

TITLE: INTERFACES TO MANAGE DIRECT NETWORK PEERINGS**BACKGROUND**

[0001] Many companies and other organizations operate computer networks that
5 interconnect numerous computing systems to support their operations and the services they
provide to their end customers distributed worldwide. For example, data centers housing
significant numbers of interconnected computing systems have become commonplace, such as
private data centers that are operated by and on behalf of a single organization, and public data
centers that are operated by entities as businesses to provide computing resources to customers.
10 In many cases providers set up large networks that may logically span several regions or even
countries, and may include numerous data centers with varying levels of services and facilities
available, utilized together to provide a unified set of services to their end customers.

[0002] In some data centers that have been set up to provide computing and/or storage
facilities to remote clients, the set of computational resources at the data center may be
15 dynamically divided into resource pools, with each pool being made available for exclusive use
by a given client for designated periods of time. There are a number of alternatives available for
how the consumers of these facilities establish network connectivity to the resource pools that
have been designated for their use. The customer requests may originate from a wide variety of
devices – desktop personal computers, laptops, client-office servers, tablets, smart phones and
20 the like. These devices may use either long-lasting network links (e.g., using a client office
network with a T1 connection) to communicate with their proximate private network and/or the
public Internet, or they may have transient connectivity (e.g., in the case where the customer uses
a mobile smart phone). The proximate networks to which the customer devices are directly
connected may in turn route request traffic to the provider network's data centers over a wide
25 variety of paths. Such paths in many cases may have somewhat unpredictable performance,
reliability and security characteristics.

[0003] For some casual types of service requests, such as a request from a customer to read a
recent news article from a web-based news provider, a reasonable variation in responsiveness
and an occasional dropped connection may be acceptable. However, for many business-related
30 data transmissions, such as stock quote services and stock purchase order services provided by
online stock traders, or for high-bandwidth software package deployments originating at a
software development center, more stringent performance, reliability and security needs may
exist. In such environments, a customer of the provider network may need a greater level of
network isolation and control than is generally available over the public Internet. For example,

the customer may wish to establish, if possible, dedicated physical network links between the customer's own network and the provider network, such that the only traffic transmitted over those links is traffic generated on behalf of the customer and in accordance with policies established by the customer. Furthermore, to meet quickly changing business needs, customers
5 may want the ability to enable and disable such dedicated links dynamically and with a minimum of effort and delay.

BRIEF DESCRIPTION OF THE DRAWINGS

- [0004] Figure 1 illustrates an example system, according to at least some embodiments.
- 10 [0005] Figure 2 provides a high level overview of steps that may be taken with the help of services provided by a connectivity coordinator to establish dedicated connectivity between a client network and a resource collection, according to at least some embodiments.
- [0006] Figure 3 illustrates an example of constituent elements of a request for dedicated connectivity from a client, according to at least some embodiments.
- 15 [0007] Figure 4 illustrates an example of constituent elements of a response to a request for dedicated connectivity, according to at least some embodiments.
- [0008] Figure 5 illustrates an example of constituent elements of a request for a logically-isolated network path, according to at least some embodiments.
- [0009] Figure 6 illustrates an example of contents of a confirmation message indicating that
20 requested connectivity has been established, according to at least some embodiments.
- [0010] Figure 7 illustrates an example of two logically-isolated network paths sharing a physical link established using an interface provided by a connectivity coordinator, according to at least some embodiments.
- [0011] Figure 8 is an illustration of a portion of an exemplary web-based interface that may
25 be provided by connectivity coordinator, according to at least some embodiments.
- [0012] Figure 9 is a flowchart of a method for providing connectivity-related services, according to at least some embodiments.
- [0013] Figure 10 is a flowchart of a method for providing connectivity-related services including guidance on configuring a client network device, according to at least some
30 embodiments.
- [0014] Figure 11 illustrates an example of a system including a provider network to which connectivity may be provided from client networks via "last-mile" connectivity providers, according to at least some embodiments.
- [0015] Figure 12 illustrates an example of constituent elements of a request for dedicated
35 connectivity through a connectivity provider, according to at least some embodiments.

[0016] Figure 13 illustrates an example of constituent elements of a response to a dedicated connectivity request that includes information on available connectivity providers, according to at least some embodiments.

[0017] Figure 14 illustrates an example communication from a client identifying a selected connectivity provider, according to at least some embodiments.

[0018] Figure 15 illustrates example communications from a connectivity coordinator to a connectivity provider and a client after the client has selected the connectivity provider, according to one embodiment.

[0019] Figure 16 is an illustration of a portion of an exemplary web-based interface that may be provided for initiating connectivity provider selection, according to some embodiments.

[0020] Figure 17 is a flowchart of a method for enabling clients to select connectivity providers, according to at least some embodiments.

[0021] Figure 18 is a flowchart of a method for providing dynamic connectivity-related services, according to at least some embodiments.

[0022] Figure 19 is a flowchart of a method comprising responding dynamically to changing traffic levels, according to at least some embodiments.

[0023] Figure 20 is a block diagram illustrating an example computer system that may be used in some embodiments.

[0024] While embodiments are described herein by way of example for several embodiments and illustrative drawings, those skilled in the art will recognize that embodiments are not limited to the embodiments or drawings described. It should be understood, that the drawings and detailed description thereto are not intended to limit embodiments to the particular form disclosed, but on the contrary, the intention is to cover all modifications, equivalents and alternatives falling within the spirit and scope as defined by the appended claims. The headings used herein are for organizational purposes only and are not meant to be used to limit the scope of the description or the claims. As used throughout this application, the word “may” is used in a permissive sense (i.e., meaning having the potential to), rather than the mandatory sense (i.e., meaning must). Similarly, the words “include,” “including,” and “includes” mean including, but not limited to.

DETAILED DESCRIPTION OF EMBODIMENTS

[0025] Various embodiments of methods and apparatus for using interfaces such as application programming interfaces (APIs) to manage dedicated network connectivity between customer networks and provider networks are described. Networks set up by an entity such as a company or a public sector organization to provide one or more services accessible via the

Internet (such as various types of cloud-based computing or storage) to a distributed set of clients may be termed provider networks in this document. Such a provider network may include numerous data centers hosting various resource pools, such as collections of computer servers, storage devices, networking equipment and the like, needed to implement and distribute the
5 services offered by the provider.

[0026] In order to extend the geographical areas over which its services can be accessed with desired levels of performance, reliability and security, an operator of a provider network may establish dedicated private network paths between its data centers and one or more routers that are physically located at a facility remote from the data centers. The facilities at which these
10 routers are housed are termed “router co-location facilities” in this document, as they may sometimes house routers and other network equipment owned and/or managed by business entities other than the provider network’s operator, such as by independent network service providers or by the clients themselves. Routers owned or managed by, or on behalf of, the provider network operator at the router co-location facilities are called “endpoint” routers in this
15 document, as they may represent the furthest points to which the provider network’s control or ownership of network equipment extends. For example, only traffic that has passed through a device owned or managed by the provider network operator, and therefore complies with policies set by the provider network operator, may be allowed on the private paths between the endpoint routers and other components of the provider network. In some embodiments one or more other
20 routers at the router co-location facilities may be part of a client network – i.e., such routers may owned and/or managed by or on behalf of the clients, or the other routers may have private connectivity to the systems at which clients of the provider network generate service requests for the provider network. These other routers are termed “client-side” routers in this document.

[0027] In order to facilitate the management of network connectivity between the client
25 networks and the provider network, in some embodiments a connectivity coordinator responsible for implementing one or more programmatic interfaces through which connectivity-related requests from clients are handled may be set up. A variety of different kinds of connectivity-related requests, such as requests to discover the types of connectivity options that may be accessible, requests to select a particular connectivity option or service, requests to set up or
30 dismantle a physical link, and the like may be made available through the interface in different implementations. The interface may be exposed to the client in many different ways in various embodiments: for example, as an API, through a graphical user interface, a web page or web site, or even as a collection of commands that can be issued from a computer system’s command-line prompt.

[0028] In one embodiment, one or more collections or pools of resources at a data center may be allocated for use by a particular client, i.e., to implement functionality needed to satisfy services requested from devices of the client network. In such an embodiment, a connectivity coordinator may be operable to receive a request to establish dedicated connectivity from a client to one or more of the resource pools. The connectivity request may be generated or formatted to conform to the interface implemented by the connectivity coordinator – for example, it may be received via a web-based form submission in a case where the interface is presented to the client as a set of web pages. In response to the request for dedicated connectivity, the connectivity coordinator may select a particular endpoint router from among the set of endpoint routers of the provider network as the target router from which dedicated connectivity is to be provided to the requesting client. For example, the target router may be selected from the available endpoint routers at a router co-location facility geographically closest to the client's premises, at which the client has access to an existing client-side router. In some implementations the interface may allow the client to specify various details in the request that may help the connectivity coordinator choose an appropriate target endpoint router, such as one or more names and/or addresses of router co-location facilities, a desired bandwidth, desired price ranges, and the like.

[0029] Having selected the target endpoint router, the connectivity coordinator may generate configuration instructions for one or more physical network links to be established to provide the desired dedicated connectivity, and transmit the instructions back to the client in reply to the client's request. The reply may also be generated to conform to the interface implemented by the connectivity provider: e.g., in a case where a web page with a form was used for the client request, the configuration instructions may also be specified as one or more web pages, or as links to documents accessible from the web site. The configuration instructions may, for example, identify a physical location of the endpoint router, a physical port of the target endpoint router, the rack in which the target router is housed, the kind of connector required for the physical link, and so on. After the client sets up the physical network link in accordance with the instructions, the connectivity provider in some embodiments may verify that the link has been set up correctly and send a confirmation message to the client indicating that the requested dedicated connectivity has been established.

30

Example system environment

[0030] Figure 1 illustrates an example system environment, according to at least some embodiments. The system 100 may include a provider network 105 with one or more data centers 110 maintained to provide services to clients, such as cloud computing services or cloud storage services. The data centers 110 in turn may include resource collections such as 120A and

35

120B. Each resource collection 120 may include a set of resources (for example, computer servers, storage devices, network devices, etc.) such as resources 112A in resource collection 120A and resources 112B in resource collection 120B. The system 100 may also include a connectivity coordinator 114 configured to provide a connectivity service to the clients, a connectivity database 115, and a plurality of endpoint routers such as endpoint routers 132A and 132B in some embodiments. The endpoint routers 132 may be linked to the resource collections 120 via private network paths such as paths 170A, 170B, 170C and 170D. Direct use of a private network path 170, such as 170A-170D, may be limited to devices and servers of the provider network, i.e., a network packet may only be physically transmitted upon a link of the private network path from a device owned or managed by the owner of the provider network. The term “path” as used herein, broadly refers to the set of links and devices traversed by a network message or packet between a source device and a destination device. The set of links of a given path may, in some cases, comprise a single wire, as when the source and destination may be linked directly by a physical cable. In other cases the path may include multiple wired and/or wireless links and several intermediary devices such as switches, gateways, routers and the like. Paths may comprise unidirectional and/or bidirectional physical links.

[0031] In the illustrated embodiment, two example client networks 162A and 162B representing respective clients of the provider network 105 are shown. Each client network comprises a plurality of client devices 148 (e.g., 148A and 148B) from which requests that are ultimately serviced at resource collections 120 may be generated. Such client devices 148 may include a variety of systems such as desktop or racked computer systems, laptops, notebooks, tablets, smart phones and the like. Some client devices 148 may be housed at client data office premises, client data centers, or client home premises in various embodiments, and others may be mobile devices with no fixed physical location. In the illustrated environment, the client devices 148 of a particular client have access to a client-side router – e.g., client devices 148A of client network 162A are connected to client-side router 142A via path 160A, and client devices 148B of client network 160B are connected to client-side router 142B via path 160B.

[0032] Client-side router 142A in the example shown in Figure 1 is housed at a router co-location facility 150A, where endpoint router 132A is also present, and client-side router 142B is housed at a router co-location facility 150B where endpoint router 132B is located. In general, a number of different types of paths to various components of data center 110, such as connectivity coordinator 114 and resource collections 120, may be available from client networks 162A in various embodiments. For example, a path 175 that does not include endpoint router 132A is shown in Figure 1 between client devices 148A of client network 162A and connectivity provider 114; such a path 175 may include various routers, gateways and devices of the public

Internet, for example, which may or may not provide desired levels of performance, availability, reliability or other service characteristics required for some of the services provided by the provider network 105. Other paths similar to 175, but not illustrated in Figure 1, may also be available to the resource collections 120 from client devices 148A and/or 148B.

5 **[0033]** The connectivity service provided by connectivity coordinator 114 may include a number of techniques implemented to help clients of the provider network establish and manage dedicated network paths from the client networks 162 to the resource collections 120 in various embodiments. For example, a cross-network connection 191 including a physical link or cable between client-side router 142B and endpoint router 132B may have been established with the
10 help of some of the features of the connectivity service provided by connectivity coordinator 114. The term cross-network connection, as used herein, refers to a physical network connection set up between two autonomous networks. For example, within the Internet, an autonomous network may be identified by a unique Autonomous System (AS) identifier - a collection of connected Internet Protocol (IP) routing prefixes under the control of one or more network
15 operators that presents a common, clearly defined routing policy to the Internet. In the illustrated embodiment, the connectivity coordinator 114 may be operable to implement an interface defining various connectivity operations available to clients, including establishment of cross-network connections such as connection 191, and establishment of logically isolated connections or paths using such cross-network links. The interface may be implemented using a variety of
20 approaches: e.g., as an application programming interface (API), a web-based interface, other graphical user interfaces, or command-line interfaces, in different embodiments. Connectivity coordinator 114 may also make the interface known to clients using various mechanisms – e.g., a notification detailing the interface may be published on one or more corporate web sites in one embodiment. Connectivity coordinator 114 may itself be resident outside provider network 105
25 in some embodiments, and in other embodiments it may be incorporated within the provider network 105.

30 **[0034]** Using the interface implemented by connectivity coordinator 114, a client may generate a request for dedicated connectivity. The client may, by making such a request, wish to indicate that a physical network link (similar to cross-network connection 191 shown in Figure 1)
35 be established exclusively for a client network 162A to communicate with one or more resource pools 120, for any of a number of reasons such as better performance, higher reliability, enhanced security, or lower or more predictable cost. The exclusivity requirement may indicate that only traffic to or from a specified set or sets of network addresses in the client's network should be allowed over the requested dedicated physical network link. The set or sets of physical addresses may be identified and/or modified by the client after the physical link has been set up

in some implementations, i.e., they may not have to be specified prior to the initial establishment of the physical link. The request may provide any combination of a number of different details that may be of use to connectivity provider 114 in various embodiments – for example, the physical location of one or more client-side routers that could be used, a desired bandwidth and, or other service requirements. In some implementations specific service requirements may be provided in subsequent requests after the initial request to establish dedicated physical connectivity has been sent.

[0035] In response to receiving such a request, connectivity coordinator 114 may select a particular endpoint router 132 that may be appropriate to meet the requirements of the client. For example, the endpoint router that is physically closest to one or more client premises where the majority of client requests may be expected to originate may be chosen in one implementation. Having selected the target endpoint router to provide the dedicated connectivity, the connectivity coordinator 114 may generate a reply or notification comprising configuration instructions for a physical network link to be established to the target endpoint router 132 to provide at least a portion of the desired dedicated connectivity. The instructions may include a variety of elements in various embodiments, including for example a port number, rack identifier and the like to identify the target endpoint router. The instructions may also be generated in accordance with the interface – for example, as the contents of a web page in an implementation where the request was received as a web-based form submission. In some cases multiple responses may be sent by the connectivity coordinator – for example, a combination of one or more web responses to the requester, and one or more email messages, some of which may be sent to entities other than the client, such as the operator of the router co-location facility 150. The response or responses from the connectivity coordinator 114 may be used to set up the physical link needed. The time taken to establish the physical link may vary widely, depending on the responsiveness of the client, the responsiveness of the operator of the router co-location facility 150, and various procedures that may have to be completed for security, authorization and the like. After the physical network link has been set up, the connectivity coordinator 114 may in some embodiments transmit a confirmation message to the client indicating that the desired dedicated connectivity has been established. Various connectivity related information, including for example data identifying details of physical links such as cross-network connection 191, data identifying the clients for which such links have been established, the dates or times of establishment of dedicated connectivity, and the like, may be stored in connectivity database 115 in some embodiments.

[0036] In addition to the network paths illustrated in Figure 1, in many environments there may be several other alternative network paths available between the client networks 162 and various components of the system 100. For example, some connectivity requests may be

transmitted to the connectivity provider 114 over a path that includes public or shared links, and various services provided at resource collections 120 may be accessed over public or shared links as well. In some cases the alternative paths may serve as backups in case connectivity over the desired dedicated paths is interrupted.

5

Physical and logical connectivity establishment

[0037] Figure 2 provides a high level overview of steps that may be taken with the help of services provided by connectivity coordinator 114 in one embodiment to establish dedicated connectivity between a client network 162 and a resource collection 120. As shown in the entry
10 labeled 201 in Figure 2, the connectivity coordinator 114 may implement an interface defining a set of connectivity-related operations available to clients of the provider network 105, to other entities (such as one or more administrative servers, measurement agents, billing agents and the like) and/or other parties. The set of available operations may include, for example, operations to create, query, retrieve, update or delete connectivity records or objects in some implementations.
15 The available operations may be exposed via application programming interfaces (APIs) in any of a variety of standard specifications or programming languages, such as Web Services Description Language (WSDL), XML, Java, C, C++, Python, Perl, or derivatives thereof, in some environments, where clients may interact with the connectivity provider programmatically by issuing method calls, function calls and the like. In other environments, in addition to or
20 instead of providing a public API using which clients may write code, the connectivity coordinator may provide a more user-friendly interface such as a collection of Web pages. In one implementation, the connectivity coordinator may for example publish a set of documents (similar to Javadocs in a case where Java or a Java-like programming language is used) that provide an exhaustive list of APIs, and may expose a frequently used subset of connectivity-
25 related operations via a Web page or pages. In such an environment a client may opt to use the Web pages for common operations, and may resort to programs that invoke the API calls for more complex operations or for operations for which a web interface is not provided by connectivity coordinator 114. A particular web-based interaction with the client may result in an invocation of one or more of the APIs internally at the connectivity coordinator 114 in some such
30 embodiments. Other types of interfaces, such as command-line tools, independently installable graphical user interfaces (GUIs) (i.e., GUIs that do not rely on Web pages and HTTP-based interactions), thick clients, electronic mail, or messaging protocols, may be used in isolation or in combination to implement the services provided by connectivity coordinator 114 in various embodiments. In some cases the interface may consist of multiple layers, where one layer of

interface may invoke another, and one or more of the layers may be exposed for direct client interactions.

[0038] In one embodiment, the connectivity coordinator may provide a “Getting Started Guide” or some other similar documentation that may provide examples of how the interface may be used. The following list, with entries labeled API-1 through API-18, is a set of API call invocation examples that may be provided in such documentation for a subset of connectivity services provided by connectivity coordinator 114.

[API-1] *CustomerId customerId = createNewCustomer(CustomerInfo customerInfo);*

The createNewCustomer API may be used to create a customer account at the connectivity provider. It may take customer information (e.g., name, address, payment-related details) as input, and return a customer identifier.

[API-2] *ConnectionRequestId requestId = requestDirectConnection(CustomerId customerId, ConnectionSpecification connectionSpecification);*

The requestDirectConnection API may be used to by a customer to send a request for dedicated connectivity, with various details of the properties of the desired connectivity encapsulated in a ConnectionSpecification object.

[API-3] *RequestStatus requestStatus = getConnectionRequestStatus(CustomerID customerId, RequestId requestId);*

A customer may use the getConnectionRequestStatus API to query the current status of a connection request – for example, the connectivity provider may indicate in the returned RequestStatus object that the current state is “in-progress”, “completed” or “rejected”.

[API-4] *ConnectionId connectionId = getConnectionId(CustomerID customerId, RequestId requestId);*

If a Connection Object is created successfully by the connectivity coordinator (and for example stored in connectivity database 115), a client may use the getConnectionId API to obtain an identifier for that connection object.

[API-5] *ConnectionInfo connectionInfo = getConnectionInfo(ConnectionId connectionId);*

The getConnectionInfo API may be used to obtain the properties of the connection object, including such properties as the physical location of a router, a port number, traffic usage metrics, etc.

[API-6] *PhysicalConnectionInfo physicalInfo = getPhysicalConnectionInfo(ConnectionInfo connectionInfo);*

The getPhysicalConnectionInfo API may be used to extract the location-specific properties of the connection object from the ConnectionInfo object.

[API-7] *AuthInfo authInfo =*

getAuthInfo(PhysicalConnectionInfo(physicalConnectionInfo);

The getAuthInfo API may be used to extract authorization-related information for the connection – e.g., a document allowing a technician to enter the premises where an endpoint router 132 is housed, and make a physical network link to an endpoint router.

- 5 [API-8] ***RequestStatus modificationStatus = modifyConnection(ConnectionId connectionId, ModificationInfo modificationInfo);***

The modifyConnection API may be used to request changes to an existing Connection – e.g., to request more bandwidth.

[API-9] ***RequestStatus disableStatus = disableConnection(ConnectionId connectionId);***

- 10 The disableConnection API may be used to request that an existing connection be disabled, i.e., that no traffic be allowed to flow through the physical link previously set up for that connection.

[API-10] ***RequestStatus enableStatus = enableConnection(ConnectionId connectionId);***

The enableConnection API may be used to request that an existing (e.g., currently disabled) connection be enabled.

- 15 [API-11] ***RequestStatus deleteStatus = deleteConnection(ConnectionId connectionId);***

The deleteConnection API may be used to request that a connection be removed permanently.

[API-12] ***LogicalRequestId logicalRequestId = setUpLogicalConnection(ConnectionId connectionId, LogicalConnectionParameters lcParameters);***

- 20 The setUpLogicalConnection API may be used to request that a logically isolated network path be set up using a previously established physical connection and a set of logical connection properties encapsulated in a LogicalConnectionParameters object.

[API-13] ***LogicalConnectionId logicalConnectionId =***

getLogicalConnectionId(LogicalRequestId logicalRequestId);

- 25 A client may use the getLogicalConnectionId API to obtain an identifier for a particular logical connection.

[API-14] ***LogicalConnectionInfo logicalConnectionInfo =***

getLogicalConnectionInfo(LogicalConnectionId logicalConnectionId);

- 30 The getLogicalConnectionInfo API may be used to obtain the properties of the logical connection, including such properties as the VLAN tag being used for the logical connection, and/or other routing-related information associated with the logical connection.

[API-15] ***LogicalConnectionRequestStatus modificationStatus =***

modifyLogicalConnection(LogicalConnectionId logicalConnectionId, LogicalConnectionModificationInfo modificationInfo);

- 35 The modifyLogicalConnection API may be used to request changes to an existing logical connection – e.g., to modify the set of network prefixes associated with it.

[API-16] *LogicalConnectionRequestStatus disableLogicalConnectionStatus =
disableLogicalConnection(LogicalConnectionId connectionId);*

The disableLogicalConnection API may be used to request that an existing logical connection be disabled, i.e., that no traffic be allowed to flow through the logically-isolated path associated
5 with the logical connection.

[API-17] *LogicalConnectionRequestStatus enableLogicalConnectionStatus =
enableLogicalConnection(LogicalConnectionId connectionId);*

The enableLogicalConnection API may be used to request that an existing (e.g., currently disabled) logical connection be enabled.

10 [API-18] *LogicalConnectionRequestStatus deleteLogicalConnectionStatus =
deleteLogicalConnection(LogicalConnectionId connectionId);*

The deleteLogicalConnection API may be used to request that a logical connection be removed permanently.

[0039] Turning again to Figure 2, the next high-level step illustrated at 206 is the
15 establishment of a client account, which may be used for example for billing purposes. In some embodiments the interface provided by connectivity coordinator 114 may be used (such as via an invocation of a createNewCustomer API or via a web interface that in turn invokes a similar API) to set up the customer account. In other embodiments the connectivity coordinator 114 may not be involved in account creation directly, and some other mechanism (such as interactions
20 with an account manager component not shown in Figure 1) may be used to set up customer accounts.

[0040] A customer that has an account set up may use the interface implemented by connectivity coordinator 114 to first establish a physical link for the desired dedicated connectivity (entry 211 in Figure 2), and then establish one or more logically-isolated network
25 paths that use that physical link (entry 221). Finally, the functionality of the dedicated connectivity may be verified or validated (entry 231), e.g., in some embodiments the client and/or the connectivity coordinator 114 may perform one or more verification operations and confirm that the client's request has been satisfactorily implemented. Each of the high-level steps illustrated in entries 211, 221 and 231 of Figure 2 may involve several interactions and/or
30 operations at the client end and at connectivity provider 114, and further details of each high-level step are provided below.

Example requests and responses for connectivity establishment

[0041] Figure 3 illustrates exemplary elements of a request 351 for dedicated connectivity
35 from a client, according to one embodiment. As shown, the request, which may be generated at a

client device 148 and may be formatted in accordance with the interface provided for connectivity-related services by connectivity coordinator 114, comprises location information 360, bandwidth requirement 361, availability requirement 363, multi-path requirement 365, client network equipment information 367, and additional specifications 368. Not all these elements may be included in a connectivity request; any combination or subset of these and other elements may be included in the requests in various embodiments. In implementations where an API similar to the requestDirectConnection API described above is used, some or all of the request elements may be provided as fields of a ConnectionSpecification object or its equivalent.

[0042] The location information 360 may include details of a physical location at which the dedicated connectivity is desired: for example a street address where a client-side router 142 currently exists or where such a client-side router may need to be set up, e.g., with the help of a third-party network service provider. In some cases the client may simply list one or more cities or even states where portions of the client network 162 are located and request the connectivity coordinator 114 to provide a set of possible sites where a physical connection could be set up to serve the client network.

[0043] In some implementations the client may specify a desired bandwidth for the dedicated connectivity via bandwidth requirement 361. The interface provided to the client by connectivity provider may, for example, allow the client to choose among a discrete set of bandwidth choices such as 500 Megabits/second, 1 Gigabit/second or 10 Gigabits/second, where the choices may be derived from the details of the specific networking hardware available for establishing a physical link to an endpoint router 132. For example, at some router co-location facilities, the choices for physical links may include 1Gbps 1000BASE-LX (1310nm) single-mode fiber connections over single-mode fiber, and 10Gbps 10GBASE-LR (1310nm) single-mode fiber connections over single-mode fiber, and the connectivity coordinator 114 may allow the client to choose between the 1Gbps option and the 10Gbps option. In other cases the client may be allowed to request any arbitrary bandwidth and the connectivity coordinator 114 may respond to the request by indicating the bandwidth it is able or willing to provide. In one implementation the connectivity coordinator may not provide any guarantees of bandwidth availability and instead, for example, indicate to the client that a best-effort approach will be used – i.e., the connectivity coordinator will try to provide as much bandwidth (up to the customer's desired limit) as possible. In another implementation, the connectivity coordinator may indicate that more than one physical link may be needed – e.g., if the customer requests 20Gbps and the maximum bandwidth available over a single cable is 10Gbps. It may also be possible to set up multiple physical links distributed over different router co-location facilities 132 in response to a single request for dedicated connectivity – e.g., if a particular client has access to client-side routers 142A and 142B at

respective facilities 132A and 132B, one or more physical links may be set up at each facility if needed or requested. The interface provided by connectivity coordinator 114 may allow clients to specify whether distinct physical locations should be used to provide the desired connectivity, and if so, how many locations should be used.

5 **[0044]** The client may in some embodiments also provide an availability requirement 363 and/or a multi-path requirement 365. The availability requirement may be expressible in any of various metrics such as desired maximum network outage limits (e.g., one hour per year maximum outage time) or mean time between outages. A multi-path requirement 365 may indicate the number of physical links that should be set up between a client-side router 142 and
10 an endpoint router 132. Multiple physical links may for example be requested for performance (e.g., so that traffic from the client network 162 may be load-balanced or otherwise distributed over multiple physical paths, thereby reducing network congestion), for higher availability (e.g., by providing multiple paths, an alternate path may be available as a backup path in case of a failure at one of the physical links), or a combination of performance and availability reasons. In
15 addition to specifying how many physical links are needed, a client may also specify the manner in which traffic is to be distributed among them. In a case where two paths are requested, for example, the client may specify whether they should be established in an active/active mode (e.g., where Border Gateway Protocol (BGP) Multipathing is used to balance load across the two links, and in case of a failure, one link takes over the other's traffic), or in active/standby mode
20 where only one of the links is in use at a time, and the second link is activated only in the event of a failure at the first link. Default choices (e.g., active/active) may be indicated via the interface to the client in some implementations, so that the client need not explicitly specify the type of multi-path setup if the client does not wish to do so. In some cases, indicating a multi-path requirement 365 may negate the need for (or contradict) an availability requirement 363, so the
25 client may be allowed to specify only one of these two types of options.

[0045] In one embodiment, in order for example to further simplify the tasks that the client may need to do to establish connectivity at their end, or to optimize performance, the connectivity coordinator 114 may also be able to provide configuration instructions, suggestions, and/or preferred settings for the specific type of networking equipment that the client may have.
30 In such an environment, a client may provide client network equipment information 367 to connectivity coordinator 114, which may for example consult a database of configuration data (e.g., database 115) to look up configuration instructions for the equipment, and provide configuration suggestions or instructions to the client. If a client indicates via information 367 that they wish to use a particular type or class of router from a particular vendor (e.g., a Cisco
35 router, a Juniper router, or a Yamaha router), for example, the connectivity coordinator may be

able to provide vendor-specific configuration hints for the particular type of router or for a particular version of software running on that particular router. Such hints may include examples of how to configure or verify BGP settings, tunneling-related settings, IKE (Internet Key Exchange) settings, and may also include instructions on how to test that the particular vendor's device is operating effectively. Troubleshooting hints and/or tuning hints such as preferred buffer sizes and the like that may be vendor-specific and/or device-specific may also be provided by connectivity coordinator 114 in some embodiments. Since at least in some environments the provider network 105 may have a large number of clients using a wide variety of networking equipment, the connectivity coordinator 114 may be able to build up a knowledge base covering a wide variety of networking equipment configurations, the kinds of client-side settings that work best with the provider network's own equipment, and so on, which may be very helpful clients that are embarking on the process of linking their client networks 160 to the provider network 105. In some implementations, additional specifications 368 for the desired connectivity may also be included in a client request – e.g., specifications of a desired start time or end time for the dedicated connectivity, or an acknowledgement that a particular BGP version and/or Bidirectional Forwarding Detection (BFD) are supported in the client network 162.

[0046] In various embodiments information similar to that shown in Figure 3 may be communicated in multiple steps to the connectivity coordinator 114 – e.g., first location information and desired bandwidth may be communicated, then the connectivity coordinator may provide a response with a list of possible options, and then from among the possible options the client may choose one option and provide additional specifications in subsequent messages. The information may be transmitted to the connectivity coordinator 114 from the client (or from a third party on behalf of the client) using any available network path – for example a path 175 that may include portions of the public internet. Some or all of the interactions between the client and the connectivity coordinator 114 may be encrypted in various embodiments. In some cases where the client does not currently have a client-side router already available at an appropriate router co-location facility 150, further interactions may be required between the client and the connectivity coordinator 114, wherein for example the connectivity coordinator provides suggestions for third-party network service providers that the client may be able to use to obtain access to a suitable router.

[0047] Figure 4 illustrates an example of constituent elements of a response that may be generated to a request for dedicated connectivity from a client, according to at least some embodiments. The illustrated example shows the connectivity coordinator 114 sending a response 451 back to the requesting client device 148, and also an optional notification 452 that may be sent to an operator or manager of a router co-location facility 150 in some

implementations. Having examined the various parameters or properties of the dedicated connectivity requested by the client as illustrated in Figure 3, the connectivity coordinator 114 may eventually decide on a particular endpoint router 132 that may be appropriate for a physical link to be set up to the client's network. For example, in Figure 1, endpoint router 132A at router co-location facility 150A may be chosen to provide physical connectivity to client network 162A. Response 451 may include any combination of physical link configuration instructions 471, authorization information 482, a connection identifier 482, and device-specific configuration instructions 483. The physical link configuration instructions 471 may for example pinpoint the exact physical coordinates where a cable originating from a client-side router such as router 142A is to be attached: an identification 467 of the physical port (e.g., "port 3" or "the third port from the left"), a cage identifier 461, a rack identifier 463, and a patch panel identifier 465.

[0048] In many cases networking equipment such as routers 132 and 142 are housed in secure environments where not everyone may have physical access. In such cases, authorization information 481, which may for example comprise a legally-binding agreement to allow a technician to access the endpoint router 132A may be provided to the client. In some environments a document similar to or derived from a commonly-used standard authorization communication format called "LOA-CFA" (Letter Of Authority and Customer Facility Assignment) may be used for authorization information 481. Authorization information 481 may itself include the physical link coordinates such as port identifier 467, cage identifier 461, rack identifier 462, and patch panel identifier 465 in some cases. The response 451 may also include a connection identifier 482 corresponding to the requested dedicated connectivity, which may be used in further communications from the client to the connectivity coordinator 114, such as a request for establishment of logically-isolated paths via the setUpLogicalConnection API described earlier and discussed further in conjunction with the description of Figure 5 below.

[0049] In some embodiments the connectivity coordinator 114 may also provide configuration instructions 483 for client-side network equipment. Such instructions may be provided in cases where client network equipment information 367 was earlier provided to the connectivity coordinator 114, and may also be provided for a default set of devices (e.g., the most commonly used types of routers) even the client did not previously provide details of the client-side equipment in some implementations. Depending on the specifics of the endpoint router 132 selected for the physical connection, different sets of client-side configuration settings may in general be appropriate even for a given piece of client-side networking equipment, and the connectivity coordinator may consult its configuration knowledge base to choose the appropriate instructions after the endpoint router 132 has been selected.

[0050] As described earlier, authorization may be required to set up physical connectivity to an endpoint router 132 in some environments. In some embodiments, authorization information 481 may also (or instead) be sent to an operator 433 of the router co-location facility 150 by connectivity coordinator. In some jurisdictions legal constraints may prevent such direct communication between the connectivity coordinator 114 and co-location facility operators 433, in which case the authorization information may, if required, be provided by the client to the operator 433.

[0051] In many cases a client may be interested in using resource collections 120 for a variety of different purposes – for example, a software vendor may wish to use one set of resources 112A to set up a development and build environment for its software engineers, another set of resources 112B for an intranet for storing and sharing corporate information internally within the company, and a third set of resources 112C (not shown in Figure 1) for a web site that may be accessed by the software vendor's customers. Such a client may desire, for example for administrative purposes, accounting/billing purposes, and/or security purposes, that the network traffic for each set of resources 112 be isolated from the traffic for the other sets of resources 112. For example, the software vendor may wish to ensure that the build-related traffic be kept separate from the intranet traffic, that traffic from the build machines or resources 112A may not be permitted to reach one or more intranet servers 112B, and so on. At the same time, such a client may wish to utilize the same dedicated physical connectivity provided via an endpoint router 132 for all these different functions, i.e., the client may wish to establish multiple logically-isolated network paths that all share the same physical link similar to cross-network link 191 established for dedicated connectivity to resource collections 120. In some embodiments, the interface set up by connectivity coordinator 114 may be capable of providing support for various operations related to such logically-isolated paths, such as creating, modifying, deleting, and retrieving or querying the state of the paths.

[0052] Figure 5 illustrates an example of constituent elements of an isolation request 551 for a logically-isolated network path that may be sent to connectivity coordinator 114, according to at least some embodiments. Prior to making a request for a logically-isolated network path, a client may have established a physical link to obtain dedicated connectivity, as illustrated in high-level step 211 of Figure 2, and may have obtained a connection identifier 482 during the physical-link establishment process. That connection identifier may be included in the request 551 in the illustrated embodiment. Request 551 may also comprise various selection criteria such as any combination of a VLAN tag 501, a BGP ASN 511, a set of network prefixes 521, pairing information 531, virtual private gateway information 541, and/or other information that may be useful in network isolation in various embodiments.

- [0053]** A Virtual Local Area Network (VLAN) is a method often used for creating multiple logically-isolated networks within a single physical network. A tag or identifier called a VLAN tag may be inserted into the header of each packet being transmitted within a given VLAN environment to enable switches or other network devices to identify the VLAN to which the packet belongs. In one embodiment, connectivity coordinator 114 may require the client to provide a unique VLAN tag 501 for each logically-isolated network path that the client wishes to establish, i.e., a client may not be permitted to use the same VLAN tag for multiple logically-isolated paths. In one implementation the VLAN tag 501 may be required to comply with a standard, such as the Ethernet 802.1q standard.
- 10 **[0054]** A client may also be required to provide a BGP Autonomous System Number (ASN) 511. As noted earlier, an Autonomous System (AS) is a collection of connected Internet Protocol (IP) routing prefixes under the control of one or more network operators that presents a common, clearly defined routing policy to the Internet. A unique ASN is typically allocated to each AS for use in BGP routing. ASN 511 may be public (i.e. may be exposed to various routers of the public
- 15 Internet) or private (exposed only to routers of the provider network 100 and the client network 162), depending on the type of logical connectivity the client wishes to establish in various embodiments. The client may also provide a set of network prefixes 521 to be advertised for the logically isolated network, e.g., in accordance with BGP or another routing protocol. Pairing information 531, indicating for example whether the desired logically-isolated path is to be
- 20 paired in an active/active or active/standby mode with any other path, may also be included in request 551 in some embodiments. In some implementations the provider network may support establishment of virtual private gateways to support VPN (virtual private network) functionality between a client network 162 and resource collections 120, and request 551 may also include an identification of such a virtual private gateway to be used for the logically-isolated network path.
- 25 In some embodiments Multiple Protocol Label Switching (MPLS) techniques may be used to implement logical network isolation. While the exemplary elements illustrated in Figure 5 may be applicable in environments where BGP and related protocols are in use, in other embodiments other network isolation mechanisms (e.g. any other techniques usable for connecting to virtual private clouds or VPNs) may be supplied by the client and used by connectivity provider for
- 30 logical network isolation. In the example setUpLogicalConnection API call described earlier, some or all of the various elements of request 551 may be included for example in fields of the LogicalConnectionParameters object passed as a parameter.
- [0055]** In one embodiment, after receiving the request 551 for establishing a logically-isolated network path, connectivity coordinator 114 may perform a set of operations, such as
- 35 updating connectivity database 115, propagating appropriate routing information to various

routers of provider network 105, refreshing various routing-related caches, and the like, to complete the requested configuration. After establishing the logically-isolated network path successfully, in some embodiments the connectivity coordinator 114 may send a confirmation message back to the client indicating that the requested dedicated connectivity and/or logical isolation has been successfully provisioned. Figure 6 illustrates an example of contents of such a confirmation message 651 indicating that requested connectivity has been established, according to at least some embodiments. In the illustrated example the physical connection confirmation details 601 may confirm some of the information related to the physical link established at the request of the client, such as port identifier 467, rack identifier 463, available bandwidth, etc. Logical connection confirmation details 621 may confirm properties of the logically-isolated network paths, such as VLAN tag 501, BGP ASN 511, network prefixes 521, pairing information 531, and virtual private gateway information 541. In the illustrated example confirmation message 651 also includes the connection identifier 482 and support information 611 – e.g., information that the client may use to obtain help in case of a traffic interruption, poor performance, or other issue that may arise. Confirmation messages 651 may exclude any combination of the elements shown in Figure 6 in different embodiments, and may include additional information in some embodiments. In one embodiment multiple confirmation messages may be sent by connectivity coordinator 114 – e.g., a first confirmation message may be sent after the physical link is established, and a second confirmation message may be sent after the logically-isolated network path has been established. Connectivity coordinator 114 may also send instructions to the client to verify or validate that the desired connectivity is functioning correctly at the client end – e.g., in an embodiment where resources 112 include virtual compute servers with associated public and/or private IP addresses, such instructions may direct the client to start up a virtual compute server and ping one of its IP addresses.

25

Example of logically-isolated network paths over shared physical link

[0056] Figure 7 illustrates an example of two logically-isolated network paths 752A and 752B sharing a single dedicated physical link such as a cross-network connection established using an interface provided by connectivity coordinator 114, according to at least some embodiments. In the environment shown in Figure 2, the client requires connectivity to be established and maintained between internal network 732 and a restricted-access resource farm 712. At the same time the client has set up a de-militarized network zone (DMZ) 722 (which may also be termed a perimeter network) – a sub-network of client network 162A that may expose some of the client's services to the public or untrusted Internet via public-access resource farm 702 within provider network 105's data centers. To ensure that the traffic for both the

35

restricted-access resource farm 712 and the public-access resource farm 702 meets desired performance, security and cost requirements, the client may first use the interface provided by connectivity coordinator 114 to establish a cross-network connection 791 between client-side router 142A and endpoint router 132A, using for example the steps described in Figure 2. The client may further use other components of the interface to establish two logically-isolated network paths that share the cross-network connection 791: path 752A for traffic between DMZ 722 and the public-access resource farm 702, and path 752B for traffic between the client's internal network 732 and restricted-access resource farm 712.

[0057] In some embodiments multiple dedicated physical links such as cross-network connections 791 or 191 may be set up on behalf of a single customer, either within one router co-location facility 150 or across multiple router co-location facilities. For example, a multinational corporation may have office facilities in several different countries, which may all benefit from dedicated connectivity to a set of resource collections 120; in such a case, one or more dedicated physical links may be set up for respective geographically separated office locations. A single physical link may be shared across numerous logically-isolated paths such as paths 752 of Figure 7. Furthermore, a given resource collection such as a resource farm 702 or 712 may be accessible via a plurality of logically-isolated paths 752, where some of the logically-isolated paths 752 may use different dedicated physical links 791.

20 Example of web-based interface

[0058] Figure 8 is an illustration of a portion of an exemplary web-based interface that may be provided by connectivity coordinator 114 in some embodiments. As noted earlier, the interface implemented by connectivity coordinator 114 to provide connectivity services may be exposed to clients as a set of web pages in some embodiments. Web page 800 of Figure 8 is a representation of an example of one such web page that includes several form fields that a client may fill out to provide details about desired dedicated connectivity requirements. In some implementations the submission of form data via an interface like web page 800 may result in an invocation of one or more API calls similar to those listed earlier in conjunction with the description of element 201 of Figure 2.

[0059] In area 803 of web page 800, a friendly greeting and overview message may be provided. Form fields 805 may be provided to allow the client to specify a physical location where the dedicated connectivity is desired. Using form field 807, the client may specify desired bandwidth, for which a default value of 1Gbps is shown pre-selected in Figure 8. Form fields 809 may be used to provide optional pairing or multi-path information; as shown, a default of two connections in active/active mode is pre-selected. For fields 811 may allow the client to specify a

vendor name and model for a client router to be used for a dedicated physical link. Form field 813 may allow the client to identify a network service provider that may also be involved in setting up the dedicated connectivity – for example, an operator of the router co-location facility that may be used. In some embodiments, when the client fills in the address information in form fields 805, the connectivity coordinator 114 may automatically fill out the network service provider form field 813, or may populate a set of drop-down options from which the client may select a preferred provider via form field 813. The client may submit the completed form using submit button 815 in the illustrated example. In some implementations employing a web page interface, several different web pages may be employed during the process of establishing the desired physical and logical connectivity. As the client fills out one form entry, the connectivity coordinator may be able to customize or narrow the set of options available for subsequent form entries.

[0060] Figure 9 is a flowchart of a method for providing connectivity-related services, according to at least some embodiments. As shown in element 900 in the flowchart, an interface that defines a set of connectivity operations may be implemented, for example by a connectivity coordinator 114. The connectivity operations provided via the interface may include services to set up, query, modify, disable and tear down various types of physical and logical connections in various embodiments. The interface may comprise any combination of a set of APIs, a web-based or standalone GUI, command-line tools, and the like.

[0061] A request for dedicated connectivity may be received in accordance with the interface, as shown in element 910. For example, in an environment where the interface is web-based, the request may comprise one or more HTTP or HTTPS requests, while in a different embodiment, the request may comprise one or more method calls from a program coded and executed on behalf of the client. The request may comprise an enumeration of several details that may be needed to make a decision as to where and how the dedicated connectivity may be provided, and which business entities such as third-party network service providers or network data center operators may need to be involved. For example, the request may specify a desired physical address at which a client-side router 142 is available for use, a desired bandwidth, and various other requirements.

[0062] On receiving the request, a target endpoint router 132 of a provider network 105 may be selected, through which a route to provide the desired dedicated connectivity to the client may be configurable, as shown in element 920 of Figure 9. The target endpoint router may be selected based on any of a variety of factors in different embodiments, including physical location, measured and/or expected bandwidth utilization levels, costs, previous positive or negative

experiences with the operator of the facility where the router is located, compatibility with the client's networking equipment and the like.

[0063] A set of configuration information and instructions may then be generated for setting up a physical link to the target endpoint router, as shown in element 930, and a response may then be transmitted (element 940). In some embodiments the response may be submitted only to the requesting client, while in other embodiments a response may be submitted to an operator of a router co-location facility 150 where the physical link is to be established, or responses may be submitted to both the requesting client and the facility operator. The response may include data identifying the particular physical port, cage, rack, and/or patch panel where a physical cable may be attached in some implementations. Authorization information, e.g., granting permission to a technician to access the endpoint router may be included in the response, or may be made accessible via the response.

[0064] In one implementation, after the physical link is established, a confirmation message indicating that the desired connectivity has been established successfully may be transmitted to the client (element 950 of Figure 9). In other implementations, a confirmation message may be generated after one or more logically-isolated network paths have been established using the newly established physical link.

[0065] Figure 10 is a flowchart of a method for providing connectivity-related services including guidance on configuring a client network device, according to at least some embodiments. An interface that allows a client to make a variety of connectivity-related requests, including requests for assistance in configuring one or more network devices that may be used to establish dedicated connectivity with a provider network, may be implemented as shown in element 1000. A request that provides an identification of networking equipment (e.g., any combination of a vendor name, a model name, and a software version identifier for software running on the networking equipment) available for use by the client may be received (element 1010) in accordance with the interface. Such a request may also include other details of the connectivity requested by the client, such as a desired bandwidth, availability/redundancy requirements, and the like.

[0066] In response to the request, in some implementations a connectivity coordinator 114 may query a database of configuration information, for example using a combination of vendor name, model name, software version and/or connectivity requirements (element 1020). If appropriate configuration guidance is found, e.g., based on the identification information provided in the request, a response containing the configuration information or instructions may be generated (element 1030) and transmitted to the requesting client (element 1040). In some implementations, the database of configuration information may include an inventory of where

(i.e., by which clients) different types of networking equipment are being used; in such a case, a record indicating that the requesting client uses the specified equipment may be inserted into the database (element 1050). In some environments additional information on experiences with different types of networking equipment, such as surveys of client satisfaction with their
5 networking equipment, mean times to failure, availability data, and the like may also be maintained in a knowledge base by connectivity coordinator 114, and some or all of this additional information may also be made available via the interface.

Interactions with last-mile connectivity providers

10 [0067] Figure 11 illustrates an example of a system 1105 including a provider network 1100 to which connectivity may be provided from client networks such as 1162A and 1162B through “last-mile” connectivity providers (e.g., 1150A, 1150B, and 1150C), according to at least some embodiments. In many environments, client devices such as 1148A and 1148B may be provisioned within networks (e.g., 1162A and 1162B) that may not have private paths available
15 from their client-side routers 1142 to router co-location facilities (similar to facilities 150 of Figure 1) where endpoint routers such as 1132A and 1132B may be located. This may be especially likely in the case of relatively small client businesses, or when client business premises are located in areas that are somewhat remote from router co-location centers. Such client networks 1162 may have access via shared network paths (e.g., the portions of the public
20 Internet, including for example portions of path 1175) to various resource collections 1120 of the provider network 1100, but the operators of the client networks may wish to avail themselves of the advantages of dedicated paths to the resource collections. Various third party connectivity providers 1150 (i.e., business entities other than the provider network’s operator) may be capable of providing the dedicated paths to the endpoint routers 1132 – for example, in Figure 11,
25 connectivity provider 1150C is shown providing a dedicated or direct path 1149 between endpoint router 1132B and client network 1162B. Such connectivity providers may help clients to bridge the gap between the client networks 1162 and the private paths 1170 (e.g., paths 1170A, 1170B, 1170C and 1170D, similar to paths 170 of Figure 1) available between the endpoint routers 1132 and the resource collections 1120. These third party connectivity providers
30 may be referred to as “last-mile” connectivity providers (or “last-kilometer” connectivity providers in environments where metric distance units are more popular), as they are often responsible for implementing physical network connectivity closest to the client premises, and therefore furthest from the premises of major network infrastructure providers. In this document, last-mile connectivity providers may also be referred to using the abbreviation “LMCP”.

[0068] Identifying which, if any, last-mile connectivity providers may be available and willing to link a client's network to the provider network 1100 may often be cumbersome from a client's perspective. In some cases, a number of LMCPs may operate in the vicinity of the client's premises, but only a subset may be supported or preferred by the operator of provider network 1100. In the embodiment illustrated in Figure 11, connectivity coordinator 1114 may be operable to implement an interface defining a variety of connectivity-related services, which may allow clients to easily determine which LMCPs 1150 can be used to connect to provider network 1100. Such an interface may further allow the clients to establish desired dedicated connectivity (e.g., over a direct path 1149) to resource collections 1120 with the help of selected last-mile connectivity providers. Connectivity coordinator 1114 may implement one or more databases 1115 to store connectivity-related information, including for example a directory of last-mile connectivity providers 1150 and their offerings. The interface may be published or made available to clients by connectivity coordinator 1114 using any appropriate technique, such as one or more email messages to all the clients of the provider network 1100, establishing a web site or web page with the details of the interface, and so on. The interface itself may, for example, be programmatic, and may comprise any combination of a collection of APIs, one or more web pages, command-line tools, an installable graphical user interface, or the like. Connectivity coordinator 1114 may itself be resident outside provider network 1100 in some embodiments, and in other embodiments it may be incorporated within the provider network 1100.

[0069] Using the interface, e.g., from one of the client devices 1142A, a client may submit a request for dedicated connectivity, e.g., over a path 1175 that may include links of the public Internet. The request may, for example, include the physical address or addresses at which the client desires dedicated connectivity. In response to the request, the connectivity coordinator may be operable to identify one or more LMCPs 1150 that may be available to establish dedicated connections between the provider network 1100 and the requesting client's network (e.g., 1162A), and generate and transmit a response that lists the selected LMCP or LMCPs. The selected LMCP may operate or manage one or more routers that happen to be co-located with one of the endpoint routers 1132 of the provider network 1100, or may have the ability to set up such routers if they are not already available. In some embodiments the connectivity coordinator 1114 may allow the client to select one LMCP from among a set of available LMCPs, while in other embodiments the connectivity coordinator 1114 may determine the specific LMCP or LMCPs that should be used, and inform the client of the determination. Later, after the dedicated connectivity has been set up, for example via steps similar to those outlined in Figure 2, the connectivity coordinator 1114 may in some embodiments provide a confirmation to the client

indicating that the desired connectivity has been verified. The interface may be used for communication between the client and the connectivity coordinator 1114 during any of the stages of dedicated connectivity establishment and use – e.g., a client may query the state of a requested connection or an established connection using the interface, and may request various connectivity modifications, disabling and enabling of connectivity, and the like. Responses to the client requests may also be formatted in accordance with the interface.

[0070] The connectivity coordinator 1114 may, for example, look up LMCP information in database 1115 to respond to the initial request for dedicated connectivity. In cases where multiple LMCPs 1150 are available, the connectivity coordinator 1114 may in some implementations provide an unordered enumeration of all available LMCPs to the client. In other implementations the available LMCPs may be sorted according to any of a variety of criteria based on details of the client's request and on the connectivity coordinator's LMCP knowledge base. For example, if the connectivity provider 1114 is aware of service quality rankings or ratings of the various LMCPs, it may sort the LMCPs from highest to lowest quality. If the connectivity provider 1114 has pricing information available for the different LMCPs it may rank them according to price, and so on. Connectivity coordinator 1114 may in some implementations periodically query clients for quality rankings or ratings of different LMCPs and store the results of such surveys in its database 1115, or it may monitor outages or support requests to establish its own quality rankings. In one implementation in which clients may specify desired connectivity establishment times (e.g., the logical equivalent of "I need this connectivity by August 1, 2011 at 8AM EST"), the connectivity coordinator may be able to eliminate some LMCPs from the list of available LMCPs based on how quickly the LMCPs have been known to establish connectivity in the past. In some embodiments the interface supported by the connectivity coordinator 1114 may allow clients to query its knowledge base of LMCPs. In addition to the paths illustrated in Figure 11, in many environments there may be several other alternative network paths available between the client networks 1162 and various components of the system 1105 – for example, the connectivity requests may be transmitted to the connectivity provider 1114 over a path that includes public or shared links, and various services provided at resource collections 1120 may be accessed over public or shared links as well.

30

LMCP-related communications with the connectivity coordinator

[0071] Figure 12 illustrates an example of constituent elements of an initial request 1251 for dedicated connectivity through a connectivity provider 1150, according to at least some embodiments. As shown the request comprises location information 1260 for the client network 1162, and optional connectivity details 1261, an optional start time 1268, and an optional end

35

time 1269. The connectivity coordinator 1114 may use location information 1260 as the primary criterion for querying its LMCP database to identify available LMCPs. Optional connectivity details 1261 may include requirements similar to those shown in Figure 3, for example, bandwidth requirement 361, availability requirement 363, and/or multi-path requirement 365. In some embodiments the client may also specify a desired start time 1268 and/or a desired end time 1269 – for example, indicating that the dedicated connectivity is only going to be needed for 3 months starting on January 1, 2011. In some cases the start and end times may indicate that the client only desires to use the dedicated connectivity part of the time – e.g., the start time and end time may be specified as “8AM – 8PM, Monday through Friday”. In some implementations where desired start times 1268 are indicated by the client, end times 1269 may not be required. Timing requests comprising desired start and/or end times may be sent separately from the initial request 1251 in some embodiments.

[0072] Figure 13 illustrates an example of constituent elements of a response 1301 to a dedicated connectivity request that includes information on available connectivity providers 1150, according to at least some embodiments. The response 1301 may comprise a list of one or more LMCP details records 1361, e.g., 1361A and 1361B, that the connectivity coordinator 1114 may have found to meet the request 1251 from the client. Different types of information regarding the available LMCPs may be provided to the client in various embodiments. For example, LMCP details record 1361A may comprise an identification (e.g., name and contact information) 1311A of the LMCP 1150A, pricing information in fields 1321A and 1321B, estimated connectivity establishment time 1341A, and/or a satisfaction rating 1351A. Pricing information may be split into a recurring price component 1321A (e.g., “X dollars per month, independent of the actual usage”) and a non-recurring price component 1331A (e.g., based on measured bandwidth usage by the client). In some implementations the pricing information may be broken down further into components that are to be paid to the LMCP 1150A directly by the client, and components that are to be paid to the operator of provider network 1100. The interface supported by coordinator 1114 may allow clients to submit pricing-related queries as separate requests in some implementations. The earliest time at which the LMCP 1150A and/or the provider network 1100 operator may be able to establish the desired dedicated connectivity may be indicated via field 1341A. In some cases a satisfaction rating 1351A (e.g., based on surveys of the clients of the LMCP 1150A) may be included, which may be helpful to the client in selecting among available LMCPs. Details record 1361B may include similar fields as record 1361A for a different LMCP, e.g., for LMCP 1150B.

[0073] Figure 14 illustrates an example communication from a client identifying a selected connectivity provider 1150A, that may be generated by the client after receiving a response 1301,

according to at least some embodiments. Selection notification 1451 may also be formatted in accordance with the interface implemented by connectivity coordinator 1114, e.g., as an API call or a web form selection. Figure 15 illustrates example communications from connectivity coordinator 1114 to an LMCP 1150 and the client after the client has selected the LMCP, according to one embodiment. As shown, in response 1551 to the client, the connectivity coordinator 1114 may provide confirmation 1583 of the selection of the LMCP. In one embodiment, the connectivity coordinator 1114 may determine the LMCP 1150 to be used, e.g., based on the client's location information, and may not require that the client make a selection; in other embodiments, the connectivity coordinator 1114 may wait for the client's selection or confirmation before making the determination of the LMCP. In some implementations the client may suggest or recommend one or more LMCPs in the initial request for dedicated connectivity (e.g., request 1251 of Figure 12), and the connectivity coordinator 1114 may determine the LMCP to be used based on the initial request. A connection identifier 1581 may also be provided to the client in response 1551. In some embodiments authorization information 1582 allowing physical access to the network provider's endpoint router 1132, similar to the authorization information 481 shown in Figure 4, may be provided to the client as well. In notification 1552 sent to the selected LMCP 1150, connectivity coordinator 1114 may also provide authorization information 1582, as well as physical link configuration instructions 1571 (similar to physical link configuration instructions 471 of Figure 4) which may include port, cage, rack and/or patch panel identifiers 1567, 1561, 1563 and 1565 respectively for the endpoint router 1132 to which a physical link may be established by or on behalf of the selected LMCP. In some embodiments the authorization information 1582 may be sent to either the client or the LMCP, but not both.

[0074] After communications similar to those shown in Figure 15 are received by the client and/or the selected LMCP 1150, a path (similar to direct path 1149 of Figure 11) comprising a dedicated physical link may be established between equipment of the provider network (such as an endpoint router 1132) and the client network 1162 using the selected LMCP's network and/or equipment in one embodiment. As desired, one or more logically-isolated paths, similar to those discussed in conjunction with the description of Figures 5 and 7 may then be established using the newly-established dedicated physical link. In some implementations, the connectivity coordinator 1114 may verify, e.g., by exchanging one or more network packets or messages with the client and examining the routes taken by the messages, that the desired dedicated connectivity has been provided, and may send a confirmation message to the client and/or the LMCP 1150 indicating this accomplishment.

[0075] In environments where timing constraints (such as start times and/or stop times) were requested by the client, connectivity provider 1114 may also be operable to implement those

timing constraints, e.g., by scheduling routing changes or making other configuration changes at the appropriate times. For example, in one environment where BGP is used as a routing protocol, a set of network prefixes for the client network may be advertised at or shortly prior to a desired start time to enable or disable the dedicated connectivity. In some implementations, the dedicated

5 connectivity provided to a client with the help of an LMCP may have an associated bandwidth limit, and the interface supported by the connectivity coordinator 1114 may also enable clients to make bandwidth modification requests – e.g., to request a higher or lower rate of traffic than initially agreed upon. In response to such requests the connectivity coordinator 1114 may dynamically change configuration settings at one or more devices to comply with the new

10 bandwidth requirement. In one embodiment the connectivity coordinator 1114 may itself monitor the rate at which traffic flows to a client network 1162. If traffic over a measured period of time reaches a threshold (e.g., 80% or more of the maximum allowed traffic), the connectivity coordinator 1114 may inform the client that a bandwidth modification may be advisable, and the client may in turn request a bandwidth increase using the interface. In some implementations the

15 connectivity coordinator 1114 may also be configured to notify a client if it may be advisable to lower the bandwidth limit associated with a dedicated direct path, e.g., if measurements indicate that the client appears to use only a small fraction of the bandwidth it requested.

Example web interface for initiating LMCP selection

20 **[0076]** Figure 16 is an illustration of a portion of an exemplary web-based interface that may be provided for initiating connectivity provider selection, according to some embodiments. As shown, the web-based interface may comprise a web-page 1600 with several form fields that may be presented to the client by connectivity coordinator 1114. The web page may include a welcome message area 1603, and form fields 1605 for the client to specify a physical address

25 where dedicated connectivity is desired. In fields 1607, the client may indicate whether assistance in selecting a connectivity provider is desired. Bandwidth requirements may be specified in fields 1609, and timing-related requirements for the dedicated connectivity, such as a desired start time and/or a desired end time, may be specified in field 1611. The submit button 1615 may be used to submit the completed form to the connectivity coordinator 1114.

30 **[0077]** In one embodiment, the submission of such a form may result in the invocation of one or more APIs at the connectivity coordinator similar to APIs API-1 through API-18 listed in conjunction with the description of Figure 2. Some additional APIs, including some LMCP-specific APIs and APIs to provide dynamic modification of existing connections and/or pricing-related operations may also be supported in some implementations, for which example

35 invocations may include the following:

[API-21] ***ProviderList providerList = getConnectionProviders(CustomerID customerId, CustomerLocationRecord location);***

The getConnectionProviders API may be used to find available LMCPs based on the location information specified in a CustomerLocationRecord object.

- 5 [API-22] ***RequestStatus status = setConnectionProvider(ProviderId providerId, CustomerID customerId, RequestId requestId);***

The setConnectionProvider API may be used to specify that a particular LMCP identified by its ProviderId has been selected by the customer.

[API-23] ***RequestStatus status = setConnectionStartTime(ConnectionId connectionID);***

- 10 The setConnectionStartTime API may be used to specify a start time for the dedicated connectivity.

[API-24] ***RequestStatus status = setConnectionEndTime(ConnectionId connectionID);***

The setConnectionEndTime API may be used to specify a desired end time for dedicated connectivity.

- 15 [API-25] ***PricingInfo pricingInfo = getConnectionPricingInfo(ConnectionId connectionID);***

The getConnectionPricingInfo API may be used to query pricing-related information for an existing connection or for a connection that has not yet been established.

- [0078] In some embodiments, APIs such as those for which examples are provided above may be available for use directly by clients, LMCPs 1150 and/or other network providers, or by facility providers such as operators of router co-location facilities 150. In some implementations multiple layers of interfaces may be supported, allowing clients to request some connectivity-related operations using a web interface, for example, and to perform or request other operations using an API.

- [0079] Figure 17 is a flowchart of a method for enabling clients to select connectivity providers, according to at least some embodiments. The method (as shown in element 1700 of Figure 17) comprises implementing an interface that defines a set of connectivity operations made available to clients of a provider network 1100 by a connectivity coordinator 1114. The interface may comprise an API, a command-line interface, a web-based interface, some other GUI, or any other programmatic interface, for example. A request for dedicated connectivity, formatted in accordance with the interface, may be received (element 1702). The request may in some cases directly indicate that the requester requires assistance in selecting a connectivity provider. In other cases the request may simply contain some information (such as a physical address where a client wishes to obtain dedicated connectivity) from which the connectivity coordinator 1114 may infer, based on its knowledge of where endpoint routers 1132 of provider network 1100 are physically located, that a connectivity provider may be needed to fulfill the

client's request. As indicated in element 1704, one or more connectivity providers may then be selected, and a response identifying the selected connectivity provider or providers may be generated (element 1706 of Figure 17) and transmitted (element 1708). In some implementations the response may be transmitted to just the requesting client, while in other implementations a
5 response and/or notification may also or instead be transmitted to the selected connectivity provider.

[0080] When the client receives the information identifying candidate connectivity providers such as LMCPs 1150, it may select one (if more than one were identified by the connectivity coordinator 1114) and notify the connectivity coordinator of its choice. The connectivity
10 coordinator 1114 may then communicate with the selected LMCP 1150, and with the client, to coordinate the establishment of a physical link (and in some cases one or more logical connections that use the physical link) to fulfill the client's connectivity needs. After connectivity has been successfully established, the connectivity coordinator 1114 may in some implementations send a confirmation of the completion of the establishment (element 1710 of
15 Figure 17).

[0081] In some embodiments a number of additional capabilities, in addition to connectivity provider selection and connectivity establishment, may be supported via the interface provided by a connectivity coordinator 1114. Figure 18 is a flowchart of a method for providing dynamic connectivity-related services, according to at least some embodiments. The connectivity
20 coordinator 1114 may wait for connectivity-related requests from clients, as shown in element 1800. When such a request is received, if the request is a supported type of request, the connectivity coordinator 1114 may take the requested action. For example, if a request for dynamic enablement of connectivity is received and supported in the embodiment (element 1810), the connectivity coordinator may enable traffic flow, as shown in element 1815. In some
25 cases enabling or disabling traffic may require interaction or coordination between the connectivity coordinator 1114 and an LMCP 1150. Similarly, if the request is for disabling connectivity (element 1820), the desired change may be put into effect by connectivity coordinator 1114 (element 1825). If the request is for a change to the bandwidth limits associated with a client's dedicated connection path (element 1830), the requested change may be
30 implemented, e.g., by making configuration changes at one or more network devices of the provider network 1100 and/or an LMCP. If the request comprises a pricing query (element 1840), the connectivity coordinator 1114 may provide the requested pricing information (element 1845), which may for example include recurring and/or non-recurring cost information that the provider network operator and/or the LMCP may charge the client. If an invalid or unsupported
35 request is received, as shown in element 1850, the connectivity coordinator may transmit a

response indicating that an unexpected request was received. In each case, as the arrows leading back to element 1800 indicate, the connectivity coordinator eventually resumes waiting for connectivity-related requests. While the determination of the type of request is shown as a series of checks in Figure 1800 for simplicity (first checking for dynamic enablement requests, then for disablement requests, and so on), in various implementations the request type may be determined in a single step using logic similar to a “case” or “switch” statement in C or Java

[0082] Figure 19 is a flowchart of a method comprising responding dynamically to changing traffic levels, according to at least some embodiments. As shown in element 1900, the connectivity coordinator may monitor traffic rates over the dedicated path set up for a client. If some threshold of traffic is reached or sustained over a period of time (as detected in element 1910), the connectivity coordinator may provide an indication to the client that a bandwidth change may be appropriate (element 1915). If a request to change the bandwidth is received in response to the indication (element 1920), the connectivity coordinator may implement the requested change (element 1925). The connectivity coordinator may then resume monitoring the traffic. Each of the interactions illustrated in Figures 18 and 19 between the connectivity coordinator 1114 and the client may be implemented using the interface or interfaces (such as one or more web pages) provided by the connectivity coordinator.

Example use cases

[0083] The techniques described above of providing easy-to-use interfaces for dedicated connectivity operations may be used in a variety of environments. For example, if the provider network is expanding quickly across new geographical regions where the reliability, performance and/or security of publicly available networking facilities is limited, more and more clients may wish to utilize dedicated connectivity, especially if it is provided at a reasonable price point. In addition, in cases where a provider network operator may already provide a set of interfaces for managing computation and/or storage resources (such as resource collections 120 or 1120) that are currently accessed via shared (non-dedicated) paths, the provision of additional interfaces to manage dedicated connectivity options may significantly increase the adoption rate of the dedicated connectivity services in which the operator has invested.

[0084] Example embodiments can be described in view of the following clauses:

1. A system, comprising:

a data center including a resource collection designated to respond to service requests received from a client;

a plurality of endpoint routers linked to the data center by one or more private network paths; and

a connectivity coordinator;

wherein the connectivity coordinator is operable to:

implement an interface defining connectivity operations available to the client;

receive a connectivity request from the client for dedicated connectivity to the

5 resource collection, wherein the connectivity request is formatted in accordance with the interface;

in response to the connectivity request,

select a target endpoint router of the plurality of endpoint routers, wherein the target endpoint router is configurable to provide a route over a private network path of the

10 one or more private network paths in accordance with the connectivity request;

generate a reply comprising configuration instructions for a physical network link to be established to the target endpoint router to provide at least a portion of the dedicated connectivity; and

transmit the reply to the client.

15

2. The system as recited in clause 1, wherein the interface comprises at least one of: an application programming interface (API), a graphical user interface (GUI), or a command-line interface.

20 3. The system as recited in clause 1, wherein the target endpoint router is housed within a facility requiring authorization for physical access, and wherein the reply includes an indication of an authorization of physical access to the target endpoint router at the facility.

4. The system as recited in clause 1, wherein the connectivity coordinator is further operable to:

25 receive an isolation request from the client to establish a logically isolated network path to the resource collection via the physical link; and

implement a network isolation mechanism to establish the logically isolated network path in accordance with the isolation request.

30 5. The system as recited in clause 4, wherein the network isolation mechanism comprises at least one of: a virtual local area network (VLAN) mechanism or a Multi-Protocol Label Switching (MPLS) technique.

6. The system as recited in clause 1, wherein the connectivity coordinator is further operable to:

receive, from the client, identification information of a network device to be used to transmit network traffic of the client to the resource collection; and

provide, to the client, one or more configuration instructions for the network device based on the identification information.

5

7. A method, comprising:

presenting a connectivity service to a client of a provider network, wherein said connectivity service includes a connectivity coordinator implementing a programmatic interface defining connectivity operations available to the client;

10 receiving, at the connectivity coordinator, a connectivity request for dedicated connectivity to a resource collection of the provider network, wherein the connectivity request is formatted in accordance with the interface;

in response to the connectivity request, selecting a target endpoint router of a plurality of endpoint routers of the provider network, wherein the target endpoint router is configurable to provide a route over a private network to the resource collection in accordance with the connectivity request;

15 generating a notification comprising configuration information for a physical network link to be established to the target endpoint router to provide at least a portion of the dedicated connectivity; and

20 transmitting the notification.

8. The method as recited in clause 7, further comprising:

receiving identification information of a network device to be used to transmit network traffic of the client to the resource collection; and

25 providing one or more configuration instructions for the network device based on the identification information.

9. The method as recited in clause 7, wherein the resource collection comprises a plurality of resources, further comprising:

30 receiving one or more selection criteria identifying a subset of the plurality of resources to which an isolated connection is to be provided over the physical network link; and

routing network traffic in accordance with the selection criteria.

10. The method as recited in clause 9, wherein a selection criteria of the one or more selection criteria comprises a virtual local area network (VLAN) tag.

35

11. The method as recited in clause 7, wherein the target endpoint router is housed within a facility requiring authorization for physical access, and wherein the notification includes an indication of an authorization of physical access to the target endpoint router at the facility.
- 5
12. The method as recited in clause 7, wherein transmitting the notification comprises sending the notification to an operator of a facility at which the target endpoint router is housed.
13. A non-transitory computer-accessible storage medium storing program instructions that when
- 10 executed on one or more processors:
- implement a programmatic interface defining connectivity operations available to a client of a provider network;
- receive a connectivity request from the client for dedicated connectivity to a resource collection of the provider network, wherein the connectivity request is formatted in accordance
- 15 with the interface;
- in response to the connectivity request, generate a notification comprising configuration information for a physical network link to be established to a target endpoint router of the provider network to provide at least a portion of the dedicated connectivity, wherein the target endpoint router is configurable to provide a route over a private network path to the resource
- 20 collection in accordance with the connectivity request; and
- transmit the notification.
14. The non-transitory computer-accessible storage medium as recited in clause 13, wherein the program instructions when executed on the one or more processors:
- 25 after the physical network link has been established, transmit a confirmation message indicating that the dedicated connectivity has been provided.
15. The non-transitory computer-accessible storage medium as recited in clause 13, wherein the configuration information comprises at least one of: a physical port of the target endpoint router,
- 30 a rack identifier, a cage identifier or a patch panel identifier.
16. The non-transitory computer-accessible storage medium as recited in clause 13, wherein the connectivity request comprises one or more of: a bandwidth requirement, an availability requirement, or a requirement for a plurality of physical paths to the resource collection.

17. The non-transitory computer-accessible storage medium as recited in clause 13, wherein the program instructions when executed on the one or more processors:

receive identification information of a network device to be used to transmit network traffic of the client to the resource collection; and

5 provide one or more configuration instructions for the network device based on the identification information.

18. The non-transitory computer-accessible storage medium as recited in clause 13, wherein the resource collection comprises a plurality of resources, wherein the program instructions when
10 executed on the one or more processors:

receive one or more selection criteria identifying a subset of the plurality of resources to which an isolated connection is to be provided over the physical network link; and

generate routing information to route network traffic in accordance with the selection criteria.

15

19. The non-transitory computer-accessible storage medium as recited in clause 18, wherein a selection criteria of the one or more selection criteria comprises a virtual local area network (VLAN) tag.

20 20. The non-transitory computer-accessible storage medium as recited in clause 13, wherein the target endpoint router is housed within a facility requiring authorization for physical access, and wherein the notification includes an indication of an authorization of physical access to the target endpoint router at the facility.

25 21. The non-transitory computer-accessible storage medium as recited in clause 13, wherein the configuration information comprises an identification of a physical location of the target endpoint router.

22. A system, comprising:

30 a plurality of resource collections of a provider network, including a first resource collection within a first geographical zone of the provider network and a second resource collection within a second geographical zone of the provider network;

an endpoint router within the first geographical zone, linking the first resource collection to a client network of a client via a dedicated physical network link; and

35 a connectivity coordinator;

wherein the connectivity coordinator is operable to:

implement an interface to receive connectivity requests from the client;

receive a connectivity request from the client to establish a logically isolated network path to the second resource collection, wherein the connectivity request is

5 formatted in accordance with the interface; and

perform one or more configuration operations to enable traffic to flow from the client network to the second resource collection over a logically isolated network path using the dedicated physical network link.

10 23. The system as recited in clause 22, wherein the connectivity coordinator is further operable to:

in response to the connectivity request, send instructions to the client to transmit connectivity metadata associated with the logically isolated network path to a destination address within the second geographical zone; and

15 prior to performing the one or more configuration operations, verify that the connectivity metadata was transmitted in accordance with the instructions.

24. The system as recited in clause 22, wherein the connectivity coordinator is further operable to:

20 implement a first pricing policy for the first geographical zone, and a second pricing policy for the second geographical zone; and

provide an indication of the first and second pricing policies to the client in accordance with the interface.

25 25. The system as recited in clause 22, wherein the connectivity coordinator is further operable to:

provide an enumeration to the client of one or more resource collections to which logically isolated connections may be established over the dedicated physical network link, wherein the enumeration is formatted in accordance with the interface.

30

26. The system as recited in clause 22, wherein the interface comprises one or more of: an application programming interface, a command line interface, a graphical user interface or a web interface.

35 27. A method, comprising:

- presenting a connectivity service to a client of a provider network, wherein the provider network comprises a first geographical zone including a first resource collection allocated to the client, and a second geographical zone including a second resource collection allocated to the client, wherein said connectivity service includes a connectivity coordinator implementing a
- 5 programmatic interface defining connectivity operations available to the client;
- receiving a connectivity request from the client to establish a logically isolated network path to the second resource collection using a dedicated physical link established on behalf of the client to an endpoint router within the first geographical zone, wherein the connectivity request is formatted in accordance with the interface; and
- 10 performing a configuration operation to enable traffic to flow along the second logically isolated network path via the dedicated physical network link.

28. The method as recited in clause 27, further comprising:

- in response to the connectivity request, sending instructions to the client to transmit
- 15 connectivity metadata associated with the logically isolated network path to a destination address within the second geographical zone; and
- prior to performing the configuration operation, verify that the connectivity metadata was transmitted in accordance with the instructions.

29. The method as recited in clause 28, wherein the connectivity metadata is encoded in
- 20 accordance with an encryption algorithm set up to secure communication between the connectivity coordinator and a network device in the second geographical zone.

30. The method as recited in clause 27, further comprising:

- implementing a first pricing policy for the first geographical zone, and a second pricing
- 25 policy for the second geographical zone; and
- providing an indication of the second pricing policy to the client in accordance with the interface.

31. The method as recited in clause 30, wherein at least one pricing policy of the first and second
- 30 pricing policies comprises a pricing indicator based on at least one of: an amount of network traffic generated, a distance over which network traffic is transmitted, a use of a load balancing mechanism, or a use of a wide area network (WAN) scaling mechanism.

32. The method as recited in clause 27, further comprising:

validating, prior to performing the configuration operation, that the configuration operation is in compliance with one or more access policies associated with the second resource collection.

5 33. The method as recited in clause 27, further comprising:

providing an enumeration to the client of one or more resource collections to which logically isolated connections may be established over the dedicated physical link, wherein the enumeration is formatted in accordance with the interface.

10 34. The method as recited in clause 27, further comprising:

providing an indication to the client of a first performance service level for traffic within the first geographical zone, and a second performance service level for traffic between the first geographical zone and the second geographical zone, wherein the indication is formatted in accordance with the interface.

15

35. The method as recited in clause 27, wherein the configuration operation comprises a routing change at the endpoint router.

20 36. A non-transitory computer-accessible storage medium storing program instructions that when executed on one or more processors:

implement a programmatic interface defining connectivity operations available to a client of a provider network, wherein the provider network comprises a first geographical zone comprising a first resource collection allocated to the client, and a second geographical zone comprising a second resource collection allocated to the client;

25 receive a connectivity request from the client to establish a logically isolated network path to the second resource collection using a dedicated physical link established on behalf of the client to an endpoint router within the first geographical zone, wherein the connectivity request is formatted in accordance with the interface; and

perform a configuration operation to enable traffic to flow along the second logically
30 isolated network path via the dedicated physical network link.

37. The non-transitory computer-accessible storage medium as recited in clause 36, wherein the instructions when executed on the one or more processors:

35 in response to the connectivity request, send instructions to the client to transmit connectivity metadata associated with the logically isolated network path to a destination address

within the second geographical zone over a path that excludes the dedicated physical network link; and

prior to performing the configuration operation, verify that the connectivity metadata was transmitted in accordance with the instructions.

5

38. The non-transitory computer-accessible storage medium as recited in clause 37, wherein the connectivity metadata is encoded in accordance with an encryption mechanism.

39. The non-transitory computer-accessible storage medium as recited in clause 36, wherein the instructions when executed on the one or more processors:

10

implement a first pricing policy for the first geographical zone, and a second pricing policy for the second geographical zone; and

provide an indication of the second pricing policy to the client in accordance with the interface.

15

40. The non-transitory computer-accessible storage medium as recited in clause 36, wherein the instructions when executed on the one or more processors:

provide an enumeration to the client of one or more resource collections to which logically isolated connections may be established over the dedicated physical link, wherein the enumeration is formatted in accordance with the interface.

20

41. The non-transitory computer-accessible storage medium as recited in clause 36, wherein the instructions when executed on the one or more processors:

provide an indication to the client of a first performance service level for traffic within the first geographical zone, and a second performance service level for traffic between the first geographical zone and the second geographical zone, wherein the indication is formatted in accordance with the interface.

25

42. The non-transitory computer-accessible storage medium as recited in clause 36, wherein the instructions when executed on the one or more processors, wherein the configuration operation comprises a routing change at the endpoint router.

30

43. The non-transitory computer-accessible storage medium as recited in clause 36, wherein the interface comprises at least one of: an application programming interface, a command line interface, a graphical user interface, or a web interface.

35

44. A system, comprising:

a resource collection allocated to a first client of a provider network and linked to a first client network of the first client via a first dedicated physical link; and

5 a connectivity coordinator;

wherein the connectivity coordinator is operable to:

implement an interface defining connectivity operations available to a plurality of clients of the provider network, including the first client and a second client;

transmit a notification formatted in accordance with the interface to the second client
10 indicating that access to a service implemented by the first client at the resource collection is configurable via a path comprising a second dedicated physical link, wherein the second dedicated physical link connects a second client network of the second client to the provider network; and

in response to a subscription request from the second client formatted in accordance with
15 the interface, perform one or more configuration operations to enable a request for the service from the second client network to be routed to the resource collection using the second dedicated physical link.

45. The system as recited in clause 44, wherein the connectivity coordinator is further operable
20 to:

receive a service discovery request from the second client, formatted in accordance with the interface;

wherein the notification formatted in accordance with the interface is generated in response to the service discovery request.

25

46. The system as recited in clause 44, wherein the connectivity coordinator is further operable to:

receive a service advertisement request from the first client, formatted in accordance with the interface, indicating that the service is available for subscription;

30 wherein the notification formatted in accordance with the interface is generated after the service advertisement request is received.

47. The system as recited in clause 44, wherein the notification comprises a pricing indicator for the service.

35

48. The system as recited in clause 44, wherein the interface comprises at least one of: an application programming interface, a command line interface, a graphical user interface, or a web interface.

5 49. A method, comprising:

presenting a connectivity service to a plurality of clients of a provider network, including a first client and a second client, wherein said connectivity service includes a connectivity coordinator implementing a programmatic interface defining connectivity operations available to the plurality of clients;

10 transmitting, to the second client, a notification formatted in accordance with the interface, indicating that a service implemented by the first client at a first resource collection of the provider network is available for subscription; and,

in response to a subscription request from the second client formatted in accordance with the interface, performing one or more configuration operations to enable a request for the service
15 from a client network of the second client to be routed to the first resource collection using a dedicated physical link established between the client network and the provider network.

50. The method as recited in clause 49, wherein the notification comprises one or more pricing indicators for the service.

20

51. The method as recited in clause 50, wherein the one or more pricing indicators for the service include a first pricing indicator for a first time period, and a second pricing indicator for a second time period.

25 52. The method as recited in clause 50, wherein the one or more pricing indicators for the service include a subscription fee charged by the first client, and a network usage fee charged by an operator of the provider network.

53. The method as recited in clause 50, wherein a pricing indicator of the one or more pricing
30 indicators is based on at least one of: an amount of network traffic associated with the service, or a distance over which network traffic associated with the service is transmitted.

54. The method as recited in clause 49, further comprising:

validating, prior to transmitting the notification, that the notification is in compliance with
35 one or more access policies associated with the first resource collection; and

validating, prior to performing the one or more configuration operations, that the one or more configuration operations are in compliance with the one or more access policies.

55. The method as recited in clause 49, further comprising:

5 prior to performing the one or more configuration operations in response to the subscription request, verifying, using one or more communications with the first client formatted in accordance with the interface, that the subscription request is acceptable to the first client.

56. The method as recited in clause 49, further comprising:

10 receiving a service advertisement request from the first client, formatted in accordance with the interface, indicating that the service is available for subscription.

57. The method as recited in clause 49, further comprising:

15 receiving a subscription slot availability indicator from the first client, formatted in accordance with the interface, comprising an indication of number of subscriptions available to the service.

58. The method as recited in clause 49, further comprising:

20 receiving a service discovery request from the second client, formatted in accordance with the interface,
 wherein the transmitting the notification is responsive to the service discovery request.

59. The method as recited in clause 49, wherein the one or more configuration operations comprise establishing a logically isolated network path over the dedicated physical link.

25

60. The method as recited in clause 49, wherein the one or more configuration operations comprise a routing change at a router of the provider network.

61. A non-transitory computer-accessible storage medium storing program instructions that when executed on one or more processors:

30

 implement a programmatic interface defining connectivity operations available to a plurality of clients of a provider network, including a first client and a second client;

 transmit, to the second client, a notification formatted in accordance with the interface, indicating that a service implemented by the first client at a first resource collection of the provider network is available for subscription; and,

35

in response to a subscription request from the second client formatted in accordance with the interface, perform one or more configuration operations to enable a request for the service from a client network of the second client to be routed to the first resource collection using a dedicated physical link established between the client network and the provider network.

5

62. The non-transitory computer-accessible storage medium as recited in clause 61, wherein the notification comprises one or more pricing indicators for the service.

63. The non-transitory computer-accessible storage medium as recited in clause 61, wherein the instructions when executed on one or more processors:

10

prior to performing the one or more configuration operations in response to the subscription request, verify, using one or more communications formatted in accordance with the interface, that the subscription request is acceptable to the first client.

64. The non-transitory computer-accessible storage medium as recited in clause 61, wherein the instructions when executed on one or more processors:

15

receive a service advertisement request from the first client, formatted in accordance with the interface, indicating that the service is available for subscription.

65. The non-transitory computer-accessible storage medium as recited in clause 61, wherein the instructions when executed on one or more processors:

20

receive a subscription slot availability indicator from the first client, formatted in accordance with the interface, comprising an indication of number of subscriptions available to the service.

25

66. The non-transitory computer-accessible storage medium as recited in clause 61, wherein the instructions when executed on one or more processors:

receive a service discovery request from the second client, formatted in accordance with the interface,

wherein the notification is transmitted in response to the service discovery request.

30

67. The non-transitory computer-accessible storage medium as recited in clause 61, wherein the one or more configuration operations comprise establishing a logically isolated network path over the dedicated physical link.

35

68. The non-transitory computer-accessible storage medium as recited in clause 61, wherein the one or more configuration operations comprise a routing change at a router of the provider network.

5 Illustrative computer system

[0085] In at least some embodiments, a server that implements a portion or all of one or more of the technologies described herein, including the techniques to implement an interface that defines various connectivity services and operations, and to receive and respond to various types of connectivity requests via the interface, may include a general-purpose computer system that
10 includes or is configured to access one or more computer-accessible media, such as computer system 2000 illustrated in Figure 20. In the illustrated embodiment, computer system 2000 includes one or more processors 2010 coupled to a system memory 2020 via an input/output (I/O) interface 2030. Computer system 2000 further includes a network interface 2040 coupled to I/O interface 2030.

15 [0086] In various embodiments, computer system 2000 may be a uniprocessor system including one processor 2010, or a multiprocessor system including several processors 2010 (e.g., two, four, eight, or another suitable number). Processors 2010 may be any suitable processors capable of executing instructions. For example, in various embodiments, processors 2010 may be general-purpose or embedded processors implementing any of a variety of
20 instruction set architectures (ISAs), such as the x86, PowerPC, SPARC, or MIPS ISAs, or any other suitable ISA. In multiprocessor systems, each of processors 2010 may commonly, but not necessarily, implement the same ISA.

[0087] System memory 2020 may be configured to store instructions and data accessible by processor(s) 2010. In various embodiments, system memory 2020 may be implemented using
25 any suitable memory technology, such as static random access memory (SRAM), synchronous dynamic RAM (SDRAM), nonvolatile/Flash-type memory, or any other type of memory. In the illustrated embodiment, program instructions and data implementing one or more desired functions, such as those methods, techniques, and data described above, are shown stored within system memory 2020 as code 2025 and data 2026.

30 [0088] In one embodiment, I/O interface 2030 may be configured to coordinate I/O traffic between processor 2010, system memory 2020, and any peripheral devices in the device, including network interface 2040 or other peripheral interfaces. In some embodiments, I/O interface 2030 may perform any necessary protocol, timing or other data transformations to convert data signals from one component (e.g., system memory 2020) into a format suitable for
35 use by another component (e.g., processor 2010). In some embodiments, I/O interface 2030 may

include support for devices attached through various types of peripheral buses, such as a variant of the Peripheral Component Interconnect (PCI) bus standard or the Universal Serial Bus (USB) standard, for example. In some embodiments, the function of I/O interface 2030 may be split into two or more separate components, such as a north bridge and a south bridge, for example.

5 Also, in some embodiments some or all of the functionality of I/O interface 2030, such as an interface to system memory 2020, may be incorporated directly into processor 2010.

[0089] Network interface 2040 may be configured to allow data to be exchanged between computer system 2000 and other devices 2060 attached to a network or networks 2050, such as other computer systems or devices as illustrated in Figures 1 through 20, for example. In various
10 embodiments, network interface 2040 may support communication via any suitable wired or wireless general data networks, such as types of Ethernet network, for example. Additionally, network interface 2040 may support communication via telecommunications/telephony networks such as analog voice networks or digital fiber communications networks, via storage area networks such as Fibre Channel SANs, or via any other suitable type of network and/or protocol.

15 [0090] In some embodiments, system memory 2020 may be one embodiment of a computer-accessible medium configured to store program instructions and data as described above for Figures 1 through 19 for implementing embodiments of methods and apparatus for interfaces to manage direct network peerings. However, in other embodiments, program instructions and/or data may be received, sent or stored upon different types of computer-accessible media.
20 Generally speaking, a computer-accessible medium may include non-transitory storage media or memory media such as magnetic or optical media, e.g., disk or DVD/CD coupled to computer system 2000 via I/O interface 2030. A non-transitory computer-accessible storage medium may also include any volatile or non-volatile media such as RAM (e.g. SDRAM, DDR SDRAM, RDRAM, SRAM, etc.), ROM, etc, that may be included in some embodiments of computer
25 system 2000 as system memory 2020 or another type of memory. Further, a computer-accessible medium may include transmission media or signals such as electrical, electromagnetic, or digital signals, conveyed via a communication medium such as a network and/or a wireless link, such as may be implemented via network interface 2040.

30 Conclusion

[0091] Various embodiments may further include receiving, sending or storing instructions and/or data implemented in accordance with the foregoing description upon a computer-accessible medium. Generally speaking, a computer-accessible medium may include storage media or memory media such as magnetic or optical media, e.g., disk or DVD/CD-ROM, volatile
35 or non-volatile media such as RAM (e.g. SDRAM, DDR, RDRAM, SRAM, etc.), ROM, etc, as

well as transmission media or signals such as electrical, electromagnetic, or digital signals, conveyed via a communication medium such as network and/or a wireless link.

[0092] The various methods as illustrated in the Figures and described herein represent exemplary embodiments of methods. The methods may be implemented in software, hardware, 5 or a combination thereof. The order of method may be changed, and various elements may be added, reordered, combined, omitted, modified, etc.

[0093] Various modifications and changes may be made as would be obvious to a person skilled in the art having the benefit of this disclosure. It is intended to embrace all such modifications and changes and, accordingly, the above description to be regarded in an 10 illustrative rather than a restrictive sense.

WHAT IS CLAIMED IS:

1. A system, comprising:

a data center including a resource collection designated to respond to service requests received from a client;

a plurality of endpoint routers linked to the data center by one or more private network paths; and

a connectivity coordinator;

wherein the connectivity coordinator is operable to:

implement an interface defining connectivity operations available to the client, wherein the connectivity operations pertain to network connectivity enabling the client to access the resource collection in the data center, wherein the connectivity operations include at least one of an operation to set up a connection, an operation to query about a connection, an operation to modify a connection, or an operation to take down a connection;

receive a connectivity request from the client for dedicated connectivity to the resource collection, wherein the connectivity request is formatted in accordance with the interface;

in response to the connectivity request,

select a target endpoint router of the plurality of endpoint routers, wherein the target endpoint router is configurable to provide a route over a private network path of the one or more private network paths in accordance with the connectivity request to provide the client access to the resource collection in the data center;

generate a reply comprising configuration instructions for a physical network link to be established to the target endpoint router to provide at least a portion of the dedicated connectivity for the client to access the resource collection in the data center; and

transmit the reply to the client including the configuration instructions for the physical network link.

2. The system as recited in claim 1, wherein the interface comprises at least one of: an application programming interface (API), a graphical user interface (GUI), or a command-line interface.

3. The system as recited in claim 1, wherein the target endpoint router is housed within a facility requiring authorization for physical access, and wherein the reply includes an indication of an authorization of physical access to the target endpoint router at the facility.
4. The system as recited in claim 1, wherein the connectivity coordinator is further operable to:
 - receive an isolation request from the client to establish a logically isolated network path to the resource collection via the physical link; and
 - implement a network isolation mechanism to establish the logically isolated network path in accordance with the isolation request.
5. The system as recited in claim 4, wherein the network isolation mechanism comprises at least one of: a virtual local area network (VLAN) mechanism or a Multi-Protocol Label Switching (MPLS) technique.
6. The system as recited in claim 1, wherein the connectivity coordinator is further operable to:
 - receive, from the client, identification information of a network device to be used to transmit network traffic of the client to the resource collection; and
 - provide, to the client, one or more configuration instructions for the network device based on the identification information.
7. A method, comprising:
 - presenting a connectivity service to a client of a provider network, wherein said connectivity service includes a connectivity coordinator implementing a programmatic interface defining connectivity operations available to the client, wherein the connectivity operations pertain to network connectivity enabling the client to access a resource collection of the provider network, wherein the connectivity operations include at least one of an operation to set up a connection, an operation to query about a connection, an operation to modify a connection, or an operation to take down a connection;
 - receiving, at the connectivity coordinator, a connectivity request for dedicated connectivity to the resource collection of the provider network, wherein the connectivity request is formatted in accordance with the interface;

in response to the connectivity request,

selecting a target endpoint router of a plurality of endpoint routers of the provider network, wherein the target endpoint router is configurable to provide a route over a private network to the resource collection in accordance with the connectivity request to provide the client access to the resource collection;

generating a notification comprising configuration information for a physical network link to be established to the target endpoint router to provide at least a portion of the dedicated connectivity for the client to access the resource collection; and

transmitting the notification including the configuration information for the physical network link.

8. The method as recited in claim 7, further comprising:

receiving identification information of a network device to be used to transmit network traffic of the client to the resource collection; and

providing one or more configuration instructions for the network device based on the identification information.

9. The method as recited in claim 7, wherein the resource collection comprises a plurality of resources, further comprising:

receiving one or more selection criteria identifying a subset of the plurality of resources to which an isolated connection is to be provided over the physical network link; and

routing network traffic in accordance with the selection criteria.

10. The method as recited in claim 9, wherein a selection criteria of the one or more selection criteria comprises a virtual local area network (VLAN) tag.

11. The method as recited in claim 7, wherein the target endpoint router is housed within a facility requiring authorization for physical access, and wherein the notification includes an indication of an authorization of physical access to the target endpoint router at the facility.

12. The method as recited in claim 7, wherein transmitting the notification comprises sending the notification to an operator of a facility at which the target endpoint router is housed.

13. A non-transitory computer-accessible storage medium storing program instructions that when executed on one or more processors:

implement a programmatic interface defining connectivity operations available to a client of a provider network, wherein the connectivity operations pertain to network connectivity enabling the client to access a resource collection of the provider network, wherein the connectivity operations include at least one of an operation to set up a connection, an operation to query about a connection, an operation to modify a connection, or an operation to take down a connection;

receive a connectivity request from the client for dedicated connectivity to the resource collection of the provider network, wherein the connectivity request is formatted in accordance with the interface;

in response to the connectivity request,

generate a notification comprising configuration information for a physical network link to be established to a target endpoint router of the provider network to provide at least a portion of the dedicated connectivity for the client to access the resource collection, wherein the target endpoint router is configurable to provide a route over a private network path to the resource collection in accordance with the connectivity request to provide the client access to the resource collection; and

transmit the notification including the configuration information for the physical network link.

14. The non-transitory computer-accessible storage medium as recited in claim 13, wherein the program instructions when executed on the one or more processors:

after the physical network link has been established, transmit a confirmation message indicating that the dedicated connectivity has been provided.

15. The non-transitory computer-accessible storage medium as recited in claim 13, wherein the configuration information comprises at least one of: a physical port of the target endpoint router, a rack identifier, a cage identifier or a patch panel identifier.

16. The non-transitory computer-accessible storage medium as recited in claim 13, wherein the connectivity request comprises one or more of: a bandwidth requirement, an availability requirement, or a requirement for a plurality of physical paths to the resource collection.

17. The non-transitory computer-accessible storage medium as recited in claim 13, wherein the program instructions when executed on the one or more processors:

receive identification information of a network device to be used to transmit network traffic of the client to the resource collection; and

provide one or more configuration instructions for the network device based on the identification information.

18. The non-transitory computer-accessible storage medium as recited in claim 13, wherein the resource collection comprises a plurality of resources, wherein the program instructions when executed on the one or more processors:

receive one or more selection criteria identifying a subset of the plurality of resources to which an isolated connection is to be provided over the physical network link; and

generate routing information to route network traffic in accordance with the selection criteria.

19. The non-transitory computer-accessible storage medium as recited in claim 18, wherein a selection criteria of the one or more selection criteria comprises a virtual local area network (VLAN) tag.

20. The non-transitory computer-accessible storage medium as recited in claim 13, wherein the target endpoint router is housed within a facility requiring authorization for physical access, and wherein the notification includes an indication of an authorization of physical access to the target endpoint router at the facility.

21. The non-transitory computer-accessible storage medium as recited in claim 13, wherein the configuration information comprises an identification of a physical location of the target endpoint router.