Contents

[Project Vision 2](#_Toc517012297)

[Introduction 2](#_Toc517012298)

[Positioning 3](#_Toc517012299)

[Problem Statement 3](#_Toc517012300)

[Product Position Statement 3](#_Toc517012301)

[Stakeholder Descriptions 4](#_Toc517012302)

[Stakeholder Summary 4](#_Toc517012303)

[User Environment 5](#_Toc517012304)

[Example Scenario 1: The On-Prem Password Manager company 5](#_Toc517012305)

[Example Scenario 2: The database company 5](#_Toc517012306)

[Example Scenario 3: The Video Game company 5](#_Toc517012307)

[Product Overview 6](#_Toc517012308)

[Needs and Features (Functional requirements) 6](#_Toc517012309)

[Other Product Requirements (Non-functional requirements) 7](#_Toc517012310)

[Business Justification for Functional and Non-functional Requirements 8](#_Toc517012311)

[Justification of ordering of requirements 9](#_Toc517012312)

# Project Vision

## Introduction

In today’s modern world computing is becoming more fundamental to everything we do, and with that comes a focus on creating highly available and distributed services. With high availability services, many computers work together to make a service achieve the greatest uptime possible, this includes being fault-tolerant to various outage scenarios. Consensus algorithms are a foundational part of building these systems.

Consensus algorithm work by ensuring all nodes in a cluster are in agreeance on the current state of the service and continues uptime of the service during node failure or network link failures between the nodes.

The vision for this project is to create and publish a fully featured implementation of a Consensus Algorithm as an open source code library which would allow for application developers to implement consensus/fault-tolerance into their services as easily as possible.

The benefit to an implementing company could be twofold; firstly, solving the problem of increasing uptime of a critical service, and secondly by taking advantage of consensus distributed consistent log feature they can maintain service uptime whilst offloading server computation to the end user devices, directly saving hosting costs for the company.

*Create a fully featured open source code library to allow the most amount of developers access to adding consensus features as easily as possible to their projects*

## Positioning

### Problem Statement

Consumer driven requirements for always available services has pushed companies to eliminate any downtime from their offerings. These downtimes impact user experience as well as trust in their platform.

When consumers have poor experiences and lose trust in a service they may consider alternatives solutions for their needs. This impacts the growth of a company, as well as their ability to extract the financial benefits of scale from having a large user base.

Companies which minimise or eliminate any impact from downtime for their end users may better compete in the marketplace for their services.

### Product Position Statement

Our proposed project targets developers who are responsible for implementing services which require high uptime. This open source library aims to fill the market position of a consensus algorithm which focuses on ease of integration and which is written in the language the developer is most likely to be using.

The benefits to a company would include reducing overall product development time, reducing the total cost of developing a highly available service, and by reducing the need/costs for consensus specialists to implement the desired feature. Reducing these costs will also lower the barrier for entry for companies which previously could not have reasonably implemented this.

Driving down development overhead costs can not only improve a company’s competitiveness in their market and allow them to capture the cost of unnecessary expenditure, the reduction of development time can also reduce the financial risk that companies accept during their projects which implement this technical feature.

## Stakeholder Descriptions

### Stakeholder Summary

|  |  |  |
| --- | --- | --- |
| **Name** | **Description** | **Responsibility** |
| **Project Team** | Our team; functioning as developers, as well as project sponsors and management | * Ensuring system is maintainable * Ensuring market demand * Ensuring features are working as desired * Ensuring system is reasonably secure * Monitoring of project development progress Investing their time into development |
| **Team Leader/CTO** | The individual who is responsible for selecting and assessing technologies that may be beneficial for inclusion into their services | * Assess the amount of time and cost of implementation of this library as their feature * Confirmation of the benefits of the library matching the business requirements * Comparison of alternative solutions for their requirements * Proposing library as solution for their requirements to project sponsor * Managing the project of implementation |
| **Project Sponsor** | The individual who has requested the inclusion of the feature into their service, and provide the financing for implementation | * Assess the proposal by the Team Leader/CTO in the context of cost to benefit to the company and realization of benefits * Responsible to the business for the success of the project and reasonable use of expenditure * Responsible for sponsoring the resolution of the business case |
| **Developer - Consensus Specialist** | The individual in the Project Team whose responsibility is to learn | * Responsible for the intimate knowledge of the library * Responsible for technical practical knowledge to understand consensus algorithm and implementation * Liaising as the specialist consult inside the Project Team for practical implementation of the library into the service |
| **Developer** | The individuals responsible for the integration of the library into their service | * Understanding their existing service such that they can liaise with the Consensus Specialist to plan how to implement * Implementing the feature into their service |
| **End user** | The end users who is pressuring the company for highly available service | * Not applicable |

## User Environment

As the reliance on always available services grows, there becomes more companies and individuals who rely on their services having high uptime. These companies have sometimes many thousands of users relying on the availability of their services to perform important roles in their lives.

Some examples of the users would be:

* Companies who offload server session to customers, such as multiplayer games.
* Companies offering As A Service (aaS) products needing high uptime
* Companies who want to add high reliability to their services
* Database programmers

### Example Scenario 1: The On-Prem Password Manager company

This company produces a Linux .NET core web server that their business clients run on premises, this server is a Password Manager for all of their employee’s company logins details and supplier logins in. Their biggest customer, a 1000 seat enterprise, has complained of the risk of reliability in this password web server and would like to mitigate the risk of failure of the server as it would stop hundreds of people from conducting their daily duties. They suggested that they were interested in a more reliable solution which mirrors itself between sites to maintain its uptime, is transaction safe, and importantly never loses data or returns incorrect results. On-Prem Password Manager is looking to implement a consensus algorithm at the web server level, and then allow its customers to install multiple instances to achieve this higher level of reliability.

### Example Scenario 2: The database company

Consensus algorithms are complicated, and databases are a critical part of almost all web services. In order to bring consensus to the masses, a database company is looking to implement clustering at the transaction log level between multiple instances of their database. This would allow web servers to failover to other nodes in the cluster if one fails. The job of the developers is greatly simplified as they only need to write heartbeat checks to their database, and upon failure select another node in the cluster.

### Example Scenario 3: The Video Game company

There is a video game company who runs a popular multiplayer turn-based strategy mobile game, many thousands of players are playing at any one time. Analysing their company costs, hosting costs accounts for around 40% of yearly expenditure, which is second only to staff costs. Each game server hosts up to 10 players at a time, meaning the company needs scale up to run hundreds and potentially thousands of these servers during peak playing times. The idea came up about offloading running the server onto one of the players playing in each game, however it was shot down due to the ability for players to leave at any moment causing bad experiences for the other players in the servers. They’ve now had the bright idea to employ a consensus algorithm which synchronises game state among their players, so if someone leaves, the game server starts running on another player. This will dramatically cut hosting costs, while not impacting uptime for the players.

## Product Overview

### Needs and Features (Functional requirements)

Priority scale is between 1 and 12, with 1 being the most important

|  |  |  |  |
| --- | --- | --- | --- |
| **Need** | **Priority** | **Features** | **Planned Release** |
| Consensus between distributed systems | **1** | Replicated log, with consensus algorithm | [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/) |
| Fault tolerant distributed service | **2** | Consensus algorithm allows for a fault tolerant distributed system | [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/) |
| Improved reliability of existing service | **3** | System is fault tolerance, so it will improve reliability | [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/) |
| Complete proven reliability | **4** | Based on proven algorithm | [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/) |
| Minimal additional surface area for failure | **5** | Complete coverage unit testing | Version 1.0 |
| Cross Platform | **6** | Targeting .NET standard framework | [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/) |
| Mitigate project abandonment | **7** | Licensing allow for profit | [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/) |
| Minimal overhead/impact to service performance | **8** | Equivalent to leading consensus algorithm, Paxos in performance | [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/) |
| Minimal resource usage | **9** | Consensus Log compaction | Version 1.0 |
| Ability to attempt to designate a node to run the UAS | **10** | Add method to API to allow for attempting to become leader of the cluster, so as to start UAS | Version 1.0 |
| Upgrade path | **11** | Versioning built in, backwards compatibility minor releases and single major | Final |

### 

### Other Product Requirements (Non-functional requirements)

Priority scale is between 1 and 14, with 1 being the most important

|  |  |  |  |
| --- | --- | --- | --- |
| **Requirements** | **Solution** | **Priority** | **Planned Release** |
| Reliability | Full coverage unit testing | **1** | Final |
| Usability | Designed to be as simple as possible to integrate.  Released as Nuget package | **2** | [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/)  [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/) |
| Documentation | Full coverage documentation for algorithm and API | **3** | Final |
| Quality | Full coverage unit testing  Strict adherence to style guide | **4** | Final  [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/) |
| Performance | Matches Paxos in performance of consensus  Own thread with ASYNC/non-blocking operations  Performance analysis | **5** | [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/)  [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/)  Version 1.0 |
| Compatibility | Written in .NET the second most popular language  Minimal dependencies  Written in .NET standard, cross platform  Designed to be as simple as possible to port languages | **6** | [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/)  [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/)  [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/)  [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/) |
| Availability | Can be run between servers locally or across Internet | **7** | [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/) |
| Security | Network level authentication | **8** | [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/) |
| Privacy | Security measures to join cluster | **9** | [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/) |
| Scalability | Dynamic cluster membership, horizontal scaling | **10** | Version 1.0 |
| Testability | Open source code, unit tests provided | **11** | Final |
| Extendability | Open source code | **12** | Final |
| Auditability | Open source code  Logging | **13** | Final  [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/) |
| Troubleshooting | Verbose logging | **14** | [Prototype](https://bitbucket.org/teamdecided/raftprototype/src/master/) |

### Business Justification for Functional and Non-functional Requirements

First and foremost, the most important feature of this project is increasing the reliability of a User Application Service by maintaining a replicated log; hence the top 5 functional requirements link directly to this. Not only does the desired system require to spread a replicated log amongst nodes, it also needs to be fault tolerant in the case of failure of a minority of nodes and manage the start/stop of a UAS in cases where only one instance should run. An unconventional part of reliability is usability, therefore there is a focus on the design of the usability of the library to ideally eliminate issues to reliability caused during implementation. To reinforce usability for the product, there will also be complete documentation available.

Companies looking to run this software may look to do so on any operating systems (e.g. Windows, Linux and mobile), so it’s important that the code is written in a way that allows for this portability. To implement this, the project will be written in the [.NET Standard Framework](https://msdnshared.blob.core.windows.net/media/2016/09/dotnet-tomorrow.png), which allows portability over all platforms .NET is able to run on.

Part of the benefit of releasing all of the code for this project as open source is that the business risk of project abandonment by ourselves could be picked up by the client, or even a potential open source community. This would enable security updates, bug fixes, and new features to continually be added to the project; and the licensing the project is released under allows for companies to profit from the code.

A fact of implementing a consensus under the hood of a service is that there will always be some overhead; the consensus algorithm being implemented has the same performance as the more popular/complicated and error prone Paxos algorithm. So, using Raft gives us all of the benefits, without the understandability/usability downsides of Paxos. This minimal performance overhead allows companies implementing the library to not suffer from unreasonable latency overhead while waiting for servers to reach consensus, which would’ve caused a direct impact to their user experience. With careful design, and appropriate performance analysis, it’s even be possible to for this library to be used in a real time 30-60 tick game.

As a single instance of the User Application Server (UAS) runs at a time it is important to the user experience that the most ideal node in terms of latency, or hardware performance runs the UAS. Although we’ve removed implementing this feature ourselves from our own scope, we are adding the ability for the user of our library to pick which node runs the UAS. They are able to implement their own fitness calculations, with flap detections for flaky nodes and such if they’d like.

As companies require always on services, an important part is that the consensus library also include backwards compatibility, so allow for rolling software updates through the nodes in a cluster. This allows for the updating of UAS without causing service downtime to the service, avoiding any customer impact.

Security and privacy are also both highly important since these nodes will connect across the network, in a modern-day service there is an expectation of a duty of care to users for the protection of their potentially sensitive information or data.

Part of the ease of code maintainability is facilitating the recreation of any possible bugs in the library; as such, to enable the troubleshooting of this the library will output at multiple debugging levels based on desired configuration.

Implementing scalability in the cluster by enabling the dynamic adding and removing of nodes allows the integration of this library with the horizontal scalability necessary for modern infrastructure design.

### Justification of ordering of requirements

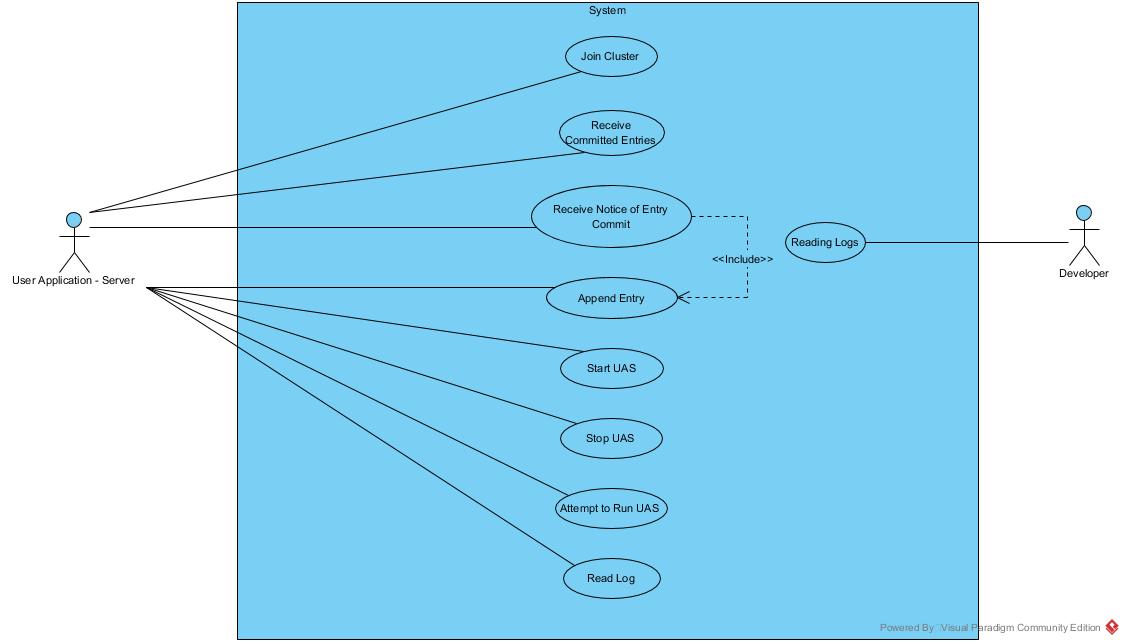
The library being quite atomic in nature (i.e. it either works or it doesn’t when ready for integration) has impacted the ordering of the requirements priorities. They have primarily been focused on functionality first before the extra “nice to haves”, so that the product can get out the door as fast as possible, while never compromising on the reliability. For example, Security is absolutely necessary in a production release of this library as it may be conducted over the Internet (and other publicly known vulnerabilities are exploited in similar software, discussed elsewhere in this document), however it’s technically not required to have a running consensus algorithm.

## 

## Requirement Model

### Use Case Model

After investigation of the requirements the following [Use Case diagram](https://cdn.discordapp.com/attachments/418342556800385048/457780209610129410/unknown.png) has been developed to outlining the initial set of interactions required

.

## Short Use Case Descriptions - In Table Form

### Actor Key

UAS = User Application Server

Dev = Developer

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Use Case Name** | **Precondition** | **Actor** | **Need** | **Do Something** | **Basic Intent/Goal** |
| Join Cluster | Has information about the cluster | UAS | Join the cluster | Join Cluster | Become a member of the cluster |
| Receive Committed Entries | The current node has joined the cluster | UAS | Bring itself up to date | Receive Committed Entries | Can maintain consensus |
| Receive Notice of Entry Commit | The current node has attempted to append an entry | UAS | To know if the message they requested to be committed, was committed | Receive Notice of Entry Commit | So that they may update the running UAS state |
| Append Entry | The current node should be running a UAS | UAS | To add a message into the distributed log and have it protected by consensus | Append Entry | Message gets committed into the consensus |
| Start UAS | The current node was previously not running the UAS | UAS | To be notified when they need to start running the UAS | Start UAS | They may go ahead and start up their UAS to provide services |
| Stop UAS | The current node was previously running the UAS | UAS | To be notified when they need to stop running the UAS | Stop UAS | They stop acting as a UAS and providing service |
| Attempt to Run UAS | The current node currently doesn’t run UAS | UAS | Attempt to designate a specific node to run the UAS | Attempt to Run UAS | The identified nodes will try to become leader and hence be responsible for running the UAS |
| Read Log | The current node has joined the cluster | UAS | Read log entries that have been committed | Read Entry Value | Can reference the committed data |
| Read Logs | At least 1 node has been instantiated | Dev | Understand what the underlying consensus algorithm is doing, perhaps tracking a bug | Read Logs | Understand what the underlying consensus algorithm is doing |

### 

### Domain model

#### 

https://cdn.discordapp.com/attachments/418342556800385048/457782561985527818/unknown.png

### Discussion regarding domain model

We’ve found this is easiest to explain using an example of a distributed multiplayer game running a which is in a server/client setup. In this example the server is also running on one of the clients. In that example, the UAC would be the game client which the player’s would be interacting with, and UAS would be the game server all the game clients are talking back to. Our library would be integrated into the UAS.

Reading from the top left of the diagram it’s shown how the UAC (i.e. game clients) talk to the active UAS (i.e. game server) in the Cluster Application Servers(CAS). Then it’s shown how the CAS is made up of UASs. It’s the library which links these UACs together into a cluster, and the UACs simply talk to whomever is the ‘leader’ UAS in the CAS. The UAC can simple just try each of the IPs in the CAS until they find the leader who responds, this cuts down on the complexity in the UAC software.

Next, the UAS uses the Consensus API to talk to its internal Consensus Node. This consensus node is the part that handles cluster operations, and part of that is instructing the UAS to let it know when it should and shouldn’t be running the active game server. The nodes all communicate to each other through messages which are sent and encrypted by the Networking Library. Consensus between the nodes is through a shared distributed log (DistributedLog) which the active UAS commits entries into, and the follower nodes are updated with. Each node stores information about the other nodes which are in the cluster with it. And each node does verbose logging for troubleshooting reasons.

### Does the analysis of non-functional requirements demonstrate an ability to identify, quantify, prioritise, and communicate required system qualities? (LO2)

This is the part of the document where the above would be analysed, however after discussing with Jim, he has confirmed that this is covered by our “Business Justification for Functional and Non-functional Requirements” section in the Project Vision above.

### Discussion regarding CCRD

The CCRD Use Case is the list of the most important (critical, core, risky, difficulty) Functional/Non-functional requirements, and as per our project milestones, must be implemented into the Prototype. As our project fundamentally is very atomic (i.e. it either works or it doesn’t), we’ve had to include all but a single Use Case (Attempt To Run UAS) into our Prototype to mitigate the risk of implementation issues later on. This means our CCRD Use Case includes 7 of the 11 requirements of the Prototype (As outlined in Project Vision), and as stated above includes all but 1 of our above outlined use cases.