Network to Cloud DC (Net2Cloud) Problem statements and Gap Analysis Soliciting comments for WG Last Call

draft-ietf-net2cloud-problem-statement-14
draft-ietf-net2cloud-gap-analysis-08

Linda.Dunbar@futurewei.com
Andy Mails (aqmalis@gmail.com)
Christianjacquenet@orange.com
Mehmet.toy@verizon.com
kmajumdar@microsoft.com
Problem Statement Update Since IETF 110

- Add co-author Kausik Majudmar from Azure
- Added the network problems encountered by Azure
- Revised the gap analysis accordingly
More eBGP peering errors for Cloud DC.

- Traditional network operators establish long term peers whom they have prior negotiated peering policies.
- Those long-term peering relationships resolve issues early on. Less likely to have on-going issues.

- Cloud DC GW: more peers & shorter durations without prior agreed policies.

More BGP peering errors, such as capability mismatch, BGP cease notification, unwanted route leaks, missing Keepalives, etc.
Large number of routes change triggered by a site/pod failure or degradation

- One of the sites/pods have failure or reduced capacity. A group of routes (not entire AFI/SAFI) need to be switched
  - VPN GW or Cloud GW are good; therefore, BFD doesn’t detect the failures.
5G Edge Cloud: Multiple Instances at different Edge Cloud DC

Key issues:
- The difference of routing distances is relatively small
- Capacity at the Edge DC plays bigger role for performance
- Source (UEs) can ingress from different Ingress routers
Connect the on-demand, elastic and 3rd party hosted workloads

- **NAT for Cloud Services**
  - Different Cloud operators support different levels of NAT functions.
  - Proper configuration of NAT has to be performed in Cloud DCs and in their own on-premise DC

- **Cloud Discovery**
  - Location of workloads and connectivity are not easily visible
  - Desirable to have tools to discover cloud services in much the same way as you would discover your on-premises infrastructure
Issues with the DNS in the Context of Routing

- DNS for Cloud Resources
  - Need to establish policies and rules on how/where to forward DNS queries to
    Cloud’s DNS can be configured to forward queries to customer managed authoritative DNS servers hosted on-premises, and to respond to DNS queries forwarded by on-premises DNS servers.
  - Collisions can still occur. Better to use the global domain name even when an organization does not make all its namespace globally resolvable

- DNS based solution to reach App Instances in the optimal Cloud DC locations (Cloud discovery)
  - Dependent on client behavior
    • Client can cache results indefinitely
    • Client may not receive service even though there are servers available (before cache timeout) in another Cloud DC
  - No inherent leverage of proximity information present in the network (routing) layer, resulting in loss of performance
    • Client on the west coast can be mapped to DC on the east coast
  - Inflexible traffic control:
    • Local DNS resolver become the unit of traffic management
Network: Site <-> Cloud & Cloud <-> Cloud

- Problems associated with Multiple Cloud DC Interconnection
  - Different Cloud providers have different access method.
    - Today you have to hairpin the traffic to customer GWs
  - Different Cloud providers have different APIs for calling security functions, the NAT, etc.

- Multiple types of connections to workloads in a Cloud DCs
  - it is not visible to Apps in a Cloud DC what type of network access is used.

- IPsec P2P doesn’t scale well with Multipoint mesh connection & poor performance.

- unknown segments → difficult to collect end to end performance metrics

- Problems of MPLS based VPN extending to Hybrid Cloud DC
  - PE might not have direct connections to Cloud DCs
  - Most Cloud DCs don’t expose their internal network. Difficult to extend MPLS VPN into Cloud DCs
  - Most Cloud Operators use Ipsec VPN to connect to their clients
Net2Cloud Gap Analysis
Gap analysis update since IETF 110

• Application Based Forwarding may require different forwarding topologies based on Application identifiers
• Need simpler way to switch large number of services when a site/pod failure or degradation is detected
• Doesn’t yet have ways to indicate one client route can be carried by multiple underlay paths.
Gap Summary
specifically for Routing Area

• For Accessing Cloud Resources
  a) Traffic Path Management: when a remote vCPE can be reached by multiple PEs of one provider VPN network, it is not straightforward to designate which egress PE to the remote vCPE based on applications or performance.
  b) BGP between PEs and remote CPEs via untrusted networks: BGP over TCP over IPsec

• Missing control plane to advertise network condition connecting the virtual nodes within Cloud DCs.

• Issues of Aggregating traffic over private paths and Internet paths
  a) Control plane messages for different overlay segmentations needs to be differentiated. User traffic belonging to different segmentations need to be differentiated.
  b) BGP Tunnel Encap doesn’t have ways to indicate a route or prefix that can be carried by both IPsec tunnels and VPN tunnels
  c) Missing clear methods in preventing attacks from Internet-facing ports
Next Step

- Request Working Group Last Call