



UNIVERSIDADE DO ESTADO DE MINAS GERAIS
FUNDAÇÃO EDUCACIONAL DE ITUIUTABA
CURSO DE SISTEMAS DE INFORMAÇÃO



Apostila de Modelagem e Avaliação de Desempenho

Prof. Walteno Martins Parreira Júnior

www.waltenomartins.com.br

waltenomartins@yahoo.com

2010

SUMÁRIO

1.	Apresentação.....	2
1.1	Abordagens para avaliar o desempenho de sistemas.....	2
1.2	O processo de modelagem e suas fases componentes.	4
1.3	Bibliografia citada:	4
2.	Introdução.....	6
2.1	Considerações preliminares	6
2.2	Considerações acerca da implantação e operação de sistemas computacionais em uma empresa	6
2.3	Resultados.....	8
3.	Áreas de Aplicação e objetivos.....	8
4.	Técnicas	10
4.1	kernel	10
4.2	Benchmark	11
4.3	Programas sintéticos	12
4.4	Modelos Analíticos.....	12
4.4.1	Um sistema de fila e seus elementos	13
4.5	Simulação.....	14
4.6	Monitoramento	14

1. Apresentação

1.1 Abordagens para avaliar o desempenho de sistemas

Restringindo-se aos sistemas que geram algum tipo de bem de consumo ou de serviços, tais como os de manufatura, de transporte ou os computacionais, pode-se ter a noção da complexidade do tema. Os sistemas são os mais diversos possíveis, e cada qual possui as suas peculiaridades, sensibilidades e problemas, o que gera um conjunto de incertezas sobre como tratar o assunto.

Com a evolução tecnológica experimentada nas últimas décadas, outros tipos de requisitos foram acrescentados aos sistemas que fornecem serviços aos seus usuários, tal como desempenho, aqui com a conotação de algum tipo de satisfação gerada pela eficiência na realização da tarefa. De maneira que a execução dessa tarefa possa ser mais ou menos eficiente, tomando-se algum parâmetro como critério.

A maneira pela qual será avaliado o desempenho depende diretamente das características do sistema envolvido. Tratando-se de sistemas computacionais, pode-se dividir a avaliação em dois paradigmas distintos:

- Aquele que precisa obter medidas no próprio sistema, e a partir delas o avaliador pode estabelecer uma interpretação em relação ao desempenho,
- Aquele que se baseia em criar-se uma abstração do sistema, através da qual se pode estabelecer um certo prognóstico (com um grau de incerteza associado) a respeito do desempenho desse sistema.

Logicamente, há uma relação custo/benefício embutida na questão dos paradigmas, que está relacionada ao quanto se pode permitir de intrusão e qual o grau de imprecisão que é aceitável. Contudo, há outra questão que deve ser levantada: quando se pretende ter a idéia daquilo que o sistema pode proporcionar? Essa talvez seja a questão que norteie a escolha por um dos paradigmas. Se uma avaliação *post-mortem* é admitida, então pode ser mais natural utilizar-se a idéia da instrumentação do sistema. Entretanto, o atual estado da arte de certa forma leva à noção de que é necessário um conhecimento antecipado daquilo que está por acontecer, ou seja, de tudo que um determinado sistema pode oferecer aos seus potenciais usuários, antes mesmo que esse sistema exista. De certa forma, esse grau de "maturidade" se deve ao montante de custos e esforços envolvido, na tentativa de minimizar as perdas.

No paradigma da abstração, comumente referenciado como modelagem, que em uma definição simplista, baseia-se em uma associação de uma representação gráfica, denominada de modelo, e um método de cunho aleatório, que forneça medidas para avaliar-se o desempenho do sistema que foi modelado. O termo modelagem possui várias conotações nas subáreas da ciência da computação, tal qual a modelagem de dados utilizada na engenharia de software. Entretanto, modelagem para fins de avaliação de desempenho, como assumido nesta disciplina, é um processo complexo que fornece medidas pelas quais pode-se ter uma idéia do comportamento do sistema, geralmente a longo prazo.

Talvez pelo fato de a avaliação de desempenho ser uma área proveniente da teoria das probabilidades, de certa forma, tenha-se negligenciado em relação à representação do

modelo, dando-se uma ênfase maior à solução matemática aplicada a este modelo. Fazendo-se conjecturas, talvez pelo grau de especialização das pessoas envolvidas com o tema, um conjunto de equações seja tão (ou mais) claro que uma representação gráfica do mesmo problema. Essa possibilidade, apesar de compreensível, pode gerar um certo distanciamento entre o desenvolvimento do conhecimento e sua efetiva utilização pela comunidade menos especializada, que em tese deveria ser o alvo dos benefícios desse desenvolvimento.

Para delimitar-se o problema a ser estudado, o contexto de avaliação de desempenho que se pretende investigar no decorrer destas aulas está relacionado ao fato de que o ingrediente determinante no desempenho de um sistema está associado à disputa dos recursos disponíveis pelos clientes que utilizam os serviços prestados por esses recursos. Assim, o sistema (para fins de avaliação de desempenho) pode ser visualizado como um conjunto de recursos que prestam serviços (comumente chamados de servidores) e de clientes que solicitam algum tipo de atendimento nesses recursos (eventualmente gerando filas de solicitações em um determinado servidor). Para estudar o relacionamento entre clientes e servidores, existe uma vertente dos processos estocásticos, denominado de teoria das filas.

Partindo da premissa de que a disputa pelo recurso é determinante no desempenho e ainda que tão importante quanto a solução dada ao modelo é a sua especificação, veremos que existem várias formas de obter-se a união especificação/solução.

No que concerne à idéia de vislumbrar-se um sistema como um conjunto de estados, aos quais são aplicados certos eventos, que fazem com que haja transições entre esses estados, à medida que o tempo passa. É importante ressaltar que essa é a abordagem usada nas técnicas de especificação referenciadas e, via de regra, na literatura pertinente. Dessa forma, tanto redes de filas, redes de Petri e Statecharts, diferem basicamente na funcionalidade gráfica oferecida.

Entretanto, o fato das técnicas de especificação serem conceitualmente semelhantes pode levar à suposição equivocada que elas possuem o mesmo poder de representação, o que não se constitui em uma regra. Na verdade, originalmente, as redes de filas foram idealizadas para um propósito bem mais específico (o de representar sistemas de filas), daí advém o seu tipo de representação peculiar. Já redes de Petri e Statecharts foram concebidos inicialmente para representar sistemas genéricos (não necessariamente sistemas de filas), o que proporcionou um poder de representação relativamente mais amplo. Contudo, redes de Petri e Statecharts possuem representações diferenciadas, cada uma com suas características próprias - diferenças enfatizadas no decorrer do curso.

As redes de Petri, idealizadas em (Petri, 1966), já possuem suas extensões voltadas à avaliação de desempenho, tais como as redes de Petri estocásticas (Moloy, 1982) e redes de Petri estocásticas generalizadas (Chiola et al., 1993). Já os Statecharts não tinham o caráter estocástico contemplado. Essa situação começou a ser modificada com alguns trabalhos que apontaram para essa direção, tais como (Francês, 1998) e (Vijaykumar, 1999). Esses trabalhos associavam os Statecharts às cadeias de Markov a tempo discreto e contínuo, respectivamente, provendo dessa maneira uma solução analítica aos Statecharts.

Esta tese apresenta uma generalização dos trabalhos citados anteriormente, os Statecharts Estocásticos, que oferecem templates (conjuntos de estados e eventos pré-definidos para designar fontes de geração de clientes, filas, servidores e sorvedouros), através dos quais se pode descrever o funcionamento básico de um sistema de fila, sem uma modificação gráfica dos elementos básicos dos Statecharts. Essa especificação é genérica o suficiente de tal forma que seja possível a representação de elementos distintos

às filas e aos servidores. Além dos Statecharts Estocásticos, também é trabalhada uma representação bem mais específica voltada exclusivamente aos sistemas de filas, os Queuing Statecharts. Essa especificidade é, na verdade, uma aglutinação entre as representações dos Statecharts e das redes de filas, herdando destas a estrutura diagramática e daqueles a abordagem de propagação de eventos e ações entre os diversos componentes do modelo (centros de serviço).

O ponto-chave da análise está no fato de que sistemas de filas possuem um comportamento relativamente uniforme, quando observados em função dos eventos que são relevantes à avaliação de desempenho. Assim, para descrever-se um sistema de filas há alguns eventos básicos que devem ser levados em consideração, tais como a geração dos clientes a uma determinada taxa, a chegada desses clientes à fila, a tomada e a liberação de um determinado servidor por um cliente. De certa forma, há uma relação, nem sempre tão explícita, entre esses eventos e os eventos constantes nos métodos de solução (tanto analíticos quanto por simulação).

1.2 O processo de modelagem e suas fases componentes.

As técnicas de especificação que podem ser utilizadas: redes de filas, redes de Petri e Statecharts. Cada técnica é observada sob o prisma de sua representação. Sobre cada uma das técnicas, são feitas considerações sobre vantagens e desvantagens da representação.

Redes de filas são abordados aspectos como elementos básicos de sua representação, distribuições de probabilidades associadas a cada elemento, composição e estruturas possíveis de centros de serviço e redes abertas, fechadas e mistas.

Redes de Petri são apresentados os elementos básicos, uma definição formal, as redes elementares através das quais são compostas todas as demais, as redes coloridas e, sucintamente, as estocásticas.

Statecharts são apresentados os elementos básicos (principalmente estados e eventos), algumas características peculiares, tais como paralelismo e hierarquia entre componentes, entrada por condição, estados parametrizados e delay e timeout.

Nesta disciplina, necessita-se da definição de conceitos básicos de estatísticas, tais como: população e amostra, medidas de locação (por exemplo, médias, mediana e moda) e medidas de dispersão (tais como, variância e desvio padrão). Da teoria das probabilidades, usando as suas definições, tais como: variáveis aleatórias, definições de distribuições discretas e contínuas, algumas das principais distribuições discretas (binomial e de Poisson) e contínuas (normal, exponencial, hiperequencial e de Erlang). Como extrapolação da teoria das probabilidades, são ilustrados conceitos de processos estocásticos e, em particular, o processo estocástico denominado de cadeias de Markov (tanto a parâmetro discreto quanto a contínuo).

1.3 Bibliografia citada:

CHIOLA, G., MARSAN M. A., CONTE, G. Generalized Stochastic Petri Nets: A Definition at the Net Level and Its Implications. *IEEE Transactions on Software Engineering*, vol. 19, n. 2, p. 89-106, 1993.

FRANCÊS, C.R.L. *Stochastic Feature Charts - Uma Extensão Estocástica para os Statecharts*. São Carlos, 1998. 125p. Dissertação (Mestrado) - Instituto de Ciências Matemáticas e de Computação (ICMC), Universidade de São Paulo (USP).

MOLLOY, M.K. Performance Evaluation Using Stochastic Petri Nets. *IEEE Trans. Comput.*, v. C-31, n. 9, p. 913-17, 1982.

PETRI, C.A. *Kommunikation mit Automaten*. Schriften des IIM Nr. 2, Institut für Instrumentelle Mathematik, Bonn, 1962. English Translation: Technical Report RADC-TR-65-377, Griffiths Air Force Base, New York, v. 1, Suppl.1, 1966.

SANTANA, M.J., SANTANA, R.H.C., FRANCES, C.R.L. Avaliação e Análise de Desempenho de Sistemas Computacionais: Técnicas e ferramentas, São Carlos, 1997. Instituto de Ciências Matemáticas e de Computação (ICMC), Universidade de São Paulo (USP).

VIJAYKUMAR, N.L., CARVALHO, S.V., ABDURAHIMAN, V. *Statecharts: Their Use in Specifying and Dealing with Performance Models*. São José dos Campos, 1999. Tese (Doutorado) – Instituto Tecnológico da Aeronáutica (ITA).

2. Introdução

2.1 *Considerações preliminares*

A crescente demanda de recursos financeiros utilizados na instalação, manutenção, operação e programação de equipamentos de computação tem revelado a fundamental importância de uma utilização racional e efetiva das facilidades que tais sistemas proporcionam.

Sabe-se então que o estudo de desempenho de equipamentos de computação constitui-se tão- somente um aspecto do problema global de controle de custos de processamento de dados. A necessidade permanente de se alocar recursos financeiros limitados invoca o uso de técnicas de avaliação de desempenho de modo a tornar possível uma análise de custo-eficiência que abranja o sistema de computação como um todo e não apenas a máquina propriamente dita.

Apesar da contínua melhoria da relação preço/performance dos sistemas modernos, não parece que a necessidade de otimização desses sistemas desaparecerá. Pelo contrário, quanto mais poderosos for o hardware e mais sofisticado o software de um sistema, tanto maior será a distância entre boa e má performance, e mais amplos os objetivos de um estudo de desempenho.

2.2 *Considerações acerca da implantação e operação de sistemas computacionais em uma empresa*

A única justificativa para a utilização de um computador por parte de uma empresa é a certeza de que sua implantação possibilitará sejam atingidos benefícios bem definidos no contexto das atividades da empresa, os quais seriam impossíveis ou anti-econômicos sem o seu uso.

Antes da decisão de adoção do processamento eletrônico de dados de uma empresa, seus administradores devem se conscientizar de que o computador é uma ferramenta poderosa e cara dentro de uma organização.

Sugere-se que o dimensionamento da configuração a ser adquirida seja feito de um modo conservativo, o sistema escolhido deve ser tal que, se mantido totalmente ocupado, completará sua carga de trabalho em tempo (e com uma margem de segurança).

Decisões puramente emotivas(e por isso sujeitas a pressões) e/ou baseadas em justificativas tecnicamente mal fundamentadas resultam, muitas vezes, numa “pressa” de gastar dinheiro que está longe de ser uma demonstração de bom senso.

Em muitos casos, a saturação prematura de um sistema pode ser facilmente explicada a luz de alguns fatos, tais como:

- a) a ineficiência ou mesmo total inexperiência de escalonamento de serviços, levando a carga do sistema a composição altamente desfavoráveis ao rendimento do sistema;
- b) a existência de programas codificados ineficientemente e com alta frequência de submissão e/ou consumo de unidade central de processamento (UCP);

- c) o mau dimensionamento e/ou seleção de equipamentos tendo em vista outras opções, compatíveis e mais vantajosas;
- d) a inadequação de vários parâmetros do sistema operacional atribuídos quando da geração do sistemas ou de suas posteriores inicializações;
- e) a inexistência de uma mentalidade voltada para melhor utilização dos sistemas nos vários níveis de atividade da instalação (operadores, programadores, analistas, usuários e gerentes)

A análise de desempenho de um sistema de computação pode resultar em grande economia para uma instalação.

De um modo geral, o ciclo de vida de um centro de processamento de dados pode ser descrito como se vê na figura 1.

As decisões inerentes a cada uma das quatro fases são de estrita responsabilidade da administração da empresa não devendo, ser delegadas a organização externa e, especialmente, ao vendedor.

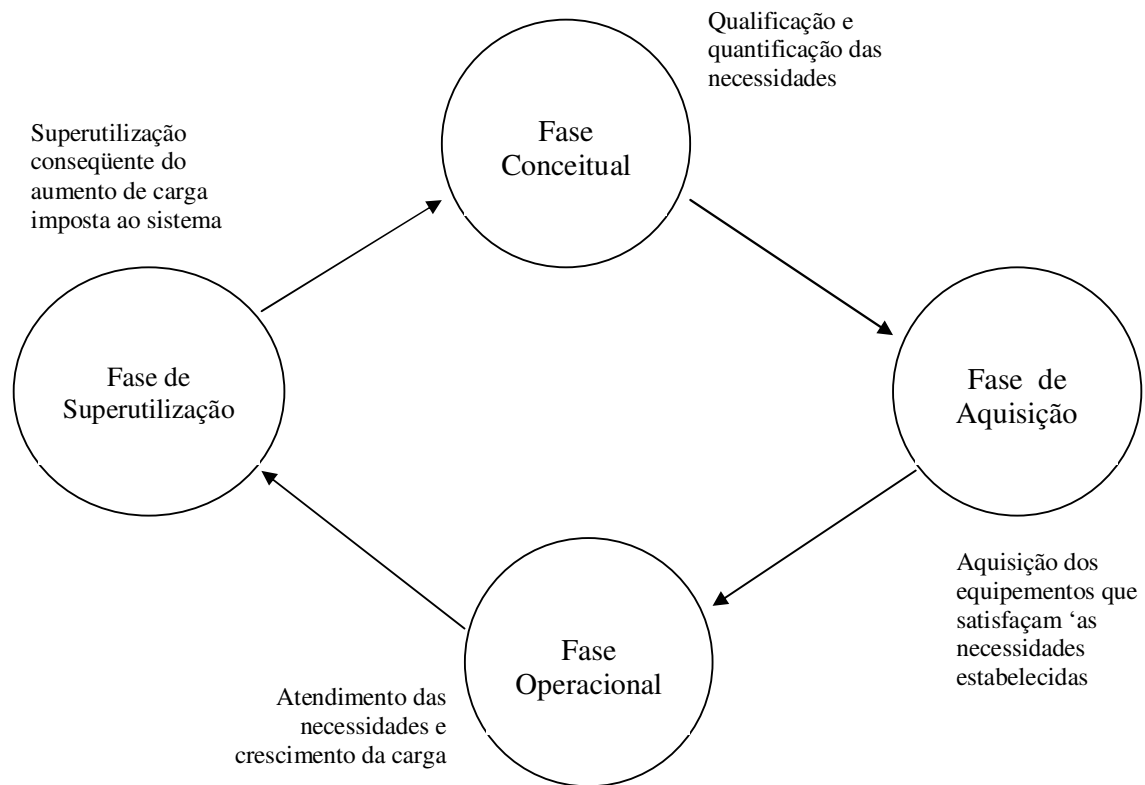


Figura 1 – Ciclo de vida de um CPD

A resolução dos problemas de casa instalação requer compreensão de suas dificuldades potenciais e a realização de medições específicas que indiquem quais problemas estão realmente ocorrendo. Essas atividades poderão proporcionar benefícios em cada uma das quatro fases mostradas na Figura 1, quais sejam:

- a) Fase conceitual – A qualificação e a quantificação das necessidades são feitas de uma maneira mais racional e objetiva.
- b) Fase de aquisição – O dimensionamento dos dispositivos a serem adquiridos é feito com maior sensibilidade, permitindo menores margens de tolerância.
- c) Fase operacional – Atinge-se um completo domínio do sistema do acompanhamento do crescimento de sua carga, resultando uma utilização mais eficiente e permitindo seja adiada ao máximo a entrada na fase de superutilização.
- d) Fase de superutilização – Seu início pode ser previsto com exatidão, de modo a serem evitadas as crises operacionais decorrentes da falta de conhecimento acerca da capacidade residual de processamento do sistema.

2.3 Resultados

Há aparentemente um sentimento generalizado – detectado principalmente em níveis gerenciais – de que avaliação e otimização de desempenho são assuntos extremamente sofisticados, que acarretam custos consideráveis e com resultados nem sempre convincentes.

O conceito de sistema de computação, é aqui utilizado como extensivo a toda a instalação, abrangendo, pois dos aspectos gerenciais à configuração do hardware, passando pelas áreas de operação, sistemas operacionais e programação.

3. Áreas de Aplicação e objetivos

Há três grandes áreas nas quais a análise de desempenho de sistemas de computação deve ser aplicada:

- a) no **projeto** de componentes de hardware e/ou software de novos sistemas, como ferramenta de avaliação progressiva durante o decorrer do projeto.
- b) Na **seleção** de sistemas, como instrumento de decisão na aquisição ou substituição de componentes de hardware e/ou software.
- c) Na **avaliação** de sistemas já implantados, para otimização, previsão e planejamento de futuras alterações.

No caso A, o objetivo é determinar o desempenho de um sistema que ainda não existe fisicamente por achar-se em fase de projeto.

No caso B, o aspecto mais importante é a determinação do desempenho relativo de todas as opções possíveis para aquisição de novos sistemas.

No caso C, enfoque principal é o da obtenção de dois grandes objetivos: extrair a máxima produtividade do sistema e oferecer altos padrões de serviços aos usuários. A coleta e a análise de dados minuciosos permitirão que se efetuem com segurança a previsão do impacto de mudanças no sistema, a projeção de sua carga à vista do crescimento vegetativo e/ou da implantação de novas aplicações, e ainda a determinação de um perfil do sistema.

As técnicas existentes podem mostrar diferentes graus de adequação a cada uma das três áreas definidas, conforme a tabela 1.

O valor pode variar de zero a três de conformidade com a sua aplicabilidade, onde:

0 – Técnica não aplicável na área;

1 – Técnica aplicável, porém inadequada;

2 – Técnica aplicável porém insuficiente, devendo ser utilizada em conjunto com outras técnicas;

3 – Técnica satisfatória.

Técnicas	Área de Aplicação					
	Projeto		Seleção		Avaliação	
	Hardware	Software	Hardware	Software	Hardware	Software
Kernel	3	1	2	1	0	2
Benchmark	0	1	3	3	2	2
Programas Sintéticos	2	2	3	3	2	2
Modelos Analíticos	3	1	2	1	2	1
Simulação	3	3	2	2	2	2
Monitoramento	2	2	2	2	3	3

Tabela 1 – Graus de adequação de cada técnica a área de aplicação

A figura 2 ilustra de forma gráfica a aplicabilidade das várias técnicas quanto a cada área de aplicação. As linhas cheias indicam que a técnica é satisfatória para a área apontada enquanto as linhas tracejadas indicam que a técnica deve ser utilizada apenas como ferramenta auxiliar na área correspondente.

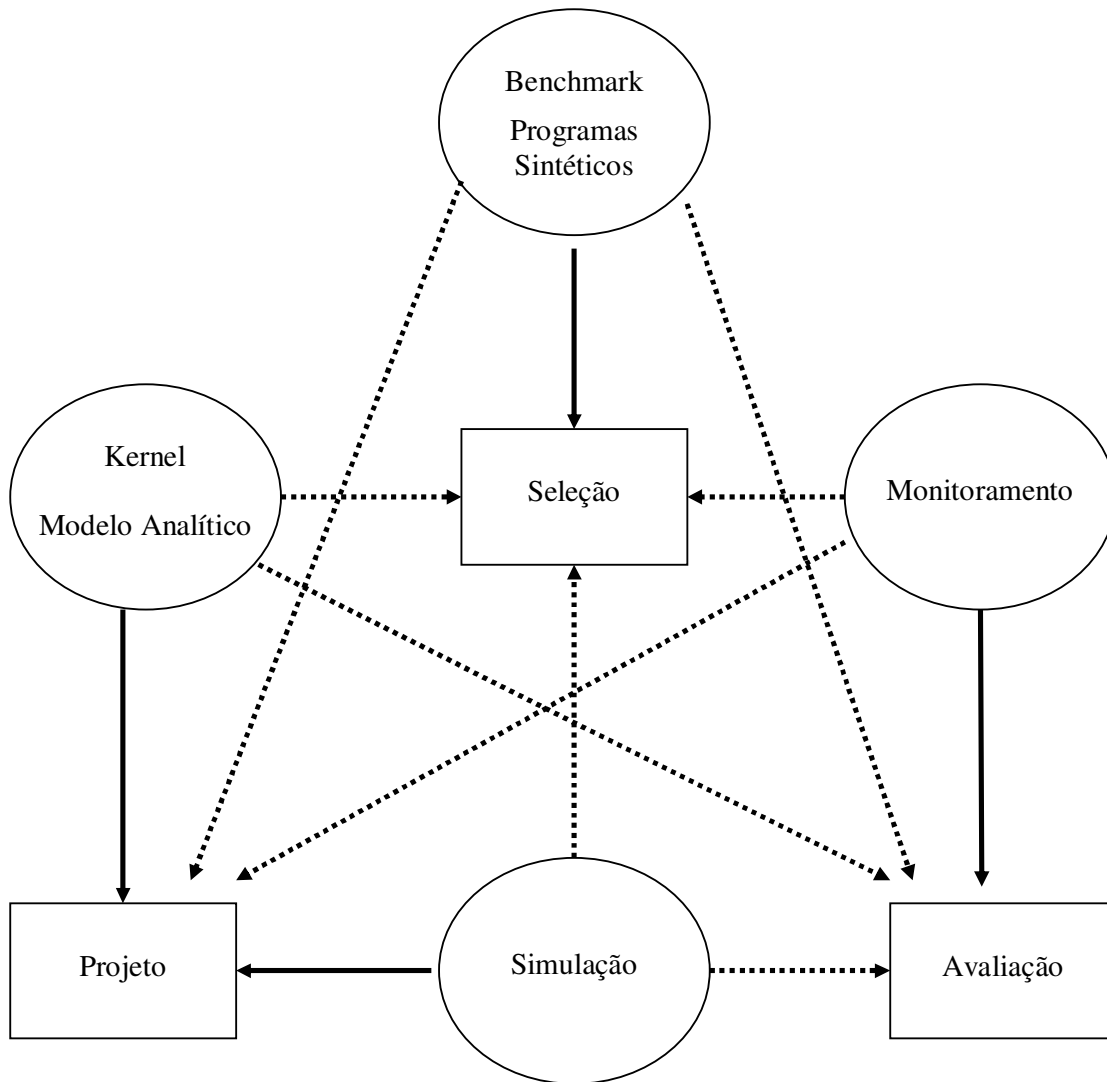


Figura 2 – Aplicabilidade das técnicas para cada área de aplicação

4. Técnicas

4.1 *kernel*

Esta técnica é assim chamada porque consiste no exame da parte central, ou essencial, de uma aplicação. As partes mais frequentemente usadas de uma aplicação são determinadas e a seguir programadas. É possível, avaliar-se qual delas executa o programa de maneira mais eficiente, é elaborado um conjunto de programas kernel, atribuindo-se a cada um deles um peso relativo, determinado em função de sua importância, do quão crítico ele é, etc.

Um programa Kernel é, uma mistura de instruções que compõem um programa ou parte de um programa e cujo tempo de execução é determinado com base nos tempos de instrução fornecidos pelo fabricante.

Seja, por exemplo, fluxo de uma aplicação tal como o que se descreve na figura 3.

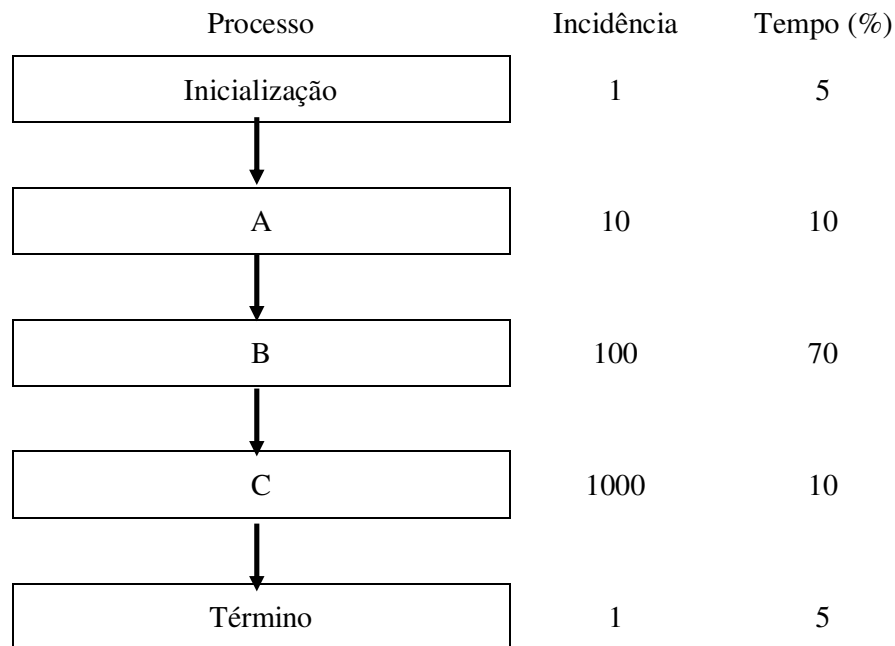


Figura 3 – Fluxo de uma aplicação

Embora a seção C seja executada mais vezes que as demais, a escolha deve recair sobre a seção B, a qual, em termos do percentual do tempo consumido, demonstra ser a verdadeira porção crítica

A análise do kernel é especialmente útil na comparação de sistemas de arquiteturas diferentes já que ela permite aplicar a um problema particular o repertório de instruções de cada sistema em sua totalidade.

4.2 Benchmark

Um Benchmark consiste em um conjunto de programas selecionados de maneira a construir uma composição representativa de carga de uma instalação e processados num sistema que se deseja avaliar.

Essa técnica permite determinar se um sistema particular pode processar adequadamente uma carga real, se as características de preço X desempenho do sistema proposto são suficientes, e quais as melhores opções de configuração, tanto de hardware quanto de parâmetros do sistema operacional, constituindo-se, pois, em uma técnica fundamentalmente voltada para a área de seleção de sistemas.

Ela pode também ser útil na avaliação de sistemas já existentes, quando se deseja verificar se uma alteração produzirá os efeitos esperados sobre o seu desempenho no processamento de carga corrente.

No que concerne á composição do benchmark, podem-se conceituar dois tipos: as cargas naturais e as cargas artificiais.

As cargas naturais correspondem exatamente à carga real sem quaisquer limitações. As cargas artificiais, por sua vez, podem ser de duas espécies: as cargas sintéticas, projetadas e implementadas independentemente da carga real e as cargas híbridas, provenientes de partes de carga real.

De um modo geral, as cargas naturais necessitam de períodos de medição muito longos para que se possa garantir um alto grau de representatividade, o que não ocorre como as cargas artificiais. Isso constitui uma das razões básicas pela qual as cargas artificiais são geralmente preferidas.

4.3 Programas sintéticos

A carga real de um sistema, isto é , a carga gerada pelos usuários do environment normal, é geralmente irreproduzível em sua composição exata. Entretanto, se as propriedades estatísticas de carga não mudam com o tempo, ela é estatisticamente reproduzível. Por outro lado, embora a carga real possa ser usada em experiências de avaliação, os intervalos de medição necessários podem mostrar-se excessivamente longos, resultando numa grande quantidade de dados coletados. No caso do Benchmark, a seleção de Jobs pode constituir-se numa tarefa árdua. Em certos casos, pode haverá necessidade de se codificar programas complexos e converter seus arquivos de dados de modo a se poder comparar sistemas diferentes. Os custos de tais atividades podem tornar-se proibitivos. Por essas razões é preferível, e em alguns casos essencial, construir uma carga-piloto, mesmo quando é fisicamente factível medir o sistema sob carga real.

Um programa sintético simula a demanda de utilização dos recursos do sistema de maneira requerida pela carga, mas não executa em verdade trabalho “útil”. Ele é usado em um sistema como carga-piloto reproduzível, podendo ter suas características modificadas por meio de um conjunto de parâmetros de entrada.

Um exemplo clássico de programa sintético é o proposto por Buchholz para fins de comparação de diferentes sistemas. Na ausência de uma definição teórica de desempenho, o autor propõe que se use como grandeza comparativa entre sistemas o tempo de execução despendido. Programa consiste essencialmente em um procedimento simplificado de manutenção de arquivos no qual, a cada match encontrado entre um registro do arquivo mestre e um registro do arquivo de transações, é executado um loop totalmente limitado por UCP. Tendo que o numero de registros dos arquivos mestre e de transações, e o numero de vezes em que o loop deve ser percorrido, são parâmetros de entrada do programa, este comporta-se à como um job totalmente limitado por UCP, totalmente limitado por E/S ou como um Job localizado em qualquer ponto entre esses dis extremos.

Assim, qualquer job constituinte de carga real de uma instalação pode ser representado pelo programa de Buchholz.

4.4 Modelos Analíticos

Um modelo analítico trata de representação matemática de um sistema, sendo particularmente numerosos os baseados na teoria das filas. Em geral sã utilizados para fornecer dados relativos ao desempenho de um componente particular do sistema, tal como

uma unidade de disco, de UCP ou de certos algoritmos do sistema operacional. Alguns modelos podem ser facilmente alterados de modo a refletir, com novos, parâmetros, as modificações desejadas.

O aspecto mais desfavorável dessa técnica consiste em sua dificuldade de retratar em fidelidade os efeitos das interações existentes entre os componentes de hardware, os módulos do sistema operacional e os programas de aplicação. Além disso, há casos em que o modelo pode-se tornar muito complexo. Nestes casos, as hipóteses simplificadoras tendem a reduzir sua validade.

4.4.1 Um sistema de fila e seus elementos

A Análise de filas é um ramo da teoria da probabilidade que estuda as características de sistemas sujeitos a fluxos aleatórios de sua variável de entrada. O sistema estudado pode ser um supermercado, um processador de mensagens em tempo real, uma estação de telefônica, etc. idealmente, o comportamento de cada um desses vários sistemas pode ser representado matematicamente e seus efeitos esperados, determinados a partir de uma análise adequada dos vários modos de operação.

Uma fila ocorre quando um usuário chega a uma facilidade de serviço e a encontra ocupada, sendo obrigado a esperar. Três elementos podem ser caracterizados para que o sistema de fila possa ser descrito matematicamente: O tráfego de entrada, a linha de espera e a facilidade de serviço, conforme mostra a figura 4.

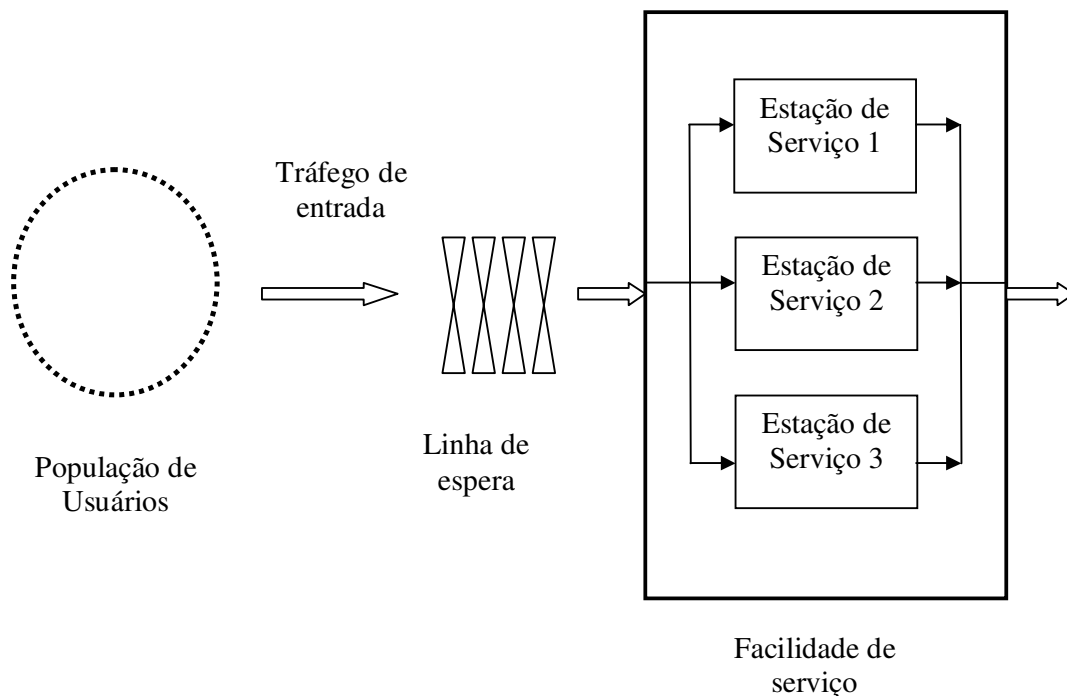


Figura 4 – Elementos de um sistema de fila

4.5 Simulação

A técnica de simulação envolve a construção de modelos da estrutura do sistema a ser estudado e a posterior aplicação de tais modelos. Num modelo de simulação, o hardware, o software, os programas de aplicação e os mecanismos internos do sistema devem ser especificados. Uma vez desenvolvido o modelo, a simulação é exercitada pela ocorrência de eventos consecutivos, tais como atualizações, transações, consultas, interrupções, retardos, esperas, etc. Blocos do simulador são usados para gerar tais eventos com quaisquer taxas de chegada, rotas ou distribuições. A cada evento, por meio de parâmetros a ela associado, podem ser impostas características, tais como propriedades, comprimentos de mensagem, origem, destino, etc.

Um estudo de simulação pode ser dividido em seis partes:

- Enunciado do problema
- Desenvolvimento do modelo
- Estimativa do desempenho esperado
- Construção de modelo de simulação
- Validade do modelo
- Execução de avaliação, usando simulação.

Alguns aspectos a serem considerados quando se aplica são:

A simulação é uma técnica dispendiosa devido à grande quantidade de tempo e recursos humanos gastos na implantação e execução.

Uma das limitações da técnica é a eventual necessidade de adoção de hipóteses simplificadoras para fins de exercitação do modelo.

A técnica possibilita serem conseguidos bons resultados na seleção de sistemas, podendo ser de grande utilidade para prever o comportamento resultante de alterações propostas.

4.6 Monitoramento

A técnica do monitoramento consiste na coleta de dados sobre o desempenho de um sistema existente e sua posterior análise. É usado em geral tanto para a localização de pontos de saturação, que produzem a eficiência de uma instalação, quanto para a obtenção de um perfil operacional de um sistema.

Uma definição mais precisa diz ser o monitoramento um processo de medir o sistema e entender os eventos internos e suas inter-relações, de modo a permitir que pequenas alterações no hardware e/ou no software possam acarretar considerável melhoria em seu desempenho. É, uma técnica voltada inteiramente para a avaliação de sistemas existentes.

Três fases podem ser definidas no processo de monitoramento:

- a) **Coleta de dados:** Executada durante a operação normal do sistema e que gera grande quantidade de dados.

- b) **Redução de dados:** Constitui-se no tratamento dos dados coletados de modo a apresentá-los sob a forma de médias, desvio-padrão, máximos, mínimos, histogramas de distribuição, gráficos de tempo, etc.
- c) **Análise de dados:** Desta fase resultam os verdadeiros benefícios de técnica e nela assumem fundamental importância a criatividade, a experiência, e o bom senso do analista responsável pelo monitoramento. É feita *a posteriori* na maioria dos casos.

As técnicas de monitoramento podem ser implementadas de duas maneiras: por software ou por hardware.

No primeiro caso, são utilizados programas executados concorrentemente com os programas de usuários e que perfazem a coleta de dados gravando-os. Este caso é denominado **Monitoramento por software** e pode ser implementado por contabilizadores ou por monitores de software.

No segundo caso, é utilizado um dispositivo de hardware que, ligado a vários pontos do sistema por conexões elétricas, detecta e grava sinais indicadores dos níveis de atividade dos diversos recursos do sistema. A esse tipo de atividade dá-se o nome de **Monitoramento por hardware**, sendo os dispositivos conhecidos como monitores de hardware.