Redbooks
ibm.com/redbooks

Four Ways to Transform Your Mainframe for a Hybrid Cloud World

IBM Garage Technical Enablement Series

Guillaume Arnould

Guillaume Hoareau

Herve Sabrie

Sebastien Llaurency

Yann Kindelberger

IBM

Redpaper

Mainframe application modernization: The four strategic dimensions

The IBM® mainframe remains a widely used enterprise computing workhorse, hosting essential IT for the majority of the world's top banks, airlines, insurers and more. As the mainframe continues to evolve, the newest IBM Z® servers offer solutions for AI and analytics, blockchain, cloud, DevOps, security and resiliency, with the aim of making the client experience similar to that of using cloud services.

Many organizations today face challenges with their core IT infrastructure:

- Complexity and stability

  An environment might have years of history and be seen as too complex to maintain or update. Problems with system stability can impact operations and be considered a high risk for the business.

- Workforce challenges

  Many data center teams are anticipating a skills shortage within the next 5 years due to a retiring and declining workforce specialized in the mainframe, not to mention the difficulty of attracting new talent.

- Total cost of ownership

  Some infrastructure solutions are seen as too expensive, and it's not always easy to balance up-front costs with the life expectancy and benefits of a given platform.

- Lack of speed and agility

  Older applications can be seen as too slow and monolithic as organizations face an increasing need for faster turnaround and release cycles.

Some software vendors suggest addressing these challenges with the "big bang" approach of moving your entire environment to a public cloud. But public cloud isn't the best option for every workload, and a hybrid multicloud approach can offer the best of both worlds. IBM Z is constantly being developed to address the real challenges businesses face today, and every day we're helping clients modernize their IT environments.

Consider the following strategic elements when modernizing your mainframe environment:

- Infrastructure
- Applications
- Data access
- DevOps chain

This IBM Redpaper® publication focuses on these modernization dimensions.

Infrastructure modernization

Most IBM clients' mainframe systems are operating on the latest IBM Z hardware, but some are using earlier systems. The first step in updating your mainframe environment is adopting the newest features that can help you get the most from your infrastructure. Many technical innovations were introduced in IBM z15™. This platform has been engineered to encrypt data everywhere, provide for cloud-native development, and offer a high level of stability and availability so workloads can run continuously.

Application modernization

Core system applications — implemented as monolithic applications — form the backbone of many enterprises' IT. The key characteristic of these monolithic applications is the hyper integration of the main components of the application, which makes it difficult to understand and update them. Modernizing your mainframe applications starts with creating a map to identify which applications should follow a modularization process and which should be refactored. This implies working on APIs and microservices for better integration of the mainframe with other IT systems and often redefining the business rules. You might also move some modules or applications to the cloud using containerization.

Data access modernization

For years, some businesses have chosen to move their sensitive data off IBM Z to platforms that include data lakes and warehouses for analytics processing. Modern businesses need actionable and timely insight from their most current data; they can't afford the time required to copy and transform data. To address the need for actionable insights from data in real time and the cost of the security exposures due to data movement off the mainframe, IBM Z offers modern data management solutions, such as production data virtualization, production data replication in memory, and data acceleration for data warehouse and machine learning solutions.

DevOps chain modernization

The pressure to develop, debug, test, release and deploy applications quickly is increasing. IT teams that don't embrace DevOps are slower to deliver software and less responsive to the business's needs. IBM Z can help clients learn how to modernize through new DevOps tools and processes to create a lean and agile DevOps pipeline from modern source-code management to the provisioning of environments and the deployment of the artifacts.

This paper will highlight how to modernize through new DevOps tools and processes, to create a lean and agile continuous integration / continuous development (CI/CD) DevOps pipeline from a modern source code management, to the provisioning of the environments and the deployment of the artifacts. We will go through these four different pillars and explain how to modernize the mainframe.

Infrastructure modernization

The world has changed, and we are living in a dangerous world. In every international crisis, there is now a risk of cyber attack in the background. Cyber activities orchestrated by very organized entities demonstrate that day after day they develop new and sophisticated types of attacks to slow down, to destroy, and ultimately to remove from the market targeted organization.

This is one of the reasons why most of security standards and regulations raised the bar of their expectations to a higher level, and faster than before. They focused only in a recent past on data breach and security, and nowadays they extend to data privacy and cyber resilience.

In September 2019, IBM introduced the IBM z15, the latest evolution of the IBM Z technologies, designed for cyber resilience, security and privacy. It introduces leading-edge new technologies to comply with today's and tomorrow's challenges, helping organization to better protect their data at speed and volume, in motion and at rest, on and off the IBM Z platform.

Hardware evolution to tackle new security challenges

At first with the launch of IBM z15, IBM announced the important evolution of hardware encryption features. The CP Assist Cryptographic Functions (or simply CPACF) embedded inside each processor chip, supports new encryption algorithms that matter to encrypt data faster, and to reduce the increase in CPU utilization. The 7th generation of the hardware encryption features in PCI Express (the Crypto Express 7S cards or simply CEX7S) reach new performance records and exist in new packaging to comply more than ever with the client's infrastructure requirements. These improvements are an opportunity for organizations to encrypt more data thanks to the incredible encryption bandwidth available by design and by default in every IBM z15 configuration.

The second noticeable improvement from the hardware, the new IBM z15, embeds a new integrated accelerator for zEnterprise® Data Compression (or simply zEDC), inside each processor chip, very much like the encryption is with CPACF. Now customers can have the best of both worlds with compression and encryption (in that order) right on the processor chip. Encryption becomes even less expensive, since after compression, there is much less data to encrypt.

A significant step forward preparing the security of tomorrow is done with the new IBM z15. Quantum computing capabilities and their use, are growing—and will explode over the next 10 - 20 years. We all know that a key quantum computing use case is code breaking, and that includes intuiting encryption keys or breaking cryptographic algorithms in a very fast way. This is why, to make the data secure today and tomorrow, IBM Z is starting down the path for crypto agility by providing quantum-safe digital signing algorithms as part of the base system. As an initial use case, z/OS® audit logs can be dual signed with one National Institute of Standards and Technology (NIST) certified digital signature and one quantum-same digital signature in order to provide clients an early view of this new technology.

Data security journey to protect the data on the platform

In July 2017, IBM announced its new IBM z14® mainframe server, which combined both traditional mainframe hardware and new capabilities in areas such as cloud, cognitive, analytics and more. Most importantly, the z14 included a strategy security feature named Pervasive Encryption for IBM Z or simply PE to help clients stay one step ahead of cyber threats.

Pervasive Encryption for IBM Z is a consumable approach to enable extensive encryption of data in-flight and at-rest to substantially simplify encryption, and reduce costs associated with protecting data.

With the z14, for the first time in the 50-year history of mainframe technology, encryption became pervasive. This feature added software-based security intelligence to the mainframe's robust encryption mechanism allowing security solutions to leverage hardware-based cryptography like never before.

To protect the data at rest, Pervasive Encryption for IBM Z can be used in conjunction with full disk and tape encryption, database encryption, and application encryption. This multi layers encryption approach helps to address main enterprise data security risks:

- Full disk and tape encryption

  Full disk and tape encryption protects against intrusion, tamper or removal of physical infrastructure.

- File or data set encryption

  Volumes encryption on LinuxONE and Linux on IBM Z. Data set encryption for z/OS, managed through z/OS and providing simple policy controls that allow clients to protect data in mission critical databases including IBM DB2®, IMS, and VSAM. Additionally, z/OS data set encryption gives clients the ability to eliminate storage administrators from the compliance scope.

- Database encryption

  Database encryption provides selective encryption and granular key management control of sensitive data.

- Application encryption

  Application encryption is used to encrypt sensitive data when lower levels of encryption are not available or suitable.

Figure 1 summarizes the value of the multi-layer encryption approach and addresses different risks.

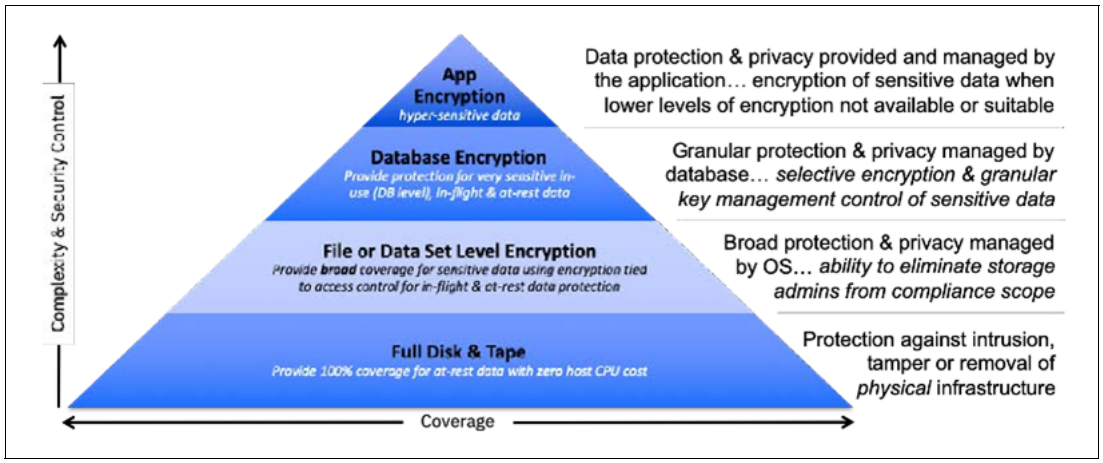

Figure 1 The pervasive encryption pyramid and multi-layer approach

To protect the data in motion, Pervasive Encryption for IBM Z relies on an early announced set of functionalities to protect the network sessions according to today's standards. This includes the following:

- Traditional secured network protocol support

  The objective is to secure the data from/to the IBM Z thanks to the secured network protocol implementation. This includes the support of SSL/TLS, SSH, IPSec.

- zERT

  zERT stands for z Encryption Readiness Tool. It is literally a z/OS network scanning tool to help z/OS administrators to be sure that the network sessions and connection are secured and if secured, indicates the quality of the network security.

- z/OS Coupling Facility Encryption

  Coupling Facility (CF) encryption helps to protect z/OS Coupling Facility data. CF encryption processes are transparent to applications leveraging CF structures. Encryption is based on policies which are established on a workload and Coupling Facility structure basis, to identify data that is to be encrypted before being sent to the Coupling Facility.

- End-Point Fiber Channel Encryption

  Fiber Channel is the premier transport for Storage Area Networks. In September 2019, IBM introduced the IBM z15, to extend the IBM Z position as with a new feature named “End Point Fiber Channel Encryption”. It better protects data circulating from the storage to the OS encrypting the network flow at hardware level. This offering provides in-flight protection for all data, independent of the operating system, file system, or access method in use.

Pervasive Encryption for IBM Z consumable features help many pain points associated with the EU’s upcoming General Data Protection Regulation (GDPR), which governs how companies around the world handle personal data belonging to EU residents.

We can encrypt today’s data on the IBM Z platform with no application changes, and no impact in SLA. The schema in Figure 2 summarizes the value of the pervasive encryption approach, protecting both data at rest and in motion on IBM Z Platform.

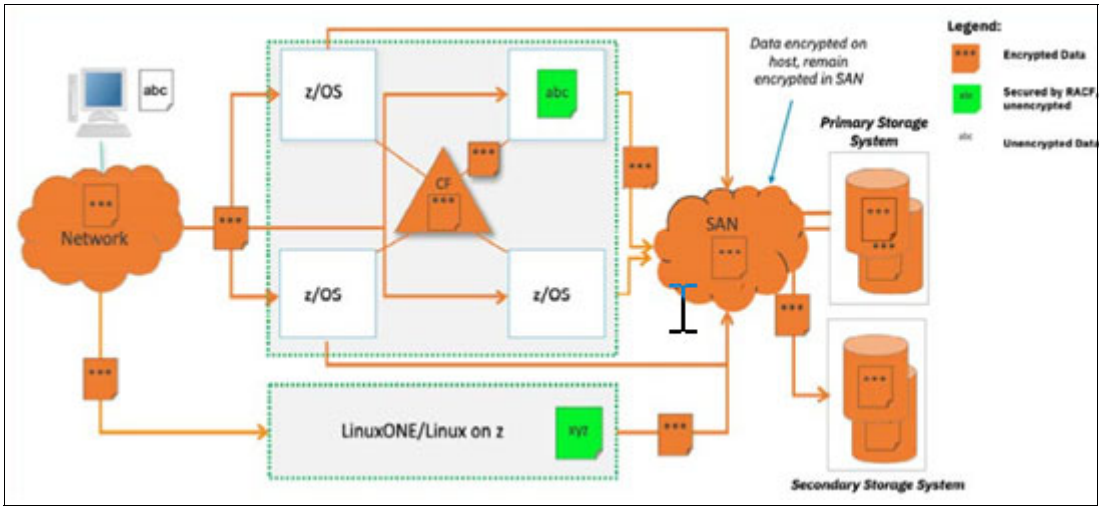

Figure 2 Protected data on the IBM z infrastructure

Data privacy journey to protect the data off the platform

Businesses and organizations are very concerned with ensuring that data shared with their own networks, that often include third-party partners, remains protected, accessible, and private. That data is protected today within IBM Z thanks to pervasive encryption. The next step with the IBM z15 is to protect that data even as it moves throughout the enterprise (see Figure 3 on page 6).

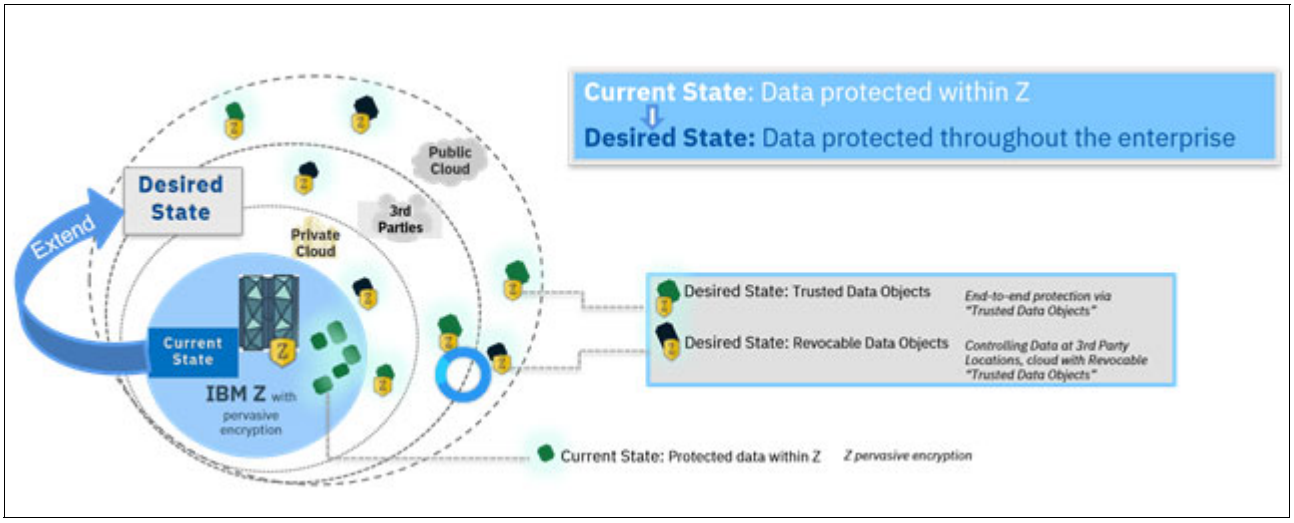

Figure 3 Data protection beyond IBM Z

In September 2019, IBM introduced the IBM z15, to extend the IBM Z position as the industry-leading platform for mission-critical hybrid cloud, with new innovations across security, data privacy, and resilience.

Data security requires a multi-level approach in order to effectively reduce the risk of breaches among businesses. Pervasive encryption encrypts all data associated with an application, database, or cloud service whether on premises, at rest or in flight. The IBM z15 extends this beyond the border of the IBM Z environment. The data-centric IBM z15 offers security solutions to simultaneously address breaches and provide privacy and ease of operations for any business operating in the connected IBM Z landscape.

The data privacy journey to protect data off the IBM Z platform is based on the two following technologies introduced with the IBM z15:

- IBM Z Data Privacy for Diagnostics

  IBM Data Privacy Passports provides clients with the capability to protect sensitive data from source DBMS. But there is also the use case of data that organizations have to care for regarding the exfiltration risk: dumps. IBM Z Data Privacy for Diagnostics provides clients with the capability to protect sensitive data that may be included in diagnostic dumps. Now sensitive data can be tagged such that it can be identified in dumps with no impact to dump capture times. Tagged sensitive data in dumps can be secured and redacted before sending to third-party vendors.

- IBM Data Privacy Passports

  The IBM Z Data Privacy Passports, in conjunction with IBM z15, is designed to enforce security and privacy protections to data not only on IBM Z but across platforms (including cloud and distributed environments).

There are two privacy services delivered by IBM Data Privacy Passports:

- Data protection: Protecting the data at the exfiltration point.
- Data enforcement: Enforcing the data at the consumption point.

Data protection

Data protection is about protecting the data at the exfiltration point. To do so, IBM Data Privacy Passport provides a data-centric security (data protection mechanism stays with the data) solution that enables data to play an active role in its own protection across the enterprise.This offering is the next logical step from the IBM z14 Pervasive Encryption for IBM Z offering, now extending the IBM Z leadership in security and data protection to data not only resident on the IBM Z platform but also as it moves throughout the enterprise and beyond.

The concept of Data Centric Audit and Protection (DCAP) is a transition from the current model most enterprises have become accustomed to, as illustrated in Figure 4.

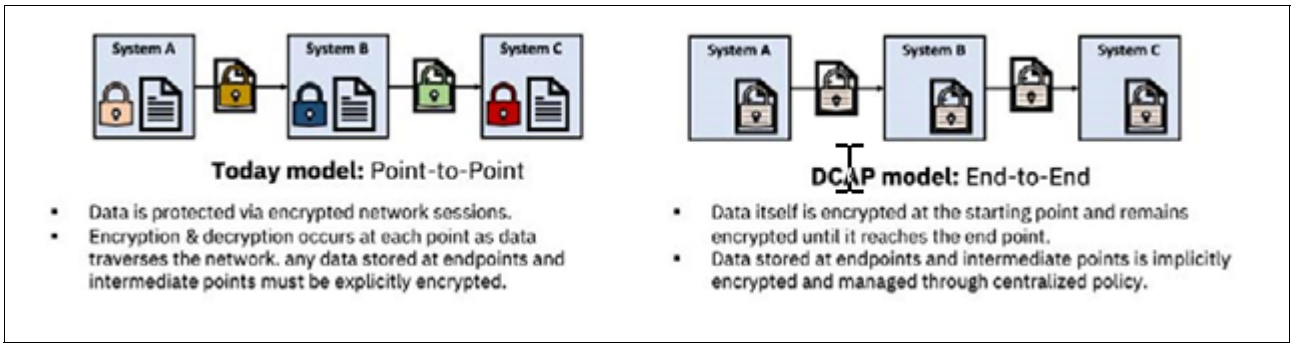

Figure 4 Current model of data protection vs. DCAP

In the DCAP model, before the data is moved around the enterprise, it will be repackaged into a secure object. In the case of Data Privacy Passports, this is call the Trusted Data Object (TDO).

The Data Privacy Passport Offering does this protection at a field level, which means that there is a level of granularity to this protection that cannot be obtained from more broad protection techniques. Once the field is wrapped in a TDO, that TDO moves throughout the enterprise.

Figure 5 shows the lifetime of a data source protected with IBM Data Privacy Passports thanks to the Data Protection function.

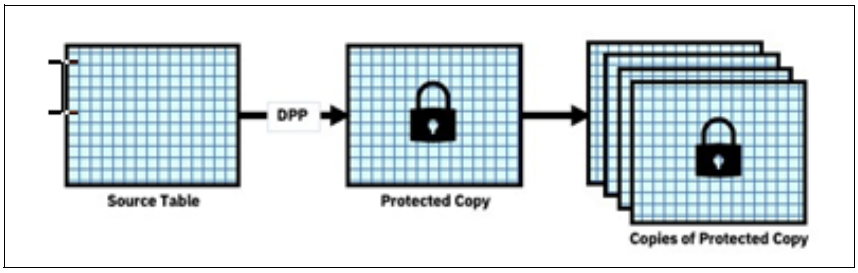

Figure 5 Copies of trusted data objects

A source table to be shared is copied and then protected via encryption according to the defined fine grain privacy policy. A protected copy and its copies are called a TDO and comply with the DCAP concept that the data stays with the security mechanism. Only IBM Data Privacy Passports has the required keys to decrypt the content of a TDO and its copies.

When any original TDO or its copies need to be decrypted, the TDO must come back to the IBM Data Privacy Passports infrastructure. Otherwise, it is impossible to read the encrypted data of a TDO. This is how a lost, an exfiltrated, or a breach copy of a data source remains protected by default.

Data enforcement

Data enforcement is about enforcing the data at the consumption point. To do so, IBM Data Privacy Passports provides an Enforced View of the data upon SQL queries. The Enforced View will depend on the user's credentials that is provisioning the data. The data as a source may be in the clear (original source of data), or a Trusted Data Object. In this last case, we can combine both data protection and data enforcement together to secure the data at the point of extraction and consumption as illustrated in Figure 6.

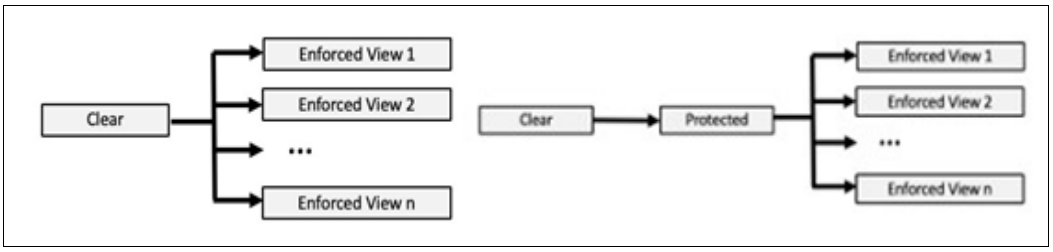

Figure 6 Data enforcement: One source, multiple views according to the need to know

The IBM Z is weaponized to protect its data, and the data of others at the point of extraction, and the point of consumption. IBM Z environment can reduce the risk associated with data exfiltration covering IBM Z DBMS, and non-IBM Z DBMS. This is an important factor of synergies and integration as expected to confirm its role of security hub, in the multi-cloud driven enterprise IT infrastructure.

Application modernization

The fundamental challenges of legacy applications to support the new markets forces and business requirements is the lack of modularity. This situation is the result of decades of fixes and patching creating layers over layers of an old-core system, resulting in an extremely complex monolithic structure that has created an uncontrolled and unknown hyper-integrated model of processes, data, and transactionality.

Due to this lack of modularity, any change requires a deep and long analysis and costly regression tests. It definitively impacts the agility to move the customer needs and requirements to the system in production. This also results in high costs of development and maintenance, as well as obsolescence risks, due to the difficulties to update outdated middleware.

From hyper-integration to hyper-modularity

To address the challenges due to lack of modularity, the recommendation is to modernize those core legacy applications by shifting to a new paradigm from hyper-integration to hyper-modularity.

The modularity does have several dimensions that are crucial for an agile modular system:

- Structural modularization

  The system is composed of modules, each one with a clear responsibility and interface. The system is structured as a set of collaborating Modules in a loosely coupled way.

- Development modularity

  The strict encapsulation and interface allow very loosely coupled design and development life cycles, extremely useful for agile development techniques.

- Operational modularity

  Each module is operated independently of other modules, running on different and isolated runtime environments depending on the implementation technology.

A strict modularized approach has the following advantages:

- Less maintenance and evolution costs. Since the development of each module is very isolated and the inter-module interfaces are clearly defined, much less regression testing is needed, being the maintenance and evolution are much simpler.
- Much faster response to the business needs. This is related to the previous point.
- System robustness. Module fencing and isolation helps to prevent failure cascade propagation.
- Future proof architecture. Modules can be renewed or replaced by new technologies with low impact on the rest of the system. Technical obsolescence can be managed easily as a technological replacement only impacts the specific module.
- Modularity is a key condition in moving workloads to the cloud. The lack of modularity will prevent the cloud journey.

The application modernization approach

The modularization or refactorization paradigm will allow the modernization of each module independently using different technologies and IBM solutions, as shown in Figure 7.

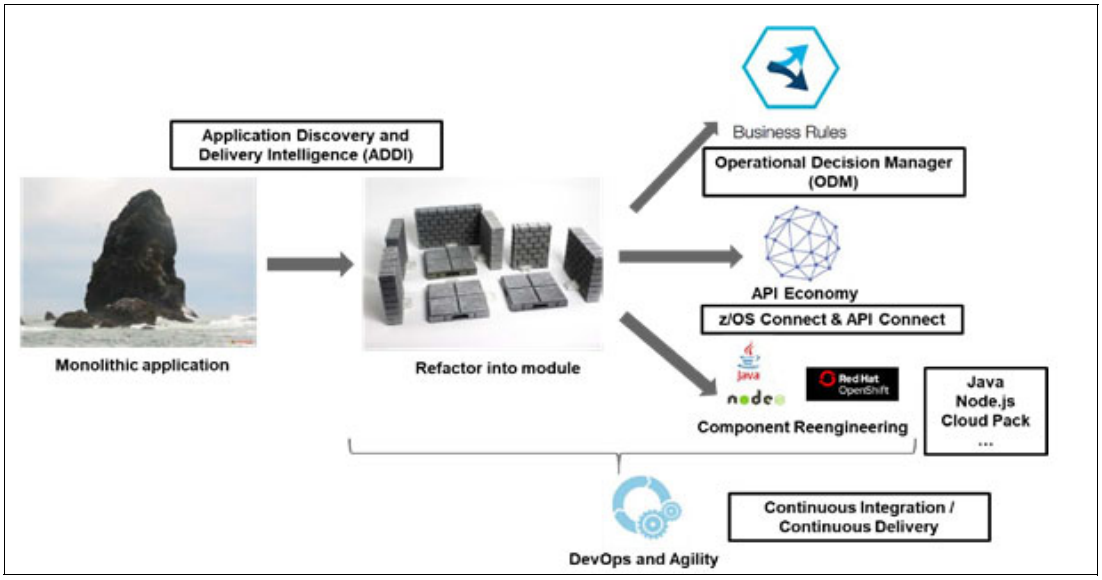

Figure 7 Approach for application modernization

The first step consists in the refactorization into modules. Then multiple options are available to modernize each module of the application:

- Externalize decision logic to a business rules management system.
- Expose the business logic through APIs.
- Re-engineer some modules and consider a move to cloud approach for some of them.

Lastly, all those applicative components have to be managed through new modern DevOps tools and processes in order to create an agile CI/CD pipeline.

Application assessment

In order to refactor the monolith application into modules, the first step will be to analyze the application to discover and understand the interdependencies between the components. A manual analysis of these applications is time-consuming and limits the capacity for modernization. Tooling will definitively be helpful in this area.

IBM Application Discovery and Delivery Intelligence is an analytical platform that helps developers rapidly discover and analyze relationships between application components, data, and jobs to identify the scope of each module and to make changes safely and efficiently.

The broad support for IBM and non-IBM languages and environments, as well as a visual analysis of the application dependencies will help to quickly understand the applications and to define the future modules.

Modernize with business rules

The first option to modernize the application is the externalization of the decision logic to a business rules management system. Many traditional mainframe applications (COBOL, PL/1) have hard-wired business rules or decision logic encoded into the application making it complex and hard to update.

Moving and maintaining rules outside the application in a rule's engine can dramatically improve business responsiveness. The externalization of the decision logic to a business rules management system will bring the following advantage:

- Significant cost savings from reducing the change cycle of an application.
- Removal of duplicated business logic by sharing the rules across platforms and languages.
- Increased agility by giving the business direct access to defining and updating rules and therefore removing the need for IT application changes.

Operational Decision Manager is a set of tooling and interfaces that provide a means to encapsulate decision logic away from program code into rules. These rules are rendered as natural language. When decisions are exposed in this format, they become more agile, adaptable to change, and visible to those who need them. By having a rich set of tooling, the change cycle of a rule and its deployment is much less than the traditional code development and deployment route.

Operational Decision Manager consists of two components, which together form a platform for the management and execution of business rules:

- The Decision Center, which provides an integrated repository and management components for line-of-business subject matter experts to directly participate in the governance of business rule-based decision logic.
- The Decision Server, which provides the runtime components to execute rule-based decision logic on mainframe systems.

The Decision Server for z/OS provides the ability to author business decisions based on COBOL copybook or PL/I include files. It then offers a number of z/OS based execution run times in which the business decisions can be executed from COBOL and PL/I applications. There are a number of options to execute decisions on z/OS, which allow IBM CICS®, IMS, and batch COBOL or PL/I applications to call the Decision Server for z/OS.

Expose through REST APIs

The second option to modernize the application is the exposition of the business logic through APIs. New cloud native applications place new requirements for a successful hybrid integration architecture in which services from the mainframe have to be discovered and consumed by using REST APIs. IBM has developed technologies to make sure that mainframe applications and data can be accessed by means of a REST APIs. This technology is called z/OS Connect Enterprise Edition (z/OS Connect EE) and runs under an IBM WebSphere® Liberty profile inside z/OS.

z/OS Connect EE provides a single common gateway for REST HTTP calls to reach business assets and data on z/OS operating systems. Where these assets run is specified in the z/OS Connect configuration, which relieves client applications in the cloud, mobile, and web worlds of the need to understand the details about how to reach them and how to convert payloads to and from the formats that the applications require. Services can be enabled without writing code and tooling is provided for creating the data transformation artifacts. With z/OS Connect EE, mobile and cloud application developers can incorporate z/OS data and transactions into their applications, whether they work inside or outside the enterprise, without needing to understand z/OS subsystems. Using z/OS Connect EE, the z/OS resources appear now as any other REST API.

Re-engineer some modules for a move to cloud approach

As a third option some customers may also decide to re-engineer some modules. They can take benefits of using Linux containers to modernize a part of their application.

These Linux containers can be run under:

- z/OS by using IBM z/OS Container Extensions (IBM zCX)
- IBM z/VM® by using Red Hat OpenShift Container Platform
- A Hyper Protected Appliance (Hyper Protect Virtual Servers)

In all cases, the modules/functions will need to be created as Linux container images specifically for the s390x architecture.

Let’s review those three ways of running Linux containers on IBM Z.

The first way to run Linux containers is to use IBM z/OS Container Extensions (IBM zCX) environments. It allows you to integrate Linux on Z with z/OS applications. Application developers can build on popular open-source solutions and deploy their Linux applications, IBM software, and third-party software in zCX environments that are as close to z/OS applications and data as possible, as shown in Figure 8 on page 12.

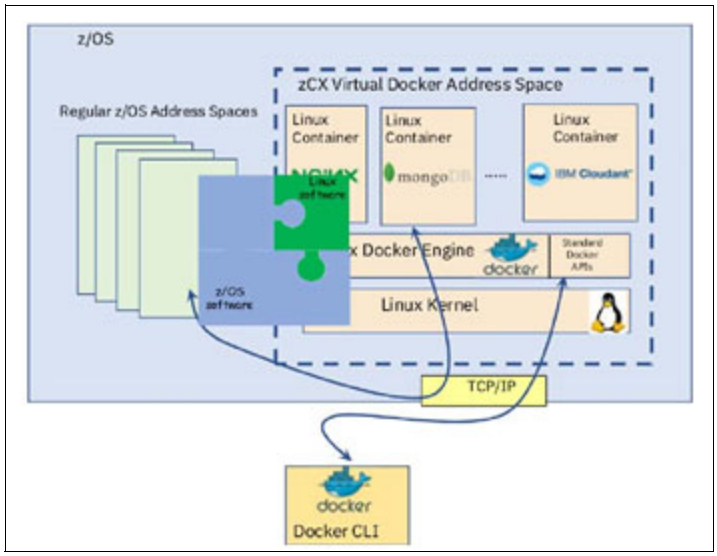

Figure 8 Architecture of a zCX environment

zCX is capable of running Linux containers (Docker containers as of today) but from a development point of view, it does not require a different source code repository or a different application build engine than the ones used for x86.

There are a number of benefits of using zCX:

- Develop and run microservices applications inside z/OS.
- Facilitate adoption and Linux usage inside z/OS.
- Allow a common CI/CD process for development.
- Benefits from z/OS quality of service for Linux workloads.

Note that at the time of this publication, zCX does not integrate with the standard container orchestration engine (namely: Kubernetes) to help to handle the containers.

The second way to run Linux containers is with Red Hat OpenShift Container Platform. The numerous innovations of the last years on IBM Z have today allowed Linux to be positioned as a strategic choice for many customers. It is the engine of growth to host a growing choice of modern applications, with or without affinities with z/OS environments.

The ability to run a Red Hat OpenShift Container Platform cluster will help to:

- Redefine the security perimeter for the next generation of applications by protecting against:
  - Threats related to data privacy
  - The theft of identifiers
  - Malware and ransomware
  - Fraudulent or unauthorized manipulation of databases
- Maintain reduced operational costs by benefiting from performance enhancements, scalability and simple environmental management.
- Integrate into multiple hybrid cloud environments.
- To be able to reuse the DevOps tools already in place for distributed and / or public cloud environments.

As a strategic investment in Linux and the IBM Z platform, the announcement of IBM Cloud® PAKs support on Linux on IBM Z reinforces IBM's importance and willingness to provide the latest cloud technologies on the platform. The IBM Cloud PAKs are IBM containerized software packaged with open source components, pre-integrated with the common operational services, and secure by design. They are IBM Certified, up-to-date software that secures the entire stack, from hardware to applications and portable, running on-premises, on public clouds, or in a pre-integrated system.

The benefits of running OpenShift on Linux IBM Z are:

- Perfect fit for dynamic workloads due to vertical scalability.
- Can support unpredicted microservices growth and spin.
- Fine granular capacity allocation and sharing through virtualization.
- Massive number of secured microservices on one physical server.
- Advantage of scale up vs. scale out on x86.
- Securable server and high-speed encryption.
- Certified multi-tenant security and fast crypto.
- Easy failover in case of errors or service crash.

As a third way of running Linux containers which may contain very sensitive information, is to use IBM Hyper Protect Virtual Servers. IBM Hyper Protect Virtual Servers protect critical Linux workloads during build, deployment, and management on-premise. Note that there is not, at the time of this publication, an integration with Red Hat OpenShift Container Platform. Customers can create secured virtual servers in which they can deploy docker and containerized images.

The benefits of running IBM Hyper Protect Virtual Servers are:

- Developers can securely build their applications in a trusted environment with integrity.
- IT infrastructure providers can manage the servers and virtualized environment where the applications are deployed without having access to those applications or their sensitive data.
- Application users can validate that those securely built applications originate from a trusted source by integrating this validation into their own auditing processes.
- Chief Information Security Officers (CISOs) can be confident that their data is both protected and private from internal and external threats.
- Security measures are delivered to address threat vectors that appear at different phases of an application’s lifecycle: build, deployment, and management.
- IBM Hyper Protect Virtual Servers are designed to uniquely protect such workloads that are deployed on IBM Z and LinuxONE servers in hybrid, multi-cloud environments.

Additionally, it is important to mention the capability to use multi-architecture Linux container images. This capability provides a way for the developer not to bother about the target hardware server platform where the application will be deployed to. The right Linux container image for the platform is picked up automatically (of course, it has first to be created once for the server’s hardware architecture) during the deployment phase of the application, *i.e.*, at the creation of the required Linux containers for the application on the target server.

IBM's CI / CD solutions address both distributed and mainframe developments. As described in “DevOps chain modernization” on page 18, a single orchestrator can be configured to manage the production and deployment of mainframe applications or managed microservices in containers across all the previously discussed solutions.

Data access modernization

As introduced earlier, we will discuss digital transformation and data modernization. We will share innovative IBM technologies that can help you leverage the valuable data that originates on the IBM Z platform to support that effort.

Digital transformation, modernization, and becoming more data-driven

As we all know, business cycles are accelerating, and new technologies are rapidly improving the way businesses compete, and disrupting customary business approaches. Disruptive technology is at the core of business upheaval. At the core of the disruptive technology is data. Organizations that leverage data and analytics to gain insight into their business processes and data tend to fare better in serving their customers and clients.

Data latency

Data, its use, and the insights it provides have a shelf life. Some decisions require the most current data and real-time insight (at the point of interaction or transaction) improving customer interactions and decision making in areas, such as fraud detection and up sell or cross sell efforts, and supporting real-time opportunities, such as digital moments.

This requires organizations to deliver data and insight when and where it is needed and in real-time. Organizations need to embed data and actionable insights into their modern engagement and transactional applications.

A common issue within organizations is that data is not being used effectively and large portions of data go unused. A large majority of executives are fearful of disruption by data-driven start-ups and few executives are completely confident in their future data strategy. An architectural option is to access enterprise data where it originates, so critical data driven decisions can be made before an interaction or transaction completes. Traditional data movement approaches can limit the opportunity to benefit from data where and when it is needed.

Unfortunately, many organizations still use the same tired-old architectural pattern of replicating data from its point of origin. Any action based on replicated data can never be real time and can often be sub-optimal.

Every copy of data has its unique cost. Data movement can impact data usage, time to insight, and potentially undermine some of the value of the modernization efforts. The time between when data originates and when it can be used by a downstream system matters. We call this *the data latency gap*.

Data gravity

Several years ago, the term “data gravity” was offered as an analogy to the concept of gravity in physics. Data gravity implies that data has mass, and therefore attracts other objects, like physical objects do. The more data there is, the greater the mass, the more likely it is to attract other objects such as applications, services and other data. Although the cloud often consumes most of everyone’s mindshare, most essential transactional data is still generated on the mainframe. From a cost, efficiency, regulatory, security and practical perspective, this moving data off of the mainframe just does not make sense. And as mentioned earlier, this data movement can impact customer success.

The IBM journey to modernize the data access

IBM Z technology can help access IBM Z data in place, with minimal latency and in a security-rich environment. Data and insight can be embedded in transactional applications for better customer engagement and real-time decision making that drives better business outcomes. The best strategy is to interact with the data where it originates or is stored. This aligns with the data gravity concept: applications and services are drawn to the mass of data.

Data virtualization with IBM Data Virtualization Manager for z/OS

Data virtualization has emerged as a cost effective, substitute for, and augmentation to, traditional data collection. With data virtualization, you can access data where it originates to reduce the time and resources used to combine data from multiple systems. Less time and resources can translate into savings.

IBM Data Virtualization Manager provides virtual, integrated views of data residing on IBM Z, and enables users and applications read/write access to IBM Z data in place, without having to move, replicate or transform data. And it performs these tasks with minimal additional processing costs. By unlocking IBM Z data using popular, industry-standard APIs such as JDBC, ODBC and REST, Data Virtualization Manager can save time and money.

Developers can readily combine IBM Z data with other enterprise data sources to gain real-time insight, accelerate deployment of new web and mobile applications, modernize the enterprise, and take advantage of today's API economy.

Data Virtualization Manager for z/OS (see Figure 9 on page 16) allows:

- Ready access to relational and non-relational IBM Z transactional data.
- Access and updates to live IBM Z data via modern APIs (when combined with z/OS Connect).
- Reduction of the cost and delay of moving data to non-Z platforms.
- Simplification of the development of applications accessing relational and non-relational data types including VSAM, IMS, ADABAS, IDMS, SMF and non-IBM Z data sources.
- Modernization of existing applications and reduction of data movement off platform.
- Reduction of data copies that result in latency, cost and risk (security, governance, decision latency).

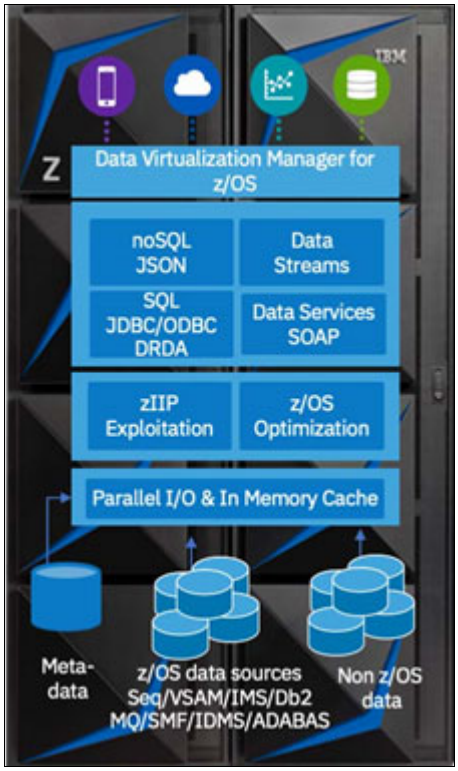

Figure 9 Data virtualization manager

Data replication in memory with IBM Open Data Analytics for z/OS

Apache Spark opens up access to a 100's of data sources to deliver analytics in real time in memory. Spark enables data scientists, developers and data engineers to work together to access all data, build analytic models quickly, iterate fast in a unified programming model, and deploy those analytics everywhere.

Apache Spark is a big data processing framework built for speed, ease of use, and sophisticated analytics that enables applications in Hadoop clusters, or stand alone, to run many times faster in memory and on disk. It is an in-memory compute engine that works with data; not a data store. Users can write applications quickly in a variety of languages (Java, Scala, Python, R). It also combines SQL, streaming, and complex analytics together. It runs everywhere while enabling highly iterative analysis on large volumes of data at scale. Apache Spark runs on Hadoop, Mesos, standalone, in the cloud or on IBM Z.

In 2019, IBM released IBM Open Data Analytics for z/OS (IzODA), depicted in Figure 10 on page 17, which is a bundle of Apache Spark libraries and packages running on z/OS. IBM z Systems® clients can take advantage of IzODA on both Linux on z Systems® and z/OS to leverage Apache Spark for a holistic view of their enterprise data.

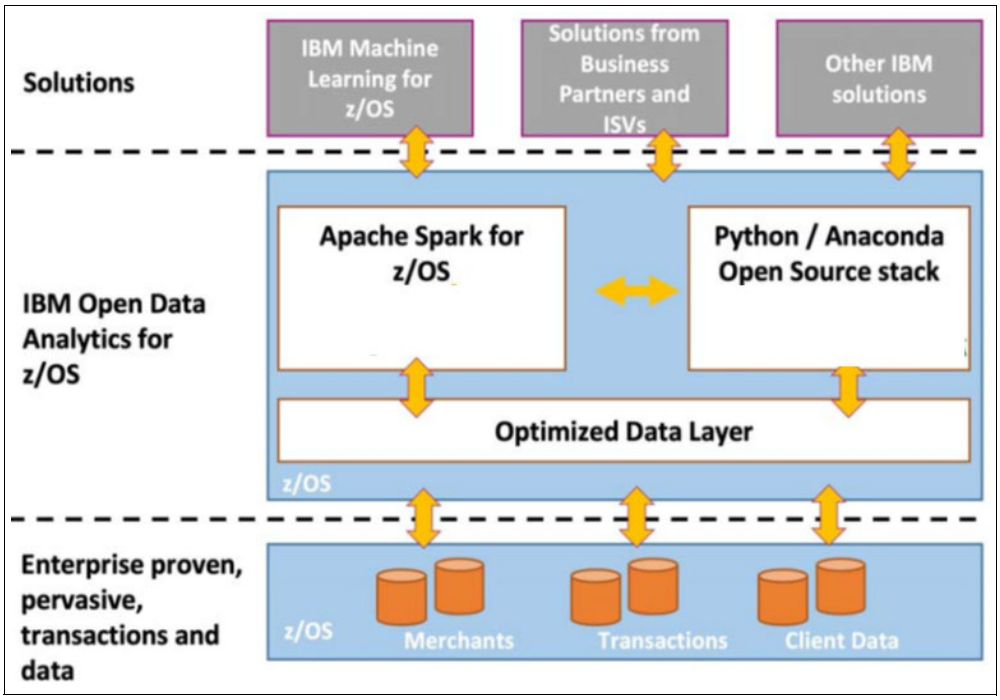

Figure 10 IBM Open Data Analytics for z/OS

IzODA leverages source data's original security specification. Because IzODA is co-located with data sources, it is optimized, uses parallel load of data resulting in faster end-to-end analytic performance and leverages Java on z/OS performance features. IzODA makes optimal use of zIIP for Spark & Optimized Data Layer. On IBM Z systems, you can scale-up and benefit from larger in-memory frameworks for optimal analytic performance. Using IzODA, you will also benefit through selected reduction of ETL and associated costs for analytics. A consistent access to a wide variety of data environments on and off the platform is made possible (DB2, VSAM, IMS, PS, PDSE, SMF, SYSLOG, ADABAS, IDMS, Oracle Enterprise, Teradata, HDFS, and so on).

AI with IBM Watson Machine Learning for z/OS

IBM Watson® Machine Learning for z/OS (WMLz) enables you to run machine learning on the IBM Z platform while keeping mission-critical data where it is and using the existing Z capabilities.

WMLz essentially provides IBM Z clients with a complete enterprise machine learning platform that leverages valuable business data to infuse mission-critical applications with intelligence. Specifically, WMLz provides an enterprise grade platform with the following capabilities:

- Secure IBM Z platform for running machine learning with data in place.
- Full lifecycle management of models.
- Enterprise grade performance and high availability.
- Flexible choice of machine learning languages.
- Intuitive self-guided model development.
- Developer-friendly API interfaces for applications on the Z platform.

WMLz leverages proven IBM machine learning capabilities, including IBM Open Data Analytics for z/OS (IzODA). Running on Z, IzODA serves as the data processing cluster for WML for z/OS and delivers advanced data analytics through z/OS Spark, z/OS Anaconda, and Mainframe Data Service (MDS). While it has been said that Spark provides the best-of-breed analytics engine for large-scale data processing and Anaconda supplies a wide range of popular Python packages for model training, MDS connects those data processing engines to your enterprise data sources on Z, such as IBM Db2®, IMS, VSAM, and SMF.

Finally, together with other IBM cognitive technologies, including IBM Db2 Analytics Accelerator for z/OS, WMLz can also help transform the mainframe platform into a highly efficient, hybrid transaction, and analytic processing environment. It can build, deploy, and manage behavioral models in real time by directly consuming data that is stored in Db2 for z/OS (or other sources) and transformed in the Accelerator. The data stays securely in the Accelerator as it is being ingested into WMLz, removing the latency between data creation and transactions.

IBM Db2 Analytics Accelerator on z/OS slashes the cost and complexity of analytics processing by centralizing analytics where the data resides, enabling analytics on data as it is generated. IBM Db2 Analytics Accelerator on z/OS is a high-performance database warehouse that processes Db2 queries at exceptional speed. IBM Db2 Analytics Accelerator for z/OS extends the Db2 for z/OS database system to accelerate processing of the complex, resource-intensive Db2 queries vital to organizations' mission-critical reporting and analytic workloads. Coupling the Accelerator with Db2 for z/OS delivers high-speed performance for mixed transactional and analytic workloads, allowing organizations to exploit the high-value, sensitive IBM Z data in realtime.

The Hybrid Transactional and Analytic Processing or HTAP allows analytics to be run against transactional data at the source, optimizing real-time decision making. This HTAP approach fully isolates transaction and analytic processing - each run in separate resource pools so analytic processing has no impact on transactional performance.

IBM Db2 Analytics Accelerator on z/OS is deployed in two different form factors, either as an IBM Integrated Analytics System appliance or directly on IBM Z in a virtual appliance form factor exploiting Integrated Facility for Linux (IFL) processors.

DevOps chain modernization

Let's start by explaining quickly the meaning of DevOps. To thrive in today's fast-moving economy, customers must accelerate their application development lifecycle. The pressure to develop, debug, test, release, and deploy quickly is increasing. DevOps can be described as the combination of development and IT operations through tools (the toolchain) and processes (the pipeline) to shorten the software development lifecycle to deliver features, fixes and updates quickly.

Still today, some mainframe development practices do not embrace modern approaches because they rely on rigid practices (silos, no automation…). In the meantime, some development tools have still archaic interfaces that limit the developer productivity. In some cases the current mainframe development processes and tools cannot keep up with the increase in volume and frequency of business changes requested by the line of business.

A modernization of the DevOps toolchain and pipeline is required. This modernization process begins by trying to narrow the gap between the two worlds: the mainframe and the cloud-friendliness. It can be achieved by the standardization of the delivery processes and practices across the two worlds and by leveraging synergistic effects across all development teams. It consists in integrating Z into a generic DevOps toolchain and pipeline, generally well adopted by the cloud-friendliness applications.

A generic DevOps pipeline is made up of several building blocks, each serving a dedicated purpose. Figure 11 shows the main building blocks.

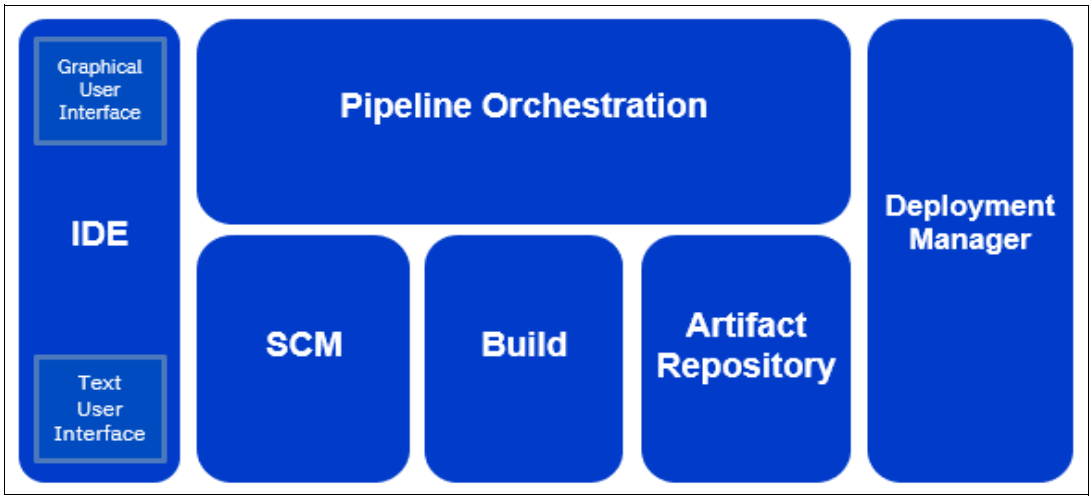

Figure 11 CI/CD pipeline building blocks

The integrated development environment (IDE) allows developers to make code changes. The files are managed by a source configuration manager (SCM), which provides configuration and version control management. These source files are built, and the binaries and outputs are created. These outputs can be managed, as part of an archive that we will call a “package”, in an artifact repository. The artifact repository makes outputs available to the deployment manager. Then, the deployment to various environments can begin. The automation and its rules during all these steps are provided by an orchestrator.

A good place to start with DevOps modernization is probably with the source-code management. A modern SCM should have the following capabilities:

The ability to truly provide full isolation by managing the full set of source code (Cobol, PL/I, Java, Node.js, etc.) and not a concatenation of data sets.

The ability to build artifacts that are stored and versioned in an artifact repository. Git is currently considered to be the de facto standard SCM in the distributed software development. Compared to a library manager on z/OS, which relies on concatenation for compilation and linking, Git provides complete isolation and remove dependencies on the areas under current development by other development teams.

IBM provides solutions such as IBM Dependency Based Build (DBB) that enables a true parallel development and a full integration with a modern Source Configuration Manager like GitHub and a CI/CD toolchain.

Adopting both a modern SCM like Git, DBB and an artifacts repository such as JFrog or an IBM UrbanCode® CodeStation ensure a complete decoupling of the build and deployment phases.

Another feature of this decoupling is that it enables developers to adopt provisioning practices such as spinning up quickly an isolated test execution environment.

IBM provides also various solutions in this area such as:

- IBM Z Development and Test Environment
- Allows you to develop and test mainframe applications on x86 hardware. It enables any z/OS software to run on a x86-compatible on-premise system or cloud instance by emulating the IBM Z instruction sets. The major advantage is to gain agility by enabling earlier and more frequent testing while freeing up mainframe capacity for more valuable production workloads.
- z/OS Cloud Broker
- Provides direct, self-service access to z/OS computing resources by end users through a supported cloud platform like OpenShift. It allows for a full self-service model of z/OS resource consumption and empowers developers to use z/OS in a cloud-native way.

Another key component in a modern DevOps toolchain is the deployment manager. The deployment manager is responsible for understanding the execution environments and maintains an inventory of the environment's deployed content. It is used to rollout application packages stored in the artifact's repository. A solution such as Urban Code Deploy is a software product for automating application deployments across the various IT environments. Urban Code Deploy automates the application delivery pipeline, while providing the audit trails, versioning and approvals needed.

Lastly, the pipeline orchestrator is where all the automation happens. The pipeline orchestrator provides connectors to the version control, the build systems, and the packaging and deployment manager. Its goal is to remove manual and repetitive tasks as much as possible. Jenkins is currently considered to be one of the standard orchestrators in distributed software development. Thanks to a z/OS agent, Jenkins can also orchestrate some actions on the mainframe.

In addition to the key building blocks described above, other technologies like Zowe or Ansible can be part of customer DevOps toolchain.

- Zowe is a new open source software framework that provides solutions that allow development and operations teams to securely manage, control, script and develop on the mainframe like any other cloud platform. Zowe is the first open source project based on z/OS.
- IBM Z Ansible content helps enable development and operations automation through unified workflow orchestration with configuration management, provisioning, and application deployment in one, easy-to-use platform.

As a summary, an enterprise can become more agile and shorten their application development lifecycle on z/OS by implementing a modern DevOps pipeline and toolchain. They can then provide a development experience that is familiar to all of the developers within their enterprise.

Conclusion

The mainframe continues to deliver the things that it is famous for, like security, resiliency, high performance and throughput. But organizations must adapt the way in which they use it, build applications, and manage it, making the mainframe a more agile platform for mission-critical data and applications.

The mainframe modernization journey is a long process and may require fundamental changes and the implementation of new solutions. It's also a real opportunity for the customers to integrate the mainframe across a hybrid multi-cloud environment and to incorporate it into projects running modern workloads using technologies like Machine Learning and other AI techniques.

By breaking up this modernization journey into different initiatives such as the modernization of the infrastructure, the application, the data access and the DevOps toolchain is clearly the most realistic approach.

Authors

This paper was produced by a team of specialists from the IBM Garage™ working at IBM Montpellier, France.

Guillaume Arnould joined IBM in 1996 and started in the IBM z System Manufacturing Test Engineering before spending two years in Poughkeepsie, NY. In 2001, he joined the IBM Client Center in Montpellier, to work as a Performance Expert on DB2 for z/OS client benchmarks. After 10 years as the Technical Team Leader in the Smarter Banking Showcase, leading new developments and production platforms as well as engaging with customers, Guillaume is now Advanced Technical Sales Expert and leads the Analytics on z Solutions team. He is working on solutions such as IBM Db2 Analytics Accelerator, Watson Machine Learning on Z, Db2 for z/OS, IBM Data Gate and Data Virtualization Manager for z/OS.

Guillaume Hoareau Hoareau is an IBM Certified Expert Integration Architect. He joined IBM in 2006. He is currently the CTO IBM Systems Lab Services for IBM Z and LinuxONE, covering EMEA as a pre-sales consultant. Guillaume spent the last 15 years to encrypt what can be on IBM Z and LinuxONE through leading edge client security projects. His areas of expertise cover security for both IBM Z and LinuxONE Systems and including: Data Privacy Passports, Hyper Protect Virtual Server on-premises, Pervasive Encryption, SIEM, Security Dashboards, Linux Containers, FHE, PQC, general security and compliance. Guillaume is also active from the innovation side, has he is a senior inventor, and demo maker to promote the latest technologies available on security and privacy.

Herve Sabrie is an upline manager and he leads the mainframe activities as part of the IBM Garage for Systems in EMEA. The IBM Z and IBM LinuxONE team is at the cutting edge of current and next generation technology focusing on Mainframe Application Modernization, Artificial Intelligence, Hybrid Cloud, Security, Resiliency or Blockchain. Clients and partners using IBM Garage for Systems are in a “show me, don't tell me” culture leveraging our deep technical experts. Hervé joined IBM several years ago and got multiple positions within Supply Chain, Services and Systems organizations. He is certified as an IBM Executive Project Manager.

Sébastien LLaurency is an IBM Certified Expert Integration Architect.He joined IBM in 1999. He is currently working at the New Technology Center, on Linux and Cloud Computing engagements on IBM Z and LinuxONE Systems supporting customers Europe. He is based in Montpellier, France. He made several publications, obtained a number of IBM products certifications and got 3 patents filed, among which one granted. His areas of expertise cover IBM Z & LinuxONE Systems, virtualization, linux, open-stack/shift, containers, devops, performance, security and cognitive technologies for business innovation. He is actively engaged on hybrid cloud computing architectures projects with several customers from different industries. He is now leading the Global Technical Council for WW IBM Garage for Systems.

Yann Kindelberger Yann is a lead Mainframe architect in the IBM Garage for Systems located in Montpellier, south of France. As an executive integration Architect, he is in charge of designing advanced IT solutions to meet customer challenges, mainly in the retail banking and payment industries. Over the last 7 years, Yann has been involved in several pre-sales and post sales projects around legacy application modernization leveraging API Management, hybrid Cloud and DevOps on IBM Z.

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an IBM Redbooks® residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Stay connected to IBM Redbooks

- Find us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

Notices

This information was developed for products and services offered in the US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, MD-NC119, Armonk, NY 10504-1785, US

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to actual people or business enterprises is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at “Copyright and trademark information” at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks or registered trademarks of International Business Machines Corporation, and might also be trademarks or registered trademarks in other countries.

CICS®	IBM Z®	z Systems®
Db2®	IBM z Systems®	z/OS®
DB2®	IBM z14®	z/VM®
IBM®	Redbooks®	z15™
IBM Cloud®	Redbooks (logo) ®	zEnterprise®
IBM Garage™	UrbanCode®	
IBM Watson®	WebSphere®	

The following terms are trademarks of other companies:

The registered trademark Linux® is used pursuant to a sublicense from the Linux Foundation, the exclusive licensee of Linus Torvalds, owner of the mark on a worldwide basis.

Zowe, are trademarks of the Linux Foundation.

Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Ansible, OpenShift, Red Hat, are trademarks or registered trademarks of Red Hat, Inc. or its subsidiaries in the United States and other countries.

Other company, product, or service names may be trademarks or service marks of others.

REDP-5639-00

ISBN 0738459763

Printed in U.S.A.