

Measuring and Reducing Postgres Transaction Latency

Fabien Coelho

MINES ParisTech, PSL Research University

pgDay Paris – March 23, 2017



Postgres Latency

Talk Outline

2 Performance Comparisons

- Two Connection Costs
- Latency Pitfalls
- Benchmarking with Rate and Limit
- Three Storage Options
- Two Protocol Impacts
- Four Query Combination Tricks
- Reducting Server Distance
- Performance Scalability
- Miscellaneous Settings

Introduction

- Subject
- Typical Web Application
- Transaction Performance Definitions
- pgbench

3 Conclusion

- Latency and Throughput Wrap-Up
- Lessons Learned
- Contributions to Postgres



Postares Latency

Subject

Small OLTP	
 CRUD queries data fit in share RW, RO 	
Focus	

performance with emphasis on latency

shared buffers

- OnLine Transaction Processing
 - WHERE pk=? . . . small, few GB pabench builtins

Wrap-Up



Subject

Postgres	
Latency	

F. Coelho

ntroduction

- Subject
- Application Definitions pabench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

Small OLTP

- CRUD queries
- data fit in shared buffers
- RW, RO

Focus

OnLine Transaction Processing

... WHERE pk=? small, few GB pgbench builtins

and Motivation

performance with emphasis on latency

interactive web app



Subject

Small OLTP

Postgres	
Latency	

F. Coelho

Introduction

- Subject
- Application Definitions

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

CRUD	alleries
ONOD	quenes

- data fit in shared buffers
- RW, RO

Focus

OnLine Transaction Processing

... WHERE pk=? small, few GB pgbench builtins

and Motivation

- performance with emphasis on latency
- experiments & measures

interactive web app do not assume!



Subject

S

F. Coelho

Introduction

- Subject
- Application Definitions

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

mall OLTP	OnLine Transaction Processing
 CRUD queries data fit in shared buffers RW. RO 	WHERE pk=? small, few GE pabench builtins
, -	10
ocus	and Motivation

- performance with emphasis on latency
- experiments & measures

interactive web app do not assume!

latency performance : RW \times 63, RO \times 219



Typical Web Application

Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions Client user acts on user-agent, sends to Server process request, database operations to Database stores and retrieves data



Database Operations

Connection

3-Tier Architecture

TCP/IP, SSL & AAA

Request-Response cycles

transfer, parse, plan, execute, transfer back



Typical Web Application

Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions Client user acts on user-agent, sends to Server process request, database operations to Database stores and retrieves data



Database Operations

Connection

3-Tier Architecture

TCP/IP, SSL & AAA

Request-Response cycles

transfer, parse, plan, execute, transfer back



Postgres Latency	Definitions	time & operations
F. Coelho roduction bject plication finitions bench rformance	Throughputoperations per time unit usual approach, load measured in tpsLatencytime for one operation must fit application requirements	tx/s ms/tx
tency		
prage ptocol		
mbinations stance alability scellaneous onClusion rap-Up ssons intributions		



Postgres Latency	Definitions	time & operations
F. Coelho troduction ubject pplication terinitions gbench erformance ionnection atoncy	Throughputoperations per time unit usual approach, load measured in tpsLatencytime for one operation must fit application requirements	tx/s ms/tx
tate & Limit itorage irotocol	Comments	
combinations histance scalability	correlated	and contradictory
	max vs enough	and vice-versa
Vrap-Up essons	sensitive to many settings	net, soft & hard
ontributions	throughput bottleneck & latency additivity	



Postgres Latency	Definitions	time & operations
F. Coelho troduction ubject oplication efinitions upench	Throughput operations per time unit usual approach, load measured in tps Latency time for one operation	tx/s ms/tx
erformance onnection atency ate & Limit torage	must fit application requirements Comments	
oncolusion stance calability iscellaneous onclusion rap-Up	 correlated max vs enough sensitive to many settings 	and contradictory and vice-versa
assons ontributions	 throughput bottleneck & latency additivity 	not, son a nara



Postgres Latency	Definitions	time & operations
F. Coelho troduction ubject polication	Throughput operations per time unit usual approach, load measured in tps	tx/s
efinitions gbench erformance onnection atency	Latency time for one operation must fit application requirements	ms/tx
ate & Limit torage rotocol	Comments	
ombinations istance calability	correlated	and contradictory
iscellaneous onclusion	max vs enough	and vice-versa
Irap-Up essons	sensitive to many settings	net, soft & hard
ontributions	throughput bottleneck & latency additivity	



Postgres Latency	Definitions	time & operations
F. Coelho	Throughput operations per time unit usual approach, load measured in tps	tx/s
Definitions rgbench erformance Connection Latency	Latency time for one operation must fit application requirements	ms/tx
Rate & Limit Storage Protocol	Comments	
Combinations Distance Scalability Aliscellaneous Conclusion Vrap-Up Lessons Contributions	 correlated max vs enough sensitive to many settings throughput bottleneck & latency additivity 	and contradictory and vice-versa net, soft & hard deep voodoo!



Postgres Performance Swiss Army Knife



Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusior

Wrap-Up Lessons Contributions input SQL-like scripts with minimal client-side language options time to run, prepared, reconnections, ... parallelism threads, clients, asynchronous calls output statistical performance data

aveats

long enough

Available Features

- several times
- pedal-to-the-metal max speed test

warm-up, checkpoint and vacuum reproducibility not representative



Postgres Performance Swiss Army Knife



Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions input SQL-like scripts with minimal client-side language options time to run, prepared, reconnections, ...
 parallelism threads, clients, asynchronous calls output statistical performance data

Caveats

Iong enough

Available Features

- several times
- pedal-to-the-metal max speed test

warm-up, checkpoint and vacuum reproducibility not representative



Default TPC-B-like Transaction

pgbench -b tcpb-like

Postgres Latency

F. Coelho

ntroduction Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusior

Wrap-Up Lessons Contribution

TPC-B-like banking transaction

```
-- random ids and amount
\set aid random(1, 100000 * :scale)
 set bid random(1, 1 * :scale)
set tid random(1. 10 * : scale)
set delta random(-5000, 5000)
-- actual transaction
BEGIN;
UPDATE pgbench_accounts
  SET abalance = abalance + :delta WHERE aid = :aid:
SELECT abalance
  FROM pgbench_accounts WHERE aid = :aid:
UPDATE pgbench_tellers
  SET thalance = thalance + :delta WHERE tid = :tid:
UPDATE pgbench_branches
  SET bbalance = bbalance + :delta WHERE bid = :bid:
INSERT INTO pgbench_history (tid, bid, aid, delta, mtime)
  VALUES (:tid, :bid, :aid, :delta, CURRENT_TIMESTAMP);
END:
```

Pattern 3 updates 1 insert



Default TPC-B-like Transaction

pgbench -b tcpb-like

Postgres Latency

F. Coelho

ntroductio Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusio

Wrap-Up Lessons Contributior

TPC-B-like banking transaction

```
-- random ids and amount
\set aid random(1, 100000 * :scale)
 set bid random(1, 1 * :scale)
set tid random(1. 10 * : scale)
set delta random(-5000, 5000)
-- actual transaction
BEGIN;
UPDATE pgbench_accounts
  SET abalance = abalance + :delta WHERE aid = :aid:
SELECT abalance
  FROM pgbench_accounts WHERE aid = :aid:
UPDATE pgbench_tellers
  SET thalance = thalance + :delta WHERE tid = :tid:
UPDATE pgbench_branches
  SET bbalance = bbalance + :delta WHERE bid = :bid:
INSERT INTO pgbench_history (tid, bid, aid, delta, mtime)
  VALUES (:tid, :bid, :aid, :delta, CURRENT_TIMESTAMP);
END:
```

Pattern



- 1 insert
- 1 select



Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performanc

Connection

Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

Performance Comparisons

Two Connection Costs



pgbench -C



F. Coelho

pabench

Connection Rate & Limit Storage Miscellaneous

Wrap-Up

pgbench 	postgres →	ClientLANServer	8 cores, 16 GB 1 Gbps 16 cores, 32 GB, HDD
pgbench -i -s	100		1.5 GB



pgbench -C



F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

pgbench	$\begin{array}{c} \text{postgres} \\ \rightarrow \boxed{ \\ Server} \end{array}$	ClientLANServer	8 cores, 16 GB 1 Gbps 16 cores, 32 GB, HDD
Initialization			Postgres 9.6.1
pgbench -i -s 1	100		1.5 GB



pgbench -C

Postgres Latency

F. Coelho

ntroduction

Subject Application Definitions pgbench

Performance

Connection

Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

pgbench 	LAN Server	ClientLANServer	8 cores, 16 GB 1 Gbps 16 cores, 32 GB, HDD
Initialization	Postgres 9.6.1		
pgbench -i	-s 100		1.5 GB
pgbench -T	2000 -C "host=server	sslmode=require'	' 36.1 tps



pgbench -C

Postgres Latency

F. Coelho

ntroduction

Subject Application Definitions pgbench

Performance

Connection

Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

pgbench LAN Client	postgres →	ClientLANServer	8 cores, 16 GB 1 Gbps 16 cores, 32 GB, HDD
Initialization and	Benchmarks		Postgres 9.6.1
pgbench -i -s	100		1.5 GB
pgbench -T 200 pgbench -T 200	0 -C "host=server 0 -C "host=server	sslmode=require sslmode=disable	" 36.1 tps " 56.4 tps



pgbench -C



F. Coelho

ntroduction

Subject Application Definitions pgbench

Performance

Connection

Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

pgbench	postgres	Client	8 cores, 16 GB
	\rightarrow	LAN	1 Gbps
Client	Server	Server	16 cores, 32 GB, HDD
Initialization and	Benchmarks		Postgres 9.6.1
pgbench -i -s	100		1.5 GB
pgbench -T 200	0 -C "host=server	sslmode=require	' 36.1 tps
pgbench -T 200	0 -C "host=server	sslmode=disable'	' 56.4 tps
pgbench -T 200	0 "host=server	sslmode=disable'	' 105.4 tps



pgbench -C

9/40



F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusior

	pgbench j	postgres → Server	ClientLANServer	8 cores, 16 GB 1 Gbps 16 cores, 32 GB, HDD
	Initialization and B	Benchmarks		Postgres 9.6.1
	pgbench -i -s 1	00		1.5 GB
	pgbench -T 2000 pgbench -T 2000 pgbench -T 2000	-C "host=server -C "host=server "host=server	sslmode=require' sslmode=disable' sslmode=disable'	36.1 tps 56.4 tps 105.4 tps
	connection A	AA		8.2 ms
SSL negociation				10.0 ms
transfers and transactions			9.5 ms	



Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection

Latency Bate & Limit

Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contribution

Performance Comparisons

Latency Pitfalls



pgbench -j 4 -c 8

Version 9.6.1

Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection

Latency

Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contribution

Version 9.5.5

throughput

average latency

329.4 tps throughput 24.3 ms average latency 326.4 tps 24.4 ms



Postares

Latency F. Coelho

Latency

Latency Comparison - 9.5 vs 9.6 pgbench -j 4 -c 8 Version 9.6.1 Version 9.5.5 throughput 329.4 tps throughput average latency average latency 24.3 ms 600 600 500 500 chousand transactions housand transactions: 400 400 300 300 200 200 100 100 0 n 2 з 2 0 1 Λ 0 з transaction latency in seconds transaction latency in seconds

E,

4

326.4 tps

24.4 ms







Instant TPS



What is happening?

Miscellaneous

Wrap-Up Lessons Contributions

transaction surges are absorbe then data are written disk

in-memory + WAL checkpoint



What is happening?

Buy Now, Pay Later!

Latency

- transaction surges are absorbed
- then data are written disk

in-memory + WAL checkpoint



Checkpointing

Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection

Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

Postgres 9.5 Checkpoint

- data writes spread over some time
- OS choose when to actually write
- until fsync is called...

random I/O 30s delay on Linux

Postgres 9.6 Checkpoint



Checkpointing

Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol

Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

Postgres 9.5 Checkpoint

- data writes spread over some time
- OS choose when to actually write
- until fsync is called...

random I/O 30s delay on Linux

ostgres 9.6 Checkpoint



Checkpointing

Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol

Combinations Distance Scalability

Conclusion

Wrap-Up Lessons Contributions

Postgres 9.5 Checkpoint

- data writes spread over some time
- OS choose when to actually write
- until fsync is called...

random I/O 30s delay on Linux I/O storm – on low-end HDD

ostgres 9.6 Checkpoint



Checkpointing

Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection Latency

Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusior

Wrap-Up Lessons Contributions

Postgres 9.5 Checkpoint

- data writes spread over some time
- OS choose when to actually write
- until fsync is called...

random I/O 30s delay on Linux I/O storm – on low-end HDD

Postgres 9.6 Checkpoint

- **sorted** data writes spread over some time
- flush instructions sent regularly (256 kB)

sequential I/O

 $checkpoint_flush_after$



Checkpointing

Postgres Latency

F. Coelho

- Introduction
- Subject Application Definitions pgbench

Performance

Connection Latency Bate & Limit

Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

Postgres 9.5 Checkpoint

- data writes spread over some time
- OS choose when to actually write
- until fsync is called...

random I/O 30s delay on Linux I/O storm – on low-end HDD

Postgres 9.6 Checkpoint

- **sorted** data writes spread over some time
- flush instructions sent regularly (256 kB)
- when fsync is called

sequential I/O

```
checkpoint\_flush\_after
```

ok!


Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contribution

Performance Comparisons

Benchmarking with Rate and Limit



Postgres Latency F. Coelho

pabench

Rate & Limit

Wrap-Up

Rate (tps) and Limit (ms)

pgbench -R 100 -L 100 -N

slow & sllatency	kipped	
Pg 9.6		sorted checkpoint
slow & sllatency	kipped	
slow & sllatency	kipped	



Postares

Latency

E Coelho

Rate & Limit

Wrap-Up

Rate (tps) and Limit (ms)







Postares

Latency

E Coelho

Rate & Limit

Wrap-Up

Rate (tps) and Limit (ms)







pgbench -R 100 -L 100 -N





slow & skipped



Wrap-Up Lessons Contributions



pgbench -R 100 -L 100 -N





pgbench -R 100 -L 100 -N





Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability

Conclusion

Wrap-Up Lessons Contribution

Performance Comparisons

Three Storage Options



FILLFACTOR Storage Parameter

Postgres Latency

F. Coelho

ntroduction

Subject Application Definitions pgbench

Connection Latency Rate & Limit Storage

Combinations Distance Scalability Miscellaneous

Conclusio

Wrap-Up Lessons Contributions

CREATE TABLE pgbench_accounts(...) WITH (FILLFACTOR = 100);

FILLFACTOR Usage

- MVCC: UPDATE = DELETE + INSERT
- some free space available in page
- **but** more pages/costs for other operations

up to 3 pages changes 1 inside page change trade-off

throughputlatency	406.9 tps 19.7 ± 12.3 ms	throughputlatency	



FILLFACTOR Storage Parameter

Postgres Latency

F. Coelho

ntroduction

Subject Application Definitions pgbench

Connection Latency Rate & Limit Storage Protocol

Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

CREATE TABLE pgbench_accounts(...) WITH (FILLFACTOR = 100);

FILLFACTOR Usage

- MVCC: UPDATE = DELETE + INSERT
- some free space available in page
- but more pages/costs for other operations

up to 3 pages changes 1 inside page change trade-off

FILLFACTOR = 10	0			
throughputlatency	406.9 tps 19.7 ± 12.3 ms	throughputlatency		



FILLFACTOR Storage Parameter

Postares Latency

E Coelho

Storage

CREATE TABLE pgbench_accounts(...) WITH (FILLFACTOR = 100);

FILLFACTOR Usage

- MVCC[·] UPDATE = DELETE + INSERT
- some free space available in page
- **but** more pages/costs for other operations

up to 3 pages changes 1 inside page change trade-off

FILLFACTOR = 10	00	FILLFACTOR = 95	
throughputlatency	406.9 tps 19.7 ± 12.3 ms	throughputlatency	416.8 tps 19.2 ± 8.3 ms



pabench

Storage

Miscellar Wrap-Up

Hardware

HDD vs SSD

Postgres Latency	Hard Disk Drive			Solid State Disk	
F. Coelho roduction ubject uplication finitions bench	 mechanics fast sequential I/O slow random I/O 		vs	 electronics fast sequential I/O fast random I/O 	
tency tency	pgbench -j 4 -c 8 -T 2500 -M prepared				
otocol ombinations					
alability scellaneous	HDD 406.9 tps	19.7 ± 12.3	3 ms		
	SSD				



Hardware

HDD vs SSD

HDD

2000



Combinations
Distance
Scalability
Miscellaneous

Rate & Limit Storage

pabench

Wrap-Up

 fast sequential I/O slow random I/O 							fast s fast	sequer randor	itial I n I/O
pgbenc	:h −j 4 −c 8	3 -T 2500 -N	۹ pre	epar	ed				
Postgr	es 9.6				600 500 400	-			instation of a line
HDD SSD	406.9 tps	19.7 ± 12.3	3 ms	tps	300 200 100 0	- - 0	500	- - 1000 seco	1500
								seco	nas

2500



Storage

Hardware

HDD vs SSD



Postgres 9.6						
HDD	406.9 tps	19.7 \pm 12.3 ms				
SSD	4,764.9 tps	$1.7\pm2.4~ms$				





Storage

Hardware

HDD vs SSD



Postgres 9.6						
HDD	406.9 tps	19.7 ± 12.3 ms				
SSD 4,764.9 tps 1.7 ± 2.4 ms						
checkpoint full page write effect						





Can you loose your data?

Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability

Conclusion

Wrap-Up Lessons Contributions

CREATE UNLOGGED TABLE pgbench_accounts(...);

throughputlatency	406.9 tps 19.7 ± 12.3 ms	throughputlatency

. . .



Can you loose your data?

Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contribution

CREATE UNLOGGED TABLE pgbench_accounts(...);

Standard	ACID	
throughputlatency	406.9 tps 19.7 ± 12.3 ms	throughputlatency

. . .



Can you loose your data?

Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

CREATE UNLOGGED TABLE pgbench_accounts(...);

Standard	ACID	UNLOGGED	good luck!
throughputlatency	406.9 tps 19.7 ± 12.3 ms	throughplatency	ut 5,310.7 tps 1.5 ± 0.3 ms

. . .



Can you loose your data?

Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

CREATE UNLOGGED TABLE pgbench_accounts(...);

Standard	ACID	UNLOGGED	good luck!
throughputlatency	406.9 tps 19.7 ± 12.3 ms	throughputlatency	5,310.7 tps 1.5 ± 0.3 ms

NO!



Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance

Miscellaneous

Conclusion

Wrap-Up Lessons Contribution

Performance Comparisons

Two Protocol Impacts



Read-Only In-Cache Test



Postgres Latency

F. Coelho

ntroduction Subject Application Definitions

Performance

Connection Latency Rate & Limit Storage Protocol Combinations

Distance Scalability Miscellaneous

Conclusior

Wrap-Up Lessons Contributions

```
\set aid random(1, 100000 * :scale)
\set tid random(1, 10 * :scale)
\set bid random(1, :scale)
BEGIN;
SELECT abalance FROM pgbench_accounts WHERE aid=:aid;
SELECT tbalance FROM pgbench_tellers WHERE tid=:tid;
SELECT bbalance FROM pgbench_branches WHERE bid=:bid;
COMMIT;
```

1 transfers	network protocol
2 parse query	syntax analysis
3 plan query	optimization
4 execute query	cheap if in cache



Read-Only In-Cache Test



Postgres Latency

F. Coelho

ntroduction Subject Application Definitions pgbench

Performance Connection

Latency Rate & Limit

Protocol

Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

```
\set aid random(1, 100000 * :scale)
\set tid random(1, 10 * :scale)
\set bid random(1, :scale)
BEGIN;
SELECT abalance FROM pgbench_accounts WHERE aid=:aid;
SELECT tbalance FROM pgbench_tellers WHERE tid=:tid;
SELECT bbalance FROM pgbench_branches WHERE bid=:bid;
COMMIT;
```

Operations	Queries on 3 tables
1 transfers	network protocol
2 parse query	syntax analysis
3 plan query	optimization
4 execute query	cheap if in cache



Destares

Storage Protocol

Miscellaneous

Wrap-Up

Protocol

SSL or not

Latency F. Coelho	SSL Costs time $\& \in$	
ntroduction	negotiation and re-negotiation	Confidentiality
Subject Application Definitions	cryptographic functions	Integrity
pgbench	certificate	Authentication
Connection		

throughputlatency	709.7 tps $.407 \pm 0.132$ ms	throughputlatency	



SSL or not

Postgres
Latency

F. Coelho

ntroduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage **Protocol** Combinations

Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

SSL Costs	time & €	Benefits		
 negotiation and re-negotiation cryptographic functions certificate 		ConfidentialityIntegrityAuthentication		
throughputlatency 1.402	709.7 tps 7 ± 0.132 ms	throughputlatency		



SSL or not

Postgres Latency

F. Coelho

pabench

Storage Protocol

Miscellaneous

Wrap-Up

SSL Costs	time & €	Benefits	Snake Oil!
 negotiation and re-negotiation cryptographic functions certificate? 		ConfidentialityIntegrityAuthentication	y n
 throughput latency 	709.7 tps 1.407 ± 0.132 ms	throughputlatency	



pgbenc

Protoco

Miscell Wrap-U

Protocol

SSL or not

Postgres Latency F. Coelho	SSL Costs	time & €	Benefits	Snake Oil!
troduction bject vplication vfinitions bench	 negotiation and rest cryptographic fun certificate 	e-negotiation actions	ConfidentialityIntegrityAuthentication	
onnection tency tte & Limit orage otocol	pgbench -j 1 -c 1 -D) scale=100 -f ro	3.sql -T 30 "host=serve:	r"
ombinations stance alability	sslmode=require	SSL		
scellaneous onclusion rap-Up ssons ontributions	throughputlatency 1.4	709.7 tps 07 ± 0.132 ms	throughputlatency	



pgben

Protoco

Wrap-I

Protocol

SSL or not

Postgres Latency F. Coelho	SSL Costs	time & €	Benefits	Snake Oil!
plication solution offinitions bench	 negotiation and r cryptographic fur certificate 	e-negotiation actions	ConfidentialIntegrityAuthentication	ty on
onnection tency tte & Limit orage otocol	pgbench -j 1 -c 1 -E) scale=100 -f ro	3.sql -T 30 "host	=server"
stance alability	sslmode=require	SSL	sslmode=disabl	e <i>clear</i>
scellaneous onclusion rap-Up ssons ontributions	throughputlatency 1.4	709.7 tps 07 ± 0.132 ms	throughputlatency	781.6 tps 1.277 ± 0.034 ms



Simple vs Prepared

Postgres Latency

F. Coelho

ntroduction

Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance

Miscellaneous

Conclusior

Wrap-Up Lessons Contributions

prepare once in session
PREPARE Abal(INT) AS
SELECT abalance
FROM pgbench_accounts
WHERE aid=\$1;
execute multiple times
EXECUTE Abal(1);
EXECUTE Abal(5432);
EXECUTE Abal(18);

Prepare

- temporary one-cmd function
- factor out *parse* cost
- keep plan and execute
- pgbench -M prepared...

throughputlatency	709.7 tps 1.407 ± 0.132 ms	througlatence	jhput y	



Simple vs Prepared

Postgres Latency

F. Coelho

ntroduction

Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage **Protocol** Combinations Distance

Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions -- prepare once in session PREPARE Abal(INT) AS SELECT abalance FROM pgbench_accounts WHERE aid=\$1; -- execute multiple times... EXECUTE Abal(1); EXECUTE Abal(5432); EXECUTE Abal(18);

Prepare

- temporary one-cmd function
- factor out *parse* cost
- keep plan and execute
- pgbench -M prepared...

ro3.sql	simple		
throughput	709.7 tps	throughput	
latency	1.407 \pm 0.132 ms	latency	



Simple vs Prepared

Postgres Latency

F. Coelho

ntroductior

Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance

Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

prepare once in session
PREPARE Abal(INT) AS
SELECT abalance
$FROM pgbench_accounts$
WHERE aid=\$1;
execute multiple times
<pre>EXECUTE Abal(1);</pre>
EXECUTE Abal(5432);
EXECUTE Abal(18);

Prepare

- temporary one-cmd function
- factor out *parse* cost
- keep plan and execute
- pgbench -M prepared...

ro3.sql	simple	ro3.sql	prepared
throughput	709.7 tps	throughput	860.0 tps
latency	1.407 \pm 0.132 ms	latency	1.161 \pm 0.082 ms



Postares Latency

F. Coelho

pabench

Rate & Limit Combinations

Wrap-Up

Performance Comparisons

Four Query Combination Tricks



UPDATE & SELECT

Postgres Latency

F. Coelho

ntroductior Subject

Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations

Scalability

Miscellaneous

Conclusior

Wrap-Up Lessons Contributions

-- update table UPDATE pgbench_accounts SET abalance = abalance + :delta WHERE aid = :aid;

-- get updated data SELECT abalance FROM pgbench_accounts WHERE aid = :aid;

Standard

throughputlatency

406.9 tps 19.7 ± 12.3 ms --combined JPDATE pgbench_accounts SET abalance = abalance + :delta WHERE aid = :aid RETURNING abalance;

UPDATE RETURNING Option

return updated rows

one parse, plan, execute

Combined Update

- throughput
- latency



UPDATE & SELECT

Postgres Latency

E Coelho

Combinations

-- update table UPDATE pgbench_accounts

SET abalance = abalance + \cdot delta WHERE aid = :aid:

-- get updated data SELECT abalance FROM pgbench_accounts WHERE aid = :aid;

throughput

-- combined

UPDATE pgbench_accounts SET abalance = abalance + :deltaWHERE aid = \cdot aid **RETURNING** abalance:

UPDATE RETURNING Option

- return updated rows
- one parse, plan, execute

- throughput
- latencv



UPDATE & SELECT

Postgres Latency

F. Coelho

ntroductio Subject Application Definitions

Performance

Connection Latency Rate & Limit Storage Protocol

Combinations

Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions -- update table

UPDATE pgbench_accounts SET abalance = abalance + :delta WHERE aid = :aid;

-- get updated data SELECT abalance FROM pgbench_accounts WHERE aid = :aid; -- combined

UPDATE pgbench_accounts SET abalance = abalance + :delta WHERE aid = :aid RETURNING abalance;

UPDATE RETURNING Option

- return updated rows
- one parse, plan, execute

Standard

throughputlatency

Combined Update

- throughput
- latency

406.9 tps

 $19.7 + 12.3 \, ms$



UPDATE & SELECT

Postgres Latency

F. Coelho

ntroductio Subject Application Definitions

Performance

Connection Latency Rate & Limit Storage Protocol Combinations

Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions -- update table UPDATE pgbench_accounts SET abalance = abalance + :delta

WHERE aid = :aid;

-- get updated data SELECT abalance FROM pgbench_accounts WHERE aid = :aid; -- combined

UPDATE pgbench_accounts SET abalance = abalance + :delta WHERE aid = :aid RETURNING abalance;

UPDATE RETURNING Option

- return updated rows
- one parse, plan, execute

406.9 tps
19.7 \pm 12.3 ms

Combined Update

- throughput 408.2 tps
- latency 19.6 ± 8.7 ms


Client-combined SQL Queries

Postgres Latency

F. Coelho

ntroduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations

Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

```
-- "ro3c.sql" pgbench script
\set aid random(1, 100000 * :scale)
\set tid random(1, 10 * :scale)
\set bid random(1, :scale)
BEGIN \;
SELECT abalance FROM
pgbench_accounts WHERE aid=:aid \;
SELECT tbalance FROM
pgbench_tellers WHERE tid=:tid \;
SELECT bbalance FROM
pgbench_branches WHERE bid=:bid \;
COMMIT;
```

Combine

with \backslash ;

- embedded semi-colon ;
- request with multiple queries
- response with list of results
- avoid request-response loop

throughput	709.7 tps	throughput	
latency	1.407 ± 0.132 ms	latency	



Client-combined SQL Queries

Postgres Latency

F. Coelho

ntroduction

Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations

Combinations

Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

```
-- "ro3c.sql" pgbench script
\set aid random(1, 100000 * :scale)
\set tid random(1, 10 * :scale)
\set bid random(1, :scale)
BEGIN \;
SELECT abalance FROM
pgbench_accounts WHERE aid=:aid \;
SELECT tbalance FROM
pgbench_tellers WHERE tid=:tid \;
SELECT bbalance FROM
pgbench_branches WHERE bid=:bid \;
COMMIT;
```

Combine

with \backslash ;

- embedded semi-colon ;
- request with multiple queries
- response with list of results
- avoid request-response loop

ro3.sql	standard		
throughput	709.7 tps	throughput	
latency	1.407 \pm 0.132 ms	latency	



Client-combined SQL Queries

Postgres Latency

F. Coelho

ntroduction

Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol

Combinations

Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

```
-- "ro3c.sql" pgbench script
\set aid random(1, 100000 * :scale)
\set tid random(1, 10 * :scale)
\set bid random(1, :scale)
BEGIN \;
SELECT abalance FROM
pgbench_accounts WHERE aid=:aid \;
SELECT tbalance FROM
pgbench_tellers WHERE tid=:tid \;
SELECT bbalance FROM
pgbench_branches WHERE bid=:bid \;
COMMIT;
```

Combine

embedded semi-colon ;

- request with multiple queries
- response with list of results
- avoid request-response loop

ro3.sql	standard	ro3c.sql	combined
throughput	709.7 tps	throughput	1,311.5 tps
latency	1.407 \pm 0.132 ms	latency	$\textit{0.748} \pm \textit{0.132}~ms$

with $\:$



Postgres Latency

F. Coelho

ntroductior

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol

Combinations

Distance Scalability Miscellaneou

Conclusion

Wrap-Up Lessons Contributions

CREA	ΤΕ ΤΥΡ	E Bal	ances			
AS	(abal	INT,	tbal	INT,	bbal	INT);

CREATE FUNCTION getBalSQL(INT, INT, INT)
 RETURNS Balances AS \$\$
 SELECT
 (SELECT abalance
 FROM pgbench_accounts WHERE aid=\$1),
 (SELECT tbalance
 FROM pgbench_tellers WHERE tid=\$2),
 (SELECT bbalance

FROM pgbench_branches WHERE bid=\$3)

```
$$ LANGUAGE SQL;
```

-- "ro3sf.sql" pgbench script \set aid random(1, 100000 * :scale) \set tid random(1, 10 * :scale) \set bid random(1, :scale) SELECT getBalSQL(:aid, :tid, :bid);

throughputlatency	709.7 tps 1.407 ± 0.132 ms	throughputlatency	



Postgres Latency

F. Coelho

ntroductior

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol

Combinations

Distance Scalability Miscellaneou

Conclusion

Wrap-Up Lessons Contributions CREATE TYPE Balances AS (abal INT, tbal INT, bbal INT);

CREATE FUNCTION getBalSQL(INT, INT, INT) RETURNS Balances AS \$\$ SELECT (SELECT abalance FROM pgbench_accounts WHERE aid=\$1), (SELECT tbalance FROM pcbench tallang WHERE tid=\$2)

FROM pgbench_tellers WHERE tid=\$2),

(SELECT bbalance FROM pgbench_branches WHERE bid=\$3)

\$\$ LANGUAGE SQL;

-- "ro3sf.sql" pgbench script \set aid random(1, 100000 * :scale) \set tid random(1, 10 * :scale) \set bid random(1, :scale) SELECT getBalSQL(:aid, :tid, :bid);

ro3.sql	standard		
throughput	709.7 tps	throughput	
latency	$1.407\pm0.132~ms$	latency	



Postgres Latency

F. Coelho

ntroductior

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage

Protocol

Distance Scalability Miscellaneous

Conclusion

Lessons Contributions CREATE TYPE Balances AS (abal INT, tbal INT, bbal INT);

CREATE FUNCTION getBalSQL(INT, INT, INT)
 RETURNS Balances AS \$\$
 SELECT
 (SELECT abalance
 FROM pgbench_accounts WHERE aid=\$1),
 (SELECT tbalance
 FROM pgbench_tellers WHERE tid=\$2),
 (SELECT bbalance

FROM pgbench_branches WHERE bid=\$3)

\$\$ LANGUAGE SQL;

-- "ro3sf.sql" pgbench script \set aid random(1, 100000 * :scale) \set tid random(1, 10 * :scale) \set bid random(1, :scale) SELECT getBalSQL(:aid, :tid, :bid);

ro3.sql	standard	ro3sf.sql	SQL call
throughput	709.7 tps	throughput	1,395.4 tps
latency	1.407 \pm 0.132 ms	latency	$\textit{0.712}\pm\textit{0.075}\ \textit{ms}$



Postgres Latency

F. Coelho

ntroduction

Subject Application Definitions pgbench

Performance

Latency Rate & Limit Storage Protocol Combinations Distance

Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

CREATE FUNCTION getBalPL(a INT, t INT, b INT) RETURNS Balances AS \$\$ DECLARE abal INT; tbal INT; bbal INT; BEGIN SELECT abalance INTO abal

SELECT abalance INTO abal FROM pgbench_accounts WHERE aid=a; SELECT tbalance INTO tbal FROM pgbench_tellers WHERE tid=t; SELECT bbalance INTO bbal FROM pgbench_branches WHERE bid=b; RETURN (abal, tbal, bbal)::Balances; END; \$ LANGUAGE PLpgSQL;

-- "ro3pf.sql" pgbench script \set aid random(1, 100000 * :scale) \set tid random(1, 10 * :scale) \set bid random(1, :scale) SELECT getBalPL(:aid, :tid, :bid);

throughput	709.7 tps	throughput	
latency	$1.407\pm0.132~ms$	latency	29



Postgres Latency

F. Coelho

ntroduction

Subject Application Definitions pgbench

Connection

Latency

```
Storage
```

Protocol

Combinations

Distance Scalability Miscellaneou

Conclusion

Wrap-Up Lessons Contributions

CREATE FUNCTION getBalPL(a INT, t INT, b INT) RETURNS Balances AS \$\$ DECLARE abal INT; tbal INT; bbal INT; BEGIN SELECT abalance INTO abal

FROM pgbench_accounts WHERE aid=a;

SELECT tbalance INTO tbal

FROM pgbench_tellers WHERE tid=t;

SELECT bbalance INTO bbal

FROM pgbench_branches WHERE bid=b; RETURN (abal, tbal, bbal)::Balances;

END;

\$\$ LANGUAGE PLpgSQL;

-- "ro3pf.sql" pgbench script \set aid random(1, 100000 * :scale) \set tid random(1, 10 * :scale) \set bid random(1, :scale) SELECT getBalPL(:aid, :tid, :bid);

ro3.sql	standard		
throughput	709.7 tps	throughput	
latency	1.407 \pm 0.132 ms	latency	00



Postgres Latency

F. Coelho

ntroduction

Subject Application Definitions pgbench

Connection

Latency Rate & Lin

```
Storage
```

Protocol

Combinations

Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

CREATE FUNCTION getBalPL(a INT, t INT, b INT) RETURNS Balances AS \$\$ DECLARE abal INT; tbal INT; bbal INT; BEGIN SELECT abalance INTO abal FROM pgbench_accounts WHERE aid=a; SELECT tbalance INTO tbal FROM pgbench_tellers WHERE tid=t; SELECT bbalance INTO bbal FROM pgbench_branches WHERE bid=b;

RETURN (abal, tbal, bbal)::Balances; END:

\$\$ LANGUAGE PLpgSQL;

-- "ro3pf.sql" pgbench script \set aid random(1, 100000 * :scale) \set tid random(1, 10 * :scale) \set bid random(1, :scale) SELECT getBalPL(:aid, :tid, :bid);

?

ro3.sql	standard	ro3pf.sql	PL/pgSQL call
throughput	709.7 tps	throughput	2,485.5 tps
latency	1.407 \pm 0.132 ms	latency	$0.400 \pm 0.055~ms$



Postgres Latency

F. Coelho

ntroduction

Subject Application Definitions pgbench

Connection

Latency Rate & Lin

Storage

Protocol

Combinations

Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

CREATE FUNCTION getBalPL(a INT, t INT, b INT) RETURNS Balances AS \$\$ DECLARE abal INT; tbal INT; bbal INT; BEGIN SELECT abalance INTO abal FROM pgbench_accounts WHERE aid=a; SELECT tbalance INTO tbal

FROM pgbench_tellers WHERE tid=t;

SELECT bbalance INTO bbal

\$\$ LANGUAGE PLpgSQL;

FROM pgbench_branches WHERE bid=b; RETURN (abal, tbal, bbal)::Balances; END:

-- "ro3pf.sql" pgbench script \set aid random(1, 100000 * :scale) \set tid random(1, 10 * :scale) \set bid random(1, :scale) SELECT getBalPL(:aid, :tid, :bid);

PL/pgSQL caches plans!

ro3.sql	standard	ro3pf.sql	PL/pgSQL call
throughput	709.7 tps	throughput	2,485.5 tps
latency	1.407 \pm 0.132 ms	latency	$0.400 \pm 0.055~ms$



Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability

Conclusion

Wrap-Up Lessons Contribution

Performance Comparisons

Reducting Server Distance



Postgres Latency

F. Coelho

ntroduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability

Conclusio

Wrap-Up Lessons Contributions

Intercon	nection
LAN	Local Area Network
LO	loopback interface

IPC Inter-Process Communication

localhost Unix domain socket

Ethernet

LAN LO IPC	100.3 tps	9.9 ms	LAN LO IPC	709.7 tps	1.4 ms



Interconnection

Postgres Latency

F. Coelho

ntroduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability

Conclusic

Wrap-Up Lessons Contributions

LAN	Local Area Network
LO	loopback interface

IPC Inter-Process Communication

Ethernet localhost

Unix domain socket

TPC-B-Like		on HDD			
LAN	100.3 tps	9.9 ms	LAN	709.7 tps	1.4 ms
LO	114.5 tps	8.7 ms			
IPC	113.5 tps	8.8 ms			



Interconnection

Postgres Latency

F. Coelho

ntroduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability

Conclusic

Wrap-Up Lessons Contributions

LAN	Local Area Network
LO	loopback interface

IPC Inter-Process Communication

Ethernet localhost

Unix domain socket

TPC-B-Like	on SS	SD			
LAN 403. LO 1,133. IPC 1,243.	8 tps 2.4 n 3 tps 0.9 n 1 tps 0.8 n	ns ns ns	LAN LO IPC	709.7 tps	1.4 ms



Interconnection

Postgres Latency

F. Coelho

ntroduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability

Conclusio

Wrap-Up Lessons Contributions

LAN	Local Area Network
LO	loopback interface
	later Decessor Occurrent's attent

IPC Inter-Process Communication

Ethernet

localhost

Unix domain socket

TPC-B-Like on SSD		Read-Only 3			
LAN	403.8 tps	2.4 ms	LAN	709.7 tps	1.4 ms
LO	1,133.3 tps	0.9 ms	LO	2,515.3 tps	0.4 ms
IPC	1,243.1 tps	0.8 ms	IPC	3,607.6 tps	0.3 ms



Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contribution

Performance Comparisons

Performance Scalability



Postgres

Latency

Clients Scalability



Read-Only 3 – remote SSL simple queries

F. Coelho

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

Best Throughput

Best Latency

compromise



Base

Postgres Latency

F. Coelho

Introduction

Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

Best Throu	ighput	•
37,639 tps	4.103 ms	156/4

Best Latency

Compromise

Read-Only 3 – remote SSL simple queries





Base

Postgres Latency

F. Coelho

Introductio Subject Application Definitions pabench

Performance Connection

Latency Rate & Limit Storage Protocol Combinations Distance Scalability

Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

Best Throughput				
37,639 tps	4.103 ms	156/4		
Best Latency				
5,748 tps	1.042 ms	6/1		

Read-Only 3 – remote SSL simple queries





Base

Postgres Latency

F. Coelho

Introductio Subject Application Definitions pabench

Performanc Connection Latency Rate & Limit Storage Protocol Combinations

Scalability

Conclusion

Wrap-Up Lessons Contributions

Best Throughput				
37,639 tps	4.103 ms	156/4		
Best Laten	су	•		
5,748 tps	1.042 ms	6/1		
Compromi	se	•		

Read-Only 3 – remote SSL simple queries





Postares

Latency F. Coelho

Clients Scalability



Read-Only 3 – remote noSSL prepared PL call

Subject

Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

Best Throughput

Best Latency

romise



Best

Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

Best Throughput **181,503 tps** 0.766 ms *140/4*

Best Latency

Compromise

Read-Only 3 – remote noSSL prepared PL call





Best

Postgres Latency

F. Coelho

Subject Application Definitions

pgbench

Performance

Latency Rate & Limit Storage Protocol Combinations Distance

Scalability

Miscellaneous

Conclusion

Wrap-Up Lessons Contributions









Best

Postgres Latency

F. Coelho

Introductio Subject Application Definitions

Performanc

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

Best Throughput				
181,503 tps	140/4			
Best Latence	су	•		
39,232 tps	0.254 ms	10/2		
39,232 tps	0.254 ms	10/2		
39,232 tps Compromis	0.254 ms e	10/2		

Read-Only 3 – remote noSSL prepared PL call





Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability

Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

Performance Comparisons

Miscellaneous Settings



App & Postgres

Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

Δn	nli	02	tic	۱n
¬μ	ρī	ua	uc	Л

connection persistence cache Memcached Redis

ostgres configuration

disk block_size random_page_cost

ch

nemory shared_buffers effective_cache_size huge_p

eckpoint _timeout _completion_target _flush_after

wal max_wal_size

framework?

change defaults





Postares Latency

E Coelho

Miscellaneous

Application

connection persistence cache Memcached Redis

Postgres configuration

disk block_size random_page_cost

memory shared_buffers effective_cache_size huge_pages checkpoint _timeout _completion_target _flush_after wal max wal size

framework?

change defaults





tweak and choose

Postgres Latency

F. Coelho

ntroduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability

Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

FS XFS ext4 Btrfs ZFS, mount options IO io scheduler, queue length, write delay, dirty bytes... others NUMA....

lardware

OS

expensive is (probably) better

diskS tables wal logs, HDD-with-cache, SSI

tweaking read ahead, write flush

RAID with large caches, BBU





Postgres Latency

F. Coelho

ntroduction

Subject Application Definitions pgbench

Performance

Latency Rate & Limit Storage Protocol Combinations Distance

Scalability

Miscellaneous

Conclusion

Wrap-Up Lessons Contributions

DS		tweak and choose
FS	XFS ext4 Btrfs ZFS, mount options	
IO	io scheduler, queue length, write delay, dirty bytes	
thers	NUMA,	

Hardware

diskS tables wal logs, HDD-with-cache, SSD

tweaking read ahead, write flush

RAID with large caches, BBU

expensive is (probably) better

36/40



Postgres Latency

F. Coelho

Introduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusion

Wrap-Up Lessons Contribution

Conclusion



Pc

E

Wrap-I

Wrap-up

pgbench -j 1 -c 1 ...

stgres		TPC-E	3-like	Read-Ol	nly 3
atency		tps	ms	tps	m
Coelho	HDD -C SSL	36.1	27.7	52.7	18.
luction	HDD -C noSSL	56.4	17.7	110.1	9.
t	HDD SSL	105.4	9.5	709.7	1.
ation	SSD SSL	403.8	2.47	695.1	1.
sh	SSD noSSL	465.4	2.15	820.1	1.
rmance	+ prepared	548.1	1.82	974.0	1.
ction y	 returning 	529.4	1.89	—	
Limit	+ prepared	681.2	1.47	_	
ol	- combined	857.8	1.15	1,536.4	0.
nations ce	 – SQL func 	940.3	1.06	1,818.1	0.
ility	+ prepared	957.9	1.04	2,144.7	0.
lucion	- PL func	1,279.4	0.78	2,778.0	0.
Jp	+ prepared	1,323.2	0.75	3,040.4	0.
ns	localhost	1,907.6	0.52	10,006.8	0.
	socket	2,273.1	0.44	11,545.5	0.

connection

ms

18.96

9.08

1.41

1.44

1.22

1.02

0.64

0.55

0.46

0.36

0.33

0.10

0.09

_

- simple to prepared
- combinations...
- remote to local



Postares

Wrap-up

pgbench -j 1 -c 1 ...

Latency					
F. Coelho					
atroduction					
Subject					
Application					
Definitions					
ogbench					
erformance					
Connection					
_atency					
Rate & Limit					
Storage					
Protocol					
Combinations					

Combination
Distance
Scalability
Miscellaneou

Conclusio

Wrap-Up Lessons Contributions

	TPC-B-like		Read-O	nly 3
	tps	ms	tps	ms
HDD -C SSL	36.1	27.7	52.7	18.96
HDD -C noSSL	56.4	17.7	110.1	9.08
HDD SSL	105.4	9.5	709.7	1.41
SSD SSL	403.8	2.47	695.1	1.44
SSD noSSL	465.4	2.15	820.1	1.22
+ prepared	548.1	1.82	974.0	1.02
 returning 	529.4	1.89	-	-
+ prepared	681.2	1.47	-	-
 combined 	857.8	1.15	1,536.4	0.64
 SQL func 	940.3	1.06	1,818.1	0.55
+ prepared	957.9	1.04	2,144.7	0.46
 PL func 	1,279.4	0.78	2,778.0	0.36
+ prepared	1,323.2	0.75	3,040.4	0.33
localhost	1,907.6	0.52	10,006.8	0.10
socket	2,273.1	0.44	11,545.5	0.09

connection

HDD to SSD

- SSL to none
- simple to prepared
- combinations...
- remote to local

 $\times 3$ to $\times 14$



Wrap-up

pgbench -j 1 -c 1 ...

Postgres
Latency

F. Coelho

ntroduction

Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability

Conclusio

Wrap-Up Lessons Contributions

	TPC-B-like		Read-Only 3	
	tps	ms	tps	ms
HDD -C SSL	36.1	27.7	52.7	18.96
HDD -C noSSL	56.4	17.7	110.1	9.08
HDD SSL	105.4	9.5	709.7	1.41
SSD SSL	403.8	2.47	695.1	1.44
SSD noSSL	465.4	2.15	820.1	1.22
+ prepared	548.1	1.82	974.0	1.02
 returning 	529.4	1.89	—	-
+ prepared	681.2	1.47	-	-
 combined 	857.8	1.15	1,536.4	0.64
 SQL func 	940.3	1.06	1,818.1	0.55
+ prepared	957.9	1.04	2,144.7	0.46
 PL func 	1,279.4	0.78	2,778.0	0.36
+ prepared	1,323.2	0.75	3,040.4	0.33
localhost	1,907.6	0.52	10,006.8	0.10
socket	2,273.1	0.44	11,545.5	0.09

connectionHDD to SSD

- SSL to none
- simple to prepared
- combinations...
- remote to local

imes4 to =



Wrap-up

pgbench -j 1 -c 1 ...

Postgres
Latency

F. Coelho

	4.0	Ч		÷	
1		u			

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability

Conclusio

Wrap-Up Lessons Contributions

	TPC-B-like		Read-O	nly 3
	tps	ms	tps	ms
HDD -C SSL	36.1	27.7	52.7	18.96
HDD -C noSSL	56.4	17.7	110.1	9.08
HDD SSL	105.4	9.5	709.7	1.41
SSD SSL	403.8	2.47	695.1	1.44
SSD noSSL	465.4	2.15	820.1	1.22
+ prepared	548.1	1.82	974.0	1.02
 returning 	529.4	1.89	—	-
+ prepared	681.2	1.47	-	-
 combined 	857.8	1.15	1,536.4	0.64
 SQL func 	940.3	1.06	1,818.1	0.55
+ prepared	957.9	1.04	2,144.7	0.46
 PL func 	1,279.4	0.78	2,778.0	0.36
+ prepared	1,323.2	0.75	3,040.4	0.33
localhost	1,907.6	0.52	10,006.8	0.10
socket	2,273.1	0.44	11,545.5	0.09

connectionHDD to SSD

- SSL to none
- simple to prepared
- combinations...
- remote to local

+15% to +18%



Postares

Wrap-up

pgbench -j 1 -c 1 ...

Latency	
F. Coelho	
ntroduction	
Subject	
Application	
Definitions	
ogbench	
Performance	
Connection	
Latency	
Rate & Limit	
Storage	
Protocol	
Combinations	
Distance	
Scalability	
Miscellaneous	

Conclusio

Wrap-Up Lessons Contributions

	TPC-B-like		Read-O	nly 3
	tps	ms	tps	ms
HDD -C SSL	36.1	27.7	52.7	18.96
HDD -C noSSL	56.4	17.7	110.1	9.08
HDD SSL	105.4	9.5	709.7	1.41
SSD SSL	403.8	2.47	695.1	1.44
SSD noSSL	465.4	2.15	820.1	1.22
+ prepared	548.1	1.82	974.0	1.02
 returning 	529.4	1.89	—	-
+ prepared	681.2	1.47	-	-
 combined 	857.8	1.15	1,536.4	0.64
– SQL func	940.3	1.06	1,818.1	0.55
+ prepared	957.9	1.04	2,144.7	0.46
– PL func	1,279.4	0.78	2,778.0	0.36
+ prepared	1,323.2	0.75	3,040.4	0.33
localhost	1,907.6	0.52	10,006.8	0.10
socket	2,273.1	0.44	11,545.5	0.09

connectionHDD to SSD

SSL to none

simple to prepared

combinations...

remote to local

+2% to +28%



Postares

Wrap-up

pgbench -j 1 -c 1 ...

Latency			
F. Coelho			
atroduction			
Subject			
Application			
Definitions			
ogbench			
erformance			
Connection			
atency			
Rate & Limit			
Storage			
Protocol			
Combinations			
Distance			
Scalability			
Ale and the second			

Conclusior

Wrap-Up Lessons Contributions

	TPC-B-like		Read-Only 3	
	tps	ms	tps	ms
HDD -C SSL	36.1	27.7	52.7	18.96
HDD -C noSSL	56.4	17.7	110.1	9.08
HDD SSL	105.4	9.5	709.7	1.41
SSD SSL	403.8	2.47	695.1	1.44
SSD noSSL	465.4	2.15	820.1	1.22
+ prepared	548.1	1.82	974.0	1.02
 returning 	529.4	1.89	-	-
+ prepared	681.2	1.47	-	-
- combined	857.8	1.15	1,536.4	0.64
– SQL func	940.3	1.06	1,818.1	0.55
+ prepared	957.9	1.04	2,144.7	0.46
– PL func	1,279.4	0.78	2,778.0	0.36
+ prepared	1,323.2	0.75	3,040.4	0.33
localhost	1,907.6	0.52	10,006.8	0.10
socket	2,273.1	0.44	11,545.5	0.09

connection

- HDD to SSD
- SSL to none
- simple to prepared
- combinations...

remote to local

 $\times 3$ to $\times 4$


Destause

Wrap-up

pgbench -j 1 -c 1 ...

rosigres	
Latency	
F. Coelho	H
traduction	HD
ubject	H
pplication	S
gbench	S
erformance	
onnection atency	— r
late & Limit	
rotocol	— c
ombinations listance	- 5
calability	
iiscellaneous	— F

Wrap-Up Lessons Contributions

	TPC-B-like		Read-O	nly 3
	tps	ms	tps	ms
HDD -C SSL	36.1	27.7	52.7	18.96
HDD -C noSSL	56.4	17.7	110.1	9.08
HDD SSL	105.4	9.5	709.7	1.41
SSD SSL	403.8	2.47	695.1	1.44
SSD noSSL	465.4	2.15	820.1	1.22
+ prepared	548.1	1.82	974.0	1.02
 returning 	529.4	1.89	-	-
+ prepared	681.2	1.47	—	_
 combined 	857.8	1.15	1,536.4	0.64
 SQL func 	940.3	1.06	1,818.1	0.55
+ prepared	957.9	1.04	2,144.7	0.46
 PL func 	1,279.4	0.78	2,778.0	0.36
+ prepared	1,323.2	0.75	3,040.4	0.33
localhost	1,907.6	0.52	10,006.8	0.10
socket	2,273.1	0.44	11,545.5	0.09

connection

- HDD to SSD
- SSL to none
- simple to prepared
- combinations...
- remote to local

 $\times 1.7$ to $\times 3.7$



Wrap-up

pgbench -j 1 -c 1 ...

Postgres		IPC
Latency		tps
F. Coelho	HDD -C SSL	36.1
and and for a	HDD -C noSSL	56.4
Iroduction	HDD SSL	105.4
oplication	SSD SSL	403.8
bench	SSD noSSL	465.4
erformance	+ prepared	548.1
onnection Itency	- returning	529.4
ate & Limit	+ prepared	681.2
orage otocol	- combined	857.8
ombinations stance	– SQL func	940.3
calability	\dots + prepared	957.9
scellaneous	– PL func	1,279.4
nclusion rap-Up	+ prepared	1,323.2
issons	localhost	1.907.6

	TPC-В	8-like	Read-Only 3	
	tps	ms	tps	ms
HDD -C SSL	36.1	27.7	52.7	18.96
HDD -C noSSL	56.4	17.7	110.1	9.08
HDD SSL	105.4	9.5	709.7	1.41
SSD SSL	403.8	2.47	695.1	1.44
SSD noSSL	465.4	2.15	820.1	1.22
+ prepared	548.1	1.82	974.0	1.02
 returning 	529.4	1.89	-	-
+ prepared	681.2	1.47	-	-
 combined 	857.8	1.15	1,536.4	0.64
 SQL func 	940.3	1.06	1,818.1	0.55
+ prepared	957.9	1.04	2,144.7	0.46
 PL func 	1,279.4	0.78	2,778.0	0.36
+ prepared	1,323.2	0.75	3,040.4	0.33
localhost	1,907.6	0.52	10,006.8	0.10
socket	2,273.1	0.44	11,545.5	0.09

- connection
- HDD to SSD
- SSL to none
- simple to prepared
- combinations...
- remote to local

\times 63 to \times 219

and scaling effects



Destares

Latency F. Coelho	Things to Bri	ng Home	in-memory OLTP load
Introduction Subject Application	NoTPS	not only TPS latency-throughput compromise	latency matters!
pgbench Performance Connection		experiment and measure pgbench is improving	do not assume!
Latency Rate & Limit Storage Protocol		version sorted and flushed checkpoints	9.6!
Distance Scalability		costs	network, parse & plan
Miscellaneous Conclusion		ACID	$SSD \gg HDD$
Wrap-Up		pg as a cache manager	SSD = HDD



Storage

Miscellane Wrap-Up Lessons

Postgres Latency F. Coelho	Things to Brin	ng Home	in-memory OLTP load
plication	NoTPS	not only TPS latency-throughput compromise	latency matters!
bench erformance	Performance	experiment and measure pgbench is improving	do not assume!
tency ite & Limit orage otocol ombinations		version sorted and flushed checkpoints	9.6!
stance alability		costs	network, parse & plan
scellaneous		ACID	$SSD \gg HDD$
rap-Up ssons		pg as a cache manager	SSD = HDD



Miscellar Wrap-Up Lessons

Postgres Latency F. Coelho	Things to Brin	ng Home	in-memory OLTP load
plication	NoTPS	not only TPS latency-throughput compromise	latency matters!
bench erformance	Performance	experiment and measure pgbench is improving	do not assume!
tency te & Limit orage otocol ombinations	Postgres	version sorted and flushed checkpoints	9.6!
stance		costs	network, parse & plan
scellaneous		ACID	$SSD \gg HDD$
rap-Up ssons		pg as a cache manager	SSD = HDD



Storage Protocol

Miscellar Wrap-Up Lessons

Postgres			
Latency	Things to Brir	ng Home	in-memory OLTP load
F. Coelho	Ŭ	3	
plication	NoTPS	not only TPS latency-throughput compromise	latency matters!
bench erformance	Performance	experiment and measure pgbench is improving	do not assume!
tency ite & Limit orage otocol ombinations	Postgres	version sorted and flushed checkpoints	9.6!
stance	High	costs	network, parse & plan
scellaneous		ACID	$SSD \gg HDD$
rap-Up ssons		pg as a cache manager	SSD = HDD



Storage

Wrap-Up Lessons

Postgres			
Latency	Things to Bri	ng Home	in-memory OLTP load
F. Coelho	0	5	· · · , · · · · · ·
plication	NoTPS	not only TPS latency-throughput compromise	latency matters!
bench erformance	Performance	experiment and measure pgbench is improving	do not assume!
tency ite & Limit orage otocol ombinations	Postgres	version sorted and flushed checkpoints	9.6!
stance	High	costs	network, parse & plan
scellaneous	RW load	ACID	$SSD \gg HDD$
rap-Up ssons		pg as a cache manager	SSD = HDD



Storage

Miscellane Wrap-Up Lessons

Postgres			
Latency	Things to Brin	na Home	in-memory OI TP load
F. Coelho	Thinge to Bh		
plication	NoTPS	not only TPS latency-throughput compromise	latency matters!
bench erformance	Performance	experiment and measure pgbench is improving	do not assume!
tency ite & Limit orage otocol ombinations	Postgres	version sorted and flushed checkpoints	9.6!
stance alability	High	costs	network, parse & plan
scellaneous	RW load	ACID	$SSD \gg HDD$
rap-Up ssons	RO load	pg as a cache manager	SSD = HDD



Contributions

provided or provoked

Postgres Latency

F. Coelho

ntroduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance Scalability Miscellaneous

Conclusio

Wrap-Up Lessons Contributions

About Core	& Andres Freund
 sorted checkpoints flushed checkpoints 	
 expressions mixed and weighted scripts and builtins better statistics improved usability rate and limit load 	\set -b/-f@ <i>stddev, per script</i> -c/-j -P -R -L
debug	



Contributions

provided or provoked

Postgres Latency

F. Coelho

ntroduction

Subject Application Definitions pgbench

Performance

Connection Latency Rate & Limit Storage Protocol Combinations Distance

Scalability

Conclusio

Wrap-Up Lessons Contributions

About Core	& Andres Freund
sorted checkpointsflushed checkpoints	
About pgbench	& Robert Haas
expressions	$\setminus \texttt{set}$
mixed and weighted scripts and builtins	-b/-f@
better statistics	stddev, per script
improved usability	-c/-j -P
rate and limit load	-R -L
debug	



Measuring and Reducing Postgres Transaction Latency

Fabien Coelho

MINES ParisTech, PSL Research University

pgDay Paris – March 23, 2017