Mainframe workload Modernization Patterns for Financial Services

It’s our observation that financial services organizations have been pioneers in adopting technology to help reduce operating costs, bring innovations faster and provide better customer experience. For many, IBM Mainframes have been part of their core IT platform for more than 5 decades due to powerful processing power, reliability and security. Even today, 85 of the top 100 banks use mainframes to run their mission critical IT workloads.

The new competition from FinTechs and BigTechs, evolving business models and regulatory and compliance risks are driving changes in financial services industry’s technology needs. They are looking for a modern platform which can provide agility, security and scalability at a lower cost. They are investing in cloud computing and have started moving their non-critical workloads to cloud. Some of the organizations are already underway moving critical user facing applications and data to cloud to provide faster time to market, better user experience and improved scalability without comprising on security. However, some may still be reluctant to migrate mission-critical mainframe workloads to hybrid cloud due to risks of migration failure, very long migration timelines and little or no know-how of the old processes which may make the migration difficult.

Some of the challenges that organizations may face with the mainframe platform motivating them to look for alternatives are:

- **Shortage of skilled resources** –The existing mainframe skilled staff is retiring and is not being replaced by new talent due to lack of interest by younger generation engineers to learn mainframe skills.
- **Outdated tools and development environments** – Many mainframe developers still use Interactive System Productivity Facility (ISPF) green screens to write, debug and test code, elongating the overall development and product release cycles, which could be a barrier for onboarding new staff who are used to work with modern development environments.
- **Manual Dev and Operations processes** – Low penetration of DevOps practices for mainframe application development, causing two-speed IT where distributed development platform is more agile than the mainframe
- **Limited access to application and data causing interoperability issues** – The business logic hidden in legacy programs is not easily accessible to distributed applications, making it difficult for mainframe to operate in modern IT ecosystem
- **Inability to unlock the value of massive business data stored on mainframes** – The mainframe data stored on disparate database management systems is not accessible in real-time for analytics and machine learning processes
- **High cost of operation** – The technical debt accumulated over decades and unoptimized processes are contributing to higher costs of operations on mainframe platform

Several of these challenges are perceived rather than real. Many of the financial services organizations have not refreshed their mainframe tooling and technology stack over a decade and continue to run it as a legacy platform. With the release of z13 in 2015, the mainframe platform was redesigned to bring many innovations and modern tooling. IBM has been continuously investing in modernizing this platform with the release of z14 in 2017, z15 in 2019, z16 in 2022 and z17 in 2025. With each version, the modern mainframe hardware and software was designed to provide enhanced security, speed, resiliency and modern tools to interoperate in a hybrid cloud ecosystem.

The hybrid cloud infrastructure is a combination of one or more public cloud, private cloud and on-premises infrastructure. The hybrid cloud environment enables adoption of technologies which provide consistent development experience with modern developer tools, automation and DevSecOps practices. We find that most organizations are embracing hybrid cloud strategy as their preferred cloud deployment model. The mainframe platform will continue to be an important part of this strategy and organizations need to start thinking long term on what should move to cloud and what can stay and be modernized on mainframe.

Capgemini and IBM help financial services customers navigate through their mainframe modernization journey. Capgemini provides a balanced approach towards mainframe modernization by choosing the right workloads to migrate to public/private cloud along with identifying in-place modernization patterns to modernize mainframe workloads.

The Capgemini in-place mainframe modernization solution focuses on modernizing developer experience, bringing agility and speed of delivery, providing real-time access to application business logic and data, and bringing Artificial Intelligence and Machine Learning and data analytics closer to source data. Capgemini has created a mainframe modernization Experience Zone where customers can experience these modernization use cases in action and co-develop their use cases with Capgemini mainframe experts before implementing it in their environment.

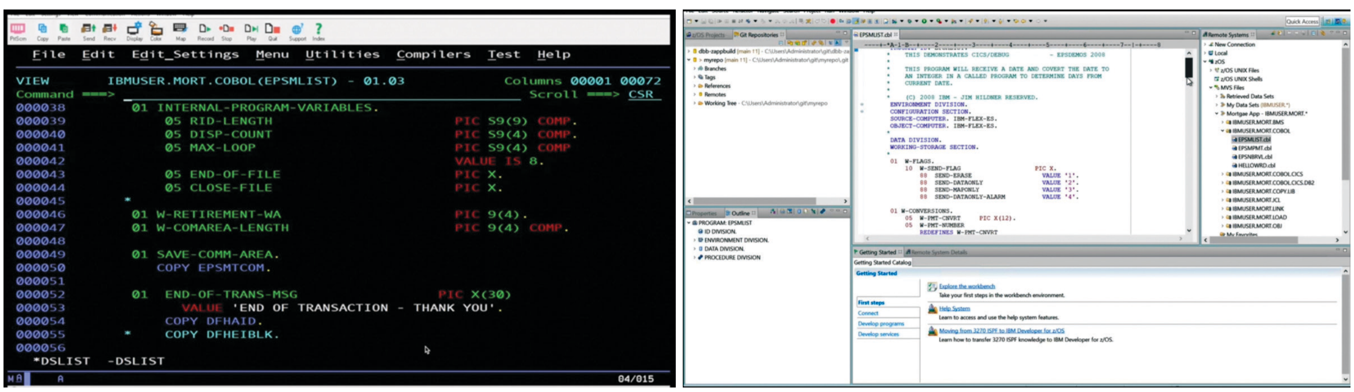
Below is a brief summary of each of these use cases.

Modernizing developer experience

Mainframe developers have not seen significant changes to their development environments in the last few decades and they are largely based on 3270 emulator screens using the ISPF editor. Customers can embrace the modern integrated development environments (IDEs) of their choice based on industry preferred platforms such as Eclipse or VSCode to integrate with the mainframe. These are well known platforms used for developing applications written in newer generation languages and now the same platform can be used to develop applications on mainframe in COBOL, PL1, JCL and many other supported languages. These are feature-rich IDEs and provide powerful tools to code, debug and test applications which are not available on traditional 3270 based mainframe development environments.

Implementing modern IDEs brings the experience of developing mainframe applications closer to distributed application development. It is easier to find new talent who have worked on these environments and can easily learn COBOL to support mainframe applications.

Figure 1. Side-by-side view of two IDEs: ISPF editor v/s Eclipse based IBM Developer for z/OS (IDz)



The success of this use case relies essentially on adoption of these tools by mainframe development staff. Many of the mainframe developers are using ISPF editors for decades and onboarding them to new IDEs will require training and coaching followed by regular feedback sessions. Once they see the power of these editors and efficiency gain they bring in their day-to-day work, they will adopt quickly.

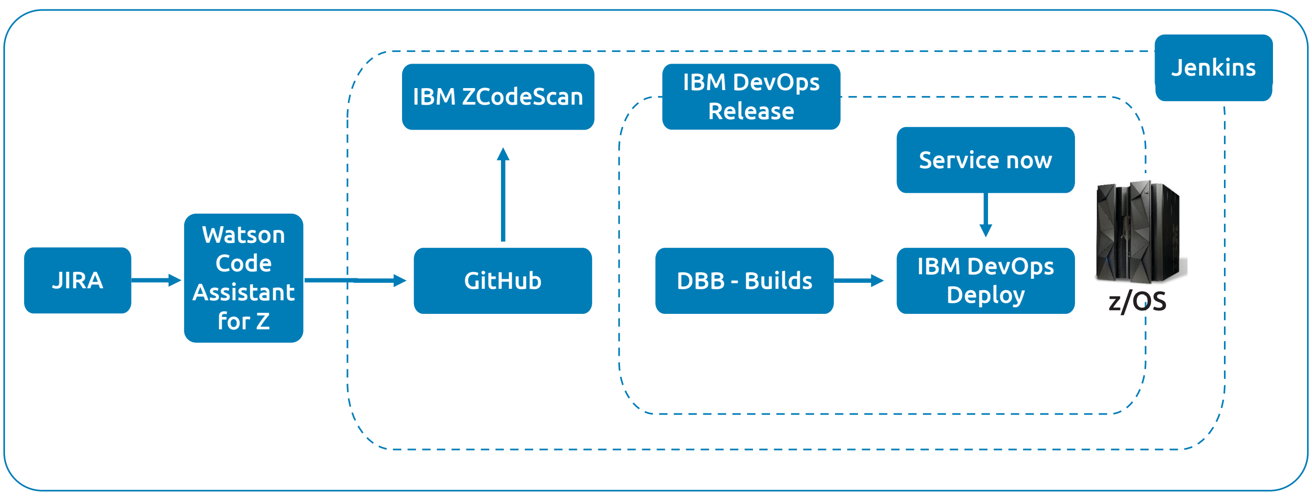
Agility and speed of delivery using DevOps tools and practices

Organizations are adopting agile software development life cycle to deliver more functionalities at a faster rate. DevOps has been at the heart of this strategy and it has helped to speed-up the delivery cycles by shifting the role of delivery and operations from multiple teams to software engineers, providing them tools to automate the entire lifecycle using continuous integration and continuous delivery/ deployment, CI/CD toolchain.

Many organizations today either already have matured DevOps practice or in-process of implementing one, but the majority of these implementation only focus on distributed and cloud native technologies leaving the mainframe as an island to operate in the old world. This creates two-speed IT situation where one team is delivering faster than other.

Many of the open-source CI/CD toolchains can now be configured to work with the modern mainframe such as GitHub and Jenkins. Organizations can leverage these tools to create common DevOps framework to orchestrate the development, integration and deployment of an application to multiple target platforms including cloud and mainframe.

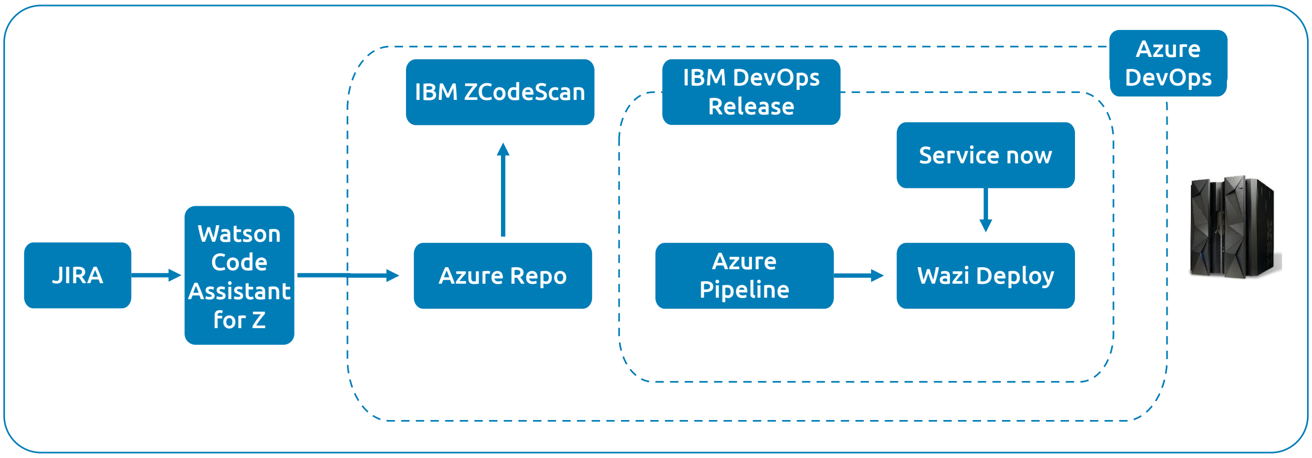
Figure 2. Sample reference architecture to implement DevOps pipeline for mainframe technology stack



Activities	Requirement Management	Development Modernization	Source Control Management	Code Quality	Automated Unit Test	Build	Deploy	Release Management	Continuous Monitoring & Reporting	CICD	ITSM Integration
Tools Used	JIRA	Watson Cod Assistant for Z	GitHub	IBM ZCodeScan	TA4Z	Dependency Based Build (DBB)	(IBM DevOps Deploy)	IBM DevOps Release	Omegamon / Instana	Jenkins	ServiceNow

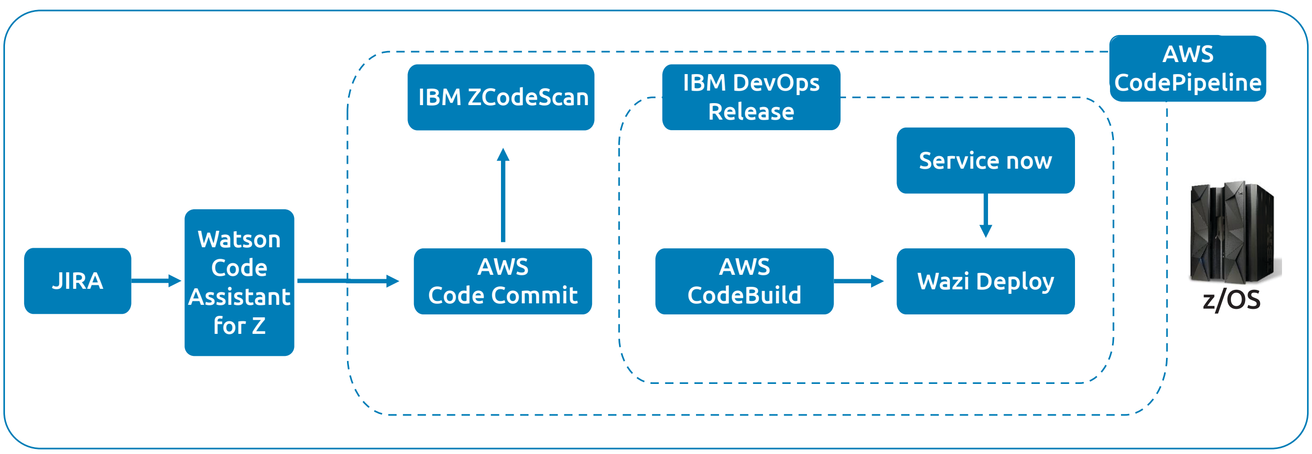
For customers with hybrid cloud ecosystem, the DevOps frameworks can also be implemented on public platform, enabling integrated pipeline and single-speed IT across cloud and Z applications. Below diagrams show reference architecture of implementing DevOps for Z applications on Azure and AWS:

Figure 3. CI/CD Pipeline on Azure for Z applications



Activities	Requirement Management	Development Modernization	Source Control Management	Code Quality	Automated Unit Test	Build	Deploy	Release Management	Continuous Monitoring & Reporting	CICD	ITSM Integration
Tools Used	JIRA	Watson Code Assistant for Z	Azure Repo	IBM ZCodeScan	TA4Z (Test Accelerator for Z)	Azure Pipeline	Wazi Deploy	IBM DevOps Release	Omegamon / Instana	Azure DevOps	ServiceNow

Figure 4. CI/CD Pipeline on AWS for Z Applications



Activities	Requirement Management	Development Modernization	Source Control Management	Code Quality	Automated Unit Test	Build	Deploy	Release Management	Continuous Monitoring & Reporting	CICD	ITSM Integration
Tools Used	JIRA	Watson Code Assistant for Z	AWS CodeCom mit	IBM ZCodeScan	TA4Z (Test Accelerator for Z)	AWS CodeBuild	Wazi Deploy	IBM DevOps Release	Omegamon / Instana	AWS CodePipelin	ServiceNow

Like the developer experience use case, the adoption is a key to success of any DevOps implementation. Start with the proof of concept to ensure that your setup works as expected, and then onboard the applications to DevOps pipeline based on the maturity of an application and development team.

DevOps is not only about tooling but also about ways of working. DevOps promotes a culture of continuous learning, innovation, and trust, requiring large enterprises to adopt new practices and tools. IBM Z can be part of this transformative strategy to meet evolving market demands.

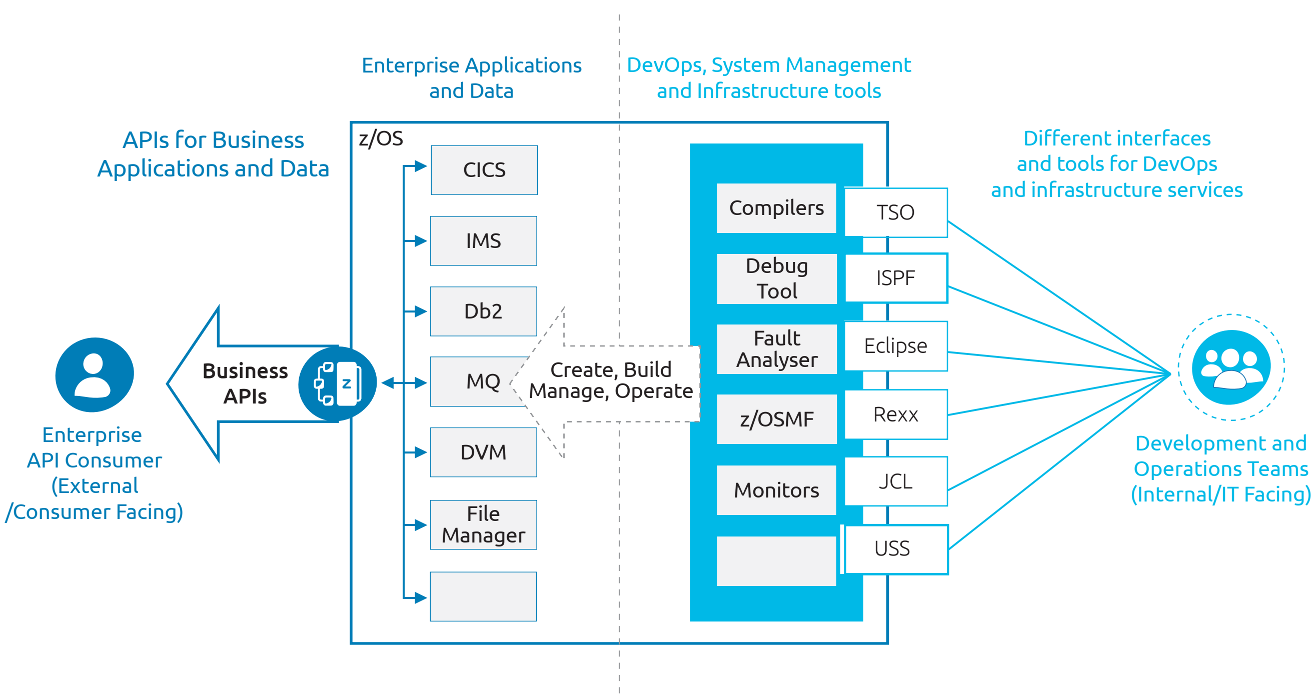
Real-time access to mainframe applications and data

Digital transformation requires access to core applications and data to operate on omnichannel platform. Majority of financial services organizations have their core business logic on the mainframe written in old programs with little or no documentation. Re-engineering and rewriting these programs on a different platform is effort-intensive and error prone.

This modernization use case transforms mainframe COBOL programs to RESTful services which can be accessed by internal or external applications using open API standards. This not only ensures the seamless exchange of information across hybrid cloud ecosystem but also creates ways to monetize the organization's valuable intellectual property buried in these legacy programs.

IBM's z/OS Connect accelerates the creation of these APIs by providing intuitive and user-friendly interface "Z/OS Connect Designer , an IDE for developers to define the REST services and APIs in minutes. These APIs can be registered and discovered from an enterprise API Gateway or management platform by any of the cloud applications running on public or private cloud.

Figure 5. z/OS connect EE for transforming mainframe assets to business APIs

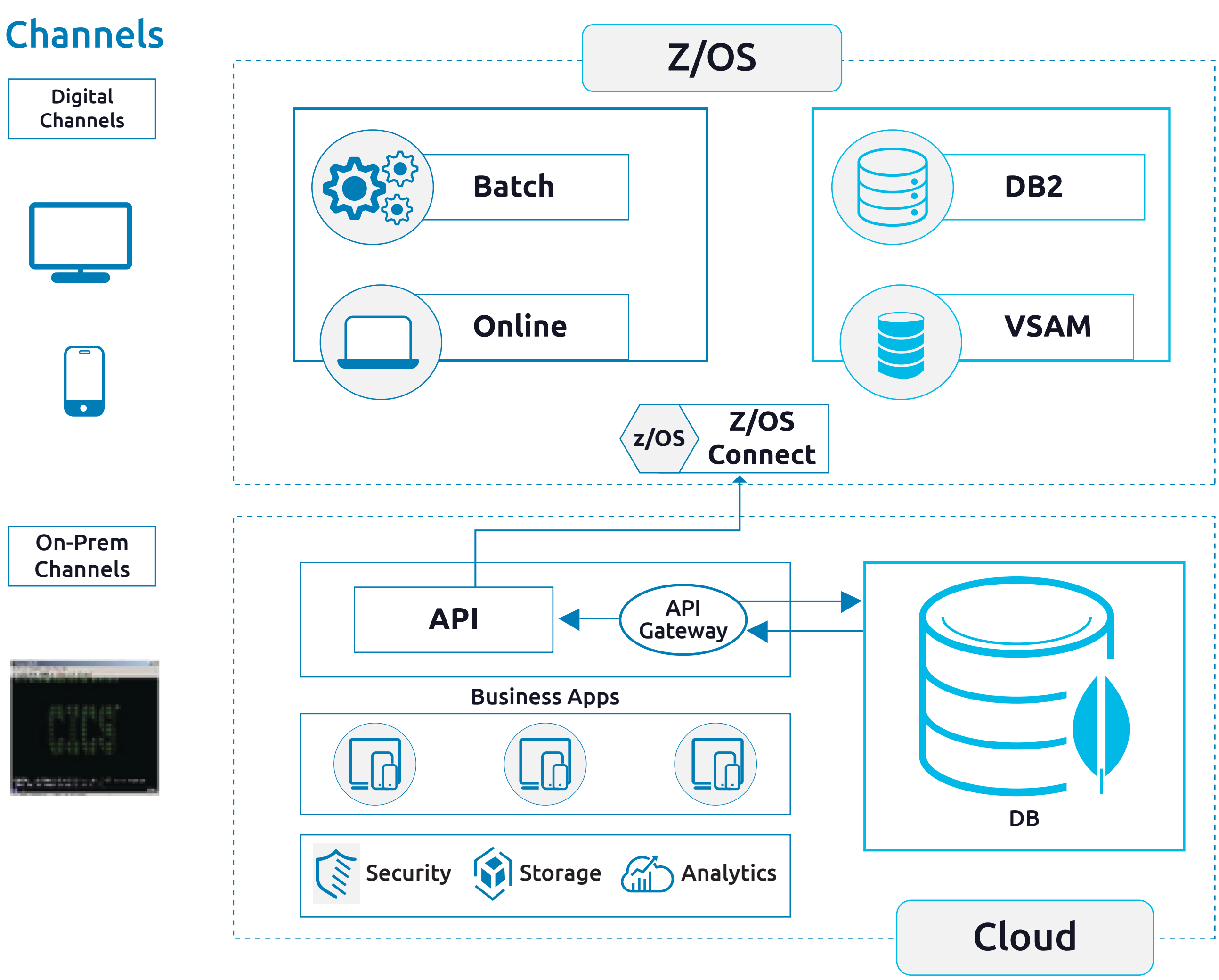


Real-time access to business logic or data stored on Cloud for Z applications

Mainframe applications can access the business logic or data stored on cloud with the help of Open APIs built and deployed with z/OS connect interface. This opens-up many possibilities and use cases while modernizing the mainframe apps:

- Access to archival data stored on Cloud
- Flexibility to choose the right platform for modernizing the apps and providing seamless integration between different platforms (cloud, on-premise distributed and mainframe) .
- Integrating legacy mainframe apps with insights generated by new cloud-native apps

Figure 6. APIs with z/OS connect to access data stored on MongoDB on Cloud by Z applications



Real-time access to event driven data

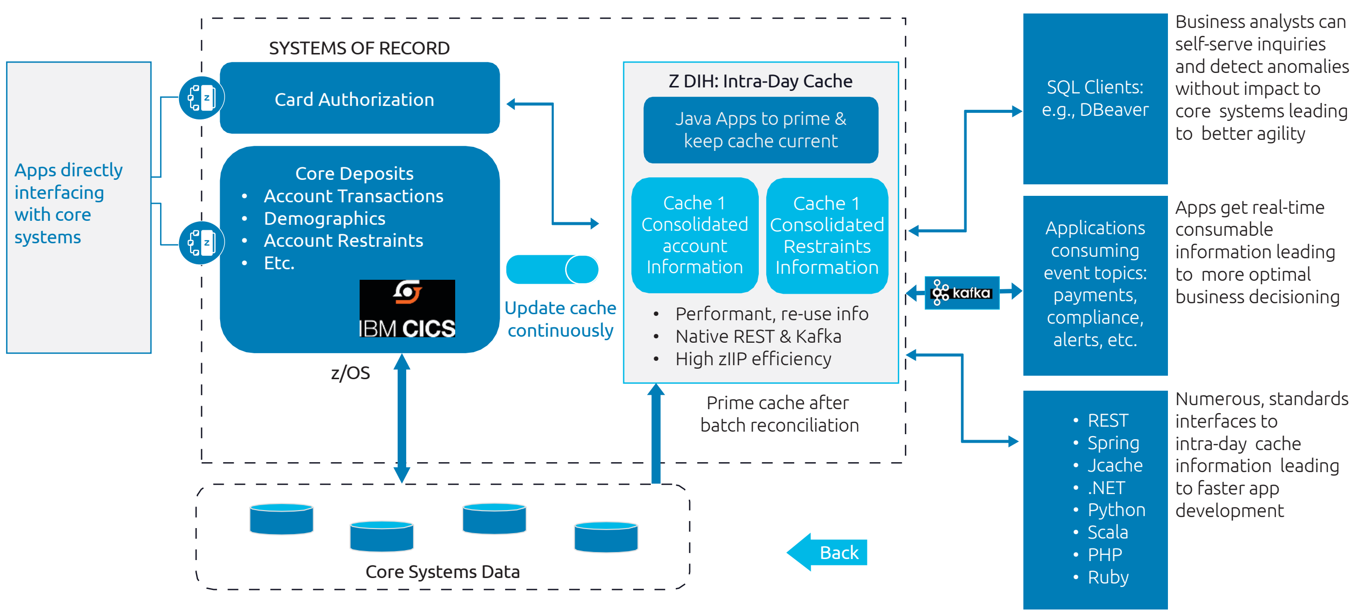
Mainframe usage is specifically high for data-centric industries who want fast and secure processing of data. The vast majority of credit card transactions are still processed on mainframes and many of the core banking transactions still run on mainframes. Majority of organizations are using mainframe for system of records and generating huge amount of transaction data daily. This data is locked-in on mainframes and organizations lack holistic strategy to make this data available for cloud and data analytics applications. Many of the organizations have implemented enterprise Datalike or Datawarehouse where the mainframe data is curated and loaded to run analytics and business reporting, but this data is mostly a day old and does not support real-time data driven decisioning.

Organizations can solve this problem by creating a Digital Integration Hub exposing core mainframe data elements in real-time to all consumers. Digital Integration Hub is an advanced application architecture that aggregates multiple back-end system of record data sources into a low-latency and scale-out, high-performance data store (Gartner definition).

IBM's z Digital Integration Hub (zDIH) provides flow of information between system of records on mainframe and cloud environment with high through-put and low latency. This is not raw data but curated information which can be accessed in real-time by cloud applications. The information can be curated from various mainframe data sources such as DB2, VSAM and IMS or directly from business applications while processing data in real-time. The data is stored in an in-memory cache which can be easily accessed using either RESTful API, Kafka topics or SQL clients. The goal of the DIH is to:

- Prevent the core applications from being overwhelmed with excessive API requests by routing majority of read-only calls to DIH cache
- Provide ability for cloud applications to consume real-time information by subscribing to Kafka topics
- Provide real-time data access through SQL client to business analysts, developers and production support team

Figure 7. z Digital Integration Hub to provide access to event driven information on mainframe

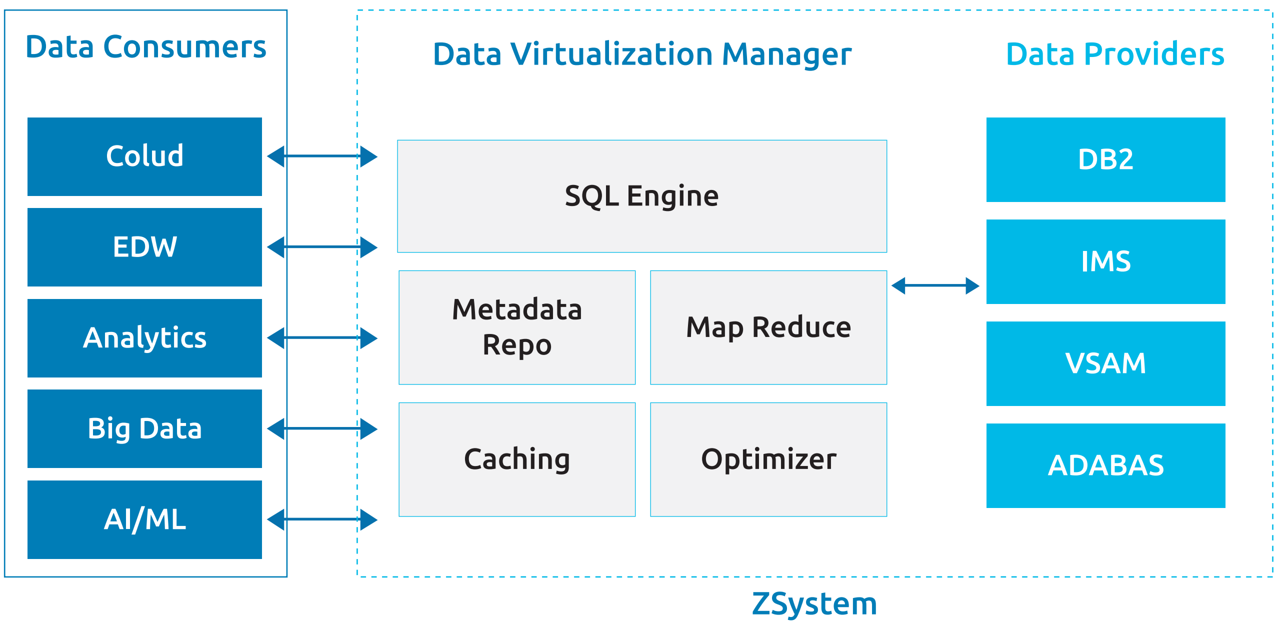


Virtualized integrated access to data stored in disparate database management systems

Mainframes have data stored in various formats, from modern relational database like DB2 to purpose specific non-relational database management systems like IMS, VSAM, IDCAMS or ADABAS. Due to the differences in storage and data retrieval mechanism, it becomes complex to have single integrated view of the mainframe data for business applications. The traditional ETL process of transforming and curating this data in consistent format in separate data store is time consuming, involves storage , processing cost and requires periodic administrative maintenance.

IBM's Data Virtualization Manager (DVM) allows to access data in real-time, without relocating the data from source and also to work with different sets of data structures stored on a mainframe virtually. It allows to create data views that can integrate the data elements from variety of database management systems such as DB2, VSAM or IMS.

Figure 8. Data Virtualization Manager to virtualize integrated access to disparate Z data stores



Organizations can leverage Data Virtualization Manager (DVM) to:

- Make it easier to read from and write to both relational and non-relational mainframe data source
- Create combined views of data from various sources for business applications
- Cut down on the number of data copies, keeping applications up-to-date and consistent by using the original mainframe data directly.
- Connect mainframe data with either custom-built or off-the-shelf cloud applications through standard

Bring AI/ML to data at source

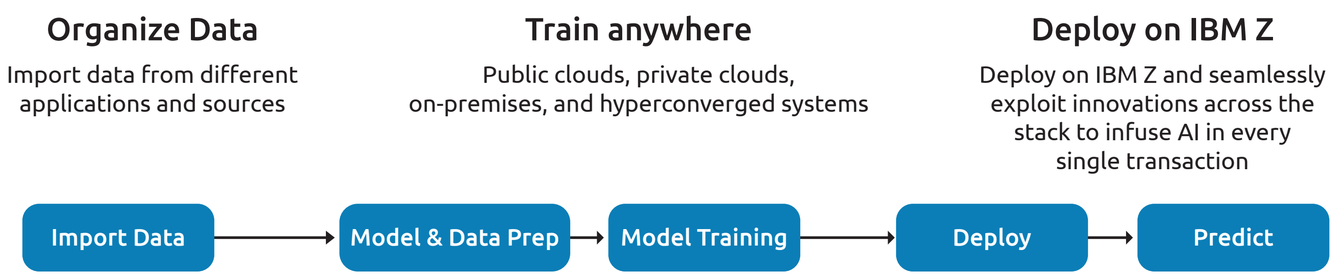
Organizations are rapidly adopting usage of AI/ML to extract value out of a staggering amount of data being collected due to digital transformation. The AI/ML workloads are currently run on data curated on either on-premise or cloud hosted Datawarehouse which generally does not have real-time data.

The value extracted from real-time data will be significantly higher for some of the popular AI/ML processes like fraud detection, sanctions and anti-money laundering. Financial organizations are currently only able to score 10% of transactions for fraud detection in real-time, due to latency involved in accessing data across servers via network to make accurate decisions. Moving data from origination platform (mainframe) to processing platform (cloud).

After launch of z16 machines, it is possible to execute AI/ML models on mainframe using Machine Learning for z (MLz). Also, with the launch of z17, telum II and AI hardware accelerator can handle 450 billion inference requests per day with 1 ms latency. This makes it possible for credit card companies to detect fraud in real-time and make a timely detection.

This modernization approach focuses on executing already (externally) trained AI models on mainframe, the development and training of models can happen on any of the cloud platform. This could have game changing implications on how AI/ML models are developed, trained and executed in future for Mainframe Applications.

Figure 9. Train anywhere and deploy on z workflow for AI/ML models



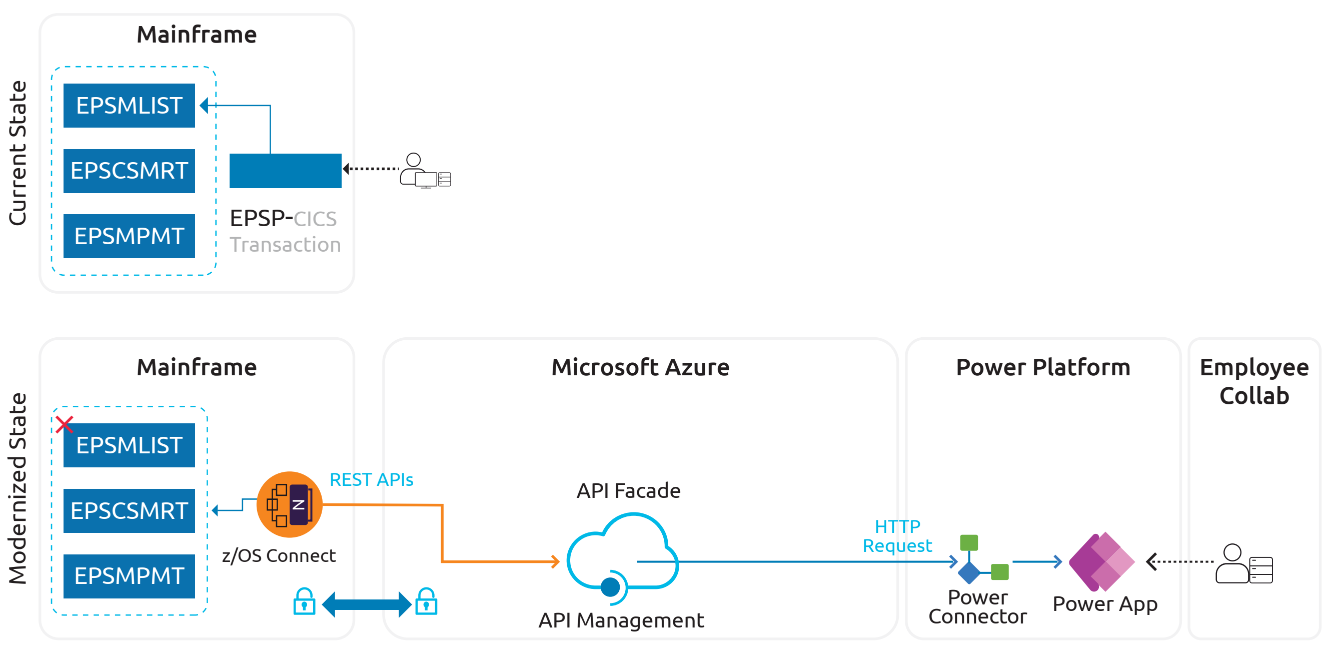
Modernizing legacy CICS / IMS interface with Low-code / No-code platform

Most of the organizations have modernized their customer facing applications to digital channels over the last two decades but the internal applications used by front and back office staff are still written in legacy CICS or IMS-DC interfaces. These interfaces have limited modern features and have steep learning curve for new staff. In addition, it poses a great challenge for the organizations to integrate these assets with rest of their enterprise platform. Hence, Organizations are looking to modernize these applications with modern user interface to improve staff productivity and shorten the learning curve for new staff.

In addition, Organizations could extend their business critical applications to the modern interfaces .

Microsoft power platform provides low-code/no-code tools to quickly modernize the legacy screens to modern user interface. The z/OS connect is IBM's low-code API generation platform for creating Open APIs with existing CICS application which can be discovered with any compatible Open API connecter. The below diagram shows the modernization of one such application from CICS to Microsoft Power Platform:

Figure 10. UI Modernization with Microsoft Power Platform



Benefits of using low-code / No-code programming platform:

- Rapid app development with low cost and low risk
- Provide business agility for digital transformation
- Re-imagine with new ways of working
- Integrate with modern cloud based applications

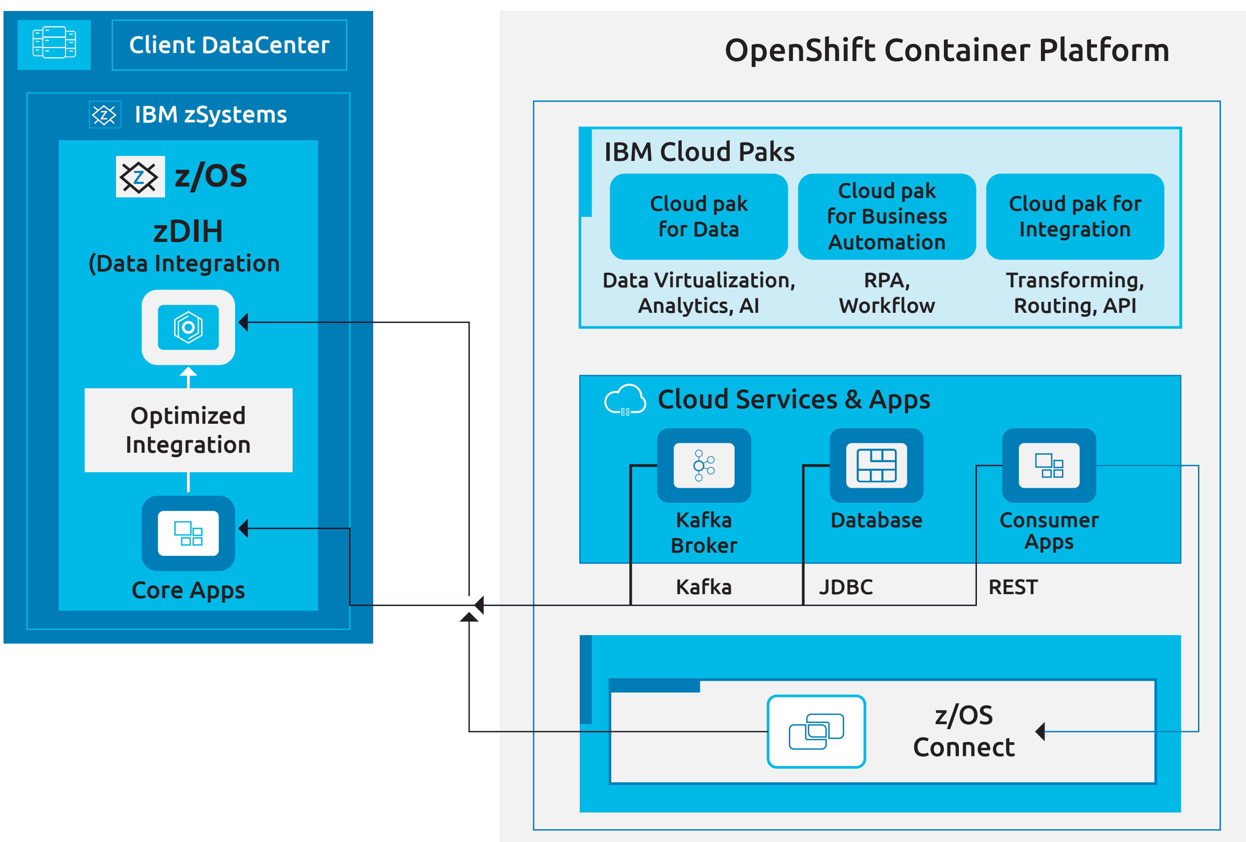
Modernizing in Hybrid Cloud with Red Hat OpenShift

Red Hat's OpenShift is a container application platform built on Kubernetes, designed to simplify the deployment, scaling, and management of containerized applications. It provides a robust set of tools and services for developers and IT operations teams, making it easier to adopt containerized workflows and leverage the power of Kubernetes for enterprise applications.

Red Hat OpenShift platform is being widely adopted as a popular hybrid cloud platform by organizations due to its ability to seamlessly operate either on-premise or any public cloud platform. After acquisition of Red Hat, IBM has extended this platform to operate on Z, enabling organizations to modernize the Z applications using cloud native technologies.

The below diagram shows the reference architecture of how the Z applications can be modernized with OpenShift platform:

Figure 11. Modernizing Z apps in hybrid cloud with Red Hat OpenShift



The reference architecture shown above uses IBM technologies that have been already discussed in this paper:

- Z/OS Connect - Create Open APIs to provide access to business logic and data
- Z Digital Integration Hub – Event driven flow of information from mainframes to cloud apps
- Cloud Pak for Data – Data from mainframes and cloud can be virtualized for 360-dgree views
- Cloud Pak for Business Automation – Define workflows based on new data from mainframe apps
- Cloud Pak for Integration - Provides transformation, routing and API services to integrate mainframe data with other cloud services

Using Red Hat OpenShift on IBM Z architecture for mainframe modernization offers a range of compelling benefits. This combination leverages the strengths of IBM Z's robust hardware and OpenShift's container orchestration capabilities to modernize legacy systems while maintaining performance, security, and scalability.

Some of the key advantages of using this architecture are:

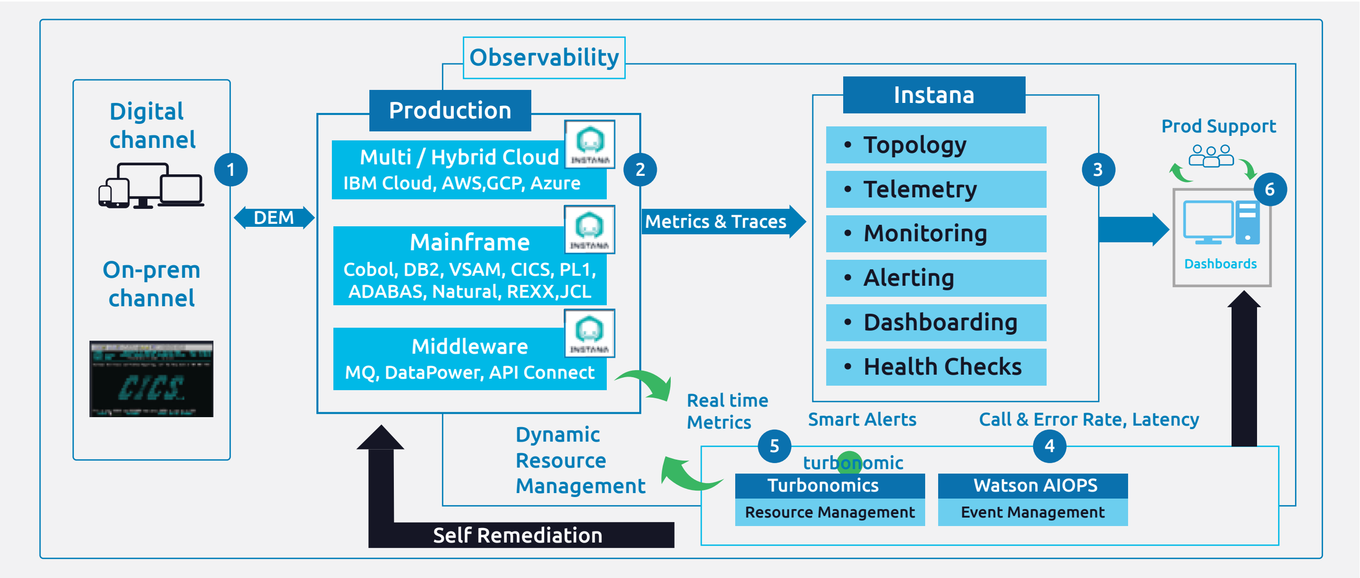
- Enhanced Modernization with Cloud-Native Capabilities by bringing Kubernetes based architecture to mainframe
- Cost Efficiency with zIP Utilization as the workloads running on OpenShift on Z are zIP enabled
- Seamless Hybrid Cloud Integration by providing a consistent platform across IBM Z, Linux on Z, and other public cloud environments
- Performance optimization and low latency by co-locating modern containerized applications with legacy z/OS workloads on the same IBM Z hardware
- Simplified developer experience with tools like CodeReady workspaces and integrated development environments

End to End Observability for hybrid cloud applications

As mentioned earlier in this document, financial services organizations are increasingly adopting hybrid cloud platform to run their IT workloads. The hybrid cloud provides them flexibility and agility along with regulatory compliance around data sovereignty but also comes-up with the challenge of managing and monitoring diverse and interconnected environment.

The organizations are planning for observability of cloud applications from the inceptions, but these implementations generally lack the end to end insights, especially from mainframe components. IBM's Instana, Turbonomics and Watson AIOps address many of these end to end observability challenge as shown below:

Figure 12. End to End Enterprise Observability stack with Mainframe



The components involved in this solution are:

Digital Experience metrics (DEM) Consist of on-perm channel and digital channel real user monitoring metrics

Agent Instrumentation: Instana agent instrumentation on mainframe, hybrid cloud and middleware which sends logs, metrics and traces to Instana

IBM Instana: A powerful application performance management (APM) tool that provides observability, monitoring, and automation for applications

IBM Cloud Pak for AIOps: Ingests data from various sources, detects anomalies, identifies root causes, and automates responses

IBM Turbonomic: Provides Dynamic Resource management using real time metrics for resource and cost optimization

IBM Z APM: Provides OpenTelemetry to deliver transaction trace for mainframe and remove blind-spots within application operations.

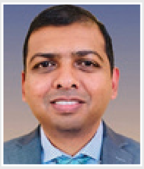
Single Pane of view: Provides a unified view of the entire IT landscape and improve root cause analysis

CONCLUSION

Mainframe modernization does not have to be migrating away from mainframe. Capgemini and IBM are bringing a balanced approach to help customers on their modernization journey to choose the right path for them between migration v/s in-place modernization.

• Performance result is extrapolated from IBM internal tests running local inference operations in an IBM z16 LPAR with 48 IFLs and 128 GB memory on Ubuntu 20.04 (SMT mode) using a synthetic credit card fraud detection model (https://github.com/IBM/ai -on-X-fraud-detection) exploiting the Integrated Accelerator for AI. The benchmark was running with 8 parallel threads each pinned to the first core of a different chip. The scpu command was used to identify the core-chip topology. A batch size of 128 inference operations was used. Results were also reproduced using a z/OS V2R4 LPAR with 24 CPs and 256GB memory on IBM z16. The same credit card fraud detection model was used. The benchmark was executed with a single thread performing inference operations. A batch size of 128 inference operations was used. Results may vary.

Authors



Nitin Chavan
is a Senior Director and Cloud Champion with Capgemini. His focus area is to help customers to adopt to hybrid cloud journey using IBM and Red Hat technologies including Z system modernization. He has 19 years of experience in providing technology solutions and has led multiple large transformation engagements for Z application and data modernization for financial services customers.

Dr. Gururaj Rao
is an IBM Fellow & GSI Ecosystem Executive at IBM Systems Group leading partnerships with worldwide Systems Integrators to expand the coverage to grow z Systems. He has been a technology and technical leader establishing the technical direction for IBM's Enterprise Systems working with inter-divisional teams of hardware, software and Research engineers and working with enterprise customers and IBM's client facing units.

About Capgemini

Capgemini is a global business and technology transformation partner, helping organizations to accelerate their dual transition to a digital and sustainable world, while creating tangible impact for enterprises and society. It is a responsible and diverse group of 340,000 team members in more than 50 countries. With its strong over 55-year heritage, Capgemini is trusted by its clients to unlock the value of technology to address the entire breadth of their business needs. It delivers end-to-end services and solutions leveraging strengths from strategy and design to engineering, all fueled by its market leading capabilities in AI, generative AI, cloud and data, combined with its deep industry expertise and partner ecosystem. The Group reported 2024 global revenues of €22.1 billion.

Get the future you want | www.capgemini.com