SECTION 4. Computer science, computer engineering and automation.

Olga Viktorovna Ustimenko
PhD student, junior researcher, software engineer, assistant
Joint Institute for Nuclear Research, Dubna, Russia
ustimenko@jinr.ru

DEVELOPMENT PROSPECTS OF DISTRIBUTED STORAGE OF BIG DATA

Abstract: *The report presents an analysis of Big Data storage solutions in different directions. The purpose of this paper is to introduce the technology of Big Data storage, prospects of storage technologies, for example, the software DIRAC. The DIRAC is a software framework for distributed computing.*

The report considers popular storage technologies and lists their limitations. The main problems are the storage of large data, the lack of quality in the processing, scalability, the lack of rapid availability, the lack of implementation of intelligent data retrieval.

The device features, functionality and operation methods for high -energy physics experiments DIRAC software are presented in the main part of the paper. Experimental computing tasks demand a wide range of requirements in terms of CPU usage, data access or memory consumption and unstable profile of resource use for a certain period. The DIRAC Data Management System (DMS), together with the DIRAC Storage Management System (SMS) provides the necessary functionality to execute and control all the activities related with data.

Key words: *distributed storage systems, Big Data, software framework, Pilot Jobs.*

Citation: Ustimenko OV (2014) DEVELOPMENT PROSPECTS OF DISTRIBUTED STORAGE OF BIG DATA. ISJ Theoretical & Applied Science 9 (17): 23-27. doi: http://dx.doi.org/10.15863/TAS.2014.09.17.3

Introduction

The very notion of Big Data has appeared relatively recently and has not been clearly defined yet. Big Data is huge volume (hundreds of petabytes) of information that can be useful to solve scientific and business problems. Standard classic storage technology, data processing and analysis are not suitable for the use of such information. However, in our work specific information such as text, images, graphics, video and audio readings is actively used [2].

Processing experiment BESIII [9] needs improving management and storage of data within the project DIRAC [10]. Commercial organizations identify the need to develop new systems of customer specific task. Quantity and accessibility are the main aspects of the commercial software. However, data sets exceed the volume significantly in the project. Maintaining the quality in processing, scalability, fast accessibility is needed for experiments. Software should be able to change the information and to implement intelligent search data.

An important aspect is adaptation of the existing services of data access and management, working primarily with local file systems and partly with remote data sets, which use cloud storage and a very large volume of data with distributed computing environments. Grid computing was developed to provide users with uniform access to large-scale distributed resources [5].

Analysis of existing distributed storage systems Big Data

Despite the fact that there is a sufficient variety of distributed storage systems, nowadays the requirements for them are not implemented completely. We face the following challenges [1]:

- ease of deployment, operational changes of data;
- scaling control system;
- simplicity, flexibility and depth of configuration;

- quick access to a huge volume of data sets and geographically distributed resources;
- availability and depth of monitoring;
- reliability of the network infrastructure in a heterogeneous environment [4].

The most promising technologies for storage and data management are all the object cloud data storage (Amazon S3, MS DataMarket, OpenStack Swift, DropBox), specialized storage system adapted for storing scientific data (RasDaMan and SciDB), hybrid storage systems (ActiveStorage), distributed system with streaming treatment (Twitter Storm, Yahoo S7, IBM Streams, Hadoop).

Specificity of problems with such products is that these aspects are not fully analyzed and considered:

- Variety of data formats and their increase.
- Heterogeneity of computing resources in the cloud composition and use of previously established grid computing environments.
- Uniqueness and diversity of application services for science and education (diversification of subject areas).
- Architectural binding application services to the computing infrastructure (infrastructure optimization for the developer).
- User participation in the creation and development of new services and composite applications based on them (self-organization of clouds).
- Virtual professional community (stable feedback).
- Support of interactive services in real time (services of access to the equipment, imaging, etc.).
- The lack of universal and high-level query language that takes into account the specifics of storage (e.g., in the cloud) and data models (primarily multi-dimensional arrays).
- The lack of scalable, distributed, open platform for cloud storage of scientific data.
- The lack of a framework, which provides distributed parallel processing of streams of scientific data.

Thus, each of the above systems has a priority of tasks being processed.

Development of distributed storage systems of Big Data

New trends of development of distributed storage systems include the systems that will solve the problems of creating composite applications, intelligent search support and application of services, dynamic management of service performance, flexible integration with real-time systems.

Formalized knowledge is necessary for the formation of intelligent storage systems for a very large volume of data. The development of intelligent technologies annotates, searches, and application services refer to this notion. Management of composite application performance should be made in the executive environment. It is important to be able to plan the execution, using a general-purpose communication network. With all the features above the system should work in real time (the existence of decision-making centers).

An example of a well-organized distributed storage architecture for a large volume of data is the monitoring system of file transfer service (FTS) developed with the support of the JINR, as well as monitoring system of Tier3-centers of data analysis, the global system of monitoring of data transfer in the infrastructure WLCG [6].

Different grid systems and services are used for organizing data storage. In the Grid environment the following systems are the most commonly used: Castor, dCache, DPM. For system interoperability a Storage Resource Manager (SRM) service has been designed.

dCache system is aimed at storing large amounts of experimental data. In order to access files dCache uses proprietary protocols (DCAP), gridFTP or any other file access protocol.

Xrootd provides a user-friendly and fast access to data of any kind. The data must be organized in a hierarchical file system, such as catalog- based namespaces.

Big Data in the project DIRAC

DIRAC (Distributed Infrastructure with Remote Agent Control) INTERWARE is a software framework for distributed computing providing a complete solution to one (or more) user community requiring access to distributed resources (Figure 1). The DIRAC builds a layer between the users and the resources offering a common interface to a number of heterogeneous providers, integrating them in a seamless manner, providing interoperability, at the same time as an optimized, transparent and reliable usage of the resources [10].

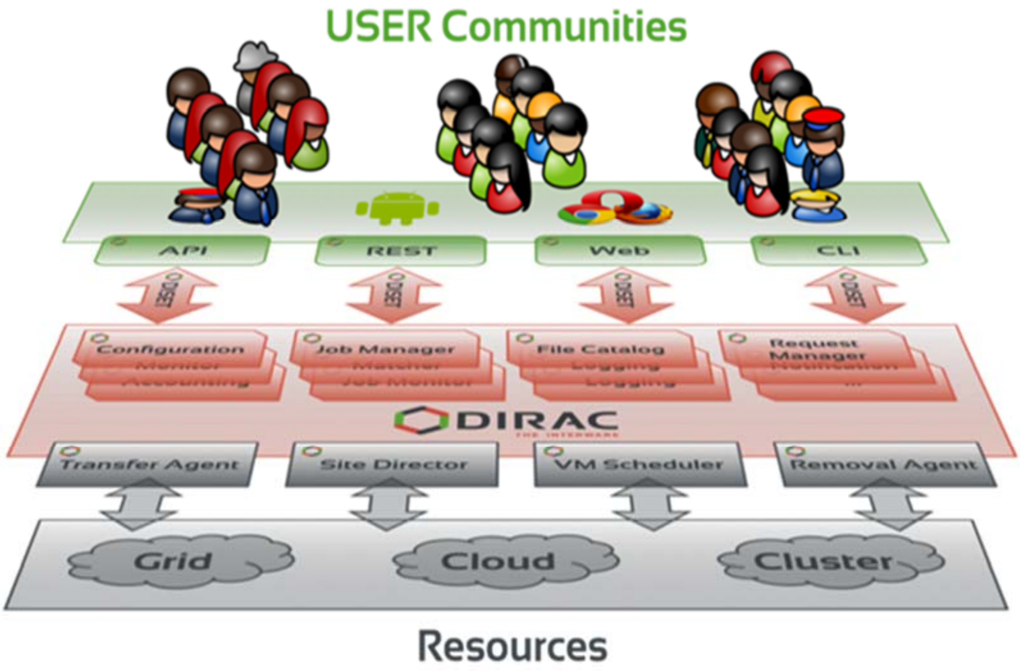

Figure 1 - Virtual structure of the DIRAC.

The Workload Management System (WMS) with Pilot Jobs introduced by the DIRAC project is now widely used in various grid infrastructures. This concept allows one to aggregate computing resources of different source and nature, such as computational grids, clouds or clusters in a single system, transparently for the end users.

The DIRAC implements WMS able to manage simultaneously computing tasks, where not all tasks are of the same nature [8]. They can range from very high priority short payloads to check the current situation at the different computing resources available to the system, to chaotic end-user payloads. And they also include activities like Monte Carlo simulations and real-time detector data processing.

In order to optimize the access to the computing resources, the DIRAC, early in its development phase, chooses a solution with a single central server holding the complete queue of pending payloads, late binding of resources to payloads and the usage of the Pilot Job paradigm. For the implementation of this design choice three key components are necessary: the Pilot Jobs, the TaskQueues, and the TaskQueue Directors.

They are described in detail in the following:

1. Pilot Jobs are nothing more than empty resource reservation containers that are sent to the available computing resources with the final aim of executing the most appropriate pending payload in the central WMS queue.
2. As soon as payloads are introduced into the DIRAC WMS system and after a consistency check and a proper resolution of possible target computing resources in case there are Input Data requirements new tasks are organized into TaskQueues. TaskQueues are nothing more than sorted groups of payloads waiting for execution with identical requirements to the possible Pilot Job requesting a match.
3. TaskQueue Director Agents is the way chosen by the DIRAC to populate the available computing resources, the worker nodes, with Pilot Jobs. This logic is the following: Query the central WMS server to get a list of TaskQueues.

With all these capabilities the DIRAC allows an overall optimization of the resource usage for the complete user community, including a large variety of different payloads and use patterns. The resulting system has shown very nice scalability [7]. In order to submit a job to the DIRAC one needs to describe task requirements. There are several requirements that can be specified, for instance, the name of the executable program, the location of the output of this program, etc. the DIRAC Workload Management System, with the DIRAC system memory management (SMS), provides the necessary functionality for the implementation and monitoring of all activities related to your data.

Conclusion

One of the main methods for solving the problem of Big Data is the use of «cloud» computing by which users are able to remotely access information arrays and the use of distributed computing resources to handle them. There is a possibility to increase productivity through high concurrency and multiplicity of access points in cloud computing systems.

At the same time computer systems for the storage and processing of data sets can be placed in specialized centers, data processing and storage. These centers are used not only for remote storage and backup of large amount of scientific data, but also for remote access to applications that provide analysis of these data.

The cloud paradigm is wide-spread in the IT world, and the DIRAC is ready to provide seamless integration of these resources to its existing grid and cluster users today. High Energy Physics is in the position to take the leading role in the integration, however, the scalability must be the key factor in the process of suitability identification.

References

1. Barseghyan AV, Kupriyanov MS, Stepanenko VV, Cold II (2007) Data analysis technology. DataMining, VisualMining, TextMining, OLAP/AA, St.Petersburg: BHV-Petersburg, pp.384.
2. Bashmakov AI, Bashmakov IA (2005) Intelligent Information Technologies. Moscow: Publishing House of the MSTU. NE Bauman, pp.304.
3. Casajus A, Graciani R, Paterson S, Tsaregorodtsev A (2010) DIRAC Pilot Framework and the DIRAC Workload Management System. 17th International Conference on Computing in High Energy and Nuclear Physics (CHEP09). IOP Publishing Ltd. Journal of Physics: Conference Series 219.
4. Demchev AP, Ilyin VA, Kryukov AP, etc. (2008) Problems of open networks in distributed processing systems, very large volumes of data. Institute of Nuclear Physics. DV Skobel'tsyn Moscow State University. MV Lomonosov, RRC "Kurchatov Institute", Moscow.

5. Fifield T, Carmona1 A, Casajs A, Graciani R (2010) Integration of cloud, grid and local cluster resources with DIRAC. The University of Melbourne, Australia.
6. Korenkov VV, Uzhinskiy AV (2009) Monitoring system data service (FTS) project EGEE/WLCG. Numerical Methods and Programming, new computing technologies, Volume 10, pp.96-100.
7. Paterson S, Closier J (2009) Computing in High-Energy Physics and Nuclear Physics.
8. Tsaregorodtsev A, et al. (2003) Computing in High-Energy Physics and Nuclear Physics.
9. (2014) The site of the experiment BESIII. Available: http://bes3.ihep.ac.cn/. Accessed: 2014 Sep 10.
10. (2014) The site of the project DIRAC. [Electronic resource] Available: http://diracgrid.org. Accessed: 2014 Sep 12.