Taking research experiments to another level with superior data management services

The most demanding, data-intensive research experiments in Europe and beyond require services that are capable of operating at an unprecedented scale. An EU initiative developed scalable technologies to meet the challenge of supporting petascale data management in modern e-Infrastructures.

For many years, managing extremely large and growing volumes of data has been a challenge for scientific experiments that use distributed e-Infrastructures for their computing needs. New features and functionalities need to be developed and made available to the research community to cope with the dynamic nature and flexibility of these powerful resources.

Managing data in highly distributed computing environments

The EU-funded XDC project developed and released enhanced data management services that can be coherently harmonised with current and next-generation e-Infrastructures deployed throughout Europe, such as the European
Open Science Cloud (EOSC) and the Worldwide LHC Computing Grid coordinated by the European Organization for Nuclear Research. These open, interoperable and easy-to-use services will help to build global infrastructure for distributed computing.

XDC team members improved existing federated data management services by adding missing functionalities. “New users mean new requested functionalities,” comments project coordinator Daniele Cesini. “Significantly extending the provided functionalities is of utmost importance in building infrastructures that can be exploited by user communities different from those that historically founded their computing models on distributed systems.” Team members enhanced user experience in accessing such data management services by providing more user-friendly interfaces.

The scientists provided adaptable functionalities to address modern e-Infrastructures’ increasing dynamic nature and flexibility. “Due to the advent of virtualisation techniques, cloud computing paradigms, and Infrastructure as a Service and Platform as a Service orchestration tools, resources once identified as ‘sites’ in e-Infrastructures have become ‘liquid’ and highly dynamic,” Cesini explains. Sites can be created, destroyed, attached and detached from infrastructure with a few mouse clicks in a time period that was inconceivable just a few years ago.

Furthermore, the resources created or attached can be heterogeneous in nature, without a predefined architecture. “However, when it comes to data management, a high dynamicity poses huge challenges with respect to efficiency, transparency and reliability,” he adds. XDC delivered data management solutions to dynamically extend a computing centre to a remote site that provides transparent bidirectional access to data stored in both locations. It also offered solutions to dynamically include sites with limited storage capacity, thus providing transparent access to data stored remotely.

Open-source platforms available for widespread use

To facilitate interoperability, standardisation and adoption, the XDC architecture uses open standards and protocols available on state-of-the-art distributed computing ecosystems to guarantee that the released components can be easily plugged into European e-Infrastructures and cloud-based computing environments overall. Project partners created two open-source software releases that can be deployed on public and private cloud infrastructures: XDC-1 (code name Pulsar) and XDC-2 (Quasar). Both are based on existing production quality services that were enriched with new functionalities and usability improvements to make complex infrastructures exploitable by an increasing number of user communities. They organised these building blocks in a coherent architecture and provided several contributions.
catalogue describes the services and the new related functionalities developed and enhanced during the project.

“XDC delivered important and innovative services that have been proposed as candidates for inclusion in the EOSC-hub Service Catalogue,” concludes Cesini. The EOSC-hub project simplifies access to a comprehensive suite of products, resources and services supplied by major pan-European and international organisations.

Keywords

XDC, data, computing, data management, e-Infrastructure, distributed computing, XDC-1, XDC-2

This project is featured in...
Discover other articles in the same domain of application

SCIENTIFIC ADVANCES

Getting down to the core of future earthquakes in Europe

3 June 2022

SCIENTIFIC ADVANCES

Need help preparing scientific documents?

27 December 2023