



GENERATIVE AI FOR HYBRID CLOUD

Generative AI solutions to solve current hybrid cloud management challenges

Generative artificial intelligence (GenAI) has demonstrated its ability to create new data, content, and code, which has propelled it to the forefront of global news headlines and modern technological advancement. As a key component of the larger AI ecosystem, GenAI plays a unique role, filling gaps and enhancing capabilities of other AI technologies. GenAI interfaces with predictive models, reinforcement learning algorithms, and other AI elements to form a cohesive, integrated system that drives innovation and efficiency.

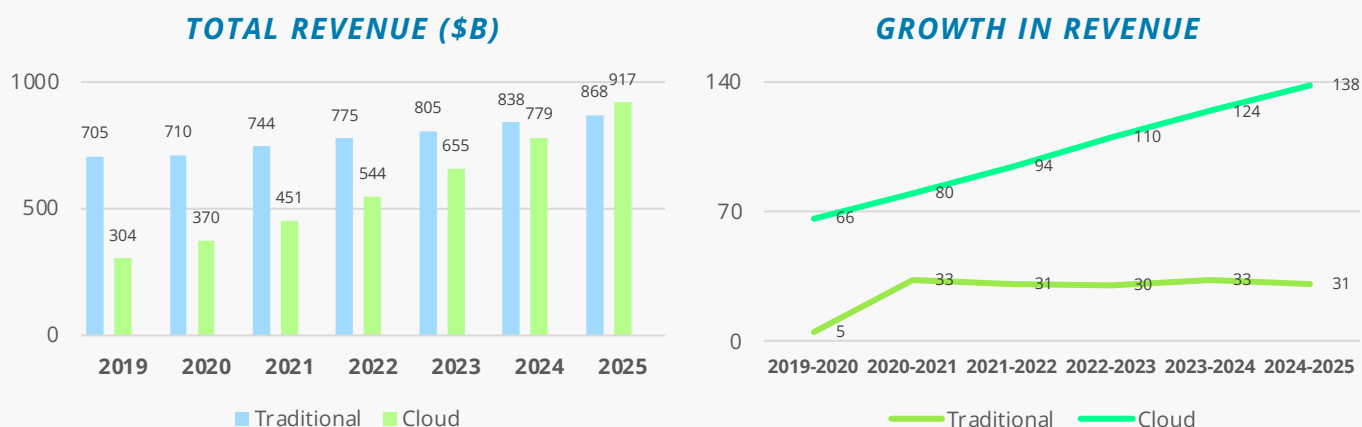
Organizations, both commercial and federal, are operating in a paradigm filled with new capabilities and opportunities, and IT departments in particular face the task of having to do more with less. This includes tasks such as optimizing existing processes and decreasing total cost of ownership for IT solutions. For these organizations, harnessing the power of GenAI presents a unique opportunity to streamline processes, optimize resource allocation, and unlock new possibilities for IT without needing to significantly change their existing workforce.

Challenges with current state hybrid cloud management

The modern IT landscape is characterized by constant change and the need for rapid adaptation. This necessitates that organizations integrate infrastructure that is not only robust and secure, but also adaptable and cost-effective. Hybrid cloud environments have continued to grow in popularity over the past several years, as organizations have recognized the value of utilizing cloud storage technology while also understanding that legacy IT infrastructure offers the advantage of pure ownership.

Figure 1 exemplifies this rise of cloud adoption in comparison to legacy IT infrastructure's persistent revenue usage and revenue generation: the effective use of working capital and deploy an organization's workforce more efficiently.

ENTERPRISE SPENDING ON TRADITIONAL STORAGE VS. CLOUD STORAGE



Source: Gartner
758067_C

Figure 1. Enterprise spending on traditional (on-premises) storage vs. cloud storage. <https://www.gartner.com/en/newsroom/press-releases/2022-02-09-gartner-says-more-than-half-of-enterprise-it-spending>

Hybrid cloud environments combine the strengths of on-premise and private/public cloud technologies to realize the best of both solutions. However, this added flexibility comes with additional complexity, leaving many hybrid cloud solutions with additional overhead and consumption for duplicating services across environments.

The adoption of hybrid cloud environments has been rapid and commonly performed through a phased approach where legacy systems are augmented with cloud functionalities or services. However, many decisions are required as organizations are tasked with choosing which cloud provider to use, determining whether they need private or public cloud solutions, making budget considerations, and avoiding vendor "lock-in." Additionally, when combining legacy systems, private clouds and one or multiple public cloud offerings, many IT departments find that their users and applications are now spread across several storage methods without the means to centrally view, secure, or manage these environments effectively. At the same time, IT leaders are expected to comply with a growing and evolving set of mandates, including cloud policies that address significant concerns, such as regulatory compliance, risk, and security.

Deloitte's Multi-Cloud Command Center, or "DMC3" can resolve these concerns by providing a unified service catalog of multiple cloud providers, such as Amazon Web Services (AWS), that can be utilized and managed through a single multi-tenant portal. DMC3 comes with pre-built APIs to all major cloud providers, on-premises virtual environments, and bare metal solutions to enable maximum flexibility. Organizations can utilize DMC3's powerful blueprint systems to achieve automation and standardization of policy, procedure, governance, and cybersecurity workflows and controls for hybrid- and multi-cloud environments. The entire platform is installed with real-time security and threat monitoring that is unified through each component of the client's hybrid environment. For several years, this platform has saved large government organizations millions in cost savings and significant time reductions, and Deloitte continues to invest in additional capabilities to further expand its value.

While current hybrid cloud management tools such as DMC3 offer robust functionality, they can traditionally require maintenance and code considerations. This reflects the complexity of managing disparate environments, and skilled IT professionals are needed to ensure their smooth operation.



GenAI for hybrid cloud

Hybrid cloud management tools have become a necessity for IT leaders to regain proper management and governance over an abundance of public and private cloud providers under one organizational hood. As an example of a centralized cloud management solution, DMC3 allows for the application of standardized policy, procedure, governance, and cybersecurity workflows and controls for disparate system and application stakeholders supporting the mission. The tool also enables an organization to collect, analyze, consolidate, and report performance and utilization metrics that allow for easier service usage and cost transparency of all cloud services to which it is connected.

As GenAI has developed abilities related to tasks highlighted in DMC3, organizations like Deloitte are assessing their current suite of tools and assets to consider where this exciting new capability can further accelerate time to deliver cloud management solutions to clients. Some areas already being explored include:

Enhancing development and operational efficiency with GenAI

GenAI tools can automate tasks like code generation, function suggestion, and code review in building cloud applications. This streamlines development processes, reduces development time, and frees up developers to focus on more complex tasks.

Cost-savings via increased efficiency:

- *Cost management:* By leveraging predictive analytics and machine learning algorithms, GenAI can detect cost optimization opportunities by identifying the most cost-effective deployment options, thereby balancing on-premises and cloud resources.
- *Reduce manual work:* GenAI automates repetitive and time-consuming tasks associated with hybrid cloud management. This includes provisioning resources, configuring deployments, and generating reports.
- *Automated scaling:* GenAI can automate the scaling of resources based on workload demand, ensuring that applications and services can effectively handle fluctuations in traffic and demand across hybrid cloud environments.

Adaptive security response: GenAI can analyze security threats and automatically generate tailored security configurations or mitigation strategies that adapt to the specific cloud environment where the threat is detected. This helps create a more dynamic and responsive security posture across multiple cloud platforms.

- *Security and compliance:* GenAI can assist in ensuring security and compliance standards are maintained across hybrid cloud deployments by analyzing data patterns, identifying security threats, and recommending appropriate security measures.
- *Disaster recovery and high availability:* GenAI can help in designing and implementing effective disaster recovery and high availability strategies by analyzing data replication, failover mechanisms, and recovery time objectives across hybrid cloud infrastructure.

Personalized cloud service recommendations: GenAI can analyze an organization's specific needs and workloads, then recommend the most suitable cloud services from different providers within a multi-cloud setup. This goes beyond basic optimization and leverages generative capabilities to suggest tailored combinations of cloud resources.

- *Resource optimization:* GenAI technology can analyze workloads across on-premises and cloud environments to optimize resource allocation, facilitating efficient utilization of computing resources.
- *Performance monitoring and management:* GenAI can continuously monitor the performance of applications and services across hybrid cloud environments, identifying bottlenecks and optimizing performance through dynamic resource allocation.



Generative cloud migration strategies: GenAI can analyze an organization's existing IT infrastructure and propose migration strategies for transitioning workloads to a multi-cloud environment. This includes generating code for workload migration and configuration across different cloud platforms.

- *Workload placement and migration:* GenAI can analyze workload characteristics and requirements to recommend placement of workloads between on-premises and cloud environments, as well as assist in effective workload migration.

Synthetic data generation: GenAI can create realistic synthetic data sets that mimic real-world data and simulate thousands or millions of real-world scenarios to train AI models without privacy concerns. This allows for the training and testing of AI in a multi-cloud setting without needing access to sensitive data across different cloud providers.

GenAI offers a multitude of benefits for organizations managing hybrid cloud environments. Overall, the range of benefits that Gen AI technology offers provides a massive opportunity for developers to accelerate tools and products offered by their organization.



Extending Deloitte Multi-Cloud Command Center (DMC3) to GenAI

Deloitte identified two key areas to accelerate hybrid cloud tools, such as DMC3, through the integration of GenAI into day-to-day development tasks:

Automated infrastructure standardization and deployment:

The current method of constructing infrastructure as code (IaC) through DMC3 includes automating workloads and repository development. Using GenAI, the DMC3 team can accelerate automation through generating configuration “snapshots.” This makes creating IaC for new cloud environments even faster, as shown in Figure 2:

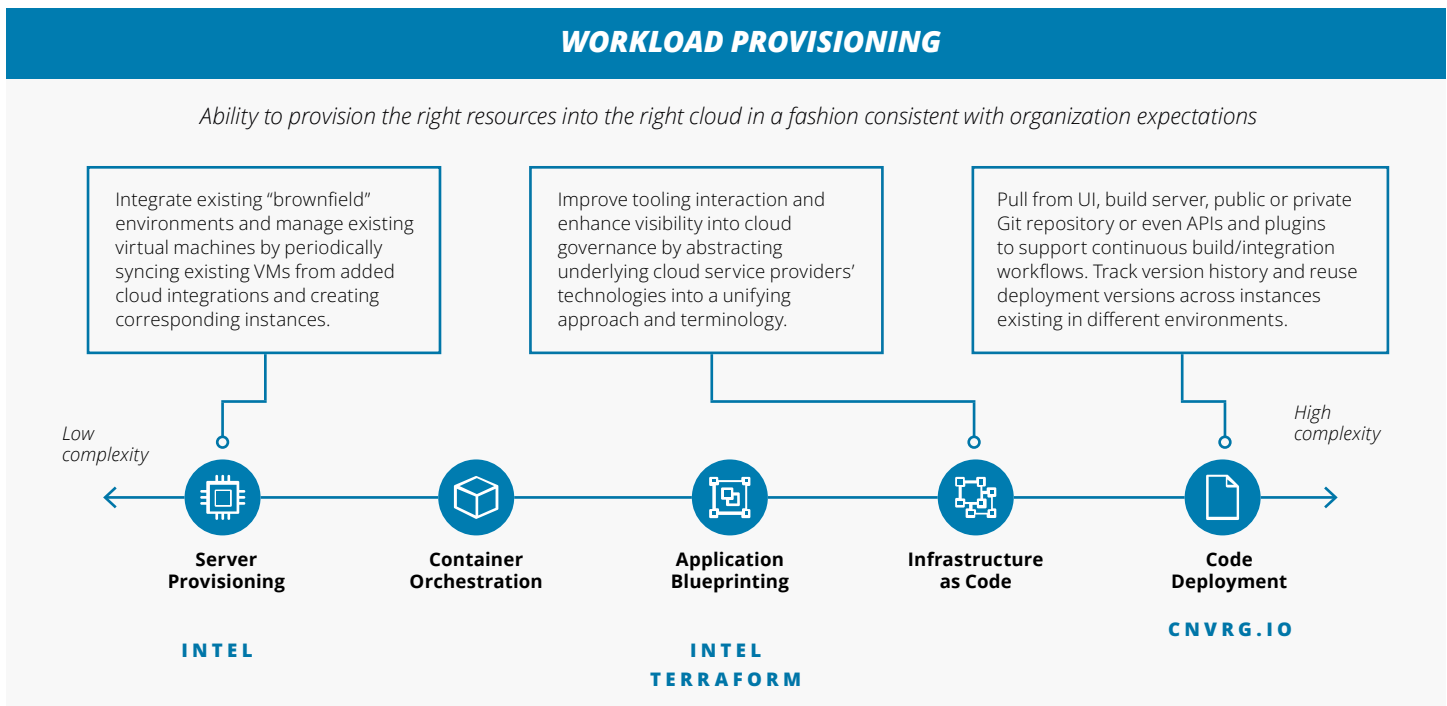


Figure 2: Workload provisioning innovation flow chart

With growing demand for environments capable of handling GenAI workloads, it is important that these blueprints and workflows are optimized for AI methods such as training and inference. Providing exact requirements and documentation can give additional complexity to help AI generate IaC in a manner that is compliant with the current system or specific requirements. By utilizing this IaC “snapshot” generation capability, it would be possible to create a “build one, deploy repeatedly” AI-optimized cloud ecosystem. Through this and related automation, efficiencies can be introduced into the DevOps application release lifecycle, such as those depicted in the workload provisioning flow of Figure 2 and beyond.

Code conversion: DMC3 uses languages like Python, C, and JSON to deploy workloads in various cloud environments (such as AWS). GenAI enables developers to rapidly translate existing code between languages to deploy the same workload in multiple environments. This drastically reduces the time it takes to deploy code scripts for large, complex organizational IT structures.

The workflow in Figure 3 illustrates DMC3's ability to drive efficiency of IT operations across multiple cloud platforms within a range of complexity. By enabling an AI-driven cloud environment, DMC3 is able to conduct even more efficient operations by leveraging AI to automate complex processes and optimize resource allocation, effectively reducing the reliance on siloed tools and processes.

This not only streamlines workflows but also enhances the system's ability to handle diverse AI workloads effectively across different environments, thereby boosting overall operational efficiency.

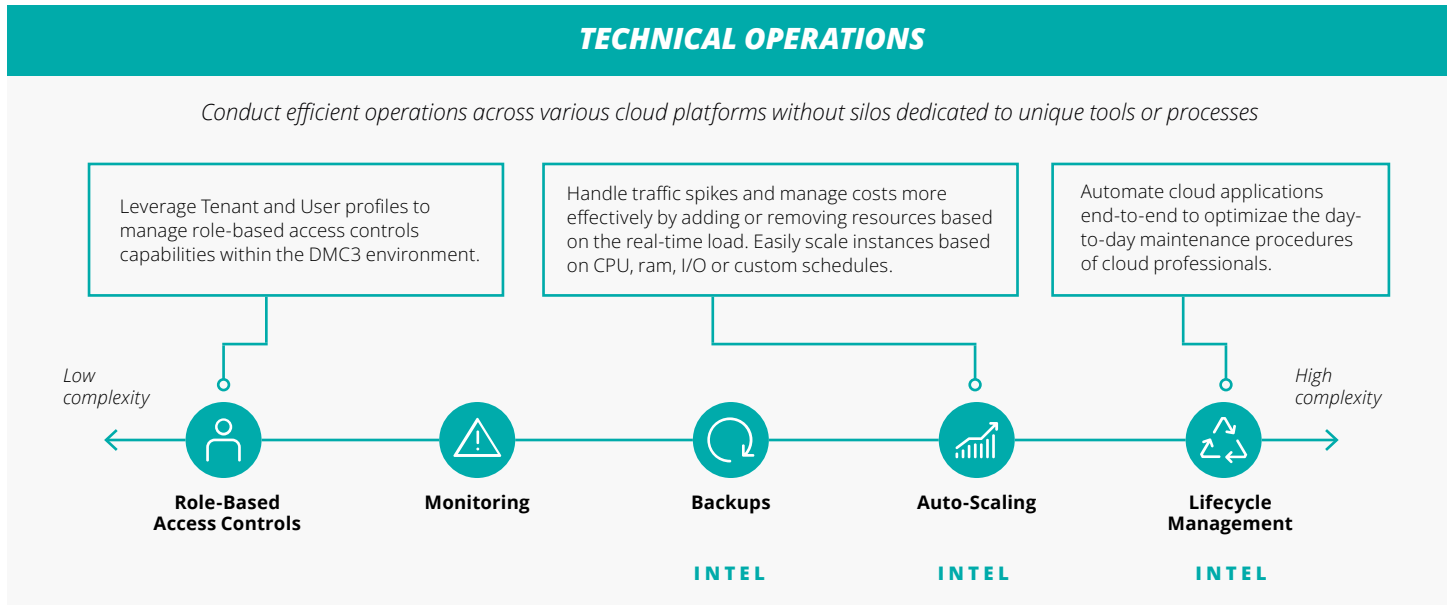


Figure 3: Technical operation innovation flow chart

On the front of organizational management, an AI-optimized environment can provide advanced analytics and predictive capabilities to better understand and manage cloud financials. AI algorithms can analyze spending patterns and predict future costs with greater accuracy, enabling proactive budgeting and cost management.

This empowers business units with precise, real-time financial insights, fostering accountability and preventing cost overruns. The organizational management workflow for DMC3 in Figure 4 is greatly enhanced by AI's ability to autonomously make cost-savings decisions and provide organizational insights without needing to be asked.

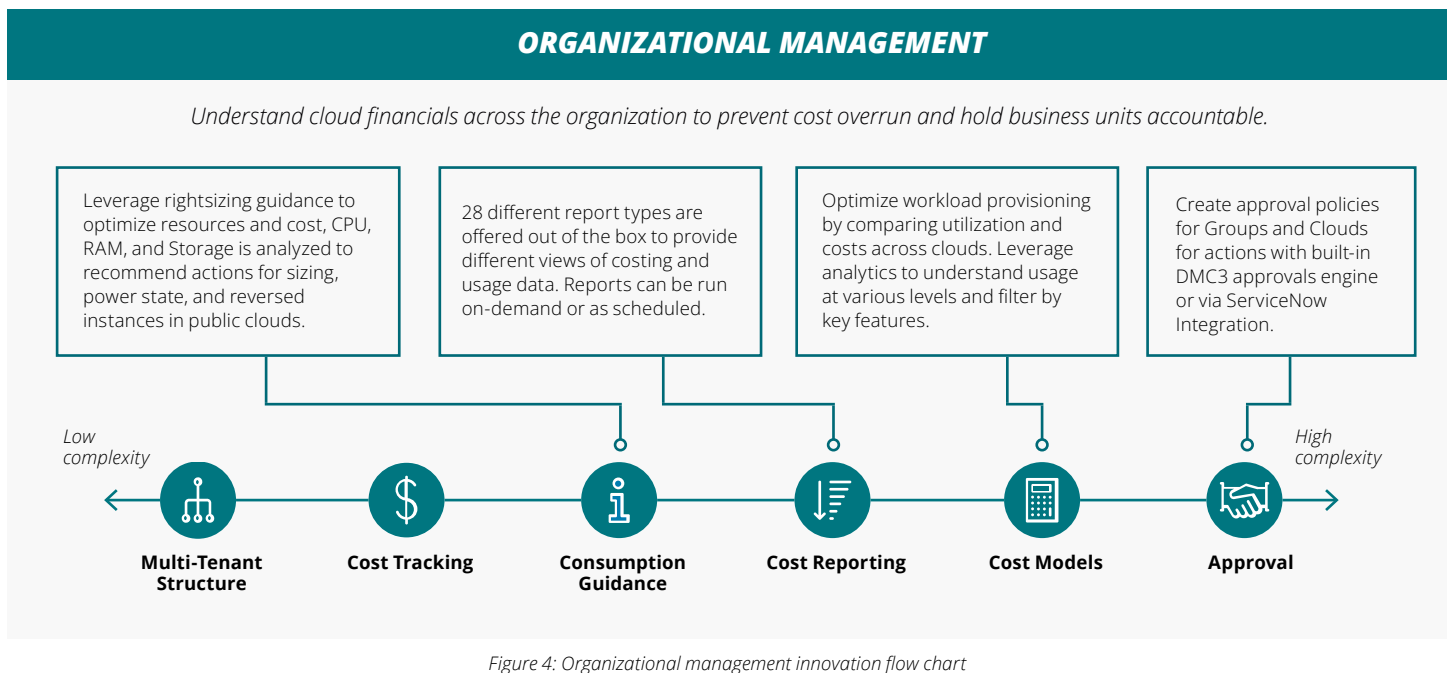


Figure 4: Organizational management innovation flow chart

Applying DMC3 with GenAI to real-world clients

GenAI clearly shows potential to enhance the effectiveness of complex systems like DMC3. Its application is predicted to yield significant benefits to organizations, fostering an environment that enables growth and operational efficiency.

One such example is the GenAI benefits of DMC3 realized for a civil government organization that has relied on DMC3 for their IT infrastructure over several years with Deloitte. With the expected growth of organizations like these, there will be an additional increase in data and operational requirements, leading to increased need for new cloud deployments and additional storage provisions. In these scenarios, the implementation of GenAI can prove beneficial. The technology can enable DMC3 developers to work at a faster pace, efficiently managing the agency's multi-cloud environment and decreasing the time needed to manage extensive data and operational expansions.

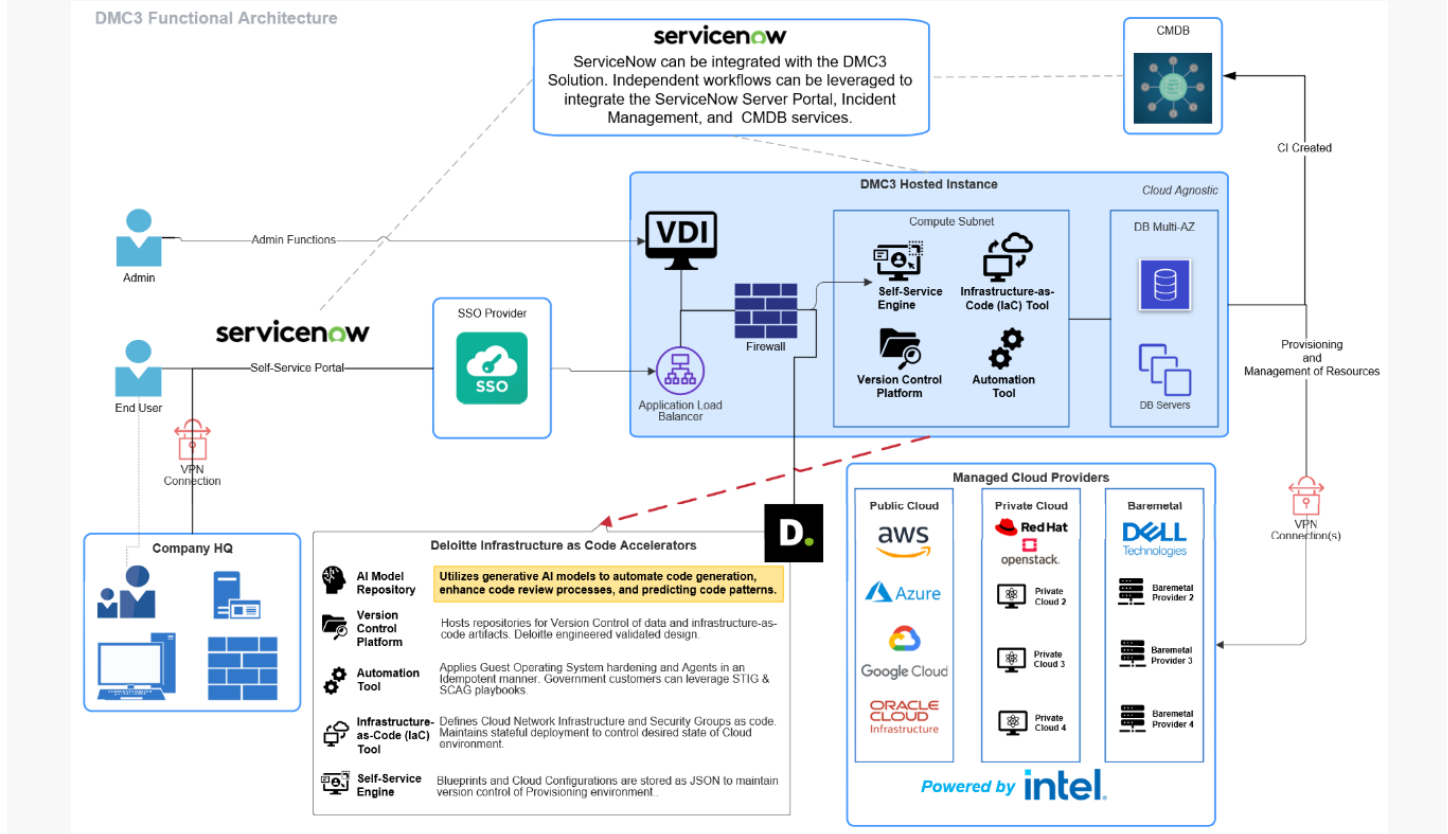
But the impact of GenAI extends beyond enhancing speed and efficiency. Its potential to automate lengthy code translation and deployment processes can also save substantial time for organizations. By expediting processes that previously took days, weeks, or even months, such as cloud migrations or new environment setups, the technology can result in significant cost savings.

Take for instance a specific civil agency with a large amount of IT data and the capacity to spend nearly \$6 million hosting legacy infrastructure within a data center; the integration of GenAI within the DMC3 asset presents a substantial value proposition, potentially saving millions of dollars for such clients through faster cloud provisioning while enhancing operational efficiency and scalability.

How GenAI fits into a hybrid cloud architecture

The integration of GenAI within hybrid cloud architectures can yield significant benefits. By leveraging GenAI models, like those obtained from online AI communities such as Hugging Face, DMC3's suite of code accelerators can be significantly enhanced. Furthermore, hybrid cloud architectures allow for the centralization of data generated by different cloud providers, leading to an insight-rich environment for GenAI-optimized workloads. This consolidated environment empowers GenAI users to perform in-depth analyses and draw insightful conclusions.

EXHIBIT 1 | GENAI FOR HYBRID CLOUD REFERENCE ARCHITECTURE



While several integration points for Gen AI within this architecture have been identified, these are merely starting points. As GenAI technology continues to evolve, more opportunities for integration will be discovered, driving further advancements in the field.

Conclusion

A paradigm shift is occurring through the advent of GenAI. The increased ability for automation allows hybrid cloud developers to focus their work on more complex and technical architecture challenges, while GenAI capabilities focus on repetitive tasks, enabling both faster and more effective solution delivery. As GenAI continues to evolve, the industry could not only benefit from faster automation, but also from uncovering machine-driven insights. The evolution of DM3 through GenAI will continue to position the product for increased market relevance as well as increasing Deloitte's serviceable market at large.

Both GenAI and hybrid cloud management are ever-changing fields of technology that change at a rapid pace. The confluence of two or more impactful technologies such as these require a diverse set of skills and perspectives to collaborate and find ways to build whole capabilities for the future. System integrators like Deloitte, alongside global technology leaders such as AWS and Intel Corp, continue to join forces to generate elevated value and spark progress for clients by leveraging the cutting-edge technologies of today to build solutions for tomorrow.

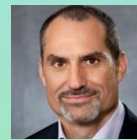
As used in this document, "Deloitte" means Deloitte Consulting LLP, a subsidiary of Deloitte LLP. Please see www.deloitte.com/us/about for a detailed description of our legal structure. Certain services may not be available to attest clients under the rules and regulations of public accounting.

This publication contains general information only and Deloitte is not, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor. Deloitte shall not be responsible for any loss sustained by any person who relies on this publication.

Copyright © 2024 Deloitte Development LLC. All rights reserved.

MEET THE DELOITTE THOUGHT LEADERS

Key points of contact for any questions regarding the content of this paper.



Doug Bourgeois

Managing Director

Deloitte Consulting LLP

GPS CBO Cloud Engineering



Vish Chanda

Senior Manager

Deloitte Consulting LLP

GPS CBO Cloud Engineering



Ryan Luckay

Specialist Leader

Deloitte Consulting LLP

GPS Analytics and Cognitive

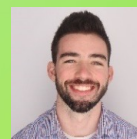


Vijay Mosur

Senior Manager

Deloitte Consulting LLP

GPS CBO Cloud Engineering



Matt Sheerin

Consultant

Deloitte Consulting LLP

GPS CBO Core Technology

Operations