

Construindo um Route Collector no PTT

PTT-Metro de Curitiba/PR

LACNIC XII - NAPLA 2009 - Cidade do Panamá - Panamá
26 de Maio 2009

Introdução

Ambiente de um PTT

Motivação

Coleta de dados

Coleta e armazenamento de informações

Relatórios

Geração de Relatórios

Conclusão

Pedro R. Torres Jr. `torres@pop-pr.rnp.br`
PTT-Metro de Curitiba

Agenda

- 1 Introdução
- 2 Coleta de dados
- 3 Relatórios
- 4 Conclusão

Introdução

Ambiente de um PTT

Motivação

Coleta de dados

Coleta e armazenamento
de informações

Relatórios

Geração de Relatórios

Conclusão

Equipamentos

- Switch Ethernet: Vlans
- Route Servers: BGP Speakers
- Gerência: Looking Glass, Estatísticas, Monitores, etc.

[Introdução](#)

[Ambiente de um PTT](#)

[Motivação](#)

[Coleta de dados](#)

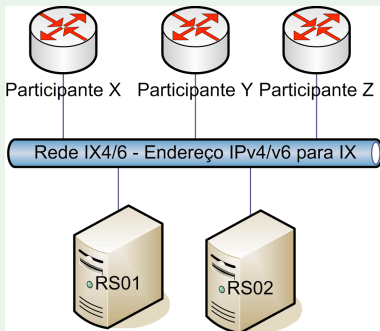
[Coleta e armazenamento
de informações](#)

[Relatórios](#)

[Geração de Relatórios](#)

[Conclusão](#)

Topologia Simplificada

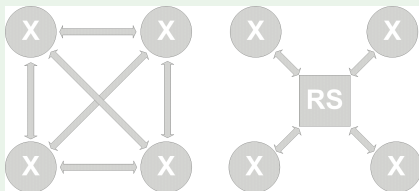


Cenário Padrão em um PTT-Metro

Políticas que podem ser aceitas

- Acordo de Troca de Tráfego Multilateral - ATM
- Acordos Bilaterais
- Compra/Venda de Trânsito

Cenário Padrão em um PTT-Metro



- ATM: Uso dos Route Servers para estabelecer as sessões BGP
- Route Server é um AS/Next-Hop transparente
- Comportamento de full-mesh de sessões BGP entre os participantes do ATM
- Políticas Bilaterais requerem: $n * (n - 1) / 2$ sessões!

Route-Server - RS

- Hub de sessões BGP
- Concentrador de informações de roteamento

Ambiente

- Participantes adotam políticas de roteamento exclusivas no PTT
- Administrador de peering conhece bem o contato de cada prefixo aprendido no RS
- Excelente local para monitorar!

[Introdução](#)

[Ambiente de um PTT](#)

[Motivação](#)

[Coleta de dados](#)

[Coleta e armazenamento de informações](#)

[Relatórios](#)

[Geração de Relatórios](#)

[Conclusão](#)

O que coletar ou monitorar?

- Mensagens de controle do BGP
- Updates e mais updates!
- RIBs

Introdução

Ambiente de um PTT

Motivação

Coleta de dados

Coleta e armazenamento de informações

Relatórios

Geração de Relatórios

Conclusão

RS implementado com Quagga Software Suite

- dump de mensagens BGP
- dump de updates recebidos
 - Código modificado para garantir a coleta de todos os updates!
- dump da tabela de roteamento

Formato MRT

- Routing information export format: MRT
- draft-ietf-grow-mrt-09
- Suporta informações IPv6 e AS4

Introdução

Ambiente de um PTT

Motivação

Coleta de dados

Coleta e armazenamento
de informações

Relatórios

Geração de Relatórios

Conclusão

Storage

- dumps são escritos para o sistema de arquivos do RS
- cópia dos dumps é feita para o Route Collector (storage)
 - Temos todos os dados desde 01/2008
- Trabalho semelhante ao do RouteViews

Introdução

Ambiente de um PTT

Motivação

Coleta de dados

Coleta e armazenamento
de informações

Relatórios

Geração de Relatórios

Conclusão

Possibilidades de uso

- Histórico
- Facilmente integrável com ferramentas como BGPlay
- Possibilidade de explorar as informações

Ferramentas

- Route Collector armazena todos os dumps originais
- Dumps MRT são convertidos para formato SQL
 - Criamos os scripts: `mrt2sql`
- Importamos todos os dumps para o banco de dados

Exemplo de consulta SQL:

- ```
select count(1) from routes where peerasn = 10881;
```

### Introdução

Ambiente de um PTT

Motivação

### Coleta de dados

Coleta e armazenamento de informações

### Relatórios

Geração de Relatórios

### Conclusão

# Relatórios Gerados (até o momento...)

## Updates

### Diversas estatísticas

- Instabilidade
- Maiores ASPATH
- Maiores Prependes

### Introdução

Ambiente de um PTT

Motivação

### Coleta de dados

Coleta e armazenamento  
de informações

## Routes

- Tamanho da RIB
- Número de prefixos por ASN
- Distribuição de prefixos IPv4 e IPv6
- Proposta de agregação
- Prefixos mais específicos
- Maiores ASPATH e Prependes

### Relatórios

Geração de Relatórios

### Conclusão

Emitidos diariamente, com dados do dia anterior!

## RFC2439:

*...BGP implementation may tend to fail due to high routing update volume*

## Instabilidade

- Cada mensagem de update de um prefixo é considerado uma instabilidade
- Relatório separado para mensagens de update do tipo announce e withdraw
  - Descobrimos que recebemos mensagens de withdraw de um full-routing!
  - Possivelmente uma implementação BGP com defeito

### Introdução

Ambiente de um PTT

Motivação

### Coleta de dados

Coleta e armazenamento de informações

### Relatórios

Geração de Relatórios

### Conclusão

## Exemplo de Consulta Diária

```
peerasn | msg | nlri | msgpernlri | instabilidade
65001 | 396 | 54 | 7.0 | 64.18
65002 | 112 | 5 | 22.0 | 18.15
65003 | 56 | 26 | 2.0 | 9.07
65004 | 53 | 53 | 1.0 | 8.58
...
```

## Resultados

- Já descobrimos ASN verbosos por erro de configuração
- Já descobrimos ASN danificados em provedores internacionais
- Já descobrimos problemas com implementações do BGP

### Introdução

Ambiente de um PTT

Motivação

### Coleta de dados

Coleta e armazenamento de informações

### Relatórios

Geração de Relatórios

### Conclusão

## Exemplo de Consulta Diária

| nlri            |  | peerasn |  | aspath            |  | msg |  | instab. |
|-----------------|--|---------|--|-------------------|--|-----|--|---------|
| 192.168.10.0/21 |  | 65001   |  | 65001 64999 64999 |  | 420 |  | 67.12   |
| 172.16.120.0/19 |  | 65002   |  | 65002 65003       |  | 15  |  | 7.12    |
| 10.120.0.0/20   |  | 65003   |  | 65003             |  | 10  |  | 5.26    |
| 201.55.128.0/21 |  | 65004   |  | 65004             |  | 10  |  | 5.26    |
| ...             |  |         |  |                   |  |     |  |         |

## Resultados

- Já descobrimos redes em router BGP instável

### Introdução

Ambiente de um PTT

Motivação

### Coleta de dados

Coleta e armazenamento de informações

### Relatórios

Geração de Relatórios

### Conclusão

## Prefixos exportados por cada ASN

```
peerasn | prefix
65001 | 1274
65002 | 384
65003 | 346
65004 | 223
65005 | 23
65006 | 15
65007 | 10
65008 | 9
65009 | 2
```

...

## Resultados

- Acompanhamos o crescimento da RIB
- Detectamos crescimentos súbitos

### Introdução

Ambiente de um PTT

Motivação

### Coleta de dados

Coleta e armazenamento  
de informações

### Relatórios

Geração de Relatórios

### Conclusão

## Distribuição de prefixos IPv4

| mask | totalprefix |
|------|-------------|
| 15   | 4           |
| 16   | 50          |
| 17   | 36          |
| 18   | 201         |
| 19   | 93          |
| 20   | 372         |
| 21   | 252         |
| 22   | 210         |
| 23   | 291         |
| 24   | 801         |
| 25   | 6           |
| 26   | 7           |
| 31   | 1           |

## Distribuição de prefixos IPv6

| mask | totalprefix |
|------|-------------|
| 32   | 3           |
| 48   | 1           |
| 64   | 2           |

## Resultados

- Uso comum de /24
- Participante que só anunciam /24 em uso
- Administradores preferem máscara par :)

### Introdução

Ambiente de um PTT

Motivação

### Coleta de dados

Coleta e armazenamento  
de informações

### Relatórios

Geração de Relatórios

### Conclusão

### Prefixos IPv4 mais específicos que /24

```
peerasn| prefix
```

|       |  |    |
|-------|--|----|
| 65001 |  | 11 |
| 65002 |  | 2  |
| 65003 |  | 1  |

### Prefixos IPv6 mais específicos que /48

```
peerasn| prefix
```

|       |  |   |
|-------|--|---|
| 65001 |  | 2 |
|-------|--|---|

## Resultados

- Não filtramos prefixos mais específicos que /24
- É melhor ter este tráfego no PTT :)
- Algum participante pode estar filtrando!

### Introdução

Ambiente de um PTT

Motivação

### Coleta de dados

Coleta e armazenamento de informações

### Relatórios

Geração de Relatórios

### Conclusão

### Ganho obtido com a agregação

| asn   | total | aggr | gain | gainper100 |
|-------|-------|------|------|------------|
| 65001 | 1274  | 424  | 850  | 66.72      |
| 65002 | 346   | 137  | 209  | 60.40      |
| 65003 | 384   | 103  | 281  | 73.18      |
| 65004 | 223   | 163  | 60   | 26.91      |
| 65005 | 23    | 7    | 16   | 69.57      |
| 65006 | 21    | 6    | 15   | 71.43      |
| 65007 | 12    | 4    | 8    | 66.67      |
| 65008 | 10    | 5    | 5    | 50.00      |
| 65009 | 15    | 12   | 3    | 20.00      |
| 65010 | 3     | 1    | 2    | 66.67      |
| 65011 | 4     | 2    | 2    | 50.00      |
| 65012 | 2     | 1    | 1    | 50.00      |
| 65013 | 4     | 4    | 0    | 0.00       |
| 65014 | 2     | 2    | 0    | 0.00       |
| 65015 | 3     | 3    | 0    | 0.00       |
| 65016 | 2     | 2    | 0    | 0.00       |
| 65017 | 2     | 2    | 0    | 0.00       |

### Resultados

- Agregação sugerida que preserva a política de trânsito
- Ganho elevado para a maioria das redes

#### Introdução

Ambiente de um PTT

Motivação

#### Coleta de dados

Coleta e armazenamento  
de informações

#### Relatórios

Geração de Relatórios

#### Conclusão

# RIB - Tamanho do ASPath e Prepends

## Tamanho do ASPath

peerasn | tammaxaspath

|       |    |
|-------|----|
| 65001 | 11 |
| 65002 | 10 |
| 65003 | 8  |
| 65004 | 4  |
| 65005 | 2  |
| 65006 | 2  |
| 65007 | 2  |
| 65008 | 1  |
| 65009 | 1  |
| 65010 | 1  |
| 65011 | 1  |
| 65012 | 1  |
| 65013 | 1  |
| 65014 | 1  |
| 65015 | 1  |
| 65016 | 1  |
| 65017 | 1  |

## Tamanho do Prepend

peerasn | tammaxprepend

|       |    |
|-------|----|
| 65001 | 10 |
| 65002 | 9  |
| 65003 | 7  |
| 65004 | 2  |
| 65005 | 0  |
| 65006 | 0  |
| 65007 | 0  |
| 65008 | 0  |
| 65009 | 0  |
| 65010 | 0  |
| 65011 | 0  |
| 65012 | 0  |
| 65013 | 0  |
| 65014 | 0  |
| 65015 | 0  |
| 65016 | 0  |
| 65017 | 0  |

## Resultados

- Vemos prepends longos para prefixos mais específicos!
- Junto vemos desagregações "estranhas"

### Introdução

Ambiente de um PTT

Motivação

### Coleta de dados

Coleta e armazenamento  
de informações

### Relatórios

Geração de Relatórios

### Conclusão

## RIB do RouteServer vs. visão global

- Existe prefixo de um participante visto globalmente mas não visto no PTT?
- Existe prefixo mais específico de um participante do PTT visto globalmente?

## Solução

- Comparativo da RIB do RouteServer com RIB do projeto RouteViews
- RIB do RouteViews também é inserida no banco de dados
- As perguntas acima são respondidas diariamente
- Levado em consideração AS com perfil "descentralizado"

## Lista de prefixos mais especificos

| MENOS ESPECIFICO | AS PATH |  | MAIS ESPECIFICO | AS PATH     |
|------------------|---------|--|-----------------|-------------|
| 172.16.192.0/18  | 65002   |  | 172.16.192.0/24 | 65001 65002 |

## Lista de prefixos nao divulgados

| PREFIXO         |  | AS PATH |
|-----------------|--|---------|
| 172.16.196.0/24 |  | 65002   |
| 172.16.0.0/20   |  | 65002   |
| 172.16.55.0/24  |  | 65002   |

## Resultados

- Solicitado ao participante verificar se existe algum erro em filtro com PTT
- Motivado a manter anúncios completos para o PTT

### Introdução

Ambiente de um PTT

Motivação

### Coleta de dados

Coleta e armazenamento de informações

### Relatórios

Geração de Relatórios

### Conclusão

### Conclusão

Estes relatórios auxiliam o Coordenador de Peering a propor melhorias de roteamento aplicadas ao PTT, diretamente a um participante, otimizando toda a infra-estrutura de troca de tráfego.

#### Introdução

Ambiente de um PTT

Motivação

#### Coleta de dados

Coleta e armazenamento de informações

#### Relatórios

Geração de Relatórios

#### Conclusão

## Uso atual

Em uso no PTT-Metro de Curitiba

## Trabalhos Futuros

- Criação de novos relatórios.
- Encaminhamento automático a lista de participantes (ou envio privado)
- Expandir isso para outros PTT
- Fazer análise mais elaborada (vide Geoff Huston)

### Introdução

Ambiente de um PTT

Motivação

### Coleta de dados

Coleta e armazenamento  
de informações

### Relatórios

Geração de Relatórios

### Conclusão

- RouteViews: <http://www.routeviews.org/>
- Potaroo: <http://bgp.potaroo.net/>
- Quagga: <http://www.quagga.net>
- PTT-Metro Curitiba: <http://pr.ptt.br>

## Introdução

Ambiente de um PTT

Motivação

## Coleta de dados

Coleta e armazenamento  
de informações

## Relatórios

Geração de Relatórios

## Conclusão

# Construindo um Route Collector no PTT

PTT-Metro de Curitiba/PR

*LACNIC XII - NAPLA 2009 - Cidade do Panamá - Panamá*  
26 de Maio 2009

Introdução

Ambiente de um PTT

Motivação

Coleta de dados

Coleta e armazenamento  
de informações

Relatórios

Geração de Relatórios

Conclusão

Pedro R. Torres Jr. [torres@pop-pr.rnp.br](mailto:torres@pop-pr.rnp.br)  
PTT-Metro de Curitiba