

Approach and tools for working with Big Data on the “Govorun” supercomputer

Wednesday, 13 October 2021 15:45 (15 minutes)

It is noteworthy that modern HPC systems are used not only as traditional computing environments for performing massively parallel calculations, but also as systems for Big Data analysis and artificial intelligence tasks that arise in different scientific and applied problems. At the same time, despite the increase in supercomputer performance, memory and data storage bandwidths are becoming bottlenecks. To accelerate work with data for tasks of different types, solved on the “Govorun” supercomputer, a hierarchical hyperconverged data processing and storage system with a software-defined architecture was developed and implemented. This system is based on the disaggregated RSC solution for data processing and storage, which enhances the efficiency of solving user problems, as well as increases the efficiency of using both computational resources and data storage resources. It realizes a new paradigm for working with data, i.e., the integration of computing elements and novel types of data storage elements (Intel Optane PMem, Intel SSD) into a unified computing environment.

The technologies implemented on the “Govorun” supercomputer are used for efficient data processing of the MPD experiment of the NICA megaproject and in the joint projects of Meshcheryakov Laboratory of Information Technologies and Laboratory of Radiation Biology.

The studies in this direction were supported by the RFBR special grant (“Megascience –NICA”), No.18-02-40101.

Primary author: Mr ZUEV, Maxim

Presenter: Mr ZUEV, Maxim

Session Classification: Information Technologies

Track Classification: Information Technology