Análise de Tráfego Externo em Situações de Roteamento Parcial

Eduardo Ascenço Reis
<eduardo@intron.com.br>
<eascenco@iqaratelecom.com.br>

Agenda

- Caracterização de Troca de Tráfego Externo
- Objetivos da Análise de Troca de Tráfego Externo
- Contabilidade de Tráfego Externo – Netflow
- Análise com BGP Full Routing e Agregação de Tráfego
- Análise sem BGP Full Routing e Agregação de Tráfego
- Oregon Route Views Project
- Sistema de Análise de Tráfego Externo

Caracterização de Troca de Tráfego Externo

Sistemas Autônomos (AS)

Provedores
- Acesso
- Serviços

Instituições
- Acadêmicas
- Governamentais
- Comerciais

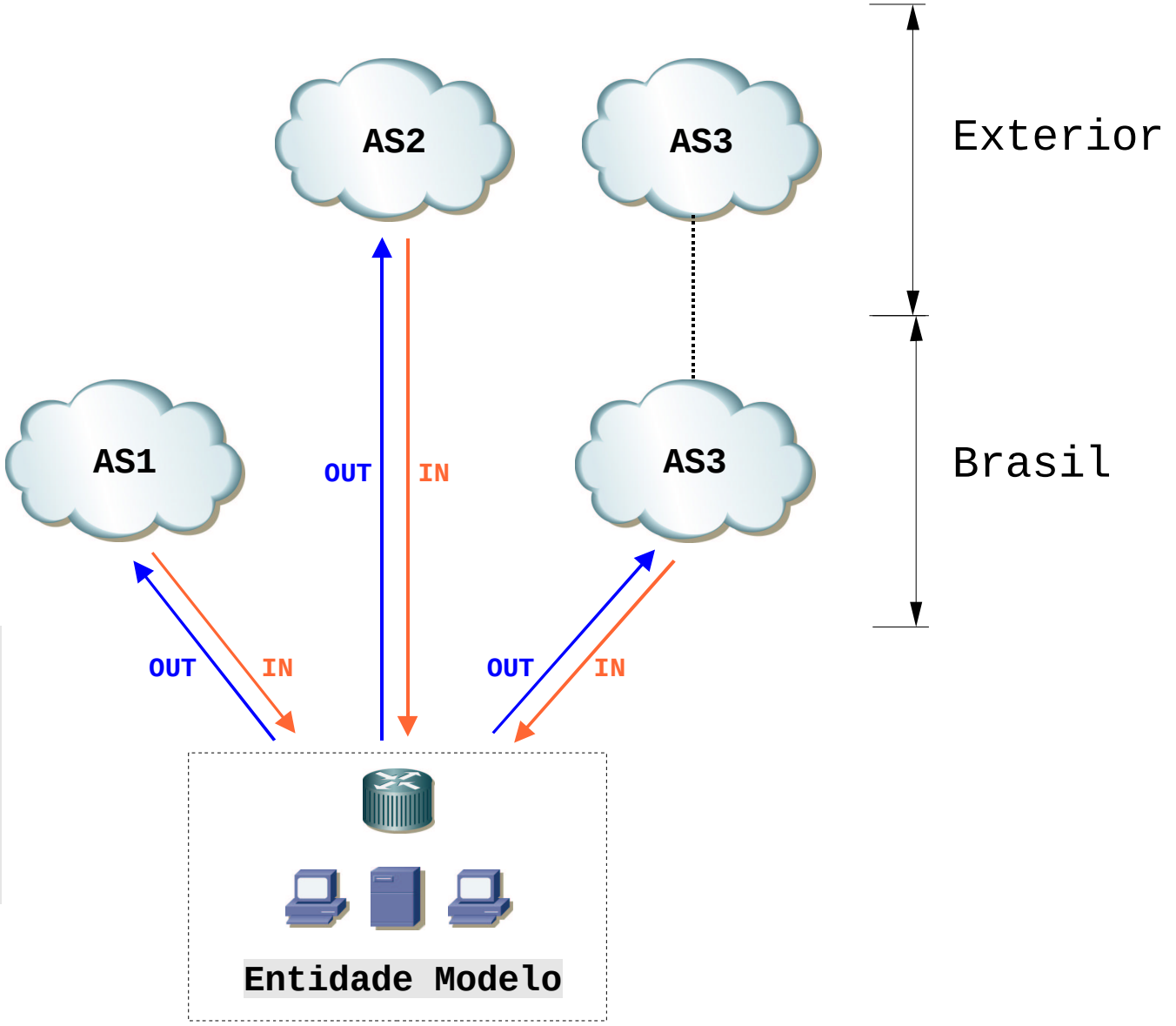
Grandes Corporações

Et al

Análises

Contabilidade
- Acumulada (GB / dia)
- Momento (Mbps)

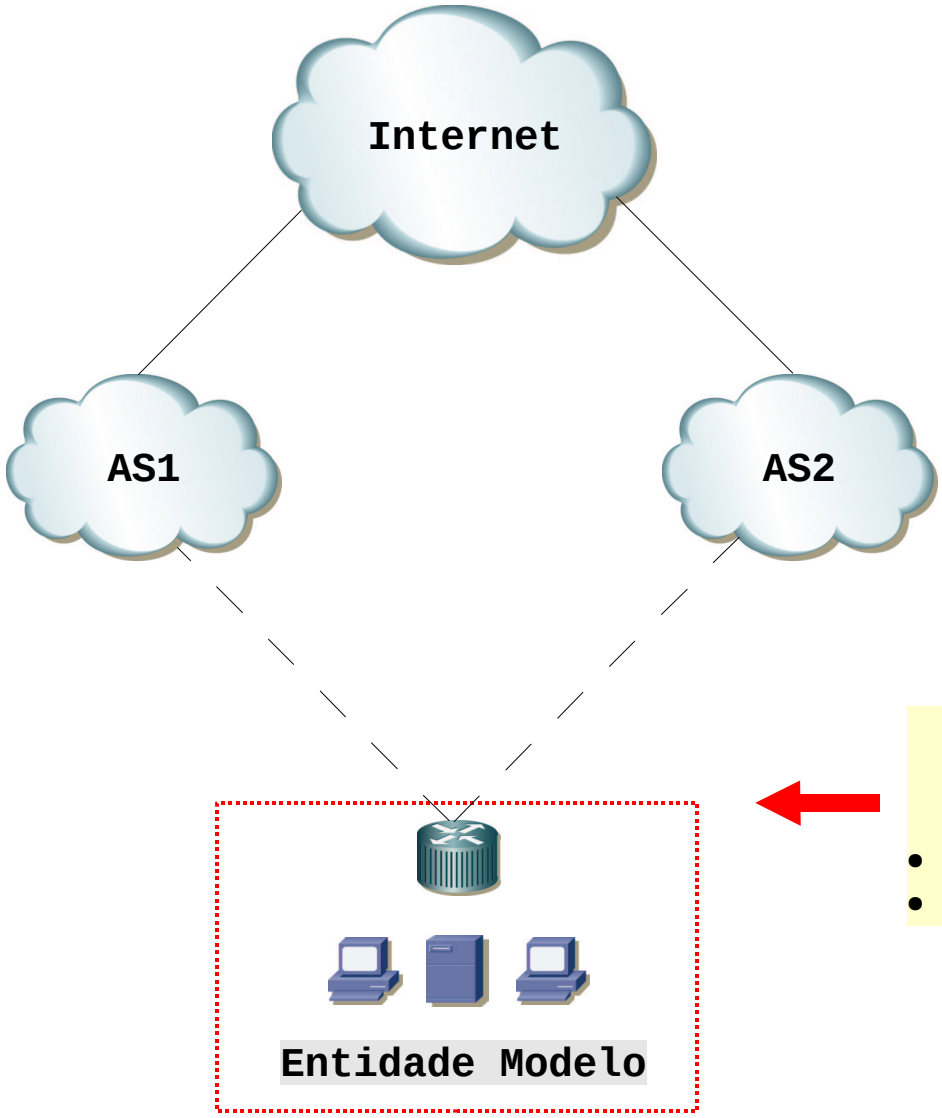
Relação OUT / IN
- Fornecedor
- Consumidor



Objetivos da Análise de Troca de Tráfego Externo

- Melhor conhecimento da rede, sistemas, serviços, etc
- Engenharia de Tráfego Externo – (Manobras de Tráfego)
- Otimização de conexão com ISP preferenciais (menor latência)
- Análise da Relação OUT/IN
  Relação de Fornecedor # Consumidor de Serviços (banda)
- Estudo de viabilidade
  - Estabelecimento de peering
  - Participação em pontos de troca de tráfego
  - Para se tornar AS – ativação de BGP
- Quantificação de utilização de serviços por determinados ISP
- Redução de custos

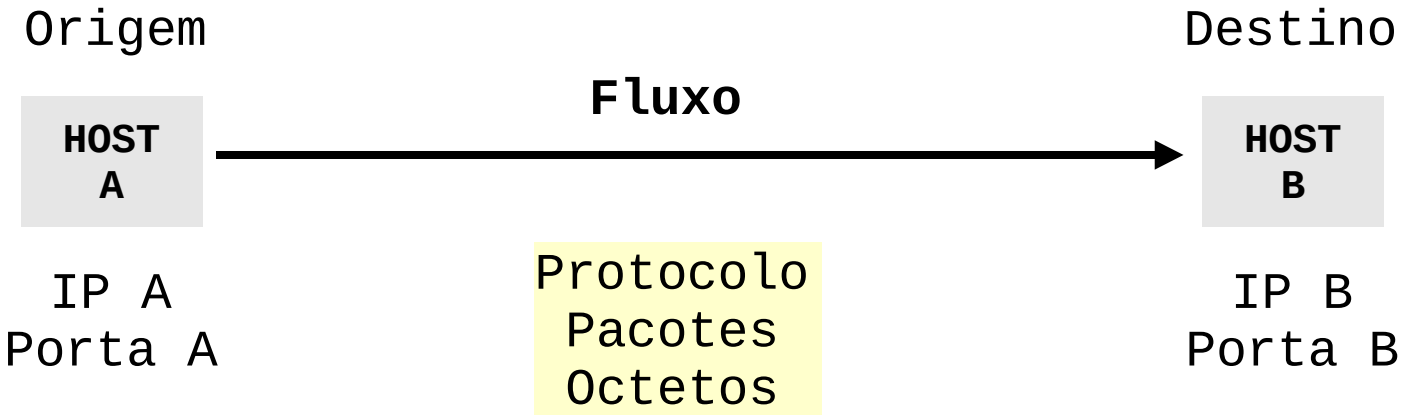
Ponto Típico de Contabilidade de Tráfego Externo (Borda)



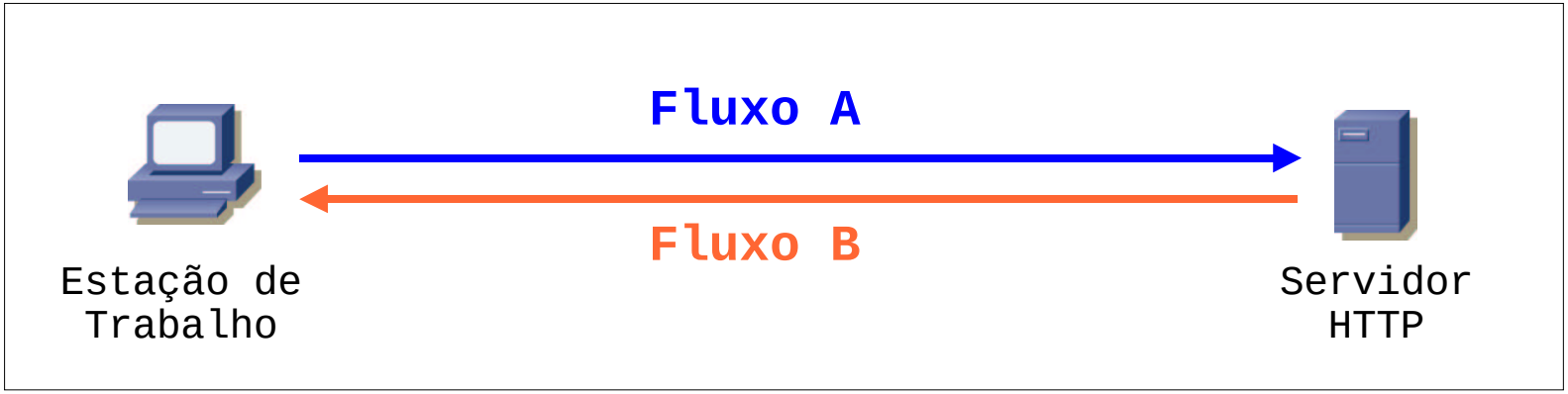
Elementos L3 de Borda de Rede

- Roteadores
- Multi-Layer Switches

Contabilidade de Tráfego Externo – Netflow – Fluxos



Exemplo de Duplo Fluxo no Acesso a um Serviço TCP



Contabilidade de Tráfego Externo − Netflow − Campos

Field	Description
srcaddr	Source IP address
dstaddr	Destination IP address
srcport	TCP/UDP source port number or equivalent
dstport	TCP/UDP destination port number or equivalent
protocol	Name or label assigned to a protocol definition in the nfknown.protocols file
protocol byte (prot)	IP protocol type (for example, TCP = 6; UDP = 17)
ToS	IP type of service
input interface	SNMP index of input interface
output interface	SNMP index of output interface
nexthop	IP address of next hop export device
src_as	Autonomous system number of the source, either origin or peer
dst_as	Autonomous system number of the destination, either origin or peer
masked srcaddr	Source IP address masked with the source netmask (src_mask)
masked dstaddr	Destination IP address masked with the destination netmask (dst_mask)
src_mask	Source IP address prefix mask bits
dst_mask	Destination IP address prefix mask bits
packet count	Packets counted as part of this record
byte count	Total number of Layer 3 bytes counted as part of this record
flow count	Total number of flows aggregated into this record
firstTimeStamp	Time, in UTC seconds, of the first packet summarized into this record
lastTimeStamp	Time, in UTC seconds, of the last packet summarized into this record
totalActiveTime	Sum of individual active time for all the flows summarized into the current record

http://www.cisco.com/en/US/products/sw/netmgtsw/ps1964/products_installation_and_configuration_guide_chapter09186a00800fea59.html#wp30537

Contabilidade de Tráfego Externo − Netflow − Config

Exemplo de Configuração Roteador Cisco

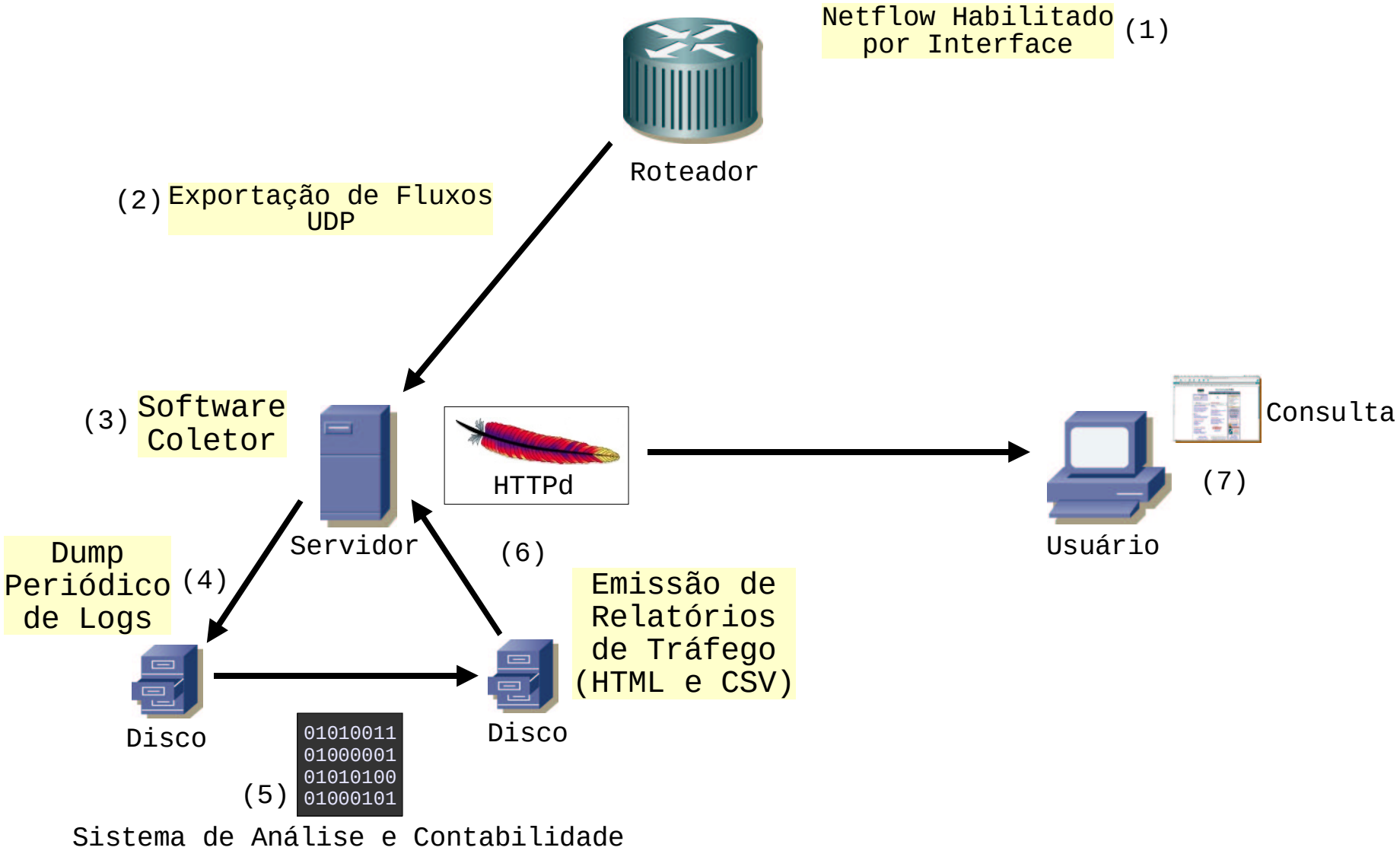
```
!
interface ATM1/0/0
ip route−cache flow
!
interface FastEthernet1/1/0
 ip route−cache flow
!
ip flow−export source Loopback0
ip flow−export version 5
ip flow−export destination 10.10.10.10 3000
!
```

*** Outros Fabricantes (e.g. Juniper e Foundry) também possuem o recurso contabilidade de tráfego com Netflow etc.

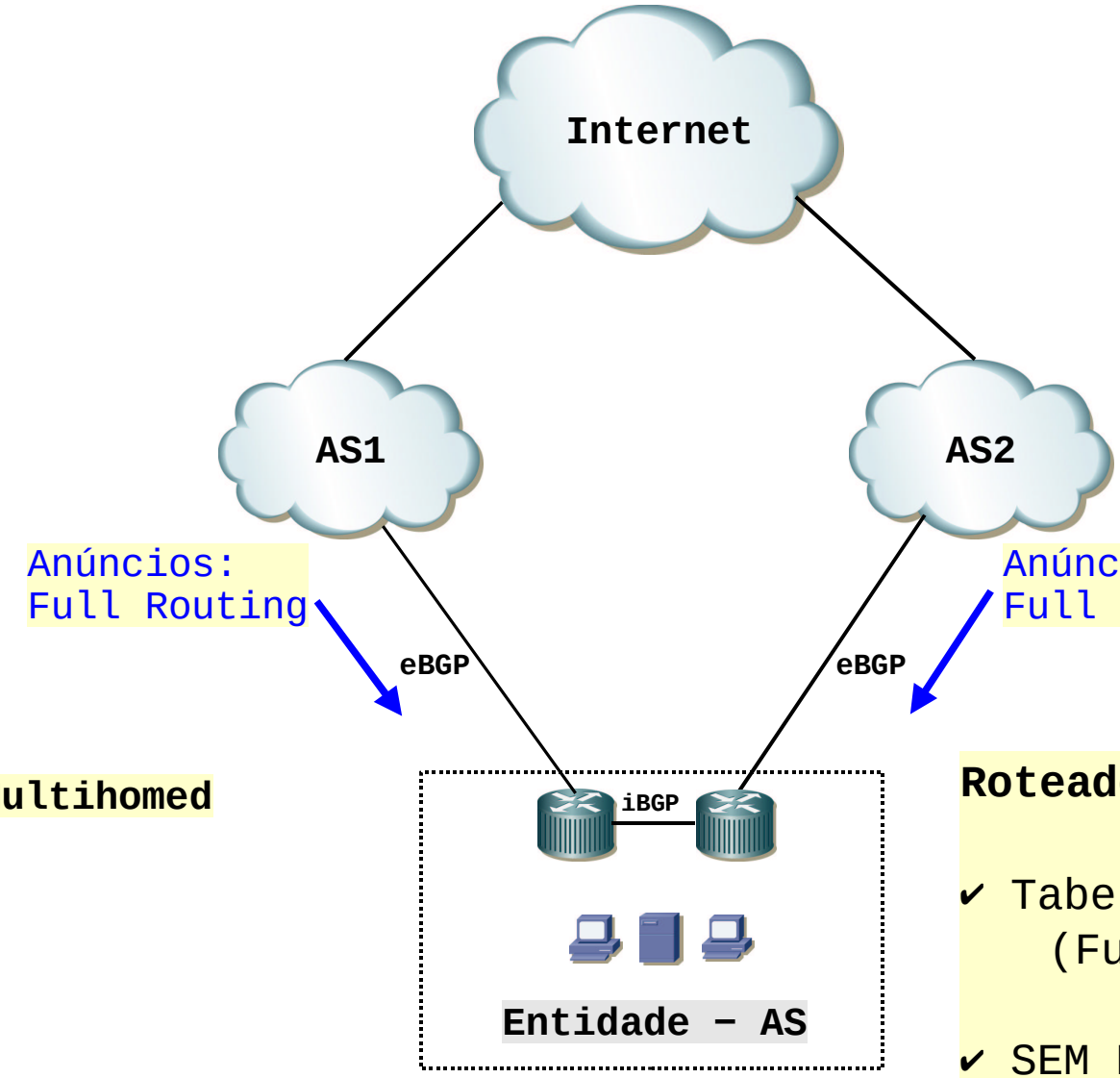
Contabilidade de Tráfego Externo − Netflow − CLI

```
Router#show ip cache flow
IP packet size distribution (99107M total packets):
   1-32   64   96  128  160  192  224  256  288  320  352  384  416  448  480
   .001 .538 .031 .024 .026 .010 .008 .005 .003 .004 .004 .004 .004 .003 .003

    512  544  576 1024 1536 2048 2560 3072 3584 4096 4608
   .003 .008 .018 .031 .263 .000 .000 .000 .000 .000 .000

IP Flow Switching Cache, 4456704 bytes
  65524 active, 12 inactive, 1835045597 added
  2994593322 ager polls, 0 flow alloc failures
  Active flows timeout in 30 minutes
  Inactive flows timeout in 15 seconds
  last clearing of statistics never
Protocol         Total    Flows   Packets Bytes  Packets Active(Sec) Idle(Sec)
--------         Flows     /Sec     /Flow  /Pkt     /Sec     /Flow     /Flow
TCP-Telnet      936823      0.2        51   767     11.1      21.9      30.6
TCP-FTP        5614171      1.3         9   410     12.3       9.8      31.6
TCP-FTPD       1460780      0.3       129   794     44.0      41.0      29.4
TCP-WWW     2238875777    521.2        11   605   6027.7       6.7      33.5

(...)

IGMP                1      0.0         2    40      0.0       2.0      18.4
IPINIP            222      0.0         1    49      0.0       0.4      38.5
GRE             27299      0.0       205    47      1.3     849.5      13.8
IP-other       400321      0.0      1002   146     93.4     208.1      28.3
Total:     10424914665   2427.2         9   469  23074.4      13.4      32.2

SrcIf         SrcIPaddress    DstIf         DstIPaddress    Pr SrcP DstP  Pkts
AT1/0/0.1     143.107.254.11  Fa1/1/0       200.218.224.249 06 0050 E28F     1
Fa1/1/0       200.218.224.249 AT4/0/0.1     143.107.254.11  06 E28F 0050     2

(...)
```

Contabilidade de Tráfego Externo – Netflow – Modelo de Sistema



Situação com BGP Full Routing – ISP Tradicional (AS)



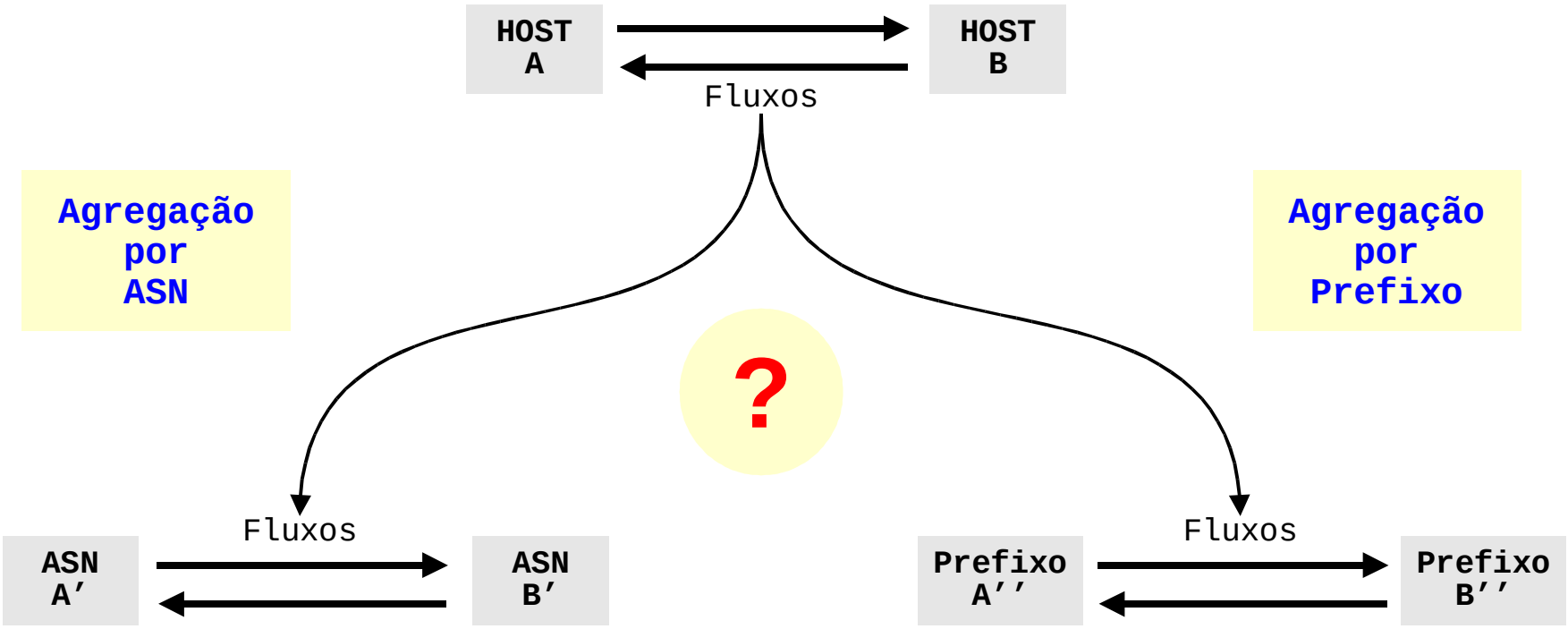
eBGP Multihomed

Roteadores

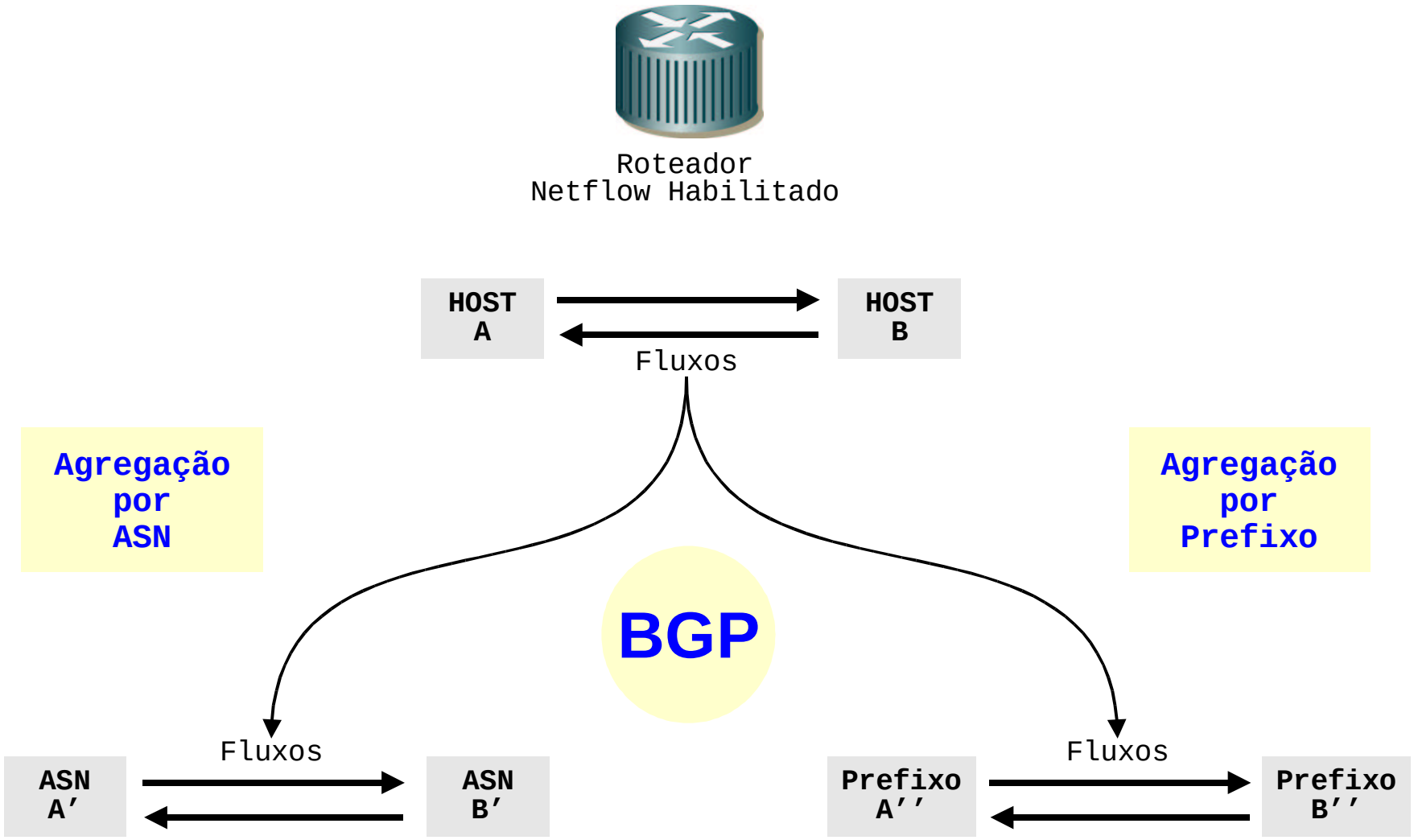
- ✔ Tabela BGP Completa (Full Routing)
- ✔ SEM Rota Default Aponta para Null0

Situação com BGP Full Routing – Netflow – Agregação (1/3)

Roteador
Netflow Habilitado



Situação com BGP Full Routing – Netflow – Agregação (2/3)



Situação com BGP Full Routing − Netflow − Agregação (3/3)

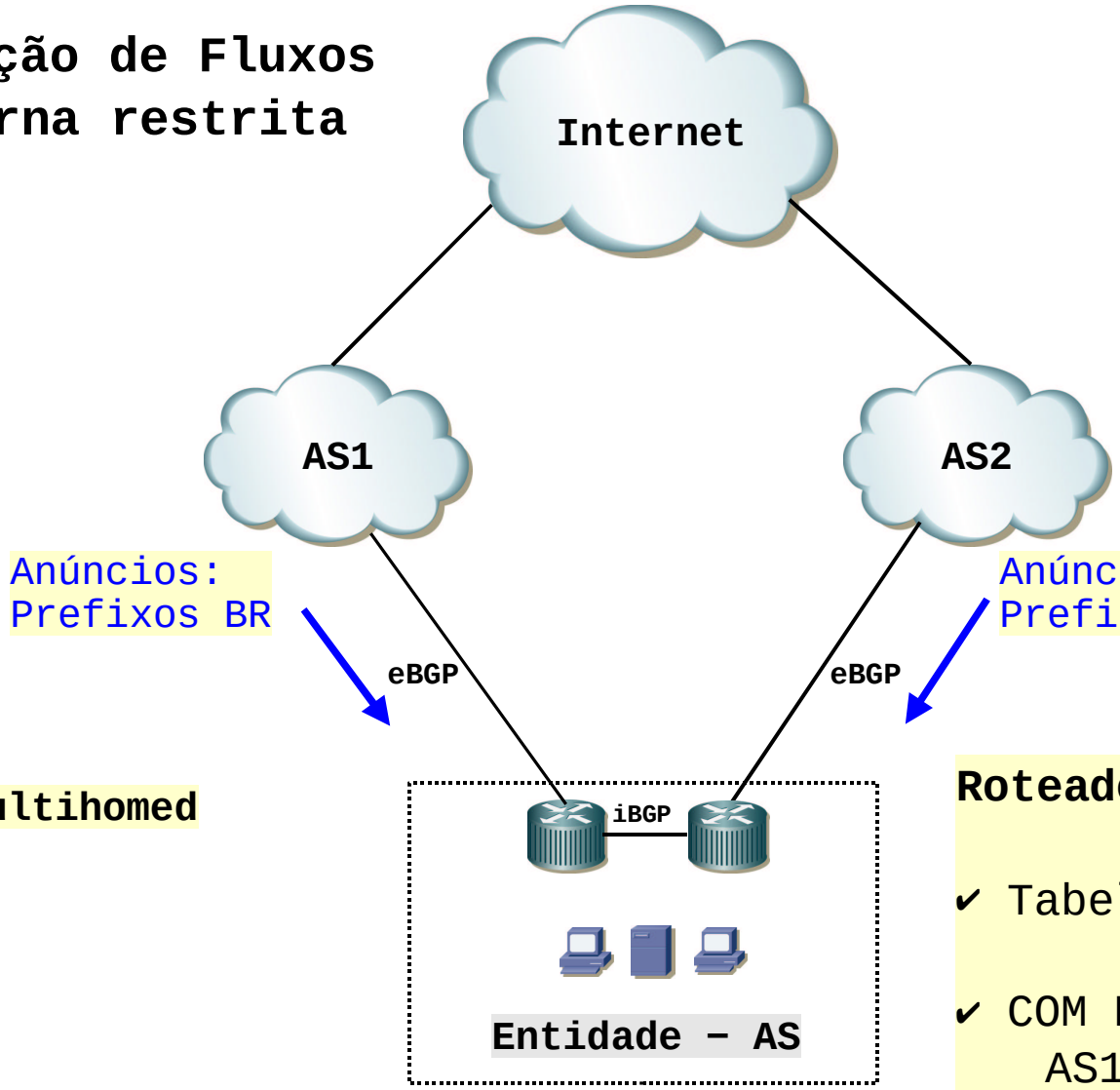
O processo de agregação de tráfego por prefixos ou ASN depende da tabela BGP completa como base de informaçoes.

Exemplo de configuração em Cisco

```
!
ip flow-aggregation cache as|prefix
 export destination 10.10.10.10 9999
 enabled
!
```

Situação de BGP com Tabela Parcial & Rota Default

✗ **Agregação de Fluxos se torna restrita**



eBGP Multihomed

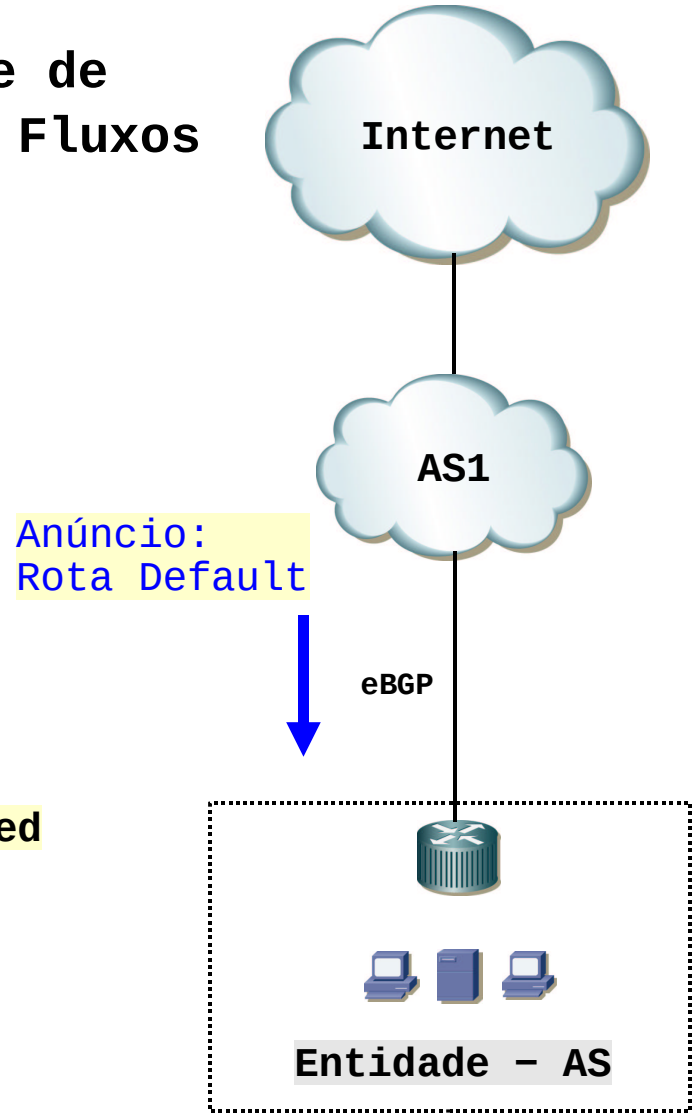
Roteadores

✔ Tabela BGP Parcial

✔ COM Rota Default para AS1 e AS2 com pesos diferentes

Situação de BGP com Tabela Local & Rota Default

✗ **SEM Capacidade de Agregação de Fluxos**



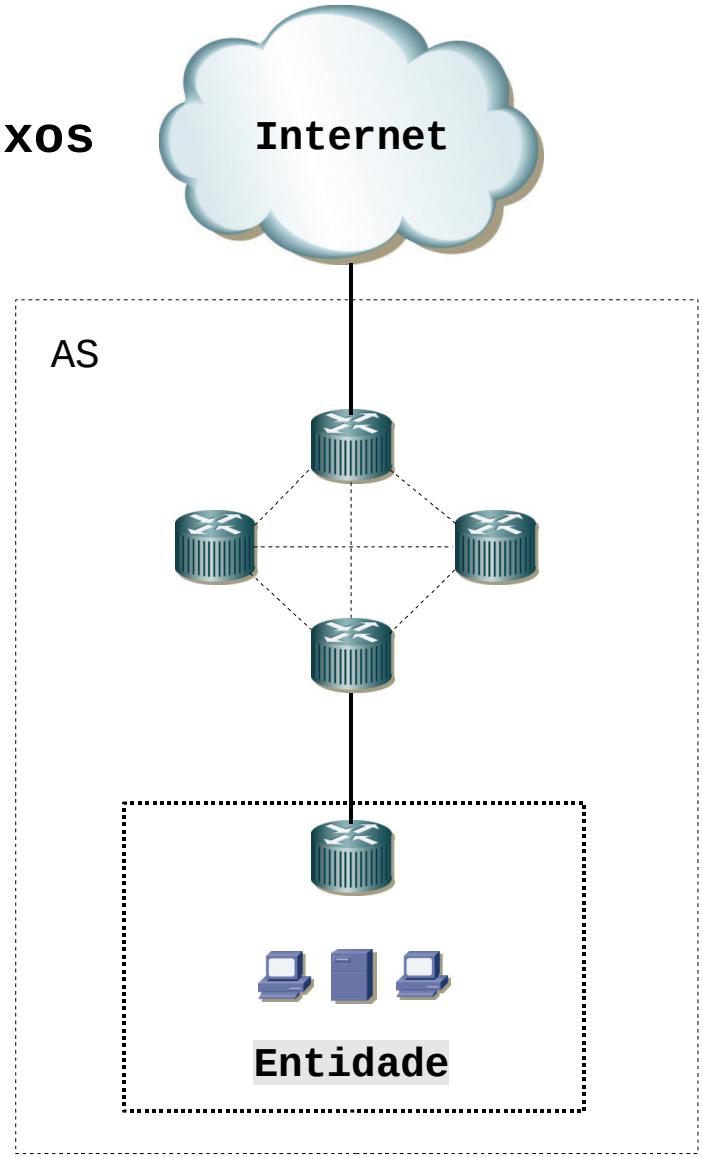
Roteador

✔ Tabela BGP Local

✔ COM Rota Default para AS1

Situação SEM BGP – Apenas Rota Default

✗ **SEM Capacidade de Agregação de Fluxos**



Entidade faz parte de um AS que lhe provê acesso

Exemplos de AS:
- Provedor Acesso
- Data Center
- Campus Universitário

Roteador ou Switch L3

✔ COM Rota Default

Análise de Tráfego Externo SEM o Recurso de Agregação

Como realizar a análise de tráfego externo em situações nas quais a capacidade de agregação de fluxos se torna restrita ou aparentemente não possível nos elementos de rede (roteadores e switches L3) ?

Alternativa:

Uso de base externa com informações da tabela completa BGP !

Route Views Project – Oregon – Base Externa BGP (1/2)

University of Oregon
Route Views Project

http://www.route-views.org/
http://www.routeviews.org/

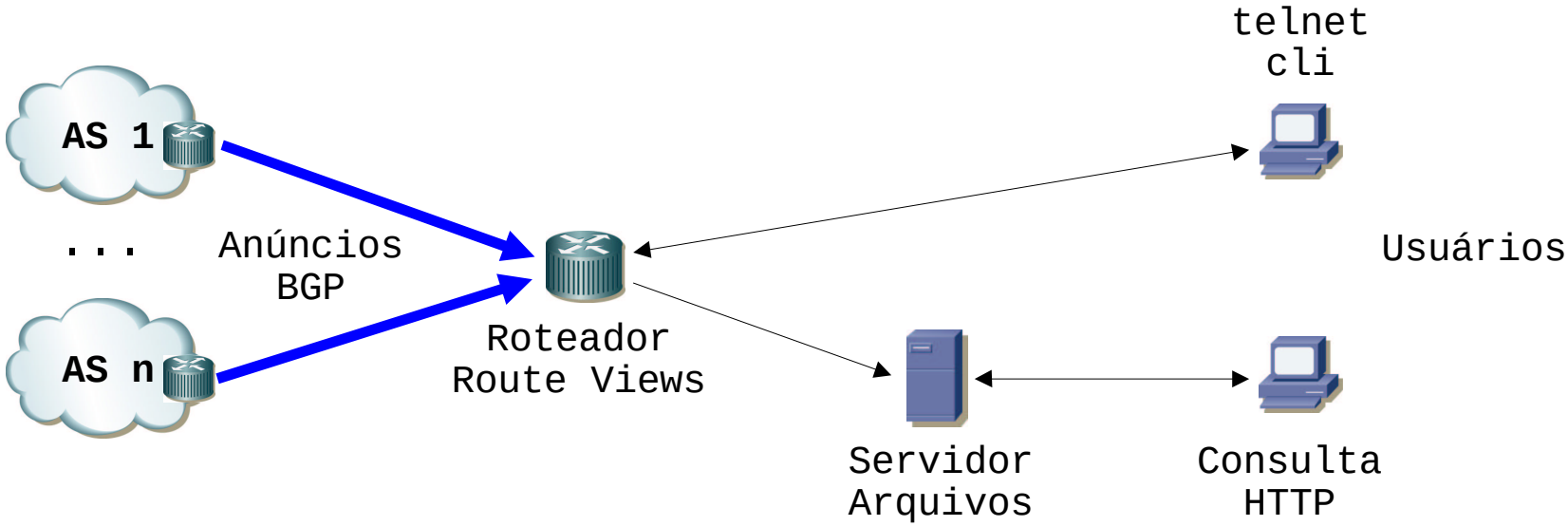
Projeto provê acesso em tempo real ao sistema de roteamento global pela visão de diferentes operadoras / backbones espalhados pela Internet.

Os participantes estabelecem sessões eBGP multi-hop com os roteadores do projeto e anunciam todos os seus prefixos conhecidos (Full Routing).

Os roteadores do projeto não anúnciam nenhum prefixo, dessa forma os anúncios recebidos não são repassados e nem utilizados (forwarding).

Participantes: 58 no roteador principal
Sprint, UUNET/MCI, Level3, SAVVIS, ATT, Telia, XO, Telefonica, Global Crossing, France Telecom, LINX, *et al*

Route Views Project – Oregon – Base Externa BGP (2/2)



Route Views Archive Project
http://archive.routeviews.org/

- MRT format RIBs (zebra)
- 'sh ip bgp' format RIBs
- UPDATES e dampening data

SATE – Sistema de Análise de Tráfego Externo

SATEparc

Sistema de Análise de Tráfego Externo
em Situações de Roteamento Parcial

Recursos

Hardware:

- PC Desktop Intel Pentium 4 2GHz
  512MB RAM – 30GB disco

Software:

- Debian GNU/Linux 3.1 (sarge)
- Flow-tools
- Bind9
- Apache
- Perl Scripts

SATEparc – Modelo do Sistema

Roteador
com Netflow

Usuário

Fluxos
(udp)

Consulta
(http)

SATE

Coletor

flow-capture

- ✔ recebimento dos fluxos
- ✔ dump fluxos em disco (arquivos)

Analisador

flow-cat|flow-print
bind9
Perl Scripts

- ✔ montagem da base local de mapeamento Prefixo<=>ASNori<=>BR
- ✔ análise fluxos (parse)
- ✔ emite logs análise

Emissor de Relatórios

Apache
Perl Scripts

- ✔ análise logs (parse)
- ✔ emite relatórios consolidados em HTML e CSV
- ✔ Interface de consulta

SATEparc – Base Local de Mapeamento de Prefixos (1/3)

oix-full-snapshot-latest
'sh ip bgp' format RIBs
[1]

Fonte Oficial de Identificação
de blocos CIDR BR
|lacnic|BR|ipv4|
[2]

Tabela BGP Completa IPv4

Lista Blocos CIDR BR
Alocados

Zona DNS Local
Base de Dados de Mapeamento de Prefixos
Modelo do Projeto Route Views [3]

[1] http://archive.routeviews.org/oix-route-views/2005.07/oix-full-snapshot-latest.dat.bz2
[2] ftp://ftp.lacnic.net/pub/stats/lacnic/delegated-lacnic-latest
[3] http://archive.routeviews.org/dnszones/

SATEparc – Base Local de Mapeamento de Prefixos (2/3)

Mecanismo de Mapeamento

Endereço IP →

- Prefixo
- ASNorigem
- BR|NoBR

Zona de DNS local (rviews-local)
Registros TXT

Exemplo de Consulta

```
$ host registro.br
registro.br has address 200.160.2.3
$
$ host -t TXT 3.2.160.200.rviews-local
3.2.160.200.rviews-local text "22548" "200.160.0.0" "20" "BR"
$
$ host www.route-views.org
www.route-views.org has address 128.223.61.18
$
$ host -t TXT 18.61.223.128.rviews-local
18.61.223.128.rviews-local text "3582" "128.223.0.0" "16" "NOBR"
$
```

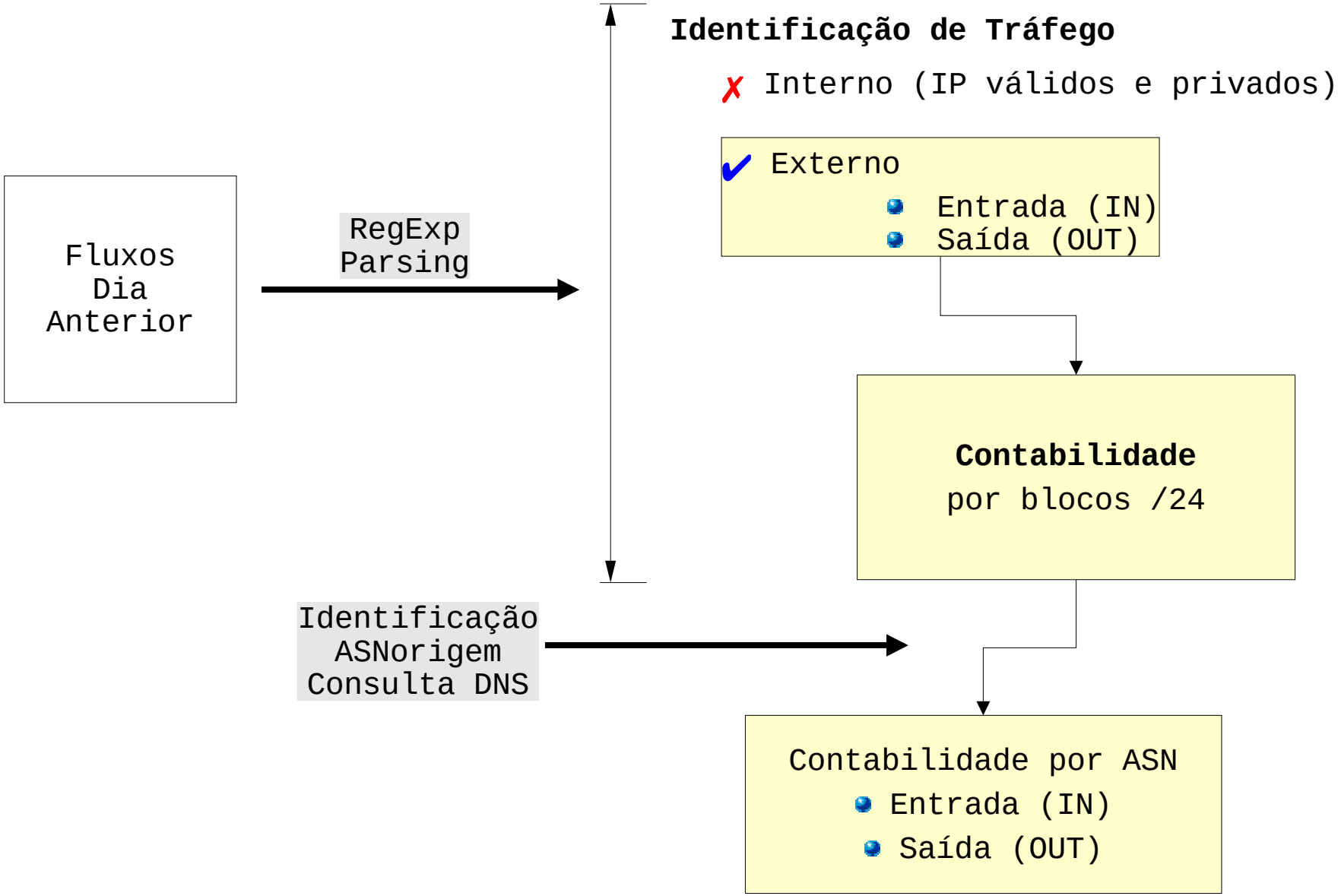
SATEparc – Base Local de Mapeamento de Prefixos (3/3)

Exemplo de Construção da Zona DNS

```
Prefixo|ASNorigem|BR/NoBR
200.160.0.0/20|22548|BR
```

```
;
0.160.200                IN TXT  "22548" "200.160.0.0" "20" "BR"
*.0.160.200              IN TXT  "22548" "200.160.0.0" "20" "BR"
;
1.160.200                IN TXT  "22548" "200.160.0.0" "20" "BR"
*.1.160.200              IN TXT  "22548" "200.160.0.0" "20" "BR"
;

(...)

;
14.160.200               IN TXT  "22548" "200.160.0.0" "20" "BR"
*.14.160.200             IN TXT  "22548" "200.160.0.0" "20" "BR"
;
15.160.200               IN TXT  "22548" "200.160.0.0" "20" "BR"
*.15.160.200             IN TXT  "22548" "200.160.0.0" "20" "BR"
;
16.160.200               IN TXT  "22148" "200.160.16.0" "20" "BR"
*.16.160.200             IN TXT  "22148" "200.160.16.0" "20" "BR"
;
```

SATEparc – Resumo do Processo de Análise e Contabilidade



SATEparc – Saída do Sistema – Relatórios

Relatórios: Diário, Mensal e Anual

Relação Tráfego Brasileiro # Internacional

- Total
- Entrada (IN)
- Saída (OUT)

TOP20 ASs
(por Tráfego Total)

Global
(Todos AS)

ASN: aut-num
Nome: owner

→

Tráfego Total
(IN+OUT)
GB e %

Tráfego Entrada
(IN)
GB e %

Tráfego Saída
(OUT)
GB e %

Relação
OUT/IN

SATEparc – Alguns Números (aproximados)

Oregon Route Views

oix-full-snapshot-latest.dat

680 MB

8 M linhas

Zona DNS de Mapeamento de Prefixos

db.rviews-local

150 MB

2.8 M registros

Exemplo de Caso (site)

Links Externos: 2x E3 ATM

Média: 22 Mbps [*]

Pico: 34 Mbps [*]

* Considerando valores mensais e o maior valor entre IN e OUT

Fluxos: 85 M dia

Total (IN + OUT): 330 GB

Tempo de Execução da Análise (dia): 3h

Agradecimentos

Iqara Telecom
Divisão de Engenharia e Tecnologia

Oregon Route Views Project
David Meyer, John Heasley,
Joel Jaeggli & Project fellows

LacNic
Ricardo Patara

http://www.intron.com.br/doc/gter19.sate.parc-route.ear.20050704.pdf