ARCHIPELAGO
Internet Measurement Infrastructure
http://www.caida.org/projects/ark/

caida
www.caida.org

Why should my network host an Ark Node?

When your network hosts a measurement node that participates in CAIDA's Archipelago (Ark) infrastructure, it broadens the view of the global Internet for the network research community. Network researchers use CAIDA topology data to conceive, develop, and test their models and methods. Participating networks agree to our Memorandum of Cooperation (MoC) [1] and install a Raspberry Pi, 1-U server, virtual server, or software container dedicated to Ark measurement.

Once deployed, the Ark node conducts continuous measurements of the routed IPv4 (and IPv6 when the hosting network supports it) address space. Ark aggregates the resulting data on a server at UC San Diego's Supercomputer Center. Each additional node contributes to – and increases the completeness and accuracy of -- data representing the topological structure of the Internet core.

CAIDA uses these continuous measurements, as well as sophisticated Internet-scale alias resolution methods developed in-house, to build the Internet Topology Data Kit (ITDK) – a heavily annotated router-level graph of the Internet to support data science on Internet topology. CAIDA further annotates each router with its inferred geolocation, and inferred operator, to support sophisticated analyses of the router-level Internet topology by the Internet research community.

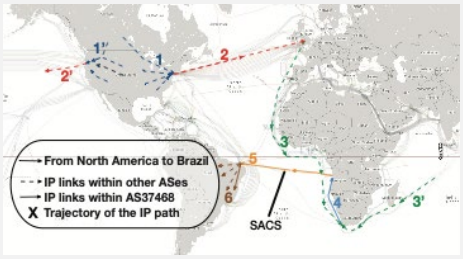
As of 2024, the Ark measurement platform has supported the Internet research community for more than 15 years. Recent work has investigated the unintended consequences of submarine cable deployment on Internet routing (PAM'18), persistent interdomain congestion (SIGCOMM'18), the impact on performance and resilience of regional access network topology structure (IMC'21), and methods to automatically extract meaning from Internet router hostnames (CoNEXT'21, IMC'20, IMC'19). As of 2024, CAIDA is enhancing the infrastructure with a domain-specific language to allow researchers to quickly and correctly build and execute experiments.

This range of scientific experiments has successfully demonstrated our vision of a metaphorical distributed measurement "operating system" to support empirical Internet science.

[1] https://www.caida.org/projects/ark/moc/

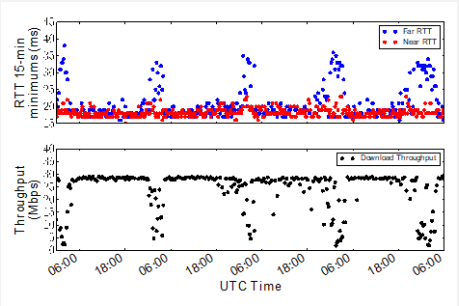
If interested, please contact
ark-info@caida.org

Figure 1: North America Paths routed through Africa to Brazil



Paths taken from North American Ark nodes to Brazil that transited Africa after deployment of the SACS Brazil-Angola submarine cable., Unexamined routing configurations induced surprising performance impairments, i.e., the cable resulted in *longer* RTTs for these paths.

Figure 2: Internet interdomain congestion



Time series of latency probes (top panel) and throughput measurements (bottom panel). Lightweight latency probes identified interdomain links between networks with evidence of congestion (grey portions). Throughput measurements established the effect of the congestion on performance.

Figure 3: Internet router geolocation

xo.iad02.atlas.cogentco.com
as2828.was14.ip4.gtt.net
te9-2-0d0.cir1.ashburn-va.us.xo.net

vodafone.iad02.atlas.cogentco.com
ae-0.vodafone.asbnva02.us.bb.gin.ntt.net
zayo.vodafone.er2.iad10.us.zip.zayo.com

usqas1-rt002i.i3d.net
interactive.edge1.washington111.level3.net
ce-0-4-0-2.r05.asbnva02.us.ce.gin.ntt.net

level3-as3356.e0-51.switch2.ash1.he.net
ae-1-3510.edge1.washington111.level3.net
level3.ashburn2.ash.seabone.net

Hostnames of 4 routers located in/near Washington D.C., with router interface names assigned by nine operators. Ark enabled development of a technique to automatically learn the naming convention of each operator using the ITDK as a primary input. CAIDA provides the inferred rules, and a public API, as a service to the research community. Note that ash, in red, is an airport in Nashua, NH; the technique learns that the operators used 'ash' to mean Ashburn, VA.

UC San Diego SDSC
San Diego Supercomputer Center, University of California San Diego
9500 Gilman Drive, mc0505, La Jolla, CA 92093-0505 | 858-534-5000 | www.caida.org