



CEFET/RJ



BENCHMARK TPC

Eduardo Ogasawara
eogasawara@ieee.org
<https://eic.cefet-rj.br/~eogasawara>

Sumário

- Benchmarks
- TPC
 - TPC-C
 - TPC-H
 - TPC-App
- Download da especificação

Benchmarks

- O que é um benchmark?
 - Carga usada no processo de avaliação de algum sistema
- Atributos desejáveis de um benchmark
 - Relevante
 - Entendível
 - Métricas bem definidas
 - Escalável
 - Alta cobertura
 - Bem aceitável

Benefícios dos Benchmarks TPC

- Definição do que deve ser avaliado
- Incentivo ao desenvolvimento
- Evidências comparativas
- Bons benchmarks guiam o avanço tecnológico

Conselho de desempenho do benchmark TPC

- Organização sem fins lucrativos
- Define regras e testes para avaliação de bancos de dados
- Composta por uma série de funcionários de empresas de informática
- Provê dados para teste de desempenho relevante para sistemas computacionais

Tipos de Benchmarks TPC

- TPC-C
 - Benchmark OLTP
- TPC-H
 - Benchmark OLAP
- TPC-App
 - Servidor de aplicação e Serviços Web

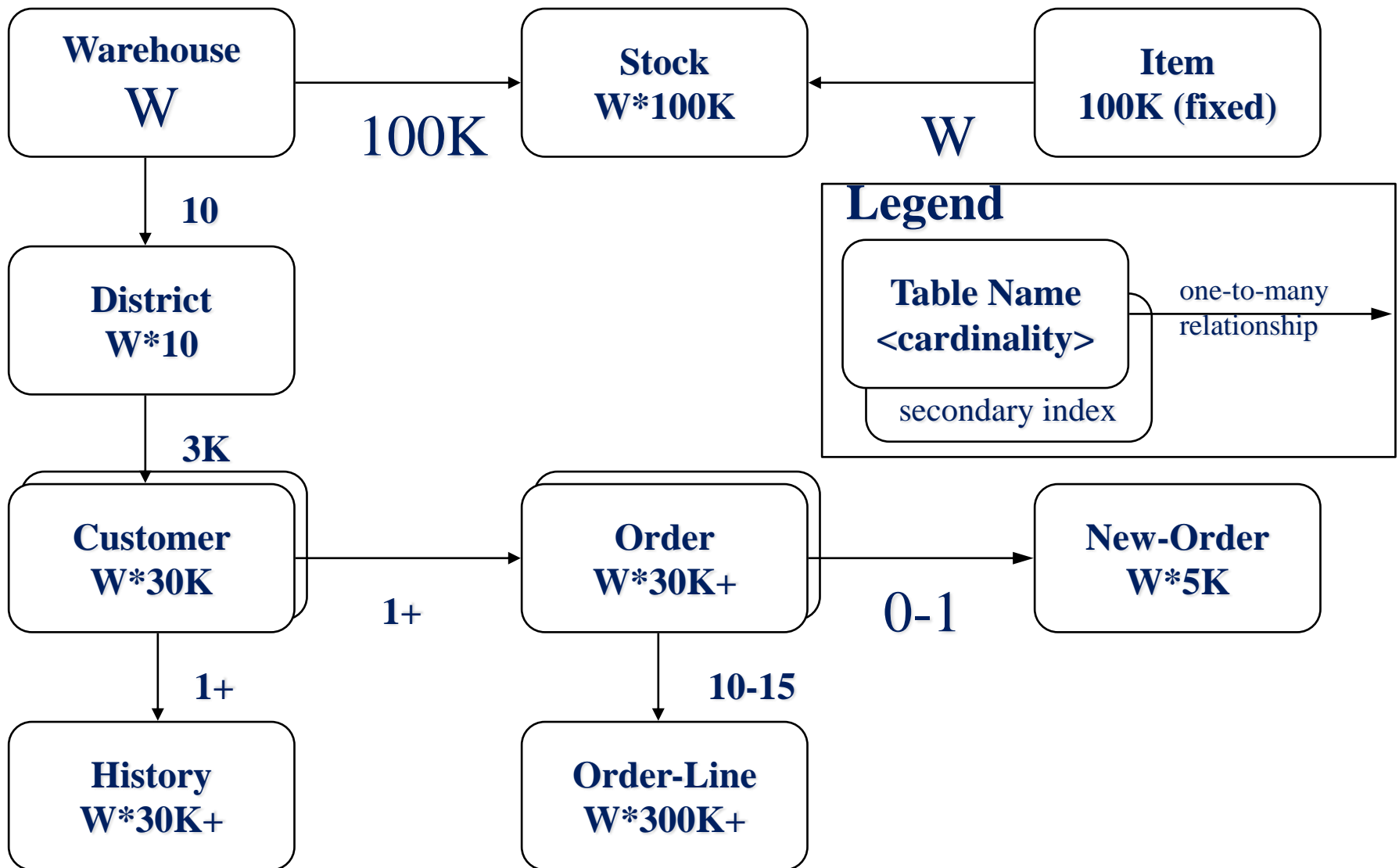
Benchmarks TPC-C (OLTP)

- Comparação de desempenho OLTP em diversos hardwares e softwares
- Simula um ambiente de computação completo
 - Quantidade grande usuários executam transações contra um banco de dados
 - Transações concorrentes de complexidade e tipos diferentes
- Banco de dados composto de nove tabelas e um grande número de registros
- Métricas
 - Transações por minuto

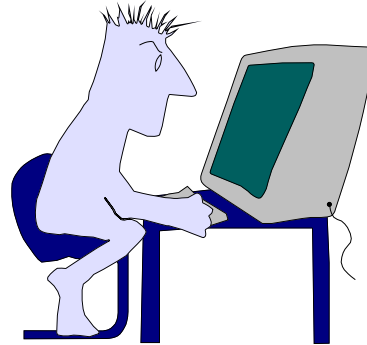
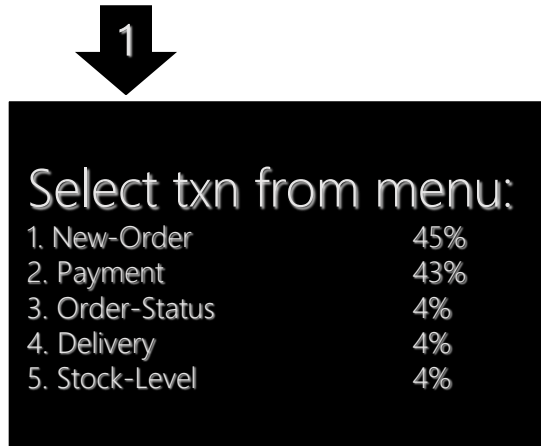
Transações do Benchmarks TPC-C

- New-order
 - Insere uma nova ordem de compra de um cliente
- Payment
 - Atualiza o saldo do cliente refletindo um pagamento
- Delivery
 - Atende ordens de compra
- Order-status
 - Recupera o status das ordens de compra mais recente dos clientes
- Stock-level
 - Monitora a quantidade de itens em estoque

Modelo do Benchmark TPC-C



Processo do Benchmark TPC-C



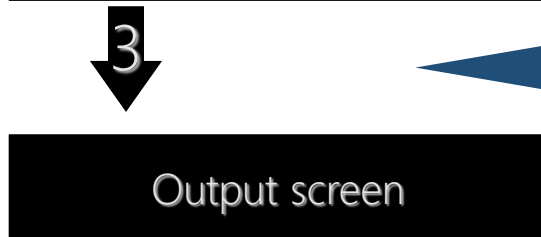
Cycle Time Decomposition
(typical values, in seconds,
for weighted average txn)



Measure menu Response Time

Menu = 0.3
Keying = 9.6

Keying time



Measure txn Response Time

Txn RT = 2.1

Think time

Think = 11.4

Average cycle time = 23.4

Go back to 1

Teste ACID do Benchmark TPC-C

- TPC-C procura demonstrar que as propriedades ACID estão sendo atendidas
- Atomicidade
 - Confirma que todas as mudanças de uma transação são confirmadas ou abortadas
- Consistência
- Isolamento
- Durabilidade
 - Demonstra a recuperação do sistema em:
 - Faltas de energia
 - Falta de memória
 - Problemas na mídia de armazenamento

Benchmark TPC-H (OLAP)

- Avalia o desempenho de banco de dados para tomadas de decisão
 - Consultas OLAP
- Consiste de um conjunto de consultas ad-hoc orientadas a negócio e modificações concorrentes de dados
- Consultas
 - Examinam grande volume de dados
 - Alto grau de complexidade
 - Fornecem respostas questões críticas de negócio

Características do benchmark TPC-H

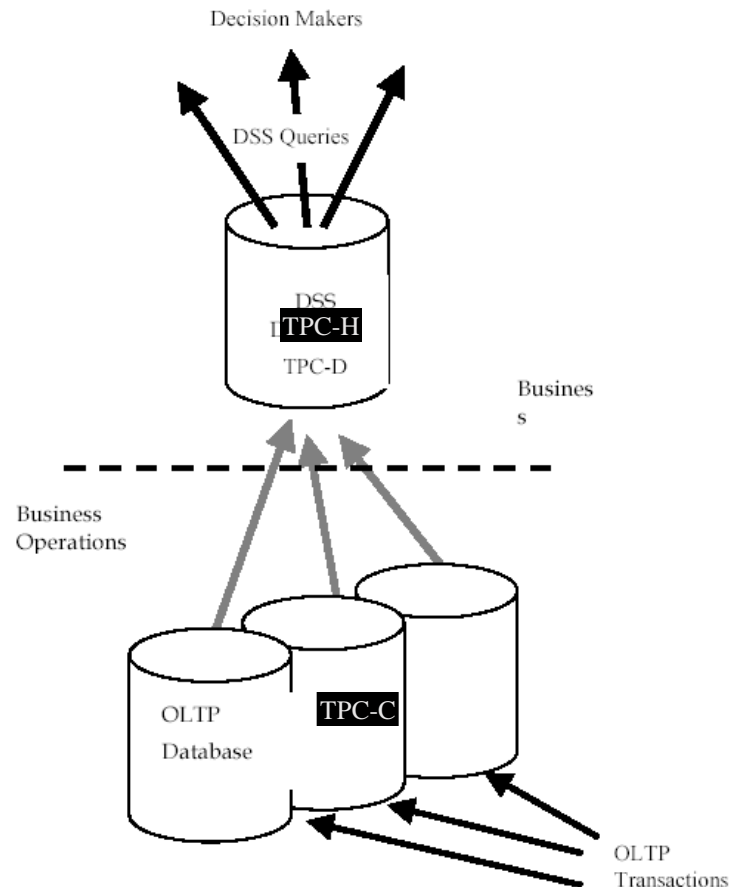
- Detalhes sobre as consultas
 - Alto grau de complexidade
 - Vários tipos de acesso aos dados
 - Natureza ad-hoc
 - Examinam uma grande porcentagem dos dados disponíveis
 - Diferentes umas das outras
 - Possuem parâmetros que mudam durante a execução da consulta

Tipos de consultas OLAP do benchmark TPC-H

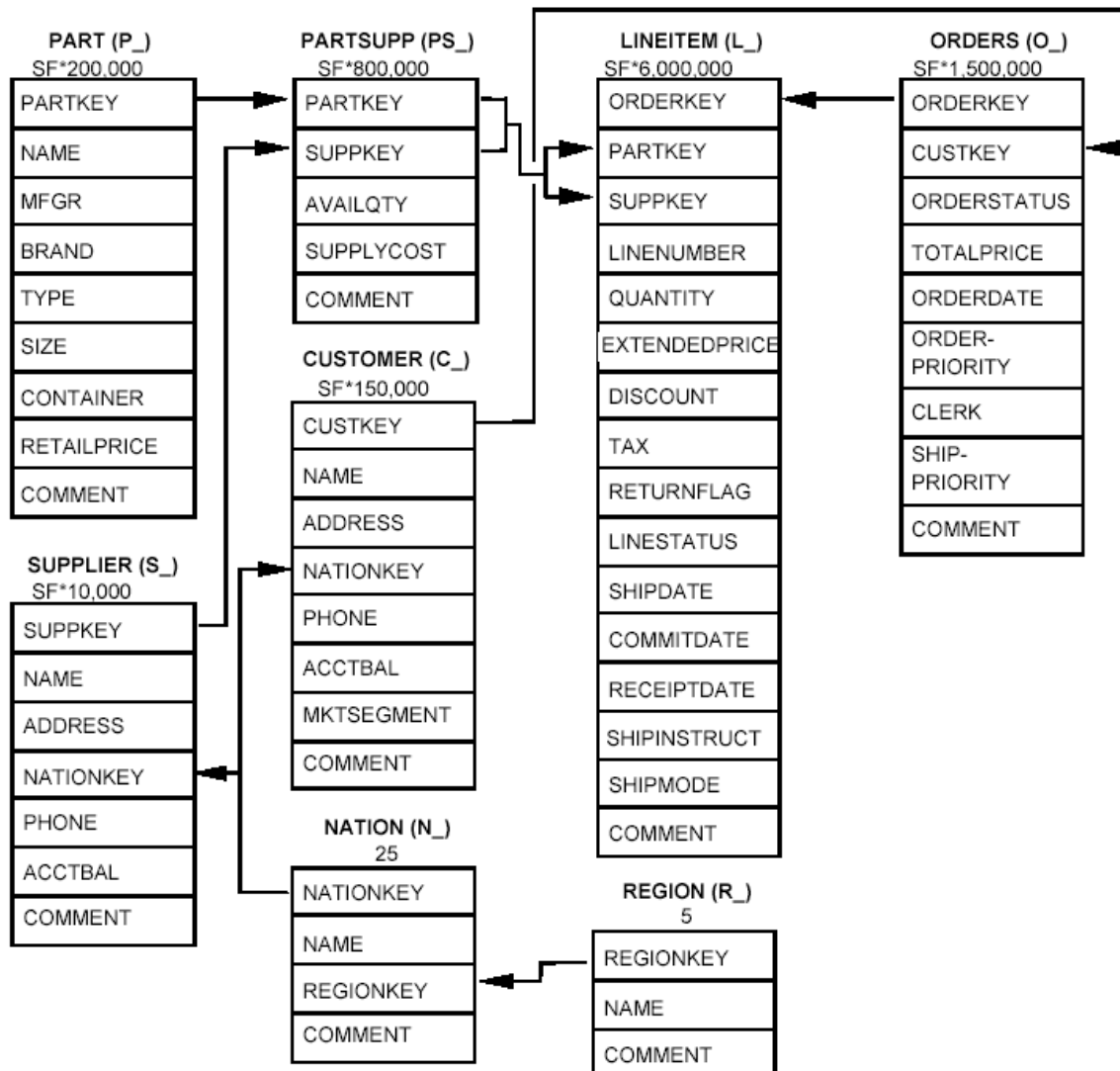
- Preço e promoção
- Gerenciamento de oferta e procura
- Gerenciamento de lucros e rendimentos
- Estudo de satisfação do cliente
- Gerenciamento de remessa

Cenário OLAP do benchmark TPC-H

- O processo de construção da base não é o objetivo do estudo, mas a base não pode usar nenhum mecanismo que atenda apenas este benchmark



Esquema do benchmark TPC-H



Execução do benchmark TPC-H

- Preparação da base
 - DBGen
 - Gerador da base de dados
- Preparação das consultas
 - QGen
 - Gerador das consultas
- Site do TPC-H
 - <http://www.tpc.org/tpch/default.asp>
- Link do DBGen
 - http://www.tpc.org/tpch/spec/tpch_2_8_0.zip

Benchmark TPC-H – Dicas para carga

- Criação do esquema: Deve-se manter a ordem das colunas para cada tabela indicada no TPCH.pdf
- Existem duas opções para carga:
 - Funções do banco de dados que permitam a carga direta de arquivo com separadores
 - Criar um programa (Java ou .Net) que leia o arquivo e faça a carga se conectando a base

Benchmark TPC-H – Carga via SGBD

- PostgreSQL
 - <http://www.postgresql.org/docs/8.2/interactive/sql-copy.html>
 - alter table supplier add stub char(1);
 - COPY supplier FROM '/data/tpch/supplier.tbl' USING DELIMITERS '|' WITH NULL AS 'NULL';
 - alter table supplier drop column stub;
- MySQL
 - load data infile "/var/lib/mysql-files/region.tbl" IGNORE INTO TABLE region fields terminated by '|';
- SQL Server - Bulk insert
 - <http://msdn.microsoft.com/en-us/library/ms188365.aspx>
 - BULK INSERT customer FROM '/data/tpch/customer.tbl' WITH (FIELDTERMINATOR = '|', ROWTERMINATOR = '\n')
- Oracle - Oracle Bulk Loader
 - <http://infolab.stanford.edu/~ullman/fcdb/oracle/or-load.html>
 - set NLS_LANG=AMERICAN_AMERICA.WE8MSWIN1252
 - sqlldr tpc control=customer.ctl log=customer.log bad=customer.bad

Referências

