

Adaptable Distributed Storage

BigDataStack Software Component developed by LeanXcale

The Adaptable Distributed Storage component of BigDataStack consists of a novel mechanism that allows the data storage layer to be adapted to diverse workloads during the run-time. It allows LeanXcale (LXS) datastore to partition its datasets to smaller fragments and split, move or merge those fragments effectively among the available data nodes in order to achieve the balance of the load, both in terms of incoming workloads and stored data load. As scaling in/out the nodes of a datastore is not an easy task, databases generally either sacrificing the provision of transactional semantics in order to be able to scale effectively on the run-time or have to suffer from long periods of downtime or decreased performance during the scaling process. The Adaptable Distributed Storage on the other hand will allow the LXS datastore to scale effectively during runtime, without downtimes or increased performance overheads, while at the same time, ensuring transactional semantics. This will allow its storage to become truly elastic.

Input

Information regarding how a dataset can be fragmented (split points) that can be provided by the system administrator.

Output

This component produces no output. The output could be an optimized allocation of the resources of the storage layer.

Adaptable Distributed Storage

Initial TRL

TRL 2

Final TRL

TRL 5

End Users

Any enterprise that deals with Big Data and needs a storage layer that can handle diverse workloads. Increasing data at a high rate while the need for data consistency and transactional insurance is crucial.

Key Features and Benefits

This technology can allow the storage layer to scale out automatically, without downtime or decreased performance. Moreover, data allocation is achieved automatically, without the need of a DB administrator to explicitly perform any corrective action.

Essential Information for Users

A domain expert may have to define the split points for the data to be fragmented, by defining additional indexes on specific columns.



Adaptable Distributed Storage

Standards involved in the development of the component

None

Implementation in BigDataStack Use Cases

Use Cases with an increased data volume will be a candidate.

Differentiators from competitors in the market

The Adaptable Distributed Storage can scale out to numerous nodes, with no decreased performance. Moreover, the novel algorithm for resource allocation can allow the data to be fragmented/split and moved across data nodes, thus providing an optimal deployment that is expected to reduce the overall cost for resources.