Sandish Kumar H N

Greater Minneapolis-St. Paul Area

Big Data Solutions Engineer at phData

sandishsany@rediffmail.com - 9008990742

Summary

I'm Senior Big Data Solution Engineer with 6+ years of experience in multiple business domains using latest technologies and platforms. The focus is on Cloud, Big Data, Machine Learning, Data Science and Data Mining. A skilled developer, an architect with strong problem solving, debugging and analytical capabilities, who creates the technical vision and actively engages in understanding customer requirements. I focus, particularly on software performance and efficiency. Result oriented and hands-on, who skillfully balances between meeting resource and time constraints while doing it right.

CERTIFICATION:

- Cloudera Certified Hadoop and Spark Developer.
- AWS Certified Solutions Architect - Associate.
- DataBricks Certified Apache Spark Developer.
- DataStax Certified Apache Cassandra Developer.
- MapR Certified Hadoop Developer.
- Cloudera Certified Hadoop Administrator.

- Languages: Proficient in Core-Java, Scala, Familiar with Python;
- Technologies: Apache Hadoop (HDFS, YARN, MR), Apache Spark (Core, SQL, Streaming), Apache - Kafka (Core, KSQL, Streams), Akka (Actor Model, Http, Streams, Cluster), StreamSets, Nifi
- Data Warehouse/MPP: Apache Hive, Impala, Presto, Redshift
- Columnar DataBases: Kudu, Cassandra, Hbase
- Cloud Platforms: Amazon Web Services (EC2, S3, Cloudformation, EMR, Redshift, Kinesis, Glue, API Gateway, VPN, IAM, Athena), Google Cloud Platform(Cloud storage, BigQuery, Compute Engine, Container Services)
- Container Services: Docker, Kubernetes
- Development Env Tools: Jenkins, Gerrit, Git
- Workflow: Oozie, Airflow, Rundeck
- Search: Lucene, Elasticsearch, Solr
- Data Formats: Avro, Parquet, Json, CSV, XML
- Security: Sentry, Kerberos, SSL/SAL,

- Others: MySql, Bash, JavaScript

Experience

Apache Contributor at The Apache Software Foundation
November 2016 - Present (2 years 11 months)

Open Source work:

Scala and Akka:

Apache Kudu, Akka Alpakka, Apache Gearpump

Java:

Apache NIFI , Apache Sqoop, Streamsets, Apache Kudu

Links to Opensource work.

https://github.com/apache/nifi/commits?author=SandishKumarHN

https://github.com/pulls?&q=is:pr+author:SandishKumarHN+archived:false

https://reviews.apache.org/users/sany/

https://issues.apache.org/jira/browse/GEARPUMP-376?jql=assignee%20in%20(%22sanysandish%40gmail.com%22)%20order%20by%20created%20DESC

https://gerrit.cloudera.org/#/q/owner:%22Sandish+Kumar+HN+%253Csanysandish%2540gmail.com%253E%22+status:Merged

Big Data Solutions Engineer at phData
July 2016 - Present (3 years 3 months)

- Building and managing solutions for phData customers

- Working on Hadoop, Apache Spark, Kafka, Impala, Python, Java, Scala, Akka Streams, Akka Http, Docker, AWS, Apache Kudu, StreamSets, Kerberos, Docker, Ansible, Cloudera, Spring-Hbase-Boot

Senior Big Data Consultant at Third Eye Consulting Services & Solutions LLC.
December 2013 - July 2016 (2 years 8 months)

Third Eye Consulting Services & Solutions LLC, Provides Big Data analytics solutions that can be easily customized to the customer’s specific business needs and seamlessly integrated into their existing infrastructure. -- Online retail analytics -- Digital advertisements analytics -- Traffic Data analytics

I have worked on multiple BigData projects with different clients like Intel, Hortonworks, GridGain, Microsoft

Technologies used : Apache Spark, Hadoop, MapReduce, Storm, Kafka, HDFS, Hbase,Hive, Shark, BigQuery, Spring Data, Cassandra, Social Media Data, Amazon Web Services, Rackspace, Google Cloud

Big Data Consultant at Gimmie.io
January 2015 - March 2015 (3 months)

* Flow design for click stream data

* Installation of Storm and Kafka on multi-node cluster

* Written Kafka Rest API to collect events from front-end

* Written Storm topology to accept the events from Kafka producer and emit into Cassandra DB

* Written Junit test cases for Storm Topology.

Big Data Developer at Positive Bioscience
February 2013 - January 2014 (1 year)

- Design and development of software for Bioinformatics, Next Generation Sequencing (NGS) in Hadoop MapReduce framework, MangoDB using Amazon S3, Amazon EC2, Amazon Elastic MapReduce(EMR).

- Developed Hadoop MapReduce program to perform custom Quality Check on genomic data. Novel features of the program included capability to handle file-format/sequencing-machine errors, automatic detection of base-line PHRED score and being platform agnostic (Illumina, 454 Roche, Complete Genomics, ABI Solid input format data).

- Developed a Hadoop MapReduce program to perform sequence alignment on sequencing data. The MapReduce program implements algorithms such as Borrows-Wheeler Transform (BWT), Ferragina-Manzini Index (FMI), Smith-Waterman dynamic programming algorithm using Hadoop distributed cache.

- Configured and ran all MapReduce programs on 20-30 node cluster (Amazon EC2 spot instances) with Apache Hadoop-1.4.0 to handle 600GB/sample of NGS genomics data.

- Configured a 20-30 node (Amazon EC2 spot Instance) Hadoop cluster to transfer the data from Amazon S3 to HDFS and HDFS to Amazon S3 and also to direct input and output to the Hadoop MapReduce framework.

- Successfully ran all Hadoop MapReduce programs on Amazon Elastic MapReduce framework by using Amazon S3 for Input and Output.

- Developed java RESTfull webservices to upload data from local to Amazon S3, listing S3 objects and file manipulation operations.

- Developed MapReduce programs to perform Quality Check, Sequence Alignment, SNP calling, SV/CNV detection on single-end/paired-end NGS data.

- Designed and transmitted a RDBMS(SQL) Database to NOSQL MangoDB Database.

Software Developer (Hadoop) at PointCross
December 2011 - February 2013 (1 year 3 months)

-Implemented Java Hbase MapReduce paradigm to load raw seismic and sensor data onto Hbase database on a 4 node Hadoop cluster.
-Developed RESTfull web service application to fetch values from Hbase to UI in XML and JSON Response format.
-Scripted separately in Hive and Pig to aggregate multiple drilling parameter values for downstream
processing on a 4 node cluster.
-Programmed in Java MapReduce to port Indexed data from Hbase to Solr search engine
-Implemented cross-connection settings between Hive and Hbase.
-Installed and configured Hadoop 1.X, Hive, pig, Hbase on a 4 node local cluster.
-Customized OpenLayers maps to display KML information stored in Habse database to display oil well Information such as geo-location data.

Education

Visvesvaraya Technological University
Bachelor's Engineering, Computer Science, 2008 - 2011
School of Mines
Diploma in Computer Science, Computer Science, 2005 - 2008
Activities and Societies: ISQLProject Internet Structure query language

Sandish Kumar H N

Greater Minneapolis-St. Paul Area

Big Data Solutions Engineer at phData

sandishsany@rediffmail.com - 9008990742

LinkedIn Recruiter

4 people have recommended Sandish Kumar

"Sandish is one of the most talented big data software engineers I have ever met. He worked his way up from the bottom and has acquired an amazing set of data wrangling, data cleaning and filtering and end to end big data machine learning pipelines. I highly recommend Sandish and would hire him again in a second!"

—David Johnston, Sr AI/ML Cloud Architect, Wipro Limited, managed Sandish Kumar at The Apache Software Foundation

"Good,hardworker"

—Rajesh D S, Senior Associate (BigData), Cognizant Technology Solutions, managed Sandish Kumar indirectly at PointCross

"I have worked with Sandish on several genomics projects to build proprietary software for interpreting Next Generation Sequencing data. Although he was not from a biotechnology background, he was quick to pick up new terminologies and concepts. He was diligent and committed to getting a quality product out on time. He is hard-working, self-motivated and enthusiastic. On a personal front, he gets along well in a team, respects good management and aspires to emulate quality leaders. In short, if you want a guy to get your work done, I would recommend Sandish."

—Prabhasa R., Scientific Manager, Excelra Knowledge Solutions Pvt Ltd, worked directly with Sandish Kumar at Positive Bioscience

"Sandish is very hard-worker and opened to new ideas and technologies"

—Leandro Cavaleri Gerhardinger, Digital Transformation Enabler & Product Engineering Leader, NTUC Enterprise Co-operative Limited, managed Sandish Kumar at Gimmie.io

Sandish Kumar H N
sanysandish@gmail.com

Ph: +919008990742
skype: sandishhadoop

SUMMARY:

I'm Big Data Solutions Engineer Consultant with 5+ years of experience in multiple business domains using latest technologies and platforms. Focus is on Big Data, Cloud, Machine Learning, Data Science and Data Mining. A skilled developer, architect with strong problem solving, debugging and analytical capabilities, who creates the technical vision and actively engages in understanding customer requirements. I focus particularly in software performance and efficiency. Result oriented and hands on, who skillfully balances between meeting resource and time constraints, while doing it right.

- Incube **Open Source Committer** at StreamSets Data Collector Project.
- Strong experience creating real time data streaming solutions using Apache Spark Core, Spark SQL & Data Frames, Spark Streaming, Apache Storm, Kafka.
- Experience in building Data-Pipelines using Big Data Technologies
- Hands-on experience in writing MapReduce programs and user-defined functions for Hive and Pig
- Experience in NoSQL technologies like HBase, Cassandra
- Excellent understanding /knowledge on Hadoop (Gen-1 and Gen-2) and various components such as HDFS, Job Tracker, Task Tracker, Name Node, Data Node, Resource Manager (YARN).
- Excellent understanding and knowledge of NOSQL databases like HBase, and Cassandra.
- Experience in importing and exporting data using Sqoop from HDFS to Relational Database Systems (RDBMS) and from RDBMS to HDFS.
- Proficient at using Spark APIs to cleanse, explore, aggregate, transform, and store machine sensor data
- Configured a 20-30 node (Amazon EC2 spot Instance) Hadoop cluster to transfer the data from Amazon S3 to HDFS and HDFS to Amazon S3 and also to direct input and output to the Hadoop MapReduce framework.
- Hands-on experience with systems-building languages such as Scala, Java
- Hands-on experience with message brokers such as Apache Kafka and RabbitMQ.
- Worked extensively with Dimensional modeling, Data migration, Data cleansing, Data profiling, and ETL Processes features for data warehouses.
- Implemented Hadoop based data warehouses, integrated Hadoop with Enterprise Data Warehouse systems
- Built real-time Big Data solutions using HBASE handling billions of records.
- Involved in designing the data model in Hive for migrating the ETL process into Hadoop and wrote Pig Scripts to load data into Hadoop environment.
- Expertise in writing Hive UDF, Generic UDF's to incorporate complex business logic into hive queries in the process of performing high level data analysis.
- Worked on Spark Machine Learning library for Recommendations, Coupons Recommendations, Rules Engine.
- Experience in working with various Cloudera distributions (CDH4/CDH5) and have knowledge on Hortonworks and Amazon EMR Hadoop Distributions.

- Experience in administering large scale Hadoop environments including design, configuration, installation, performance tuning and monitoring of the cluster using Cloudera manager and ganglia.
- Worked extensively with Dimensional modeling, Data migration, Data cleansing, Data profiling, and ETL Processes features for data warehouses
- Experience in Object Oriented Analysis Design (OOAD) and development of software using UML Methodology, good knowledge of J2EE design patterns and Core Java design patterns.
- Experience in designing both time driven and data driven automated workflows using Oozie.
- Experience in writing UNIX shell scripts.

CERTIFICATION:

- DataBricks Certified Apache Spark Developer
- Cloudera CCA Spark and Hadoop Developer
- Cloudera Certified Administrator for Apache Hadoop (CCAH)
- MapR Certified Apache Hadoop Developer
- DataStax Certified Apache Cassandra Developer

TECHNICAL HIGHLIGHTS

- **Big Data Technologies**: Hadoop (Hortonworks, Cloudera, MapR), Spark, Spark Streaming, Spark Sql, Spark ML, Mapreduce, HDFS, Cassandra, Storm, Apache Kafka, Streamsets, Flume, Oozie, Solr, Zookeeper, Solr, Tez, Data Modelling, Pig, Hive, Impala, Drill, Sqoop and RabbitMQ.
- **NOSQL Database**: Hbase, Cassandra,
- **SQL DB's**: Hive, Pig, PrestoDB, Impala, SparkQL
- **Search:** Hsearch, Apache Blur, Lucene, Elasticsearch, Nutch
- **Programming Languages**: Java, Scala, Basic's (Python, Clojure)
- **Cloud Platform**: Amazon Web Services (EC2, Amazon Elastic Mapreduce, Amazon S3), Google Cloud Platform (Bigquery, App Engine, Compute Engine, Cloud SQL), Rackspace (CDN, Servers, Storage), Linode Manager
- **Monitoring and Reporting**: Ganglia, Nagios, Custom Shell scripts, Tableau, D3.js, Google Charts
- **Data:** E-Commerce, Social Media, Logs and click events data, Next Generation Genomic Data, Oil & Gas, Healthcare, Travel
- **Other:** HTML, JavaScript, Extjs, CSS, JQuery

WORK EXPERIENCE:

Big Data Solutions Engineer
PhData Inc
Bangalore, India August 2016 to Present

- Building and managing solutions for phData customers
- Working on Hadoop, Apache Spark, Impala, Apache Kudu, StreamSets, Spring-Hbase-Boot, Kerberos, Docker, ansible, Cloudera, Java, Scala, C++
- Successfully developed custom Origin, Destinations and processors for Hadoop, Spark, Amazon S3 and Kafka in Streamsets open source project.
- Incube Open Source committer at Streamsets

Data Engineer (R&D)
Intel, Bangalore, India Dec 2015 to September

Project Name: OSINT (Open Source Intelligent Tool)
Project Description:

An Open source platform for all Malware related information, Raise cross-organizational situational awareness. Manage risks across your internal org and supply chain. Speed response (and pre-response) to incidents. Prioritize effective use of tactical cyber solutions. Collaborate, budget and strategize around cyber defense, Educate and inform your workforce.

The OSINT Platform includes - Websites, blogs, forums, breach databases, exploit databases, malware, vulnerability, Dark Web sources and myriad others. Your Own Evaluated Threat Data - Your own team's low-level data that's confirmed as "real" or relevant and diligently recorded, analyzed (e.g. found Trojans or confirmed SNORT hits). Highly-Focused, Highly-Relevant Data Feeds - Commercial phishing feeds, patch management updates, Spam analytics, AV/AM, Government alerts and indicators. Partner and Supply Chain Data - Extending reporting, sharing and data collection from your own "Private ISAC". SIEM Information - Evaluated events from SIEM analysis and exploration. TIP Data - HUMINT analysis and other alerts, low-level data from "traditional" Threat Intelligence Platforms that can be confirmed as relevant threat events

Senior Big Data Consultant
Third Eye Consulting Services & Solutions -Bangalore, India Dec 2013 to Dec 2015

Project Name: Yardstick Spark (Open Source)
Client: Gridgain
Project Description: Yardstick Apache Spark is a set of apache Spark and apache Ignite comparative benchmarks written on top of Yardstick framework.
Responsibilities:

- I have successfully written Spark CoreRDD application to read auto generated 1 billion records and compare with IgniteRDD in Yardstick framework to measure performance of Apache Ignite RDD and Apache Spark RDD.
- I have successfully written Spark DataFrame application to read from HDFS and analyze 10 million twitter records using Yardstick framework to measure performance of Apache Ignite SQL and Apache Spark DataFrame.
- I have successfully written Spark Streaming application to read streaming twitter data and analyze twitter records in real time using Yardstick framework to measure performance of Apache Ignite Streaming and Apache Spark Streaming.
- Implemented test cases for Spark and Ignite functions using Scala as language.

- Hands-on experience in setting up 10 node Spark cluster on Amazon Web Service's using Spark EC2 script.
- Implemented D3.js and Tableau charts to show performance difference between Apache Ignite and Apache Spark.

Environment: Spark, Spark Core, Data Frame's, Spark Streaming, Scala, HDFS, Apache Ignite, Yardstick Tool, D3js, Tableau, AWS, 10 million twitter data records and 1 billion auto generated records.

Project Name: E-Commerce Data Pipeline
Client: Obsessory.com
Project Description: Obsessory is a technology company that provides a web and mobile platform to assist shoppers in discovery, search, comparison, and tracking of items across the Internet. Obsessory's powerful search engine catalogs millions of products from online stores on a daily basis and uses proprietary algorithms to enhance the depth and breadth of the user's search. Obsessory employs adaptive and social learning to continuously refine the search results and present the user with the most relevant selection of items. Furthermore, Obsessory helps users keep track of desired items across the Internet and get notified of price changes, and availability of tracked items as well as sales, store events and promotions **Responsibilities:**

Pre-Processing:

- Crawling of 100+ sites Data using Nutch
- Fashion based ontology maintenance
- Using Scala, Spark & echo system to enriched given data using Fashion Ontology to Validation/Normalizing the data
- Designed schema and modeling of data and Written algorithm to store all validated data in Cassandra using Spring Data Cassandra REST
- Programs for Validation/Normalizing/Enriching and REST API to Develop UI Based on manual QA Validation. Used SparkSQL, Scala to running QA based SQL queries.

Indexing:

MR Programs on top of Hbase:

- To standardize the Input Merchants data
- To upload images to RackSpace CDN
- To index the given Data sets into HSearch
- To MR programs on Hbase to extract the color information from Images including density.
- To MR programs on Hbase to persist the data on Hbase tables
- above MR jobs will run based on timing and bucketing. **Color-Obsessed:**

Using Image color and density data Users are allowed to select 1,2.. colors with different densities and result will be a list of products where each product image contains all give colors with exact density this has been implemented on top Hbase using Spring REST web service for color Obsessed search API.

Post-Processing:

- Setting up the Spark Streaming and Kafka Cluster
- Developed a Spark Streaming Kafka App to Process Hadoop Jobs Logs
- Kafka Producer to send all slaves logs to Spark Streaming App
- Spark Streaming App to Process the Logs with given rules and produce the Bad Images, Bad records, Missed Records etc.
- Spark Streaming App collect user actions data from front end
- Kafka Producer based Rest API to collect user events and send to Spark Streaming App
- Hive Queries to Generate Stock Alerts, Price Alerts, Popular Products Alerts, New Arrivals for each user based on given likes, favorite, shares count information

- Worked on SparkML library for Recommendations, Coupons Recommendations, Rules Engine.

Environment: H Search (Hbase+lucene), Cassandra, Hive, Spark (Core, SQL, ML, Streaming), Hadoop, MapReduce, Amazon Webservice, Linode, CDN, Scala, Java, Affiliates feeds Rakuten, CJ, Affiliate window, Webgains.

Project Name: Cimbal/MobApp Pay
Client: Intel
Project Description: Cimbal is a mobile promotion and payment network designed to increase business sales and deliver targeted deals to consumers **Responsibilities:**

- Written MapReduce programs to validate the data
- Written more than 50 Spring Data Hbase rest API's in Java
- Schema design on Hbase and cleaning data
- Written Hive queries for analytics on user's data.

Environment: Hadoop MapReduce, Hbase, Spring Data Rest Web Service, CDH, Users Payment Data

---

Project Name: Truck Events Analysis
Client: HortonWorks
Project Description: The Trucking business is a high-risk business in which truck drivers venture into remote areas, often in harsh weather conditions and chaotic traffic on a daily basis. Using this solution illustrating Modern Data Architecture with Hortonworks Data Platform, we have developed a centralized management system that can help reduce risk and lower the total cost of operations.
Responsibilities:

- Written a simulator to send/emit events based on NYC DOT data file.
- Written Kafka Producer to accept/send events to Kafka Producer which is on Storm Spout
- Written Storm topology to accept events from Kafka Producer and Process Events
- Written Storm Bolt to Emit data into Hbase, HDFS, Rabbit-MQ Web Stomp
- Hive Queries to Map Truck Events Data, Weather Data, Traffic Data

Environment: Hadoop, HDFS, Hive, HBase, Kafka, Storm, Rabbit-MQ WebStormp, Google Maps, New York City Truck Routes from NYC DOT. -Truck Events Data generated using a custom
simulator. - Weather Data, collected using APIs from Forcast.io. -Traffic Data, collected using APIs from MapQuest.

Project Name: Comparative Analysis of Big Data Analytical Tools – (Hive, Hive on Tez, Impala, SparkQL, Apache Drill, BigQuery, PrestoDB running on the Google Cloud and AWS)
Client: ThirdEyeCss.com **Responsibilities:**

- Installation of Hive, Hive on Tez, Impala, SparkQL, Apache Drill, BigQuery, PrestoDB, Hadoop, Cloudera CDH, Hortonworks HDP
- Schema design for data sets on all Hive, Hive on Tez, Impala, SparkQL, Apache Drill, BigQuery, PrestoDB
- Query design for given data set
- Debugging on Hive, Hive on Tez, Impala, SparkQL, Apache Drill, BigQuery, PrestoDB, Hadoop, Cloudera CDH, Hortonworks HDP
- Time Comparison of each Hive, Hive on Tez, Impala, SparkQL, Apache Drill, BigQuery, PrestoDB
- Time comparison between different cloud platforms
- Times metrics web based visualization design on google charts

Environment: Hive, Hive on Tez, Impala, SparkSQL, Apache Drill, BigQuery, PrestoDB,

Hadoop, Cloudera
CDH, Hortonworks HDP, Google Cloud Platform, Amazon web service, Twitter streaming data

---

Senior Big Data Consultant
Positive Bioscience – Mumbai, India Jan-2013 to Dec-2013

Project Name: Next Generation DNA Sequencing Analysis **Client:** Positive Bioscience **Responsibilities:**

- Developed a Hadoop MapReduce program to perform sequence alignment on NGS data.
- The MapReduce program implements algorithms such as Borrows-Wheeler Transform (BWT), Ferragina-Manzini Index (FMI), Smith-Waterman dynamic programming algorithm using Hadoop distributed cache.
- Design and development of software for Bioinformatics, Next Generation Sequencing (NGS) in Hadoop MapReduce framework, Cassandra using Amazon S3, Amazon EC2, Amazon Elastic MapReduce(EMR).
- Developed Hadoop MapReduce program to perform custom Quality Check on genomic data. Novel features of the program included capability to handle fileformat/sequencing-machine errors, automatic detection of base-line PHRED score and being platform agnostic (Illumina, 454 Roche, Complete Genomics, ABI Solid input format data).
- Developed a Hadoop MapReduce program to perform sequence alignment on sequencing data. The MapReduce program implements algorithms such as Borrows-Wheeler Transform (BWT), Ferragina-Manzini Index (FMI), SmithWaterman dynamic programming algorithm using Hadoop distributed cache.
- Configured and ran all MapReduce programs on 20-30 node cluster (Amazon EC2 spot instances) with Apache Hadoop-1.4.0 to handle 600GB/sample of NGS genomics data.
- Configured a 20-30 node (Amazon EC2 spot Instance) Hadoop cluster to transfer the data from Amazon S3 to HDFS and HDFS to Amazon S3 and also to direct input and output to the Hadoop MapReduce framework.
- Successfully ran all Hadoop MapReduce programs on Amazon Elastic MapReduce framework by using Amazon S3 for Input and Output.
- Developed java Restful web services to upload data from local to Amazon S3, listing S3 objects and file manipulation operations.
- Developed MapReduce programs to perform Quality Check, Sequence Alignment, SNP calling, SV/CNV detection on single-end/paired-end NGS data.
- Designed and transmitted a RDBMS(SQL) Database to NOSQL Cassandra Database.

Hadoop Developer
PointCross.com - Bangalore, India Nov-2011 to Jan-2012

Project Name: DDSR (Drilling Data Search and Repository)

This project aims to provide analytics for Oil and Gas exploration data. This DDSR repository build by using HBase, Hadoop and its sub projects. We are collecting thousands of wells data from across the globe. This

data is stored in Hbase and Hive by using Hadoop MapReduce jobs. On top of this data we are building analytics for search and advanced search.

Project Name: Seismic Data Server & Repository (SDSR™)

Our Seismic Data Server & Repository solves the problem of delivering, on demand, precisely cropped SEG-Y files for instant loading at geophysical interpretation workstations anywhere in the network. Based on Hadoop file storage, Hbase™ and MapReduce technology, the Seismic Data Server brings fault-tolerant petabyte-scale store capability to the industry. Seismic Data Server supports post-stack traces now with pre-stack support to be released shortly.

ADDITIONAL INFORMATION:

EDUCATION:

- Bachelor of Engineering in Computer Science and Engineering, University of VTU Bangalore, Karnataka, India, 2011.
- Diploma in Computer Science and engineering University of KAE, Bangalore, Karnataka, India, 2008

REFERENCE LINKS:

- LinkedIn: https://in.linkedin.com/in/sandishkumar
- Twitter: https://twitter.com/sandishsany
- GitHub: https://github.com/SandishHadoop
- Skype: sandishhadoop