

DBM041: Implement Apache Hadoop for Big Data

หลักการและเหตุผล:

This course provides the knowledge and skills to use the tools and functions needed to work within Apache Hadoop and designed for the beginner on this product.

วัตถุประสงค์:

- Overview of Big Data
- Describes about ecosystem of Apache Hadoop
- How to install and configure Apache Hadoop
- Working with HDFS
- Describes concept of MapReduce
- Working with Hive, Pig and Impala
- Using Data transfer with Sqoop and Flume

หลักสูตรนี้เหมาะสำหรับ:

The primary audience for this course is individuals who implement Apache Hadoop in organization for Big Data technology and prepare next step to Data Analytics.

ความรู้พื้นฐาน:

- Working knowledge of SQL
- Working knowledge of relational databases
- Some experience with database design

เนื้อหาการอบรม:

Module 1: Introduction Apache Hadoop

- What is Big Data
- Big Data in The Modern World
- How to implement Big Data
- Compare Database Technology
- Hadoop Overview
- Concept for implement Hadoop
- Compare Hadoop Distributions
- About Data Lake

Module 2: Hadoop Basics

- HDFS and MapReduce
- Hadoop Run Modes and Job Types
- Hadoop Software Requirement and Recommendations
- Hadoop in the Cloud
- Hadoop Installation
- About Hadoop Cluster Management
- About HUE

Module 3: Hadoop Distributed File System (HDFS)

- What's HDFS?
- HDFS Core Concept
- HDFS Daemons (Name Node\ Data Node)
- Name Node's Meta Data
- Function of Secondary Name Node
- HDFS HA Architecture
- HDFS File Write\Read Walkthrough
- Manage HDFS with Shell Command
- Manage HDFS with HUE

Career for the Future Academy: CFA

Module 4: MapReduce

- MapReduce Explained
- Map Concept
- Reduce Concept
- MapReduce Job Walkthrough

Module 5: Hive, Pig, And Impala

- Comparing Hive, Pig and Impala
- Hive Overview
- Using HiveQL
- Load Data to Hive's Table
- Internal\ External Hive's table
- Query data with HiveQL
- Pig Overview
- Using Pig Latin
- Databags, Tuples and Atoms
- Constructs on Pig
- Pig with HCatalog
- Impala Overview
- Using Impala-shell
- Load Data to Impala's Table
- Query data with Impala's command

Module 6: Data Import and Export

- About Data Analysis Workflow
- Sqoop overview
- Sqoop use case
- How Sqoop works
- Using Sqoop's import
- Sqoop Parallel import concept
- Sqoop's option File
- Using Sqoop's Hive Import
- Using Sqoop's Incremental Import
- Using Sqoop's Export
- Intro Oozie Workflow for import\export
- Flume Overview
- How Flume works
- Flume's Agent and Collections
- Using Flume's for real-time data

วิทยากร:



อาจารย์สรรัตน์ เกษมบุญศิริ

- วิทยากรรับเชิญประจำสถาบันพัฒนาบุคลากรแห่งอนาคต
- Managing Director (Born2Learn Co., Ltd)
- Microsoft Certified Trainer (MCT)
- Microsoft Certified System Engineer (MCSE)
- Microsoft Certified Solution Associate -Database
- Microsoft Certified Professional (MCP)
- Certified ITIL Foundation v.3
- Certified CompTIA Security +
- Cisco Certified Network Associate (CCNA)
- Certified AOTS of Japan Members

จำนวนชั่วโมงในการฝึกอบรม: 3 วัน (18 ชั่วโมง)

ช่วงเวลาฝึกอบรม: 9.00 - 16.00 น.

Career for the Future Academy: CFA

กำหนดการอบรม: ตามตารางปฏิทินอบรมประจำปี <https://www.career4future.com/trainingprogram>

ค่าลงทะเบียนอบรม: ท่านละ 11,500 บาท (ราคารวมภาษีมูลค่าเพิ่มแล้ว)

** สถาบันฯ เป็นหน่วยงานราชการ จึงไม่อยู่ในเกณฑ์ที่ต้องถูกหักภาษี ณ ที่จ่าย

สถานที่ฝึกอบรม:

สถาบันพัฒนาบุคลากรแห่งอนาคต

เลขที่ 73/1 อาคารสำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ (สวทช.) ชั้น 6

ถนนพระรามที่ 6 แขวงทุ่งพญาไท เขตราชเทวี กรุงเทพฯ 10400

วิธีการสำรองที่นั่ง:

ติดต่อสำรองที่นั่งล่วงหน้า ในวัน-เวลาราชการ

โทรศัพท์: 0 2644 8150 ต่อ 81886, 81887

โทรสาร: 0 2644 8110

Website: www.career4future.com

E-mail: training@nstda.or.th