

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
COMPUTAÇÃO**

André Barros de Sales

**MEDIDAS DE LATÊNCIA EM AMBIENTES DE
PROCESSAMENTO DE ALTO DESEMPENHO**

Dissertação submetida à Universidade Federal de Santa Catarina como parte dos requisitos para a obtenção do grau de Mestre em Ciência da Computação.

Orientador: Prof. Dr. Carlos Becker Westphall

Florianópolis, julho de 2000.

MEDIDAS DE LATÊNCIA EM AMBIENTES DE PROCESSAMENTO DE ALTO DESEMPENHO

André Barros de Sales

Esta Dissertação foi julgada adequada para a obtenção do título de Mestre em Ciência da Computação, na Área de Concentração de Sistemas de Computação e aprovada em sua forma final pelo Programa de Pós-Graduação em Ciência da Computação.



Prof. Dr. Carlos Becker Westphall

Orientador



Fernando A. Ostuni Gauthier

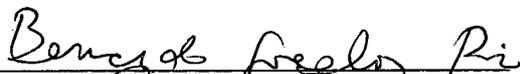
Coordenador do Curso

Banca Examinadora



Prof. Dr. Carlos Becker Westphall

Presidente



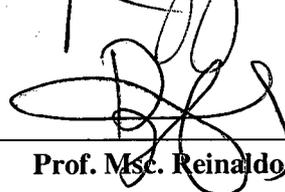
Prof. Dr. Bernardo Gonçalves Riso



Prof. Dra. Mirela S. M. Annoni Notari



Prof. Dr. Ricardo Felipe Custódio



Prof. Msc. Reinaldo Haas

AGRADECIMENTOS

A Deus.

A minha família
pelo apoio e incentivo pela busca de novos caminhos.

A família Aguiar Velho
pelo carinho com que me receberam em sua família.

Aos professores e aos colegas do Laboratório de Redes e Gerência (LRG),
principalmente, à Profa. Mirela e ao amigo Abiel,
que ao longo do curso contribuíram para a realização deste trabalho.

Aos colegas do Núcleo de Processamento de Dados,
Solange, Melo, Vilson, Kátia, Gentil, Jaime, Gustavo, Walter, Hélio e Richard
pelos momentos passados juntos e por tudo que aprendi de cada um.

Aos professores integrantes da banca examinadora.

Aos amigos Andreлина, Bel, Cláudio, Flávia, Marlene e Saulo.

Ao Rick Jones da Hewlett-Packard Company
pela ajuda e contribuição científica.

A Valdete e Vera da secretaria do CPGCC da UFSC.

A Companhia de Processamento de Dados de Goiânia – Comdata
e aos amigos da Comdata.

A todas as outras pessoas não citadas aqui que contribuíram
direta ou indiretamente para o desenvolvimento desse trabalho.

Em especial, ao meu orientador, Prof. Dr. Carlos Becker Westphall,
pela orientação, apoio, estímulo e amizade durante a realização deste trabalho.

SUMÁRIO

LISTA DE FIGURAS	vi
LISTA DE QUADROS	viii
LISTA DE SIGLAS	ix
RESUMO	x
ABSTRACT.....	xi
1 INTRODUÇÃO	12
1.1 MOTIVAÇÃO DO TRABALHO	13
1.2 OBJETIVO	14
1.3 ESTRUTURA DO TRABALHO.....	15
2 PROCESSAMENTO DE ALTO DESEMPENHO E AVALIAÇÃO DE SISTEMAS COMPUTACIONAIS	16
2.1 PROCESSAMENTO DE ALTO DESEMPENHO – PAD	16
2.1.1 Formas de Obtenção do Alto Desempenho.....	16
2.1.2 Principais Tipos de Paradigmas de Computação Paralela	22
2.2 AVALIAÇÃO DE SISTEMAS COMPUTACIONAIS	24
2.2.1 Técnicas de Avaliação	24
2.2.2 Métricas de Performance	25
3 MATERIAIS E MÉTODOS	29
3.1 AMBIENTE DE AVALIAÇÃO	29
3.2 MÉTODO DE MEDIÇÃO DE LATÊNCIA EM AMBIENTES COMPUTACIO- NAIS	31
3.2.1 Ferramentas de Medição	32
3.2.2 Netperf	34
3.3 COLETA DE DADOS DENTRO DO CLUSTER	34
3.3.1 Coleta de dados das estações	36
4 AVALIAÇÃO DOS RESULTADOS	39
4.1 CLUSTER AIX	40
4.1.1 Cluster Aix utilizando o Protocolo TCP	40
4.1.2 Cluster Aix utilizando o Protocolo UDP	42
4.2 CLUSTER LINUX	45

4.2.1 Cluster Linux utilizando TCP.....	47
4.3 SP 2	49
4.3.1 SP 2 utilizando protocolo TCP.....	49
4.3.2 SP 2 utilizando protocolo UDP.....	51
4.4 COMPARAÇÃO DOS RESULTADOS DAS LATÊNCIAS.....	53
5 CONCLUSÃO	55
5.1 TRABALHOS FUTUROS.....	57
REFERÊNCIAS BIBLIOGRÁFICAS	59
ANEXO 1.....	65
ANEXO 2.....	75

LISTA DE FIGURAS

Figura 2.1- Arquitetura Von Neumann.	18
Figura 2.2 - Processadores matriciais organizados em malha bidimensional.....	19
Figura 2.3 - Modelo MIMD fracamente Acoplado.....	21
Figura 2.4 - Modelo de MIMD com forte acoplamento.....	22
Figura 2.5 - Construção de um muro sobre o paradigma de paralelismo de dados.	23
Figura 2.6 - Construção de um muro sobre o paradigma do paralelismo funcional.	23
Figura 2.7 - Construção de um muro sobre o paradigma do paralelismo pipelining.	24
Figura 2.8 - Vazão x eficiência.	26
Figura 2.9 - Latência x eficiência.....	27
Figura 3.1 - Topologia da rede Cluster UFSC.	30
Figura 3.2 - Formato do arquivo do Script.....	35
Figura 3.3 - Formato do arquivo do Script para medição da latência da estação.	37
Figura 4.1 - Neste diagrama são mostrados os valores da latência (ms) por tamanho de segmento TCP (bytes) para o Cluster Aix. As barras verticais correspondentes de latência máxima, mínima para cada segmento, as dentro das caixas maiores concentram-se de 25% até 75% dos valores e a caixa menor indica o valor da mediana. Note-se que a tendência exponencial apresentada pelos valores da latência em função do tamanho do segmento é devidos ao fato de que o tamanho segmento não seguem uma relação linear.....	40
Figura 4.2 - Gráfico dos valores de latência (ms) pelo tamanho do segmento TCP (bytes) para o Cluster Aix. A linha reta representa a regressão linear efetuada. Os círculos representam os valores amostrais. Note-se que a reta ajusta-se bem aos dados.....	42
Figura 4.3 - Neste diagrama são mostrados os valores de latência (ms) pelo tamanho do segmento UDP (bytes) para o Cluster Aix. Idem a figura 4.1.	43
Figura 4.4 - Gráfico da latência (ms) pelo segmento UDP (bytes) no Cluster Aix e a equação de regressão linear. Idem Figura 4.2.	44
Figura 4.5 - Neste diagrama são mostrados os valores da latência (ms) por tamanho de segmento TCP (bytes) para o Cluster Linux – Note-se o	

comportamento irregular da latência média para segmento de 1.518 bytes.....	45
Figura 4.6 - Neste diagrama são mostrados os valores da latência (ms) por tamanho de segmento TCP (bytes) para a latência de Loop Back na estação Recope02 – Idem a figura 4.5.....	46
Figura 4.7 - Neste diagrama são mostrados os valores da latência (ms) pelo tamanho do Segmento TCP (bytes) para o Cluster Linux. Idem a figura 4.1.....	47
Figura 4.8 - Gráfico da latência (ms) pelo segmento TCP (bytes) no Cluster Linux e a equação de regressão linear. Idem Figura 4.2.....	48
Figura 4.9 - Neste diagrama são mostrados os valores da latência (ms) pelo tamanho do Segmento TCP (bytes) para o SP 2. Idem a figura 4.1.....	49
Figura 4.10 - Gráfico dos valores de latência (ms) pelo tamanho do segmento TCP (bytes) para o SP 2. A linha reta representa a regressão linear efetuada. Os círculos representam os valores amostrais. Note-se que existe uma grande dispersão dos dados em relação a reta.....	50
Figura 4.11 - Neste diagrama são mostrados os valores da latência (ms) pelo tamanho do Segmento UDP (bytes) para SP 2. Idem a figura 4.1.....	51
Figura 4.12 - Gráfico da latência (ms) pelo segmento UDP (bytes) no Cluster Linux e a equação de regressão linear. Idem Figura 4.10.....	52

LISTA DE QUADROS

Quadro 2.1 - Seqüência de Instruções X Seqüência de Dados.....	17
Quadro 2.2 - Critérios de seleção da técnica de Avaliação de Desempenho.	25
Quadro 3.1 - Estações/Clusters utilizados na Avaliação.....	31
Quadro 3.2 - Redes indicadas para utilização do Netperf.	33
Quadro 3.3 - Instalação da Ferramenta para avaliação da latência dentro do cluster.....	34

LISTA DE SIGLAS

ATM	- Modo de transmissão Assíncrono
CISC	- <i>Complex Instruction Set Computers</i>
CPU	- Unidade Central de Processamento
FDDI	- <i>Fiber Distribution Data Interface</i>
GigaFLOPS	- Bilhões de instruções de ponto flutuante por segundo (1 Giga = 1.000.000.000)
ICMP	- <i>Internet Control Message Protocol</i>
IETF	- <i>Internet Engineering Task Force</i>
HPC	- <i>High Performance Computing</i>
HPS	- <i>High Performance Switch</i>
MIMD	- Múltiplas Instruções com Dados Múltiplos
MISD	- Múltiplas Instruções com Dados Simples
MPI	- <i>Message Passing Interface</i>
MTU	- Unidade Máxima de Transmissão
PAD	- Processamento de Alto Desempenho
PVM	- <i>Parallel Virtual Machine</i>
RFC	- <i>Request for Comments</i>
RISC	- <i>Reduced Instruction Set Computers</i>
SISD	- Simples Instruções com Dados Simples
SIMD	- Simples Instruções com Dados Múltiplos
TCP	- <i>Transmission Control Protocol</i>
TeraFLOPS	- Trilhões de instruções de ponto flutuante por segundo (1 Tera = 1.000.000.000.000)
UDP	- <i>User Datagram Protocol</i>

RESUMO

Este trabalho compara alguns ambientes de Processamento de Alto Desempenho - PAD, considerando a latência (tempo em que um pacote trafega de um ponto para outro, fim-a-fim, na rede) como o critério para avaliar o desempenho desses ambientes. Este trabalho, cujo título é “Medidas de Latência em Ambientes de Processamento de Alto Desempenho”, apresenta 3 agrupamentos (*clusters*) da UFSC nos quais são realizadas as medidas de latência. Esses *clusters* são dispostos da seguinte maneira: um com sistema operacional Linux interligado com uma rede Fast-Ethernet e os outros dois com sistema operacional AIX, um interligado por uma rede ATM e o outro interligado por uma rede Ethernet. O método para realizar as medidas nesses ambientes é baseado, principalmente, na “Metodologia de *Benchmarking* para equipamentos de Interconexão” do *Request for Comments* - RFC: 2544, que é definido pelo *Internet Engineering Task Force* – IETF. A ferramenta Netperf, da *Hewlett-Packard Company*, é o programa para medir a latência nesses ambientes; os motivos da seleção, as características e o manuseio do Netperf também são descritos neste trabalho. Após definidos o objetivo, o critério, os ambientes, o método e a ferramenta é realizada a medição nesses *clusters* e são apresentadas as análises das medidas de latência. Como resultado, é realizada a comparação das latências desses *clusters*. Um outro resultado é a definição de expressões matemáticas para cada *cluster*, relacionando a variabilidade da latência com o tamanho do segmento e apresentando o seu respectivo coeficiente de determinação. Como conclusão, este trabalho mostra que o *cluster* com o sistema operacional Linux interligado por uma rede Fast-Ethernet possui a menor latência, seguido pelo *cluster* com o sistema operacional AIX interligados por uma rede ATM, e pelo SP 2, com o sistema operacional AIX interligado por uma rede Ethernet. Uma outra conclusão deste trabalho é que, para ambientes que não possuem colisão, a variabilidade da latência pode ser explicada pelo tamanho do segmento com um alto coeficiente de determinação, no entanto, para as redes que possuem colisão, a variabilidade da latência também dependerá da taxa de colisão.

ABSTRACT

This work compares some High Performance Computing environments – HPC, considering latency (the time needed for a package to go from one point to another, end-to-end, in a network) as the criterion for evaluating the performance of those environments. This work, titled “Latency Measurement in High Performance Computing Environments”, presents 3 groupings (clusters) from the UFSC in which the latency measurements are done. Those clusters are disposed in the following manner: one with the Linux operating system interconnected by a Fast-Ethernet network and the two others with an AIX operating system, one interconnected by an ATM network and the other interconnected by an Ethernet network. The method for carrying out measurements in those environments is mainly based on the “Benchmarking Methodology for Interconnection equipment” from Request for Comments – RFC: 2544, which is defined by the Internet Engineering Task Force – IETF. The Netperf tool, from *Hewlett-Packard Company*, is the program for measuring the latency in those environments; the reasons for choosing, the characteristics and the manipulation of Netperf are also described in this work. After defining the aim, the criterion, the environments, the method and the tool, the measurement is carried out on those clusters and the latency measurement analyses are presented. As a result, the latency comparison for those clusters is carried out. Another result is the definition of mathematical expressions for each cluster, establishing a relationship between the latency variability and the segment size and presenting its respective determination coefficient. In conclusion, this work shows that the Linux operating system cluster interconnected by a Fast Ethernet network has the smallest latency, followed by the AIX operating system cluster interconnected by an ATM network, and then by the SP 2, with the AIX operating system interconnected by an Ethernet network. Another conclusion of this work is that, for collision-free environments, the latency variability can be explained by the segment size with a high determination coefficient. However, for networks that have collision, latency variability will depend on the collision rate as well.

1 INTRODUÇÃO

Na área científica as simulações computacionais têm sido utilizadas como alternativa aos experimentos, os quais são, muitas vezes, dispendiosos ou até impossíveis de serem realizados.

Na maioria dos casos, as simulações exigem a execução de algoritmos complexos bem como envolvem grande volume de dados, os quais necessitam de elevada capacidade computacional; isso tem contribuído para um rápido crescimento do Processamento de Alto Desempenho (PAD) nos últimos anos. Além disso, várias áreas dos setores comercial e industrial mundiais estão utilizando aplicações de PAD como um instrumento para o ganho de competitividade.

O Processamento de Alto Desempenho é composto por um conjunto de *hardware*, *software* e técnicas de processamento cujo principal objetivo é reduzir ao máximo o tempo de execução das aplicações.

Existe um custo elevado e algumas restrições nas tecnologias para obtenção do PAD em um único processador, envolvendo pesados sistemas de refrigeração e consumo de energia elétrica. Uma das alternativas é a utilização da computação paralela, utilizando um conjunto de processadores do CISC e RISC, os quais possuem custo reduzido e não requerem pesados sistemas de refrigeração. Neste tipos de PAD, os processadores são interconectados por uma rede que permite a cooperação e o compartilhamento de dados entre os mesmo.

Para as redes de computadores um dos fatores mais importante de desempenho é a latência, que é definida como o tempo em que um pacote trafega de um ponto para outro, fim-a-fim, na rede [SHU99].

A avaliação de desempenho de sistemas computacionais pode ser feita por meio de 3 técnicas: simulação, modelo analítico e medida [JAI91]. Este trabalho avalia a latência pela técnica de medida sobre os ambientes PAD interconectados por uma rede.

No Brasil, a Financiadora de Estudos e Projetos (FINEP) criou o Sistema Nacional de Processamento de Alto Desempenho (SINAPAD), baseado em universidades, para dar acesso e difundir o uso de PAD, visando ao desenvolvimento econômico do País.

A Universidade Federal de Santa Catarina (UFSC) tem assumido um papel importante e de liderança no incentivo, definição, estudo e pesquisas de soluções, formação, geração de propostas e apoio à implementação nas áreas de PAD.

Na UFSC, o estudo do PAD e de suas aplicações é realizado pelo Núcleo de Redes de Alta Velocidade e Computação de Alto Desempenho – NURCAD. Dentro de suas atividades o NURCAD coordena o projeto de Rede Metropolitana de Alta Velocidade de Florianópolis – RMAV-FLN.

O projeto RMAV-FLN faz parte dos projetos RMAVs nacional, criado pela Rede Nacional de Pesquisa (RNP) e pelo Programa Temático Multiinstitucional em Ciência da Computação (ProTeM-CC), tendo como objetivo promover, em diversas regiões do País, a criação de infra-estrutura e serviços de redes de alta velocidade.

Este trabalho faz parte do projeto RMAV-FLN e são apresentados os ambientes de PAD da UFSC, avaliando a latência. As atividades desta pesquisa estão sendo realizadas pelo Laboratório de Redes e Gerência do Departamento de Informática e Estatística em cooperação com o Laboratório de Interoperabilidade do Núcleo de Processamento de Dados, ambos da UFSC.

Alguns trabalhos de pesquisas relacionados em medida de latência podem ser encontrados em [BER96], [BRA99], [LEO00], [WEL96] e [WEL97].

A principal diferença desta pesquisa em relação às citadas acima está no tipo de latência analisado. Neste trabalho, é analisada a latência fim-a-fim e não apenas a latência da rede ou do equipamento de interconexão. Além disso, a variabilidade da latência é determinada em relação ao tamanho do segmento.

1.1 MOTIVAÇÃO DO TRABALHO

Cada vez são mais comuns as aplicações que precisam de alta velocidade de execução como, por exemplo, o mapeamento do Genoma do ser humano. Nessa aplicação, para se realizar o mapeamento desejado, foram utilizados cerca de 700 processadores. Os cálculos complexos foram “quebrados” em operações mais simples e executados por vários processadores, em diferentes computadores, isto é, foi utilizado um ambiente de Processamento de Alto Desempenho- PAD [MAD00].

Os ambientes de PAD na UFSC têm sido utilizados para a realização de várias pesquisas, tais como:

- impacto de eventos extremos do clima em escala global na região Sul do Brasil, pelo Centro Integrado de Meteorologia e Recursos Hídricos de Santa Catarina - CLIMERH - SC;
- cálculo numérico de autovalores de matrizes esparsas aplicado ao problema de estabilidade de sistemas elétricos de potência, pelo Departamento de Matemática - CFM – UFSC;
- análise de desempenho do esquema de linearização de Newton na discretização das equações do movimento com diagramas de Voronoi, pelo Laboratório de Computação Algébrica e Simbólica (LabCAS) do Departamento de Informática e de Estatística – CTC – UFSC;
- reconfiguração em tempo real de redes de distribuição de energia com o auxílio de redes neurais artificiais, pelo Laboratório de Planejamento de Sistemas de Energia Elétrica (Labplan) do Departamento de Engenharia Elétrica – UFSC; etc.

Algumas aplicações PAD possuem comunicação, em determinados instantes, entre os processadores. O número de comunicações dependerá de cada aplicação. A comunicação entre os processadores das estações se efetua através da rede para as aplicações que utilizam ambientes PAD interconectados por essa rede.

Para essas aplicações, o tempo total de execução consiste do tempo computacional somado ao tempo de comunicação através da rede [JOH94].

São fatores como esses colocados acima que motivaram a pesquisa dos ambientes de PAD e, especificamente, a latência.

1.2 OBJETIVO

O objetivo desta dissertação é avaliar o desempenho da latência de ambientes computacionais de Processamento de Alto Desempenho por meio da técnica de medida. Com isso, é possível determinar qual é o ambiente que obtém a menor latência para a

utilização das aplicações de PAD, ou mesmo para planejar a construção de novos ambientes de PAD.

Para essa avaliação, devem ser alcançados os seguintes objetivos específicos:

- apresentar os principais conceitos relacionados com a avaliação de sistemas computacionais e de Processamento de Alto Desempenho;
- avaliar diferentes agrupamentos (*clusters*) de processamento de alto desempenho quanto aos seus valores de latência;
- relacionar o parâmetro latência com o tamanho dos *segmentos* (*Transmission Control Protocol* - TCP e *User Datagram Protocol* - UDP) usando expressões matemáticas, para que o valor da latência possa ser usado pelas outras técnicas para avaliação de desempenho de sistemas computacionais; e
- definir quais são as melhores combinações de características (sistema operacional, rede, etc.) a serem escolhidas para se obter a menor latência nas redes de Processamento de Alto Desempenho.

1.3 ESTRUTURA DO TRABALHO

O trabalho está organizado em 5 (cinco) capítulos:

- Capítulo 1: Introdução, onde são apresentados a motivação do trabalho e os objetivos a serem atingidos com a realização do mesmo;
- Capítulo 2: Processamento de Alto Desempenho e Avaliação de Sistemas Computacionais, onde é fornecido o embasamento teórico sobre Processamento de Alto Desempenho e Avaliação de Sistemas Computacionais;
- Capítulo 3: Materiais e Métodos, apresenta o ambiente de avaliação, a ferramenta utilizada e o método usado para realizar a medição;
- Capítulo 4: Avaliação dos Resultados, onde são apresentadas as análises feitas sobre os dados da técnica de medida, mostrando os resultados obtidos e a discussão dos mesmos.

Em seguida são apresentadas a Conclusão e as Referências Bibliográficas.

2 PROCESSAMENTO DE ALTO DESEMPENHO E AVALIAÇÃO DE SISTEMAS COMPUTACIONAIS

Este capítulo descreve no item 2.1, o Processamento de Alto Desempenho e, no item 2.2, a Avaliação de Sistemas Computacionais.

2.1 PROCESSAMENTO DE ALTO DESEMPENHO - PAD

Considerando a capacidade da tecnologia atual, o processamento de alto desempenho pode se apresentar como sendo qualquer sistema computacional com uma capacidade de operações aritméticas por unidade de tempo maior que 1 Gigaflops (um bilhão de operações de ponto flutuante por segundo) [MAR98], [CHO90] ou 1 Teraflops [HAL99], [MAC97]. Outras referências apresentam a definição de PAD como sendo qualquer computador que tenha uma ordem de magnitude de performance muito maior do que as estações de trabalhos normais disponíveis.

As principais aplicações do PAD têm sido nas áreas científica e de engenharia, como:

- áreas que necessitam de um avançado visualizador de imagens, Sistemas de Informação Geográfica (GIS), previsão do tempo a longo prazo etc.;
- realidade virtual, trabalho e lazer;
- banco de dados, para aplicações que precisam processar rapidamente grande quantidade de dados, como, por exemplo, dados de vídeo; e
- mecânica quântica, etc. [CAS97], [MAR98].

2.1.1 FORMAS DE OBTENÇÃO DO ALTO DESEMPENHO

Para os computadores seqüenciais convencionais obterem um maior desempenho, é necessária a utilização de um sofisticado projeto de *chips*. A tecnologia necessária para a melhoria de um único processador torna-se cada vez mais cara, uma vez que os limites de velocidade de processamento estão sendo alcançados [CAS97].

As restrições impostas pela tecnologia e o custo elevado para a obtenção do alto desempenho em computadores com um único processador, conduzem a uma alternativa que é a utilização da computação paralela. Os modos para processar dados com múltiplos processadores em paralelo são:

- executar o mesmo programa em todos os processadores; ou
- executar diferentes programas que rodam em cada processador.

As principais características para realizar um processamento de alto desempenho devem contemplar os seguintes requisitos [MAR95]:

- processadores, podendo cooperar concorrentemente;
- memória, podendo ser local ou compartilhada pelos processadores; e
- rede, interconectada entre os processadores, permitindo o compartilhamento de dados e sincronização.

Taxonomia das Arquiteturas do Processamento de Alto Desempenho

Os sistemas de computação paralela classificam-se dentro da seqüência de instruções e da seqüência de dados a serem processados. [REA99] estabelece que as arquiteturas de computadores podem ser agrupados em quatro categorias, conforme Quadro 2.1:

Quadro 2.1 - Seqüência de Instruções X Seqüência de Dados.

		Seqüência de instruções	
		Simple Instruções (SI)	Múltiplas Instruções (MI)
Seqüência de Dados	Dados Simples (SD)	SISD (Von Neuman)	MISD (Pipeline)
	Dados Múltiplos (MD)	SIMD (vetores, matrizes)	MIMD (múltiplos computadores)

Conforme [MAR95] e [MAR98] serão apresentados com mais detalhes os conceitos para cada uma dessas categorias.

Categoria Simples Instruções com Dados Simples - SISD

A categoria SISD é implementada pela tradicional arquitetura Von Neuman. Essa arquitetura computacional foi desenvolvida em 1945 pelo matemático John Von Neumann. Ela é composta por 5 partes: equipamentos de entrada de dados, equipamentos de saída de dados, unidade de memória (UM), unidade aritmética (UA) e unidade de controle (UC), que está representada na Figura 2.1 por [MAC97].

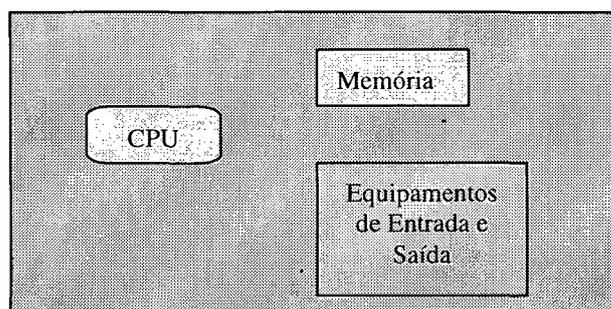


Figura 2.1- Arquitetura Von Neumann.

A única diferença entre as primeiras máquinas, conforme a Figura 2.1, e as atuais é o acoplamento da unidade aritmética com a unidade de controle, o que resulta na conhecida Unidade Central de Processamento - CPU, podendo possuir arquiteturas de seus processadores *Reduced Instruction Set Computers* - RISC ou *Complex Instruction Set Computers* - CISC [VIA98]. Exemplo de máquinas que implementam essa arquitetura [HAL99]: Pentium, Macintosh, RS6000 etc.

Categoria Simples Instruções com Dados Múltiplos - SIMD

Uma característica dessa categoria é que possui muitos e simples processadores. Todos os processadores possuem um módulo de memória local e próprio, e executam simultaneamente a mesma tarefa com diferentes fluxos de dados, conforme é apresentado na Figura 2.2.

O SIMD será o padrão no projeto da nova geração de computadores. A padronização para um sistema de computador consiste em entradas/saídas, unidade de controle, n processadores (cada um com sua memória), e a interconexão com a rede. Todos os processadores ativos executam as mesmas instruções no mesmo tempo, por isso é simples instrução. Como cada processador ativo executa a instrução em seus dados no seu módulo de memória, por isso são múltiplos dados [CHO90].

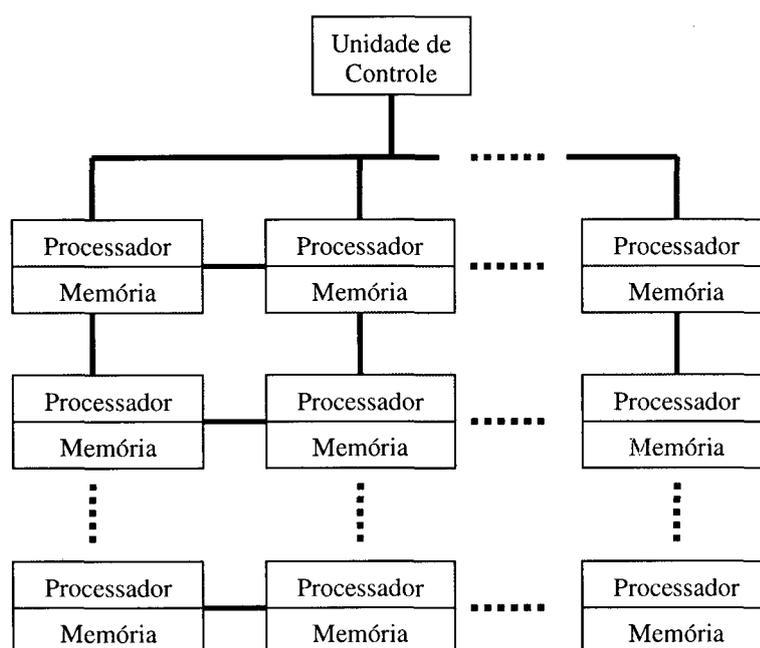


Figura 2.2 - Processadores matriciais organizados em malha bidimensional.

É apresentado na Figura 2.2 um conjunto de processadores matriciais organizados em uma malha bidimensional; uma característica importante dessa técnica é a execução de um processo síncrono (*lock step*). A sincronização entre os nós é mantida por uma unidade de controle que determina o que cada processador irá executar. A eficiência depende do número de dados que podem ser processados em paralelo e, por esse motivo, o modelo SIMD é conhecido na literatura como programação de dados paralela [CAS97].

Categoria Múltiplas Instruções com Dados Múltiplos - MIMD

Esta categoria caracteriza-se por possuir duas ou mais CPUs interligadas, trabalhando em conjunto que:

- pode possuir diferentes processadores (Alpha, SPARC, PowerPC etc.);
- pode possuir uma arquitetura síncrona ou não para a execução das tarefas;
- cada processador pode executar tarefas que podem ser diferentes; e
- dados que podem ser diferentes em cada processador.

Não existe uma sincronização global, o que requer processadores autônomos como se fossem computadores distintos. Esta categoria pode ser subdivida em duas novas diferentes categorias MIMD [MAR98], [MAC97]:

- com fraco acoplamento;
- com forte acoplamento.

A categoria MIMD com fraco acoplamento:

O fraco acoplamento caracteriza-se pela utilização de duas ou mais estações de trabalho e/ou computadores pessoais, conectados através de uma rede local de comunicação, conforme a Figura 2.3, para solucionar problemas com elevados requisitos computacionais, possuindo as seguintes características [CAS97], [MAC97]:

- possui seu próprio processador, memória e equipamentos de entrada e saída;
- a comunicação entre os processadores é feita pela troca de dados, utilizando bibliotecas de troca de mensagens, como por exemplo, MPI (*Message Passing Interface*), PVM (*Parallel Virtual Machine*) etc.; e
- máquinas fisicamente localizadas em ambientes distantes e diferentes.

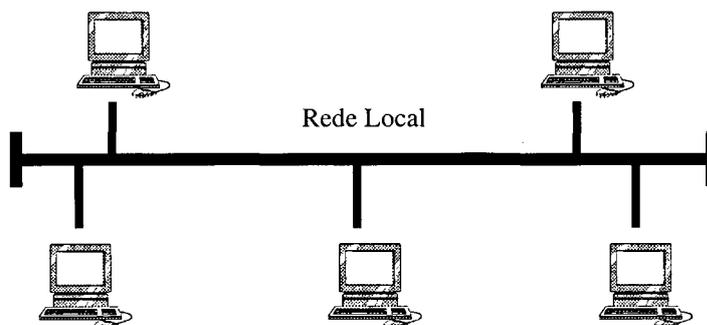


Figura 2.3 - Modelo MIMD fracamente Acoplado.

A rede local para essa categoria geralmente é uma Ethernet (10 Mbps), ou *Fiber Distribution Data Interface* - FDDI (100 Mbps), ou Modo de transmissão Assíncrono ATM (155 - 622 Mbps). Alguns exemplos de equipamentos que possuem essa configuração são Meiko CS-2, IBM SP-2, Cray T3D etc. [MAR95].

A Figura 2.3 apresenta um exemplo de estações de trabalho que poderiam ser estações RISC/6000 43P e interligados por uma rede local, podendo ser um barramento Ethernet.

Este trabalho analisa a latência sobre ambientes nessa categoria MIMD com fraco acoplamento.

A categoria MIMD com forte acoplamento:

As principais características da categoria com forte acoplamento são:

- a memória é compartilhada pelos processadores; e
- o sistema operacional e o *hardware* são projetados para garantir que os múltiplos processadores não armazenem simultaneamente e erroneamente no mesmo endereço na memória, comprometendo o conteúdo da memória.

Conforme ilustra a Figura 2.4 pode não ser possível aumentar indefinidamente o número de processadores, uma vez que os caminhos de acesso à memória podem formar um gargalo no sistema.

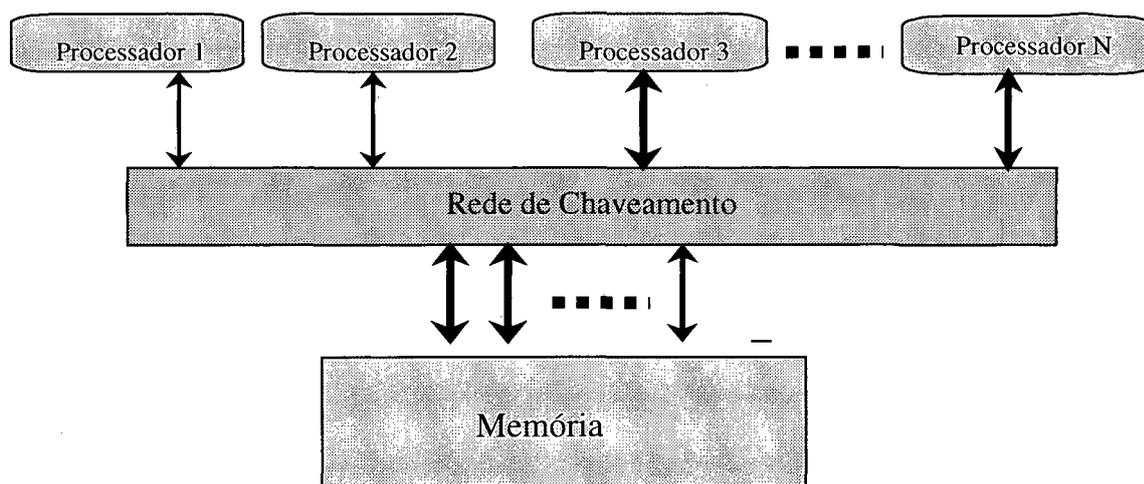


Figura 2.4 - Modelo de MIMD com forte acoplamento.

A Figura 2.4 [CAS97] apresenta a categoria MIMD com forte acoplamento, enquadrando-se nos sistemas computacionais com memória compartilhada, onde existem dois ou mais processadores (multiprocessadores) que compartilham uma única memória e são controlados por apenas um único sistema operacional [CAS97].

Alguns exemplos dessa categoria são SG PowerChallenge, Sun SPARCserver 1000, Cray 90 etc. E a rede de comunicação costuma ser de velocidade entre 200 - 400 Mbps [CAS97].

2.1.2 PRINCIPAIS TIPOS DE PARADIGMAS DE COMPUTAÇÃO PARALELA

Paradigma pode ser entendido como um modelo a ser usado, abstraindo detalhes e facilitando o entendimento do problema real. Os principais paradigmas na computação paralela são [CAS97]:

- i. paralelismo de dados;
- ii. paralelismo funcional; e
- iii. paralelismo *pipelining*.

Será usado um exemplo simples para elucidar esses paradigmas como em [CAS97], [MAR98].

(i) Paralelismo de dados

Nesse paralelismo é realizada uma decomposição dos dados em partições e posterior distribuição entre os processadores. A Figura 2.5 apresenta um exemplo da construção de um muro.

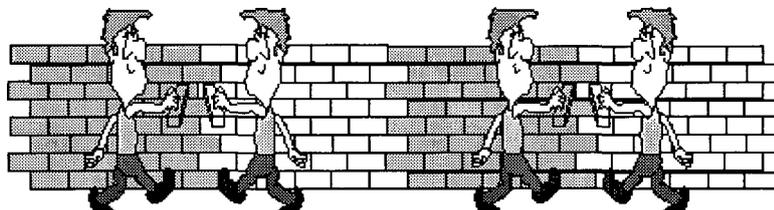


Figura 2.5 - Construção de um muro sobre o paradigma de paralelismo de dados.

Na construção de um muro da Figura 2.5 pode-se fazer uma analogia com o paralelismo de dados, na medida em que todos os trabalhadores são responsáveis pela construção de diferentes seções do muro, e todos os trabalhadores realizando a mesma tarefa.

(ii) Paralelismo funcional

Para o paralelismo funcional é realizada a decomposição do algoritmo em diferentes seções e cada processador é responsável por uma seção. Na Figura 2.6 é apresentado o exemplo da construção do muro para demonstrar o paralelismo funcional.

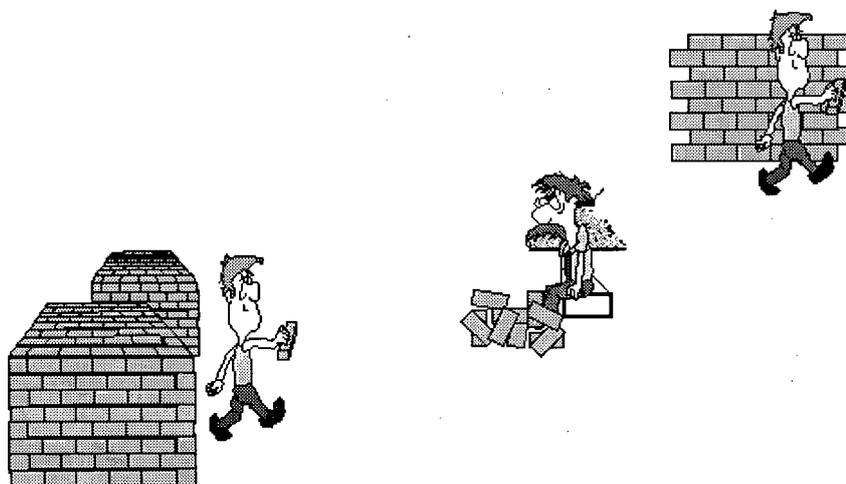


Figura 2.6 - Construção de um muro sobre o paradigma do paralelismo funcional.

Observa-se na Figura 2.6 que são atribuídas diferentes tarefas para cada trabalhador. É feita a divisão das tarefas entre os trabalhadores, onde um é responsável por carregar os tijolos, o outro por passar o cimento e outro por assentá-lo na parede.

(iii) Paralelismo *pipelining*

No paralelismo *pipelining*, um conjunto de tarefas semelhantes é realizado nos processadores. Na Figura 2.7 é apresentado o exemplo da construção de um muro sobre o paradigma do paralelismo *pipelining*.

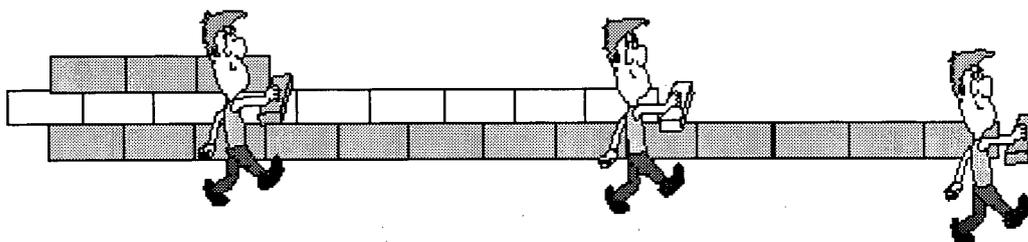


Figura 2.7 - Construção de um muro sobre o paradigma do paralelismo *pipelining*.

A Figura 2.7 apresenta um exemplo, no qual a tarefa é decomposta em seções horizontais, e cada trabalhador é responsável por uma seção.

2.2 AVALIAÇÃO DE SISTEMAS COMPUTACIONAIS

2.2.1 TÉCNICAS DE AVALIAÇÃO

Avaliação do desempenho de Sistemas Computacionais pode ser requerida a partir do momento em que apresente a necessidade de obter a melhor performance do sistema, ou de como manter a melhor performance na rede. Em [JAI91] são apresentadas três técnicas para avaliação de desempenho de sistemas: **modelo analítico**, **simulação** e **medida**. Algumas das características dessas técnicas são mostradas no Quadro 2.2.

Quadro 2.2 - Critérios de seleção da técnica de Avaliação de Desempenho.

Critério	Modelo Analítico	Simulação	Medida
1. Estágio	Qualquer	Qualquer	Depois do protótipo
2. Tempo necessário	Pequeno	Médio	Variado
3. Ferramentas	Analistas	Linguagem de simulação	Instrumentação
4. Confiabilidade	Baixa	Moderada	Variada
5. Custo	Baixo	Médio	Alto

Observa-se no Quadro 2.2 os critérios de seleção da técnica de avaliação de desempenho. O modelo analítico é uma técnica que exige muito do analista, ele pode ser realizado por meio de equações matemáticas, teoria de filas etc. Trabalhos de pesquisas com essa técnica podem ser vistos em [COS99], [SAL00a], [SAL00b] e [SAL00c].

A simulação é uma técnica que frequentemente está mais próxima da realidade, possibilitando procurar combinações mais adequadas para os parâmetros de um determinado sistema. Trabalhos de pesquisas relacionados com essa técnica podem ser encontrados em [RUB99a], [RUB99b], [RUB99c] e [SAL99].

A técnica de medida é a realidade propriamente dita. Com os dados coletados, é possível concluir sobre o comportamento do ambiente, detectar falhas e prevenir futuros problemas. Trabalhos de pesquisas relacionados com essa técnica em redes podem ser encontrados em [BER96], [BRA99], [LEO00], [WEL96] e [WEL97].

2.2.2 MÉTRICAS DE PERFORMANCE

Métricas Individuais

- Vazão

Vazão (ou *throughput*) é a taxa na qual os pedidos são atendidos (servidos) pelo sistema. Por exemplo, em caso de uma interface de rede é a quantidade de pacotes por

segundo transmitidos por unidade de tempo. Essa taxa pode ser medida em pacotes por segundo (pps) ou bits transmitidos por segundo (bps) [FRA96].

A vazão pode, normalmente, aumentar ou diminuir conforme a carga do sistema. A capacidade nominal é a vazão máxima sob condições de carga ideal; em redes de computadores é conhecido como largura de banda. A capacidade útil do sistema é determinada definindo um tempo de resposta máximo e calcula-se a vazão para este limite.

A eficiência do sistema é determinada pela razão entre a capacidade útil do sistema e a capacidade nominal (largura de banda):

$$\text{eficiência} = \frac{\text{capacidade útil do sistema}}{\text{largura de banda}}$$

Por exemplo, se a vazão de uma rede Fast-Ethernet a 100Mbps é de 28Mbps, então a sua eficiência é de 28%.

$$\text{eficiência} = \frac{28 \text{ Mbps}}{100 \text{ Mbps}} = 28\%$$

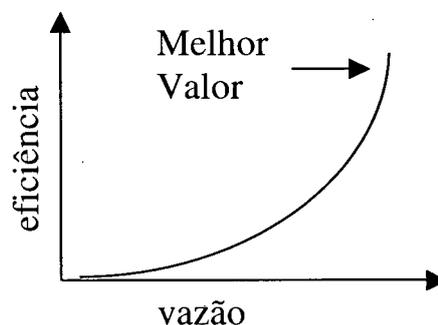


Figura 2.8 - Vazão x eficiência.

O gráfico da Figura 2.8 [JAI91] apresenta o comportamento dessa técnica, que é a relação entre vazão e sua eficiência. Quanto maior for a vazão, melhor será a eficiência do sistema, ou seja, a vazão é diretamente proporcional a sua eficiência.

- Latência

Latência é o tempo em que um pacote trafega de um ponto para outro, fim-a-fim, em uma rede [SHU99]. O termo latência é definido na RFC 1242 [BRA91], intitulado “Terminologia de Benchmarking para Equipamentos de Interconexão de Rede”.

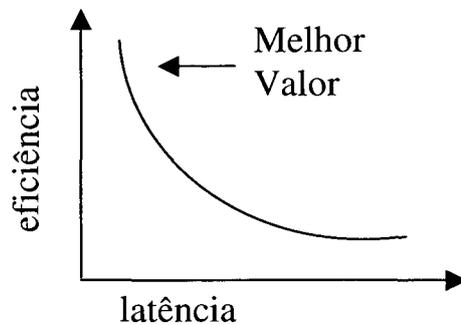


Figura 2.9 - Latência x eficiência.

Pelo gráfico da Figura 2.9 [JAI91] pode-se observar o comportamento da relação entre a latência e sua eficiência. Quanto menor for a latência, mais eficiente será o sistema. A latência é inversamente proporcional à eficiência do sistema.

Vários fatores influenciam para determinar a latência [SHU99]:

- distância geográfica;
- vazão;
- o tamanho do pacote;
- a carga de trabalho na rede;
- protocolos [TAN96]; e
- interfaces de rede [TAN96].

- Medidas de Latência X PAD

O tempo de execução das aplicações paralelas dependerá também do tempo de comunicação na rede, pois no tempo de transmissão dos dados as aplicações ficam paradas esperando estes dados chegarem.

As equações de regressão linear a serem obtidas neste trabalho permitem extrapolações lineares, as quais viabilizam a obtenção de boas aproximações do tempo de transporte de qualquer tamanho de segmento. Através deste tempo medido, pode-se avaliar, por meio de modelos analíticos, o comportamento das aplicações para prever o tempo de utilização da rede. Dependendo da aplicação pode-se ajustá-la de modo a obter um melhor aproveitamento da rede e de estações disponíveis. Além disso, os resultados de modelos analíticos podem permitir o planejamento de ambientes de *cluster*, ajudando a decidir quais os tipos de rede (Ethernet, FastEthernet, ATM, Miranet) deve-se utilizar [HAA00a].

Modelos analíticos para avaliar uma aplicação que utiliza ambientes PAD, por exemplo, podem ser encontrados em [HAA00] e [JOH94].

3 MATERIAIS E MÉTODOS

Um ambiente de Processamento de Alto Desempenho é formado por um contexto de aplicações, estações de trabalho, sistemas operacionais, interfaces de redes e equipamentos de interconexão. Este trabalho apresenta um estudo sobre diferentes estações, sistemas operacionais e redes, evidenciando somente o tempo de latência, que é muito importante para o PAD, de algumas partes isoladamente e num conjunto, para apoiar a decisão de escolha de qual *cluster* a ser utilizado ou adquirido.

Este capítulo apresenta o ambiente de trabalho e a maneira de como serão realizados os testes para verificar as latências das estações e dos *clusters*.

3.1 AMBIENTE DE AVALIAÇÃO

O ambiente de pesquisa está baseado na sub-rede *cluster* da rede UFSC da Universidade Federal de Santa Catarina, conforme Figura 3.1. O *cluster* é composto por um conjunto de estações de trabalho interligadas por uma rede de alta velocidade e que possui uma base comum de usuários, limitando o nosso contexto de trabalho em:

- 6 estações IBM 43P;
- 6 Pentium II-400MHz;
- 1 IBM SP2 com 9 nós;
- 7 Switches sendo:
 - 4 ATM IBM 8265 (CPSW)
 - 1 CoreBuilder 7000 da 3COM (cb7000);
 - 1 IBM 8271 modelo 712 (sb01-UFSC); e
 - 1 HPS (o switch interno do SP2).

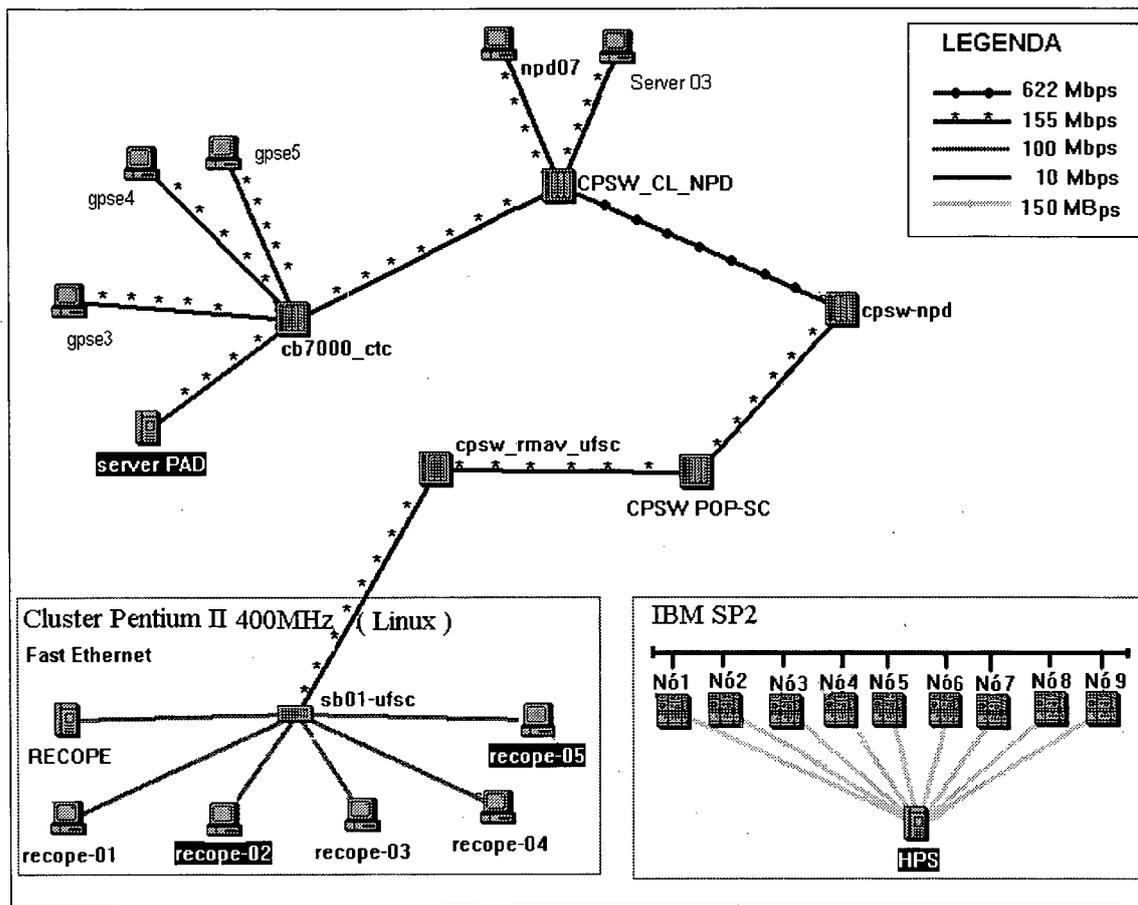


Figura 3.1 - Topologia da rede *Cluster UFSC*.

Os equipamentos estão interligados por uma rede ATM de 155 e 622 Mbps, Fast Ethernet de 100 Mbps e Ethernet de 10 Mbps.

O SP2 da UFSC pode ser visto como um conjunto de 9 nós (ou estações) agrupados dentro de um gabinete interligado por um *switch* (HPS) de alta velocidade (150 MBps) e por uma rede Ethernet. O nó 1 e 2 possuem 128MB de RAM e 4.5GB de disco. O nó 3 e 4 possuem 256MB de RAM 4.5GB de disco. O nó 5 possui um disco de 4.5GB e 512MB de RAM. O nó 6, 7, 8 e 9 possuem 256MB de RAM e 4.5GB de disco.

O Quadro 3.1 apresenta as características das estações como o sistema operacional, o nome do agrupamento (*cluster*) que elas são referenciadas e o tipo de rede utilizada por esse *cluster*.

Quadro 3.1 - Estações/Clusters utilizados na Avaliação.

Agrupamento	Estações	Rede	Hardware	Sistema Operacional
Cluster AIX	Gpse3, Gpse4, Gpse5 e Server Pad	ATM 155Mbit/s	Risc 6000 - 43P	AIX versão 4.3
Cluster Linux	Recope, Recope-1, Recope-2, Recope-3, Recope-4, Recope-5	Fast-Ethernet	Pentium II-400MHz	Linux – Red Hat Versão: 6.1 Kernel 2.12
SP2	SP2 (com 9 Nós)	Ethernet	IBM 9076, Risc 6000	AIX versão 4.1
-	Srv03	ATM 155Mbit/s	RISC 6000 – E30	AIX versão 4.3
-	Npd07-Lane	ATM 155Mbit/s	RISC 6000 - 43P	AIX versão 4.3

Na topologia da rede *cluster* UFSC a rede ATM é utilizada com serviço LANE simulando a Ethernet.

O Serviço LANE é um serviço de conectividade de rede que permite sistemas finais conectar-se a uma rede ATM como se fosse conectado a uma rede local tradicional (Ethernet e Token Ring). Um rede emulada, denominada ELAN (*Emuled LAN*), consiste de sistema finais com interfaces ATM: roteadores, dispositivos de porta (comutadores e *bridges*) e estações [MEL00].

Nesta pesquisa o Serviço LANE da rede ATM permitirá que os sistemas finais que utilizam a rede Ethernet possam usufruir as vantagens da rede ATM com velocidade de 155 Mbit/s.

3.2 MÉTODO DE MEDIÇÃO DE LATÊNCIA EM AMBIENTES COMPUTACIONAIS

Para medir a latência, fim-a-fim, em ambientes computacionais foram utilizados a RFC 2544 (*Request for Comments*) [BRA99], que faz recomendações para medição de latência em equipamentos de interconexão e algumas considerações de [TAN96] para avaliação de desempenho.

Para mensurar a latência são considerados os seguintes parâmetros:

- o tamanho dos segmentos é de 64, 128, 256, 512, 1024, 1280 e 1518 bytes, que é definido para as redes Ethernet pela RFC 2544;
- a duração do teste é de 70 segundos, de acordo com a RFC 2544. O tempo de duração do teste foi determinado próximo ao valor mínimo definido pela RFC 2544 para reduzir o tempo de utilização dos testes nesses ambientes;
- o número de repetições é de 20 vezes como a quantidade representativa para obtenção da média, de acordo com a RFC 2544;
- as medições são sequenciais durante a semana e finais de semana para que a amostra seja significativa, conforme em [TAN96];
- o MTU (unidade máxima de transmissão) usada é de 1500 *octetos* (padrão); e
- o tamanho da janela para o protocolo TCP é o definido pela RFC1323 “*Large Windows*”, padrão para os sistemas operacionais em análise.

3.2.1 FERRAMENTAS DE MEDIÇÃO

No início deste trabalho, foi pensado em utilizar o *ping Internet Control Message Protocol* - ICMP como a ferramenta para medir a latência, mas este procedimento não foi seguido pois os dados obtidos por esse utilitário talvez não sejam os mais exatos (os resultados dependerão da sobrecarga do dispositivo) [FRA96]. As unidades de medidas esperadas neste trabalho sejam da ordem de milisegundos (e frações dessa unidade) e as unidades de medidas do *ping* ICMP são mensuradas apenas em milisegundos conforme [DIS00].

Foi feito um levantamento de informações sobre as ferramentas na RFC 2398 [PAR98], “*Algumas Ferramentas Implementadas para Teste em TCP*”, que apresenta algumas ferramentas implementadas para estes tipos de testes. Dentre as ferramentas citadas pela RFC 2398 o Netperf é utilizado para mensurar a largura de banda de uma conexão ou a latência para TCP, UDP e DLPI, além de poder mensurar a utilização da CPU. A ferramenta também é referenciada pelo Centro de Super Computação de Pittsburgh – USA.

Outro ponto observado da ferramenta é o seu funcionamento nas seguintes plataformas:

- AIX;
- BSD
- FreeBSD
- HP-RT
- HP-UX
- IRIX
- Linux
- MPE/Ix
- MS-DOS
- NetBSD
- NT
- OSF
- SCO
- Solaris
- SunOS
- Unixwave
- Windons

Além de ser indicada para mensurar nas redes do Quadro 3.2:

Quadro 3.2 - Redes indicadas para utilização do Netperf.

ATM	Taxa desconhecida 25Mbit/s 51Mbit/s 100Mbit/s (TAXI) 155Mbit/s (SONET) 622Mbit/s
Ethernet	10Mbit/s
TokenRing	
FastEthernet	100Mbit/s CSMA/CD (100 BaseT, 100 BaseX etc.)
FDDI	
Token Ring	~ 4 Mbit/s (IEEE 802.5-4) 16 Mbit/s (IEEE 802.5-16)

O Netperf é o escolhido para medir a latência fim-a-fim, considerando que os sistemas operacionais dessa pesquisa são o AIX e o LINUX, e as redes em estudos são o ATM, Ethernet e Fast-Ethernet, os quais fazem parte das especificações da ferramenta.

3.2.2 NETPERF

O Netperf é uma ferramenta de *benchmark* para performance de rede de computadores. Ela foi desenvolvida pela Divisão de Informação de Redes da Hewlett-Packard Company . A versão utilizada nesse trabalho foi a 2.1.3 de setembro de 1997.

A ferramenta é composta de dois programas básicos – netperf e netserver. Baseado no modelo cliente-servidor, nesse caso o modelo monitor/refletor [LEO00], onde o tráfego é gerado de uma estação monitora, com o programa netperf, até a estação refletora, com o netserver, no qual o tráfego é refletido e retornado à estação monitora.

A terminologia do Netperf é a definida pela “Terminologia de *Benchmarking* para Equipamentos de Interconexão de Redes Computadores” definida pela RFC 1242 [BRA91]. Este documento da IETF é o primeiro produto do Grupo de Trabalho de Metodologia de *Benchmarking*.

3.3 COLETA DE DADOS DENTRO DO CLUSTER

A ferramenta Netperf permite medir a latência, fim-a-fim, das estações. A instalação da ferramenta Netperf, e netserver, será efetuada nas seguintes máquinas, conforme o Quadro 3.3, que apresenta entre quais estações serão realizadas as medições das latências.

Quadro 3.3-Instalação da Ferramenta para avaliação da latência dentro do cluster.

Netserver	Netperf
Gpse3	Gpse4
Recope	Recope-01, Recope-02, Recope-03, Recope-04, Recope-05
Recope-01	Recope-03
Recope-04	Recope-05
SP2-N64	SP2-N61, SP2-N62, SP2-N63, SP2-N65, SP2-N66, SP2-N67, SP2-N68 e SP2-N69.
SP2-N61	SP2-N62
SP2-N65	SP2-N66
ServerPad	Gpse3, Gpse4, Gpse 5, Server 03

O Quadro 3.3 apresenta que, para cada *cluster* é definido uma estação na qual todas as outras irão enviar os segmentos. Como também são apresentadas outras estações para verificar a latência das mesmas dentro do *cluster*.

O teste é realizado da seguinte maneira: ativar o programa netserver da ferramenta na máquina que será a refletora, e após executa-se o *script* que chama o programa netperf passando as seguintes informações:

- a duração;
- o tipo do teste definindo o protocolo (TCP ou UDP);
- o endereço IP, da máquina remota executando o netperf; e
- o tamanho do segmento;

Um exemplo da configuração para a estação Gpse3, é executar o seguinte comando para ativar a ferramenta no modo refletor:

```
/urs/local/netserver
```

Um outro exemplo para as estações executarem o programa netperf é utilizar um *script*, conforme é ilustrado na Figura 3.2.

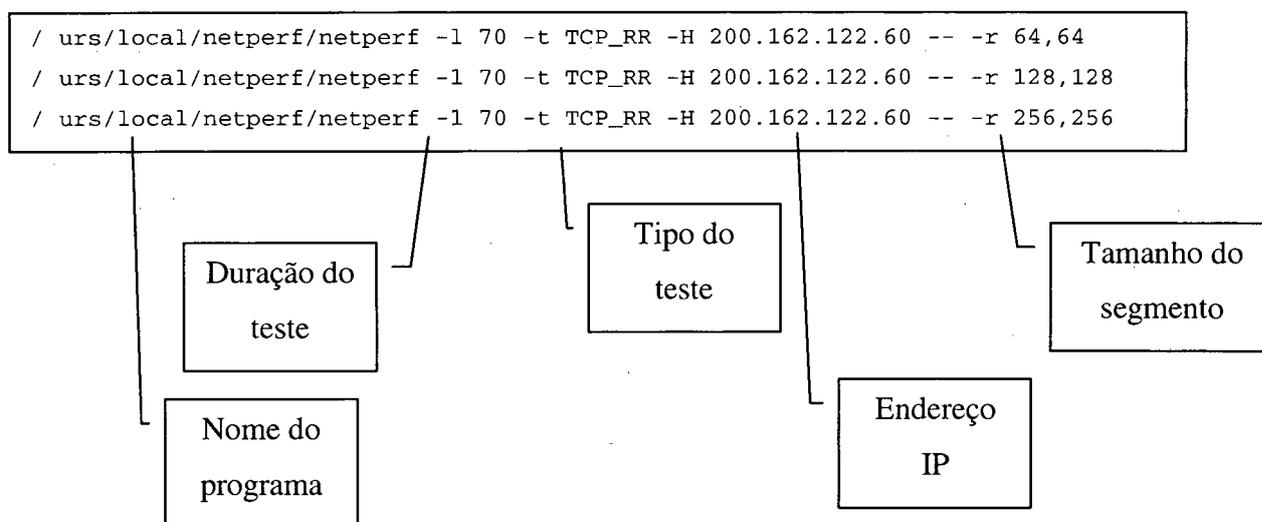


Figura 3.2 - Formato do arquivo do Script.

Na Figura 3.2 pode-se observar o *script* para a execução de três testes com o netperf. Na primeira linha pode-se visualizar os seguintes parâmetros:

- l, que define o tempo de execução do teste, no exemplo de 70 segundos;
- t, que define em qual protocolo será analisado a latência, no exemplo o TCP;
- H, que define o número IP da máquina remota que deverá estar executando o programa netserver; e
- r, que define o tamanho do segmento que será realizado o teste, no exemplo de 64 bytes.

O netperf permite a configuração de diversos parâmetros, como, por exemplo, o tamanho do *buffer* das estações, local e remota. Nesta pesquisa serão considerados todos os parâmetros *default*, do sistema operacional.

A ferramenta irá fornecer o número de transações realizadas em 1 segundo. A partir dessa informação pode-se inferir o valor da latência.

$$latência = \frac{1}{\text{no. de transações/s}}$$

Um exemplo pode ser apresentado da seguinte maneira, com a execução do *script* da Figura 3.2, para o segmento TCP de tamanho de 64 bytes obtém-se um resultado de 250 transações/s, então o valor da latência será:

$$latência = \frac{1}{250} = 0,004 \text{ segundos}$$

3.3.1 COLETA DE DADOS DAS ESTAÇÕES

Esta medição é realizada para verificar a influência da latência interna da estação sobre a latência fim-a-fim de cada *cluster*.

Esse teste apresenta a latência que ocorre para que um segmento saia de dentro da estação até a interface de rede e retorne à CPU. Mas devido à utilização da ferramenta

Netperf, a medição enviará o segmento duas vezes à interface de rede e para cada vez que isso ocorra, ela retornará à CPU. Ou seja, representa a latência em que um segmento é enviado pela CPU da estação até a interface de rede, depois da interface até a CPU, e o seu retorno, de ida até a interface e de volta à CPU.

No *cluster* Aix essa medição será na estação *gps5*, no *cluster* Linux será na estação *recope 2* e no SP 2 na estação nó 7.

O teste será realizado da seguinte maneira: ativa o programa *netserver* da ferramenta na máquina que será monitorada. Após isso, executa-se um *script* que executa o programa *netperf*, passando as seguintes informações:

- a duração;
- o tipo do teste definindo o protocolo (TCP ou UDP);
- o endereço IP; e
- o tamanho do segmento.

Os protocolos utilizados para os testes são o TCP e o UDP. O endereço IP para este teste será o de *loop back* (127.0.0.1), conforme exemplo da Figura 3.3. Os testes foram realizados nas estações em períodos em que ela não estava sendo utilizada por nenhum usuário.

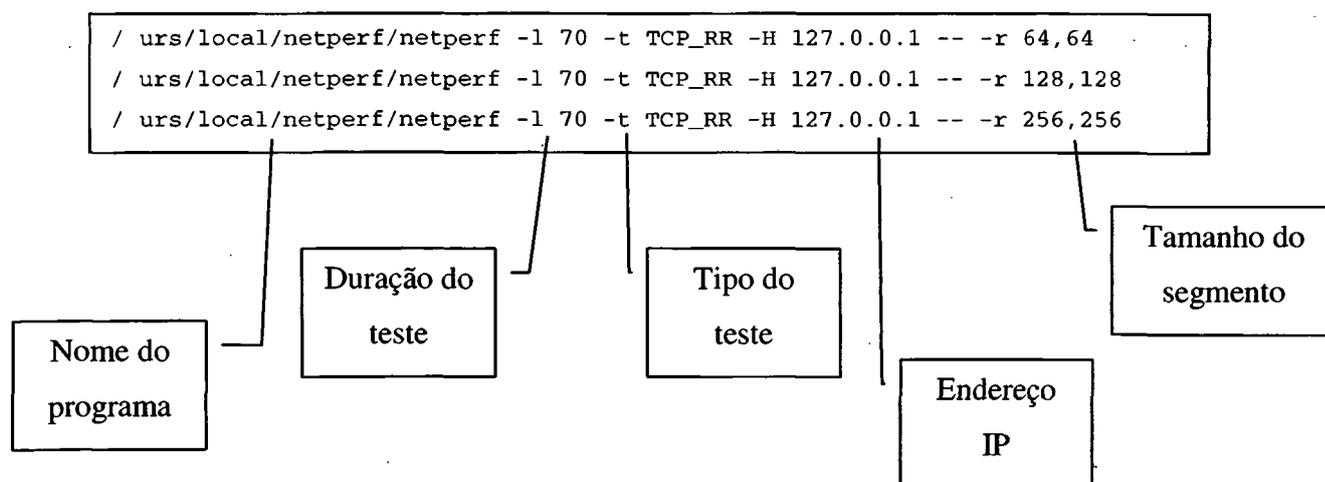


Figura 3.3 - Formato do arquivo do *Script* para medição da latência da estação.

Na Figura 3.3 pode-se observar o *script* para a execução de três testes com o *netperf*. Na primeira linha pode-se observar os seguintes parâmetros:

- l, que define o tempo de execução do teste, no exemplo de 70 segundos;
- t, que define em qual protocolo será analisado a latência, no exemplo o TCP;
- H, que define o número IP de *loop back*; e
- r, que define o tamanho do segmento que será realizado o teste, no exemplo de 64 bytes.

Neste capítulo foi descrito o ambiente de trabalho, o método para medir a latência e a ferramenta a ser utilizada.

4 AVALIAÇÃO DOS RESULTADOS

A partir da descrição do ambiente, da ferramenta e do método de medição, os dados foram coletados e posteriormente tabulados e analisados, através do programa de planilha eletrônica (Excel) e programa de estatística (Statistica).

Os dados medidos são apresentados num diagrama de caixa para cada protocolo, onde pode-se ver o comportamento deles. O diagrama apresenta o valor mínimo (L) e o valor máximo (T) dos dados coletados. O diagrama também mostra a posição dos quartis inferior e superior (\square) dos dados coletados, assim como a posição da mediana (\square).

Além disso, é verificada uma relação matemática entre a latência e o tamanho dos segmentos. Através da análise de regressão é apresentada uma equação para relacionar a variabilidade da latência em função do tamanho do segmento e do seu respectivo coeficiente de determinação. Esses valores são apresentados em gráficos.

Para cada equação definida será apresentado o coeficiente de determinação R^2 [BAR98], que representa a seguinte razão:

$$R^2 = \frac{\sum \left(\hat{Y} - \bar{Y} \right)^2}{\sum \left(Y - \bar{Y} \right)^2} = \frac{\text{variação explicada}}{\text{variação total}}$$

Onde:

\hat{Y} = valor predito

\bar{Y} = média aritmética

Y = valor observado

Como já foi descrito, o tamanho dos segmentos variam entre 64 e 1.518 bytes e as latências serão apresentadas em milisegundos.

A latência interna da estação, que ocorre quando um segmento sai da camada de aplicação até chegar à interface de rede da estação, foi analisada também. Como a

latência interna foi muito pequena em relação à latência fim-a-fim desses *clusters*, a mesma não será apresentada.

4.1 CLUSTER AIX

Pela medição da latência no *cluster* Aix pode-se realizar uma análise das latências para os segmentos nos protocolos TCP e UDP.

4.1.1 CLUSTER AIX UTILIZANDO O PROTOCOLO TCP

Os valores da latência por tamanho de segmento TCP para o Cluster Aix são apresentados no digrama da Figura 4.1. O diagrama apresenta o valor mínimo (⊥) e o valor máximo (⊤) de latência para cada segmento. O diagrama também mostra a posição dos quartis inferior e superior () da latência, assim como a posição da mediana (□).

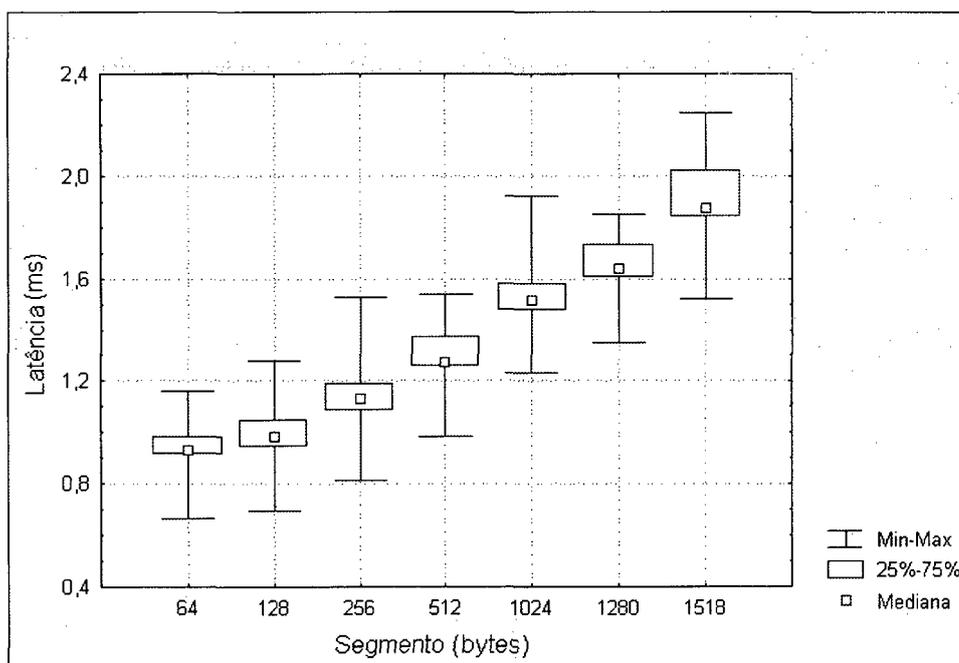


Figura 4.1 – Neste diagrama são mostrados os valores da latência (ms) por tamanho de segmento TCP (bytes) para o Cluster Aix. As barras verticais correspondentes a latência máxima, mínima para cada segmento, o das caixas maiores concentram-se de 25% até 75% dos valores e a caixa menor indica o valor da mediana. Note-se que a tendência exponencial apresentada pelos valores da latência em função do tamanho do segmento é devido ao fato de que o tamanho segmento não segue uma relação linear.

No diagrama da Figura 4.1 pode-se observar que para segmentos de TCP de tamanho de 64 bytes o valor mínimo da latência é de aproximadamente de 0,6 milissegundos e o valor máximo é de 1,2 milissegundos. Pode-se observar também que os quartis inferior e superior localizam-se próximos de 0,9 e 10 milissegundos. Para o tamanho do segmento TCP de 128 bytes o valor mínimo é de aproximadamente 0,7 milissegundos e o valor máximo é de aproximadamente 1,3 milissegundos. Analisando o diagrama observa-se um comportamento de crescimento da latência em relação ao segmento TCP.

É importante constatar que a latência está relacionada diretamente com o tamanho do segmento, através do qual pode-se estabelecer uma relação entre ambos os valores. Foi feita a análise de regressão e definida a equação linear (1) com o seu respectivo coeficiente de determinação.

$$Lat_TCP = 0,925607 + 0,000615 * tamanho_segmento_TCP \quad (1)$$

$$R^2 = 0,834197$$

A partir do valor de R^2 obtido, próximo de 1, pode-se afirmar que essa equação representa o comportamento da variabilidade da latência da rede, o que pode ser explicada em 83% pelos diferentes tamanhos de segmentos para intervalos entre 64 a 1.518 bytes e os outros 17% devido a outros fatores.

Os tamanhos dos segmentos estão de acordo com a RFC 2544, para segmentos maiores pode ser afetado por outros parâmetros, como o valor do MTU etc.

Na Figura 4.2 observa-se o gráfico dos valores da latência e do tamanho do segmento TCP com a equação de regressão para o *cluster* Aix de todos os dados medidos. O modelo matemático reafirma a relação que existe entre a latência e o tamanho do segmento.

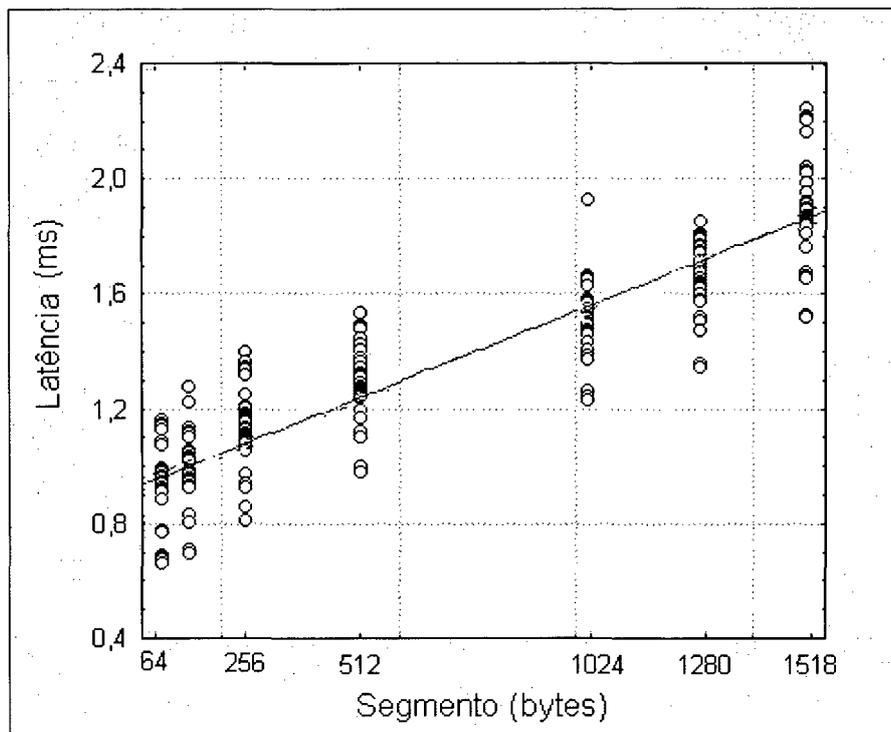


Figura 4.2 – Gráfico dos valores de latência (ms) pelo tamanho do segmento TCP (bytes) para o Cluster Aix. A linha reta representa a regressão linear efetuada. Os círculos representam os valores amostrais. Note-se que a reta ajusta-se bem aos dados.

Também observa-se na Figura 4.2 a pequena variação da latência em relação ao tamanho do segmento sobre a reta da equação (1), justificando o alto coeficiente de determinação.

4.1.2 CLUSTER AIX UTILIZANDO O PROTOCOLO UDP

Na Figura 4.3 é apresentado o diagrama da latência versus o tamanho do segmento UDP do *cluster* Aix dos dados medidos. O diagrama apresenta o valor mínimo (L) e o valor máximo (T) de latência para cada segmento. O diagrama também mostra a posição dos quartis inferior e superior (□) dos dados coletados, assim como a posição da mediana (◻).

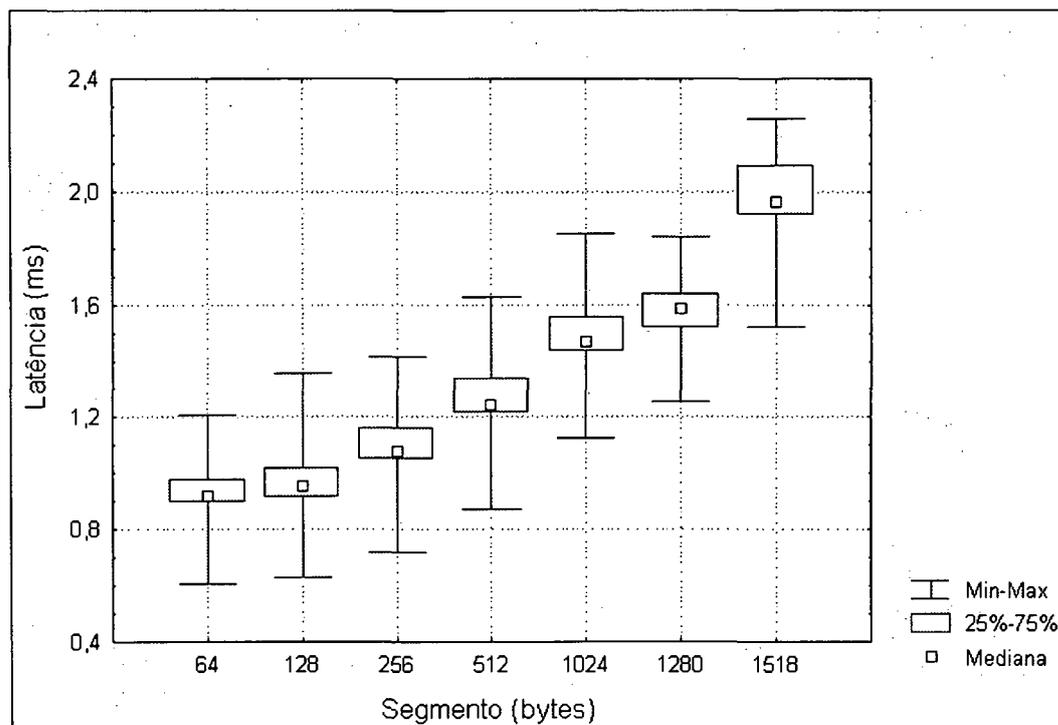


Figura 4.3 – Neste diagrama são mostrados os valores de latência (ms) pelo tamanho do segmento UDP (bytes) para o Cluster Aix. Idem a figura 4.1.

É apresentado no diagrama da Figura 4.3 o comportamento dos dados medidos. Para o segmento UDP de tamanho de 64 bytes a latência mínima medida foi de aproximadamente 0,6 milissegundos e a máxima de 1,2 milissegundos, os quartis inferior e superior encontram-se na faixa de 0,9 e 1,0 milissegundos, apresentando um comportamento regular.

Para os segmentos de tamanho de 128 bytes a latência mínima é de aproximadamente 0,6 milissegundos e o valor máximo é de aproximadamente 1,4 milissegundos e também apresenta pequena variação entre o quartis inferior e superior, apresentando um comportamento de pequena variação. Pode-se observar que há um crescimento da latência em relação ao tamanho do segmento UDP.

Para constatar a relação entre esses valores foi feita uma análise de regressão linear, definindo a equação (2) e o seu respectivo coeficiente de determinação para o cálculo da latência versus o tamanho dos segmentos UDP.

$$Lat_UDP = 0,894871 + 0,000636 * tamanho_segmento_UDP \quad (2)$$

$$R^2 = 0,83279$$

O coeficiente de determinação $R^2 = 0,83279$ é estatisticamente significativo, confirmando que a equação (2) é representativa, e que essa equação representa o comportamento da variabilidade da latência da rede que pode ser explicada em 83% pelos diferentes tamanhos de segmentos para intervalos entre 64 e 1.518 bytes.

Na Figura 4.4 observa-se os dados medidos em relação à equação (2). O modelo matemático reafirma a relação existente entre a latência e o tamanho do segmento.

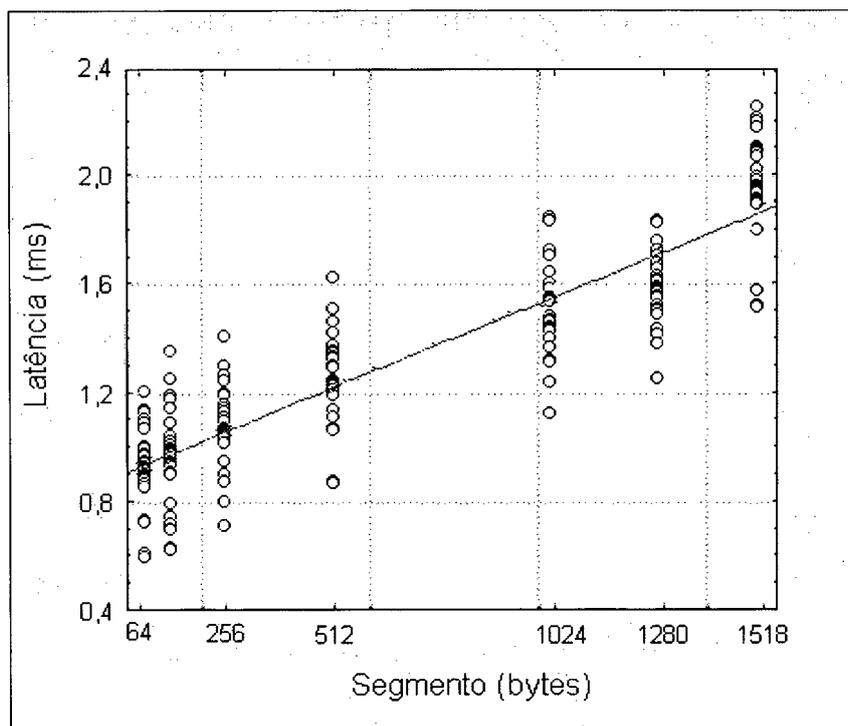


Figura 4.4 – Gráfico da latência (ms) pelo segmento UDP (bytes) no Cluster Aix e a equação de regressão linear. Idem Figura 4.2.

Pode-se observar no gráfico da Figura 4.4 que há uma pequena variação da latência em relação ao tamanho do segmento UDP em relação à equação (2).

4.2 CLUSTER LINUX

Devido a uma incompatibilidade da versão do sistema operacional com a versão da ferramenta Netperf utilizada para obter as informações não foi coletada a latência para segmento UDP do *cluster* Linux.

Na análise dos dados coletados do cluster Linux observou-se um comportamento irregular da latência para o segmento de tamanho 1.580 bytes. Como é apresentado no diagrama da latência média medida no *cluster* Linux na Figura 4.5.

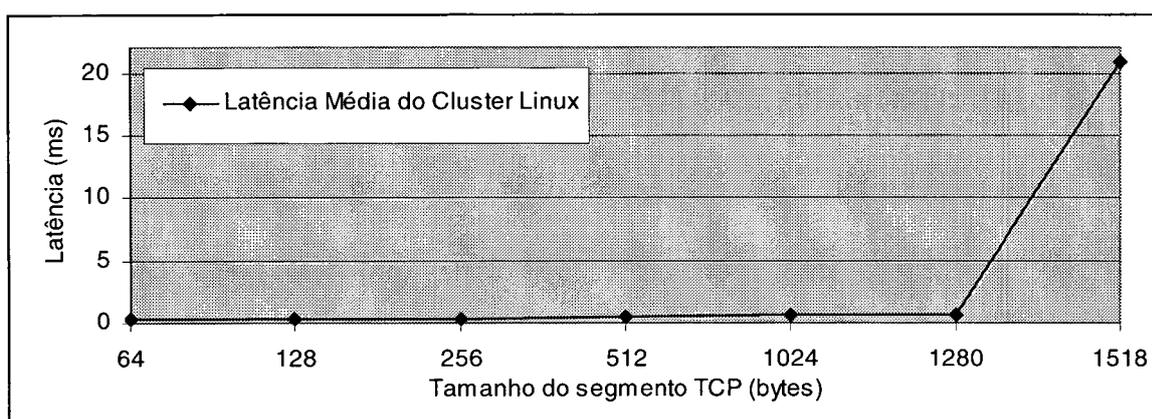


Figura 4.5 – Neste diagrama são mostrados os valores da latência (ms) por tamanho de segmento TCP (bytes) para o Cluster Linux – Note-se o comportamento irregular da latência média para segmento de 1.518 bytes.

Observa-se no diagrama da Figura 4.5 o comportamento entre os segmentos e a latência para o *cluster* Linux que apresenta pequeno valores de latência para segmentos até 1.280 bytes, havendo um aumento exponencial para o segmento de tamanho de 1.518 bytes. As causas desse comportamento poderiam ser:

- um erro de configuração do tamanho de MTU nas estações;
- um erro de configuração do tamanho de MTU o *switch* (sb01-ufsc)
- algum erro na implementação do protocolo com as estações do *cluster* Linux; ou
- algum erro com o *switch*.

Então, para verificar esse comportamento, foram observados os tamanhos dos MTU, e chegou-se à conclusão de que as configurações permaneciam de acordo com as especificações.

Para verificar onde ocorreu o erro na estação ou no *switch* foram realizados alguns testes de *loop back* nas estações, e obteve-se um resultado semelhante entre as estações, que é apresentado no diagrama da Figura 4.6.

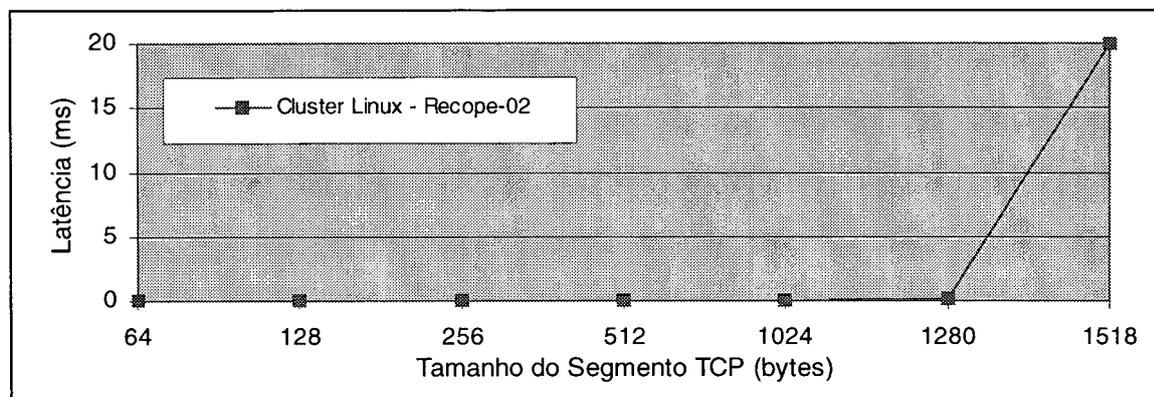


Figura 4.6 – Neste diagrama são mostrados os valores da latência (ms) por tamanho de segmento TCP (bytes) para a latência de *Loop Back* na estação Recope02. – Idem a figura 4.5.

No diagrama da Figura 4.6 observa-se que o problema está dentro da estação. É reconhecida a existência de um erro na implementação nesse sistema operacional do algoritmo do Nagle no protocolo TCP do Linux, o qual possui alguns problemas em relação ao desempenho do TCP com o Kernel 2.12 [BRI96]. Este mesmo autor disponibiliza um programa para correção desse problema. Essa informação foi notificada à equipe responsável pelo *cluster* Linux.

Com a correção desse erro espera-se diminuir o tempo de comunicação, reduzindo o tempo de processamento ao menor período de tempo possível, o que é um dos objetivos do PAD.

Para continuar a comparação entre os *clusters* das latências, no *cluster* Linux a análise das latências é feita para o tipo de segmento TCP, variando de 64 até 1.280 bytes.

4.2.1 CLUSTER LINUX UTILIZANDO TCP

É apresentado na Figura 4.7 a representação dos dados medidos e agrupados pelo tamanho de segmento TCP para o *cluster* Linux.

A Figura 4.7 mostra no diagrama uma pequena variação da latência para cada tamanho de segmento TCP no *cluster* Linux, como também a posição da mediana e os quartis inferior e superior dos dados coletados.

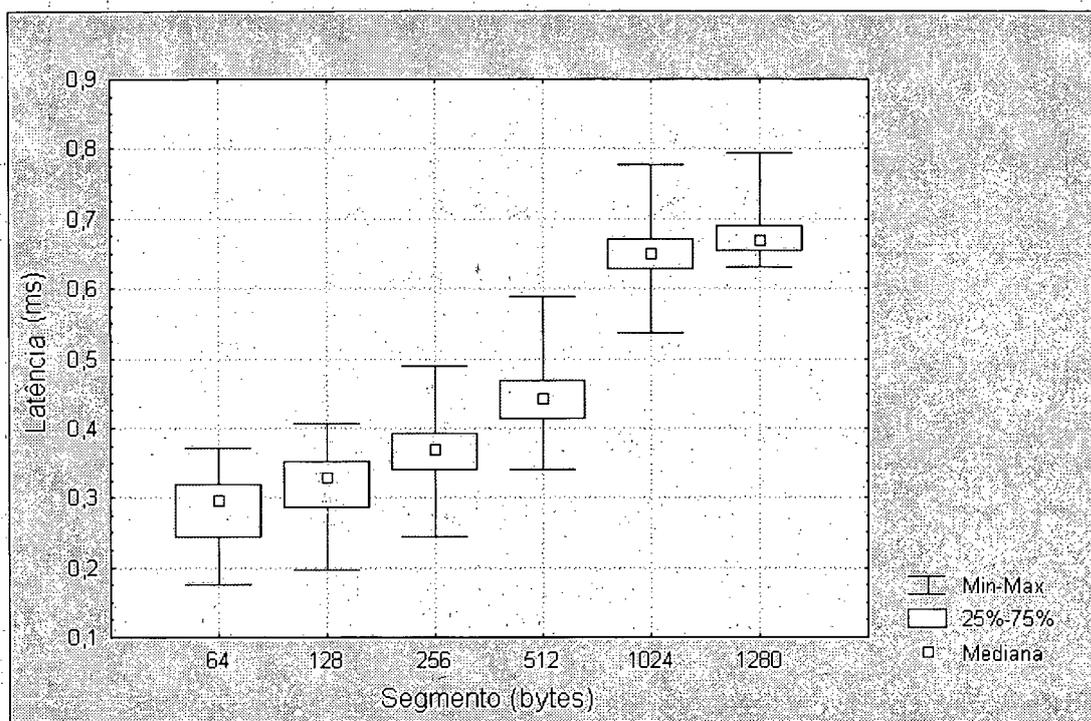


Figura 4.7– Neste diagrama são mostrados os valores da latência (ms) pelo tamanho do Segmento TCP (bytes) para o Cluster Linux. Idem a figura 4.1.

Observa-se que o valor de latência para o segmento TCP de tamanho de 64 bytes mínimo é de aproximadamente de 0,175 milissegundos e o máximo é de 0,375 milissegundos; os quartis mínimo e máximo variam em torno de 0,1 milissegundos.

Para segmentos de tamanho de 128 bytes, o valor mínimo medido é de aproximadamente 0,2 milissegundos e o máximo é de aproximadamente 0,4 milissegundos. Nota-se que existe um crescimento direto da latência em relação ao tamanho do segmento, pelo qual pode-se estabelecer uma relação linear, à qual foi

obtida por meio da análise de regressão sobre a latência monitorada, definindo a equação (3) com o seu respectivo coeficiente de determinação.

$$Lat_TCP = 0,273835 + 0,000336 * tamanho_segmento_TCP \quad (3)$$

$$R^2 = 0,913747$$

O valor do coeficiente R^2 mostra que a variabilidade da latência pode ser explicada em relação ao tamanho do segmento, nesse modelo ela cresce linearmente para intervalos de segmentos de 64 até 1.280 bytes.

Na Figura 4.8 é apresentado o gráfico da latência pelo tamanho do segmento em relação à equação (3). Os círculos representam os valores amostrais. Note-se que a reta ajusta-se bem aos dados.

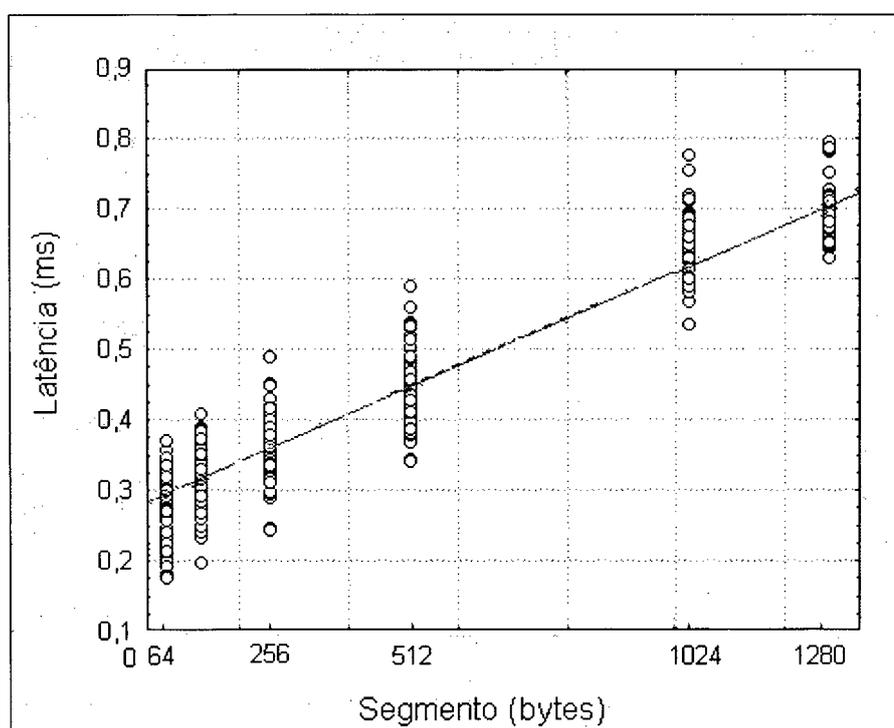


Figura 4.8 – Gráfico da latência (ms) pelo segmento TCP (bytes) no Cluster Linux e a equação de regressão linear. Idem Figura 4.2.

4.3 SP 2

Pelas medidas da latência no SP2 pode-se realizar uma análise das latências para cada tipo de segmento TCP e UDP. É importante salientar aqui, que as medidas no SP2 foram feitas na rede Ethernet pois o HPS apresentava problemas na época em que foram realizadas estas medidas.

4.3.1 SP 2 UTILIZANDO PROTOCOLO TCP

Na Figura 4.9 é apresentado o diagrama que mostra o valor mínimo (⊥) e o valor máximo (⊤) da latência para cada tamanho de segmento. Também mostra os quartis inferior e superior (□) da latência medida e a posição da mediana (◻).

Pode-se observar na Figura 4.9 que existe um crescimento da mediana da latência em relação ao tamanho dos segmentos, mas que existe uma grande variação da latência mínima e máxima. Note-se que o tamanho segmento não seguem uma relação linear.

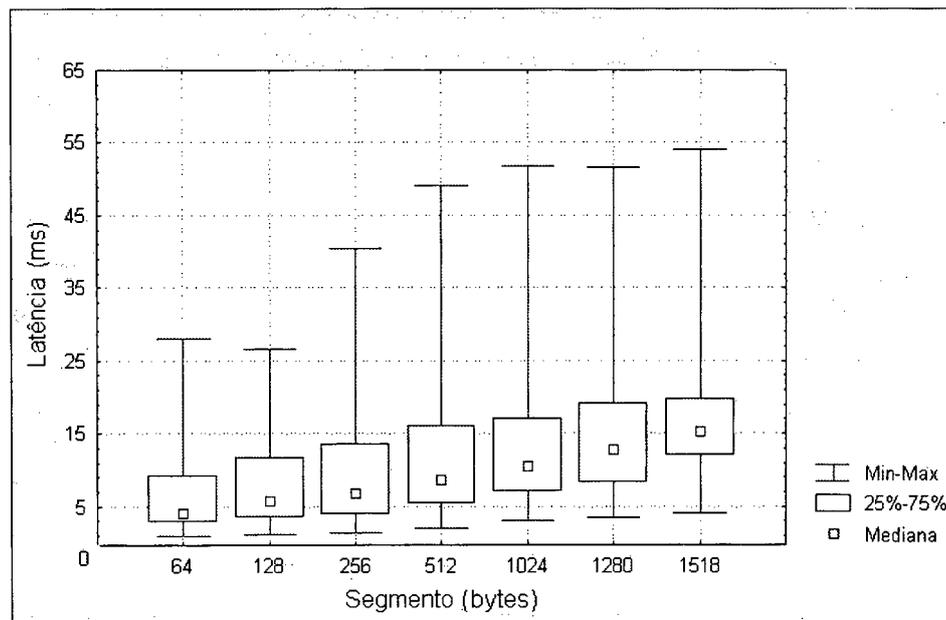


Figura 4.9 – Neste diagrama são mostrados os valores da latência (ms) pelo tamanho do Segmento TCP (bytes) para o SP 2. Idem a figura 4.1.

Observa-se na Figura 4.9 o segmento TCP de 64 bytes apresentando um valor mínimo de aproximadamente 1 milisegundos e máximo de aproximadamente 27 milisegundos. Também é apresentado para os segmento TCP de tamanho de 128 bytes essa grande variação da latência em relação ao tamanho do segmento; indicando a possibilidade da existência de um outro fator, além do tamanho dos segmentos, que está afetando o valor da latência no SP 2.

Mesmo assim, foi feita uma análise de regressão linear entre a latência e o tamanho do segmento, definindo a equação (4) com o seu respectivo coeficiente de determinação, sendo para esse caso um valor baixo, provando a existência de um outro fator que não foi analisado no SP2 para determinar a latência.

$$Lat_TCP = 1,735339 + 0,000677 * tamanho_segmento_TCP \quad (4)$$

$$R^2 = 0.231976$$

Pela Figura 4.10 pode-se observar o gráfico dos valores da latência pelo tamanho do segmento TCP e a equação de regressão no SP 2 dos dados medidos. Verifica-se a grande variação do comportamento da latência em relação à reta da equação (4).

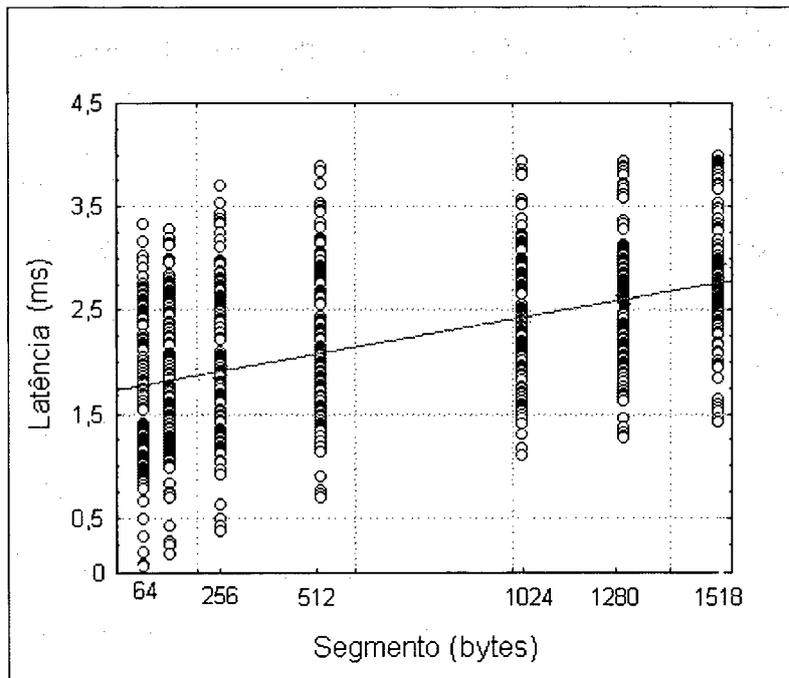


Figura 4.10 – Gráfico dos valores de latência (ms) pelo tamanho do segmento TCP (bytes) para o SP 2. A linha reta representa a regressão linear efetuada. Os círculos representam os valores amostrais. Note-se que existe uma grande dispersão dos dados em relação à reta.

O gráfico da Figura 4.10 confirma a existência de um ou mais fatores influenciando na variabilidade da latência. Com o objetivo de determinar qual ou quais fatores que estão influenciando na variabilidade da latência, foram realizados outros testes no SP 2.

O teste principal, baseado no método heurístico de tentativa e erro, foi a medição da latência entre dois nós (estações) quando o SP 2 não estava sendo utilizado. O resultado apresentou um crescimento linear da latência em relação ao tamanho do segmento, provando que o outro fator que influi na latência é a taxa de colisões, existentes no barramento Ethernet.

4.3.2 SP 2 UTILIZANDO PROTOCOLO UDP

O diagrama da Figura 4.11 apresenta uma grande variação dos valores mínimos (L) e dos valores máximos (T) para as latências medidas, em relação ao tamanho dos segmentos. Como aconteceu com os segmentos TCP aconteceu também com os segmentos UDP, apresentando um comportamento irregular dos dados medidos.

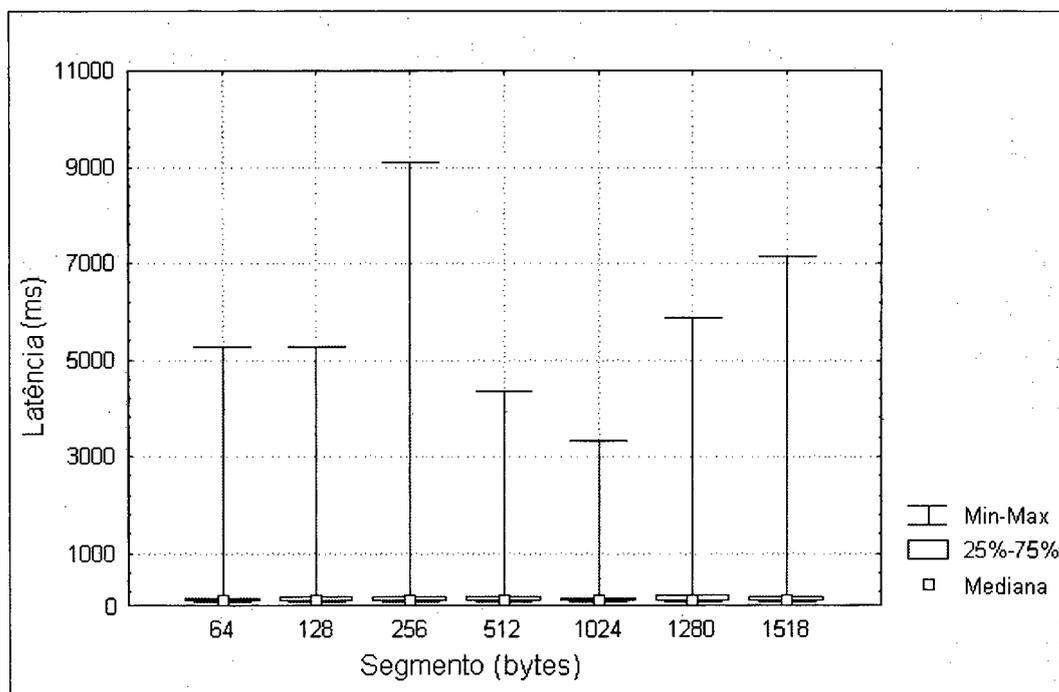


Figura 4.11 - Neste diagrama são mostrados os valores da latência (ms) pelo tamanho do Segmento UDP (bytes) para SP 2. Idem a figura 4.1.

O comportamento irregular indica a existência de um ou mais fatores influenciando sobre a variabilidade da latência. Para definir qual ou quais são esses fatores foram realizados os mesmos os mesmos testes para o segmento TCP, onde ficou provado que o outro fator também é a taxa de colisão para o segmento UDP.

Pode-se observar na Figura 4.11 a posição dos quartis inferior e superior, apresentando um comportamento que indica o efeito das colisões no valor da latência. Para esse caso, também foi realizada a análise de regressão linear definindo a equação (5) com o seu respectivo coeficiente de determinação.

$$Lat_UDP = 2,977304 + 0,000409 * tamanho_segmento_UDP \quad (5)$$

$$R^2 = 0.01587371$$

E como se esperava, o coeficiente de determinação da equação (5) também é pequeno e pode ser observado na Figura 4.12 o comportamento dos dados em relação à equação (5).

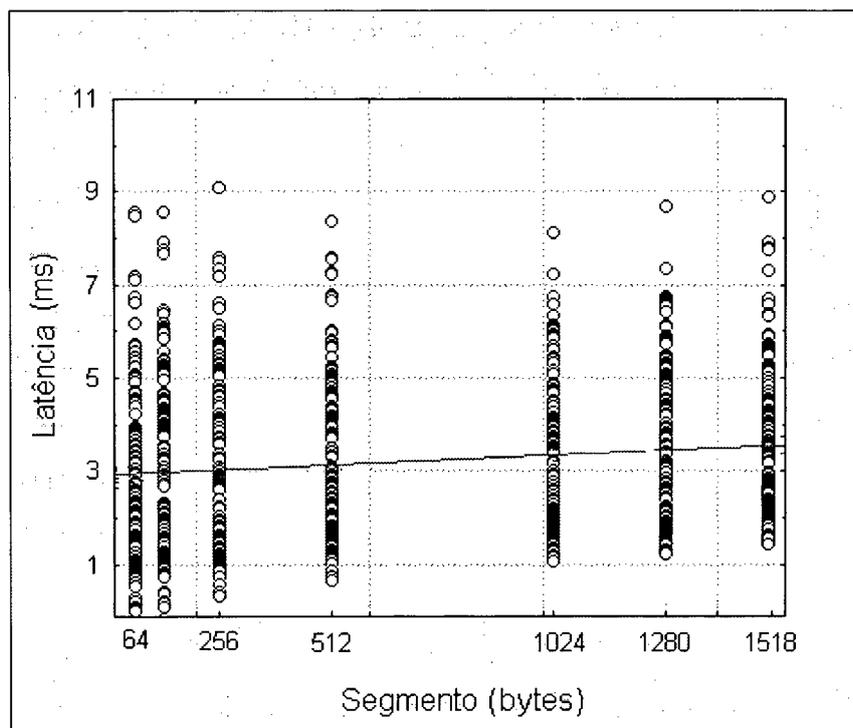


Figura 4.12 Gráfico da latência (ms) pelo segmento UDP (bytes) no Cluster Linux e a equação de regressão linear. Idem Figura 4.10.

No gráfico da Figura 4.12 é confirmado o fato de que no barramento Ethernet a variabilidade da latência não depende apenas do tamanho do segmento, como também das colisões que acontecem no acesso ao meio.

4.4 COMPARAÇÃO DOS RESULTADOS DAS LATÊNCIAS

Na Tabela 4.1 é apresentada a latência de cada *cluster* em relação ao tipo de protocolo, para o respectivo tamanho do segmento. É interessante observar o coeficiente de determinação para os *clusters* respectivos.

Tabela 4.1 - Latência versus Tamanho de segmentos.

Segmento/Protocolo	Cluster					
	Linux		Aix		SP 2	
	TCP	UDP	TCP	UDP	TCP	
64	0.295339	0.964967	0.935575	1.778718	3.00348	
128	0.316843	1.004327	0.976279	1.822046	3.029656	
256	0.359851	1.083047	1.057687	1.908702	3.082008	
512	0.445867	1.240487	1.220503	2.082014	3.186712	
1.024	0.617899	1.555367	1.546135	2.428638	3.39612	
1.280	0.703915	1.712807	1.708951	2.60195	3.500824	
1.580	0.783883	1.859177	1.860319	2.763076	3.598166	
R ²	0,913747	0,83279	0,8534197	0,0158737	0,231976	

Segundo a Tabela 4.1, o *cluster* Linux possui a menor latência entre os *clusters*; após, o *cluster* Aix; e em seguida o SP 2, para todos os tamanhos de segmento em análise.

Os baixos valores de latência do *cluster* Linux são devidos à combinação das características do sistema operacional com a rede Fast-Ethernet (sem colisão),

* Tempo previsto através da equação (3).

garantindo menor tempo de tráfego do segmento na rede, confirmando as recomendações feitas por [DIE98]. Um aspecto a ser considerado é o tamanho do segmento, já que para valores maiores de 1.500 bytes é necessário fazer a fragmentação do segmento, o que afetaria diretamente a latência.

Mesmo que o *cluster* Aix utilize as vantagens da tecnologia ATM (155 Mbit/s) superior à velocidade da rede Fast-Ethernet (100 Mbit/s) do *cluster* Linux, o *cluster* Aix possui uma latência maior, isso é devido ao tempo de encapsulamento da mensagem e da negociação existente utilizando o serviço LANE na rede ATM, o que afeta a latência para tamanhos de segmentos entre 64 e 1.580 bytes.

O baixo valor do coeficiente de determinação entre a variabilidade da latência e o tamanho do segmento para o SP 2 indica que existe um outro fator não considerado na medição. Esse fator é a taxa de colisão que ocorre no barramento Ethernet. Deve-se ressaltar aqui, que o teste realizado no SP 2 ocorreu utilizando a interface de rede Ethernet, e os valores serão melhores se for utilizada a interface de rede do HPS interno do SP2. O HPS não foi utilizado para a realização dos testes, pois na época em que eles foram iniciados não estava ativado, e é indicado nesta pesquisa para trabalhos futuros.

5 CONCLUSÃO

Ambientes de Processamento de Alto Desempenho são formados por um contexto de aplicações, estações de trabalhos, sistemas operacionais, interfaces de redes e equipamentos de interconexão que são utilizados para resolver grandes problemas computacionais no menor período de tempo possível. O fator essencial para esse ambiente em relação à rede é o tempo de tráfego de um segmento.

Este trabalho apresenta uma pesquisa sobre o tempo em que um segmento trafega de um ponto a outro, fim-a-fim, em função do seu tamanho, para a avaliação de alguns ambientes de Processamento de Alto Desempenho (*cluster* Aix, *cluster* Linux e o SP 2).

Após a análise dos resultados obtidos, pode-se afirmar que, para o *cluster*:

- Aix foi realizada uma aproximação ao modelo linear entre o tamanho do segmento e a latência da rede, tanto para o protocolo TCP como para o UDP. Considerando esse modelo, foi demonstrado que a variabilidade da latência da rede pode ser explicada em 85% e 83% pelos diferentes tamanhos do segmento TCP e UDP, respectivamente;
- Linux, também foi realizada uma aproximação ao modelo linear entre o tamanho do segmento e a latência da rede. Neste caso só foi analisado o protocolo TCP. Considerando esse modelo, foi demonstrando que a variabilidade da latência da rede pode ser explicada em 91% pelo tamanho do segmento; e
- SP 2 teve um comportamento irregular da variabilidade da latência quanto aos diferentes tamanhos de segmentos que foram usados, tanto para segmentos TCP quanto para UDP. Neste caso, a variabilidade da latência pode ser explicada em 23% e 1,5% pelo tamanho do segmento TCP e UDP, respectivamente, demonstrando o efeito de outras causas no tempo do tráfego, neste caso, o efeito das colisões que acontecem devido ao barramento Ethernet.

Por meio da análise de regressão foram obtidas as expressões matemáticas para relacionar a latência com o tamanho do segmento. Ficou determinado para o *cluster*:

- Aix, as expressões (1) e (2), ver itens 4.1.1 e 4.1.2, com um alto valor do coeficiente de determinação para TCP e UDP, respectivamente;
- Linux, a expressão (3), ver item 4.2.1, também com um valor do coeficiente de determinação muito alto para o TCP; e
- SP 2, as expressões (4) e (5), ver itens 4.3.1 e 4.3.2, mas o valor do coeficiente de determinação é muito baixo e não são consideradas representativas, devido às colisões que acontecem no barramento.

Um alto valor do coeficiente de determinação permite afirmar que o modelo matemático é adequado.

Essas expressões matemáticas para determinar a latência poderão ser usadas pelas outras técnicas para avaliação de desempenho de sistemas computacionais, simulação ou modelo analítico, para avaliação desses ambientes.

Para o modelo analítico, um exemplo pode ser encontrado em [JOH94], que utiliza como uma das variáveis o tempo de comunicação na rede. Esta aplicação consiste de um modelo de previsão de tempo numérica com a utilização do modelo analítico fornecido por [JOH94], em um caso padrão definido por [SAT98], e os valores de latência medidos nesta dissertação, [HAA00a] obteve valores analíticos do tempo de comunicação de 186 segundos para o Cluster Linux com 04 PCs. Este valor aproxima-se dos 190 segundos (10% do tempo total de processamento) estimados por [HAA00a] com base na porcentagem de utilização da CPU.

A partir das análises feitas nos *clusters*, fica demonstrado o alto grau de relação que existe entre o tamanho do segmento e a latência, para ambientes que não possuem colisões na rede.

No sistema operacional Linux, foi encontrado um erro de implementação no algoritmo Nagle do Protocolo TCP, o qual provoca altos valores de latência para segmento de 1.580 bytes. É disponibilizado pelos fornecedores do Linux um programa para a correção do problema.

Com relação à comparação feita entre os diferentes *clusters*, aquele que teve menor latência no tráfego dos segmentos de modo geral foi o *cluster* Linux, isso ocorre devido à combinação do sistema operacional Linux e a rede Fast-Ethernet.

Apesar da rede ATM possuir maior vazão, ela perde muito tempo em encapsulamentos, negociações etc., motivo pelo qual o *cluster* Aix obteve valores de latência maiores, para esses tamanhos de segmentos, do que o *cluster* Linux.

Recomendações:

Para se obter valores de latência entre 0,2953 e 0,7838 milisegundos com segmentos de tamanho entre 64 e 1.518 bytes, com o protocolo TCP deve-se utilizar o sistema operacional Linux, com barramento Fast Ethernet, usando um *switch* (com o objetivo de evitar colisões).

Assim sendo, não é recomendada a utilização do barramento de rede que permite a colisão em ambiente de Processamento de Alto Desempenho.

5.1 TRABALHOS FUTUROS

Como perspectivas de trabalhos futuros sugere-se as seguintes atividades:

Verificar o comportamento da latência dos *clusters* para tamanhos de segmentos maiores que 1.500 bytes em redes Fast-Ethernet e ATM, para verificar o desempenho da fragmentação da mensagem, como também alterar o tamanho do MTU.

Verificar a eficiência do protocolo TCP no *cluster* Linux após as correções do Kernel 2.12.

Realizar a regressão múltipla nos dados coletados para definir uma única equação de latência para os *clusters* Aix e Linux para segmentos de 64 a 1.280 bytes.

Determinar o efeito das colisões na variabilidade da latência para as redes com barramento Ethernet.

Verificar a latência do SP 2, utilizando a interface do HPS entre as estações (nós) em vez de interface Ethernet com colisão.

Especificar os detalhes do método usado neste trabalho para a medição da latência, fim-a-fim, em redes de computadores sem colisão.

Definir expressões matemáticas para determinar a latência em relação ao tamanho do segmento em outros ambientes de PAD, para que possam ser usadas pelas outras técnicas de avaliação do desempenho de sistemas computacionais.

Por intermédio das equações da latência obtidas, avaliar o tempo de latência para as técnicas de gerenciamento de redes (SNMP, Agentes Móveis, Agentes Autônomos etc.), em um ambiente de PAD.

Em paralelo as atividades apresentadas nesta dissertação de mestrado também foram desenvolvidos trabalhos relacionados que podem ser observados nos anexos 1 e 2. Estes trabalhos são:

SALES, André Barros, SARI, Solange, WESTPHALL, Carlos Becker. Análise da Eficiência de Técnicas de Gerenciamento de Rede em Um Ambiente de Processamento Paralelo e Distribuído. II Simpósio de Informática do Planalto Médio. Passo Fundo/RS, Mai/2000.

SALES, André Barros, SARI, Solange, WESTPHALL, Carlos Becker. Eficiência do Gerenciamento Descentralizado com Agentes Autônomos no Gerenciamento de Redes. XVIII Simpósio Brasileiro de Telecomunicações. Gramado (RS), 03 a 06 de setembro de 2000.

REFERÊNCIAS BIBLIOGRÁFICAS

- [BAR98] BARBETTA, Pedro Alberto, Estatística aplicada às Ciências Sociais, 2 ed. – Florianópolis: UFSC, 1998.
- [BER96] BERRY, Noémi, *IP Switching*, NASA/Ames Research Center, Moffett Field, CA, May 21, 1996. Url: http://www.nas.nasa.gov/Groups/LAN/nren/s_pres.html, página visitada em 17 de novembro de 1999.
- [BRA91] BRADNER, S., *Benchmarking Terminology for Network Interconnection Devices*, Network Working Group, Request for Comments: 1242, July 1991. Url: <http://WWW.ietf.org/rfc/rfc1242.txt?number=1242>, página visitada em 3 abril de 2000.
- [BRA99] BRADNER, S., *Benchmarking Methodology for Network Interconnect Devices*, Network Working Group, Request for Comments: 2544, March 1999. Url: <http://WWW.ietf.org/rfc/rfc2544.txt?number=2544>, página visitada em 30 abril de 2000.
- [BRI96] BRIGDGES, P. ET ALLI. Users Guide do mpich, *A Portable Implementation of MPI 1.2.0*, 1996
- [CAS97] CASTRO, Maria Cláudia de Almeida, Avaliação e Melhoria da Segurança Dinâmica utilizando Computação de Alto Desempenho, Tese de Dissertação de Mestrado, Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina – UFSC, Brasil, 1997.
- [CHO90] CHOROFAS, Dimitris N., STEINAMANN, Heinrich, SUPERCOMPUTÇÃO, McGraw-Hill, 1990.

- [COS99] COSTA, Taís Freire da Silva, *Avaliação Analítica de Agentes Móveis na Gerência de Redes*, Tese de Dissertação de Mestrado em Ciência da Computação – Departamento de Informática e Estatística – Universidade Federal de Santa Catarina – UFSC, Brasil, 1999.
- [DIE98] DIETZ, Hank, *Linux Parallel Processing HOWTO*, janeiro 1998. Url <http://www.how2linux.com/>, página visitada em 15 de junho de 2000.
- [DIS00] DISMAN Working Group, W. Kenneth, *Definitions of Managed Objects for Remote Ping, Traceroute, and Lookup Operations*, Internet-Draft March 2000, Expiration Date: September 2000. Url: <http://www.ietf.org/internet-drafts/draft-ietf-disman-remops-mib-08.txt>, página visitada em 30 abril de 2000.
- [FAR99] FARAZDEL, Abbas, ARCHONDO-CALLAO, Gonzalo R., HOCKS, Eva, SAKACHI, Takaaki, VAGNINI, Federico, *Understanding and Using the SP Switch*, *redbook da IBM*, 1999. Url: <http://www.redbooks.ibm.com/pubs/pdfs/redbooks /sg245161.pdf>, página visitada em 11 de outubro de 1999.
- [FRA96] FRANCESCHI, Analúcia Schiaffino Morales de, *Aplicação de Desempenho para Validar a Gerência Pró-ativa de Redes*, Tese de Dissertação de Mestrado em Ciência da Computação – Departamento de Informática e Estatística – Universidade Federal de Santa Catarina – UFSC, Brasil, 1996.
- [HAL99] HALLOY, C., YARKHAN, A., *Current Trends in Supercomputers and Scientific Computing*, November 1997. Url: <http://www-jics.cs.utk.edu/MMC/sld001.htm>, página visitada em 31 de outubro de 1999.

- [HAA00] HAAS, Reinaldo, AMBRIZZI, Tércio, FILHO, Augusto José Pereira, Comparação de Desempenho entre um Cluster PC-Linux e um SP 2 em simulações com o Modelo ARPS, XI Congresso Brasileiro de Meteorologia, Centro Cultural da UERJ - RJ, Outubro de 2000.
- [HAA00a] HAAS, Reinaldo, Comunicação Pessoal.
- [JAI91] JAIN, Raj. “*The Art of Computer Systems Performance Analysis*”. EUA : John Wiley E Sons, 1991.
- [JOH94] JOHNSON, Kenneth W. BAUER, Jeff, RICCARDI, Gregory A., XUE, Ming, DROEGEMEEIER, Kelvin K., *Distributed Processing of a Regional Prediction Model*, Monthly Weather Review, 122: 2558-2572, 1994.
- [LEO00] LEOCÁDIO, Márcio Augusto Parente, RODRIGUES, Paulo Henrique de Aguiar, Uma Ferramenta para Geração de Tráfego e Medição em Ambiente de Alto Desempenho, SBRC2000, Brasil, maio 2000, pp. 321-336.
- [MAC97] MACHADO, Francis Berenger, Maia, Luiz Paulo, Arquitetura de Sistemas Operacionais, Rio de Janeiro : LTC Livros Técnicos e Científicos Editora S.A., 1997.
- [MAD00] MADUREIRA, Francisco, Computadores utilizados em alguns Centro de Estudos de Genomas”. **Folha de São Paulo**, 5 julho de 2000. Caderno de Informática, p. F8.
- [MAR95] MARK, J., Introductio to HPC, abril de 1995. Url: <http://www.hpcc.ecs.soton.ac.uk/EandT/courseware/IntroHPC/index.html>, página visitada em 23 de outubro de 1999.

- [MAR98] MARK, J., D. NICOLE, *An Overview of HPC*, January 1998. Url: <http://www.hpcc.ecs.soton.ac.uk/EandT/courseware/HPCoverview/contents.html>, página visitada em 23 de outubro de 1999.
- [MEL00] MELO, Edson, SARI, Solange, SIQUEIRA, Walter, Avaliação de Desempenho do Serviço LANE sobre ATM, II Workshop RNP2, Anais 2000. Minas Gerais: 23 e 24 de maio de 2000, pp. 82-93.
- [PAR98] PARKER, S., SCHMECHEL, C., Network Working Group, *Some Testing Tools for TCP Implementors*, Request for Comments: 2398, August 1998. Url: <http://WWW.ietf.org/rfc/rfc2398.txt?number=2398>, página visitada em 23 abril de 2000.
- [REA99] REA, Alan, *An Introduction to Parallel Computing*, Versão 2.2, agosto de 1995. Url: <http://www.pcc.qub.ac.uk/tec/courses/intro/ohp/intro22.doc1.html> , página visitada em 30 de outubro de 1999.
- [RUB99a] RUBINSTEIN, M. G. and Duarte, O. C. M. B, *Evaluating Tradeoffs of Mobile Agents in Network Management*, Networking and Information Systems Journal, Hermes Science Publications, vol. 2, no. 2, pp. 237-252, 1999.
- [RUB99b] RUBINSTEIN, M. G. and Duarte, O. C. M. B. – “Análise da Eficiência de Agentes Móveis no Gerenciamento de Redes”, XIX Congresso Nacional da Sociedade Brasileira de Computação, XXVI Seminário Integrado de Software e Hardware - SEMISH'99, vol. 1, pp. 167-178, Rio de Janeiro, RJ, Brasil, julho 1999.

- [RUB99c] RUBINSTEIN, M. G. and Duarte, O. C. M. B., *Evaluating the performance of mobile agents in network management*. Relatório Técnico, Grupo de Teleinformática e Automação COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil, março de 1999.
- [SAL99] SALES, André Barros de, WESTPHALL, Carlos B., FILHO, Paulo José de Freitas, Modelando e Simulando uma Intranet usando Arena, Livro de Resumos da VII Semana de Pesquisa da UFSC. Santa Catarina: 20 a 23 de setembro 1999, pp. 140.
- [SAL00a] SALES, André Barros, SARI, Solange, WESTPHALL, Carlos Becker, Análise da Eficiência de Técnicas de Gerenciamento de Rede em Um Ambiente de Processamento Paralelo e Distribuído. II Simpósio de Informática do Planalto Médio. Passo Fundo/RS, Mai/2000.
- [SAL00b] SALES, André Barros, SARI, Solange, WESTPHALL, Carlos Becker, Avaliação de Desempenho de Agentes Autônomos no Gerenciamento de Redes de Computadores. V Workshop de TMN, Belo Horizonte/MG, Mai/2000, pg.67.
- [SAL00c] SALES, André Barros de Sales, WESTPHALL, Avaliação de Técnicas de Gerenciamento de Redes de Computados, I Workshop da Pós-Graduação em Ciência da Computação, Cadernos de Resumos. Universidade Federal de Santa Catarina. Santa Catarina: 16 e 17 de maio de 2000, pp. 41.
- [SHU99] SHURAN Software Ltd, *IP Product Testing*, 1999. Url: http://www.tcpip.com/support/w_papers.htm, página visitada em 12 de julho de 2000.
- [TAN96] TANENBAUM, Andrew S. Redes de Computadores, tradução [ds 3. Ed. Original] Insight Serviços de Informática. Rio de Janeiro: Campus, 1997.

[VIA98] VIANNA, Filipe, FINK, Daniel, INÁCIO, José, *A General Model for a Digital System*, março de 1998. Url: <http://www.geocities.com/ResearchTriangle/Lab/6866/index.html>, página visitada em 30 de outubro de 1999.

[WEL96] WELSH, Matt, BASU, Anindya, EICKEN Thorsten von, *Low-Latency Communication over Fast Ethernet*, EuroPar '96, Lyon, France, august 1996.

[WEL97] WELSH, Matt, BASU, Anindya, EICKEN Thorsten von, *ATM and Fast Ethernet Network Interfaces for User-Level Communication*, Proceedings of High-Performance Computer Architecture 3, San Antonio, february 1997.

ANEXO 1**ANÁLISE DA EFICIÊNCIA DE TÉCNICAS DE GERENCIAMENTO DE REDES EM UM
AMBIENTE DE PROCESSAMENTO PARALELO E DISTRIBUÍDO**

André Barros de Sales*,**
andreb@lrg.ufsc.br

Solange Sari**
solange@npd.ufsc.br

Carlos B. Westphall*
westphal@lrg.ufsc.br

*Laboratório de Redes e Gerência
Departamento de Informática e Estatística

**Laboratório de Interoperabilidade
Núcleo de Processamento de Dados

Universidade Federal de Santa Catarina
Campus Universitário – Trindade
Caixa Postal 476 – CEP: 88010-970
Florianópolis – SC - Brasil

Resumo

Este artigo apresenta uma avaliação baseada no modelo analítico de algumas técnicas de gerenciamento de redes para um ambiente de processamento paralelo e distribuído. Cada técnica de gerenciamento (SNMP, Agentes Móveis e Agentes Autônomos) é representada por uma equação analítica considerando a largura de banda, latência, número de estações e características específicas da própria técnica. Os equipamentos desse ambiente estão interligados por uma rede heterogênea composta de ATM, Fast Ethernet e Ethernet. É apresentada uma comparação entre as técnicas, relacionando o número de estações com o tempo de resposta da coleta de informação dos elementos gerenciados. Embora questões de implementação e administração da rede sejam consideradas, esta avaliação analítica contribui para a escolha adequada da técnica de gerenciamento.

Abstract

This article presents an evaluation based on the analytical model of some network management techniques for a parallel and distributed processing environment. Each management technique (SNMP, Mobile Agents and Autonomous Agents) is represented by an analytical equation which considers bandwidth, latency, number of stations and specific characteristics of the own technique. The equipments of this environment are

interconnected with an heterogeneous network composed by ATM, Fast Ethernet and Ethernet. A comparison of techniques relating the number of stations with the response time of information collection of managed elements is presented. Though considering implementation and network management questions this analytical evaluation contributes to the adequate choice of the management technique.

1 Introdução

1.1 Visão Geral

A Computação de Alto Desempenho (HPC – *High Performance Computing*) pode ser definida como os recursos computacionais mais potentes disponíveis em cada época [CAS97] isto inclui computadores, redes, algoritmos e os ambientes necessários para tornar tais sistemas disponíveis. Estes sistemas abrangem desde um cluster de estações de trabalho até os grandes supercomputadores.

As Redes de Computadores, seus recursos e aplicações, por mais simples que sejam, tornaram-se indispensáveis às operações das organizações. Eventuais problemas no seu funcionamento podem implicar em uma degradação inaceitável de desempenho ou mesmo parada parcial ou total, da rede. Uma rede gerenciada oferece uma garantia maior aos usuários da disponibilidade de seus serviços em um nível de desempenho aceitável, além de oferecer recursos para um gerenciamento pró-ativo.

O Gerenciamento de Redes é composto por elementos ou entidades de gerenciamento, um ou mais objetos gerenciados (agente), ao menos uma estação de gerenciamento da rede (gerente), um protocolo de gerenciamento e informações de gerenciamento [ROS94].

As técnicas de gerenciamento de redes em investigação neste trabalho apresentam os seguintes modelos:

- SNMP (*Simple Network Management Protocol*) - baseia-se no modelo cliente-servidor, e todas as informações relacionadas ao gerenciamento são processadas na estação de gerenciamento, conforme [ROS94];

- gerenciamento com Agentes Móveis - o processamento e o controle da estação de gerenciamento são descentralizados e executados nos elementos gerenciados, conforme [RUB99b];

- gerenciamento com Agentes Autônomos - oferece uma gerência distribuída no gerenciamento da rede, onde existe uma estação de gerenciamento principal e outras estações de gerenciamento na rede, conforme o modelo apresentado por Koch [KOC97]; e

- gerenciamento com Agentes Autônomos como Agentes Adaptativos – oferece também uma gerência distribuída no gerenciamento da rede, onde todos os elementos gerenciados desejados no gerenciamento de redes possuem um Agente Adaptativo, conforme o modelo apresentado por Vieira [VIE97].

1.2 Objetivos e Metas

O objetivo principal desse trabalho é avaliar as técnicas de gerenciamento de redes, na rede cluster, de processamento paralelo e distribuído. A fim de atingir este objetivo são estabelecidas algumas metas, as quais direcionaram o trabalho de pesquisa e avaliação dos resultados:

- definir expressões analíticas das técnicas de gerenciamento de rede;
- selecionar os objetos gerenciáveis na rede cluster;
- definir a variável da MIB (Management Information Base) para avaliação;
- realizar uma avaliação baseada no modelo analítico das técnicas de gerenciamento sobre uma rede de processamento paralelo e distribuído;
- avaliar os resultados; e
- inferir informações.

1.3 Ambiente de Avaliação

O ambiente de pesquisa está baseado na sub-rede cluster da rede UFSC da Universidade Federal de Santa Catarina. O cluster é composto por um conjunto de estações de trabalho interligadas por uma rede de alta velocidade e que possuem uma base comum de usuários, limitando o nosso contexto de trabalho em:

- 8 estações IBM 43P;
- 6 Pentium;
- 1 IBM SP2 com 9 nós; e
- 7 *Switches* sendo:
 - 4 ATM IBM 8265;
 - 1 CoreBuilder 7000 da 3COM;
 - 1 IBM 8271 modelo 712; e
 - 1 HPS (o *switch* interno do SP2).

Os equipamentos estão interligados por uma rede ATM de 155 e 622 Mbps, Fast Ethernet de 100 Mbps e Ethernet de 10 Mbps, como apresentado na Figura 1.

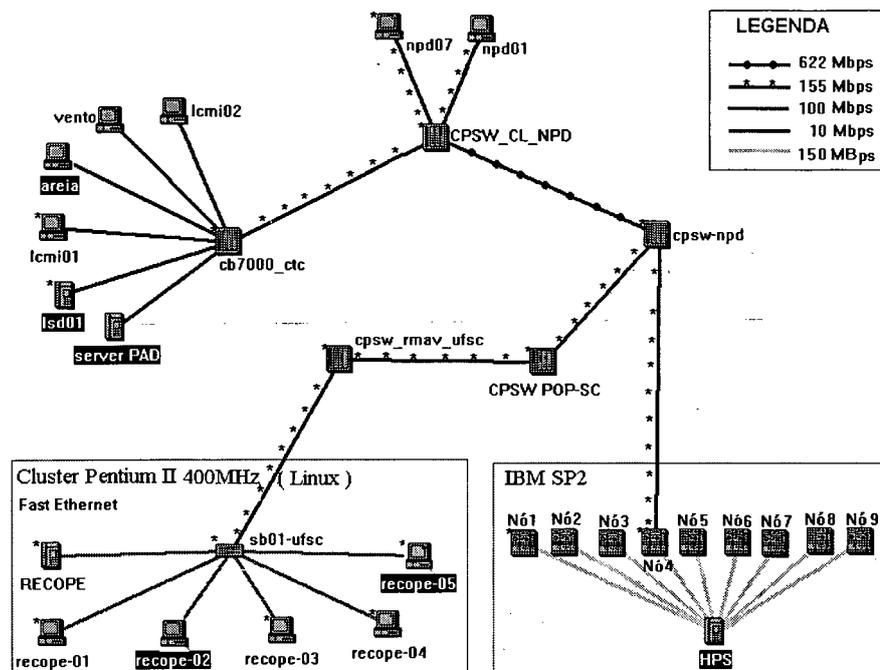


Figura 1- Topologia da Rede

A Figura 1 apresenta a topologia onde será realizada o experimento de investigação das técnicas de gerenciamento de rede.

1.4 Apresentação do Trabalho

Nesta primeira seção está descrita a problemática abordada, bem como os elementos de resolução. Na seção 2 estão descritas as equações do Modelo Analítico das Técnicas de Gerenciamento, sendo subdividida em Agentes Autônomos, Agentes Adaptativos, SNMP e Agentes Móveis. A seção 3 descreve o Ambiente para Análise. A seção 4 apresenta a Comparação das Técnicas de Gerenciamento, onde é apresentado um gráfico obtido através do experimento comparando as técnicas de gerenciamento para o ambiente de processamento paralelo e distribuído. A seção 5 apresenta as conclusões desse trabalho.

2 Modelo Analítico das Técnicas de Gerenciamento

As variáveis utilizadas nas expressões analíticas são apresentadas na Tabela 1.

Tabela 1 - Variáveis utilizadas no modelo analítica

Variável	Unidade de medida	Descrição
L_i	Segundos	Latência da rede i
B_i	Bps	Largura de banda da rede i
K	Bits	Tamanho do código do agente
P	Bits	Tamanho do pedido

R	Bits	Tamanho da resposta
n_u	Inteiro	Número total de elementos gerenciados na rede u
Q	Inteiro	Número total de sub-redes
Rlog	Inteiro	Tamanho da resposta de um Agente Coletor de Dados

Um dos critérios para avaliação das técnicas de gerenciamento é o tempo de resposta resultante da coleta de informação da MIB. Na Tabela 1 apresenta as variáveis utilizadas neste modelo analítico e suas respectivas unidades de medida.

2.1 Agentes Autônomos

Koch [KOC97] apresenta o seguinte conceito “Agente Autônomo é um sistema computacional que, baseado em um conjunto de objetivos ou comportamento determinado, extrai informações do ambiente e conclui ações ou reações de interação com o ambiente ou com outros agentes a fim de obter cooperação”.

Com a utilização de Agentes Autônomos no gerenciamento de rede consegue-se montar uma estrutura semelhante a um gerenciamento distribuído na rede. A rede possui uma estação de gerenciamento e definem-se outras estações, com Agentes Autônomos, que executem tarefas de gerenciamento, incluindo a coleta de dados.

O tempo de resposta gasto na consulta às MIBs dos diversos recursos distribuídos por um ou mais segmentos deve considerar como o envio do pedido (P) da estação de gerenciamento até os Agentes Coletores de Dados, Agentes Autônomos, que responderão com as informações de seus elementos gerenciados contidas em um arquivo de *log* ($Rlog$). São considerados, também, o tamanho do pedido (P), a latência da rede (L), a largura de banda (B) e os elementos gerenciados (n). Esse trabalho propõe a seguinte equação para calcular o *taa* (tempo de coleta de dados):

$$taa = \sum_{z=1}^q n_z \left(\left(\sum_{i=0}^z (2L_i) \right) + \left(\sum_{i=0}^z \frac{P + R \log}{B_i} \right) \right) \quad [1]$$

Caso os Agentes Coletores de Dados não possuam as informações de seus elementos gerenciados, então para calcular o *taa*, equação [1], deve-se calcular o tempo *taal*, equação [2], tempo para o Agente Coletor de Dados obter as informações de seus elementos gerenciados, para cada Rlog:

$$taal = \sum_{z=1}^q n_z \left(\left(\sum_{i=0}^z (2L_i) \right) + \left(\sum_{i=0}^z \frac{P + R}{B_i} \right) \right) \quad [2]$$

O tempo *taal* é o tempo de coleta de dados pelos Agentes Coletores de Dados aos seus elementos gerenciados, que após realizar a coleta irá enviar o arquivo de *log* ($Rlog$) à estação de gerenciamento.

2.2 Agentes Adaptativos

Em Vieira [VIE97] são apresentados os Agentes Adaptativos que sendo treinados, podem ser autônomos na gerência. A inclusão desses Agentes Adaptativos

aos agentes dos elementos gerenciados desejados no gerenciamento de redes pode ser através do protocolo SMUX, definido pela RFC1227.

Pode-se treinar esses agentes para enviarem a informação da variável desejada para a estação de gerenciamento podendo, assim, reduzir o tempo de desempenho.

Considera-se que todos os elementos da rede gerenciada, dos quais deseja-se coletar as informações, possuem um Agente Adaptativo para enviar à estação de gerenciamento a variável desejada e este envio seja sincronizado. Esse trabalho propõe a seguinte equação para calcular o tempo *taad*, tempo gasto para obter a informação dos elementos gerenciados no gerenciamento por Agentes Adaptativos, como:

$$taad = \sum_{z=1}^q n_z \left(\left(\sum_{i=0}^z (L_i) \right) + \left(\sum_{i=0}^z \frac{R}{B_i} \right) \right) \quad [3]$$

onde o cálculo do *taad* considera o tamanho da resposta (*R*), a latência da rede (*L*), a largura de banda (*B*) e os elementos gerenciados (*n*).

2.3 SNMP

O SNMP é um protocolo desenvolvido para facilitar a troca de informações de gerenciamento entre dispositivos de rede. Essas informações transportadas pelo SNMP, como número de pacotes por segundo, permitem aos administradores da rede gerenciar o desempenho da rede de forma remota, encontrando e solucionando os problemas da rede.

No gerenciamento SNMP, o gerente solicita informações aos agentes e também solicita aos agentes que executem ações corretivas sobre os elementos gerenciados; os agentes coletam as informações armazenadas na MIB e as envia para a estação de gerenciamento quando solicitadas pelo gerente, além disso, o agente executa as ações solicitadas pelo gerente.

Costa [COS99] apresenta a expressão analítica para calcular o *tsnmp*:

$$tsnmp = \sum_{z=1}^q n_z \left(\left(\sum_{i=0}^z (2L_i) \right) + \left(\sum_{i=0}^z \frac{P+R}{B_i} \right) \right) \quad [4]$$

A equação [4] apresenta o, *tsnmp*, tempo de resposta resultante das consultas às MIBs, dos diversos recursos distribuídos por um ou mais segmentos através do protocolo SNMP pela estação de gerenciamento. A equação *tsnmp* considera a latência da rede (*L*), a largura de banda (*B*), os elementos gerenciados (*n*), o tamanho do pedido (*P*) e da resposta (*R*).

2.4 Agentes Móveis

Agentes Móveis são programas criados para executar suas tarefas em diferentes computadores da rede e possuem a capacidade de mover-se pela rede, transportando-se entre os equipamentos, levando consigo dados e códigos.

Em Costa [COS99] é apresentada a expressão para calcular o *tAg_Movel*, equação [5], tempo de resposta resultante das consultas às MIBs dos diversos recursos distribuídos por um ou mais segmentos através de um agente móvel. A estação de gerenciamento enviará o agente móvel na rede, ele coletará as informações de todos os

elementos gerenciados especificados pela estação de gerenciamento, e depois voltará a estação de gerenciamento. Onde o cálculo do tAg_Movel considera a latência da rede (L), o tamanho do agente móvel (k), o tamanho da resposta (R), a largura de banda (B) e os elementos gerenciados (n).

$$tAg_Movel = \sum_{j=1}^q \left(\sum_{i=j-1}^j L_i + \sum_{i=j-1}^j \frac{k + R \left(\sum_{u=1}^{j-1} n_u \right)}{B_i} + \sum_{i=2}^{n_j} \frac{k + R \left(\sum_{u=1}^{j-1} n_u \right) + (i-1)R}{B_j} + (n_j - 1)L_j \right) + t_Volta \quad [5]$$

O tempo de t_Volta da equação [6] é calculado como:

$$t_Volta = \left(\sum_{i=0}^q L_i + \sum_{i=0}^q \frac{k + R \left(\sum_{u=1}^q n_u \right)}{B_i} \right) \quad [6]$$

3 Ambiente para Análise

3.1 Caracterização do Ambiente

A comparação de performance entre as técnicas de gerenciamento foi realizada na topologia apresentada na Figura 1. Para manter homogeneidade com os trabalhos de [COS99], [RUB99a] e [RUB99b] será considerado que os enlaces e os nós não possuem carga, não há perdas nos enlaces e não é considerado o tempo de processamento na camada de aplicação.

O valor da variável de latência será considerado à latência do *Switch*, 65 microsegundos para ATM e 60 microsegundos para *Fast Ethernet*, como apresentado em [WEL98]; a latência do HPS do IBM SP2 é de 1,2 microsegundos conforme [FAR99]; e sendo considerada apenas uma latência de 1 milisegundos para a rede Ethernet [COS99], [RUB99a].

A largura de banda será de 155 ou 622 Mbps para a rede ATM, 100 Mbps para a rede Fast Ethernet e 10 Mbps para a rede Ethernet. Para o HPS do IBM SP2 será de 150MBps conforme [FAR99].

Para analisar o desempenho da atividade de gerência das técnicas de gerenciamento, utilizará a variável *ifInErrors* da interface de cada estação da rede, sendo que o tamanho do pedido é de 71 octetos e a resposta é de 76 octetos. A MTU (Unidade Máxima de Transmissão) usada será de 1500 octetos para não haver fragmentação de mensagens SNMP por estas serem pequenas, conforme [RUB99b]. O tamanho do agente móvel será considerado de 5 Kbytes conforme [BAL98].

A estação de gerenciamento será considerada como a estação *npd07*, conforme Figura 1. Para o método de gerenciamento com Agentes Autônomos será considerado como Agentes Coletores de Dados as estações *npd01*, *nó 4* e *recope-01*, conforme Figura 1.

4 Comparação das Técnicas de Gerenciamento

Os resultados obtidos na comparação entre as técnicas de gerenciamento são apresentados na Figura 2, sendo considerada para os Agentes Autônomos o modo em que os Agentes Coletores de Dados não possuem as informações dos seus elementos gerenciados.

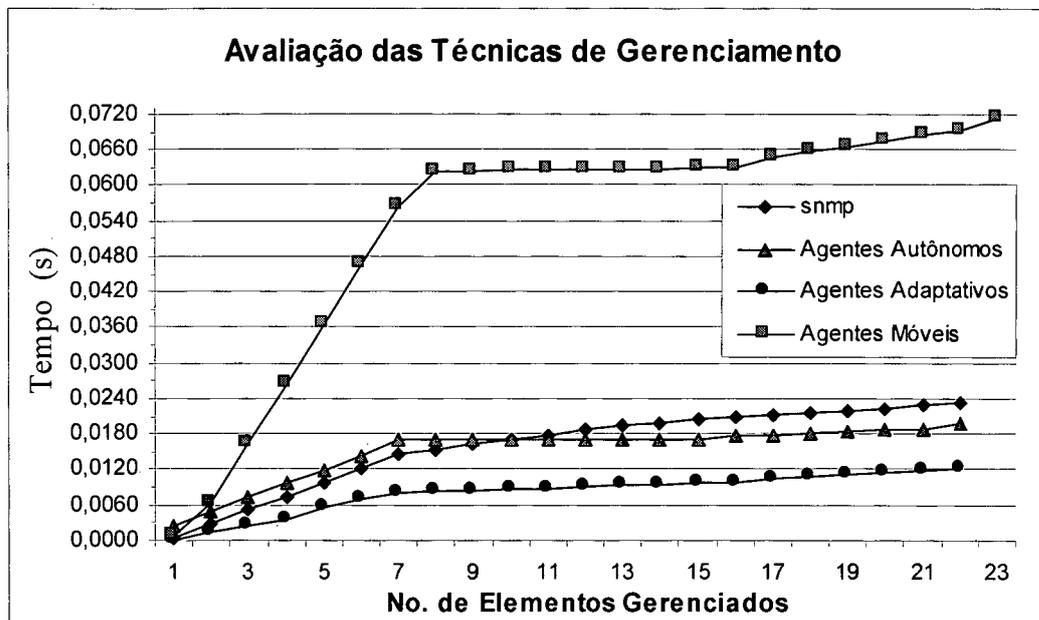


Figura 2– Comparação das técnicas de gerenciamento

O número de elementos gerenciados para a técnica de Agentes Móveis é maior em um elemento gerenciado, devido à volta do Agente Móvel para estação de gerenciamento.

Os primeiros elementos gerenciados estão na rede Ethernet, observa-se duas relações básicas: quanto maior a latência maior será o tempo de resposta, e quanto menor a largura de banda maior será o tempo de resposta. Estas relações estão mais acentuadas para os Agentes Móveis, devido ao seu tamanho em relação ao tamanho do pedido das outras técnicas.

Os Agentes Autônomos tem como tempo de resposta 17,09 milisegundos, sendo maior ao do SNMP com 17,07 milisegundos somente para os 10 primeiros elementos gerenciados. Mas para as largura de banda maiores e latência menores os Agentes Autônomos tornaram-se melhor que o SNMP obtendo o resultado de 19,84 milisegundos para os Agentes Autônomos e 23,23 milisegundos para o SNMP. Não se pode deixar de atentar o posicionamento dos Agentes Coletores de Dados na rede, eles encontram-se apenas nas redes ATM e *Fast Ethernet*.

O método de Agentes Adaptativos apresenta o tempo de resposta de 12,29 milisegundos obtendo o melhor tempo de resposta em relação as outras técnicas comparadas.

5 Conclusão

O gerenciamento SNMP, acrônimo de *Simple Network Management Protocol*, pelo que o seu próprio nome diz Protocolo Simples para Gerência de Redes, é devido aos agentes SNMP serem simples e executarem operações elementares. O SNMP possui uma gerência centralizada, o que pode gerar problemas de tráfego e congestionamento dos meios de transmissão. A estação de gerenciamento é responsável por receber todas as informações de gerenciamento, e é provável que os segmentos de rede aos quais a estação de gerenciamento está ligada sustente tráfegos altos estando sujeitos a colisões. A gerência centralizada pode tornar-se inviável para lidar com o crescente tamanho da rede, e contraditoriamente, contribuir para o seu mau funcionamento ao gerar maior quantidade de tráfego na rede.

O gerenciamento com Agentes Móveis descentraliza o processamento e o controle da estação de gerenciamento, tornando assíncrona a comunicação com a estação de gerenciamento, permitindo um balanceamento de carga de processamento. Mas os Agentes Móveis em comparação com as outras técnicas para executar a coleta de uma informação nos elementos gerenciados possui o maior tempo de resposta, principalmente para as menores larguras de banda.

Os Agentes Autônomos apresentam algumas vantagens como a diminuição do tráfego entre a estação de gerenciamento e os elementos gerenciados, uma vez que o processo de aquisição e análise das informações é levado para mais perto do local dos elementos gerenciados. Uma outra vantagem é maior agilidade na tomada de decisões, as quais podem ser tomados mais próximos dos elementos gerenciados, evitando-se a necessidade de comunicação com a estação de gerenciamento.

Os Agentes Adaptativos oferecem uma gerência automatizada e pró-ativa, possuem a vantagem de tomarem as decisões nos próprios elementos gerenciados, solucionando os problemas rapidamente e diminuindo o tráfego na rede devido ao seu gerenciamento distribuído. Possui a desvantagem de ter seu treinamento demorado, e seu uso é indicado onde a precisão das decisões não é um requisito forte. Devido ao seu grau de distribuição, em todos os elementos gerenciados, torna-se de difícil sua manutenção.

A avaliação das técnicas de gerência baseada no modelo analítico oferece um outro critério a ser considerado, além das características específicas de cada técnica. Em um ambiente de processamento paralelo e distribuído, onde é avaliado o desempenho de uma tarefa de gerenciamento de rede é comparado o desempenho das técnicas de gerenciamento SNMP, Agentes Móveis e Agentes Autônomos.

Os resultados indicam que os valores de largura de banda e latência refletem diretamente no desempenho na coleta das informações das técnicas avaliadas.

Através deste estudo conclui-se que analiticamente o melhor tempo de resposta da coleta para a variável *ifInErrors* são dos Agentes Adaptativos, com 12,29 milisegundos. Seguido dos Agentes Autônomos, com 19,84 milisegundos, SNMP, com 23,23 milisegundos, e Agentes Móveis, com 71,22 milisegundos, para esse ambiente e com essas características. A escolha da técnica deve também levar em consideração a sua implementação, tirando proveito da simplicidade de algumas técnicas. Além disso, a escolha deve atender aos objetivos estabelecidos pelo gerente humano da rede.

Como proposta para trabalhos futuros têm a investigação de outros critérios para suporte a escolha da técnica gerenciamento e medir o valor real da latência da rede para aplicar nas equações apresentadas.

6 Bibliográficas

- [BAL98] BALDI, M., e PICCO, G.P. Evaluating the tradeoffs of mobile code design paradigms in network management applications. In Proceedings of the 20th International Conference on Software Engineering (ICSE'97), Kyoto, Japan, April de 1998.
- [CAS97] CASTRO, Maria Cláudia de Almeida, Avaliação e Melhoria da Segurança Dinâmica utilizando Computação de Alto Desempenho, Tese de Dissertação de Mestrado, Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina – UFSC, Brasil, 1997.
- [COS99] COSTA, Taís Freire da Silva, Avaliação Analítica de Agentes Móveis na Gerência de Redes– Tese de Dissertação de Mestrado em Ciência da Computação – Departamento de Informática e Estatística – Universidade Federal de Santa Catarina – UFSC, Brasil, 1999.
- [FAR99] FARAZDEL, Abbas, ARCHONDO-CALLAO, Gonzalo R., HOCKS, Eva, SAKACHI, Takaaki, VAGNINI, Federico, Understanding and Using the SP Switch, <http://www.redbooks.ibm.com/pubs/pdfs/redbooks/sg245161.pdf>, redbook da IBM, 1999.
- [KOC97] KOCH, Fernando Luiz, Agentes Autônomos para Gerenciamento de Redes de Computadores, Tese de Dissertação de Mestrado em Ciência da Computação – Departamento de Informática e Estatística – Universidade Federal de Santa Catarina – UFSC, Brasil, 1997.
- [ROS94] ROSE, Marshall T., The Simple Book: an introduction to Internet management – 2. ed. New Jersey – USA: PTR Prentice – Hall, 1994.
- [RUB99a] - RUBINSTEIN, M. G. and Duarte, O. C. M. B., “Evaluating Tradeoffs of Mobile Agents in Network Management”, Networking and Information Systems Journal, Hermes Science Publications, vol. 2, no. 2, pp. 237-252, 1999.
- [RUB99b] RUBINSTEIN, M. G. and Duarte, O. C. M. B. – “Análise da Eficiência de Agentes Móveis no Gerenciamento de Redes”, XIX Congresso Nacional da Sociedade Brasileira de Computação, XXVI Seminário Integrado de Software e Hardware - SEMISH'99, vol. 1, pp. 167-178, Rio de Janeiro, RJ, Brasil, Julho 1999.
- [VIE97] VIEIRA, Elvis Melo, Método para Desenvolver Agentes Adaptativos em Gerência de Redes usando Redes Neurais, Tese de Dissertação de Mestrado em Ciência da Computação – Departamento de Informática e Estatística – Universidade Federal de Santa Catarina – UFSC, Brasil, 1997.
- [WEL98] WELSH, Matt, ANNOUNCEMENT: Release of U-Net 2.1 software distribution, University of California, Berkeley <http://www.cs.berkeley.edu/~mdw/projects/unet>, 1998.

ANEXO 2

SBrT 2000 – XVIII Simpósio Brasileiro de Telecomunicações, 3 a 6 de Setembro, 2000, Gramado – RS.

EFICIÊNCIA DO GERENCIAMENTO DESCENTRALIZADO COM AGENTES AUTÔNOMOS NO GERENCIAMENTO DE REDES

A. B. Sales, S. Sari**, C. B. Westphall**

*Laboratório de Redes e Gerência - Departamento de Informática e Estatística

**Laboratório de Interoperabilidade - Núcleo de Processamento de Dados

Universidade Federal de Santa Catarina

Florianópolis – SC, Brasil

RESUMO

Este artigo apresenta uma avaliação baseada no modelo analítico de duas técnicas de gerenciamento de redes para um ambiente de processamento paralelo e distribuído. As técnicas de gerenciamento SNMP e Agentes Autônomos são representadas por uma equação analítica considerando a largura de banda, latência, número de estações e características específicas da própria técnica. Os equipamentos desse ambiente estão interligados por uma rede heterogênea composta de ATM, Fast Ethernet e Ethernet. É apresentada uma comparação entre as técnicas, relacionando o número de estações com o tempo de resposta da coleta de informação dos elementos gerenciados.

ABSTRACT

This article presents an evaluation based on the analytical model of two network management techniques for a parallel and distributed processing environment. The management technique SNMP and Autonomous Agents are represented by an analytical equation which considers bandwidth, latency, number of stations and specific characteristics of the own technique. The equipments of this environment are interconnected with an heterogeneous network composed by ATM, Fast Ethernet and Ethernet. A comparison of techniques relating the number of stations with the response time of information collection of managed elements is presented.

1. Introdução

Computação de alto desempenho também pode ser definida como os recursos computacionais mais potentes disponíveis em cada época [CAS97] isto inclui computadores, redes, algoritmos e os ambientes necessários para tornar tais sistemas disponíveis. Estes sistemas abrangem desde um cluster de estações de trabalho até os grandes supercomputadores.

As Redes de Computadores, seus recursos e aplicações, por mais simples que sejam, tornaram-se indispensáveis às operações das organizações. Eventuais problemas no seu funcionamento podem implicar em uma degradação inaceitável de desempenho ou mesmo parada parcial ou total, da rede. Uma rede gerenciada oferece uma garantia maior aos usuários da disponibilidade de seus serviços em um nível de desempenho aceitável, além de oferecer recursos para um gerenciamento pró-ativo.

O objetivo principal desse trabalho é avaliar duas técnicas de gerenciamento de redes, para verificar qual é a mais apropriada para um ambiente de processamento paralelo e distribuído. As técnicas de gerenciamento de redes em investigação apresentam os seguintes modelos:

- SNMP (*Simple Network Management Protocol*) - baseia-se no modelo cliente-servidor, e todas as informações relacionadas ao gerenciamento são processadas na estação de gerenciamento, conforme [ROS94]; e

- o gerenciamento com Agentes Autônomos - oferece uma gerência distribuída no gerenciamento da rede, onde existe uma estação de gerenciamento principal e outras estações de gerenciamento na rede, conforme o modelo apresentado por Koch [KOC97].

Nesta primeira seção está descrita a problemática abordada, bem como os elementos de resolução. Na seção 2 estão descritas as equações do Modelo Analítico das Técnicas de Gerenciamento, sendo subdividida em SNMP e Agentes Autônomos. A seção 3 descreve o Ambiente para Análise. A seção 4 apresenta a Comparação das técnicas de gerenciamento. A seção 5 apresenta a Conclusão. Para finalizar a seção 6 apresenta as referências bibliográficas.

2. Modelo Analítico das Técnicas de Gerenciamento

As variáveis utilizadas nas expressões analíticas são apresentadas na Tabela 1.

Tabela 1 - Variáveis utilizadas para avaliação analítica

Variável	Unidade de medida	Descrição
L_i	Segundos	Latência da rede i
B_i	Bps	Largura de banda da rede i
P	Bits	Tamanho do pedido
R	Bits	Tamanho da resposta
N_u	Inteiro	Número total de elementos gerenciados na rede u
Q	Inteiro	Número total de sub-redes
$Rlog$	Inteiro	Tamanho da resposta de um Agente Coletor de Dados

2.1 SNMP

O SNMP é um protocolo desenvolvido para facilitar a troca de informações de gerenciamento entre dispositivos de rede. Essas informações transportadas pelo SNMP, como número de pacotes por segundo, permitem aos administradores da rede gerenciar o desempenho da rede de forma remota, encontrando e solucionando os problemas da rede.

Costa [COS99] apresenta a expressão analítica para calcular o $tsnmp$:

$$tsnmp = \sum_{z=1}^q n_z \left(\left(\sum_{i=0}^z (2L_i) \right) + \left(\sum_{i=0}^z \frac{P+R}{B_i} \right) \right) \quad [1]$$

A equação [1] apresenta o, $tsnmp$, tempo de resposta resultante das consultas às MIBs, dos diversos recursos distribuídos por um ou mais segmentos através do protocolo SNMP pela estação de gerenciamento. Onde a equação $tsnmp$ considera a latência da rede (L), a largura de banda (B), os elementos gerenciados (n), o tamanho do pedido (P) e da resposta (R).

2.2 Agentes Autônomos

Koch [KOC97] apresenta o seguinte conceito "Agente Autônomo é um sistema computacional que, baseado em um conjunto de objetivos ou comportamento determinado, extrai informações do ambiente e conclui ações ou reações de interação com o ambiente ou com outros agentes a fim de obter cooperação".

Com a utilização de Agentes Autônomos no gerenciamento de rede consegue-se montar uma estrutura semelhante a um gerenciamento distribuído na rede. A rede possui uma estação de gerenciamento e definem-se outras estações, com Agentes Autônomos, que executem tarefas de gerenciamento, incluindo a coleta de dados.

O tempo de resposta gasto na consulta às MIBs dos diversos recursos distribuídos por um ou mais segmentos deve considerar como o envio do pedido (P) da estação de gerenciamento até os Agentes Coletores de Dados, Agentes Autônomos, que responderão com as informações de seus elementos gerenciados contidas em um arquivo de \log ($Rlog$). São considerados, também, o tamanho do pedido (P), a latência da rede (L), a largura de banda (B) e os elementos gerenciados (n). Esse trabalho propõe a seguinte equação para calcular o taa (tempo de coleta de dados):

$$taa = \sum_{z=1}^q n_z \left(\left(\sum_{i=0}^z (2L_i) \right) + \left(\sum_{i=0}^z \frac{P+R \log}{B_i} \right) \right) \quad [2]$$

Caso os Agentes Coletores de Dados não possuam as informações de seus elementos gerenciados, então para calcular o taa , equação [2], deve-se calcular o tempo $taal$, equação [3], tempo para o Agente Coletor de Dados obter as informações de seus elementos gerenciados, para cada $Rlog$:

$$taal = \sum_{z=1}^q n_z \left(\left(\sum_{i=0}^z (2L_i) \right) + \left(\sum_{i=0}^z \frac{P+R}{B_i} \right) \right) \quad [3]$$

O tempo $taal$ é o tempo de coleta de dados pelos Agentes Coletores de Dados aos seus elementos gerenciados, que após realizar a coleta irá enviar o arquivo de \log ($Rlog$) à estação de gerenciamento.

3. Ambiente de Análise

O desempenho foi avaliado através do número de octetos transmitidos e recebidos pela estação de gerenciamento e do tempo de resposta na obtenção de uma variável da MIB do elemento gerenciado.

3.1 Infra-Estrutura da Rede

O ambiente de pesquisa está baseado na sub-rede cluster da rede UFSC da Universidade Federal de Santa Catarina. O cluster é composto por um conjunto de estações de trabalho, interligados por uma rede de alta velocidade e que possuem uma base comum de usuários, limitando o nosso contexto de trabalho em:

- 8 estações IBM 43P;

- 6 Pentium;

- 1 IBM SP2 com 9 nós;

- 7 Switches sendo: 4 ATM IBM 8265; 1

CoreBuilder 7000 da 3COM; 1 IBM 8271 modelo 712; e 1 HPS (o switch interno do SP2).

Os equipamentos estão interligados por uma rede ATM de 155 e 622 Mbps, Fast Ethernet de 100 Mbps e Ethernet de 10 Mbps, como apresentado na Figura 1.

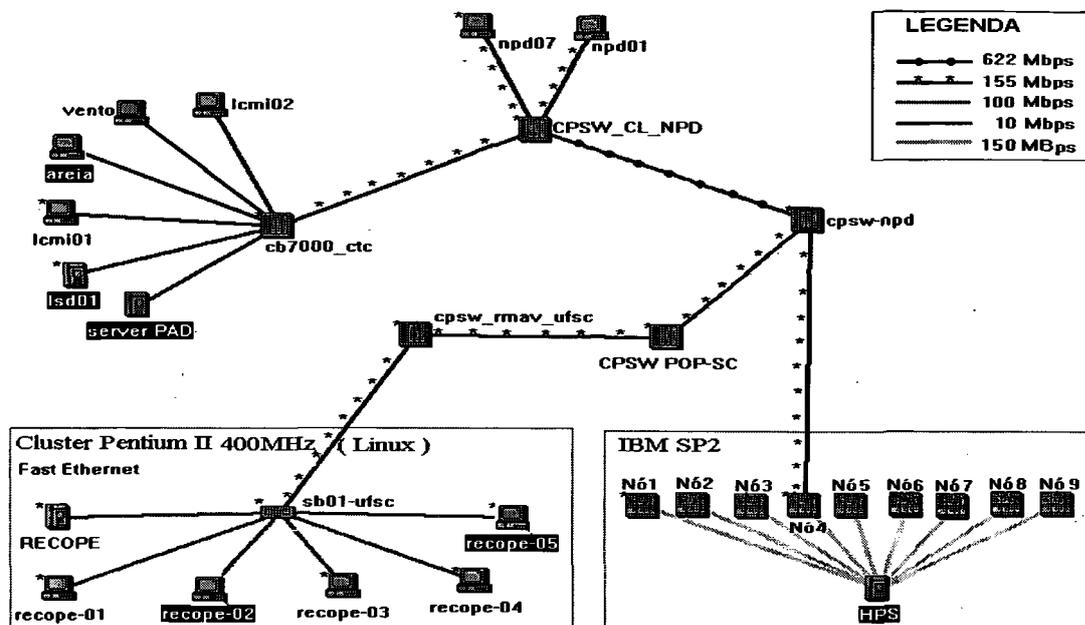


Figura 1- Topologia da Rede

3.2 Caracterização do Ambiente

A comparação de performance entre as técnicas de gerenciamento foi realizada na topologia apresentada na Figura 1. Para manter homogeneidade com os trabalhos de [RUB99a] e [RUB99b] será considerado que os enlaces e os nós não possuem carga, não há perdas nos enlaces e não é considerado o tempo de processamento na camada de aplicação.

O valor da variável de latência será considerado à latência da rede, 65 microsegundos para ATM e 60 microsegundos para Fast Ethernet, como apresentado em [WEL98]; a latência do HPS do IBM SP2 é de 1,2 microsegundos conforme [FAR99]; e a latências de 1 milissegundo é utilizado para a rede Ethernet como em [COS99], [RUB99a].

A largura de banda será de 155 ou 622 Mbps para a rede ATM, 100 Mbps para a rede Fast Ethernet e 10 Mbps para a rede Ethernet. Para o HPS do IBM SP2 será de 150MBps conforme [FAR99].

Para analisar o desempenho da atividade de gerência das técnicas de gerenciamento esta utilizará a variável *ifInErrors* da interface de cada estação da rede, sendo que o tamanho do pedido é de 71 octetos e a resposta é de 76 octetos. A MTU (Unidade Máxima de Transmissão) usada será de 1500 octetos para não haver fragmentação de mensagens SNMP por estas serem pequenas, conforme [RUB99b].

A estação de gerenciamento será considerada como a estação *npd07*, conforme Figura 1. Para o método de gerenciamento com Agentes Autônomos será considerado como Agentes Coletores de Dados as estações *npd01*, *nó 4* e *recope-01*, conforme Figura 1.

4. Comparação das Técnicas de Gerenciamento

Os resultados obtidos na comparação entre as técnicas de gerenciamento são apresentados na Figura 2, sendo considerada apenas uma latência da rede Ethernet de 1 milissegundo. Foi considerado, também, para os Agentes Autônomos o modo em que os Agentes Coletores de Dados não possuem as informações dos seus elementos gerenciados.

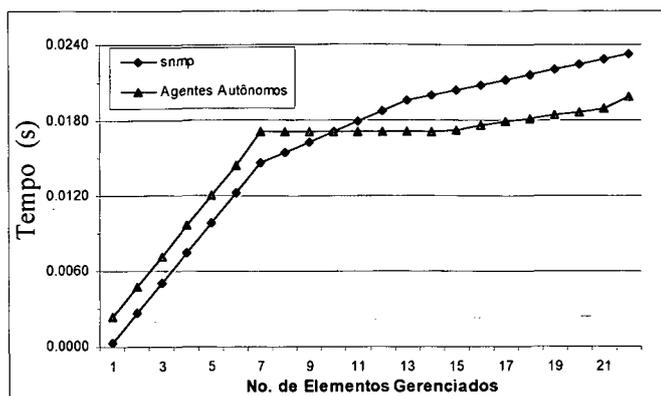


Figura 2 - Comparação das técnicas de gerenciamento

Os primeiros elementos gerenciados estão na rede Ethernet, observa-se duas relações básicas: quanto maior a latência maior será o tempo de resposta, e quanto menor a largura de banda maior será o tempo de resposta.

Os Agentes Autônomos tem como tempo de resposta 17,09 milisegundos, sendo maior ao do SNMP com 17,07 milisegundos somente para os 10 primeiros elementos gerenciados. Mas para as largura de banda maiores e latência menores os Agentes Autônomos tornaram-se melhor que o SNMP obtendo o resultado de 19,84 milisegundos para os Agentes Autônomos e 23,23 milisegundos para o SNMP. Não se pode deixar de atentar o posicionamento dos Agentes Coletores de Dados na rede, eles encontram-se apenas nas redes ATM e *Fast Ethernet*.

5. Conclusões

O gerenciamento SNMP pelo que o seu próprio nome diz Protocolo Simples para Gerência de Redes, isso é devido aos agentes SNMP serem simples e executarem operações elementares. O SNMP possui uma gerência centralizada, o que pode gerar problemas de tráfego e congestionamento dos meios de transmissão. A estação de gerenciamento é responsável por receber todas as informações de gerenciamento, e é provável que os segmentos de rede aos quais a estação de gerenciamento está ligada sustente tráfegos altos estando sujeitos a colisões. Com o crescente tamanho da rede a gerência centralizada pode contribuir para o seu mau funcionamento ao gerar maior quantidade de tráfego na rede.

Os Agentes Autônomos apresentam algumas vantagens como a diminuição do tráfego entre a estação de gerenciamento e os elementos gerenciados, uma vez que o processo de aquisição e análise das informações é levado para mais perto do local dos elementos gerenciados. Uma outra vantagem é maior agilidade na tomada de decisões, as quais podem ser tomadas mais próximos dos elementos gerenciados, evitando-se a necessidade de comunicação com a estação de gerenciamento.

A avaliação analítica das técnicas de gerência oferece um outro critério a ser considerado na escolha da técnica apropriada, além das características específicas de cada técnica. Em um ambiente

de processamento paralelo e distribuído, é avaliado o desempenho de uma tarefa de gerenciamento de rede.

Os resultados indicam que os valores de largura de banda e latência refletem diretamente no desempenho na coleta das informações das técnicas avaliadas.

Foi observado que o melhor tempo de resposta da coleta da variável *ifInErrors* são dos Agentes Autônomos, com 19,84 milisegundos, depois com o SNMP, com 23,23 milisegundos, para esse ambiente e com essas características. A escolha da técnica deve também levar em consideração a sua implementação, tirando proveito da simplicidade de algumas técnicas. Além disso, a escolha deve atender aos objetivos estabelecidos pelo gerente humano da rede.

6. Referências

- [BAL97] BALDI, M., GAI, S. e PICCO, G.P. Exploiting code mobility decentralized and flexible network management. In First International Workshop on Mobile Agents, Berlim, Alemanha, abril de 1997, pp.13-26.
- [BAL98] BALDI, M., e PICCO, G.P. Evaluating the tradeoffs of mobile code design paradigms in network management applications. In Proceedings of the 20th International Conference on Software Engineering (ICSE'97), Kyoto, Japão, abril de 1998.
- [CAS97] CASTRO, Maria Cláudia de Almeida, Avaliação e Melhoria da Segurança Dinâmica utilizando Computação de Alto Desempenho, Tese de Dissertação de Mestrado, Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina – UFSC, Brasil, 1997.
- [COS99] COSTA, Taís Freire da Silva, Avaliação Analítica de Agentes Móveis na Gerência de Redes– Tese de Dissertação de Mestrado em Ciência da Computação – Departamento de Informática e Estatística – Universidade Federal de Santa Catarina – UFSC, Brasil, 1999.
- [FAR99] FARAZDEL, Abbas, ARCHONDO-CALLAO, Gonzalo R., HOCKS, Eva, SAKACHI, Takaaki, VAGNINI, Federico, Understanding and Using the SP Switch, http://www.redbooks.ibm.com/pubs/pdfs/redbooks/s_g245161.pdf, redbook da IBM, 1999.
- [KOC97] KOCH, Fernando Luiz, Agentes Autônomos para Gerenciamento de Redes de Computadores, Tese de Dissertação de Mestrado em Ciência da Computação – Departamento de Informática e Estatística – Universidade Federal de Santa Catarina – UFSC, Brasil, 1997.
- [ROS94] ROSE, Marshall T., The Simple Book: an introduction to Internet management – 2. ed. New Jersey – USA: PTR Prentice – Hall, 1994.
- [RUB99a] - RUBINSTEIN, M. G. and Duarte, O. C. M. B, “Evaluating Tradeoffs of Mobile Agents in Network Management”, Networking and Information Systems Journal, Hermes Science Publications, vol. 2, no. 2, pp. 237-252, 1999.

[RUB99b] RUBINSTEIN, M. G. and Duarte, O. C. M. B. – “Análise da Eficiência de Agentes Móveis no Gerenciamento de Redes”, XIX Congresso Nacional da Sociedade Brasileira de Computação, XXVI Seminário Integrado de Software e Hardware - SEMISH'99, vol. 1, pp. 167-178, Rio de Janeiro, RJ, Brasil, Julho 1999.

[WEL97] WELSH, Matt, BASU, Anindya, EICKEN Thorsten von, ATM and Fast Ethernet Network Interfaces for User-Level Communication, Proceedings of High-Performance Computer Architecture 3, San Antonio, February 1997.

[WEL98] WELSH, Matt, ANNOUNCEMENT: Release of U-Net 2.1 software distribution, University of California, Berkeley <http://www.cs.berkeley.edu/~mdw/projects/unet>, 1998.