Illumio® Adaptive Security Platform® 18.1 PCE Supercluster Reference Implementation

Last Updated: 07/02/2018

Table of Contents

Controlled Document - Not for Public Release

This document must be used only by Illumio® employees. Public distribution outside of Illumio is not allowed without approval of Product Management.

Product Versions

This reference implementation is based on the following hardware and software versions.

Product	Version
Illumio® Adaptive Security Platform® PCE Supercluster	18.1
F5 Global Traffic Manager (GTM)	13.1.02
F5 Load Traffic Manager (LTM)	13.1.02

About Illumio®

Illumio® products and services are built on our patented technologies. For information on Illumio® patents and patent applications, see https://www.illumio.com/patents.

Illumio® Adaptive Security Platform® Training

Illumio® offers a wide yet focused training curriculum for Illumio® Adaptive Security Platform®, from beginning to advanced topics.

To see available courses, log into your Illumio® support account and select the **Training** tab.

Search Knowledge Base and Documentation

For useful short articles about Illumio® Adaptive Security Platform®, log into your Illumio® support account and select the **Knowledge Base** or **Documentation** tabs.

Illumio® Adaptive Security Platform®Support

If you cannot find what you are looking for in this document or the support knowledge base and documentation, contact us at:

- support@illumio.com
- +1-888-631-6354
- +1-408-831-6354

Recommended Skills

Illumio® recommends that you be familiar with the following:

- Your organization's security goals
- Illumio® Adaptive Security Platform®, including basic deployment of the PCE and VEN, the *Supercluster Deployment and Usage* guide, and related documentation.
- General computer system administration, including startup/shutdown, common processes or services
- Linux shell (bash), Windows PowerShell, or both
- TCP/IP networks, including protocols, well-known ports, and the Domain Name System (DNS)
- General knowledge of Global Services Load Balancing (GSLB)
- F5 Networks BIG-IP® Global Traffic Manager (GTM)
- Server-side SSL certificates

How to use this guide

The paper has several major divisions:

- For brief background information on the PCE and GTM, see Background.
- The reference network designs with logical architecture are described in PCE Supercluster Network Design with Global Traffic Manager (GTM).
- Deeper detail is in DNS and GTM configurations, including Step-by-step configuring the GTM and DNS.
- Other Supercluster deployment considerations includes recommendations about health monitoring of the Supercluster via GTM, auto-deployment of GTM configurations, Kerberos with the PCE Supercluster, and other topics.

Related Documentation

Illumio® Adaptive Security Platform® documentation is available from the Support portal.

- *PCE Web Console* guide: working with Illumination®, designing policy, creating labels, and provisioning and administering managed workloads.
- *PCE Deployment* guide: requirements, planning, and installing the Policy Compute Engine (PCE).
- *PCE Operations* guide*: common* operational tasks on the Policy Compute Engine (PCE).
- *Supercluster Deployment and Usage* is available from your Illumio® representative.
- *PCE REST API* guide: Programming Illumio® Adaptive Security Platform®.
- *VEN Deployment* guide*:* installing and activating the Virtual Enforcement Node (VEN) on workloads.
- *VEN Operations* guide*:* administering the Virtual Enforcement Node (VEN) directly on managed workloads.

PCE Supercluster network design with GTM, LTM, and DNS

This paper presents a geographically distributed network architecture for deployment of the Illumio® Adaptive Security Platform® PCE Supercluster in conjunction with regional corporate DNS servers and the F5 BIG-IP Domain Name System (DNS), formerly known as Global Traffic Manager (GTM).

This reference implementation was designed by Illumio® Technical Marketing and has been deployed in the field.

Problem to solve: One purpose of this architecture is to geographically constrain VEN-to-PCE communications with the PCEs in the regional data center closest to the VEN. Another goal is to avoid having to specify the names of those regional PCEs in VEN activation/pairing. A single pairing script can be used on all workloads regardless of location.

- Regardless of geographic region, the pairing commands for all VENs remain virtually identical. The management server specified at activation (the name of the PCE with which to pair) is always the same: a global fully-qualified domain name defined in the Domain Name System (DNS). This name is called a *Supercluster global FQDN*. The Supercluster global FQDN is understood by all members in the Supercluster.
- The GTM/DNS servers resolve the Supercluster global FQDN of the Supercluster to a regional PCE Supercluster member. The VEN pairs with the regional Supercluster member.
- The GTM allows for additional features such as health monitoring of the Supercluster PCEs, name resolution is processed only for PCE nodes that are healthy.
- Optionally, an intermediary layer between the PCE and the GTM comprised of an Application Delivery Controller (ADC), such as the F5 Local Traffic Manager (LTM), can be used to load-balance traffic to the PCE and perform health monitoring instead of the GTM. The LTM can integrate with the GTM to report the availability status of the configured virtual servers allowing the GTM to make intelligent name resolution decisions.

Supercluster reference network implementation -- logical architectures

This section presents several architectures for a hypothetical network of distributed data centers. In addition, this section decomposes the flow of a request from a VEN to pair with its local PCE cluster.

Alternative logical architectures

This paper presents three alternative logical architectures, from simplest to most complex:

1. A Supercluster deployment with a two-tier network implementation:
    a. A Global Traffic Manager (GTM) tier.
    b. The Supercluster tier with a single PCE cluster in each distributed data center.
2. A variation on the first architecture: A Supercluster deployment with a Supercluster deployment with a three-tier network implementation:
    a. A Global Traffic Manager (GTM) tier.
    b. An intervening Load Traffic Manager (LTM) tier.
    c. The Supercluster tier.
3. A variation on the second architecture: a Supercluster deployment with Supercluster deployment with a three-tier network implementation with multiple PCE clusters in distributed data centers:
    a. A Global Traffic Manager (GTM) tier.
    b. A Load Traffic Manager (LTM) tier.
    c. The Supercluster tier with multiple PCE clusters in each distributed data center.

Variations and additional features or functions that can be considered to this design are discussed in Other Supercluster deployment considerations.

Hypothetical distributed data centers

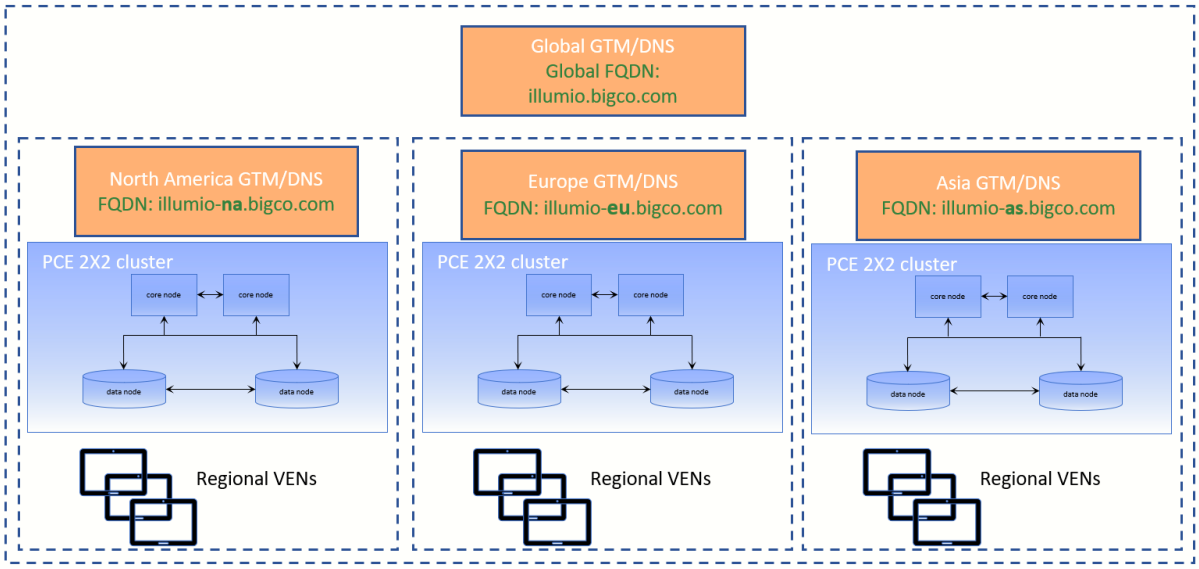
All three architectures imagine data centers for for a fictitious company named BigCo.com with distributed across regions in North America, Europe, and Asia.

- Each regional data center has a single 2x2 PCE cluster that is a Supercluster member.
- The regional Supercluster members are joined with the Supercluster leader.
- Each regional data center has both a GTM and a local DNS server. The GTM relies on the local DNS server to resolve hostnames at the regional level.
- Each regional GTM/DNS has a local FQDN for its PCE cluster:
    - North America: illumio-**na**.bigco.com
    - Europe: illumio-**eu**.bigco.com
    - Asia: illumio-**as**.bigco.com

The PCE `runtime_env.yml` of all the PCEs in the Supercluster optionally includes a supercluster FQDN understood by all the member PCEs of the supercluster. This

Architecture 1 - two-tier Supercluster with F5 GTM

In this diagram the Supercluster global FQDN is illumio.bigco.com. For clarity, this is shown as a separate GTM, but the name is actually defined on all the regional GTMs. All the member of the PCE respond to the Supercluster global FQDN.



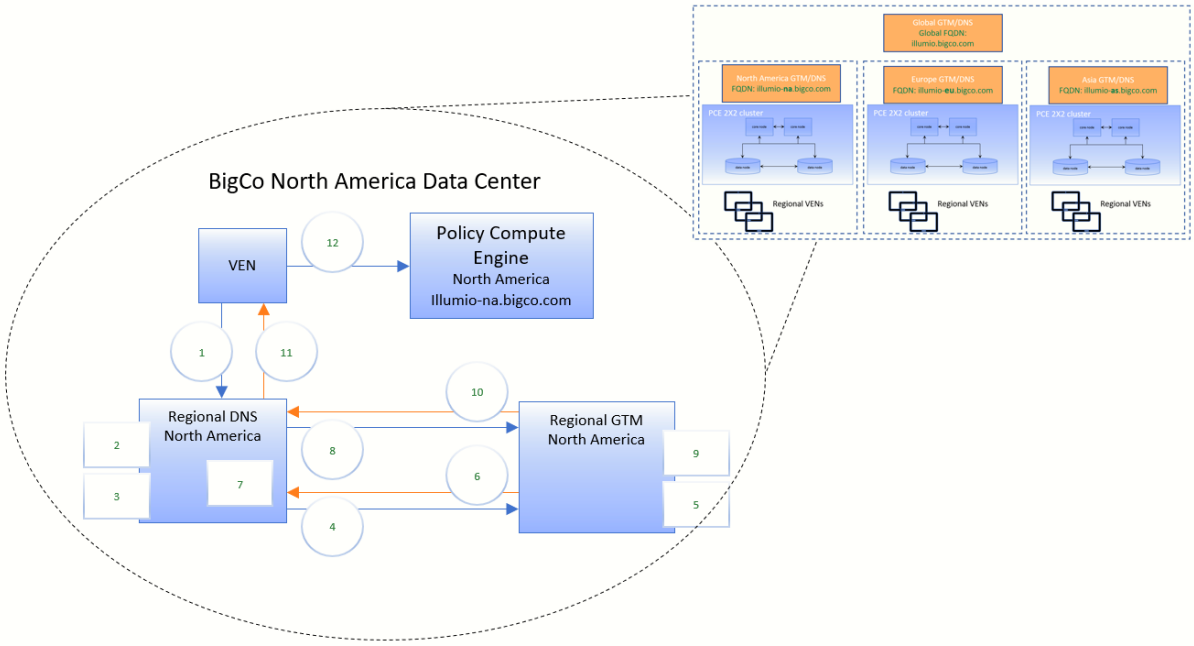
One problem BigCo.com needs to solve is to make the VENs communicate for pairing with their local regional PCE without having to maintain many different scripts or other mechnisms to pair in the various regions.

The solution described here is to program the GTMs *to resolve the Supercluster global FQDN of the Supercluster Leader to the FQDN of the local regional PCE*.

FQDN flow via GTM and DNS

The architecture shown below focuses on the VEN-to-PCE interaction in a single regional data center, North America. The same architecture applies to the other data centers.

Summary of flow: A regional VEN in North America makes a request to pair with the Supercluster global FQDN illumio.bigco.com but is "rerouted" via a series of DNS name changes to the North America regional FQDN of the PCE Supercluster member `illumio-`**`na`**`.bigco.com`. The numbered steps below annotate the lower level interactions between the VEN, the DNS, the GTM, and the Supercluster member.



Step	Description	Sequence of Name Changes	Notes
1	The VEN makes a request to the North America regional DNS server for the IP address of the Supercluster leader global FQDN illumio.bigco.com.	illumio.bigco.com	
2	A DNS CNAME (alias) entry for the global FQDN illumio.bigco.com is illumio.**wip**.bigco.com.		The term *wip* is a naming convention from F5 that signifies the "wide IP" feature managed by the GTM. For more details, see the F5 documentation.
3	The wip.bigco.com subdomain is serviced by the regional GTM.		
4	The original request from the VEN for illumio.bigco.com is changed to illumio.**wip**.bigco.com and sent to the regional GTM	illumio.**wip**.bigco.com	

Step	Description	Sequence of Name Changes	Notes
5	Based on the IP/Subnet of the DNS server and the availability of the local PCE the regional DNS CNAME (alias) for illumio.wip.bigco.com is converted to illumio-**na**.wip-**na**.bigco.com.		
6	The original request from the VEN for illumio.bigco.com is now changed to illumio-**na**.wip-**na**.bigco.com. and sent back to the DNS server	illumio-**na**.wip-**na**.bigco.com	
7	The DNS server is configured so that wip-**na**.bigco.com subdomain is serviced by the **regional** GTM.		
8	A request is made to resolve illumio-**na**.wip-**na**.bigco.com to the Regional GTM.		
9	To distribute the load, the GTM uses round-robin name resolution among the illumio-**na**.bigco.com PCE core nodes, based on PCE health check and availability.		
10	The original request from the VEN for the *global* illumio.bigco.com is changed to the IP of *regional* illumio-**na**.bigco.com core node	illumio-**na**.bigco.com	
11	The DNS response for illumio.bigco.com is now the IP of illumio-**na**.bigco.com.		
12	The VEN receives the IP address of the regional PCE from the DNS server and connects to start pairing.		Regular heartbeats between the VEN and PCE follow the same pattern. Caching on the DNS server of the previous response prevents the full DNS-GTM from being necessary for every heartbeat interval from the VENs.

DNS configuration

This section presents configuration details of the Domain Name System (DNS) for this Supercluster reference implementation:

- Regardless of region, the DNS servers have virtually identical configurations for the Supercluster.
- The DNS redirects Supercluster-specific requests to another device for intelligent name resolution.
- No additional changes are required at the workload level.

DNS zone definition

This is the DNS zone definition for BigCo.com.

In a DNS zone definition, DNS A records define the canonical IP address of a domain name. CNAME records are domain aliases that are equated with those A records. In the configuration below the CNAME resolves the DNS request to a different sub-domain rather than a specific IP address. The sub-domain resolutions are provided by the specified GTMs in the second half of the configuration. This allows the DNS server to proxy the request to the appropriate GTM for more intelligent name resolution.

These definitions assume that all domains are internal to an enterprise: the 10.x.x.x IP address ranges.

This zone definition is referred to in the named.conf configuration file detailed below.

Example DNS zone definition

```
; First half of DNS configuration
$ORIGIN bigco.com.
$TTL 1D
@ IN SOA dns-na.bigco.com. hostmaster.bigco.com.  (
                                        2018022618      ; serial
                                        1D              ; refresh
                                        1H              ; retry
                                        1W              ; expire
                                        3H )            ; minimum

; main domain name servers
                            IN      NS      dns-na.bigco.com.
dns-na                      IN      A       10.1.1.8
illumio.bigco.com.          IN      CNAME   illumio.wip.bigco.com.
illumio-na.bigco.com.       IN      CNAME   illumio-na.wip-na.bigco.com.
illumio-eu.bigco.com.       IN      CNAME   illumio-eu.wip-eu.bigco.com.
illumio-as.bigco.com.       IN      CNAME   illumio-as.wip-as.bigco.com.

; Second half of DNA configuration
```

```
; sub-domain definitions with name servers (gtm) that service the subdomains
; The name server (NS) for the "wip.illumio.com" subdomain should reflect the local GTM
; This top line is expected to differ on each LDNS (local DNS) server.
wip                         IN      NS      big-dns-na
wip-na                      IN      NS      big-dns-na
big-dns-na                  IN      A       10.0.1.20
wip-eu                      IN      NS      big-dns-eu
big-dns-eu                  IN      A       10.1.1.20
wip-as                      IN      NS      big-dns-as
big-dns-as                  IN      A       10.2.1.20
```

named.conf

The `named.conf` file establishes a DNS server as a master, slave, or cache-only name server. It also specifies the zones over which the server has authority and the data files it reads for its initial data.

In the named.conf shown below, the zone bigco.com is included in the configuration, as defined above for the DNS zone.

For more details, see Oracle documentation.

Example named.conf

```
options {
listen-on port 53 { 127.0.0.1; 10.1.1.8; };
listen-on-v6 port 53 { ::1; };
forwarders { 8.8.8.8; 8.8.4.4; };
directory "/var/named";
dump-file "/var/named/data/cache_dump.db";
statistics-file "/var/named/data/named_stats.txt";
memstatistics-file "/var/named/data/named_mem_stats.txt";
allow-query {
  localhost; 10.0.1.0/24; 10.1.1.0/24; 10.2.1.0/24; 10.3.1.0/24;};  #Optional parameter to limit subnets which
can query the DNS server
  recursion yes;
  dnssec-enable yes;
  dnssec-validation yes;
  /* Path to ISC DLV key */
  bindkeys-file "/etc/named.iscdlv.key";
  managed-keys-directory "/var/named/dynamic";
  pid-file "/run/named/named.pid";
  session-keyfile "/run/named/session.key";
};

logging {
  channel default_debug {
    file "data/named.run";
```

```
    severity dynamic;
    };
};
zone "bigco.com" IN {
  type master;
  file "bigco.com.zone";
  notify yes;
  forwarders {
  };
};
zone "." IN {
  type hint;
  file "named.ca";
};

include "/etc/named.rfc1912.zones";
include "/etc/named.root.key";
```

Step-by-step configuring the F5 GTM for first architecture – No F5 LTM

This section presents configuration details of the Global Traffic Manager (GTM) that correspond to this Supercluster reference implementation:

- Regardless of region, the GTM configurations are exactly the same.
- The GTM performs DNS based load-balancing as well as the Health Monitoring

The GTM configuration described in this document relies on a common configuration that is identical on all GTMs. To simplify the implementation, you can configure a BIG-IP synchronization group among the regional GTMs. This allows for configuration changes on one GTM to be synchronized across the sync group.

For an overview, see BIG-IP synchronization group requirements.

For details, see enabling BIG-IP synchronization.

These are step-by-step details for configuring the F5 GTM.

Configure F5 GTM

1. **Create the DNS listeners (UDP and TCP)**
   a. *DNS > Delivery > Listeners > Listener List.*
   b. *Click **Create**.*
   c. Enter a listener name for UDP (for example, **gtm_listener_us-west-2_udp**).
   d. In **Destination Address** field, enter private IP address of the F5 GTM.
   e. For **Protocol** under **Service** section, select **UDP**.

f. Click **Finished**.

2. **Repeat above step to create a listener for TCP protocol.**

3. **Create Data Center Definitions**
   This is a required step in F5 configurations.
   a. DNS > GSLB > Data Centers > Data Center List
   b. Click **Create**
   c. Create different Data Centers for ALL your regional PCE

4. **Create the custom Health Monitors for each PCE core node**
   a. *DNS > GSLB > Monitors.*
   b. Enter a Name
   c. Select Type '**HTTPS**'.
   d. In the **Send String** field, enter:

```
GET /api/v1/node_available HTTP/1.1\r\nHost: fqdn_of_the_PCE_core_node\r\nConnection: Close\r\n\r\n
```

   e. In **Receive String** field, enter:

```
HTTP/1.1 200 OK
```

   f. Alias Service Port: 8443
   g. Click Finished.

5. **Define the Server (GTM Itself)**
   This step is necessary to prevent the monitors from failing.
   a. *DNS > GSLB > Servers > Server List.*
   b. Click **Create**.
   c. Enter a name
   d. *Product:***BIG-IP System**
   e. Select the data center where this server resides
   f. Choose '**bigip**' from available Monitors as the "*Selected*" Health Monitor

6. **Define the Server (PCE Core nodes)**
   a. DNS > GSLB > Servers > Server List.
   b. Enter a name to identify a PCE core node
   c. *Product:* **Generic Host**.
   d. Enter the IP address of the PCE core node
   e. Select the data center where this server resides
   f. Apply the Health Monitor created in **Step 4** specific to the core node
   g. **Define the Virtual Server Resource**
   This step is necessary in order to have the PCE core node as a selectable drop down option in later steps.

* Enter the same name as server for easy identification
* Enter the IP address.
* Select **Finished**

7. **Repeat for each PCE core node in ALL regions**

8. **Define a Pool for each PCE**
Create a Pool for each regional PCE that contains its respective core nodes configured in the previous Step as its member.
a. DNS > GSLB > Pools > Pool List.
b. Click **Create**
c. Enter a Name.
d. Select Type 'A' (e.g. DNS 'A' record)
e. Load Balancing Method
* Preferred: Round Robin
* Alternate: Round Robin
* Fallback: Return to DNS
f. From Virtual Server Drop select and add virtual servers in **Step 6(g)** for all core nodes of the regional PCE.

9. **Create Wide IPs**
Define the FQDNs the GTM will resolution
a. *DNS > GSLB > Wide IPs > Wide IP List.*
b. Select Create
c. Name should be the FQDN defined as the CNAME for the regional PCE FQDN in the DNS server's zone file
d. Select Type 'A',
e. Under 'Pools' select the pool created for the regional PCE.
* There should be one pool, with the regional PCE cores nodes as its members.
f. Select Add

10. **Create the CNAME pools**
These will be used by the Supercluster global FQDN to determine which PCE to redirect the name resolution.
Previous steps need to be completed before you can proceed to this step.
a. **DNS > GSLB > Pool List**
b. Create a new Pool.
c. Enter a name that will map a Wide IP created in the previous step
d. Select Type CNAME
e. Load Balancing Method
* Preferred: Round Robin
* Alternate: Round Robin
* Fallback: Return to DNS
f. From Wide IP dropdown, select a corresponding Wide IP created in **Step 9**
* Each CNAME Pool will have only one member corresponding to an already created Wide IP
g. Select 'Add'.
h. Click Finished.

11. **Create the supercluster Wide IP**
    a. *DNS > GSLB > Wide IPs > Wide IP List.*
    b. Select Create
    c. Name should be the FQDN defined as the CNAME for the Supercluster global FQDN in the DNS server's zone file
    d. Select Type 'A'
    e. Load Balancing Method: **Topology**
    f. From the 'Pool List' Select and add ALL the CNAME pools configured the previous step
    g. Click Finished

12. **Define the Topology Region**
    a. Create a Region
    b. Use a name representative the PCE region
    c. Member Type:
    You will be adding at least to characteristics to define the Region
    * IP Subnet 'is'
        i. use subnet/IP address of the DNS servers in the region
        ii. Click Add
    * Pool 'is'
        i. Use the predefined PCE pool created in **Step 8** that corresponds to the Region being defined.
        ii. Click Add
    * Pool 'is'
        i. Use the predefined CNAME pool created in **Step 10** that corresponds Region being defined
        ii. Click Add
    d. Repeat for each PCE region

13. **Define the Topology Record**
    The topology Record is defined in this way to allow for automatic failover in the future.
    a. Request source
        * Region is <region>
    b. Destination
        * Region is <same region>
    c. Weight:
        * 100
    d. Repeat for each Region

14. **Failover**
    In the future we may append this guide to include dynamic and deterministic failover based on additional topology record definitions.
    Today, we are highly recommending PCE failover/failback require human intervention.
    If a PCE fails change the Last Resort Pool under the regional Wide IP to pool representing non-local PCE core nodes.
    a. DNS > GLSB > Wide IPs > Wide IPs list > {select regional Wide IP} > Pools > Last Resort Pool
    b. Select an available 'A' type non-local pool
    c. Select Update.

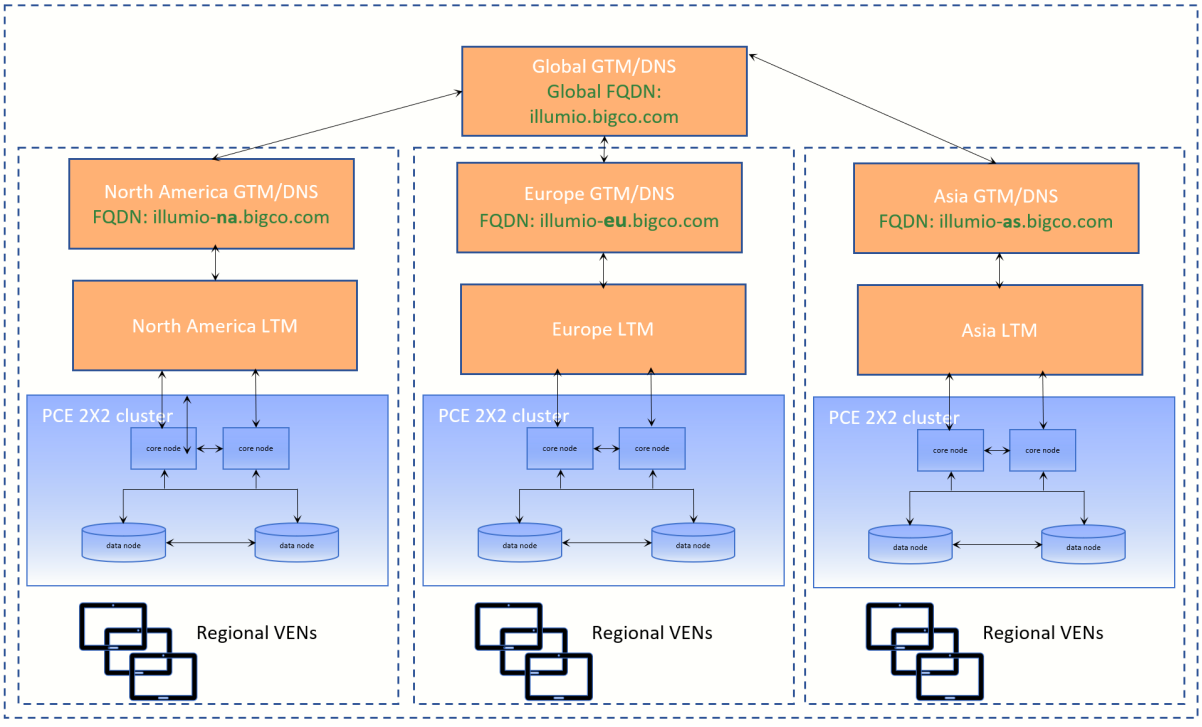
15. (a) Repeat above steps to configure F5 GTM instances in all other regions where Supercluster PCEs reside.
    (b) Optionally, enable 'Configuration Synchronization' in the GTM

Architecture 2 – three-tier with F5 GTM and F5 LTM integration

The GTM is fully capable of monitoring and handling the load-balancing of traffic between the PCE core nodes in a supercluster. Operationally, organizations choose that the individual PCE load-balancing and monitoring be performed by an Application Delivery Controller (ADC), commonly referred to as load balancer, such as the F5 LTM. The GTM configuration monitors the availability of the LTM PCE virtual-servers to indirectly get the health status of the physical PCE core nodes. This section describes the two-tier GTM-to- LTM relationship and the necessary step-by-step configurations.

This GTM configuration differs from the configuration of the two-tier architecture. The two-tier architecture does not have an intermediate LTM between the GTM and the PCE.

As in the two-tier architecture the Supercluster global FQDN is defined on the regional GTMs. This logical architecture shows a separate GTM, which in the physical architecture does not exist.



Step-by-step configuration of F5 LTM

These are step-by-step details for configuring the F5 LTM.

Configure F5 LTM (single PCE per region)

1. **Create the custom Health Monitor**
   a. Local Traffic > *Monitors*.
   b. Select Create
   c. Enter a Name
   d. Select Type 'HTTPS'.
   d. In the Send String field, enter:

      GET /api/v1/node_available HTTP/1.1\\r\\nHost: *replace_with_domain_dot_com***\\r\\n**

   e. In Receive String field, enter:

      200 OK

   f. Alias Service Port: 8443* (Note: Review Separating the Northbound and Southbound API Ports)
   g. Select Finished

2. **Define the Nodes**
   a. Local Traffic > Nodes
   b. Select Create
   c. Enter a Name to reference a local PCE Core node
   d. Enter the Address of the PCE Core node
   e. Repeat of each Core Node
   f. Select Finished

3. **Define the Pool**
   a. Local Traffic > Pools
   b. Enter a Name for the Pool which collectively represents the local PCE core nodes
   c. Apply the Custom Health Monitor created in step 1
   d. Load Balancing: Round Robin
   e. Select Node List
   f. Add each of the Nodes created in Step 2, Service Port: "* All Services"

4. **Define the Virtual Servers**
   a. Local Traffic > Virtual Servers
   b. Select Create
   c. Create PCE specific Virtual Servers
      - *PCE-8443*
        i. Name: **PCE-8443**
        ii. Destination Address/Mask: **Virtual IP**

iii. Service Port: **8443/HTTPS**
iv. Protocol: **TCP**
v. Source Address Translation: **AutoMap** (or SNAT to have greater control of the source IP)
vi. Default Pool: ***{Pool Created in Step 3}***

- *PCE-8444*
  i. Name: **PCE-8444**
  ii. Destination Address/Mask: **Virtual IP**
  iii. Service Port: **8444/HTTPS**
  iv. Protocol: **TCP**
  v. Source Address Translation: **AutoMap** (or SNAT to have greater control of the source IP)
  vi. Default Pool: ***{Pool Created in Step 3}***

- *PCE-9443***** (optional, use if separating the Northbound port to 9443)

  i. Name: **PCE-9443**
  ii. Destination Address/Mask: **Virtual IP**
  iii. Service Port: 9**443/HTTPS**
  iv. Protocol: **TCP**
  v. Source Address Translation: **AutoMap** (or SNAT to have greater control of the source IP)
  vi. Default Pool: ***{Pool Created in Step 3}***

d. Create Supercluster Specific Virtual Servers

- *Supercluster-5432 (Database Synchronization)*
  i. Name: supercluster**-5432**
  ii. Destination Address/Mask: **Virtual IP**
  iii. Service Port: 5432**/Other**
  iv. Protocol: **TCP**
  v. Source Address Translation: **AutoMap** (or SNAT to have greater control of the source IP)
  vi. Default Pool: ***{Pool Created in Step 3}***
- *Supercluster-8300-tcp (Service Discovery)*
  i. Name: **supercluster-8300-tcp**
  ii. Destination Address/Mask: **Virtual IP**
  iii. Service Port: 8300**/Other**
  iv. Protocol: **TCP**
  v. Source Address Translation: **AutoMap** (or SNAT to have greater control of the source IP)
  vi. Default Pool: ***{Pool Created in Step 3}***

- *Supercluster-8302-tcp (Service Discovery: Intra-Cluster)*

  i. Name: **supercluster-8302-tcp**
  ii. Destination Address/Mask: **Virtual IP**
  iii. Service Port: 8302**/Other**
  iv. Protocol: **TCP**
  v. Source Address Translation: **AutoMap** (or SNAT to have greater control of the source IP)
  vi. Default Pool: ***{Pool Created in Step 3}***
- *Supercluster-8302-udp* (Service Discovery: Intra-Cluster Alternative)
  i. Name: **supercluster-8302-udp**

ii. Destination Address/Mask: **Virtual IP**
iii. Service Port: 8302**/Other**
iv. Protocol: **UDP**
v. Source Address Translation: **AutoMap** (or SNAT to have greater control of the source IP)
vi. Default Pool: ***{Pool Created in Step 3}***

Step-by-step configuration of F5 GTM with F5 LTM integration

These are step-by-step details for configuring the F5 GTM with an integrated LTM.

Configure F5 GTM with F5 LTM (single PCE per region) integration

1. **Run the bigip_add script in the GTM to add the LTM**
   a. From the GTM cli "run gtm bigip_add -a <username>@<ip_addr of LTM>"
      - Please refer to https://support.f5.com/csp/article/K14495 for additional information

2. **Add each F5 LTM to the F5 GTM**
   a. DNS > GSLB > Servers
   b. Select Create
   c. Enter a Name
   d. Product: BIG-IP System
   e. Data Center: { create new or select existing, different for each TLM}
   f. Devices > Add
      - Enter Device Name
      - Enter Address
      - Select Add
   g. Health Monitors: bigip
   h. Resources: Virtual Server Discovery: Enabled
   i. DNS > GSLB > Server List > {select Server} > Virtual Servers
      - Select Update
        - This will automatically populate the virtual servers configured on the LTMs in the GTM
        - This informs the GTM of the availability of those services based on Health Checks performed by the LTM

3. **Define a Pool for each LTM front-ending the PCE**
   Create a Pool for each regional PCE that contains its respective core nodes configured in the previous Step as its member.
   a. DNS > GSLB > Pools > Pool List.
   b. Click **Create**
   c. Enter a Name.
   d. Select Type 'A' (e.g. DNS 'A' record)
   e. Load Balancing Method
      * Preferred: Round Robin
      * Alternate: Round Robin
      * Fallback: Return to DNS

f. From Virtual Server drop-down select the regional virtual server running on port 8443, populated from Step 2.

4. **Repeat Step 3 in each region**

5. **Create Wide IPs**
   Define the FQDNs the GTM will resolution
   a. *DNS > GSLB > Wide IPs > Wide IP List.*
   b. Select Create
   c. Name should be the FQDN defined as the CNAME for the regional PCE FQDN in the DNS server's zone file
   d. Select Type 'A',
   e. Under 'Pools' select the pool created for the regional PCE.
      * There should be one pool, with the regional LTM virtual server on 8443 as its members.
   f. Select Add

6. **Create the CNAME pools**
   These will be used by the Supercluster global FQDN to determine which PCE to redirect the name resolution.
   Previous steps need to be completed before you can proceed to this step.
   a. **DNS > GSLB > Pool List**
   b. Create a new Pool.
   c. Enter a name that will map to a Wide IP created in the previous step
   d. Select Type CNAME
   e. Load Balancing Method
      * Preferred: Topology
      * Alternate: Return to DNS
      * Fallback: None
   f. From Wide IP dropdown, select ALL of the Wide IPs created in Step 4
      * Each CNAME Pool will have multiple members corresponding to already created Wide IPs
   g. Select 'Add'.
   h. Click Finished.

7. **Create the supercluster Wide IP**
   a. *DNS > GSLB > Wide IPs > Wide IP List.*
   b. Select Create
   c. Name should be the FQDN defined as the CNAME for the Supercluster global FQDN in the DNS server's zone file
   d. Select Type 'A'
   e. Load Balancing Method: **Topology**
   f. From the 'Pool List' Select and add ALL the CNAME pools configured the previous step
   g. Click Finished

8. **Define the Topology Region**
   a. Create a Region
   b. Use a name representative the PCE region
   c. Member Type:
   You will be adding at least to characteristics to define the Region
      * IP Subnet 'is'

i. use subnet/IP address of the DNS servers in the region
ii. Click Add
* Pool 'is'
i. Use the predefined LTM pool created in **Step 2** that corresponds to the Region being defined.
ii. Click Add
* Pool 'is'
i. Use the predefined CNAME pool created in **Step 5** that corresponds Region being defined
ii. Click Add
d. Repeat for each PCE region

9. **Define the Topology Record**
The topology Record is defined in this way to allow for automatic failover in the future.
a. Request source
* Region is <region>
b. Destination
* Region is <same region>
c. Weight:
* 100
d. Repeat for each Region

10. **Failover**
In the future we may append this guide to include dynamic and deterministic failover based on additional topology record definitions.
Today, we are highly recommending PCE failover/failback require human intervention.
If a PCE fails, change the Last Resort Pool under the regional Wide IP to a pool representing a LTM front-ending non-local PCE core nodes.
a. DNS > GLSB > Wide IPs > Wide IPs list > {select regional Wide IP} > Pools > Last Resort Pool
b. Select an available 'A' type non-local pool
c. Select Update.

11. (a) Repeat above steps to configure F5 GTM instances in all other regions where Supercluster PCEs reside.
(b) Optionally, enable 'Configuration Synchronization' in the GTM

Architecture 3 – three-tier architecture with GTM, LTM, and multiple PCEs in a single region

A key assumption in this reference architecture thus far has been each regional site has VEN managed workloads within the capacity limits of a single PCE. A region is largely defined by the local DNS, rather than geographically. The GTM does not see the client source address in the forwarded requests from the local DNS, only the local DNS IP address itself. For the GTM to make topology-based decisions in this architecture, the local DNS servers in each region are expected to be different. If your regional workload capacity requirements exceed the capacity of a single regional PCE, an additional regional PCE can be added, provided the following conditions are met:

- ADCs, such as the F5 LTM, are required to front-end the PCEs in the same region.

        - The ADCs can make load balance decision based on the source IP of the client.
        - Stickiness while load-balancing VEN traffic to a specific PCE is paramount to maintaining consistency in reporting and other unintended consequences of constant flipping back and forth between two regional PCEs
    - The API ports for the REST API and the PCE web console must be separated.
        - All PCEs in the Supercluster must use the same port for VEN-to-PCE communications.
        - VEN load-balancing should only be applied to the virtual servers for ports 8443 and 8444.
        - The northbound port will continue to be used by the supercluster for regular communication.
    - The configuration represented here assumes there is no preference as to which PCE in the region the VEN establishes a persistent relationship.

In this architecture:

1. There are two regional PCEs:
    - US-PCE-A
    - US-PCE-B
2. Each PCE has its own ADC with one VIP for each ADC
3. The GTM resolves US-PCE as part of the Supercluster global FQDN and round robins between the two ADCs VIPs
4. Each ADC has two pools representing the core nodes of each regional PCE:
    - Pool-US-PCE-A
    - Pool-US-PCE-B
5. An iRule is applied to the Virtual Servers on port 8443 and 8444 to direct traffic based on the source subnet:
    - Even subnets are directed to Pool-US_PCE-A
    - Odd subnets are directed to Pool-US_PCE-B
6. Traffic for other Virtual Servers are are resolved as usual to the individual PCE FDQNs.

DNS zone for three-tier architecture with multiple regional PCEs

DNS multiple PCEs per Region

```
$ORIGIN bigco.com.
$TTL 1D
@ IN SOA dns-na.bigco.com. hostmaster.bigco.com.  (
                                        2018022618      ; serial
                                        1D              ; refresh
                                        1H              ; retry
                                        1W              ; expire
                                        3H )            ; minimum

; main domain name servers
                            IN      NS      dns-na.bigco.com.
dns-na                      IN      A       10.1.1.8
illumio.bigco.com.          IN      CNAME   illumio.wip.bigco.com.
illumio-na.bigco.com.       IN      CNAME   illumio-na.wip-na.bigco.com.
```

```
illumio-na-A.bigco.com.                IN                CNAME        illumio-na-A.wip-na.bigco.com.
illumio-na-B.bigco.com.                IN                CNAME        illumio-na-B.wip-na.bigco.com.
illumio-eu.bigco.com.       IN      CNAME   illumio-eu.wip-eu.bigco.com.
illumio-as.bigco.com.       IN      CNAME   illumio-as.wip-as.bigco.com.

; sub-domain definitions with name servers (gtm) that service the subdomains
; The name server (NS) for the "wip.illumio.com" subdomain should reflect the local GTM
; This top line is expected to differ on each LDNS (local DNS) server.
wip                                                IN      NS      big-dns-na
wip-na                                            IN      NS      big-dns-na
big-dns-na                                       IN                A                10.0.1.20
wip-eu                                    IN      NS      big-dns-eu
big-dns-eu                                    IN      A       10.1.1.20
wip-as                                            IN                NS                big-dns-as
big-dns-as                                    IN      A       10.2.1.20
```

Step-by-step configuration of F5 LTM with multiple PCEs per region(?? Anand)

These are step-by-step details for configuring the F5 LTM in the sub-regions.

> F5 LTMs with single PCEs per region should be configured following the instructions provided in the Architecture 2 section.

Configure F5 LTM (Multiple PCEs in Sub-Regions)

1. **Create the custom Health Monitor**
   a. Local Traffic > *Monitors*.
   b. Select Create
   c. Enter a Name
   d. Select Type 'HTTPS'.
   d. In the Send String field, enter:

      GET /api/v1/node_available HTTP/1.1\\r\\nHost: *replace_with_domain_dot_com***\\r\\n**

   e. In Receive String field, enter:

      200 OK

   f. Alias Service Port: 8443* (Note: Review Separating the Northbound and Southbound API Ports)
   g. Select Finished

2. **Define the Nodes of regional PCE-A**
   a. Local Traffic > Nodes
   b. Select Create
   c. Enter a Name to reference PCE-A Core node
   d. Enter the Address of the PCE-A Core node
   e. Repeat of each Core Node
   f. Select Finished

3. **Define the Nodes of regional PCE-B**
   a. Local Traffic > Nodes
   b. Select Create
   c. Enter a Name to reference PCE-B Core node
   d. Enter the Address of the PCE-B Core node
   e. Repeat of each Core Node
   f. Select Finished

4. **Define the Pool for the regional PCE-A core nodes**
   a. Local Traffic > Pools
   b. Enter a Name for the Pool which collectively represents the regional PCE-A core nodes
   c. Apply the Custom Health Monitor created in step 1
   d. Load Balancing: Round Robin
   e. Select Node List
   f. Add each of the Nodes created in Step 2, Service Port: "* All Services"

5. **Define the Pool for the regional PCE-B core nodes**
   a. Local Traffic > Pools
   b. Enter a Name for the Pool which collectively represents the regional PCE-B core nodes
   c. Apply the Custom Health Monitor created in step 1
   d. Load Balancing: Round Robin
   e. Select Node List
   f. Add each of the Nodes created in Step 2, Service Port: "* All Services"

6. **Define an Even/Odd subnet iRule**
   a. Local Traffic > iRules > iRules List
   b. Select Create
   c. Enter a Name
   d. Enter the iRule

iRule - Even/Odd subnet

```
when CLIENT_ACCEPTED {
     set thirdoctet [lindex [split [IP::client_addr] "."] 2]
     if { ($thirdoctet%2 == 0) } {
          pool PCE-A
     }
     else {
          pool PCE-B
```

```
    }
}
```

e. Select Finished

7. **Define the Virtual Servers**
   a. Local Traffic > Virtual Servers
   b. Select Create
   c. Create PCE specific Virtual Servers
      - *PCE-8443*
        i. Name: **PCE-8443**
        ii. Destination Address/Mask: **Virtual IP**
        iii. Service Port: **8443/HTTPS**
        iv. Protocol: **TCP**
        v. Source Address Translation: **AutoMap** (or SNAT to have greater control of the source IP)
        vi. iRules: ***{Enable iRule created in Step 6}***

      - *PCE-8444*
        i. Name: **PCE-8444**
        ii. Destination Address/Mask: **Virtual IP**
        iii. Service Port: **8444/HTTPS**
        iv. Protocol: **TCP**
        v. Source Address Translation: **AutoMap** (or SNAT to have greater control of the source IP)
        vi. iRules: ***{Enable iRule created in Step 6}***

      - *PCE-9443** (**Required**: port 9443 can be any number, but must be same on all PCEs in the supercluster)

        i. Name: **PCE-9443**
        ii. Destination Address/Mask: **Virtual IP**
        iii. Service Port: 9**443/HTTPS**
        iv. Protocol: **TCP**
        v. Source Address Translation: **AutoMap** (or SNAT to have greater control of the source IP)
        vi. Default Pool: ***{Pool Created in Step 2}***

   d. Create Supercluster Specific Virtual Servers

      - *Supercluster-5432 (Database Synchronization)*
        i. Name: supercluster**-5432**
        ii. Destination Address/Mask: **Virtual IP**
        iii. Service Port: 5432**/Other**
        iv. Protocol: **TCP**
        v. Source Address Translation: **AutoMap** (or SNAT to have greater control of the source IP)
        vi. Default Pool: ***{Pool Created in Step 2}***
      - *Supercluster-8300-tcp (Service Discovery)*
        i. Name: **supercluster-8300-tcp**
        ii. Destination Address/Mask: **Virtual IP**
        iii. Service Port: 8300**/Other**

iv. Protocol: **TCP**
v. Source Address Translation: **AutoMap** (or SNAT to have greater control of the source IP)
vi. Default Pool: ***{Pool Created in Step 2}***

- *Supercluster-8302-tcp (Service Discovery: Intra-Cluster)*

i. Name: **supercluster-8302-tcp**
ii. Destination Address/Mask: **Virtual IP**
iii. Service Port: 8302**/Other**
iv. Protocol: **TCP**
v. Source Address Translation: **AutoMap** (or SNAT to have greater control of the source IP)
vi. Default Pool: ***{Pool Created in Step 2}***

- *Supercluster-8302-udp* (Service Discovery: Intra-Cluster Alternative)
  i. Name: **supercluster-8302-udp**
  ii. Destination Address/Mask: **Virtual IP**
  iii. Service Port: 8302**/Other**
  iv. Protocol: **UDP**
  v. Source Address Translation: **AutoMap** (or SNAT to have greater control of the source IP)
  vi. Default Pool: ***{Pool Created in Step 2}***

Step-by-step configuration of F5 GTM and F5 LTM with multiple PCEs in the same region(??Anand) <WIP>

These are step-by-step details for configuring the F5 GTM.

Configure F5 GTM with F5 LTM (Single Region & Multi Sub-Region model)

1. **Run the bigip_add script in the GTM to add the LTM**
   a. From the GTM cli "run gtm bigip_add -a <username>@<ip_addr of LTM>"
   - Refer to this knowledge base article for additional information

2. **Add each F5 LTM to the F5 GTM**
   a. DNS > GSLB > Servers
   b. Select Create
   c. Enter a Name
   d. Product: BIG-IP System
   e. Data Center: { create new or select existing, different for each TLM}
   f. Devices > Add
   - Enter Device Name
   - Enter Address
   - Select Add

   g. Health Monitors: bigip
   h. Resources: Virtual Server Discovery: Enabled
   i. DNS > GSLB > Server List > {select Server} > Virtual Servers
   - Select Update

- This will automatically populate the virtual servers configured on the LTMs in the GTM
- This informs the GTM of the availability of those services based on Health Checks performed by the LTM

3. **Define a Pool for each region and sub-region**
   a. DNS > GSLB > Pools > Pool List.
   b. Click **Create**
   c. Enter a Name.
   d. Select Type 'A' (e.g. DNS 'A' record)
   e. Load Balancing Method
      * Preferred: Round Robin
      * Alternate: Round Robin
      * Fallback: Return to DNS
   g. For each region **without** sub-regions, from the Virtual Server drop-down, select the regional virtual server running on port 8443, populated from Step 2.
   h. For each sub-region, from Virtual Server drop-down, select the sub-region virtual server running on port 8443, populated from Step 2.
   i. For each region with sub-regions, from the Virtual Server drop-down, select the sub-region virtual servers on port 8433 populated from Step 2.
      - These ALL the same virtual servers added in step 3(h), i.e. multiple virtual-servers.

4. **Repeat Step 3 in each region**

5. **Create Wide IPs**
   Define the FQDNs the GTM will resolution
   a. *DNS > GSLB > Wide IPs > Wide IP List.*
   b. Select Create
   c. Name should be the FQDN defined as the CNAME for the PCE FQDNs in the DNS server's zone file
   d. Select Type ‘A’,
   e. Under 'Pools' select the pool created for the regional PCE
      - Note: the regional PCE with sub-regions will also only have one pool, but this pool will include both sub-regional LTM VIPs
   f. Select Add
   g. repeat for the sub-regional PCEs
      - Note: the sub-regional PCEs will also have only one pool, but this pool will only include the sub-region specific LTM VIPs
      - This step is taken so that the sub-regional PCEs can also individually be resolved by the GTM for other supercluster functionality unrelated to the VEN-PCE communication.

6. **Create the CNAME pools for the Regional PCEs**
   These will be used by the Supercluster global FQDN to determine which PCE to redirect the name resolution.
   Previous steps need to be completed before you can proceed to this step.
   The Sub-regions should not be involved in this step
   a. **DNS > GSLB > Pool List**
   b. Create a new Pool.
   c. Enter a name that will map to a Wide IP created in the previous step for the regional PCEs
   d. Select Type CNAME
   e. Load Balancing Method
      * Preferred: Topology
      * Alternate: Return to DNS
      * Fallback: None
   f. From Wide IP dropdown, select ALL of the Regional Wide IPs created in Step 5
      - Each CNAME Pool will have multiple members corresponding to already created Wide IPs
   g. Select 'Add'.
   h. Select Finished.

7. **Create the supercluster Wide IP**
   a. *DNS > GSLB > Wide IPs > Wide IP List.*
   b. Select Create
   c. Name should be the FQDN defined as the CNAME for the Supercluster global FQDN in the DNS server's zone file
   d. Select Type 'A'
   e. Load Balancing Method: **Topology**
   f. From the 'Pool List' Select and add ALL the CNAME pools configured the previous step
   g. Select Finished

8. **Define the Topology Region**
   a. Create a Region
   b. Use a name representative the PCE region
   c. Member Type:
   You will be adding at least to characteristics to define the Region
      * IP Subnet 'is'
         i. use subnet/IP address of the DNS servers in the region
         ii.Select
      * Pool 'is'
         i. Use the predefined LTM pool created in **Step 2** that corresponds to the Region being defined and if applicable the LTM pools of sub-Regions as well.
         ii. Select
      * Pool 'is'
         i. Use the predefined CNAME pool created in **Step 6** that corresponds Region being defined
         ii. Click Add
   d. Repeat for each PCE region

9. **Define the Topology Record**
   The topology Record is defined in this way to allow for automatic failover in the future.
   a. Request source
      * Region is <region>

b. Destination
* Region is <same region>
c. Weight:
* 100
d. Repeat for each Region

10. **Failover**
In the future we may append this guide to include dynamic and deterministic failover based on additional topology record definitions.
Today, we are highly recommending PCE failover/failback require human intervention.
If a PCE fails, change the Last Resort Pool under the regional Wide IP to a pool representing a LTM front-ending non-local PCE core nodes.
a. DNS > GLSB > Wide IPs > Wide IPs list > {select regional Wide IP} > Pools > Last Resort Pool
b. Select an available 'A' type non-local pool
c. Select Update.

11. (a) Repeat above steps to configure F5 GTM instances in all other regions where Supercluster PCEs reside.
(b) Optionally, enable 'Configuration Synchronization' in the GTM

Other Supercluster deployment considerations

Supercluster SSL certificates and Subject Alternative Name (SAN)

In addition to the SSL certificates that secure communication among the nodes of a single PCE cluster, the members of the Supercluster also require valid certificates for the Supercluster global FQDN. Because there is no optional configuration to specify the use of a separate certificate for the Supercluster global FQDN, consider using a certificate with Subject Alternative Names (SANs). This certificate should at a minimum contain an entry that specifies the name of the regional PCE FQDN in addition to the name of the Supercluster global FQDN.

Kerberos Considerations

VEN pairing/authentication using Kerberos is not a wide-spread practice. Kerberos configuration is complex and beyond the scope of this paper. Customers that rely Kerberos for VEN pairing/authentication with the Supercluster should be sure that the Supercluster global FQDN is defined in the Kerberos `keytab` file on all core nodes of the supercluster, in addition to the regional PCE FQDN.

Separating ports for PCE web console/REST API and VEN traffic

Network traffic for the VENs and PCE web console (UI) by default use the same port: 8443. You can redefine the ports needed for these different kinds of traffic. This recommended configuration change enables greater control of access to the PCE because you can separate user access to the PCE on a port different from the port for VEN traffic.

> If you separate these ports, be sure to appropriately update the health monitor configuration of the `/node_availability` configured on the GTM or LTM.

Port separation is controlled by the `runtime_env.yml` parameters `front_end_management_https_port` and `front_end_https_port`, which are detailed in Illumio® Adaptive Security Platform® PCE Deployment.

Network Address Translation (NAT)

If you run NAT between your between PCE, make sure to change value of the `supercluster.node_public_ip` parameter in the `runtime_env.yml` for all nodes in the PCE cluster. The value of this parameter must match the external address of the node; that is, the post-NAT value. The PCE's Service Discovery component typically informs other nodes of it real, non-NATted IP address. If the addresses between the nodes are NAted, your PCE cluster might fail to join the Supercluster. Using the `supercluster.node_public_ip` parameter defines the actual address to use for connectivity between the nodes in the Supercluster.

Considerations of VEN pairing profile without Supercluster global FQDN

Setting the Supercluster global FQDN is an optional configuration parameter in the `runtime_env.yml` file. By default, the VEN pairing script from the web console uses the FQDN of the Supercluster leader PCE when the Supercluster global FQDN is not specified in the configuration. The pairing script can still be used to pair a VEN to any PCE in the supercluster only if the FQDN is manually edited to reflect the FQDN of the individual PCE.

In this scenario. the VEN is programmed to always try to communicate with the individual PCE FQDN it paired with, instead of the Supercluster global FQDN. In this type of scenario you do not need to use an intelligent DNS resolvers such as the F5 GTM.

The implication for PCE failover is that you must use the REST API to re-home each VEN individually to a different PCE.

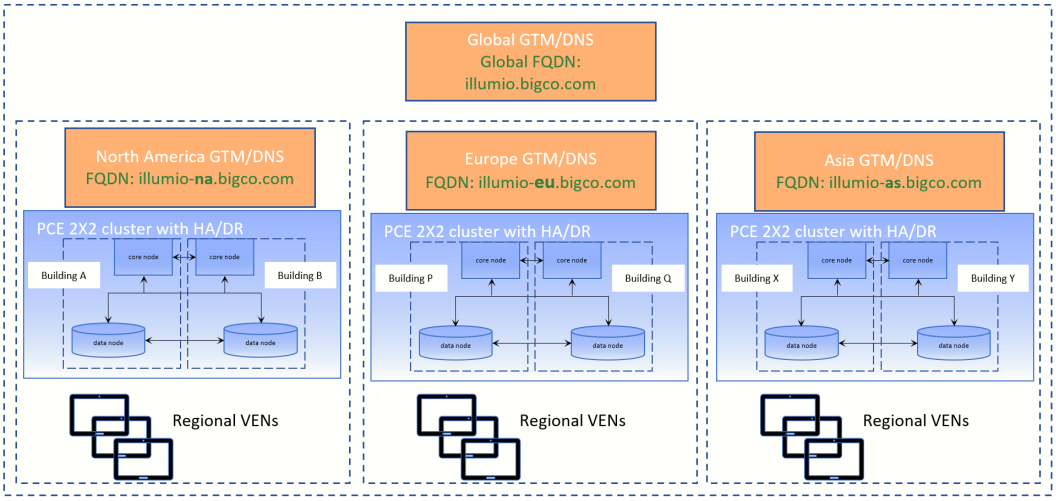
Alternatively, if the SSL certificate uses SAN and includes the FQDN of the all PCEs in the Supercluster, DNS can be manually changed or the GTM can be configured to resolve the IP address of another PCE in the

Supercluster. The VEN heartbeats and relays traffic flow information to the failover PCE without any re-configuring to communicate with a different PCE in the Supercluster.

High Availability and Disaster Recovery

Consider distributing your PCE cluster nodes across separate locations to ensure continued operations in case of a disaster.

Not shown in the Supercluster reference network implementation -- logical architectures is distribution of the nodes of the PCE cluster among different physical locations. This is distribution of nodes at the lowest level of the Supercluster leader and members. The diagram of the distributed data centers in Hypothetical distributed data centers for BigCo.com can be redrawn to show that the various nodes of a cluster have been physically located in different buildings:



You must make sure that network latency is no more than 15ms. (??Anand - this may be 10ms, unless it has changed with a newer version of consul: Alex – this spec is not called out anywhere I can find. Also we do not document "consul" per se; we call it "discovery service")

Standby PCE in case of failure

Consider keeping as a standby an unused but fully configured PCE cluster that you can use in case any member of the Supercluster fails. If this PCE is already a member of the cluster, you can change name resolution either through manual manipulation of DNS A records or "Pool of last resort" in the GTM to direct the VENs to the standby PCE. If the failed PCE is the leader in the supercluster, the standby PCE can also be promoted to the supercluster leader.

See the details on promoting a member and other considerations about restoring in the *Supercluster Deployment and Usage* guide.

VEN-less Supercluster Leader for policy management

Consider *not* pairing VENs with the Supercluster leader. This allows the Supercluster leader resources to be centered exclusively on policy generation and management. In addition, the "VEN-less" leader can be used as a backup in case a member fails.

Background to load balancing (L4/DNS), GSLB, and GTM

As with a single PCE, all PCEs in the Supercluster must be front-ended with a load balancer (DNS or Layer 4) to distribute requests across the PCEs' Core nodes.

Global Server Load Balancing (GSLB) or a manual DNS update can be used to failover VENs to a different PCE in case of an extended PCE outage. See GSLB Requirements and Supercluster HA in the Supercluster Deployment and Usage guide

The original Server Load Balancing (SLB) uses a farm of servers to deliver application and services on the network. Load Balancing servers listen for requests and they allocated the request to the server in the farm that is best placed to fulfill it at this time.

By contrast, Global Server Load Balancing (GSLB) refers to the intelligent distribution of traffic across server resources located in multiple geographies. The servers can be on premises in a company's own data centers, or hosted in a private cloud or the public cloud.

For background on the GTM, see this documentation.

Revision History: Illumio® Adaptive Security Platform® 18.1 SUpercluster Reference Implementation

Date	Description
2018-07-09	Update with additional architectures, step-by-step instructions for the F5 LTM, and other considerations.
2018-06	First published.