



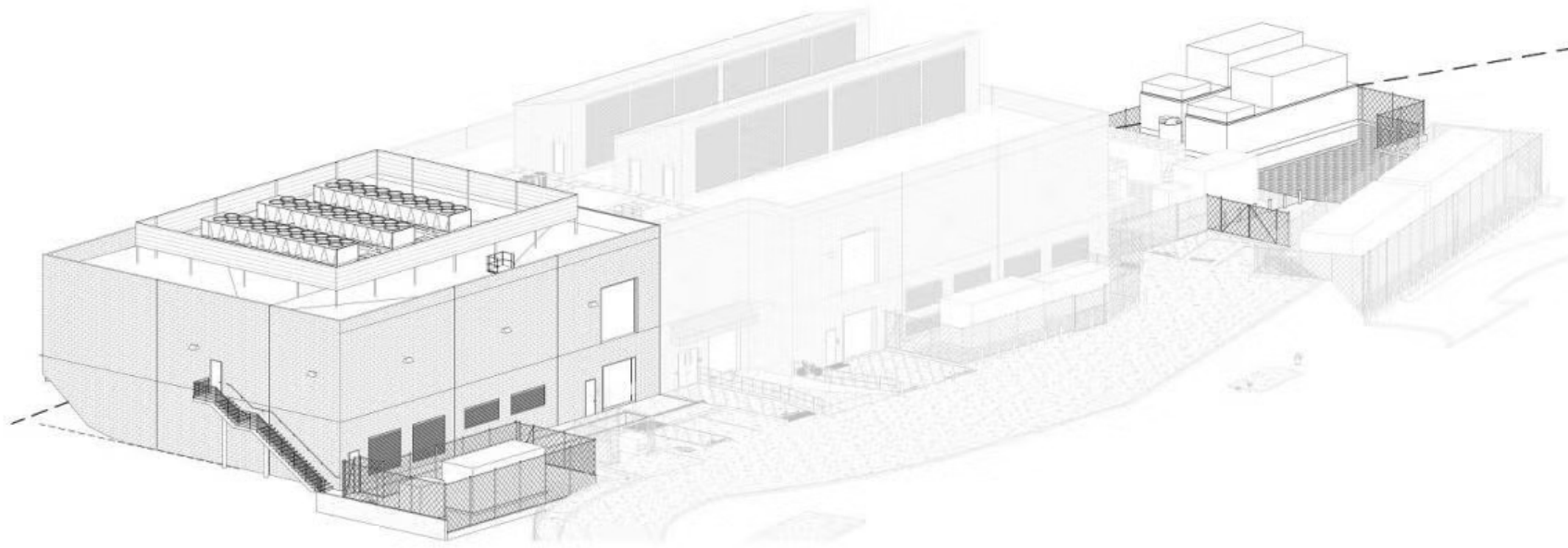
USDF Infrastructure

Riccardo Veraldi

Agenda

- Background
 - SLAC/S3DF overall network overview
 - SRCF-I SRCF-II and S3DF
- S3DF Network
 - S3DF network layout
 - S3DF network performance
- Embargo Network
 - Layer2/3
 - IPsec Tunnels and routing
 - Embargo Network monitoring

SRCF

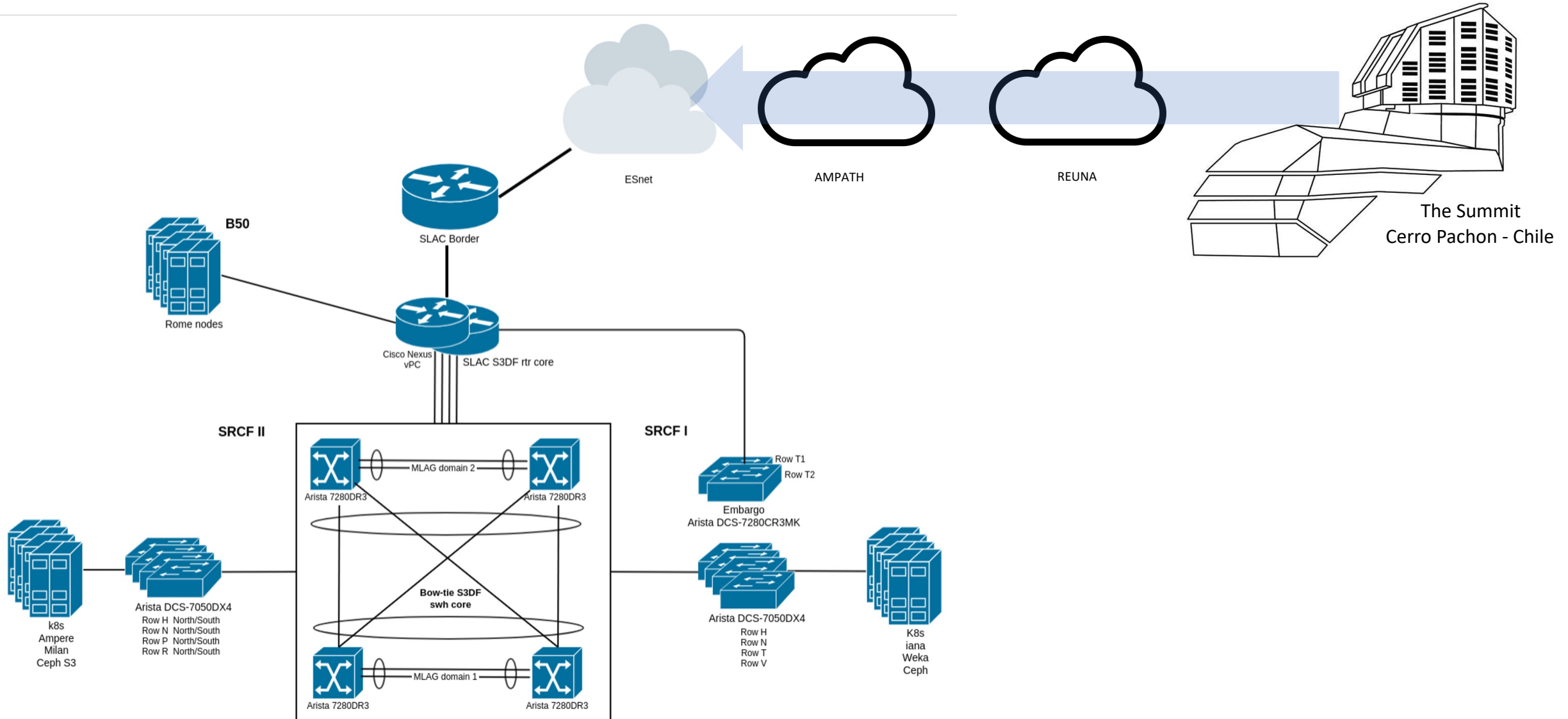


- SRCF-II doubled the data center capabilities, for a total of 6 megawatts of power capacity.

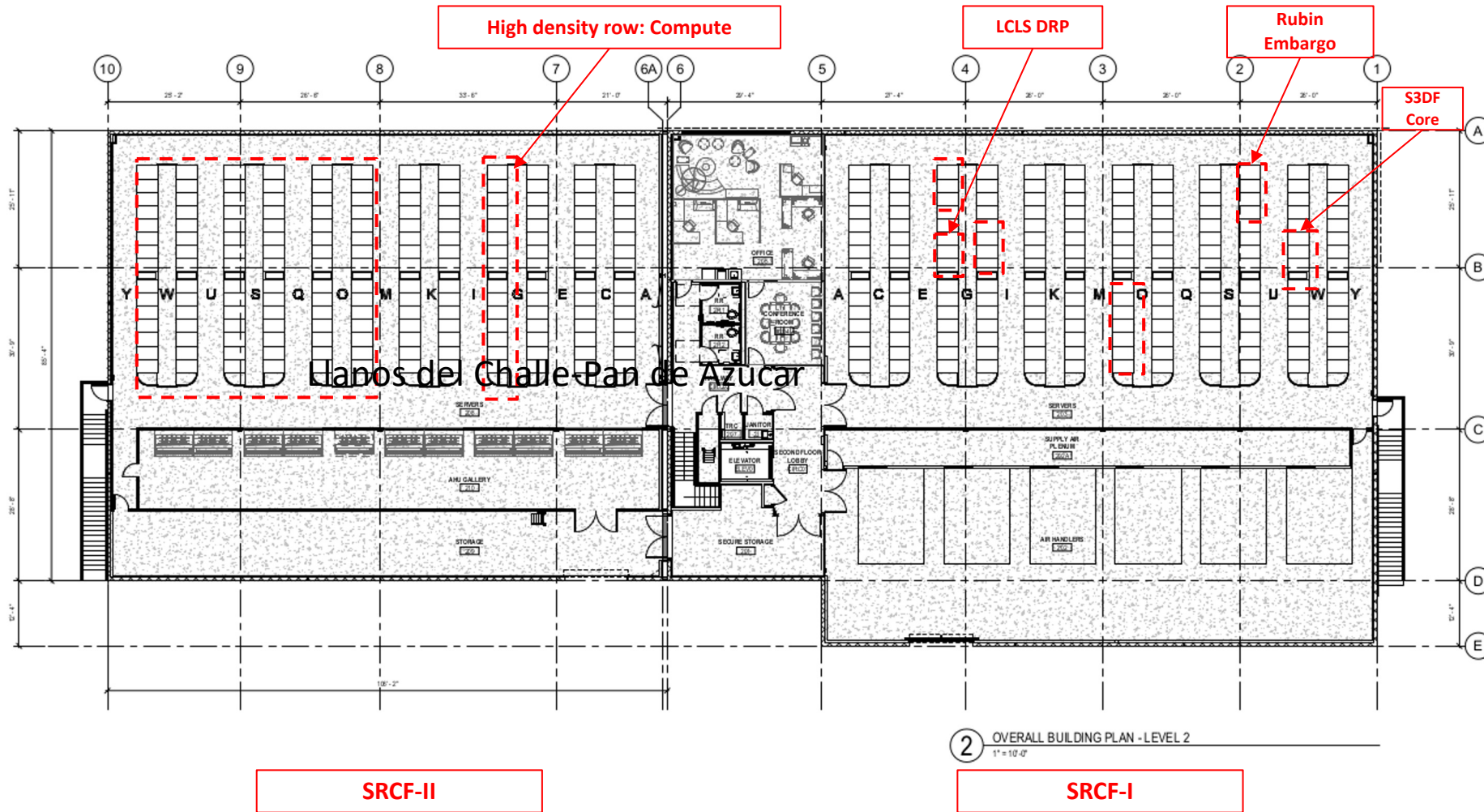
S3DF

- S3DF is a compute, storage, and network architecture designed to support massive scale analytics required by SLAC experimental facilities and programs, including LCLS/LCLS-II, Vera C. Rubin Observatory, UED, and the Stanford-SLAC cryoEM Center (S2C2)
- Rubin USDF is hosted at SLAC within S3DF

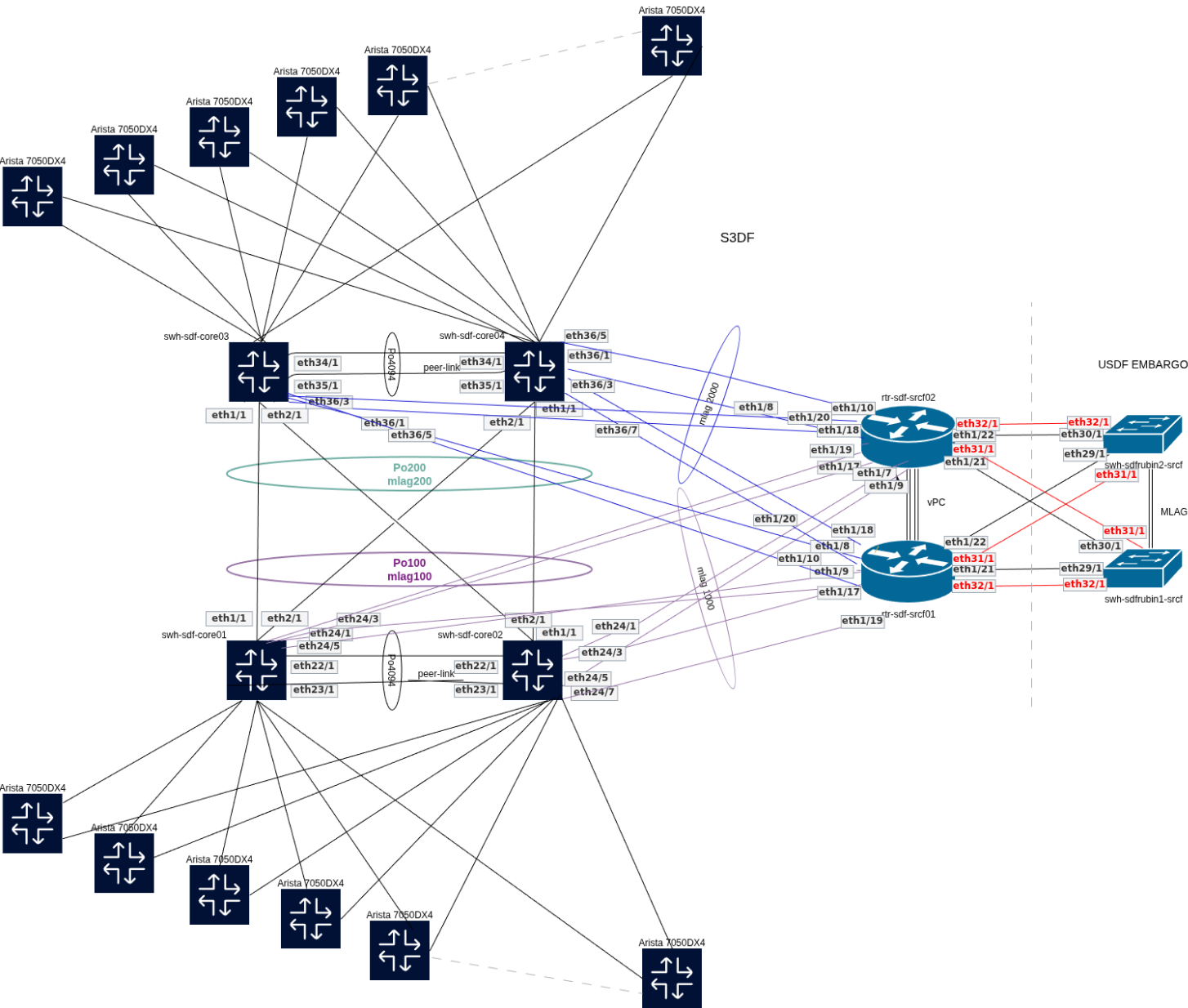
SLAC/USDF network overview



SRCF/S3DF



S3DF Network Layout



- 4x Arista DCS-7280R3 (Jericho)
 - 24x to 36x 400G ports - ultra deep buffer
- 33x Arista DCS-7050DX4 (Trident)
 - 32x 400G ports
- BowTie Model
 - Layer 2
 - Allows connections between two different MLAG switch pairs - 4x400G LACP port channel
 - 8x100G to S3DF SLAC routers
 - Layer 3:
 - For each SVI define a **ip virtual-router address**
 - Inter-S3DF VLAN routing will be performed using the virtual ipv4 address (Assigned to all of the 4 cores)
 - Based on VARP
 - Allows multiple switches to simultaneously route packets from a common IP address in an active-active router configuration
 - Unlike VRRP, VARP avoids extra hops by allowing *local routing*, reducing latency and simplifying the path.
 - Performance
 - 1600Gb backbone across cores
 - Leaf switches connected with 400Gb to each core (2x400G)

Network hardware



Arista 7280DR3-24

7280DR3-24 series: 24 port 400G QSFP-DD

- 9.6 Tbps wire speed performance with 16GB buffer
- Supports 24 x 400G QSFP-DD or up to 96 x 100G



*Arista 7050DX4-32S and 7050DX4M-32S:
32x 400G QSFP-DD ports, 2 SFP+ ports*

Transceivers and cables

FS.COM, Arista:

400G QSFP-DD to 4 x 100G QSFP56 Passive
Direct Attach Copper Breakout Cable



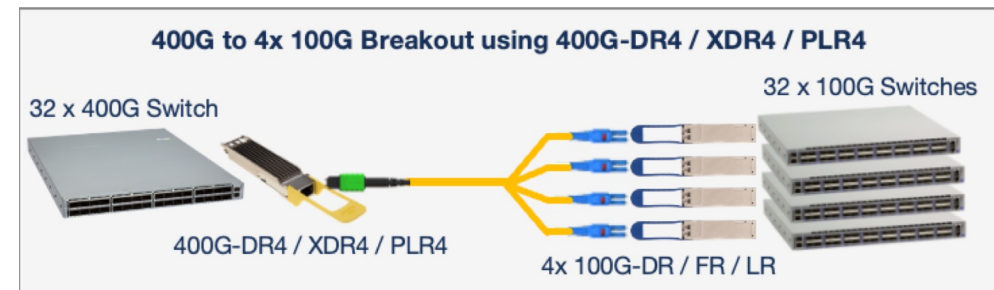
Compatible Cisco to Arista:

Cisco: QSFP28-100G-DR
Arista: QSFP-DD-400G-DR4



Spine to Leaf:

Arista Networks QDD-400G-FR4
Compatible QSFP-DD 400GBASE-FR4 PAM4
1310nm 2km DOM Duplex LC/UPC SMF
Optical Transceiver Module



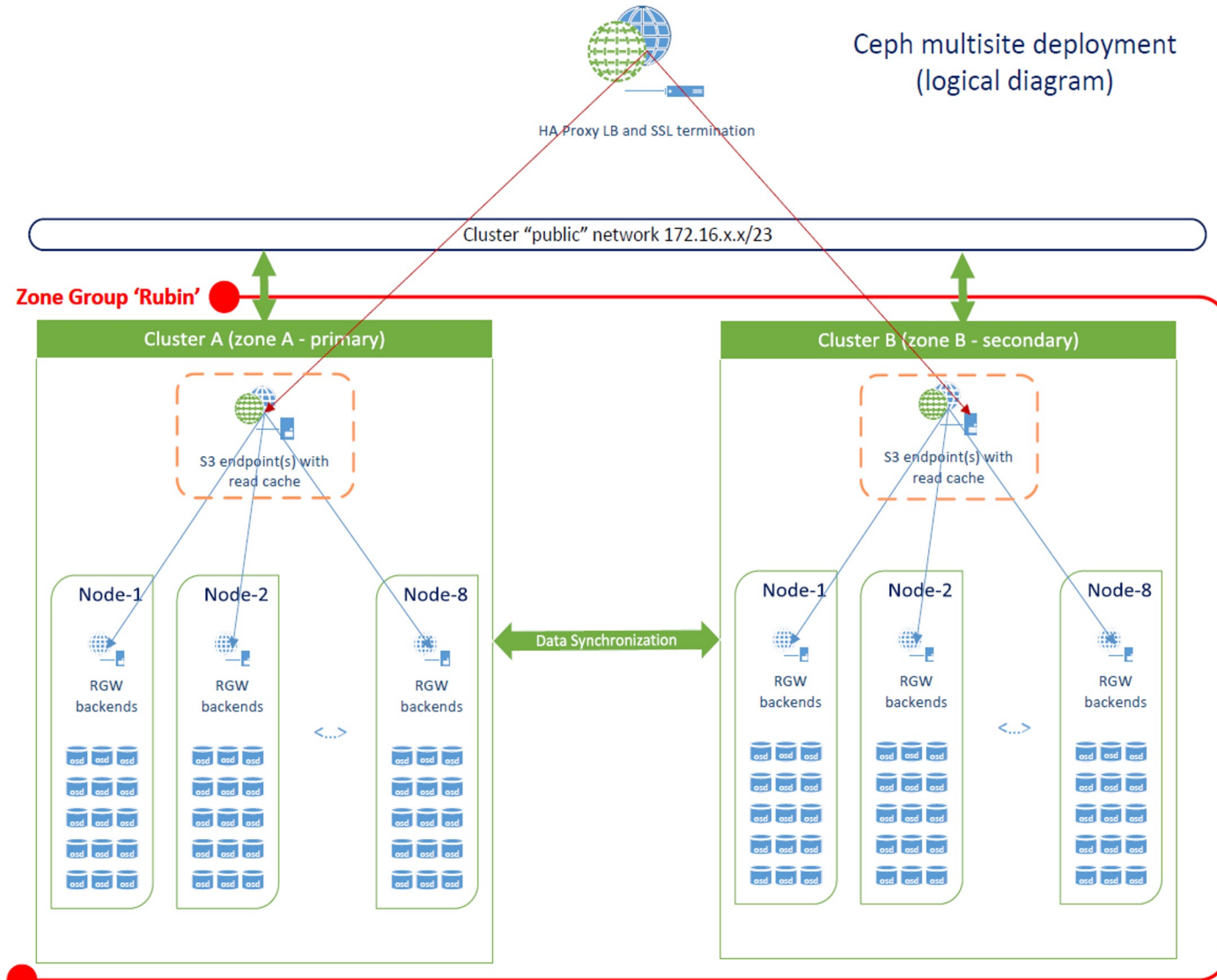
S3DF Network Performance

1.6Tb/s max capacity between SRCF-I and SRCF-II

Highest Bow-Tie traffic reached on Feb 11, 2026 ~900Gb/s



Embargo Layout



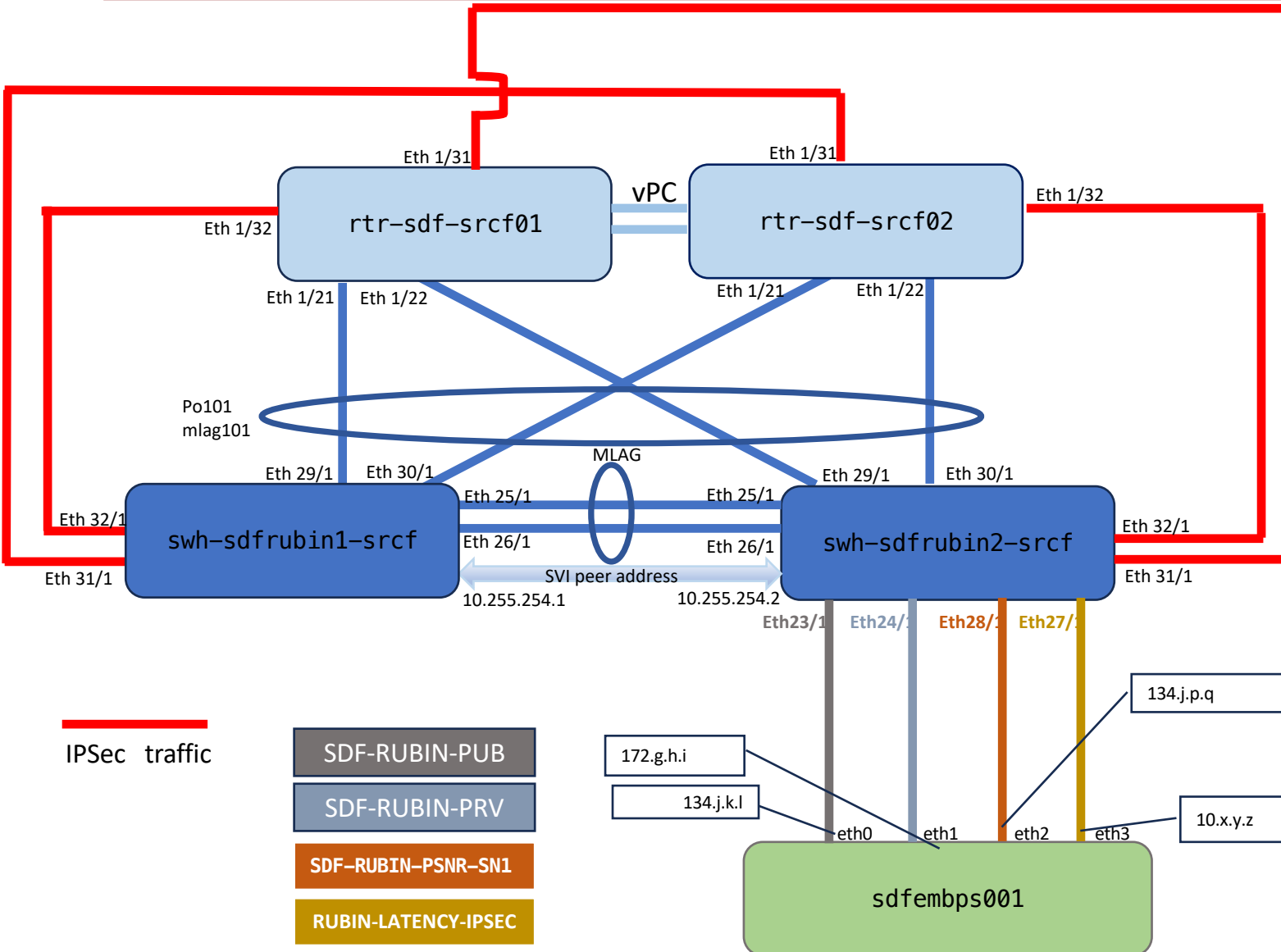
- Requirements
 - Resiliency (Maximize Uptime)
 - Steady State Capacity: 800 TiB
 - Encryption
 - Performance: 7s Latency
- Delivered
 - Usable 1.8 PiB

	T1		T2	
42		42		42
41	sdfemb101	41	sdfemb201	41
40	sdfemb102	40	sdfemb202	40
39	sdfemb103	39	sdfemb203	39
38	sdfemb104	38	sdfemb204	38
37	sdfemb105	37	sdfemb205	37
36	sdfemb106	36	sdfemb206	36
35	sdfemb107	35	sdfemb207	35
34	sdfemb108	34	sdfemb208	34
33		33		33
32		32		32
31		31		31
30		30		30
29		29		29
28		28		28
27		27		27
26	sdfembs3[101-102]	26	sdfembs3[201-202]	26
25		25		25

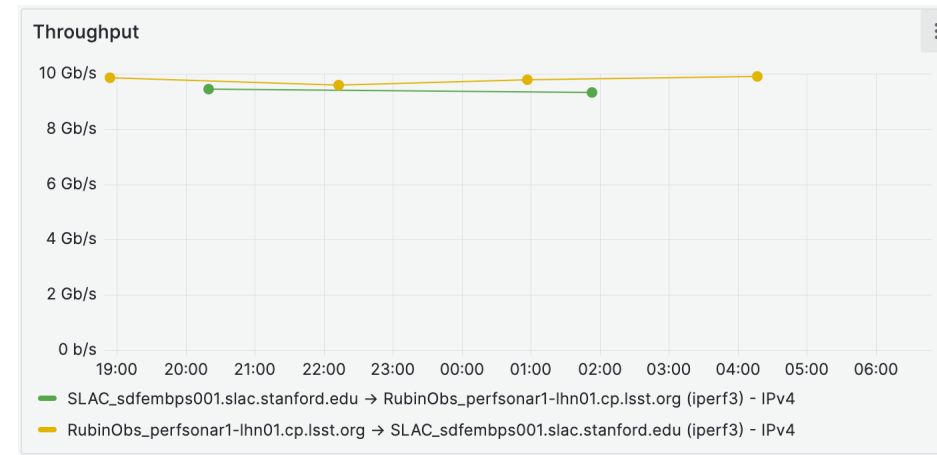
Embargo network – Layer 2/3

- 2x Arista DCS-7280CR3MK
 - 32x 100G ports
 - 4x 400G ports
- VRRP
- 30 servers connected
 - Two redundant zones for Ceph S3 storage
- 1 perfsonar node
 - 4 interfaces
 - Outside tunnel Throughput
 - Outside tunnel Latency
 - Inside tunnel Throughput
 - Inside Tunnel Latency

Embargo network layout



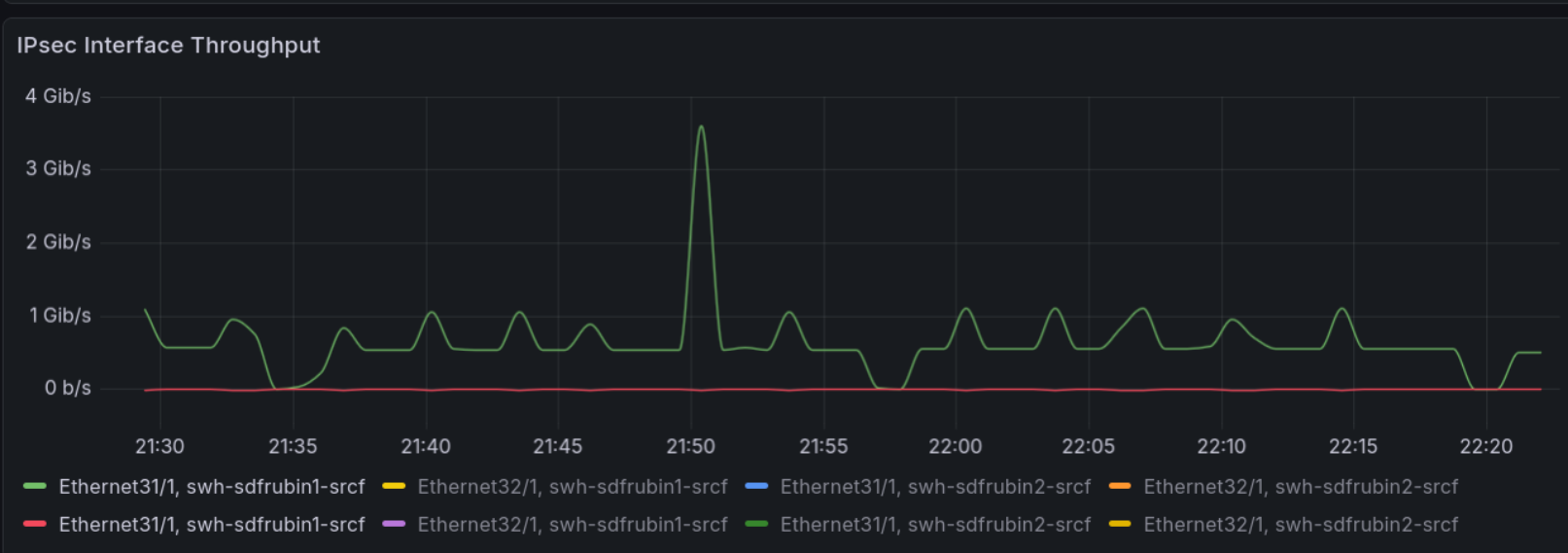
perfSonar IPsec throughput test



Embargo network monitoring



Rubin First Alerts Day – Feb 24th



Security - Network ACL

IP Access List rubin-tunnel101

counters per-entry

10 permit ip host **139.x.y.z** host **134.a.b.d** [match 13242001111 packets, 0:00:00 ago]

11 permit ip host **134.a.b.c** host **134.a.b.d** [match 10229 packets, 0:05:03 ago]

30 permit ahp host 139.x.y.z host 134.a.b.d

40 permit esp host 139.x.y.z host 134.a.b.d

50 deny ahp any any

60 deny esp any any

70 deny ip any any [match 1066731 packets, 0:00:00 ago]

(implicit) deny ip any any

S3DF Network evolution

- Just implemented:
 - New hardware accelerated wire-speed NAT
 - 2x Arista DCS-7050CX3
 - MLAG All Active Multipath (layer 2)
 - Redundant HA setup
 - This will allow better and improved performance of all network connections exiting SLAC from the compute nodes
- Longer term:
 - Next generation network for the whole data center 800/1600Gbit port speed
 - 800Gb/s 802.3df
 - 1.6Tb/s 802.3dj (July 2026)
 - Investigating Overlay networking, MP-BGP eVPN VxLAN
- Cybersecurity:
 - We are constantly working with SLAC Cyber to be compliant with requirements and DoE mandates.

