Kevin A. McGrail
kmcgrail@InfraShield.com

INFRASHIELD

openexpo virtual experience

Using Open Source Software to Solve Cyber Security Issues

Introduction

MAPLE LEAF
READY TO EAT · PRÊT À MANGER
KAM
LUNCHEON MEAT / PAIN DE VIANDE
455
TENEUR ÉLEVÉE EN PROTÉINES
340 g
SUGGESTED SERVING
PRÉSENTATION SUGGÉRÉE

https://www.linkedin.com/in/kmcgrail

INFRASHIELD

Security Expertise

Apache SpamAssassin

KAM.cf

Helped with the first IDS

MIMEDefang

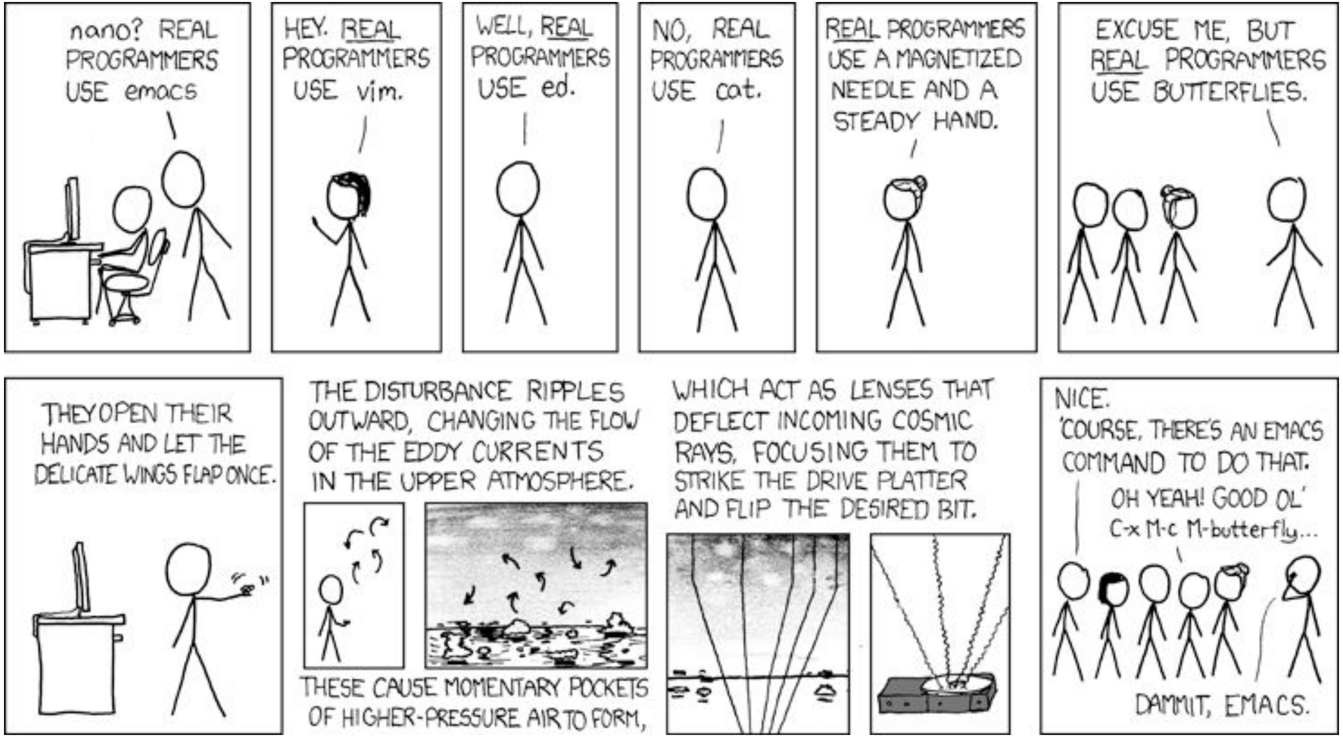
Phone Lines Cut

Apology from the FBI

9
THUNDERBIRD

Tone Setters

xkcd is a source of great knowledge and humor


nano? REAL PROGRAMMERS USE emacs
HEY. REAL PROGRAMMERS USE vim.
WELL, REAL PROGRAMMERS USE ed.
NO, REAL PROGRAMMERS USE cat.
REAL PROGRAMMERS USE A MAGNETIZED NEEDLE AND A STEADY HAND.
EXCUSE ME, BUT REAL PROGRAMMERS USE BUTTERFLIES.
THEY OPEN THEIR HANDS AND LET THE DELICATE WINGS FLAP ONCE.
THE DISTURBANCE RIPPLES OUTWARD, CHANGING THE FLOW OF THE EDDY CURRENTS IN THE UPPER ATMOSPHERE.
THESE CAUSE MOMENTARY POCKETS OF HIGHER-PRESSURE AIR TO FORM,
WHICH ACT AS LENSES THAT DEFLECT INCOMING COSMIC RAYS, FOCUSING THEM TO STRIKE THE DRIVE PLATTER AND FLIP THE DESIRED BIT.
NICE.
'COURSE, THERE'S AN EMACS COMMAND TO DO THAT.
OH YEAH! GOOD OL' C-x M-c M-butterfly...
DAMMIT, EMACS.

Cyber Security includes:

Information Technology (IT) and...

Operational Technology (OT)

CIA Triad

Confidentiality, Integrity and Availability

"The only secure computer is one that's unplugged, locked in a safe, and buried 20 feet under the ground in a secret location... and I'm not even too sure about that one."
Dennis Hughes, FBI.

Cyber Security Requires a Moral Compass

USENIX / Systems Administrators’ Code of Ethics (Previously SAGE/Lisa)
https://www.pccc.com/base.cgim?template=sage_code_of_ethics

usenix
THE ADVANCED
COMPUTING SYSTEMS
ASSOCIATION

Q/EH® Qualified/ Ethical Hacker
CEH Certified Ethical Hacker

Rules of Engagement or ROE

- Have an ROE and do NOT breach it!
- Don’t have an ROE? Don’t do it!
- Not sure what something does? DO NOT RUN IT.

KAM’s Three Favorite Open Source Projects for Cyber Security

#1 Logging (& Time Series Databases)

Challenges with Logging

Produce Data 7*24 with High Frequency and Large Volume

Big Data

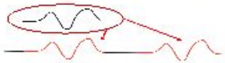
50Hz, 500points/machine,
20K wind-turbines machines,
totally up to 500 million points/sec

□ Feature of Data Ingestion

- ➢ Out-of-order sometimes
- ➢ Sparse Table (different machine has different sensors)

□ Feature of Data Query

- ➢ Time Dimension is always accessed
- ➢ Aggregation is the first-class citizen
- ➢ Time-series-specific query and analysis



□ Challenges

- ➢ Large Volume
- ➢ High Throughput
- ➢ Low Cost (historical data)
- ➢ Low Latency for Query
- ➢ Fast Aggregation
- ➢ Query-Analysis hybrid workloads

#1 Trait during
incident responses:

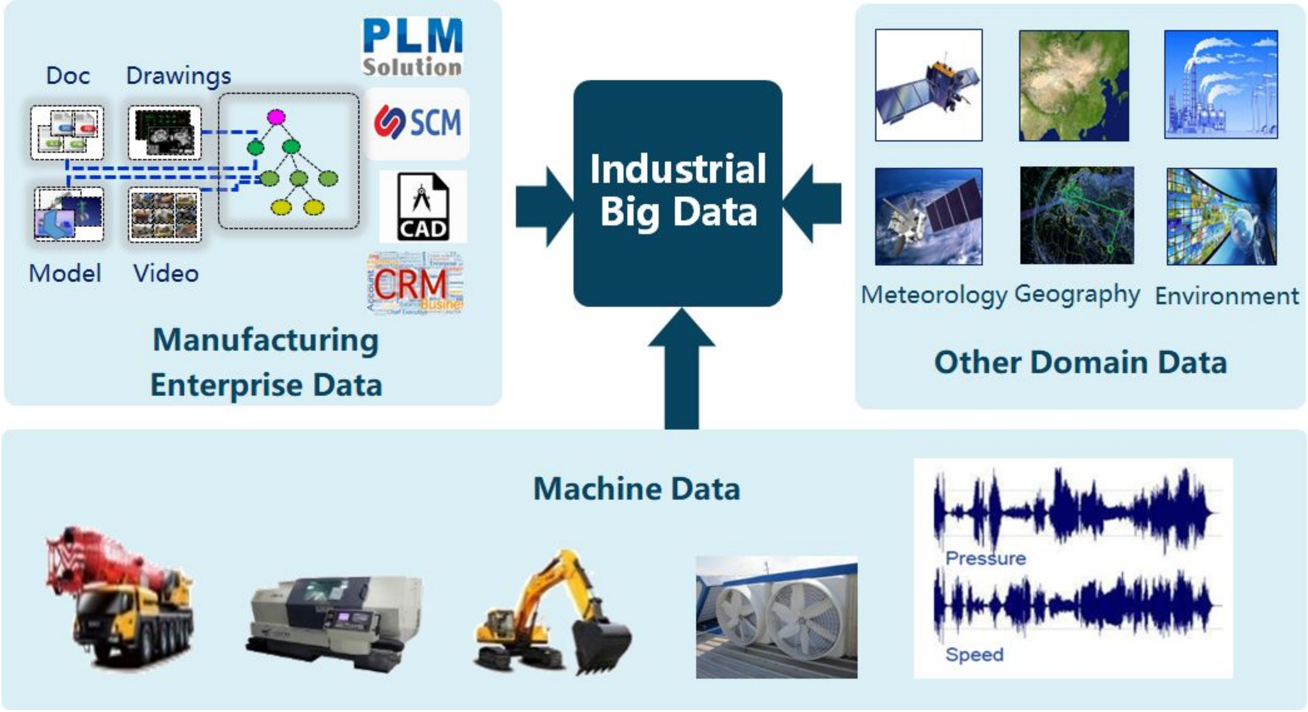
NO centralized logging

What is a Time Series Database?

Apache IoTDB

(Incubating)

Industrial Big Data



Doc

Drawings

Model

Video

PLM Solution

SCM

CAD

CRM

Manufacturing Enterprise Data

Industrial Big Data

Meteorology

Geography

Environment

Other Domain Data

Machine Data

Pressure

Speed

Apache IoTDB Features

Persist data efficiently

- Millions points ingestion per sec per node
- Tens of millions of time series

Query data with low latency

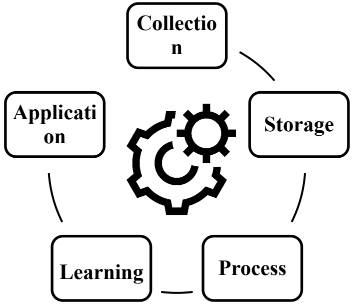
- Efficiently filter data: millions of points per sec
- Aggregation: tens of ms latency on billions of points

Exclusive operations of time series

- Segmentation
- Representation
- Subsequence matching
- Time-frequency transform
- Visualization

Integration with existing ecosystem

- Kafka
- MatLab
- Spark
- MapReduce
- Grafana



Cover the life cycle of data

- Connecting Edge to the Cloud
- Powerful query engine
- User Friendly analytics

Apache Hadoop

Quiz #1: What’s in a name?

Quiz #2: Why did they really call it Big Data?

Big Data

“Big data is data sets that are so voluminous and complex that traditional data-processing application software are inadequate to deal with them. “ Wikipedia

Apache has 48 projects under Big Data!

Apache Metron

Metron provides a Cyber Security Application Framework to detect anomalies

Ingest
Parse
Add Intel
Profile
Alert
Index

Stored in Hadoop Distributed File System (HDFS)

Shorter period of time into Solr for visualization and insight.

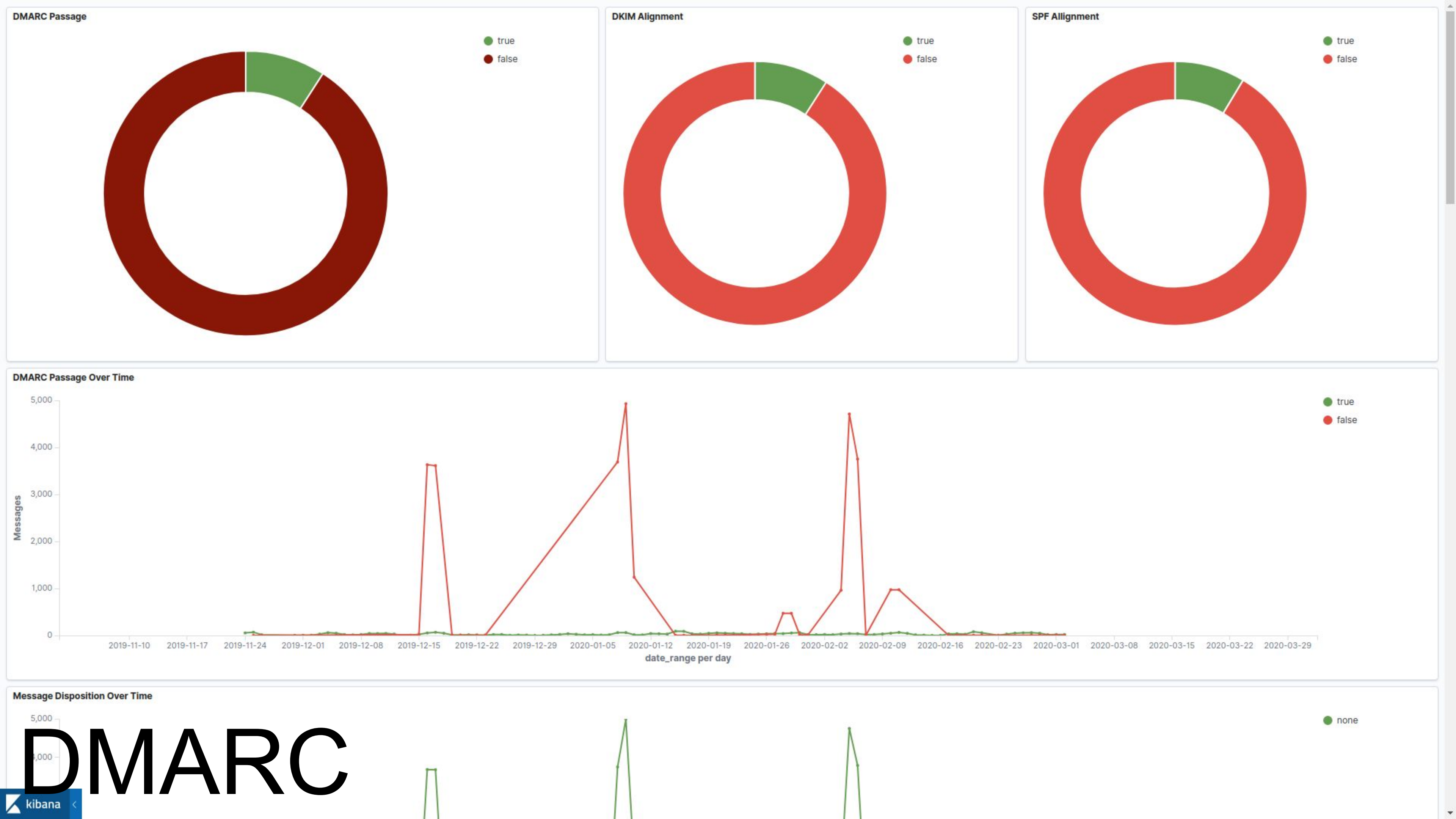
Why? “Know about a problem before a customer”

The ELK Stack

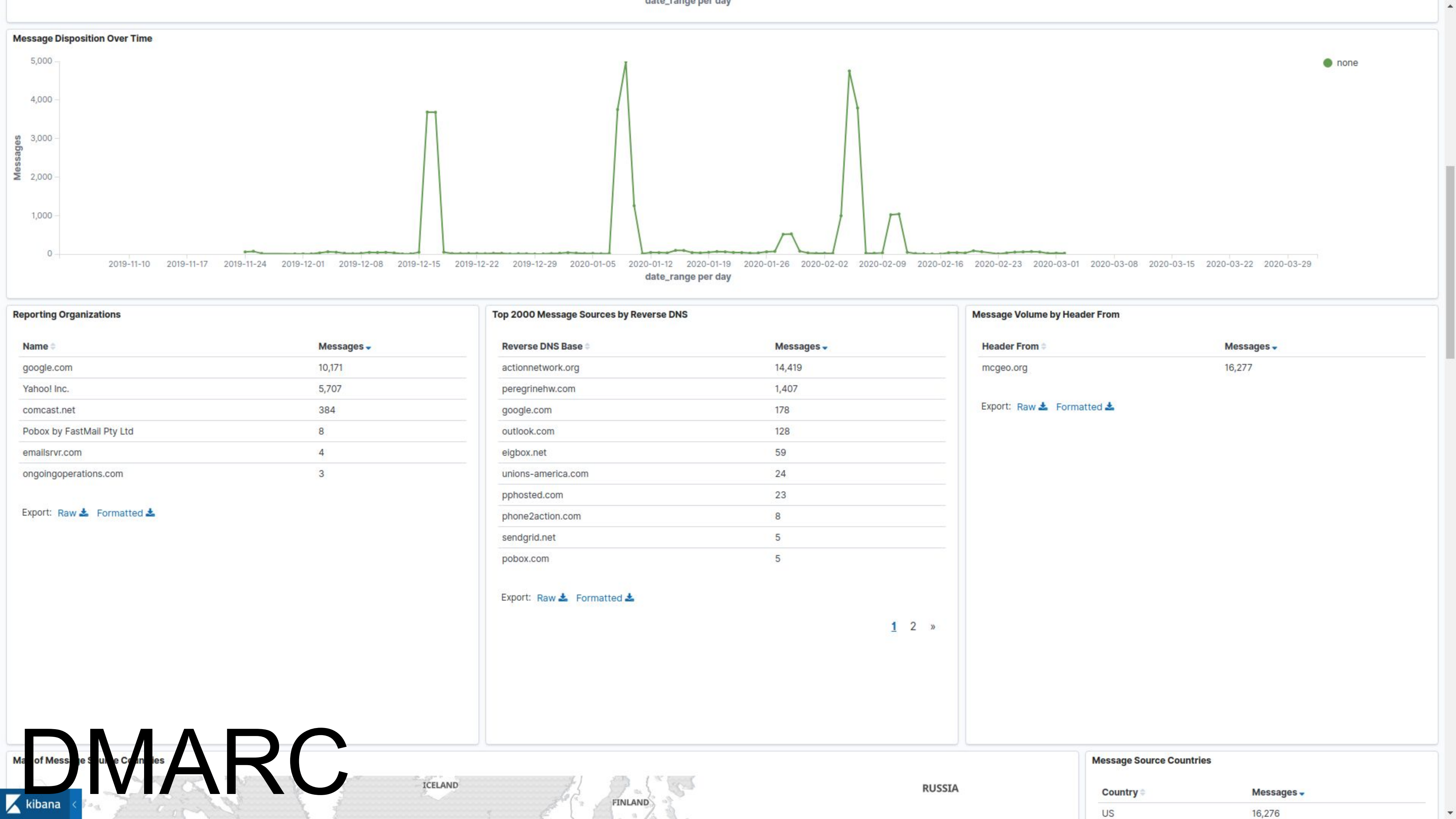
Logstash to ingest the data

Elasticsearch to store the data

Kibana to visualize the data


DMARC Passage
true
false
DKIM Alignment
true
false
SPF Allignment
true
false
DMARC Passage Over Time
5,000
4,000
3,000
2,000
1,000
0
Messages
2019-11-10
2019-11-17
2019-11-24
2019-12-01
2019-12-08
2019-12-15
2019-12-22
2019-12-29
2020-01-05
2020-01-12
2020-01-19
2020-01-26
2020-02-02
2020-02-09
2020-02-16
2020-02-23
2020-03-01
2020-03-08
2020-03-15
2020-03-22
2020-03-29
date_range per day
true
false
Message Disposition Over Time
5,000
4,000
none
kibana

DMARC



Reporting Organizations

Name	Messages
google.com	10,171
Yahoo! Inc.	5,707
comcast.net	384
Pobox by FastMail Pty Ltd	8
emailsrvr.com	4
ongoingoperations.com	3

Export: Raw Formatted

Top 2000 Message Sources by Reverse DNS

Reverse DNS Base	Messages
actionnetwork.org	14,419
peregrinehw.com	1,407
google.com	178
outlook.com	128
eigbox.net	59
unions-america.com	24
pphosted.com	23
phone2action.com	8
sendgrid.net	5
pobox.com	5

Export: Raw Formatted

1 2 »

Message Volume by Header From

Header From	Messages
mcgeo.org	16,277

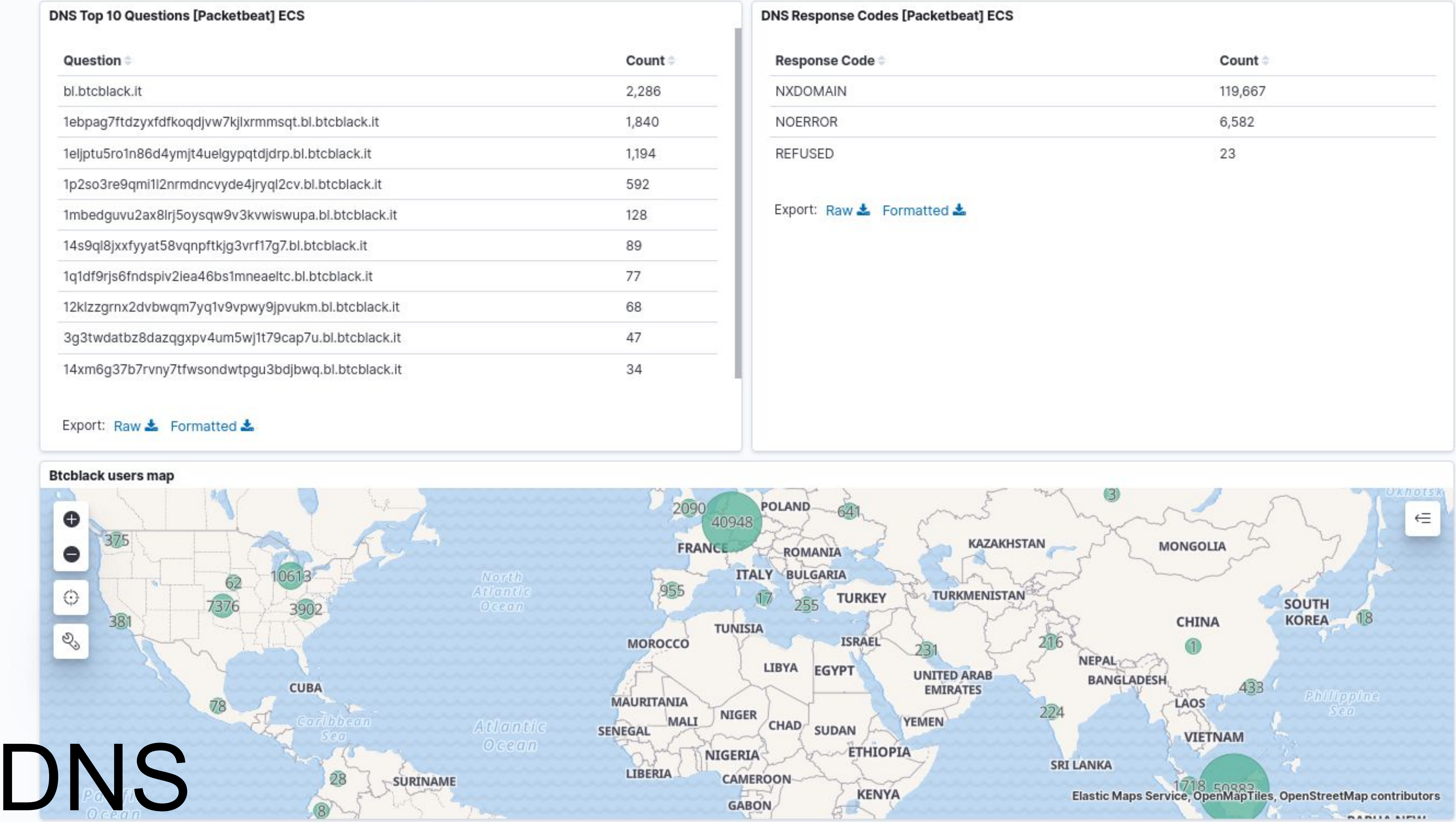
Export: Raw Formatted

DMARC

Map of Message Source Countries

Message Source Countries

Country	Messages
US	16,276



DNS Top 10 Questions [Packetbeat] ECS

Question	Count
bl.btcblack.it	2,286
1ebpag7ftdzyxfdfkoqdjvw7kjlxrmmsqt.bl.btcblack.it	1,840
1eljptu5ro1n86d4ymjt4uelgypqtdjdrp.bl.btcblack.it	1,194
1p2so3re9qmi1l2nrmdncvyde4jryql2cv.bl.btcblack.it	592
1mbedguvu2ax8lrj5oysqw9v3kvwiswupa.bl.btcblack.it	128
14s9ql8jxxfyyat58vqnpftkjg3vrf17g7.bl.btcblack.it	89
1q1df9rjs6fndspiv2iea46bs1mneaeltc.bl.btcblack.it	77
12klzzgrnx2dvbwqm7yq1v9vpwy9jpvukm.bl.btcblack.it	68
3g3twdatbz8dazqgxpv4um5wj1t79cap7u.bl.btcblack.it	47
14xm6g37b7rvny7tfwsondwtpgu3bdjbwq.bl.btcblack.it	34

Export: Raw Formatted

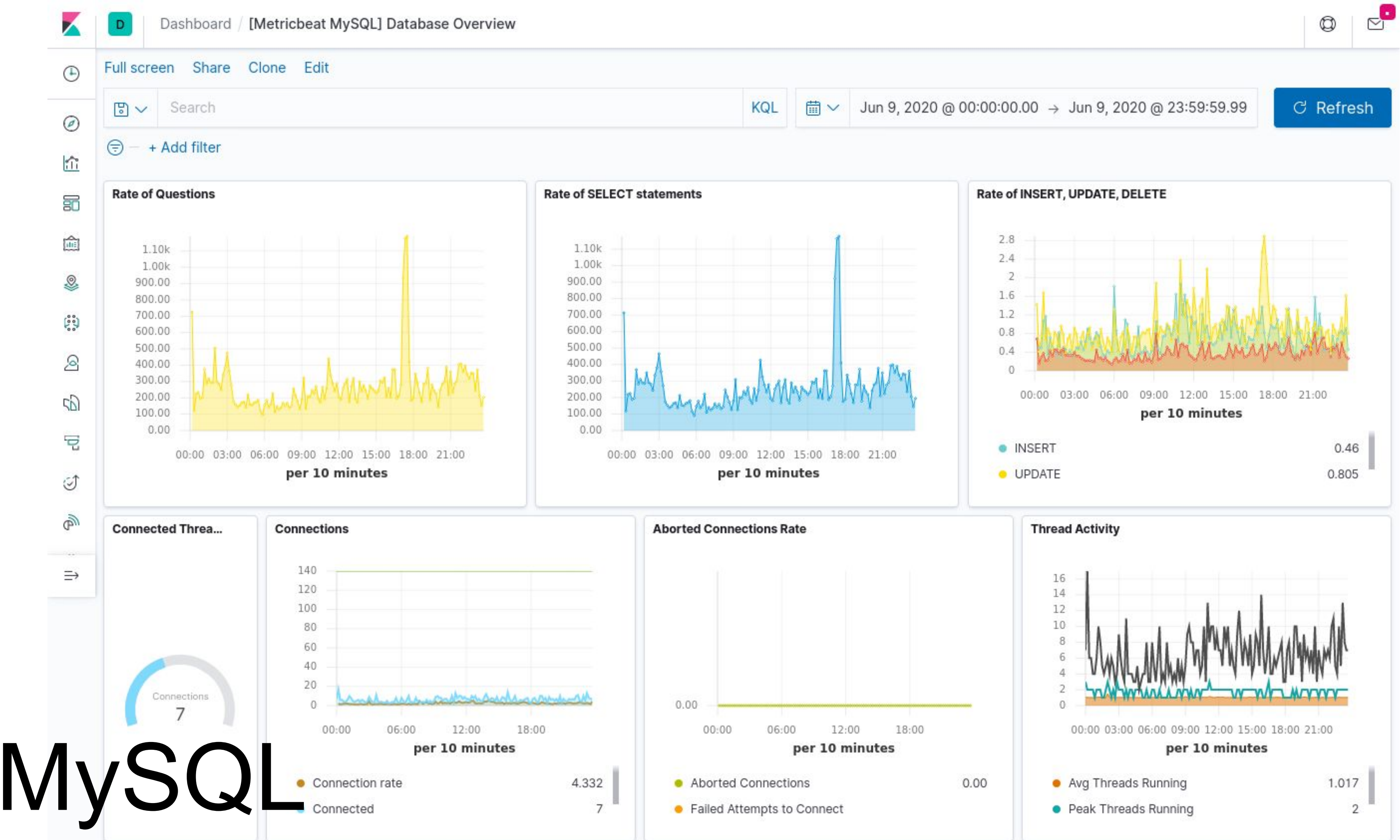
DNS Response Codes [Packetbeat] ECS

Response Code	Count
NXDOMAIN	119,667
NOERROR	6,582
REFUSED	23

Export: Raw Formatted

Btcblack users map

DNS



Dashboard / [Metricbeat MySQL] Database Overview

Full screen Share Clone Edit

Search KQL Jun 9, 2020 @ 00:00:00.00 → Jun 9, 2020 @ 23:59:59.99 Refresh

+ Add filter

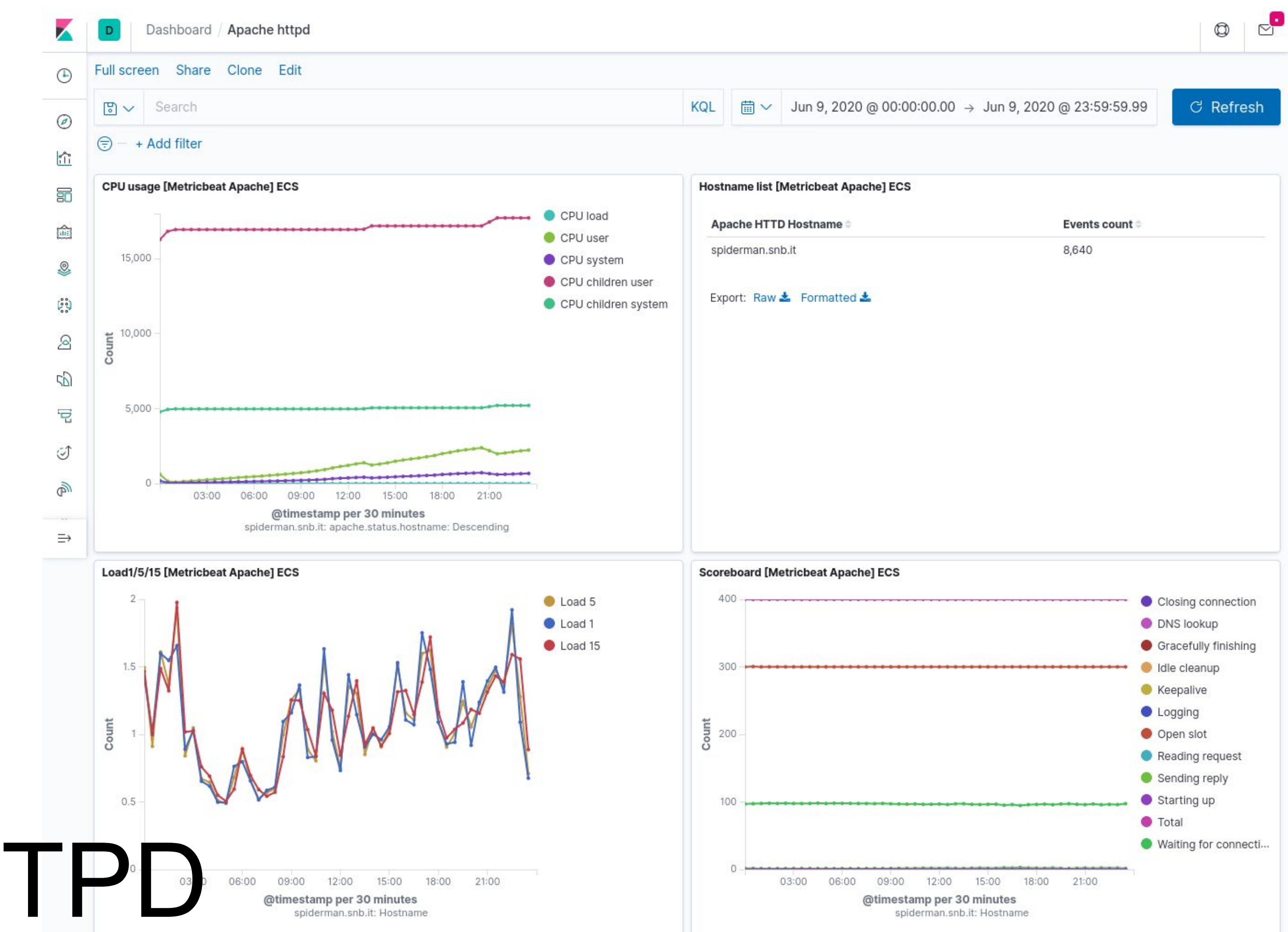
● INSERT 0.46
● UPDATE 0.805

● Connection rate 4.332
● Connected 7

● Aborted Connections 0.00
● Failed Attempts to Connect

● Avg Threads Running 1.017
● Peak Threads Running 2

MySQL



Apache HTTD Hostname	Events count
spiderman.snb.it	8,640

HTTPD

Quiz Break

What is the name of the First all-electronic Computer?

Quiz Break

A: ENIAC

(Electronic Numerical Integrator And Computer)

#2 Email Security

What is Apache SpamAssassin?

- A Scoring Framework
- An API & a Program
- Stops Spam
- Spam Markers:
  - Consent vs. Content i.e. what is Spam?
  - Transactional/Relational
  - RFC-Compliance

What is MIMEDefang?

E-Mail filtering framework using Sendmail's Milter interface

MIMEDefang provides mechanism; you provide policy - Dianne F. Skoll

Perform midstream commands on SMTP conversations

MIMEDefang & KAM.cf (and more)...

The McGrail Foundation's mission is to provide services, education and advocacy for private, secure and unimpeded business and communications.

THE
McGRAIL
FOUNDATION

Which reminds me…

The Apache Software Foundation

Who is the ASF?

The Apache Software Foundation is a 501(c)(3) Charity often referred to as just Apache or the ASF.

501(c)(3) Charity not a 501(c)(6) Trade Organization

We’re known for the HTTP server and the Apache Software License.

What is the ASF’s Mission?

To provide software for the public good.

We do this by providing services and support for many diverse software project communities of individuals AT NO CHARGE.

The Apache Software License

The ASLv2 is known for its permissive, business-friendly stance with patent grants and without copyleft provisions.

Powered by Apache

80% of the world's websites use our software

Every Smartphone in the world uses our software

Every plane in US airspace is tracked w/our software

Projects.Apache.org & The Incubator

There are currently 387 open source initiatives at the ASF:

- 202 committees managing 339 projects
- 5 special committees
- 43 incubating podlings

#3 Linux Distros

KALI and Raspberry Pi OS (Raspbian)

So much more I could talk about...

https://mcgrail.com/downloads

The Perl Conference 2019 - Fighting Spam with Perl using Apache SpamAssassin & MIMEDefang

ThunderBay GDG 2020 - How to Build a Cyber Range

webpros Summit 2019 - Linux System Swiss Army Tools for Administrators

Currently working on an open source curriculum for a cybersecurity practicum

Thanks!

Image Credits:

KAM photo taken by Ted King, used with permission.
xkcd Comic https://xkcd.com/378/ & https://xkcd.com/327/ (CC BY-NC 2.5)
Company & Project Logos are Brand Resources of their Respective Companies

Appreciation:

Thanks to James Thompson at cPanel for the Apache SpamAssassin Logo
Thanks to Josh Audette.for the MIMEDefang Logo
Thanks to Jianmin Wang for lending some of his slides for IoTDB
Thanks to Giovanni Bechis & Georgia Smith @ PCCC for the Kibana visualizations.

Thank you to the MANY open source projects and their efforts in changing the way we compute.

INFRASHIELD

Kevin A. McGrail
www.linkedin.com/in/kmcgrail