Effective version controlled digital object access and sharing over a networked environment using DOIP and NDN

Cas Fahrenfort and Dr. Zhiming Zhao

System and Networking Lab, University of Amsterdam, the Netherlands

Introduction: FAIRness and data infrastructure

Data are playing an increasingly important role in scientific research. However, those data objects often
- lack globally resolvable names (or identifiers)
- have different versions which makes the discovery, access and citation of those data objects difficult
- are large in size, making performance of the object sharing critical
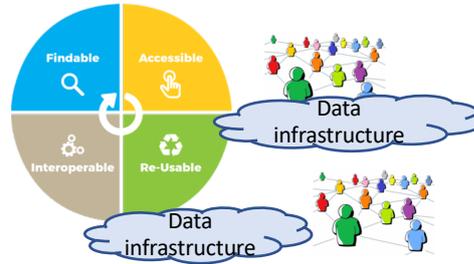
Advanced data infrastructures are thus required to effectively manage the dynamic evolution (versioning), identification and citation (globally resolvable names) of digital objects, and enable their high find-ability, accessibility, interoperability and re-usability (namely **FAIR**) among distributed communities.

Proposed solution: NDN4DI

An on demand NDN environment manager (NDN4DI) allows
1) a data manager to publish data objects and obtain the PID for them,
2) plan a customized NDN environment at Cloud providers, and
3) automate the provision and deployment of the planned environment.

After the NDN environment is in operation, the NDN4DI also provides functions for handling PID repositories from different publishers
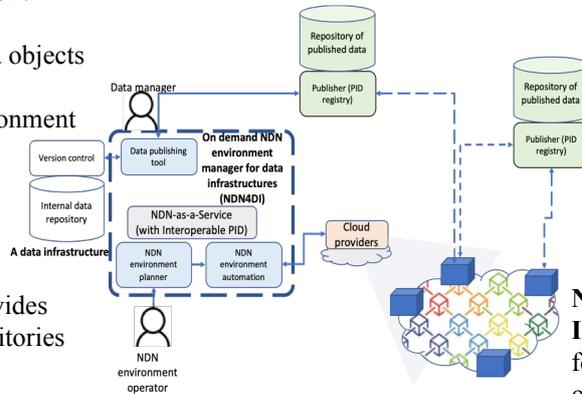

Digital Object Interface Protocol (DOIP) defines
1. an Identifier/Resolution system for allotment and resolution of Persistent Identifiers (PIDs)
2. a Repository System for storing digital objects
3. a search-able registry for all digital objects stored in the Repository System

Named Data Networking (NDN) is an **Information Centric Networking (ICN)** solution for routing objects through a network based on object name instead of location.

Prototype and use case

An initial prototype which offers functionality to persistently publish version control repositories (Git and SVN) to a number of currently available, online, freely accessible publishing systems has been developed in relation to this project.
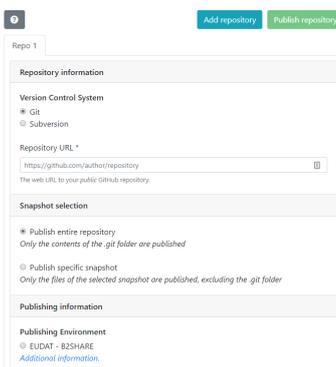
The digital content from the version control system can be published as a **Digital Object** with a **PID**.
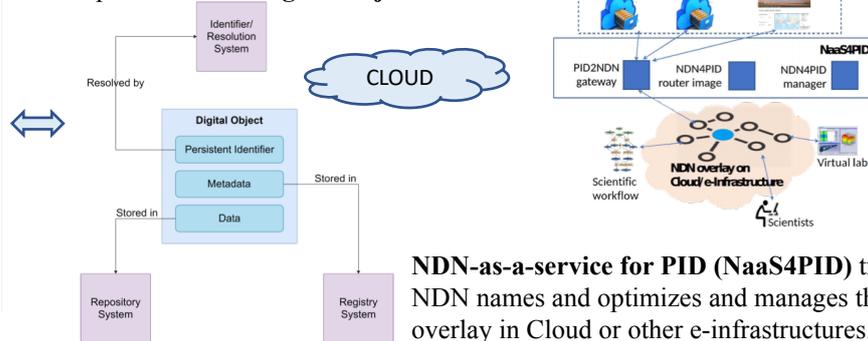
NDN-as-a-service for PID (NaaS4PID) translates PIDs to NDN names and optimizes and manages the virtual NDN overlay in Cloud or other e-infrastructures.

The prototype is built using Angular (TypeScript) and ASP .NET Core (C#). Scan the QR code below to check out a live demo or source code. The prototype allows a user publish versioned objects from a repository, with a standardized metadata schema and publishing of multiple, dependent repositories simultaneously.

Acknowledgement

References:
[1] Spiros Koulouzis, Rahaf Mousa, Andreas Karakannas, Cees de Laat, Zhiming Zhao (2018) *Information Centric Networking for Sharing and Accessing Digital Objects with Persistent Identifiers on Data Infrastructures*, in the 3rd Int'l Workshop on Distributed Big Data Management, in the context of IEEE CCGrid, Washington, US [doi 10.1109/CCGRID.2018.00098].