

Avaliação de Desempenho de Sistemas Computacionais - ADS

Prof. Dr. Paulo Antonio Leal Rego
paulo@dc.ufc.br

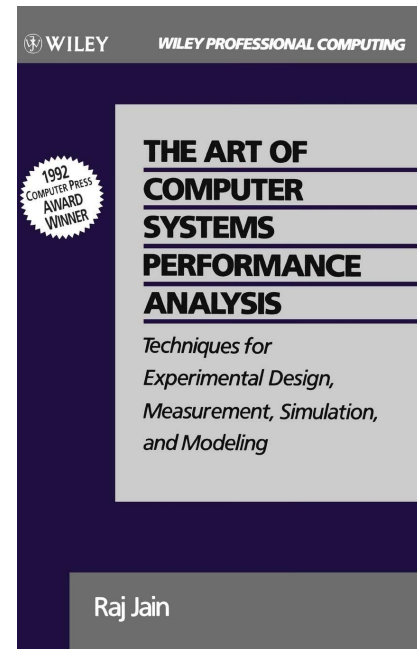


Universidade Federal do Ceará

Agenda

1. Introdução
2. Conceitos Básicos de ADS
3. Erros Comuns em uma ADS
4. Abordagem Sistemática para ADS
5. Seleção de Técnicas e Métricas

Capítulos 1, 2 e 3



Introdução

Introdução

- O que é desempenho?



Introdução

- O que é desempenho?



Ação de cumprir um trabalho, obrigação, ...

Introdução

- A busca por "Bom Desempenho" e "Baixo Custo"



Introdução

- Cost-performance trade-off

Samsung unveils the Galaxy S6

Comparison of select flagship smartphones by major manufacturers



	Samsung Galaxy S6 Edge	Samsung Galaxy S6	HTC M9	Google/Motorola Nexus 6	Apple iPhone 6 Plus	Apple iPhone 6	Motorola MotoX	Sony Xperia Z3	LG G3
Display size (diagonal)	5.1 in	5.1 in	5.0 in	6.0 in	5.5 in	4.7 in	5.2 in	5.2 in	5.5 in
Resolution (pixels)	577ppi	577ppi	441ppi	493ppi	401ppi	326ppi	423ppi	424ppi	538ppi
Weight	132g	138g	157g	184g	172g	129g	144g	152g	149g
Size, in mm (H x W x D)	142.1 x 70.1 x 7.0	143.4x 70.5 x 6.8	144.6 x 69.7 x 9.6	159.3 x 83 x 10.1	158 x 77.8 x 7.1	138 x 67 x 6.9	140.8 x 72.4 x 9.9	146 x 72 x 7.3 (mm)	146.3 x 74.6 x 8.9
Operating System	Android 5	Android 5	Android 5	Android 5	iOS 8	iOS 8	Android 5	Android 4.4.4	Android 4.4.2
RAM	3GB	3GB	3GB	3GB	1GB	1GB	2GB	3GB	2/3GB
Processor	2.1/1.5GHz, octa-core	2.1/1.5GHz, octa-core	2.1/1.5GHz, octa-core	2.7GHz quad-core	1.4GHz, dual-core	1.4GHz, dual-core	2.5GHz quad-core	2.5GHz quad-core	2.5 GHz quad-core
Internal storage	32/64/128GB	32/64/128GB	32GB	32/64GB	16/64/128GB	16/64/128GB	16/32GB	16/32GB	16/32GB
Storage expansion	none	none	Up to 128GB, MicroSD	none	none	none	none	Up to 128GB, MicroSD	Up to 128GB, MicroSD
Camera (rear/front)	16MP/5.0MP	16MP/5.0MP	20.7MP/4.0MP	13MP/2MP	8MP/1.2MP	8MP/1.2MP	13MP/2MP	20.7MP/2.2MP	13MP/ 2.1 MP
Video recording	2160p at 30fps	2160p at 30fps	1080p at 30fps	2160 at 30fps	1080p at 60 fps	1080p at 60 fps	2160p at 30fps	2160p at 30fps	2160p at 30fps
Battery life/ capacity	2600 mAh	2550 mAh	2840 mAh	3220 mAh	2915 mAh	1810 mAh	2300 mAh	3100 mAh	3000 mAh
NFC	yes	yes	yes	yes	yes	yes	yes	yes	yes
Wireless charging	yes	yes	no	yes	no	no	no	no	yes
Water/dust resistant	no	no	no	yes	no	no	yes	yes	no
Launch date	April 2015	April 2015	March 2015	October 2014	September 2014	September 2014	October 2014	September 2014	May 2014

Source: Companies. *Not officially.

C. Inton, 02/03/2015

REUTERS

Introdução

- O que vocês analisam ao escolher
 - Remédio?
 - Empresa de streaming de vídeo?
 - Provedor de Internet?
 - Tênis?
 - Roupas?
 - Notebook?
 - Smartphone?
 - Carro?
 - Casa?



Introdução

- Comparar "coisas"

O que é mais rápido?



X



Introdução

- O que é avaliação de desempenho?
 - De pessoas
 - De sistemas

Introdução

- O que é avaliação de desempenho?
 - **De pessoas**
 - De sistemas



Refere-se a um mecanismo ou ferramenta que visa conhecer e medir o desempenho dos indivíduos na organização, estabelecendo uma comparação entre o desempenho esperado e o apresentado por esses indivíduos.

Introdução

- O que é avaliação de desempenho?
 - De pessoas
 - **De sistemas**



Refere-se a um conjunto de métodos que possibilita investigar o comportamento temporal de sistemas; conhecer e medir o desempenho de sistemas; estabelecer uma comparação entre diferentes sistemas e/ou configurações dos mesmos.

Introdução

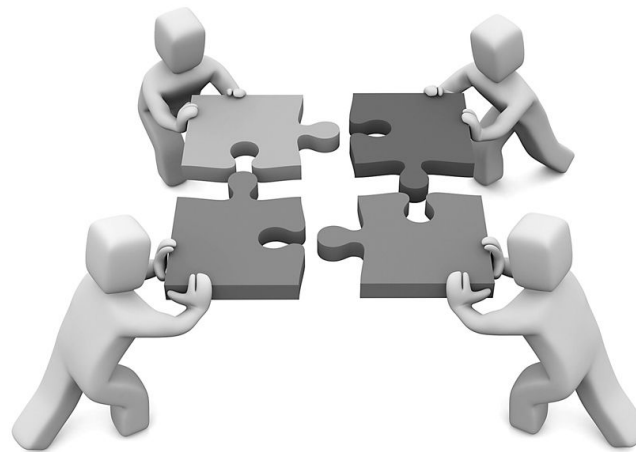
- Mesmo que não haja alternativas e tenhamos apenas um sistema, a avaliação de desempenho ajuda a determinar quão bem ele está desempenhando seu papel e se alguma melhoria precisa ser feita.



Conceitos Básicos de ADS

Conceitos Básicos de ADS

- Sistema
 - Coleção de entidades que interagem entre si e das quais podemos determinar relações e comportamentos.



Conceitos Básicos de ADS

- Métrica
 - Critério para comparar o desempenho de um sistema.
 - Ex: tempo de execução, taxa de transferência, delay, taxa de erro, disponibilidade, etc.



Conceitos Básicos de ADS

- Carga de Trabalho (Workload)
 - Conjunto de requisições feitas a um sistema.
 - Ex: lista de instruções a serem executadas em uma CPU, requisições a serem executadas em um serviço Web, método de um aplicativo a ser executado num smartphone.
 - Pode ser **real** ou **sintética**.

Conceitos Básicos de ADS

- Parâmetros
 - São configurações/informações.
 - Do sistema:
 - Ex: modelo da CPU, quantidade de memória RAM, tipo de conexão à Internet.
 - Da carga de trabalho:
 - Ex: tipo de requisição, quantidade de requisições.

Conceitos Básicos de ADS

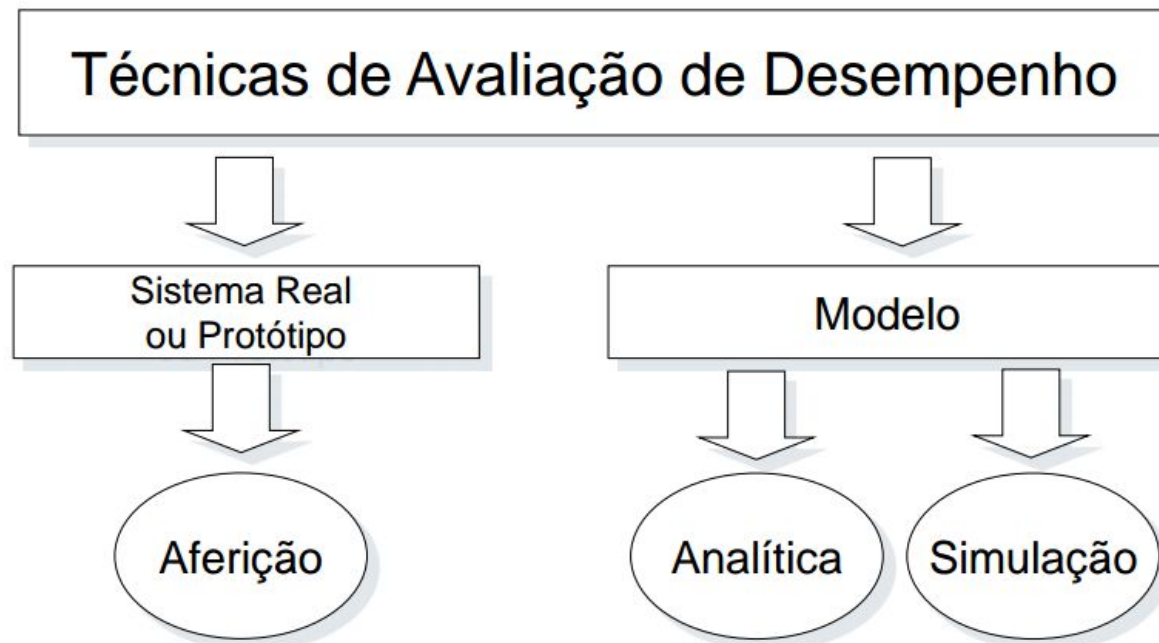
- Parâmetros podem ou não variar durante a ADS
- Fatores
 - São parâmetros que variam
 - Seus valores são chamados de **Níveis**
 - Ex: Web Service
 - Fator: tipo de requisição
 - Níveis: GET, POST, DELETE
 - Ex: Método que processa uma imagem
 - Fator: tamanho da imagem
 - Níveis: 500kB, 1MB, 2MB

Conceitos Básicos de ADS

- Exercício:
 - Quais métricas de desempenho devem ser usadas para comparar o desempenho dos seguintes sistemas?
 - Duas unidades de disco
 - Dois sistemas de gerenciamento de banco de dados
 - Dois algoritmos de retransmissão de pacotes

Conceitos Básicos de ADS

- Técnica de Avaliação



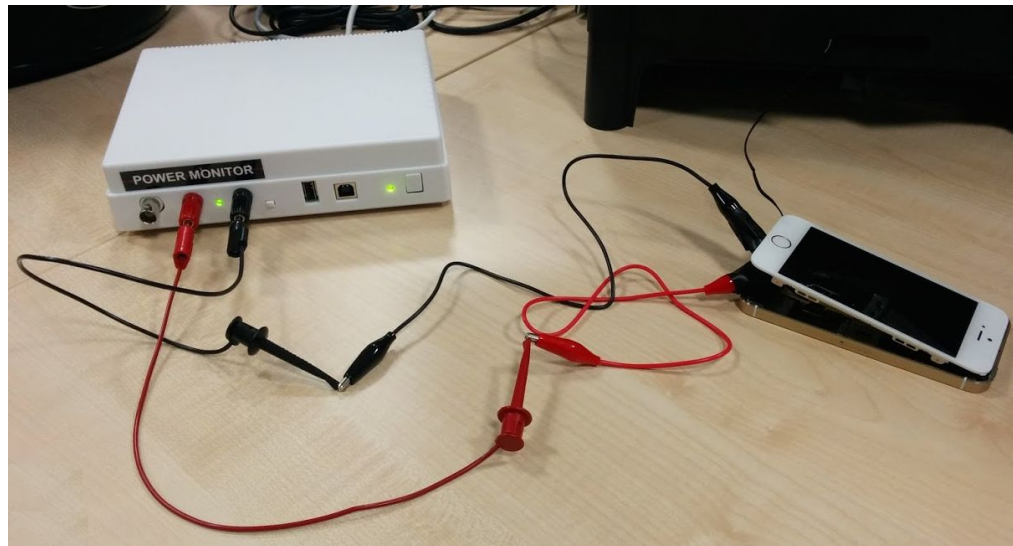
Conceitos Básicos de ADS

- Técnica de Avaliação
 - Aferição (medição, experimentação direta)
 - Provê resultados mais precisos.
 - Depende do sistema real ou protótipo.
 - Consequências: custo, tempo, falta de flexibilidade para verificar diferentes alternativas.
 - Benchmarks.
 - Monitor.

Conceitos Básicos de ADS

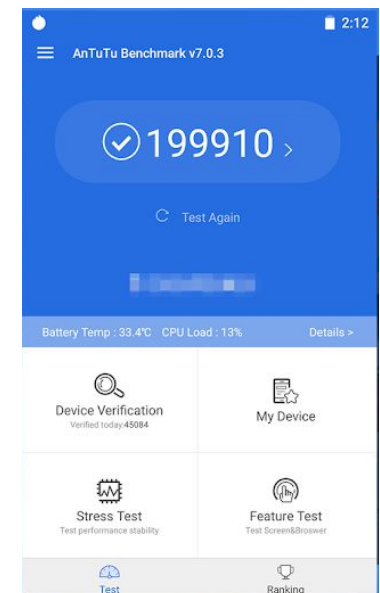
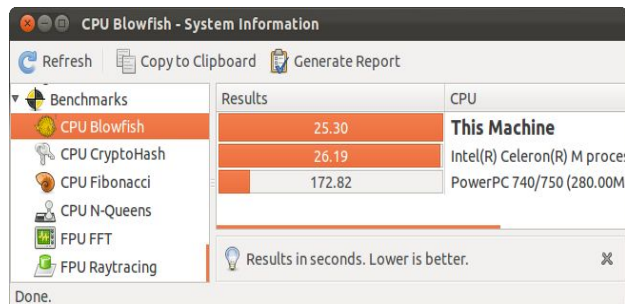
- Técnica de Avaliação
 - Aferição (medição, experimentação direta)
 - Exemplo de Monitor.

Monsoon Power Monitor



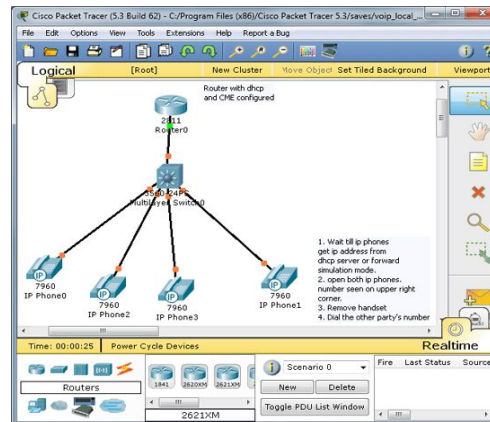
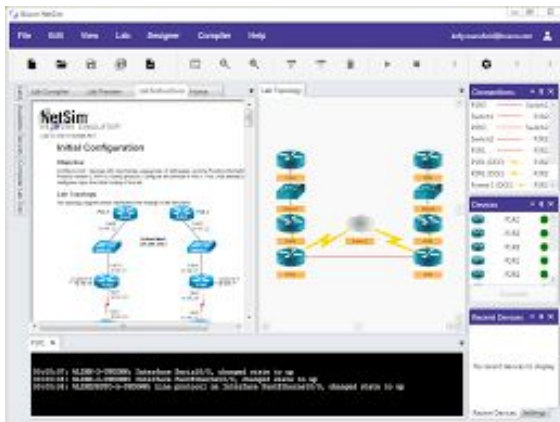
Conceitos Básicos de ADS

- Técnica de Avaliação
 - Aferição (medição, experimentação direta)
 - Benchmark: ato de executar um programa ou conjunto de operações para avaliar o desempenho de um sistema.



Conceitos Básicos de ADS

- Técnica de Avaliação
 - Simulação
 - Depende de um simulador: aparelho ou software capaz de reproduzir e simular o comportamento de algum sistema.

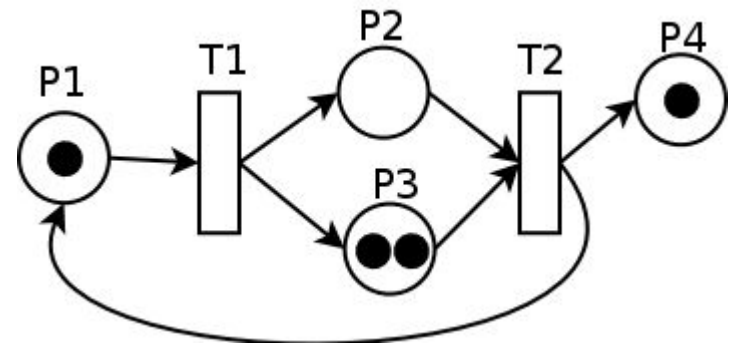


Conceitos Básicos de ADS

- Técnica de Avaliação
 - Simulação
 - Não depende do sistema real.
 - Baixo custo.
 - Precisão moderada.
 - Alta flexibilidade.
 - Muda parâmetros e já vê os resultados.
 - "Fácil" de implementar.

Conceitos Básicos de ADS

- Técnica de Avaliação
 - Modelagem Analítica
 - Descrição matemática de sistemas.
 - Não depende do sistema real.
 - Baixo custo.
 - Tende a ser menos preciso que as outras técnicas.
 - Alta flexibilidade.



Conceitos Básicos de ADS

- Objetivos da Análise de Desempenho
 - Comparar alternativas.
 - Determinar o impacto de um parâmetro do sistema.
 - Otimizar um sistema.
 - Caracterizar a carga do sistema.
 - Predição do desempenho de cargas futuras.

Conceitos Básicos de ADS

- Raj Jain considera que fazer avaliação de desempenho de sistemas é uma arte:
 - A avaliação bem-sucedida não pode ser produzida mecanicamente.
 - Toda avaliação requer um conhecimento íntimo do sistema que está sendo modelado e uma seleção cuidadosa da metodologia, carga de trabalho e ferramentas.

Erros Comuns em uma ADS

Erros Comuns em uma ADS

- Falta de Metas/Objetivos
 - Métricas, cargas de trabalho e metodologia dependem do objetivo.
 - Definir os objetivos antes de escrever a primeira equação do modelo analítico ou primeira linha de código de uma simulação.

Erros Comuns em uma ADS

- Objetivos Tendenciosos
 - O objetivo é "mostrar que o NOSSO sistema é melhor do que os OUTROS"
 - O problema torna-se o de encontrar as métricas e carga de trabalho "corretas" para comparar os sistemas e mostrar o que se deseja

Erros Comuns em uma ADS

- Objetivos Tendenciosos
 - "Mesmo" desempenho.

Métrica: Transações por Segundo

	Carga de Trabalho A	Carga de Trabalho B	Média
Sistema A	10	20	15
Sistema B	20	10	15

Erros Comuns em uma ADS

- Objetivos Tendenciosos
 - Usando B como base: Sistema A é "melhor".

Métrica: Transações por Segundo com base em B

	Carga de Trabalho A	Carga de Trabalho B	Média
Sistema A	0,5 (10/20)	2 (20/10)	1,25
Sistema B	1 (20/20)	1 (10/10)	1

Erros Comuns em uma ADS

- Objetivos Tendenciosos
 - Usando A como base: Sistema B é "melhor".

Métrica: Transações por Segundo com base em A

	Carga de Trabalho A	Carga de Trabalho B	Média
Sistema A	1 (10/10)	1 (20/20)	1
Sistema B	2 (20/10)	0,5 (10/20)	1,25

Essa técnica é conhecida como the Ratio Game

Erros Comuns em uma ADS

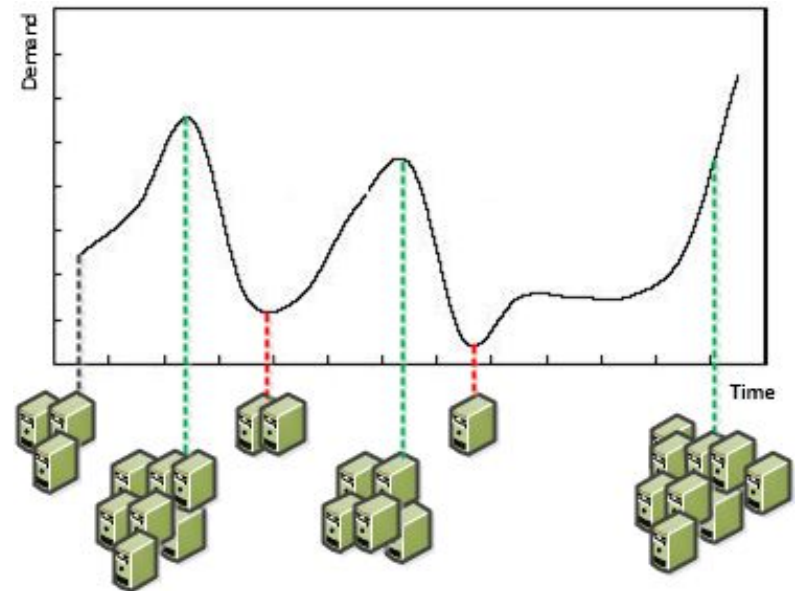
- Métricas de Desempenho Incorretas
 - A escolha de métricas de desempenho corretas depende dos serviços fornecidos pelo sistema ou subsistema a ser modelado.
 - Analistas escolhem métricas fáceis de calcular ou medir, em detrimento das relevantes.

Erros Comuns em uma ADS

- Métricas de Desempenho Incorretas
 - Exemplo: O desempenho de CPUs é comparado com base em seus throughputs, que são geralmente medidos em termos de MIPS (millions of instructions per second). No entanto, comparar MIPS de duas diferentes arquiteturas de CPU, tais como Reduced Instruction Set Computers (RISCs) e Complex Instruction Set Computers (CISCs), não faz sentido pois as instruções são diferentes.

Erros Comuns em uma ADS

- Carga de Trabalho não Representativa
 - Leva a conclusões imprecisas.
 - A carga deve representar a utilização real do sistema.



Erros Comuns em uma ADS

- Ignorar Parâmetros/Fatores Significativos
 - Leva a conclusões imprecisas.
 - É importante identificar os fatores que terão um impacto significativo sobre o desempenho.



Erros Comuns em uma ADS

- Escolha da Técnica Errada
 - Analistas tendem a escolher a técnica de sua preferência.



Erros Comuns em uma ADS

- Utilizar uma Abordagem não Sistemática
 - A seleção arbitrária de parâmetros, métricas, fatores e cargas de trabalho leva o analista a cometer erros durante a ADS.

Erros Comuns em uma ADS

- Design de experimentos (DoE) inadequado
 - DoE está relacionado ao número de aferições ou experimentos de simulação que precisam ser conduzidos e os valores dos parâmetros usados em cada experimento.
 - A seleção de valores impróprios pode resultar em perda de tempo e recursos do analista.

Erros Comuns em uma ADS

- Nível de detalhes não apropriado na modelagem
 - O nível de detalhes tem um grande impacto na formulação do problema. Evitem modelos que sejam muito amplos ou simplificados demais.
 - Modelo detalhado pode ser mais útil ao comparar pequenas variações de uma abordagem comum
 - Modelos mais simples permitem comparar alternativas que são muito diferentes de forma mais rápida.

Erros Comuns em uma ADS

- Falta de análise dos resultados
 - Problema comum, principalmente quando falta ao analista conhecimentos sobre *data analysis* ou conhecimento sobre o sistema em questão.
 - Apresentar gráficos e tabelas não é suficiente.
 - É essencial que os resultados sejam explicados



Erros Comuns em uma ADS

- Análise errada
 - Analistas geralmente cometem erros na aferição, simulação e modelagem analítica
 - Exemplo:
 - calculando a média de proporções
 - simulações muito curtas
 - falta de análise estatística

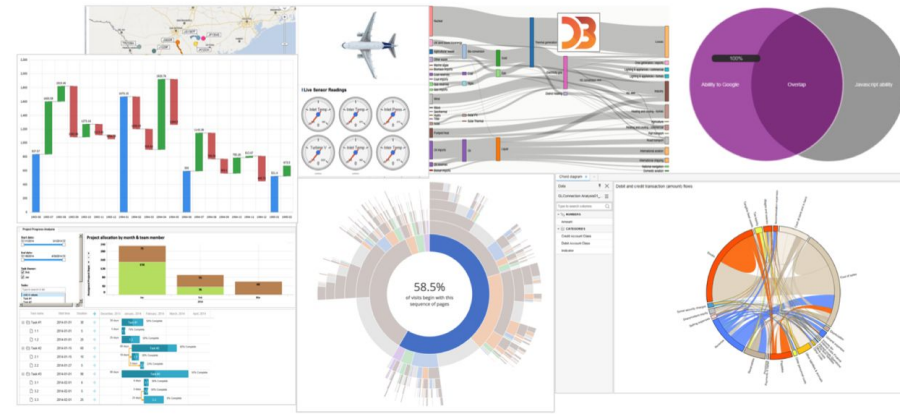


Erros Comuns em uma ADS

- Falta de análise de sensibilidade
 - Análise de sensibilidade é uma técnica que visa prever o resultado gerado por alterações nos parâmetros ou nas atividades em um processo.
 - O resultado de uma avaliação de desempenho é uma evidência, não verdade absoluta.
 - O resultado pode ser sensível à carga de trabalho e aos parâmetros do sistema.

Erros Comuns em uma ADS

- Apresentação não adequada dos resultados
 - Transmitir/vender os resultados da análise aos tomadores de decisão/revisores/leitores é de responsabilidade do analista/pesquisador.
 - Requer uso prudente/adequado de palavras, figuras e gráficos para explicar os resultados e as análises.



Erros Comuns em uma ADS

- Omissão de suposições e limitações
 - Tal omissão pode levar usuários a aplicarem a mesma análise em outro contexto, onde as suposições não são válidas.
 - É importante listar suposições e premissas no início do relatório, mas também apresentar as limitações no final do mesmo.
 - Sempre que possível, apresentar conclusões sobre os ambientes aos quais a análise não se aplica.

Abordagem Sistemática para ADS

Abordagem Sistemática para ADS

- Segundo Raj Jain, o processo de análise de desempenho é uma arte
 - Não deve ser realizado de forma mecânica.
- Passos comuns ajudam a evitar os erros durante a ADS
 - Abordagem sistemática



Abordagem Sistemática para ADS

1. Estabelecer Objetivos e Definir o Sistema
 - Primeiro passo em qualquer projeto de ADS.

2. Listar os Serviços e Saídas
 - Todo sistema oferece um serviço.
 - Quando um usuário solicita um serviço há uma série de possíveis resultados (uns desejáveis outros não).

Abordagem Sistemática para ADS

3. Selecionar as Métricas

- Selecionar o(s) critério(s) usado(s) para análise e comparação.
- Métricas são geralmente relacionadas a:
 - velocidade;
 - precisão; e
 - aos serviços disponíveis.

Abordagem Sistemática para ADS

4. Listar os Parâmetros

- Todos os parâmetros devem ser listados a fim de identificar aqueles que afetam o sistema.

5. Selecionar os Fatores e os Níveis

- Quanto mais fatores e níveis, mais tempo vai levar os experimentos.

Abordagem Sistemática para ADS

6. Selecionar a Técnica de Avaliação

- Depende do tempo, recursos disponíveis e nível desejado de precisão.

7. Selecionar a Carga de Trabalho

- Depende da técnica.
- Pode-se utilizar traces, benchmarks, etc.
- Deve ser representativa.

Abordagem Sistemática para ADS

8. Planejar e Executar o Experimento

- Definir experimentos que ofereçam máxima informação e mínimo esforço.
- Estudar o efeito relativo dos fatores.

9. Analisar e Interpretar Dados

- Considerar variabilidade dos resultados.
- Escolher técnicas estatísticas apropriadas.
- Resultados são evidências, não conclusões.

Abordagem Sistemática para ADS

10. Apresentar os Resultados

- Membros da equipe utilizam os resultados para tomada de decisão.
- Usar gráficos adequados.



Considerações Finais

- Importância da ADS
- Erros podem acontecer durante o processo de ADS, principalmente devido a **negligências simples, equívocos e falta de conhecimento** sobre técnicas de avaliação de desempenho
- Abordagem sistemática ajuda a evitar os erros
- Nas próximas aulas...

Exercícios

- E1) Selecione um artigo que apresenta resultados de uma avaliação de desempenho e faça uma lista dos pontos fortes e pontos fracos do estudo. O que você faria diferente?

Exercícios

- E2) Escolha um sistema para fazer análise de desempenho e descreva brevemente:
 - Os serviços providos pelo sistema.
 - As métricas de desempenho
 - Os parâmetros do sistema
 - A carga de trabalho e seus parâmetros
 - Os fatores e seus níveis
 - A técnica de avaliação a ser utilizada
- Justifique suas escolhas.

Seleção de Técnicas de Avaliação e Métricas

Seleção da técnica de avaliação

- A fase do ciclo de vida em que o sistema está é a chave para a escolha!
 - A aferição só é possível quando já existir algo semelhante ao sistema proposto, como ao projetar uma versão aprimorada de um produto.
 - Se for um novo conceito, modelagem analítica e simulação são as únicas técnicas possíveis.

Seleção da técnica de avaliação

- Tempo necessário para fazer a avaliação
 - A modelagem analítica é a opção mais rápida.
 - A simulação demoram muito.
 - Geralmente:
 - modelagem analítica < aferição < simulação
 - A lei de Murphy afeta a aferição com mais frequência do que outras técnicas. Se alguma coisa pode dar errado, vai dar.
 - Tempo necessário para a aferição é o que mais varia entre as três técnicas.

Seleção da técnica de avaliação

- Critérios para a seleção

Criterion	Analytical Modeling	Simulation	Measurement
1. Stage	Any	Any	Postprototype
2. Time required	Small	Medium	Varies
3. Tools	Analysts	Computer languages	Instrumentation
4. Accuracy ^a	Low	Moderate	Varies
5. Parameters trade-off	Easy	Moderate	Difficult
6. Cost	Small	Medium	High
7. Saleability	Low	Medium	High

Seleção da técnica de avaliação

- Cada vez mais, múltiplas técnicas têm sido usadas para validar os resultados. Regras de validação:
 - Não confie nos resultados de uma simulação até que tenham sido validados por modelagem analítica ou aferição.
 - Não confie nos resultados de um modelo analítico até que tenham sido validados por um simulação ou aferição.
 - Não confie nos resultados de uma aferição até que tenham sido validados por simulação ou modelagem analítica.

Seleção da técnica de avaliação

- É importante mencionar que:
 - Aferição é tão suscetível a erros/bugs quanto as outras duas técnicas.
 - Duas ou mais técnicas também podem ser utilizadas de forma sequencial.
 - Por exemplo: usar um modelo analítico simples para encontrar um conjunto adequado de parâmetros do sistema; e depois usar simulação para fazer a avaliação de desempenho.
 - Isso reduz o número de simulações e uso de recursos.

Seleção de métricas de desempenho

- Para qualquer avaliação de desempenho, é essencial escolher as métricas a serem avaliadas.
- Para ajudar nessa escolha, pode-se preparar uma lista com os serviços fornecidos pelo sistema.

É o entendimento do sistema que vai permitir a definição das métricas de desempenho mais adequadas!

Seleção de métricas de desempenho

- Lista com os serviços fornecidos pelo sistema
- Para cada requisição feita a um serviço, podem ocorrer três coisas:
 - O serviço responder corretamente;
 - O serviço responder incorretamente; ou
 - O serviço não responder.

Seleção de métricas de desempenho

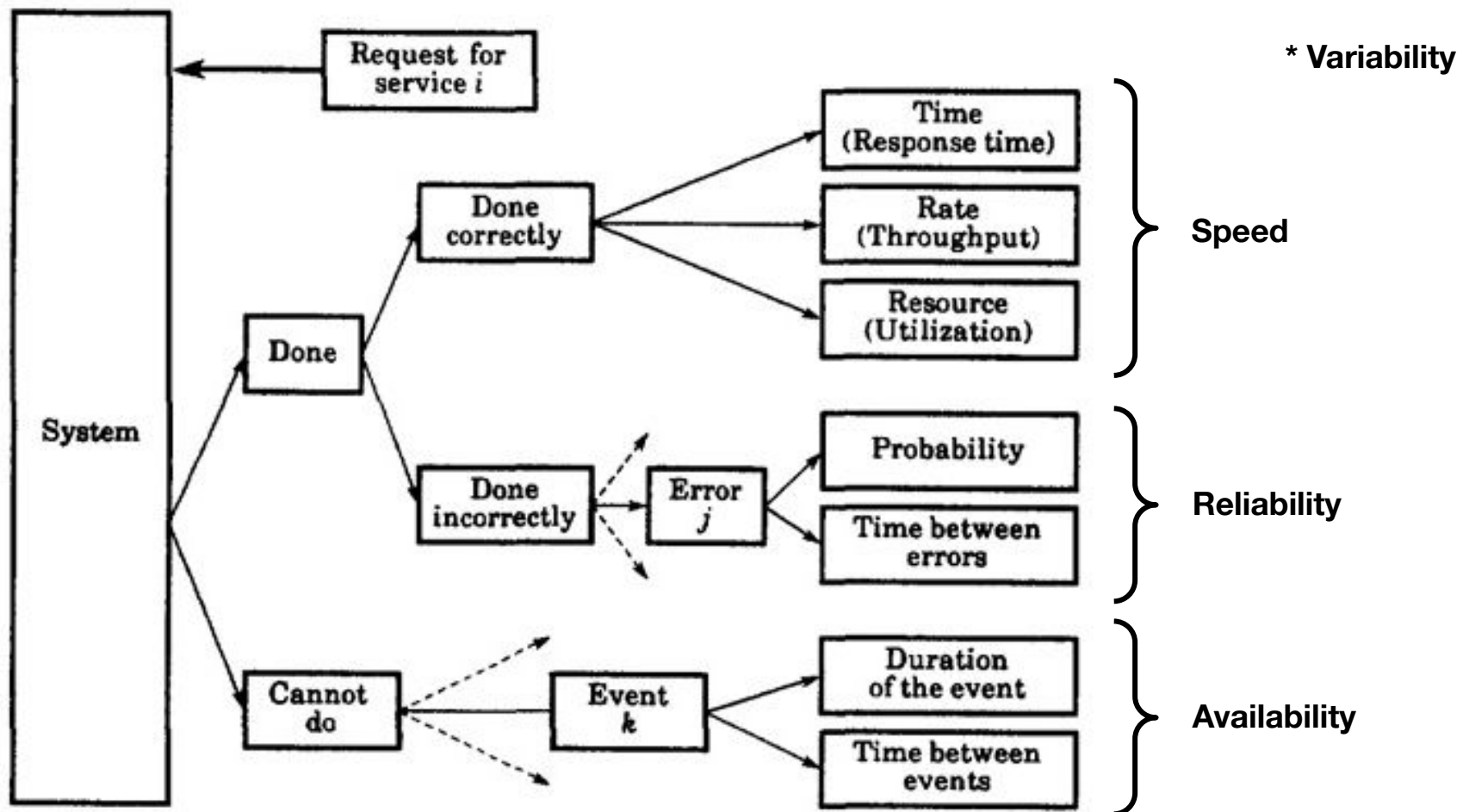
- Lista com os serviços fornecidos pelo sistema
- Para cada requisição feita a um serviço, podem ocorrer três coisas:
 - O serviço responder corretamente;
 - O serviço responder incorretamente; ou
 - O serviço não responder.

Faça uma análise dessas possíveis situações para:

(a) um sistema de encaminhamento de pacotes

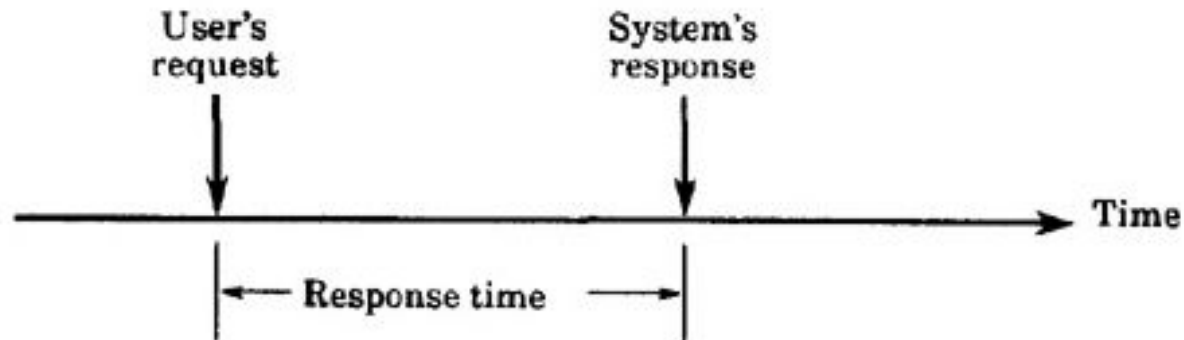
(b) um sistema de banco de dados

Seleção de métricas de desempenho



As métricas mais comuns

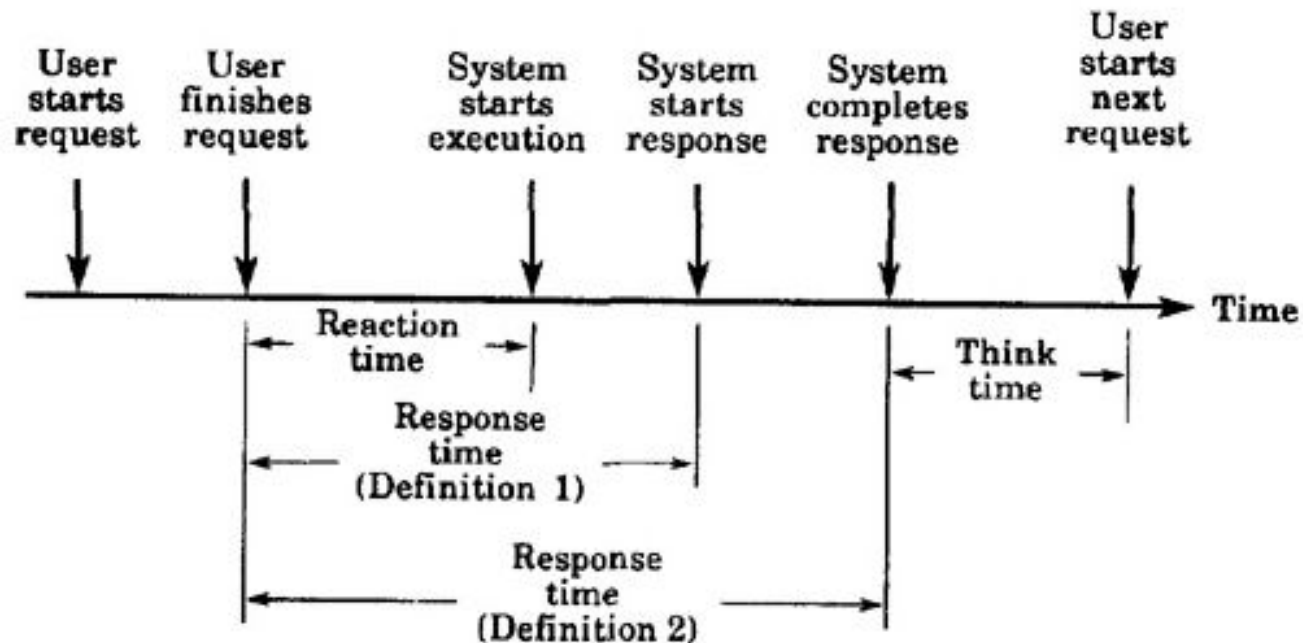
- Response time (tempo de resposta)



(a) Instantaneous request and response.

As métricas mais comuns

- Response time (tempo de resposta)



(b) Realistic request and response.

As métricas mais comuns

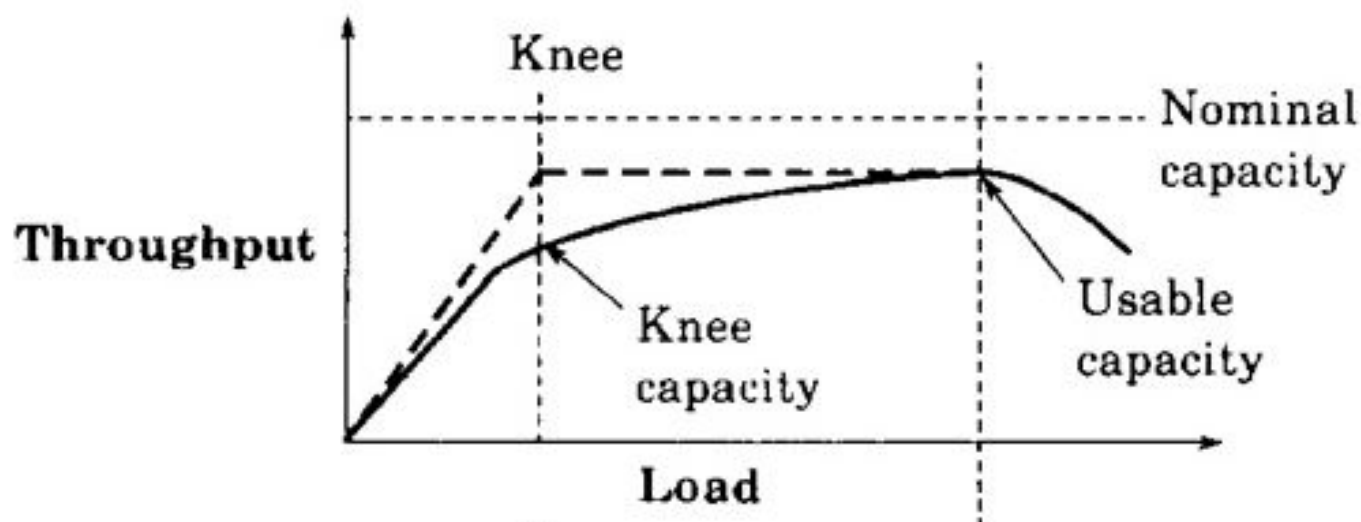
- Throughput (vazão) é definida como a taxa (requisições por unidade de tempo) em que o sistema consegue atender às requisições.
 - Para operações em lote (batch), é medida em jobs por unidade de tempo.
 - Para sistemas interativos, é medida em requisições por unidade de tempo.
 - Para CPUs, é medida em Millions of Instructions Per Second (MIPS) ou Millions of Floating-Point Operations Per Second (MFLOPS).

As métricas mais comuns

- Throughput (vazão) é definida como a taxa (requisições por unidade de tempo) em que o sistema consegue atender às requisições.
 - Para redes, é medida como packets per second (pps) ou bits per second (bps).
 - Para sistemas de processamento de transações, é medida como Transactions Per Second (TPS).

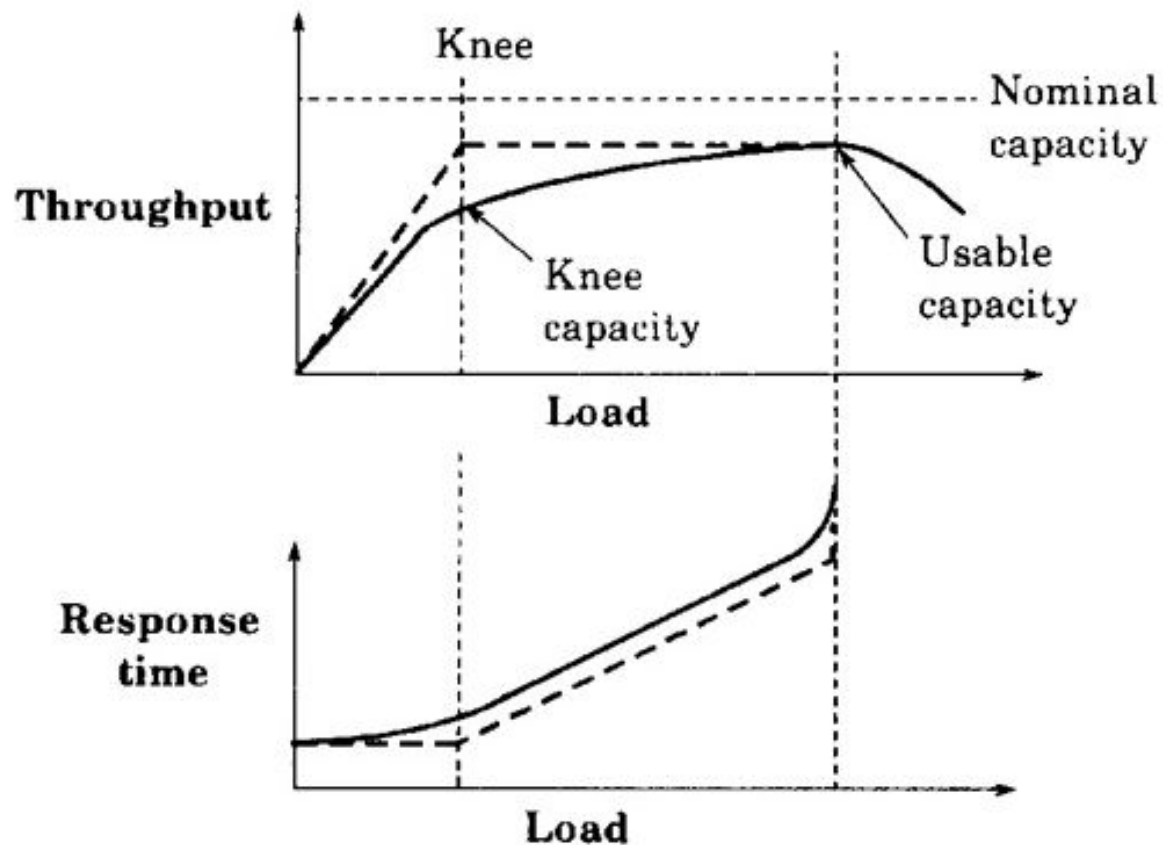
As métricas mais comuns

- Throughput (vazão) geralmente aumenta com a carga do sistema. O aumento contínuo da carga leva ao aumento da vazão, podendo inclusive começar a diminuir em certo ponto.



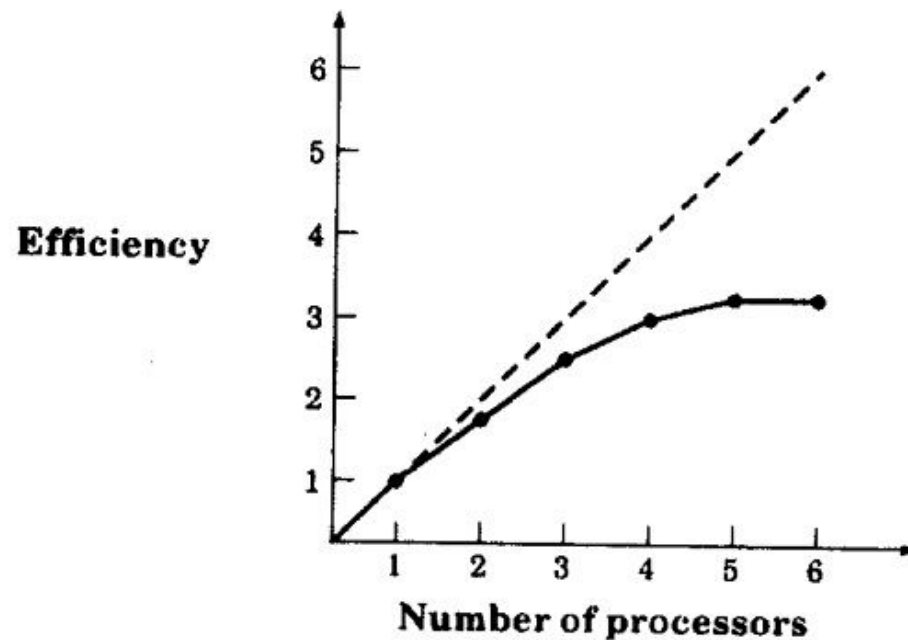
As métricas mais comuns

- Throughput (vazão)



As métricas mais comuns

- Efficiency (eficiência) é a razão entre a vazão máxima alcançável (usable capacity) e a capacidade nominal (nominal capacity)



As métricas mais comuns

- Utilization (utilização) de um recurso é medida como a fração do tempo em que o sistema está ocupado atendendo às requisições.
- Para alguns recursos, faz sentido falar em Busy Time e Idle Time num certo intervalo de tempo, como CPU.
- Já para outros recursos, como memória, apenas uma fração desta pode ser utilizada em certo intervalo de tempo.
 - Usa-se média, máximo ou mínimo de utilização.

As métricas mais comuns

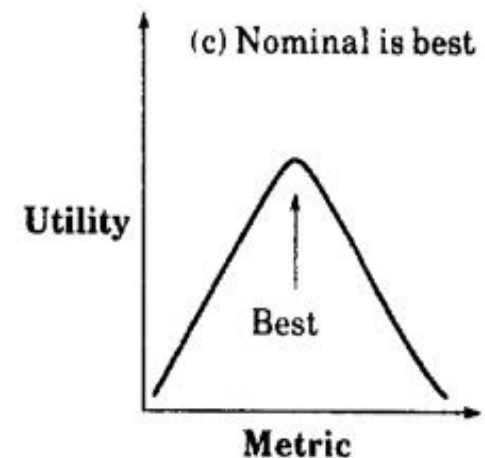
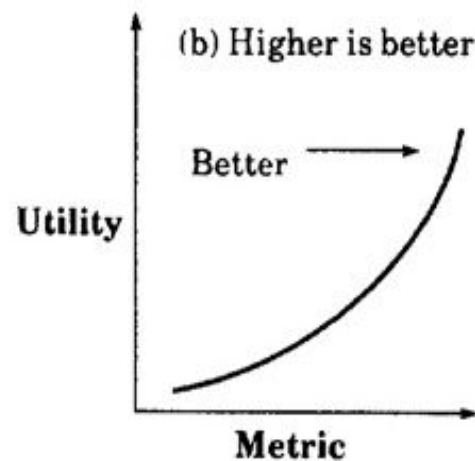
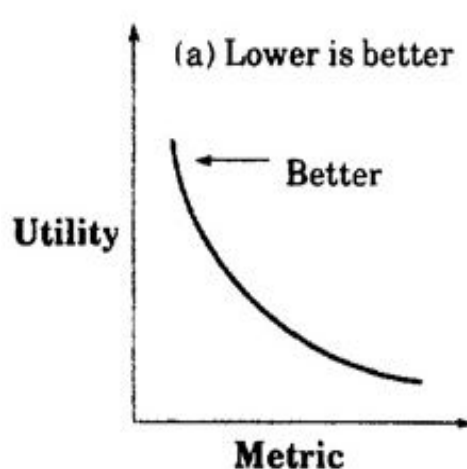
- Reliability (confiabilidade) de um sistema é usualmente medida pela probabilidade de erros ou pela média do tempo entre os erros.
- Availability (disponibilidade) de um sistema é definida pela fração de tempo em que o sistema está disponível para atender às requisições.
 - *Downtime* mede o tempo em que o sistema não está disponível. *Uptime* mede o tempo em que o sistema está disponível. A média do uptime é chamada de Mean Time To Failure (*MTTF*)

As métricas mais comuns

- A métrica cost/performance ratio (custo-benefício) é geralmente utilizada para comparar custos de dois ou mais sistemas.
 - O custo inclui hardware, licença de software, instalação e manutenção por um certo período de tempo.
 - Geralmente o desempenho (performance) é medido em termos de vazão (throughput).

As métricas mais comuns

- Utility Classification
 - Higher is Better
 - Lower is Better
 - Nominal is Best



Exercícios

- E1) Qual metodologia você escolheria?
 - Para selecionar 1 notebook
 - Para selecionar 1000 desktops para sua empresa
 - Para comparar 2 Manipuladores de Planilhas

Exercícios

- E2) Faça uma lista completa de métricas para comparar:
 - Dois computadores pessoais
 - Dois sistemas de banco de dados
 - Duas unidades de disco
 - Dois smartphones
 - Dois provedores de computação em nuvem

Bibliografia

- **Básica:**

- JAIN, Raj. The art of computer systems performance analysis: techniques for experimental design, measurement, simulation, and modeling . New York, NY: John Wiley & Sons, 1991. xxvii, 685 p. ISBN 9780471503361.
- KANT, K. Introduction to computer system performance evaluation. New York: McGraw-Hill, 1992.

Até a próxima aula!

Paulo Antonio Leal Rego
pauloalr@ufc.br



Universidade Federal do Ceará