**Gaurav Bansal**

H.No 72, Sector 26, Noida (U.P), 201301

Mobile: +91-9818690858 ~ Email: gaurav3ansal@gmail.com

**PROFESSIONAL SNAPSHOT**

**•** Having **16+ years** of experience in complete Software Life Cycle in Agile environment including Business Analysis, Architecture & Design, Development, Testing, Technical Support & Solution, Process Automation, Team, Project & Service Management and Client servicing abilities.

**• Big Data Advanced Analytics** - Experienced in working with large datasets using advanced data analysis and processing frameworks.

* **Hadoop, Spark** and **Beam** Big Data Ecosystems

**• Data Modeling** - Experienced in using statistical tools & techniques and numerical modeling with real-world data.

**•** Experienced in providing recommendations & guidance for client’s technology approach across the IaaS, IaC, PaaS, SaaS, Digital Platform and Cloud Native landscape. Incorporating:

* Architecture and Design of multi-cloud infrastructure, digital platform, SaaS and Cloud Native environments.
* Cloud Native technology recommendations.
* Assessment and recommendation of SaaS services.
* Creation of technology roadmaps.
* Application readiness assessments.
* Migration & Modernization
  + On-prem data warehouses to cloud Big Data Analytical services.
  + Monolithic Web applications (dotnet, java, python) to Microservices architecture.

**•** Experienced in defining architecture roadmaps and journeys for the clients and in responding to RFPs and SoWs.

**•** Experienced in designing, creation and curation of Cloud Reference architectures, blueprints, reusable framework/components to accelerate development on projects.

**•** Experienced in maximizing the productivity of the team by establishing programming, design and other technical & coding standards. Followed best practices for operational efficiency and cost reduction with Cloud technologies.

|  |  |
| --- | --- |
| Programming & Scripting Languages | Core Java, Python / Exposure to RBash, JavaScript |
| Platforms/ Frameworks | Apache Beam, Hadoop, Spark Apache Atlas (Data Governance and Metadata) and Apache Ranger (Data Security) Apache Griffin (Big Data Quality Solution for Batch and Streaming) Google Cloud DLP (Sensitive data inspection, classification, and de-identification platform) Sisense (Business Intelligence, Discovery and Analysis), Google Data Studio |
| Databases & Data Warehouses | Relational Database: MSSQL, MySQL, PostgreSQL Data Warehouse: Hive, Redshift, BigQuery |
| Machine Learning Libs | BigQuery ML, Redshift ML |
| Tools | Workflow Scheduler System: Airflow/Composer Data Management Tool: NiFi, Sqoop Collaborative R[evision Control](http://en.wikipedia.org/wiki/Revision_control): Bitbucket & SourceTree, Google Source Repositories, GitLab/GitHub [Project Management Software](http://en.wikipedia.org/wiki/Project_management_software)/Bug Tracking System: JIRA |

**WORK EXPERIENCE**

EXL Service Pvt. Ltd.: January 2021 - Present

Telus International India: August 2018 – January 2021

PST Analytics: April 2017 – August 2018

FIS Global Business Solutions India Pvt. Ltd.: January 2017 – April 2017

Omnicom Media Group India Pvt. Ltd.: July 2014 –January 2017Samsung R&D Institute India – Noida: September 2011 – July 2014  
HCL Technologies Pvt. Ltd.: August 2009 – Septembe 2011  
R Systems International Ltd.: February 2007 – July 2009

**PROJECTS HANDLED**

**Client: EXL Service (India)**

* Westcor Land Title Insurance Company (WLTIC) / Xtrakto– Cloud Architect
* Insurance Australia Group (IAG) / Smart Data Signal – Cloud Architect
* Hollard / Smart Data Signal – Cloud Architect
* Tokio Marine Kiln (TMK) / Smart Data Signal – Cloud Architect
* CUNA Mutual Finance Group (CMFG) – Cloud Architect
* Liberty Mutual Re-insurance – Cloud Architect

|  |  |  |
| --- | --- | --- |
| Synopsis | : | Westcor Land Title Insurance Company (WLTIC) |
| Responsibility | : | * Created secure, scalable, multi-account hybrid environments that implements the security baseline compliant with customer standards. * Hybrid and multi-cloud deployments with centralized and distributed network designs with network and application-level firewall configurations. * Implemented further protections for data with IPSec tunnels and managed SSL certificates. * Architecture & Design of multi-cloud infrastructure: * Tools used: Global Accelerator, VPC, Route 53, Endpoints, Peering Connections, AWS Backup, EFS, S3, S3 Glacier * Integration of hybrid cloud networks: * Tools used: Gateways (Internet, NAT, Virtual Private, Customer, Transit), Site-to-Site VPN connections, * Implementation of enterprise data warehouse and data lakes for insights & reporting: * Tools used: Athena, Redshift * CICD pipelines: * Tools used: AWS Developer Tools (Cloud9, CodeArtifact, CodeBuild, CodeCommit, CodeDeploy, CodePipeline, CodeGuru) * ETL/ELT pipelines: Experienced in working with large datasets using advanced data analysis. * Tools used: Redshift, S3 * Data & Database migrations (Homo & Heterogeneous): * Tools used: AWS DMS, Debezium, DataSync, AWS Transfer Family * Data modelling: Experienced in using statistical tools & techniques and numerical modelling with real-world data. * Framework used: Data Vault 2.0 * Cloud computing: * Tools used: Lambda, Textract, Comprehend, Transcribe, Sagemaker, EC2, Elastic Container Registry, SNS, SQS, EventBridge * Management & Governance: * Tools used: CloudFormation, CloudTrail, CloudWatch, Config, Systems Manager, * Security, Identity and Compliance: * Tools used: Certificate Manager, Firewall Manager, IAM, Inspector, Key Management Service, Network Firewall, Secrets Manager, Security Hub, WAF, |
| Frameworks & Languages | : | Python 3x |
| Data Warehouse | : | Redshift |
| Database | : | MySQL, MSSQL |

**Client: Telus (Canada)**

* Telus International Data Lake (TIDL) – Big Data AA
* Edmonton Global Data Lake (EGDL) – Big Data AA
* Simplot Grower Solutions Data Lake (SGSDL) – Big Data AA
* SickKick Data Commons (SKDC) – Big Data AA
* Telus Cattle Trace System (TCTS) – Cloud Migration Microservices Architecture
* Toronto Star Corporation (Torstar) – Cloud Migration Hybrid

|  |  |  |
| --- | --- | --- |
| Synopsis | : | **TIDL** – is an enterprise level data analytics platform used to perform advanced analytics on historical data (Batch and Streaming) for HRM and WFM systems.  **EGDL** –Data Analytics platform for Data n’ Research for Economic Development.  **SGSDL** **–** a scalable and operationally flexible data analytics platform to store and access data including SGS agronomic data, financial, operational, and strategic solutions data.  **SKDC** – An intelligent data analytics platform which enables access to distributed data assets. The platform describes the data itself, as well as the data ownership, access, and usage rights. |
| Responsibility | : | * **Creation** of Security Configuration benchmark intended for system and application administrators, security specialists, auditors, platform deployment, and DevOps personnel. * **Organizing** Data and Analytics Kickoff (**DAK**) workshops and technical deep dive sessions. * **Creation** ofTechnical Design Documents (**TDD**) with summary of the cloud data and analytics elements, including current architectural and data engineering plans, identified technical solutions, and technical recommendations. * **Setting** up and customizing Data Provenance and Governance (Discovery, Lineage and Security) using NiFi, **Atlas** and **Ranger** for Industry and regulatory compliance (PII, PCI, HIPAA, etc.) * **Setting** up a CI/CD pipeline for data-processing workflows using **Cloud Build**, Source Repositories and Composer on Google **Kubernetes** Engine using **Docker** and **Terraform**. * **Setting** up and customizing Big Data Quality Solution using Apache **Griffin** for Data Quality control and usage. * **Integration** of Sisense and Data Studio with BigQuery, Google Sheets & Drive APIs using **Cloud Functions**, **Workflows.** * **Design/Development** of data pipelines using NiFi/Airflow/PubSub/Beam/Dataflow/Sqoop/Spark/Dataproc/Cloud DLP/Cloud Functions/BigQuery Streaming APIs/Python/Pandas. * **Analysis** of Avro, Parquet and NLD JSON formatted nested data using BigQuery (partitioning, clustering, and the usage of nested fields). * **Orchestration** of workflows using Cloud **Composer**/Airflow. * **Optimization** of data pipelines using Cloud storage and efficient data compression file formats such as Avro and Parquet. |
| Frameworks & Languages | : | Avro, NiFi, Atlas, Ranger, Airflow, Griffin, Dataflow, Pub/Sub, Dataproc, Sqoop, Sisense, Pandas, DLP, Spark Python 3x |
| Data Warehouse | : | BigQuery |
| Database | : | MySQL, MSSQL |

|  |  |  |
| --- | --- | --- |
| Synopsis | : | **TCTS** –Tracking cattle for the certified sustainable beef framework as defined by CRSB, helping the client meet requirements of Certified Sustainable beef by different Retailers and consumers.  **Torstar** -migrated Sales applications should be able to perform standardization and automation (high value repetitive human tasks in the Sales Process) functionalities in the Sales cycle right from the Layer 1 till final campaign reporting. |
| Responsibility | : | * **Due diligence** processes (Business Systems, Proprietary Technology, Infrastructure, Organization & Process, Data Analytics, Security & Compliance) for Application support, Cloud services, Cloud support/maintenance, DBA, Platforms, Risk & cost exposure from security deficiencies and non-compliance. * **High-level phase wise scope estimation** (Determine and Discover/Migration Planning and Application Portfolio/Migration Execution) * **Managed** Cloud Security through entire information processing lifecycle including hardware infrastructure, service deployment (service account credentials management), user identity, storage services and operational security. |
| Frameworks & Languages | : | Dotnet, Flask, Node  PHP, Node.js, Java, Python 2x/3x |
| Databases | : | MySQL, MSSQL, PostgreSQL |

**Client: PST Analytics**

* **IoTDP (IoT Data Predictor) –** Big Data Advanced Analytics & ML and Cloud Computing Senior Consultant

|  |  |  |
| --- | --- | --- |
| Synopsis | : | **IOTDP –** is a scalable data analyzer from sources such as IoT devices & sensors. Integrated cloud-based solutions to track unused IoT sensors capacity. |
| Responsibility | : | * **Design/Development** of data pipelines using serverless architecture (AWS Lambda and Google Cloud Functions and BigQuery) on hybrid cloud network securing and restricting cloud APIs with VPN Tunneling. * **Design/Development** of event-driven data pipeline solutions using Cloud Pub/Sub and data feed automations using Cloud Logging & Dataflow. * **Setting** up container management platform for multi-architecture container images for different IoT devices runtime environments. * **Management** of Infrastructure as Code with Terraform, Cloud Build and Gitops. |
| Frameworks & Languages | : | Hadoop/Spark/Beam Python 3x, Scala |
| Data Warehouse | : | Hive, BigQuery |
| Database | : | MySQL |

**Client: FIS Global (United States)**

* **Scalable Data Platform –** Senior Technical Lead I Big Data & Machine Learning Solutions

|  |  |  |
| --- | --- | --- |
| Synopsis | : | **Scalable Data Platform –** is a highly scalable analytics and machine learning platform used to get insights out of various banking and financial data. |
| Responsibility | : | • **Involved in:**   * Agile Development and Testing practices. * Data Pipeline architecture and development using Apache Beam and AWS Kinesis Firehose. * Serverless Lambda architecture design layers using AWS Lambda with Python. |
| Frameworks & Languages | : | Spark, Beam Python 2x |
| Data Warehouse | : | Hive, Redshift |
| Database | : | MySQL |

**Client: Omnicom Media Group (United States)**

* **Agile Data Xplorer –** Technical Group Lead Big Data Solutions

|  |  |  |
| --- | --- | --- |
| Synopsis | : | **Agile Data Xplorer –** is a cross-channel analytical engine that allows users to explore very large data sets to provide clients with data-driven insights to increase the overall customer satisfaction and revenue. |
| Responsibility | : | • **Involved in:**   * Software Architecture Document, High & Low Level Design document, Functional requirements for the project. * Data Pipelines development using Python/Pandas * Integration of **Selenium** Test Automation Framework to perform data ingestion process automation and validation of large and complex datasets on Flask based Web Application. |
| Frameworks & Languages | : | MapReduce, Selenium, Flask Python 2x |
| Data Warehouse | : | Hive, Redshift |
| Database | : | MySQL, MSSQL |

**PROFESSIONAL QUALIFICATION**

|  |  |  |  |
| --- | --- | --- | --- |
| **Examination** | **Year** | **Board/University** | **%** |
| M.C.A | 2007 | Guru Nanak Dev University, Amritsar (Punjab) | 68 |

**PERSONAL DETAILS**

|  |  |
| --- | --- |
| PERMANENT ADDRESS | Nahan, H.P |
| MARITAL STATUS |  |
| NATIONALITY | Indian |
| LANGUAGES | Hindi, English |
| DOB | 20-01-1983 |
| PASSPORT NO. (VALID UPTO) | N7219614 (31-03-2026) |
| VISA TYPE (VALID UPTO) |  |