

Pick the right PostgreSQL for a **10X** difference

A price-performance analysis of PostgreSQL 15 on five clouds



Google Cloud



Tencent
Cloud



HUAWEI



NineData



玖章算术(浙江)科技有限公司

www.ninedata.cloud

Shopping for Cloud Database – still the old fashion way

Surprise, Surprise!!! In the same US region, the same PostgreSQL, **10X price-perf difference**

The cutting-edge cloud business still run like travel agents 30 years ago!

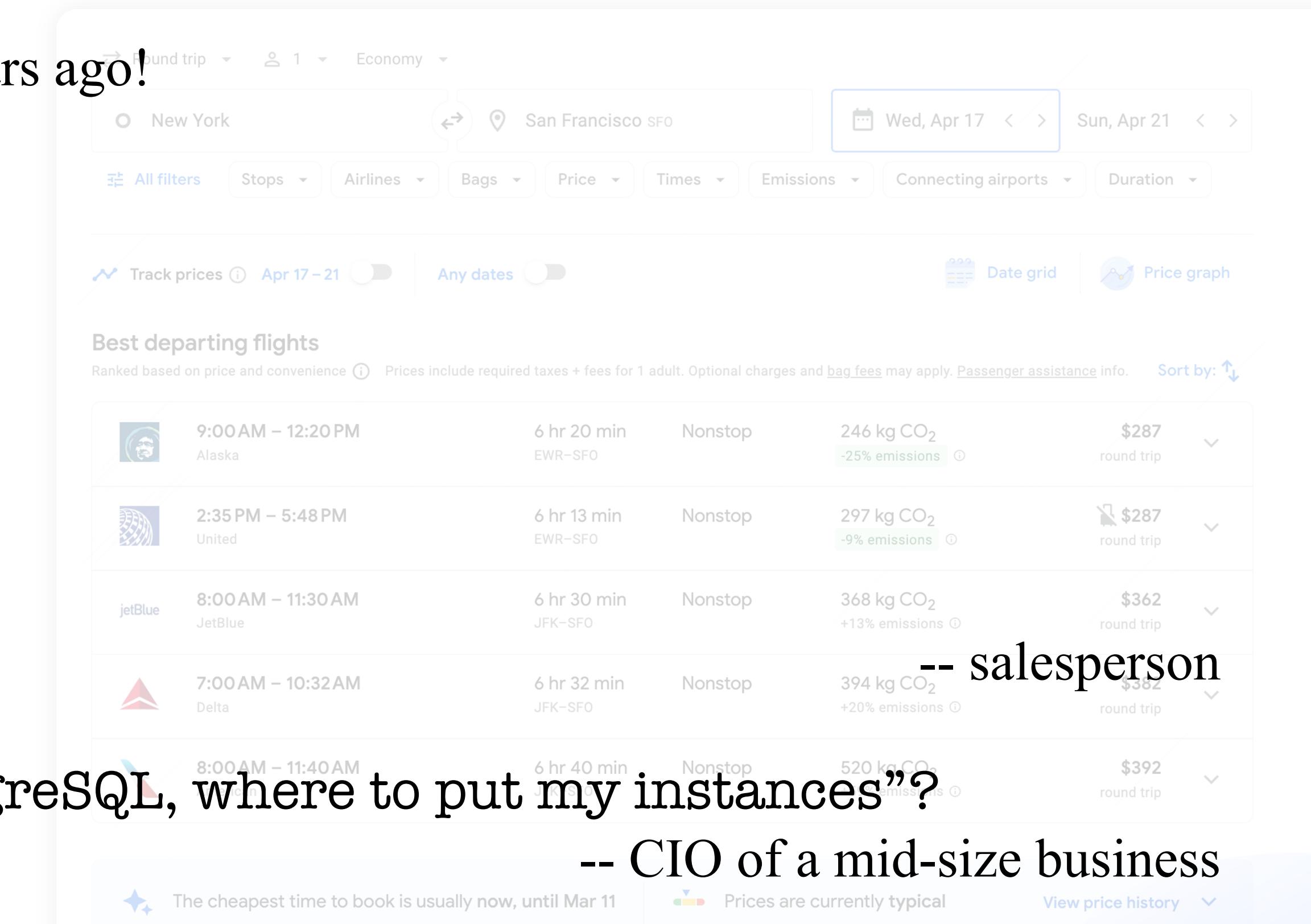
Where is my “**google flights**” app for Database shopping?

“My (cloud-based) PostgreSQL is better than others.”

“My application will run on cloud, a linux EC2 + a PostgreSQL, where to put my instances?”

“Save up to 52 percent on your database compute costs with ... ”

-- the title of a blog about a “fully managed, PostgreSQL-compatible database service”



The Game is ready for the Change – just like the airline business matured

Standardized Product

While the premium seats(Oracle, Aurora, AlloyDB) needs agent's help, standard seats(open-source RDS, MySQL, postgreSQL) are pretty much the same.

Price is “transparentized”

A side-effect to Cloud Database. The cost must be transparent to void the cost of salesperson.

Easy to jump the ship with multi-cloud strategy

No one like to be locked in. Not the insurance company, not the airline, not the small-mid tier business running on Database. Cloud Database provides understand a lock will hurt their business and embrace the multi-cloud world.

“Cloud switching just got easier: Removing data transfer fees when moving off Google Cloud”

- Amit Zavery, GM/VP, Head of Platform, Google Cloud

Case Study – PostgreSQL 15 on Clouds

[Use Case]

A CIO decides to use a compute engine(such as EC2) and a RDS (such as PostgreSQL) for one of his company's business division, which is a small-middle size e-commerce. The workload shows typical OLTP patterns and the budget is \$10K/year.

[Setup]

Database: PostgreSQL 15 (OLTP for daily transactions and OLAP for monthly reporting)

Architect on Cloud: A Linux + a Database instance in U.S

Benchmark: Sysbench oltp_read_write

Run options: threads = 2, 4, 8, 16, 32, 64, 128 (256, 512), 10 tables with size 10M

Run report: TPS, P95 latency and CPU utilization

[Key Question]

Which Cloud Provider to choose? Alibaba Cloud, AWS, GCP, Huawei and Tencent (large airline carrier ≠ better choice)

Disclaimer:

The engineer/author responsible for this analysis is a former employee of both Huawei and Alibaba. They currently have no business affiliations with any of the cloud providers mentioned, aside from incurring costs or applying for free credits to conduct the benchmark tests. He was as surprised by the data as anyone else and had a hard time explaining the huge gap.

```
#!/bin/bash
...
...
for ((i=2; i<=128; i=i*2))
do
    echo "Thread = $i"
    echo "#####----- $(date) -----#####" >> "$ofile"
    echo "Thread = $i" >> "$ofile"
    sysbench --db-driver=pgsql --pgsql-host="$dbip" --
pgsql-port="$dbport" --pgsql-user="$dbuser" --pgsql-
password="$pw" --pgsql-db=t_sysbench --
table_size=10000000 --tables=10 --events=0 --time=600 -
-threads=$i --percentile=95 --report-interval=60
oltp_read_write run >> "$ofile"
done
```

Overview of Price Tag

PostgreSQL 15	Alibaba	AWS	GCP_EP	Huawei	Tencent
Monthly Cost(\$)	762.93	751.01	770.40	675.06	826.85
Region	US (Virginia)	us-east-1a (N. Virginia)	us-central1-c (Iowa)	LA-Mexico City1	US (Virginia)
HA	Yes(multi-zone)	No	No	Yes (same zone)	Yes (same zone)

Notes

All DB instances are 8CPU 32GB memory and 250 GB SSD (except for GCP_EP, GCP enterprise Plus, only offer 8C64GB model)

- High availability:** AWS and GCP's instances were single-DB instance(otherwise, cost will be double); Alibaba, Huawei and Tencent are Primary-Secondary HA(the cost will be half if choose single-DB instance). This factor alone put AWS and GCP at twice expensive comparing to Alibaba Huawei and Tencent.
- DB model:** 8C32GB is chosen so that concurrency can scale up to 64~128 threads in sysbench before performance degradation. In another word, the DB instance should be sufficient for most small business and some mid-tier enterprise too.
- Region:** the exercise is focused on regions in US. However, Huawei doesn't offer cloud service in the States, so Mexico City is picked.
- Pay-per-use:** all monthly costs calculated based on hourly on-demand price model.
- Discounted Price:** some cloud providers often discounted price as part of sales strategy. For example, at the time of benchmark, Alibaba Cloud offers discount at 78%, so the actual cost was \$595.09/month. On one hand, Alibaba always carries some kind of discount both for pay-as-you-go(in this case) and subscription by month or years, so discounted price is indeed normal price; on another hand, there is no guarantee of the discounted price for pay-as-you-go. Because this exercise is to provide a guideline for performance and business planning, the full price is used to avoid un-planned cost spike. Hence, \$762.93/month is used.
- Exchange rate:** 7.21(CNY:USD = ¥ 7.21 : \$1) as of mid of January, 2024.

Three Dimension of Evaluation

Performance centric

- sysbench's oltp_read_write for PostgreSQL 15
- TPS and P95 latency in high concurrent environment(scale from 2 to 128, a couple cases up to 512)
- CPU utilization and latency as indicators for workload stability.

Price/Performance Ratio

- Database instances **monthly cost** between \$650~850
- Powerful enough for **small business and some mid-tier** enterprises.
- A '**TPS/month-cost**' metric to gauge the output efficiency of a production system.
- **Impossible triangle effect**: monthly cost, DB instance(CPU/Memory) and HA. GCP and AWS's environments are single node instead of HA setup.

User friendly

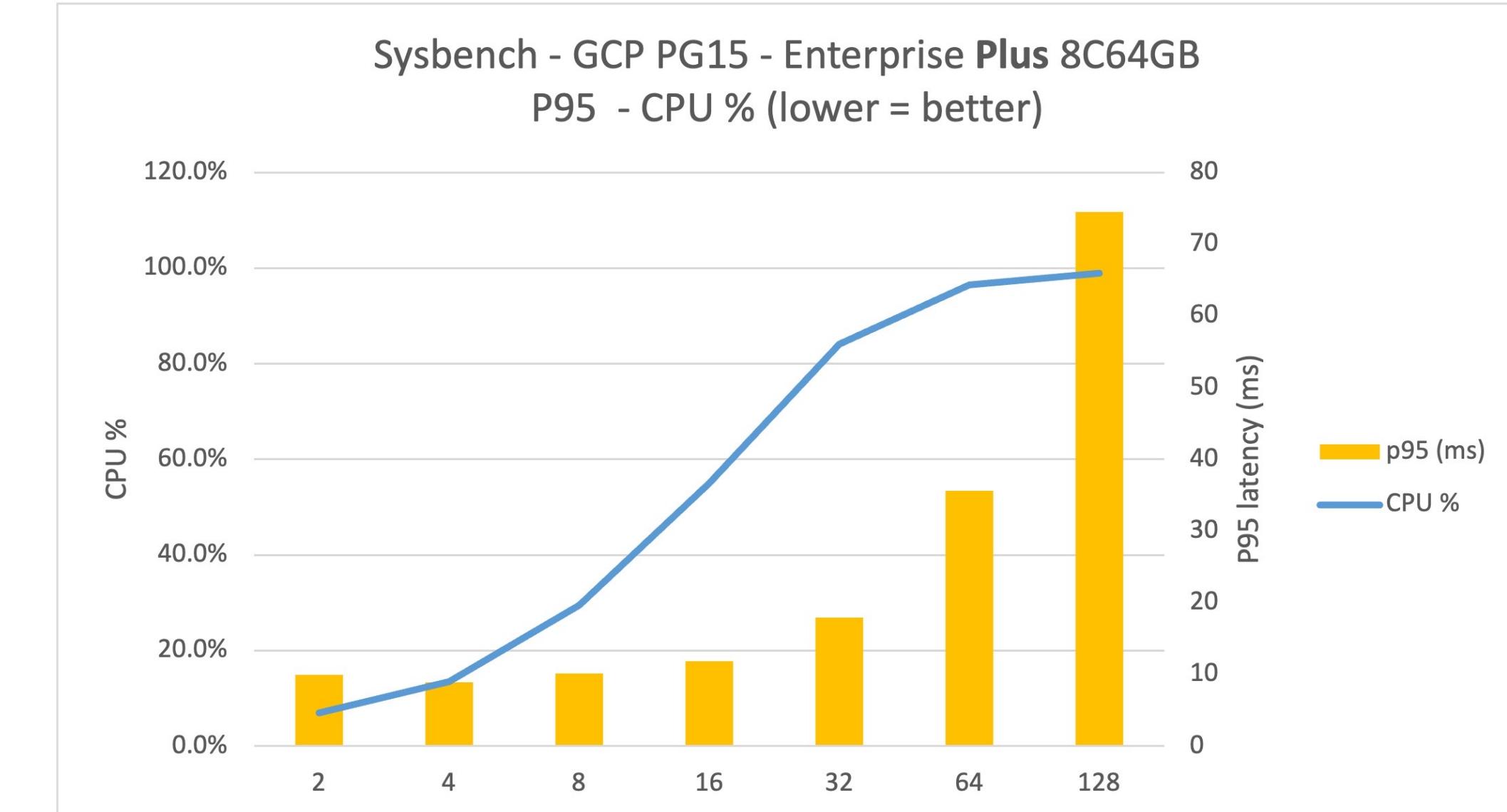
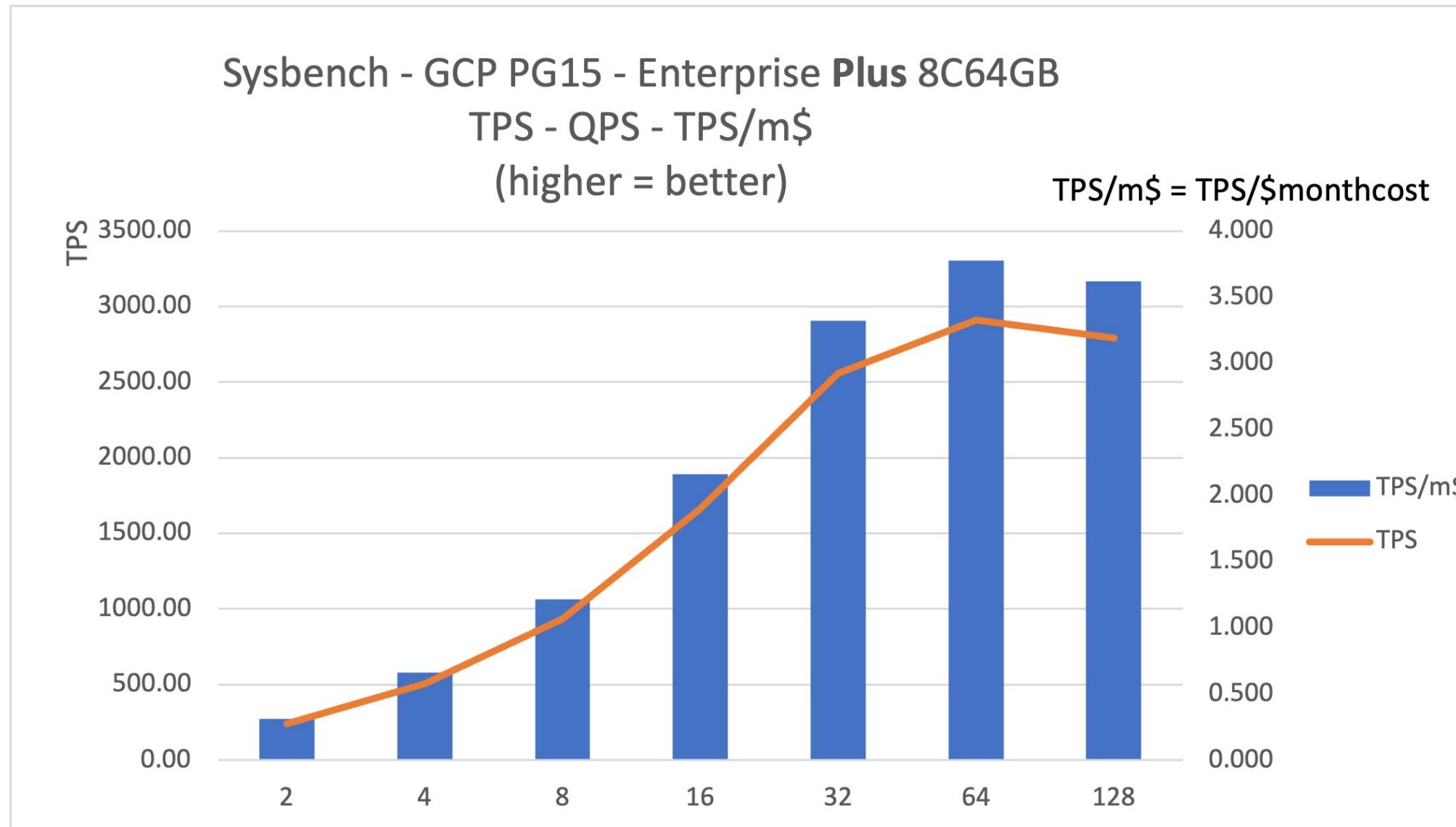
, an subjective evaluation with human aspect and skill-level in mind.

Assume the user is an entry-level Admin with knowledge. The report also includes observations during the performance analysis. While challenging to quantify, feedback from this evaluation is crucial to identify and eliminate the least user-friendly options.



GCP – A reliable large company with reasonable price tag

Google Cloud



- Concurrency scale to
- Highest TPS
- **Sweat spot** for productivity
- TPS/monthcost @ sweat spot

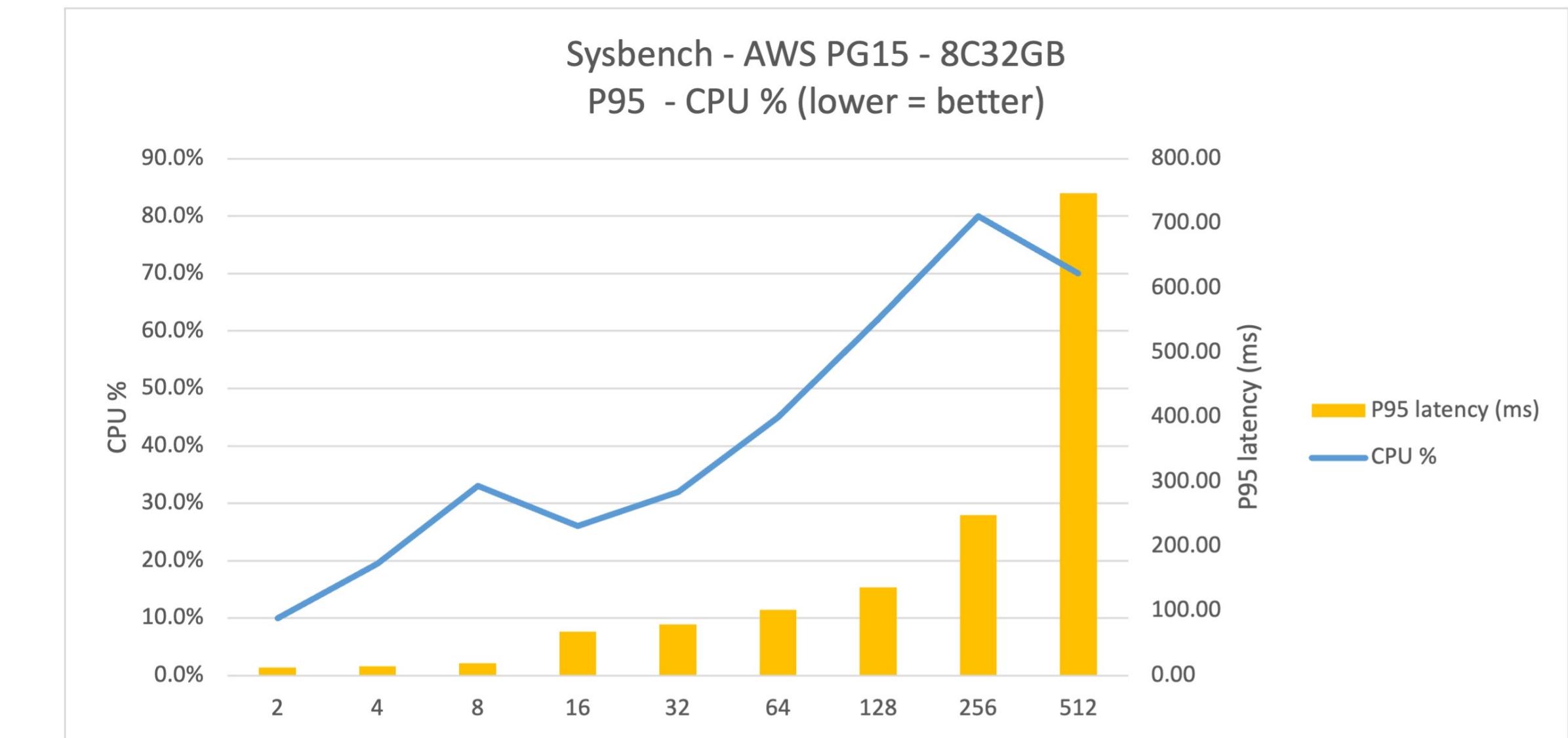
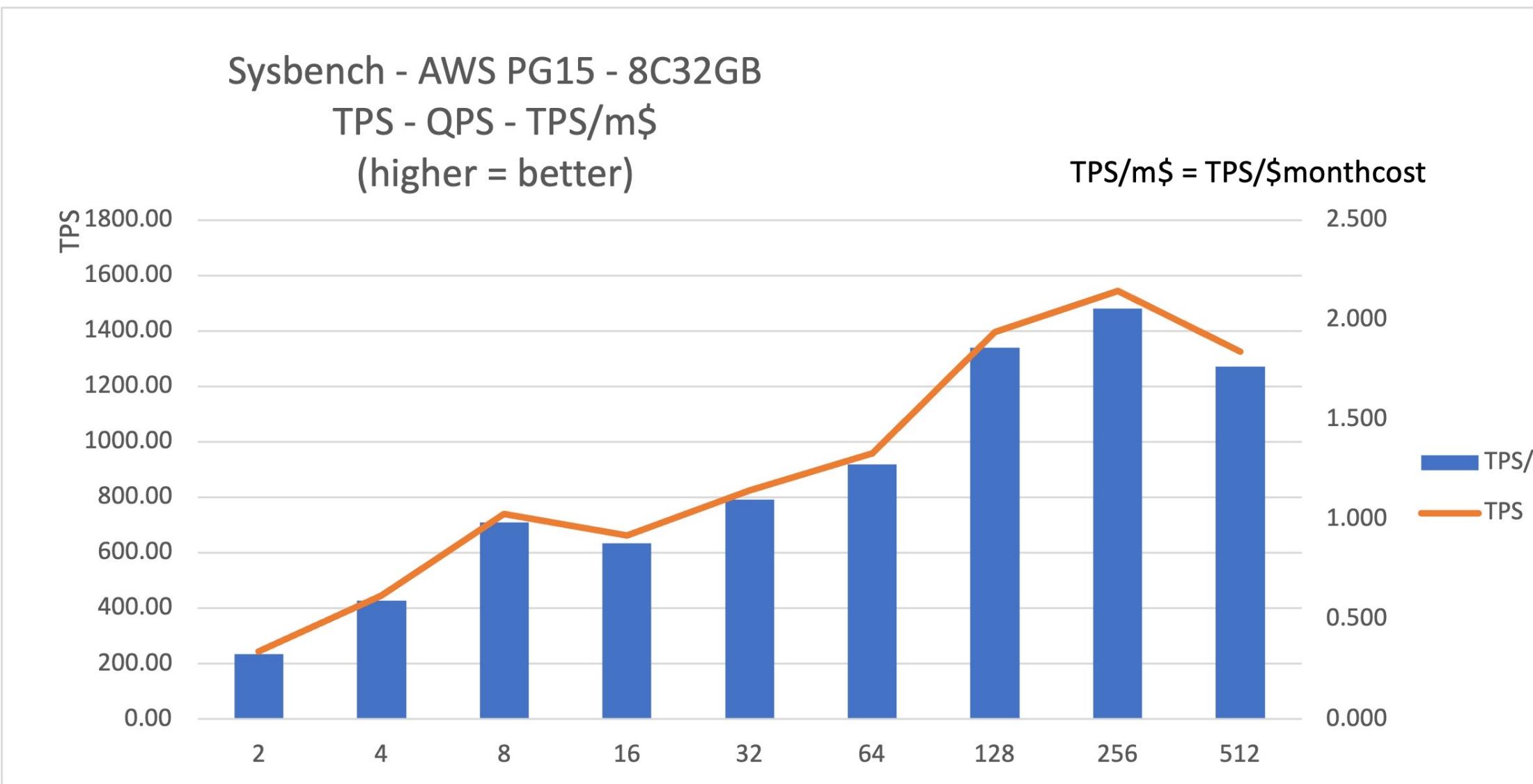
64 threads
2910@64 threads
64 Threads
P95 latency = 35.59ms
TPS = 2910

3.778 TPS/dollar-month

Middle Tier
Middle Tier
Middle Tier
Middle Tier



AWS – matured but expensive



- Concurrency scale to
- Highest TPS
- Sweat spot for productivity
- TPS/monthcost @ sweat spot

256 threads
1543@256 threads
128 Threads
P95 latency = 136ms
TPS = 1396

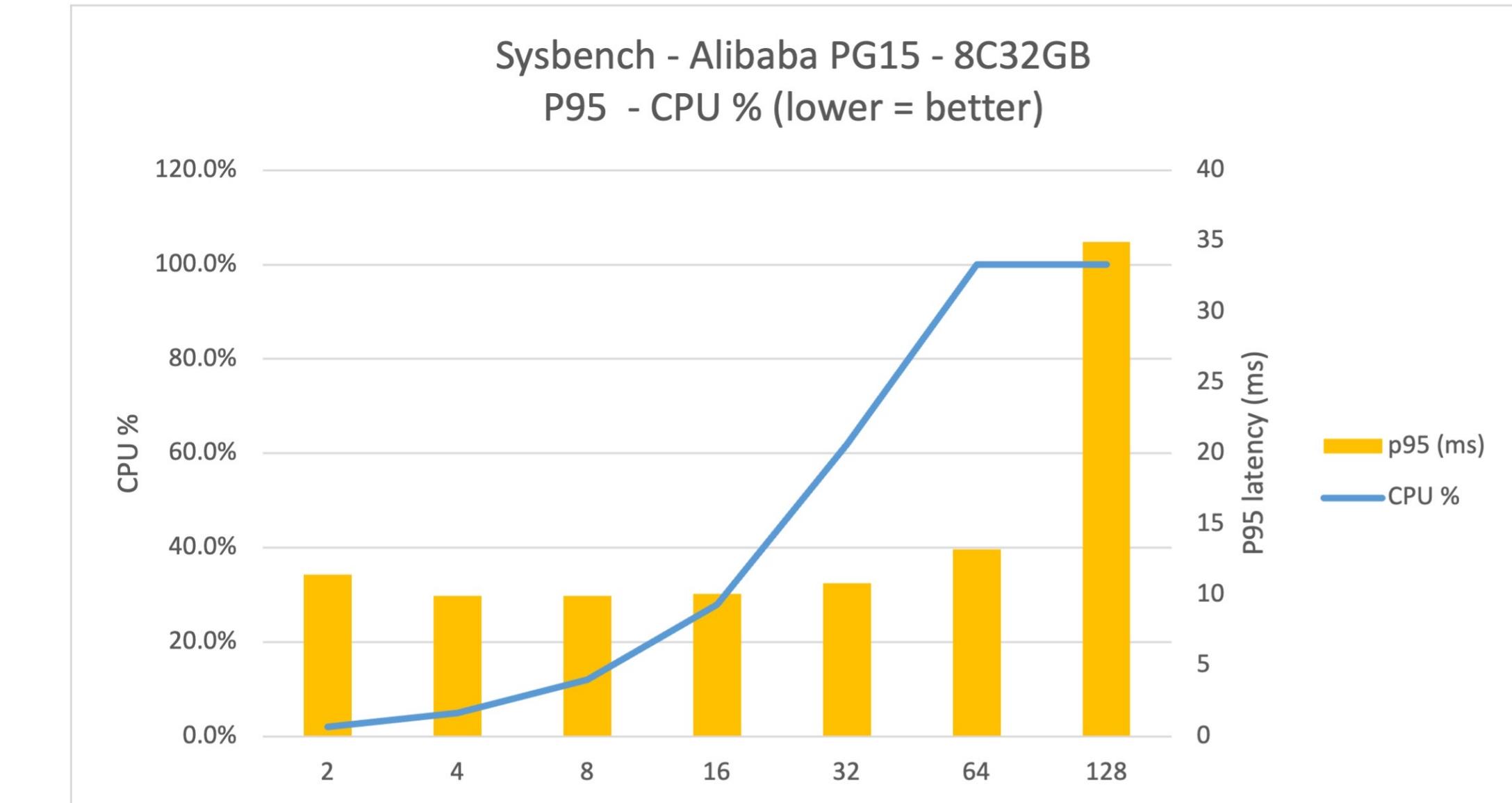
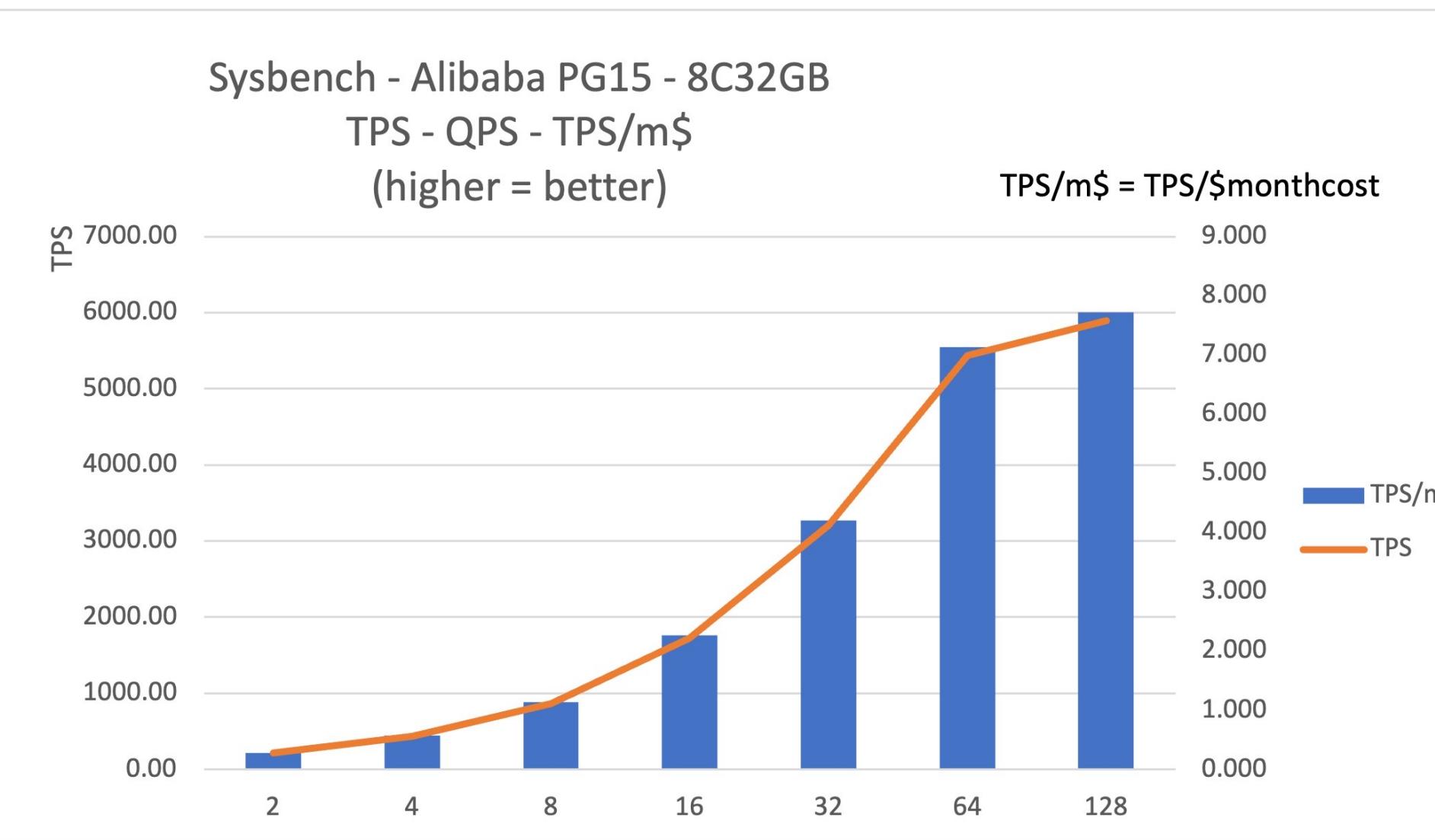
Highest of all five clouds
Lowest of all five clouds, $\frac{1}{2}$ of GCP's
Worst of all, P95 is 4X of GCP and only product 50% TPS
Worst of all, $\frac{1}{2}$ of GPS's

One thing I hate: Too expensive



Alibaba Cloud – Surprised winner

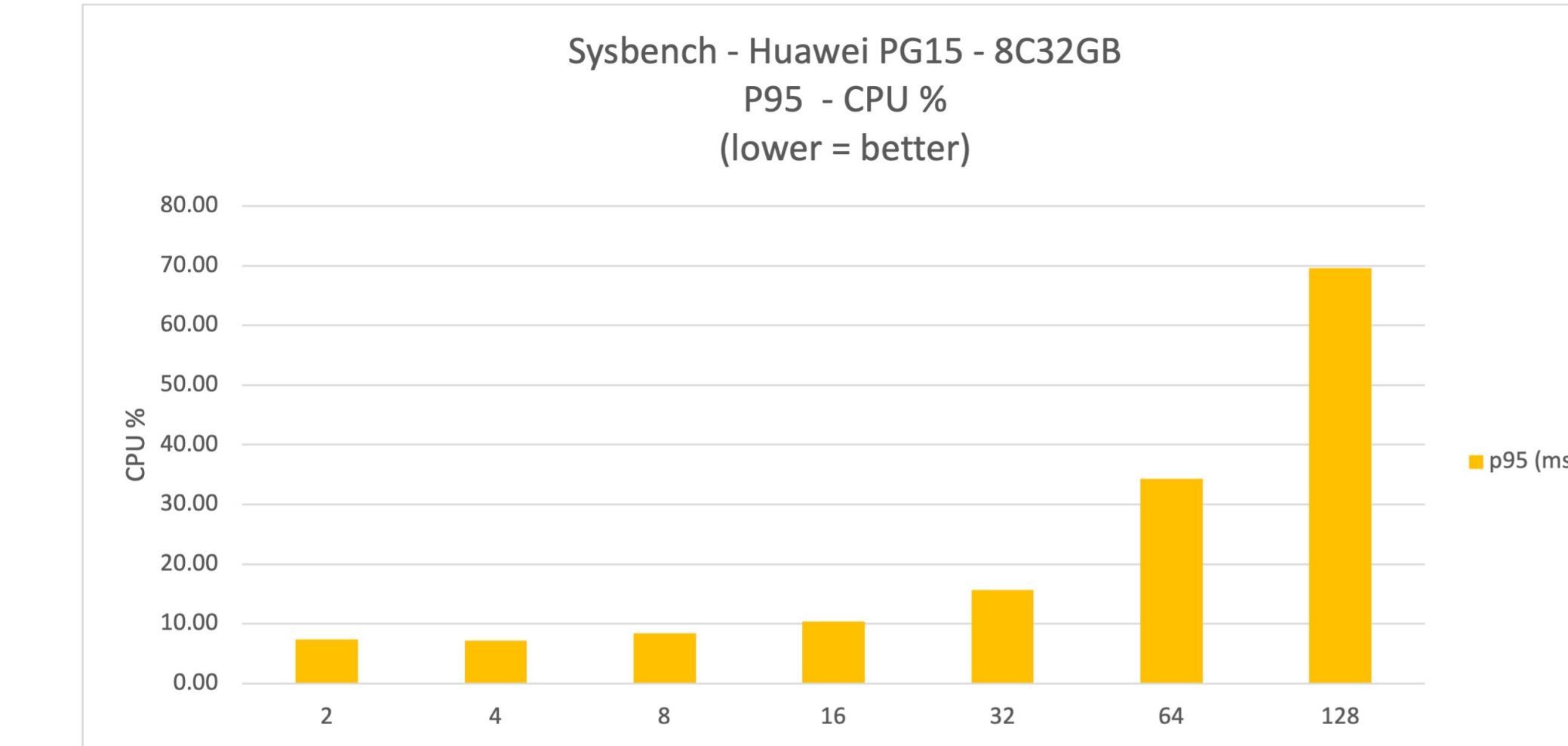
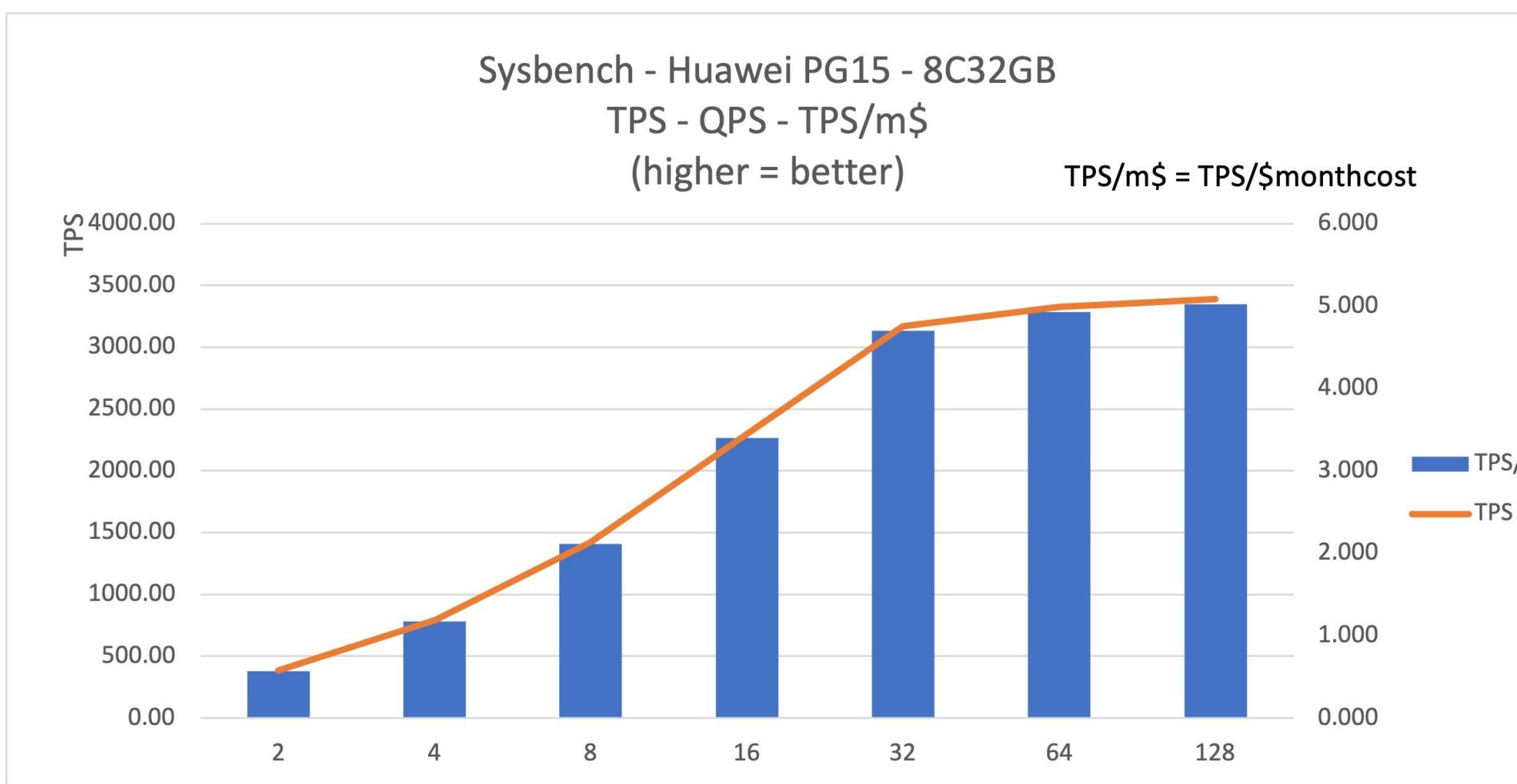
Alibaba Cloud



➤ Concurrency scale to	64 threads	Middle Tier
➤ Highest TPS	5892@128 threads	Highest of all five clouds, 2X of GCP's
➤ Sweat spot for productivity	64 Threads P95 latency = 13.22ms TPS = 5440	Best of all, P95 is 37% of GCP and produce 87% extra TPS
➤ TPS/monthcost @ sweat spot	7.13 TPS/dollar-month	Best of all, 89% higher than GCP's and almost 4X of AWS



Huawei Cloud – Reasonable Performance without US regions



➤ Concurrency scale to	32 threads	Lower Tier
➤ Highest TPS	3388@128 threads	Middle Tier
➤ Sweat spot for productivity	32 Threads P95 latency = 15.70ms TPS = 3172	P95 is 44% of GCP (@ 64 threads) and produce 9% extra TPS (GCP @ 64 thds)
➤ TPS/monthcost @ sweat spot	4.699 TPS/dollar-month	24% higher than GCP's (@64 thds)

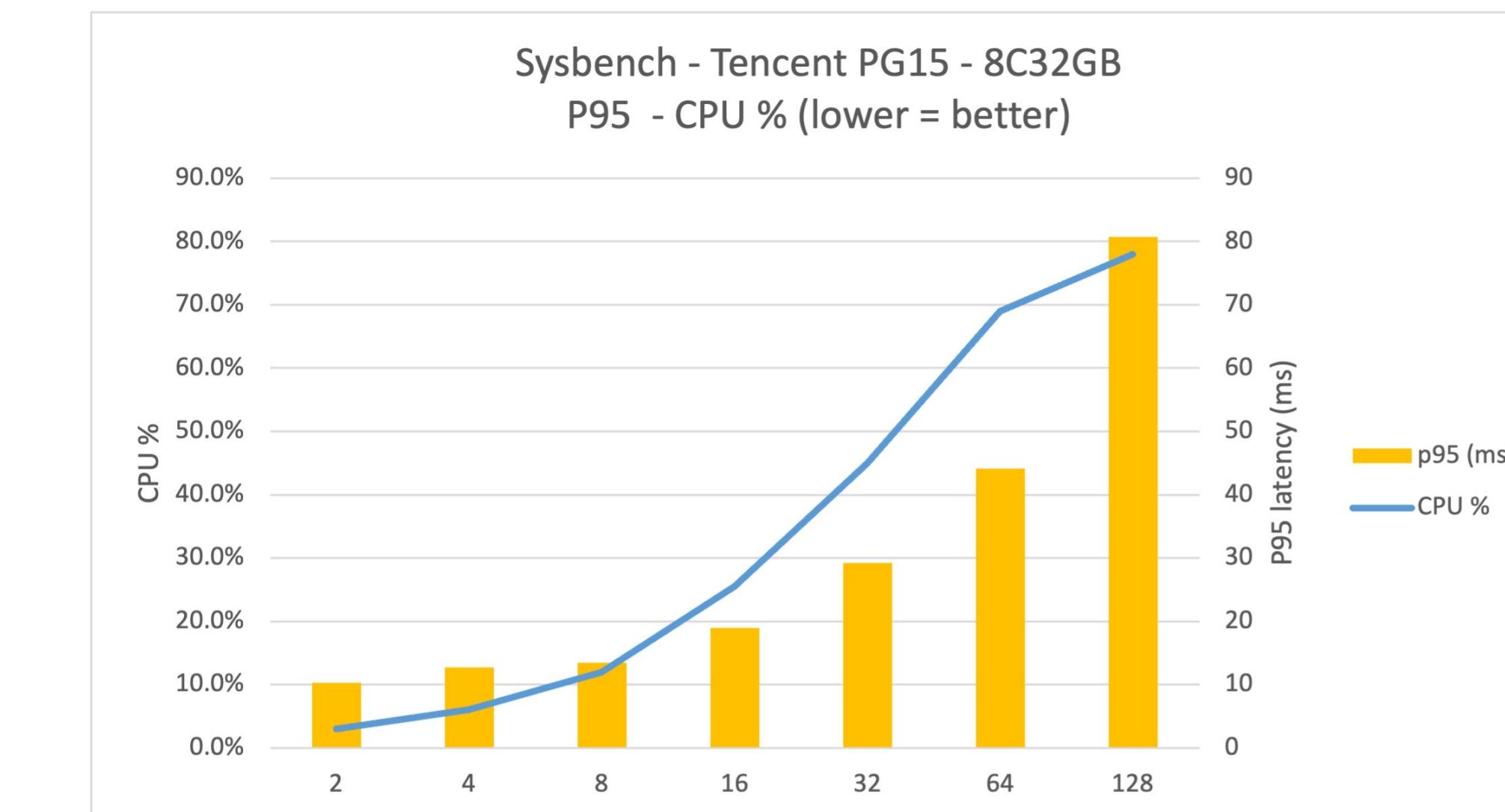
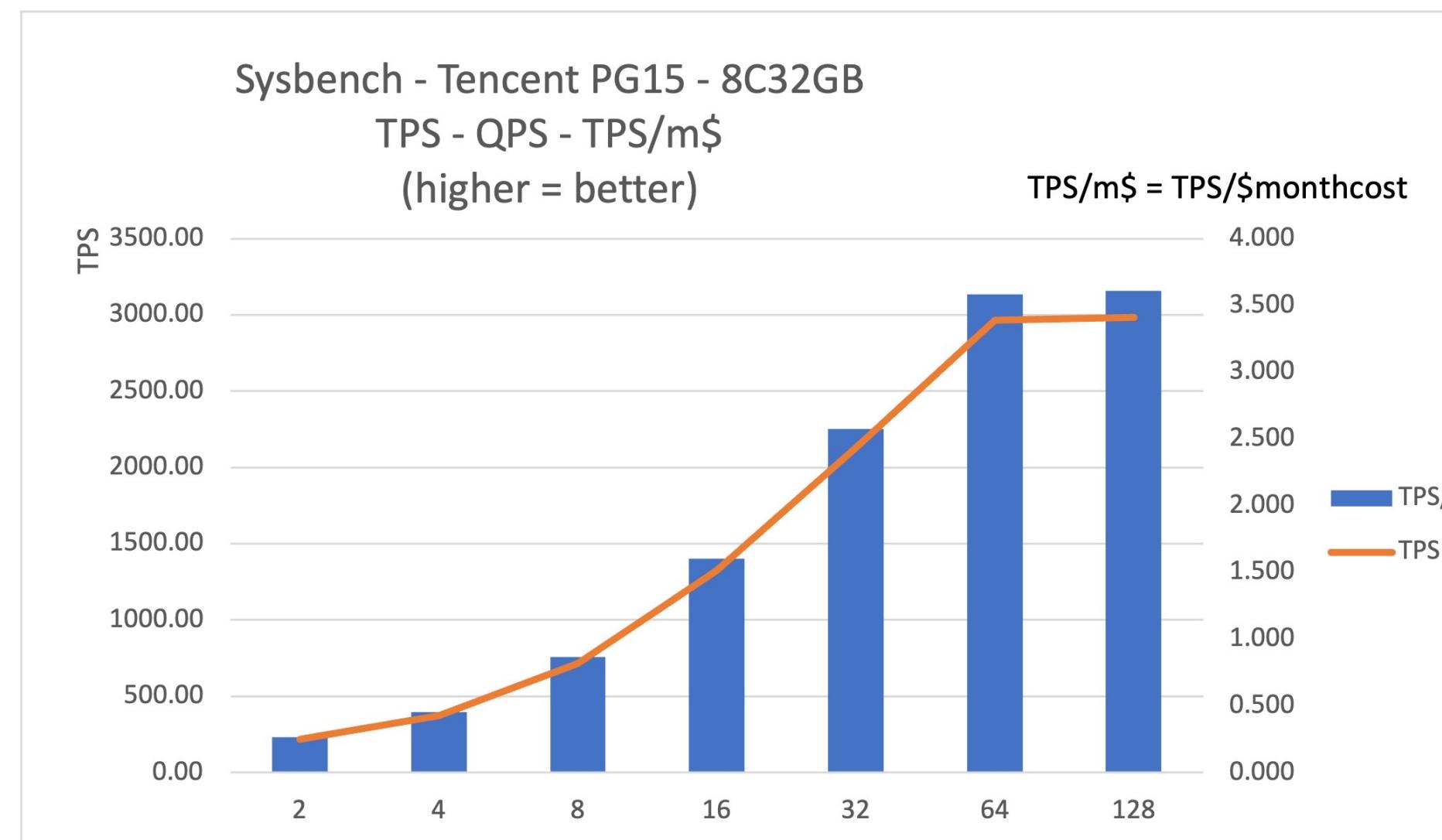
One thing I hate: No performance information. Hence, the report doesn't contain CPU utilization.

After 3+ hours' struggling on the console with authorization and permission, I cannot figure out how to get the monitor/performance information. On the other four clouds(GCP, AWS, Ali, Tencent) only 10seconds~5 minutes was necessary.



Tencent
Cloud

Tencent Cloud – Reasonable Performance without US regions

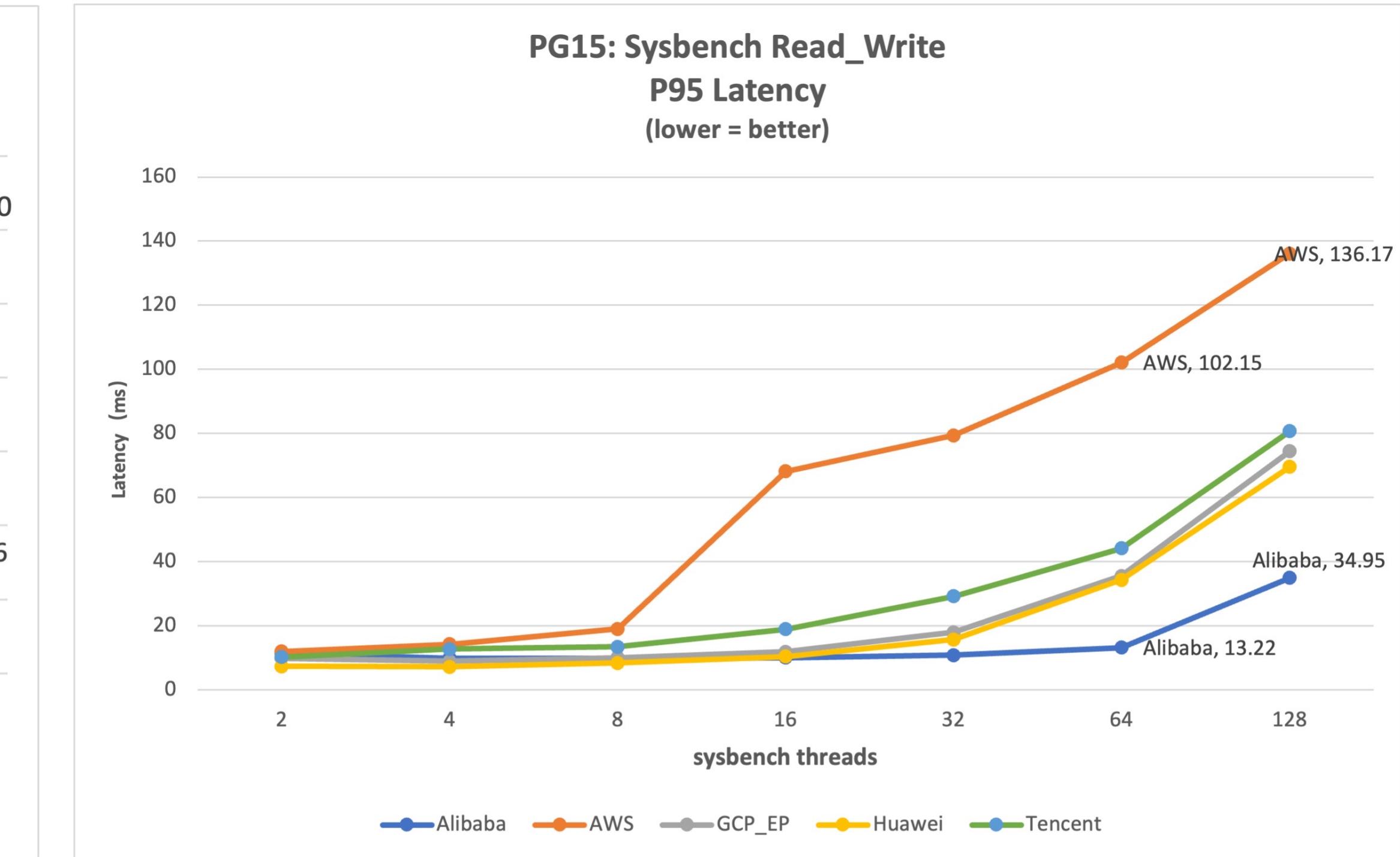
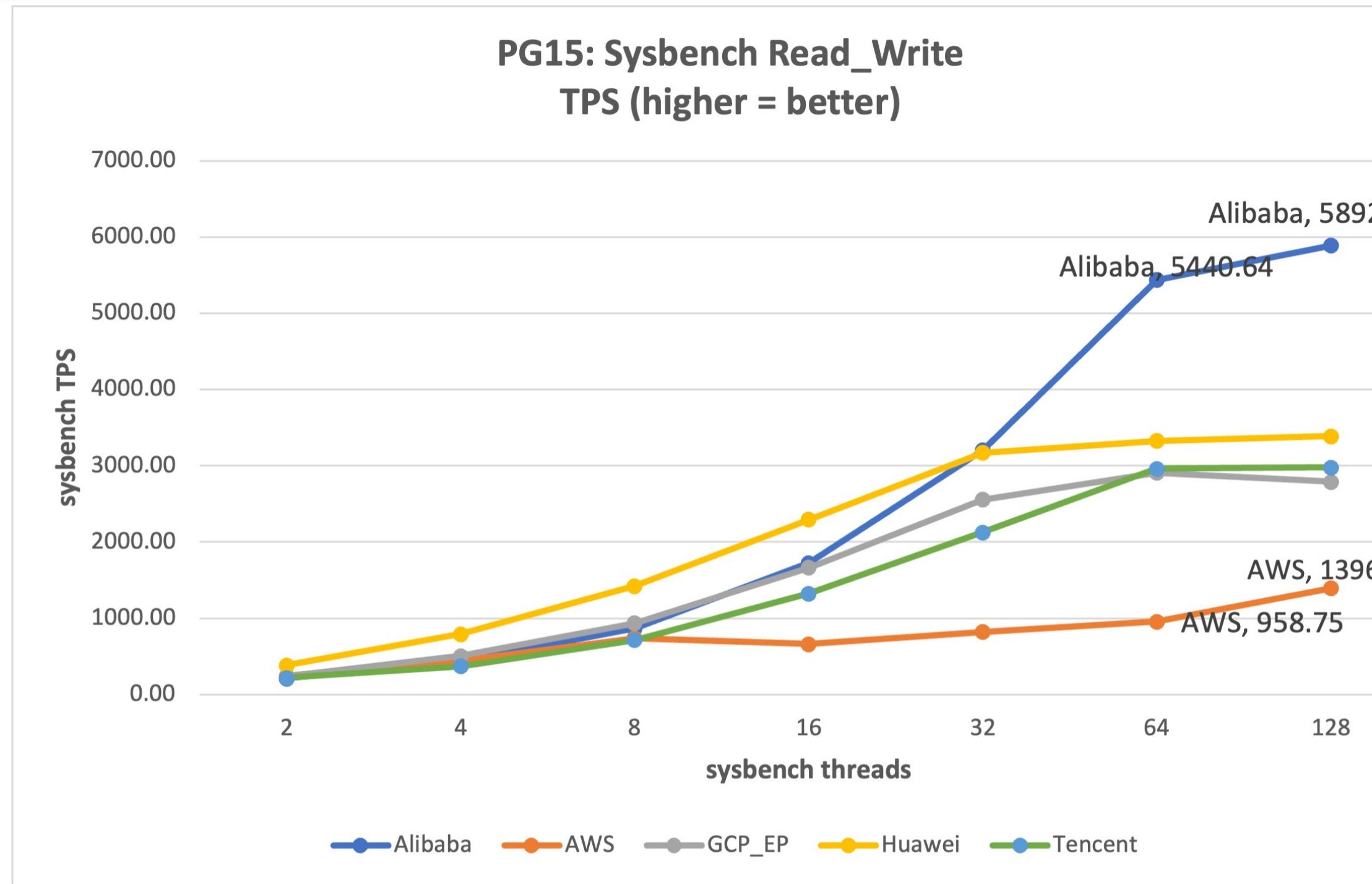


➤ Concurrency scale to	64 threads	Middle Tier
➤ Highest TPS	2982@128 threads	Middle Tier
➤ Sweat spot for productivity	64 Threads P95 latency = 44.17ms TPS = 2964	P95 is 25% more than GCP with the same TPS
➤ TPS/monthcost @ sweat spot	3.585 TPS/dollar-month	Slightly lower than GCP's

One thing I hate: If account setup on “China site”, the control console carries no English option.

“China site” is an unique and complex concept for Chinese cloud providers

Performance centric: Alibaba >> Huawei ≈ GCP ≈ Tencent >> AWS

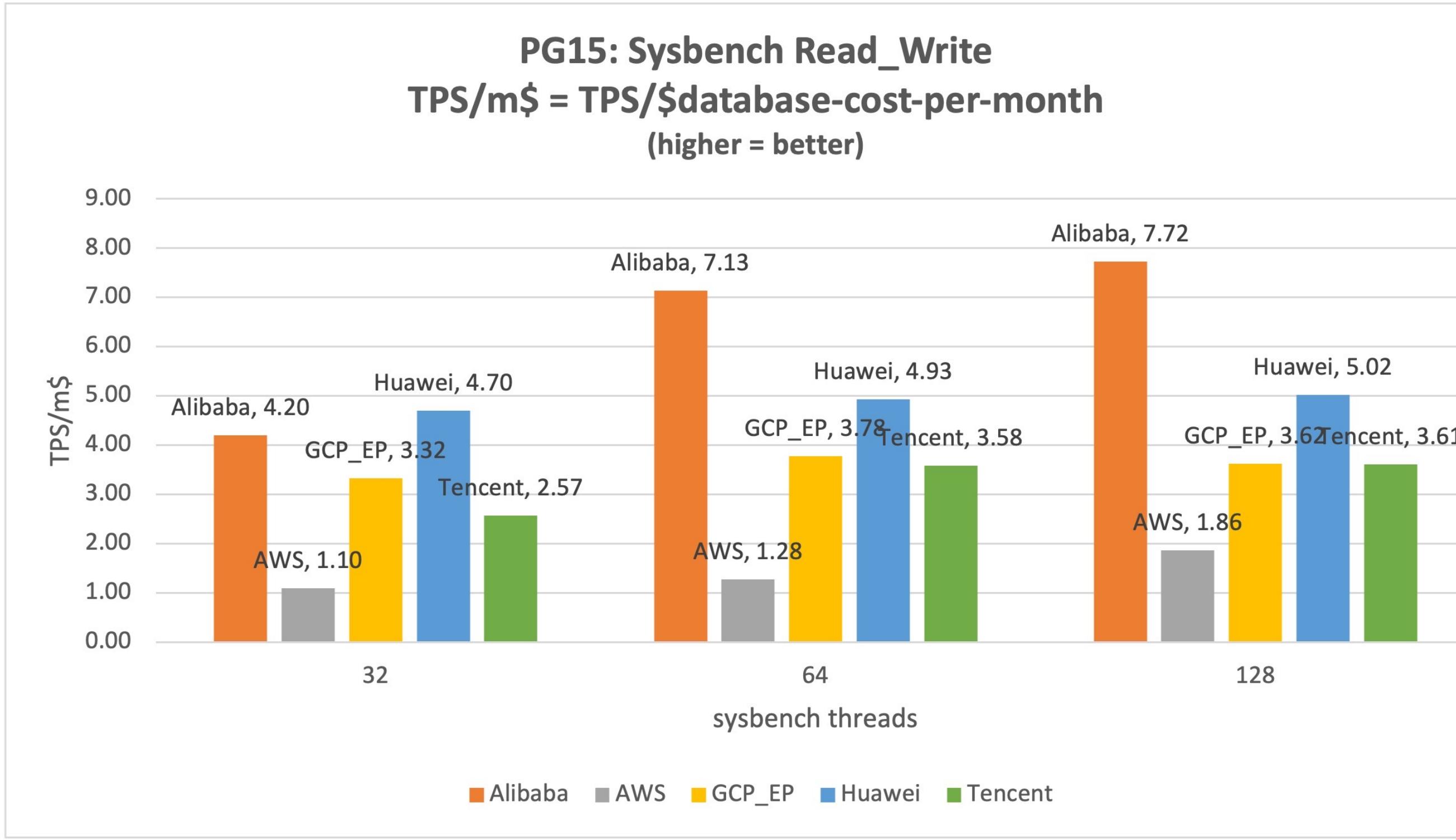


TPS	2	4	8	16	32	64	128
Alibaba	212.86	437.77	869.43	1726.85	3204.61	5440.64	5892.50
AWS	244.96	445.22	739.99	661.48	825.02	958.75	1396.26
GCP_EP	240.90	508.23	935.88	1664.64	2559.39	2910.95	2790.33
Huawei	386.99	793.30	1424.76	2296.64	3172.38	3328.92	3388.67
Tencent	219.97	373.31	714.87	1326.32	2127.34	2964.04	2982.50

P95(ms)	2	4	8	16	32	64	128
Alibaba	11.45	9.91	9.91	10.09	10.84	13.22	34.95
AWS	11.96	14.22	19.00	68.16	79.42	102.15	136.17
GCP_EP	9.91	8.90	10.09	11.87	17.95	35.59	74.46
Huawei	7.34	7.17	8.37	10.40	15.70	34.34	69.57
Tencent	10.27	12.75	13.46	18.95	29.19	44.17	80.76

Price/Performance: Alibaba >> Huawei > GCP ≈ Tencent >> AWS

ATTN: Alibaba Huawei and Tencent (2 instances HA) vs. GCP and AWS (Single Instance)



- **Alibaba:** the clear winner
- **GCP, Huawei and Tencent:** the middle pack
- **AWS:** the surprised bottom

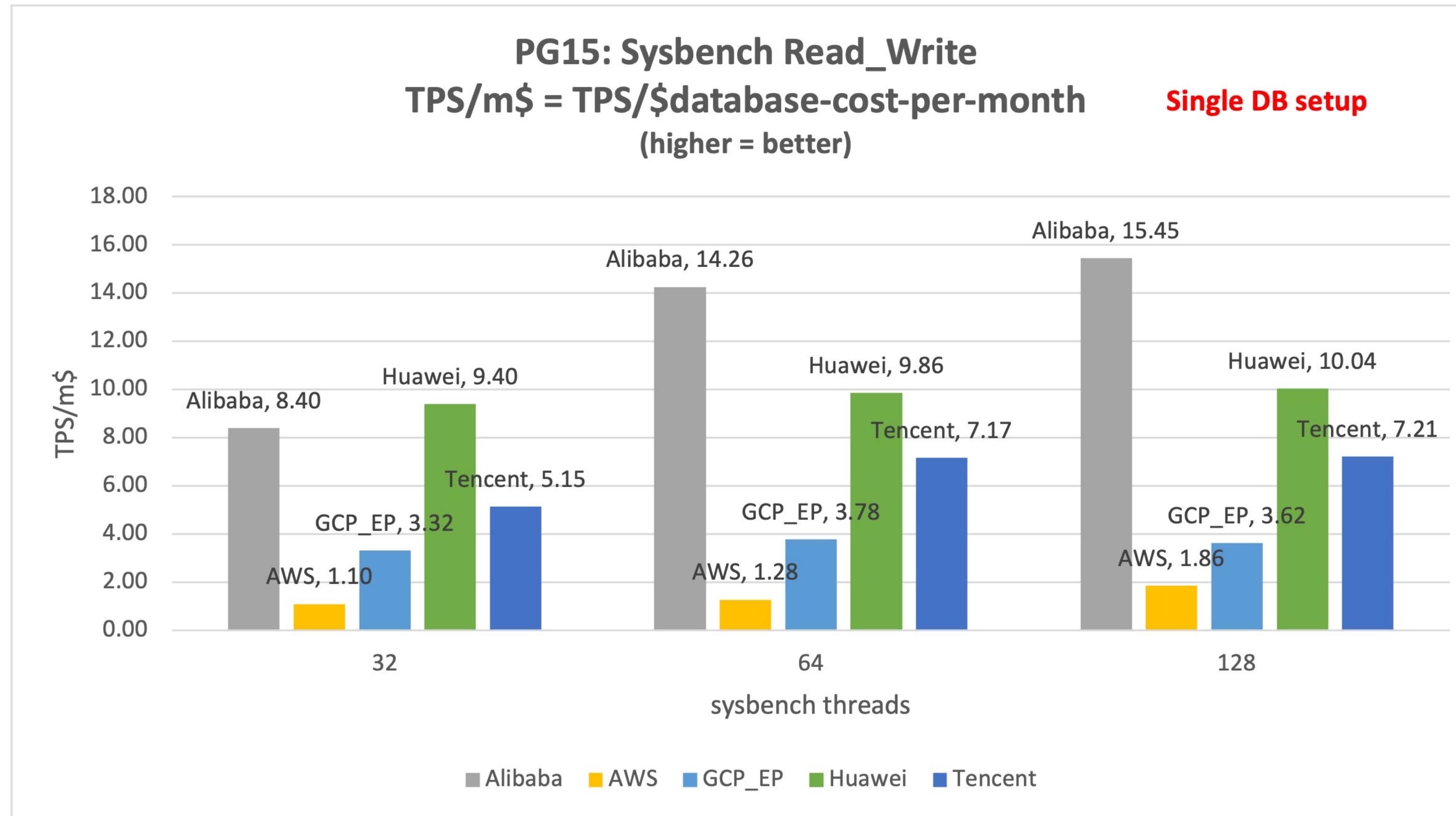
Notes:

1. Currency exchange 7:21 CNY : 1 USD
2. Huawei, Tencent and Alibaba provide same zone standby(may not be the same model as master)
3. AWS charge IOPS, roughly 25% of the total cost

TPS/m\$	2	4	8	16	32	64	128
Alibaba	0.28	0.57	1.14	2.26	4.20	7.13	7.72
AWS	0.33	0.59	0.99	0.88	1.10	1.28	1.86
GCP_EP	0.31	0.66	1.21	2.16	3.32	3.78	3.62
Huawei	0.57	1.18	2.11	3.40	4.70	4.93	5.02
Tencent	0.27	0.45	0.86	1.60	2.57	3.58	3.61

Price/Performance: Alibaba = 10X of AWS

For fair comparison, use the cost of single DB for All five clouds



Notes:

1. Double the TPS/m\$ for Alibaba, Huawei and Tencent
2. Alibaba produce **7X(32 thds), 11X(64thds) and 8X(128thds)** of AWS's

User friendly: GCP ≈ AWS >> Alibaba >> ~~Huawei > Tencent~~

- **Both GCP and AWS** : matured with Infrastructure as Code/OpenAPI interface
- **Alibaba** : less user friendly, but passable
- **Huawei**: with many issues, 1) some key operation menu are still in English, 2) browser-based SSH lacks key feature, 3) fail to turn on monitor/performance information; 4) no region in US
- **Tencent**: no English support, a deal killer

Conclusion

Limitation

- All Data and benchmark were collected and executed in the 2nd half of January, 2024. Price and system may change afterward.
- PostgreSQL 15: the minor versions of the databases from different vendors may vary.
- CPU, disk types, network, and other aspects: vary among different vendors, and even among different regions of the same vendors
- Default parameters for RDS instances: different vendors also vary, and even the same vendor may adjust parameters at different model.
- Over- and under-provisioning in cloud computing
- A variety of disk storage options. For example, Alibaba Cloud supports ESSD PL1/2/3, and AWS supports storage choices like gp3 and io1.
- Price structure and discount: Tencent has ladder price : (0, 4] day ¥15.96/hour; (4, 15] days ¥ 12.12/hour; (15, ∞) days ¥8.28/hour;
AWS charges IOPS, Alibaba always carries some discount

**Although the design is imperfect,
the report points to the two winners with
confidence**

[Cost sensitive user]



[Stay with US-based large company]



Google Cloud

Disclaimer:

The engineer/author responsible for this analysis is a former employee of both Huawei and Alibaba. They currently have no business affiliations with any of the cloud providers mentioned, aside from incurring costs or applying for free credits to conduct the benchmark tests. He was as surprised by the data as anyone else and had a hard time explaining the huge gap.