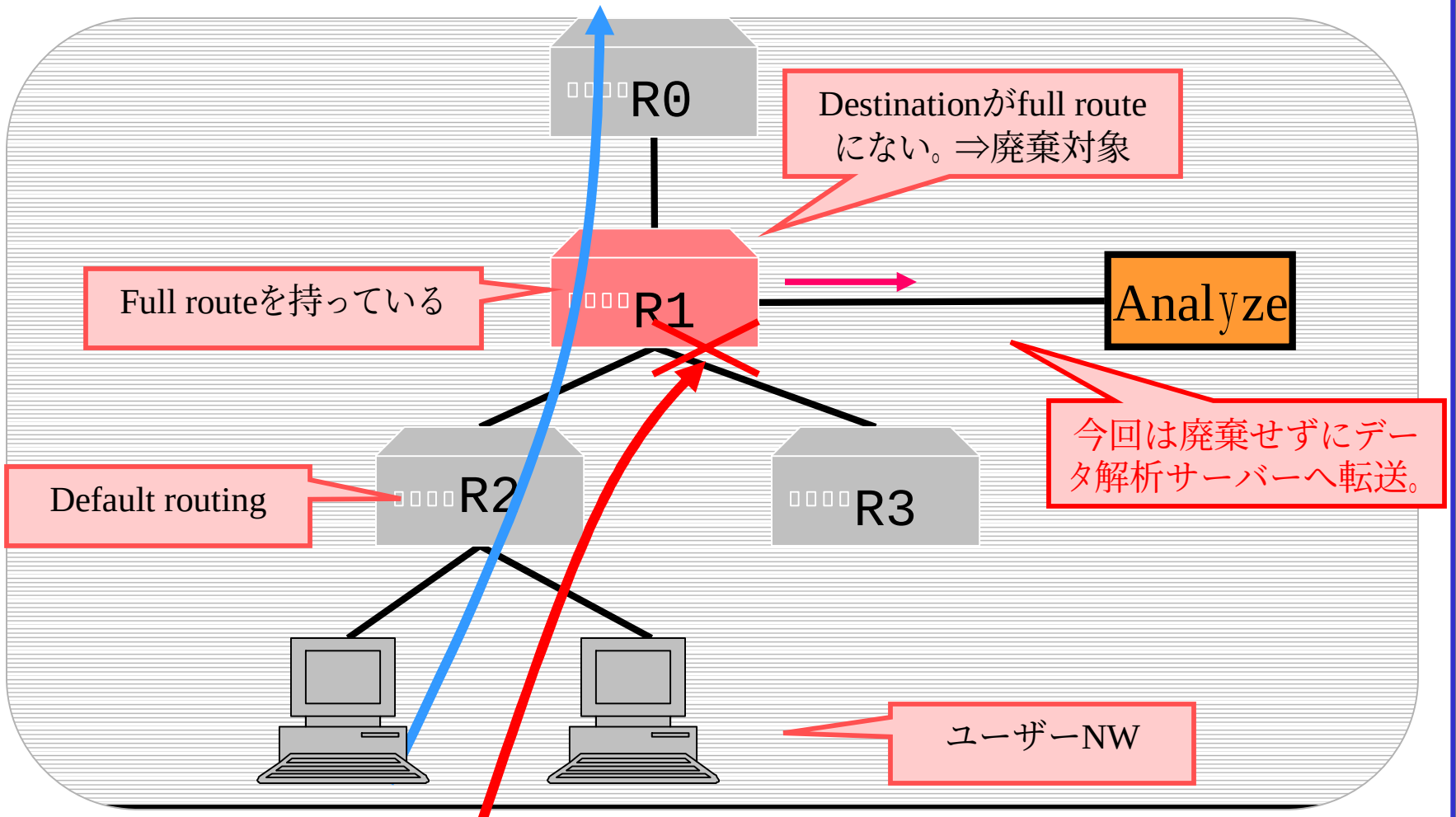


Destination address
1.A.A.0
1.A.A.1
1.A.A.2
⋮
192.168.0.1
192.168.0.2
192.168.0.3
⋮

ランダム送信
Slammer等

シーケンシャル送信
Blaster等

モニタリング原理



ユーザーから出された無駄なトラフィックを上位NWへ送らないようある程度集約された途中ノードで廃棄している。この廃棄ポイントで不要パケットを解析する。

不正宛先向けトラフィック解析

★どのような切り口でデータをながめるか？

- 総量の変化をながめるだけでも

  そもそも不正な宛先向けの廃棄されるパケット。その量の増減傾向をみるだけで異常度合いがわかるのでは？

  「急に増えた」「急に減った」ということで何か異常が発生しているという初動判断が可能？

- 宛先port

  「感染しているWormの識別」「新種のWormの発生」

- 送信元IP

  感染者数、蔓延傾向、沈静化傾向

- TCP/UDP

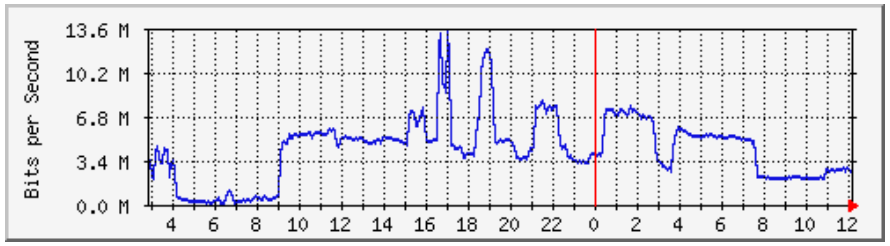
- IDSによる解析

  多面的に解析するために

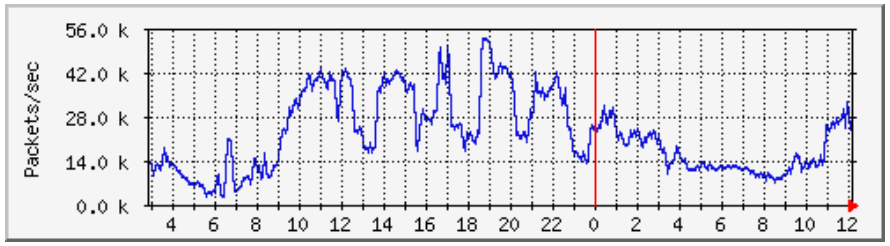
トラフィック量を見る

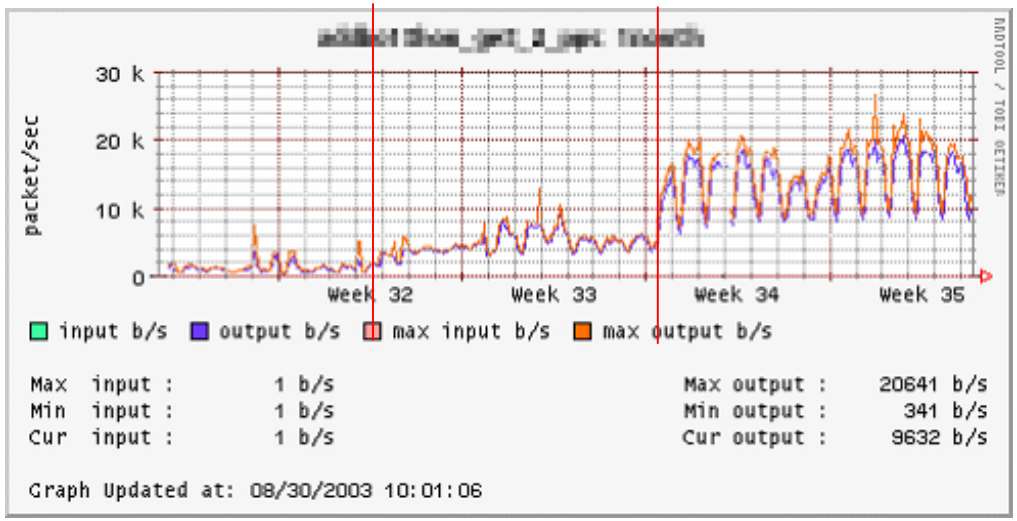
★全体量の変化を眺めてみる。

Bit/sec



Packet/sec





←過去の例)
Blaster/Nachiの発生による
不要トラフィックの増大状況。

顕著に今までとは違う傾向があらわれていることが観測できる。

トラフィック量を見る

- 全体のトラフィックに対する不要トラフィックの割合は0.25%程度。
- 過去の大規模なWorm発生時に、ほぼ発生と同時と思われるタイミングで量的増加を確認できている。

同時に他の場所に仕掛けてあったIDSでの顕著な状況の検知は6～8時間後。
IDSの監視対象NWの規模に依存するため、監視対象とするNW、アドレスブロックの規模を大きくすることで検知率/速度を上げることは可能、、、

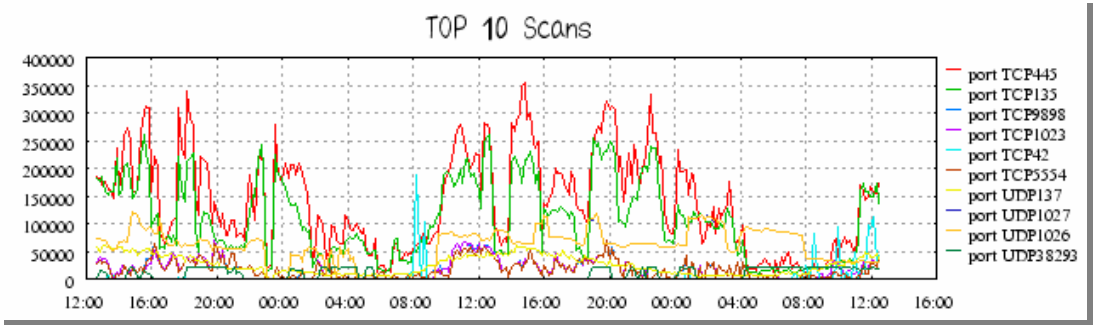
不正宛先向けトラフィックをみることにより、少ない観測対象から迅速に網内のWormの拡散トラフィック、蔓延状況を観測することができる。

「届くパケットよりも送出されるパケットを観測したほうが効率的？」

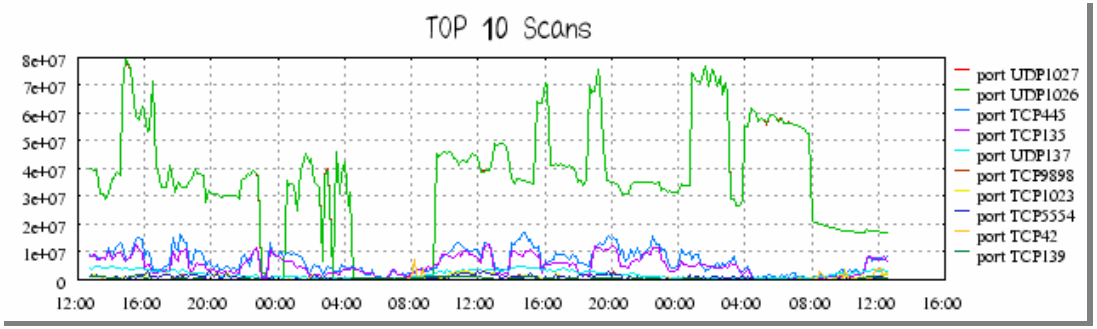
宛先PORTを見る。

★プロトコル、宛先ポート別のトラフィックを眺めてみる。

Packet count



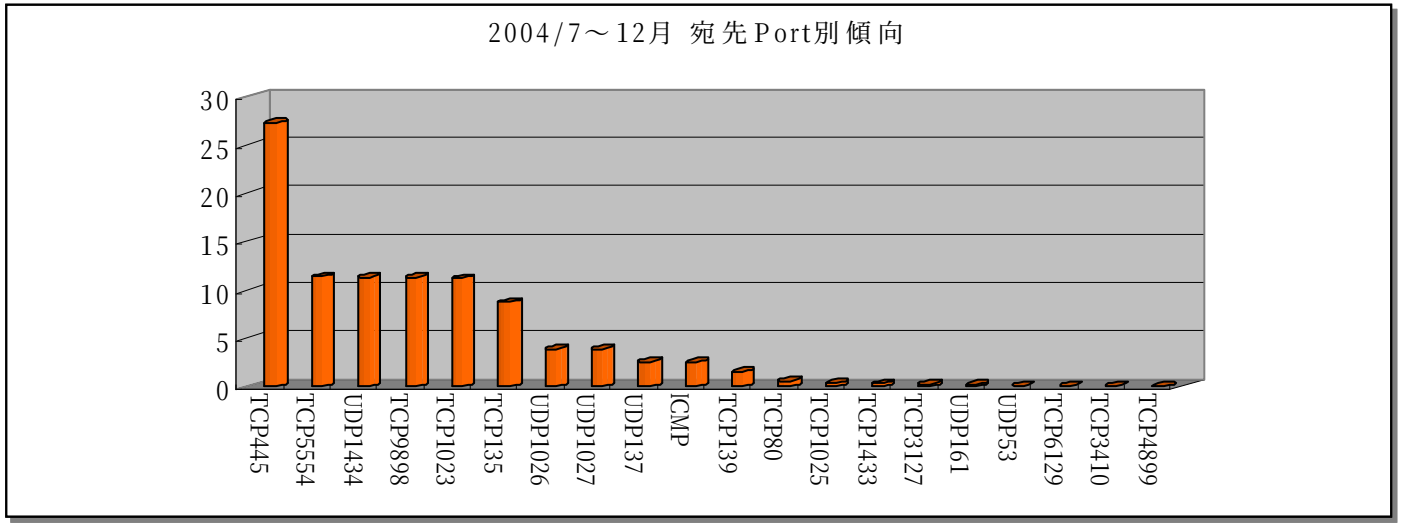
bit count



代表的なWormとその利用するport

Protocol /Port	Packet Size	Service	Worm/Virus
UDP1434	404	MS-SQL-Monitor	Slammer
TCP135	48	Location server	Blaster
TCP445	48	Microsoft-ds	Sasser, Dabber
TCP137	48	NetBios	Blaster,etc
TCP9898	48		Dabber
TCP1023	48		Dabber
TCP5554	48		Dabber
UDP1026	553	Messenger Service	Spam広告
UDP1027	553	Messenger Service	Spam広告

宛先PORTを見る。

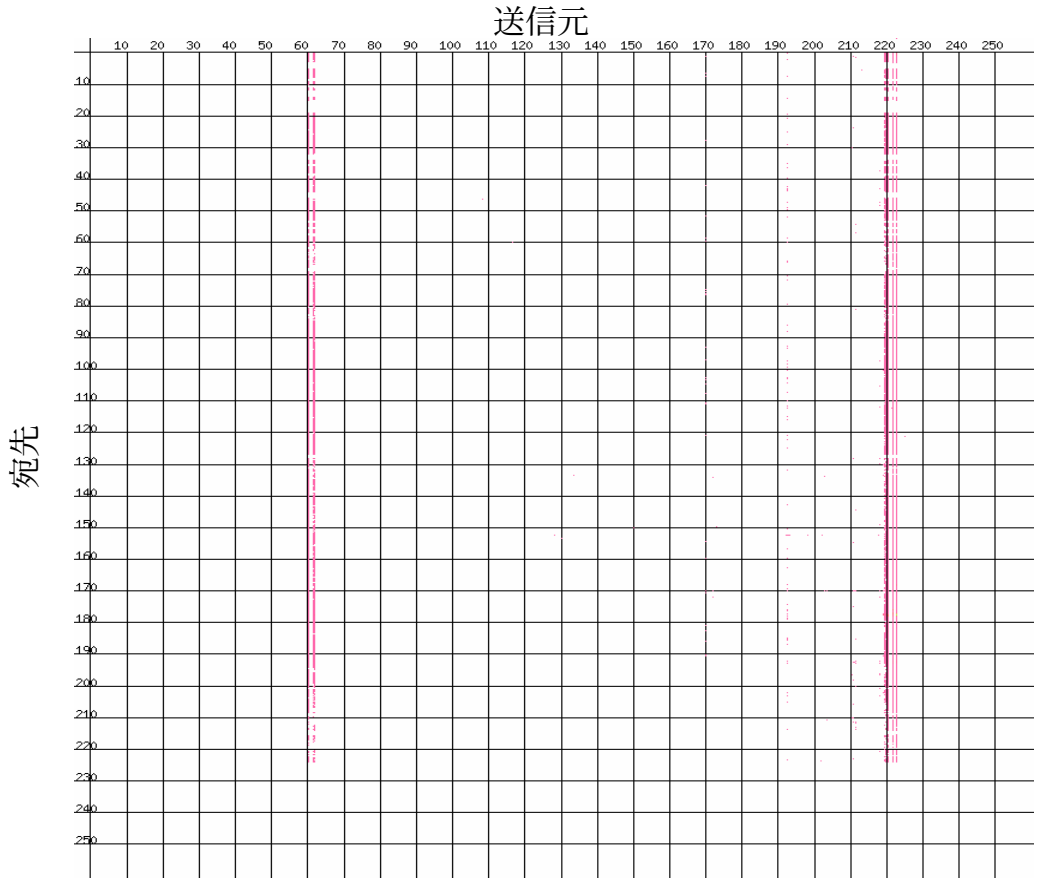


- 傾向
  ・Slammer等も未だ多量に存在し、その爆発力は大きい。
  ・Sasser/Dabber（TCP445,1023,5554,9898）、Messenger Spam（UDP1026,1027）の活動が顕著

- 宛先Portに関しては想定どおりWormの感染活動に使用されるものがほとんど。
  送信元IPを集計することにより網内でどの程度のユーザーがWormに感染しているか推定できる。

送信元IPについて

★送信元IPを眺めてみる。



■傾向

・送信元を改竄されているパケットは多数あるが全体に対する比率的には非常に小さい。

蔓延しているWormの挙動的には送信元を改竄しないものが多いため？

・プライベートアドレスはある程度見られる。

送信元IPについて

★もう一歩すすめて

Worm感染したPCはbackdoor活動により特定ポート待ち受けや任意のircサーバーへ接続しZombieホストとしてDDosやSpamの発信源となりうる。

そのようなZombie化によるSpamの発生源となっていないか状況を確認してみる。

やり方

・DNSBLを使用する。

SpamcopやSpamhaus等のDNSBLへのDNS問い合わせ等によりSpam Source、Zombie、HijackedIP情報を提供されている。

・Virus DNS LookUpの調査結果とのマッチング

backdoor活動を行うWormの中には、特定のPORTを開いて待ち受けているものの他に、任意のIRCサーバーへの接続をしたり任意のサーバーへ感染した旨を伝えるものがある。
その際に接続先URLをDNSにて確認するため、そのURLを調べるqueryを追いかければその種のWormの感染者を推定できる。
ただし、irc等に接続しているzombieとそうでないzombieのどちらの影響が大きいかは定量的にはわからないけど。。。。。

今回はDNSBLとの対応のみで。。。

送信元IPについて

★DNSBLの情報との比較

- 登録/削除ポリシー(登録理由、30分でupdate、2日で削除、依頼がくるまで削除されない等)が各リストで違いがあるため複数のリストで比較。

理由	list1	list2	list3	list4	list5	list6
DynamicIP	67.6%	0.5%	-	-	-	-
Form mail	3.8%	-	-	-	-	-
Open Sock	0.3%	-	-	-	-	-
Spam Source	0.1%	1.3%	1.0%	57.8%	88.2%	-
Open Relay	-	-	-	8.1%	-	-
Zombie/hacked ip	-	-	-	-	-	0.1%

- 傾向

・何れかのlistに登録されているかどうかという観点からみると99.9%登録されている。
・約60%程度は複数のlistでspamの送信源と認識されているとみえる。
・実際には送信されるパケットの宛先port別に比較する必要があるかも。
・BLの鮮度が高いために、比較はリアルタイムにおこなうべき。

TCP/UDP解析

★TCP/UDP比率

パケット数比率

TCP	80.5%
UDP	17.2%
ICMP	2.1%

ここで観測されているのは廃棄されてしまうパケットであるため、TCPについてはシーケンスは成立しないため見たままでは少ないが、フラグ状態を見ることによってわかることもある。

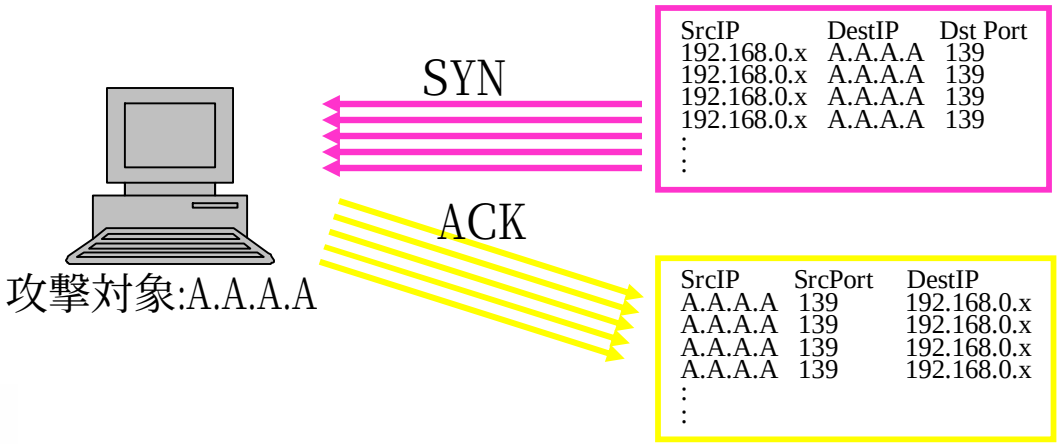
TCP/UDP解析

★TCPについてみる。

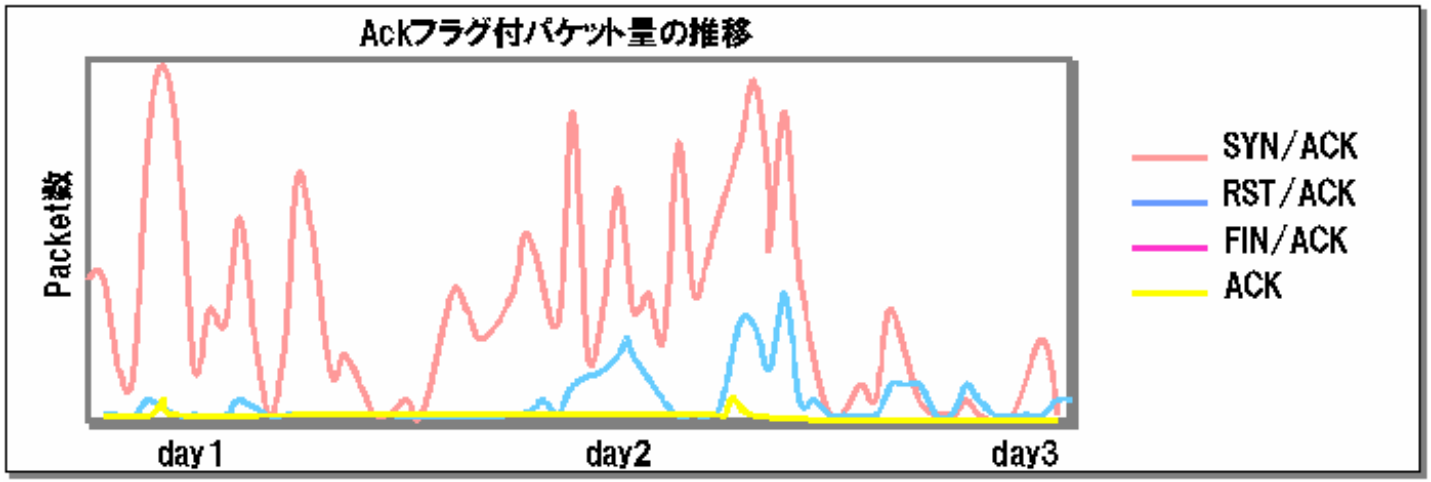
- TCPパケット中99.8%がSYN
  - ・Wormの感染活動に使われるもの。
  - ・Portscan

- ACKパケット
  - ・攻撃を受けた結果(backscatter)。
  - ・感染活動の送信元IPがプライベートアドレス等無効なものへの応答。

ACKパケットの送信元portに着目することでユーザーが受けている各種攻撃、感染活動を観測することができる。



TCP/UDP解析



SYN/ACK		RST/ACK		FIN/ACK		ACK	
TCP139	22.5%	TCP139	21.9%	TCP4662	0.1%	TCP139	1.2%
TCP80	1.0%	TCP135	1.7%	TCP721	0.1%	TCP80	0.1%
TCP4662	0.7%	TCP445	0.8%	TCP80	0.1%	TCP6346	0.1%
TCP60121	0.2%	TCP4899	0.3%	TCP722	0.1%	TCP25	0.1%
TCP135	0.1%	TCP61157	0.1%	TCP25	0.1%	TCP32656	0.1%
TCP25	0.1%	TCP80	0.1%	TCP723	0.1%	TCP4662	0.1%
TCP445	0.1%	TCP4662	0.1%	TCP6349	0.1%	TCP721	0.1%

TCP/UDP解析

★パケットサイズを眺めてみる。

Size	%
0-31	0.85%
32-127	92.47%
128-255	6.06%
256-511	0.13%
512-1023	0.05%
1024-	0.44%

Size	%	Protocol/Port
48	94.49%	TCP
404	1.55%	UDP1434
553	1.21%	UDP1026,1027
28	1.00%	Icmp
78	1.00%	UDP137
56	0.22%	Icmp
110	0.17%	UDP161

■ 傾向

・48byteのpacketはTCPパケット。それ以外はほとんどがUDPパケット。

・UDPパケットについては**プロトコル、宛先ポート、パケットサイズ**からWormを推測できる。

IDSで解析すると

★IDSで眺めてみる。

- TCPに関してはシーケンスが進まないためsignature matchは上手く機能しない。（SCAN系はわかる。）
- UDP/ICMPに関してはサイズまたはDATAから識別がつくのでsignature matchの利用は可能。

検知例

1 BAD-TRAFFIC loopback traffic	266587	371177
2 SNMP public access udp	46700	371177
3 SCAN SOCKS Proxy attempt	27821	371177
4 SCAN Squid Proxy attempt	27225	371177
5 SNMP request udp	1432	371177
6 SHELLCODE x86 NOOP	725	371177
7 SCAN Proxy Port 8080 attempt	224	371177
8 SNMP request tcp	222	371177
9 SCAN FIN	97	371177
0 SNMP trap udp	41	371177

- 宛先ポート別の解析ならportscan detection機能が使える。

  検知条件によりログ量の削減にもなる。

トラフィック解析時の問題点

一般的なセキュリティ関連のトラフィック解析時の問題点について

- 誤検知
- ダイナミックIPの問題
  - ・送信源の特定が困難
  - ・トラフィック変動

Dynamic IPで発信源がころころ変わればそのIPのトラフィック変動は特殊なことになるためトラフィックパターンの変動のみを見るシグネチャへの影響が？

実際の解析、検知においてこんなことが、、、

- TCP/UDP High-port同士の多量トラフィックが長期間にわたり発生
  - 独自アプリケーション
    不要ポート、Wormのポートやその他の条件でフィルタリングを実施すること困難

- DNS多量queryが1IPから1000q/s以上
  - そういうユーザーは結構いる。。。

- 多量SYNが1IPに対して大量に。
  - spamじゃない正規なメールの利用。。。
    上り下りトラフィックが別経路を流れるとsyn floodに見えることも？

- Netflowのsamplingデータ
  - sample rateの問題
    検知できないものがある。挙動の正確な把握が難しい場合がある。
    証拠としての価値。

    低いsample rateの場合は長期期間のデータ収集により精度を上げる必要がある。

- ダイナミックアドレス
  - ダイナミックIPに改竄して数万IP。ユーザー特定が困難

まとめ

- 全体のトラフィックではなく不要トラフィックを解析することにより自網のユーザーのWormの感染活動の状況の把握が可能。
- 網内の感染活動の状況を迅速に把握することが可能。
- 解析対象トラフィックが0.25%程度まで削減できるので解析システムに求められる性能、コストを抑えることができる。
- もともと不要トラフィックの解析であるため誤検知の確率が下げられるのでは。

まとめ(つづき)

- トラフィックアナライザやIDS、市販のアプライアンスを使用すればできることも多いが、やりたいことが微妙にできなかったりする。

  他のさまざまなDatabaseとの照合等も困難。

  今回の解析は独自に開発した解析システム(Mozakin)にてリアルタイム解析を行っている。

- 定点観測、状況観測という意味では十分だが、具体的なセキュリティ対策に利用するにはダイナミックIPユーザーの特定が必要。

  ⇒NetflowとAAAが連動した課金システム等もあることからこういったものの応用でユーザー特定へとつなげられる。

- Virus DNS lookupによるZombie関連の解析等さらに別視点によるトラフィックデータ解析については別の機会があれば。