

Avaliação de desempenho da arquitetura Wolf

João Angelo Martini^{1*}, Marcos Antônio Cavenaghi² e Claudinei Geraldo¹

¹Departamento de Informática, Universidade Estadual de Maringá, Av. Colombo, 5790, 87020-900, Maringá, Paraná, Brasil.

²Departamento de Computação, Universidade Estadual Paulista, Bauru, São Paulo, Brasil. *Author for correspondence.
e-mail: jangelo@din.uem.br

RESUMO. Este trabalho apresenta a avaliação de desempenho da arquitetura Wolf em função das redes ópticas de interconexão. Com base nos requisitos de comunicação, foi proposto um modelo de rede óptica de interconexão para a arquitetura e também foi proposta uma alteração arquitetural, um novo modelo para a arquitetura Wolf, a fim de se investigar a possibilidade de melhoria de desempenho da arquitetura. Os dois modelos da arquitetura foram simulados usando o simulador SAW. SAW é um simulador dirigido a tempo que implementa as características do Wolf. Os parâmetros distribuição de carga e tempo de execução foram usados para investigar o desempenho da arquitetura. Os resultados obtidos através das simulações realizadas mostram que a segunda versão da arquitetura Wolf apresentou melhores resultados para os parâmetros investigados.

Palavras-chave: redes de interconexão, arquiteturas paralelas, análise de desempenho, arquitetura Wolf.

ABSTRACT. Performance of Wolf Architecture. Performance analysis of interconnection networks for Wolf architecture is provided. Based on communication requirements, an optical interconnection network model for the architecture, an architectural modification, a new model for the Wolf architecture were suggested to investigate the possibilities for performance gain. Both architecture models were simulated with SAW simulator. SAW is a time-driven simulator that implements the Wolf characteristics. Parameters such as load distribution and execution time were dealt with so that performance of the architecture might be investigated. Results by simulations showed that the new model of Wolf architecture gives better results for the parameters investigated.

Key words: interconnection networks, parallel architectures, performance analysis, Wolf architecture.

Redes de interconexão desempenham um papel fundamental no desenvolvimento de um Sistema Paralelo, sendo um fator determinante para o seu desempenho, confiabilidade e custo. As limitações conhecidas das redes de interconexão eletrônicas, tais como largura de banda e desempenho não satisfatórios (Stone e Cocke, 1991; Louri e Sung, 1994; Borella *et al.*, 1994), levam à investigação de outras tecnologias capazes de suportar estes requisitos. Estas investigações trazem à luz uma promissora tecnologia: comunicação óptica.

Esta tecnologia pode satisfazer os requisitos de largura de banda e desempenho impostos pelas altas taxas de comunicação observadas nesses sistemas. Redes de interconexão baseadas nesta tecnologia (redes de interconexão óptica) são amplamente estudadas. Neste contexto, *Wavelength Division Multiplexing* (WDM) (Willner, 1997; Borella *et al.*, 1997; Senior *et al.*, 1998) pode ser usado para implementar tais redes.

WDM divide a largura de banda da fibra óptica em canais assíncronos, cada um com um diferente comprimento de onda e taxa de transmissão. Este mecanismo permite a implementação de taxas da ordem de Tb/s (Terabits/s). As redes baseadas em WDM têm algumas vantagens sobre as redes eletrônicas convencionais (Borella *et al.*, 1994; Chatterjee e Pawlowski, 1999):

- Largura de Banda: a largura de banda de uma fibra é de 25 THz e há sistemas comerciais operando em Gb/s e sistemas experimentais em Tb/s. A tecnologia comercial atual está muito próxima dos 40 Gbps por canal e com a evolução observada nos últimos anos; acredita-se que velocidades de Tbps serão alcançadas em breve.
- Baixo Nível de Perdas: em seu comprimento de onda de atenuação mínima, 1,55 μm , a perda em uma fibra é da ordem de 0,2 dB/Km, enquanto que num cabo coaxial de cobre a perda é de 2,5 dB/Km em um sinal de

1 MHz e mais de 50 dB/Km em 1 GHz. Essa atenuação é muito mais relevante quanto mais alta for a taxa de transmissão.

- Escalabilidade: a perda de potência em um barramento comum cresce linearmente com o número de processadores, já num acoplador estrela passiva (PSC - *Passive Star Couple*) cresce logaritmicamente. Isso permite acoplar uma quantidade razoável de nós à rede sem a utilização de regeneradores.
- Modularidade: na maioria dos sistemas *single-hop* um nó pode ser acrescentado à rede sem necessidade de se redimensionar o sistema todo. Isso é ideal para Sistemas de Alto Desempenho, em que a adição de um nó ao sistema deve ser o mais simples possível.
- *Multicasting*: *broadcasting* e *multicasting* são viabilizados em uma rede WDM devido à natureza *broadcast* do acoplador PSC.

O presente trabalho avalia o desempenho da arquitetura Wolf em função das redes ópticas de interconexão, considerando os parâmetros distribuição de carga e tempo de execução. Duas abordagens foram adotadas neste trabalho: uma propôs um modelo de rede de interconexão para a arquitetura e a outra propôs uma alteração arquitetural, isto é, uma nova abordagem para a arquitetura Wolf, a fim de se investigar a possibilidade de melhoria de desempenho da arquitetura. Os dois modelos da arquitetura foram simulados, e os resultados obtidos mostram que a segunda versão da arquitetura Wolf apresentou melhores resultados para os parâmetros investigados.

Nas seções subseqüentes, são apresentados os seguintes pontos: conceitos de redes de interconexão em arquiteturas paralelas, situando a fotônica como forte candidata ao papel de rede para arquiteturas paralelas; introduz-se a arquitetura Wolf, seus requisitos de comunicação e a nova abordagem proposta para a arquitetura Wolf, então renomeada para arquitetura Wolf 2; análise de desempenho dos dois modelos da arquitetura Wolf em função dos parâmetros considerados neste estudo e também em função da taxa de transmissão da rede de interconexão; por fim apresentam-se as conclusões obtidas através deste trabalho de investigação.

Contexto de redes de interconexão em arquiteturas paralelas

O desenvolvimento de redes de comunicação e arquiteturas paralelas está relacionado desde os projetos pioneiros. Muitos projetos experimentais em arquiteturas paralelas foram investigados a partir de então e têm servido como base para subseqüentes

desenvolvimentos na área de redes de alta velocidade baseadas em comunicação óptica.

Porém, o interesse em relação às pesquisas na área de Arquiteturas Paralelas tem tido altos e baixos, apresentando um comportamento cíclico de problemas e soluções em que novos experimentos e estudos trazem à tona problemas anteriormente não detectados e suscitam a busca de soluções através do desenvolvimento de novas pesquisas.

O interesse pela área diminuiu quando se identificou que o subsistema de interconexão era tão ou mais complexo do que todo o restante da arquitetura. Conforme a velocidade dos elementos computacionais foi aumentando, cada vez mais o gargalo foi recaído sobre a interconexão entre os elementos (Stone e Cocke, 1991; Louri e Sung, 1994). Dentro desse ciclo problema-solução, inserem-se dois trabalhos de investigação na área de redes de interconexão. O primeiro, desenvolvido por Martini (Martini, 1994), discute a influência das redes de interconexão na arquitetura Wolf (Cavenaghi et al., 1998, 2000a e b). Nesse trabalho, implementaram-se duas redes de interconexão para as quais foram testadas diversas topologias: *crossbar*, *GSN* e *Delta*. Foram analisadas todas as combinações desses pares de redes de interconexão para uma série de *benchmarks*.

Esse trabalho mostrou que não há uma única combinação de topologias para as redes de interconexão da arquitetura que seja eficiente para todos os casos. Esse é o principal resultado do trabalho de Martini (Martini, 1994), colocando a fotônica numa perspectiva de ser a abordagem que poderia viabilizar a arquitetura. Também neste contexto, insere-se outro trabalho desenvolvido por Martins (Martins, 1998), que mostra a viabilidade e o potencial de exploração e investigação dessa nova tecnologia (fotônica) para arquiteturas paralelas em geral.

Arquitetura Wolf

Esse processo de investigação de rede de interconexão para a arquitetura Wolf conduziu a uma nova proposta que emprega redes ópticas de interconexão na arquitetura. A arquitetura a fluxo de dados Wolf, apresentada na Figura 1, é baseada no modelo de fluxo de dados dinâmico. Para situar a arquitetura Wolf dentro do contexto de fluxo de dados, faz-se necessário apresentar uma descrição sucinta dos conceitos fundamentais relacionados a esse paradigma de computação.

Os princípios fundamentais do modelo de fluxo de dados foram propostos na década de 70. O conceito fundamental do modelo de fluxo de dados

baseia-se na chamada *regra de disparo*. Segundo essa regra, uma instrução é escalonada para execução quando todos os operandos de entrada necessários para sua execução ficarem disponíveis. Quando essa condição é satisfeita, diz-se que a instrução foi disparada. Os dados de entrada são então consumidos e geram-se outros dados de saída.

A seqüência de execução das instruções está intrinsecamente relacionada à dependência dos dados. Um programa em fluxo de dados não é representado por uma seqüência de instruções com controle central que impõe a ordem de execução, como no modelo de von Neumann, mas é representado através de um grafo bidimensional. No grafo, cada instrução é representada por um nó e os arcos que ligam os nós representam os caminhos que os dados percorrem até a instrução subsequente.

Os dados são transmitidos através de fichas¹ que percorrem os arcos entre os nós. Como os nós podem possuir mais de uma entrada, sua execução só é disparada após a chegada de todas as fichas nas correspondentes entradas. O nó que executa a operação somente é executado a partir da chegada das fichas correspondentes aos seus dados em suas entradas. Supondo que apenas uma ficha esteja presente na sua respectiva entrada, então essa ficha deve esperar pelo seu par correspondente, que é denominada ficha parceira ou ficha par. A partir da chegada da parceira, as fichas são emparelhadas e o nó pode ser executado.

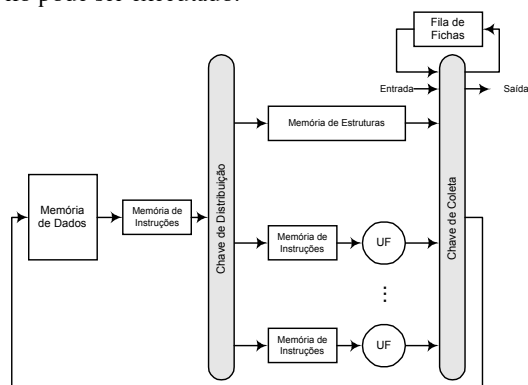


Figura 1. Arquitetura Wolf

O funcionamento da arquitetura é descrito resumidamente a seguir: as fichas são inseridas na arquitetura através da Chave de Coleta, esta chave transmite as fichas para a Memória de Dados, se esta unidade estiver disponível, ou então transmite para a Fila de Fichas para posterior distribuição. A

¹ Em arquiteturas a fluxo de dados os dados e as informações de controle referentes a eles são representados por entidades denominadas fichas.

Memória de Dados gera um pacote formado por um par de fichas dirigidas para um mesmo nó. Esse pacote é enviado para a Memória de Instruções Externa. Dependendo do tipo do pacote recebido por essa memória, ela pode gerar fichas destinadas a operações de tratamento de estruturas de dados ou transmitir o pacote inalterado para a Chave de Distribuição. A Chave de Distribuição pode transmitir o pacote para a Memória de Estruturas ou para uma das Memórias de Instruções Internas. No caso de pacotes destinados à Memória de Instruções Interna, esses pacotes são enviados para a primeira memória que estiver disponível. A partir do pacote recebido, a Memória de Instruções Interna cria um pacote executável com todas as informações necessárias para a execução da instrução enviando-o à sua correspondente Unidade Funcional, a qual processa esse pacote executável, produzindo novos pacotes de dados. A Memória de Estruturas é responsável pelo armazenamento de dados estruturados.

Requisitos de comunicação da Arquitetura Wolf e proposta de rede de interconexão

O estudo dos requisitos de comunicação da arquitetura Wolf permite que se conheça o fluxo das fichas na arquitetura e, portanto, viabiliza uma proposta de modelo de rede de interconexão para a arquitetura. A rede de interconexão deve interligar as unidades da arquitetura obedecendo os requisitos de comunicação apresentados por elas. Deve haver algum tipo de controle que faça o gerenciamento quando dois ou mais nós tentam transmitir simultaneamente para o mesmo destino, evitando a ocorrência de colisões. Portanto, é necessário um protocolo de controle que arbitre o acesso ao meio de transmissão para permitir que as unidades efetuem suas transmissões.

Baseado na análise de requisitos de comunicação da arquitetura, definiram-se os parâmetros de projeto para as redes de interconexão da arquitetura Wolf. A arquitetura da rede é baseada no sistema WDM *single-hop*. No modelo proposto, cada nó da rede recebe dados através de um canal (comprimento de onda) específico. Dessa forma, um nó transmissor que queira transmitir para um determinado destino deve ter seu transmissor sintonizado no comprimento de onda específico do receptor de destino desejado. Portanto, o acesso de cada nó ao canal de transmissão de dados é realizado através de um transmissor sintonizável e um receptor fixo. A recepção de cada nó é realizada através de uma onda portadora independente. Isso evita a dependência do sistema em relação à

sintonização dos receptores e também evita colisões de receptores em que o receptor não está sintonizado no comprimento de onda de um determinado nó transmissor. Esse controle evita colisões de transmissores, pois os nós só transmitem para os destinos que estão livres e nunca iniciam um ciclo de transmissão para o mesmo destino concorrentemente.

O estudo da arquitetura viabilizou uma nova proposta de alteração arquitetural, que é mostrada na Figura 2. Esta é uma tentativa de se investigar uma possível distribuição de tráfego mais uniforme para a arquitetura. Esta nova proposta para a arquitetura Wolf também foi simulada. As duas versões da arquitetura Wolf foram investigadas, e os resultados foram analisados comparativamente. Doravante, a arquitetura Wolf com redes ópticas, mostrada na Figura 1, será chamada de Wolf 1 e a arquitetura Wolf, com a modificação, mostrada na Figura 2, será chamada de Wolf 2.

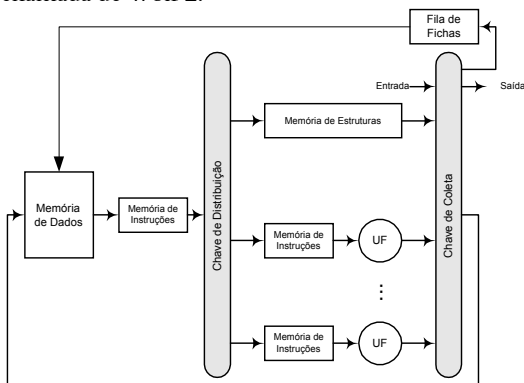


Figura 2. Nova proposta para a Arquitetura Wolf

Avaliação de desempenho da Arquitetura Wolf

Esta seção apresenta a avaliação de desempenho da arquitetura Wolf. Para realizar esta avaliação, utilizaram-se os seguintes parâmetros: distribuição de carga e tempo de execução.

Para investigar o comportamento da arquitetura Wolf utilizou-se um simulador dirigido a tempo desenvolvido em GNU C++. O simulador implementa todas as características da arquitetura descrita anteriormente.

O simulador implementa um mecanismo de temporização denominado de relógio global, o qual é incrementado a cada vez que todas as unidades da arquitetura são chamadas no simulador. No final da execução de um programa, o valor final apresentado pelo relógio global representa o número total de ciclos consumidos para a simulação do referido programa. A menor unidade de tempo gerenciada

pelo simulador é denominada de *time step* ou de ciclo.

As simulações foram realizadas com a utilização de *benchmarks* que geram tráfego intenso na rede de interconexão: Somatorial, Integração Binária, Gauss, Interpolação, Laplace, Multiplicação de Matrizes, N-Queens e Simple (Cavenaghi et al., 1998). Foram usadas diversas taxas de transmissão durante as simulações. As taxas de transmissão usadas foram: $T_{x_{2,5}}$ para 2,5 Gbps, T_{x_5} para 5 Gbps, $T_{x_{10}}$ para 10 Gbps, $T_{x_{20}}$ para 20 Gbps, $T_{x_{40}}$ para 40 Gbps e $T_{x_{400}}$ para 400 Gbps.

A faixa de taxas de transmissão abrange taxas disponíveis comercialmente até as taxas mais elevadas factíveis em ambientes de laboratórios de pesquisas ([Willner, 1997; Senior et al., 1998; Shi et al., 1998; Terahara et al., 1998; Aisawa et al., 1998; Chen et al., 1998]). Os valores das taxas de transmissão são multiplicados por dois até a taxa $T_{x_{40}}$, e a seguir multiplicou-se a taxa por um fator de dez, a fim de se investigar a possibilidade de encontrar uma margem superior limitante para a taxa de transmissão.

Parâmetros de investigação

A simulação da arquitetura é complexa e gera resultados volumosos e de difícil interpretação. Para facilitar a análise, a apresentação desses resultados será segmentada.

Os parâmetros analisados foram:

- **distribuição de carga:** investiga a eficiência do subsistema de chaveamento na transferência de atividade computacional entre os módulos do sistema. Esse aspecto é importante porque a arquitetura paralela investigada é de granulosidade fina, gerando um fluxo intenso de pacotes em todos os canais.
- **tempo de execução:** investiga o impacto da eficiência da rede de interconexão na execução global do algoritmo. Esse aspecto é importante porque se sabe que há grande variabilidade da carga computacional em arquiteturas paralelas, e as redes de interconexão apresentam comportamentos distintos sob diferentes cargas.

Benchmarks

Esta subseção apresenta os programas de teste utilizados para se investigar a rede de interconexão da arquitetura Wolf. Foram adotados como programas de teste alguns algoritmos já extensivamente usados na pesquisa de arquiteturas paralelas por outros grupos de pesquisa.

Os programas utilizados são os seguintes:

- **somatorial:** esse programa calcula o somatório de N pela fórmula Soma $(N, (N+1)/2+1)$ + Soma $((N+1)/2, 1)$. É um algoritmo duplamente recursivo, que gera uma árvore binária de processos em que folhas de um mesmo nível podem ser executadas em paralelo.
- **integração binária:** esse programa integra a função $y = x^2 - 6x - 10$, num intervalo $[x_1, x_2]$ através do método dos trapézios. É um programa com alto grau de recursividade, que cria uma árvore de processos com 2^{n-1} processos, em que n é o número de divisões do intervalo.
- **eliminação de Gauss:** esse algoritmo é extensivamente usado em soluções de sistemas de equações lineares e usa o método de triangularização com pivotamento parcial.
- **interpolação:** esse programa usa o método de interpolação polinomial para efetuar aproximações.
- **Laplace:** esse algoritmo resolve uma equação simples de Laplace através do método de relaxação.
- **multiplicação de matrizes:** esse programa cria e multiplica duas matrizes quadradas.
- **N-Queens:** esse programa consiste na solução recursiva do problema de colocar N rainhas em um tabuleiro $N \times N$ de forma que nenhuma ataque a outra. É um algoritmo típico de estratégia dividir-para-conquistar.
- **simple:** é um programa de simulação hidrodinâmica simples codificado pelo *Lawrence Livermore National Laboratory* que tem sido extensivamente usado por grupos de pesquisa em arquiteturas a fluxo de dados.

Resultados e discussão

Esta seção apresenta os resultados obtidos através deste projeto. Os resultados são apresentados em função dos parâmetros investigados: distribuição de carga e tempo de execução.

Distribuição de carga

A Figura 3 mostra o coeficiente de variação (C.V.) para as distribuições de carga apresentadas pelas duas versões da arquitetura Wolf para os *benchmarks* testados. O desvio padrão representa a variação dos valores em torno da média. Geralmente o desvio padrão é apresentado de forma relativa ao valor da média, essa relação é denominada de

coeficiente de variação. Valores menores do coeficiente representam um sistema mais homogêneo, ou seja, há uma distribuição de carga mais uniforme. O que é uma característica desejável para um sistema de computação paralelo, visto que, se o sistema é homogêneo, seus processadores têm carga de trabalho uniforme. Não há sobrecarga de uns e ociosidade de outros. Se a distribuição de carga não é uniforme, ela ocasiona um desequilíbrio, isto é, provoca sobrecarga em alguns processadores, enquanto outros processadores trabalham muito pouco, tornando o sistema desbalanceado.

As figuras de distribuição de carga apresentam, no eixo vertical, o coeficiente de variação e , no eixo horizontal, as taxas de transmissão simuladas para a rede de interconexão, para as duas versões da arquitetura. Em suma, as figuras permitem identificar se há algum impacto da taxa de transmissão sobre a distribuição de carga, além de permitir verificar se a alteração arquitetural implementada tem também alguma influência sobre a distribuição de carga. Nas legendas das figuras onde se lê Nova Wolf, deve-se entender Wolf 1 e em Nova WolfTQ, deve-se entender Wolf 2.

A distribuição de carga para as duas versões da arquitetura Wolf apresentaram semelhanças de comportamento para a maioria dos *benchmarks* testados. Para os *benchmarks*: Somatorial, Integração Binária e N-Queens, a Wolf 2 apresentou distribuição de carga muito melhor que a Wolf 1, para a região que abrange as taxas de transmissão de 2,5 Gbps a 20 Gbps. Para as taxas de 40 Gbps e 400 Gbps, as distribuições de carga dos dois modelos da arquitetura Wolf ficaram muito próximas e foram piores que para a região de $T_{x_{2,5}}$ a $T_{x_{20}}$.

Já para os *benchmarks*: Gauss, Interpolação, Laplace e Simple, o comportamento foi semelhante aos anteriores, mas as diferenças entre os coeficientes de variação das duas versões da arquitetura não foram tão significativas como nos casos anteriores.

E o *benchmark* Multiplicação de Matrizes foi o que apresentou um comportamento distinto dos demais, no sentido de que as diferenças entre os dois modelos da arquitetura foram bem menores que nos casos anteriores. Além disso, a região com melhor distribuição de carga foi obtida para as taxas de $T_{x_{40}}$ e $T_{x_{400}}$, ao contrário dos demais *benchmarks*, mostrando que esse tipo de algoritmo apresenta peculiaridades que afetam o desempenho da arquitetura de modo diferente dos demais.

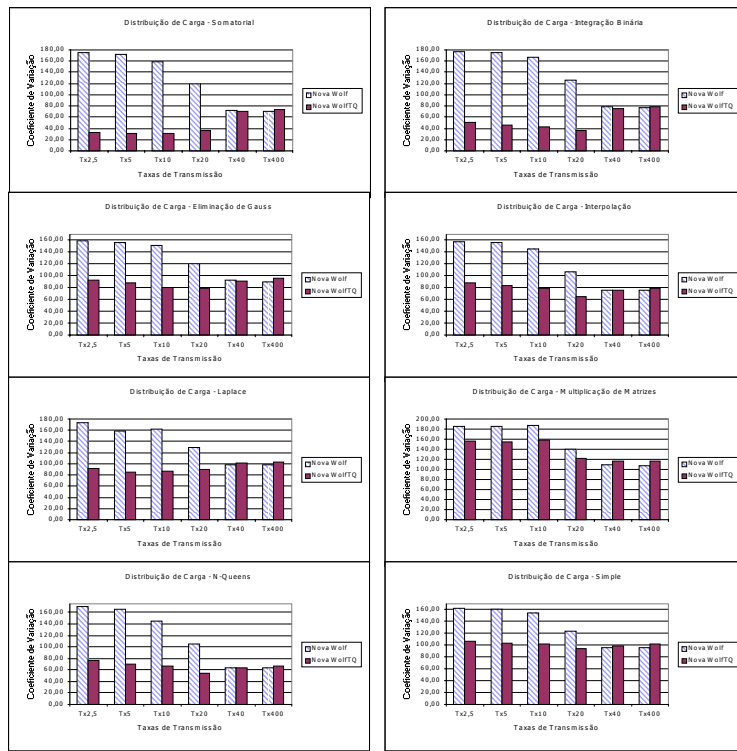


Figura 3. Coeficiente de variação para distribuição de carga

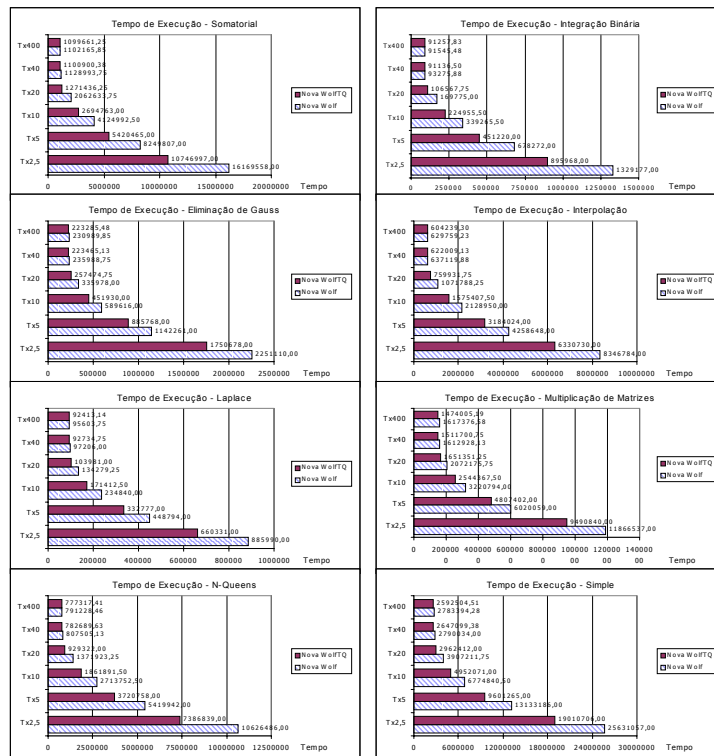


Figura 4. Tempos de execução para as duas versões da Wolf

De modo geral, pode-se concluir que para a distribuição de carga o modelo da arquitetura Wolf 2 apresentou melhores resultados e que a taxa de transmissão tem influência sobre a distribuição de carga na arquitetura, sendo a região de $T_{x_{2,5}}$ a $T_{x_{20}}$ a que proporciona melhor distribuição.

Tempo de execução

Esta seção apresenta os resultados dos tempos de execuções dos *benchmarks* em função das taxas de transmissão e dos dois modelos da arquitetura Wolf. Esse estudo é importante para permitir o balanceamento da velocidade da rede de interconexão em relação à velocidade das demais unidades. Essa investigação possibilita detectar se há limitantes inferiores e superiores para a velocidade da rede em função dos tempos de execução.

A Figura 4 mostra os tempos de execução para as duas versões da arquitetura Wolf em função das variações das taxas de transmissão. O eixo das ordenadas representa as velocidades denominadas de $T_{x_{2,5}}$ até $T_{x_{400}}$ e o eixo das abscissas representa o tempo de execução² simulado em função da taxa de transmissão da rede.

Os resultados apresentados por esses gráficos mostram que, para todas as taxas de transmissão, ocorreram melhorias muito significativas de desempenho para a versão Wolf 2. Para ilustrar esse ganho de desempenho exemplifica-se o ganho obtido para alguns *benchmarks* simulados.

Para o *benchmark* Somatorial, obteve-se uma redução de aproximadamente 33% no tempo de execução para a Wolf 2, com a taxa de 2,5 Gbps. Outro exemplo que quantifica bem o ganho expressivo obtido através da modificação arquitetural implementada é o caso do *benchmark* Integração Binária. Para esse programa, obteve-se um ganho de aproximadamente 32% no tempo de execução para a taxa de transmissão de 2,5 Gbps.

O tempo de execução é um dos parâmetros mais utilizados para avaliação de desempenho de arquiteturas paralelas e de alto desempenho. A partir daí pôde-se avaliar o quão importante foi o desenvolvimento da modificação arquitetural proposta neste projeto, pois os resultados mostram que para este parâmetro o desempenho obtido com a nova versão da arquitetura Wolf foi extremamente expressivo.

Os dados obtidos permitiram concluir que arquiteturas paralelas constituem uma alternativa para a obtenção de sistemas computacionais de alta performance. No entanto, um dos principais desafios de projeto e frequentemente o gargalo de tais sistemas são as redes de interconexão. Diante das limitações das redes de interconexão eletrônicas tradicionais e do desenvolvimento e amadurecimento da tecnologia de comunicação óptica, o interesse de pesquisa na área de redes de interconexão migrou para a tecnologia WDM.

No espectro de modelos de arquiteturas paralelas propostas, o modelo de fluxo de dados tem sido questionado por exigir grande capacidade de chaveamento. Essa característica parece complementar uma das principais características dos sistemas de comunicação óptica que é disponibilizar uma grande capacidade de comunicação. Surge então uma promissora possibilidade que é a de se investigar a implementação de comunicação óptica em arquiteturas a fluxo de dados. Isso tem o potencial de viabilizar arquiteturas baseadas no paradigma de fluxo de dados e outras arquiteturas paralelas.

Este estudo apresentou a arquitetura Wolf, analisou os requisitos de comunicação apresentados pela arquitetura e, a partir dessa investigação, definiu-se um modelo de rede de interconexão para essa arquitetura. Uma vantagem apresentada pelo modelo de rede proposta é a inexistência de colisões, o que evita o *overhead* de comunicação, que seria gerado caso houvesse necessidade de se tratar colisões.

Outra vantagem do modelo proposto é que a rede apresenta independência de receptores sintonizáveis, isto é, evita as colisões de receptores que ocorrem quando o receptor não está sintonizado no comprimento de onda do transmissor que está lhe enviando o dado.

Trabalhos de investigação na área de arquiteturas paralelas têm sido extensivamente realizados. Entretanto, uma parcela significativa dos modelos de investigação são comumente baseados em hipóteses que nem sempre são realistas, principalmente na área de redes de interconexão em arquiteturas paralelas. Em alguns trabalhos de investigação são consideradas algumas hipóteses que ajudam a simplificar o modelo teórico, mas a simplificação pode resultar em discrepâncias na predição do desempenho da arquitetura. O presente trabalho é importante, porque apresenta medidas quantitativas das estratégias de

² A unidade de tempo de execução é denominada de *time step* conforme apresentado anteriormente e é utilizada como unidade do eixo das abscissas para todos os gráficos desta seção.

investigação propostas para a arquitetura Wolf (objeto deste estudo).

O comportamento da nova versão da arquitetura Wolf foi analisado comparativamente à versão original da arquitetura em função da variação da taxa de transmissão da rede de interconexão para os seguintes parâmetros: Distribuição de Carga e Tempo de Execução.

Em relação a esses parâmetros investigados na avaliação de desempenho da arquitetura pode-se chegar às seguintes conclusões:

Distribuição de Carga: para todos os *benchmarks* simulados, a versão 2 da Wolf apresentou uma melhor distribuição de carga principalmente para as taxas de transmissão na região abrangendo a faixa de 2,5 Gbps a 20 Gbps. Isso suscita maiores investigações abrangendo a região acima de 20 Gbps a fim de se identificar qual o comportamento da arquitetura para uma faixa de transmissão mais refinada, ou seja, há possibilidade de se investigar a faixa acima de 20 Gbps com variações menores na taxa de transmissão.

Tempo de Execução: os resultados de investigação do comportamento da arquitetura em termos de tempos de execução mostraram que a versão 2 da arquitetura Wolf apresentou desempenho melhor que a versão original para todas as taxas de transmissão. A implementação da modificação arquitetural proposta trouxe melhorias muito significativas para os tempos de execução dos *benchmarks* confirmando a aposta inicial de que a versão 2 da arquitetura poderia apresentar melhores resultados de desempenho.

Agradecimentos

Os autores agradecem ao CNPq (Programa PIBIC-CNPq/UEM), à Fundação Araucária, à FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), à FUNDUNESP (Fundação para o Desenvolvimento da UNESP) e ao Infomar - Centro Softex Genesis de Maringá pelo apoio financeiro.

Referências

AISAWA, S. et al. Ultra-wideband, long distance WDM demonstration of 1 Tbit/s (50x20 Gbit/s), 600 Km transmission using 1550 and 1580 nm wavelength bands. *Electron. Lett.*, Stevenage, v. 34, n. 11, p. 1127-1129, 1998.

BORELLA, M. et al. Optical interconnects for multiprocessor architectures using wavelength-division multiplexing. In: ANNUAL HAWAII INTERNATIONAL CONFERENCE ON SYSTEM

SCIENCES, 27, 1994, Honolulu. *Proceedings...* Honolulu, 1994, p. 499-508.

BORELLA, M. et al. Optical components for WDM Lightwave Networks. *Proc. IEEE*, New York, v. 85, n. 8, p. 1274-1307, 1997.

CAVENAGHI, M.A. et al. Solution for the Wolf Architecture Load Balancing Problem. 2000 SYMPOSIUM ON PERFORMANCE EVALUATION OF COMPUTER AND TELECOMMUNICATION SYSTEMS (SPECTS 2000), 2000, Vancouver, *Proceedings...* Vancouver, 2000. p. 103-107.

CAVENAGHI, M.A. et al. Proposal to Solve Workload Problems in the Simulator of the Wolf Architecture. INTERNATIONAL CONFERENCE/EXHIBITION ON HIGH PERFORMANCE COMPUTING IN ASIA-PACIFIC REGION, 4, 2000, Beijing. *Proceedings...* Beijing: The IEEE Computer Society, 2000. p. 315-318.

CAVENAGHI, M.A. et al. A Simple Mechanism to Deal with Sequential Code in Dataflow Architectures. INTERNATIONAL CONFERENCE/EXHIBITION ON HIGH PERFORMANCE COMPUTING IN ASIA-PACIFIC REGION, 4, 2000, Madras. *Proceedings...* Madras: The IEEE Computer Society, 1998. p. 188-193.

CHATTERJEE, S.; PAWLOWSKI, S. Enlightening the effects and implications of nearly infinite bandwidth. *Communications of the ACM*, v. 42, n. 6, p. 75-83, 1998.

CHEN, C. D. et al. 1.2 Tbit/s (30 ch x 40 Gbit/s) WDM transmission over 85 Km fibre. *Electron. Lett.*, vol. 34, n. 10, p. 1002-1004, 1998.

LOURI, A.; SUNG, H. 3D Optical interconnects for high-speed interchip and interboard communications. *IEEE Computer*, v. 27, n. 10, p. 27-37, 1994.

MARTINI, J.A.. *Estudo de desempenho de redes de interconexão para a arquitetura a fluxo de dados Proto-Wolf*. 1994. Dissertação (Mestrado) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 1994.

MARTINS, J. E. M. P. *Investigação de técnicas fotônicas de chaveamento aplicadas em arquiteturas paralelas*. 1998. Tese (Doutorado) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 1998.

SENIOR, J. M. et al. Developments in wavelength division multiple access networking. *IEEE Communic. Magazine*, Piscataway, v. 36, n. 12, p. 28-36, 1998.

SHI, H. et al. 20x5 Gbit/s optical WDM transmitter using single-stripe multiwavelength modelocked semiconductor laser. *Electron. Lett.*, v. 34, n. 2, p. 179-181, 1998.

STONE, H. S.; COCKE, J. Computer architecture in the 1990s. *IEEE Computer*, v. 24, p. 30-38, 1991.

TERAHARA, T. et al. 0.7 Tbit/s (66x10.66 Gbit/s) WDM transmission over 2212 Km using broadband, high-power EDFA with pump reflector. *Electron. Lett.*, Stevenage, v. 34, n. 10, p. 1001-1002, 1998.

WILLNER, A. E. Mining the optical bandwidth for a terabit per second. *IEEE Spectrum*, New York, v. 34, n. 4, p. 32-41, 1997.

Received on August 18, 2001.

Accepted on October 31, 2001.