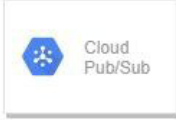

















Data Engineering On Cloud

Vineet Jaiswal

Category	AWS	Azure	Google Cloud	Open Source
Real time data stream for real time applications	 Kinesis Family	 Events Hubs	 Cloud Pub/Sub	 Kafka
NoSQL	 Dynamo DB	 Cosmos DB	  Cloud Firestore Cloud Bigtable	  Mongo DB Cassandra
Data Warehouse	 Redshift	 Azure Synapse Analytics	 BigQuery	  SQLite PostgreSQL
Managed Hadoop/Spark service	 EMR	  DataBricks HDInsight	 Cloud Dataproc	  Hadoop Hops Apache Beam
Extract Transform and Load (ETL)	 Glue	 Data Catalog	 Cloud Data Catalog	  Splunk Apache Spark
Data orchestration / ETL	  Data Pipeline AWS Step Functions	 Data Factories	 Cloud Dataflow	 Apache Airflow
Analytics	  Kinesis Analysis Athena	   Data Lake Analysis Data lake Store Stream Analysis	  BigQuery Cloud Dataflow	  Splunk Flink
Relational Database	 RDS	 SQL Server	 Cloud SQL	   MongoDB MySQL SQL Server Express
Cache	 Elastic Cache	 Cache for Redis	 Cloud Memorystore	 Redis