



AcceleRoute™

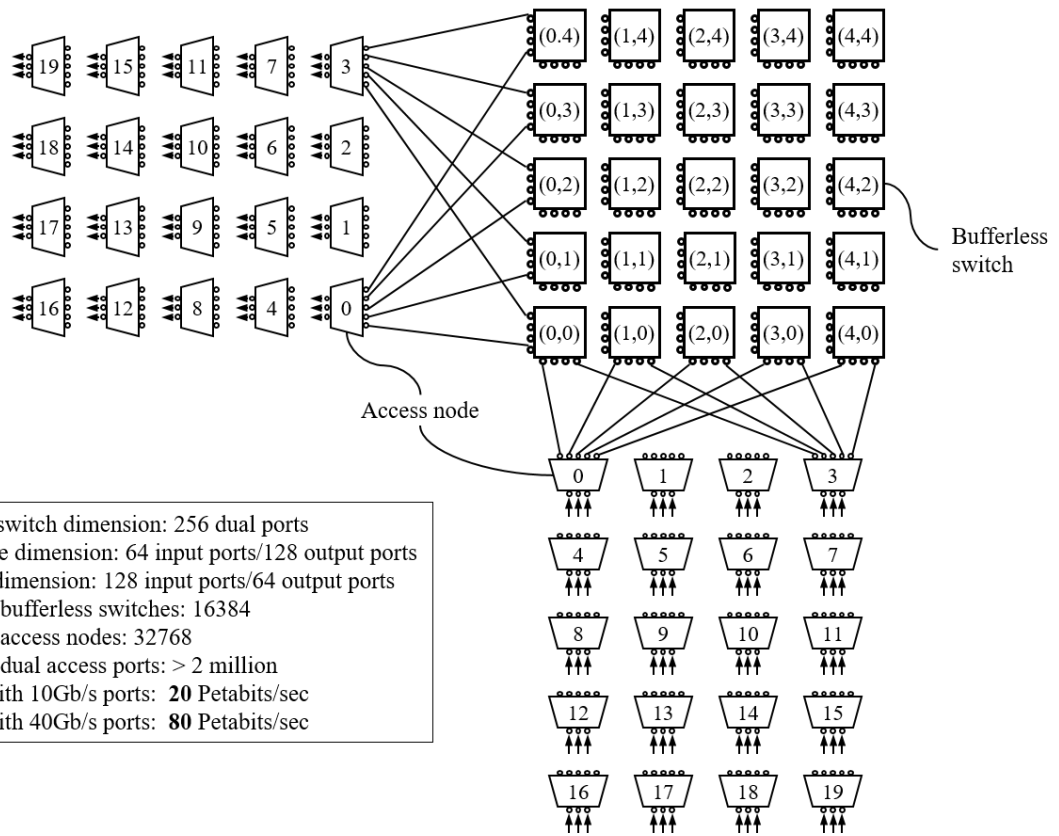
A lightning-fast network architecture for Cloud Data Centers

Family I

Scalable network architecture based on bufferless switches.

The patents of **Family I** target a large-scale data center providing a switching system that scales to a capacity of tens of thousands of terabits per second (Tb/s). Independent space 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 is devised to naturally equalize data flow rates through the switches. Each access node has upstream channels to a selected set of bufferless switches and downstream channels from a different set of bufferless switches which is orthogonal to the selected set. Optionally, the input sides of all of the bufferless switches connect to mutually orthogonal sets of access nodes.

Family-I Data Center Application



A data center of a capacity of 80 Pb/s is realizable; this capacity is more than an order of magnitude larger than the envisaged capacity of current designs. All control data and the bulk of payload data traverse one

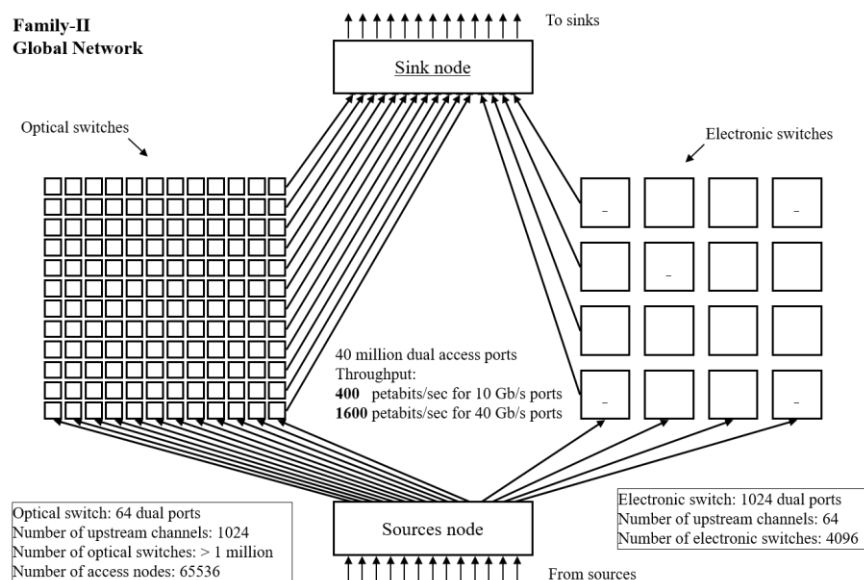
hop. Payload data is buffered only at the edge; *en route* buffering – a major source of routing problems and latency – is eliminated. The network structure is SDN-friendly and facilitates both distributed and centralized control.

Family II

A Time Coherent Global Network

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; this capacity is more than two orders of magnitude larger than the envisaged capacity of current designs. As in **Family I**, all control data and the bulk of payload data traverse one hop. Payload data is buffered only at the edge; *en route* buffering – a major source of routing problems – is eliminated. The network structure is SDN-friendly and facilitates both distributed and centralized control.



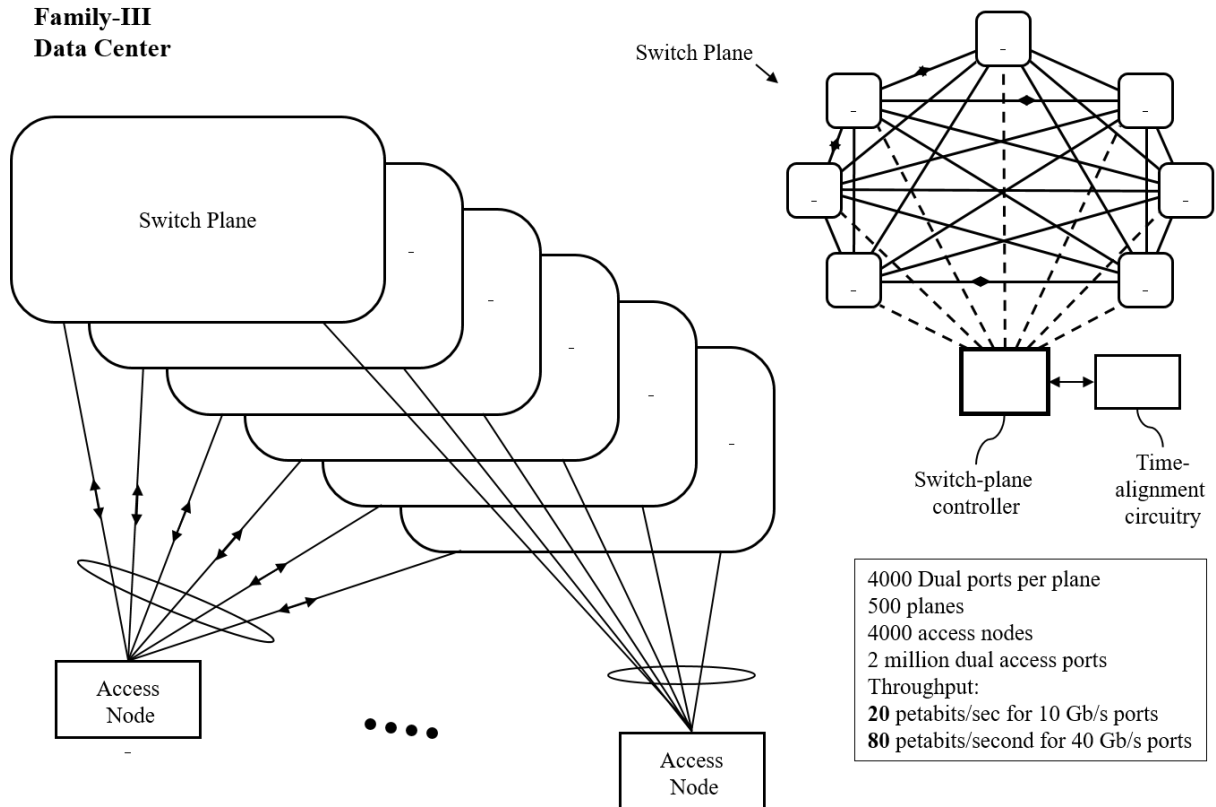
Optionally, the bufferless switches may be organized into two planes; a first plane comprising independent fast optical switches, and a second plane comprising bufferless electronic switches.

Family 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 thousands of terabits per second (Tb/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 or spread across a continent. The connectivity pattern of source nodes and sink nodes to the switch planes is devised so that each pair of a source node and a sink node connects only once to a common switch. 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 one switch.

Family-III Data Center



A data center or a continental-coverage network of a capacity of 80 Pb/s is realizable; this capacity is more than an order of magnitude larger than the envisaged capacity of current designs. All control data and the

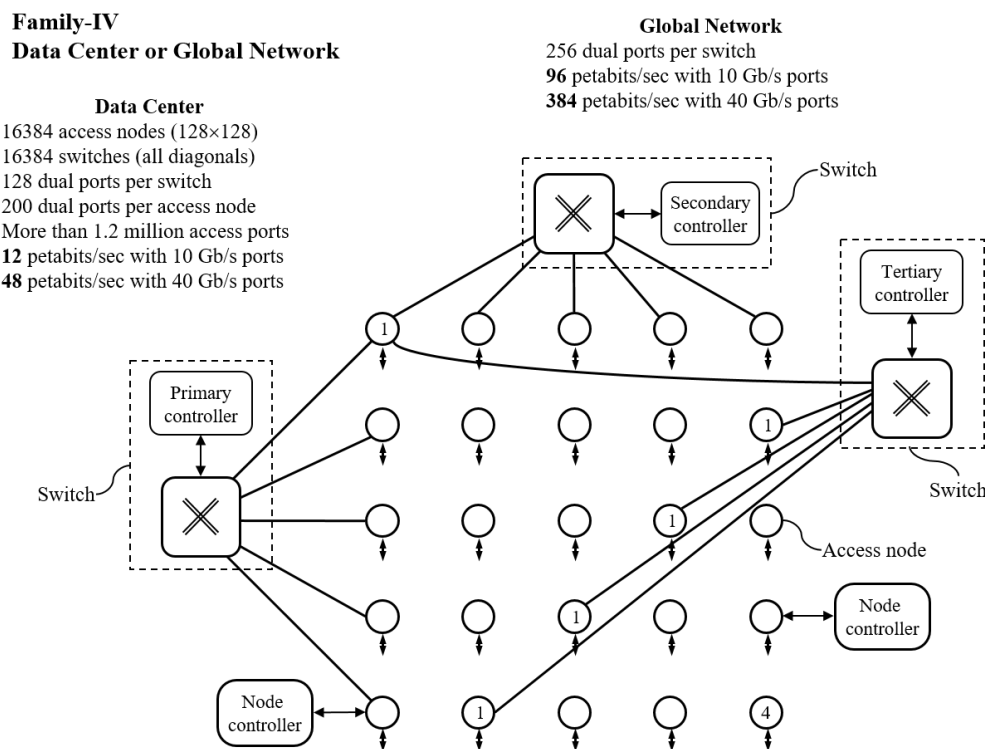


bulk of payload data are transferred through a single switch plane. Payload data is buffered only at the edge; *en route* buffering – a major source of routing problems – is eliminated. The network structure is SDN-friendly and facilitates both distributed and centralized control.

Family IV

Large-scale data center or geographically distributed network of global coverage

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. With a sufficient number of formations, each access node may have a path to any other access node traversing only one switch.



A data center of a capacity of 48 Pb/s is realizable; this capacity is more than an order of magnitude larger than the envisaged capacity of current data-center designs. A wide-coverage network of a capacity exceeding 380 Pb/s is realizable; this capacity is more an order of magnitude larger than the envisaged capacity of current network designs.



With a sufficient number of formations, all control data and the bulk of payload data traverse one hop. Payload data is buffered only at the edge; *en route* buffering – a major source of routing problems – is eliminated. The switching-system structure is SDN-friendly and facilitates both distributed and centralized control.

Family V

Electronic Large-scale data center or geographically distributed network of global coverage

The patents of **Family V** target 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 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 one-hop paths. The switching-system structure is SDN-friendly and facilitates both distributed and centralized control.

For a data-center application, using access nodes of 128 dual ports each and central switches of 16384 dual ports each, 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.

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 access capacity exceeds 320 Pb/s. This capacity is significantly larger than the envisaged capacity of current



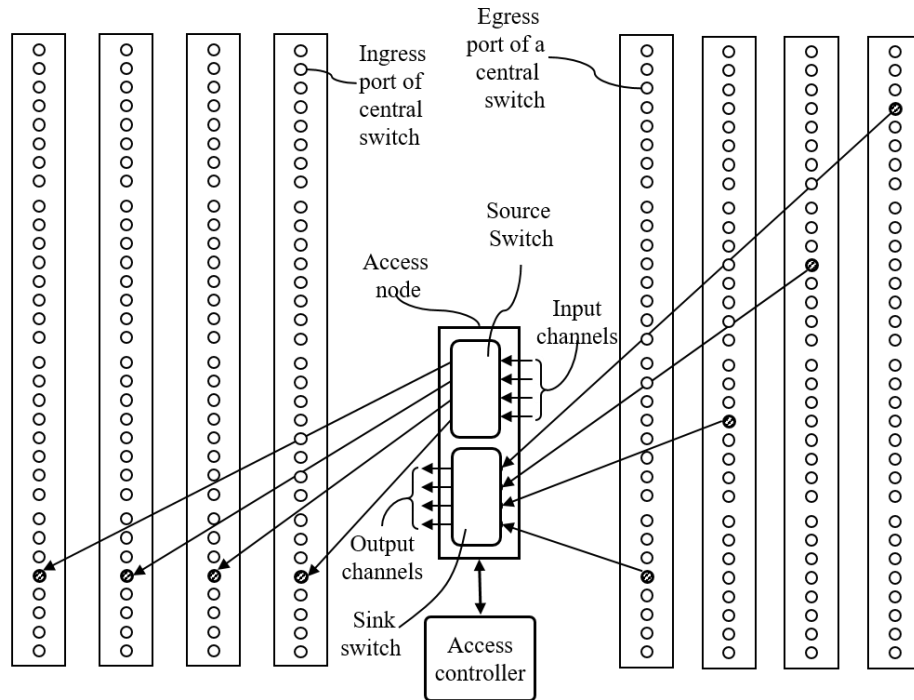
network structures.

Family-V

Data Center or Global network

Data Center
128 Central switches
16000 dual ports each
16000 access nodes
2 million dual access ports
Throughput:
20 Pb/s for 10 Gb/s ports
80 Pb/s for 40 Gb/s ports

Global Network
512 Central switches
16000 dual ports each
16000 access nodes
8 million dual access ports
Throughput:
80 Pb/s for 10 Gb/s ports
320 Pb/s for 40 Gb/s ports



Family VI

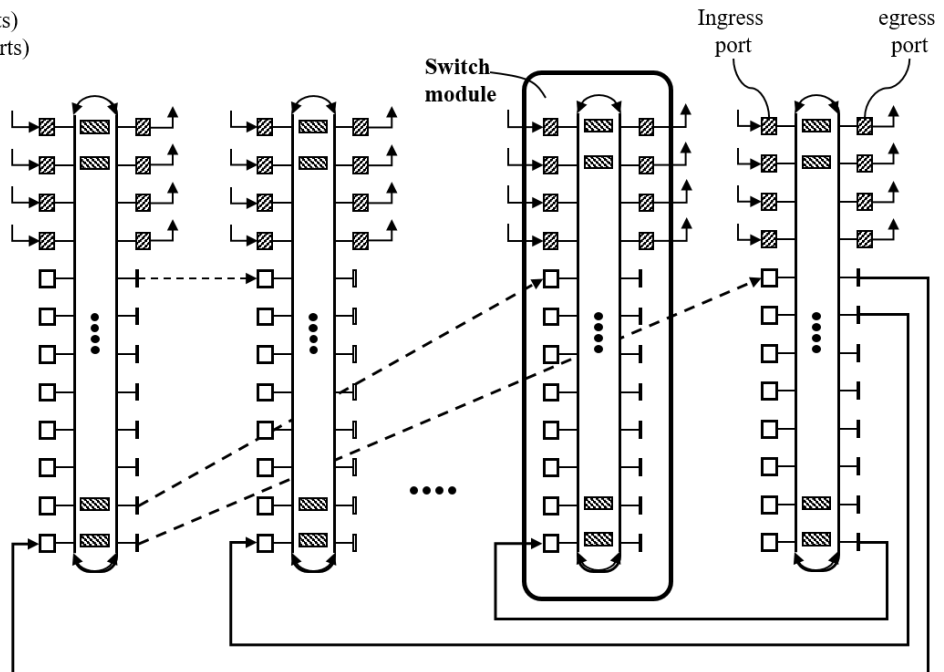
Packet switches based on contention-free switch modules

The patents of Family VI disclose various packet switching systems based on interconnecting contention-free switch modules. In one implementation (illustrated 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.



Family-VI
Switching system using contention-free switch modules

984 access ports
Throughput:
9.8 Terabits/sec (10 Gb/s ports)
39.2 Terabits/sec (40 Gb/s ports)



Family VII

A contiguous switching system for cloud communications or Large-Scale Data Centers

The patents of **Family VII** disclose a contiguous switching system for a large-scale cloud communications network or a 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 large-scale data center.

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

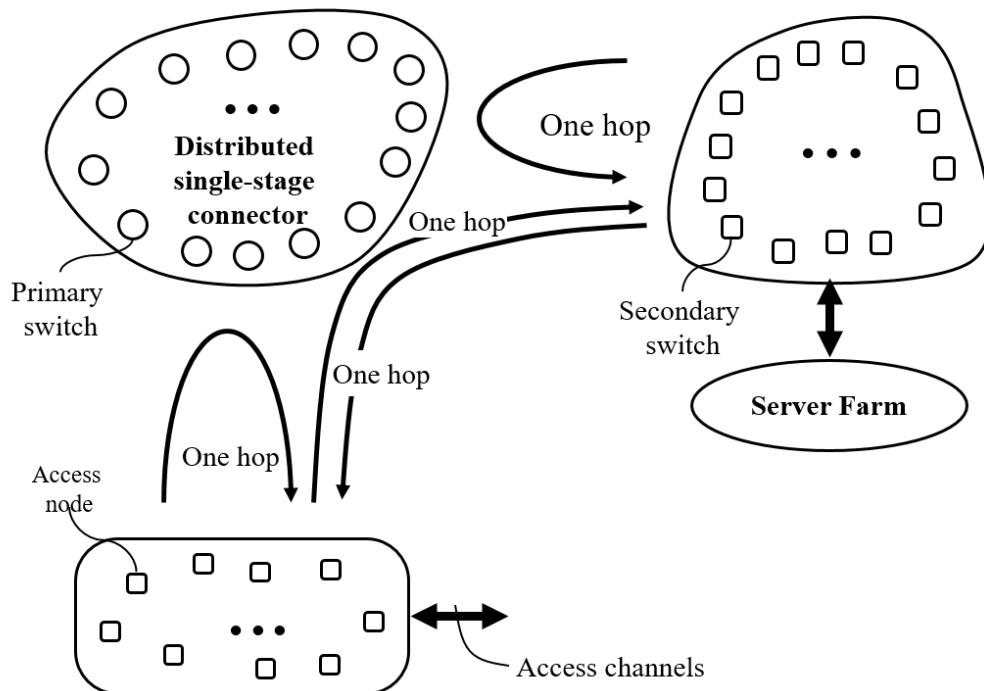


each access node to each secondary switch, and from each secondary switch to each access node. The connectivity pattern enables network scalability to accommodate hundreds of thousands of access nodes while permitting a significant proportion of traffic to be routed through single-hop paths.

The structure of the switching system greatly facilitates accommodating numerous virtual networks.

Family-VII
Contiguous Network
Global network and Large-Scale Data Center

Throughput:
400 Pb/s with 10 Gb/s ports
1600 Pb/s with 40 Gb/s ports



Family VIII

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

Family VIII of pending patent applications discloses 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.

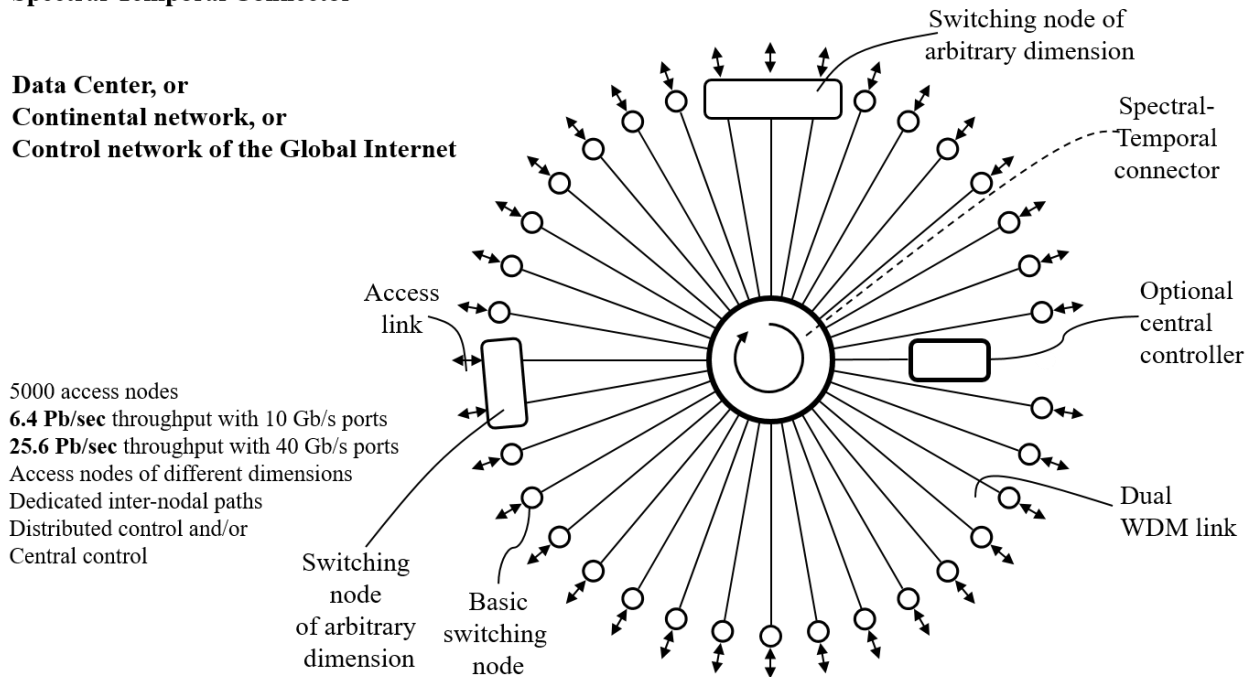
Each access node has a direct permanent path to each other access node of a capacity of the order of 1 Gb/s. Data may be transferred from an access node to another through a respective direct path and numerous other paths each traversing one intermediate access node. A configuration supporting 5000 access 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.

Family-VIII
Spectral-Temporal Connector

**Data Center, or
Continental network, or
Control network of the Global Internet**



Family IX

Distributed Control of a Vast Communication Network

Family IX of pending patent applications discloses 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. 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
Routing control of a legacy network

Spectral-Temporal Connector of Family-VIII facilitates both distributed and central control of the global Internet

