INVENTIONSHARE®

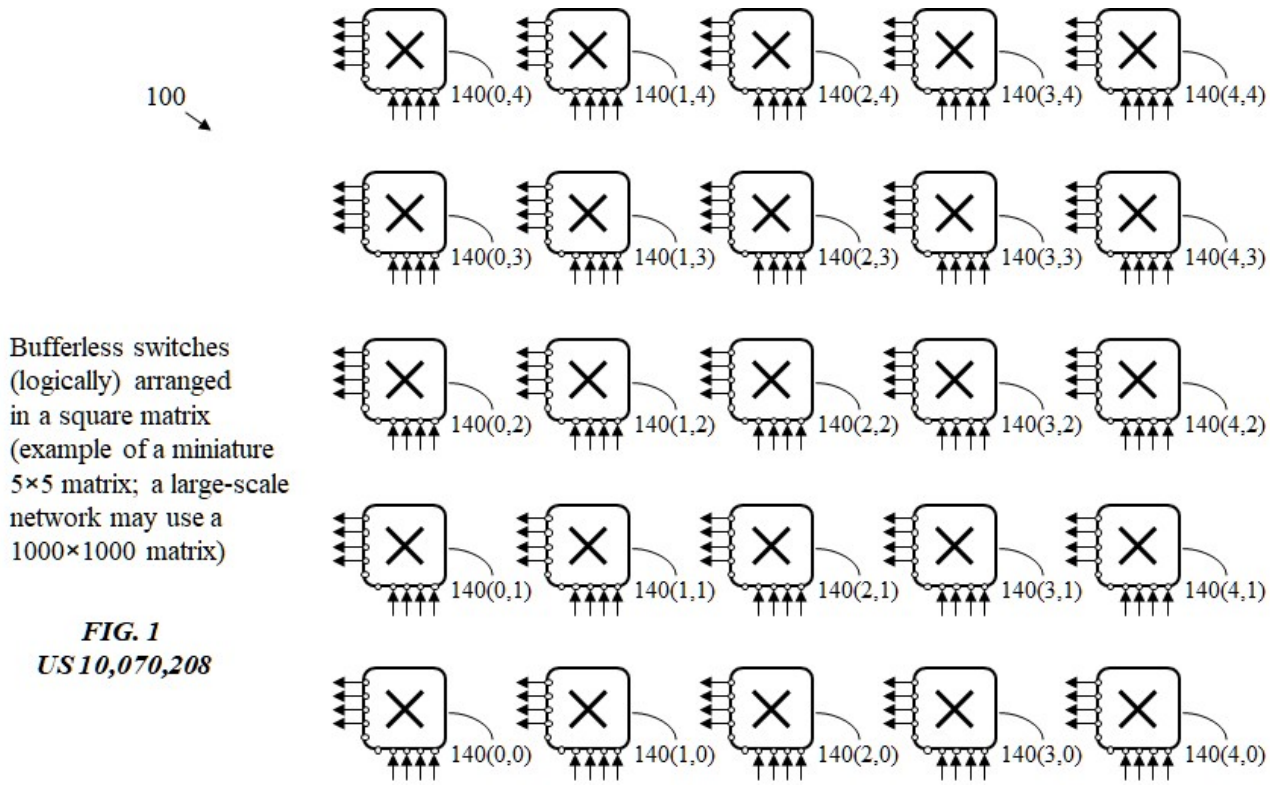
AcceleRoute™

A lightning-fast network architecture for Cloud Data Centers

Family I
Generic Scalable Network architecture based on Bufferless Core Switches.

The patents of Family I target a large-scale data center having a switching system that scales to a capacity of hundreds of petabits per second (Pb/s). Independent bufferless switches interconnect access nodes coupled to data sources and data sinks, where each access node has a path to each other access node traversing only one intermediate switch. The connectivity pattern of access nodes to switches may be devised in a variety of ways to naturally equalize data flow rates through the switches. Each access node has upstream channels to a selected set of the bufferless switches and downstream channels from another, or same, set of the bufferless switches. Optionally, the input sides of all of the bufferless switches may connect to mutually orthogonal sets of access nodes.

All control data and a portion of payload data traverse one intermediate switch. Payload data is buffered only at the edge; core buffering – a major source of routing problems and data-transfer latency – is eliminated. The network structure facilitates both distributed and centralized control.



Bufferless switches (logically) arranged in a square matrix (example of a miniature 5×5 matrix; a large-scale network may use a 1000×1000 matrix)

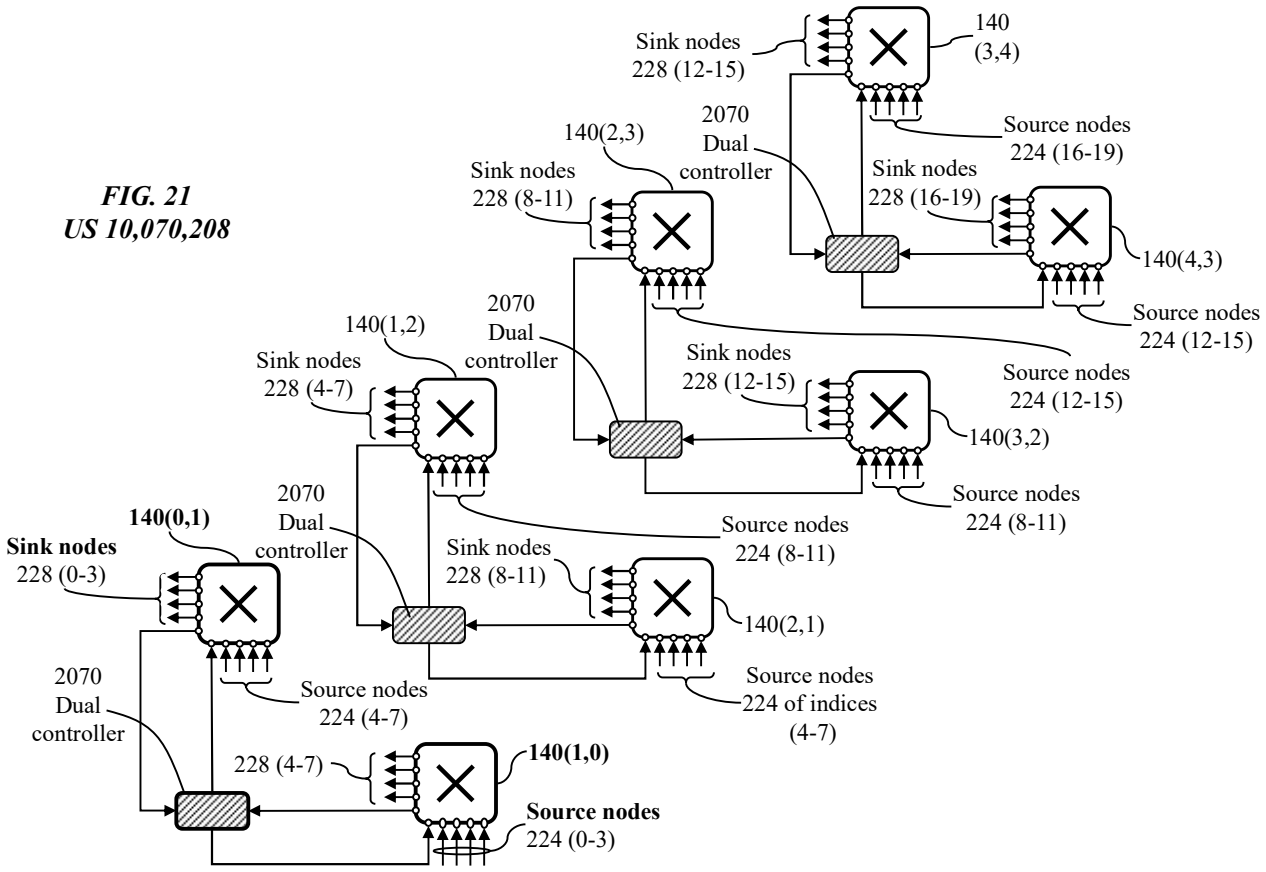
FIG. 1
US 10,070,208

Family I includes U.S. patents (7,760,716), (7,817,627), (8,422,508), (8,605,716), (9,148,370), (9,277,302), and (9,706,274), and (10,070,208).

US 10,070,208 discloses a configuration where the bufferless switches are arranged in a square matrix with integrated pairs of switches as illustrated in FIG. 1 and FIG. 21 of US *10,070,208.*

FIG. 21 of US 1010,208 illustrates diagonal switch pairs of the matrix of switches of FIG. 1. A switch 140(1,0) connects to source nodes 224 of indices {0, 1, 2, 3} and sink nodes 228 of indices {4, 5, 6, 7} while a complementary switch 140(0,1) connects to source nodes 224 of indices {4, 5, 6, 7} and sink nodes 228 of indices {0, 1, 2, 3}. An integrated pair of switches share a respective controller and, preferably, a respective switching mechanism.

FIG. 21
US 10,070,208



InventionShare®

InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

Family II

A Time Synchronized Global Network Using Spectral Routers

The patents of Family II target a data network of global coverage, scalable to an access capacity of hundreds of petabits per second with coverage of the order of 100,000 access nodes (edge nodes). The structure of the data network permits gradual growth from a very small fraction of the attainable capacity to full capacity and full coverage without disrupting the network during transition. The structural simplicity and ease of control facilitate partitioning the network among numerous constituent networks.

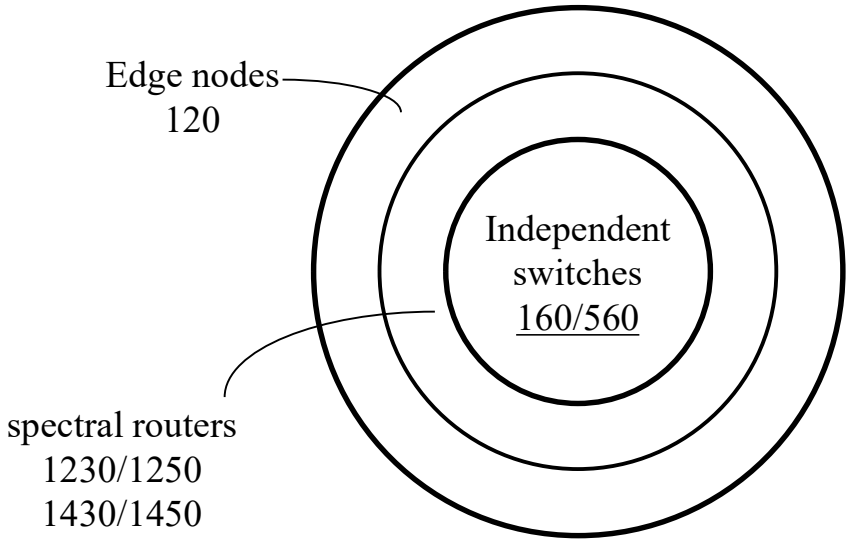
Electronic access nodes may connect to independent bufferless (preferably optical) switches through upstream and downstream spectral routers. Optionally, all sets of access nodes connecting to the upstream spectral routers may be selected to be mutually orthogonal. A significant proportion of traffic from any access node to any other access node of a global network would be routed through a single intermediate stage. A bufferless switch is preferably a fast optical switch to avoid the complexity, cost, and power consumption associated with optical-to-electrical and electrical-to-optical conversions. With the current state of the art, electronic switches can be realized to have a transient time of the order of a few nanoseconds. Current optical switches of large dimensions have a latency of several microseconds which precludes their use in a data network. However, a low-latency optical switch of a relatively small dimension (64 ports, for example) is realizable and may be used to interconnect access nodes.

A network of a capacity exceeding 1000 Pb/s is realizable. As in Family I, all control data and a portion of payload data undergo a single core switching process. Payload data is buffered only at the edge; core buffering – a major cause of routing problems – is eliminated. The network structure facilitates both distributed and centralized control. Optionally, the bufferless switches may be organized into two planes; a first plane comprising independent fast optical switches, and a second plane comprising bufferless electronic switches.

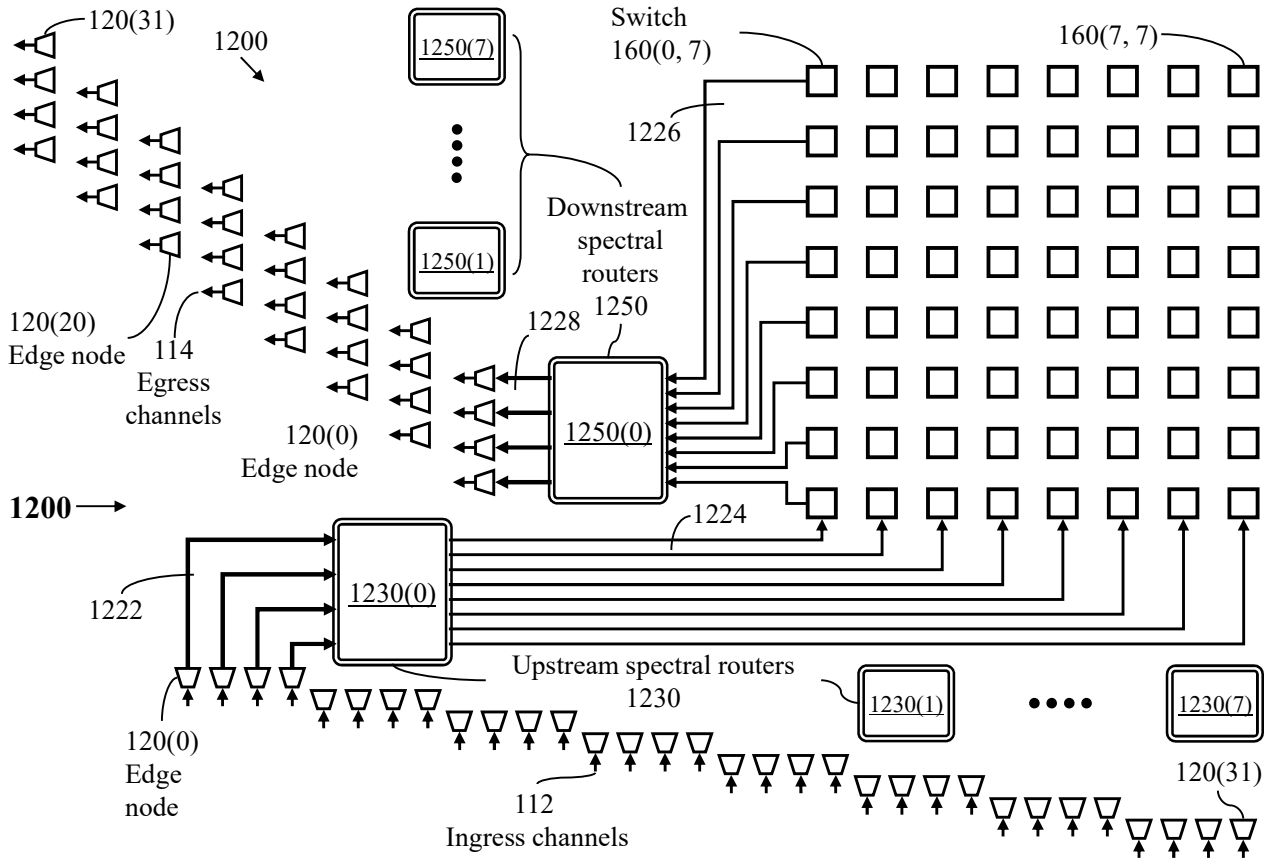
Family II includes U.S. patents (8,295,698) and (9,596,524).

FIG. 12 and FIG. 27 of US9,596,524 (reproduced below) illustrate a set of access nodes 120 interconnected through a matrix of independent (i.e., not directly interconnected) bufferless (preferably optical) switches 160. Passive upstream spectral routers 1230 (only one is illustrated)

connect subsets of the set of access nodes to respective rows of switches of the matrix of switches. Passive downstream spectral routers 1250 (only one is illustrated) connect columns of switches of the matrix of switches to respective subsets of the set of access nodes.



2700
FIG. 27 of
US 9596524



InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

FIG. 12 of US 9596524

Family III

High-capacity network of multiple switch planes

The patents of Family III target a large-scale data center or a network of continental coverage (4000 access nodes or so) providing a switching system that scales to a capacity of tens of petabits per second (Pb/s). Multiple switch planes, each switch plane having meshed bufferless, preferably fast optical, switches interconnect source nodes to sink nodes in a pattern which minimizes processing effort and requisite switching hardware. The components of the switching system may be localized in a data-center campus or spread across a continent. The connectivity pattern of source nodes and sink nodes to the switch planes is devised so that each access-node pair (a source-node and a sink-node) connects only once to a common switch.

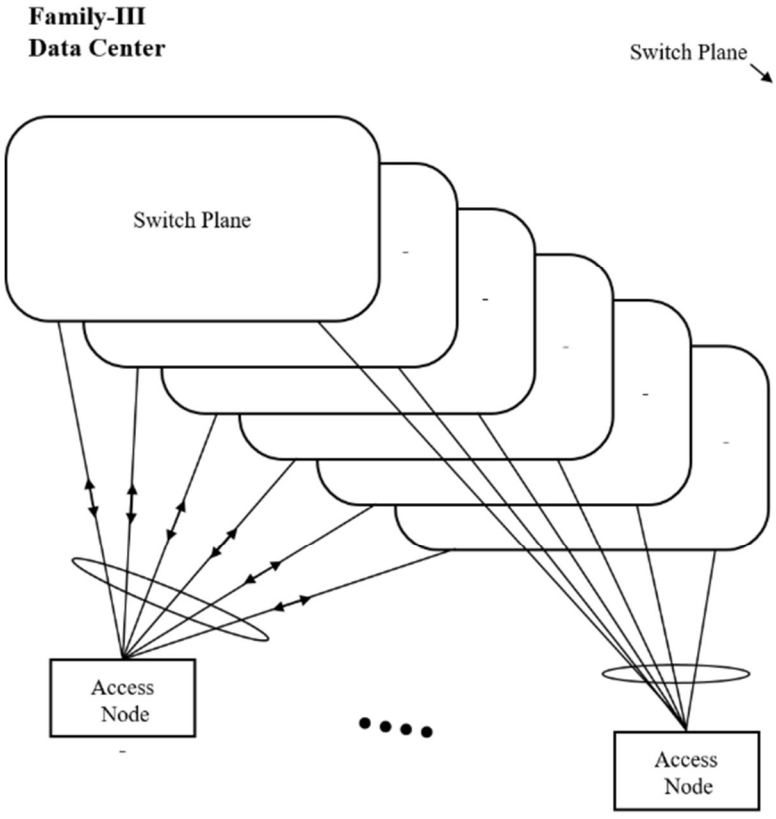


FIG. 1 of US 9,565,487

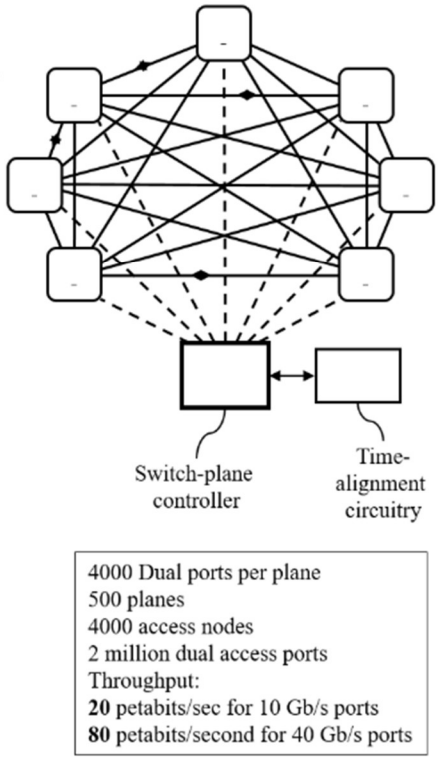


FIG. 59 of US 9,565,487

InventionShare®

InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

Family III includes U.S. patents (8,050,257), (8,774,200), and (9,565,487). FIG. 1 of U.S. 9,565,487 (reproduced above) is an overview of the structure of the data center. FIG. 59 (reproduced above) illustrates a switch plane having switch units, interconnected in a full mesh structure, and a switch-plane controller coupled to each switch unit of the switch plane

Upstream wavelength channels from each source node connect to different switch planes in a manner which ensures orthogonal connectivity of the source nodes to the bufferless switches where a set of source nodes connecting to any switch and another set of source nodes connecting to any other switch of any switch plane have at most one source node in common. This property results in creating a path from any source node to any sink node traversing a respective switch.

FIG. 32 of US 9,565,487 (reproduced below) illustrates a network comprising centralized switch planes with symmetrical connectivity in the upstream and downstream directions.

A data center or a continental-coverage network of a capacity of 80 Pb/s is realizable. All control data and a portion of payload data are transferred through a single switch plane. Payload data is buffered only at the edge; core buffering - a major source of routing problems - is eliminated. The network structure is SDN-friendly and facilitates both distributed and centralized control.

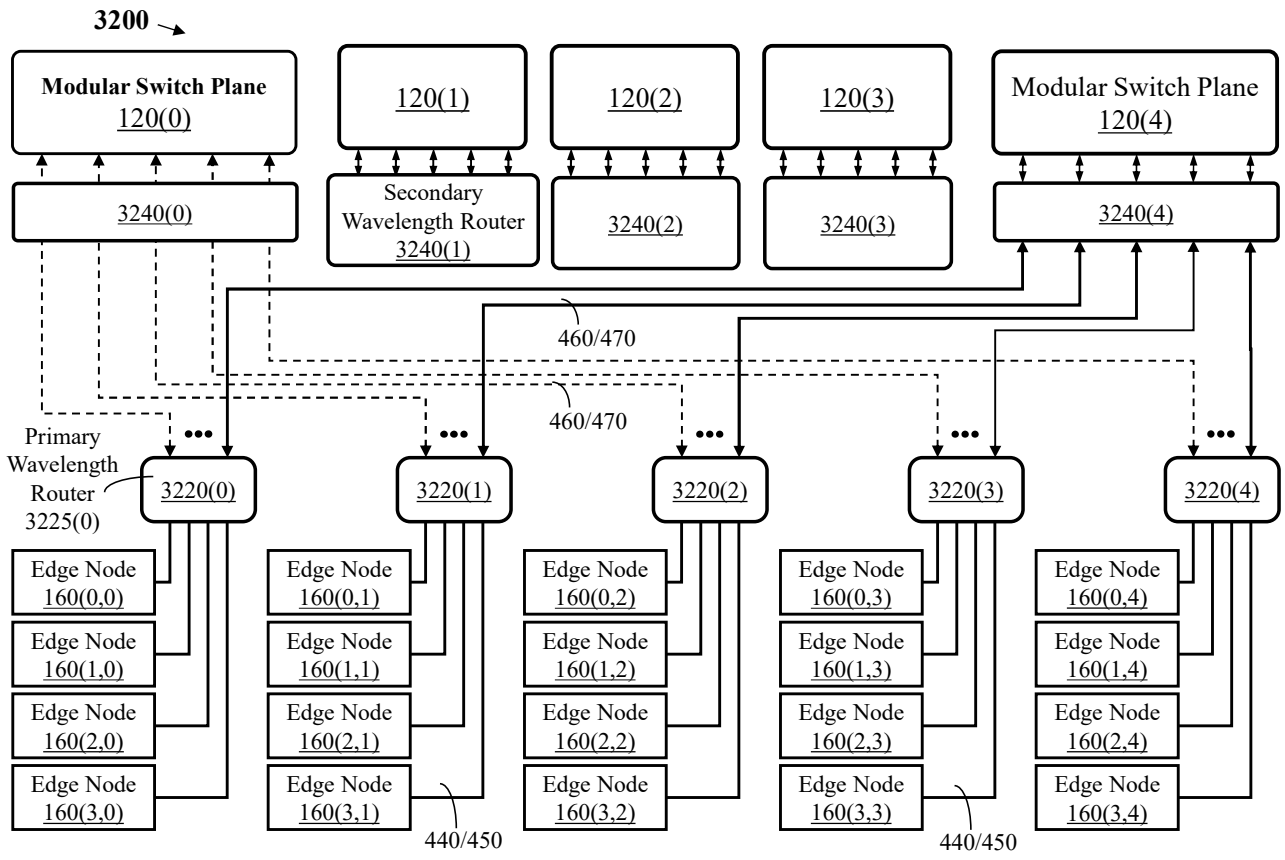


FIG. 32 of US 9,565,487

Family IV

Multiple petabit-per-second switching system based on orthogonal formations of sets of access nodes

The patents of Family IV target a switching system which may serve as a large-scale data center or a geographically distributed network of global coverage. Access nodes interfacing with external network entities (data sources/sinks) are interconnected through independent switches. The access nodes are arranged into a number of formations; each formation encompassing all of the access nodes. Within each formation, the access nodes are divided into a number of disjoint sets of nodes. The access nodes of each set are interconnected through a respective switch and are selected so that each set of access nodes of any formation is orthogonal to each set of access nodes of each other formation. With such a structure, each access node has a set of routes to each other access node, each route of which traversing at most two switches and an intermediate node.

INVENTIONSHARE®

InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

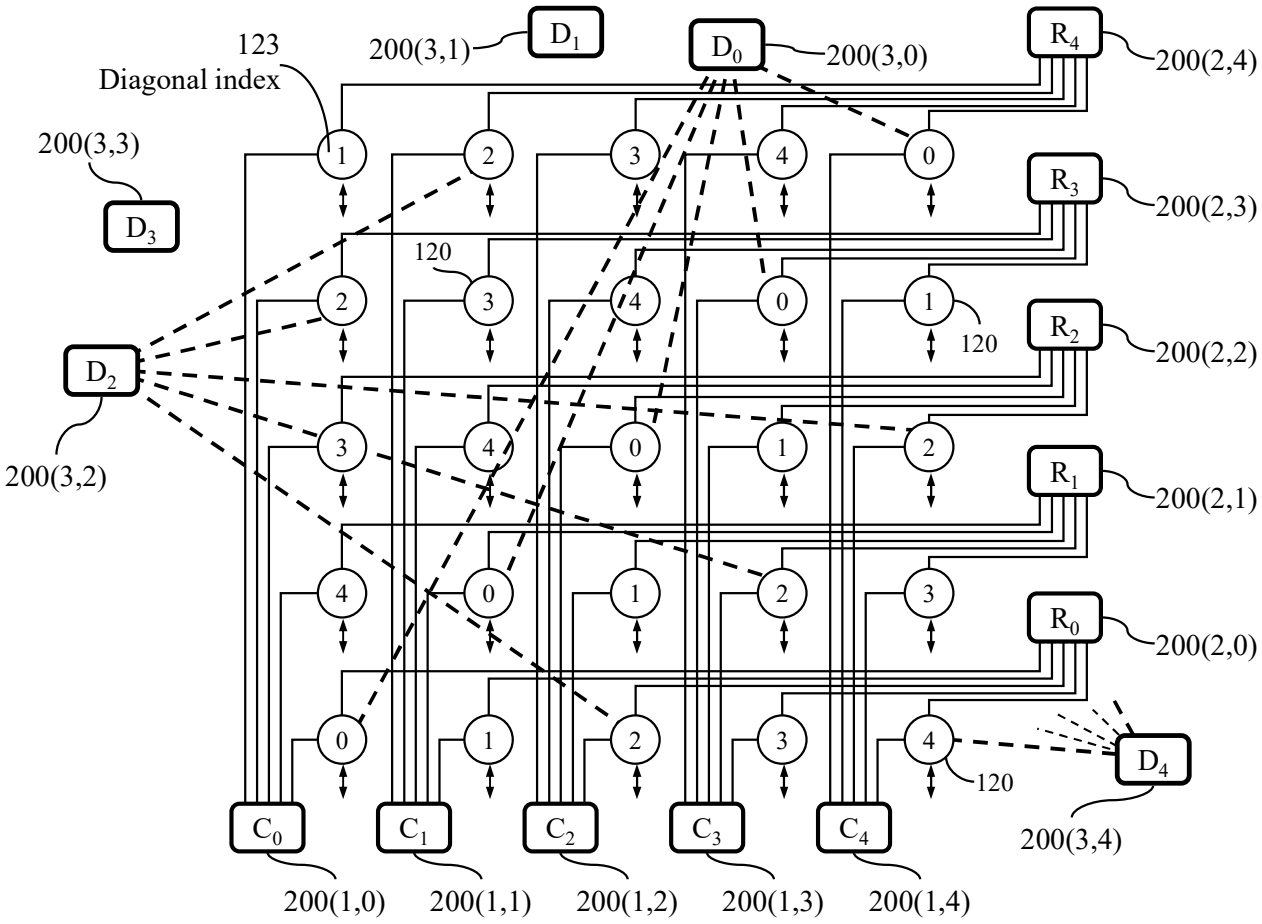


FIG. 15 of US 9,860,132

InventionShare®

InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

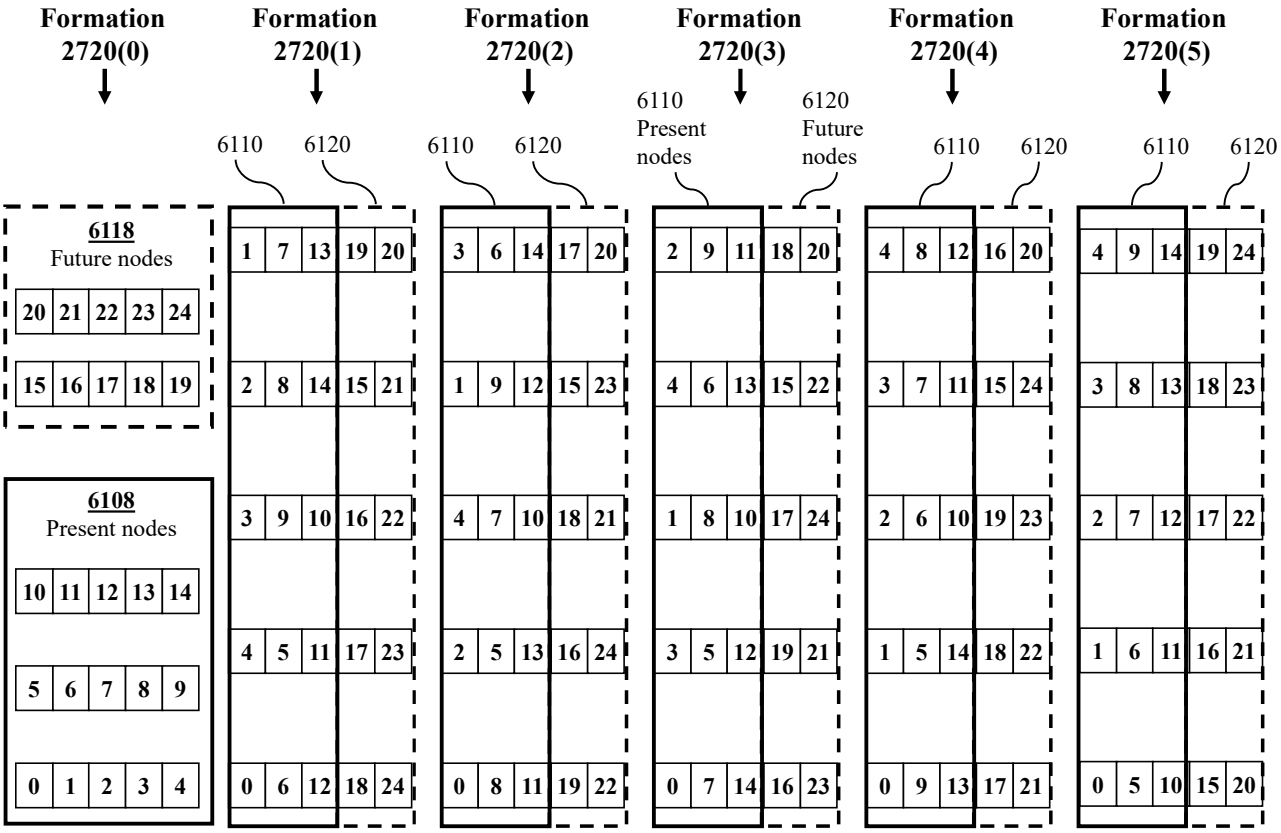


FIG. 61 of US 9,860,132

Family IV includes U.S. patents (9,054,979) and (9,860,132). FIG. 15 of US 9,860,132 (reproduced above) illustrates a configuration of only three formations. With a sufficient number of formations, each access node may have a path to any other access node traversing only one switch. A wide-coverage network of a capacity exceeding 380 Pb/s is realizable.

FIG. 61 of US 9,860.132 (reproduced above) illustrates exemplary formations of present and anticipated future nodes. The first formation 2720(0) includes both present nodes 6108 and anticipated future nodes.6118. Each of formations 2720(γ), $0<\gamma<6$, includes a formation 6110 of present nodes and a formation 6120 of the anticipated future nodes.

All control data and a portion of payload data are communicated through paths each traversing one intermediate switch. Payload data is buffered only at the edge; core buffering – a major source of routing problems – is eliminated. The switching-system structure facilitates both distributed and centralized control.

INVENTIONSHARE®

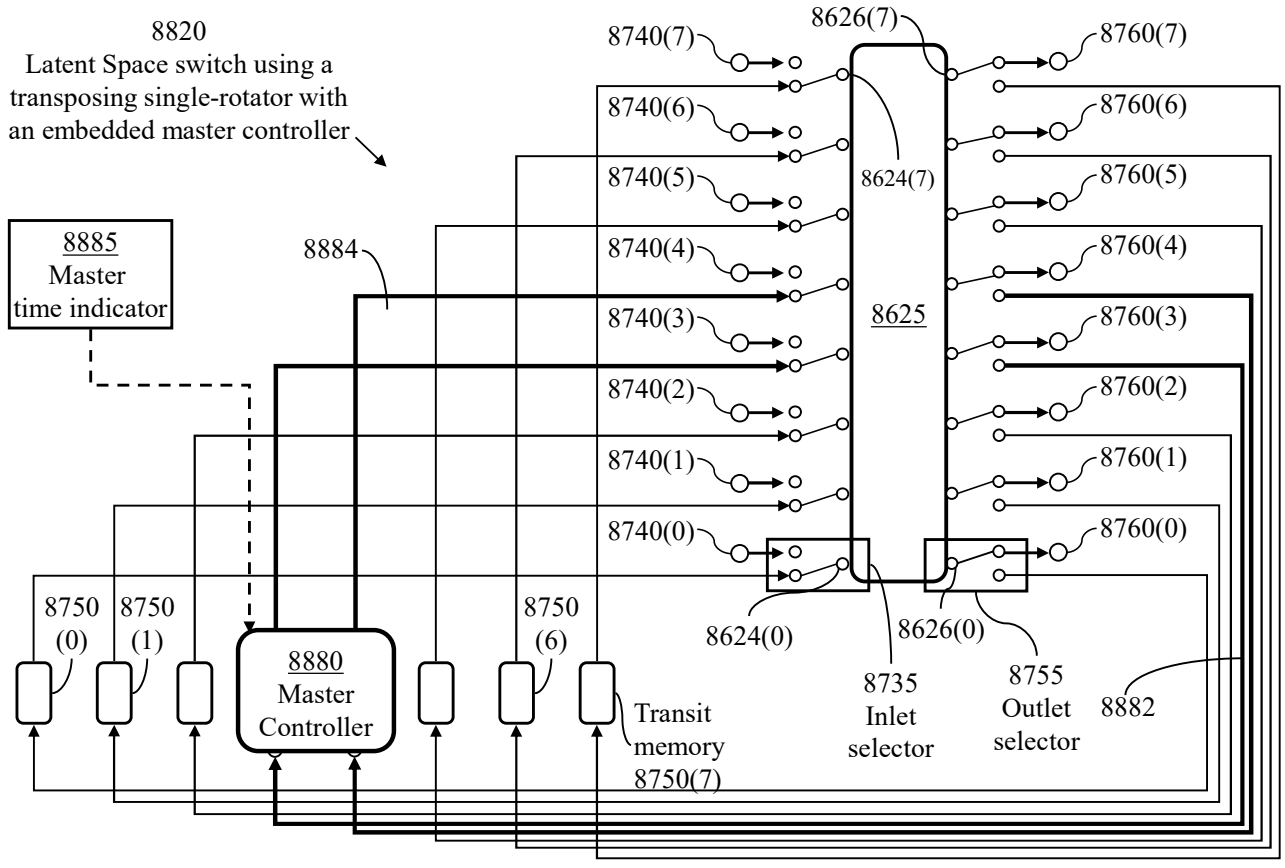
InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

Family V

Very-high-throughput latent space switches each using a single rotator

The disclosed configurations of latent space switches enable the simplest possible realization of a scalable data switch having a number of access ports that may vary, for example, between 32 dual ports and 16384 dual ports. Family V includes U.S. patents (8,204,050), (8,971,340), (9,154,255), and (9.252, 909).

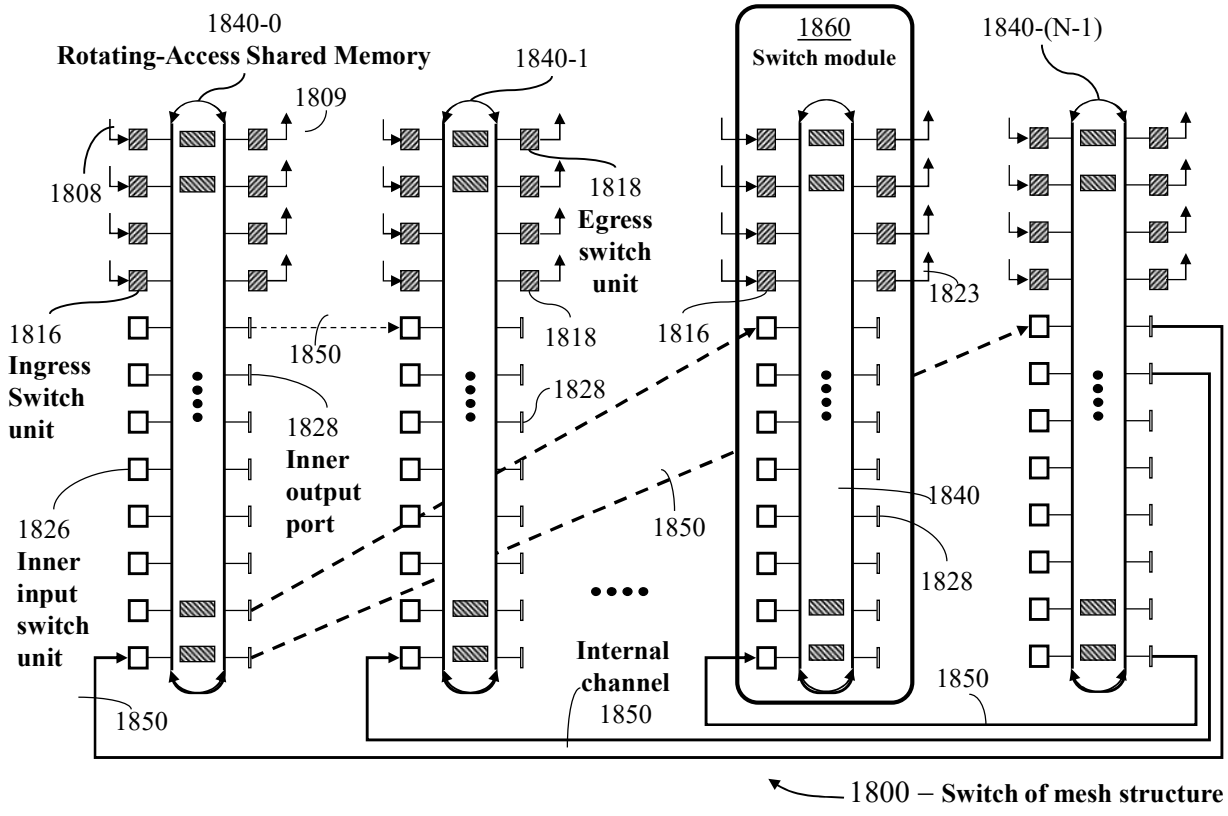
A latent space switch has a number, N, exceeding 2, of memory devices, N ingress ports, N egress ports, and a rotator. A rotator may be a conventional rotator or a “transposing rotator” defined in US 8,971,340. The drawing below illustrates the configuration using a transposing rotator. A transposing rotator having N inlets and N outlets is configured to cyclically connect each inlet to each outlet, during a predefined time frame, starting, for each inlet, with a respective transposed outlet. The time frame is organized into N time slots. During each time slot, each inlet alternately connects to a respective ingress port and a respective memory device while a peer outlet of the inlet alternately connects to the respective memory device and a respective egress port.



InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

Family VI
Packet switches based on contention-free switch modules

The patents of Family VI disclose various packet switching systems based on interconnecting contention-free switch modules. Family VI includes U.S. patents (7,693,142), (8,223,759), (8,576,839), (8,774,176), and (9,634,960). In one implementation (FIG. 18, US 9,634,960, reproduced below), each switch module employs a rotating-access shared memory and the contention-free switch modules are interconnected in a full mesh structure. The use of contention-free switch modules simplifies data routing across the switch but limits the overall capacity of the switching system. Thus, the switching systems of Family VI target packet switches of moderate capacities. For example, using switch modules of 64 dual ports each, of which 24 dual ports are access ports connecting to external data sources and sinks and the remaining 40 dual ports are inner ports linked to neighboring switch modules, the resulting packet switch would have 984 access ports. With each dual port (access port or inner port) operating at 40 Gb/s, the capacity of the packet switch would be approximately 39 Tb/s.



InventionShare®

InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

Family VII
A contiguous switching system for cloud communications or a Large-Scale Data Center

The patents of Family VII disclose a contiguous switching system for a large-scale cloud communications network or a localized large-scale data center. The switching system is configured to enable routing a significant proportion of incoming data through a distributed single-stage connector. The disclosed switching system may serve as a shared cloud network supporting numerous constituent networks of global coverage or a campus large-scale data center.

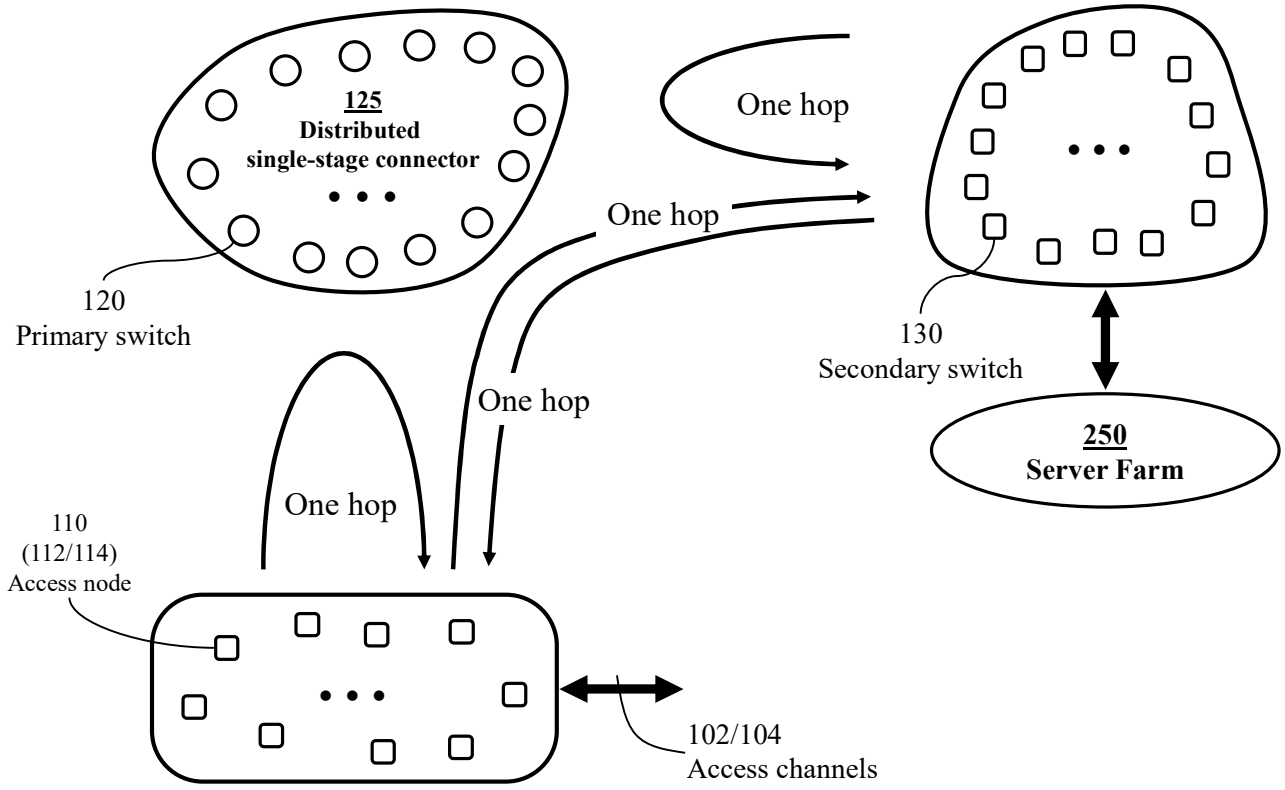
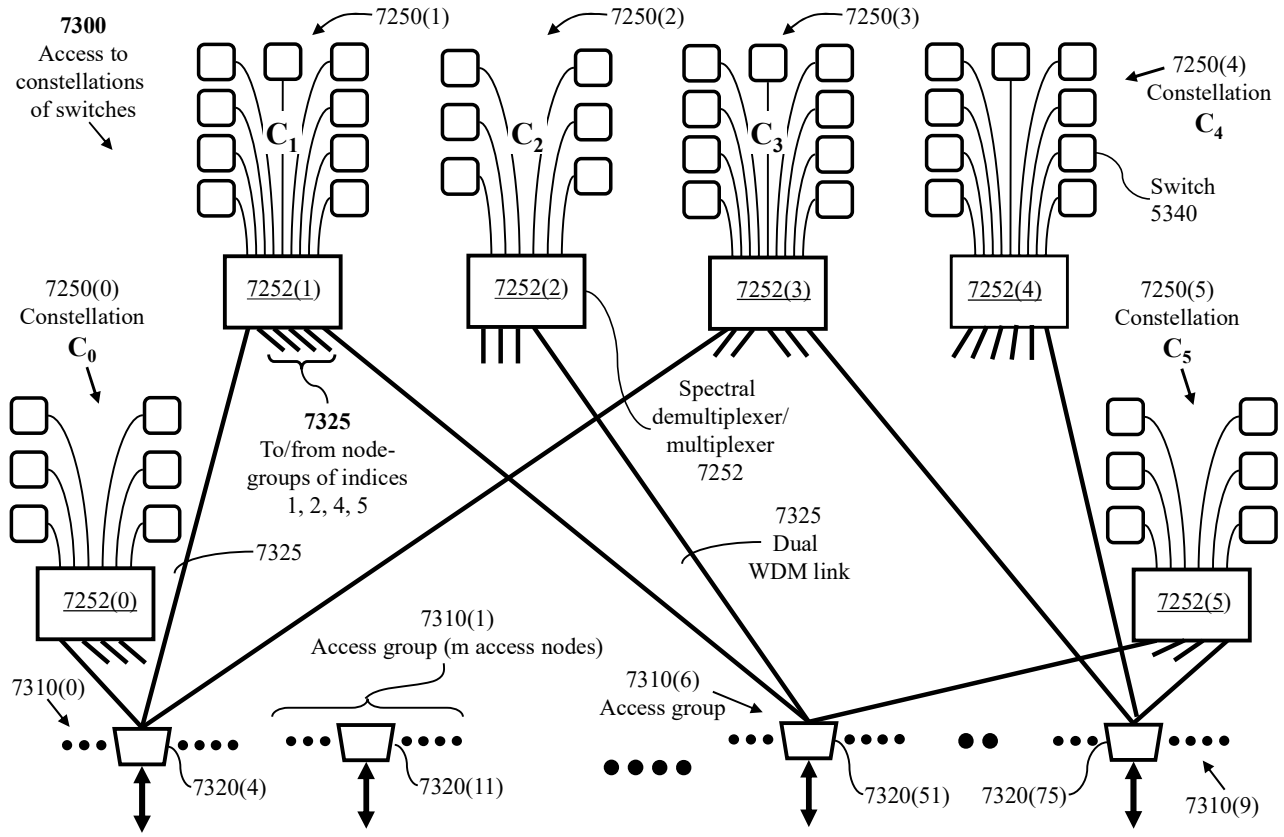


FIG. 3 of US 10,027,510

The switching system is based on orthogonal connectivity of a distributed single-stage connector, having a large number of disjoint primary switches, to a large number of access nodes and a large number of secondary switches. The secondary switches may be coupled to a large number of servers forming a server farm. The orthogonal-connectivity scheme yields a contiguous network providing single-hop paths from each access node to each other access node, from each secondary switch to each other secondary switch, from each access node to each secondary switch, and from

each secondary switch to each access node. The connectivity pattern enables network scalability to accommodate hundreds of thousands of access nodes while permitting a significant proportion of traffic to be routed through single-hop paths.

Family VII includes U.S. patents (10,027,510), (10,129,050), and (10, 412, 472). FIG. **3** of US 10,129,50 (reproduced above) illustrates a contiguous network, serving external traffic and providing access to a server farm. In a *contiguous* network, having a set of access nodes and a set of switches, each access node connects to each other access node through a respective single switch. The contiguous network of the present invention accommodates a large number of access nodes well beyond what can be realized with the conventional single Clos Star Network. FIG. 73 of US 10,412,472 (reproduced below) illustrates a large-scale contiguous network having access nodes arranged into access groups and distributors arranged into constellations of collocated distributors. The distributors may comprise switches, rotators, or a mixture of switches and rotators. Each access group connects to each distributor of a respective set of distributors selected so that each pair of access groups connects once to a respective distributor. At least one access group comprises a global controller.



InventionShare®

InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

FIG. 73 of US 10,412,472

InventionShare®

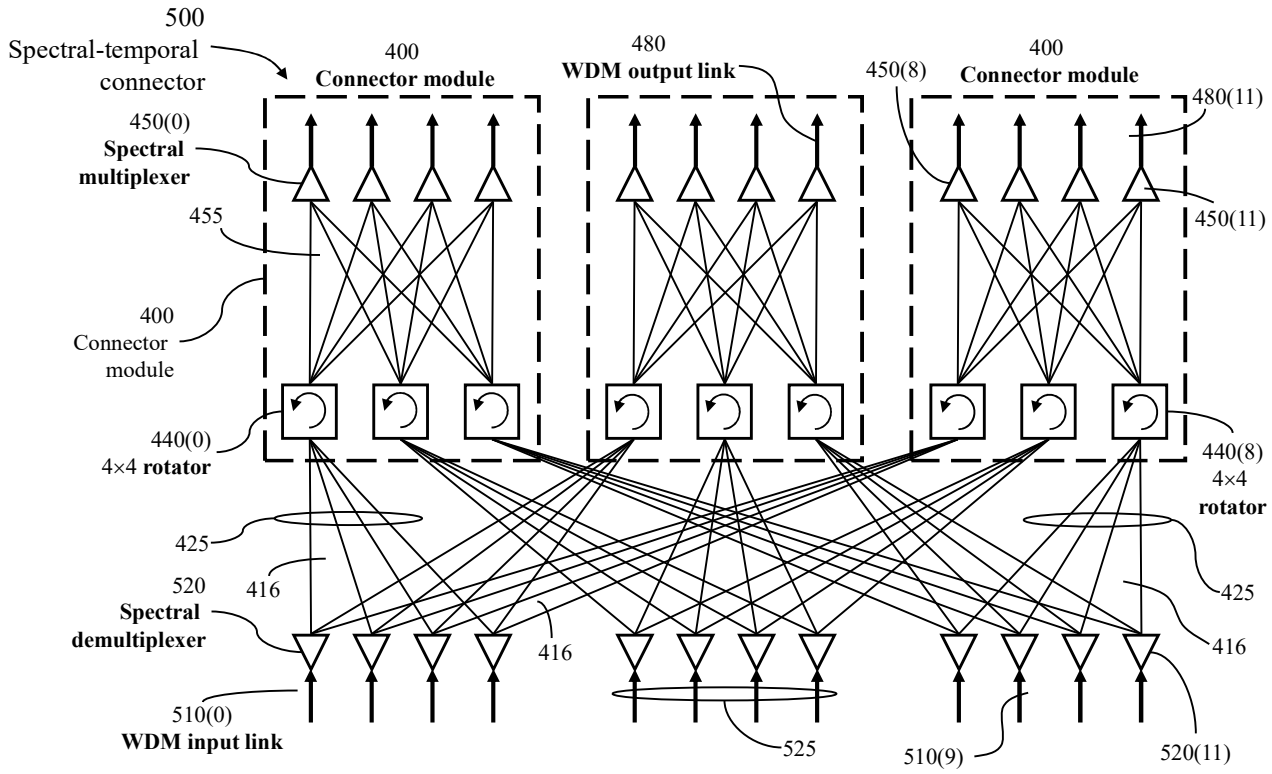
InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

Family VIII

Optical spectral-temporal connector interconnecting a large number of nodes in a full-mesh structure

The patents of Family VIII include U.S. Patents (9,509,432), (9,647,792), and (10,003,865) which disclose an optical spectral-temporal connector interconnecting a large number of access nodes in a full-mesh structure. Optical signals occupying multiple spectral bands carried by a link from an access node are de-multiplexed into separate spectral bands individually directed to different connector modules. Each connector module has a set of temporal rotators and a set of spectral multiplexers. A temporal rotator cyclically distributes segments of each signal at each inlet of the rotator to outlets of the rotator. Several arrangements for time-aligning all the access nodes to the connector modules are disclosed.

FIG. 5 of US 9,509,432 (reproduced below) illustrates a spectral-temporal connector employing an array of temporal rotators for transferring signals from each wavelength-division-multiplexed (WDM) input link of a plurality of WDM input links to each WDM output link of a plurality of WDM output links.



InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

FIG. 5 of US 9,509,432

FIG. 6 of US 9,509,432 (reproduced below) illustrates a set of timing circuits individually coupled to respective temporal rotators of the spectral-temporal connector of FIG. **5**. All timing circuits obey a single master time indicator.

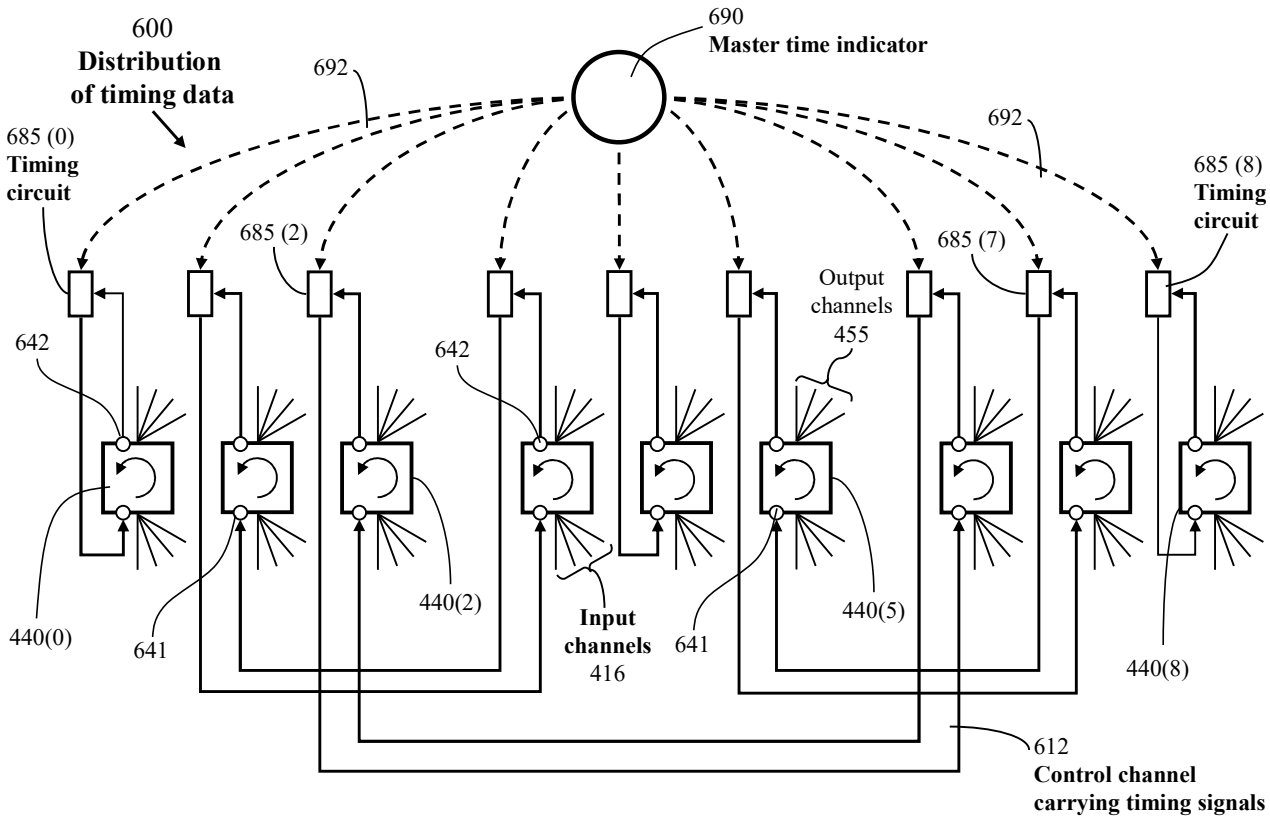


FIG. 8 of US 9,509,432

FIG. 44 of US 9,509,432 (reproduced below) illustrates a spectral-temporal connector interconnecting switching nodes of different dimensions and an optional central controller.

Each access node has a direct permanent path to each other access node of a capacity of the order of 1 Gb/s. Data may be transferred from an access node to another through a respective direct path and numerous other paths each traversing one intermediate access node. A configuration supporting 5000 access nodes and a throughput of 16 Pb/s would be suitable for a large-scale high-performance data center or a geographically distributed network. The extreme simplicity of this architecture greatly facilitates both distributed control and SDN-style central control.

INVENTIONSHARE®

InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

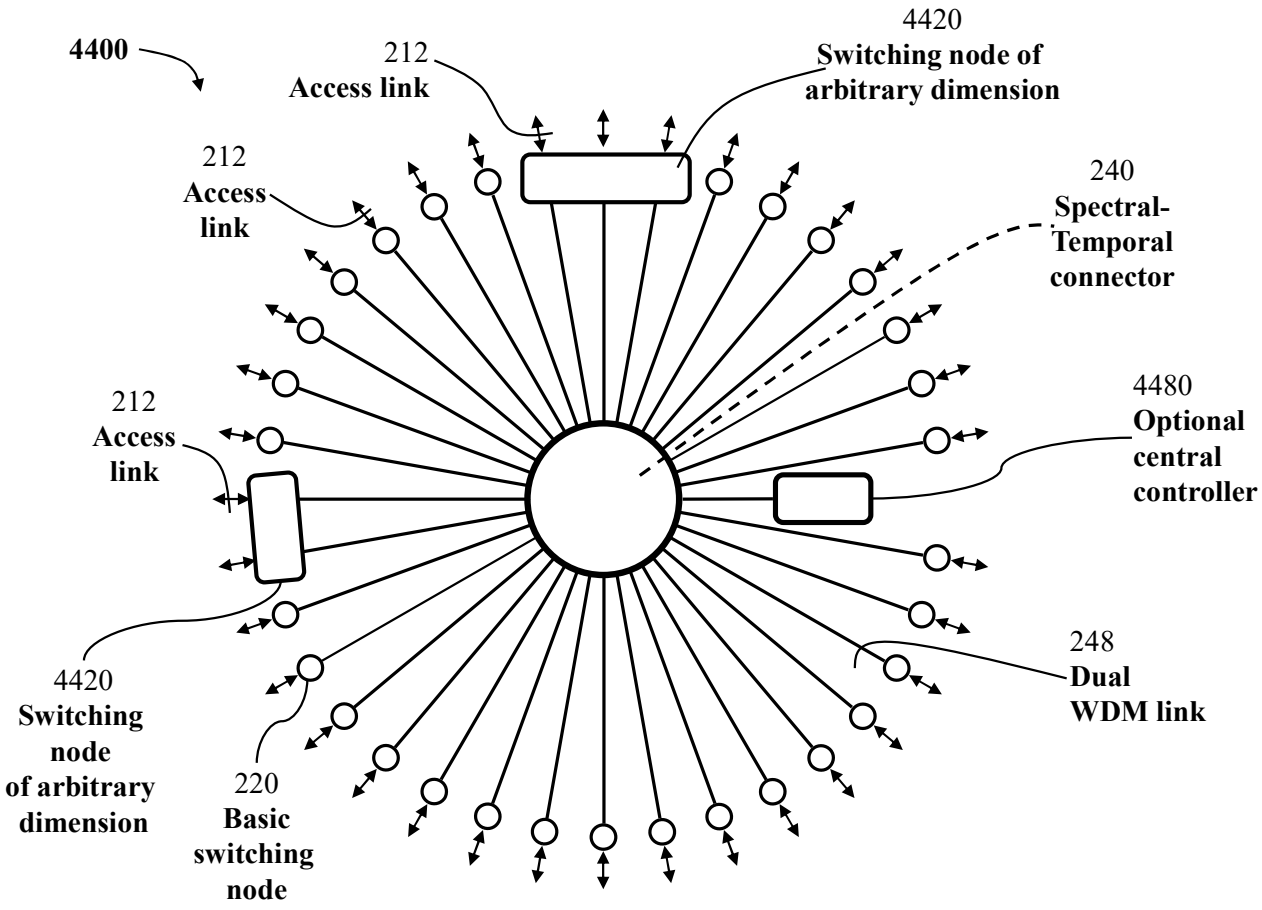


FIG. 44 of US 9,509,432

The spectral-temporal connector of family VIII is used in family IX and family XII

InventionShare®

InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

Family IX

Distributed Control of a Vast Communication Network

The patents of Family IX disclose a method and a system for distributed computation of a routing table for a vast **legacy** communication network. The processing effort of route generation and tracking network-elements' states is distributed among multiple coordinated network controllers. The network controllers may be interconnected in a full mesh structure, preferably through a spectral-temporal connector, disclosed in Family VIII, to form a distributed control system for a network of a large number of nodes.

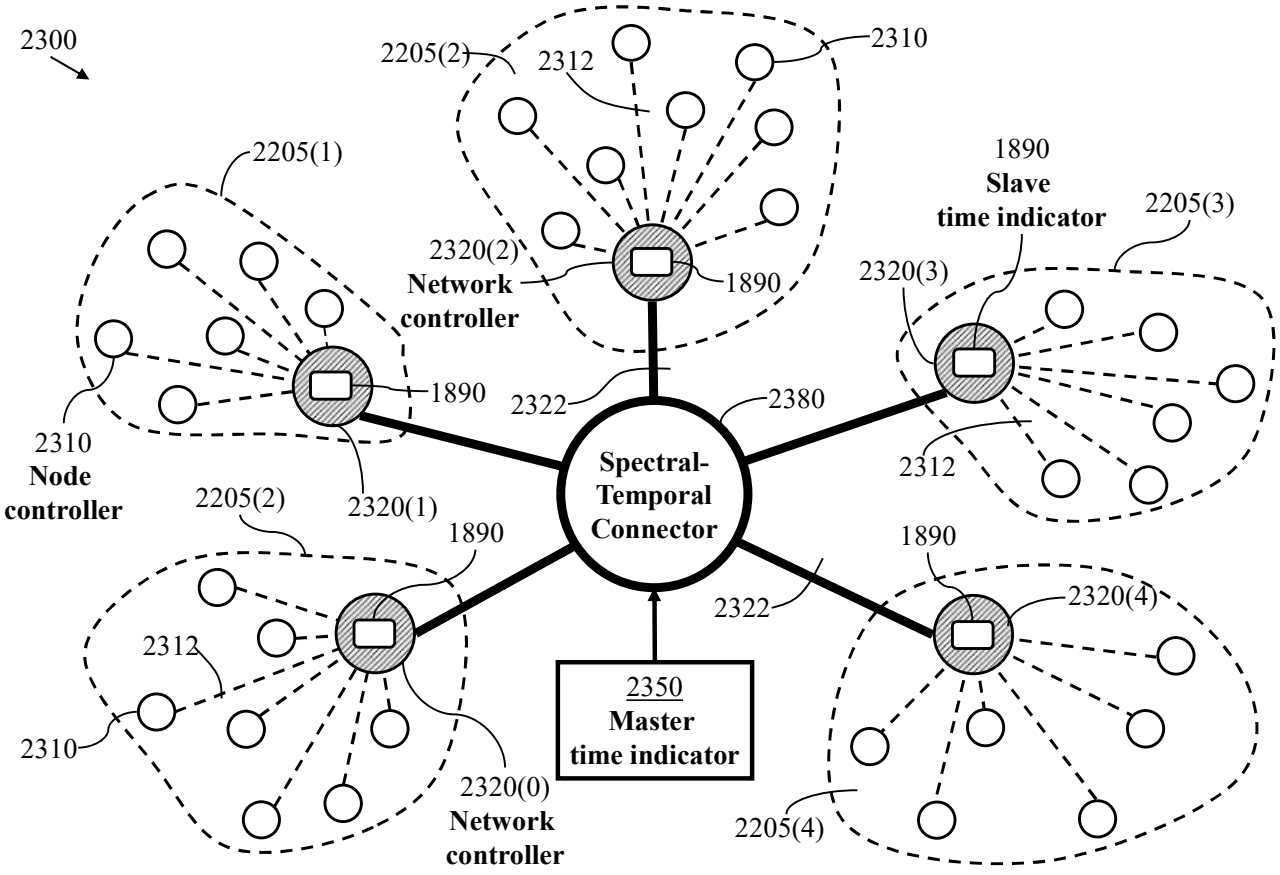


FIG. 23 of US 10,021,025

A network controller acquires characterizing information of links emanating from a respective set of nodes, communicates the information to each other network controller, and determines a route set from each node of the respective set of nodes to each other node of the network. The network controller may determine, for each link included in the route set, identifiers of specific route sets which traverse the link. Accordingly, a state-change of any link in the network

can be **expeditiously communicated to network controllers** to take corrective actions where necessary. A network controller may rank routes of a route set according to some criterion to facilitate selection of a favourable available route for a connection.

The disclosed control system remedies major routing problems of legacy networks.

Family IX includes U.S. patents (9,762,479) and (10,021,025). FIG. 23 of US 10,021,025 (reproduced above) illustrates connection of node controllers to a respective network controller, and interconnection of network controllers of the network of FIG. 22 through a spectral-temporal connector. FIG. 28 of US 10,021,025 (reproduced below) illustrates selective topology-data dissemination from a network controller to other network controllers according to inter-domain affinity.

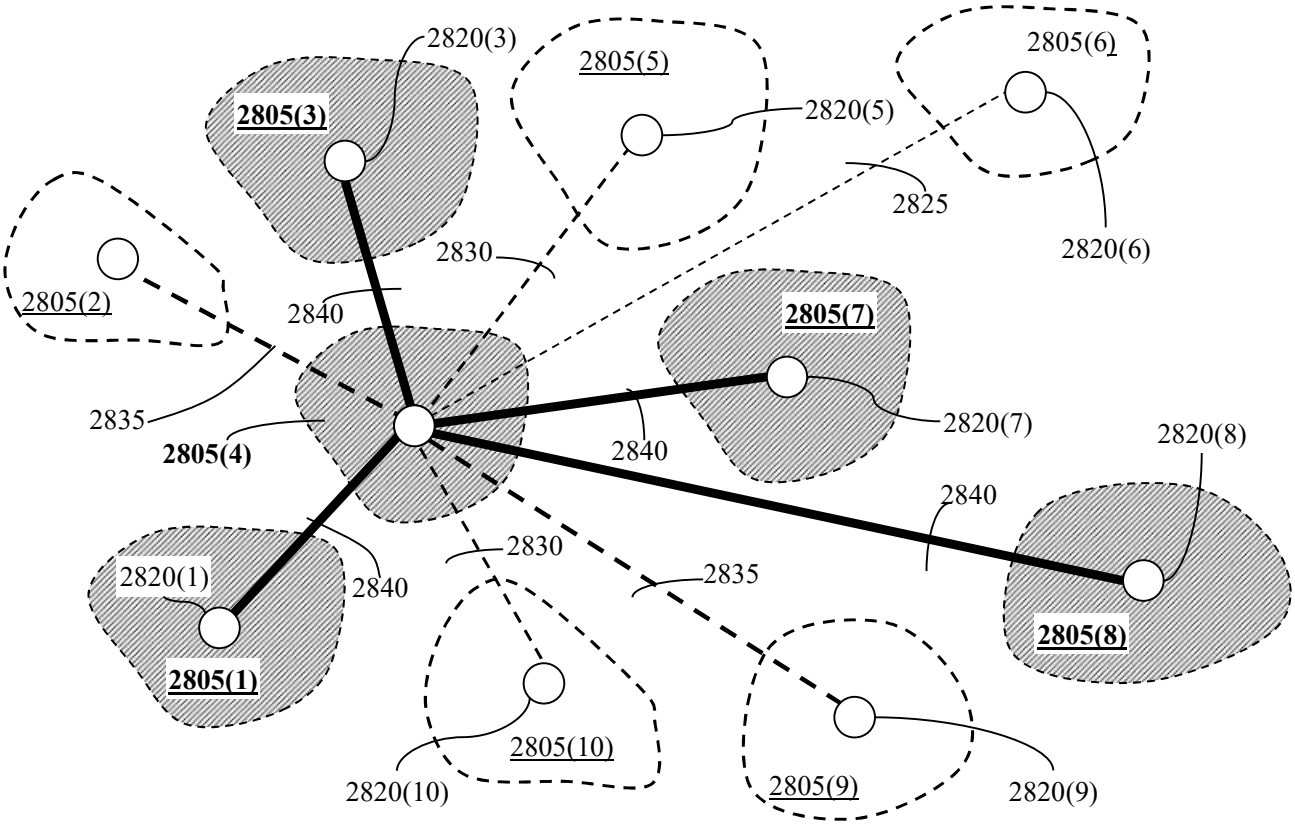


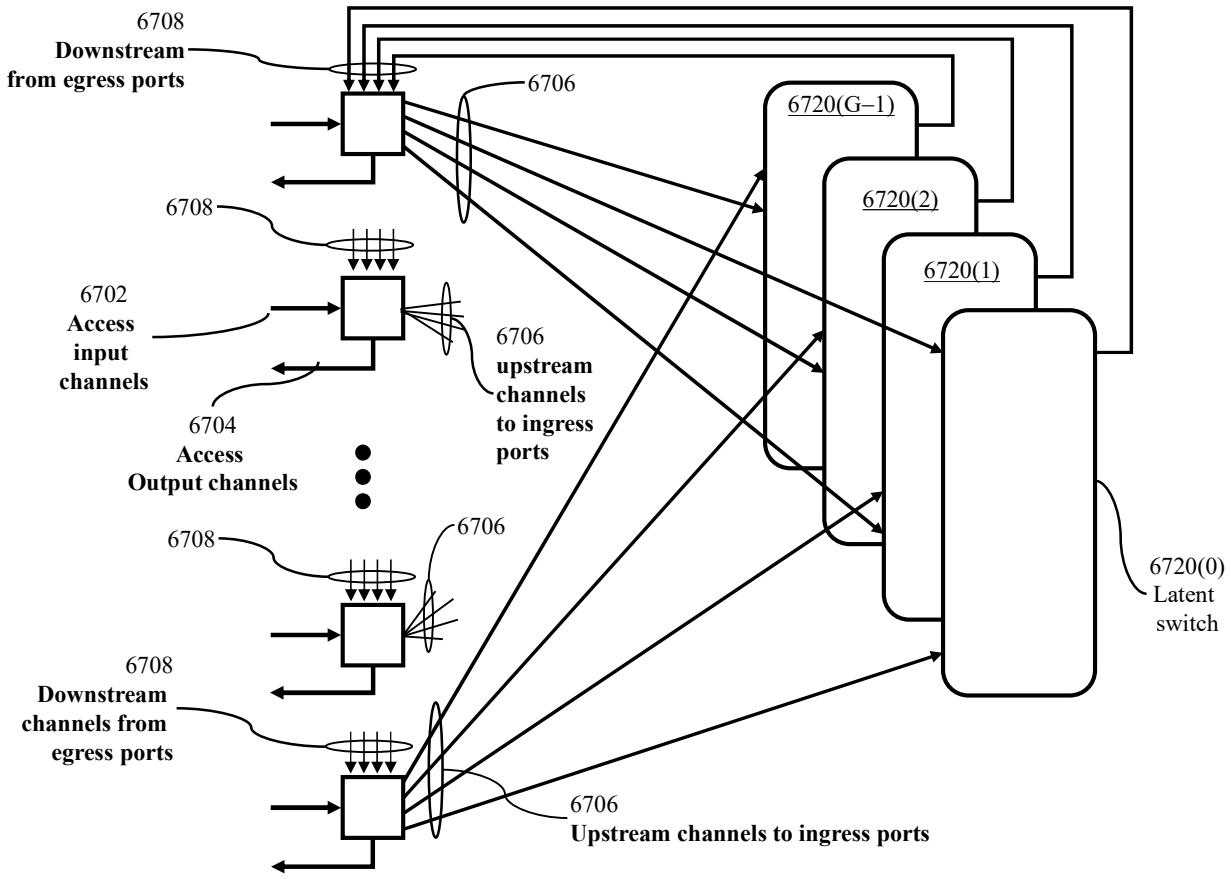
FIG. 28 of US 10,021,025

InventionShare®

InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

Family X
An electronic large-scale datacenter
or a geographically spread network of global coverage

US patent 9,219,697 of **Family X** targets a switching system which may serve as a large-scale data center or a geographically distributed network of global coverage. Access nodes of moderate dimensions are interconnected through central switches of large dimensions to form a large-scale switching system. The central switches are configured as latent space switches (Family V) which scale easily to large dimensions. Each access node has asymmetric connections to the input sides and output sides of the central switches so that paths from an originating access node to a destination access node through the central switches are subject to staggered switching delays permitting an access controller of any access node to select an available path of minimum switching delay for a given flow. All of the control data and payload data traverse respective one-hop paths. The switching-system's structure facilitates both distributed and centralized control.

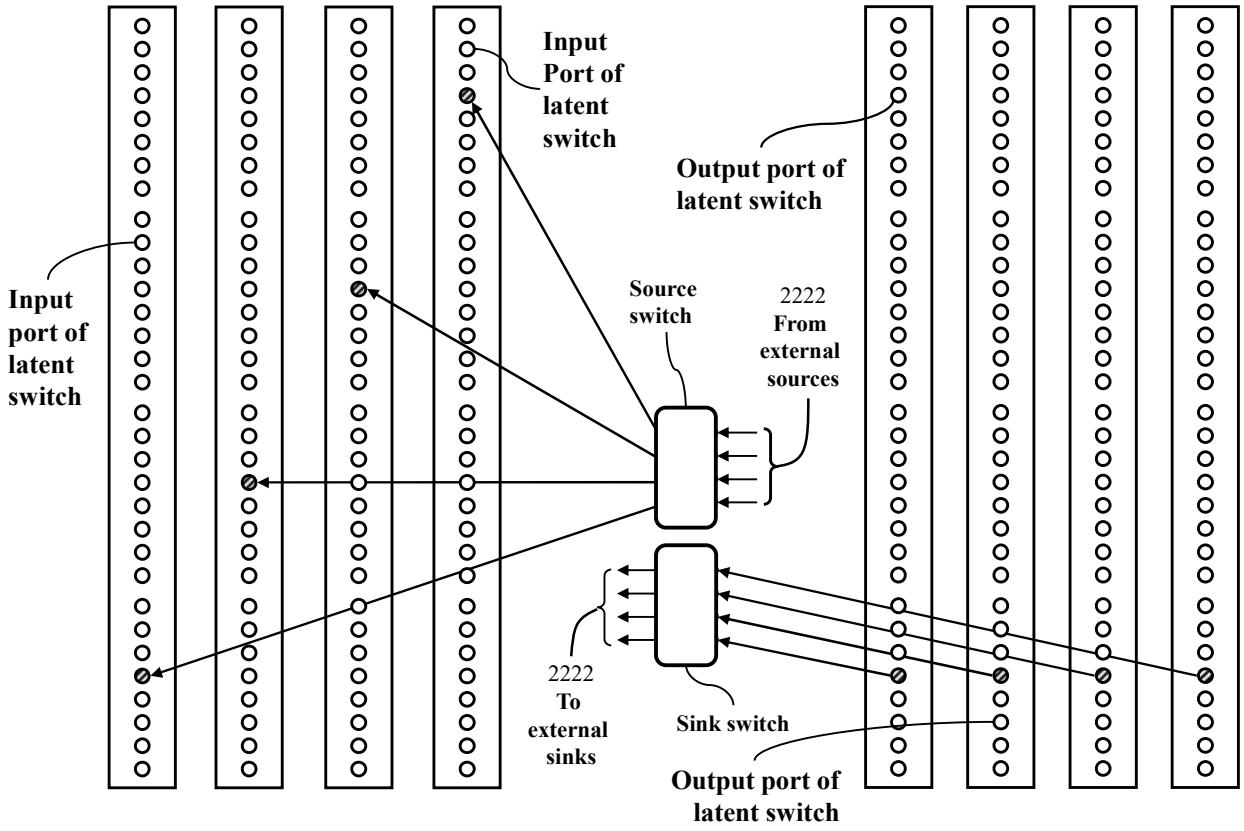


InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

FIG. 67 of US 9,219,697

For a data-center application, using access nodes of 128 dual ports each and central switches of 16384 dual ports each, for example, a switching system of more than 2 million dual ports is realized. At a port capacity of 40 Gb/s, the access capacity exceeds 80 Pb/s and the bulk of traffic experiences a switching delay below five microseconds. This capacity is significantly larger than the envisaged capacity of current data-center designs.

FIG. 67 of US 9,219,697 (reproduced below) illustrates a switching system 6700 having a number of independent central switches 6720 interconnecting a plurality of access switches 6710. Each central switch 6720 has a respective master controller (not illustrated in FIG. 67) and each access switch 6710 has a respective access controller (not illustrated in FIG. 67). Each access switch 6710 receives data from external data sources through access input channels 6702 and transmits data to external data sinks through access output channels 6704. Data received at input ports of the access switches 6710 may include packets of arbitrary lengths and, to facilitate switching, the data may be segmented into data segments of equal size and switched as such through the switching system 6700. The switched data segments are re-assembled at output ports of access switches 6710 to reproduce the packets in the forms in which they were received.



InventionShare®

InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

FIG. 85 of US 9,219,697

FIG. 85 of US 9,219,697 illustrates a first connectivity scheme of an access switch to the central switches. The illustrated connectivity scheme of an access switch to four central switches aims at minimizing switching latency. In the illustrated example, each central switch has only 32 ingress ports and 32 egress ports for ease of illustration; for a large-scale data center, the number of dual ports of each central switch may exceed 10,000. Each central switch is preferably configured as a single-rotator latent space switch (Family V) employing a transposing rotator alternatively connecting ingress ports to a bank of transit-memory devices and the bank of memory devices to egress ports.

Access nodes of 512 dual ports each and central switches of 16384 dual ports each may be deployed to create a switching system of more than 8 million dual ports. At a port capacity of 40 Gb/s, the throughput exceeds 320 Pb/s.

FIG. 86 of US 9,219,697 illustrates a second connectivity scheme of an access switch to the central switches.

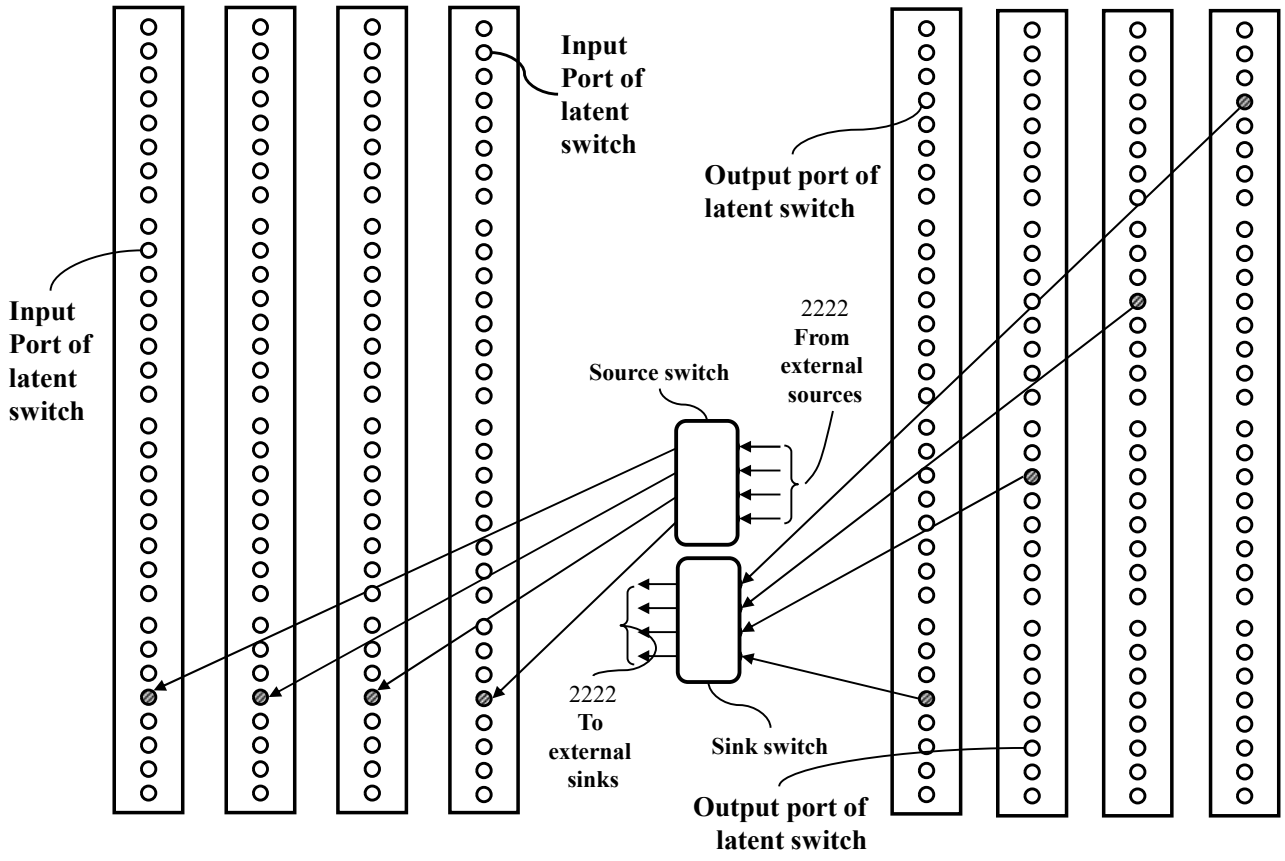


FIG. 86 of US 9,219,697

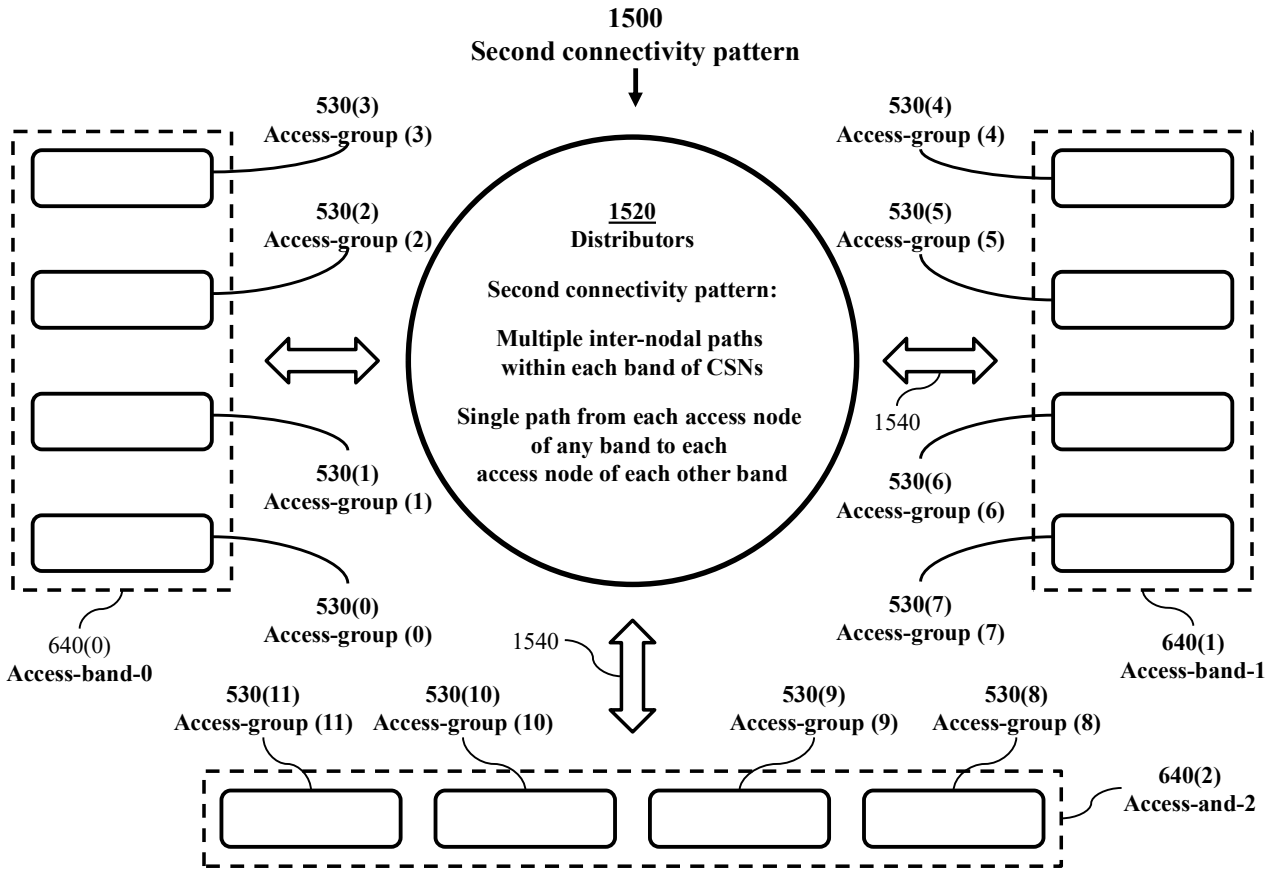
InventionShare®

InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

Family XI

Exabits-per-second global-coverage low-latency data network based on network fusion

Family XI includes U.S. patents (10,757,488), (11,206,467), (11,356,240), (11,616,735), and (12,273,228). A technique of "fusing" multiple Clos star networks (CSNs), for example, fusing 16 to 1024 CSNs, which preserves the conventional CSN properties while providing at least a single-hop path from each access node of each constituent CSN to each access node of each other constituent CSN is disclosed in U.S. Patents 10,757,488 and 11,206,467. An enhanced "deep fusing" technique (US Patents 11,616,735 and 12,273,228) divides the constituent CSNs into multiple CSN bands providing multiple **parallel** single-hop paths from each access node to each other access node of a **same CSN band** and a single-hop path from each access node of any CSN band to each access node of each other CSN band. FIG. 15 of US12,273,228 (reproduced below) illustrates a connectivity pattern of the access groups to the distributors to form a contiguous network based on second-order fusing of Clos star networks.

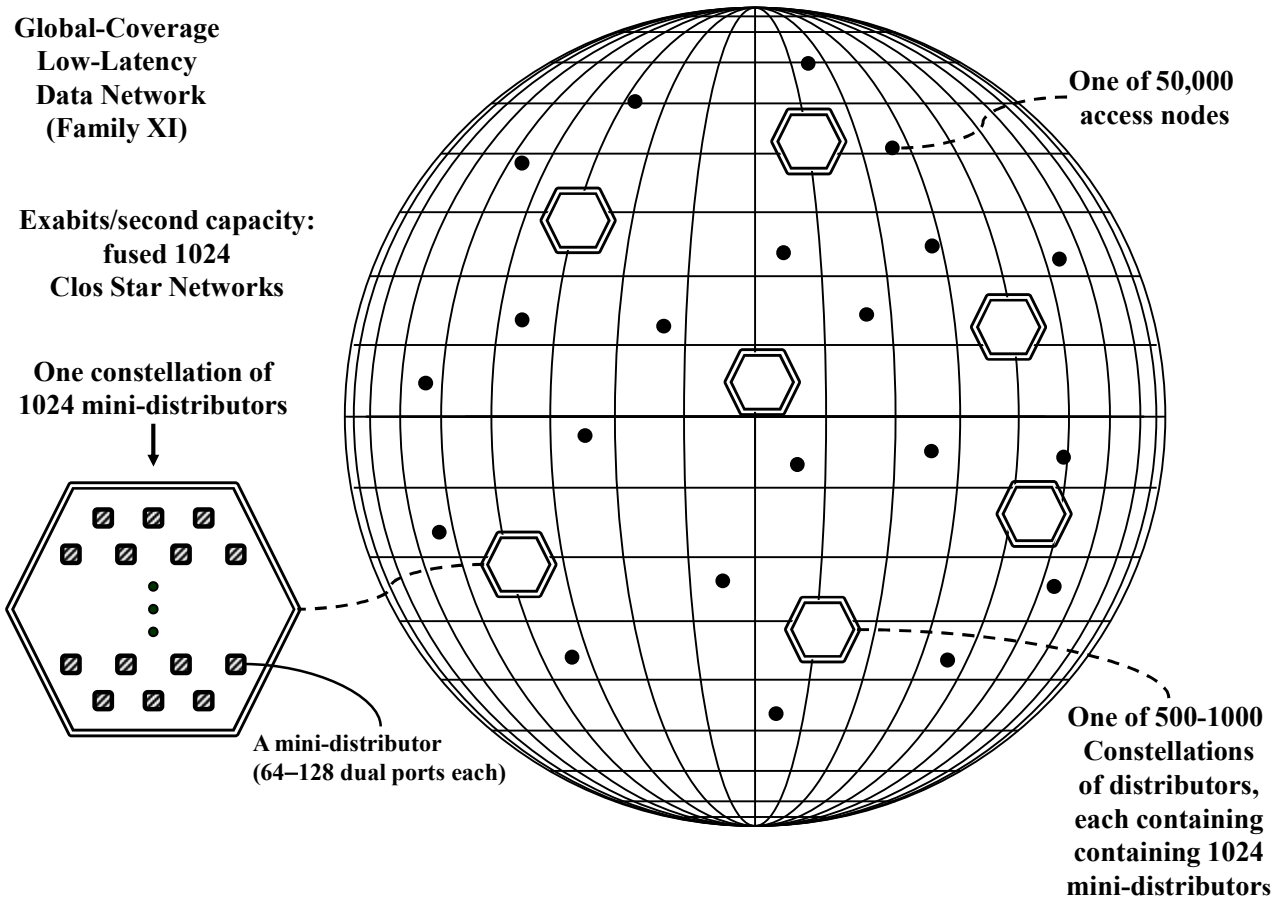


InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

FIG. 15 of US 12,273,228

Dividing 1024 CSNs, for example, into 32 CSN bands, each CSN band containing 32 CSNs, creates 32 subnetworks, each subnetwork having 1024 access nodes that are heavily interconnected through parallel single-hop paths. Each access node of each subnetwork still has a single-hop path to each access node of the other 31 subnetworks.

An exemplary global-coverage data network, illustrated below, supports numerous access nodes (50,000, for example) interconnected through 500-1000 ***constellations of electronic or optical distributors***, with each access node connecting, and ***time aligned***, to only 32 constellations yet having at least one single-hop path to each other access node. The throughput of the network is of the order of an exabit/second (10^{18} bits/second). A technique of time aligning an access node to any distributor is disclosed in US Patent 11,356,240. The traditional network layer, *colloquially* known as "cross connects", is eliminated. The Drawing below illustrates a global network functioning as a single Clos Star Network.



InventionShare®

InventionShare® is the exclusive global agency for the licensing or acquisition of AcceleRoute Inventions

Family XII

Exabits-per-second global-coverage, low-latency, data network built on constellations of spectral-temporal multiplexers

The conventional Clos star network (CSN) provides a number of parallel paths from each access node to each other access node of a set of access nodes through a set of distributors (switches/routers) with each path traversing only one distributor. A technique for extending the coverage (number of access nodes) and throughput (bits/second) of a single CSN, connects the access nodes of the single CSN to the distributors (switches/routers) through passive ***spectral-temporal multiplexers*** (US Patents 9,509.432 and 12,015,887) arranged in constellations. An overview is illustrated below.

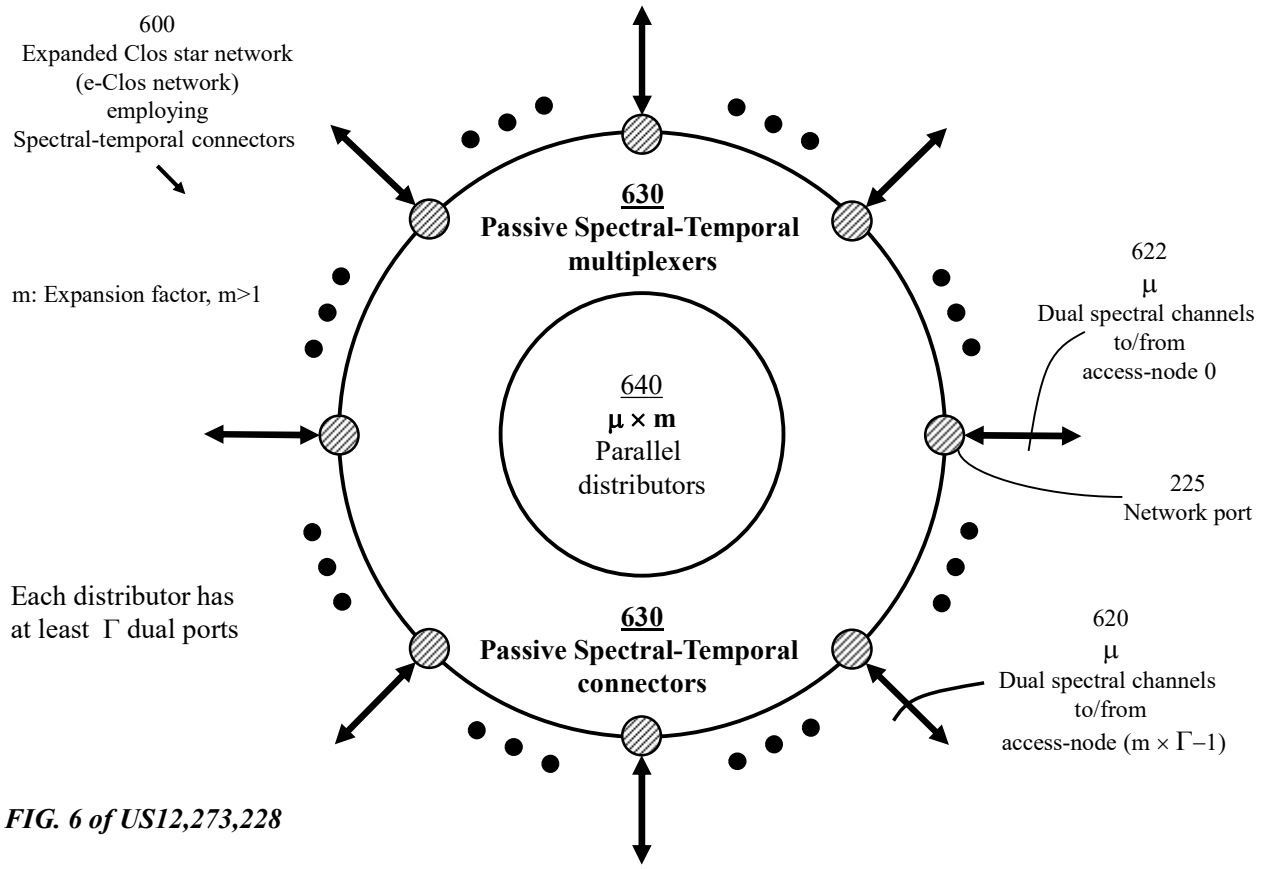


FIG. 6 of US12,273,228

An exemplary global-coverage network interconnecting 32,768 access nodes through 32,768 distributors is illustrated below. The access nodes are connected to the distributors through 32 ***constellations*** of passive spectral-temporal multiplexers. Each constellation has 1,048,576 dual access ports

facing the access nodes and a same number of dual core ports facing the distributors. Two arrangements of connecting access nodes and distributors to the spectral-temporal connectors are described in US Patent 12,015,887. The access nodes and the distributors are generally geographically spread.

The resulting network is **still a single CSN** but at a (much) larger scale. The network gracefully handles time-varying spatial traffic distribution so that, in general, source data received at the access ports of any access node may be directed to inner ports of numerous other access nodes, or, whenever needed, the entire source data at an access node may be directed to inner ports of only one other access node. This property **eliminates the need for conventional traffic engineering** since any spatial distribution of data traffic is automatically accommodated. The realizable throughput of the network is of the order of an exabit/second (10^{18} bits/second).

A first arrangement (FIG. 24 of US 12,015,887) divides the **access nodes** into 32 groups of 1024 access nodes each, and connects the access nodes of each group to access ports of a respective **constellation of spectral-temporal multiplexers**. Each distributor of the 32,768 distributors has a dual multichannel link to respective dual core ports of each of the 32 constellations.

A second arrangement (FIG. 27 of US 12,015,887) divides the **distributors** into 32 sets of 1024 distributors each and connects the distributors of each set to core ports of a respective **constellation of spectral-temporal connectors**. Each access node of the 32,768 access nodes has a dual multichannel link to respective dual access ports of each of the 32 constellations.

Variations of the above first arrangement and the second arrangement where the number of access nodes equals the number of distributors and each access node is integrated with a respective distributor produce mesh structures illustrated in FIG. 36 and FIG. 39, respectively, of US 12,015,887. An advantage of either of the mesh structures is the availability of a direct path (**zero hop)** from each access node to each other access node in addition to the numerous parallel single-hop paths. This is realized at the expense of scalability; for example, under the same scalability constraints of individual components, a coverage of 32768 access nodes with a throughput of approximately 1000 petabits/second of the arrangement of FIG. 24 would be reduced to a coverage of 16384 with a throughput of 500 petabits/second (at a higher efficiency). However, a direct path for each pair of access nodes brings in tremendous control advantages. Figures 40, 41, and 42 of US12,015,887 illustrate exemplary schemes of establishing paths through the mesh network of FIG. 36 or FIG. 39.

The Drawing below illustrates a global network accommodating thousands of globally spread access nodes yet operating as a single Clos Star Network.

