

Métricas de Gerenciamento de Redes

Prof. Michel Bonfim (michelsb@ufc.br)

Disciplina: Gerência de Redes

Métricas

- Em uma rede operacional, existem várias quantidades relacionadas com o desempenho e com a confiabilidade do seus componentes;
- **Métrica:** especificação destas quantidades;
- **Medida:** valores quantificados de uma métrica;
- O processo de quantificar uma medição de uma métrica é definido como uma **Metodologia de Medição** de uma métrica.

Métricas

- O objetivo principal para a definição de métricas é prover uma base para a avaliação do desempenho de diferentes componentes desta rede.
 - Auxiliar na detecção e resolução de problemas;
 - Permitir o planejamento e previsões de crescimento das redes

Métricas

Grupos de Métricas:

- Disponibilidade;
- Desempenho;
- Confiabilidade.

Disponibilidade

- Refere-se a quantidade de tempo que uma rede está disponível para os usuário.
- Disponibilidade pode ser expressa como um percentual por ano, mês, semana, dia ou hora que a rede estará disponível em relação ao tempo total (**uptime X downtime**)

$$Disponibilidade(\%) = \frac{(PeríodoTotalMonitoramento - TempoIndisponível)}{PeríodoTotalMonitoramento} * 100\%$$

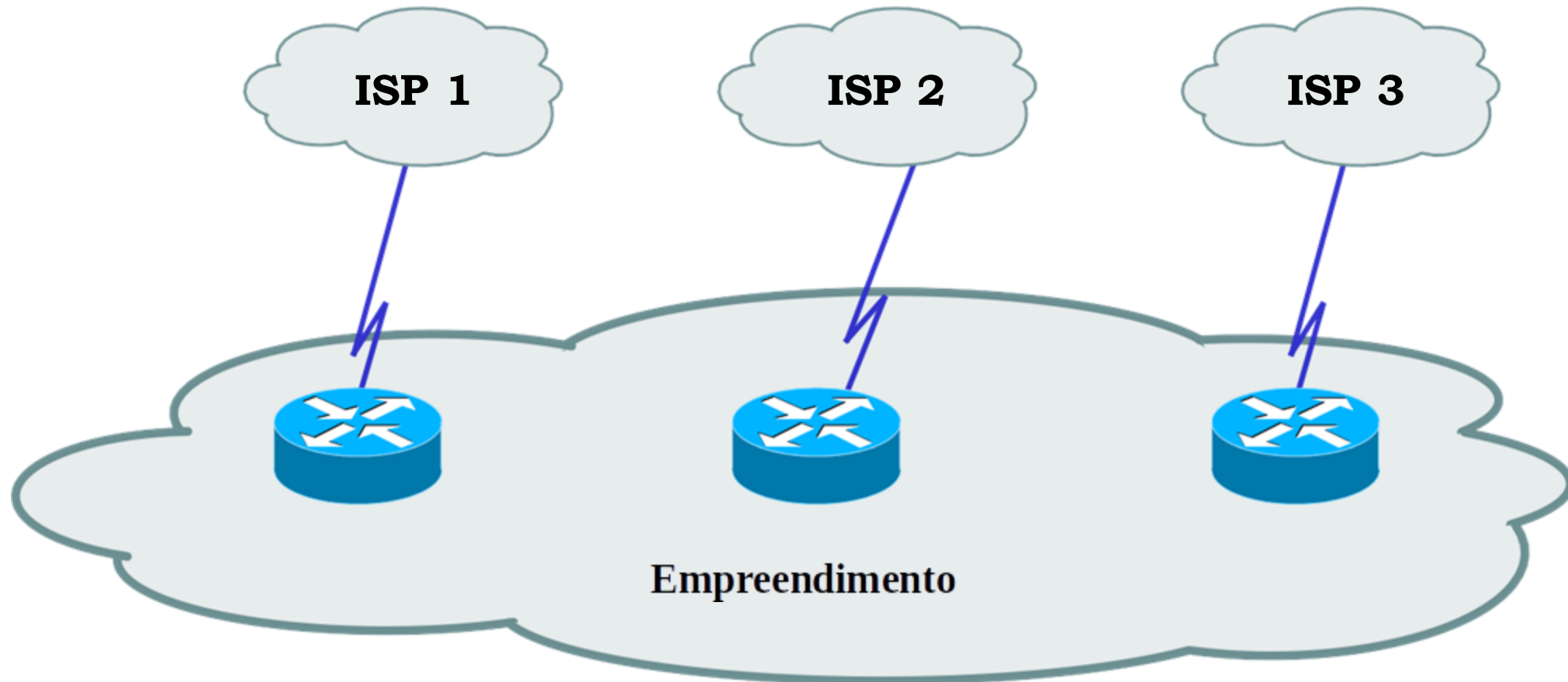
- Por exemplo:
 - Operação 24/7 = 24 horas, 7 dias da semana = 100% de disponibilidade
 - 165 horas em 168 horas semanais = 98,21% de disponibilidade

Disponibilidade

Disponibilidade (%)	<i>Downtime/ano</i>	<i>Downtime/mês</i>
95%	18 dias 6:00:00	1 dias 12:00:00
96%	14 dias 14:24:00	1 dias 4:48:00
97%	10 dias 22:48:00	0 dias 21:36:00
98%	7 dias 7:12:00	0 dias 14:24:00
99%	3 dias 15:36:00	0 dias 7:12:00
99,9%	0 dias 8:45:35.99	0 dias 0:43:11.99
99,99%	0 dias 0:52:33.60	0 dias 0:04:19.20
99,999%	0 dias 0:05:15.36	0 dias 0:00:25.92

Disponibilidade

Quanto maior a disponibilidade, maior a redundância e custo das soluções

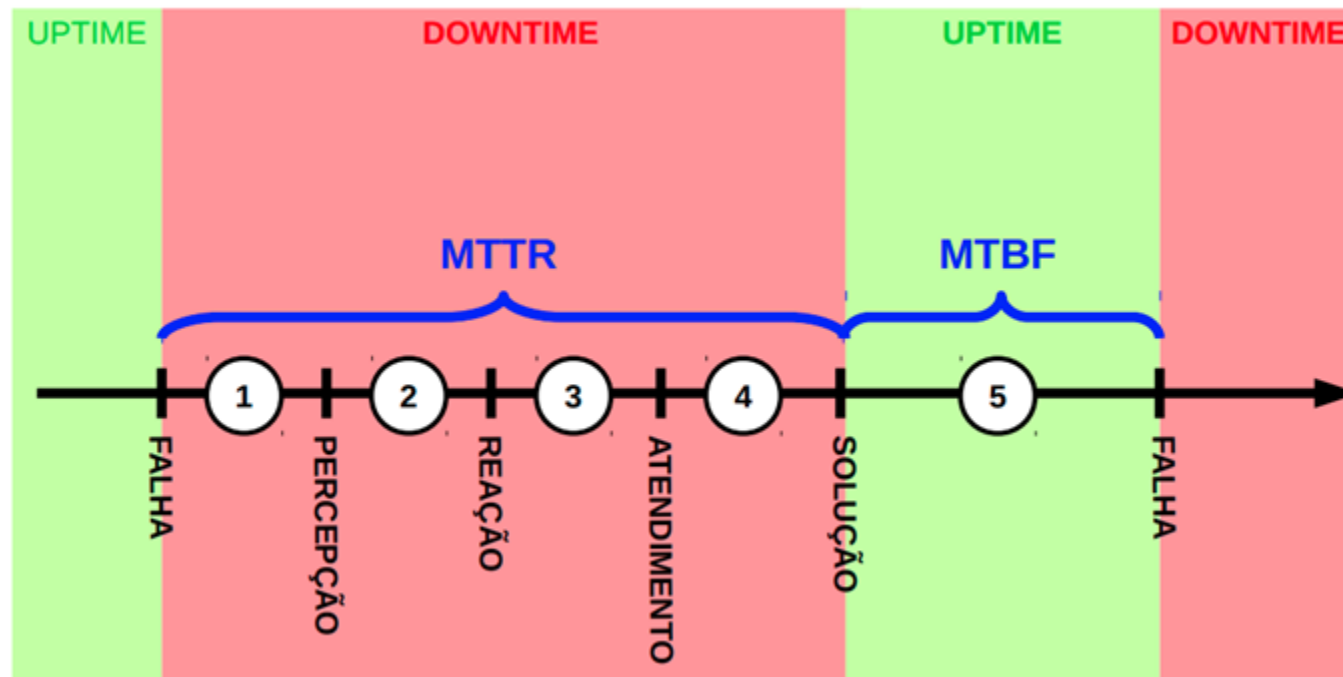


Disponibilidade

- Disponibilidade também pode ser expressado por:

$$\text{Disponibilidade} = \text{MTBF} / (\text{MTBF} + \text{MTTR})$$

- Sendo:
 - MTBF - Mean Time Between Failures**(tempo medio entre falhas)
 - MTTR - Mean Time To Repair** (tempo médio de recuperação)



Disponibilidade

Por exemplo:

A rede não deve falhar mais do que 1 vez em 166 dias (4000 horas) e o tempo de reparo não deve exceder 1 hora

$$\textit{Disponibilidade} = 4.000/4.001 = 99,98\%$$

Desempenho da Rede

Métricas de desempenho comuns incluem:

- Vazão (**Throughput**);
- **Goodput**;
- **Largura de Banda**;
- Atraso (latência - **delay**);
- Tempo de resposta (Round-trip Time - **RTT**);
- Variação do atraso (**jitter**);
- **Utilização de Recursos**.

Vazão ou Throughput

- **Vazão ou Throughput** é a quantidade de dados transmitidos sem erro por unidade de tempo, medida em:
 - bps (bits por segundo), Kbps, Mbps e Gbps;
 - B/s (Bytes por segundo);
 - pps (pacotes por segundo);
- O throughput pode ser traduzido como a taxa de **transferência efetiva de um sistema**:
 - A taxa de transferência efetiva de um determinado sistema (uma rede de roteadores por exemplo) pode ser menor que a taxa de entrada devido às perdas e atrasos no sistema.

Vazão ou Throughput

- Fatores que podem afetar a Vazão:
 - Tamanho dos pacotes;
 - Espaço inter-quadros entre dois pacotes;
 - Taxa de encaminhamento de pps do dispositivo;
 - Capacidade do cliente, servidor e dispositivos de rede (CPU, memória e velocidade de acesso ao dispositivo de armazenamento);
 - Protocolos;
 - Distância;
 - Erros;
 - Hora do dia.

Throughput versus **Goodput**

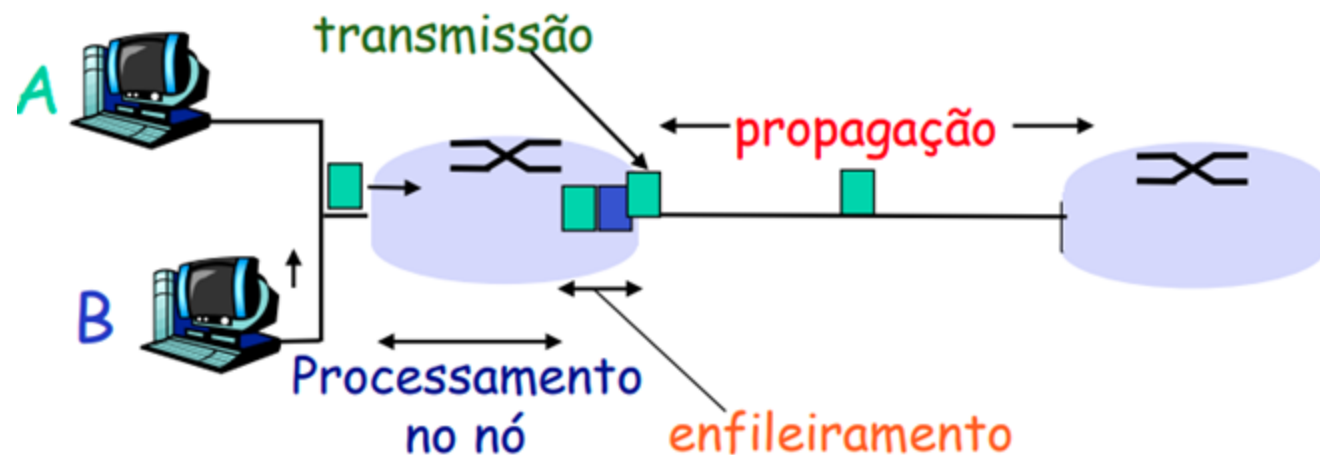
- **Goodput** é a vazão na camada de aplicação
- É necessário especificar:
 - Throughput ou goodput?
 - B/s inclui somente os dados da aplicação ou também os bytes do cabeçalho ?

Largura de Banda

- Largura de Banda e vazão não são a mesma coisa
- Largura de Banda é a capacidade de transporte de dados da rede
 - Expressa em bps (bits por segundo), Kbps, Mbps e Gbps;

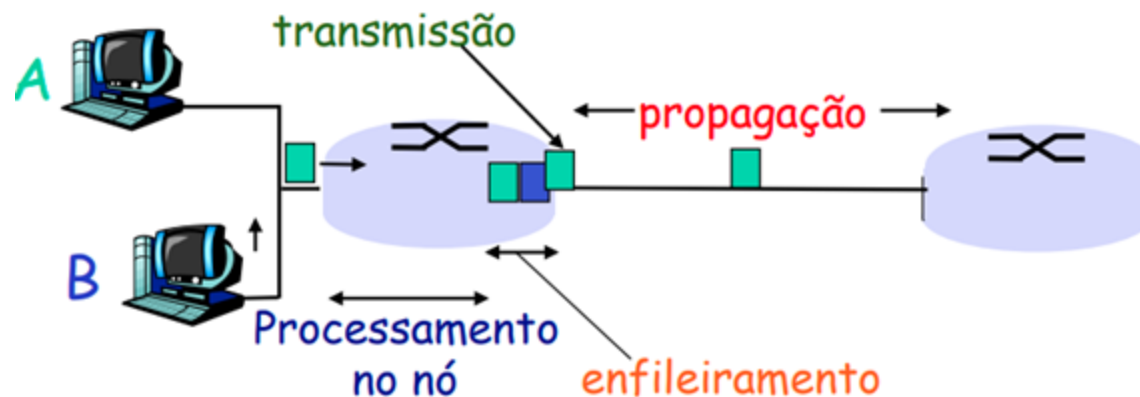
Atraso - Latência - Delay

- Atrasos de entregas de pacotes em uma rede podem ser ocasionados por diversos fatores.
- Há quatro tipos principais de atraso:
 - Atraso de processamento;
 - Atraso de fila;
 - Atraso de transmissão;
 - Atraso de propagação.



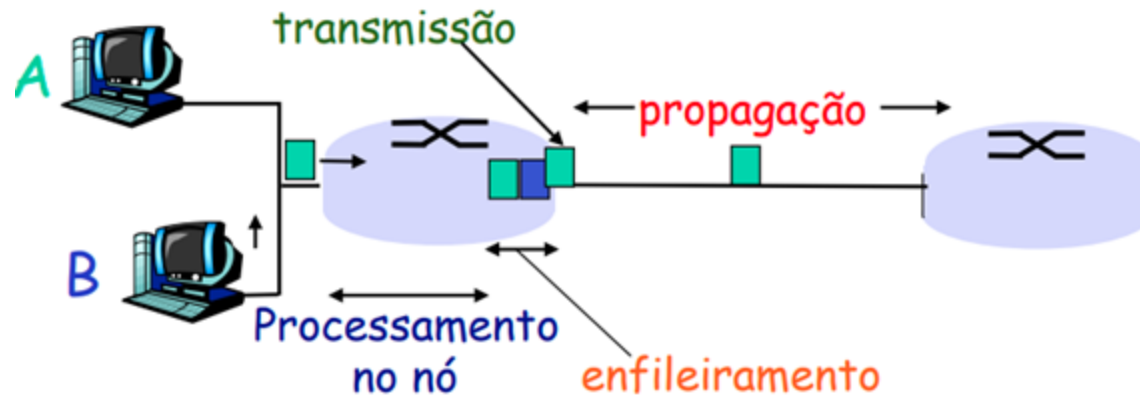
Atraso de Processamento

- O tempo necessário para um equipamento:
 - Ler e processar o cabeçalho do pacote
 - Decidir para onde este deve ser encaminhado
- Normalmente, poucos microssegundos ou menos.



Atraso de fila

- Tempo em que um pacote fica na fila aguardando ser enviado ao próximo nó da rede
- O tempo dependerá do congestionamento



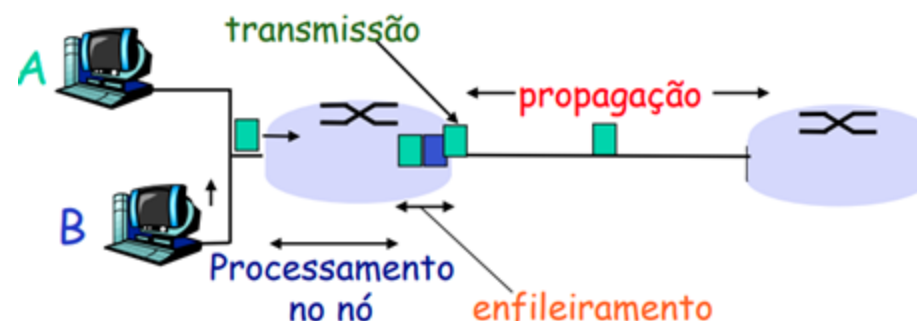
Atraso de Transmissão

- É o tempo necessário para transportar todos os bits do quadro, de dentro do equipamento para o meio de transmissão.

R = banda do enlace (bps)

L = tamanho do pacote (bits)

Atraso de Transmissão = **L/R**



- Normalmente será de alguns micro a milissegundos
- Por exemplo, é necessário cerca de 5 ms para transmitir um pacote de 1024 bytes a uma taxa de 1,544 Mbps

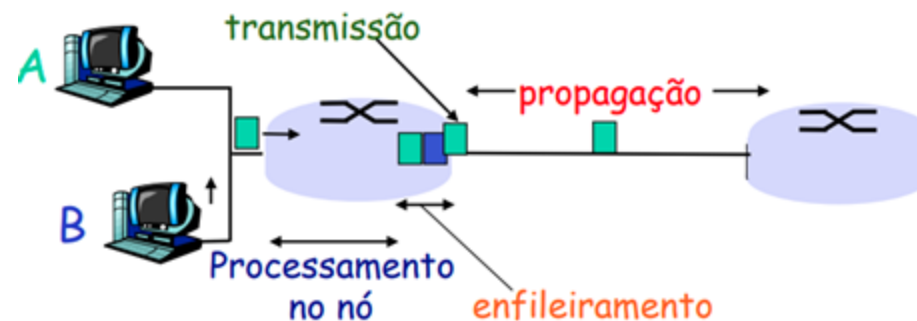
Atraso de Propagação

- Uma vez que um bit seja empurrado no link, ele precisa se deslocar para o elemento seguinte. O tempo gasto para propagar do começo do link até o router B é o **atraso da propagação**.
- **A velocidade da propagação depende do meio físico do link** (isto é, fibra ótica, cabo coaxial, cabo par-trançado, etc) e está na escala de 2×10^8 m/s a 3×10^8 m/s (\cong velocidade de luz).

S = distância entre dois elementos de rede

v = velocidade de propagação do sinal

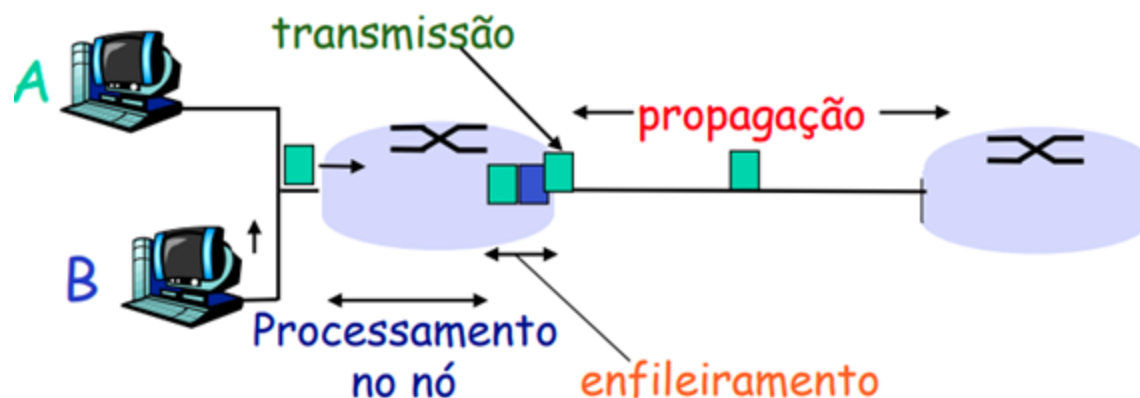
$$\text{Atraso da Propagação} = S / v$$



- Em redes WAN, os atrasos da propagação estão na ordem dos microsegundos.

Atraso - Latência - Delay

- A soma dos atrasos nos dá o **Atraso Total**;
- Quanto maior este valor menor será a velocidade de transmissão de dados na sua rede.



Tempo de Resposta

- Tempo decorrido entre a requisição de uma ação ao sistema e a resposta completa à sua requisição;
- Quanto menor a latência, melhor o tempo de resposta da rede.
- Depende de diversas variáveis e suas interações:
 - Função da aplicação e do equipamento que roda a aplicação, não apenas da rede;
 - Usuários desejam uma resposta em 100 a 200 milisegundos;

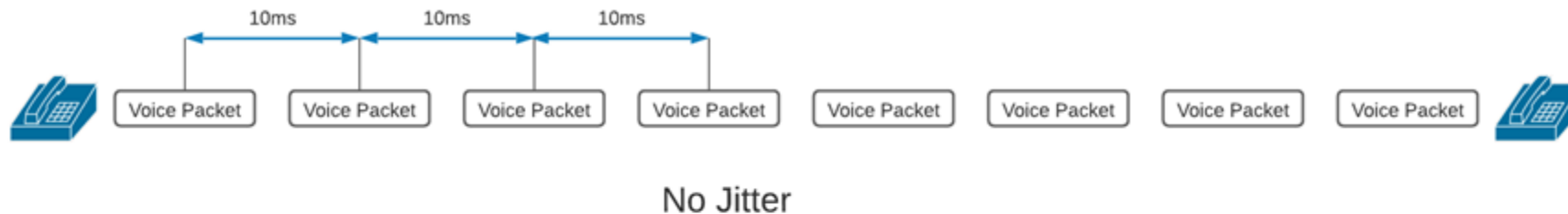


Tempo de Resposta

```
[root@madeira ~]# ping 8.8.8.8
PING 8.8.8.8 (8.8.8.8) 56(84) bytes of data.
64 bytes from 8.8.8.8: icmp_req=1 ttl=54 time=93.9 ms
64 bytes from 8.8.8.8: icmp_req=2 ttl=54 time=155 ms
64 bytes from 8.8.8.8: icmp_req=3 ttl=54 time=173 ms
64 bytes from 8.8.8.8: icmp_req=4 ttl=51 time=91.2 ms
64 bytes from 8.8.8.8: icmp_req=5 ttl=52 time=82.4 ms
64 bytes from 8.8.8.8: icmp_req=6 ttl=54 time=146 ms
^C
--- 8.8.8.8 ping statistics ---
6 packets transmitted, 6 received, 0% packet loss, time 5005ms
rtt min/avg/max/mdev = 82.499/123.993/173.882/35.823 ms
```

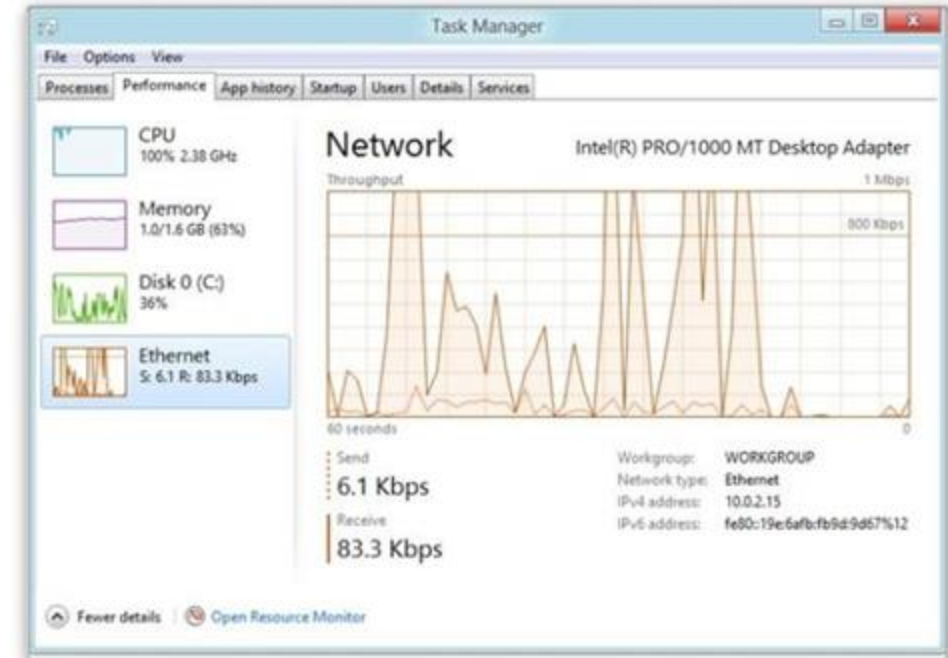
Variação do atraso - Jitter

- Quantidade de tempo médio que o atraso sofre variação
- Voz, vídeo, e áudio não toleram variação do atraso
- Tamanho dos pacotes menor é mais eficiente nesse caso
- Situação problema:
 - Eficiência para grande volume de aplicações vs. pequena variação do atraso requerida por aplicações multimídia
 - Pacotes maiores têm melhor eficiência e goodput



Utilização de Recursos

- É a **quantidade em uso** de um determinado recurso (**porcentagem**), em relação à sua capacidade total de atendimento;
- Auxilia na **avaliação e detecção de gargalos**, uma vez que influenciam diretamente no tempo de resposta da rede e das aplicações envolvidas;
- Monitorando-se a utilização, ao longo de um determinado período de tempo, é possível **observar os intervalos de pico** na utilização de um serviço ou recurso;
- Exemplos: ***memória disponível, leitura/escrita em disco, processador (CPU), largura de banda da rede, quantidade de máquinas virtuais***, etc.



Confiabilidade

- Confiabilidade de um sistema é uma função do tempo definida como uma probabilidade condicional que um sistema funcione em perfeitas condições em um intervalo de tempo;
- Diferente de disponibilidade:
 - confiabilidade depende de um intervalo de tempo e a disponibilidade é obtida em um instante de tempo específico.
- Métricas de confiabilidade comuns incluem:
 - Perda de Pacotes/Quadros;
 - Precisão.

Perda de Pacotes/Quadros

- A perda de pacotes/quadros analisa o número de pacotes/quadros que foram transmitidos pelo transmissor e que nunca foram recebidos em seu destino.
 - Razões: erros e atraso excessivo.
- Problema crítico quando os pacotes/quadros não podem ser retransmitidos, como nos protocolo UDP e RTP;
- A métrica normalmente é chamada de Taxa de Perda de Pacotes/Quadros.
 - Expressa como uma porcentagem do número total de pacotes/quadros transmitidos.
- Por exemplo:
 - Se 1000 quadros foram transmitidos, mas somente 900 recebidos, a taxa de perda seria:

$$((1000 - 900) / 1000) \times 100\% = 10\%$$

Perda de Pacotes/Quadros

- **No caso do atraso excessivo:**
 - Quando o atraso se torna muito alto é possível que uma das filas do roteador se torne cheia;
 - Um roteador tem um espaço de memória para armazenar esta fila;
 - Se esta fila se tornar tão grande que preencha toda a memória;
 - Pacotes começaram a ser descartados;
 - Esses pacotes serão considerados perdidos.

Precisão

- O objetivo geral da precisão é que os dados recebidos no destino devem ser os mesmos os dados enviados pela fonte.
- Em todo tipo de comunicação à distância, existe a possibilidade de ocorrer erro na hora de se interpretar os dados:
 - As causas típicas de erros de dados incluem picos de energia, conexões físicas pobres , dispositivos falhando e ruídos causados pela maquinaria elétrica .
 - Às vezes, erros de software podem causar erros de dados também, apesar de ser menos comum.
 - Pode ser especificado como uma **Taxa de Erro de Bit (BER)**.
- Precisão geralmente refere-se ao **número de quadros transmitidos sem erros em relação ao número total de quadros transmitidos**.

Precisão

- A maioria dos dispositivos de camada 2 (incluindo ethernet) descartaram um quadro que tiver o valor de verificação do quadro incorreto (FCS - Frame Check Sequence);
- Portanto, o BER ou a Precisão só pode ser averiguada em alguns protocolos, como o SONET/SDH.

QoS

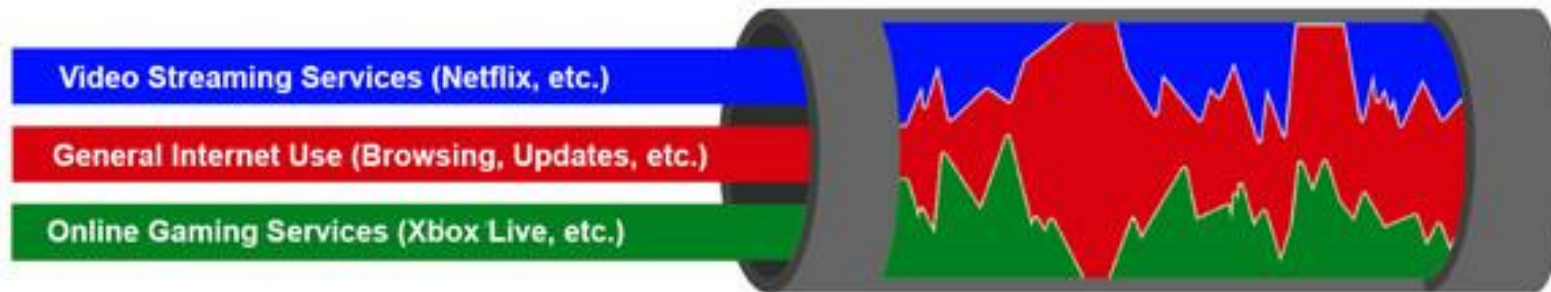
- **QoS (Qualidade de Serviço):** garantias de transmissão para certos fluxos de dados
- Estabelecimento de métricas e parâmetros de **QoS (Qualidade de Serviço)**
 - **Atraso:** É o tempo necessário para um pacote ser passado do emissor, através da rede, até o receptor.
 - **Variação do atraso (jitter):** É a variação no atraso fim-a-fim.
 - **Largura de banda:** É a taxa de transmissão de dados máxima que pode ser sustentada entre dois pontos finais.
 - **Confiabilidade:** Como uma propriedade dos sistemas de transmissão, pode ser vista como a taxa de erros do meio físico.
- Modelos de serviços e mecanismos:
 - Arquitetura de Serviços Integrados (Intserv)
 - A Arquitetura de Serviços Diferenciados (Diffserv)
 - Multiprotocol Label Switching (MPLS)
 - Roteamento com QoS;
 - Engenharia de Tráfego.

Requisitos de QoS

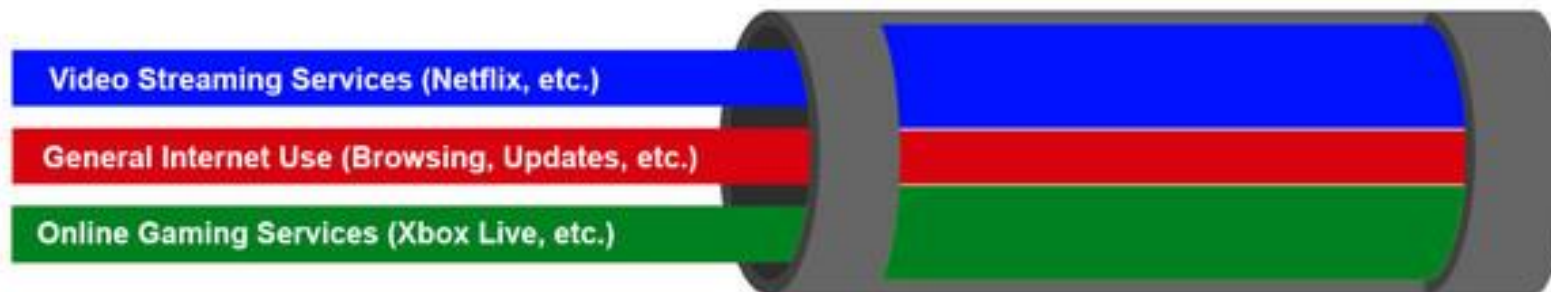
APLICAÇÃO	CONFIABILIDADE	RETARDO	FLUTUAÇÃO	LARG. DE BANDA
Correio eletrônico	Alta	Baixa	Baixa	Baixa
Transf. de Arquivos	Alta	Baixa	Baixa	Média
Acesso a Web	Alta	Média	Baixa	Média
Login remoto	Alta	Média	Média	Baixa
Áudio por demanda	Baixa	Baixa	Alta	Média
Vídeo por demanda	Baixa	Baixa	Alta	Alta
Telefonia	Baixa	Alta	Alta	Baixa
Videoconferência	Baixa	Alta	Alta	Alta

QoS

Bandwidth with no Quality of Service rules applied



Bandwidth with Quality of Service rules applied



Acordos de Nível de Serviço

Definição e cumprimento de **Acordos de Nível de Serviço (ANS)**

- Documento que define níveis de serviços acordados entre o cliente e o provedor de serviços.
- Descreve o serviço, os níveis de qualidade que devem ser garantidos, as responsabilidades das partes e eventuais compensações quando os níveis de qualidade não são cumpridos;
- Alguns indicadores comumente utilizados são a **Disponibilidade (Service Availability)**, **Largura de Banda**, o **Tempo de Resposta**, o **MTBF (período médio entre falhas)**.
- Deve ser escrito em linguagem que o negócio entenda de forma clara, concisa e livre de jargões.

