CALIS-based Cloud Library Services Platform Model

Ying CHEN, Runtong ZHANG, Haoyue LI, Ruixia LI, Yuhong GAO
Institute of Information Systems, Beijing Jiaotong University, Beijing, China
chenying@bjtu.edu, rtzhang@bjtu.edu.cn, lihaoyue8894@126.com,
10120785@bjtu.edu.cn, 10120766@bjtu.edu.cn
doi: 10.4156/aiss.vol3.issue6.25

Abstract

Cloud computing has been applied in many fields at home and abroad as a rapidly developing new information technology. This paper introduces cloud computing which is applied in digital libraries, analyzes current situation and existing problems of the cloud computing in digital library. On this basis, on the combination of cloud computing, SaaS, web2.0, SOA and other technologies, this paper proposes a CALIS-based cloud service strategy and the corresponding cloud library services platform (i.e. Nebula platform) model. The model is suitable for constructing large-scale distributed network of public digital library services. All library resources and service distributed on the Internet can be integrated as a whole, which forms a new type of adaptive control service system supporting interlibrary collaboration and service access, as well sharing resources from different libraries.

Keywords: Cloud Computing, Digital Library, Cloud Services, CALIS

1. Introduction

"Cloud computing" has becomes one of the hottest keywords in the research field. The principle of cloud computing is that the user program processing request is automatically split into countless smaller subroutines, then it searches, calculates and analyzes the data through the "cloud" constituted by network batch server (hundreds of thousands of computers); at last it returns the result to the users.

With the gradually increasing information the library have, it need more information storage space and the information resources reconstruction among libraries, not only cause a large number of redundant information, also increased investment cost. The emergence of “Cloud computing” has brought the construction of the digital library new opportunities with the internet. How to use the thoughts and the principle of "computing clouds" to construct digital library, which lead to the library digital resources can be fully sharing, has become a research hot spot in library field and the development tendency of digital library.

CALIS (China Academic Library & Information System), is one of three public service systems in China's higher education planning. CALIS aims to integrate the state investment, modern library conception, advanced technical means, abundant literature resources in colleges and manpower resources, and construct CALIS as the core of education literature joint security system, and realize information resources sharing, co-constructs to maximize social benefit and economic benefit and to serve higher education in Chinese.

Based on the above problems, and combined with cloud computing, SaaS, web2.0, SOA technology, this paper puts forward the cloud strategy based on CALIS and the corresponding cloud library service platform (Nebula platform) model. This model is suitable for building a large distributed public digital library service network, and it can integrate the resources and services distributed in each library of the internet for a whole, and then form a controllable adaptive new service system, and support transparent interlibrary cooperation and service access, and support interlibrary resources construction and sharing, and have the adaptive extended ability.[1]

2. Cloud Computing in Digital Library

2.1. The Status of Digital Library

After Grid era, Web2.0 experience, digital library slowly entered cloud computing era. The application of cloud computing technology in digital library will bring a real transition, which will bring about profound impacts on digital library services.

After nearly 10 years of construction, in most of the digital libraries, software and hardware infrastructure have a good development, and the literature resources service made a solid pace to the
digital direction, which in certain degree satisfied people's personalized, better educated literature
information needs. However, at present, the construction of the digital library and literature services
also exists some prominent problems[2][3][4]: (1) the literature guarantee ability is insufficient; (2) the
overall benefit is poor, and the use costs of is overall high; (3) repetitive resources construction is
serious, regions, there are lots of similar digital library, and from the region perspective, the repetitive
software and hardware construction situation is more serious; (4) concepts backward, as the network
technology and digital publishing technology development, the common library literature services
should change into primarily "To use give priority to” to the library, but many construct digital library
in the idea of a collection agency.

2.2. The Status of Cloud Computing in Digital Library

With OCLC (on-line Computer Library Center) announcing the launch of "based on cloud
services", the IT application in libraries also entered "cloud" era. For example[5][6][7], some world
famous libraries including university Libraries has begun to adopt "cloud services" to cut there IT
technology departments and increase efficiency.

For example[8] OhioLINK library alliance are using amazon.com’s Webservices to hosting some of
public digital resources, and are testing the limitation of DSpace Collection software and server
management in the clouds. Eastern Kentucky University Library is using Google Docs to collect
replied from website tables, and they take Google Calendar as training and conference Calendar, and
they also use Google Analytics to collect website, the Library catalog and blogging data. Western State
College, located in iGunnison, Colorado, is using Google’s App Engine for their ELibrary, and they
also put two Microsoft Access database circulation, and move government publication management to
this service[2][9] At present, the cloud services provided by OCLC are more focus in the dynamic
data search, so, it belongs to cloud services based on dynamic data, which largely marks a more
substantial applications including cloud services in library field. it is visible that the idea of cloud
computing are being accepted and approved by the library domain and it has important practical
significance and development tendency.

The practical application of cloud computing in domestic libraries is still in the theoretical
problem seminar about digital libraries” gave high attention on cloud computing. in the report of "cloud
computing technology in the application of CALIS project” from Wenqing Wang, it gave a detailed
explanation of building process of CALIS cloud platform model and the key technologies used, and
also the significance of cloud computing, that is the resource development tendency of the socialization,
service station and specialization. Xiaojing Hu's "cloud computing @ library" report, briefly introduces
library examples involved in cloud computing at home and abroad. The report of "environmental
scanning about computing clouds in libraries" from Liang Zhao analyzed the internal and external
environment of cloud computing in libraries, which appealed to the attention of cloud computing in
library world.[10]

3. Application Challenges of Cloud Computing in Digital Library

Although cloud computing can lead to strong, flexible and low cost collaboration and innovation
platform on library's infrastructure construction and service, but there are some obstacles about its
application in library field needed to be overcome.[11][12][13]
(1) The lack of standards and data
At present, cloud computing has been provided by Google, Amazon, Microsoft and some domestic
companies, and IBM has formed WuXi cloud computing center. But overall, the lack of uniform
standards, may lead to re-write applications when the cloud service providers in change, and the
disunity of the technology maintaining the normal operation of the cloud computing could cause
confusion when users choose a specific plan, which greatly limits the popularization and promotion of
cloud computing.
(2) Interoperability of data and application
It is very important that Data and application provides standard access interface. Organizations often
want to flexibly create new data and application, and they could interoperate, regardless of
infrastructure providers (whether public cloud, private cloud, traditional IT environment, or some
combination of the foregoing). Cloud vendors need support interoperability standards, which help the
organization to make any cloud providers ability into its solution.
CALIS-based Cloud Library Services Platform Model
Ying CHEN, Runtong ZHANG, Haoyue LI, Ruixia LI, Yuhong GAO
Advances in Information Sciences and Service Sciences, Volume3, Number6, July 2011

(3) Security
Although cloud computing service providers have professional technical teams and management teams to ensure service security, but there are many potential security risks in cloud computing services.
① Priority access to risk: after enterprise data delivered to clouds service providers, who has priority access to data, it cannot rule out the possibility of data leak. ② Management rights risk: although the enterprise users pass data to cloud computing services to manage, but the enterprise itself is responsible for data security and integration issues, etc. ③ Data premises risk: when enterprise customers using cloud computing services, they do not know which server places their own data, and they even don’t understand which countries this server is placed in. In cloud computing services platform, large of enterprises are in sharing environment, even if adopting data encryption method, it cannot ensure foolproof, and it will reduce the efficiency in using data if encrypted. ④ Data recovery risk: enterprise users not only need to know whether providers have the data recovery ability, but also must know how long service providers take to finish data recovery. ⑤ Long-term development risk, if the enterprise users selected a computing clouds service providers, when computing clouds service providers bankruptcy or purchased by others, enterprise customer service will be interrupted or become unstable.

(4) Bandwidth Shortage
Because cloud computing services are based on the Internet, all of the application data will be transmission based on the telecommunication network, which cause increasing transmission quantity and raise very tall requirement for bandwidth. In addition, data long-distance transmission causing delay worries, and adding the stability problem of internet, cover certain shadows on the development of cloud computing services. But with the development of network technology, the gradually improved bandwidth, and the application of technical means such as Users visit nearby calculated, distributed storage, load balance, re-direction of network request, network access optimization, SSL application accelerated, it can largely protect users timely, stability and safely to use concerning services of cloud computing.

(5) Cloud Software
Cloud software is a kind of SaaS online service mode, and enterprise can rent service to get information service responsible by the third-party professional company. SaaS has become a new trend in the development of software industry, and it will even be the commanding heights in future software industry. But because the Cloud Software has high requirements of reliability and expansibility of service quality, software technology designed to support the SaaS services must have multiple SaaS provider collaborative solution. Meanwhile, related standards of the SaaS service, the integration standards, Service Level Agreement, the migration among different providers and data integration all directly influence the user’s confidence index.

(6) The Concept of Library
Along with the entity library atrophying, the identity and belonging of librarian job will decline. Whether the librarians can adopt a positive attitude to meet new technology will deeply affect cloud computing related technology and service in library application.

4. CALIS-based Cloud Library Services Platform Model

Based on the above problems, and combined with cloud computing, SaaS, web2.0, SOA technology, this paper puts forward the cloud strategy based on CALIS and the corresponding cloud library service platform (Nebula platform) model. This model is suitable for building a large distributed public digital library service network, and it can integrate the resources and services distributed in each library of the internet for a whole, and then form a controllable adaptive new service system, and support transparent interlibrary cooperation and service access, and support interlibrary resources construction and sharing, and have the adaptive extended ability.

4.1. Overall Platform Model

Based on construction requirements of three stages, and combined with cloud computing, SaaS, web2.0, SOA technology, we put forward the framework of CALIS cloud library service platform (called Nebula platform), it includes the following four aspects:
(1) CALIS digital library public service platform facing the library, which is used to construct CALIS cloud service center;
(2) CALIS digital library SaaS service platform facing the library, which is used to provide SaaS service;
(3) digital library local service platform facing the library: including local application foundation platform and local application system
(4) CALIS cloud federal service platform facing the library, which aim to integrate local library service, CALLS public service and the third-party service.

This platform model is shown in figure 4.

4.1.1. Infrastructure Services and Platform Services

These two types of services not only can be provided by third party service providers, but also can be provided by nebula computing centers of by their own. CALIS will build one or more than one nebula service center and their own infrastructure platform SaaS. At the same time, CALIS will appropriately use some of SaaS services provided by third parties.

4.1.2. Public service platform of CALIS digital library

The public service platform (known as Nebula Central Platform) constitutes by a set of software. It can be used in the nebula and can provide basic services including unified certification services, accounting services, joint resources search services, data services, knowledge services and digital object storage and download services, metadata unified catalog services, document combine subscription services, global resource scheduling services and so on. Services above are for the libraries directly or are available to the library by a group of Open API. CALIS public service platform can be deployed in the CALIS national center and also can be deployed in some provincial centers for the establishment of CALIS digital library nebula service center.

4.1.3. SaaS Platform of CALIS Digital Library

The SaaS platform provides the final application services directly to library. Such services include interlibrary loan SaaS services, reference and consultancy SaaS services, special database processing SaaS services, teach participant management SaaS services, degree thesis SaaS services and so on.
Every library can rent parts or all services on demand. SaaS platform of CALIS Digital Library can be deployed in CALIS National Centers and various provincial centers.

4.1.4. CALIS local service platform

CALIS local service platform consists of CALIS local application basic platform, CALIS local applications system and third-party local application system. The first and the second categories systems can provide localized nebula solutions for libraries. CALIS local application basic platform (called Nebula Main Server) with a unified services registration and management, unified monitoring and log management, local unified authentication or authorization, single sign-on, public services release, external services subscription and other core functions. In addition, the platform also provides state management, load management and other real-time services and provides simplify and automate way of deployment and management to ensure availability and scalability of services. The basic platform is responsible for interacting with CALIS digital library public service platform and CALIS nebula federation services platform to achieve the automatic registration and discovery of distributed resources and services. CALIS local application basic platform is the key to local nebula solution and play a core role of linking the preceding and the following, integrating in the library, connecting between libraries and so on.

CALIS local application system (collectively known as Nebula Application Server) is used to provide specific business functions for the library. Such applications include the local version of the portal system (Portal), interlibrary loan systems, document delivery tools, reference and consultancy system, special database processing systems, resource harvesting system, resource distribution system, resource storage service system, a unified retrieval system, resource scheduling system and so on. CALS local application system is embedded standardized interfaces and can seamlessly integrate with CALIS local basic platform.

Third-party local applications become an integral part of local digital library through supporting for the standardized unified authentication, monitoring, logging and other interfaces and seamless integrate with CALIS local basic platform.

4.1.5. Nebula Federation Services Platform

Federal nebula services platform (known as the Nebula Federation Platform, NFP) is not only for realizing the integration of services across nebula, but also is the key to forming mixed nebula for private nebula and public nebula interoperability. It provides an abstraction of the application for external services by Open API mean, and also provides hosting services for the Open API of CALS centers.

This platform integrates the different library local service platforms, CALIS public service platforms and third party public service platforms to provide a unified service interface to the library, for further mash-up of library, CALIS, and ISV and reduce the difficulty of service integration, which can form a high integrated digital library system by unify the interface and operation for the reader.

4.2. Service Architecture

4.2.1. SOA Architecture

Service Oriented Architecture (SOA) is effective modeling ideas and software construction methods of the current web application makes systems’ reusability, reliability, scalability and development speed higher by modular, integrated, dynamic assembly.

The OSGi specifications[14] define a generic, standard, component-oriented computing environment for network service. Software components can dynamically discover and use other databases or applications. OSGi defines component standards and service standards. Services are interactive interface between components. Components can pluggable, change behavior dynamically and have greater multiplexing capability. Some middleware and applications such as IBM, SUM, BEA, open source community are beginning to build with OSGi standard.

Architecture of CALIS digital library system nebula services platform based on SOA specifications makes uniform package for the each service using OSGi standard to provide integrated services publishing for service providers and to provide a unified communication between service consumers and providers.
4.2.2. Service Model

CALIS Digital Library nebula services platform is composed by a series of service functions. These functions are divided into five levels: basic level, development level, core services and common service layer, application layer, portal layer. As shown in Figure 5.

**Figure 5.** CALIS Nebula Services Platform Architecture

CALIS digital library nebula services platform makes services of all levels be built on the OSGi framework. These services are packaged as OSGi service for other components calling.

Applications of new development are developed based on Nebula development platform, which can call the existing core or general services. The existing application integrates into the entire platform by the interface provided.

Some applications (such as interlibrary loan, reference and consultancy, etc.) need to be modified and access unified internal data based on Multi-Tennant to be able to provide SaaS services.

Portal systems use portlet as a unified applications front-end component, call other services, provide a united user login or logout, provide customer self-service centers and present the various business functions in a form of unified interface and operation mode to the user.

Services above may constitute a different CALIS application platform or application system through combination.

Applications’ operating condition monitoring, service switching and service distribution implement by the core service of nebula service platform in order to achieve the system’s autonomy. Core Services also provides security management, transaction management, fault tolerance, persistence, cache management, load management, service scheduling, large data...
distributed storage and other services. These core services ensure the entire nebula services platform can provide unified and large-scale services as “nebula”.

4.3 Application methods

CALIS Digital Library nebula platform can provide a standardized, low-cost, adaptive, and scalable digital library unified services and integrated solutions. It provides flexible deployment and application methods for CALIS and its member libraries, which not only can meet the need of CALLS building the public nebula service center, but also can meet needs of building a private nebula services for the library, and achieve the integration of these two types of nebula services

4.3.1. Local Nebula Platform (Private Nebula)

Library can easily build their own digital libraries nebula platform and support integration with third-party applications by using CALIS local platform (including CALIS local basic platform Main Server and applications AppServer). This local nebula platform can be installed on the local HaaS or PaaS and also can be partly installed in a remote public PaaS or HaaS platform. These library nebula platforms can access and integrate services provided by CALIS public nebula.

Part of the library nebula platforms can provide services for open part. These services can be accessed and integrated by CALIS nebula service centers, CALIS nebula federation service platform and other libraries in order to realize sharing and integration of various resources and services between libraries.

The whole library nebula platform can be deployed in more than one cluster server (as shown in Figure 6). That means the Main server, Nebula App Server (referred to as the N-Server) can be deployed in one or more servers; The server can flexibly change with change of visit and demand on storage capacity. At the same time, the external service and the unity authentication service of the local platform are all unity provided by the Main Server. In addition, third-party applications (referred to as the X-Server) can be integrated with Main Server and N-Server.

Figure 6. Local Nebula Platform of Library

4.3.2. CALIS Public Nebula

Each center of CALIS can effectively build digital library public nebula of university using Paas and SaaS CALIS public service platform. Public nebula both could build in basis environment of nebula computing centers CALIS their own (HaaS or PaaS) and can put part of function on remote public
infrastructure HaaS or PaaS foundation platform provided by other nebula computing centers, and its
deployment manner similar to Figure 3.
CALIS public nebula provides their services to every library by the way of SaaS or PaaS for using
their own.
The deployment of CALIS nebula federal service platform in CALIS National nebula center can
collect Open API providing by all CALIS public nebula, serve outside in a uniform and transparent
way and provide integrated method of all public and private nebula for deep integration of libraries, so
form digital libraries mixed nebula (hybrid or federated nebula).

5. Conclusion

This paper introduces cloud computing which is applied in digital libraries, analyzes current
situation and existing problems of the cloud computing in digital library . On this basis, on the
combination of cloud computing, SaaS, web2.0, SOA and other technologies, this paper
proposes a CALIS-based cloud service strategy and the corresponding cloud library services
platform (i.e. Nebula platform) model. The model is suitable for constructing large-scale
distributed network of public digital library services. All library resources and service
distributed on the Internet can be integrated as a whole, which forms a new type of adaptive
control service system supporting interlibrary collaboration and service access, as well sharing
resources from different libraries.

But in practice, the cloud computing is facing the large number of technical problems and
engineering problems. Although the proposal of IPv6 has improved the shortcoming of IPv4 and
remedy a lot of security leakage. However, wireless networking around the city in anytime and
anywhere will generate more information security problem than before[15]. Cloud computing is
associated with a range of severe and complex privacy issues. The main issues of cloud security
are all related to data security which is the basic issue of cloud security. We need a measure to
prevent our data from being obtained or damaged by some people who harbor vicious intentions.
Therefore, it is necessary to encrypt data and make that the data obtained illegally can not be
deciphered[16].

6. Acknowledgement

This work is supported by Beijing Natural Science Foundation(On the Key Technologies by
Seamless Connection between Wireless Sensor Networks and Internet Based on the Applications
to the Internet of Things).

7. References

[2] Lingchao Zhang, “Primary research on the construction of the digital library Based on cloud
2010.
under cloud computing environment”, Journal of the Library Science of Sichuan, vol. 17,
[8] Nai Zhang, “Information resource sharing model under cloud computing environment”,