

Alibaba Cloud CDN

Product Introduction

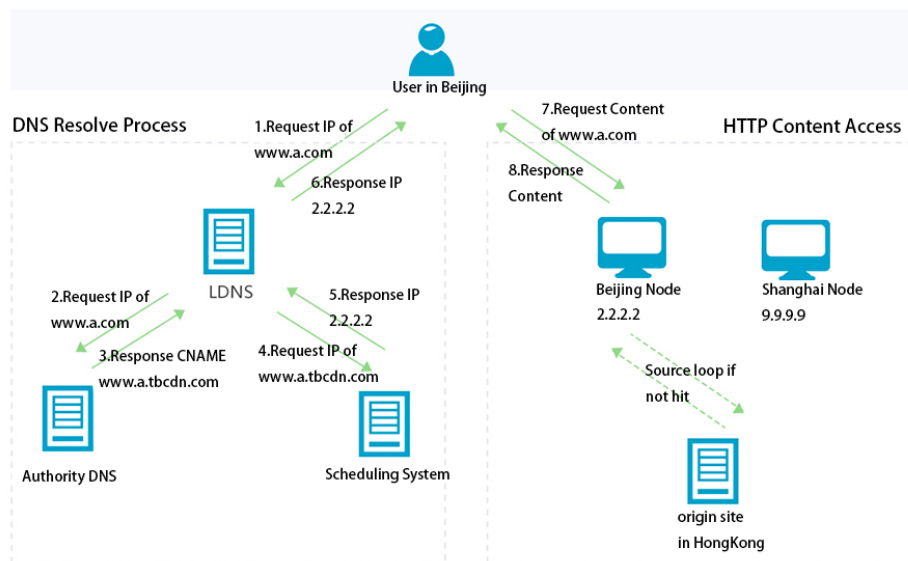
Product Introduction

Introduction

Alibaba Cloud Content Delivery Network, shortened to CDN, is a distributed network built on, and overlaying, the bearer network, and is composed of edge node server clusters distributed across different regions. A CDN replaces the traditional data transmission mode centered on web servers.

Alibaba Cloud CDN delivers the source content to edge nodes, and works to a precise scheduling system. It distributes user requests to the most suitable nodes, allowing the user to retrieve the content they need as quickly as possible, reducing congestion of traffic across the Internet, and increasing the response speed of users visiting websites.

The process for http request handling, after using CDN is as follows:



Function introduction

Node cache

The object heat algorithm and hierarchical HOT cache resources allow for precise resource acceleration.

The high-performance cache system design, use of multiple, balanced CPU cores, and efficient and control of memory maximizes SSD IOPS and throughput.

Each node provides high-speed read/write SSD storage. In combination with the SSD acceleration capabilities, this greatly reduces user access waiting time and improves availability.

Smart compression effectively reduces the size of the content transmitted by users, accelerating distribution.

Page optimization removes spaces, line breaks, TABs, annotations, and other redundant page content, reducing page sizes.

Multiple JavaScript/CSS files are combined into a single request, reducing the number of requests.

Precise scheduling

Alibaba Cloud CDN allocates scheduling domains that improve the overall speed of sites. It caters for a variety of sites such as portal information site, multimedia audio and video site, game site, mobile apps and more. CDN benefits include:

A scheduling system that supports scheduling for millions of domain names with a single machine.

Finer control capabilities and protocol scalability in order to help reduce costs.

Multi-level scheduling policies so that node failure will not affect user availability.

Multi-system interaction that seamlessly coordinates with security defense systems, refresh systems, and content management systems.

Real-time data scheduling and support for node-level traffic prediction which improves the quality and accuracy of scheduling.

Business support for multiple scenarios

The audio and video incremental on-demand service provides low buffering times, a smooth playback experience, and supports MP4 and FLV video formats.

Resource link authentication and custom authentication keys are supported to ensure the security of your media resources and provide protection against leeching.

Through integration with multiple Alibaba Cloud services, CDN works seamlessly with different services to increase cloud resource access and download speeds. Such services include:

Object Storage Service (OSS), which works with CDN to accelerate website access and effectively reduce OSS Internet traffic fees.

Elastic Compute Service (ECS) works with CDN to increase site availability, protect the origin site information on the server, and reduce bandwidth usage cost.

Server Load Balancer for the origin site address to go back to the source for data, reducing the retrieve from the origin bandwidth pressure.

CDN also supports non-Alibaba Cloud origin sites and provides unobstructed access and rapid deployment of acceleration services following resource reviews.

Alibaba Cloud live streaming media service provides an integrated solution for media asset storage, slicing and transcoding, access authentication, and content delivery acceleration.

Self-management

Use the self-help console for minute-level deployment of all nodes using custom configurations.

Use the CDN console to add, delete, modify, and query self-configured domain names, as well as set node cache acceleration policies, anti-leeching measures, http header information, etc. You can also enable different types of acceleration and optimization functions as needed. [Learn more.](#)

The open and scalable CDN APIs enable flexible deployment, fast operations, precision usage, and timely monitoring of CDN domains, distribution resources, and monitoring data. It can also be used with the APIs of other Alibaba Cloud products for a custom, multi-platform portal. [Learn more.](#)

Real-time monitoring

The all-round information monitoring feature provides multi-dimensional support for resource distribution.

Full network monitoring, data analysis, and resource report downloads, provide a wide range of monitoring information, including information on bandwidth traffic, access quality, visitor data, popularity analysis, and security protection.

The advantages of applying CDN are described in the following section.

Stable and fast

Advanced distributed system architecture: almost 1000+ nodes in China and over 200+ abroad.

Adequate bandwidth and storage resources: a single node provides a bandwidth of more than 40 Gbps and a storage capacity of 40 TB to 1.5 PB.

Stable and efficient performance indicators: greater than 95% hit rate and video fluency rate, as well as millisecond-level response time.

Mature monitoring and service systems: 24/7 network-wide monitoring, smart monitoring and scheduling based on service quality.

Cost effective

With scalable resources, you are charged only for resources you use, and can achieve cross-carrier, cross-region network-wide coverage.

Two billing methods are provided, PayByTraffic and PayByBandwidth, to satisfy different business needs.

The service automatically responds to site traffic spikes and makes adjustments without user intervention, reducing pressure on the origin site.

Easy to use

You can add, delete, modify, and query domain names on your own by using a wide range of simple custom configuration options. CDN supports customizing anti-leech measures, cache policies, HTTP request headers, and other functions.

The open API interfaces provide functions such as service activation, content refreshing, monitoring data retrieval, and distribution log downloads.

This section describes different CDN application scenarios.

Website/application acceleration

To accelerate the distribution of resources for websites, or applications, with a large volume of static resources, you can separate the dynamic and static contents. The dynamic files can be stored on ECS. For large volumes of static resources such as image, HTML, CSS, and JS files, we recommend that you store them on OSS. This can accelerate content download speeds and make it easy to perform distribution for images, videos, and other media content.

Acceleration of on-demand audio and video/large file downloads and distribution

Alibaba Cloud CDN supports the downloading and distributing of various types of files. It also supports the acceleration of online on-demand streaming services, for example, MP4 and FLV videos, or scenarios where the average size of a single file is greater than 20 MB. The primary service scenarios are on-demand video/audio and large file downloads (for example, installation packages). You can combine CDN with OSS to increase origin retrieval speeds to reduce origin retrieval bandwidth cost by nearly 60%.

Acceleration of live streaming media

The live streaming media service provides an integrated solution for media asset storage, slicing and transcoding, access authentication, and content delivery acceleration. Combined with Alibaba Cloud Auto Scaling, CDN can promptly adjust server bandwidth and respond to sudden access traffic bursts. When combined with the media transcoding service, CDN can provide high-speed and stable concurrent transcoding and seamless task scaling.

Mobile application acceleration

For the distribution of mobile app update files (APK files), CDN delivers optimized and accelerated distribution of in-app images, pages, short videos, UGC, and other media content. The http DNS

service prevents DNS hijacking and retrieves precise DNS resolution results in real time, effectively reducing user access time and improving user experience.

Architecture

The following section details key components featured in the CDN architecture.

Key components

LVS performs Layer-4 server load balancing

DR mode

Dual-LVS performs active-active mutual backup

WRR is the algorithm used for server load balancing

Tengine performs Layer-7 server load balancing

The Alibaba Cloud high-performance HTTP servers developed on Nginx are open-source. For more details, refer to the [official website](#).

Proactive health checks

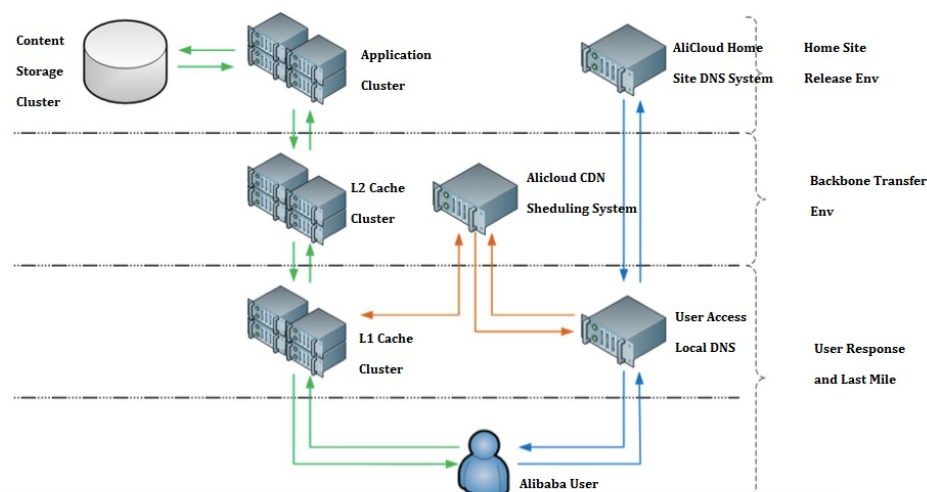
SPDY v3 support

Swift performs HTTP caching

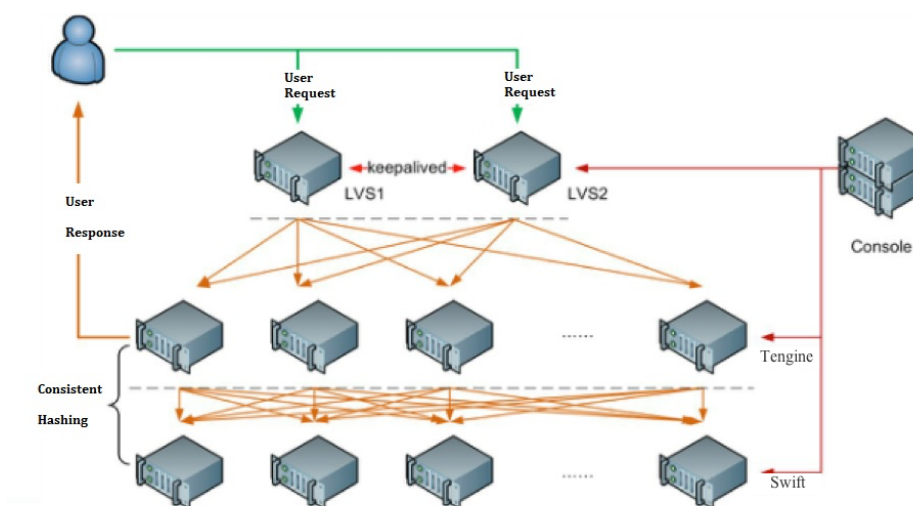
High-performance cache

Disk (SSD/SATA)

Architecture diagram



Deployment architecture diagram



Domain name

A domain name is a server or network system name that identifies IP resources connected to the Internet. All domain names are unique worldwide.

CNAME record

A Canonical Name (CNAME) record is a type of resource record in the Domain Name System (DNS) used to specify that a domain name is an alias for another domain, which is the "canonical" domain.

CNAME domain name

CDN domain acceleration needs to use the CNAME records.

You will receive an accelerated domain name which is the CNAME domain name (this domain name should be *.kunlun.com) after CDN acceleration on the Alibaba Cloud console is configured. The domain name resolution is then formally transferred to Alibaba Cloud after you reference the domain name *.kunlun.com following the CNAME operation.

All requests for this domain name will be transferred to the Alibaba Cloud CDN nodes.

DNS

DNS stands for Domain Name System. It refers to the domain name resolution service. Its function on the Internet is to convert a domain name to an IP address that can be recognized by the network.

The domain name and IP address are synonymous to each other and the task of converting between the two is called domain name resolution. Domain name resolution requires a dedicated domain name resolution server to complete the task, and the entire process runs automatically.

For example, the domain name www.baidu.com is automatically converted to 220.181.112.143.

Edge node

Edge node is also referred to as CDN node or cache node.

Edge node is a concept proposed in contrast to the complex structure of the network. It refers to network nodes with a smaller number of accessible links. This improves the response capability and connection speed for the end user.

Edge nodes are used to store webpage contents and objects with a higher traffic volume in specialized cache machine on the front-end of the server. This improves the speed and quality of website access.

origin retrieval host

origin retrieval host: The origin retrieval host determines the site to which a origin retrieval request is sent.

Example 1: The origin site is a domain name.If the origin site is www.a.com and the origin retrieval host is www.b.com,the actual origin retrieval request is sent to the IP address resolved from www.a.com corresponding to the site www.b.com on the host.

Example 2: The origin site is an IP address.If the origin site is 1.1.1.1 and the origin retrieval host is www.b.com,then, the actual origin retrieval request is sent to the site www.b.com on the host

corresponding to 1.1.1.1.

Protocol-based origin retrieval request

The same protocol is used for sending origin retrieval requests and accessing resources from the client. That is, if the client initiates an HTTPS request for resources but the resources are not cached on the CDN node, the node initiates a origin retrieval HTTPS request for the resources. The same is true for HTTP requests.

Filter parameters

When a URL request includes a question mark (?) and request parameters are sent to a CDN node, the CDN node determines whether to send the request to the origin site. If the "Filter Parameter" function is enabled, after the request arrives at the CDN node, the URL without parameters will be intercepted and requested against the origin site. In addition, the CDN node keeps only one copy. If the "Filter Parameter" function is disabled, different copies are cached on the CDN node for different URLs.

Example of use

The `http://www.abc.com/a.jpg?x=1` URL request is sent to a CDN node. After the "Filter Parameter" function is enabled, the CDN node initiates the `http://www.abc.com/a.jpg` request (ignore the parameter `x=1`) to the origin site. After the origin site returns a response, the CDN node keeps a copy. Then, the origin site sends a response on `http://www.abc.com/a.jpg` to the terminal. For all requests similar to `http://www.abc.com/a.jpg?` parameters, the origin site responds to the content of CDN copy `http://www.abc.com/a.jpg`. When the "Filter Parameter" function is disabled, different copies are cached on the CDN node for different URLs. For example, different contents are returned by the origin site in response to `http://www.abc.com/a.jpg?x=1` and `http://www.abc.com/a.jpg?x=2`.

Date	Event Description
2008~2011	Taobao CDN was launched in order to serve the Taobao site.
2011~2014-02	Taobao CDN evolved into Alibaba Cloud CDN, which was to provide services for all subsidiaries of the Alibaba Group.
2014-03-21	The Alibaba Cloud CDN service was officially launched and provided for external sales
2015-05-22	Responding to the call of the premier, Alibaba Cloud reduced the full price of CDN by 21%.

2015-06-04	Customization of the 404 Page function was deployed.
2015-06-18	Alibaba Cloud released the new OpenAPI to support the addition, deletion, modification, and query of CDN domains.
2015-07-31	The on-demand streaming media acceleration solution was deployed, along with the on-demand authentication function.
2015-08-27	The Set HTTP Request Header function was deployed.
2015-09-24	Support for resource monitoring traffic report export was added to provide multi-dimensional support for resource distribution.
2015-10-13	The priority function for custom cache configurations was deployed.
2015-12-29	The domain name configuration and resource monitoring APIs were deployed.
2016-02-02	The prepay CDN resource package was deployed.