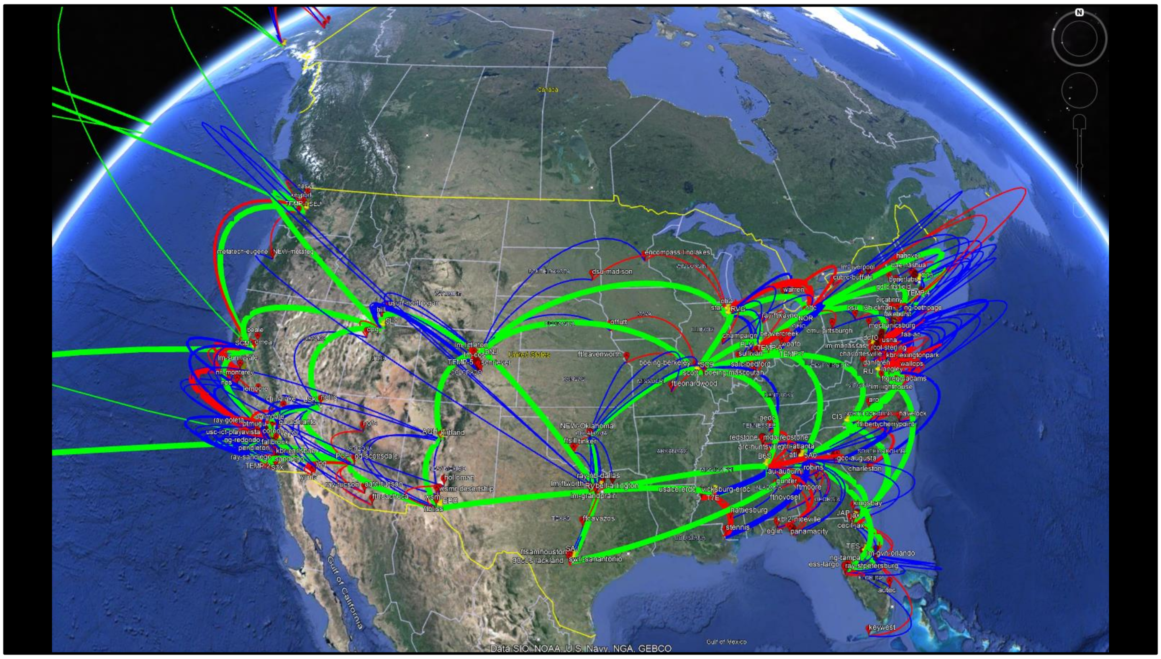
Measurement Infrastructure on the Defense Research and Engineering Network (DREN)

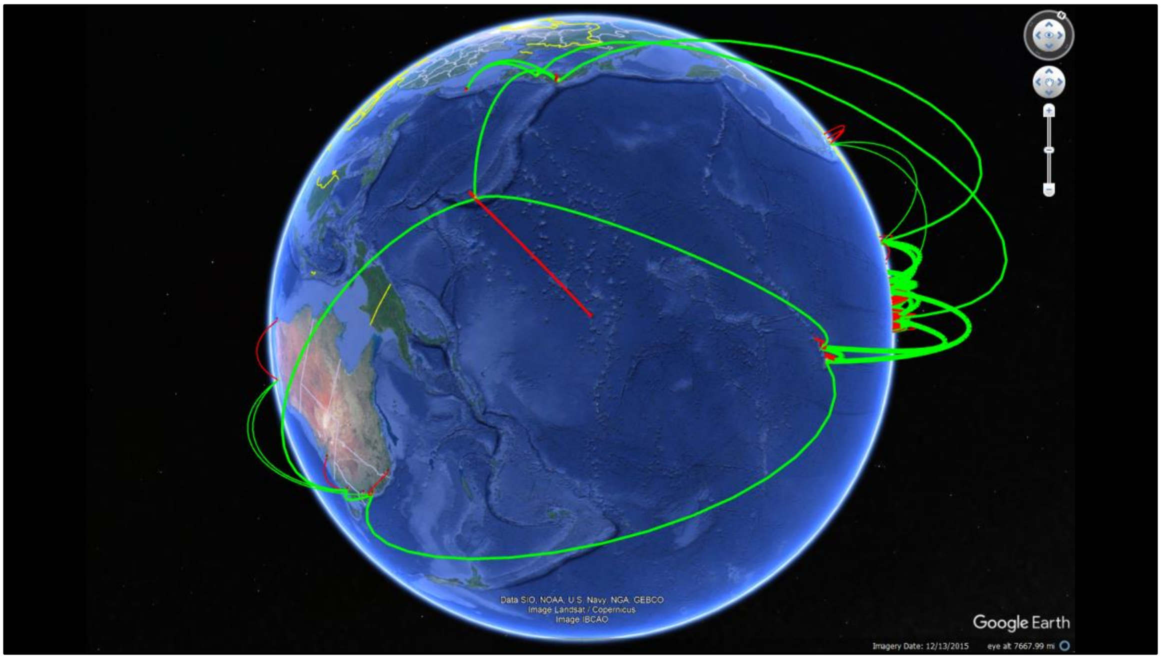
Phil Dykstra

phil@pdykstra.com

CAIDA Workshop, 24 June 2024



A visualization of DREN in Google Earth. Green are core circuits. Red and Blue are primary and secondary access circuits connecting each site to the core. The length of each arc is the equivalent length of fiber to produce the measured round trip time. A perfect point to point fiber would not be raised above the surface of the earth.

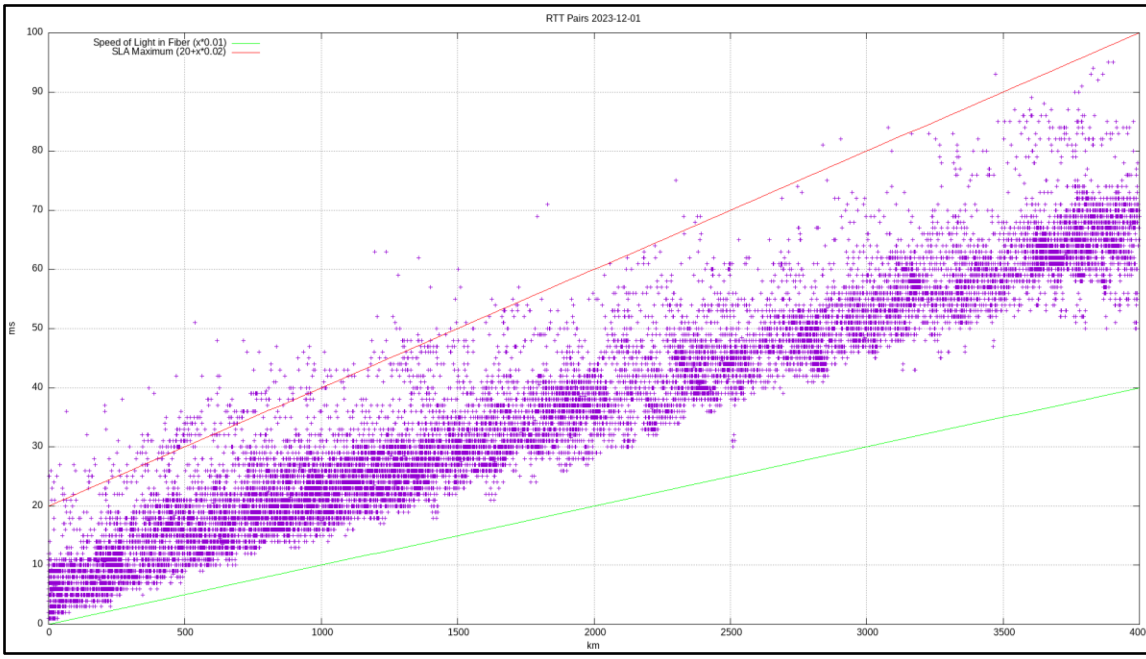


DREN has been expanding into the Pacific. Australia is planned but not up yet. In general, trans oceanic circuits have less “excess” latency because they are relatively straight lines.

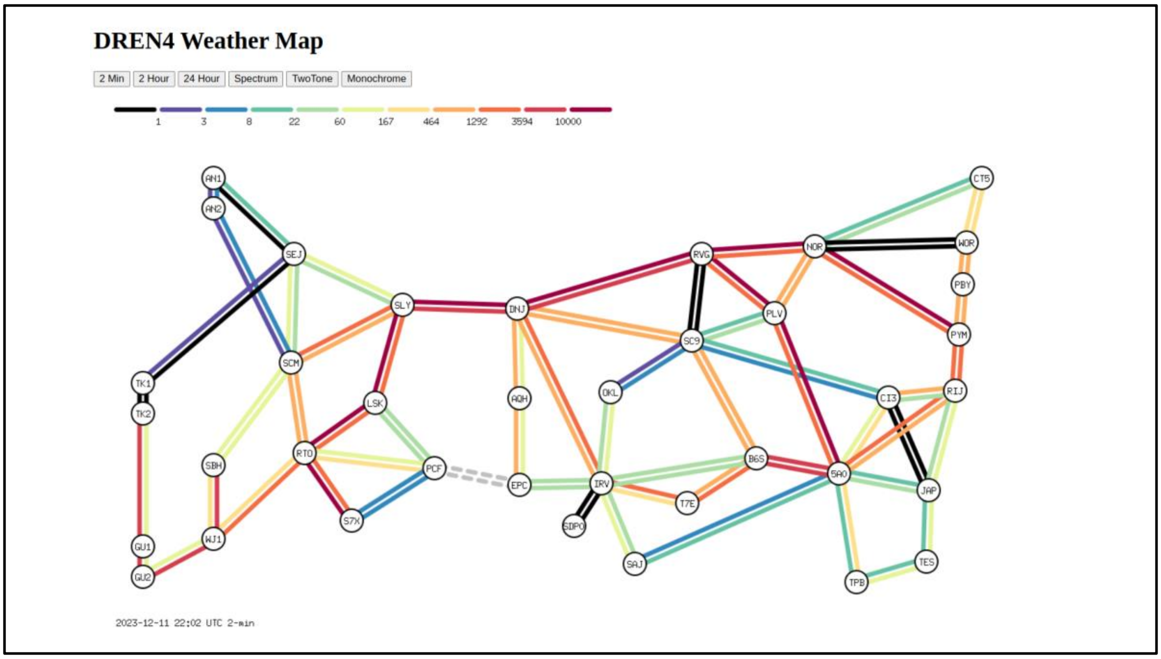
Remember Mathis?

$$Throughput \propto \frac{MTU}{RTT \times \sqrt{loss}}$$

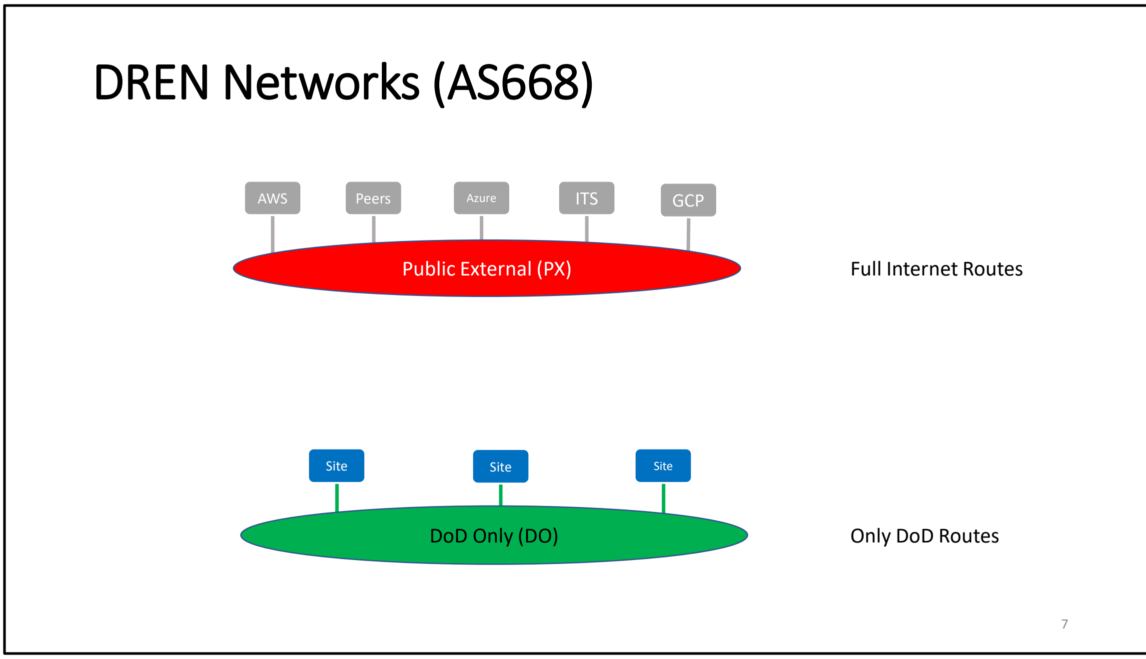
It has proven hard to raise MTU over 1500 or 9000 if you are lucky. We generally throw bandwidth at loss to make it approach zero. This leaves RTT as the performance parameter we have some control over. Engineer to minimize latency!



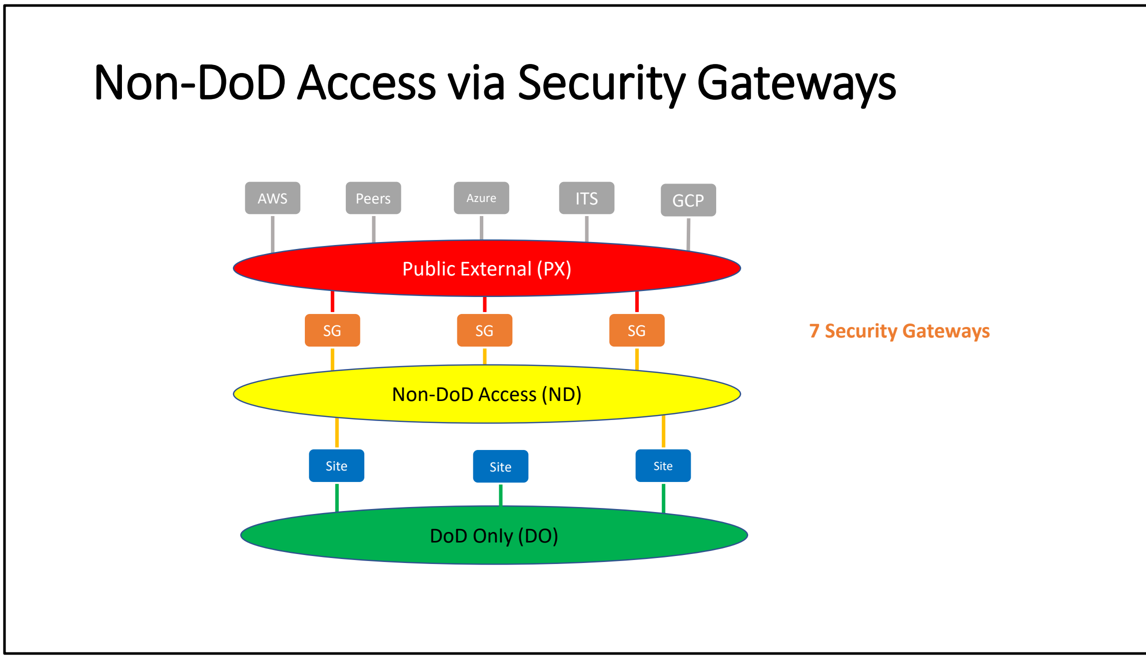
All pairs of DREN Nodes showing the physical distance between them (x) and the measured RTT between them (y). The green line slope is the delay caused by the speed of light in fiber. The red line is twice the light-in-fiber delay plus 20ms. We engineer DREN to keep RTTs below this threshold. When this data was collected roughly 1% of all ~30000 node pairs exceeded this limit, most often due to high latency on access circuits.



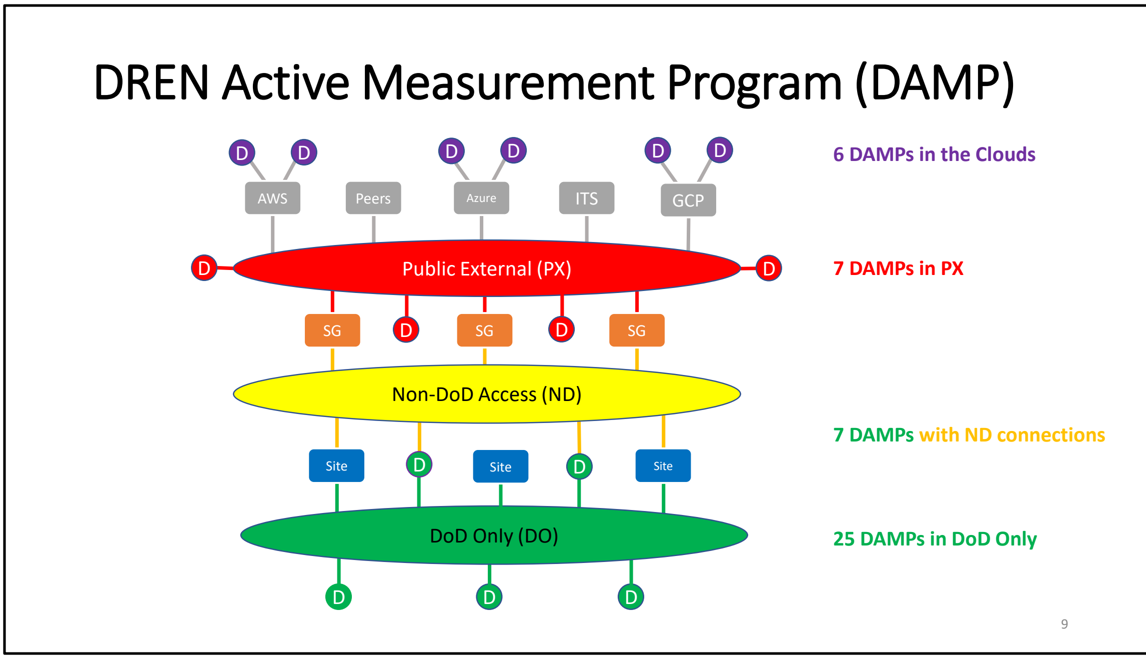
A weather map of the core circuits. PCF to EPC was down at this time leaving us within one circuit failure of partitioning the network!



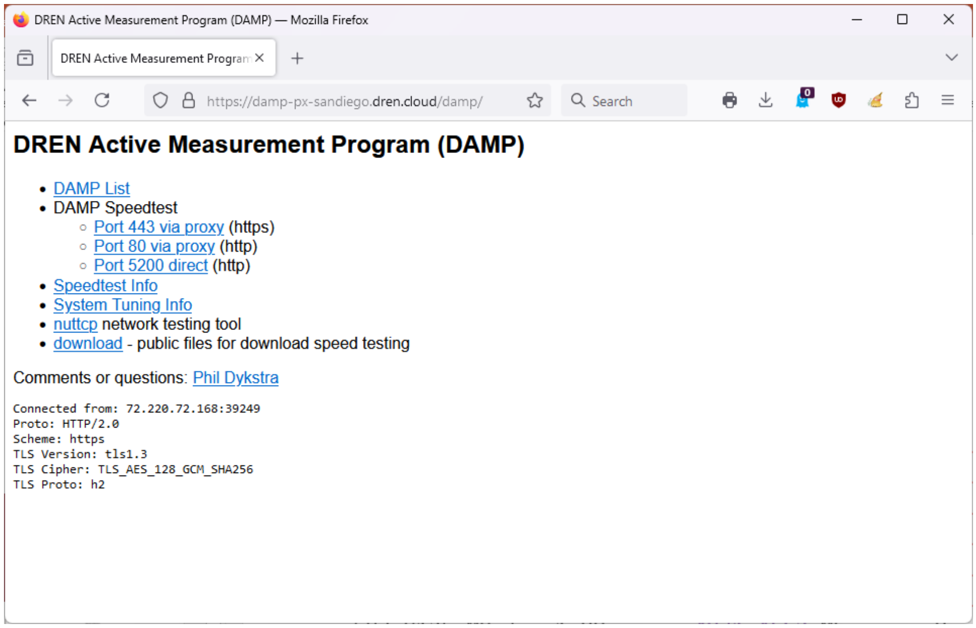
There are roughly 200 sites on DREN connected to the “DoD Only” (DO) network. There are no routes to/from the internet on this network. DREN’s peering and transit happens on a separate Public External (PX) network with multiple sets of full internet routes.



If a site needs to access the Internet, they get a SECOND interface to the Non-DoD access network (ND). This ND network connects to Public External via seven Security Gateways. Careful routing ensures that a Security Gateway will see both directions of any connection. The split networks (DO+ND) at the sites allow the implement of different security policy / firewalls on each connection. From a security perspective PX sees constant scans and attacks, ND is much more quiet, and DO sees nearly zero malicious activity.

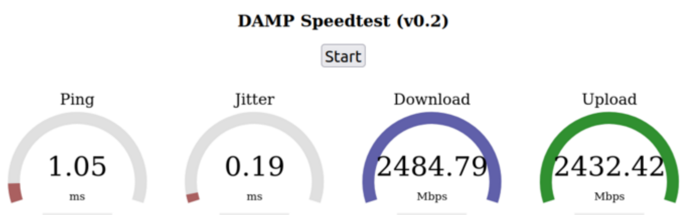


DAMP systems are Linux servers running performance testing tools. For many years, these were only on the DO network so could not be accessed / used by anyone outside of DREN. Recently we began placing DAMPs in the Public External Network and in Cloud provider networks. These are open for researchers to use, and for end users to test their performance.



What you see when you connect to a Public External or Cloud DAMP (with /damp/ in the URL). We try to make them self documenting and offer both simple and advanced performance testing tools.

Web based speed test



- Server side is a single Go program (multi platform, no files)
- Client side is JavaScript
- Might be a good place for WebAssembly

A simple web based speed test. Can also be used via curl (see on DAMP docs). With curl on an LAN it runs up to 8 Gbps. In a browser (Chrome) up to 2.5 Gbps.

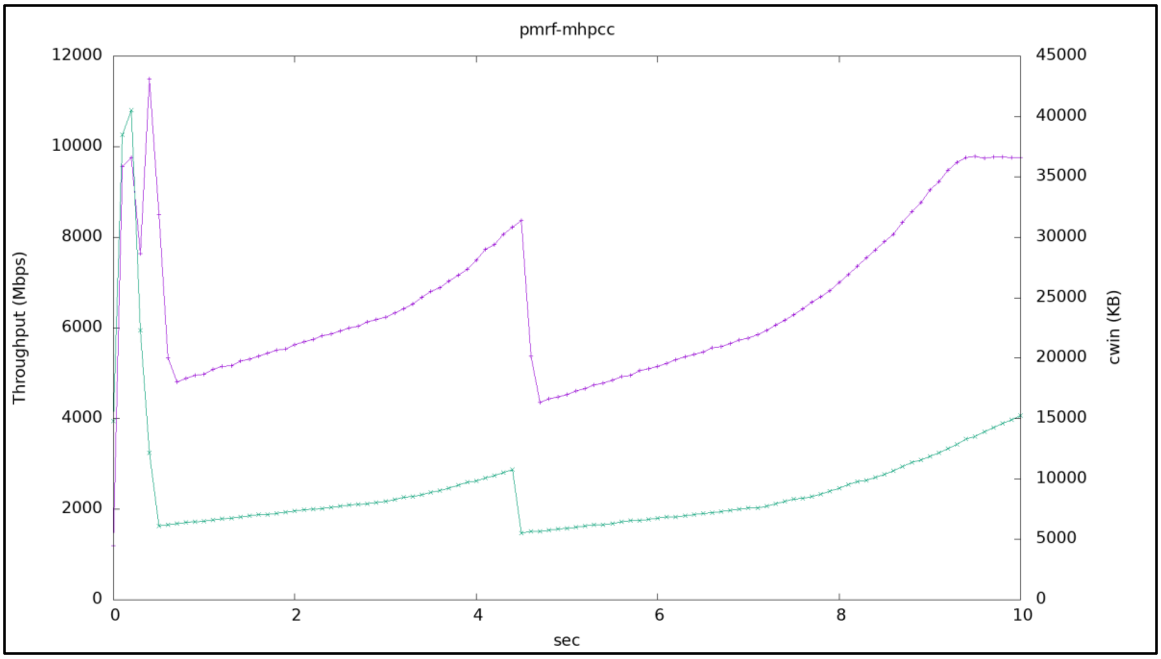
nuttcp

A screen log of running nuttcp, a TCP and UDP tool similar to iperf. Take away is it is real time, interruptible and supports third party operation without accounts. This allows rapid debugging. I find it far more productive than iperf and batch based testing services.

nuttcp throughput tests

```
$ nuttcp -i1 damp-pmrf.dren.mil damp-mhpcc.dren.mil
  997.5625 MB /   1.00 sec = 8367.9516 Mbps    55 retrans  12040 KB-cwnd
  834.4375 MB /   1.00 sec = 6999.1602 Mbps     2 retrans   6608 KB-cwnd
  645.9375 MB /   1.00 sec = 5418.6845 Mbps     0 retrans   7402 KB-cwnd
  721.1875 MB /   1.00 sec = 6050.1283 Mbps     0 retrans   8213 KB-cwnd
  831.7500 MB /   1.00 sec = 6977.3224 Mbps     0 retrans   9982 KB-cwnd
  749.1250 MB /   1.00 sec = 6283.4876 Mbps     1 retrans   6042 KB-cwnd
  587.8125 MB /   1.00 sec = 4931.4662 Mbps     0 retrans   6774 KB-cwnd
  662.1250 MB /   1.00 sec = 5554.1349 Mbps     0 retrans   7611 KB-cwnd
  768.8750 MB /   1.00 sec = 6449.8555 Mbps     0 retrans   9390 KB-cwnd
  969.3750 MB /   1.00 sec = 8131.8858 Mbps     0 retrans  11997 KB-cwnd
 7887.4121 MB /  10.12 sec = 6539.0847 Mbps 37 %TX 47 %RX 58 retrans
12328 KB-cwnd 10.57 msRTT
```

Example run with per-second output. Purple (throughput) and green (TCP congestion window) are plotted on the next slide.



A 10 second nuttcp test plotted. I would also like to get the dynamic receive window on these plots.

Cloud Connectivity, Throughput (Mbps)

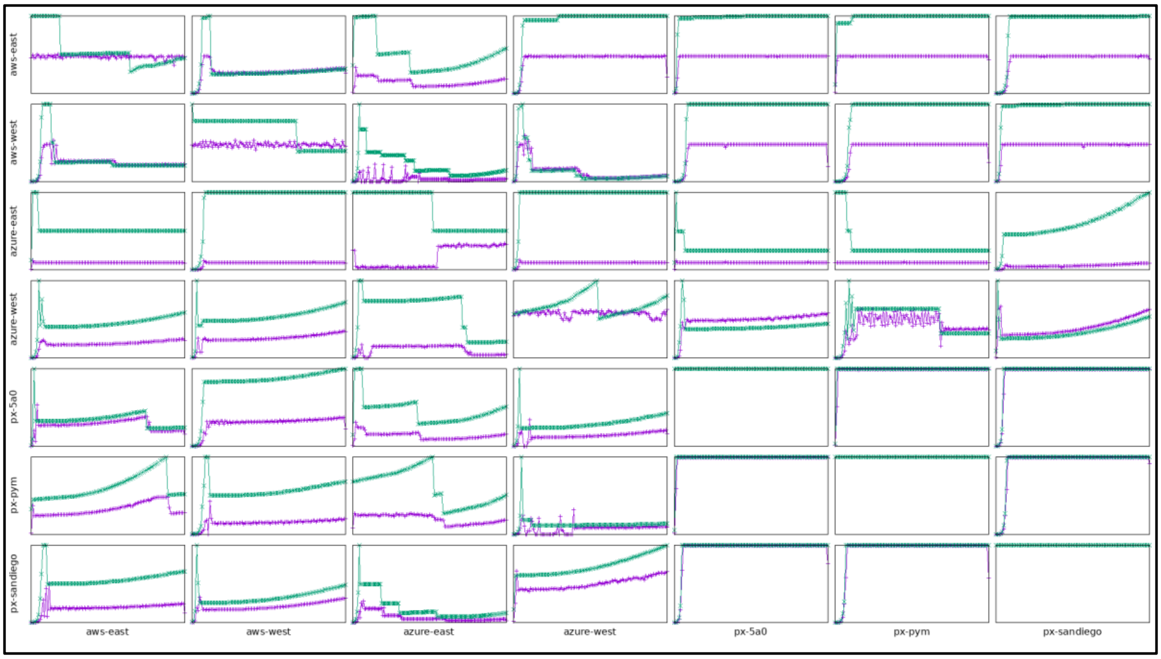
src\dst	aws-east	aws-west	azure-ea	azure-we	px-5a0	px-pym	px-sandi	TOTAL
aws-east	4736	2761	1461	4482	4651	4706	4366	22427
aws-west	2426	4766	313	1266	4363	4299	4600	17266
azure-ea	945	881	1638	920	945	952	525	5168
azure-we	1828	2513	1085	5747	4811	4137	3964	18337
px-5a0	2718	2965	1340	1361	20862	9740	9358	27482
px-pym	3206	1489	1983	740	9748	20878	9115	26280
px-sandi	1913	2062	667	4968	9357	9150	20936	28116
TOTAL	13036	12671	6849	13735	33875	32984	31928	145077

An example matrix of tests. Such test sets are great at identifying whether problems are transmitters, receivers, or both.

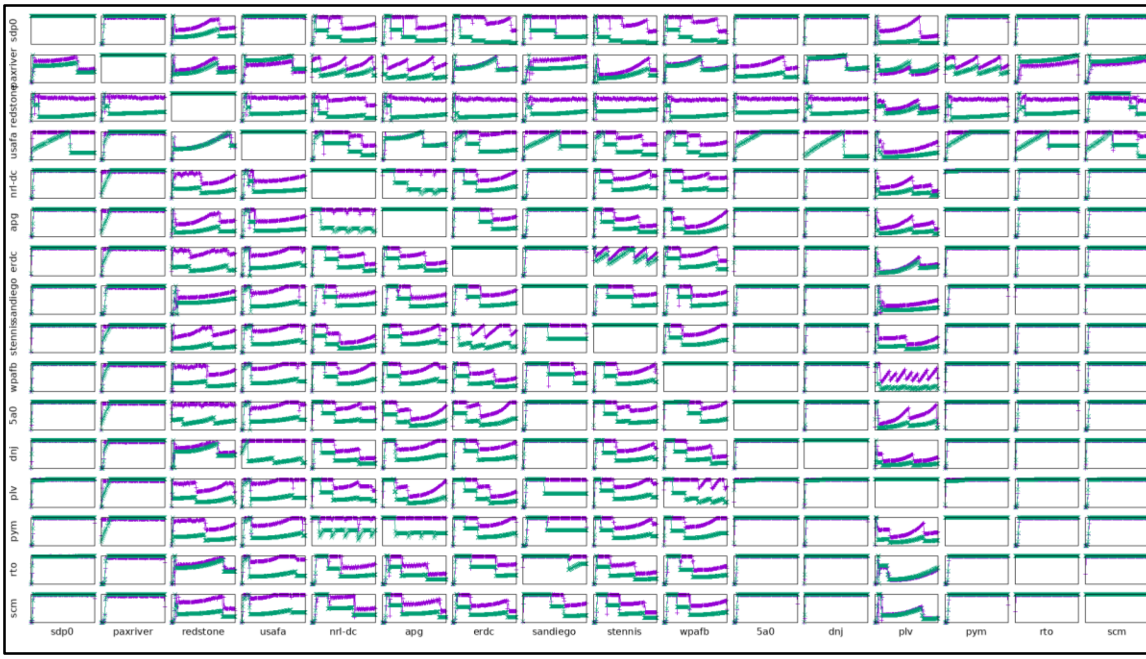
Cloud Connectivity, Retransmits (packets)

src\dst	aws-east	aws-west	azure-ea	azure-we	px-5a0	px-pym	px-sandi	TOTAL
aws-east	194	208	3804	0	2	2	2	4018
aws-west	562	564	181504	5	13	5	0	182089
azure-ea	72	0	168	0	18	78	0	168
azure-we	20887	28200	25192	3354	668	1108	1250	77305
px-5a0	11620	11544	3065	39237	0	0	0	65466
px-pym	1995	8610	3986	154485	2	0	0	169078
px-sandi	13761	4959	12975	5492	0	0	0	37187
TOTAL	48897	53521	230526	199219	703	1193	1252	535311

Even more than throughput, TCP retransmits often make problems stand out.



More details in a matrix. You start to see patterns.



Even more (256 TCP tests). You start to recognize specific test patterns after a while.
Singular network tests are always suspect: is it repeatable, is my test host bad, etc.
When everyone starts pointing a finger at you, it's probably you.

Takeaways

- Latency matters
- People time matters
- **nuttcp** is great
- Get more than one opinion

https://damp-px-sandiego.dren.cloud/damp/
phil@pdykstra.com