

Pontos de Troca de Tráfego na Internet - Projeto PTTmetro SET 2007

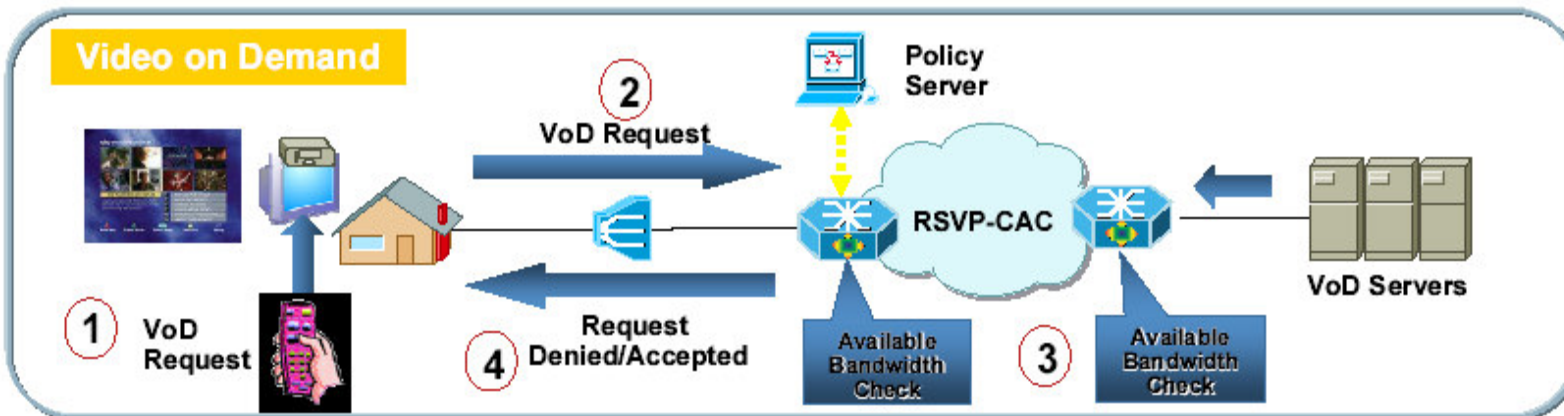
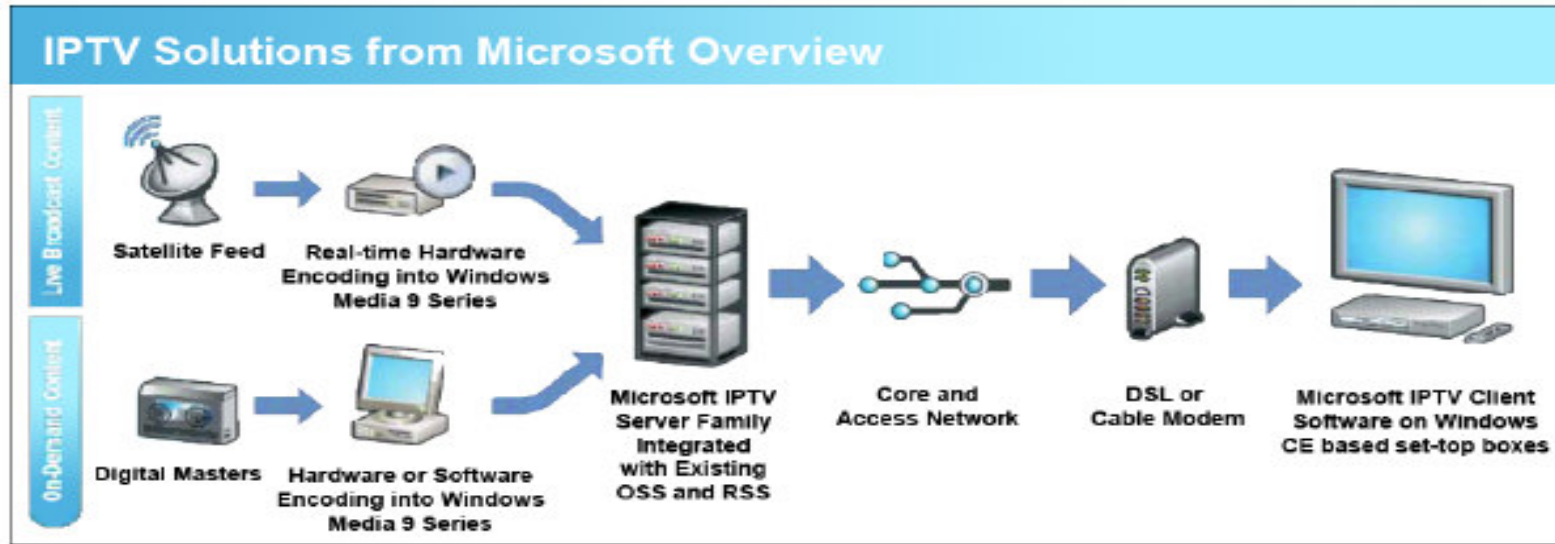
24/08/2007

Oripide Cilento Filho
(oripide@nic.br)

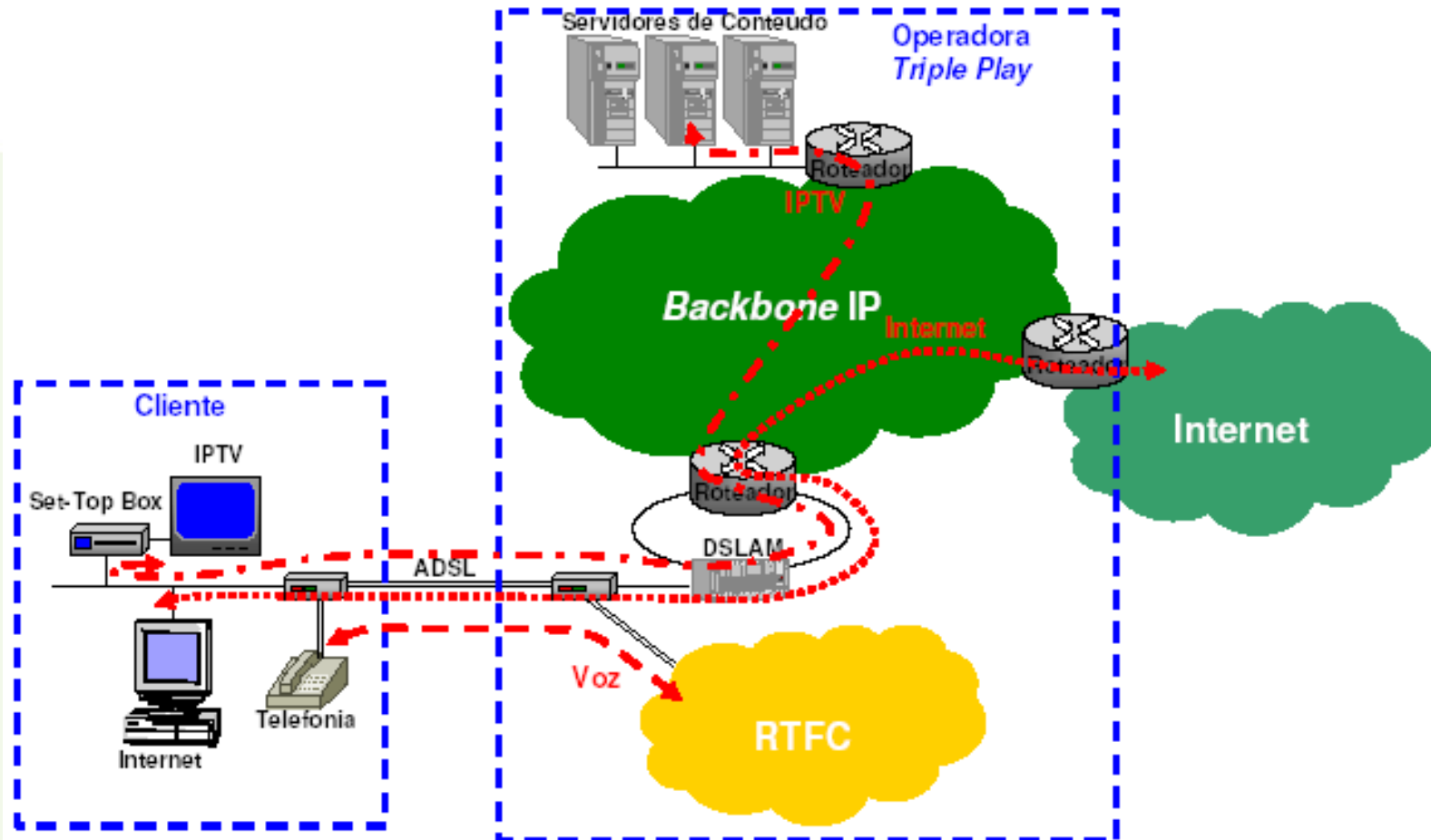
Agenda

- IPTV - Arquitetura
- Sistemas Autônomos
- Princípios do protocolo BGP
- Internet - Arquitetura
- Trânsito e *Peering*
- Motivação para implantação de PTTs
- Detalhes técnicos do projeto PTTmetro
- Regras de adesão
- Utilização do PTTmetro na distribuição de IPTV

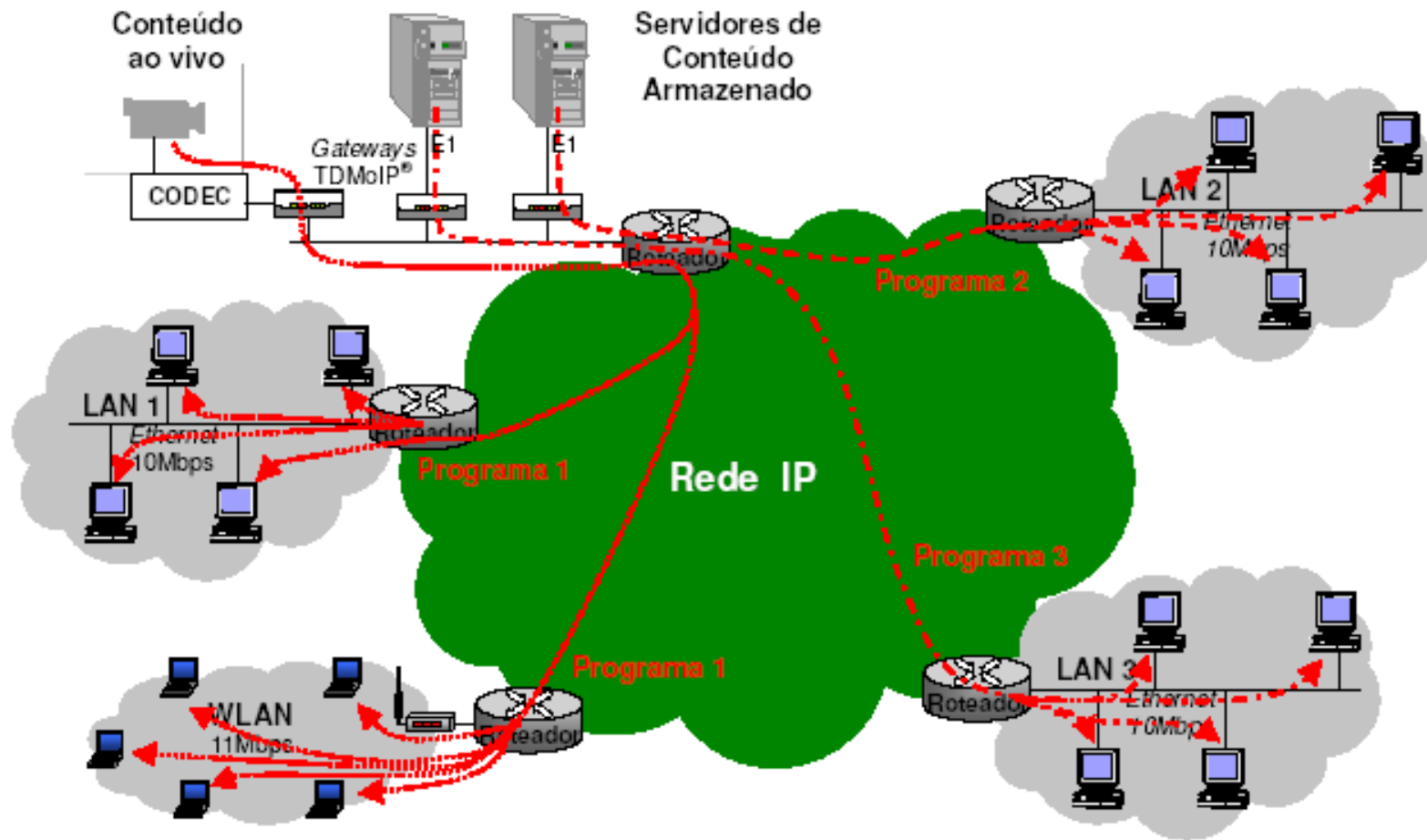
IPTV Arquiteturas



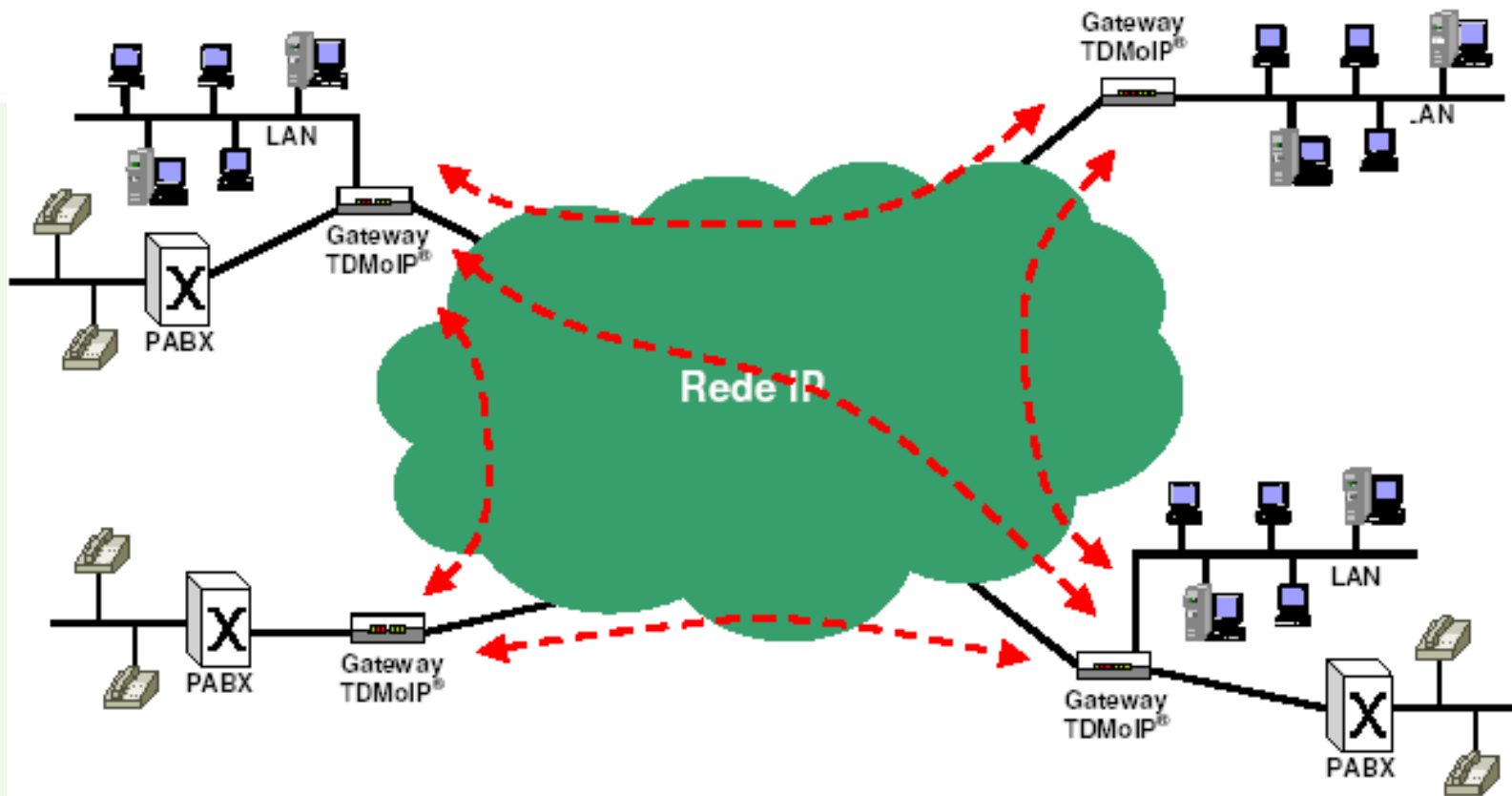
IPTV numa configuração *Triple Play*



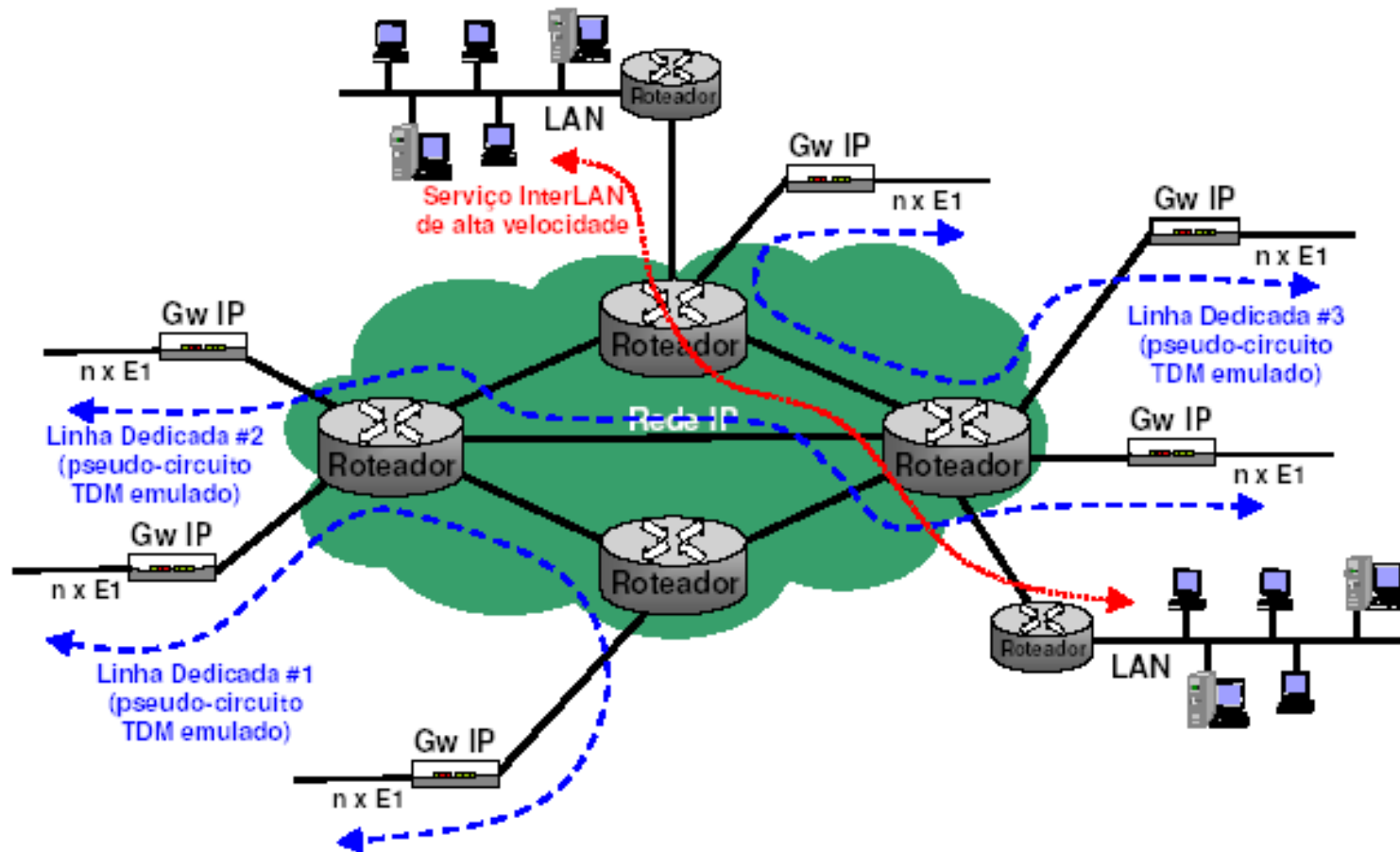
Broadcasting IPTV



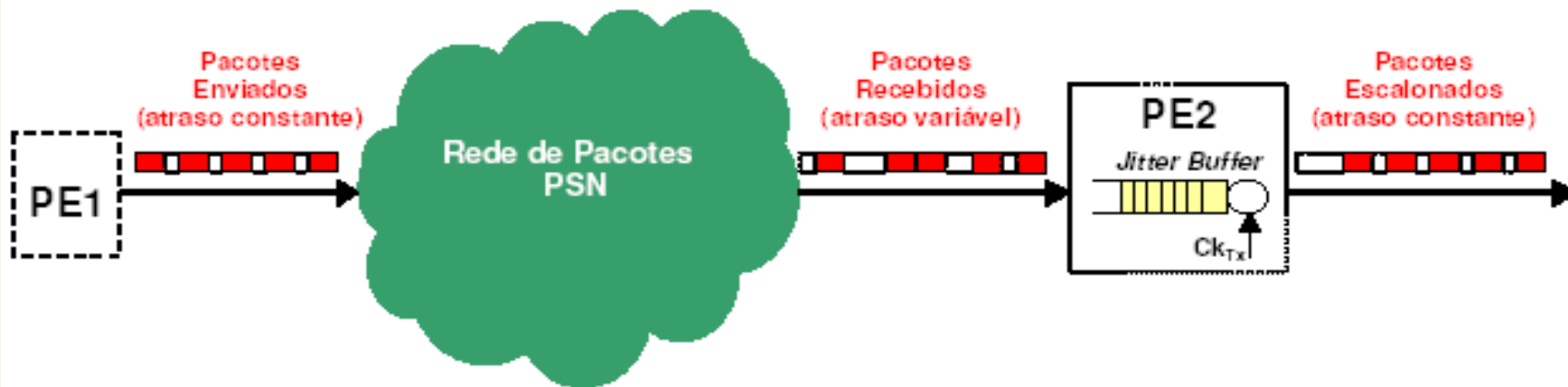
VoIP e Telefonia IP



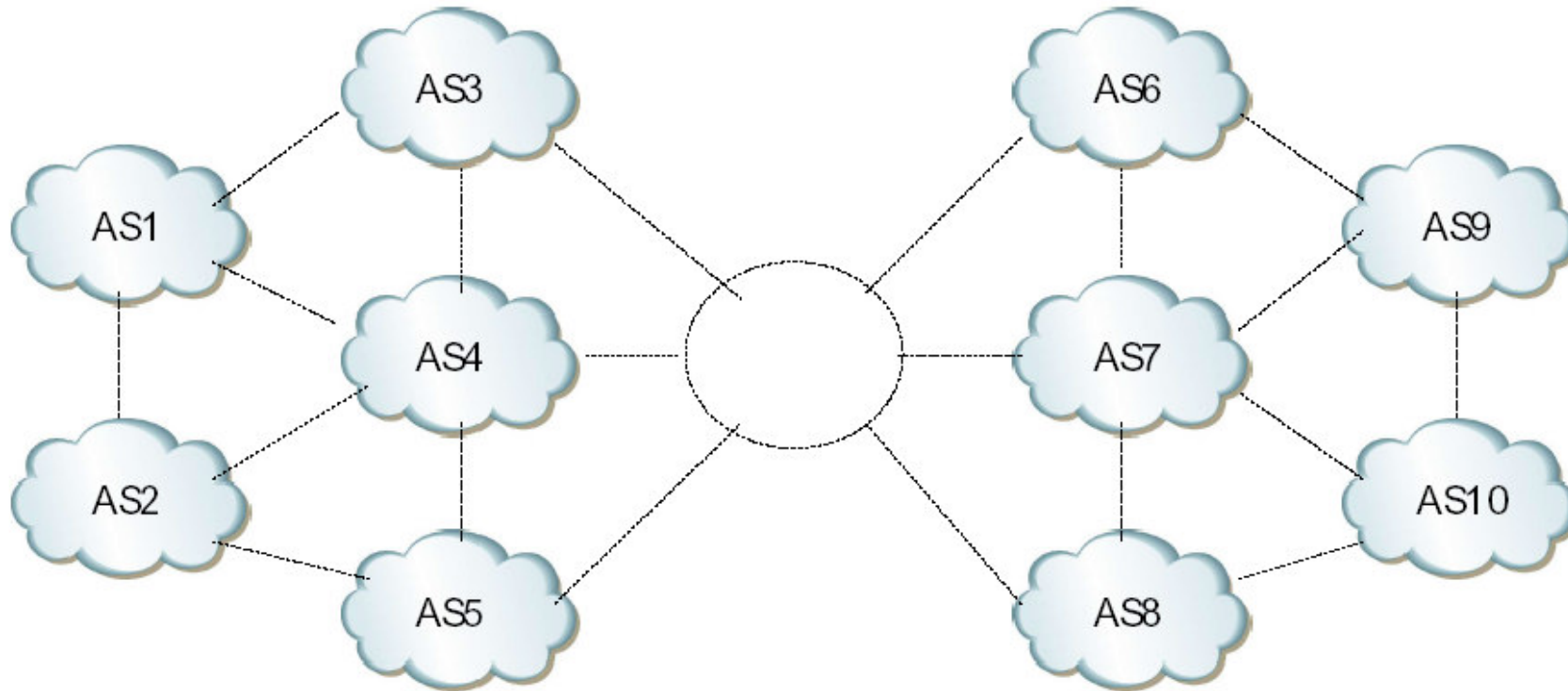
Detalhe da rede IP



Um dos problemas da Internet: Deficiência na qualidade de serviço na entrega dos pacotes



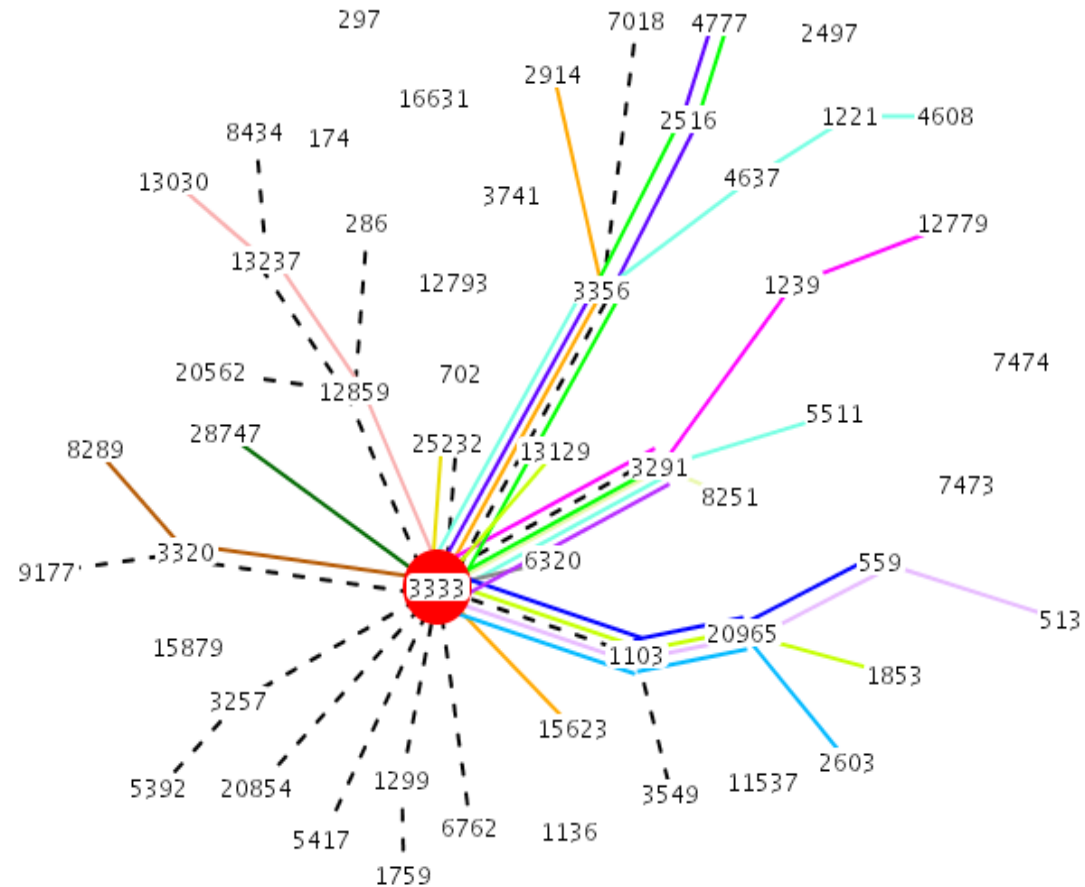
Visão da Internet com um conjunto de sistemas autônomos (AS) inter-conectados



AS (*Autonomous System*)

- De acordo com a RFC 1930 (*Requests for Comments*), um sistema Autônomo (AS) é um grupo de redes IP gerenciados por uma ou mais operadoras de rede, que compartilham uma única e mesma política de roteamento.

Mapa de Topologia - Nivel AS exemplo

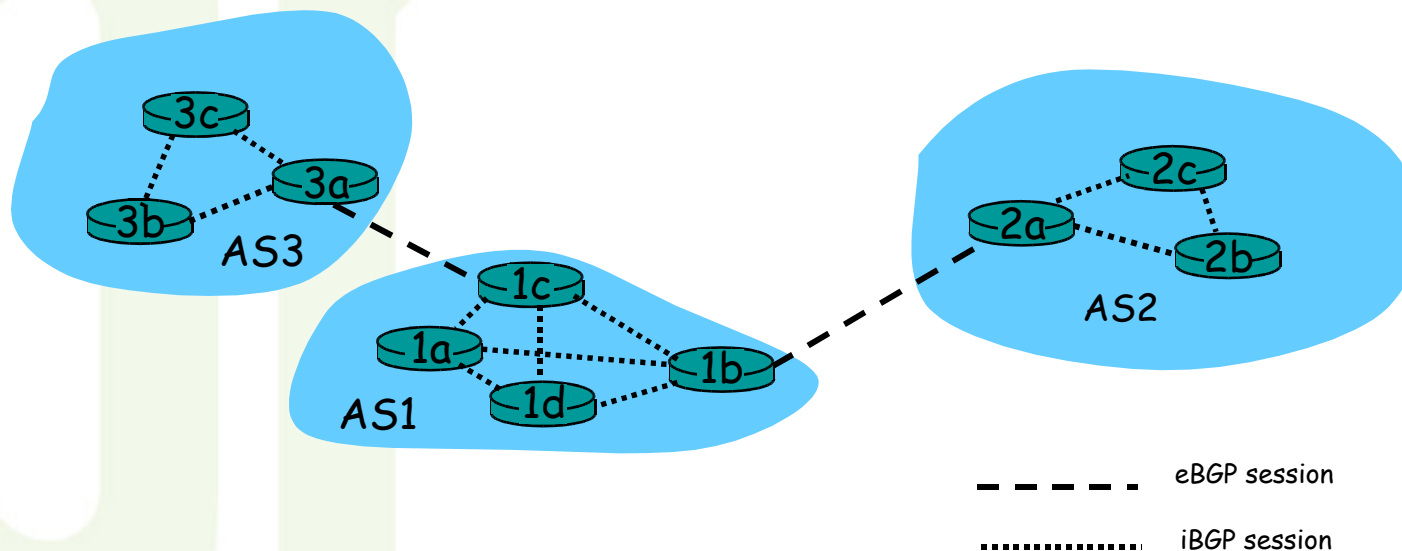


Roteamento inter-AS na Internet: BGP

- **BGP (*Border Gateway Protocol*)**: o padrão “de fato”
- BGP provê para cada AS meios de:
 - Obter informação de alcance de sub-redes (prefixos) a partir de ASs vizinhos.
 - Propagar informação de alcance para todos os roteadores internos ao AS .
 - Determinar “boas” rotas para sub-redes a partir de informação de alcance e políticas.
- Permite que uma sub-rede anuncie a sua existência para o resto da Internet: *“Estou aqui!”*

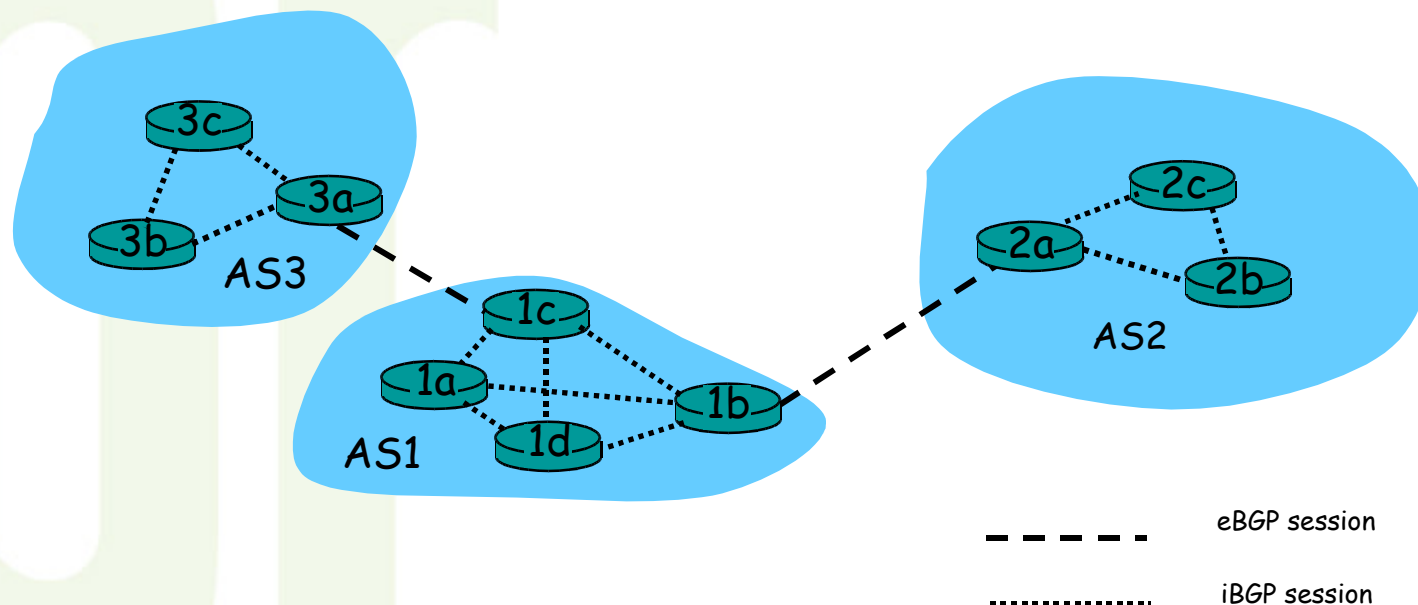
Fundamentos do BGP

- Par de roteadores (pares BGP) trocam informação de roteamento através de conexões TCP semi-permanentes::
sessões BGP
- Note que sessões BGP não correspondem a enlaces físicos.
- Quando um AS2 anuncia um prefixo para AS1, AS2 está **prometendo** que vai enviar à aquele prefixo quaisquer datagramas destinados ao mesmo.
 - AS2 pode agregar prefixos nos seus anúncios



Distribuindo informação de alcance

- Com a sessão eBGP 3a-para-1c, AS3 envia informação de alcance de prefixos para AS1.
- 1c pode usar iBGP ou outro protocolo interno para distribuir esta nova informação de alcance de prefixo para todos os roteadores em AS1.
- 1b pode então re-anunciar a nova informação de alcance para AS2 através da sessão eBGP 1b-para-2a.
- Quando um roteador aprende sobre um novo prefixo, ele cria uma entrada para o prefixo na sua tabela de encaminhamento.



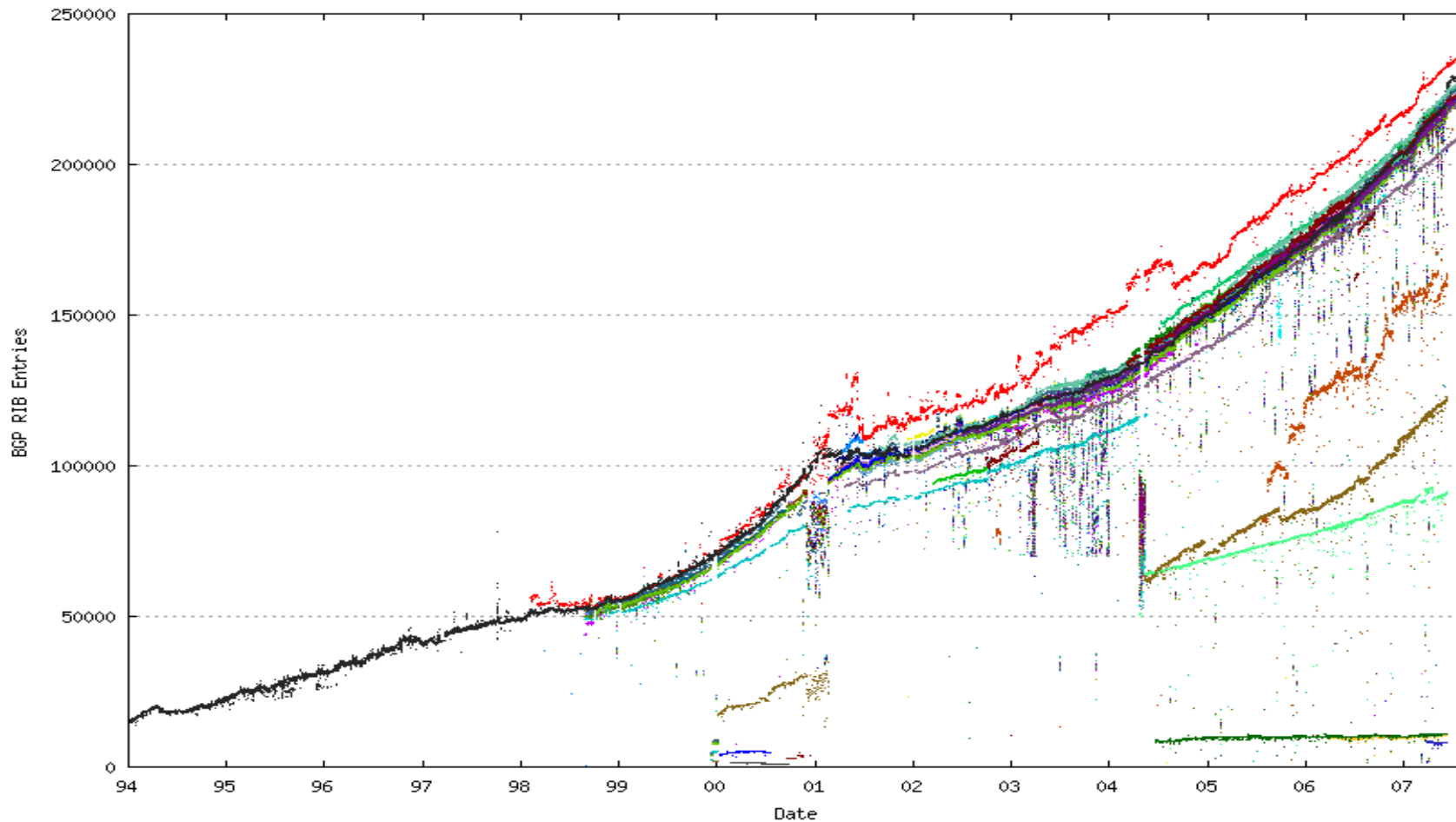
Atributos de caminho & rotas BGP

- Quando um prefixo é anunciado, o anúncio inclui atributos BGP.
 - prefixo + atributos = “rota”
- Dois atributos importantes:
 - **AS-PATH**: contém os ASs pelos quais o anúncio para o prefixo passou: AS 67 AS 17
 - **NEXT-HOP**: indica o roteador específico, interno ao AS, que leva ao AS do próximo salto. (Pode haver múltiplos enlaces do AS atual para o AS do próximo salto)
- Quando um roteador de borda recebe um anúncio de rota, usa a **política de importação** para aceitar/declinar.

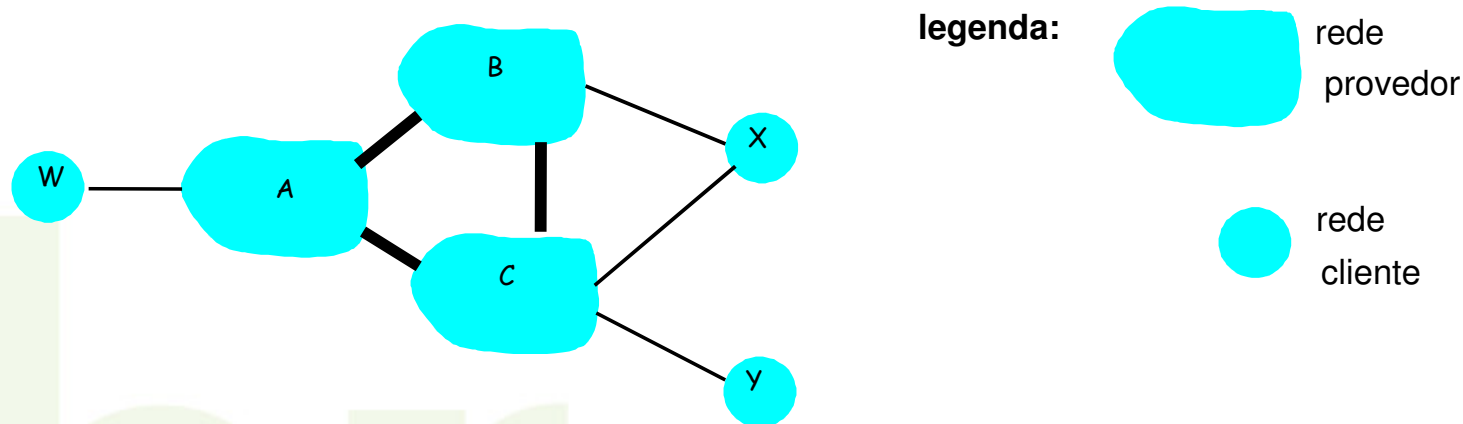
Seleção de rota do BGP

- Roteador pode aprender sobre mais de 1 rota para algum prefixo. Ele deve selecionar a rota.
- Regras de eliminação:
 - Valor do atributo *preferência local* associado à rota: decisão política
 - Menor AS-PATH
 - Roteador NEXT-HOP mais próximo: roteamento “batata quente”
 - Critérios adicionais

Evolução da tabela de roteamento na Internet



Políticas de roteamento BGP



- A,B,C são **redes de provedores**
- x,w,y são clientes (das redes de provedores)
- x com **duas interfaces** conectadas a duas redes
 - x não quer rotear de B para C
 - .. então x não vai anunciar para B a rota para C

Por quê há diferenças entre roteamento Intra- e Inter-AS?

Políticas:

- Inter-AS: administração quer controle sobre como tráfego roteado, quem transita através da sua rede.
- Intra-AS: administração única, logo são desnecessárias decisões políticas

Escalabilidade:

- roteamento hierárquico economiza tamanho de tabela de rotas, reduz tráfego de atualização

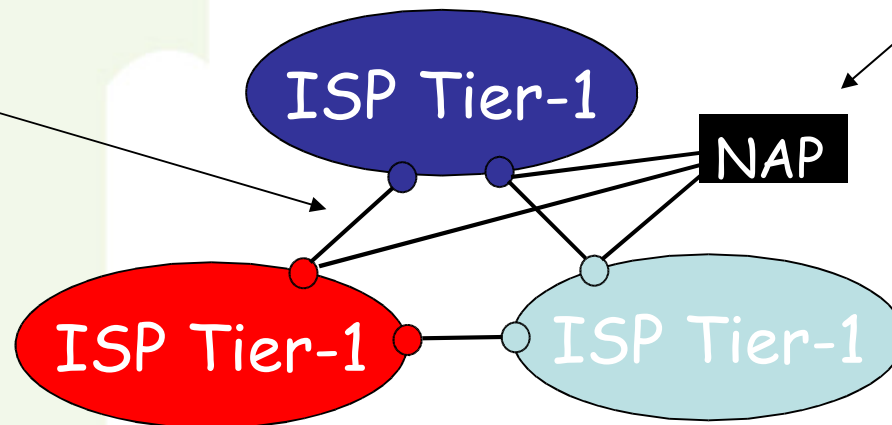
Desempenho:

- Intra-AS: pode focar em desempenho
- Inter-AS: políticas podem ser mais importantes do que desempenho

Arquitetura da Internet: rede de redes

- Parcialmente Hierárquica
- **No centro: ISPs de "tier-1"** (ex.: UUNet/MCI, BBN/Genuity, Sprint, AT&T, Global Crossing, Level 3, Cogent), cobertura nacional/international
 - não compram "transito" de ninguém

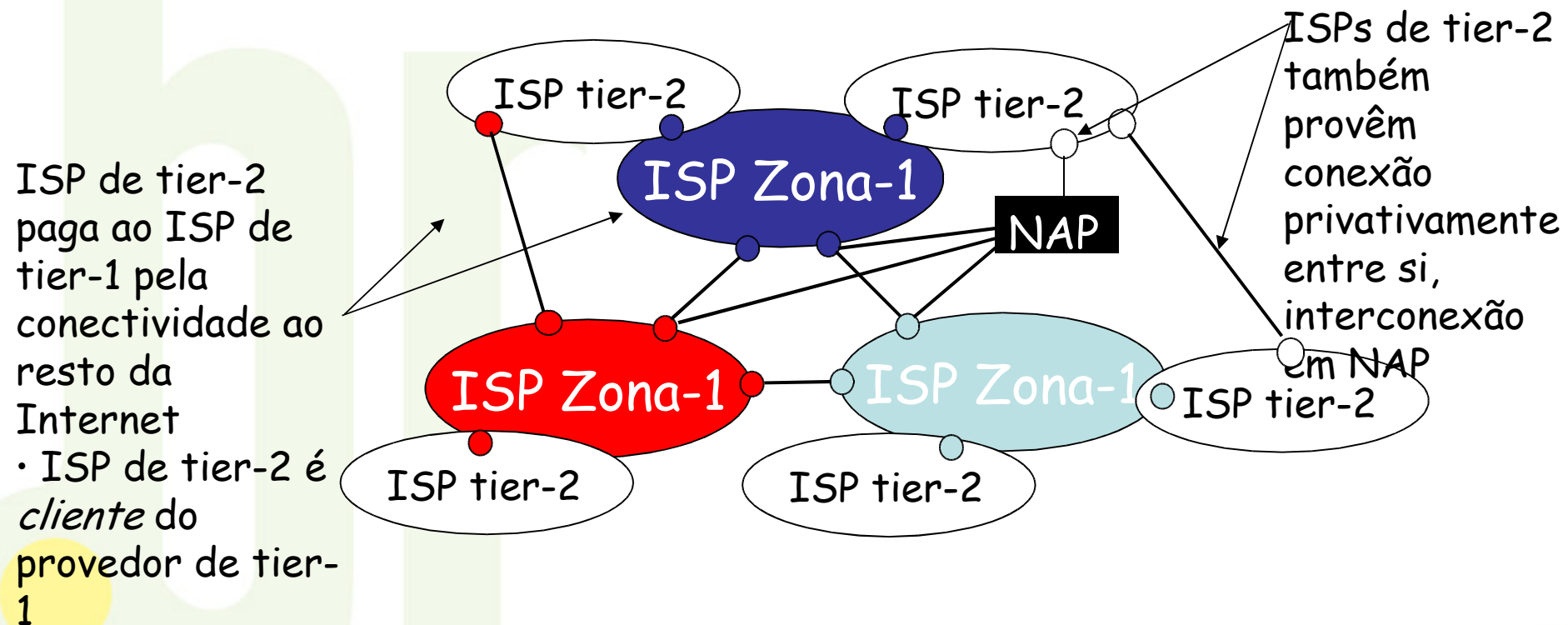
A "tier-1" provê interconexão (*peer*) de modo privativo



A "tier-1" também provê interconexão nos pontos de acesso (NAPs) da rede pública

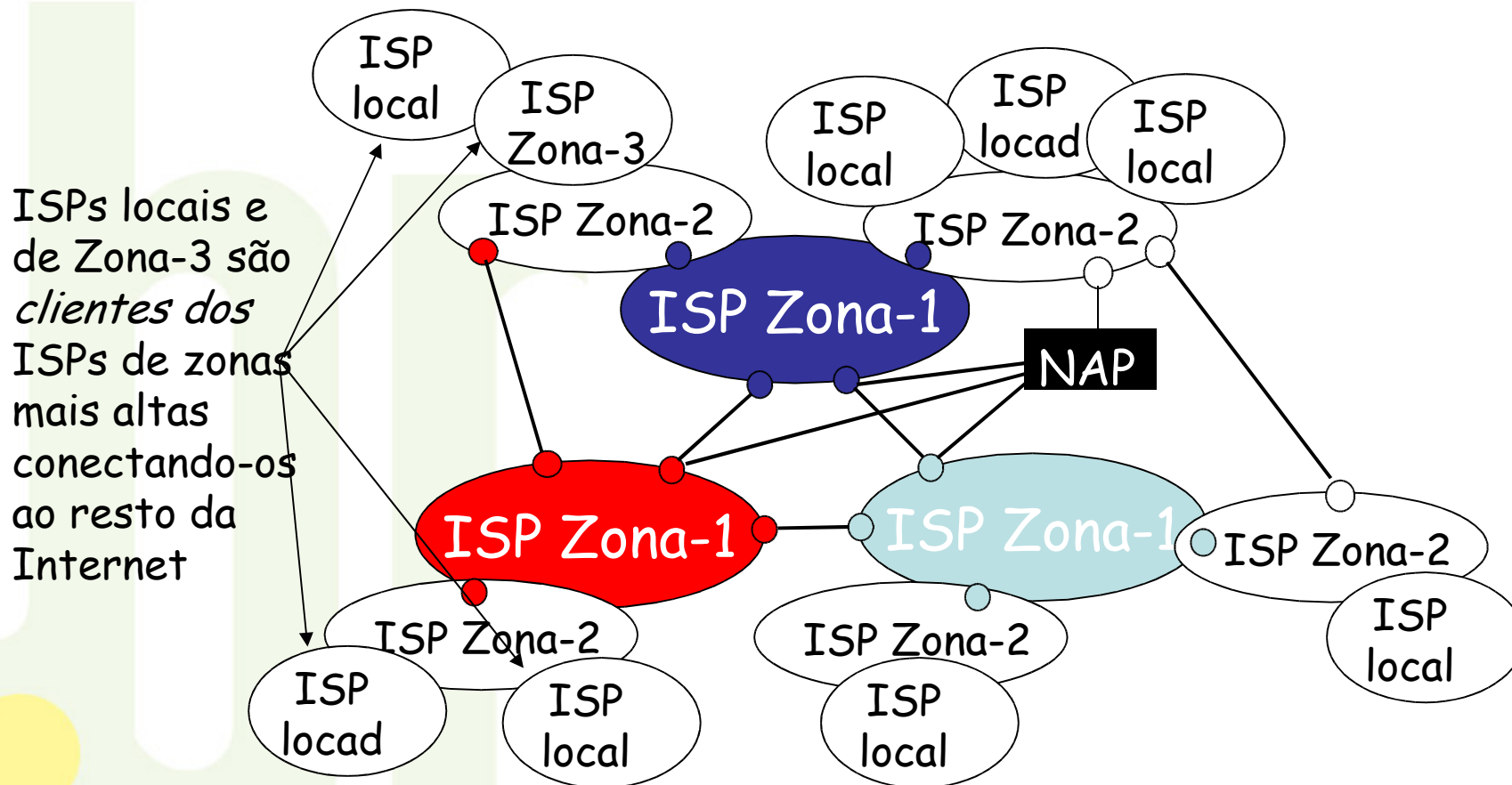
Estrutura da Internet: rede de redes

- **ISPs de "tier-2": ISPs menores (freqüentemente regionais)**
 - Conectam-se a um ou mais ISPs de tier-1, as vezes a outros ISPs de tier-2
 - Nossos principais ISPs (Embratel, Oi (Telemar), Brasil Telecom, Telefônica, Intelig, ...) estão nesta categoria



Estrutura da Internet: rede de redes

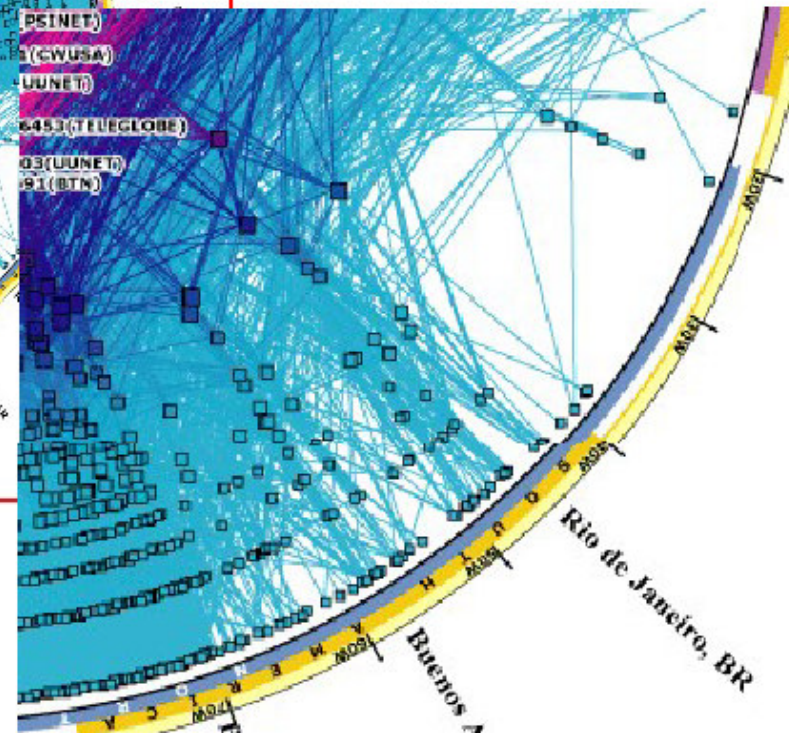
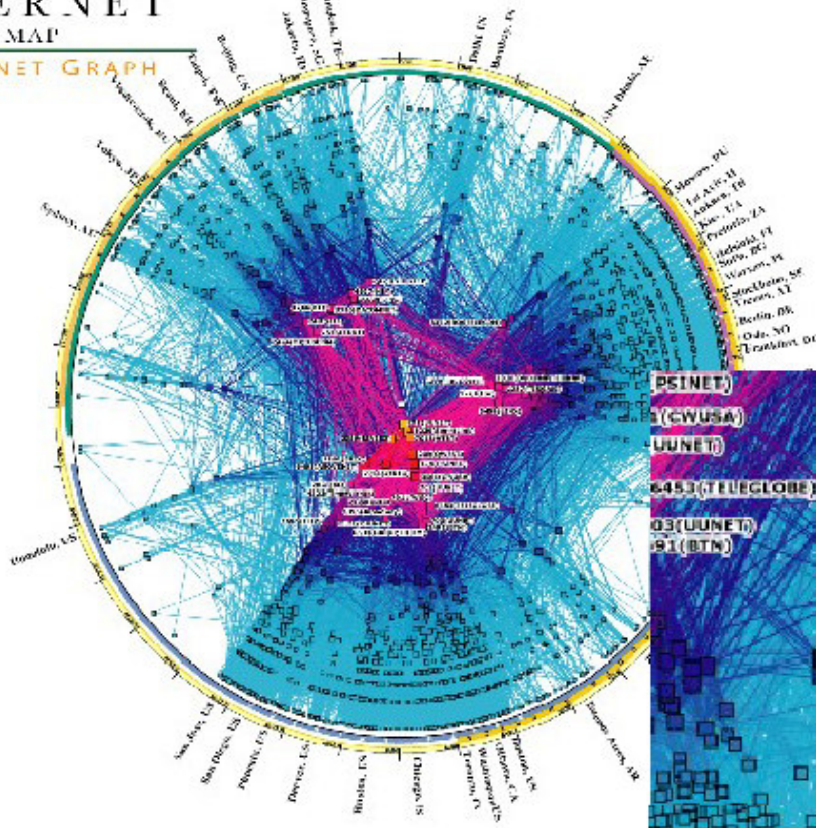
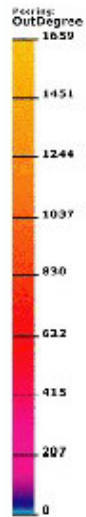
- ISPs de “Zona-3” e ISPs locais
 - Última rede de acesso (“hop”) (mais próxima dos sistemas finais)



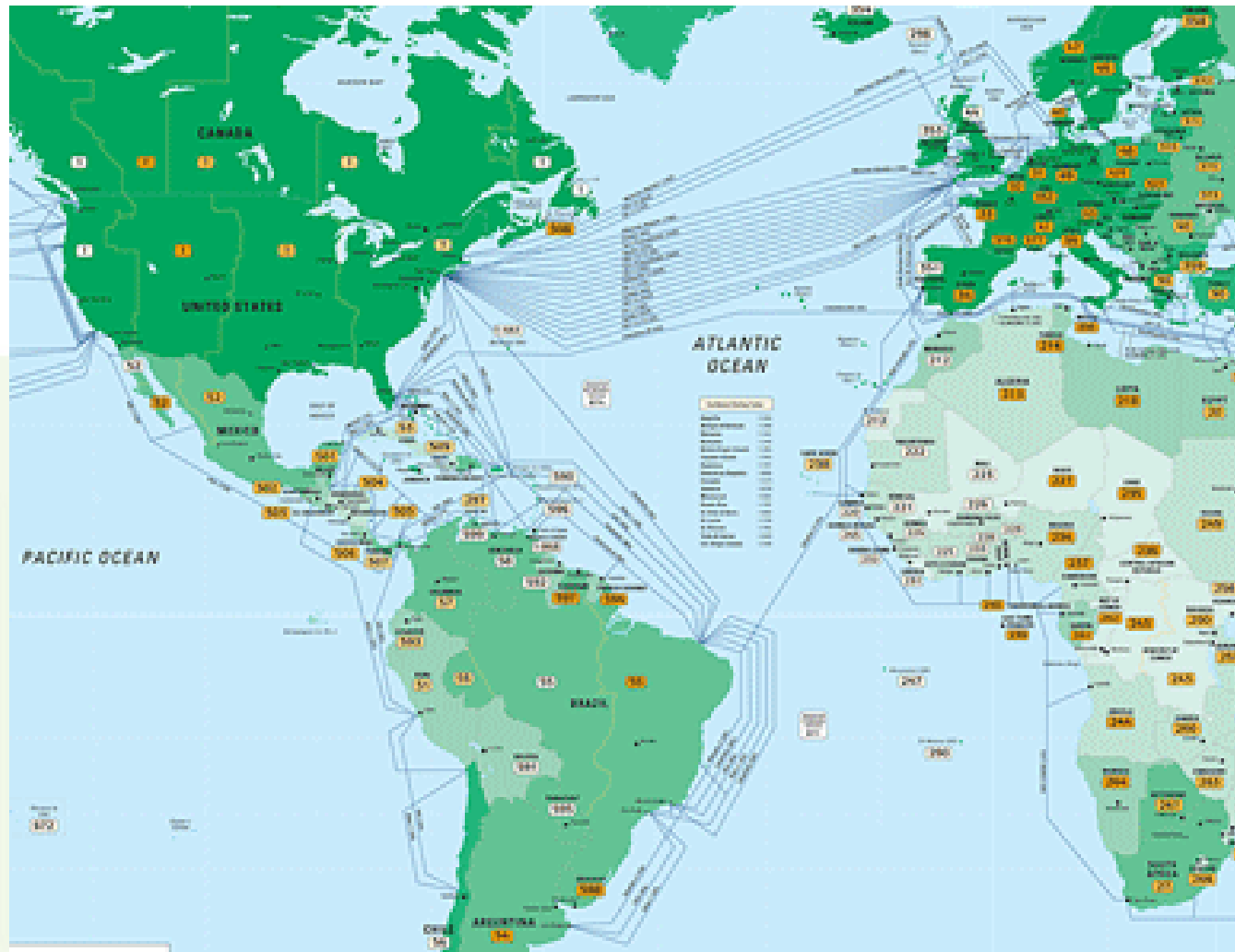
IPv4 INTERNET TOPOLOGY MAP

AS-level INTERNET GRAPH

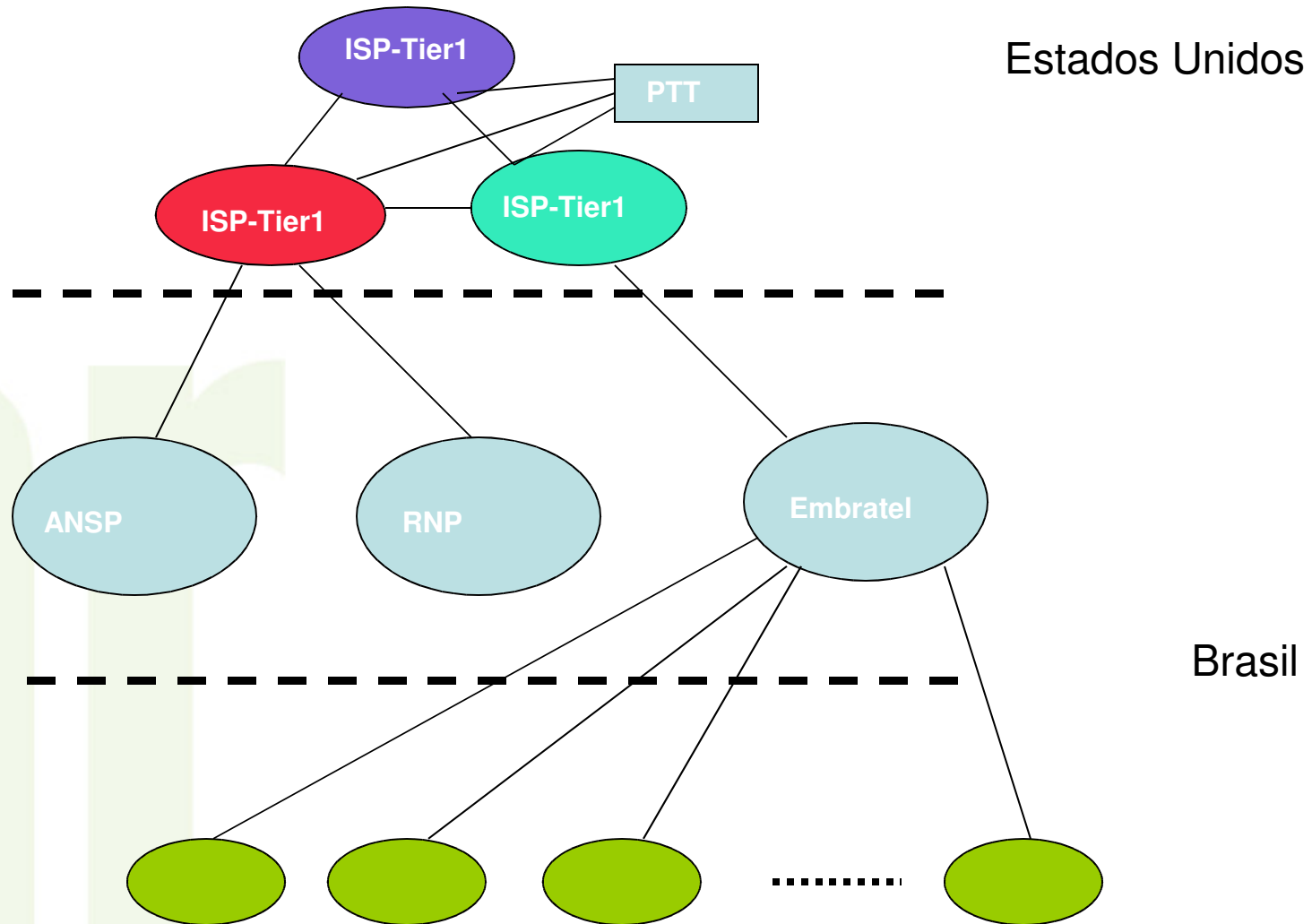
copyright ©2005 UC Regents. all rights reserved.



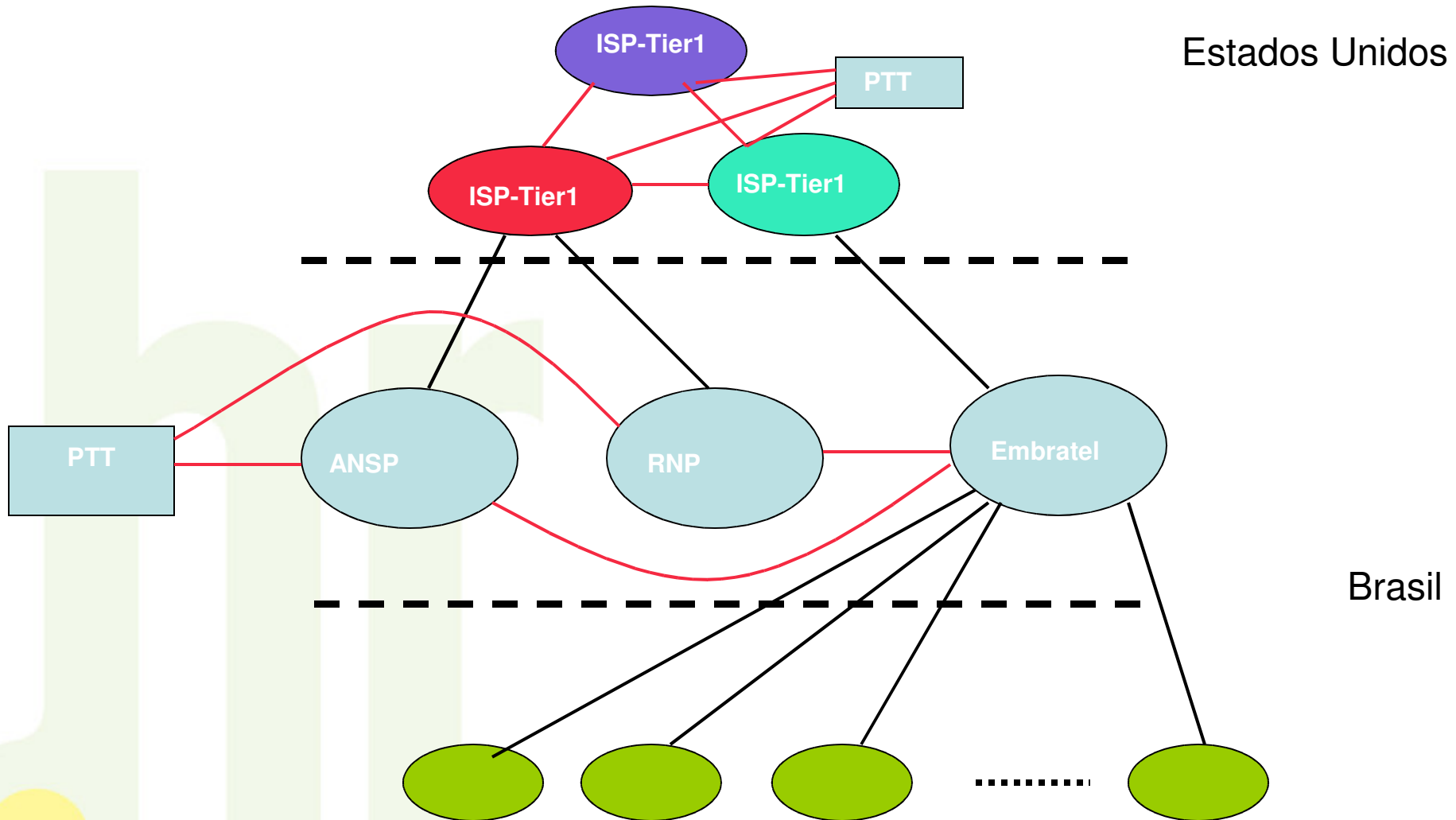
Motivação inicial: reduzir atraso e custo



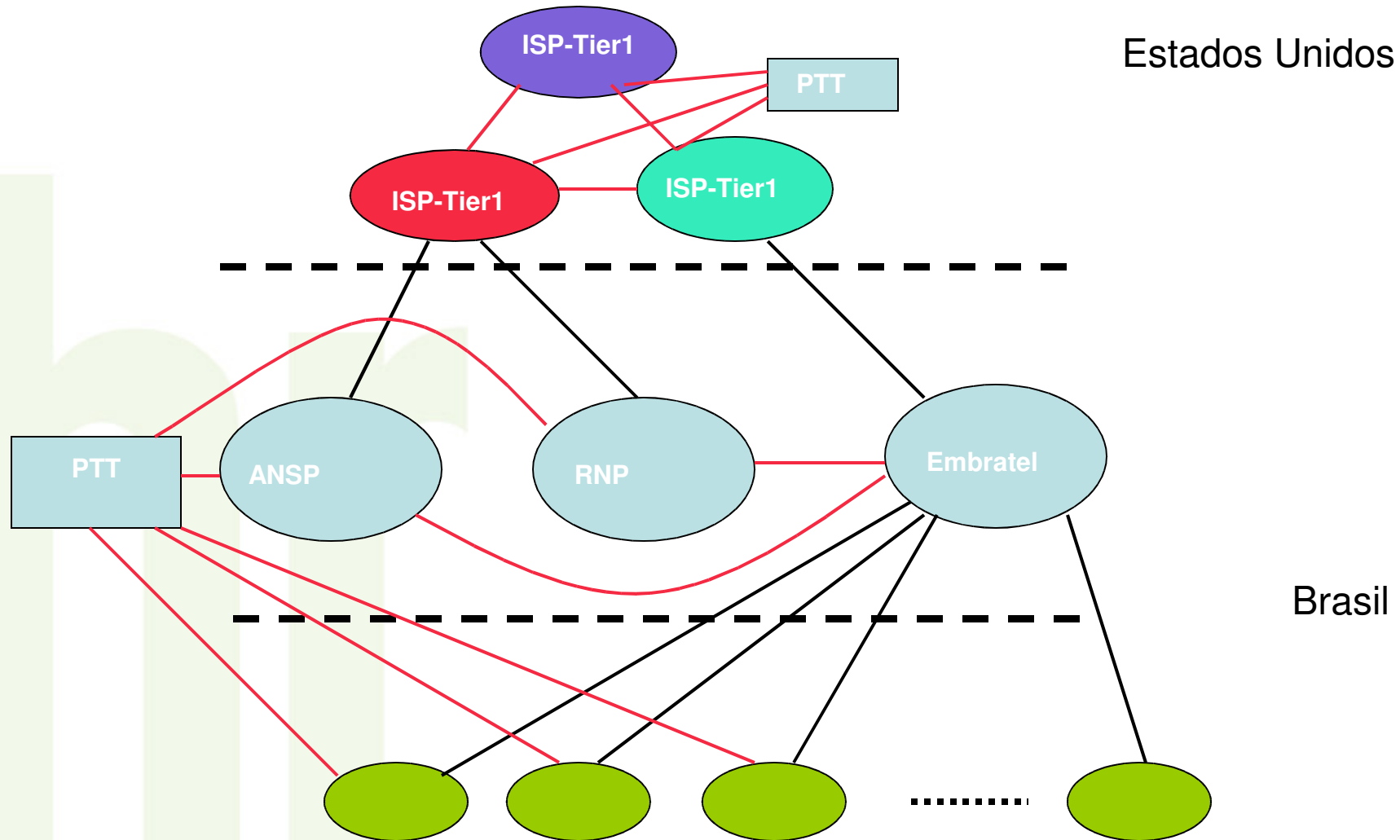
Motivação inicial: reduzir atraso e custo



Motivação inicial: reduzir atraso e custo



Motivação inicial: reduzir atraso e custo



Peering, de acordo com o IETF (Internet Engineering Task Force)

- “Negociações de arranjos de interconexão recíprocos, de livre estabelecimento ou outra maneira, entre provedores de serviço operacionalmente independentes.”

Tipos de acordos entre ISPs (*Internet Services Providers*)

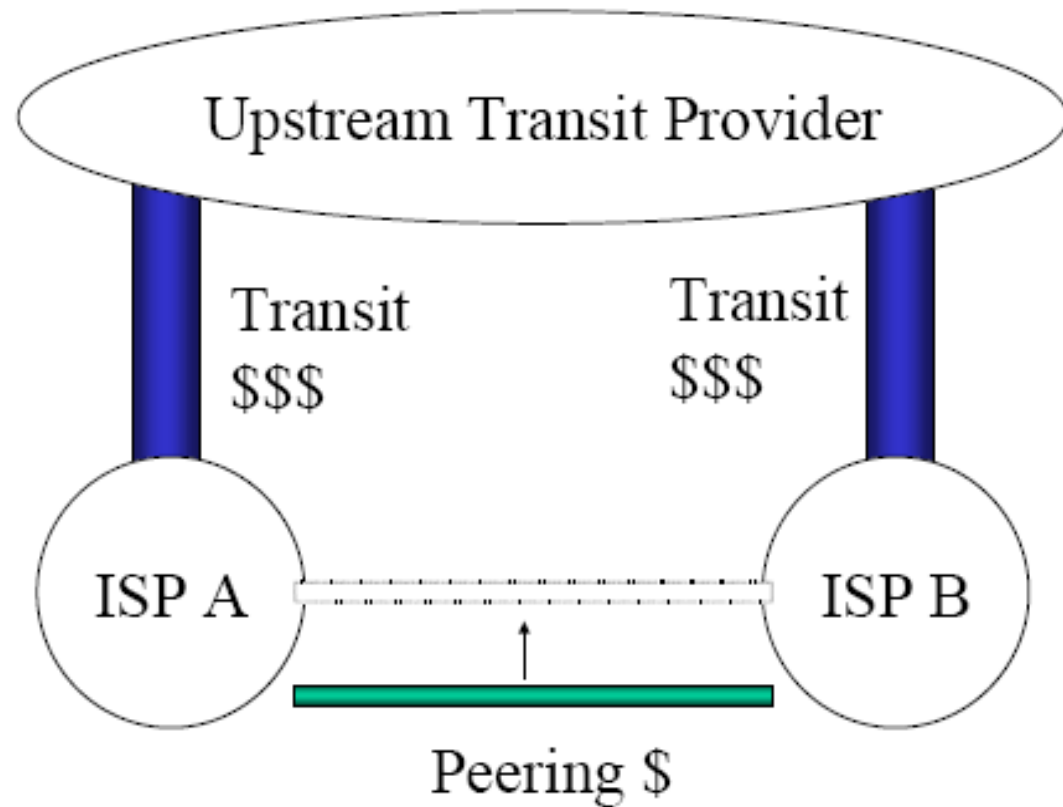
- *Peering* (parceria)
 - acordo entre dois AS
 - aplica-se apenas quanto o tráfego originado de um usuário do AS 1 destina-se ao usuário do AS 2
 - normalmente, não existem ajustes financeiros entre os participantes do *peering*. Os custos restringem-se aos equipamentos de cada um e ao custo de trazer o tráfego de um local ao outro.
 - Os parceiros podem combinar de trocar tráfego em diversas localidades geográficas. Adota-se, normalmente, o *hot potato routing mechanism*, que consiste em repassar o pacote ao parceiro tão logo possível

Tipos de acordos entre ISPs (*Internet Services Providers*)

- Trânsito

- é um acordo entre dois AS (1 e 2)
- quando o AS-1 oferece trânsito ao AS-2 os usuários do AS-2 poderão acessar os usuários de **todos os outros AS** com quem o AS-1 tem *peering*.
- um acordo de trânsito envolve, normalmente, o pagamento de taxa de transporte de AS-2 para AS-1
- neste acordo, todo o tráfego de requisitado pelos usuários de um AS passa “por dentro” da rede de um outro AS para alcançar outros usuários .
- Os “outros usuários” podem ser, no limite, toda a Internet !

Um ponto de Troca de Tráfego (PTT) possibilitou migrar tráfego de “trânsito” para “peering”



PTTMetro

- PTTMetro é o nome dado ao projeto do Comitê Gestor da Internet no Brasil (CGI.br) que promove e cria a infraestrutura necessária (Ponto de Troca de Tráfego - PTT) para a inter-conexão direta entre as redes (*Autonomous Systems* - ASs) que compõem a Internet Brasileira.
- A atuação do PTTMetro volta-se às regiões metropolitanas no país que apresentam grande interesse de troca de tráfego Internet comercial e acadêmica sob uma gerência centralizada.

PTTmetro - vantagens

- Racionalização dos custos, uma vez que os balanços de tráfego são resolvidos direta e localmente e não através de redes de terceiros, muitas vezes fisicamente distantes.
- Para o usuário final também ficaria mais barato vários serviços na Internet como hospedagem de site, por exemplo.
- Outra grande vantagem é o maior controle que uma rede pode ter com relação a entrega de seu tráfego o mais próximo possível do seu destino, o que em geral resulta em melhor desempenho e qualidade para seus clientes e operação mais eficiente da Internet como um todo.
- PTTMetro é um projeto aprovado pelo CGI.br e possui recursos constantes para a atualização tecnológica
- Tem por finalidade criar Pontos de Troca de Tráfego em regiões metropolitanas que tenham ISPs interessados em trocar tráfego Internet

Projeto PTTMetro

- PTTMetro tem por projeto as seguintes características:
 - Neutralidade
 - Qualidade
 - Baixo custo das alternativas, com alta disponibilidade
 - Matriz de troca de tráfego regional única
 - Melhoria na interconexão das redes
- PTT's em funcionamento / previstos:
 - **São Paulo**
 - **Brasília**
 - **Rio de Janeiro**
 - **Florianópolis**
 - Londrina
 - Fortaleza
 - **Curitiba**
 - **Porto Alegre**
 - **Belo Horizonte**
 - Salvador
 - Recife

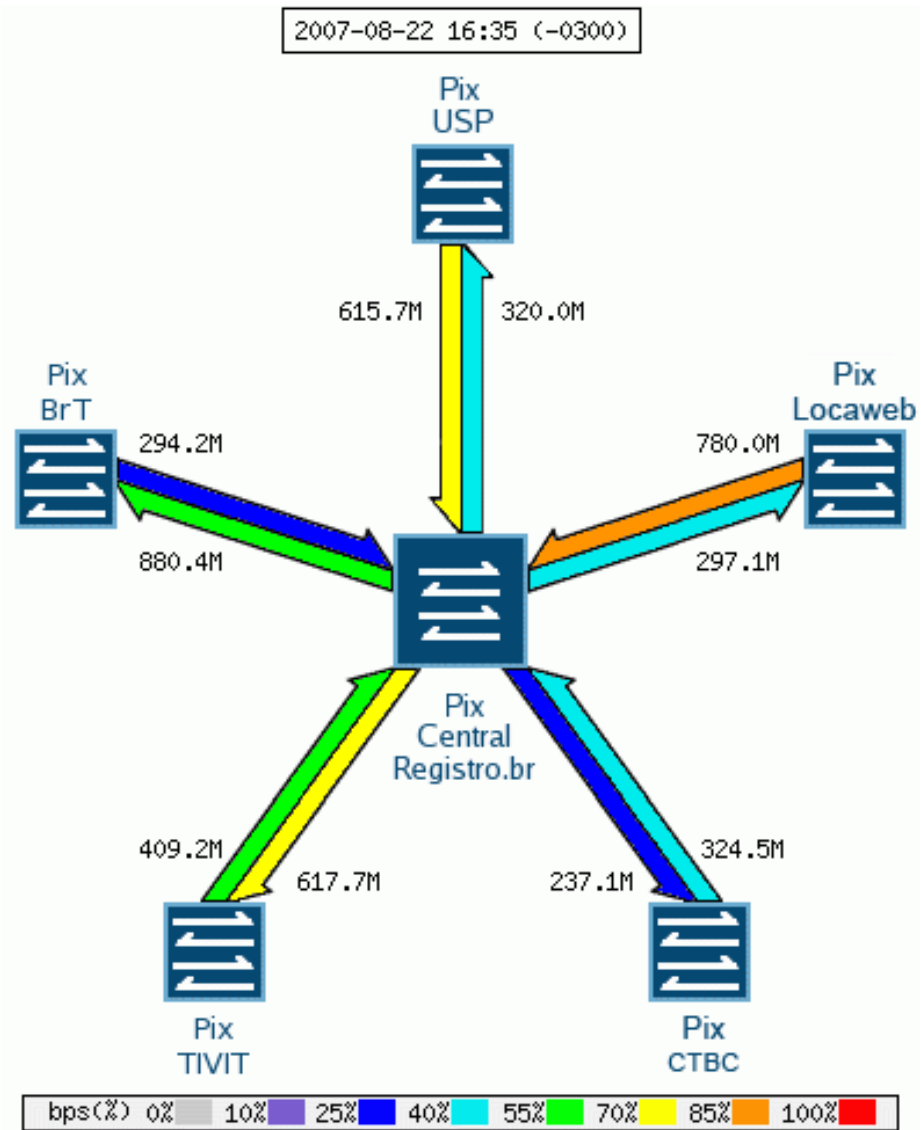
Projeto PTTMetro - Alguns números atuais

| Local | # Participantes | # PIXes | Tráfego (pico) Mbps |
|---------------|-----------------|---------|------------------------|
| São Paulo | 41 | 5 | 3.120 |
| Porto Alegre | 19 | 3 | 414 |
| Brasília | 7 | 2 | 217 |
| Curitiba | 13 | 2 | 360 |
| Florianópolis | 6 | 2 | 92 |
| Total | 86 | 14 | 4.200 |

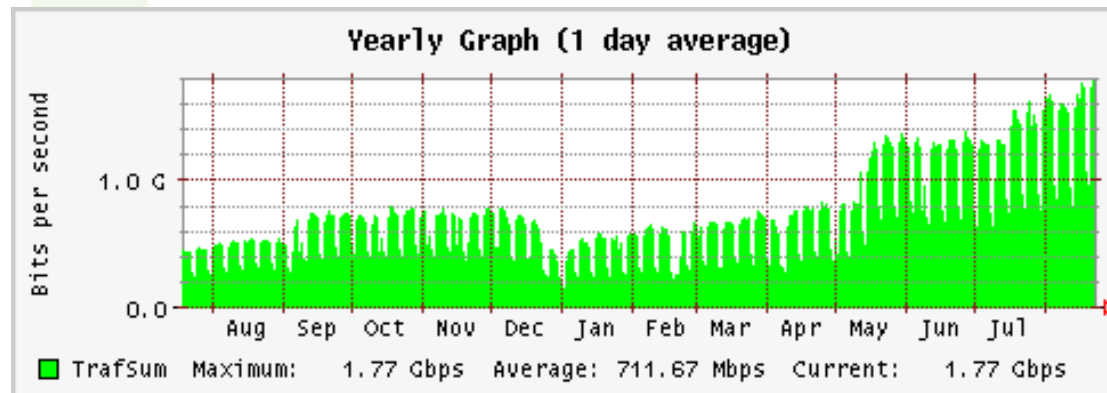
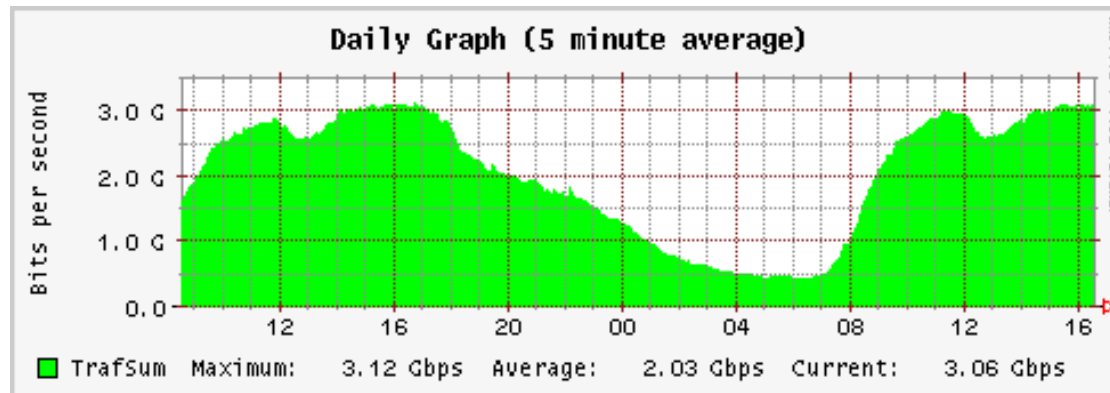
Participantes (São Paulo)

- 42WoodyNet
- 3856PCH
- 8055Value4Net
- 10954 Serpro
- 12956 TIWS
- 16685Optiglobe
- 19089DH&C
- 22548Registro.br
- 27664CTBC Multimidia
- 28572Apoiocom
- 28590Neovia
- 28669America-Net
- 1916RNP
- 6140Impsat
- 8167Brasil Telecom
- 11271COMSAT
- 14886Atos Origin
- 16735CTBC
- 21911Dualtec
- 26107*BSA Brasil
- 27715LocaWeb
- 28573NET
- 28607Easytone
- 28671Conecta
- 2688AT&T
- 7738Oi
- 10429Telefonica
- 12654RIPE NCC RIS
- 15201UOL
- 18881GVT
- 22356Durand
- 26598COMPUTEASY
- 28571USP
- 28587Yahoo!
- 28630Ascenty
- 11706Terra Networks

PTTMetro São Paulo

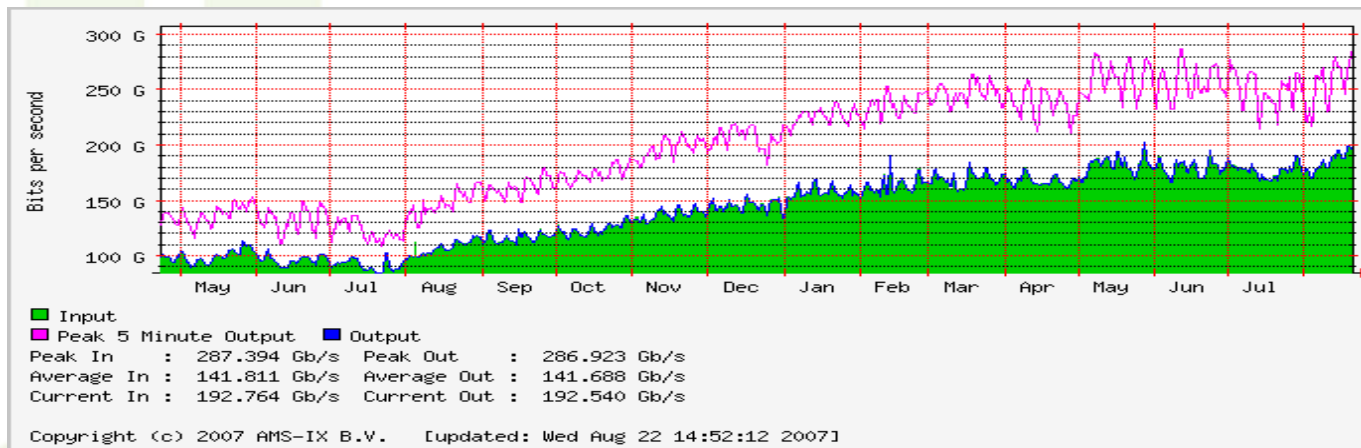
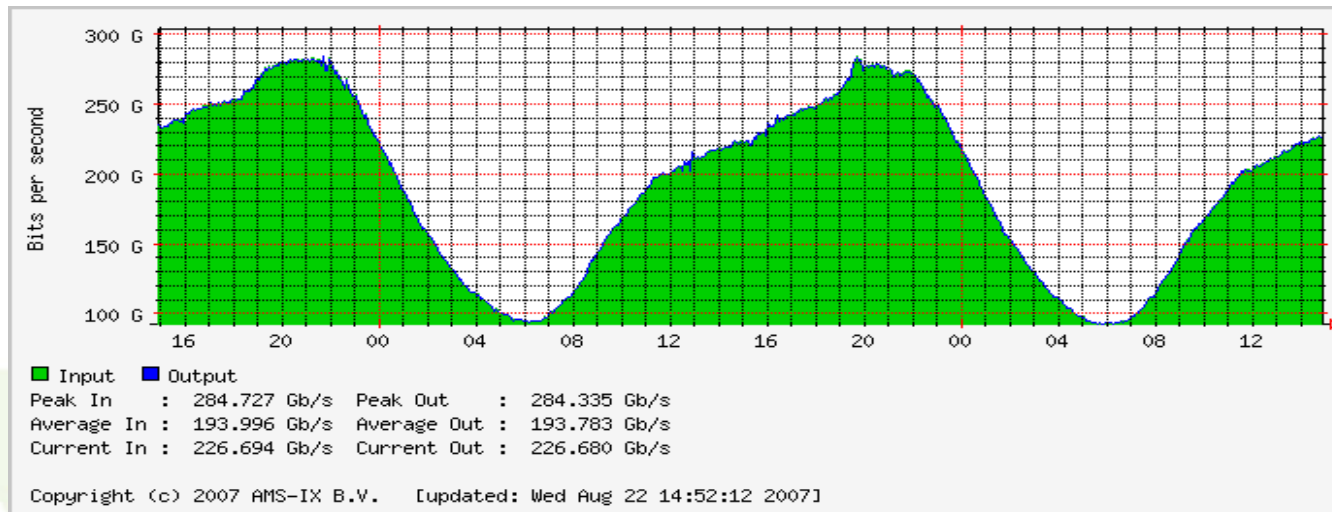


PTTMetro São Paulo - Evolução do tráfego



AMS-IX

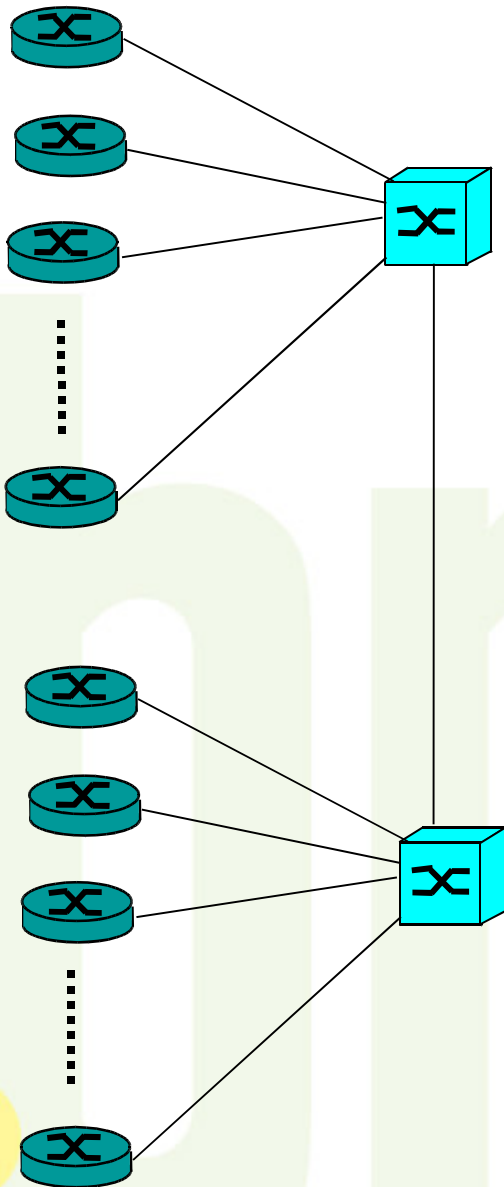
Amsterdam Internet Exchange



Detalhes do Projeto

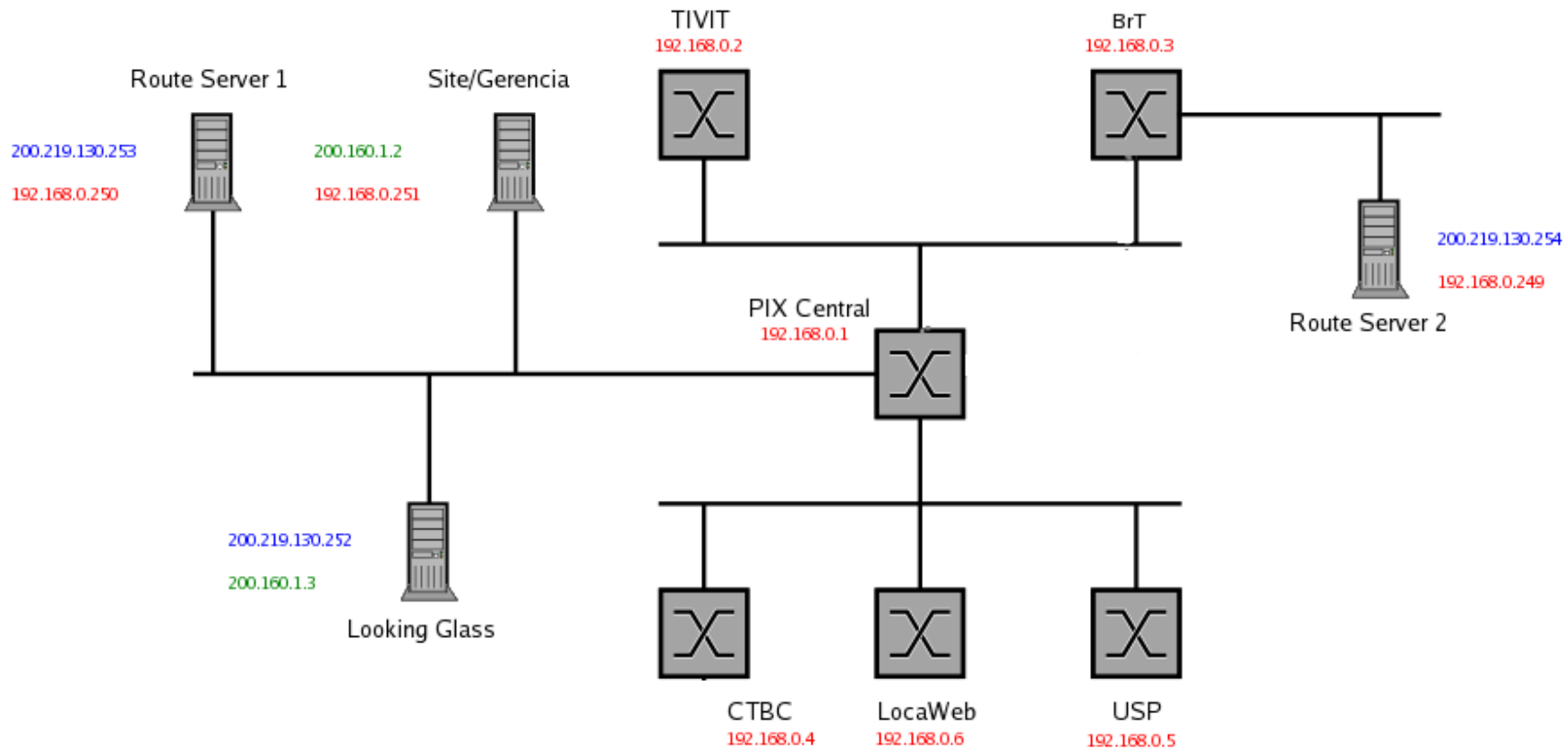
- Existe um switch metro ethernet por PIX (Ponto de Interconexão) com:
 - VLAN IPv4
 - VLAN IPV6
 - VLAN Gerência
- 4 servidores (S.O FreeBSD 6.x):
 - Estação de gerência e site
 - 2 *Route Servers* (*quagga* com suporte MD5)
 - *Looking Glass* (*quagga* com suporte MD5)
- Prefixos IPv4 e IPv6 exclusivos para cada localidade e não roteável globalmente (200.219.x.y/24)
- Prefixos IPv4 públicos para o *site* e *looking glass*
- Prefixo reservados (RFC 1918) para rede de gerência

PTTMetro – Aspectos Técnicos



- Conecta todos os PIXs através de fibra óptica apagada
- Arquitetura – Switch Ethernet camada 2
- VLAN (IPv4, IPv6, Multicast, ...)
- AS Privado (AS 64512 to AS 65535) para “route server” e “looking glass”
- Telefones IP interligados ao INOC-DBA (Inter-**N**etwork **O**perations **C**enter **D**ial-**B**y-**A**SN (<http://www.pch.net/inoc-dba/>)

PTTMetro - Arquitetura São Paulo



Regras técnicas para adesão

- ASN - possuir e operar um sistema autônomo.
- Acordo multilateral de tráfego (ATM) via RS, ou relações bilaterais diretas. Estabelecer acordos de troca de tráfego com outros participantes.
- Conexão a um único PIX por localidade.
- BGP4
- Caso anuncie prefixos de outros participantes é obrigatório marcar NEXT_HOP_SELF;
- Tráfego Broadcast deve estar limitado exclusivamente a resolução ARP;
- Participantes não podem apontar rota default ou se utilizar de recursos de outros sem a devida autorização.

Utilização do PTT em ambiente da TV digital

- **Canal de retorno:** Na proposta do documento Interozes - Coletivo Brasil de Comunicação Social (www.interozes.org.br) “TV Digital - princípios e propostas para uma transição baseada no interesse público” no seu item 11.4, apresenta a proposta de utilizar a estrutura dos PTTmetro nas principais cidades para implementar os canais de retorno.
- A qualidade de serviço assegurada no canal de retorno deve-se ao seu uso com interatividade em tempo real tais como: jogos, VoIP, gerenciamento *out band*, medição de qualidade da transmissão, envio de arquivos etc...
- **IPTV e VOIP:** assegura os requisitos de QoS (banda passante, latência, *jitter*, etc..) para a transmissão de conteúdos multimídias

Referências

- Comitê Gestor da Internet no Brasil – CGI.br <http://www.cgi.br>
- Núcleo de Informação e Coordenação do Ponto BR – NIC.br <http://www.nic.br>
- Ponto de Troca de Tráfego <http://www.ptt.br>
- TDMoIP como alternativa para Broadcasting em Aplicações IPTV - Tese de Dissertação de Mestrado de Wilson Dutra UnB Outubro de 2006
- Redes de Computadores e a Internet James F. Kurose Keith W. Ross
- Amsterdam Internet Exchange (www.mas-ix.net)

OBRIGADO !

oripide@nic.br