Contribution ID: **428** Type: **Plenary talk**

GRID Didn't Take Off, But Is There a Chance?

Monday 7 July 2025 16:30 (30 minutes)

The GRID concept, as a computing infrastructure, according to its authors, has not been implemented by 2022 [1]. This concept was proposed by Foster and Kesselman [2] in 1999. It should be noted that this concept best suits the needs of modern applications such as on-demand computing, on-demand data storage, etc. [4]. Here is how the authors themselves formulated the main properties of the GRID infrastructure (cited follow [1]):
1. "[C]ordinate resources that are not subject to central control⋯(The Grid integrates and coordinates resources and users that are in different control domains—for example, a user's desktop versus a central computing domain; different administrative units of the same company; or different companies; and solves the problems of security, policy, payment, membership, etc. that arise in these conditions. Otherwise, we are dealing with a local control system.)
2. ⋯use of standard, open, general-purpose protocols and interfaces⋯(The Grid is built from multi-purpose protocols and interfaces that solve such fundamental problems as authentication, authorization, resource discovery, and resource access. ⋯[I]t is important that these protocols and interfaces be standard and open. Otherwise, we are dealing with an application-specific system.)
3. ⋯provide non-trivial qualities of service. (A Grid allows the use of its constituent resources in a coordinated manner to provide different qualities of service, such as response time, throughput, availability, and security, and/or the coordination of multiple types of resources to meet complex user requirements, so that the combined system provides value that is greater than the sum of its parts.)"
Foster and Kesselman in [1] have stated that none of the numerous attempts to implement the GRID concept have yet been successful. None of the projects listed in this paper have been able to implement the above requirements, to automatically minimize the execution time of an application in a computing infrastructure given its time constraints through efficient automatic management of computing load balancing between available computing devices, "non-trivial quality of service", efficient management of data transfer between application components, and the ability to increase computing resources on demand without limitation [1]. The presentation discusses the reasons for this disappointing conclusion that are computer performance, available data rates, and available mathematical methods for managing computing infrastructure resources in the early 2000s were insufficient to achieve these properties.
However, it has been significant breakthrough in the period 2014 and 2024, the server performance increased by about 7 times, supercomputer performance increased by 4 orders of magnitude, and
the maximum data rate increased from several Tbit/s to 1.1 Pbit/s over the same period [3]. These advances have stimulated a breakthrough in mathematical optimization methods based on machine learning. The point is that, given the above-mentioned speed of computation and data transfer, as well as the fact that network services and application components are easily scalable and operate in real time, to optimize resource management, data flows and computations in the next-generation computing infrastructure, let's call it Computing Centric Network (CCN), control algorithms with low time complexity are required, since time delays in decision-making become critical for the efficient operation of CCN. Classical optimization methods [3] are not suitable for this purpose, since they are based on centralized decision-making, i.e., centralized combinatorial enumeration of solution options is carried out using a (deterministic) algorithm capable of finding the best solution to the problem. In addition to computational complexity, this approach is associated with high overhead costs for collecting, processing and transmitting data between components of the computing infrastructure. The report shows that under these conditions, distributed multi-agent optimization (MAO) methods are the preferred choice. In these methods, the solution to the problem is obtained through the self-organization of a distributed set of algorithm agents capable of competition and/or cooperation and having their own criteria, preferences, and constraints. It is considered the solution is found when, the agents reach consensus (temporary equilibrium or balance of interests).

References:

1. Foster and. C. Kesselman, “The history of the grid,”arXiv preprint arXiv:2204.04312 (2022).
2. I. Foster and. C. Kesselman, “The Grid 2: Blueprint for a New Computing Infrastructure”(Elsevier, Amsterdam, 2003).
3. R.Smeliansky, E.Stepanov Machine Learning to Control Network Powered by Computing Infrastructure. ISSN 1064-5624, Doklady Mathematics, 2024. © Pleiades Publishing, Ltd., 2024.
4. R. Smeliansky, “Network powered by computing: Next generation of computational infrastructure,”in Edge Computing—Technology, Management and Integration (IntechOpen, 2023), pp. 47–70.

Author: SMELYANSKIY, Ruslan (Moscow State University)

Presenter: SMELYANSKIY, Ruslan (Moscow State University)

Session Classification: Plenary