

Benchmarking Transaction Processing Systems – Past, Present and Future

Trish Hogan, IBM Corporation

June 2008

© 2008 TPC

Overview of Transaction Processing

- Past
- Present
- Future

Past: Pre-TPC TP1 and DebitCredit

- TP1 – batch debit/credit, no network, no user interaction, no run rules
- DebitCredit – April 1985
 - Total System Cost + 5 years Maintenance
 - Scale up rules
 - 95% of transactions complete in 1 second
 - No validation that it was run correctly

Past: TPC-A

- Nov. 1989 TPC-A
 - Based on DebitCredit
 - 90% of transactions complete in 2 seconds
 - Added some ACID requirements
 - Publicly disclosed FDR
- July 1990 first TPC-A results
- ~ 300 TPC-A results published

Past: TPC-B

- Aug. 1990 TPC-B
 - TPC-A results just started being published
 - Based on TP1 so no end users at a terminal going over a network
 - Mid 1991 first TPC-B result
 - ~ 130 TPC-B results

Past and Present: TPC-C

- July 1992 TPC-C
 - Wholesale Supplier model
 - Started with terminals moved to html
 - Sep. 1992 first TPC-C result
 - Nearly 700 TPC-C results
 - Number of publications beginning to drop
 - Used for engineering work and big splash launches

Present

- End Users on Web, phone, appliance...
- Global access
- More complex transactions
- Bigger databases
- RAID protected
- Database constraints enforced

Present and Future: TPC-E

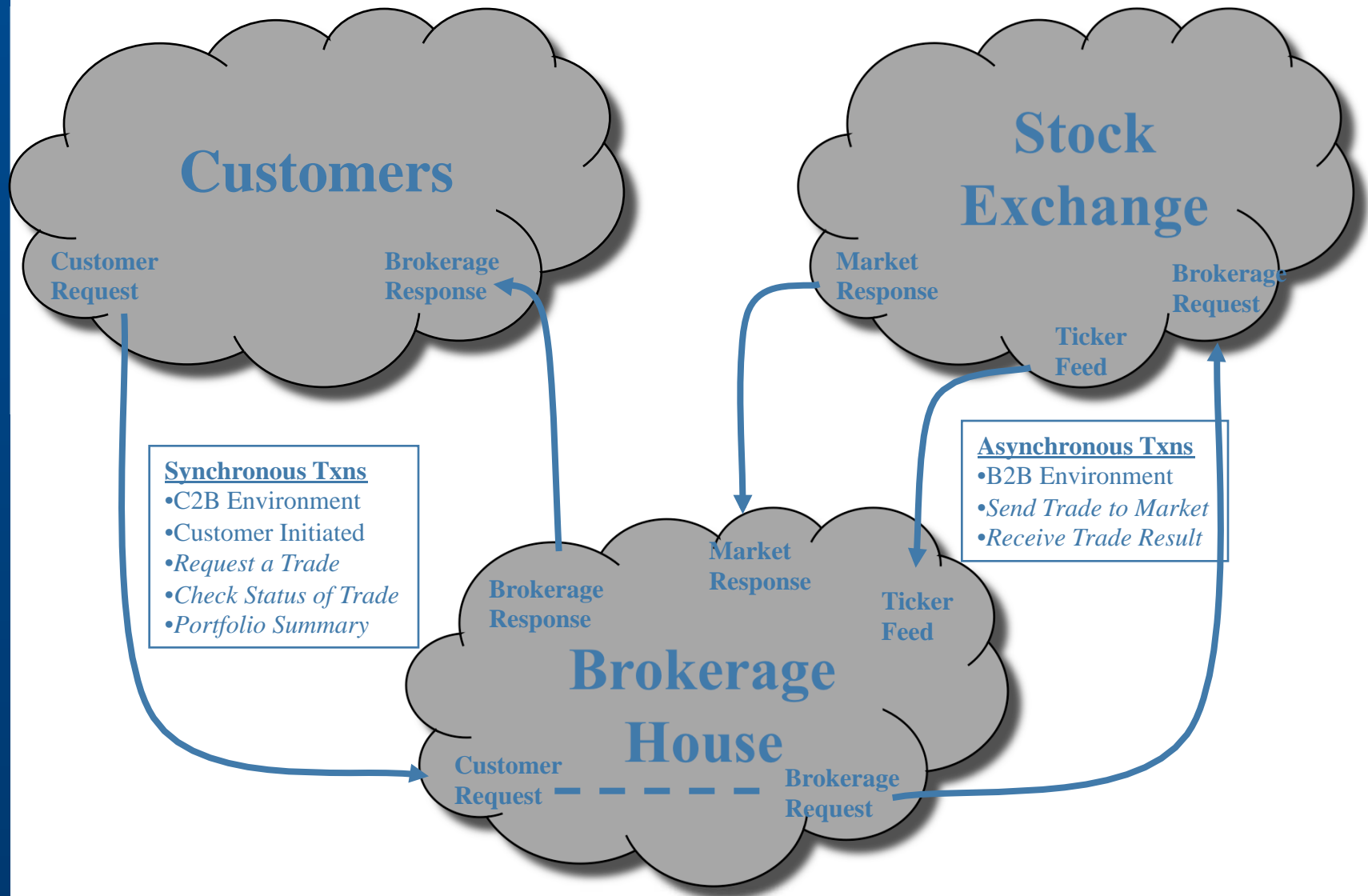
- Feb. 2007 TPC-E
 - Brokerage house model
 - OLTP Database-centric workload
 - Comparability of results (*)
 - Reduce cost/complexity of running benchmark
 - Enhance schema complexity
 - Jul. 2007 first TPC-E result

(*) TPC-E results are intended for OLTP database testing and are not an indicator for past or future market performance, nor should they be used as such.

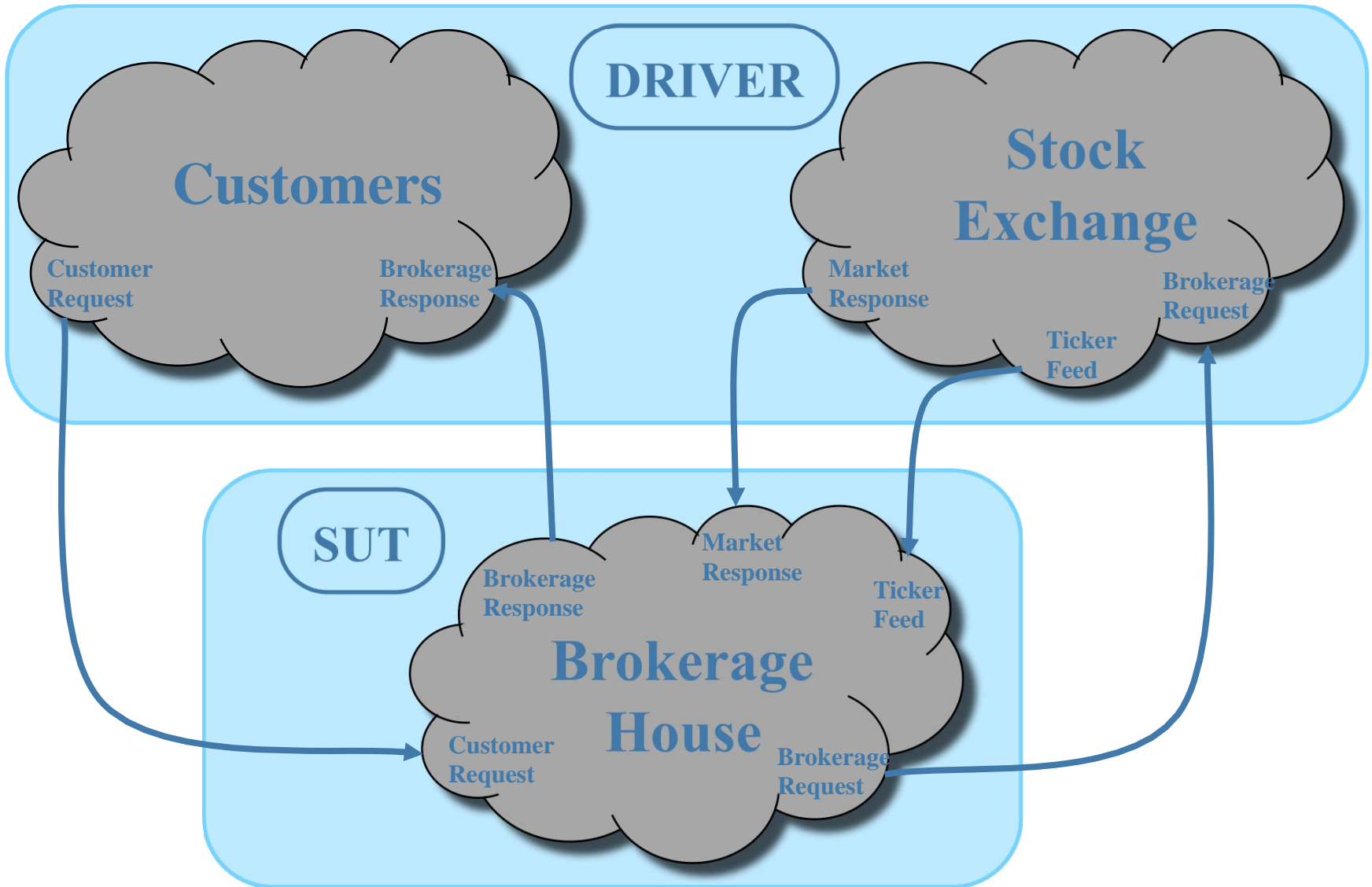
Present and Future

Characteristics	TPC-C	TPC-E
Business model	Wholesale supplier	Brokerage house
Number of database tables	9	33
Number of database columns	92	188
Minimum columns per table	3	2
Maximum columns per table	21	24
Datatype count	4	Many
Primary keys	8	33
Foreign keys	9	50
Tables with foreign keys	7	27
Check constraints	0	22
Referential integrity	No	Yes
Database roundtrips per transaction	One	One or many
Number of transactions	5	10
Number of physical I/Os	5x	x
RAID requirements	Database log only	Everything
Timed database recovery	No	Yes
TPC-provided code	No	Yes

TPC-E Business Model – Financial Market



TPC-E Business Model – Financial Market



TPC-E Database Scaling

Customer	Broker	Market	
ACCOUNT_PERMISSION	BROKER	COMPANY	SECTOR
CUSTOMER	CASH_TRANSACTION	COMPANY_COMPETITOR	SECURITY
CUSTOMER_ACCOUNT	CHARGE	DAILY_MARKET	
CUSTOMER_TAXRATE	COMMISSION_RATE	EXCHANGE	
HOLDING	SETTLEMENT	FINANCIAL	Dimension
HOLDING_HISTORY	TRADE	INDUSTRY	ADDRESS
HOLDING_SUMMARY	TRADE_HISTORY	LAST_TRADE	STATUS_TYPE
WATCH_ITEM	TRADE_REQUEST	NEWS_ITEM	TAXRATE
WATCH_LIST	TRADE_TYPE	NEWS_XREF	ZIP_CODE
Legend:			
	Fixed Tables	Growing Tables	Scaling Tables

Database – Content

- Populated with pseudo-real data
- Distributions based on:
 - 2000 U.S. and Canada census data (*)
 - Used for generating name, address, gender, etc.
 - Introduces natural data skew
 - Actual listings on the NYSE and NASDAQ
- Benefits
 - Realistic looking data
 - Compressible for backup testing, etc.
 - Closer match to actual customer databases
 - Anticipate usage well beyond benchmark

(*) only names of 2000 census have been used – all other data are fictional and any similarities are purely coincidental

Database – Content

- Sample data from TPC-C CUSTOMER table

C_FIRST	C_MIDDLE	C_LAST	C_STREET_1	C_CITY
RONpTGcv5ZBZO8Q	OE	BARBARABLE	bR7QLfDBhZPHlyDXs	OmWlmezIJ0GeP kYM
e8u6FMxFLtt6p Q	OE	BARBARPRI	eEbgKxoIzx99ZTD S	4V1t1VmdVcXyoTOMwpPz
bTUkSuVQGdXLjGe	OE	BARBARPRES	QCGLjWnsqSQPN D S	jVHBwIGFh2k oTOMwpPz
18AEf3ObueKvubUX	OE	BARBARESE	JnBSg4RtZbALYu S	5g8XMnlegn oTOMwpPz
mFFsJYeYE6AR bUX	OE	BARBARANTI	MLEwwdy3dXfqngFcE	yVVR4iEtj0ADEwe wpPz

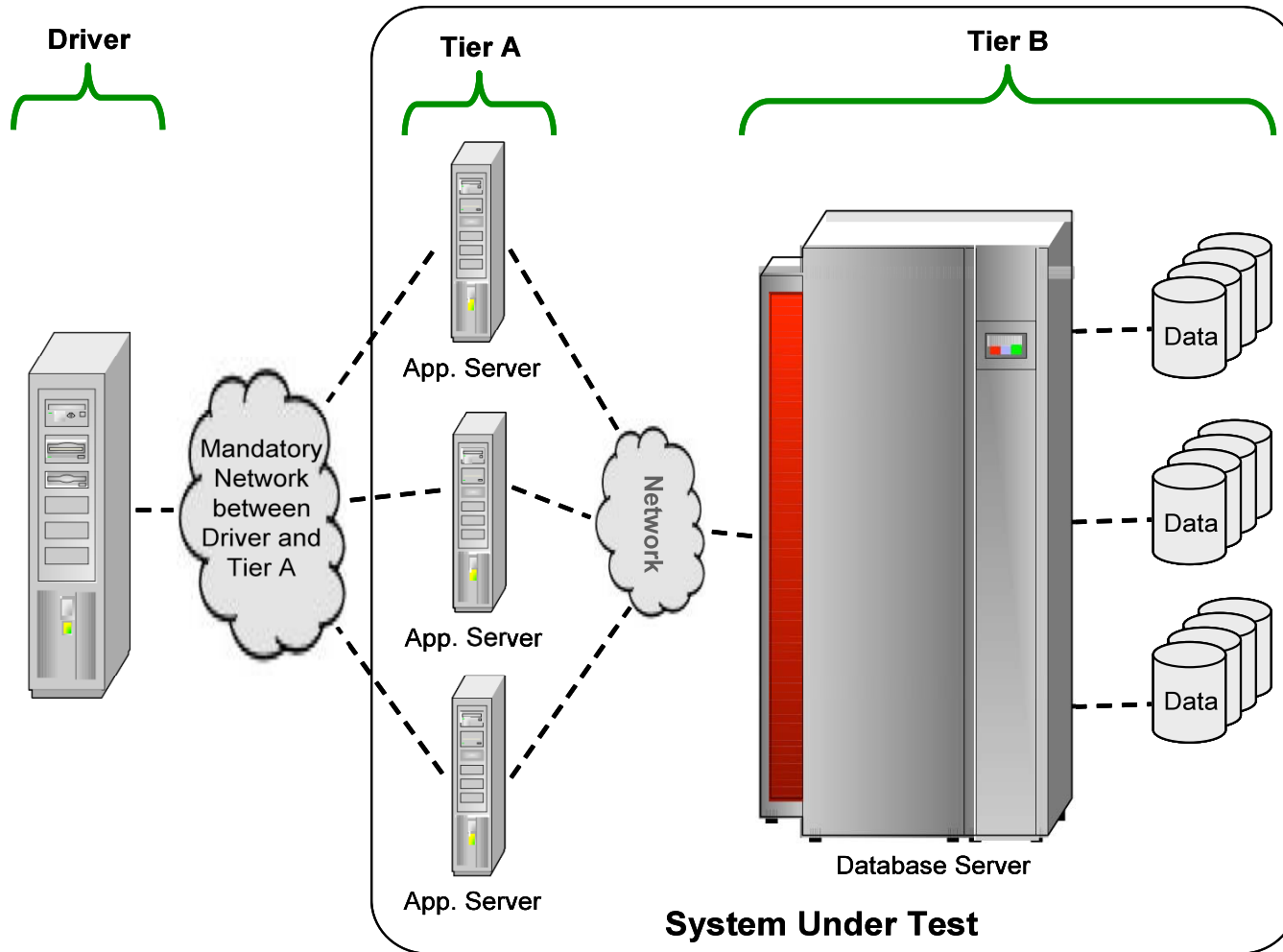
- Sample data from TPC-E CUSTOMER table

C_TAX_ID	C_L_NAME	C_F_NAME	C_M_NAME	C_GNDR	C_DOB	C_EMAIL_1
757FI2006HD923	Mexicano	Courtney	T	F	1997-11-30	CMexicano@hotmail.com
922SN3775RQ823	Udley	Judith	F	F	1954-09-27	JUdley@earthlink.com
006GT3444BE624	Buchanan	John	R	M	1971-06-13	JBuchanan@msn.com
181UZ4114LR434	Soloman	Clinton	D	M	1938-02-27	CSoloman@rr.com
355IE4773VF335	Orner	Harry	P	M	1974-11-15	Horner@attbi.com

Transactions - Overview

Name	Symbol	Access	Description
Broker-Volume	BV	RO	DSS-type medium query
Customer-Position	CP	RO	“What am I worth?”
Market-Feed	MF	RW	Processing of Stock Ticker
Market-Watch	MW	RO	“What’s the market doing?”
Security-Detail	SD	RO	Details about a security
Trade-Lookup	TL	RO	Look up historical trade info
Trade-Order	TO	RW	Enter a stock trade
Trade-Result	TR	RW	Completion of a stock trade
Trade-Status	TS	RO	Check status of trade order
Trade-Update	TU	RW	Correct historical trade info

Sample Implementation



TPC-E Summary - Highlights

- Financial business model
- Rich transaction set
- Major components provided
- Diverse, realistic schema
 - Extensive use of non-primary key access
 - Foreign key relationships
 - Complex DML
 - Referential integrity
- Availability requirements
 - Storage media must be fault tolerant
 - Practical implication: RAID-1 or RAID-5

TPC-E Summary - Benefits

- Server-centric workload with strong DB focus
- Broader coverage: database functions, schema, features
- Realistic application transaction model
- Ease of benchmarking: quicker startup, lower cost
- Reduced I/O requirements
- Comparability of results
- Realistic database schema, population, and transactions
- Specification provides code where sponsor creativity not being tested

Trademarks and Copyright

- TPC, TPC Benchmark, TPC-E and TPC-C are trademarks of the Transaction Processing Performance Council. All other materials are © 2008 TPC. All rights reserved.