Андрей Николаенко
IBS

PGDAY'17 RUSSIA
КОНФЕРЕНЦИЯ ПО БАЗАМ ДАННЫХ

DBMS Benchmarks in a Nutshell

концентрированное обозрение сферы эталонного тестирования СУБД

IBS

INTERNET ARCHIVE WayBackMachine

http://www.tpc.org/tpcc/results/c-db-results.idc Go

1 captures
28 Feb 01 – 28 Feb 01

JAN FEB MAR
28
2000 2001 2002

PGDAY'17 RUSSIA

TPC Transaction Processing Performance Council

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
  - - FAQ
  - TPC-H

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
[illegible]	Fujitsu/ICL	teamServer M830ic	[illegible]	[illegible]	[illegible]	[illegible] NT Enterprise	BEA Tuxedo 6.4 CFS	09/22/98

296 результатов

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

Sybase Adaptive Server Enterprise [illegible]	[illegible]	[illegible]	9853	64.22US $	02/12/98	IBM AIX 4.[illegible]	[illegible]	[illegible]
Sybase Adaptive Server Enterprise 11.5.1	SGI	Origin2000 (c/s)	6040	121.98US $	07/31/98	SGI IRIX 6.5	BEA Tuxedo 6.3	04/23/98
Sybase Adaptive Server Enterprise 11.9.3	Sun	Enterprise 4500	50268	49.88US $	03/30/00	Sun Solaris 7	BEA Tuxedo 6.3	11/23/99
Sybase ASE 12.0.0.2	Sun	Starfire Enterprise 10000	156873	48.81US $	02/28/01	Sun Solaris 7	BEA Tuxedo 6.3	08/29/00
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	12/21/95

TPC-C, апрель 2017

PGDAY'17 RUSSIA

TPC

PC-C - Top Ten Performance Results
ersion 5 Results As of 29-Jun-2017 at 1:11 PM [GMT]

ote 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 All

Rank	Company	System	Performance (tpmC)	Price/tpmC	Watts/KtpmC	System Availability	Database	Operating System	TP Monitor	Date Submitted
1	TTA	Jet-speed HHA2212	139,909	1,730.00 KRW	NR	05/09/17	Goldilocks v3.1 Standard Edition	CentOS 6.6	Red Hat JBOSS Web Server	05/08/17
2	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

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

W3C CSS

SQL Anywhere

Goldilocks

PGDAY'17 RUSSIA

КОНФЕРЕНЦИЯ ПО БАЗАМ ДАННЫХ

РАННЯЯ ИСТОРИЯ

TP1

PGDAY'17 RUSSIA

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

PGDAY' RUSSIA '17

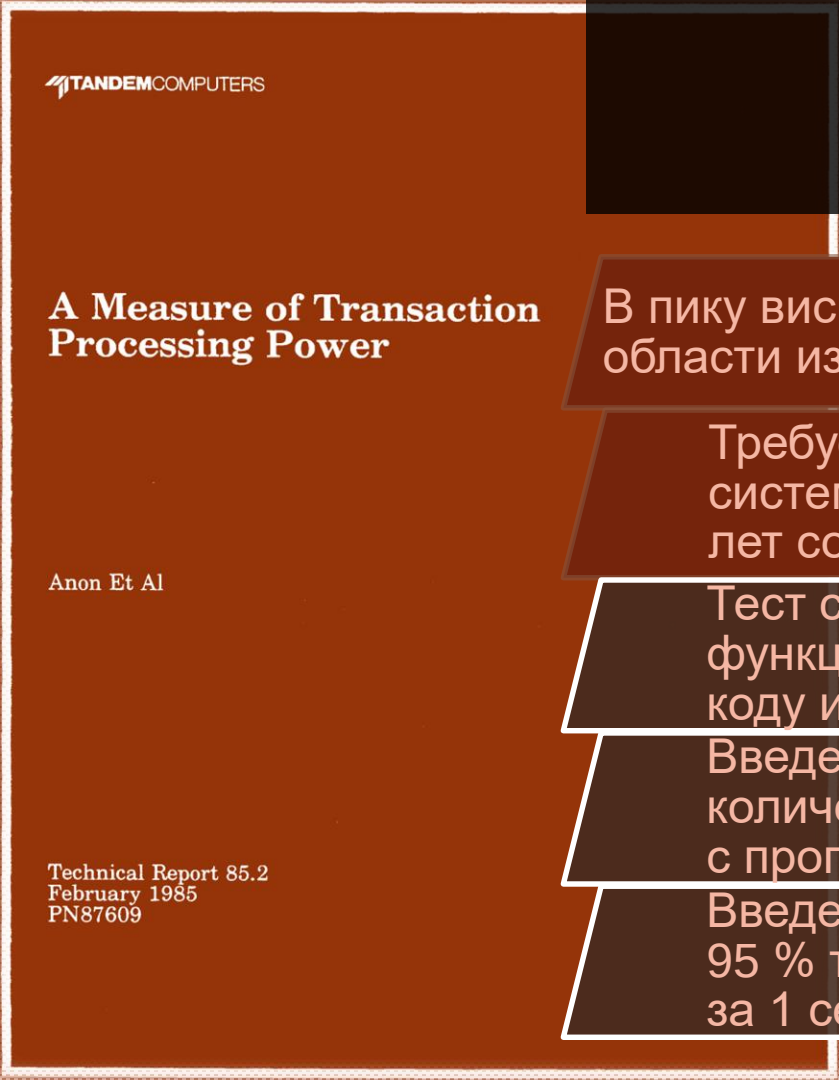
Альтернативный, более строго детерминированный бенчмарк

Лицензия на СУБД запрещает публикацию каких-либо тестов производительности СУБД

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

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

Джим Грей: DebitCredit



В пику висконсинскому бенчмарку, на базе предметной области из TP1

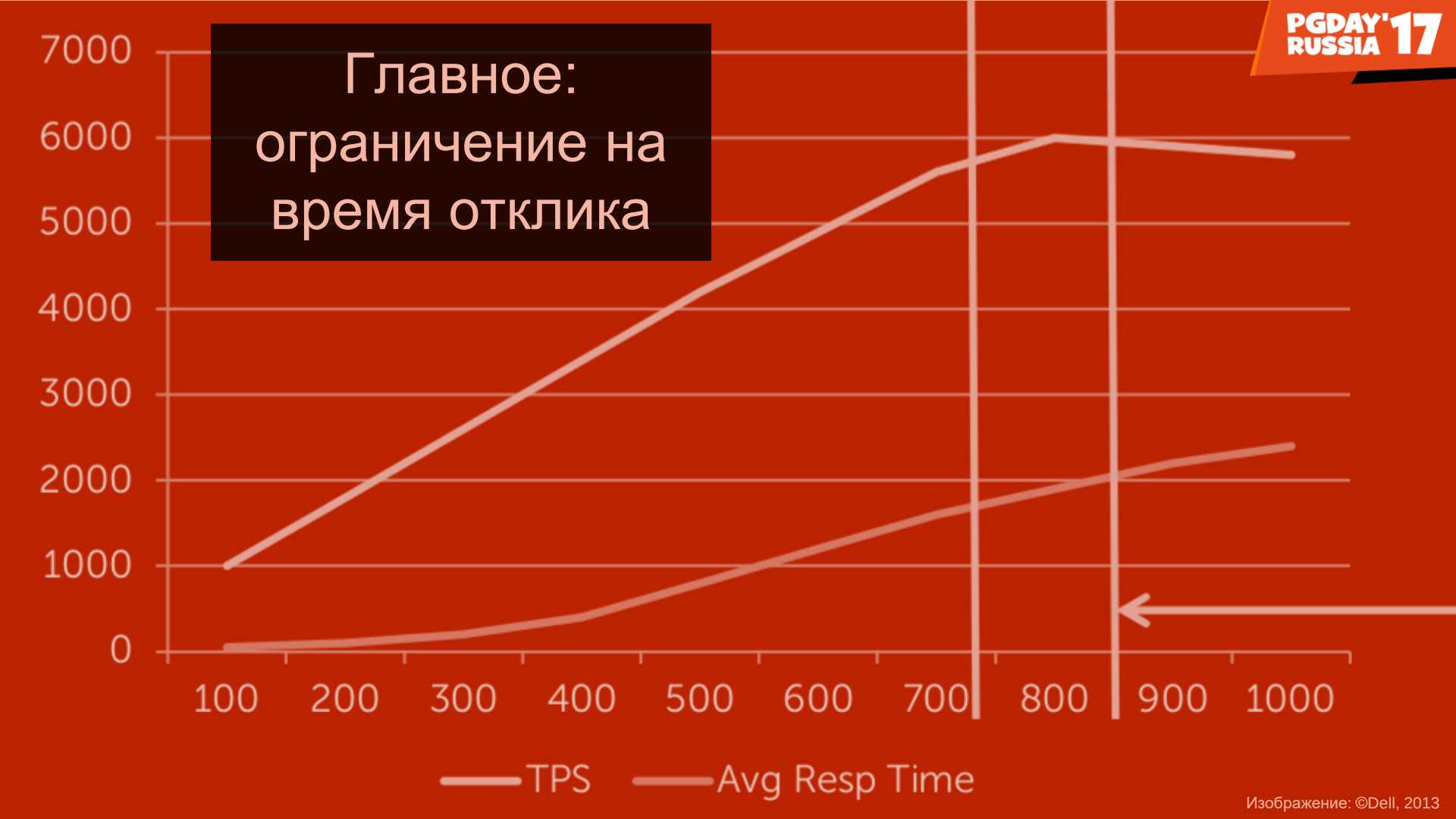
Требуется публикация суммарной стоимости системы, включая оборудование, лицензии, 5 лет сопровождения

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



Изображение: ©Dell, 2013

ЭПОХА СОВЕТА

TPC.org

PGDAY' RUSSIA 17

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

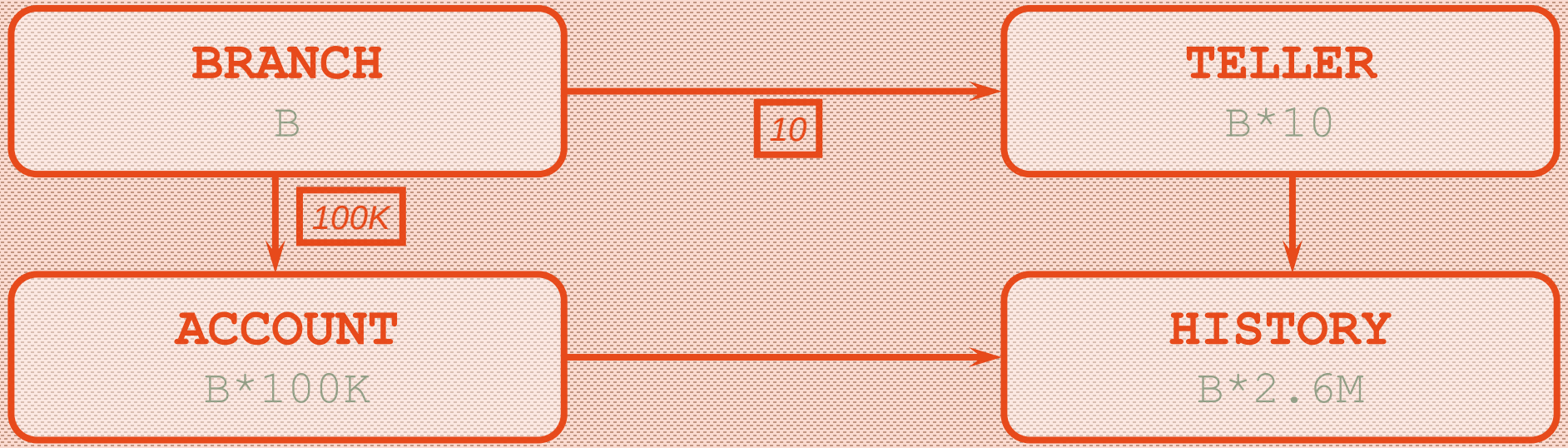
TPC-A и TPC-B: тест

PGDAY'17 RUSSIA

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



Цикл в 10 с 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

PGDAY RUSSIA '17

Наследие TP1

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

PGDAY' RUSSIA 17

5 типов транзакций

NEW-ORDER	PAYMENT	DELIVERY	ORDER-STATUS	STOCK-LEVEL
• новая заявка от клиента	• факт платежа • обновление баланса клиента	• заявка на доставку • *(пакетный процесс)*	• проверка статуса последней заявки от клиента	• проверка уровня запасов на складе
45%	43%	4%	4%	4%

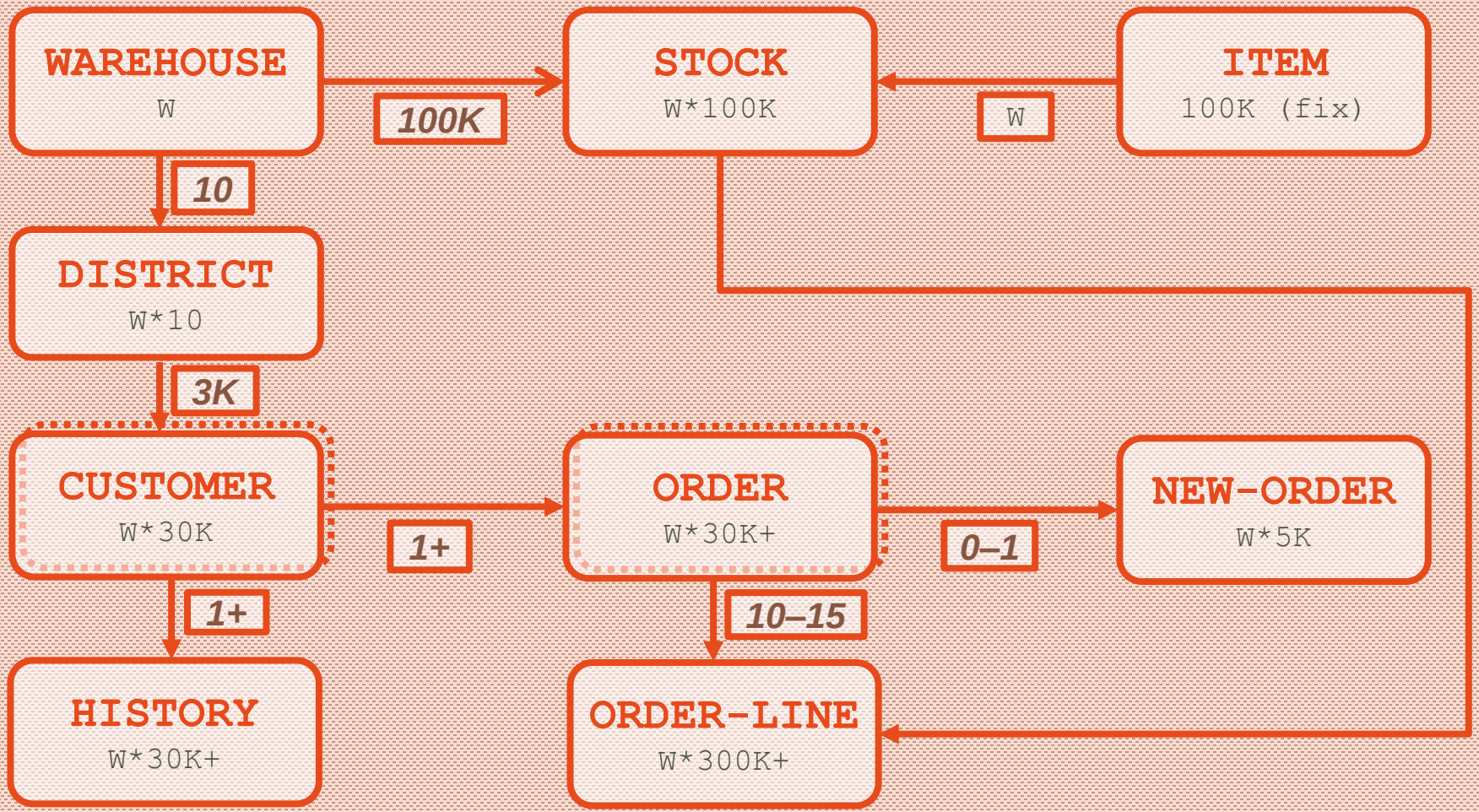
Больше вариаций

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

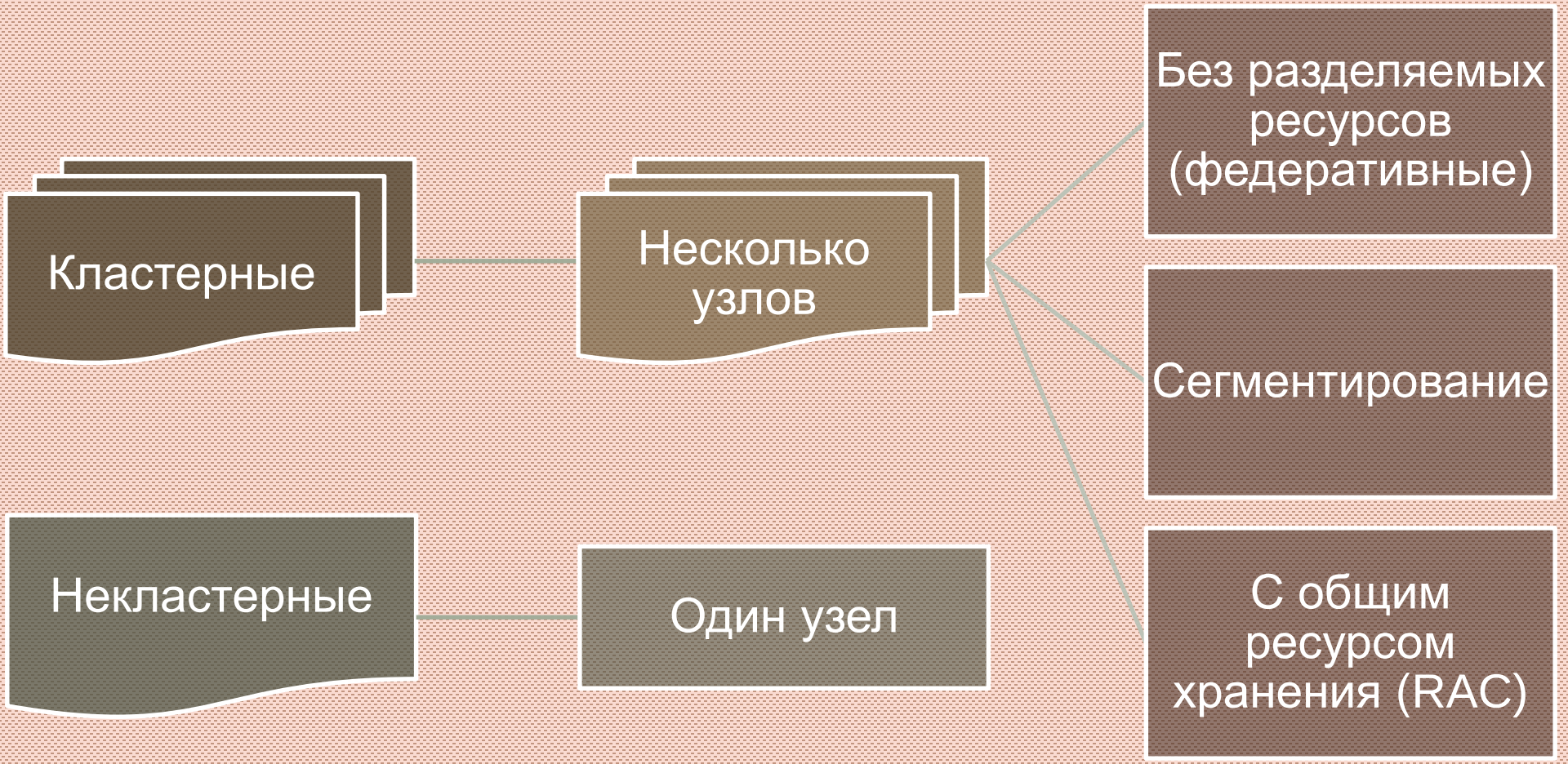
↓

Максимум – 1,2 tpmC на терминал

↓

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

TPC-C: кластерные и некластерные



TPC-C: метрики

PGDAY'17 RUSSIA

tpmC

Транзакций в минуту

tpmC / $

Стоимость транзакции

Полная стоимость оборудования

Лицензии и поддержка на 3 года

W / ktpmC

Энергозатраты на [кило] транзакцию

Интерпретация TPC-C

PGDAY'17 RUSSIA

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

tpmC × 2	• …не только создание заявок
tpmC / 2	• …если нет монитора транзакций
tpmC / 3	• …если на стороне оператора клиент типа Oracle Forms
tpmC × 2 / 3	• …если используются легковесные экранные формы (*curses* и т. п.)
tpmC / 2	• …если в приложении не было тонкой настройки SQL
tpmC / 2	• …если параллельно нагрузке формируются отчёты или выполняются пакетные задания

Просто ли обмануть TPC-C?

PGDAY'17 RUSSIA

```
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

PGDAY'17 RUSSIA

Фиксация транзакции = записано на энергонезависимый 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

TPC-C: критика 1990-х

PGDAY RUSSIA '17

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

PGDAY' RUSSIA 17

TPC-D	1995–1999	Первая попытка OLAP-теста
TPC-R	2001–2005	Отчётность
TPC-W	2001–2005	Онлайн-веб-коммерция

TPC-H

PGDAY' RUSSIA 17

1999 год: OLAP vs OLTP

“Ad-hoc decision support”

Вместо признанного неудачным TPC-D

«Весовые категории»

100 ГБ

300 ГБ

1 ТБ

3 ТБ

10 ТБ

30 ТБ

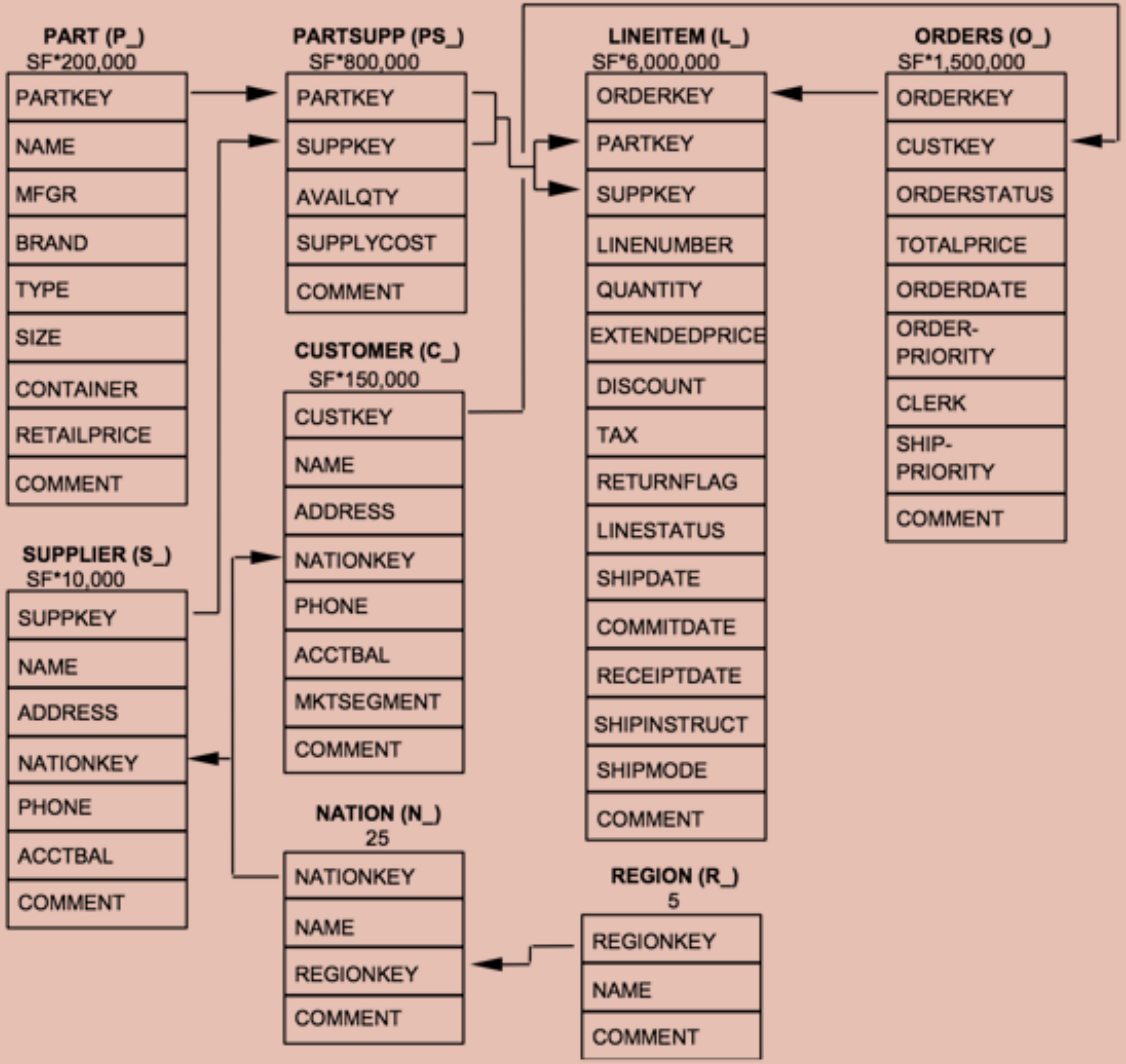
100 ТБ

Параллельная нагрузка

22 вида сложных запросов

2 типа обновления хранилища

TPC-H: модель



1 SF ~ 1 GB

Ограничения целостности

Не звезда

TPC-H Q1

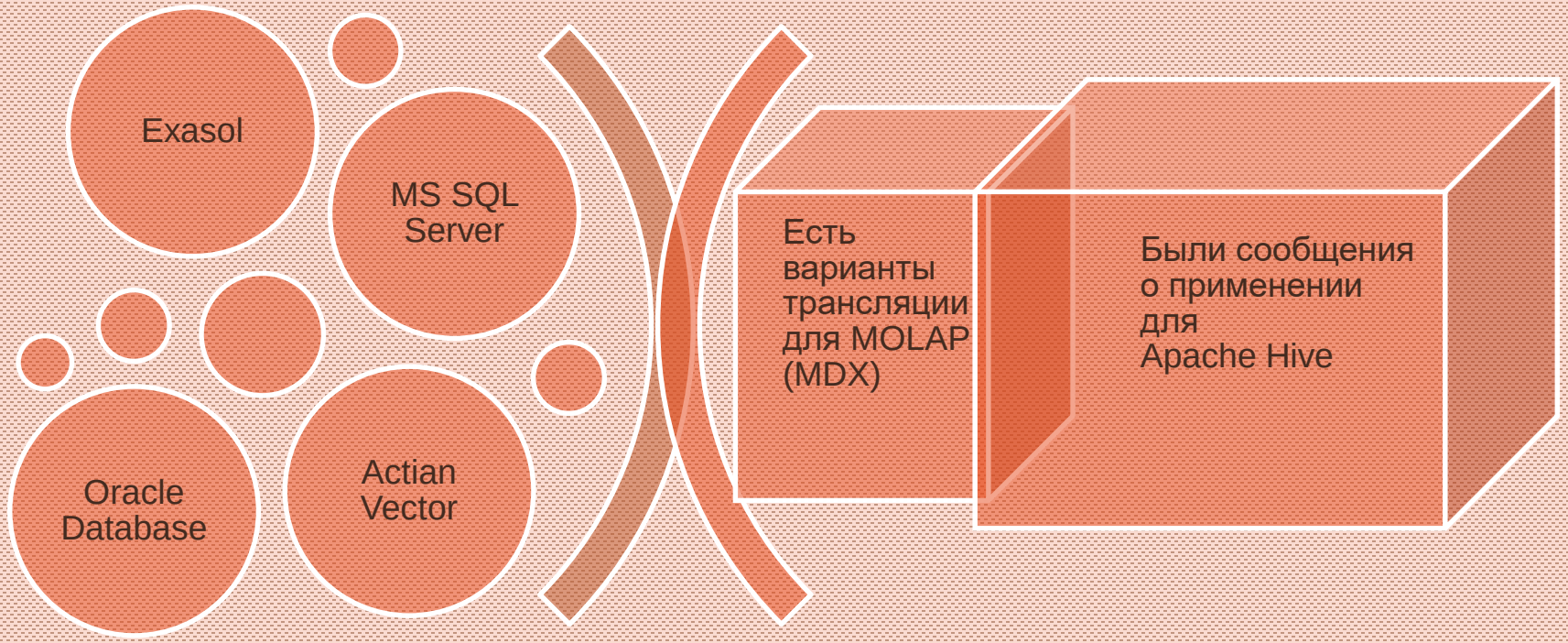
PGDAY'17 RUSSIA

“Functional Query Definition”

```
SELECT
   l_returnflag,  l_linestatus,
   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_linestatus
 ORDER BY l_returnflag,  l_linestatus
```

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»

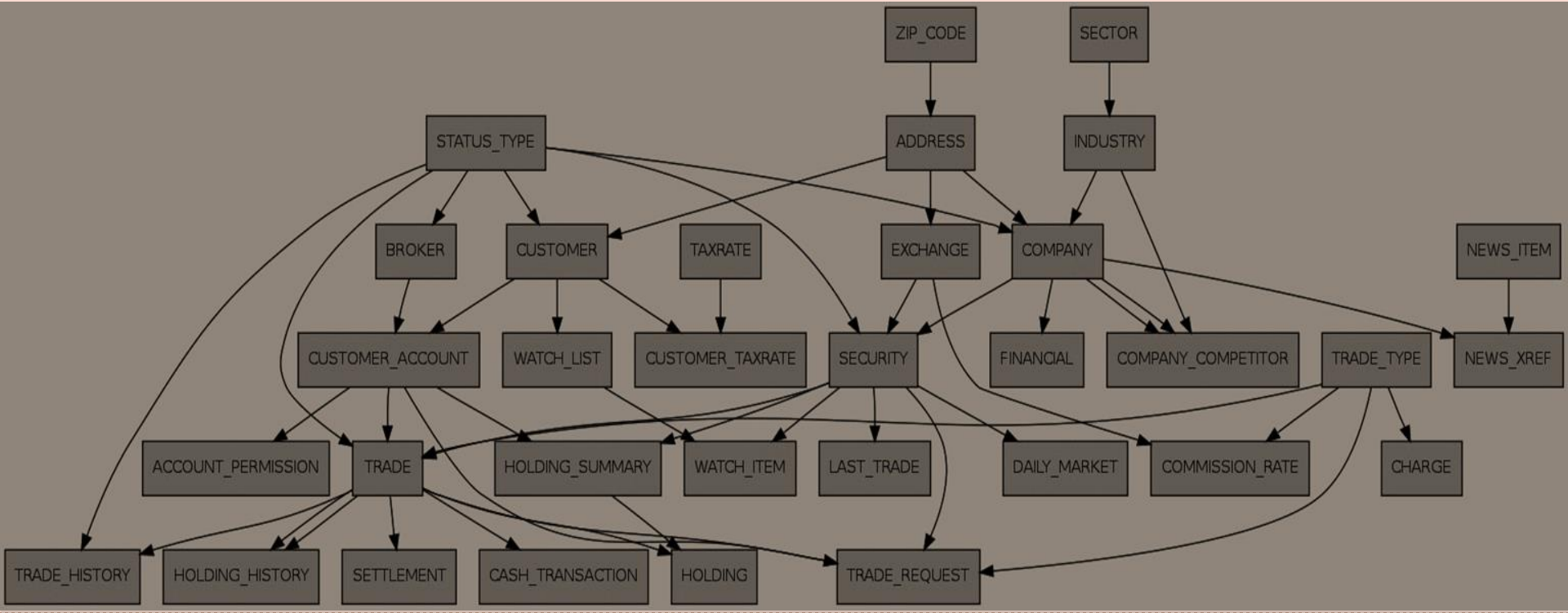
- Тоже OLTP, но более «гибридный»
- Включённые ограничения целостности
- Без мониторов транзакций
- Больше чтений
- Больше видов нагрузок

TPC-E: сложный поток операций

PGDAY' RUSSIA 17

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



Изображение: ©Transaction Processing Council, 2009

PGDAY'17 RUSSIA

TPC

TPC-E - All Results - Sorted by Performance
Version 1 Results As of 26-Nov-2016 at 6:02 PM [GMT]

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.

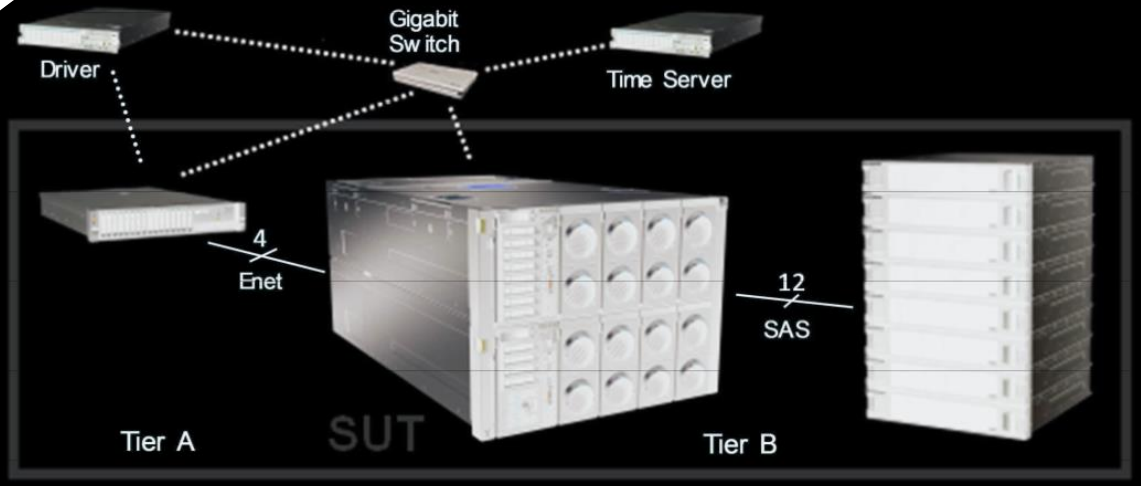
Hardware Vendor	System	v Performance (tpsE)	Price/tpsE	Watts/tpsE	System Availability	Database	Operating System	Date Submitted
Lenovo	Lenovo System x3950 X6	11,059	143.91 USD	NR	12/17/15	Microsoft SQL Server 2014 Enterprise Edition	Microsoft Windows Server 2012 R2 Standard Edition	12/17/15
FUJITSU	FUJITSU 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

78 результатов

Все – некластерные, MS SQL Server на Windows Server x64

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.



System x3650 M5, with:
- 2 x Intel® Xeon® Processor E5-2699 v3 2.30GHz (2 Procs/36 Cores/72 Threads)
- 64GB Memory
- 2 x 300GB SFF SAS (RAID-1)
- 1 x ServeRAID M5210
- Onboard Quad Gb Ethernet
- 2 x Dual 10Gb-T Ethernet

System x3950 X6, with:
- 8 x Intel Xeon Processor E7-8890 v3 2.50GHz (8 Procs/144 Cores/288 Threads)
- 4096GB Memory
- 2 x 900GB 10K SAS (RAID-1)
- 6 x 800GB SAS SSD (RAID-10)
- 1 x ServeRAID M5210
- 8 x ServeRAID M5225
- 4 x Dual 10Gb-T Ethernet

12 x Lenovo E1024 JBOD Enclosures, with:
- 104 x 800GB 2.5" SAS SSD (6 x 17-drive RAID-5) (1 x 2-drive RAID-1)

104 Total External Drives

Initial Database Size	Redundancy Level: 1	Storage
47,222 GB	**RAID-10 Log** **RAID-5 Data** **RAID-1 tempdb**	**2 x 900GB 2.5" 10K SAS** **6 x 800GB 2.5" SAS SSD** **104 x 800GB 2.5" SAS S**

11 ktpse,
$144/tpse
(1,6 MIO)

Изображение: ©Lenovo, © TPC, 2015

TPC-ABCEH

PGDAY'17 RUSSIA

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

Новые TPC-x и TPCx-X

PGDAY' RUSSIA 17

TPC-DI

ETL-нагрузки

TPC_DI_RPS

0 результатов

TPC-DS

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

QphDS@Size

0 результатов

TPCx-BB

Экспресс-бенчмарк для аналитических больших данных на базе BigBench

BBpm@Size

0 результатов

TPCx-HS

Экспресс-бенчмарк для HDFS-совместимых систем по типу TeraSort

HSph@Size

От 1 до 4 результатов в разных «весовых категориях»

Публикация?

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

Публикация на TPC.org

Обязательный аудит – sizing.com

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?

PGDAY'17 RUSSIA

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

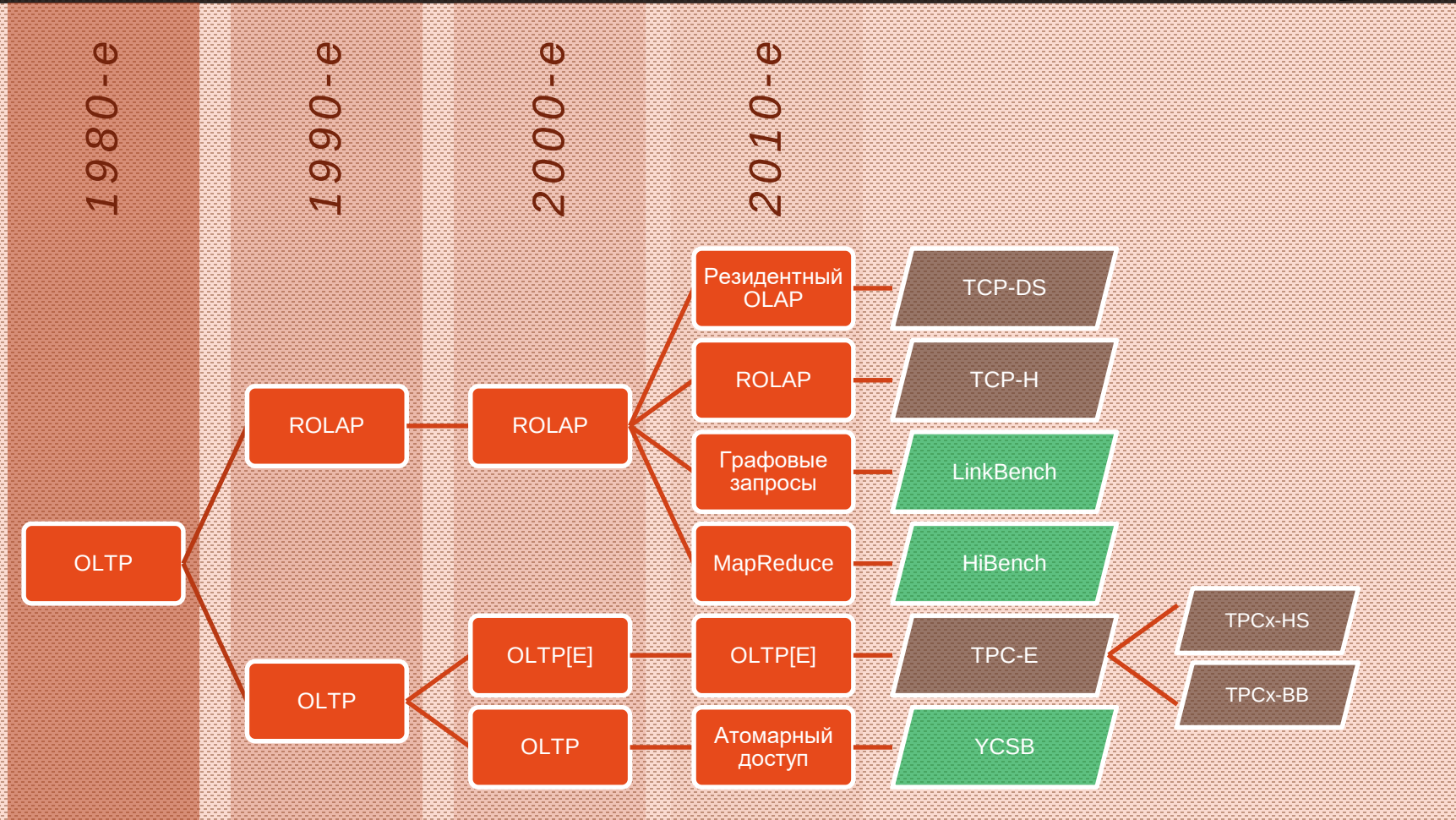
- *“Licensee may disclose the results of any benchmark test of the Program or its subcomponents to any third party provided that Licensee, if* … A) … B) … C)…”

PGDAY'17 RUSSIA

КОНФЕРЕНЦИЯ ПО БАЗАМ ДАННЫХ

НОВАЯ ВОЛНА

Новым нагрузкам – новые бенчмарки



Основные бенчмарки «новой волны»

PGDAY'17 RUSSIA

Terrasort Benchmark	• *Воплощён в TPCx-HS*
BigBench	• *Воплощён в TPCx-BB*
Intel HiBench	• *Серия тестов для Hadoop*
Yahoo! Cloud Services Benchmark	• *Набор «эталонных обстрелов» для NoSQL-систем*
Linkbench	• *Графообразная нагрузка на РСУБД*

Linkbench

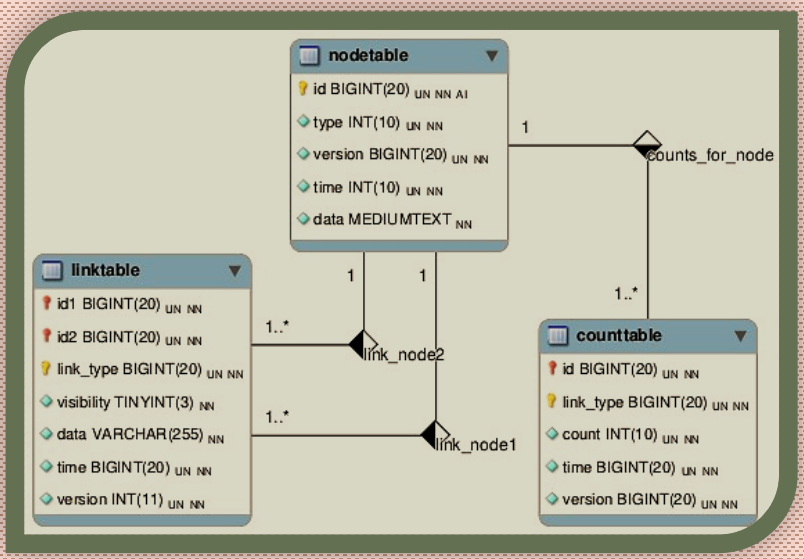
PGDAY'17 RUSSIA

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 для каждой из нагрузок
- 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

Intel Hibench для Hadoop

Workload	System Resource Utilization	Data Access Patterns	Map/Reduce Stage Time Ra
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 — Map → Red →	
Bayesian Classification (1st & 2nd job)	I/O bound, with high CPU utilization in map stage in the 1st job	— Map ⟫ Red → → Map → Red —	
K-means Clustering	CPU bound in iteration; I/O bound in clustering	→ Map — Red — → Map → ⊘ no reducer	
Enhanced DFSIO	I/O-bound	trivial	trivial

→ data → less data — even less data — compressed

Изображение: Intel, 2013

Yahoo! Cloud services benchmark

PGDAY'17 RUSSIA

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	Scalaris	Tarantool	Voldemort

Обстрел с одной нагрузочной станции

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

YCSB: нагрузки

PGDAY'17 RUSSIA

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

Пример «бенчмарктенинга» в новом классе

PGDAY'17 RUSSIA

The Register®
Biting the hand that feeds IT

DATA CENTRE | SOFTWARE | SECURITY | TRANSFORMATION | DEVOPS | BUSINESS | PERSONAL TECH | SCIENCE | EMERGENT TECH | BOOTNOTES

Data Centre ▸ Storage

1

Storage bunfight: Hazelcast says Gridgain faked benchmark result

In-memory startup denies Hazelcast's alleged porkies

By Chris Mellor 22 Feb 2016 at 15:32

SHARE ▼

In-memory software startup GridGain has denied an allegation by Hazelcast that it committed benchmark fraud.

This comes after GridGain published benchmark data showing it is faster than Hazelcast, which is one of its competitors.

Most read

PGDAY'17 RUSSIA

КОНФЕРЕНЦИЯ ПО БАЗАМ ДАННЫХ

ЗАМЕР «КАРМАННЫМИ СРЕДСТВАМИ»

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 – 4K или 8K?

Sysbench

PGDAY'17 RUSSIA

Утилита MySQL (автор – Алексей Копытов)

fileio	cpu	memory	threads	mutex	oltp

Своеобразные тесты, не похожие на какие-либо стандартные бенчмарки

Широко используется для внутренних сравнений MySQL, MariaDB и их форков

TPCE-MySQL

PGDAY' RUSSIA 17

TPC-E-нагрузчик для MySQL от Percona

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

Shell

```
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

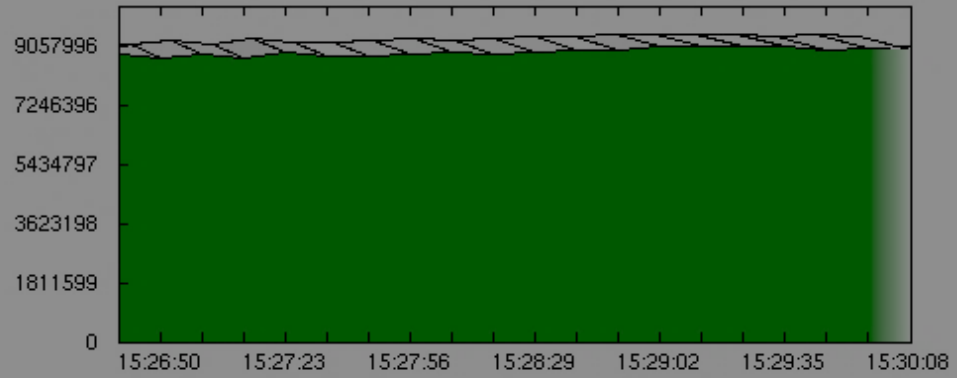
File Edit Options Help

RUNNING - TPC-C

Benchmark
- DB2
  - TPC-C
    - Schema Build
    - Driver Script
      - Options
      - Load
    - Virtual User
    - Autopilot
    - Transactions
    - Metrics
    - Mode
- Oracle
- SQL Server

Script Editor | Virtual User Output | Transaction Counter | Metrics | Autopilot

9057996 tpm



Virtual User	Iterations	Complete	Status
1	1	0	
2	1	0	
3	1	0	
4	1	0	
5	1	0	
6	1	0	
7	1	0	

HammerDB: TPC, но не совсем…

PGDAY' RUSSIA 17

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

Quest Benchmark factory for databases

PGDAY' RUSSIA 17

Коммерческий инструмент

Oracle Database

MySQL

MS SQL Server

SQLite

SQL Anywhere

Поддерживает архитектуру с несколькими нагрузочными станциями (Windows)

TPC-C ?

TPC-D

TPC-E

TPC-H

ASP3AP

osdldbt.sourceforge.net

PGDAY' RUSSIA 17

MySQL

PostgreSQL

…расширяемо

While the inspiration for these workloads are the TPC-<*x*…>, 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

PGDAY' RUSSIA 17

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
SIBench (Snapshot Isolation)	SmallBank	LinkBench			

TPC tools

Средства от Совета…

Исходные тексты на Си

Вся соединяемость и прочая работа – своими руками

Нет для TPC-C: только «сэмпл» в спецификации

TPC-* By Example

PGDAY' RUSSIA '17

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

```
#define
OBNDRAA(lda,cursor,sqlvar,progv,progvl,ftype,indp,alen,arcode,ms,cs)\
   if (obndra((cursor),(text*)(sqlvar),NA,(ub1*)(progv),(progvl),(ftype),NA,\
      (indp),(alen),(arcode),(ub4)(ms),(ub4*)(cs),(text*)0,NA,NA))\
      {errrpt(lda,cursor,sqlvar);return(-1);}\
   else\
      DISCARD 0

#define
ODEFIN(lda,cursor,pos,buf,bufl,ftype,scale,indp,fmt,fmtl,fmtt,rlen,rcode)\
   if (odefin((cursor),(pos),(ub1*)(buf),(bufl),(ftype),(scale),(indp),\
      (text*)(fmt),(fmtl),(fmtt),(ub4*)(rlen),(rcode))\
      {errrpt(lda,cursor,(text *)f
   else\
      DISCARD 0

#define OEXFET(lda,cursor,nrows,cancel,exact)\
   if (oexfet((cursor),(nrows),(cancel),(exact)))\
      {if ((cursor)->rc == 1403) DISCARD 0; \
       else if (errrpt(lda,cursor,(text *)"OEXFET")==RECOVERR) \
         {orol(lda);return(RECOVERR);} \
       else {orol(lda);return(-1);}}\
   else\
      DISCARD 0

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

#define OPARSE(lda,cursor,sqlstm,sqll,defflg,lngflg)\
   if (oparse((cursor),(sqlstm),(sb4)(sqll),(defflg),(ub4)(lngflg)))\
      {errrpt(lda,cursor,sqlstm);return(-1);}\
   else\
```

```
      DISCARD 0

#define OEXN(lda,cursor,iters,rowoff)\
   if (oexn((cursor),(iters),(rowoff))) \
      {if (errrpt(lda,cursor,(text *)"OEXN")==RECOVERR) \
         {orol(lda);return(RECOVERR);} \
       else {orol(lda);return(-1);}}\
   else\
      DISCARD 0

#endif
/* additions done for 814 -shishir */
#define OCI_ATTR_SRVRCTXT          OCI_ATTR_SERVER
#define OCI_ATTR_USERCTXT          OCI_ATTR_USER
#define OCI_ATTR_ROWCNT            OCI_ATTR_ROW_COUNT
#define OCI_HTYPE_ERR              OCI_HTYPE_ERROR
#define OCI_HTYPE_STM              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
```

```
#endif /* The bind and define handles for each transaction are included in
their respective header files. */ extern OCIEnv *tpcenv; extern OCIServer
*tpcsrv; extern OCIError *errhp; extern OCISvcCtx *tpcsvc; extern
OCISession *tpcusr; extern OCIStmt *curi; extern char *uid; extern char
*pwd; /* for stock-level transaction */ extern int w_id; extern int d_id; extern
int c_id; extern int threshold; extern int low_stock; /* for delivery transaction
*/ extern int del_o_id[10]; extern int carrier_id; extern double qtime; extern
int retries; extern int retries_serial; extern int retries_recoverr; extern int
retries_snapshot; /* for payment transaction */ extern int hvlastrow;
extern char c_last[17]; extern char c_first[17]; extern char c_middle[3];
extern double c_balance; extern int o_id; extern text o_entry_d[20]; extern int
carrier_id; extern int o_ol_cnt; extern int o_all_local; extern int supply_w_id[15]; extern int
ol_amount[15]; ub4
ol_del_len[15]; extern text ol_delivery_d[15][11]; /* xmie - begin */ extern
OCIRowid *o_rowid; /* xmie - end */ /* for payment transaction */ extern int
c_id; extern int 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_quanti10[15]; extern int nol_quanti91[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 status; extern int tracelevel; /* Miscellaneous */ extern OCIDate
cr_date; extern OCIDate c_since; extern OCIDate o_entry_d_base; extern
```

Benchware

PGDAY'17 RUSSIA

«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_$%`

Прогон реальной нагрузки и бенчмарки приложений

PGDAY'17 RUSSIA

КОНФЕРЕНЦИЯ ПО БАЗАМ ДАННЫХ

Самый практичный метод?

Воспроизведение нагрузки: точный опыт

PGDAY'17 RUSSIA

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»

PGDAY' RUSSIA 17

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

PGDAY'17 RUSSIA

SAPS [SAP Application Performance Standard] на модуле SD

Oracle E-Business Suite Standard Benchamrks

Order-to-Cash	OLTP	Payroll	…

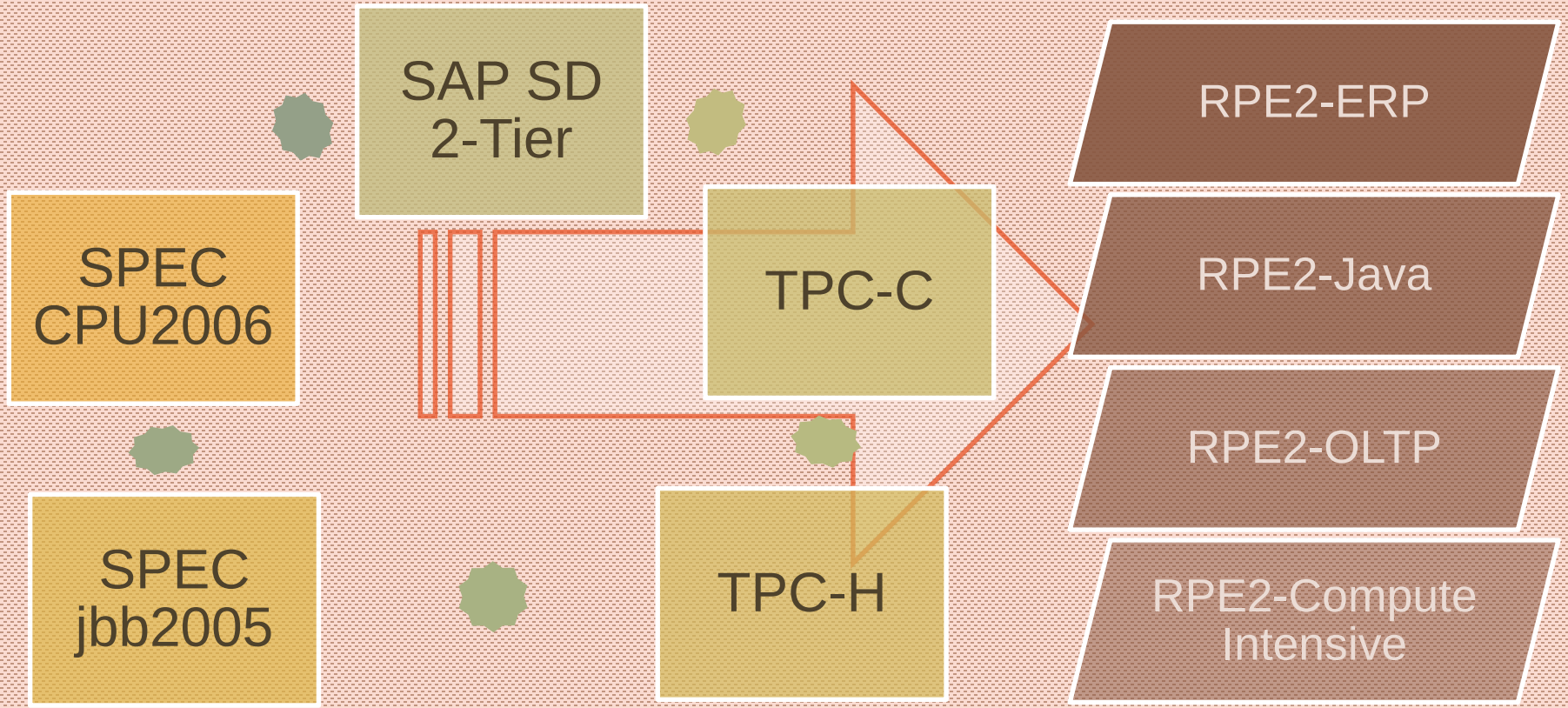
Microsoft Dynamics AX Application Benchmark Toolkit

1С: типовая конфигурация
«Стандартный нагрузочный тест»

Показатель – Apdex (число одновременно работающих пользователей при заданном максимальном времени отклика)

RPE2

Суперкомпозитный показатель от *Gartner* (*Ideas*)



Атомизация показателей

Машины баз данных?

PGDAY'17 RUSSIA

Предконфигурированные аппаратно-программные комплексы для СУБД

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

Teradata

Последние публикации Teradata с TPC-H:

Teradata a division of NCR	WorldMark 5250	18,542	638.37 USD	NR	07/27/01	Teradata V2R4.1	MP-RAS 3.02.00	10/09/01
Teradata a division of NCR	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

PGDAY'17 RUSSIA

Не публикуются tps, Qph…

«Паспортные показатели» (X6-8)

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

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

IBM Pure Data for Operational analytics

PGDAY' RUSSIA 17

Не публикуются Qph…

«Паспортные данные» – про ввод-вывод

	Extra Small	Small	Medium	Large
Пропускная способность, Гбай	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”

PGDAY' RUSSIA '17

IOPS со стороны Oracle Database

Orion	SLOB	Benchware	DBMS_RESOURCE_MANAGER .CALIBRATE_IO
(Oracle IO Numbers)	(рекомендуют EMC, Flashgrid)	(?)	

Получение SQL IOPS для других СУБД?

Статистические представления (…IO_STATS…)

Атомизация агрегатных показателей

tpmC → SQL IOPS

QpmH → SQL bandwidth

Атомизация показателей: Pro

PGDAY' RUSSIA 17

Независимость от заданных моделей и схем	• Показательны не только для 3NF, снежинок, звёзд
Признаются производителями машин баз данных	• Включены в «паспортные показатели»
Доступны для «карманной инструментовки»	• Могут быть запущены со стороны БД
Показательны для широкого класса СУБД и СУБД-подобных систем	• Могут быть интересны для NoSQL

Атомизация показателей: …et contra

PGDAY' RUSSIA 17

Вопросы к методологии

Результаты CALIBRATE_IO и Benchware – разные

Неочевидный прикладной смысл результата

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

Нет готового инструмента для большинства известных СУБД

Есть для Oracle DB, IBM DB2, MS SQL

Дискуссионные вопросы

PGDAY' RUSSIA 17

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

PGDAY'17 RUSSIA
КОНФЕРЕНЦИЯ ПО БАЗАМ ДАННЫХ

Спасибо за внимание!

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

IBS