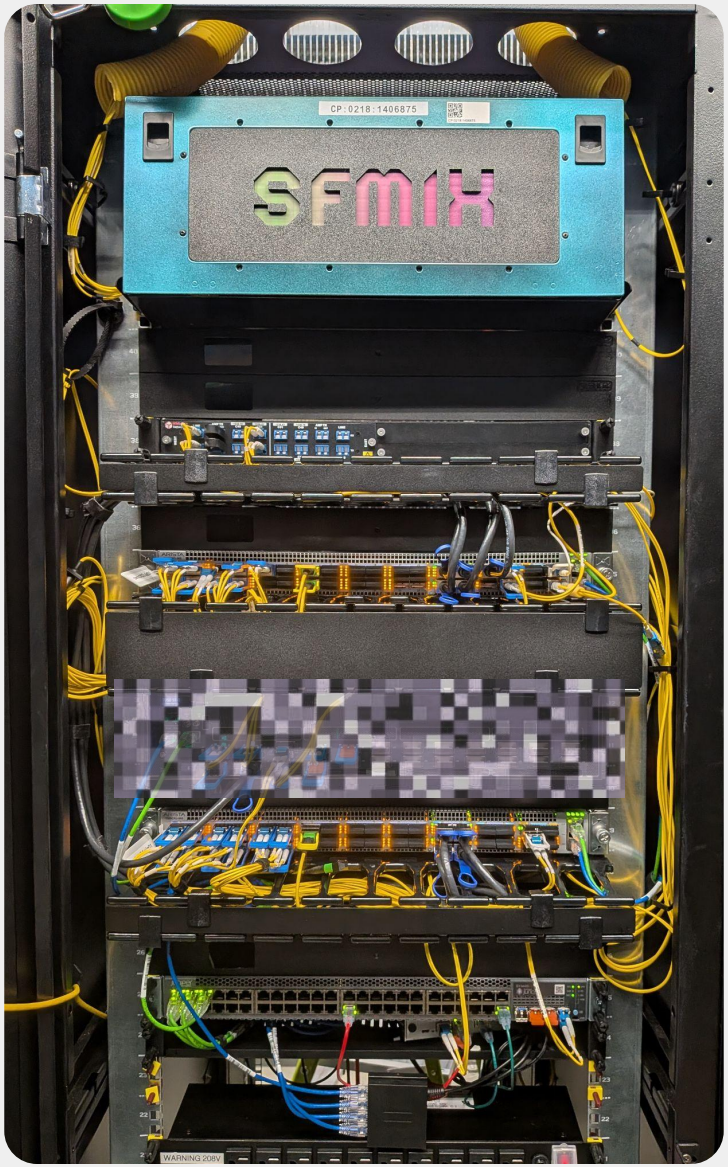


Mutual Aid

L3 Services in a L2 World

Euro-IX

Vienna, Austria
May 5th, 2026

Highlights

ESTABLISHED

2006

Non-profit co-operative

DATACENTERS

9 Today

+2 By End of Year

SCALE

325G Peak

92 Participants

Infrastructure

NETWORKING

VXLAN: Arista 7280-36S & -48C6

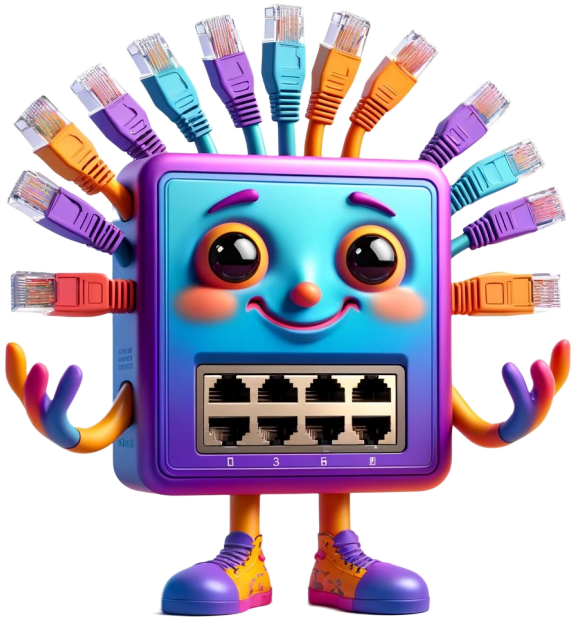
INTER-SWITCH LINKS

Nx 400G or 2x 100G BiDi, ring design
Paths range from <1 to 85 km

Operations

- Almost all volunteers, inc. board
- Matt = ¼ time General Manager
- “Self-serve”; mantra to expose telemetry & debug capabilities to all participants

Mutual Aid



… is a form of voluntary reciprocal exchange of resources and services for mutual benefit, often within a community

Organization Structures

- Independent regional NetOps community, minimal to zero legal entity
- Datacenter and/or MMR operator, value add
- Legally established non-profit/co-op
- Franchise IX’s
- IX-as-a-{*ancillary*}-service

Route Servers

- Limited / if any filtering
- Max. accepted prefixes
  - *Static → PeeringDB API*
- Blackhole sink
- IRR filtering
  - “Authenticated” IRR’s
- MANRS adoption
  - Known IRR and/or RPKI
- Web interface w/ annotations “reason(s)”
- Published RS filter policies, ARouteServer
- Tag prefixes w/ geographical identifier

Learned from Colo: scl02 (2972 Stender Way) (63055:1984:5)

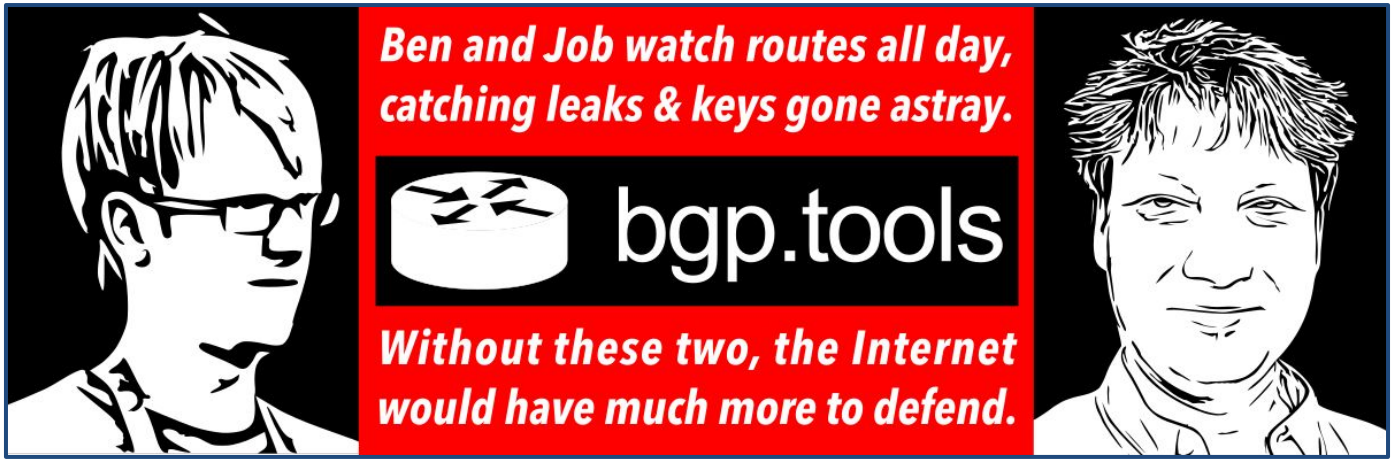
Participant Type: Exempt (63055:1900:2)

Origin ASN is included in client's AS-SET (63055:64512:21)

Value Add Infrastructure

{DNS} Caches

- PCH, ISC, ICANN, Netnod, RIPE, etc
- Versign, Affilias, etc



Troubleshooting tools

- RouteViews, NLNOG RING, bgp.tools, etc

Summary: IX's already support
***more than* forwarding L2 ethernet frames**

Avg. 2025 🇺🇸 🇨🇦 *{non-commercial}* IX

- Robust L2 fabric w/ VXLAN+EVPN, no STP
- Multi-site fabrics *(campus and/or region)*
- Cross-x $$$ rising, more value desired
- ISP specific caches can be tweaked to serve multiple AS's, facilicated over an IX
  - BYO caches (MCC), no traffic requirements
- Cloud & content steady; largest growth from regional/multi-state ISP "access networks"

SFMIX 2024 Situation

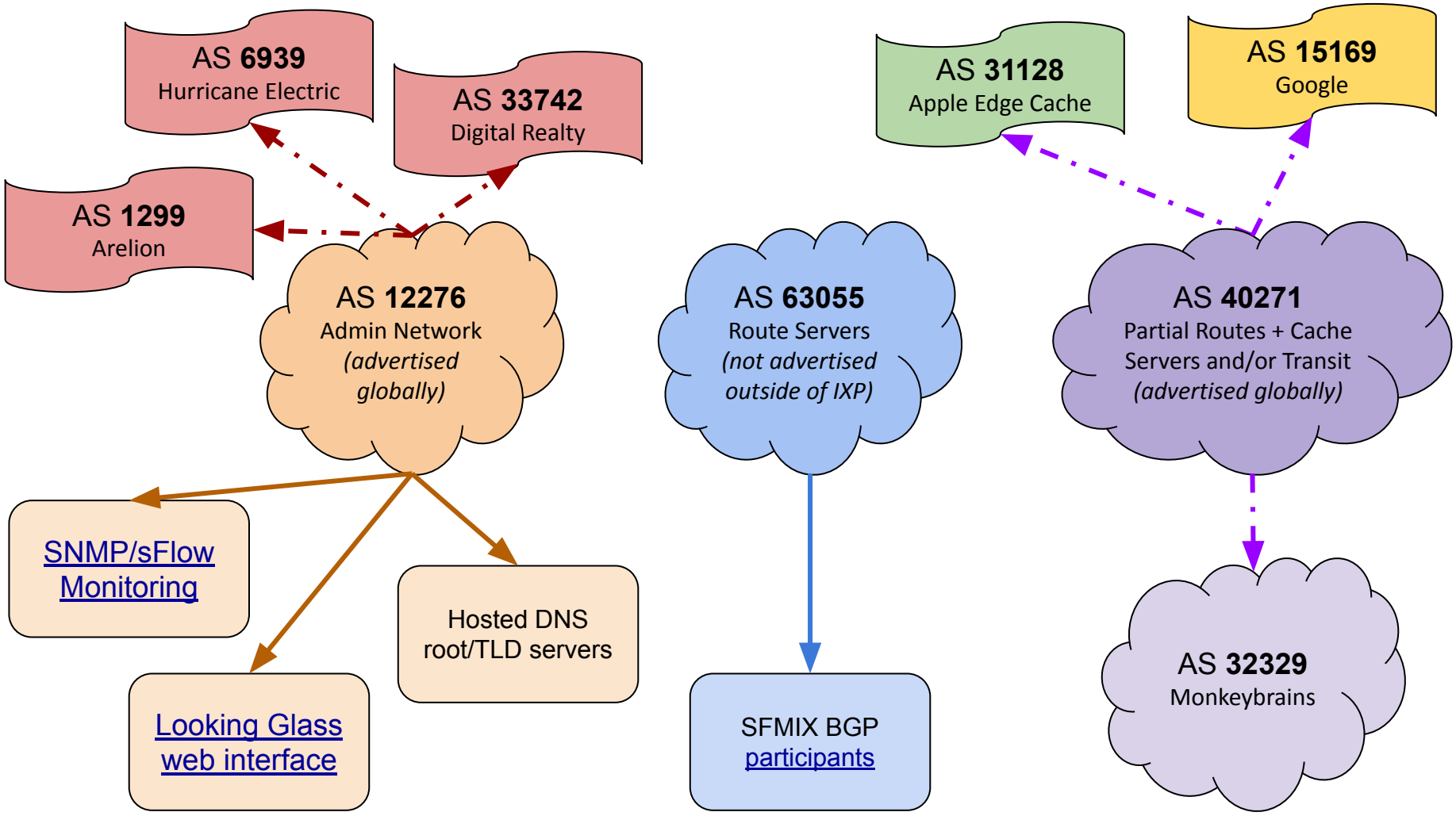
Content moves to {off IX} peering and/or caches

- Regional ISP's lost major content peer over IX
  - ISP's triggered undisclosed de-peer criteria
  - Content de-peer RS's a year prior
  - Peer is valuable to ISP, vocal customers
  - Content offers PNI, no mutual location
  - ISP escalates to IX for help
- IX not geared to transport PNI's
  - No SLA, tech support overhead
  - Compete against membership
- Content peer did offer PNI to the IX, *if an sourced with an ASN not on the IX…*

2024 Proof of Concept

- Spin up “aggregator” ASN, get some v4/v6 PI
- Recycle existing Arista, limitations found
  - EOL Juniper MX150’s *(rib-groups w/ BGP communities for selective VRF leaking)*
- Private peering w/ Google VPP *(Oct ‘24)*, then added caches:
  - Apple, Netskrt, Valve, Microsoft

ASN Usage


AS 6939
Hurricane Electric
AS 33742
Digital Realty
AS 1299
Arelion
AS 12276
Admin Network
(advertised globally)
SNMP/sFlow Monitoring
Hosted DNS root/TLD servers
Looking Glass web interface
AS 63055
Route Servers
(not advertised outside of IXP)
SFMIX BGP participants
AS 31128
Apple Edge Cache
AS 15169
Google
AS 40271
Partial Routes + Cache Servers and/or Transit
(advertised globally)
AS 32329
Monkeybrains

Disperse VRF leaking

Content “🍏 cache”:

- Receive opted-in-ISP routes
- Not interested in other caches
- Static default route *(no DFZ)*

Content “B”

- Wants cache(s) + full table?
- No {duplicate} private peering

ISP “A”

- Wants cache(s) + private peering

ISP “C”

- Wants private peering
- No cache(s)

AS 40271 {Large} Communities

Informational (RX)		Traffic Engineering (TX)	
`40271:1500:1`	Internal Routes from AS 40271	`40271:2000:15169`	Don't send this route to AS15169 (Google)
`40271:1902:1`	RPKI OV State Unknown	`40271:2001:5`	Don't send this to caches *(select peers by type)*
`40271:1984:1`	Route learned in Site ID 1 *(200 Paul Ave, SF)*		

POC Results

- Value {*for subset of IX users*} is real:
  - Single ISP filled 20Gbps day 1
  - Convention center: *“This solves so many problems, been waiting for this..”*
  - ISP: *“We already trust you, the IX is more reliable then our transits, do it..”*
- Board feedback
  - Limit blast radius & admin overhead

Jan 2026 Production Deployment

- 2x Nokia 7750 SR-1-24D *(datasheet)*
  - 2RU; 24x 800G QSFP [+|56-DD|112]
  - VRF leaking by BGP communities

- Validated interop configs for current *(VXLAN + static flood)* & future *(EVPN)* w/ our existing Arista platform - *before sale*

- Existing IX cross-x w/ VLAN tag or new cable, new VNI defined across backbone

Operational model

- AS 40271 “partial routes + caches” is opt-in
  - Importantly, only ISP’s which consume are underwriting the service *(router investment, cross-x’s, power, transit to feed caches, etc)* - **not** entire membership
  - Survey *(response rate ~17% of total IX)* interest, commits cover CapEx in ~2 yrs
  - $300/MRC for 10Gbps of capacity
- 95th percentile 85Gbps w/ 5 ISP’s *(April ‘26)*

{Technical} **Concerns**

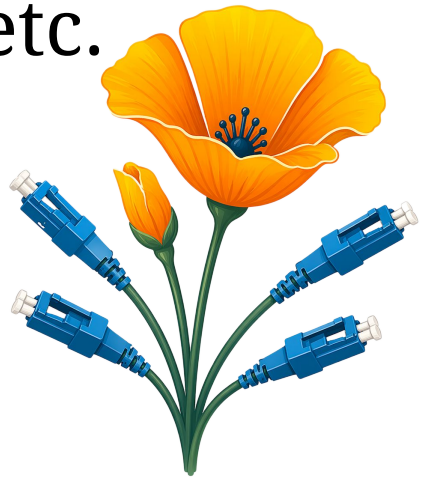
- Arista L2 sub-interface mode caveats
  - 😒 storm-control limited to 7280R3
  - BPDU guard seems to work
  - Remember to enable L2 hardware counters

- SR-OS : IPFIX w/ IE315 or proprietary UDP stream
  - *tbd* sFlow-RT pipeline Prometheus + Grafana

- Optics Compatibility
  - No 4x 100G LR4 optics - only 4x LR1 or 2x LR4

{Operational} **Realities**

- >3 months from PO to arrival
- Monthly / quarterly billing
  - “Traditional” IX LAN annual, low admin time
- Does limited service become majority headache
  - Education on as-path, local-pref, etc.
  - End-user DNS w/ ECS
- Branding / service != traditional IX



matt@sfmix.org

thank you

special thanks

Sam, Saju, Amer, Greg, & all of the Nokia Cloud Majors team

Jof's technical prowness from rapid prototype to production grade reality

AS 40271 Provisioning steps

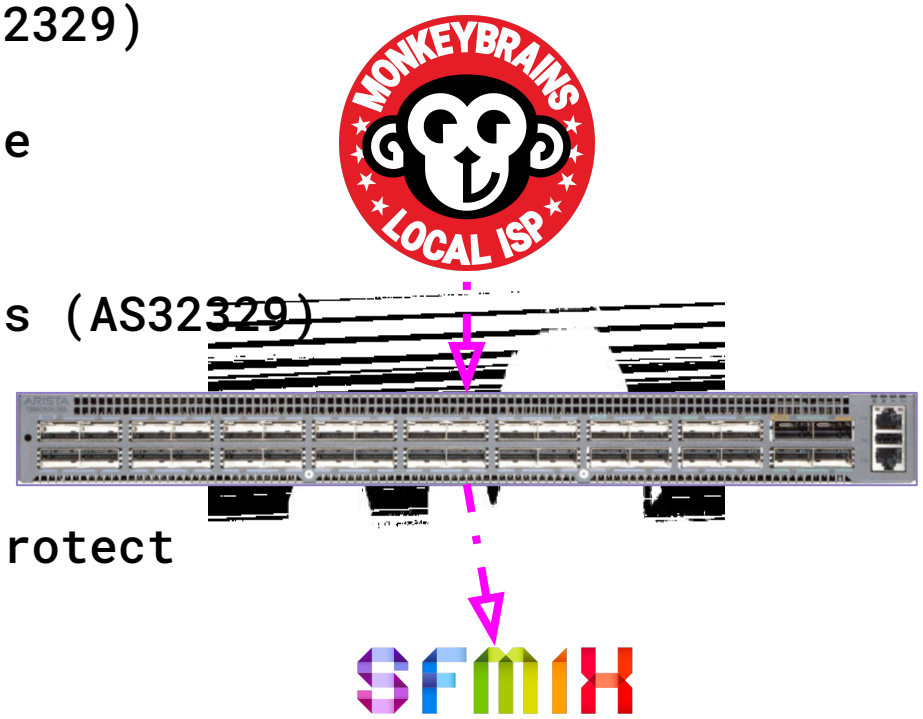
- Allocate VLAN / VNI, /31 & /127 in NetBox
- Push new VNI mesh wide, w/ Vxlan1 intf
- Update member port to include VLAN tag
- Define VPRN /31 & /127 w/ SAP VLAN id
- Query PeeringDB for as-set
- Spawn bgpq4 to generate IRR prefix-lists
- Define member BGP sessions and AS specific communities
- Update “AS-SFMIX-TRANSIT” ARIN as-set
- Validation of sessions & route filtering
- *Profit? Debug? Cry?*

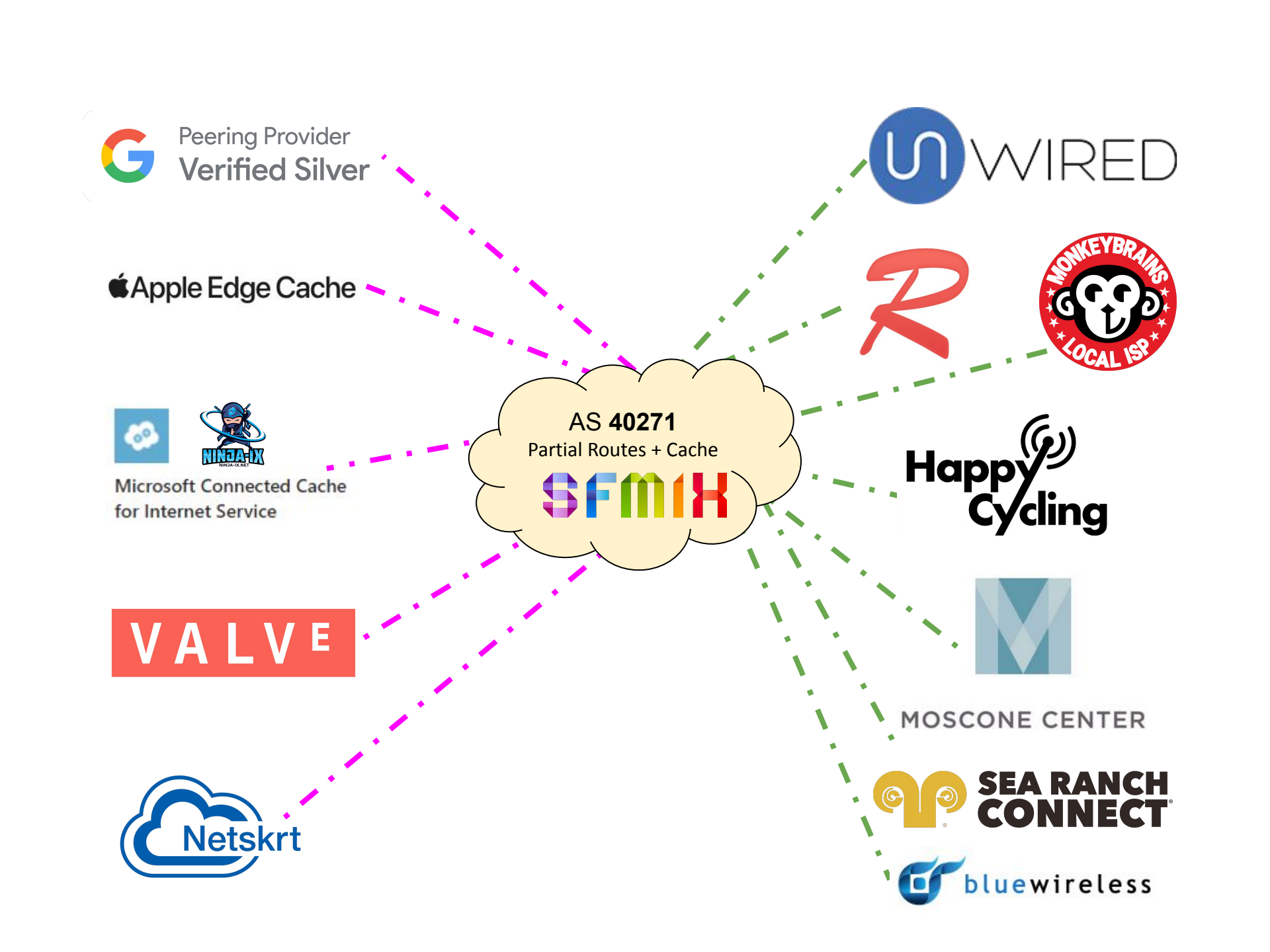
EOS config {participant}

```
interface Ethernet8/1
   !! CP:0218:1406875 Port 127 (Slot K - MPO B - Ports 13 & 14)
   !! SN:382948 SO:1347674139 Cust-CID:SFM-32329-02
   description Monkeybrains (AS32329)
   channel-group 106 mode on
   spanning-tree bpduguard enable

interface Port-Channel106.998
   description Peer: Monkeybrains (AS32329)
   vlan id 998
   encapsulation vlan
     client untagged
   mac address limit violation protect
   mac address limit 2

interface Port-Channel106.1421
   description Partial: Monkeybrains (AS32329)
   encapsulation dot1q vlan 1421
   vlan id 1421
```




Peering Provider
Verified Silver
Apple Edge Cache
Microsoft Connected Cache
for Internet Service
NINJA-IX
VALVE
Netskrt
AS 40271
Partial Routes + Cache
SFMIX
UNWIRED
MONKEYBRAINS
LOCAL ISP
Happy Cycling
MOSCONE CENTER
SEA RANCH CONNECT
bluewireless