**Sector/Sphere** is a software platform that supports very large distributed data storage and simplified distributed data processing. The system consists of Sector, a distributed storage system, and Sphere, a runtime middleware to support simplified development of distributed data processing.

In the Sector system, a master node maintains the file system, while the data is stored on the slave nodes, possibly across multiple data centers. A security server provides user account verification, access control IP list, etc. UDT is used for high speed data transfer between slaves and between slaves and clients.

**Sphere** allows developers to write certain distributed data parallel applications with several simple APIs. User-define functions can be independently applied to each element in the dataset, while the result can be written to either the local disk or common destination files on other nodes. Data locality is a key factor for the performance of Sphere. In addition, sphere provides transparent load balance and fault tolerance.

The **Terasort** benchmark sorts \(N \times 10GB\) data, where \(N\) is the number of slave nodes in the system.

The **CreditStone** application computes the fraudulent rate for each merchant in the credit card transaction dataset.