BUILDING BLOCK WHITE PAPER

# DATA INTEGRITY

## Reducing the impact of an attack

Michael J. Stone
Donald Tobin
National Cybersecurity Center of Excellence
National Institute of Standards and Technology

Harry Perper
Sarah Weeks
The MITRE Corporation

DRAFT
November 23, 2015
di-nccoe@nist.gov

National Institute of Standards and Technology
U.S. Department of Commerce
NCCoE
NATIONAL CYBERSECURITY CENTER OF EXCELLENCE

The National Cybersecurity Center of Excellence (NCCoE) at the National Institute of Standards and Technology (NIST) works with industry, academic, and government experts to find practical solutions for businesses' most pressing cybersecurity needs. The NCCoE demonstrates how standards and best practices established by NIST and other organizations can be applied in technical reference architectures and serves as a collaboration hub where small businesses, market-leading companies, government agencies, and individuals from academia work together to address broad cybersecurity problems. To learn more about the NCCoE, visit https://nccoe.nist.gov. To learn more about NIST, visit http://www.nist.gov.

NCCoE building blocks address technology gaps that affect multiple industry sectors. They represent core capabilities that can and should be applied across industry cybersecurity and business use cases.

## Abstract

Constant threats of destructive malware, malicious insider activity, and even honest mistakes create the imperative for organizations to be able to quickly recover from an event that alters or destroys any form of data (database records, system files, configurations, user files, application code, etc.). Further, businesses must be confident that recovered data is accurate and safe. The National Cybersecurity Center of Excellence (NCCoE)—in collaboration with members of the business community and vendors of cybersecurity solutions—is creating an example solution to address these complex data integrity challenges.

Multiple systems need to work together to prevent, detect, notify, and recover when data integrity is jeopardized. This project explores methods to effectively monitor and detect data corruption in commodity components (server, operating system, applications, and software configurations) as well as custom applications and data. It also explores issues of auditing and reporting (user activity monitoring, file system monitoring, database monitoring, scanning backups/snapshots for malware and rapid recovery solutions) to support recovery and investigations. To address real-world business challenges around data integrity, the resulting example solution will be composed of open-source and commercially available components. Ultimately, this project will result in a publicly available NIST Cybersecurity Practice Guide—a description of the solution and practical steps needed to implement an example solution that addresses these existing challenges.

## Keywords

*business continuity, malware, ransomware, integrity, attack vector, data recovery, malicious actor*

## DISCLAIMER

Certain commercial entities, equipment, or materials may be identified in this document in order to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by NIST or NCCoE, nor is it intended to imply that the entities, materials or equipment are necessarily the best available for the purpose.

## COMMENTS ON NCCOE DOCUMENTS

Organizations are encouraged to review all draft publications during public comment periods and provide feedback. All publications from NIST's National Cybersecurity Center of Excellence are available at http://nccoe.nist.gov.

Comments on this publication may be submitted to: DI-NCCoE@nist.gov

Public comment period: *November 23, 2015 to January 22, 2016*

## CONTRIBUTORS

We gratefully acknowledge the contributions of: Ted Kolovos and Leslie Anderson

## Table of Contents

## List of Figures

## List of Tables

# 1. Executive Summary

The NCCoE is responding to industries that have identified the problem of data corruption by malicious actors.  In order to remain operational, an organization should be able to quickly recover from a data integrity attack and trust that the recovered data is accurate, complete, and free of malware.

Malicious actors are intent on disrupting operations or achieving financial gain and will corrupt critical information maintained by an organization to achieve their goals. Information such as customer data, transaction records, and correspondence are typically the targets for unauthorized insertion, modification or deletion. These types of data integrity attacks, especially when they target an entire organization, can lead to catastrophic impacts and impair the company's ability to operate. There is evidence of malicious actors attempting and successfully corrupting high-value data across various industries.  In other cases, systems are held hostage by a specific type of malware called ransomware, which encrypts various types of data files on a system so the users can no longer use them, and then demands a payment for the decryption keys for the files.

There are many attack vectors that a data integrity attack can utilize to gain access to corporate systems. Typical attack vectors include phishing (email), drive-by website downloads, unmitigated vulnerabilities on externally facing resources, and malicious/infected attachments. Once the malware is operational it can use multiple techniques to spread throughout an organization, exfiltrate data, and corrupt it.  The data at risk includes but is not limited to: active, current data, back-up data, system configurations, and baseline system operating systems.

The project described in this document will help organizations address the issue of detecting and recovering from a data integrity attack on its data. The data includes databases, stored files, configurations, operating system files, as well as other types of files. These types of files can be corrupted before or after they have been stored to a back-up system. One example is undetected malware that is stored in system back-ups prior to detection.

The project will include an architectural depiction and example solution that can reduce the risk and impact of a data integrity attack.  The solution will integrate commercial and open source products.  The project will result in a NIST Cybersecurity Practice Guide.  The practice guide provides a description of the practical steps needed to implement the proposed architecture.  Organizations will be able to use the practice guide to influence architectural changes that enhance their ability to recover from data corruption attacks.

## 2. Business Value

Corporate resilience against data corruption is critical to business continuity, cost avoidance and regulatory compliance.  The potential business benefits of the data integrity solution explored by this project include:

- Detecting back-up data tampering attempts
- Reducing the impact of a data corruption attack
- Reducing downtime caused by data corruption
- Improving IT resource efficiency through automated testing
- Improving trustworthiness of back-up data
- Reducing the negative impact to the reputation of an organization due to data corruption events
- Providing management with overall health and status of the organization's data and continuity of operations

## 3. Description

### Who should read this document?

The intended audience for this document is CIO, CISO and IT management personnel interested in mitigating the threat of data corruption caused by malicious actors as well as unintentional human or computer error.

### Goal

The goal of this project is to help prevent the use of corrupted data when recovering systems from back-up storage. The solution will provide guidance for incorporating data corruption prevention, detection, and recovery into a corporate IT architecture. In addition to protecting critical enterprise information, corruption alerts and activity logs provide the security analysts with indictors of malicious activity. The project will explore methods to monitor and protect commodity components (operating systems, applications, and software configurations), custom applications, and data (database and files). It will produce an architecture that includes components that will integrate detection and notification of data corruption events as well as approaches to automation for recovery from such events.

The project will also include auditing and reporting (user activity monitoring, file system monitoring, database monitoring, scanning backups/snapshots for corruption or malware) to support investigations.

### Background

The NCCoE, working with the organizations across the set of critical infrastructure industries, including information sharing and analysis centers (ISACs) identified the need for a data integrity solution. The center held a workshop to identify key issues that affect consumer data protection, encapsulated in NISTIR 8050. This document identified data integrity (among other items) as a key cybersecurity issue that needs to be addressed. The need arises from the recognition that malicious actors are devising methods of corrupting data within organizations. The data corruption includes data modification as well as data destruction. In addition the center met with representatives the financial sector ISAC (FS-ISAC) for guidance, and has been working with the FS-ISAC Destructive Malware Data Integrity Task Force.

## Scope

This project will answer specific questions pertaining to data integrity and recovery such as:

- What data was corrupted? When was it corrupted? How it was corrupted? Who corrupted it?
- Do any other events coincide with this corruption?
- What was the impact of the data corruption? (Systems affected, timelines, etc.)
- Which backup version should be used to recover data?

This project will address three solution areas:

1) File system integrity solution to allow recovery from trusted backups and snapshots.
2) Database integrity solution with transactions and versioning to allow for rollbacks to a known good state.
3) An overall automated system that incorporates the previous two areas and includes the following:

    a. Activity logging and monitoring.
    b. Versioning and journaling file system solutions that incorporate formal change management procedures covering both normal and emergency changes to systems.
    c. Restoration of desktops, applications, and critical services quickly after cyber incidents.
    d. Alert systems to notify administrators when baseline controls are changed on critical systems.

## Assumptions

This project assumes data integrity needs to be addressed for each of the following components:

- Operating Systems

- File System
- Applications (including custom code)
- Databases
- Virtual machines (including software defined networks)

## 4. Example Scenarios

The example scenarios below illustrate some of the challenges that this project will address. The relevant functions and categories from the NIST Framework for Improving Critical Infrastructure Cybersecurity (CSF) that can be employed to mitigate the events throughout the attack are listed below.

### Scenario 1 - Ransomware

For financial gain, an organized crime group has set up multiple seemingly legitimate domains that contain destructive malware to be automatically downloaded and discreetly/silently installed, without the user's knowledge, when a website on the domain is visited. Once the malware is installed it can encrypt the organization's file system and require a ransom payment in order to unencrypt the files to be restored. Left unmitigated, the malware on one system is designed to move laterally within the network to other client and server systems within an organization's network, encrypting those systems and demanding ransom before access to those systems can be restored.

The project addresses respond and recover CSF categories

- Malware encrypts files and displays notice demanding payment for decryption
  - Respond/Recover:
    - notify security (DE.DP-4, RS.CO-2, DE.EA-5)
    - file integrity monitor (PR.DS-1, PR.DS-6, PR.PT-1)

The project does not address these protect and detect CSF categories

- User receives phishing email with executable attachment
  - Protect/Detect: email security and attachment scanning
- User runs the attachment containing malware which installs and infects the user's machine
  - Protect/Detect: Host-based Anti-malware, application whitelisting, EMET, sandboxing/virtualization
- Malware sets up command and control where it receives instructions and cryptographic keys
  - Protect/Detect: Host-based firewall/IDS, network-based firewall/IDS

### Scenario 2 - Data destruction

An adversary wishing to impact the operations of a major lending or banking institution launches a spear-phishing campaign against individuals in the target corporation.  Once any of the human targets clicks on a link or attachment, the malware downloads and installs itself on that user's machine, and immediately starts looking to infect other systems across the enterprise.  At a predetermined time, the malware starts encrypting all data on the infected machines.  Then it writes over the original unencrypted content and deletes the encryption keys.

The project addresses respond and recover CSF categories

- Malware destroys data on user's machine
  - Respond/Recover:
    - back-ups(PR.DS-1, PR.IP-4)
    - file integrity monitor (PR.DS-1, PR.DS-6, PR.PT-1)

The project does not address these protect and detect CSF categories

- User receives phishing email with executable attachment
  - Protect/Detect: email security and attachment scanning
- User runs the attachment containing malware which installs and infects the user's machine
  - Protect/Detect: Host-based Anti-malware, application whitelisting, EMET, sandboxing/virtualization
- Malware performs reconnaissance and attempts to spread throughout the enterprise.
  - Protect/Detect: network-based firewall/IDS, use of P-VLANs

### Scenario 3 - Data Manipulation (insider threat)

A disgruntled employee seeks to harm his employer by damaging its business operations, brand, or reputation, and launches a campaign to modify company records. Using authorized credentials he has, or acquires, he modifies database entries over time.  Because the modifications he makes appear to be legitimate, a significant amount of data is corrupted before he is discovered.

The project address respond and recover CSF categories

- User modifies a configuration file in violation of established baselines
    - Protect/Detect:
        - **file integrity monitor** (PR.DS-1, PR.DS-6, PR.PT-1)
        - **user activity auditing** (DE.CM-3, PR.PT-1)
- Administrator modifies a user’s file
    - Protect/Detect:
        - **file integrity monitor** (PR.DS-1, PR.DS-6, PR.PT-1)
        - **user activity auditing** (DE.CM-3, PR.PT-1, DE.AE-1)
- Administrator and/or script modifies data in a database
    - Protect/Detect:
        - **database transaction auditing** (PR.DS-1, PR.PT-1, DE.CM-1)

# 5. CURRENT CHALLENGES

## Detecting Data Corruption in Back-ups

Data back-up software and systems focus on accurately restoring data as originally stored.  This approach is effective for data that is known to be 100% error free and un-corrupted.  These systems generally do not provide a retroactive data testing scheme to test data for corruption by insiders or malicious applications while in storage.

## Detecting malware in back-up data

Data back-up software and systems generally do not have manual or automated testing capabilities to identify and remediate malware in backed up data.  Malware detection is typically done at runtime in operational systems by anti-virus/anti-malware software.  In addition the software is not designed to test data in non-realtime.  Malware that is designed to be dormant for periods of time may not be detectable until active with current anti-virus/anti-malware software. A time-shifting, self-contained testing environment that can emulate the passage of time may be able to detect time-sensitive or time-delayed malware activity in addition to malware with signatures for activity monitoring that was unknown at the time the backup was completed.

## Automation of Backup Data Testing

Back-up data testing is typically used to verify that back-up data can be used to restore systems to operational readiness.  Data back-up software and systems generally do not offer automated backup data integrity or malware testing capabilities.

## 6. Relevant Standards and Guidelines

NIST SP 800-14, Generally Accepted Principles and Practices for Securing Information Technology Systems

http://csrc.nist.gov/publications/nistpubs/800-14/800-14.pdf

NIST SP 800-27A, Engineering Principles for Information Technology Security (A Baseline for Achieving Security) Revision A

http://csrc.nist.gov/publications/nistpubs/800-27A/SP800-27-RevA.pdf

NIST SP 800-33, Underlying Technical Models for Information Technology Security

http://csrc.nist.gov/publications/nistpubs/800-33/sp800-33.pdf

NIST SP 800-34, Contingency Planning Guide for Federal Information Systems

http://csrc.nist.gov/publications/nistpubs/800-34-rev1/sp800-34-rev1_errata-Nov11-2010.pdf

NIST SP 800-86, Guide to Integrating Forensic Techniques into Incident Response

http://csrc.nist.gov/publications/nistpubs/800-86/SP800-86.pdf

NIST SP 800-53, Security and Privacy Controls for Federal Information Systems and Organizations

http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-53r4.pdf

NIST SP 800-160, Systems Security Engineering, An Integrated Approach to Building Trustworthy Resilient Systems

http://csrc.nist.gov/publications/drafts/800-160/sp800_160_draft.pdf

ISO/IEC 27001, Information Technology – Security Techniques – Information Security Management Systems

http://www.iso.org/iso/home/search.htm?qt=27001&sort=rel&type=simple&published=on

ISO/IEC 15408-1, Information technology – Security Techniques – Evaluation Criteria for IT Security – Part 1: Introduction

http://www.iso.org/iso/home/search.htm?qt=15408-1&sort=rel&type=simple&published=on

ISO/IEC 15408-2, Information technology – Security Techniques – Evaluation Criteria for IT Security – Part 2: Security

http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=46414

ISO/IEC 15408-3, Information technology – Security Techniques – Evaluation Criteria for IT Security – Part 3: Security Assurance Components

http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=46413

# 7. Desired Solution Characteristics

To address the three scenarios, this project will use a collection of commercially available technologies to demonstrate security and functional characteristics of a data integrity solution. The data integrity solution shall include the following characteristics:

- Automated data corruption testing
- Automated data corruption detection
- Automated data corruption event logging
- Secure data integrity monitoring and alerting information (checksums, off-site, hard-copy)
- Automated detection and reporting of all file modifications / creations / deletions
- Automated detection and reporting of all database modifications / creations / deletions
- Automated correlation of file changes and users
- Automated user activity recording
- Automated anomalous user activity detection
- Automated configuration management monitoring

# 8. Security Control Map

Table 1 maps the characteristics of the applicable standards and best practices described in the Framework for Improving Critical Infrastructure Cybersecurity (CSF), and other NIST activities. This exercise is meant to demonstrate the real-world applicability of standards and best practices, but does not imply that these characteristics will meet your industry's requirements.

| Solution Characteristic | NIST CSF Category | Informative References |
|---|---|---|
| Automated data corruption testing | PR.DS-1<br>PR.DS-6 | **NIST SP 800-53 Rev. 4** SC-28, SI-7<br>**ISO/IEC 27001:2013** A.8.2.3, A.12.2.1, A.12.5.1, A.14.1.2, A.14.1.3 |
| Automated data corruption detection | PR.DS-1<br>DE.CM-1 | **NIST SP 800-53 Rev. 4** AC-2, AU-12, CA-7, CM-3, SC-5, SC-7, SC-28, SI-4<br>**ISO/IEC 27001:2013** A.8.2.3 |
| Automated data corruption event logging | PR.DS-1<br>PR.PT-1 | **NIST SP 800-53 Rev. 4** AU Family, SC-28<br>**ISO/IEC 27001:2013** A.8.2.3, A.12.4.1, A.12.4.2, A.12.4.3, A.12.4.4, A.12.7.1 |
| Data integrity information must be secure | PR.DS-1<br>PR.DS-6 | **NIST SP 800-53 Rev. 4** SC-28, SI-7<br>**ISO/IEC 27001:2013** A.8.2.3, A.12.2.1, A.12.5.1, A.14.1.2, A.14.1.3 |
| Back-ups must be secure | PR.DS-1<br>PR.IP-4 | **NIST SP 800-53 Rev. 4** CP-4, CP-6, CP-9, SC-28<br>**ISO/IEC 27001:2013** A.8.2.3, A.12.3.1, A.17.1.2A.17.1.3, A.18.1.3 |
| Ability to detect and report on all file modifications/creations/deletions | PR.DS-1<br>PR.PT-1<br>DE.CM-1 | **NIST SP 800-53 Rev. 4** AC-2, AU Family, CA-7, CM-3, SC-5, SC-7, SC-28, SI-4<br>**ISO/IEC 27001:2013** A.8.2.3, A.12.4.1, A.12.4.2, A.12.4.3, A.12.4.4, A.12.7.1 |
| Ability to detect and report on all database modifications/creations/deletions | PR.DS-1<br>PR.PT-1<br>DE.CM-1 | **NIST SP 800-53 Rev. 4** AC-2, AU Family, CA-7, CM-3, SC-5, SC-7, SC-28, SI-4<br>**ISO/IEC 27001:2013** A.8.2.3, A.12.4.1, A.12.4.2, A.12.4.3, A.12.4.4, A.12.7.1 |
| Ability to correlate file change with user | PR.PT-1<br>DE.CM-1<br>DE.CM-3 | **NIST SP 800-53 Rev. 4** AC-2, AU Family, CA-7, CM-3, CM-10, CM-11, SC-5, SC-7, SI-4<br>**ISO/IEC 27001:2013** A.12.4.1, A.12.4.2, A.12.4.3, A.12.4.4, A.12.7.1 |
| User activity recording | PR.PT-1<br>DE.CM-3 | **NIST SP 800-53 Rev. 4** AC-2, AU Family, CA-7, CM-10, CM-11<br>**ISO/IEC 27001:2013** A.12.4.1, A.12.4.2, A.12.4.3, A.12.4.4, A.12.7.1 |
| User activity anomaly detection | PR.PT-1<br>DE.CM-1<br>DE.CM-3 | **NIST SP 800-53 Rev. 4** AC-2, AU Family, CA-7, CM-3, CM-10, CM-11, SC-5, SC-7, SI-4<br>**ISO/IEC 27001:2013** A.12.4.1, A.12.4.2, A.12.4.3, A.12.4.4, A.12.7.1 |

| Solution Characteristic | NIST CSF Category | Informative References |
|---|---|---|
| Configuration management (install, monitor, recover) | PR.DS-1<br>PR.IP-3<br>PR.IP-9<br>PR.PT-1<br>DE.AE-4 | **NIST SP 800-53 Rev. 4** AU Family, CA-7, CM-3, CM-4, CP-2, IR-4, IR-5, IR-8, SA-10, SC-28, SI-4<br>**ISO/IEC 27001:2013** A.8.2.3, A.12.1.2, A.12.4.1, A.12.4.2, A.12.4.3, A.12.4.4, A.12.5.1, A.12.6.2, A.12.7.1, A.14.2.2, A.14.2.3, A.14.2.4, A.16.1.1, A.17.1.1, A.17.1.2 |

Table 1: Solution to security category map

The list of characteristics and corresponding capabilities is not exhaustive. Furthermore, capabilities are listed to provide context for the characteristics and are not meant to be prescriptive.

## 9. High-Level Architecture

The figure below depicts the proposed high-level environment and architecture to help ensure data integrity within the enterprise.

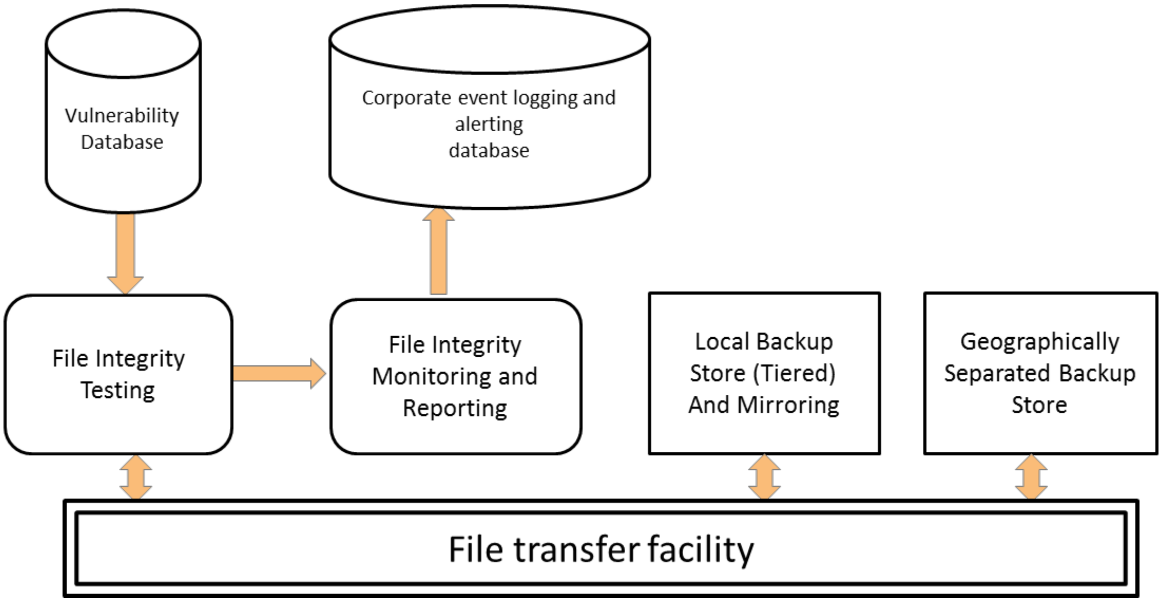
Vulnerability Database
Corporate event logging and alerting database
File Integrity Testing
File Integrity Monitoring and Reporting
Local Backup Store (Tiered) And Mirroring
Geographically Separated Backup Store
File transfer facility

Figure 1. Data Integrity Building Block high-level architecture

# 10. Component List

Data integrity solutions include but are not limited to the following components:

- File integrity monitors
- File versioning
- File integrity testing
- User activity monitoring
- Configuration management
- Database rollbacks
- Virtual machine integrity/snapshots/versioning
- Versioning file systems
- Journaling file systems

Some of these are subcomponents of the components shown in the architecture in section 9.

# Appendix A – References

[1] NIST Framework for Improving Critical Infrastructure Cybersecurity, Version 1, February 12, 2014 http://www.nist.gov/cyberframework/upload/cybersecurity-framework-021214.pdf

[2] "Joint Statement: Destructive Malware." Ffiec.gov. March 30, 2015. Accessed July 10, 2015. https://www.ffiec.gov/press/PDF/2121759_FINAL_FFIEC Malware.pdf.

[3] Karen Scarfone, Murugiah Souppaya, Paul Hoffman. "NIST Special Publication 800-125 Guide to Security for Full Virtualization Technologies" NIST.gov January 2011 http://csrc.nist.gov/publications/nistpubs/800-125/SP800-125-final.pdf

[4] Executive Order no. 13636, *Improving Critical Infrastructure Cybersecurity*, DCPD-201300091, February 12, 2013. http://www.gpo.gov/fdsys/pkg/FR-2013-02-19/pdf/2013-03915.pdf

[5] Financial Sector Information Sharing and Analysis Center Best Practices for U.S. Financial Institutions, Reducing Risks Associated with Destructive Malware, November 2, 2015

[6] Souppaya, Murugiah, and Karen Scarfone. "NIST Special Publication 800-83 Rev 1 Guide to Malware Incident Prevention and Handling for Desktops and Laptops." NIST.gov. July 1, 2013. Accessed August 29, 2015. http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-83r1.pdf

[7] "Spotting the Adversary with Event Log Monitoring". NSA.gov. December 16, 2013. Accessed August 17,2015. https://www.nsa.gov/ia/_files/app/spotting_the_adversary_with_windows_event_log_monitoring.pdf

[8] "Eight Best Practices for Disaster Recovery." CIO.com. November 18, 2004. Accessed August 29, 2015. http://www.computerworld.com/article/2568383/disaster-recovery/eight-best-practices-for-disaster-recovery.html

[9] Ron Ross, Janet Carrier Oren, Michael McEvilley "NIST Special Publication 800-160 draft Systems Security Engineering: An Integrated Approach to Building Trustworthy Resilient Systems" NIST.gov May 2014 http://csrc.nist.gov/publications/drafts/800-160/sp800_160_draft.pdf

[10] Arnold Johnson, Kelley Dempsey, Ron Ross, Sarbari Gupta, Dennis Bailey "NIST Special Publication 800-128  Guide for Security-Focused Configuration Management of Information Systems" NIST.gov August 2011 http://csrc.nist.gov/publications/nistpubs/800-128/sp800-128.pdf