What Does “Big Data” Even Mean?

Josh Goldfarb
FloCon 2014

nPULSE TECHNOLOGIES

Agenda

- Introduction
- Big Substance or Big Hype?
- Process
- Data Value vs. Data Volume
- Streamlining Workflow
- Integrating Actionable Intelligence
- Communal Presence
- Remembering the User
- Money Shot
- Summary
- Q&A

Introduction

- “Big Data” seems to be everywhere
- What does it mean?
- A fad or something we can capitalize on?
- How can we exploit it?

Big Substance or Big Hype?

- Probably both
- Unique opportunity
- Collection and analysis
  - One requires the other to maximize value
  - Collection is fairly mature
  - Analysis is maturing
- Examples from experience

Process

- It’s not sexy, but it’s important
- Data is plentiful, but resources and budget are not
- Work smarter, not harder
- Develop efficient processes for all functions
- Automate within the process where possible (not for automation’s sake)
- Enrich data where possible to further optimize the process (not for enrichment’s sake)

Data Value vs. Data Volume

- Collection and storage are both expensive
- Difficulty in assessing data value causes organizations to collect everything
- Data volume is overwhelming, cutting down on retention
- Retention need not be uniform
- Often helpful to consider the value of the data collected to security operations/incident response
- Netflow as an example
- Layer 7 enriched netflow as an example

Streamlining Workflow

- Data nearly limitless
- Processing power somewhat limited
- Technology/tools maturing
- Number of analysts very limited
- Focus analysts on one stream of alerting/one work queue
- Enable/facilitate quick and efficient investigation/analysis with the end goal of resolution (no analysis for analysis' sake)
- Re-read process slide (nothing worse than a wild goose chase)

Integrating Actionable Intelligence

- Intelligence can greatly add to the detection capabilities of organizations, as well as the response capabilities
- Must be timely, reliable, high fidelity, and actionable
- Integrate into the alerting stream (keep the analysts focused on one work queue)
- Example

Communal Presence

- The best organizations maintain a strong presence in the broader community
- Give and take
- Sometimes the best intelligence sources are free
- Be remembered

Remembering the User

- Systems get infected, but users create, edit, use, and share data
- Users sometimes do weird/dangerous things
- World is very IP-based right now
- Perhaps user-based analysis is the next frontier
- If we look at the data from a user-based perspective, the world looks a bit different
- Can begin to build a much richer picture

Money Shot

- Lots of talk about Indicators of Compromise (IOCs) and/or other forms of intelligence
- Just another “jumping off point” or pivot
- Easy to generate a high volume of false positives with the wrong “jumping off point”
- Choose wisely
- Go for the “Money Shot”

nPULSE TECHNOLOGIES

Summary

- “Big Data” maturing
- Has promise
- Offers unique capabilities
- Leveraging the tools for both collection and analysis is key
- Smooth and seamless integration of people, process, and technology/tools is still as important as always

nPULSE TECHNOLOGIES

Questions?