Milestone 03: Research Methodology

Revisions

After re-evaluating the existing research aim, hypothesis, and research questions, it has been concluded that the title will remain as proposed originally. The hypothesis will be slightly modified to exclude the security aspect of the research, due to the vast research implications that it imposes it is not suitable given the time frame.

**Research Hypothesis:**

It is hypothesised that availability **~~and security~~** of will be improve significantly due to containerization.

The research questions remain targeted towards availability improvements using orchestration platforms, however different strategies of orchestration will not be compared in the current study and have been excluded given the available timeframe.

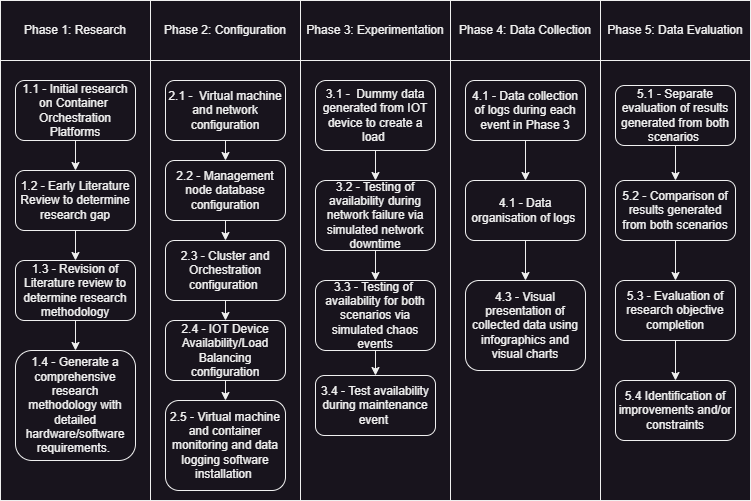
**Research Questions:**

1. How does containerization affect the availability and performance of IoT applications?
2. How do container orchestration platforms contribute to enhancing availability in IoT deployments?

~~What container orchestration strategies can be used to ensure high availability for IoT devices with intermittent or unreliable network conditions?~~

1. How do container orchestration platforms contribute to enhancing availability in IoT with intermittent or unreliable network conditions?

Research Pipeline

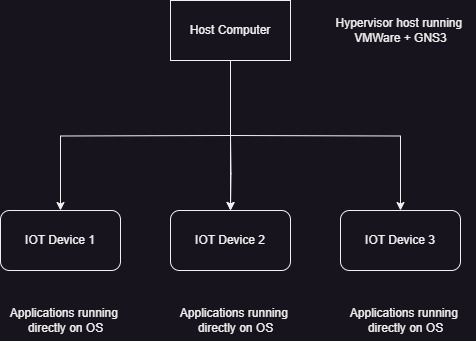
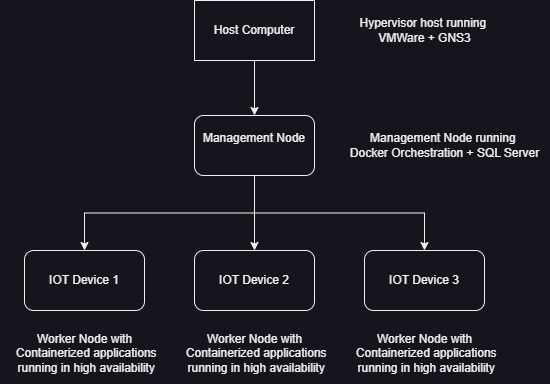


Phase 1

This phase consists of the pre-experimentation research which is required to determine which resources should be used to conduct the experiment with. A literature review is compiled by evaluating similar studies in the field and their respective findings.

Phase 2

This phase consists of implementing resources inside a scenario with configurations according to the desired datasets to be acquired from experimentation. Two scenarios are required for this specific research aim due to the comparative nature of the desired data. Scenario 1 will be used as a baseline for availability, while scenario 2 will provide comparative data with regards to availability through the use of failovers and load balancing features boasted by Container Orchestration Platforms.

 *Scenario 1: Non-Containerized Solution Scenario 2: Containerized Solution*

Phase 3

This phase consists of the experimentation necessary for the research data to be discovered. Various simulations and tests will be conducted while live data is logged to determine the resilience of the systems for both scenarios accordingly. The systems will also be tested under intermittent network to simulate Agricultural IOT devices accordingly.

Phase 4

This phase consists of the collection and organisation of data generated from data logging, it will be strategically sorted to provide a structured comparative analysis for both Containerized and Non-Containerized solutions.

Phase 5

This phase consists of the data collected in Phase 4 being evaluated for the purpose of answering the Research questions initially proposed. Analysis will take place to determine if the research objective has been satisfied, and any improvements or limitations will be identified accordingly.

Research Methodology

Various experiments related to the research objective will be conducted to determine availability variables in both scenarios and provide a cross comparison of the two. Containerized applications with functional IOT capabilities such as data logging and periodic synchronisation will be configured on each virtual worker node to simulate a containerized IOT Device. Both configurations will be handling dummy data related to Agricultural IOT sensors, such as humidity and temperature to simulate a typical working load. Critical network and hardware failures will be simulated to analyse how availability is affected in relation to containerization. Availability will also be tested under simulated maintenance conditions. A tool such as Prometheus is ideal for monitoring the Virtual Machine data and Ansible is a potential tool that may be useful for certain configurations.

The data generated from these experiments is quantitative in nature, as it deals with values such as downtime recovery time, failover delay, load balancing performance, and status, which can all be objectively quantified.