

OWEN RUMNEY

Email	owen@owenrumney.co.uk
Phone	07955 316 910
LinkedIn	www.linkedin.com/in/owenrumney
GitHub	github.com/owenrumney

PROFILE

Adaptable skill set	I maintain a high-level of competency across my existing skillset and possess an ability to assimilate new skills and knowledge; this enables me to be a valuable resource that can cover many aspects of a project. AWS Certified Architect, Developer and Big Data Specialist.
Creative approach	Tackling all problems, no matter how insurmountable they seem, with an open mind and measured approach, I am able to resolve issues in a way that ensures the most pragmatic solution is used within the constraints of the situation.
Delivery focused	I take a holistic approach to the delivery of technical solutions, ensuring that I fully understand business needs, I recognise the role of IT is a facilitative function. This means that my commitment to a project does not end when the testing cycle is complete.

TECHNICAL SKILLS

Languages	Strong: Java, Python, Bash Scripting, T-SQL Intermediate: Scala, PowerShell
Data and Storage	HBase, Hive, HDFS, S3, PostgreSQL, MySQL, MSSQL, ElasticSearch
Streaming	Kafka, AMQ, SQS, SNS, StreamSets, NiFi, Spark, PySpark, Samza, Storm, Flume
Platforms and Tooling	Amazon Web Services, Cloudera CDH, Hortonworks HDP, Docker, Minikube, ELK stack, Ansible, Cloud Formation, SBT, Maven, Guice, JUnit, Mockito, Scala Test, Nexus, SonarQube, Swagger, Flask, IIS, Tomcat
OS	Linux [RHEL, CentOS, Oracle Linux, Ubuntu], OSX, Windows
Methodologies	Agile/SCRUM/SAFE, Distributed Systems, Event Driven Development, Micro Services, Design Patterns, Mocking, REST Services, SOLID, TDD, Unit Testing, Source Control [ClearCase/Git/SVN/TFS], CI and CD [Jenkins/Travis]

WORK EXPERIENCE

Lead Platform Engineer | BP Downstream, Sunbury, UK *09/2018 – present*

Implementing a proof of concept data protection solution aiming to provide data tokenisation for the existing BP Data Lake. This system uses Privitar Publisher for tokenisation and a Cloudera CDH cluster to provide the compute power for the Privitar Spark job.

- Developed AWS Cloud Formation scripts to automate the build of the Cloudera cluster with an additional EC2 instance on which to run Privitar Publisher application
- Wrote Bash and Python bootstrap scripts to install Cloudera and configure TLS/SSL for secure inter node communication. Both components are fully integrated with LDAP and secured with Kerberos
- Worked with Privitar consultants and business to implement the data protection rules and policies for our use cases

Senior Data Engineer | 6.6 – Home Office, East Croydon, UK *08/2017 – 06/2018*

Returned to the project to develop a REST API to return associated entities from within HBase and develop the supporting stream processing application to link these entities. Following successful completion of this work I joined the three-man Data Ops team. We were tasked with the enhancement and operations of the Platform.

- Developed Samza stream processing application to identify associated entities on Kafka topics and writing the algorithm to store the links in HBase
- Supported the enhancement of Hortonworks HDP cluster in AWS. This work included everything from adding functionality for YARN application monitoring in Zabbix, aggregated logging to an ELK stack maintained by Data Ops Team and performance tuning of Kafka to the deployment of data application releases into the platform
- Developed a Hadoop in a box solution to run locally in Docker but could be scaled up to run in Kubernetes as a fully functional cluster. This is now being used for dev and test environments which is saving the project a generally estimated £250/month in always on EC2 costs

Platform Engineer | BP IST, Canary Wharf, UK**02/2017 – 08/2017**

Working in the Data Lake platform team within BP Integrated Supply and Trading segment; supporting the ongoing enhancement and day to day operations.

- Wrote the management tooling to maintain the Users and Roles. This was a PowerShell application which took an CSV file with the required entities and ran them under a role with delegated access to maintain the cluster's Active Directory Organisational Unit
- Developed the Ansible roles to deploy secure, kerberised, Kafka and StreamSets into the Cloudera CDH data lake.
- Implemented example StreamSets data flows to validate the product and supported a US based team in their adoption of the tool

Senior Data Engineer | 6.6 – Home Office, London, UK**03/2016 – 02/2017**

Working in a small Big Data team developing Tomcat hosted REST APIs to support the ingestions of entity change snippets to be saved into HBase and combined to form a complete entity. The microservice architecture meant the system was operating as eventually consistent, leading to a need to support the application teams using the data in their approach

- Developed robust Samza streaming applications to write CDC events from Kafka topics to HBase and rebuild the entity to be pushed to HBase indexes
- Developed new REST APIs for the ingestion CDC events from and AMQ/Storm process and surfacing the indexed entities
- Evaluated Open Source technologies for submit to the architectural review board for approval
- Technical lead for the eight-man Ingestion team accountable for ongoing enhancement of the Storm bolts restructuring data coming out of Postgres CDC and pushing into Tomcat REST APIs. Working with the wider SAFE release train in prioritising work and ensuring the other teams' dependency on us was managed

Senior Data Engineer | comparethemarket.com, Peterborough, UK**04/2015 – 03/2016**

As a member of the BI team I was the only person with Java and Scala programming experience so took on the code development tasks of the team. My experience deploying Cloudera CDH and Amazon Web Services lead to me taking ownership of the migration of the cluster from on-premises to AWS hosted

- Designed the AWS cluster in conjunction with the central Architecture team. As the licences had already been purchased this was essentially a lift and shift process
- Implemented the ETL solution for getting the website and mobile app's events into HDFS. This solution used Camus to consume from Kafka topics and route the events using custom partitioners
- Developed a large number of Hive and Pig scripts, most notably the ETL process behind the Meerkat Movies campaign
- Wrote a Spark application for bridging the gap between website quotes and real partner sales. This is critical to determine sales that have originated from comparethemarket.com. Basic fuzzy logic was required to accommodate user's tendency to use false varying information for comparethemarket.com and their chosen insurer

Big Data Engineer | BP IST, Canary Wharf, UK**04/2013 – 04/2015**

Realised the first production installation of Hadoop in BP – the system tracked Market Data usage by processing application and server logs to ensure it was correctly paid for; this was to mitigate the risk of fines similar to those incurred by other banks in the region of \$80m

Previous roles – more information available on request**05/1999 – 04/2013**

SharePoint Developer - Shell, London, UK

Framework Developer - BP IST, Canary Wharf, UK

Front Office Developer - UBS, London, UK

SharePoint Developer - IBM (DEFRA Project), London, UK

Analyst Developer - Red Bull Racing (F1), Milton Keynes, UK

Solution Developer - MBRT, Milton Keynes, UK

Solution Developer - Paradise Computing, Northampton, UK

Electronic Engineer - Royal Air Force, UK