|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **SUMIT GANGULY**  **~ Hadoop Developer ~**  **E-Mail:** sumitganguly85@gmail.com, **Phone:** +91-9874323700, +91-8910897710  **Linkedin:** [www.linkedin.com/in/sumit-ganguly-281085](http://www.linkedin.com/in/sumit-ganguly-281085)   |  |  | | --- | --- | | **Profile Snaps :** | * **1.5 Year of Experience As Hadoop Developer on Cloudera and Hottonworks distribution with both on premises and cloud platform , IBM Big Data University certification** on **Hadoop Foundation Level 2** & **Hadoop Programming Level 2** along with **Oracle Certified Associate** having overall **7.5 Years of Industry Experience in Application Software Development** with **Oracle PL/SQL, UNIX Shell Scripting** and **Microsoft Excel Macro VBA Programming** * **Having experience with Cloudera distribution of 7 Months and 10 Months with Hottonworks distribution.** * **Currently working on Hottonworks (HDP 2.6) distribution hosted in private cloud environment named JAWS (JIO Advanced Web Services)** * Sound knowledge in **Apache Hadoop Components like Edge Node, Active Name Node, Data Node, Secondary Name Node, Passive Name Node, Block Placement Policy & Rack Awareness, Anatomy of File write/read operation in HDFS, Sync Marker and Secondary Checkpoint Backup Mechanism** * Good understanding in **YARN** like **Resource Manager, Application Manager, Data Node Manager, Application Master, Container, Heartbeat Mechanism, Life Cycles of a submitted Job in YARN, YARN Child, Job History Server** * Good knowledge in difference between **Generation I & Generation II Hadoop in terms of HDFS Federation, Fail Over Fencing and STONITH Process under High Availability** * Good knowledge in **Map Reduce Paradigm like Input Split, Record Boundary, TextInputFormat, Mapper Class, Mapper Output Collector, Circular Memory, Spilled File, Partitioner, Combiner, Shuffle Phase, Reducer, Reducer Output Collector** * Hands on Knowledge in writing **Basic Map Reduce Programme in Java version 1.7.0\_67** * Hands on exposure of **Data Import and Export done through SQOOP**. **Performed Incremental data importing into HDFS with append and lastmodified mode** * Good knowledge of **Apache Flume agent**. **Performed log import into HDFS with Netcat Source, Memory Channel and HDFS Sink** * Understanding of different components of **Apache Kafka like Producer, Topics, Kafka Brokers, Partitioner, Replication, Offset Number, Consumer Groups** * Perusing **Apache Spark with Scala** | |  |  | | **Technical Competencies :** | * **HDFS, Map Reduce & YARN with Apache Hadoop Big Data** * **HIVE** * **Apache Sqoop** * **Apache Flume** * **Apache Kafka** * **Apache Spark with Scala** * **Apache PIG** * **Core Java** * **Unix Shell Scripting** * **RDBMS, Oracle PL/SQL 11g** * **Excel VBA Macro Programming** * **Microsoft Silverlight 3.0** | |  |  | | **Certifications :** | * **Big Data 101 (Big Data University – Cognitive Class AI)** * **Hadoop 101 (Big Data University – Cognitive Class AI)** * **MapReduce and YARN (Big Data University – Cognitive Class AI)** * **Accessing Hadoop Data Using Hive (Big Data University – Cognitive Class AI)** * **Moving Data into Hadoop using SQOOP & Flume (Big Data University – Cognitive Class AI)** * **Simplifying data pipelines with Apache Kafka (Big Data University – Cognitive Class AI)** * **Apache Pig 101 (Big Data University – Cognitive Class AI)** * **SQL Access for Hadoop (Big Data University – Cognitive Class AI)** * **Oracle Certified Associate (Oracle Corporation)** | |  |  | | **Organisational Experience :** | **Since Nov’17 with Tech Mahindra, Mumbai, India as Tech Lead**   |  |  | | --- | --- | | **Project Handled:** |  | | Nov’17- Till Date: | As Hadoop Consultant working at customer location | | Project Location: | Navi Mumbai, India |  |  |  | | --- | --- | | **Title:** | Data Lake for Petrochemical Refinery and Marketing | | **Customer:** | Giant in Petrochemical Industry – Mumbai, India | | **Technology Used**: | Hadoop Generation 2, Hottonworks HDP 2.6, JIO Advanced Web Services Cloud, Sqoop, Hive, Shell Scripting, Kylo, Spark, Kafka, Java, Python |   **Roles & responsibilities:**   * Defining and configuring Hottonworks Master and Slave nodes in Cloud * Install and configure Hadoop eco system components available with Hottonworks HDP 2.6 package on cloud named JIO Advanced Web Services * Interacting with Business Analyst to get the business requirement on Petrochemical Refinery and Marketing vertical * Developing Framework/Tool for data ingestion into Hive using Sqoop. The core engine of this framework/tool is being developed with Shell script using Xmllint API and wrapper of that framework tool would be developed in Java or Python * Loading data using that Framework/Tool into Hive from RDBMS systems like Microsoft SQL Server and Oracle * Developing scripts in Hive for data modelling and transformation to implement different KPI provided by Business Users * Contributing in performing unit and SIT testing and providing UAT support for business users * Administering the preparation of various types of Technical and functional documents as per project needs’ * Upcoming Milestone – The data ingestion/extraction/near real time transformation frameworks/tools for unstructured data like log files, image, pdf, Data of Petrochemical R&D Laboratory, various digital signal received from retail petrol pump etc are to be developed in Kafka, Spark, Python     **Other responsibilities apart from project:**   * Mentoring technically on Hadoop to few employees of Tech Mahindra as a part of reskilling program.   **Jul’14- Sep’17 with Cognizant Technologies Solutions India Pvt. Ltd, Kolkata, India/ Kowloon, Hong Kong as Associate**   |  |  | | --- | --- | | **Project Handled:** |  | | Jun’15- Jan’17: | As Onsite Developer& Coordinator | | Feb’17- till date: | As Offshore Developer | | Onsite Location: | Kowloon, Hong Kong | | Offshore Location: | Kolkata, India |  |  |  | | --- | --- | | **Title:** | Client Administration System | | **Customer:** | Reputed Life Insurance Company– Kowloon, Hong Kong | | **Technology Used**: | Apache Hadoop, HIVE, Sqoop, Cloudera distribution 5.10, Oracle Database 11g, Oracle Forms 11g, Oracle Reports, Unix Shell Scripting, Windows Scripting, Control – M Job Scheduler, Microsoft Excel VBA Programming, Windows 2008 Server, Unix Server |   **Roles & responsibilities:**   * Performing various activities like requirement gathering; performing functional & impact analysis of the requirements and coding for implementing the same * Extending support in preparing functional specifications and walkthrough with business analyst and business users of Manulife and finalizing the requirement * Facilitating tuning of SQL, PL/SQL in existing package &procedures * Load data into Hive using Sqoop and perform analysis in HIVE with help of HIVE inbuilt functions and few UDF developed in Java to meet requirements as per KPI provided by Business Users and BA * Administering the process of reviewing coding as per Manulife IT guidelines * Contributing in performing unit and SIT testing while providing UAT support for business users * Administering the preparation of mod list, SIT test case and functional documents * Providing support for post-production implementation   **Significant Accomplishments:**   * Bagged Difference Maker Award * Played a pivotal role in implementing process enhancement initiatives through measures such as performance tuning, code and auditing of the project  |  |  | | --- | --- | | **Title:** | Manulife Mainframe Re Platform | | **Customer:** | Reputed Life Insurance Company– Kowloon, Hong Kong | | **Period:** | Jul’14- May’15 | | **Technology Used**: | Oracle Database 12c, Oracle SQL Loader, Aix Shell Scripting, Control – M job scheduler, Microsoft Excel VBA Programming, Windows 2008 Server, AIX Server |   **Roles & responsibilities:**   * Performed various activities like fixing data migration strategies, preparing scripts for Oracle SQL Loader for data loading into Oracle tables from data files generated by Mainframe Cobol Program from Mainframe VSAM File * Perform the process of writing Oracle stored procedures and AIX Shell Scripting for data migration * Extended support in preparing tools in Excel VBA Macro and AIX to automate the migration process like SQL loading control file and low value update after migration * Contributed in reviewing coding as per Manulife IT guidelines * Facilitated unit testing & SIT testing while providing UAT support to the business users   **Significant Accomplishments:**   * Bagged Outstanding Contribution Award   **Jul’10- Jul’14 with CMC Limited, Kolkata as IT Engineer**   |  |  | | --- | --- | | **Title:** | Product Development & Customization – General Reinsurance | | **Customer:** | Reputed General Insurance Company– Mumbai, India | | **Period:** | Apr’12 to Jul’14 | | **Technology Used**: | Oracle Database 10g, Microsoft Silverlight 3.0, WCF Service with VB.NET, Windows 2008 Server |  |  |  | | --- | --- | | **Title:** | Product Development & Customization – General Reinsurance | | **Customer:** | Reputed General Insurance Company– Chennai | | **Period:** | Jul’10 to Mar’12 | | **Technology Used**: | Oracle Database 10g, Microsoft Silverlight 3.0, WCF Service with VB.NET, Windows 2008 Server |   **Significant Accomplishment:**  Bagged Best Employee Award in two consecutive years | |  |  | | **Education :** | * **B.TECH**. (Electronics & Instrumentation Engineering) with **74.9%** Marks from West Bengal University of Technology, Kolkata in **2008** * **XII** with **63%** Marks from West Bengal Council of Higher Secondary Education in **2003** * **X** with **75.75%** Marks from West Bengal Board of Secondary Education in **2001** | |  |  | | **Tools & environments :** | * **Oracle Virtual Box** * **Cloudera– 5.10** * **Hottonworks HDP 2.6** * **Cloud – JIO Advanced Web Services** * **UNIX – Cent OS** * **Web Application - Cloudera Hue** * **Windows – Windows 7, Windows 10, Windows 2003 Server, Windows 2008** * **Eclipse Indigo & Eclipse Mars** * **Oracle PL/SQL Developer** * **Microsoft Excel** * **Visual Studio 2008** | |  |  | | **Personal Details :** | |  |  | | --- | --- | | **Date of Birth:** | 28th October 1985 | | **Languages Known:** | English, Bengali and Hindi | | **Address:** | Village – Kanaipur, P.O. – Dakshin Basudebpur, P.S – Dhaniakhali, District- Hooghly, 712301, West Bengal, India | | **Passport:** | R4893018 | |   \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_    SUMIT GANGULY |