
Data-Centric Security

Key to Cloud and Digital Business

Ulf Mattsson
CTO, Protegrity

Ulf.Mattsson AT protegrity.com

Mary Ann Davidson, Chief Security Officer, Oracle Corporation



ORACLE



Is the Cloud the Answer? What Coke's Breach Teaches Us



Coke recently disclosed that sensitive information belonging to approximately 70,000 current and former North American employees was compromised because the data hadn't been encrypted on company

Industry Involvement

○ PCI Security Standards Council

- Encryption & Tokenization Task Forces
- Cloud & Virtualization SIGs



○ International Federation for Information Processing

- WG 11.3 Data and Application Security



○ Cloud Security Alliance



○ American National Standards Institute



○ User Groups

- Security: ISSA & ISACA
- Databases: IBM & Oracle



Agenda

- Exponential growth of data generation
 - New business models fueled by Big Data, cloud computing and the Internet of Things
 - Creating cybercriminal's paradise
- Challenge in this interconnected world
 - Merging data security with data value and productivity.
- Urgently need a data-centric strategy
 - Protect the sensitive data flowing through digital business systems
- Solutions to bring together data insight & security
 - Safely unlock the power of digital business

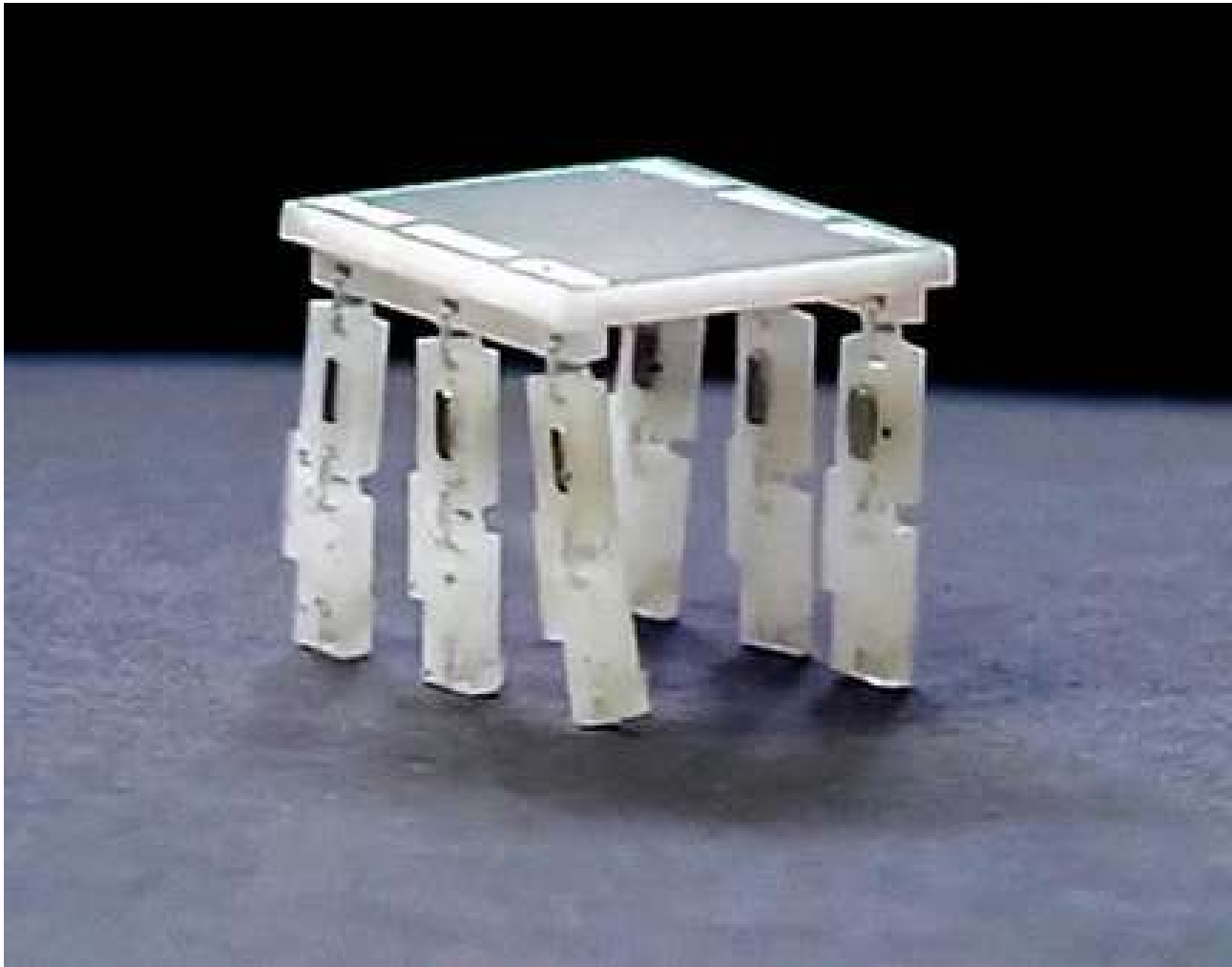


Are you ready for a big change revolution?



Source: www.firstpost.com

Micro-robots, the size of a grain of rice



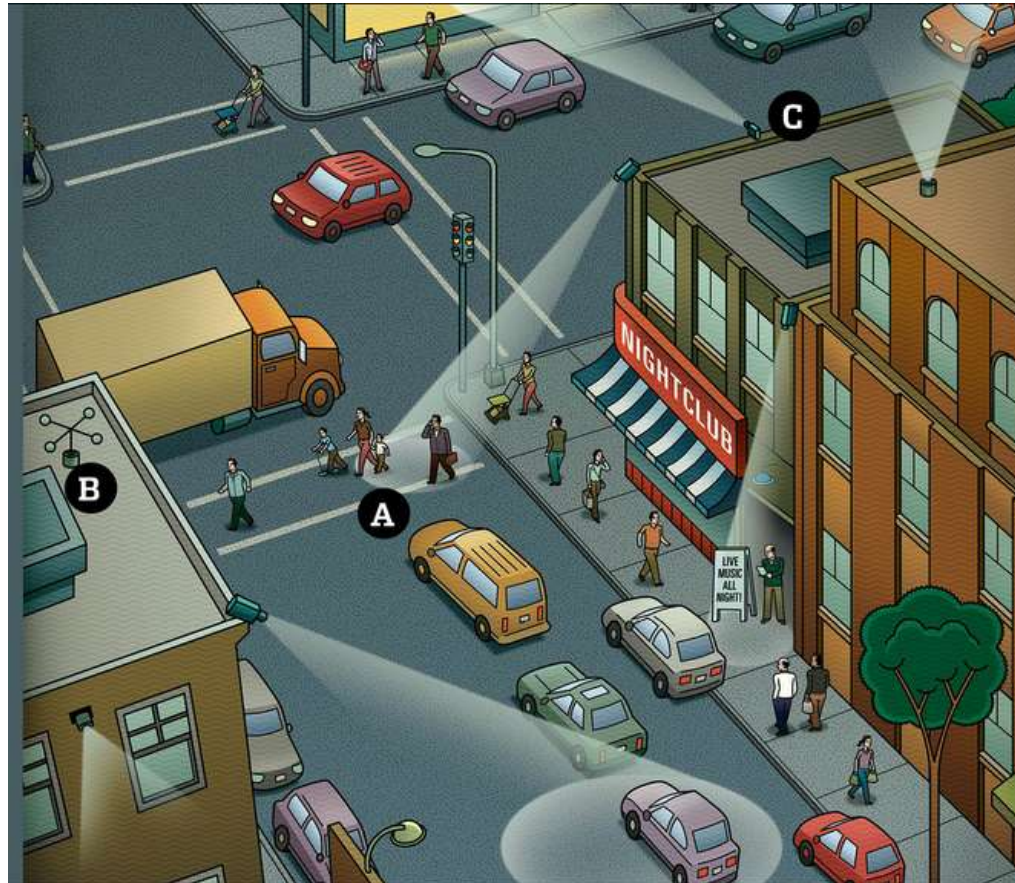
Source: www.ted.com/talks/sarah_bergbreiter

AVATAR - The Internet Of Things?



Source: thesocietypages.org/socimages/2009/12/28/on-avatar-the-movie-spoiler-alert/

They're Tracking When You Turn Off the Lights



Sensors to capture data on environmental conditions including sound volume, wind and carbon-dioxide levels, as well as behavioral data such as pedestrian traffic flow

Samsung engineers are working on wearable for early stroke detection



Source: Early Detection Sensor and Algorithm Package (EDSAP)

FTC Wants a Trusted, Secure Internet of Things



The Federal
Trade
Commission
(FTC)
Looking
At Apple
HealthKit

Source: www.cio-today.com

Security Threats of Connected Medical Devices

○ The Department of Homeland Security

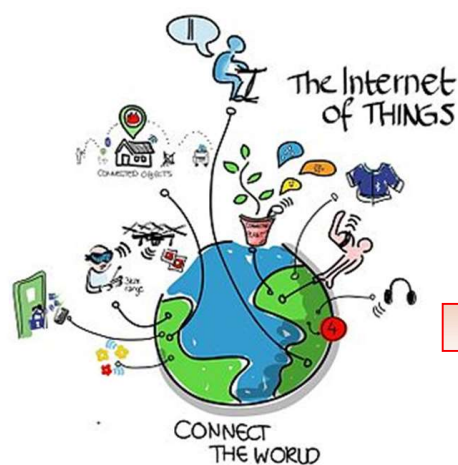
- Investigating 2 dozen cases of suspected cyber security flaws in medical devices that could be exploited
- Can be detrimental to the patient, creating problems such as instructing an infusion pump to overdose a patient with drugs or forcing a heart implant to deliver a deadly jolt of electricity
- Encrypt medical data that's stored

○ PricewaterhouseCoopers study

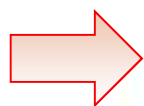
- \$30billion annual cost hit to the U.S. healthcare system due to inadequate medical-device interoperability

IoT is a Paradise for Hackers

- Almost 90 percent of the devices collect personal information such as name, address, date of birth, email, credit card number, etc.
- Un-encrypted format on to the cloud and big data, thus endangering the privacy of users



Source: wikipedia.org



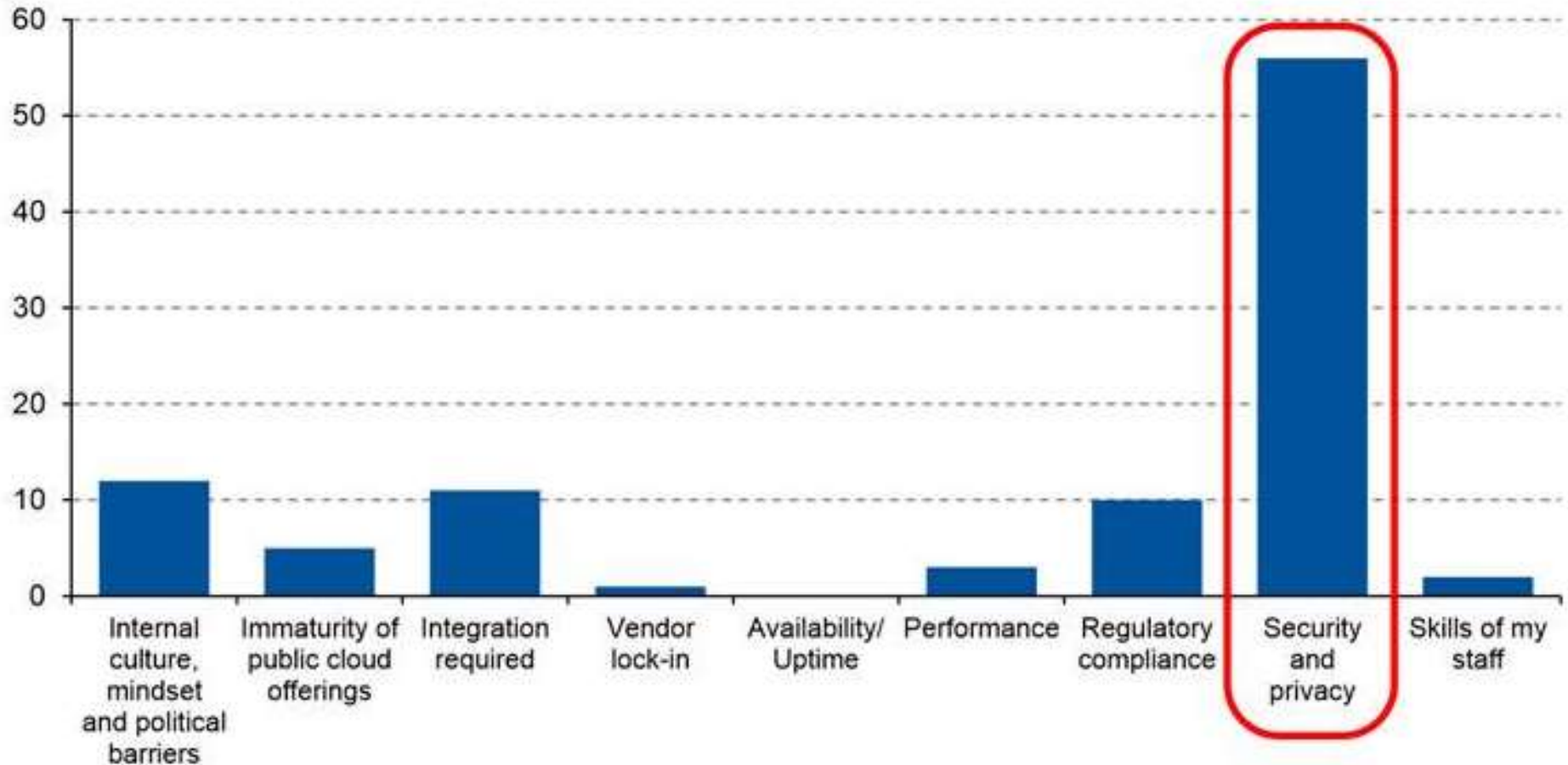
Source: HP Security Research

Cloud Security

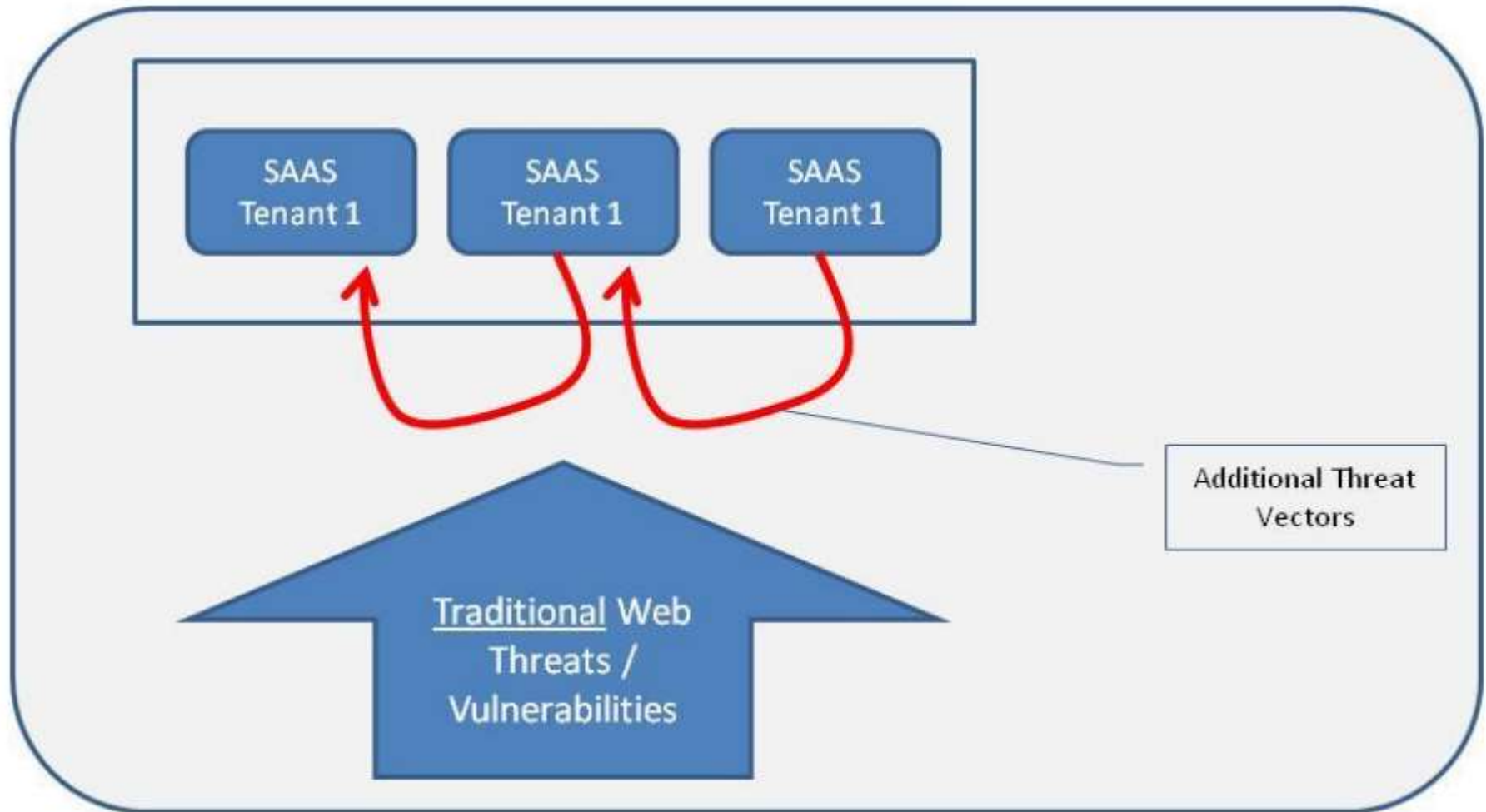
82%

Of organizations currently (or plan to) transfer sensitive/confidential data to the cloud in next 24 mo.

What Is Your No. 1 Issue Slowing Adoption of Public Cloud Computing?



Threat Vector Inheritance



Amazon AWS Breach

- Researchers again finding similar issues that six years ago were demonstrated by Ristenpart about concrete evidence for sensitive information leakage on a commercial cloud. A 2015 research paper presents a full-fledged attack that exploits leakages of decryption keys and concluded that the cross-VM leakage is present in public clouds and can become a practical attack vector for both co-location detection and data theft.
- Gartner is recommending to "understand when data appears in clear text, where keys are made available and stored, and who has access to the keys," and recommending to "apply encryption or tokenization."
- Gartner released the report "Simplify Operations and Compliance in the Cloud by Protecting Sensitive Data" in June 2015 that highlighted key challenges as "cloud increases the risks of noncompliance through unapproved access and data breach."
- The report recommended CIOs and CISOs to address data residency and compliance issues by "applying encryption or tokenization," and to also "understand when data appears in clear text, where keys are made available and stored, and who has access to the keys."

New Challenges when Moving Data to Cloud

- Shared infrastructures
- Multi-tenant environments
- Liability of data loss or security breach
- Data reside issues
- Data purge/deletion issues
- Shadow IT
- Key management model

49% recommended Database security

40% of budget still on Network security
only

19% to Database security

Conclusion: Organizations have traditionally spent money on network security and so it is earmarked in the budget and requires no further justification

CHALLENGE

**How can we
Secure Data
in the new
Perimeter-less
Environments?**

SOLUTION

Fine Grained Data Security

Gartner®

Data–Centric Audit and Protection (DCAP)

Organizations that have not developed data-centric security policies to coordinate management processes and security controls across data silos need to act

By 2018, data-centric audit and protection strategies will replace disparate siloed data security governance approaches in 25% of large enterprises, up from less than 5% today



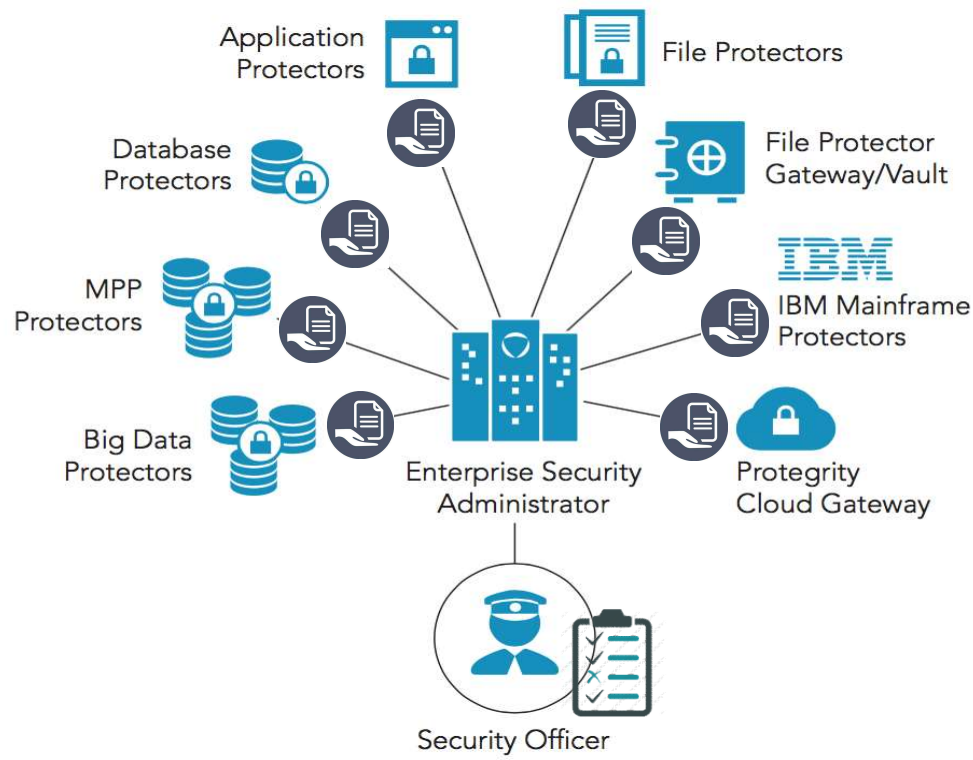
Data–Centric Audit and Protection (DCAP)

- Centrally managed security policy
- Across unstructured and structured silos
- Classify data, control access and monitoring
- Protection – encryption, tokenization and masking
- Segregation of duties – application users and privileged users
- Auditing and reporting

Gartner[®]

Source: Gartner – Market Guide for Data – Centric Audit and Protection (DCAP), Nov 21 2014

Central Management, Policy Deployment – Hadoop + Beyond



Enterprise Data Security Policy

What

What is the sensitive data that needs to be protected.

How

How you want to protect and present sensitive data. There are several methods for protecting sensitive data.

Who

Who should have access to sensitive data and who should not. Security access control.

When

When should sensitive data access be granted to those who have access. Day of week, time of day.

Where

Where is the sensitive data stored? This will be where the policy is enforced.

Audit

Audit authorized or un-authorized access to sensitive data.

Securing Cloud Data

Enterprises Are Leaving Cloud Security Policies To Chance

- Only a third have a strategy for securing a mix of different data center and cloud deployment scenarios
- 75 percent of organizations utilize identity and access management tools on premises, only 31 percent use it in the cloud
- 63 percent of organizations use a SIEM to track security events across traditional data center assets, just 25 percent do the same with cloud assets
- A huge issue for many organizations today, given the fact that many public cloud providers don't currently offer or support many security tools considered standard by most security teams

Data-Centric Protection Increases Security in Cloud Computing


- Rather than making the protection platform based, the security is applied directly to the data
- Protecting the data wherever it goes, in any environment
- Cloud environments by nature have more access points and cannot be disconnected
- Data-centric protection reduces the reliance on controlling the high number of access points

- Gartner released the report “Simplify Operations and Compliance in the Cloud by Protecting Sensitive Data” in June 2015
- Cloud increases the risks of noncompliance through unapproved access and data breach
- CIOs and CISOs to address data residency and compliance issues by applying encryption or tokenization
- Understand when data appears in clear text, where keys are made available and stored, and who has access to the keys
- Cloud Data Protection Gateways” provides a “High Benefit Rating” and “offer a way to secure sensitive enterprise data and files”



INFOWORLD TECH WATCH

By [Serdar Yegulalp](#) | [Follow](#)

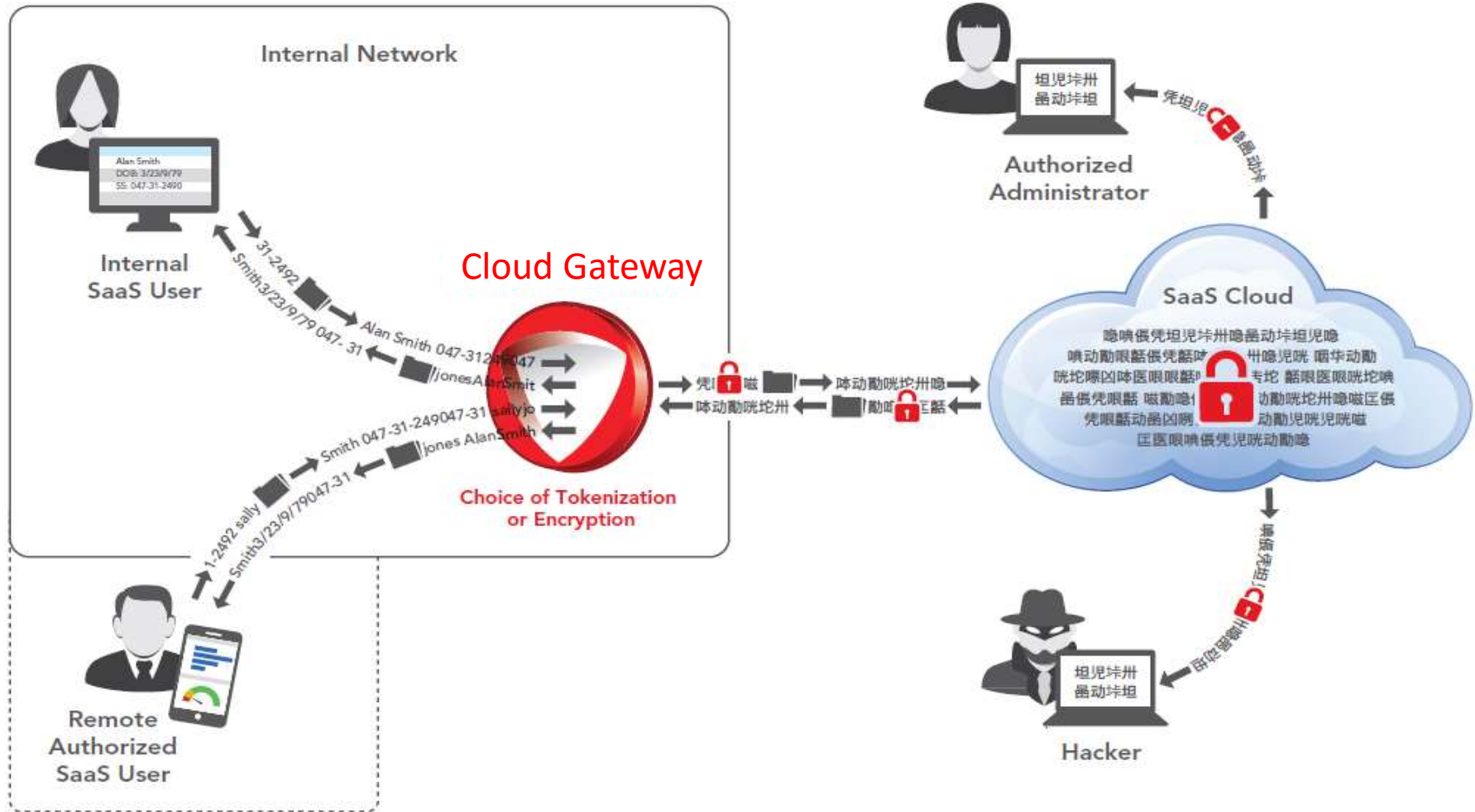
[About](#) 

Informed news analysis every weekday

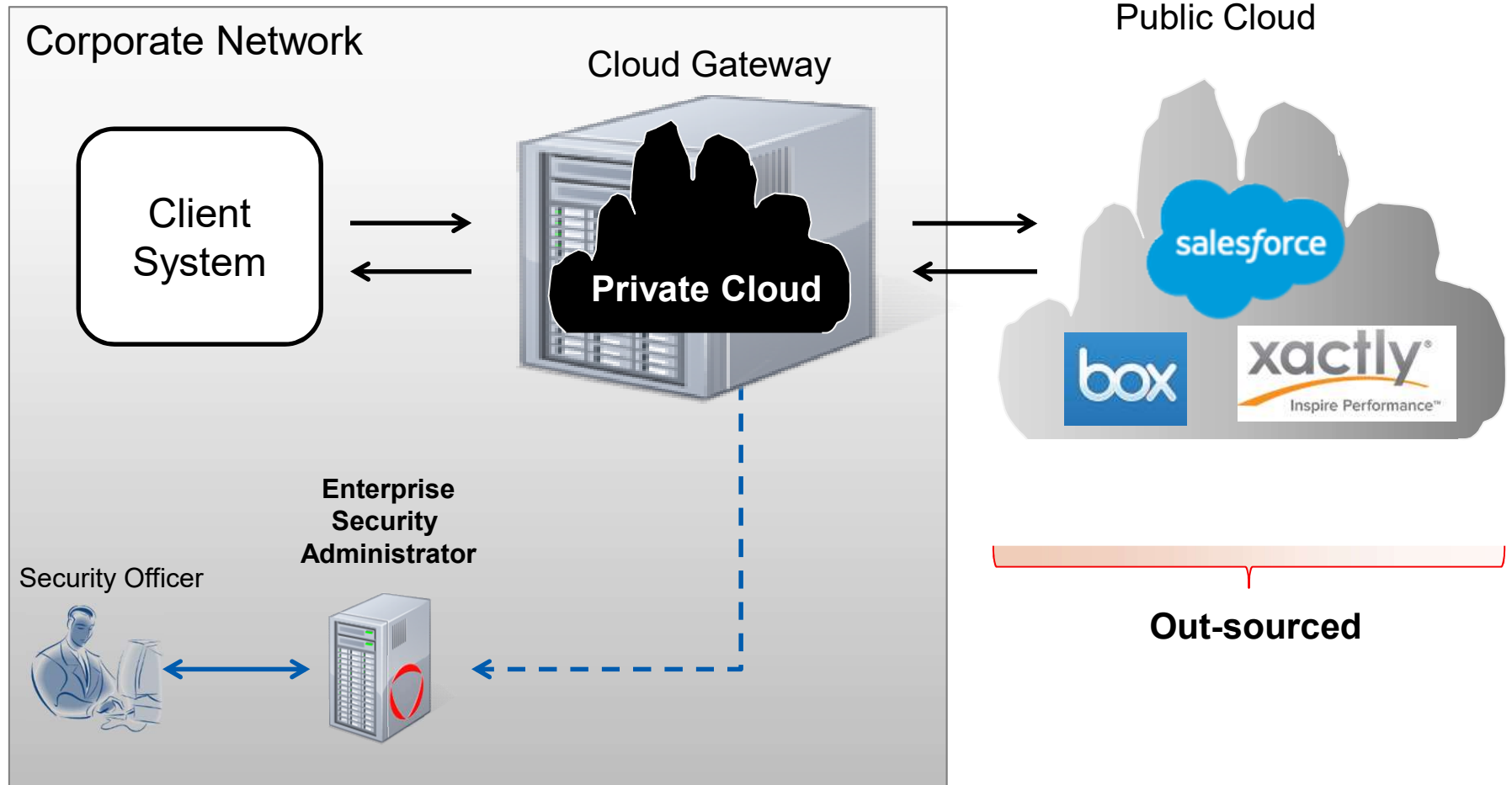
Google Compute Engine now supports bring-your-own-key encryption

Other big cloud providers already offer BYOK, but Google claims its option, available at no additional cost, is more comprehensive

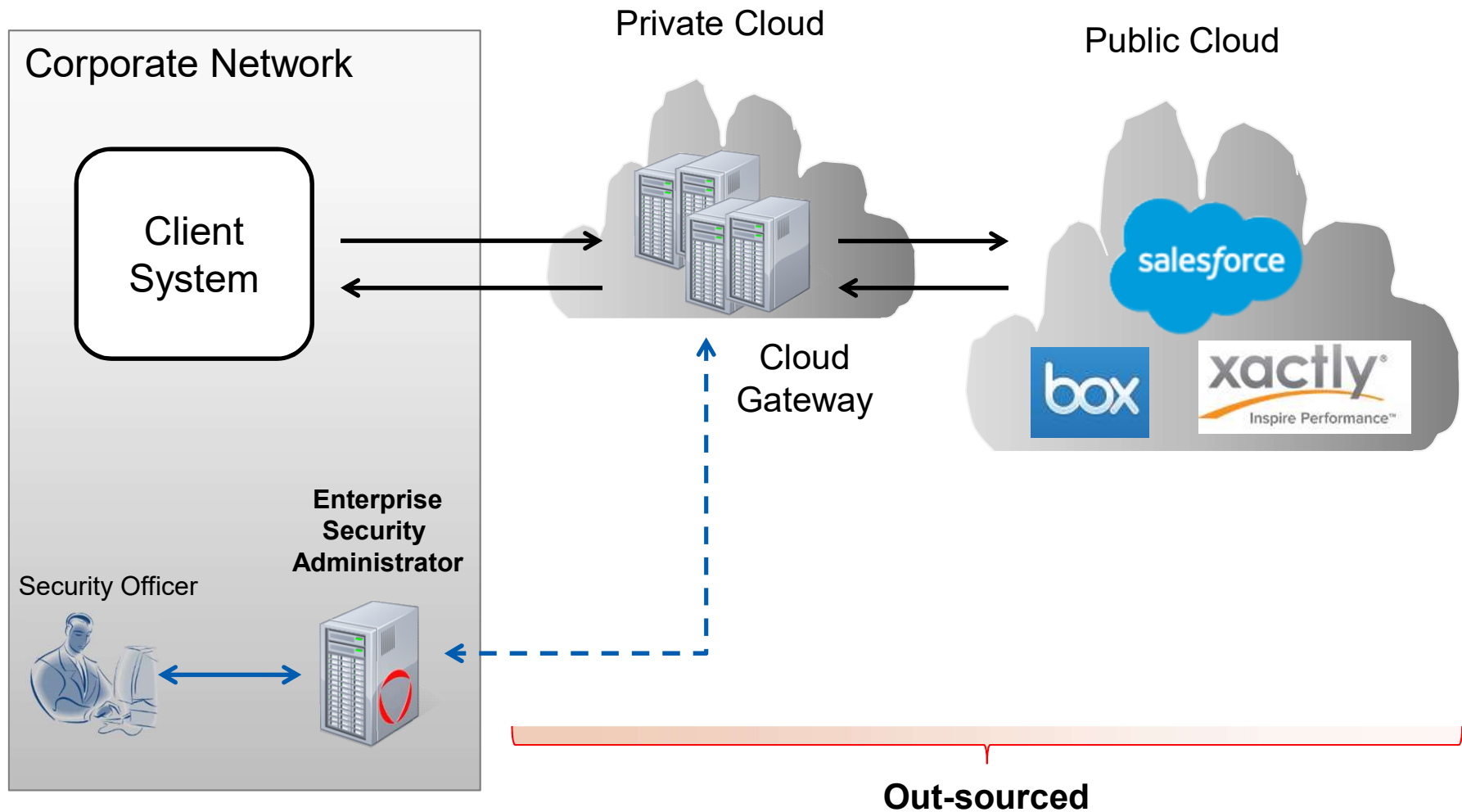
Protect the Entire Flow of Sensitive Data



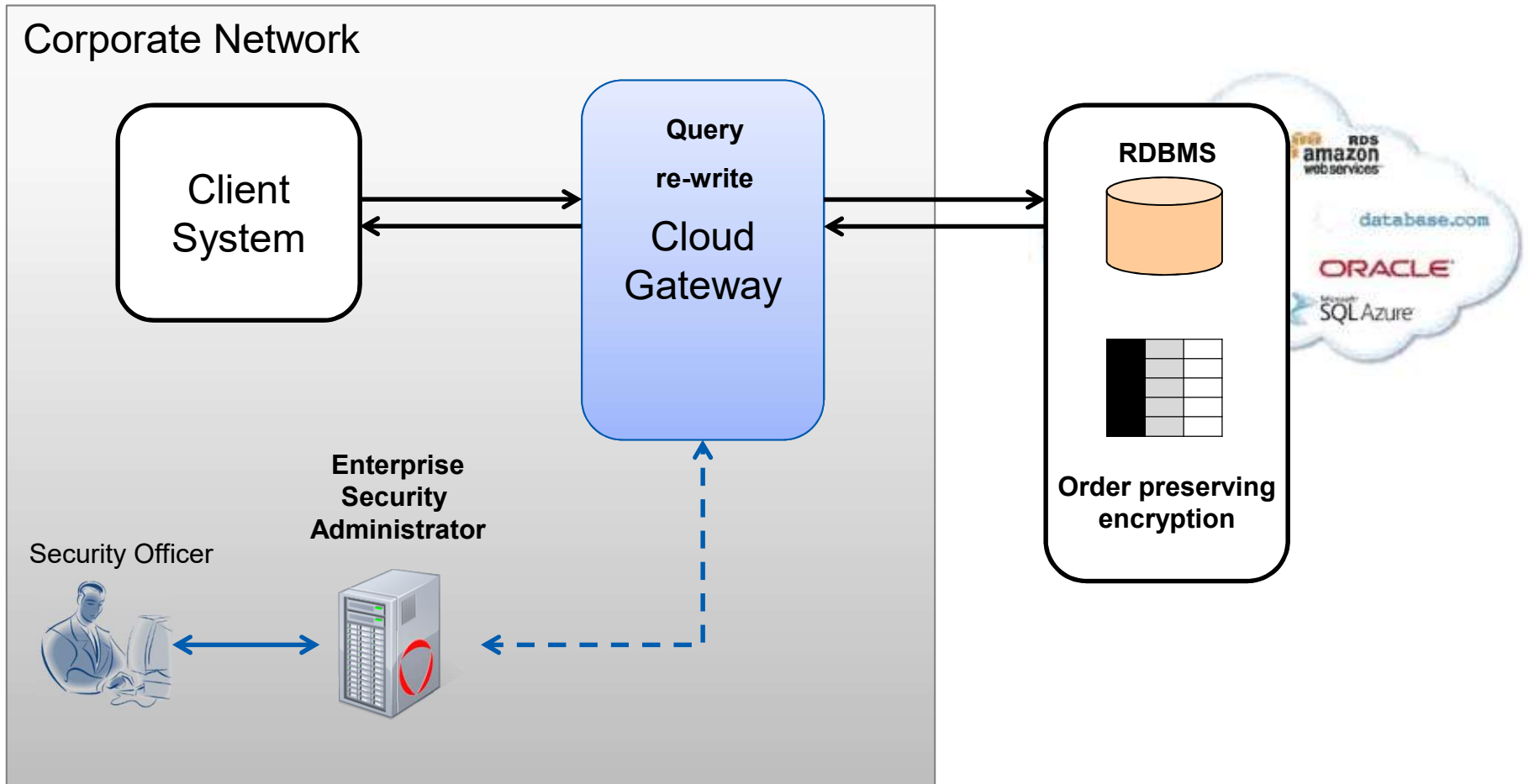
Security Gateway Deployment – Hybrid Cloud



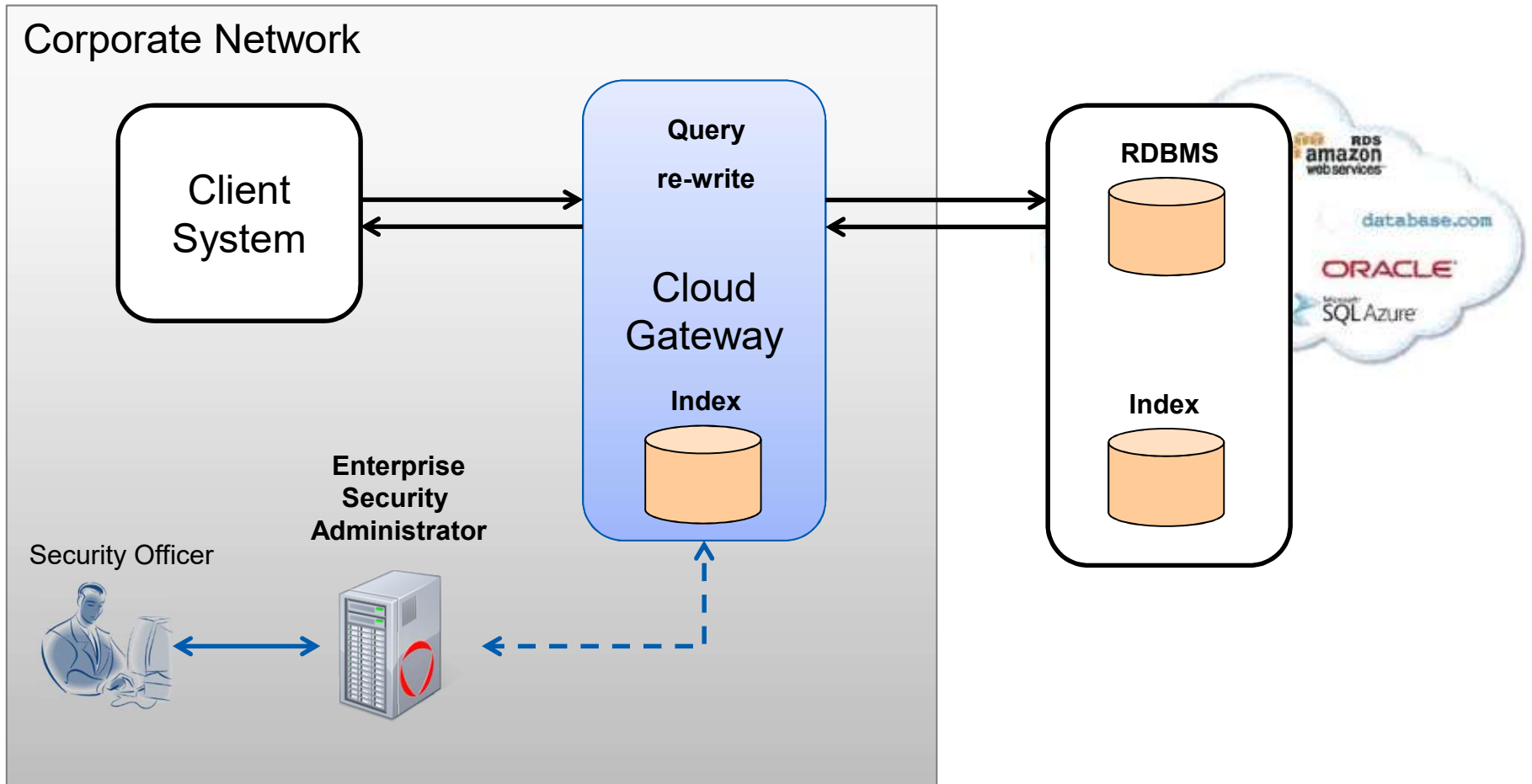
Security Gateway Deployment – Hybrid Cloud



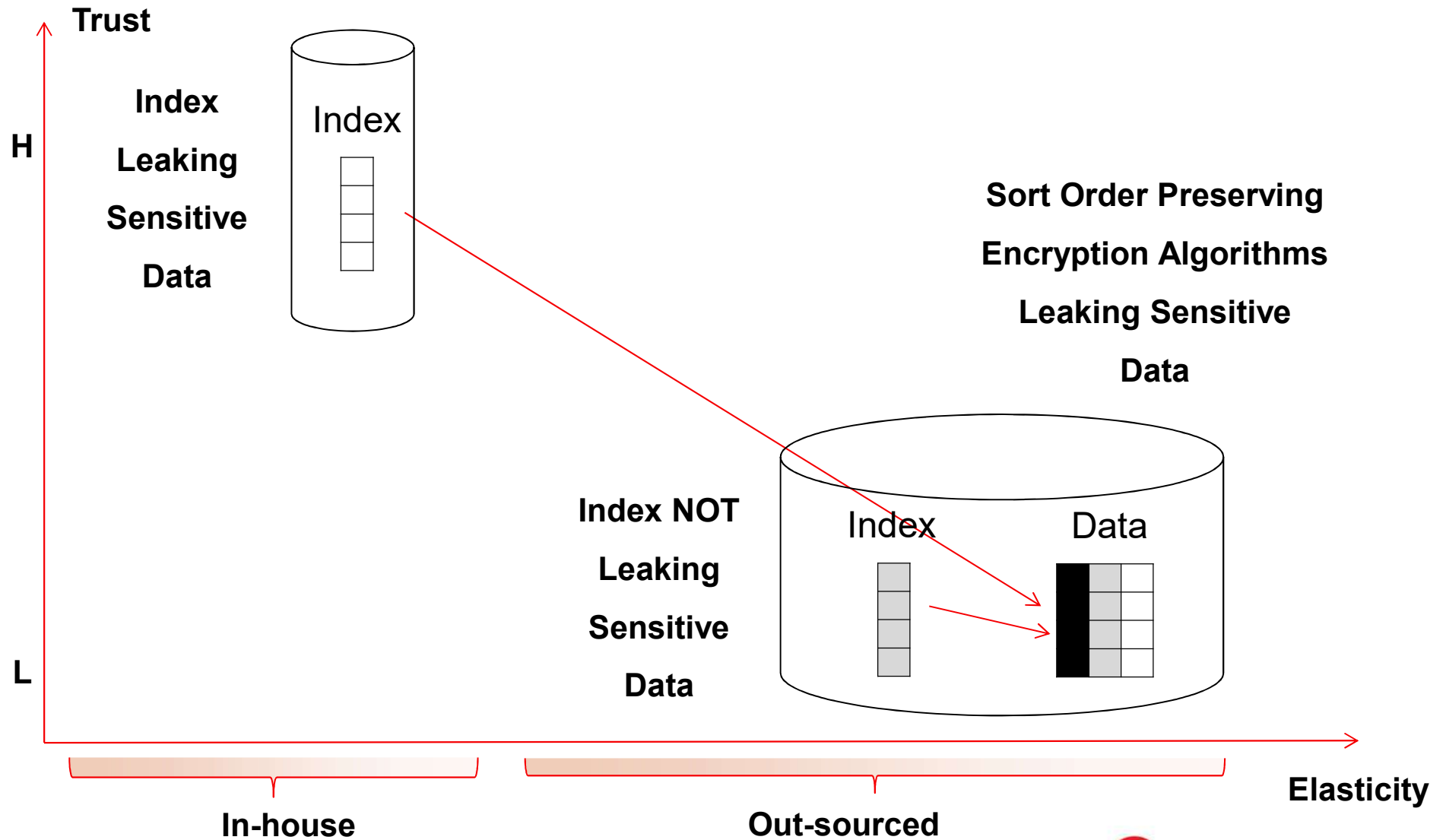
Security Gateway – Searchable Encryption



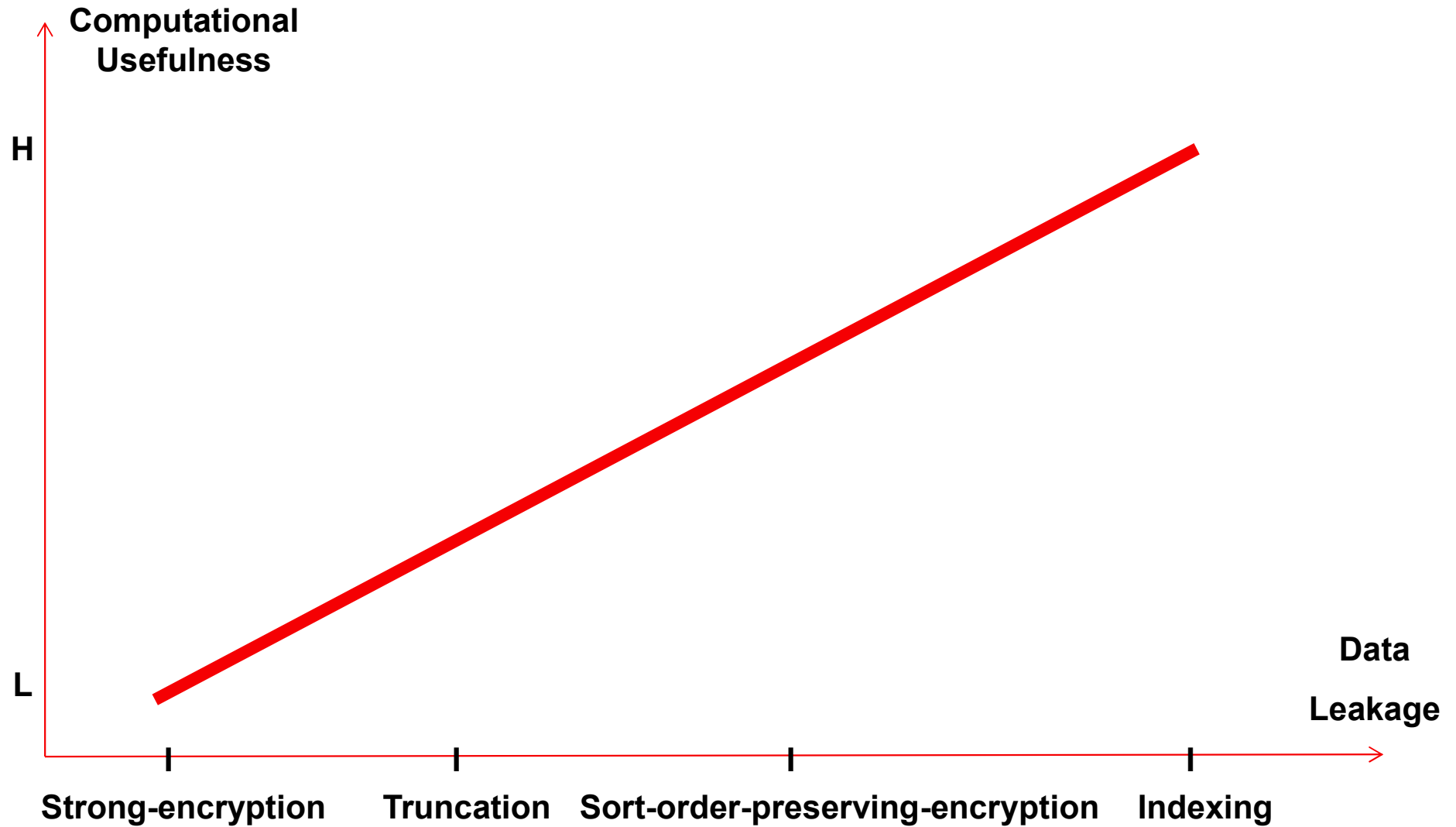
Security Gateway – Search & Indexing



Risk Adjusted Data Leakage

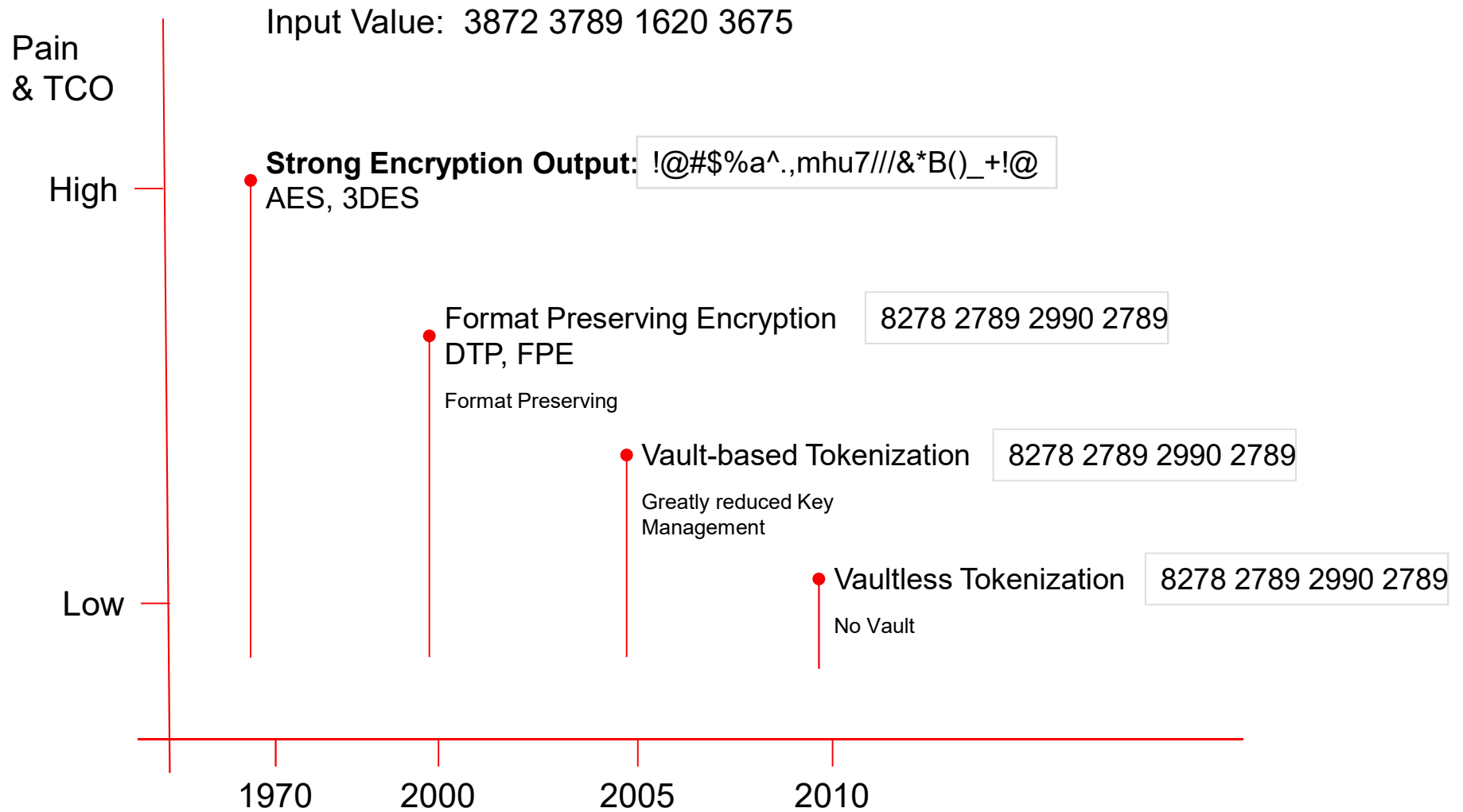


Risk Adjusted Storage – Data Leaking Formats



Comparing Fine Grained Data Protection Methods

Reduction of Pain with New Protection Techniques

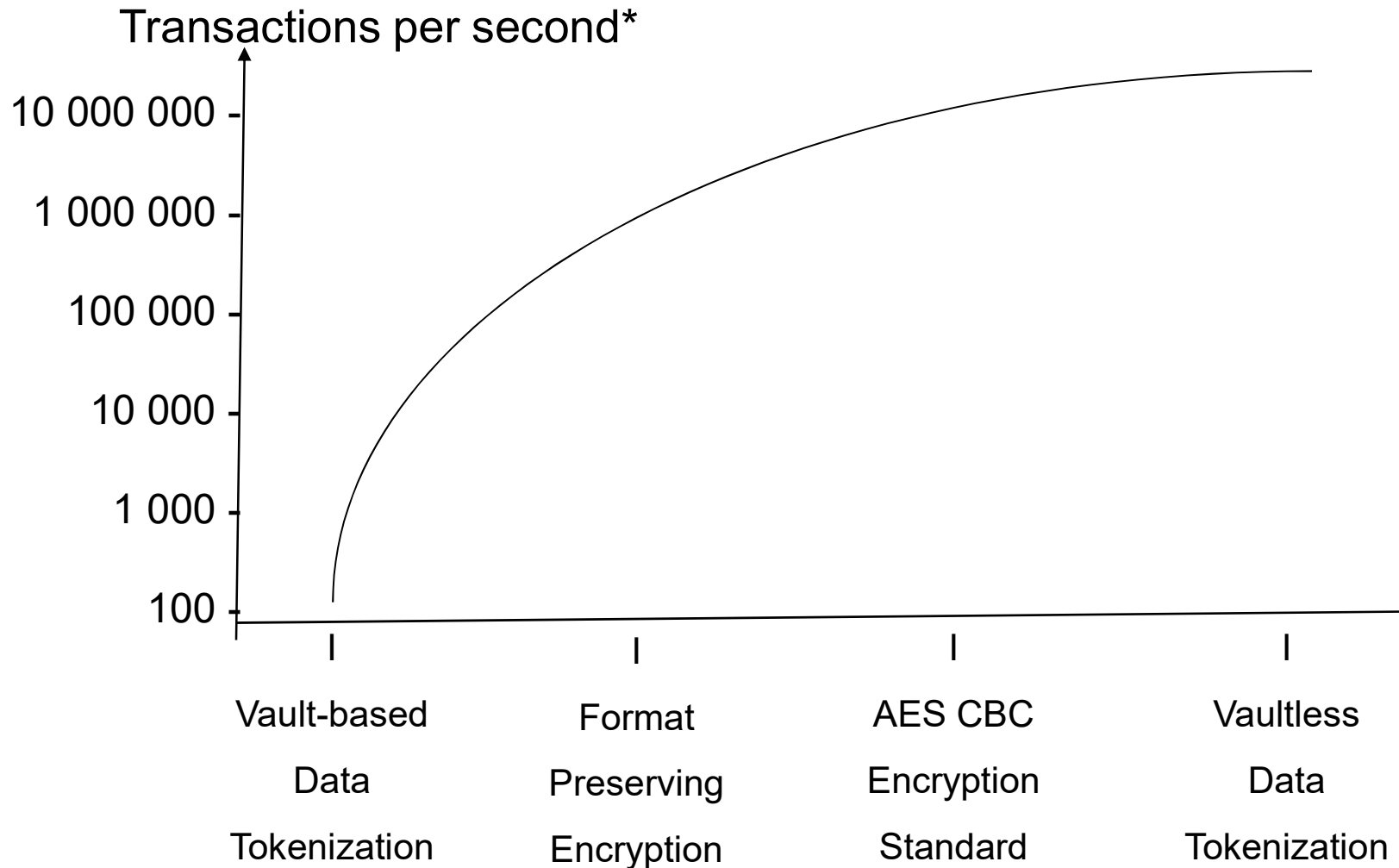


Cloud Gateway - Requirements Adjusted Protection

Data Protection Methods	Scalability	Storage	Security	Transparency
System without data protection				
Weak Encryption (1:1 mapping)				
Searchable Gateway Index (IV)				
Vaultless Tokenization				
Partial Encryption				
Data Type Preservation Encryption				
Strong Encryption (AES CBC, IV)				

Best Worst

Speed of Fine Grained Protection Methods



*: Speed will depend on the configuration

What is Data Tokenization?

Fine Grained Data Security Methods

Tokenization and Encryption are Different

	Encryption	Tokenization
Used Approach	Cipher System	Code System
Cryptographic algorithms	●	
Cryptographic keys	●	
Code books		●
Index tokens		●




Source: McGraw-HILL ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY



Significantly Different Tokenization Approaches

Property	Vault-based		Vaultless
	Dynamic	Pre-generated	
Footprint	Large, Expanding	Large, Static	Small, Static
Replication	Complex replication required	No replication required	No replication required
Collisions	Prone to collisions	No collisions	No collisions
Latency / Performance	Will impact performance and scalability	Will impact performance and scalability Faster than the traditional dynamic approach	Little or no latency Fastest tokenization in the industry
Tokenizing many data categories	Potentially impossible	Potentially impossible	Can tokenize many data categories with minimal or no impact on footprint or performance

Examples of Protected Data

Field	Real Data	Tokenized / Pseudonymized
Name	Joe Smith	csu wusoj
Address	100 Main Street, Pleasantville, CA	476 srta coetse, cysieondusbak, CA
Date of Birth	12/25/1966	01/02/1966
Telephone	760-278-3389	760-389-2289
E-Mail Address	joe.smith@surferdude.org	eo.e.nwuer@beusorpdqo.org
SSN	076-39-2778	076-28-3390
CC Number	3678 2289 3907 3378	3846 2290 3371 3378
Business URL	www.surferdude.com	www.sheyinctao.com
Fingerprint		Encrypted
Photo		Encrypted
X-Ray		Encrypted
Healthcare / Financial Services	Dr. visits, prescriptions, hospital stays and discharges, clinical, billing, etc. Financial Services Consumer Products and activities	Protection methods can be equally applied to the actual data, but not needed with de-identification

Protection Granularity: Field Protection

Production Systems

Encryption

- Reversible
- Policy Control (authorized / Unauthorized Access)
- Lacks Integration Transparency
- Complex Key Management
- Example !@#\$%a^.,mhu7///&*B()_+!@

Non-Production Systems

Masking

- Not reversible
- No Policy, Everyone can access the data
- Integrates Transparently
- No Complex Key Management
- Example 0389 3778 3652 0038

Protection Granularity: Field Protection

Production Systems

Encryption

- Reversible
- Policy Control (authorized / Unauthorized Access)
- Lacks Integration Transparency
- Complex Key Management
- Example `!@#%a^.,mhu7///&*B()_+!@`

Vaultless Tokenization / Pseudonymization

- Reversible
- Policy Control (Authorized / Unauthorized Access)
- Integrates Transparently
- No Complex Key Management
- Business Intelligence Credit Card: `0389 3778 3652 0038`

Non-Production Systems

Masking

- Not reversible
- No Policy, Everyone can access the data
- Integrates Transparently
- No Complex Key Management
- Example `0389 3778 3652 0038`

Harvard Professor Re-Identifies Anonymous Volunteers In DNA Study

- New technology or analytics tools may open routes to re-personalise anonymous data and may require further measures to keep data sets anonymous
- More data will over time make the issue worse
- Sweeney's now famous re-identification of Weld's hospitalization data using voter list information in his 2010 paper

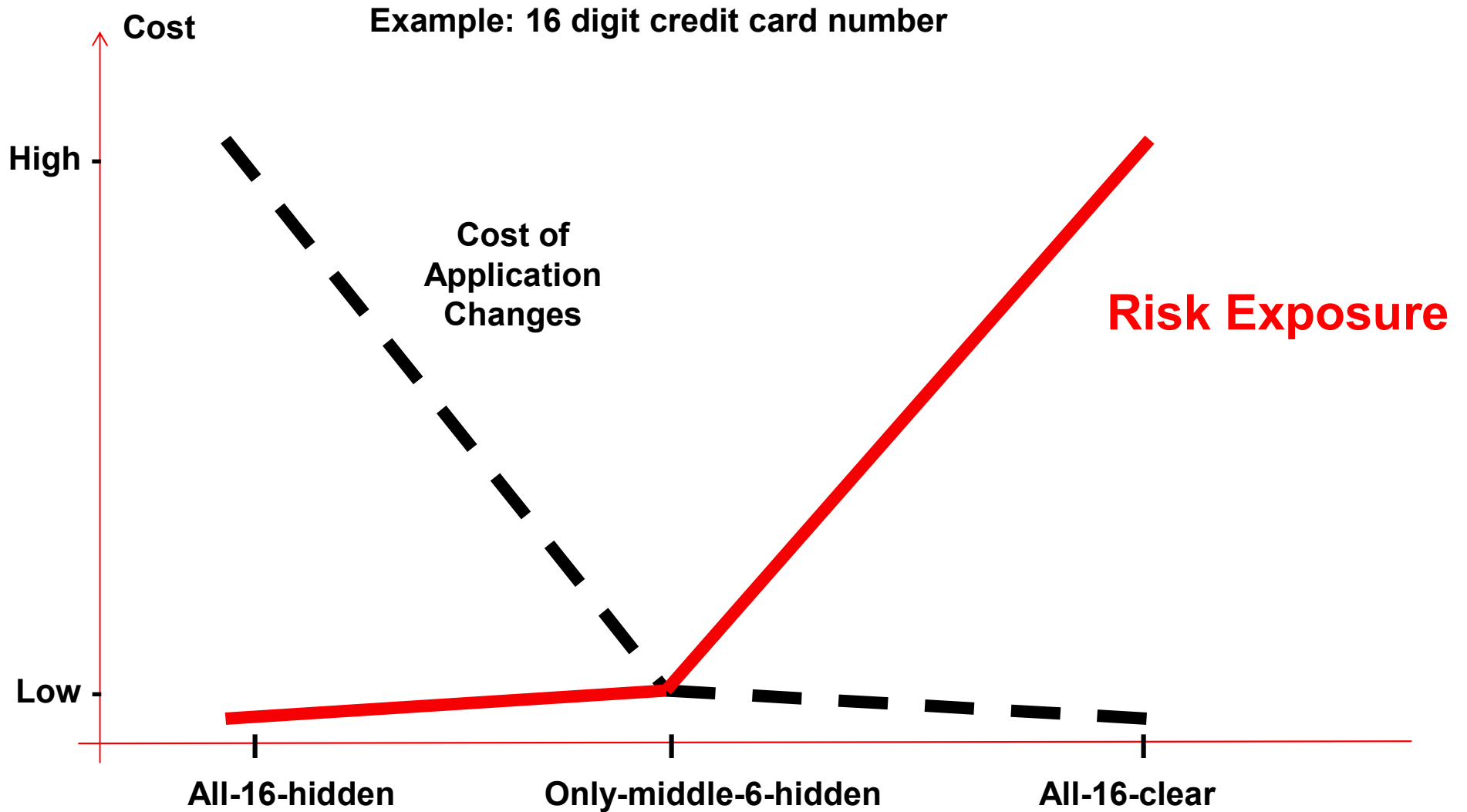


Data Masking Considerations

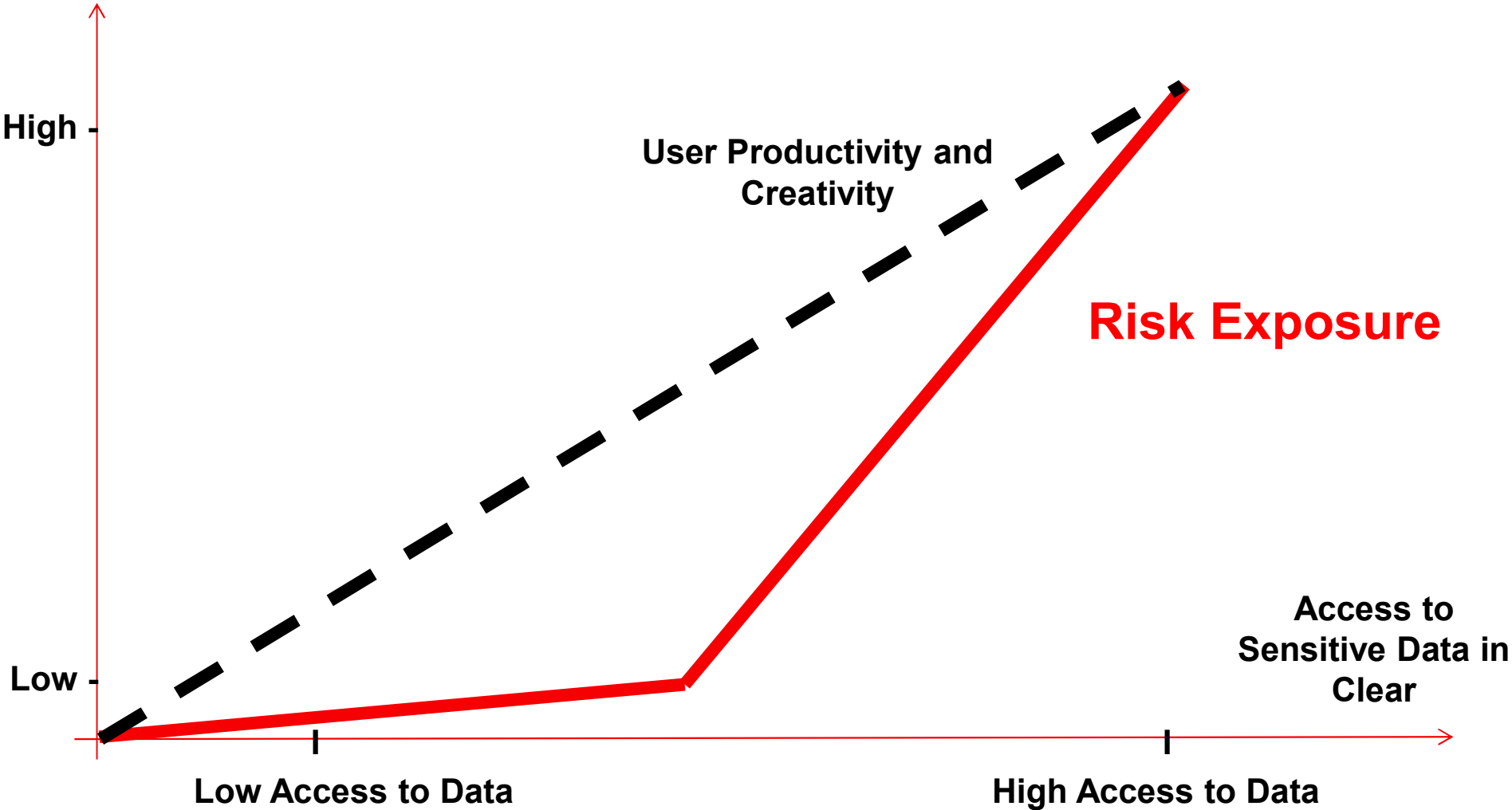
NIST IR 8053 - De-Identification of Personal Information

- De-identified data can be re-identified
- Disagreement regarding re-identification risk
- HIPAA's Safe Harbor de-identification is not firmly rooted in theory
- No accepted standards for testing the effectiveness of a de-identification
- Use technical controls

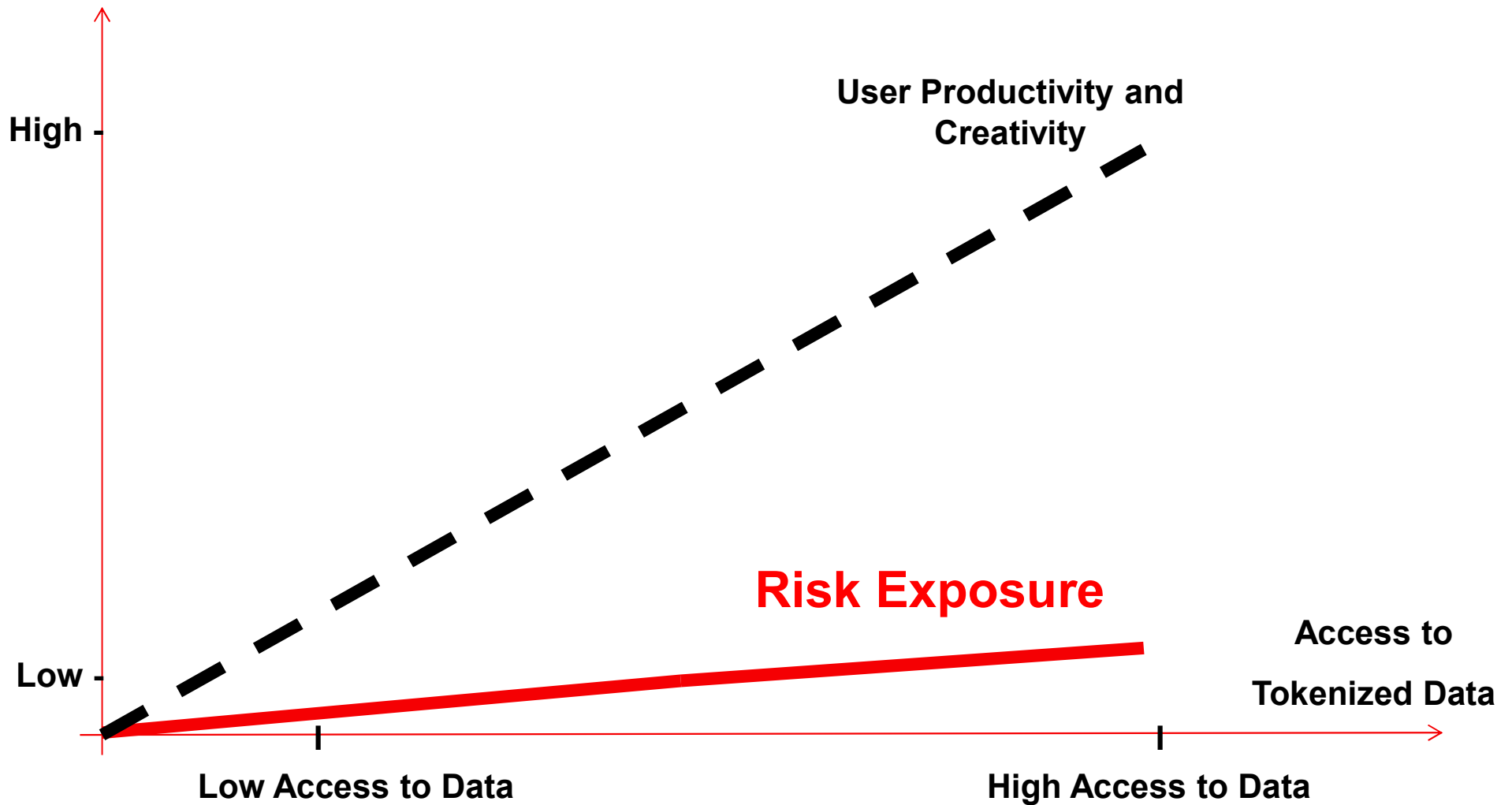
Partial Protection of Data Fields



Traditional Access Control



Fine Grained Protection of Data Fields



Securing Big Data

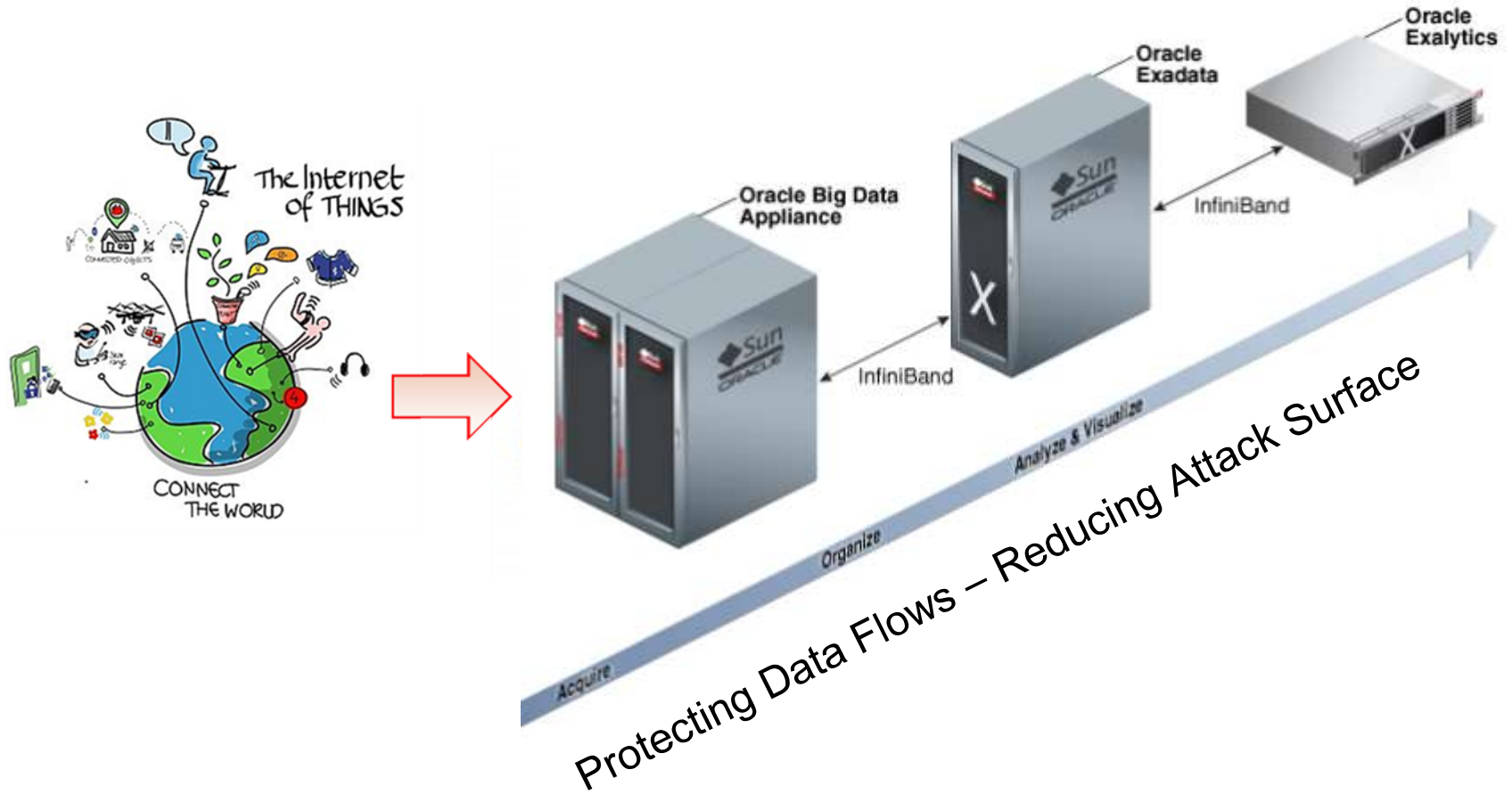
Big Data Needs a Data-Centric Security Focus

- CISOs should not treat big data security in isolation, but require policies that encompass all data
- New data-centric audit and protection solutions and management approaches are required
- Big data initiatives require data to move between structured and unstructured data silos, exposing incoherent data security policies that CISOs must address to avoid security chaos

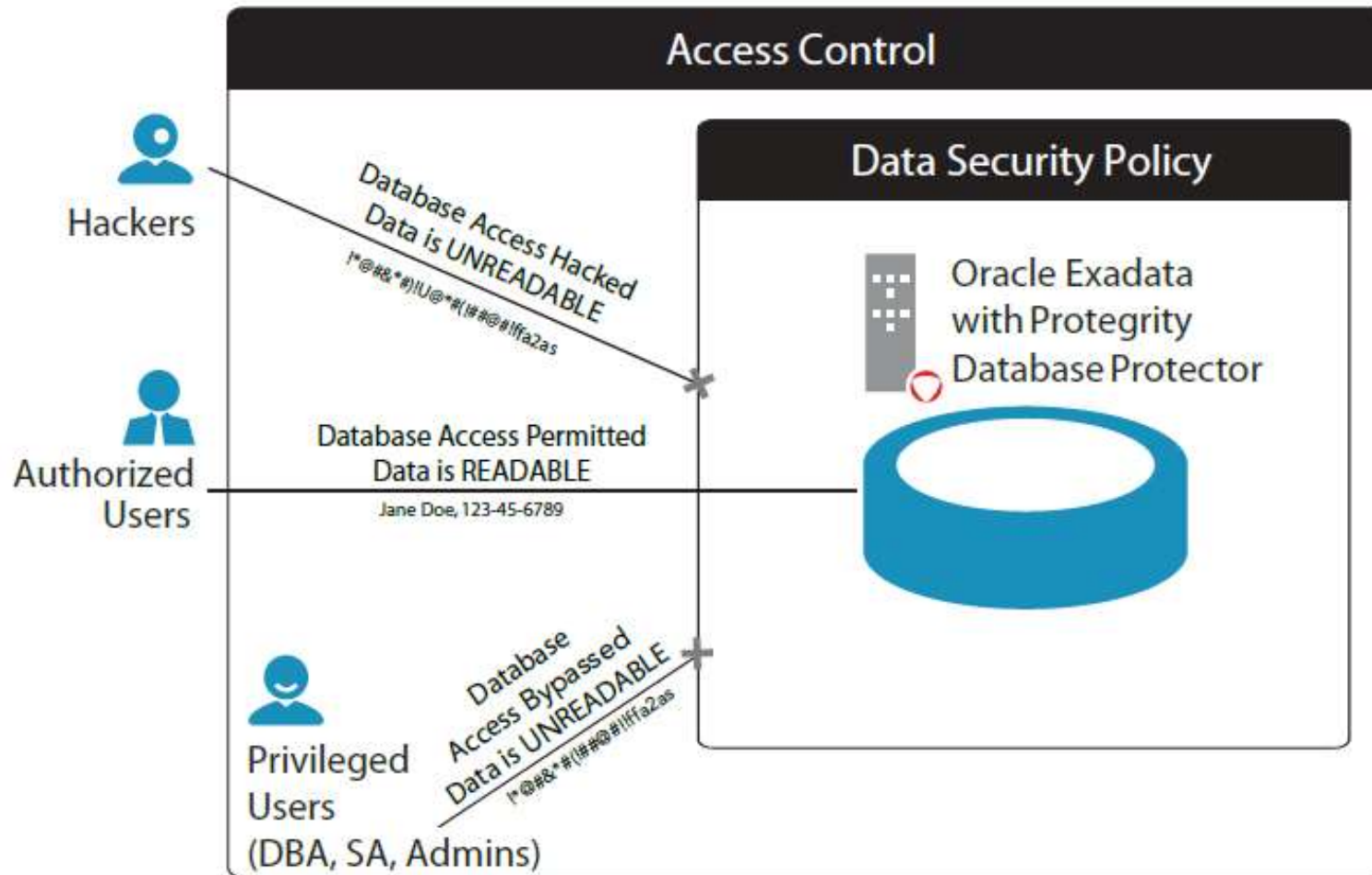
Gartner[®]

Source: Gartner – **Big Data Needs a Data-Centric Security Focus**, 2014

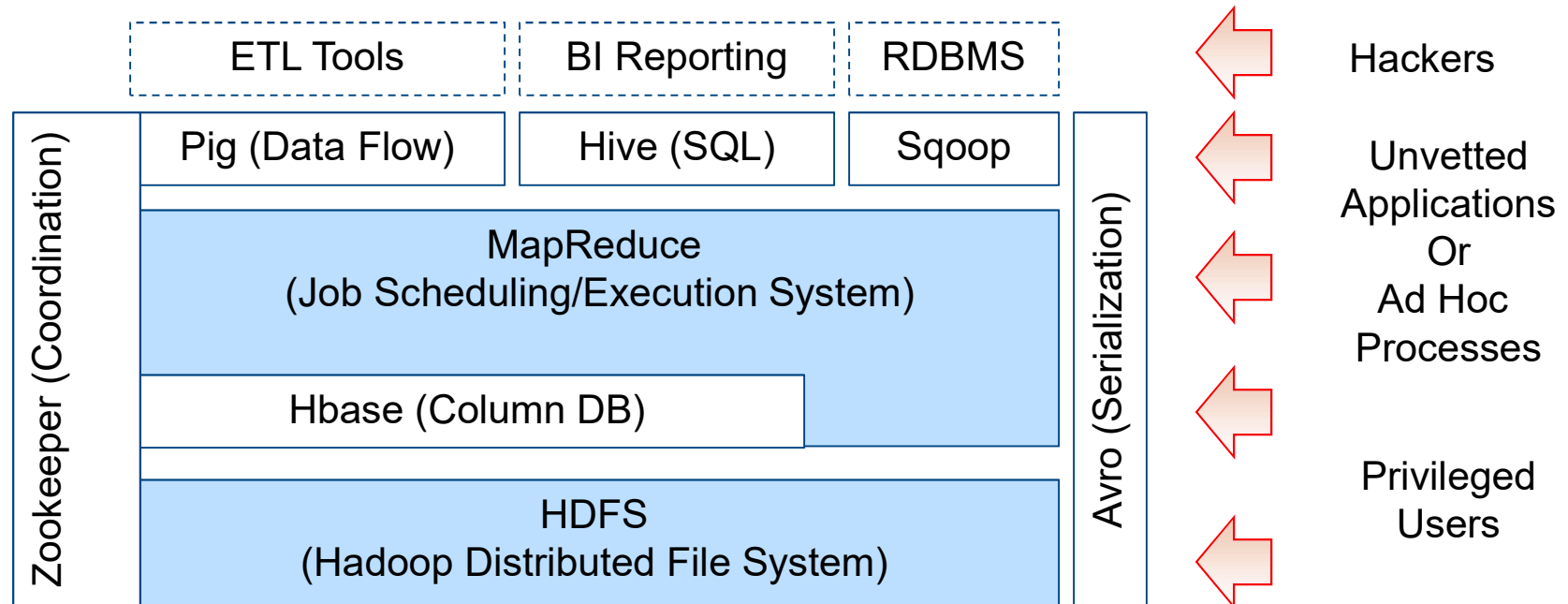
Oracle's Big Data Platform



Oracle's Exadata

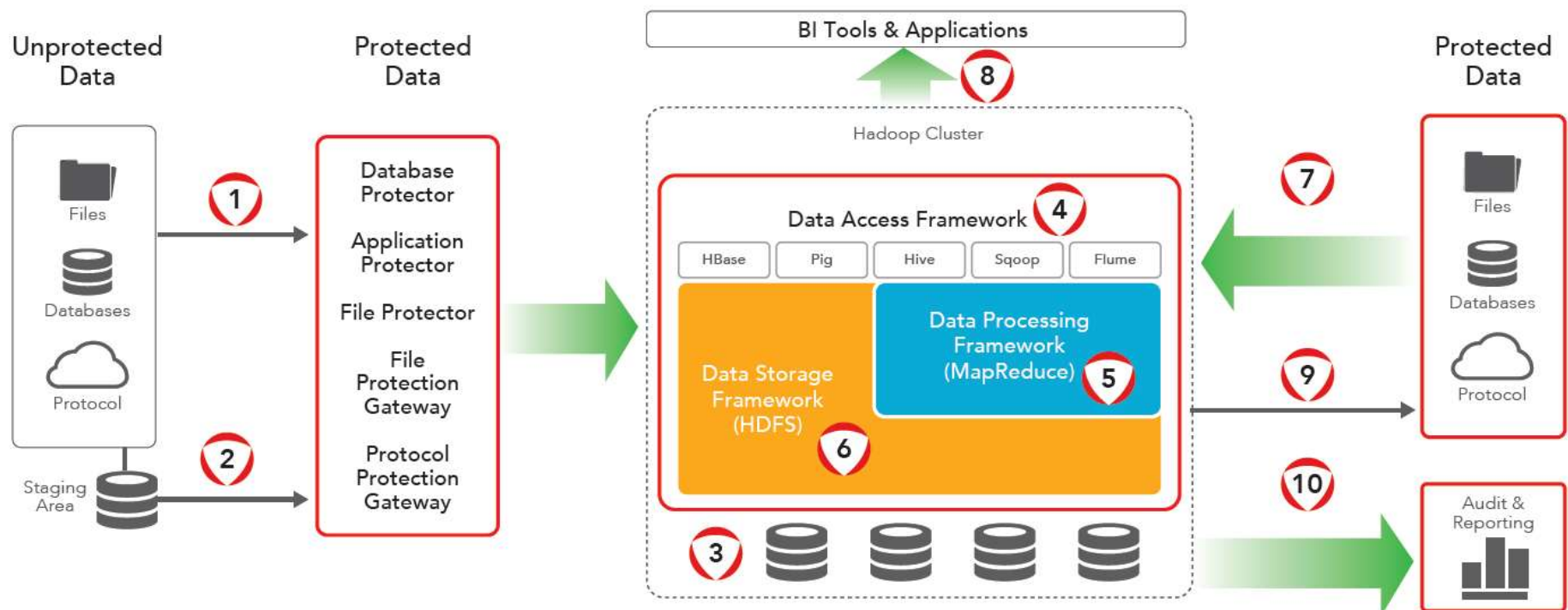


Many Ways to Hack Big Data



Source: <http://nosql.mypopescu.com/post/1473423255/apache-hadoop-and-hbase>

Securing Big Data



1. Data protection at database, application or file
2. Data protection in a staging area

3. Volume encryption in Hadoop
4. Hbase, Pig, Hive, Flume and Scope using protection API
5. MapReduce using protection API
6. File and folder encryption in HDFS
8. Export de-identified data

7. Import de-identified data
9. Export identifiable data
10. Export audit s for reporting

Summary

- Exponential growth of data generation
 - New business models fueled by Big Data, cloud computing and the Internet of Things
 - Creating cybercriminal's paradise
- Challenge in this interconnected world
 - Merging data security with data value and productivity.
- Urgently need a data-centric strategy
 - Protect the sensitive data flowing through digital business systems
- Solutions to bring together data insight & security
 - Safely unlock the power of digital business



Thank you!

Questions?

Please contact us for more information

www.protegrity.com

Ulf.Mattsson AT protegrity.com

