# Cloud Native Approach to Automate Implementation of Network Strategy

A technical paper prepared for presentation at SCTE TechExpo24

**Richard Brown**
Director
Cox Communications
richard.brown@cox.com

**Pavan Chandrashekar**
Lead Data Engineer
Cox Communications
pavan.chandrashekar@cox.com

# Table of Contents

**Title** **Page Number**



## List of Figures

**Title** **Page Number**

# 1. Introduction

In the rapidly evolving telecommunications industry, effective network planning is essential for maintaining a competitive edge and ensuring agile service delivery. As the third-largest Multiple System Operator (MSO) in the United States, our company is uniquely positioned to leverage cutting-edge technologies to enhance its network planning processes. One such transformative technology is cloud data services.

Cloud Services offers a scalable, flexible, and robust solution for managing the vast amounts of data generated by telecommunications networks. By utilizing cloud-based platforms, we can streamline our data management processes, improve the accuracy and speed of network planning, and ultimately deliver superior services to our customers. This paper explores the strategic benefits of adopting cloud services for network planning and provides a roadmap for successful implementation within our company.

The traditional approach to network planning involves significant manual effort and is often constrained by the limitations of on-premises data storage and processing capabilities. As network demands continue to grow, these limitations can hinder the ability to make timely and informed decisions. Cloud services, on the other hand, offer virtually unlimited storage and computing power, enabling us to process large datasets in real-time and gain valuable insights into network performance and capacity requirements.

Moreover, cloud-based data services facilitate advanced analytics and machine learning applications, which can predict network issues before they occur, optimize resource allocation, and identify opportunities for network expansion. By harnessing the power of the cloud, we can enhance our predictive capabilities, reduce operational costs, and improve overall network reliability and efficiency.

In this paper, we will examine the specific advantages of cloud services for network planning, including scalability, agility, enhanced analytics, and improved collaboration. We will also address potential challenges and considerations, such as costs and integration with existing systems, and provide best practices for a successful transition to a cloud-based network planning solution.

As our company continues to innovate and grow, embracing cloud-based data services for network planning represents a critical step towards futureproofing our network infrastructure and maintaining our position in the telecommunications industry. By leveraging the full potential of cloud technology, we can ensure that our network remains robust, agile, and capable of meeting the evolving needs of our customers.

On-premises big data platforms, while powerful, come with their own set of challenges. Managing and maintaining an on-premises big data platform cluster requires substantial hardware investments and a dedicated team of skilled professionals. These on-premises infrastructures often face limitations in scalability and can struggle with the dynamic demands of network data processing. Furthermore, the operational overhead associated with Hadoop can divert resources from core business activities. In contrast, cloud services provide a fully managed cloud environment that eliminates the need for physical hardware and complex system administration, allowing our teams to focus on leveraging data insights for strategic network planning.

# 2. How It Is Built

## 2.1. The Migration

Migrating our on-premise network planning processes to the cloud involved a meticulous and well-orchestrated approach, leveraging a suite of cloud services to ensure a seamless transition while enhancing the scalability, flexibility, and performance of our network planning operations. The process began with a comprehensive assessment of our existing infrastructure, which included evaluating our hardware and software configurations, understanding our data workflows and dependencies, and identifying the volumes of data that needed to be migrated.

The first phase involved detailed planning and requirement gathering. We conducted a thorough analysis of our current infrastructure, identifying the data pipelines, the nature of our workloads, and the interdependencies of various applications. This assessment helped us map out a migration strategy that minimized downtime and ensured continuity of operations. Our goal was to create a cloud environment that not only replicated our on-premises setup but also took full advantage of the cloud's capabilities to enhance our operations.

Based on our assessment, we designed a robust cloud architecture centered on a serverless approach. We selected key cloud services to replace our on-premises components, ensuring a smooth and efficient transition.

- A scalable and durable storage solution was chosen for our raw and processed data. Its integration with other cloud services made it an ideal choice for our data lake, allowing us to store vast amounts of data with high durability and availability. The scalability ensured that we could handle our growing data needs without concern for infrastructure limitations.

- An automated service was used to streamline our ETL (Extract, Transform, Load) processes. This service's ability to automatically generate code for ETL jobs based on data sources significantly reduced our development time. By using this service, we could streamline our data processing pipelines, ensuring that data was cleaned, transformed, and made available for analysis in a timely manner. Its serverless architecture also meant that we did not have to manage any underlying infrastructure, further simplifying our operations.

- A serverless compute service enabled us to execute code in response to events without the need to manage servers, making it perfect for automating various aspects of our data processing workflows. With this service, we could trigger specific processes based on changes in our data, ensuring real-time processing and reducing latency. This approach allowed us to build a highly responsive and scalable architecture.

- A workflow orchestration service was used to manage complex workflows, ensuring that each step of our ETL processes was executed in the correct sequence. By using this service, we could define and visualize our workflows, making it easier to manage and debug our processes. This orchestration approach ensured that our data pipelines were robust and reliable, with clear visibility into each stage of the process.

In addition to these core services, we also integrated interactive query services and managed NoSQL databases into our architecture. The interactive query service allowed us to analyze data directly in storage using standard SQL, enabling us to perform ad-hoc queries on our data lake without the need for complex ETL processes, providing quick and flexible access to our data for analysis and reporting. The managed NoSQL database service was used to store metadata and other structured data that required low-latency access.

To facilitate the data migration process, we developed a custom framework using a combination of database migration, data transformation, and serverless compute services. This framework allowed us to transfer data from our on-premises systems to the cloud efficiently and securely. The database migration service enabled us to migrate our databases with minimal downtime, while the data transformation and serverless compute services handled the data transformation and loading processes. This approach ensured data integrity and reduced manual intervention, allowing us to focus on optimizing our new cloud environment.

A crucial aspect of our migration strategy was setting up a robust CI/CD pipeline to ensure smooth and automated deployment processes. We utilized a version control system for managing and tracking changes to our infrastructure and application code effectively. Infrastructure as code (IaC) was employed to define and provision our cloud resources in a consistent and repeatable manner. By using IaC, we could automate the creation and management of our cloud infrastructure, ensuring that our environments were always in sync and reducing the risk of configuration drift.

An automation server was employed to orchestrate our CI/CD pipeline. We set up the server to automate the build and deployment processes, ensuring that our code changes were continuously integrated and delivered. Pipelines were configured to pull the latest code from the version control system and deploy the changes to our cloud environment using IaC. This setup allowed us to deliver new features and updates rapidly, with high confidence in the stability and reliability of our deployments.

Throughout the migration, we maintained a strong focus on security and compliance. The robust security features of the cloud services, including encryption at rest and in transit, identity and access management with the least privilege principle, and comprehensive logging and monitoring, provided the necessary controls to protect our data and ensure regulatory compliance.

By leveraging a powerful suite of cloud services and integrating CI/CD practices, we successfully migrated our on-premises network planning processes to the cloud, achieving significant improvements in scalability, flexibility, and performance. This migration not only modernized our infrastructure but also positioned us to take full advantage of the innovative capabilities offered by the cloud, driving greater efficiency and value for our organization.

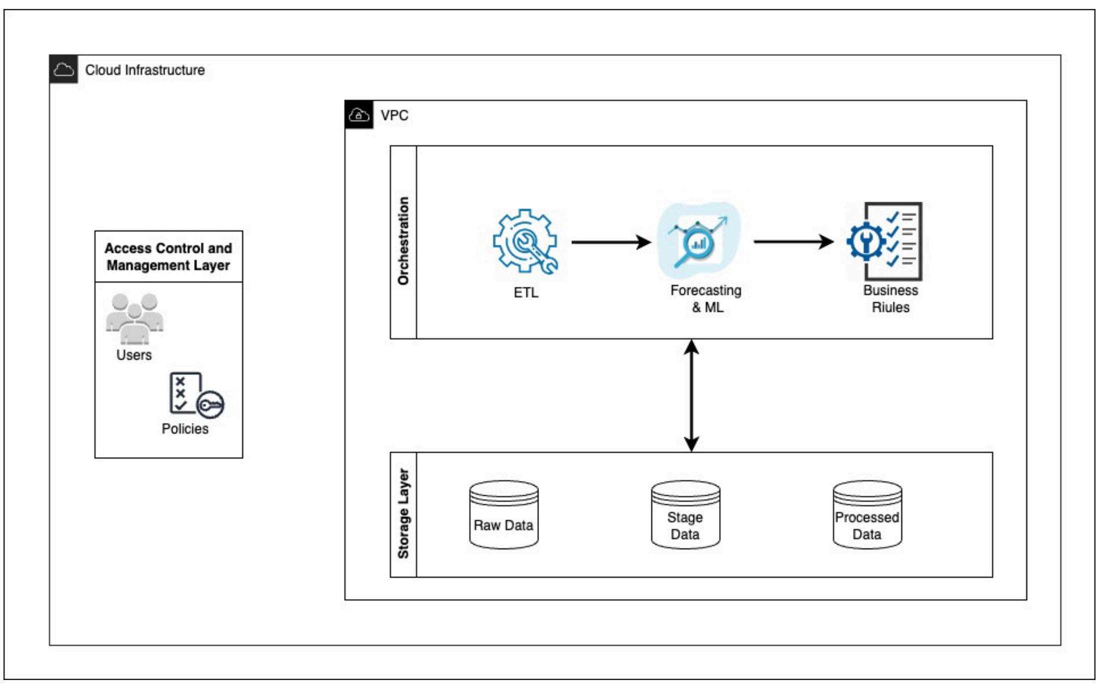


**Figure 1 - Cloud Services and Design**

## 2.2. Linking Data

### 2.2.1. ETL

With data securely stored in a scalable and durable storage solution, we harnessed the power of an automated ETL (Extract, Transform, Load) service to streamline our processes. The serverless architecture of this ETL service eliminated the necessity of managing underlying infrastructure, allowing us to concentrate on developing ETL jobs rather than on platform maintenance. This shift significantly reduced operational overhead, improved scalability, and leveraged distributed computing capabilities for efficient data processing.

To ensure seamless integration of data from various sources, we employed a serverless interactive query service. Its SQL-based querying allowed us to analyze data directly in the storage solution without needing to load it into a database. By using this query service, we could easily ingest data from other models and incorporate it into our process.

To begin, we cataloged our diverse data sources using an integrated data catalog service. This cataloging step was crucial for organizing our data assets and enabling seamless discovery and query. The data catalog maintained metadata about our data, which simplified data management and enhanced data governance.

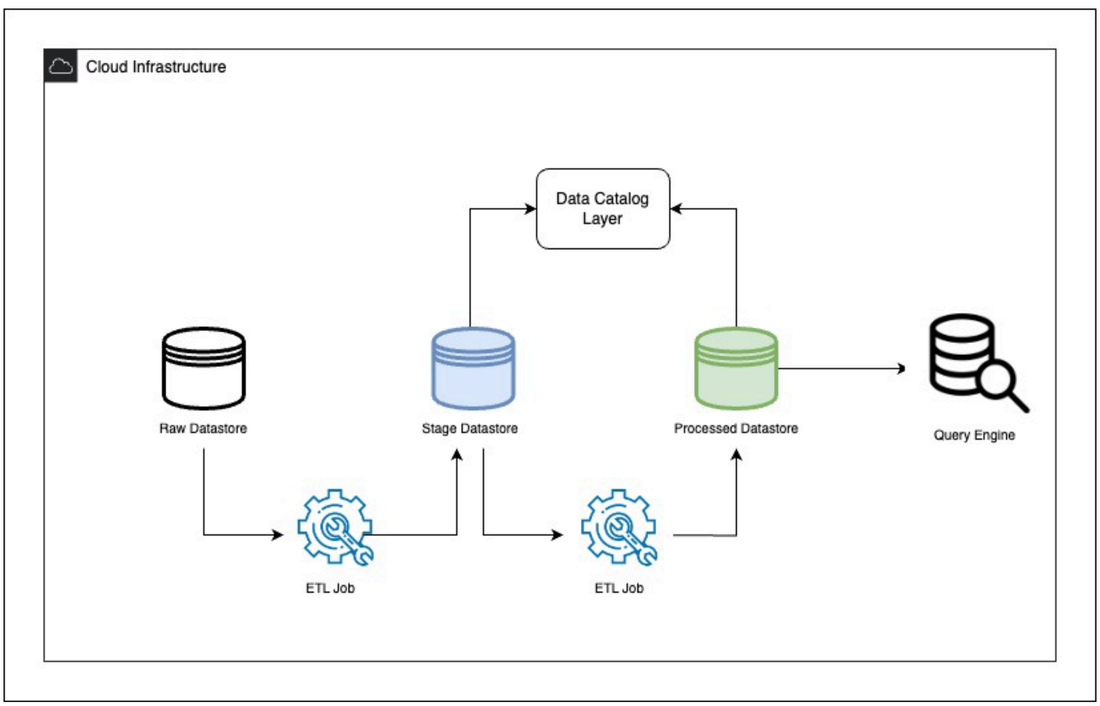


**Figure 2 - ETL Workflow**

Our ETL jobs were scripted to perform complex data transformations, including data cleansing, normalization, and enrichment. The flexibility of our scripting approach allowed us to incorporate custom logic and advanced transformation techniques to ensure our data was accurately processed and ready for downstream analysis.

A key component of our ETL process was the use of automated crawlers. These crawlers traversed our data stores, inferred schema, and updated the data catalog. This automation ensured that our metadata remained current, reflecting any changes in the underlying data. By maintaining up-to-date metadata, we facilitated more accurate data querying and reporting.

### 2.2.2. Event-Driven Processing

Automated functions played a pivotal role in streamlining our data workflows. The event-driven architecture enabled us to trigger processing jobs in response to specific events, such as the arrival of new data in cloud storage. This approach ensured that our ETL processes were initiated promptly, reducing latency and enhancing the timeliness of data availability.

For instance, when new data was uploaded to the cloud storage, an automated function was invoked. This function validated the incoming data, ensuring it met predefined quality standards. Upon successful validation, the function triggered the appropriate job to process the new data. This seamless integration provided an efficient and robust mechanism for handling real-time data ingestion and processing.

Additionally, automated functions were employed for monitoring and notifications. After the completion of ETL tasks, these functions evaluated the success or failure of the processes. Notifications were sent via a messaging service to alert stakeholders about the status of ETL jobs, ensuring timely awareness and response to any issues.

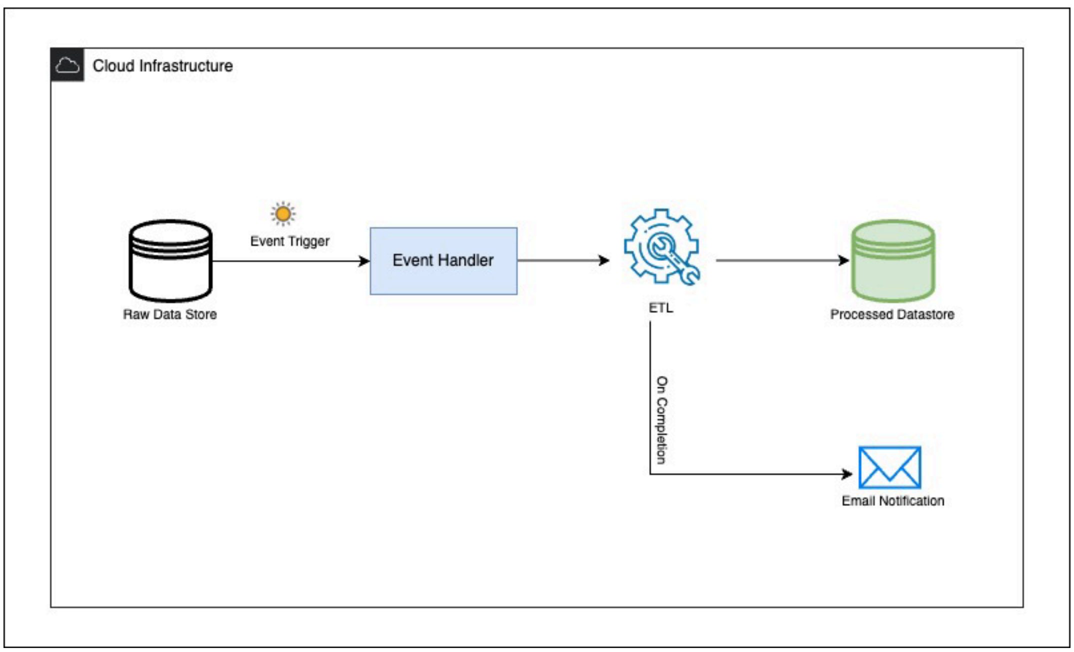


**Figure 3 - Event based ETL**

### 2.2.3. Orchestration

Managing the complex sequence of ETL operations required a sophisticated orchestration mechanism, and a state machine orchestration service provided the ideal solution. This service allowed us to design and implement state machines that coordinated various data processing tasks, ensuring that each step was executed in the correct order.

Using this orchestration service, we built visual workflows that clearly illustrated the flow of our ETL processes. These workflows included conditional branches, retries, and parallel execution paths, accommodating the diverse requirements of our data pipelines. The visual representation made it easier to understand, monitor, and troubleshoot the ETL processes, enhancing overall operational efficiency.

To further enhance our automation capabilities, we developed an automation document to enable the synchronous execution of a custom process on a virtual machine instance, which was integrated as a step in our state machine workflow. By incorporating this custom process, we extended the functionality of our data pipelines, allowing for complex and diverse data processing tasks that required additional computational power. This seamless integration ensured that all components of our ETL process were effectively managed and executed within a unified workflow.

The orchestration service also provided detailed logging and metrics, offering insights into the performance and status of each workflow. This visibility enabled us to quickly identify and address any bottlenecks or failures, ensuring the reliability and robustness of our ETL operations.

In conjunction with the orchestration service, automated workflows provided a robust framework for managing the lifecycle of our ETL jobs. These workflows allowed us to define and monitor a series of interconnected jobs, ensuring that each job was executed in the correct order. This capability was essential for managing dependencies between tasks and coordinating the overall ETL process.

#### *2.2.4. Scheduled and On-Demand Pipelines*

To meet various business requirements, we established both scheduled and ad-hoc ETL pipelines. Scheduled pipelines operated at regular intervals, like weekly or monthly, ensuring consistent and periodic data processing. These pipelines were managed using workflows that orchestrated the sequence of ETL jobs and managed dependencies between tasks.

Conversely, ad-hoc pipelines were triggered in response to specific events or user requests. This flexibility proved vital for scenarios in network planning where immediate data processing was essential, such as for ad-hoc analysis or urgent requests from leadership.

### 2.3. Output and Accessibility of Network Plans

Our processes generated valuable network plans that needed to be accessible to various downstream systems and stakeholders. To achieve this, we implemented multiple strategies for output and data accessibility, ensuring that the network plans were easily available for analysis, integration, and decision-making.

#### *2.3.1. API for Programmatic Access*

We utilized RESTful APIs as a key method to make our processed network plans accessible. This enabled seamless programmable access for other systems and applications to retrieve our data. The use of secure and scalable endpoints ensured that our network plans could integrate smoothly into various workflows and applications. This approach supported real-time data access, enabling external systems to fetch current network plans as required. Furthermore, implementing access controls and monitoring mechanisms ensured that data access was restricted to authorized users, maintaining robust levels of security and performance.

#### 2.3.2. Pushing Data to Relational Databases

To integrate with stakeholder applications already operating on-premises, we employed data connections to push transformed data into their relational databases. This step was essential for ensuring stakeholders could incorporate our outputs into their workflows without disruption. The relational databases offered a dependable environment for storing network plans, enabling efficient data retrieval and analysis. Through these connections, we maintained smooth data transfer, ensuring external databases remained updated with the latest processed network plans.

#### 2.3.3. Data Visualization

To support data visualization and business intelligence efforts, we made our processed network plans accessible through a structured data layer. This allowed for dynamic and interactive dashboards, aiding stakeholders in gaining real-time insights into the network plans. Visualizing the data in this way enabled users to spot trends, patterns, and anomalies easily, which in turn supported informed decision-making based on data. The interactive features of these dashboards allowed users to delve deeper into specific details and tailor views according to their requirements, thereby enhancing the overall utility and value of the network plans.

#### 2.3.4. Centralized Data Lake and Access Management

We utilized a centralized data lake to store all our network plan data in a unified repository. To manage access and ensure data security, we employed a comprehensive access management system. This system streamlined the setup of secure data lakes by automating data cataloging, access control, and data ingestion processes.

Facilitating cross-account data sharing proved especially advantageous for distributing network plans across external partners or different departments within our organization. This approach streamlined and secured data sharing, fostering collaboration and maximizing data utilization across various teams and partners.

## 3. Advanced Functionality

Cloud services offer a comprehensive suite of cloud services that cater specifically to the needs of large-scale data analytics and storage. Services such as elastic clusters allow for big data processing, and services like serverless computing allow for seamless, scalable data management. These services are designed to handle varying workloads and can automatically scale up or down based on demand, ensuring optimal performance, cost efficiency, and agility. By moving to cloud services, our network planners can access real-time data processing capabilities, enabling quicker decision-making and more responsive network management. As well, we can meet the ever-changing demands as we shift strategic focus based on external factors, such as competitive pressure or population surges.

Moving from our on-premises big data clusters to cloud services, our team was able to leverage PySpark, leaving little development once the data sources, code, and processes were migrated. Utilizing and deploying our existing statistical and machine learning models saved time in the ability to migrate and created a faster path to a full end to end deployment. As well, we were able to see instant results in efficiency of the runs with ability to scale more resources to support the large-scale modeling.

Additionally, cloud services offer advanced analytics and machine learning tools, such as cloud-based machine learning libraries, that provide powerful capabilities to expand our predictive analytics capabilities. These services can be deployed to analyze vast amounts of network data to predict load, optimize network upgrades, and help us to enhance overall network performance. We are also finding the flexibility and innovation offered by cloud services can drive significant improvements in network reliability and efficiency as we partner with our analytics team's cloud services assets for collaboration in solving complex network problems.

In the highly competitive telecommunications industry, optimizing network performance and making swift, informed decisions are critical for maintaining service quality, a competitive advantage, and operational efficiency. Cloud services offer a robust suite of services that we leverage for rules-based optimization and decision making, enhancing network planning capabilities and implementing strategic policies as we plan years into the future. By integrating cloud services powerful cloud infrastructure with advanced rules-based systems, MSOs can automate complex processes, optimize resource allocation, and enhance the speed of large-scale decision-making.

## 4. What Is Next

Transitioning from on-premises computing platforms to cloud services also enhances collaboration and integration with the enterprise. Cloud services offered as part of a cloud-based environment facilitates easier sharing of data assets and resources across teams and departments, fostering a more collaborative approach to network planning and offered end-to-end automation. Additionally, cloud services integrate seamlessly with a wide array of third-party tools and services, allowing for a more flexible and interconnected ecosystem. This interoperability is crucial for leveraging best-of-breed solutions and staying ahead in a competitive market.

We can also explore integrating AI with cloud services for reporting and automation, which will transform the network planning process by providing advanced, real-time analytics and insights. With the size of our models, and the vast output, reporting can be complex. With AI-driven automation, we can reduce the burden of manual data analysis, while increasing the functional delivery of data to outside stakeholders and executive teams. The ability to create detailed, interactive dashboards democratizes data access across the organization, fostering a data-driven culture. By harnessing the power of AI and cloud services, telecommunications companies can achieve more accurate, efficient, and strategic network planning, ultimately delivering superior service to their customers and maintaining a competitive edge in the industry.

## 5. Conclusion

The scalability of cloud services ensures that as our data volumes grow, the infrastructure and reporting can scale accordingly without compromising performance. This scalability, combined with the cost-efficiency of cloud computing, means that MSOs like ourselves, can handle increasing amounts of data without significant additional investments in physical infrastructure.

In the ever-evolving telecommunications industry, utilizing cloud services for effective network planning is paramount for maintaining service quality, ensuring operational efficiency, and staying competitive. As the complexity and scale of network data continue to grow, traditional approaches to data management and decision-making are just no longer sufficient. Transitioning to cloud-based solutions, like cloud services for data modeling, predictive analytics, and rules-based optimization and decisioning, offers transformative benefits that can drive significant improvements in network planning.

Cloud services provide a robust and scalable cloud infrastructure that facilitates the automation of complex processes and enhances decision-making through real-time data insights and advanced analytics. By utilizing services, MSOs can automate responses to network events, optimize resource allocation, and predict potential issues before they impact service quality. This integration not only reduces operational overhead but also ensures a more agile and responsive network management approach, while providing more accuracy for planned activity.

The scalability and flexibility of cloud services allow for dynamic resource scaling based on real-time demands, ensuring that network resources are used efficiently and cost-effectively. This is particularly crucial for handling varying traffic loads and preventing congestion, thereby maintaining high standards of network reliability and customer satisfaction. Additionally, the ability to create interactive dashboards and to utilize comprehensive data processing with cloud services enables stakeholders at all levels to make informed, data-driven decisions.

By embracing cloud services for rules-based optimization and decisioning, MSOs can futureproof their network infrastructure, streamline their operations, and achieve a more proactive and strategic approach to network planning. This transition not only enhances operational efficiency but also positions companies to better meet the growing demands of their customers, ensuring sustained competitiveness in a rapidly changing industry landscape.

As well, with cloud services and other cloud-based services, this opens the door for faster access and more streamlined integration for emerging AI technologies. Adding the efficiency of AI into processes and reporting can lead to more understanding of the complex data and provide a faster ability to make decisions impacting years of network growth.

In conclusion, the integration of cloud-based technologies provided by cloud services represents a critical advancement in network planning. The shift from traditional on-premises solutions like Hadoop to cloud services cloud platform offers unparalleled opportunities for optimization, automation, and intelligent decision-making. As the telecommunications industry continues to evolve, leveraging cloud services capabilities will be essential for delivering superior network performance and maintaining a leading edge in the market.

# Abbreviations

| | |
|---|---|
| IaC | Infrastructure as Code |
| CLOUD SERVICES | |
| MSO | |
| SCTE | Society of Cable Telecommunications Engineers |

# Bibliography & References


1. Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., ... & Zaharia, M. (2010). *A View of Cloud Computing*. Communications of the ACM, 53(4), 50-58..

    This paper provides an overview of cloud computing, discussing its definition, benefits, and challenges. It introduces the key concepts of cloud services, deployment models, and the economic implications of cloud computing.

2. Mattingly, M. (2022). *Best Practices for ETL in the Cloud*. O'Reilly Media.

    This book offers best practices for designing and implementing ETL processes in cloud environments. It covers a range of topics, including data integration, cloud storage, and orchestration tools, with practical examples from leading cloud providers.

3. Buyya, R., Broberg, J., & Goscinski, A. M. (Eds.). *Cloud Computing: Principles and Paradigms*. Wiley Press.

    This book provides a comprehensive introduction to cloud computing principles and paradigms. It covers various aspects of cloud computing, including architecture, service models, deployment models, and the challenges associated with cloud adoption.

4. Rittinghouse, J. W., & Ransome, J. F *Cloud Computing: Implementation, Management, and Security*. CRC Press.

    This book focuses on the practical aspects of implementing, managing, and securing cloud computing environments. It includes case studies, best practices, and guidelines for effective cloud adoption and management.