WAN’s (R)Evolution

robert@raszuk.net
Credit’s

- Main inspiration (and some content) came from Geoff Huston - Chief Scientist @ APNIC

- SD WAN information, running code, global network etc .. donated by Sproute Networks Inc.
- Some SR related drawings come from Kris Michielsen & Clarence Filsfils @ Cisco
- Experience and future extensions of TCP Analyzer is a work with Prof. Alejandro Popovksy

... everyone who I worked with for nearly 25 years in global networking industry
A little bit of history - roots of networking

Circuit switching - yes that how it all started ...

Then came tone dialing and digital phone exchanges setting up p2p connections aka circuits between two phones (end points).
You must be thinking now ...

what is he talking about - it is so boring !!!
A little bit of history - roots of networking

Well now imagine that those were not humans, but computers on both ends trying to communicate with each other using packets.
A little bit of history - roots of networking

Then came routers (aka gateways) and telco circuits as well as satellite links to build Internet ... still the fundamental goal was to connect everyone on Earth - for free!
Protocols and the way they help to run networks ...
Telco’s p2p circuits and links ...

Physical & data link WAN layers:

- Copper cross connects
- Digital containers (SDH/SONET)
- Frame Relay
- ATM
- Satellite
- Dark fiber

- Optical channels (\(\lambda\) aka wavelength)
- Wireless
- Ethernet

*Winner of the day: Ethernet (de-facto today’s standard)*
Protocols...

IGPs (Interior Gateway Protocols)
IS-IS (ISO10589:2002, RFC1195)
OSPFv2 (RFC2328)
OSPFv3 (RFC5340)

BGP (Border Gateway Protocol)
BGPv4 - RFC4271
(with RFC 4760 MP-BGP)

Failure detection
LOS (use on optical interfaces)
BFD (RFC5880..5)

"MPLS TRANSPORT" vs "MPLS+[BGP] APPLICATIONS"
- LDP (RFC5036)
- Labeled BGP (RFC3107)
- RSVP-TE (RFC3209)
- L3VPNs (RFC4364)
- L2VPNs (RFC6624)
- VPLS (RFC4761)
- EVPN (RFC7432)

TUNNELING vs ENCAPSULATION
IDENTIFIER - LOCATOR SPLIT
(ILNP RFC6740 vs LISP RFC6830)
See also IRTF RRG (RFC6227, RFC6115)

TRAFFIC ENGINEERING
- BGP policies (interdomain)
- RSVP-TE (mainly intradomain)
Segment Routing (same administration)

CONVERGENCE vs PROTECTION
FAST CONNECTIVITY RESTORATION
FRR: LFA (RFC5286) vs TI-LFA (rtgwg draft)

Failure detection
LOS (use on optical interfaces)
BFD (RFC5880..5)
Overhead ...
TE - current approach (data plane based) - BEFORE

Cars:
- SFO-LAX
- SAN-SMF
- LAX-SFO
- SMF-SAN

No Traffic Engineering analogy to Human Drivers

SRC - Feb’2000: [https://www.nanoq.org/meetings/nanog18/presentations/raszuk.ppt](https://www.nanoq.org/meetings/nanog18/presentations/raszuk.ppt)
TE - current approach (data plane based) - AFTER

Cars:
- SFO-LAX
- SAN-SMF
- LAX-SFO
- SMF-SAN

Traffic Engineering analogy to Auto Pilot inserted into each car!

SRC - Feb'2000: https://www.nanog.org/meetings/nanog18/presentations/raszuk.ppt
Vector Routing - a control plane approach 0 overhead

TE with Segment Routing

segment-routing
traffic-eng
policy GREEN
color 30 end-point ipv4 1.1.1.4
candidate-paths
  preference 100
dynamic
metric
type delay

[REF: www.segment-routing.net]
Resilience with Segment Routing via TI-LFA

route isis 1
   interface Gi0/0/0/0
   address-family ipv4 unicast
       fast-reroute per-prefix
       fast-reroute per-prefix ti-lfa

router ospf 1
   area 0
       interface GigabitEthernet0/0/0/0
           fast-reroute per-prefix
           fast-reroute per-prefix ti-lfa enable
Load balancing & resilience via SR Anycast SID
How to interconnect enterprise sites ...
WAN Physical Topologies - circuits between sites

!!! Warning - Caution !!!

A lot of circuits sold these days are emulated circuits over IP share backbones
L3VPNs - evolution or customer lock?

- $$$ expensive 300-600 USD per 1M/month
- locked to single SP
- up to 3 months of provisioning time per site
- not extendable to public clouds
- not extendable to mobile users
- no real guarantees of anything
- same for L2VPNs customers still fully responsible of their own routing
- no e2e dynamic path selection
SD-WAN innovation

- very affordable
- secure
- web GUI+API provisioning
- new site provisioning in seconds
- integrated with both public and private clouds
- seamless remote access
- both software and hardware based
- true zero touch provisioning (ZTP)
- e2e SLA based path monitoring and selection
- application aware policies
- automated sites mgmt
SD-WAN innovation - example Sproute Networks

Highly recommended → If you want to try it for free email them with discount code: RR@47JAIIO

SRC: http://www.sproute.com/
Back to bigger picture

Internet vs Services/Content
Internet tiers (traditional model)
But where users go and what they do on the net?

They (incl. enterprises) are accessing content and services!
How content/service providers change Internet?

Regional ISP

Tier 1’s Transit

Google, Facebook, Amazon, Microsoft, Apple, Tencent ...
Global Private Networks

Regional ISP

Local ISP

Local ISP

BUE IX

DE-CIX

https://www.peeringdb.com/search?q=CABASE shows 25 IXes in Argentina
Where is the content? In major global DCs?
Google - 15

**Americas**
- Berkeley County, South Carolina
- Council Bluffs, Iowa
- Douglas County, Georgia
- Jackson County, Alabama
- Lenoir, North Carolina
- Mayes County, Oklahoma
- Montgomery County, Tennessee
- Quilicura, Chile
- The Dalles, Oregon

**Asia**
- Changhua County, Taiwan
- Singapore

**Europe**
- Dublin, Ireland
- Eemshaven, Netherlands
- Hamina, Finland
- St Ghislain, Belgium
Role Reversal

Service portals are increasingly located adjacent to users

And that means changes to the network:

- Public (Global) Networks no longer carry users’ traffic to/from service portals via ISP carriage services
- Instead, Private Networks carry content to service portals via CDN services and their internal replications

This shift has some profound implications for the Internet
Who is building real transit now?

Almost all new submarine international cable projects are heavily underwritten by content providers, not carriers.

Microsoft and Facebook’s 160Tbps transatlantic undersea cable carries more data than any other.

SoftBank, Facebook, and Amazon commit to 8,700-mile transpacific subsea cable system.
Content providers becoming Internet owners ...
Content & Service really is the King of the Internet

- None of these seven technology companies are a telephone company, or even a transit ISP, or even an ISP at all!

- All of them have pushed aside carriage networks in order to maintain direct relationships with billions of consumers

- These valuable consumer relationships are based on content services, not carriage
Suggestion for Palermo Analyzer ...

predict the future

- Let’s add as a default graph showing the ratio of total traffic going from/to Argentina to top 7 content/service providers

- Imagine the effects on your country top transit ISP’s traffic or revenue when any of the 7 providers open a local POP in Argentina or direct most traffic to other POPs (ex: Google’s DC in Quilicura/Chile ?

- And perhaps this is just a matter of time.
Competition or Cartel

At some point in the past decade or so the dominant position across the entire Internet has been occupied by a very small number of players who are moving far faster than the regulatory measures that were intended to curb the worst excesses of market dominance by a small clique of actors.

These actors have enough market influence to set their own rules of engagement with:

- Users
- Each other
- Third party suppliers/vendors
- Regulators and Governments
Competition or Cartel

At some point in the past decade or so the dominant position across the entire Internet has been occupied by a very small number of players moving far faster than the regulatory measures were intended to curb the worst excesses of market dominance by a small clique of actors.

These actors have enough market influence to set their own rules of engagement with:

- Users
- Each other
- Third parties/vendors
- Regulators and Governments

Is this the Internet we were dreaming of / hoped for?
Few not so well known networking features …
More different BGP paths is good for you!


DIVERSE PATH - RFC 6774

ADD PATHS - RFC 7911

- Add-Path will signal diverse paths from 2 to X paths
- Required all Add-Path receiver BGP router to support Add-Path capability.
BGP Optimal Route Reflection

Problem statement

- RRs as control plane only platforms – departure from classic POP to Core location due to end to end encapsulation in networks and emerging Internet free core
- Suboptimal best/2nd best path selection for clients – difficult to ensure hot potato routing
- Position of control plane RRs should not play any role in path selection for clients.

REF: draft-ietf-idr-bgp-optimal-route-reflection
eiBGP load balancing

For L3VPNs in the VRF context

For multiple CEs in the global RIB**

** Cisco only
Smart Edge Routing

Google’s Espresso Project

Facebook’s Edge Fabric

CISCO Performance Routing (PFR)
http://docwiki.cisco.com/wiki/PFR:Technology_Overview

Nuage Intelligent Traffic Steering

Hopefully soon Univ. of Palermo SER Controller
MP-TCP

Data transmission with plain TCP

Alice

3G (10.11.12.13)

Plain TCP connexion

data

WiFi (10.11.12.14)

One link is available but unused

Bob

Eth (20.21.22.23)

Data transmission with MPTCP

Alice

3G (10.11.12.13)

Plain TCP connexion (subflow 1)

data

WiFi (10.11.12.14)

Plain TCP connexion (subflow 2)

Bob

Eth (20.21.22.23)

All the available links are used

Transmission of "Hi, Hello world" using MPTCP
(on port 80 for each subflow)

Legend and concepts

drum, dack* = sequence number and acknowledgement of the data of the whole MPTCP session (snagym and snagym are related to the data transmitted in each subflow)
token = number used to identify a MPTCP session (TCP subflows would be initialized with the same token)

* represents the next expected byte (that will be indicated by DACK field)

Bandwidth aggreg.
TCP Resilience
Handover
Transparency
Questions, comments, discussion welcome!