Trabalho de Análise de Desempenho

Adriana Cesário de Faria Alvim
Profs. Noemi e Celso
Departamento de Informática
Pontifícia Universidade Católica do Rio de Janeiro

Período 99.1

1 Envio com Bloqueio em MPI

No modelo de programação MPI [1], uma execução compreende um ou mais processos que se comunicam chamando rotinas de uma biblioteca para receber e enviar mensagens para outros processos.

Uma operação de envio/recebimento com bloqueio suspende a execução do programa até que o buffer da mensagem sendo recebida/enviada possa ser reutilizado. Operações sem bloqueio retornam o controle para o programa imediatamente após terem iniciado a comunicação, separando comunicação e computação.

O modo de comunicação é determinado pela operação de envio. Existem quatro operações de envio com bloqueio e quatro operações de envio sem bloqueio, que correspondem aos quatro modos de comunicação existentes. A operação de recebimento não especifica o modo de comunicação, simplesmente é com bloqueio ou sem bloqueio.

A diferença básica entre os modos de comunicação reside em dois pontos, que por sua vez, estão relacionados à duas fontes de *overhead*:

Fontes de *overhead*:

- **Overhead de Sistema**: acontece quando os dados da mensagem são copiados do buffer da tarefa emissora para a rede, e copiados da rede para o buffer da tarefa receptora.

- **Overhead de Sincronismo**: é o tempo gasto pela espera do acontecimento de um evento em uma outra tarefa.

Diferenças entre os modos de comunicação:

- no fato da operação de envio precisar ou não que um recebimento combinado seja executado. Pode incorrer em *overhead* de sincronismo.
- no fato dos dados da mensagem serem bufferizados pelo sistema ou só copiados pela tarefa receptora quando este estiver executando o recebimento. Pode incorrer em *overhead* de sistema.

A bufferização de mensagens pelo sistema desassocia as operações de envio e recebimento, uma vez que um envio com bloqueio pode se completar tão logo a mensagem seja bufferizada, mesmo que o receptor não tenha executado uma operação de recebimento combinada com esta. Por outro lado, a bufferização pode ser bastante cara, pois acarreta cópias e alocações de memória adicionais.

1.1 Modos de Comunicação Ponto-a-Ponto com Bloqueio

- **Envio com bloqueio em modo síncrono** (MPI_SSEND): Uma operação de envio que use o modo síncrono pode ser iniciada com uma operação de recebimento combinada executada ou não. No entanto, o envio só será completado com sucesso se a execução de um recebimento combinado ocorrer e a operação de recebimento tiver iniciado o recebimento da mensagem. Quando a operação MPI_SSEND é executada, uma notificação "pronto para enviar" é enviada pela tarefa emissora. Quando a tarefa receptora executa uma rotina de recebimento uma notificação "pronto para receber" é enviada. Os dados são então transferidos. Portanto, o fim de um envio síncrono não só indica que o buffer da tarefa emissora pode ser reutilizado, mas indica também que a tarefa receptora iniciou o recebimento da mensagem. Não importa a ordem relativa das mensagens, para que uma comunicação no modo síncrono se complete é necessário que ambas as tarefas atinjam a comunicação, isto é, estejam sincronizadas ("aperto de mãos"). Assim, algum *overhead* de sincronismo incorre tanto na tarefa emissora como na receptora, e ele tende a ser maior naquela que precede a outra. Este modo é o mais seguro, pois seu correto funcionamento não depende da ordem das mensagens nem do tamanho das mensagens.
- **Envio com bloqueio em modo pronto** (MPI_RSEND): Uma operação de envio neste modo de comunicação só pode ser iniciada se uma operação de recebimento combinada tiver sido executada. De outra forma, a operação

está errada e seu resultado é indefinido. A operação MPI_RSEND simplesmente envia a mensagem pela rede. No entanto é necessário que a notificação "pronto para receber" tenha chegado, indicando que a tarefa receptora enviou uma mensagem de recebimento. No caso de operação com bloqueio, a única espera na tarefa emissora é que todos os dados tenham sido transferidos do buffer da mensagem. Dependendo do quanto antes aconteceu a operação de recebimento, algum *overhead* de sincronismo pode incorrer na tarefa receptora. Este modo visa minimizar o *overhead* de sistema e *overhead* de sincronismo na tarefa emisssora. Neste modo a ordem das mensagens é fundamental e não deve ser utilizado ao menos que o usuário tenha certeza que um recebimento combinado tenha sido executado.

- **Envio com bloqueio em modo bufferizado** (MPI_BSEND): Neste modo o usuário garante a bufferização, pelo sistema, das mensagens enviadas. A operação MPI_RSEND copia os dados do buffer da mensagem para um buffer fornecido pelo usuário, e retorna. A tarefa emissora pode então seguir com computação e alterar o conteúdo original do buffer da mensagem, sabendo que estas alterações não irão refletir nos dados efetivamente enviados. A operação de envio pode então completar antes que uma operação de recebimento combinada seja executada. Neste caso, quando a operação de envio termina, a tarefa emissora não sabe se a mensagem por ela enviada foi recebida ou não. Quando a notificação "pronto para receber" chega, os dados do buffer fornecido pelo usuário são copiados através da rede. Este modo incorre em um *overhead* de sistema adicional ocasionado pela cópia dos dados do buffer da mensagem para o buffer do usuário. *Overhead* de sincronismo é eliminado na tarefa emissora, para ela é irrelevante o momento em que a tarefa receptora executa a operação de recebimento. Algum *overhead* de sincronismo pode acontecer na tarefa receptora, ela precisa esperar a mensagem chegar antes de retornar. Um erro ocorrerá na ausência de espaço no buffer fornecido pelo usuário.

- **Envio com bloqueio no modo padrão** (MPI_SEND): Neste modo o MPI determina o melhor comportamento do sistema que funcione para a maioria dos usuários. Duas situações podem acontecer, dependendo se o tamanho da mensagem é maior ou menor que um determinado valor de TRESHOLD.

  Quando o tamanho da mensagem é menor ou igual ao TRESHOLD a operação MPI_SEND copia a mensagem através da rede para um buffer do sistema na tarefa receptora. A operação de envio retorna e a tarefa emissora pode continuar com computação. Na inicialização do programa, um buffer do sistema por tarefa é disponibilizado. Quando a operação de recebimento for executada, os dados serão copiados do buffer do sistema para a tarefa receptora.

  Da mesma forma que no modo bufferizado, a utilização de um buffer diminui o *overhead* de sincronismo na tarefa emissora; por outro lado, aumenta o *overhead* de sistema resultante do trabalho extra de copiar para o buffer. De todas as maneiras, é possivel incorrer *overhead* de sincronismo na tarefa receptora se esta precede a tarefa emissora. Diferentemente como ocorre no modo bufferizado, não acontecerá um erro se não houver espaço no buffer do sistema. Neste caso, a tarefa emissora ficará bloqueada enquanto a tarefa receptora não executar uma operação de recebimento que retire dados suficientes do buffer do sistema. Desta forma algum *overhead* de sincronismo pode incorrer na tarefa emissora.

Quando o tamanho da mensagem é maior do que o TRESHOLD, o comportamento da operação MPI_SEND é essencialmente o mesmo do modo síncrono.

A idéia por trás dessa mudança de comportamento de acordo com o tamanho da mensagem é que mensagens pequenas se beneficiam da menor chance de *overhead* de sincronismo utilizando o buffer do sistema. Por outro lado, conforme o tamanho da mensagem aumenta, o custo de copiar para o buffer também aumenta e além disso pode ser impossível contar com espaço suficiente no buffer do sistema. Desta forma, o modo padrão tenta executar a melhor estratégia.

2 Definição do Sistema

A motivação deste trabalho surgiu de um conjunto de três fatores:

- interesse em se relacionar e avaliar o desempenho dos quatro tipos de modo de comunicação disponíveis em MPI para a operação de envio com bloqueio.
- interesse em se estudar o desempenho da biblioteca MPI no ambiente de programação paralela recentemente criado no Laboratório de Informática da PUC - um cluster de processadores Intel.
- interesse em se avaliar o desempenho deste novo ambiente em relação a outro ambiente de programação paralela, normalmente utilizado pelo grupo de paralelismo - o SP2 do LNCC.

Pensando assim, dois estudos de caso serão apresentados neste trabalho:

- Estudo MPI NO CLUSTER
- Estudo CLUSTER $\times$ SP2

3 Métrica

O tempo mínimo gasto por uma operação de envio com bloqueio para um determinado tamanho de mensagem. É válido dizer que esta não é uma medida do tempo necessário para completar a comunicação, nem uma medida do tempo total de *overhead* do sistema. O programa irá reportar o tempo gasto (*wallclock*) somente pelas **operações de envio** com bloqueio para os modos de comunicação selecionados. Todas as operações de recebimento serão colocadas antes que qualquer operação de envio seja executada.

4 Parâmetros

A variável que representa o TRESHOLD (Seção 1.1) é uma variável de ambiente que afeta o comportamento da troca de mensagens em implementações MPI. Dependendo do seu valor uma operação no modo padrão pode ter um comportamento bufferizado ou síncrono. Para as duas implementações de MPI utilizadas neste trabalho, os valores dos TRESHOLDS são: 4k para a versão da IBM e 16k para a versão LAM/MPI.

A carga na rede e a ordem das tarefas de envio e recebimento são parâmetros conhecidos que também afetam o comportamento da troca de mensagens. A fim de delimitar o objeto de estudo deste trabalho estes dois parâmetros foram muito bem controlados: utilizou-se máquinas em modo exclusivo e as tarefas de recebimento sempre antecedem as de envio.

5 Fatores

A escolha dos níveis de cada fator para os dois estudo foi realizada tendo como modelo as características das aplicações paralelas desenvolvidas no curso até então.

1. Modo de Comunicação

    - síncrono
    - pronto
    - bufferizado

    Uma análise preliminar dos dados mostrou que o envio padrão apresentou um comportamento exatamente como definido em [1]: mensagens de tamanho igual ou menor do que o TRESHOLD levam um tempo parecido com o que levaria uma mensagem no modo bufferizado e mensagens maiores do que este TRESHOLD levam um tempo parecido com o que levaria uma mensagem no modo síncrono. Por isso, este modo de comunicação não foi incluído em nenhum estudo, seu comportamento está bem definido e conhecendo-se o comportamento dos modos síncrono e bufferizado é possível estimá-lo.

2. Tamanho da Mensagem

   - real (4 bytes) de tamanho 0 = 0 bytes. Considerado como o tamanho mínimo que uma mensagem pode assumir (*header*).
   - real (4 bytes) de tamanho 5120 = 20480 bytes = 20k. Considerado, para a maioria das aplicações de interesse, um tamanho "médio/grande", pode-se dizer realista.
   - real (4 bytes) de tamanho 262144 = 1048576 bytes = 1M. Considerado, também para a maioria das aplicações de interesse, um tamanho "muito grande", quase irreal.

3. Ambientes Paralelos

   - Cluster Intel do Laboratório de Informática da PUC (**Cluster**): *workstation* de 32 processadores pentium II 400 Mhz com 32 Mb de memória cada e um disco de 6 Gb. As estações estão conectadas por uma rede ethernet através de um *switch* IBM 8274 modelo W93 na placa ESM-C-32W que permite 32 segmentos ethernets a 10 Mbps. O sistema operacional é o LINUX e a implementação MPI é a versão LAM/MPI 6.1. Os dois processadores utilizados nos testes estavam em modo exclusivo.
   - Máquina paralela IBM SP2 do LNCC (**SP2**): sistema de multiprocessadores de memória distribuída que utiliza a troca de mensagens como esquema de comunicação. Os dois processadores do cluster do SP2 utilizados nos testes deste trabalho são do tipo THIN (RS6000 mod. 390), estavam em modo exclusivo, e possuem as seguintes características:
     - sistema operacional AIX 4.1.4
     - 256 Mb RAM
     - desempenho de 266 MFlops
     - 16 procs com 2.2 Gb de disco local e /tmp 150 Mb
     - implementação MPI versão IBM

     Para realizar a troca de mensagens pela rede entre os processadores, utilizou-se o dispositivo HPS (*High Performance Switch*, 80 Mbytes/s) com o protocolo TCP/IP.

6 Carga de Trabalho

O programa EnvioComBloqueio será executado em exatamente dois processadores, identificados como nó 1 e nó 2. O processo no nó 1 será responsável por executar três operações de envio relacionadas aos três modos de comunicação selecionados e monitorar o tempo. O processo no nó 2 irá executar três operações de recebimento. É importante dizer que as operações de recebimento antecedem as operações de envio. Isto porque deseja-se avaliar somente o desempenho das operações de envio. A Figura 1 descreve o pseudo-código do programa EnvioComBloqueio.

```
procedure EnvioComBloqueio(tamanho_msg)
Inicializa MPI;
Inicializa variáveis;
se processador = nó 1 então
   Inicializa conteúdo das mensagens;
   Recebimento com bloqueio de msg. "pronto para receber" do nó 2;
   Marca tempo inicial do envio sincronizado;
   Realiza envio com bloqueio sincronizado para nó 2;
   Marca tempo final do envio sincronizado;
   Imprime diferença entre tempos inicial e final do envio sincronizado;
   Marca tempo inicial do envio pronto;
   Realiza envio com bloqueio pronto para nó 2;
   Marca tempo final do envio pronto;
   Imprime diferença entre tempos inicial e final do envio pronto;
   Aloca buffer para envio bufferizado;
   Marca tempo inicial do envio bufferizado;
   Realiza envio com bloqueio bufferizado para nó 2;
   Marca tempo final do envio bufferizado;
   Imprime diferença entre tempos inicial e final do envio bufferizado
senão se processador = nó 2 então
   Coloca 3 recebimentos sem bloqueio;
   Envio com bloqueio de msg. "pronto para receber" para nó 1;
   Espera todos os recebimentos finalizarem;
fim se
fim EnvioComBloqueio.
```

Figura 1: Pseudo-código do programa EnvioComBloqueio

7 Projeto Experimental

7.1 Estudo de Caso: MPI no Cluster

Existem diversos tipos de projeto experimental. O primeiro experimento realizado neste trabalho utiliza o *Projeto Completo de Dois-Fatores com Replicações*, conforme apresentado em Jain [2], para analisar o desempenho do estudo de caso MPI no Cluster. Neste tipo de projeto dois parâmetros (fatores) são variados e controlados de forma cuidadosa a fim de se estudar o impacto que eles causam no desempenho do sistema. Além disso, repetir os experimentos possibilita determinar se as interações entre os fatores são significantes ou se, simplesmente, as diferenças observadas são devido às variações aleatórias devido a erros de medida e parâmetros não controlados.

Um projeto experimental fatorial completo com dois fatores A e B com a e b níveis e r repetições para cada uma das ab combinações fator/nível demanda abr experimentos.

Os fatores do estudo MPI no Cluster são: *modo de comunicação* com 3 níveis e *tamanho da mensagem* com 3 níveis. A Tabela 1 apresenta os fatores e seus níveis. Cada experimento será repetido 10 vezes. O projeto consiste de um total de (3 modos de comunicação) X (3 tamanhos de mensagem) = 9 experimentos com 10 repetições cada, totalizando 90 observações, para as quais o tempo decorrido em microsegundos (μs) será medido.

Símbolo	Fator	Nível 1	Nível 2	Nível 3
A	Modo de Comunicação	síncrona	pronta	bufferizada
B	Tamanho da Mensagem	0 bytes	20k	1M

Tabela 1: Fatores e seus níveis utilizados no estudo MPI no Cluster

Para os dois fatores A e B do estudo MPI no Cluster nos níveis a e b respectivamente, e r repetições, o modelo é

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk} \tag{1}$$
$$i = 1, \ldots, a; j = 1, \ldots, b; k = 1, \ldots, r;$$

onde

y_{ijk} = resposta (observação) na k-ésima replicação do experimento com os fatores A e B nos níveis a e b respectivamente.

μ = média do tempo decorrido (observação).

α_i = efeito do fator A (*modo de comunicação*) no nível i.

β_j = efeito do fator B (*tamanho da mensagem*) no nível j.

γ_{ij} = efeito da interação entre os fatores A e B nos níveis i e j.

e_{ijk} = erro experimental na k-ésima replicação do experimento com os fatores A e B nos níveis a e b.

Os efeitos dos fatores são calculados de forma que sua soma seja igual a zero:

$$\sum_{i=1}^{a} \alpha_i = 0$$
$$\sum_{j=1}^{b} \beta_j = 0$$

Os efeitos das interações são calculados de forma que a soma dos valores dos efeitos para um mesmo fator A ou um mesmo fator B seja igual a zero:

$$\sum_{j=1}^{b} \gamma_{1j} = \sum_{j=1}^{b} \gamma_{2j} = \ldots = \sum_{j=1}^{b} \gamma_{aj} = 0$$
$$\sum_{i=1}^{a} \gamma_{i1} = \sum_{i=1}^{a} \gamma_{i2} = \ldots = \sum_{i=1}^{a} \gamma_{ib} = 0$$

E desejamos que o erro médio em cada experimento seja igual a zero

$$\sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{r} e_{ijk} = 0$$

É importante dizer que este modelo assume que os efeitos dos dois fatores somam e os erros são aditivos.

7.1.1 Cálculo dos Efeitos

Os dados medidos no estudo MPI NO CLUSTER consistem de r observações para cada uma das ab alternativas do modelo. A Tabela 2 relaciona estes dados. Observa-se que os valores da tabela (tempos em microsegundos) possuem um grande espalhamento, a razão entre o maior e o menor valor y_{max}/y_{min} é 1189302/33, aproximadamente 36039. Segundo Jain [2], não é apropriado, estatisticamente, adicionar respostas que variam tanto como estas. Neste caso, se faz necessário o uso de um modelo multiplicativo para a análise dos dados. Utilizando-se uma transformação logarítmica dos valores observados, os cálculos para o modelo multiplicativo são feitos exatamente da mesma forma como no modelo aditivo. A Tabela 3 * relaciona a transformação logarítmica dos dados da Tabela 2.

Para cada uma das $abr = 3 * 3 * 10 = 90$ observações existe uma equação do tipo

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}$$

Adicionando-se estas equações obtemos

$$\sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{r} y_{ijk} = rba\mu + rb\sum_{i=1}^{a} \alpha_i + ra\sum_{j=1}^{b} \beta_j + r\sum_{i=1}^{a} \sum_{j=1}^{b} \gamma_{ij} + \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{r} e_{ijk}$$

*Utilizou-se, de forma equivocada, ponto ao invés de vírgula para separar casas decimais para todos os valores numéricos deste trabalho.

Uma vez que, pelo modelo, a soma dos efeitos α_i, β_j e das interações entre os efeitos γ_{ij} é igual a zero; e queremos que o erro médio também seja zero, a equação anterior fica

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r} y_{ijk} = rba\mu + 0 + 0 + 0 + 0$$

Desta forma o parâmetro μ do modelo é calculado da seguinte forma

$$\mu = \frac{1}{rba}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r} y_{ijk}$$

O valor do lado direito da equação é conhecido como a média geral dos valores observados de todas as replicações de todos os experimentosde (abr observações) e é denotado por $\overline{y}_{...}$. Os pontos no subscrito indicam a dimensão (fator A ou B, ou replicação) em torno da qual a média é calculada. Desta forma, a média de todas as observações onde o nível do fator A igual a i é denotada por $\overline{y}_{i..}$ e é calculada da seguinte forma

$$\overline{y}_{i..} = \mu + \alpha_i + \frac{1}{br}\sum_{j=1}^{b}\beta_j + \frac{1}{br}\sum_{j=1}^{b}\gamma_{ij} + \frac{1}{br}\sum_{j=1}^{b}\sum_{k=1}^{r} e_{ijk}$$

Uma vez que os últimos três termos são iguais a zero, temos

$$\overline{y}_{i..} = \mu + \alpha_i$$

De forma similar, temos

$$\begin{aligned} \overline{y}_{.j.} &= \mu + \beta_j \\ \overline{y}_{...} &= \mu \end{aligned}$$

Desta forma os valores estimados dos parâmetros do modelo são

$$\begin{aligned} \mu &= \overline{y}_{...} \\ \alpha_i &= \overline{y}_{i..} - \overline{y}_{...} \\ \beta_j &= \overline{y}_{.j.} - \overline{y}_{...} \\ \gamma_{ij} &= \overline{y}_{ij.} - \overline{y}_{i..} - \overline{y}_{.j.} + \overline{y}_{...} \end{aligned}$$

A partir das médias dos logarítimos de todas as observações de cada experimento é possível calcular os efeitos multiplicativos dos vários fatores e suas interações. A Tabela 4 apresenta as médias dos logarítimos dos 9 experimentos e os efeitos multiplicativo dos fatores. Verifique que a soma dos efeitos é igual a zero. Um envio em um *modo de comunicação* do tipo médio de uma mensagem de tamanho médio é dado por $\overline{y}_{...}$ e necessita de um tempo aproximado de log 3.5370 (3444 microsegundos).

		tamanho da mensagem			
		0 bytes	20k	1M	
Com.	i	$j=1$	$j=2$	$j=3$	k
		688	32203	1084000	1
		684	32890	1075834	2
		689	37574	1115845	3
		680	29895	1056873	4
sínc.	1	683	29871	1045385	5
		683	34793	1115845	6
		689	30569	1056873	7
		688	34026	1045385	9
		685	30139	1018173	9
		688	37896	1072839	10
		33	362	1073145	1
		33	384	1189302	2
		33	354	1183247	3
		33	374	1103370	4
pronto	2	33	350	1187491	5
		33	357	1183247	6
		32	354	1103370	7
		33	364	1187491	8
		33	360	1139768	9
		33	357	1128463	10
		40	361	15365	1
		40	352	15317	2
		41	337	15301	3
		40	335	15396	4
buff.	3	41	333	15180	5
		40	327	15301	6
		40	341	15196	7
		40	338	15180	8
		40	343	15303	9
		39	325	15167	10

Tabela 2: Tempos em microsegundos para cada combinação fator/nível, replicada 10 vezes

- Fator Modo de Comunicação: Um envio no *modo de comunicação* síncrono necessita aproximadamente de um tempo log 0.92 μs a mais do que um envio médio (isto é um fator de 8.39 vezes), como um envio médio necessita de aproximadamente 3444 μs, um envio síncrono necesssita aproximadamente de 28874 μs, 25430 μs a mais do que um envio médio. De forma similar, um envio nos modos de comunicação pronto e bufferizado necessitam de um tempo log 0.16 e log 0.76 μs a menos do que um envio médio (fatores 0.69 e 0.17) que correspondem respectivamente a um tempo aproximado de 2389 e 592 μs (respectivamente 1054 e 2852 μs a menos do que o envio médio). Em outras palavras, o tempo de envio no *modo de comunicação* síncrono é aproximadamente 49 vezes ($10^{\alpha_1-\alpha_3}$) maior do que o tempo no *modo* bufferizado e aproximadamente 12 vezes ($10^{\alpha_1-\alpha_2}$) maior do que o tempo no *modo* pronto.

- Fator Tamanho da Mensagem: Um envio de uma mensagem de 0 bytes necessita de um tempo log 1.55 μs a menos do que um envio médio (isto é,

		tamanho da mensagem			
		0 bytes	20k	1M	
Com.	i	$j=1$	$j=2$	$j=3$	k
sínc.	1	2.8376	4.5079	6.0350	1
		2.8351	4.5171	6.0317	2
		2.8382	4.5749	6.0476	3
		2.8325	4.4756	6.0240	4
		2.8344	4.4752	6.0193	5
		2.8344	4.5415	6.0476	6
		2.8382	4.4853	6.0240	7
		2.8376	4.5318	6.0193	8
		2.8357	4.4791	6.0078	9
		2.8376	4.5786	6.0305	10
pronto	2	1.5185	2.5587	6.0307	1
		1.5185	2.5843	6.0753	2
		1.5185	2.5490	6.0731	3
		1.5185	2.5729	6.0427	4
		1.5185	2.5441	6.0746	5
		1.5185	2.5527	6.0731	6
		1.5051	2.5490	6.0427	7
		1.5185	2.5611	6.0746	8
		1.5185	2.5563	6.0568	9
		1.5185	2.5527	6.0525	10
buff.	3	1.6021	2.5575	4.1865	1
		1.6021	2.5465	4.1852	2
		1.6128	2.5276	4.1847	3
		1.6021	2.5250	4.1817	4
		1.6128	2.5224	4.1813	5
		1.6021	2.5145	4.1847	6
		1.6021	2.5328	4.1817	7
		1.6021	2.5289	4.1813	8
		1.6021	2.5353	4.1848	9
		1.5911	2.5119	4.1809	10

Tabela 3: Transformação logarítmica dos resultados da Tabela 2

um fator de 0.03 vezes), que corresponde a um tempo de 97 μs (3347 μs a menos). Um envio de uma mensagem de 20k necessita de um tempo log 0.34 μs a menos do que um envio médio (fator 0.46) que corresponde a um envio de 1591 μs aproximadamente (1853 μs a menos do que o envio médio). Um envio de uma mensagem de 1M necessita de um tempo log 1.89 μs a mais do que um envio médio (fator 77.06), logo um envio de uma mensagem de 1M necessita aproximadamente de 265376 μs (261933 μs a mais do que um envio médio). De outra forma, o tempo de um envio de uma mensagem de 1M é 167 vezes ($10^{\beta_3-\beta_2}$) maior do que o tempo de envio de uma mensagem de tamanho 20k e 2744 vezes ($10^{\beta_3-\beta_1}$) maior do que o tempo de envio de uma mensagem de tamanho 0 bytes.

Além desses efeitos, existem nove interações de fatores dois a dois (Tabela 5), calculadas a seguir:

$$\gamma_{11} = \overline{y}_{11.} - \overline{y}_{1..} - \overline{y}_{.1.} + \overline{y}_{...} = 2.8361 - 4.4605 - 1.9855 + 3.5370 = -0.0728$$

$$\gamma_{21} = \overline{y}_{21.} - \overline{y}_{2..} - \overline{y}_{.1.} + \overline{y}_{...} = 1.5172 - 3.3783 - 1.9855 + 3.5370 = -0.3096$$

			tamanho da mensagem				
			0 bytes	20k	1M		
Com.	i	k	$j=1$	$j=2$	$j=3$	$j=.$	$\alpha_i=\overline{y}_{i..}-\overline{y}_{...}$
sínc.	1	.	$\overline{y}_{11.}=2.8361$	$\overline{y}_{12.}=4.5167$	$\overline{y}_{13.}=6.0287$	$\overline{y}_{1..}=4.4605$	$\alpha_1=0.92$
pronto	2	.	$\overline{y}_{21.}=1.5172$	$\overline{y}_{22.}=2.5581$	$\overline{y}_{23.}=6.0596$	$\overline{y}_{2..}=3.3783$	$\alpha_2=-0.16$
buff.	3	.	$\overline{y}_{31.}=1.6031$	$\overline{y}_{32.}=2.5305$	$\overline{y}_{33.}=4.1833$	$\overline{y}_{3..}=2.7722$	$\alpha_3=-0.76$
	.	.	$\overline{y}_{.1.}=1.9855$	$\overline{y}_{.2.}=3.2017$	$\overline{y}_{.3.}=5.4239$	$\overline{y}_{...}=3.5370$	$\sum\alpha_i=0$
$\beta_j=\overline{y}_{.j.}-\overline{y}_{...}$			$\beta_1=-1.55$	$\beta_2=-0.34$	$\beta_3=1.89$	$\sum\beta_j=0$	

Tabela 4: Cálculo dos efeitos multiplicativos dos fatores para o estudo MPI no Cluster

$$\gamma_{31}=\overline{y}_{31.}-\overline{y}_{3..}-\overline{y}_{.1.}+\overline{y}_{...}=1.6031-2.7722-1.9855+3.5370=0.3824$$
$$\gamma_{12}=\overline{y}_{12.}-\overline{y}_{1..}-\overline{y}_{.2.}+\overline{y}_{...}=4.5167-4.4605-3.2017+3.5370=0.3915$$
$$\gamma_{22}=\overline{y}_{22.}-\overline{y}_{2..}-\overline{y}_{.2.}+\overline{y}_{...}=2.5581-3.3783-3.2017+3.5370=-0.4849$$
$$\gamma_{32}=\overline{y}_{32.}-\overline{y}_{3..}-\overline{y}_{.2.}+\overline{y}_{...}=2.5305-2.7722-3.2017+3.5370=0.0934$$
$$\gamma_{13}=\overline{y}_{13.}-\overline{y}_{1..}-\overline{y}_{.3.}+\overline{y}_{...}=6.0287-4.4605-5.4239+3.5370=-0.3187$$
$$\gamma_{23}=\overline{y}_{23.}-\overline{y}_{2..}-\overline{y}_{.3.}+\overline{y}_{...}=6.0596-3.3783-5.4239+3.5370=0.7945$$
$$\gamma_{33}=\overline{y}_{33.}-\overline{y}_{3..}-\overline{y}_{.3.}+\overline{y}_{...}=4.1833-2.7722-5.4239+3.5370=-0.4758$$

		tamanho da mensagem			
		0 bytes	20k	1M	
Com.	i	$j=1$	$j=2$	$j=3$	$\sum\gamma ij$
sínc.	1	$\gamma_{11}=$ -0.0728	$\gamma_{12}=$ 0.3915	$\gamma_{13}=$ -0.3187	0
pront	2	$\gamma_{21}=$ -0.3096	$\gamma_{22}=$ -0.4849	$\gamma_{23}=$ 0.7945	0
buff.	3	$\gamma_{31}=$ 0.3824	$\gamma_{32}=$ 0.0934	$\gamma_{33}=$ -0.4758	0
$\sum\gamma_{ij}$		0	0	0	

Tabela 5: Cálculo dos efeitos multiplicativos das interações entre os fatores

Os efeitos das interações são interpretados da seguinte forma. Considere a interação entre os fatores *modo de comunicação* síncrono e *tamanho da mensagem* de 1M, um envio leva log 0.3187 μs a menos do que levaria um envio em um modo comunicação médio de uma mensagem de tamanho 1M, ou log 0.3187 μs a menos do que levaria um envio de uma mensagem de tamanho médio no *modo de comunicação* síncrono. De outra forma, o envio de uma mensagem de tamanho 1M no *modo de comunicação* síncrono leva um tempo log 0.3187 μs a menos do que se esperaria de um envio considerando uma mensagem de tamanho 1M e o *modo de* *comunicação* síncrono. Quer dizer, existe uma interação entre os dois fatores cujo efeito desta combinação faz com que o tempo seja menor do que o previsto pelo modelo.

7.1.2 Estimativa dos Erros Experimentais

Após calcularmos os parâmetros do modelo, a resposta estimada para todas as observações no (i, j)-ésimo experimento é dada por

$$\hat{y}_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} = \overline{y}_{ij.}$$

Assim, o erro na k-ésima replicação do experimento é a diferença entre a resposta obtida e a resposta esperada (média de todas as replicações do experimento)

$$e_{ijk} = y_{ijk} - \overline{y}_{ij.}$$

Por exemplo, considere os dados das Tabelas 3 e 4, a média de todas as replicações do experimento (1,1) é 2.8361. Assim, os erros de cada replicação do experimento são $e_{111} = y_{111} - \overline{y}_{11.} = 2.8376 - 2.8361 = 0.0015$, $e_{112} = y_{112} - \overline{y}_{11.} = 2.8351 - 2.8361 = -0.0011$, $\ldots$, $e_{1110} = y_{1110} - \overline{y}_{11.} = 2.8376 - 2.8361 = 0.0015$. Os erros para as demais observações são calculados de forma análoga. Verifica-se que a soma dos erros para cada experimento é igual a zero, logo a soma total dos erros também é zero.

7.1.3 Alocação de Variação

A variação total de y é devida aos dois fatores, às interações entre eles e aos erros experimentais. Para calcular o seu valor eleva-se ao quadrado os dois lados da equação (1) do modelo e adiciona-se todas as observações. Os produtos cruzados dos termos se cancelam (somam zero) devido à restrição de que efeitos somam zero e erros somam zero, e obtém-se

$$\sum_{ijk} y_{ijk}^2 = abr\mu^2 + br\sum_i \alpha_i^2 + ar\sum_j \beta_j^2 + r\sum_{ij} \gamma_{ij}^2 + \sum_{ijk} e_{ijk}^2$$

A equação anterior pode ser escrita como

$$SSY = SS0 + SSA + SSB + SSAB + SSE$$

A variação total é

$$SST = SSY - SS0 = SSA + SSB + SSAB + SSE$$

Desta forma a variação total pode ser dividida em partes, explicadas pelos fatores A e B, à interação AB, e uma parte não explicada devido a erros experimentais. A percentagem da variação explicada por um fator ou interação pode ser usada para medir a importância do efeito correspondente. Para o estudo MPI NO CLUSTER as diversas somas de quadrados são:

$$SSY = \sum_{ijk} y_{ijk}^2 = (2.8376)^2 + (2.8351)^2 + \ldots + (4.1809)^2 = 1368.2796$$

$$SS0 = abr\mu^2 = 3 * 3 * 10 * (3.5370)^2 = 1125.9353$$

$$SSA = br \sum_i \alpha_i^2 = 3 * 10 * [(-1.55)^2 + (-0.34)^2 + (1.89)^2] = 43.8886$$

$$SSB = ar \sum_j \beta_j^2 = 3 * 10 * [(0.92)^2 + (-0.16)^2 + (-0.76)^2] = 182.398$$

$$SSAB = r \sum_{ij} \gamma_{ij}^2 = 10 * [(-0.0728)^2 + (-0.3096)^2 + \ldots + (-0.4758)^2] = 16.0362$$

$$SST = SSY - SS0 = 242.3443$$

A soma SSE pode ser obtida utilizando-se as outras somas como mostrado a seguir

$$SSE = SSY - SS0 - SSA - SSB - SSAB$$

$$= 1368.279 - 1125.9353 - 43.8886 - 182.398 - 16.0362 = 0.0215$$

A percentagem de variação explicada pelos vários fatores é

- explicada pelo *modo de comunicação*

$$100 \frac{SSA}{SST} = 100 \frac{43.8886}{242.3443} = 18.11\%$$

- explicada pelo *tamanho da mensagem*

$$100 \frac{SSB}{SST} = 100 \frac{182.398}{242.3443} = 75.26\%$$

- explicada pela interação *modo de comunicação/tamanho da mensagem*

$$100 \frac{SSAB}{SST} = 100 \frac{16.0362}{242.3443} = 6.62\%$$

- inexplicada

$$100\frac{SSE}{SST} = 100\frac{0.0215}{242.3443} = 0.01\%$$

Tamanho da mensagem é o fator que explica a maior variação, 75.26%. O fator *modo de comunicação* é menos importante pois explica 18.11% da variação. Os 6.62% de variação devido à interação não é desprezível. De alguma forma pode-se dizer que a seleção do *modo de comunicação* depende do *tamanho da mensagem.* Como o modelo multiplicativo é capaz de explicar 99.99% da variação, este parece ser um modelo adequado.

7.1.4 Análise de Variância

Na seção anterior, foi utilizada uma abordagem informal (porém útil, na prática) que aloca a variação a diferentes fatores. Neste tipo de abordagem qualquer fator que possua uma percentagem alta de variação é considerado importante. Esta importância, entretanto, deve ser distinguida da significância, que é uma abordagem estatística. Para verificar se um fator possui um efeito estatisticamente significante na resposta, compara-se a sua contribuição para a variação com a contribuição dos erros. Se a variação inexplicada (devido aos erros) for alta, um fator que explique uma grande fração da variação pode ser considerado estatisticamente insignificante. O procedimento estatístico utilizado para analisar a significância dos vários fatores é conhecido como *Analysis Of Variance* (ANOVA).

Para entender ANOVA, considere as somas de quadrados SSY, $SS0$, SSA, SSB, $SSAB$ e SSE. Cada uma das somas de quadrados possui um grau de liberdade associado que corresponde ao número de valores independentes necessários para o seu cálculo. Os graus de liberdade para as somas de quadrados são os seguintes:

$$SSY = SS0 + SSA + SSB + SSAB + SSE$$

$$abr = 1 + (a-1) + (b-1) + (a-1)(b-1) + ab(r-1)$$
$$90 = 1 + 2 + 2 + 4 + 81$$

A soma SSY consiste da soma de abr termos, todos podendo ser escolhidos de forma independente, portanto o grau de liberdade de SSY é abr. A soma $SS0$ consiste de um único termo μ^2 repetido abr vezes. Uma vez que um valor para μ tenha sido escolhido, $SS0$ pode ser calculado. Desta forma, $SS0$ possui grau de liberdade um. A soma SSA consiste da soma de a termos α_i^2, porém somente $a-1$ termos são independentes, já que os α_i's somam zero. A soma SSE consiste de abr fatores dos quais apenas $ab(r-1)$ podem ser escolhidos independentemente.

Isto porque os r erros devem somar zero. Assim, somente $(r-1)$ erros em cada um dos ab experimentos são independentes. Os graus de liberdade para as demais somas são justificados de forma similar. Observe que os graus de liberdade dos dois lados da equação somam zero.

Dado duas somas de quadrados SSi e SSj com v_i e v_j graus de liberdade, a razão $(SSi/v_i)/(SSj/v_j)$ possui uma distribuição F com numerador v_i graus de liberdade e denominador v_j graus de liberdade. A hipótese de que a soma SSi seja menor ou igual a soma SSj no nível de significância α é rejeitada se a razão for maior do que o quantil $(1-\alpha)$ da variável F. Assim, a razão calculada é comparada com $F[1-\alpha, v_i, v_j]$ (disponível de uma tabela de quantis de F) e as somas de quadrados são consideradas significantemente diferentes se o F calculado for maior do que o da tabela. Este procedimento é conhecido como Teste-F.

Assumindo-se que os erros possuem distribuição normal, SSE, SSA, SSB e $SSAB$ possuem distribuição *chi-square* †, o Teste-F pode ser usado para verificar se SSA, SSB e $SSAB$ são significantemente maiores que SSE. O valor (SSi/v_i) é chamado de Média do Quadrado de i. A seguir, o cálculo das Médias dos Quadrados para o estudo MPI NO CLUSTER

$$MSA = \frac{SSA}{a-1} = \frac{43.8886}{3-1} = 21.9443$$

$$MSB = \frac{SSB}{b-1} = \frac{182.398}{3-1} = 91.199$$

$$MSAB = \frac{SSAB}{(a-1)(b-1)} = \frac{16.0362}{(3-1)(3-1)} = 4.009$$

$$MSE = \frac{SSE}{ab(r-1)} = \frac{0.0215}{3*3*(10-1)} = 0.00026$$

$$\frac{MSA}{MSE} = \frac{21.9443}{0.00026} = 82554.03$$

$$\frac{MSB}{MSE} = \frac{91.199}{0.00026} = 343088.9$$

$$\frac{MSAB}{MSE} = \frac{4.009}{0.00026} = 15081.94$$

Para um nível de significância de 0.10, as razões MSA/MSE, MSB/MSE e $MSAB/MSE$ possuem distribuição $F[0.90, n, m]$ onde 90% é o nível de confiança desejado, n é o grau de liberdade do fator/interação correspondente e m é o grau de liberdade do erro. A Tabela de quantis de F disponível em Jain [2] não possui

† A soma de quadrados de variáveis independentes, normalmente distribuídas, com v graus de liberdade possui distribuição *chi-square* com v graus de liberdade.

valores para $m = 80$, neste caso será utilizado o valor mais alto disponível que é para $m = 40$.

A Tabela 6 sumariza os valores calculados. Observa-se que todas as razões calculadas (coluna F-Calculado) são maiores do que os valores correspondentes na coluna F-Tabela, então SSA, SSB e $SSAB$ são significantemente maiores que SSE e por isso assume-se que todos os fatores e suas interações são significantes no nível 0.10.

Componente	Soma dos Quadrados	Perc. da Variação	Graus de Liberdade	Quadrado Médio	F-Calculado	F-Tabela $F[0.90, n, m]$ $m = 40$
y	1368.279					
$y_{...}$	1125.9353					
$y - y_{...}$	242.3443	100				
Modo Com.	43.8886	18.11	2	21.9443	82554.03	2.44
Tam. Msg.	182.398	75.26	2	91.199	343088.9	2.44
Interação	16.0362	6.62	4	4.009	15081.94	2.09
Erros	0.0215	0.01	81	0.00026		

Tabela 6: Tabela ANOVA para o estudo MPI no Cluster

7.1.5 Intervalos de Confiança para os Efeitos

Parâmetro	Estimativa	Variância
μ	$\overline{y}_{...}$	s_e^2/abr
α_i	$\overline{y}_{i..} - \overline{y}_{...}$	$s_e^2(a-1)/abr$
β_j	$\overline{y}_{.j.} - \overline{y}_{...}$	$s_e^2(b-1)/abr$
γ_{ij}	$\overline{y}_{ij.} - \overline{y}_{i..} - \overline{y}_{.j.} + \overline{y}_{...}$	$s_e^2(a-1)(b-1)/abr$
s_e^2		$SSE/ab(r-1) = MSE$
$s_e = \sqrt{MSE} = \sqrt{0.00026} = 0.0163$		
Grau de liberdade para os erros $= ab(r-1)$		

Tabela 7: Estimativa dos parâmetros do modelo

Os valores estimados dos parâmetros do modelo são variáveis aleatórias baseadas em uma amostra. Certamente, valores estimados baseados em outra amostra gerariam resultados diferentes. A variância dos parâmetros do modelo pode ser calculada a partir da variância dos erros utilizando as expressões da Tabela 7. A variância dos erros s_e^2 é o já visto quadrado médio MSE. O desvio padrão dos erros s_e simplesmente é a raiz quadrada da sua variância MSE. Como o grau de liberdade dos erros é superior a 30, o intervalo de confiança para os efeitos/interações pode ser calculado utilizando-se os quantis de uma distribuição normal para $ab(r-1)$ graus de liberdade. Considerando-se um intervalo de confiança de dois lados de

95%, tem-se $\alpha = 0.05$, $1 - \alpha/2 = 0.975$ e o quantil $z_{0.975} = 1.960$. Por exemplo, o intervalo de confiança para o fator *modo de comunicação* síncrona (efeito $\alpha_1 = 0.92$ e variância $s_{\alpha_1} = 0.024$) é

$$\alpha_1 \mp (z_{0.975} * s_{\alpha_1}) = 0.92 \mp (1.960 * 0.024) = 0.92 \mp 0.00476 = (0.9187, 0.9283)$$

Os demais intervalos são calculados de forma similar e estão relacionados nas Tabelas 8 e 9.

Parâmetro	Efeito Médio	Desvio Padrão	Intervalo de Confiança
μ	3.5370	$s_\mu = s_e\sqrt{\frac{1}{abr}} = 0.0163\sqrt{\frac{1}{90}} = 0.0017$	$(3.5336, 3.5404)$
Modo de Comunicação			
sinc.	0.92	$s_{\alpha_i} = s_e\sqrt{\frac{(a-1)}{abr}} = 0.0163\sqrt{\frac{2}{90}} = 0.0024$	$(0.9228, 0.9242)$
pronto	-0.16		$(-0.1635, -0.1540)$
buff.	-0.76		$(-0.7696, -0.7600)$
Tamanho da Mensagem			
0 bytes	-1.55	$s_{\beta_j} = s_e\sqrt{\frac{(b-1)}{abr}} = 0.0163\sqrt{\frac{2}{90}} = 0.0024$	$(-1.5563, -1.5468)$
20k	-0.34		$(-0.3401, -0.3306)$
1M	1.89		$(1.8821, 1.8916)$
$z_{0.975} = 1.960$			

Tabela 8: Intervalo de confiança para os efeitos no estudo MPI no Cluster

Modo/Tam.	0 bytes	20k	1M
sinc.	$(-0.0796, -0.0661)$	$(0.3848, 0.3983)$	$(-0.3254, -0.3119)$
pronto	$(-0.3163, -0.3028)$	$(-0.4916, -0.4782)$	$(0.7877, 0.8012)$
buff.	$(-0.3757, 0.3892)$	$(0.0866, 0.1001)$	$(-0.4825, -0.4691)$
$s_{\gamma_{ij}} = \sqrt{s_e^2\frac{(a-1)(b-1)}{abr}} = s_e\sqrt{\frac{(a-1)(b-1)}{abr}} = 0.0163\sqrt{\frac{4}{90}} = 0.0034$			
$z_{0.975} = 1.960$			

Tabela 9: Intervalo de confiança para as interações no estudo MPI no Cluster

Como nenhum intervalo inclui o zero, todos os efeitos são considerados significantes. Além disso, percebe-se que os intervalos são bastante estreitos. Devido às replicações foi possível isolar as interações dos erros experimentais. Ao mesmo tempo, os erros experimentais são muito pequenos e pode-se confiar nos limites dos efeitos previstos pelo modelo.

7.1.6 Testes Visuais

Na análise dos dados, as seguintes suposições foram feitas:

1. Os erros do modelo são estatísticamente independentes.
2. Os erros possuem uma distribuição normal.
3. Os erros possuem um desvio padrão constante.

Se alguma dessas suposições for violada, as conclusões baseadas no modelo não são válidas. Ao contrário dos testes estatísticos, os testes visuais são aproximados. Entretanto, podem ser úteis para identificar algum problema no modelo.

- Teste da Independência dos Erros: para cada experimento y_{ijk}, deve-se computar o erro e_{ijk} e a seguir plotar em um gráfico e_{ijk} versus a resposta estimada $\hat{y}_{ij}$. Qualquer tendência visível no gráfico é um indicador de que os erros são dependentes e que o modelo é inadequado.

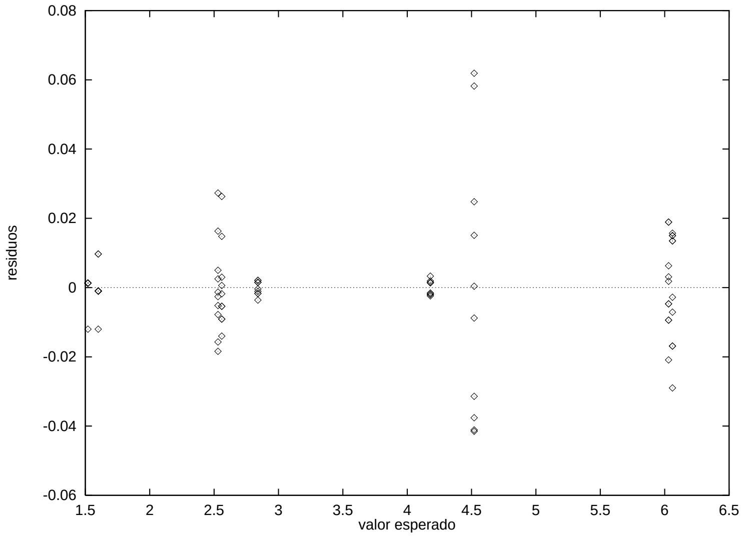

Figura 2: Gráfico dos resíduos e_{ijk} versus o valor esperado $\hat{y}_{ij}$ para o modelo multiplicativo

A Figura 2 mostra o resultado do teste de independência dos erros para o estudo MPI no Cluster. Observa-se que os valores dos eixos vertical e horizontal diferem de duas ordens de magnitude e os resíduos são pequenos se comparados com os valores esperados. É possível dizer que não parece existir nenhuma tendência definida nos pontos do gráfico, e, portanto, afirmar a suposição de que os erros são independentes, possuem a mesma distribuição (o que também implica na mesma variância de todos os fatores).

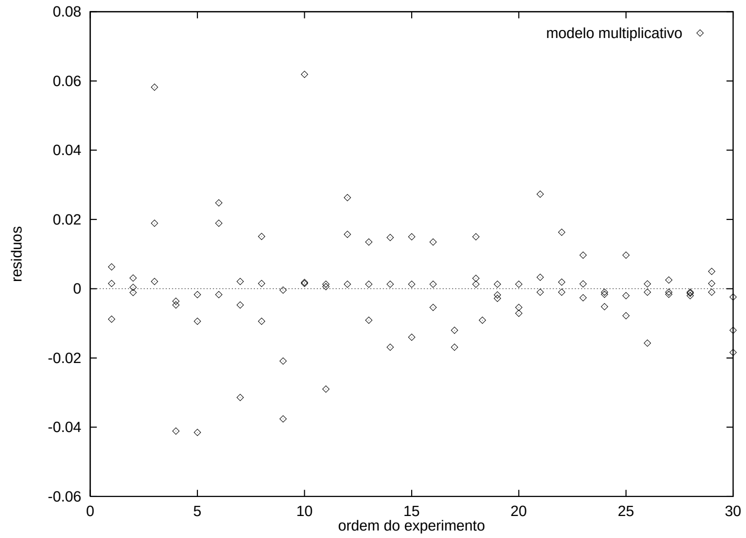

Figura 3: Gráfico dos resíduos e_{ijk} versus a ordem de execução do experimento para o modelo multiplicativo

- Teste para verificar efeitos colaterais: um segundo teste é plotar os resíduos em função do número do experimento pela ordem em que foram conduzidos. Qualquer tendência verificada é um indicador da presença de fatores não considerados pelo sistema ou efeitos colaterais (inicialização incorreta) que variam de um experimento para o outro e afetam a resposta. Para os testes realizados a ordem disponível é a ordem para cada tamanho de mensagem, como existem três tamanhos e 90 observações, o eixo horizontal do gráfico da Figura 3 varia de 1 a 30, e também parece não mostrar nenhuma tendência.

Para demais testes visuais veja em Jain [2].

7.1.7 Conclusão

O primeiro objetivo deste trabalho foi analisar o desempenho dos diferentes modos de comunicação para um envio com bloqueio em MPI no Cluster Intel do Laboratório de Informática (estudo MPI NO CLUSTER).

Inicialmente foram coletados dados usando a operação de envio com bloqueio para todos os tipos de modo de comunicação, nos dois ambientes paralelos, e para uma grande variedade de tamanhos de mensagens (0 bytes, 4k, 16k, 20k, 100k, 1M, 30M e 50M). Uma análise informal dos dados mostrou que o envio bufferizado sempre era mais rápido do que o envio síncrono, exceto para mensagens de 50M (tamanho não considerado em nenhum estudo por ser "irreal"). Considera-se que o *modo de comunicação* pronto, no total, possui o menor *overhead*. Primeiro, não é necessário acontecer o “aperto de mãos” entre as tarefas emissora e receptora (como no modo síncrono), além disso, não existe a necessidade da cópia extra para um buffer (como nos modos bufferrizado e padrão). No entanto, esta vantagem não teve o efeito esperado.

Na análise realizada para o estudo MPI NO CLUSTER foi possível observar muitas dessas tendências discutidas acima. Os efeitos dos fatores principais confirmam que, comparando-se com um envio médio, o modo de comunicação bufferizado é o mais rápido e o síncrono é o mais lento. Também parece óbvio dizer que, comparando-se com um envio médio, o envio de uma mensagem de 0 bytes é mais rápido e o envio de uma mensagem de 1M é mais lento. Quanto ao modo pronto, este é mais rápido do que o síncrono e mais lento do que o bufferizado se comparado com o envio médio. O que inicialmente parecia incoerente com a afirmação de que este modo é o que possui menor *overhead*, na verdade tem uma outra explicação: a cópia para a rede (*modo* pronto) é mais demorada do que a cópia para o buffer fornecido pelo usuário (*modo* bufferizado) [1]. De alguma forma o efeito da interação modo de comunicação e tamanho da mensagem foi explicado pelo modelo, pois de fato existe uma relação forte entre estes dois fatores. E pode-se dizer que esta relação está ligada à relação entre o *overhead* de sincronismo e o *overhead* de sistema, explicados na Seção 1. Provavelmente, o efeito da interação seria melhor avaliado para tamanhos de mensagem maiores do que os selecionados no estudo.

Quer dizer, apesar do modo síncrono ser o mais “seguro” [‡], e por isso mesmo o mais portável, verificou-se que o *overhead* de sincronismo incorrido por este modo é significativo. O modo síncrono é dito “seguro” pois não depende da ordem em que as tarefas emissora e receptora são executadas (diferentemente do modo pronto), nem da quantidade de espaço do buffer (diferentemente dos modos bufferizado e padrão). Por outro lado, o modo bufferizado desassocia as tarefas receptora e emissora eliminando assim, o *overhead* de sincronismo na tarefa emissora e garantindo que a ordem das tarefas não faça diferença no correto funcionamento do programa (diferentemente do modo pronto). O *overhead* de sistema necessário para copiar os dados da mensagem do buffer da tarefa de envio para a rede e da rede para o buffer da tarefa de recebimento só será sentido pela tarefa de recebimento. Esta talvez seja a principal razão do modo de comunicação bufferizado ter apresentado o melhor desempenho na análise realizada para o estudo MPI NO CLUSTER. Apesar do modo pronto não ter demonstrado a vantagem de tempo esperada, avaliar o seu desempenho é bastante útil pois ele dá uma noção do tempo mínimo de *overhead* de sistema. É interessante observar que como, para mensagens grandes, o comportamento do modo pronto foi parecido com o do modo síncrono, isto é uma indicação de que nestes casos o *overhead* de sincronismo não é tão significativo quanto o tempo gasto com a cópia dos dados através da rede.

[‡] “Seguro” significa que se um código executa sobre determinadas condições (tamanho de mensagem, arquitetura, etc.) ele executará sobre todas as condições.

7.2 Estudo de Caso: Cluster $\times$ SP2

O segundo experimento realizado neste trabalho utiliza o *Projeto Completo de k-Fatores com Replicações*, conforme apresentado em Jain [2], para analisar o desempenho do estudo de caso CLUSTER $\times$ SP2.

Um projeto experimental fatorial completo utiliza todas as possíveis combinações de todos os fatores em todos os níveis. Um estudo de desempenho com k fatores, com n_i níveis para o i-ésimo fator, demanda n experimentos, onde

$$n = \prod_{i=1}^{k} n_i$$

Para ser possível estimar erros experimentais se faz necessário repetir os experimentos. Se tivermos r repetições para cada uma das n combinações fator/nível, teremos um projeto com nr observações.

Os fatores do estudo CLUSTER $\times$ SP2 são: *modo de comunicação* com 2 níveis, *tamanho da mensagem* com 2 níveis e *ambiente paralelo* com 2 níveis. A Tabela 10 apresenta os fatores e seus níveis. Cada experimento será repetido 10 vezes. O projeto consiste de um total de (2 modos de comunicação) X (2 tamanhos de mensagem) X (2 ambientes paralelos) = 8 experimentos com 10 repetições cada, totalizando 80 observações, para as quais o tempo decorrido em μs será medido.

Símbolo	Fator	Nível 1	Nível 2
A	Modo de Comunicação	síncrona	bufferizada
B	Tamanho da Mensagem	0 bytes	20k
C	Ambiente Paralelo	Cluster	SP2

Tabela 10: Fatores e seus níveis utilizados no estudo CLUSTER $\times$ SP2

O modelo genérico para um projeto k-fatorial completo possui $2^k - 1$ efeitos. Esses efeitos incluem k efeitos principais, $\binom{k}{2}$ interações entre fatores dois a dois, $\binom{k}{3}$ interações entre fatores três a três, e assim por diante. Para os três fatores A, B, C do estudo CLUSTER $\times$ SP2 nos níveis a, b e c respectivamente, e r repetições, o modelo é

$$\begin{gathered} y_{ijkl} = \mu + \alpha_i + \beta_j + \epsilon_k + \gamma_{ABij} + \gamma_{ACik} + \gamma_{BCjk} + \gamma_{ABCijk} + e_{ijkl} \\ i = 1, \ldots, a; j = 1, \ldots, b; k = 1, \ldots, b; l = 1, \ldots, r \end{gathered}$$

onde

y_{ijkl} = resposta (observação) na l-ésima replicação do experimento com os fatores A, B e C nos níveis a, b e c respectivamente.

μ = tempo decorrido (observação) médio.

α_i = efeito do fator A (*modo de comunicação*) no nível i.

β_j = efeito do fator B (*tamanho da mensagem*) no nível j.

ϵ_k = efeito do fator C (*ambiente paralelo*) no nível k.

γ_{ABij} = interação entre os fatores A e B nos níveis i e j.

γ_{ABCijk} = interação entre os fatores A, B e C nos níveis i, j e k.

e_{ijkl} = erro experimental na l-ésima replicação do experimento com os fatores A, B, e C nos níveis a, b e c respectivamente.

e assim por diante.

7.2.1 Cálculo dos Efeitos

		Cluster $k=1$		SP2 $k=2$		
		0 bytes	20k	0 bytes	20 k	
Com.	i	$j=1$	$j=2$	$j=1$	$j=2$	l
sínc.	1	688	32203	747	1597	1
		684	32890	834	1639	2
		689	37574	746	1768	3
		680	29895	814	1594	4
		683	29871	750	2366	5
		683	34793	752	1656	6
		689	30569	733	1607	7
		688	34026	798	2321	9
		685	30139	788	1601	9
		688	37896	730	1577	10
buff.	3	40	361	172	446	1
		40	352	173	471	2
		41	337	169	494	3
		40	335	172	564	4
		41	333	235	480	5
		40	327	169	469	6
		40	341	175	567	7
		40	338	173	527	8
		40	343	174	482	9
		39	325	174	476	10

Tabela 11: Tempos em microsegundos para cada combinação nível/fator, replicada 10 vezes

Os dados observados para o estudo CLUSTER $\times$ SP2 encontram-se na Tabela 11. Assim como no estudo MPI NO CLUSTER, observa-se um grande espalhamento nos valores da tabela, a razão entre o maior e o menor valor y_{max}/y_{min} é 37896/39, aproximadamente 972, o que indica a necessidade de um modelo multiplicativo para este estudo. A Tabela 12 relaciona a transformação logarítmica dos dados da Tabela 11.

A análise de um projeto fatorial com k fatores é análoga àquela com dois fatores apresentada na Seção 7.1. A seguir as fórmulas para cálculo das médias das várias dimensões da tabela.

		Cluster $k=1$		SP2 $k=2$		
		0 bytes	20k	0 bytes	20k	
Com.	i	$j=1$	$j=2$	$j=1$	$j=2$	l
sínc.	1	2.8376	4.5079	2.8733	3.2033	1
		2.8351	4.5171	2.9212	3.2146	2
		2.8382	4.5749	2.8727	3.2475	3
		2.8325	4.4756	2.9106	3.2025	4
		2.8344	4.4752	2.8751	3.3740	5
		2.8344	4.5415	2.8762	3.2191	6
		2.8382	4.4853	2.8651	3.2060	7
		2.8376	4.5318	2.9020	3.3657	8
		2.8357	4.4791	2.8965	3.2044	9
		2.8376	4.5786	2.8633	3.1978	10
buff.	3	1.6021	2.5575	2.2355	2.6493	1
		1.6021	2.5465	2.2380	2.6730	2
		1.6128	2.5276	2.2279	2.6937	3
		1.6021	2.5250	2.2355	2.7513	4
		1.6128	2.5224	2.3711	2.6812	5
		1.6021	2.5145	2.2279	2.6712	6
		1.6021	2.5328	2.2430	2.7536	7
		1.6021	2.5289	2.2380	2.7218	8
		1.6021	2.5353	2.2405	2.6830	9
		1.5911	2.5119	2.2405	2.6776	10

Tabela 12: Transformação logarítmica dos resultados da Tabela 11

$$\begin{aligned}
\overline{y}_{i...} &= \mu + \alpha_i \\
\overline{y}_{.j..} &= \mu + \beta_j \\
\overline{y}_{..k.} &= \mu + \gamma_k \\
\overline{y}_{....} &= \mu
\end{aligned}$$

Os valores estimados dos parâmetros do modelo são

$$\begin{aligned}
\mu &= \overline{y}_{...} \\
\alpha_i &= \overline{y}_{i...} - \overline{y}_{....} \\
\beta_j &= \overline{y}_{.j..} - \overline{y}_{....} \\
\epsilon_k &= \overline{y}_{..k.} - \overline{y}_{....} \\
\gamma_{ABij} &= \overline{y}_{ij..} - \overline{y}_{i...} - \overline{y}_{.j..} + \overline{y}_{....} \\
\gamma_{ACik} &= \overline{y}_{i.k.} - \overline{y}_{i...} - \overline{y}_{..k.} + \overline{y}_{....} \\
\gamma_{BCjk} &= \overline{y}_{.jk.} - \overline{y}_{.j..} - \overline{y}_{..k.} + \overline{y}_{....} \\
\gamma_{ABCijk} &= \overline{y}_{ijk.} - \overline{y}_{ij..} - \overline{y}_{i.k.} - \overline{y}_{.jk.} + \overline{y}_{i...} + \overline{y}_{.j..} + \overline{y}_{..k.} - \overline{y}_{....}
\end{aligned}$$

A Tabela 14 apresenta os efeitos principais, calculados a partir das médias dos logarítimos de todas as observações de cada experimento (Tabela 13). Um envio em um *modo de comunicação* do tipo médio de uma mensagem de tamanho médio

			Cluster $k=1$		SP2 $k=2$	
			0 bytes	20k	0 bytes	20k
Com.	i	l	$j=1$	$j=2$	$j=1$	$j=2$
sínc.	1	.	$\overline{y}_{111.}=2.8361$	$\overline{y}_{121.}=4.5167$	$\overline{y}_{112.}=2.8856$	$\overline{y}_{122.}=3.2435$
buff.	2	.	$\overline{y}_{211.}=1.6031$	$\overline{y}_{221.}=2.5305$	$\overline{y}_{212.}=2.2498$	$\overline{y}_{222.}=2.6956$
			$\overline{y}_{11..}=2.8609$	$\overline{y}_{12..}=3.8801$	$\overline{y}_{1.1.}=3.6764$	$\overline{y}_{1.2.}=3.0645$
			$\overline{y}_{21..}=1.9265$	$\overline{y}_{22..}=2.6129$	$\overline{y}_{2.1.}=2.0667$	$\overline{y}_{2.2.}=2.4727$
			$\overline{y}_{.11.}=2.2196$	$\overline{y}_{.12.}=2.5677$	$\overline{y}_{.21.}=3.5235$	$\overline{y}_{.22.}=2.9695$
			$\overline{y}_{1...}=3.3705$		$\overline{y}_{2...}=2.2697$	
			$\overline{y}_{.1..}=2.3927$		$\overline{y}_{.2..}=3.2465$	
			$\overline{y}_{..1.}=2.8715$		$\overline{y}_{..2.}=2.7686$	
			$\overline{y}_{....}=2.8201$			

Tabela 13: Médias dos valores para as várias dimensões da Tabela 12

Fator	Nível 1		Nível 2		
A	$\alpha_1=\overline{y}_{1...}-\overline{y}_{...}=$	0.5504	$\alpha_2=\overline{y}_{2...}-\overline{y}_{...}=$	-0.5504	$\sum \alpha_i=0$
B	$\beta_1=\overline{y}_{.1..}-\overline{y}_{...}=$	-0.4264	$\beta_2=\overline{y}_{.2..}-\overline{y}_{...}=$	0.4264	$\sum \beta_j=0$
C	$\epsilon_1=\overline{y}_{..1.}-\overline{y}_{...}=$	0.0515	$\epsilon_2=\overline{y}_{..2.}-\overline{y}_{...}=$	-0.0515	$\sum \epsilon_k=0$

Tabela 14: Efeitos dos três fatores principais para o estudo CLUSTER × SP2

em um *ambiente paralelo* do tipo médio é dado por $\overline{y}_{...}$ e necessita de um tempo aproximado de log 2.8201 (634 μs).

- Fator Modo de Comunicação: Um envio no *modo de comunicação* síncrono necessita aproximadamente de um tempo log 0.5504 μs a mais do que um envio médio (isto é um fator de 3.55 vezes), como um envio médio necessita de aproximadamente 634 μs, um envio síncrono necesssita de, aproximadamente, 2252 μs, 1618 μs a mais do que um envio médio. Um envio no *modo de comunicação* bufferizado necessita de um tempo log 0.5504 μs a menos do que um envio médio (fator 0.2816) que corresponde a um tempo aproximado de 179 μs (455 μs a menos do que o envio médio). Comparando-se os dois níveis entre si, o tempo de envio no *modo de comunicação* síncrono é aproximadamente 12.6 vezes ($10^{\alpha_1-\alpha_2}$) maior do o tempo de um envio no *modo de comunicação* bufferizado.

- Fator Tamanho da Mensagem: Um envio de uma mensagem de 0 bytes necessita de um tempo log 0.4264 μs a menos do que um envio médio (isto é, um fator de 0.37 vezes), que corresponde a um tempo de 238 μs (397 μs a menos do que o envio médio). Um envio de uma mensagem de 20k necessita de um tempo log 0.4264 μs a mais do que um envio médio (fator 2.67) que corresponde a um envio de 1692 μs aproximadamente (1058 μs a mais do que o envio médio). O tempo de envio de uma mensagem de tamanho 20k é aproximadamente 7 vezes ($10^{\beta_2-\beta_1}$) maior do que o tempo de envio de uma mensagem de 0 bytes.

- Fator Ambiente Paralelo: Um envio no cluster necessita aproximadamente de um tempo log 0.0515 μs a mais do que um envio médio (isto é um fator de 1.12 vezes), como um envio médio necessita de aproximadamente 634 μs, um envio síncrono necesssita de, aproximadamente, 714 μs, 80 μs a mais do que um envio médio. Um envio no SP2 necessita de um tempo log 0.0515 μs a menos do que um envio médio (fator 0.8883) que corresponde a um tempo aproximado de 563 μs (71 μs a menos do que o envio médio). O tempo de envio no Clsuter é aproximadamente 1.2 vezes ($10^{\epsilon_1-\epsilon_2}$) maior do que o tempo de envio no SP2.

Os efeitos multiplicativos das interações dois a dois são apresentados nas Tabelas 15 a 17 e calculados conforme a seguir:

$$\gamma_{AB11} = \overline{y}_{11..} - \overline{y}_{1...} - \overline{y}_{.1..} + \overline{y}_{....} = 2.8609 - 3.3750 - 2.3927 + 2.8201 = 0.0832$$

$$\vdots$$

$$\gamma_{BC22} = \overline{y}_{.22.} - \overline{y}_{.2..} - \overline{y}_{..2.} + \overline{y}_{....} = 2.9695 - 3.2465 - 2.7686 + 2.8201 = -0.2255$$

$$\gamma_{ABC111} = \overline{y}_{111.} - \overline{y}_{11..} - \overline{y}_{1.1.} - \overline{y}_{.11.} + \overline{y}_{1...} + \overline{y}_{.1..} + \overline{y}_{..1.} - \overline{y}_{...} =$$
$$= 2.8361 - 2.8609 - 3.6764 - 2.2196 + 3.3705 + 2.3927 + 2.8715 + 2.8201 = -0.1052$$

$$\vdots$$

E assim por diante [§].

Modo/Tamanho	0 bytes	20 k
síncrono	$\gamma_{AB11} =$ -0.0832	$\gamma_{AB12} =$ 0.0832
bufferizado	$\gamma_{AB21} =$ 0.0832	$\gamma_{AB22} =$ -0.0832

Tabela 15: Efeitos da interação entre os fatores Modo de Comunicação e Tamanho da Mensagem

Por exemplo, o tempo de um envio síncrono no Cluster é 3.23 vezes ($10^{\gamma_{AC11}-\gamma_{AC12}}$) um envio síncrono no SP2 e o tempo de um envio bufferizado no SP2 é 3.23 vezes ($10^{\gamma_{AC22}-\gamma_{AC21}}$) um envio bufferizado no Cluster.

Por último, os efeitos das oito interações de fatores três a três são relatados na Tabela 18.

[§] Algumas diferenças de valores calculados pode ser devida a arredondamentos.

Modo/Ambiente	Cluster	SP2
síncrono	$\gamma_{AC11} =$ 0.2545	$\gamma_{AC12} =$ -0.2545
bufferizado	$\gamma_{AC21} =$ -0.2545	$\gamma_{AC22} =$ 0.2545

Tabela 16: Efeitos da interação entre os fatores Modo de Comunicação e Ambiente Paralelo

Tamanho/Ambiente	Cluster	SP2
0 bytes	$\gamma_{BC11} =$ -0.2255	$\gamma_{BC12} =$ 0.2255
20 k	$\gamma_{BC21} =$ 0.2255	$\gamma_{BC22} =$ -0.2255

Tabela 17: Efeitos da interação entre os fatores Tamanho da Mensagem e Ambiente Paralelo

7.2.2 Alocação de Variação

A variação total de y é calculada de forma análoga à apresentada na Seção 7.1.3.

$$SST = SSY - SS0 = SSA + SSB + SSC + SSAB + SSAC + SSBC + SSABC + SSE$$

$$SSY = \sum_{ijkl} y_{ijkl}^2 = (2.8376)^2 + (2.8351)^2 + \ldots + (2.6776)^2 = 686.002$$

$$SS0 = abcr\mu^2 = 2 * 2 * 2 * 10 * (2.8201)^2 = 636.2303$$

$$SSA = bcr \sum_i \alpha_i^2 = 2 * 2 * 10 * [(0.5504)^2 + (-0.5504)^2] = 24.2348$$

$$SSB = acr \sum_j \beta_j^2 = 2 * 2 * 10 * [(-0.4264)^2 + (0.4264)^2] = 14.5468$$

$$SSC = abr \sum_k \epsilon_k^2 = 2 * 2 * 10 * [(0.0515)^2 + (-0.0515)^2] = 0.2119$$

$$SSAB = cr \sum_{ij} \gamma_{ij}^2 = 2 * 10 * [(-0.0832)^2 + (0.0832)^2 + (-0.0832)^2 + (0.0832)^2] = 0.5537$$

$$SSAC = br \sum_j \gamma_{ik}^2 = 2 * 10 * [(0.2545)^2 + (-0.2545)^2 + (0.2545)^2 + (-0.2545)^2] = 5.1805$$

	Cluster		SP2	
Modo Com.	0 bytes	20k	0 bytes	20k
sínc.	$\gamma_{ABC111} =$ -0.1052	$\gamma_{ABC121} =$ 0.1052	$\gamma_{ABC112} =$ 0.1052	$\gamma_{ABC122} =$ -0.1052
buff.	$\gamma_{ABC211} =$ 0.1052	$\gamma_{ABC221} =$ -0.1052	$\gamma_{ABC212} =$ -0.1052	$\gamma_{ABC222} =$ 0.1052

Tabela 18: Efeitos da interação entre os fatores Modo de Comunicação, Tamanho da Mensagem e Ambiente Paralelo

$$SSBC = ar\sum_{j}\gamma_{jk}^{2} = 2*10*[(-0.2255)^2+(0.2255)^2+(-0.2255)^2+(0.2255)^2] = 4.0684$$

$$SSABC = r\sum_{j}\gamma_{ijk}^{2} = 10*[(-0.15052)^2+\ldots+(0.15052)^2] = 0.8848$$

$$SST = SSY - SS0 = 49.7699$$

$$SSE = SSY-SS0-SSA-SSB-SSC-SSAB-SSAC-SSBC-SSABC = 0.0891$$

A percentagem de variação explicada pelos vários fatores é

- explicada pelo *modo de comunicação*

$$100\frac{SSA}{SST} = 100\frac{24.2348}{49.7699} = 48.69\%$$

- explicada pelo *tamanho da mensagem*

$$100\frac{SSB}{SST} = 100\frac{14.546}{49.7699} = 29.23\%$$

- explicada pelo *ambiente paralelo*

$$100\frac{SSC}{SST} = 100\frac{0.2119}{49.7699} = 0.43\%$$

- explicada pela interação *modo de comunicação/tamanho da mensagem*

$$100\frac{SSAB}{SST} = 100\frac{0.5537}{49.7699} = 1.11\%$$

- explicada pela interação *modo de comunicação/ambiente paralelo*

$$100\frac{SSAC}{SST} = 100\frac{5.1805}{49.7699} = 10.41\%$$

- explicada pela interação *tamanho da mensagem/ambiente paralelo*

$$100\frac{SSBC}{SST} = 100\frac{4.0684}{49.7699} = 8.17\%$$

- explicada pela interação *modo de comunicação/tamanho da mensagem/ambiente paralelo*

$$100\frac{SSABC}{SST} = 100\frac{0.8848}{49.7699} = 1.78\%$$

- inexplicada

$$100\frac{SSE}{SST} = 100\frac{0.0891}{49.7699} = 0.18\%$$

Percebe-se que, diferentemente do estudo MPI NO CLUSTER, nesse estudo o fator *modo de comunicação* teve uma maior contribuição para a variação do que o fator *tamanho da mensagem*. Quer dizer, considerando-se apenas os níveis de *tamanho da mensagem* 0 bytes e 20k foi possível verificar a importância do fator *modo de comunicação* no desempenho do envio. Curiosamente, outro fator principal, o *ambiente paralelo*, teve uma importância desprezível de menos de 1%, inferior, inclusive, à contribuição de fatores secundários. Do ponto de vista das interações dois a dois percebe-se claramente a dependência dos fatores *ambiente* versus *modo de comunicação* e *ambiente* versus *tamanho da mensagem*. Para finalizar, o modelo consegue explicar 99.82% da variação. A primeira análise indica que este o modelo multiplicativo é um bom modelo para o estudo CLUSTER × SP2.

7.2.3 Análise de Variância

Os testes para verificar estatisticamente a significância dos fatores são realizados de forma similar àquela utilizada na Seção 7.1.4.

Os graus de liberdade para as somas de quadrados são os seguintes:

$$SSY = SS0 + SSA + SSB + SSC + SSAB + SSAC + SSBC + SSABC + SSE$$

$$\begin{aligned} abcr = 1 + (a-1) + (b-1) + (c-1) + (a-1)(b-1) + (a-1)(c-1) + \\ +(b-1)(c-1) + (a-1)(b-1)(c-1) + abc(r-1) \end{aligned}$$

$$80 = 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 72$$

A seguir, o cálculo das médias dos quadrados:

$$MSA = \frac{SSA}{a-1} = \frac{24.2348}{2-1} = 24.23483$$

$$MSB = \frac{SSB}{b-1} = \frac{14.546}{2-1} = 14.546$$

$$MSC = \frac{SSC}{c-1} = \frac{0.2119}{2-1} = 0.2119$$

$$MSAB = \frac{SSAB}{(a-1)(b-1)} = \frac{0.5537}{(2-1)(2-1)} = 0.5537$$

$$MSAC = \frac{SSAC}{(a-1)(c-1)} = \frac{5.1805}{(2-1)(2-1)} = 5.1805$$

$$MSBC = \frac{SSBC}{(b-1)(c-1)} = \frac{4.0684}{(2-1)(2-1)} = 4.0684$$

$$MSABC = \frac{SSABC}{(a-1)(b-1)(c-1)} = \frac{0.8848}{(2-1)(2-1)(2-1)} = 0.8848$$

$$MSE = \frac{SSE}{abc(r-1)} = \frac{0.0891}{2*2*2*(10-1)} = 0.0012$$

A seguir o cálculo das razões com distribuição $F[0.90, n, m]$ onde 90% é o nível de confiança desejado, n é o grau de liberdade do fator/interação correspondente e m é o grau de liberdade do erro. Assim como na Seção 7.1.4 será utilizado o valor de m mais alto disponível na tabela.

$$\frac{MSA}{MSE} = \frac{24.23483}{0.0012} = 19576.4949$$

$$\frac{MSB}{MSE} = \frac{14.546}{0.0012} = 11750.6428$$

$$\frac{MSC}{MSE} = \frac{0.2119}{0.0012} = 171.1493$$

$$\frac{MSAB}{MSE} = \frac{0.5537}{0.0012} = 447.2327$$

$$\frac{MSAC}{MSE} = \frac{5.1805}{0.0012} = 4184.6823$$

$$\frac{MSBC}{MSE} = \frac{4.0684}{0.0012} = 3286.3541$$

$$\frac{MSABC}{MSE} = \frac{0.8848}{0.0012} = 714.6931$$

A Tabela 19 sumariza os valores calculados assumindo-se um nível de significância de 0.10. Como as razões (F-Calculado) são maiores do que os valores da tabela (F-tabela), assume-se que todos os efeitos são estatísticamente significantes no nível de significância de 0.10.

Componente	Soma dos Quadrados	Perc. da Variação	Graus de Lib.	Quadrado Médio	F-Calculado	F-Tabela $F[0.90, n, m]$ $m = 40$
y	686.002					
$y_{...}$	636.2303					
$y - y_{...}$	49.7699	100				
Modo Comunicação	24.2348	48.69	1	24.23483	19576.4949	2.84
Tamanho Mensagem	14.5468	29.23	1	14.546	11750.6428	2.84
Ambiente Paralelo	0.2119	0.43	1	0.2119	171.1493	2.84
Com./Tamanho Msg.	0.5537	1.11	1	0.5537	447.2327	2.84
Com./Ambiente Par.	5.1805	10.41	1	5.1805	4184.6823	2.84
Tam./Ambiente Par	4.0684	8.17	1	4.0684	3286.3541	2.84
Com./Tam./Par.	0.8848	1.78	1	0.8848	714.6931	2.84
Erros	0.0891	0.18	72	0.0012		

Tabela 19: Tabela ANOVA para o estudo CLUSTER × SP2

7.2.4 Intervalos de Confiança para os Efeitos

As Tabelas 20 a 24 mostram os intervalos de confiança para os diversos fatores e suas interações. Observa-se, por exemplo, que o intervalo do fator *modo de comunicação* não inclui zero e portanto o efito deste fator é considerado significativo. O mesmo pode ser dito para os demais fatores isolados e para as interações entre eles. Assim como no estudo anterior, a replicação dos experimentos permitiu isolar as interações dos erros experimentais que, por sua vez, são muito pequenos. Pode-se confiar nos limites dos efeitos previstos pelo modelo.

Parâmetro	Efeito Médio	Desvio Padrão	Intervalo de Confiança
μ	2.8201	$s_\mu = s_e\sqrt{\frac{1}{abcr}} = 0.0352\sqrt{\frac{1}{80}} = 0.0039$	$(2.2724, 2.2878)$
Modo Comunicação			
sinc.	0.5504	$s_{\alpha_i} = s_e\sqrt{\frac{(a-1)}{abcr}} = 0.0352\sqrt{\frac{1}{80}} = 0.0039$	$(0.5427, 0.5581)$
buff.	-0.5504		$(-0.5581, -0.5427)$
Tamanho da Mensagem			
0 bytes	-0.4264	$s_{\beta_j} = s_e\sqrt{\frac{(a-1)}{abcr}} = 0.0352\sqrt{\frac{1}{80}} = 0.0039$	$(-0.4341, -0.4187)$
20k	0.4264		$(0.4187, 0.4341)$
Ambiente Paralelo			
Cluster	0.0515	$s_{\epsilon_k} = s_e\sqrt{\frac{(a-1)}{abcr}} = 0.0352\sqrt{\frac{1}{80}} = 0.0039$	$(-0.0438, 0.0592)$
SP2	-0.0515		$(-0.0592, 0.0438)$
$s_e = \sqrt{MSE} = \sqrt{0.0012} = 0.0352$			
$z_{0.975} = 1.960$			

Tabela 20: Intervalo de confiança para os efeitos principais no estudo CLUSTER × SP2

Modo/Tamanho	0 bytes	20 k
síncrono bufferizado	(-0.0909,-0.0755) (0.0755,0.0909)	(0.0755,0.0909) (-0.0909,-0.0755)
$s_{\gamma_{AB}} = s_e\sqrt{\frac{(a-1)(b-1)}{abcr}} = 0.0039$		
$z_{0.975} = 1.960$		

Tabela 21: Intervalos de confiança para os efeitos das interações entre os fatores Modo de Comunicação e Tamanho da Mensagem

Modo/Ambiente	Cluster	SP2
síncrono bufferizado	(0.2468,0.2622) (-0.2622,-0.2468)	(-0.2622,-0.2468) (0.2468,0.2622)
$s_{\gamma_{AC}} = s_e\sqrt{\frac{(a-1)(c-1)}{abcr}} = 0.0039$		
$z_{0.975} = 1.960$		

Tabela 22: Intervalos de confiança para os efeitos das interações entre os fatores Modo de Comunicação e Ambiente Paralelo

7.2.5 Testes Visuais

O resultado do teste visual da independência dos erros é mostrado na Figura 4. A maioria dos valores dos eixos vertical e horizontal diferem de duas ordens de magnitude. Apenas poucos pontos diferem de uma ordem de magnitude, que já é considerado uma boa diferença. Não parece existir nenhuma tendência definida nos pontos do gráfico, e, portanto, pode-se dizer que os erros são independentes.

A Figura da Tabela 5 permite visualizar que não há indicação da existência de fatores não considerados pelo sistema ou efeitos colaterais (inicialização incorreta). Para os testes realizados a ordem disponível é a ordem para cada tamanho de mensagem de cada ambiente paralelo, por isso o eixo horizontal do gráfico varia de 1 a 20.

Tamanho/Ambiente	Cluster	SP2
0 bytes 20 k	(-0.2332,-0.2178) (0.2178,0.2332)	(0.2178,0.2332) (-0.2332,-0.2178)
$s_{\gamma_{BC}} = s_e\sqrt{\frac{(b-1)(c-1)}{abcr}} = 0.0039$		
$z_{0.975} = 1.960$		

Tabela 23: Intervalos de confiança para os efeitos das interações entre os fatores Tamanho da Mensagem e Ambiente Paralelo

Modo Com.	Cluster 0 bytes	20k
sínc.	(-0.1129,-0.0975)	(0.0975,0.1129)
buff.	(0.0975,0.1129)	(-0.1129,-0.0975)
	SP2	
sínc.	(-0.1129,-0.0975)	(0.0975,0.1129)
buff.	(0.0975,0.1129)	(-0.1129,-0.0975)
$s_{\gamma_{ABC}} = s_e\sqrt{\frac{(a-1)(b-1)(c-1)}{abcr}} = 0.0039$		
$z_{0.975} = 1.960$		

Tabela 24: Intervalo de confiança para as interações entre os fatores Modo de Comunicação, Tamanho da Mensagem e Ambiente Paralelo

7.2.6 Conclusão

Sobre o estudo CLUSTER × SP2 a análise preliminar dos dados (vide Seção 7.1.7) mostrava uma tendência do SP2 a ser mais rápido para mensagens iguais ou maiores que 20k e para os modos síncrono e bufferizado. Quer dizer, o efeito do fator *ambiente paralelo* depende muito da seleção dos níveis para os fatores *modo de comunicação* e *tamanho da mensagem*. Neste estudo, houve uma preocupação em se escolher dois *tamanhos de mensagens* e dois *modos de comunicação* considerados mais usuais. Levando-se em conta esta escolha, concluí-se que o efeito do fator *ambiente paralelo*, muito pequeno, não é relevante. Diferentemente do estudo MPI NO CLUSTER o fator *modo de comunicação* teve uma maior contribuição para a variação do que o fator *tamanho da mensagem*. Mais uma vez, sendo os níveis das mensagens escolhidos para este estudo mais realistas, tem-se aqui uma melhor avaliação da importância deste fator.

8 Considerações Finais

Para sintetizar, em relação ao desempenho de um envio com bloqueio em MPI, supondo que exista um recebimento combinado colocado e máquinas exclusivas

- Faz diferença significativa escolher o modo de comunicação. Dê preferência ao bufferizado. É importante lembrar que esta preferência está considerando somente o ponto de vista da tarefa de envio.

- Deve-se considerar a interação *modo de comunicação/tamanho da mensagem*, o *modo de comunicação* padrão utiliza uma boa política: mensagens "pequenas" bufferizadas e mensagens "grandes" sincronizadas.

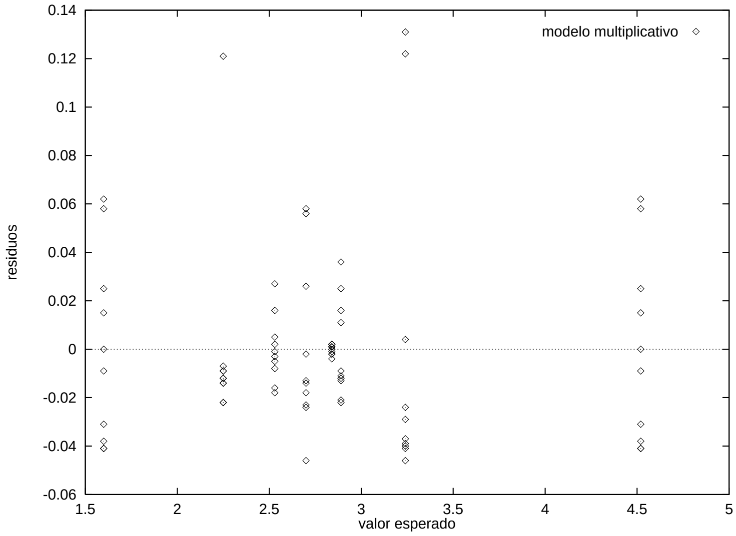

Figura 4: Gráfico dos resíduos e_{ijk} versus o valor esperado $\hat{y}_{ij}$

- Para os dois modos de comunicação considerados de maior interesse (bufferizado e sincronizado) e dois tamanhos de mensagens considerados "pequeno" e "médio para grande" (0 bytes e 20k) não faz diferença significativa executar no Cluster ou SP2.

Referências

[1] Cornell Theory Center, *MPI: Message Passing Interface Library*, disponível em ˘http://www.tc.cornell.edu/Edu/Talks/topic.html#mess.

[2] R. Jain, *The Art of Computer Systems Performance Analysis Algorithms*, John Wiley & Sons, 1991, New York.

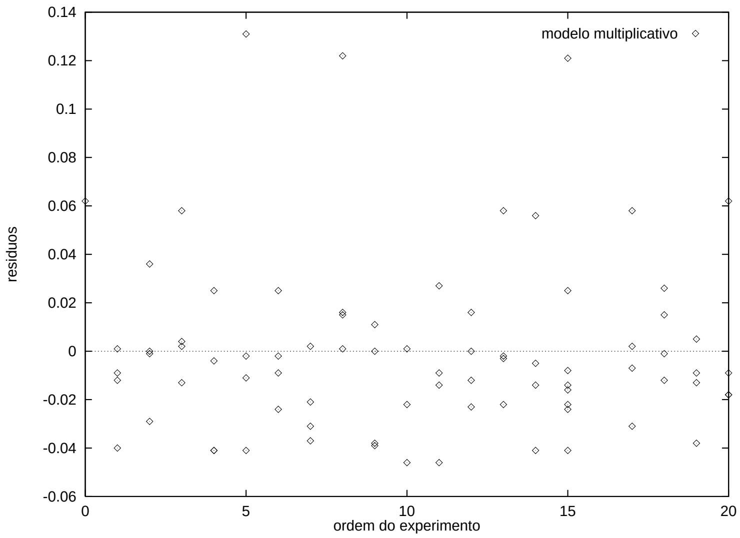

Figura 5: Gráfico dos resíduos e_{ijk} versus a ordem de execução do experimento