

ЭТАЛОННЫЕ ТЕСТЫ ПРОИЗВОДИТЕЛЬНОСТИ СУБД

обзор и
тенденции



The TPC defines transaction processing and database benchmarks and delivers trusted results to the industry.

Home**Results**TPC-C
TPC-H
TPC-R
TPC-W**Benchmarks**TPC-C
- **Results**
- Description
- FAQTPC-H
- Results
- SpecificationTPC-R
- Results
- SpecificationTPC-W
- Results
- SpecificationObsolete
- TPC-A
- TPC-B
- TPC-D**Technical Articles****Related Links****Complete TPC-C Results List - Sorted by Database Vendor**

Print Version

Database	Hardware Vendor	System	tpmC	Price/tpmC	System Availability	Operating System	TP Monitor	Date Submitted
SymfoWARE Server Enterp. Ed. VLM 3.0	Fujitsu	PRIMEPOWER 2000 c/s w/32 Front-Ends	183770	56.16US \$	06/30/01	Sun Solaris 7	BEA Tuxedo 6.4 CFS	01/12/01
Oracle8i Ent. Edition 8.1.5	Fujitsu/ICL	GP7000F Model 2000 c/w	138735	74.32US \$	01/14/01	Sun Solaris 7	BEA Tuxedo 6.4 CFS	07/21/00
Oracle8 Enterprise Edition v8.1.7.1	HP	HP 9000 Superdome Enterprise Server	197024	66.27US \$	05/01/01	HP HP-UX 11i	BEA Tuxedo 6.4	01/18/01
Fujitsu/ICL SymfoWARE Server for UltraSPAR V11L20	Fujitsu/ICL	GRANPOWER 7000 Model 200 c/s	5738	69585.00Yen	11/12/97	Fujitsu/ICL UXP/DS Basic Software V20	TP-Base V20	05/12/97
Fujitsu/ICL SymfoWARE Server for UltraSPAR V11L20	Fujitsu/ICL	GRANPOWER 7000 Model 200 c/s	5738	69585.00Yen	11/12/97	Fujitsu/ICL UXP/DS Basic Software V20	TP-Base V20	05/12/97
Fujitsu/ICL SymfoWARE Server for VLM 2.0	Fujitsu/ICL	GP 7000 Model 600 (c/s)	20170	110.30US \$	01/21/99	Sun Solaris 2.6	BEA Tuxedo 4.2	07/31/98
Fujitsu/ICL SymfoWARE Server 1.1	Fujitsu/ICL	teamServer M830ic	17057	36.24US \$	03/01/99	Microsoft Windows NT Enterprise Edition 4.0	BEA Tuxedo 6.4 CFS	09/22/98
Fujitsu/ICL SymfoWARE Server 1.1	Fujitsu/ICL	GRANPOWER 5000 Model 680	17057	36.24US \$	03/01/99	Microsoft Windows NT Enterprise Edition 4.0	BEA Tuxedo 6.4 CFS	09/22/98
Fujitsu/ICL SymfoWARE Server for VLM 2.0	Fujitsu/ICL	GP7000F Model 600 c/s	33113	83.74US \$	12/23/99	Sun Solaris 2.6	BEA Tuxedo 6.4 CFS	06/24/99
Fujitsu/ICL SymfoWARE Server 1.1	Fujitsu/ICL	GRANPOWER 5000 Model 680	25440	28.57US \$	01/16/00	Microsoft Windows NT Enterprise Edition 4.0	BEA Tuxedo 6.4 CFS	07/19/99
Fujitsu/ICL SymfoWARE Server 1.1	Fujitsu/ICL	teamserver M890ic	25440	28.57US \$	01/16/00	Microsoft Windows NT Enterprise Edition 4.0	BEA Tuxedo 6.4 CFS	07/19/99
Fujitsu/ICL SymfoWARE Server for Workgroup 2.0	Sun	Enterprise 450	20123	27.70US \$	07/28/00	Sun Solaris 7	BEA Tuxedo 6.3	02/02/00
Fujitsu/ICL SymfoWARE Server 2.0.1	Sun	Enterprise 450	25375	21.35US \$	10/31/00	Sun Solaris 8	BEA Tuxedo 6.3	08/24/00
IBM DB2 for AS/400 V4R1	IBM	AS/400e Server Model s40-2261 c/s	25149	138.06US \$	08/29/97	IBM OS/400 V4R1	CICS for AS/400 V3R6	08/18/97
IBM DB2 for AS/400 V4R3	IBM	AS/400e Server Model S40 2208	43160	138.06US \$	08/29/97	IBM OS/400 V4R3	CICS for AS/400 V4R3	09/01/98
IBM DB2 for AS/400 V4R3	IBM	AS/400e Model 740 Feature 2070	43160	138.06US \$	08/29/97	IBM OS/400 V4R3	BEA Tuxedo 6.4	06/07/99
		Netfinity 8200					Microsoft COM+	07/02/99

296 результатов

Simfoware, Oracle DB 7-8, DB2/400 и UDB, Informix, MS SQL Server, Sybase ASE ...

Sybase ASE 11.5	SGI	Origin2000 (c/s)	9853	64.22US \$	02/12/98	IBM AIX 4.2.0	BEA Tuxedo 6.3	04/23/98
Sybase Adaptive Server Enterprise 11.5.1	SGI	Enterprise 4500	6040	121.98US \$	07/31/98	SGI IRIX 6.5	BEA Tuxedo 6.3	11/23/99
Sybase Adaptive Server Enterprise 11.9.3	Sun	Starfire Enterprise 10000	50268	49.88US \$	03/30/00	Sun Solaris 7	BEA Tuxedo 6.3	08/29/00
Sybase ASE 12.0.0.2	Sun	Digital AlphaServer 8400 5/300 c/s	156873	48.81US \$	02/28/01	Sun Solaris 7	BEA Tuxedo 6.3	12/21/95
Sybase SQL Server 11.0	Sybase	Digital AlphaServer 8400 5/300 c/s	11014	222.02US \$	03/01/96	Digital UNIX V3.2C	ITI Tuxedo 4.2	07/02/99

TPC-C, АПРЕЛЬ 2017

TPC

TPC-C - Top Ten Performance Results
Version 5 Results As of 24-Apr-2017 at 2:33 AM [GMT]

Note 1: The TPC believes it is not valid to compare prices or price/performance of results in different currencies.

All Results Clustered Results Non-Clustered Results Currency

Rank	Company	System	Performance (tpmC)	Price/tpmC	Watts/KtpmC	System Availability	Database	Operating System	TP Monitor	Date Submitted
1	 SAP	Dell PowerEdge T620	112,890	.19 USD	NR	11/25/14	SQL Anywhere 16	Microsoft Windows 2012 Standard x64	Microsoft COM+	11/25/14

the Watts/Ktpmc column indicates that no energy data was reported for that benchmark.

 ©2001 - 2017 TPC All rights reserved. Web-Design and Maintenance by [Hotea Solutions](#)

SQL
Anywhere

ПЛАН

Что случилось с TPC-C?

- Ранняя история TPC
- TPC-A/B, -C, -H, -E
- TPC: устаревшие и новые
- Особенности публикации

Бенчмарки новой волны

- MapReduce (Hadoop)
- Графовые
- Метауниверсальные, атомарные

«Карманные средства» для замера

Ввод-вывод

Прогон нагрузки и бенчмарки приложений

TP1

Середина
1970-х,
бенчмарк IBM

обработка
банковских
транзакций

Idefix: 100 tps

(1973, «банк с
1 тыс. отделений и
10 тыс.
операционистов»)

Пакетный режим

без сети

без времени на
реакцию оператора

Начало 1980-х:
фантастические
победные
реляции на
10 ktps

Рынок вышел на
миллиардные
обороты

... с самым
быстрорастущим
сегментом с
фокусом на OLTP

... но конечные
заказчики не
получали
и 1 ktps ...

ДЭВИД ДЕВИТТ: ВИСКОНСИНСКИЙ БЕНЧМАРК

Альтернативный, более
строго
детерминированный
бенчмарк

Напечатан на борьбу с
«бенчмарковыми
войнами», но разжёг их
ещё сильнее!

Лицензия на СУБД
запрещает публикацию
каких-либо тестов
производительности
СУБД

Из-за низких
результатов Oracle DB
появилась «оговорка
Девитта»
(DeWitt clause)

ДЖИМ ГРЕЙ: DEBITCREDIT

TANDEM COMPUTERS

A Measure of Transaction Processing Power

Anon Et Al

Technical Report 85.2
February 1985
PN87609

В пику висконсинскому бенчмарку, на базе предметной области из TP1

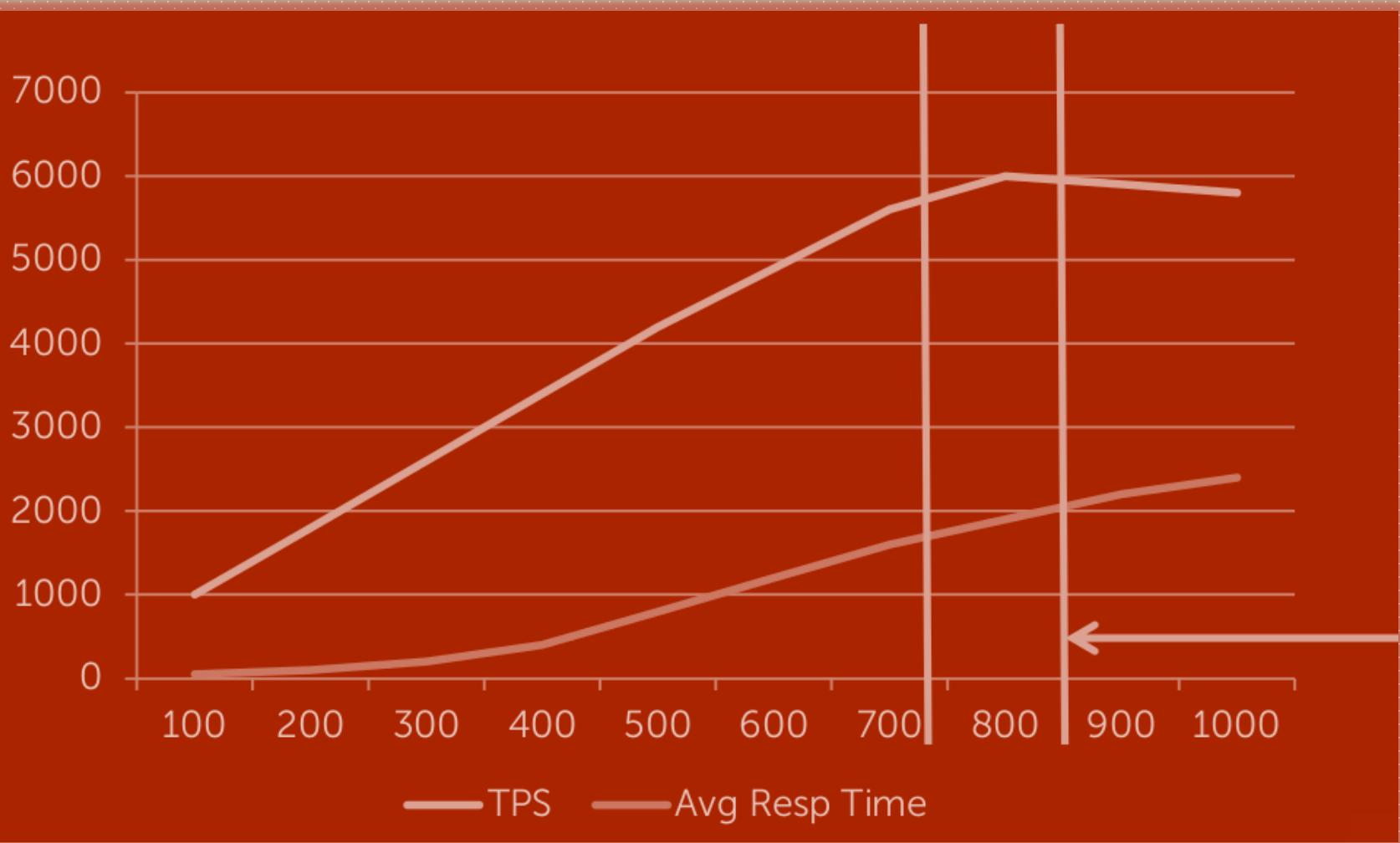
Требуется публикация суммарной стоимости системы, включая оборудование, лицензии, 5 лет сопровождения

Тест специфицирован на языке функциональных требований, без требований к коду или фрагментов кода

Введены правила масштабирования теста по количеству пользователей и размерам таблиц с пропорциональным ростом

Введено ограничение на время отклика: 95 % транзакций должно завершиться за 1 сек.

ГЛАВНОЕ: ОГРАНИЧЕНИЕ НА ВРЕМЯ ОТКЛИКА



TPC.ORG

Бенчмарковые войны продолжались: как проверить результат?

Нужна некоммерческая организация

1988: Transaction Processing Performance Council –
Совет по оценке производительности транзакционной обработки

Основатель – Омри Серлин и 8 согласных вендоров

Члены на 2017 год:

Actian

Cisco

Cloudera

Dell

DataCore

Fujitsu

HPE

Hitachi

Huawei

IBM

Inspur

Intel

Lenovo

Microsoft

Oracle

Pivotal

Red Hat

SAP

Teradata

VMWare

TPC-A И TPC-B: TECT

На основе TP1 – обработка розничнобанковских транзакций

READ 100 bytes from TTY (AID, TID, BID, DELTA)

BEGIN TRANSACTION

UPDATE ACCOUNT **WHERE** ACCOUNT_ID = AID:

READ ACCOUNT_BALANCE FROM ACCOUNT

SET ACCOUNT_BALANCE = ACCOUNT_BALANCE + DELTA

WRITE ACCOUNT_BALANCE TO ACCOUNT

WRITE TO HISTORY:

 AID, TID, BID, DELTA, TIME_STAMP

UPDATE TELLER **WHERE** TELLER_ID = TID:

SET TELLER_BALANCE = TELLER_BALANCE + DELTA

WRITE TELLER_BALANCE TO TELLER

UPDATE BRANCH **WHERE** BRANCH_ID = BID:

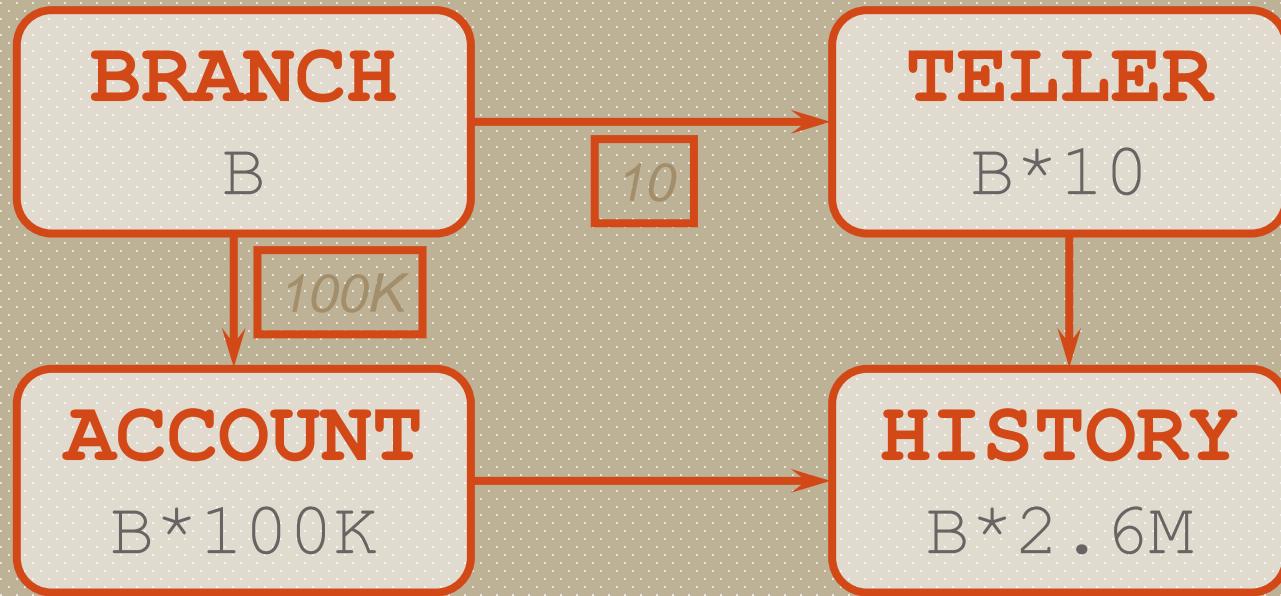
SET BRANCH_BALANCE = BRANCH_BALANCE + DELTA

WRITE BRANCH_BALANCE TO BRANCH

COMMIT TRANSACTION

WRITE 200 bytes to TTY (AID, TID, BID, DELTA)

TPC-A И TPC-B: МОДЕЛЬ



Десятисекундный
цикл у каждого
терминала

1 транзакция в
секунду на
каждое
отделение

Отклик от 90%
транзакций –
менее 2 сек

Считается:
средний
показатель
транзакций в
секунду за 15 мин

TPC-A И TPC-B: РАЗЛИЧИЯ

TPC-A

Терминалы

Время на реакцию
пользователя

TPC-B

Серверный тест

Сокращённая
история (30 дней)

TPC-A И TPC-B: КРИТИКА

Наследие TPC1

Слишком простые, чтобы исключить подстройки

Невероятные разбросы результатов

Первый результат (HP, 1990)

38.2 tpsA

29200 \$/tpsA

Последний результат (DEC, 1994)

3700 tpsA

4800 \$/tpsA



1995: признаны несостоятельными

TPC-C: усложнение

5 типов транзакций

NEW-ORDER	PAYMENT	DELIVERY	ORDER-STATUS	STOCK-LEVEL
<ul style="list-style-type: none">• новая заявка от клиента	<ul style="list-style-type: none">• факт платежа• обновление баланса клиента	<ul style="list-style-type: none">• заявка на доставку• (пакетный процесс)	<ul style="list-style-type: none">• проверка статуса последней заявки от клиента	<ul style="list-style-type: none">• проверка уровня запасов на складе
45%	43%	4%	4%	4%

Больше вариаций

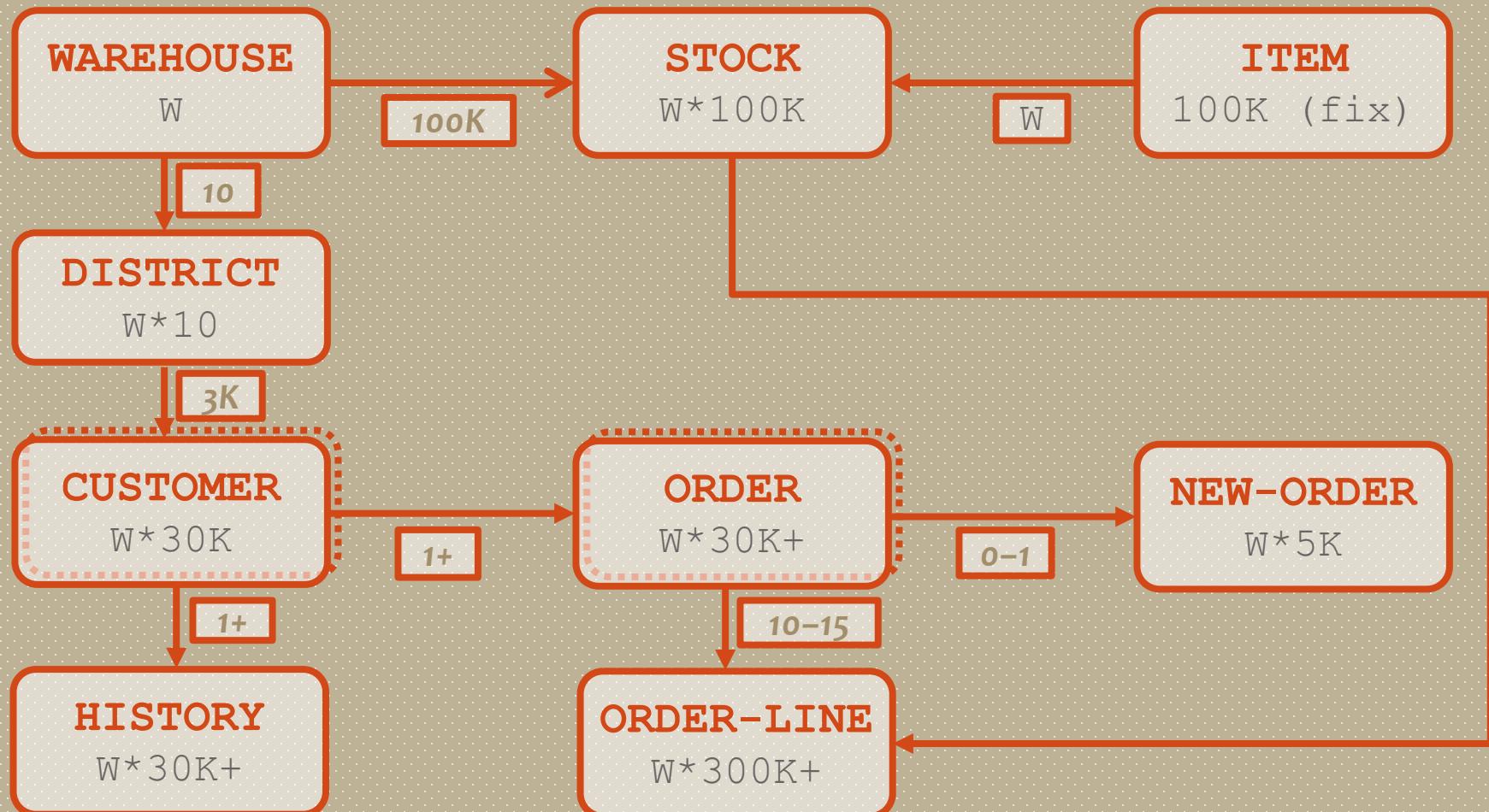
9 таблиц	Вставки, обновления, удаления, отмены операций	Доступ по первичными и вторичным ключам
----------	--	---

Пороги отклика для 90% транзакций

Менее 5 сек для интерактивных операций

Менее 20 сек для пакетных процессов

МОДЕЛЬ ДАННЫХ TPC-C



TPC-C: МАСШТАБИРОВАНИЕ

Квант масштабирования – склад (W)



10 терминалов на склад

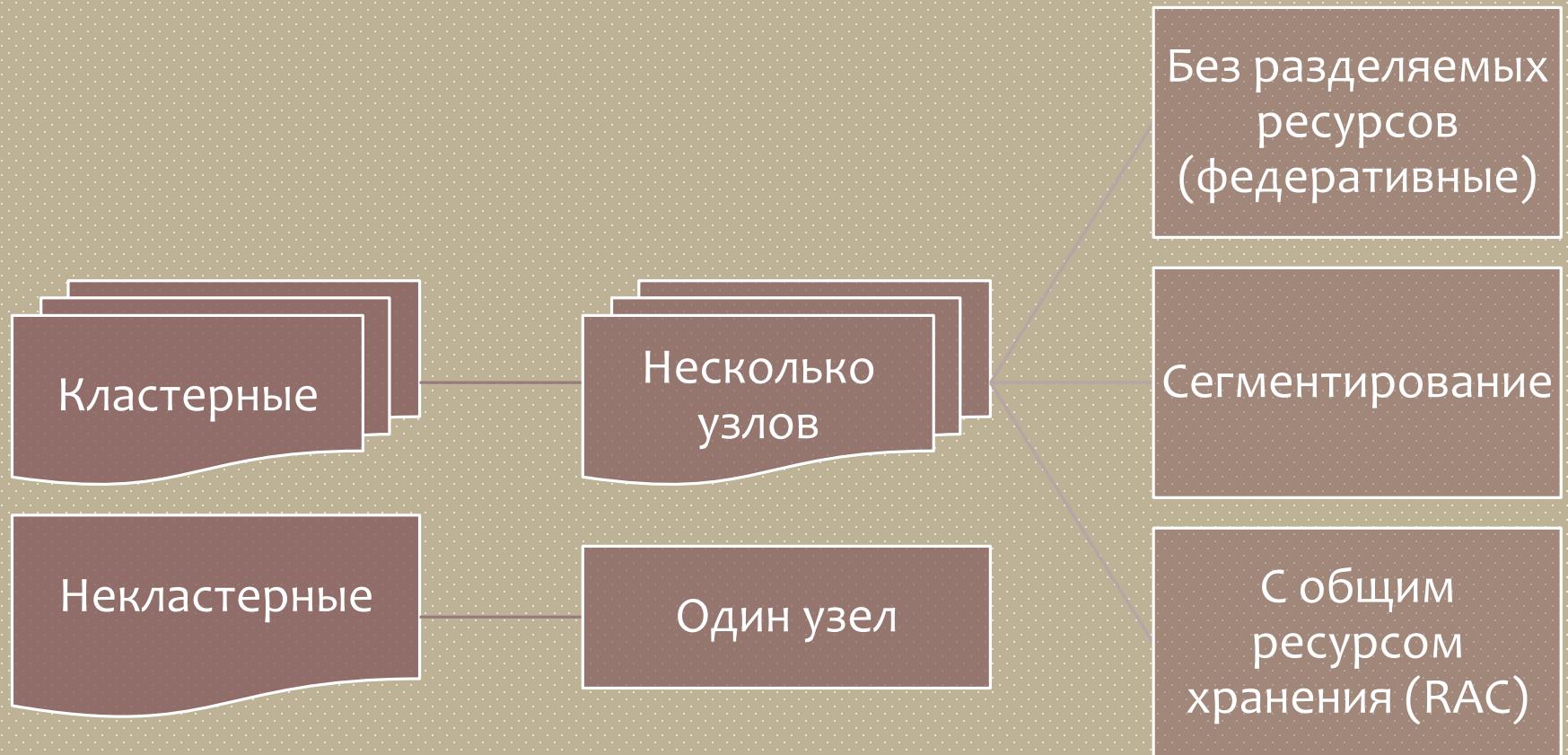


Максимум – 1,2 тpmC на терминал



Каждый новый склад →
+10 участков, + 100 единиц хранения, +300 тыс.
клиентов

TPC-C: КЛАСТЕРНЫЕ И НЕКЛАСТЕРНЫЕ



TPC-C: МЕТРИКИ

tpmC

Транзакций в
минуту

tpmC / \$

Стоимость
транзакции

Полная стоимость
оборудования

Лицензии и
поддержка на 3
года

W / ktpmC

Энергозатраты на
[кило]
транзакцию

ИНТЕРПРЕТАЦИЯ ТРС-С

По Аллану Паркеру (Alan Parker. Tuning databases on Solaris platform. Prentice-Hall, 2002)

$\text{tpmC} \times 2$

- ... не только создание заявок

$\text{tpmC} / 2$

- ... если нет монитора транзакций

$\text{tpmC} / 3$

- ... если на стороне оператора клиент типа Oracle Forms

$\text{tpmC} \times 2 / 3$

- ... если используются легковесные экранные формы (curses и т. п.)

$\text{tpmC} / 2$

- ... если в приложении не было тонкой настройки SQL

$\text{tpmC} / 2$

- ... если параллельно нагрузке формируются отчёты или выполняются пакетные задания

ПРОСТО ЛИ ОБМАНУТЬ TPC-C?

```
mount -t tmpfs -o size=2048g tmpfs /u01 tablespaces
```

CREATE UNLOGGED TABLE ...

_ALLOW_RESETLOGS_CORRUPTION = TRUE

_IN_MEMORY_UNDO = TRUE

_DB_BLOCK_HASH_LATCHES = 32768

...

ОТВЕТ ОТ СОВЕТА

Фиксация транзакции = записано на постоянный носитель

Защита «одиночного сбоя носителя»

Зеркальная синхронная фиксация на раздельно
запитанные носители

Журналы упреждающей записи на независимом
носителе и восстановление

Защита от перезагрузки одного узла

Зеркальная синхронная фиксация
на отдельном узле

Журналы упреждающей записи и
восстановление

Работоспособность при
перезагрузке

Тесты на все аспекты ACID закреплены в спецификации в
весьма жёсткой трактовке

ТРС-С: КРИТИКА 1990-Х

Нетипичные нагрузки

А как же отчётность?

А как же системы поддержки принятия решений?

Простая логика

Нет декларативных ограничений целостности

Нет триггерной логики

Даже оптовики работали по-другому!

Заявка: несколько неуспешных поисков с подстановочными символами

После каждой заявки печатается бумажный отчёт (иногда трижды)

«Обновление балансов» на каждой онлайн-операции при больших нагрузках невозможно → вставка + пакетный досчёт остатков

TPC: ОТМЕНЁННОЕ

TPC-D

1995–1999

Первая
попытка
OLAP-теста

TPC-R

2001–2005

Отчётность

TPC-W

2001–2005

Онлайн-
веб-
коммерция

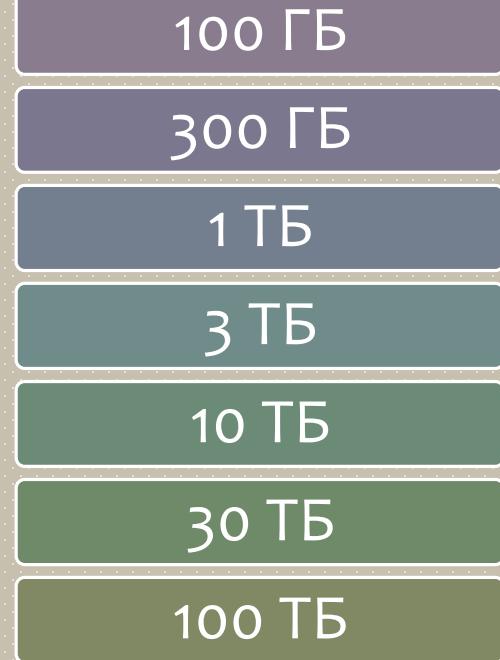
TPC-H

1999 год:
OLAP vs OLTP

“Ad-hoc
decision
support”

Вместо
признанного
неудачным
TPC-D

«Весовые
категории»

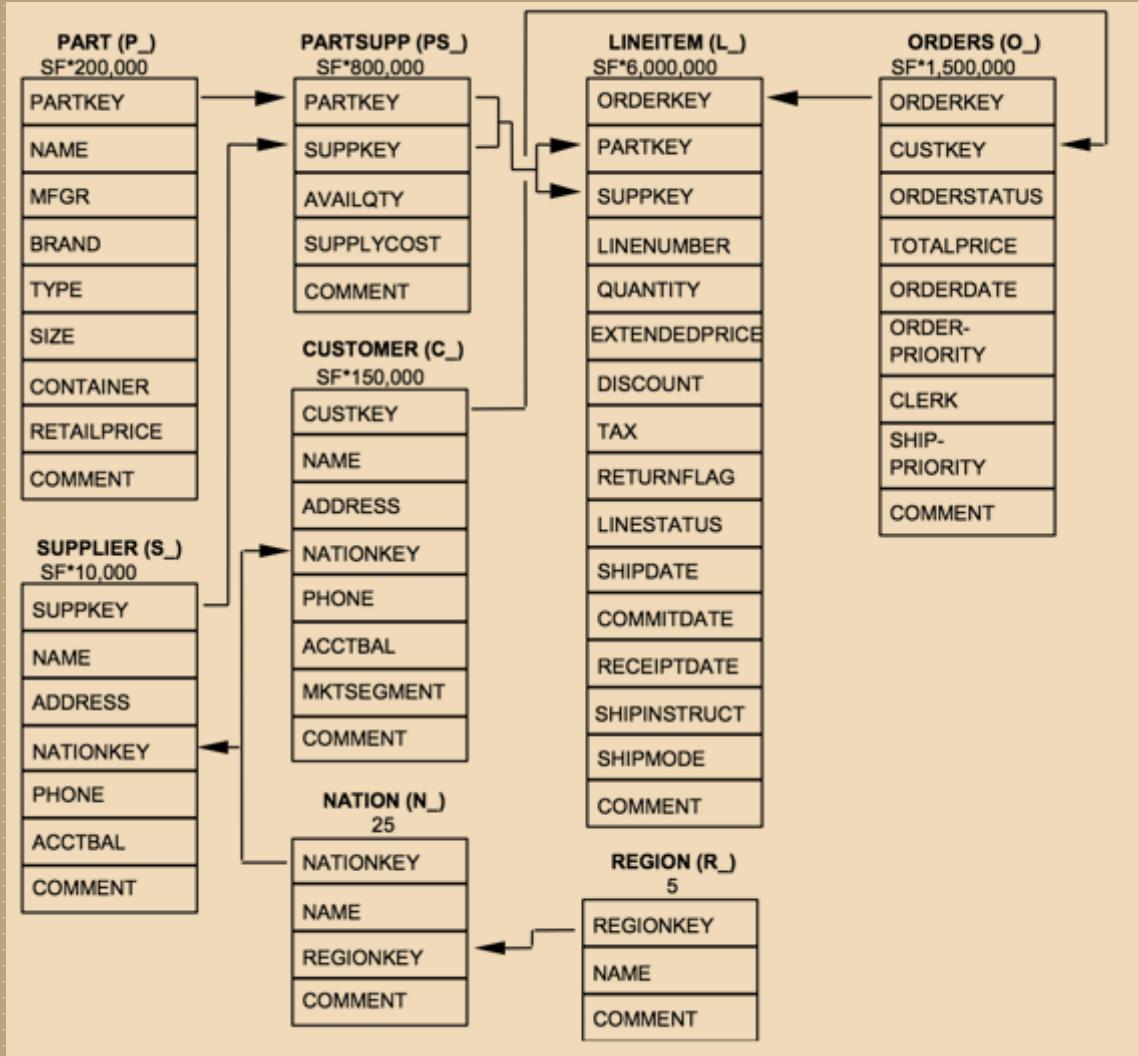


Параллельная
нагрузка

22 вида
сложных
запросов

2 типа
обновления
хранилища

TPC-H: МОДЕЛЬ



1 SF ~ 1 GB

Ограничения
целостности

Не звезда

TPC-H Q1

“Functional Query Definition”

SELECT

```
l_returnflag, l_linenstatus,  
sum(l_quantity) as sum_qty,  
sum(l_extendedprice) as sum_base_price,  
sum(l_extendedprice*(1-l_discount)) as sum_disc_price,  
sum(l_extendedprice*(1-l_discount)*(1+l_tax)) as sum_charge,  
avg(l_quantity) as avg_qty,  
avg(l_extendedprice) as avg_price,  
avg(l_discount) as avg_disc,  
count(*) as count_order
```

FROM lineitem

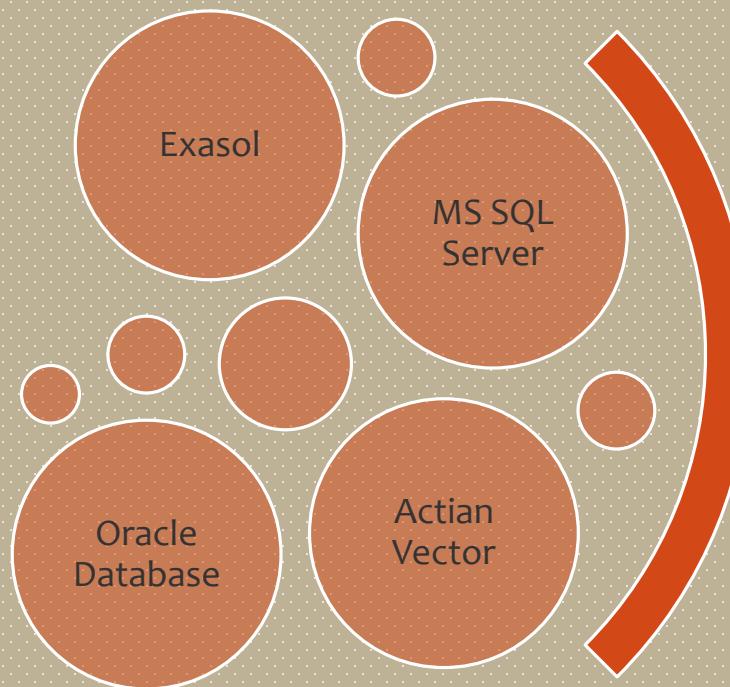
WHERE l_shipdate <= date '1998-12-01' - interval Δ day (3)

GROUP BY l_returnflag, l_linenstatus

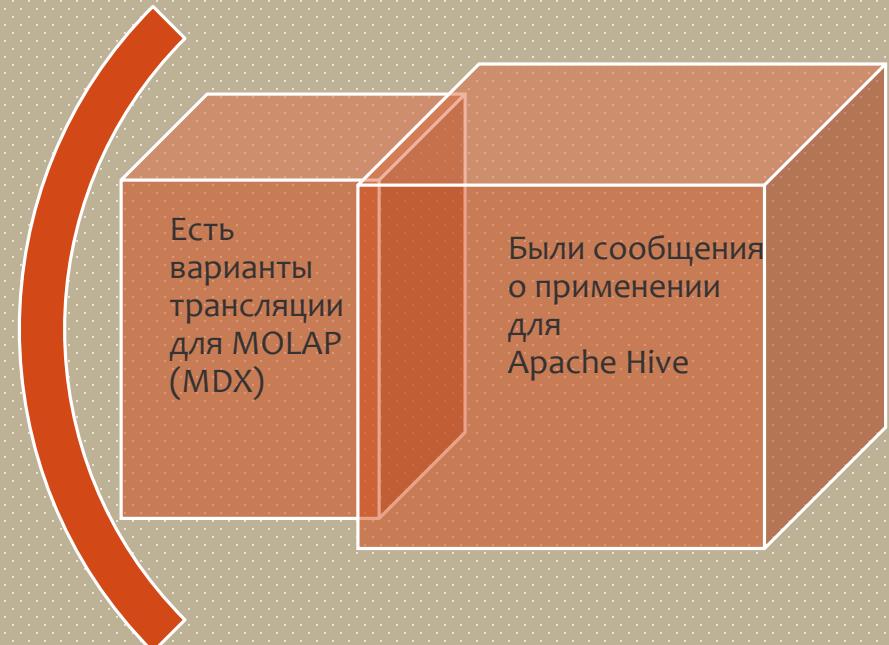
ORDER BY l_returnflag, l_linenstatus

Δ = random [60...120]

TPC-H: РЕЗУЛЬТАТЫ-2017



От 1 до 10 результатов в
каждой «весовой категории»



Напрямую трудно применимо
к нереляционным СУБД

TPC-E: «ПРАКТИЧНЫЙ OLTP»

Тоже OLTP, но более «гибридный»

Включённые ограничения целостности

Без мониторов транзакций

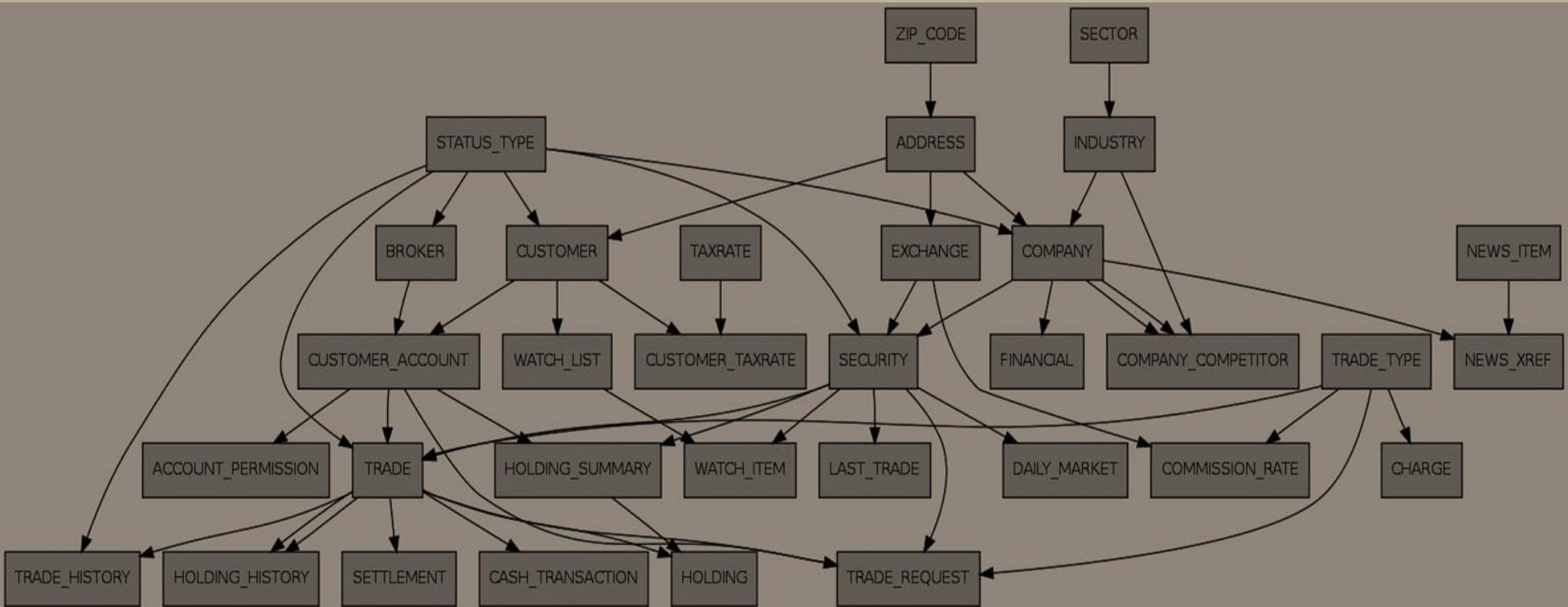
Больше чтений

Больше видов нагрузок

TPC-E: СЛОЖНЫЙ ПОТОК ОПЕРАЦИЙ

Транзакция	Сложность	Доступ	Инициатор	% успешных	% в нагрузке
Broker-Volume	4	Чтение	Брокер	100%	4,90%
Customer-Position	4	Чтение	Клиент	67%	13%
Market-Feed	3	Чтение-запись	Триггер	100%	1%
Market-Watch	3	Чтение	Клиент	100%	18%
Security-Detail	3	Чтение	Клиент	100%	14%
Trade-Lookup	3	Чтение	Брокер, клиент	100%	8%
Trade-Order	5	Чтение-запись	Клиент	33%	10,10%
Trade-Result	5	Чтение-запись	Триггер	83%	10%
Trade-Status	2	Чтение	Клиент	100%	19%
Trade-Update	3	Чтение-запись	Брокер, клиент	33%	2%

TPC-E: МОДЕЛЬ ДАННЫХ



Results displayed with a grey background are ***Historical Results***, which might not be up to date with regards to pricing and/or availability of HW or SW.

Note 1: The TPC believes it is not valid to compare prices or price/performance of results in different currencies.

Click on the column header to sort on that column; click again to reverse the sort order.

Hardware Vendor	System	v Performance (tpsE)	Price/tpsE	Watts/tpsE	System Availability	Database	Operating System	Date Submitted
Lenovo	Lenovo System x3950 X6	11,059	143.91 USD	NR	12/17/15	Microsoft SQL Server 2014 Enterprise Edition	Microsoft Windows Server 2012 R2 Standard Edition	12/17/15
FUJITSU	FUJITSU Server PRIMEQUEST 2800E2	10,058	187.53 USD	NR	11/11/15	Microsoft SQL Server 2014 Enterprise Edition	Microsoft Windows Server 2012 R2 Standard Edition	11/11/15
Lenovo	System x3950 X6	9,145	192.38 USD	NR	11/25/14	Microsoft SQL Server 2014 Enterprise Edition	Microsoft Windows Server 2012 Standard Edition	11/25/14
Lenovo	Lenovo System x3850 X6	9,068	139.85 USD	NR	07/31/16	Microsoft SQL Server 2016 Enterprise Edition	Microsoft Windows Server 2012 R2 Standard Edition	05/31/16

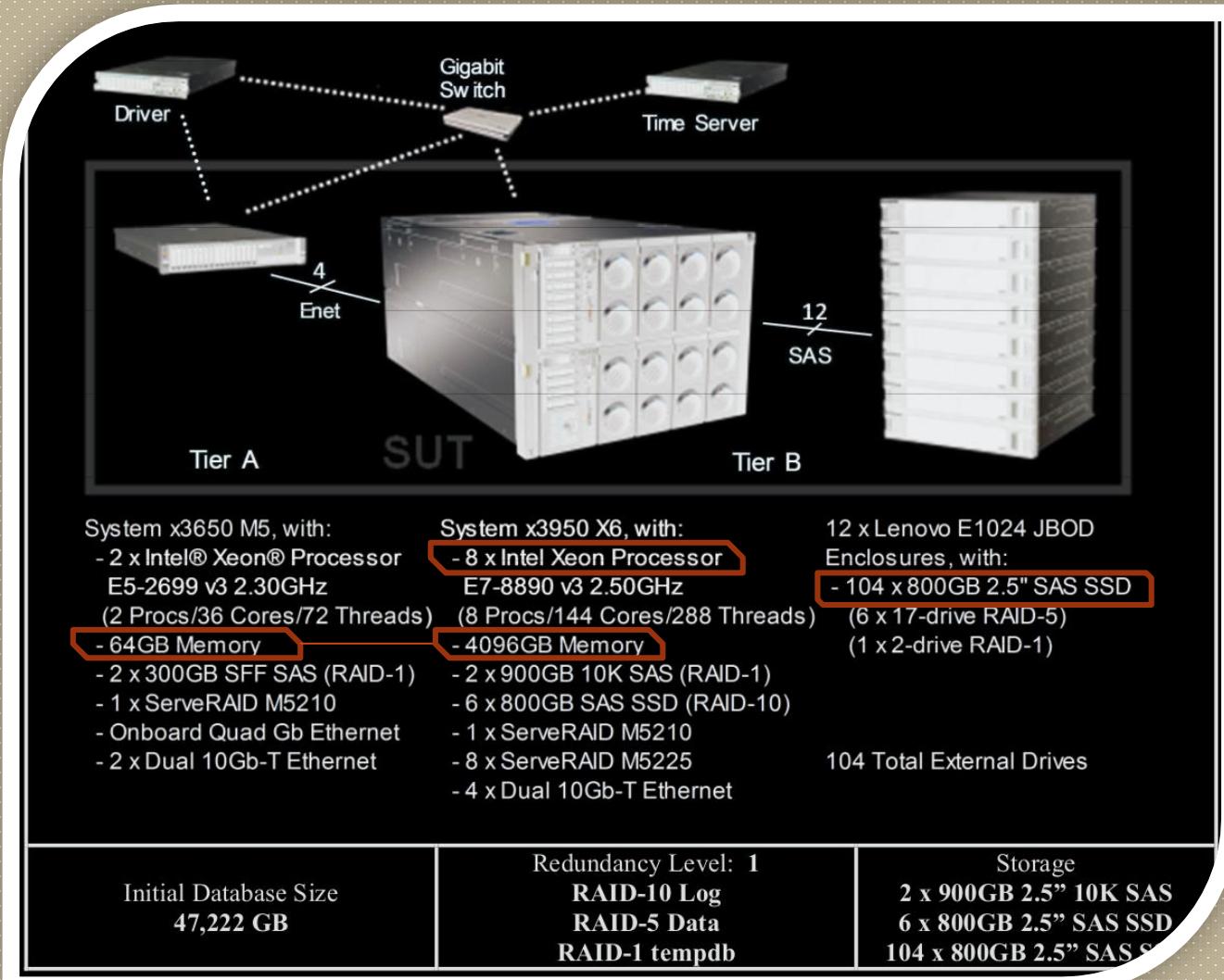
78 результатов

Все – некластерные, MS SQL Server на Windows Server x64

IBM	IBM BladeCenter HS22	8,796	116.62 USD	NR	08/10/07	Microsoft SQL Server 2005 Enterprise Edition x64 SP2	Microsoft Windows Server 2003 Enterprise Edition x64 SP1	12/11/07
DELL	Dell PE2900	145	788.69 USD	NR	12/11/07	Microsoft SQL Server 2005 Enterprise Edition x64 SP2	Microsoft Windows Server 2003 Enterprise Edition x64 SP1	12/11/07

'NR' in the Watts/tpsE column indicates that no energy data was reported for that benchmark.

11 KTPSE, \$144/TPSE (1,6 MIO)



TPC-ABCEN

	TPC-A/B	TPC-C	TPC-E	TPC-H
<i>Предметная область</i>	Розничный банк	Оптовый склад	Брокерский дом	Дискретный сбыт
<i>Таблиц</i>	4 ⬇	9 ⬇	33 ⬆	8 ⬇
<i>Чтение-запись</i>	100%	92%	23,10%	8%
<i>Чтение</i>	0%	8%	76,90%	92%
<i>Ограничения целостности</i>	0 ⬇	0 ⬇	22 ⬆	17 ⬆
<i>Генерация данных</i>	Случайная	Случайная	Псевдо-действительная	Случайная
<i>Метрика</i>	tpsA	tpmC	tpsE	QphH@Size

НОВЫЕ TPC-X И TPCx-X

TPC-DI

ETL-нагрузки

TPC_DI_RPS

о результатах

TPC-DS

Поддержка
принятия
решений,
«включая большие
данные»
«включая
Hadoop/Spark»)

QphDS@Size

о результатах

TPCx-BB

Экспресс-
бенчмарк для
аналитических
больших данных
на базе BigBench

BBpm@Size

о результатах

TPCx-HS

Экспресс-
бенчмарк для
HDFS-
совместимых
систем по типу
TeraSort

HSph@Size

От 1 до 4
результатов в
разных «весовых
категориях»

ПУБЛИКАЦИЯ?

Обязательный аудит – sizing.com

Запрещено публиковать результаты TPC-тестов, полученные по средствами TPC, кроме следующих случаев:

Публикация на TPC.org

Академическая или исследовательская публикация, не применимая с маркетинговыми целями

С однозначным утверждением, что полученный результат не является сравнимым с результатами TPC.org

С разрешения TPC.org

«ОГОВОРКА ДЕВИТТА» В 2017 ГОДУ?

MSFT
EULA

- “You may not disclose the results of any benchmark test ... without **Microsoft’s** prior written approval”

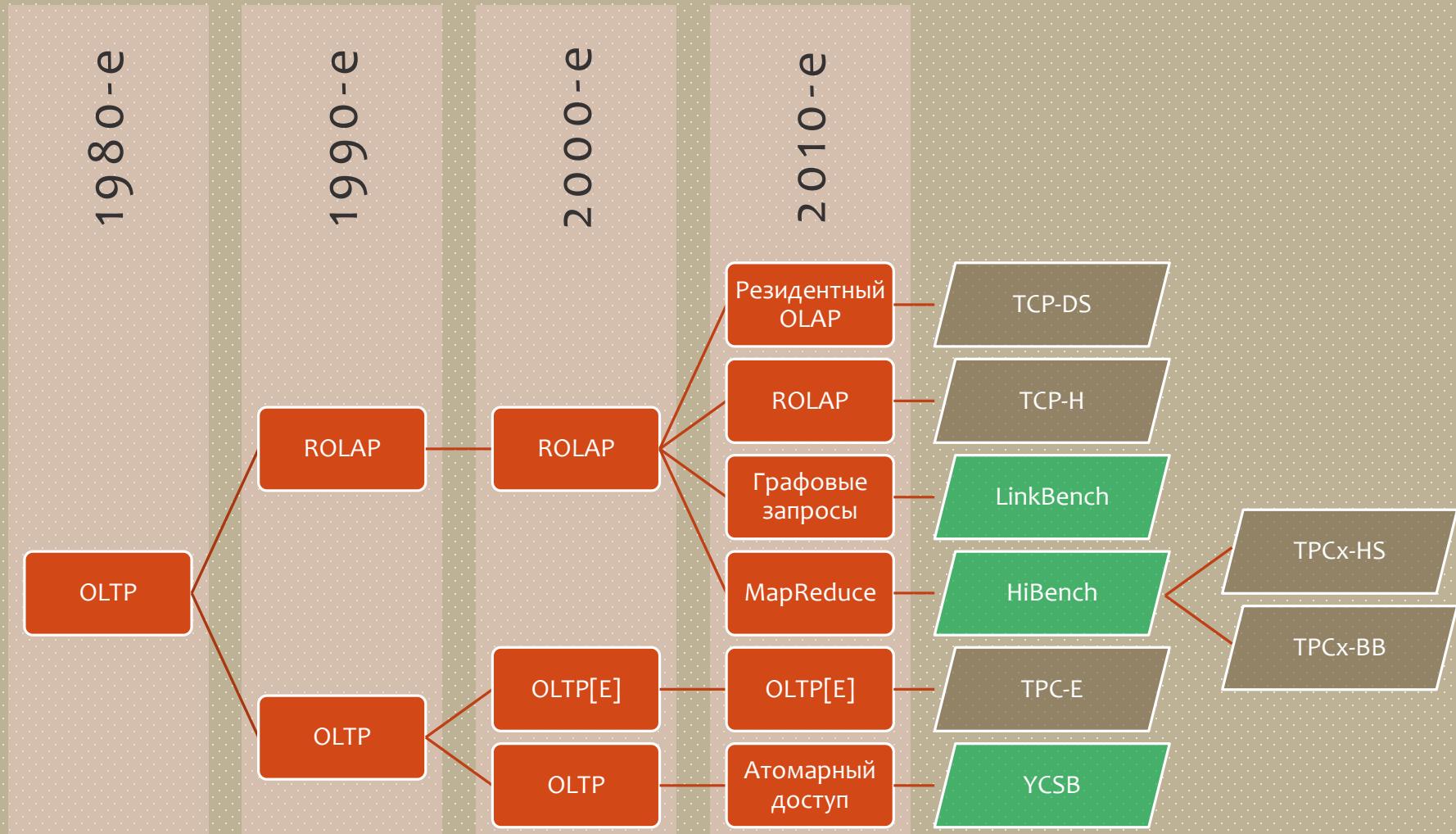
OTN
Lic.

- “You may not disclose results of any Program benchmark tests without **Oracle’s** prior consent”

IBM
IPLA

- “Licensee may disclose the results of any benchmark test of the Program or its subcomponents to any third party provided that Licensee, if ... A) ... B) ... C) ...”

НОВЫМ НАГРУЗКАМ – НОВЫЕ БЕНЧМАРКИ



ОСНОВНЫЕ БЕНЧМАРКИ «НОВОЙ ВОЛНЫ»

Terrasort Benchmark

- Воплощён в TPCx-HS

BigBench

- Воплощён в TPCx-BB

Intel HiBench

- Серия тестов для Map-Reduce

Yahoo! Cloud Services Benchmark

- Набор «эталонных обстрелов» для NoSQL-систем

Linkbench

- Графообразная нагрузка на РСУБД

INTEL HIBENCH ДЛЯ HADOOP

Characterization Results

Workload	System Resource Utilization	Data Access Patterns	Map/Reduce Stage Time Ratio
Sort	I/O bound	Map Red	
WordCount	CPU bound	Map Red	
TeraSort	Map stage : CPU-bound; Red stage : I/O-bound	Map Red	
Nutch Indexing	I/O bound, high CPU utilization in map stage	Map Red	
Page Rank (1st & 2nd job)	CPU-bound in all jobs	Map Red	
Bayesian Classification (1st & 2nd job)	I/O bound, with high CPU utilization in map stage in the 1 st job	Map Red	
K-means Clustering	CPU bound in iteration; I/O bound in clustering	Map Red	
Enhanced DFSIO	I/O-bound	trivial	trivial

6 data less data even less data compressed



YAHOO! CLOUD SERVICES BENCHMARK

Новый инструмент «бенчмаркового маркетинга»

Смело используется исследователями

(V. Abramova et al. Experimental Evaluation of NoSQL Databases // IJDMS Vol.6, №3, 2014)

Cassandra

HBase

Elastic
search

MongoDB

Oracle
NoSQL

OrientDB

Redis

Scalarm

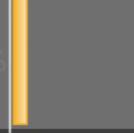
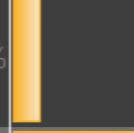
Tarantool

Voldemort

Обстрел с одной нагрузочной станции

Вместо транзакции – атомарная операция
(возможно, чтение нескольких записей)

YCSB: НАГРУЗКИ

Нагрузка	Тип	Пример	% чтений	% записей
A	Атомарные чтения и записи	<i>запись действий в веб-сессии</i>	 50%	 50%
B	Преимущественное чтение	<i>тегирование фото</i>	 95%	 5%
C	Атомарные чтения	<i>доступ к профилю пользователя</i>	 100%	 0%
D	Чтение новых записей	<i>обновление стаутса пользователя и его показ</i>	 95%	 5%
E	Запрос коротких интервалов	<i>запрос потока дискуссии</i>	 90%	 10%
F	Чтение - обновление - запись	<i>изменение профиля пользователем</i>	 50%	 50%

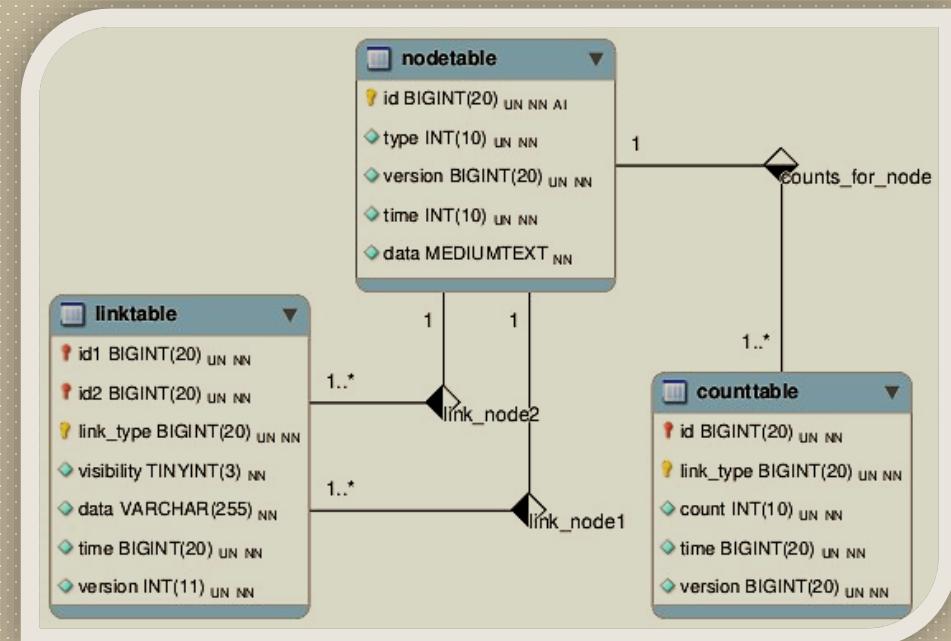
LINKBENCH

Нагрузка Facebook

- Ранние графовые тесты были ориентированы на граф-анализ
- Реальная нагрузка крупной интернет-компании
- Транзакционность (MVCC)

Методологично

- Статистические законы при генерации
- Показатели – средние времена отклика на одной из
- Средний отклик в 99-м процентиле
- Документов в секунду – на запись
- Запросов в секунду – на чтение



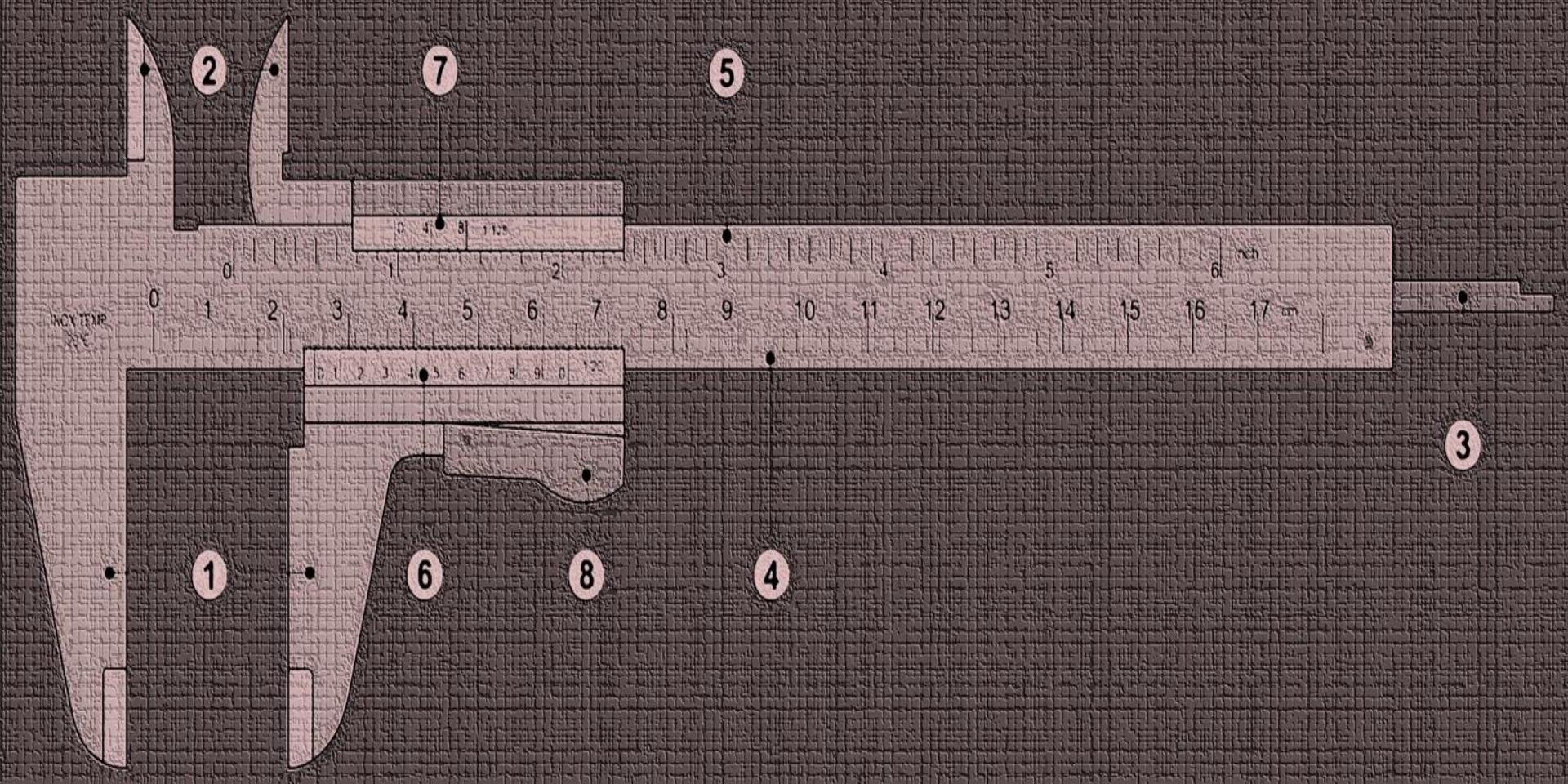
Вопрос Facebook

MySQL
(InnoDB vs
TokuDB)

HBase

MongoDB /
TokuMX

OrientDB



ЗАМЕР «КАРМАННЫМИ СРЕДСТВАМИ»

Возможен?
Корректен?
Репрезентативен?
Сравним?
Повторяем?

PGBENCH

TPC-B

Стандартная
часть
PostgreSQL

Запуск
одной
строкой

Фактический стандарт для нужд
PostgreSQL

Что лучше – XFS или
ext4?

Таблицы на SSD,
индексы на HDD, или
наоборот?

Размер блока – 4К
или 8К?

SYSBENCH

Утилита MySQL (автор – Алексей Копытов)

fileio

cpri

memory

threads

mutex

oltp

Свообразные тесты, не похожие на какие-либо
стандартные бенчмарки

Широко используется для внутренних сравнений MySQL,
MariaDB и их форков

TPCE-MYSQL

TPC-E-нагрузчик для MySQL от Percona

You can download it from our Launchpad Percona-tools project, it's

```
1 | bzr branch lp:~percona-dev/perconatools/tpcemysql
```

Important **DISCLAIMER**:

Using this package you should agree with **TPC-E License Agreement**,

which in human words is:

- You can't name results as "TPC Benchmark Results"
- You can't compare results with results published on <http://www.tpc.org/> and you can't pretend the results are compatible with published by TPC.

HAMMERDB

Свободная Java-GUI-программа, запускающая
TPC-C (?) и TPC-H (?) с поддержкой множества СУБД:

Oracle
Database

Microsoft
SQL Server

IBM DB2

TimesTen

MySQL

MariaDB

Postgre
SQL

Postgres
Plus AS

Greenplum
DB

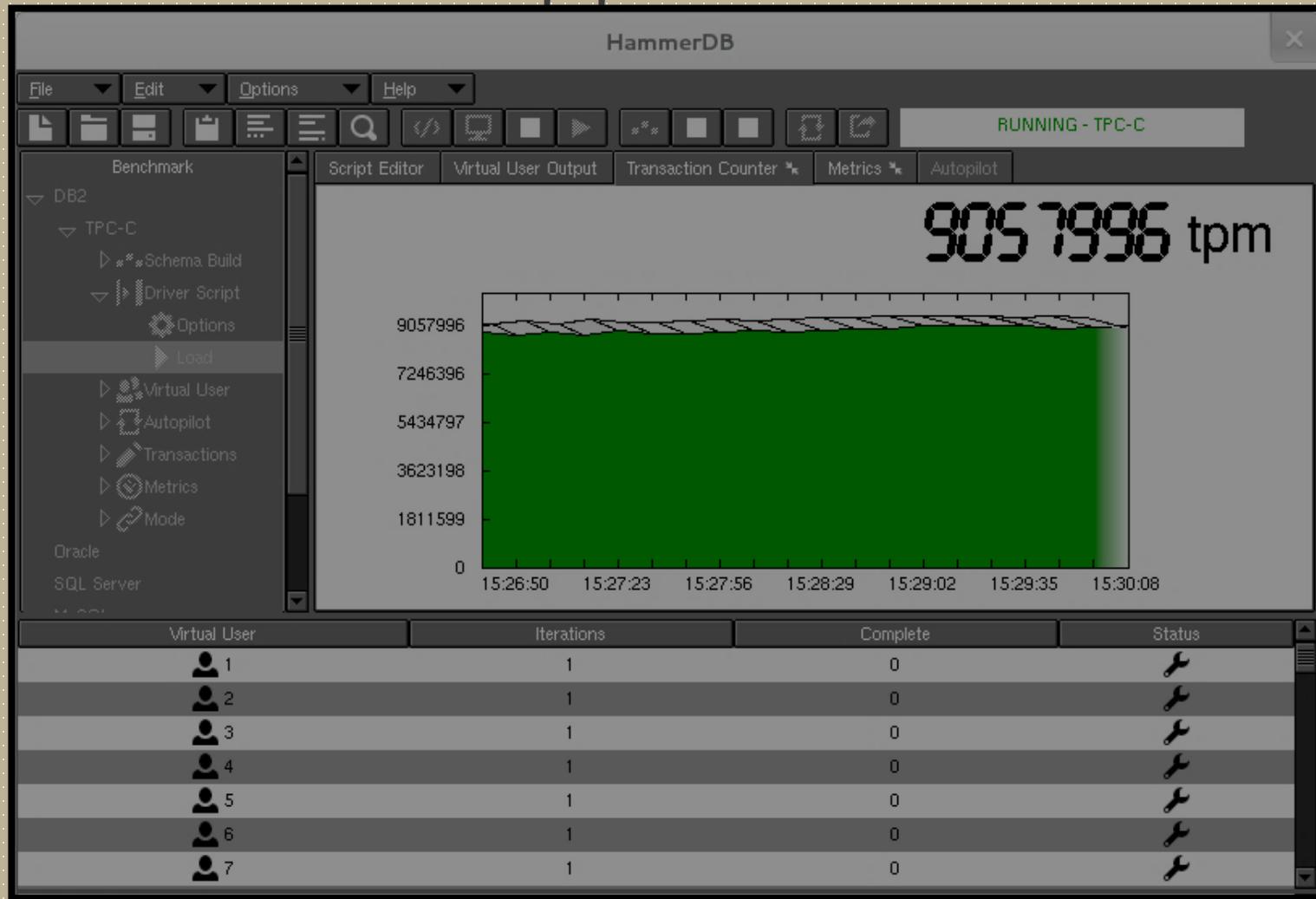
Redis

Amazon
Aurora
(MySQL)

Amazon
Redshift
(ParAccel)

Trafodion
SQL on
Hadoop

HAMMERDB – TPC-C В «ОДИН КЛИК»?



HAMMERDB: TPC, НО НЕ СОВСЕМ...

TPC-C

Не эмулируется нагрузка всего
окружения

Нагрузочная станция – в роли
единственного монитора транзакций

TPC-H

Единственная нагрузочная станция

При этом нет востребованных в современных бенчмарков

TPC-E

TPC-DS

HAMMERDB: ПОПУЛЯРНОСТЬ

Используется
вендорами
оборудования
и платформ

Но аккуратно,
в формате
«блог нашего
сотрудника»

Хаб с
результатами
тестов

Раздел
«*Performance
Data*»

SWINGBENCH

TPC-C-подобная нагрузка на Oracle Database и Timesten

+ Своеобразный OLTP- тест с большой PL/SQL-логикой

Специфичные для Oracle DB средства мониторинга и анализа (AWR, etc.)

Поддержка координатора

GUI и командная строка

Сравнимость только внутри инструмента

Нерепрезентативные результаты в условиях RAC

DELL BENCHMARK FACTORY FOR DATABASES

Коммерческий инструмент (наследие Quest Software)

Oracle
Database

MySQL

MS SQL
Server

SQLite

SQL
Anywhere

TPC-C ?

TPC-D

TPC-E

TPC-H

ASPЗAP

Поддерживает архитектуру
с несколькими нагрузочными станциями
(Windows)

OSDLDBT.SOURCEFORGE.NET

MySQL

PostgreSQL

...расширяемо

While the inspiration for these workloads are the TPC-*<...>*, workloads are entirely different and results obtained from them should not and can not be compared to TPC results.

The use of any supplied results of these tests for commercial purposes is expressly prohibited.

DBT-1

- TPC-W

DBT-2

- TPC-C

DBT-3

- TPC-H

DBT-4

- TPC-App

DBT-5

- TPC-E

OLTPBENCH

github.com/oltpbenchmark/oltpbench

Java-средство для командной строки

Для любой РСУБД с поддержкой JDBC

Особая версия для Hstore (VoltDB)

TPC-C

Wikipedia

Synthetic
Resource
Stresser

Twitter

Epinions.com

TATP

AuctionMark

SEATS

YCSB

JPAB
(Hibernate)

CH-
benCHmark

Voter

SI Bench
(Snapshot
Isolation)

SmallBank

LinkBench

TPC TOOLS



Исходные
тексты на Си

Нет для ТРС-С:
только «сэмпл» в
спецификации

Вся
соединяемость
и прочая
работа –
своими руками

TPC-* BY EXAMPLE

По «отчётам о полном раскрытии информации» на TPC.org

```

{text*),(fmt),(fmt1),(fmt2),(rlen),(rcode)))))

{errpt(lda, cursor, (text *)ftype); return(-1);}\

else\

    DISCARD 0

#define OEXFET(lda, cursor, nrows, cancel, exact)\

if (oexfet((cursor),(nrows),(cancel),(exact)))\

{if ((cursor)>rc == 1403) DISCARD 0; \
else if (errpt(lda, cursor, (text *)"OEXFET") == RECOVERR) \
    {rol(lda); return(RECOVERR);}\ \
else {rol(lda); return(-1);}\ \
else\

    DISCARD 0

#define OOPEN(lda, cursor)\

if (oopen((cursor),(lda),(text*)0,NA,NA,(text*)0,NA))\
{errpt(lda, cursor, (text *)"OOPEN"); return(-1);}\ \
else\

    DISCARD 0

#define OPARSE(lda, cursor, sqlstm, sql, defflg, lngflg)\

if (oparse((cursor),(sqlstm),(sb4)(sql),(defflg),(ub4)(lngflg)))\
{errpt(lda, cursor, sqlstm); return(-1);}\ \
else\

    DISCARD 0

/* additions done for 814 -shishir */

#define OCI_ATTR_SERVER          OCI_ATTR_SERVER
#define OCI_ATTR_USERCTXT        OCI_ATTR_SESSION
#define OCI_ATTR_ROWCNT          OCI_ATTR_ROW_COUNT
#define OCI_HTYPE_ERR            OCI_HTYPE_ERROR
#define OCI_HTYPE_STMT           OCI_HTYPE_STMT

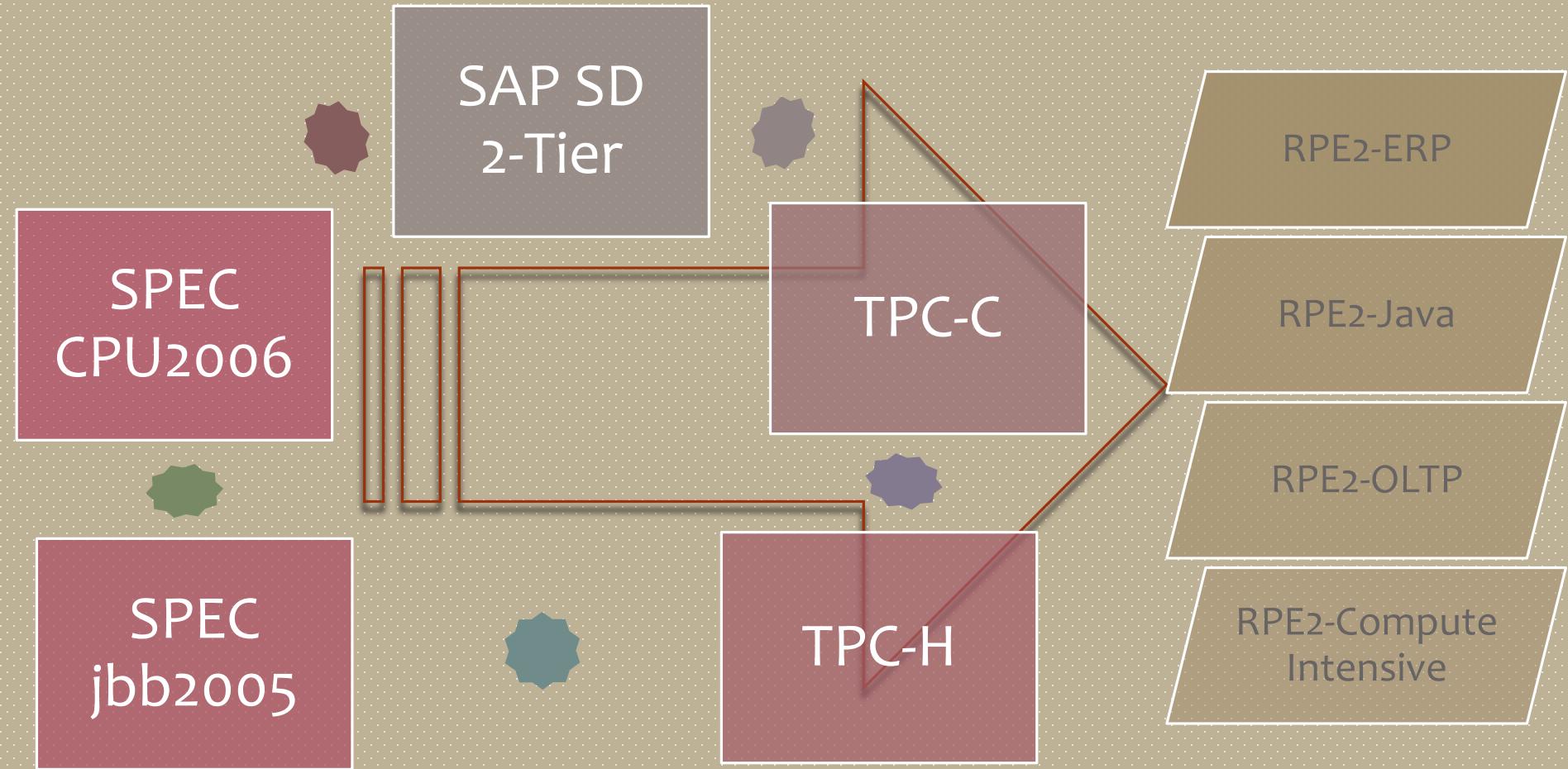
tpcc.h
/* $Header: tpcc.h 7030100.1 95/07/19 15:10:55 plai Generic $ Copyr (c)
1993 Oracle */
/*
=====
+-----+ | Copyright (c) 1995 Oracle Corp, Redwood Shores, CA
| OPEN SYSTEMS PERFORMANCE GROUP || All Rights Reserved |
+-----+
=====+ | FILENAME | tpcc.h | DESCRIPTION | Include file for
TPC-C benchmark programs.
+-----+
=====+*#ifndef TPCC_H #define TPCC_H #ifndef FALSE #
define FALSE 0 #endif #ifndef TRUE # define TRUE 1 #endif #include
#include #include #include #ifndef boolean #define boolean int #endif
#include #include #include /* #ifdef __STDC__ #include "ociapr.h" #else
#include "ocikpr.h" #endif */ typedef struct cda_def csrdef, typedef struct

```

```
o_carrier_id, e....  
ol_i_id[15]; extern int ol_quantity[15];  
ol_del_len[15]; extern text ol_delivery_d[15][11]; /* xnie - begin */ extern  
OCIRowid *o_rowid; /* xnie - end */ /* for payment transaction */ extern int int  
c_w_id; extern int c_d_id; extern int h_amount; extern char w_street_1[21];  
extern char w_street_2[21]; extern char w_city[21]; extern char w_state[3];  
extern char w_zip[10]; extern char d_street_1[21]; extern char  
d_street_2[21]; extern char d_city[21]; extern char d_state[3]; extern char  
d_zip[10]; extern char c_street_1[21]; extern char c_street_2[21]; extern char  
c_city[21]; extern char c_state[3]; extern char c_zip[10]; extern char  
c_phone[17]; extern text c_since_d[11]; extern char c_credit[3]; extern int  
c_credit_lim; extern float c_discount; extern char c_data[201]; extern text  
h_date[20]; /* for new order transaction */ #ifndef TUX extern int  
nol_i_id[15]; extern int nol_supply_w_id[15]; extern int nol_quantity[15];  
extern int nol_qtyint0[15]; extern int nol_qtyint1[15]; extern int  
nol_ytdqty[15]; extern int nol_amount[15]; extern int o_all_local; extern float  
w_tax; extern float d_tax; extern float total_amount; extern char i_name[15]  
[25]; extern int i_name_strlen[15]; extern ub2 i_name_strlen_len[15]; extern  
ub2 i_name_strlen_rcode[15]; extern ub4 i_name_strlen_csize; extern int  
s_quantity[15]; #if 0 extern char brand_gen[15]; #endif extern ub2  
brand_gen_len[15]; extern ub2 brand_gen_rcode[15]; extern ub4  
brand_gen_csize; extern int i_price[15]; extern char brand_generic[15][1];  
extern int statust; extern int tracelevel; /* Miscellaneous */ extern OCIDate  
cr_date; extern OCIDate c_since; extern OCIDate o_entry_d_base; extern
```

RPE2

Суперкомпозитный показатель от Gartner (Ideas)



BENCHWARE

«Швейцарский нож» от Манфреда Дрозда: *Peakmarks Benchware*

OraCPU

PL/SQL op
• [ops]

PL/SQL alg
• [ops]

OraSRV

In-memory
SQL
• [ms]
• [dbps]
• [tps]
• [rps]

OraSTO

SeqIO
• [GBps]
• [iops]

RandIO
• [GBps]
• [iops]

OraOLTP

OLTP Select
• [rps]
• [tps]

OLTP Update
• [rps]
• [tps]

OraLoad

TransLoad
• [rps]
• [tps]

BulkLoad
• [rps]
• [tps]

OraAgg

OraAgg &
Rep
• [rps]
• [tps]

Только Oracle Database, только PL/SQL и SYS.V_\$%

МАШИНЫ БАЗ ДАННЫХ?

Предконфигурированные
аппаратно-программные
комpleксы для СУБД

*Кто, как не они должны
быть «проверены и
отмерены» для
практических целей?*

TERADATA

Последние публикации Teradata с TPC-H:

 a division of	WorldMark 5250	18,542	638.37 USD	NR	07/27/01	Teradata V2R4.1	MP-RAS 3.02.00	10/09/01
 a division of	NCR 5350	79,528	212.97 USD	NR	12/20/02	Teradata V2R5.0	MP-RAS 3.02.00	01/06/03

Лицензируется по
«внутреннему Qph»
– **tPerf** [Traditional Performance]

EXADATA

Не публикуются tps, Qph...

«Паспортные показатели» (X6-8)

	НС	EF
<i>Флэш</i>		
Пропускная способность SQL, Гбайт/с	300	350
SQL kIOPS	4140	4130
Ёмкость, ТБ	179,2	358,4
<i>Диски</i>		
Пропускная способность SQL, Гбайт/с	25	-
SQL kIOPS	36	-
Ёмкость, ТБ	1344	-
<i>Общее</i>		
Скорость загрузки, Тбайт/ч	21	21

$$\sum V \times \sqrt{\text{IOPS}} \approx \text{Const}$$

IBM PURE DATA FOR OPERATIONAL ANALYTICS

Не публикуются Qph...

«Паспортные данные» – про ввод-вывод

	Extra Small	Small	Medium	Large
Пропускная способность, Гбайт/с	3,2	9,6	16	22,4
kIOPS	34	57	148	205
Скорость загрузки, Тбайт/ч	1,161	3,484	5,807	8,13

“SQL IOPS”

IOPS со стороны Oracle Database

Orion

(Oracle IO Numbers)

SLOB

(рекомендуют EMC,
Flashgrid)

Benchware

(?)

DBMS_RESOURCE_MANAGER.
.CALIBRATE_IO

Получение SQL IOPS для других СУБД?

Статистические представления (... IO_STATS...)

АТОМИЗАЦИЯ АГРЕГАТНЫХ ПОКАЗАТЕЛЕЙ

tpmC

SQL IOPS

QpmH

SQL
bandwidth

АТОМИЗАЦИЯ ПОКАЗАТЕЛЕЙ: PRO

Независимость от заданных
моделей и схем

Признаются
производителями машин
баз данных

Доступны для «карманной
инструментовки»

Показательны для широкого
класса СУБД и СУБД-
подобных систем

- Показательны не только
для 3NF, снежинок, звёзд
- Включены в «паспортные
показатели»
- Могут быть запущены со
стороны БД
- Могут быть интересны для
NoSQL

АТОМИЗАЦИЯ ПОКАЗАТЕЛЕЙ: ... ET CONTRA

Вопросы к
методологии

Неочевидный
прикладной смысл
результата

Нет готового
инструмента для
большинства
известных СУБД

Результаты
CALIBRATE_IO и
Benchware – разные

Для разных СУБД
одно и то же число
операций – разное
количество
транзакций

Есть для Oracle DB,
IBM DB2, MS SQL



ПРОГОН РЕАЛЬНОЙ НАГРУЗКИ И БЕНЧМАРКИ ПРИЛОЖЕНИЙ

Самый
практичный
метод?

ВОСПРОИЗВЕДЕНИЕ НАГРУЗКИ: ТОЧНЫЙ ОПЫТ

Средства на стороне СУБД

Oracle Real Application Testing
(Database Replay)

MS SQL Server Distributed Replay

Средства эмуляции нагрузки

JMeter

LoadRunner

...

Системы с полным API-доступом (обычно JSON via HTTP)

Журналирование и
«самозахват»

Повтор с интервалами

Разделение типов
нагрузки
(категорий
пользователей)

Масштабирование по
данным?

ТОНКОСТИ «ЗАШНУРОВКИ»

Как
размножить
базу данных?

«Вегетативное
размножение»

Повторные со
случайным
сдвигом?

Наложение
реальных данных
(из открытых
данных)

Влияние на
аналитику

Предсказуемые
выборки

Пониженная
селективность

Влияние на
OLTP

Ключи, индексы...

ТИРАЖИРУЕМЫЕ ПРИЛОЖЕНИЯ

SAPS [SAP Application Performance Standard] на модуле SD

Oracle E-Business Suite Standard Benchmarks

Order-to-Cash

OLTP

Payroll

...

Microsoft Dynamics AX Application Benchmark Toolkit

Серия бенчмарков для 1С

Показатель – число одновременно работающих пользователей
при заданном максимальном времени отклика

ДИСКУССИОННЫЕ ВОПРОСЫ

TPC-E и TPC-DS

- Смогут ли стать столь же популярными, как TPC-C и TPC-H в своё время?
- Проблема мегаконфигураций

«Карманный» инструментарий

- TPC-B повторить просто
- Остальные – сложно или с неполным соответствием

SQL IOPS & Bandwidth

- Признание, адаптация и стандартизация, интерпретация
- Возможны ли другие показатели подобного рода?

Стандартизация и универсализация тестов для новых нагрузок

- Атомарные (YCSB-тип)
- Графовые (LinkBench-тип)

Перспективы эталонных тестов на основе реальной нагрузки нового поколения

- Новые типы приложений – порталы, системы групповой работы
- Новые архитектуры слоя БД – сегментация, резидентные СУБД, резидентные гриды



СПАСИБО ЗА
ВНИМАНИЕ!

<mailto:anikolaenko@ibs.ru>
<mailto:anikolaenko@acm.org>