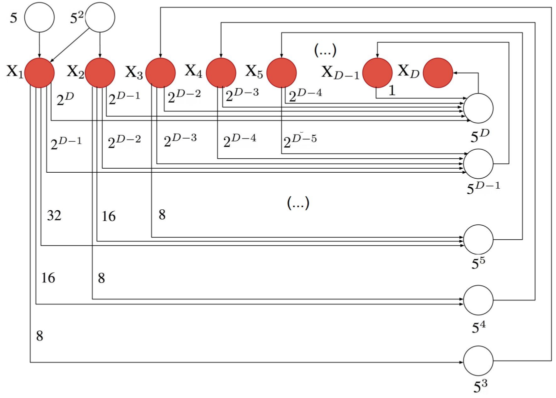
GradientGraph™ (G2) Analytics: High-Precision Analysis of Flow and Bottleneck Performance in Data Networks

HPC User Forum

Jordi Ros-Giralt, Sruthi Yellamraju, Noah Amsel, Brendan Von Hofe, Jamie Nachbar, Lucy O'Toole, James Ezick, Alison Ryan, Richard Lethin

Technology supported by the Department of Energy under SBIR Award DE-SC0019523

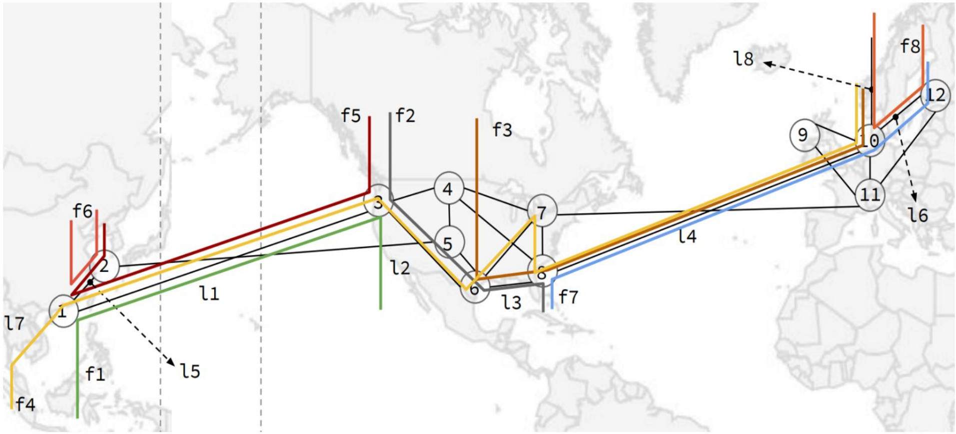


Reservoir Labs

632 Broadway 803, New York,
NY 10012,
United States

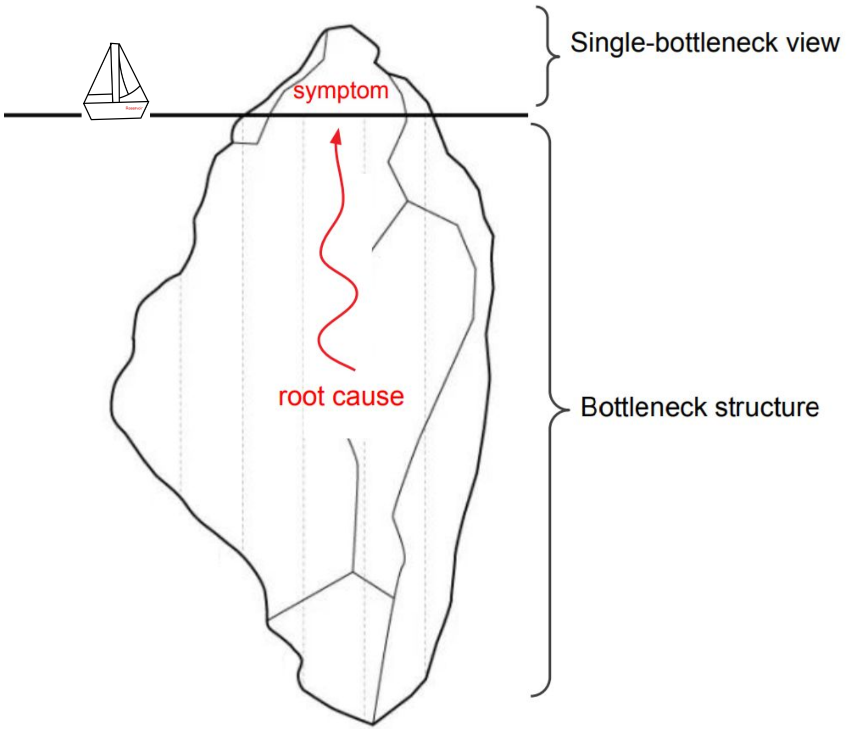
Growing demands on networks - flow optimization

- Scientific collaborations and data-centric research at the "regional, national and global scale."
- Ever increasing rates of data production from scientific experiments (LHC, genomics, astronomy, physics, …)
- Commercially, movement of logs (transactions, cyber) and training data
- Adoption of 5G and new applications (telemetrics, AI on demand, xR)

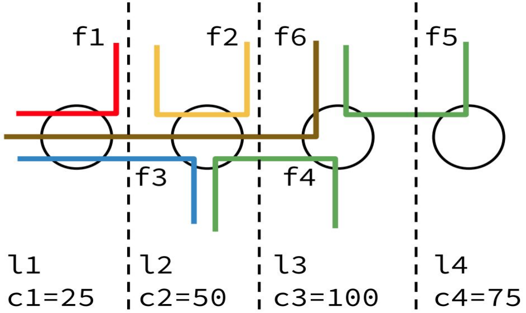


Problem Positioning: The Root Cause Was Hiding

- Structure of the congestion problem in data networks:
  - The single-bottleneck problem is the tip of the iceberg.
  - The bottleneck structure is the submerged portion.

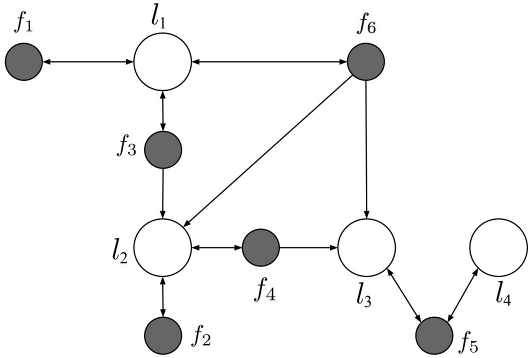


Theory of Bottleneck Structures



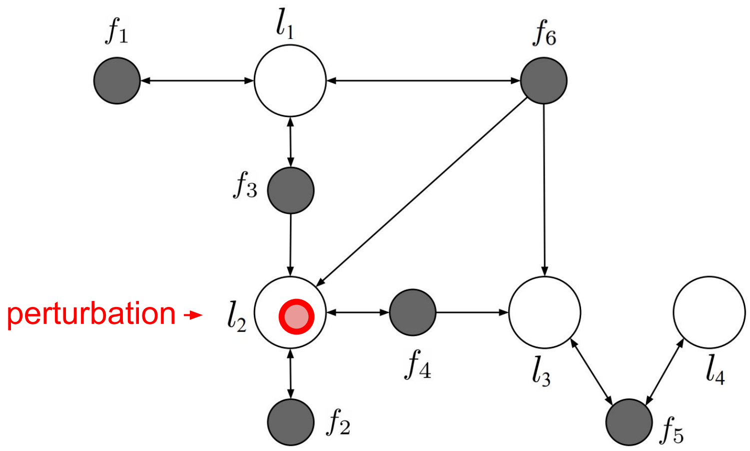
G2

Bottleneck structure:



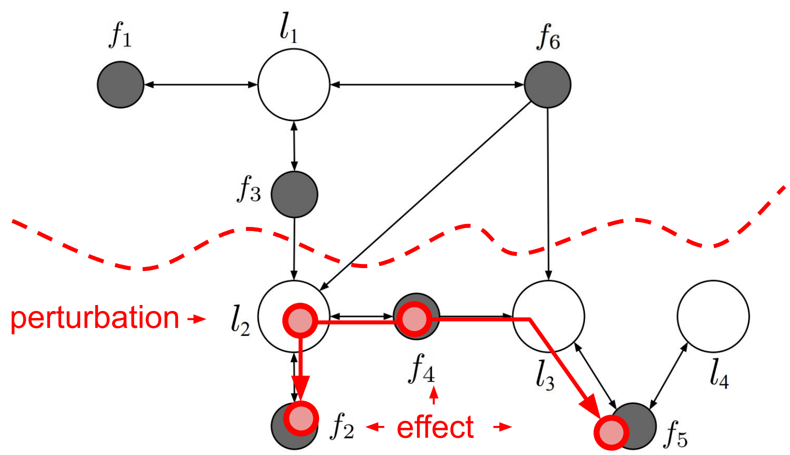
[1] Mathematics presented at ACM SIGMETRICS (June 2020) and ACM SIGCOMM (August 2021): https://bit.ly/3eGOPrb

Theory of Bottleneck Structures



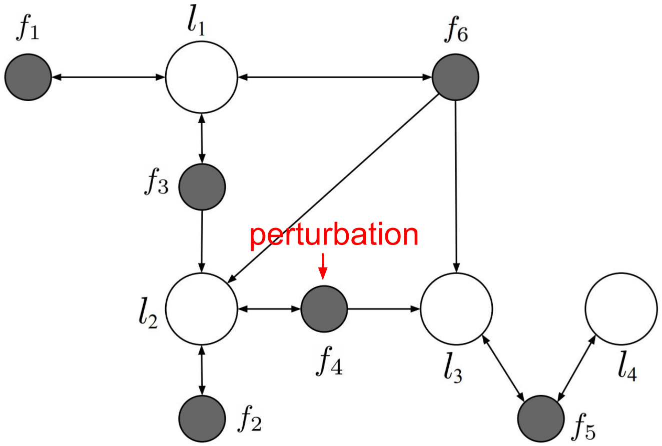
[1] Mathematics presented at ACM SIGMETRICS (June 2020) and ACM SIGCOMM (August 2021): https://bit.ly/3eGOPrb

Gradients Through Bottlenecks - Estimate Effects of Changes



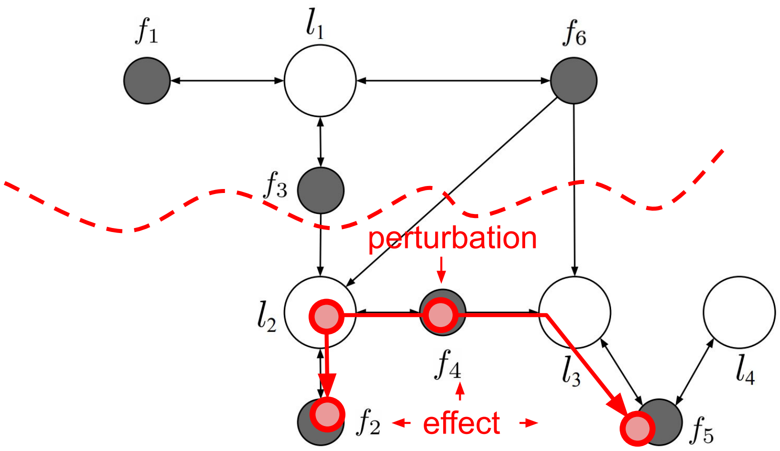
[1] Mathematics presented at ACM SIGMETRICS (June 2020) and ACM SIGCOMM (August 2021): https://bit.ly/3eGOPrb

Gradients Through Bottlenecks - Estimate Effects of Changes



[1] Mathematics presented at ACM SIGMETRICS (June 2020) and ACM SIGCOMM (August 2021): https://bit.ly/3eGOPrb

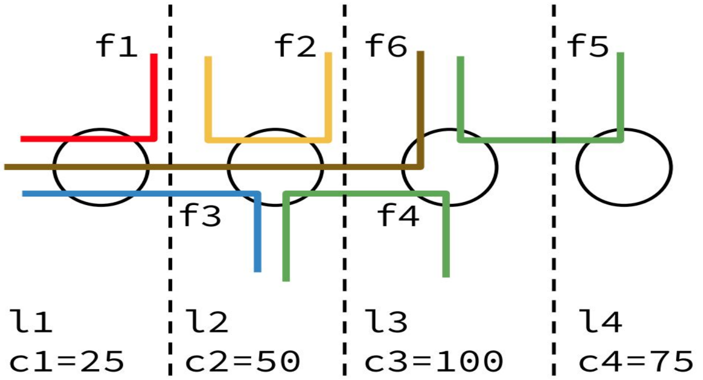
Gradients Through Bottlenecks - Estimate Effects of Changes



[1] Mathematics presented at ACM SIGMETRICS (June 2020) and ACM SIGCOMM (August 2021): https://bit.ly/3eGOPrb

G2: Analytical Environments and REST API Programmable Interface

Application 1: Elephant Flow Detection

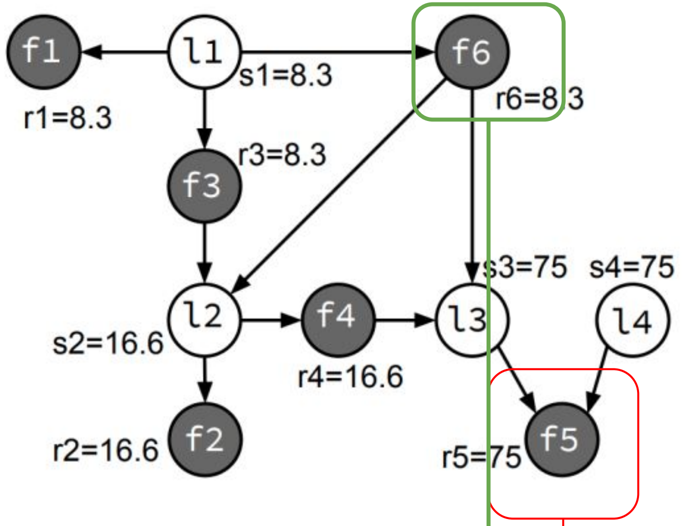


Flow transmission rate					
r1	r2	r3	r4	r5	r6
8.3	16.6	8.3	16.6	75	8.3

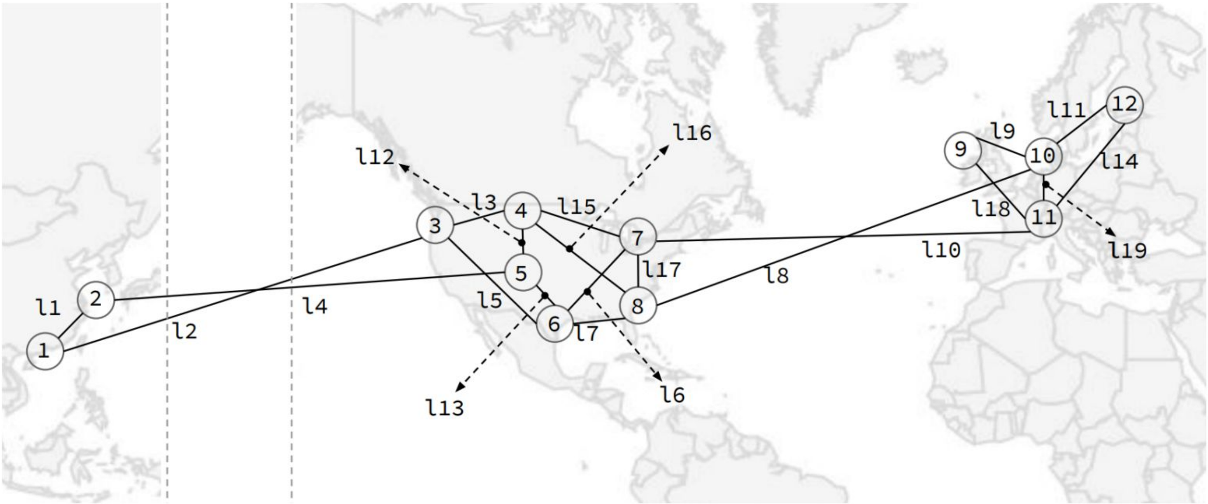
true elephant flow

heavy hitter flow

Flow Gradient Graph:



Application 2: Identification of Max-Throughput Paths



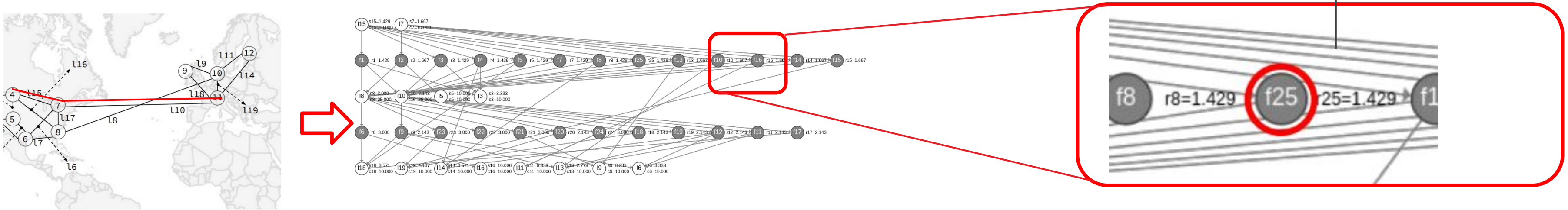
* Google's B4 network

- Background traffic: flows connecting every pair of datacenters between the US and Europe (both directions).
- Goal: transfer 1 terabyte of data from datacenter 11 (EU) to datacenter 4 (US) using the highest-throughput path available.

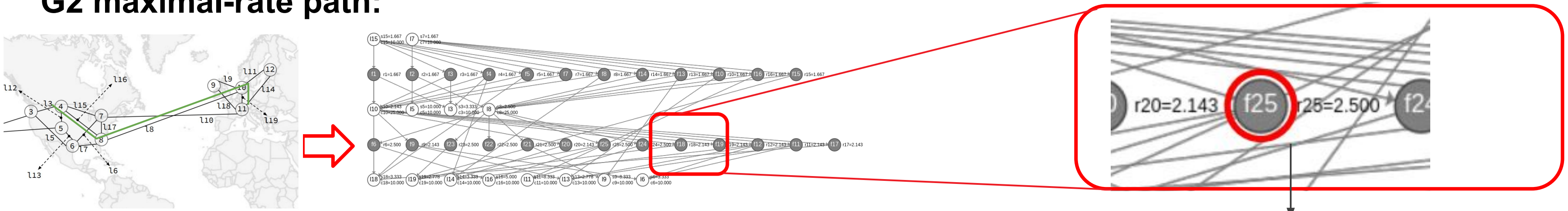
Application 2: Identification of Max-Throughput Paths

$r_{25} = 1.429$

Traditional lowest-cost path:



G2 maximal-rate path:



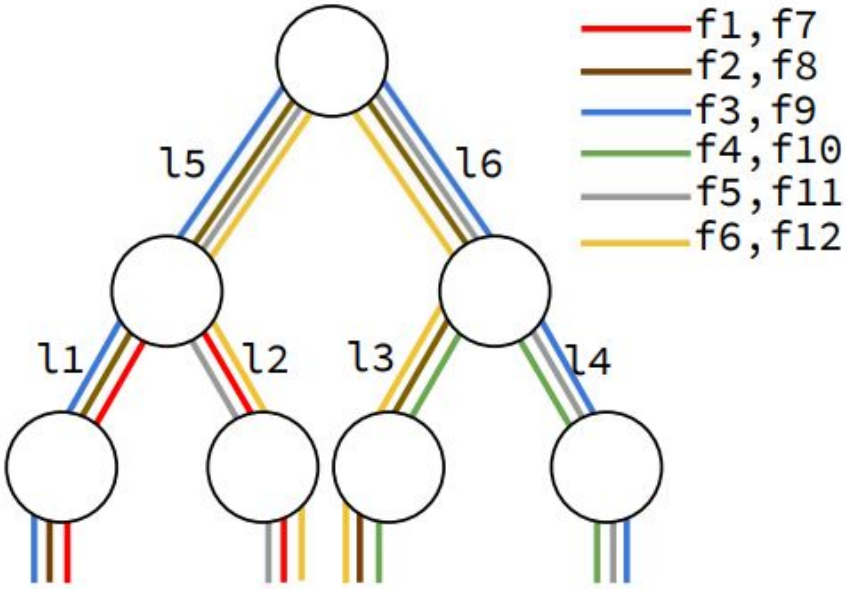
$r_{25} = 2.5$

- Flow performance improvement of **74.95%** without affecting lower throughput flows

Application 3: Capacity Planning in Fat-Tree (Clos) Networks

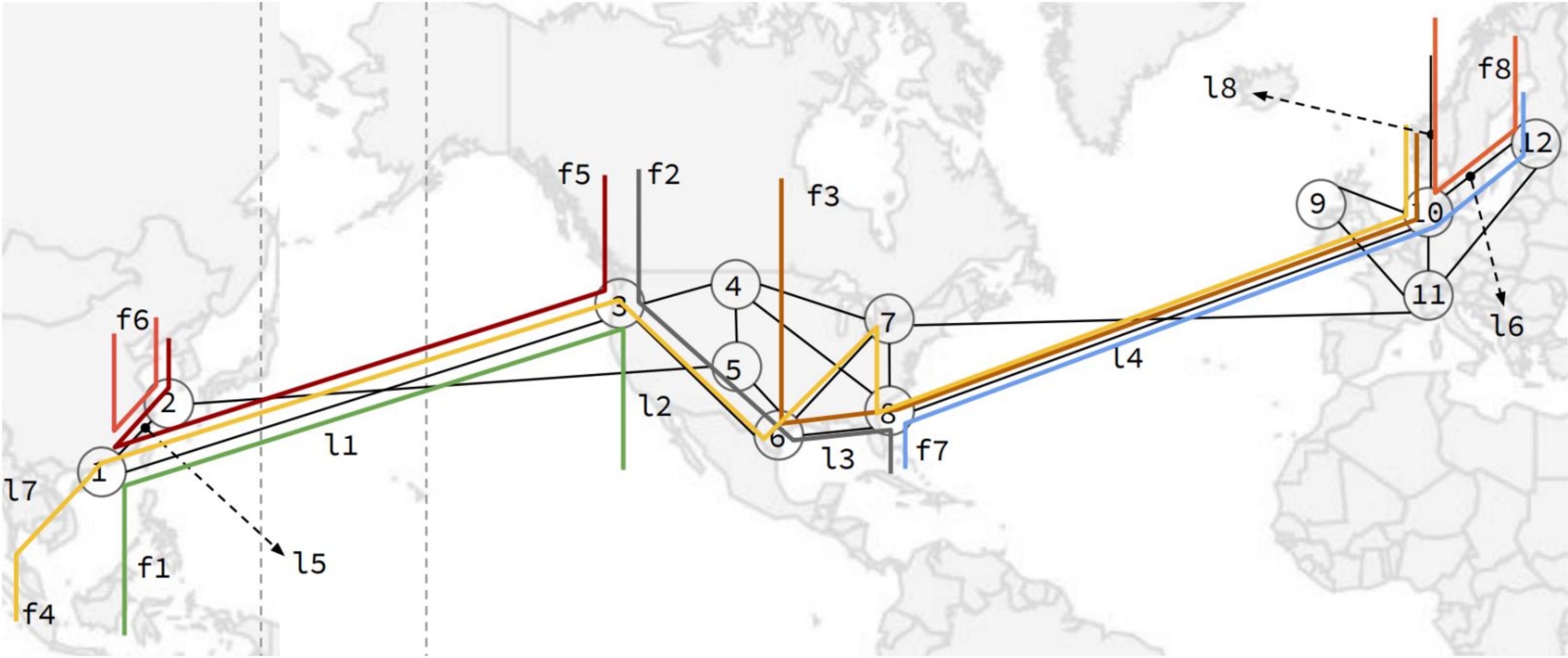
$$c_{l_5} = c_{l_6} = \lambda \times \tau$$

$$c_{l_1} = c_{l_2} = c_{l_3} = c_{l_4} = \lambda$$

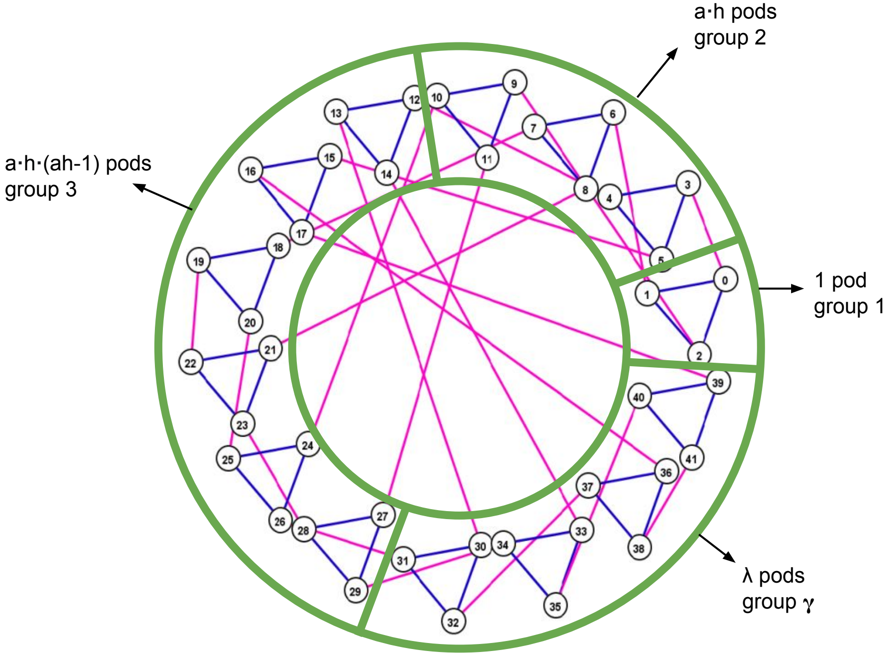


Application 4: Acceleration of Time-Bound Constrained Flows

Google's B4 SDN WAN Network (SIGCOMM 2017)



Application 5: Optimal Design of Data Centers and Supercomputers



Let α be the skeweness parameter of a dragonfly topology with parameters a, g and h. Then, the following design minimizes the maximum flow completion time of the network:

$$c_1 = \frac{\rho_1}{\alpha + \rho_2 - 1} c_2$$

where

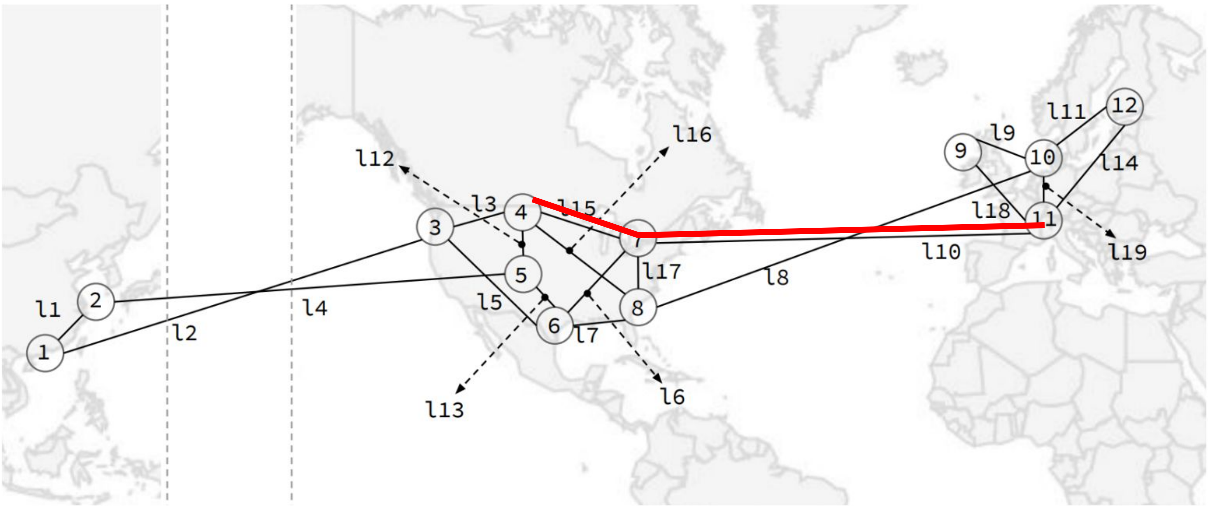
$$\rho_1 = \frac{1}{h}\left[a(g-1) + \sum_{i=0}^{\gamma-2}\left(a(g-1) - \sum_{j=0}^{i} a^2 h(ah-1)^j\right)\right]$$

$$\rho_2 = g + \lambda + \frac{1}{a-1}\sum_{i=0}^{\gamma-2}\left[ah(a-1)(ah-1)^i + \frac{(a-1)h}{ah-1}\left(a(g-1) - \sum_{j=0}^{i} a^2 h(ah-1)^j\right)\right]$$

$$\gamma = max\left[x \mid a(g-1) - \sum_{i=0}^{x-2} a^2 h(ah-1)^i \geq 0\right]$$

$$\lambda = g - 1 - \sum_{i=0}^{\gamma-1} ha(ah-1)^i$$

Application 6: SLA Breach Detection and Recovery

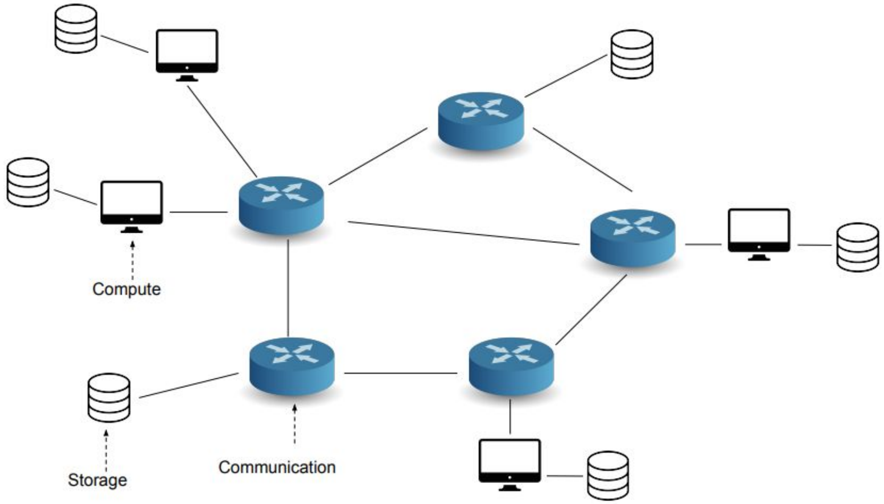


* Google's B4 network

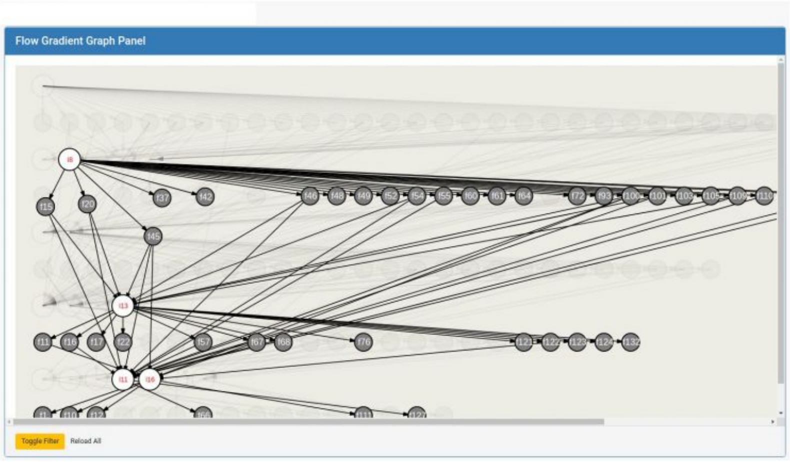
- **Situation**: Red flow performs at 1.429Gbps, but its SLA requires a minimum of 2Gbps.

Bottleneck Structures and Gradients Can Be Computed in Real Time

Network:



Model:



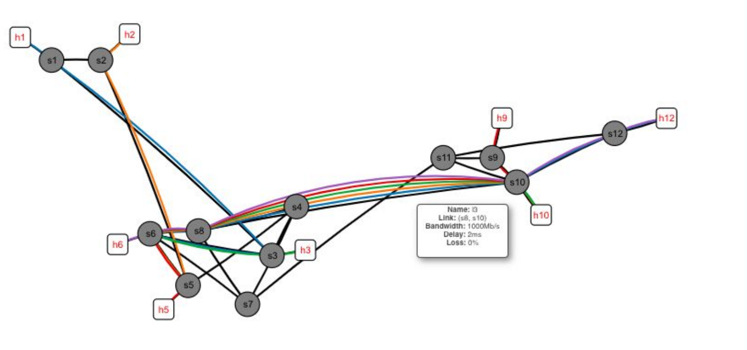
GradientGraph™ Analytics: Platform

Reservoir Labs, Inc.

Network Topology Dashboard

Dashboard

Topology Panel



Toggle Filter | Toggle Flows | Reload Flows | Reload All

Switches

Switches | Hosts | Links | Flows

Type in element to search

Show 10 entries

ID	Links Attached
s1	(s1 , s3) (s1 , s2)
s10	(s9 , s10) (s8 , s10) (s10 , s12) (s10 , s11)
s11	(s9 , s11) (s10 , s11) (s7 , s11) (s11 , s12)
s12	(s10 , s12) (s11 , s12)
s2	(s2 , s5) (s1 , s2)
s3	(s3 , s6) (s3 , s4) (s1 , s3)
s4	(s3 , s4) (s4 , s8) (s4 , s7) (s4 , s5)
s5	(s2 , s5) (s5 , s6) (s4 , s5)
s6	(s3 , s6) (s6 , s8) (s6 , s7) (s5 , s6)
s7	(s7 , s11) (s7 , s8) (s6 , s7) (s4 , s7)

Showing 1 to 10 of 12 entries

Previous | 1 | 2 | Next

Dashboard Configuration

Automated Refresh

Off

Refresh Rate (s)

10

Time to refresh (positive integer)

Overlap Flows

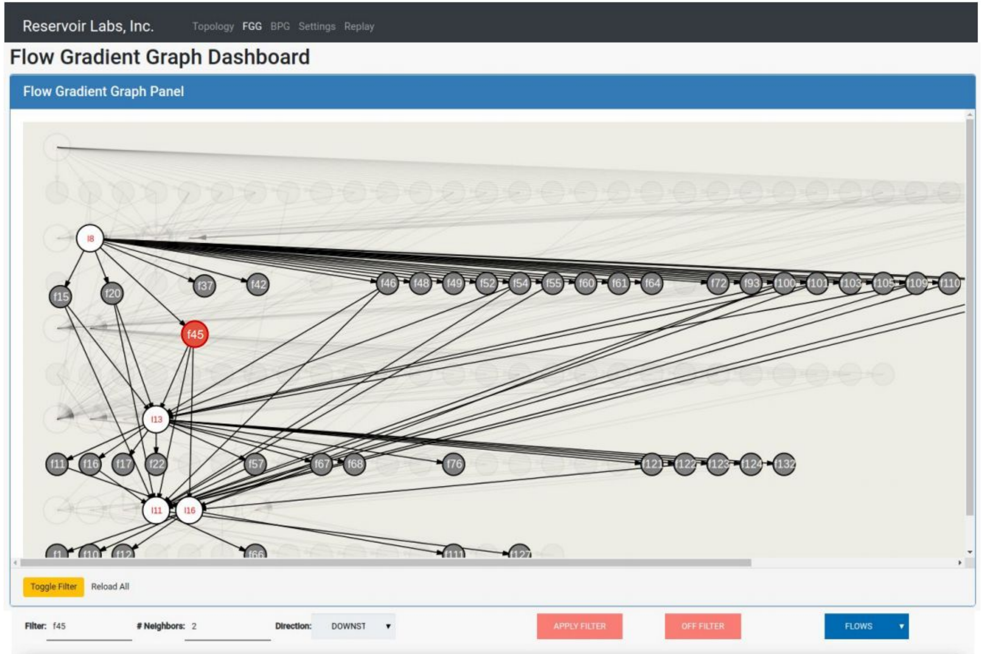
No

Automated Refresh Stops

☐ **Export Configuration**

Apply

GradientGraph™ Analytics: Platform



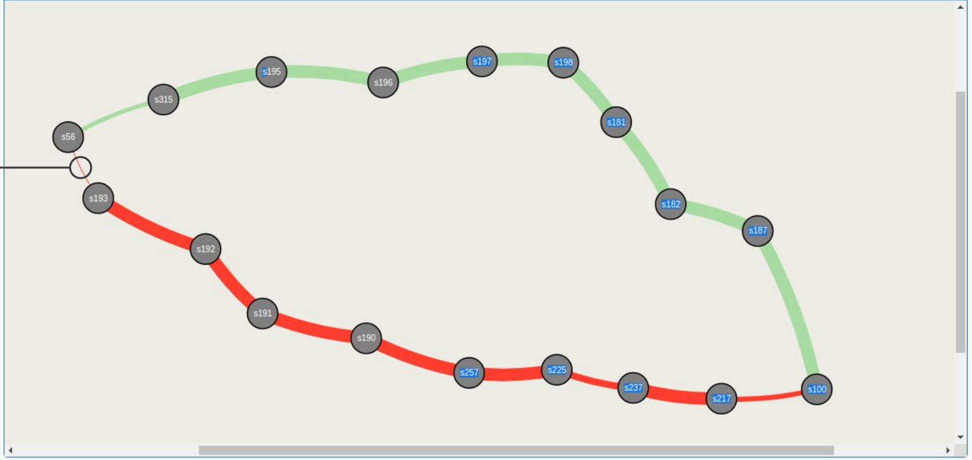
ID.	Start Host	End Host	Flow Size (bytes)	Gradient ↓	Transmission Rate (Mb/s)	Fairshare (Mb/s)	In Nodes	Out Nodes
f45	h5	h1	256000000	1.123	3.21	3.21	l8	l16, l11, l13
f111	h11	h1	256000000	0.667	18.97	18.97	l16, l19	None
f42	h4	h10	256000000	0.307	3.21	3.21	l8	l3, l1

GradientGraph Analytics: Routing through Low Congestion Paths

Avoiding high-congestion path

Routing Dashboard

Routing Panel



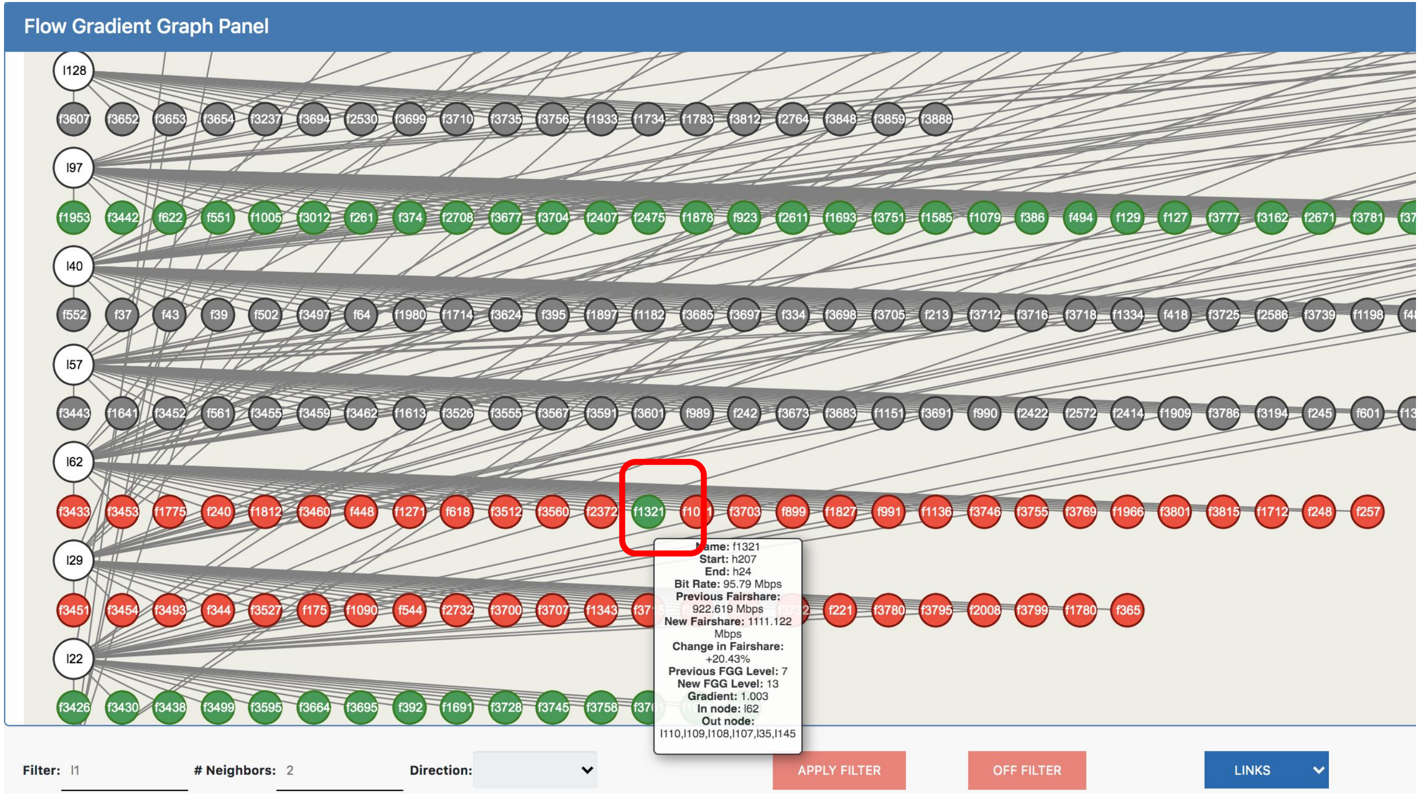
Filter: f14004

ROUTE FLOW | APPLY FILTER | OFF FILTER | OFF ROUTING

ID	Path	Flow Rate (% Change)	FGG Level (Δ)
f14004 (active)	s100→l81→s217→l82→s237→l83→s225→l114→s257→l120→s190→l16→s191→l17→s192→l18→s193→l121→s56	434.953 (0%)	7 (0)
f14004	s56→l233→s315→l211→s195→l22→s196→l23→s197→l24→s198→l25→s181→l4→s182→l11→s187→l209→s100	2456.14 (+464.69%)	35 (+28)

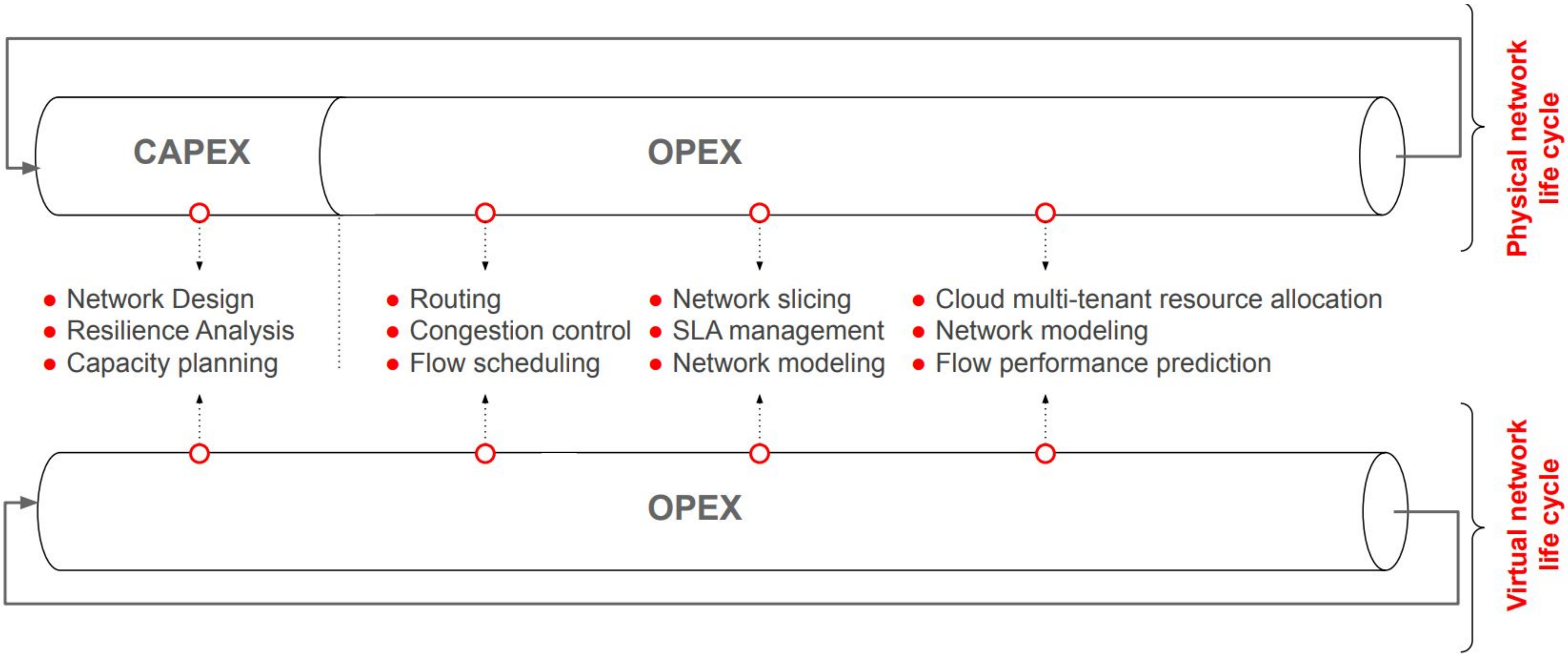
Identification of a path 464% faster

GradientGraph™ Analytics: Platform



Total system throughput increases by 3.925Gbps

G2: Optimizing the Entire Network Lifecycle



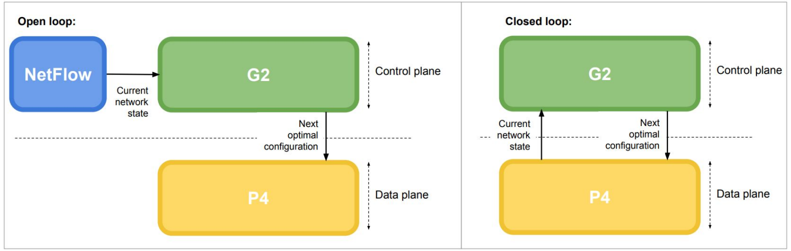
SC21 Demo: SENSE + G2 + NetPredict + P4 Use Case

"Laboratory use case" using ESnet SENSE/NetPredict services, the PRP federated k8s clusters, and Reservoir Labs G2 instance to resolve congestion in real-time

(1) Generate long-lasting flows with background traffic

(2) Create congestion on one or more segments

(3) Compute the bottleneck structure and identify the impactful flows using G2



(4) Predict next state of the network using NetPredict,

(5) Use G2 Flow Gradient Graph to compute best alternative path recommendations

(6) Re-route flows onto their best alternative paths using P4

(7) Validate the impact of re-routing flows is as predicted and that congestion has been resolved

(8) Validate the impact of re-routing flows is as predicted

Near-term operational goals:

- Embed the 4-step sequence in production operational networks: (1) congestion detection, (2) impactful flow-group prediction and identification, (3) agile flow steering, (4) verification of congestion mitigation
- Tune sequence of steps and decision parameters
- Develop success metrics
- Predict and optimize using machine learning

Current Status

- Initial deployment with Research and Educational Networks and ESnet
- Looking forward to collaborations to test the technology in data center and supercomputer fabrics
- See you at SC21 - stop by SCinet to talk with us!
- More information at http://www.reservoir.com/gradientgraph