

Características do Padrão IEEE 802.17 e Implementação Utilizando Gigabit Ethernet

Caciano dos Santos Machado, Daniel Seelig, Roberto Ely Scherer, Tórgan F. Siqueira¹

¹Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)

{caciano,danielseelig}@gmail.com, {rescherer,torgan}@inf.ufrgs.br

Resumo. *O IEEE 802.17 é o padrão que introduz uma nova arquitetura de rede com topologia em anel, denominada RPR (Resilient Packet Ring), destinada a redes metropolitanas e de longa distância. Este artigo apresenta uma visão geral do funcionamento e características desta nova tecnologia, tais como a arquitetura, reuso espacial, descoberta de topologia, resiliência e nível físico utilizado. São apresentadas as idéias básicas, com vistas a ilustrar o potencial de uso de Gigabit Ethernet (GE) e 10 GE sobre RPR, com ênfase no suporte físico provido por estas duas tecnologias.*

1. Introdução

A topologia em anel é muito utilizada em redes metropolitanas (*Metropolitan Area Networks* - MANs). Exemplo disso são as redes SONET/SDH (*Synchronous Optical Networks*), que trabalham com circuitos dedicados garantindo largura de banda, atraso e variação de atraso, além de recuperação rápida (<50 ms) e transparente em caso de falhas. Utilizar um anel duplo, onde um deles atua exclusivamente como *backup* do ativo e permanece inativo até uma ocorrência de falha no primário, torna ineficiente o uso da banda [Davik et al. 2004]. Outro exemplo são as MANs Gigabit Ethernet (GE) em anel, que oferecem baixo custo, facilidade de gerenciamento e simplicidade de integração com os equipamentos existentes. Entretanto, não permitem a recuperação automática do enlace (por exigirem reconstrução da árvore do algoritmo *Spanning Tree*) e nem laços fechados. Também não permitem um sistema de distribuição justa de largura de banda (*fairness*) e priorização de tráfego, pois não garantem banda, atraso e variação de atraso, tal qual SONET [Gambiroza et al. 2004].

Neste contexto, foi desenvolvida a tecnologia *Resilient Packet Ring* (RPR) [IEEE 802.17 2004], que busca agregar as melhores características e resolver alguns problemas das redes em anel existentes. As redes RPR são otimizadas para transmitir pacotes em anel e para otimizar o uso da banda através de reuso espacial. Elas suportam a transmissão de dados em *unicast*, *multicast* e *broadcast* e garantem a banda, atraso e variação de atraso através da definição de classes de dados. Além disso, distribuem justamente a banda não alocada, descobrem automaticamente a topologia e recuperam-se rapidamente em caso de falhas.

Neste artigo, são apresentadas as características do padrão IEEE 802.17 e a viabilidade de utilizar a tecnologia GE para implementação de anéis RPR. A seção 2 apresenta algumas características da arquitetura do padrão. Na seção 3 é descrita subcamada MAC (*Media Access Control*) e na seção 4, a camada física. Finalizando, são apresentadas algumas conclusões com respeito à utilização sobre GE.

2. Arquitetura do Padrão IEEE 802.17

As redes RPR são constituídas por, no máximo, 255 estações interligadas por dois anéis (*ringlets*) de fibra óptica, sendo que os dados trafegam em direções opostas em cada um. A estação transmissora decide em qual *ringlet* os dados serão transmitidos para o receptor, geralmente utilizando o menor caminho [Spadaro et al. 2004]. Se uma estação não reconhece o endereço de destino de um quadro, então ele é repassado para a estação seguinte, pois o *Time-To-Live* (TTL) limita a circulação de quadros com endereços de destino desconhecidos.

No RPR, os métodos de transmissão de quadros suportados são o *cut-through*, no qual eles começam a ser repassados antes de serem completamente recebidos, e o *store-and-forward*, que os armazena antes de repassar. Quando uma estação RPR recebe um quadro, ele é removido, diferentemente de outras redes em anel, nas quais seu conteúdo é copiado e ele percorre o anel até retornar à estação de origem. Esta remoção permite que outras estações transmissoras aproveitem a banda que seria ocupada pelo quadro no caminho de volta. Esta característica é conhecida como reuso espacial. A figura 1 ilustra um cenário onde o reuso espacial é obtido em um dos anéis. A estação 2 está transmitindo para a estação 4 enquanto a estação 6 está transmitindo para a estação 9.

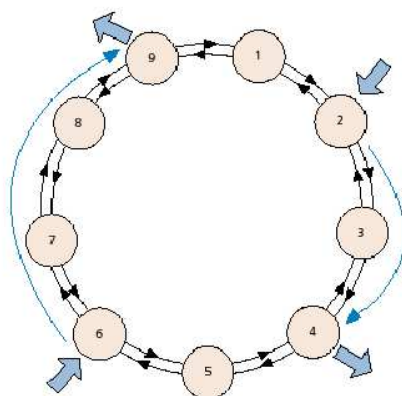


Figura 1. Reuso espacial em anéis [Davik et al. 2004]

3. Subcamada MAC

A subcamada MAC controla o acesso dos clientes ao anel, provendo pontos de acesso para três diferentes classes de serviços para troca de dados. Destas três classes, duas ainda se subdividem para formação de subclasses. Cada classe possui conformadores de tráfego separados para cada anel, objetivando o controle de fluxo. As características de cada classe e suas respectivas subclasses são apresentadas na tabela 1.

Para garantir as características atribuídas às subclasses A0, A1 e B-CIR, a largura de banda de cada uma delas deve ser pré-alocada, nunca excedendo o limite máximo provido pelo enlace de cada estação. A banda reservada às aplicações da subclasse A0 lhes é exclusiva e será desperdiçada se não for completamente usada. As subclasses A1 e B-CIR terão banda disponível quando requisitarem. Se não for utilizada, estará disponível para o tráfego de dados da subclasse B-EIR e classe C. Estas duas classes de tráfego são denominadas *fairness eligible*, pois são controladas pelo algoritmo *fairness* [Davik et al. 2004].

Tabela 1. Classes de serviços do RPR e suas respectivas características [IEEE 802.17 2004]

classe de serviço			qualidades do serviço				fairness eligible
classe	exemplo de uso	subclasse	banda garant.	variação de atraso	banda		
					tipo	subtipo	
A	tempo real	A0	sim	baixo	alocado	reservado	não
		A1	sim	baixo	alocado	requisitado	não
B	quase tempo real	B-CIR	sim	vinculado	alocado	requisitado	não
		B-EIR	não	não vinculado	oportunistico	requisitado	sim
C	melhor esforço	—	não	não vinculado	oportunistico	requisitado	sim

As subclasses A0, A1 e B-CIR não possuem reuso espacial definido no padrão, mas nada impede que isto seja provido.

A subcamada MAC também provê filas para entrada de dados e de trânsito. O padrão define três filas de entrada de dados (uma para cada classe) e uma ou duas de trânsito - *Primary Transit Queue* (PTQ) e *Secondary Transit Queue* (STQ). Na solução com duas filas de trânsito, os quadros de alta prioridade (classe A) são enfileirados na PTQ e os de menor prioridade (classes B e C) são enfileirados na STQ. O repasse de quadros da PTQ tem prioridade sobre o da STQ e sobre o da maioria dos tipos de quadros introduzidos pela estação. Assim, um quadro da classe A não vai sofrer muito mais do que o atraso de propagação e alguns atrasos ocasionais nas filas de trânsito. Já os da STQ tem prioridade sobre os das filas de entrada das classes B e C. A figura 2 mostra uma interface do anel com três filas de egresso de quadros e duas para trânsito. Os números indicam a prioridade que as filas teriam no enlace de transmissão [Davik et al. 2004].

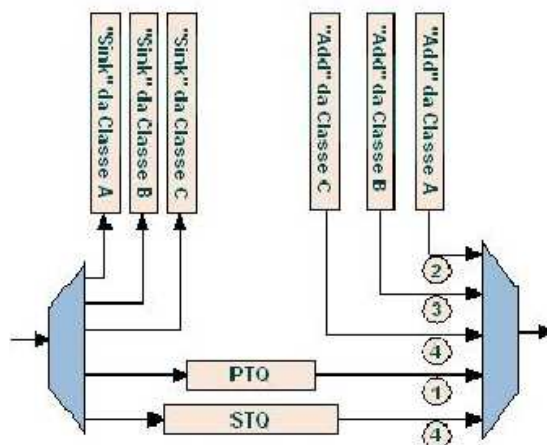


Figura 2. Fluxos de dados e a respectiva prioridade das fila [Davik et al. 2004]

3.1. Algoritmo *Fairness*

O objetivo do algoritmo *fairness* é distribuir de forma justa a banda não utilizada (com exceção da banda alocada para a subclasse A), entre as estações que compõem o anel

RPR, para o tráfego de dados das classes B-EIR e C. Quando o enlace de saída de uma estação estiver congestionado, o algoritmo envia uma mensagem de *fairness* para o enlace oposto do anel. As estações que receberem esta mensagem devem diminuir a sua taxa de envio de acordo com o que está especificado na mensagem. Esta taxa foi calculada para uma justa distribuição da banda do enlace na estação onde ocorreu o congestionamento.

Existem dois métodos previstos no padrão para estimação da taxa de envio: o conservador e o agressivo. No conservador, quando uma estação está congestionada, mensagens de *fairness* são enviadas em intervalos dinâmicos. A taxa informada nestas mensagens é dada pela divisão da banda disponível pelo número de estações que estão transmitindo ou pela utilização das propriedades da estação para inserir dados no anel. Neste método, o intervalo dinâmico é calculado de acordo com o tempo que as outras estações do anel levam para adequar suas taxas de transmissão de forma que isto possa ser percebido pela estação que está congestionada. Desta forma, mensagens informando que não há mais congestionamento fazem com que o tráfego vá aumentando gradativamente.

No método agressivo, quando uma estação está congestionada, mensagens *fairness* são enviadas em intervalos pré-determinados, com uma taxa aproximada à que a estação pode adicionar dados ao anel. Quando cessa o congestionamento, mensagens *fairness* são enviadas pela mesma estação informando que não há congestionamento. Isto faz com que as estações que recebam esta mensagem aumentem sua taxa de transmissão gradativamente, porém de forma mais rápida, já que a estação antes congestionada não espera para ver o resultado antes de enviar outras mensagens [Davik et al. 2004]. Este rápido ajuste é menos restritivo sobre a taxa de transmissão em relação ao método conservador. Estações com uma única fila de trânsito precisam de um método mais restritivo, sendo normalmente utilizado o método conservador neste caso. Quando estações possuem duas filas de trânsito, normalmente o método agressivo é utilizado. Independente do método escolhido, o algoritmo não descarta pacotes.

3.2. Descoberta de Topologia

A descoberta de topologia determina a conectividade e o ordenamento das estações em torno do anel [Davik et al. 2004]. De acordo com o padrão [IEEE 802.17 2004], a descoberta ocorre na inicialização, quando entra uma nova estação no anel, periodicamente e quando é detectada uma falha em um enlace (ou estação). O protocolo de descoberta recolhe informações sobre as estações e as armazena na base de dados de cada estação. Quando o sistema inicializa, cada estação envia um quadro de controle, a mensagem de descoberta de topologia, que contém inicialmente só o seu próprio estado. Estas mensagens são enviadas por todo o anel em ambos os *ringlets*, com um TTL inicial de 255, que é decrementado antes de serem repassadas nas demais estações. Quando uma nova estação é inserida, ou quando uma detecta falha no enlace, ou quando ela recebe uma mensagem de topologia inconsistente com sua imagem, ela imediatamente envia uma mensagem de descoberta. A base de dados inclui a ordem das estações e o estado da proteção, isto é, enlaces conectados e seu estado: *fail*, *degrade* ou *idle*. A estabilização da informação de topologia ocorre quando a imagem não muda dentro de um certo intervalo de tempo. Mesmo estáveis e consistentes, as estações de tempos em tempos transmitem mensagens de descoberta de topologia, provendo robustez à operação do anel. A base de dados tem outros dois usos: o MAC a usa para encontrar o caminho mais curto quando um cliente envia um quadro sem especificar o *ringlelet* a ser usado, e também para calcular o *Frame*

Round Trip Time (FRTT) no modo conservativo do algoritmo *fairness*. O mecanismo de descoberta de topologia e o de proteção compartilham recursos de controle de mensagens.

3.3. Mecanismo de Proteção

A resiliência do RPR está baseada no mecanismo de proteção que, segundo o padrão, é rápido, robusto, flexível e eficiente. O anel duplo provê um caminho alternativo no caso de colapso de um enlace ou de uma estação, e os métodos de resposta a falhas incluem *steering*, *wrapping* e *passthrough*. O *steering* é obrigatório, e no caso de falhas, redireciona o tráfego protegido pelo laço ainda conectado ao destino. O *wrapping* é opcional, e redireciona o tráfego protegido pelo *ringlet* oposto em caso de falhas, fechando o laço no ponto de ruptura. O *passthrough* é opcional e torna uma estação que tenha se auto-detectado internamente falha num repetidor transparente, isto é, numa simples fila FIFO. O *wrapping* pode ser usado como uma medida imediata, para minimizar a perda de pacotes, enquanto que o *steering* pode ser usado para otimizar o *ringlet* a longo prazo. Quadros não marcados elegíveis para *wrapping* são sempre descartados num *wrap point*.

Os quadros que o RPR transmite em modo *strict* devem chegar na mesma ordem em que foram enviados. Para manter a entrega ordenada após uma falha, todas as estações devem parar de transmitir pacotes e descartar os que estejam em trânsito, até que a nova imagem da topologia estabilize e esteja consistente novamente. O tempo entre a detecção da falha e a estabilização é chamado tempo de restauração do anel ou tempo de convergência do algoritmo. Antes deste tempo, que o RPR estabelece como sendo inferior a 50 ms¹, a estação não pode fazer *steering* de novos quadros. Para minimizar o descarte de pacotes, é definido um modo opcional de transmissão de pacotes, o relaxado. Neste modo, pacotes que chegam fora de ordem são tolerados, podem ser enviados logo após uma falha e antes da consistência e não são descartados das filas de transmissão.

4. Camada Física

Para que o RPR fosse rapidamente desenvolvido e utilizado, não foi criada nenhuma nova tecnologia para a camada física. Para suporte nesta camada, o padrão pode utilizar tanto SONET/SDH como GE e 10 GE, ambos com fibra óptica. Para isto, foram apenas definidas subcamadas de reconciliação, que fazem o mapeamento destas tecnologias da camada física com a de controle de acesso ao meio, conforme demonstrado na figura 3.

Como SONET/SDH são atualmente as tecnologias mais utilizadas nas redes MAN, é possível que a adoção do RPR sobre elas seja mais comum inicialmente, pois não é necessária a substituição da estrutura existente. Porém, deve-se destacar a possibilidade da utilização de RPR sobre as redes nativas GE e 10 GE (que utilizam fibra óptica), baseadas nos padrões 802.3 e 802.3ae. Para isto, são necessárias algumas exceções e mudanças em relação às especificações nativas [IEEE 802.17 2004]:

- Não podem ser utilizados repetidores;
- quadros com tamanho mínimo de 16 e máximo de 9.216 bytes;
- não é utilizada a autonegociação (transmissão é sempre *full-duplex*); e
- controle de fluxo é desabilitado ([O'Connor and Bruckman 2002]).

Considerando estes pontos (e a subcamada de reconciliação), as tecnologias GE e 10GE fornecem suporte adequado para RPR funcionar de acordo com suas especificações.

¹Não incluso o tempo de detecção de falha nem o de configuração

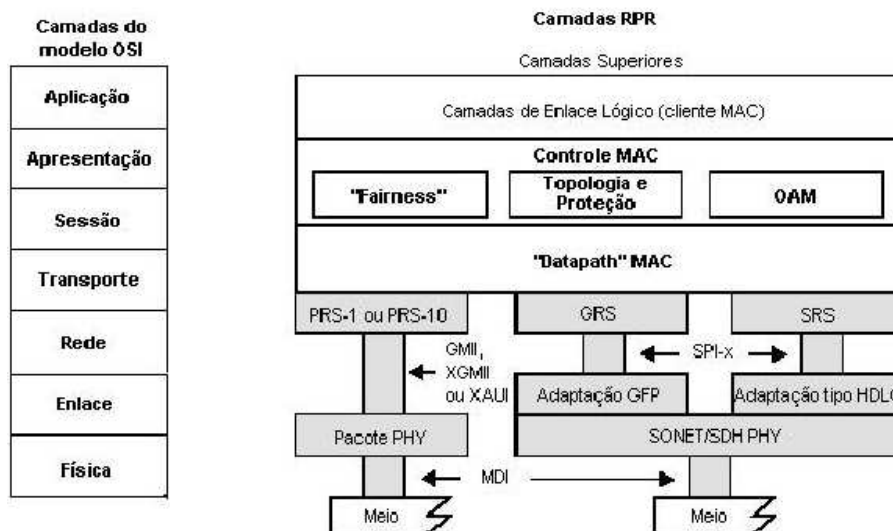


Figura 3. Camada física RPR relacionada ao modelo de referência OSI

5. Conclusão

Neste artigo foram descritas diversas características do RPR que justificam a sua adoção em relação a outras tecnologias com topologia em anel. Como exemplo, pode-se citar o melhor aproveitamento da banda e sua distribuição justa para serviços do tipo melhor esforço. Embora SONET/SDH seja o suporte comumente utilizado nos locais onde o RPR já foi adotado, espera-se que novas estruturas sejam desenvolvidas, principalmente para ambientes GE e 10 GE, que trazem vantagens como menor custo e complexidade. Neste contexto, estendem-se as possibilidades de uso RPR. Nas WANs é possível manter um nível de serviço semelhante ao já provido com uma estrutura bem menos complexa. Nas LANs corporativas o RPR possibilita agregar proteção e garantias de serviços para aplicações de tempo real a custos compatíveis.

Referências

- Davik, F., Yilmaz, M., Gjessing, S., and Uzun, N. (2004). IEEE 802.17 Resilient Packet Ring Tutorial. *IEEE Communications Magazine*, 42(3):112–118.
- Gambiroza, V., Yuan, P., and Knightly, E. (2004). IEEE 802.17 Media Access Protocol for High-Speed Metropolitan-Area Resilient Packet Rings.
- IEEE 802.17 (2004). IEEE Std 802.17 – 2004, IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements - Part 17: Resilient packet ring (RPR) access method and physical layer specifications.
- O'Connor, M. and Bruckman, L. (2002). Corrigent Systems: IEEE 802.17 - Resilient Packet Rings - Part 2 of 4 - The PHY, Protection and Transit Paths.
- Spadaro, S., Sole-Pareta, J., Careglio, D., and Szymanski, K. W. A. (2004). Positioning of the RPR standard in contemporary operator environments. *IEEE Network*, 18(2):35 – 40.