EIPAAS-5801 Implementation plan(Docker upgrade to 20.10.6 due to cve)

Overview

This guide outlines the plan for the Docker Upgrade to 20.10.6 from 19.03.5 due to CVE-2021-21284.

Kubernetes Checks

- Validate the nodes within the cluster ensuring version and status.

```
kubectl get nodes
```

- Check the current cluster capacity

```
kubectl top nodes
```

- Check Container Deployments
  - Check for failed, evicted, or errored containers.

```
kubectl get pods -A | grep -v "Running\|Completed"
```

  - Review overall cluster deployment.

```
kubectl get all -A
```

  - Collect some basic stats for the Kubernetes cluster as a comparison.

```
for i in $(kubectl api-resources  | awk '(NR>1) {print $1}'); do value=$(kubectl get $i -A 2>/dev
/null | wc -l); if [[ $value -gt 0 ]]; then echo "$i: $value"; fi;done
```

Assumptions

A nKaas cluster has already been deployed.

Hot Fixes

Ensure the following commits are merged into the project branches

Repo	Commit
Control	45c69a4cbdf

Implementation steps

1. From the AWS console, start the Bootstrap VPN and Bootstrap Server and log into the VPN.

2. Browse to https://<url-prefix>bs-go.<domain>
3. Run the **r10k_deploy** pipeline to pull in the latest code.
4. **Image Builds - Dry-Run Mode**
    a. Click into the ADMIN - ENVIRONMENTS - IMAGE_FACTORY environment and ensure IMAGE_BUILD_DRY_RUN is set to true.
    b. Then run the below pipeline(schedule_container_images) to see the changes are reflected with respect to the docker upgrade on the output of the pipelines. Once you are happy with the changes/output then proceed with the below.
    c. Click into the ADMIN - ENVIRONMENTS - IMAGE_FACTORY environment and ensure IMAGE_BUILD_DRY_RUN is set to **false**.
5. **Prepare new IMAGE_SET_ID within Consul KV.**
    a. Update the IMAGE_SET_ID within the IMAGE_FACTORY environment to the new value. Note. Updates to this variable must always be managed at the environment level and not the pipeline level from Pipelines - Pipeline Groups - **configure_platform**, filter on **generate and** click the 'play +' button on the **generate_image_sets**
    b. Select the Environment variables tab and set the SOURCE_IMAGE_SET_ID to the previous IMAGE_SET_ID being used. You should also see the inherited IMAGE_SET_ID variable in the list set to the new value.
    c. Kick off the pipeline which will export all image references in Consul KV under the SOURCE_IMAGE_SET_ID and create a new folder named after the IMAGE_SET_ID in Consul KV with an identical set of entries.
    d. This will automatically trigger a set of downstream IMAGE_SET pipelines that create the image_set references in consul KV based on the images within the same folder.
6. **Build required images for new IMAGE_SET_ID.**
    a. Kick off the **schedule_container_images** build pipeline to build your new images which will update Consul KV under the IMAGE_SET_ID folder.
    b. When complete kick off the affected **image_set** capability pipelines(**caas_container_image_set**) to pick up the newly built image references which in turn updates the image_set references in Consul KV under the IMAGE_SET_ID folder.
    c. From Pipelines - Pipeline Groups - **configure_platform**, filter on **push and** kick off the **push_image_sets** This will push a configuration file to cloud storage containing the image_sets generated for the new IMAGE_SET_ID.
    d. Click into the ADMIN - ENVIRONMENTS - IMAGE_FACTORY environment and ensure IMAGE_BUILD_DRY_RUN is set to **false**.
    e. From Pipelines - Pipeline Groups - **configure_platform**, filter on **backup** and kick off the **backup_bootstrap_server** pipeline.
    f. Disconnect from the VPN and from the AWS console shutdown the Bootstrap VPN and Bootstrap Server.
7. **Pull newly created image_set from cloud to the consul.**
    a. Connect to the Ops VPN and access Rundeck.
    b. Pull down the newly created image_set via the sync job.
8. **Upgrading Docker on servers.**
    a. Cordon all the worker and master nodes in the cluster.
    b. Go to Rundeck, Run **Provision Capability** with the below values.
        i. capability Kubernetes/workers
        ii. image_set newly_created_image_set
        iii. provision 1-plan
        iv. environment_name reuquired_environment
    c. Run the job and see the output it will change the launch configuration with the new AMI but it doesn't affect the currently running instances.
    d. Once the output is fine and proceeds with the Apply.
    e. Increase the ASG for master and worker nodes(It is recommended to have an odd number of members in a cluster. An odd-size cluster tolerates the same number of failures as an even-size cluster but with fewer nodes).
    f. Drain and terminate the worker nodes one by one.
        i. Once the new node is created, make sure it joins the cluster.
        ii. Check if that node is healthy completely.
        iii. Once the node is completely healthy, then go for the second node.
    g. Repeat the same steps for master nodes.
        i. Before you start to master, first check which is etcd master and remove that node at last (The leader can change over time, but too frequent changes can impact the performance of the etcd itself. This can also be a signal of the leader being unstable because of connectivity problems, or maybe etcd has too much load).

```
$ k logs etcd-ip-10-20-5-153.eu-west-1.compute.internal -n kube-system | grep -i leader
raft2022/03/21 13:23:48 INFO: raft.node: 426c2702e23353e2 elected leader 86517905758dfc1b
at term 6

##Here  86517905758dfc1b is master ID.

$ k logs etcd-ip-10-20-5-153.eu-west-1.compute.internal -n kube-system | grep -i cluster
2022-03-21 13:25:02.186547 I | etcdserver/membership: added member 86517905758dfc1b
[https://10.20.5.6:2380] to cluster f01a25b0bc63408a from store
2022-03-21 13:25:02.186553 I | etcdserver/membership: added member 1cda1c5968ca2be7
[https://10.20.4.221:2380] to cluster f01a25b0bc63408a from store
2022-03-21 13:25:02.186559 I | etcdserver/membership: added member 426c2702e23353e2
[https://10.20.5.153:2380] to cluster f01a25b0bc63408a from store

##From above output we can conclude that  86517905758dfc1b(10.20.5.6) is master.
```

        ii. We can drain and terminate 86517905758dfc1b(10.20.5.6) at last.
        iii. Once you drain and terminate any master node.
        iv. Make sure the newly created etcd pod is up and running.
        v. Check if that pod is joined etcd cluster or not by checking the etcd pod logs.

```
$ k logs etcd-ip-10-20-5-153.eu-west-1.compute.internal -n kube-system | grep -i cluster
2022-03-21 13:25:02.186547 I | etcdserver/membership: added member 86517905758dfc1b
[https://10.20.5.6:2380] to cluster f01a25b0bc63408a from store
2022-03-21 13:25:02.186553 I | etcdserver/membership: added member 1cda1c5968ca2be7
[https://10.20.4.221:2380] to cluster f01a25b0bc63408a from store
2022-03-21 13:25:02.186559 I | etcdserver/membership: added member 426c2702e23353e2
[https://10.20.5.153:2380] to cluster f01a25b0bc63408a from store
## with the above info we can confirm that newly created etcd node is joined cluster

$ k logs etcd-ip-10-20-5-153.eu-west-1.compute.internal -n kube-system | grep -i leader
raft2022/03/21 13:23:48 INFO: raft.node: 426c2702e23353e2 elected leader 86517905758dfc1b
at term 6
## with the above info we can confirm that newly created etcd node is elected master.

embed: ready to serve client requests
etcdserver/api/etcdhttp: /health OK (status code 200)
## with the above info we can confirm that newly created etcd node ready to server client
requests.
```

vi. Once the newly created CP node is joined the cluster, then proceeds with the next node.

h. Once all nodes are sorted then modify the ASG back to normal.

Post-implementation Checks

1. Check the updated docker version and services should be up and running
2. Check the cluster status.
3. Check the consul services.
4. Make sure Prometheus/kibana/Grafana are working fine.

Backout plan

1. Cordon all the worker and master nodes in the cluster.
2. Go to Rundeck, Run **Provision Capability** with the below values.
    a. capability Kubernetes/workers
    b. image_set old_image_set
    c. provision 1-plan
    d. environment_name reuquired_environment
3. Run the job and see the output it will change the launch configuration with the old AMI but it doesn't affect the currently running instances.
4. Once the output is fine and proceeds with the Apply.
5. Increase the ASG by one for master and worker nodes.
6. Drain and terminate the worker nodes one by one.
7. Repeat the same steps for master nodes.
8. Once all nodes are sorted then modify the ASG back to normal.

Restore Cluster

nKaaS: Backup and Restore

Risk to service

The current version(**19.03.5)** which we are using on nKaas is impacted with **CVE-2021-21284.**

Service impact

The entire cluster will restart since all the servers(master and workers) are restarting, so there is a small downtime we can expect. And if there are no replicas, that service will get affected for a short amount of time.

Reference

https://etcd.io/docs/v3.3/faq/

https://sysdig.com/blog/monitor-etcd/

Removing Older ETCD Members from ETCD cluster:

Install etcdctl and kubectl on WSL using the below references.

https://craftbakery.dev/kubectl-on-windows-10-machine/

https://computingforgeeks.com/how-to-install-etcd-on-ubuntu-linux/

Using the below command we can get the list of nodes that are in etcd cluster.

```
etcdctl --endpoints=https://${etcd_addr}:2379 --cacert=./etcd/ca.crt --cert=./etcd/healthcheck-client.crt --
key=./etcd/healthcheck-client.key member list

etcdctl --endpoints=https://10.20.4.178:2379 --cacert=./ca.crt --cert=./healthcheck-client.crt --key=.
/healthcheck-client.key member list
```

Note: We can get the above certs from CP nodes in **/etc/kubernetes/pki/etcd/** location.

Using the below command we can remove the older etcd members from the cluster

```
etcdctl --endpoints=https://${etcd_addr}:2379 --cacert=./etcd/ca.crt --cert=./etcd/healthcheck-client.crt --
key=./etcd/healthcheck-client.key member remove $etcd_node_id

etcdctl --endpoints=https://10.20.4.178:2379 --cacert=./ca.crt --cert=./healthcheck-client.crt --key=.
/healthcheck-client.key member remove 930cd531db387944
```

Note

Don't terminate/stop any CP node from AWS Console, if we terminate any CP node from AWS console then the dependencies won't remove from the etcd cluster which cause issues in the future. To terminate any CP/worker node, always use rundeck job(Drain and Terminate) so that it will remove the dependencies by creating an appropriate Kubernetes Job.