ROBOCOAST™ DIH Digital Innovation HUB

Aarila Dots Oy

12.04.2019

v 1.0.1

Study of the Cloud Platform Ecosystem in China

REPORT

European Union
European Regional Development Fund

Leverage from the EU 2014–2020

SATAKUNTALIITTO
Regional Council of Satakunta

Prizztech

Table of Contents

1. Description of The Study

1.1. General Overview

This study was conducted to research and document the different options available for cloud infrastructure and service providers in Mainland China as well as their special requirements and differences from the perspective of European industries.

In order to confirm the technical requirements, some of the services and tools provided by the local cloud service providers were tested on small scale by setting up test accounts with different providers and researching virtual server provisioning APIs.

The main focus of the study was to research the use of cloud platforms from the perspective of the European industries working in China, taking into account the tools, software, services and methods commonly used by these industries, and the use cases they would have when setting similar cloud infrastructures in Mainland China.

1.2. Study Methods

In order to deploy servers and use cloud services in Mainland China, even for the purposes of testing them for this study, ICP was required. We were able to use an existing Chinese company and their ICP to fully test the services outlined in this study.

Due to the large quantity of the available services and service providers, in some cases applied testing was omitted in favor of researching through the available service documentation and API documentation listing all the available options and functionalities.

In order to achieve a comprehensive study of the differences in practice, a basic reference design of a cloud infrastructure was constructed to go through the different requirements and challenges as well as to test solutions for designing an infrastructure capable of being operated both in Europe and Mainland China with minor differences.

Many of the practical tests for this study were conducted on Alibaba Cloud services, while others were mostly researched on their APIs, documentation, available libraries, etc.

1.3. Disclaimer

The contents of this study, including but not limited to the information about the regulations, laws, available services, functionality and features of services, limitations, restrictions, instructions and other advice have been provided by Aarila Dots Oy for informational purposes only and are not intended as and do not constitute legal advice or opinion. We have made efforts to ensure the accuracy of the information provided, however, we do not promise or guarantee that the information is correct, complete or up

to date. Aarila Dots Oy expressly disclaims all liability in respect to actions taken or not taken based on any or all the contents of this study.

1.4. Authors of The Study

This study has been conducted by Aarila Dots and the study was led by the Chief Systems Architect of Aarila Dots; Toni Lähdekorpi. The extensive technical expertise of the rest of the Aarila Dots team was also utilized during this study.

The team members of Aarila Dots have worked in various different cloud platform development and research projects as well as have a good understanding of the Chinese culture, with most members having completed a course in the basics of the Chinese language, as well as one member having previously lived and worked in China.
This China-specific expertise includes some preliminary work done in investigating the different technical and legal requirements for operating cloud-based systems in Mainland China.

It is notable that one of the specific expertise of the team in cloud platform building extends beyond the use of commercial “off-the-shelf” solutions, to developing vendor agnostic custom solutions for different kinds of use cases, including self-hosted ones.
In addition to server and cloud expertise, the team also has full proficiency in multiple programming languages and different software platforms, databases and libraries, including but not limited to; JavaScript, TypeScript, Node.js, Go, Java, Kotlin, Objective-C, Swift, Python, PHP, C, C++, Bash, Delphi, Lua, C#, XSL, Angular, Vue, PostgreSQL, MySQL, MSSQL, SQLite, MongoDB, and Redis.

2. Cloud Service Providers

2.1. Overview and Findings

Overall the cloud services available in Mainland China are on par with the similar services in Europe with similar API structures and tools and sharing many functionalities.

The biggest differences come from restrictions placed on content hosting and the purchase of these services as well as the fact that the majority of these services are either mostly or fully in Chinese, including the technical documentation and API documentation needed for the implementation.

Another of the differences in the use of Chinese in most documentation, user interfaces, and other legal texts. With service providers offering services in Europe, these are usually all available in English.

2.2. Chinese Providers

This study focuses on the 14 Chinese cloud service providers with the most visibility online, and ones that provided most of the services listed in the study requirements.

These services researched for this study were: Alibaba Cloud (阿里云), Baidu Cloud (百度云), Baidu SU (百度云加速), Meituan Open Services (美团云), Tencent Cloud (腾讯云), UCloud (与云), QingCloud (青云), Huawei Enterprise Cloud (华为企业云), NetEase Cloud (网易蜂巢), Western Digital (西部数码), Elephant Cloud (象云), E Cloud (天翼云), Grand Cloud (盛大云), KS Cloud (金山云).

2.3. Western Providers

From the Western Providers, 2 of the only cloud service providers that provide some of the services listen in the study requirements were: Azure (operated by 21Vianet) and Amazon (operated by NWCD and Sinnet).

Due to China's legal and regulatory requirements, foreign-owned companies have either licensed their technologies or formed different kinds of collaborations with Chinese companies in order to be able to provide services in China. [1] [2] [3]

According to the People's Republic of China Telecom Regulation (中华人民共和国电信条例), cloud service providers must have telecommunication permits in order to provide cloud services in China. According to Article 10 of the regulation only, local companies with less than 50 percent foreign investment qualify for said permits. [4]

Western providers are a relatively recent addition to the list of service providers in China. [5] [6] While these providers might be the largest providers outside China, their number of services, tools, scale, and server locations are still not at the same level as local providers. [7] [8]

---

[1] Stanford Management Science and Engineering. (2018) Why China is taking over the cloud computing market
https://mse238blog.stanford.edu/2018/07/sebgomez/why-china-is-taking-over-the-cloud-computing-market/ [Accessed 22.12.2018]
[2] AWS. (2018) 在中国简介
https://www.amazonaws.cn/about-aws/china/?nc1=h_ls [Accessed 22.12.2018]
[3] Azure China Documentation. (2018) Azure operations in China vs. Global Azure
https://docs.microsoft.com/en-us/azure/china/china-overview-operations [Accessed 09.01.2019]
[4] 中华人民共和国工业和信息化部. (2016) 中华人民共和国电信条例
http://www.miit.gov.cn/n1146295/n1146557/n1146619/c4860613/content.html [Accessed 09.01.2019]
[5] AWS News Blog. (2017) Now Open – AWS China (Ningxia) Region
https://aws.amazon.com/blogs/aws/now-open-aws-china-ningxia-region/ [Accessed 05.01.2019]
[6] Microsoft Azure. (2018) Welcome to Azure China 21Vianet
https://docs.microsoft.com/en-us/azure/china/china-welcome [Accessed 06.11.2018]
[7] 微软云技术. (2018) 智能安全可信混合全球的 Azure 平台
https://www.microsoft.com/china/azure/ [Accessed 06.11.2018]

These restrictions have given an edge for the local providers, giving them more time to build their cloud platforms locally.

For the purpose of European industries, using these providers may speed up the development time when using the same tools and services as they may already use outside China. There are however major obstacles and the services may be completely different from how they work outside of China, even if they share the same name. [9] Customer support, server access, and other key parts may be operated by the local Chinese partner company. And some of the key features, like server root access, AWS KMS, AWS Route53 and others may not work at all. [10]

2.4. Providers Outside Mainland China

Most of the cloud service providers commonly used in Europe offer their services from datacenters geographically close to Mainland China, without being inside and subject to the same jurisdiction and restrictions. While connecting to these servers usually work, in our testing we have found these connections to be too unreliable for business-critical use.

Additionally, when deploying databases or other data storages or data processing services that handle personal data of people living inside Mainland China, juridical restrictions may also apply.

[8] Bloomberg. (2018) Google Is In China Cloud Talks With Tencent, Others https://www.bloomberg.com/news/articles/2018-08-03/google-is-said-to-be-in-china-cloud-talks-with-tencent-others [Accessed 05.11.2018]
[9] AWS Documentation. (2018) Documentation by Service https://docs.amazonaws.cn/en_us/aws/latest/userguide/services.html [Accessed 05.11.2018]
[10] The Register. (2018) No root for you, or how to stop worrying and love AWS China https://www.theregister.co.uk/2018/05/18/china_cloud_setup/ [Accessed 05.01.2019]

2.5. Provided Services

List of providers and the availability of services researched in this study can be found below. More details about the providers, website links and the documentation languages are listed in attachment 1 of this study.

Company name		**List of services required in the study**	
English	**Chinese**	**Provided**	**Missing**
Alibaba Cloud	阿里云	Machine Learning Platform For AI Web Application Firewall Cloud CDN Elastic Computing - Scalable Virtual Servers - Auto Scaling - Server Load Balancer - Various Security and Monitoring Resources Log Service CloudMonitor Database Services - ApsaraDB	
Baidu Cloud	百度云	AI development platform Infinite Baidu machine learning BML Load balancing BLB Application firewall WAF Content Distribution Network CDN Cloud server BCC Document database MongoDB	Centralized logging service Metrics and performance tools
Baidu SU	百度云加速	Application firewall WAF Content Distribution Network CDN	
Meituan Open Services	美团云	Load balancing Web application firewall WAF Content Distribution Network CDN Cloud host MongoDB	Artificial Intelligence Platform Centralized logging service Metrics and performance tools
Tencent Cloud	腾讯云	Cloud Load Balancer Web Application Firewall (WAF) Content Delivery Network Cloud Virtual Machine - Auto Scaling Cloud Log Service (CLS) Cloud Monitoring and Warning Cloud MongoDB Service	Artificial Intelligence Platform

UCloud	与云	AI online service UAI-Inference AI training service UAI-Train Load balancing ULB Web application firewall UWAF Cloud distribution UCDN Cloud host UHost Cloud monitoring UMon Cloud database MongoDB UDB	Centralized logging service
QingCloud	青云	Deep learning platform Load balancing Web application firewall CDN service Virtual host - Automatic expansion Operation and maintenance and monitoring MongoDB	Centralized logging service
Huawei Enterprise Cloud	华为企业云	Machine Learning Service Deep Learning Service Elastic Load Balance Web Application Firewall Content Delivery Network Elastic Cloud Server - Auto Scaling Log Tank Service Open platform for real-time resource monitoring, alarming, and notification Document Database Service (DDS) is a MongoDB-compatible database	
NetEase Cloud	网易蜂巢	Load balancing Web application firewall Content Delivery Network (CDN) Cloud Server Log service Application monitoring Alarm management MongoDB	Artificial Intelligence Platform
Western Digital	西部数码	Web Application Firewall (WAF) Cloud Server	
Elephant Cloud	象云	Load balancing CDN acceleration Public cloud Real-time monitoring and alarm services Cloud database service (MySQL)	Artificial Intelligence Platform Web Application Firewall Centralized logging service

E Cloud	天翼云	Elastic load balancing Web application firewall CDN content distribution Elastic cloud host Log audit Cloud audit Cloud monitoring service Document database service (MongoDB protocol)	Artificial Intelligence Platform
Grand Cloud	盛大云	Load balancing Cloud host Cloud monitoring Database MySQL	Artificial Intelligence Platform Web Application Firewall Distributed Content Delivery Network Centralized logging service
KS Cloud	金山云	Deep Learning Platform (KDL) Machine Learning Platform (KML) Load balancing (SLB) Web Application Firewall (WAF) Content Distribution Network (CDN) Cloud Server (KEC) Cloud Monitoring Cloud database (MongoDB)	Centralized logging service
Azure	.	Cognitive Services Load Balancer Application Gateway Content Delivery Network (CDN) Virtual Machines Azure Monitor Azure Cosmos DB (MongoDB 3.2 compatible)	Centralized logging service
Amazon	-	AWS Deep Learning AMIs Elastic Load Balancing Amazon CloudFront (CDN) Amazon EC2 Amazon Elasticsearch Service Amazon CloudWatch Amazon Relational Database Service (RDS) Amazon DynamoDB	Web Application Firewall

3. Technical Requirements

3.1. General Requirements

From a technical perspective, the Chinese cloud services researched in this study operate in a very similar way to western service providers. There are however some requirements that are either mandatory or at least make the process of using the services easier.

Many of the services use WeChat or Baidu accounts for authentication, need a Chinese phone number for verification or two-factor authentication tokens or they need an address in China. And most provide only Chinese payment methods or restrictions in place for international payment cards.

When researching the available services, we found that some service providers are not able to provide even pre-sale support without at least some of these verifications being done beforehand.

3.2. Software Mirrors and Package Repositories

There are restrictions when installing software packages to servers inside Mainland China. Most of the Western software mirror servers are not accessible or don't always perform as expected. One of the de-facto standard mirrors is from Taobao. [11] [12] [13]
It provides mostly Node.js packages (NPM) but also mirrors for things like Node.js, Python, PhantomJS, electron, atom-shell, git-for-windows, atom, yarn, etc. Another popular mirror for RubyGems is hosted by Ruby China and sponsored by UpYun. [14]

The service providers offering application hosting usually come with preconfigured Chinese mirrors and may not have an option to directly change the defaults. Or they may not allow the use of package managers directly but instead require the dependencies to be packaged when deploying. [15] [16]

When deploying virtual servers, some service providers are offering server images preconfigured to work well in China, including preconfigured package managers.

---

[11] TaoNPM. (2018) 淘宝 NPM 镜像
https://npm.taobao.org/ [Accessed 05.01.2019]
[12] StackOverflow. (2014) `npm install` goes [~~to~~] dead in China
https://stackoverflow.com/questions/22764407/npm-install-goes-to-dead-in-china [Accessed 06.01.2019]
[13] Screenshot 2019-03-12 10.41.05
[14] Ruby China. (2018) RubyGems 镜像
https://gems.ruby-china.com/ [Accessed 06.01.2019]
[15] Screenshot 2019-03-14 13.56.08
[16] Node.js 运行环境
https://help.aliyun.com/document_detail/58011.html [Accessed 14.03.2019]

3.3. Domain and Hostname Restrictions

The Ministry of Industry and Information Technology requires for ICP applicants to provide proof of ownership, using a domain ownership certificate (域名证书), of a domain name when linking them to a license. Most domain registrars do not provide a valid certificate and companies may have to transfer their domains to local companies or alternatively register new domain names for these purposes. Some TLDs are not available for use in China at all.

The use of subdomains for the purposes of linking to a load balancer or CDN is not possible unless the acquired ICP covers the whole domain name. In practice, setting up a separate domain name for the purposes of deploying cloud infrastructure in China is needed.

4. Legal Requirements

4.1. Internet Content Provider License (备案)

The Internet Content Provider License, commonly known as the ICP license is a permit issued by the Ministry of Industry and Information Technology of the People's Republic of China (中华人民共和国工业和信息化部) for companies publishing online content that is available in China. Including operating a domain name or having a publicly accessible IP address. [17] The regulation states that the license is required when publishing content available online. These licenses are issued at the provincial level and in addition to statewide rules, provincial requirements may also apply. Depending on the situation, foreign companies may want to choose their jurisdiction based on local requirements.

There is also a grace period for companies when they can operate without a valid license while the registration process is in progress. In practice, however, it was found during the research for this study, that the license is required immediately when a company registers as a client for any of the cloud service providers researched in this study.

In cases where the grace period has expired or the ICP is not valid for some other reason, the IP addresses and domain names may be blocked. When using Alibaba Cloud, the blocked resource returns an HTML page with an explanation in English and Chinese:

"*Sorry, the website is unable to access for the moment. According to the filing requirements of China's Ministry of Industry and Information Technology (MIIT), the website is accessible*

[17] 江苏省通信管理局. (2016) 中华人民共和国电信条例
http://www.jsca.gov.cn/xxgk/zcfg/xzfg/201611/t20161114_45619.html [Accessed 06.01.2019]

only if the ICP information is accurate and the ICP license is filed. Please contact the person in charge of the website for assistance." [18]

The Chinese version of this message is more informative and provides possible reasons for the block and instructions. [19]

In order to use a domain name with the cloud services, a valid domain ownership certificate is also required.

Most of the cloud service companies provide tools and services to help with the whole ICP registration process.

There are two types of ICP licenses granted by the ministry:

4.1.1. ICP license

For commercial websites directly selling goods or services online. Represented in the ICP number with the characters "ICP证".

4.1.2. ICP filing

For non-commercial websites which are purely informational and do not directly sell goods or services online. Represented with the characters "ICP备".

4.2. Cyber Security Law of the People's Republic of China (中华人民共和国网络安全法)

The Cyber Security Law was enacted to *"increase cybersecurity and national security, safeguard cyberspace sovereignty and public interest, protect the legitimate rights and interests of citizens, legal persons and other organizations and promote healthy economic and social development."* [20]

It requires network operators, service providers, content providers, and many other companies to store select data within China and allows Chinese authorities to conduct spot-checks on a company's network operations. [21]

The law includes, but is not limited to internet security, but also communication security, computer security, information security, and may in some cases even include industrial automation and control system security. Significantly, the companies affected are not

[18] Screenshot 2019-03-20 14.48.38
[19] Screenshot 2019-03-20 14.48.35
[20] 中国人大网. (2015) 中华人民共和国网络安全法
http://www.npc.gov.cn/npc/xinwen/lfgz/flca/2015-07/06/content_1940614.htm [Accessed 09.01.2019]
[21] The Diplomat. (2017) China's Cybersecurity Law: What You Need to Know
https://thediplomat.com/2017/06/chinas-cybersecurity-law-what-you-need-to-know/ [Accessed 09.01.2019]

limited to content providers or network operators. It may, however, be difficult for companies to identify and verify their category and what obligations and responsibilities they have.

Many of the cloud service providers of this study do provide services to help with the implementation of this law and automation tools for their servers and services, as well as consultation for foreign and domestic companies.

4.3. Telecom Regulation of the People's Republic of China (中华人民共和国电信条例)

The telecom regulation is a permit issued by the Ministry of Industry and Information Technology of the People's Republic of China and Information Technology to companies providing telecom services.

The telecom requirement mostly applies to companies providing telecommunication services, hosting services, etc. In some cases, if the provided service is a service classified in Chapter 2 of the regulation, providing platform as a service (PaaS) or even software as a service (SaaS) may fall under the regulation. [22]

According to Article 10 of the regulation, only local companies with less than 50 percent foreign investment qualify for said permits.

5. Cloud Services

5.1. HTTP Traffic Load Balancing

All of the researched cloud service providers provide HTTP traffic load balancing that is integrated into their systems, similarly to the tools and services of Western service providers.

Cloudflare provides its services through Baidu with their service: SU (百度云加速). This service includes load balancing, WAF, etc., that can be used in conjunction with other cloud service providers in a way that is similar to how Cloudflare services are used internationally.

The key difference is that load balancing between multiple service providers, the ICP restrictions apply to all of the servers and their providers and the load balancer service providers all have different methods for validation and requirements for linking external addresses.

---

[22] 中华人民共和国工业和信息化部. (2016) 中华人民共和国电信条例 http://www.miit.gov.cn/n1146295/n1146557/n1146619/c4860613/content.html [Accessed 09.01.2019]

5.2. Web Application Firewall

With the exception of Elephant Cloud, E Cloud, and Amazon, all of the researched cloud service providers provide an integrated Web Application Firewall as a service.

When distributing infrastructure between multiple providers, a load balancer providing WAF, like Baidu SU can be used even with those service providers that are not providing this service.

5.3. CDN (Distributed Content Delivery Network)

With the exception of Grand Cloud, all of the researched cloud service providers provide a CDN service that can be used to distribute content from various geographical locations.

The functionality of CDNs is critically dependant on the geographical location of the edge servers. The server location availability of the service providers is a key part and should be taken into account when selecting a provider.

The two Western service providers researched in this study both had only two data centers in China.

5.4. Provisioning of Virtual Machines Using APIs

All of the researched cloud service providers have APIs for provisioning and deploying virtual machines and some have the functionality to automatically scale servers based on usage.

5.5. Centralized Logging Options

Most of the researched cloud service providers do not have their own centralized logging services. However, the largest service providers have logging options available where the data can be fed into their service from all the other providers using standard protocols.

Our testing focused on the logging service provided by Alibaba Cloud. The service can be fed raw log data from common logstores, syslog, web tracking, their API and using SDKs provided for multiple platforms (Java, .NET, Python, PHP, C, Android, iOS and Go).[23]

They can also get data from third-party software including NGINX, Docker, Apache, IIS, and MySQL with pre-built bindings and/or configuration files.[24]

[23] Screenshot 2019-03-20 12.06.13
[24] Screenshot 2019-03-20 12.06.13

5.6. Centralized Metric Tools

While most of the researched cloud service providers do have their own metrics and performance tools, some also provide options for aggregating data from multiple providers into one, making the use of a distributed infrastructure possible, where the services are distributed between different service providers.

In the testing done of the Alibaba Cloud logging services, the aggregated data can be used in various ways to get metrics, graphs and other insights beyond just log monitoring. [25]

The output of SQL queries on the data can be plotted in multiple ways. [26]

When running virtual server instances from most service providers, the image is customized to include binaries to output data to the metric tools built into the service. These provide the basic metrics for CPU, memory, networking, and others.

The binary services included in the Alibaba Cloud images also contain other safety features like brute force preventions, intrusion detection, etc. [27] Their binaries communicate with the host machine via either a virtual socket on the virtual machine or a private network. [28] These proprietary monitoring tools cannot be easily audited.

In our testing and our basic cloud architecture reference design, we came to the conclusion that the built-in metric features provide a good starting point when running the infrastructure under one service provider. When aggregating from multiple service providers or using automated deployments, the use of third-party tools offers more flexibility, centralization, a standardized format, and easier auditing.

5.7. Geographically Distributed and Replicated Database Options

Most of the researched cloud service providers have distributed and replicated database options that have their own proprietary modifications to the database software.

The provided database options are compatible with the common protocols, like MongoDB, MySQL, PostgreSQL, etc., and work as drop-in replacements using existing drivers and libraries. Depending on the use case, there might still be additional requirements in case of importing data into these databases. There may also be differences in the replication and geographical distribution options between the providers.

[25] Screenshot 2019-03-20 12.17.43
[26] Screenshot 2019-03-20 12.17.49
[27] Screenshot 2019-03-12 09.58.01
[28] Screenshot 2019-03-12 09.20.21

In addition to the technical differences, services related to the legal requirements, like the Cyber Security Law may not be available or may differ between providers.

These differences should be taken into account when selecting the database service provider.

6. Operational Guidelines

6.1. General

Most of the commonly used proprietary (and services that may even be considered de-facto standards) by European industries are not directly available in Mainland China.

The Chinese service providers tested in this study provide all of the commonly used functionality in similar or in some cases even drop-in replacements.

The use of standardized or open source platforms and software, like Linux based operating systems and databases, help deal with the differences as they usually allow for operating a cloud infrastructure that is vendor-agnostic and can be spread over multiple service providers. This allows for the shared use of many configuration files, tools, and software across countries, as well as an easy migration between Mainland China and Europe.

One of the key requirements when working with these services is a Chinese speaking technical person or systems administrator. Our research found that the top largest service providers have some documentation and user interfaces available in English, but all of them had many critical and important parts available only in Chinese. This includes most of the China-specific technical documentation, contracts, terms of service and other legal documents.

6.2. Server Migration

Most of the service providers offering basic virtual server hosting have dedicated options to import operating system images that can be exported from third-party providers allowing for a way to migrate from and to China. Other providers only allow for their own server images to be run or to create images from existing virtual machines in their infrastructure. [29]

[29] Screenshot 2019-04-04 14.54.44

We've successfully tested a workaround for these restrictions using recovery options provided for booting the virtual machines from an external rescue operating system with networking and full raw disk access to the startup disk.

In this method, copying a raw disk image over the network using common tools (SSH and dd) makes migrating from and to these service providers possible.

Some service providers have created migration tools, like Alibaba Cloud Migration Tool [30] that provides a migration client software that migrates data from a virtual machine to a virtual machine. In our testing, we found the use of these tools to have issues if the migration is done into or from Mainland China and service providers in the West.

Since these migrations result in a virtual machine that is not a bit by bit copy, they may only be used in cases where the operating system, software, and tools can be handled by the migration tool.

6.3. Data Migration

We found some connectivity issues while testing the Chinese service providers that provide tools for directly connecting to the APIs and file protocols, like S3 [31], WebDAV, SFTP, FTP, SCP, rsync, and others to migrate data between their services and third-parties.

The connections between some services in Mainland China and Western service providers occasionally had timeout issues. We determined that they were most likely due to restrictions placed on shared IP addresses or subnets since the connections were cut between the networks themselves.

While our testing shows these issues to be seldom persistent, we tested different approaches to mitigate these issues and found that, as all of the larger service providers also provide services in Western data centers and that their interoperability provided a more stable way for data migration and connections. Deploying proxy machines to these locations and using them while connecting between providers, or at first migrating the data to a Western data center of the same service provider and then to Mainland China, are viable options.

[30] Cloud Migration tool overview
https://www.alibabacloud.com/help/doc-detail/62349.htm [Accessed 04.04.2019]
[31] Migrate data from Amazon S3 to Alibaba Cloud OSS
https://www.alibabacloud.com/help/doc-detail/64919.htm [Accessed 04.04.2019]

There are migration tools with options to directly migrate from services like Azure Blob and Amazon S3 to a service provider in Mainland China like the Data online migration tool from Alibaba Cloud that is currently in beta. [32] [33] [34]

While we found these tools to make the migration to Mainland China easy, they do not provide a direct way to export the data back to these service providers.

The use of official VPN providers is also an option when the service provider in China allows this and/or provides tools for this. This is the easiest option when working with AWS S3 and Alibaba Cloud OSS. [35]

The method and mode of operation with this service is similar to the workaround. The network architecture used with this method accelerates the transmission speed of cross-country data migration.

6.3.1. Amazon S3

From the large Chinese providers, Tencent Cloud provides a storage service called Cloud Object Storage with a web-based management interface, multilingual SDK as well as command-line and graphical tools and a fully S3 compatible with API that can be used in most services and tools as a drop-in replacement allowing to migrate without having to refactor applications for compatibility.

While Alibaba Cloud has an “Object Storage Service” that provides similar functionality and additional features to the storage service provided by Amazon. [36] [37] They don’t, however, provide an S3 compatible API that would work as a drop-in replacement. Migration to their APIs or the provided SDKs is required.

6.3.2. Azure

There are no direct and compatible alternatives to the proprietary APIs of the different storage options of the Azure cloud services. Similar functionality is available but the tools and software using these APIs and SDKs need to be migrated.

As for the data migration between the data storage service, like Azure Blob, there are options like the Data online migration tool described in the main section.

---

[32] Screenshot 2019-03-13 10.12.34
[33] Screenshot 2019-03-13 10.11.40
[34] Screenshot 2019-03-13 10.11.25
[35] Migrate data from AWS S3 to Alibaba Cloud OSS by using IPSec VPN and Express Connect https://www.alibabacloud.com/help/doc-detail/100497.htm [Accessed 04.04.2019]
[36] Screenshot 2019-03-14 10.45.31
[37] Screenshot 2019-03-14 10.28.02

Our testing found the most effective solution for two-way data migration between Azure Blob and service providers in Mainland China to be the manual use of separate tools like AzCopy, running on a virtual machine and transferring the data to separate storage like Alibaba Cloud Object Storage Service.

6.3.3. Google

Similar to Azure, Google Cloud Storage has no direct compatible alternatives, yet similar functionality is available but the tools and software using these APIs and SDKs need to be migrated.

Our testing found the most effective solution for two-way data migration between Google Cloud Storage and service providers in Mainland China to be the manual use of the official gcloud tool, running on a virtual machine and transferring the data to separate storage like Alibaba Cloud Object Storage Service.

7. Cloud Architecture Reference Design

7.1. General

European companies used to running cloud services and tools are restricted by various factors outlined in this study. The inability to use the same proprietary services and tools leads into either using different infrastructures, tools for Europe and in Mainland China, or use tools available in both.

One of the goals of this study was to build a reference design of a cloud architecture based on open standards and open-source technologies that is vendor-independent.

Using a vendor-independent approach allows for easier migrations between service providers, even between Europe and Mainland China as well as building an infrastructure distributed across multiple service providers.

7.2. SaltStack for State Management and Orchestration

This reference design uses the open-source orchestration and configuration management tool SaltStack. It was configured to use a central repository in GitLab to store its configurations (pillars) and instructions (states).

SaltStack is a Python-based open-source configuration management software and remote execution engine. Supporting an “Infrastructure as Code” approach to deployment and cloud management, similar to Puppet, Chef, and Ansible.

SaltStack can provision new virtual machines, set up those servers based on their roles, install all the required dependencies and libraries, deploy software, set up configuration files dynamically and using templates as well as maintain those servers and their states. [38]

In this reference design, the SaltStack states were configured to set up a basic system configuration using SaltStack's "file.managed" state and Jinja templates with custom data for each server role. These states were set up to configure:

- Alibaba Log Service as the endpoint for the syslog daemon.
- Firewalling based on role.
- Basic system configurations: timezone, tmpfs, backups, ssh rules and keys.
- System-wide security settings and brute force protection using commonly available tools (rkhunter, denyhosts, HubbleStack).
- HTTP server as a reverse proxy for the APIs.
- Fetch latest CI/CD build artifacts from GitLab, trigger after a successful pipeline in a specific branches.
- Configure package managers (NPM) Node.js, (pip) Python, (gem) Ruby to use mirrors.
- Set up sample API software to listen to a local socket accessible by the Nginx reverse proxy.

These states are vendor-agnostic and can be deployed for multiple service providers at once with the only requirement of having to set up the "salt-minion" service to connect to a central "salt-master" machine.

Out of the tested providers, all have either the standard cloud-init script capability (including Alibaba Cloud, Tencent Cloud, Baidu Cloud, Huawei Enterprise Cloud) or a way to run custom scripts on new machine startup making setting up the initial "salt-minion" service possible.

7.3. Continuous Integration and Continuous Deployment

The continuous integration and deployment services commonly used in Europe have their infrastructure outside of Mainland China, making using them unreliable or impossible.

There are several open source options for this functionality. For this reference design, we set up a virtual server running the open source project management tool GitLab with multiple distributed GitLab runners, providing the backend for the continuous integration

[38] Orchestrate Runner
https://docs.saltstack.com/en/latest/topics/orchestrate/orchestrate_runner.html#orchestrate-runner
[Accessed 10.04.2019]

and continuous deployment functionalities. We configured the runners to use Docker as the container provider using the official Chinese Docker mirror. [39]

7.4. Provisioning of Servers

This research focused on the feasibility of using the APIs of the service providers for creating autoscaling and automatically provisioned cloud infrastructure, and the possible differences to the service providers used by European industries.

While extending the server orchestration functionality of SaltStack or use them with custom software is fairly simple, out of the tested service providers, Alibaba Cloud has a pre-made driver for SaltStack. [40]

Our testing showed that provisioning servers using these APIs and the SaltStack driver work similarly to other service providers in Europe with some minor exceptions related to some translations. Some of the API responses contain Chinese characters representing details that may be used programmatically and should be taken into account. One of these examples is the use of the Chinese word 位, translating to “bit” in the name of the operating systems (Like “Ubuntu 18.04 64位”).

In this reference design, the main focus was in the use of the Alibaba Cloud driver for SaltStack. A simple cloud-init script was used when manually deploying new servers and Saltify to configure servers with SSH access.

7.5. Database

The clustered and geographically distributed database options provided by most of the tested service providers are similar or work as drop-in replacements for those provided by service providers offering databases as service in Europe.

For the reference design, we used a MongoDB cluster with Alibaba Cloud. [41] [42]

7.6. Data Storage and CDN

Tencent Cloud COS (Cloud Object Storage) was used as the data storage as it provided an API that is fully compatible with S3 and can be used as a drop-in replacement. [43]

---

[39] Docker 中国官方镜像加速
https://www.docker-cn.com/registry-mirror [Accessed 10.04.2019]
[40] Getting Started With Aliyun ECS
https://docs.saltstack.com/en/latest/topics/cloud/aliyun.html [Accessed 10.04.2019]
[41] Screenshot 2019-03-12 09.51.02
[42] Screenshot 2019-03-12 09.51.08
[43] Cloud Object Storage
https://intl.cloud.tencent.com/product/cos [Accessed 10.04.2019]

For large datasets and server images, Alibaba Cloud's OSS was used with the migration tool to migrate data from data centers in Europe.

These storages were also tested for their CDN capabilities. All the commonly used functionalities, like CORS, tokens, etc work similarly.

7.7. Load Balancing, WAF and CDN

The reference design uses a simple HTTP load balancer, set up in Alibaba Cloud with a health check configured on the sample API using the Node.js library provided by Alibaba (also available for Python, Java, Go and PHP). A WAF is linked in front of the balancer with basic security features like protecting against common OWASP threats [44], proxying the traffic to hide the load balancers IP, etc. [45]

7.8. Use Case

The use cases of every company differ and there is no one solution that would directly work with all situations. The reference design here provides a basic cloud architecture that is common in many use cases, where there is a replicated and geographically distributed database and data storage, multiple distributed servers hosting an application or API, CDN and a load balancer in front of the servers and a centralized deployment for these.

[44] Web应用防火墙 - 功能特性
https://www.alibabacloud.com/help/zh/doc-detail/28518.html [Accessed 10.04.2019]
[45] Screenshot 2019-03-12 10.52.50