128 TECHNOLOGY

NG-WAN GUIDE

CONTENTS

NOTICE

THE SPECIFICATIONS AND INFORMATION REGARDING THE PRODUCTS IN THIS GUIDE ARE SUBJECT TO CHANGE WITHOUT NOTICE. ALL STATEMENTS, INFORMATION, AND RECOMMENDATIONS IN THIS MANUAL ARE BELIEVED TO BE ACCURATE BUT ARE PRESENTED WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. USERS MUST TAKE FULL RESPONSIBILITY FOR THEIR APPLICATION OF ANY PRODUCTS.

THE SOFTWARE LICENSE AND LIMITED WARRANTY FOR THE ACCOMPANYING PRODUCT ARE SET FORTH IN THE INFORMATION PACKET THAT SHIPPED WITH THE PRODUCT AND ARE INCORPORATED HEREIN BY THIS REFERENCE. IF YOU ARE UNABLE TO LOCATE THE SOFTWARE LICENSE OR LIMITED WARRANTY, CONTACT YOUR 128 TECHNOLOGY, INC. REPRESENTATIVE FOR A COPY.

NOTWITHSTANDING ANY OTHER WARRANTY HEREIN, ALL DOCUMENT FILES AND SOFTWARE OF THESE SUPPLIERS ARE PROVIDED “AS IS" WITH ALL FAULTS. 128 TECHNOLOGY DISCLAIMS ALL WARRANTIES, EXPRESSED OR IMPLIED, INCLUDING, WITHOUT LIMITATION, THOSE OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE.

IN NO EVENT, SHALL 128 TECHNOLOGY OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THIS MANUAL, EVEN IF 128 TECHNOLOGY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

ABOUT THIS GUIDE

PURPOSE

This document identifies the design considerations and efforts required to design and deploy a Next-Generation Wide Area Network (NG-WAN) that serves to meet legacy WAN and traditional Software-defined WAN (SD-WAN) requirements in addition to preparing the network for future applications. Every deployment has unique characteristics and associated needs. Readers of this document should consult their technical advisors and adapt to their specific deployment scenarios as needed.

AUDIENCE

This guide is for internal use only. The document assumes that the reader has an architectural understanding of the 128T Networking Platform and associated benefits of session-oriented networking.

RELATED DOCUMENTS

The following documents will provide additional information:

128T Reference Guide

128T Installation Guide

128T Configuration Guide

128T Berkeley Packet Filter Syntax Guide

128T Networking Platform License and Entitlement Document

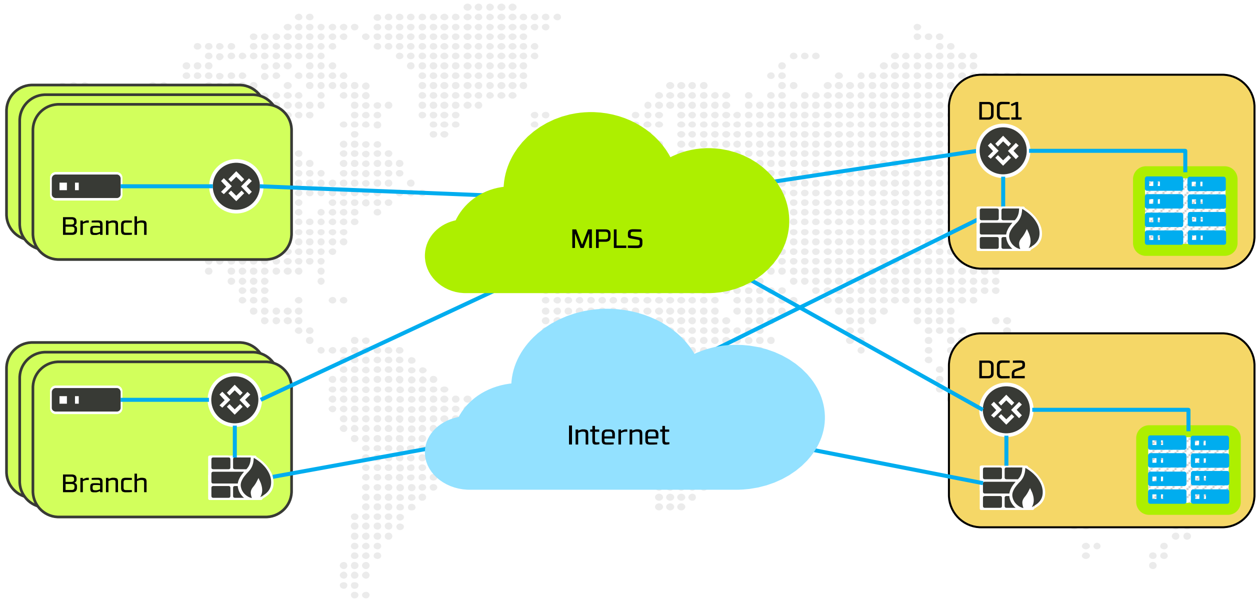
INTRODUCTION

The 128 Technology solution utilizes an innovative service-centric, session-oriented, and security-infused routing paradigm for building context aware networks. It enables centralized control, simplifies deployment of context aware networks, introduces intelligent service routing with in-band signaling, provides fine grained micro-segmentation, and infused security based on a zero-trust model.

This document describes the different aspects of deploying the 128T NG-WAN solution. It is not a comparison of benefits versus traditional SD-WAN solutions. There are different deployment options discussed in this document. It is meant to be a guideline on what all can/may be done as opposed to being a gospel or some absolute criteria on what should be done in all possible cases.

SOLUTION ARCHITECTURE

This section explains the basic concepts of the 128T NG-WAN solution architecture. The solution enables an enterprise or service provider to combine traditional WAN private links such as Multiprotocol Label Switching (MPLS) links with other low-cost alternatives such as Internet and/or 4G/LTE. The solution maximizes performance while reducing costs and improving reliability by making optimal use of link resources and making intelligent path selection decisions. We consider a simple WAN deployment scenario for illustration purposes.



This typical existing WAN deployment has the following characteristics before the 128T solution is introduced:

- Data Centers have connectivity via the private MPLS link and public Internet link
- Most branches have connectivity via the private MPLS link and public Internet link
- Some branches only have connectivity via the private MPLS link
- 4G/LTE does not provide any connectivity
- Applications are hosted at the Data Centers
- DC1 and DC2 share loads among themselves
- Branches are configured to split loads between DC1 and DC2

Introducing the 128T Networking solution will bring the following additional benefits:

- Seamless session based failover between MPLS and Internet paths to ensure near 100% application uptime without any changes to applications
- Immediate augmentation of MPLS only branches with 4G/LTE to improve resiliency
- Session duplication for high-value traffic to ensure near zero-packet loss
- Session maximization to augment speeds of all branches to aggregated capacity of all available paths

- Application awareness and SLA guarantees using session migration to specific traffic
- Intelligent load balancing of flows between data centers
- No overhead from IPSec, VPNs, or other overlay technologies leading to bandwidth savings
- Adaptive encryption savings to prevent double encryption
- Integrated load balancing, firewall, WAN optimization, and other functionality

TOPOLOGY

This section describes the physical topology options for inserting the 128T Networking Platform into an existing WAN deployment or building a new NG-WAN deployment. For simplicity, the network options shown consider the data center and the branches of a network. The following two topology options can be utilized:

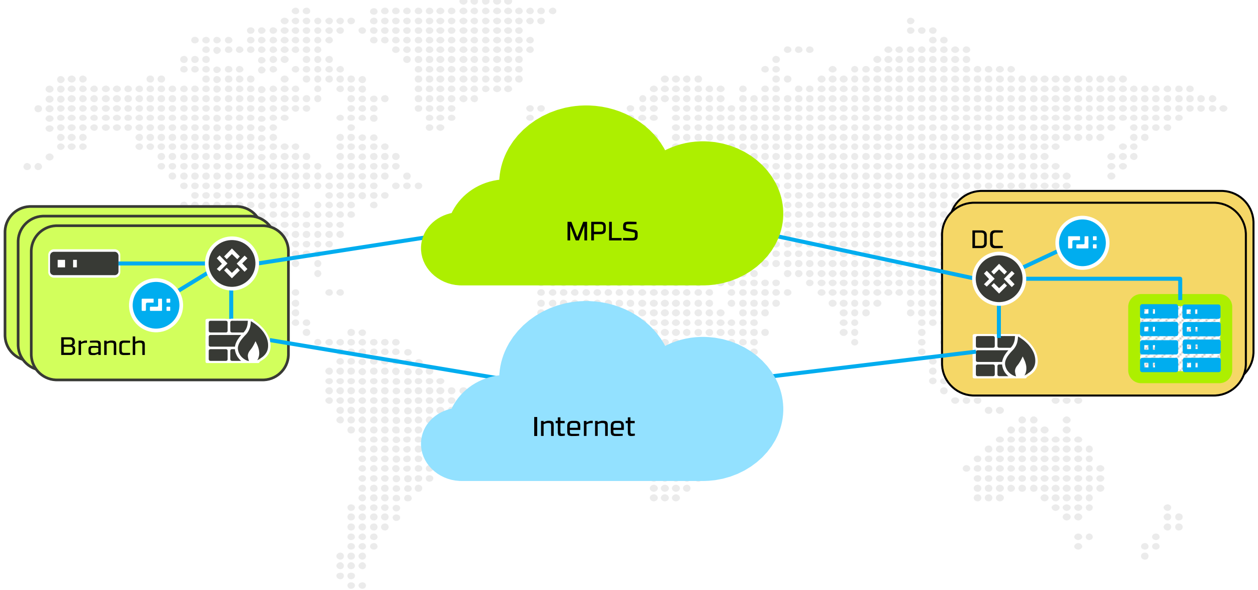
- One-Arm
- In-Line

Note: In the One-Arm topology the 128T router is logically in-line with one arm of the 128T router physically connected to the existing in-line router.

One-Arm topology is recommended for existing deployments to migrate seamlessly to the 128T solution. In-Line topology is recommended for new sites in existing deployments or for new deployments. One-arm topology has the advantage of being able to support migration of services to the new paradigm when desired while in-line topology provides the migration of all services immediately. The choice of topology depends on deployment requirements.

One-Arm

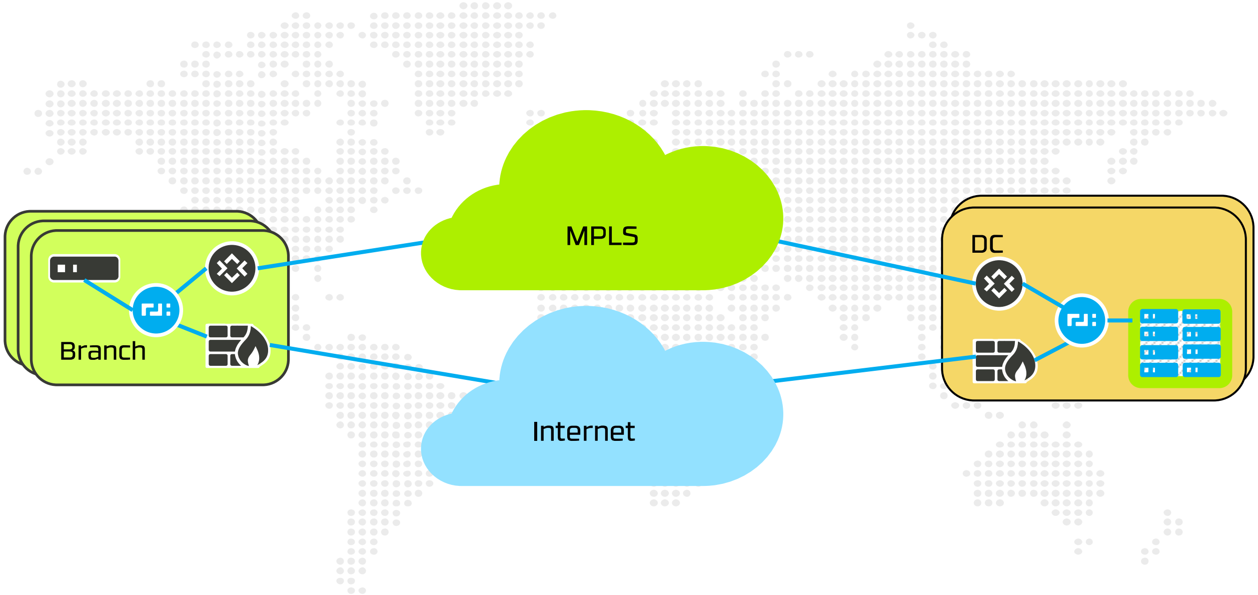
In one-arm (also known as stub router or "router on a stick") topology, the 128T router has a single physical connection to the network. This deployment is advantageous to introduce the 128T solution incrementally.



Services can be incrementally migrated to the 128T router. The existing router redirects all relevant traffic to the 128T router in both directions.

In-Line

In in-line topology, the 128T router sits directly inline at the edge of the network. It can be placed behind the existing firewall facing the Internet and behind the existing Customer Edge (CE) router facing the MPLS link.



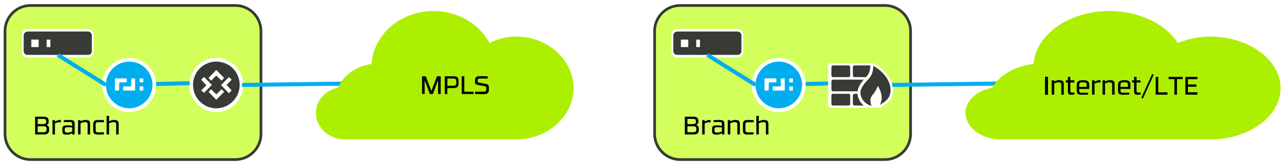
BRANCH SITES

Routers in an NG-WAN deployment can have various WAN transport combinations as well as router configuration modes to achieve high degree of service resiliency. In this document branch sites refer to the different remote locations that need to be interconnected with each other and with larger central office or data center sites.

Design	WAN Routers	WAN Transports	Primary	Secondary
Single Transport	Single	Single	MPLS/Internet/LTE	
Dual Transport	Single	Dual	Internet	MPLS/Internet/LTE
Dual Transport	Dual	Dual	Internet	MPLS/Internet/LTE

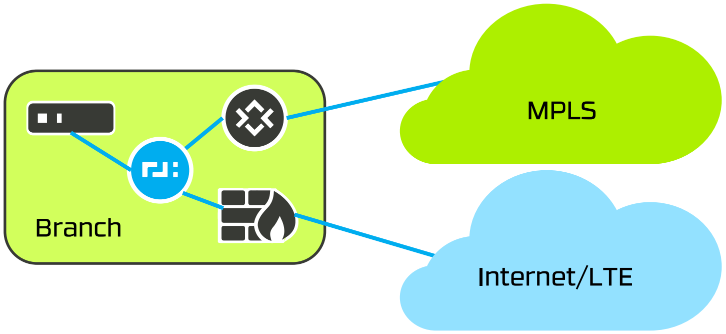
There may be tertiary transport available in many sites. The 128T routers can integrate WAN transports as long as there are available ports on the router.

Single Router Single Transport WAN Edge



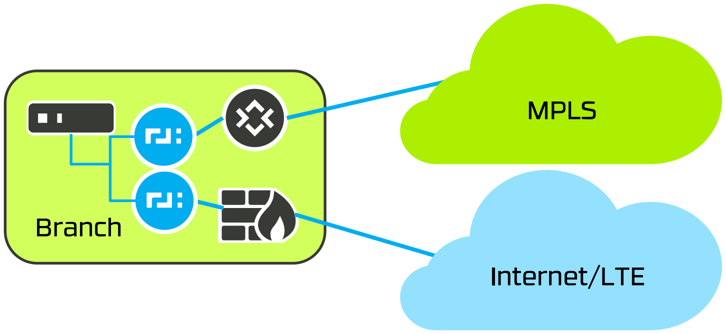
Most single transport WAN edge sites can be realized with a single router. Dual routers may be used for LAN redundancy. The diagrams show in-line topology however the design can accommodate one-arm topology too. This can easily be extended to include other hybrid transport models as they become available. Single router site deployments do not support link or router redundancy. They are suitable for small or remote sites where there is no business justification to have a second WAN transport. 4G/LTE can easily be added to these sites to have a path of last resort if desired.

Single Router Dual Transport WAN Edge



Most sites are expected to have dual WAN transports. These sites can be realized using a single router. This type of deployment supports link resiliency. It also ensures that a low-cost Internet path is used as the primary WAN interface. MPLS can serve as the secondary WAN interface. It can also serve as the primary WAN interface for high value traffic. LTE can serve as the path of last resort. The diagram shows in-line topology however the design can accommodate one-arm topology too. Existing routers and firewalls that are already present can continue to be used as shown. They can also be eventually replaced with the 128T router itself to provide integrated functionality.

Dual Router Dual Transport WAN Edge



Certain sites require a dual router WAN design including regional offices and campus locations with large number of users. The dual router dual transport WAN edge design ensures complete link and router redundancy removing all single points of failure. Similar to single router dual transport WAN edge deployments the WAN links can be used in various degrees of preference. The diagram shows in-line topology however the design can accommodate one-arm topology too.

DATA CENTER SITES

Data Center or larger central office sites are expected to have multiple WAN transports. They are also expected to have a few sites for redundancy and load balancing. The 128T NG-WAN solution allows organizations to scale their network and load balance traffic to maximize efficiency and lower costs. These data center sites are expected to have connectivity to all branch sites. They can also be split by geographic regions or proximity as desired. Data center sites communicating with different branch sites having overlapping IP addresses can distinguish between branch sites by relying on a unique router/peer identifier. It is expected for router redundancy that the data center sites would follow the dual router dual transport site design. Quality points and server loads for different services can be used to balance data center usage or to intelligently redirect traffic to different server applications resident on different data centers to maximize efficiency.

DATA MODEL

The 128T router offers a new approach to IP networking by placing services at the heart of the NG-WAN design. Using the 128T data model, administrators describe the services that their NG-WAN solution will offer, and the properties of those services. The 128T router software delivers traffic to these services while abiding by the policies, constraints, and path properties that those services require. This data model is the language for describing the network, service, and policy architectures.

The data model is used to describe network, service, and policy behavior. The topmost configuration container in the 128T data model is called the authority, which is where system-wide global data is stored. Conceptually, the authority represents the complete set of all 128T routers managed under a single organizational entity. For example, if an organization called Acme Packet decides to deploy an NG-WAN solution for interconnecting different sites, the authority can be “Acme Packet”.

The global data within the “Acme Packet” authority container includes service-layer and policy-layer configuration that applies to all of the 128T routers within the Acme Packet organizational entity.

Service configuration, which represents the cornerstone of the 128T router’s worldview, is part of the set of global data within an authority. Services represent specific applications that a network delivers; e.g., Acme Packet may support web services, database services, or voice/video services. Using a top-down approach, the 128T data model asks that administrators define the services that their network will deliver, the requirements that the service demands (in terms of latency, packet loss, jitter, etc.), and the network topology – and the 128T router will deliver traffic to the service using the optimal paths through the network.

Services are said to reside within tenants, a term used to represent a segmented partition within a L2/L3 network. Unlike other networking paradigms, where segmentation is done using overlay networking techniques (such as VLANs, VxLANs, etc.), the 128T router uses a novel tenancy model to place traffic sources and routes to their services into logical partitions within the underlay network itself. A rich set of hierarchical access control policies built into the tenancy model ensures that network traffic flows along prescribed paths, and only from eligible sources. Tenants, like the services that they contain, are also part of the global data within an authority. A tenant defined within a 128T authority is said to “stretch” across all 128T routers that are members of that authority, and tenant information is shared between 128T router instances. For example, Acme Packet may have different tenants such as marketing, engineering, and others having access to different services depending upon their authorization.

A set of global policies rounds out the data model; complementing the router-specific policies, the global policies describe the treatment of traffic that flows between 128T routers. This includes information on how packets are classified into their various types (e.g., how to differentiate between web traffic, voice traffic, proprietary application traffic, etc.) and the requirements that those traffic varieties have from a networking perspective.

Qualified Name Services

Qualified Service Names, or QSNs, are a novel mechanism employed by the 128T router to describe collections of users, tenants, service agents (end devices that deliver services), or groups of service agents with a human-readable, URL-like syntax.

The 128T router uses QSNs as descriptive routing and access statements, where administrators can define policies related to these collections of users, tenants, service agents, and services and allow or deny them access to the resources of the 128T router.

QSNs are effectively a syntax for describing your network's resources. They have the following format:

qsn://tenant.authority/service-group/service

The most common application for QSNs is to define access-policy statements referencing other tenants, or service agents. For example, Acme Packet employees in the engineering department may be further split into development and test organizations. Development group may have access to IT tools like JIRA, bitbucket, etc.

qsn://dev.engineering.acmepacket/ittools/jira

Service Routes

Service routes, associated with service configuration elements, allow administrators to influence the 128T router's egress interface for that service's traffic. Unlike service configurations themselves (which are shared among all 128T routers that comprise an authority), service-route configuration elements are local data. This means that each router within an authority can have its own unique service-route configuration to affect traffic distinctly, per instance.

Conceptually, a service-route is akin to a traditional static route, but only for traffic specific to the associated service. That is, when packets arrive at a 128T router and match a known service (irrespective of whether that service was explicitly provisioned, or learned), then the 128T router uses its route selection algorithm to choose the most appropriate service-route for bearing that session-oriented traffic. This algorithm will consider the current traffic load, the provisioned capacities for the route(s) that match the service, the traffic distribution policy for this service, and the availability of the next-hop elements derived via Bidirectional Forwarding Detection (BFD).

Each service-route represents a single /32 host, and is colloquially referred to as a service agent. Conceptually, a service agent can be thought of as a discrete end device (host computer) capable of delivering a service; this may be a single computer running a web server within a web server farm, an individual PBX that is part of a larger call control group, a single database node that is a member of a database cluster, etc.

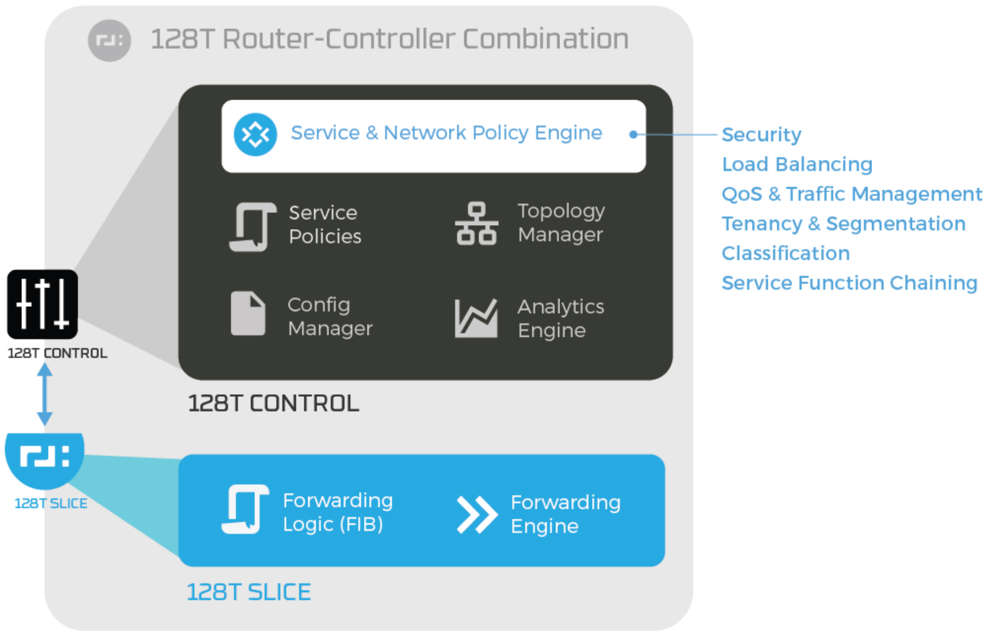
PRODUCTS

The NG-WAN solution requires the following products:

- 128T Software
- White-box Hardware
- Third-Party DevOps

128T SOFTWARE

The 128T Networking Platform is a new breed of software-based networking infrastructure. The 128T Networking Platform delivers control, virtualization, security, network services, and visibility across networks in the WAN, datacenter, and branch office/campus environments. The 128 Technology mission is to eliminate complexity and bring about a new realm of service agility and innovation for IP services.



The 128T Networking Platform is entirely software-based. It is comprised of two key building blocks: the 128T Control and the 128T Slice.

The 128T Control is a centralized operation resource that performs all centralized router and control functions. This includes computing and preparing routing tables, managing service policies, collecting analytics, and router management. Specific capabilities include:

- Centralized routing stack
- Centralized service and policy definition
- Distribution of the routing information base and service policy to each Slice
- Analytics engine and database
- Consolidated management and operational platform

The 128T Slice is a Software LIne Card Engine that performs all of the packet forwarding services utilizing the routing information provided by the 128T Control. Specific capabilities include:

- Low latency routing and packet forwarding

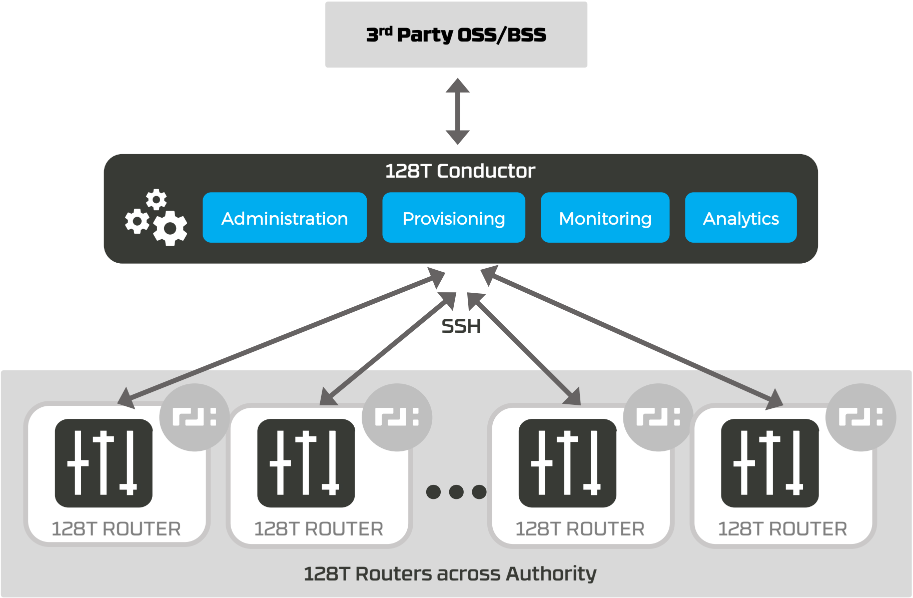
- Slices maintain a complete copy of the distributed routing information base
- Stateful session detection, classification, routing, and traffic management
- Application specific routing and QoS treatment
- Integrated load balancing
- Integrated ACLs, DoS protection, and session-based traffic shaping

The 128T Slices, taken together with the 128T Control, form a single distributed routing and service delivery system. Existing Ethernet or IP routed networks become network fabrics for interconnecting the Slices.

The 128T Control and 128T Slice can be deployed on any of the following platforms:

Category	Platform
Server	Bare Metal/Commercial off-the-shelf (INTEL based COTS)
Virtual	VMware, Kernel-based Virtual Machine (KVM)
Cloud	OpenStack, vCloud Director

The 128T Conductor is a platform to provide central administration, provisioning, monitoring, and analytics. The Conductor provides a single point of management of multiple, geographically-dispersed routers across an authority. It can also serve as the northbound interface to an OSS/BSS.



The 128T Networking Platform can be downloaded from the Yellowdog Updater, Modified, installation software (YUM) server. This requires a certificate for downloading that can be acquired by contacting your local account representative.

The 128T Networking Platform is available in 1-year and 3-year production software subscriptions. The subscription metric is the combination of project-wide bandwidth utilization and the number of sites where software is deployed. For a complete definition of subscription license entitlements, refer to the 128T Networking Platform License and Entitlement Document.

NOTE: 128 Technology's bandwidth utilization license defines an actual utilization rather than physical or allocated capacity. The "TYPICAL CAPACITY" is provided in the price book as a sizing guide and assumes a 25% utilization rate.

The following speeds may be used to calculate average bandwidth across all sites to arrive at the chose for utilization license:

Site Type	No. of Users	Connection Capacity	Utilization
Small Branch	0-50	10Mbps	1Mbps
Medium Branch	51-500	25Mbps	2.5Mbps
Large Branch	501-2500	50Mbps	5Mbps
Central Office	2500-10000	100Mbps	10Mbps
Small Data Center	10000-50000	150Mbps	15Mbps
Large Data Center	50000+	200Mbps	20Mbps

For an organization that has 2000 small branches and 2 central offices, the overall utilization is 2000 * 1 + 2 * 10 = 2020Mbps. Thus, the license required is **2.5Gbps**.

WHITE-BOX HARDWARE

The 128T Networking Platform can be implemented on any INTEL based Commercial off-the-shelf (COTS) platform when physical servers are required. Here are the recommendations for the different hardware to be used depending on the sites being served:

Site Type	No. of Users	Cores	Processor	Memory	Hard Disk
Small Branch	0-50	4	Intel Atom C2558	8GB	128GB
Medium Branch	51-500	8	Intel Atom C2758	16GB	128GB

Site Type	No. of Users	Cores	Processor	Memory	Hard Disk
Large Branch	501-2500	8	Intel Atom C2758	32GB	128GB
Central Office	2500-10000	12	Intel Xeon E5-2600 v4	128GB	500GB
Small Data Center	10000-50000	14	Intel Xeon E5-2600 v4	256GB	500GB
Large Data Center	50000+	22	Intel Xeon E5-2600 v4	256GB	500GB

The platforms must have Intel DPDK enabled NIC ports for LAN and WAN interfaces. In addition, two management ports are recommended. These should be chosen based on interface speed requirements. It is recommended to choose Intel Quick Assist enabled processors onboard or as daughter cards to get high encryption speed throughputs. Dual power, monitoring interface (with riser cards), LTE, Wi-Fi, and rack mountable options can be chosen as necessary.

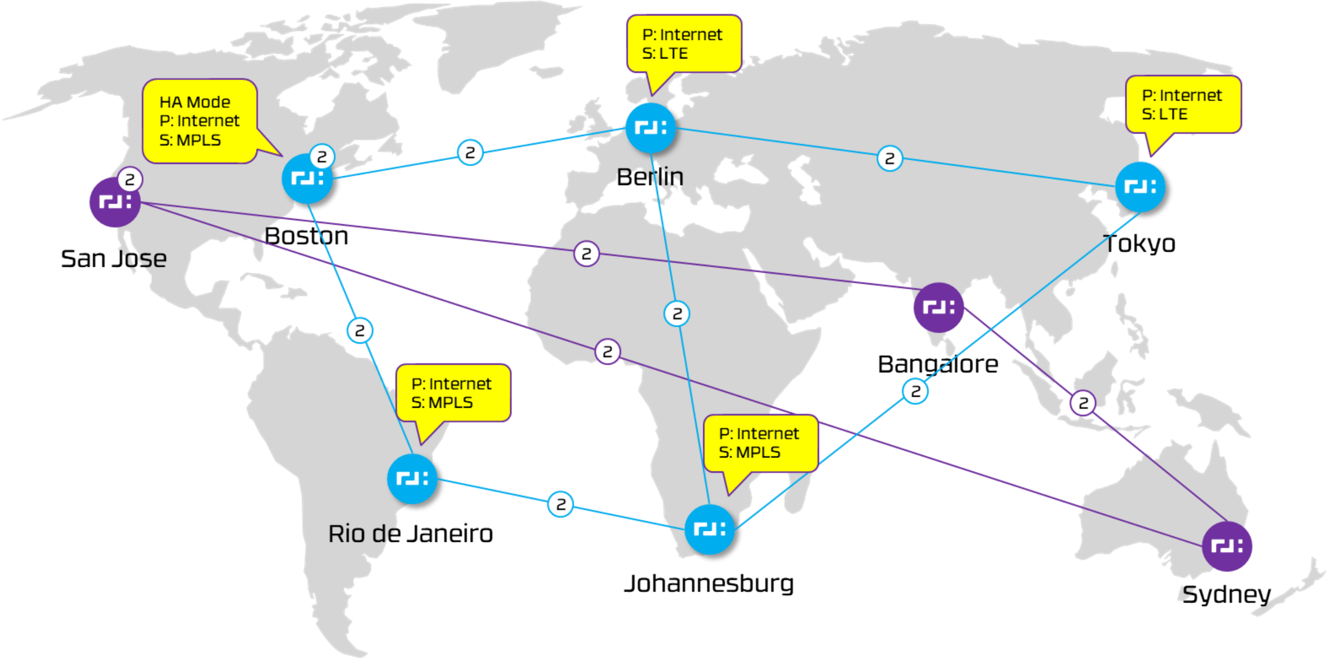
NOTE: 128 Technology has certified vendors from whom the hardware can be purchased. There is no restriction on using the certified vendor. The 128T Networking Platform should function on most hardware platforms having the recommended specifications.

BEST PRACTICES

The following section describes best practices and functionality of features on the 128T router that will provide the ability to deliver an uncompromised experience over a variety of WAN links.

Example: Sample Network

The following sample network will be utilized to demonstrate the concepts involved in the NG-WAN deployment. The blue routers represent a customer network belonging to a Bank. The purple routers represent another customer network belonging to a software development enterprise. All the sites have dual connections with a combination of MPLS, Internet, and 4G/LTE links. There are dual routers in Boston and San Jose for High Availability.



ROUTING

The 128T router supports explicit, static routing as well as dynamic routing via BGPv4.

The 128T router supports BGPv4 allowing it to be deployed as a peering router (iBGP and eBGP). It can also operate as a route reflector client (as is common in iBGP environments). BGPv4 is a venerable, mature protocol with many decades of features and enhancements. Many aspects of the 128T router's implementation of BGPv4 are tunable via a wide variety of configuration options exposed to the administrator. In addition to learning routes by way of BGP peering associations, administrators can configure static routes that are effectively inserted directly into the 128T router's Routing Information Base (RIB).

Example: Border Gateway Protocol

128T routers can BGP peer with traditional routers. As an example, consider a software company that has a datacenter in San Jose. The enterprise is hosting a Git repository at the datacenter. They have developers at an offshore office in Bangalore. The datacenter and the office have a Broadband link.

The BGP routing configuration is performed at the router level. The BGP configuration required at Bangalore 128T Router to peer with the 128T Router in San Jose CA is as follows:

```
router   office_site_bangalore_branch_128t_router
…
    routing                   default-instance
        type                  default-instance

        routing-protocol  bgp
            type                         bgp
            local-as                     1921722434
            …
            router-id                    192.172.24.34
            …

            address-family               ipv4-unicast
                afi-safi                ipv4-unicast
                …

                network                 10.2.21.0/24
                    network-address  10.2.21.0/24
                exit
            exit

            neighbor                     192.168.20.14
                neighbor-address  192.168.20.14
                neighbor-as       1921722430
                …

                transport
                    local-address  192.168.22.23
                exit

                multihop
                    ttl           64
                exit

                address-family    ipv4-unicast
                    afi-safi      ipv4-unicast
                    …
                exit
            exit
        exit
        static-route      0.0.0.0/0 100
            destination-prefix  0.0.0.0/0
            distance            100
            next-hop            192.168.22.1
        exit
    exit
exit
```

Similarly, the BGP configuration required at the 128T router in San Jose to peer with the 128T router in Bangalore would be as follows:

```
router   office_site_sanjose_dc_128t_ha_router
…
    routing               default-instance
        type              default-instance

        routing-protocol  bgp
            type                     bgp
            local-as                 1921722430
            …
            router-id                192.172.24.30
            …

            address-family          ipv4-unicast
                afi-safi            ipv4-unicast
                …

                network             10.2.1.0/24
                    network-address  10.2.1.0/24
                exit
            exit

            neighbor                 192.168.22.23
                neighbor-address  192.168.22.23
                neighbor-as       1921722434
                …

                transport
                    local-address  192.168.20.14
                exit

                multihop
                    ttl           64
                exit

                address-family    ipv4-unicast
                    afi-safi      ipv4-unicast
                    …
                exit
            exit
        exit
        static-route      0.0.0.0/0 100
            destination-prefix  0.0.0.0/0
            distance            100
            next-hop            192.168.20.1
        exit
    exit
exit
```

The two configuration snippets shown above demonstrate how an eBGP configuration between two 128T Routers, or with a traditional router is created. BGP neighbors are created and networks 10.2.1.0/24 and 10.2.21.0/24 are advertised respectively.

DHCP Client for WAN Links

128T routers support WAN link IP address learning through DHCP Client. This eliminates the need for manual configuration to deploy 128T routers and associated costs with purchasing static IP addresses. The 128T router can dynamically obtain IP addresses for WAN links.

Example: DHCP Client for WAN Links

Consider a Bank which is planning to bring in service to thousands of ATMs around the world. ATMs require access to Internet on a daily basis to function properly. Each ATM has dual transport with a Broadband link and a LTE link. To simplify, reduce costs, and improve the pace of deploying ATMs, DHCP client can be configured per network-interface as follows:

```
router        bank_site_berlin_branch_128t_router
…
    node                  bank_site_berlin_branch_128t_router
        …
        device-interface  1
            …
            type               ethernet
            pci-address        0000:00:09.0
            enabled            true

            network-interface  broadband
                name        broadband
                …
                dhcp        v4
            exit
        exit
    exit
exit
```

It is assumed that there is a *service* that would model open Internet which could look like:

```
service      internet
```

```
            name     internet
            …
            tenant   atm.sales.bank
            …
            address  0.0.0.0/0
            …
        exit
```

Lastly in order to use the DHCP interface for the *Internet* service, a *service-route* with an empty gateway needs to be created as follows:

```
            service-route     internet
                name          internet
                service-name  internet

                next-hop     bank_site_berlin_branch_128t_router broadband
                  node-name  bank_site_berlin_branch_128t_router
                  interface  broadband
            exit
```

Once the DHCP address is learned all traffic to Internet would be routed to the gateway learnt from DHCP via the DHCP interface, the network-interface called broadband.

FAILSAFE DELIVERY

Multiple paths often exist between peers in most large enterprise and service provider networks. These multiple paths can be used to reroute traffic in case of failures or link performance degradation. Multiple paths can be utilized simultaneously to improve user experience through improved resilience to network failures and higher throughput. The 128T solution combines Optimized Heuristics, Intelligent Path Monitoring, and Lossless Application Delivery to form a failsafe delivery model that ensures application traffic is delivered despite network failures. In addition, innovative quality of service and traffic engineering enhancements ensure optimal non-stop application performance and superior end-user experience with the 128T solution.

The 128 Technology solution combines three different key technologies related to delivery of application flows to ensure optimized traffic delivery in the face of all odds. These three technology areas include:

- Intelligent Path Monitoring
- Lossless Application Delivery
- Optimized Heuristics

Intelligent Path Monitoring

The 128T routers based on session-oriented networking paradigms connect endpoints to services. For effective delivery of traffic loads it is important to monitor network paths between 128T router instances.

Out-of-Band Monitoring

BFD is used to gauge the health of the distributed components as well as the integrity of the network connectivity between them. BFD [RFC5880] is traditionally used to ensure path connectivity between two routers' forwarding planes.

In addition to basic BFD, the 128T routers use an enhanced version of BFD to measure latency, packet loss, and jitter between themselves. This data is used to supply each 128T router with real-time link attributes, that can affect how traffic is delivered.

When a 128T router sends BFD packets to a peer, it includes in the BFD payload the quality point value configured for that link to its adjacency. This is retained by the receiving 128T router and used when making service routing decisions. Dynamic updates to these quality points can also be done to effect changes in traffic patterns.

BFD control packets are also enhanced with "BFD metadata". The metadata inserted by a 128T router can optionally be encrypted. The 128T implementation of BFD does not change the protocol's behavior, messages, or encoding; the only difference is the addition of 128T specific metadata when a 128T router transmits BFD to another 128T router.

Each 128T router uses both BFD's asynchronous mode and echo mode. Asynchronous mode is used for liveness checks and exchanging packet loss data. Echo mode is used for determining path latency and jitter.

Connectivity

Assuming BFD connections are established between two 128T routers, if a 128T router fails to receive a number of consecutive control packets from a counterpart, it treats that peer as unreachable. This has the effect of removing that 128T router as a potential target for new session assignments.

Quality

Unique to the 128T solution is the use of BFD for measuring link quality (latency, jitter, packet loss) between 128T routers. It is understood that packet queuing on either the transmitting or receiving 128T router can skew the results of the test; if BFD packets are treated with a higher priority than session traffic, the BFD test will produce results that are more favorable than the results that the session traffic will experience. 128T proposes a conservative approach and to treat BFD with a relatively low priority.

A 128T router will use echo mode BFD to periodically test each destination 128T router; the process for doing so is for a transmitting 128T router to send a series of BFD packets to the destination spaced evenly. As the packets return back, the round trip time is halved to estimate the one-way latency, and the variation in inter-arrival time of the return packets estimates jitter. The number of packets sent in each series and the number received is used to determine loss.

This enhanced version of BFD running between 128T routers helps in network path monitoring when there are no flows. The data gathered from network path monitoring is used as cost metric in the algorithms used for distributing server loads. This ensures that only those servers which are reachable via paths that meet the application demands are used.

Lossless Application Delivery

The 128T solution can utilize innovative server load monitoring and intelligent network path monitoring to ensure that the algorithms have the best possible heuristics available to choose the most appealing paths and servers.

Redundant or alternate paths between nodes in a network can be used to reroute traffic, improve resiliency, and maximize throughput. These maximally diverse paths can provide link and node protection for 100% of paths and failures as long as the failure does not cut the network into multiple pieces.

Example: Secure Vector Routing

Secure Vector Routing or SVR constitutes the main building block on which the 128T router bases many of its features on. SVR is the technology that two 128T Routers use to peer with one another. A 128T Router defines adjacencies pointing to other 128T routers' network interfaces, it also defines a peer configuration for each remote 128T router specifying whether or not traffic needs to be encrypted and/or authenticated, and one or more service-routes that describe which traffic is routed to each 128T router. One of the premises of SVR is its intelligent path monitoring and lossless application delivery. BFD is one of the fundamental pieces that SVR makes use of to achieve its goals.

Consider two 128T routers, one located in Johannesburg and the other in Berlin. The example shows how to create an SVR relationship between these two routers.

SVR is a routing technology by which 128T Routers can encrypt and authenticate all traffic routed among themselves. A *security* profile defines encryption and authentication algorithms, key length, etc. The *security* profile is defined at the 128T Authority level, which means that the *security* profile is shared among all 128T routers within the Authority they belong. The *security* profile can be created as follows:

```
        security        Peer-Johannesburg-Berlin
            name                 Peer-Johannesburg-Berlin
            description          "inter-router Johannesburg Berlin security"
            hmac-cipher          sha256
            hmac-key             4a656665
            encryption-cipher    aes-cbc-256
            encryption-key
603deb1015ca71be2b73aef0857d77811f352c073b6108d72d9810a30914dff4
            encryption-iv        f0f1f2f3f4f5f6f7f8f9fafbfcfdfeff
            encrypt              true
            hmac                 true
            …
        exit
```

This *security* profile "Peer-Johannesburg-Berlin" above is going to be used to encrypt and authenticate the metadata that is shared between the 128T routers in Johannesburg and Berlin using aes-cbc-256 and sha256 respectively.

Each 128T router needs to create a *peer* for each remote 128T router. Shown next is the *peer* in Johannesburg for the 128T router located in Berlin:

```
router            bank_site_johannesburg_branch_128t_router
…
    peer                       bank_site_berlin_branch_128t_router
        name                   bank_site_berlin_branch_128t_router
        authority-name         Authority128
        router-name            bank_site_berlin_branch_128t_router
        inter-router-security  Peer-Johannesburg-Berlin
    exit
exit
```

Similarly, it is shown next the *peer* to Johannesburg of the Berlin 128T router:

```
router            bank_site_berlin_branch_128t_router
…
    peer                       bank_site_johannesburg_branch_128t_router
        name                   bank_site_johannesburg_branch_128t_router
        authority-name         Authority128
        router-name            bank_site_johannesburg_branch_128t_router
        inter-router-security  Peer-Johannesburg-Berlin
    exit
exit
```

As it can be observed in the two snippets above, a *security* profile is referenced from the *peer* for inter-router security purposes.

The network interfaces and IP addresses that are going to be used for the SVR path have not been specified yet. Each 128T router should have an *adjacency* referring to the other 128T router's IP address where the SVR path terminates. This IP address is known as "waypoint":

```
router            bank_site_johannesburg_branch_128t_router
…
    node                  bank_site_johannesburg_branch_128t_router
        name              bank_site_johannesburg_branch_128t_router
        id                1
        …

        device-interface  1
            id                 1
            type               ethernet
            pci-address        0000:00:05.0
            enabled            true

            network-interface  mpls
                name        mpls
                …

                address     192.168.12.32
                    ip-address     192.168.12.32
                    prefix-length  24
                    gateway        192.168.12.1
                exit
                …

                adjacency   10.171.11.4
                    ip-address  10.171.11.4
                    peer        bank_site_berlin_branch_128t_router
                    cost        0
                    qp-value    0
                exit
                icmp        allow
            exit
        exit

        device-interface  2
            id                 2
            type               ethernet
            pci-address        0000:00:06.0
            enabled            true

            network-interface  broadband
                name        broadband
                …

                address     192.168.13.32
                    ip-address     192.168.13.32
                    prefix-length  24
                    gateway        192.168.13.1
                exit

                …

                adjacency   50.66.1.4
                    ip-address  50.66.1.4
                    peer        bank_site_berlin_branch_128t_router
                    cost        0
                    qp-value    0
                exit
                icmp        allow
            exit
        exit
    exit
```

Similarly, it is shown next the *adjacencies* to Johannesburg of the Berlin 128T router:

```
router            bank_site_berlin_branch_128t_router
…
    node                  bank_site_berlin_branch_128t_router
        name              bank_site_berlin_branch_128t_router
        …

        device-interface  2
```

```
            id                 2
            type               ethernet
            pci-address        0000:00:14.1
            enabled            true

            network-interface  broadband
                name        broadband
                …

                address     50.66.1.4
                    ip-address     50.66.1.4
                    prefix-length  24
                    gateway        50.66.1.2
                exit

                …

                adjacency   192.168.13.32
                    ip-address  192.168.13.32
                    peer        bank_site_johannesburg_branch_128t_router
                    cost        0
                    qp-value    0
                exit
                icmp        allow
            exit
        exit

        device-interface  3
            id                 3
            type               ethernet
            pci-address        0000:00:14.1
            enabled            true

            network-interface  mpls
                name        mpls
                …

                address     10.171.11.4
                    ip-address     10.171.11.4
                    prefix-length  24
                    gateway        10.171.11.2
                exit

                …

                adjacency   192.168.12.32
                    ip-address  192.168.12.32
                    peer        bank_site_johannesburg_branch_128t_router
                    cost        0
                    qp-value    0
                exit
```

```
                        icmp        allow
                    exit
                exit
            exit
```

Once one or more SVR path exists between two 128T Routers, then features such as Multi-Path Session Migration, Multi-Path Session Redundancy, Session Load Balancing, Session Security (Encryption, Authentication) can be enabled on a per service basis.

Multi-Path Session Migration

Multi-path session migration refers to the ability to migrate an existing session to an alternate path between two 128T routers. Multi-path session migration requires that packets can be forwarded not only on the shortest-path tree but on another maximally redundant path. This guarantees 100% recovery for single failures when the paths are completely disjoint.

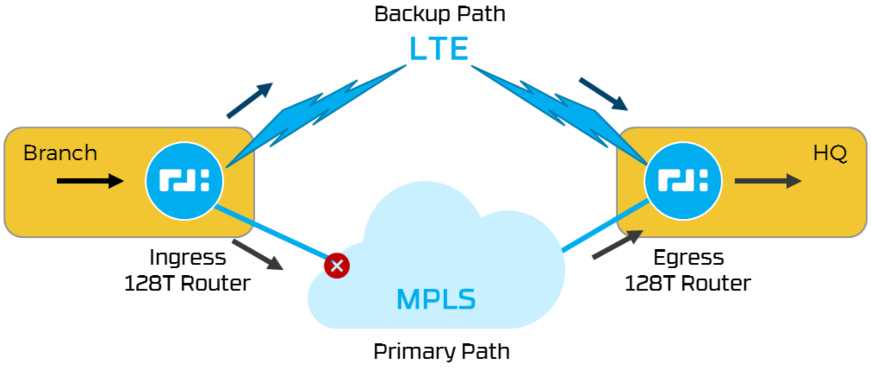
The administrator can configure multiple paths between two 128T routers. These paths can be redundant (completely disjoint) or maximally redundant (as disjoint as possible).

The administrator can configure the sessions traversing these routers to work in primary/backup mode or in load balancing mode. There may be sessions which the administrator may configure as not to take the alternate path.

The 128T router will switch traffic to the alternate path when it detects that the existing path has a failure or a link gradation that renders it unfit for use for the application.

LTE/Wired Connections

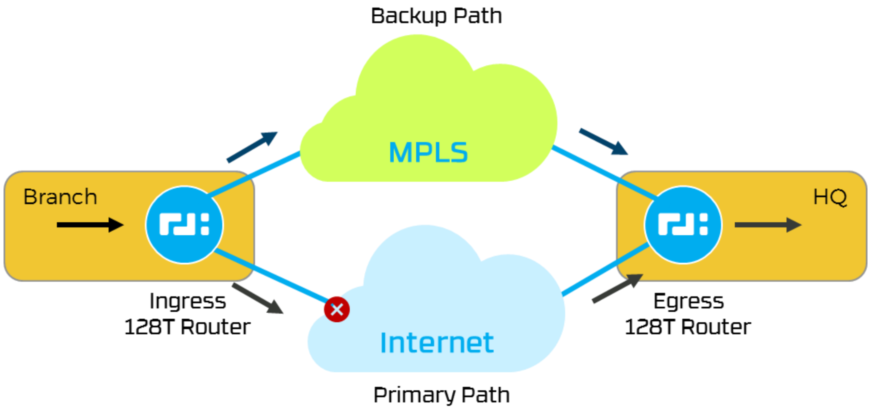
Service providers or large enterprises can provide their subscribers or branch offices with access to fixed and mobile networks. It has become desirable to use these heterogeneous networks as backups in case of failures.



In most cases the fixed wired connection is used as the primary. LTE (Long Term Evolution) or 3G connection is used as the backup. The traffic always flows over the wired connection.

In case of failure the sessions are moved to the wireless connection. When the wired connection is restored the sessions are moved back to the wired connection.

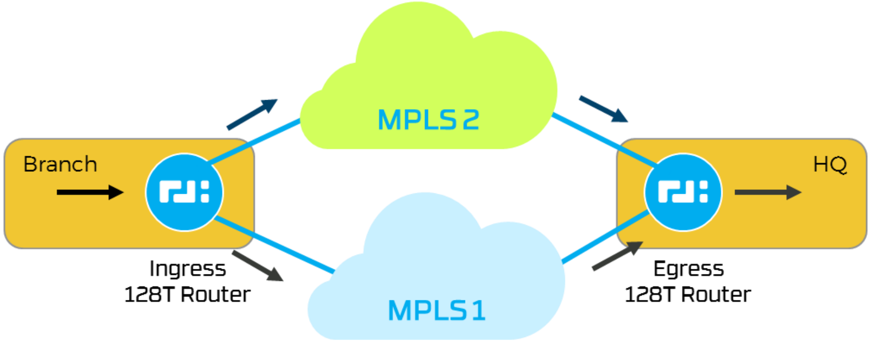
Internet/MPLS Connections



Service providers or large enterprises can provide their subscribers or branch offices with access to connections via dedicated MPLS circuits and Internet connection. It has become desirable to use the Internet connection to save costs while the MPLS circuits act as backup if needed.

The traffic always flows over the Internet connection. In case of performance degradation, the sessions are moved to the MPLS connection. When the Internet connection is restored to acceptable performance, the sessions are moved back to the Internet connection.

Wired/Wired Connections



Service providers or large enterprises can provide their subscribers or branch offices with access to two different fixed connections for diversity. It has become desirable to use both these networks and they act as backups of each other in case of failures.

It is also possible to have two paths within the same network. In this case, it is desirable that both paths be used in load balancing mode.

Example: Multi-Path Session Migration

Multi-Path Session Migration functionality is enabled on a per service level, so that each service can be configured with the session migration policy that better fits its needs. Consider a software company that has a datacenter in San Jose and a branch in Sydney. The enterprise is hosting a Git repository at the datacenter for its customers located at each branch. The datacenter has a Broadband link, while the branch in this example has dual transport: a Broadband link and a LTE link. The desired Session Migration behavior for their hosted Git service is at each Branch to always prefer the Broadband link over the LTE link and only failover to the LTE link in cause of:

- Broadband link failure: the network interface operationally going down due to physical conditions of the Ethernet link.
- BFD determines that there is no suitable path to the remote 128T Router at the datacenter from the Broadband interface. Some reasons for such decision are: no path meets the expected SLA defined in its *service-policy* (packet loss, latency, jitter), the configurable amount of BFD control packets in *bfd* have not received response.

Multi-Path Session Migration due to connectivity loss

The session migration policy for a service is defined in a *service-policy* as follows:

```
service-policy  sw_repo
            name                         sw_repo
            qp-preference                highest
            session-resiliency           revertible-failover
exit
```

There are two policies for Multi-Path Session Migration as listed in the parameter *session-resiliency*:

- Failover: in case the preferred link is not suitable, migrate all new and existing sessions of the service to the next available most preferred link.
- Revertible-failover: same as the failover policy but once the preferred link is restored, then all new and existing sessions are migrated back to it.

The parameter *qp-preference* determines what is the static criteria to select the most preferred link of a 128T router. Quality points (*qp-value* parameter of a *network-interface*) are assigned to each link of the 128T according to our preference given by its bandwidth, reliability, etc. In our example, the administrator of the 128T Router at the branch decided to assign the highest *qp-value* to the most preferred link, and as such the *qp-preference* has been set to highest as shown above. This information is distributed and therefore known by all 128T Routers that are part of the same Authority.

The Quality Points Values are configured in *qp-value* under either the *network-interface* or the *adjacency* of the ingress 128T Router making the decision instead, in our example, the 128T Router located at the branch:

```
device-interface  2
     …
     network-interface  broadband
         name        broadband
         qp-value    10
         address     192.168.21.22
             …
         exit

         adjacency   192.168.19.12
             ip-address  192.168.19.12
             peer        office_site_sanjose_dc_128t_ha_router
             qp-value
             …
         exit
         …
     exit
 exit

 device-interface  3
     …
     network-interface  lte
         name        lte
         …
```

```
                qp-value    5
                address     192.168.22.22
                    …
                exit

                adjacency   192.168.20.12
                    ip-address  192.168.20.12
                    peer        office_site_sanjose_dc_128t_ha_router
                    qp-value
                    …
                exit
                …
            exit
        exit
```

In the example above *qp-values* are configured at the *network-interface* level as opposed to at the *adjacency* level. The reason is that per the example, the policy is always to prefer the Broadband link over the LTE link and such decision it is not dependent of which far end remote 128T Router destination the 128T Router in Sydney is going to send traffic to.

Lastly the *service-policy* needs to be associated with the *service* that models the Git service:

```
service      sw_repo
    name     sw_repo
    …
    service-policy  sw_repo
```

Multi-Path Session Migration is performed among two 128T routers via SVR as mentioned in the previous section. In this example, it is assumed that in the 128T router at the branch, *adjacencies*, a *peer*, and a *service-route* pointing to the 128T router at the datacenter exist. The 128T router configuration of the Sydney branch is shown above, which has an adjacency pointing to the waypoint IP addresses of the San Jose 128T router for both network interfaces, the Broadband and LTE interface. The "office_site_sanjose_dc_128t_ha_router" peer along with the *service-route* are shown next:

```
peer                       office_site_sanjose_dc_128t_ha_router
    name                   office_site_sanjose_dc_128t_ha_router
    authority-name         Authority128
    router-name            office_site_sanjose_dc_128t_ha_router
    inter-router-security  peer_site_sanjose_dc_128t_ha_router
```

And the *service-route* pointing to the peer hosting the Git service:

```
service-route             sw_repo
     name          sw_repo
     service-name  sw_repo
     peer          office_site_sanjose_dc_128t_ha_router
 exit
```

With the configuration shown above, sessions corresponding to the Git service originated from the branch will make use of the Broadband link first, and only use the LTE link in case of:

- Broadband link failure: the network interface operationally going down due to physical conditions of the Ethernet link.
- The configurable amount of BFD control packets (in *bfd*) send from the Broadband link have not received response, effectively determining that connectivity from the Broadband link to the 128T Router at the datacenter has been lost.

Existing and new Git sessions will be migrated back to the most preferred link as soon as both failed conditions above are no longer true.

Multi-Path Session Migration due to quality degradation

As listed above in "Multi-Path Session Migration due to connectivity loss", the migration of sessions of the Git service are triggered due to physical/operational status of the Broadband link, or when the 128T Router at the branch, via BFD, determines that it cannot reach the 128T Router at the datacenter from its Broadband link. However, the migration of the Git service will not additionally be triggered by a quality degradation/SLA violation of the Broadband link. The configuration for migrating sessions across different links due to link quality is shown next:

```
service-policy  sw_repo
    name                         sw_repo
    …
    path-quality-filter          true
    max-loss                     0.5
    max-latency                  250
    max-jitter                   100
    …
exit
```

After enabling *path-quality-filter* parameter every 128T router within the Authority, including the 128T router at the branch, will start measuring the quality of all its links via BFD. Those links which do not meet any of the requirements specified in the configured SLA above: a packet loss above 0.5%, or with a round-trip latency above 250ms, or with a max jitter of 100ms, will be filtered out and considered not suitable for use, so that new and existing sessions of the Git service will migrate to the next most preferable link which does.

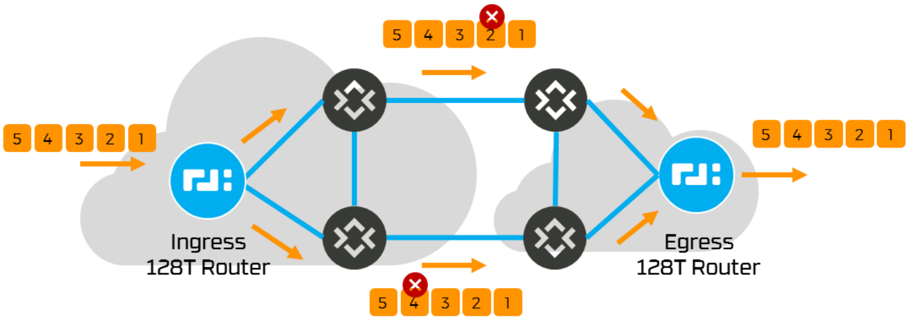
It is possible to configure Multi-Path Session Migration which combines both "Multi-Path Session Migration due to connectivity loss" and "Multi-Path Session Migration due to quality degradation" as explained in this document. In this case all the traffic which corresponds to the Git service will make use of the Broadband link first, and only use the LTE link in case of:

- Broadband link failure: the network interface operationally going down due to physical conditions of the Ethernet link.
- The configurable amount of BFD control packets (in *bfd*) send from the Broadband link have not received response, effectively determining that connectivity from the Broadband link to the 128T Router at the datacenter has been lost.
- No path originated from the Broadband link meets the agreed SLA defined in the *service-policy* (packet loss, latency, jitter) of the Git service, effectively determining that the quality of the Broadband link is degraded to steer sessions of the Git service.

Existing and new Git sessions will be migrated back to the most preferred link as soon as the three failed conditions above have been solved.

Multi-Path Session Redundancy

Packet loss is not avoidable in large networks. This loss might be due to congestion; it might also be a result of an unplanned outage caused by a flapping link, a link or interface failure, a software bug, or a maintenance person accidentally cutting the wrong fiber. Since UDP/IP flows do not provide any means for detecting loss and retransmitting packets, it is left up to the higher layer and the applications to detect, and recover from, packet loss.



Existing loss mitigation techniques such as retransmission, forward error correction (FEC) – both media independent and media specific FEC, and interleaving have proven to be successful in limited scenarios only. One technique to recover from packet loss without incurring unbounded delay is to duplicate the packers and send them in separate redundant streams. The probability that two copies of the same packet are lost in quite small. This scheme has comparatively high overhead in terms of bandwidth as everything is sent twice. The recommendation will be to use this scheme for high-value traffic. For example, a service provider may use this scheme to provide a superior user experience for telepresence traffic for VIP customers.

The administrator can configure dual paths for a high-value session. The 128T ingress router will duplicate the packets arriving from the session and send it over the dual paths. Both streams carry the same payload with identical sequence numbers. This allows the 128T egress router to identify and suppress the duplicate packets, and subsequently produce a loss-free and duplicate-free output stream.

This reduces the delay when packet loss occurs. An unrecoverable loss happens only when two network failures happen in such a way that the same packet is affected on both paths.

This technique requires the forwarding delay of the network paths to be more or less the same to ensure that the removal of duplicates and the application succeed. The 128T egress router monitors the delays over the dual paths and report whether they are acceptable for this scheme to function. If the delays are above the acceptable limits, the 128T ingress router will stop duplicating the stream after a wait period until the delays are below the acceptable limit before resuming the duplicate the stream.

This technique ensures lossless transport of traffic. These techniques ensure the best use of network resources and the lossless delivery of application traffic for applications.

Example: Multi-Path Session Redundancy

Multi-Path Session Redundancy functionality is enabled on a per service level. Consider a Bank which has a datacenter in Boston MA and a branch in Johannesburg. The datacenter hosts a VoIP UC/PBX server to offer communications to all of its branches. VoIP Unified Communications phones are located at each branch, and these require connectivity to the UC/PBX server located at the datacenter to operate. The datacenter has two links: a MPLS link and a Broadband link. Let's assume that each branch has dual transport as well: a MPLS link and a Broadband link. The Bank needs a high performance and reliable communication system to operate quickly and efficiently. The VoIP UC service is of high-value. Given such requirements, modeling the Unified Communications as a service and enabling Multi-Path Session Redundancy is recommended.

Session redundancy policy for a service is defined in a *service-policy* such as:

```
service-policy  unified_communications
            name                         unified_communications
            session-resiliency           packet-duplication
            path-quality-filter          false
exit
```

The *session-resiliency* policy for Multi-Path Session Duplication is packet-duplication as shown above. It is also important to note that *path-quality-filter* parameter needs to be set to false.

Lastly the *service-policy* needs to be associated with the *service* that models the Git service:

```
        service      uc_pbx
            name     uc_pbx
            …
            service-policy  unified_communications
```

Multi-Path Session Duplication is performed among two 128T routers via SVR as mentioned previously. In this example, it is assumed that in the 128T router at the Johannesburg branch, *adjacencies*, a *peer*, and a *service-route* pointing to the 128T router at the datacenter exist. The 128T router in the branch in Sydney must have adjacencies pointing to the waypoint IP addresses of the Boston 128T router within at least two network interfaces. Each network interface that contains and adjacency towards Boston 128T router would be used for packet duplication. For a configuration example of *adjacency*, *peer* and *service-route* please refer to previous section "Example: Multi-Path Session Migration".

With the configuration shown above all packets corresponding to VoIP sessions originated from the branch will be duplicated across both links: the MPLS link and the Broadband link of the branch.

Optimized Heuristics

The 128 Technology solution operates on the notion of sessions which are targeted to deliver traffic to service agents representing application servers. These servers are assigned loads and quality points.

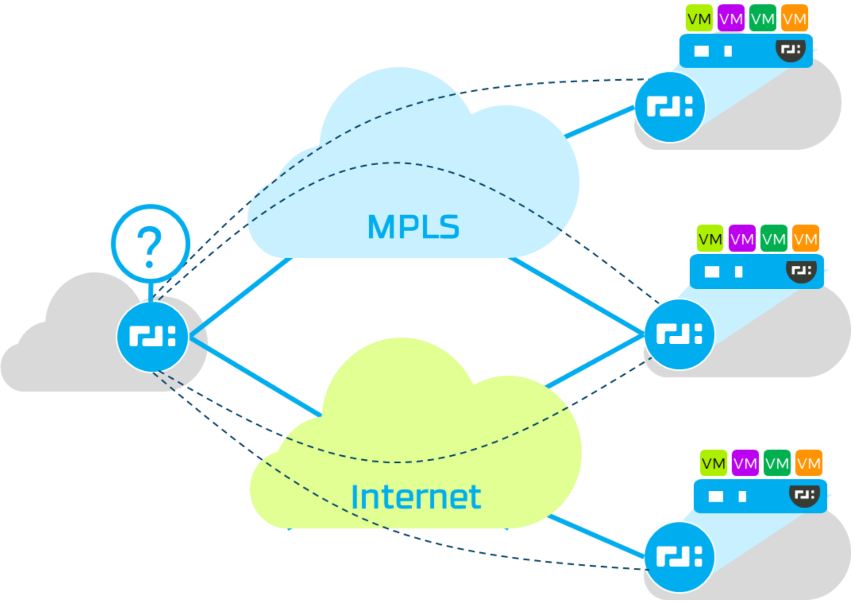
Server Loads

Server loads define the ratio of load that a particular server can take compared to other servers for the same application.

The load can specify the maximum number of sessions that a server can support or the maximum bandwidth of the traffic that can be sent to the server.

The load balancing algorithm at the 128T router will take this into account while distributing application flows across the servers.

For example, if there are three servers which can serve a particular application and the loads assigned to them are 100, 50, and 50 then a proportional distribution algorithm would spread the application flows in the ratio 2:1:1 among them in the absence of other factors. The 128T routers monitor the loads in use to a particular server and these are used for spreading loads to ensure that the servers are not overwhelmed.



Quality Points

Quality points are administrator assigned cost metrics to a 128T router adjacency that indicates path preference useful for ensuring that traffic flows to locations using links that meet service level agreement thresholds. For example, if two server locations that are reachable via quality points 200 and 100 then the one with quality point 200 will be preferred over quality point 100 in the absence of other factors.

In addition to quality points, other real-time cost criteria such as maximum session rate, packet loss, latency, and jitter can also be used to determine if the path to a particular server is suitable for use by the application requirements. An administrator can ensure that traffic flows to the preferred server over the preferred link by assigning appropriate server loads and quality points. These metrics will change dynamically during the course of operation depending on traffic flows leading to server loads being distributed over different WAN links.

Strategies

The 128T router can distribute traffic based on different criteria in two different strategies. These algorithms attempt to maximize the loads to agents based on available capacities while minimizing the cost of the paths to the selected agents or servers.

PROPORTIONAL

The Proportional algorithm distributes sessions to agents weighted on the relative available capacity. For example, if an agent has double the amount of available capacity as another, it will receive twice the number of sessions. This is good for distributing loads over all paths that meet the application SLA and there is no restriction on the path being used. For example, sending all traffic to multiple servers over an MPLS network.

HUNT

The Hunt algorithm routes all sessions sequentially to the agent with the highest capacity until a load threshold is reached for the agent, then switch and repeat to the agent with the next highest capacity. For agents with the same capacity, the agent's name is used as a tiebreaker, lexicographically lowest first. This is good for maximizing usage of a lower cost link say Internet over using an MPLS link. The higher cost link will only be used when the lower cost link usage has reached a threshold. This is preferred for the SD-WAN case to use up lower cost links before higher cost links.

These mechanisms ensure that the best possible or preferred server is chosen for application traffic delivery.

Example: Session Load Balancing

Session Load Balancing functionality is enabled on a per service level. Consider a bank in Germany which has many ATMs deployed across Berlin. For accounting and compliance purposes, the bank has datacenter in Boston and a disaster recovery backup site in Tokyo. The datacenter as well as the disaster recovery site hosts a farm of servers that receive, process and store all accounting records sent from each ATM. The datacenter has two links: a MPLS link and a Broadband link. Each ATM has dual transport as well: a Broadband link and a LTE link. Given the number of ATMs deployed currently and many more expected to be deployed across the world, the bank needs a high performance and reliable network that spreads the load across all transaction servers hosted in Boston and Tokyo. Under such circumstances, Session Load Balancing is the right network functionality that this sales accounting service demands.

Session load balancing for a service is enabled in the *service-policy* associated with such service:

```
service-policy  accounting
    name                         accounting
    lb-strategy                  proportional
    …
exit
```

There are currently two strategies for session load balancing:

- Hunt: given a pool of possible destinations to route the traffic to, sessions are initially routed to one destination only. Only when the first target destination becomes unresponsive or its capacity has been reached, then new sessions are routed to the next available destination.
- Proportional: sessions are distributed and balanced and therefore routed across a pool of target destinations, according to the capacity each can process.

The service-policy is then associated with the service "atm_accounting" as shown below:

```
service      atm_accounting
    name     atm_accounting
    service-policy accounting
exit
```

Once the *service* and its desired *service-policy* has been associated, the next step consists of defining the capacity constraints of each destination target. Each destination is represented using a *service-route*, and the constraints of each destination are specified via a *service-route-policy* as it will be shown shortly. From the perspective of the 128T Router in Berlin, there are two different destinations or available routes, for the *atm_accounting* service: one route to the datacenter in Boston MA and another route to the backup datacenter in Tokyo. It is assumed that the 128T Router in Berlin has already adjacencies and is peering with each 128T Router located in Boston and Tokyo. Additionally, the 128T Router in Berlin must define the following configuration in order to define the constraints of each route:

```
router bank_site_berlin_branch_128t_router
    …
    service-route     atm_accounting_servers_boston
        name          atm_accounting_servers_boston
        service-name  atm_accounting
        peer          bank_site_boston_dc_128t_ha_router
        service-route-policy  atm_accounting_servers_boston
    exit

    service-route     atm_accounting_servers_tokyo
        name          atm_accounting_servers_tokyo
        service-name  atm_accounting
        peer          bank_site_tokyo_branch_128t_router
        service-route-policy atm_accounting_servers_tokyo
    exit

    service-route-policy      atm_accounting_servers_boston
        name                     atm_accounting_servers_boston
        max-sessions             50
        session-high-water-mark  95
        session-low-water-mark   90
    exit

    service-route-policy      atm_accounting_servers_tokyo
        name                     atm_accounting_servers_tokyo
        max-sessions             25
        session-high-water-mark  95
        session-low-water-mark   90
    exit
exit
```

The policy of the 128T Router in Berlin listed above describes that there are two SVR routes for the "atm_accounting" service:

- A SVR route called "atm_accounting_servers_boston" to route all traffic from the ATMs to the Boston datacenter (through the peer 128T Router in Boston), which has a total capacity of 50 sessions as denoted by the associated *service-route-policy* "atm_accounting_servers_boston".
- A SVR route called "atm_accounting_servers_tokyo" to route all traffic from the ATMs to the Tokyo datacenter, which has a total capacity of 25 sessions as described by the associated *service-route-policy* "atm_accounting_servers_tokyo".

With the three snippets of configuration above, the 128T Router in Berlin is going to distribute all traffic from the ATMs in Berlin to Boston and Tokyo datacenter locations in a distributed fashion. This is indeed a use case of session load balancing among 128T Routers at different locations. However, a scenario where session load balancing is performed across a pool servers instead is supported the same way as discussed next.

Considering that the datacenter in Boston has a pool of two accounting servers to process transactions, the configuration that the 128T Router located in Boston needs to perform proportional session load balancing across both the servers is the following:

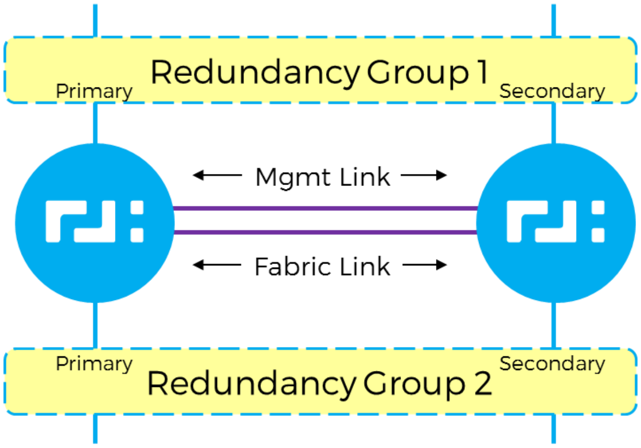
```
router bank_site_boston_dc_128t_ha_router
    …
    service-route     atm_accounting_server_1
        name          atm_accounting_server_1
        service-name  transaction_servers
        destination   10.1.13.10
        service-route-policy  atm_accounting_servers
    exit

    service-route     atm_accounting_server_2
        name          atm_accounting_server_2
        service-name  transaction_servers
        destination   10.1.13.11
        service-route-policy  atm_accounting_servers
    exit

    service-route-policy      atm_accounting_servers
        name                     atm_accounting_servers
        max-sessions             25
        session-high-water-mark  95
        session-low-water-mark   90
    exit
exit
```

Each accounting server above has been limited to a capacity of 25 maximum sessions per server, which accounts for a total aggregate of 50 sessions the Boston datacenter can handle simultaneously.

HIGH AVAILABILITY

The 128T resiliency solution offers the industry's most comprehensive, advanced, and innovative high availability and disaster recovery solution along with stateful failover. This guarantees virtually zero downtime, multi-site failover, and a scale out architecture with N+M redundancy. This is a must-have for organizations to maintain connectivity during planned and unplanned downtime. It prevents revenue loss, improves productivity, reduces security risks, improves customer satisfaction, and guarantees regulatory compliance.

The solution operates in Active/Active clustering mode. Multiple routers are grouped together as clusters, with multiple Active units processing traffic and sharing the network load. Each cluster node contains a minimum of two units acting as a Stateful HA pair. Active/Active clustering provides Stateful failover in addition to load sharing. The customer may choose to pass all traffic through one of the routers in the cluster. In this case the remaining routers in the cluster will not be processing traffic but they are all in Active mode with ability to process traffic if required.



Interfaces on different routers can be configured as redundancy groups. These redundancy groups are collection of resources that need to failover between the routers. An interface in a redundancy group is chosen as the primary and another as the secondary. This is done via a leader election or based on user defined priorities. Primary interfaces are used to route traffic through the cluster. In case of failure the traffic from the primary interface is switched to the secondary interface in the redundancy group via Gratuitous Address Resolution Protocol (GARP) or other routing protocol exchange.

A fabric link between the routers is used to route traffic between them in case of failure. In the diagram these are shown as directly connected links but they do not have to be. Also the diagram shows two routers in a cluster for ease of understanding however there can be multiple routers in the cluster.

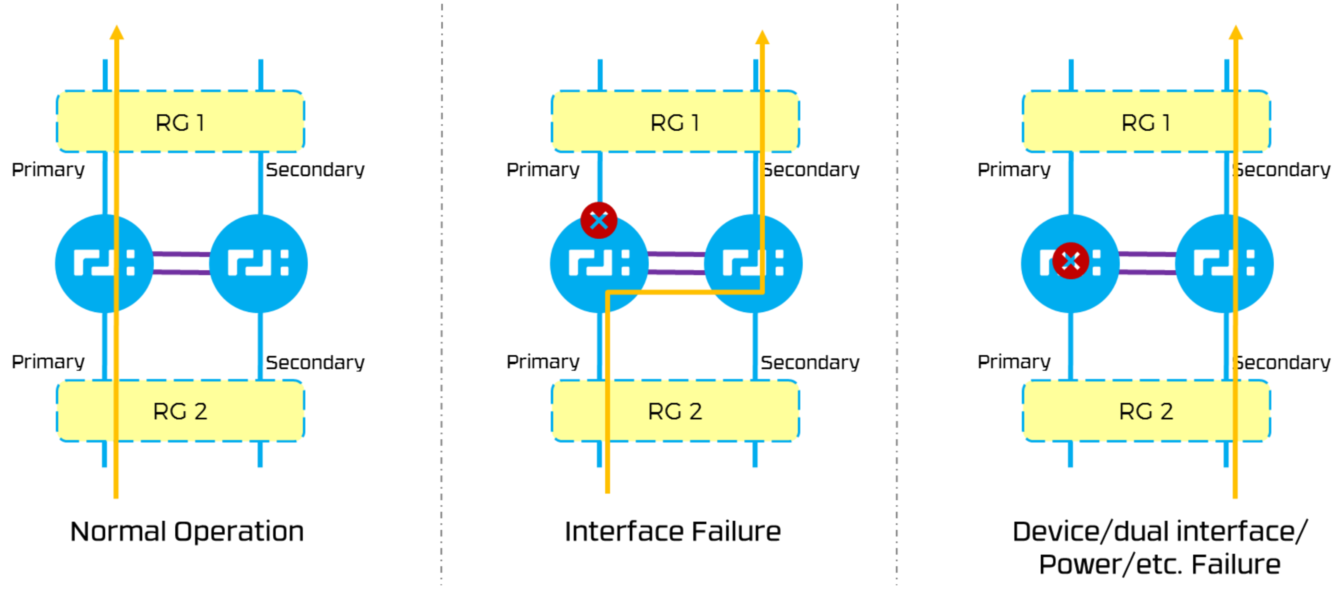
The management link between the routers is used to exchange routing and flow information between the routers. This is shown as a separate directly connected link between the routers in the diagram however it can share any link. All information between the routers are shared using highly efficient in-memory databases to minimize bandwidth usage and to enable instantaneous information exchange.

All processes in a 128T router are self-resilient. They can regenerate themselves independently in case of process failures or exceptions. Unless there is a dependency that requires other processes to restart or the device to switchover, the process will rebuild itself and establish communications to the existing processes. If a process failure requires another process to be restarted due to a dependency, then that process is restarted automatically. In-built self-checking mechanisms managed with software diagnostics ensure the integrity of the entire system.

Failure Scenarios

The 128T resiliency solution can enable zero downtime failure protection for all types of planned and unplanned network outages. It ensures that the end user application is oblivious to any network outage. The solution can provide high availability protection and switchover in case of process, interface, component, device, and cluster failovers. These may be caused by power outages, human errors, or other factors.

In case of normal operation all traffic is forwarded through the primary interfaces of the redundancy groups. These may be on a single router or may be spread across the router resulting in different flows to go through either router.



In case of a single primary interface failure, traffic will be routed through the fabric link to the secondary interface on the other router. This fabric link does not have to be a dedicated directly connected link. In case of process, device, component, or dual interface failures that completely disables the router passing traffic then the traffic switches over completely to the other router.

In Service Software Upgrades

The 128T resiliency solution enables the ability to support in-service software upgrades. This reduces downtime due to planned upgrades. The 128T system provides information to switch traffic to different routers in a cluster. Any router within the cluster can be isolated to prevent traffic forwarding through it while other routers in the cluster continue to forward all the traffic. This isolated router can then be upgraded. The upgraded router can then continue to participate in the cluster. Traffic can be switched back to this router or it can remain dormant in the cluster while fully Active to process traffic if required.

Example: High Availability

It is always recommended to deploy 128T as a cluster of a minimum of two nodes for high redundancy. Network interfaces of each node are grouped into redundant groups as shown in the following example, a 128T Router HA made of two 128T combo nodes. All network interfaces of each node are within the same redundant group, which effectively triggers a failover when any of the network interfaces of the active node go into a failed state. It is also worth mentioning that a critical system failure of the active node will result in a failover as well as denoted by the picture above.

Let’s consider the case of a cloud provider which has a datacenter in San Jose which has dual transport connections: MPLS link and a Broadband link. Additionally, there are three internal LAN networks as well. The HA configuration required to make a 128T Router highly redundant is as follows:

```
router                bank_site_boston_dc_128t_ha_router
…
    node                  bank_site_boston_dc_128t_ha_router_ha_1
        name              bank_site_boston_dc_128t_ha_router_ha_1
        id                1
        enabled           true
        role              combo

        device-interface  0
            id                 0
            …
            shared-phys-address fa:16:3e:12:07:85

            network-interface  lan1
                name        lan1
                global-id   1
                …

                address     10.1.11.2
                    ip-address     10.1.11.2
                    prefix-length  24
                exit
                …
            exit
        exit

        device-interface  1
            id                 1
            …
            shared-phys-address fa:16:3e:11:5a:d4

            network-interface  lan2
                name        lan2
                global-id   2
                …

                address     10.1.12.2
                    ip-address     10.1.12.2
                    prefix-length  24
                exit
                …
            exit
        exit

        device-interface  2
            id                 2
            …
            shared-phys-address fa:16:3e:5d:ea:15

            network-interface  lan3
                name        lan3
                global-id   3
                …

                address     10.1.13.2
                    ip-address     10.1.13.2
                    prefix-length  24
                exit
                …
            exit
        exit

        device-interface  3
            id                 3
            …
            shared-phys-address fa:16:3e:60:c5:bd

            network-interface  mpls
                name        mpls
                global-id   4
                …

                address     192.168.10.12
                    ip-address     192.168.10.12
                    prefix-length  24
                    gateway        192.168.10.1
                exit
                …
            exit
        exit

        device-interface  4
            id                 4
            …
            shared-phys-address fa:16:3e:f6:7d:5d

            network-interface  broadband
                name        broadband
                global-id   5
                …

                address     192.168.11.12
                    ip-address     192.168.11.12
                    prefix-length  24
                    gateway        192.168.11.1
                exit
                …
            exit
        exit

    exit

    node                 bank_site_boston_dc_128t_ha_router_ha_2
        name             bank_site_boston_dc_128t_ha_router_ha_2
        id               2
        enabled          true
        role             combo

        device-interface  0
            id                 0
            …
            shared-phys-address fa:16:3e:12:07:85

            network-interface  lan1
                name        lan1
                global-id   1
                …

                address     10.1.11.2
                    ip-address     10.1.11.2
                    prefix-length  24
                exit
                …
            exit
        exit

        device-interface  1
            id                 1
            …
            shared-phys-address fa:16:3e:11:5a:d4

            network-interface  lan2
                name        lan2
                global-id   2
                …

                address     10.1.12.2
                    ip-address     10.1.12.2
                    prefix-length  24
                exit
                …
            exit
        exit

        device-interface  2
            id                 2
            …
            shared-phys-address fa:16:3e:5d:ea:15
```

```
    network-interface  lan3
        name        lan3
        global-id   3
        …

        address     10.1.13.2
            ip-address     10.1.13.2
            prefix-length  24
        exit
        …
    exit
exit

device-interface  3
    id                 3
    …
    shared-phys-address fa:16:3e:60:c5:bd

    network-interface  mpls
        name        mpls
        global-id   4
        …

        address     192.168.10.12
            ip-address     192.168.10.12
            prefix-length  24
            gateway        192.168.10.1
        exit
        …
    exit
exit

device-interface  4
    id                 4
    …
    shared-phys-address fa:16:3e:f6:7d:5d

    network-interface  broadband
        name        broadband
        global-id   5
        …

        address     192.168.11.12
            ip-address     192.168.11.12
            prefix-length  24
            gateway        192.168.11.1
        exit
        …
    exit
exit
```

```
            exit

            redundancy-group          group-a
                name      group-a

                member    bank_site_boston_dc_128t_ha_router_ha_1 0
                    node       bank_site_boston_dc_128t_ha_router_ha_1
                    device-id  0
                exit

                member    bank_site_boston_dc_128t_ha_router_ha_1 1
                    node       bank_site_boston_dc_128t_ha_router_ha_1
                    device-id  1
                exit

                member    bank_site_boston_dc_128t_ha_router_ha_1 2
                    node       bank_site_boston_dc_128t_ha_router_ha_1
                    device-id  2
                exit

                member    bank_site_boston_dc_128t_ha_router_ha_1 3
                    node       bank_site_boston_dc_128t_ha_router_ha_1
                    device-id  3
                exit

                member    bank_site_boston_dc_128t_ha_router_ha_1 4
                    node       bank_site_boston_dc_128t_ha_router_ha_1
                    device-id  4
                exit
                priority  50
            exit

            redundancy-group          group-b
                name      group-b

                member    bank_site_boston_dc_128t_ha_router_ha_2 0
                    node       bank_site_boston_dc_128t_ha_router_ha_2
                    device-id  0
                exit

                member    bank_site_boston_dc_128t_ha_router_ha_2 1
                    node       bank_site_boston_dc_128t_ha_router_ha_2
                    device-id  1
                exit

                member    bank_site_boston_dc_128t_ha_router_ha_2 2
                    node       bank_site_boston_dc_128t_ha_router_ha_2
                    device-id  2
                exit

                member    bank_site_boston_dc_128t_ha_router_ha_2 3
```

```
                    node       bank_site_boston_dc_128t_ha_router_ha_2
                    device-id  3
                exit

                member    bank_site_boston_dc_128t_ha_router_ha_2 4
                    node       bank_site_boston_dc_128t_ha_router_ha_2
                    device-id  4
                exit
                priority  20
            exit
        exit
```

The configuration shown above corresponds to a 128T HA Router called “bank_site_boston_dc_128t_ha_router” which is made up of two 128T combo nodes respectively “bank_site_boston_dc_128t_ha_router_ha_1” and “bank_site_boston_dc_128t_ha_router_ha_2”. It is important to observe the following facts in the configuration:

- Each network-interface of “bank_site_boston_dc_128t_ha_router_ha_1”, the first 128T combo node, has a unique *global-id* (even across diferent *device-interfaces*).
- Each network-interface has a new parameter called *shared-phys-interface* which consist of a physical MAC address. It is the virtual MAC address of the network-interface that is advertised by the 128T combo node that is active for the redundancy group the network interface belongs to.
- Each *device-interface* and *network-interface* of the second 128T combo node, “bank_site_boston_dc_128t_ha_router_ha_2”, is exactly the same as the first 128T combo node, including: IP addresses, netmask, gateways, global-id, etc… even those parameters and values omitted in this example need to be the identical.

Lastly, the two *redundant-group* need to be configured. In the snippet above, all network interfaces are part of the same redundancy-group, resulting in a HA cluster where only node is active while the other stays standby. How a 128T combo node is considered active or standby for a given redundancy-group is by the redundancy-group parameter *priority*. The redundancy-group “group-a”, which groups all network interfaces of the first 128T combo node and has a priority of 50 is therefore selected and the combo node for this redundancy group.

SECURITY

Per-session encryption and per packet authentication is supported between all 128T routers. Encryption is performed using AES256 and per packet authentication is performed using HMAC-SHA256-128. Session based encryption and authentication inherently supported by 128T platform completely eliminates the need for standalone solution providing secure internetworking and multi-site VPN.

128T adopts a Zero-Trust security model which guarantees that only authorized flows traverse the network. This ensures:

- Access control for each route and authentication of all communications
- Policy based inter-router traffic encryption
- Fully distributed stateful firewall protection

128T Slices exchange metadata in the first packet as a part of the flow setup process. The metadata exchanged is signed using HMAC-SHA256-128. Optionally, the metadata could be encrypted using AES256. By signing and optionally encrypting the metadata exchanged in the first packet, 128T creates a secure environment in which 128T system's routing fabric is reserved for its own exclusive use and protecting it from insiders and eavesdroppers. The keys for encryption and per packet authentication are dynamically generated by the 128T Slices at boot time and are securely stored on the 128T Control. The 128T Control distributes these keys to the 128T Slices on an as-needed basis.

Per-session encryption is supported between all instances of 128T Slices. 128T Slice-to-128T Slice communication is protected using FIPS 140-2[1] level AES256 encryption and HMAC-SHA256-128 based per-packet authentication.

The 128T Slice-to-128T Slice encryption is done in a stateless manner by explicitly carrying Initialization Vector (IV) in each packet. The IV is generated using the FIPS140-2[2] DRBG method, The DRBG method of generating IV allows 128T platform to generate a true random number thus providing complete protection from the Man in the Middle and replay attacks.

While performing encryption of the application traffic, because of the session oriented nature of 128T routers, 128T can detect whether the traffic is already encrypted using TLS/HTTPS or by IPsec. If the application traffic is already encrypted using IPsec or TLS, 128T will not re-encrypt the packet thus eliminating the overhead associated with double encryption.

Example: Session Security

Session authentication and packet encryption is enabled on a per tenant or service level. The 128T Router configuration example shown next belong to a bank with many branches across the world. The surveillance of each branch is delegated to a surveillance company in Boston CA where a team of security and monitoring experts are video monitoring each branch 24/7. In order to reduce costs, the branches only dispose of a Broadband connection. Therefore, all video traffic must be encrypted. Under these circumstances, 128T routers can authenticate and encrypt all traffic across the Internet. A configuration example is listed below:

```
security        bank_security_cameras
      name                     bank_security_cameras
      …
      hmac-cipher              sha256
      …
      encryption-cipher        aes-cbc-256
      …
      encrypt                  true
      hmac                     true
      …
exit
```

[1] http://csrc.nist.gov/publications/fips/fips140-2/fips1402.pdf
[2] http://csrc.nist.gov/publications/nistpubs/800-90A/SP800-90A.pdf

The *security* policy above corresponds to the policy which demands to encrypt the payload of each packet using aes-cbc-256 and authenticate each packet using sha256. Otherwise, if the *encrypt* parameter is set to a value of “false”, the payload of each packet will not be encrypted. Similarly, if the *hmac* parameter is set to “false”, only the first packet of a session will be authenticated.

Once the desired security policy is created, it has to be associated with either the tenant or service we want it to apply it to:

```
    tenant        surveillance.bank
          name      surveillance.bank
          …
          security  bank_security_cameras
    exit
```

In the configuration snippet above the security policy is applied to a tenant, in this example to tenant “surveillance.bank” which includes all surveillance traffic of the bank. If the business policy is to only encrypt the traffic corresponding to the security cameras, but to not encrypt and authenticate other surveillance type traffic:

```
      service      security_cameras
          name     security_cameras
          …
          tenant   surveillance.bank
          …
          security bank_security_cameras
          …
```

Applying a security policy to a service will takes precedence over any security applied at the tenant at the same time if any.

LIFECYCLE MANAGEMENT

128T NG-WAN solution supports all stages in the deployment lifecycle as follows:

- Set-up: Create the underlying OS and configure the OS appropriately
- Install: Install 128T software
- Provision: Create and modify the configuration
- Monitor: Monitor the status of the running system
- Support: Methods to assist customers with issues
- Upgrade: Upgrade the 128T software to a new version
- De-provision: Shut-down a 128T router in the network.

For each of the stages, we may also support several different tools, including

- DevOps: A variety of development and operations tools for automation
- Installer: A 128T tool for installing software
- Conductor: The 128T authority-wide management tool
- PCLI: The 128T Router Programmable Command Line Interface
- Salesforce: The 128T Customer Relationship Management tool

SUMMARY

The 128 Technology solution utilizes an innovative service-centric, session-oriented, and security-infused routing paradigm for building context aware networks. It enables centralized control, simplifies deployment of context aware networks, introduces intelligent service routing with in-band signaling, provides fine grained micro-segmentation, and infused security based on a zero-trust model.

The 128 Technology solution provides a NG-WAN solution that goes above and beyond traditional SD-WAN offerings by solving underlying network issues and delivering unparalleled experiences.

128 TECHNOLOGY

www.128technology.com | info@128technology.com