OSKR LLC

Posted: 08/09/2023
OSKR LLC is Emeryville, California based Limited Liability Firm which assists you with Economic analysis for complex litigation. OSKR LLC have been serving growing companies since 2020.

CUSTOMER

OSKR

PRODUCT & SERVICES
- Azure DataLake
- Azure DataBricks
- Python
- PySpark
- T-SQL

INDUSTRY
Professional Services

ORGANIZATION SIZE
Over 100 employees

COUNTRY
USA

BUSINESS NEEDS
OSKR LLC needed assistance with complex analysis involving data to Petabyte scale.

SOLUTION
G5 Data & Azure Architect designed & implemented complex solution to perform following:
- Transfer data from one Azure DataLake to another with the ability to continue where the copy operation left off.
- Custom developed code while being performant was able to transfer terabytes of compressed data from one DataLake to another with great performance.
- Developed custom code to extract, transform and analyze data. This data was in-turn used to generate complex reports used in litigation. Some of the salient feature of this implementation was:
  o Infusing data with custom values
  o Decompressing and using Petabytes of data to extract required information.
  o Setup DataBricks environment backed by Microsoft Azure running 100 worker Node to process the data.
  o Optimized and high performant custom-built code to perform feature extraction.
  o Data processing based on custom timeline and schedule.
  o Setup execution environment in under 10 minutes and execute code to process terabytes of weekly data in hours.

BENEFITS
1. Portable and re-usable code. This solution can be reconstructed in another environment with very little modification.
2. Pay for use. Environment was setup to scale up and scale down as per demand saving cost.
3. Process terabytes scale of data in hours

RESULT: HIGHLY PERFORMANT & COST-EFFECTIVE SOLUTION TO PROCESS & ANALYZE PETABYTE SCALE OF DATA