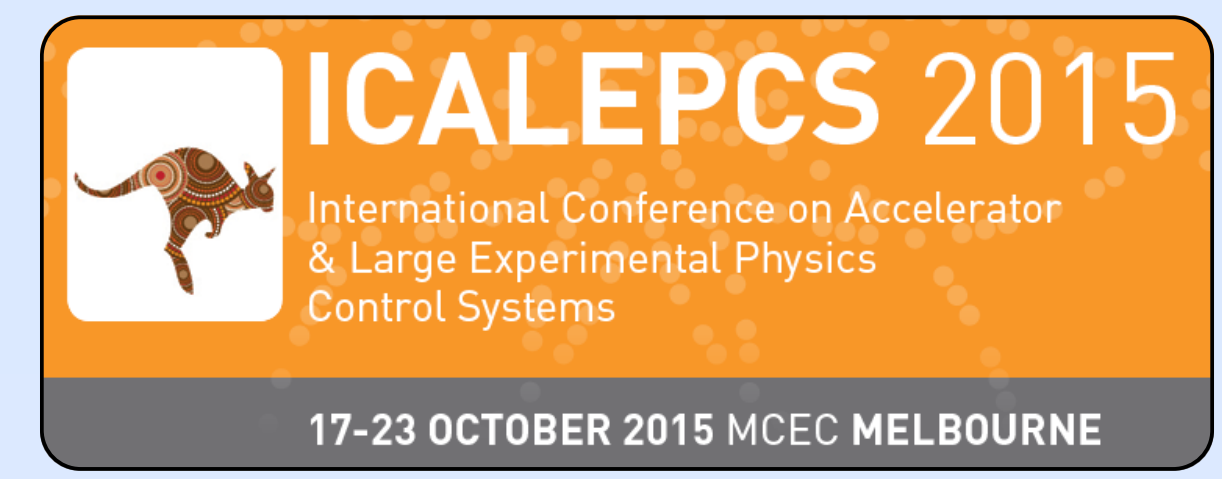




# Towards a Second Generation Analysis Framework for LHC Transient Data Recording

S. Boychenko, C. Aguilera-Padilla, M. Dragu, M.A. Galilee, J.C. Garnier, M. Koza, K. Krol, R. Orlandi, M.C.Poeschl, T.M. Ribeiro, M. Zerlauth, CERN, Geneva, Switzerland



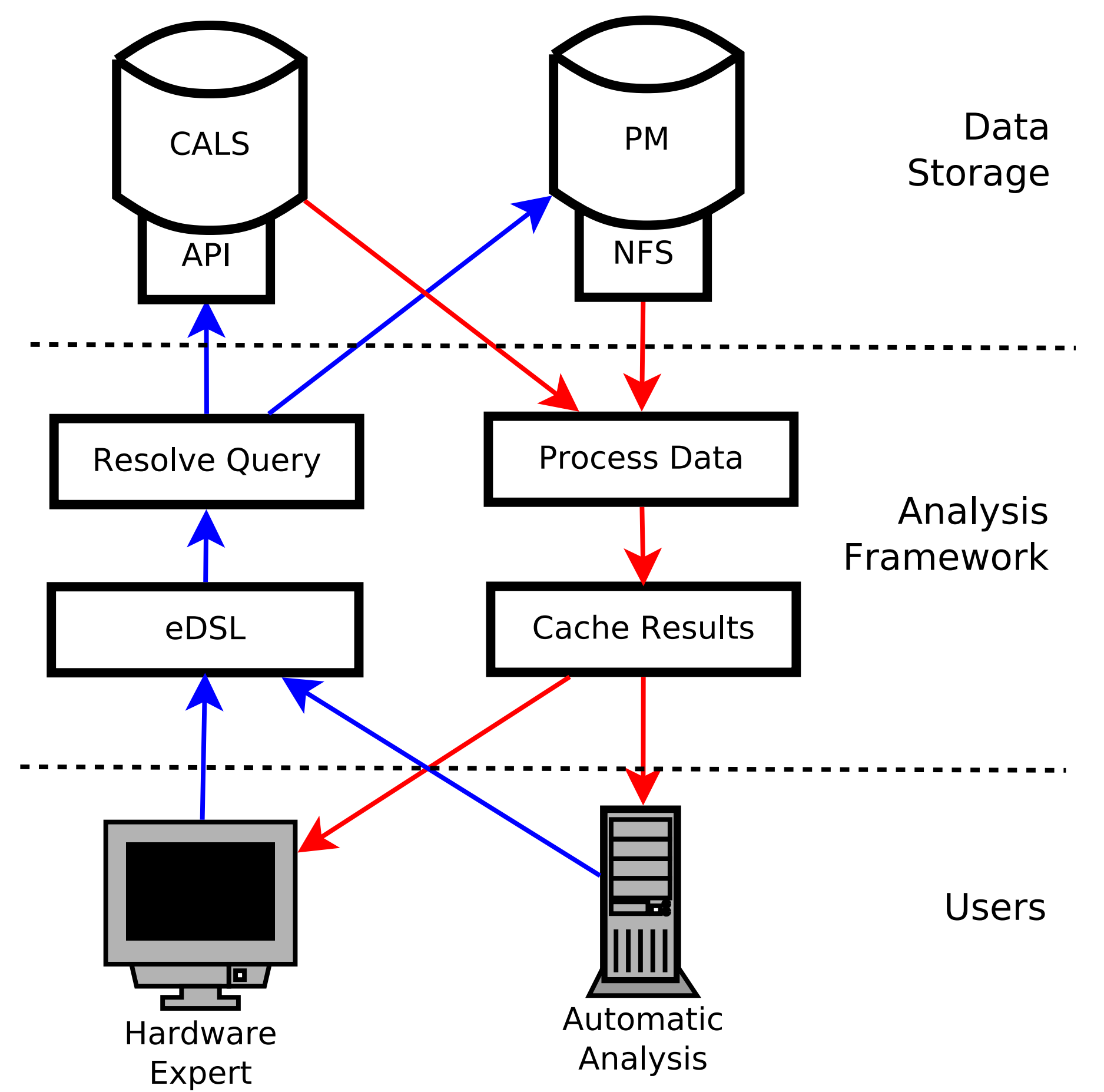
## Motivation

Based on experience of the previous hardware commissioning we have identified the main shortcomings of the current data analysis approach. Combined with the use cases collected over several operational years and the feedback provided by operators and hardware experts, we were able to define the main goals and requirements for the next generation accelerator data analysis framework:

- Improve data extraction throughput and make parallel processing efficient
- Correlate the data from multiple storage systems in a transparent way for users
- Perform calculations close to data to reduce unnecessary data transfer overheads
- Provide operators and hardware experts with a simple language to define analysis processes
- Handle heterogeneous workloads efficiently

Some of the aforementioned requirements are already satisfied by the currently deployed analysis framework (portrayed in the "Analysis Framework Workflow"). The overall performance of the implemented data processing engine depends on the efficiency of the underlying data storage services.

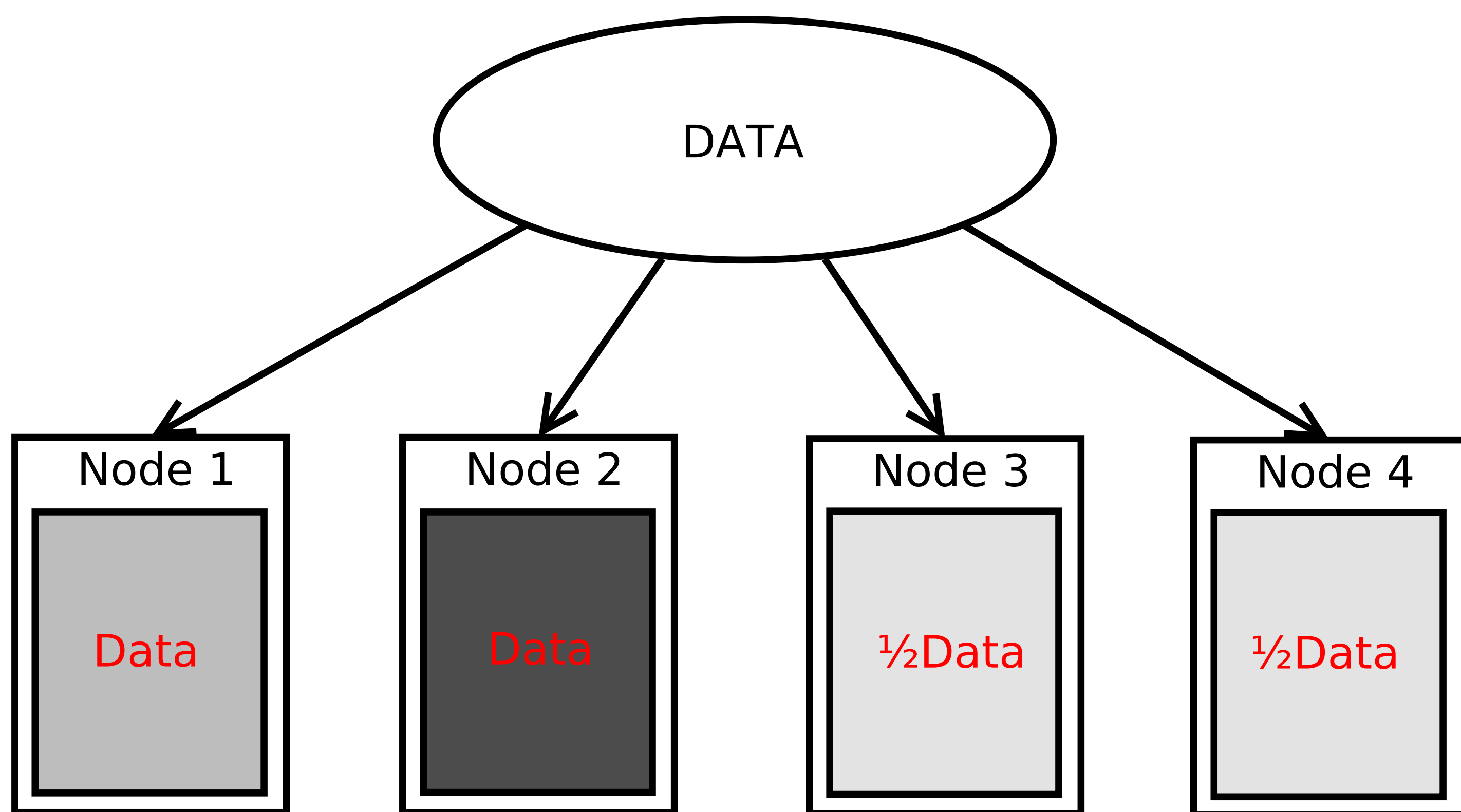
## Analysis Framework Workflow



## Mixed Partitioning Scheme Replication

To overcome the currently existing limitations and satisfy defined requirements, we propose a new strategy for data storage, named Mixed Partitioning Scheme Replication. The main idea behind the proposed approach is to replicate the data partitioned according to different criteria. Both partitioning and replication are workload-driven and aware of each others existence and responsibilities, so the overall data storage strategy can be optimized for heterogeneous workloads.

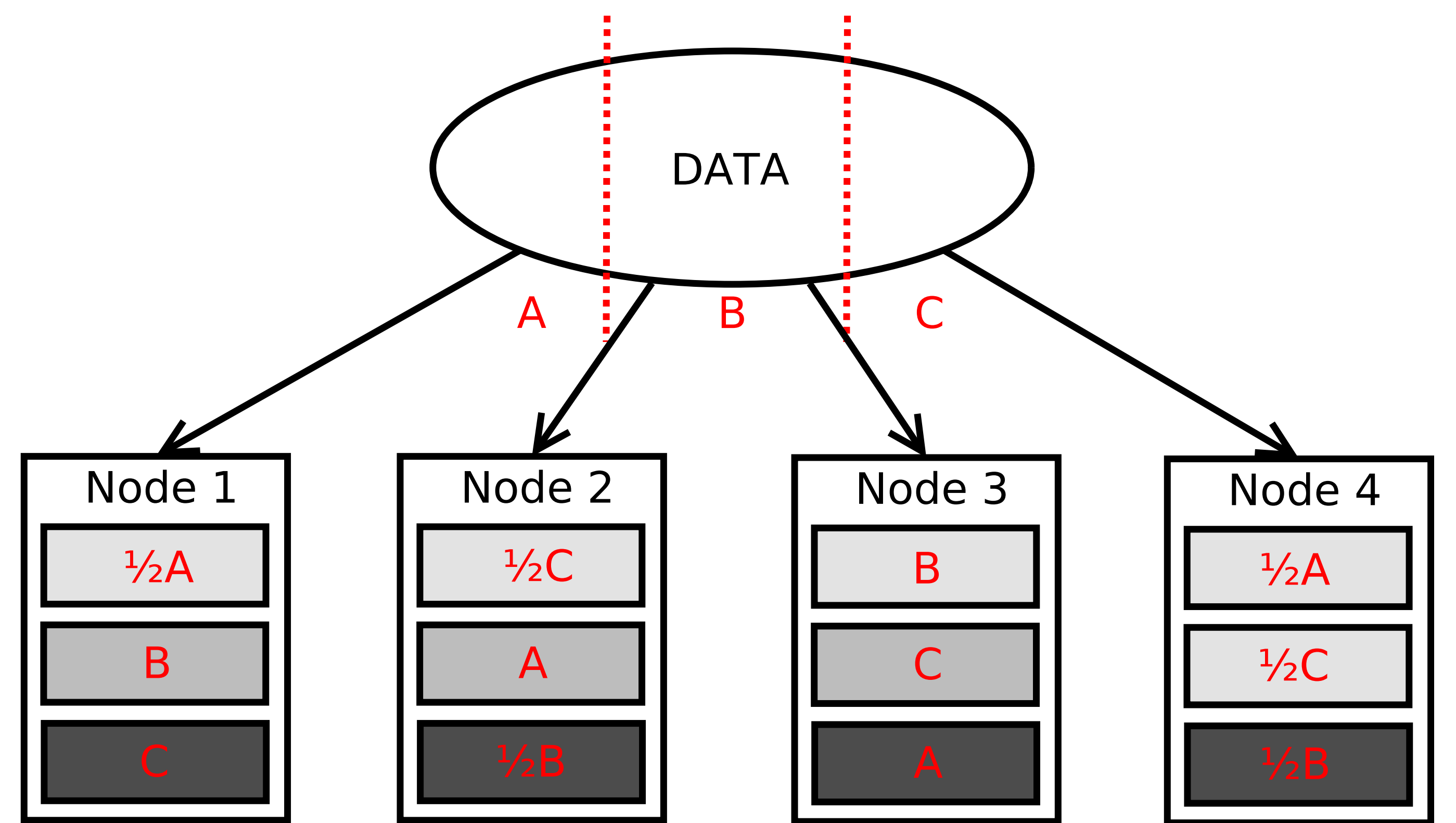
### Simple Mixed Partitioning Scheme Replication



Each replica has full copy of the data partitioned using different criteria.

- ✓ Efficient handling of highly heterogeneous workloads
- ✗ Might create very unbalanced situations (some nodes might be underused, while others under heavy load)

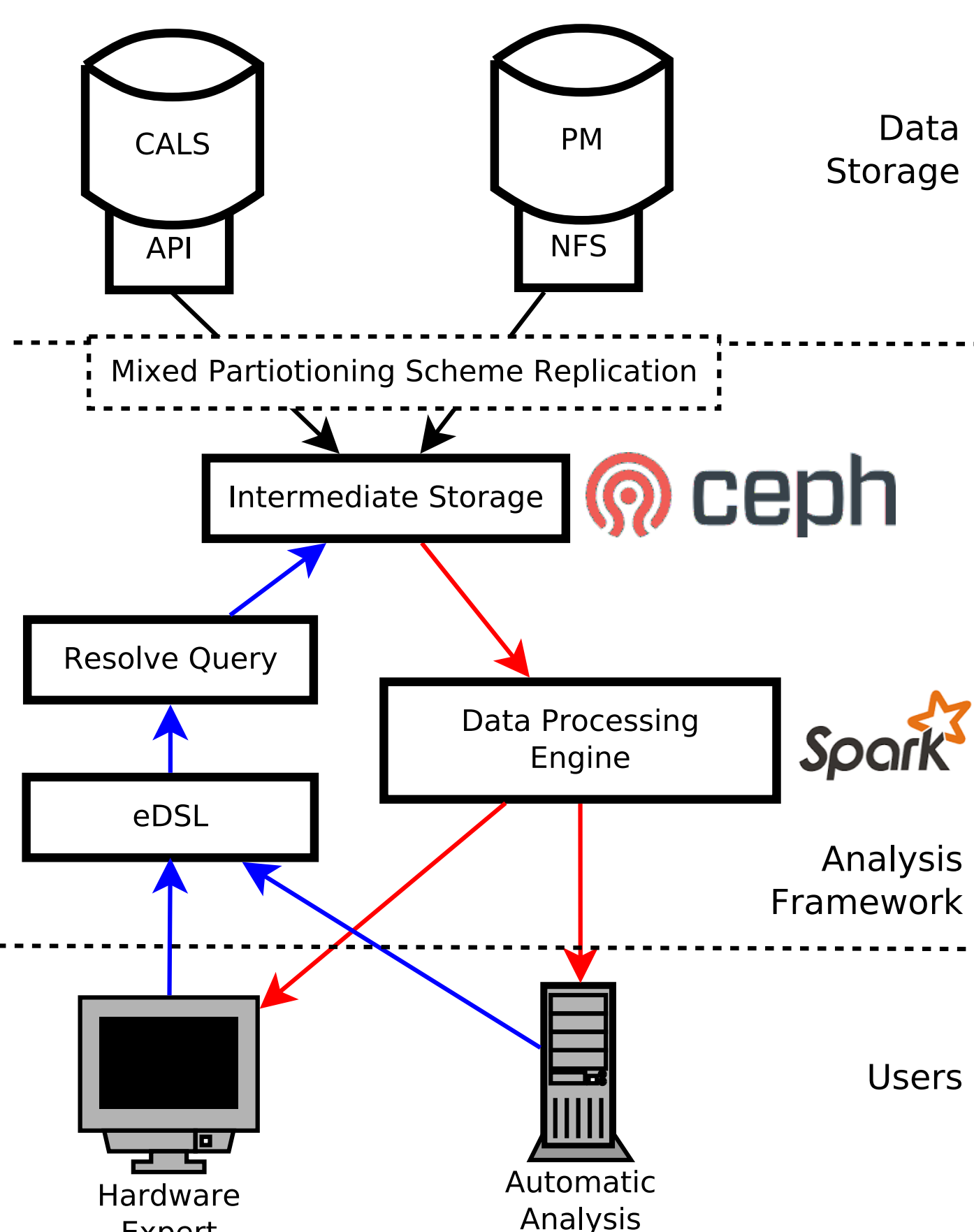
### Complex Mixed Partitioning Scheme Replication



Data is partitioned in two phases, each replica uses different partitioning criteria to partition the data segments.

- ✓ Optimization of storage strategies for the load distribution throughout the whole infrastructure
- ✗ Complex configuration, implementation and maintenance

## Improved Analysis Framework



## Proposed Solution

The architecture of the currently deployed analysis framework enables straightforward replacement and integration of new modules, thus we plan the enrichment of the existing solution with new features, rather than the replacement of the entire code base.

- The proposed solution could be initially deployed on an intermediate storage component enhanced with the functionalities of modern engines for large-scale data processing
- In addition to the possibility of working with data through eDSL, users will have the possibility to execute arbitrary requests on the data using the data processing engines native query language
- The choice of the storage will depend on its compatibility with different data processing engines and its performance with heterogeneous workloads
- Initially the data will be fetched from the CERN Accelerator Logging Service and the Post Mortem framework
- Before being available to the users, the data will be pre-processed, correlated and merged, to abstract the data source it was originally fetched from