

How Data Centers Provide Consumer Services

Original author: Tom Coughlin

November 2, 2017

Introduction

Why we need data centers

Mobile devices have the ability to store and process data, and we can do these locally. But

Introduction

Why we need data centers

Mobile devices have the ability to store and process data, and we can do these locally. But

- the storage capacity is very limited (usually less than 128 GB)
- the processing power is also limited

Introduction

Why we need data centers

Mobile devices have the ability to store and process data, and we can do these locally. But

- the storage capacity is very limited (usually less than 128 GB)
- the processing power is also limited

So, we can transfer and store the data in a cloud server, process on the cloud server, and get back the result.



Figure: Mobile devices depend upon cloud servers

Content Delivery

Disadvantages of single content server

It is of low efficiency if all content is delivered from a central server (or cluster), especially for those users far from server:

- the high latency may lower user experience
- the cost of data transmission may be large

Content Delivery

Disadvantages of single content server

It is of low efficiency if all content is delivered from a central server (or cluster), especially for those users far from server:

- the high latency may lower user experience
- the cost of data transmission may be large

We can use edge servers to help deliver the content, which are geographically distributed all over the world. Edge servers store copies of frequently visited content, while central server provide those that are less frequently visited.

Content Delivery

Content distribution system

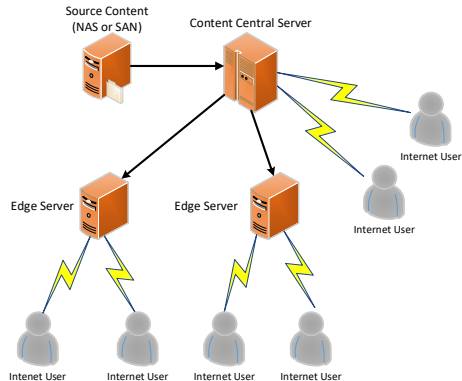


Figure: An Internet content distribution system

Content Delivery

Comparison between single content server and CDN

	Single Server	CDN
throughput	low	high
latency	high	low
cost	high	low
availability	low	high
reliability	low	high
update	easy	difficult

Table: Comparison between single content server and CDN

Content Delivery

Grid network

- In grid network, content is available on much more closely managed independent storage nodes.

Content Delivery

Grid network

- In grid network, content is available on much more closely managed independent storage nodes.
- Grid network are used in peer-to-peer content delivery, where all peers are equal.

Content Delivery

Grid network

- In grid network, content is available on much more closely managed independent storage nodes.
- Grid network are used in peer-to-peer content delivery, where all peers are equal.
- All nodes in grid network are independent, but well coordinated.

Content Delivery

Grid network

- In grid network, content is available on much more closely managed independent storage nodes.
- Grid network are used in peer-to-peer content delivery, where all peers are equal.
- All nodes in grid network are independent, but well coordinated.
- Grid network can make full use of existing bandwidth, and dynamically optimize the route of delivery.

Content Delivery

Grid network

Although grid network have so many advantages, hierarchical CDN are still widely used.

Content Delivery

Grid network

Although grid network have so many advantages, hierarchical CDN are still widely used.

- There exists asymmetry between servers and clients: servers send large amount of data, while clients often receive little.

Content Delivery

Grid network

Although grid network have so many advantages, hierarchical CDN are still widely used.

- There exists asymmetry between servers and clients: servers send large amount of data, while clients often receive little.
- Not all nodes are willing to help transfer the data, nor to store the content.

Content Delivery

Grid network

Although grid network have so many advantages, hierarchical CDN are still widely used.

- There exists asymmetry between servers and clients: servers send large amount of data, while clients often receive little.
- Not all nodes are willing to help transfer the data, nor to store the content.
- Grid network may face several security problems.

Content Storage

Storage media

Two types of storage media are used in content delivery system:
hard disk drives and flash memory.

Content Storage

Storage media

Two types of storage media are used in content delivery system:
hard disk drives and flash memory.

- Hard disk drives are cheap, while flash memory is expensive;

Content Storage

Storage media

Two types of storage media are used in content delivery system:
hard disk drives and flash memory.

- Hard disk drives are cheap, while flash memory is expensive;
- Hard disk drives have larger storage capacity than flash memory;

Content Storage

Storage media

Two types of storage media are used in content delivery system:
hard disk drives and flash memory.

- Hard disk drives are cheap, while flash memory is expensive;
- Hard disk drives have larger storage capacity than flash memory;
- Flash memory are much faster than hard disk drives.

Content Storage

Heterogeneous storage systems

Because both two types of storage media have their own advantages, we can combine these two media to better the performance. This lead to the idea of heterogeneous storage.

Content Storage

Heterogeneous storage systems

Because both two types of storage media have their own advantages, we can combine these two media to better the performance. This lead to the idea of heterogeneous storage.

- NAS, SAN, or object storage system, likely uses hard disk drives, to store massive data.
- Central server or edge servers, may use both hard disk drives and flash memory, to accelerate the delivery of the content.

Data Compression

Definition

Data compression is the process to reduce the number of bits used to represent the information. Decompression is the inverse process of data compression.

Data Compression

Definition

Data compression is the process to reduce the number of bits used to represent the information. Decompression is the inverse process of data compression.

Data compression can be either lossless or lossy:

- Lossless compression reduces the number of bits by eliminating statistical redundancy. It can be used for general purpose.
- Lossy compression reduces the number of bits by removing unnecessary or less important information. It is usually used in compressing multimedia content.

Data Compression

Data compression in content delivery

Since bandwidth is expensive, data compression is very meaningful in content delivery. However, compression and decompression also take a lot of time, which may increase the latency.

Data Compression

Data compression in content delivery

Since bandwidth is expensive, data compression is very meaningful in content delivery. However, compression and decompression also take a lot of time, which may increase the latency.

Some standardized compression algorithms, such as H.264 and H.265, balance the compression rate and time cost, so they are suitable for Internet content delivery.

Summary

- Hierarchical CDN
- Heterogeneous storage system
- Data compression

Q & A