AstroCloud, a Cyber-Infrastructure for Astronomy Research: Cloud Computing Environments

Changhua Li,1 Jiawei Wang,1 Chenzhou Cui,1 Boliang He,1 Dongwei Fan,1 Yuecheng Yang,2 Junyi Chen,2 Hailong Zhang,3 Ce Yu,4 Jian Xiao,4 Chuanjun Wang,2 Zhihuang Cao,1 Yufeng Fan,2 Zhi Hong,4 Shanshan Li,1 Lining Mi,1 Wanghui Wan,1,5 Jianguo Wang,2 and Shucheng Yin4

1National Astronomical Observatories, Chinese Academy of Sciences (CAS), 20A Datun Road, Beijing 100012, China
2Yunnan Astronomical Observatory, CAS, P.O.Box110, Kunming 650011, China
3Xinjiang Astronomical Observatory, CAS, 150 Science 1-Street, Urumqi, Xinjiang 830011, China
4Tianjin University, 92 Weijin Road, Tianjin 300072, China
5Central China Normal University, 152 Luoyu Road, Wuhan 430079, China

Abstract. AstroCloud is a cyber-Infrastructure for Astronomy Research initiated by Chinese Virtual Observatory (China-VO) under funding support from NDRC (National Development and Reform commission) and CAS (Chinese Academy of Sciences). Based on CloudStack, an open source software, we set up the cloud computing environment for AstroCloud Project. It consists of five distributed nodes across the mainland of China. Users can use and analysis data in this cloud computing environment. Based on GlusterFS, we built a scalable cloud storage system. Each user has a private space, which can be shared among different virtual machines and desktop systems. With this environments, astronomer can access to astronomical data collected by different telescopes and data centers easily, and data producers can archive their datasets safely.

1. Introduction

Recent years, with more the large telescope start to use, the astronomy research has stepped the big data era. but the network bandwidth, especially the Internet, is still the 1000/100Mbps. it is difficult to download the big data to astronomer’s machine. Based on this reason, if astronomers have enough private and safe computing machines in each data center, it will avoid a lot of data downloaded to the local machine. This flow will be more effective. Fig.1 shows this two different workflow, the left is the traditional flow, the right is the new flow based on Astrocloud.

In order to implement this workflow, cloud computing technology give us a perfect solution. Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction(Mell & Grance 2011). Ac-
According to the provisioned resource type, Cloud computing includes three service models, Software as a Service (SaaS), Platform as a Service (PaaS), Infrastructure as a Service (IaaS). At present, IaaS is the main service model of AstroCloud computing environments.

2. Architecture

The astronomy data centers are geographical distribution in China. AstroCloud computing environments will integrate all this distributive computing, network and storage resources, and give astronomer a centralized resource view and operation platform. Each data center will become a cloud computing node. Figure 2 shows the architecture of one cloud node.

This architecture contains three layers. The top layer is the user layer, it is the user interface, including browsers and software clients for many kinds of terminal like desktops, mobiles. The bottom layer is the physical layer, it is the physical resource. The middle layer is the virtualization layer, it is the key content of Astrocloud platform. The Cloudstack is an open source cloud computing management middleware software. Based on this, we implement the cloud virtual machine creation and access, cloud monitoring, cloud storage, cloud authentication and single sign on. In this cloud computing environment, we integrate many IT technology like VLAN, VPC, Firewall, Port forward, KVM, NFS, Samba, etc.

3. Main Function Components

In AstroCloud platform, it includes cloud machine creation and management, cloud storage system of myospace, cloud monitor and cloud usage log and analysis components. Figure 3 shows the overview of function components and relative technology.
Function 1: Only spend five minutes, You can get a virtual machine (VM) with most of astronomy software. By Internet, you can access it by SSH or remote desktop at any location. According to your requirement, you can choose the suitable hardware configuration of VM.
Function 2: Provider a private, safe and large capacity storage space, which share and auto-mount to your different VMs.
Function 3: Cloud monitoring system, realtime status collection and notification.
Function 4: Detailed usage log and analysis, including the usage of resource, user billing.

4. MyVOSpace in AstroCloud

MyVOSpace is the private, safe user storage space, which is one of the main contents of AstroCloud computing environment. In AstroCloud, MyVOSpace can be shared in different VMs and remote desktop of one user. Fig.4 shows the framework of MyVOSpace, cloud storage system and the relation with other components, technology.

![Figure 4. The framework of MyVOSpace of Cloud Computing Environments](image)

GlusterFS is an unified, poly-protocol, scale-out filesystem serving many PBs of data. Based on this, we built a large capacity and easy extended, stable storage system.

Acknowledgments. This paper is funded by National Natural Science Foundation of China (U1231108), Ministry of Science and Technology of China (2012FY120500), Chinese Academy of Sciences (XXH12503-05-05). Data resources are supported by Chinese Astronomical Data Center.

References