The challenge that organizations face is how best to protect their sensitive data while balancing ease of use, cost, and scale.

### Introduction

Over the last few years, audit controls, compliance regulations, and damage to reputations due to data breaches have continued to increase and show no signs of slowing down. As a result, organizations have experienced difficulties with managing change and developing strategies that mitigate risks associated with accidental data disclosure, while enabling cost-effective IT operations. Additionally, analysis of attack trends over recent years indicates that sensitive data continues to be the focus of modern cybercrime, making that a top priority for where to focus enterprise IT security efforts. The challenge that organizations face is how best to protect their sensitive data while balancing ease of use, cost, and scale.

Another related trend in protecting sensitive data is the rapid adoption of cloud services within organizations, whether at the platform, infrastructure, or product level. Depending on the size, sophistication, and function of an organization's traditional physical infrastructure, the path to a cloud-enabled or cloud-based operating environment may vary significantly from organization to organization. As assets, applications, and data flows move into, through, and around the cloud environment, new challenges emerge around ensuring the security of sensitive data. By taking a data-centric approach to security, organizations can experience the operational benefits of cloud infrastructure and maintain best-practice data security within a single construct and at significant cost savings.

The International Data Corporation (IDC) predicts that, by 2016, there will be an 11 percent shift of IT budget away from traditional in-house IT delivery toward various versions of cloud computing. By 2017, 35 percent of new applications will use cloud-enabled, continuous delivery, allowing faster DevOps life cycles to streamline rollout of new features and business innovation. By 2018, more than 60 percent of enterprises will have at least half of their infrastructure on cloud-based platforms.¹

A data-centric approach to securing cloud infrastructure—defining technical solutions based on the format and life cycle of data—can save organizations capital expenditures and the costly man-hours required to track down existing infrastructure complexities, nuances, and out-of-date systems. The ability to leverage a low-cost platform to pool and aggregate security monitoring metadata facilitates the ability to perform advanced analytics within a cloud infrastructure environment.

This paper describes the current state and challenges faced by organizations that rely on traditional on-premise systems in today’s rapidly shifting and increasingly complex environment. In response to these challenges, we describe the motivation, technical approach, and business advantages associated with Booz Allen Hamilton's experience with migrating traditional IT environments into Amazon Web Services, while not only maintaining compliance with common frameworks (e.g., PCI, HIPAA, FISMA), but also achieving best-in-class data protection.

---

Current State

Corporate network infrastructures typically consist of multiple devices for different functions, with heavy dependency on legacy systems to support critical business functions. For example, an organization might have multiple applications supported by a series of different databases and mainframes depending on the organizational function or level of sensitivity. Typically, an organization has distinct sets of systems that are used for internal finance, marketing functions, and core business processes. Although all of these functions fall under the umbrella of the company, the organization’s leaders might not want all three groups to have access to each other’s information. In addition to internal controls, the protection of all three groups entails mechanisms that safeguard external threats; examples include firewalls, DMZs, intrusion protection systems, and enough IT personnel to effectively administer these servers. Figure 1 illustrates the complexities of a typical networking infrastructure for a production environment either in the cloud or on-premises.
Although there are potentially different sets of hardware and software for each line of business within the organization, there is a large amount of redundancy within the organization’s overall infrastructure. In addition, because the traditional form of corporate IT infrastructure is to have all hardware assets on premise, an organization will also have to account for power, space, and cooling in addition to the cost of the physical devices. A realistic cost estimation for an IT infrastructure is nearly impossible to calculate without knowing the specific goals of an organization. However, a 2013 worldwide cost breakdown of non-personnel IT expenditures was $2.06 trillion dollars, with the US spending a majority of that money.

Even with such large expenditures for resources, especially if the equipment is top of the line when purchased, the anticipated life cycle is between 3 and 4 years, based on a static R&D slope.

**Lowering CapEx in favor of Opex: Transitioning to the Cloud**

According to an article on Infoworld.com, “The debate between operating expense (aka opex, the cloud’s approach) and capital expense (aka capex, the on-premise approach) is waged daily at companies. Although there are trade-offs no matter what a firm chooses, it’s clear that opex is increasingly favored. In the age-old rent-versus-buy debate, the cloud is making rental very compelling, especially as managed cloud environments begin to implement tools that provide automated spin-up/spin-down to avoid excessive consumption of resources (and higher costs).”

Clearly, cloud computing can save time and money by reducing the costs and complexity of an infrastructure with the ability to adjust capacity for memory and processing power on-demand. Cloud computing also increases opportunities for innovation by reducing time to market all while enhancing the security of the environment and the data it contains. According to Andrew McAfee in an article from the Harvard Business Review, in today’s IT infrastructure, only about 11 percent of the IT budget is used for innovation, and the rest is spent on maintenance and infrastructure. The transition to a cloud infrastructure frees up money for organizations to be more innovative. However, migration to a cloud infrastructure is often met with increased security scrutiny due to the perception of less control and insight over management of security controls and infrastructure.

**A Data-Centric Approach to Security**

The traditional approach to securing the network is to focus on the enterprise boundaries and endpoints, an effort that boils down to essentially trying to “keep the attackers out.” With the proliferation of data beyond the traditional corporate border, such as through a cloud migration, the mindset has to change to accommodate the fact that sensitive corporate data can reside and be utilized almost anywhere, and that data owners can’t control infrastructure around data in order to protect it. Security practitioners have begun to develop strategies and technologies to provide security at the data element level to help reduce the impact of an infrastructure compromise.

Securing data at granular levels within a massively scalable cloud comes with a few challenges. The most pressing is the ability to discover sensitive elements within a largely unstructured repository (in the form of archived and indexed files, or in the form of data

---

parsed into key-value pairs for analytic processes). Although there are challenges, solutions exist in the form of controls that can be applied to secure data that is considered at risk. Once sensitive elements are identified, then appropriate controls can be applied, including format preserving encryption, tokenization, and masking, all of which provide strong security without impacting the structure of the data (and hence any database schemes or application design).

Another challenge is that for analytics purposes, data can rarely be encrypted without disrupting analysis. In these cases, leveraging systems to carefully tag data ensures that data access is selectively granted only to entities (which may include humans or applications) with the correct roles or attributes. Tags can be further added to help support discovery and analytics processes for data monitoring and other uses. Tagging approaches vary widely. Historical approaches leverage native operating system level attributes such as Alternate Data Streams and Extended Attributes; however, these approaches pose challenges because they are not applicable to emerging cloud analytic systems such as Hadoop. More advanced approaches leverage metadata wrappers to apply a virtually limitless number of tags to data at the file level that do not affect key file properties (such as size and time stamps), but can be securely bound to the data assets to ensure secure data management and processing. These tags can then be leveraged to provide ultra-granular encryption, redaction and access control to the files, regardless of the infrastructure. Furthermore, these tags provide fidelity to ingest processes for analytics platforms (such as Hadoop) to embed security at the data level for downstream analytics.

**Implementing a Data-Centric Security Solution in the Cloud**

Based on consistent discussions with Fortune 500 companies regarding the challenges with securely leveraging public cloud infrastructures such as Amazon Web Services (AWS) for business process while still maintaining security programs that satisfied customers and auditors alike, Booz Allen embarked on an 8-month journey to prove out a data-centric security program in AWS.

As a first step, it was critical to understand the different security and management responsibilities between AWS and the cloud customer before discussing the solution. AWS is responsible for core foundational services of the infrastructure such as computing, storing, database, and network functionality. These foundational services are segmented within AWS’s global infrastructure to provide the speed, availability, and security that cloud customers are looking for. AWS also provides identity access management but limited only for accounts used to manage the AWS infrastructure. This leaves the cloud customers with the responsibilities for configuration, management, and protection of the infrastructure to keep it functional while still protecting customer data and meeting compliance regulations. Figure 3 illustrates the responsibilities that are managed by AWS and those that the customer is responsible for managing.

Booz Allen used a phased approach to address the need to maintain a functional, secure, and compliant infrastructure.

The first phase gathered information that consisted of the following components:

1. Proper scoping of the intended solution and identification of any currently deployed components
2. Conducted subject matter expert interviews from all key areas of the organization, including cross-domain departments and executive management
3. Reviewed relevant technical implementations to ascertain what the current configurations look like
The next phase consisted of evaluating existing management and technical infrastructures against industry best practices, regulatory requirements, and established corporate baselines.

Using all of this detailed information, our team designed not only the required technical components, but also a roadmap of how the solution should be implemented to ensure that all business and regulatory requirements are satisfied.

**Booz Allen’s Transition Strategy**

Booz Allen has developed a phased cloud computing transition methodology designed to address the issues and risks associated with migrating an existing system to the cloud. Deploying cloud-computing solutions requires both a short-term and a long-term strategy.

One of the cloud solutions Booz Allen Hamilton has implemented is the protection of credit card information through commercial off the shelf software solutions. The concept for this use case is, once a customer’s credit card information is entered into the client’s system, one of two protection methodologies is employed. The first method is called tokenization, where the credit card string is recognized and then tagged within the server.

The second method is obfuscation, which occurs when the leading digits of the credit card number are recognized by the system and blocked out (obfuscated). Only the last digits (normally the last 4) will be visible; all the other digits will appear as an “x”. There are two types of solutions used in this model. One is a “data discovery” service used to aggregate Active Directory user and group details, Access Control List (ACL) information and all data access events—without requiring native OS auditing—to build a complete picture of who can and who is accessing data, and who should have their access revoked. It also leads IT to rightful data owners, so the right people can ensure appropriate access and usage. The key is to use a solution that is specifically designed for unstructured data. The other solution is

---

![Diagram: Shared Responsibility Model for Infrastructure Services](image)

data encryption and tokenization for structured and unstructured data types that enables cost-effective payment card industry (PCI) compliance, scope reduction, and secure analytics. The flow chart in Figure 4 illustrates our use case.

Booz Allen’s cloud computing offerings focus on the following areas:

• Cloud Strategy and Planning—Includes portfolio management, transition plans and sequencing, return on investment and risk management, cloud policy consulting, cloud governance policies, cloud strategic planning, cloud data strategies, technology forecasting, and a range of related strategy services

• Cloud Application Development—Includes the development, deployment, and management of applications in the cloud; the re-engineering of legacy systems so they are cloud enabled; data management; and application and service integration

• Cloud Infrastructure Services—Includes infrastructure assessment and analysis, design consultation, and private cloud building

• Cloud Security—Includes certification and accreditation of cloud solutions, identity management, cloud segmentation, security audit, application and data obfuscation, and security integration

**Conclusion**

While many organizations experience difficulty keeping pace with environmental changes, solutions for optimizing spend by using cloud technologies are beginning to emerge as cost-optimal solutions, with a measured data-centric approach, that represent a secure, realistic alternative to traditional infrastructure scaling practices.

A data-centric approach to securing cloud infrastructure can save organizations vast capital expenditures and countless man-hours required to manage traditional infrastructure, and also provide the ability to leverage a low-cost platform that pool and aggregate data in order to perform advanced, secure analytics for the business.

Figure 4: PCI Use Case
To learn how Booz Allen Hamilton can help your business thrive, contact:

**Steve Coraggio**  
Principal  
coraggio_stephen@bah.com  
Tel +1 917-305-8004

**Chris White**  
Senior Associate  
white_chris@bah.com  
Tel +1 917-305-8082

[www.boozallen.com/cybersolutions](http://www.boozallen.com/cybersolutions)

---

Booz Allen Hamilton has been at the forefront of strategy and technology for more than 100 years. Today, the firm provides management and technology consulting and engineering services to governments in the civil, defense, and intelligence markets, global corporations, and not-for-profit organizations. Booz Allen partners with private and public sector clients to solve their most difficult challenges. Headquartered in McLean, Virginia, the firm employs more than 22,500 people globally, and had revenue of $5.27 billion for the 12 months ended March 31, 2015. To learn more, visit [www.boozallen.com](http://www.boozallen.com). (NYSE: BAH)