Market Requirements Document

Feature Name: ENHANCED DATA AND SERVICE AVAILABILITY PACKAGES

Description of the Problem

Technical issues

Background: Objectivity has offered a synchronous data replication capability (originally called the Data Replication Option, or Objectivity/DRO) and a limited fault tolerance solution (Fault Tolerance Option, or Objectivity/FTO) since Release 3. The options were repackaged as Objectivity High Availability (Objectivity/HA) at Release 8.

The data replication feature uses a quorum-based algorithm to permit clients to write to multiple database copies (images). Each database image must be in a subdivision of a Federated Database termed an Autonomous Partition (AP). A logical Database may only be a member of a single AP. Creation, updating and deletion of partitions can be done online. Each AP has its own equivalent to the Federated Database database, containing Schemas and Database catalogs. Each AP also has its own Lock Server, which administers locks for all of the Databases in its partition.

Networking Issues: Although the quorum-based algorithm is still one of the most theoretically sound replication mechanisms, it only works well in a high bandwidth, low latency, reliable network. CERN and CalTech benchmarked it on local networks and on a high bandwidth link between Geneva, Switzerland and Pasadena, CA and it performed well. A similar test on a high latency, unreliable satellite link between KEK in Japan and CalTech was abandoned. In reality, we have had very few successful deployments of Objectivity/HA, mainly because of this shortcoming and complex recovery issues.

Single Points of Failure: The other major problem is that although there is no single point of failure for a federation when multiple Lock Servers are deployed, loss of a Lock Server takes all of the databases that it controls offline. Recovery from this state generally requires manual intervention. Worse still, if the Lock Server administering the Federated Database goes down it becomes impossible to create new databases, run backups or update schemas. In short, there is no Lock Server redundancy, introducing single points of failure.

Market Issues

The high availability features were introduced in order to address telecom and process control equipment manufacturers' requirements, though the first deployment was in a proof of concept for an airport control system being build by Xwave in Canada. Developers of applications for these markets were highly competent in implementing fault tolerant solutions and could avoid or work around our product's shortcomings.

The rapid growth in NoSQL and Big Data solutions has introduced new paradigms and expectations in our current and projected markets. Data sharding (which our Federated Database – Database – Container – Physical File storage hierarchy is an example of) is commonplace, with Hadoop HDFS and many alternatives offering fully redundant data replication and data servers.

Asynchronous replication of small or large objects has also become more commonplace. Some of the newer solutions also work in hybrid dedicated, geographically distributed and cloud environments, where the networks involved have variable latency and bandwidth and are much more likely to fail during a transaction. Objectivity/HA cannot be easily and reliably deployed in those environments.

Definitions

Mean Time Between Failures (MTBF) is the predicted elapsed time between inherent failures of a system during its intended operation. As an example, suppose that a system is meant to operate over a 24 hour period, but that it goes down for 1 hour after 12 hours and again for 3 hours after the first 18 hours. It has been operational for 20 hours and has failed 2 times, so the MTBF is 10 hours.

Mean Time to Recovery (MTTR) is the average time that a system will take to recover from any failure. "Recover" may mean repair, replace, resolve or be partially or fully available. We will set goals for restoring partial and full access to the data in a Federated Database later in this MRD, expressed as a ***Recovery Time Objective (RTO)***.

Recovery Point Objective (RPO) is defined by business continuity planning. It is the maximum tolerable period in which data may be unavailable due to a major incident.

Total Cost of Ownership (TCO) is a financial estimate aimed at helping implementers determine the direct and indirect costs of a product or system. Using "free" software to protect data containing valuable financial transactions may actually increase the TCO if some or all of them are lost. The Citibank currency exchange system (CAOS) that used Objectivity/DB had all of its data replicated three times, with one copy over 100 kilometres from the main data processing center. An average transaction was worth many tens of thousands of Dollars, so the equipment and licensing costs were miniscule in comparison to the value of the data.

High Availability is a system design approach and associated service implementation that ensures a prearranged level of operational performance will be met during a contractual measurement period. Contracts often associate a percentage value to this parameter. "Five Nines", or 99.999% availability, is the minimum for most telecom equipment, allowing no more than a total of 5.26 minutes per year of unscheduled downtime. Process control, aviation, medical and battlefield systems may require "Six Nines", or 99.9999% availability, i.e. no more than 31.5 seconds of unscheduled downtime per year. Both requirements usually involve solutions that need additional or specialized hardware or that reduce the impact of outages by eliminating single points of failure.

Continuous Operation refers to a system's ability to avoid planned outages. There must be ways to perform administrative work, such as hardware and software maintenance and changes, without affecting the services available to users.

Continuous Availability combines the characteristics of high availability and continuous operation to provide the ability to keep a system running without any noticeable downtime. This may be expressed in terms of ***Service Availability***.

Many other terms, such as Fault Tolerant Systems and Non-Stop Systems, are also still in common use, but we will use the above definitions for our purposes.

Description of the Requested Feature

There are two groups of requirements, expressed as:

- ➢ **1. Data Availability Packages**, expressed in terms of Mean Time Between Failures, Mean Time To Recovery and Recovery Time Objectives.
- ➢ **2. Service Availability Packages**, expressed in terms of Recovery Point Objectives, Service Availability and Total Cost of Ownership.

1. Data Availability Packages

There will be a range of offerings that cater to a spectrum of data availability requirements. There are two main packages:

- ➢ **The Data Availability Base Package** will offer no guarantees on MTBF, MTTR or achieving RTOs. If a piece of hardware or software fails then data and/or services may become unavailable and uncompleted transactions will be lost. Data will be recoverable if online, incremental backups have been maintained correctly. Recovery may require manual intervention in some cases, such as loss of a machine running a Lock Server. This equates to the standard Objectivity/DB and InfiniteGraph products today. It does not include Objectivity/HA features.
- ➢ **The Enhanced Data Availability Package** will provide customizable data replication (synchronous and asynchronous), service protection, load balancing and online maintenance features. It is aimed at deployments ranging from local clusters to geographically dispersed or hybrid (private plus cloud) environments.

2. Service Availability Packages

There will be two main packages that will enable customers to progressively increase Service Availability according to their RPO and TCO requirements:

- ➢ **The Base Service Availability Package** will offer no guarantees on Service Availability. The only choices that affect TCO will involve our standard licensing terms. Recovery of data and services may require manual intervention, or simple automation. The Recovery Point Objective is to restart all server processes and restore all committed data. There is no commitment to a minimum or maximum time to partially or fully achieve this objective. Data availability is as for the Data Availability Base Package.

- **The Enhanced Service Availability Package** will offer Five Nines or Six Nines guarantees for Objectivity server and data availability. These requirements will dictate specific Recovery Point Objectives. This feature will minimize disruption to active transactions by providing standby or load balancing servers. All maintenance, including upgrading Objectivity software, must be possible without service disruption. It will also require deployment of the Enhanced Data Protection Package.

How is this problem being solved now, and why isn't that acceptable?

- Prospects are choosing competing Big Data solutions, so we are losing revenue and traction in the market.
- We don't offer “jump start” capabilities that would make our products more easily accessible by builders of continuously available services.

What languages must support this capability?

- C++ and Java (Essential)
- Net for C# (Possibly)

Which platforms must be supported?

- Linux, Solaris, Windows and Mac OS X, in that order.

Do any competitors already have this feature?

- Most mature RDBMS products have many of the features but lack resilience in hybrid (cluster plus cloud) environments.
- The NoSQL and analytics platform vendors have some of the features. None of them has all of the features.
- Versant Object Database used to have a High Availability option based on Windows NT failover products and a RAID. Their new analytics packages can use Hadoop MapReduce and deploy Hadoop HDFS. It is unclear how those capabilities relate to the standard VOD offering.

- ➢ Neo4J recently introduced a solution that replicates some memory caches and files and has standby data servers. Its effectiveness is unknown. There are no published data or service availability commitments.

Customers who require this feature

- ➢ Big Data Analytics application/framework providers.
- ➢ Enterprise systems.
- ➢ Equipment vendors and systems built for the medical and defense markets.

Revenue at risk, or which could be won

- ➢ This capability could revive existing markets, particularly in the equipment space and open new markets in conventional IT shops that are evaluating Big Data solutions. It will also enhance our offerings to the Intelligence Community.

Related Material

We will also need:

- ➢ New Quality Assurance material.
- ➢ Updated documentation, training and web based training.
- ➢ Collateral.

Notes

1. It may be possible to harden the current Objectivity/HA features to provide some or all of the Enhanced Data and Service Availability Package features, but it will probably be more effective to leverage some currently available, mostly Open Source, solutions. For example:

- ➢ Distributed, parallel, high availability file systems, such as ZFS (from Oracle/ Sun), GPFS (IBM), Lustre, Ceph, GlusterFS, Starfish or XtreemFS.
- ➢ Asynchronous, cloud-capable, file replication solutions, such as Aspera – http:// asperasoft.com.
- ➢ High availability cluster middleware, such as Linux HA – http://www.linux-ha.org.
- ➢ Open Source service availability middleware, including virtual server capabilities, such as OpenSaf.

  "The goal of the OpenSAF project is to develop middleware based on open and industry standard interfaces for applications requiring uninterrupted 24x7 service. OpenSAF is actively supported by leading companies in the communications and enterprise computing industries and is focused on becoming the premier SA solution for commercial products and converged applications for all markets." - http://www.opensaf.org/link/linkshow.asp?link_id=151213.

2. Both the Enhanced Data and Service Availability Packages can be developed and sold in modules, e.g.:
    a) A module that incorporates flexible file system capabilities, with support for data redundancy, server load balancing, automated backups and geographic data replication.
    b) Another module could incorporate service availability middleware to provide server standby/failover and capacity expansion capabilities.

3. Any new features should be implemented with awareness of potential security concerns. No system can be resilient if it is open to attack.