Implementing a parallel processing architecture embedded into a distributed object storage system

This technology is for a novel “in-storage” parallel computing platform that processes various types of media data (images, voice, music, videos) stored in OpenStack Swift, a distributed object storage system. The platform provides users with the computing resources of the storage system and a job management server as a one-stop operation tool. Additionally, the platform automatically deploys applications and performs parallel computation inside of the storage system by uploading both user-defined program code and its data.

Features

- **Computing Efficiency**: Effective resource usage for computational parallelism on distributed storage systems.
- **Computing Availability**: High availability from a redundancy of computing resources and stored data.
- **Scalability and Durability**: Highly scalable capacity and highly durable data protection from data loss as media storage with OpenStack Swift.
- **Operation Efficiency**: One stop service with a job management server for code deployment, execution, and monitoring.

Application Scenarios

- Parallel computing of a huge amount of media content. (e.g., feature extraction of images, sounds, and videos.)
- Applications required security/privacy (e.g. masking and filtering images, sounds, videos, etc.)
- AW3D*, which is a service providing global 3D topographic map data generation, can be a use case of this technology (under a proof of concept study)

*AW3D is a service provided the NTT DATA Corp. that provides 3D topographic map data of the world.