

Proposta de GT da RNP

1 Título do Projeto

GT-ATER (Aceleração do Transporte de Dados com o Emprego de Redes de Circuitos Dinâmicos)

2 Coordenador

Nome: Kleber Vieira Cardoso

Instituição: Instituto de Informática (INF) - Universidade Federal de Goiás (UFG)

Currículo Lattes: <http://lattes.cnpq.br/0268732896111424> Dados de contato:

3 Resumo

Em NRENs (National Research and Education Networks) como APAN, Internet2, ESnet, GÉANT, CLARA e RNP, é crescente a necessidade de transportar grandes volumes de dados de maneira confiável e eficiente. Por outro lado, é bastante conhecida a dificuldade do TCP em acompanhar a ampliação da capacidade de transmissão da rede, sobretudo devido aos altos valores do produto largura de bandaatraso. Uma das abordagens para melhorar o desempenho do TCP em redes de alta capacidade é a criação de circuitos dinâmicos. Os circuitos dinâmicos podem ser usados como atalhos dentro da rede, onde o produto largura de banda-atraso é reduzido pela remoção de parte dos atrasos de enfileiramento. Esta proposta tem como objetivo desenvolver e implantar um serviço para identificar fluxos de pacotes que transportam grande volume de dados e criar circuitos dinâmicos dedicados para enviá-los. Através desse serviço, os usuários da RNP que precisam transportar grandes volumes de dados de maneira confiável poderão fazê-lo em menos tempo.

4 Abstract

In National Research and Education Networks (NRENs) such as APAN, Internet2, ESnet, GÉANT, CLARA e RNP, there is a growing need to transport large volumes of data reliably and efficiently. On the other hand, it is recognized the difficulty of TCP to follow the network capacity extension, especially due to the high values of bandwidth-delay product. One approach to improve the TCP performance in high-capacity networks is the use of dynamic circuits. Dynamic circuits can be used as shortcuts within the network, reducing the bandwidth-delay product through removing part of the queue delays. This proposal aims to develop and deploy a service to identify packet flows which transport large volume of data and send them through dedicated dynamic circuits. This service will provide to RNP users a reduction in the time spent on reliable transport of large data volumes.

5 Parcerias

Esta proposta tem com parceiros o professor José Ferreira de Rezende (COPPE/UFRJ) e o pesquisador Antônio Tadeu Azevedo Gomes (MARTIN/LNCC).

O professor José Ferreira de Rezende atua na área de Redes de Computadores desde 1985. Obteve o título de Doutor em Informática pela Université Paris VI (Université Pierre et Marie-Curie) em 1997. Ingressou como professor da COPPE/UFRJ em 1998. Já orientou 35 teses de mestrado e doutorado e tem mais de 150 artigos publicados em conferências e periódicos nacionais e internacionais. Participou dos seguintes projetos de P&D nas áreas de redes fixas e móveis: FuturaRNP (RNP), GT-Travel (RNP), InterTV (MCT/FINEP/CPqD

SBTVDigital RFP14), Taquara (Projeto Giga/RNP), GigaBot (Projeto Giga/RNP), VIMOS (CT-INFO) e Quaresma (PROTEM-RNP). Nesses projetos, coordenou diversas atividades de pesquisa. Teve ainda os seguintes projetos de pesquisa individuais aprovados: CoVER (CNPq Edital Universal 2009), ECCRAM (CNPq Edital Universal 2006), UnFADiNG (bolsista do CNPq Nível 1D), RADHIOS (CNPq Edital Universal 2004), QoSmo (bolsista do CNPq Nível 2), QoS na Internet (CNPq Edital Universal 2001), e Redes ATM (FAPERJ Instalação 1999).

Antônio Tadeu A. Gomes possui graduação em Informática - Magna Cum-Laude - pela UFRJ (1995), mestrado e doutorado em Informática pela PUC-Rio (1999 e 2005). Atualmente é pesquisador do LNCC/MCTI (Laboratório Nacional de Computação Científica), onde coordena, juntamente com o pesquisador Artur Ziviani, o grupo de pesquisa MARTIN (Mecanismos e Arquiteturas em Teleinformática - <http://martin.lncc.br>). Também é secretário executivo do SINAPAD (Sistema Nacional de Processamento de Alto Desempenho - <http://www.lncc.br/sinapad>). Coordena o Grupo de Trabalho *mc²* da RNP. Leciona na Pós-graduação em Modelagem Computacional do LNCC. Atua na área de Ciência da Computação, principalmente nos seguintes temas: Arquitetura e Modelagem de Software, Redes de Computadores e Computação de Alto Desempenho. É membro da Sociedade Brasileira de Computação (SBC) e da Association for Computing Machinery (ACM).

6 Duração do Projeto

12 meses

7 Sumário Executivo

Esta seção apresenta a motivação do projeto, as características do serviço proposto, a viabilidade técnica e o escopo do trabalho. Ao longo desta seção, serão apresentados também os detalhes do protótipo a ser demonstrado ao final dos 12 meses.

7.1 Motivação

Microscópios digitais, telescópios, aceleradores de partículas, simulações meteorológicas e imagens médicas são apenas alguns exemplos de equipamentos ou aplicações que geram grandes volumes de dados. Além disso, à medida que os equipamentos e aplicações evoluem, aumenta a quantidade de informação gerada. Para ilustrar, o Tevatron do Fermilab gerou aproximadamente 20 PB de dados ao longo de 26 anos de operação, enquanto seu sucessor, o LHC do CERN, deve gerar em torno de 15 PB por ano. Para atender essa demanda sempre crescente de grande volume de dados, a comunidade de redes tem desenvolvido várias soluções em todas as camadas da rede, desde a física até a de aplicação. Como há um interesse em manter a compatibilidade com as tecnologias utilizadas na Internet, o uso da pilha TCP/IP é mandatório. Nesse contexto, é notória a dificuldade em evoluir o TCP para aproveitar de maneira efetiva a capacidade crescente das redes de comunicação, mantendo o protocolo compatível com as tecnologias de rede tradicionais que oferecem taxas de transmissão muito inferiores.

A literatura de redes é farta em trabalhos que apresentam diferentes mecanismos de controle de congestionamento para o TCP que buscam utilizar de maneira mais eficiente a largura de banda disponível de redes de alta capacidade, mas que funcionam adequadamente em redes convencionais [1, 2, 3, 4]. Em geral, o TCP não recebe informações explícitas sobre o nível de ocupação da rede, logo os mecanismos de controle de congestionamento precisam inferir a situação da rede através dos reconhecimentos (ACKs) que retornam do receptor. Em redes com grande memória, ou seja, que possuem um alto produto largura de banda-atraso (Bandwidth-Delay Product ou BDP), há uma grande quantidade de ACKs em trânsito e, portanto, é mais complexo estimar a ocupação da rede. Por essa razão, se torna difícil ajustar o tamanho da janela de transmissão do TCP para o valor ideal que permita usufruir da capacidade total disponível e também reagir a eventos de perda de pacotes. Nesse contexto, o que se observa é que os mecanismos de controle de congestionamento disponíveis para o TCP demoram para ocupar toda a capacidade disponível e são severamente afetados por perda de pacotes, ainda que a taxa de perda seja baixa. É importante lembrar que as perdas em redes cabeadas,

sobretudo em redes óticas, são geralmente causadas por transbordo de buffer nos roteadores. A principal causa dos transbordos é o tráfego em rajada, comumente gerado pelos mecanismos de controle de congestionamento do TCP. Em roteadores de backbone, onde é comum haver uma grande quantidade de fluxos TCP compartilhando a fila [5], é necessário utilizar buffers que sejam grandes o suficiente para amenizar o efeito das rajadas, mesmo nas abordagens mais otimistas [6, 7]. Naturalmente, quanto maiores são os buffers dos roteadores, mais elevada é a memória da rede, ou seja, maior é o BDP.

Além de mecanismos de controle de congestionamento especializados em rede com alto BDP, outras abordagens têm sido empregadas para tratar as limitações do transporte confiável de dados, tais como o UDT, conexões TCP em paralelo e logística em rede. O UDT (UDP-based Data Transfer) [8] combina dois mecanismos: 1) controle de janela, que limita o número de pacotes não reconhecidos ainda em trânsito, e 2) controle de taxa, que configura o intervalo entre pacotes. Os parâmetros dos dois mecanismos são ajustados dinamicamente por uma técnica de estimação de banda baseada em pares de pacotes. O UDT ganhou notoriedade ao vencer três vezes a competição de transferência de grande volume de dados chamada Supercomputing Bandwidth Challenge. O UDT utiliza algoritmos complexos e é implementado na camada de aplicação, por essa razão exige mais capacidade computacional que o TCP. No entanto, a principal desvantagem do UDT é a sua inadequação a enlaces que possuam grande número de fluxos concorrentes [9]. Em enlaces de backbones de NRENs, como a RNP, é comum ter dezenas ou centenas de milhares de fluxos concorrendo pelos recursos de cada roteador.

O uso de conexões TCP em paralelo [10, 11] é uma abordagem antiga, mas bastante efetiva. Com múltiplas conexões TCP, uma transferência passa a ter múltiplas janelas de transmissão independentes que conseguem ocupar mais rapidamente a capacidade disponível na rede, assim como reagir a perda de pacotes de maneira mais eficiente. No entanto, essa abordagem beneficia quem a utiliza em detrimento dos demais, pois consiste em competir de maneira desigual pelos recursos da rede [12, 13]. Em enlaces compartilhados, o uso de conexões TCP em paralelo deve ser limitado.

A logística em rede [14] consiste em utilizar equipamentos intermediários ao longo de um caminho fim-a-fim e segmentar uma conexão TCP em múltiplas conexões em série passando por esses equipamentos. Assim, uma rede com alto BDP pode ser tratada como múltiplas redes com valores menores de BDP. Cada conexão passa a operar em um cenário no qual o TCP apresenta desempenho adequado com relação à ocupação da rede e reação a perda de pacotes. Uma solução baseada em logística em rede já foi implementada na RNP a partir do GT-Travel [15], apresentando resultados satisfatórios [16]. A principal dificuldade dessa abordagem é a adesão ao serviço, uma vez que pelo menos um dos lados da conexão TCP precisa ser modificado, mesmo que minimamente, para utilizar a solução.

Todas as abordagens comentadas têm em comum o fato de lidar com a restrição apresentada pela rede, ou seja, minimizar os efeitos dos altos valores de BDP. A partir do momento em que a RNP oferece um serviço de redes de circuitos dinâmicos, surge a oportunidade de tratar o problema em sua origem. Através de um circuito dinâmico é possível remover, ainda que apenas logicamente, roteadores intermediários. Isso significa reduzir o atraso de enfileiramento nesses equipamentos, sendo que esse atraso tem comumente peso significativo no BDP. Além disso, a perda por transbordo de buffer passa a praticamente não existir ao longo do circuito dinâmico. O uso de circuitos dinâmicos pode ainda ser combinado com soluções como UDT e conexões TCP em paralelo, uma vez que o tráfego que flui em circuito pode ser isolado do restante da rede. O uso de circuitos dinâmicos permite implementar um serviço que acelere o transporte confiável de dados sem alterar as aplicações que se encontram nas bordas da rede e com pouca, ou até nenhuma, participação dos usuários beneficiados. Essa é a proposta do GT-ATER que será detalhada nas seções que seguem.

7.2 Características do Serviço

Atualmente, a RNP oferece, através do serviço MEICAN, uma maneira dos usuários solicitarem a criação de circuitos sob demanda. No MEICAN, o usuário acessa uma interface Web na qual indica os equipamentos de origem e de destino do circuito, assim como os horários em que deseja o estabelecimento desse circuito. Esse serviço atende uma ampla gama de necessidades, porém alguns cenários importantes não são contemplados. Eventualmente, o usuário pode não saber o horário em que uma determinada aplicação precisa realizar a

transferência de dados. Por exemplo, o mecanismo de sincronização de grandes storages pode ser ativado de maneira não determinística. Além disso, o MEICAN estabelece um circuito entre duas redes, ficando a cargo do usuário definir o que deve fluir através do circuito dinâmico. A princípio, todo o tráfego entre as duas redes pode ser enviado pelo circuito, mas o usuário pode ter interesse em encaminhar apenas fluxos específicos. Essa filtragem deve ser implementada pelo usuário.

A proposta do GT-ATER é oferecer também um serviço que permita aos usuários da RNP solicitarem circuitos dinâmicos, porém com enfoque em uma demanda diferente, observando as questões apresentadas anteriormente. Assim, o usuário acessará uma interface Web na qual informará o perfil do tráfego que deseja enviar por um circuito dinâmico. Caso queira, o usuário poderá indicar o destino do circuito, mas essa informação é opcional, uma vez que poderá ser extraída dos próprios pacotes. O perfil do tráfego será descrito através da escolha de regras de filtragem, ou seja, o usuário escolhe um protocolo (ex.: TCP) e indica as eventuais informações extras (ex.: porta). Não é necessário indicar quando o circuito deve ser ativado, pois a ativação ocorrerá sob demanda, quando o tráfego começar a fluir. A duração de um circuito será em fatias (slots) fixas de tempo, cujo tamanho poderá ser indicado pelo usuário ou definido pelo administrador. Quando uma fatia de tempo for encerrada e ainda existir tráfego a ser enviado, o serviço poderá solicitar um novo circuito por mais uma fatia de tempo. O serviço proposto será capaz de encaminhar o tráfego pela rede convencional sempre que não conseguir reservar um novo circuito, de maneira transparente e sem perda de dados. Quando o usuário não desejar mais utilizar uma regra, pode simplesmente removê-la. A regra garante que apenas o perfil do tráfego definido será transportado através do circuito. Como o esquema de regras é flexível, o usuário pode escolher enviar desde todo o tráfego entre duas redes até um fluxo bastante específico.

Como há uma grande quantidade de aplicações que utilizam o HTTP, será oferecida uma opção para ativar a criação de circuitos dinâmicos a partir da inspeção do cabeçalho desse protocolo. Inicialmente, o intuito é verificar apenas o campo que indica o tamanho do arquivo a ser transportado (Content-Length) e permitir a escolha de um limiar. Assim, o usuário poderá também indicar a partir de qual tamanho de arquivo, enviado através de HTTP, deseja a ativação de circuitos dinâmicos. Nesse modo de operação, o serviço poderá estimar a duração da transferência, solicitando a reserva do circuito dinâmico pelo tempo correspondente.

A autorização para que uma regra seja efetivamente ativada deverá ser fornecida pela equipe de operação da RNP, semelhante ao que é realizado no serviço MEICAN. Assim, um administrador acessará a interface Web oferecida pelo serviço proposto, porém com privilégios especiais. A interface oferecida ao administrador o permitirá autorizar regras solicitadas e também criar regras. A criação de regras pelo administrador permite que o serviço seja mais transparente para o usuário final, pois seu tráfego pode ser transportado por um circuito dinâmico sem que ele tome conhecimento.

O administrador poderá ativar uma regra ou apenas solicitar seu monitoramento. No modo de monitoramento, a regra não pode dar origem a um circuito dinâmico, mas todo tráfego que casar com a regra será contabilizado. Assim, após um tempo determinado pelo administrador, ele poderá verificar a quantidade de fluxos, número de pacotes por fluxo e número de pacotes totais, duração de cada fluxo e duração média, vazão por fluxo e vazão média. Essas informações podem ser usadas para subsidiar a decisão de ativar ou não a regra solicitada. Essas informações também estarão disponíveis quando a regra for ativada, podendo ser utilizadas para decidir pela continuidade ou remoção da regra. Além disso, caso observe algum indício de tráfego que envolva grande volume de dados, o administrador pode criar regras de monitoramento e usar as informações para antecipar a demanda dos usuários. Por fim, o administrador pode também mudar o nível de um usuário para gerente local, permitindo que suas regras sejam ativadas sem a necessidade de intervenção do administrador.

Uma evolução do serviço proposto prevê a melhoria de dois itens. O primeiro seria o suporte a um conjunto de regras ainda mais flexíveis que permitiriam aos usuários (ou administradores) descreverem suas necessidades através de uma pequena linguagem de propósito específico. O segundo seria o uso de um esquema de permissões mais elaborado, no qual usuários e regras poderiam ser associados pelo administrador com o intuito de aumentar a flexibilidade do controle de criação de regras.

Caso a proposta do GT-ATER venha a ser aprovada, a RNP passará a contar com informações sobre a efetiva demanda de tráfego de alto volume de dados. Essas informações poderão ser usadas para justificar a implantação de soluções mais complexas e de alto custo como a que já teve início na Internet2, chamada

Science DMZ. Basicamente, a Science DMZ consiste em um perímetro de rede que utiliza equipamentos, configurações e políticas de segurança otimizadas para aplicações científicas de alto desempenho [17]. Assim, a demanda especializada por alto desempenho passa a ser separada da demanda convencional, podendo contar com soluções customizadas como sistemas dedicados para transferência de dados e rede definida por software.

7.3 Solução Proposta

O serviço proposto pelo GT-ATER fará uso de dois tipos de equipamentos: um servidor e um intermediário. No servidor, será hospedado o núcleo do serviço, ou seja: o controle de acesso dos usuários e dos administradores, a solicitação para criação de circuitos dinâmicos, o armazenamento do resumo das estatísticas de tráfego, a comunicação com o serviço de circuitos dinâmicos da RNP e o controle remoto dos equipamentos intermediários. Nos equipamentos intermediários, estarão as regras e a configuração para estabelecer a conectividade entre as extremidades dos circuitos dinâmicos. A Figura 1 ilustra os componentes do serviço que serão detalhados a seguir.

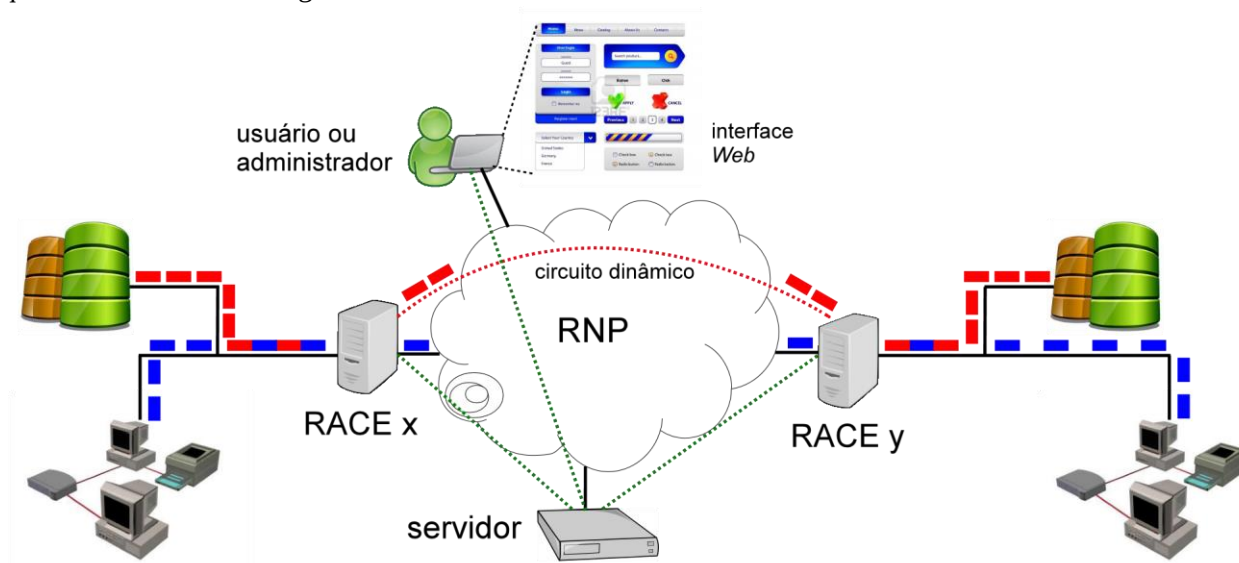


Figura 1: Arquitetura do serviço proposto pelo GT-ATER.

Nomeamos os equipamentos intermediários com base em um resumo de suas atribuições: RACE (Rule Applier and Circuit Endpoint). Um RACE pode ser implementado em um computador convencional com algumas interfaces de rede. Inicialmente, duas interfaces do equipamento são colocadas em modo ponte (bridge), fazendo com que o encaminhamento seja realizado na camada de enlace. Essa abordagem ajuda a tornar um RACE transparente aos demais equipamentos no mesmo segmento de rede. Normalmente, os RACEs trabalharão aos pares, sendo coordenados pelo servidor para formar as extremidades dos circuitos dinâmicos. Quando um circuito dinâmico se torna disponível, o RACE de origem inicia o processo de reencaminhamento do tráfego de acordo com a regra que recebeu do servidor. O reencaminhamento pode ser realizado logicamente, acrescentando a etiqueta (tag) de VLAN que identifica o circuito, ou fisicamente, através de envio por outra interface de rede, conforme ilustrado na Figura 1. Em ambas as soluções, o endereço MAC dos pacotes também será alterado para refletir o RACE que está na outra extremidade do circuito. O RACE de destino realiza a tarefa complementar de remover a etiqueta de VLAN adicionada pelo serviço ou simplesmente encaminhar o pacote para o destino final. Novamente, é necessário substituir o endereço MAC para o valor adequado. Assim, os RACEs funcionam como comutadores de nível 2 (switch) que escondem os circuitos dinâmicos das redes que estão nas bordas.

Cada RACE mantém uma conexão segura com o servidor do GT-ATER a partir da qual recebe toda a configuração necessária e pela qual informa sobre o tráfego que casa com alguma regra, envia estatísticas de uso e relatórios de operação. O canal seguro também é usado para monitorar a disponibilidade de todos os componentes do serviço. Caso algum RACE não consiga estabelecer a comunicação segura com o servidor

após algum tempo, todas as regras são desativadas e o equipamento passa a funcionar com um comutador (switch) simples, reencaminhando de maneira convencional todos os pacotes. Do lado do servidor, a ausência de algum RACE significa que novos circuitos não podem ser iniciados ou terminados naquela localidade. Após a aprovação de uma regra, é tarefa do servidor estabelecer um circuito dinâmico assim que é informado sobre o surgimento de tráfego que atenda à referida regra. Para isso, o servidor entra em contato com o OSCARS (On-demand Secure Circuits and Advance Reservation System) e solicita a criação do circuito. O intuito será usar uma abordagem semelhante a utilizada pelo serviço MEICAN, o qual possui uma conta dentro do sistema OSCARS com privilégio para gerenciar reservas, atualizar e visualizar topologias.

7.4 Viabilidade Técnica

O serviço proposto pelo GT-ATER se baseia no Serviço Experimental CIPÓ (SE-CIPÓ), o qual está em operação e teste em mais de uma dezena de PoPs da RNP. Recentemente, o coordenador desta proposta e a equipe do PoP-GO solicitaram que o serviço fosse implantado também no ponto de presença do Estado de Goiás. Os responsáveis pelo SE-CIPÓ já autorizaram essa implantação e os procedimentos para a efetivação do processo estão em andamento.

O coordenador desta proposta tem experiência na implantação de serviços e desenvolvimento de ferramentas e sistemas, tendo participado inclusive das duas fases do GT-Travel. Além disso, o coordenador conta com o apoio de seu grupo de pesquisa formado por pesquisadores da área de sistemas distribuídos, estudantes de pós-graduação e graduação. A equipe tem proficiência em diferentes linguagens de programação e fundamentos sólidos de redes de computadores, sistemas operacionais e sistemas distribuídos.

Por fim, é importante destacar o apoio das parcerias que incluem uma equipe da UFRJ, liderada pelo prof. José Rezende, e outra equipe do LNCC, liderada pelo pesquisador Antônio Tadeu. A equipe do prof. José Rezende foi uma das que trabalhou na concepção e implantação do SE-CIPÓ e terá papel fundamental nas decisões de projeto do GT-ATER. A equipe do pesquisador Antônio Tadeu tem grande experiência em computação de alto desempenho e possui interesse em utilizar o serviço proposto pelo GT-ATER tanto no SINAPAD quanto no GT-*mc*².

7.5 Escopo do Trabalho

Esta proposta abrange o desenvolvimento do software necessário para criação de um serviço para aceleração do transporte confiável de dados através do uso de redes de circuitos dinâmicos. O serviço proposto será baseado no SE-CIPÓ, oferecendo aos usuários uma interface simples e o maior nível de transparência possível.

O software a ser desenvolvido possui duas partes principais, a que é executada no servidor e a que é executada nos RACEs. A parte do software presente no servidor terá como principais componentes: Web services, OpenSSL e banco de dados MySQL. A parte do software presente nos RACEs terá como principais componentes: OpenSSL, Linux bridge, ebttables e código em raw sockets.

Todo o software desenvolvido neste projeto será devidamente documentado e disponibilizado para uso da RNP e de seus usuários.

8 Recursos Financeiros

Os recursos necessários à realização do projeto compreendem equipamentos, viagens e pagamento de pessoal.

8.1 Equipamentos

Para desenvolvimento e testes do protótipo:

- 1 switch gerenciável de 24 portas com slots de expansão para interface ótica – switch para testes no laboratório e interligação direta em fibra ótica do laboratório para a rede da UFG: R\$ 4.000,00
- 4 interfaces GBIC – conexão através de fibra ótica do laboratório para a rede da UFG: R\$ 4.000,00

- 1 servidor de rack – equipamento para implantação e testes do protótipo do núcleo do serviço: R\$ 8.200,00
- 3 computadores com 3 interfaces de rede cada um – serão utilizados para atuar como RACEs nos testes em laboratório: R\$ 9.000,00

Total: R\$ 25.200,00

De fato, o protótipo será constituído por mais equipamentos que os listados anteriormente, sendo os adicionais oferecidos como contrapartida dos pesquisadores e de suas instituições.

8.2 Viagens

Estão previstas viagens entre Goiânia, Petrópolis e Rio de Janeiro com o objetivo de realizar reuniões do projeto, num total de 5 viagens (com 2 diárias cada uma) orçadas em R\$ 1.500,00 cada.

Total: R\$ 7.500,00

8.4 Resumo do Orçamento

Equipamentos para desenvolvimento - R\$ 25.200,00

Viagens Nacionais (5 * R\$ 1.500,00) - R\$ 7.500,00

Total - R\$ 162.496,08

9 Ambiente para Testes do Protótipo

Inicialmente, o protótipo do serviço será desenvolvido em um laboratório de pesquisa do Instituto de Informática da UFG. Posteriormente, serão realizados testes entre os PoPs de Goiás e do Rio de Janeiro.

Todo o processo de prototipagem será realizado sobre a plataforma GNU/Linux.

Referências

- [1] D. Katabi, M. Handley, and C. Rohrs, “Congestion control for high bandwidth-delay product networks,” *ACM SIGCOMM Computer Communication Review*, vol. 32, no. 4, pp. 89–102, 2002.
- [2] L. Xu, K. Harfoush, and I. Rhee, “Binary increase congestion control (BIC) for fast long-distance networks,” in *INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 4, pp. 2514–2524, mar 2004.
- [3] C. Jin, D. Wei, and S. Low, “FAST TCP: motivation, architecture, algorithms, performance,” in *INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 4, pp. 2490–2501, mar 2004.
- [4] S. Ha, I. Rhee, and L. Xu, “CUBIC: a new TCP-friendly high-speed TCP variant,” *ACM SIGOPS Operating System Review*, vol. 42, pp. 64–74, jul 2008.
- [5] D. Wischik and N. McKeown, “Part I: buffer sizes for core routers,” *SIGCOMM Computer Communication Review*, vol. 35, pp. 75–78, jul 2005.
- [6] G. Raina, D. Towsley, and D. Wischik, “Part II: control theory for buffer sizing,” *SIGCOMM Computer Communication Review*, vol. 35, pp. 79–82, jul 2005.
- [7] M. Enachescu, Y. Ganjali, A. Goel, N. McKeown, and T. Roughgarden, “Part III: routers with very small buffers,” *SIGCOMM Computer Communication Review*, vol. 35, pp. 83–90, jul 2005.

- [8] Y. Gu and R. L. Grossman, "UDT: UDP-based data transfer for high-speed wide area networks," *Computer Networks*, vol. 51, pp. 1777–1799, may 2007.
- [9] Y. Zhu, A. Bassi, P. Massonet, and D. Talia, "Mechanisms for High Volume Data Transfer in Grids," tech. rep., Institute on Knowledge and Data Management, 2007.
- [10] H. Sivakumar, S. Bailey, and R. L. Grossman, "PSockets: The Case for Application-level Network Striping for Data Intensive Applications using High Speed Wide Area Networks," in *ACM/IEEE Conference Supercomputing*, p. 38, nov. 2000.
- [11] T. Hacker, B. Athey, and B. Noble, "The end-to-end performance effects of parallel TCP sockets on a lossy wide-area network," in *International Parallel and Distributed Processing Symposium*, pp. 434–443, 2002.
- [12] T. J. Hacker, B. D. Noble, and B. D. Athey, "Improving Throughput and Maintaining Fairness using Parallel TCP," in *IEEE International Conference on Computer Communications (INFOCOM)*, vol. 4, pp. 2480–2489, mar 2004.
- [13] B. Tuffin and P. Maillé, "How Many Parallel TCP Sessions to Open: A Pricing Perspective," in *International Workshop on Internet Charging and QoS Technology (ICQT)*, pp. 2–12, 2006.
- [14] D. M. Swamy and R. Wolski, "Data Logistics in Network Computing: The Logistical Session Layer," in *IEEE International Symposium on Network Computing and Applications*, vol. 2, pp. 174–185, oct 2001.
- [15] RNP, "GT-Travel 2 - Transporte em Alta Velocidade." http://www.rnp.br/pd/gts2008-2009/gt_travel2.html, 2009. [Último acesso: 05-Setembro-2012].
- [16] C. H. P. Augusto, M. W. R. da Silva, K. V. Cardoso, A. C. Mendes, R. M. Guedes, and J. F. de Rezende, "Implementação e Avaliação de Encadeamento de Conexões TCP em Redes de Grande Memória e Altas Perdas," in *Simpósio Brasileiro de Redes de Computadores (SBRC)*, pp. 89–102, 2009.
- [17] ESnet, "Science DMZ – A Scalable Network Design Model for Optimizing Science Data Transfers." <http://fasterdata.es.net/science-dmz/>, 2009. [Último acesso: 05-Setembro-2012].