
Redes IP e tecnologias de nível 2

Arquitecturas / MPLS

FEUP/DEEC
Redes de Computadores
MIEEC – 2010/11
José Ruela

Evolução – X.25, Frame Relay e ATM

- A aprovação, em 1976, da primeira versão (*draft*) da recomendação X.25 criou as condições para a oferta de um serviço público de comutação de pacotes, que era baseado no modelo de circuitos virtuais
- O modelo de circuitos virtuais foi também adoptado no serviço *Frame Relay* (ou mais genericamente nos designados serviços de suporte em modo trama)
 - » O serviço *Frame Relay* surgiu como uma evolução (e um substituto) natural do X.25 e apresentava características que o recomendavam como alternativa vantajosa a circuitos dedicados para interligação de LANs
- O passo seguinte nesta evolução conduziu ao desenvolvimento da tecnologia ATM (igualmente baseada no modelo de circuitos virtuais), escolhida pelo ITU-T como solução alvo para a designada RDIS de Banda Larga
 - » Embora o conceito de RDIS de Banda Larga, como preconizado pelo ITU-T, não tenha chegado a concretizar-se, a tecnologia ATM foi adoptada por muitos operadores de Telecomunicações e fornecedores de serviços IP (ISP – *Internet Service Providers*) no núcleo das suas redes

Arquitecturas baseadas em ATM

- As principais características do ATM (elevada capacidade de comutação e baixa latência) justificaram assim a exploração das infraestruturas de rede instaladas por fornecedores de serviços IP para interligação de *routers* IP colocados na sua periferia
 - » Uma solução simples, do tipo *overlay* (redes sobrepostas), especificada pelo IETF, consistia em considerar o ATM como uma mera tecnologia de nível 2 (a exemplo do serviço MAC em LANs) – o que justificou a designação CLIP (*Classical IP and ARP over ATM*)
 - » Esta solução é igualmente aplicável a redes IP em ambiente LAN
- A introdução da tecnologia ATM em LANs foi também objecto de duas soluções especificadas pelo ATM Forum – LANE (*LAN Emulation*) e MPOA (*Multiprotocol over ATM*) – que igualmente se baseavam no modelo *overlay*
 - » Estas soluções pretendiam ter carácter genérico, não se restringindo a redes IP
- CLIP, LANE e MPOA, pela reduzida importância actual, são apenas referidas, a título informativo, no Anexo

ATM e IP – evolução

- A exploração de forma combinada das melhores características das redes IP e ATM motivou a pesquisa de soluções que promovessem a respectiva integração (e não a sua simples co-existência)
- Entre as soluções propostas, destaca-se a arquitectura MPLS (*Multiprotocol Label Switching*), especificada pelo IETF
- O MPLS combina técnicas de encaminhamento de nível 3 (típicas das redes IP) com técnicas de comutação de nível 2 baseadas em etiquetas (mas não necessariamente ATM), próprias da comutação de circuitos virtuais
 - » O MPLS tem um mecanismo próprio de etiquetagem que permite o transporte de tráfego IP sobre diversas tecnologias de nível 2 (ATM, LAN, PPP, etc.), mesmo que estas não sejam nativamente baseadas em etiquetas
- O MPLS introduz um plano de controlo unificado em redes IP e permite dotá-las com mecanismos de Engenharia de Tráfego, inexistentes nas redes IP convencionais

MPLS

Multiprotocol Label Switching

MPLS – Multiprotocol Label Switching

- O IETF desenvolveu uma arquitectura designada *Multiprotocol Label Switching* (MPLS), que é objecto de vários documentos, nomeadamente
 - » RFC 2702 – *Requirements for Traffic Engineering Over MPLS*
 - » RFC 3031 – *Multiprotocol Label Switching Architecture*
 - » RFC 3032 – *MPLS Label Stack Encoding*
 - » RFC 3035 – *MPLS using LDP and ATM VC Switching*
 - » RFC 3036 – *LDP Specification*
 - » RFC 3209 – *RSVP-TE: Extensions to RSVP for LSP Tunnels*
 - » RFC 3212 – *Constraint-Based LSP Setup using LDP*
 - » RFC 3270 – *MPLS Support of Differentiated Services*
 - » RFC 3564 – *Requirements for Support of Differentiated Services-aware MPLS Traffic Engineering*
 - » RFC 3945 – *Generalized Multi-Protocol Label Switching (GMPLS) Architecture*

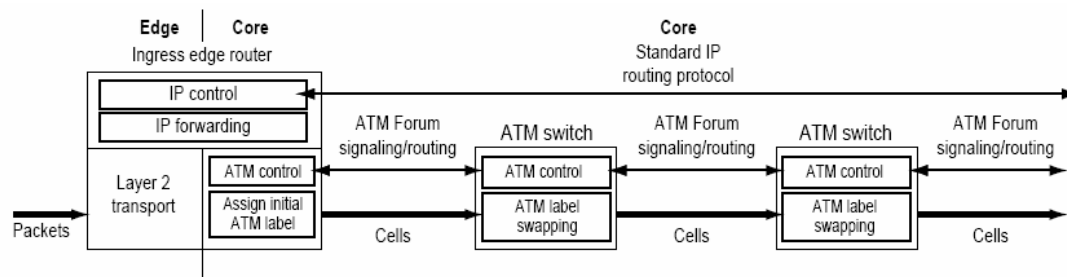
Necessidade

- Com o crescimento explosivo da Internet, os fornecedores de serviços IP necessitam de dispor de uma infra-estrutura de rede (*backbone*) de elevada capacidade e pequena latência, flexível, escalável, e que permita oferecer Qualidade de Serviço diferenciada
 - » A gestão destas redes requer mecanismos poderosos de Engenharia de Tráfego, isto é, a capacidade de mapear fluxos de dados na topologia física, de forma a distribuir o tráfego de forma equilibrada na rede, e assim conseguir utilização eficiente de recursos, evitar congestionamento e melhorar o desempenho global
- Em muitos casos as redes de fornecedores de serviço IP evoluíram para uma estrutura em dois níveis – um núcleo (*core*) de alta velocidade que interliga *routers* IP localizados na periferia (*edge*), onde se situa a inteligência de processamento
 - » A arquitectura CLIP (*Classical IP over ATM*), do tipo *overlay*, segue este modelo mas apresenta diversas limitações, o que justificou a pesquisa de novas soluções baseadas no conceito de comutação multi-camada (*multilayer switching*), de que o MPLS se constitui actualmente como referência
 - » As arquitecturas de comutação multi-camada integram técnicas de comutação baseadas em etiquetas (*label swapping and forwarding*), realizadas na camada de ligação de dados, com técnicas de encaminhamento características da camada de rede (*network layer routing*)

Modelo overlay – Classical IP over ATM (CLIP)

- Em ambiente WAN, a arquitectura CLIP caracteriza-se por dois tipos de equipamentos – *routers* IP e comutadores ATM – e duas topologias segregadas (topologia física ATM na qual é sobreposta a topologia lógica IP)
 - » Os *routers* e os comutadores não partilham informação topológica, o que impede uma optimização conjunta de recursos
 - » Cada rede tem os seus protocolos próprios de encaminhamento e sinalização e o seu espaço de endereçamento, com naturais duplicações e ineficiências
- Ao organizar os *routers* IP numa única LIS (*Logical IP Subnetwork*), para evitar múltiplos *routers* no percurso, cada *router* é adjacente lógico (*logical peer*) de qualquer outro
 - » O número de Circuitos Virtuais cresce com o quadrado do número de *routers*, o que levanta problemas de escalabilidade
 - » O conseqüente aumento do tamanho das tabelas de encaminhamento torna a sua gestão e actualização mais complexa
 - » O tráfego de controlo cresce com o aumento do número de adjacências entre *routers* e com a replicação com efeitos multiplicativos dos anúncios que recebem dos seus vizinhos
 - » Em caso de reconfigurações podem criar-se situações de instabilidade com tempos elevados de convergência para um estado estável

Classical IP over ATM



Redes IP – decisão de encaminhamento

- Em redes IP, por cada pacote que recebe, um nó (*router*) toma uma decisão independente quanto ao próximo passo do percurso (*hop-by-hop routing*), com base no endereço de destino (*destination based routing*) e no conteúdo actual da sua tabela de encaminhamento
- Este método (*destination based routing*) tem algumas limitações
 - » A decisão é tomada exclusivamente com base em informação presente no cabeçalho do pacote, excluindo-se assim critérios baseados noutra tipo de informação para além da explicitamente transportada no pacote
 - » O cabeçalho transporta mais informação do que a que é utilizada para seleccionar o percurso (*next hop*), o que significa que nem mesmo é explorada a possibilidade de usar outra informação disponível no cabeçalho (e.g., campo ToS – *Type of Service*)
- O processamento do cabeçalho é relativamente complexo (quando comparado com o processamento de um identificador de circuito virtual, ou de etiquetas na terminologia MPLS)
 - » *Look-up* da tabela de encaminhamento
 - » Decremento do campo TTL (*Time To Live*)
 - » Computação dum novo *checksum*

Redes IP – algoritmos de encaminhamento

- Os algoritmos de encaminhamento interior usados em Sistemas Autónomos IP (*Autonomous Systems – AS*) são baseados em métricas de menor custo
 - » No caso habitual em que se usa uma métrica aditiva baseada em pesos atribuídos às ligações que formam o percurso, os algoritmos calculam os percursos mais “curtos” (*shortest path*) ou de menor “custo”
 - » Os percursos (*shortest path routes*) de vários nós de origem para um dado nó de destino formam uma árvore com raiz no nó de destino
- Uma consequência é que as ligações que fazem parte de uma árvore podem ficar congestionadas, enquanto outras ligações se podem manter subutilizadas
 - » Isto acontece porque não é possível explorar percursos alternativos (a alteração dinâmica de métricas poderia ser uma solução, mas é demasiado complexa e pode causar efeitos indesejáveis, por exemplo, instabilidade do algoritmo)
 - » A utilização de percursos alternativos permitiria reencaminhar tráfego de modo a evitar zonas congestionadas, distribuir carga de forma mais equilibrada pelas ligações e realizar a transferência de tráfego para percursos de reserva em caso de falhas
- Nas redes IP convencionais não existem mecanismos de Engenharia de Tráfego e não é possível providenciar rotas que suportem requisitos específicos de fontes ou serviços ou sujeitas a restrições de Qualidade de Serviço (QoS)

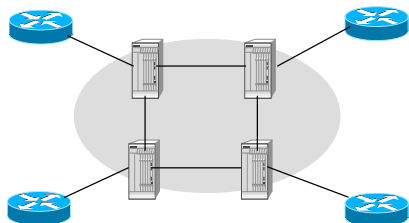
MPLS – objectivos

- O MPLS incorpora princípios básicos e ideias adoptadas em arquitecturas multi-camada desenvolvidas por fabricantes, tais como o *IP Switching* (Ipsilon) e o *Tag Switching* (Cisco), sendo muito influenciado por este
 - » Embora o objectivo de integração de IP com ATM esteja na origem do desenvolvimento destas arquitecturas, o MPLS tem um carácter mais genérico, permitindo o suporte de múltiplos protocolos de rede (para além do IP) sobre um mecanismo de comutação de etiquetas (não necessariamente baseado em ATM)
 - » Na prática a adopção de MPLS tem sido orientada ao suporte de IP, sendo o ATM um candidato (não único) para providenciar a comutação de etiquetas
- Ao contrário dos modelos *overlay*, nas arquitecturas multi-camada existe uma única topologia e uma família de equipamentos que partilham informação topológica e executam o mesmo conjunto de protocolos, reduzindo-se assim o número de adjacências entre nós, que é agora da ordem de grandeza do que é habitual em redes de comutação de pacotes parcialmente emalhadas

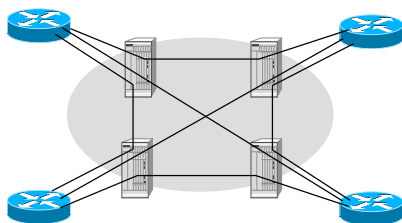
CLIP versus MPLS – topologias física e lógica

CLIP

- Topologia física

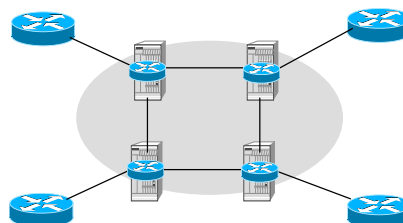


- Topologia lógica (IP)

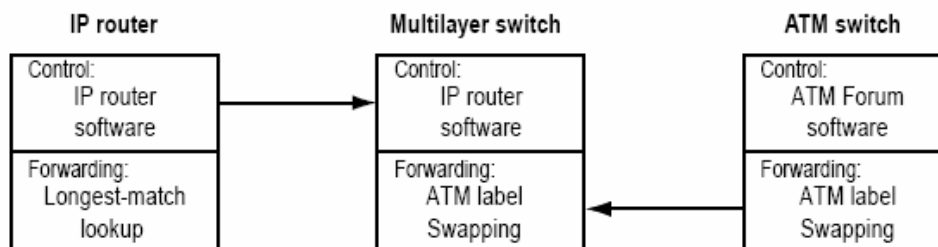


MPLS

- Topologia física e lógica

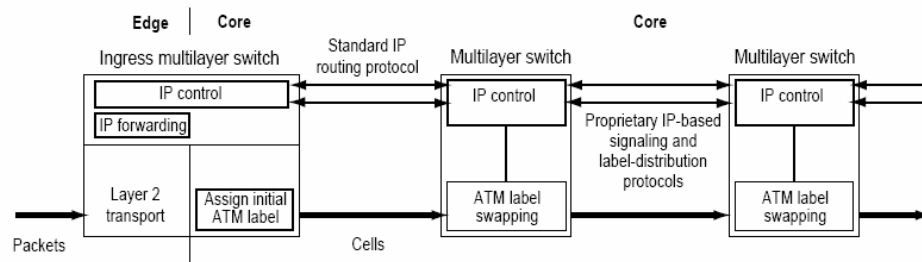


Comutação multi-camada – perspectiva

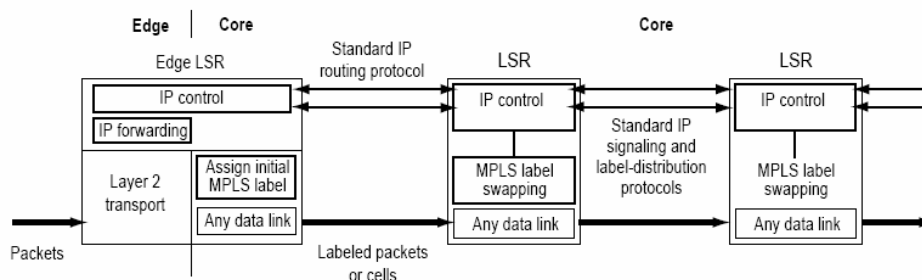


Comutação multi-camada – evolução

- Soluções de fabricantes



- MPLS



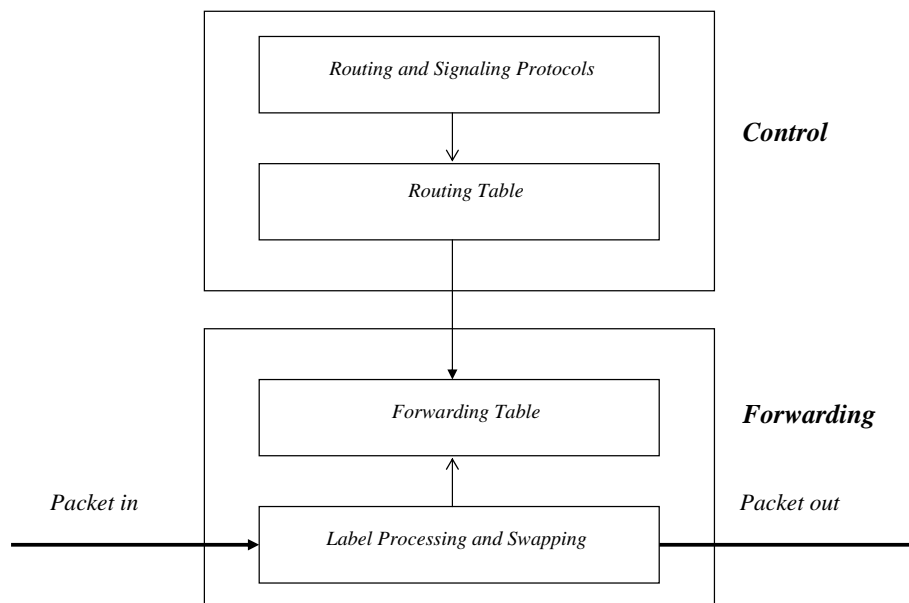
MPLS – princípios arquitetónicos

- As arquiteturas de comutação multi-camada (*multilayer switching*), de que o MPLS se constitui como modelo, têm como ideia base combinar
 - » Técnicas simples e robustas de encaminhamento na camada de rede (de que o paradigma é o IP e protocolos associados) – *Layer 3 Routing / Control*
 - » Técnicas de comutação rápida, eficientes e escaláveis, na camada de ligação de dados (de que o ATM é a principal referência) – *Layer 2 Forwarding / Switching*
- Em MPLS as funções de Controlo e de Transporte de dados são separadas
 - » As funções de Controlo, realizadas em *software*, baseiam-se em algoritmos e protocolos de encaminhamento convencionais (normalizados) e em protocolos de sinalização adequados aos requisitos da arquitetura (podendo ser adaptações ou extensões de protocolos existentes)
 - » As funções de Transporte (*forwarding / switching*), realizadas em *hardware*, baseiam-se em técnicas de comutação de etiquetas
- Torna-se assim necessário estabelecer uma associação entre informação de nível 3 (rotas) e informação de nível 2 (etiquetas) – *label binding* – o que justifica a designação comutação multi-camada

Componentes de Controlo e de Transporte

- Num domínio MPLS os pacotes são comutados exclusivamente com base numa etiqueta que transportam num cabeçalho adicional, sendo este processo realizado pela componente de Transporte
 - » Nos nós de entrada (*ingress nodes*), a cada pacote é apenas uma etiqueta; nos nós internos a etiqueta é processada e trocada (*label swapping*); a etiqueta é removida nos nós de saída (*egress nodes*)
- A componente de Controlo, com recurso a protocolos de encaminhamento e sinalização, mantém informação topológica necessária para a construção e manutenção das tabelas de encaminhamento e de comutação
 - » As rotas usadas para o transporte de pacotes no interior dum domínio MPLS, entre nós de entrada e de saída, são determinadas com base em protocolos de encaminhamento ou por outros meios (e.g., encaminhamento explícito)
 - » Em cada nó, a informação sobre rotas mantida numa tabela de encaminhamento (*routing table*) é usada para construir a tabela de comutação (*forwarding table*) que associa um par <porta, etiqueta> na entrada a um par <porta, etiqueta> na saída, realizando-se deste modo o mapeamento entre rotas e etiquetas (*label binding*)
 - » Para manter a consistência das tabelas de comutação dos nós, é necessário um protocolo de distribuição de etiquetas (LDP – *Label Distribution Protocol*)

Componentes de Controlo e de Transporte



Conceitos e definições

- *Label* (etiqueta) – identificador de comprimento fixo e pequeno, definido num espaço de valores contíguos, com significado local, usado para identificar fluxos de pacotes (designação equivalente – *tag*)
- *Forwarding Equivalence Class* (FEC) – grupo de pacotes (nível 3) transportados de forma idêntica (pelo mesmo percurso, sujeitos ao mesmo tratamento no que se refere ao despacho), pelo que podem ser mapeados na mesma etiqueta
- *Label Switched Hop* – salto (*hop*) entre dois nós MPLS adjacentes entre os quais a transferência de pacotes é baseada numa mesma etiqueta
- *Label Switched Path* (LSP) – percurso (rota) criado pela concatenação de *Label Switched Hops*, isto é, um percurso entre nós de entrada e saída num domínio MPLS, ao qual está associado uma sequência ordenada de etiquetas, o que permite transporte de pacotes entre nós MPLS pela simples troca de etiquetas
- *Stream* – agregado de um ou mais fluxos tratados de forma semelhante para efeito de transporte na rede e descritos por uma mesma etiqueta
- *Stream merge* – agregação de fluxos (*streams* / LSPs) num ponto de confluência (*merge point*), a partir do qual o agregado partilha o mesmo percurso e etiqueta
- *Label stack* – conjunto ordenado de etiquetas no qual se realizam funções *Push* e *Pop* a que corresponde um modo de operação do tipo LIFO (*Last In First Out*)

Critérios para definição de FEC

São possíveis vários critérios para definir um FEC; estes critérios deverão ser tidos em consideração no processo de classificação de pacotes e respectiva associação a um FEC

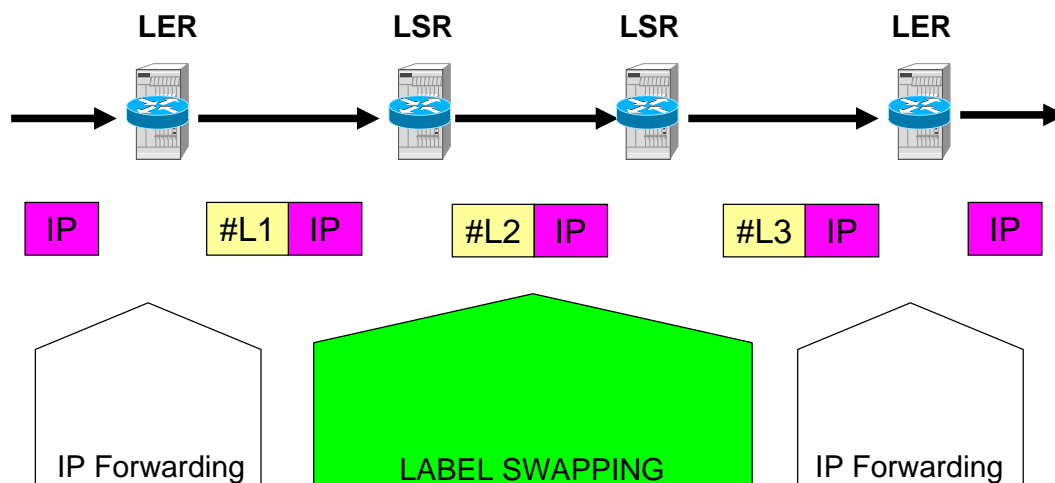
- Prefixo do endereço de destino – corresponde ao critério actualmente usado em redes IP (*destination based routing*)
- Classe de Serviço (CoS) – baseado em campos do cabeçalho IPv4 (e.g., *IP Precedence, Type of Service*) ou IPv6 (e.g., *Flow*)
- Fluxo de Aplicação (baseado em endereços de origem e destino e informação adicional das camadas de Rede e de Transporte)
- Grupo *multicast*
- Encaminhamento explícito
- *Virtual Private Network* (VPN)

Label Switching Routers

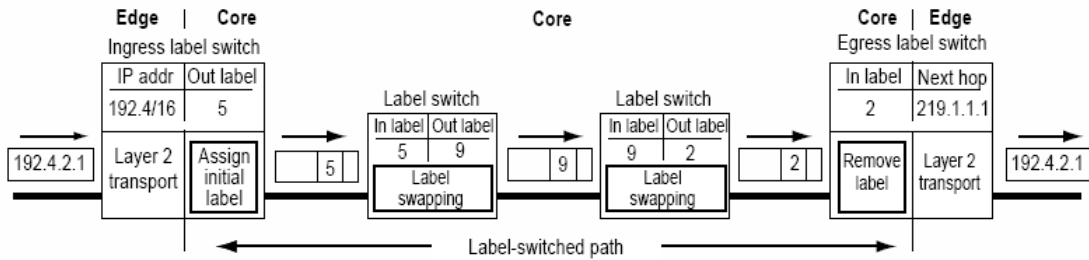
- Os nós de um domínio MPLS designam-se genericamente *Label Switching Routers* (LSR) – é, no entanto, habitual designar os nós de entrada e de saída no domínio por *Label Edge Routers* (LER)
- Todos os nós participam nos protocolos de encaminhamento, de distribuição de etiquetas e outros protocolos de controlo
- Os pacotes de um mesmo FEC são transportados no mesmo LSP
 - » O LSP é definido pela sequência de etiquetas associadas ao FEC em cada nó
- As funções mais complexas associadas ao transporte de pacotes são realizadas pelo LER de entrada, nomeadamente
 - » Interface a redes externas (terminação do protocolo IP)
 - » Classificação dos pacotes IP (associação a um FEC e portanto a um LSP)
 - Esta associação é feita uma única vez, no LER de entrada (nas redes IP convencionais cada nó toma uma decisão independente para associar um pacote a uma rota)
 - » Atribuição da etiqueta (associada ao FEC / LSP)
 - » Adição da etiqueta ao pacote e determinação da porta de saída
- Os nós interiores realizam a função de comutação (processamento e troca de etiquetas), enquanto o LER de saída remove a etiqueta e processa o pacote IP, decidindo sobre o respectivo encaminhamento

Comutação de Etiquetas

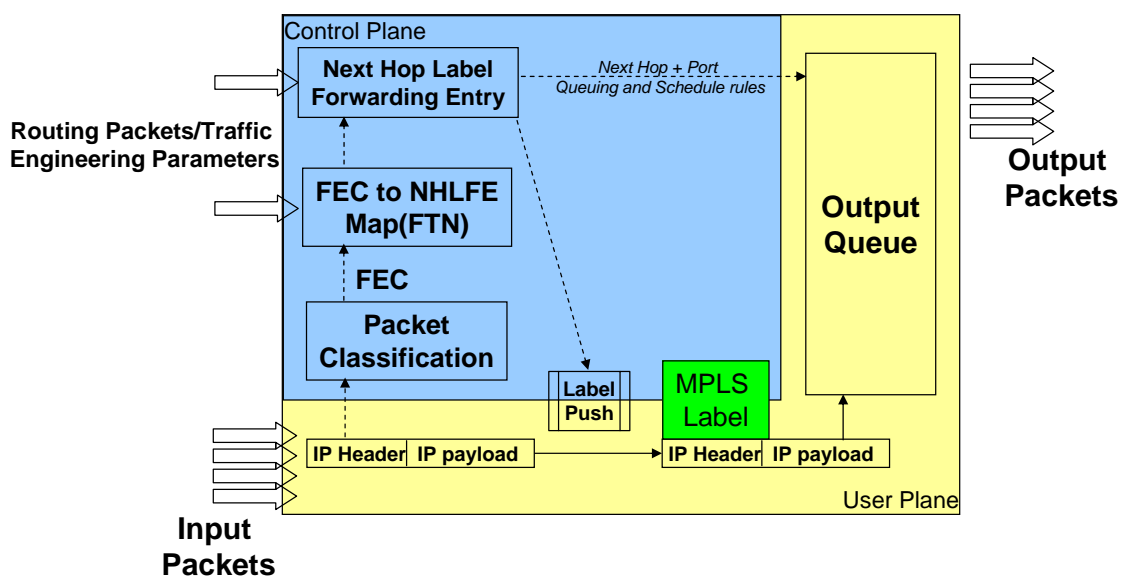
- O LER na entrada do domínio MPLS adiciona uma etiqueta ao pacote IP
- Os LSR ao longo do LSP realizam comutação com base na etiqueta (*label swapping*)
- O LER de saída remove a etiqueta, recuperando o pacote inicial



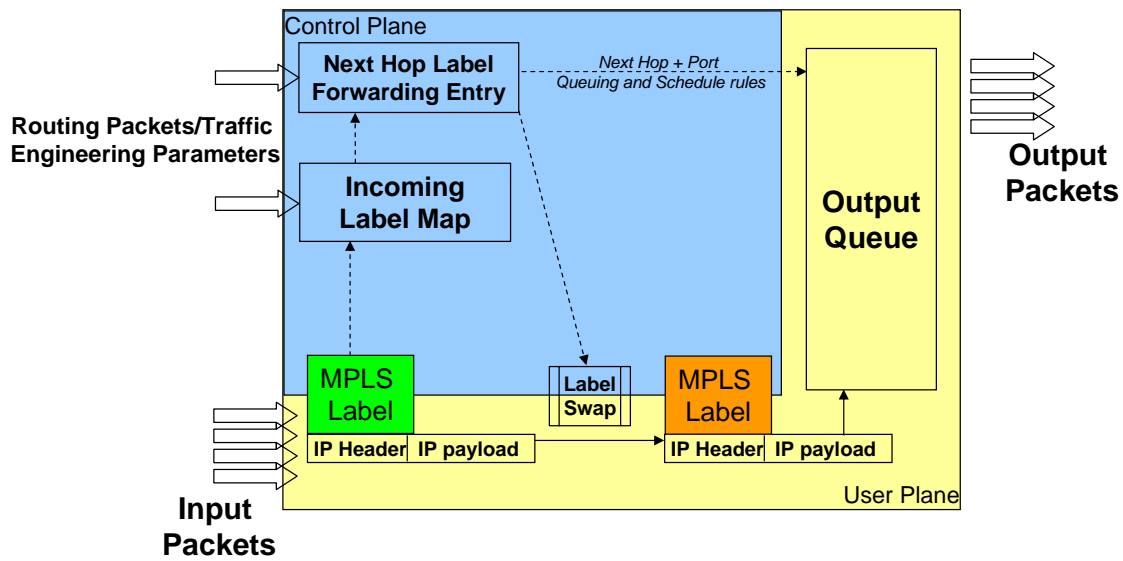
LSP – Label Switched Path



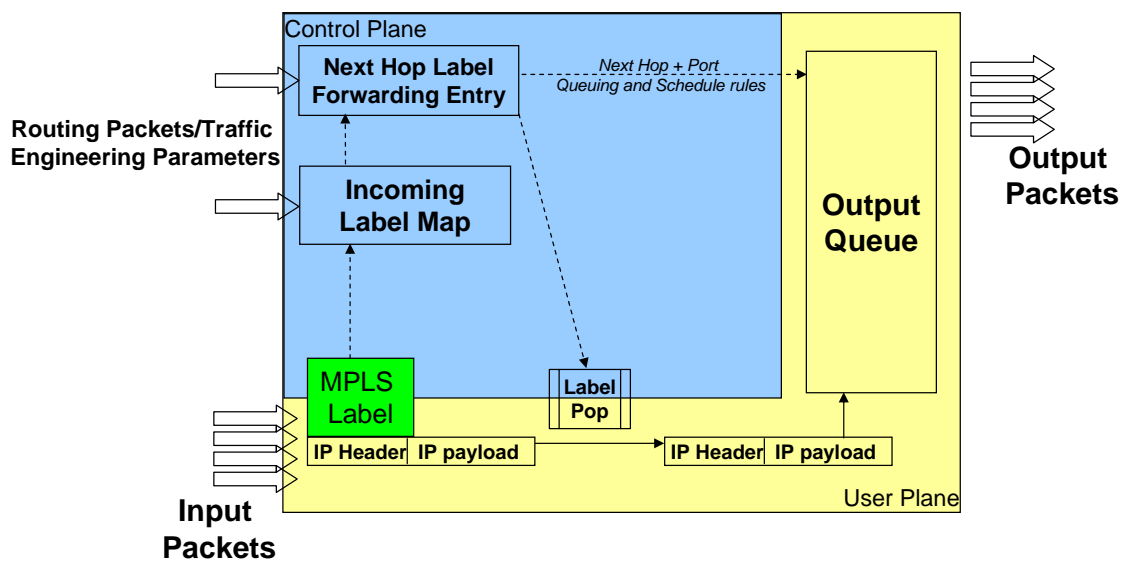
Nó de entrada (ingress node)



Nó intermédio



Nó de saída (egress node)



Estabelecimento de Label Switched Paths

O estabelecimento de *Label Switched Paths*, condição para se poder realizar o transporte de pacotes num domínio MPLS, exige um conjunto de passos que podem ser realizados de acordo com várias estratégias

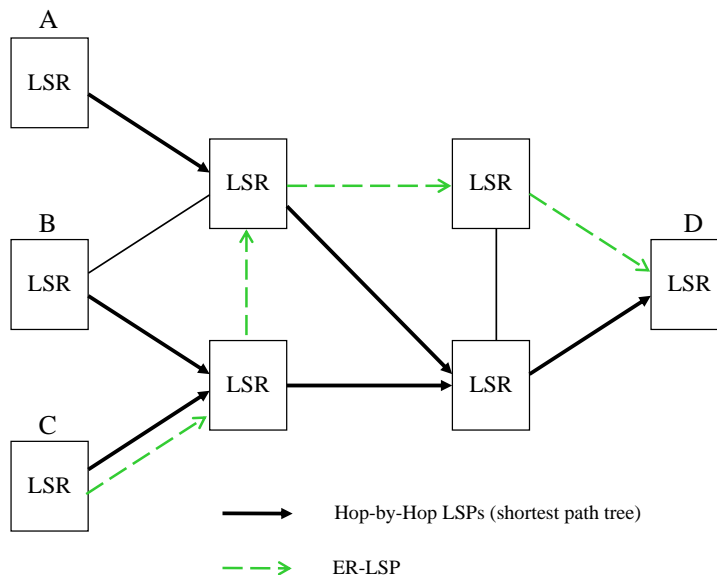
- Descoberta e selecção de rotas
- Criação de etiquetas
- Associação de etiquetas a cada FEC (*label binding*) ao longo da rota (LSP)
- Distribuição da informação sobre a associação etiqueta / FEC pelos LSR que fazem parte do LSP a estabelecer (*Label Distribution Protocol*)
 - » Nós adjacentes passam a partilhar essa informação
 - » Um nó completa a informação presente na sua tabela de comutação relativa a um LSP, com base na informação partilhada com os nós adjacentes nesse LSP

Seleccção de rotas

- Em MPLS são adoptados dois métodos para descoberta e selecção de rotas
- O método básico é baseado em algoritmos de encaminhamento passo a passo (*hop-by-hop routing*), tal como os usados actualmente em redes IP
 - » Cada LSR determina de forma independente o nó adjacente (*next hop*) para cada FEC (*hop-by-hop routed LSP*)
 - » Este método é baseado na topologia e permite descoberta de *shortest path routes*
- O segundo método é baseado em encaminhamento explícito e pode reservar recursos ao longo da rota (ER-LSP – *Explicitly Routed Label Switched Path*)
 - » Um LSR (normalmente o LER de entrada ou de saída) especifica alguns ou todos os LSR que fazem parte do LSP (*loosely / strictly routed LSP*, respectivamente)
 - » A rota explícita (ER-LSP) pode ser escolhida por configuração ou dinamicamente determinada, por exemplo, com base em encaminhamento sujeito a restrições
 - » Duas estratégias podem ser seguidas para estabelecimento de um ER-LSP
 - Adoptar um protocolo de reserva de recursos (e.g., RSVP) e estendê-lo de modo a suportar encaminhamento explícito e distribuição de etiquetas (MPLS-RSVP)
 - Adoptar um protocolo usado para distribuição de etiquetas e estendê-lo de modo a suportar encaminhamento explícito e reserva de recursos (CR-LDP – *Constraint-based Routed Label Distribution Protocol*)

Exemplo de Label Switched Paths

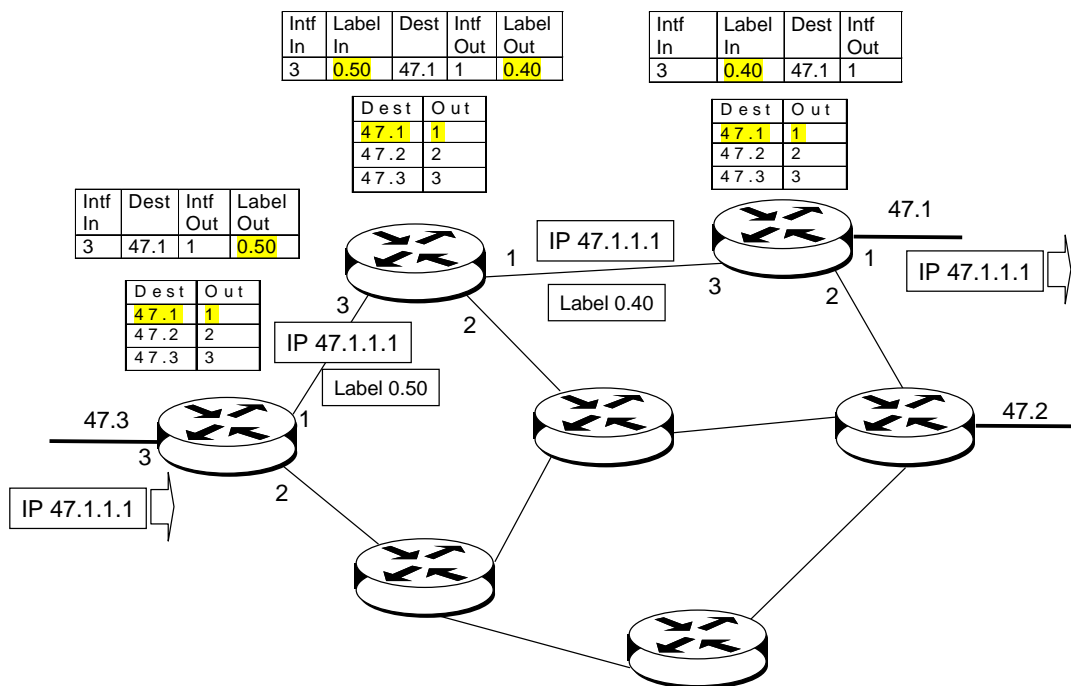
- Árvore de LSPs com origem em A, B e C e destino em D
- ER-LSP entre C e D



Criação de etiquetas

- A necessidade de criação de etiquetas, respectiva associação a um FEC (rota) e estabelecimento de um LSP pode ser determinada de acordo com dois tipos de critérios
 - » O processo pode ser despoletado pela chegada de pacotes associados a um FEC para o qual não existe ainda um LSP estabelecido; este método baseado em tráfego é do tipo *data driven (flow based)*
 - » O processo pode ser desencadeado em resposta a tráfego de controlo (*control driven*), com duas variantes relacionadas com o método de selecção de rotas
 - A descoberta de rotas em resultado da execução de protocolos de encaminhamento permite iniciar o processo de atribuição de etiquetas, associação de etiquetas a FECs (rotas) e distribuição desta informação (protocolo de distribuição de etiquetas); este método baseia-se na partilha de informação topológica pelos nós da rede (*topology based*)
 - O processo pode iniciar-se com pedidos para o estabelecimento de rotas explícitas (ER-LSP), o que permite criar rotas alternativas às descobertas pelos algoritmos de encaminhamento, de acordo com políticas de carácter administrativo (*policy routing*) ou objectivos de Engenharia de Tráfego
- Critérios do tipo control driven são mais facilmente escaláveis, uma vez que se baseiam em rotas estabelecidas antes (e independentemente) de qualquer fluxo de tráfego ser submetido à rede

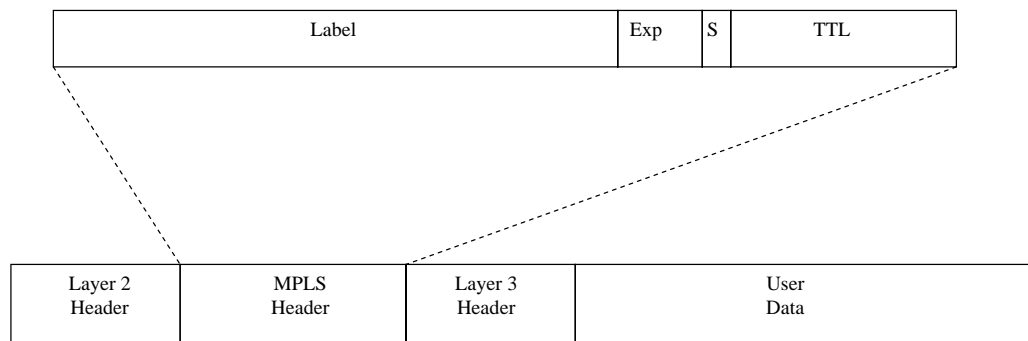
Tabelas de encaminhamento e de comutação



Codificação de etiquetas

- Em MPLS é possível usar diferentes tecnologias de nível 2 – ATM, *Frame Relay*, PPP, *LAN Switching*
 - » As tramas *Frame Relay* e as células ATM transportam identificadores que podem desempenhar a função de etiquetas MPLS – embora permitam apenas um ou quando muito dois níveis de etiquetagem
 - » Por outro lado tramas PPP e MAC não transportam quaisquer identificadores
- O IETF definiu uma forma genérica de codificação de etiquetas, que se torna particularmente útil em PPP e LANs
 - » Uma pilha de etiquetas (*label stack*) é representada como uma sequência de entradas, cada uma constituída por quatro octetos e quatro campos
 - Label Valor da etiqueta (20 bits)
 - EXP Para uso experimental (3 bits)
 - S Indicação de *bottom of stack* (1 bit)
 - TTL *Time To Live* (8 bits)
 - » Quando necessário um cabeçalho MPLS (*shim header*) é inserido entre os cabeçalhos de nível 2 e nível 3
 - » Em ATM a etiqueta no topo da pilha pode ser transportada no VCI e/ou VPI, mas etiquetas adicionais são transportadas num *shim header* na trama AAL5

Cabeçalho MPLS



Túneis MPLS

- Conceptualmente um LSP é um túnel MPLS, isto é, um pacote IP é encapsulado com uma etiqueta (transportada num cabeçalho de nível 2 ou num *shim header* entre cabeçalhos de nível 2 e 3) e viaja inalterado no interior do domínio MPLS
- Em MPLS é possível usar vários níveis de encapsulamento, recorrendo à pilha de etiquetas (*label stack*), ou seja, é possível criar túneis MPLS (LSPs) dentro de um LSP de nível superior, o que permite fazer encaminhamento hierárquico
- Genericamente um *router* R_u pode forçar um pacote a ser enviado para um *router* R_d , mesmo que R_u e R_d não sejam adjacentes no percurso passo-a-passo que seria normalmente seguido pelo pacote e que R_d não seja o destino final; isto pode ser feito encapsulando o pacote dentro de outro pacote endereçado a R_d – cria-se assim um túnel para transporte do pacote original (*tunneled packet*)

Engenharia de Tráfego com MPLS

- A possibilidade de estabelecer ER-LSP permite suportar técnicas de Engenharia de Tráfego e diferenciação de serviços num domínio MPLS
 - » O tráfego pode ser enviado por percursos diferentes dos baseados em métricas do tipo *shortest path*, com base em políticas administrativas, de Qualidade de Serviço (QoS) ou requisitos de Engenharia de Tráfego
- Algoritmos de encaminhamento sujeito a restrições (*constraint-based routing*) têm em conta características das ligações físicas, como largura de banda, atraso, número de saltos (*hops*) e parâmetros de QoS
 - » O cálculo de rotas sujeitas a restrições requer a extensão dos actuais protocolos de encaminhamento, o que deve ser considerado no âmbito mais geral do suporte de Engenharia de Tráfego na Internet
- O encaminhamento explícito é um caso particular de encaminhamento sujeito a restrições, em que a restrição é o percurso explícito
 - » MPLS permite estabelecer e enviar tráfego em ER-LSP, mas não providencia os meios para descobrir percursos sujeitos a restrições
 - » Uma vez que MPLS permite Engenharia de Tráfego e encaminhamento explícito, suscita o interesse em técnicas de encaminhamento que seleccionem rotas baseadas em requisitos de QoS (*QoS routing*) ou políticas administrativas (*policy routing*)

MP λ S – Multiprotocol Lambda Switching

- O paradigma MPLS pode ser estendido ao domínio óptico, com a adopção das técnicas do plano de controlo MPLS para controlar um *cross-connect* óptico (OXC – *Optical Cross-Connect*) e usando comprimentos de onda em lugar de etiquetas numéricas (conceptualmente equivalentes), o que é traduzido na designação *Multiprotocol Lambda Switching*
 - » Funções do plano de controlo incluem descoberta de recursos, controlo distribuído de encaminhamento e gestão de conexões
 - » Estas funções devem permitir estabelecimento expedito de canais ópticos, suporte de funções de Engenharia de Tráfego e esquemas de protecção e restauro
- Um canal óptico (*lightpath*) pode consistir num único comprimento de onda ao longo do percurso na rede óptica de transporte (propriedade de continuidade de comprimento de onda) ou numa concatenação de comprimentos de onda (o que determina a necessidade de conversão)
- A mesma arquitectura de controlo pode ser usada para controlar LSRs e OXCs, que deste ponto de vista apresentam algumas propriedades comuns
 - » Em ambos os casos existe separação do plano de controlo do plano de dados
 - » As relações estabelecidas num LSR/OXC e mantidas numa tabela entre pares <porta, etiqueta / canal óptico> na entrada e na saída não são alteradas pelo *payload* dos pacotes (plano de dados)

Conclusões

- O MPLS oferece múltiplas vantagens sobre soluções anteriores, permitindo suportar infra-estruturas de redes com capacidade de integração e diferenciação de serviços
 - » Algumas das vantagens não são exclusivas do MPLS e, por outro lado, MPLS tem de ser combinado com outros mecanismos (e.g., QoS) para se tirar total partido dos seus atributos
- Um argumento forte a favor do MPLS é a combinação das seguintes propriedades
 - » Escalabilidade do ponto de vista do número de nós e fluxos de tráfego
 - » Flexibilidade, uma vez que não restringe a tecnologia de comutação
 - » Simplicidade e rapidez da comutação de etiquetas, o que garante elevado desempenho
 - » Capacidade de suportar percursos definidos por critérios de Engenharia de Tráfego e diferenciação de serviços, essencial para a provisão de QoS
- As propriedades de escalabilidade e Engenharia de Tráfego são importantes para a implantação do MPLS, pois permitem criar uma infra-estrutura de elevado desempenho (elevada capacidade e pequena latência), possibilitando economias de escala, oferta de novos serviços (e.g., VPNs, segregação de tráfego por CoS) e restauro rápido de rotas; por outro lado, os operadores podem tirar partido dos investimentos feitos na tecnologia ATM, usando LSRs baseados em ATM ou MPLS sobre ATM num modelo *overlay* (neste caso, LSRs MPLS comunicam através de uma rede ATM)
- A generalização do MPLS (e em particular a sua extensão ao domínio óptico) permite aos operadores investir de forma segura numa solução de longo prazo

Anexo

CLIP

Classical ATM and ARP over ATM

CLIP – Classical IP and ARP over ATM

- CLIP está definido no RFC 2225 (Abril de 1998), que tornou obsoleto o RFC 1577 (Janeiro de 1994)
- CLIP adota o modelo *overlay* (redes sobrepostas) – o ATM é considerado como qualquer outra tecnologia de subrede (LANs ou MANs IEEE 802.x, *Frame Relay*, etc.)
- CLIP baseia-se no RFC 1483 (*Multiprotocol Interconnect over ATM*)
 - » Usa AAL5
 - » Especifica como identificar / multiplexar diferentes protocolos em ATM, existindo duas alternativas
 - Multiplexagem por Circuito Virtual – cada Circuito Virtual transporta um único protocolo, identificado durante o estabelecimento da ligação; esta solução é pouco escalável
 - Multiplexagem por Pacote – as unidades de dados são encapsuladas em LLC ou LLC/SNAP
 - » Em CLIP é adoptado, por omissão, encapsulamento LLC/SNAP, para pacotes IP e ARP (transportados em tramas AAL5)
 - OUI = 0x 00 00 00; PID = 0x 08 00 (IP); PID = 0x 08 06 (ARP)

LIS – Logical IP Subnetwork

- Um conceito básico em CLIP é o de *Logical IP Subnetwork* (LIS), isto é, um grupo de estações (*hosts* ou *routers*) que constituem uma subrede IP, independentemente da respectiva localização física na rede ATM
 - » Estações na mesma LIS partilham o mesmo *netid/subnetid*
 - » Estações na mesma LIS comunicam directamente através de um VCC ATM (permanente ou comutado), isto é, sem qualquer *router* no percurso dos dados
 - » Numa rede ATM podem ser configuradas múltiplas LIS que operam independentemente
 - » A comunicação entre estações em LIS diferentes, ainda que ligadas à mesma rede ATM, realiza-se através de um ou mais *routers* (excepto no caso de se usar NHRP – *Next Hop Resolution Protocol*)
 - Para interligar duas ou mais LIS, um *router* tem de ser membro dessas LIS
 - Qualquer estação pode igualmente pertencer a mais do que uma LIS, o que nalguns casos pode evitar o recurso a *routers* para comunicação com essa estação

Aspectos de configuração

- Requisitos de máquinas IP (*hosts*, *routers*) que operam numa LIS ATM
 - » Todos os membros da LIS têm o mesmo *netid/subnetid* e a mesma máscara de endereço
 - » Todos os membros da LIS estão directamente ligados à rede ATM
 - » Todos os membros da LIS têm um mecanismo para resolver endereços IP em endereços ATM via *ATMARP* e inversamente via *InATMARP*, quando usam SVCs
 - » Todos os membros da LIS têm um mecanismo para resolver VCs em endereços IP, via *InATMARP*, quando usam PVCs
 - » Todos os membros de uma LIS devem poder comunicar entre si via ATM
- Parâmetros ATM que devem ser configurados em cada estação IP
 - » Endereço ATM (endereço de *hardware*) da estação IP
 - » Lista de endereços ATM de servidores *ATMARP* localizados na LIS

Serviço ATMARP

- Uma LIS oferece um único serviço ATMARP (centralizado ou distribuído) a todos os seus membros que usam SVCs

- Um servidor ATMARP realiza as seguintes funções
 - » Regista pares de endereços (IP, ATM) dos membros da LIS
 - O processo de registo é diferente em RFC 1577 e em RFC 2225
 - » Mantém actualizada a tabela que relaciona pares de endereços (IP, ATM) dos membros de uma LIS
 - » Resolve endereços a pedido

Pacotes ATMARP e InATMARP

- Os pacotes *ATMARP* e *InATMARP* contêm quatro endereços
 - » *Source hardware address* – endereço ATM da estação de origem
 - » *Source protocol address* – endereço IP da estação de origem
 - » *Target hardware address* – endereço ATM da estação alvo
 - » *Target protocol address* – endereço IP da estação alvo

- Estão definidos cinco tipos de operação
 - » *ATMARP_Request*
 - » *ATMARP_Reply*
 - » *ATMARP_NAK*
 - » *InATMARP_Request*
 - » *InATMARP_Reply*

Resolução de endereços

- Circuitos Virtuais Permanentes (PVCs) e encapsulamento LLC/SNAP
 - » Uma estação envia *InATMARP_Request* no PVC para obter o endereço IP da estação remota (no outro extremo do PVC)
- Circuitos Virtuais Comutados (SVCs)
 - » Uma estação envia *ATMARP_Request* ao servidor, para obter o endereço ATM da estação alvo, indicando o respectivo endereço IP
 - » Uma vez resolvido o endereço ATM, a estação pode estabelecer o Circuito Virtual com a estação alvo
 - » Deste processo resulta uma dupla associação
 - Associação dos endereços IP e ATM da estação alvo, necessária ao processo de estabelecimento do Circuito Virtual – essa informação é guardada na tabela ATMARP da estação
 - Associação do endereço IP da estação alvo com um Circuito Virtual (VPI/VCI), necessária para transferência de pacotes IP (encapsulados em tramas AAL5, antes da fragmentação em células ATM) no Circuito Virtual

Registo de uma estação

- O registo de uma estação (cliente) inicia-se após o estabelecimento de um Circuito Virtual para troca de mensagens com o servidor
- Em RFC 2225, cada estação deve registar inicialmente cada par de endereços (IP, ATM) que lhe foi atribuído (pelo menos um par)
 - » O registo inicia-se com o envio de *ATMARP_Request* em que o *Target IP Address* é o seu endereço IP (igual ao *Source IP Address*), o *Source ATM Address* é o seu endereço ATM e é solicitado o *Target ATM Address*
 - » Em caso de sucesso, o servidor envia uma resposta em que o *Target ATM Address* e o *Source ATM Address* são iguais ao endereço ATM da estação
 - » Cada estação deve refrescar a sua informação ATMARP no servidor pelo menos uma vez em cada 15 minutos (repetindo o processo de registo)
- Em RFC 1577 é o servidor que toma a iniciativa do registo
 - » O registo inicia-se com o envio de *InATMARP_Request*, solicitando os endereços da estação – *Target IP Address* e *Target ATM Address*
 - » Em caso de sucesso, a estação envia uma resposta que inclui os seus endereços IP e ATM
 - » O processo de refrescamento é também, neste caso, idêntico ao de registo

Tabela ATMARP do servidor

- Um servidor ATMARP mantém uma tabela que relaciona pares de endereços (IP, ATM) dos membros de uma LIS
- Uma entrada que associa um par de endereços (IP, ATM) de um cliente é criada pelo processo de registo inicial do cliente
- Cada entrada da tabela é válida por um período de 20 minutos, sendo eliminada se entretanto não for refrescada (actualizada) pelo cliente, independentemente do estado de qualquer VC associado a essa entrada
- O refrescamento (actualização) de uma entrada ocorre em dois casos
 - » Quando o cliente refresca a sua informação ATMARP, por repetição do processo de registo
 - » Quando o cliente envia um pedido normal de resolução de endereços
 - Se existir uma entrada correspondente ao par de endereços (IP, ATM) da estação (presentes no *ATMARP_Request*), o temporizador é reiniciado, desde que tenham passado mais de 10 minutos desde a última actualização
 - A entrada é criada, se não existir, permitindo que antigos clientes RFC 1577 se registem com o simples envio de pedidos de resolução de endereços

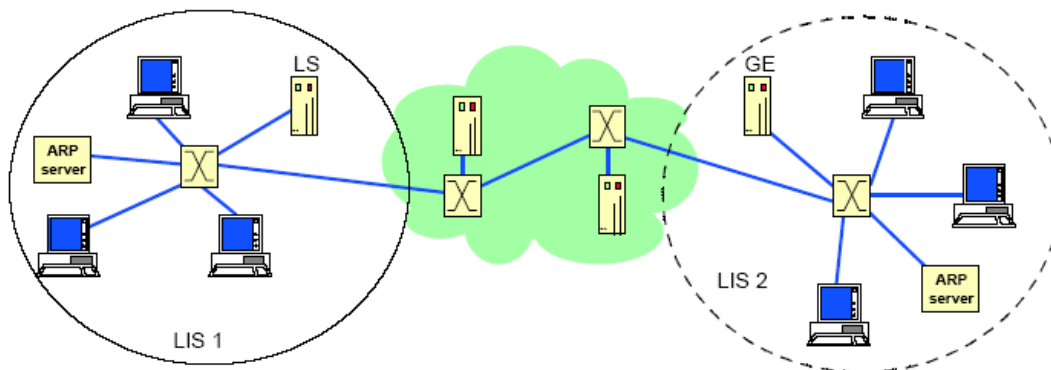
Cientes ATMARP

- Um cliente contacta o serviço ATMARP para
 - » Se registar e posteriormente refrescar essa informação
 - » Resolver endereços de outras estações e revalidar essa informação
- Os clientes devem
 - » Responder a pacotes *ATMARP_Request* (por razões de robustez) e *InATMARP_Request* recebidos em qualquer VC
 - » Enviar pacotes *ATMARP_Request* ao servidor e usar as respostas positivas (*ATMARP_Reply*) para criar / actualizar as entradas da sua tabela ATMARP
 - » Enviar pacotes *InATMARP_Request*, conforme necessário, e usar as respostas (*InATMARP_Reply*) para criar / actualizar as entradas da sua tabela ATMARP
 - » Refrescar no servidor a sua informação ATMARP, pelo menos uma vez em cada 15 minutos (repetindo o processo usado para registo)
- As entradas da tabela ATMARP de um cliente são válidas por um período máximo de 15 minutos (*ageing*), após o que devem ser removidas ou revalidadas

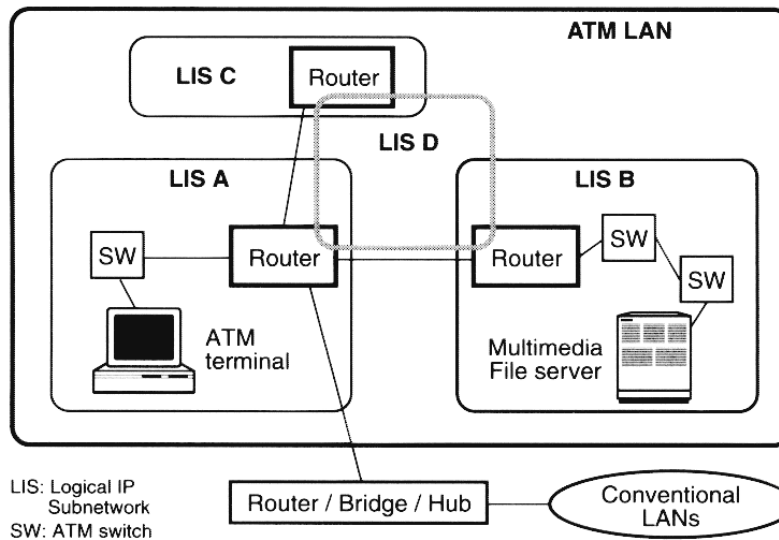
Tabela ATMARP dos clientes

- Uma entrada da tabela ATMARP de um cliente é (temporariamente) invalidada quando expirar o seu período de validade
- Se já não existir qualquer VC associado à entrada, esta é eliminada
- Se existir um VC aberto (PVC ou SVC), a entrada deve ser revalidada antes de se poder continuar a enviar tráfego de dados (no caso de a revalidação se realizar antes de expirar o período de validade, o temporizador é reiniciado e pode ser enviado tráfego de dados durante o processo de revalidação)
 - » No caso de um PVC, a entrada é revalidada (actualizada) com o envio de *InATMARP_Request* e a recepção de *InATMARP_Reply*
 - » No caso de um SVC, a entrada é revalidada inquirindo o serviço ATMARP
 - Se for recebida uma resposta válida (*ATMARP_Reply*) a entrada é actualizada
 - Se o serviço não conseguir resolver o endereço (resposta *ATMARP_NAK*), o SVC deve ser fechado e a entrada removida
 - Se o serviço não estiver disponível, o cliente tenta revalidar a entrada enviando *InATMARP_Request* no SVC, sendo a entrada actualizada se for recebido *InATMARP_Reply*; caso contrário, o SVC deve ser fechado e a entrada removida
- Uma entrada invalidada é removida se entretanto o VC associado for fechado

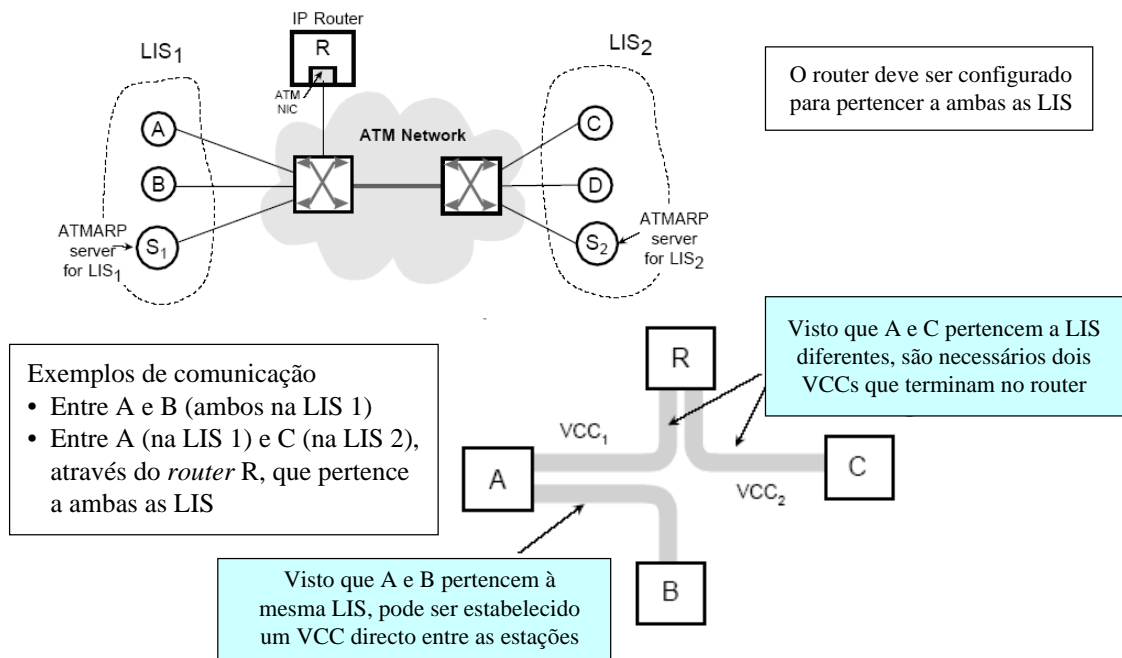
Múltiplas LIS



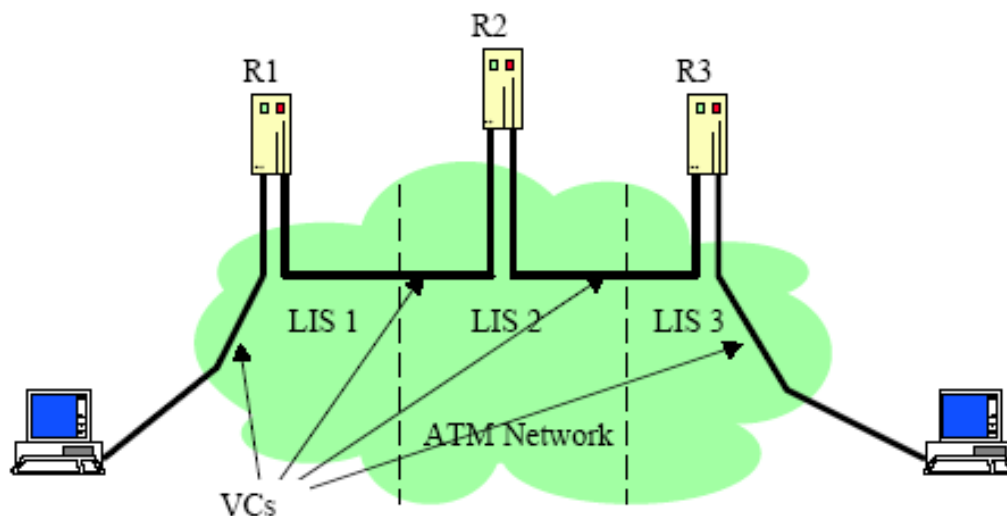
Exemplo – LAN ATM com múltiplas LIS



Comunicação na mesma LIS e entre LIS



Múltiplas LIS – percurso dos dados



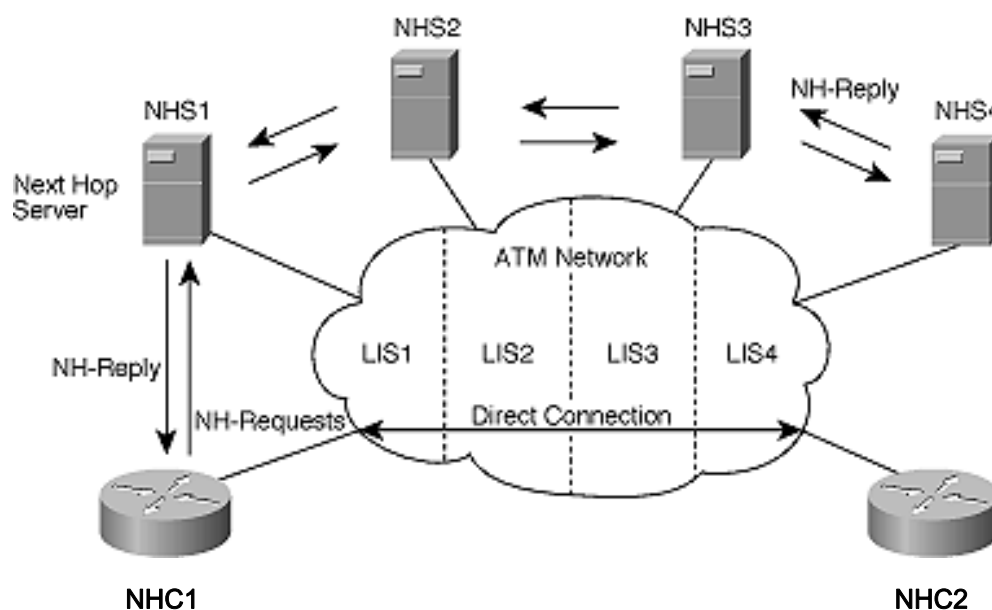
Generalização – uso de NHRP em redes NBMA

- Em CLIP, o serviço ATMARP é definido num ambiente ATM configurado como uma LIS, isto é, ATM é um substituto de segmentos LAN que ligam *hosts* e *routers* IP que operam segundo o paradigma clássico de uma LAN
- Em redes NBMA (*Non Broadcast Multiple Access*), de que o ATM é um exemplo, a organização em várias LIS é útil, pois limita o número de clientes por servidor ARP e permite separar domínios por razões administrativas, de segurança, etc., mas a comunicação entre estações em LIS diferentes tem de ser realizada através de *routers*, mesmo sendo possível estabelecer uma ligação directa (por exemplo, um circuito virtual ATM) entre as estações
- A generalização do suporte de IP (e outros protocolos de *internetworking*) sobre redes NBMA com o objectivo de permitir comunicação directa entre sistemas em LIS diferentes, baseia-se num protocolo de resolução de endereços – NHRP (*Next Hop Resolution Protocol*) – concebido para operar em redes NBMA, e descrito no RFC 2332

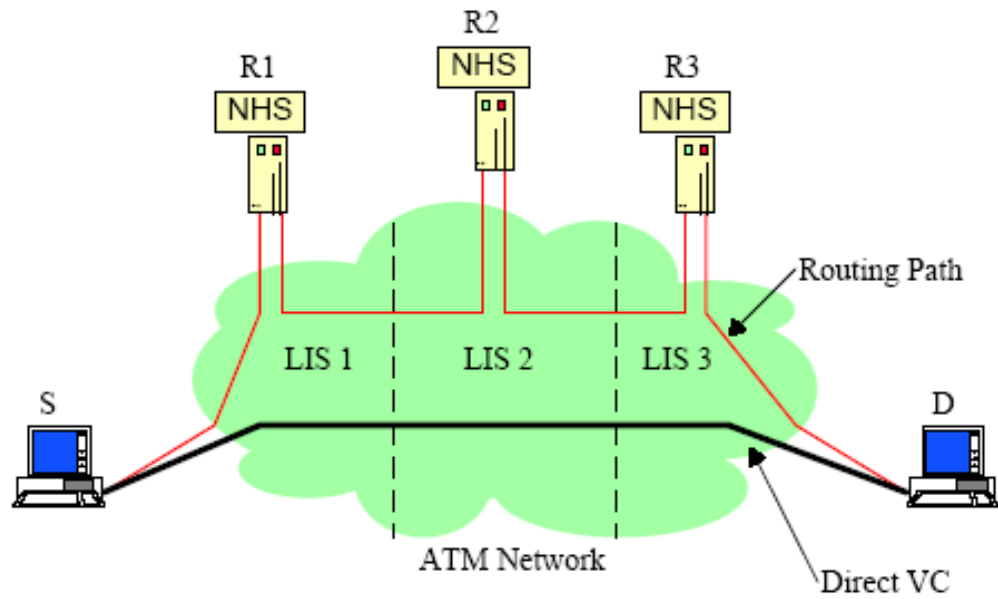
NHRP – Next Hop Resolution Protocol

- NHRP permite a uma estação que queira comunicar através de uma rede NBMA determinar endereços NBMA e endereços de *internetworking* (por exemplo, IP) de “*NBMA next hops*” no percurso até uma estação de destino
 - » Se a estação de destino estiver ligada à rede NBMA e se, do ponto de vista administrativo, for permitida comunicação directa, então o “*NBMA next hop*” é a própria estação de destino; de contrário, o “*NBMA next hop*” é o *router* de saída da rede NBMA “mais próximo” da estação de destino
- NHRP baseia-se no modelo cliente-servidor – NHC (*Next Hop Client*) e NHS (*Next Hop Server*), residindo este num *router*
 - » Um NHC regista-se inicialmente num *router* / NHS e emite pedidos de resolução de endereços NBMA (*NHRP Resolution Request*) ao seu NHS
 - » Os pedidos são “encaminhados” através de NHSs até ao NHS “mais próximo” do destino; este NHS faz a resolução do endereço, sendo a informação enviada pelo percurso inverso (*NHRP Resolution Reply*)
 - » Completado o processo, é estabelecido um circuito virtual directo, que funciona como um atalho (*short-cut*) entre os NHCs correspondentes
 - » As mensagens do protocolo NHRP e os pacotes de dados iniciais (antes do estabelecimento do *short-cut*) seguem o percurso *default* determinado pelo protocolo de encaminhamento

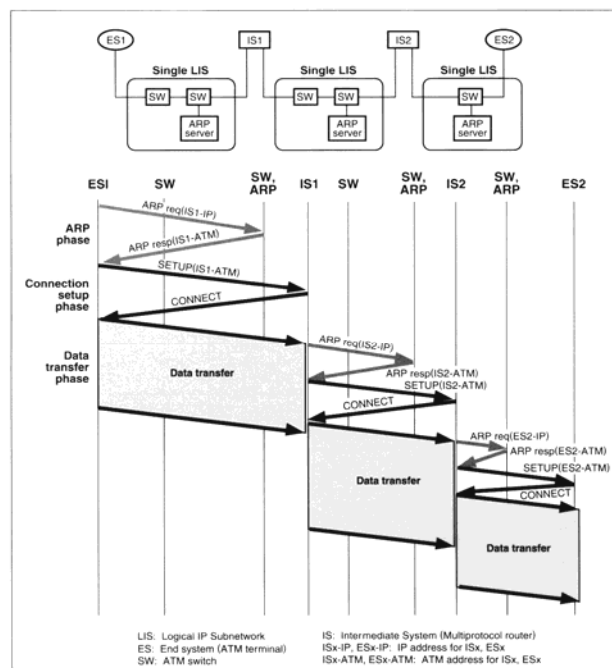
NHRP – mensagens de controlo



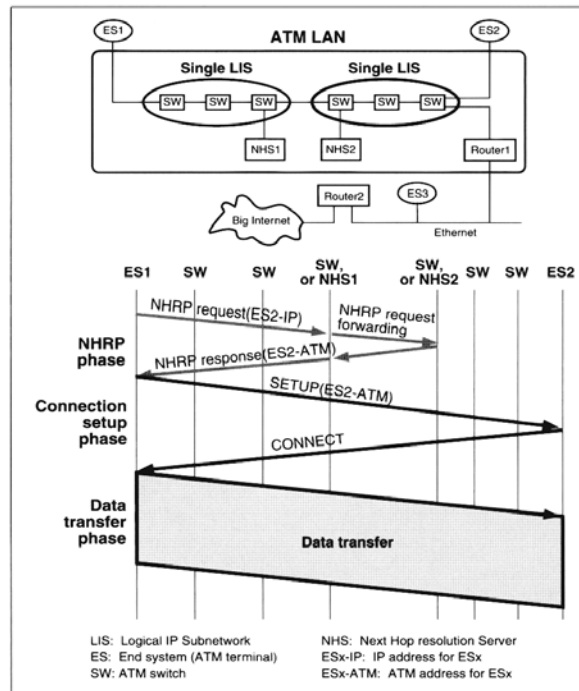
NHRP – percurso dos dados (default e short-cut)



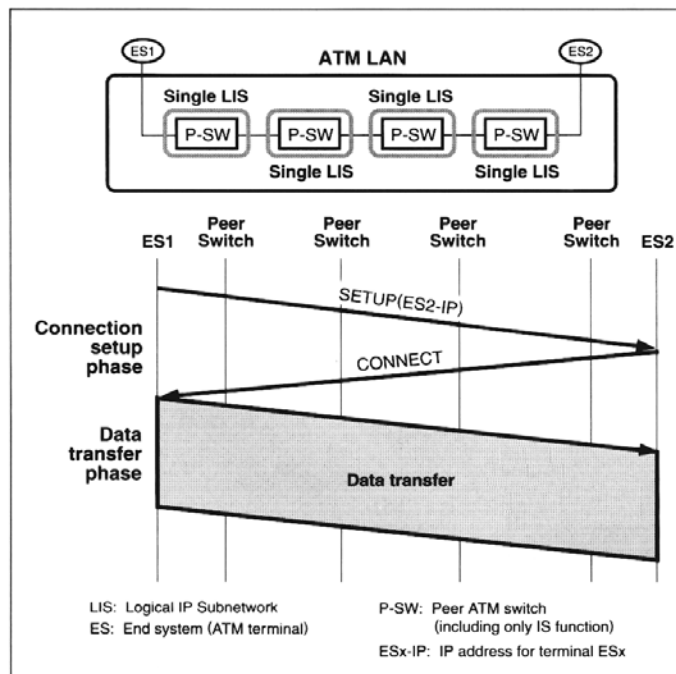
Exemplo – CLIP



Exemplo – CLIP e NHRP



Exemplo – possível evolução (peer model)



CLIP em WANs IP – organização

- Com a implantação da tecnologia ATM, os fornecedores de serviços IP passaram a poder dispor de infra-estruturas de rede (*backbones*) de elevada capacidade e pequena latência, flexível e escalável
 - » A tecnologia ATM permite explorar mecanismos de Engenharia de Tráfego (com recurso a *Virtual Paths*) e suporta QoS diferenciada
- Estas redes organizaram-se numa estrutura em dois níveis – um núcleo (*core*) de alta velocidade interligando *routers* IP localizados na periferia (*edge*), onde se situa a inteligência de processamento
- A adopção do modelo CLIP neste tipo de redes caracteriza-se por dois aspectos que o distinguem da forma como é usado em LANs
 - » A conectividade entre *routers* IP é assegurada por Circuitos Virtuais Permanentes, que definem a topologia lógica da rede IP
 - » Os *routers* são tipicamente organizados numa única LIS, de forma a evitar *routers* intermédios no percurso de dados (malha lógica completa)

CLIP em WANs IP – análise

- A arquitectura CLIP caracteriza-se por dois tipos de equipamentos – *routers* IP e comutadores ATM – e duas topologias segregadas (topologia física ATM na qual é sobreposta a topologia lógica IP)
 - » Os *routers* e os comutadores não partilham informação topológica, o que impede uma optimização conjunta de recursos
 - » Cada rede tem os seus protocolos próprios de encaminhamento e sinalização e o seu espaço de endereçamento, com naturais duplicações e ineficiências
- Ao organizar os *routers* IP numa única LIS cada *router* é adjacente lógico (*logical peer*) de qualquer outro
 - » O número de Circuitos Virtuais cresce com o quadrado do número de *routers*, o que levanta problemas de escalabilidade
 - » O consequente aumento do tamanho das tabelas de encaminhamento torna a sua gestão e actualização mais complexa
 - » O tráfego de controlo cresce com o aumento do número de adjacências entre *routers* e com a replicação com efeitos multiplicativos dos avisos que recebem dos vizinhos
 - » Em caso de reconfigurações podem criar-se situações de instabilidade com tempos elevados de convergência para um estado estável
- A arquitectura MPLS resolve estes problemas do modelo CLIP

LANE

LAN Emulation Over ATM

LAN Emulation

“In order to use the vast base of existing LAN application software, the ATM Forum has defined an ATM service, herein called “LAN Emulation”, that emulates services of existing LANs across an ATM network and can be supported via a software layer in end systems. The LAN Emulation service enables end systems (e.g. workstations, servers, bridges, etc.) to connect to the ATM network while the software applications interact as if they are attached to a traditional LAN. Also, this service supports interconnection of ATM networks with traditional LANs by means of IEEE bridging methods. This allows interoperability between software applications residing on ATM-attached end systems and on traditional LAN end systems.”

LAN Emulation – objectivos

- O ATM Forum definiu um serviço de emulação de LANs em ATM (LANE – *LAN Emulation*) com as mesmas características do serviço MAC das LANs IEEE 802
- O serviço LANE tinha como objectivo permitir reutilizar, de forma transparente, o *software* então disponível em LANs (*legacy LANs*) e executar as aplicações existentes sem necessidade de alterações
 - » LANE permite tornar ATM transparente em subredes de nível 2 (*bridged LANs*)
 - » LANE é independente dos protocolos de nível 3 (transparência)
- O serviço LANE requer uma camada adicional de *software* nos *end systems* (*hosts*) ATM e em *intermediate systems* (*bridges* e *routers*) com interfaces ATM, que permita oferecer o serviço MAC; são também necessárias funções adicionais nos equipamentos de rede (comutadores ATM e/ou servidores dedicados) com o objectivo de emular o comportamento de uma LAN e os respectivos serviços

Características das LANs e ATM – comparação

- Características das LANs (serviço MAC)
 - » Serviço sem conexão (*connectionless*)
 - » *Broadcast / multicast* nativo
 - Automático em meios partilhados (uma única trama difundida no meio)
 - Replicação simples de tramas em comutadores (sem necessidade de estabelecer e gerir conexões multiponto), na mesma VLAN
 - » Endereços MAC não estruturados (*flat*) e independentes da topologia da rede ou da localização física dos sistemas
 - » *Bridging* transparente e configuração de LANs Virtuais (VLANs)
- Características do ATM
 - » Serviço orientado à conexão (*connection oriented*)
 - » *Broadcast / multicast* não nativo
 - Replicação simples de células em comutadores, mas gestão de conexões multiponto complexa (inserção e remoção de *end-points*)
 - » Endereços ATM estruturados e associados a portas de comutadores

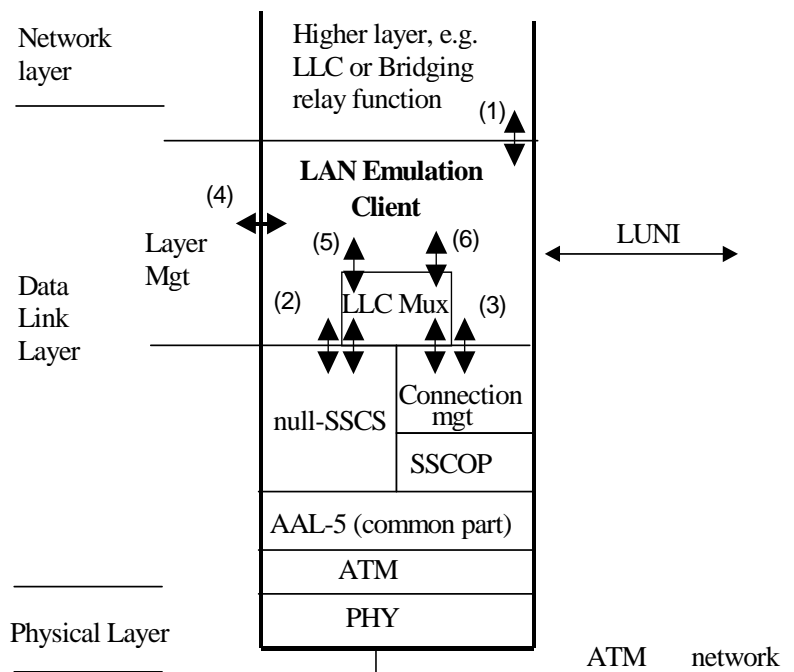
LANs Emuladas e LANs Virtuais

- Uma LAN emulada em ATM (ELAN – *Emulated LAN*) é de facto uma LAN Virtual, oferecendo um mecanismo de *bridging* entre os membros da ELAN (tal como em LANs comutadas)
 - » Numa ELAN é possível executar sem modificações as mesmas aplicações que correm numa LAN
- Uma vez que uma ELAN é uma LAN Virtual é possível constituir múltiplas ELANs numa rede ATM
 - » A comunicação entre ELANs realiza-se através de *routers*
- A necessidade de garantir interfuncionamento de LANs emuladas em ATM com LANs legadas (*legacy LANs*) requer a extensão a ELANs dos mecanismos de *bridging*
 - » Uma ELAN pode coincidir com uma VLAN ou fazer parte duma VLAN que inclui também estações ligadas a uma LAN convencional
 - » As aplicações podem ser executadas indistintamente em *hosts* ligados através de ATM ou duma LAN convencional

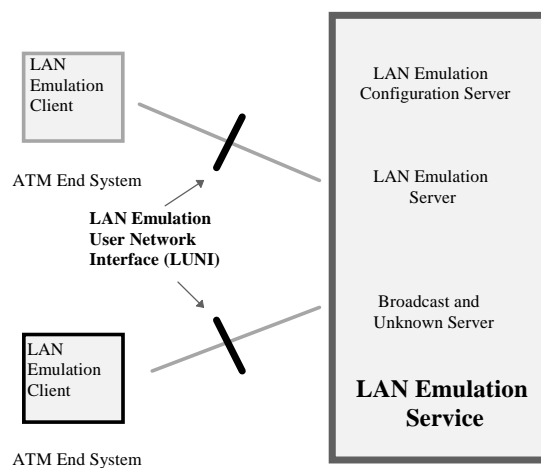
LANE – funções, modelo e componentes

- O ATM Forum definiu a emulação de LANs IEEE 802.3 e 802.5
- A emulação do serviço MAC requer as seguintes funções
 - » Resolução de endereços
 - Determinação do endereço ATM de uma estação, conhecido o seu endereço MAC (LE_ARP)
 - » Difusão de tramas
 - Necessário para tramas com endereço de destino *broadcast* ou *multicast* ou com endereço *unicast* mas localização desconhecida (endereço MAC ainda não resolvido num endereço ATM)
 - » Encapsulamento
 - Tramas MAC são encapsuladas com um cabeçalho adicional (que na versão 2.0 inclui LLC/SNAP) e transportadas em AAL5
- O serviço LANE é baseado no modelo cliente-servidor
 - » Os clientes residem em *hosts*, *bridges* / comutadores ou *routers* e incluem um componente designado LEC (*LAN Emulation Client*)
 - » O serviço (*LAN Emulation Service*) é realizado por três componentes: LECS (*LAN Emulation Configuration Server*), LES (*LAN Emulation Server*) e BUS (*Broadcast and Unknown Server*)

Arquitetura – interfaces e camadas protocolares



Interface de acesso ao serviço



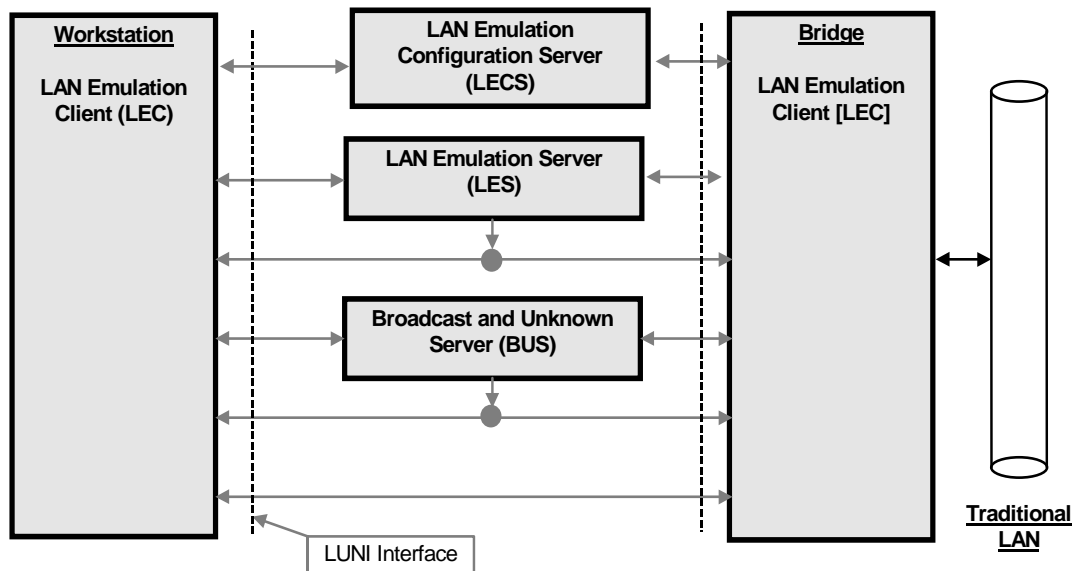
LEC – LAN Emulation Client

- Um LEC (*LAN Emulation Client*) está associado a uma única ELAN
- Uma máquina pode pertencer a mais do que uma ELAN, mas necessita de ter um LEC por cada ELAN a que se associe
- Um LEC suporta as seguintes funções
 - » Oferece interface LAN *standard* às camadas protocolares superiores
 - » Efectua o seu registo no LES da respectiva ELAN
 - Um LEC pode registar todos os endereços MAC pelos quais é responsável ou então associar-se como *proxy* (esta é a situação típica de uma *bridge* ou de um comutador entre uma LAN e a rede ATM)
 - Um LEC pode registar endereços *multicast* para recepção selectiva de tramas (serviço *multicast* selectivo)
 - » Realiza pedidos de resolução de endereços, dirigidos ao LES
 - » Encapsula as tramas MAC e estabelece os circuitos virtuais ATM para comunicação com outros LECs ou com o BUS

LAN Emulation Service – componentes

- LECS – *LAN Emulation Configuration Server*
 - » Existe um único por domínio, para configuração de ELANs
 - » Associa cada LEC à respectiva ELAN, fornecendo-lhe o endereço do LES correspondente
- LES – *LAN Emulation Server*
 - » Existe um único por ELAN
 - » Aceita registos dos LECs pelos quais é responsável
 - » Mantém associações entre endereços MAC e ATM
 - » Responde a pedidos de resolução de endereços ou transfere esses pedidos (em particular se não for capaz de resolver os endereços)
- BUS – *Broadcast and Unknown Server*
 - » Existe um único por ELAN
 - » Emula o mecanismo de difusão de tramas típico das LANs
 - » A difusão pode ser realizada para todos os endereços ATM conhecidos (registados) na ELAN ou pode ser selectiva

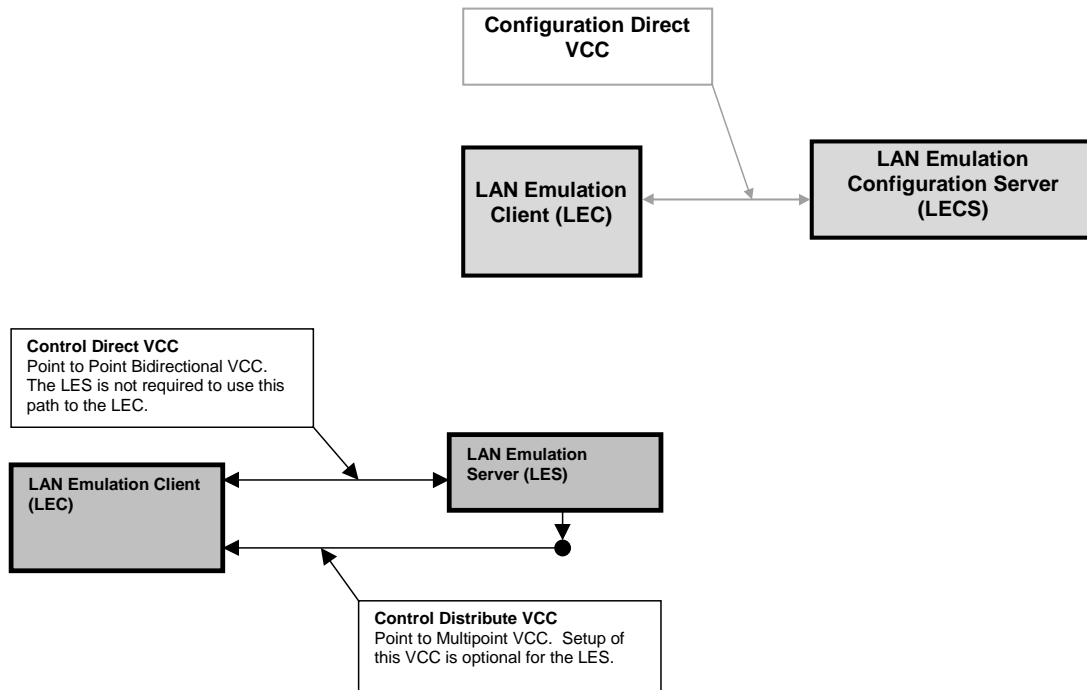
Conexões de dados e de controlo



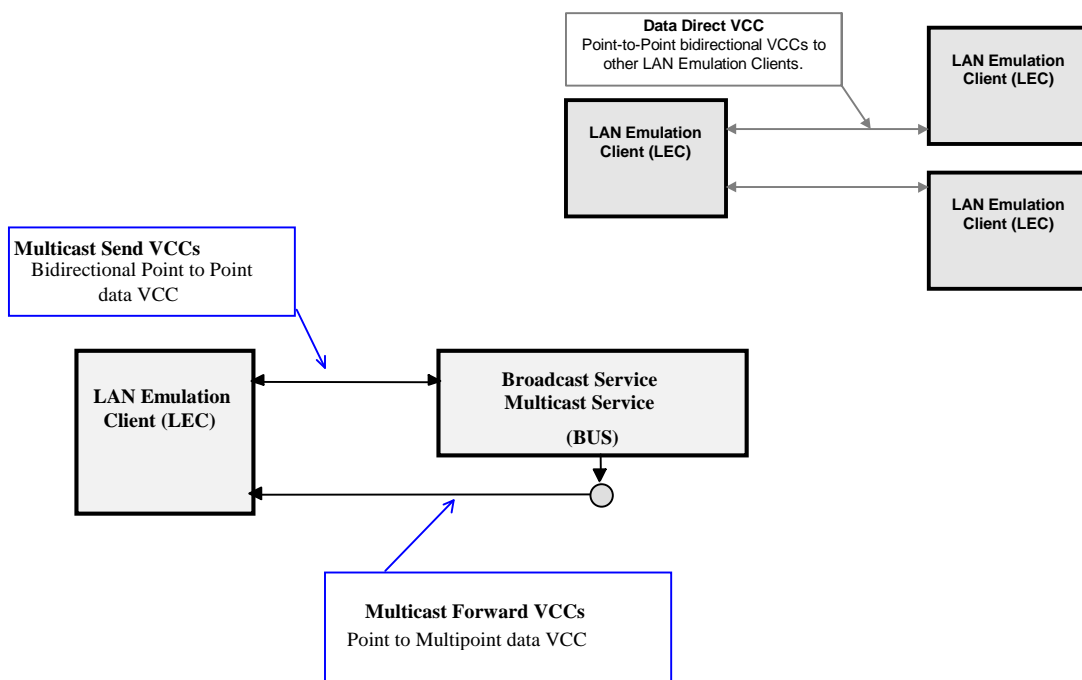
Conexões de dados e de controlo

- O serviço LANE requer o estabelecimento de conexões ATM (VCC) para suportar fluxos de controlo e de dados
- A transferência de dados entre LECs é suportada em VCCs ponto a ponto bidireccionais designados *Data Direct VCC*
- São necessários vários tipos de VCCs para comunicação do LEC com os servidores (dados ou controlo)
 - » Com o LECS
 - *Configuration Direct VCC* – ponto a ponto, bidireccional
 - » Com o LES
 - *Control Direct VCC* – ponto a ponto, bidireccional
 - *Control Distribute VCC* – ponto a multiponto, unidireccional (opcional)
 - » Com o BUS
 - *Multicast Send VCC* – ponto a ponto, bidireccional
 - *Default Multicast Send VCC* – associado ao endereço *broadcast*
 - *Selective Multicast Send VCC* (opcionais)
 - *Multicast Forward VCC* – ponto a multiponto (pelo menos um)

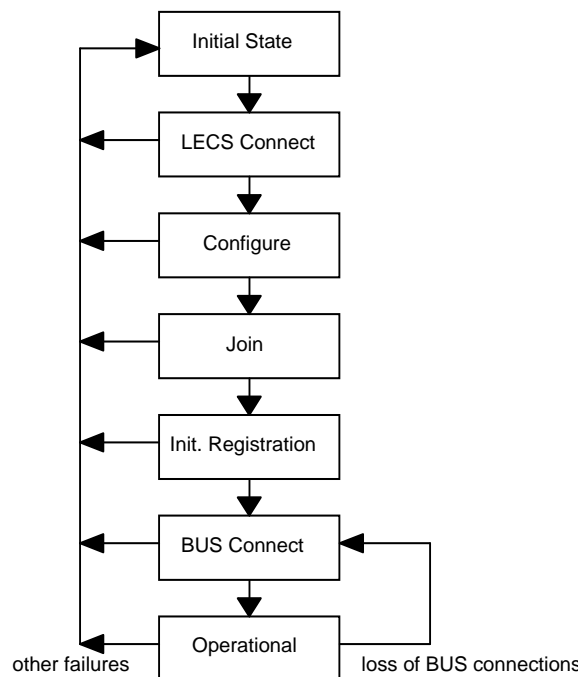
Conexões de controlo



Conexões de dados



Inicialização do serviço LANE



Inicialização do serviço LANE

A Inicialização do serviço envolve várias fases e funções

- *Initial State* – neste estado o LES e os LECs devem conhecer alguns parâmetros (endereços, nome da ELAN, etc.)
- *LECS Connect* – LEC estabelece *Configuration Direct VCC* com LECS
- *Configuration* – LEC descobre o endereço do LES na respectiva ELAN (por pré-configuração ou obtido do LECS)
- *Join* – LEC estabelece *Control Direct VCC* com LES; no final tem atribuído um identificador único (LECID), conhece o identificador da ELAN respectiva (ELAN_ID), o tipo da LAN e o tamanho máximo das tramas
- *Initial Registration* – LEC pode registar qualquer número de endereços MAC (*unicast* ou *multicast*) para os quais pretende receber tramas (para além de um endereço *unicast* que pode ser registado na fase *Join*); permite verificar se os endereços registados são únicos
- *BUS Connect* – LEC estabelece *Default Multicast Send VCC* com BUS (após resolver o endereço *broadcast*); o BUS estabelece o primeiro *Multicast Forward VCC*

Resolução de endereços

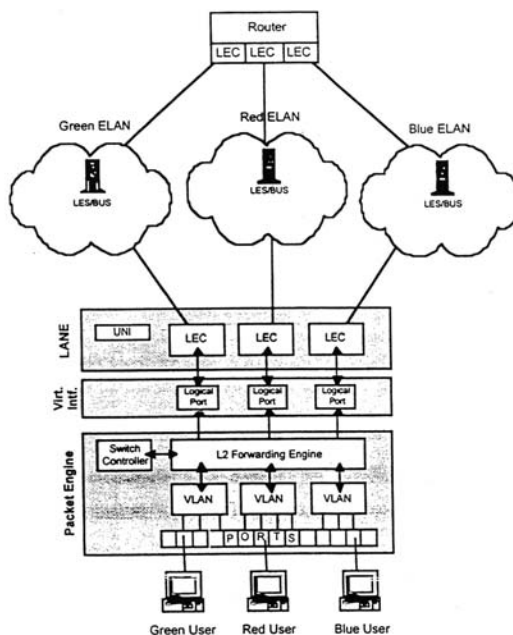
- A função de resolução de endereços consiste na determinação do endereço ATM (de outro LEC ou do BUS) associado a um endereço MAC (*unicast* ou *multicast*)
 - » No caso de uma *bridge* LAN/ATM, vários endereços MAC de estações na LAN são associados a um único endereço ATM
- Um LEC invoca a resolução de endereços (envia *LE_ARP request* no *Control Direct VCC*) para conhecer o endereço ATM de outro LEC ou do BUS (associado ao endereço *broadcast* ou a endereços *multicast*)
- O LES reage a um pedido do seguinte modo
 - » Se conhecer o endereço ATM correspondente ao endereço MAC, responde com *LE_ARP reply* ou opcionalmente envia o pedido ao LEC respectivo (e eventualmente aos restantes); a resposta é obrigatória na resolução do endereço *broadcast* do BUS
 - » Se não conhecer, envia o pedido para todos os LECs registados como *proxies* e eventualmente para os restantes (pode usar *Control Distribute VCC*); se receber uma resposta, reenvia-a ao LEC que invocou a resolução do endereço
- Um LEC deve responder a pedidos relativos a um endereço que registou ou para o qual é um *proxy*

Tráfego unicast

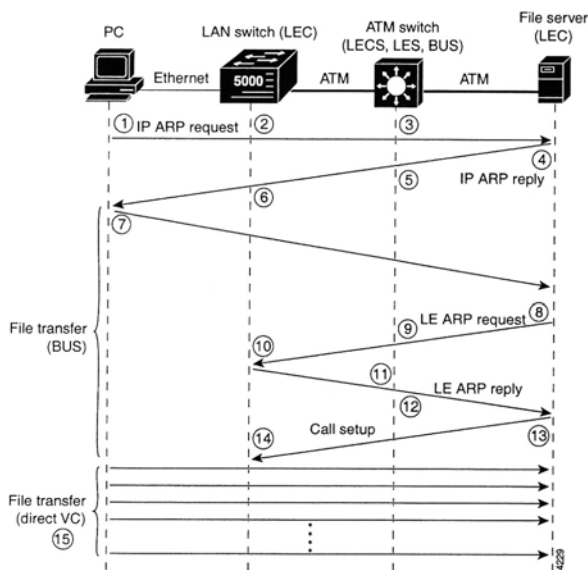
- O tráfego *unicast* entre dois LECs deve ser enviado de preferência num *Data Direct VCC*
 - » O estabelecimento de um *Data Direct VCC* pressupõe o conhecimento prévio do endereço ATM de destino (obtido pelo mecanismo de resolução de endereços)
 - » O endereço ATM de destino pode ser conhecido mas ainda não (ou já não) existir um *Data Direct VCC* estabelecido
- Se o endereço ATM de destino ainda não tiver sido resolvido ou se for conhecido mas ainda não tiver sido estabelecido o *Data Direct VCC* correspondente, o LEC pode enviar tráfego *unicast* para o BUS usando o *Multicast Send VCC*
 - » O BUS transmite pelo menos para o LEC de destino, se registado
 - » Se o LEC de destino não estiver registado, o BUS difunde o tráfego pelo menos para os clientes *proxy* (tipicamente *bridges* LAN/ATM) e opcionalmente para os restantes clientes

Bridge LAN/ATM

- A figura representa uma *bridge* LAN/ATM (*edge device*)
 - » Cada uma das ELANs faz parte de uma VLAN que integra estações de uma LAN convencional comutada
 - » Na *bridge*, cada ELAN é representada por um LEC, que deve ser associado à respectiva VLAN
- A comunicação entre estações de uma mesma VLAN é realizada por meio de mecanismos de nível 2 (*bridging*)
- A comunicação entre estações que pertencem a VLANs diferentes exige mecanismos de nível 3 (*routing*)
 - » Na figura é representado um único *router* ligado à rede ATM, pelo que deverá incluir um LEC por cada ELAN/VLAN a interligar



Exemplo



- Step 1 PC—Before starting the file transfer the PC must locate the file server on the network. To find the file server's MAC address, the PC broadcasts an ARP request with the file server's IP address.
- Step 2 LEC on Catalyst 5000 switch—Receives ARP requests and forwards to the BUS configured on the ATM switch router.
- Step 3 BUS on ATM switch router—Broadcasts the ARP request to all members of the emulated LAN using a point-to-multipoint VCC.
- Step 4 LEC on file server—Receives the ARP request, recognizes its own IP address and responds with an ARP reply back to the BUS in the ATM switch router.
- Step 5 BUS on ATM switch router—Forwards the ARP reply to the Catalyst 5000 switch.
- Step 6 LEC on Catalyst 5000 switch—Forwards the ARP reply to the originating PC.
- Step 7 PC—Starts sending the packets of the file transfer using the multicast send VCC from the Catalyst 5000 to the BUS on the ATM switch router, which forwards the packets over the multicast forward VCC to the file server. This gets the data moving in the interim until the data direct VCC is set up.
- Step 8 LEC on file server—Starts to set up the direct VCC to the Catalyst 5000 switch using an LE_ARP request to the LES. This request asks for the ATM address that corresponds to the PC's MAC address. (The PC's MAC address was obtained from the original ARP request in Step 4.)
- Step 9 LES on ATM switch router—Looks up the PC's MAC address in its look-up table and multicasts the LE_ARP request to all LECs.
- Step 10 LEC on Catalyst 5000 switch—Receives the LE_ARP request and finds the PC's MAC address in its look-up table. (It learned the PC's MAC address in Step 2.)
- Step 11 LEC on Catalyst 5000 switch—Adds its own ATM address into the LE_ARP request and returns it to the LES in the ATM switch router.
- Step 12 LES on ATM switch router—Multicasts the LE_ARP reply to all members of the emulated LAN, including the file server.
- Step 13 LEC on File Server—Receives the LE_ARP as part of the emulated LAN and signals for a data direct VCC to the Catalyst 5000 using the ATM address.
- Step 14 ATM switch router—Sets up a data direct VCC between the Catalyst 5000 and the file server.
- Step 15 PC—The file transfers directly from the PC using the direct data VCC from the Catalyst 5000 to the ATM-attached file server.

MPOA

Multiprotocol Over ATM

MPOA

“The primary goal of MPOA is the efficient transfer of inter-subnet unicast data in a LAN environment. MPOA integrates LANE and NHRP to preserve the benefits of LAN Emulation, while allowing inter-subnet, Internetwork Layer protocol communication over ATM VCCs without requiring Routers in the data path. MPOA provides a framework for effectively synthesising bridging and routing with ATM in an environment of diverse protocols, network technologies, and IEEE 802.1 Virtual LANs. This framework is intended to provide a unified paradigm for overlaying Internetwork Layer protocols on ATM. MPOA is capable of using both routing and bridging information to locate the optimal exit from the ATM cloud.”

Objectivos

- MPOA define uma arquitectura que permite o transporte eficiente de protocolos de nível 3 directamente sobre ATM, em ambientes LAN
- MPOA é uma extensão de LANE e, para além de *bridging*, fornece uma solução para o transporte de pacotes entre subredes, com eliminação de *routers* no percurso dos dados
- MPOA mapeia protocolos de nível 3 em ATM e permite comunicação directa entre dispositivos MPOA em diferentes subredes, através de atalhos (*short-cuts*) providenciados por Circuitos Virtuais (VCs) ATM
- MPOA permite a criação de subredes virtuais na camada de rede

MPOA como extensão de LANE

- LANE torna ATM transparente em subredes de nível 2 (LANs)
 - » Emula LANs em ATM (ELANs)
 - » Permite, através de *bridges* / comutadores ATM/LAN, generalizar o conceito de VLAN, mapeando ELANs ATM em VLANs
 - » Fornece serviço MAC (transparente a protocolos de nível 3)
- LANE realiza *bridging* de tráfego na mesma subrede virtual de nível 2 (VLAN); tráfego entre subredes é transferido através de *routers*, o que torna a solução dificilmente escalável, para além da degradação de desempenho e da não exploração de QoS ATM
- MPOA generaliza o conceito, tornando ATM transparente em redes de nível 3 – emula *routing* sobre ATM e oferece um modo unificado para executar protocolos de nível 3 sobre ATM

Princípios arquitectónicos

- MPOA estende o conceito de subredes virtuais de nível 2 (VLANs), baseadas em LANE, a subredes virtuais de nível 3 (sem restrições quanto à localização física); elimina a necessidade de *routers* para interligar VLANs, mas permite interfuncionamento com *routers* convencionais
- MPOA permite estabelecer *short-cuts* ATM entre sistemas que pertencem a diferentes subredes virtuais de nível 2 (VLANs), enquanto LANE permite estabelecer VCs ATM entre sistemas que pertencem à mesma ELAN
- Do ponto de vista dos protocolos de nível 3 existe um único salto lógico na rede ATM (conectividade extremo a extremo directamente sobre ATM), sem existência de *routers* no percurso, com a inerente melhoria de desempenho, visto evitar-se o processamento pacote a pacote nos *routers*
- MPOA opera nos níveis 2 e 3, usando LANE para comunicação no nível 2 (mesma ELAN) e estabelecendo VCs ATM directos para comunicação no nível 3 entre subredes virtuais

Funções MPOA

- MPOA baseia-se na separação das funções de
 - » *Routing* – execução de protocolos de encaminhamento, criação e manutenção de tabelas de encaminhamento
 - » *Switching (forwarding)* – transporte de pacotes de dados em VCs ATM (sinalização e comutação ATM)
- MPOA permite a coexistência de diferentes protocolos de rede (IP, IPX, ...) e de encaminhamento (OSPF, RIP, BGP,...)
- Para o estabelecimento de *short-cuts* ATM, MPOA usa NHRP (*Next Hop Resolution Protocol*), que permite estender os protocolos de resolução de endereços a redes NBMA (*Non Broadcast Multiple Access*)
- MPOA baseia-se num modelo cliente / servidor suportado nas funções
 - » MPOA *Client* (MPC)
 - » MPOA *Server* (MPS)
- Visto que MPOA é uma extensão de LANE, os dispositivos MPOA incluem o componente LEC; os MPS suportam a função NHS – *Next Hop Server*

Comunicação entre dispositivos MPOA

- Um ambiente MPOA é constituído pelos seguintes dispositivos
 - » *Hosts*
 - » *Routers*
 - » *Edge Devices*
- Dispositivos MPOA podem estabelecer VCs ATM entre si
 - » Para transporte de dados (em tramas de nível 2) na mesma subrede (ELAN)
 - » Para transporte de dados (em pacotes de nível 3) entre subredes (*short-cuts*)
- Os dispositivos MPOA suportam LANE, embora possam existir no ambiente MPOA dispositivos que suportam LANE mas não MPOA
- A comunicação de dispositivos MPOA com dispositivos que apenas suportam a função LEC é baseada em LANE, sendo realizada directamente (*bridging*) se pertencerem à mesma subrede (VLAN), ou obrigatoriamente através de *Routers* MPOA se pertencerem a subredes diferentes (neste caso não é possível estabelecer *short-cuts*)

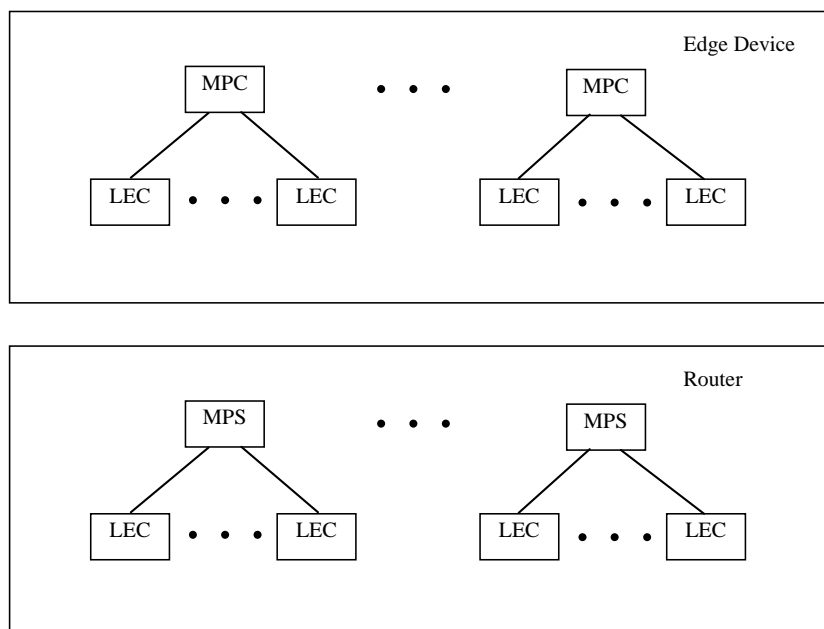
MPOA Client (MPC)

- Reside em *Hosts* e em *Edge Devices* MPOA – pode servir um ou mais LECs e comunicar com um ou mais MPSs
- Termina *short-cuts* ATM estabelecidos para comunicação entre subredes
- Realiza funções de transferência de pacotes (*forwarding*) mas não executa funções de encaminhamento (*routing*)
- Em nós de ingresso, antes de decidir estabelecer um *short-cut*, detecta fluxos de pacotes enviados para um *router default* (com função MPS) através da ELAN correspondente (*bridging*)
 - » Um fluxo pode ser definido com base em endereços, tipo de protocolo, *sockets*, etc., e caracteriza-se por um número pré-definido de pacotes de um dado tipo observados num determinado intervalo de tempo (e.g., 10 pacotes durante 1 s)
- Para estabelecer um *short-cut* um MPC interacciona com um MPS para descobrir o endereço ATM do dispositivo MPOA de destino

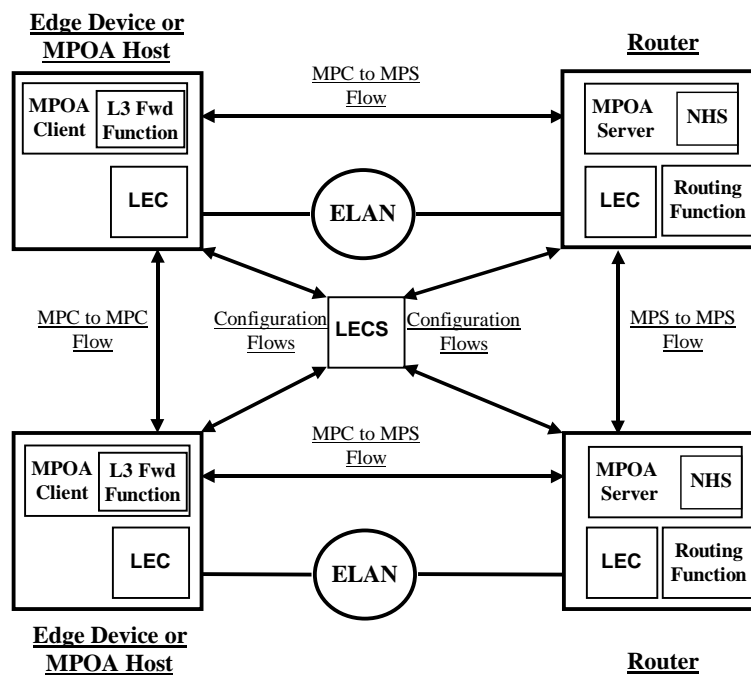
MPOA Server (MPS)

- É um componente lógico de *Routers* MPOA – equipamento com as funções de um *router* convencional e com funcionalidade MPS
- Um MPS inclui um NHS
- Um MPS interactiva com o NHS local e com as funções de encaminhamento para responder a pedidos MPOA de MPCs
 - » Converte os pedidos / respostas MPOA em pedidos / respostas NHRP
 - » Realiza resolução de endereços de nível 3 (e.g., um endereço IP) em endereços ATM
- Fornece aos MPCs a informação que permite o estabelecimento de *short-cuts* e subsequente transferência de pacotes no VC respectivo (*forwarding*)

MPC, MPS e LEC em Edge Devices e Routers



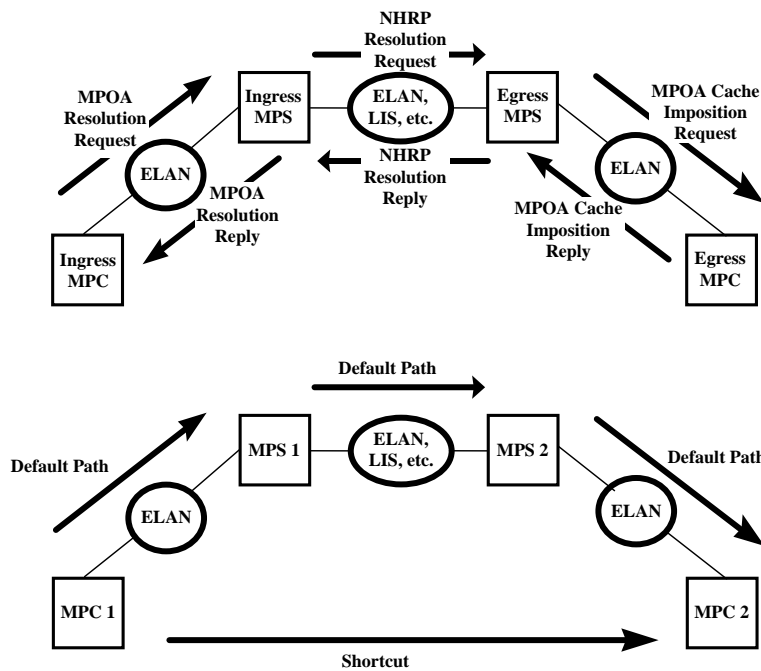
Fluxos de informação entre sistemas MPOA



Estabelecimento de atalhos (short-cuts)

- MPOA permite estabelecer atalhos (*short-cuts*) para comunicação entre subredes
- Antes de ser estabelecido um *short-cut*, um MPC transfere de forma transparente pacotes entre portas lógicas (que representam subredes) e LECs associados; é usado LANE, quer para comunicação na mesma subrede quer para enviar para um *default router* o tráfego entre subredes
- Durante este processo o MPC monitora o fluxo de pacotes enviado para o *router* (via ELAN) e determina se (e quando) um fluxo pode beneficiar de um *short-cut*
- No caso de não existir já um *short-cut* estabelecido, interroga o MPS para resolver o endereço de destino num endereço ATM; quando obtém uma resposta positiva, o MPC guarda a informação em *cache*, estabelece o VC (*short-cut*) e passa a transferir os pacotes nesse VC (*forwarding*)

Resolução de endereços e encaminhamento



Resolução de endereços e encaminhamento

MPOA Terms

MPOA Term	Definition
MPOA resolution request	A request from an MPC to resolve a destination protocol address to an ATM address to establish a shortcut VCC to the egress device.
NHRP resolution request	An MPOA resolution request which has been converted to an NHRP resolution request.
MPOA cache-imposition request	A request from an egress MPS to an egress MPC providing the MAC rewrite information for a destination protocol address.
MPOA cache-imposition reply	A reply from an egress MPC acknowledging an MPOA cache-imposition request.
NHRP resolution reply	An NHRP resolution reply that eventually will be converted to an MPOA resolution reply.
MPOA resolution reply	A reply from the ingress MPS resolving a protocol address to an ATM address.
Shortcut VCC	The path between MPCs over which Layer-3 packets are sent.

Ingress e Egress Caches

Ingress Cache

Keys		Contents		
MPS Control ATM Address	Internetwork Layer Destination Address	Dest. ATM Address or VCC	Encapsulation Information	Other information needed for control (e.g. Flow Count and Holding Time)

Egress Cache without Tags

Keys		Contents		
Internetwork Layer Destination Address	Source/Dest. ATM Addresses	LEC	DLL header	Other information needed for control (e.g. Holding Time)

Egress Cache with Tags

Keys			Contents		
Internetwork Layer Destination Address	Source/Dest. ATM Addresses	Tag	LEC	DLL header	Other information needed for control (e.g. Holding Time)

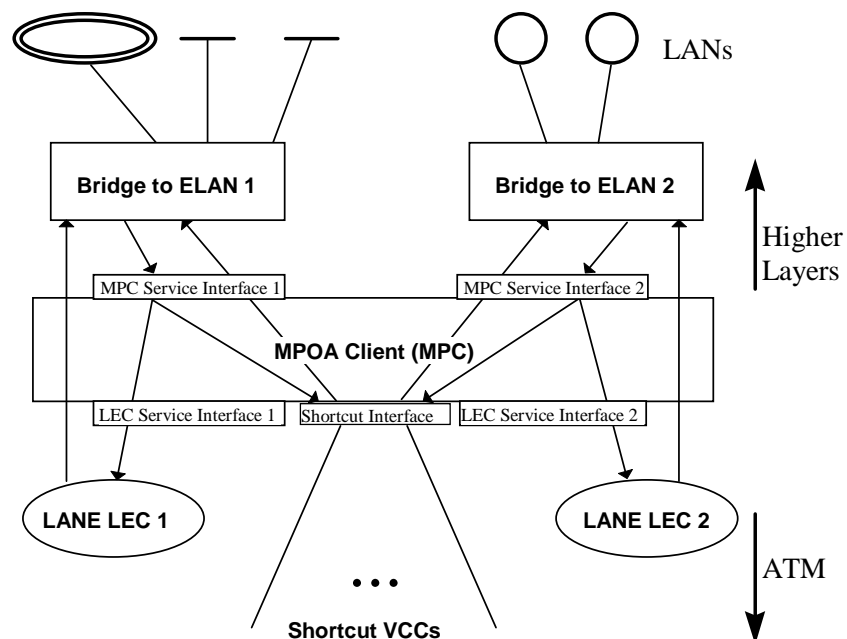
Host MPOA

- Um *Host* MPOA é um dispositivo MPOA apenas ligado à rede ATM e inclui
 - » Um ou mais LECs (permite LANE)
 - » Um ou mais MPCs (permite comunicação com outras subredes)
 - » Uma ou mais pilhas protocolares completas (*end system*)
- Pode pertencer a uma ou mais ELANs (ou, mais geralmente, VLANs)
- A comunicação com outros dispositivos MPOA é baseada em LANE, se pertencerem à mesma subrede (VLAN), ou em MPOA, se pertencerem a subredes diferentes

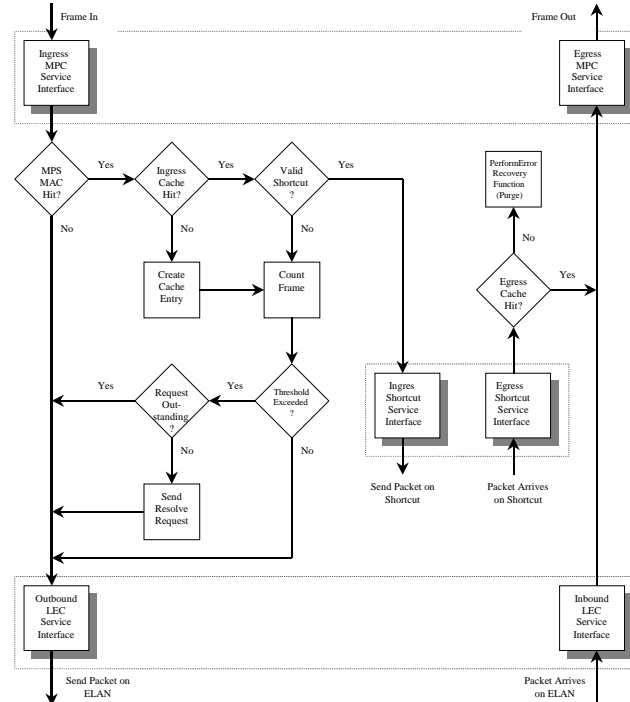
Edge Device

- É um dispositivo na fronteira entre a rede ATM e uma ou mais LANs e inclui
 - » Um ou mais LECs (permite LANE)
 - » Um ou mais MPCs (permite comunicação com outras subredes)
 - » Uma ou mais portas para LANs (*bridge ports*)
- Um *Edge Device* estabelece VCs com outros dispositivos MPOA mas não participa em protocolos de encaminhamento
- Um *Edge Device* comporta-se como uma *bridge* para tráfego numa mesma subrede (VLAN), envolvendo um LEC próprio, que representa estações não ATM nessa subrede (VLAN), e outros LECs que podem residir em *Hosts* ou *Routers* MPOA, outros *Edge Devices* ou dispositivos LANE não MPOA
- Para tráfego entre subredes (VLANs) é usado MPOA

Interfaces de um MPC num Edge Device



Fluxos de dados num Edge Device



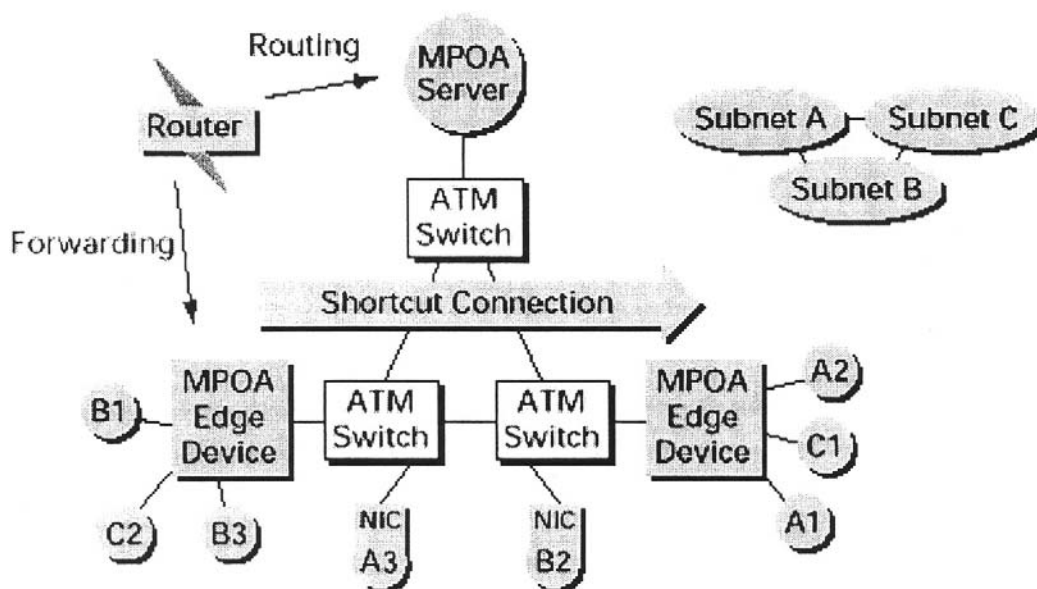
Router MPOA

- É um dispositivo que permite comunicação entre subredes virtuais com base em protocolos de nível 3 – pode ser baseado em *hardware* dedicado ou em *software* instalado em *routers* ou comutadores
- Um *Router MPOA* inclui
 - » Um ou mais LECs
 - » Um ou mais MPSs (e respectivos NHSs)
 - » Funções de encaminhamento
- Executa protocolos de encaminhamento e mantém tabelas de encaminhamento
 - » Funciona como *router* convencional na comunicação com dispositivos que não suportam MPC (apenas LEC) ou para encaminhar pacotes entre subredes antes do estabelecimento de *short-cuts*
 - » Funciona como *route server* (MPS), realizando a resolução de endereços de nível 3 em endereços ATM, o que permite o estabelecimento de *short-cuts* pelos MPCs

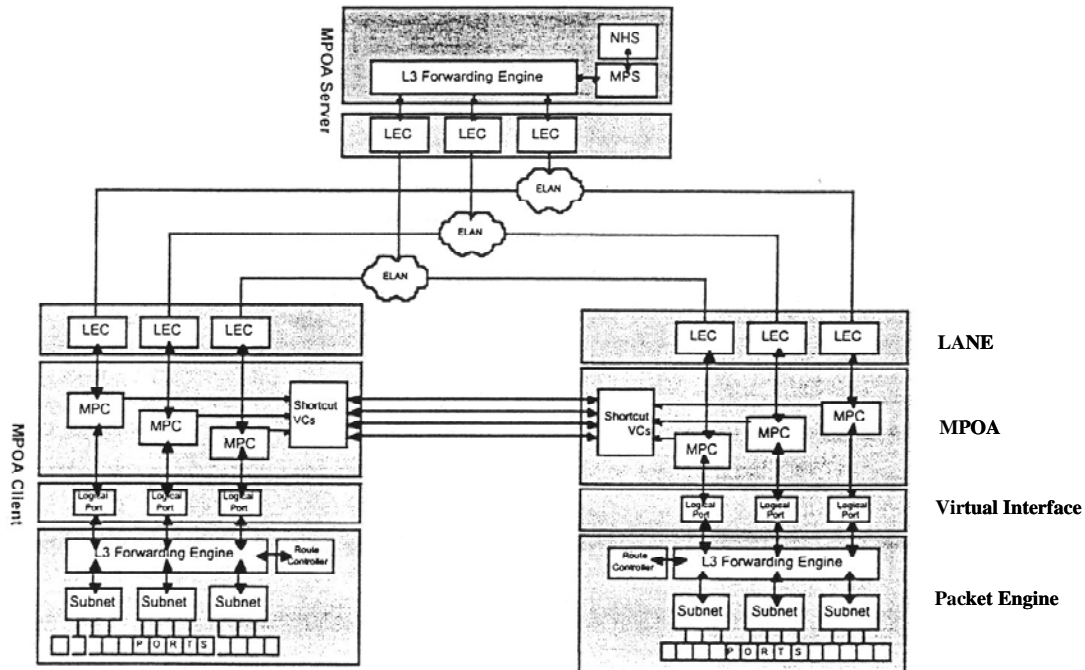
Routing Virtual

- O conceito de *Routing Virtual* baseia-se na separação das funções de *Routing* e *Forwarding* – respectivamente da responsabilidade de MPSs e de MPCs
- As funções de um *router* convencional estão distribuídas; por analogia
 - » Os *Edge Devices* constituem as cartas de interface do *router*
 - » A rede ATM constitui o *backplane* do *router*
 - » O *route server* (MPS) constitui o processador de controlo
- Em MPOA é adoptado um modelo designado *overlay layered routing* – as funções de *routing* e de *switching* (ATM) não estão integradas, isto é, não partilham informação topológica nem usam os mesmos protocolos de encaminhamento
 - » Os *routers* executam protocolos de encaminhamento associados aos protocolos de nível 3 a transportar na rede
 - » Os comutadores ATM estabelecem VCs com base em procedimentos de sinalização e de protocolos de encaminhamento específicos do ATM (por exemplo, PNNI)

Routing Virtual – separação de funções



Exemplo



Conclusões

MPOA tem as seguintes características

- Suporta VLANs (LANE) e garante conectividade global característica de protocolos de nível 3 (*internetworking*)
- Oferece uma forma universal de comunicação para protocolos de nível 3 sobre ATM
- Unifica os mecanismos de *bridging* e *routing* em ambientes caracterizados pela existência de VLANs, múltiplos protocolos de rede e várias tecnologias de subrede
- Permite explorar ATM de forma eficiente, uma vez que suporta comunicação directa sobre VCs (sem *routers* no percurso de dados), com possibilidade de negociação de QoS por VC
- Separa e distribui as funções de *routing* e *switching* por vários dispositivos
- Requer modificações mínimas em *hubs*, *bridges* e *routers*
- É vocacionado para LANs, mas não para WANs