

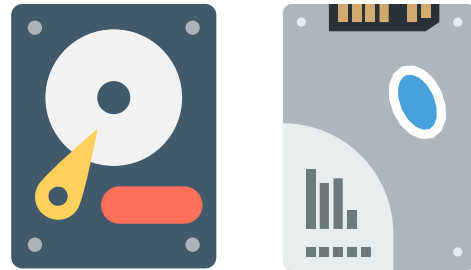
Make Your PostgreSQL 10x Faster on Cloud in Minutes

Sangwook Kim

app^osha

Storage Diversity

All storages are not created equal



≠



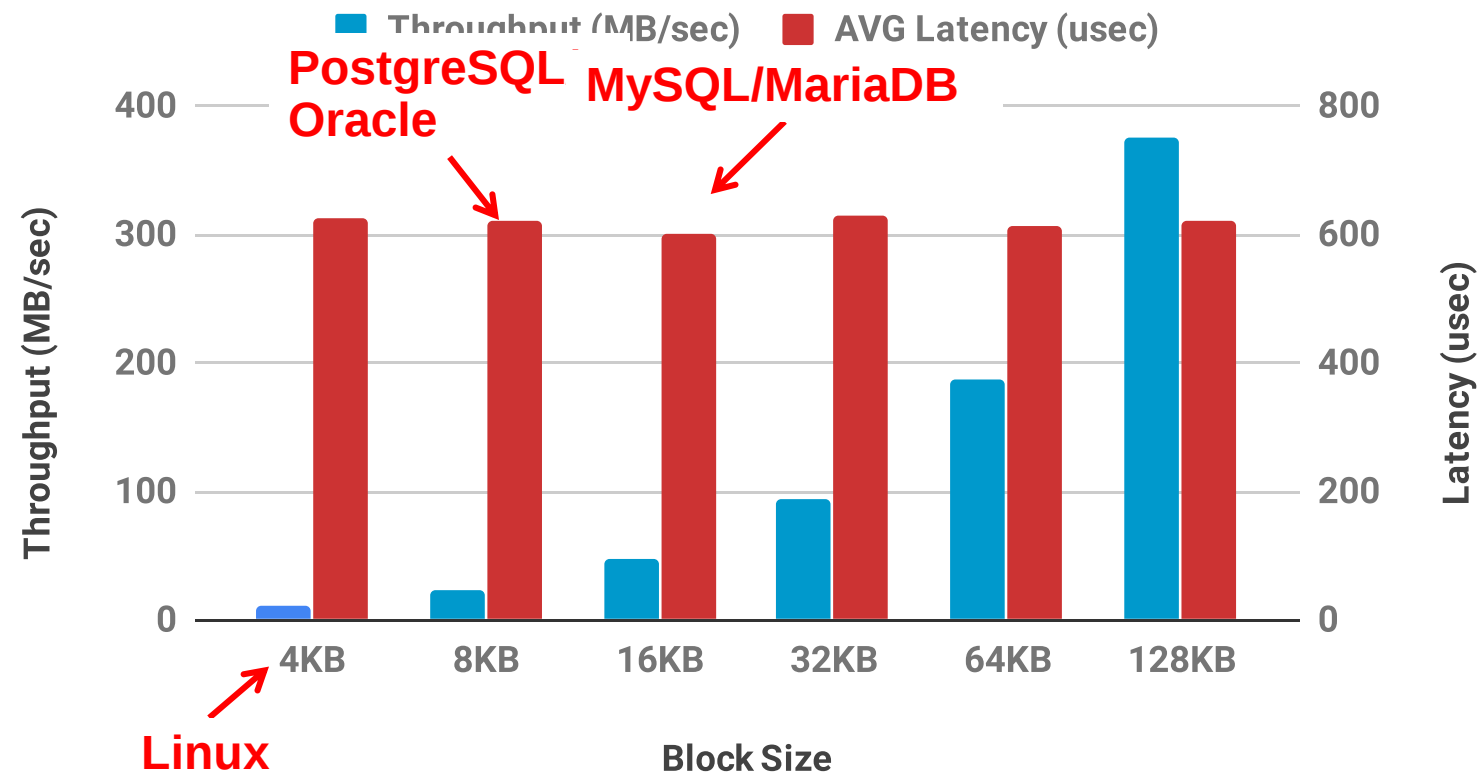
Own H/W
HW-controlled

Pay-per-use
SW-controlled

Example: AWS SSD Storage

I/O below 256 KiB consumes one IOPS

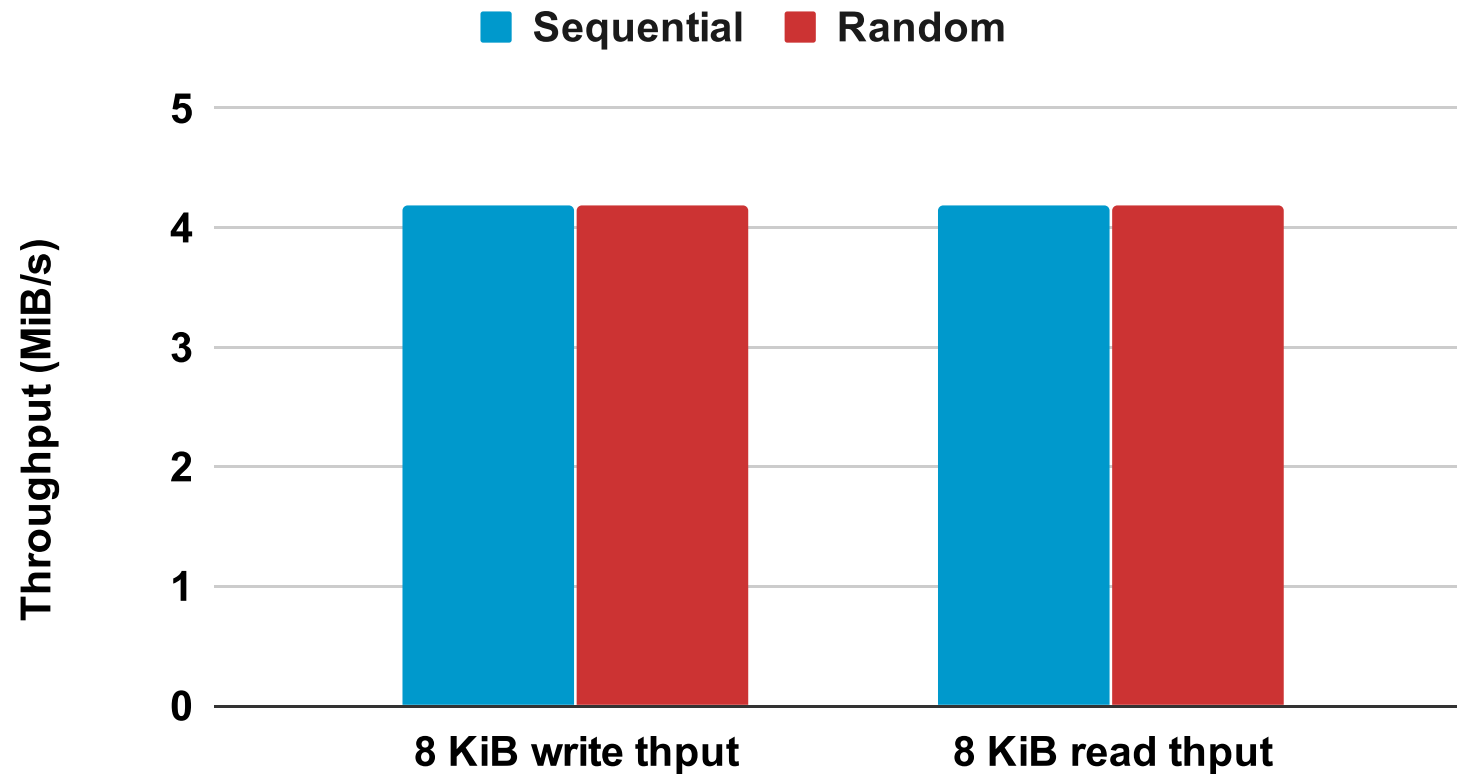
FIO random I/O test with 3000 IOPS AWS EBS



Example: AWS HDD Storage (1)

Sequential and random I/O show identical performance

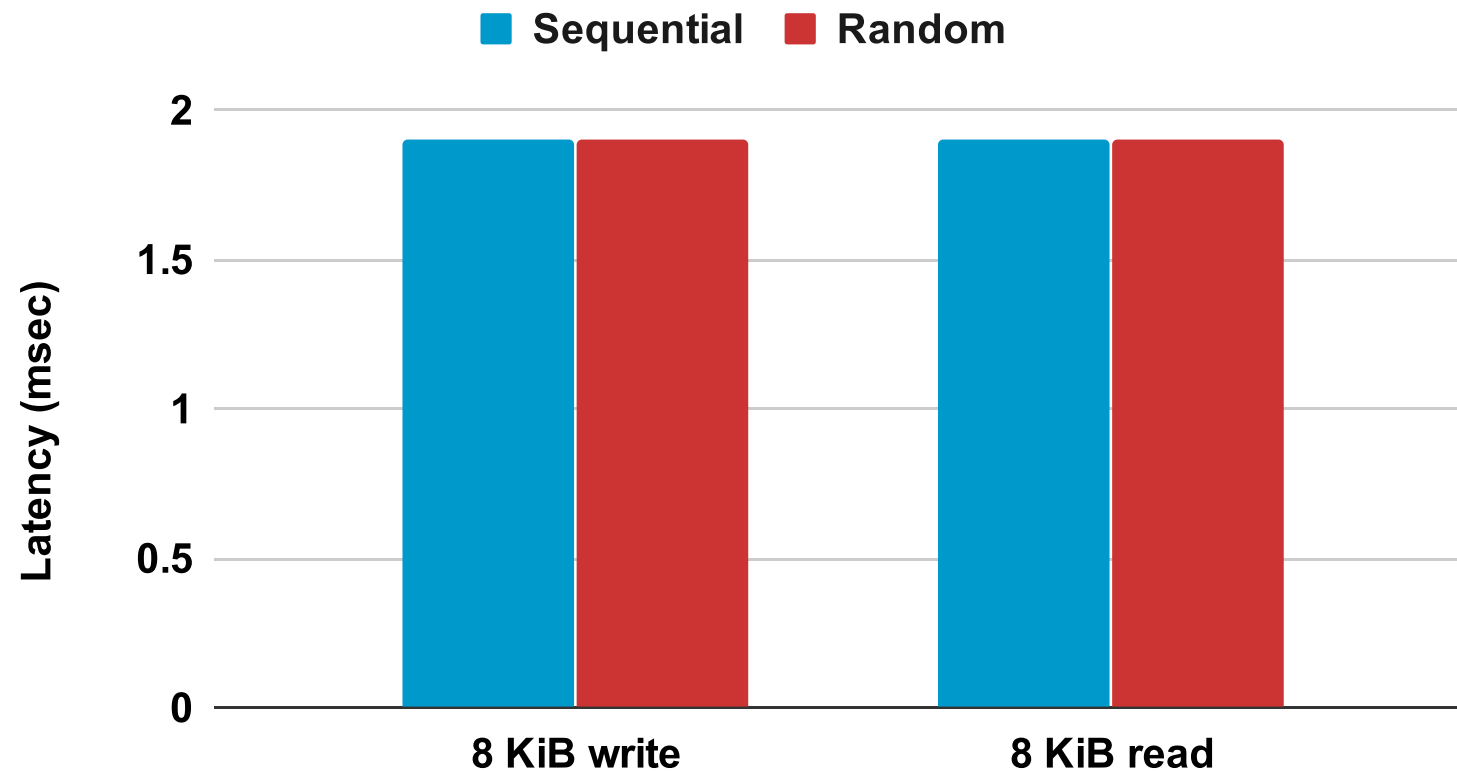
FIO throughput with 16 TiB AWS HDD EBS



Example: AWS HDD Storage (2)

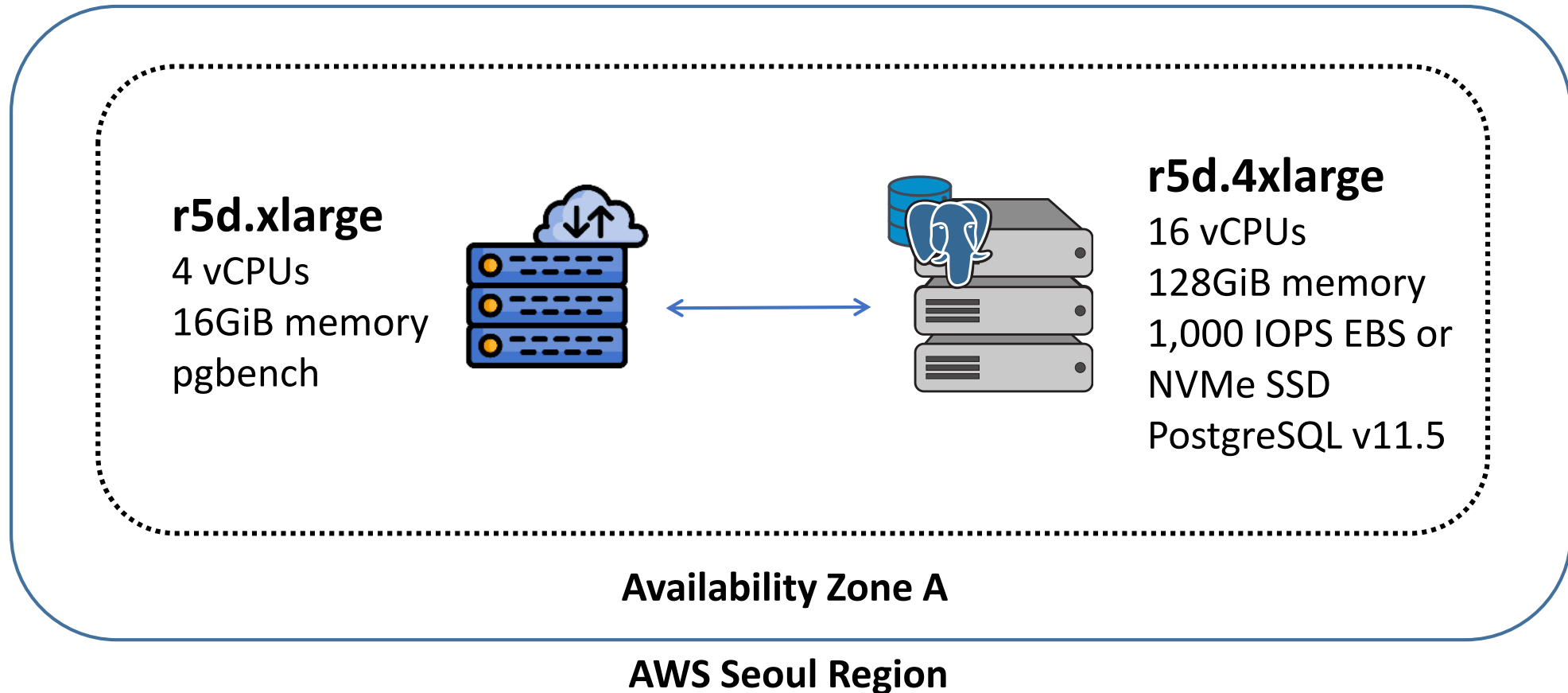
Sequential and random I/O show identical performance

FIO latency with 16 TiB AWS HDD EBS



Illustrative Experiment (1)

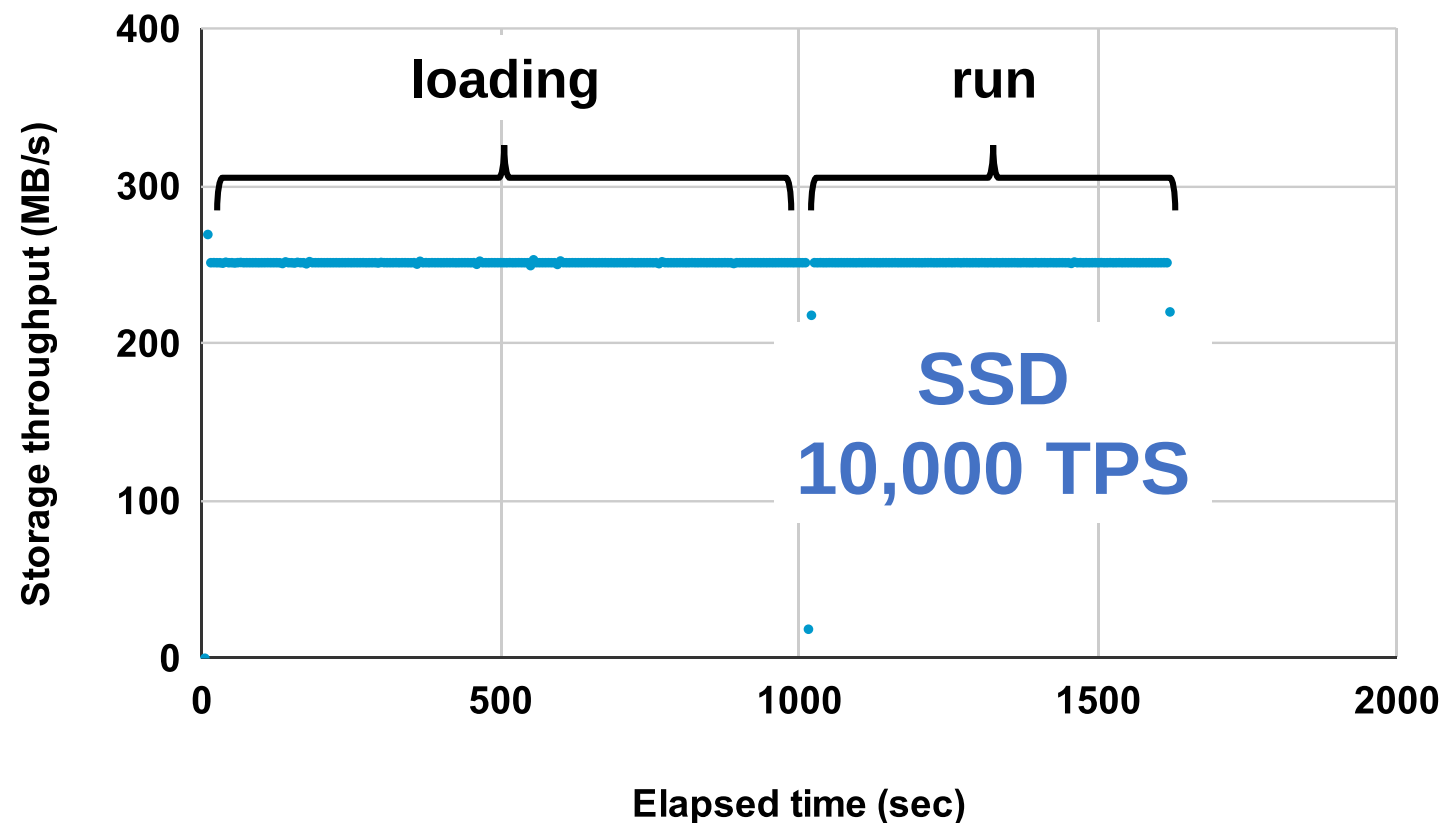
Test setup on AWS cloud



Illustrative Experiment (2)

PostgreSQL can fully utilize SSD

pgbench, 5K scale, 100 clients, 300 GiB NVMe SSD



Illustrative Experiment (3)

PostgreSQL cannot fully utilize cloud storage

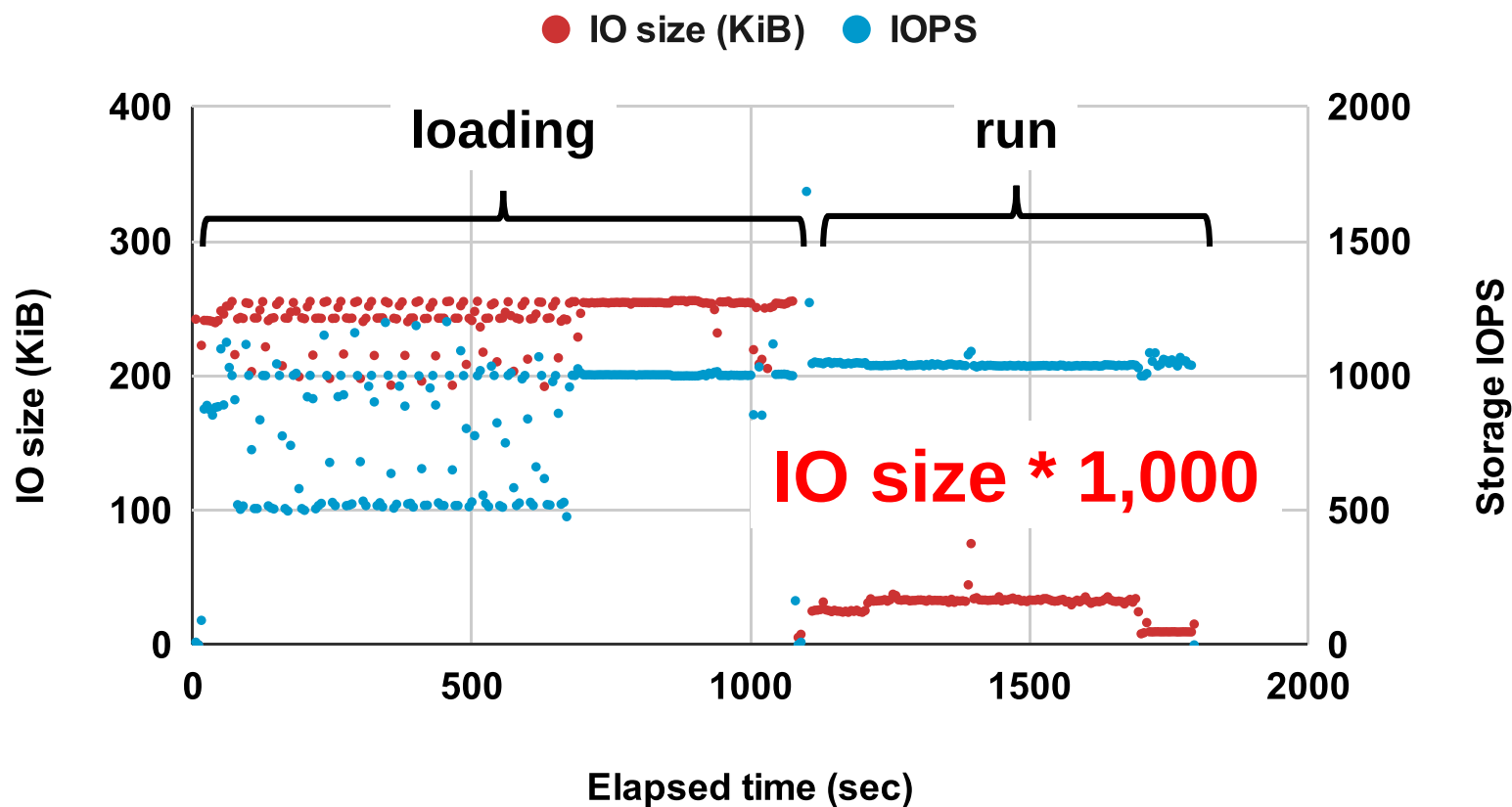
pgbench, 5K scale, 100 clients, 300 GiB EBS SSD (1K IOPS)



Illustrative Experiment (4)

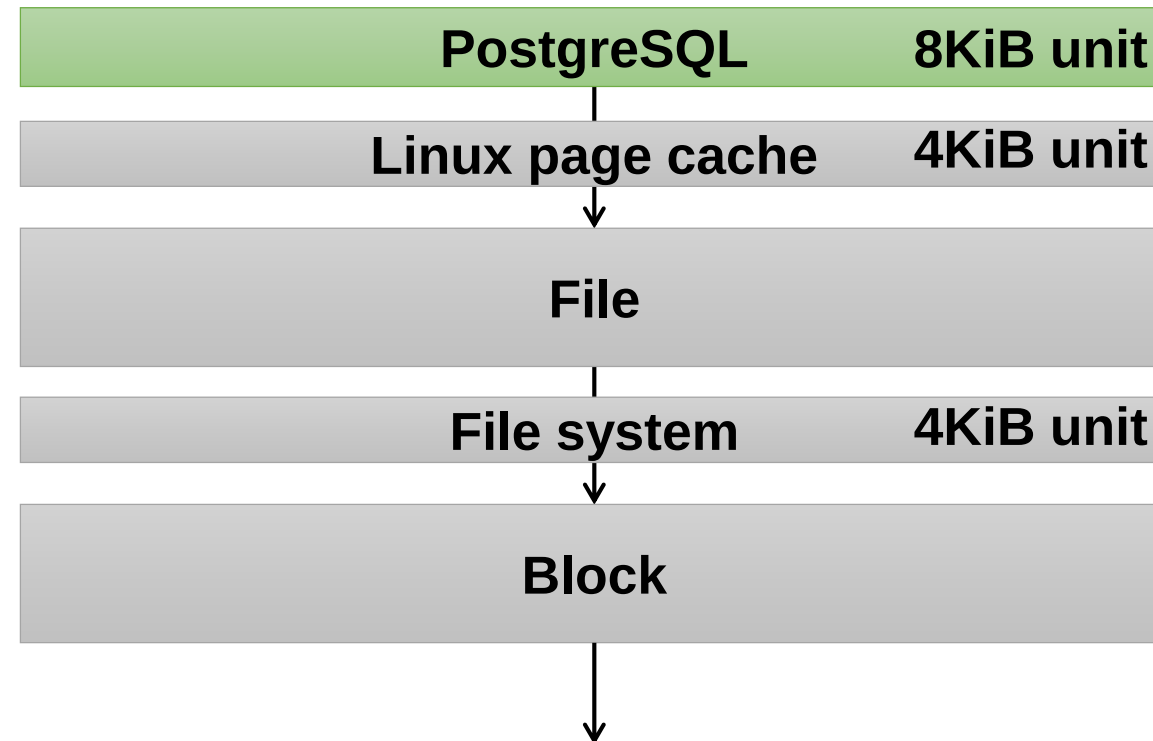
PostgreSQL cannot fully utilize cloud storage

pgbench, 5K scale, 100 clients, 300 GiB EBS SSD (1K IOPS)



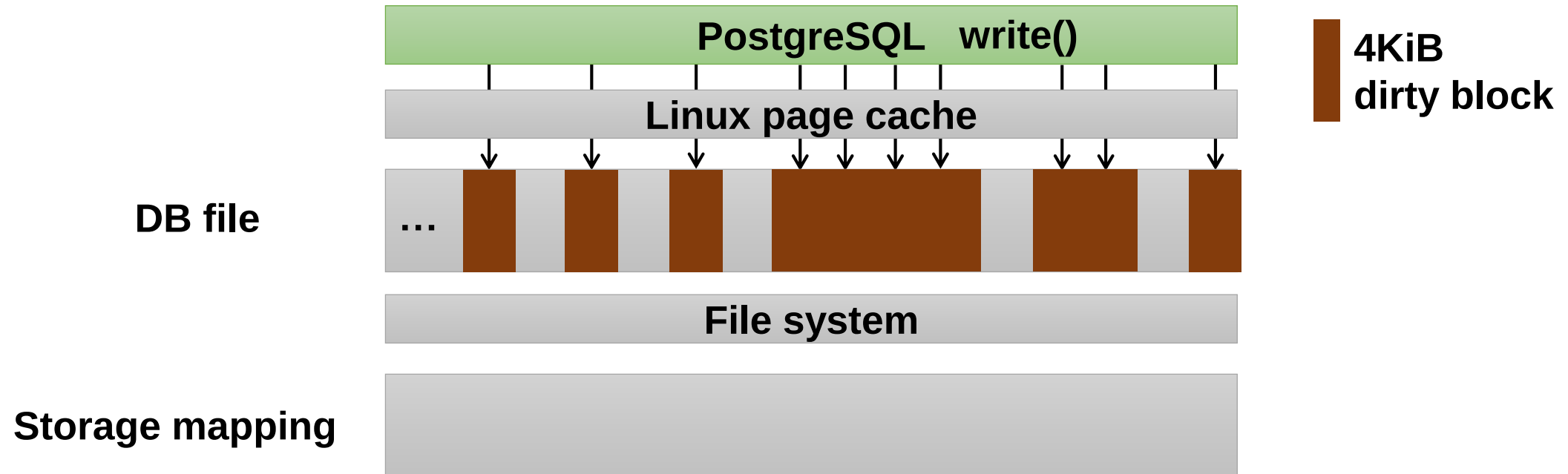
What's the Problem? (1)

Linux issues I/Os in 4 KiB unit



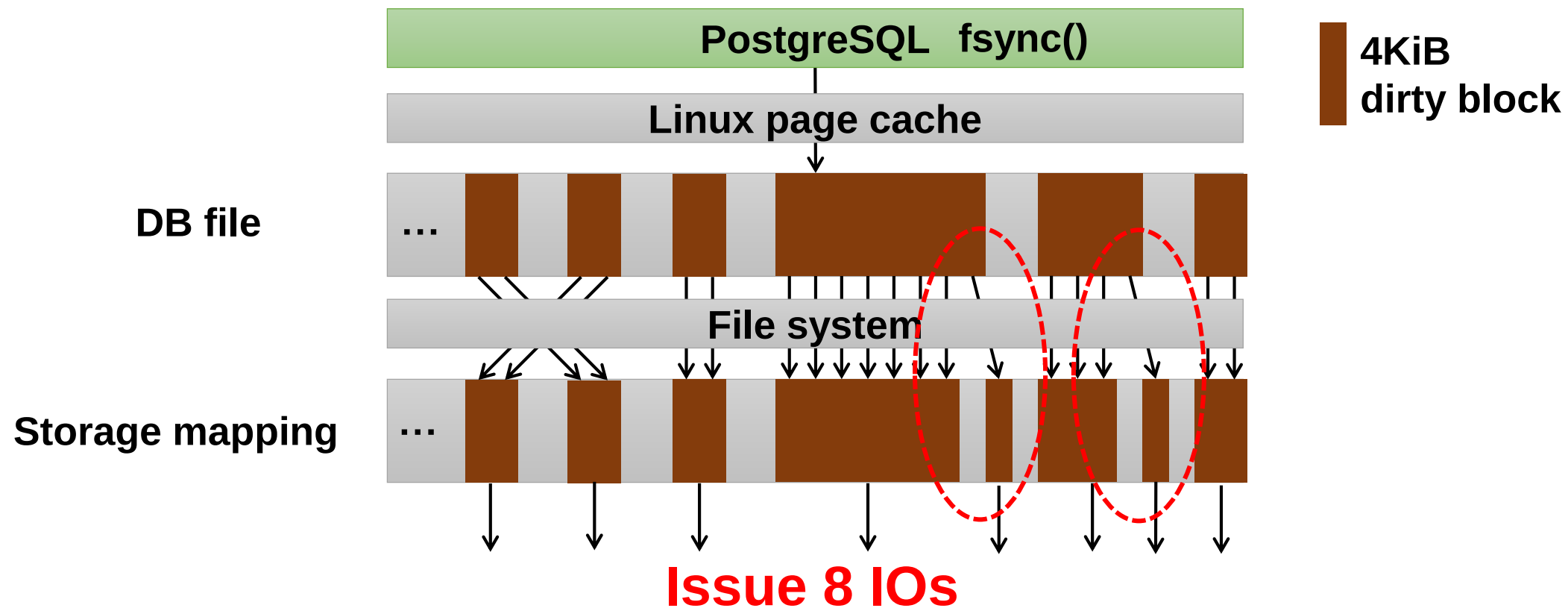
What's the Problem? (2)

Linux issues I/Os in 4 KiB unit



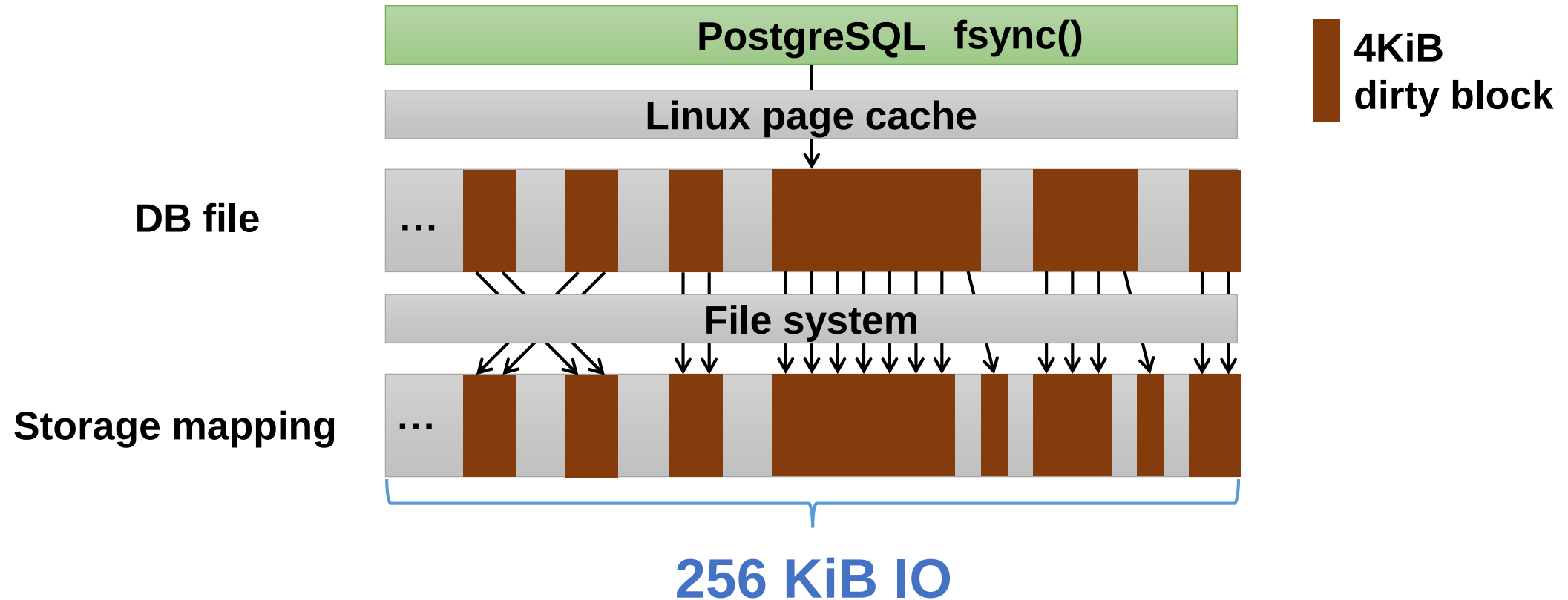
What's the Problem? (3)

Linux issues I/Os in 4 KiB unit



What's the Problem? (4)

What if 8 IOs can be a single IO?



What's the Problem? (5)

Each cloud has different policy for its storage



	IOPS	Thput	Note
ebs-gp2	3/GiB	250	Burstable
ebs-io1	\$67/1K/M	500	
ebs-st1	500	40/TiB	Burstable
ebs-sc1	250	12/TiB	Burstable



	IOPS	Thput	Note
ultra	\$49/1K/M	\$1/MB/M	
pre-ssd	20,000	900	Burstable
std-ssd	6,000	750	Pay # IOs
std-hdd	2,000	500	Pay # IOs

■ ■ ■



Google Cloud

	IOPS	Thput	Note
pd-ssd	30/GiB	0.48/GiB	
pd-ssd-16kb	30/GiB	0.48/GiB	16 KiB
pd-std	1.5/GiB	0.12/GiB	
pd-std-16kb	1.5/GiB	0.12/GiB	16 KiB

AppOS Extension

AppOS is a file engine built for diverse types of storage

Bare metal



Virtual machine



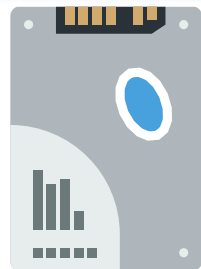
Container



Container

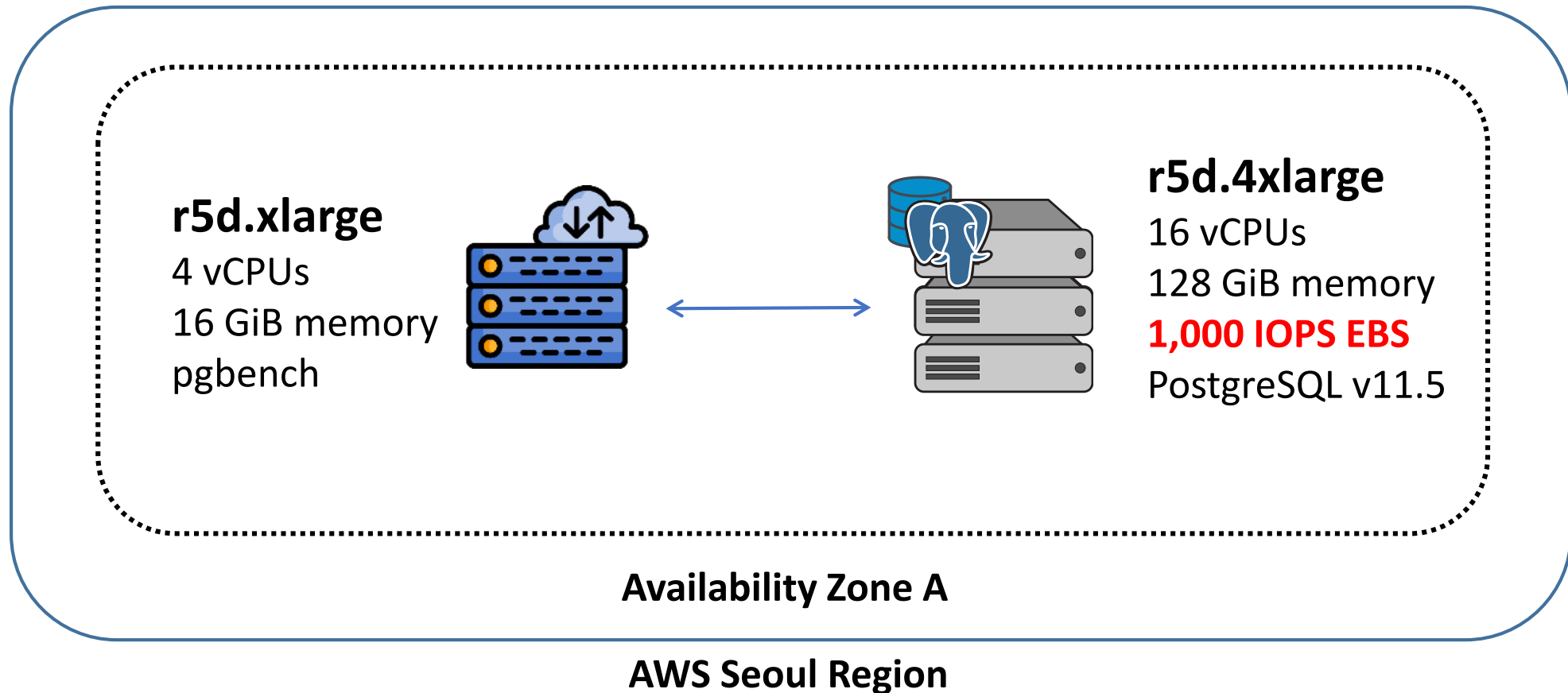


AppOS extension



Performance on AWS Storage (1)

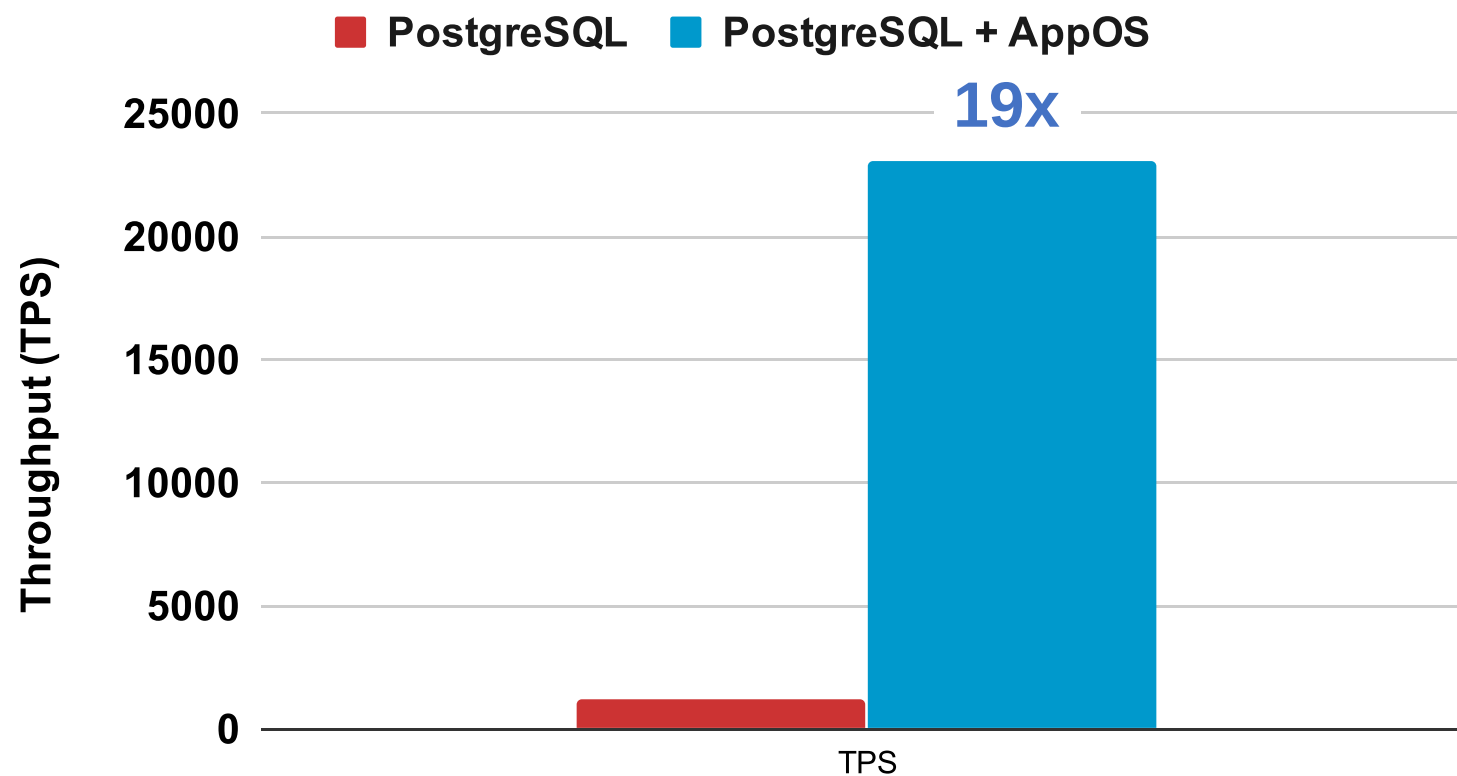
Test setup on AWS



Performance on AWS Storage (2)

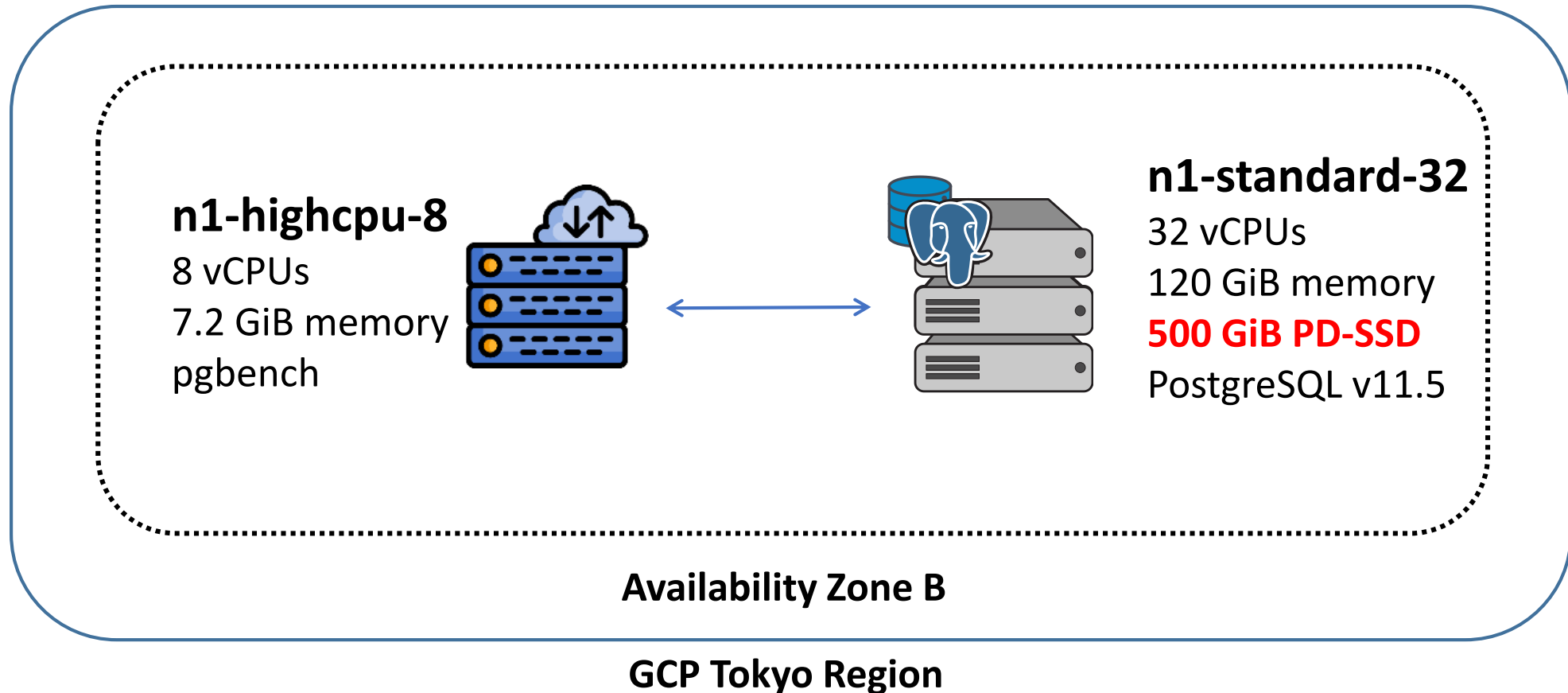
pgbench shows 19x better TPS with AppOS

pgbench, 5K scale, 100 clients, 300 GiB EBS (1K IOPS)



Performance on GCP Storage (1)

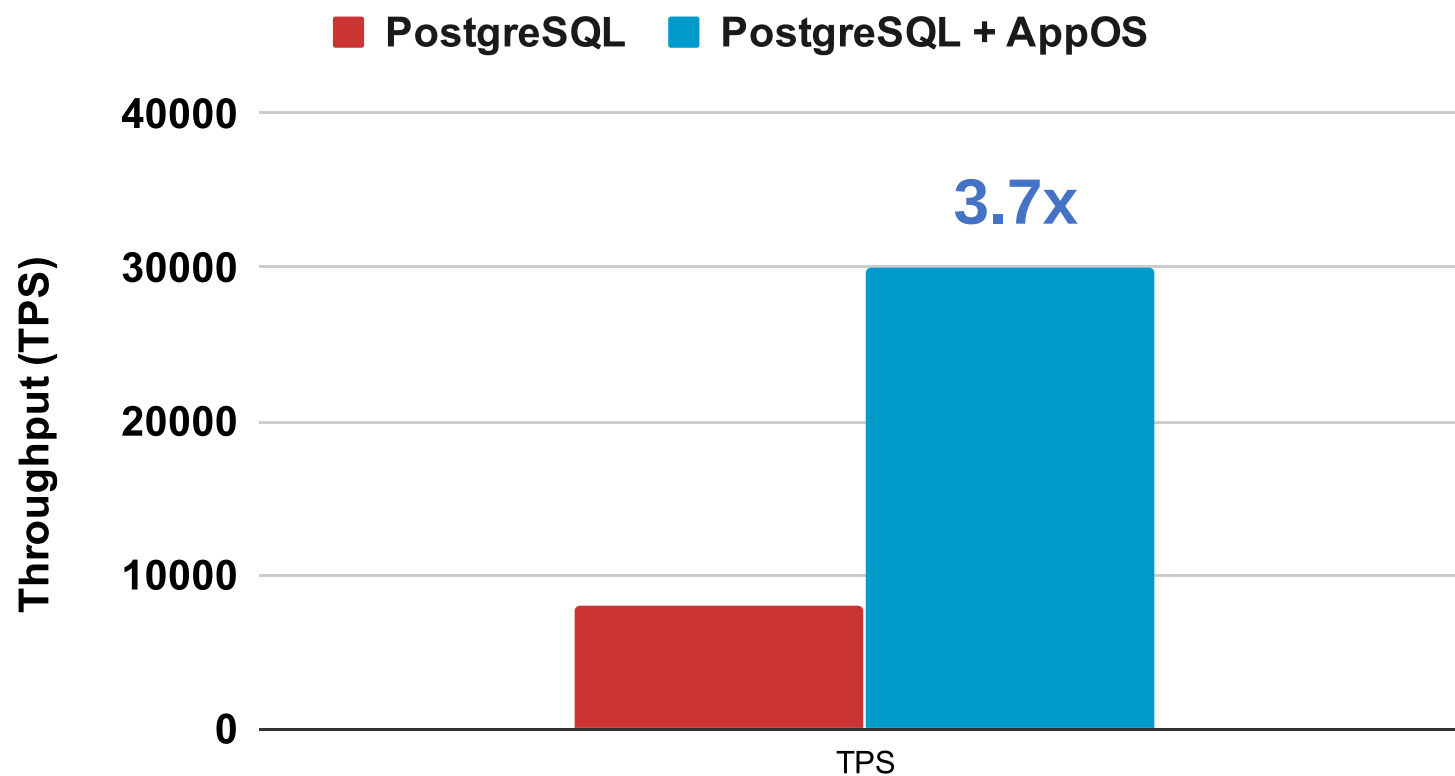
Test setup on GCP



Performance on GCP Storage (2)

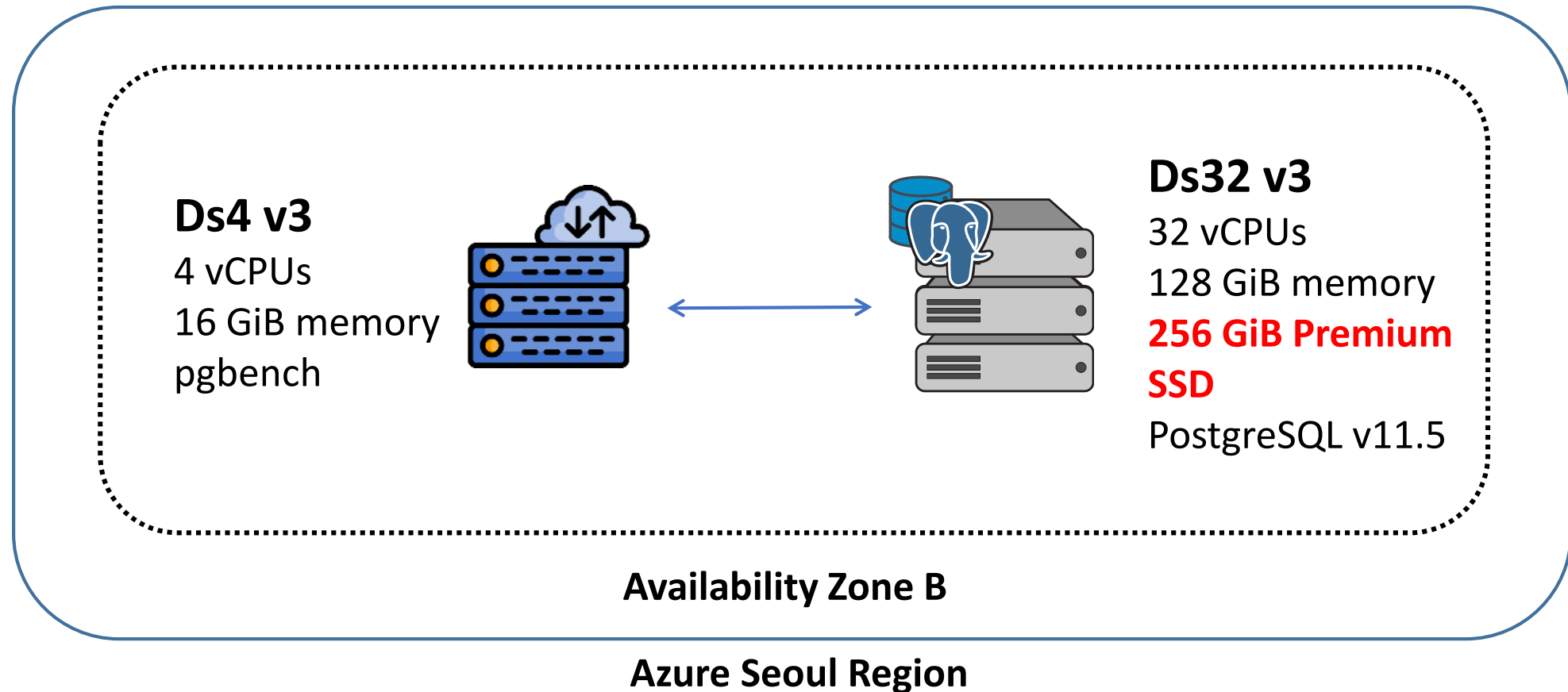
pgbench shows 3.7x better TPS with AppOS

pgbench, 5K scale, 100 clients, 500 GiB PD SSD



Performance on Azure Storage (1)

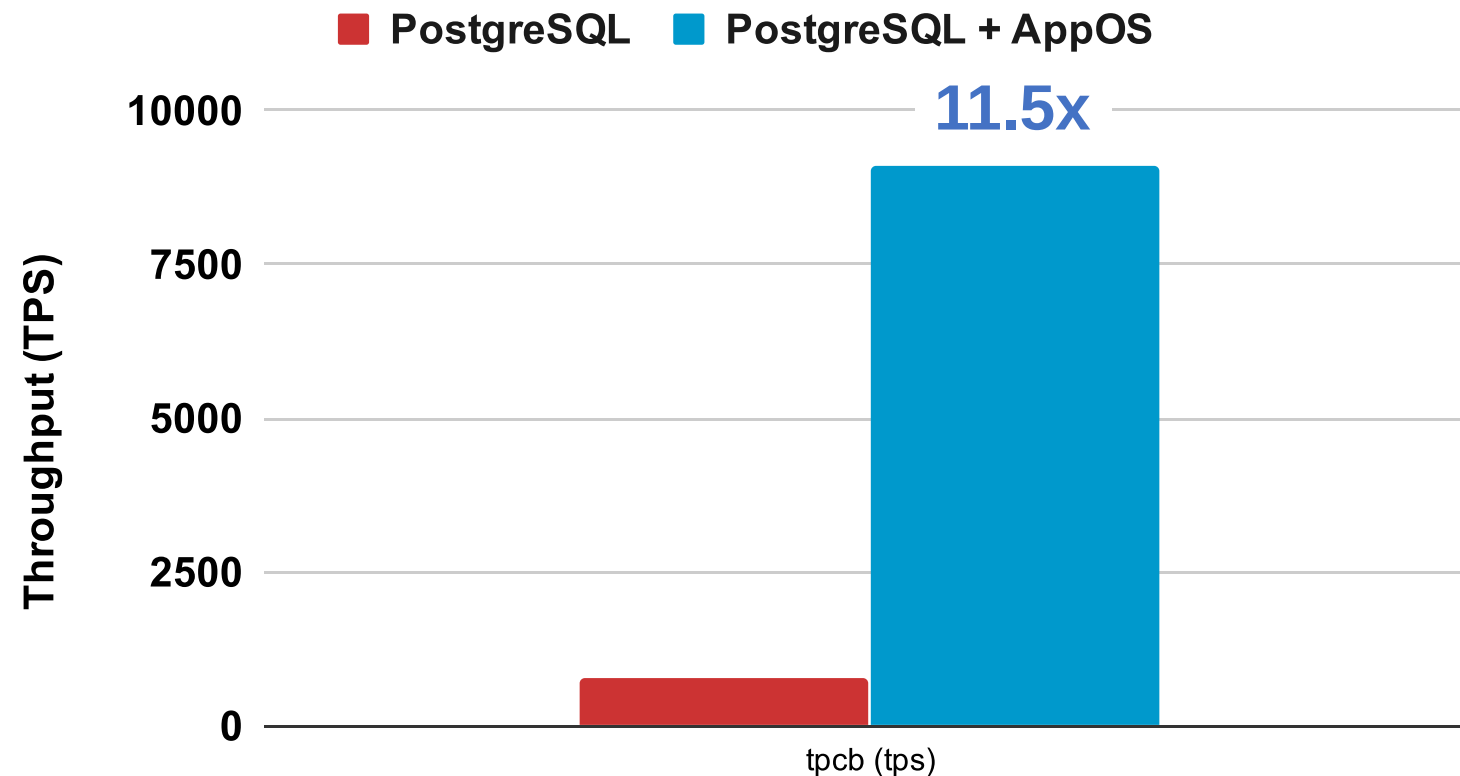
Test setup on Azure



Performance on Azure Storage (2)

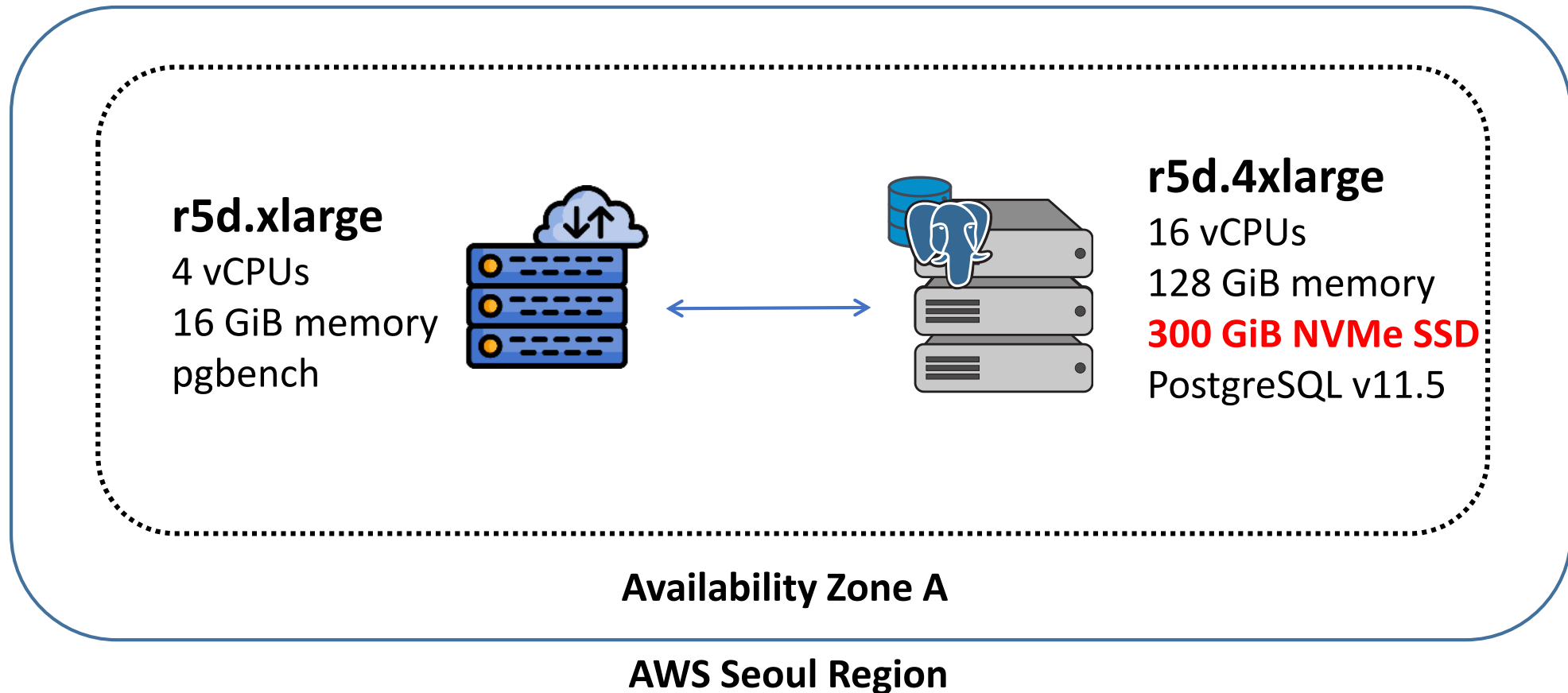
pgbench shows 11.5x better TPS with AppOS

pgbench, 5K scale, 100 clients, 256 GiB Premium SSD



Performance on SSD (1)

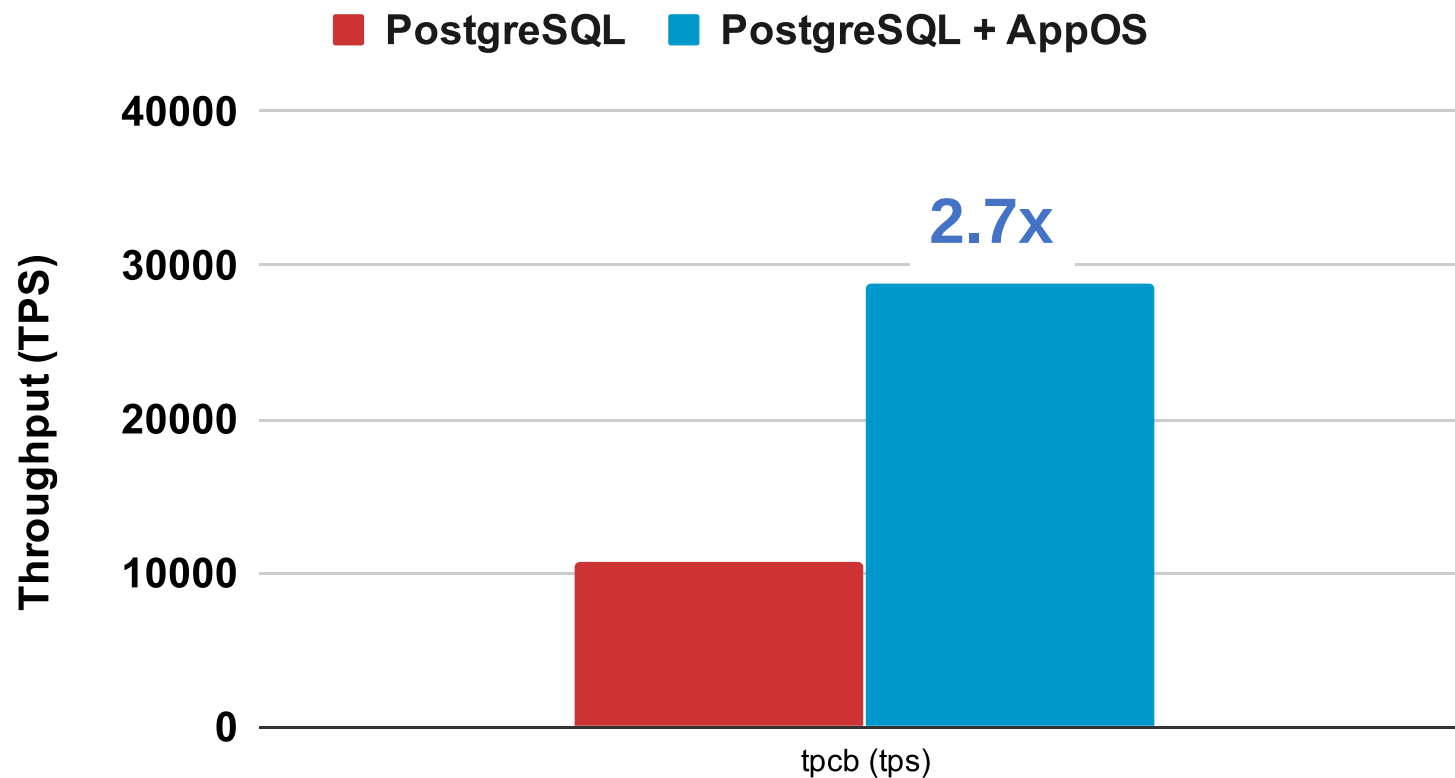
Test setup on AWS



Performance on SSD (2)

Pgbench shows 20x better TPS with AppOS

pgbench, 5K scale, 100 clients, 300 GiB NVMe SSD



AppOS Design Goals

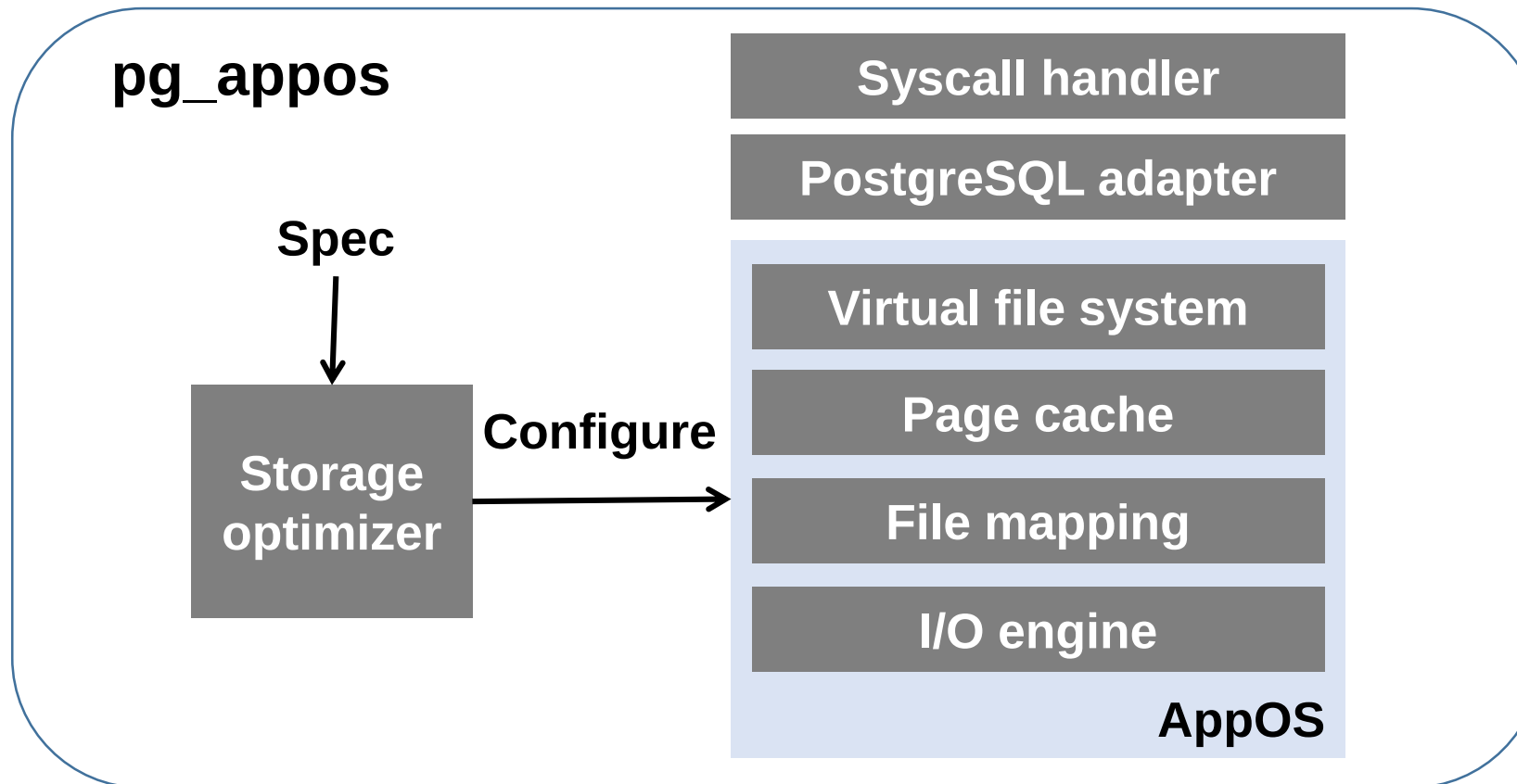
**No PostgreSQL
code changes**

**PostgreSQL-aware
I/O processing**

**Storage-aware
I/O processing**

AppOS Architecture

AppOS extension is a collection of libraries



AppOS Internals (1)

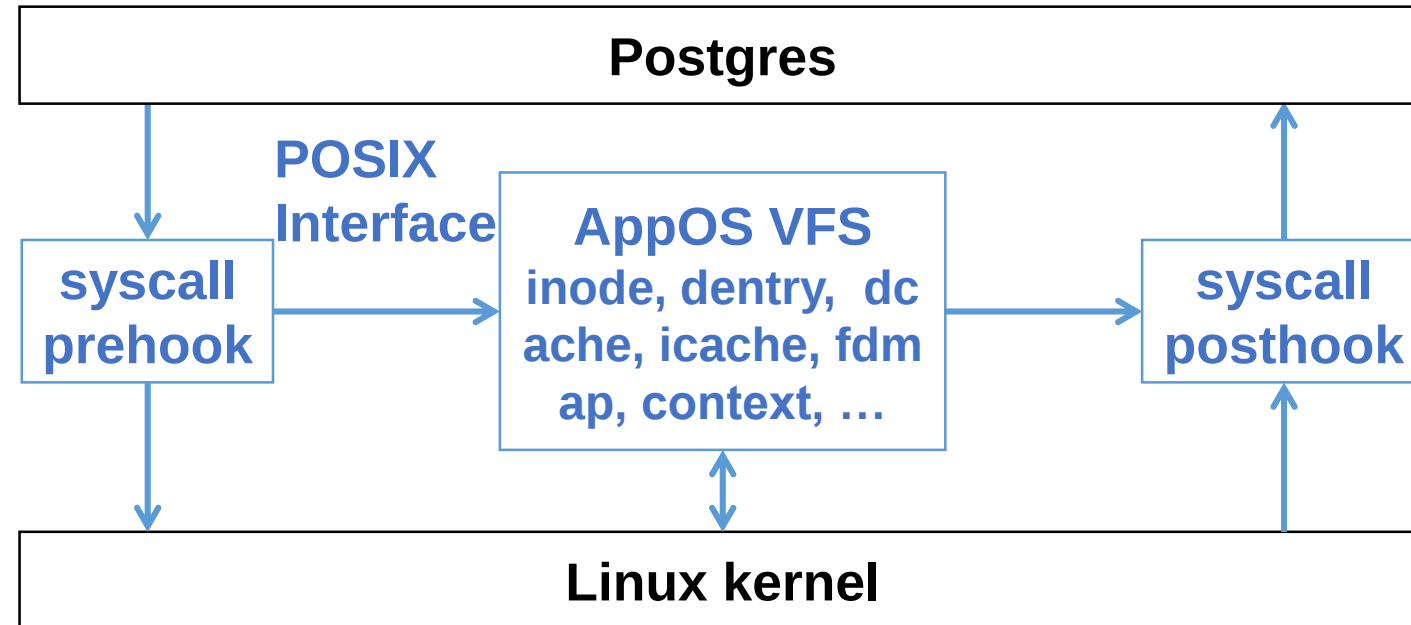
AppOS does not require code changes

Physical Virtual

Ring 3 Ring 3

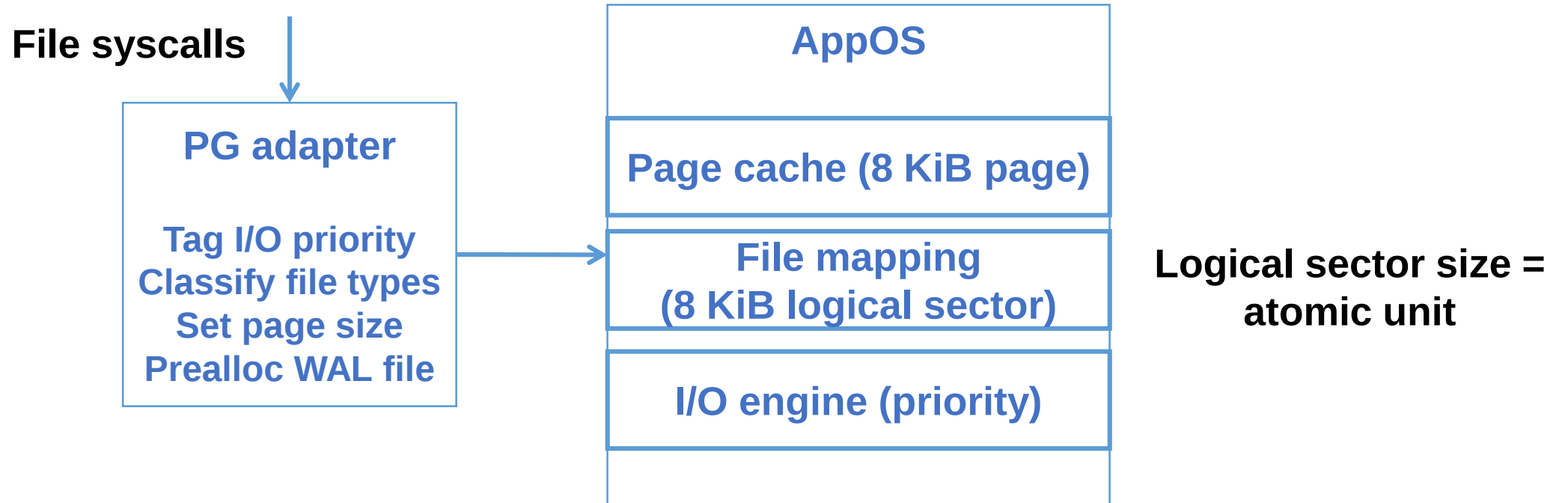
Ring 0

Ring 0



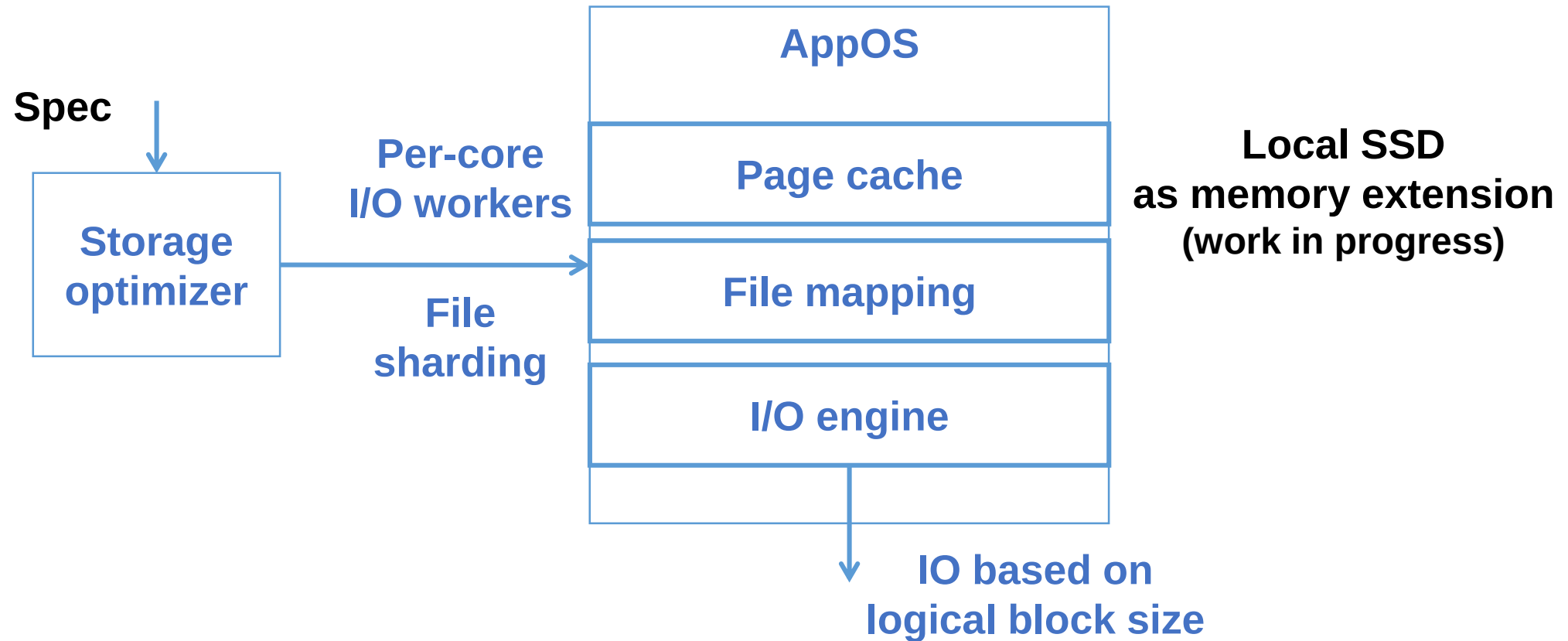
AppOS Internals (2)

AppOS provides PostgreSQL-aware I/O processing



AppOS Internals (3)

AppOS provides storage-aware I/O processing



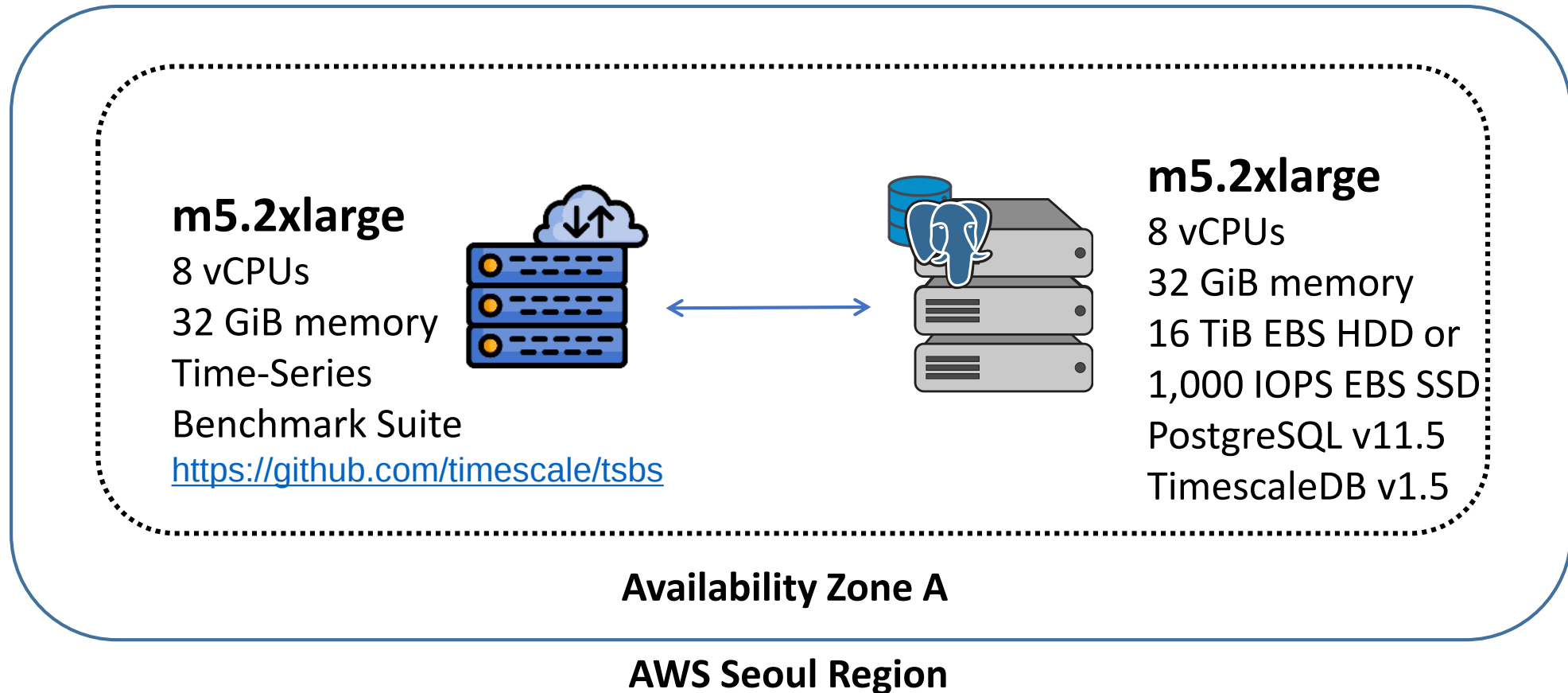
Extensibility

AppOS can seamlessly work with PostgreSQL-derived DBs



TimescaleDB Performance (1)

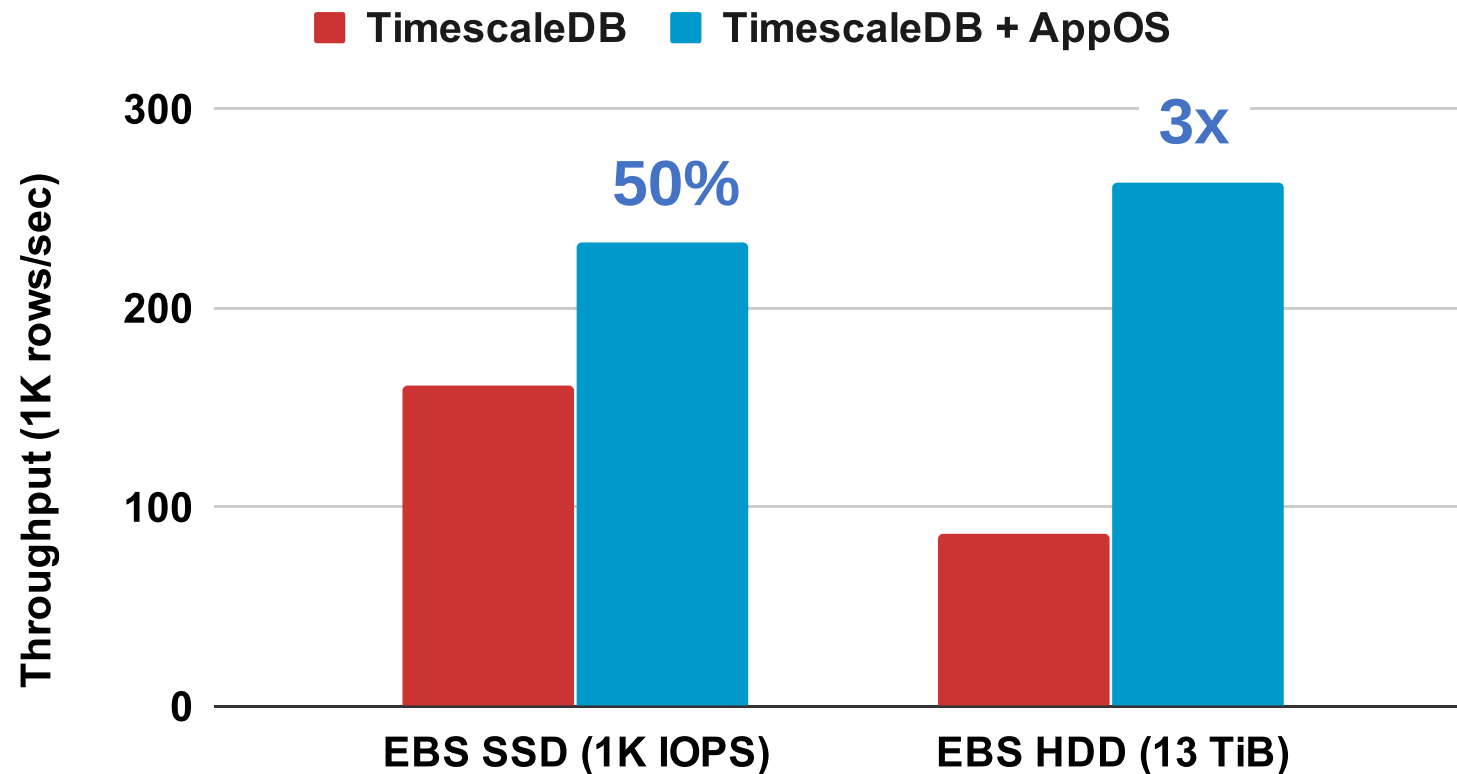
Test setup on AWS



TimescaleDB Performance (2)

TimescaleDB inserts using TSBS on AWS

TimescaleDB inserts throughput for 1B rows





apposha



<https://apposha.io>
shawn@apposha.io