Cloud Trends 2025: Unveiling the Future Of Cloud Technology

Dr. Magesh Kasthuri, Vishnu Vardhan Yalala, Veerabadrachari R, Vasudha Mishra, Anuj Mehra and Sourabh

December 2024

Trend report from DMTS Community

Table Of Contents

Introduction

The ever-evolving landscape of cloud technology continues to shape the future of businesses, driving digital transformation and innovation. As we approach 2025, several key trends are poised to dominate the cloud technology arena, emerging with renewed vigor from 2024. Cloud technology is not new age and its one of the technology which emerges in the last decade continuously to keep itself up-to-the-mark of the expectations of the industrial need and digital transformation trends.

Cloud trends is a three dimensional approach such as

- Hyperscalers introduced new services and enhancements in the existing services in the last few years to keep them up in the market demand and meet customer needs at scale
- Domain expectations builds usecases which are industry specific and thus introducing domain centric advancements like Industrial cloud and
- Combinational technology solutions like Data centric cloud services, Generative AI cloud services and Quantum cloud services.

This research report delves into the top ten cloud trends expected to make a significant impact, supported by insights from leading analyst reports such as Gartner, Forrester, Everest, and IDC. This report is distinct in that the authors bring the trend analysis from Analyst reports point of view as horizontal view and understanding on business trends from customers as vertical view.

Multi-Cloud and Hybrid Cloud Solutions (Future is Hybrid)

We can understand that there are many factors influencing the decision including Geo spread Customer portfolio and IT teams, distributed technology stack of applications, reusability of hybrid services for gradual cloud adoption of multiple applications in the estate to name a few.

In a Multi-cloud architecture, three most commonly used pattern, which can influence to choose Multi-cloud solutions as explained below three most commonly:

Data oriented multi-cloud architecture – This is a heterogeneous pattern, where enterprise choose multi-cloud based on suitability of data sources. For example, if the application estate has Oracle, Postgre, MySQL, MSSQL and customer prefers not to do actual DB transformation (Oracle to MSSQL or MSSQL to Postgre), then we can choose applications with Oracle, Postgre, MySQL to go to AWS (where Oracle to PSQL is quite easy in AWS/Azure) and MSSQL based apps can go to Azure.

For applications (batch, real-time analytics) using NoSQL database platforms are quite suitable to move to any CSP (eg: MongoDB, Hadoop Cluster) and only cost comparison (for selected region(s), zone(s)) can be compared for multi-cloud adoption.

Service oriented multi-cloud architecture – This is a homogeneous pattern, where enterprise choose multi-cloud based on business applications including services, API layers and processing logic (functions) where cross service communication across CSP and centralized service availability for multi-cloud is also possible. For example, a polyglot Microservice groups can be migrated to AWS, Azure, GCP depending on service usage and need (eg: Apigee services can go to GCP, Service orchestration using Fabric can go to Azure and Hub services can goto AWS).

Process oriented multi-cloud architecture – This is a hybrid pattern, where enterprise choose multi-cloud based on combination of Data oriented apps or service oriented apps or in combination of both. They are only concerned about the ease of portability to cloud and flexibility in implementation.

For example, ERP or SAP Hana apps can be migrated to AWS along with services, Frontend apps can run on Azure and at the same time, Batch processing with Analytics and reporting can run on GCP.

Distributed Cloud vs Multi-cloud solutions

Distributed cloud refers to the deployment of cloud services across multiple geographic locations, while being managed centrally from a single platform. This architecture allows for data and applications to be physically closer to end-users, reducing latency and improving performance. By spreading out the resources, distributed cloud enhances reliability and scalability, ensuring continuous service availability even if one location faces disruptions.

Service deployment in a distributed cloud environment involves strategically placing computational resources, storage, and networking capabilities at various edge locations. These resources are then orchestrated and managed by a central platform that oversees operations, updates, and security policies. This central management simplifies the complexity of handling multiple sites while leveraging the benefits of localized computing.

Multi-cloud solutions, on the other hand, involve using multiple cloud service providers to host different parts of an application or workload. This approach helps organizations avoid vendor lock-in, increase redundancy, and optimize cost by selecting the most suitable services from various providers. Multi-cloud strategies are particularly beneficial in addressing regulatory requirements across different regions and industries.

For instance, data sovereignty laws may dictate that certain data must remain within specific geographic boundaries. By utilizing a multi-cloud approach, companies can ensure compliance by storing and processing data in regions that meet regulatory standards. Additionally, multi-cloud solutions offer flexibility to adopt best-of-breed technologies tailored to particular industrial applications, enhancing overall efficiency and innovation.

In summary, distributed cloud and multi-cloud strategies are reshaping cloud computing by offering improved performance, compliance, and flexibility. These approaches enable businesses to better meet the demands of modern applications and regulatory landscapes, ensuring robust and adaptable cloud infrastructures.

A recent case study on Wipro's customer in multi-cloud and hybrid cloud solutions involves a European client who embarked on a multi-cloud transformation journey. The client adopted a business case-driven application rationalization approach, transforming 77% of their workload into the cloud with the adoption of container and application modernization.

Benefits to the Customer:

Seamless Support: Wipro supported the application portfolio without any impact on the applications migrated to the public cloud under the Re-Host model.

Upskilled Workforce: For applications that required re-factoring and architectural changes, Wipro upskilled the existing application support team, ensuring continuous application services in the transformed environment.

Efficient Transformation: The collaborative exercise between Wipro and the client facilitated a seamless transition of application support without creating any business impacts.

Wipro's Unique Proposition:

- Studio Framework and Accelerators: Wipro leveraged its Studio framework, accelerators, BOTS, and predefined migration run books along with ITIL-led operating procedures to carry out the migration efficiently.
- Cloud Business and Migration Office: Wipro's Cloud Business and Migration office strategized and planned for large-scale execution, ensuring a smooth migration process.
- Automated Migration and Testing: The use of pipeline-driven automated migration and testing in an agile mode helped fast-track the deployment cycle.
- Detailed Assessment: Performing a detailed assessment of infrastructure and applications helped derive the right business use cases, ensuring a tailored approach to the client's needs.

Edge Computing

Edge computing is set to revolutionize the cloud landscape by bringing data processing closer to the source of data generation. This trend is driven by the growing demand for low-latency applications and real-time data processing. According to IDC, the global edge computing market is projected to reach $250 billion by 2025, reflecting a compound annual growth rate (CAGR) of 37.4% from 2024. The proliferation of Internet of Things (IoT) devices and the need for real-time analytics are key factors propelling the adoption of edge computing.

As per Gartner’s Strategic Technology trend report 2024, one of the technology that can drive significant disruption and futuristic opportunity for IT services for the next five to ten years is Distributed Cloud.

Currently most Cloud Service Providers (CSPs) which runs centralized Cloud service model and offers On-demand and metered access to the computing services like database, virtual machines and storage services across different regions, availability zones and edge locations. For better RPO and RTO, site redundancy and Multi-AZ services are provided to ensure resilience and high availability.

But the future of Cloud service model breaks this centralized model and move towards distributed Cloud model where CSPs data center will not be a barrier for storage locations. This will be one more level

higher to Hybrid cloud where Hybrid cloud is combination of Private data center based Public cloud farms and Distributed cloud moves towards decentralized model in which various micro-cloud located outside the centralized cloud farms in which each micro-cloud can serve the purpose for Storage, compute and network services in its control by logically disbursed between each other.

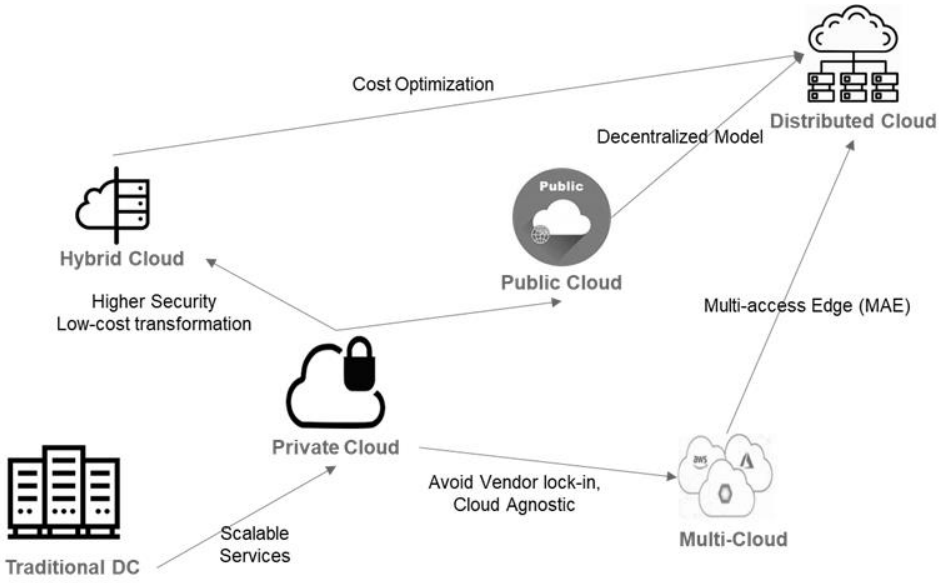


Figure: Cloud migration paths and cross roads

Edge Computing and Fog Computing already moves the cloud model towards Distributed Cloud and with the intent to reduce network latency by keeping the micro-cloud closer to user regions (edge caching based on usability, accessibility).

With the advantage of Distributed Cloud and the entry of 5G architecture where distributed location access is very much needed for requirements like Regulatory (cross-border data access), Location (multi-regional service access with low latency and higher throughput), Security (data privacy) and Redundancy (to handle situation like large scale traffic and outages).

For large enterprises, Distributed cloud can assure shifting from CAPEX to OPEX using on-demand micro-cloud usage with geo spreaded servers and use intelligent transport of data across locations with less latency and intelligent caching for large volume of data like video streaming and Data lake solutions.

For Cloud Architects and Solution consultants, how this differs they Day in the life is quite interesting as Distributed Cloud makes them to think solution coherent to the offerings of micro-cloud services and Non-functional requirement gets addressed in a much sophisticated way in Distributed Cloud as compared to siloes in Centralized cloud model in terms of performance of service communication, network model, API gateway, CDN usage and messaging services.

As per Gartner and NextPlatform, Distributed Cloud doesn't mature over night and take over but it can potentially occupy majority of cloud investment business by 2025. Security features in cloud platform differs from traditional data security in multi-aspect as they are divergent in platform architecture and

also new-age hybrid cloud architecture needs to protect application, data and services in on-premises data center (eg: Data security), cloud services (eg: Server security) and mobile users (eg: client security).

Gartner predicts that Secure Access Service Edge (SASE) could be transformational and next wave enterprise security architecture. SASE is an amalgamation of Network as a Service and Network security as a Service to combine the power of Service Delivery Wide Area Network (SD-WAN), network routing, network segmentation, Firewall as a Service, cloud based security services (Cloud Security as a Service) and WAN optimization.

Figure: SASE on Edge computing

SASE converges the security architecture with Zero Trust Network Access (ZTNA), Cloud Secure Web Gateway (SWG) and Web application and API protection as a Service (WAAPaaS) to deliver network and security features for new age business scenarios like AI enabled services, 5G network slicing and edge computing to improve global scale and operational resilience.

There are four characteristics that is enabled by SASE such as Identity-Driven to enable user and resource (network, cloud) identity, cloud-native architecture to leverage cloud capabilities like agility, scalability and elasticity, supports Edge computing to reduce network latency and better throughput and Globally distributed network to deliver low-latency network services to enterprise edges.

Edge computing gives a distributed cloud computing facility to cloud operated applications thus reducing network latency and improves customer experience. Since Edge computing gives highly-reliable cloud services, there are many usecases across different sectors like traffic management system, autonomous vehicle for real-time vehicle control, Predictive maintenance in Manufacturing, improved Supplychain management in retail, precision agriculture using Drones, live security monitoring in country borders using IoT camera devices are some of the usecases which gets the full benefit of edge computing.

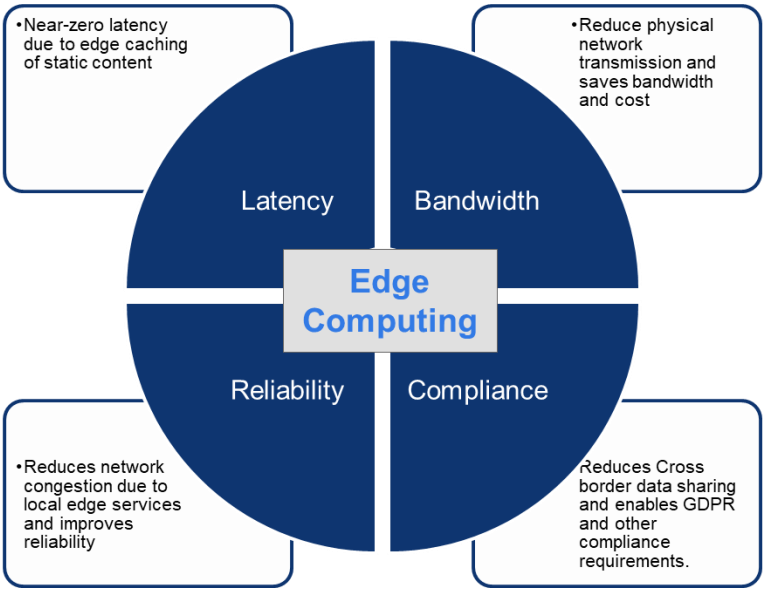


Figure: Edge computing features and benefits

Banking sector is one of the fastest technology adapter as it targets improved customer experience with high-reliable and low-latency services. Hence, Edge computing enables many usecases in Banking industry. Some of them are discussed here:

- Move critical IT operations to edge computing to create lean and portable branches for banks.
- Personalization of customer engagement through improved customer experience through reliable and fast content delivery of static content helps seamless user experience and end-user gets attracts to it in retail banking.
- Fraud detection in user transaction is an important usecase in Banking sector which requires high attention and quicker resolution. Using Edge services, this can be achieved at a faster rate through continuous real-time monitoring and analytics on transaction and detect anomaly in behaviour and alert for fraud transaction faster.
- Enable reliable payment facilities in POS terminals and ATM branches through local edge services.
- For wealth management and asset management, reliable high-frequency algorithmic trading (HFT), plays a critical role as it relies on disparities in price for assets and trade products on different exchanges. Here, edge plays important role to manage trading from local edge locations.

Enabling Ultra Low-Latency Applications, IoT, and Edge Deployments using 5G in Cloud solutions

The integration of 5G technology with cloud computing has the potential to revolutionize various industries by facilitating the development of ultra low-latency applications, the Internet of Things (IoT), and edge computing deployments. This synergy not only enhances the performance of existing cloud-based solutions but also paves the way for innovative applications and services.

Ultra Low-Latency Applications

The advent of 5G technology brings unprecedented speed and latency improvements to wireless communications. With latency as low as 1 millisecond, 5G enables real-time processing and responsiveness, which are critical for applications requiring instantaneous data transmission and

processing. Cloud computing, with its vast computational resources and scalable infrastructure, complements 5G by providing the necessary backend support for these applications.

One of the most notable beneficiaries of this synergy is the field of augmented reality (AR) and virtual reality (VR). The high data transfer rates and low latency of 5G ensure a seamless and immersive experience for users, while cloud computing handles the intensive processing and data storage requirements. This combination allows for more complex and interactive AR/VR applications, such as remote surgeries, virtual training simulations, and real-time collaboration tools.

Internet of Things (IoT)

The proliferation of IoT devices has led to an exponential increase in the volume of data generated. Managing and processing this data efficiently is a significant challenge for traditional computing infrastructures. However, the combination of 5G and cloud computing offers a robust solution to this problem.

5G networks provide the necessary bandwidth and connectivity to support a massive number of IoT devices, ensuring reliable and consistent data transmission. At the same time, cloud computing platforms offer scalable storage and processing capabilities, allowing for real-time analysis and decision-making. This synergy is particularly beneficial for industries such as smart cities, healthcare, agriculture, and manufacturing, where IoT devices play a crucial role in monitoring and optimizing various processes.

Edge Computing Deployments

Edge computing, which involves processing data closer to the source rather than relying on centralized cloud servers, is gaining traction due to its ability to reduce latency and bandwidth consumption. The integration of 5G with edge computing further enhances these benefits, enabling more efficient and responsive applications.

In this context, 5G serves as the backbone for edge computing deployments, providing high-speed connectivity and low-latency communication between devices and edge nodes. Cloud computing, on the other hand, offers the necessary infrastructure to support edge deployments, including data storage, analytics, and machine learning capabilities. This combination is particularly advantageous for applications that require real-time processing and decision-making, such as autonomous vehicles, industrial automation, and remote monitoring systems.

Developing Cloud Technologies by Industries

Industries across various sectors are increasingly recognizing the potential of combining 5G and cloud computing to drive innovation and enhance operational efficiency. By leveraging the strengths of both technologies, businesses can develop new cloud-based solutions that cater to the evolving demands of the market.

For instance, in the healthcare industry, the synergy between 5G and cloud computing enables remote patient monitoring, telemedicine, and real-time diagnostics, improving patient care and reducing healthcare costs. In the manufacturing sector, this combination facilitates predictive maintenance, smart factory automation, and supply chain optimization, leading to increased productivity and reduced downtime.

Similarly, the retail industry can benefit from enhanced customer experiences through personalized recommendations, seamless omnichannel integration, and efficient inventory management. The financial sector, too, stands to gain from improved fraud detection, real-time transaction processing, and enhanced cybersecurity measures.

The synergy between 5G and cloud computing is a game-changer for industries looking to leverage the power of ultra low-latency applications, IoT, and edge deployments. By combining the high-speed, low-latency capabilities of 5G with the scalable and flexible infrastructure of cloud computing, businesses can unlock new opportunities for innovation and growth. As these technologies continue to evolve, their integration will undoubtedly lead to the development of even more advanced and impactful cloud-based solutions.

Data Security and compliance to get data sovereignty to avoid cross border data access from cloud data center and enable data management through local edge computing and storage services.

Wipro has recently implemented Edge solution for a global leader in the life sciences industry. The client aimed to enhance their device connectivity cloud platform and edge gateway to integrate with on-premise medical devices, leveraging cloud PaaS components for improved time to market, quality, and reduced dependency on custom components.

Wipro Solution:

- Re-architecting and Re-designing: Wipro re-architected and re-designed the entire cloud and edge platform to leverage cloud-native PaaS.
- Complex Multi-component Solution: The solution was Azure-based and devoid of proprietary elements, ensuring flexibility and scalability.
- IoT Edge Protocol Translation: Deployed IoT edge protocol translation to cover interfaces like FHIR, HL7, and other medical equipment-specific interfaces.
- Communication Security: Implemented secure communication between medical devices, edge, and cloud using X.509 attestation with an external CA.

Benefits Delivered:

- Accelerated Time to Market: The complex multi-component solution was delivered with an accelerated time to market, leveraging a quick ramp-up of a team of experts.
- Secure and Scalable Architecture: The solution provided a secure, scalable architecture with Microsoft Assurance, managing a volume of around 2000 messages per second.
- Improved Operational Efficiency: The integration of cloud-native PaaS components and IoT edge protocol translation improved the overall operational efficiency and reduced dependency on custom components.

Wipro's Unique Proposition:

- Cloud-First Approach: Wipro's cloud-first approach, based on workload categorization and assessment, ensured a tailored solution for the client.
- Hyper-Converged Solutions: Provided hyper-converged solutions for remote sites, making the edge lighter and more efficient.
- Validated Blueprints and Standardization: Delivered standardized stacks and validated blueprints across regions, ensuring consistency and reliability.
- Strategic Partnerships: Collaborated with strategic partners like Nutanix, Dell, VMware, HPE, Cisco, and hyperscalers (AWS, Microsoft Azure, Google) to enhance the solution's value proposition.

Serverless Computing

Serverless computing, also known as Function as a Service (FaaS), is gaining traction as organizations seek to streamline their development processes and reduce operational overhead. Gartner's analysis indicates that by 2025, serverless computing will be a standard practice for over 50% of global enterprises, compared to 30% in 2024.

The serverless model allows businesses to focus on writing code without worrying about infrastructure management. With serverless applications, code is packaged into small units called “functions” , which gets executed as a managed service. The most commonly used expression of serverless computing is FaaS, such as AWS (Amazon Web Services) Lambda, or Microsoft Azure Functions or GCP (Google Cloud Platform) cloud functions which are being as managed service and computing resources are provisioned specifically for execution purpose and there is a minimal operation controls required on computing resources; and with the growing adoption of cloud computing, customer are looking for ways to consume and operate resources effectively; have agility, scalability and speed to accelerate go to market strategy.

There are many benefits to use serverless computing:

1) Easy to use and consumption based computing resources; and all CSPs provide these services on time metrics ie milliseconds
2) Compute resources can be shut-down to zero when not being used and can be up and running on the fly which enable flexibility and scalability for customers
3) There are very minimal or no close to no efforts required to manage computing platform; resources get provisioned and released automatically during the function execution.
4) Enhance developer experience, wherein developers can focus more on application design, new functionalities creation and configurations than managing infrastructure related management or provisioning etc.
5) Many CSPs provide free-tier which helps developers to experiment new features, functionalities and do rapid experiments with a low entry cost barrier.

While there are many benefits; there are potential areas for consideration:

1) Vendor lock-in – Some of these solution are proprietary to specific providers; in case of application to be moved from one platform to another; this could potentially incur significant efforts for re-engineering and additional cost associated with it.
2) Going monolithic to microservices and adopt serverless architecture required deep technical expertise and experience in redesign architectures, even developing the entirely new code-set for application which can be time-consuming and potentially expensive from TCO prospective as compare to benefits expected.
3) Applications with high CPU and Memory needs could be potentially expensive as compare to other alternates like IaaS or CaaS.
4) The cost associated with API gateway use could go northbound with high rates of function invocation.
5) Customer with fixed budgets need to implement adequate controls for scalability and ensure cost is controlled; as running future cost is depend on the application usage.

It is recommended to engage the expert in transformation initiatives and run through proof of concept (POC) to validate use case and assumptions on application scalability, performance and cost ownership before placing them into production.

Wipro has successfully implemented serverless solutions including a recent Australia-based telecom service provider. The client aimed to create a serverless version of their Incident and Roster Management solution using open-source tools Oncall and Iris provided by LinkedIn.

Benefits to the Customer:

- Cost Efficiency: The serverless solution resulted in no hosting costs, as the client only paid when the system was used.
- Scalability and Flexibility: The solution was designed to support future integration with analytics and monitoring tools like New Relic and repositories like Splunk.
- Enhanced Security: The system was accessed through LDAP/SAML for authentication, ensuring secure access.

Wipro's Unique Proposition:

- Expertise in Serverless Technologies: Wipro leveraged AWS Lambda, SQS, and API Gateway to deploy the solution, showcasing their deep knowledge in serverless technologies.
- Integration Capabilities: The solution was designed to support future integrations with various analytics and monitoring tools, demonstrating Wipro's ability to create flexible and extensible architectures.
- Cost-Effective Solutions: By utilizing a serverless architecture, Wipro provided a cost-effective solution that minimized hosting costs, aligning with the client's budgetary constraints.

Artificial Intelligence and Machine Learning Integration

The integration of artificial intelligence (AI) and machine learning (ML) within cloud platforms is transforming how businesses operate. Cloud providers are enhancing their offerings with advanced AI and ML capabilities, enabling organizations to harness the power of data-driven insights. Forrester's report suggests that by 2025, AI-driven cloud services will see a 40% increase in adoption, as enterprises leverage these technologies for predictive analytics, automation, and personalized customer experiences.

AI-based services on cloud platforms are set to experience remarkable advancements in 2025 and with the development and demand in Generative AI based solutions across multiple industrial applications, these advancements are high in demand from many customers. AWS, Azure, GCP, OCI, and IBM Cloud are all investing heavily in AI, driving innovation and expanding their capabilities. Analyst reports from Gartner and Forrester highlight the potential for these platforms to revolutionize industries, enhance security, and democratize access to AI. As these trends continue to unfold, businesses can look forward to leveraging more powerful, efficient, and ethical AI solutions in the cloud. Some of the key points are discussed below:

AIML services on AWS

Amazon Web Services continues to leverage its robust infrastructure to deliver a broad array of AI services. From machine learning tools like SageMaker to AI-driven analytics, AWS is expected to enhance its offerings, focusing on ease of use and accessibility. Gartner predicts that AWS will maintain its leadership in the cloud AI market, driven by continuous innovation and a strong developer community.

Azure's AI Expansion

Microsoft Azure is anticipated to further integrate AI into its core services. With advancements in natural language processing, computer vision, and predictive analytics, Azure aims to provide businesses with

powerful tools to derive actionable insights from data. Forrester analysts highlight Azure's commitment to democratizing AI, making it more accessible for enterprises of all sizes.

Google Cloud Platform's AI Prowess

Google Cloud Platform is set to capitalize on its strength in data-driven AI. By enhancing its AI and machine learning suite, including AutoML and TensorFlow, GCP is expected to offer more automated and scalable solutions. Gartner's reports suggest that GCP will focus on customization and flexibility, allowing organizations to tailor AI solutions to their specific needs.

Oracle Cloud Infrastructure's Strategic Moves

Oracle Cloud Infrastructure is making strategic investments to bolster its AI capabilities. With a focus on enterprise applications, OCI aims to integrate AI into its core offerings, enhancing performance and efficiency. Forrester notes that Oracle's approach will likely emphasize seamless integration with existing enterprise systems, driving adoption among large corporations.

IBM Cloud's AI Vision

IBM Cloud is leveraging its rich history in AI research to offer cutting-edge services. Watson, IBM's flagship AI, is expected to become even more sophisticated, providing advanced analytics and cognitive computing capabilities. Gartner's analysis indicates that IBM will prioritize AI-driven innovation in healthcare, finance, and other critical industries, reinforcing its position as a leader in AI for specialized applications. Across these major cloud providers, several common trends are emerging for 2025 such as

Increased Automation

Automation will be at the forefront, with AI playing a crucial role in managing and optimizing cloud resources. Automated machine learning (AutoML) and AI-driven operational tools will reduce the need for manual intervention, streamlining processes and improving efficiency.

Enhanced Security

AI will be instrumental in enhancing security measures, detecting threats, and mitigating risks in real-time. AI-driven security solutions will become standard offerings, providing robust protection for cloud-based applications and data.

Greater Accessibility and Democratization

Cloud providers will continue to democratize AI, making it more accessible to businesses of all sizes. User-friendly interfaces, pre-built models, and comprehensive documentation will lower the barriers to entry, enabling more organizations to harness the power of AI.

Industry-Specific Solutions

Tailored AI solutions for specific industries will become more prevalent. Cloud providers will develop specialized AI models and applications to address unique challenges in sectors such as healthcare, finance, retail, and manufacturing.

Focus on Ethical AI

Ethical considerations will play a significant role in the development and deployment of AI services. Cloud providers will prioritize transparency, fairness, and accountability, ensuring that AI solutions align with ethical standards and regulatory requirements.

There are three trends in the AI world which will become prominent in the next year - Agentic AI, AI Governance Platforms and Disinformation Security.

1. **Agentic AI**: Agentic AI's ability to act autonomously or semi autonomously has the potential to help CIOs realize their vision for generative AI to increase productivity across the organization.

Some of the use cases are given below:

- Empowering workers to develop and manage more complicated, technical projects — whether microautomations or larger projects — through natural language.
- Automating customer experiences by using data analysis to make highly calculated decisions at each step.
- Changing decision making and improving situational awareness in organizations through quicker data analysis and prediction intelligence.

By 2028, at least 15% of day-to-day work decisions will be made autonomously through agentic AI. The usage of Agentic AI is expected to go up from 0% in 2024 to 15% in 2028 as per the Gartner **Report.**

Unlike Gen-AI, which primarily responds to user prompts, Agentic AI systems are designed to be proactive partners in problem-solving and decision-making, capable of driving significant business outcomes.

Key Features of Agentic AI

Agentic AI is characterized by its ability to act independently and make decisions to achieve specific goals. The design of effective Agentic AI systems, according to Andrew Ng, involves four main components:

1. Reflection: The AI examines its outputs and processes, learning from them to improve future performance.
2. Tool Use: The AI uses external tools, such as web searches or code execution, to enhance its capabilities.
3. Planning: The AI develops and executes multi-step strategies to achieve complex goals.
4. Multi-agent Collaboration: Multiple AI agents collaborate to achieve better outcomes than a single agent could.

Agentic AI vs. Gen-AI Chatbots and Copilots

While Gen-AI chatbots and copilots assist with specific tasks, they are largely reactive, responding to user inputs based on predefined logic. In contrast, Agentic AI systems are proactive, planning and adapting to new challenges. For example, where a chatbot might answer a query, an Agentic AI system would anticipate the next question, plan a response, and collaborate with other tools or AI systems to provide a comprehensive solution.

AI Governance Platforms: AI governance platforms help manage and control AI systems by ensuring they are used responsibly and ethically.

They allow IT leaders to make sure AI is reliable, transparent, fair and accountable while also meeting safety and ethical standards. This ensures that AI aligns with the organization's values and broader societal expectations.

Attributes of Governance are Transparency, Accountability, Fairness, Privacy. AI governance platforms consist of **Ethics, Responsible AI policies and AI Technology.**

AI is being used in more areas, especially in industries with strict regulations. As AI spreads, so do risks like bias, privacy issues and the need to align with human values. It's crucial to ensure AI doesn't harm certain groups, manipulate markets or control important systems. Some of the use cases are given below:

- Assessing potential risks and harms that AI systems may pose, such as bias, privacy violations and negative societal impacts.
- Guiding AI models through the model governance process to ensure all appropriate gates and controls are followed during the model life cycle.
- Tracking usage, monitoring AI system performance, auditing decision-making processes and ensuring AI systems remain aligned with governance standards over time.

Spatial Computing

Spatial computing is revolutionizing the way we interact with digital and physical worlds by seamlessly integrating them. It encompasses technologies such as augmented reality (AR), virtual reality (VR), mixed reality (MR), and other forms of immersive experiences. At the recent CES event, approximately 70% of the coverage was dedicated to advancements in spatial computing, underscoring its growing significance and impact on the consumer market.

Spatial computing has evolved from basic 3D graphics to sophisticated systems that can understand and interpret the physical environment. This progression has been driven by advancements in hardware, software, and connectivity. Devices such as AR glasses, VR headsets, and spatial sensors are becoming more accessible and affordable, paving the way for widespread adoption.

Cloud computing plays a crucial role in the development and deployment of spatial computing applications. By providing scalable and flexible infrastructure, cloud services enable the processing power, storage, and connectivity required for sophisticated spatial computing experiences.

Processing Power and Storage

Spatial computing applications often require significant computational resources to render complex graphics and process large amounts of data. Cloud platforms offer virtually unlimited processing power and storage capacity, allowing developers to create more detailed and immersive experiences without being constrained by local hardware limitations.

Scalability and Flexibility

Cloud computing allows spatial computing applications to scale dynamically based on demand. This means that a high number of users can access AR or VR experiences simultaneously without experiencing performance issues. The cloud's flexibility also enables developers to quickly deploy updates and new features, ensuring that users always have access to the latest advancements.

Connectivity and Collaboration

Cloud services facilitate seamless connectivity and collaboration in spatial computing environments. For example, multiple users can interact in a shared virtual space, regardless of their physical location. This capability is particularly valuable in remote work and education scenarios, where teams can collaborate on projects in real-time using VR or MR tools.

Data Analytics and Machine Learning

Cloud platforms also provide powerful data analytics and machine learning capabilities that enhance spatial computing applications. By analyzing user interactions and environmental data, these technologies can optimize and personalize experiences. For instance, AR navigation apps can learn users' preferences and habits to provide more accurate and relevant directions.

Spatial computing is poised to become a cornerstone of the digital landscape, with cloud computing playing a pivotal role in its growth and accessibility. As seen at the CES event 2024, the consumer market is embracing these technologies, driving innovation and creating new experiences. By leveraging the power of the cloud, spatial computing applications can scale, evolve, and deliver unparalleled value to users across various industries.

Wipro's provided AI/ML solution in cloud adoption for a British multinational oil and gas company. The company had decades of essential information and insights locked into content that was difficult to search and retrieve for the BP Upstream community. Wipro's solution significantly improved the ability to intelligently search for documents by creating an intuitive search capability via a web-based portal.

Solution Details:

Platform: Azure AI/ML

Techniques Used: Azure Cognitive Search service, custom skill for PII detection, security group for access control, and NER for entity extraction.

Benefits to Customer:

- Enhanced Search Capabilities: The solution allowed users to find valuable information quickly, significantly improving the ability to search for documents intelligently.
- Improved Employee Experience: The solution provided a one-stop platform for all employee services, leading to a multifold improvement in employee experience.
- Operational Efficiency: The intuitive search capability and discovery tool enabled users to find essential information quickly, enhancing operational efficiency.

Wipro's Unique Proposition:

- Integration with Multiple Backend Systems: The solution integrated multiple backend systems with the service desk for auto-ticketing, live agent system for human intervention, SSO, and other systems.
- Multi-Language Support: The service desk and chat-based solution supported multiple languages, catering to a diverse workforce.
- AI-Driven Solutions: Leveraging Azure AI and ML algorithms, Wipro provided advanced search and knowledge management capabilities, enhancing the overall user experience.

Cloud Security, Compliance and Resiliency

As cloud adoption continues to grow, so does the focus on security and compliance. Ensuring data protection and regulatory compliance remains a top priority for businesses. Gartner predicts that by 2025, cloud security will account for 20% of total cybersecurity budgets, up from 15% in 2024. The rise in cyber threats and the implementation of stringent data privacy regulations are driving organizations to invest in robust cloud security solutions.

Following are the key Cloud Compliance, Security and Resiliency solutions and frameworks for robust security in 2025.

1. AI-Driven Threat Detection and Response - Leveraging Artificial Intelligence for Enhanced Security

Artificial Intelligence (AI) and Machine Learning (ML) are revolutionizing cloud security by enabling proactive threat detection and automated responses. In 2025, AI-driven threat detection will become a cornerstone of cloud security strategies, allowing organizations to identify and mitigate threats in real-time.

2. Zero Trust Architecture - Adopting a Zero Trust Approach for Comprehensive Security

Zero Trust Architecture (ZTA) operates on the principle of "never trust, always verify," ensuring that every access request is authenticated, authorized, and encrypted. By 2025, zero trust architectures will be widely adopted as a fundamental component of cloud security frameworks.

3. Enhanced Data Encryption Techniques - Advancing Encryption to Protect Sensitive Data

Data breaches remain a significant concern for enterprises, making robust encryption techniques essential. In 2025, advancements in data encryption will provide stronger protection for sensitive information stored and transmitted in the cloud.

Cloud compliance refers to the practice of complying with national and international regulatory standards for cloud usage. The applicability of these standards varies depending on your industry, business location, and the location of your users.

- Ensure compliance with industry-specific regulations like GDPR, HIPAA, PCI-DSS, etc.
- Monitor cloud infrastructure in real time for compliance violations, misconfigurations, and security gaps
- Streamline the audit process and automate the process of report generation
- Enable cloud compliance across multiple cloud environments to have a unified security posture view

Enforce data protection standards to detect vulnerabilities that could expose sensitive data.
Cloud compliance market is estimated at USD 112.10 billion by 2030 compared to USD 39.23 billion in 2023. Cloud Compliance Tools streamline adherence to the compliance standards.

Cloud Service providers offer the native compliance services such as:

- AWS Security Hub
- Azure Security Center
- Google Cloud Security Command Center

There are many Cloud Compliance Tools available that the enterprises can utilize to ensure they are meeting the compliance standards.

Chaos engineering has emerged as a discipline to enable resiliency in the cloud. Chaos engineering is the discipline of experimenting on a distributed system in order to build confidence in its capability to withstand turbulent conditions in production.

At the core of chaos engineering is the idea of conducting experiments to either affirm or disprove hypotheses. Here, a hypothesis refers to an expected or assumed behavior of a system, under specific scenarios. During chaos engineering experiments, hypotheses are tested by injecting turbulence e.g. faults, under real situations, while observing system behavior. The observed behavior is new knowledge, as it affords insights to how the system will fail or withstand (confirm or disprove the defined hypotheses).

When you are talking about SRE with High availability and Scalability, you should understand the application architecture on how it is built and the deployment architecture on how it is setup in order to decide how you can innovate in Availability and scalability solution.
Operations team is moving towards Engineering route with Infra as a code and Site Reliability Engineering (SRE) in which Deployment architecture is suitable for Application, Data platforms, Services (API) and more. Deployment architecture involves Infrastructure provisioning, Packaging, Deploying and Runbook automation as a unified package to build infrastructure, platform and application components.

For example, Canary deployment is used to deploy to a subset of servers first called as 'canary'. In African mines, Canary bird is used to send to the mine first to find if there is any toxic gas before sending actual miners. Canary deployment saves time in deploying to multiple servers without validating and it is suitable for clustered deployment. It is similar to staging except that canary is a production server and used as Live system after complete deployment.

On the other side, developing a blue green deployment architecture is also popular in cloud based application development. Blue Green deployment is the most commonly preferred approach in target platform due to flexibility and benefits that it provides in the deployment architecture. Typically, we will have two set of deployment instances one called as Blue and the other called as Green such as:

- Blue is serving as ACTIVE node and Green is serving as STANDBY. At one time, blue will have one version (n) and Green will have its previous version (n-1). After deploying n-version in Blue, it will be tested and if there is any issue, blue services will be isolated and green services will become active. This makes sure that you will have better turn-around to rollback during failures.
- If blue service installation is stable, we can convincingly use Green services as DR standby by having same version so that if Blue service goes down, Green service can be used as DR recovery site.

In modern cloud lifecycle, SRE and cloud operations aims at operational excellence and moving from timed wait activities in operations, monitoring and management to self-healing services to autonomously manage cloud operations. This will lead CloudOps to move towards intelligent operations through NoOps.

Wipro recently provided cloud security solutions for a pharmaceutical manufacturer. The organization needed to establish its own infrastructure independent of its previous parent company. Wipro was selected to configure a new data center (DC) in AWS within a few weeks, ensuring it was a well-architected, multi-account AWS environment based on security and compliance best practices.

Challenges:

- Establishing a new DC in AWS quickly.
- Adopting an effective Cloud Governance Strategy.
- Handling governance and audit-related tasks.
- Implementing security policies impacting cloud governance adoption.

Solution Highlights:

- Wipro proposed AWS Control Tower to establish a landing zone based on a well-architected, multi-account baseline.
- Implemented Guardrails to enforce governance rules for security, compliance, and operations.
- Deployed Checkpoint CloudGuard NGFW for access control and deep packet inspection.
- Assessed and deployed MS Defender for endpoint security on server workloads in AWS.
- Enhanced governance posture and mitigated challenges in managing a multi-account environment.
- Enabled auto-remediation of cloud misconfigurations with preventive and detective controls.

Benefits to Customer:

- Unified visibility and continuous monitoring of the cloud environment.
- Identification and remediation of misconfigurations.
- Improved governance and security posture.
- Cost savings through efficient cloud management.
- Enhanced data risk assessment and mitigation.

Wipro's Unique Proposition:

- Expertise in designing and implementing cloud governance frameworks.
- Ability to quickly configure and deploy secure, compliant cloud environments.
- Comprehensive security solutions integrating multiple tools and services.
- Strong focus on continuous monitoring and auto-remediation to maintain security and compliance.

Cloud-Native Applications

The shift towards cloud-native applications is transforming software development and deployment. Cloud-native development leverages microservices architecture, containerization, and DevOps practices to build scalable and resilient applications. Everest Group's research indicates that by 2025, cloud-native applications will constitute 80% of all new software development projects, up from 60% in 2024. This trend is fueled by the need for faster time-to-market and enhanced application performance.

The potential trends in Cloud Native applications in 2025:

1. **Increased Adoption of Kubernetes**: Kubernetes has become the de facto standard for container orchestration, and its adoption is expected to continue growing. More organizations will likely leverage Kubernetes for managing their microservices and cloud-native applications.
2. **Serverless Computing**: The trend towards serverless architectures is expected to continue. More companies will adopt serverless platforms like AWS Lambda, Azure Functions, and Google Cloud Functions to achieve better scalability, reduced operational overhead, and lower costs.
3. **Service Meshes**: Tools like Istio, Linkerd, and Consul are becoming more popular for managing microservices communication. Service meshes provide observability, security, and traffic management, which are crucial for complex cloud-native applications.
4. **Edge Computing**: With the growth of IoT and the need for real-time data processing, edge computing will likely become more prevalent. Cloud-native applications will extend to the edge, enabling faster data processing and reduced latency.
5. **Security and Compliance**: As cloud-native applications become more widespread, the focus on security and compliance will intensify. Developers will need to integrate security practices (DevSecOps) into their CI/CD pipelines and ensure compliance with regulations.
6. **AI and Machine Learning Integration**: Cloud-native platforms will increasingly integrate AI and machine learning capabilities. This will enable more intelligent applications that can leverage data for real-time decision-making and automation.
7. **Multi-cloud and Hybrid Cloud Strategies**: Organizations will continue to adopt multi-cloud and hybrid cloud strategies to avoid vendor lock-in, improve resilience, and optimize costs. Tools and platforms that facilitate seamless management across different cloud environments will gain popularity.
8. **Developer Experience**: Improving developer experience will be a key focus area. Tools and platforms that simplify the development, deployment, and management of cloud-native applications will continue to evolve, making it easier for developers to build and maintain applications.
9. **Observability and Monitoring**: As cloud-native applications become more complex, the need for advanced observability and monitoring tools will grow. Solutions that provide deep insights into application performance, distributed tracing, and anomaly detection will be in high demand.
10. **Green Computing**: There will be a growing emphasis on sustainability and energy efficiency in cloud-native applications. Organizations will seek to minimize their carbon footprint by optimizing resource usage and adopting eco-friendly practices.

Cost Optimization Using FinOps

Financial Operations (FinOps) is crucial for managing cloud costs effectively. Cloud native computing allows organizations to implement FinOps principles by providing granular visibility into resource usage and expenditure. Cloud native tools enable real-time monitoring and automation of cost management, ensuring that resources are used efficiently. By adopting a cloud native approach, companies can scale resources dynamically, reducing waste and optimizing spending. The ability to allocate costs to specific departments or projects also enhances accountability and financial transparency.

Composable Architecture

Composable architecture is a design principle that emphasizes the assembly of modular and reusable components. Cloud native computing supports this architecture by enabling the creation, deployment, and management of microservices. Each microservice can be developed, scaled, and maintained independently, providing greater flexibility and agility. This modularity allows organizations to innovate rapidly, adapt to changing requirements, and deploy new features without disrupting existing services. Cloud native platforms facilitate the seamless integration of various services, fostering a truly composable IT environment.

Containerization Beyond Kubernetes

While Kubernetes is a widely adopted container orchestration platform, cloud native computing extends containerization capabilities beyond Kubernetes. Emerging technologies such as serverless computing and container-as-a-service (CaaS) offer additional options for deploying and managing containers. Serverless computing allows developers to focus on writing code without managing the underlying infrastructure, while CaaS provides a managed environment for container deployment. These advancements enhance the flexibility and scalability of containerized applications, enabling organizations to choose the best-fit solutions for their specific needs.

Data Gravity Challenges

Data gravity refers to the tendency of data to attract applications and services, making it difficult to move large datasets across different environments. Cloud native computing addresses data gravity challenges by providing tools and frameworks for efficient data management and migration. Data replication, synchronization, and storage solutions are integral to cloud native platforms, ensuring data accessibility and consistency across distributed environments. By leveraging cloud native practices, organizations can overcome data silos, optimize data workflows, and ensure seamless access to critical information.

Resiliency and Self-Healing Capabilities

Resiliency and self-healing are fundamental aspects of cloud native solutions. Cloud native architectures are designed to withstand failures and recover automatically, ensuring high availability and reliability. Techniques such as automated scaling, load balancing, and redundancy are built into cloud native platforms, minimizing downtime and service disruptions. Self-healing capabilities enable systems to detect and resolve issues proactively, often without human intervention. These features enhance the overall stability and performance of applications, making them robust and resilient in the face of unexpected challenges.

Cloud native computing is pivotal in modern IT infrastructure, offering significant benefits in cost optimization, architectural flexibility, containerization, data management, and resilience. By embracing cloud native practices, organizations can achieve greater efficiency, agility, and reliability in their operations. The integration of FinOps for cost control, the adoption of composable architecture, the exploration of containerization beyond Kubernetes, the mitigation of data gravity challenges, and the implementation of resiliency and self-healing capabilities collectively contribute to a robust and future-proof IT ecosystem. As the cloud native landscape continues to evolve, its importance will only grow, driving innovation and excellence in the digital era.

Wipro had a strategic cloud adoption partnership for a global leader in the Life Sciences industry in the development of an Intelligent Patient Care and Hemodynamic Monitoring System. This solution was built on Azure PaaS services and utilized a cloud-native microservices architecture to provide individualized patient management across diverse patient profiles and care settings.

Benefits to the Customer:

- Enhanced Device Connectivity: The solution integrated on-premise medical devices with cloud PaaS components, improving time to market and quality while reducing dependency on custom components.
- Predictive Decision Support: Advanced monitoring and analytics solutions enabled the development of a predictive decision support system.
- Scalability and Security: The platform was designed for horizontal scaling and met HIPAA compliance, ensuring security at multiple levels.
- Automation: Deployment of all components was automated, enhancing efficiency.

Wipro's Unique Proposition:

- Domain and Technology Expertise: Wipro leveraged its talent in both the domain and necessary technologies, ensuring a quick ramp-up of the expert team.
- IoT Edge Protocol Translation: The solution covered interfaces like FHIR, HL7, and other medical equipment-specific interfaces, ensuring comprehensive connectivity.
- Secure Communication: Implemented X.509 attestation with external CA for secure communication between medical devices and the cloud, managing a high volume of messages per second.
- Automation Framework: Wipro provided an Azure ARM template-based automation framework for provisioning, configuring, and lifecycle management of cloud resources.

Quantum Computing in the Cloud

Quantum computing, once considered a futuristic concept, is gradually becoming a reality with cloud-based quantum computing services. Leading cloud providers are offering quantum computing platforms, enabling researchers and businesses to experiment with quantum algorithms. IDC's forecast suggests that by 2025, the quantum computing market will reach $10 billion, driven by advancements in quantum hardware and software. The potential to solve complex problems faster than classical computers makes quantum computing a game-changer for various industries.

Cloud Service Providers like AWS, GCP, Azure, IBM and Oracle have already enabled Quantum computing services for performing Research on quantum computing algorithms. It helps in accelerating scientific discoveries, running simulations, build quantum software faster and test different quantum hardware.

Quantum simulators are being used across a broad range of scientific research and industrial use cases, including for generation and distribution of quantum-secure random numbers for use in entropy-as-a-service and for quantum key distribution.

More use cases include testing and validation of post-quantum cryptographic algorithms for application and database security for highly regulated enterprise applications, and simulating molecules for drug discovery use cases in the life Sciences and pharma industry.

For the financial industry, this appliance is used for portfolio optimization and faster risk and fraud analysis use cases. Quantum machine learning (QML) algorithms are going to be significant in the development of the next generation of AI systems and large language models using GPU-based QML trainings. Following are the Quantum computing services being offered by various cloud service providers.

1. **Amazon: Amazon Braket**

   Amazon Braket is a fully managed quantum computing service designed to help speed up scientific research and software development for quantum computing. Braket helps in accelerating scientific discoveries, run simulations, build quantum software faster, test different quantum hardware.

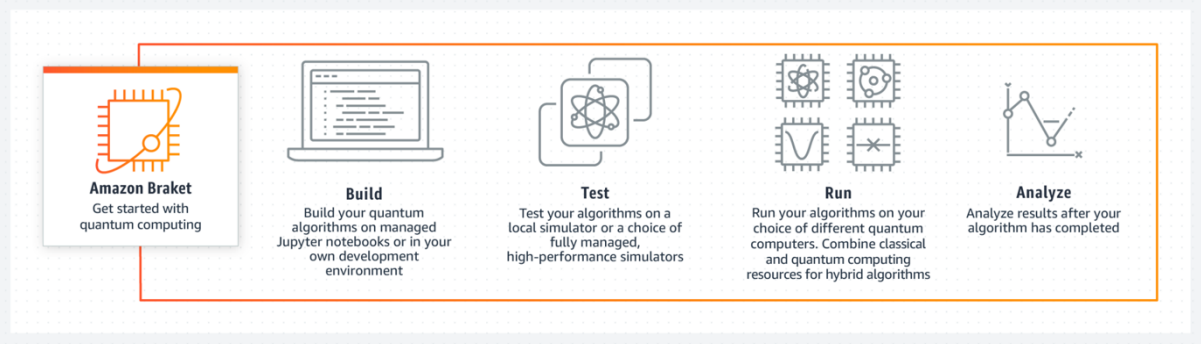


Image Source: https://aws.amazon.com/braket/

2. **Google Cloud Platform: IonQ on GCP**

   Google cloud offers IonQ service in its Market Place. Developers, researchers, and enterprises alike can now access IonQ's high-fidelity, 11-qubit system via Google Cloud in just a few clicks, with billing and provisioning handled via their existing Google Cloud accounts.

   Official Reference:
   https://cloud.google.com/blog/products/compute/ionq-quantum-computer-available-through-google-cloud

3. **Azure: Azure Quantum**

   Provides a diverse portfolio of quantum hardware to build towards the emergence of fault tolerant quantum systems. It needs no installation and provides an open-source SDK that works with modern environments supporting Q#, Qiskit, and Cirq library.

   Azure quantum resource estimator tool to estimate the logical and physical qubits and runtime required to execute quantum applications on future scaled quantum computers.

Official Reference: https://azure.microsoft.com/en-in/products/quantum/

4. **IBM: Qiskit Runtime service**

   Qiskit Runtime is a service and programming model for building, optimizing, and executing quantum workloads. Get pay-as-you-go access to IBM quantum processing units (QPUs) via a low-latency containerized execution environment.

   Official Reference: https://cloud.ibm.com/quantum

5. **Oracle: NVIDIA cuQuantum Appliance on OCI**

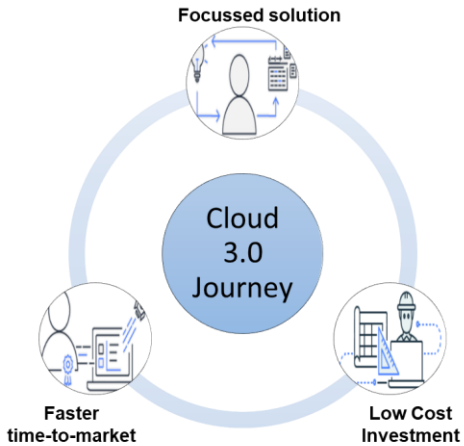
   Oracle Cloud Infrastructure offers NVIDIA cuQuantum Applicance on its Market place. The NVIDIA cuQuantum Appliance is a highly performant multi-GPU solution for quantum circuit simulation. It contains the NVIDIA cuStateVec and cuTensorNet libraries, which optimize state vector and tensor network simulation, respectively. The cuTensorNet library functionality is accessible through Python for tensor network operations.

   Official Reference: https://blogs.oracle.com/cloud-infrastructure/post/oracle-marketplace-nvidia-cuquantum-appliance

Cloud-Based DevSecOps and Automation

The adoption of cloud-based DevOps and automation tools is streamlining software development and IT operations. These tools facilitate continuous integration and continuous delivery (CI/CD), enabling faster and more reliable software releases. Forrester's analysis indicates that by 2025, 70% of enterprises will have fully integrated cloud-based DevOps practices, up from 50% in 2024. The drive towards digital transformation and the need for operational efficiency are key factors behind this trend.

For every Enterprise IT landscape, Operational activity to handle infrastructure, application management, monitoring and service related activities is very vital and backbone of effective Governance. With change in IT operation landscape and lifecycle stages to define better agility, cost effectiveness, smooth transitions, well-defined activities, there are many improvements we have seen in IT Operations activities like ITOps, CloudOps, DevOps, NoOps, AIOps, BizDevOps, DevSecOps, SysOps and DataOps.



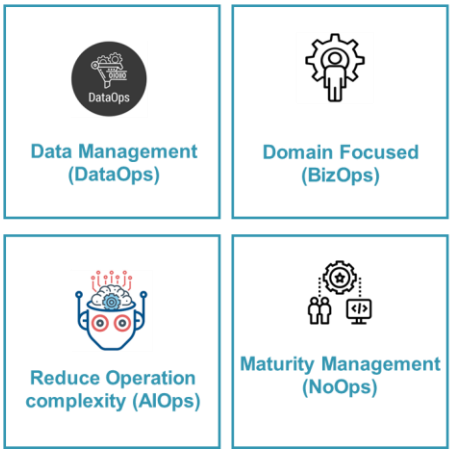


Figure: Operations in Next generation Cloud transformation (Cloud 3.0)

These Ops are not grown to next stage in sequential manner but you can logically arrange them in order as shown in above figure.

ITOps is traditional IT operations for IT management, network, infrastructure, application management and technical support or Help desk operations.

SysOps is more of infrastructure activities (or traditionally Software Configuration Management) to handle build, deployment, migration and managing systems in Cloud platforms and remains early stage of CloudOps.

DevOps is the smooth relation to create an agile team for Development and Operations to have handshake in frequent iterations. Improved version of DevOps 2.0 is also called as **BizDevOps** to enable business team (eg: Business analyst, consultant) to collaborate with DevOps team and more suitable for product development.

SecOps is similar to DevOps but DevOps is between Development team and Operations team whereas SecOps is between Security and Risk management team and Operations team to strengthen various security activities like infrastructure security, network security, data security, application security etc.,

DevSecOps is combining above two of DevOps and SecOps to have smooth interaction between Development team, Security team and Operations team to design, develop, security standard implementation, build, deployment and maintenance activities.

CloudOps is similar to ITOps but used for Cloud platform and have different flavours like Native Operations, third party operations and Agnostic operations. CloudOps differs from SysOps in that CloudOps concentrates more on task automation and cost optimization factors.

AIOps was originally defined by Gartner to handle Big data platforms, machine learning platforms and analytics platforms and also aims at automation activities to reduce IT operation complexity.

DataOps is sliced down version of AIOps to handle Data platforms to include data management, migration, validation, analytics and setup process and standards for data management activities.

Above all, Modern IT operations and Next generation IT operations are aiming at **NoOps** where it leads to automate everything and spend nothing for operations team. For example, Netflix or Facebook aims on development and not setting up an Operations team to support.

IT operations sometimes gets overloaded with volume and complexities in variances of operations and hence simplifying the operational overhead is always an ever demanding activity for an enterprise. In Cloud world, we use the terms DevOps, DevSecOps, SecOps, SysOps, AIOPs, MLOps and NoOps and invariably we have to live with some of them in day-to-day life of cloud transformation journey.

Out of this, it is always confusing to know the difference between AIOps and MLOPs as they are usually used interchangeably without understanding the technical similarities and differences between them. Though MLOPs is closely associated with Azure platform for Machine learning related services and AIOps is technically termed by Gartner to define Automation of IT operations in general and cloud operations in specific, they have some common line and some major differences between them.

MLOPs or Machine Learning Operations is a technical implementation process to create, deploy and maintain production ready machine learning models. They can be Machine Learning workflows or Data science to combine the power of deep learning to train and develop a machine-learning model and use them for real-time machine learning activities like celebrity face detection and voice recognition. MLOps also includes developing a Machine-learning pipeline to build and deploy workflows and model deployment for end user consumption and it standardizes processes in routine activities.

AIOps or Artificial Intelligence based Operations is a functional implementation in IT operations to improvise the processes and efficiency in activities by automating them and commonly used for performance monitoring and anomaly detection in system behaviour and provides business benefits such as increased Return of Investment (ROI) like increase productivity and efficiency and reduce downtime in services (meantime to recovery or MTTR) and improved collaboration through automation services.

AIOPs is typically useful for support operations like automation of incident management, monitoring and management of services and process efficiency. MLOps is useful to develop business models using Machine learning algorithms and build a model for standardizing processes in repeated activities.

MLOps combines Machine learning, Data science, DevOps and data engineering activities and AIOps combines AI based activities and big data to improve operational efficiency.

Gartner defines an adaptable strategy for digital transformation such as cloud adoption journey as Bimodal IT strategy or practice. This could be a game changer model in elevating Cloud operations for enterprises in 2025. Bimodal means two modes of work, which are separate, but they are coherent style of operations and work. This Bimodal IT strategy is two-dimensional where mode-1 is sequential in nature and predictable through best practices and mode-2 is dynamic and agile in nature and aims at delivering results to drive business through exploratory methods.

Figure: Bimodal IT approach

In Cloud adoption, drives Innovation and speed of delivery through technology-centric approach and on the otherside implements business centric approach for agile and reliable solution approaches in IT transformation through digital innovation and cloud adoption.

Gartner also predicts that the top three cloud purchase decisions that drives Digital transformation and cloud journey in next three years are Cost efficiency, Sustainability and Operational Excellence. To enable a strategic approach to handle these services are:

- Innovation led approach but more focus on cost benefits and cloud service expense management through FinOps driven strategy.
- Waste tolerance in IT resource through GreenOps, which focuses on carbon emission control and management for any hyperscalers and cloud services.
- Improve operational efficiency in regular IT operations including DevSecOps, CloudOps and AIOps to lead to operational nirvana state through NoOps.

The key benefits in adopting IT strategy leading to Cloud journey and Digital transformation are Faster time to market, flexibility in adoption strategy, cultural transformation to simplify operational activities and above all innovation led approach to tailor best practices and at the same time introduce new ways of working in the entire lifecycle of transformation.

When looking into the trends of CloudOps in 2025, one should look at the new age terms like CapOps and RevOps. As per Densify's definition, CapOps is coined as a branch of operational activities which bridges the gap between DevOps and FinOps. DevOps handles operational delivery and automation in continuous delivery activities. FinOps handles cost management and cloud service optimization through pricing efficiency and predictability analytics. But there is a gap between them where it primarily address how cloud compute and storage services can be optimized and well-handled to save cost and reduce wastage.

This kind of CapOps helps to handle right sizing, auto-scaling of fan-in and fan-out of container services including container's node optimization and storage optimization to utilize right capacity (eg: database storage size management). In Traditional IT operations, we used to have Capacity Management as a branch of activity who works closely with Infrastructure architect to handle capacity of IT resources. With Cloud platform, there is additional flavour added to resource management through dynamic capacity management where one has to work on continuous improvement in capacity planning and dynamically use the optimization options like spot instances, committed user discounts and savings plan. Hence there would be a growing demand for CapOps in order to get Operational Efficiency and Resource optimization in cloud resources.

RevOps (Revenue Operations) is an extension of FinOps and CapOps which is the goal driven from CIO team to focus on Revenue cycle management in cloud operations and cost management to bridge the gap between increasing CAPEX and lowering the recurring OPEX in cloud platforms. Cloud operations focussing on Cost transparency (CapOps and Cloud governance), cost management (Centralized vs Federated CloudOps) and cost optimization (FinOps) are the driving factor for RevOps approach and this would have a greater focus in 2025 for large enterprise to contain their ongoing costs in cloud journey.

A Wipro's customer case study in Cloud DevSecOps and automation solution involves a leading automotive manufacturer seeking a digital transformation partner to enable end-to-end agile processes, automation capabilities, and operational efficiency.

Challenge:
The client faced challenges with siloed security processes and tools spread across on-premise environments. They aimed to set up an application security program before embarking on cloud transformation, practicing Agile/DevSecOps delivery.

Wipro proposed a solution to address these challenges by:

- Embracing development teams onto a secure agile journey.
- Deploying a building-breaker strategy to make development teams security advocates.
- Enabling continuous planning and security test execution powered by NextGen SAST and ASOC platforms.
- Providing proactive visibility, prioritization, and feedback on threats and vulnerabilities within each phase of the SDLC.

Benefits:

- Increased visibility on the IT landscape, improving observability regarding risk vs. vulnerability scoring.
- Enhanced operational efficiency, delivering 100 Pentesting engagements annually and testing 120 SAST applications quarterly.
- Enabled readiness for DevSecOps scale by introducing a secure-by-design culture.
- Enhanced ability to capture security assurance metrics, boosting customer confidence and trust among regulators.

Wipro's Unique Proposition:

- The Wipro ASOC platform, which supports proactive visibility and feedback on threats and vulnerabilities.
- A hybrid approach to verifying control effectiveness across various applications, including web, mobile, IoT, cloud, and COTS products.
- The deployment of a continuous planning and security test execution model, reducing the mean time for vulnerability remediation.

Green Cloud Computing (Sustainability or Green IT)

Sustainability is becoming a critical consideration in cloud computing. Green cloud computing focuses on minimizing the environmental impact of cloud operations through energy-efficient data centers and renewable energy sources. Gartner's report highlights that by 2025, 60% of cloud providers will have committed to achieving carbon neutrality, compared to 40% in 2024. The emphasis on sustainability is driven by corporate social responsibility initiatives and the growing awareness of climate change.

A key observation is that the granularity of data being collected is not available, and as the Cloud platform becomes the default choice, there will be pressing need to provide granular information in terms of various cloud constructs and a better accounting system for enterprises to be able to better understand their

carbon emission consumptions. All the major players have already started indicating the regions that operate *Greener* than other regions. This is currently linked to the electricity generation that is predominantly the reason for emissions, however as we learn more about the emissions in general, there is going to be a better understanding of the emissions across the value chain, which in cloud computing will include various aspects, building and destroying servers, building data centre facilities, upgrading hardware components, emissions linked to storage of data, and many other aspects that are currently being researched.

Google has a target to run all its data centers with clean and green energy by 2030, Amazon has set a target for the same at 2040, and Azure has a target to be carbon negative by 2030. The 3 providers currently utilise renewable energy purchases, or credits, to compensate for emissions they make. This basically offsets the emissions in the accounting systems, where in by investing more into sustainable power generation projects and initiatives, a company can achieve carbon neutral position. One of the side-effects this brings in is an imbalance of emissions across the globe which eventually negatively impacts bio-diversity and human living. It would be a key focus of the industry in general to focus on keeping this balance, by achieving green energy based "cloud regions" across the globe.

With GenAI becoming the new normal for all enterprises, the required data storage and computation capacity is going to ever increasing in nature in next few years. It is natural to focus on Sustainability of these GenAI initiatives as they move from sideline to main stream customer IT landscapes. The AI and GenAI based on Cloud Computing resources will not only help Enterprises drive their Sustainability initiatives, but also there is expected to be more focus on running the AI and GenAI processes themselves in Greener manner.

One of the key trend we could observe is that GreenOps and FinOps are going to be more closely integrated. An enterprise can benefit from cost optimisations' benefits as well as meeting their Sustainability goals at the same time when these initiatives are driven in an integrated manner. The key objective of GreenOps and FinOps is to run leaner, as the underlying energy consumption is the driving factor for costs as well as carbon emissions. An integrated strategy incorporating Sustainability initiatives when organisation tries to maximize business value at minimum cloud consumption costs will be a driving factor for a more sustainable enterprise.

Sovereign Cloud

The popular term "sovereign cloud" describes a cloud architecture to provide security and data privacy adhering to local laws and regulations such as GDPR, CCPA etc and multiple countries have enacted and planning data protection and sovereignty laws.

Gartner defines Sovereign Cloud as "*Sovereign cloud is the provision of cloud services within a jurisdiction meeting data residency requirements and operational autonomy. It is intended to ensure that data, operations infrastructure and technology are free from control by external jurisdictions and protected from foreign government influence and access*".

According to the 2022 Gartner Digital business buying behaviour survey, nearly one-forth of IT services buyers consider digital and data sovereignty as a key external trends shaping their organization's business

strategy. Organizations are becoming deeply concerned about their dependence on foreign cloud providers and as alternate instead evaluating sovereign clouds.

As per one of the recent Gartner report 97.2 of the cloud infrastructure and platform services consumed worldwide are provided by the US. Growing geo-political regulatory fragmentation, protectionism and industry compliance are driving the demand of regional and sovereign clouds. Public sectors, energy and utilities and financials organizations are keen to limit the critical lock in and single point of failure with their cloud providers outside their base country.

There is a wide range of criteria for a sovereign cloud and nations and governments have rigorous data protection rules in place. Data that is sensitive or private is protected by sovereign cloud laws and regulations to make sure that it is always local in the region and secure. The data set that stored in cloud platforms affects the data protection requirements. For example- restrictions governing the use of government agencies and financial and medical information are more stringent than those governing statistical analysis of user and traffic data.

Cloud sovereignty demands validation on two fronts: the enterprise monitoring its cloud and data storage services and the capability to demonstrate compliance with local data privacy and security laws and regulations. As per Gartner, by 2025, 30% of the multinational organizations will experience revenue loss, brand damage or legal action due to unmanaged digital sovereign risk; and by 2028, over 50% of the multinational enterprises will have digital sovereignty strategies, up from less than 10% today.

Regular evaluations of cloud records that track data transfer and access rights over a predetermined period are used to support sovereignty claims. If a cloud fails its sovereignty tests, its owner may be required to pay a fine and occasionally make up for any harm caused to users by malicious data. Laws and regulations that keep data safe are more critical than ever as data transforms into a valuable asset rather than a trail consumers leave online.

For all businesses, for-profit and nonprofit alike, that gather user data and information, data sovereignty and the cloud are crucial. There are multiple ways how CSPs can help clients achieving their cloud sovereignty requirements: by capitalize on emerging opportunity and assist clients by creating a framework for addressing sovereignty requirements and comprehensive set of complementary services that supplement leading global cloud IaaS, PaaS and Saas providers cloud offerings.

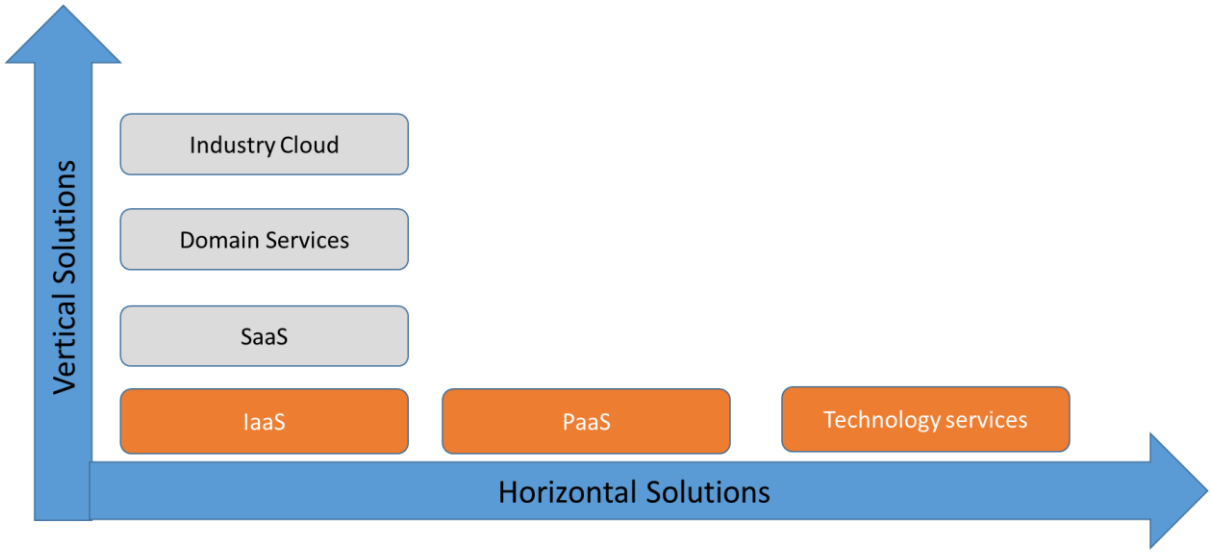
Industrial Cloud

Industrial cloud is getting popular in recent times and Analyst like Forrester predicts that Industry cloud accelerate their momentum in 2022 and Industry solutions increase business agility and resiliency. But, the question is why Industry cloud is important and why we are inclined towards vertical solutions with industry cloud?

Traditionally, cloud solutions are horizontal in nature to move from infrastructure (IaaS), runtime platform (PaaS) and application stack (technology services like AIML, Data lake). But, when there is industry specific architecture pattern possible which could be common for many industries, then we should think on vertical cloud solutions.

Initially, vertical solution started with Software with a specific business functionality (eg: SaaS) to provide stipulated business activities. For example, core banking solutions, KYC services, Insurance claim

verification to name a few. An extension to this is domain services in cloud platforms for specific business capabilities coupled with pre-loaded reference data. For example, FOREX conversion or retail supplychain services. They help to build quick domain enabled applications for faster time to market.



Now, what is trending now is Industrial cloud where we can see a lot of new services introduced by popular hyperscalers. For example, Microsoft Cloud for Retail, Amazon FinSpace, Google Healthcare API are recent additions to Industrial cloud services. They have some standard features to provide domain centric services like

- They comply with pillars of Well-architected framework like scalability, security, availability, performance and operational resilience.
- They are designed to support compliance to regulatory requirements depending on supported regions.
- They are flexible in architecture to integrate or extend to other native services.
- They have security best practices included in its design.
- They provide some reference architectures for standard implementation through Infra-as-Code (IaC) templates.

Industrial cloud and Industry specific 'as a service' are trending in cloud platforms and popular Cloud Service providers (CSP) like Azure, Google and AWS are leading this space by introducing various Industry cloud and domain specific cloud services.

Compared to other sectors, Banking/Financial Services and Healthcare is leading this area like cloud API and domain services are introduced to attract customers for quick integration with their domain applications. For example, Google Cloud's Healthcare solutions are very attractive to have domain compliance (HIPAA, HITRUST) API services for data processing, data analytics, FHIR and DICOM data handling (import and export) to dataflow, Healthcare domain datasets available for data analytics and building trained model for AI enabled services (EHR, EMR data processing).

Microsoft and other Cloud Service provider like Amazon and Google, targets industrial cloud solutions for smaller and mid-size enterprise to quickly adapt cloud solutions for their enterprise solutions. Earlier in the year 2000, Microsoft announced Microsoft Cloud for Healthcare which aims for one-stop healthcare solution in Cloud platform (both SaaS and PaaS solutions) combining multiple solution services like Telemedicine, AI based diagnostics to name a few. On the other hand, AWS has Transcribe Medical

service to handle Medical services related functions. In the same way, Google has a feature called Healthcare API by providing DICOM, FHIR, and HL7v2 APIs for handling healthcare functions and data store. It is HIPAA compliance and caters to privacy and security requirements for data security and transaction security.

As per HealthTech Magazine, more than 90% health data is digitized and processed through cloud platforms in US which accelerates exchange of Electronic Health Record (EHR) between health organizations and improve patient monitoring activities.
Healthcare solutions in Cloud platforms requires lot of foundational building blocks like HIPAA compliant, data security, AI based solutions for training and executing models, inter-operability of services and improved customer experience. With Healthcare cloud and Healthcare Bot, now Microsoft has acquired Nuance to strengthen its cloud solutions in Healthcare industry.

In Microsoft Build 2020, Microsoft unleashed its 'game changer' Cloud offering called Microsoft Cloud for Healthcare as first of its kind as its Industry specific cloud offering which can be used for integrated solution for Hospitals and Healthcare units to provide a unified system by utilizing the flexibility and workload management of Azure Cloud infrastructure for

- Single pane of support system for enhanced patient data management
- Collaboration system for Health care groups and other systems for insurance, Government support and Payment gateway.
- Data insights and Analytics in operational and clinical data.
- Security and inter-operability of Cloud architecture.
- Flexibility to extend the ecosystem for healthcare partners.
- Integrated Bot Services for Contact center automation and customer support
- HIPAA certified services and HITRUST certified solution.

Conclusion

Cloud technology is not new but it is also stagnated when we see the growth and abundance of improvements from various hyperscalers. With Adaptive cloud from Microsoft, Quantum and Sagemaker with Q services from Amazon, Big data services from Google, Generative AI services from Oracle Cloud Infrastructure (OCI) we are seeing cloud is in its peak state to grow and take customers to next level of digital adoption. Technologists rightly predict that combinational technology like Cloud with AI, Cloud with Data, Cloud with Quantum, Cloud with GenAI and Cloud with Blockchain will help implementing new business cases more efficiently. The growth of cloud technology equally helps the growth of IT transformation and business agility across various industries.

References

1. https://emt.gartnerweb.com/ngw/globalassets/en/information-technology/documents/trends/2025-top-tech-trends-ebook.pdf
2. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9133399
3. https://community.nasscom.in/index.php/communities/cloud-computing/future-trends-cloud-computing-what-expect-2025

4. https://www.cigniti.com/blog/future-cloud-infrastructure-trends-2025/
5. https://www.gartner.com/en/articles/top-technology-trends-2025
6. https://www.researchnester.com/reports/public-cloud-in-bfsi-market/6482
7. https://www.clouddefense.ai/future-of-cloud-computing/
8. https://www.cloudzero.com/blog/cloud-computing-statistics/
9. https://www.forrester.com/blogs/category/cloud-computing/
10. https://www.gartner.com/en/documents/5556895
11. https://www.gartner.com/en/newsroom/press-releases/2023-11-29-gartner-says-cloud-will-become-a-business-necessity-by-2028
12. https://www.grandviewresearch.com/industry-analysis/cloud-computing-industry
13. https://www.everestgrp.com/webinars/mid-market-digital-transformation-insights-and-outlook-for-2025-webinar.html
14. https://es.nttdata.com/documents/banking/everest-group-maximizing-value-from-cloud-investments-in-banking-and-financial-services.pdf
15. https://www.linkedin.com/posts/everest-group_amazon-web-services-aws-services-state-activity-7131302243747655682-ZCXo
16. https://www.seagate.com/files/www-content/our-story/trends/files/dataage-idc-report-final.pdf
17. https://www.idc.com/getdoc.jsp?containerId=prUS52460024
18. https://www.businesswire.com/news/home/20210914005759/en/IDC-Forecasts-Worldwide-Whole-Cloud-Spending-to-Reach-1.3-Trillion-by-2025
19. https://www.gartner.com/en/documents/5210663

About Wipro Limited

Wipro Limited (NYSE: WIT, BSE: 507685, NSE: WIPRO) is a leading technology services and consulting company focused on building innovative solutions that address clients' most complex digital transformation needs. Leveraging our holistic portfolio of capabilities in consulting, design, engineering, and operations, we help clients realize their boldest ambitions and build future-ready, sustainable businesses. With over 230,000 employees and business partners across 65 countries, we deliver on the promise of helping our customers, colleagues, and communities thrive in an ever-changing world. For additional information, visit us at www.wipro.com.