on an open platform OpenVz. And on another – system virtualization based on KVM (Kernel-based Virtual Machine) because it is more powerful (Intel Xeon 22.53-2.80 GHz and 24 GB of RAM, Adaptec RAID controller, 3.6 Tb disk space) and support hardware virtualization. KVM unlike Openvz allows one to run virtual servers running different operating systems, not only Linux. Last year we tried to use a virtualization system based on Linux container (LXC), but because it was unstable, we were forced to go to the KVM virtualization.

We have two file storage servers in the data center. One of which is used for backup storage project "Radioastron" (20 Tb disk space) and second storage data of radio astronomy observations PRAO and of its processing (48 Tb). All servers are connected to a local network with two or more Ethernet cards to increase the speed of information exchange between servers and a local area network, and also for increase redundancy. This is possible thanks to the ability of our Gigabit switch D-link DGS-3420 to combine multiple twisted-pair or fiber Ethernet links into one fault-tolerant and load balanced logical link. In Linux systems this feature is called bonding, in switches its commonly referred to as a port channel or link aggregation. Either way, its using the LACP (802.1ad) protocol for that.

PROSPECTS OF CLOUDY TECHNOLOGIES IN THE SOLUTION OF THE TASK OF THE ANALYSIS OF LARGE VOLUMES OF THE DATA OBTAINED IN ASTRONOMICAL SUPERVISION

E.A. Isaev [1,2], V.V. Kornilov [2,3], V.A. Samodurov [1,2], A.A. Tarasova [4]

[1] Pushchino Radio Astronomy Observatory ASC LPI, Pushchino, Russia, *is@itaec.ru*
[2] National research university Higher school of economics, Moscow, Russia
[3] Institute of Mathematical Problems of Biology RAS, Pushchino, Russia
[4] FIAN, Moscow, Russia

ABSTRACT. In modern astronomy the problem of the big data obtained during scientific experiments is very actual. One of the perspective approaches to the solution of the problem of processing of superlarge volumes of experimental data in real time is the use of the technology of cloud computing which assumes ensuring remote dynamic access of users to services, computing resources and appendices on the Internet.

Key words: Big Data, Cloud Computing

The success in the development of modern computer technologies and electronics led to possibility of creation of scientific experimental installations of the new generation which characteristic is the large volumes of the data obtained during scientific experiments. Especially it is actual for astronomy and astrophysics. Devices for astronomical observing allow to obtain data with more and more high resolution, the observations of astronomical objects are conducted not only in visible light, and in all range of an electromagnetic range, thus the only observation which can last from several seconds to several minutes, gives from several megabytes to several gigabytes of information. Such astronomical projects as "Radioastron", "Millimetron", "Kvazar-KVO", "RATAN-600", PAN-STARRS, SDSS, LOFAR, ASKAP, SKA and others are capable to generate tens and hundreds gigabytes of the supervision given for every second. So, for example, the data archive of the Telescope of Hubble for 15 years of its work is about 25 Tb [1]; the Large Survey Telescope (LSST) with a diameter of mirror of 8.4 meters and a 3 Gigapixels matrix will make 30 Tb of data only for one night, and the full volume of supervisory archive is estimated at 200 Petabyte [2].

The data obtained during experiments should be able to be stored, processed, transferred and analyzed to receive the new knowledge from these data. Enormous volumes of these data and high speed of their increasing do these tasks rather difficult for the effective decision.

The term Big data usually use for data sets with sizes beyond the ability of commonly used software tools to capture, curate, manage, and process the data within a tolerable elapsed time. In 2012, Gartner did the definition of Big data as follows: "Big data is high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization." [3]. Additionally, the new V: "Veracity" and "Visualization" now are often added to describe Big data [4].

The ability to extract knowledge from such huge data with the help of special analytical methods and modern computer technologies predetermines the success of modern science.

One of the perspective approaches to the solution of the problem of processing of superlarge volumes of experimental data in real time is the use of the technology of cloud computing which assumes ensuring remote dynamic access of users to services, computing resources and appendices on the Internet. Cloud computing is a type of computing that relies on sharing computing resources. To do this, cloud computing uses networks of large groups of servers typically running low-cost consumer PC technology with specialized connections to spread data-processing chores across them. This shared IT infrastructure contains large pools of systems that are linked together [5].

So, for processing of big data the huge computing capacities are required, and the cloudy computing is capable to present almost unlimited processor resources for processing of these volumes of data, and on very effective and flexible scheme – the necessary computing resources can be quickly provided with the minimum operational expenses. The Cloud Computing model is a perfect match for big data since cloud computing provides unlimited resources on demand. Thus, the role of cloudy technologies consists in the possibility of the effective analysis of large volumes of data, by their selection from various places of storage and use for their processing of necessary computing power.

The problem of big data in modern astronomy compels to look for non-standard ways for work with them. Transfer, storage and, the most important – the analysis of these data, in an ideal in real time, demands fantastic computer capacities. Alternative to creation of expensive supercomputer complexes of each experimental installation are the cloudy technologies which can successfully solve problems of processing of big data. Cloud computing provides the tools and technologies to build data/compute intensive parallel applications with much more affordable prices compared to traditional parallel computing techniques. But for successful implementation of this concept it is necessary to solve the number of problems, including the shortage of the qualified specialists for cloudy technologies and to the analysis of big data.

References

1. *http://hubblesite.org/*
2. *http://www.lsst.org/*
3. Laney D.: *http://www.gartner.com/resId=2057415*
4. Mayer-Schonberger V., Cukier K. "Big Data: A Revolution That Will Transform How We Live, Work, and Think". Eamon Dolan/Houghton Mifflin Harcourt, 2013.
5. Erl T., Puttini R., Mahmood Z. "Cloud Computing: Concepts, Technology & Architecture". Prentice Hall, 2013.