CLAIMS

Claim 1: A system comprising:

a first cloud exchange network for a first cloud exchange, the first cloud exchange network comprising one or more network devices located within a first data center, the one or more network devices configured with a routing instance to implement 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 the routing instance of the first dedicated virtual gateway is configured to process network traffic among the customer network, the first cloud service provider network, and the second cloud service provider network routed through the routing instance of the first dedicated virtual gateway to dynamically police the network traffic according to a destination-based rate limit for rate-limiting the network traffic received from the customer network at the first cloud service provider network and the second cloud service provider network, to provide an aggregate bandwidth subscription for the customer network; and

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.

Claim 2: The system of claim 1, wherein the one or more network devices comprise first one or more network devices, the system further comprising:

a second cloud exchange network for a second cloud exchange, the second cloud exchange network comprising second one or more network devices located within a second data

center geographically remote from the first data center, the second one or more network devices configured with a routing instance to implement a second dedicated virtual gateway, wherein the first dedicated virtual gateway is configured to interface with a fourth virtual connector to the second dedicated virtual gateway.

Claim 3: The system of claim 2, wherein the routing instance of the first dedicated virtual gateway dynamically polices the network traffic based on the aggregate bandwidth subscription that 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.

Claim 4: The system of claim 2, wherein the routing instance of the first dedicated virtual gateway exchanges routes with the routing instance of the second dedicated virtual gateway.

Claim 5: The system of claim 1,

wherein to dynamically police the network traffic according to the destination-based rate limit, the first dedicated virtual gateway performs destination-based rate limiting such that the network traffic is rate-limited to a set of one or more destinations, and

wherein the first dedicated virtual gateway uses destination class usage to perform the destination-based rate limiting by monitoring the network traffic from the customer network by performing lookups of a destination address of the network traffic destined for at least one of the first cloud service provider network or the second cloud service provider network.

Claim 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.

Claim 7: The system of claim 6, wherein the routing instance of the first dedicated virtual gateway does not directly shape network traffic between first cloud service provider network and the customer network.

Claim 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.

Claim 9: The system of claim 1, wherein the one or more network devices comprise first one or more network devices, the system further comprising:

a second cloud exchange network for a second cloud exchange, the second cloud exchange network comprising second one or more network devices located within a second data center geographically remote from the first data center, the second one or more network devices configured with a routing instance to implement a second dedicated virtual gateway, wherein when the customer network is subscribed to a third cloud service provider connected to the second cloud exchange network, the routing instance of the first dedicated virtual gateway automatically establishes a fourth virtual connector with the routing instance of the second dedicated virtual gateway.

Claim 10: A system comprising:

a first cloud exchange network of a first cloud exchange, the first cloud exchange network comprising first one or more network devices located within a first data center, the first one or more network devices configured with a first routing instance to implement a first dedicated virtual gateway storing a first route to a customer network;

a second cloud exchange network of a second cloud exchange, the second cloud exchange network comprising second one or more network devices located within a second data center geographically remote from the first data center, the second one or more network devices configured with a second routing instance to implement a second dedicated virtual gateway storing a second route to a cloud service provider network,

wherein the first routing instance of the first dedicated virtual gateway is configured to police 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,

wherein the second routing instance of the second dedicated virtual gateway is configured to police network traffic over the second route based on a second aggregate bandwidth subscription configured in the second cloud exchange network and network traffic between the first dedicated virtual gateway and the second dedicated virtual gateway, and

wherein the first routing instance is configured to police network traffic over the first route according to a first destination-based rate limit for rate-limiting the network traffic between the first dedicated virtual gateway and the second dedicated virtual gateway and the first aggregate bandwidth subscription, and wherein the second routing instance is configured to police network traffic over the second route according to a second destination-based rate limit for rate-limiting the network traffic between the first dedicated virtual gateway and the second dedicated virtual gateway and the second aggregate bandwidth subscription; and

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.

Claim 11: The system of claim 10, wherein to police network traffic according to the first destination-based rate limit, 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 using destination class usage to perform the destination-based rate limiting by monitoring the network traffic from the customer network by performing lookups of a destination address of the network traffic destined for at least one of the customer network or the cloud service provider network.

Claim 12: The system of claim 10, 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.

Claim 13: A method comprising:

configuring, by a programmable network platform, a first dedicated virtual gateway in a first cloud exchange network comprising one or more network devices located within a first data center, the one or more network devices configured with a routing instance to implement the first dedicated virtual gateway 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 the routing instance of the first dedicated virtual gateway is configured to process network traffic among the customer network, the first cloud service provider network, and the second cloud service provider network routed through the routing instance of the first dedicated virtual gateway to dynamically police, by the first dedicated virtual gateway, the network traffic according to a destination-based rate limit for rate-limiting the network traffic received from the customer network at the first cloud service provider network and the second cloud service provider network, to provide 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;

obtaining, by the programmable network platform, an aggregate bandwidth for the aggregate bandwidth subscription;

obtaining, by the programmable network platform, a local bandwidth for network traffic forwarded by the first dedicated virtual gateway to the first cloud service provider network;

setting, by the programmable network platform, 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

configuring, by the programmable network platform, 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 customer network.

Claim 14: The method of claim 13, wherein the one or more network devices comprise first one or more network devices, wherein the first dedicated virtual gateway is configured to interface with a fourth virtual connector to a second dedicated virtual gateway comprising a routing instance configured on second one or more network devices of 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,

Claim 15: The method of claim 14, wherein dynamically policing the network traffic based on the aggregate bandwidth subscription comprises limiting, by the routing instance, 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.

Claim 16: The method of claim 14, further comprising:

receiving, by the routing instance of the first dedicated virtual gateway, routes from the routing instance of the second dedicated virtual gateway that indicate networks reachable via the second dedicated virtual gateway.

Claim 17: The method of claim 16, wherein dynamically policing the network traffic according to the destination-based rate limit comprises performing destination-based rate limiting based on the routes from the second dedicated virtual gateway such that the network traffic is rate-limited to a set of one or more destinations, the method further comprising:

using, by one of the first dedicated virtual gateway or the second dedicated virtual gateway, destination class usage to perform destination-based rate limiting by monitoring the network traffic from the customer network by performing lookups of a destination address of the network traffic destined for at least one of the first cloud service provider network or the second cloud service provider network.

Claim 18: The method of claim 17, 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.

Claim 19: The method of claim 13, 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.