

Redes de Alta Velocidade e as Aplicações Multimídia

Autora: Regina Melo Silveira.

e-mail: regina@larc.usp.br

Neste pequeno tutorial é feita uma descrição da evolução e da situação atual das redes de alta velocidade e das aplicações multimídia implementadas sobre tais redes. Será dado um enfoque especial à rede ATM, à aplicação de Vídeo sob Demanda e ao sistema LDMD desenvolvido no LARC.

1. [Introdução](#)
2. [A rede ATM](#)
3. [Aplicações Multimídia](#)
4. [Video sob Demanda](#)
5. [O Laboratório Distribuído de Multimídia sob Demanda - LDMD](#)

1. Introdução

O desenvolvimento das tecnologias de redes de computadores tem se dado em ritmos acelerados nos últimos trinta anos. Tal desenvolvimento iniciou-se com um experimento acadêmico, o ARPANET em 1969, e desde então surgiram várias topologias e tecnologias que deram origem às redes de diferentes portes (LANs, MANs e WANs) e diferentes arquiteturas. Evoluções mais recentes em relação ao meio físico para a transmissão digital (fibras óticas) e equipamentos de conexão de rede computadorizados, criou subsídios para novas técnicas de transmissão que apresentam um desempenho muito melhor, maximizando a velocidade e minimizando as perdas de dados durante a transmissão. Essas novas tecnologias deram origem as redes de alta velocidade ou, como algumas vezes são chamadas, redes Gigabits e nos últimos três anos, as

Terabits.

As redes de alta velocidade surgiram principalmente devido a forte influência das redes públicas de telefonia e de telecomunicação. Tais redes sofreram um processo de migração da tecnologia analógica para digital, devido a necessidade de fornecer serviço de transmissão de voz a custos menores. No primeiro passo para essa evolução foi feito uma integração entre tecnologia de transmissão digital e chaveamento de circuito. Por isso, tais redes foram denominadas "Redes Digitais Integradas" ("Integrated Digital Networks" - IDNs). No entanto, a distribuição geográfica das redes telefônicas aliada a maior capacidade de transmissão, deixou clara a possibilidade de aproveitar essa mesma estrutura para transmissão de dados. Surgiu assim a idéia das redes ISDN ("Integrated Services Digital Network"), cujo nome se refere desta vez, à integração de serviços de comunicação.

A ITU-T (International Telecommunications Union), na época CCITT (Comité Consultatif International Télégraphique et Téléphonique), foi o órgão responsável pela padronização da estrutura e dos protocolos para a implementação do ISDN. Em 1988, nas primeiras etapas de elaboração das recomendações, surgiu um novo conceito com a proposta do ISDN Faixa Larga ou B-ISDN (Broadband ISDN), que foi definido pela ITU-T como "uma estrutura capaz de atender serviços que necessitem canais com taxas de transmissão superiores aos do canal primário estabelecido pelo ISDN". Essa nova estrutura foi proposta tendo em vista:

- a crescente demanda de serviços de banda larga;
- a disponibilidade de novas tecnologias de transmissão de alta velocidade, de chaveamento e de processamento de sinal;
- a melhoria da capacidade de processamento de dados e imagens na máquina cliente;
- o avanço no processamento de aplicações em software;
- a necessidade de integrar serviços interativos e de distribuição.

A proposta do B-ISDN veio, sem dúvida alguma, causar um impacto muito maior em termos de novas aplicações para redes de computadores, principalmente em aplicações que envolvam imagens ou vídeos, sejam elas interativas ou não. A figura 1[31] mostra a evolução das tecnologias de transmissão em relação as aplicações que utilizam mídias visuais. É possível perceber que a tecnologia digital fez crescer as possibilidades de aplicações que envolvem imagem e vídeo. Inicialmente havia recursos suficiente para

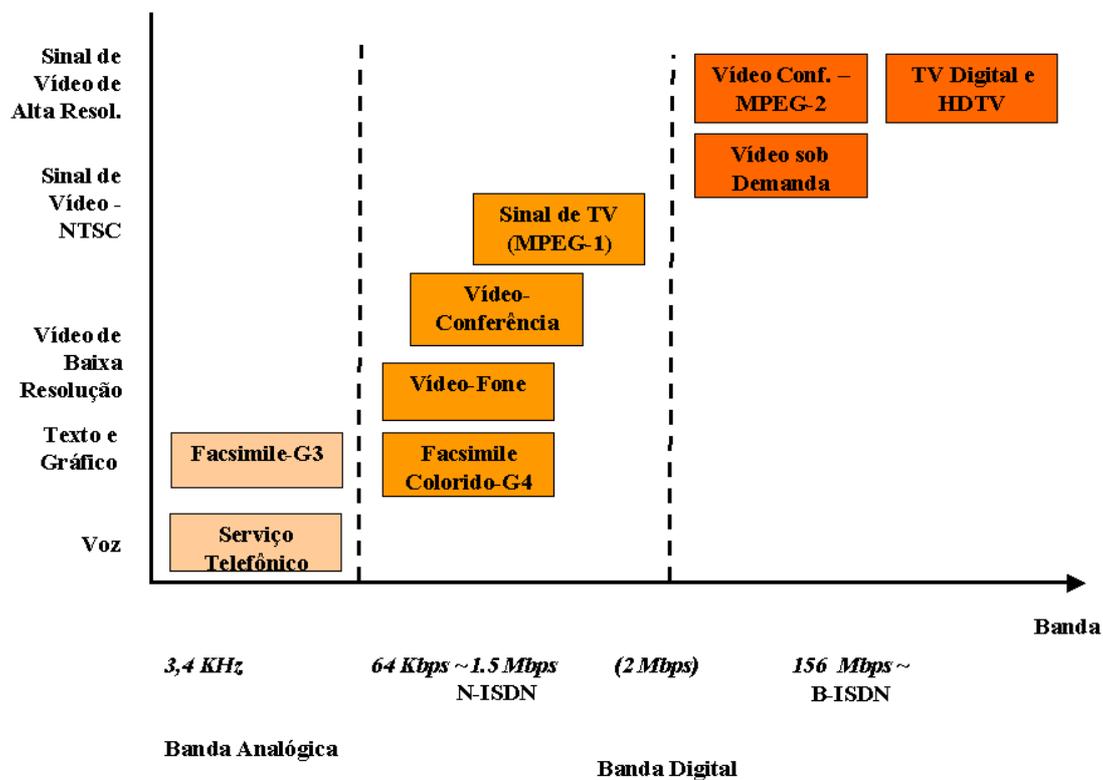


Figura 1: Evolução das tecnologias de transmissão em relação às mídias visuais.

transmitir somente vídeos de baixa resolução. Com a evolução da tecnologia ISDN para B-ISDN torna-se viável transmitir até mesmo vídeo ou sinal de TV de alta resolução, ou seja, HDTV (Hight Density Television). A partir do surgimento do B-ISDN, o conceito original do ISDN passou a ser referido como ISDN Faixa Estreita ou N-ISDN ("Narrowband ISDN"). Assim como a tecnologia de rede Frame Relay foi desenvolvida para trabalhar sobre o N-ISDN, a tecnologia ATM surgiu para atender aos requisitos de transmissão para o B-ISDN.



2. A Rede ATM

ATM (*Asynchronous Transfer Mode*) tem claramente um papel importante na evolução das redes "workgroup" e "enterprise". As redes baseadas em tecnologia ATM apresentam várias vantagens sobre as tecnologias utilizadas em redes LAN, MAN e WAN, dentre elas podemos citar a largura de banda escalável à custos menores, melhor desempenho e garantia de qualidade de serviço (QoS)[16]. A evolução causada por esta nova tecnologia de redes de computadores reside no fato de que tais características promoverem a utilização de novas classes de aplicações multimídia. Aplicações como vídeo-conferência, vídeo sob demanda,

trabalho compartilhado, telemedicina, tele-educação, entre outras aplicações para rede, impõem a utilização de uma quantidade grande de banda e, simultaneamente, é necessário ter uma infra-estrutura de rede que suporte uma grande gama de serviços.

É possível implantar enlaces ATM com diferentes larguras de banda, a partir de 1,5 Mbps, ou com valores bem maiores saltando para 25 Mbps, 155 Mbps (OC3), 622 Mbps (OC12) ou 2,488 Gbps (OC48), dependendo do meio de transmissão, que pode ser fibra monomodo ou multimodo[20]. As aplicações que utilizam ATM, podem ter requisitos de transmissão bem diferentes umas das outras, ou seja, enquanto uma aplicação requer latência muito baixa, outras não toleram perda de informações mas suportam atrasos.

A tecnologia ATM envolve a transferência de dados em pequenos pacotes (células) de 53 bytes, permitindo múltiplas conexões lógicas multiplexadas sobre uma interface física única [3]. A rede ATM necessita de um controle de erro e de fluxo mínimo, diminuindo o "overhead" em cada célula, assim como o "overhead" de processamento em cada nó da rede. Isso faz com que seja possível operar com uma taxa de transmissão de dados elevada. A taxa de perda de células na rede que utiliza meios físicos óticos é estimada entre 10^{-8} e 10^{-10} , dependendo da confiabilidade do sistema de transmissão, do comprimento do cabo, do tipo de técnica empregada pelo "switch" ATM e do algoritmo que ele utiliza para evitar congestionamento[23][50]. Apesar do bom desempenho obtido na rede, ATM é uma tecnologia extremamente complexa, devido ao conflito criado pela necessidade de otimizar a banda enquanto assegura diferentes QoS. Para isso, é necessário uma infra-estrutura baseada em protocolos que garantam o bom desempenho da rede e permitam sua interconexão com as demais tecnologias de redes já existentes[16]. Esses problemas são resolvidos com uma combinação de técnicas de gerenciamento de tráfego, que resultam em um controle bastante sofisticado. As funções do Gerenciador de Tráfego ATM (TM), foram criadas para prevenir a sobrecarga tanto da rede como do sistema final, e se isso ocorrer, as funções TM poderão reagir ao congestionamento de forma a conseguir manter os requisitos de QoS ao mesmo tempo que maximiza a utilização dos recursos da rede[19].

O TM implementa um conjunto de funções para o controle do tráfego e congestionamento das conexões. Os três módulos principais que constituem o TM são[3]:

- Gerenciamento de recursos utilizando caminhos virtuais;
- Controle de Admissão de Conexão (CAC);
- Controle dos Parâmetros de Utilização (UPC).

O gerenciamento de recursos utilizando vias virtuais comanda as conexões lógicas na rede ATM. Essas conexões lógicas são chamadas de "Conexões de Canal Virtual" (VCCs). Uma aplicação negocia um contrato de tráfego com a rede para cada conexão virtual. O contrato de tráfego é um acordo a respeito do comportamento do tráfego e o tipo de serviço que a conexão deve suportar, e seu elemento chave é a categoria de serviço. Os Canais Virtuais (VC) são alocados em uma "Conexão de Via Virtual" (VPC), onde todos os VCCs de um mesmo VPC tem o mesmo nó de origem e de destino dentro da rede. Para chegar ao seu destino final, normalmente, uma célula ATM percorre vários VPCs[3]. O mecanismo de canais virtuais tem como vantagem simplificar a arquitetura, aumentar o desempenho e

reduzir o processamento na rede.

No estabelecimento da conexão usando VPCs, é feita uma negociação entre as aplicações das pontas, ou seja, dos usuários finais, onde é verificado se a rede pode satisfazer a qualidade de serviço desejada. A rede procura alocar VCCs que possibilitem o estabelecimento da conexão, ou quando necessário, solicita mais capacidade. Caso não seja possível atender a QoS desejada a conexão é rejeitada como mostra o diagrama de conexão da figura 2.

Em outras palavras, uma conexão entre o ponto de origem e o ponto de destino do tráfego passa por vários "switchs" ou nós da rede. Em cada um desses nós as células devem entrar em uma fila para serem processadas no "switch". Para estabelecer uma conexão que trace um determinado caminho na rede, é necessário fazer a reserva de recursos em cada um dos pontos de fila, para garantir a QoS especificado no contrato de tráfego. Tal reserva de recursos geralmente compreende alocação de banda e de espaço no "buffer" nos "switchs".

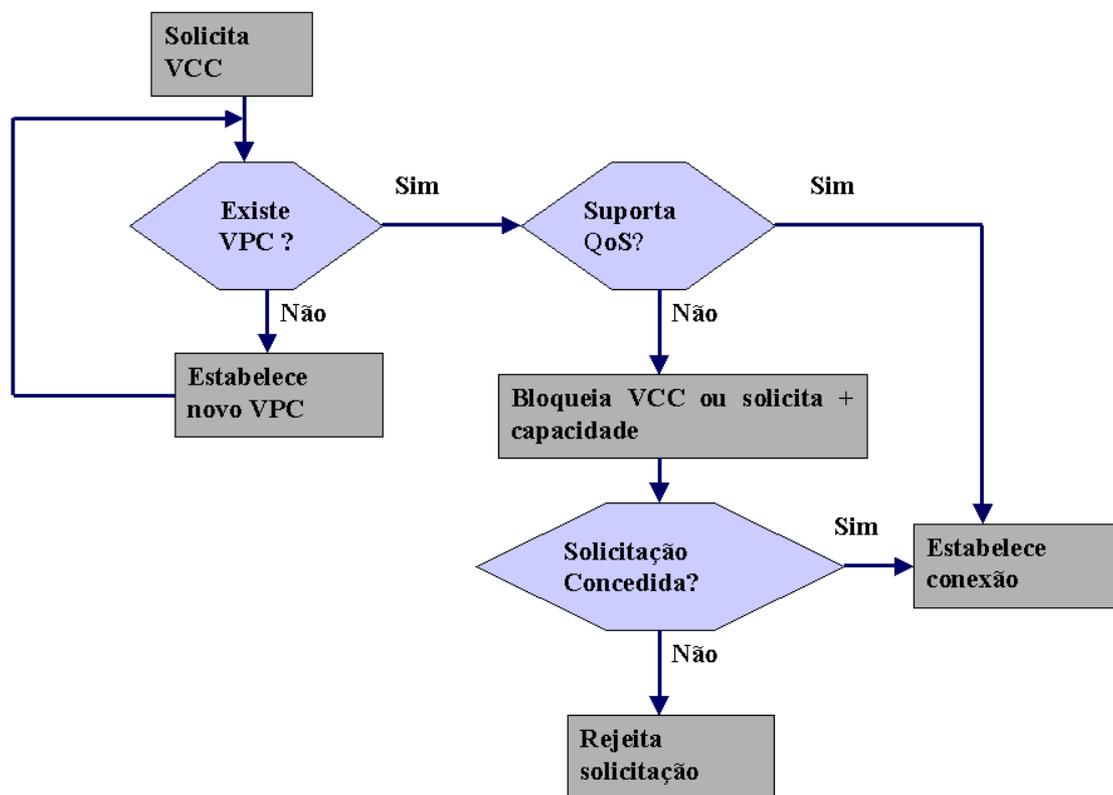


Figura 2: Diagrama de estabelecimento de conexão utilizando VP.

A sinalização é o processo utilizado para o estabelecimento, monitoramento e rompimento de conexões ATM. Através desse processo são enviadas mensagens para todos os nós da rede. Essas mensagens podem ser por exemplo SETUP, CONNECT ou DISCONNECT. As informações de sinalização são transportadas em conexões próprias, separadas das conexões utilizadas para transporte de informações de usuários. Esse tipo de procedimento, comumente

denominado "out of band signalling", é uma característica herdada das redes de comutação rápida de pacotes[30].

O conjunto de regras (ou procedimentos) que determina se a conexão será ou não estabelecida é comandado pelo CAC. Como foi mencionado anteriormente, quando um usuário requisita um novo VPC ou VCC, ele deve especificar (explícita ou implicitamente) o serviço que está sendo requisitado. Esta requisição, controlada pelo CAC, consiste nas seguintes indicações:

- Categoria de serviço em que a conexão deverá se enquadrar;
- Descrição do tráfego a ser transmitido;
- Parâmetros de QoS.

A rede ATM é capaz de transmitir diferentes tipos de tráfego, como voz, vídeo e dados, simultaneamente. No entanto, a maneira como a rede lida com cada um desses tráfegos depende das características do fluxo de dados e dos requisitos da aplicação. A classificação dos possíveis tipos de tráfegos na rede, ou seja, das categorias de serviço, é feita pela "Interface Usuário-Rede" (UNI). Assim, são definidas as seguintes categorias de serviço para a rede ATM:

- Serviços em tempo-real: aplicações que envolvem informações que pretendem ser reproduzidas de maneira contínua, portanto são sensíveis a atrasos e variação de atrasos:
 - Taxa de transmissão constante (CBR): utilizado em aplicações que requisitem a transmissão de uma taxa de dados fixa, que é continuamente disponibilizada durante a existência da conexão. Ideal para aplicações de áudio e vídeo não comprimidos, transmissão de voz, "*multicast*" e "*broadcast*" de vídeo e áudio não comprimido ou comprimidos com taxa constante.
 - Taxa de transmissão variável em tempo real (rt-VBR): aplicações que são sensíveis a atrasos estão nesta categoria. Sinais de áudio e vídeo comprimidos apresentam taxa de dados variáveis. Neste tipo de serviço a banda de transmissão é alocada dinamicamente, de maneira assíncrona, de acordo com a demanda requerida no momento. Desta forma, é possível que o transporte das células seja feito através da multiplexação estatística ou determinística com células de outros canais virtuais de mesma rota na rede, aproveitando melhor os recursos disponíveis.
- Serviços em tempo-não-real: serviços que atendem aplicações cujo tráfego se caracteriza pela presença de rajadas e que não sejam sensíveis a atrasos.
 - Taxa de transmissão variável em tempo-não-real (nrt-VBR): Este tipo de serviço é utilizado por aplicações que gerem tráfego variável na rede mas que não são sensíveis a atrasos. Assim, a rede pode disponibilizar banda dinamicamente garantindo relativamente pequeno atraso e baixa taxa de erro.
 - Taxa de transmissão disponível (ABR): a aplicação que utiliza esta categoria de

serviço especifica a taxa máxima de células (PCR) e a taxa mínima de células (MCR), assim, a rede aloca recursos que atenda pelo menos o MCR.

- Taxa de transmissão não especificada (UBR) – Serviço utilizado por aplicações que tolerem variações de atraso e perda de células. É conhecido também por serviço do melhor esforço e é utilizado para transmissão de textos, dados, imagens, mensagens e para terminais remotos.

Portanto, a categoria de serviço mais adequada para transmissão de sinal de vídeo é o rt-VBR, já que por se tratar de transmissão de sinais digitais de vídeo comprimido, seus quadros não apresentam volume constante de dados.

A qualidade de serviço é a principal característica implementada pela rede ATM. Nesta tecnologia é possível definir diferentes QoS para os diversos serviços (ou aplicações). Um serviço é considerado como tendo uma determinada qualidade quando ele atende os pré-requisitos de um conjunto de parâmetros mensuráveis. A rede deve ser capaz de dividir a banda entre os usuários e assegurar que o tráfego solicitado por qualquer dos usuários da rede não afete a QoS de outro usuário[16]. Dependendo do modelo de tráfego utilizado, os procedimentos do CAC podem ser bastante conservativos em relação a alocação de recursos, o que pode reduzir o ganho estatístico. Um CAC eficiente produz o máximo de ganhos estatísticos sem violar a QoS. Essa eficiência dependerá de quão próximos da realidade chegam os modelos de tráfego e de fila.

Um usuário de uma VCC especifica a QoS e a descrição de tráfego através de parâmetros pré-definidos. Para o serviço de Vídeo sob Demanda (VoD) esses parâmetros são:

- Parâmetros de Qualidade de Serviço (QoS):
 - Variação de atraso da célula (CDV) – variação entre o tempo de chegada de uma célula em um nó da rede;
 - Máximo atraso de transmissão da célula (CTD) – tempo entre a transmissão do último bit de uma célula na UNI da fonte e o recebimento da primeiro bit da mesma célula na UNI destino.
 - Taxa de perda de células (CLR) – a perda de células ocorre quando há "overflow" no buffer interno de um swicht da rede ATM ou quando ocorrem erros de transmissão. O CAC deve ter estratégias para evitar que a taxa de perda seja grande. A taxa é calculada através da relação (núm. de cél. perdidas/ total de cél. transmit.).
- Descritores de tráfego (DoT):
 - Taxa Máxima de células (PCR) – define o limite máximo de células que pode ser submetida por uma fonte de tráfego a conexão ATM. O PCR é definido em termos da variável T, que expressa o mínimo espaço entre células, portanto $PCR=1/T$.
 - Taxa média de células (SCR) – define o limite superior da taxa média de uma

conexão ATM, calculado sobre uma escala de tempo que é bem grande em relação a T. O descritor SCR só faz sentido se $SCR < PCR$.

- Tamanho máximo da rajada (MBS) – é o número máximo de células que pode ser enviado continuamente para uma dada taxa máxima de células.

Após ter sido definido o contrato de tráfego para a conexão, a rede aplica um algoritmo de admissão de conexão para avaliar se a conexão pode ser admitida respeitando a QoS desejada. Se a conexão é finalmente aceita e estabelecida, então as células começam a ser transmitidas e as funções do UPC passam a monitorar a rede para garantir que o tráfego está de acordo com o contrato. Neste trabalho essas funções não serão detalhadas, pois fogem ao enfoque principal. Maiores informações a respeito destas funções podem ser encontradas em [19]e[3].

Até aqui foi tratada a negociação de parâmetros, isto é, QoS e os descritores de tráfego, para o estabelecimento da conexão inicial na transmissão de dados sobre rede ATM. Mas, devido ao tráfego VBR apresentar características não estacionárias, como por exemplo, em função das variações do grau de atividade nas cenas de um vídeo ou da interatividade, geralmente não é possível definir um único conjunto de parâmetros UPC que resulte na qualidade uniforme do vídeo durante toda a sessão[29]. Normalmente o que ocorre é um super dimensionamento dos parâmetros UPC inicialmente declarados. Para evitar essa situação, o serviço oferecido para esse tipo de aplicação deve suportar renegociação de banda. Existem algumas propostas de esquemas para renegociação da banda, ou alocação dinâmica de banda durante as transmissões sobre ATM. Apesar desses esquemas não estarem ainda implementados nos equipamentos comercialmente disponíveis, eles criam meios poderosos de otimização da utilização dos recursos da rede. As duas principais propostas nesse sentido são:

- Alocação dinâmica estatística: o gerenciamento de tráfego (TM) utiliza um algoritmo estatístico para verificar a utilização da rede. Caso as características do tráfego de uma conexão mude significativamente, é feita uma renegociação baseado na nova caracterização do tráfego.
- Alocação dinâmica determinística: o TM utiliza uma função previamente definida para descrever o comportamento daquele tráfego em questão, e a partir desta faz-se a renegociação[21][22][8].

Vários trabalhos definem modelos estatísticos ou determinísticos para tais implementações.

2.1. Admissão de Conexão VBR

Como os fatores de admissão de conexão da categoria de serviço VBR irão influenciar a especificação do SoQ proposto neste trabalho, será um pouco mais detalhado os mecanismos e critérios utilizados.

A categoria de serviços VBR é principalmente especificada para um suporte mais eficiente às aplicações de vídeo[19], ou qualquer outra aplicação que tenha um comportamento de tráfego conhecido, ou previsto, por características de rajadas. Como já foi mencionado, o tráfego VBR é caracterizado pela taxa de células média (SCR), pela taxa de pico de células

(PCR) e pelo tamanho máximo da rajada (MBS). O SCR é medido sobre um período de tempo pré-definido e representa a taxa de transmissão média. O PCR representa o espaço mínimo entre as células, que é o pico de largura de banda necessária. A transmissão de vídeo pode aumentar devido a uma variação da ação em uma cena do vídeo. Neste caso, a flutuação dinâmica da taxa de transmissão não deve exceder aos valores pré-determinado de PCR e SCR. Geralmente, a taxa de transmissão média permite a realização de ganhos estatísticos. Isso pode ser feito alocando estatisticamente uma quantidade de largura de banda menor que o pico de banda e maior ou igual a largura de banda média. O tráfego rt-VBR apresenta restrições nos requisitos de atraso e, portanto, não pode ser extensivamente "bufferizado" na rede. As diferenças na quantidade de "buffer" disponível em cada nó da rede tem grande impacto sobre a quantidade de banda alocada.

A unidade de medida dos parâmetros PCR e SCR é células/segundo, e do MBS é segundos[30]. Se, em um determinado nó da rede, o CAC não consegue alocar a largura de banda necessária para suportar a descrição de tráfego requisitado, o nó pode reduzir tais valores de descrição. Entretanto, o valor dos descritores de tráfego não podem ser reduzidos a valores menores que os especificados como o mínimo aceitável na mensagem de sinalização. Se esses valores, quando manipulados pela rede, ultrapassam os valores mínimos especificados, então a conexão não é efetuada, sendo abortada ao ocorrer o "time-out".

Os parâmetros de QoS CDV e Max-CTD são expressos em microsegundos, enquanto o CLR é expresso em termos da magnitude de "n" ($1 < n < 15$), onde o CLR assume o valor de 10^{-n} . Como a mensagem de sinalização atua em cada ponto da rede, o Max-CTD e o CDV são acumulados e carregados na mensagem de sinalização, por todo o percurso de alocação de recurso. Quando a mensagem de sinalização alcança o último nó, os parâmetros de QoS acumulados representam os valores fim a fim que a rede pode garantir para a conexão. O estabelecimento da conexão será rejeitada se esses valores excederem os valores requisitados.

O parâmetro CLR não é acumulado explicitamente. Cada elemento da rede aceita ou rejeita a chamada de conexão dependendo do compromisso entre a taxa de perda suportada por ele e o CLR requisitado. A taxa de erro de uma conexão é definida pela maior taxa de erro de seus enlaces.



3. Aplicações Multimídia

As aplicações multimídia têm sido desenvolvidas nos últimos anos graças ao grande desenvolvimento ocorrido na área da computação. Alguns dos fatores que fortemente influenciaram tais aplicações são o aumento da capacidade de processamento, o aumento da capacidade de armazenamento, o surgimento de mídias digitais como o CD-ROM e o DVD e a melhoria da qualidade de resolução dos monitores, tudo isso aliado a custos mais baixos desses equipamentos eletrônicos. No entanto, até pouco tempo, as aplicações multimídia se limitavam a aplicações "stand alone", a redes locais ou a CD-ROMs. Com as redes de alta velocidade, especialmente com as redes ATM, que implementam garantia de QoS, as aplicações multimídia de acesso remoto ganharam novo estímulo. Aqui será discutido quais

são os requisitos para a transmissão e os critérios de desempenho para tais aplicações.

3.1. Requisitos para transmissão de material multimídia

Para que seja possível disponibilizar material multimídia online em ambiente de rede corporativa, é necessário criar uma estrutura cliente/servidor capaz de, no servidor, armazenar, transmitir e integrar todas as mídias envolvidas, e no cliente receber e reproduzir adequadamente tais mídias de forma sincronizada. Para o servidor é necessário arquitetar um sistema que comporte todo o conteúdo multimídia.

As mídias podem ser divididas em dois grupos[23]:

- Mídias discretas, onde estão envolvidas apenas dimensões espaciais (textos, hipertextos e imagens);
- Mídias contínuas, que envolvem tanto dimensões espaciais como temporais (áudio e vídeos).

O armazenamento e distribuição das mídias discretas não acarretam dificuldades devido ao pequeno volume de dados que essas mídias apresentam. Para as mídias contínuas, devemos considerar que o sistema pretende implementar transmissão em tempo real, isso é, os dados de áudio e vídeo serão transmitidos "por demanda". O áudio, apesar de ser uma mídia contínua, apresenta baixas taxas de transmissão, principalmente quando empregado uma técnica de compressão, como mostra a tabela 1. Já os vídeos digitalizados, sejam imagens em movimento ou animações gráficas, apresentam grandes volumes de dados, como ilustra a tabela 2.a, e, mesmo utilizando técnicas de compressão (tabela 2.b), a dificuldade em seu armazenamento e transmissão requer certos cuidados.

Por esse motivo, ou seja, pelo fato de haver dificuldades na distribuição e armazenamento de vídeos digitais, neste trabalho será tratado exclusivamente dos sinais de vídeo, não entrando no mérito do sinal de áudio ou de qualquer mídia discreta.

Qualidade de Áudio	Especificação	Taxa de Tx.
Áudio com Qualidade de Voz	1 canal, amostras de 8-bit / 8kHz	64 Kbps
Voz Digitalizada	Padrão G.728, 3.4 kHz	16 Kbps
Áudio Monofônico	1 canal, amostras de 16-bit/44.1kHz	705.6Kbps
Áudio com qualidade de CD	2 canais, amostras de 16-bit/44.1kHz	1.411Mbps
Áudio codificado com	Equivalente a qualidade de CD	384 Kbps

MPEG		
------	--	--

Tabela 1 : Taxa de transmissão para áudio e áudio comprimido.

	Resolução	1 minuto	1 hora
a	640 x 480	1.6 GB	97 GB
	320 x 240	400 MB	24 GB
b	640 x 480	16 MB	970 MB
	320 x 240	4 MB	240 MB

Tabela 2: Valores estimados do volume do sinal de vídeo (a) não comprimido e (b) comprimido, em bytes.

Com os valores apresentados na tabela 2 fica clara a necessidade da utilização de uma rede que disponibilize banda larga de transmissão para aplicações que tenham sido elaboradas para a distribuição de vídeo com boa qualidade, isto é, qualidade igual ou superior ao do aparelho de vídeo cassete (VCR).

3.2. Critérios do Desempenho da Transmissão de Vídeo

A transmissão de vídeo apresenta muitas propriedades em comum com a transmissão de outras mídias, no entanto, como já mencionado, ela possui dois aspectos específicos, que são o grande volume de dados e a necessidade da transmissão síncrona e em tempo-real. Baseado nestas características, podemos estabelecer quatro critérios principais para avaliarmos quantitativamente o desempenho da comunicação estabelecida pela rede. Esses critérios são:

- atraso total, devido a busca no servidor e a transmissão;
- variação do atraso;
- taxa de transmissão;
- taxa de erro de transmissão.

Esses critérios de desempenho da rede influenciam fortemente a qualidade de um sistema de distribuição de vídeo. Para que o desempenho relacionado a esses critérios seja aceitável, é necessário empregar algumas estratégias como veremos a seguir.

Atraso Total: Quando um cliente deseja receber *online* um vídeo remoto, ele faz uma requisição ao servidor. O servidor de conteúdo, então, em resposta a essa requisição, faz uma busca em sua estrutura interna verificando onde se encontra armazenado tais dados. A partir daí, o servidor de vídeo começa a transmitir esses dados para o cliente. Tanto a busca do servidor, quanto a transmissão, causarão um certo atraso, que deve ser considerado. Para evitar interrupções na reprodução do vídeo, um buffer é alocado no cliente e a reprodução só

é iniciada quando esse buffer já estiver completo. Dimensionando o buffer adequadamente, é possível minimizar a perda de qualidade na reprodução devido ao atraso.

Varição do Atraso: A transmissão de dados em uma rede, esta sujeita a variação de atrasos que são introduzidos principalmente pelos equipamentos utilizados na transmissão e pelo compartilhamento de recursos realizado pela rede. Cabos, repetidores, chaveadores e roteadores inserem atrasos a transmissão do sinal digital. Portanto, a variação do atraso depende da tecnologia utilizada e da distância coberta pela rede. Como no caso do atraso total, o "buffer" alocado na máquina cliente, quando bem dimensionado, também atenuará tal efeito.

Taxa de Transmissão: Várias técnicas e padrões de compressão tem sido desenvolvidos nos últimos anos com a intenção de reduzir a taxa de transmissão de sinal de vídeo. Dentre eles podemos citar o Motion JPEG[23], MPEG-1[7], MPEG-2[10] e o H.261[24], que têm se mostrado os mais promissores. As técnicas de compressão utilizadas por esses padrões são, em sua maioria, com perdas, eliminando as redundâncias espaciais e temporais existentes. Portanto há uma redução da qualidade da imagem, muitas vezes não perceptível, para reduzir o volume de dados. Apesar de alguns deles conseguirem altas taxas de compressão e consequentemente, taxas de transmissões bem baixas comparadas às dos sistemas analógicos convencionais, é possível verificar que essas taxas de transmissão ainda são muito altas para aplicações que utilizem a Internet ou redes convencionais. A tabela 3 mostra as taxas obtidas com os padrões mencionados.

Padrão	Banda	Taxa Compressão
Motion JPEG	10-20 Mbps	7-27:1
MPEG-1	1.2-2.0 Mbps	100:1
H.261	64Kbps-2 Mbps	24:1
MPEG-2	4-60 Mbps	30-100:1

Tabela 3 : Requisito de banda para transmissão de vídeo.

Taxa de Erro: A taxa de erro diz respeito ao comportamento da rede em relação a alteração, perda, duplicação e entrega fora de ordem de dados ao serem transmitidos. Normalmente, esses prejuízos são evitados com a implementação, nos protocolos da tecnologia da rede utilizada ou na aplicação, de funções de detecção, notificação e correção de erro.



4. Vídeo sob Demanda

O serviço de Vídeo sob demanda (VoD) foi definido pelo "ATM Forum" como [1]:

"Serviço assimétrico que envolve várias conexões, transferindo informações de vídeo digital, comprimido e codificado, de um servidor (tipicamente um servidor de vídeo), para um cliente (tipicamente um "Set Top Terminal" ou um PC). No destino o vídeo é descomprimido, decodificado, convertido de digital para analógico e apresentado em um monitor."

Esse tipo de sistema pode ser considerada a aplicação mais promissora entre todas da nova geração de aplicações multimídia. Com esse sistema é possível acessar uma base de dados, escolher um filme, e assisti-lo, com transmissão em tempo real. O sistema deve permitir a utilização das funções típicas de um equipamento de vídeo cassete, ou seja, pausar, retornar, ir para frente, sem perder a característica de sincronização e de tempo-real. Além disso, o VoD deve disponibilizar vídeos de qualidade igual ou superior ao do material gravado em fitas de vídeo (VCR). Por enquanto esse sistema tem sido implementado de forma experimental em laboratórios de pesquisa e em pequenos domínios como hotéis e condomínios, simulando uma vídeo locadora digital. Com a implantação de redes de alta velocidade de longa distância, este sistema deverá nos próximos anos ser implementado em redes metropolitanas corporativas, sendo aplicadas não só com funções de entretenimento, como também em divulgações mais ampla de material de vídeo com outras características, como por exemplo de ensino remoto.

Tecnicamente, trata-se de um sistema bastante complexo, não só pelo fato de transmitir vídeo em tempo-real, como também pelo fato de manter sua eficiência ao atender um grande número de usuários ao mesmo tempo. Atendendo inclusive, vários usuários que acessem simultaneamente o mesmo vídeo.

Para isso é necessário criar uma arquitetura hierárquica de servidores associado a sistemas de armazenamento de forma eficiente, levando-se em consideração vários aspectos, que impõem um grande número de desafios tecnológicos. Muitos pesquisadores aceitaram esse desafio e tem publicado trabalhos mostrando pesquisas que procuram definir uma arquitetura adequada para o sistema[32], desenvolver mecanismos de recuperação mais eficiente dos dados armazenados[33], melhorar o desempenho do sistema[34], criar mecanismos de gerenciamento específico[35], estudar a otimização da transmissão de vídeo[36][37] ou criar mecanismos mais eficientes para a indexação de vídeo[38].



5. O Laboratório Distribuído Multimídia sob Demanda – LDMD

O sistema de vídeo sob demanda do LARC foi arquitetado de forma a apresentar uma distribuição hierárquica dos recursos utilizados, garantindo melhor desempenho. Sendo assim, foi elaborado a partir de uma rede ATM, utilizando um switch ATM ligado a dois

servidores, um servidor de arquivo e um servidor de vídeo, como é mostrado na fig.4. O servidor de arquivos, ou de conteúdo, gerencia o acesso a base de dados e localiza o conteúdo escolhido pelo cliente. Após a localização, caso o arquivo esteja armazenado na memória terciária, este é transferido para o servidor de vídeo para que então seja transmitido ao cliente.

O servidor de vídeo armazena os vídeos mais acessados, enquanto os demais ficam armazenados em uma memória terciária. Em nosso sistema um *juke-box* será utilizado com essa função. Quando um vídeo solicitado por um usuário se encontra neste armazenador, o servidor de conteúdo se incumbem de verificar a posição do vídeo no juke-box e recuperá-lo, colocando-o no servidor de vídeo para que então possa ser transmitido. Toda essa estrutura está ligada por meio de links de 155 Mbps.

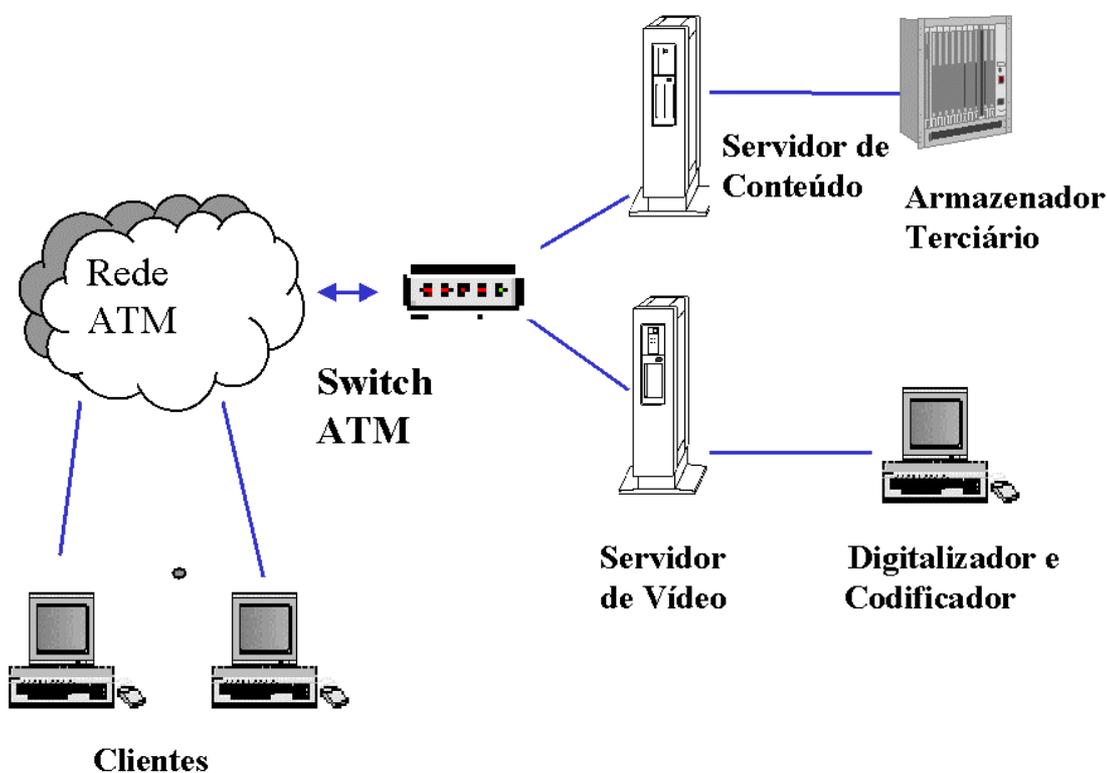


Figura 4: Arquitetura do sistema LDMD com distribuição hierárquica.

Se os arquivos de vídeo forem colocados em um único disco, seu acesso ficará restrito a capacidade de vazão do disco, por isso, o servidor de vídeo possui "*disk array*" (múltiplos discos). Para que seja possível um número maior de acessos concorrentes deve ser utilizada uma técnica que faça a divisão dos arquivos de vídeo em múltiplos blocos alocados em diferentes discos. O tamanho do bloco deve ser escolhido de forma a balancear a eficiência dos discos e a transmissão mas, uma vez determinada, esta dimensão será aplicada a todo o sistema. Os blocos são armazenados nos *slots* sequencialmente. Numa exibição típica, os

discos lêem os blocos numerados e os entregam à rede para o envio ao respectivo cliente, essa técnica é denominada "data interleaving"[2], como mostra a tabela 2.4 permitindo que vários usuários acessem o mesmo vídeo em tempos subsequentes,

Volta	Disco 1	Disco 2	Disco 3
1	Arq. A, bloco 1	Arq. B, bloco 1	Arq. C, bloco 1
2	Arq. C, bloco 2	Arq. A, bloco 2	Arq. B, bloco 2
3	Arq. B, bloco 3	Arq. C, bloco 3	Arq. A, bloco 3
4	Arq. A, bloco 4	Arq. B, bloco 4	Arq. C, bloco 4

Tabela 4: Método de leitura de arquivos armazenados com a técnica "data interleaved".

O cliente, tem acesso ao conteúdo do sistema a partir de uma interface com a base de dados, utilizando um browser, permitindo que seja feita a busca e seleção do material desejado.

Para que a transmissão e a reprodução do vídeo no equipamento do cliente seja bem sucedida, é necessário definir a taxa com que devem ser transferidas as células ATM, a quantidade de memória requerida no equipamento do usuário para armazenar as células já transmitidas, e o número de células que devem ser transmitidas antes do início da reprodução do vídeo. Para isso, algumas vezes é necessário um equipamento adicional no cliente, denominado "set-top box". Este equipamento possui memória (buffer) para o armazenamento temporário de algumas seqüências do vídeo escolhido, um decodificador e um software de reprodução do sinal de vídeo decodificado. O sucesso da transmissão e reprodução do vídeo depende principalmente dos fatores:

- Tipo de transmissão na rede ATM;
- Qualidade de serviço negociado na conexão ATM;
- Tipo de codificação do vídeo a ser transmitido.

Como já foi mencionado, a rede ATM possui opções para que a transmissão e a qualidade de serviço sejam escolhidos de acordo com a aplicação.