

CLAIMS

What is claimed is:

1. A system comprising:
 - a first cloud exchange network for a first cloud exchange, the first cloud exchange network located within a first data center and configured with a first dedicated virtual gateway, the first dedicated virtual gateway configured to interface with a first virtual connector to a customer network, with a second virtual connector to a first cloud service provider network, and with a third virtual connector to a second cloud service provider network,
 - wherein network traffic among the customer network, the first cloud service provider network, and the second cloud service provider network is routed through the first dedicated virtual gateway, and
 - wherein the first dedicated virtual gateway dynamically polices the network traffic based on an aggregate bandwidth subscription configured in the first cloud exchange network that limits a total bandwidth that may be used over the first cloud exchange network between the customer network, the first cloud service provider network, and the second cloud service provider network.
2. The system of claim 1, further comprising a second cloud exchange network for a second cloud exchange, the second cloud exchange network located within a second data center geographically remote from the first data center, wherein the first dedicated virtual gateway includes a fourth virtual connector to a second dedicated virtual gateway configured on the second cloud exchange network.
3. The system of claim 2, wherein the first dedicated virtual gateway dynamically polices the network traffic based on the aggregate bandwidth subscription limits the total bandwidth that may be used over the first cloud exchange network between the customer network, the first cloud service provider network, the second cloud service provider network, and the second dedicated virtual gateway.
4. The system of claim 2, wherein the first dedicated virtual gateway exchanges routes with the second dedicated virtual gateway.

5. The system of claim 1, wherein to dynamically police the network traffic, the first dedicated virtual gateway performs destination-based rate limiting.
6. The system of claim 5, wherein the first dedicated virtual gateway is configured, based on the destination-based rate limiting, with a first policer to shape network traffic between the first dedicated virtual gateway and the first cloud service provider network and a second policer to shape network traffic between the first dedicated virtual gateway and the customer network.
7. The system of claim 6, wherein the first dedicated virtual gateway does not directly shape network traffic between first cloud service provider network and the customer network.
8. The system of claim 1, further comprising a second cloud exchange network for a second cloud exchange, the second cloud exchange network located within a second data center geographically remote from the first data center, wherein the first dedicated virtual gateway is configured to dynamically shape network traffic over the first cloud exchange network based on inter-exchange traffic between the first dedicated virtual gateway and a second dedicated virtual gateway operating in the second cloud exchange network.
9. The system of claim 1, further comprising a second cloud exchange network for a second cloud exchange, the second cloud exchange network located within a second data center geographically remote from the first data center, wherein when the customer network is subscribed to a third cloud service provider connected to the second cloud exchange network, the first dedicated virtual gateway automatically establishes a fourth virtual connector with a second dedicated virtual gateway configured on the second cloud exchange network.
10. The system of claim 1, further comprising:
 - a programmable network platform comprising processing circuitry and configured to:
 - obtain an aggregate bandwidth for the aggregate bandwidth subscription;

obtain a local bandwidth for network traffic forwarded by the first dedicated virtual gateway to the first cloud service provider network;

set a remote bandwidth limit for network traffic forwarded by the first dedicated virtual gateway to the second cloud service provider network or the customer network, the remote bandwidth limit based on a difference between the aggregate bandwidth and the local bandwidth; and

configure the first dedicated virtual gateway with the remote bandwidth limit to cause the first dedicated virtual gateway to limit the remote bandwidth for network traffic forwarded to the second cloud service provider network or the customer network.

11. A system comprising:

a first cloud exchange network of a first cloud exchange, the first cloud exchange network located within a first data center and configured with a first dedicated virtual gateway storing a first route to a customer network; and

a second cloud exchange network of a second cloud exchange, the second cloud exchange network located within a second data center geographically remote from the first data center, and configured with a second dedicated virtual gateway storing a second route a cloud service provider network,

wherein the first dedicated virtual gateway polices network traffic over the first route based on a first aggregate bandwidth subscription configured in the first cloud exchange network and network traffic between the first dedicated virtual gateway and the second dedicated virtual gateway, and

wherein second dedicated virtual gateway polices network traffic over the second route based on a second aggregate bandwidth subscription established by the second cloud exchange network and network traffic between the first dedicated virtual gateway and the second dedicated virtual gateway.

12. The system of claim 11, further comprising:

a programmable network platform comprising processing circuitry and configured to:

obtain a first aggregate bandwidth for the first aggregate bandwidth subscription;

obtain a local bandwidth for network traffic forwarded by the first dedicated virtual gateway to networks colocated in the first data center;

set a remote bandwidth limit for network traffic forwarded by the first dedicated virtual gateway to the second dedicated virtual gateway based on a difference between the first aggregate bandwidth and the local bandwidth; and

configure the first dedicated virtual gateway with the remote bandwidth limit to cause the first dedicated virtual gateway to limit the remote bandwidth for network traffic forwarded to the second dedicated virtual gateway.

13. The system of claim 12, wherein the first dedicated virtual gateway configured with the remote bandwidth limit performs destination-based rate limiting to police the network traffic to limit the network traffic forwarded to the second dedicated virtual gateway.

14. The system of claim 12, wherein the programmable network platform is configured to:

map respective Border Gateway Protocol communities for one or more networks, advertised by the second dedicated virtual gateway, to a single policer for the first dedicated virtual gateway; and

configure, in the first dedicated virtual gateway, the single policer for the first dedicated virtual gateway with a remote bandwidth limit that is based on the first aggregate bandwidth subscription.

15. A method comprising:

configuring, by a programmable network platform, a first dedicated virtual gateway in a first cloud exchange network located within a first data center, to interface with a first virtual connector to a customer network, with a second virtual connector to a first cloud service provider network, and with a third virtual connector to a second cloud service provider network, wherein network traffic among the customer network, the first cloud service provider network, and the second cloud service provider network is routed through the first dedicated virtual gateway; and

dynamically policing, by the first dedicated virtual gateway, the network traffic based on an aggregate bandwidth subscription that limits a total bandwidth that may be used over the first cloud exchange network between the customer network, the first cloud service provider network, and the second cloud service provider network.

16. The method of claim 15, wherein the first dedicated virtual gateway includes a fourth virtual connector to a second dedicated virtual gateway configured on a second cloud exchange network of a second cloud exchange, the second cloud exchange network located within a second data center geographically remote from the first data center,

17. The method of claim 16, wherein dynamically policing the network traffic based on the aggregate bandwidth subscription comprises limiting the total bandwidth that may be used over the first cloud exchange network between the customer network, the first cloud service provider network, the second cloud service provider network, and the second dedicated virtual gateway.

18. The method of claim 16, further comprising:
receiving, by the first dedicated virtual gateway, routes from the second dedicated virtual gateway that indicate networks reachable via the second dedicated virtual gateway.

19. The method of claim 18, wherein dynamically policing the network traffic comprising performing destination-based rate limiting based on the routes from the second dedicated virtual gateway.

20. The method of claim 19, further comprising:
configuring, by the programmable network platform, the first dedicated virtual gateway with a first policer to perform destination-based rate limiting to shape network traffic between the first dedicated virtual gateway and the first cloud service provider network and a second policer to perform destination-based rate limiting to shape network traffic between the first dedicated virtual gateway and the customer network.

21. The method of claim 15, further comprising:
configuring, by the programmable network platform, the first dedicated virtual gateway to dynamically shape network traffic over the first cloud exchange network based on inter-exchange traffic between the first dedicated virtual gateway and a second dedicated virtual gateway operating in a second cloud exchange network.