

CDO-2

A PROPOSAL TO DETERMINE AND ACHIEVE THE AMOUNT OF RESOURCES NECESSARY AND SUFFICIENT FOR DATA PROCESSING IN MAINFRAMES

Antonio Cesar Sartoratto Dias (State University of Campinas, São Paulo, Brazil) -

acsdias@uol.com.br

Marcus Fabius H. de Carvalho (Renato Archer Research Center, São Paulo, Brazil) -

marcius.carvalho@cenpra.gov.br

This thesis has two objectives: (1) determine the elapsed time and amount of resources necessary and sufficient for data processing in mainframes, and (2) identify the constraints that impair their achievement. We believe that mainframes can be compared to machines which transform data into information and we applied concepts of industrial productivity and efficiency to obtain the best mainframe use level. Although we use concepts, rules of thumbs and classical formulas of IT, it was important to use methods of the industrial processes to structure this methodology. Data were obtained in 4 companies in the financial and government sectors with similar situations of processor consumption, efficiency and response time. Only one of them authorized the completion of the project. The following results were obtained: the average transaction time was reduced from 29 s. to 0.3 s., the use of processor was reduced from 91% to 43%.

Keywords: IT Management, Efficiency, Mainframes.

UMA PROPOSTA PARA DETERMINAR E ATINGIR A QUANTIDADE DE RECURSOS NECESSÁRIA E SUFICIENTE PARA O PROCESSAMENTO DE DADOS EM COMPUTADORES DE GRANDE PORTE

Antonio Cesar Sartoratto Dias
e-mail: acsdias@uol.com.br
telefones: 11-5584-8255 e 11-9972-3077
Universidade Estadual de Campinas
Faculdade de Engenharia Mecânica
Departamento de Engenharia de Fabricação
Comissão de Pós-Graduação em Engenharia Mecânica
Curso: Engenharia Mecânica
Área de Concentração: Engenharia de Fabricação

Orientador: Prof. Dr. Marcius Fabius Henriques de Carvalho
área temática de submissão desta proposta:
Métricas Aplicadas à Tecnologia de Informação

1 Breve Introdução

O consumo de MIPS tem aumentado em níveis superiores ao crescimento da economia, segundo pesquisas da Computerworld (2007) e Arcati (2007). Na Figura 1.1 é possível observar que, entre 2003 e 2006, o setor bancário brasileiro cresceu 39% enquanto que o consumo de MIPS cresceu 112%, segundo dados da Febraban (2007) que declara: a duplicação da capacidade de processamento dos bancos em 3 anos reflete a necessidade de tratamento de maiores volumes de transações, de novas aplicações desenvolvidas, da instalação de sofisticados controles e de instrumentos de gestão dos bancos.

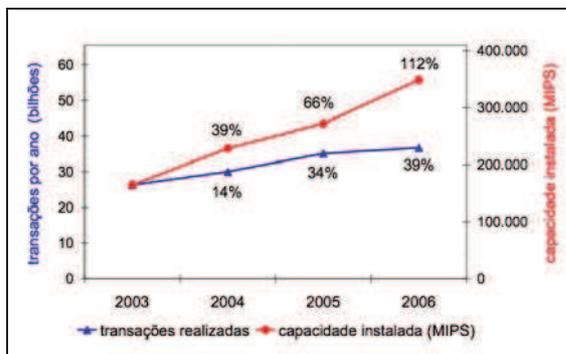


Figura 1.1 – Crescimento do mercado bancário X crescimento de MIPS do setor

Apesar da alta de consumo de MIPS ser atribuída ao aumento da complexidade do negócio, foi observado, em instituições governamentais e financeiras, que os computadores de grande porte estão atuando com baixo nível de eficiência. As Figuras 1.2 a 1.5 registram a diferença entre o nível observado de utilização de processador e o nível de consumo equilibrado proposto por Amdahl (1970). A distância entre as linhas superiores e inferiores representa a perda de eficiência nos processos realizados pelos equipamentos analisados.

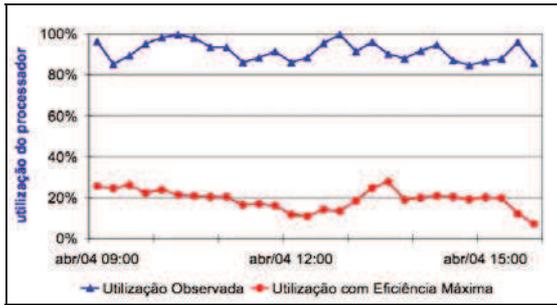


Figura 1.2 – nível observado X nível ideal do ambiente 1

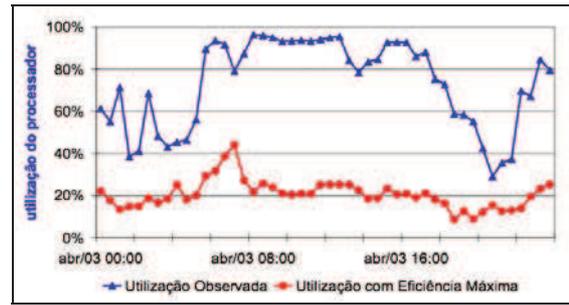


Figura 1.3 – nível observado X nível ideal do ambiente 2

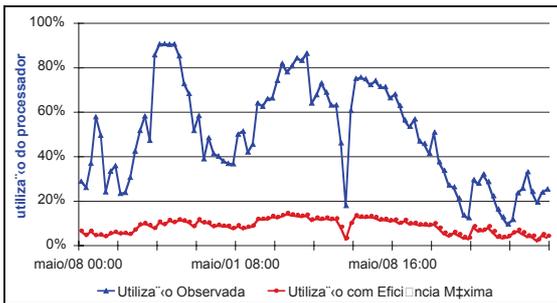


Figura 1.4 – nível observado X nível ideal do ambiente 3

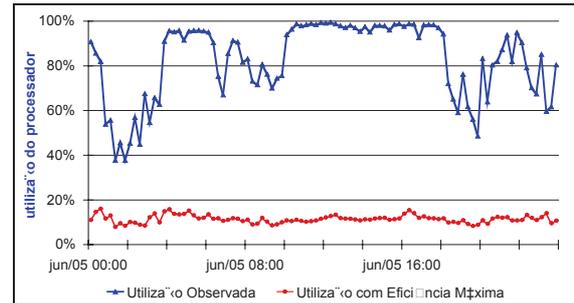


Figura 1.5 – nível observado X nível ideal do ambiente 4

1.1 Problema de Pesquisa

Este trabalho tem o objetivo de determinar e atingir o tempo de duração e a quantidade de recursos necessários e suficientes para o processamento de dados realizado por computadores de grande porte. O problema de pesquisa é formalizado pela questão “qual a duração e a quantidade de recursos ideais para processar uma determinada quantidade de dados em uma determinada configuração de equipamento”?

O ponto de partida para as pesquisas são os conceitos de eficiência utilizados pela indústria, cuja representação gráfica de um processo é semelhante ao utilizado pelo processamento de dados, conforme pode ser observado na Figura 1.6.

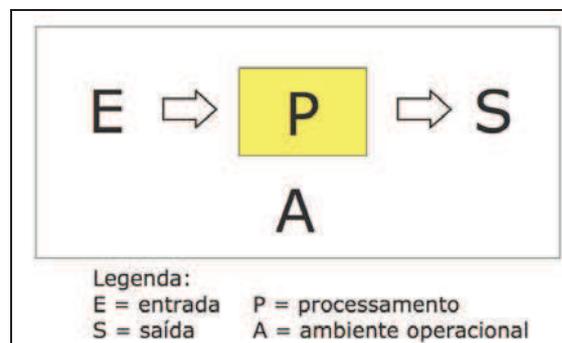


Figura 1.6 – Representação de um processo

1.2 Objetivo

O objetivo é encontrar o ponto de eficiência para o processamento de dados realizado por computadores de grande porte, assim como estabelecer o tempo ideal de processo.

1.3 Justificativa

Os parâmetros propostos por este trabalho estabelecem uma meta para projetos de melhoria de tempo de resposta e aprimoram os resultados de planejamento de capacidade.

A Figura 1.7 registra a variação do tempo de processamento de 1 milhão de registros em dois ambientes operacionais. Foi observado que existem parametrizações que permitem que o tempo de duração dos processos sejam menores, mais estáveis e menos dependentes da carga de trabalho realizada pelo equipamento.

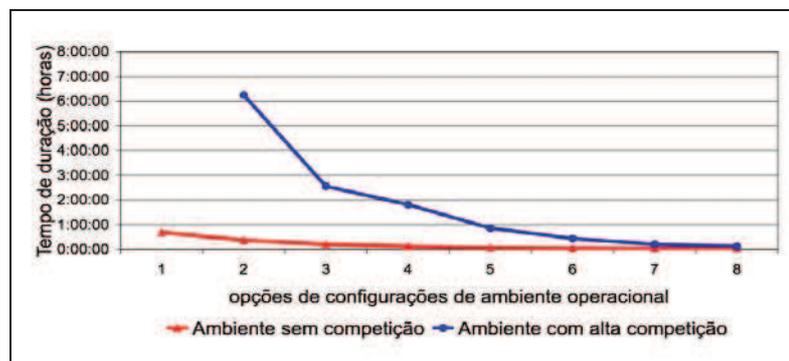


Figura 1.7 – Tempo de processamento de 1 milhão de registros em função das configurações utilizadas e desempenho do sistema de armazenamento

A Figura 1.8 registra variação nos tempos das transações de um sistema crítico em um ambiente operacional de baixa eficiência. Serviços similares tiveram tempo de processamento variando entre 0,1 seg. e 20 seg. Existem conhecimentos que agrupados auxiliarão a responder qual o tempo ideal para uma transação.

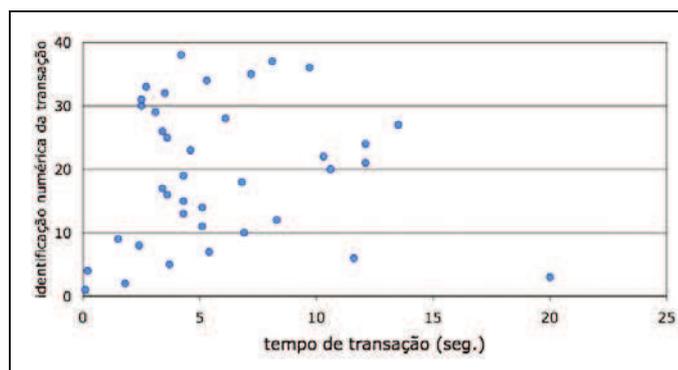


Figura 1.8 – Variação final do tempo de transação em um sistema crítico

2 Breve Quadro Teórico de Referência

A seguir serão discutidos alguns pontos que sustentam a proposta deste trabalho.

A Figura 2.1 resume os resultados das pesquisas de Spruth (2007) que permitem observar: (1) os *mainframes* atuais são capazes de processar 57 milhões de transações por hora, volume que grandes bancos brasileiros processam por dia, (2) oferecem potencial para tempos de

respostas de transações na ordem de 0,12 segundos, valor inferior, por exemplo, aos resultados de *benchmarks* publicados pela SAP, em torno de 1,76 seg. (SAP, 2007).

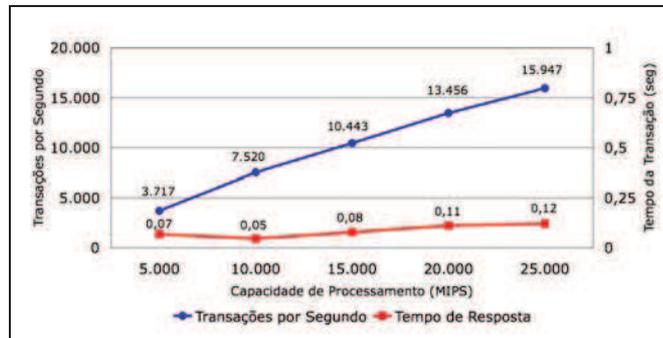


Figura 2.1 – Quantidade de transações e tempo de resposta em função da capacidade de MIPS do *mainframe*, adaptada de Spruth (2007)

O trabalho original de Spruth (2007) mostra, através da Figura 2.1, que o tempo de resposta aumenta à medida que aumenta a quantidade de MIPS, quando é esperado o oposto. Ou seja, que maior capacidade de MIPS reduza o tempo do processamento de dados. A divergência apontada por Spruth (2007), juntamente com aumento do consumo de MIPS demonstrado pelas pesquisas da Computerworld (2007), Arcati (2007) e Febraban (2007), são temas de estudo deste trabalho, pois os resultados na prática podem ser melhores que os observados.

Os fatores que influenciam no desempenho do processamento de dados são:

- Aplicativos

O papel dos desenvolvedores de TI é essencialmente centrar-se na eficiência da entrega e manutenção de aplicações em larga escala (van der Zee, 2003).

O foco na produtividade e eficiência durante a fase de desenvolvimento do aplicativo tem transferido custos para o período de sua vida útil. Ainda que a eficiência de execução fosse a principal preocupação no projeto das primeiras linguagens, considera-se isso menos importante agora. Os esforços para projetar e implementar linguagens de alto nível são motivados pelo compromisso de diminuir os custos da criação de aplicativos. (Sebesta, 2007)

A complexidade dos aplicativos serão objeto de estudo. Este trabalho fará a revisão bibliográfica das métricas de complexidade para em seguida realizar a análise da relação entre complexidade de aplicativos e consumo de MIPS. Pandian (2003) concluiu que o desenvolvimento de uma métrica de complexidade tem sido um exercício científico há décadas. Soluções foram encontradas e publicadas por pesquisadores, mas não encantaram todos os profissionais, algumas delas são apoiadas por ferramentas, a fim de facilitar a aplicação. As métricas de complexidade estão relacionadas na Figura 2.2.

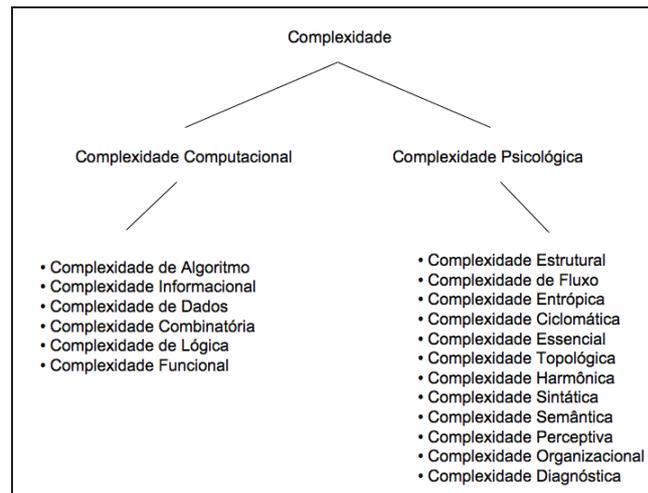


Figura 2.2 – Métricas de complexidade, adaptada de Pandian (2003)

Outra fonte proposta como referência para o estudo da complexidade de aplicativos é Laird e Brennan (2006). Durante o desenvolvimento deste trabalho haverá pesquisa sobre McCabe (1976) e trabalhos mais recentes. Ele desenvolveu trabalhos pioneiros nessa área e atualmente produz softwares para medir complexidade em aplicativos escritos em algumas linguagens, como o Cobol, presente nos setores bancário e governamental.

- Armazenamento

A configuração do sistema de armazenamento de dados tem um grande impacto sobre o desempenho. (Hennessy e Patterson, 2007)

Os sistemas de armazenamento tem na teoria das filas a base para o bom desempenho, equilibrando o *throughput* em relação ao tempo de resposta. A espera pela transferência de dados pode interferir com a execução de outros processos. Também existe uma sobrecarga devido ao tratamento de interrupções de leituras de gravações. (Hennessy e Patterson, 2003)

- Balanceamento

Um computador é considerado balanceado quando existe uma distribuição equilibrada entre as diversas cargas de trabalho de seus componentes. Isto quer dizer que todos os dispositivos devem ser utilizados igualmente. Um sistema balanceado não é aquele formado por dispositivos de mesmas velocidades e capacidades. Também não é aquele onde os componentes mais rápidos são visitados mais vezes ou que cada visita seja mais demorada. Um sistema equilibrado implica na eliminação de gargalos. São considerados dispositivos gargalos aqueles onde os impactos das melhoras são percebidas com maior intensidade que o restantes dos demais dispositivos. De qualquer forma, os sistemas balanceados são os ideais, sem que nenhum dispositivo ou recurso seja um fator limitador do bom desempenho. É importante considerar o equilíbrio dos computadores já que eles são os responsáveis pelos padrões de desempenhos elevados. (Menascé e Almeida, 2004)

Segundo Fortier e Michel (2003), otimizar é utilizar 100% do recurso. Entretanto, recursos com excesso de uso geram longas filas de espera, aumentando o tempo de processo e diminuindo o nível de serviço oferecido. Para Fortier e Michel (2003), a eficiência está relacionada com a máxima utilização e a máxima vazão. Por exemplo, se existe um processador avaliado em 100 *megaflops* e durante um teste é medido 90 *megaflops*, a eficiência desse processador é de 90%. Sob essa interpretação, um

aplicativo que ocupe 100% do processador, sem produzir resultados, pode ser considerado eficiente.

A Figura 2.3 ilustra a relação da frente de atendimento com os serviços de retaguarda, onde estão incluídos os serviços de processamento de dados. Durante o desenvolvimento deste trabalho será avaliada a contribuição que a melhoria da eficiência do processamento de dados representa na melhoria do desempenho da organização. Para conhecer essa relação, haverá consulta à uma determinada instituição bancária que iniciou um projeto de BPM (*Business Performance Measurement*) e a uma entidade que agrupa profissionais de planejamento de capacidade.

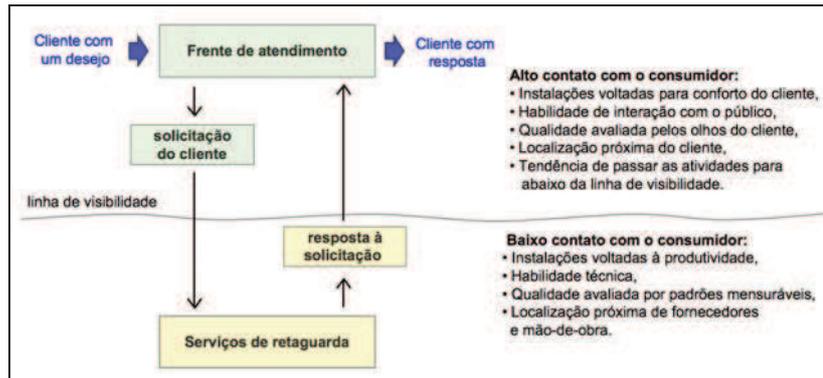


Figura 2.3 – Relação dos serviços de retaguarda com a frente de atendimento, adaptada de Fitzsimmons e Fitzsimmons (2005)

Ainda que os dados nacionais não sejam disponibilizados pelas entidades citadas, ou que a busca desses dados não seja aconselhada pela banca deste Consórcio, é sabido que à medida que a organização direciona-se para a ênfase em resultados de negócios, distancia-se da ênfase na eficiência das tarefas de T.I. (van der Zee 2003)

Segundo Murphy (2002), existem 4 níveis de impacto da T.I. na estrutura de negócios. No nível 1, onde predominam os mainframes, os controles de custos de desenvolvimento e manutenção de sistemas são pouco rigorosos e nas empresas que utilizam controles financeiros rigorosos, esses custos são atribuídos aos prováveis benefícios diretos e indiretos.

Como este trabalho propõe algumas formas de mensuração, é importante mencionar o estudo de Albertin e Albertin (2005): Os investimentos e operações de TI não devem ser medidos caso o benefício esperado exceda os custos de medir. O valor da mensuração, percebido pela empresa, é diferença entre o custo da mensuração e valor dos benefícios recebidos, diretos e indiretos. Este cenário está representado na Figura 2.4.

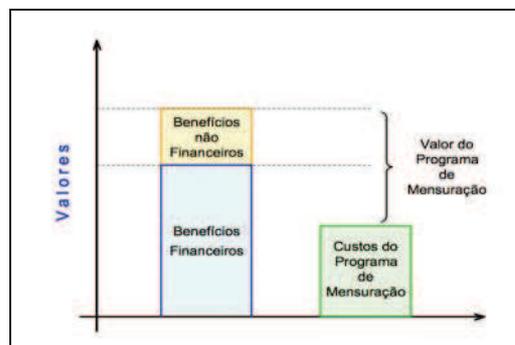


Figura 2.4 – Valor do investimento em mensuração, adaptada de (Albertin e Albertin, 2005)

As pesquisas utilizadas para comparar o crescimento do consumo de MIPS com o crescimento da economia são:

- Febraban: publica anualmente os resultados de pesquisas realizadas junto a CIOs durante o CIAB (Febraban, 2007)
- FGV: pesquisa Anual de Administração de Recursos de Informática (Meirelles, 2007)
- Computerworld: pesquisa realizada durante o ano de 2007 junto a CIOs do mundo sobre tendências, práticas e uso de recursos.
- Arcati Mainframe Yearbook 2007: pesquisa anual realizada junto a CIOs do mundo, mostra as tendências, práticas e uso de recursos.

3 Procedimentos Metodológicos

Para que este trabalho cumpra seu objetivo de “determinar e atingir o tempo de duração e a quantidade de recursos necessários e suficientes para o processamento de dados realizado por computadores de grande porte”, este capítulo é composto de 3 itens principais:

- Determinar o menor tempo de processo para uma configuração de equipamento,
- Localizar o ponto de máxima eficiência do processo realizado,
- Implementar as modificações necessárias no ambiente operacional.

3.1 Composição do tempo de resposta

A Figura 3.1 ilustra a composição do tempo percebido por um usuário para obter a resposta de uma requisição de um processo informatizado. O tempo de resposta ponta a ponta pode ser dividido em dois componentes principais: tempo de rede externa e tempo no local de processamento. (1) O tempo de rede externa compreende o tempo gasto pelo tráfego das mensagens entre o local de trabalho do usuário e o local de processamento. (2) O tempo no local de processamento pode ser decomposto em dois componentes principais: o tempo de serviço e o tempo de fila. Tempo de serviço é o período de tempo durante o qual um pedido está recebendo serviço de um recurso, como processador, disco e rede local. Tempo de fila é o tempo gasto pelo pedido aguardando para ter acesso a um desses recursos. (Menascé e Almeida, 2002)

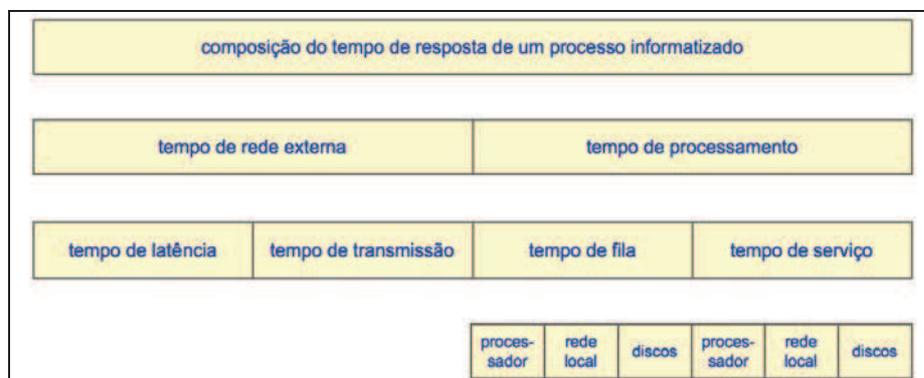


Figura 3.1 – Composição do tempo de resposta de um processo de tecnologia da informação

Uma requisição pode ter que visitar um recurso mais de uma vez antes que seja concluída. Por exemplo, mais de uma leitura ou gravação em disco ou muitas visitas ao processador, resultando em uma soma de tempos entre receber o recurso e esperar por ele.

A Equação 3.1 apresenta a composição numérica do tempo de resposta (R) de um processo da Tecnologia da Informação. O tempo de serviço (S) é aquele onde a requisição recebe tratamento do processador, discos e rede local ou linha de comunicação. A notação S_i^j é usada para indicar o tempo de serviço no recurso i durante a j -ésima visita ao recurso. O tempo de fila (F) é aquele onde a requisição aguarda para receber o atendimento de um recurso. A notação F_i^j representa o tempo gasto por um pedido aguardando para ter acesso ao recurso i durante a j -ésima visita (Menascé e Almeida, 2002). Para que a equação seja utilizada para a minimização do tempo de resposta, foram acrescentados os conceitos propostos por Goldberg e Luna (2000) e as limitações dadas pelo fabricante do equipamento utilizado no ambiente de estudo (IBM, 2005).

Equação 3.1 – minimização do tempo de resposta

Minimizar a função objetivo:

$$R = \sum_{\substack{\text{visitas } j \\ \text{recursos } i}} (S_i^j + F_i^j)$$

Sujeito às restrições:

- Política de localização de dados
- Política de desempenho
- Política de disponibilidade
- Política de nível de serviço
- Tecnologia utilizada

3.2 O tempo de resposta da tecnologia utilizada

Os valores do tempo de resposta esperado para a tecnologia utilizada no ambiente operacional serão obtidos através de benchmarks publicados por organizações criadas para essa finalidade, como por exemplo:

TPC - Transaction Performance Council: mede o tempo de transações para as principais cargas de trabalho utilizadas pelo mercado.

SPC - Storage Performance Council: mede o tempo de resposta de discos disponibilizados ao mercado. A Figura 3.2 ilustra o comportamento do tempo de resposta do disco IBM DS-8300 Turbo (SPC, 2006).

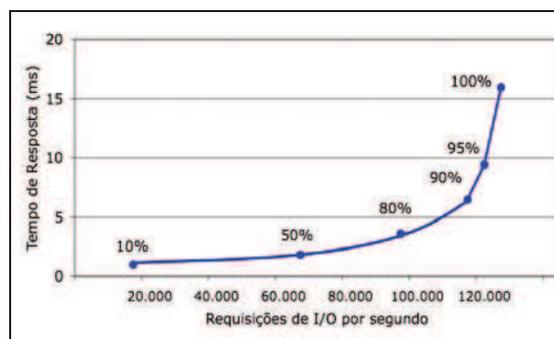


Figura 3.2 – Tempo de resposta do disco IBM DS-8300 em função da utilização, adaptada de (SPC, 2006)

É possível observar que ao atingir 100% da capacidade de utilização, os sistemas de armazenamentos de dados elevam em muitas vezes o tempo de resposta. A metodologia apresentada neste trabalho propõe que os dados sejam redistribuídos entre os discos existentes para obter tempo de resposta menor, sem que haja necessidade de aquisições ou atualizações tecnológicas do ambiente operacional.

Devido às similaridades entre os processos industriais e os da Tecnologia da Informação, algumas técnicas e conceitos desenvolvidos para a área industrial, como o conceito de eficiência, podem ser aplicados para direcionar o desempenho do processamento de dados realizados por computadores de grande porte.

3.3 Os conceitos de produtividade e eficiência utilizados na indústria

O alvo da eficiência é a eliminação do desperdício e tem recebido maior atenção e obtido maior avanço no setor de manufatura. Os desperdícios são encontrados em todo tipo de empresa e sua eliminação significa redução de custos e aumento da eficiência que é o maior interesse de todo homem de negócios, independente do ramo de sua atividade. É representada pela relação entre o desempenho real de uma operação e o desempenho ideal, máximo possível de ser alcançado, conforme a Equação 3.2. (Brisco, 2007)

Equação 3.2 – Índice de Eficiência (Brisco, 2007)

$$\text{Eficiência} = \frac{\text{Desempenho real}}{\text{Desempenho ideal}}$$

O desempenho real pode facilmente ser mensurado através da observação. A dificuldade vem em descobrir o desempenho ideal para ser colocado no denominador, que deve ser determinado por um perito que investigue máquinas, ferramentas, materiais, métodos, arranjos e verifique as melhores configurações para a execução de uma tarefa. (Brisco, 2007)

A eficiência pode também ser medida em relação ao tempo total para processamento, conforme indicado pela Equação 3.3 (Best, 2001).

Equação 3.3 – Índice de Eficiência de Processamento (Best, 2001)

$$EP = \frac{AV}{AV + NAV}$$

Onde: EP = eficiência de processamento medida em tempo,
AV = tempo de adição de valor,
NAV = tempo de não-adição de valor.

O tempo total que o material está no sistema de produção é dividido em 5 categorias: Operação, Transporte, Inspeção, Espera e Armazenamento. Somente a categoria “Operação” adiciona valor para o cliente, as outras 4 categorias não adicionam valor e reduzem a eficiência de processamento. (Best, 2001)

Ao discutir o desempenho de unidades de produção, é comum descrevê-las como sendo mais ou menos “eficiente”, ou mais ou menos “produtivo” (Lovell, 1993). Os termos, produtividade e eficiência, vêm sendo utilizados freqüentemente nos últimos anos como sinônimos apesar de não serem precisamente as mesmas coisas. A produtividade de um processo é a relação da saída que produz com os insumos que utiliza, conforme descrito pela Equação 3.4 (Coelli, 2006).

Equação 3.4 – Índice de produtividade (Coelli, 2006)

$$\text{Produtividade} = \frac{\text{Saídas}}{\text{Entradas}}$$

Entre duas empresas, a produtividade varia devido a diferenças na tecnologia da produção, a diferenças na eficiência do processo de produção e a diferenças no ambiente em que a produção ocorre. (Lovell, 1993)

Para ilustrar a distinção entre os termos, é útil considerar um processo de produção simples em que uma única entrada (x) é usada para produzir uma única saída (y). Na Figura 3.3, a linha OF' representa uma fronteira da produção que pode ser usada para definir o relacionamento entre a entrada e a saída. A fronteira da produção representa a saída máxima atingível de cada nível de entrada. O ponto A representa um ponto ineficiente e os pontos B e C representam pontos eficientes. Uma firma que opera no ponto A é ineficiente porque tecnicamente poderia aumentar a saída ao nível associado com o ponto B sem requerer mais entrada. (Coelli, 2006)

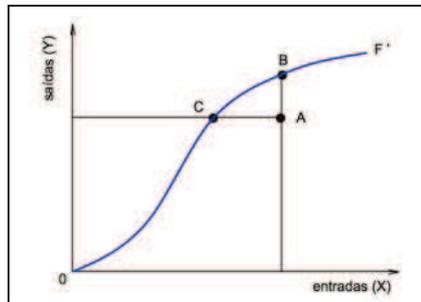


Figura 3.3 – Fronteiras da produção e eficiência técnica, adaptada de (Coelli, 2006)

A Figura 3.4 ilustra a distinção entre a produtividade, eficiência do processo de transformação e escala econômica. De acordo com a Equação 3.4, a inclinação do raio y/x fornece a medida de produtividade do ponto. Se a firma que opera no ponto A se deslocasse para o ponto B, tecnicamente eficiente, a inclinação do raio seria maior, implicando uma produtividade mais elevada no ponto B. Entretanto, quando move-se para o ponto C, encontra o ponto de eficiência máxima. O ponto C é o ponto de máxima eficiência. Este último movimento é um exemplo de exploração de economia de escala. A operação em qualquer outro ponto na fronteira da produção resulta em uma produtividade mais baixa. Dessa observação é possível concluir que uma firma pode ser tecnicamente eficiente mas sua produtividade pode ser melhorada através da exploração do ponto de máxima eficiência. (Coelli, 2006)

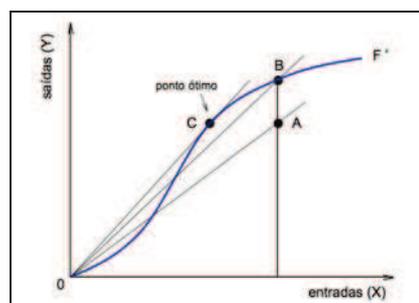


Figura 3.4 – Produtividade, eficiência e escala econômica, adaptada de (Coelli, 2006)

O ponto ótimo para o processamento de dados é encontrado na Lei de Amdahl para sistemas equilibrados. (Amdahl, 1970; Gray & Shenoy, 2000; Gray et al, 2006)

3.4 A Lei de Amdahl para sistemas equilibrados

De acordo com Brisco (2007), um perito deve estabelecer um padrão a ser seguido nas futuras execuções. Gene Amdahl estabeleceu esses parâmetros enquanto atuou como engenheiro chefe dos projetos de criação dos mainframes IBM /360 e /370. Durante as décadas de 60 e 70, ficou conhecido por muitas regras práticas para a engenharia de dados. Dentre elas, destaca-se a Lei dos Sistemas Equilibrados exposta a seguir.

A Figura 3.5 e a Equação 3.5 representam os valores definidos pela Lei de Amdahl: Um sistema equilibrado consumirá 8 MIPS (milhões de instruções por segundo) para processar cada MB/s (megabytes por segundo) transferidos de/para discos. (Gray e Shenoy, 2000; Gray et al, 2006)

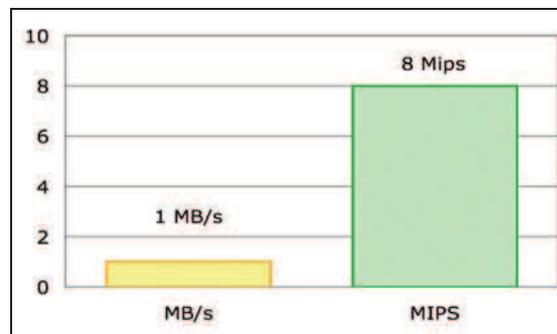


Figura 3.5 – Representação da Lei de Amdahl para sistemas equilibrados

Equação 3.5 – Consumo ideal de MIPS em um sistema equilibrado

$$\left[\begin{array}{c} \text{Consumo ideal de} \\ \text{processador em um} \\ \text{sistema equilibrado} \\ \text{(MIPS)} \end{array} \right] = 8 \times \left[\begin{array}{c} \text{Quantidade de} \\ \text{dados (MB/s)} \\ \text{lidos e gravados} \end{array} \right]$$

Essa regra prática, formulada para computadores de aplicação comercial fabricados na década de 60, permanece válida para os equipamentos atuais, de acordo com trabalhos realizados e publicados pelo laboratório de pesquisas da Microsoft. (Gray et al, 2006)

3.5 Relação entre complexidade e eficiência

A complexidade dos sistemas foi apontada pela Febraban (2007) como uma das causas do aumento do consumo de MIPS no sistema bancário brasileiro. Por isso, cabe analisar a validade da Lei de Amdahl para sistemas equilibrados (Amdahl, 1970; Gray 2006) em relação à complexidade dos sistemas. Para atingir essa meta existem duas alternativas: localizar trabalhos que já exploram essa relação ou, na falta destes, criar um ambiente de pesquisa para observar a relação entre o aumento da complexidade do sistema e o aumento do consumo de MIPS.

Para relacionar complexidade e consumo de MIPS será utilizado um grupo de aplicativos presente no ambiente de estudo deste trabalho. A complexidade será mensurada através de softwares disponíveis no mercado para essa finalidade. Eles serão relacionados e escolhidos

durante o período de desenvolvimento deste trabalho. Caso não seja possível incorporá-los ao ambiente de estudo, a complexidade será mensurada através de conceitos de FPA.

3.6 Descrição do ambiente operacional

Após a troca da versão do sistema operacional, o equipamento entrou em processo de saturação e não conseguiu atender todos os seus usuários. A solução de contorno, criada pela gerência de tecnologia da informação, foi dar prioridade ao sistema de atendimento ao público, transferir os lançamentos do sistema de manutenção de equipamentos para a baixa plataforma e atualizar o banco de dados durante o período da madrugada. Apesar das providências tomadas, a utilização do processador atingiu a média de 91%. Esse evento incentivou a busca de novo caminho pela implantação da metodologia e obter os resultados apresentados neste trabalho.

3.7 Fonte dos dados para análise e desenvolvimento do projeto

Os dados necessários para este trabalho são obtidos através dos relatórios RMF (*Resource Measurement Facility*), presentes em ambiente operacionais IBM z/OS. Eles permitem analisar o ambiente operacional, sem que seja necessária a aquisição de novos *softwares* ou novos produtos de monitoramento. Parte dos dados coletados foram utilizados para construir as Figuras 1.2 a 1.5.

Durante o desenvolvimento foram realizadas coletas de dados iniciais, finais e intermediárias para medir os resultados parciais e auxiliar na localização de novas tarefas.

Os gráficos 1.2 a 1.5 foram construídos dessa forma: (1) as linhas superiores representam a quantidade de MIPS utilizados durante o período e registrados nos relatórios RMF, (2) as linhas inferiores representam os resultados da multiplicação de 8 pela quantidade de MB/s transferidos de/para disco ao longo do período, segundo a Lei de Amdahl (1970). Para descaracterizar a identificação das empresas, a quantidade de MIPS foi trocada por porcentagem da capacidade de processamento.

A fim de cumprir o compromisso de número de páginas deste Consórcio, alguns gráficos foram retirados deste relatório para que fosse possível apresentar e obter contribuições da banca a respeito dos seguintes itens: (1) influência da complexidade sobre o consumo de MIPS, (2) contribuição do desempenho do processamento de dados sobre o desempenho da empresa. Os gráficos retirados foram: (1) Eficiência (antes e depois) dada pela Equação 3.2, (2) Eficiência de Processamento (antes e depois) dada pela Equação 3.3, representando as parcelas de tempo das operações em discos que adicionam e não adicionam valor para o cliente, (3) Índice de Produtividade (antes e depois) dada pela Equação 3.4 e que representa a relação de registros de entrada e registros de saída, foi possível concluir que existe maior quantidade de registros lidos que gravados e a produtividade pode ser aumentada com o uso adequado de índices de acesso, que permitam a localização de registros com menor quantidade de leituras.

3.8 Análise dos resultados

A Figura 3.6 registra a distribuição dos tempos das transações até o momento: 76% delas estão com até 0,3 segundos de duração. 24% das restantes, com tempo superior a 0,3 serão alvo do projeto de continuidade. No início do projeto o tempo médio de transação era de 29 segundos.

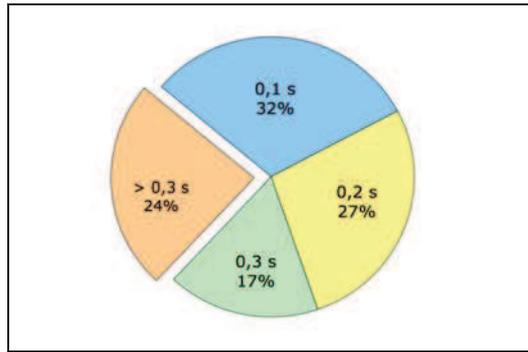


Figura 3.6 – Distribuição dos tempos de transações

As Figuras 3.7 a 3.8 registram a distribuição de carga de trabalho nos sistemas de armazenamentos de dados no início e no final deste trabalho. O círculo interno representa a meta para a distribuição homogênea da carga de trabalho: quando todos os discos trabalham igualmente, os acessos são uniformemente distribuídos, sem que haja sobrecarga de trabalho e excesso de filas em poucos deles.

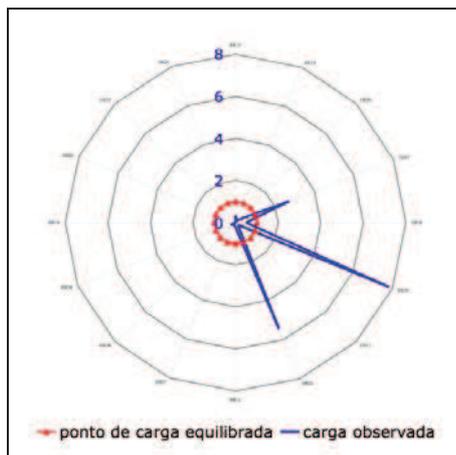


Figura 3.7 – Distribuição inicial da carga dos discos no ambiente 1

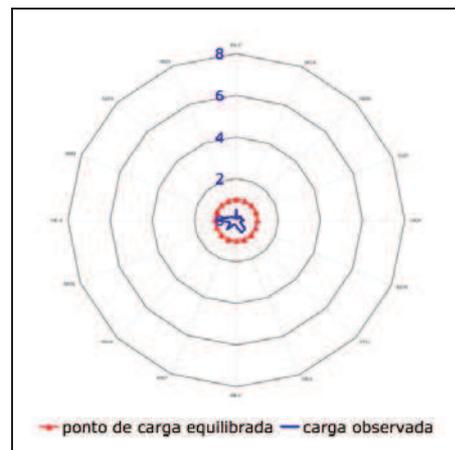


Figura 3.8 – Distribuição final da carga dos discos no ambiente 1

O tempo de fila para acesso aos discos é o componente que não agrega valor para o cliente, segundo o conceito de Best (2001) simbolizado pela Equação 3.3.

A Figura 3.9 permite observar que, após a redução das filas em discos, o tempo máximo das transações de um sistema crítico foi reduzido de 20 segundos para 4 segundos.

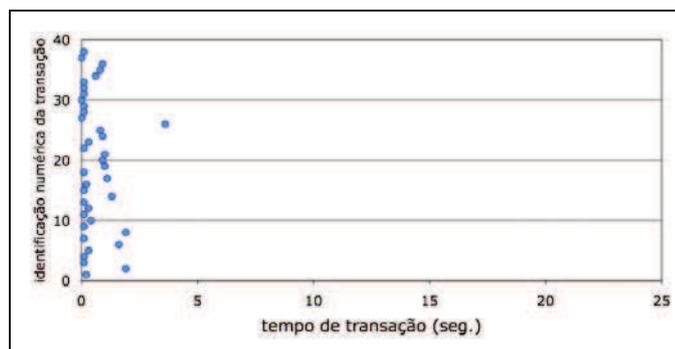


Figura 3.9 – Variação final do tempo de transação em um sistema crítico

As Figuras 3.10 e 3.11 registram a redução do consumo médio de processador de 91% para 43%, que resultou em melhor nível de serviços aos usuários finais e ao público atendido pela organização, além de redução em despesas de hardware e softwares vinculados ao consumo de MIPS.

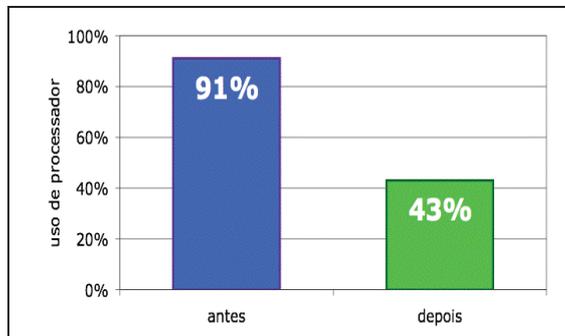


Figura 3.10 – Comparação do consumo de processador inicial e final

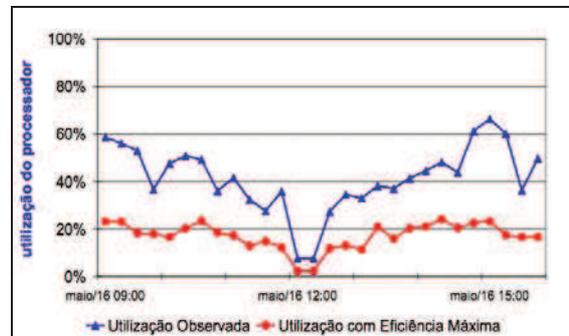


Figura 3.11 – nível observado X nível ideal do ambiente 1

3.9 Conclusões

A metodologia proposta permitiu: (1) localizar o ponto de eficiência máxima de um *mainframe*, (2) medir o potencial de ganho dos trabalhos de melhoria de desempenho a partir da relação entre a quantidade de dados processados e a quantidade de recursos computacionais utilizados, (3) aumentar a eficiência do processamento de dados de um *mainframe* através da redistribuição da carga de trabalho e redução das filas em discos.

4 Referências

- ALBERTIN, A. L.; ALBERTIN, R. M. M. Tecnologia da informação e desempenho empresarial: as dimensões de seu uso e sua relação com os benefícios de negócio. São Paulo: Atlas, 2005.
- BRISCO, N. A. Economics of efficiency. Whitefish, MT: Kessinger Publishing, 2007.
- COELLI, T. ET AL. An introduction to efficiency and productivity analysis. 2nd ed. New York: Springer, 2006.
- COMPUTERWORLD. MIPS management at mainframe organizations. Aug-2007. Disponível em : www.cio.co.uk/cmsdata/whitepapers/5595/Computerworld_Mainframe_management_research.pdf. Acesso em 07/01/2008.
- FEBRABAN. XVII Congresso e Exposição de Tecnologia das Instituições Financeiras. São Paulo, 2007. Disponível em: www.ciab.org.br/portugues/dados_setor/recursos_comput.asp. Acesso em 08/01/08.
- FITZSIMMONS, J. A.; FITZSIMMONS, M. J. Service Management: operations, strategy, information technology. 5th ed. New York: McGraw-Hill, 2005.
- FORTIER, P. J.; MICHEL, H. E. Computer System Performance Evaluation and Prediction. Burlington: Digital Pres, 2003.
- GOLDBARG, M. C.; LUNA H. P. L. Otimização combinatória e programação linear: modelos e algoritmos. Rio de Janeiro: Campus, 2000.

- GRAY, J.; SHENOY P. *Rules of thumb in data engineering*. 16th International Conference on Data Engineering, pp 3-12, San Diego, 2000.
- GRAY, J.; BELL, G.; SZALAY, A. Petascale computational systems. *Computer*, vol. 39, no. 1, pp. 110-112, Jan., 2006.
- HENNESSY, J.; PATTERSON D. *Computer architecture: a quantitative approach*. 4th ed. San Francisco: Morgan Kaufmann, 2007.
- IBM. Redbook *ABC of z/OS system programming V. 11: capacity planning and performance management*. Manual técnico SG24-6327-00. Poughkeepsie, NY: IBM ITSO, 2005.
- LAIRD, L. M.; BRENNAN, M. C. *Software measurement and estimation: a practical approach*. Hoboken, NJ: John Wiley & Sons, 2006.
- LOVELL, C. A. K. Production frontiers and productivity efficiency. In: Fried, H. O.; Lovell, C. A. K.; Schmidt, S. S. (ed.) *The measurement of productivity efficiency: techniques and applications*. New York: Oxford University Press, 1993.
- MEIRELLES, F. S. *Pesquisa anual: administração de recursos de informática*, 18^a ed. São Paulo: CIA - Centro de Informática Aplicada da Eaesp-FGV, 2007.
- MCCABE, T. J. A complexity measure. *Proceedings of the IEEE Transaction on Software Engineering*, Volume SE-2, Number 4, Dec 1976, pp 308-320.
- MENASCÉ, D.; ALMEIDA, V. *Capacity planning for web services: metrics, models and methods*. Upper Saddle River: Prentice Hall Inc, 2002.
- PANDIAN, C. R. *Software metrics: a guide to planning, analysis and application*. Boca Raton, FL: CRC Press, 2003.
- PATTERSON, D. A.; CHEN, P. M. Storage performance: metrics and benchmarks. *Proceedings of the IEEE* Volume 81, Issue 8, Aug 1993, pp 1151-1165.
- SAP. SD Parallel Standard Application Benchmark Results - With Round Robin. Disponível em: http://www.sap.com/solutions/benchmark/sd1_results.htm. Acesso em: 16/10/2007.
- SEBESTA, R. W. *Concepts of programming languages*. 8th ed. Boston: Addison-Wesley, 2007.
- SPC. Benchmark executive summary: IBM System Storage DS8300 Turbo. Dec 5, 2006. Disponível em: http://www.storageperformance.org/results/a00049_IBM-DS8300-Turbo_SPC1_executive-summary.pdf. Acesso 16/10/2007.
- SPRUTH, W. The future of mainframe. Euro CMG 2007, Nürnberg 23-25 Mai, 2007.
- VAN DER ZEE, H. *Measuring the value of information technology*. Hershey, PA: Idea Group Publishing, 2003.