

Implementando IPv6 no RSiX

Andrey Vedana Andreoli, Leandro Márcio Bertholdo,

Liane M. R. Tarouco

POP-RS / UFRGS / RSiX

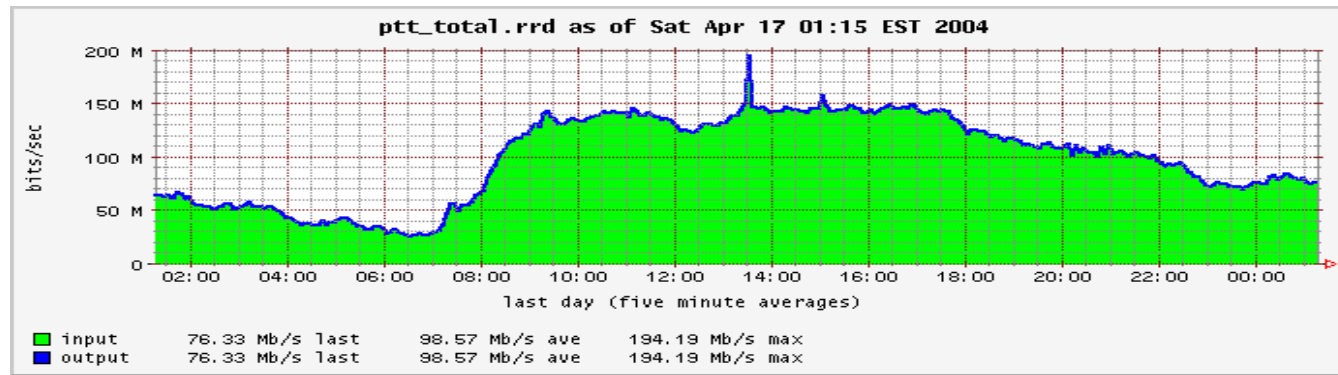
Rua Ramiro Barcelos, 2574 - Porto Alegre – RS

{andrey, berthold, liane}@penta.ufrgs.br

Sumário

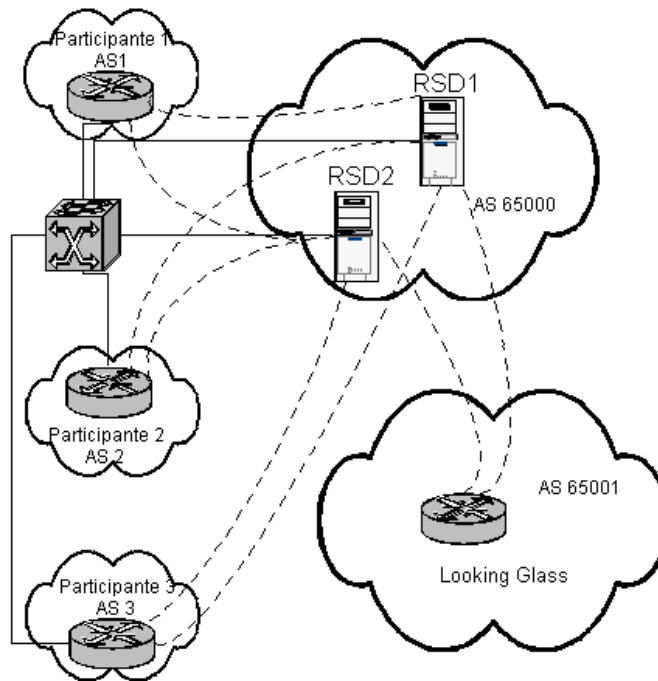
- O RSiX sobre IPv4
- Motivação para uso de IPv6
- Iniciativas em IX v6 no exterior
- Dimensionamento e endereçamento v6
- Mudanças em Route Servers
- Roteamento v6 em PTTs
- Configurações iniciais e situação atual
- Conclusões e trabalhos futuros
- Perguntas

O RSiX baseado em IPv4



- Início das operações no ano de 2000
- Reúne atualmente 15 participantes
- Média de tráfego agregado de 150 Mbps/s
- Operação exclusivamente sobre IPv4
- Agrega participantes acadêmicos e comerciais

O RSiX baseado em IPv4



- Baseia-se em acordos de tráfego multilaterais (ATM)
- Utiliza o recurso de Route Servers
- Todos os peers fazem uso do endereçamento local exclusivo ao PTT (200.132.1.0/24)

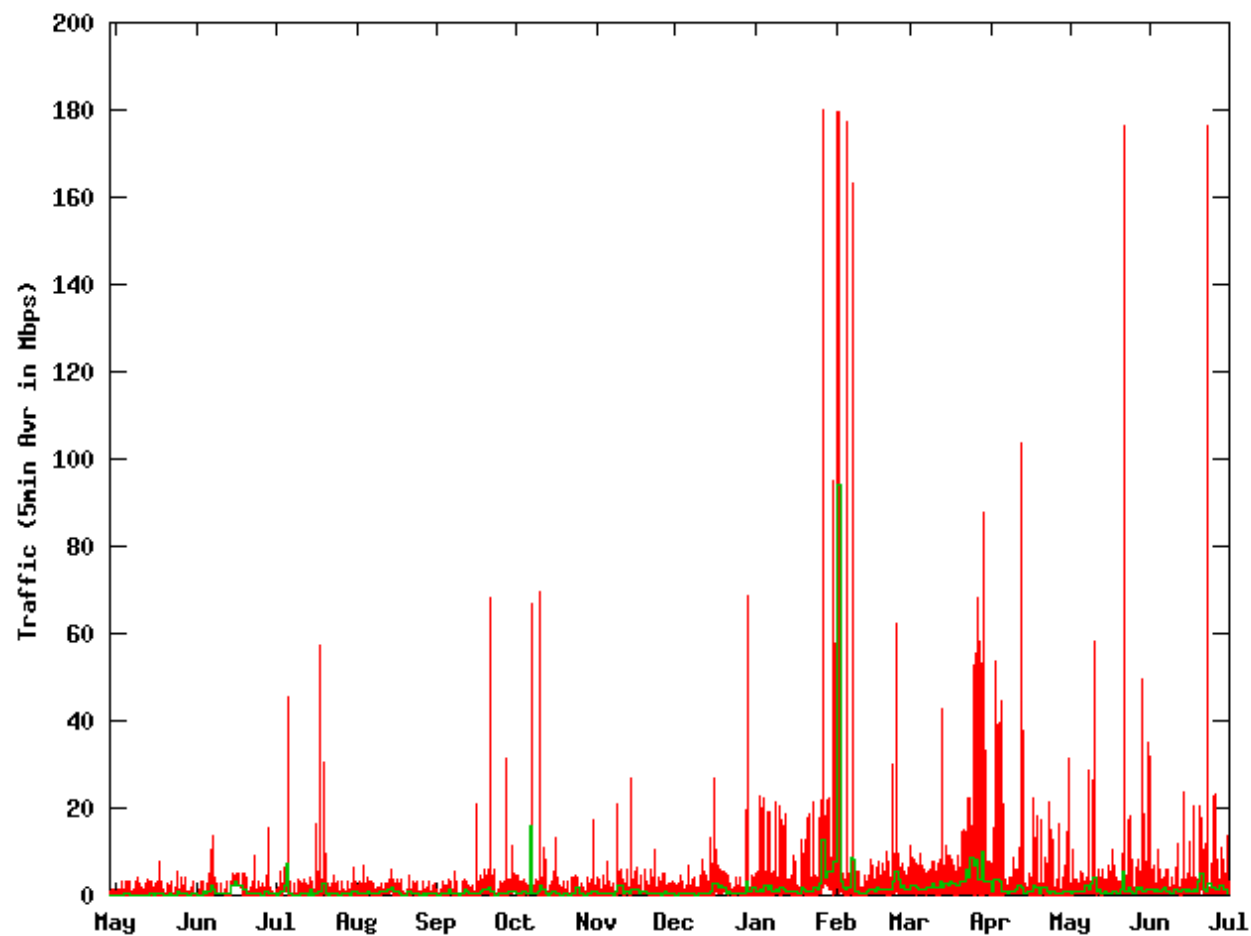
Motivação para utilização de IPv6

- Consolidação do RSiX sobre IPv4
- Histórico de utilização de IPv6 no POP-RS
 - Ativação túnel IPv6 via ANSP (2000)
 - Projeto piloto de IPv6 na RNP (POP-RS) (2001)
- Alocação de bloco de produção /42 ao POP-RS
- Demandas pela utilização de IPv6 por parte de instituições da Rede Tchê
- Atividades de pesquisa envolvendo IPv6
- Possibilidade de expansão em redes comerciais
- Iniciativas de IX v6 no exterior

Iniciativas Mundiais (NSPIX-6)

- IPv6-based Internet Exchange in Tokyo
- Início das atividades em 2001
- IX v6 only
- Conecta cerca de 20 provedores
- Proporção de 1/1000 do tráfego IPv4

Iniciativas Mundiais (NSPIXP-6)

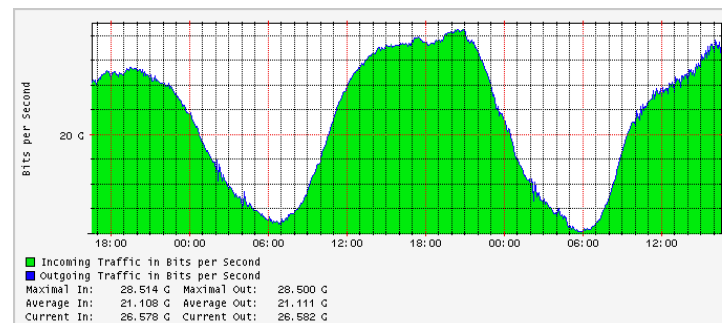


Iniciativas Mundiais (6TAP)

- Sede em Chicago/USA
- IX v6 only
- Conecta cerca de 22 provedores
- IX público
- Envolve cerca de 450 prefixos v6

Iniciativas Mundiais (AMS-IX)

- Sede em Amsterdam
- Faz parte do Europe-IX Association
- Fornece acesso IPv4 e IPv6
- IX público
- Envolve 191 participantes (IPv4 e IPv6)
- Alcança a média de 21Gbps/s (IPv4 e IPv6)



Proposta de mudança para IPv6

- Manter a estabilidade de operação sobre IPv4
- Manter a mesma topologia e equipamentos
- Permitir que a estrutura possa ser utilizada para peering IPv4, IPv6 e se desejado, ambos (Dual stack)
- Manter um ponto para exploração das áreas de gerência, segurança e engenharia sobre IPv6
- Fornecimento de estrutura capaz de expandir a conectividade IPv6 em redes acadêmicas e comerciais

Dimensionamento e endereçamento

- Existe alguma nova forma de IX IPv6?
 - Router Advertisement (RA) poderá fornecer mudanças em trocas de tráfego
- Utilização de bloco exclusivo para o PTT?
 - Facilidade na gerência e trouble shooting
 - Forma convencional de IX
 - Empregada em outros IX no exterior

Dimensionamento e endereçamento

- Alocação de bloco IPv6 exclusivo para o PTT (2001:12f0:300:1::/64)
- Como distribuí-lo?

Route Server fazendo Router Advertisement? Não

- Dificulta o controle e troubleshooting de peers
 - Mudanças de endereçamento de peers pode ocorrer já que parte do endereço é formado pelo MAC.
- Solução: Definição estática de endereços aos participantes

Dimensionamento e endereçamento

- **Determinação da regra de endereçamento**

- Regra aplicável para peers e Route Server

- Endereço base: 2001:12f0:300:1:xxxx:xxxx:xxxx:xxxx

- Converter o ASn do participante para hexadecimal

- Coloca o valor no último grupo de 16bits

Dimensionamento e endereçamento

- Exemplo:

Universidade Federal do Rio Grande do Sul

AS 19200

Decimal: 19200 = Hexadecimal: 4b0

Endereço base:

2001:12f0:300:1:xxxx:xxxx:xxxx:xxxx

Endereço resultante:

2001:12f0:300:1::4b0

Route Servers

- Ativando IPv6 na interface do Route Server (FreeBSD)

```
ipv6_enable="YES"
```

```
ipv6_ifconfig_xl0="2001:12f0:300:1::aaaa:fde8 prefixlen 64"
```

- Ativando IPv6 na interface do Route Server (FreeBSD)

```
xl0: flags=8843<UP,BROADCAST,RUNNING,SIMPLEX,MULTICAST> mtu 1500
```

```
inet 200.132.1.249 netmask 0xfffff00 broadcast 200.132.1.255
```

```
inet6 fe80::210:4bff:fec9:6cd8%xl0 prefixlen 64 scopeid 0x1
```

```
inet6 2001:12f0:300:1::aaaa:fde8 prefixlen 64
```

```
ether 00:10:4b:c9:6c:d8
```

```
media: Ethernet autoselect (100baseTX <full-duplex>)
```

```
status: active
```

Route Servers

- Ativando roteamento BGP com suporte a IPv6

Diversos softwares podem ser utilizados para este fim

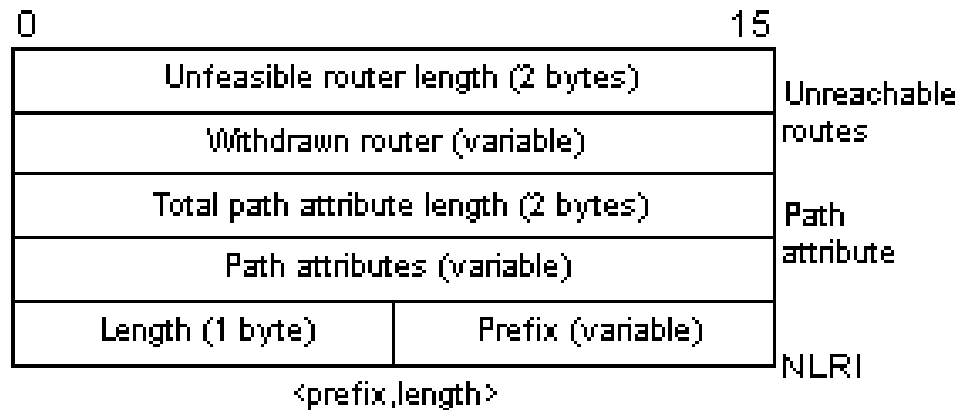
Utilização do pacote GNU Zebra no RSiX

Suporte ao MP-BGP, OSPFv3, RIPns, RA

Nenhuma modificação foi necessária na versão atual de binários, visto que já havia suporte ao MP-BGP

Como rotear IPv6 entre ASes?

- Chamado de Multiprotocol BGP ou MP-BGP;
- Baseia-se no BGP-4, incluindo funcionalidades para carregar informações de outros protocolos;
- Acrescenta identificadores, que atuam na identificação do protocolo a ser suportado:
 - Address Family Identifier (AFI)
 - Subsequent Address Family Identifier (Sub-AFI)
 - Negociadas no início da sessão BGP, na forma de “capability”

O que muda? Mensagem UPDATE

- Os campos de Path Attribute carregam informações de atributos como: Origin, Next-hop, AS-PATH, entre outros;
- Passam a ser utilizados mais dois atributos novos:
 - MP_REACH_NLRI
 - MP_UNREACH_NLRI

Códigos mais comuns para AFI e Sub-AFI

| Código AFI | Código Sub-AFI | Significado |
|-------------------|-----------------------|------------------------------|
| 1 | 1 | IPv4 Unicast |
| 1 | 2 | IPv4 Multicast |
| 1 | 3 | IPv4 based VPN |
| 2 | 1 | IPv6 Unicast |
| 2 | 2 | IPv6 Multicast RPF |
| 2 | 3 | IPv6 Unicast e Multicast RPF |
| 2 | 4 | IPv6 Label |
| 2 | 128 | IPv6 VPN |
| | | ... |

Exemplo de mensagem de UPDATE (MP_REACH_NLRI)

```
⊞ Frame 100 (165 bytes on wire, 165 bytes captured)
⊞ Ethernet II, Src: 00:10:7b:1c:fa:e0, Dst: 00:60:3e:2f:de:81
⊞ Internet Protocol Version 6
⊞ Transmission Control Protocol, Src Port: bgp (179), Dst Port: 11010 (11010)
⊞ Border Gateway Protocol
  ⊞ UPDATE Message
    Marker: 16 bytes
    Length: 87 bytes
    Type: UPDATE Message (2)
    Unfeasible routes length: 0 bytes
    Total path attribute length: 64 bytes
  ⊞ Path attributes
    ⊞ ORIGIN: IGP (4 bytes)
    ⊞ AS_PATH: 65501 65503 (9 bytes)
    ⊞ MP_REACH_NLRI (51 bytes)
      ⊞ Flags: 0x80 (Optional, Non-transitive, Complete)
        Type code: MP_REACH_NLRI (14)
        Length: 48 bytes
        Address family: IPv6 (2)
        Subsequent address family identifier: Unicast (1)
      ⊞ Next hop network address (16 bytes)
        Next hop: 3ffe:2620:7:801::1 (16)
        Subnetwork points of attachment: 0
      ⊞ Network layer reachability information (27 bytes)
        ⊞ 3ffe:2620:11:1::/64
          MP Reach NLRI prefix length: 64
          MP Reach NLRI prefix: 3ffe:2620:11:1::
        ⊞ 3ffe:2620:11:2::/64
        ⊞ 3ffe:2620:11:3::/64
```

Conexão Inicial IPv6 no RSiX

Inicialmente 2 participantes se tornaram aptos para conexão:

Participante 1: **UFRGS** (2001:12f0:301::/48)

AS 19200 = Hexadecimal: 4b0

Participante 2: **Unisinos** (2001:12f0:302::/48)

AS 19611 = Hexadecimal: 4c9b

Configuração de sessões MP-BGP (Route Server)

```
router bgp 65000
  neighbor 2001:12f0:300:1::4b0 remote-as 19200
  neighbor 2001:12f0:300:1::4b0 description UFRGS-ipv6
  !
  address-family ipv6
  neighbor 2001:12f0:300:1::4b0 activate
  neighbor 2001:12f0:300:1::4b0 maximum-prefix 10
  neighbor 2001:12f0:300:1::4b0 soft-reconfiguration inbound
  neighbor 2001:12f0:300:1::4b0 attribute-unchanged as-path next-hop
exit-address-family
!
```

Configuração de sessões MP-BGP (UFRGS - cisco)

```
ipv6 unicast-routing
!
interface Ethernet1/0
ipv6 address 2001:12F0:300:1::4B0/64
!
router bgp 19200
 neighbor 2001:12F0:300:1::AAAA:FDE8 remote-as 65000
!
 address-family ipv6
  neighbor 2001:12F0:300:1::AAAA:FDE8 activate
  neighbor 2001:12F0:300:1::AAAA:FDE8 next-hop-self
  neighbor 2001:12F0:300:1::AAAA:FDE8 soft-reconfiguration inbound
 network 2001:12F0:301::/48
 exit-address-family
```

Situação atual do participantes IPv6

```
rsd1.rsix.tche.br> sh bgp ipv6 sum
```

```
BGP router identifier 200.132.1.xxx, local AS number 65000
```

```
104 BGP AS-PATH entries
```

```
0 BGP community entries
```

```
Neighbor V AS MsgRcvd MsgSent TblVer InQ OutQ Up/Down Sate/PfxRcd
2001:12f0:300:1::4b0 4 19200 3500 3499 0 0 0 1d22h17m 1
2001:12f0:300:1::4c9b 4 19611 2818 2823 0 0 0 1d22h55m 1
```

```
Total number of neighbors 2
```

Situação atual do participantes IPv6

```
rsd1.rsix.tche.br> sh bgp ipv6
```

```
BGP table version is 0, local router ID is 200.132.1.xxx
```

```
Status codes: s suppressed, d damped, h history, * valid, > best, i -  
internal
```

```
Origin codes: i - IGP, e - EGP, ? - incomplete
```

| Network | Next Hop | Metric | LocPrf | Weight | Path |
|-----------------------|-----------------------|--------|--------|--------|---------|
| *> 2001:12f0:301::/48 | 2001:12f0:300:1::4b0 | 0 | 0 | | 19200 i |
| *> 2001:12f0:302::/48 | 2001:12f0:300:1::4c9b | 0 | 0 | | 19611 i |

```
Total number of prefixes 2
```

```
rsd1.rsix.tche.br>
```

Conclusões

- O passo mais complicado na ativação de peer IPv6 tem sido a obtenção de versões com suporte a MP-BGP
 - Versão nova não significa suporte a IPv6!
- A operação com IPv6 não aumenta a complexidade em relação ao IPv4
- A forma de IX empregada pelo IPv4 continua sendo a mais usual para IPv6
- O aumento de IX v6 poderá contribuir fortemente para a expansão de redes IPv6, aumentando sua abrangência.

Trabalhos futuros

- Ativação de IPv6 em outros peers já presentes sobre IPv4
- Estudo sobre a gerência sobre IPv6
- Aprofundamento no impacto na segurança em IX do IPv6
- Implementação de formas de visualização das rotas IPv6
- Disponibilização de estatística global de conectividade, envolvendo IPv6 de forma separada ao IPv4

Dúvidas, questionamentos, sugestões...



Site do RSiX:
<http://www.rsix.tche.br>

Obrigado!