



BOAS PRÁTICAS DE INFRA-ESTRUTURA DE REDE PARA SUPORTE A VOZ-SOBRE-IP

Histórico de Revisões

Data	Versão	Autor	Descrição
27/08/2008	1.0	LCGL ¹	Versão Inicial
10/10/2008	1.1	PHM ²	Revisão Inicial

¹ Leandro C.G. Lustosa é Gerente de Desenvolvimento da Instant Solutions, e possui Mestrado em Redes de Computadores, com ênfase em Voz-sobre-IP, qualidade de serviço e avaliação de qualidade de voz pela Universidade Federal do Rio de Janeiro (UFRJ).

² Paulo Haroldo Mannheimer é Sócio-Fundador e Diretor Executivo da Instant Solutions, e é Bacharel em Informática (“Software Básico”) pela Universidade Federal do Rio de Janeiro (UFRJ). Paulo é também “Master in Business Administration” por U.C. Berkeley.

ÍNDICE

1. Introdução	3
2. Requisitos para tráfego de voz	3
2.1. Banda	3
2.2. Perdas na rede	4
2.3. Atraso	4
2.4. Variação do atraso (Jitter)	4
3. Políticas de priorização de tráfego	4
3.1. Identificando onde as políticas de priorização de tráfego devem ser aplicadas	5
3.1.1. Cenário 1 – Acesso a WAN	5
3.1.2. Cenário 2 – Interligação entre sites	6
3.1.3. Cenário 3 – Rede local	6
4. Suportando tráfego de voz - Nível 3	7
4.1. Marcação de pacotes de voz e desmarcação de pacotes não legítimos	7
4.2. Política de gerencia de fila	8
4.3. Fragmentação para enlaces de baixa velocidade	9
4.4. Atenção ao uso de mecanismos de tratamento de congestionamento	9
4.5. Coerência com o provedor de acesso	9
5. Suportando tráfego de voz – Nível 2 (Rede local)	9
6. Conclusão	12



1. Introdução

Este documento relaciona boas práticas de infra-estrutura de rede para suporte a voz. As recomendações apresentadas devem ser encaradas como uma diretriz básica para adequação da rede a fim de obtenção do esperado nível de qualidade que a plataforma *InstantVoice*® oferece.

2. Requisitos para tráfego de voz

Para se transportar voz interativa com adequado nível de qualidade deve-se atentar para os requisitos de rede exigidos em termos de consumo de banda, perdas na rede, atraso e variação de atraso. Esta seção descreve estes requisitos.

2.1. Banda

A banda requisitada para cada fluxo de voz depende diretamente do codificador (“codec”) e do meio de transporte utilizados. A tabela 1 apresenta o consumo de banda requisitado, por canal de voz, utilizando codificadores usuais nos principais meios de transmissão:

Codificador	Meio	Banda (kbps)
G.711 (todas as variações)	Ethernet	91.56
	PPP	86.52
	MPLS	87.36
	FrameRelay	86.94
G.729	Ethernet	32.76
	PPP	27.72
	MPLS	28.56
	FrameRelay	28.14
GSM	Ethernet	38.22
	PPP	33.18
	MPLS	34.02
	FrameRelay	33.60

Tabela 1: Consumo de banda para codificadores usuais

A plataforma *InstantVoice*® utiliza preferencialmente os seguintes codificadores para o tráfego de voz (consulte a equipe de suporte para alterações da configuração padrão):

- “Softphone” IVPhone: preferencialmente GSM, com opção de G.711 e iLBC (wideband);



- Ramais analógicos: preferencialmente G.711 em rede local, com opção de G.729 para ramais remotos;
- Ramais IP (telefone IP): preferencialmente G.711 em rede local, com opção de G.729 para ramais remotos;

2.2. Perdas na rede

A perda de pacotes, normalmente decorrente de congestionamentos na rede, é uma das principais causas da má qualidade de voz. Pontos de congestionamento (gargalos) devem ser identificados e tratados com priorização do fluxo de voz.

A perda de pacote não deve ultrapassar a taxa de 1% no pior caso. É importante notar que codificadores mais robustos (como, por exemplo, o G.711) possuem maior tolerância a perdas que codificadores menos robustos (como, por exemplo, o G.729 e o GSM), entretanto, codificadores mais robustos normalmente requisitam maior consumo de banda.

2.3. Atraso

O atraso demasiado acarreta na perda de interatividade da conversa. Recomenda-se que o atraso máximo na rede em um sentido não ultrapasse 100 ms.

2.4. Variação do atraso (Jitter)

A variação do atraso pode acarretar em perdas nos *buffers* de compensação de *jitter* dos dispositivos VoIP e alta degradação na qualidade da voz. É recomendado que a variação do atraso não ultrapasse 20 ms.

Uma variação de até 30 ms é aceitável em algumas situações onde o atraso da rede é baixo, possibilitando que os *buffers* de compensação de *jitter* trabalhem com tamanhos maiores.

3. Políticas de priorização de tráfego

Políticas de priorização de tráfego são comumente chamadas de “políticas de gerenciamento de filas”, já que agem nas filas dos equipamentos de rede de forma a favorecer fluxos de maior prioridade em detrimento aos fluxos de menor prioridade.

3.1. Identificando onde as políticas de priorização de tráfego devem ser aplicadas

É importante notar que tais políticas só são efetivas quando aplicadas em pontos onde filas são geradas. Não faz sentido priorizar tráfego onde não há fila.

Logo, ao se projetar uma rede com priorização de tráfego, um dos passos mais importantes é a identificação dos pontos de gargalo onde as políticas de gerenciamento de filas devem ser aplicadas.

Nas figuras nas seções 2.1.1, 2.1.2 e 2.1.3 são ilustrados três cenários usuais de redes. Em cada um dos cenários são apontados os pontos de gargalo onde a aplicação de políticas de priorização de tráfego é efetiva.

3.1.1. Cenário 1 – Acesso a WAN

No cenário ilustrado na figura 1 podemos observar um roteador com uma interface de rede local (*Fast Ethernet* com 100 Mbps) e uma interface WAN de baixa velocidade (2 Mbps).

É importante observar que no sentido WAN para LAN a geração de filas não é caracterizada, já que o roteador recebe o tráfego da WAN em uma velocidade menor que é capaz de enfiar para a LAN.

Entretanto, no sentido LAN para WAN, é possível observar uma tendência a geração de filas, já que o roteador tende a receber da LAN mais tráfego do que é capaz de repassar para a WAN.

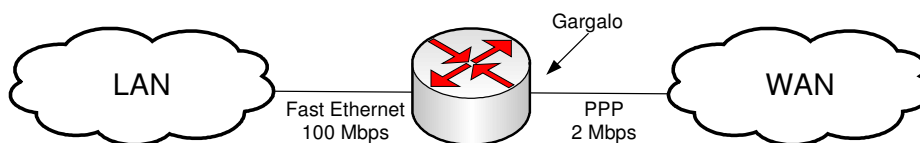


Figura 1: Cenário 1 - Acesso a WAN

Há um gargalo na interface WAN de saída do roteador. Assim, caso se queira priorizar fluxos saíntes, é necessário que se aplique uma política de priorização de tráfego no sentido de saída da interface WAN do roteador.

3.1.2. Cenário 2 – Interligação entre sites

O segundo cenário, descrito na figura 2, é muito parecido com o primeiro cenário. De fato, o segundo cenário é a junção de dois cenários 1. No primeiro cenário foi apontado onde se deve controlar a prioridade dos fluxos saíntes da LAN para a WAN. Podemos imaginar o Cenário 1 como o acesso a Internet.

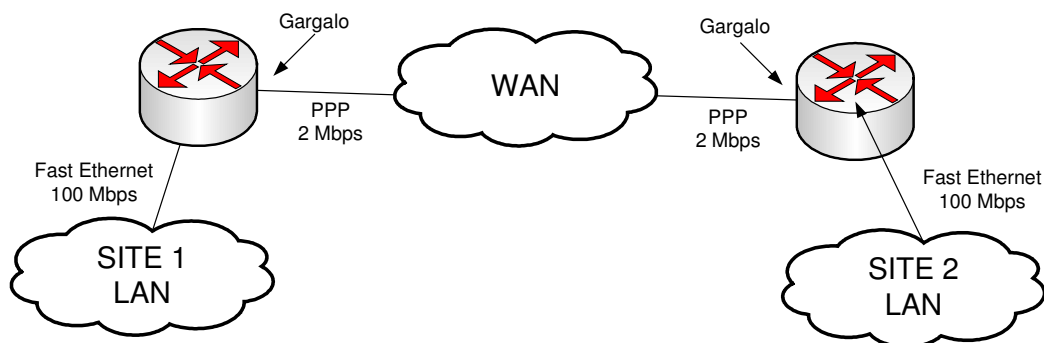


Figura 2: Cenário 2 - Interligação entre sites

O cenário 2 pode ser imaginado como uma ligação entre dois *sites* de uma mesma empresa. Neste caso, iremos controlar a prioridade dos fluxos em ambos os sentidos. Na interface de saída da interface WAN do roteador do SITE 1 controlamos a prioridade dos fluxos destinados ao SITE 2, e na interface de saída da interface WAN do roteador do SITE 2 controlamos a prioridade dos fluxos destinados ao SITE 1.

3.1.3. Cenário 3 – Rede local

O terceiro cenário, ilustrado na figura 3, possui uma característica diferente dos dois cenários apresentados anteriormente. Nos cenários 1 e 2 são apontados os pontos onde as políticas de priorização de tráfego devem ser habilitadas em um roteador. Nestes cenários a priorização é sempre baseada no nível 3, ou seja, na marcação de pacotes IP.

O cenário 3 (figura 3) não há roteador. Há apenas *switches*. Logo, a priorização de tráfego deve ser feita em nível 2.

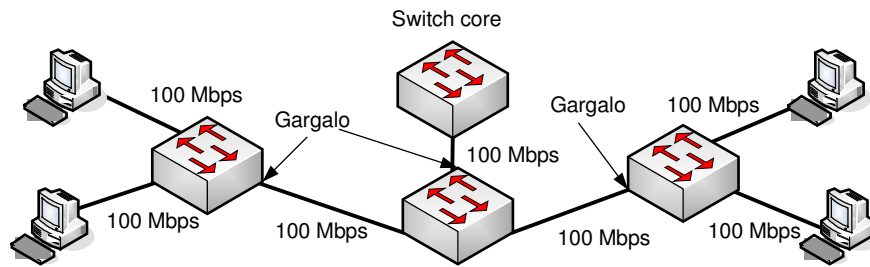


Figura 3: Cenário 3 - Rede local

Entretanto, a premissa de identificar pontos de gargalo, onde um equipamento de rede recebe mais banda por uma interface (ou um conjunto de interfaces) do que consegue transmitir por outra interface continua valendo.

Seguindo esta premissa, é possível identificar três pontos de gargalo no cenário 3. Note que se cada um dos computadores gerarem 100 Mbps destinados a algum elemento conectado diretamente ao *switch core*, os pontos indicados as interfaces de *uplink* dos *switches* ficarão engarrafadas. Logo, é possível identificar que tais pontos são efetivos para a aplicação de políticas de priorização de tráfego.

4. Suportando tráfego de voz - Nível 3

As rotas de nível 3 por onde a voz trafega devem ser identificadas. Os pontos onde as políticas de priorização de voz são efetivas (gargalos) devem ser identificados, e as recomendações descritas nos sub-tópicos desta seção devem ser averiguadas.

4.1. Marcação de pacotes de voz e desmarcação de pacotes não legítimos

- Todos os pacotes referentes aos fluxos de voz devem ser marcados;
- Recomenda-se o uso de DSCP EF;
- Os pacotes de voz podem ser caracterizados pelos endereços IP dos equipamentos envolvidos e pelas portas de transporte utilizadas pelas aplicações VoIP:
 - Porta UDP 5060 – SIP
 - Porta UDP 4569 – IAX
 - Porta UDP 10000 a 20000 – RTP (padrão *InstantVoice*®)
 - IP dos *MediaGateways* (E1, FXO e FXS)
 - IP das estações que com *softphone* VoIP
 - IP dos *softswitches*

- Pacotes não associados aos fluxos de voz marcados como pacotes de voz (DSCP EF) devem ser desmarcados (pacotes não legítimos);
- A marcação deve ser realizada em qualquer equipamento que esteja localizado anteriormente ao primeiro roteador que opera alguma política de priorização de fluxos de voz, incluindo o este próprio roteador (desde que a marcação seja realizada anteriormente ao mecanismo de priorização);
- A desmarcação de pacotes não legítimos deve ser realizada em qualquer equipamento que esteja localizado anteriormente ao primeiro roteador que opera alguma política de priorização de fluxos de voz, incluindo o este próprio roteador (desde que a marcação seja realizada anteriormente ao mecanismo de priorização).

4.2. Política de gerencia de fila

- Os fluxos referentes aos pacotes de voz devem ser gerenciados por uma fila que garanta a banda que se deseja reservar para voz (olhar seção 2.1);
- Recomenda-se uma política de fila que sirva os pacotes de voz com prioridade máxima até a exaustão (atendendo o limite de banda reservado). Esta política é comumente conhecida como “*Priority Queue*” (ou simplesmente PQ) e é oferecida pela maioria dos equipamentos de rede que suportam priorização de tráfego IP;
- Recomenda-se que os fluxos menos prioritários, que não são tratados pela PQ, sejam tratados por uma política de filas do tipo “*Fair Queue*”, de modo que aplicações interativas de menor consumo de banda (como SSH, TELNET, HTTP, entre outras) não sejam atrapalhadas por aplicações não interativas de maior consumo de banda (download FTP, download HTTP, entre outras). O uso de *Fair Queue* também diminui a probabilidade de que a voz seja degradada caso a banda reservada pela PQ seja extravasada (o tráfego que extravasar da PQ deve ser direcionado para a *Fair Queue*);
- Alguns fabricantes alertam que a reserva de mais de mais de 33% da banda de um enlace para PQ pode causar inanição aos fluxos de menor prioridade;
- Deve-se verificar o tamanho do buffer das interfaces dos equipamentos que estão priorizando os fluxos de voz. Os equipamentos costumam utilizar um buffer de saída do tipo FIFO (*First In First Out*), que quando configurado com valores elevados, podem diminuir a eficiência do mecanismo de priorização e acrescentar muito atraso na rede.
- É importante alertar que é comum que, após a habilitação de mecanismos de priorização, os buffers de saída sejam configurados automaticamente para valores padrões que

usualmente são demasiadamente altos. Estes buffers precisam ser ajustados levando-se em conta as necessidades de atraso exigidas pela voz (olhar seção 2.3).

4.3. Fragmentação para enlaces de baixa velocidade

Em enlaces com velocidades inferiores a 1024 kbps, torna-se necessária a utilização de técnicas como LFI (*Link Fragmentation and Interleaving*). O tempo de serialização de um pacote em uma interface é inversamente proporcional à velocidade do *link* desta interface; um pacote de voz pode sofrer um atraso inaceitável aguardando o término da transmissão de um pacote de dados em um link de baixa velocidade. O LFI faz com que um quadro sendo enviado seja fragmentado em tamanhos menores, permitindo a inserção de pacotes prioritários entre eles.

A tabela 2 apresenta o tamanho de fragmentos recomendados para enlaces de baixa velocidade.

Velocidade do enlace	Fragmentação
512 kbps	512 bytes
768 kbps	1024 bytes
1024 kbps (ou maior)	não necessário

Tabela 2: Fragmentação em enlaces de baixa velocidade

4.4. Atenção ao uso de mecanismos de tratamento de congestionamento

Não deve-se utilizar de algoritmos para tratar congestionamentos nas filas associadas aos fluxos de voz, como RED (Random Early Detection) ou WRED (Weighted Random Early Detection). Estes algoritmos causam descartes de pacotes que afetarão a qualidade da voz.

4.5. Coerência com o provedor de acesso

É importante ter em mente que o provedor de acesso também deve priorizar o tráfego de voz, e que deve existir uma coerência entre a configuração de priorização de tráfego adotada internamente e a política de priorização adotada pelo provedor.

O provedor deve ser informado como os pacotes de voz estão sendo marcados para que as regras de priorização sejam mapeadas para dentro da rede do provedor.

5. Suportando tráfego de voz – Nível 2 (Rede local)



Uma vertente para suporte de voz em rede local é a segmentação em nível 2 baseada em VLAN com uso de priorização segundo o padrão 802.1Q/p. Entretanto, este tipo de infra-estrutura acrescenta muita complexidade à solução.

A priorização segundo o padrão 802.1Q/p deve ser configurada pela aplicação, e a placa de rede dos equipamentos VoIP (estações de trabalho que operam *SoftPhone*, Telefones IP, *MediaGateways*, *Softswitches*, entre outros), assim como seus respectivos sistemas operacionais, devem suportar o 802.1Q/p. Também há a necessidade de configuração individual de cada um destes equipamentos.

A segmentação por VLAN exige ainda que as portas dos switches onde estes equipamentos se interligam também sejam identificadas e configuradas para segmentação da VLAN.

O roteamento de nível 3, para encaminhamento entre VLANs também se torna necessário, adicionando assim, mais custo para solução, já que normalmente há a necessidade de investimento em *switches* de nível 3 de alta performance.

Outra vertente, que em alguns casos também exige investimentos, mas que por outro lado não aumenta complexidade da rede, é a aproximação dos equipamentos VoIP do *core* da rede. Esta vertente segue as seguintes premissas:

- Os pontos de gargalo de nível 2 (como ilustrado na seção 2.1.3) devem ser minimizados nas rotas de voz;
- Cada equipamento VoIP (estações de trabalho que operam *SoftPhone*, Telefones IP, *MediaGateways*, *Softswitches*, entre outros) deve ter uma interface *Full Duplex* de ao menos 100 Mbps dedicada até o *switch* de acesso. O cascadeamento de uma estação de trabalho a um Telefone IP com *switch* interno é permitido;
- Equipamentos que concentram muitas chamadas de voz, como *MediaGateways* e *Softswitches* (E1 ou FXS) devem estar ligados diretamente ao *core* da rede;
- Estações de trabalho que utilizam voz como aplicação de missão crítica (estações de operadores de um *CallCenter*) devem possuir um cascadeamento máximo de 1 nível (um *switch*) até o *core* da rede, e devem estar conectadas a *switches* dedicados;
- O cascadeamento entre os *switches* devem ser realizados através de enlaces de maior velocidade como, por exemplo, *Gigabit Ethernet* ou *FastEthernet Channel* (junção de grupos de interfaces *FastEthernet* operando em paralelo).

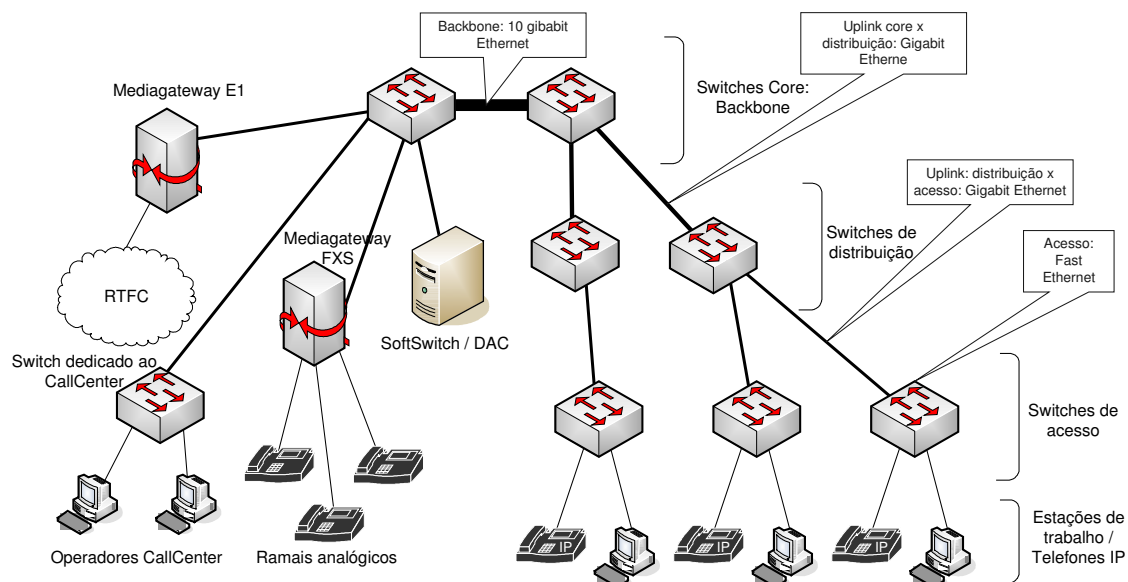


Figura 4: Infra-estrutura de rede local

A figura 4 ilustra uma infra-estrutura de rede local que segue as premissas descritas acima. Os *switches* estão classificados em três grupos: *Core*, distribuição e acesso.

Os *switches Core* estão conectados a alta velocidade e abrigam os elementos que concentram muitas chamadas de voz, como os *Mediagateways* e o *SoftSwitch/DAC*, assim como, os *switches* dedicados às estações dos operadores de *CallCenter* que operam *SoftPhones*.

Os *switches* de distribuição interligam os *switches* de acesso em alta velocidade. Os *switches* de acesso por sua vez, oferecem portas *FastEthernet* dedicadas a cada um dos Telefones IP e as estações de trabalho que operam *SoftPhones*.



6. Conclusão

As recomendações apresentadas neste material visam nortear os princípios básicos de infraestrutura para a garantia da qualidade de voz trafegando sobre o protocolo IP.

Este documento foi desenvolvido a partir de nossa experiência em casos reais de uso, e está em constante aprimoramento à medida que novos cenários exijam sua atualização. Acesse www.instant.com.br e verifique, sempre que possível, a última versão disponível.

A equipe técnica da Instant Solutions está, como sempre, à disposição para discutir detalhes específicos de sua implantação, através dos seguintes canais:

Email: tech@instant.com.br

Instant Solutions Tecnologia e Informática Ltda.
Rua da Quitanda, 62 – 12º andar
Centro – Rio de Janeiro
CEP: 20011-030
Telefone: (011/021) 3956.6100

===