e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 8, August 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.521

© 2024 IJMRSET | Volume 7, Issue 8, August 2024| DOI:10.15680/IJMRSET.2024.0708081

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Streamlining ROP Metrics and Reporting through Cloud Migration and Automation

Venkat Sunil Kumar Indurthy

Software Developer, Compunnel Software Group Inc., USA

ABSTRACT: This document outlines an innovative methodology for modernizing ETL (Extract, Transform, Load) pipelines by embracing the use of automation, cloud migration methods, performance monitoring, and data architecture with respect to business requirements. The document describes how scalable cloud-native ETL operations using Informatica PowerCenter and Azure Data Factory can replace outdated SSIS packages, under the automation interface of Control-M. The document discusses how to support high-quality data delivery in a timely manner with structured ETL testing methodology and CI/CD processes using Jenkins and GitHub. The next part of the document recommends the benefits of using Power BI for the purpose of analytics, reporting and visualization in a data enabled decision-making process. Effective collaboration with global teams is established to successfully deliver using agile and scrum project management methodologies. Key performance indicators such as throughput, latency, error rates and resource utilization provide insight to tunnel health and where there could be improvements. The paper outlines a general framework for creating dependable and intelligent data pipelines in a volatile, data-centered economic context, and outlines trending on the horizon trends, such as AI driven automation, real-time streaming ETL models and advances in engineering and governance.

KEYWORDS: ETL, Informatica PowerCenter, Azure Data Factory, CI/CD Processes, Jenkins, GitHub

I. INTRODUCTION

To achieve the best performance, companies today need to manage their operations effectively and monitor and control the supply chain. Supply chains consist of interrelated procurement, production, warehousing, shipping, and delivery operations across multiple partner companies and locations. A company can continually monitor the supply chain to obtain and maintain a real-time and operational perspective on the supply chain; this monitoring will allow the company to identify inefficiencies and disruptions that may disrupt both service level and/or cost. Advanced monitoring systems that utilize key performance indicators allow procurement, production, distribution, and customer service to assess the efficiency of the supply chain to enhance efficiency and reduce lead times, maximizing efficiency while minimizing waste and inventory costs while helping to prevent stock outs. Real-time tracking of performance also allows the supply chain to give faster responses to unexpected challenges, strengthening the resilience of the overall supply chain to deliver customer satisfaction. This data driven process will allow continued improvement and ensure alignment to the overall business as it is trying to manage cost and resources and competing business in the market place. Effective monitoring of a supply chain will be critical to continue to grow and achieve operational excellence in terms of quality service and being cost efficient [1].

In a supply chain the monitoring of the following educational topics are key areas of focus: sorting and procurement of materials, Labor productivity, manufacturing efficiency, construction and distribution of materials, performing inventory management, transportation time and delivery of products to stock, and customer service and satisfaction. Each of these areas can work together to streamline the procurement or resources and receive and sort the inventory before d/c or distribution. Operational efficiency of these areas of procurement, production, warehousing, and customer service would use a similar processes to improve customer service and effective delivery. Overall, organizing the sequence in which those items could be organized or tracked could help with efficinecy and monitoring.To deal with these pressing issues, businesses can enhance performance and lower risk to support enhanced decision-making and total visibility of the supply chain. [2]

Metrics are vital in managing a supply chain by recognizing failures and areas that need attention. Metrics provide an objective basis for easily assessing performance across all supply chain activities, including order accuracy, lead-time,

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

inventory turns, and cost effectiveness. Visibility can identify bottlenecks and some variances to output that affect productivity. Metrics allows you to track problems back through the chain to identify root causes of poor process - supplier delays, production - versus simply identifying symptoms. Metrics can also facilitate continuous improvement by establishing primary objectives, tracking results as improvements are achieved, and then direct resource allocation and decision-making. Ultimately, metrics provide the foundation for a data-driven approach that is critical for achieving operational excellence in the long-term and enhancing supply chain productivity. [3]

Effective supply chain management finds a balance between a strategy to control costs and increase efficiency and the provision of a service quality. Successful balancing will address customer needs of time to deliver and product availability without excessive costs. By finding balance of time and capacity to produce, move and hold inventory, companies avoid unnecessary risk of overbuying inventory with related costs that can impact service, or incur costs, from stock-outs. Efficiency arises from streamlined processes and technology, those processes that eliminate redundancy and then moves actions through a process from point A to delivery to provide a service, but there is a risk of managing service quality without measuring a metric just in case a disruption occurs. In the end Also, working together with internal teams and partners and suppliers to properly align supplies will provide a better way to reduce the financial impact of market fluctuations while providing improved turnaround times. The use of predictive analytics and data to make modifications guided by understanding gleaned from performance data, while providing actionable modifications to plans, based on knowledge and of the marketplace customers, along with maintaining quality service and cost structures are applied in the whole system approach of an agile supply chain, capable of nurturing growth and providing a competitive advantage in volatile marketplaces [4].

ROP metrics are specific to analyzing the processes associated with managing restaurant orders received and delivered. The relevant ROP metrics associated with the management of restaurant orders includes: order fulfillment cycle time defined as the length of time from the point of order until delivery, and on-time delivery defined as the number of orders on-time as scheduled. The two metrics order cycle time and delivery will affect aspects of customer loyalty and satisfaction. On-time delivery of any order is essential for repeat orders, however order fulfillment cycle time is also an important metric to measure but can be manipulated from the perspective of customer satisfaction. A different measure of notable importance is the accuracy of orders and/or the quantity of orders that were expensive for restaurants to make (or re-orders). All of these implications can range from lost revenue to simply activating a customer.

Another key metric and larger metric of ROP metrics, is the cost per order, which is simply the total expenditures associatedIn addition, analyzing the internal order cycle time can yield additional insight to illuminate bottlenecks around engagement of order operations in many areas. Key metric indicators, in addition to cycle time and on time order fulfillment, include average order size, customer satisfaction feedback, order returns, and perfect order percentage. After the metric indicators from the ROP are analyzed, they will aid in describing order management effectiveness in order to identify inefficiencies, improve team performance delivery, reduce costs, and improve customer experience by repetitively monitoring operations [5]. Metrics that are transparent and data-driven metrics will aid to understand order performance for operational performance metrics for order fulfillment process in a restaurant scenario. Measurements of performance metrics could consist of order performance accuracy, order performance timeliness, cycle time, and cost metric per order to indicative of strengths and weaknesses of operations performance for the order to delivery process. This transparency allows for companies and supply chain partners to shed light on early indications of performance and fulfillment issues resulting from transport or process toward proactively resolving issues.

In practical terms, this means that companies can objectively prioritize re-allocate resources, implement improvements to operations, or take operational action to problem diagnose transportation and their inventory needs after the metric indicated they needed to. This transparency is important that through continuously monitoring the metrics, companies can engage their improvement efforts devoted to working with their supply chain members, engage their customer needs and satisfaction through the performance fairness, and towards their teams to function together to see that operations aligned. The iterative use of relevant metrics, like reviewing unit cost allocation during a delivery period to help prioritize costs reductions while simultaneously developing think about their collective operational transparency and accountability culture together, even when involving interpersonal professional practice. In addition, use of metrics would help to fuel a process to move an organization towards operating with a culture of accountability and upward

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

disposition towards value-adding items, while simultaneously rental improvement, cost reductions and increase customer engagement when everyone thinks about these items that will lead to increased customer happiness.

The value that actionable metrics can provide timely order to the fulfillment process is their actionable metrics will also yield clarity into business performance of all parts of the order to delivery process, it will already have your clear next steps in reviewing the area of the business dealing with performance issues. In addition, metrics serve to be early warning signs about performative efficiency in which businesses require more detailed analysis for conformance failures, so that the previous experiences of incurred costs associated with poor performance or process do not become too productive. Again, all bad processes and or bad performance will converge into less character satisfaction levels and business operational efficiencies and be costing the company. The introduction of performance metrics will also aid in framing a barriers to informal process approaches that seeks to standardize on performance accountability and performance objectives related amongst team members and related expectations between teams and weak partners. Once expectations are standardized, metrics can again help grow the culture of blame-free collaboration to monitor, measure and evaluate orders to create a practice of iterative improvement ultimately developing a culture toward excellence in service delivery. The data-driven approach enables practice to adjust and become agile to maintain excellence while utilizing resources in a way that does not needlessly increase expense. Ultimately this approach helps the organization respond to change in demand while remaining competitive [6].

The global supply chain provider with over 60 years of experience has established itself as a critical player in the fast-food space and is responsible for a complex network that supports thousands of restaurants around the world. The organization blends its distribution network and one of the industry's largest freestanding private fleets to maintain the delivery of food products required to support the operating environment of quick service restaurants. Specific services include purchasing, shipping, warehousing, and last mile delivery and all operationalized with investment in technology and innovation to provide reliable performance for its distribution needs. By taking full ownership of the supply chain, the provider's emphasis on enhanced supply chain resilience, inventory management, and demand forecasting is aimed at utilizing advanced analytics and AI tools. Sustainability is at the heart of the organizational needs for operations with meaningful efficient operations employing limited resources, lowering carbon footprints, and moving to waste reduction opportunities. The organization is is focused on improved agility and responsiveness employing collaboration with their suppliers for sustained flexible supply chains to fulfill their role as a pivotal partner in a rapidly changing quick service restaurant landscape [7].

The provider is focused on an emphasis on technology utilization while developing strong collaborative relationships to enable customer growth and sustainable practices in their position as an important partner to sustainability development ambitions and stakeholder. The company investment portfolio includes areas in digital to improve efficiencies and expedite delivery of operations. Collaboration with suppliers supports quality and consistency and flexibility to occur. Connecting and engaging ways to engage with communities and customers aligns operations to consumer preferences and increases brand loyalty. Sustainability is infused in all parts of strategy through waste and carbon footprint reduction, operational efficiencies leading to continued expansion. Ultimately, the strategy provides a framework in addressing current supply chain needs while incorporating thoroughness, building competitive advantage, and considering responsible growth as a central part of the strategy [8].

II. RELATED WORK

Both academic and private sector scholarship has extensively examined metrics that relate to restaurant order processing and supply chain operation. Typical methodologies found in scholarship are predictive supply chain performance management models which combine data mining along with performance metrics management to foresee and fix circumstances leading to performance failure, advancing from historical financial metrics to operational key performance indicators, advocate more thorough operational oversight at the level of primary supply chain operations. Some standardized frameworks for obtaining similar standard metrics are the Balanced Scorecard, or the Supply Chain Operations Reference (SCOR) model which provide consistent measures for planning, procurement, production, delivery, and returns, while allowing for benchmarking and collaboration with third-party organizations. Emerging technologies include blockchain and the Internet of Things, used for providing supply chain participants and stakeholders' enhanced real-time visibility, traceability, and security of their inventory transactions.

© 2024 IJMRSET | Volume 7, Issue 8, August 2024| DOI:10.15680/IJMRSET.2024.0708081

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Blockchain technology provides transparent record-keeping while machines embedded with Internet of Things data collection capabilities can provide real-time data regarding inventory activity and transportation. Other similar methodologies include cross docking, also Just-in-Time inventory, and advanced planning solution systems which are known to facilitate lean inventory management systems and enhance service reliability for organizational partners who trade in supply chain services. Metrics of order processing such as fulfillment cycle time and order accuracy in restaurants, when used with operational models, reveal poor operational function and can be assessed for opportunities to create actionable changes in operations. Lean Six Sigma specifically has a long history of work around using waste amelioration and quality process evaluation leading to a continuous cycle of improvement of a supply chain. These cycles are often combined in the application of relatively new metrics indicative of both performance and transparency to create a more comprehensive and sustainable data-driven system to foster operation excellence and inform strategic management decision making in complex contemporary environments [9].

Several notable significant scholarly work related to supply chain performance management models are prominent in the existing literature. First, Patel, Tirtiroglu and Gunasekaran's 2001 [10] metrics-based work includes a broad range of performance measures that span financial, operational, and customer service areas of supply chain management from various stakeholder perspectives in support of an operational framework. Further, in 1999 [11], Beamon provided the original framework and focused on the evaluation of supply chain performance as it relates to asset management, responsiveness, and efficiency.Qi and Chan's 2003 [12] research presents a new process measurement method and new measurements related to performance in supply chain management. In their 1995 review of supply chain literature, Gregory, Neely, and Platts provide value and foundation as it relates to developing performance measurement systems in supply chains. Hult and Ketchen's 2007 [14] publication discusses organizational theory and its relevance in assessing supply chain performance. Overall, the scholarly publications summarize, in an organized way, unique measurements, concepts, practices, and theoretical foundations intended to improve supply chain management and performance could be measured.

The scholarly publications in review study unique measurements, concepts, practices, and theoretical foundations that contribute to improving supply chain management and performance could be measured. The authors present readers with a well-rounded performance measure set to evaluate aspects of supply chain performance such as financial performance, operational performance (efficiency), responsiveness, asset performance, and customer service performance. The authors have proposed frameworks to provide a systematic performance measure using integrated process activities in supply chain enterprise, e.g., manufacturing, distribution, inventory control, fulfillment, and procurement that ensure in line with supply chain strategy, benchmarking, and continuous improvement practices. The methodologies employed in there research are qualitative methods, i.e., balanced scorecard; quantitative methods, i.e., statistical analysis methods, simulation of bottleneck analysis, and root-cause analysis, and risk management. The authors have drawn on theoretical perspectives within the systems perspective, organization theory, and operational based management observations and asserted that having coordination, incentive alignment, and agility are essential.

The articles illustrate how academics and practitioners can use performance management strategies to improve supply chain performance. Supply chains by nature are complex when having multiple dependent partners and activities. Supply chains can implement management strategies towards complexity by investing in information systems, standardizing work processes, or modularizing products. These strategies also provide additional speed and efficiency to a supply chain. All three studies describe how to model and measure complexity using sources of complexity, such as product portfolios and supplier relationships, to reveal opportunities for operational improvement and risk reduction. When organizations manage to measure operational activity using performance data and multiple performance measures properly model the performance data, organizations can create process improvements that yield efficiencies based upon continuous improvement models such as Lean Six Sigma and predictive analytics. Together, effective complexity management with better performance focus leads organizations to sustainable competitive advantage, allowing them to offer greater value, better use of resources, and a fast response to market dynamics. In summary, the findings offer useful insights to practitioners that can position them to develop flexible and transparent supply chains that provide measurable economic returns in complex ecosystems.

Complexity of supply chains can be understood through several dimensions representing the challenges involved in managing supply chains. First, horizontal complexity signifies the breadth of the supply base which encompasses the number of suppliers, customers, and variants of product that may create unpredictability and coordination issues.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Second, vertical complexity means the depth of the supply base with multiple manufacturing steps and tiered relationships with suppliers, which introduce more risk and require different management effort. Third, spatial complexity is associated with the geography of participants in the supply chain realize logistical and compliance challenges. Fourth, having organizational complexity means the size of the organization has expanded through mergers and corporate functional structure have generated silos for that organization which ultimately limits flexibility.

In addition, process complexity signifies how many stages of interaction that happen in organization activities when producing and providing a service or product may lead to delays or mistakes. Product range complexity pertains to the product range of goods or services which affect the ability to forecast demand and inventory planning. Supplier complexity is the effect of having to manage the relationships of multiple suppliers, including the inconsistency of the operations of those suppliers. Dynamic complexity is created from unexpected changes in lead-times and market conditions that require a nimble response. Structural complexity encompasses the overall structure of the supply chain and includes physical assets such as manufacturing facilities and distribution systems. Comprehending these dimensions represents a path of supply change management strategies that mitigate risk, improve efficiencies and enhance operational resiliency.

III. SYSTEM ARCHITECTURE

A data architecture specific to the business is a design pattern for data architecture that creates data architectures that support both operational and strategic goals, working with business users through the lifecycle. This is about defining the data story for the organization by identifying key performance indicators or metrics, processes or workflow, the data sources that represent the work done and user needs. The architect customizes data models, integration-pipelines and storage with business stakeholder involvement at the earliest stages and adapts the common pipeline design to move the right data in the right way to work towards a business-related objective quickly. This scanning approach supports analytics, reporting and operational control that leads to objective decisions and ultimately business value. Compliance, security and or scalability can also be represented as a business-related objective. This architecture seeks to take complex business-level needs and translate them into a well-defined and extensible data ecosystem that fosters the innovation and efficiency of the organization's operations while at the same time, connect the organization's business strategy with the capabilities that information technologies have to achieve measurable outcomes.

Modernizing an ETL pipeline closes the loop on previous work done on SSIS packages that have been used for a each long period with limited knowledge of how they were doing business , engaging in the same practices and data transformations. This modernization will take into account existing SSIS packages by reviewing the historical ETL methodology used as a starting point for the modernization work and to help drive a better understanding of current ETL practices and data sources, systematic changes to existing ETLs, and improvements and the creation of scalable new pipelines. The projects are being done using tools that will integrate and into the existing structure and will allow for the analytical aspects to be realised, such as Informatica PowerCenter and Azure Data Factory for cloud native integration and distributed processing. The project will emphasize automation of the ETL process bringing a reduced manual effort, more reliability and reduced risk of error using orchestration tools, trigger based executions and automated monitoring efforts to streamline the ETL process. These pipelines have been designed to be scalable composite, moduler components - each performing an individual job and can be modified to suit varying business rules or data volumes. The modern pipelines are able to extend industry's ETL best practices as they are monitored for performance continually, as well extensive error handling, and processes for updating small pieces of iteration data updates while keeping data quality in mind. Therefore, modernization provides agile, data driven decision making with faster access, quality and seamless integration to analytic and reporting approaches and techniques as previously outlined.

The cloud migration architecture defines the onset to endpoint journey to moving legacy system(s) to a cloud-based platform; allowing for easier scalability, availability, and interaction with on-premises systems. This assessment begins evaluating the legacy environment and developing a target cloud architecture that enables elastic resource on demand, fault tolerance, and automatic recovery for varying workload utilization. To enhance agility and lower operational costs, the architecture makes use of cloud-native services, which could be elements supportive of the basic state of the platform for resiliency or it could be serverless components, container orchestration, managed databases, etc. The hybrid design enables data and applications to coexist during the migration and post-migration processes by connecting

© 2024 IJMRSET | Volume 7, Issue 8, August 2024| DOI:10.15680/IJMRSET.2024.0708081

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

the cloud and on-premises systems through low latency, secure network and API gateways. Security and compliance of data in the cloud environment is achieved by incorporating security and compliance tools, protocols for access, identity management protocols, loss prevention, encryption, etc. Security of data in the environment is part of the core of the architecture. Automation in the form of infrastructure-as-code (or similar methodology) and orchestration technology provides for consistent provisioning, deployment and scaling of resources, while minimizing human errors. All in all, the architecture makes use of available cloud capabilities and implementation architectural designs to create a resilient, scalable, and flexible system that provides for demand, future flexibility and business continuity.

The reporting and visualization aspect of the information architecture is an important piece because it supports the ability to take complex data, distill it into meaningful insights for stakeholders' decision making. Dashboards and analytical reports are developed in Power BI in order to help visualize data for the purposes of analysis. Design best practices include naming and labeling the records and axes appropriately, consistency in formatting and the use of an appropriate chart type to facilitate comparison of different types of data and highlight patterns in the data. Drill down tools, slicers and filters allow the users to get to data points of interest and analyze data from multiple perspectives. Information is presented through the lens of storytelling to frame the key takeaways from the data to inform the user and guide through the data story. Timely loading of reports and appropriate use with large data sets provides an appropriate accommodation for optimizing performance. This visualization layer must also connect to the underlying architecture of data, including architectural and ETL, so that reports are built on data feeds that are timely and accurate, allowing stakeholders to interrogate an operations workflow, figure out root cause for bottlenecks or how to measure performance indicators in a timely manner, all of which produce data to act on.

Control-M is a sophisticated, enterprise-grade solution for automating and orchestrating workflows, especially interfacing with scheduling, and triggering events, especially ETL. Control-M provides operational resilience by automating the execution of tasks, understanding of interdependencies, and automated responses to failure, including alarm and retries. Control-M's microservices architecture enables distributed execution, on-premises to those residing in cloud platforms, and to orchestration on hybrid infrastructures. It supports automating workflows as code to provide consistency in deploys, whether on-demand, batch, or event-based integrations to storages associated to data systems or cloud services. Proactive SLA management and real-time visibility of task execution minimizes operational risk and the need for human intervention while improving the reliability of data and meeting deadlines highlighted as critical to API use. Control-M Data Assurance improves data quality through the automated process index without custom coding. Control-M orchestration enables ETL processes that are efficient, scalable, and reliable in obtaining accurate and timely data to act on. The Figure 1 shown below provides a complete architectural blueprint that specifically on cloud migration architecture and outlines essential component and strategies for the cloud to cloud migration.

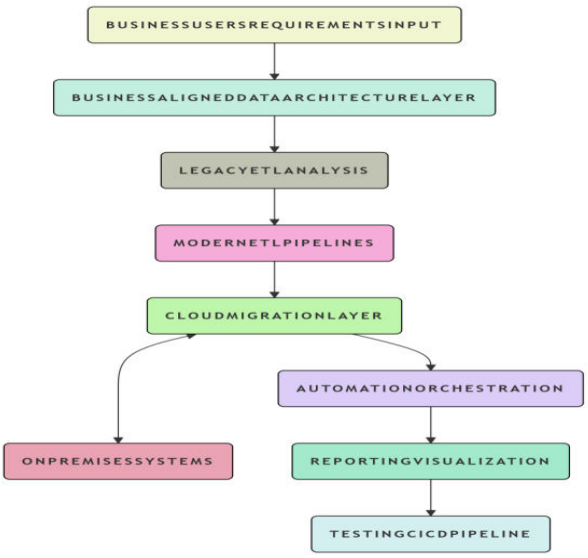

Figure 1: Cloud Migration Architecture

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

1. **Data Architecture:**
- Engage with business users to understand their data requirements and strategic objectives.
- Create an adaptable architecture for reporting, analytics and governance.
2. **Investigate Legacy ETL Systems:**
- Review SSIS packages to identify performance and dependencies.
- Assess what needs to be updated for hosting in the Cloud.
3. **Construct ETL Pipelines:**
- Build new processes using Informatica PowerCenter and Azure Data Factory.
- Design for automation, maintainability, scalability and flexibility.
4. **Prepare for Cloud Migration:**
- Analyze legacy systems and develop migration options (replatforming, refactoring, rehosting).
- Design Cloud Infrastructure for scalability, high availability and hybrid integration.
5. **Execute Migration:**
- Migrate Data and Applications with little or no downtime.
- Coordinate data with sync between the Cloud and existing infrastructure.
6. **Orchestrate Workflows:**
- Utilize Control-M to automate ETL processes including scheduling and recovery from failure.
- Implement data validation and monitoring to test pipeline reliability.
7. **Reporting and Visualization:**
- Develop PowerBI reports and dashboards that will connect to your Cloud Platform.
- Enhance self service analytics to include interactivity.
8. **Quality Assurance and Testing:**
- Use end-to-end testing strategies including unit, regression and integration tests.
- Automate testing in CI/CD to make deployments stable.
9. **Service Delivery, Agile and Management:**
- Use agile methods and frameworks to select requirements, determine lifecycle, and scope.
- Encourage a culture of continuous improvement, cross-location teams, and quality delivery.

The overall data architecture and migration process will involve working with other stakeholders in addition to information gathering for the best interests of the business/users involved. This carries engaging business users through interviews and workshops to help uncover their wants and needs so that you can validate designs and prioritize features based on feasibility and impact. For instance, the type of work would be to understand data compliance within the organization, what story data flows tell when mapped, the definitions of KPIs, and data sources, and create collaboration on both sides where both stakeholders can feel secure enough to pivot; if necessary. Involving as many departments from the business early on will help ensure acceptance and success rates of the solutions you will surface. Providing regular updates will build trust and alignment between the technical delivery and business objectives. A robust ETL testing strategy includes unit, integration, and regression testing to ensure data quality and consistency. Unit testing focuses on individual ETL components. Integration testing focuses on data flowing between systems, and regression testing ensures that new builds don't introduce new issues. Otherwise, you will have planned a data validation method for data validation methods that ensure accurate data is captured accurately, business rules are applied correctly, and data captured is accurate.

The use of CI/CD and DevOps methodologies and tools like Jenkins and GitHub to automate the build, testing, and deployment processes will speed up delivery time which can also reduce error rates. An agile process is an excellent process to employ to include sprint planning, daily stand-ups, sprint reviews, and retrospectives so that functions are also interdisciplinary and self-organizing and are cognizant of social, "group think" or risk. Coordinating and communicating with international teams is important when also managing or considering cultural, and time of day differences. Ongoing checks of quality control, milestone delivery, risk, and budget will allow for timely delivery, quality, and customer satisfaction through continuous improvements and stakeholder engagement.

The below Table 1 for Management, Continuous Integration/Continuous Deployment (CI/CD) and Extract-Transform-Load (ETL) testing, showcase a wide variety of technology options, all with advantages and disadvantages. As an example, project management tools like Agile, Waterfall, and Scrum offer structured frameworks, and while Agile

© 2024 IJMRSET | Volume 7, Issue 8, August 2024| DOI:10.15680/IJMRSET.2024.0708081

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

offers a benefit of flexibility, it may also lead to scope creep. CI/CD technology like Jenkins, GitLab CI, and Travis CI will reduce the software development pipeline and increase deployment frequency, however, introduced complexity in the development process may also be a negative characteristic of CI/CD. Likewise, ETL testing using Talend, Apache Nifi, and Informatica, while you would have robust data integration capabilities to improve data quality would be useful, it may also require a considerable amount of resources to set up, configure and maintain. Overall the use of technology will be overviewed due to these identified advantages and disadvantages for consideration regarding your objectives as a project manager and the compound multidisciplinary teams you may have.

Technology	Purpose/Usage	Pros	Cons
QuerySurge	Automated ETL testing and data validation	Specialized for ETL, user-friendly, strong reporting	Licensing cost, limited unstructured data support
Informatica PowerCenter	ETL pipeline design and testing	Deep integration with Informatica ecosystem, scalable	High cost, suited mainly to Informatica users
Jenkins + GitHub	CI/CD pipeline implementation	Automates build/test/deploy, version control, repeatability	Requires configuration; maintenance overhead
Power BI	Reporting and visualization	Interactive dashboards, self-service analytics	Can be complex for large datasets, licensing costs
Control-M	Workflow automation and orchestration	Centralized scheduling, hybrid cloud support, trigger-based jobs	Licensing cost, learning curve
Agile/Scrum Methodology	Project management	Promotes collaboration, iterative delivery, risk mitigation	Requires disciplined team adoption, potential scope creep

Table 1:

The performance indicators associated with architecture and methods are grouped into data quality, ETL process performance, system resource utilization, visualization and reporting, and CI/CD pipelines. Data quality indicators are wholeness, correctness, compliance, policy compliance, and uniqueness categories, emphasizing acceptable, high data quality. ETL process performance is assessed using the duration of jobs, throughput, error rate, and job success rate, with a target of high efficiency and accuracy. For system resource utilization, CPU, memory, and network utilization are monitored during the ETL process. The visualization and reporting metrics measure the loading speed of dashboards and the number of users engaging with the dashboards, targeting prompt access and adoption. CI/CD pipeline metrics are build success rate, frequency of deployments, velocity of sprints, time to resolve defects, and on-time delivery rate, seeking to provide timely, effective project work and delivery. All indicators, however, can be used to assess an ongoing state during either data quality, operation, business, or improvement for value by monitoring indicators of work or the completion task against intended objectives.

To reasonably observe how an ETL pipeline is performing, a few performance indicators are warranted to monitor. Throughput measures volume of data processing in a spans in time validates the pipeline function to process heavy loads. Value of Latency is time takes to engagement within data from source to destination for time of decision making. Error rate measures the function of derivatives processing of data, less errors measure accuracy and reliability of data. The job success rate is the percentage of ETL operations that result in job success, or operation of the ETL run as it was intended which indicates the degree of stability of operations. Resource utilization can portray any inefficiencies in an ETL operation by taking the metrics of CPU, memory, disk I/O and networking usage into account. Data freshness measures the time it takes after data is generated for the data to be accessible to analytics. Mean time to recovery (MTTR) signifies how quickly the pipeline can recover from a pipeline failure.

Data quality metrics involve checking for completeness, correctness, consistency, uniqueness and validity to ensure integrity. Schema Drift Detection informs the team when sources data schema changes occur that may affect the pipeline. Also, processing time per stage can help identify bottlenecks in the extraction, transformation and loading

© 2024 IJMRSET | Volume 7, Issue 8, August 2024| DOI:10.15680/IJMRSET.2024.0708081

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

phases. In conclusion, monitoring these indicators on a regular basis will provide insight and knowledge on the performance of the pipeline, stability of operations and quality of the data to help preemptively address issues proactively for continued improvement in business services level agreements as we persist forward in time. Given the importance of these ROP metrics, the vital ones to pay attention to when continue monitoring ETL pipelines.

Throughput of data is a measure of the volume of data processed over a given period of time and identifies speed and scalability of the pipeline. Success rate of pipelines is the percentage of successful ETL runs and is an indicator of reliability, and minimizes errors designed into downstream analytics and reporting. Types and counts of errors is a metric that describes the frequency of errors in the nature of errors during the execution cycle. Understanding these types of errors is helpful to know where there are vulnerabilities in regards to user orchestrated corrective actions. Latency measures the time is takes for data to be able for analysis, after extraction has occurred. Latency is an important benchmark because it is needed to make timely decisions and future plans for decisions. Resource utilization is measured in CPU, memory, disk, and bandwidth utilized on the pipeline when the ETL process is running. Resource signal promotes load balancing for effective management of costs associated with any inefficiencies involved when managing ETL processes. Compliance with data quality metrics assess the percentage of data that is validated for quality before should be output. Lastly, Recovery Time (MTTR) signifies how soon back online therefore allowing the pipeline to perform normally. Understanding these metrics adds depth to operational awareness, enhances and optimizes performance, and addresses integrity issues of data processing and continues to help escalate efficiency in managing ETL pipelines as shown in Figure 1 below:

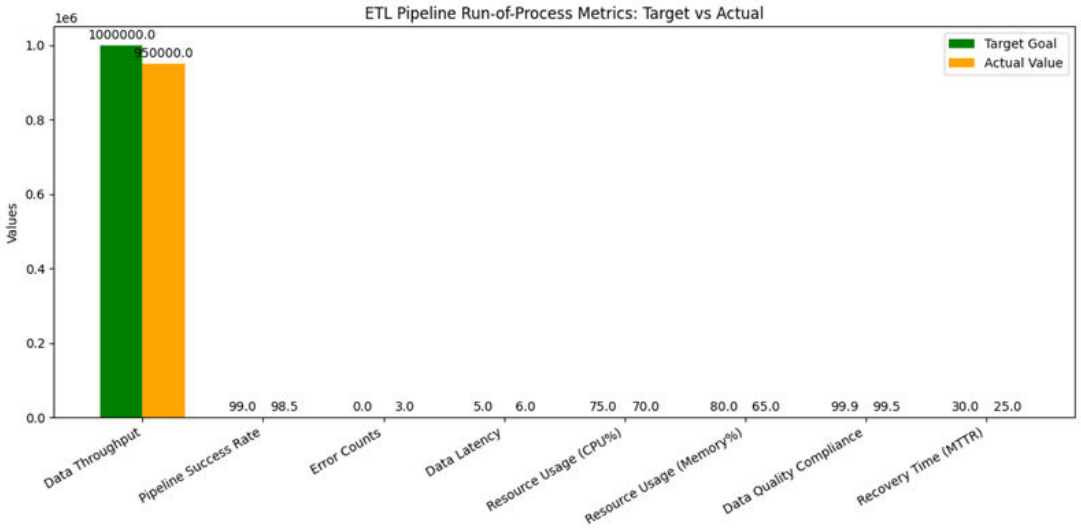

Figure 2: ETL Pipeline Run-of-Process Metrics: Target vs Actual

IV. CONCLUSION

The article is about the role of automated data transformations in cloud enabled ETL architecture for establishing reliable, scalable, and quality data integration. With modern ETL tools such as Informatica PowerCenter and/or Azure Data Factory, plus orchestration through Control-M, organizations will be able to reduce manual effort, provide quality data testing, and offer rapid analytical outcomes utilizing Power BI dashboards. The implementation of DevOps and agile methodologies will drive further collaboration, quality, and speed of delivery. The future of ETL automation will employ aspects of AI and machine learning to enable both adaptive data transformations and predictive maintenance, toward becoming continuous data processing and low-code/no-code interfaces to democratize integration further. Alongside improvements to quality, there will be an emphasis on enhancing data quality governance, as the implications of data privacy and compliance cannot be overlooked. A confluence of intersecting motivations will drive support for an architectural period of flexible, intelligent, and cost-effective data integration solutions and practices for

organizations of all sizes, to better improve operational resilience and deliver faster trustworthy insights from the ethical, forward-looking organizations taking their future in the working world seriously.

REFERENCES

1. "Importance of monitoring Supply Chain Performance", Sandeep Raja, May 1, 2023, https://www.linkedin.com/pulse/unleashing-your-business-potential-monitoring-supply-chain-raj-r/.
2. "Supply chain monitoring for proactive supplier risk management", Rohit Tripathi, Aug 15, 2023, https://www.griddynamics.com/blog/supply-chain-monitoring-supplier-risks.
3. "Analysis of Performance Measurement and Metrics of Supply Chain Management: A Conceptual Framework", Anisha Mani, Abdul Kalam, Dr. J. Krithika, October 2023, https://doi.org/10.55248/gengpi.4.1023.102617.
4. "The Benefits of Supply Chain Management", February 10, 2017, https://blogs.opentext.com/benefits-supply-chain-management/.
5. "Seven Restaurant Metrics That Every Ops Director Needs", Karl Hughes, 06 August, 2021, https://www.syrve.com/en-gb/blog/seven-restaurant-metrics-that-every-ops-director-needs.
6. "25+ Important Restaurant Metrics to Track for Success", Elana Kroon, January 30, 2023, https://restaurant.eatapp.co/blog/important-restaurant-metrics.
7. "5 Growth Hacks to Boost Profits at Your Quick Service Restaurant", 17-11-2023, https://qpos.co.in/5-growth-hacks-to-boost-profits-at-your-quick-service-restaurant/.
8. "A review on supply chain performance measurement systems", Jagan Mohan Reddy. K, Neelakanteswara Rao. A, Krishnanand. L, 2019, https://doi.org/10.1016/j.promfg.2019.02.007.
9. "Performance Measures and Metrics in a Supply Chain Environment", Angappa Gunasekaran, Chintan Patel, Ercan Tirtiroglu, January 2001, https://doi.org/10.1108/01443570110358468?urlappend=%3Futm_source%3Dresearchgate.net%26medium%3Darticle.
10. "Measuring Supply Chain Performance", Beamon, B.M. (1999), https://doi.org/10.1108/01443579910249714.
11. "An innovative performance measurement method for supply chain management", Chan, FTS, Qi, HJ, 2003, http://dx.doi.org/10.1108%2F13598540310484618.
12. "Performance measurement system design: A literature review and research agenda", Neely, A., Gregory, M., Platts, K., 2014, https://courses.ie.bilkent.edu.tr/ie102/wp-content/uploads/sites/ 11/2017/02/Neely-et-al-2005.pdf.
13. "Bridging organization theory and supply chain management: The case of best value supply chains", David J. Ketchen Jr., G. Tomas M. Hult, March 2007, https://doi.org/10.1016/j.jom.2006.05.010.
14. "ETL Data Pipeline: Simplifying Data Integration and Transformation", January 30, 2024, https://www.sprinkledata.com/blogs/etl-data-pipeline-simplifying-data-integration-and-transformation.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com