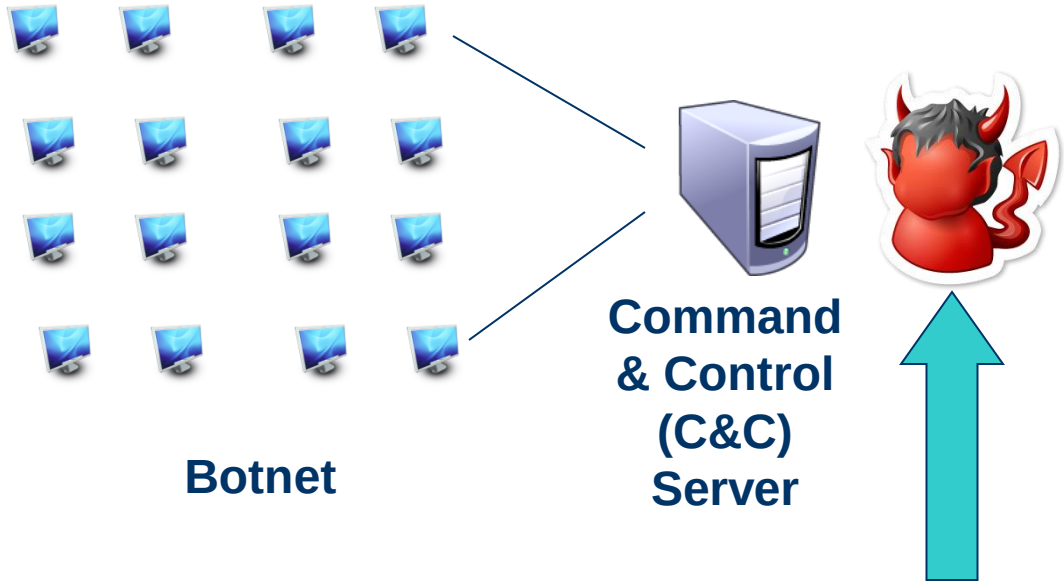
Insights from the Inside: A View of Botnet Management from Infiltration

Chia Yuan Cho[§]
Juan Caballero[§†]
Chris Grier[§]
Vern Paxson[§‡]
Dawn Song[§]

[§]University of California, Berkeley
[†]Carnegie Mellon University
[‡]International Computer Science Institute

Behind C&Cs: Botnet Management



- Management of C&C architecture?
- Response to takedown & recovery?
- Operational activities required to spam?

About MegaD

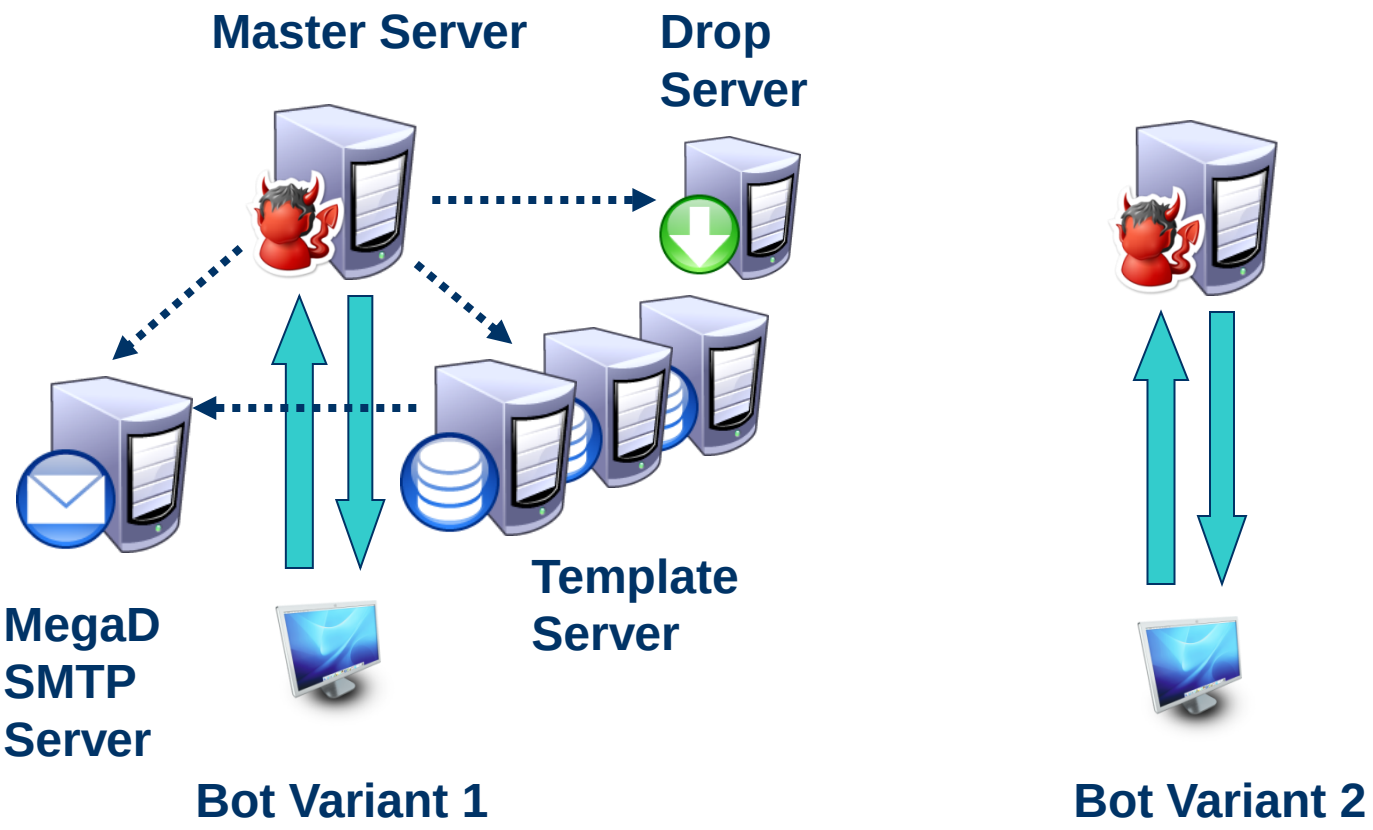
- Mass spamming botnet, appeared 2007
- 1/3 of all spam at its peak
  - 15% last week
- Survived takedown attempt
  - FireEye takedown, **Nov. 2009**
- Our 4-month infiltration
  - Oct. 27, 2009 ~ Feb. 18, 2010

Source: M86 Security Labs

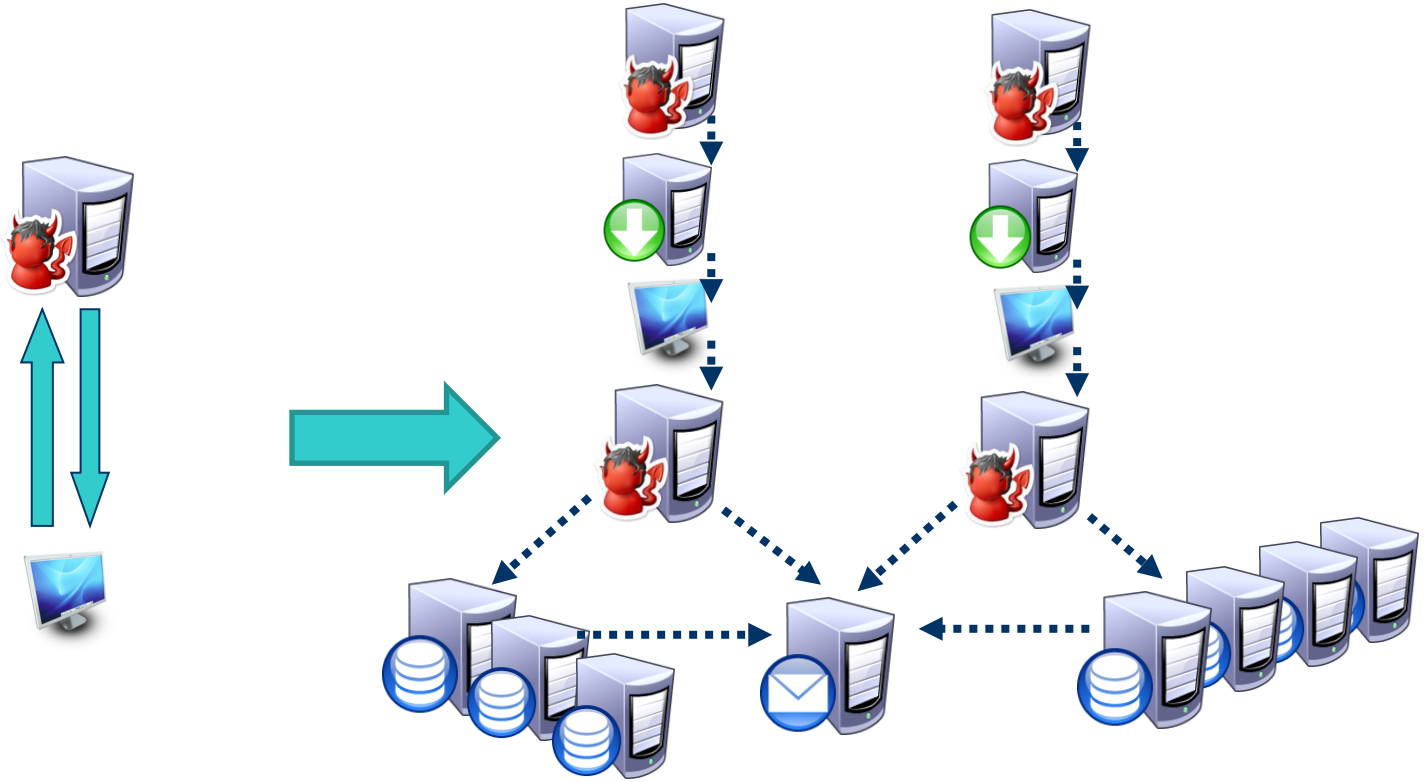
Infiltration Objectives

- Obtain insights on botnet management
  - Monitor spam activities
  - Discover C&C Architecture
    - Enumerate server types

C&C Server Types



Discover C&C Architecture



Techniques for C&C Discovery

Infiltration Techniques

- Creating *Milkers*
  - Bot emulators without malicious side effects
- *Google Hacking*
  - to discover C&C Servers

Infiltration Techniques - Milkers

- **Milkers**
  - To discover C&C architecture: *C&C Milkers*
  - To monitor spam operations: *Template Milkers*
  - IP address diversity: Tor

- **Pre-requisites**
  - C&C protocol grammar
  - Encryption/Decryption functions

Infiltration Techniques - Milkers

- **Exploit design flaws**
  - Bypass Master Servers to loot spam templates
    - Randomize 16-byte bot identifier to Template Server

Infiltration Techniques – Google Hacking

- **Intuition:**
  - Master Servers use port 80 or 443
  - *Camouflaged* as web servers by crafting response to “GET /”

  - Ubiquity of search engines on locating web servers on port 80

Infiltration Techniques – Google Hacking

- **MegaD C&C's crafted response to "GET /"**

```
HTTP/1.0 200 OK Server: Apache/1.3.37
Content-Type: text/html; charset=iso-8859-1

<html>
  <head>
    <title>  test page  </title>
  </head>
   <body>
     <a href='http://www.microsoft.com/'>microsoft.com</a>
   </body>
</html>
```

Google Hack Returns 4 Unique Results

Web Show options... Results 1 - 6 of 6 for

test page
microsoft.com.
doretorza.com/ - Cached

test page
microsoft.com.
www.doretorza.com/ - Cached

test page
microsoft.com.
selementusaks.org/ - Cached

test page
microsoft.com.
kildamindak.net/ - Cached

test page
microsoft.com.
www.kildamindak.net/ - Cached

test page
microsoft.com.
216.32.90.186/

Verified with
C&C milkers

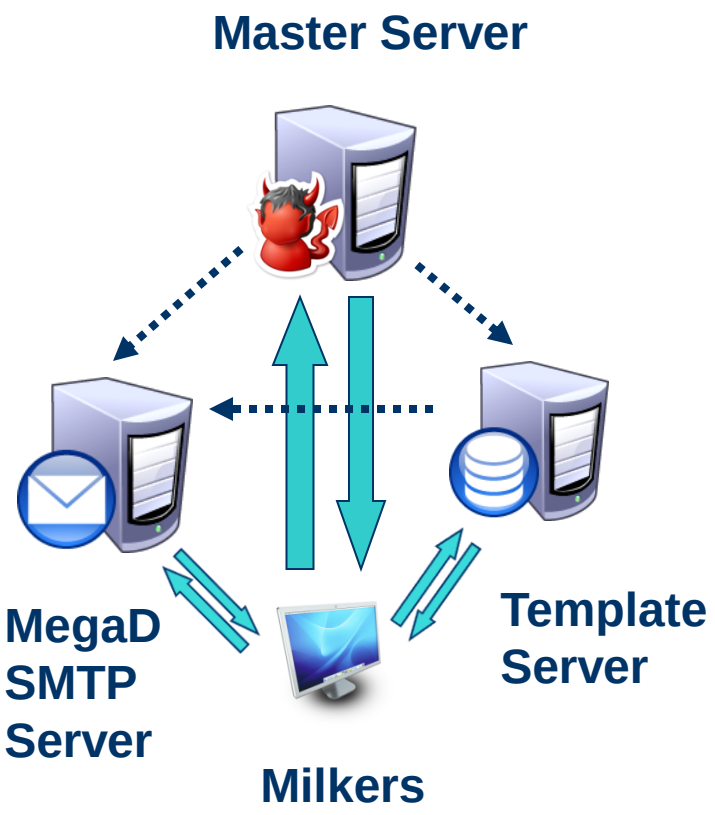
Insights from Infiltration

- Takedown and Recovery
- View of C&C Architecture
- Botnet Management Structure

Insights from Infiltration

- Takedown and Recovery
- View of C&C Architecture
- Botnet Management Structure

Start of Infiltration: Oct. 27



FireEye Takedown: Nov. 6

Inside the Takedown

- Takedown Monitoring
  - Template contents **remain unchanged** for 1 week after takedown
  - First sign of recovery: **1 week** later, on Nov. 13
    - Templates updated to point to **new** SMTP Server
  - **16 days** after takedown, MegaD's spam exceeded pre-takedown level[1]

- Inferences
  - Lack of backup hosting providers / infrastructure
  - Time taken to setup new infrastructure = 1 week

[1]Source: M86 Security Labs

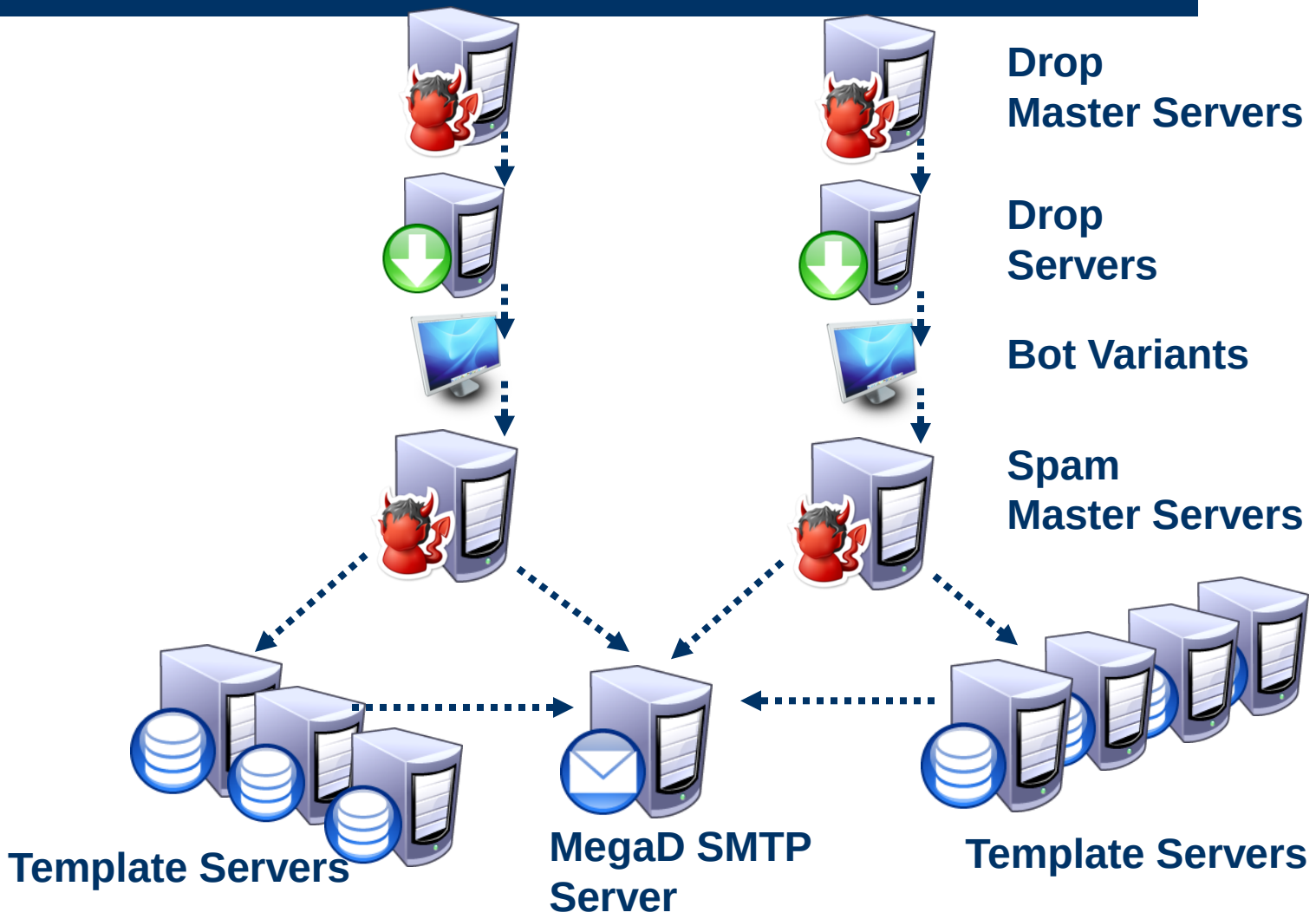
MegaD's Takedown Recovery

- Two possibilities:
  1. **Resilience:** Remnant servers redirect remaining bots to new C&C servers
  2. **New Bots:** Push out new MegaD binaries
     - MegaD known to use generic downloaders (e.g. *Piptea*)
     - Pay-Per-Installation (PPI) model
     - As cheap as $6 / 1000 installs
- Significance
  - *Did not* rely on resilience mechanisms
  - Ease of pushing out new binaries to recover within 16 days

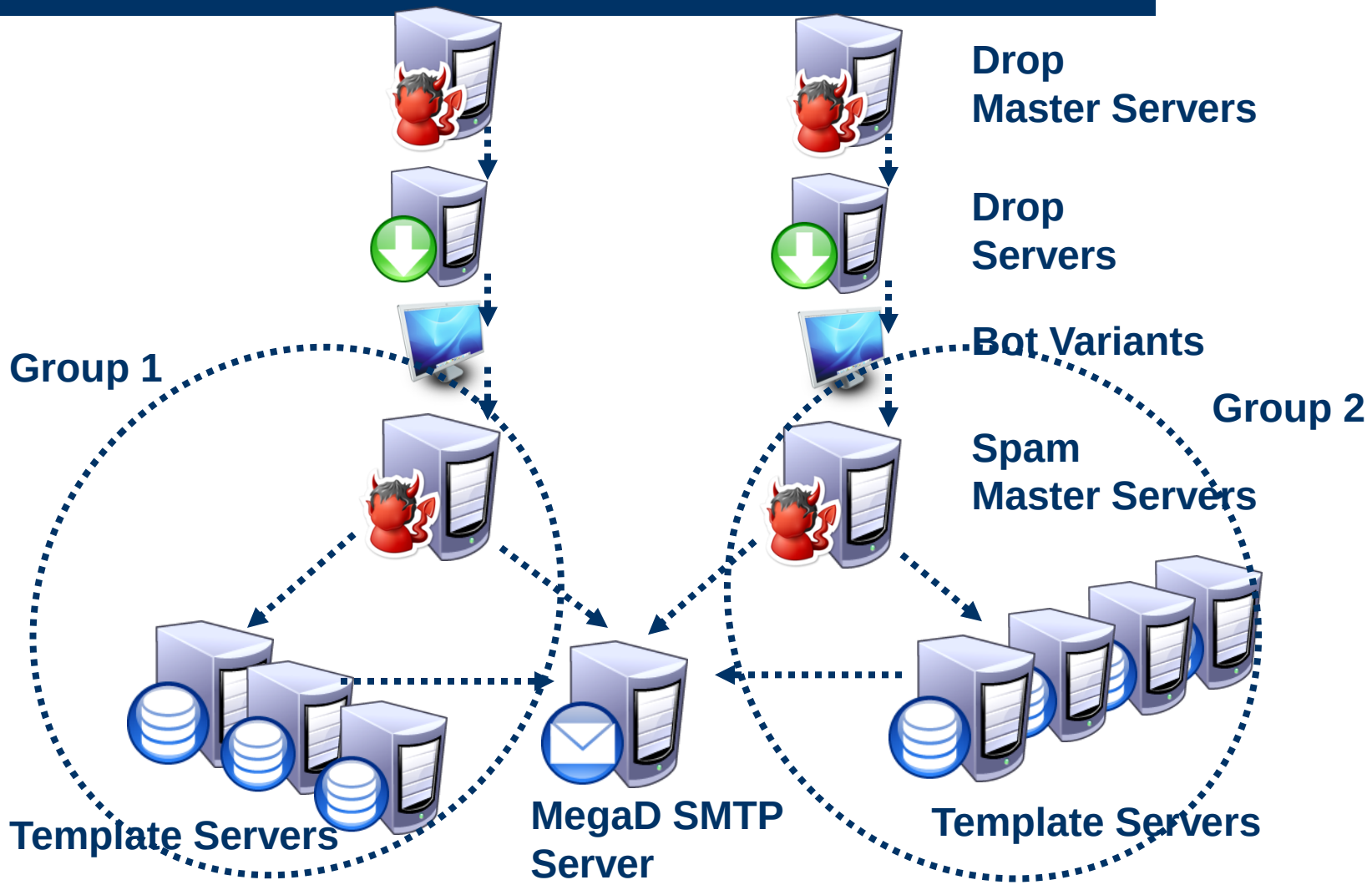
Insights from Infiltration

- Takedown and Recovery
- View of C&C Architecture
- Botnet Management Structure

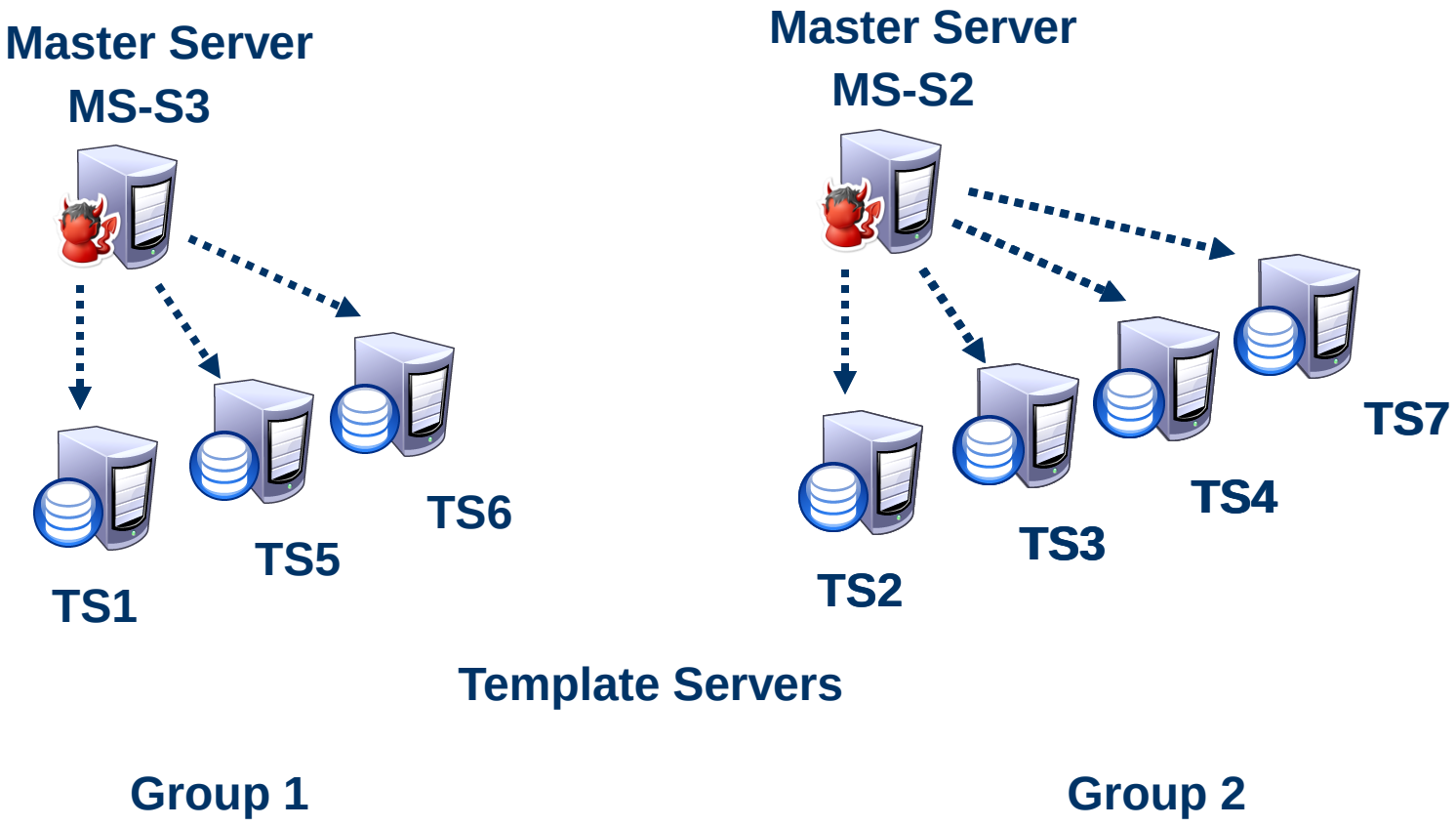
End of Infiltration: Feb. 18



Evidence #1: Differences between Groups



Evidence #1: Differences between Groups



Differences between Architecture Groups

- Possible reasons:
  - Ongoing damage from takedown in Group 2?
  - Different Botmasters?

- More clues from template analysis …

Insights from Infiltration

- Takedown and Recovery
- View of C&C Architecture
- Botnet Management Structure

Spam Template Milking Data

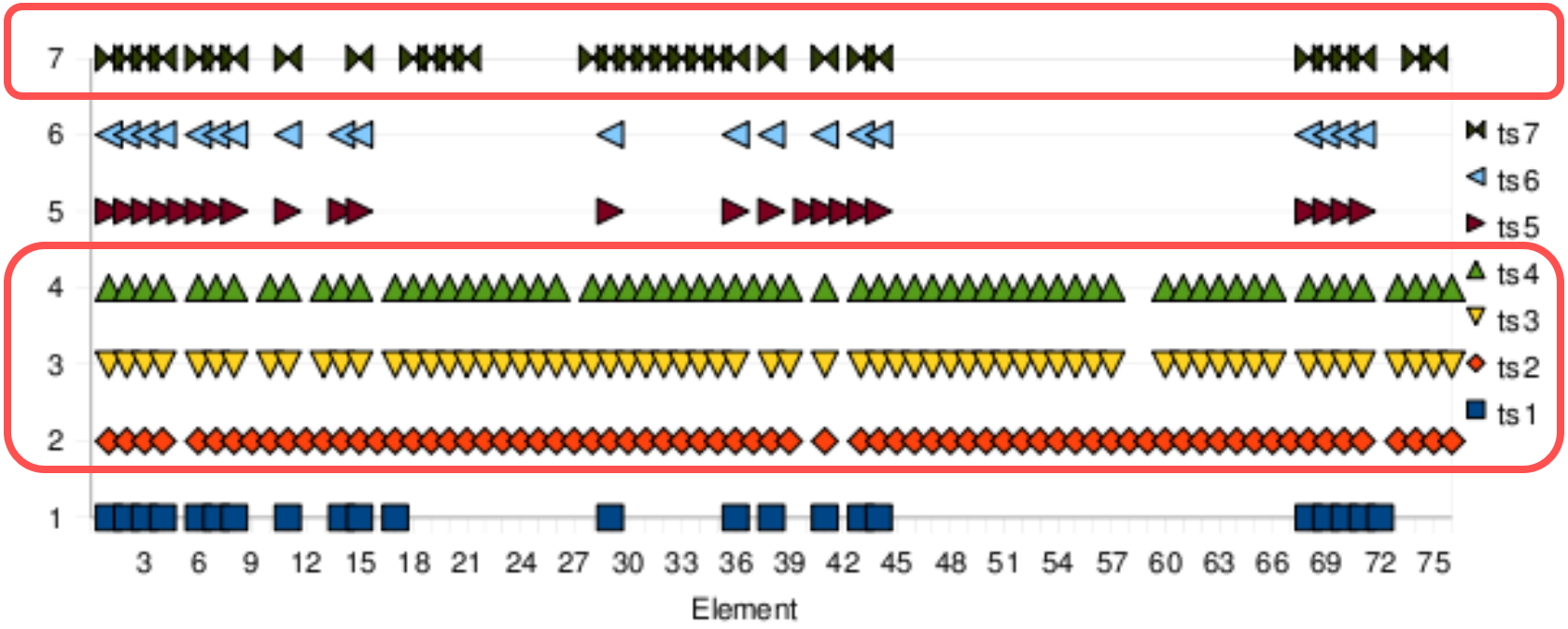
- 271K templates from the 7 Template Servers over 4 months

Template Structure

```
{TEMPLATE}
 To: <{MAILTO_NAME}>
 Subject: {_DIKSBJ_0}
  <HTML> <BODY> {_BODY_HTML} </BODY> </HTML>
{/TEMPLATE}

{TEMPLATE_DATABASE}
 {BODY_HTML}
     <br><A href="http://{_URLS_0}/">Unsubscribe</A>
 {/BODY_HTML}
 {DIKSBJ}
   Freelance Job request
   Career Advice from the experts
 {/DIKSBJ}
 {URLS}
    mainhumble.com
    farown.com
 {/URLS}
{/TEMPLATE_DATABASE}
```

Evidence #2: Differences in Template Structure



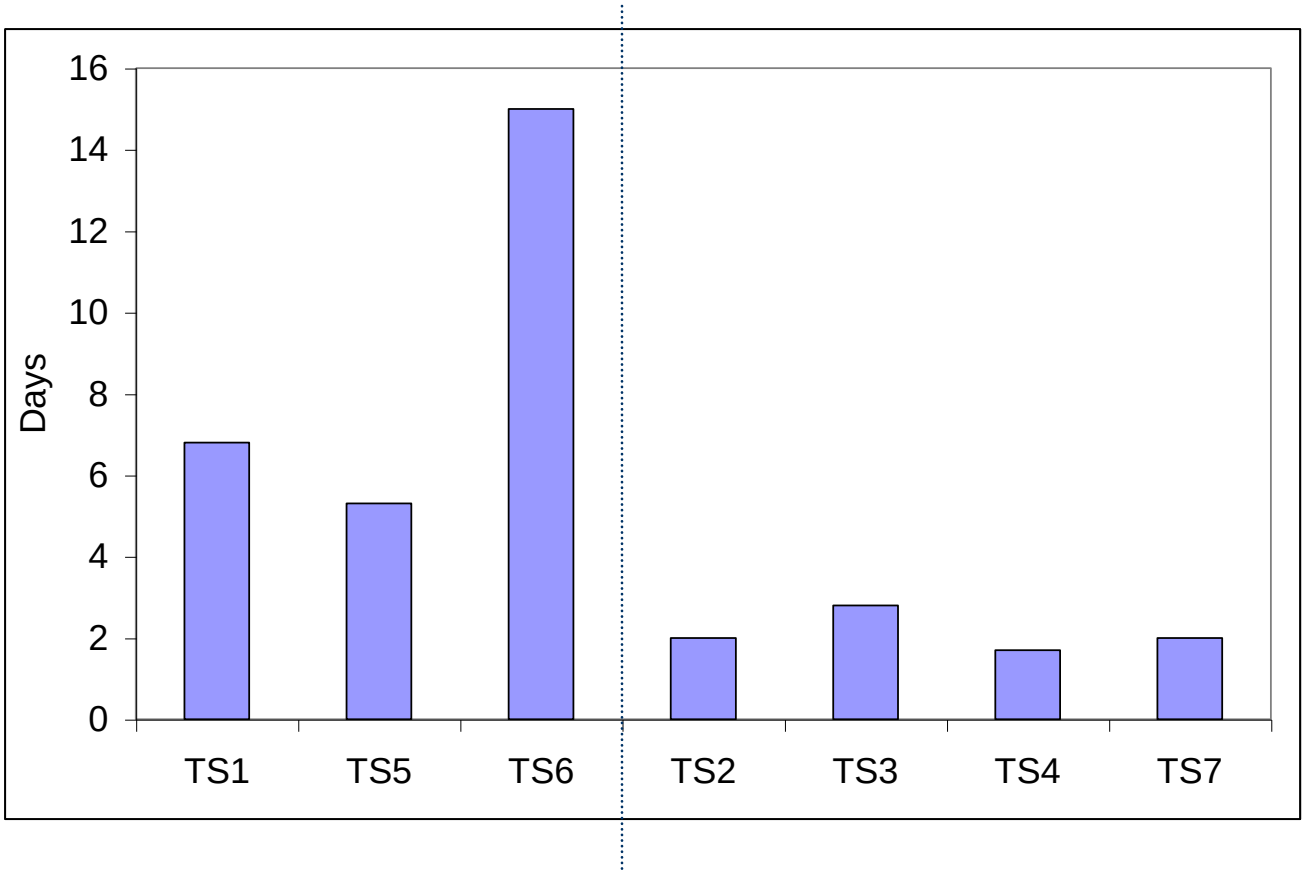
Group 2

Evidence #3: Updates to Polymorphic Elements

- **We identify 3 types of polymorphism:**
  - Single-Set Polymorphic: Fixed set
    - Eg: Outlook Express email signatures
  - Every-Set Polymorphic: Auto-updated set (by TS)
    - Eg: Image Links
  - Multi-Set Polymorphic: Fixed set for days
    - Manually-updated (by Botmaster)
    - Eg: URLs, Dynamic Subjects
- Focus on **Multi-Set Polymorphic** elements
  - Require *sustained effort* from Botmaster for continual updates

Evidence #3: Updates to Polymorphic Elements

Days between Dynamic Subject Updates



Group 1 **Group 2**

Summary of Differences between Groups

Group 1

- Architectural
  - **No** server replacement
- Templates
  - **Common** template **structure** in Group1
  - **Infrequent updates** to polymorphic elements
  - **Single** Viagra campaign

Group 2

- Architectural
  - **Frequent, planned** server replacements
- Templates
  - **Common** template **structure** in Group2
  - **Frequent updates** to polymorphic elements
  - **Diverse campaigns**: Viagra, job scams, money mules

Possible Reasons for Differences

- **Architecture:** Group 2 incurred ongoing damage from takedown?

- **Templates:** Group 2 spam campaigns are more profitable, justifying more frequent updates?

- **Architecture + Templates:** Group 1 and Group 2 are managed by different Botmasters

Related Work

- **Spamalytics**: An empirical analysis of spam marketing conversion (CCS ’08)
  - Chris Kanich et al.
- Studying spamming botnets using **Botlab** (NDSI ’09)
  - John P. John et al.
- **Spamcraft**: An inside look at spam campaign orchestration (LEET ’09)
  - Christian Kreibich et al.
- Measurements and mitigation of P2P-based botnets: A case study on Storm worm (LEET ’08)
  - **Thorsten** Holz et al.
- A multifaceted approach to understanding the botnet phenomenon (IMC ’06)
  - **Moheeb** Abu Rajab et al.

Conclusion

- Infiltration over 4 months
- Techniques:
  - C&C Milking, Template Milking
  - Google Hacking
- Insights:
  - Rich architectural view of MegaD C&C
  - How the Botnet *actually* recovers from a takedown
  - Evidence of *distinct* Botmaster management groups

Thank you!