Assignment-2

MF1632008 陈硕

1. Quality attributes for Big Data Platform

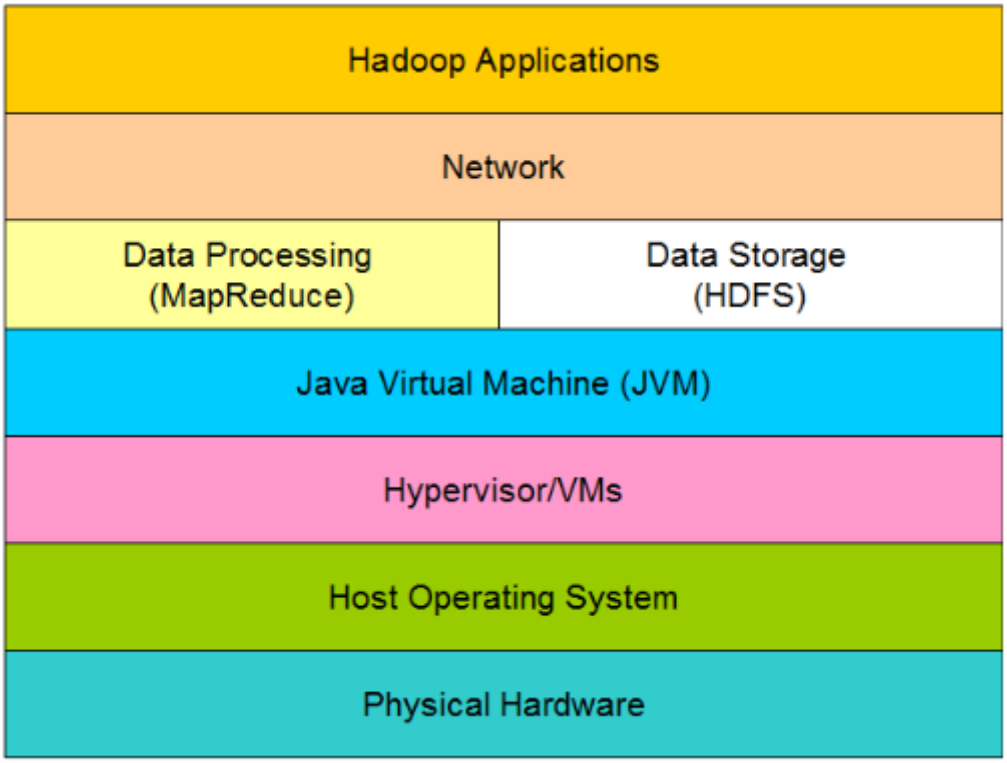
Quality attributes	Reason
Performance	The size of big data is easily recognized as an obvious challenge. Furthermore, a large percentage of the data might not be of interest. Data volume is increasing faster than computing resources and processor speeds that exist in the marketplace. Response times for results are still critical, despite the increase of data size. To ensure speed and real-time feedback from big data, We need good performance.
Security	In big data environments, data is routinely replicated and migrated among a large number of nodes. In addition, sensitive information can be stored in system logs, configuration files, disk caches, error logs, and so on. Security and privacy issues are magnified by velocity, volume, and variety of big data, such as large-scale cloud infrastructures, diversity of data sources and formats, streaming nature of data acquisition and high volume inter-cloud migration. The use of large scale cloud infrastructures, with a diversity of software platforms, spread across large networks of computers, also increases the attack surface of the entire system
Reliability	The increase of unstructured data has a large impact on scalability. Data reliability relates to the limits of data density at tolerable device-level bit error rates. Traditional RAID does not provide the levels of data durability and performance for dealing with escalating volumes of data. For disks, there is the end of life, stress modes, and overheating in data centers to consider.

2. Solutions for quality attributes

Quality Attributes	Solutions
Performance	Response times for results are still critical, despite the increase of data size. To ensure speed and real-time feedback from big data, a new approach is emerging where data sets are processed entirely within a server's memory. Several vendor solutions offer in-memory analytics to address the requirements for real-time analysis results (for example, IBM InfoSphere Streams and BigInsights, SAS Visual Analytics, and SAP HANA).

	These apply in scenarios where there is a high velocity of data or real-time return of results. The main limitation with traditional business intelligence technologies is the time that it takes to read from and store to disk (disk I/O). By storing the data in memory (RAM), this limitation is removed. However, this needs to be balanced with the higher cost of RAM, compared to disk storage, to determine whether business requirements justify the additional expense.
Security	1. Application Software Security: Use secure versions of open-source software. As described above, big data technologies weren't originally designed with security in mind. 2. Maintenance, Monitoring, and Analysis of Audit Logs 3. Secure Configurations for Hardware and Software. Build servers based on secure images for all systems in your organization's big data architecture. 4. Account Monitoring and Control

Hadoop: an open source frameword for big data distributed applications



3. Tencent Cloud

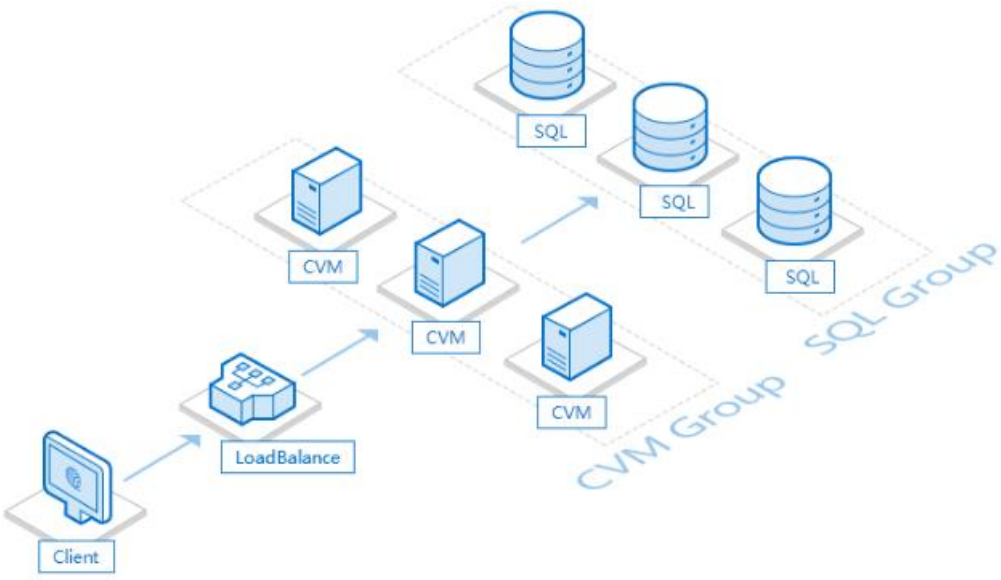
3.1. Introduction

Tencent Cloud is a secure, reliable and high-performance cloud compute service provided by Tencent. Tencent is now the largest Internet company in China, and even Asia. It's providing services for hundreds of millions of people via its flagship products like QQ and WeChat.

Services	Production
Compute	Cloud Virtual Machine, Cloud Load Balance, Virtual Private Cloud, Auto Scaling, Cloud Physical Machine, Cloud Block Storage
Storage and CDN	Cloud Object Storage, Content Delivery Network
Database	Cloud Database, Cloud Redis, Cloud MongoDB, Cloud Memcached
Security	Cloud Security, Dayu Distributed Defense, Business Security Protection

3.2. Cloud Load Balance

Cloud Load Balance (CLB) distributes access traffic from multiple public IP to CVMs, automatically detects CVM ports, and removes unavailable CVMs, thereby improving service availability.



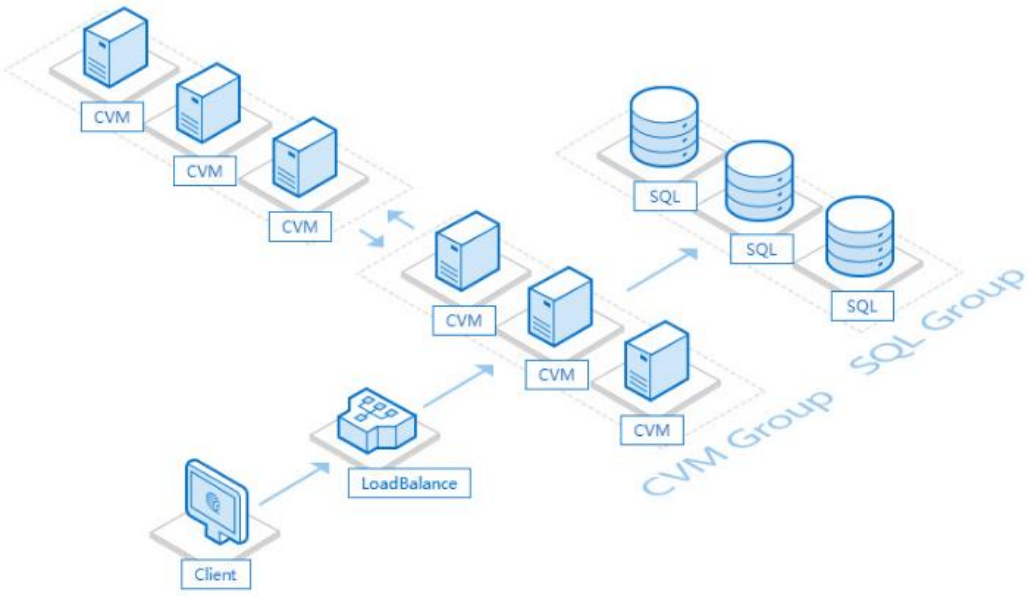
Service isolation

It supports over 100 million diversified Web services.

A Virtual Service Address (VIP) is set to virtualize multiple CVMs as a high-performance, highly available application service pool.

The CLB checks the health status of CVMs in the CVM pool and automatically isolates abnormal CVMs.

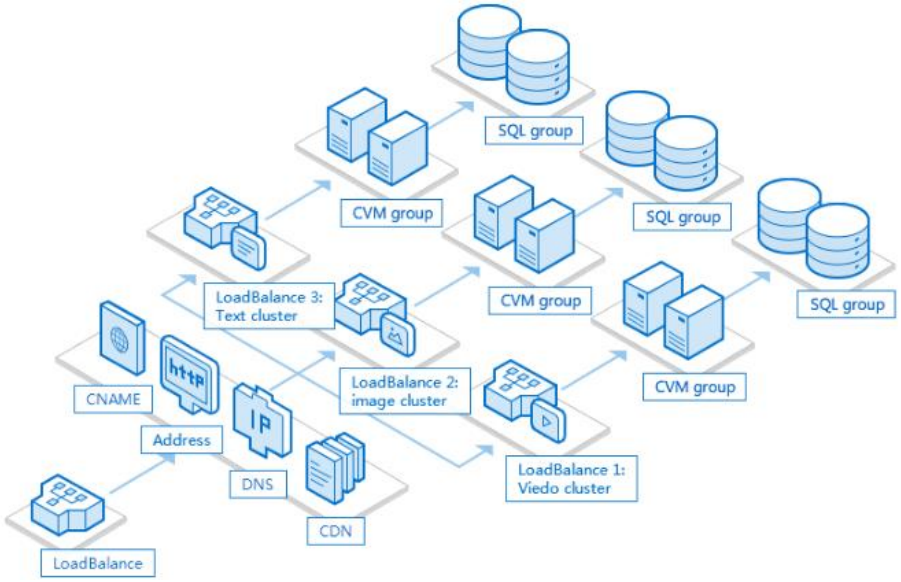
The CLB is created for clouds and eliminates CVM single-point risks.



Horizontal scaling

The CLB can expand the service capability of an application system by using the horizontal scaling function.

It combines with Auto Scaling to automatically create and release CVM instances and configure load balance, with no need to manually intervene and estimate resources, thereby reducing costs. The CLB creates and terminates CVMs based on monitoring status. Combined with 2nd-level charging, it makes full use of your computing resources.



Traffic distribution

1. For typical Web services such as forums and websites, it is recommended that different services such as pictures and texts be separated.
2. After splitting, each CLB instance and associated CVM group function as an independent service cluster.
3. The DNS service providers such as DNSpod supports CNAME switching for association between different services.

3.3. Cloud Security and Performance

Performance:
A CLB cluster supports up to 120 million simultaneous connections.It can handle heavy traffic with a peak of 40 Gbps, with a processing capability of 6 million packets per second.It is tolerant to diversified services and efficiently supports more than 100 million Web services.
The CLB can expand the service capability of an application system by using the horizontal scaling function.It combines with Auto Scaling to automatically create and release CVM instances and configure load balance, with no need to manually intervene and estimate resources. The CLB creates and terminates CVMs based on monitoring status. Combined with 2nd-level charging, it makes full use of your computing resources.

Security:
Cloud Security provides all-around security services for our customers; including network protection, intrusion detection, vulnerability protection, etc. It also produces security reports and solutions regularly, allowing for better insight into your business security status.
The DAYU system provides distributed anti-DDoS protection with TB bandwidth; high-defense zones allow for hundreds of GBs of capacity for a single IP; supports various protocols.
Nation-wide deployment and distributed detection; detects DNS hijacking quickly and accurately through comprehensive data collection and analysis.
Detect and stop malicious activities like intrusion, exploitation and web vulnerabilities; blocks malicious Web crawlers.
Simple APIs are provided to help you obtain professional security capacity for login, registration, UGC and promotions.