Network design and IP subnets

Network design

Design:

- See Logical Environmental Design for the base of the network design.
- Each logical environment has its own subnet or subnets. Reason: this makes e.g. firewalling easier and more user friendly.
- Inside each logical environment the same kind of subnets are used (the same subnet mask). Reason: using the same kind of subnets everywhere avoids confusion.
- The User Management clusters and the Workload clusters are in different subnets. Reason: clear seperation.
- For each User environment the Workload clusters are in the same subnet.
- For Sandbox, Nonprod and Prod there is a seperate NSX-T Tier-0 gateway for every Management and User logical environment. So 5 in total.
- Behind every Tier-0 gateway there is one Tier-1 gateway.
- Behind every Tier-1 gateway there are one or more Segments.
- There is a 1-to-1 relationship between segment and subnet.
- We use a /16 network range for all the networks.
- We divide this /16 network range into multiple /20 network: one for every logical environment.
- For every logical environment, there is:
  - One management cluster. This uses a /23 subnet for the nodes (vms) and a /23 subnet for Cilium Load Balancer IP pools.
  - One or more workload clusters or shared clusters. These clusters share the same /23 subnet for the nodes (vms) and another /23 subnet for the Cilium Load Balancer IP pools. Reason for not giving seperate subnets per cluster: there are not enough /23 subnets for more than 3 workload clusters in an environment. We don't know how many workload clusters there will be. This depends on whether we give out entire clusters to teams and on the IDP design. Probably there is a need for more than 3 such clusters, so share the subnets.
- The Cilium Load Balancer IP addresses are behind Cilium BGP virtual routers and there are no NSX-T Segments for these addresses. These subnets are known on the Tier-0 Gateway, so that they are known and reachable from the CBS core routers.

- On the Tier-0 Gateways, there are BGP peering relationships between the Tier-0 and the Cilium BGP virtual routers. These BGP virtual routers are reachable using the node IP addresses. On the Tier-0 these node IP addresses are configured as possible BGP routers. When a node acts as a BGP router, the relationship with Tier-0 BGP is established.
- DHCP is used in the subnets for nodes to give the nodes their IP addresses.

IP subnets for the logical environments

Rules:

- For every logical environment, we have two names:
    - Short name: a name that is used to associate with all objects in that environment. E.g. a tag of a vm.
    - Friendly name: a longer name that describes the environment.
- The names consists of several parts that describe the environment. This hierarchy is followed:
    - Management or User environment.
    - Sandbox, nonprod or prod environment. Prod is never shortened to e.g. pr, because this make prod stands out.
    - Serialnumber. The experience is that there will be a use for this in the future. E.g. a second management sandbox, a second datacenter, disaster recovery testing.
- In every subnet range (/23) these rules apply for the IP addresses:
    - 1: gateway.
    - 2 - 10: reserved for static IP addresses.
    - 11 - 250: used by DHCP for the node subnets and by Cilium for the Load Balancer IP pools.
    - 251 - 255: reserved for static IP addresses.

Networks for logical environments

Following the above rules and design, we have the following network ranges for the logical environments.

Friendly name	Short name	Network	IP range start	IP range end	Gatewa
management-sandbox-01	mgmt-sb01	10.45.16.0/20	10.45.16.1	10.45.31.254	10.45.1

Friendly name	Short name	Network	IP range start	IP range end	Gatewa
management-prod-01	mgmt-prod01	10.45.32.0/20	10.45.32.1	10.45.47.254	10.45.3
user-sandbox-01	user-sb01	10.45.48.0/20	10.45.48.1	10.45.63.254	10.45.4
user-nonprod-01	user-np01	10.45.64.0/20	10.45.64.1	10.45.79.254	10.45.6
user-prod-01	user-prod01	10.45.80.0/20	10.45.80.1	10.45.95.254	10.45.8
... 11 more ...					

For each of the above /20 network ranges, this applies to the /23 networks:

- There are 8 networks.
- The first 2 networks are used for the management cluster: 1 for the nodes and 1 for Cilium Load balancing IP pool. Names:
  - Nodes: 'k8s-' + short name + '-mgmt-nodes-01'. E.g. 'k8s-mgmt-sb01-mgmt-nodes-01'.
  - Cilium Load balancing IP pool: 'k8s-' + short name + '-mgmt-lbip-01'. E.g. 'k8s-mgmt-sb01-mgmt-lbip-01'.
- The second 2 networks are used for the user/shared clusters: 1 for the nodes and 1 for Cilium. Load balancing IP pool. Names:
  - Nodes: 'k8s-' + short name + '-workload-nodes-01'. E.g. 'k8s-mgmt-sb01-workload-nodes-01'.
  - Cilium Load balancing IP pool: 'k8s-' + short name + '-workload-lbip-01'. E.g. 'k8s-mgmt-sb01-workload-lbip-01'.
- The third and fourth 2 networks are not used right now.

For NSX-T this applies:

- Tier-0: the name is 't0-k8s-' followed by the short name.
- Tier-1: the name is 't1-k8s-' followed by the short name. Lives under Tier-0 with the same name, but staring with t0 and not t1.
- Segments: the name is the same as the network name.

Management sandbox logical environment

From the above, this applies to the management sandbox environment:

- Tier-0. Name: t0-k8s-mgmt-sb01.
- Tier-1. Name: t1-k8s-mgmt-sb01.

Segments, subnets:

Segment	Subnet	Used for
k8s-mgmt-sb01-mgmt-nodes-01	10.45.16.1/23	management sandbox - management cluster - nodes
none	10.45.18.1/23	management sandbox - management cluster - cilium
k8s-mgmt-sb01-workload-nodes-01	10.45.20.1/23	management sandbox - shared/workload clusters - nodes
none	10.45.22.1/23	management sandbox - shared/workload clusters - cilium
not used	10.45.24.1/23	
not used	10.45.26.1/23	
not used	10.45.28.1/23	
not used	10.45.30.1/23	

User sandbox logical environment

From the above, this applies to the management sandbox environment:

- Tier-0. Name: t0-k8s-user-sb01.
- Tier-1. Name: t1-k8s-user-sb01.

Segments, subnets:

Segment	Subnet	Used for
k8s-user-sb01-mgmt-nodes-01	10.45.48.1/23	user sandbox - management cluster - nodes
none	10.45.50.1/23	user sandbox - management cluster - cilium
k8s-user-sb01-workload-nodes-01	10.45.52.1/23	user sandbox - workload clusters - nodes
none	10.45.54.1/23	user sandbox - workload clusters - cilium
not used	10.45.56.1/23	
not used	10.45.58.1/23	
not used	10.45.60.1/23	
not used	10.45.62.1/23	

See kubit-ip-space Excel sheet for the original source of the subnets. This original is the source of truth and if a change is needed, this document has to be changed. After that, this page can be updated with an abstract of this Excel sheet.