Projeto SpamPots: Uso de Honeypots na Obtenção de Métricas sobre Abuso de Redes de Banda Larga para o Envio de Spam

Marcelo H. P. C. Chaves

mhp@cert.br

CERT.br – Centro de Estudos, Resposta e Tratamento de Incidentes de Segurança no Brasil

NIC.br – Núcleo de Informação e Coordenação do Ponto br

CGI.br – Comitê Gestor da Internet no Brasil

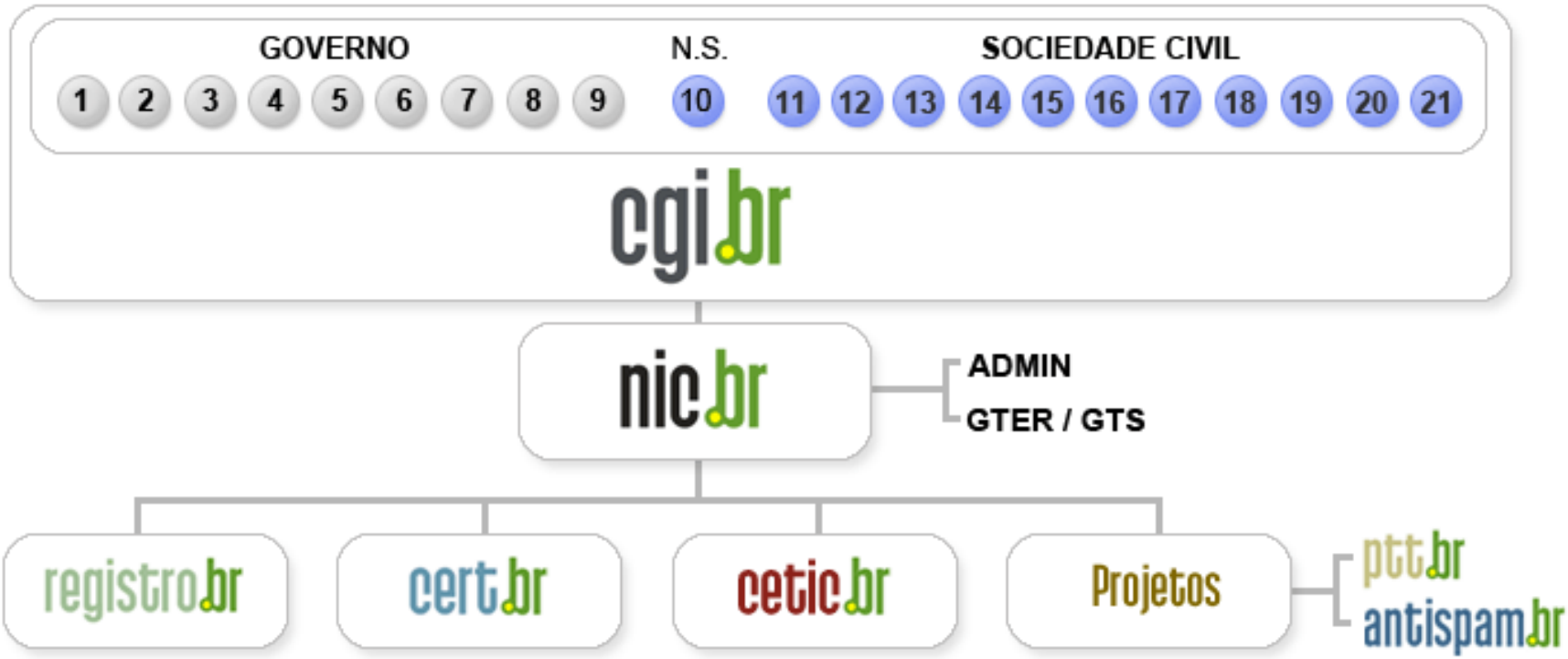
CERT.br

Criado em 1997 para tratar incidentes de segurança em computadores, envolvendo redes conectadas à Internet brasileira, exercendo as seguintes funções:

- Ser um ponto de contato nacional para notificação de incidentes de segurança
- Prover a coordenação e o apoio necessário no processo de resposta a incidentes
- Estabelecer um trabalho colaborativo com outras entidades
- Auxiliar novos CSIRTs a estabelecerem suas atividades
- Prover treinamento na área de tratamento de incidentes
- Produzir documentos de boas práticas
- Aumentar a conscientização sobre a necessidade segurança na Internet

http://www.cert.br/missao.html

Estrutura do CGI.br



01- Ministério da Ciência e Tecnologia
02- Ministério das Comunicações
03- Casa Civil da Presidência da República
04- Ministério da Defesa
05- Ministério do Desenvolvimento, Indústria e Comércio Exterior
06- Ministério do Planejamento, Orçamento e Gestão
07- Agência Nacional de Telecomunicações (Anatel)
08- Conselho Nacional de Desenvolvimento Científico e Tecnológico
09- Fórum Nacional de Secretários Estaduais para Assuntos de C&T
10- Representante de Notório Saber em Assuntos de Internet

11- Provedores de Acesso e Conteúdo
12- Provedores de Infra-estrutura de Telecomunicações
13- Indústria de Bens de Informática, Telecomunicações e Software
14- Segmento das Empresas Usuárias de Internet
15-18- Representantes do Terceiro Setor
19-21- Representantes da Comunidade Científica e Tecnológica

Atribuições do CGI.br

Entre as diversas atribuições e responsabilidades definidas no Decreto Presidencial nº 4.829, destacam-se:

- a proposição de normas e procedimentos relativos à regulamentação das atividades na internet
- a recomendação de padrões e procedimentos técnicos operacionais para a internet no Brasil
- o estabelecimento de diretrizes estratégicas relacionadas ao uso e desenvolvimento da internet no Brasil
- **a promoção de estudos e padrões técnicos para a segurança das redes e serviços no país**
- a coordenação da atribuição de endereços internet (IPs) e do registro de nomes de domínios usando <.br>
- **a coleta, organização e disseminação de informações sobre os serviços internet, incluindo indicadores e estatísticas**

Agenda

- Motivação
- O Projeto SpamPots
  - Cenário de Abuso de Proxies Abertos
  - Arquitetura
  - Honeypots
  - Servidor
- Estatísticas
- Trabalhos Futuros
- Referências

Motivação

O spam é causa de diversos problemas

- *malware*/*phishing*
- queda da produtividade (perda de mensagens, etc)
- aumento dos custos em infra-estrutura (filtros, mais banda, etc)

Politicos, provedores, teles, etc

- São pressionados pelo público em geral a “fazer algo a respeito”
- Têm diversos projetos de lei a serem apreciados
- Não têm dados que mostram o real problema

Motivação (2)

O que ouvimos

- *Proxies* abertos deixaram de ser um problema
- Hoje, apenas *botnets* são usadas no envio de spam
- O Brasil é uma das maiores fontes de spam do mundo

Nossos dados

- Notificações de spam associadas ao abuso de *proxies* abertos aumentaram nos últimos anos
- Nas estatísticas do Consórcio Brasileiro de Honeypots, varreduras por *proxies* abertos sempre fazem parte das 10 portas mais freqüentes

Motivação (3)

Muitas perguntas ainda persistem

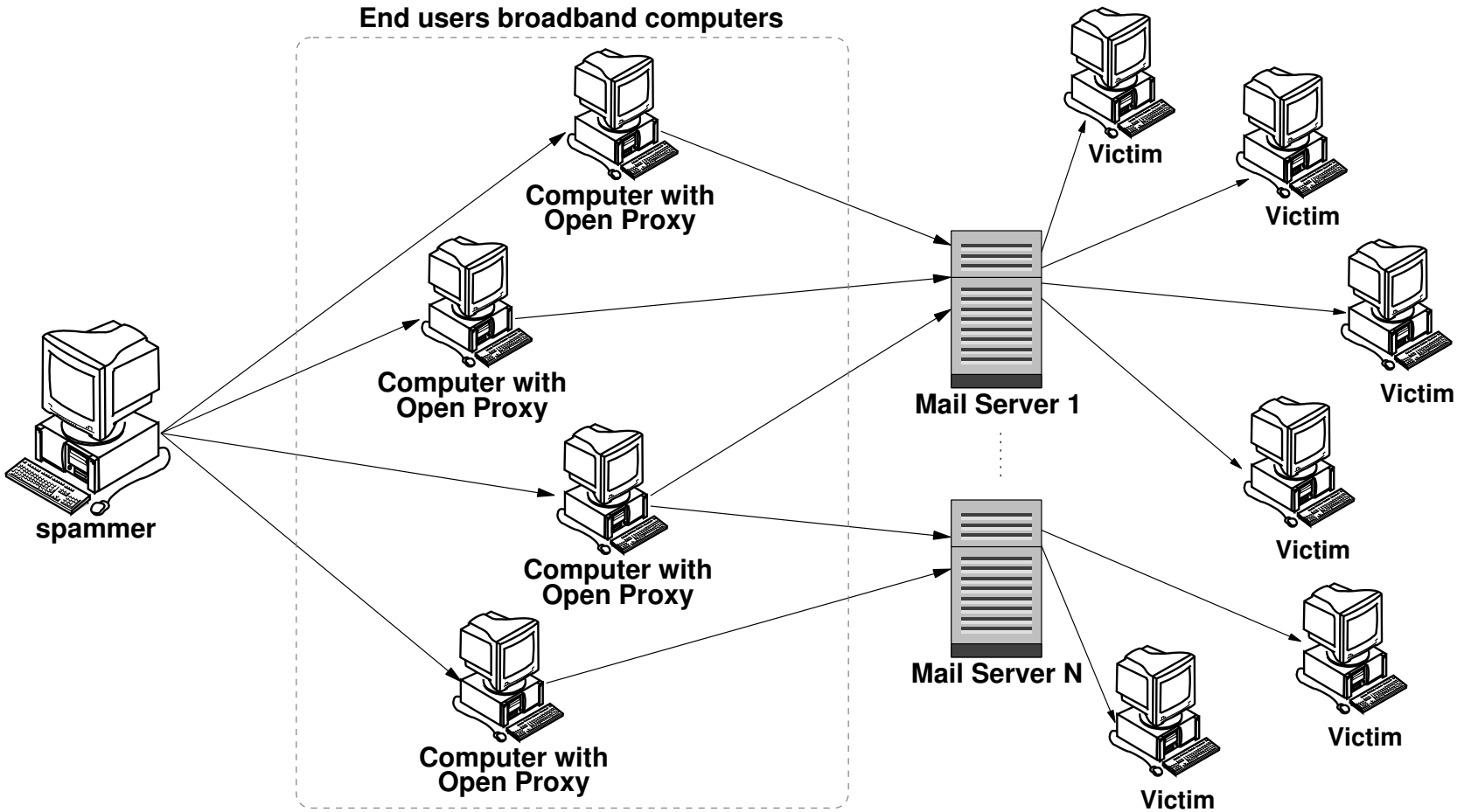
- Como convencer pessoal não ténico sobre a necessidade e efetividade das medidas de mitigação sendo sugeridas?
- Quem está abusando da nossa infra-estrutura? E como?
- Nós temos métricas nacionais ou apenas internacionais?
- Como podemos obter dados e gerar métricas para auxiliar na formulação de políticas e no entendimento do problema?

Há a necessidade de ter mais informações e de melhor entender o problema

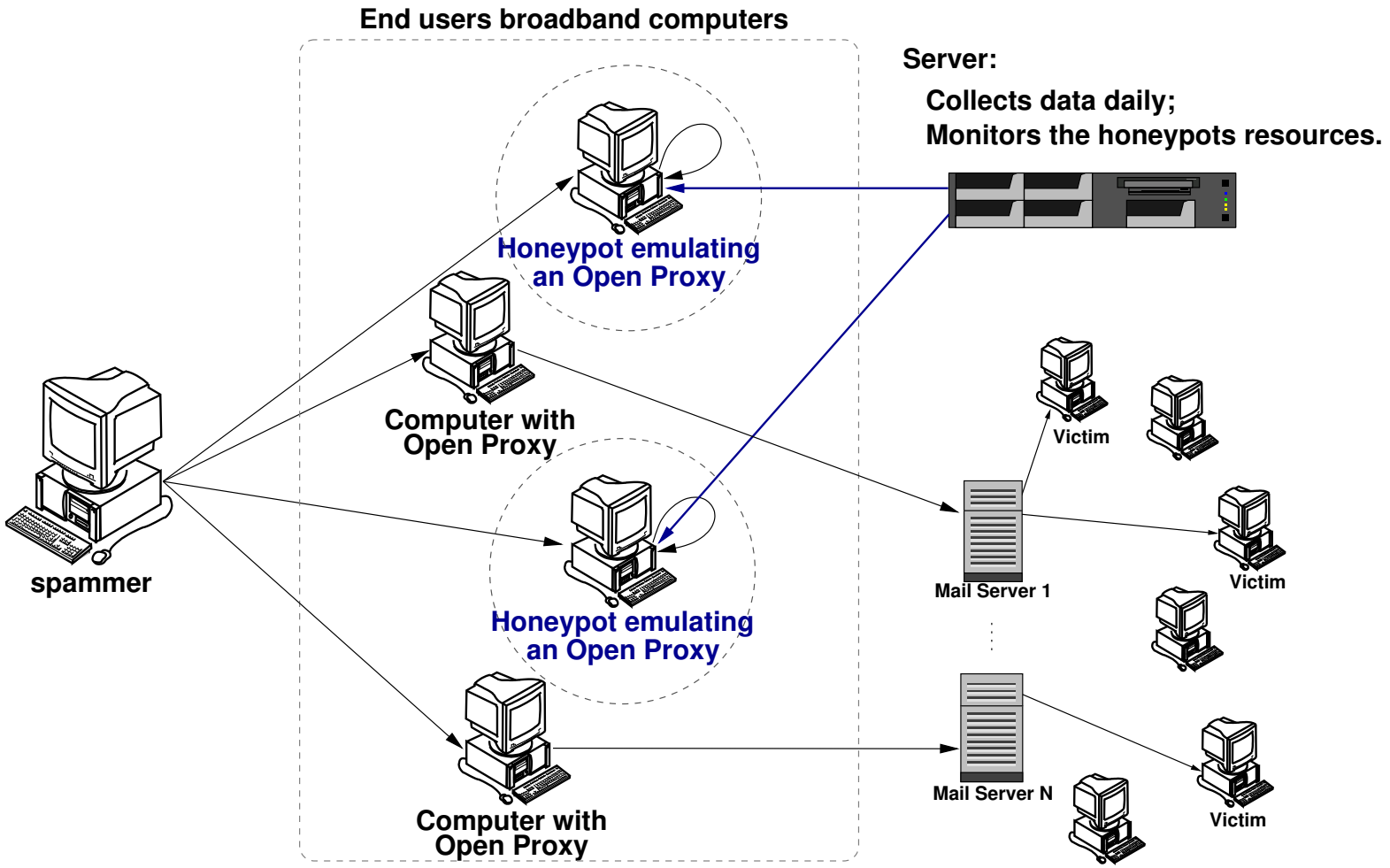
O Projeto SpamPots

- Mantido pelo CGI.br/NIC.br
  - como parte da Comissão de Trabalho Anti-Spam (CT-Spam)
- Implantação de 10 *honeypots* de baixa interatividade, emulando serviços de *proxy*/*relay* aberto e capturando spam
- Instalados em redes de banda larga (ADSL/*cable*), por um ano
  - 5 provedores de acesso banda larga – uma conexão residencial e uma comercial por provedor
- Mensurar o abuso de máquinas de usuários finais para o envio de spam

Cenário de Abuso de *Proxies* Abertos



Arquitetura



Honeypots

OpenBSD: sistema operacional (SO) adotado

- número de problemas de segurança extremamente baixo, se comparado com outros SOs
- ciclo de atualizações bem definido (2x ao ano)
- boas características proativas de segurança
  - W^X, ProPolice, systrace, *random lib loading order*
- filtro de pacotes pf: *stateful*, *queueing* (ALTQ), redireção de portas
- *logs* no formato `libpcap`: permite *fingerprinting* passivo

Honeypots (2)

Honeyd: emulação de serviços

- emulador de SMTP e *proxy* HTTP desenvolvidos por Niels Provos (com pequenas modificações)
- emulador de SOCKS 4/5 desenvolvido pela nossa equipe
- simula a conexão com o servidor SMTP destino e passa a receber os *e-mails*
- **não** entrega os *e-mails*

Responde a determinadas tentativas de confirmação enviadas pelos spammers

Servidor

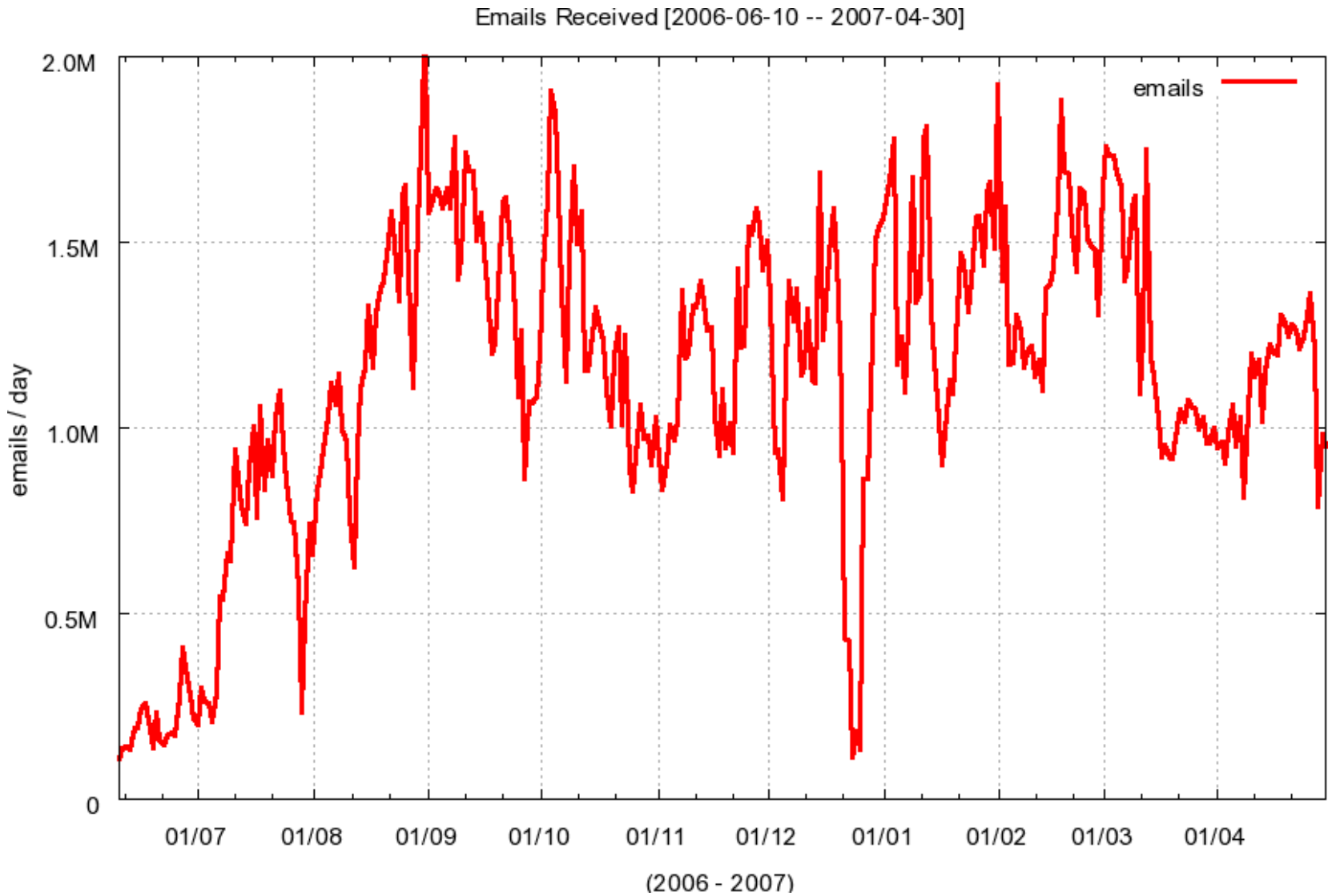
- Coleta e armazena os dados dos *honeypots*
  - inicia conexões `ssh` para a transferência de dados
  - usa `rsync` sobre `ssh` para copiar os spams coletados pelos *honeypots*
- Realiza verificações de *status* em todos *honeypots*
  - *daemons*, sincronia de relógio, espaço em disco, carga, último `rsync` bem sucedido, etc
- Fornece uma interface Web
  - *status* dos *honeypots*
  - estatísticas de *e-mails*: diária, últimos 15 minutos
  - MRTG: utilização da banda, portas usadas, *e-mails*/min, etc

Estatísticas

Estatísticas

período	10/06/2006 a 30/04/2007
dias	325
e-mails	370.263.413
destinatários	3.287.153.093
média dest./*e-mail*	$\approx$ 8,9
IPs	160.502
ASNs	2.813
CCs	157

Spams capturados / dia

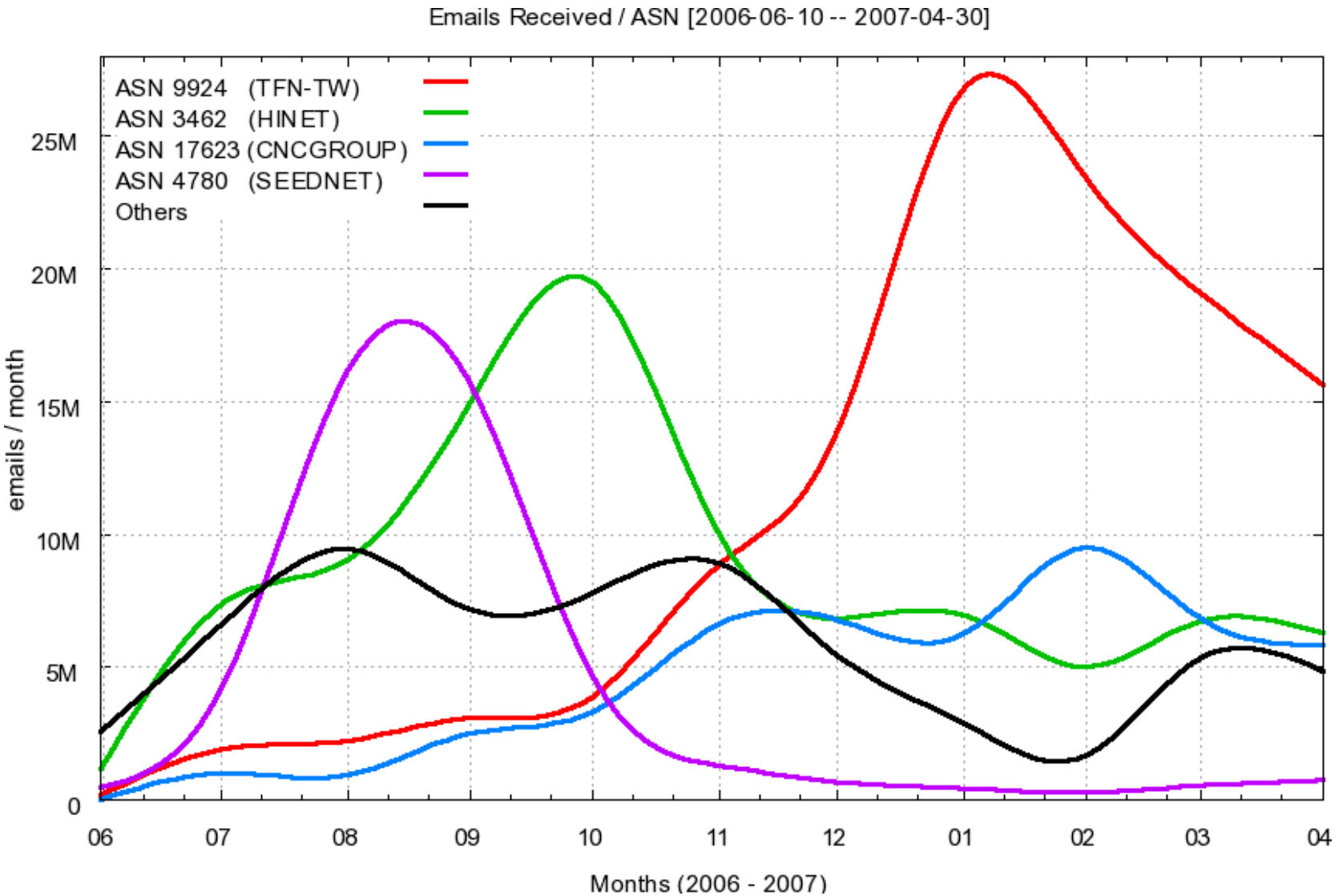


ASNs mais freqüentes

- Top10 *e-mails*/ASN:

#	ASN	Nome	%
01	9924	TFN-TW Taiwan Fixed Network	32.08
02	3462	HINET Data Communication	25.41
03	17623	CNCGROUP-SZ CNCGROUP	13.37
04	4780	SEEDNET Digital United	12.21
05	9919	NCIC-TW	02.25
06	4837	CHINA169-BACKBONE CNCGROUP	01.69
07	7271	LOOKAS - Look Communications	01.51
08	7482	APOL-AS Asia Pacific On-line	00.98
09	18182	SONET-TW Sony Network Taiwan	00.96
10	18429	EXTRALAN-TW	00.89

ASNs mais freqüentes (2)

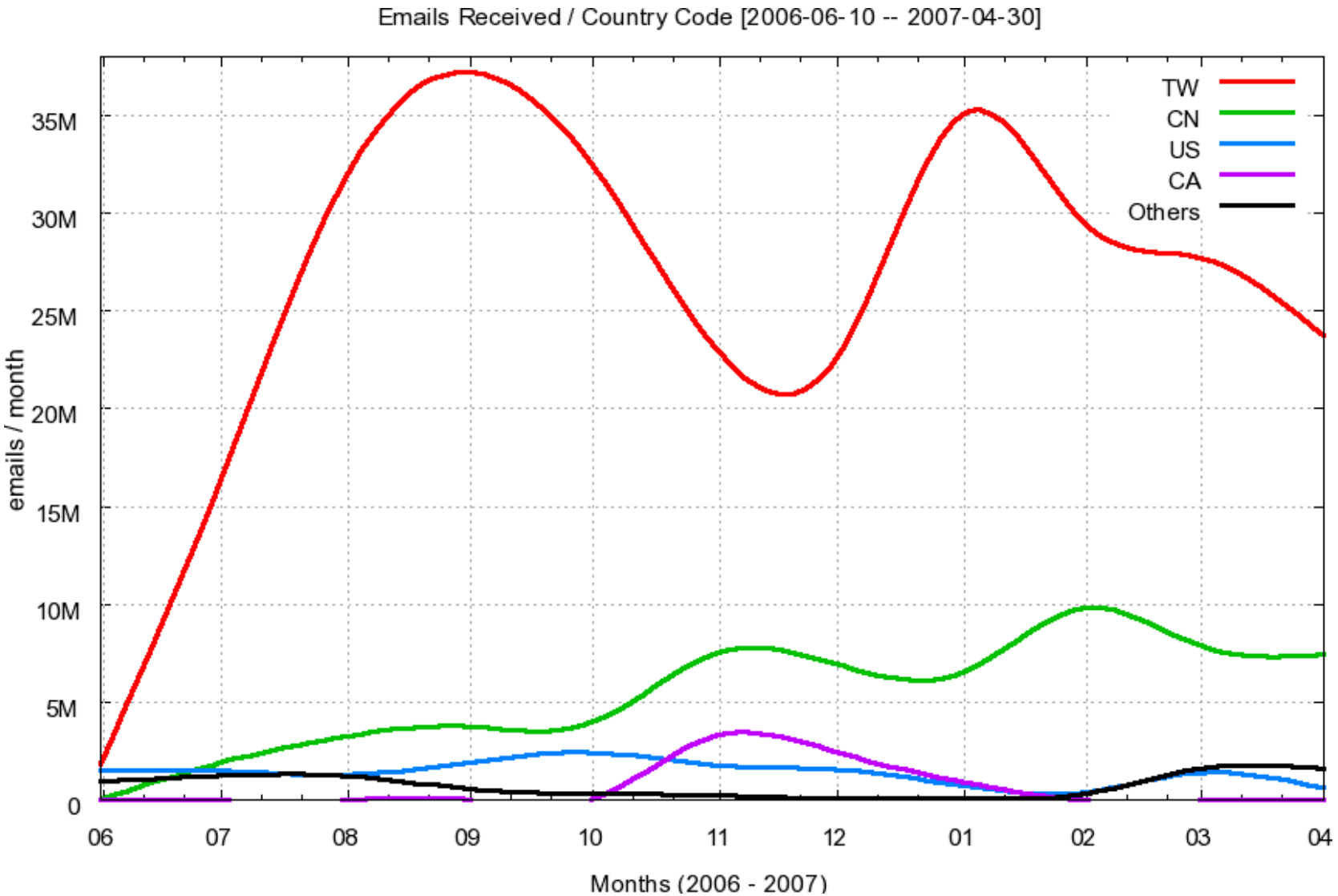


CCs mais freqüentes

- Top 10 *e-mails*/CC:

#	*e-mails*	CC	%
01	281.601.310	TW	76.05
02	58.912.303	CN	15.91
03	14.939.973	US	04.03
04	6.677.527	CA	01.80
05	1.935.648	KR	00.52
06	1.924.341	JP	00.52
07	816.072	HK	00.22
08	776.245	DE	00.21
09	642.446	BR	00.17
10	355.622	PA	00.10

CCs mais freqüentes (2)

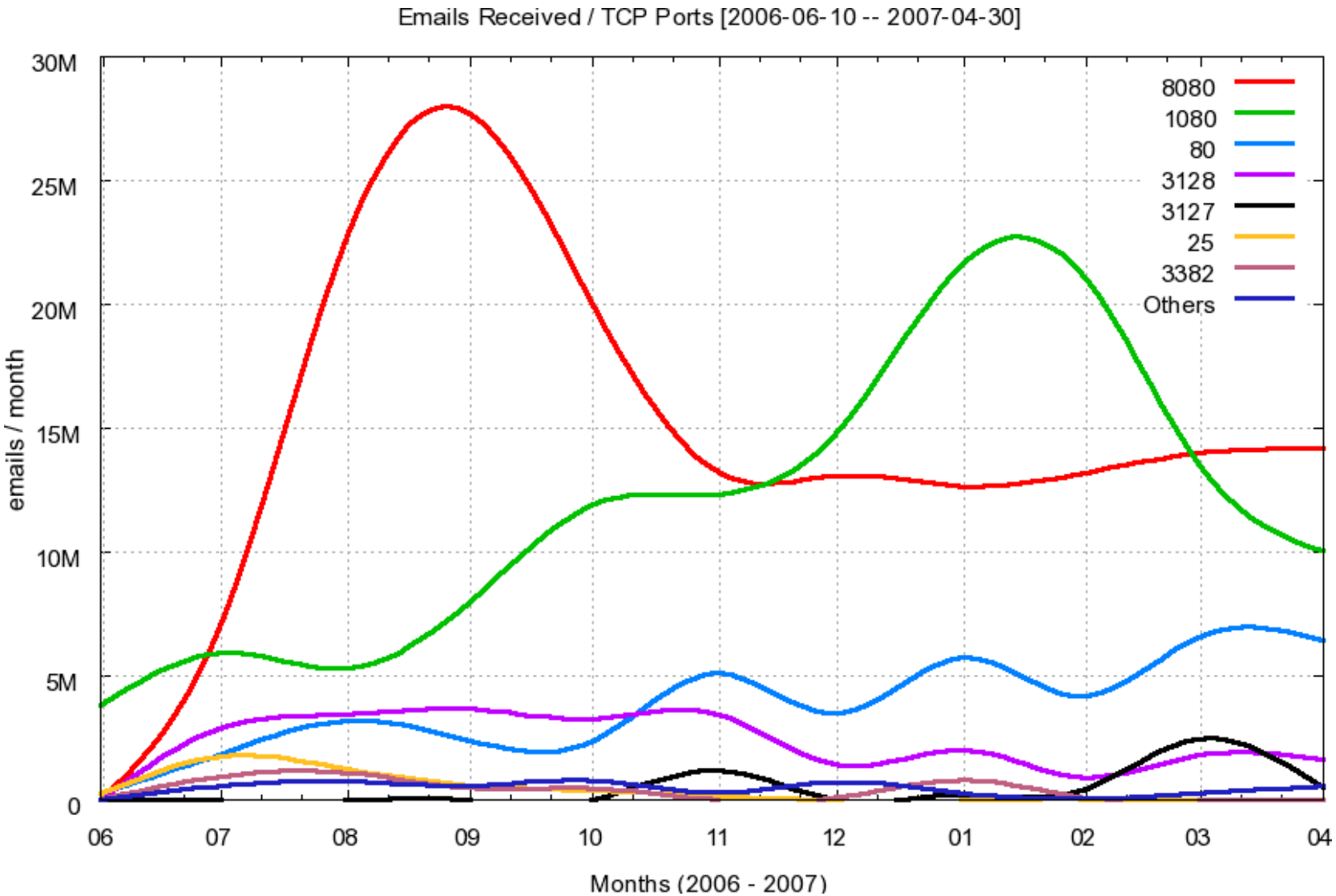


Portas TCP usadas

- *E-mails*/porta:

#	Porta TCP	protocolo	usada por	%
01	8080	HTTP	alt http	42.68
02	1080	SOCKS	socks	34.66
03	80	HTTP	http	11.22
04	3128	HTTP	Squid	06.61
05	3127	SOCKS	MyDoom	01.28
06	25	SMTP	smtp	01.18
07	3382	HTTP	Sobig.f	01.07
08	81	HTTP	alt http	00.51
09	8000	HTTP	alt http	00.37
10	6588	HTTP	AnalogX	00.27
11	4480	HTTP	Proxy+	00.15

Portas TCP usadas (2)



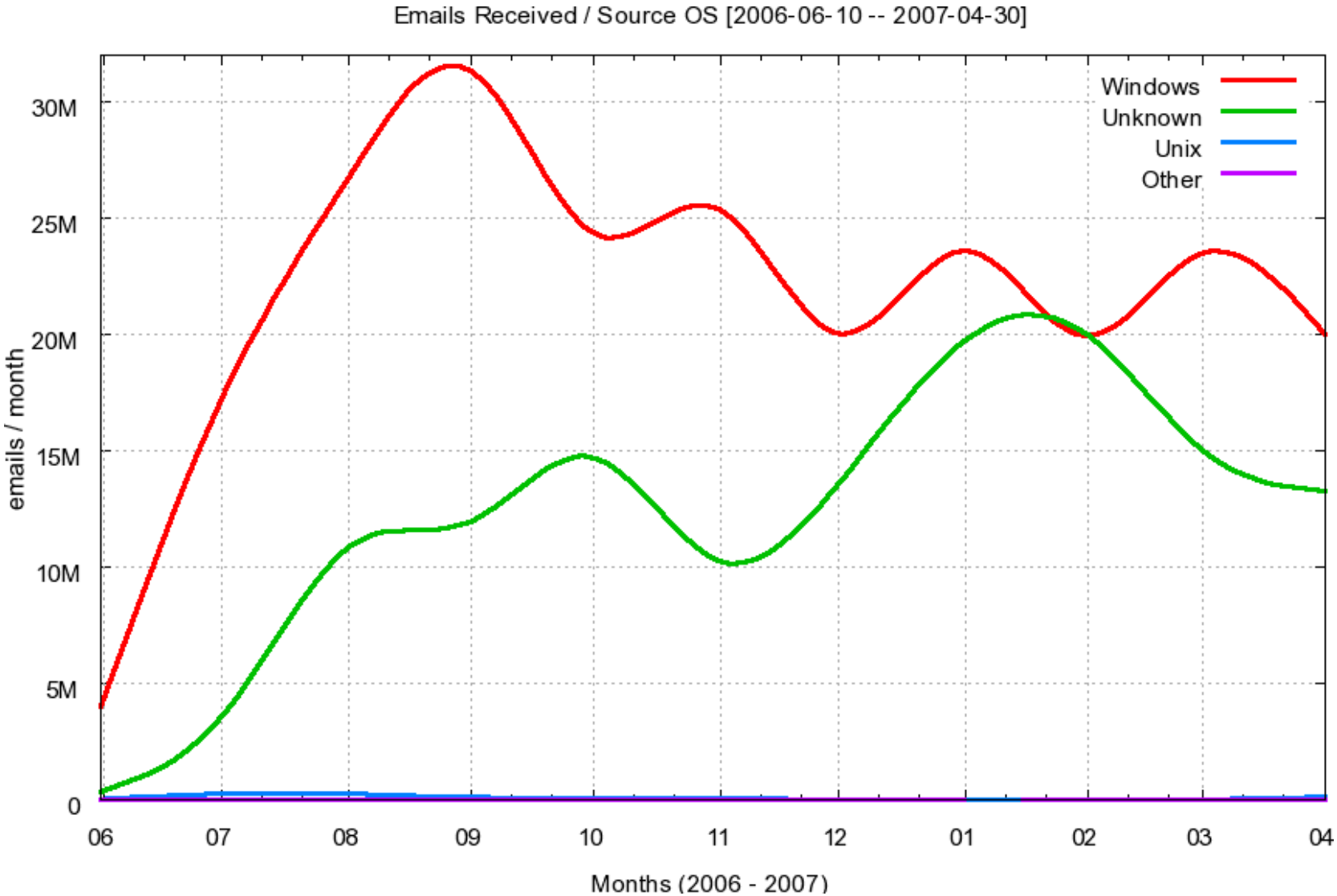
Sistemas Operacionais de Origem

- `tcpdump/pf.os` – utilizado para fazer o *fingerprinting* do SO das máquinas originando conexões TCP IPv4

#	*e-mails*	SO de origem	%
01	235.990.984	Windows	63.74
02	133.276.691	*Unknown*	36.00
03	945.642	Unix	00.26
04	50.096	Outro	00.01

http://www.openbsd.org/cgi-bin/man.cgi?query=pf.os

Sistemas Operacionais de Origem (2)



Trabalhos Futuros

Trabalhos Futuros

- Análise mais Detalhada do Problema
  - usando técnicas de *Data Mining*
  - determinação de padrões – idioma, URLs no conteúdo das mensagens, etc
  - *phishing* e outras atividades maliciosas
- Reforçar a necessidade de implementação de procedimentos e boas práticas em provedores e teles
  - gerenciamento de porta 25, SPF, DKIM, etc
  - monitoramento de abuso de *proxies*
- Cooperação internacional

Referências

- Esta apresentação pode ser encontrada em:
  http://www.cert.br/docs/palestras/
- CERT.br – Centro de Estudos, Resposta e Tratamento de Incidentes de Segurança no Brasil
  http://www.cert.br/
- NIC.br
  http://www.nic.br/
- CGI.br – Comitê Gestor da Internet no Brasil
  http://www.cgi.br/
- OpenBSD
  http://www.openbsd.org/
- Honeyd
  http://www.honeyd.org/
- Consórcio Brasileiro de Honeypots
  http://www.honeypots-alliance.org.br/