IHEP Data Center Simulation

Daria Priakhina, Andrey Nechaevskiy, Vladimir Trofimov, Gennady Ososkov, Dmitriy Marov
symsim@jinr.ru
Laboratory of Information Technologies, Joint Institute for Nuclear Research, Dubna, Russia

Dynamically developing IHEP experiments expect to deal with the Exabyte data scale and need the corresponding means of distributed computing.
The development of sophisticate grid-cloud systems intended to store, distribute and process super big volumes of experimental data inevitably demands a substantial study of their optimality by the detailed simulation of these systems [1].
The simulation program SyMSim (Synthesis of Monitoring and Simulation) [2] was developed at LIT JINR and then modified for the IHEP simulation.

Basic Concepts of Simulation

- ✓The simulation goal is to satisfy some **optimality criterion,** which minimizes the equipment cost under the unconditional fulfilment of **SLA** (Service Level Agreement).
- ✓The best way to dynamically evaluate the system functioning quality is to use its **monitoring tools.**
- ✓The simulation program is combined with a real monitoring system of the grid-cloud service through a special **database** (DB), which includes a setup of the simulated computing center, the parameters obtained from monitoring information as data flow, job stream, etc.
- ✓DB includes the simulation results.
- ✓**A web site** is needed to create a simulated infrastructure (figure 1), configure the equipment parameters and view the simulation results in the charts. [3]

Figure 1. Web site page for building a simulated infrastructure

Simulation Experience of the China IHEP Data Center

A simplified schema of the China IHEP Data Center (case 1) with two possible extensions (cases 1-2) illustrated in figure 2 was used in the simulation experiments.

The first case was simulated as the typical process of job flow in **one computing node with 500 PC**.
A file needed to perform one or more jobs should be available on the remote Disk Server and requires downloading to the local pool. **10000 jobs with the files from 0.1 GB to 100 GB** were submitted. The time, when CPUs are busy by idling jobs, because they cannot start waiting for this file, can be considered as an important characteristic of the computing system loss.

The second case includes **two computing nodes**. To analyze the usage level of these two cases, the intensity of data and job flow and the load of the communication equipment for one and two computing nodes varied in the simulation. Based on comparing the results of our simulations one can identify problems confirmed by the quantitative characteristics, which arise in the process of data processing.

Among the events occurring in the system during the simulation run, for the considered cases, one can compare such characteristics of the computing process as the job queue dynamics, the load of switches or cases of the system loss since CPUs are busy by idling jobs (table 1).

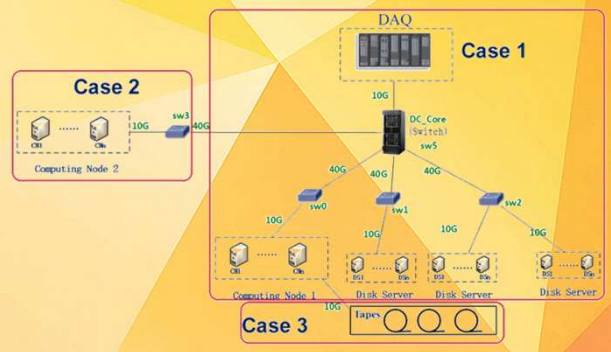

Figure 2. Simplified schema of 3 possible cases of the China IHEP Data Center with one or two computing nodes or the robotized tape library extension (cases 1-3).

Table 1. Comparison of time intervals between job submitting and execution start for cases 1 and 2.

	Average delay time (min)	Number of jobs w/o delay	Number of jobs with delay less than 60 (min)	Number of jobs with delay over 60 (min)	Total waiting time (min)
Case 1 (500 CPU)	7,2	8567	872	561	72209
Case 2 (2*500 CPU)	12,6	7598	1396	1006	126245

It was decided to base the comparison of the considered cases on system losses due to CPU load by idling jobs waiting for a file.

Simulation results

- ✓The proceeding of all jobs has quite the same time.
- ✓In case 2 (table 1) the total waiting time (time when the job was waiting for a file) increased by 75%. So 500 CPUs added in case 2 were used in not an effective way.
- ✓However, if a different way of increasing the computing power was chosen and not a computing node, but extra 500 cores to the existing computing node were added, losses on the node with 1000 cores would stay at the same level.

Data Acquisition System Simulation

The program allows one to obtain a number of important quantitative characteristics of job and data flow processes needed to see how to optimize the system.
The simulation shows that the attempts to increase the power of the computer system by enlarging the number of computer nodes lead to increasing system losses due to idle processors.
However, we can keep losses at the same level, if we increase the computing power by enlarging the number of cores in one node.

Let us consider **the third case** (figure 2), which includes **the data stream from the data acquisition** (DAQ) infrastructure to be stored on a **robotized tape library**. The DAQ system receives and stores event data from the individual detector with a frequency corresponding to the output frequency of the trigger system. Such trigger periodicity was measured by the trigger period. **The simulation run for different trigger periods** was performed, and the dependence of the total number of transferred files from this period was calculated, which is shown in figure 3.

The number of transferred files also depends on the number of drives in the robotized tape library. The following was simulated:
1. files stored from the DAQ system to the buffer without a tape copy, such files cannot be deleted (figure 4).
2. the total number of files with a tape copy is shown in figure 5.
As it seen, **with 7 drives any queue of files can be avoided because there is enough time to write all the files on tapes.**

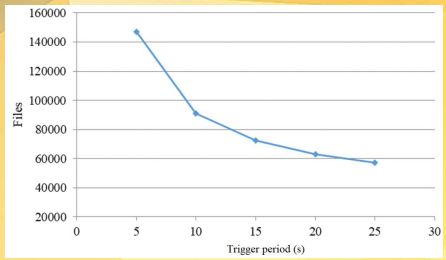

Figure 3. Dependence of the total number of files on the trigger period

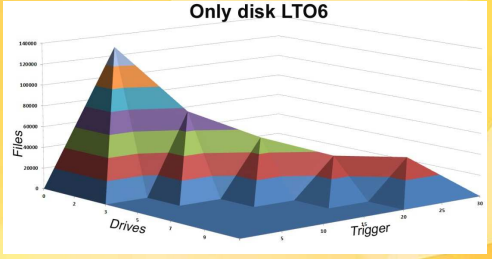

Figure 4. File number dependence on the trigger period and drive number in the robotized tape library when files are stored to the disk only, without a tape copy

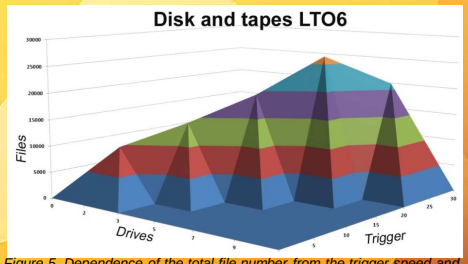

Figure 5. Dependence of the total file number from the trigger speed and drive number in the robotized tape library, when files are stored on the disk and tapes

- ✓ *The first experience with simulating IHEP computing is very preliminary and intended just to try adapting the existing simulation program to the IHEP specifics.*
- ✓ *A new version of the simulation program has been already installed in the China IHEP Data Center, adapted to some of its parameters and tested.*
- ✓ *The first attempt of simulations was performed on the quite simplified model of the IHEP Computing Center disregarding such important parts as the DAQ infrastructure, etc.*
- ✓ *Nevertheless, since a certain success of this experience has demonstrated the applicability of the simulation program, we are going to extend the IHEP Computing Center model to be simulated gradually approaching to its present and then planned structure.*
- ✓ ***Currently, the transition to a new, more promising approach to the complex computer system simulation, which takes into account the cost components of the system equipment, is in progress [4].***

References

1. *Andrey Nechaevskiy, Gennady Ososkov, Darya Pryahina, Vladimir Trofimov, Weidong Li. Simulation approach for improving the computing network topology and performance of the China IHEP Data Center // European Physical Journal (EPJ) – Web of Conferences, 2019 (accepted).*
2. *SyMSim (Synthesis Monitoring and Simulation) program web page http://symsim.jinr.ru/*
3. *Marov D.M., Pryahina D.I. Modernization of the web service for the datacenter simulation program // CEUR Workshop Proceedings. — 2018. — Vol-2267. — Pp. 573–578. – ISSN 1613-0073 (in Russian).*
4. *Priakhina Daria, Korenkov Vladimir, Nechaevskiy Andrey, Ososkov Gennady, Trofimov Vladimir. Simulation of data storage and processing centers taking into account economic components // Electronic journal "System analysis in science and education" . — 2018. — Vol-4. — P. 9 – ISSN 2071-9612 (in Russian).*