

MAC6958

Tópicos Avançados em Ciência de Dados para Redes de Computadores

- **Prof. Daniel Macêdo Batista**
Prof. Roberto Hirata
- **DCC-IME-USP**
 - 12/3/2020

Análise de Desempenho

- **Importante para mostrar que uma nova proposta avança o estado da arte**
 - Exemplo 1: Algoritmo executa mais rápido do que os outros e devolve um resultado tão bom quanto
 - Exemplo 2: Mecanismo consome menos energia do que os outros e devolve um resultado tão bom quanto
 - Exemplo 3: Sistema consegue ser implantado mais rápido do que os outros
- **É preciso realizar análise de desempenho para mensurar as melhoras dos algoritmos/mecanismos/sistemas/etc...**

Análise de Desempenho

- **Não existe um único modo de realizar análise de desempenho:**
 - Modelagem analítica
 - Simulação (com emulação sendo um caso específico de simulação)
 - Medição
- **Medição só é possível quando algo similar ao proposto já existe**
- **Simulação e modelagem analítica são as opções quando um novo conceito está sendo proposto**
- **Ideal é realizar pelo menos 2 dos modos. O ótimo seria realizar os 3. Os resultados tem que ser similares em todos os modos.**

• Modelagem analítica

- Método baseado em “mecanismos” matemáticos para descrever uma nova proposta e a partir daí concluir a respeito do seu comportamento
- Teoria das filas
 - Equações mostrando o comportamento das várias entidades do sistema a partir de modelos probabilísticos
- Programação linear
 - Restrições descrevendo o sistema e uma função objetivo buscando o valor ótimo de um parâmetro
- Útil quando se quer estimar o comportamento do sistema antes de cogitar implementá-lo em um simulador

• Simulação

- Método baseado na “imitação” da operação do mecanismo no mundo real focando nas características principais desse mecanismo
- Precisa de um modelo que represente esse mecanismo e a simulação representa a operação desse mecanismo com o passar do tempo
- Útil se a modelagem analítica é muito complexa

• Medição

- Método baseado no monitoramento do mecanismo implementado no mundo real (realista) sob uma dada carga de trabalho
- Definir a carga de trabalho é uma tarefa crucial
- Útil se a carga de trabalho real é difícil de ser modelada para a utilização dos dois métodos anteriores

Análise de Desempenho

Critério	Modelagem Analítica	Simulação	Medição
Estágio	Qualquer	Qualquer	Pós-protótipo
Tempo requerido	Pouco	Médio	Varia
Ferramentas	Analistas	Linguagens de programação	Instrumentação
Precisão	Baixa	Moderada	Varia
Custo	Baixo	Médio	Alto
“Convencimento do cliente”	Baixo	Médio	Alto
Análise de trade-offs	Fácil	Moderado	Difícil

Análise de Desempenho

- **Simulação tem sido muito utilizada em redes de computadores (embora Redes Definidas por Software tenham reduzido a justificativa para usar simulação)**
 - Não há um critério de decisão fácil para escolher como criar um modelo analítico
 - Criar a carga de entrada no modelo analítico pode ser tão difícil quanto o modelo em si, principalmente sem poder medir
 - Medir pode ser inviável principalmente pelo fato de ter que considerar o tempo no mundo real
 - Medir pode ser inviável se a proposta depende de um hardware que ainda não existe

Análise de Desempenho

- **Independente do modo escolhido, se há dependência de modelos probabilísticos é importante repetir experimentos**
 - Algum valor que resuma os experimentos deve ser usado (média por exemplo)
 - Algum valor que resuma a dispersão dos resultados deve ser usado o (desvio padrão ou intervalo de confiança por exemplo)

• Simulação

- Exige apenas a definição das principais características do mecanismo
- O sistema pode ser uma caixa se comportando como uma máquina de estados
- Múltiplas threads podem ser usadas para representar as várias partes do mecanismo se comunicando
- Útil se não puder esperar o andamento do tempo do mundo real para obter resultados

Análise de Desempenho

• Simulação (cuidados)

- Geradores de números aleatórios precisam estar testados em relação a uniformidade e independência
- Transientes iniciais precisam ser removidos antes dos cálculos das métricas
- Resultados surpreendentes precisam ser validados e não descartados
- Usar a linguagem de programação adequada (conhecimento do programador e funcionalidades para a simulação)

• Simulação (definições)

- Evento
- Modelos em tempo contínuo e em tempo discreto
- Modelos de estado contínuo e de estado discreto

Análise de Desempenho

• Simulação em redes de computadores

- Simulações baseadas em programas de computador
- Geralmente orientadas a trace (útil para reprodutibilidade)
- Geralmente baseada em simuladores existentes de código aberto que já foram validados
 - NS-2 (linguagem otcl)
 - NS-3 (linguagem C++)
 - O que se costuma fazer é implementar um módulo para alguma das versões do NS que implemente o que se pretende simular, sem se preocupar dessa forma com particularidades do simulador como gerador de números aleatórios ou com particularidades de redes não relacionadas como demais camadas da arquitetura Internet
- Se for usar seu próprio simulador, valide ele e compartilhe o código
- No geral, não há interesse em simuladores com GUI

• Validação

- Importante para confiar nos resultados do simulador
- Ideal sempre comparar com medição, mesmo que seja apenas a medição de 1 cenário e os outros apenas simulados
- Usar simuladores validados
- Divulgar todo o código para que passe pela análise de outras pessoas

- **Raj Jain. The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling. Wiley, 1991.**
- **Wikipedia, “Simulation”,**
<http://en.wikipedia.org/wiki/Simulation>