Beyond PCI – A Cost Effective Approach to Data Protection

Ulf Mattsson
CTO Protegrity
Ulf.mattsson@protegrity.com

August 5, 2010
Session 7192

SHARE in Boston

Ulf Mattsson

- 20 years with IBM Software Development
  - Received US Green Card ‘EB 11 – Individual of Extraordinary Ability’ endorsed by IBM Research
- Inventor of 21 Patents
  - Encryption Key Management, Policy Driven Data Encryption, Distributed Tokenization and Intrusion Prevention
- Research member of the International Federation for Information Processing (IFIP) WG 11.3 Data and Application Security
- Created the Architecture of the Protegrity Database Security Technology
- Received Industry's 2008 Most Valuable Performers (MVP) award together with technology leaders from IBM, Google, Cisco, Ingres and other leading companies

IBM Systems magazine

ADMINISTRATOR | TRENDS | TIPS & TECHNIQUES | CASE STUDIES | STORAGE

Administrator > Security

e-Newsletter Exclusive

Demystifying Data Security

Common-sense policies and proper tools negate security concerns

March 2010 | by Ulf Mattsson

z/JOURNAL

about us

ISSUES

z/RESOURCES

- Mainframe Buyer's Guide
- Mainframe Jobs
- Industry News
- White Papers
- Subscribe

MAINFRAME

::

Dealing With Data on the Mainframe

by Ulf T. Mattsson

> email article
> print-friendly

Data wasn't always sexy, dangerous, and madly desired by millions. Not too long ago, the only people who were regularly mucking around with databases

IBM DATABASE MAGAZINE » Powering

Home DB2 Informix U2 Application Deve

Bookstore Communi

Protecting DB2 Data

Your company's data is one of its most precious resources, an from all sides. Do you know how to protect it?

By Ulf T. Mattsson

Que

IBM DATABASE MAGAZINE »

Quarter 1, 2007 Vol. 12, Issue 1

http://www.dbmag.intelligententerprise.com/story/s

Data Security Beyond Regulatory Compliance

By Ulf Mattsson

Protecting sensitive da

In increasingly comp
approach can deal with all the new

SHARE in Boston

IBM

Demystifying Mainframe Data Security

September 23, 2009

Ulf Mattsson, *CTO Protegrity*

Gayathiri Chandran, *DB2 z/OS Security, IBM Silicon Valley Lab*

San Francisco Area DB2 Users Group

MS OS/390 & z/OS WebSphere wireless resources events services

San Franc

The next m
User Group
Wednesday

Register: click here

Topics: Wednesday September
"SOX, SOA, and SarbOx and IT

A Practical Solution that Facilit
Mandates and

Speaker: Ulf T. Mattsson – C

LA Area DB2 Users Group

Risk Management:
Understanding the New Options in Data Protection for DB2 and Files

Ulf Mattsson, *CTO Protegrity*

Ulf Mattsson, CTO, Protegrity Corporation

June 4, 2009

MasterCard Worldwide

How to Evaluate Encryption Technologies

Achieving PCI Compliance & Protecting Cardholder Data

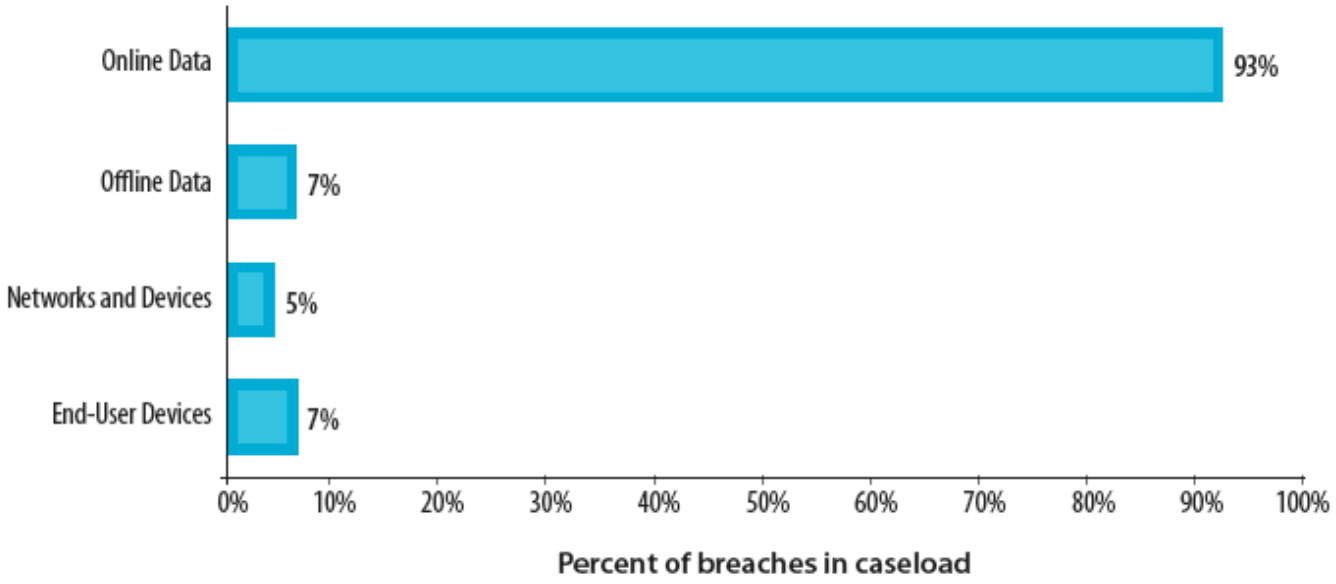
Agenda

- Review trends in data security threats
- Present case studies - protecting PCI and PII data
- Position different data security options
- Discuss how to protect the entire data flow
- Present a risk adjusted approach to data security
- Discuss data security in cloud and test environments

Online Data Under Attack – Not Laptops or Backup

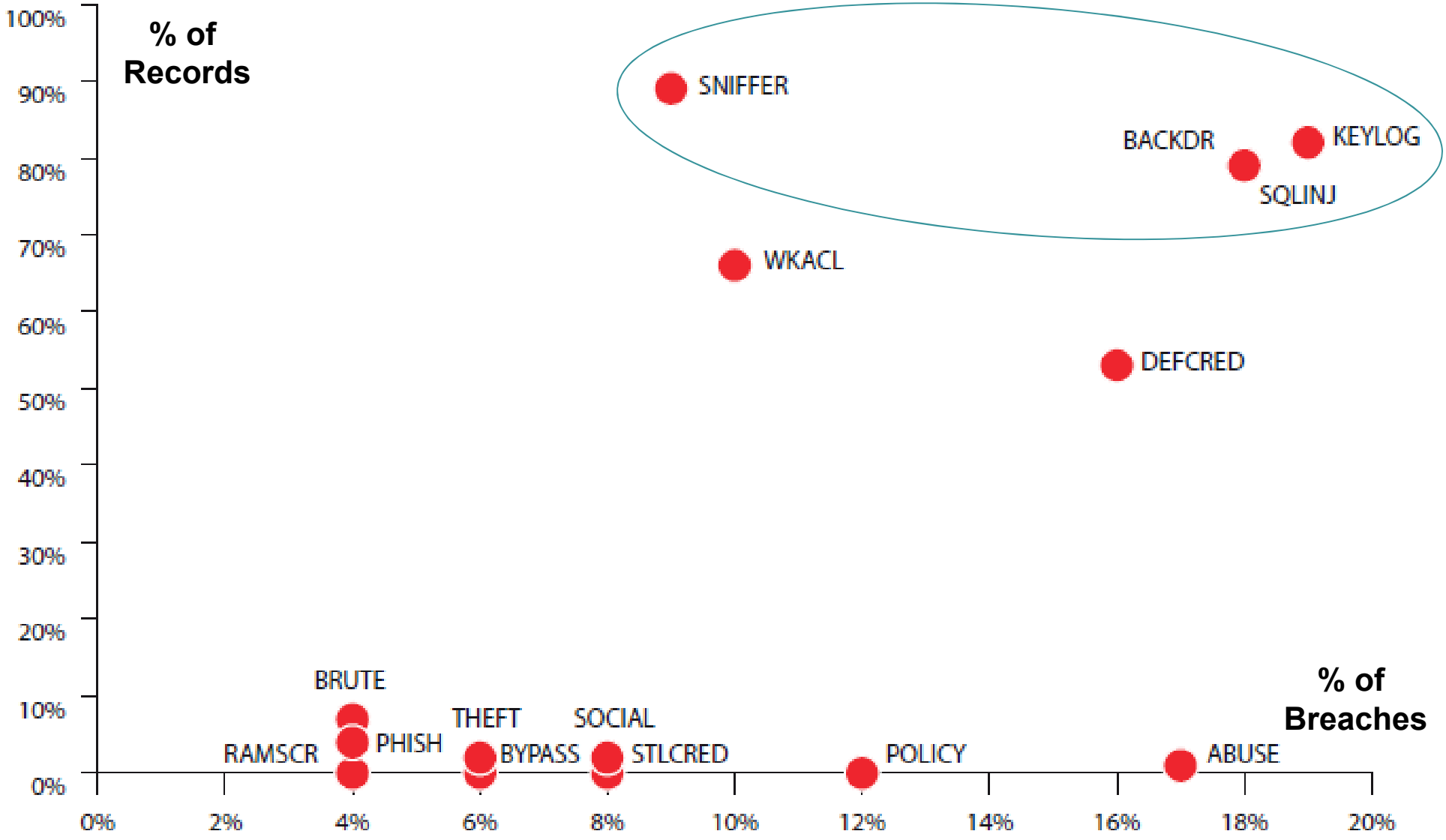
Breaches attributed to insiders are much larger than those caused by outsiders

The type of asset compromised most frequently is online data:



Slide source: Verizon Business 2008 Data Breach Investigations Report

Top 15 Threat Action Types

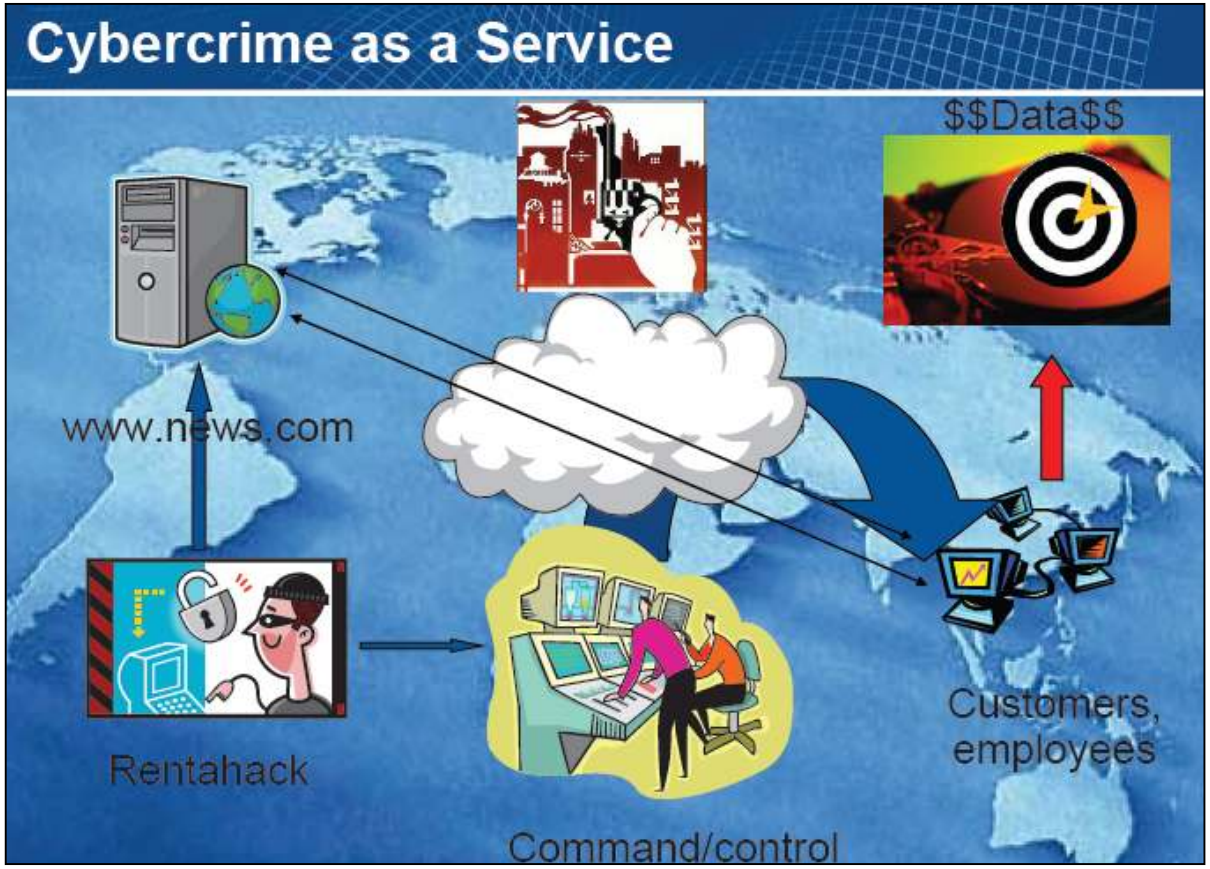


Source: 2009 Data Breach Investigations Supplemental Report, Verizon

The Gartner 2010 CyberThreat Landscape

The danger of advanced persistent threats (APTs) to enterprises.

- We have met the threat and they are us.
  - New processes
  - New technologies
  - Complacency
- You need very different armor to survive a sniper's rifle shot than you do for a hailstorm.
- Threats will always change faster than user behavior



Attacks at Different System Layers

Authorized/
Un-authorized
Users

HW Service

Contractors

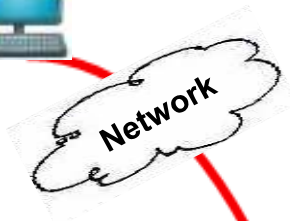
Vendors

Database Admin

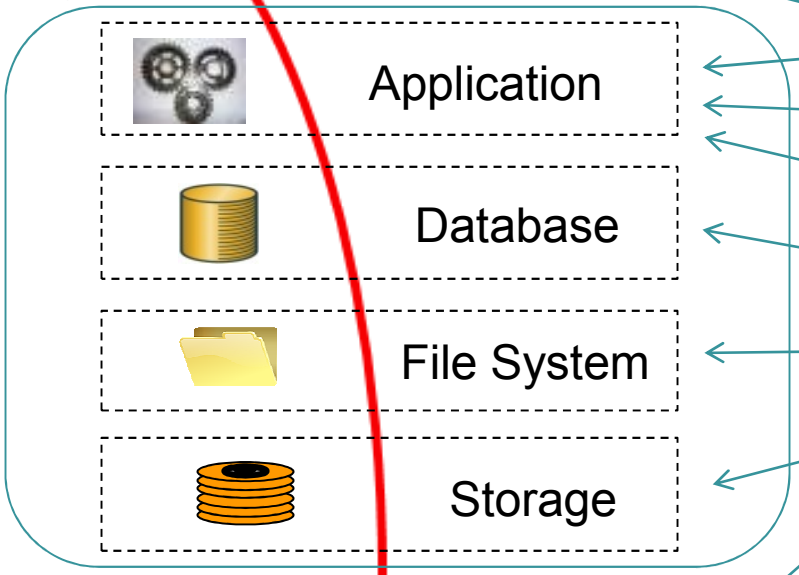
System Admin

…

Data Entry



"The perimeter is gone – need for new security approaches"



Backup

SNIFFER ATTACK

SQL INJECTION

MALWARE / TROJAN

DATABASE ATTACK

FILE ATTACK

MEDIA ATTACK

…

PCI DSS - Payment Card Industry Data Security Standard

- Applies to all organizations that hold, process, or exchange cardholder information
- A worldwide information security standard defined by the Payment Card Industry Security Standards Council (formed in 2004)
- Began as five different programs:
  - Visa Card Information Security Program, MasterCard Site Data Protection, American Express Data Security Operating Policy, Discover Information and Compliance, and the JCB Data Security Program.
- 12 requirements for compliance, organized into six logically related groups, which are called "control objectives."

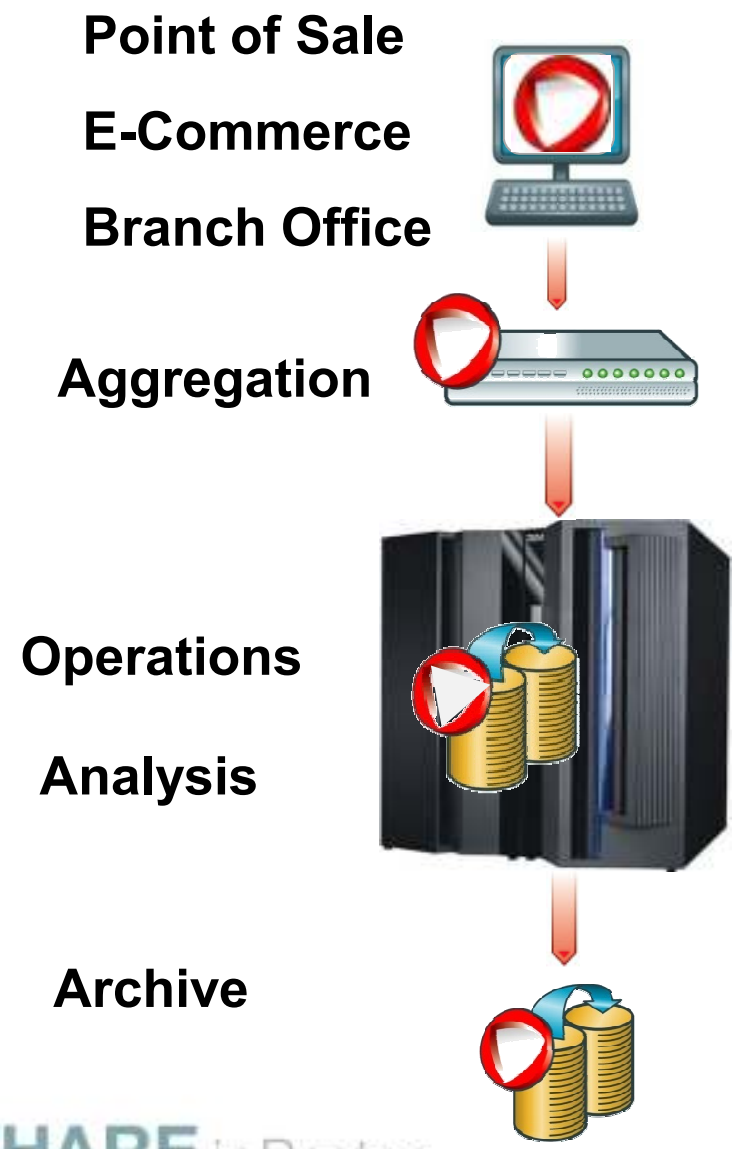
PCI DSS # 3, 6, 7, 10 & 12

Build and maintain a secure network.	1. Install and maintain a firewall configuration to protect data 2. Do not use vendor-supplied defaults for system passwords and other security parameters
Protect cardholder data.	3. Protect stored data 4. Encrypt transmission of cardholder data and sensitive information across public networks
Maintain a vulnerability management program.	5. Use and regularly update anti-virus software 6. Develop and maintain secure systems and applications
Implement strong access control measures.	7. Restrict access to data by business need-to-know 8. Assign a unique ID to each person with computer access 9. Restrict physical access to cardholder data
Regularly monitor and test networks.	10. Track and monitor all access to network resources and cardholder data 11. Regularly test security systems and processes
Maintain an information security policy.	12. Maintain a policy that addresses information security

PCI DSS #3 & 4 – Protect Cardholder Data

- 3.4 Render PAN, at minimum, unreadable anywhere it is stored by using any of the following approaches:
  - One-way hashes based on strong cryptography
  - Truncation
  - Index tokens and pads (pads must be securely stored)
  - Strong cryptography with associated key-management processes and procedures
- 4.1 Use strong cryptography to safeguard sensitive cardholder data during transmission over open, public networks.
- Comments – Cost effective compliance
  - Encrypted PAN is always “in PCI scope”
  - Tokens can be “out of PCI scope”

Case Studies – Retail Environments



'Information in the wild'
- Short lifecycle / High risk
- Databases often found at collection points

Temporary information
- Short lifecycle / High risk
- Use the transition to re-key the locks

Operating information
- Typically 1 or more year lifecycle
- Broad and diverse computing and database environment

Decision making information
- Typically multi-year lifecycle
- High volume database analysis
- Wide internal audience with privileges

Archive
- Typically multi-year lifecycle
- Preserving the ability to retrieve the data in the future is important

: Encryption service

Case Studies – PCI DSS Compliance

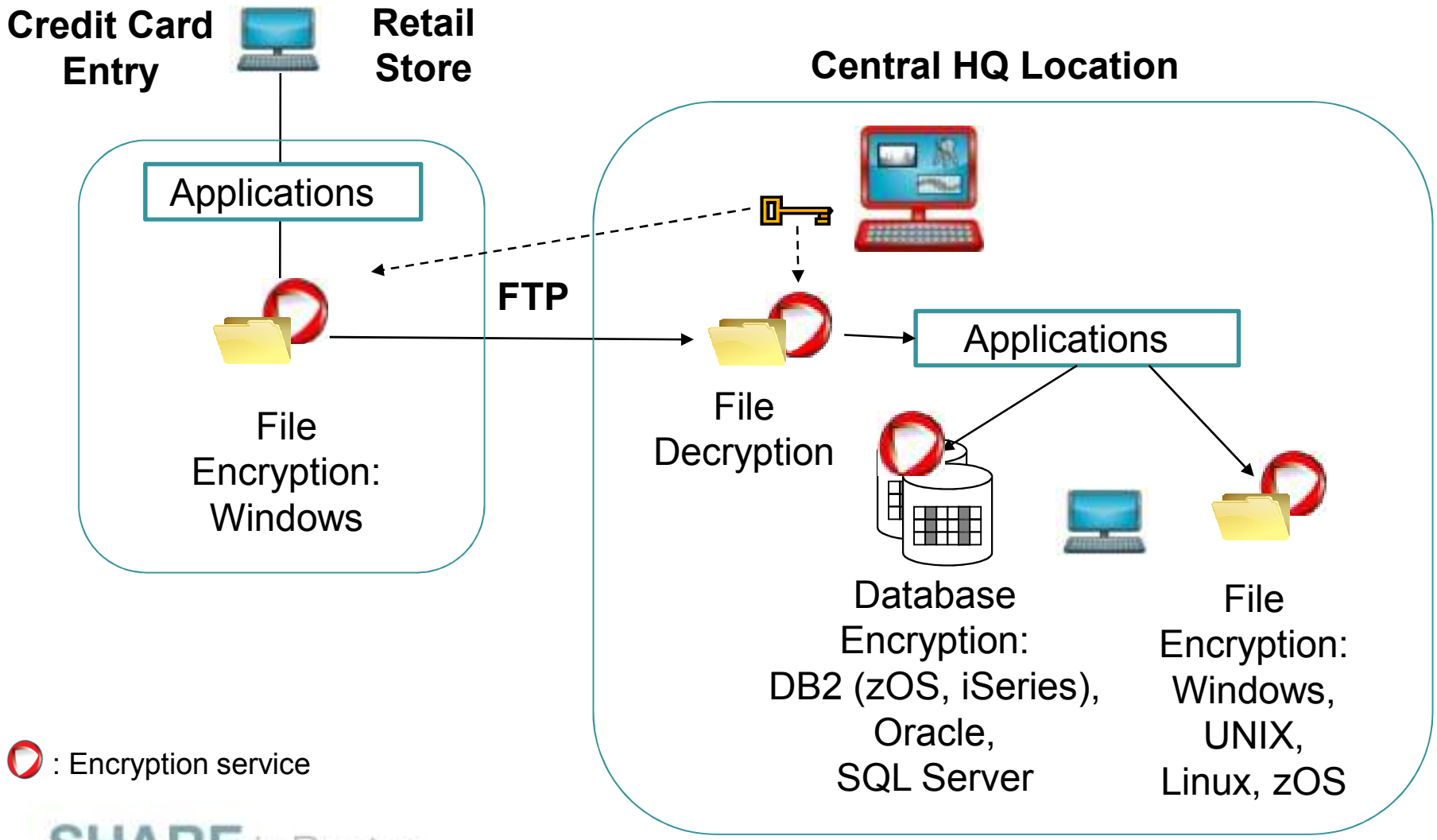
Case study #1: US Retailer

- Transparent to exiting applications
- Protect the flow of sensitive credit card information
  - From thousands of stores, Back office systems and Data warehouse
- Central key management
- Ensuring performance on the mainframe

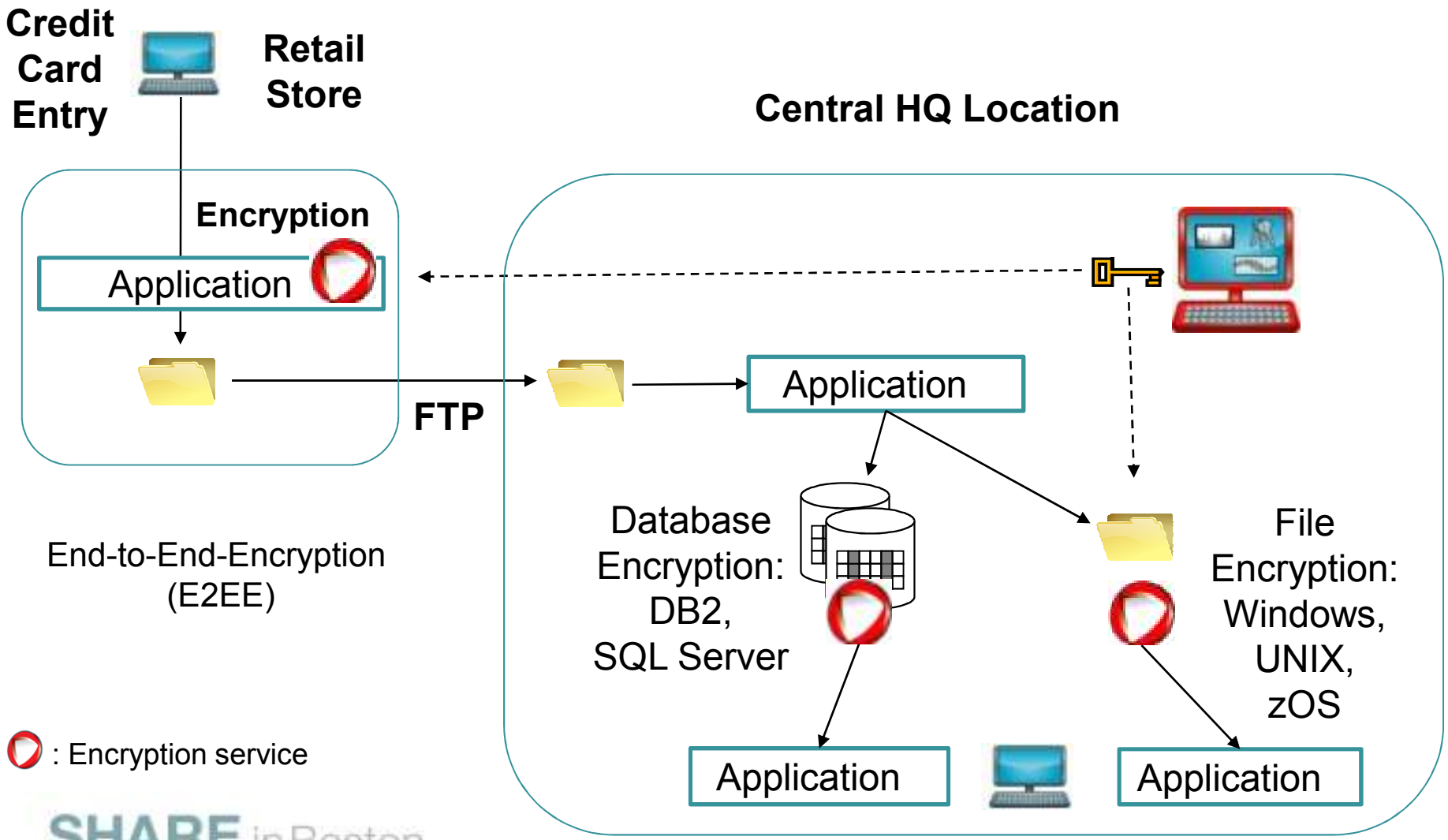
Case study #2: US Retailer

- Protection against advanced attacks
- Protect the flow of sensitive credit card information
  - From thousands of stores, Back office systems and Data warehouse
- Central key management

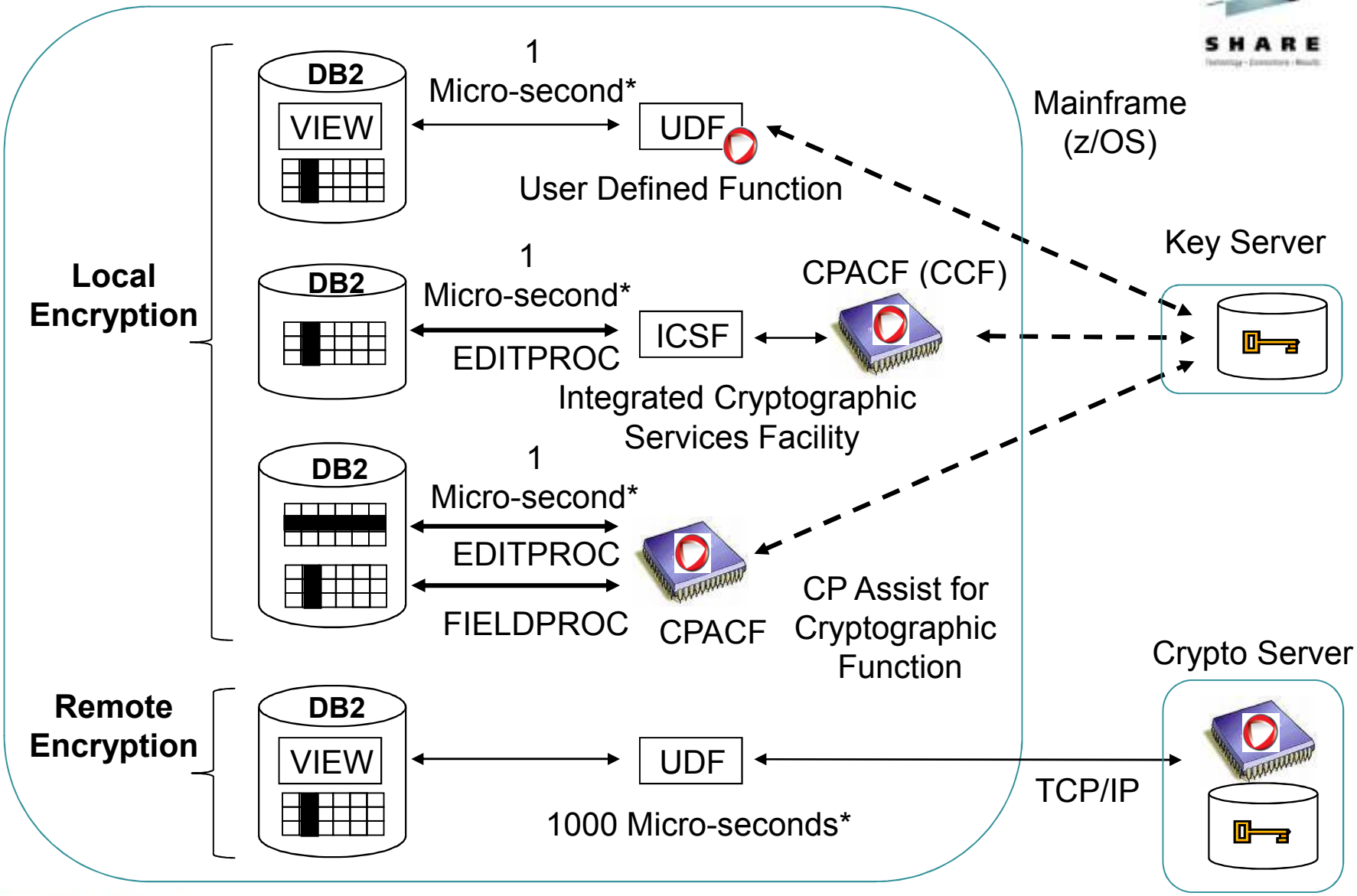
Case Study 1: Goal – PCI Compliance & Application Transparency



Case Study 2: Goal – Addressing Advanced Attacks & PCI DSS

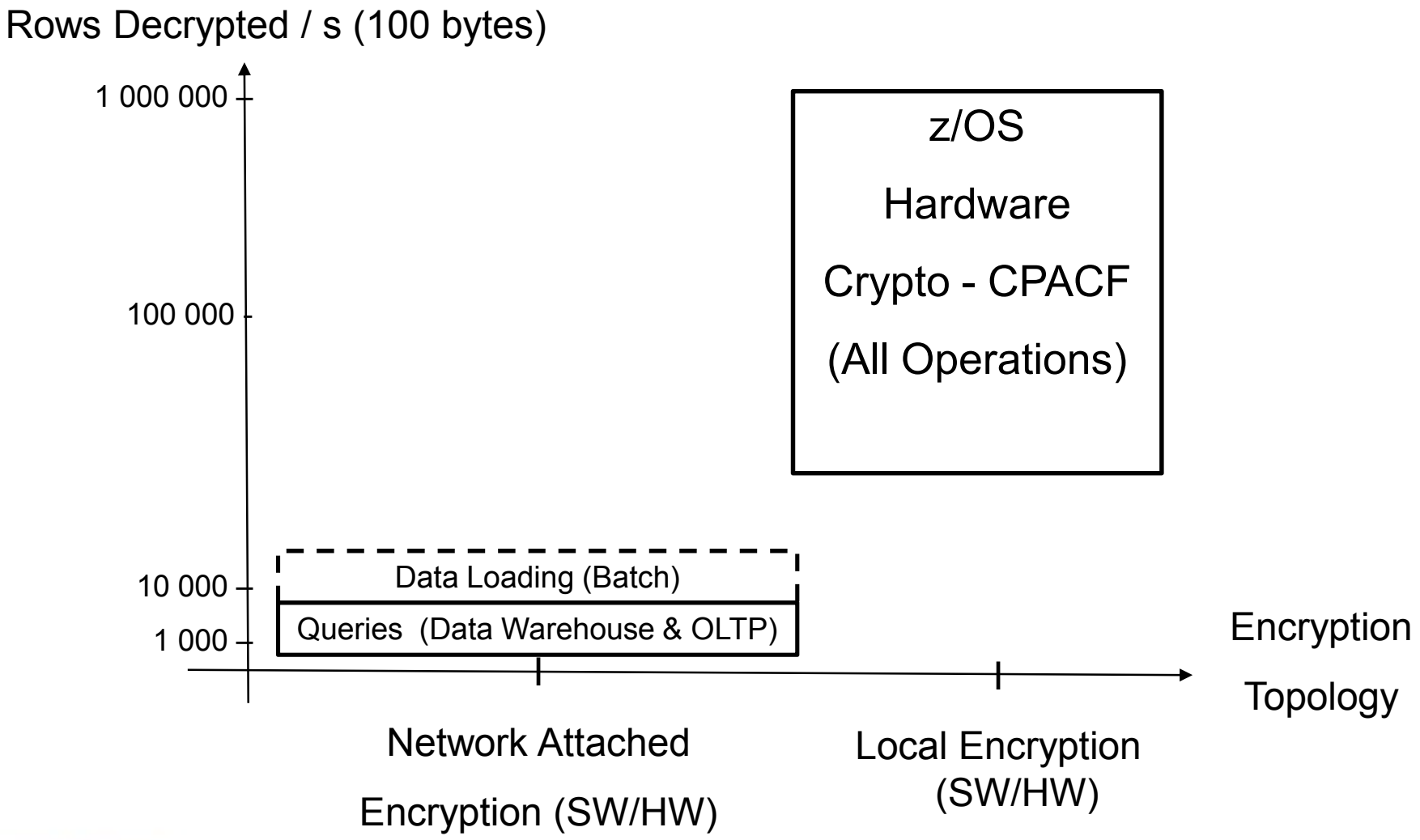


Encryption Topologies – Mainframe Example



: Encryption service * : 20 bytes

Column Encryption Performance - Different Topologies



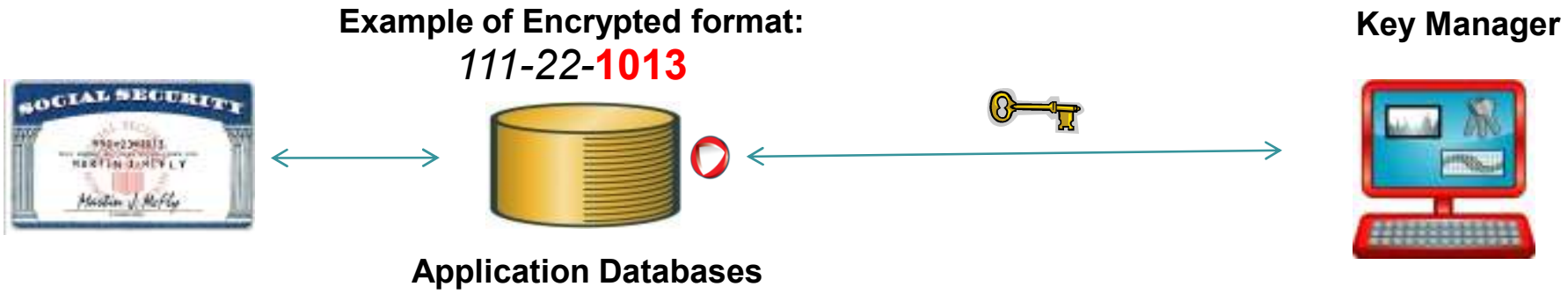
Evaluation of Encryption Options for DB2 on z/OS

Encryption Interface	Performance	PCI DSS	Security	Transparency
API				
UDF DB2 V8				
UDF DB2 V9 -				
Fieldproc				
Editproc				

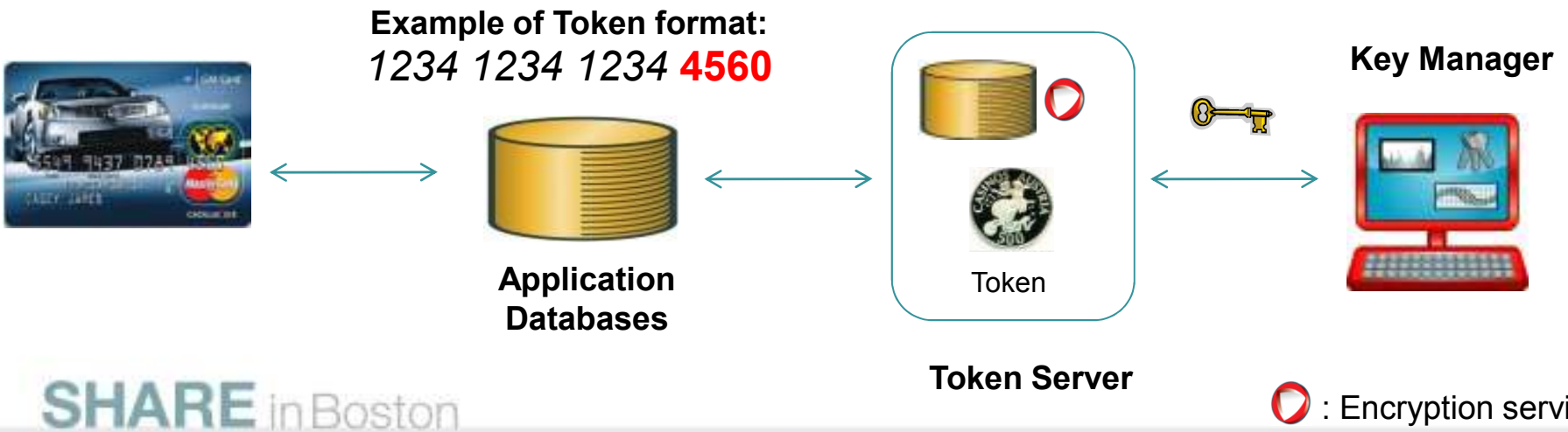
Best  Worst

Choose Your Defenses – Newer Data Security Approaches

Format Controlling Encryption



Data Tokenization



What Is Formatted Encryption?

- Where did it come from?
  - Before 2000 – Different approaches, some are based on block ciphers (AES, 3DES …)
  - Before 2005 – Used to protect data in transit within enterprises
- What exactly is it?
  - Secret key encryption algorithm operating in a new mode
  - Cipher text output can be restricted to same as input code page – some only supports numeric data
  - The new modes are not approved by NIST

Formatted Encryption - Considerations

- Unproven level of security – makes significant alterations to the standard AES algorithm
- Encryption overhead – significant CPU consumption is required to execute the cipher
- Key management – is not able to attach a key ID, making key rotation more complex - SSN
- Some implementations only support certain data (based on data size, type, etc.)
- Support for “big iron” systems – is not portable across encodings (ASCII, EBCDIC)
- Transparency – some applications need full clear text

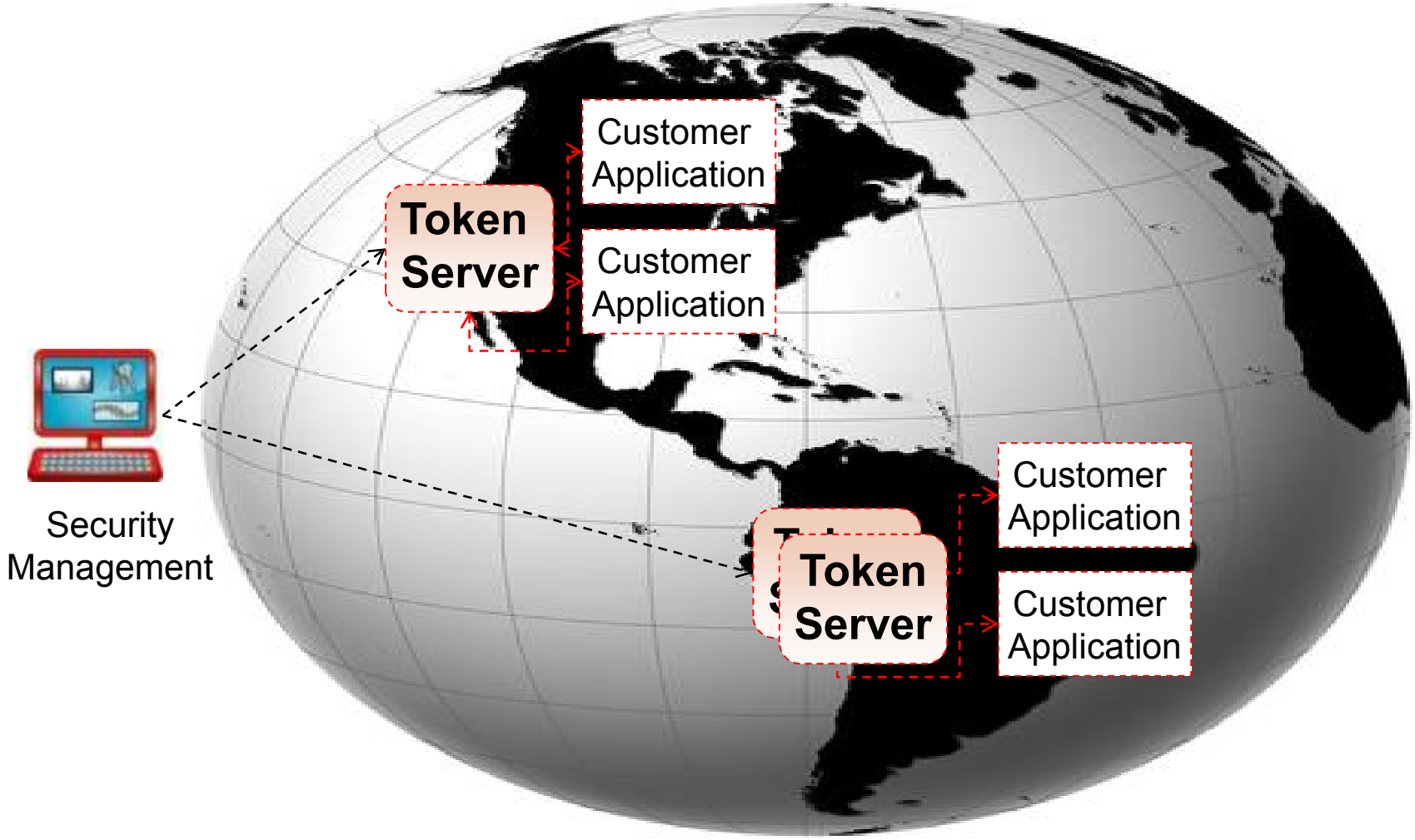
What Is Data Tokenization?

- Where did it come from?
  - Found in Vatican archives dating from the 1300s
  - In 1988 IBM introduced the Application System/400 with shadow files to preserve data length
  - In 2005 vendors introduced tokenization of account numbers
- What exactly is it?
  - It IS NOT an encryption algorithm or logarithm.
  - It generates a random replacement value which can be used to retrieve the actual data later (via a lookup)
  - Still requires strong encryption to protect the lookup table(s)

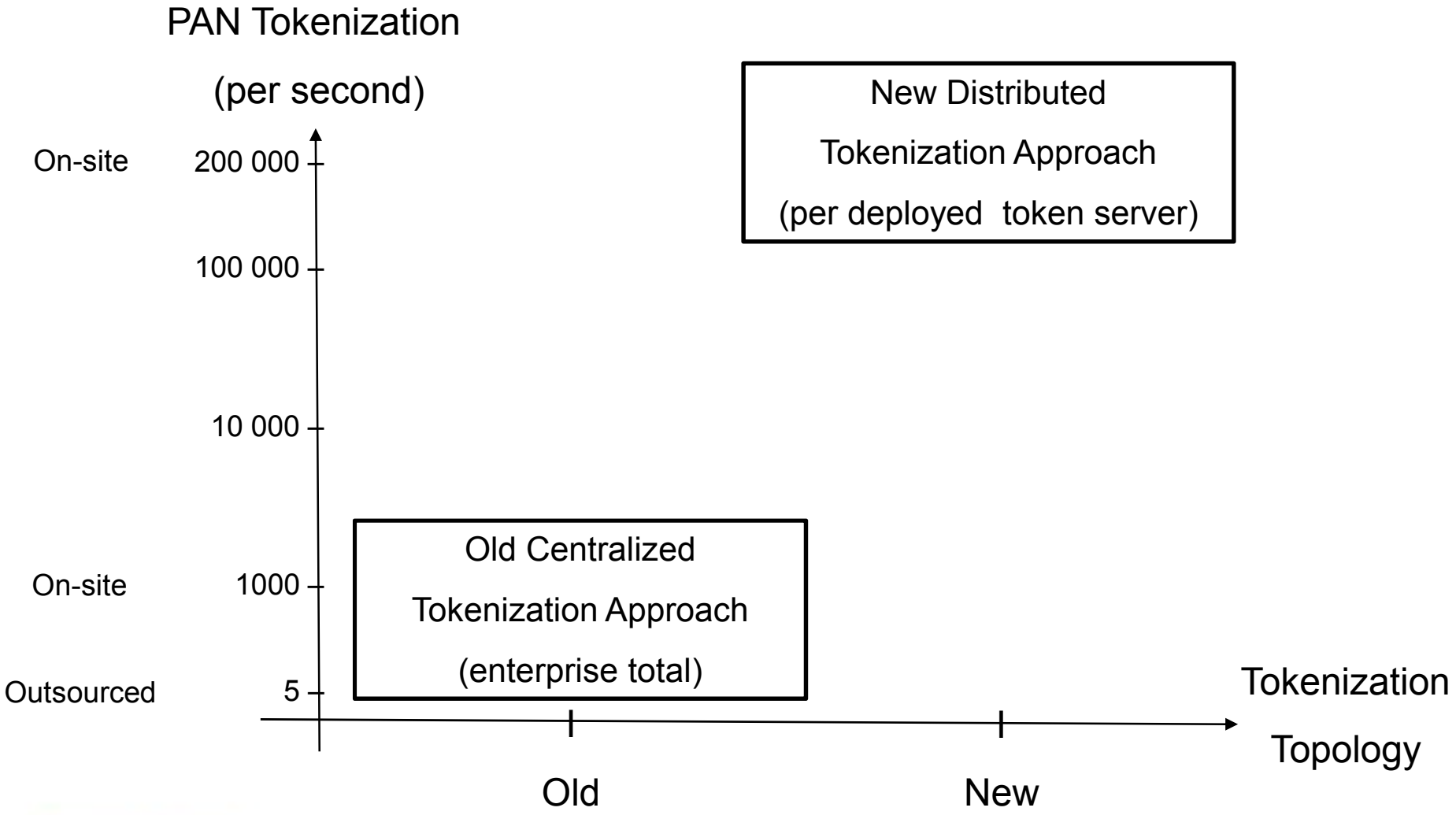
Central Tokenization - Considerations

- Transparency – not transparent to downstream systems that require the original data
- Performance & availability
  - Imposes significant overhead from the initial tokenization operation and from subsequent lookups
  - Imposes significant overhead if token server is remote or outsourced
- Security
  - Vulnerabilities of the tokens themselves – randomness and possibility of collisions
  - Vulnerabilities typical in in-house developed systems – exposing patterns and attack surfaces

New Tokenization Approach - Distributed Servers



Different Tokenization Approaches - Performance

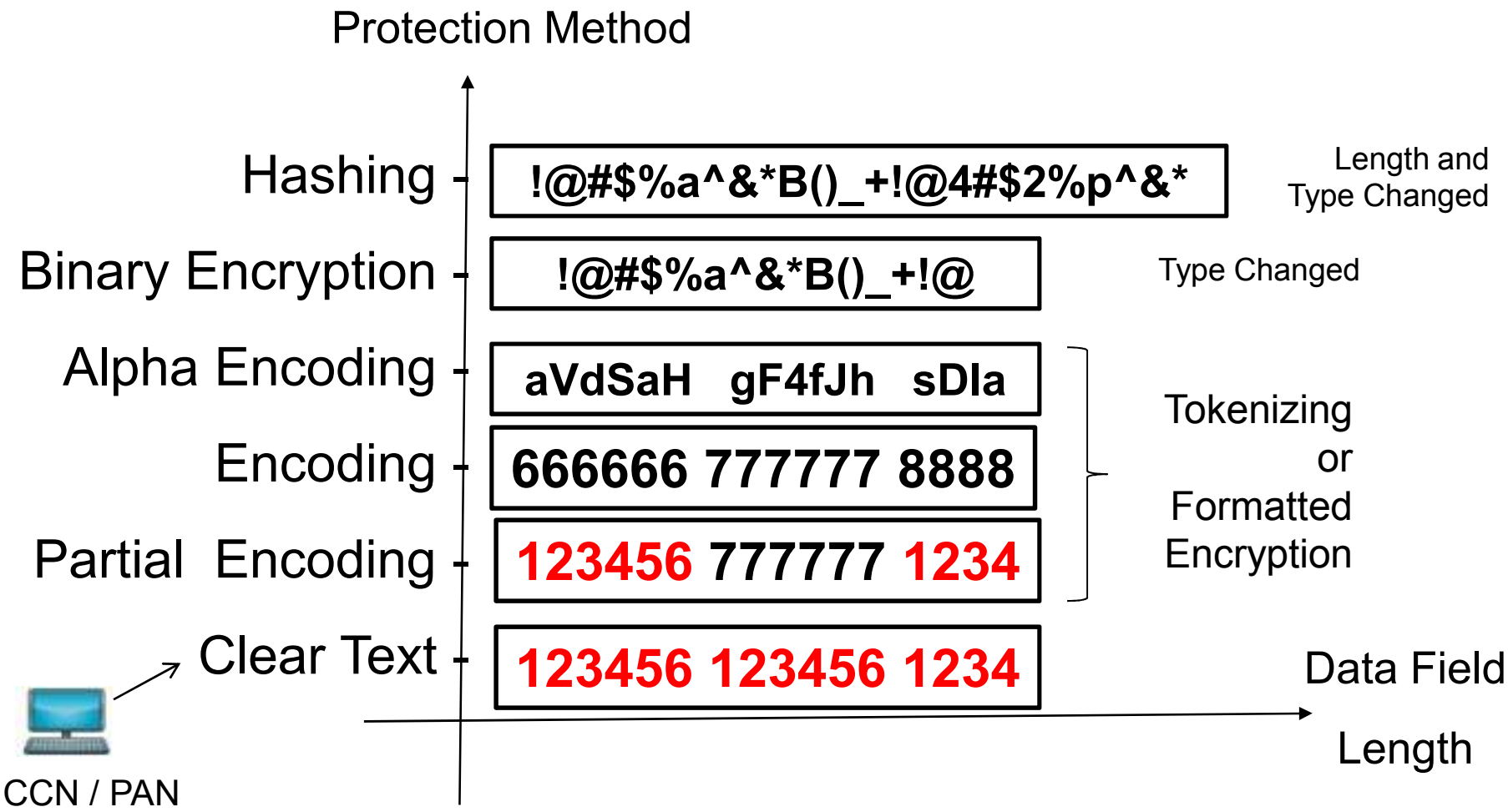


Evaluating Different Tokenization Solutions

Evaluation Area		Hosted/Outsourced		On-site/On-premises		
Area	**Criteria**	**Central (old)**	**Distributed**	**Central (old)**	**Distributed**	**Integrated**
Operational Needs	Availability	○	◔	◒	◕	●
	Scalability	○	◔	◒	◕	●
	Performance	○	◔	◒	◕	●
Pricing Model	Per Server	○	○	●	●	●
	Per Transaction	●	●	○	○	○
Data Types	Identifiable - PII	◔	◔	◒	◕	●
	Cardholder - PCI	●	●	●	●	●
Security	Separation	●	●	◕	◒	◔
	Compliance Scope	●	●	◕	◕	◕

Best ● ◕ ◒ ◔ ○ Worst

How to not Break the Data Format



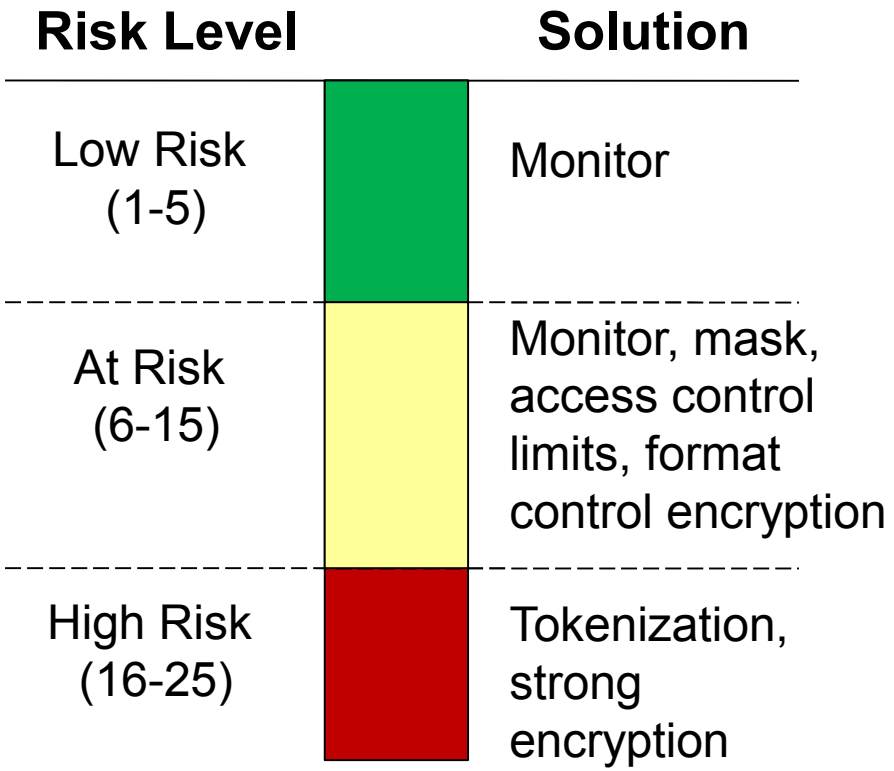
Different Security Options for Data Fields

Evaluation Criteria	Strong Encryption	Formatted Encryption	New Distributed Tokenization	Old Central Tokenization
Disconnected environments	●	●	●	○
Distributed environments	●	●	●	◔
Performance impact – data loading	●	◔	●	◔
Transparent to applications	●	◒	◒	◒
Expanded storage size	◒	●	◔	◔
Transparent to database schema	◒	●	●	●
Long life-cycle data	◔	◔	●	●
Unix or Windows &“big iron”	●	◔	●	●
Re-keying of data in a data flow	◒	◔	●	●
High risk data	●	○	●	●
Compliance to PCI, NIST	●	○	●	●

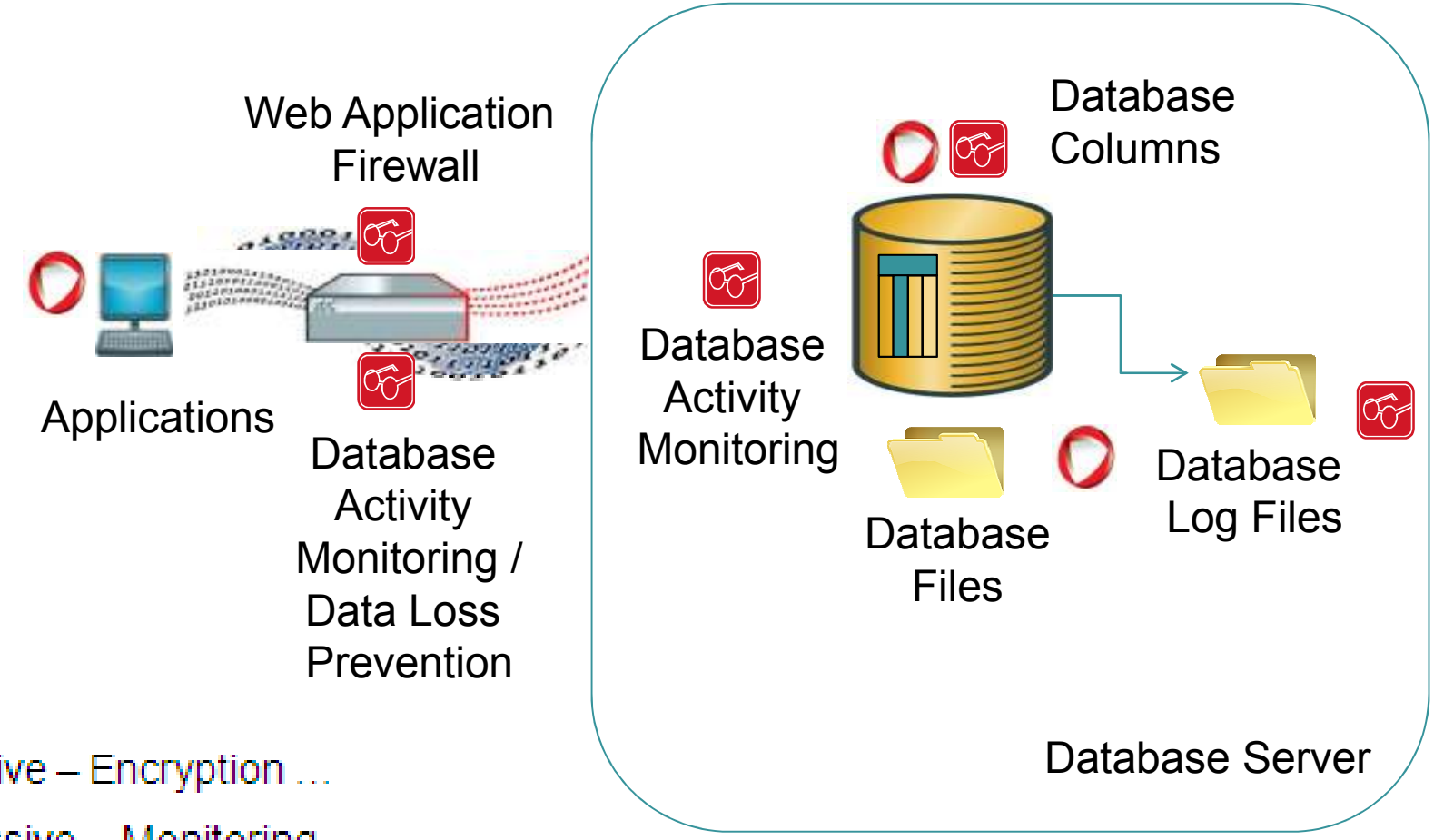
Best ● ◔ ◒ ◔ ○ Worst

Matching Data Protection Solutions with Risk Level

Data Field	Risk Level
Credit Card Number	25
Social Security Number	20
CVV	20
Customer Name	12
Secret Formula	10
Employee Name	9
Employee Health Record	6
Zip Code	3



Choose Your Defenses – A Balanced Approach



Cost Effective Technology for PCI DSS

Technologies in ascending order by average cost effectiveness rating	Pct%*
Firewalls	82%
Anti-virus & anti-malware solutions	74%
Encryption for data at rest	74%
Encryption for data in motion	71%
Access governance systems	64%
Identity & access management systems	63%
Web application firewalls (WAF)	55%
Correlation or event management systems	55%
Endpoint encryption solution	46%
Data loss prevention systems	43%
Code review	36%
Traffic intelligence systems	32%
Virtual privacy network (VPN)	26%
Intrusion detection or prevention systems	22%
Database scanning and monitoring	18%
ID & credentialing system	11%
Website sniffer or crawlers	7%
Perimeter or location surveillance systems	3%
Average	43%
Pct% defines the average percentage of respondents rating the technology as highly cost effective.	

Encryption 74%

WAF 55%

DLP 43%

DAM 18%

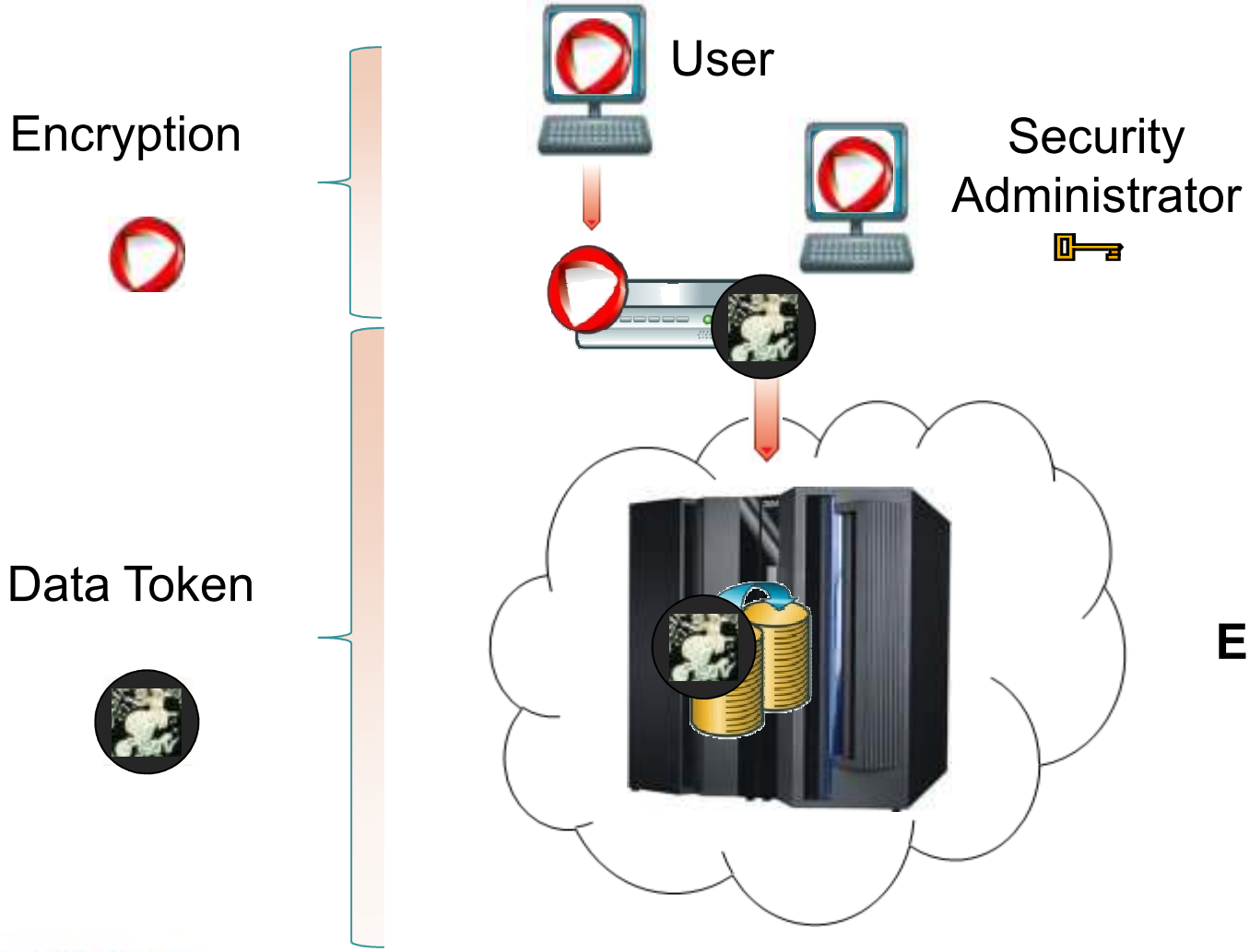
Source: 2009 PCI DSS Compliance Survey, Ponemon Institute

Choose Your Defenses – Positioning of Alternatives

Database Protection Approach	Performance	Storage	Availability	Transparency	Security
Monitoring, Blocking, Masking	●	●	●	●	○
Column Level Formatted Encryption	◒	●	●	◒	◔
Column Level Strong Encryption	◔	◕	●	◔	◕
Distributed Tokenization	◔	◕	●	◒	●
Central Tokenization	○	◔	○	◒	◒
Database File Encryption	◔	●	●	●	◔

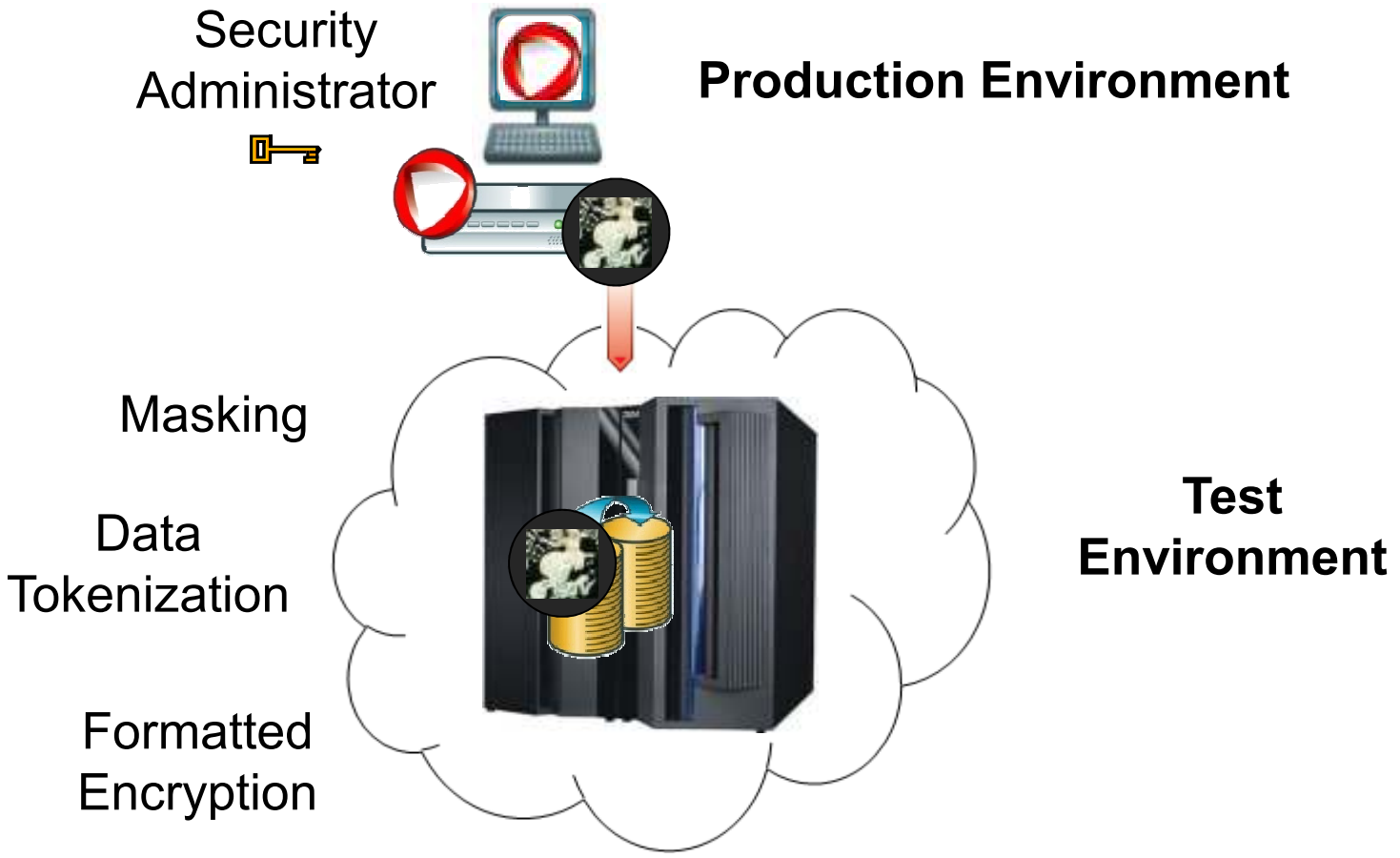
Best ● ◔ ◒ ◔ ○ Worst

Use Case –Data Protection in Cloud Environments



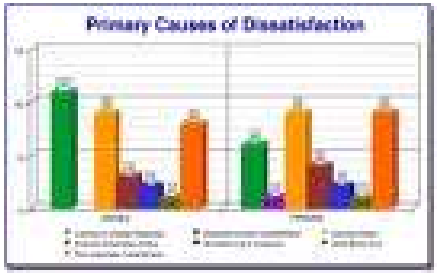
Token Encryption

Use Case – Data Protection in Test/Dev Environments

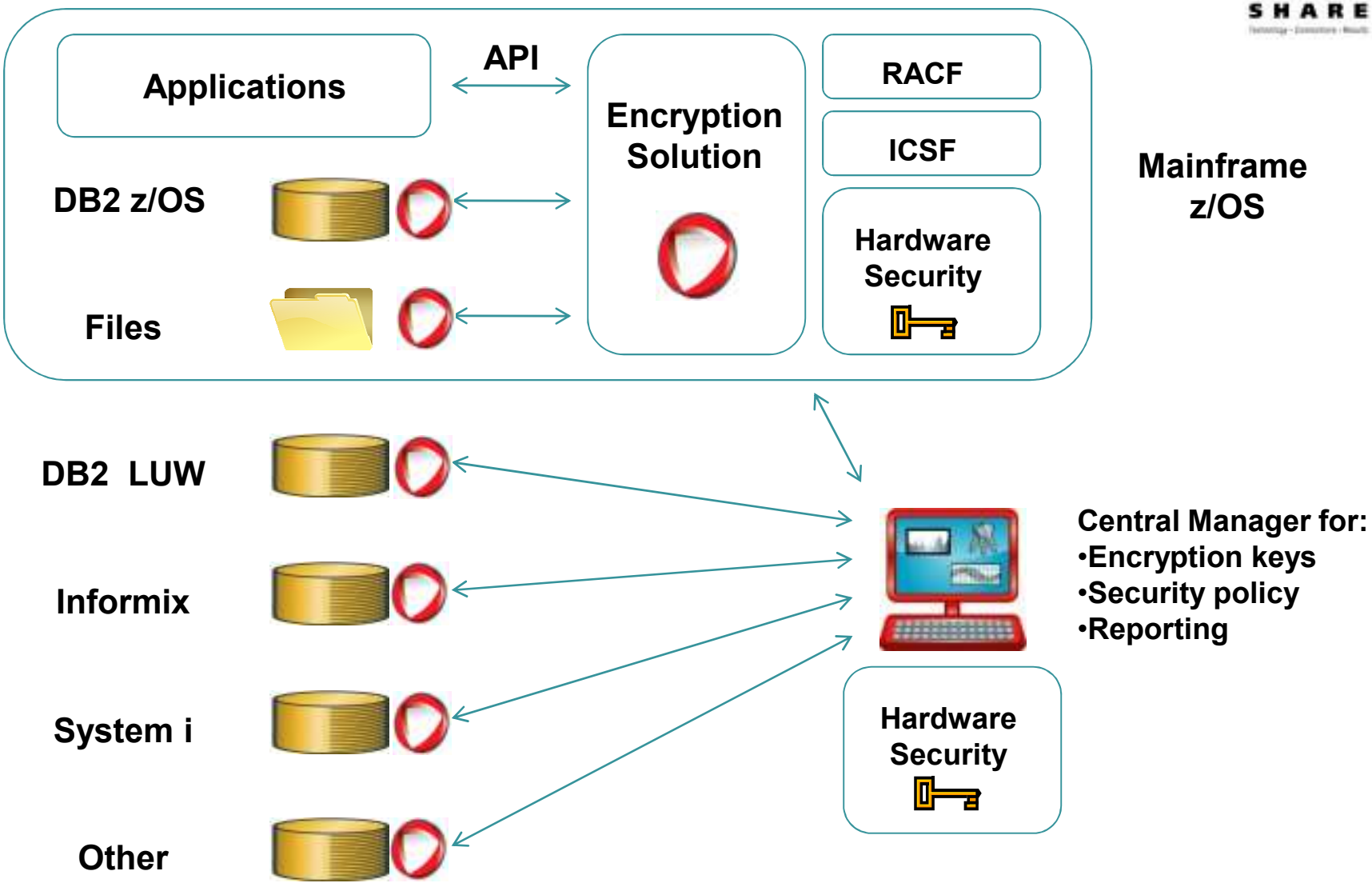


Data Protection Challenges

- Actual protection is not the challenge
- Management of solutions
  - Key management
  - Security policy
  - Auditing and reporting
- Minimizing impact on business operations
  - Transparency
  - Performance vs. security
- Minimizing the cost implications
- Maintaining compliance
- Implementation Time



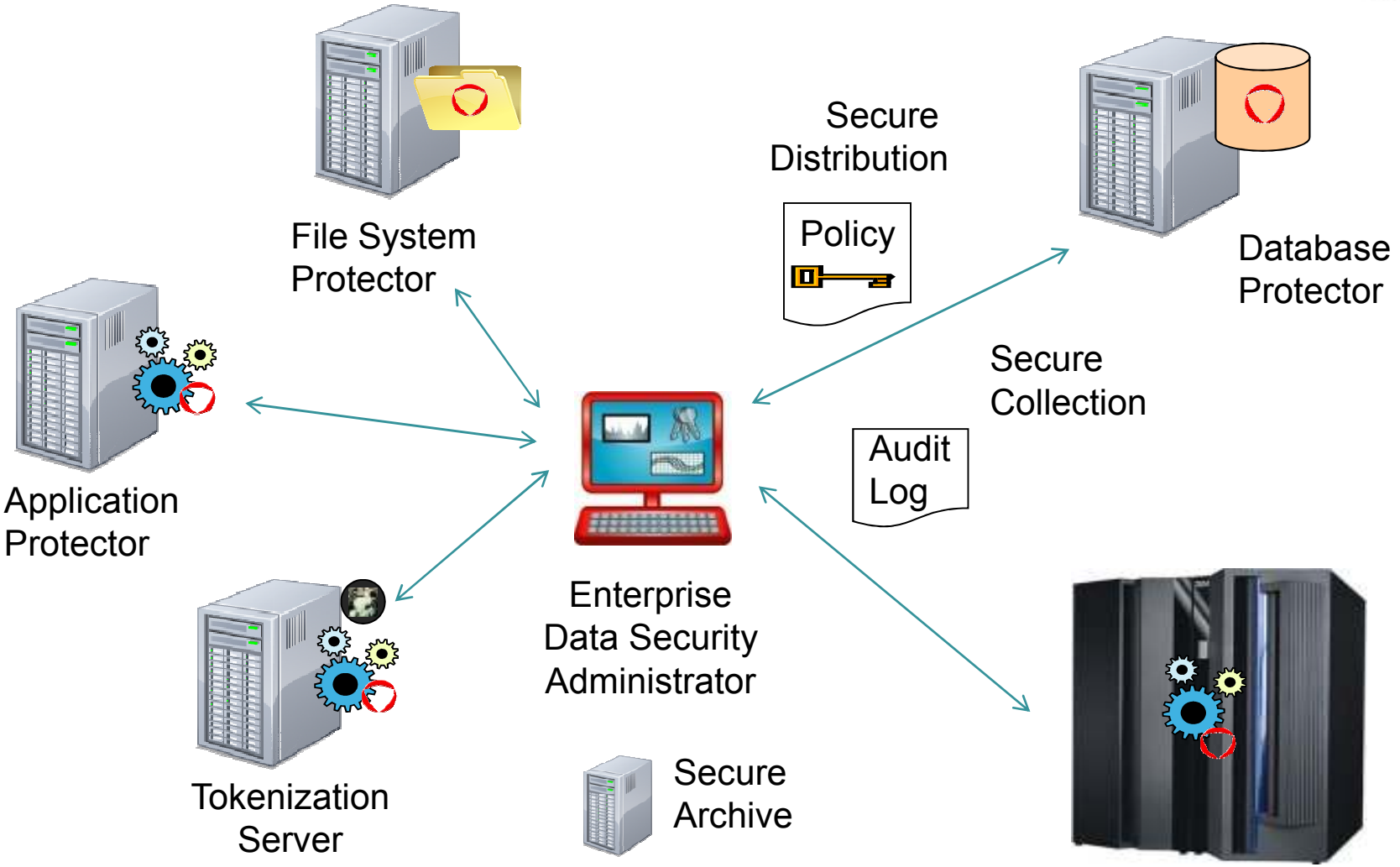
Single Point of Control for Data Encryption



Summary

- New threats to data & new regulations
- New “best practices” for data protection
- New approaches for data protection
- Protect the data flow
- Risk-adjusted approach to data security
- Centralized key management, policy and reporting

Protegrity Data Security Management



Protegrity Corporate Overview

- Enterprise Data Security Management
- Founded 1996
- 300+ customers
- Market leader in PCI DSS & PII data security
- 14 patents granted/issued
- Global reach - 60% NA, 30% EMEA, 10% Asia

Beyond PCI – A Cost Effective Approach to Data Protection

Ulf Mattsson
CTO Protegrity
Ulf.mattsson@protegrity.com

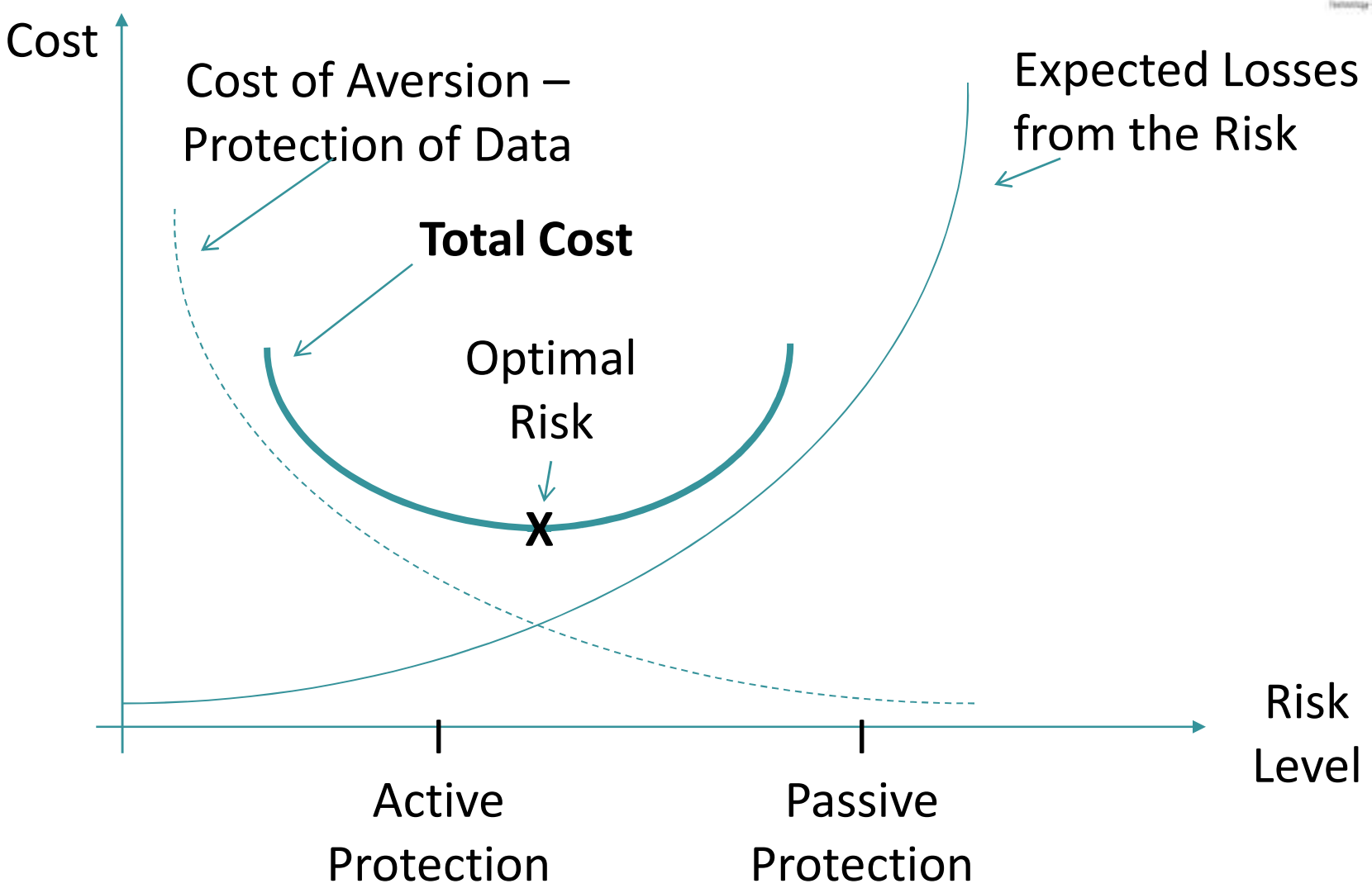
August 5, 2010
Session 7192

Appendix

A Risk-adjusted Data Security Plan

1. Know Your Data
2. Find Your Data
3. Understand Your Enemy
4. Choose Your Defenses
5. Deploy Defenses
6. Crunch the Numbers

Choose Your Defenses – Find the Balance

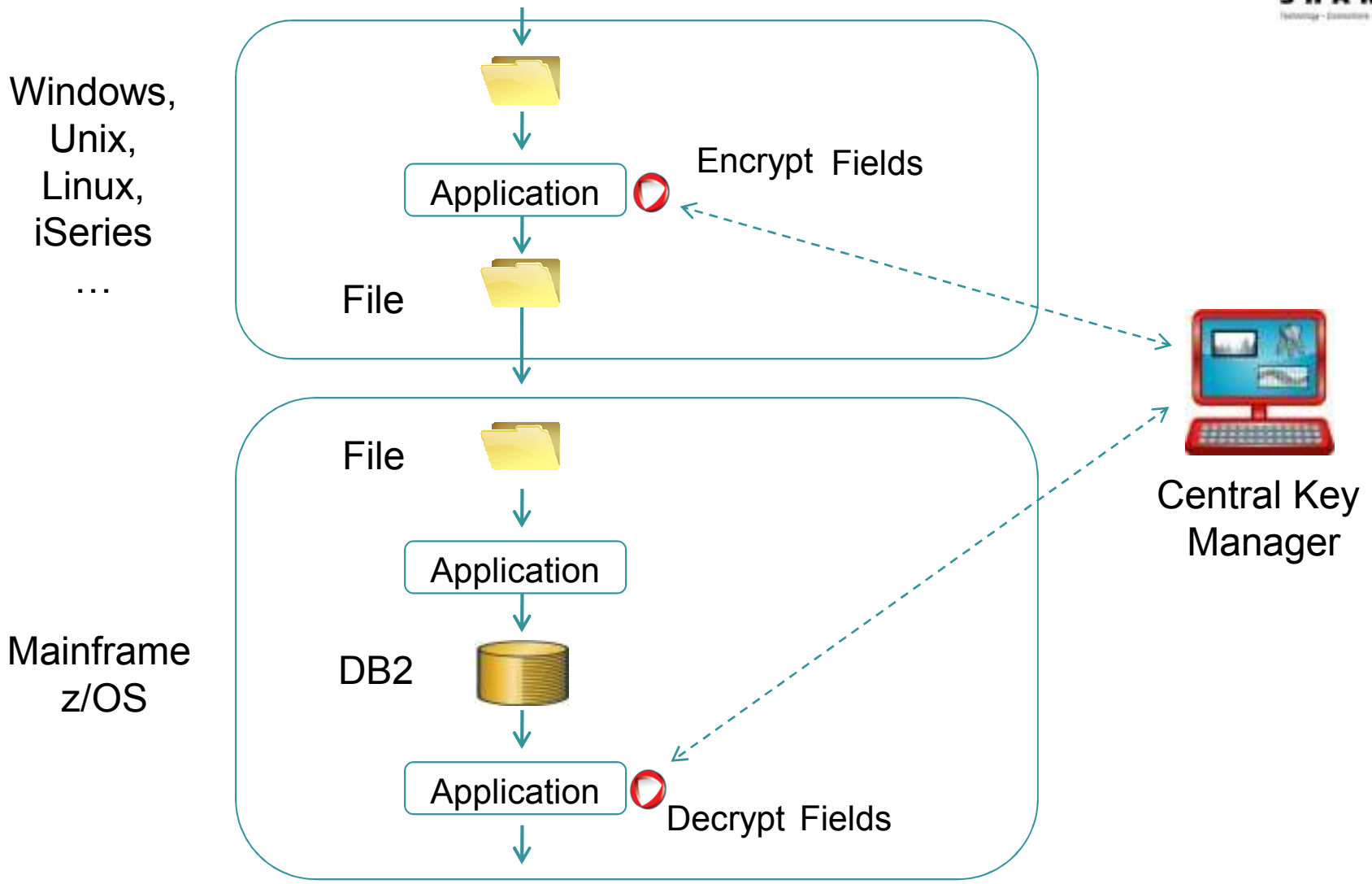


Know Your Data – Identify High Risk Data

- Begin by determining the risk profile of all relevant data collected and stored
  - Data that is resalable for a profit
  - Value of the information to your organization
  - Anticipated cost of its exposure

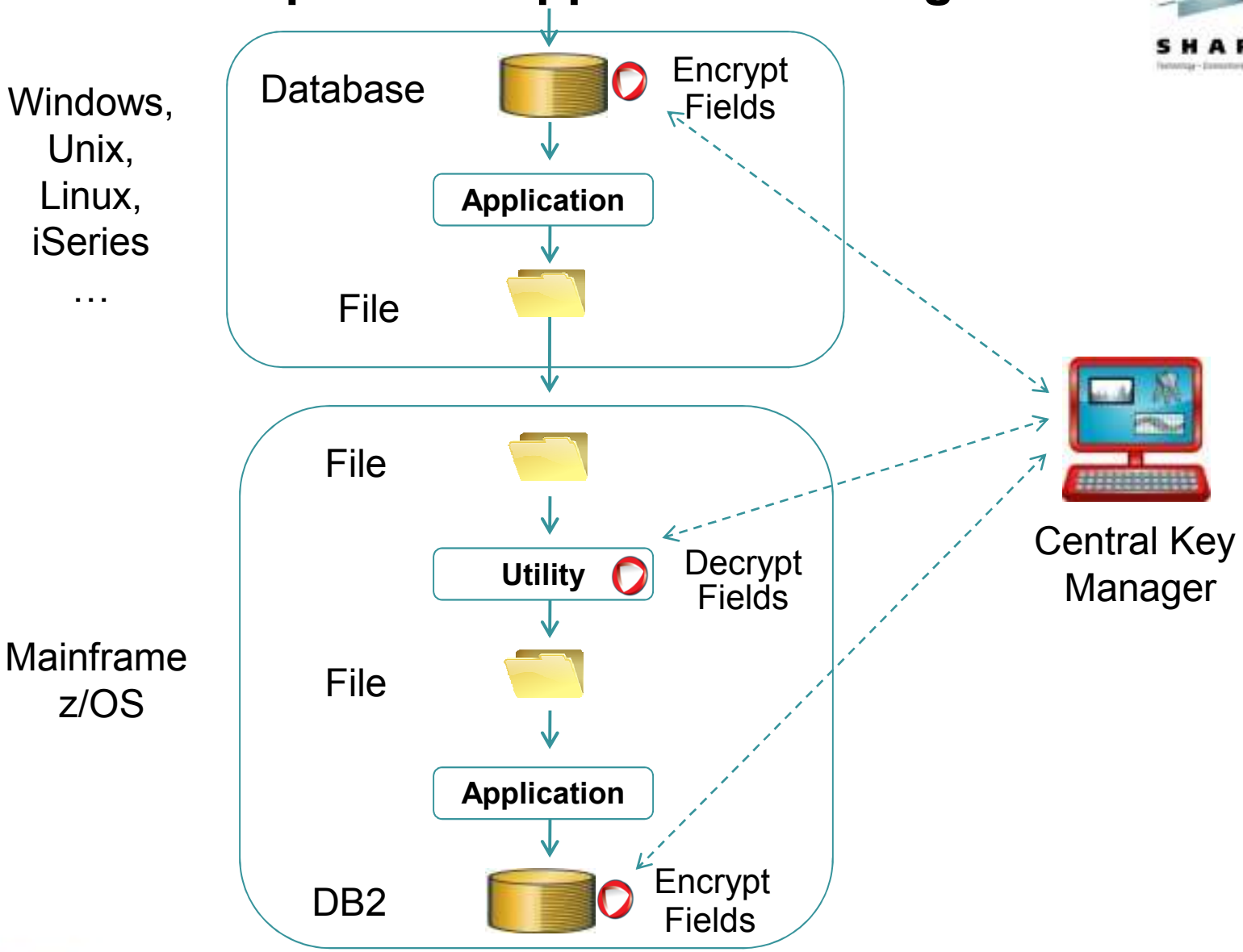
Data Field	Risk Level
Credit Card Number	25
Social Security Number	20
CVV	20
Customer Name	12
Secret Formula	10
Employee Name	9
Employee Health Record	6
Zip Code	3

Example - Protecting the Data Flow



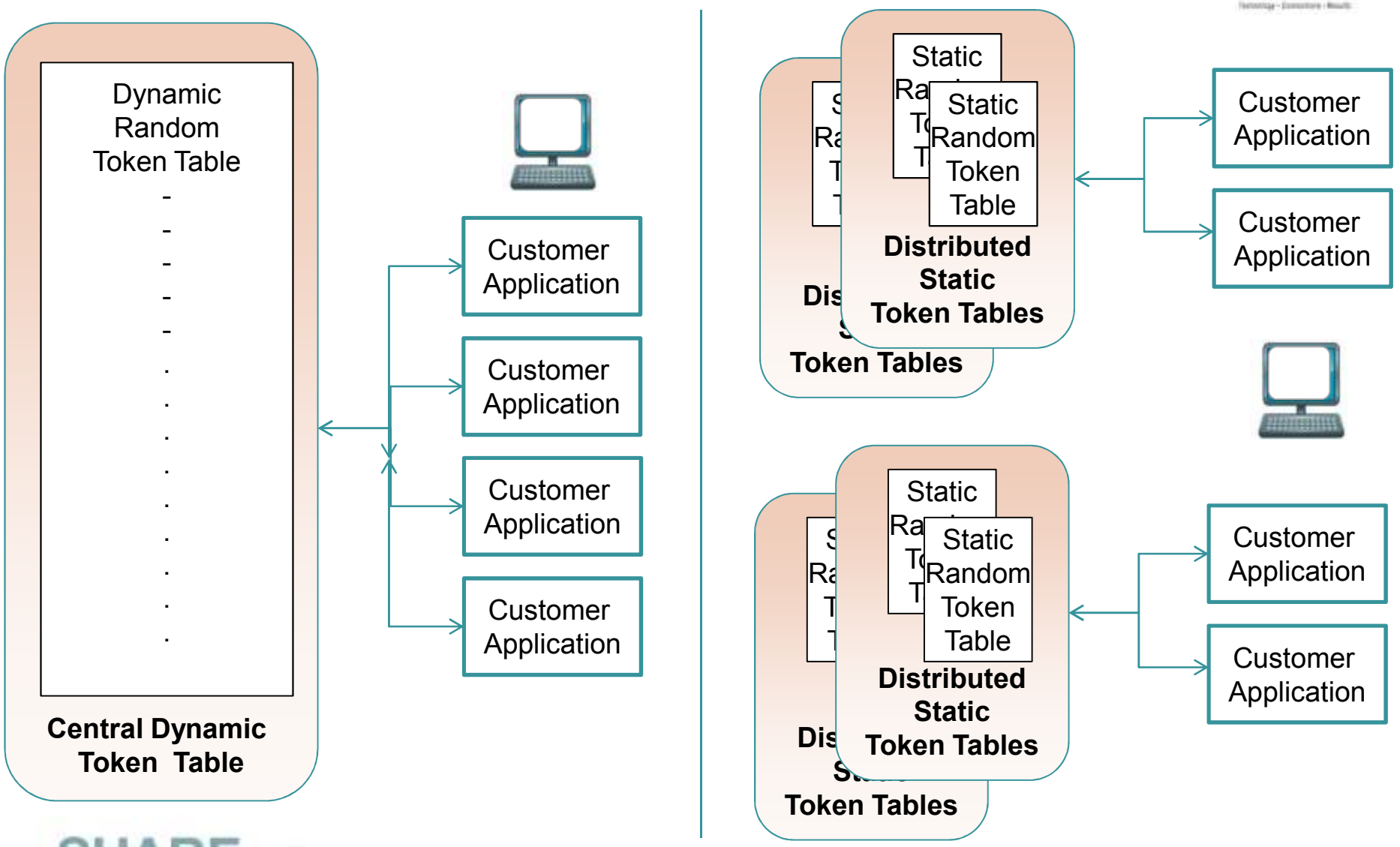
: Encryption service

Data Flow Example – No Application Changes

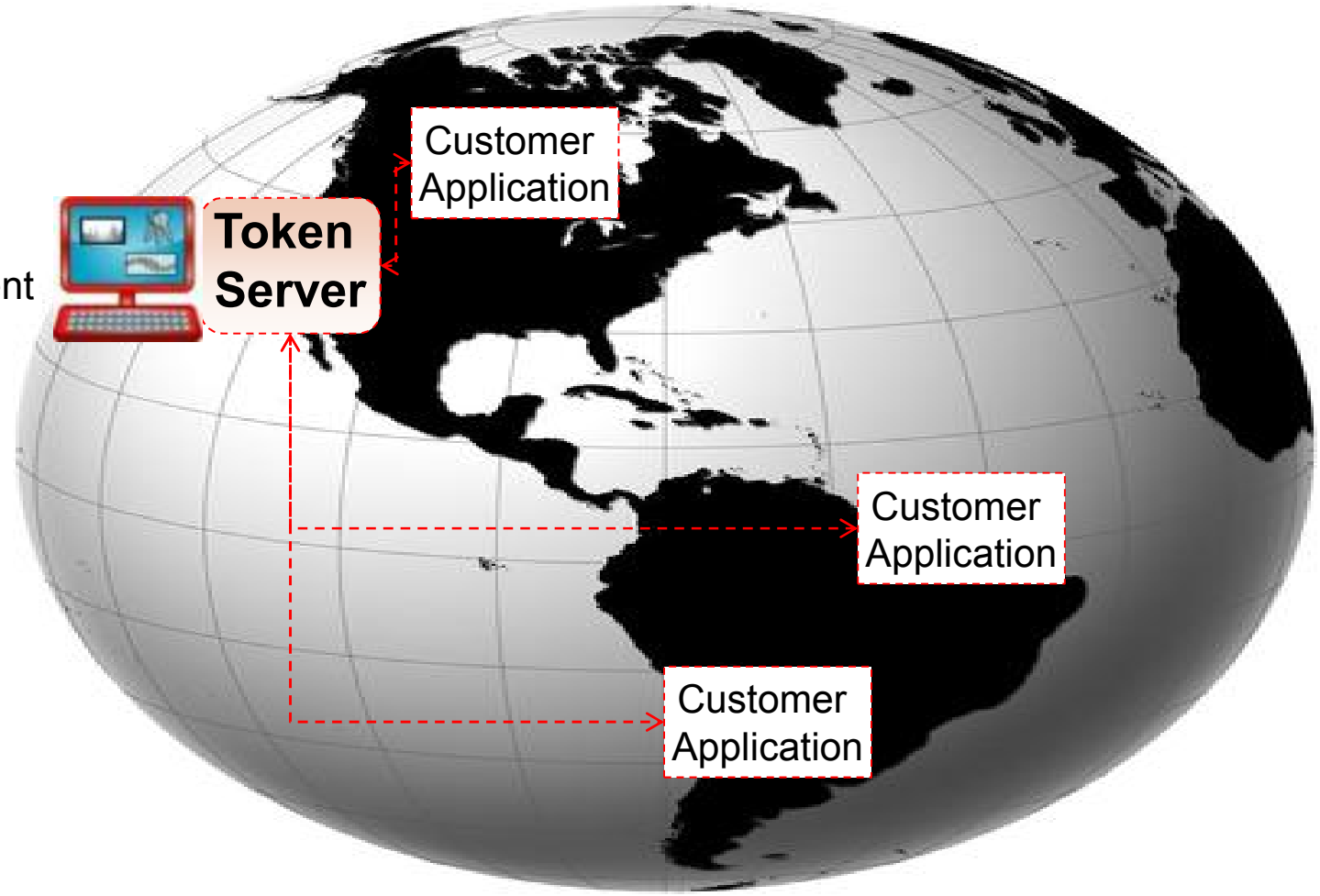


: Encryption service

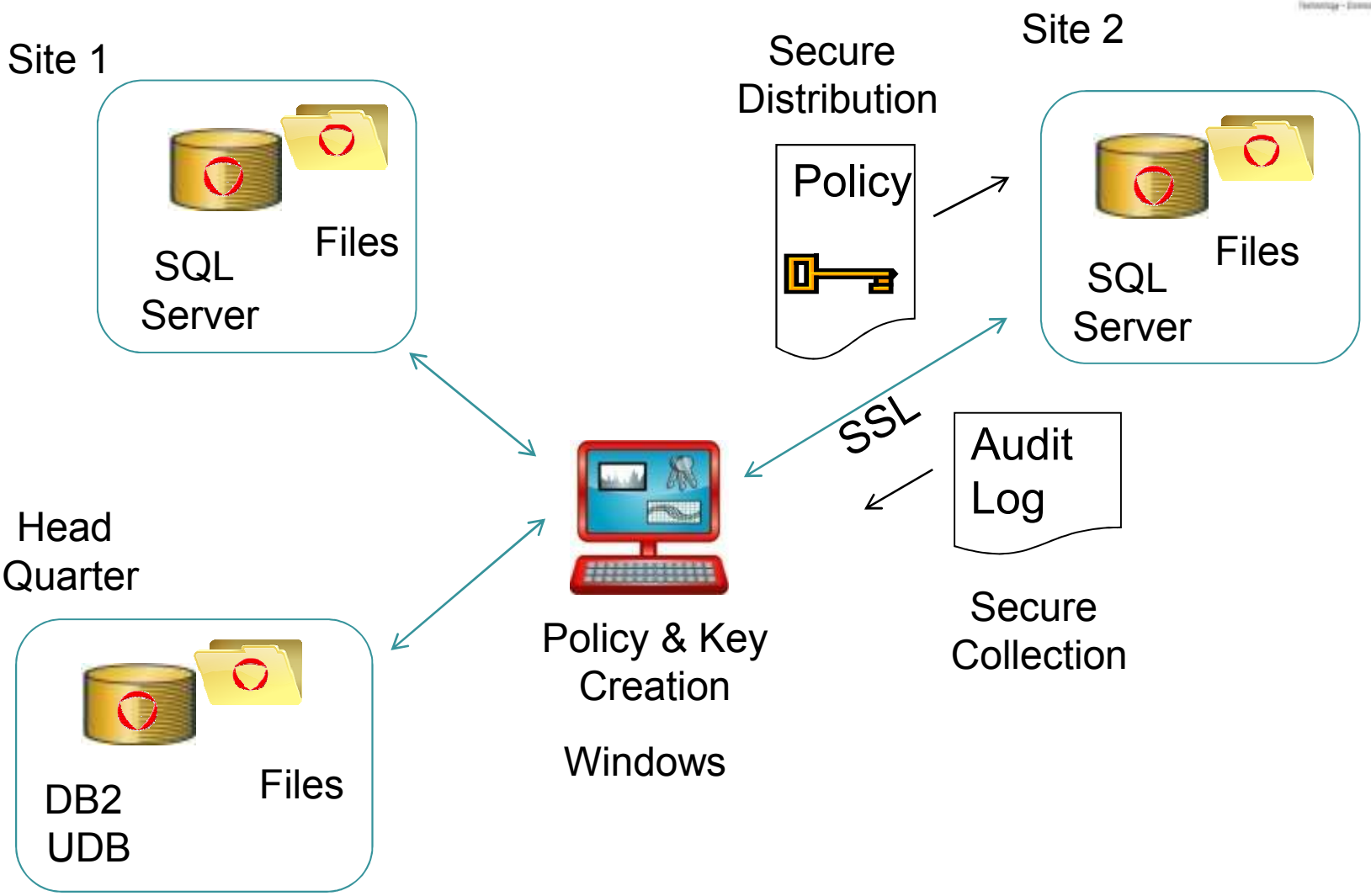
Tokenization – Central vs. Distributed



Old Tokenization Approach - One Central Server

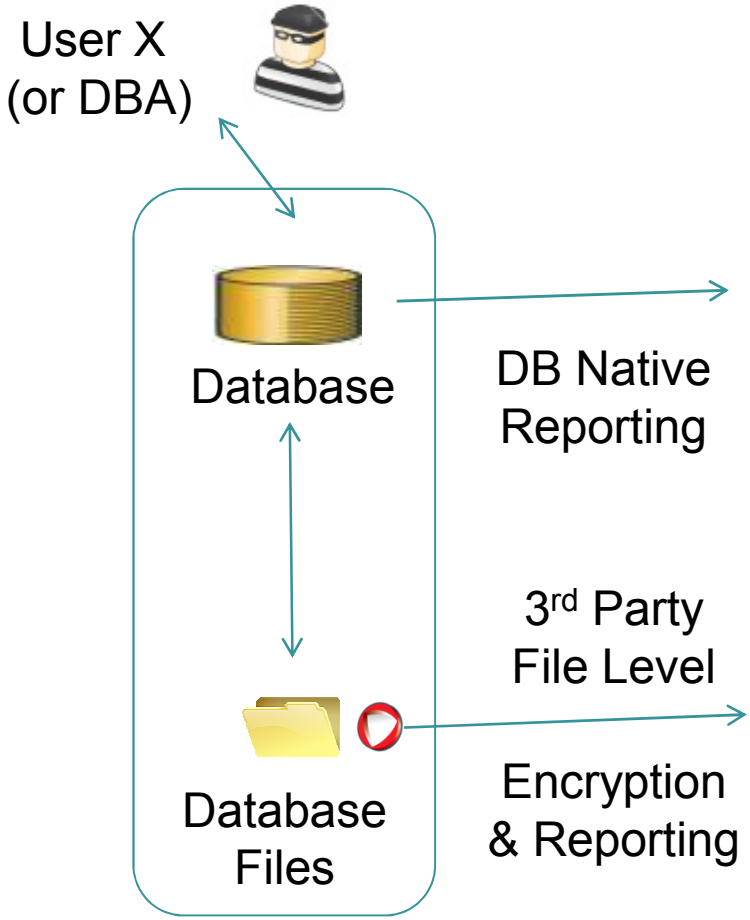


Case Study 3 – Health Information



: Encryption service

HIPAA & HITECH Act – Reporting



DB Native Reporting

3rd Party File Level

Encryption & Reporting

Not Compliant

User	Access	Patient	Health Record
z	Write	c	xxx

No Read Log

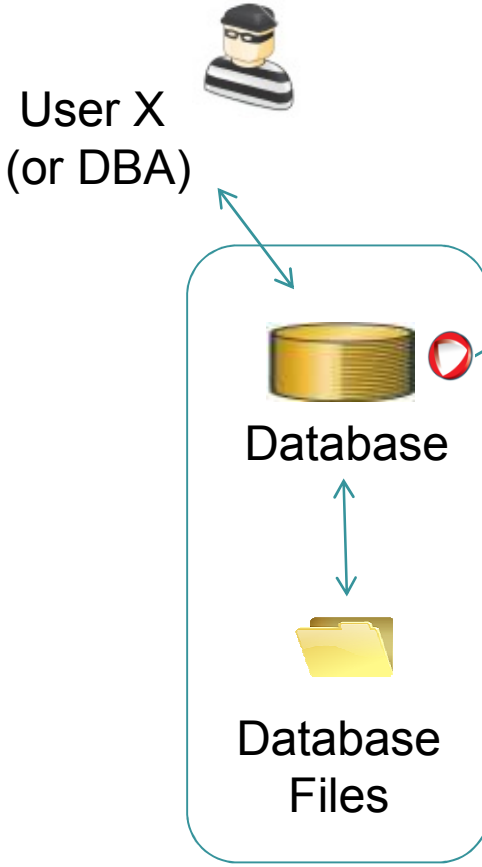
Not Compliant

User	Access	Patient	Health Data Record	Health Data File
Database Process 0001	Read	?	?	PHI002
Database Process 0001	Read	?	?	PHI002
Database Process 0001	Write	?	?	PHI002

No Information On User or Record

: Encryption and reporting service

HIPAA & HITECH Act – Reporting



3rd Party Column Level Encryption & Reporting

Security Administration

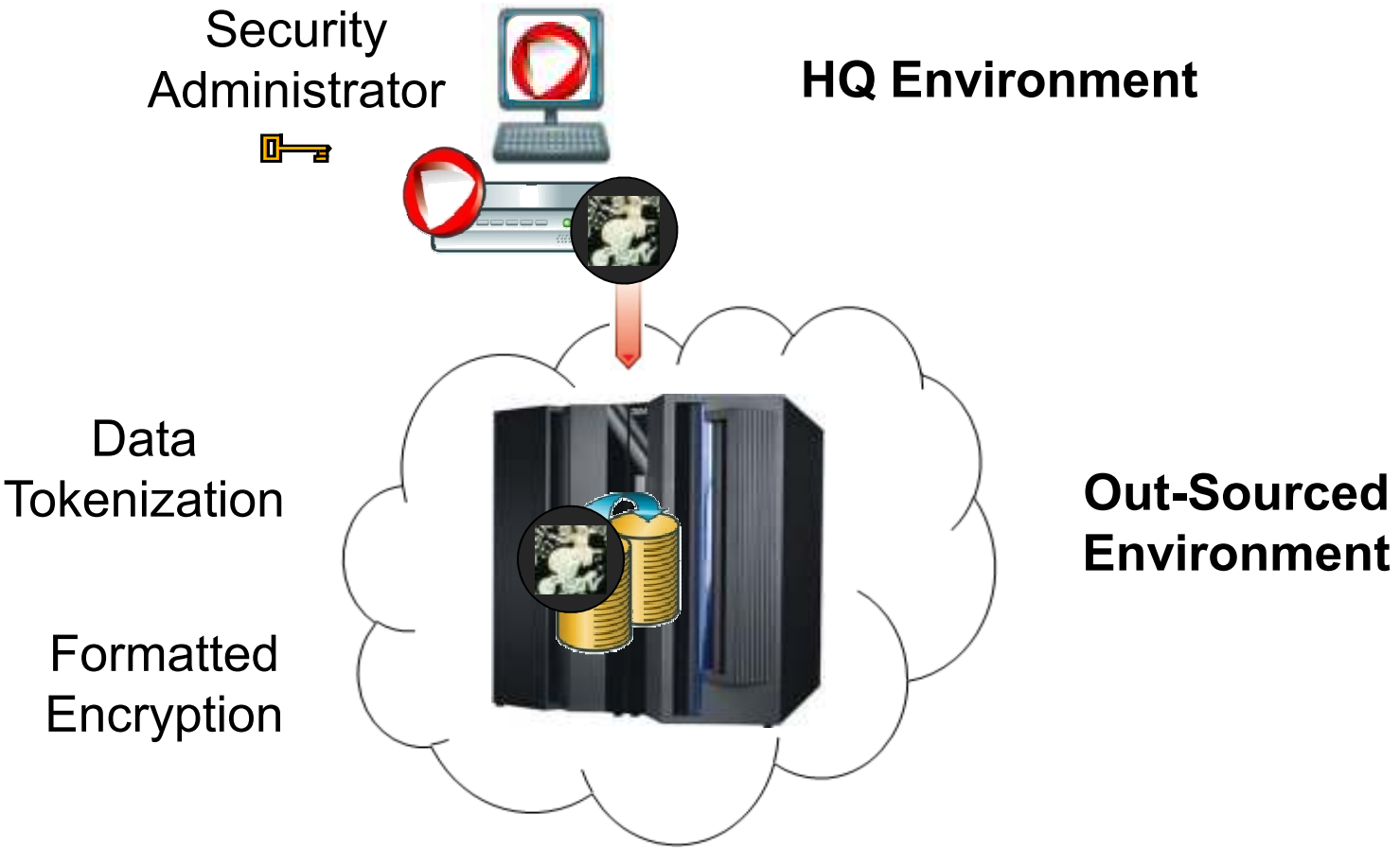
Compliant

User	Access	Patient	Health Record
x	Read	a	1234
DBA	Read	b	2345
z	Write	c	5467

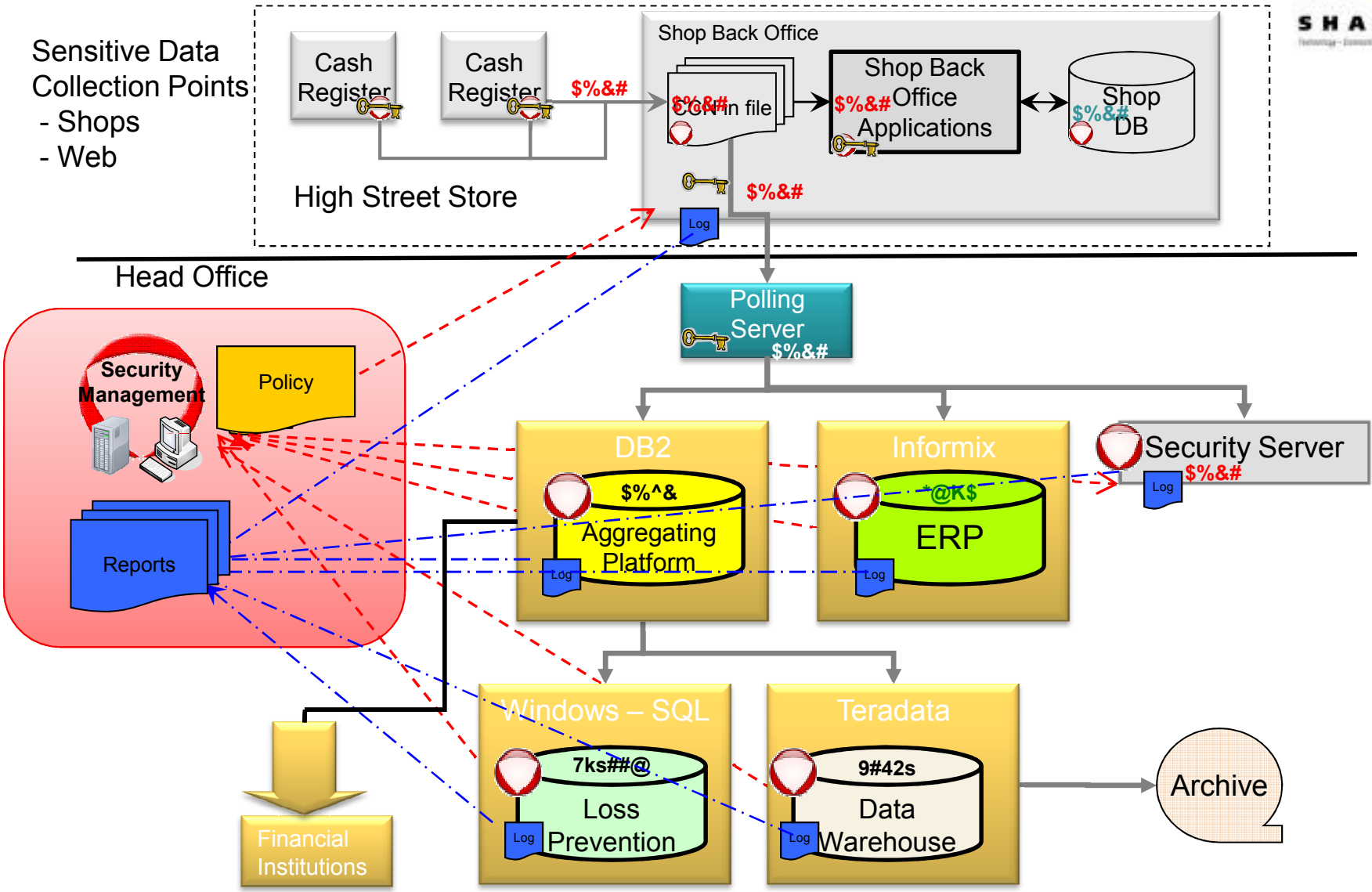
Complete & Encrypted Log

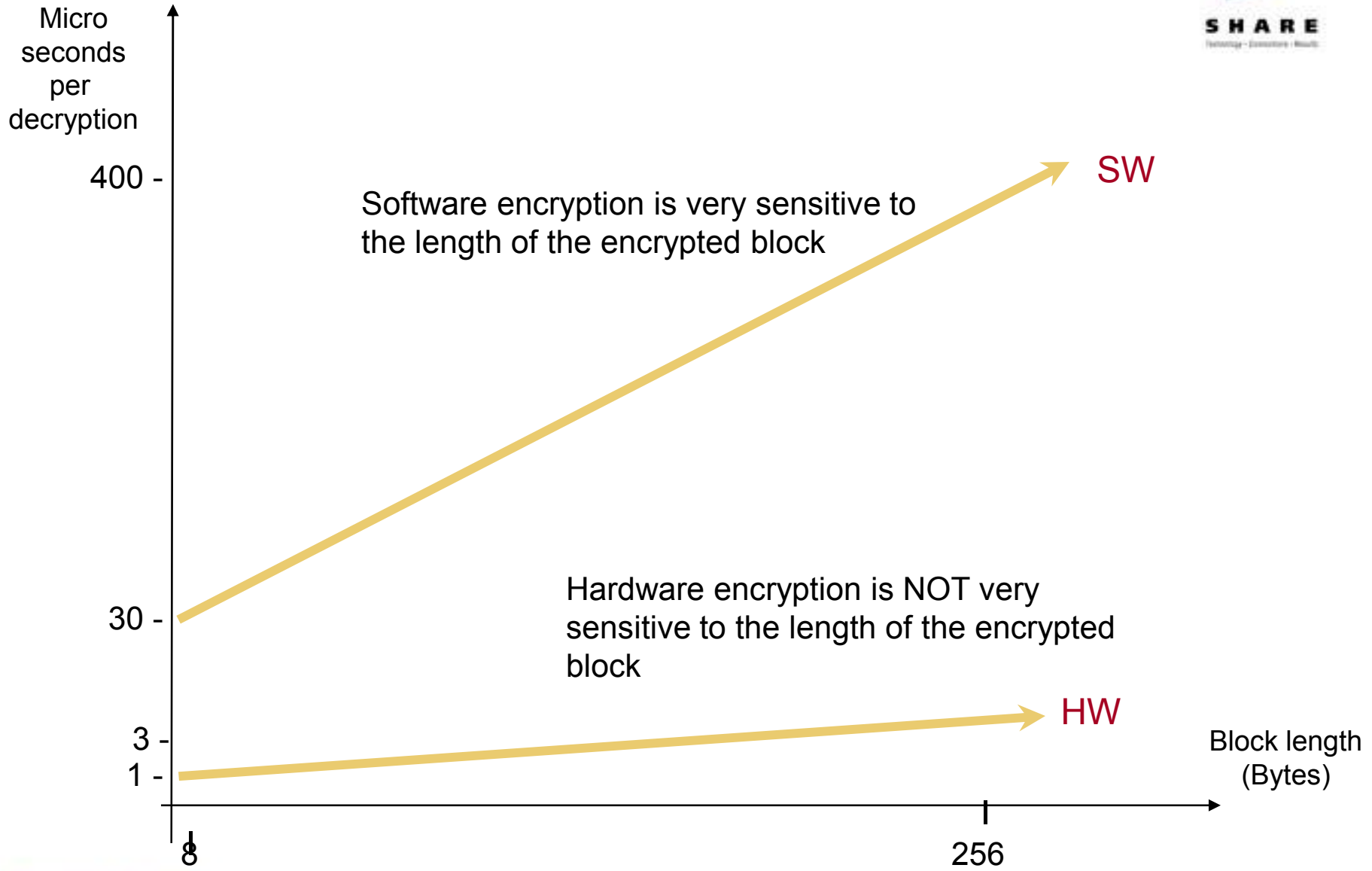
: Encryption and reporting service

Use Case – Data Protection in Out-sourced Environments

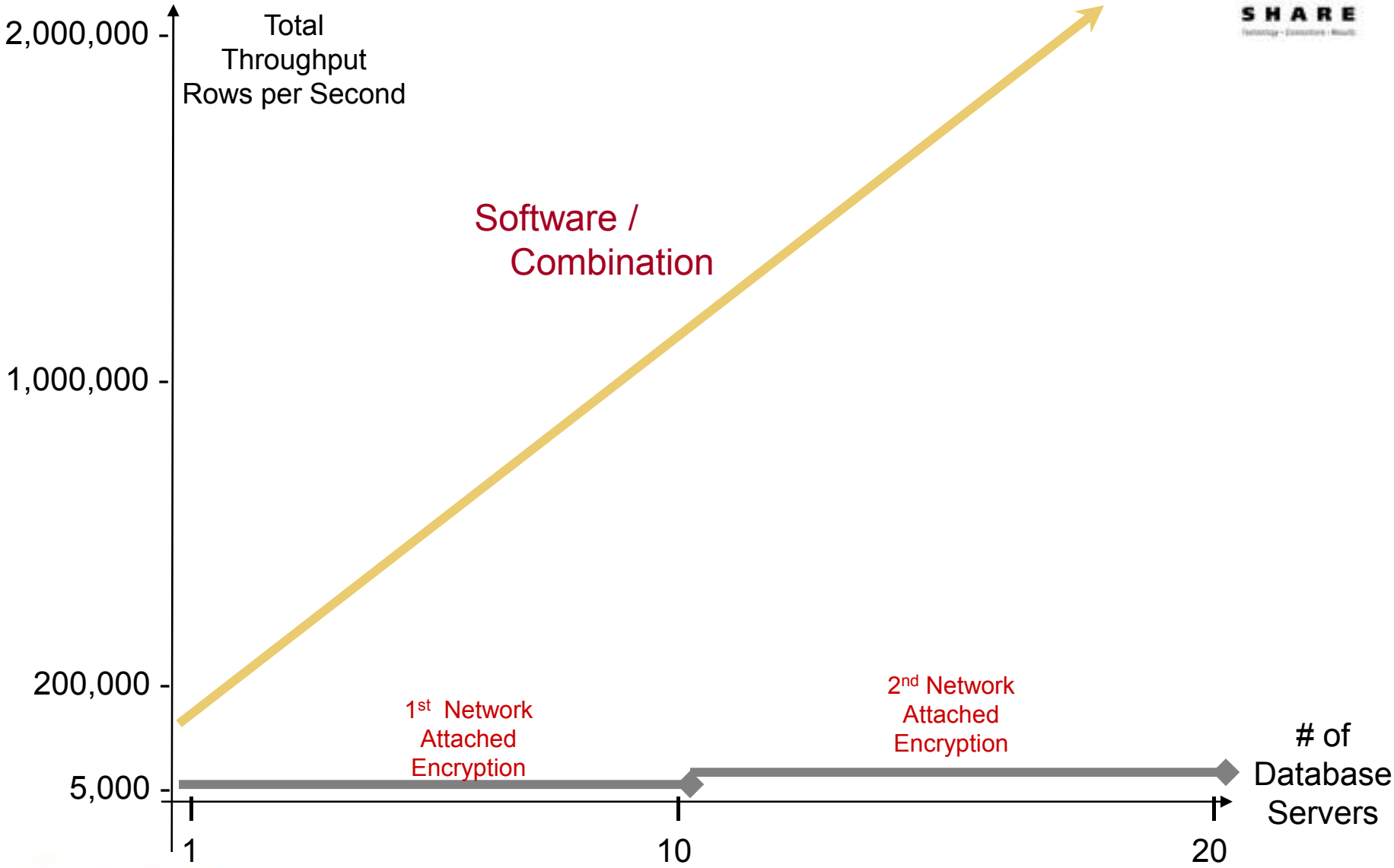


A Retail Data Flow





Test Sample with 3DES, 1 CPU


2,000,000 -
Total
Throughput
Rows per Second
Software /
Combination
1,000,000 -
200,000 -
5,000 -
1st Network
Attached
Encryption
2nd Network
Attached
Encryption
of
Database
Servers
1
10
20

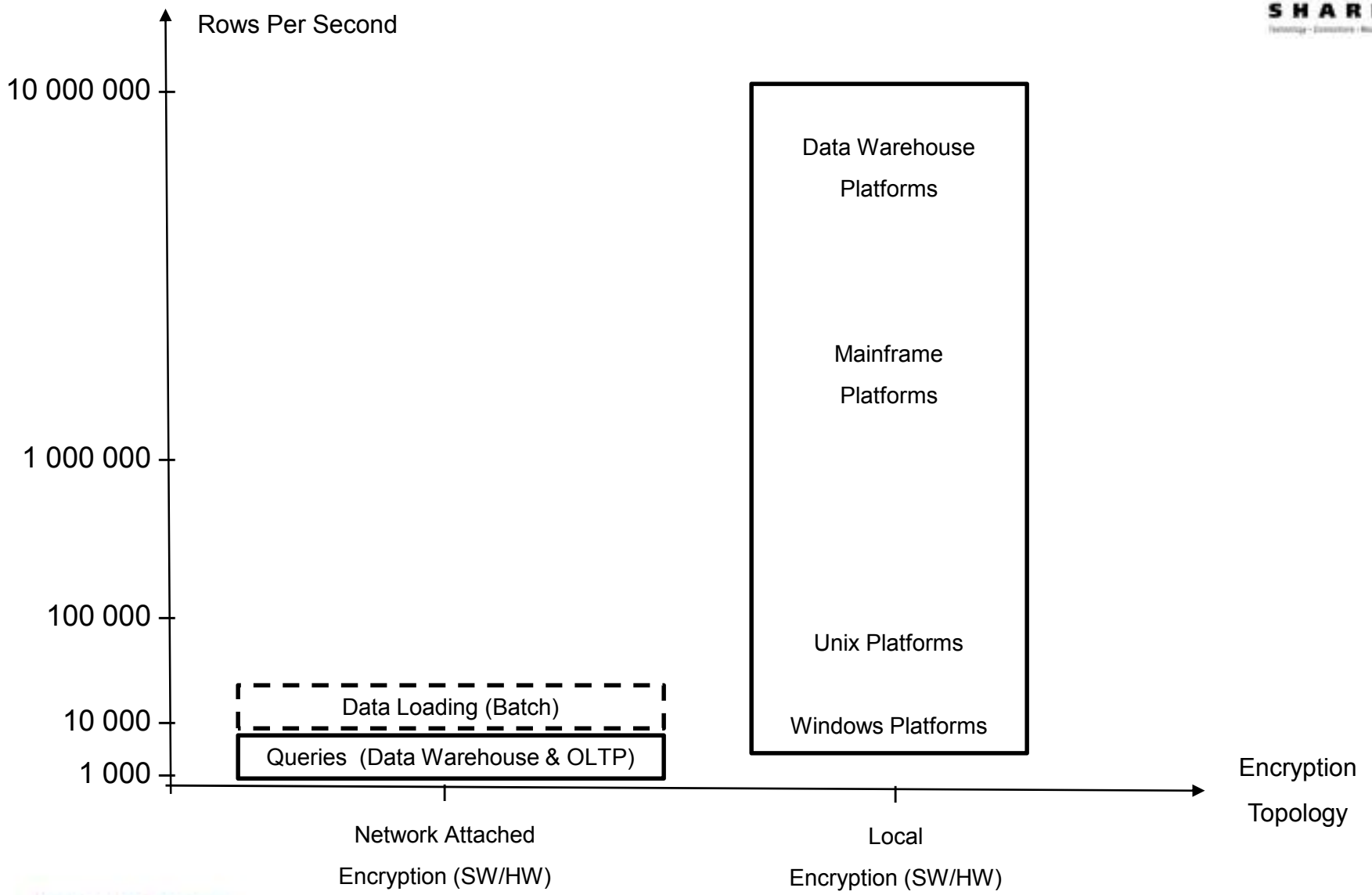
Software vs. Hardware Encryption (NAE)

- Performance and Scalability – software
  - Leverage the processing power of platform, especially advantageous in larger systems i.e. mainframe, Teradata
  - Solutions scale as data volumes grow
  - Solutions can be rolled out to thousands of remote sites using existing technology
- Performance and Scalability - hardware
  - Network Attached Encryption (NAE) devices may be shared across protection points, but network latency can be problematic
  - NAEs have set processing power. In general, to scale you must add more boxes
  - Supporting large, ‘big-iron’ systems can prove challenging

Software vs. Hardware Encryption (NAE)

- Total Cost of Ownership – software
  - Solutions scale as data volumes grow – no need to continually add devices
  - Green factor – as software leverages existing IT, no need for additional cooling, rack space or power
- Total Cost of Ownership - hardware
  - NAE devices need to be continually added or upgraded to keep up with ever growing data volumes
  - Adding devices for each protection point significantly increases deployment and maintenance costs
  - Reserve space at the local landfill

Column Encryption Performance - Different Topologies



Sample data from test cases

Dataset Comparison – Breach Source

	Verizon IR	DataLossDB
Number of breaches	592	2332
Number of compromised/lost records	516,108,232	721,657,540
Time span of dataset	2004-2008	2000-2009[8]

Breach Source	Verizon IR	DataLossDB	DataLossDB-MOD
External	73%	56%	79%
Internal	18%	35%	19%
Partner	38%	4%	0%

Source: 2009 Data Breach Investigations Supplemental Report, Verizon Business RISK team

Dataset Comparison – Industries Represented

Industry	Verizon IR	DataLossDB	DataLossDB-MOD
Retail	54%	8%	9%
Food & Beverage	19%	ND[10]	ND
Financial Services	16%	21%	21%
Technology Services	11%	6%	7%
Manufacturing	5%	6%	6%
Business Services	3%	3%	3%
Education	3%	19%	20%
Healthcare	<1%	13%	13%
Hospitality	2%	1%	<1%
Government	1%	20%	17%
Other/Misc	3%	4%	4%

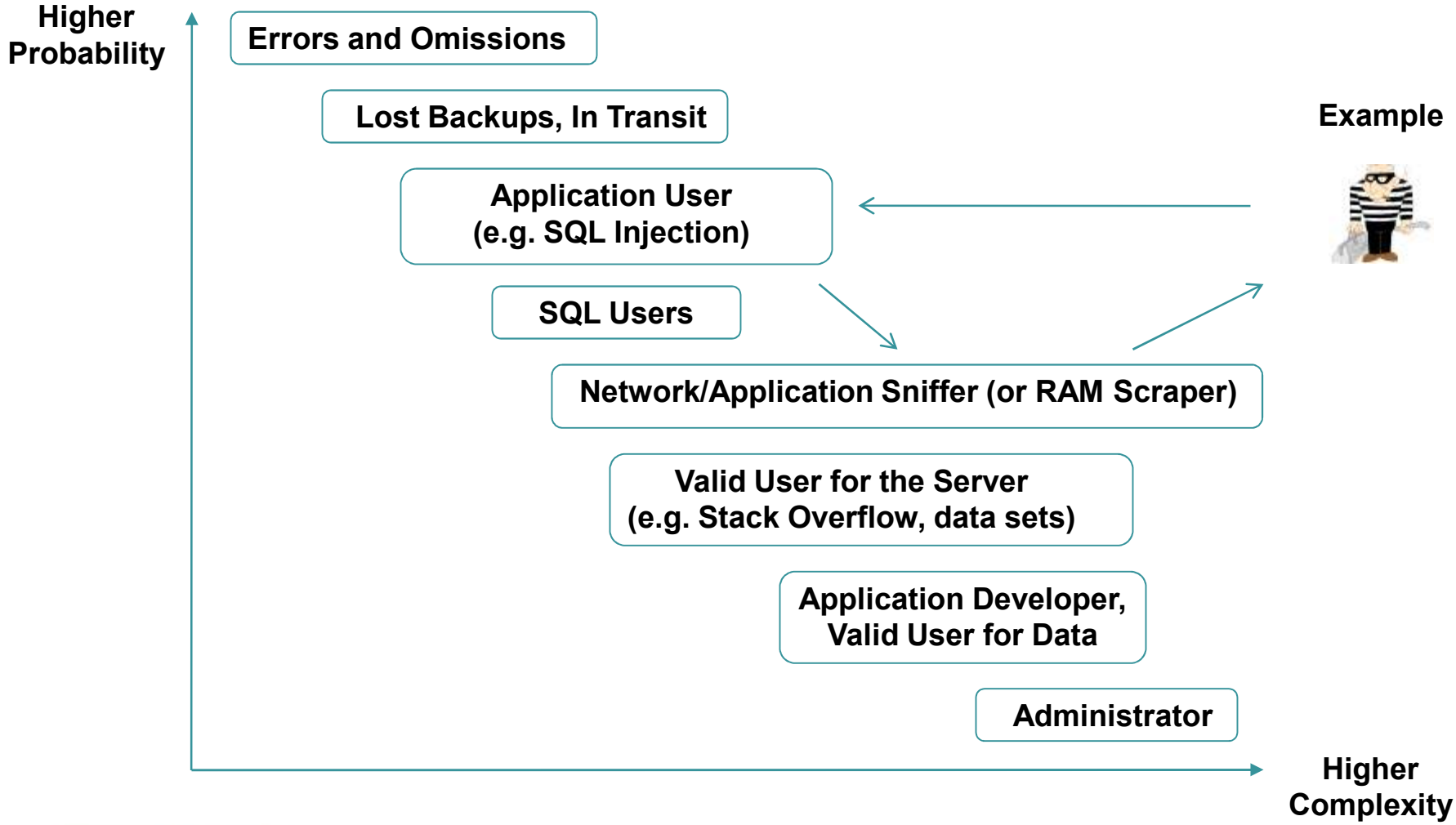
Source: 2009 Data Breach Investigations Supplemental Report, Verizon Business RISK team

Dataset Comparison – Data Type

Data Type	Verizon IR	DataLossDB	DataLossDB-MOD
Payment card data	84%	14%	17%
Personal Information	31%	89%	87%
Authentication credentials	17%	ND	ND
Account number	16%	11%	10%
Intellectual property	9%	ND	ND
Corporate Financial data	5%	11%	9%
Medical information	3%	9%	8%
Monetary Assets / Funds	11%	ND	ND
Other/Misc[11]	26%	11%	11%

Source: 2009 Data Breach Investigations Supplemental Report, Verizon Business RISK team

Step 3: Understand Your Enemy & Probability of Attacks



Source: IBM Silicon Valley Lab(2009)

Application Impact with Different Protection Options

Transparency

Type of Application	Strong Encryption	Formatted Encryption	Token
Can operate on the stored protected value	●	●	●
Need partial information in clear	○	●	●
Need full clear text information	●	○	○

Security

Type of Application	Strong Encryption	Formatted Encryption	Token
Can operate on the stored protected value	●	◒	●
Need partial information in clear	◔	◒	●
Need full clear text information	◔	◔	◔

Application Impact with Different Protection Options

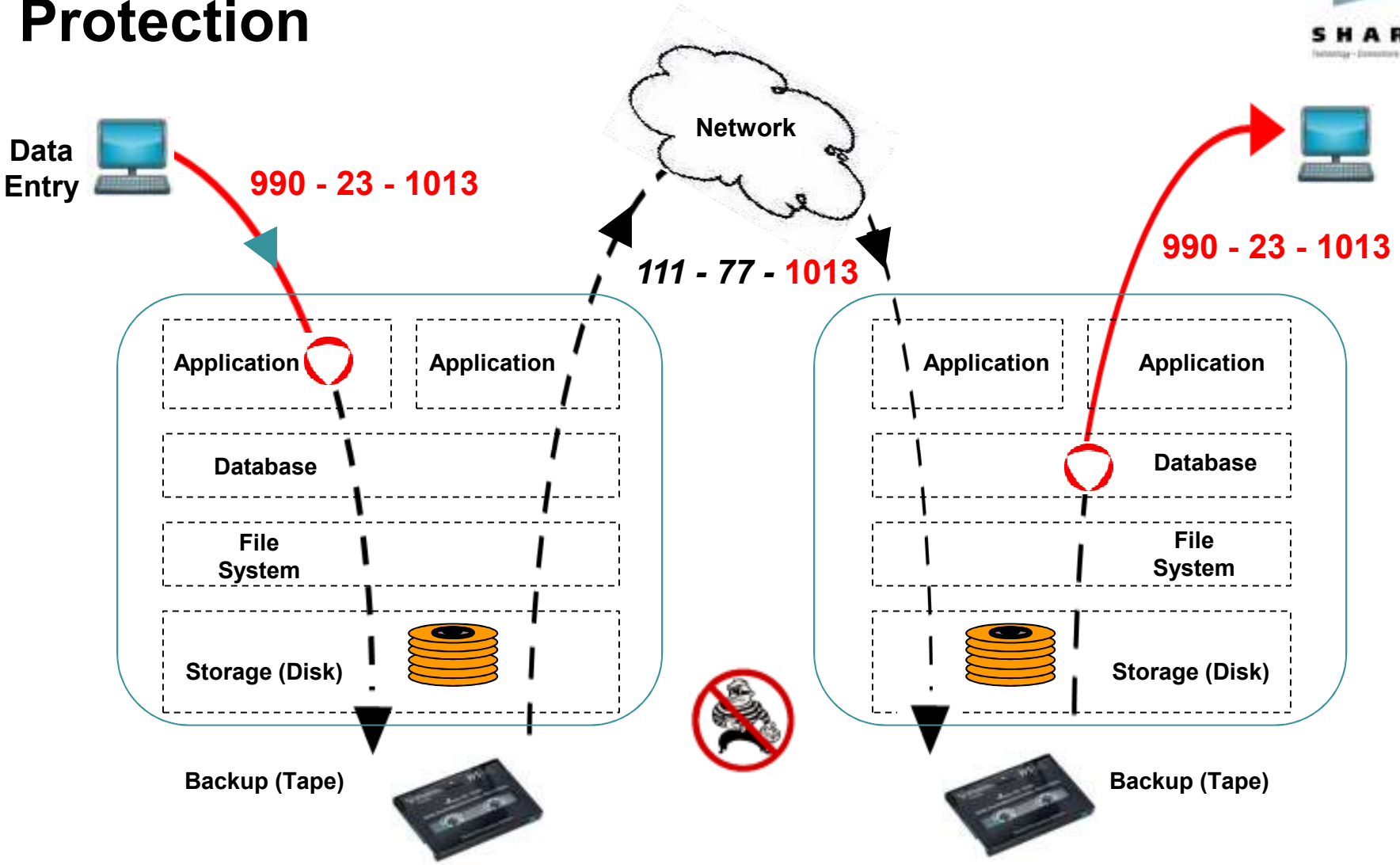
Performance and scalability

Type of Application	Strong Encryption	Formatted Encryption	Token
Can operate on the stored protected value	●	●	●
Need partial information in clear	◔	●	●
Need full clear text information	◔	◒	○

Availability

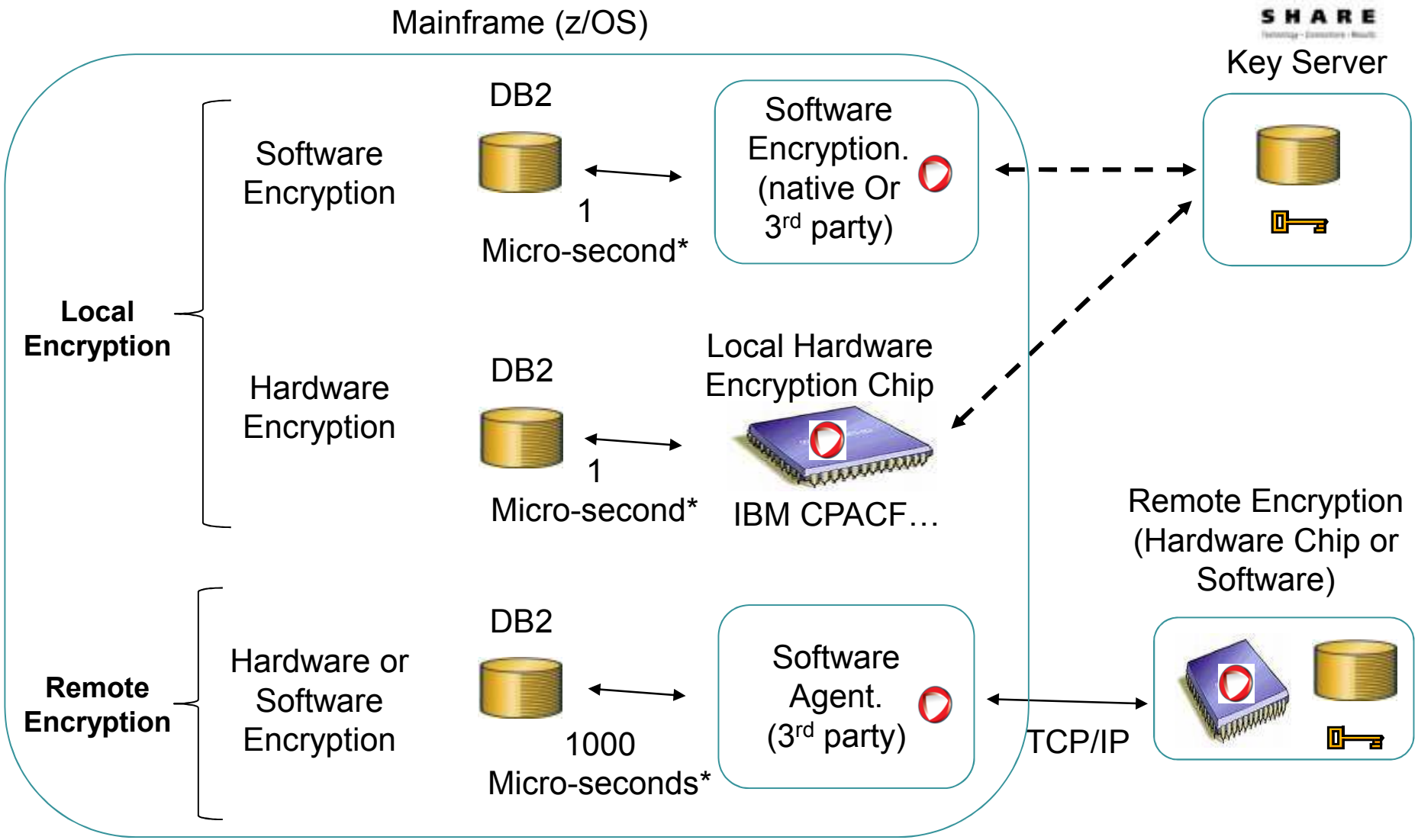
Type of Application	Strong Encryption	Formatted Encryption	Token
Can operate on the stored protected value	●	●	●
Need partial information in clear	◔	●	●
Need full clear text information	◔	◔	○

Choose Your Defenses – Data Flow Protection



Mitigation at the Right System Layer

Encryption Topologies – Mainframe Example



: Encryption service * : 20 bytes