Smart Ransomware Detection Assistance in SSDs

Andy Walls, IBM Fellow, Chief Architect and CTO IBM FlashSystems

Agenda

- Ransomware Increasing and continually Changing
- The need for fast recovery
- Immutable Snapshots
- Ransomware detection in block storage

Flash Memory Summit

17%
of Cyber Attacks are Ransomware

26%
Clients who paid the ransom still could not recover the data

23
days, average recovery after a ransomware attack

2X
Cyber Attacks YTY

21%
dormant threats, up from 5% YTY

45%
45% of production data affected

It's a Dangerous Cyber World Out There

FlashSystem Cybersecurity Strategy

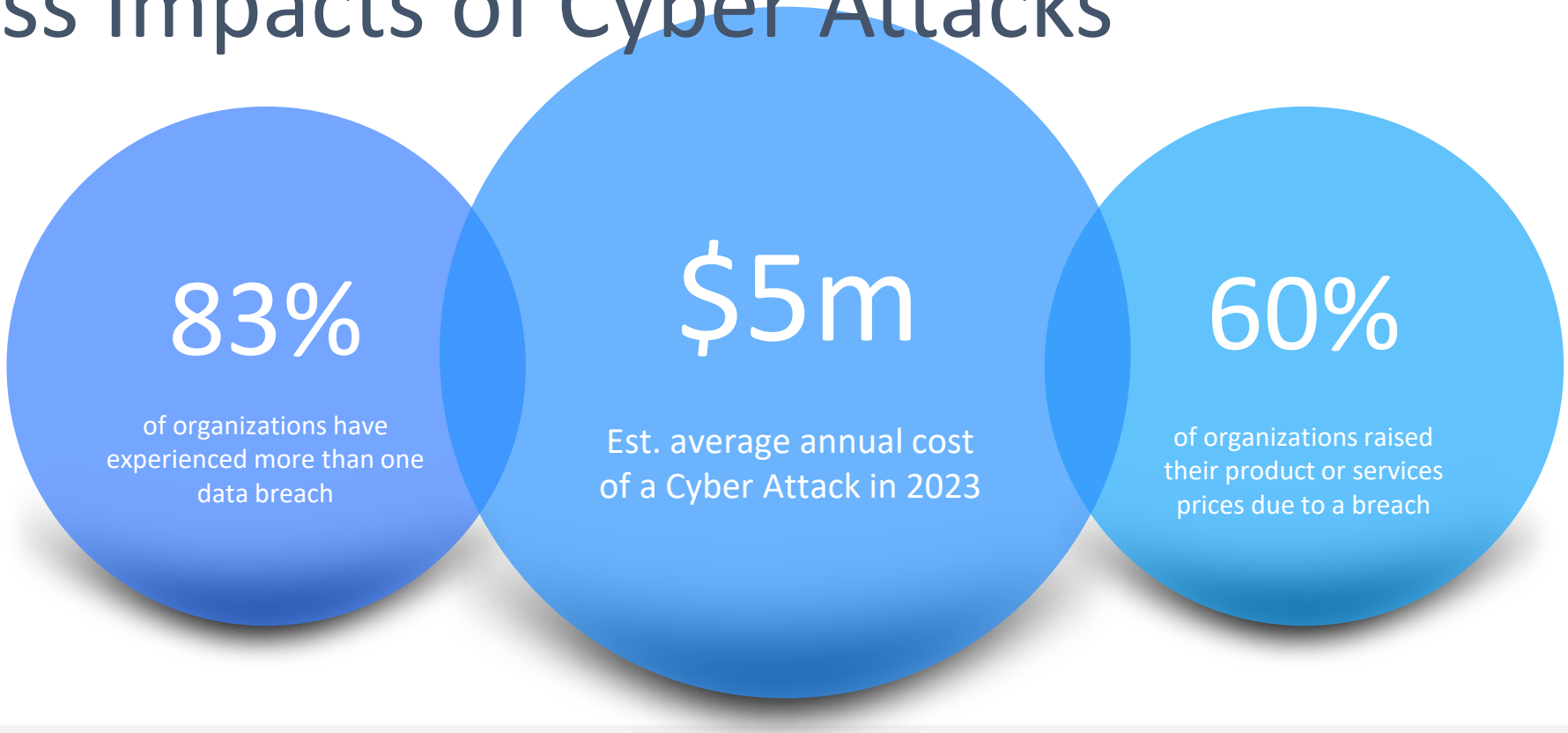
FlashSystem Strategy

- Secure by default
- Intelligence storage
- Fast recovery
- Sales Accelerator

NIST SP 800-209

- **Data Protection**
- **Restoration assurance**
- **Encryption**
- **Isolation**

Business Impacts of Cyber Attacks



$10m

for noncompliance of Operational Resilience regulations

Capability Steps to Data Resilience

Flash Memory Summit

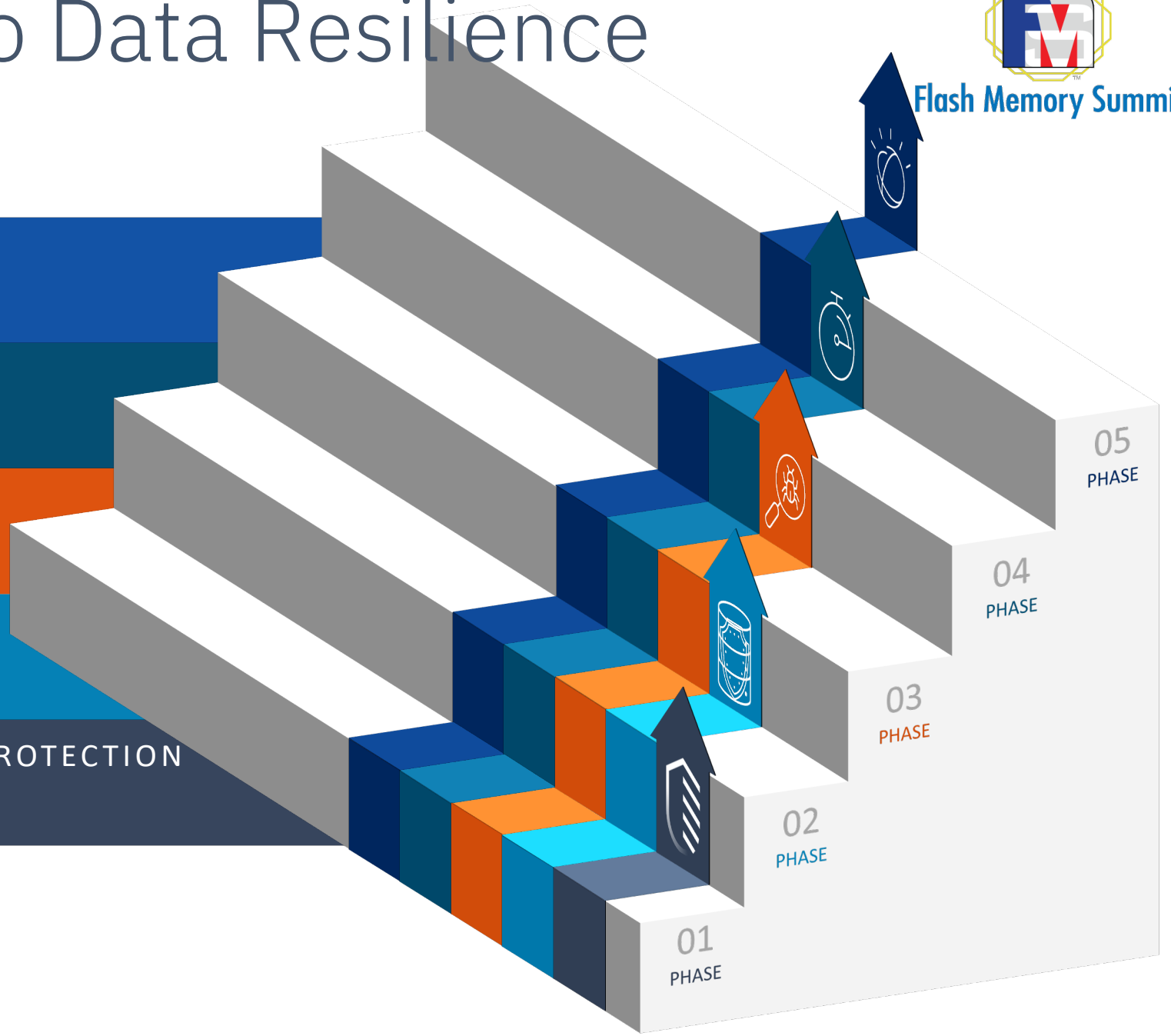
AUTOMATION
Simplified operations plus ability to test and prove recoverability
Integration between Cyber Security & Cyber Resiliency

RECOVERY
Rapid business recovery in minutes to hours
Avoid paying ransomware

DISCOVERY
Understand when defences have been compromised
Malware scanner and data pattern insights

SECURE IMMUTABILITY
Recoverable data points
Incorruptible, data can not be deleted

FOUNDATIONAL SECURITY & DATA PROTECTION
Predict, prevent, and respond SIEM Consultation, SOAR, XDR
Protect from infrastructure failures and Natural disasters



State of Security Today

Flash Memory Summit

2023 Threat Intelligence Index

41%
Of attacks, phishing was top initial access vector

27%
Of attacks saw extortion as the most common attack impact

2022 Cloud Threat Intel Report

99%
Of cases had cloud identities excessively privileged

50%
Of cloud data breaches were in shadow IT

2022 Cost of a Data Breach

$4.35 million
Average total cost of a data breach (+3%)

277 days
Average time to detect and contain (-3.5%)

Attacks are faster, better detected, but still missed

68 days down to 4 days
Average time to deploy ransomware

94% faster
Reduction in average attack duration

Top 3 Strategies to reduce the Cost of a Data Breach (same for last 3 years)

- Adopt Zero Trust based protection
- Have an IR Plan and Practice It
- Automate Detection & Response program/processes using AI/ML

The Industry has evolved to have good Disaster Recovery Methodologies

- High Availability covers for failing systems
- Disaster recovery covers large scale outage
- Backup covers for data loss - but how quick can you recover ‘everything’?
- HA / DR are mostly based upon data replication
- All these do not protect from infiltration, ransomware, or data theft

Cyber Resilience

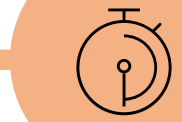
- Preparation for malicious activities during a cyber attack
- Options to contain damage
- Technologies to keep data safe even if attacked
- Priority for vital business processes
- Full recovery planning

INCREASING THREATS

Flash Memory Summit

CYBER ATTACKS

Ransomware attacks increased 95% in 2021



RANSOMWARE

"time to ransom" dropping to a matter of hours

CYBER WARFARE

Complete data wipe (production & backups)
No ability to recover

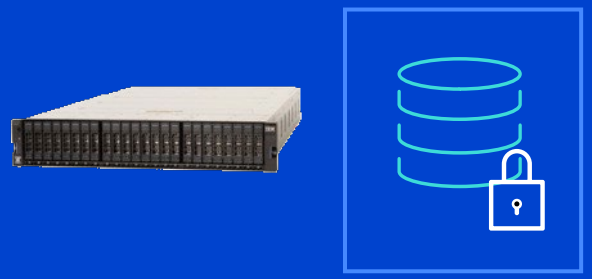
The FCM is a Computational Storage Device

- Transparent Data Reduction
- Has FPGAs to program new functions
- 4 different embedded cores to assist the offload
- Interfaces to get the results of the offload to and from the FCM

Data Resilience to accelerate recovery

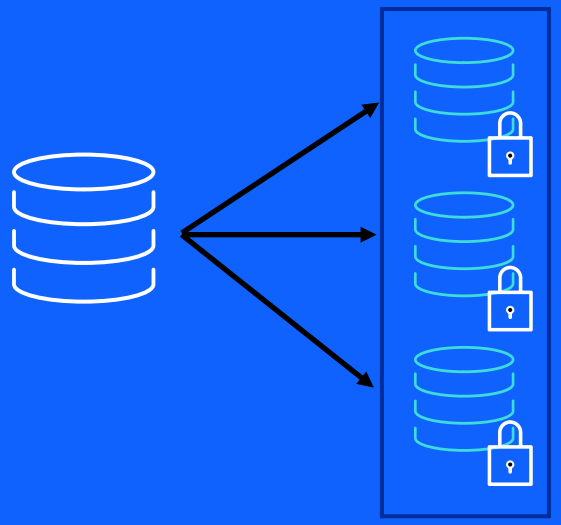
Immutable Snapshots on Array

Separation of duties



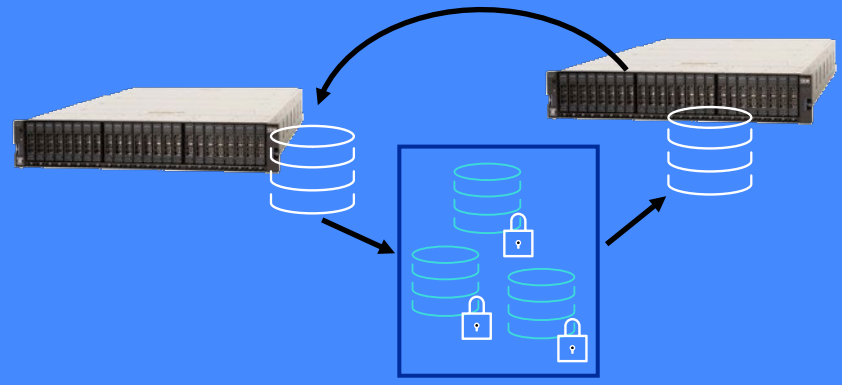
Additional security capabilities to prevent non-privileged users from compromising production data

Protected copies of data



Capabilities to regularly create secure, immutable point in time copies – Up to 15,000 copies

Speed of recovery



Functionality that enables different use cases to restore corrupted data **in minutes or hours vs days or weeks**

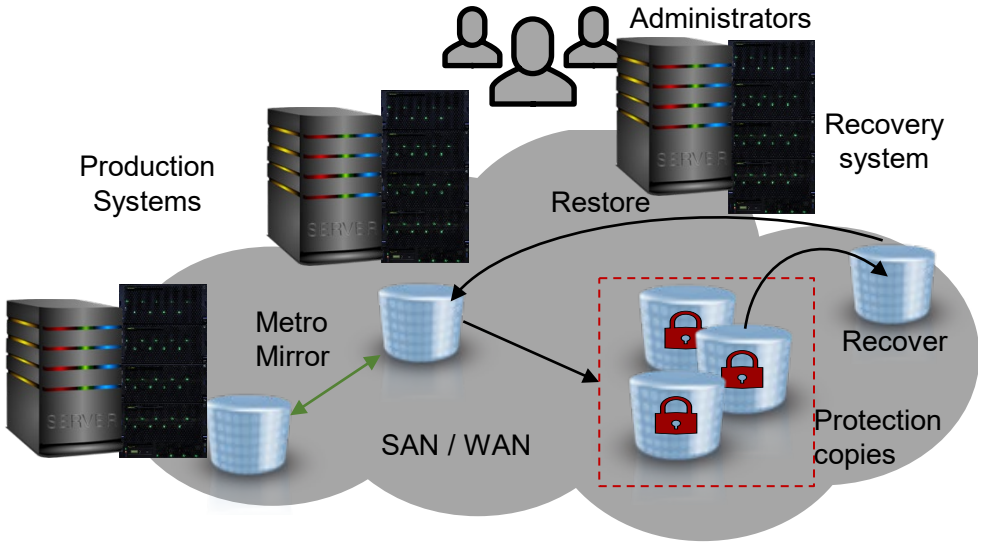
The Key to Immutable Copies

- Tighten up on security for the deletion of copies
  - Allow Expiring only
  - Two Person Integrity
  - Audit log
- Copies can not be mounted
- Having Snapshots is not enough
  - Ransomware attackers are known to infect backups
- Important to scan and use AI and analysis to ensure snapshots are not corrupted

Air gap: Virtual and physical isolation of protection copies

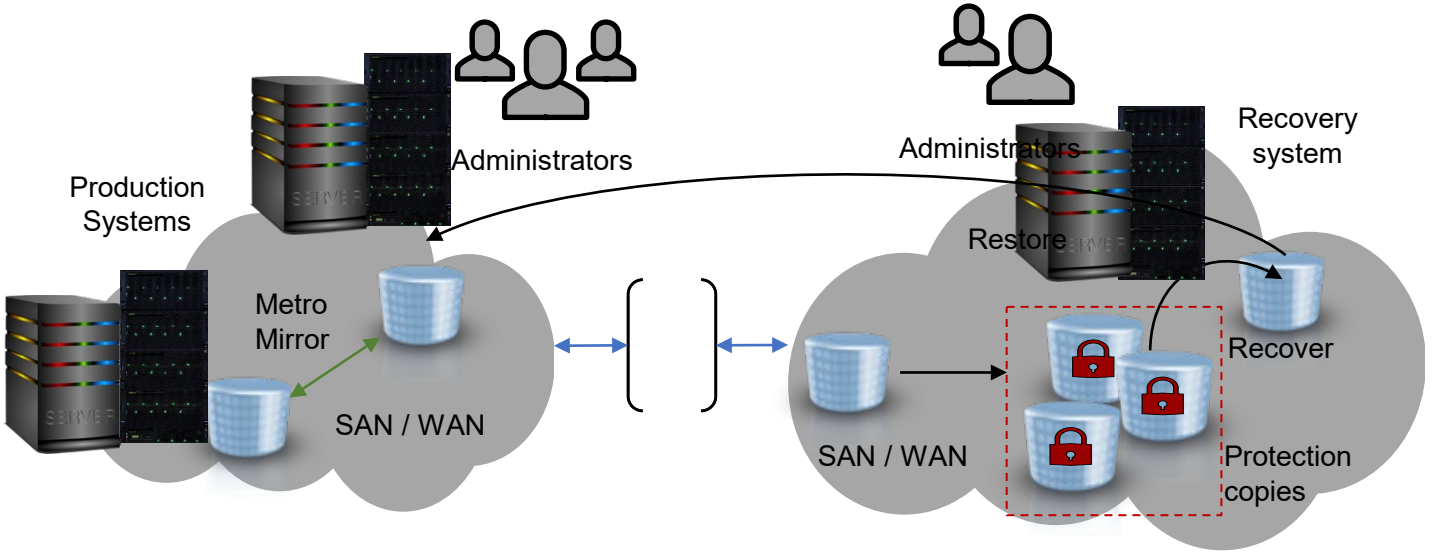
Flash Memory Summit

Virtual isolation



- The protection copies are created in one or more storage systems in the existing high availability and disaster recovery topology
- The storage systems are typically in the same SAN or IP network as the production environment

Physical isolation



- Air gapped solution
- Additional storage systems are used for the protection copies
- The storage systems are typically not on the same SAN or IP network as the production environment
- The storage systems have restricted access and even different administrators to provide separation of duties

We are developing technology not just to recover from attacks –

<u>But to detect them early!</u>

Realization #1:

Flash Memory Summit

Block Storage and SSDs are missing some context other parts of the system have

BUT: It can generate data needed for determining Ransomware attacks with less performance impact then any other part of the system

Realization #2:

To Be Truly Effective in detecting Ransomware as early as possible requires coordination between all parts of the System

- Application
- File System
- Security Software
- Block Storage

Ransomware Detection With Computational Storage Devices

Compression Statistics and deltas

Encrypted payload detection

These are per IOP statistics

Chi-Squared

Changes in Read / Write Throughput

LBA Addressing and Sequencing Patterns

(kurtosis, MAD)

The types of things to detect

- Ransomware – but not just ransomware.
- Wiperware – erases data
- Mistaken deletes
- Turning encryption or compression on in the application
- Exfiltration – stealing data but not hurting it.
- Other

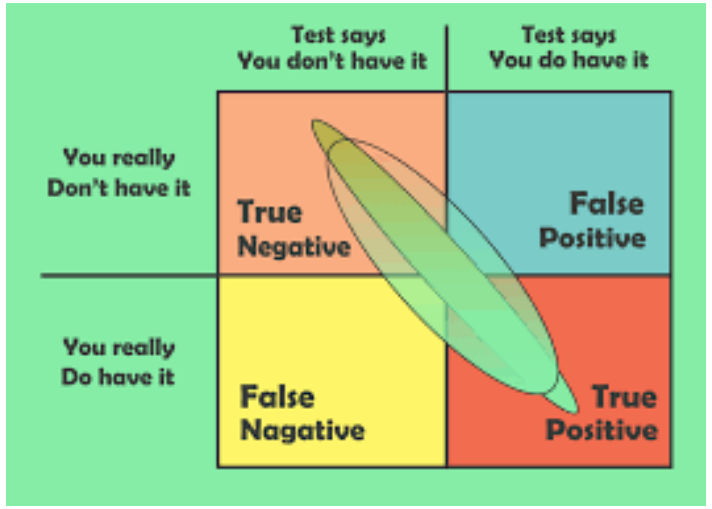
We believe Intrusion Detection is the next BIG thing in Computational Storage

How CS Devices can help

- **The CSD can collect various signals – ON THE FLY**
  - Data changes like compressibility and randomness
  - Access signals like block size and LBA, etc
- **Massive amount of telemetry data**
  - Must be summarized, grouped, preliminarily analyzed on the SSD
  - Inferencing and detailed analysis on SSD
- **For arrays of SSDs:**
  - Aggregate at system level and combine with other signals from system
  - Finer granularity grouping
- **Do supervised training with real ransomware to see how signals are affected**

False Positives are the enemy of Security alerts

- Think car alarms – annoying
- If excessive, clients will find them useless and turn them off.
- Must train in the environment of the SSD
- Must train for benign workloads that are less common
- However, False Negatives are also bad
  - Must learn
  - Being able to tune the sensitivity is good



What is the future of detection

- Getting more awareness into the block storage array
  - Volume awareness
  - VM awareness
  - Application awareness
  - File awareness
- Granularity helps with detection accuracy and isolation
- Continuous learning and upgrade

Other functions enabled by telemetry on computational storage devices

- Detailed heat information which helps reduce write amplification
  - Automatic without needing the software to note
- Colder data could use higher but slower data reduction methods
- Usage based optimizations to prepare for bursts and workloads
- Detecting and giving hints about performance anomalies

When you interact with IBM, this serves as your authorization to Flash Memory Summit or its vendor to provide your contact information to IBM in order for IBM to follow up on your interaction.

IBM's use of your contact information is governed by the IBM Privacy Policy.