Federated storage studies,
develop and implement a prototype to federate WLCG centers of different levels (T1, T2) and University clusters within one National Cloud for
the LHC data handling, storage, processing and analysis.

WLCG Demonstrator R&D Project Proposal
(draft 0.16, Feb 29, 2016)

Participants :

1. National Research Center "Kurchatov Institute" (Moscow)
    a. NRC-KI – Leading Institute
2. Joint Institute of Nuclear Research – JINR (Dubna)
3. Institute For Theoretical and Experimental Physics – ITEP (Moscow)
4. Petersburg Nuclear Physics Institute – PNPI
5. St.-Petersburg State University – SPbSU
6. Moscow Engineering Physics Institute – MEPhI
7. Skobeltsyn Institute of Nuclear Physics – SINP (Moscow)
8. The European Organization for Nuclear Research – CERN (Geneva)
9. The German Electron Synchrotron – DESY (Hamburg)

Leading PI – A.Klimentov, co-PI – A.Kiryanov, A.Zarochentsev
Participants : A.Kryukov(7), D.Krasnopevtsev(1,6), E.Lyublev(3), R.Mashinistov(1), A.Petrosyan(2), A.Poyda(1), E.Ryabinkin(1), S.Smirnov(6), D.Duellmann(8), P.Hristov(8), M.Lamanna(8), A.Peters(8), P.Fuhrmann(9) , T.Mkrtchan (9)

- We propose 1.5 years (2016-17) R&D project to develop, prototype and implement a federated data storage system between Russian WLCG sites and University clusters and to demonstrate how data reconstruction and analysis payloads (for at least two LHC experiments ATLAS and ALICE) and for bioinformatics applications (genome sequencer group of NRC-KI) running on WLCG facilities, on super-computers and Academic cloud can access federated data. This project intends to implement a federated distributed storage for all kind of operations such as read/write/transfer(3rd party)-RWT, and access via WAN from heterogeneous computing resources, such as Grid centers, University clusters, academic and commercial clouds and supercomputers.

The main objectives of research will be the development, prototyping and implementation of the federated data storage system with a single point of access and its own system of internal controls, as well as the development and implementation of a prototype of distributed data analysis center including heterogeneous computing resources: «Academic cloud», grid centers and supercomputer center. In this architecture, the system appears to the user as a single entity, while it has geographically distributed resources (such as disks or computers). The system should have the properties of fault tolerance by the vital components redundancy, scalability, with the ability to make changes in the topology without shutting down the entire system, security with mutual authentication and authorization of access to data and metadata, optimal data transfer routing providing the user access to the data directly to the nearest (optimal) target server and universality, which implies applicability to a wide range of research projects of various sizes, including, but not limited to the experiments at the Large Hadron Collider and compute-intensive bioinformatics applications.

A prototype of federated storage system will be based on computing resources located at PNPI, SPbSU, NRC KI, JINR, ITEP, MEPhI, SINP and CERN. The EOS, dCache and possibly DynaFed software packages will be exploited as the underlying data storage and management technology. We will try different configurations of computing centers, such as : T1-T1-T0, T0-T1-T2, T1-T2-T3, T2-T2-T3, etc. System efficiency and performance will be demonstrated using synthetic and experiment specific tests.

During 2016-2017 :

- The topology of the system and the general system of user authentication and authorization will be determined.
- A system of monitoring of wide area network capacity will be set up and put in place, to achieve it all federation centers have been equipped by hardware and software to measure the data transfer rate and latency (perfSONAR). The matrix containing information on wide area network capacity for the federation centers will be built. This information will be part of the future software for cyberinfrastructure, and now it is used to compare

the data access efficiency between remote access and local access.

- Together with storage system's developers synthetic tests to measure system performance will be determined and conducted.
  - At phase0 (it is in progress now) we've set up T0-T2-T2 and T1-T2-T2 configurations and we are using bonnie as a synthetic test
- Together with ATLAS and ALICE we will propose and implement a technique to check the data access efficiency for the LHC experiments ATLAS and ALICE (in the analysis of data);
  - The efficiency of the created architecture will be shown (compared with the classic grid solution) for the reconstruction and analysis of events at the ATLAS Transition Radiation Detector (ATLAS TRT, program has been provided by the ATLAS TRT team) and for the event selection from an extremely large number of events at Time Projection Chamber of ALICE.
- For each technology (EOS and dCache), different configurations will be set T1-T2-T3, T2-T2-T2, T2-T1-Tn with RWT access to the data from CERN, RF T1, HPCs and T2/ns
- We will work together with site administrators and pilot users on operational practices that will allow ease of installation, stability of the federated storage system with respect to the hardware and network failures, end-to-end traceability of the full request flow and seamless.

- operation of the storage federation as the single entity will be developed, deployed and tested.