



SIM
TELECOM



Medição da malha BGP no Brasil



Luís Felipe Balbinot
lbabinot@simtelecom.net.br

Agenda

- Estado atual da malha BGP no Brasil
- Para que medir?
- Como, onde e o que medir?
- Proposta de uma plataforma de medições
- Iniciativas similares
- Aplicação exemplo
- Perguntas e debate

Estado atual da malha BGP no Brasil

- **Algumas estatísticas (visão do AS26622 em *partial routing*):**
 - ~1570 prefixos anunciados (pouca variação)
 - ~200 sistemas autônomos ativos
 - Maior AS-Path com 10 ASs
 - Maior *prepend* com 5 ASs (AS26610)
- **Tamanho do AS-Path diretamente ligado à malha de interconexão (menor é melhor)**
 - Média do AS26622: 2,548 ASs por AS-Path

Estado atual da malha BGP no Brasil

- **Alguns dos principais pontos de troca de tráfego atualmente:**
 - RSiX (POP-RS, 13+ participantes, 150 Mbps)**
 - Terremark (HP, 30+ participantes, 1,1 Gbps)**
 - PRIX (POP-PR, 9+ participantes, 50 Mbps)**
 - OptIX-LA (Optiglobe, ? participantes, ? Mbps)**

Agenda

- Estado atual da malha BGP no Brasil
- Para que medir?
- Como, onde e o que medir?
- Proposta de uma plataforma de medições
- Iniciativas similares
- Aplicação exemplo
- Perguntas e debate

Para que medir?

- **Para lembrar: anúncios feitos com BGP dizem como os pacotes devem ingressar no AS**
Por onde e se eles chegam é outra história...
- **Então, após injetar um prefixo na malha:**
Sei qual a abrangência do meu anúncio?
Como o AS123 está chegando em meus prefixos?
Meus anúncios mais específicos estão sendo filtrados? Aonde?

Para que medir?

- **Como descobrir essas informações?**
 - bgp@noc...**
 - Telefone?**
 - iNOC BR?**
- **Automação: Looking Glass**
- **Desses ~200 ASs ativos:**
 - Quantos têm sistema de LG público?**
 - Quantos sabem o que é para que serve um LG?**
- **Maiores *carriers* hoje não possuem LG público, nem para clientes (EBT, Intelig, BrT)**

Para que medir?

- **Medir permite manter um histórico do estado da malha de roteamento**
 - Mudanças em atributos**
 - Troca de AS-Paths**
 - Correlação de eventos com mudanças na malha**
 - Detecção de falhas de implementação**
 - Etc...**
- **“Como meus anúncios eram vistos pelo AS123 no dia 20/3/2004 às 16h30?”**

Agenda

- Estado atual da malha BGP no Brasil
- Para que medir?
- Como, onde e o que medir?
- Proposta de uma plataforma de medições
- Iniciativas similares
- Aplicação exemplo
- Perguntas e debate

Como, onde e o que medir?

- **Existem várias abordagens para a medição**
 - SNMP (MIB-BGP)**
 - Somente tabela BGP
 - Scripts interativos (Tcl/Expect)**
 - Somente tabela BGP
 - Captura de pacotes (*Sniffers*)**
 - Somente mensagens BGP trocadas
 - Debug (Cisco, ???)**
 - Somente mensagens BGP trocadas
 - Participação ativa na malha**
 - Tabela BGP + Mensagens BGP


Como, onde e o que medir?

- **Mais completa: participação ativa na malha**
Mensagens BGP + Tabela BGP
- **Características da abordagem:**
Carga virtualmente nula na malha
Abordagem menos intrusiva
Estação de medição mantém conexão iBGP ou eBGP *multihop* com um ou mais roteadores do núcleo do AS (inclusive com RR)
Decisões sobre rotas aceitas feitas localmente
Possibilidade de funcionar como *Routing Beacon*
Dados coletados podem ser mantidos em histórico

Como, onde e o que medir?

- **Dados coletados exportados em *raw dump files***
 - Sistema de arquivos organizado cronologicamente
 - Pode ser exportado via NFS, HTTP, FTP, rsync, etc...
 - Viabiliza a medição com *probes* distribuídos
- **Vários processos independentes podem trabalhar com os *dumps***
 - Alimentar BDs
 - Extrair estatísticas (RRDtool)
 - Procurar por anomalias
 - Etc...

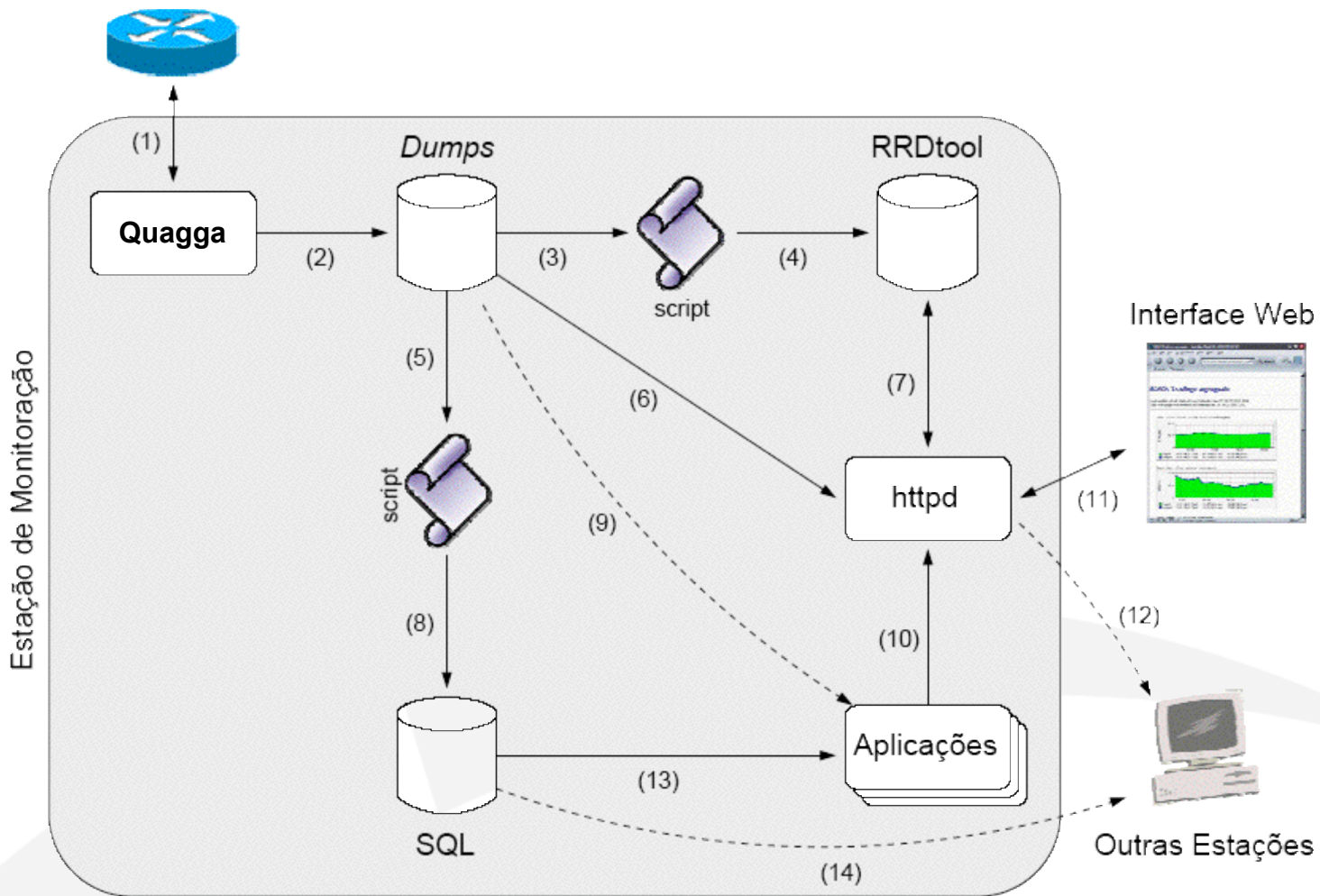
Agenda

- Estado atual da malha BGP no Brasil
- Para que medir?
- Como, onde e o que medir?
-  Proposta de uma plataforma de medições
- Iniciativas similares
- Aplicação exemplo
- Perguntas e debate

Proposta de uma plataforma de medições

- **Participação ativa na malha**
- **Estação com o software Quagga**
Evolução totalmente *free* do Zebra
Patch para geração dos *dumps* com horário certo em disco
<http://www.quagga.net/>
- **Estação entra na malha iBGP (com ou sem RR)**
- **Geração de *dumps* em intervalos programados**
Dumps são processados e vão para o BD
Raw dumps (formato MRT) ficam no sistema de arquivos

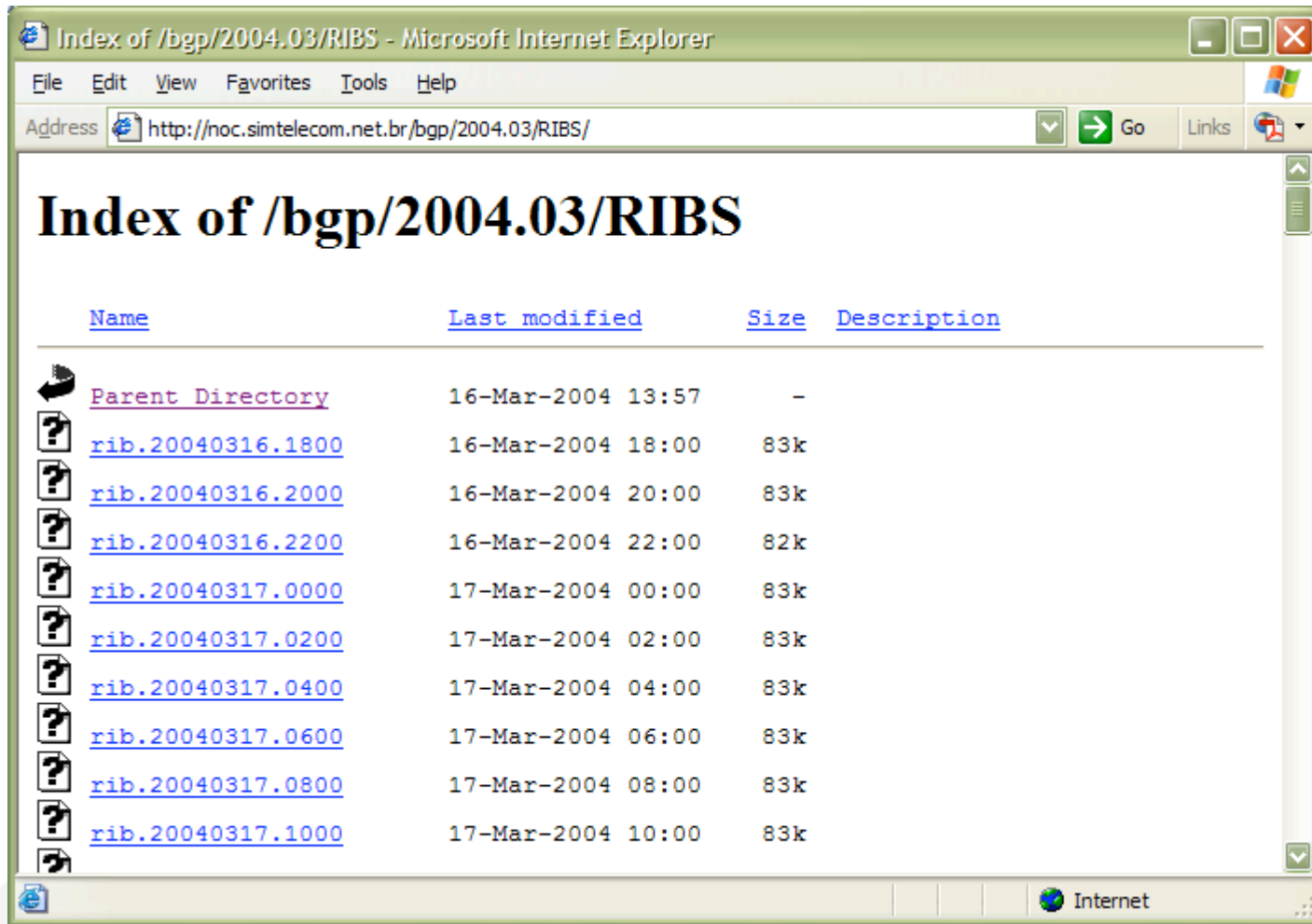
Proposta de uma plataforma de medições













Quagga: configuração

```
!  
hostname netman  
password 8 qKUmNUIkWUfI  
enable password 8 Edsk.39K8DeI  
service advanced-vty  
service password-encryption  
!  
router bgp 26622  
  bgp router-id 200.143.64.162  
  neighbor 200.143.64.193 remote-as 26622  
  neighbor 200.143.64.193 route-map nothing out  
!  
route-map nothing deny 1  
!  
dump bgp updates /data/bgp/%Y.%m/UPDATES/updates.%Y%m%d.%H%M 15m  
dump bgp routes-mrt /data/bgp/%Y.%m/RIBS/rib.%Y%m%d.%H%M 2h  
line vty  
  exec-timeout 0 0  
!
```

Estrutura dos dumps em disco



Index of /bgp/2004.03/RIBS

<u>Name</u>	<u>Last modified</u>	<u>Size</u>	<u>Description</u>
 Parent Directory	16-Mar-2004 13:57	-	
 rib.20040316.1800	16-Mar-2004 18:00	83k	
 rib.20040316.2000	16-Mar-2004 20:00	83k	
 rib.20040316.2200	16-Mar-2004 22:00	82k	
 rib.20040317.0000	17-Mar-2004 00:00	83k	
 rib.20040317.0200	17-Mar-2004 02:00	83k	
 rib.20040317.0400	17-Mar-2004 04:00	83k	
 rib.20040317.0600	17-Mar-2004 06:00	83k	
 rib.20040317.0800	17-Mar-2004 08:00	83k	
 rib.20040317.1000	17-Mar-2004 10:00	83k	

Banco de dados

Messages

Seq : MEDIUMINT UNSIGNED
Time : TIMESTAMP(14)
Type : ENUM('A','W')
PeerAddr : CHAR(15)
PeerAS : SMALLINT UNSIGNED
PrefixAddr : CHAR(15)
PrefixLen : TINYINT UNSIGNED
ASPath : TEXT
Origin : ENUM('?','I','E')
NextHop : CHAR(15)
LocalPref : SMALLINT UNSIGNED
MED : SMALLINT UNSIGNED
Community : TEXT

Tables

Seq : MEDIUMINT UNSIGNED
Time : TIMESTAMP(14)
FromAddr : CHAR(15)
FromAS : SMALLINT UNSIGNED
PrefixAddr : CHAR(15)
PrefixLen : TINYINT UNSIGNED
ASPath : TEXT
Origin : ENUM('?','I','E')
NextHop : CHAR(15)
LocalPref : SMALLINT UNSIGNED
MED : SMALLINT UNSIGNED
Community : TEXT
Agg : ENUM('Y','N')
AggAddr : VARCHAR(15)
AggAS : SMALLINT UNSIGNED

Agenda

- Estado atual da malha BGP no Brasil
- Para que medir?
- Como, onde e o que medir?
- Proposta de uma plataforma de medições
- Iniciativas similares
- Aplicação exemplo
- Perguntas e debate

Iniciativas similares

- **Route Views – University of Oregon**

Coleta de *dumps*: *scripts* interativos e participação na malha

Probes também funcionam como LG: Juniper, Zebra, Cisco, ...

Routing Beacons (bgpd.pl)

<http://www.routeviews.org/>

<ftp://archive.routeviews.org/> (dumps)

- **RIPE NCC – Routing Information Service (RIS)**

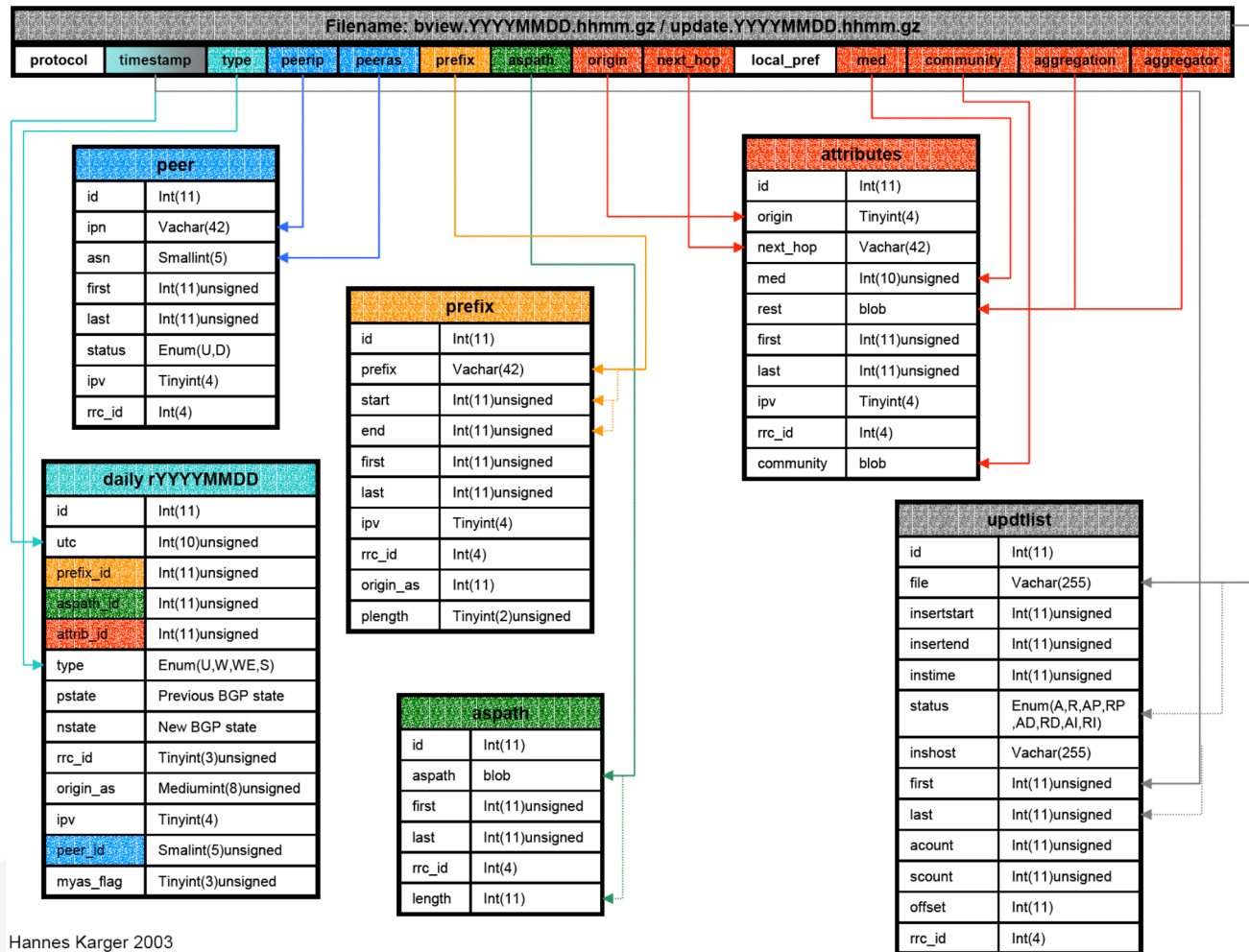
Coleta de dumps com participação na malha

Dumps vão para um BD MySQL de onde consultas são feitas

Interface web para consulta de valores (históricos inclusive)

<http://ris.ripe.net/>

RIS – Estrutura do BD



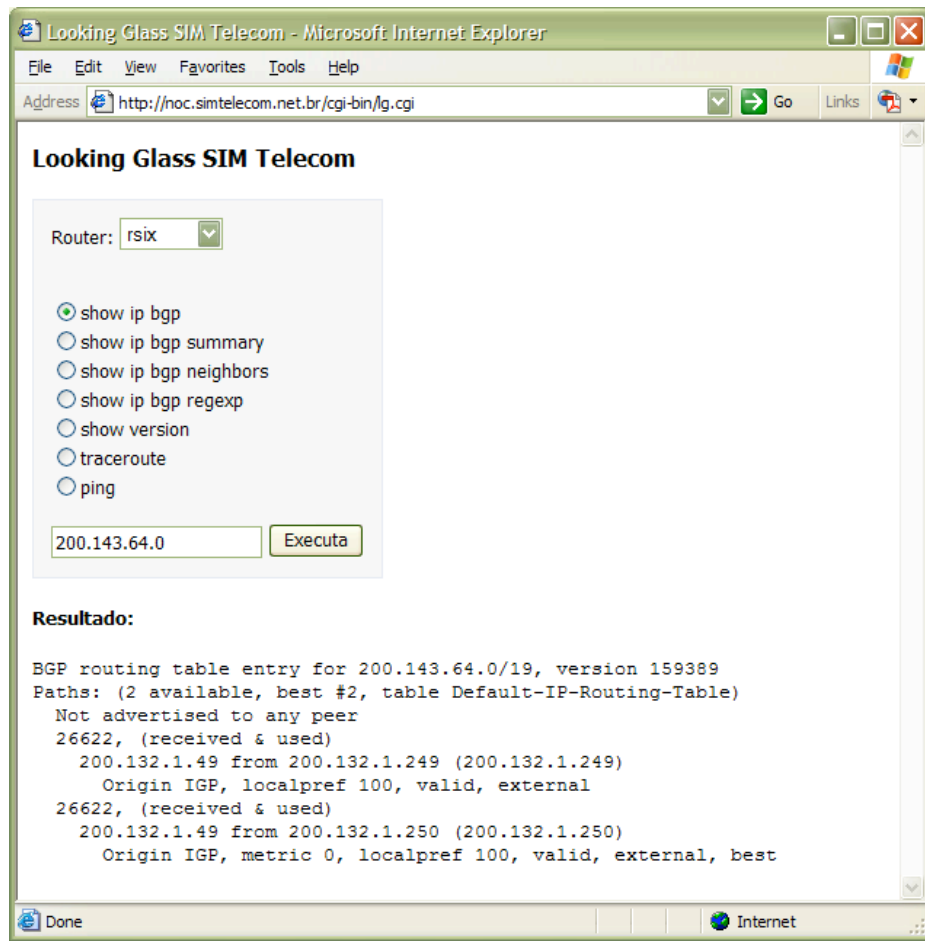
Agenda

- Estado atual da malha BGP no Brasil
- Para que medir?
- Como, onde e o que medir?
- Proposta de uma plataforma de medições
- Iniciativas similares
- Aplicação exemplo
- Perguntas e debate

Looking glass

- **Interface web feita em Perl (CGI)**
- **Consultas (ainda) feitas em Tcl/Expect**
- **Arquivo de configuração flexível**
- **Consulta interativamente roteadores**
Cisco e Zebra/Quagga somente (por enquanto)
- **Consultas são facilmente estendidas**
show ip bgp
show ip bgp regexp
show ip bgp summary
traceroute
...

Looking glass



Configuração:

```
title[]: Looking Glass SIM Telecom
```

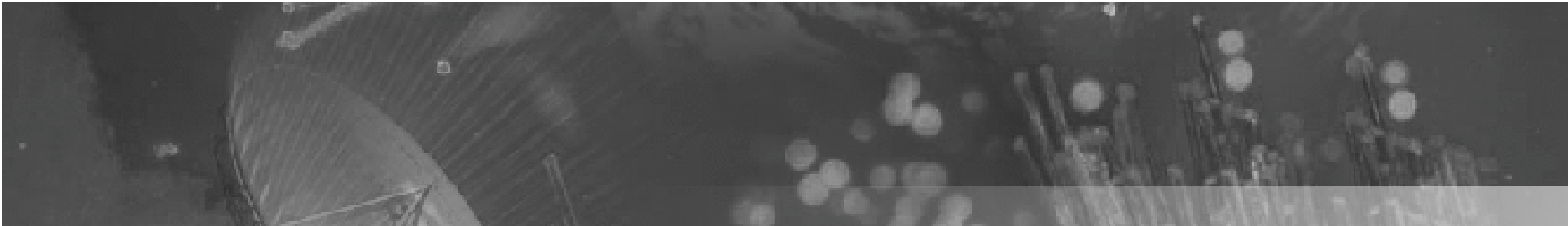
```
hostname[core01]: core01.pae.simtelecom.net.br
type[core01]: cisco
password[core01]: xxxx
prompt[core01]: core01>
```

```
hostname[netman]: netman.noc.simtelecom.net.br
type[netman]: zebra
port[netman]: 2605
password[netman]: xxxx
prompt[netman]: netman#
```

```
hostname[rsix]: lg.rsix.tche.br
type[rsix]: cisco
password[rsix]: xxxx
prompt[rsix]: looking-glass>
```

Agenda

- Estado atual da malha BGP no Brasil
- Para que medir?
- Como, onde e o que medir?
- Proposta de uma plataforma de medições
- Iniciativas similares
- Aplicação exemplo
- Perguntas e debate



Perguntas?