QIAO PENGFEI

Software/Data Engineer

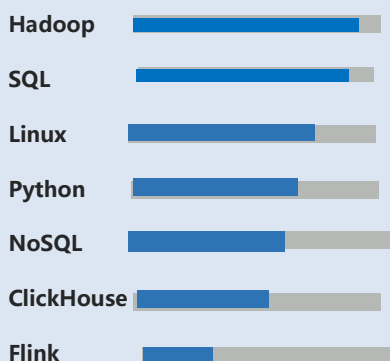## INFORMATIONS

♂: Male

@: joe19940422@gmail.com

https://blog.csdn.net/qq_33290422

☎8613219950624

SKILLS

Hadoop

SQL

Linux

Python

NoSQL

ClickHouse

Flink

LANGUAGE

Mandarin Native

English　Good

Dutch　Beginner(A1)

EDUCATION

Bachelor's in Computer science

Harbin Normal university　2013 – 2017

INTERNSHIP EXPERIENCE

2016.2 – 2017.07　　　　　　　**Like & Thupdi**　Shanghai/Beijing

① Installed and configured MapReduce,Hive and the HDFS;
② Implemented HDP, Hadoop cluster on CentOS, Data analysis
③ Using PHP/Java to process log data
④ Training Java/Go/Python/PHP/Linux/SQL

WORK EXPERIENCE

2017. 07 – 2020. 11
Beijing Tsinghua Tongheng Urban Planning & Design Institute

Thupdi is an architectural practice based in Beijing, China

① Facilitated insightful daily analyses of 50GB of website data, enterprise data, spatial data(using arcgis) and mobile data collect by external sources, installing and configuring Hadoop ecosystem components that move from individual servers to HDFS
② Apply the latest technologies and tools to implement complex Big Data platform Apache/Hortonworks/Cloudera
③ Created Hive queries that helped analysts Automated ETL processes (Nifi, Streamsets,oozie), making it easier to data
④ Worked with Sqoop to ingest & retrieve data from various RDBMS like Oracle DB & MySQL
⑤ Written Hive/impala/kudu/clickhouse/Presto queries to transform data for further downstream processing

PROFESSIONAL EXPERENCE

① Around 4 years of IT experience in dealing with **Hadoop** Ecosystem like HDFS, MapReduce, Hive, presto, Sqoop , HBase, Spark, Scala, Nifi/Streamsets ,Hue, Flume, oozie, Azkaban, Flink, **ClickHouse**

② • Excellent understanding of Hadoop Architecture and underlying Hadoop framework including Storage Management.
③ • Hands on experience in installing, configuring, and using Hadoop (Apache, Ambari, **CDH**) components like Hadoop, Map Reduce, **TEZ**, HDFS, HBase, Hive, **impala/kudu** Sqoop and Flume, Kafka, Pyspark
④ • Managed data coming from different sources and involved in HDFS maintenance and loading of structured and unstructured data.
⑤ • Experience in analyzing data using HiveSQL and custom **MapReduce** programs in Java.
⑥ • Experience in importing and exporting data using **Sqoop(**HDFS <>RDBMS),
     DataX**(**RDBMS<>Hadoop),
     waterdrop(Hive2Clickhouse),
     Streamsets (MySQL2Kudu**)** ,
     Nifi(data2Kafka)
⑦ • Experience in working with NoSQL databases like HBase and MongoDB(pyspark+ MongoDB).
⑧ • SQL-based technologies(e.g. MySQL/HiveSQL/Oracle/SQL Server/GreenPlum/Clickhouse)
⑨ • Hands on experience in **Linux**(**Centos7**/Redhat/ubuntu). Worked with Big Data distributions Cloudera.
⑩ • Expert in writing complicated **SQL** Queries and database analysis for good performance.
⑪ •Expert in Building, Deploying and Maintaining the Applications visualization tool like tableau, **Superset**

self

Excellent analytical, Interpersonal and Communication skills, fast learner, **hardworking** and good team player.

I have a legal work permit in the Netherlands

I can work more than 40 hours per week

Blog   https://blog.csdn.net/qq_33290422
LinkIn   https://www.linkedin.com/in/pengfei-qiao-a45a73178/

Clickhouse NoSQL
Pyspark Oracle Superset
QIAOPENGFEI Impala
Numpy Hadoop HBase
Azkaban Python Flink
NifiKudu
Streamsets
CentOS Kibana
Hue Kafka SQL
OOZIE Spark GreenPlum
ES CDH Sqoop