School of Computing Science and Engineering

Program:B.Tech(BA+BD)

Course Code:CSBD4070

Course Name:Big Data Security

Course Outcomes :

CO NUMBER	TITLE
CO1	Understand and apply the models of Information Security
CO2	Apply and critique strategies for personal privacy protection and Understand and build security frameworks for big data
CO3	Build security in Hadoop environment
CO4	To use big-data analytics principles to build security applications
CO5	To detect security threats and vulnerabilities using security analytic
CO6	Ability to independently carry out research / Investigations, identify problems and develop solutions to solve practical problems in **Big Data Security**. (Create)

Course Prerequisites

- **Knowledge on Big Data**

Syllabus

Unit I: Introduction to Information System Security

Information System Security: Critical characteristics of Information - NSTISSC Security Model- Components of information System SDLC Information assurance - Security Threats and vulnerabilities - Overview of Security threats- Security Standards

UNIT II: Privacy and Security of Big Data

Privacy in Big Data: Privacy need for Data Sharing Anonymization design principles Data Anonymization in multidimensional data- Data Anonymization in time series data Threats to anonymized data- Privacy preserving data mining Dynamic data Protection - Security, Compliance, Auditing and Protecting: Steps to secure big data Classifying Data Protecting Big Data Compliance

Intellectual Property Rights and challenges

UNIT III: Security Design

Security Design: Kerberos Default Hadoop Model without security - Hadoop Kerberos Security- Open source authentication in Hadoop-Log monitoring Encryption for Hadoop.

UNIT IV: INTRODUCTION TO SECURITY ANALYTICS

Introduction to Security Analytics – Techniques in Analytics – Analysis in everyday life – Challenges in Intrusion and Incident Identification – Simulation and Security Process, Analytical Softwares and tools, Malware Analysis – static and dynamic analysis - Security Intelligence –Security Breaches.

UNIT V: APPLICATIONS OF SECURITY ANALYTICS

Access Analytics – Analysis of Log file -Security analysis with text mining –Machine Learning and data mining applications for security: Intrusion detection and network anomaly detection. Big data analytics for security: Analyzing DDOS – Distributed Denial of Service attack: counter based method, and access pattern based method – Machine learning for Ransom ware detection and prevention.

Unit VI: Advances and the Latest Trends

The advances and the latest trends in the course as well as the latest applications of the areas covered in the course.

The latest research conducted in the areas covered in the course.

Discussion of some latest papers published in IEEE transactions and ACM transactions, Web of Science and SCOPUS indexed journals as well as high impact factor conferences as well as symposiums.

Discussion on some of the latest products available in the market based on the areas covered in the course and patents filed in the areas covered in the course

Text Books

1. Michael E. Whitman, Herbert J Mattord, Principles of Information Security, Sixth edition, Vikas Publishing House, 2017
2. Nataraj Venkataramanan, Ashwin Shriram, Data Privacy: Principles and Practice, First edition, Chapman and Hall/CRC, 2016
3. Mark Talabis, Robert McPherson, I Miyamoto and Jason Martin, “Information Security Analytics: Finding Security Insights, Patterns, and Anomalies in Big Data”, Syngress Media,U.S., 2014

Reference Books

1. Ben Spivey, Joey Echeverria, Hadoop Security Protecting Your Big Data Problem, OReilly Media, 2015
2. Mark Van Rijmenam, Think Bigger: Developing a Successful Big Data Strategy for Your Business, First edition, Amazon, 2014
3. Behrouz A. Forouzan, “Cryptography and Network Security”, Tata McGraw Hill Education, 2nd Edition, 2010.
4. Douglas R. Stinson ,“Cryptography Theory and Practice ”, Chapman & Hall/CRC, 3rd Edition, 2006.
5. William Stallings, “Crpyptography and Network security: Principles and Practices”, Pearson/PHI, 5th Edition, 2010.

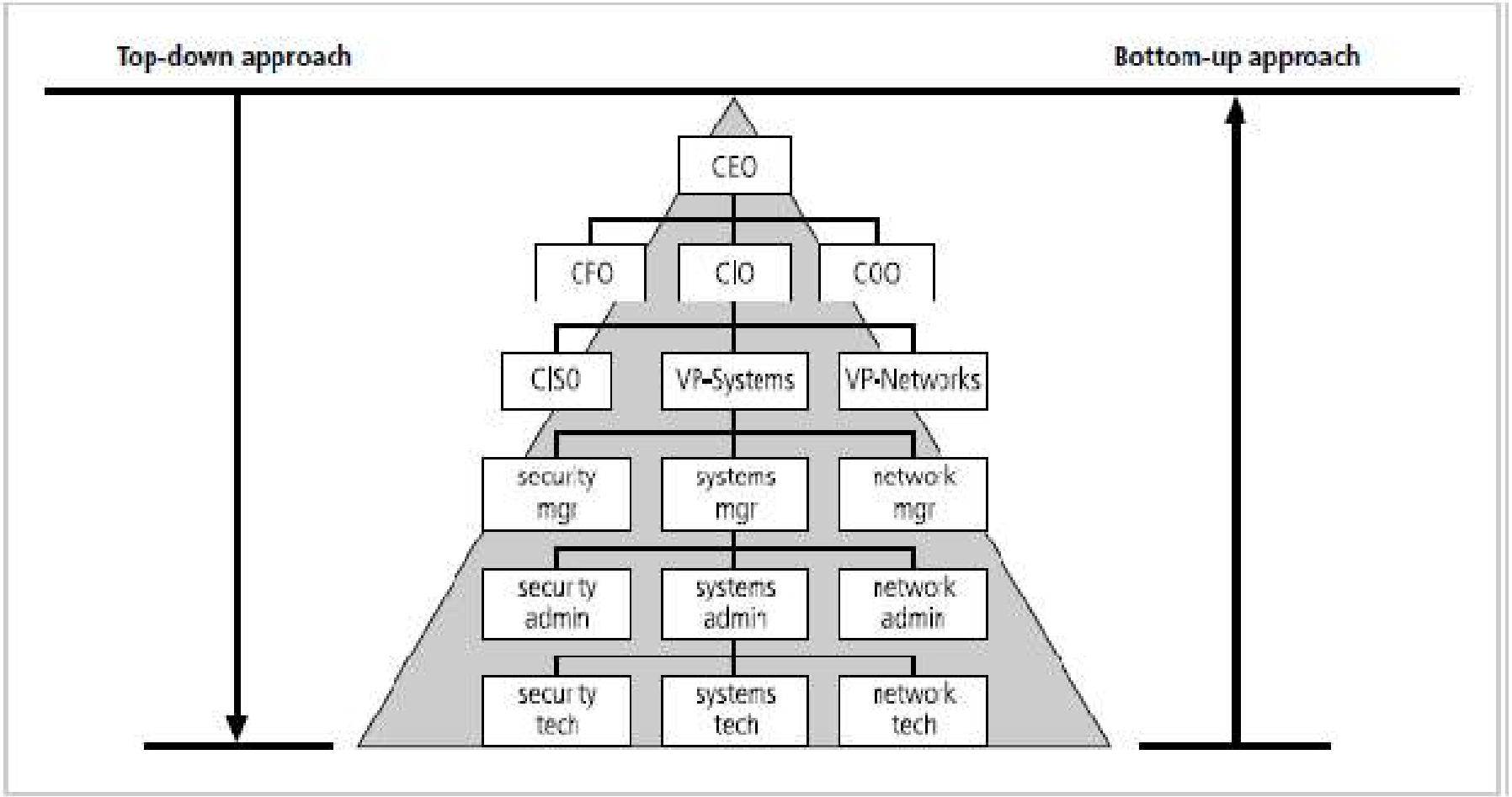
Approaches to Information Security Implementation

- Bottom up Approach
  - Information security can begin as a grassroots effort in which systems administrators attempt to improve the security of their systems
  - key advantage of the bottom-up approach is the technical expertise of the individual administrators
  - Working with information systems on a day-to-day basis, these administrators possess in-depth knowledge that can greatly enhance the development of an information security system.
  - They know and understand the threats to their systems and the mechanisms needed to protect them successfully.
  - Unfortunately, this approach seldom works, as it lacks a number of critical features, such as participant support and organizational staying power.

Approaches to Information Security Implementation

- Top-Down Approach
  - in which the project is initiated by upper-level managers who issue policy, procedures and processes, dictate the goals and expected outcomes, and determine accountability for each required action—has a higher probability of success.
  - This approach has strong upper-management support, a dedicated champion, usually dedicated funding, a clear planning and implementation process, and the means of influencing organizational culture.
  - The most successful kind of top-down approach also involves a formal development strategy referred to as a systems development life cycle

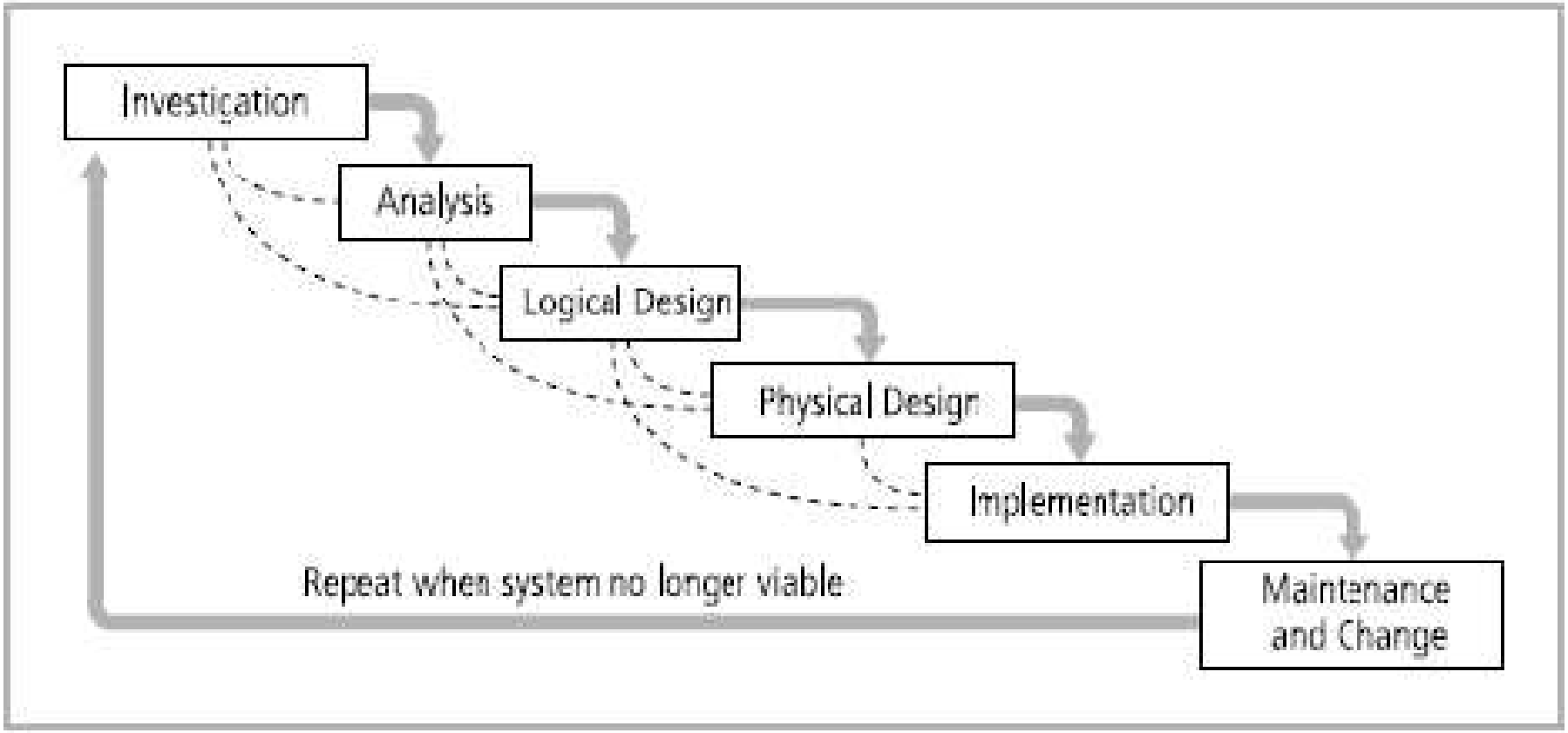
Approaches to Information Security Implementation



The Systems Development Life Cycle

- The **systems development life cycle (SDLC)** is a methodology for the design and implementation of an information system.
- A **methodology** is a formal approach to solving a problem by means of a structured sequence of procedures.

The Systems Development Life Cycle



Security Considerations in the Information SDLC

- Each of the phases of the SDLC should include consideration of the security of the system being assembled as well as the information it uses.
- Whether the system is custom and built from scratch, is purchased and then customized, or is commercial off-the-shelf software (COTS), the implementing organization is responsible for ensuring it is used securely.
- This means that each implementation of a system is secure and does not risk compromising the confidentiality, integrity, and availability of the organization's information assets.

Security Considerations in the Information SDLC

- Each of the example SDLC phases [discussed earlier] includes a minimum set of security steps needed to effectively incorporate security into a system during its development.
- An organization will either use the general SDLC described [earlier] or will have developed a tailored SDLC that meets their specific needs

Security Considerations in the Information SDLC

- **Investigation/Analysis Phases**
  - **Security categorization**—defines three levels (i.e., low, moderate, or high) of potential impact on organizations or individuals should there be a breach of security (a loss of confidentiality, integrity, or availability). Security categorization standards assist organizations in making the appropriate selection of security controls for their information systems.
  - **Preliminary risk assessment**—results in an initial description of the basic security needs of the system. A preliminary risk assessment should define the threat environment in which the system will operate.

Security Considerations in the Information SDLC

- **Logical/Physical Design Phases**
  - **Risk assessment—**analysis that identifies the protection requirements for the system through a formal risk assessment process.
  - **Security functional requirements analysis**—analysis of requirements that may include the following components: (1) system security environment (i.e., enterprise information security policy and enterprise security architecture) and (2) security functional requirements
  - **Security assurance requirements analysis**—analysis of requirements that address the developmental activities required and assurance evidence needed to produce the desired level of confidence that the information security will work correctly and effectively.
  - **Cost considerations and reporting**—determines how much of the development cost can be attributed to information security over the life cycle of the system. These costs include hardware, software, personnel, and training.
  - **Security planning** —ensures that agreed upon security controls, planned or in place, are fully documented. The security plan also provides a complete characterization or description of the information system as well as attachments or references to key documents supporting the agency's information security program
  - **Security control development** —ensures that security controls described in the respective security plans are designed, developed, and implemented.
  - **Developmental security test and evaluation** —ensures that security controls developed for a new information system are working properly and are effective.

Security Considerations in the Information SDLC

- **Implementation Phase**
  - **Inspection and acceptance** —ensures that the organization validates and verifies that the functionality described in the specification is included in the deliverables.
  - **System integration** —ensures that the system is integrated at the operational site where the information system is to be deployed for operation. Security control settings and switches are enabled in accordance with vendor instructions and available security implementation guidance.
  - **Security certification** —ensures that the controls are effectively implemented through established verification techniques and procedures and gives organization officials confidence that the appropriate safeguards and countermeasures are in place to protect the organization's information system. Security certification also uncovers and describes the known vulnerabilities in the information system.
  - **Security accreditation** —provides the necessary security authorization of an information system to process, store, or transmit information that is required. This authorization is granted by a senior organization official and is based on the verified effectiveness of security controls to some agreed upon level of assurance and an identified residual risk to agency assets or operations.

Security Considerations in the Information SDLC

- **Maintenance and Change Phase**
  - **Configuration management and control** —ensures adequate consideration of the potential security impacts due to specific changes to an information system or its surrounding environment.
  - **Continuous monitoring** —ensures that controls continue to be effective in their application through periodic testing and evaluation.
  - **Information preservation** —ensures that information is retained, as necessary, to conform to current legal requirements and to accommodate future technology changes that may render the retrieval method obsolete.
  - **Media sanitization** —ensures that data is deleted, erased, and written over as necessary.
  - **Hardware and software disposal** —ensures that hardware and software is disposed of as directed by the information system security officer.