ボットネットの多段追跡システムの構想と CCC DATAset 2008の利用手法

東京電機大学 大学院 佐々木研究室

三原元　名雲孝昭　芦野祐樹

上原　哲太郎(京都大学)　佐々木良一

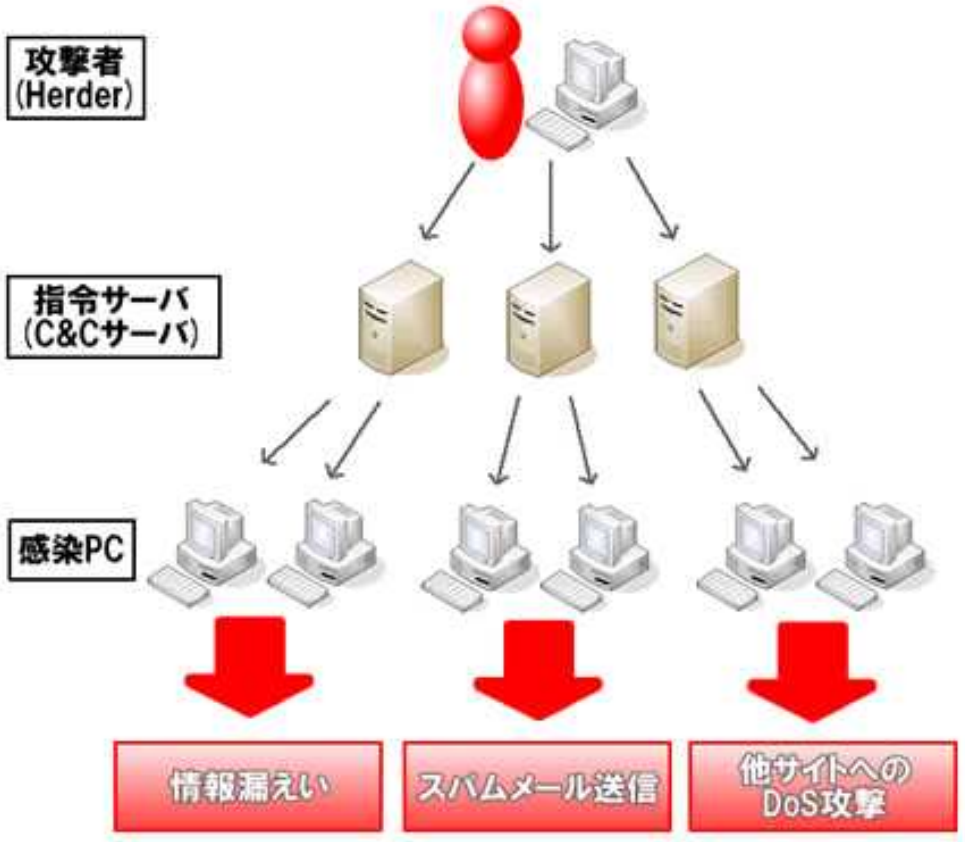
目次

1. はじめに
2. システム概要
3. CCC DATAset2008の解析結果
4. 実験
5. まとめ

目次

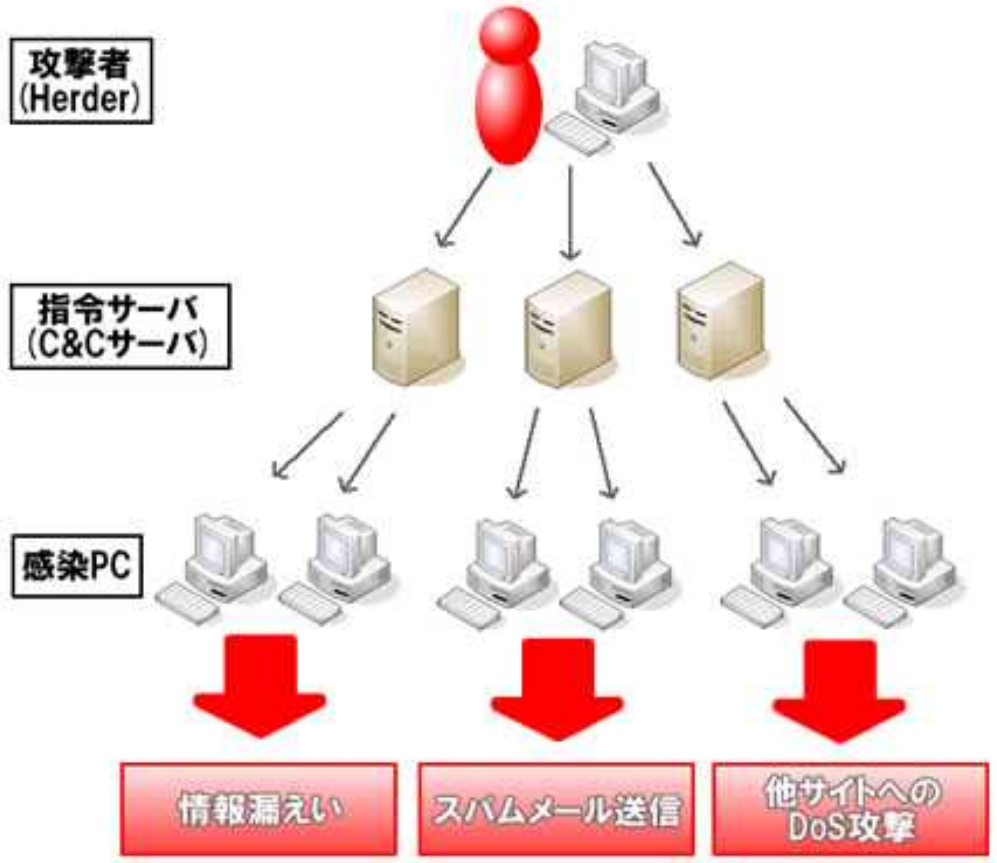
1.はじめに

- 近年ボットネットによる被害が問題となっている



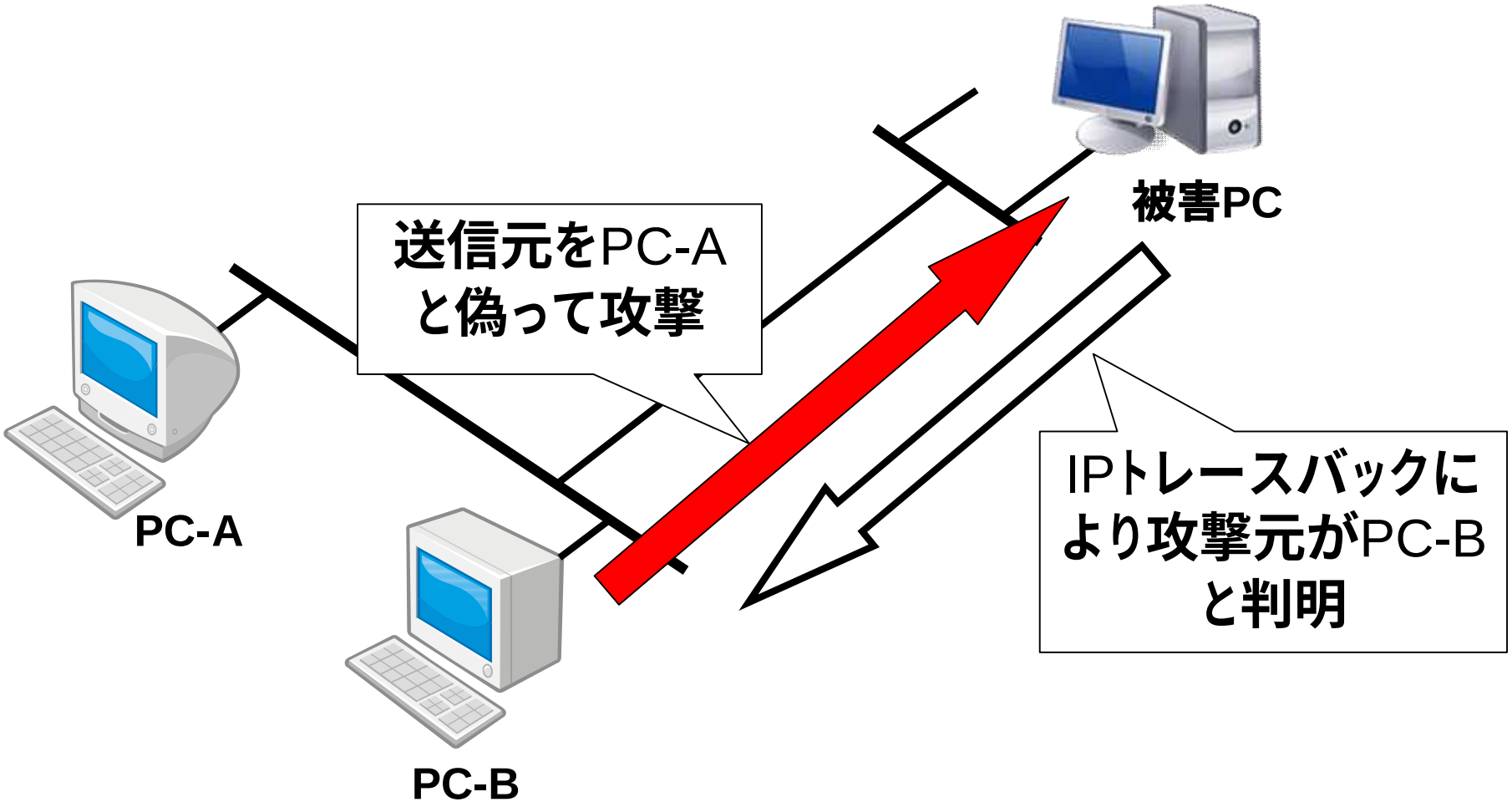
1.はじめに

- ボットネットの問題点
  - ボットPCを隔離しても解決にならない
  - ボットネットは送信元を偽装した攻撃を行うこともある



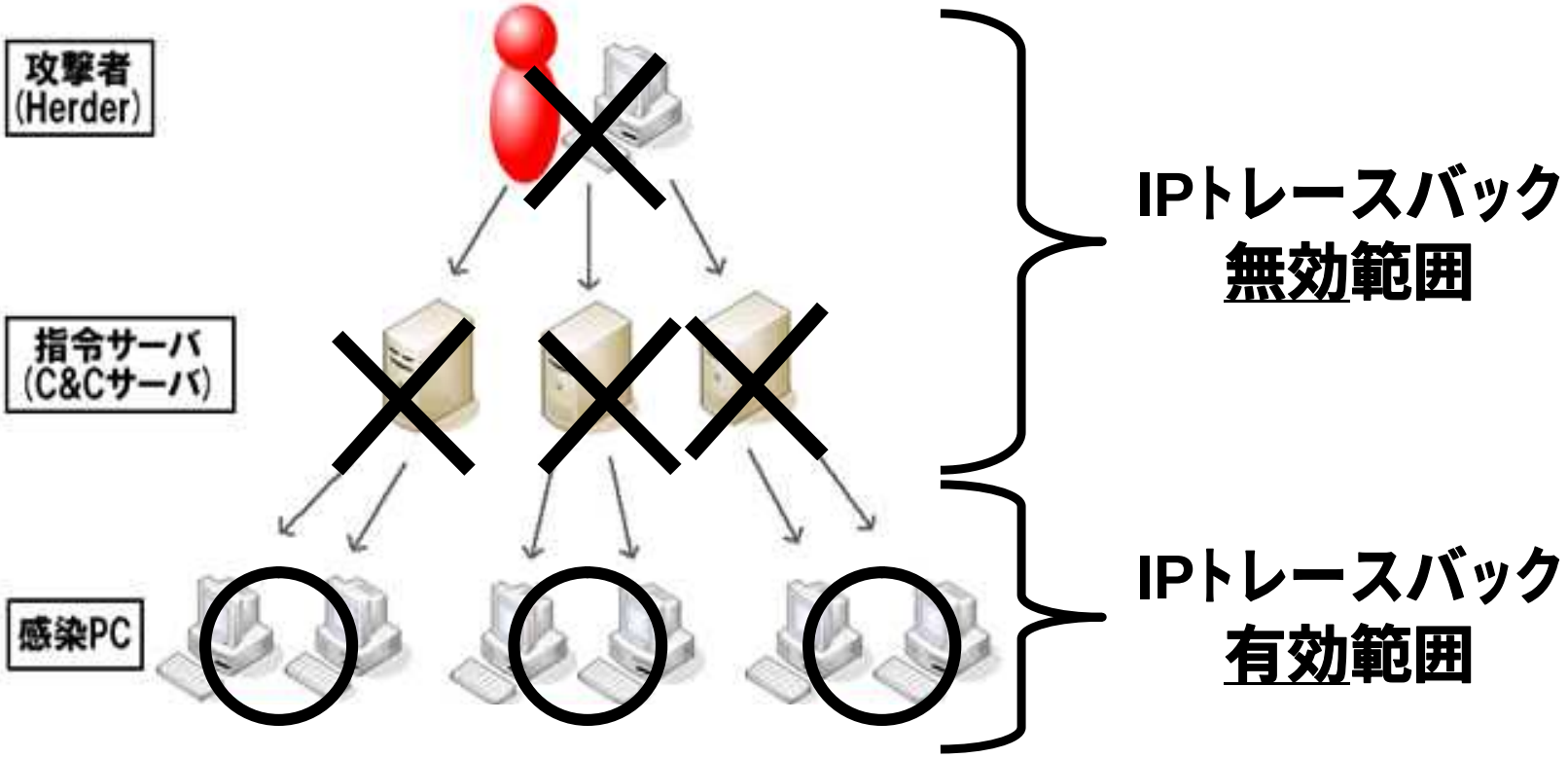
1.はじめに

- 送信元を偽装した攻撃の対策としてIPトレースバックシステムが存在



1.はじめに

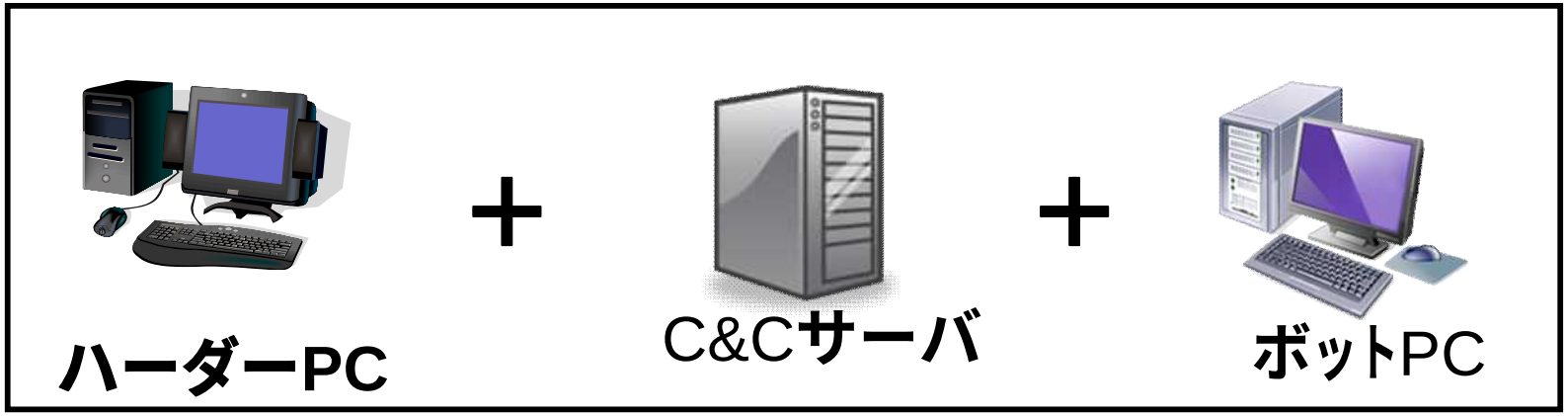
- IPトレースバックではボットPCの特定までのみ



現在，ボットネットに対する根本的な解決手法が求められている

1.目的

ネットワーク管理者同士が情報共有を行うことに
より以下の3つの特定を目的とする

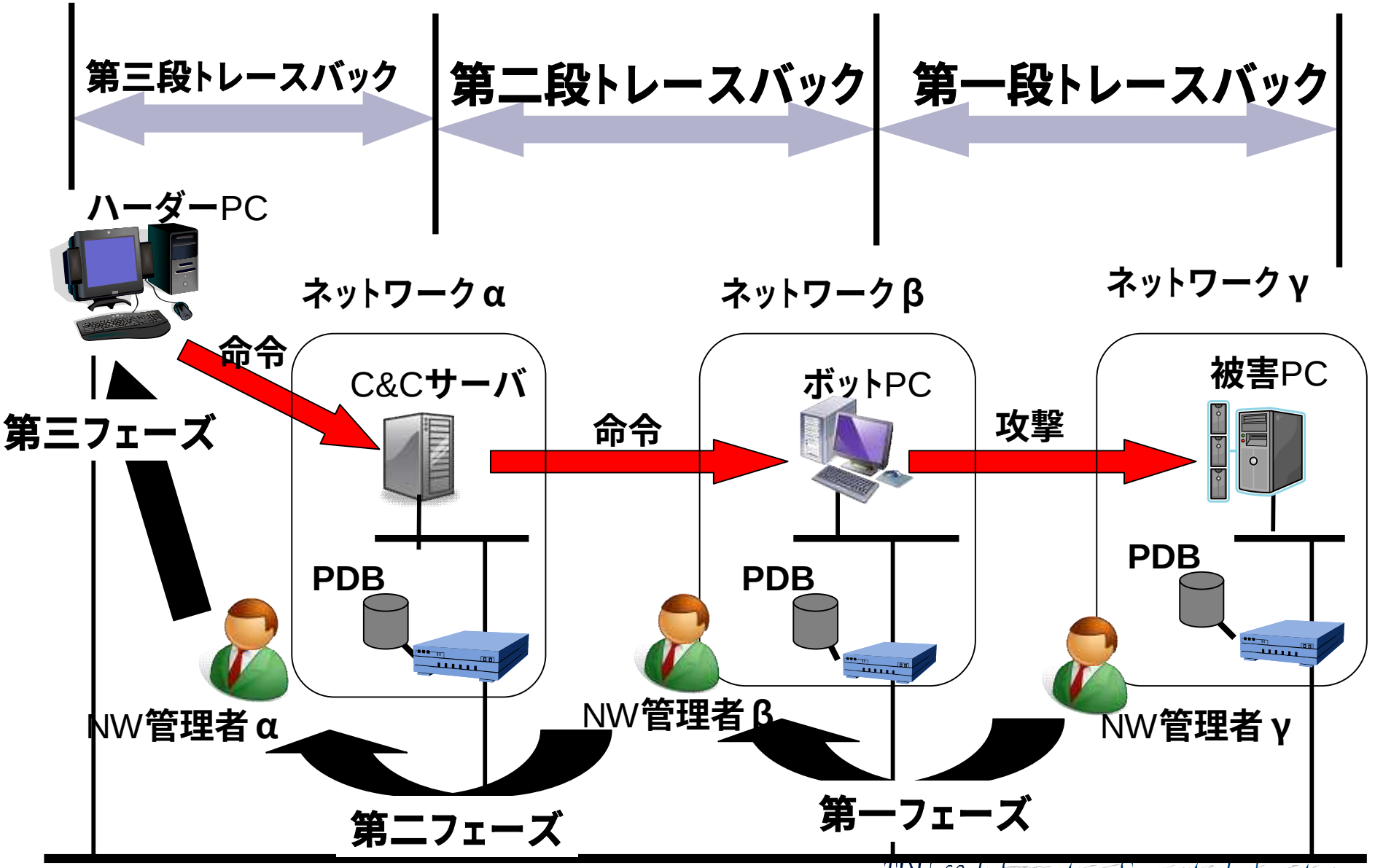


ボットネットの多段追跡システムを構想

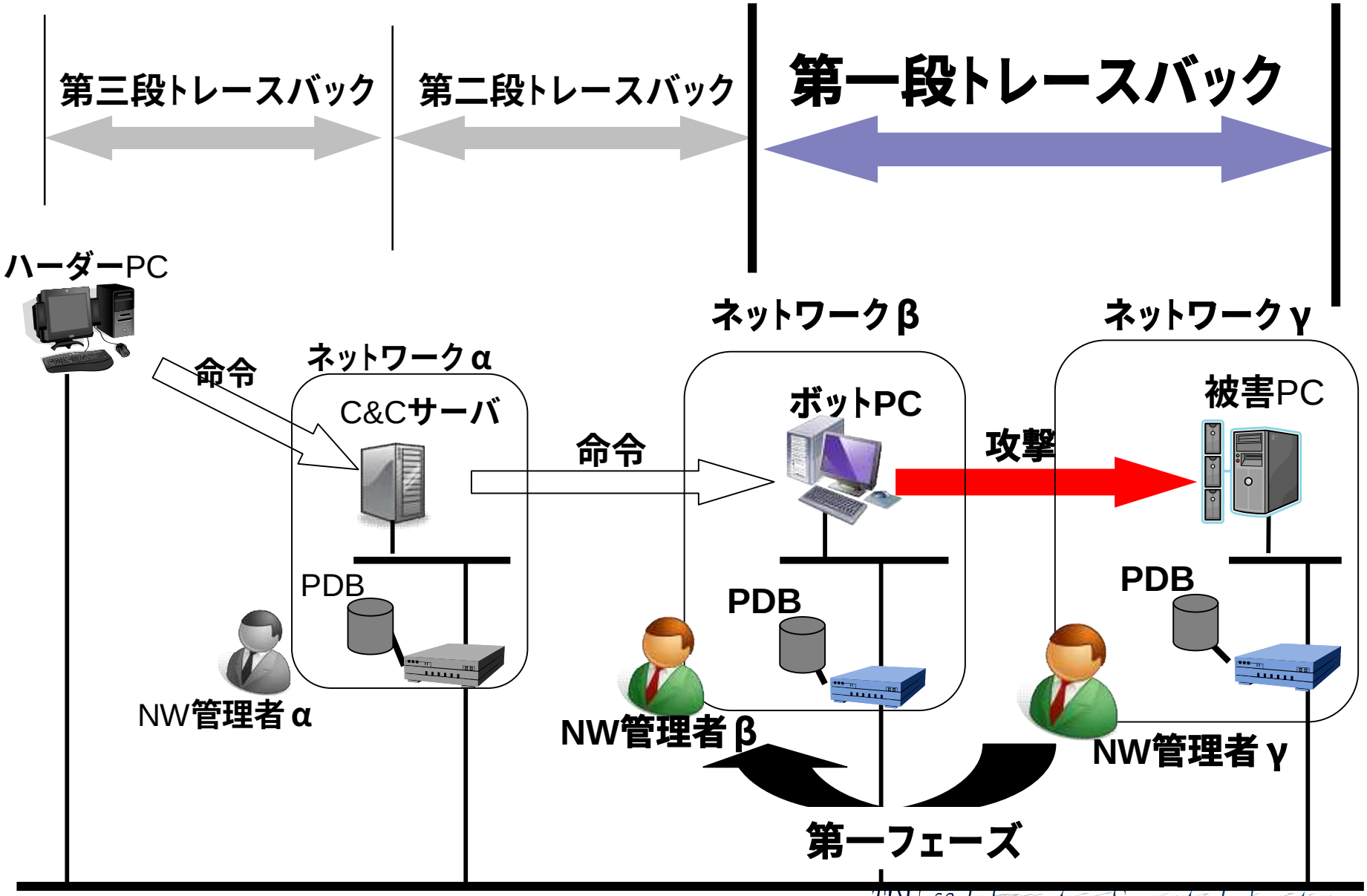
ボットネットの根本的な解決を目指す

目次

2. 多段追跡システム概要

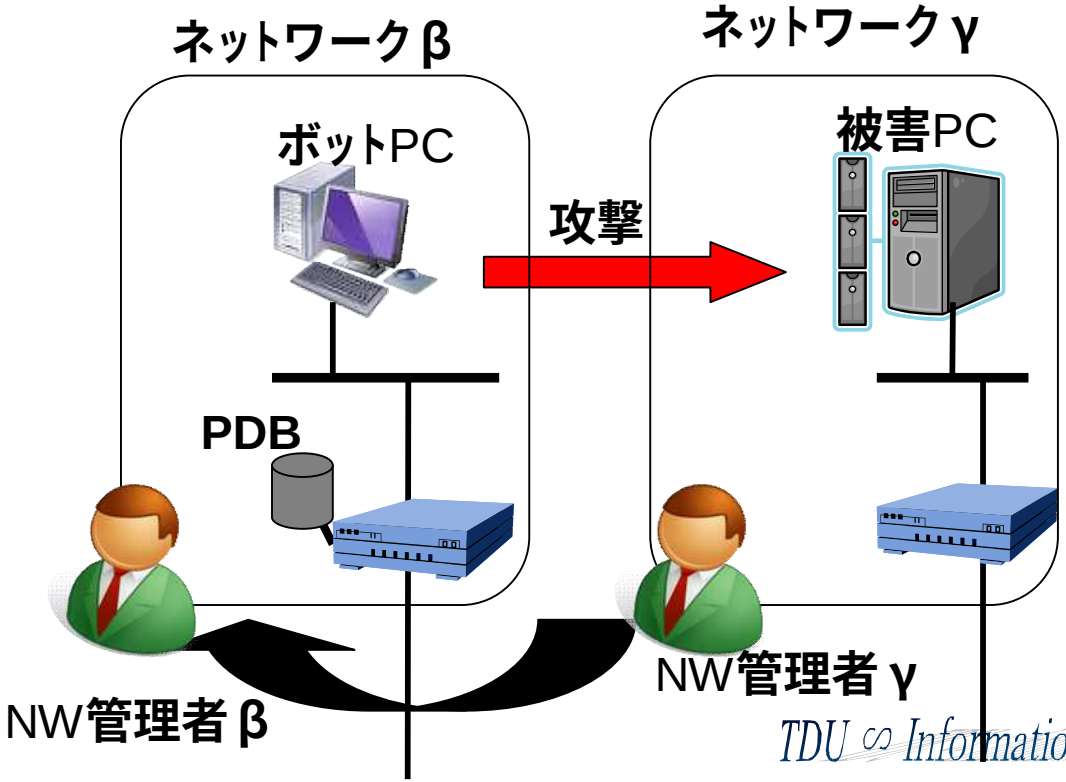


2. 多段追跡システム概要



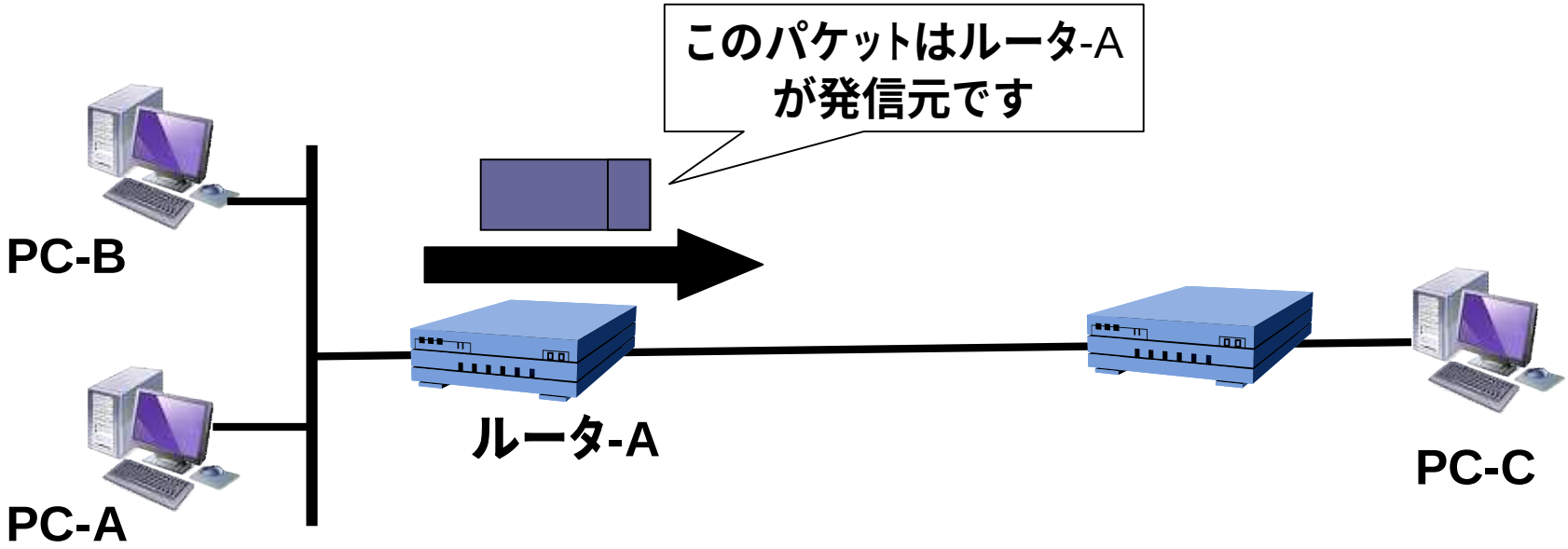
2. 第一段トレースバックシステム

- 目的
  - 攻撃を行ったネットワークを特定
  - ネットワーク管理者間で攻撃の情報共有を行う
- 手法
  - 出国印方式IPトレースバックシステム



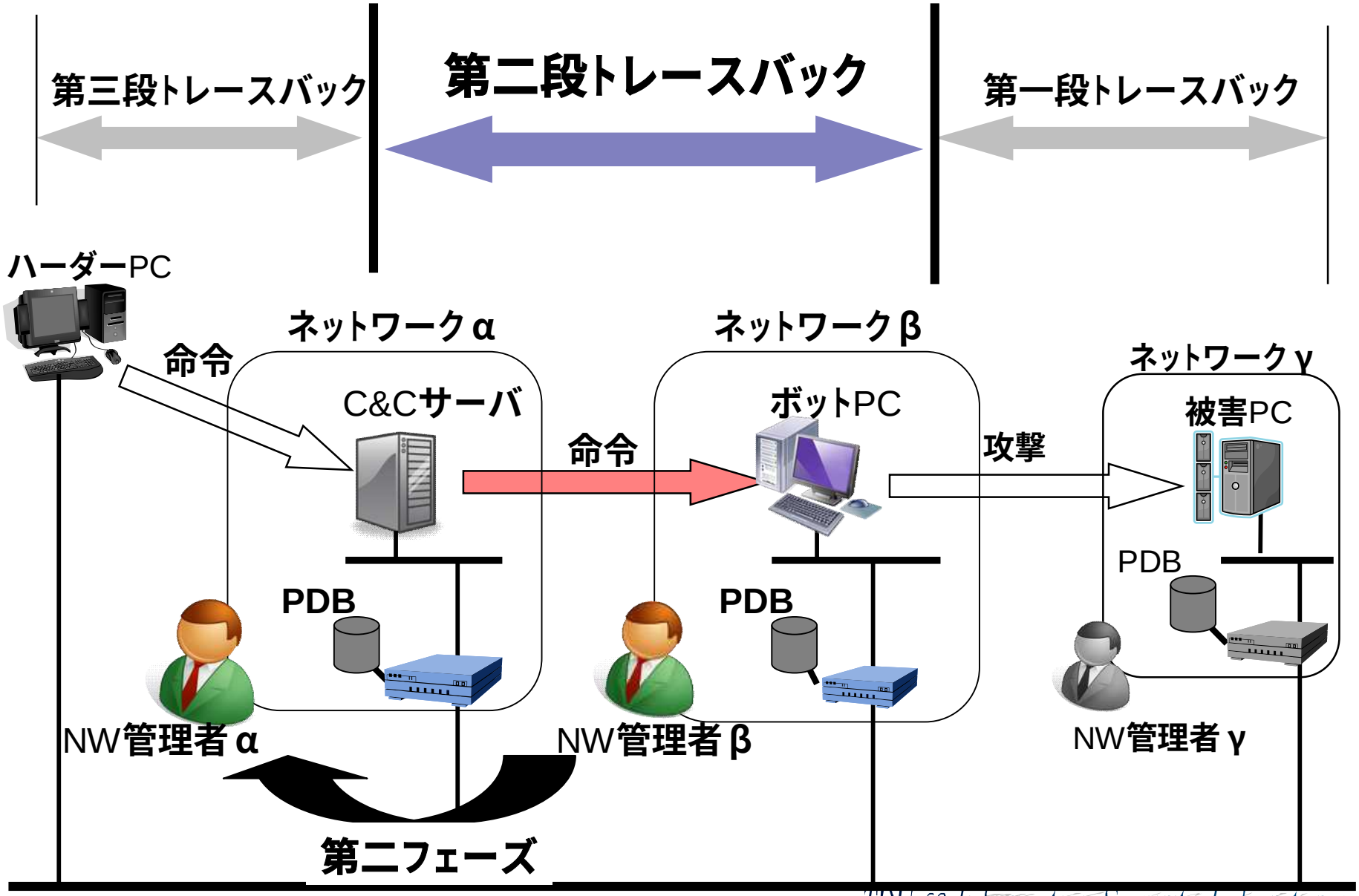
2. 出国印方式IPトレースバックシステム

- パケットにマーキングを行う方式の一つ
- エッジルータが、通過するパケットにルータ自身のIPアドレス等の情報を書き込む



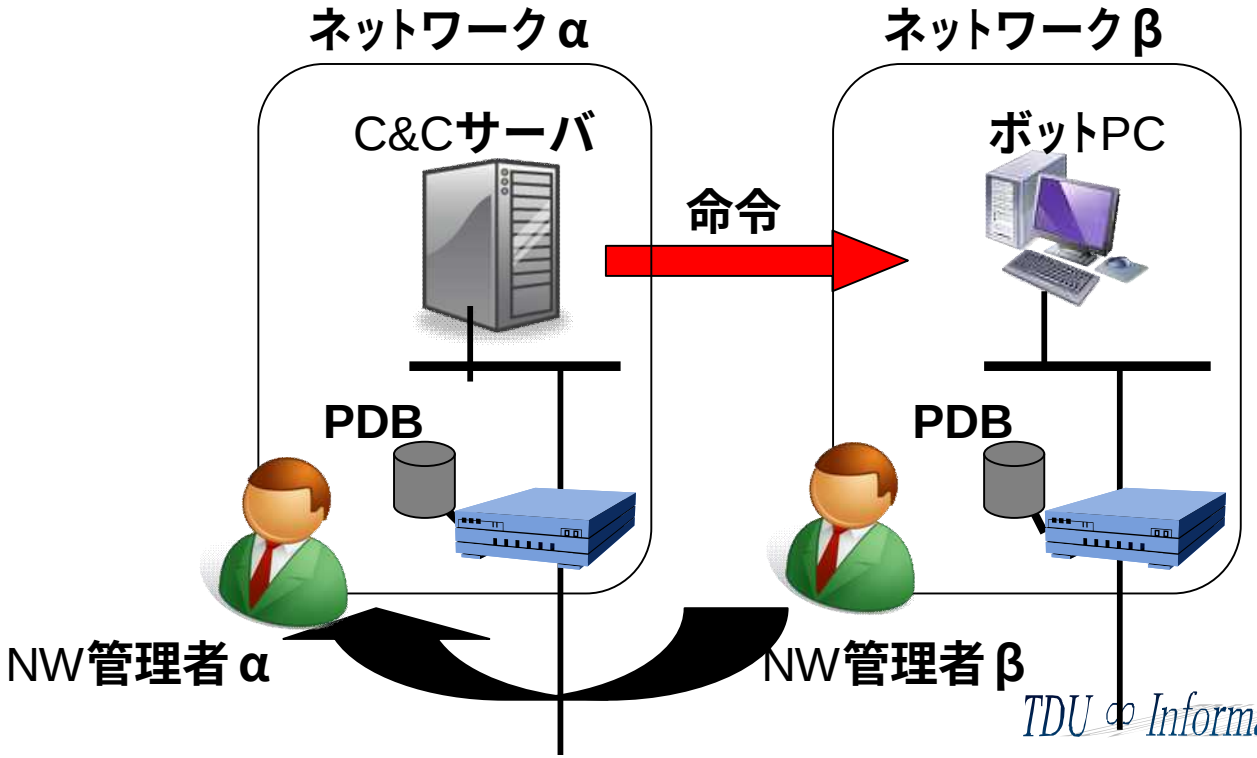
- 送信元が偽装されても真の攻撃元を特定可能

2. 多段追跡システム概要

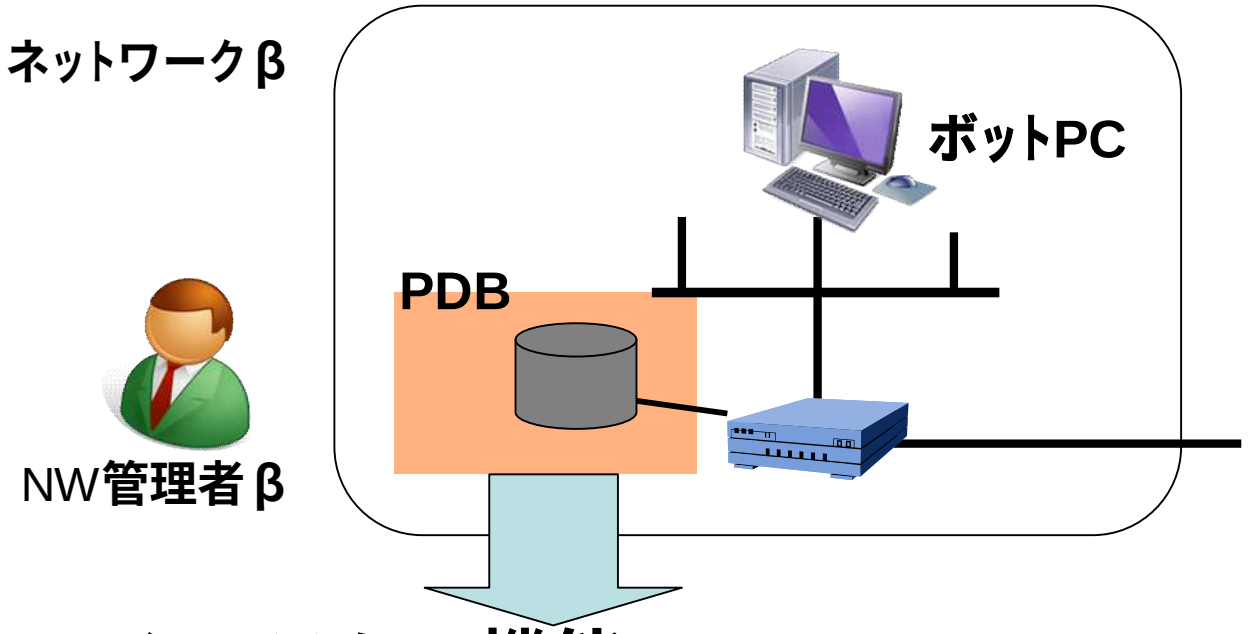


2. 第二段トレースバックシステム

- 目的
  - ボットPC，C&Cサーバ のIPアドレス特定
  - ネットワーク管理者間で情報の共有を行う
- 手法
  - パケットデータベース(PDB)システムを使用



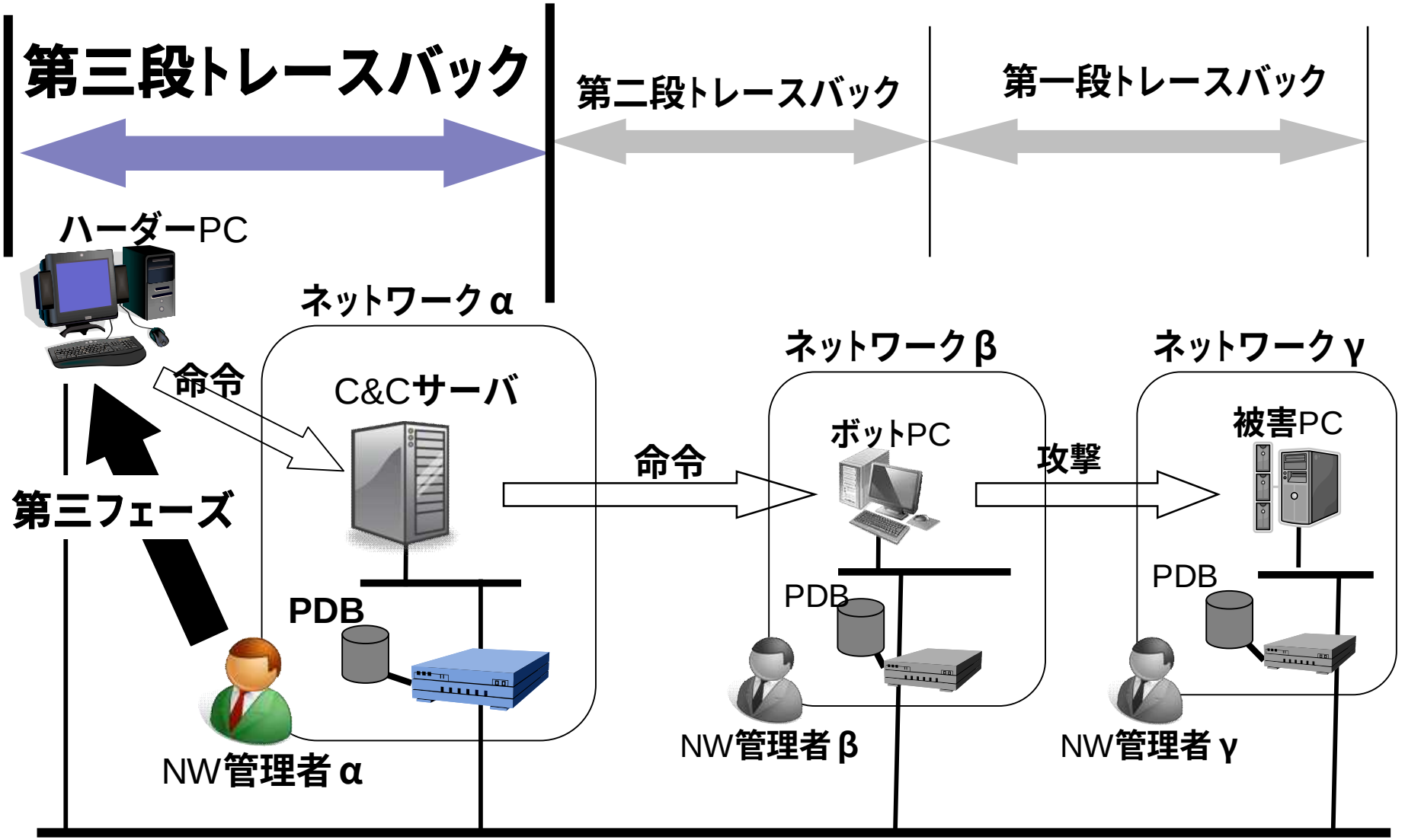
2. 第二段トレースバックシステム



- PDBシステムの機能
  1. ボットPC のIP アドレス特定
  2. C&C サーバのIP アドレス特定

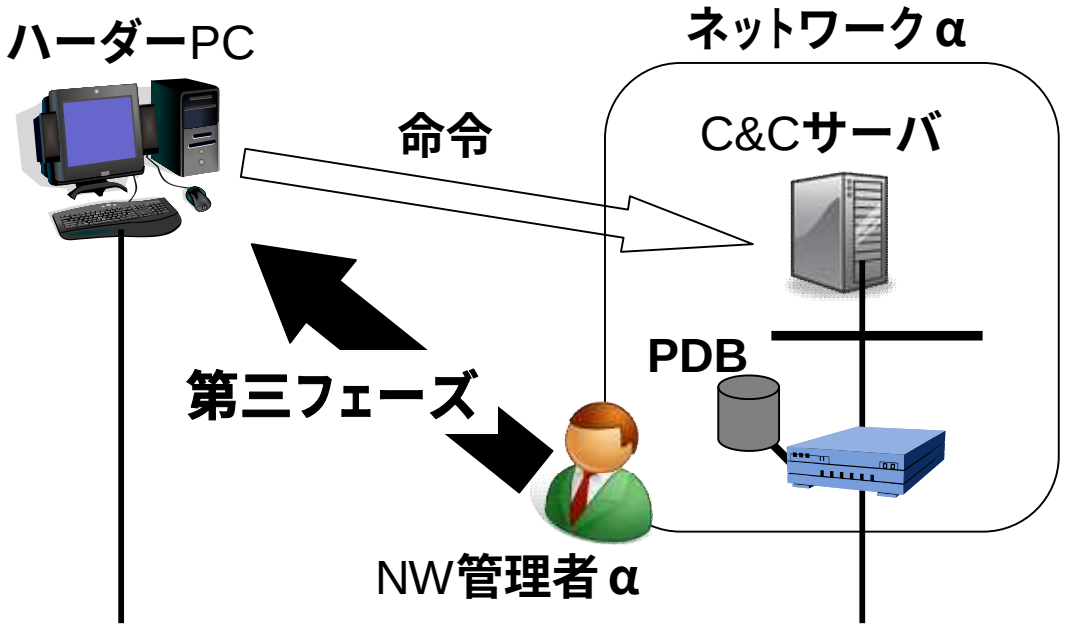
2つを特定する手法を，CCC DATAset2008の解析を用いて考案する

2. 多段追跡システム概要



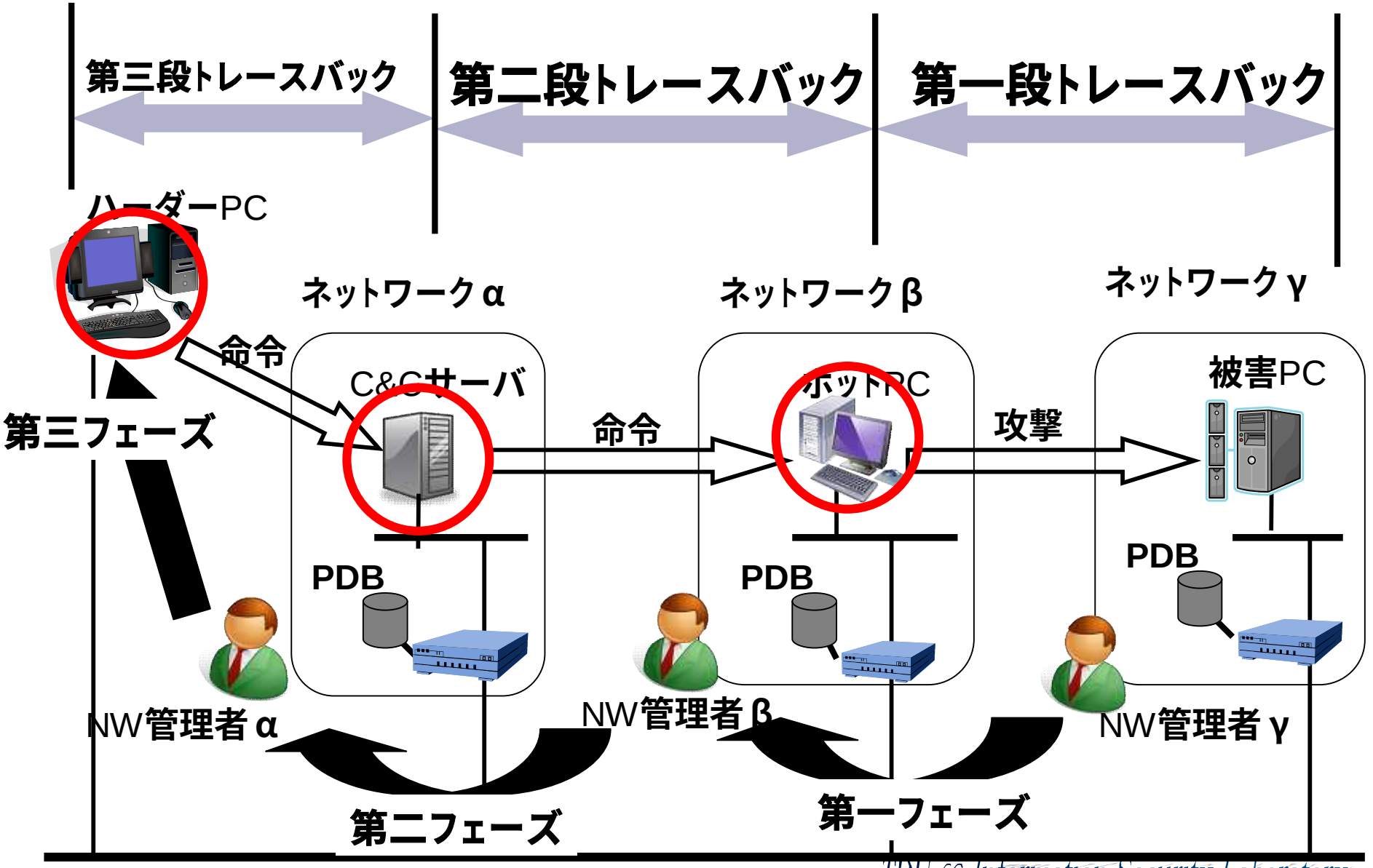
2. 第三段トレースバックシステム

- 目的
  - ハーダーが操作するPCのIPアドレスを特定

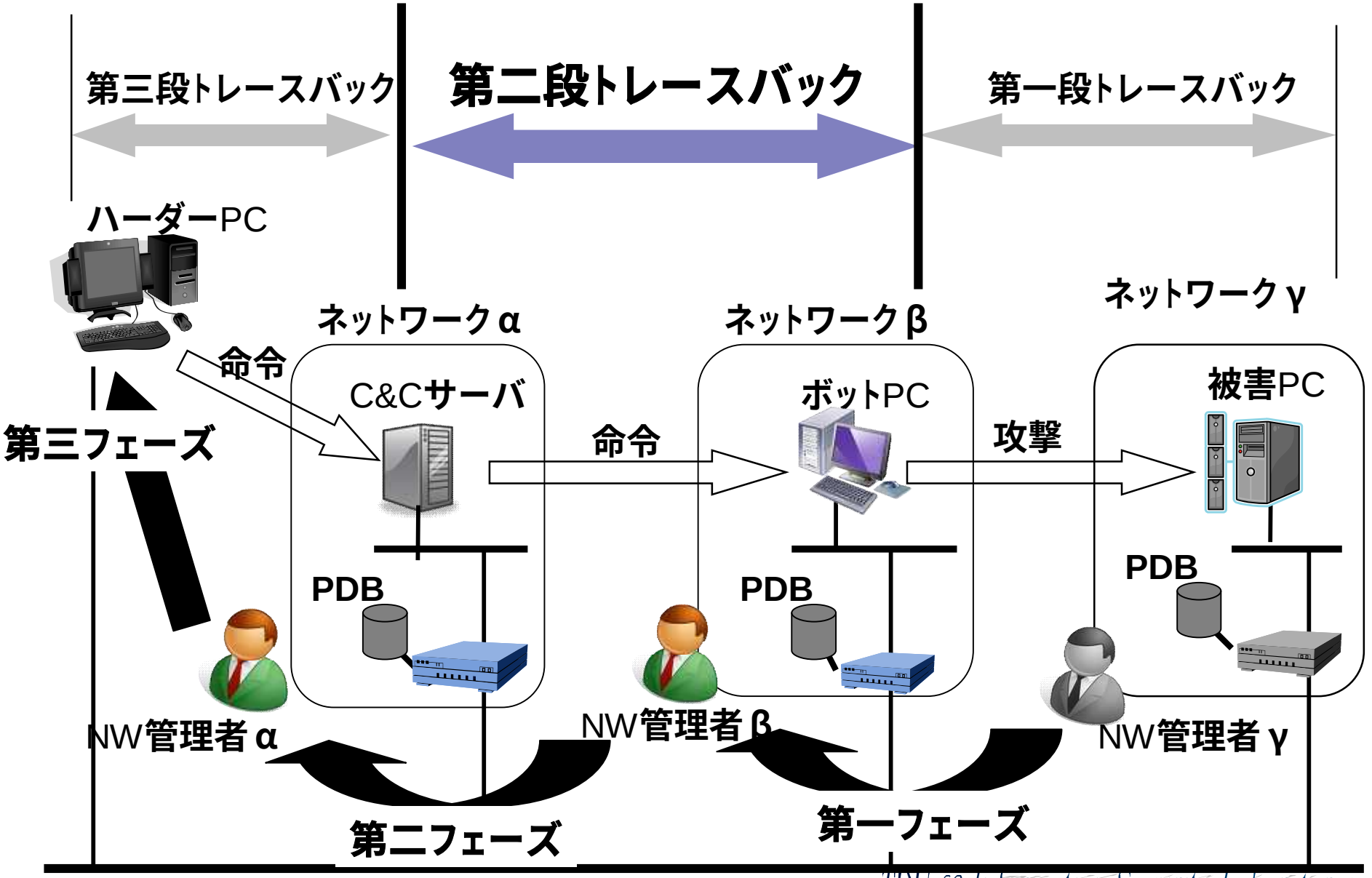


  - 今回実現手法に関しては述べない

2. 多段追跡システム概要



2. 多段追跡システム概要

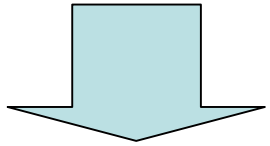


目次

1. はじめに
2. システム概要
3. CCC DATAset2008の解析結果
4. 実験
5. まとめ

3. CCC DATAset2008概要

1. **マルウェア検体**
   - **ハニーポットで収集したマルウェア検体のhash値**
2. **攻撃通信データ**
   - **ボットに感染したハニーポット（以下，ハニーポット）の2 台の通信をtcpdump形式で2 日分キャプチャしたデータ**
3. **攻撃元データ**
   - **ハニーポット112 台の6 ヶ月のマルウェア取得時のログデータ**



攻撃通信データを解析

3.CCC 2008データの解析

- **攻撃通信データから確認できた検体の種別**
  (Symantecによるウイルスチェック)

ボット名称	ボット名称
W32.IRCBot	W32.Bobax!dr
W32.IRCBot.Gen	Backdoor.IRC.Bot
W32.Spybot.Worm	未検出1
W32.Virut.A	未検出2
W32.Virut.B	未検出3
W32.Virut.H	未検出4
W32.Virut.W	未検出5
W32.Virut!gen	

3.特定手法の検討

- 攻撃通信データより確認できた15種のボットに対し
  - ボットPC のIP アドレス
  - C&C サーバのIP アドレス

を特定する要素として以下を候補とした

1. SYN パケットの再送間隔の特徴
2. C&C サーバとの通信の特徴
3. ダウンロードサーバとの通信の特徴
4. 接続先サーバの共通性

3.特定手法の検討

- 攻撃通信データより確認できた15種のボットに対し
  - ボットPC のIP アドレス
  - C&C サーバのIP アドレス

を特定する要素として以下を候補とした

1. **SYN パケットの再送間隔**
2. C&C サーバとの通信の特徴
3. ダウンロードサーバとの通信の特徴
4. 接続先サーバの共通性

3.SYN パケットの再送間隔

- 目的
  - ハニーポットが，外部のサーバに接続する際に行われるSYN パケット送信間隔に特徴はないか

ポート	1～2回目(秒)	2～3回目(秒)	3回目～ポート変更(秒)	1～3回目合計(秒)	全合計(秒)
1031	2.834	5.978	42.069	8.812	50.881
1040	3.045	6.043	42.009	9.088	51.097
1045	成功	–	–	–	–
平均	**2.940**	**6.011**	**42.039**	8.950	50.989
標準偏差	**0.106**	**0.033**	**0.030**	0.138	0.108

ポート	1～2回目(秒)	2～3回目(秒)	3回目～ポート変更(秒)	1～3回目合計(秒)	全合計(秒)
1033	2.861	6.049	12.292	8.910	21.202
1038	2.918	5.994	12.285	8.912	21.197
1039	3.055	5.953	12.168	9.008	21.176
1040	3.046	5.932	12.287	8.978	21.265
1044	3.030	5.992	12.289	9.022	21.311
1047	2.942	5.991	12.330	8.933	21.263
1052	2.964	6.048	12.170	9.012	21.182
1053	2.991	5.936	12.287	8.927	21.214
1054	2.911	5.984	12.284	8.895	21.179
1055	成功	–	–	–	–
平均	**2.969**	**5.987**	**12.266**	8.955	21.221
標準偏差	**0.060**	**0.040**	**0.053**	0.042	0.043

3.SYN パケットの再送間隔

- 結果
  - SYN パケットはどのネットワークアドレスに対しても一定間隔で行われる
  - 間隔時間は，ネットワークアドレス毎に異なる

- 特定手法の有用性
  - 宛先ネットワークアドレス毎の特徴であるため，適切ではない

3.特定手法の検討

- 攻撃通信データより確認できた15種のボットに対し
  - ボットPC のIP アドレス
  - C&C サーバのIP アドレス

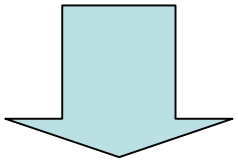
  を特定する手法として以下を候補とした
  1. SYN パケットの再送間隔
  2. **C&C サーバとの通信の特徴**
  3. ダウンロードサーバとの通信の特徴
  4. 接続先サーバの共通性

3.C&Cサーバとの通信の特徴

- 目的
  - ボットとC&C サーバ間で行われる通信に特徴はないか
- 調査項目
  - C&Cサーバが使用するポート
  - C&Cサーバが使用する命令コマンド

3.C&C サーバ使用ポート

- **結果**
  - C&C サーバが使用するポートには共通するものがない



- **特定手法としての有用性は期待できない**

TCP PORT	出現数	割合(%)
6667	51	18
8080	36	12
65520	89	32
5190	23	8
10324	30	10
1863	12	4
18067	7	2
2293	6	2
3938	6	2
7763	8	3
その他	18	7
合計	286	100

3.C&Cサーバ使用命令コマンド

- **結果**
  - **ある程度の命令コマンドを確認できた**

1) ipscan s. s. s. s dcom2 –s
2) advscan dcom135 160 5 0 –b –r –s
3): !get http://**ダウンロードサーバ**URL/~grander/unpr. exe
4) #rs2:=XIa0ZhVFU3q69d0a8Df5/betV/8WnIWAV9LI/B8t
8K9Iwq5+Ttdc7+yHIKyxzLPV6tJ

- **特定手法としての有用性**
  - **IRCを使用しない通信，暗号化された通信の存在**

⇒ 有用性は期待できない

3.特定手法の検討

- 攻撃通信データより確認できた15種のボットに対し
  - ボットPC のIP アドレス
  - C&C サーバのIP アドレス

  を特定する手法として以下を候補とした
  1. SYN パケットの再送間隔
  2. C&C サーバとの通信の特徴
  3. **ダウンロードサーバとの通信の特徴**
  4. 接続先サーバの共通性

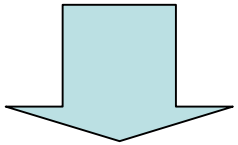
3.ダウンロードサーバとの通信の特徴

- 目的
  - ハニーポットが，ダウンロードサーバからダウンロードした各ボットプログラムとトロイの木馬のデータサイズを調査

マルウェア名称	サイズ(KB)	マルウェア名称	サイズ(KB)
W32.IRCBot	122	W32.Bobax!dr	67
W32.Spybot.Worm	62	未検出1	45
W32.IRCBot.Gen	45	未検出2	165
Backdoor.IRC.Bot	65	未検出3	104
W32.Virut.A	124	未検出4	214
W32.Virut.B	312	未検出5	87
W32.Virut.H	228	lorder.exe	27
W32.Virut.W	47	lox.exe	45
W32.Virut!gen	51	平均	106

3.ダウンロードサーバとの通信の特徴

- 結果
  - ダウンロードサーバは主に80ポートを使用する
  - ダウンロードデータサイズは小さい



- 特定手法としての有用性
  - ボットの通信とは関係ない通信とダウンロードサーバの通信の区別が困難なため、現実的ではない

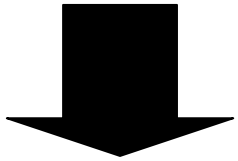
3.特定手法の検討

- 攻撃通信データより確認できた15種のボットに対し
  - ボットPC のIP アドレス
  - C&C サーバのIP アドレス

  を特定する手法として以下を候補とした
  1. SYN パケットの再送間隔
  2. C&C サーバとの通信の特徴
  3. ダウンロードサーバとの通信の特徴
  4. **接続先サーバの共通性**

3. 接続先サーバの共通性

- 目的
  - 同じ種類のボットでも，接続先のサーバが異なるものがある
  - ボットの中には，複数のC&Cサーバやダウンロードサーバに接続しているものがある



  - １５種類のボットが共通して接続するサーバがないか

3. 接続先サーバの共通性

- 結果

接続サーバ	ボット名称	接続サーバ	ボット名称
C&Cサーバ P	W32.IRCBot	ダウンロードサーバ Y	W32.Virut.A
	W32.IRCBot.Gen		W32.Virut.A
	Backdoor.IRC.Bot		W32.Virut.H
	名称不明1－5	ダウンロードサーバ N	W32.Virut.B
C&Cサーバ H	W32.Spybot.Worm		W32.Virut.W
	W32.Virut!gen		
	W32.Bobax!dr		

3. 接続先サーバの共通性

- 結果
  - 15 種類のボットは，4 つのサーバに一つ以上接続する
  - 4つのサーバに接続する際、ボットはDNS サーバに名前解決を行う

C&Cサーバ P
C&Cサーバ H
ダウンロードサーバ Y
ダウンロードサーバ N

  - ボットが行う，4つのサーバの名前解決の動作(以降，名前解決パケット)を用いた特定手法の有用性が期待できる

↓

実験による検証を行う

目次

1. はじめに
2. システム概要
3. CCC DATAset2008の解析結果
4. **実験**
5. まとめ

4.実験内容

1．検知実験

名前解決パケットを利用し，ボットPC，C&Cサーバの特定が成功するか確認

2．誤検知実験

名前解決パケットが正常通信で発生するかを確認

実験にはSnortを使用し，名前解決パケットを検知する

4.検知実験　概要

- 目的
  - ボットの検知
  - C&C サーバの特定
- 環境
  - CCC DATAset２００８
    - データの内，攻撃通信データを使用
  - Snort
    - Snort使用環境

CPU	Pentium4 3GHz
メモリ	2GB
OS	Windows XP Professional
Snort	Snort 2.8.3

4.検知実験　概要

- 概要
  1. Snortのルール に以下の2つのパケット検知ルールを設定
     1.1 名前解決パケット
     1.2 DNSサーバからの名前解決パケットの返答パケット
  2. ルールに設定したパケットを検知
     - 1.1のパケットの送信元IPアドレス
       ＝ボットPCのIPアドレス
     - 1.2のパケット内のIPアドレス
       ＝C&Cサーバの IPアドレス

4.検知実験　結果

- １５種すべての検知が行えた

ボット名称	実験１	ボット名称	実験１
W32.IRCBot	検出可	W32.Bobax!dr	検出可
W32.IRCBot.Gen	検出可	Backdoor.IRC.Bot	検出可
W32.Spybot.Worm	検出可	名称不明1	検出可
W32.Virut.A	検出可	名称不明2	検出可
W32.Virut.B	検出可	名称不明3	検出可
W32.Virut.H	検出可	名称不明4	検出可
W32.Virut.W	検出可	名称不明5	検出可
W32.Virut!gen	検出可		

4.誤検知実験　概要

- 目的
  - 誤検知の有無について確認を行う
- 環境
  - Snort
  - 非ボット感染ネットワークの通信データ

PC台数	20台 (OS:WindowsXP)
通信データ取得時間	24時間
パケット数	約50万パケット

4.誤検知実験　結果

- 4つのドメインに関して，誤検知が発生しないことが確認できた

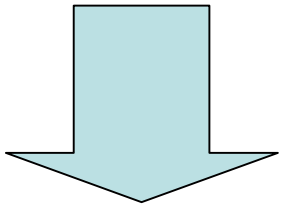
ホスト名	誤検知
C&Cサーバ P	無
C&Cサーバ H	無
ダウンロードサーバ Y	無
ダウンロードサーバ N	無

目次

1. はじめに
2. システム概要
3. CCC DATAset2008の解析結果
4. 実験
5. まとめ

5.まとめ

- 実験の結果より
  - ボットPC，C&Cサーバの特定が実環境下でも同様に行えるものと期待
  - 誤検知の可能性は少ないと期待



- 今回の方式は，第二段トレースバックシステムに有用である
- 多段追跡システム実現の見通しが立てられた

5.まとめ

- 今後は多段追跡システムの第三段の方式の提案を行う
- 引き続きボットに対する調査を行い，第二段トレースバックシステムへの適用を目指す

御清聴ありがとうございました