

Testes do IBM 8265 para Piloto RNP2

Paulo Henrique Aguiar Rodrigues [IM/NCE/UFRJ], João Carlos Peixoto de Almeida da Costa [IM/NCE/UFRJ],
Cláudia Medronho Naumann [IM/NCE/UFRJ]

Núcleo de Computação Eletrônica – UFRJ, Caixa Postal 2324, CEP 20001 – 970, Rio de Janeiro – RJ – Brasil

Email : naumann@nce.ufrj.br

Área WNRP2 : Engenharia/Medições

Antecedentes

Durante o ano 2000, o grupo acima mencionado teve participação ativa na avaliação do comutador IBM 8265, equipamento utilizado pela RNP e pelas RMAVs para a implantação do backbone RNP2. O objetivo principal foi avaliar, através de testes, algumas características básicas de funcionamento do comutador, como por exemplo, *traffic shaping*, alocação de *buffers* por VCs (Virtual Circuits), descarte seletivo, configuração de VCs com QoS em VP (Virtual Path) CBR ou VBR, funcionamento de ABR, entre outras. Para a geração e monitoração dos testes foi usada a ferramenta TMG, desenvolvida do NCE. A apresentação visa mostrar a experiência adquirida nestes testes e contribuir para boa prática.

Experimentos Realizados

O ambiente utilizado para a realização dos experimentos encontra-se abaixo:

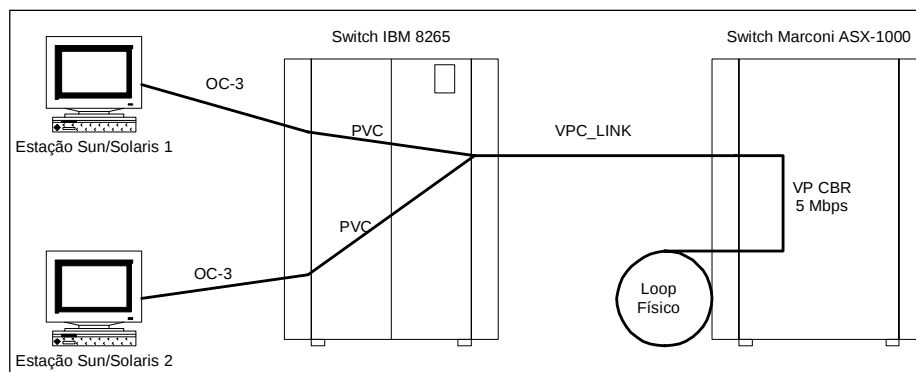


Fig. 1 – Ambiente de Testes.

Para avaliar o funcionamento do comutador, realizamos testes de estresse em um VP (VPC_LINK) com monitoramento de descarte de células. Os testes foram divididos em três etapas:

1. VC UBR estressando a capacidade do VP;
2. VC UBR na presença de um VC CBR de 2 Mbps com tráfego a 50 % do PCR (1 Mbps), com o VC UBR estressando a capacidade do VP;
3. VC UBR na presença de um VC CBR com tráfego a 100 % do PCR (2 Mbps), com o VC UBR estressando a capacidade do VP.

Os testes foram executados utilizando-se estações Solaris, um comutador IBM 8265 e um comutador FORE ASX-1000 conforme mostrado na fig.1. Todos os equipamentos foram conectados através de fibra multimodo. Foram criados inicialmente dois VPs, um no comutador IBM, chamado de VPC_LINK (que foi criado para permitir a passagem de tráfegos UBR, CBR e VBR), e outro no comutador FORE, chamado simplesmente de VP (CBR), ambos com 5 Mbps. Os PVCs (CBR e UBR) eram criados no comutador IBM de acordo com os requisitos de cada tipo de teste. Para verificarmos o comportamento do

comutador em relação à moldagem de tráfego, habilitamos o *shaping* no VP do IBM, e apenas o policiamento no VP do comutador FORE para verificar o descarte de células. Como o comutador IBM possuía dois tipos de módulos (Módulo Fibra OC-3 Nativa e Módulo WAN 8260¹), executamos cada etapa de testes duas vezes, uma para cada módulo, a fim de compararmos os resultados.

Para estressarmos o VP utilizamos a ferramenta TMG. Os testes consistiram portanto, na utilização da ferramenta TMG para a geração de tráfego na estação 1 (e na estação 2, dependendo do tipo de teste), passando pelos VPs estabelecidos nos dois comutadores, retornando à estação 1 através de um loop físico colocado no comutador FORE. Valores crescentes de taxas de tráfego foram geradas de forma a identificarmos em que ponto (valor de taxa) o descarte de células começa a ocorrer. A partir dessa informação, o limite inferior da ocupação média em células do buffer de *shaping* pode ser calculado.

Em nossa apresentação, mostraremos os detalhes da análise utilizada. Como exemplo, mostraremos a seguir os resultados da etapa 1 para as placas OC-3 e WAN 8260.

Análise para placa OC-3 nativa

Monitor_pvc2		RESUMO DAS MEDIÇÕES			
E (μs)	PCR (Kbps)	V (μs)	Perda (%)	T (atraso médio das PDU em ms)	RX_ME (taxa efetiva de recepção em kbps)
6360	1000	6370	0	1,55	997
3180	2000	3190	0	1,55	1993
1413	4500	1428	0	1,65	4452
1270	5000	1285	0	3,49	4954
1211	5250	1290	0	4,17	4940
1000	6360	1317	20	44,56	4966
636	10000	1245	48	44,70	4968

O tráfego UBR consegue excursionar até o valor máximo da banda alocada ao VP (5 Mbps), sem perdas. Quando a taxa efetiva gerada é superior a 5 Mbps, perdas ocorrem e o atraso T das PDUs elevado mostra que houve armazenamento devido ao *shaping*. Como a taxa efetiva de recepção máxima nunca ultrapassa 5 Mbps, podemos concluir que o *shaping* no VP está sendo realizado de modo correto, mesmo porque detectamos descarte ZERO no FORE. O atraso mínimo ida e volta medido é da ordem de 1,55 ms. A placa OC-3 nativa aceita a banda de 5 Mbps no VP e *shaping* com o mesmo valor.

Se analisarmos as duas últimas linhas quando foram observadas perdas na transmissão, o limite inferior da ocupação média em células do buffer de *shaping* pode ser calculado (resultado de Little vindo da teoria de processos estocásticos, que afirma que o número médio em um sistema é igual à taxa média vezes o tempo médio de permanência no sistema) multiplicando o atraso (T – 1,55) por Rx_me, taxa efetiva de recepção em kbps, e dividindo o resultado por 424 (número de bits por célula). Este cálculo nos fornece os valores de 504 e 493 células armazenadas. Este valor é um limite inferior porque o comutador trabalha com EPD e células podem ficar no buffer e serem descartadas posteriormente por este controle. Como a taxa efetiva de recepção só leva em consideração as PDUs recebidas integralmente, evidentemente a ocupação média do buffer deve ser algo maior que o indicado por este cálculo.

Para confirmar esta ocupação média, geramos um tráfego UBR de 13.144 kbps (E=1150, L=1472) e obtivemos as seguintes estatísticas: V=2642, T=46,07, P=60% e RX_ME=4968. O cálculo da ocupação média nos fornece 521 células. O acréscimo no valor de V se deve ao fator do tempo de transmissão de uma PDU com L=1472 bytes, implicando em 31 células, levar cerca de 2,628 ms a

¹ Foi verificado que a placa WAN trabalha com valores discretos de *shaping* para VP, com granularidade em torno de 512 Kbps. A tentativa de estabelecer um *shaping* de 5 Mbps no VP criado com 5 Mbps efetivamente provoca o estabelecimento do *shaping* no valor de 4.608 Kbps.

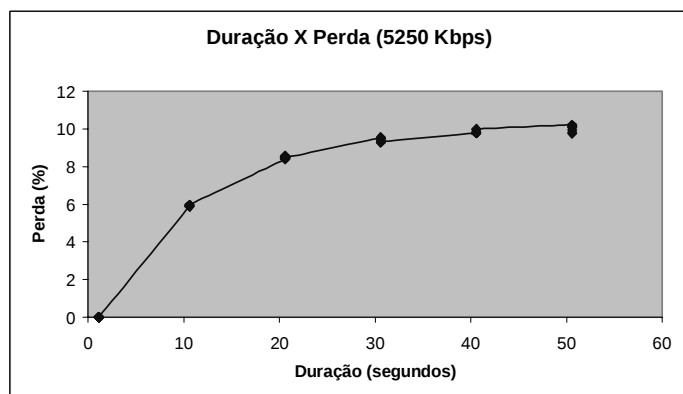
uma taxa de 5 Mbps. Este valor é maior que o E especificado. Neste exemplo específico, o programa monitor na estação foi chamado com PCR de 20000 Kbps, para garantir que estivéssemos gerando a taxa determinada pelos parâmetros L e E. Numa situação ainda mais extrema, geramos um tráfego UBR de 20.000 kbps (E=657, L=1472), disparando na estação com PCR=50.000 kbps. Obtivemos as seguintes estatísticas: V=2641, T=46,22 e RX_ME=4968. O cálculo da ocupação média nos fornece 523 células.

O número de buffers especificado pelo manual para a placa OC-3 é de 8000 (agregado para as 4 portas, possivelmente, contando entrada e saída, o que daria 1000 células por porta e sentido de transmissão). Nossa conclusão é que o número de células usadas no controle de shaping para o tráfego UBR na porta OC-3 é superior a 500. Não temos como verificar, no momento, o uso de um número próximo de 1000. Falta investigar se informações de gerência e da MIB podem gerar dados mais precisos da ocupação de buffer e possibilitar alterar esta alocação por VC e/ou VP.

Conclusões do teste 2.1 para placa WAN 8260

TEÓRICO		RESUMO DAS MEDIÇÕES			
E (μs)	PCR (kbps)	V (μs)	Perda (%)	T (atraso médio das PDU em ms)	RX_ME (taxa efetiva de recepção em kbps)
6360	1000	6363	0	1,82	999
3180	2000	3184	0	1,82	1997
1413	4500	1416	0	1,83	4490
1270	5000	1382	5,6	546	4606
1211	5250	1382	10,3	564	4606
1000	6360	1382	26	579	4606

O tráfego UBR consegue excursionar até o valor máximo da banda alocada ao VP 4,6 Mbps, sem perdas. Quando a taxa efetiva gerada é superior a 4,6 Mbps, perdas ocorrem e o atraso T das PDUs elevado mostra que houve armazenamento devido ao shaping. Como a taxa efetiva de recepção máxima nunca ultrapassa 4,6 Mbps, podemos concluir que o shaping na porta está sendo realizado de modo correto, mesmo porque detectamos descarte ZERO no FORE. O atraso mínimo de ida e volta medido é da ordem de 1,82 ms. A placa WAN só aceita shaping feito na porta. Ela não suporta shaping no VP. Os testes de shaping foram feitos à parte e comprovaram este comportamento. Interessante observar que a placa WAN 8260 possui um atraso superior ao da placa OC-3 nativa. Se analisarmos as três últimas linhas quando foram observadas perdas na transmissão, o número aproximado de células armazenadas no comutador (calculado multiplicando (T – 1,82) por Rx_me, taxa efetiva de recepção e dividindo o resultado por 424), nos fornece os valores de 5931, 6126 e 6289 células armazenadas. Este resultado é coerente com o manual da IBM que atribui buffers de 8000 células à porta da placa WAN.



Em surtos de tráfego de pequena duração acima da banda máxima do VP, a existência de buffers no controle de shaping permite que se opere com perdas baixas. Evidentemente, à medida que a duração do tráfego não conforme aumenta, a ocupação total dos buffers não impede o descarte crescente de células. A taxa de perda em % é assintótica, como mostramos no gráfico abaixo, para um tráfego de 5250 kbps num VP de 4608 kbps. Valor assintótico = $(5250 - 4608) / 5250 = 12\%$.

O atraso médio entre PDUs, V, tende a ficar constante, na medida em que o controle de shaping tem um buffer saturado e passa a gerar células com taxa constante. Eventualmente células são descartadas

por falta de buffers disponíveis na entrada. O IBM 8265 foi configurado com descarte de quadro habilitado.

Conclusões dos Testes

- Medida uma latência de comutação da ordem de 20 microssegundos entre portas OC-3 de fibra.
- VP só pode ser criado com sinalização habilitada na porta, e o tráfego reservado fica limitado a 85% da banda da porta (acordando com manual de configuração).
- Módulo de fibra OC-3 (nativo no IBM 8265) faz shaping e policiamento no VP e no VC.
- Os testes realizados nas portas OC-3 e UTP 25 Mbps indicam um buffer de shaping algo maior que 500 células, por VP. Este é um limite inferior, já que nas medidas EPD estava ligado, fazendo com que o tráfego descartado tivesse ocupação temporária nos buffers, que não é possível de ser computada. Esta informação não consta do manual de configuração.
- Módulo WAN 8260 faz shaping somente na porta, não permitindo o controle do tráfego por VC. Os testes realizados na porta WAN 8260 indicam um buffer de shaping algo maior que 6300 células. Este é um limite inferior, já que nas medidas EPD estava ligado, fazendo com que o tráfego descartado tivesse ocupação temporária nos buffers, que não é possível de ser computada. O manual indica um buffer de 8000 células para este tipo de porta.
- Módulo ATM UTP 25 Mbps não faz shaping por VC, apenas por VP.

Obs.: Na implantação do backbone RNP2 todos os tráfegos entrantes deverão sofrer preferencialmente shaping nas saídas dos próprios roteadores, já que não se consegue fazer shaping individualmente em VP na porta WAN.

Atividades Futuras

Ficaram pendentes os seguintes tópicos, e portanto sugestões para estudos futuros:

- Não foi possível configurar ABR, visto que não foi encontrado qualquer comando que habilitasse esse tipo de tráfego no PVC;
- comutador IBM não possui à primeira vista nenhuma forma de se observar descarte nos VCs individualmente, o que pode ser um problema para o futuro;
- Sinalização PNNI entre os comutadores está OK, embora o FORE possua parâmetros adicionais que afetam a configuração do domínio, e que precisam ser mais explorados;
- Não conseguimos alterar a configuração de buffers para testar a política de descarte de células. Se houver uma solução via SNMP (gerência), ela deveria ser explorada.

Reflexões

- Manuais de configuração podem não ser exatos ou fidedignos.
- Essencial estressar equipamentos em condições de teste controlado para comprovar características e levantar parâmetros.
- Testes são essenciais para comprovar efetividade de gerência, monitoração, configuração e adequação do equipamento às futuras condições de campo.
- Testes podem alterar substancialmente os procedimentos operacionais de uma rede.
- Reproduzir em laboratório as condições críticas de operação de uma rede é essencial para quantificar evolução e identificar condições de estrangulamento de performance.
- Nem sempre equipamentos são interoperáveis.