TencentDB for CTSDB

Product Introduction

Product Documentation

Contents

Product Introduction

Overview

Last updated：2023-03-02 14:45:11

Overview

TencentDB for CTSDB (CTSDB for short) is Tencent Cloud's proprietary distributed and scalable time series database that supports near real-time data search and analysis. It is a non-relational database and provides efficient read/write, cost-effective storage, powerful aggregated data analysis capabilities, and convenient features such as instance monitoring and data query result visualization. Its entire system adopts a multi-node multi-replica deployment architecture, effectively ensuring high service availability and data reliability.

Benefits

CTSDB has the following strengths in processing massive amounts of time series data:

Highly concurrent writes: CTSDB adopts the policy where data is written into the memory first and then periodically dumped as immutable files for storage. It can also write data in bulk to reduce network overheads.

Cost-effective storage: CTSDB aggregates historical data through data rollup to save storage space. Plus, it uses appropriate encoding and compression algorithms to increase the data compression ratio.

Powerful aggregated data analysis capabilities: CTSDB not only supports basic aggregate analysis functions such as Min, Max, Sum, and Avg but also supports complex ones such as Group By, Interval, Geo, and Nesting.

Product Architecture

General cluster architecture

A general cluster is a distributed cluster consisting of multiple nodes.

All nodes receive external requests, interconnect and collaborate with each other to provide services such as data storage and indexing (client requests can be distributed to appropriate nodes), and are master-eligible.

You won't need a dedicated master node if your cluster has less than 30 nodes, because the general cluster architecture will suffice.
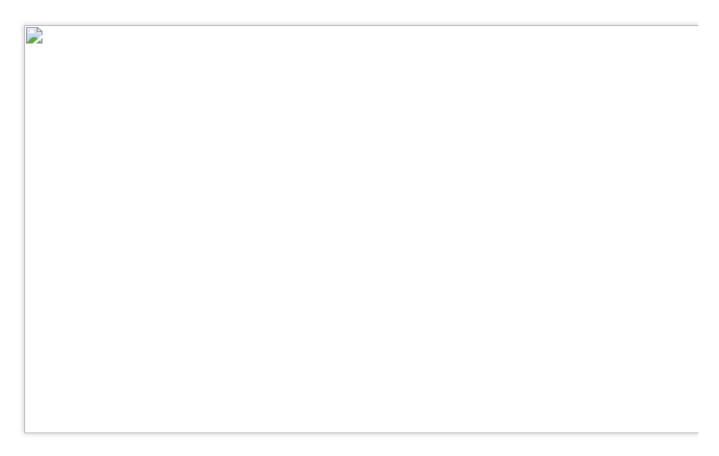
Notes

You can add a dedicate master node to a CTSDB cluster to optimize and upgrade the general cluster architecture to hybrid node cluster architecture.

Hybrid node cluster architecture

A hybrid node cluster is a distributed cluster consisting of dedicated master-eligible nodes and data nodes.

The dedicated master node is responsible for guaranteeing the health and stability of the entire cluster and doesn't provide services such as data storage. Data nodes provide services such as data storage and indexing.

As the business grows and the data volume increases, when the number of nodes exceeds 30, we recommend that you add a dedicated master node to upgrade the general cluster architecture to hybrid node cluster architecture so as to fully exert the performance of the large multi-node cluster.

Notes

A CTSDB cluster cannot be downgraded from the hybrid node cluster architecture to the general cluster architecture.

Strengths

Last updated：2021-11-05 10:31:14

High Performance

CTSDB supports bulk write, high-concurrency query, and cluster scaling to linearly improve the system performance.

Ease of Use

CTSDB offers a rich set of data types and is compatible with common APIs of Elasticsearch. Its console provides many data management and OPS features, making OPS simpler.

High Reliability

CTSDB supports multi-replica distributed deployment and automatic data balancing.

Low Costs

CTSDB uses rollup to increase the compression ratio and reduce storage costs.

Powerful Aggregate Analysis Capabilities

CTSDB supports commonly used aggregations such as Max, Min, Avg, Percentile, Sum, and Count as well as complex aggregations such as Script, Time Interval, Geo, and Nesting.

Performance

Last updated：2022-06-09 11:20:25

Performance Overview

CTSDB consists of nodes. The specification and number of single nodes determine the processing capacity of the CTSDB instance. In theory:

CTSDB instance read/write concurrency performance = (single node performance * number of nodes)

Therefore, the higher the node specification and quantity, the higher the processing capacity of the instance. The performance of a single node is mainly subject to CPU and memory. **The specific instance performance is subject to factors such as single node configuration, number of nodes, and number of written fields.** The test results listed in this document are reference values under specified parameters and only for your reference. To have a better picture, please conduct tests in your real business environment.

Performance Test

Test tool

Download the test tool.

Directions

1. Create a metric

The command is as follows:

```
# Create a metric (note that there is only one `host` in `tags`; however, if ther
e are new fields in the written data, they will be automatically put in `tags` as
tags)
curl -u {user}:{passwd} -XPUT {ctsdb_ip_port}/_metric/testa10?pretty -d '{
"tags": {
"http_code": "string"
},
"fields": {"count1":"long","count2":"long","count3":"long","count4":"long","count
5":"long","count6":"long","count7":"long","count8":"long","count9":"long","count1
0":"long"
},
"time": {
"name": "timestamp",
"format": "epoch_second"
},
```

```
"options": {
"expire_day": -1,
"refresh_interval": "10s",
"number_of_shards": 3,
"number_of_replicas": 1,
"rolling_period": -1
}
}'
# Query a metric
curl {ctsdb_ip_port}/_metric/testa10?pretty
# Delete a metric (all corresponding data will also be deleted)
curl -XDELETE {ctsdb_ip_port}/_metric/testa10
```

> Note：
>
> Here, `{ctsdb_ip_port}` is the access port of CTSDB, `{user}` the username, and `{passwd}` the
> password.

2. Write data

Write data in bulk by using the script available here. The parameters are as detailed below:

```
- db_url string
Instance VIP and Vport (example: 10.02.36.89:9200)
- metric_name string
Name of the metric to be written
- data_num int
Number of records to be written by the client at a time
- threads_nmb int
Number of concurrent writes
- counts int
Number of fields in a record
```

Modify `userpwd` in line 18 in the script before running it.

Sample code:

```
python testa.py 10.0.1.10:9200 testa10 4000 9 10
```

Script output:

It consists of two parts: `params` (option parameters) and `results` (final results).

```
-------------- params --------------
put_url http://10.0.1.10:9200/testa10/doc/_bulk
```

```
dataNum: 4000
threads_nmb: 9
counts: 10
-------------- results --------------
start all threads 2018-12-25 20:10:24
exit all threads 2018-12-25 20:10:24
startTime: 1545739824.51
endTime: 1545739824.55
diffTime: 0.0414531230927
wps: 158450.850363
```

Note：

The average write speed is the output of the `wps` field.

Performance reference values

Note：

You can select any CTSDB node configuration and number of nodes as desired. This document provides reference values of only three instance configurations. You can evaluate the performance of other configurations according to the performance evaluation formula in Performance Overview or test the performance with a test script.

Number of concurrent threads: 9

Number of written fields: 10

Single Node Configuration	Number of Nodes	Write Capacity
1 CPU core and 4 GB memory	3	30,000–50,000 data points per second
4 CPU cores and 20 GB memory	3	90,000–120,000 data points per second
8 CPU cores and 40 GB memory	3	110,000–150,000 data points per second

Regions and AZs

Last updated：2023-03-02 14:44:49

TencentDB data centers are hosted in multiple locations worldwide. These locations are known as regions. Each region contains multiple availability zones (AZs).

Each region is an independent geographic area with multiple isolated AZs. Separate AZs in the same region are connected via low-latency private networks. Tencent Cloud allows you to distribute Tencent Cloud resources across different locations. We recommend placing resources in different AZs to eliminate single points of failure which may lead to service unavailability.

Region name and AZ name can most directly embody the coverage of a data center. The following naming convention is used for your convenience:

A region name is composed of **region + city**. The `region` indicates the geographic area that the data center covers, while the `city` represents the city in or near which the data center is located.

AZ names utilize the format of **city + number**.

Regions

Tencent Cloud regions are isolated. This guarantees the maximum cross-region stability and fault tolerance. When purchasing Tencent Cloud services, we recommend selecting the region closest to your end users to minimize access latency and improve download speed. Operations such as launching or viewing instances are performed at the region level.

Private network communication:

Tencent Cloud resources in the same VPC within the same region under the same account can communicate with each other over private network. They can also be accessed at the private IP as described in Private Network Access. The networks of different regions are isolated from each other, and Tencent Cloud services in different regions cannot communicate over private network by default.

Tencent Cloud services in different VPCs can communicate with each other over CCN which is fast and stable.

AZs

An availability zone (AZ) is a physical IDC of Tencent Cloud with independent power supply and network in the same region. It can ensure business stability, as failures (except for major disasters or power failures) in one AZ are isolated without affecting other AZs in the same region. By starting an instance in an independent AZ, users can protect their applications from being affected by a single point of failure.

Supported Regions and AZs

Notes

Resources available in different regions and AZs may be sold out and become unavailable, and previously sold-out resources may be replenished. The resource availability will be assessed and adjusted based on the actual business usage as displayed on the purchase page in the console.

China

Region	AZ
South China (Guangzhou) ap-guangzhou	Guangzhou Zone 4 ap-guangzhou-4
East China (Shanghai) ap-shanghai	Shanghai Zone 2 ap-shanghai-2
	Shanghai Zone 3 ap-shanghai-3
North China (Beijing) ap-beijing	Beijing Zone 1 ap-beijing-1
Southwest (Chongqing) ap-chongqing	Chongqing Zone 1 ap-chongqing-1
Hong Kong/Macao/Taiwan (Hong Kong, China) ap-hongkong	Hong Kong Zone 3 (Hong Kong nodes cover services in the China regions of Hong Kong, Macao, and Taiwan) ap-hongkong-3

Other countries and regions

Region	AZ
Southeast Asia Pacific (Singapore) ap-singapore	Singapore Zone 2 (Singapore nodes cover services in Southeast Asia) ap-singapore-2
Southeast Asia Pacific (Bangkok) ap-bangkok	Bangkok Zone 2 (Bangkok nodes cover services in Southeast Asia) ap-bangkok-2
Southern Asia Pacific (Mumbai) ap-mumbai	Mumbai Zone 2 (Mumbai nodes cover services in South Asia) ap-mumbai-2

Northeast Asia (Seoul) ap-seoul	Seoul Zone 2 (Seoul nodes cover services in Northeast Asia) ap-seoul-2
Silicon Valley na-siliconvalley	Silicon Valley Zone 2 (Silicon Valley nodes cover services in US) na-siliconvalley-2
North America (Toronto) na-toronto	Toronto Zone 1 (Toronto nodes cover services in North America) na-toronto-1
Europe Zone (Frankfurt) eu-frankfurt	Frankfurt Zone 2 (Frankfurt nodes cover services in Europe) eu-frankfurt-2

Selection of Regions and AZs

When you purchase Tencent Cloud services, we recommend that you select the region closest to your end users to minimize access latency and improve download speed.