HI^3 HUB for Internet Incident Investigation

Alberto Dainotti

alberto@caida.org

Center for Applied Internet Data Analysis
University of California, San Diego

caida
www.caida.org

U.S. DEPARTMENT OF HOMELAND SECURITY

LARGE-SCALE INCIDENTS

a threat to private and national assets

- **Large-scale Internet incidents**
  - *BGP hijacks, connectivity outages, spam and fishing campaigns, botnet activities, large-scale bug exploitation, …*
  - A major threat to public safety and to both public and private strategic and financial assets
  - ***often unnoticed***
  - ***hard to understand*** (dynamics, motivation, infrastructure used, source, target)
    - hard to mitigate, prevent, etc.
    - hard to assess the impact
    - hard to assess restoration

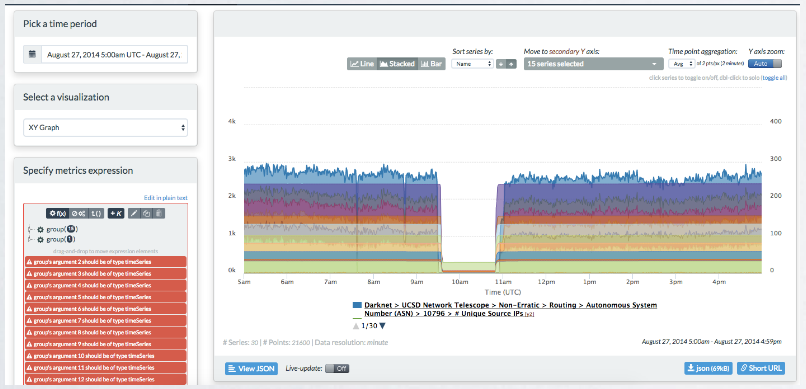
> ***DHS CISA: "there is increased risk for wide scale or high-consequence events that could cause harm or disrupt services upon which our economy and the daily lives of millions of Americans depend.“***

HUB for Internet Incident Investigation

Cybersecurity Analytics PaaS and SaaS

- Distributed infrastructure *(PaaS)* and software frameworks *(SaaS)*
  - for ingestion and correlation of streams of *diverse* cybersecurity data

- Web-based collaborative environment *(SaaS)* with trusted user groups
  - enabling analysis with interactive and visual tools, dashboards, …



Pick a time period

August 27, 2014 5:00am UTC - August 27,

Select a visualization

XY Graph

Specify metrics expression

Edit in plain text

group(15)
group(1)

drag-and-drop to move expression elements

group's argument 2 should be of type timeSeries
group's argument 3 should be of type timeSeries
group's argument 4 should be of type timeSeries
group's argument 5 should be of type timeSeries
group's argument 6 should be of type timeSeries
group's argument 7 should be of type timeSeries
group's argument 8 should be of type timeSeries
group's argument 9 should be of type timeSeries
group's argument 10 should be of type timeSeries
group's argument 11 should be of type timeSeries
group's argument 12 should be of type timeSeries

Sort series by: Name

Move to secondary Y axis: 15 series selected

Time point aggregation: Avg of 2 pts/px (2 minutes)

Y axis zoom: Auto

click series to toggle on/off, dbl-click to solo (toggle all)

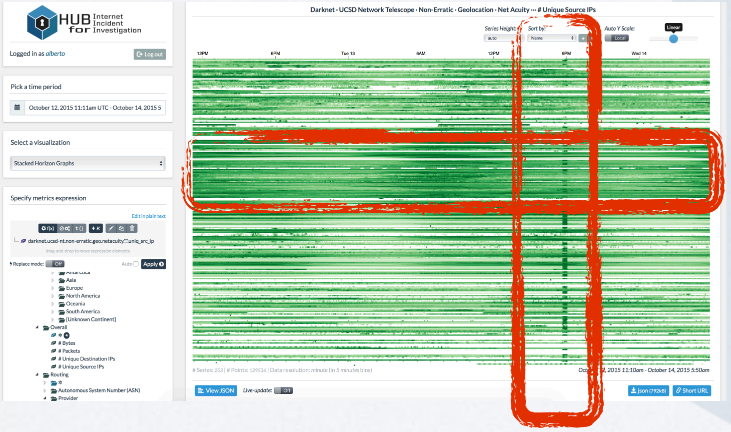
Darknet > UCSD Network Telescope > Non-Erratic > Routing > Autonomous System Number (ASN) > 10796 > # Unique Source IPs [y2]

1/30

Series: 30 | # Points: 21600 | Data resolution: minute

August 27, 2014 5:00am - August 27, 2014 4:59pm

View JSON Live-update: Off json (69kB) Short URL



HUB Internet Incident for Investigation

Logged in as *alberto* Log out

Pick a time period

October 12, 2015 11:11am UTC - October 14, 2015 5

Select a visualization

Stacked Horizon Graphs

Specify metrics expression

Edit in plain text

darknet.ucsd-nt.non-erratic.geo.netacuity.*.*.uniq_src_ip

drag-and-drop to move expression elements

Replace mode: Off Auto Apply

Antarctica
Asia
Europe
North America
Oceania
South America
[Unknown Continent]
Overall
* 4
# Bytes
# Packets
# Unique Destination IPs
# Unique Source IPs
Routing
*
Autonomous System Number (ASN)
Provider

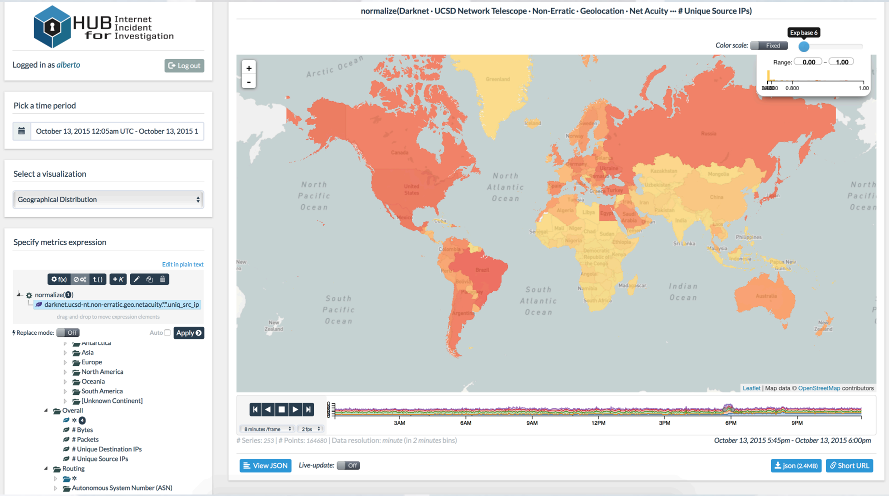
Darknet · UCSD Network Telescope · Non-Erratic · Geolocation · Net Acuity ··· # Unique Source IPs

Series Height: auto Sort by: Name Auto Y Scale: Local Linear

Series: 253 | # Points: 129536 | Data resolution: minute (in 5 minutes bins)

October 12, 2015 11:10am - October 14, 2015 5:50am

View JSON Live-update: Off json (792kB) Short URL



HUB Internet Incident for Investigation

Logged in as *alberto* Log out

Pick a time period

October 13, 2015 12:05am UTC - October 13, 2015 1

Select a visualization

Geographical Distribution

Specify metrics expression

Edit in plain text

normalize(1)
darknet.ucsd-nt.non-erratic.geo.netacuity.*.*.uniq_src_ip

drag-and-drop to move expression elements

Replace mode: Off Auto Apply

Antarctica
Asia
Europe
North America
Oceania
South America
[Unknown Continent]
Overall
* 4
# Bytes
# Packets
# Unique Destination IPs
# Unique Source IPs
Routing
*
Autonomous System Number (ASN)

normalize(Darknet · UCSD Network Telescope · Non-Erratic · Geolocation · Net Acuity ··· # Unique Source IPs)

Color scale: Fixed Exp base 6

Range: 0.00 – 1.00

Leaflet | Map data © OpenStreetMap contributors

Series: 253 | # Points: 164680 | Data resolution: minute (in 2 minutes bins)

October 13, 2015 5:45pm - October 13, 2015 6:00pm

View JSON Live-update: Off json (2.4MB) Short URL

THE HI^3 APPROACH

three key concepts

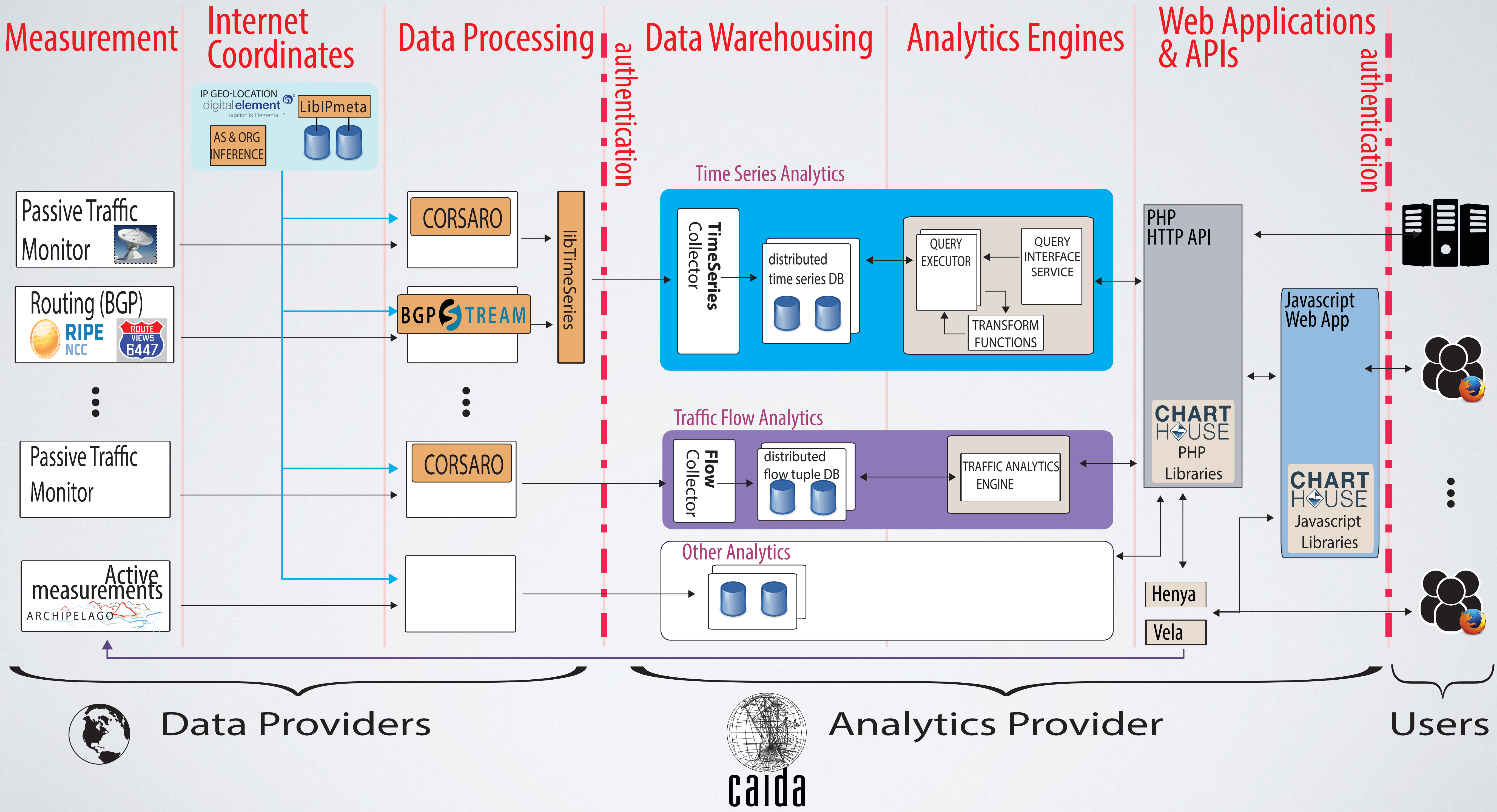
1. **Combination and correlation of diverse** Internet cyber-security data streams..
   - *Unwanted traffic, prefix hijacks, outages, DoS, spam, …*

   ..around a set of common dimensions
   - *Time, geography, and Internet Coordinates (IP, BGP, DNS, …)*

2. Data analytics in the form of interactive **exploratory data analysis** and configurable event **detection**

3. Trusted **realtime collaborative** environment

HI^3 ARCHITECTURE



Measurement

Internet Coordinates

Data Processing

Data Warehousing

Analytics Engines

Web Applications & APIs

authentication

authentication

IP GEO-LOCATION digital element Location is Elemental™

LibIPmeta

AS & ORG INFERENCE

Passive Traffic Monitor

Routing (BGP)

RIPE NCC

ROUTE VIEWS 6447

Passive Traffic Monitor

Active measurements

ARCHIPELAGO

CORSARO

BGP STREAM

CORSARO

libTimeSeries

Time Series Analytics

TimeSeries Collector

distributed time series DB

QUERY EXECUTOR

QUERY INTERFACE SERVICE

TRANSFORM FUNCTIONS

Traffic Flow Analytics

Flow Collector

distributed flow tuple DB

TRAFFIC ANALYTICS ENGINE

Other Analytics

PHP HTTP API

CHART HOUSE PHP Libraries

Javascript Web App

CHART HOUSE Javascript Libraries

Henya

Vela

Data Providers

Analytics Provider

caida

Users

"EXPLORER" INTERFACE

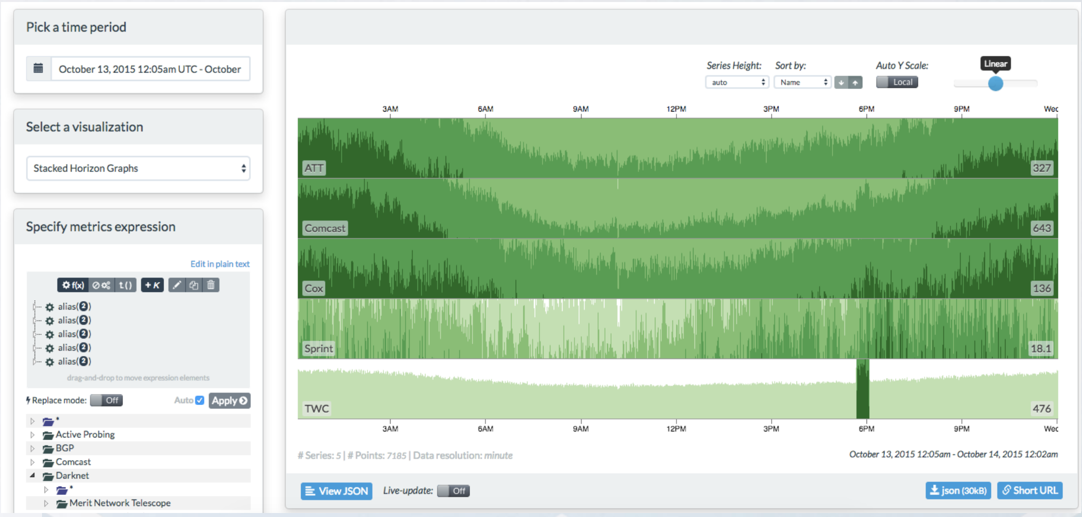
Interactively transform, aggregate, compare, and visualize time series-based Internet data

- **Last year's demo shows how it can be used to investigate a potential massive bug exploitation**

- 13th Oct 2015 - from 5.40pm UTC for about **25 minutes a large spike in source IPs from all continents**

  - but among top 5 **US providers** we see it coming **only from TWC**!

  - backscatter from UPnP machines that received spoofed packets to their **UDP port 1900 (UPnP)** originating from UDP port 80



Pick a time period

October 13, 2015 12:05am UTC - October

Select a visualization

Stacked Horizon Graphs

Specify metrics expression

Edit in plain text

alias(2)
alias(2)
alias(2)
alias(2)
alias(2)

drag-and-drop to move expression elements

Replace mode: Off   Auto ☑   Apply

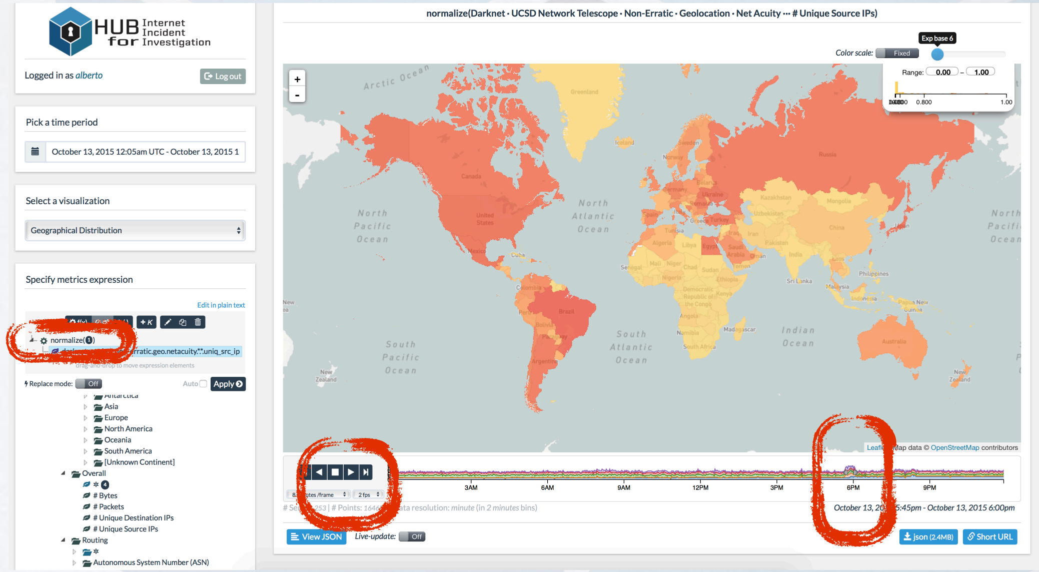
- *
- Active Probing
- BGP
- Comcast
- Darknet
  - *
  - Merit Network Telescope

Series Height: auto   Sort by: Name   Auto Y Scale: Local   Linear

Series: 5 | # Points: 7185 | Data resolution: minute

October 13, 2015 12:05am - October 14, 2015 12:02am

View JSON   Live-update: Off   json (30kB)   Short URL



HUB for Internet Incident Investigation

Logged in as *alberto*   Log out

Pick a time period

October 13, 2015 12:05am UTC - October 13, 2015 1

Select a visualization

Geographical Distribution

Specify metrics expression

Edit in plain text

normalize(1)
rratic.geo.netacuity.*.*.uniq_src_ip

drag-and-drop to move expression elements

Replace mode: Off   Auto □   Apply

- Antarctica
- Asia
- Europe
- North America
- Oceania
- South America
- [Unknown Continent]
- Overall
  - * 4
  - # Bytes
  - # Packets
  - # Unique Destination IPs
  - # Unique Source IPs
- Routing
  - *
  - Autonomous System Number (ASN)

normalize(Darknet · UCSD Network Telescope · Non-Erratic · Geolocation · Net Acuity ··· # Unique Source IPs)

Color scale: Fixed   Exp base 6   Range: 0.00 – 1.00

ates /frame   2 fps

Se 253 | # Points: 1646 ata resolution: *minute* (in *2 minutes* bins)

October 13, 20 5:45pm - October 13, 2015 6:00pm

View JSON   Live-update: Off   json (2.4MB)   Short URL

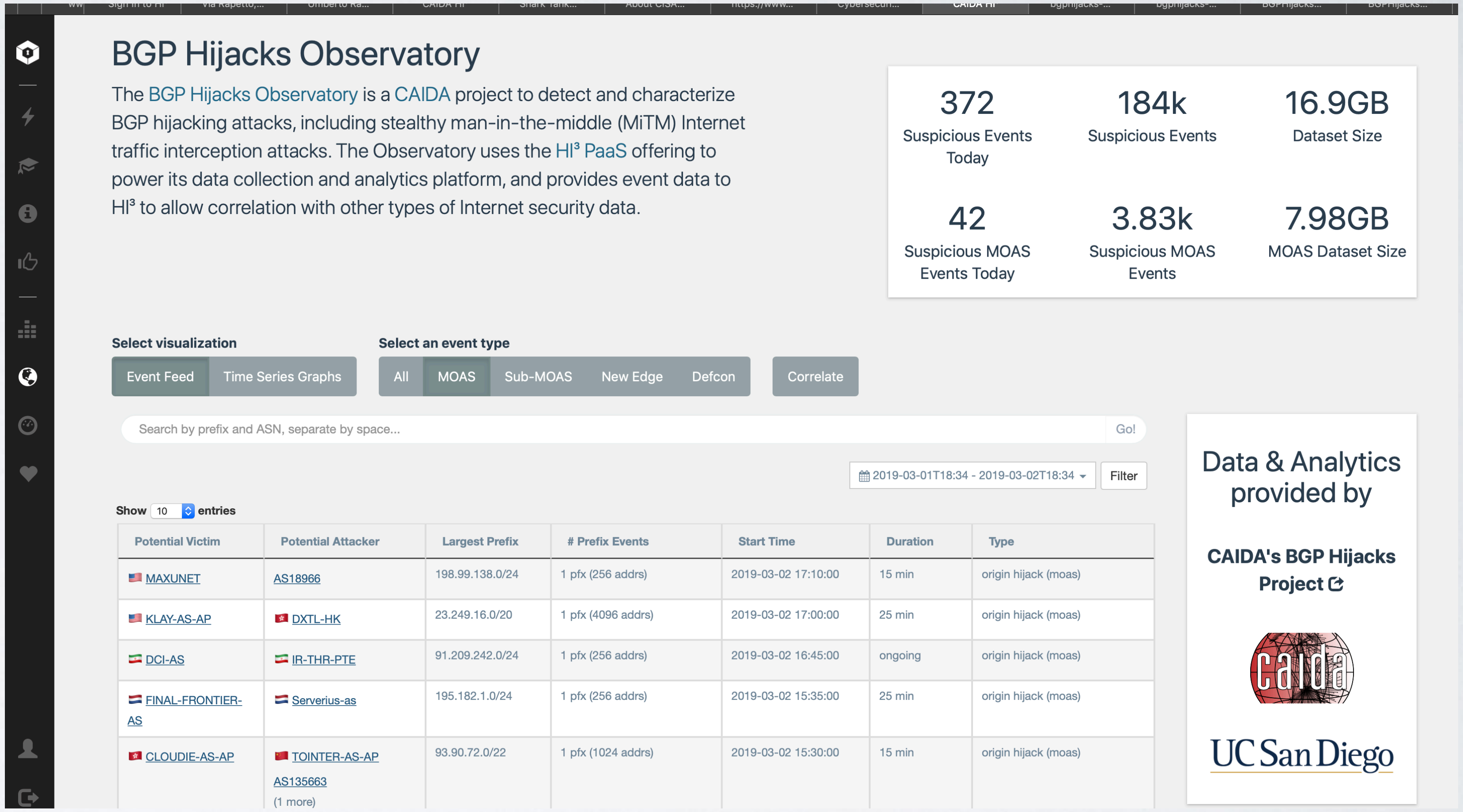
Center for Applied Internet Data Analysis
University of California San Diego

http://www.caida.org/publications/presentations/2018/hicube_impact_pi_may/hicube_impact_pi_may.pdf

FEATURED DATA FEEDS

Example: CAIDA's BGP Hijacks Observatory

- *Detects potential BGP Hijacking events by monitoring the Internet 24/7*
  - ***Detects sophisticated attacks (NewEdge, Defcon)***
  - ***Executes traceroutes during the event***
  - ***Filters out many legitimate events***
  - ***Assigns informative tags to events***
  - ***Provides visualization of AS paths and traceroutes***
- *Y2: time series for correlation w/ outages, spam*



BGP Hijacks Observatory

The BGP Hijacks Observatory is a CAIDA project to detect and characterize BGP hijacking attacks, including stealthy man-in-the-middle (MiTM) Internet traffic interception attacks. The Observatory uses the HI³ PaaS offering to power its data collection and analytics platform, and provides event data to HI³ to allow correlation with other types of Internet security data.

372 Suspicious Events Today	184k Suspicious Events	16.9GB Dataset Size
42 Suspicious MOAS Events Today	3.83k Suspicious MOAS Events	7.98GB MOAS Dataset Size

Select visualization: Event Feed | Time Series Graphs

Select an event type: All | MOAS | Sub-MOAS | New Edge | Defcon

Correlate

Search by prefix and ASN, separate by space... Go!

2019-03-01T18:34 - 2019-03-02T18:34 Filter

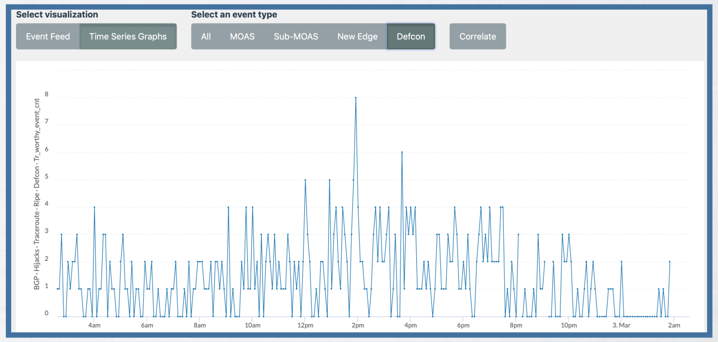
Show 10 entries

Potential Victim	Potential Attacker	Largest Prefix	# Prefix Events	Start Time	Duration	Type
MAXUNET	AS18966	198.99.138.0/24	1 pfx (256 addrs)	2019-03-02 17:10:00	15 min	origin hijack (moas)
KLAY-AS-AP	DXTL-HK	23.249.16.0/20	1 pfx (4096 addrs)	2019-03-02 17:00:00	25 min	origin hijack (moas)
DCI-AS	IR-THR-PTE	91.209.242.0/24	1 pfx (256 addrs)	2019-03-02 16:45:00	ongoing	origin hijack (moas)
FINAL-FRONTIER-AS	Serverius-as	195.182.1.0/24	1 pfx (256 addrs)	2019-03-02 15:35:00	25 min	origin hijack (moas)
CLOUDIE-AS-AP	TOINTER-AS-AP AS135663 (1 more)	93.90.72.0/22	1 pfx (1024 addrs)	2019-03-02 15:30:00	15 min	origin hijack (moas)

Data & Analytics provided by

CAIDA's BGP Hijacks Project

UC San Diego



https://dev.hicube.caida.org/feeds/hijacks

Center for Applied Internet Data Analysis
University of California San Diego

WHY "BUY"?

What does the user (e.g., CISA) get

- **Benefits**
  - Enhances our ability to detect and understand large-scale incidents
    - The whole is better than the parts
      - Ability to correlate/combine/compare multi-type data on various dimensions
    - Provides live streams of data
    - Enables collaborative analysis
    - Access through a single entry point
  - Some "exclusive" data analytics. *E.g., BGP Hijacking Observatory*

- **Risks/Challenges**
  - Developing complex infrastructure
  - Incentivizing data sharing

WHY "INVEST" YOUR DATA & TIME?

Adding Data Feeds and Analytics Platforms

- Increase the **outreach** of your project/platform
  - Critical Mass adoption model

- **Lowers costs** to for data/analytics provision (leveraging existing UI, Visualization frameworks, DB infrastructure, Auth/Auth system, …)

- Research/Investigation: leverage **combining your data with other data** and using tools to quickly correlate/compare/etc.

- Creates opportunities for **collaboration**

COMPETITION

and synergy

- Internet telemetry data analysis and event detection systems typically
  - focus on a single type of data or one class of events
  - are non-collaborative

- Potential synergy with
  - DHS IMPACT performers
  - Threat intelligence platforms

MATURITY LEVEL

hicube.caida.org — dev.hicube.caida.org

- *Prototype & development sites are both online*

- *Data Feeds:*
  - *Now:* ***Outages***, ***BGP Hijacking, Network Telescopes*** *(UCSD, MERIT)*
  - *Soon:* ***MapKIT****: relevance of Autonomous Systems in a country's Internet topology; identification of structural topological weaknesses of interest to an adversary state*
  - *Soon:* ***DoS*** *attack events*

- *Interfaces available now:*
  - *Time Series Explorer (transformation, detection, geographical viz, …)*
  - *Various project-specific interfaces*

- *Authorization and Authentication system deployed*

HUB for Internet Incident Investigation

THANKS