SINGLE & UNIFIED COURSE ...

FULLSTACK BIG DATA , CLOUD & DEVOPS ENGINEER

APACHE Spark™

inceptez
INTELLIGENCE DIGITALIZED

kafka

aws

hadoop

ACCELERATE YOUR CAREER GEAR TO EXPERIENCE, EXPLORE AND EXCEL THE CUTTING EDGE TECHNOLOGIES OF BIG DATA, CLOUD AND DEV OPS WITH REAL TIME IMPLEMENTATION .....

Just a Click Ahead to Know More About Us

Why You have To Learn from Inceptez	• http://Inceptez.in -> more -> Why I have to learn in Inceptez • https://www.shorturl.at/rsL26
FAQ about the Course	• http://Inceptez.in -> more -> Frequently Asked Questions • https://www.shorturl.at/eflqA
Our Other Value Added Services	• http://Inceptez.in -> more -> Inceptez Interview & Job Support • https://www.shorturl.at/inpKW

ABOUT INCEPTEZ

Inceptez Technologies was founded by a team of Big data Evangelists in 2014 and is one of the leading IT training, Development and staffing company specializing in Big Data, Data Science, Dev-Ops, Cloud Computing and Internet of things (IOT). Inceptez is a non money oriented training center, where we first prioritize Comprehensiveness, Engagement based, Focus based, Competitive model with high Quantity and Quality in all the training as a paramount.

Inceptez Technologies is mastered and administrated by highly skilled industry experts. We are the technology enablers committed to provide comprehensive training to the aspiring professionals in the game changing, high demanding applications such as Hadoop, Spark , Data Analytics, Data Science, DEV-OPS and AWS/Google Cloud Platforms, that are the fastest growing trend setting technologies that provide competitive advantage in the ever changing IT world.

ENQUIRE, ANALYSE, COMPARE AND FINALLY INVEST IN YOUR CAREER ACCEPTING INCEPTEZ AS YOUR CAREER PARTNER

WE CHISEL YOU TO MAKE YOU FEEL TECHNICALLY VIBRATED, MOTIVATED & OVERWHELMED

TRAINING

We are the technology leaders committed to provide comprehensive training to the aspiring professionals in the game

DEVELOPMENT

Our software engineering process collects and translates business requirements into imaginative technology solutions that become reality with custom software development.

SOLUTION

We are the technical leaders expertise in providing end to end solutions for the cutting edge technologies which Industries demands for their business growth and analytics requirements.

STAFFING

Inceptez Technologies provides flexible, innovative recruitment strategies and technologies to maximize recruiting efficiency and reduce cost.

BIG DATA **“Continuous & On Demand”**, it’s the “**DATA ENABLEMENT Strategies”** and with its the ability to *handle extremely large data sets that may be analysed computationally to reveal patterns, trends, and associations, especially relating to human behaviour and interactions*.

It is a well-proven fact that effective implementation of BigData data enablement strategies will place the implying organization way ahead of their respective competitors. So the demand for a qualified BigData - Data engineers is more in the present corporate sector. But in contrast to this demand, there is a shortage of skilled Data engineers. And so, many top notching organizations are willing to pay extremely high pay packages for the best-skilled BigData experts.

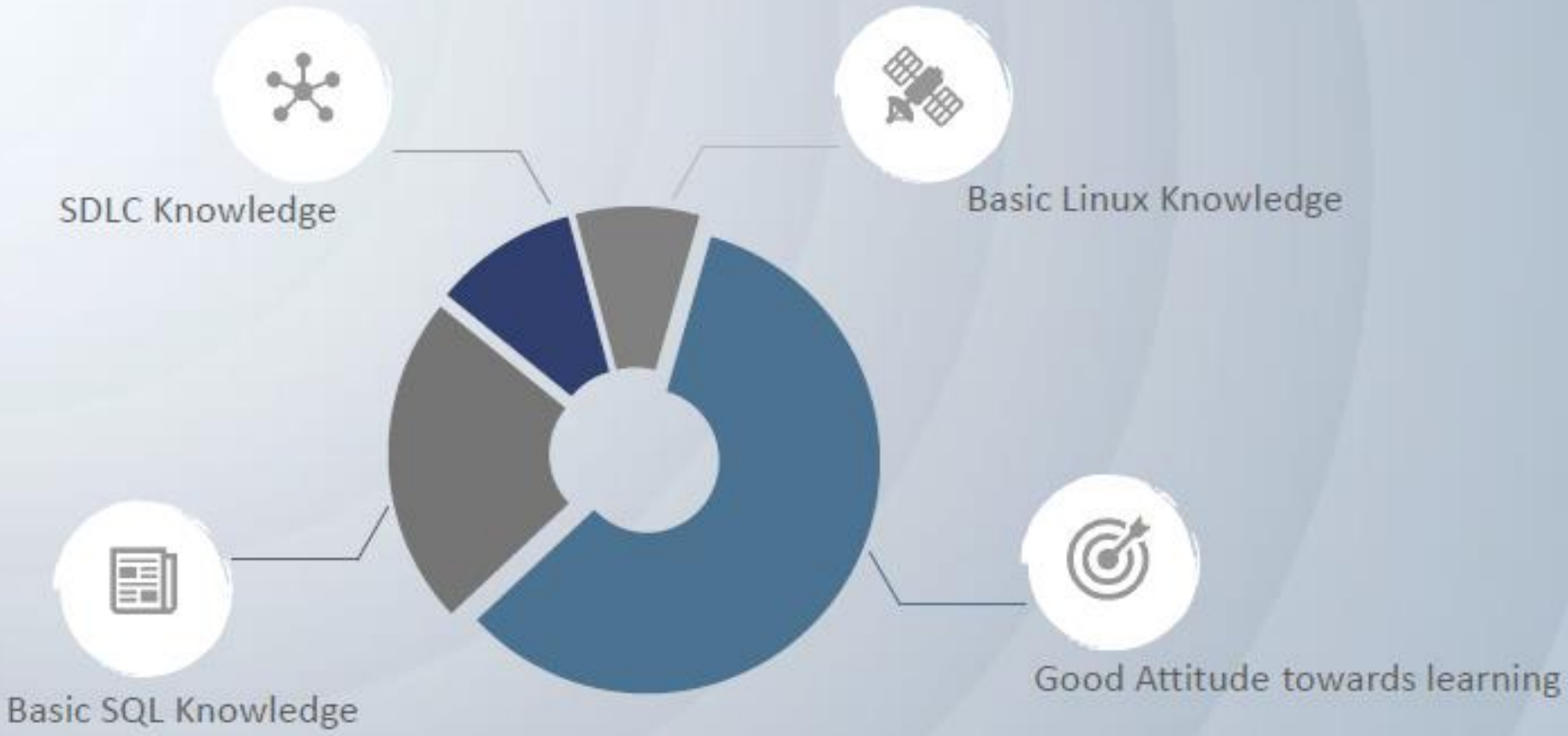
About BIG DATA

Methodology to Aquire, Clean, Process, Store, Analyze Huge Volume of Hetrogenous Data that Flows in Different Frequencies…

inceptez

Prerequisites

Good to know, anyway Don't worry we will take care of bringing IT



All Tools & Technology Stack we use in this Course

AWS ATHENA

Key Components …

HADOOP

SPARK

Cloud Services

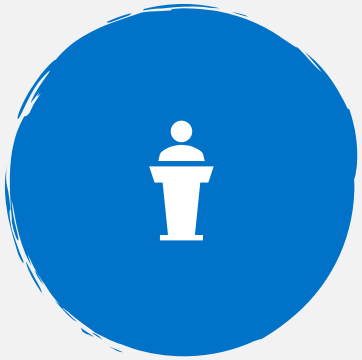
DevOps & RPA

NOSQL

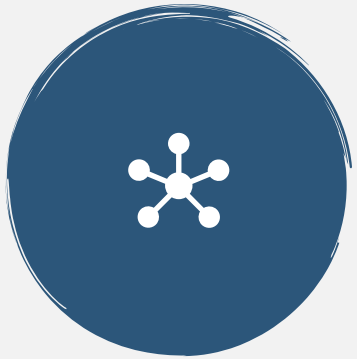
Google Cloud Platform

Focussed Teaching

Best in class Training



An extensive, comprehensive, use case and hands on oriented, Advanced topic covered course designed with distributed, In-memory, interactive, realtime and stream processing using end to end Data Engineering, Data analysis, Visualization and Dashboard **powered by Hadoop, Spark, Cloud and DevOps tools**…



Engagement based

Designed as per the Market need

What's Unique

All under ONE course Hadoop, Spark , Cloud, ELK & Cloud

Job Oriented Training, Professional Environment

All Trainings by Industry Experts, Completely Hands-on Driven

Streams of Batch, In-Memory, Real-time, Fast Data, Data lake Building

Extended Training Duration to cover topics Wide and Depth

Use cases, Case Studies, Performance Tuning, Best Practices, Benchmarking

Interview and Job Support with seasoning of Resume with Bigdata projects

End to end simplified & comprehensive learning materials

Addresses 360 degree requirements for Overwhelming Training experience

Competitive learning, Active, Comparative, Declarative learning

6 Complete Realtime Projects for gathering Realtime experience

2 Hackathons & Test with Certification Guidance

Packaging & Production Deployment Strategies

Enriched Cloud Components such as EMR, EC2, S3, IAM, ATHENA

Cloud based Hadoop/Spark Clustering

inceptez

Key Learning

- All About **BigData & Hadoop** Deep Dive
- **Linux, SQL, ETL & Datawarehouse** Refresh
- Hadoop **HDFS , Map Reduce, YARN** Distributed Framework
- **NOSQL** – for realtime data storage and search using **HBASE & ELASTIC SEARCH**
- Visualization & Dashboard – **Kibana with Elastic search** integration using Spark
- **Robotic Process Automation (RPA)** using Linux & Spark
- Inmemory stream for **Fast data, Realtime streaming & Data lake** formation using **Spark, Kafka, Nifi.**
- **Reusable Framework** creation with Logging framework.
- **Cluster formation** in Cloud environments
- **SDLC , Packaging & Deployment** in Bigdata platform
- **Six Project execution with Hackathon & Test**
- **Job submission & Orchestration** with Scheduling using Oozie

inceptez

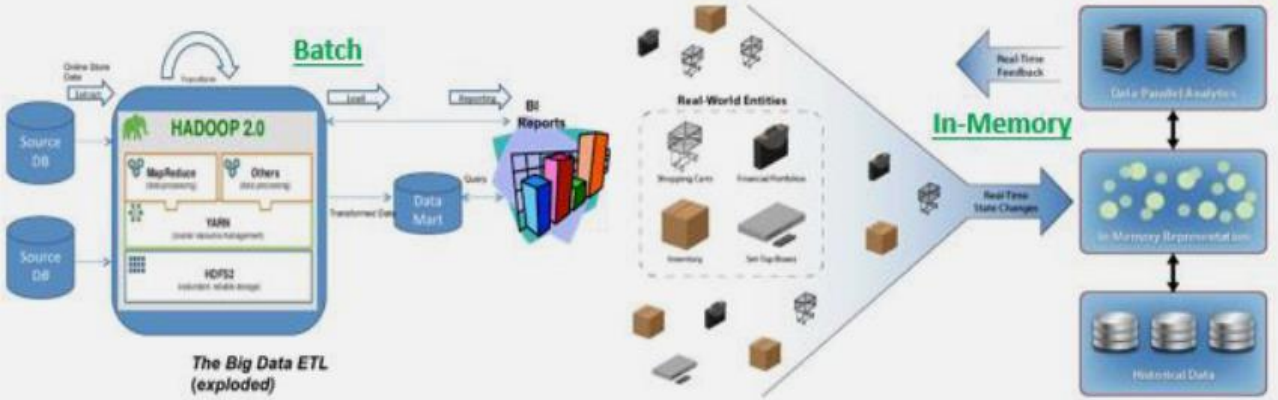
High Level Eco Systems Overview

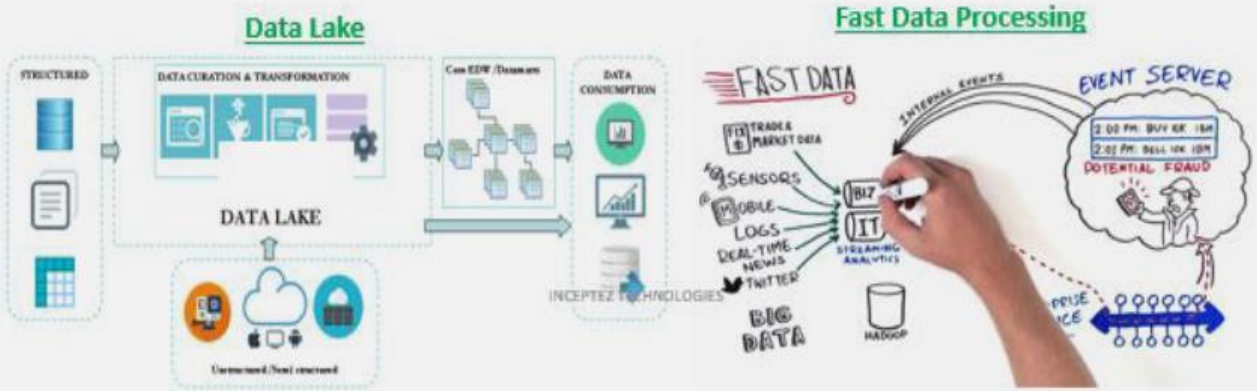
- All About BigData & Hadoop Deep Dive
- Linux, SQL, ETL & Datawarehouse Refresh
- **Hadoop** HDFS , Map Reduce, YARN Distributed Framework
- **SQOOP** – Data Ingestion Framework
- **HIVE** – SQL & OLAP Layer on Hadoop
- **HBASE & ELASTIC SEARCH** - RealTime Random Read/Write NOSQL
- **PHOENIX** – SQL Layer on top of HBASE
- **KIBANA** – Realtime Visualization on top of Elastic Search
- **OOZIE** – Workflow Scheduling & Monitoring tool
- **NIFI** – Data Flow Tool for Mediation & Routing of large dataset
- **KAFKA** – Distributed & Scalable Messaging Queue
- **SPARK** – Fast & Distributed In-Memory engine for largescale data
- **SCALA/PYTHON –** Scalable, Function based Highlevel Language
- **HUE** – GUI for Hadoop Eco System
- **AMBARI -** Provisioning, Managing and Monitoring Hadoop Clusters
- **Google Cloud** based - Hadoop & Spark Cluster setup
- **HORTONWORKS** - Distribution for provisioning Hadoop Cluster
- **AWS** Services - EMR, EC2, S3, IAM, SG, ATHENA
- MAVEN & GITHUB – DevOps Continuous Build & Version control
- Frameworks for Data Masking, Data Validation & Sanitation

inceptez

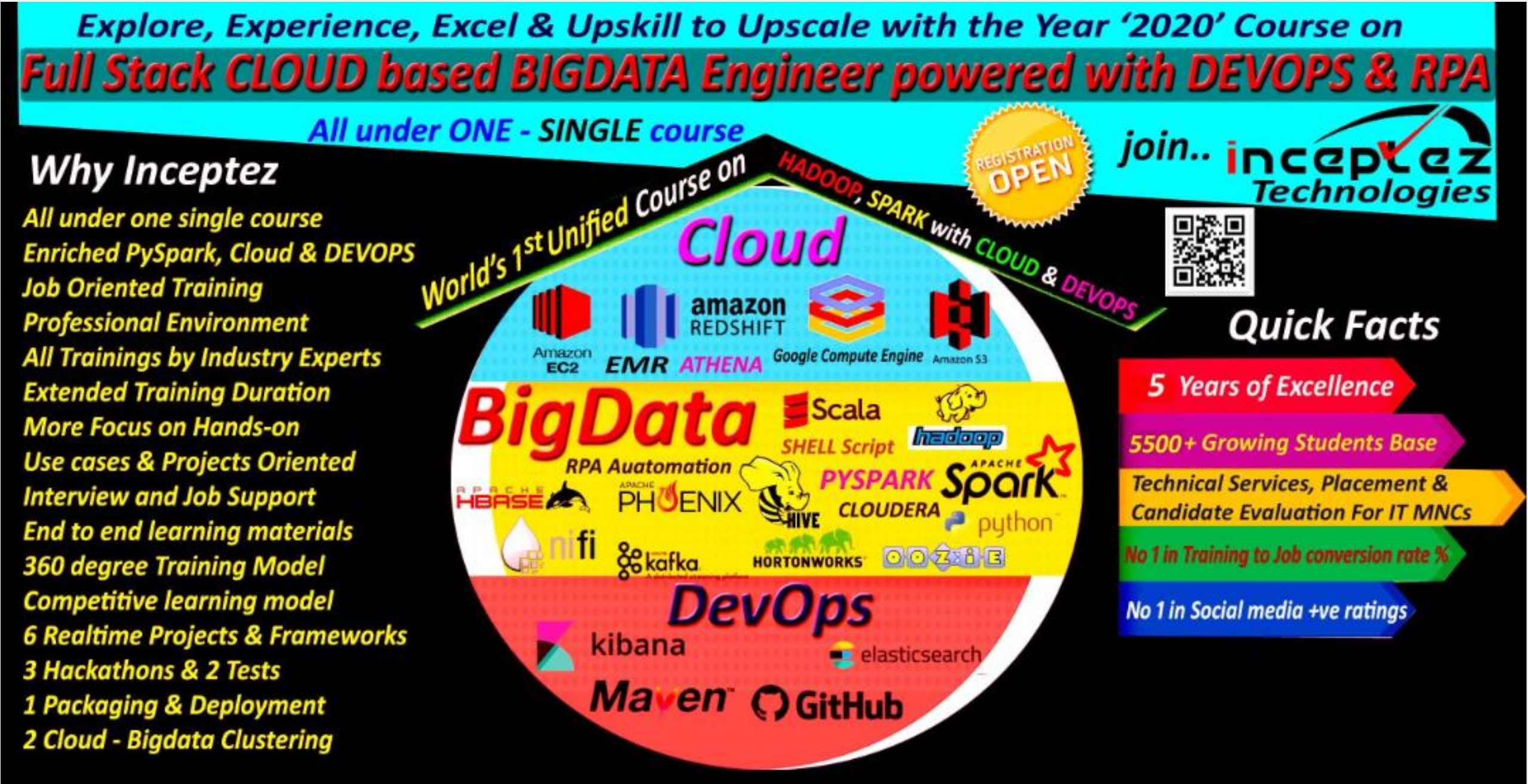
Domains Addressed

Data Engineering Techniques Covered





Course Banner



Complementary Services – We will help you to be Smart, Unique & Attractive enough to Compete...

Self Explanation:

- ✓ About You – Explain About Yourself
- ✓ Resume Building
- ✓ Project Explanation

Common Interview Questions:

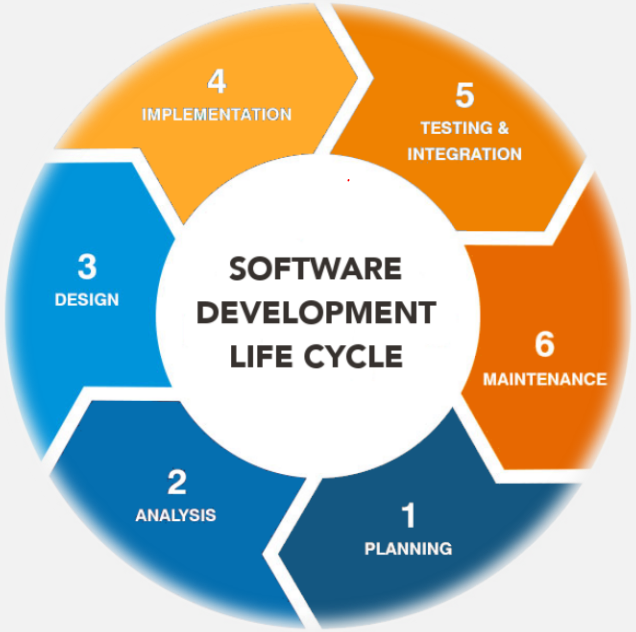
- ✓ Day to Day activities Roles
- ✓ Cluster Sizing, Versions
- ✓ Challenges you faced in your project
- ✓ Logging & UI
- ✓ Package and Deployment of Code
- ✓ Performance Tuning

Frameworks & Advanced Feature:

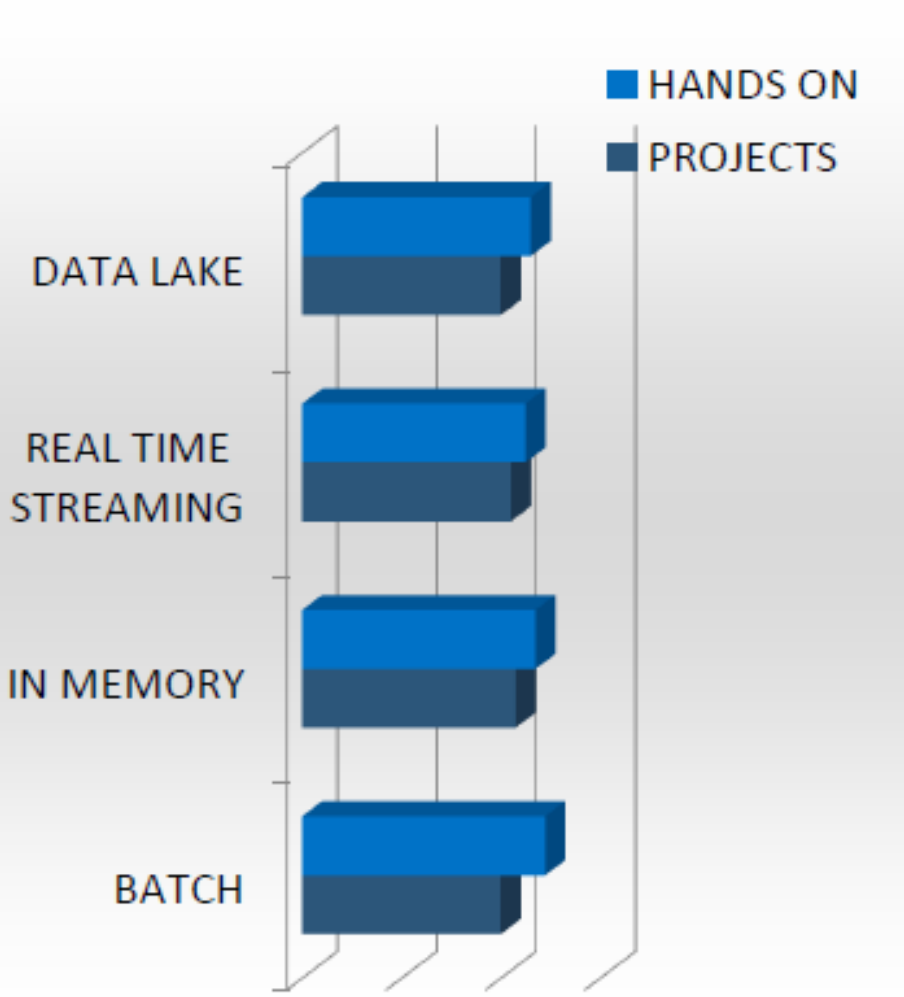
- ✓ Automation & Framework
- ✓ Data governance & Security
- ✓ Data lake Architectures

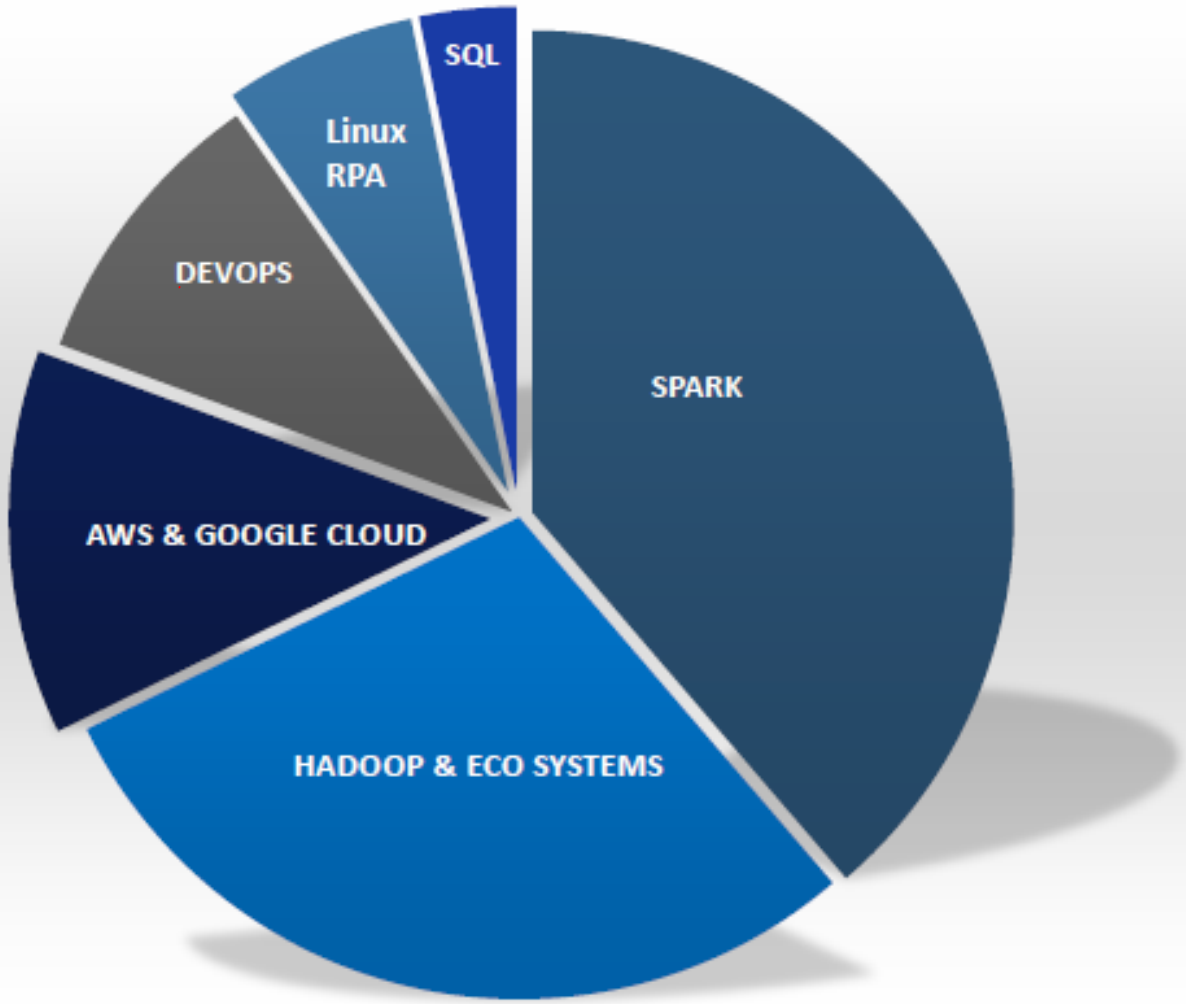
Good To Have or Additional Knowledge:

- ✓ Interview Discussion & Understand the Priority of tools.
- ✓ Basic & Advanced SQL
- ✓ Linux Shell Scripting (RPA)



Proportion of Contents





Course Details

Overview of BIGDATA

We have to first have know all about Big-Data & its Characteristics.

- Evolution of Data
- Introduction
- Classification
- Size Hierarchy
- Why Big data is Trending
- IOT, Devops, Cloud Computing, Enterprise Mobility
- Challenges in Big Data
- Characteristics
- Tools for Big Data
- Why Big Data draws attention in IT Industry
- What do we do with Big data
- How Big Data can be analyzed
- Typical Distributed System
- Draw backs in Traditional distributed Systems
- BigData tools

Linux Foundation

In this module you will be learning Introduction & Key Components of Linux Dev & Admin

- History and Evolution
- Architecture
- Development Commands
- Env Variables
- File Management
- Directories Management
- Admin Commands
- Advance Commands
- Shell Scripting
- Groups and User managements
- Permissions
- Important directory structure
- Disk utilities
- Compression Techniques
- Misc Commands
- Kernel, Shell
- Terminal , SSH, GUI
- Hands On Exercises

Linux Scripting

In this module you will be linux shell scripting and automation techniques

- Automation process using shell scripting
- Integration of hadoop eco systems with Linux scripting
- Looping, conditional, vars, methods
- Key Differences between Linux & Windows
- Kernel
  - What is the Purpose of Kernel?
  - How Kernel Works?
  - Find Kernel
- Shell
  - What is the Purpose of Shell?
  - Types of Shell
  - Environment Variables in Shell
  - Hands On Exercises

Course Details

Hadoop In Depth

In this module you will be learning all about Hadoop

- What is Hadoop?
- Evolution of Hadoop
- Features of Hadoop
- Characteristic of Hadoop
- Hadoop compared with Traditional Dist. Systems
- When to use Hadoop
- When not to use Hadoop
- Components of Hadoop (HDFS, MapReduce, YARN)
- Hadoop Architecture
- Daemons in Hadoop Version 1 & 2
- How Data is stored in Hadoop Cluster, Datacenter, Spilt, Block, Rack Awareness, Replication, Heart beat)
- Hadoop 1.0 Limitation
- NameNode High Availability

Hadoop HDFS

Hadoop distributed file system concepts with architecture, commands, options, advance options, data management

- Namenode Federation
- Hadoop version s
- Anatomy of File Read/Write
- Hadoop Cluster Formation in VM, Sandbox & GCP Cloud
- Cluster formation & sizing guide
- Hadoop Commands Handson
- Hadoop admin handson
- HDFS integration with Linux shell
- HDFS additional Usecases
- Data Integrity
- Serialization
- Compression techniques
- Data ingestion to HDFS using different ecosystems

SQOOP Data Acquisition

Data ingestion or data acquisition tool for transporting bulk data between RDBMS -> Hadoop & Vice versa

- Sqoop Introduction & History
- Technical & Business benifits
- Installation and configuration
- Why Sqoop
- Indepth Architecture
- Import & Export Properties
- Sqoop Export Architecture
- Commands (Import HDSF, HIVE, HBase from MySQL, ORACLE)
- Export Command Options
- Incremental Import
- Saved Jobs , Sqoop Merge
- Import All tables , Excludes
- Best practices & performance tuning
- Sqoop import/export use cases
- Advance sqoop commands
- Sqoop realtime usecases
- Sqoop Hive Hbase Integration

Course Details

Hive

SQL Layer on top of Hadoop for analytical and declarative queries

- Introduction to Hive
- Architecture
- Hive Vs RDBMS Vs NOSQL
- Detailed Installation (Metastore, Integrating with Hue)
- Starting Metastore and Hive Server2
- Data types (Primitive, Collection Array, Struct, Map)
- Create Tables (Managed, External, Temp)
- DML operations (load, insert, export)
- Exploring Indexes
- HQL Automation using shell scripts
- Managed Vs External tables
- HQL Queries using end to end usecases
- Hive analytical and hierarchial queries

Hive Components

Hive Components such as partition, bucketing, views, indexes, joins, handlers, udfs etc

- Hive access through Hive Client, Beeline and Hue
- File Formats (RC, ORC, Sequence)
- Partitioning (static and dynamic)
- partition with external table
- Alter, Drop, Repair Partitions
- Hive Sqoop, HBase Integration
- Hive Storage Handler implementation
- Bucketing, Partitioning Vs Bucketing
- Views, different types of joins
- Aggregation, normalization Queries
- Add files to the distributed cache, jars to the class path
- UDF using Python & Scala
- Generic UDF, UDAF

Advance Hive

Usecases & POCs on serdes, file formats, schema evolution, SCD concepts etc.,

- Optimized joins (MapSide join,SMB, Bucketing join)
- Compressions on tables (LZO, Snappy)
- Serde (XML Serde, JsonSerde, CSV, Avro, Regex)
- Parallel execution
- Sampling data
- Speculative execution
- Installation & Configuration
- Two POCs using the large dataset on the above topics
- Hive Slowly changing dimension implementation
- Hive Schema evolution use case using Avro dataset
- Hive Usecase with retail and banking dataset

Course Details

Map Reduce Framework

Hadoop Processing framework for Distributed processing with multitasking capabilities

- Introduction to MapReduce
- Hadoop Ecosystems roadmap
- Map Reduce Flow
- Types of Input and Output Format
- MapReduce in details
- Different types of files supported (Text, Sequence, map and Avro)
- MapReduce Job submission in YARN Cluster in details
- Role of Mappers and reducers
- Identity Mapper, Identity Reducer
- Zero Reducer, Custom Partitioning
- Combiner, Sequence file format
- Tweaking mappers and reducers
- Mapreduce package and deployment
- Code component, walk through
- Nline, Sequence file format

YARN

Hadoop Resource management component for containerization, scheduling with multi tenant feature

- Introduction to YARN
- YARN Architecture
- YARN Components
- YARN Longlived & Shortlived Daemons
- YARN Schedulers
- Job Submission under YARN
- Multi tenancy support of YARN
- YARN High Avalability
- YARN Fault tolerance handling
- MapReduce job submission using YARN
- YARN UI
- History Server
- YARN Dynamic allocation
- Containerization of YARN

NOSQL - HBase

Think beyond SQL with the column oriented datastore for realtime random read write of differential data sets

- Introduction to NoSQL
- Types of NOSQL
- Characteristics of NoSQL
- CAP Theorem
- Columnar Datastore
- What is Hbase
- Brief History
- Row vs Column oriented
- HDFS vs HBASE
- RDBMS vs HBASE
- Storage Hierarchy – Characteristics
- Table Design
- HMaster & Regions

Course Details

Hbase Contd

Think beyond SQL with the column oriented datastore for realtime random read write of differential data sets

- Region Server & Zookeeper
- Inside Region Server (Memstore, Blockcache, HFile, WAL)
- HBase Architecture (Read Path, Write Path, Compactions, Splits)
- Minor/Major Compactions
- Region Splits
- Installation & Configuration
- Role of Zookeeper
- HBase Shell
- Introduction to Filters
- Row Key Design
- Map reduce Integration
- Performance Tuning
- Hands on with Medical domain usecase
- Hive HBase Handler
- Sqoop HBase Integration

Phoenix

SQL Layer on top of HBASE for low latency, real time aggregation queries with joining capabilities

- Overview of Phonix
  - Introduction
  - Architecture
  - History
- Phoenix Hbase Integration
  - Hbase table, view creation
  - SQL & UDFs
  - SQL Line & PSQL Line of Phoenix
- Phoenix Load & Query engine
  - Understanding co processor Configurations
  - Hive -> Hbase -> Phoenix integration
  - Creation of views in phoenix
  - Load bulk data using psql
  - Serverlog Aggregation usecase

Oozie

In this module, you will do the Hands on and Exploration of the Integration of components

- Introduction
- History - Why Oozie
- Components
- Architecture
- Workflow Engine
- Nodes
- Workflow
- Coordinator
- Action (MapReduce, Hive, Spark, Shell & Sqoop)
- Introduction to Bundle
- Email Notification
- Error Handling
- Installation
- Workouts
- Orchestration of end to end tools
- Scheduling of data pipeline
- Invoking shell script, Sqoop, Hive & Spark

Course Details

Scala

Learn a scalable, Function based & Object oriented high level programming language

- Scala Introduction , History
- Why Scala , Scala Installation
- Function based programming features
- Variable / Values
- Conditional structure
- Looping constructs
- Execute Pattern Matching in Scala
- Exception Handling
- Method creation
- OOPs concepts (Classes, Objects, Collections, Inheritance, Abstraction and Encapsulation)
- Functional Programming in Scala (Closures, Currying, Expressions, Anonymous Functions)
- Know the concepts of classes in Scala Object Orientation in Scala (Primary, Auxiliary Constructors, Singleton Objects, Companion Objects)
- Traits, Mixins & Abstract classes

Python

In this module, you will learn about the Git Workflow and case study

- Python Introduction
- Evolution
- Application
- Features
- Installation & Configuration
- Objectives
- Flow Control
- Variables
- Data types
- Functions
- Modules
- OOPS
- Python for Spark
- Structures
- Collection types
- Looping Constructs
- Dictionary & Tuples
- File I/O

Spark

Learn the most advanced in-memory, fast, scalable market needed framework for large scale computation

- Spark Introduction
- History
- Overview
- MR vs Spark
- Spark Libraries
- Why Spark
- RDDs
- Spark Internals
- Pillars of Spark
- Transformations & Actions
- DAG , Lazy evaluation & execution
- Fault Tolerance
- Lineage
- Terminologies
- Cluster types
- Hadoop Integration
- Spark SQL
- Data frames, DataSets
- Optimizers – Catalys, Tungsten, AST

Course Details

Spark SQL & Streaming

Learn the Spark SQL & Streaming data Wrangling and Munging techniques for end to end processing framework

- Session
- Structured Streaming
- SQL Contexts
- Hive Context
- RDDs to Relations
- Spark Streaming
- Windowing function
- Why Spark Streaming
- Insurance Hackathon
- Data masking techniques
- Introduction to Spark ML
- Spark UI
- Job Submission into different cluster managers
- Reusable Framework creation
- SDLC implementation of Spark
- Building of Fat & Lean Jars

Spark Use Cases

Learn the real time data processing with different source and destination system integration

- PYSPARK integration
- Working with Pyspark functions
- Developing applications with Pyspark
- Maven Git Eclipsce integration
- Spark – NOSQL integration
- Spark options
- Integration with multiple sources & targets
- SCD implementation - Real time use cases
- Ebay auction analysis
- US customer data analysis
- End to end real-time integration with NIFI -> Kafka -> Spark Streaming -> Amazon S3 -> EC2 -> RDBMS -> Different Filesystems -> Hive -> Oozie & Hbase

Kafka

Publish – Subscriber Distributed Message Queue Cluster creation & integration

- Kafka Introduction
- Applications, Cluster Setup
- Broker fault tolerance
- Architecture
- Components
- Partitions & Replication
- Distribution of messages
- Producer & Consumer workload distribution
- Topics management
- Brokers
- Installation
- Workouts
- Console publishing
- Console Consuming
- Topic options
- Offset Management
- Cluster deployment in cloud

Course Details

NIFI

NIFI is a Data flow tool for real time data ingestion into Bigdata platform with tight integration with Kafka & Spark

- Nifi Introduction
- Core Components
- Architecture
- Nifi Installation & Configuration
- Fault tolerance
- Data Provenance Routing, mediation, transformation & routing
- Nifi -> Kafka -> Spark integration
- Workouts
- Scheduling
- Real time streaming
- Kafka producer & consumer
- File streaming with HDFS integration
- Data provenenance
- Packaging NIFI templates
- Rest Api integration
- Twitter data capture

Hue & Ambari

UI tools for working and managing Hadoop and Spark eco systems in a self driven way for development and administration

- Introduction
- Setting up of Ambari and HDP
- Cluster formation guide and implementation
- Deployment in Cloud
- Full Visibility into Cluster Health
- Metrics & Dashboards
- Heat Maps
- Configurations
- Services, Alerts, Admin activities
- Provisioning, Managing and Monitoring Hadoop Clusters
- Hue Introduction
- Access Hive
- Query executor
- Data browser
- Access Hive , HCatalog, Oozie , File Browser

Hortonworks/Cloudera

The top level distributions for managing Hadoop and spark ecosystems

- Installing and configuring HDP using Ambari
- Configuring Cloudera manager & HDP in sandbox
- Cluster Design
- Different nodes (Gateway, Ingestion, Edge)
- System consideration
- Commands (fsck, job, dfsadmin, distcp, balancer)
- Schedulers in RM (Capacity, Fair, FIFO)

Course Details

Elastic Search

Full Document search store for NOSQL solution with rich real time visualization & analytics capabilities

- Elastic Search Intro
- History
- Components
- Why ES
- Cluster Architecture/Framework
- All about REST APIs
- Index Request
- Search Request
- Indexing a Document
- Limitations
- Install/Config
- Create / Delete / Update
- Get /Search
- Realtime data ingestion with hive
- NIFI integration
- Spark streaming integration
- Hands-on Exercises using REST APIs
- Batch & Realtime usecases

Kibana

A Realtime integrated Dashboard with rich Visualization & Dashboards with creation of lines, trends, pies, bars, graphs, word cloud

- Kibana Introduction
- History
- Components
- Why Kibana
- Trend analysis
- Install/Config
- Creation of different types of visualizations
- Visualization integration into Dashboard
- Setting of indexes, refresh and lookup
- Discovery of index data with search
- Sense plugin integration
- Deep Visualizations
- Deep Dashboards
- Create custom Dashboards
- End to end flow integration with Nifi, Kafka, Spark, ES & Kibana

GitHub & Maven

Repository & Version controller for code management and package generation for dependency management & collaboration of different components used in SDLC

- DevOps Basics
- Versioning
- Create and use a repository
- Start and manage a new branch
- Make changes to a file and push them to GitHub as commits
- Open and merge a pull request
- Create Story boards
- Desktop integration
- Maven integration with Git
- Create project in Maven
- Add scala nature
- Maven operations
- Adding and updating POM
- Managing dependencies with the maven repository
- Building and installing maven
- Maven fat & lean jar build with submit

Course Details

AWS Cloud

Amazon Web Service components of EC2, S3 storage, access control, Subnets, Athena, Elastic Mapreduce components with Hadoop framework integration

- Introduction to AWS & Why Cloud
- Managing keys for password less connection
- All about EC2 instance creation till the management
- Amazon Virtual Private Cloud creation
- Managing the roles with Idendity Access management
- Amazon object simple storage service (S3) creation with static file uploads and exposure.
- Athena – SQL on top of S3 creation and managing
- Managing AWS EMR cluster with the formation.
- Spark & Hive Integration for data pipeline with S3, Redshift/Dynamo DB, EC2 instance
- Kafka integration

Google Cloud Platform

Identify the Platform as a service with the creation and management of Hadoop and Spark cluster in the Google cloud platform

- Registering and managing cloud account
- Key generation
- Cloud compute engine configuration and creation
- Enabling Ambari
- Multi Node cluster setup
- Hardware consideration
- Software Consideration
- Commands (fsck, job, dfsadmin)
- Schedulers in Resource Manager
- Rack Awareness Policy
- Balancing
- NameNode Failure and Recovery
- Commissioning and Decommissioning a Nodes
- Managing other GCP services
- Cluster health management

Value Added Services

Lets do a smart effort of learning how to prepare resume, interview, projects, answering cluster size, daily activities, roles, challenges faced, data size, growth rate, type of data worked etc.,

- Resume Building & flavoring
- Daily Roles & Responsibilities
- Cluster formation guidelines
- Interview Questions
- Project description & Flow
- Execution of end to end SDLC practices
- Framework integration with log monitor
- Data size & growth rate
- Architectures of Lambda, Kappa, Master slave, Peer to peer with types of data handled
- Datalake building guide
- Projects discussion
- Package & Deployment

Use Cases (We cover beyond this..)

- Setting up of Single node pseudo Distributed mode Cluster, Hortonworks Sandbox & Cloud based multinode Hortonworks cluster setup and Admin.
- Customer-Transaction data movement using Sqoop.
- Customer – Transaction Data analytics using Hive.
- Profession segmentation, Weblog analysis & Student career analysis using Hive
- Unstructured course data and Students processing using MapReduce
- Medical and Patient data handling using Hbase., Web Statistics low latency data processing using Phoenix.
- Web Server and HDFS data integration with Kafka using NIFI.
- Ebay Auction data analytics and Crime Department data processing using Spark Core.
- Retail Banking data processing using Spark core.
- Server Log Analysis using spark core, Sensus data analysis using Spark SQL.
- Realtime Network, HDFS and Kafka data processing using Spark Streaming.
- Create rich Visualization & Dashboard using Kibana with Ebay & Trans data
- *Managing twitter, open data, RESTAPI data using NIFI-> KAFKA->SPARK*

Use Cases

Key Stuffs behind the success that provides real experience…

inceptez

Projects (We cover beyond this..)

- *Project 1: Sentimental Analytics - Web event analytics using Linux, HDFS, Hive, Hbase & Oozie*
- *Project 2: Server log analysis for view ship pattern, threat management and error handling – Sqoop, Hive, HCatalog, HBase, Phoenix.*
- *Project3 : Datalake for Usage Pattern Analytics & Frustration scoring of customer - Data Warehouse Migration/consolidation using Sqoop, HDFS, Masking UDF Hive, Oozie, HBase, Phoenix.*
- *Project 4: Realtime Streaming analytics using Vehicle fleet data using IOT, RPA, Kafka, Spark, NIFI, Kafka, Hive, HBASE/ES, Phoenix*
- *Project 5: DataLake exploration using Spark SQL, Hive, HBASE/ES*
- *Project 6: Fast Data Processing for Customer segmentation using Oracle, XML, Kafka, Spark, NIFI, AWS S3, Hive, HBASE/ES*
- *2 Hackathons*
- *1 Exams*
- *1 Production packaging and deployment*
- *1 Cloud formation*
- *1 Live Project end to end execution and implementation*
- *1 Job support videos*
- *1 Chat & text mining*

Projects

Key Stuffs behind the success that provides real experience…

inceptez

THANK YOU

Inceptez Technologies Pvt. Ltd.

+91 70107 90330

info@inceptez.com

www.inceptez.com

www.inceptez.in

Accelerate your career gear to Big Data joining ...

inceptez

inceptez