Scaling Security Investigations With Interactive Event Graphs & Spark

Leo Meyerovich

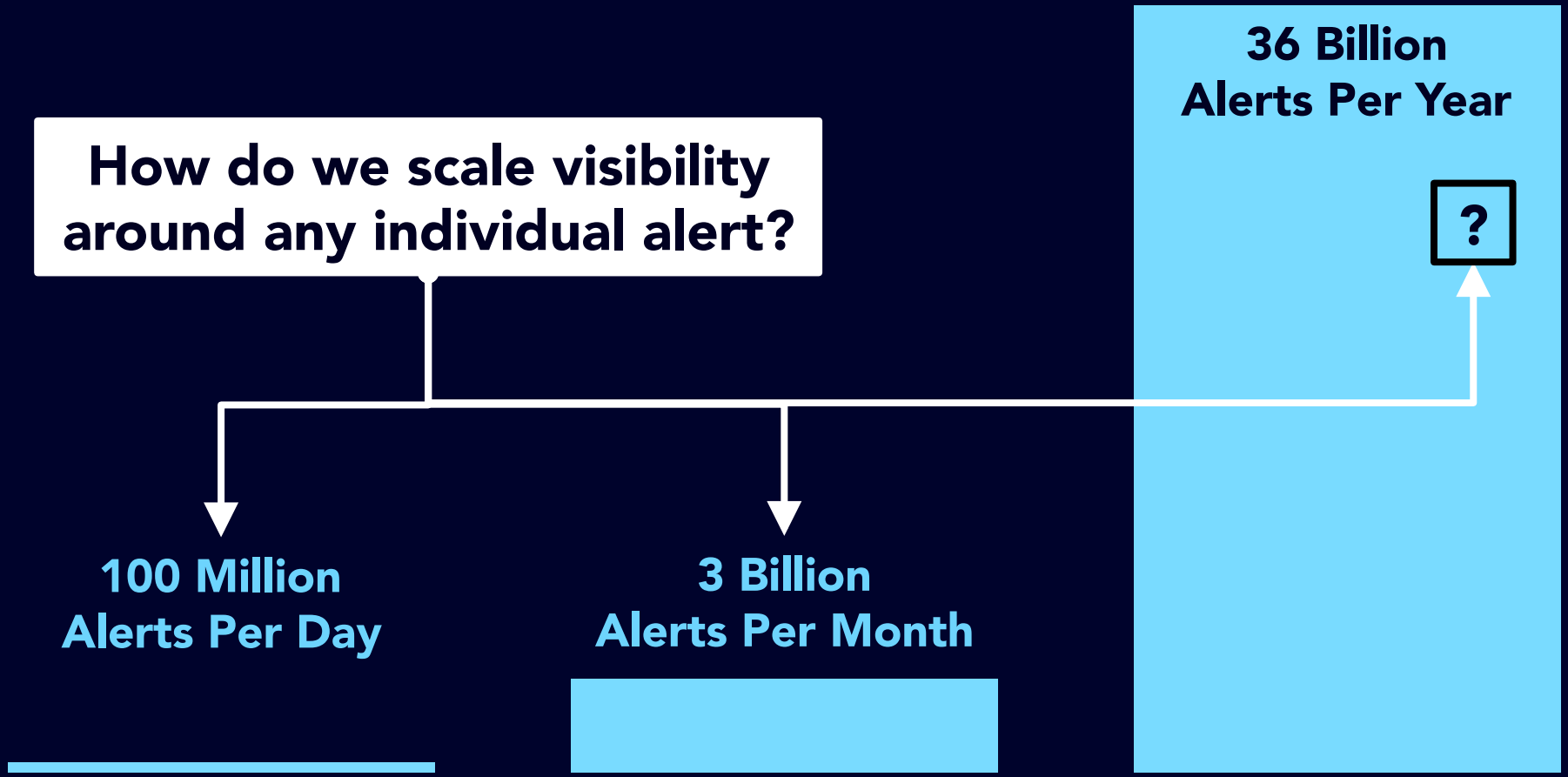
CEO/Co-founder, Graphistry

Joshua Patterson

Principal Data Scientist, Accenture Tech Labs

.conf2016

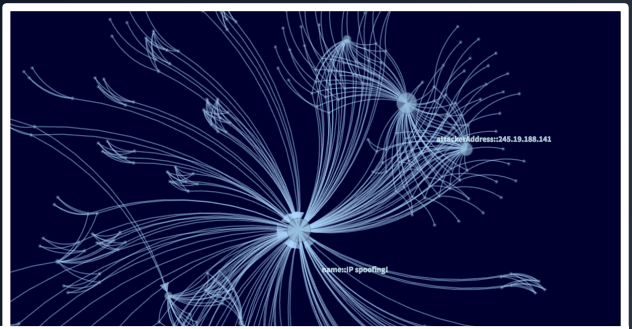
splunk>



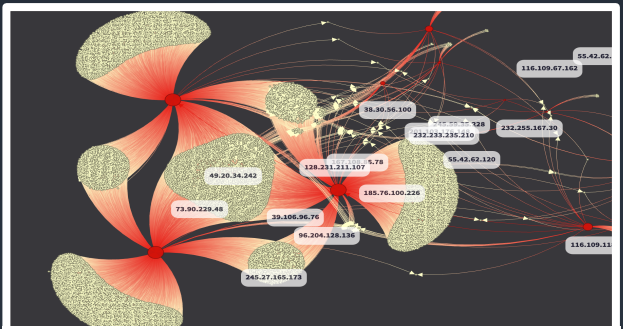
Alert counts from one enterprise security team

TALK: Architecture & Practice Of Visualizing Events @ Scale

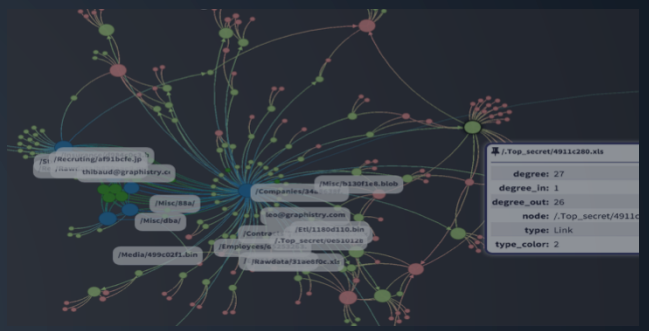
Splunk/Spark/Graphistry → Security Event Graphs



IR: Killchain Analysis

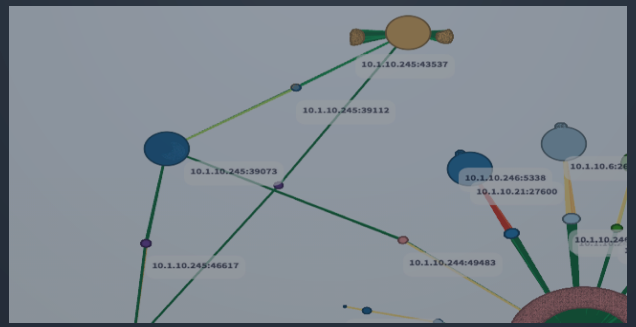


Hunting: Daily Anomalies



Shadow IT: Auditing Dropbox

Botnet Deconstruction



Ops: Outage Root Cause

Fraud: Tracking Embezzlers

QUICK DEMO

About Graphistry

Mission

Look and work across event systems with one intelligent layer

Investigator BETA

- Investigation tier with Splunk connector
- Increase visibility & automation in investigations
- Handle increasingly large and diverse data sources

Founding Team

Spun out of UC Berkeley parlab in 2014

Core Technology

Smart & scalable visual querying powered by
GPUs, language design, and unsupervised learning

Accenture Labs: Expanding Global Presence

For more than 20 years, Accenture Labs has served as the tip of the spear for technology innovation at Accenture.

Over the last 5 years Accenture Labs has:

- Supported 300+ client engagements and hosted 1100+ client workshops
- Published 200+ thought leadership pieces, filed 110+ patent applications, and garnered 350+ Tier-1 media hits

Agenda

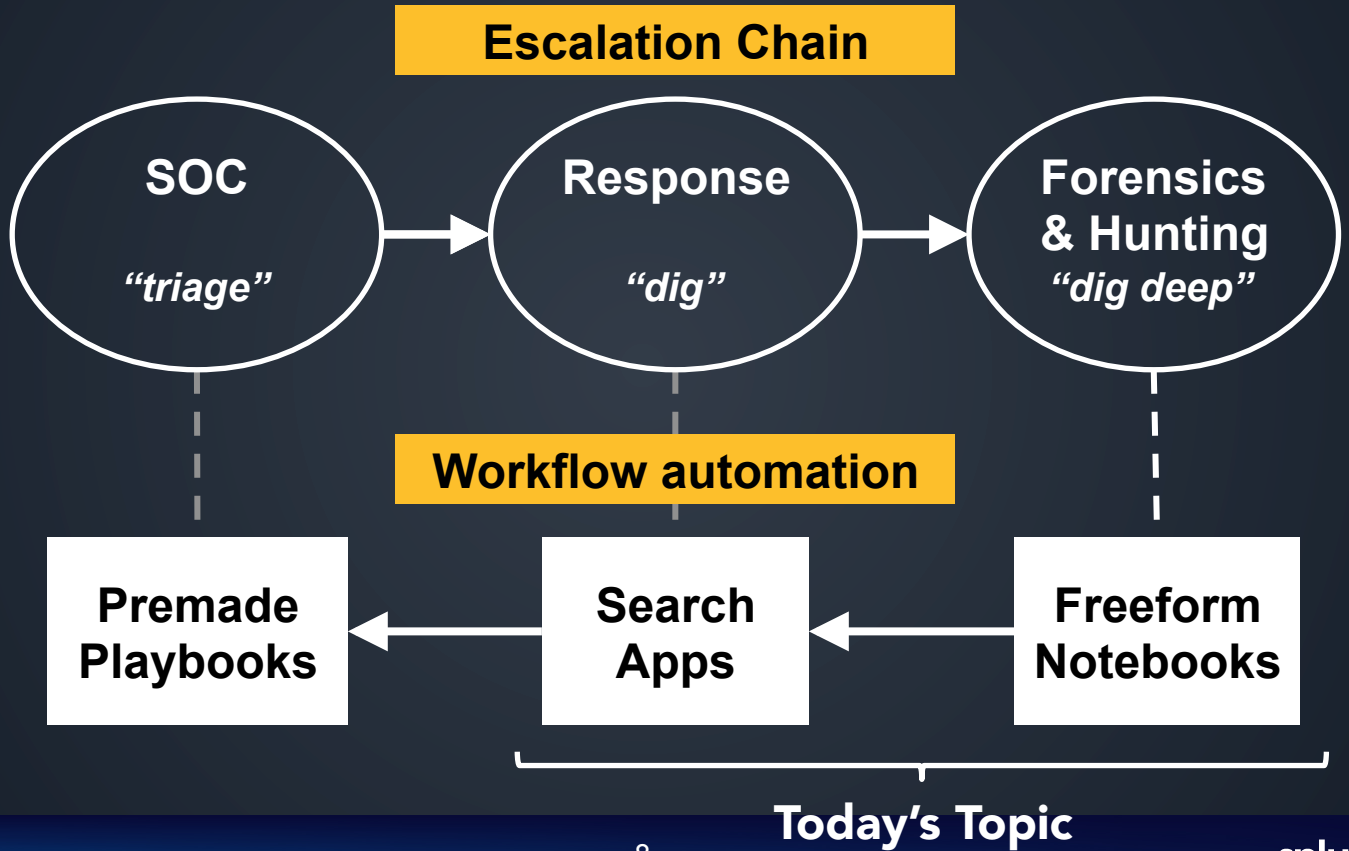
ASGARD End-to-End Architecture
Big data rethink with Splunk, Spark, & Graphistry

Hunting Demo: Notebooks for anomaly analysis

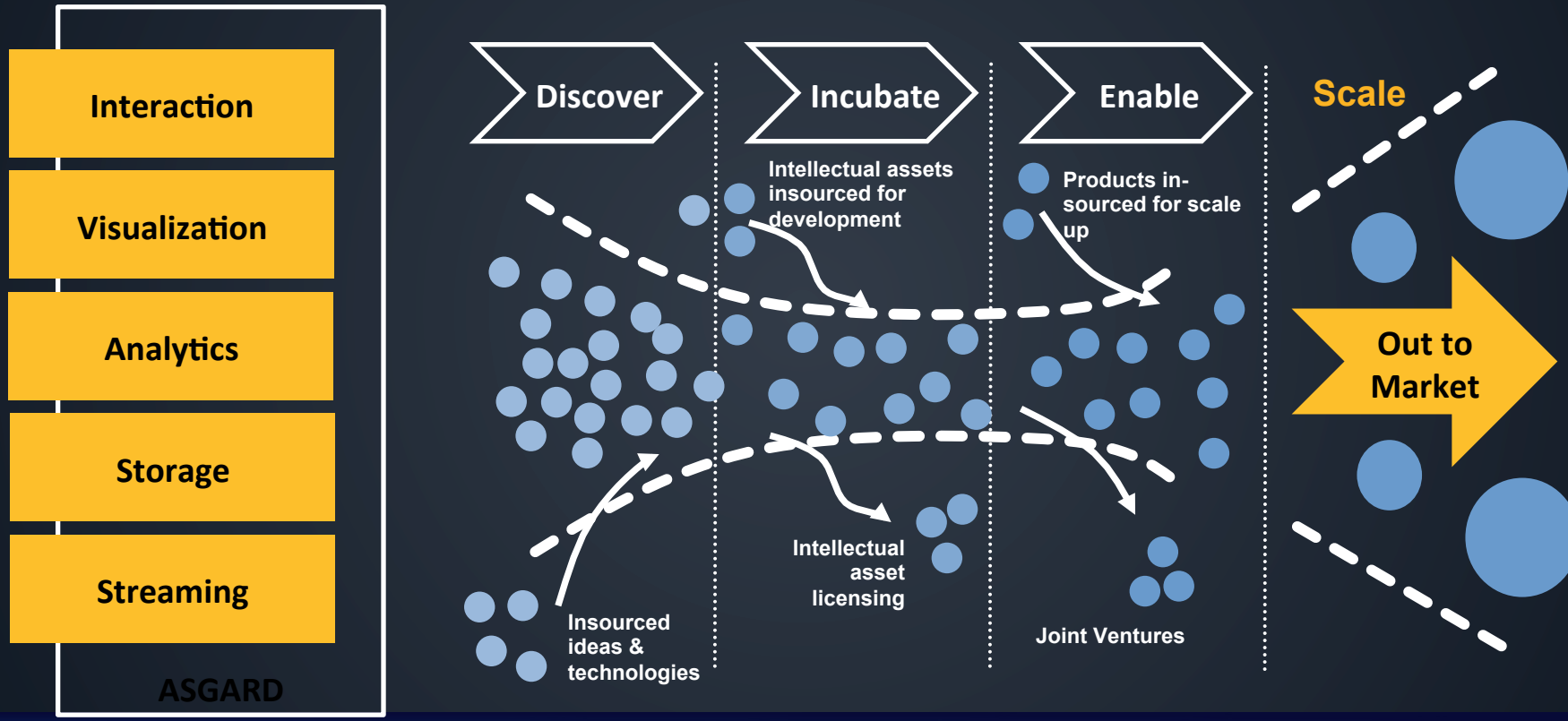
Visual Science: Event graphs for scalable views (+ GPUs!)

Incident Response Demo: Botnet outbreak

Who Visually Analyzes & How?



Accenture ASGARD: Rethinking Cyber Security ~~Analytics~~ Hunting



Accenture Labs ASGARD Platform

FIRST DEPLOYMENT

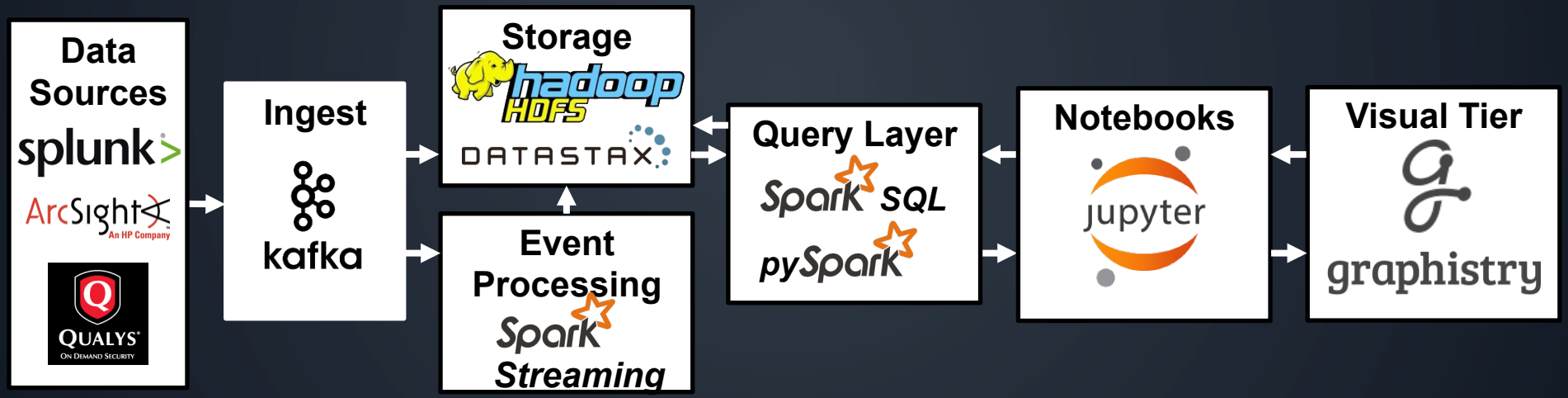
- 100-400M events/day

GOALS

- **Scalable**
- **Interactive, Real-Time**
- **Affordable**

THEMES

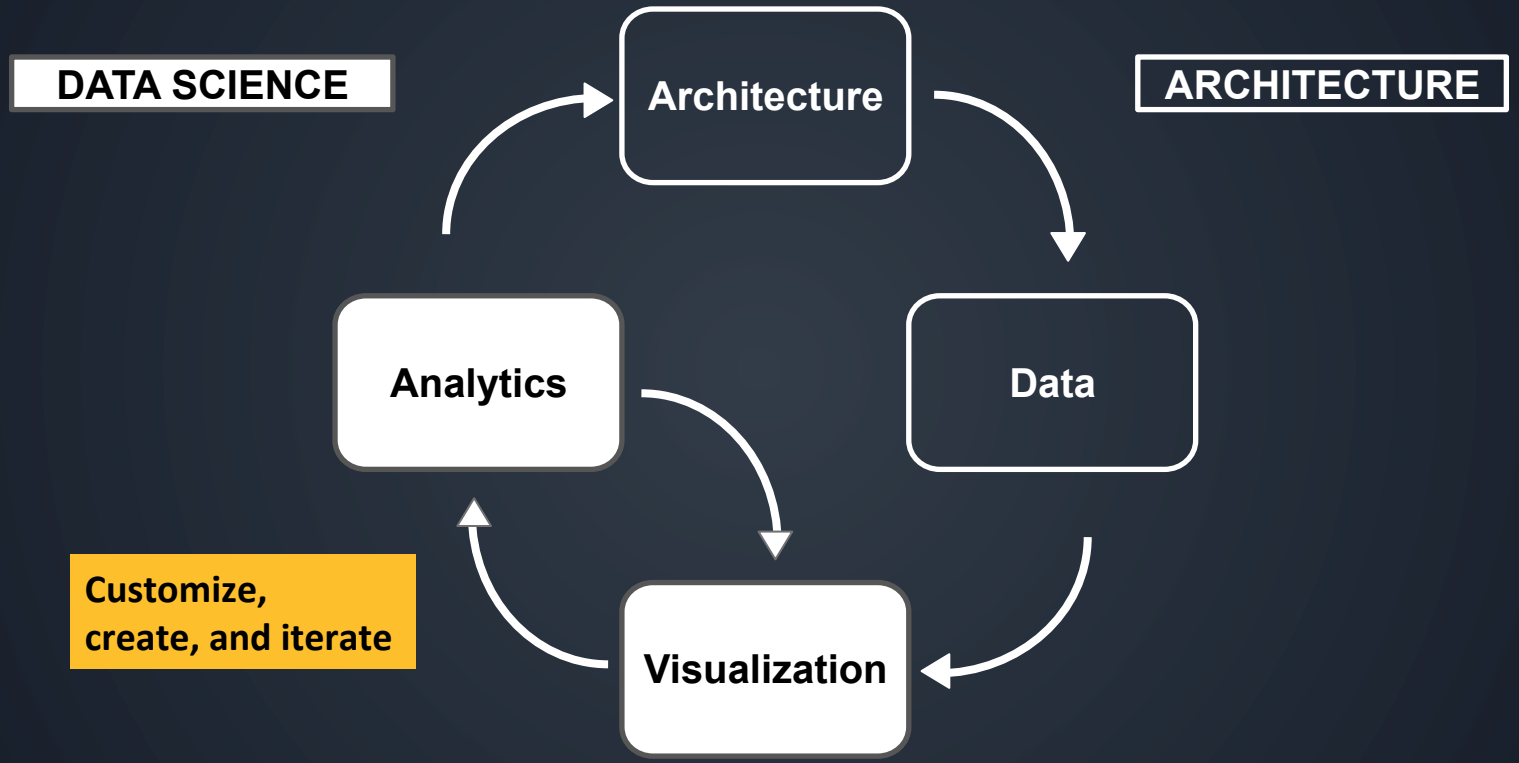
- **OSS Distributed In-Memory**
- **GPUs**
- **Events/Graphs**



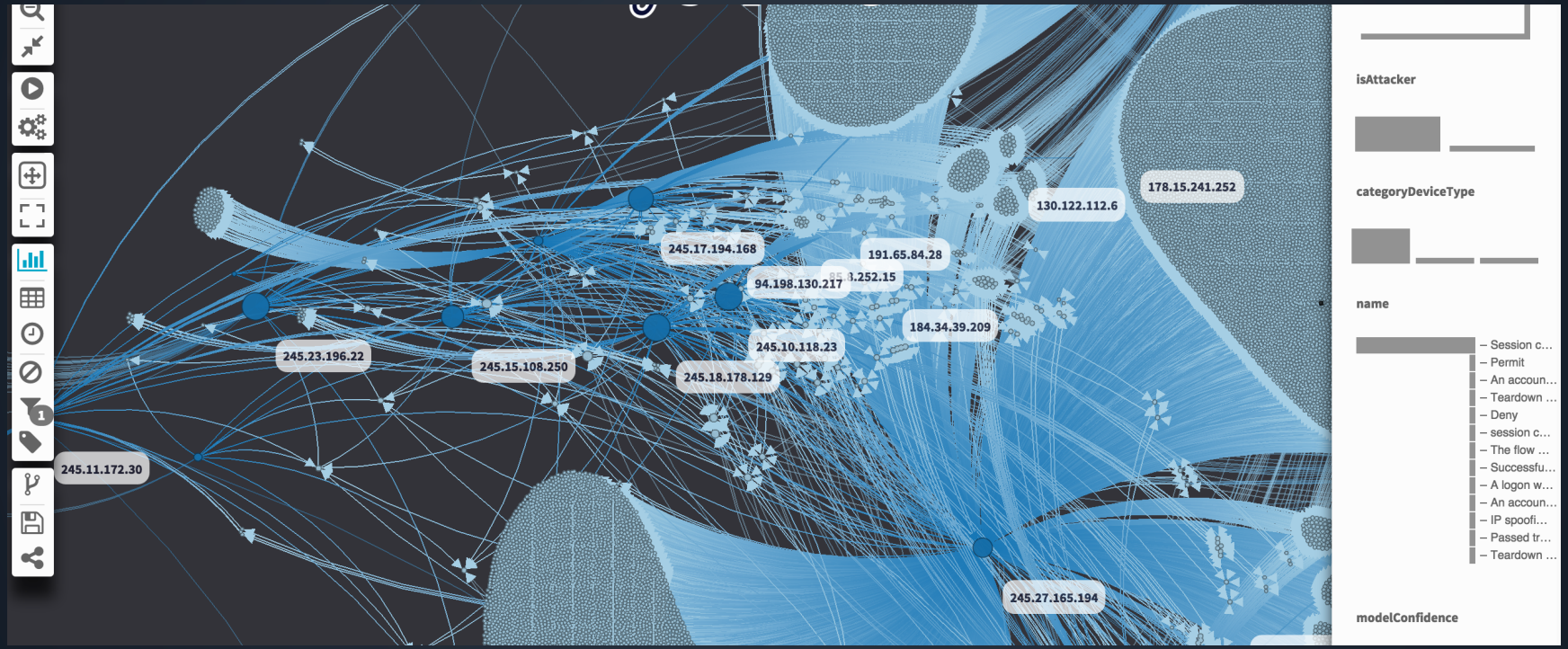
ASGARD Acceleration Benchmarks

	Everyday Scenario	Time Period	Without ASGARD	With ASGARD	ASGARD's Speed Improvement
1	Network communication lookup, from one host (IP) to multiple hosts (IPs)	1 Day	3h 20m 13s	1m 44s	114 Times Faster
		1 Week	Not Feasible	4m 05s	
2	Failed logon attempts lookup for active directory	1 Day	18m 26s	1m 37s	10 Times Faster
		1 Week	2h 13m 45s	3m 10s	41 Times Faster
3	Looking for malware (exe) in the Symantec logs	1 Day	3h 24m 36s	1m 37s	125 Times Faster
		1 Week	Not Feasible	1m 37s	
4	Proxy Logs Lookup (looking for specific domain)	1 Day	4h 30m 13s	2m 54s	92 Times Faster
		1 Week	Not Feasible	1m 09s	

Building For The Long-term: Innovation Cycle



Hunting Demo (5min): Notebook For Daily Anomalies



Agenda

ASGARD End-to-End Architecture
Big data rethink with Splunk, Spark, & Graphistry

Hunting Demo: Notebooks for anomaly analysis

Visual Science: Event graphs for scalable views (+ GPUs!)

Incident Response Demo: Botnet outbreak

Why Graph Visualizations?

GOAL: Security Visualization For The Data Era

Relevant

- **Scale visuals to modern enterprises**
- 1 million devices under management
- Billions of events between them
- Reveal patterns & outliers

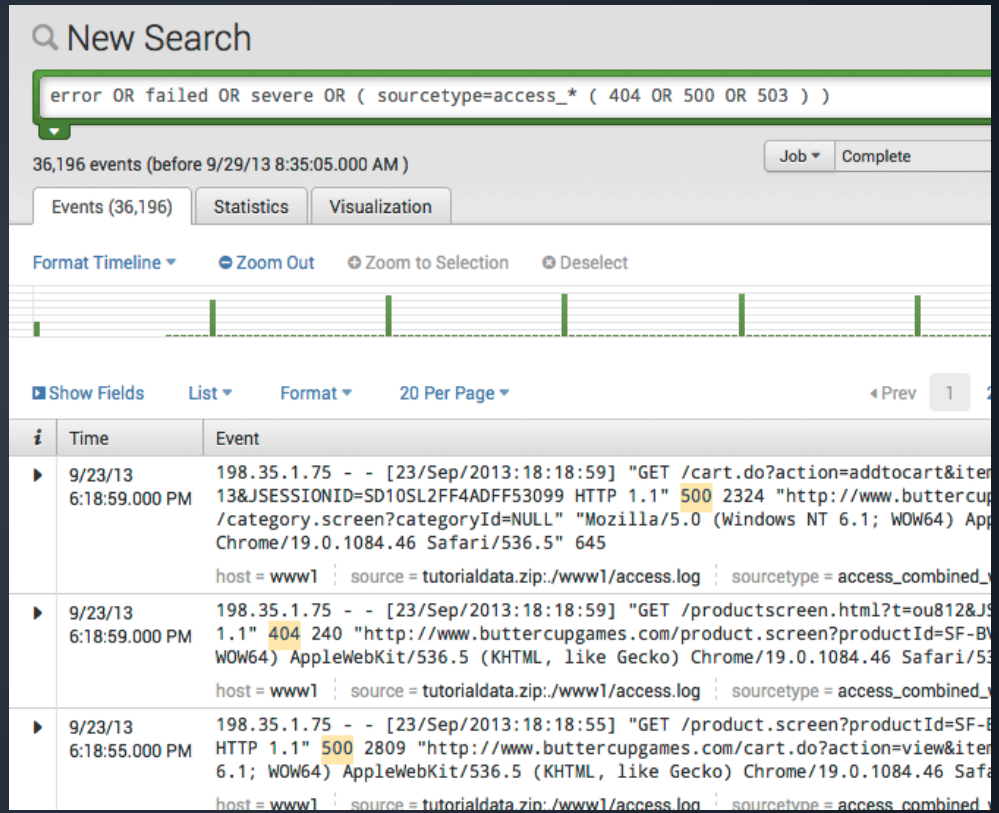
Interactive

- **Explore at speed of thought**
- Code less; easily pivot & drill
- Responsive: 10ms – 1s

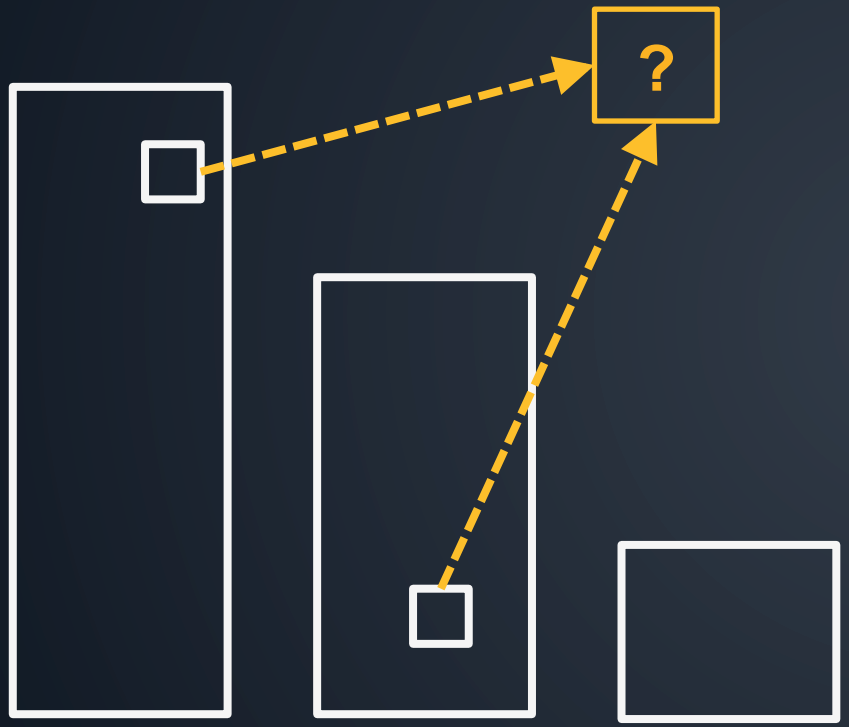
Lists Do Not Visually Scale

- Text search is a great starting point!

- **Do not scale**
  Do not see the 30K+ events nor the IPs, users, nor how they relate…



Bar Charts Hide Relationships



- Good for summaries!
- But not: relationships, patterns, outliers
- But not: individual items

Event Graphs: A Key Missing View

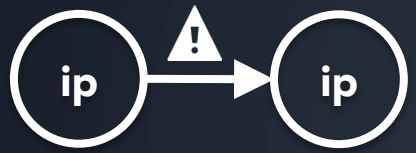
Unified Model

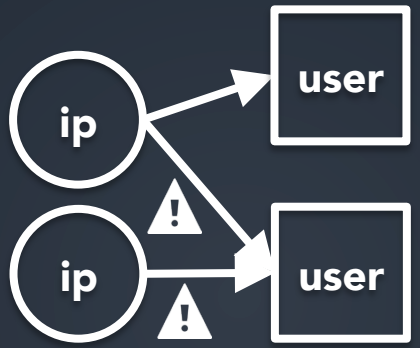
- Describes entities & links, e.g., events
- Multipurpose: connect, see, interact

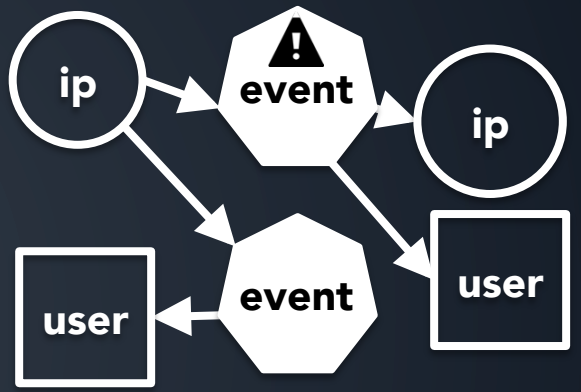
Visual

- Spot relationships, patterns, outliers
- Inspect individual items
- Work at enterprise scale

Different Graphs for Different Scales, Questions







Uni

Ex: Network mapping

"What services use this?"

Multi

Ex: SSH trails

"Is a user crossing zones?"

Hyper

Ex: Incident Response

"Did this escalate?"

Graphistry's GPU Platform: Scale & Accelerate The Visual Analytics Tier

Optimized networking

amazon web services™

GovCloud

GPU rendering

WebGL

(No JavaScript!)

iOS

GPU analysis & ML

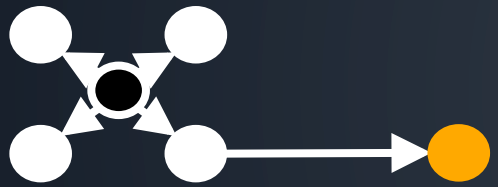
OpenCL

NVIDIA

GPUs: Accelerate Every Component 10X+

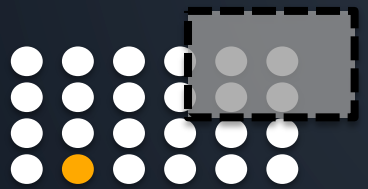
Interactive Rendering

1+ million entities: 100X+ over D3.js



Meaningful Viz: Layout & ML

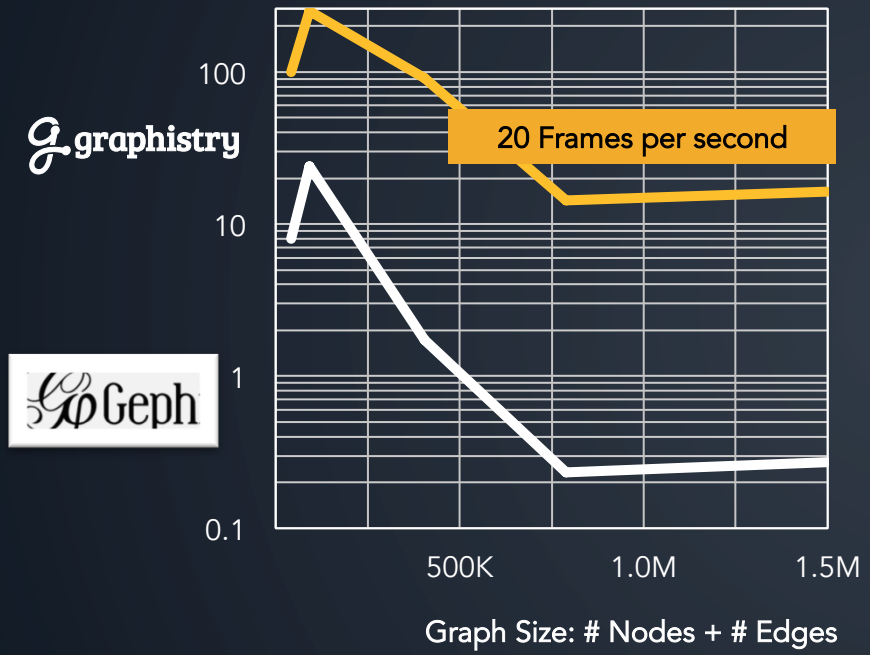
Smart clustering, coloring, sizing: 50X+ over Gephi



Interactive Analytics

Quickly drill down:
1 NVidia Tesla K80 = ~9 TFLOPS

Sample Speedup: Interactive Clustering



- 60X more data than Gephi
- Iterative clustering: pure GPU
- GPU in server via Node-OpenCL, Nvidia Docker

Demo: Botnet Investigation (7min)

Lessons Learned

ASGARD

- Rethink security platform for scale, speed, cost
- Innovation process for next-gen SIEM flow

Graphistry

- Event graphs: unify tools; explore behavior at scale
- Investigation tier: increase visibility & streamline pivots

THANK YOU

GRAPHISTRY

accenture

High performance. Delivered.

Leo Meyerovich
info@graphistry.com

Joshua Patterson
joshua.patterson@accenture.com

.conf2016

We're hiring engineers + seeking innovation partners!

splunk>