# Plan for Deployment of ElasticSearch to SKY Development Environment

**Author**: Steve Jones

**Date**: 21st January 2019

**Version**: 1.0

This document describes a proposed plan for migration of the existing SKY Project ElasticSearch Development environment from existing limited configuration to a structure which will allow storage of 1 month of data in preparation for an upcoming demonstration.

**NOTE 1**: This document refers to a spreadsheet which has been produced to document the existing and propsed configurations. It has a number of tabs which show the existing hardware available and existing/proposed configurations as well as a matrix of existing and proposed Persistent Volume definitions. The document is called - **SKY-ElasticSearch-Cluster-Kubernetes-Dev-Env.xlsx**

**NOTE 2**: ElasticSearch, Cerebro, Kibana and other utilities are deployed from a set of Helm charts which are currently stored in a private GIT repository named 'smj-kubeapps'. This repository will be zipped and delivered to the deployment server inside the SKY development environment for use deploying the various elements of the architecture. Ultimately the contents of this repository should be owned by the project and stored in a GIT repository private to the project.

**NOTE 3**: The deployment of the Kubernetes environment relies on Docker images which are stored in public DockerHub repositories under the 'smjdockerhub' account. Ultimately these images should be migrated into a private GitLab repository inside the SKY Development environment but since this has not yet happened, access to the internet is needed at the time of deployment so that the images can be fetched.

Key Tasks:-

1. Setup of Steve Jones's PC to be able to access the various systems to enable review of existing configuration. This includes Kubernetes Dashboard, Cerebro, Kibana, access to deployment server, etc...
2. Scaling Up/Out Tests on existing environment to determine if this can be done without impact to existing cluster. Some minimal tests can be attempted on laptop but these will not truly reflect performing the test in the actual environment.
3. Development of script to generate Kubernetes Persistent Volume YAML config files for creation of k8s PVs for the new configuration. This utility will be for the purposes of this deployment only, as manual creation of Persistent Volumes is not the long term plan but until auto creation is set up then this is the only viable approach.
4. Refactoring of existing Kubernetes Helm charts (smj-kubeapps) used to configure existing environment to reflect the new configuration. This will initially create a small cluster with 3 data nodes which will then be scaled dynamically once it is running.
5. Destruction of existing ElasticSearch environment.
6. Creation of required Persistent Volumes in Kubernetes environment to support new ElasticSearch cluster configuration. This requires access to the Deployment server with the **kubectl** and **helm** utilities available.
7. Creation of new ElasticSearch, Cerebro and Kibana environment from the smj-kubeapps Helm charts.
8. Testing and checking of existing environment including test ingest of data via whatever mechanisms are currently in place, e.g. NiFi...
9. Sign-off of new environment.