Alejandro D. Ramos

Data Engineer

Buenos Aires. Argentina ale3385@gmail.com
+54911-3465-3385
US Visa

SUMMARY

Hi! I'm a designer, builder and manager of Big Data infrastructures. I'm a collaborative engineer with 17 years of experience designing and executing solutions for complex business problems involving large scale data warehousing, real-time analytics, and reporting solutions. Known for using the right tools when and where they make sense, as well as creating an intuitive architecture that helps organizations effectively analyze and process terabytes of structured and unstructured data.

TECHNICAL SKILLS SUMMARY

- Operating Systems: Windows, Linux and UNIX.
- Programming Languages: Oracle PL/SQL, Hyperion Oracle, R, .NET, Java, HTML, SQL Server, Python and Data science libraries: TensorFlow, SciPy, etc.
- Databases: Spark, Oracle, DB2, MongoDB, MS SQL Server, MS Access, Oracle, MySQL, PostgreSQL, Progress. Hadoop: HBase, Cassandra, etc.
- Applications: TOAD, Erwin, Star Team, Remedy, MS Office, MS Project, MS SharePoint, Visio, Oracle Financial, Oracle PeopleSoft*, Oracle JD-Edwards*,SQL Server Management Studio, Analysis Services, Reporting Services, Oracle Workbench, Oracle RMAN, Oracle Data-guard, SQLPLUS, Oracle SQL Developer. Oracle Exadata Machine.
- Hyperion, Web-logic, OBI, ODI, OAS. Exalogic, EBS, Golden Gate.
- Cloud: Verizon cloud; Amazon AWS: S3, EC2, EBS, RDS; Terremark DC.; Windows Azure; IBM Soft-layer; Oracle Cloud: SAAS, PAAS, IAAS; IBM Bluemix. Google Cloud.
- Hadoop distribution: Cloudera, Apache Hadoop, Map R. Main Tools: Hive, Pig and Hadoop Streaming, Map-reduce, R, SPSS, SAS, Python, MATLAB, CUDA Dev NVIDIA.
- Cloud: AWS EC2, S3, ECS. Docker, Ansible, Terraform. Puppet.
- Linux: Wide variety of distributions as internet servers, file servers, etc.

EXPERIENCE

Consumer Affairs - Data Engineer - *US (Remote)*
Location: Tulsa, Oklahoma
June 2021 - Present

Summary:
As a Data Engineer specialized in Big Data at Consumer Affairs, I have spearheaded and contributed to pivotal projects aimed at managing and harnessing large-volume data. I've worked with an array of technologies, tools, and cloud services to enhance our data analytics capabilities.

Key Achievements and Responsibilities:

- Data Architecture: Designed and managed the data infrastructure to support Consumer Affairs, leveraging technologies such as Databricks, AWS Glue, Redshift, and Snowflake.
- ETL Processes: Developed and maintained robust ETL pipelines using tools like Apache Airflow, AWS Glue, and DBT in PostgreSQL, Redshift, and Snowflake environments.
- Migration to Snowflake: Successfully led the migration of the data warehouse from Redshift to Snowflake, optimizing the performance and scalability of our data platform.
- ELT Tools: Implemented ELT tools like Fivetran and Stitch to automate data loading from various sources, streamlining the Extract, Load, Transform (ELT) process.
- Collaboration with Data Science: Collaborated closely with Data Science teams to implement Machine Learning models in production using Databricks, enabling data-driven insights and recommendations.
- Data Security and Compliance: Ensured data security and compliance with relevant regulations in data handling, especially concerning consumer data.

Key Projects:

- Migration to Snowflake: Describe in detail how you successfully led the migration from Redshift to Snowflake, highlighting the benefits achieved.
- ETL Automation: Explain how you implemented automated ETL solutions using tools like AWS Glue to enhance data processing efficiency.
- Technologies and Tools Used:
- List the specific technologies, frameworks, databases, and tools you utilized during your tenure at Consumer Affairs, including AWS Glue, Redshift, Snowflake, GCP with BigQuery and more. Azure platform with Databricks and Data Factory.

***Big Data Analytics Engineer*, Bridge Corp** - *US (Remote)*

SEP 2019 – Jun 221

- Big data architecture on AWS: We extract Petabytes of data every month from external data sources from both, structured and unstructured data types.
- We use batch and streaming data processes: For batch → AWS S3 → Glue + Athena → Redshift.

  For streaming we use Kafka: AWS Lambda / AWS Step Functions → AWS MSK → AWS Kinesis → S3 → Snowflake.
- On the Data Warehouse: We use Redshift and started migration to Snowflake where we've mainly Star Schema model type and using separated databases that defines the Landing, Stage and Final zones. On the last one we store our Fact-tables and Dimensions.
- We also implemented Apache Airflow for DAG's automation that connects data from other services on AWS like Redshift – Kafka - Athena. There we store ours ETL using python, Spark, and Scala languages.
- ML Modeling: Using AWS EMR Notebooks→ time series analysis (with correlograms) to detect and assess intensity of change points or new trends in KPI's, and match them with events (such as redefining user, which impacts user count). Click fraud and Botnet detection. Taxonomy creation.
- CI/CD: We mainly use python with Alembic (SQL-Alchemy library) for SQL data migrations, and Circle CI + GitHub.
- For C-Level reports we use Power BI as main data visualization tool and started some others with Looker.

AWS (Redshift, Lambda, Athena, S3, DynamoDB, Glue), PostgreSQL, Spectrum, Snowflake Platform, Jupyter Notebook, Python 3.8, Pyspark, Scala, Apache Airflow, Kafka, Power Bi, GitHub, Jira) Azure platform : Power BI, Data factory, azure devops.

***Data Scientist with ML*, Chevron** - *US (Remote)*

NOV 2017 - AUGUST 2019

- Builder of Power BI and Qlik View reports from heterogeneous data sources.
- Mentoring of Data Science courses for internal employees.
- Creation Machine learning algorithms for reports and recommendation systems.
- Microsoft Azure architecture design for data stores and data lakes and Hybrid stores from on premises data stores.
- Currently using machine learning with python on azure and Anaconda tool for data visualization reports such as Power BI, Qlik View and Spotfire (legacy). Using GitHub and Azure DevOps as a repository.
- The main thing here is to make use of "What if" for predictive reports and dynamic dashboards reports.
- Clouds: AWS, Azure and GCP. (Depending on the business line)

Microsoft Azure, Power BI, Qlik View, GitHub, Azure DevOps, ML, Python,SQL.

***Big Data Engineer | Cloud Architect*, Oracle** - *US (Remote)*

NOV 2015 - OCT 2017

- By using the Oracle Big Data Appliance, I have created tons of reports with Python and SciPy, I have also created Machine learning algorithms for cases of study on “What if” forecast dashboards.
- Senior Principal Advanced Data Engineer for customer DIRECTV.
- Reviews customer incidents to ascertain cause, impact to the customer environment, remediation activities and identify continuous service improvements.
- Collects, creates, and provides regular reporting to the customer describing service results, value add, recommendations, and issue resolution.
- Works with the customer and internal Oracle team to ensure stability of connectivity between the customer data center and Oracle NOC.
- Monitors and tunes the system to achieve optimum performance levels in standalone and multi-tiered environments.
- Conducts system analysis, configuration management and develops improvements for system software performance, availability and reliability.
- Perform Incident resolution, Problem Determination and Root Cause Analysis in accordance with Service Level Agreements.
- Designs, develops, recommends and implements new or revised system software, utilities and automated processes as necessary.
- Provides technical expertise for system Transitions, Migrations and Consolidations.
- Oracle Cloud.
- Tools: Designed a large data warehouse using star schema, flow-flake.
- Designed and developed a Big Data analytics platform for processing customer viewing preferences and social media comments using Java, Hadoop, Hive and Pig.
- Integrated Hadoop into traditional ETL, accelerating the extraction, transformation, and loading of massive structured and unstructured data.
- Loaded the aggregate data into a relational database for reporting, and ad-hoc analyses, which revealed ways to lower operating costs and offset the rising cost of programming.

Hive, Pig and Hadoop Streaming, Map-reduce, R, SPSS, SAS, Weka, MATLAB.

***Sr. Data Lead*, IBM** - *US (Remote)*

JUN 2014 - OCT 2015

- Data leader at IBM with a team of 12 DBA’s from EEUU, India, Brazil, China and Egypt.
- We are responsible for Operation and Support of corporate financial applications for the E-pricer Project at IBM.
- Production supports readiness. Implement work and quality processes, Agile engineering methodologies and processes for Production Support services.
- Manage financial, material and human resources. Request resources and ensure they are in a timely manner.

- Ensure and maintain the daily operations of financial systems, meeting defined Service Levels Agreements.
- Collaboration with groups along all the organizations, to attend services requests or resolve problems.
- Initial implementation of work processes; ITIL engineering methodologies and processes for Maintenance, Requirement definition, Analysis and Design, Quality Assurance, Testing and Application Implementation. Additionally, configuration management and change control process.
- Assist the team members to ensure the proper implementation of internal methodology.
- Provide quality assurance for the services requested; Review functional specifications, technical specifications, unit test and handover. Ensure compliance with local and global standards into IBM standard.
- Create strategies and solutions to improve the daily operation.
- Consistently reduced operating costs.
- Support definition of new solutions and business proposals. Support for Project Estimating.
- Consolidate financial information for the collection of HP services to customers.

Main achievements:

- I performed the transition of IT services from IBM India to IBM local IT team. A smooth transition was achieved without interruptions while maintaining service levels expected by the client.
- I started the implementation of IBM Agile processes and working methods on the applications area. This will have established a standard work process under best practices in the IT industry.
- I assured the correct day to day operation of IBM applications fulfilling SLAs.

Stack, Tools & Environment: *AWS Platform, Z-OS, IBM Cloud, Redhat, Python, Bash.*

***Data Architect*, IT Resources** - *Argentina*

Jun 2011 - MAY 2014

- IT Data Architect for Comafi Bank project, responsible for Operation and Support of corporate financial applications for Comafi. Oracle RAC Implementation and tuning. Microsoft Dynamics AX ERP with SQL Server 2012.
- Lead the transition of IT services from IBM to Comafi IT Team.
- Production supports readiness. Implement work and quality processes, Bank engineering methodologies and processes for Production Support services.
- Configured SQL database to store Hive metadata.
- I performed the transition of IT services from IBM to the Comafi local IT team. A smooth transition was achieved without interruptions while maintaining service levels expected by the client.
- I started the implementation of Bank standard processes and working methods on the applications area.
- I assured the correct day to day operation of Comafi Bank applications fulfilling SLAs.

Stack, Tools & Environment: *Oracle Database 11g Cluster, SQL Server database 2012r2, Microsoft*

Dynamics AX, Intranet Integrity / Treasury, Control-M, Microsoft SharePoint, Cent-os Linux, Java, Windows Server 2012, SUN Solaris, Hadoop, MongoDB

***DBA*, Celistics Holdings** - *Argentina*

Feb 2010 - Jun 2011

- Responsible for Development and Production Support areas for Business Intelligence with technologies such as: SQL Server , MySQL , Oracle, Microsoft Dynamics AX.
- Ensure and maintain the daily operations of Microsoft Dynamics applications in production environments.
- Technical Leader, OLAP Database designer and developer.
- Responsible for the definition and design of Microsoft Dynamics application solutions for user's requirements, based on OLAP and DWH concepts. Production of solution designs and optimization.
- Implementation of work processes: ITIL engineering methodologies and processes for Maintenance and new Development, Requirement Definition, Analysis and Design, Production, Quality Assurance, Testing, Implementation and Installation.
- Definition of new solutions and business proposals. Projects Estimating.
- Execute Benchmarking and product evaluation: Microsoft Dynamics, Oracle, SAP R3, Hyperion products (Analyzer, Reports, Balance Score Card)

Stack, Tools & Environment: Hyperion Oracle. Microsoft Dynamics Brio - Hyperion Intelligence MYSQL. MS Access. MS SQL Server. UNIX, W2K, WXP, WNT

***DBA*, Gilbarco Latin America SRL** - *Argentina*

March 2005- Jan 2010

- Responsible for Development for Business Intelligence reports for the top executives of the main Gas and Oil companies in Latinamerica (Exxon mobil, Terpel, YPF , Shell, Petrobras). Technical Leader, OLAP designer and developer.
- Store Procedure development of the card Serviclub rewards systems, a software solution that rewards customers for their buys.
- Creation of models for the KPI – Key Performance Indicators for Gilbarco.
- Database administration of SQL Server - MySQL - Oracle engines.

Stack, Tools & Environment: SQL Server, MySQL, Oracle. Unix and Linux. Windows Server 20002005.

EDUCATION

Stanford University of California. US - M.S. in Statistics for Data Science 2017-2018.

Universidad Abierta Interamericana. Argentina - Master in Computer Science 2013-2015.

Universidad Tecnológica Nacional. Argentina - Bachelor in Computer Science 2006- 2011.

Luis A. Huergo High School. Argentina - Electrical Technician Engineer 2000-2005

CERTIFICATIONS AND COURSES

- GSMS, Quality Assurance, ITIL, CMMi, ISO – 9000 Gilbarco / IBM. 2006
- Gas and Oil Industry Training Gilbarco. 2007
- **Microsoft SQL Server MCDBA**: Educacion IT BA. 2007
- **MySQL DBA**: Educacion IT BA. 2008
- **ORACLE DBA**: Educacion IT BA - Proydesa Institute. 2010
- **Big Data: Hadoop Oracle Certified** (Oracle 2017)
- **Linux Administrator**: Educacion IT BA. 2010
- **Cloud Computing**: ITBA. 2014
- **Wallstreet Institute: Advanced Language English**. Advanced Technical English. 20012016
- **PMP**: Education IT BA. 2015
- **SCRUM MASTER**: UTN BA. 2016
- GCP: **Google Certified Professional Cloud Architect** 2018.
- AWS: **AWS Certified Solution Architect – Professional** 2015.
- AWS: **AWS Certified Big Data – Specialty – Professional** 2018.