

<https://events.gwdg.de/event/658/abstracts/> -> deadline 29th May 24

Cloud based flexible service infrastructure stack for the NFDI

Authors:

Jonathan Bauer^{1*}, Marcel Tschöpe¹, Paul Chr. Schnürle¹, Julian Weidhase¹, and Christian Garth², and Timo Mühlhaus²

*Lead presenter, jonathan.bauer@rz.uni-freiburg.de

¹DataPLANT, University of Freiburg

²RPTU, DataPLANT

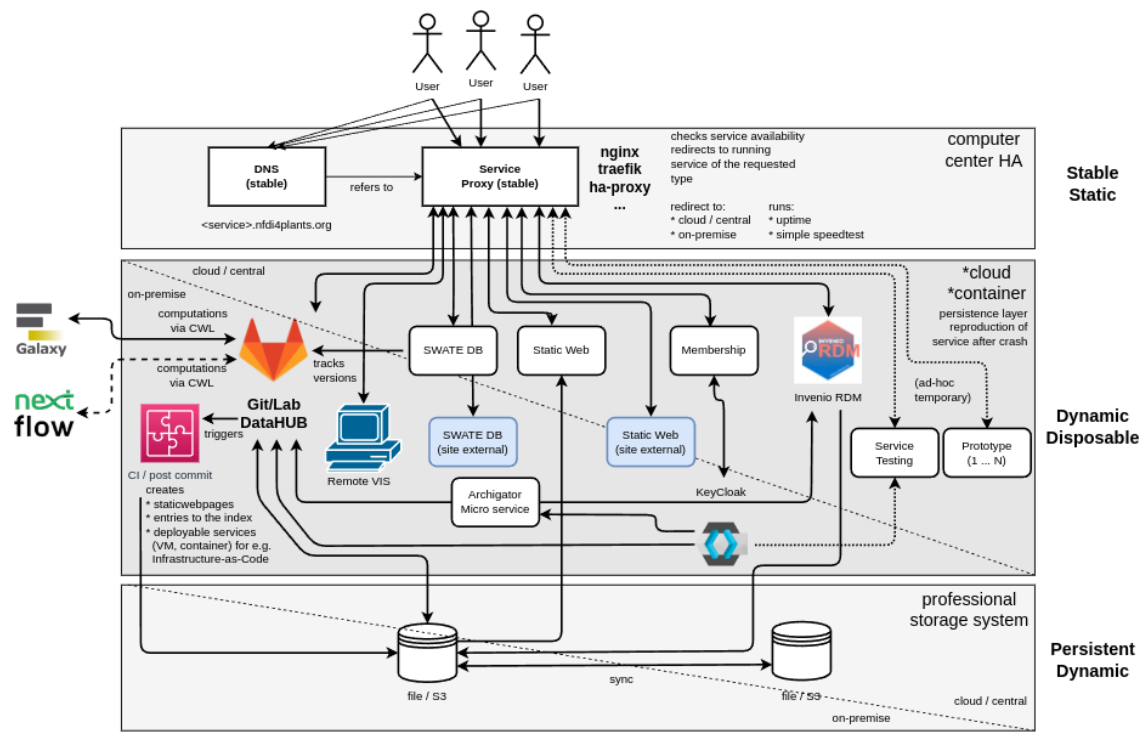
Abstract:

A core objective of the DataPLANT consortium is to provide a range of tools and services to foster its community in research data management. The offered science gateway-base for data management and publication workflow-as well as the various tool backends are designed as a set of flexible cloud-based (micro) services. The setup aims at both on-premises installations and future integration into a shared NFDI infrastructure. We will present the DataPLANT DataHUB, which provides various RDM workflows to support research data scientists at different stages of the data lifecycle - from development to publication of the results obtained. In such setups storage is usually the most local resource to consider because of the amounts of data, data protection and security considerations. Computations on the data should be done near to the storage, to achieve this remote desktop infrastructure and remote visualization would be helpful. [Fig. 1]

The aim is to exchange ideas with other NFDI consortia on the services required for RDM and the principles for service development and deployment. These considerations can be used as input for joint infrastructure development, e.g. in the context of the NFDI Common Infrastructure section or Base4NFDI. Over the past two and a half years, the DataPLANT team has developed a set of software and system components that provide services to the basic plant research community. The set of tools and microservices that have been developed and evolved to date have focused on extending the existing digital landscape of the typical plant scientist. The core services focus on data management, versioning, sharing and publishing. All services are designed as portable modules that can be integrated in the larger base infrastructure. A reverse proxy automatically manages SSL certificates for new services. It also greatly increases the flexibility of the entire infrastructure by being able to quickly change the backend services that respond to the static user-facing service hostnames. Other base services include a central authentication service and monitoring services to detect service outage and to gather performance data.

The development of applications and tools to support bottom-up community-driven research data management requires the involvement of several parties. During the development of the services, we agreed on design principles to provide high-level guidance and a set of criteria for creating desirable and maintainable applications. In DataPLANT, tool development is always motivated by community requirements, conveyed by researchers, e.g., through data stewards, to developers. Developments in DataPLANT follow an incremental and iterative approach, ensuring commitment and alignment of expectations of all stakeholders. Another aim of the service development is to enable both central and local

installation of services without divergent implementations. Thus, we hope to encourage adoption by other communities and integration into a future NFDI service infrastructure.



Version 2024-05-29

[Fig. 1] The actual service stack of DataPLANT in layer oriented representation.