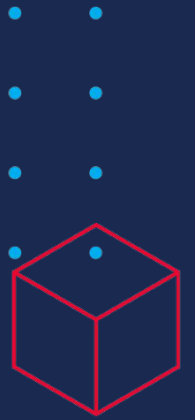


Benchbot

Benchmark as a Service for TiDB

Yuying Song, Huansheng Chen @ PingCAP

June 17 - DBTest 2022

TiDB

TiDB is an **open-source** NewSQL database that supports **Hybrid Transactional and Analytical Processing (HTAP)** workloads. It is **MySQL compatible** and features horizontal scalability, strong consistency, and high availability. The goal of TiDB is to provide users with a **one-stop database solution** that covers OLTP (Online Transactional Processing), OLAP (Online Analytical Processing), and HTAP services. TiDB is suitable for various use cases that require **high availability** and **strong consistency** with **large-scale** data.

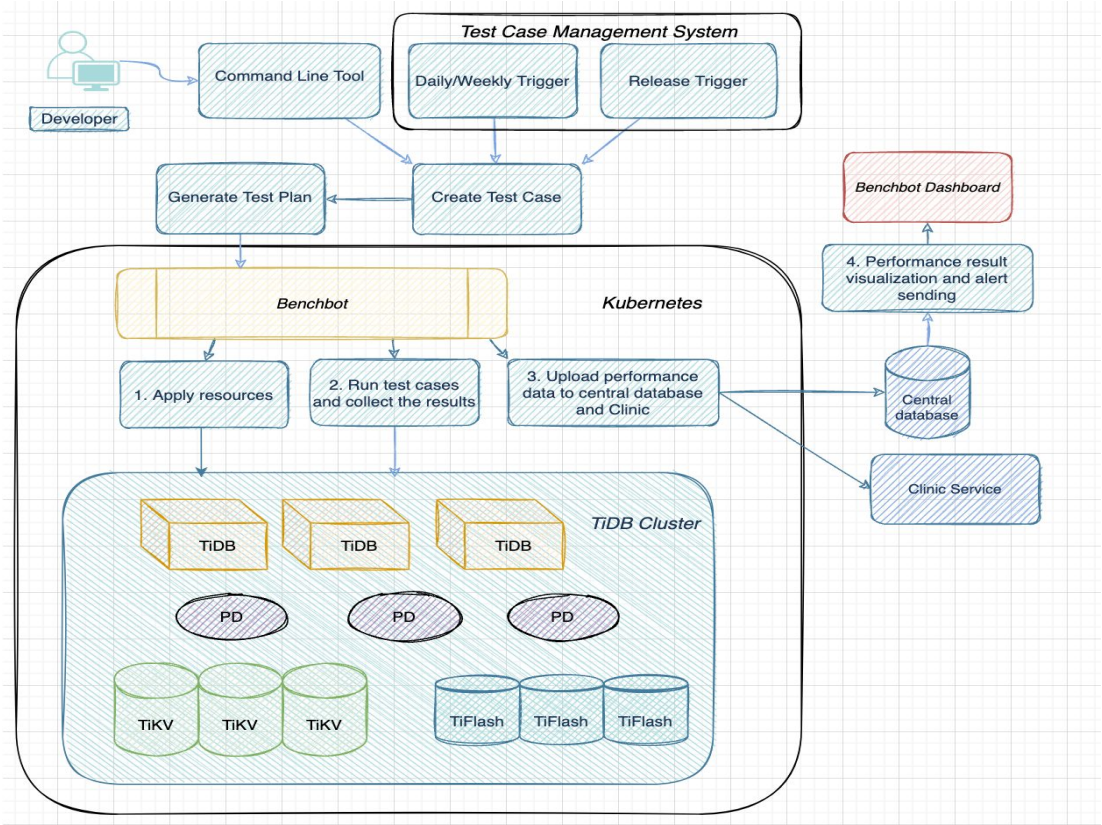
Obstacles of Benchmarking

1. Hardware Resources
2. Data Sets for Testing

What is Benchbot

- **Benchmark as a Service**
- A user-friendly performance testing self-service developed by the PingCAP Performance Testing Team to make TiDB benchmark **EASIER**, **FASTER** and **STABLER**
- Benchbot Motto: Easier, Faster, Stabler - Together

Benchbot Architecture



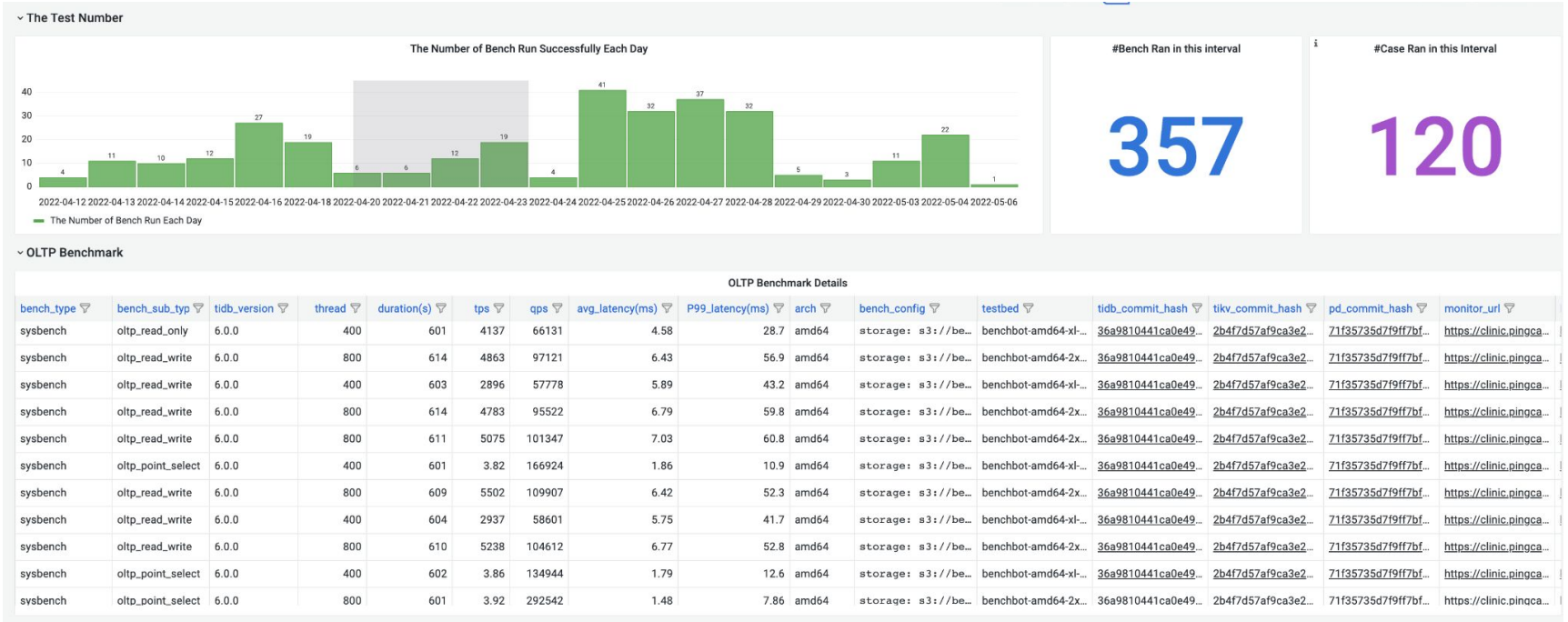
Why Benchbot - Easier

With a single Benchbot command, we can share the following resources across the whole company:

- Benchbot **hardware resource pool**
- All the **daily run benchmark pipelines** *polished up* by the Performance Testing Team
    - Nearly 30 workloads including stardard benchmarks and home-grown workloads
- All the **data sets** of these benchmark workloads

Benchbot Dashboard

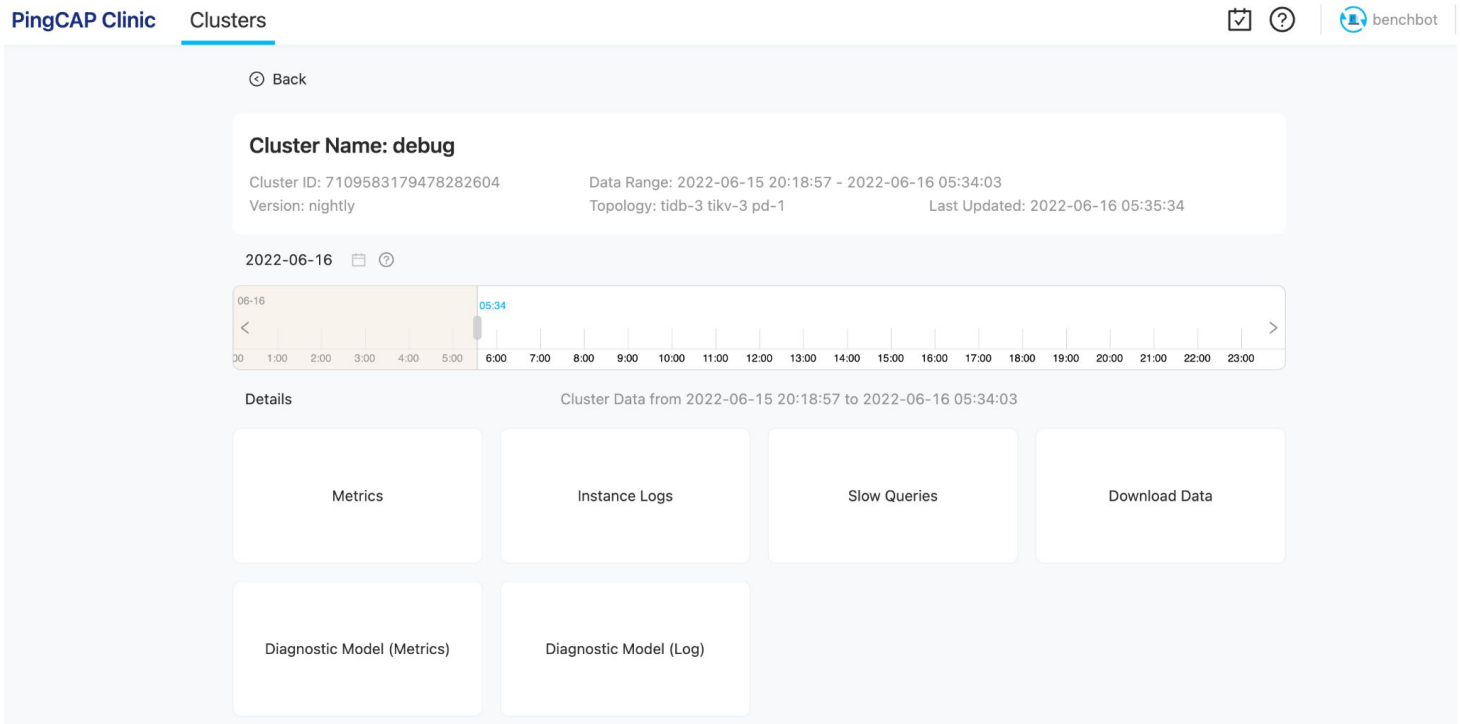
- One week before and after the Code Freeze of v6.0.0



bench_type	bench_sub_typ	tidb_version	thread	duration(s)	tps	qps	avg_latency(ms)	P99_latency(ms)	arch	bench_config	testbed	tidb_commit_hash	tikv_commit_hash	pd_commit_hash	monitor_url
sysbench	oltp_read_only	6.0.0	400	601	4137	66131	4.58	28.7	amd64	storage: s3://be...	benchbot-amd64-xl-...	36a9810441ca0e49...	2b4f7d57af9ca3e2...	71f35735d7f9ff7bf...	https://clinic.pingca...
sysbench	oltp_read_write	6.0.0	800	614	4863	97121	6.43	56.9	amd64	storage: s3://be...	benchbot-amd64-2x...	36a9810441ca0e49...	2b4f7d57af9ca3e2...	71f35735d7f9ff7bf...	https://clinic.pingca...
sysbench	oltp_read_write	6.0.0	400	603	2896	57778	5.89	43.2	amd64	storage: s3://be...	benchbot-amd64-xl-...	36a9810441ca0e49...	2b4f7d57af9ca3e2...	71f35735d7f9ff7bf...	https://clinic.pingca...
sysbench	oltp_read_write	6.0.0	800	614	4783	95522	6.79	59.8	amd64	storage: s3://be...	benchbot-amd64-2x...	36a9810441ca0e49...	2b4f7d57af9ca3e2...	71f35735d7f9ff7bf...	https://clinic.pingca...
sysbench	oltp_read_write	6.0.0	800	611	5075	101347	7.03	60.8	amd64	storage: s3://be...	benchbot-amd64-2x...	36a9810441ca0e49...	2b4f7d57af9ca3e2...	71f35735d7f9ff7bf...	https://clinic.pingca...
sysbench	oltp_point_select	6.0.0	400	601	3.82	166924	1.86	10.9	amd64	storage: s3://be...	benchbot-amd64-xl-...	36a9810441ca0e49...	2b4f7d57af9ca3e2...	71f35735d7f9ff7bf...	https://clinic.pingca...
sysbench	oltp_read_write	6.0.0	800	609	5502	109907	6.42	52.3	amd64	storage: s3://be...	benchbot-amd64-2x...	36a9810441ca0e49...	2b4f7d57af9ca3e2...	71f35735d7f9ff7bf...	https://clinic.pingca...
sysbench	oltp_read_write	6.0.0	400	604	2937	58601	5.75	41.7	amd64	storage: s3://be...	benchbot-amd64-xl-...	36a9810441ca0e49...	2b4f7d57af9ca3e2...	71f35735d7f9ff7bf...	https://clinic.pingca...
sysbench	oltp_read_write	6.0.0	800	610	5238	104612	6.77	52.8	amd64	storage: s3://be...	benchbot-amd64-2x...	36a9810441ca0e49...	2b4f7d57af9ca3e2...	71f35735d7f9ff7bf...	https://clinic.pingca...
sysbench	oltp_point_select	6.0.0	400	602	3.86	134944	1.79	12.6	amd64	storage: s3://be...	benchbot-amd64-xl-...	36a9810441ca0e49...	2b4f7d57af9ca3e2...	71f35735d7f9ff7bf...	https://clinic.pingca...
sysbench	oltp_point_select	6.0.0	800	601	3.92	292542	1.48	7.86	amd64	storage: s3://be...	benchbot-amd64-2x...	36a9810441ca0e49...	2b4f7d57af9ca3e2...	71f35735d7f9ff7bf...	https://clinic.pingca...

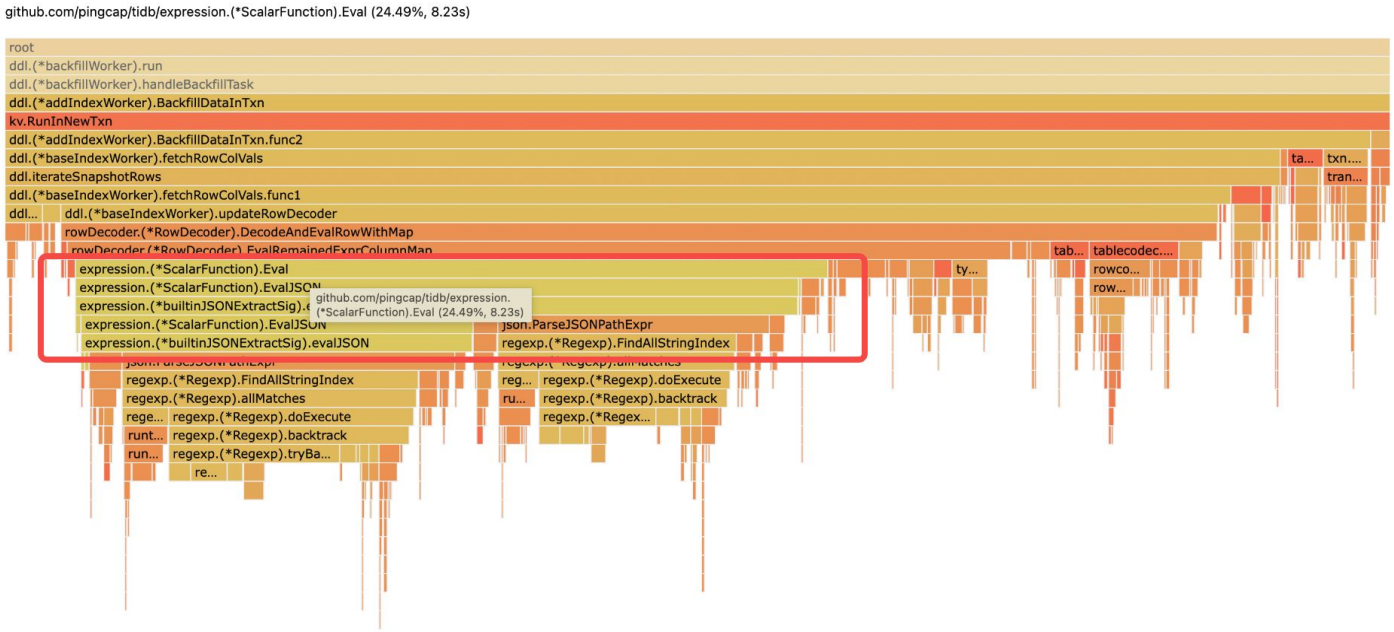
Why Benchbot - Easier

[PingCAP Clinic Diagnostic Service]



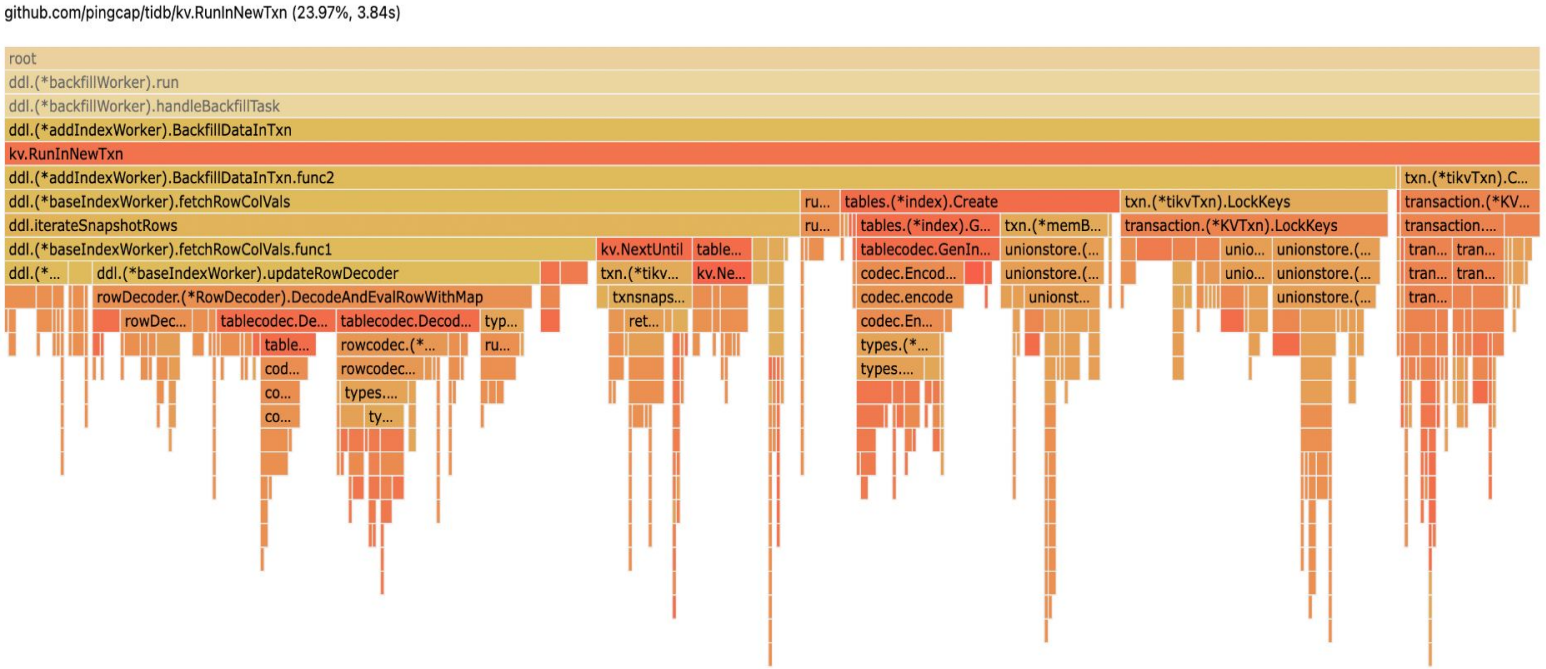
Why Benchbot - Easier

FlameGraph collected by Continuous Profiling



Why Benchbot - Easier

FlameGraph collected by Continuous Profiling



Why Benchbot - Faster

```
python3 main.py gen-benchbot-case --email "example@pingcap.com" --token 'tcmsp_xxxxx' \
--testbed_size "2xl" --arch amd64 \
--bench_type sysbench --bench_sub_types "oltp_write_only,oltp_insert" \
--duration=30m --threads 400 \
--versions v5.3.0 \
--tikv_urls ",http://fileserver.pingcap.net/download/builds/pingcap/tikv/${commit_hash}/centos7/tikv-server.tar.gz" \
--tidb_configs "{prepared-plan-cache: {enabled: true}}"  \
--tikv_configs "{raftstore: {store-io-pool-size: 1}, raft-engine: {enable: false}}" \
--tikv_configs "{raftstore: {store-io-pool-size: 1}, raft-engine: {enable: true}}"
```

Why Benchbot - Stabler

Performance tests with **high stability requirements**

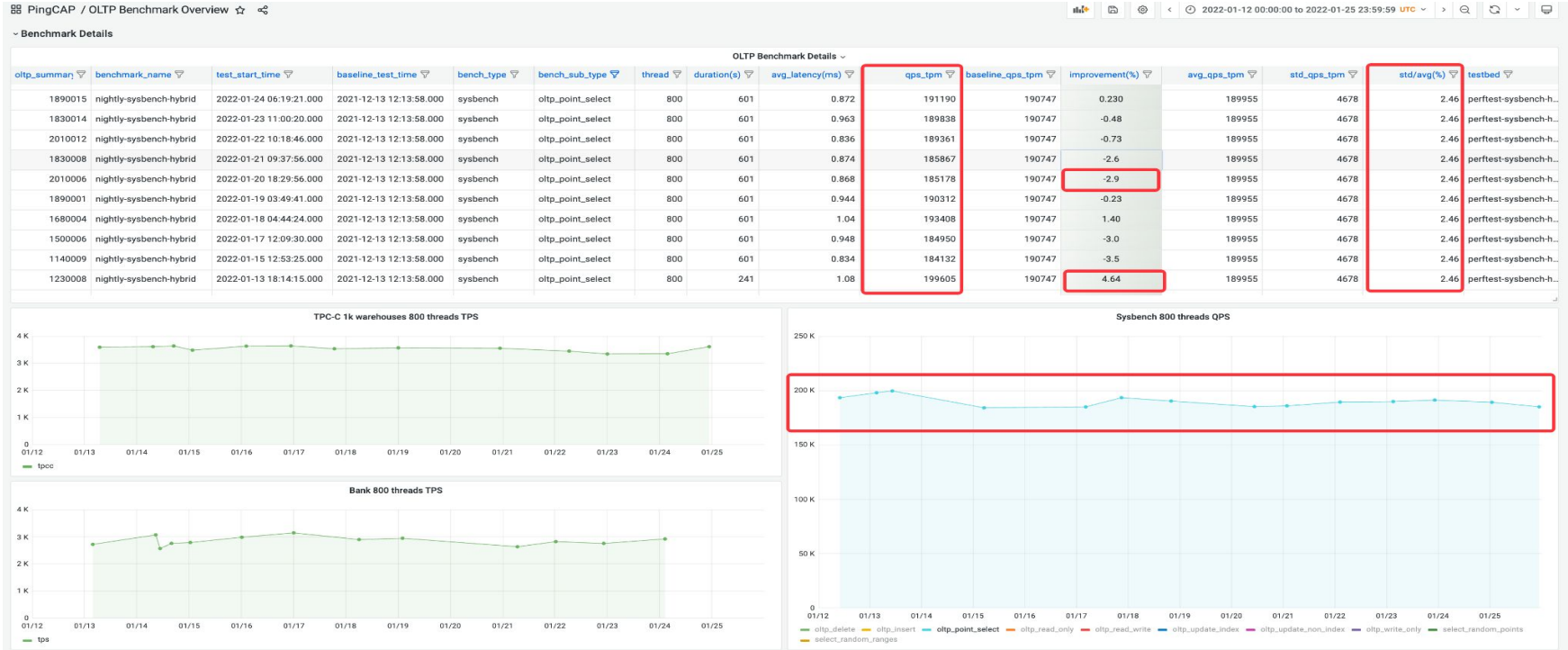
- are run on **fixed** Kubernetes nodes

All the benchmark pipelines

- are subjected to **Stability Assessment**
- are run **daily with performance regression alert**

Why Benchbot - Stabler

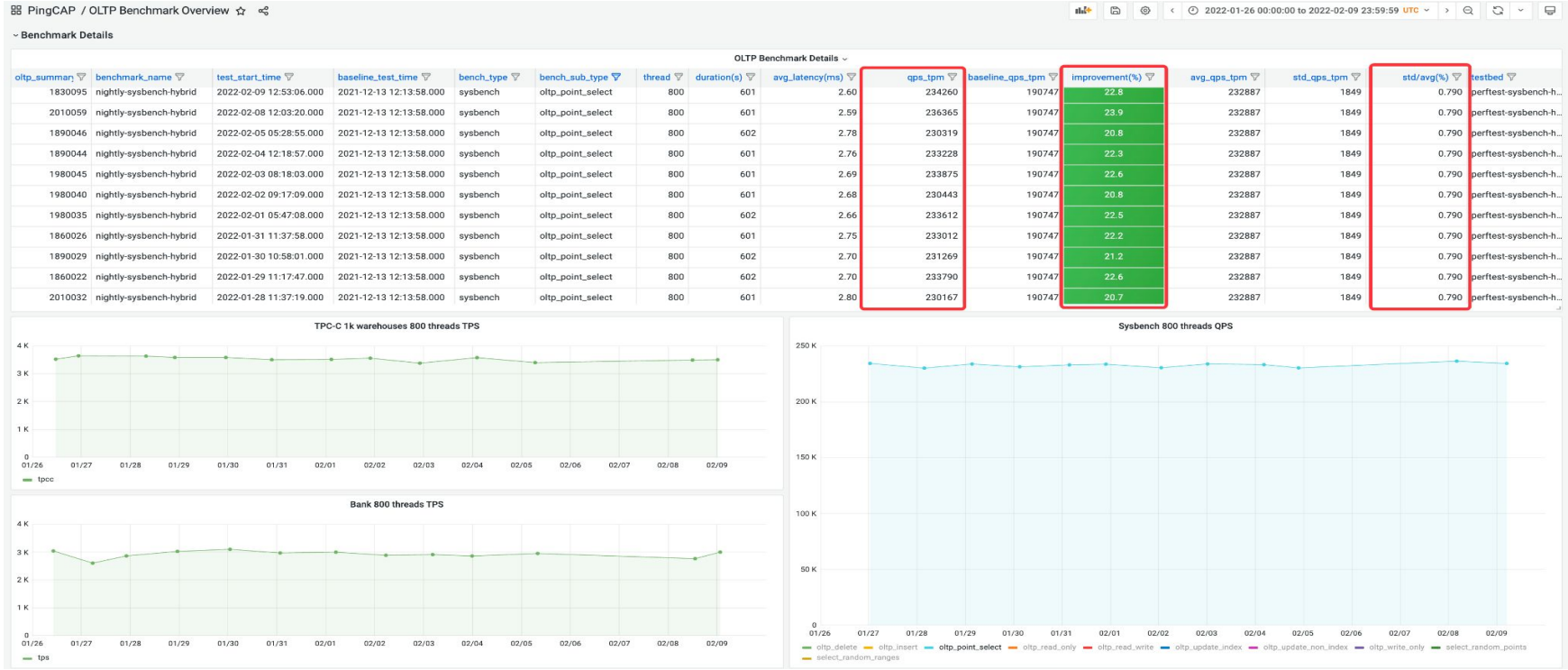
Daily Run - Before replacing HAProxy with [kube-proxy of Kubernetes]



oltp_summary	benchmark_name	test_start_time	baseline_test_time	bench_type	bench_sub_type	thread	duration(s)	avg_latency(ms)	qps_tpm	baseline_qps_tpm	improvement(%)	avg_qps_tpm	std_qps_tpm	std/avg(%)	testbed
1890015	nightly-sysbench-hybrid	2022-01-24 06:19:21.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	0.872	191190	190747	0.230	189955	4678	2.46	perftest-sysbench-h...
1830014	nightly-sysbench-hybrid	2022-01-23 11:00:20.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	0.963	189838	190747	-0.48	189955	4678	2.46	perftest-sysbench-h...
2010012	nightly-sysbench-hybrid	2022-01-22 10:18:46.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	0.836	189361	190747	-0.73	189955	4678	2.46	perftest-sysbench-h...
1830008	nightly-sysbench-hybrid	2022-01-21 09:37:56.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	0.874	185867	190747	-2.6	189955	4678	2.46	perftest-sysbench-h...
2010006	nightly-sysbench-hybrid	2022-01-20 18:29:56.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	0.868	185178	190747	-2.9	189955	4678	2.46	perftest-sysbench-h...
1890001	nightly-sysbench-hybrid	2022-01-19 03:49:41.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	0.944	190312	190747	-0.23	189955	4678	2.46	perftest-sysbench-h...
1680004	nightly-sysbench-hybrid	2022-01-18 04:44:24.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	1.04	193408	190747	1.40	189955	4678	2.46	perftest-sysbench-h...
1500006	nightly-sysbench-hybrid	2022-01-17 12:09:30.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	0.948	184950	190747	-3.0	189955	4678	2.46	perftest-sysbench-h...
1140009	nightly-sysbench-hybrid	2022-01-15 12:53:25.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	0.834	184132	190747	-3.5	189955	4678	2.46	perftest-sysbench-h...
1230008	nightly-sysbench-hybrid	2022-01-13 18:14:15.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	241	1.08	199605	190747	4.64	189955	4678	2.46	perftest-sysbench-h...

Why Benchbot - Stabler

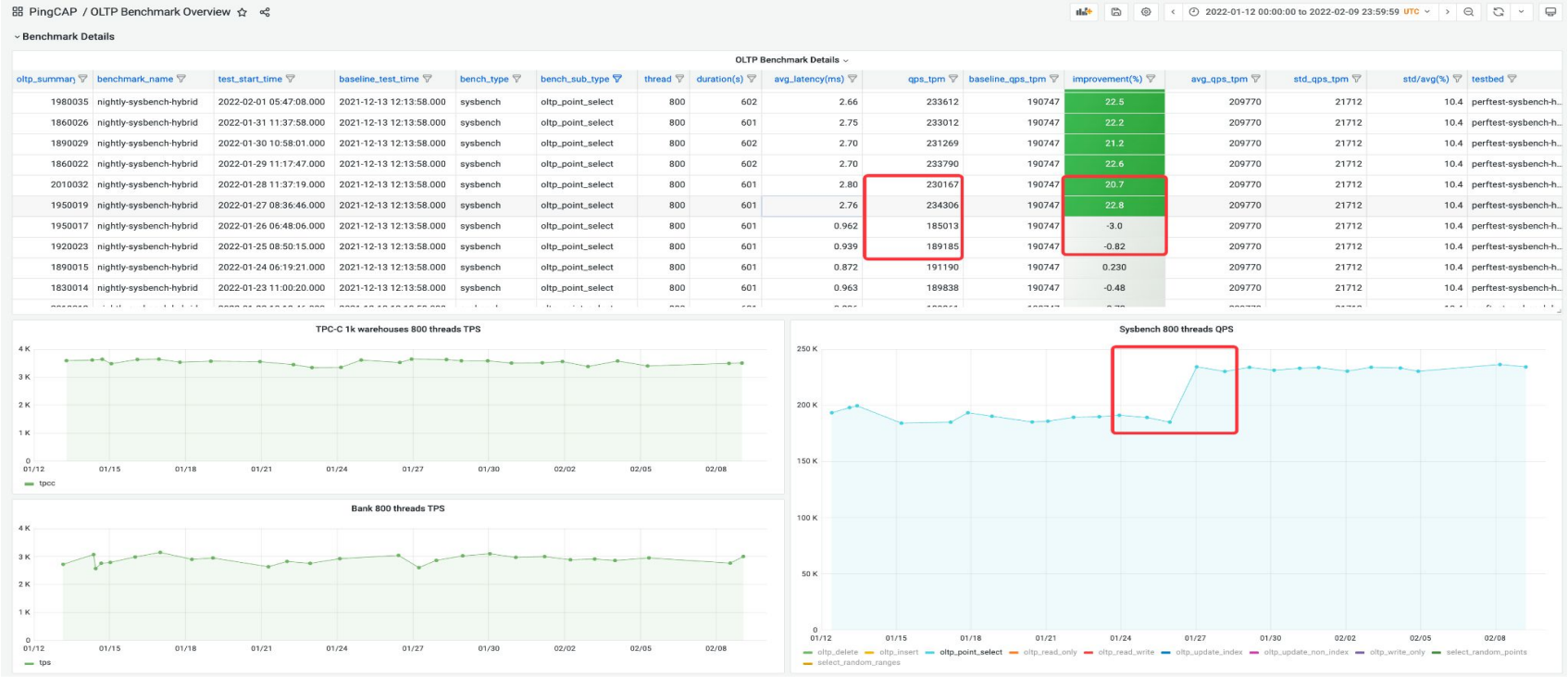
Daily Run - After replacing HAProxy with [kube-proxy of Kubernetes]



oltp_summary	benchmark_name	test_start_time	baseline_test_time	bench_type	bench_sub_type	thread	duration(s)	avg_latency(ms)	qps_tpm	baseline_qps_tpm	improvement(%)	avg_qps_tpm	std_qps_tpm	std/avg(%)	testbed
1830095	nightly-sysbench-hybrid	2022-02-09 12:53:06.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	2.60	234260	190747	22.8	232887	1849	0.790	perftest-sysbench-h...
2010059	nightly-sysbench-hybrid	2022-02-08 12:03:20.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	2.59	236365	190747	23.9	232887	1849	0.790	perftest-sysbench-h...
1890046	nightly-sysbench-hybrid	2022-02-05 05:28:55.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	602	2.78	230319	190747	20.8	232887	1849	0.790	perftest-sysbench-h...
1890044	nightly-sysbench-hybrid	2022-02-04 12:18:57.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	2.76	233228	190747	22.3	232887	1849	0.790	perftest-sysbench-h...
1980045	nightly-sysbench-hybrid	2022-02-03 08:18:03.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	2.69	233875	190747	22.6	232887	1849	0.790	perftest-sysbench-h...
1980040	nightly-sysbench-hybrid	2022-02-02 09:17:09.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	2.68	230443	190747	20.8	232887	1849	0.790	perftest-sysbench-h...
1980035	nightly-sysbench-hybrid	2022-02-01 05:47:08.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	602	2.66	233612	190747	22.5	232887	1849	0.790	perftest-sysbench-h...
1860026	nightly-sysbench-hybrid	2022-01-31 11:37:58.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	2.75	233012	190747	22.2	232887	1849	0.790	perftest-sysbench-h...
1890029	nightly-sysbench-hybrid	2022-01-30 10:58:01.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	602	2.70	231269	190747	21.2	232887	1849	0.790	perftest-sysbench-h...
1860022	nightly-sysbench-hybrid	2022-01-29 11:17:47.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	602	2.70	233790	190747	22.6	232887	1849	0.790	perftest-sysbench-h...
2010032	nightly-sysbench-hybrid	2022-01-28 11:37:19.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	2.80	230167	190747	20.7	232887	1849	0.790	perftest-sysbench-h...

Why Benchbot - Stabler

The performance of sysbench oltp_point_select workload is **stabler** and improved by **more than 20%**



oltp_summary	benchmark_name	test_start_time	baseline_test_time	bench_type	bench_sub_type	thread	duration(s)	avg_latency(ms)	qps_tpm	baseline_qps_tpm	improvement(%)	avg_qps_tpm	std_qps_tpm	std/avg(%)	testbed
1980035	nightly-sysbench-hybrid	2022-02-01 05:47:08.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	602	2.66	233612	190747	22.5	209770	21712	10.4	perftest-sysbench-h...
1860026	nightly-sysbench-hybrid	2022-01-31 11:37:58.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	2.75	233012	190747	22.2	209770	21712	10.4	perftest-sysbench-h...
1890029	nightly-sysbench-hybrid	2022-01-30 10:58:01.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	602	2.70	231269	190747	21.2	209770	21712	10.4	perftest-sysbench-h...
1860022	nightly-sysbench-hybrid	2022-01-29 11:17:47.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	602	2.70	233790	190747	22.6	209770	21712	10.4	perftest-sysbench-h...
2010032	nightly-sysbench-hybrid	2022-01-28 11:37:19.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	2.80	230167	190747	20.7	209770	21712	10.4	perftest-sysbench-h...
1950019	nightly-sysbench-hybrid	2022-01-27 08:36:46.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	2.76	234306	190747	22.8	209770	21712	10.4	perftest-sysbench-h...
1950017	nightly-sysbench-hybrid	2022-01-26 06:48:06.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	0.962	185013	190747	-3.0	209770	21712	10.4	perftest-sysbench-h...
1920023	nightly-sysbench-hybrid	2022-01-25 08:50:15.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	0.939	189185	190747	-0.82	209770	21712	10.4	perftest-sysbench-h...
1890015	nightly-sysbench-hybrid	2022-01-24 06:19:21.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	0.872	191190	190747	0.230	209770	21712	10.4	perftest-sysbench-h...
1830014	nightly-sysbench-hybrid	2022-01-23 11:00:20.000	2021-12-13 12:13:58.000	sysbench	oltp_point_select	800	601	0.963	189838	190747	-0.48	209770	21712	10.4	perftest-sysbench-h...

Why Benchbot - Together

Improve cross-team collaboration efficiency:

- Performance issues discovered in daily run and release testing can be easily reproduced with a single Benchbot command
- Since 2021, more than 80 major bugs or performance issues have been discovered via Benchbot
- Benchbot is the standard service to verify new performance features

Future Work

- Community Support: enable community contributors to submit tests to Benchbot on GitHub
- Test environment diversity: support testing on TiDB Cloud and other cloud platforms


PingCAP
TiDB
THANKS