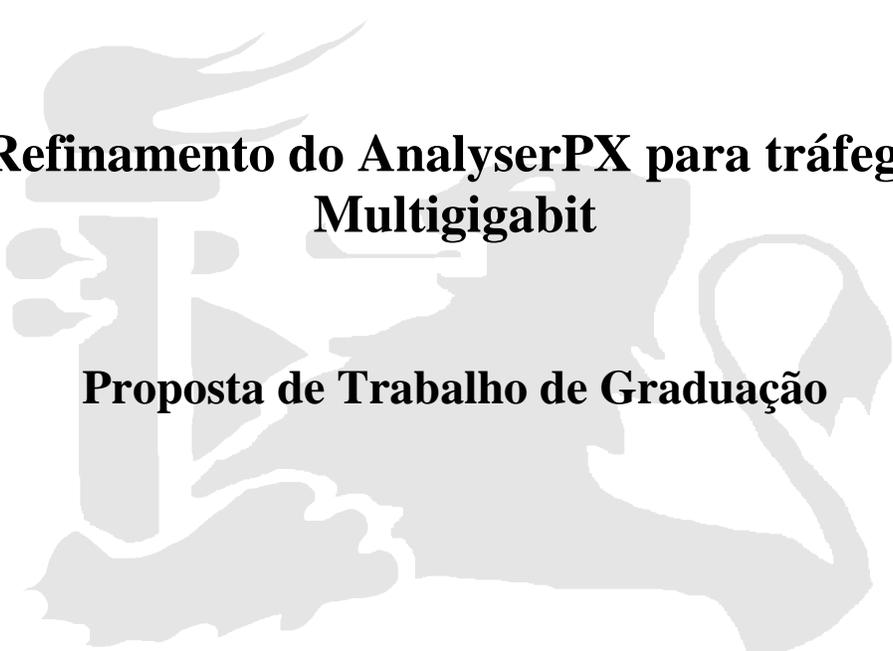


Universidade Federal de Pernambuco

GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

CENTRO DE INFORMÁTICA

2011.1



**Refinamento do AnalyserPX para tráfego
Multigigabit**

Proposta de Trabalho de Graduação

Aluno	Wesley Davison Braga Melo	{ wdbm@cin.ufpe.br}
Orientador	Djamel Fawzi Hadj Sadok	{ jamel@cin.ufpe.br}
Co-Orientador	Stênio Flávio de Lacerda Fernandes	{ sflf@cin.ufpe.br}

12 de Abril de 2011

Índice

ÍNDICE.....	2
1. RESUMO.....	3
2. CONTEXTO	3
3. ESTADO DA ARTE.....	4
4. OBJETIVOS	6
5. CRONOGRAMA.....	6
6. REFERÊNCIAS	7
7. POSSÍVEIS AVALIADORES.....	7
8. ASSINATURAS.....	8

1. Resumo

Este trabalho visa a extensão de um classificador de tráfego desenvolvido em pesquisa anterior, o AnalyserPX (1). Atualmente, este classificador atinge taxas de classificação próximas a 2.5 Gbps. Porém, para utilização em ISPs reais, faz-se necessário uma ferramenta capaz de classificar tráfego em tempo real à taxas multigigabits (10Gbps). Para atingir este propósito serão utilizadas técnicas do estado da arte em roteadores implementados em software unidas aos *hardwares* de alto desempenho disponíveis para computadores pessoais (PCs).

2. Contexto

Atualmente, o AnalyserPX é capaz de analisar tráfego a 2.5 Gbps. Para se alcançar os 10Gbps desejados, planejou-se uma arquitetura distribuída e hierárquica, composta por cinco máquinas. Quatro delas executariam o software atual de análise de tráfego, o AnalyserPX. Como cada uma suporta até 2.5 Gbps de tráfego de entrada sem perdas, temos com quatro máquinas o total de 10Gbps. Para fazer a distribuição de um tráfego vindo de um link de 10Gbps, utiliza-se a quinta máquina. Esta recebe 10Gbps e faz um balanceamento de carga do tráfego de entrada para as outras 4 máquinas.

O balanceamento de carga é realizado de forma manter a mesma distribuição de pacotes entre os módulos de análise. A máquina que realiza este trabalho será chamada de “Splitter”. Esta máquina é um tipo de roteador em software para plataformas x86 – x64 de computadores pessoais (PCs).

Abaixo podemos ver um esquema da arquitetura proposta para se alcançar os 10Gbps:

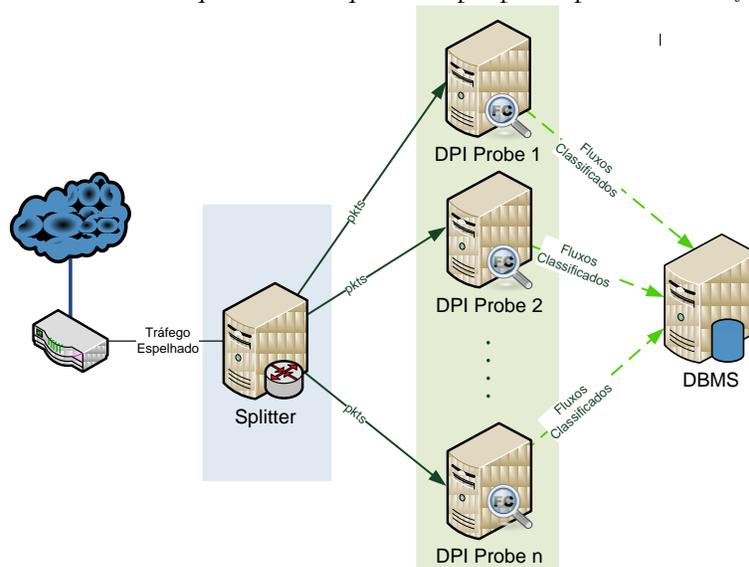


Fig 1: Diagrama da Arquitetura Distribuída

3. Estado da Arte

A fim de saber qual a viabilidade de se construir roteadores baseados em software e a que taxas estes poderiam ser utilizados, pesquisamos na literatura artigos que nos trouxessem informações relevantes para este trabalho. Abaixo trazemos um resumo sobre cada um dos artigos que fundamentaram este trabalho.

“**Can Software Routers Scale?**”, de Argyraki et Al (2). trata sobre a possibilidade de se construir roteadores baseados em software que alcancem altas velocidades de roteamento. Para isso, ele propõe uma arquitetura de roteadores baseada em cluster que permite distribuir a carga de processamento entre diversos PCs interconectados, em uma estrutura do tipo malha. Ele cita que os problemas dos roteadores são a inflexibilidade inerente àqueles que estão perto do núcleo da rede. Explica que os roteadores baseados em software construídos sobre um único PC alcançam taxas de roteamento entre 1 e 3 Gbps. Tais problemas ocorrem devido às custosas operações de processamento por pacote e o próprio chaveamento de um link de entrada para um de saída. Para explicar os problemas de processamento por pacote, eles descrevem as arquiteturas tradicionais sobre as quais os PCs são construídos, explicando os possíveis gargalos entre os barramentos de memória - CPU e PCI – CPU. Ainda mais sobre esse assunto, explica que o novo modelo de arquitetura NUMA (non-Uniform Memory Access) traz diversos benefícios a esses barramentos, eliminando os possíveis gargalos. Utilizando o modelo de Strawman, sugerem que as arquiteturas emergentes possibilitarão que esses tipos de roteadores consigam trabalhar a taxas de 40Gbps. Para verificar suas teorias, os autores realizam experimentos que apontam os gargalos previstos. Após isso, sugerem melhorias no sistema de manipulação dos pacotes dentro do sistema operacional, aconselham a serem utilizados mecanismos de acesso direto ao cache (DCA) das CPUs e arquiteturas de multi-processadores conectados em malha. Eles também propõem que um roteador em software se utilize de diversos PCs, formando uma rede de triagem. Dessa forma, um pacote recebido por uma interface de entrada chegaria em outra interface de saída passando por alguns PCs dentro da malha. Para o cálculo do caminho, ele cita algoritmos como o Valiant. Com esse sistema, ele mostra que é possível tornar roteadores desse tipo escaláveis com a quantidade de interfaces de entrada.

“**Towards Hardware Accelerated Software Routers**”, de Sarrar et Al (3). traz uma abordagem de roteadores baseados em software que mistura PCs tradicionais com Switches programáveis. Os autores citam os problemas existentes em roteadores baseados apenas em PCs e sugere que as tarefas de um roteador sejam desacopladas. O PC funciona como controlador de rotas, e o Switch programável como o encaminhador dos pacotes processados pelo PC. Os autores percebem que a maioria do tráfego da Internet possui uma previsibilidade, o que permite deixar algumas informações de rota diretamente no cache do switch. Com isso, apenas os pacotes que não tiverem suas informações de rota no switch serão processados pelo PC. Tal fato diminui os problemas de processamento por pacote, aumentando a vazão do sistema.

“**Scalable Layer-2/Layer-3 Multistage Switching Architectures for Software Routers**”, de Bianco et Al (4). propõe uma separação das funções de roteamento, sendo este feito em 2 passos. Os autores explicam a arquitetura interna dos PCs e os possíveis gargalos existentes no uso deles como roteadores em software. Destacam os problemas existentes na pilha de protocolos do Linux. Eles ainda dizem que as principais limitações de um roteador em software são o gargalo de processamento por pacote (para pacotes pequenos) e o barramento PCI (para pacotes grandes). Para superar os problemas previstos, eles sugerem que as

capacidades de roteamento da camada 2 (Ethernet) e da camada 3 (IP) sejam separadas em duas fases, com vários PCs. Na primeira, os pacotes são recebidos por PCs denominados de balanceadores. Estes apenas recebem os pacotes e os redirecionam para outros PCs, através de switches tradicionais. Na segunda fase, outros PCs denominados “roteadores” coletam os pacotes enviados pelos PCs da primeira fase, realizam o processamento por pacotes e os reenviam para os PCs da primeira fase. Estes redirecionam para as interfaces de saída. O objetivo da primeira fase é realizar um balanceamento uniforme dos pacotes de entrada entre os PCs roteadores da segunda fase. Tal mecanismo, segundo os autores, permite escalabilidade e tolerância a falhas. Para a operação de balanceamento de carga, eles mencionam 3 algoritmos possíveis para esta tarefa. Fazem avaliações de performance e mostram que é possível se chegar a uma taxa de 1Gbps, utilizando 4 PCs.

“**Towards High Performance Virtual Routers on Commodity Hardware**”, de Egi et Al.(5) traz as limitações de performance que existem na construção de roteadores virtualizados e os princípios de projeto que devem ser adotados para sua construção. Eles citam que uma das desvantagens de não virtualizar os roteadores é que servidores virtuais que rodem sobre uma arquitetura de nuvem não podem gerenciar os switches por onde trafegam os dados. Falam que as velocidades de roteamento possíveis são aceitáveis, se comparado com as soluções empresariais presentes no mercado atual. Baseado em uma configuração comum a um PC de médio porte, os autores mostram os problemas de roteamento dos pacotes utilizando um único núcleo de processamento e também com vários núcleos. Eles identificam as limitações, que estão na latência de memória, para sistemas de arquitetura uniforme de memória, e afinidade de threads e núcleos de processamento, para sistemas multi-thread com arquitetura não-uniforme de memória. Exploram a virtualização mostrando dois cenários: virtualização a nível de software (com o OpenVZ) e a nível de hardware (com o Xen). Eles citam 3 cenários de possíveis configurações de roteadores virtualizados. Para cada um dos casos, eles dizem qual é a melhor solução (se OpenVZ ou Xen). Ainda analisam especificamente as interfaces de rede e descrevem 3 opções de compartilhamento entre os roteadores virtualizados.

4. Objetivos

Como metas do projeto, temos os seguintes pontos:

- Estender o Analyser-PX e propor uma nova arquitetura de *software* para o classificador levando em consideração a modularidade da arquitetura e a facilidade de extensão em termos de funcionalidades e velocidade. Também pode ser necessária a criação de módulos do classificador dentro do *kernel* do sistema operacional visando obter o melhor desempenho possível;
- Implementar o classificador proposto em uma linguagem de alto desempenho e de fácil interação com o sistema operacional;
- Implementar o módulo “Splitter” de forma a suportar a escalabilidade do sistema proposto.

5. Cronograma

O cronograma do trabalho é apresentado na tabela logo abaixo. Foram definidas algumas tarefas genéricas de forma que estas tarefas serão divididas em subtarefas para uma melhor organização. Como qualquer atividade, este trabalho é suscetível a imprevistos e, portanto, mudanças nas datas destas tarefas.

ATIVIDADES	Abril				Maio				Junho				
Levantamento dos Requisitos Necessários Para a Extensão do Analyser PX	█	█	█	█									
Especificação da Nova Versão do Analyser PX				█	█	█							
Implementação da Nova Versão do Analyser PX						█	█	█					
Testes de Validação								█	█				
Elaboração do monografia						█	█	█	█				
Elaboração da apresentação											█	█	

6. Referências

1. *Análise de Tráfego Peer-to-Peer baseada na Carga Útil dos Pacotes*. **G. Silvestre, C. A. Kamienski, S. Fernandes, D. M. Souza, D. H. Sadok**. Curitiba, 2006. I Workshop de Peer-to-Peer (P2P).
2. *Can Software Routers Scale?* **K. Argyraki, S. Baset, B. Chun, K. Fall, G. Iannaccone, A. Knies, E. Kohler, M. Manesh, S. Nedveschi, S. Ratnasamy**. 2008. Proceedings of the ACM workshop on Programmable routers for extensible services of tomorrow .
3. *Towards Hardware Accelerated Software Routers*. **N. Sarrar, A. Feldmann, R. Sherwood, X. Huang**. 2010. Proceedings of the ACM CoNEXT Student Workshop.
4. *Scalable Layer-2/Layer-3 Multistage Switching Architectures for Software Routers*. **A. Bianco, J. M. Finochietto, G. Galante, M. Mellia, D. Mazzucchi, F. Neri**. 2009. Proceedings of the First international conference on COMmunication Systems And NETworks .
5. *Towards High Performance Virtual Routers on Commodity Hardware*. **N. Egi, A. Greenhalgh, M. Handley, M. Hoerd, F. huici, L. Mathy**. Madrid , 2008. Proceedings of the 2008 ACM CoNEXT Conference.

7. Possíveis Avaliadores

1. Judith Kelner (jk@cin.ufpe.br)

8. Assinaturas

Djamel Fawzi Hadj Sadok
Orientador

Wesley Davison Braga Melo
Aluno