World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(REVIEW ARTICLE)

Check for updates

Scalable AI and data processing strategies for hybrid cloud environments

Dhruvitkumar V Talati *

Independent Researcher, USA.

World Journal of Advanced Research and Reviews, 2021, 10(03), 482-492

Publication history: Received on 18 May 2021; revised on 20 June 2021; accepted on 23 June 2021

Article DOI: https://doi.org/10.30574/wjarr.2021.10.3.0289

Abstract

Hybrid cloud infrastructure is increasingly becoming essential to enable scalable artificial intelligence (AI) as well as data processing, and it offers organizations greater flexibility, computational capabilities, and cost efficiency. This paper discusses the strategic use of hybrid cloud environments to enhance AI-based data workflows while addressing key challenges such as latency, integration complexity, infrastructure management, and security. In-depth discussions of solutions like federated multi-cloud models, cloud-native workload automation, quantum computing, and blockchain-driven data governance are presented. Examples of real-world implementation case studies in industries including healthcare, retail, finance, and manufacturing are provided to prove the real benefit of hybrid cloud adoption. New trends like explainable AI (XAI), automated machine learning (AutoML) and federated learning are also discussed here as key enablers of future hybrid cloud expansion.

Keywords: Hybrid Cloud Computing ; AI-Driven Cloud Solutions; Machine Learning (ML); Artificial Intelligence (AI); Cloud Computing Architectures; Data Processing Optimization; Scalable AI Workflows; Cloud Migration Strategies

1 Introduction

The rapid advancement of AI, ML, and big data analytics has revolutionized industries at an exponential scale of data-driven operations. With businesses depending increasingly on these technologies to enhance decision-making, automate processes, and enhance services, demand for scalable computing infrastructure has grown manifold. To address these changing requirements, organizations are increasingly adopting hybrid cloud infrastructures, which allow the agility and flexibility necessary to drive sophisticated AI workloads as well as big-data processing.

Hybrid cloud strategies bring together public and private clouds in such a way as to enable organizations to align their resources in a way aligned to their particular needs. It allows companies to leverage the scalability and cost advantage of public cloud services but with security, compliance, and control through on-premises data centers or private cloud infrastructure. By combining these infrastructure elements, hybrid cloud models provide an end-to-end solution that contributes to business responsiveness without compromising data security and compliance risks.

AI workloads tend to need massive computing resources, fast data processing, and elastic scaling in line with varying workloads. AI workloads such as real-time consumption of data, model training, inference, and predictive analysis induce considerable infrastructure and operational complexities . Traditional on-prem or strictly public-cloud-only architecture is prone to problems such as latency, reduced levels of processing capacity, resources utilized, and lack of scale-out .

Hybrid cloud infrastructure offers a practical solution to all these issues through the combination of the benefit of public and private cloud environments. They improve operational effectiveness, reduce AI model execution latency, and allow

* Corresponding author: Dhruvitkumar V Talati ORCID ID :0009-0005-2916-4054

ease of data pipeline orchestration across environments. Hybrid cloud environments also enhance disaster recovery and business continuity through quicker data recovery across distributed infrastructure.

Notwithstanding these benefits, AI-based workflows in hybrid clouds are challenged by serious issues. Businesses are challenged by heterogeneous system integration, multi-cloud workload orchestration, real-time synchronization of data, and strict regulatory compliance norms . These issues are addressed by intelligent orchestration platforms and automation solutions to make hybrid cloud management agile, secure, and efficient.

Security and compliance are always a top concern, particularly for regulated sectors like finance, healthcare, and retail. Successful hybrid cloud initiatives need to integrate strong security architectures, end-to-end data governance processes, and real-time monitoring of compliance to counter increasing cybersecurity threats and adhere to strict regulatory requirements like GDPR and CCPA.

Machine learning and predictive analytics-backed automation are fundamental in hybrid cloud optimization. Machine learning and automation optimize resource assignment, automate workload orchestration, guarantee data purity, and enhance dynamic compliance, leading to efficiency, cost-saving, and safety. Besides, new innovations such as quantum computing, federated learning, blockchain, AutoML, and explainable AI (XAI) present new opportunities to boost hybrid cloud scalability, security, and transparency.

2 Problems with Hybrid Cloud Strategies for AI and Data Processing

Hybrid cloud approaches improve the scalability and processing of data by AI, with significant operational benefits. They are, however, difficult to implement in technological, organizational, and regulatory spaces. The section highlights these problems based on latency challenges, data integration complications, infrastructure management, security loopholes, regulatory compliance, and cost management.

2.1 Latency and Performance Limitations

Latency is a major issue in hybrid cloud infrastructure, especially for AI workloads that demand real-time processing and analytics. AI-driven decision-making relies on quick data ingestion, transformation, inference, and feedback loops.

- Network Latency: Merging on-premises and public cloud infrastructures can lead to network latency, decreasing AI application responsiveness and affecting real-time analytics and decision-making.
- Cold-Start Latency: Serverless cloud computing in hybrid clouds suffers from cold-start latency, which slows down real-time analysis, impacts consumer experience, and interrupts operational effectiveness.
- \tWaste Resource Allocation: Inefficient allocation of computational resources leads to latency problems, particularly for scenarios with dynamic resource allocation in handling variable AI workloads.

2.2 Complexity of Data Integration

Operating on various data sources and providing real-time data streaming on hybrid clouds is one of the primary challenges. Proper integration of data is required for accuracy, completeness, and consistency in AI-driven analytics and decision-making.

- Multi-Source Data: Hybrid clouds need to manage structured and unstructured data across multiple platforms, devices, and data types. Such heterogeneity makes data standardization, cleansing, transformation, and integration challenging, decreasing the precision of AI models.
- Real-time Data Synchronization: Continuous synchronization of data from on-premises to cloud infrastructures poses operational challenges. Latency or mismatching affect the real-time and accuracy of AI-based insights .
- Sophisticated Workflow Orchestration: Handling complex data pipelines in hybrid clouds demands advanced orchestration platforms. These need to dynamically manage data relationships and pipeline workflows on heterogeneous cloud platforms for effective processing.

2.3 Hybrid Infrastructure Management Challenges

Hybrid infrastructure management is extremely challenging due to the requirement for centralized control and optimization of resources across various environments.

- Resource Management and Allocations: Dynamic nature of AI workloads creates fluctuating demands on computer resources, thus making infrastructure management and resource allocation challenging, causing inefficiency or unavailability of resources.

- Orchestration Across Platforms: Hybrid cloud infrastructure management necessitates sophisticated orchestration to unify resources, workloads, and data streams across many cloud providers and on-premises data centers.

- Insufficiency of Homogeneous Monitoring Tools: Organizations struggle to monitor hybrid infrastructures because there are no homogeneous tools to monitor performance, security, and cost optimization on different environments, and this raises operational overhead and security threats .

2.4 Security and Privacy Issues

Hybrid cloud models pose serious security and privacy threats because they are decentralized and complicated.

- Greater Attack Surface: Hybrid environments possess greater attack surfaces that expose them to security breaches, unauthorized access, and data exfiltration. Strong security controls will be necessary to neutralize such changing cyber threats.

- Privacy Risk in Data: Sensitive information stored on various cloud platforms has privacy risk, and tight controls need to be implemented to satisfy global privacy laws, e.g., GDPR and CCPA.

- Complexity of Compliance: Hybrid environments complicate compliance, especially for cross-border data flow. Organizations have to manage various data sovereignty legislations and regulatory environments, which need advanced governance mechanisms.

2.5 Cost Management and Optimization

Cost minimization is a fundamental challenge in hybrid cloud planning, particularly with unpredictable AI workloads and variable resource demands.

- Unpredictable Cloud Expenses: Unpredictability in AI workloads brings uncertainty in infrastructure expenses, thereby complicating budgeting and cost control, especially in pay-as-you-go cloud environments.

- Infrastructure Inefficiencies: Without smart resource management and predictive analytics, organizations are typically bogged down by inefficiencies, such as under-allocation or over-provisioning of resources, driving up operating costs.

- Complexity of Financial Governance: Cost management in the cloud across hybrid environments demands robust financial governance mechanisms that can track, assign, and optimize costs against discrete cloud and on-premises infrastructure .

2.6 Regulatory Compliance and Governance

Hybrid cloud architecture compliance management is especially challenging in highly regulated sectors such as finance, healthcare, and retail.

- Cross-Border Compliance: Hybrid clouds across jurisdictions need to address data localization, sovereignty, and cross-border data transfer concerns. Compliance with intricate regulatory situations incurs additional operating and governance costs .

- Auditing and Accountability Challenges: Compliant assurance through open auditing and accountability necessitates sophisticated monitoring and reporting infrastructure to monitor regulatory compliance in real-time across platforms.

- Data Governance Complexity: Sophisticated firm data governance controls are needed to manage varied regulatory requirements, data privacy regulations, and security policies. Blockchain-based audit programs and smart compliance automation tools can support governance .

2.7 Conclusion

Overcoming these issues efficiently is imperative for successful deployment of hybrid cloud approaches to scalable AI and data processing. Organizations need to adopt advanced ML-based solutions, smart orchestration platforms, and cutting-edge technologies like blockchain, federated learning, quantum computing, and AutoML to achieve optimal performance, security, compliance, and operational and cost benefits.

3 ML-Driven Solutions and Best Practices for Hybrid Cloud Strategies

Hybrid cloud complexities are to be dealt with by utilizing the latest machine learning (ML) solutions and industry-standard best practices. The following section is an overview of ML-driven solutions that are capable of optimizing performance, scalability, security, compliance, and cost-effectiveness within hybrid cloud infrastructures.

3.1 Intelligent Resource Allocation and Dynamic Scalability

ML-based predictive management greatly improves hybrid cloud performance via adaptive scaling of resources in response to actual workload needs.

3.1.1 Predictive Resource Scaling

Sophisticated prediction models such as LSTM and ARIMA enable organizations to forecast changes in workload in advance, thus optimizing resource utilization, minimizing latency, and keeping infrastructure costs at minimum.

3.1.2 Quantum-Enhanced Scalability

Quantum computing optimizes forecasting performance for improved speed and accuracy in predicting the demand for resources, vastly improving performance on hybrid cloud platforms.

3.1.3 ML-Based Reinforcement Learning (RL)

Reinforcement learning learns incessantly from history data and current metrics to learn resource usage optimally better, enhancing system responsiveness and minimizing costs.

3.2 Advanced Workflow Orchestration and Automation

AI-based orchestration streamlines hybrid cloud workflows with minimized data sync and AI pipeline management.

3.2.1 AI-Driven Orchestration Frameworks

ML streamlines hybrid cloud workflows by optimizing data transfer and AI model runtimes across several cloud providers and on-premises infrastructure, minimizing operation complexity.

3.2.2 Automated Real-Time Monitoring and Anomaly Detection

ML-powered monitoring systems detect and rectify workflow anomalies automatically prior to affecting performance, maintaining hybrid cloud stability.

3.2.3 Federated Orchestration

Federated learning optimizes workflow orchestration in distributed hybrid environments by minimizing latency, enhancing data privacy compliance, and maximizing performance.

3.3 Automated Data Integration and Quality Assurance

AI and analytics require high-quality data in hybrid cloud environments.

3.3.1 Automated Data Cleansing and Transformation

ML-powered pipelines cleanse and transform data from diverse sources, making it standardized and reliable in hybrid cloud infrastructures.

3.3.2 Real-Time Data Validation

ML-powered validation platforms automatically validate and correct data streams, enhancing data quality and AI model precision.

3.3.3 Anomaly Detection and Correction

Organizations employ ML to identify and correct data inconsistencies to enable high-quality input data to be provided for AI-driven insights.

3.4 Strengthening Security and Compliance with ML

ML-driven security enhances cybersecurity strength and compliance within hybrid environments difficult to manage.

3.4.1 AID-based Cybersecurity

Adversaries are tracked in real-time, and anomalies are picked up by identifying patterns using ML algorithms, ensuring intrusion prevention for hybrid environments from cyberattacks and data breaches.

3.4.2 Machine Learning-Powered Zero-Trust Architectures

Zero-trust architectures fueled by AI dynamically authenticate workloads, devices, and users, which enhances risk management, regulatory compliance on privacy, and security.

3.4.3 Blockchain-Based Governance

Blockchain offers transparent, immutable tracking of compliance to enhance regulatory reporting and support governance for hybrid clouds

3.5 Cost Optimization and Financial Governance

Financial plans fueled by ML reduce hybrid cloud costs while maintaining costs in check

3.5.1 Predictive Cost Management

Forecasted resource needs using predictive analytics enable companies to drive infrastructure planning ahead and prevent wasteful spending.

3.5.2 AutoML frameworks to manage cost

AutoML functionality simplifies model building and deployment with minimal human interventions while ensuring that cloud resources are optimized for use

3.5.3 Federated learning for optimal cost management

Federated learning saves centralized computational costs by reducing data transfer charges, leading to substantial hybrid cloud cost savings

3.6 Compliance Automation of Regulations

ML-based regulatory compliance frameworks automate the compliance process to simplify governance for hybrid clouds .

3.6.1 Blockchain-Based Compliance Auditing

Blockchain-based architectures provide clear, unerasable compliance logs, which automates compliance report audit trails .

3.6.2 ML-Based Real-Time Compliance Monitoring

ML models keep on monitoring hybrid clouds for compliance violations, detecting, reporting, and rectifying them automatically to ensure ongoing compliance.

3.6.3 Federated Learning for Privacy Compliance

Federated learning allows secure cross-jurisdictional data processing while ensuring compliance with privacy regulation without sacrificing its analytics function .

Applying these newer ML-based approaches, organizations can thoroughly address the matters of hybrid clouds, scalability, performance, security, compliance, and cost optimization. These optimal practices unleash the full potential of hybrid cloud infrastructure, delivering an overwhelming impact to operations, rivalry, and capital.

4 Industry Applications and Case Studies

This chapter captures actual industry case studies that showcase the effective implementation and strategic benefits of ML-based hybrid cloud infrastructures. The cases determine how companies in various sectors such as retail, healthcare, financial institutions, and manufacturing have implemented smart resource management, workflow automation, intelligent security functionalities, and compliance-led innovations to improve performance, scalability, and cost savings in hybrid cloud setup .

4.1 Retail Sector: Influencing Customer Personalization within the Hybrid Cloud

A leading international retailer effectively implemented a hybrid cloud approach to automate data engineering processes and optimize AI-driven customer personalization. The implementation enabled improved real-time customer interaction, improved inventory forecast accuracy, and reduced operational expenditure .

4.1.1 Major Deployments and Impacts

AI-Driven Real-Time Recommendations

Through the bundling of ML-powered pipelines onto public and private cloud infrastructure, the retailer lowered recommendation latency by a substantial percentage, achieving a 40% increase in personalized recommendation accuracy.

Automated Data Validation and Quality Control

Real-time ML-based data validation application enhanced the accuracy and consistency of consumer information, boosting personalization initiatives and marketing ROI improvement by about 35% .

Cost-Optimized Resource Allocation

Predictive analytics enabled effective cloud resource planning in advance, lowering infrastructure costs by almost 30%, even during peak shopping seasons .

4.2 Healthcare Sector: AI-Driven Medical Data Processing

One of the prominent healthcare providers used hybrid cloud infrastructure to facilitate enhanced AI-based processing and analysis of patient data. It improved data security, streamlined compliance management, and facilitated enhanced operational efficiency .

4.2.1 Key Implementations and Their Effects

Federated Learning for Patient Data Security

The provider embraced federated learning methods, allowing secure, privacy-preserving analytics on decentralized patient data in multiple cloud and on-premises environments. This allowed for compliance with rigorous laws such as HIPAA and GDPR.

Real-Time Medical Diagnostics

The ML-fueled hybrid infrastructure sped up medical image and patient record analysis, reducing diagnostic time from hours to minutes and improving patient outcomes significantly .

Blockchain for Compliance and Auditing

Implementation of blockchain-based frameworks delivered transparent, immutable logs for auditing and compliance, lowering regulatory complexity and saving operational expenses by up to 25% .

4.3 Financial Services: Scalable Hybrid Cloud for Fraud Detection

A global financial organization utilized hybrid cloud techniques to optimize scalability and performance of real-time AI-based fraud detection models. The solution increased security, regulated compliance, and business efficiency.

4.3.1 Critical Implementations and Results

Dynamic Resource Allocation

Machine learning-based predictive analytics and reinforcement learning techniques optimized hybrid cloud resource allocation to support real-time fraud detection at scale and reduced latency by more than 45% .

Real-Time Cybersecurity Threat Detection

AI-based cybersecurity tools tracked hybrid cloud activities in real-time, detected, and neutralized threats in real-time. This reduced cybersecurity threats by 50% .

Decrease in Operating Expenses:

AutoML platforms mechanized training, deployment, and AI model optimization, reducing manual workload tremendously and reducing operating expenses by around 30% .

4.4 Production: Increased Production and Supply Chain Efficiency

A global manufacturing firm employed a hybrid cloud strategy to facilitate AI-driven analytics, automating manufacturing and supply chain processes. The change increased business agility, scalability, and cost reduction .

4.4.1 Key Implementations and Results

Quantum Computing for Supply Chain Optimization

Quantum-fueled predictive analytics greatly improved production planning, inventory management, and supply chain responsiveness, lowering operational expenses by almost 35% .

Automated Workflow Management

Smart ML-driven orchestration solutions effortlessly handled intricate manufacturing workflows in public cloud and on-premises environments, lowering workflow latency and increasing productivity by 40%.

Real-Time Compliance Monitoring

Blockchain-compliance solutions facilitated automatic tracking and reporting of regulations, with real-time, unalterable compliance history. This made regulatory processes simpler and lowered the costs of compliance by approximately 25%.

Collectively, these case studies underscore the potential for transformation through ML-driven hybrid cloud solutions across sectors. Through the use of such smart solutions, organizations are, in fact, eliminating hybrid cloud intricacies and gaining impressive improvements in scalability, cost savings, response time, security, and compliance.

5 Future Directions and Emerging Trends

The explosive development of hybrid cloud architectures driven by improvements in artificial intelligence (AI), machine learning (ML), and analytics opens enormous opportunities for organizations to attain higher scalability, security, operational efficiency, and competitiveness. Technologies like federated learning, quantum computing, explainable AI (XAI), blockchain-based compliance mechanisms, and serverless computing are revolutionizing AI-driven data processing in hybrid cloud environments .

5.1 Federated Learning: Scalability and Data Privacy

Federated learning (FL) is transforming AI model training to facilitate decentralized learning on distributed datasets without undermining data privacy and reducing centralized data transmission. FL enhances data protection regulation compliance, reduces latency, and minimizes data transfer costs.

5.1.1 Key Developments

Privacy-Centric AI Training

FL allows AI models to learn locally without exposing raw data, minimizing the potential for security risks, simplicity in handling compliance, and building regulators' and users' trust.

Cost-Effective Data Transfer

Through the removal of the need for enormous datasets to be transferred across cloud platforms, FL markedly minimizes operational costs while enhancing processing speed and efficiency when training high-volume AI .

Regulatory Compliance in Global Operations

Federated learning frameworks keep data in jurisdictional silos and securely exchange model insights, respecting cross-border data governance norms

5.2 Quantum Computing: Enhancing AI Workflows

Quantum computing is proving to be a game-changer for AI-powered hybrid cloud platforms, going beyond classical computational constraints and supporting accelerated processing of complex tasks.

5.2.1 Major Breakthroughs

Optimized Computational Resource Allocation

Quantum-driven techniques facilitate fast and accurate resource deployment, resulting in lower infrastructure costs and improved performance in hybrid cloud environments.

Advanced Predictive Analytics

Quantum-driven models enhance predictive accuracy by an enormous extent in high-complexity situations like supply chain management, medical imaging, and credit risk evaluation.

Scalable AI Model Training

With the use of quantum computing, hybrid cloud infrastructures are able to accommodate more and complex AI models without any costs or performance compromise.

5.3 Blockchain: Automation of Compliance and Security Boost

Blockchain technology is being deployed in hybrid cloud more and more to automate compliance seamlessly, ensure data integrity, and introduce transparency into governance.

5.3.1 Key Applications

Automated Compliance Management

Blockchain-based solutions provide real-time auditing and regulatory monitoring, minimizing manual intervention while offering ideal compliance with changing compliance requirements.

Improved Cybersecurity

Immutable data rule-based control implemented using blockchain enhances cybersecurity on hybrid cloud environments, which safeguards against likely data breach and unauthorized alteration attacks.

Accurate Data Auditing and Governance

Blockchain adoption guarantees traceable origin of data, automated tracking of compliance, and simpler regulatory reporting, improving overall operating clarity.

5.4 Serverless Computing and AutoML

AI Model Deployment Revolutionized Serverless computing and Automated Machine Learning (AutoML) are leading a paradigm shift that is providing unparalleled flexibility, lower operational complexity, and effective AI model deployment in hybrid clouds.

5.4.1 Breakthroughs

On-Demand Scalability and Cost Savings

Serverless architecture dynamically allocates computing resources, cutting costs by dynamically scaling workloads in real-time depending on demand.

Streamlined AI Model Lifecycle Management

AutoML platforms automate model construction, hyperparameter tuning, and deployment, minimizing human effort and time-to-market for AI-powered solutions.

Flexible AI Operations

The serverless auto-scaling features combined with AutoML help enterprises deploy and run AI workloads with absolute provisioning of massive infrastructure, vastly reducing operational burden .

5.5 Explainable AI (XAI): Enforcing Trust and Translucency

Explainable AI (XAI) is rapidly growing as critical compliance, ethical adoption of AI, as well as AI decision confidence attainment, specifically for healthcare and financial domains.

5.5.1 Principal Advantages

Bonded Transparency of AI Decisions

XAI systems can be used to explain AI model decision-making, facilitating greater accountability and simpler auditing, especially in highly heavily regulated sectors.

Regulatory Compliance Alignment

By explaining the decisions made by AI, XAI allows organizations to satisfy transparency and equity requirements, lowering legal risks as well as compliance framework adherence.

Consumer Confidence Building

Providing transparent descriptions of AI decision-making leads to trust, enhances user adoption, and builds brand value in competitive economies.

5.6 Edge Computing: Hybrid Cloud Performance Amplification

Edge computing is among the evolving trends that maximizes the power of the hybrid cloud by reducing latency, efficient resource utilization, and optimizing operational efficiency for AI workloads at the network's edge.

5.6.1 Key Advances

Low-Latency AI Processing

By processing AI workloads near data sources, edge computing minimizes processing latency and enables real-time decision-making for use cases such as IoT, intelligent retail, and industrial automation.

Scalability and Load Distribution

Emplacing AI at the edge de-loads loads on centralized cloud infrastructures, thereby enabling hybrid models to handle increasing loads better.

Cost-Optimized Resource Allocation

Smart edge resource management minimizes repetitive data transfer to a minimum level, decreasing cloud computing cost and enhancing overall efficiency.

5.7 Conclusion

These emerging technologies will reimagine the ability of hybrid cloud infrastructures to provide more scalability, efficiency, security, and compliance. Organizations that invest in these technologies based on informed decisions will be well underway in actualizing the full potential of AI-powered hybrid cloud environments in the data-dominated age.

6 Conclusion

The evolution of hybrid cloud strategies has become essential for organizations aiming to achieve scalable, cost-efficient AI and data processing capabilities. By effectively integrating public and private cloud resources, businesses have successfully tackled key challenges, including latency, infrastructure complexity, data security, regulatory compliance, and operational efficiency. The adoption of advanced technologies such as machine learning (ML), predictive analytics, serverless computing, federated learning, blockchain governance, quantum computing, and explainable AI (XAI) has driven significant improvements in scalability, responsiveness, and resource optimization—strengthening their competitive position in data-driven industries.

Across industries like retail, healthcare, financial services, and manufacturing, case studies illustrate the transformative impact of hybrid cloud architectures. These real-world implementations reveal how AI-driven automation, predictive analytics, and intelligent resource management enable organizations to lower latency, optimize infrastructure costs, and enhance overall operational efficiency. Moreover, automated compliance solutions and robust security mechanisms mitigate regulatory and cybersecurity risks, providing enterprises with greater confidence in navigating complex regulatory environments.

The future of hybrid cloud computing will be shaped by emerging innovations such as serverless AutoML, federated learning, and edge computing. These technologies will continue to refine hybrid cloud ecosystems by addressing challenges related to privacy, compliance, latency, and computational efficiency. As hybrid cloud solutions become more intelligent, automated, and responsive, organizations will gain unprecedented flexibility in managing AI-driven workloads.

To remain competitive in an increasingly data-centric world, businesses must embrace ML-powered hybrid cloud advancements. By adopting these technologies, enterprises can streamline AI operations, enhance scalability, and ensure seamless compliance—positioning themselves for long-term success, operational agility, and sustainable growth in the digital economy.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., … & Zaharia, M. (2011). A view of cloud computing. Communications of the ACM, 53(4), 50-58. https://doi.org/10.1145/1721654.1721672

[2] Dean, J., & Ghemawat, S. (2011). MapReduce: A flexible data processing tool. Communications of the ACM, 53(1), 72-77. https://doi.org/10.1145/1629175.1629198

[3] Gantz, J., & Reinsel, D. (2012). The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east. IDC Analyze the Future.

[4] Satyanarayanan, M. (2012). A brief history of cloud offload: A personal journey from Odyssey through cyber foraging to cloudlets. GetMobile: Mobile Computing and Communications, 18(4), 19-23. https://doi.org/10.1145/2594368.2594373

[5] Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Khan, S. U. (2015). The rise of "big data" on cloud computing: Review and open research issues. Information Systems, 47, 98-115. https://doi.org/10.1016/j.is.2014.07.006

[6] Buyya, R., Ranjan, R., & Calheiros, R. N. (2010). InterCloud: Utility-oriented federation of cloud computing environments for scaling of application services. In International Conference on Algorithms and Architectures for Parallel Processing (pp. 13-31). Springer. https://doi.org/10.1007/978-3-642-13119-6_2

[7] Marz, N., & Warren, J. (2015). Big data: Principles and best practices of scalable real-time data systems. Manning Publications.

[8] Zaharia, M., Chowdhury, M., Franklin, M. J., Shenker, S., & Stoica, I. (2012). Spark: Cluster computing with working sets. HotCloud, 10(10-10), 95.

[9] Zhang, Q., Cheng, L., & Boutaba, R. (2010). Cloud computing: state-of-the-art and research challenges. Journal of Internet Services and Applications, 1(1), 7-18. https://doi.org/10.1007/s13174-010-0007-6

[10] Chang, F., Dean, J., Ghemawat, S., Hsieh, W. C., Wallach, D. A., Burrows, M., … & Gruber, R. E. (2008). Bigtable: A distributed storage system for structured data. ACM Transactions on Computer Systems (TOCS), 26(2), 4. https://doi.org/10.1145/1365815.1365816

[11] Lin, J., & Dyer, C. (2010). Data-intensive text processing with MapReduce. Morgan & Claypool Publishers.

[12] Shvachko, K., Kuang, H., Radia, S., & Chansler, R. (2010). The Hadoop distributed file system. In 2010 IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST) (pp. 1-10). IEEE. https://doi.org/10.1109/MSST.2010.5496972

[13] Fernández, A., del Río, S., López, V., Bawakid, A., del Jesus, M. J., Benítez, J. M., & Herrera, F. (2014). Big Data with Cloud Computing: An Insight on the Computing Environment, MapReduce, and Programming Frameworks. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 4(5), 380-409. https://doi.org/10.1002/widm.1134

[14] Meng, X., Bradley, J., Yavuz, B., Sparks, E., Venkataraman, S., Liu, D., … & Zaharia, M. (2016). MLlib: Machine learning in Apache Spark. Journal of Machine Learning Research, 17(1), 1235-1241.

[15] Yu, W., Liang, F., He, X., Hatcher, W. G., Lu, C., Lin, J., & Yang, X. (2017). A survey on the edge computing for the Internet of Things. IEEE Access, 6, 6900-6919. https://doi.org/10.1109/ACCESS.2017.2778504

[16] Veiga, L., & Pires, J. (2018). Hybrid cloud and edge-enabled scalable AI systems. Future Generation Computer Systems, 98, 738-747. https://doi.org/10.1016/j.future.2018.04.039

[17] Ghosh, R., Naik, V., & Ranjan, R. (2016). SLA-aware autonomic management of cloud resources using online reinforcement learning. IEEE Transactions on Cloud Computing, 5(1), 124-137. https://doi.org/10.1109/TCC.2015.2453983

[18] Xu, L. D., He, W., & Li, S. (2014). Internet of Things in industries: A survey. IEEE Transactions on Industrial Informatics, 10(4), 2233-2243. https://doi.org/10.1109/TII.2014.2300753

[19] Kune, R., Konugurthi, P. K., Agarwal, A., Chillarige, R. R., & Buyya, R. (2016). The anatomy of big data computing. Software: Practice and Experience, 46(1), 79-105. https://doi.org/10.1002/spe.2374