14º GTER

“Graceful Restart” - Mais um passo em direção a redes mais estáveis

caio@juniper.net

14º GTER

Sumário

- Objetivos e Motivações
- Modificações propostas para o BGP (draft-ietf-idr-restart-03.txt)
  - End-of-RIB marker
  - "Graceful Restart Capability”
- Procedimentos para execução do “Graceful Restart”
- Exemplo prático
- Considerações

Sumário

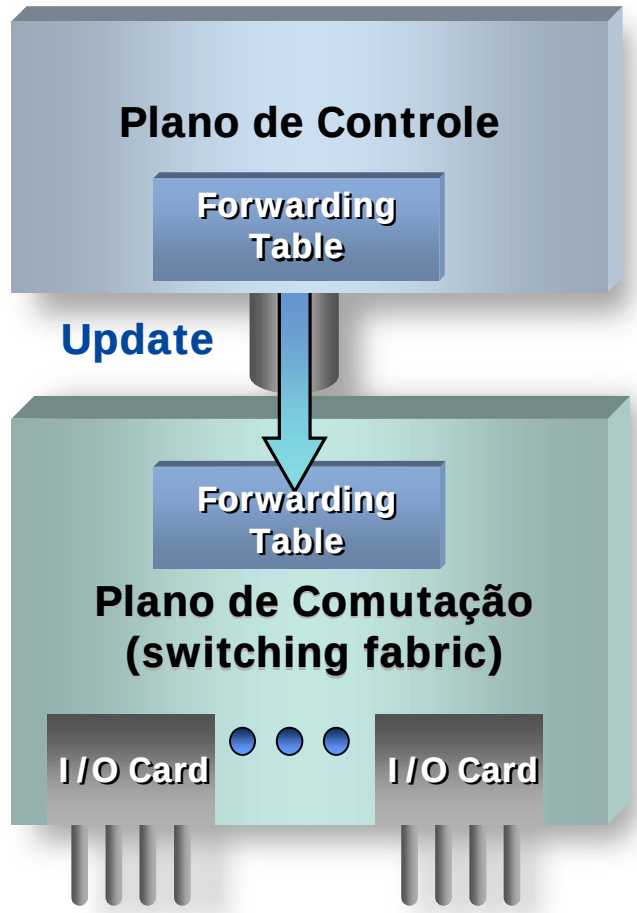
- **Objetivos e Motivações**
- Modificações propostas para o BGP (draft-ietf-idr-restart-03.txt)
  - End-of-RIB marker
  - "Graceful Restart Capability”
- Procedimentos para execução do “Graceful Restart”
- Exemplo prático
- Considerações

Efeitos negativos da reinicialização

- **Normalmente quando o BGP reinicializa em um roteador :**
  - **Todos os peers detectam que a conexão caiu**
  - **A transição “down/up” resulta em um “routing-flap” com recálculo das rotas e possível “flap” da “forwarding-table”.**
  - **Esse comportamento pode se espalhar para múltiplos domínios de roteamento.**
  - **Pode criar “buracos” e “loops” temporários**
  - **Consomem recursos do plano de controle dos roteadores afetados pelo “flap”.**

Plano de Controle e Plano de Comutação de um Roteador

- **Plano de Controle**
  - Manipula pacotes de roteamento, OSPF, IS-IS, BGP,...
  - Roda os protocolos de roteamento
  - Gera a tabela de roteamento e tabela de “forwarding”
- **Plano de Comutação**
  - Manipula todos os pacotes que atravessam o roteador
  - Executa operações sobre pacotes ex: “route-lookup”



Após dispor de uma “forwarding table”, o Plano de Comutação pode executar suas funções sem o Plano de Controle

Sumário

- Objetivos e Motivações
- **Modificações propostas para o BGP (draft-ietf-idr-restart-03.txt)**
  - **End-of-RIB marker**
  - **"Graceful Restart Capability”**
- Procedimentos para execução do “Graceful Restart”
- Exemplo prático
- Considerações

Abertura de Sessão BGP Convencional

OPEN – Capabilities
(Route refresh, Muliprotocol...)

OPEN – Capabilities
(Route refresh, Muliprotocol...)

UPDATE – NLRI (Rotas)

UPDATE – NLRI (Rotas)

Não existe sinalização explícita de “fim-de-update”

End-of-RIB

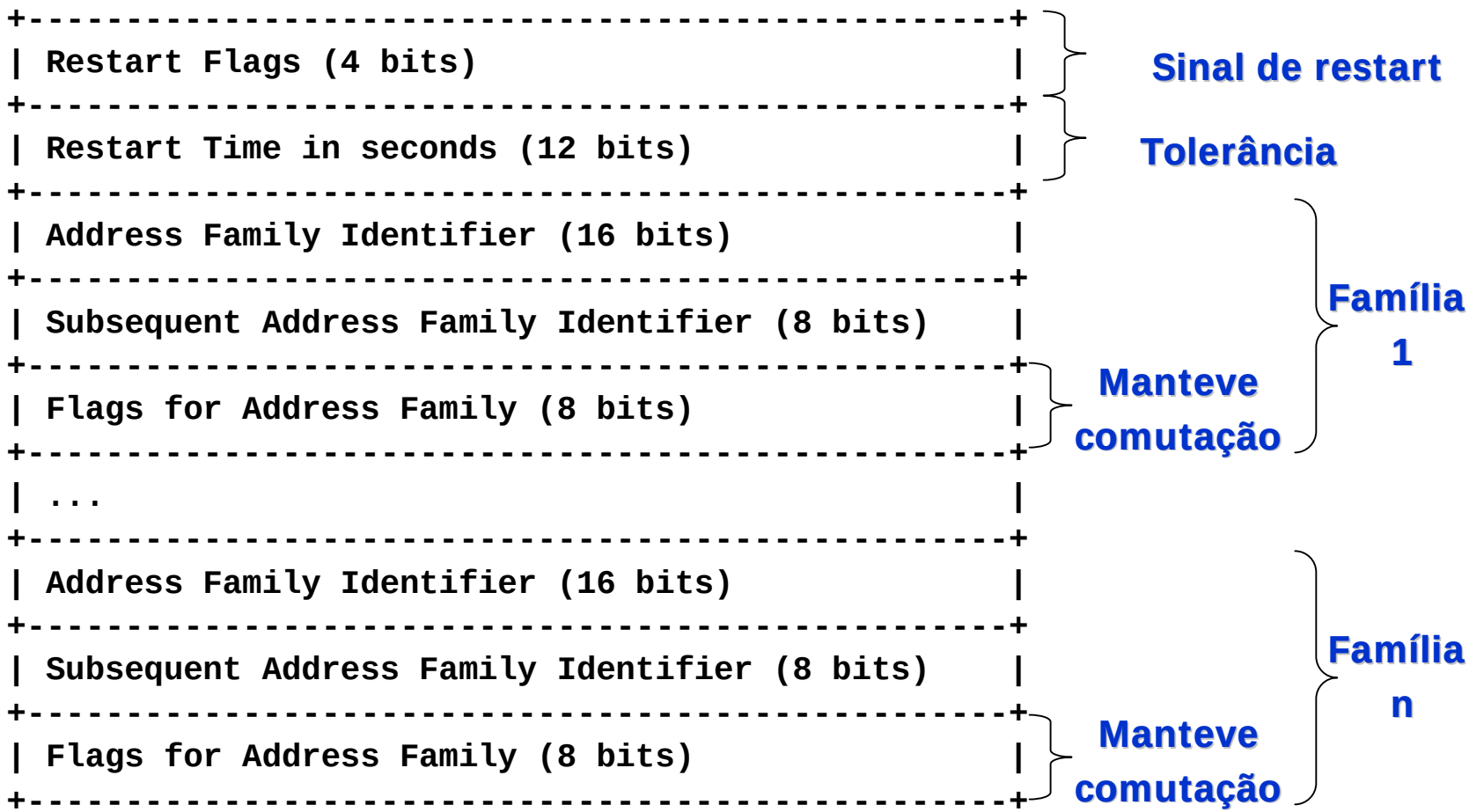
- Uma mensagem UPDATE com “withdrawn NLRI” vazio é especificada como “End-Of-RIB Marker”. Indica ao peer que o “update” de roteamento foi finalizado.
- Embora o “End-of-RIB Marker” seja especificado para o graceful restart, é recomendado o seu uso em todas as conexões pois pode ser útil na convergência do protocolo de roteamento.

Graceful Restart Capability

- **A “Graceful Restart Capability” é uma capacidade do BGP [BGP-CAP] que pode ser usada por um “BGP speaker” para indicar a sua habilidade de preservar o seu estado de “forwarding” (comutação) durante o restart do BGP. Também pode ser utilizado para sinalizar aos seus peers a intenção de gerar a “End-Of-RIB marker” ao final das atualizações iniciais de roteamento.**

Graceful Restart Capability

- A capacidade (capability) é definida abaixo:
  - Capability code: 64
  - Capability length: variable



Graceful Restart Capability - Campos

- **Restart Flags:** O bit mais significativo é definido como Restart State bit. Quando “1” indica que o “BGP speaker” reinicializou.
- **Restart Time:** Tempo estimado (s) para reestabelecimento da sessão BGP após o “restart”. É usado como “timeout” para acelerar a convergência caso o roteador não retorne após o “restart”.

Graceful Restart Capability - Campos

- **Address Family Identifier (AFI):** Identifica o protocolo suportado pela facilidade de graceful restart
- **Subsequent Address Family Identifier (Sub-AFI):** Informação adicional sobre NLRI.
- **Flags for Address Family:** Contém flags para o protocolo em questão. O bit mais significativo é definido como “Forwarding State bit” que pode ser usado para indicar se o estado de comutação para a família <AFI, Sub-AFI> foi preservado no último restart. Quando “1” indica que a comutação foi preservada.

Sumário

- Objetivos e Motivações
- Modificações propostas para o BGP (draft-ietf-idr-restart-02.txt)
  - End-of-RIB marker
  - "Graceful Restart Capability”
- **Procedimentos para execução do “Graceful Restart”**
- Exemplo prático
- Considerações

Abertura de Sessão BGP

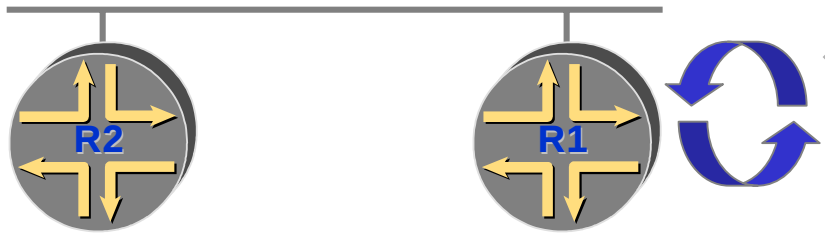
OPEN – Graceful Restart Capability

OPEN – Graceful Restart Capability

UPDATE com EOR

UPDATE com EOR

Procedimentos



- Mantém o “forwarding state” das rotas BGP na Loc-RIB.
- Marca rotas como “stale”
- Mantém comutação normalmente

TCP RST

Marca rotas como “stale” e mantém tabela de comutação
Verifica restart time

OPEN – GRCapability (Restart State = 1 e Forwarding State =1)

OPEN – GRCapability (FS = 1)

Aguarda RIB completa de todos os peers

UPDATE com EOR

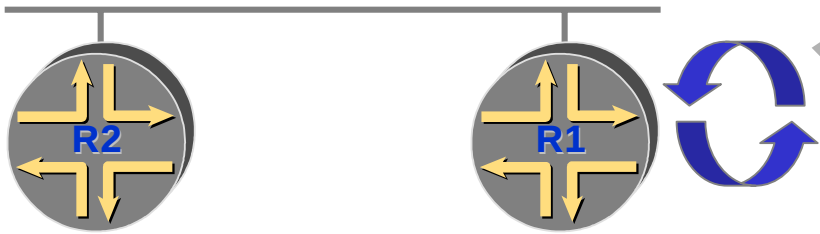
Seleciona rotas, remove “stale” e atualiza tabela

UPDATE com EOR

Remove “stale” e atualiza tabela

Procedimentos

R2 não percebe inicialmente o restart



- Mantém o “forwarding state” das rotas BGP na Loc-RIB.
- Marca rotas como “stale”
- Mantém comutação normalmente

OPEN – GRCapability (Restart State = 1 e Forwarding State =1)

OPEN ??? - Assume restart, marca rotas como “stale” e mantém tabela de comutação

OPEN – GRCapability (FS = 1)

Aguarda RIB completa de todos os peers

UPDATE com EOR

Seleciona rotas, remove “stale” e atualiza tabela

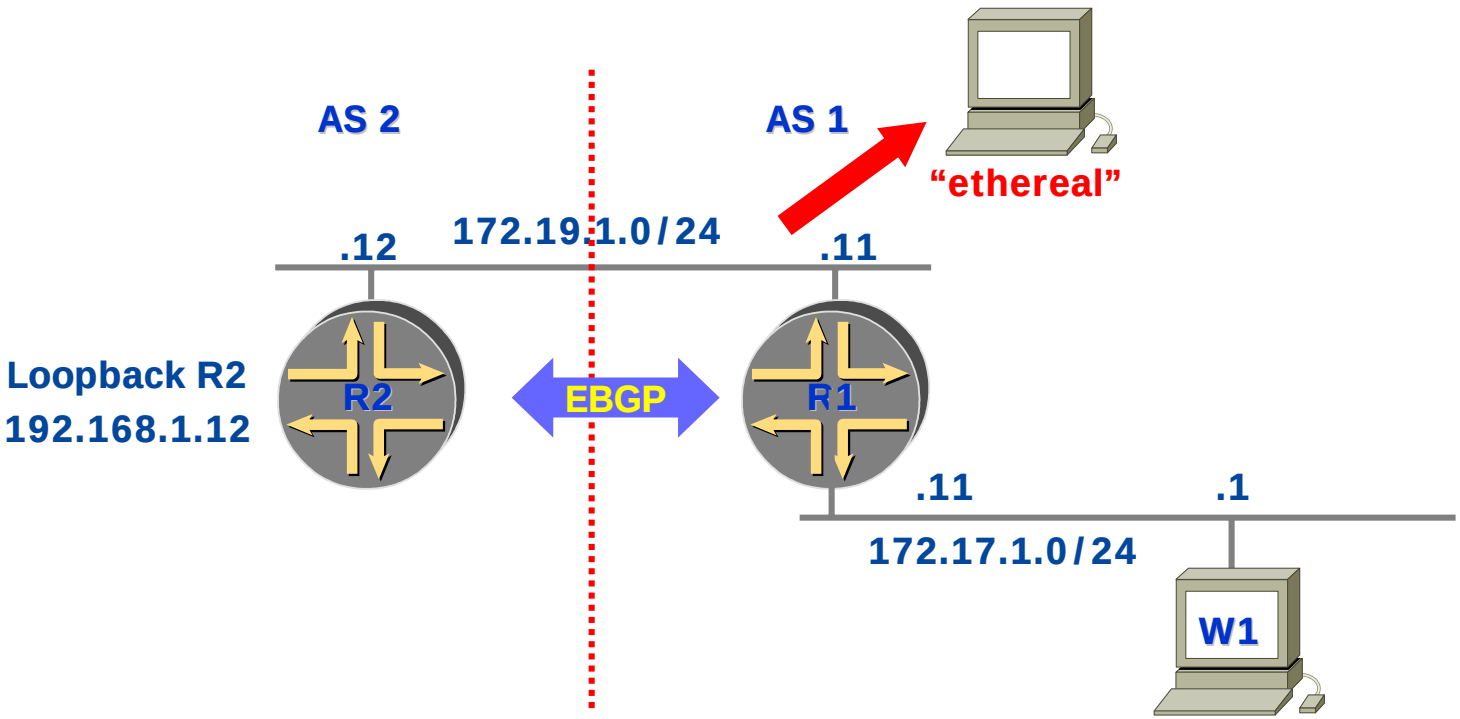
UPDATE com EOR

remove “stale” e atualiza tabela

Sumário

- Objetivos e Motivações
- Modificações propostas para o BGP (draft-ietf-idr-restart-02.txt)
  - End-of-RIB marker
  - "Graceful Restart Capability”
- Procedimentos para execução do “Graceful Restart”
- **Exemplo prático**
- Considerações

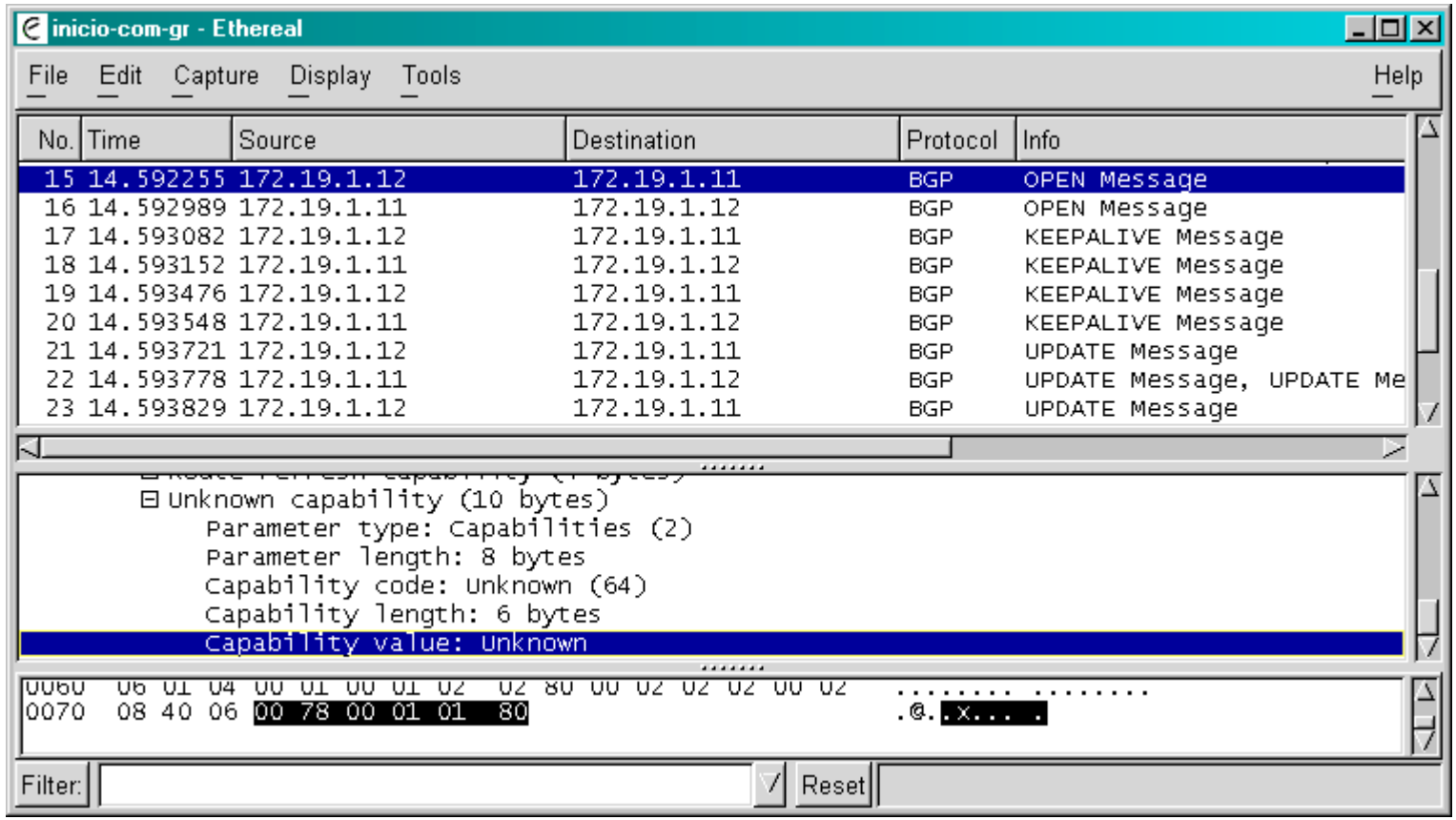
Demostração - Topologia



Para R2
172.17.1.1 via BGP

Para R1
192.168.1.12 via BGP

Início - OPEN



Restart State = 0 (início normal)

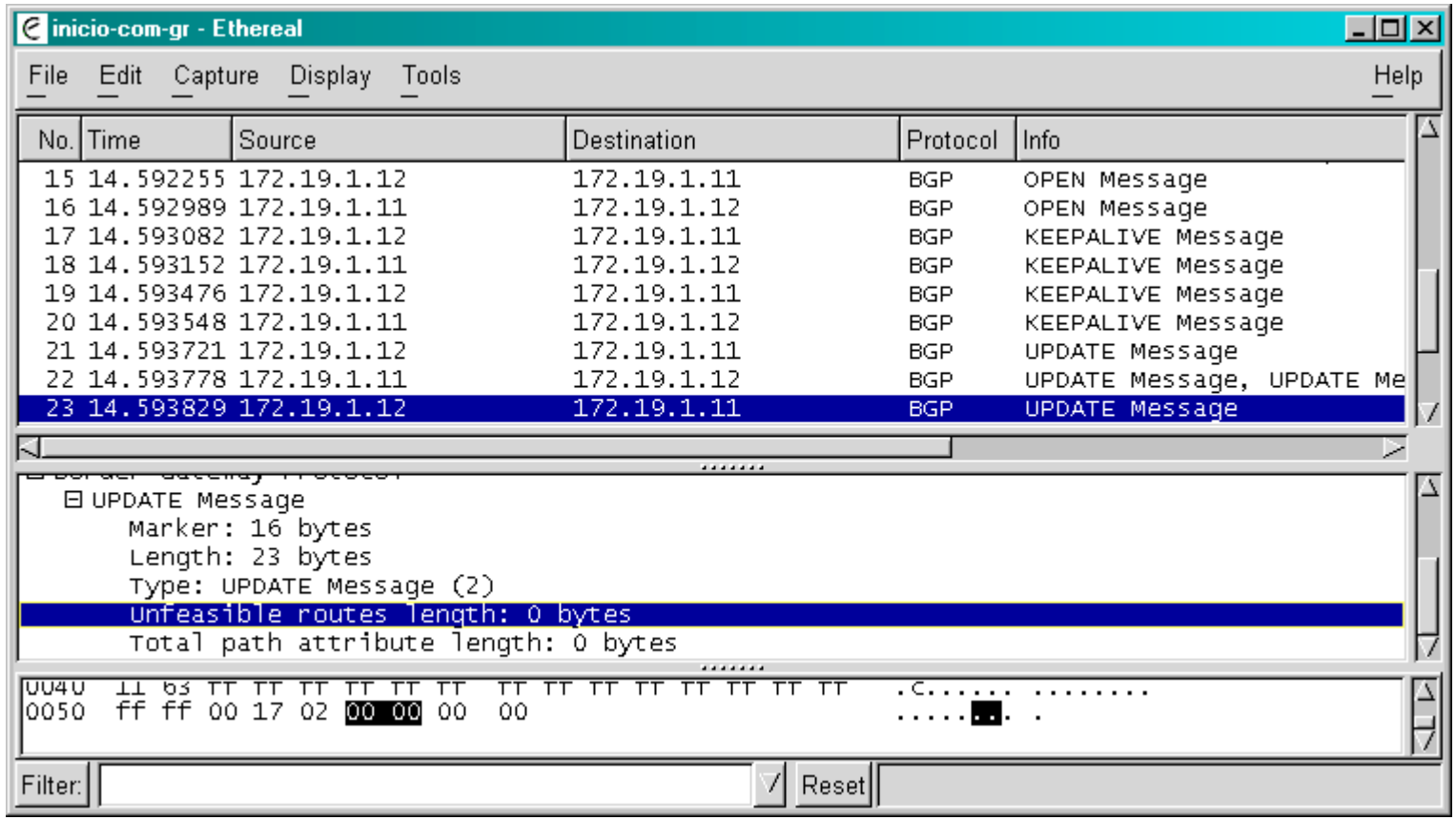
Restart Time = 120 (segundos)

AFI = 01 (IPv4)

Sub-AFI = 01

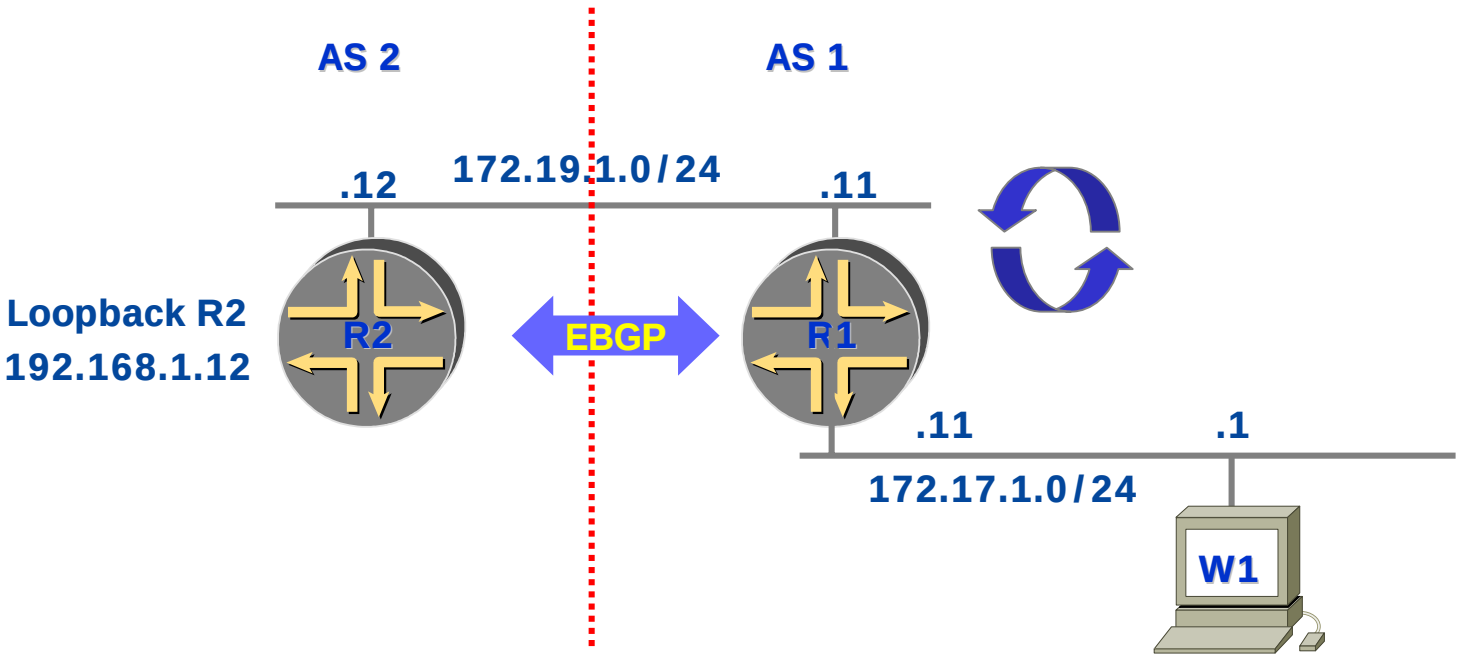
Forwarding State = 1 (manteve fowarding anterior)

Início - UPDATE

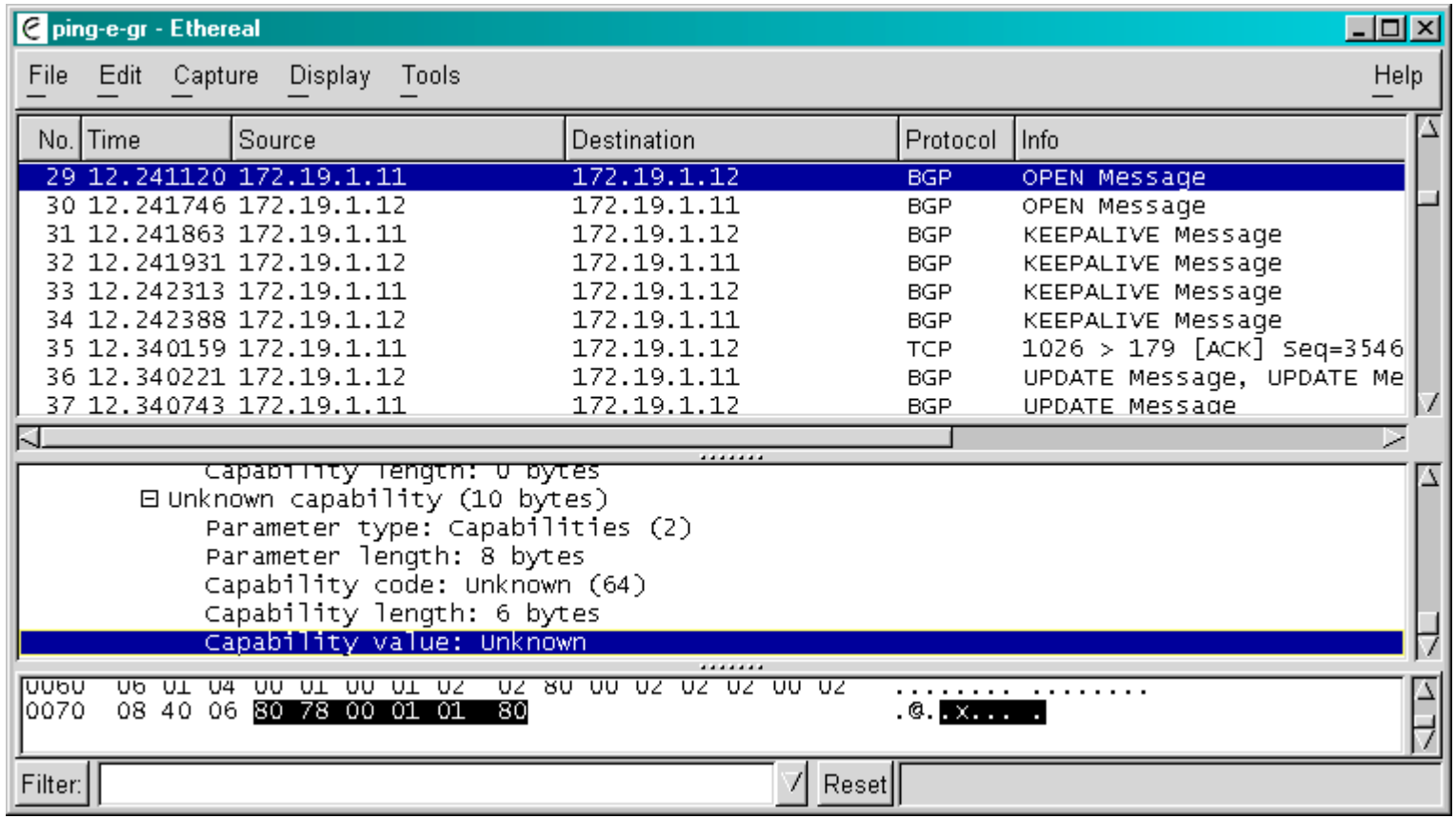


End-Of-RIB Marker

Restart em R1



Graceful Restart - OPEN



Restart State = 1 (graceful restart)

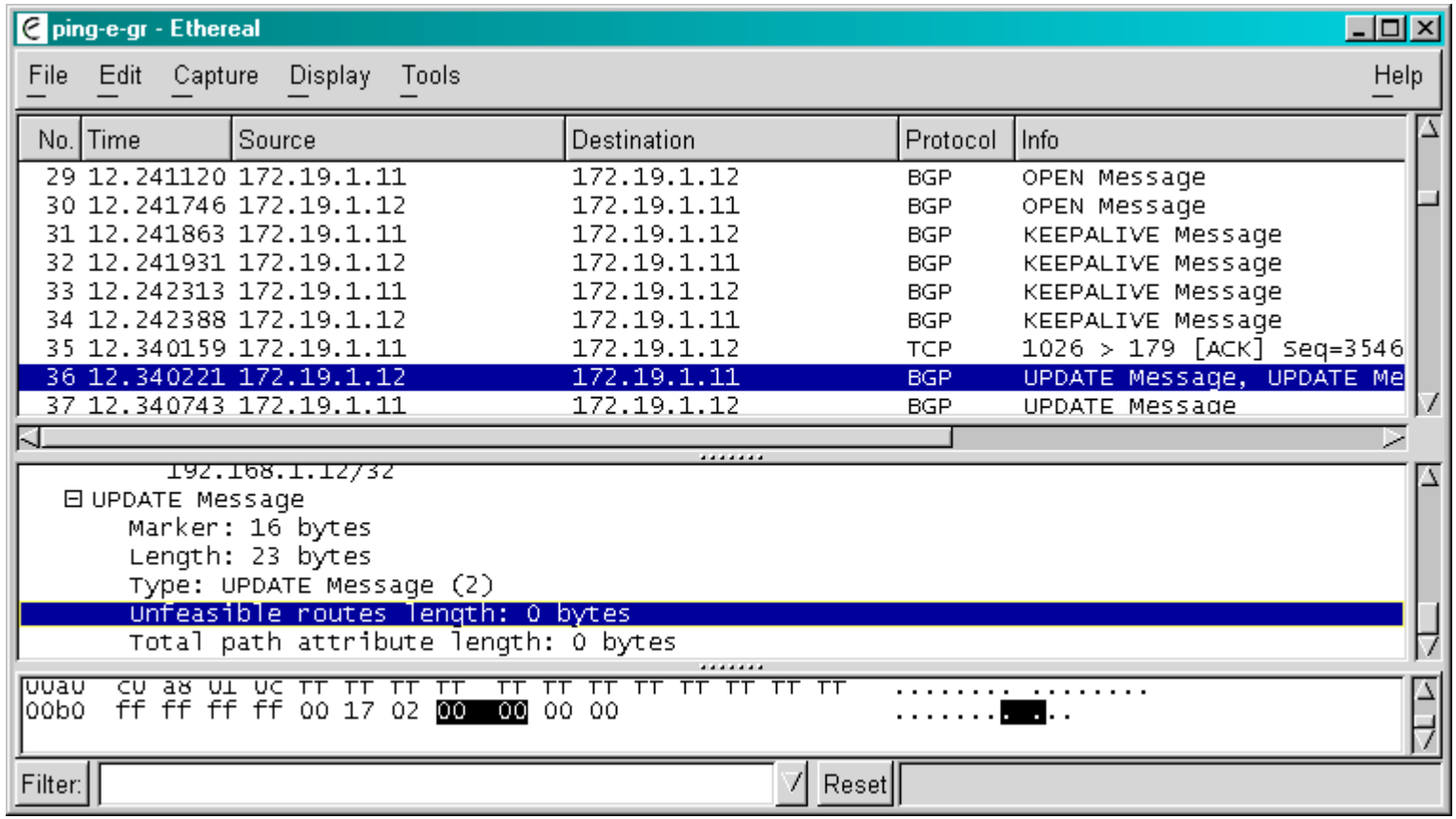
Restart Time = 120 (segundos)

AFI = 01 (IPv4)

Sub-AFI = 01

Forwarding State = 1 (manteve fowarding anterior)

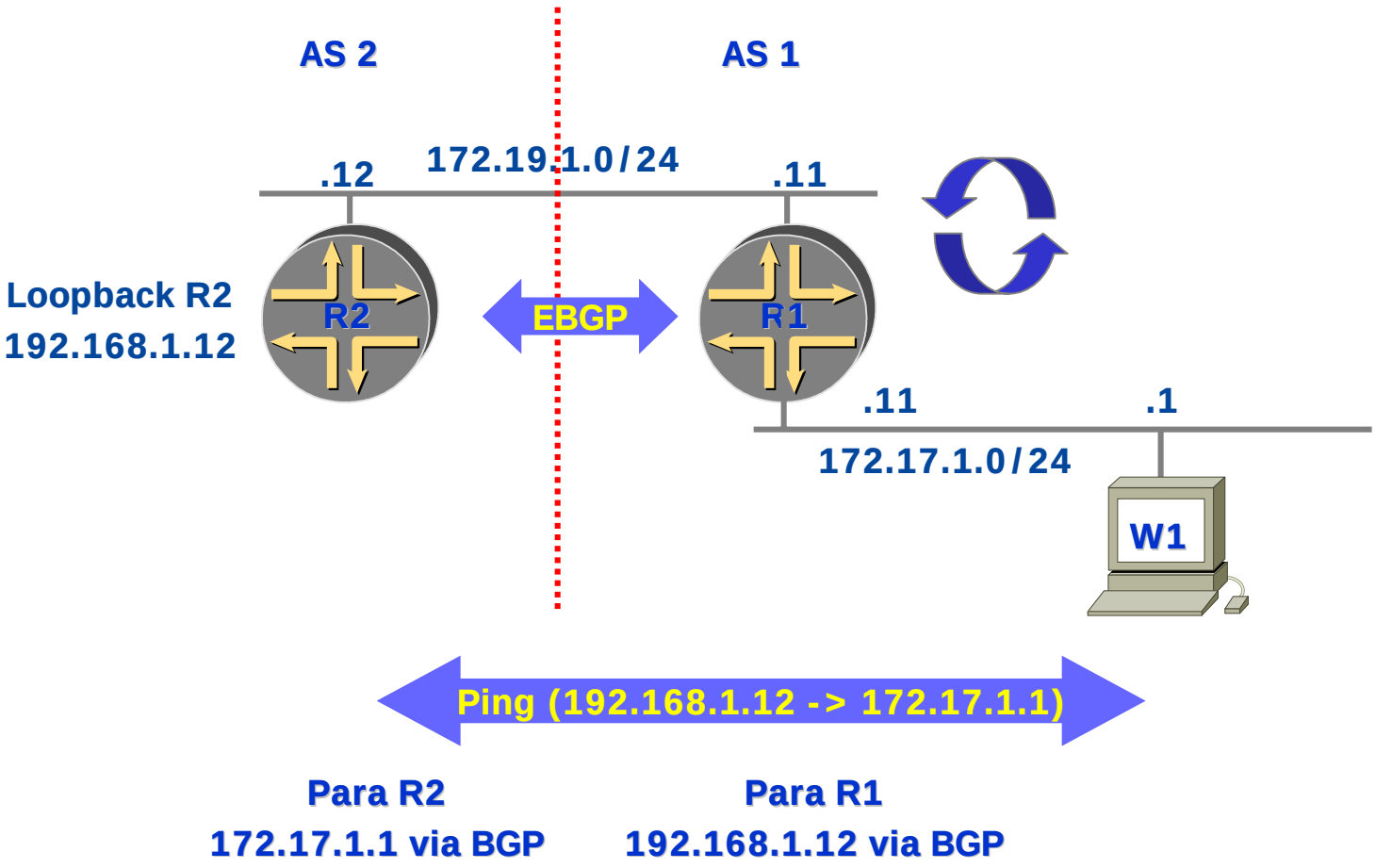
Graceful Restart - UPDATE



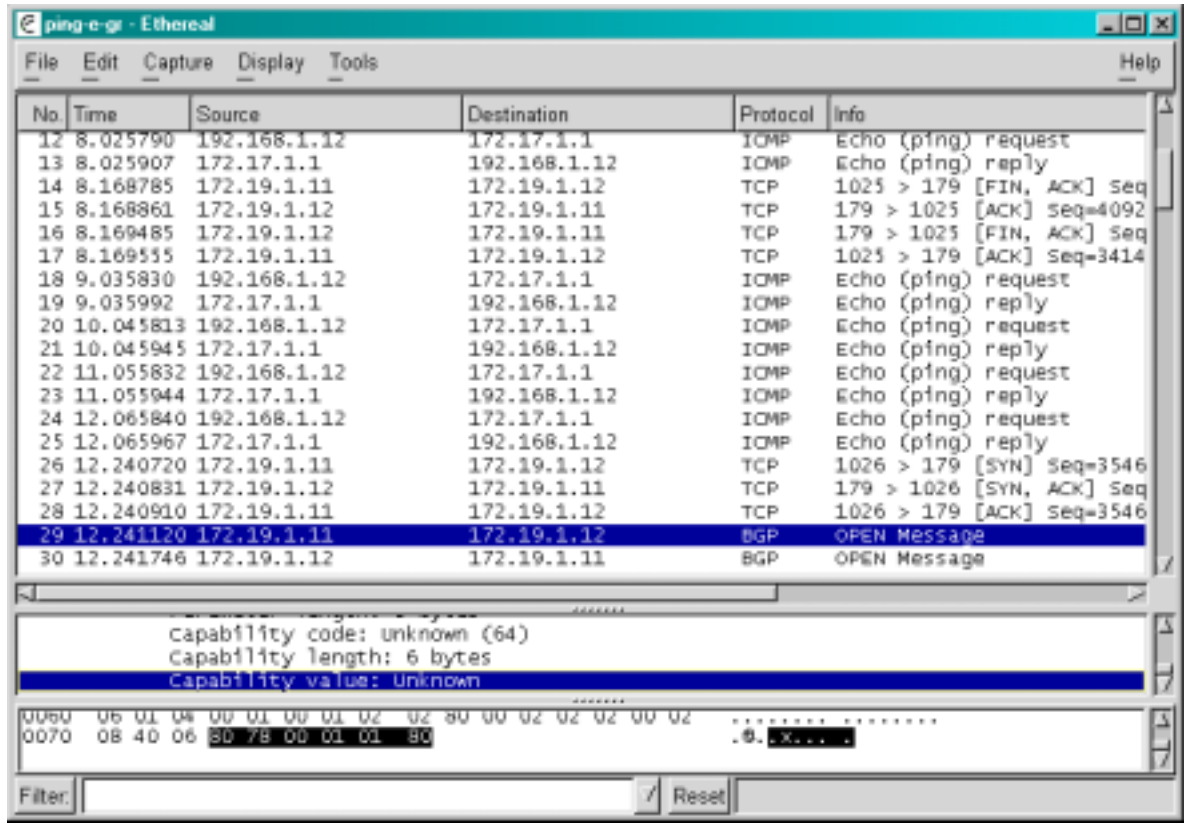
36 - R2 (receiving speaker) envia RIB e EOR

37 – Somente após receber EOR é que R1 (restarting speaker) anuncia

Trânsito de Pacotes



Graceful Restart – Comutação Initerrupta



12-13	Ping entre R2 e W1
...	user@R1> restart routing
14-17	Término de conexão TCP
18-25	Comutação não interrompida
26-28	Reinício de conexão TCP/BGP
29	OPEN com restart flag = 1 e forwarding state = 1

Sumário

- Objetivos e Motivações
- Modificações propostas para o BGP (draft-ietf-idr-restart-02.txt)
  - End-of-RIB marker
  - "Graceful Restart Capability”
- Procedimentos para execução do “Graceful Restart”
- Exemplo prático
- **Considerações**

Considerações

- Existe a probabilidade de geração de loops ou “buracos” caso ocorra modificações nas informações de rotamento antes do roteador finalizar o “restart”
- Dependendo da topologia, caso alguns roteadores rodando IBGP não possuam Graceful Restart, há uma exposição maior à geração de loops e “buracos”.
- Restart Time ainda não foi bem definido. É necessário maior experiência.

Considerações, cont

- **Há poucos benefícios em ativar o Graceful Restart para o BGP em ambientes onde existe uma interação forte entre BGP e IGP e o IGP não dispõe dessa facilidade.**

Referências

- draft-ietf-idr-restart-03.txt **Title: Graceful Restart Mechanism for BGP**
- draft-shand-isis-restart-01.txt **Title: Restart signaling for ISIS**
- draft-rekhter-bgp-mpls-restart-00.txt **Title: Graceful Restart Mechanism for BGP with MPLS**
- draft-ietf-ospf-hitless-restart-02.txt **Title: Hitless OSPF Restart**

14º GTER

Obrigado

caio@juniper.net