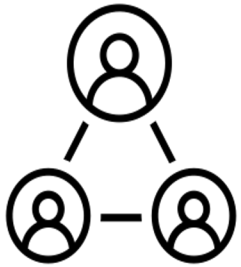
Software Defined Backbone

Experience with a deployed SD-WAN

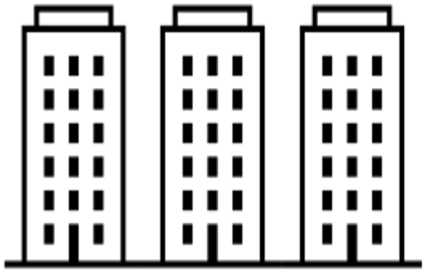
Prashanth Kumar
Jaffar Abdul

29-OCT-2025

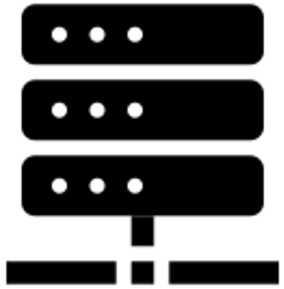
How big is LinkedIn?



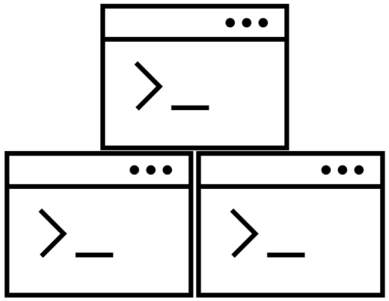
1B+ Members



Multiple Centers



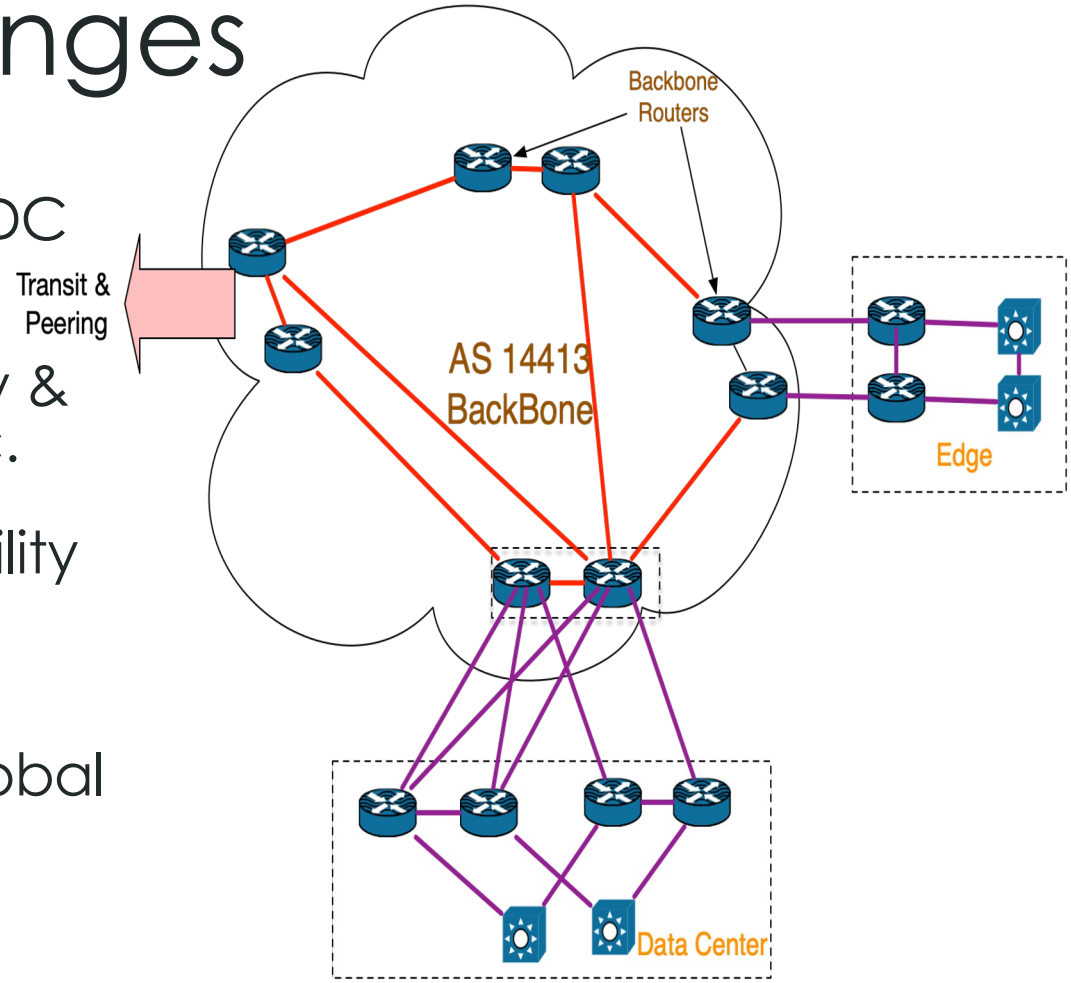
500K+ servers



3000+ services

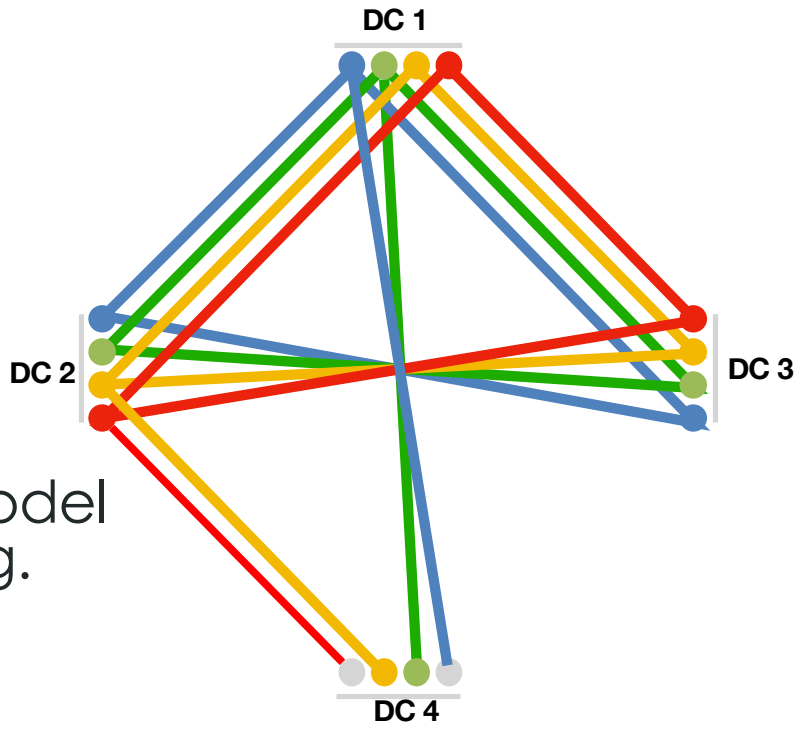
Backbone Challenges

- RSVP-TE Backbone with Auto Bandwidth for Internet & Inter-DC traffic.
- Scaling: Inter DC traffic is bursty & grew faster than Internet traffic.
- Efficiency: Prioritization and ability to use all available capacity.
- Operations: Manageability overhead and RSVP lacked global view.

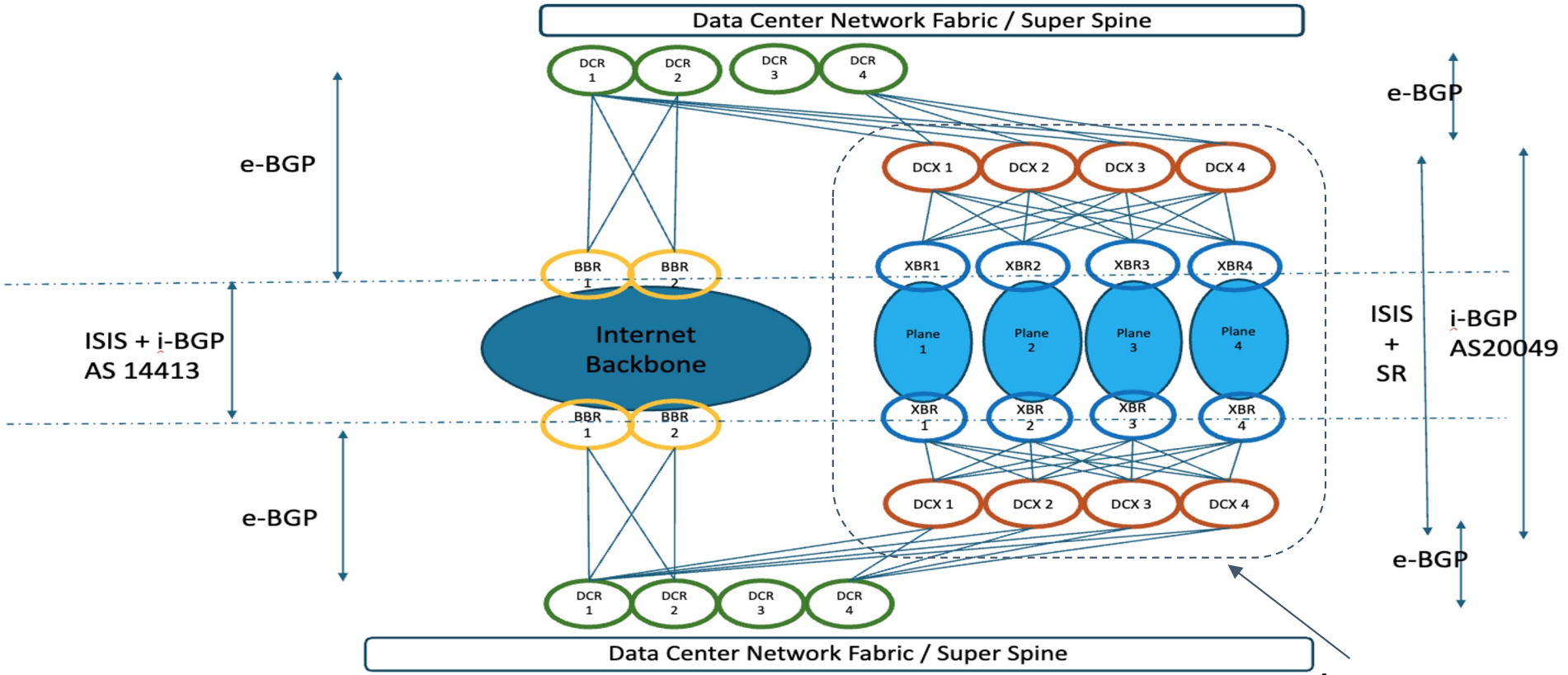


Next Gen Backbone (NGBB)

- Built separate Backbone for Inter DC Traffic.
- Scaling: Planar architecture, IS-IS SR control plane with MPLS data plane.
- Efficiency: Centralized view of traffic demands and allocation with central prioritization.
- Operations: Operated on network model on what, global management tooling.



Backbone Control Plane Design

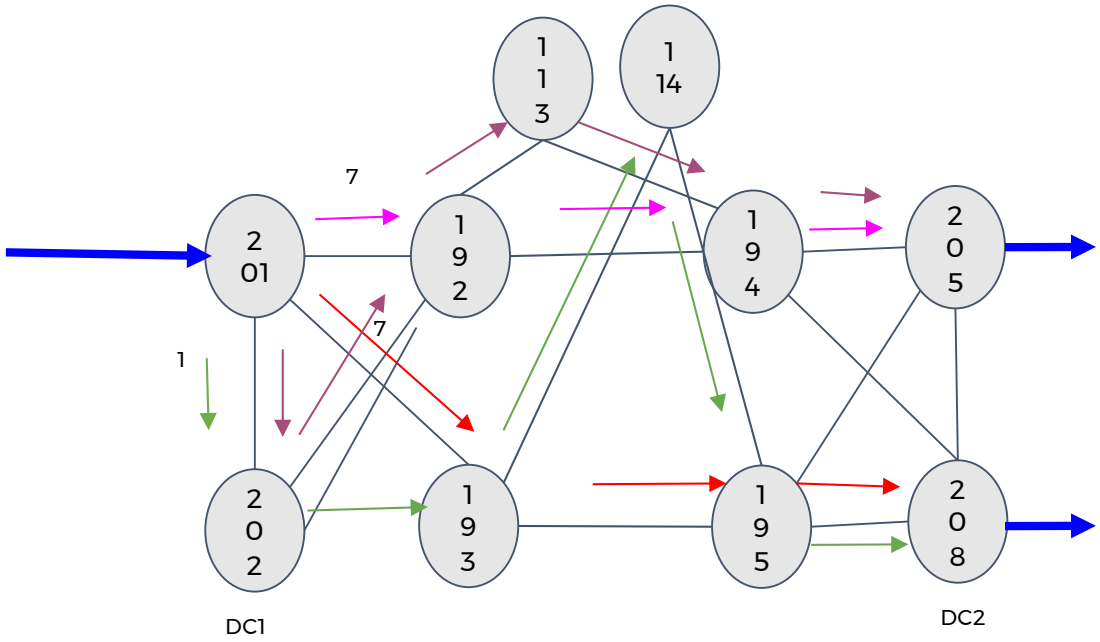


Centralized Policy Management

- Centralized controller analyzes traffic demands periodically or when topology changes.
- Priority of Traffic is market at source.
- Compute the paths to satisfy demands based on priority
- Pushes Policy Based Routes (PBR) on to ingress routers with CBF into Weighted Cost Multiple Paths (WCMP).

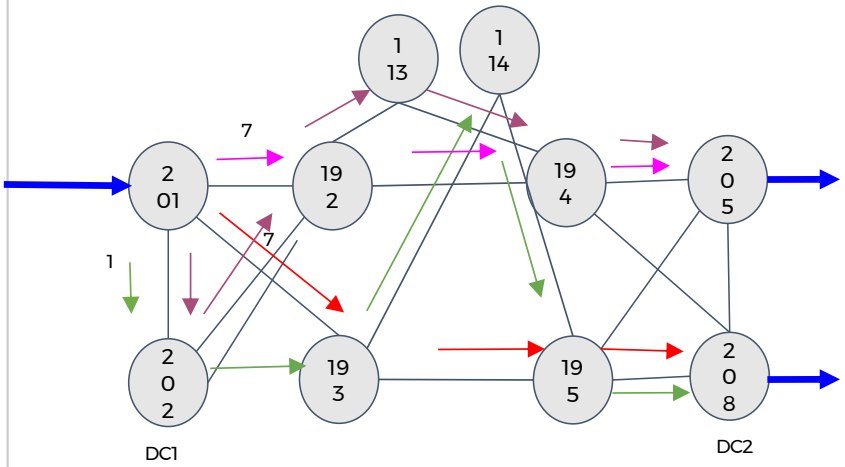
Ingress Router: Class Based Forwarding

Forwarding Equivalence Class (FEC)	Nexthop Group (NHG)
1 permit ip any 10.x.24.0/21 dscp 18	0_0_0_0_010_x_24_0_210x12
1 permit ip any 10.x.24.0/21 dscp 0	0_0_0_0_010_x_24_0_210x0



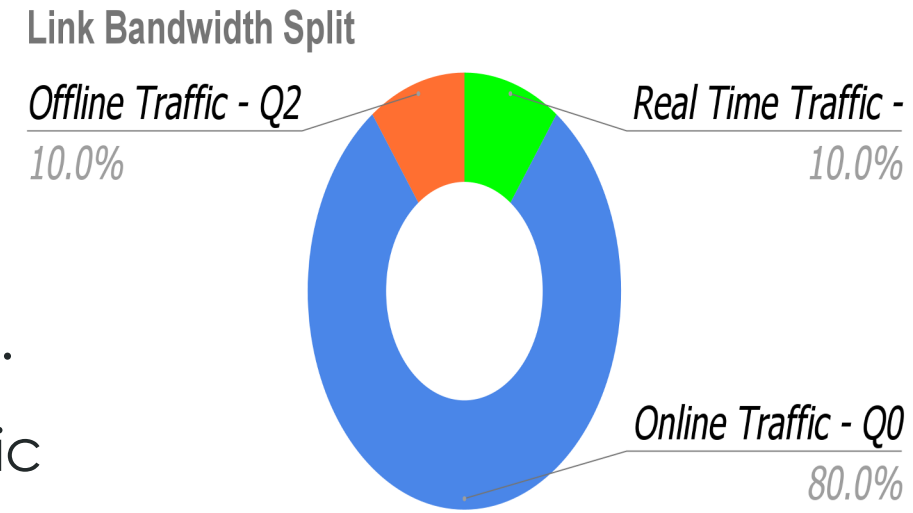
NANOG™

Ingress Router - CBF (continued)

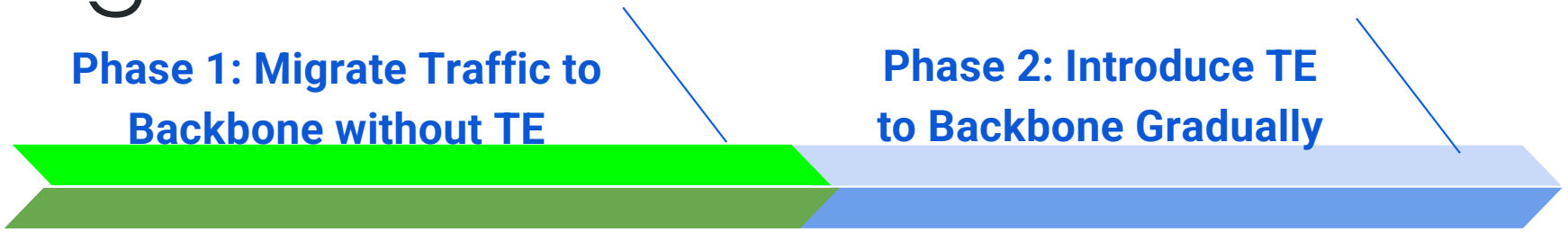
Nexthop Group (NHG)	Label Stack	Weight
0_0_0_0_010_x_24_0_210x12 	900194 900205 nexthop 10.x.1.3	7
	900195 900208 nexthop 10.x.1.4	7
	900192 900113 900194 900205 nexthop 10.x.1.2	1
	900193 900114 900195 900208 nexthop 10.x.1.2	1

NGBB QoS Design

- Hosts marks DSCP based on workload priority.
- Network Trust DSCP & different weights assigned to different FEC.
- Controller steers traffic based on FEC priority.
- Network Control traffic follows strict priority queue - Always handled first.
- Routers drop the lowest priority traffic during congestion.



Migration

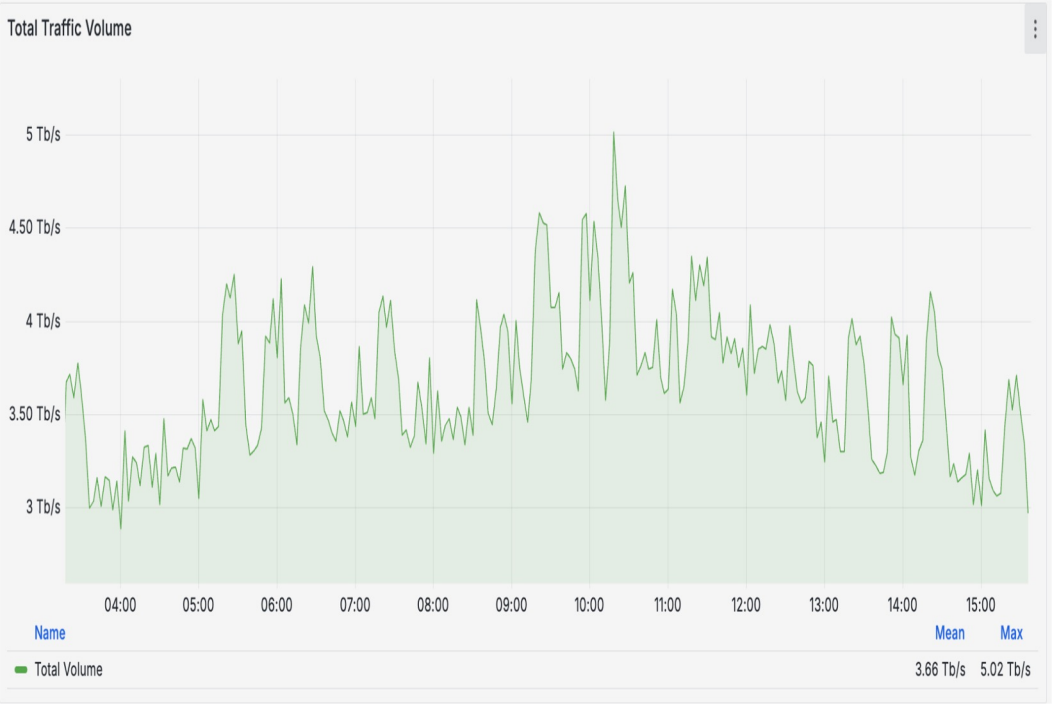


- Parallel SR Best-Effort Backbone for Inter-DC traffic.
- Migrated Traffic between 2 sites.
- Gradually onboarding rest of the sites.

- Introduced SDN based TE gradually for subset of Networks between 2 sites.
- Full SDN based TE rollout across all sites.

Impact

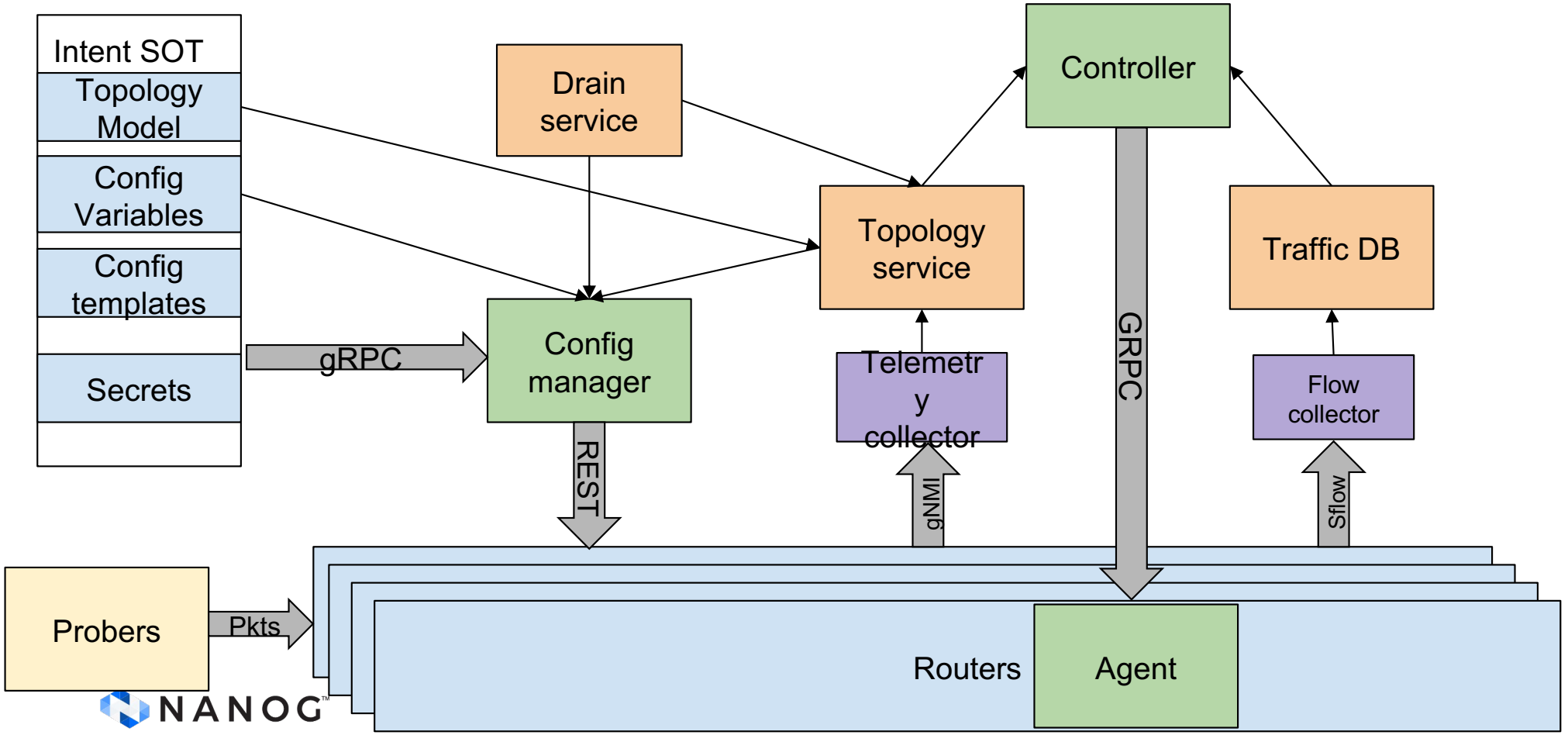
- Use up to 95% total capacity & cost efficient.
- Translate Business logic to traffic policies.
- Tailored QoS and better application performance.
- Simplified network operations.
- E2E visibility into traffic patterns & network health.



Software Design

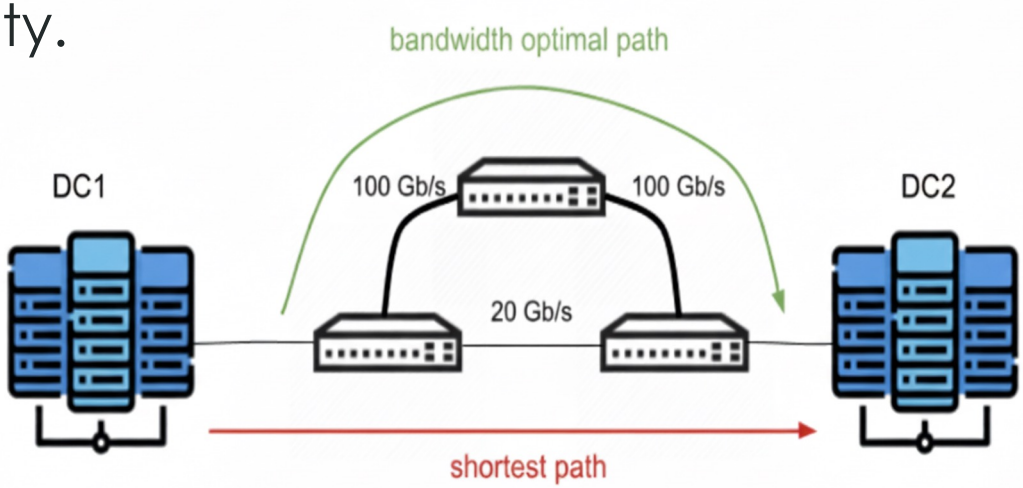
- Fully in house built, written mostly in Go.
- Microservice architecture & gRPC for inter service communications.
- REST, gRPC, & gNMI used for device interactions.
- Automated deployment of configurations, topology.
- Maintain code and SoT separately.

Software Architecture



Centralized Controller

- Leverage full network capacity.
- Paths are expressed using PBR in the routers.
- Different traffic types handled differently.
- React to changing traffic conditions.



Traffic Demand Calculation

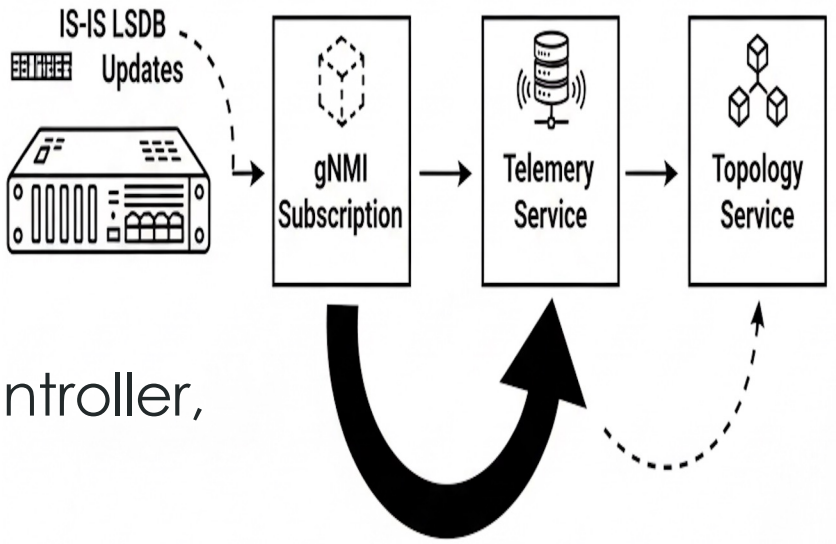
- NGBB Edge devices (DCX) exports the flow data using sFLOW.
- Sampling rate is set to 1 in 1000 packets.
- Multiple collectors instances are deployed per site for redundancy.
- Samples are aggregated by priority at each site using supernets.

	A	B	C
A		120G	70G
B	900G		10G
C	400G	100G	

NANOG™

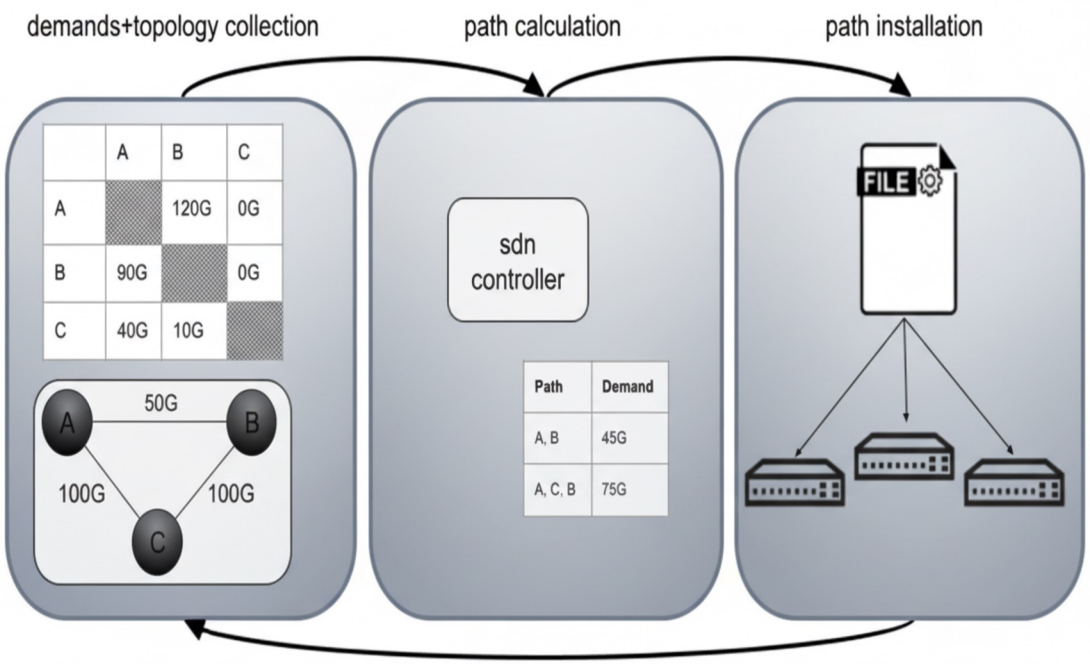
Operational Topology Acquisitions

- Operational topology built using IS-IS LSDB info.
- Routers stream LSDB info via gNMI.
- Topology service parses LSDB updates to build the network graph.
- Graph updates are pushed to the controller, initiating CSPF.



Traffic Engineering Algorithm

- TE algorithm runs every 10 seconds or on topology change.
- CSPF computes bandwidth constrained paths.
- Paths scope from ingress DCX device to destination site.

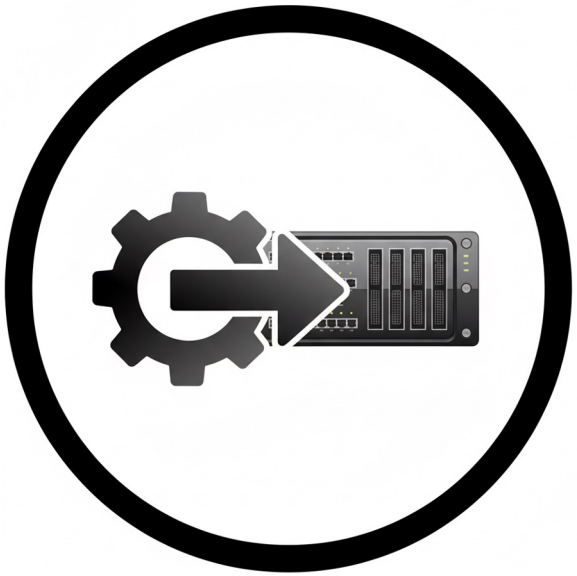


Policy computation & installation

- When multiple paths exist, they are weighted based on available capacity.
- Policies define paths as an MPLS label stack of Node SIDs.
- The controller pushes vendor-agnostic policies to routers via gRPC.
- A device-specific agent on the router translates policies into flow data.

Configuration Management

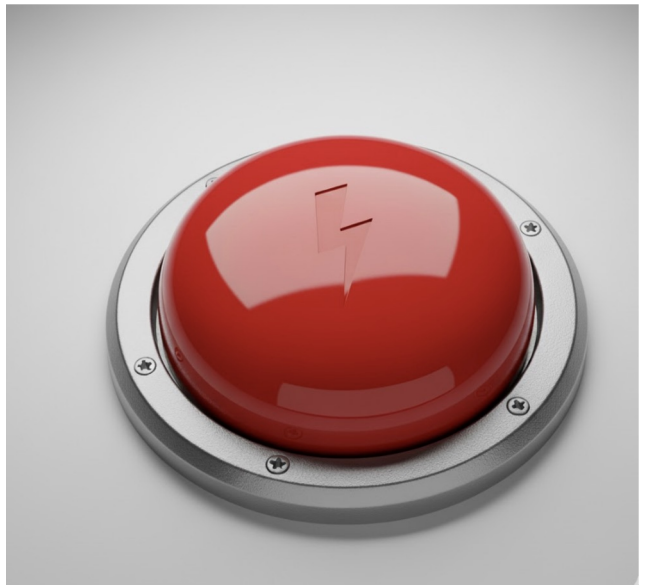
- We think about the network, not about individual devices.
- No access to CLI for normal operations.
- Software manage full configuration for the backbone devices.



- Multiple sources to build a config: topology, variables, secrets, templates.

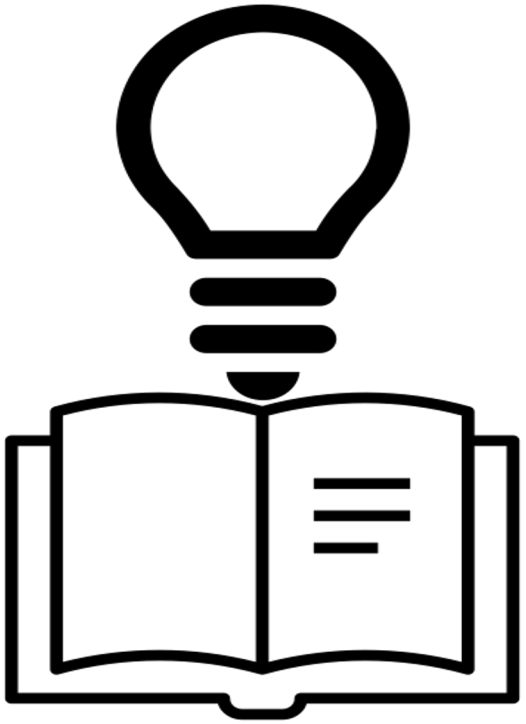
Safety and Manageability

- Fail Static: Fallback to IGP shortest path from Centralized controller during failure.
- A big red button to disable the controller for emergency.
- Black box real-time prober.
- Drain and Risk assessment with Simulation.



Lessons Learned

- Be deterministic with traffic allocation.
- Humans should express what and not how.
- Intended vs Operational.
- Rigorous simulation and lab testing are critical.
- Progressive rollout enabled early issue detection and faster iteration.



Lessons Learned (continued)

- Delta between controller traffic allocation vs real traffic in the network.
- Capacity accommodation for dark traffic is tricky.
- Vendor agent dependency for policy injection can restrict portability.
- Visibility to map the controller demand allocation to network links.

Thank you