CLABBY ANALYTICS

Research Report

The Changing Mainframe Personality: The New IBM z14

Executive Summary

In July, 2017, IBM introduced its newest generation mainframe, the z14. As could be expected, the list of hardware enhancements and software improvements is both extensive and impressive. But what really caught our attention were the improvements that IBM has made in five areas:

1. ***IBM's z14 has the ability to protect ALL data within an organization using a process called "pervasive encryption".*** This process creates a new barrier to data breaches by ensuring that only those authorized to view data are allowed to do so by using their encryption keys to decrypt data. Enterprises must continue to protect their network endpoints; but pervasive encryption creates a new defensive perimeter at the *data* level;
2. New "container-based" pricing for digital workloads. ***IBM has finally recognized that it must get more aggressive in terms of pricing "digital workloads" on the mainframe*** (new workloads that create value, growth and competitive advantage such as cloud micro-services). Competitors Google, Amazon Web Services and Microsoft Azure have been having a field day undercutting mainframe digital workload pricing. IBM's mainframe organization is now responding – and responding very aggressively.
3. ***IBM has greatly improved analytics and machine learning performance on the mainframe by greatly expanding memory*** (32 TB of main memory). z14 customers can exploit much larger in-memory computing to achieve consistently rapid response times.
4. ***IBM has modernized the z programming environment using various open architectures to integrate application program interfaces (APIs); to simplify systems management; and to dramatically improve application response times without application modifications***. Using new development features, service build time can be cut by 90% using secure APIs and advanced DevOps. New features allow services from public and private clouds to be easily connected with transactions and data on IBM Z. Thousands of open-source software packages can be exploited to create a partner ecosystem and build new applications.
5. ***IBM has fine tuned z architecture to maximize performance***. Expanded memory, a reduction in latency by a factor of 10X using zHyperLink, new instructions and other improvements all serve to speed z14 processing – setting z14 up to be a very powerful, highly scalable Blockchain secure server environment. As we stated in this report, Blockchain transaction processing is a natural for z architecture. With z14, this story has gotten even better.

All in all, IBM has taken some aggressive steps with its new z14 to competitively differentiate it from Intel cloud vendors; to open new markets; and to improve performance for its existing base. The remainder of this *Research Report* examines the new z14 architecture – and provides additional insights on pervasive encryption; on the new container-based pricing plan; on the evolving role of analytics and machine learning on z; on IBM's z modernization efforts and on Blockchain/Hyper-ledger advancements on z.

The New z14: The Architecture

When examining z architecture, we look for improvements at the CPU level, improvements in system design and throughput, and system software improvements.

> ***The big news in z14 chip design can be found in security processing. The new chip provides a four times increase in silicon area allocated to cryptographic operations. It delivers 4 to 7 times faster encryption of data than the previous generation z13.***
>
> ***Encryption on the z14 is 18 times (YES – 18X!) faster than the competition in real world scenarios. And remember, IBM offers special Crypto Express cards (current version: 6S) that further improve encryption/decryption performance.***

As for processor performance, the IBM Z still offers the fastest microprocessor performance in the industry, now at 5.2 GHz versus 5.0 on the z13. By comparison, Intel's 64-bit processors generally operate in the 2-3 GHz range – see this list). The CPU can operate in a single threaded mode, enabling it to "stack" work and rapidly execute workloads; or it can operate in a dual-thread 2-way SMT mode (this 2-way mode was introduced on the z13). Stacking refers to running multiple virtualized workloads in an OS image in cache where it can be executed more quickly than data coming in from memory. The z14 chip runs 6.1 billion transistors as compared with 3.99 billion on the z13. And there can be up to 10 active cores (physical units) per chip.

For those not familiar with cache, cache is a level of memory placed on or close to a central processing unit that stores copies of data found in frequently used locations in main memory. The z14 offers 4 levels of cache (L1 offers 96KB for instructions and 128KB for data; L2 offers 2MB for instructions and 4 MB for data; L3 offers 128 MB of store-in shared cache; and 672MB of L4 cache can also be found on chip). *The z14 has 4X more on-chip cache than typical x86 processors – and x86 processors do not have access to L4 cache*. In fairness, x86 processors offer far less cache – but use a lot of processor cores to execute threads. As an example, the 18 core E5 2699 V3 chip can access 32KB of L1 cache per core (18 X 32KB); 32KB of L2 cache per core (18 X 32KB); and 45MB of L3 shared cache. Notice, however, the differences in the cache size of each architecture. IBM measurements are mostly in megabytes; Intel measurements are – mostly in kilobytes.

With the z14, IBM has added new instructions for single instruction/multiple data (SIMD) – and has improved instruction execution bandwidth through pipeline optimization; through TLB enhancements at the translator level; and with new instructions for old and new workloads such as vector BCD arithmetic to speed COBOL processing.

At the I/O bus level, the z14 offers one InfiniBand I/O bus and two PCIe I/O busses. IBM's new zHyperLink serves to significantly lower latency, cutting I/O sensitive workload response time without requiring application changes. IBM's next-generation I/O channel hardware increases the speed of small data-transfer by a factor if 3X.

Pervasive Encryption

The concept of "pervasive encryption" needs to be put into context. Almost 5.5 million records are stolen per day – 230,367 per hour, 3,839 per minute. Of the nine billion records breached since 2013, only 4% were encrypted. Over a quarter of all corporations can expect a security breach within the next twenty-four months. Plans need to be put in place to handle these threats (see this report on IBM's Cyber Range for strategic advice on this matter).

Now consider a statement that we've been making for years: "the IBM Z is the most securable computer platforms in the industry." To back this statement up, consider that IBM Z offers:

- ***Security level EAL Level 5+*** – a security certification rating by Common Criteria based upon an international standard for hardware security. This rating is the highest rating of any commercially available system in the industry;
- ***FIPS 140-2, level 4 compliance*** – the highest compliance rating possible for cryptographic modules;
- ***Elliptic curve cryptography*** – an approach to securing public keys that greatly speeds security processing;
- ***Extensive Hardware cryptography*** – the z14 processor features logic that speeds security and hashing processing using CP Assist for Cryptographic Functions (CPACF) at the CPU level;
- ***Add-on Crypto Express 6S co-processor adapter cards*** that offer industry exclusive protected secure key functionality; and,
- ***Extensive software cryptographic features*** – including an extensive array of software security features which can be found in the operating system, throughout the infrastructure, in crypto libraries, and at the application layer.

It is also important to note that as part of z14 security, security keys are placed into protected memory and placed into hardware secure modules in order to prevent access and tampering. IBM does not make it possible for code to take control of secured modules, and focuses on encryption protection to prevent administrative abuse. IBM's Secure Service Containers help protect code – effectively encapsulating data into a virtual appliance, denying access even to privileged users.

> ***The benefits derived by using z14 security as compared with typical Intel server designs include large costs savings (because security can be off-loaded, reducing expensive CPU cycles); faster processing thanks to co-processor hardware accelerators; pre-certification to help deal with regulatory requirements; and special functionality to deal with industry specific functions such as certain secure keys required in banking and financial applications.***

Now, taking into consideration the fact that the z14 is the most securable platform in the industry, consider the effect that encrypting all data on this platform might have. IBM's new pervasive encryption strategy calls for *100% encryption of all data*. It calls for zero application changes. It calls for zero impact to service levels. And it has been designed to help protect clients and corporations from both internal and external threats.

Pervasive encryption is not free, however, The cost for this functionality can be found in systems overhead. At present, it is believed that most organizations will see mid single-digit overhead impact.

Not to be overlooked in this discussion is the impact that pervasive encryption will have on the costly and complex process of security audits. The potential exists for hundreds of thousands of dollars in security audit savings as regulators no longer will have to check what data is being protected and where it is located. This happenstance alone should help drive a lot of business IBM's way as European companies, and US businesses attempt to deal with EU privacy protection regulations as part of the General Data Protection Regulation (GDPR) compliance standards.

The New Container-based Pricing Model

As part of its z14 announcement, IBM announced a new pricing model that enables information technology executives to deploy modern, cloud-oriented applications within collocated containers on IBM Z. This "container-based" pricing model utilizes a workload definition in the z/OS Workload Manager which tracks processes separately, allowing workloads to be billed individually. By doing this, IBM can separate traditional workloads (transaction processing, custom, etc.) from new, modern cloud-oriented workloads and charge accordingly.

> ***IBM's new pricing model for digital workloads, dubbed "Container Pricing for IBM Z" simplifies pricing for new, qualifying workloads; offers flexible deployment options that are not driven by pricing; allow for price competitiveness versus other vendor's solutions; and enable different pricing metrics within the same logical partition.***

Containers can run all sorts of workloads that provide all sorts of services. For instance, a container could run a service that could provide bank customers with voice access to their account balance – as well as link to a service that could enable that banking customer to transfer money. There are hundreds of thousands of such "digital services" in the market today – and most of these "digital workloads" are not running within mainframe cloud environments (these services are being deployed in public clouds such as Amazon Web Services, Microsoft Azure, Google and other clouds). With the announcement of its new container pricing model, IBM hopes to change this.

It is also important to note that this pricing action is radically different from IBM's z System pricing actions of the past. Previous actions have involved technologies (such as discounts for deploying on Integrated Facility for Linux, on System z Integrated Information Processors [zIIP], and with packaged 'Solution Editions'); or they have involved specific pricing methods such as zCAP (capacity management), special mobile pricing, etc.. This new pricing model is broader in nature and includes hardware, software and workload foci.

New Workloads

To draw digital workloads to the mainframe, IBM is focusing on "modernizing" its application mainframe development and test environments. IBM's dev/test effort is focused on building a rich suite of modern tools that support a myriad of interfaces and that expose and exploit a wide range of application program interfaces – priced aggressively in order to compete with the above mentioned cloud vendors.

> ***The company's new "dev/test" environment encourages the development of new digital workloads. The company has put together aggressive, highly competitive pricing to encourage the development and deployment of new digital applications. And with its "payments pricing solutions" approach, the company is willing to negotiate prices, using various business metrics tied to payment volumes.***

It is important to note that the emphasis of the new digital workloads pricing initiative is not just on new generation, web-oriented applications and services. While, traditional enterprise applications will still operate under IBM's Monthly License Charge (MLC) and International Product License Agreement (ILPA) guidelines, the new container pricing can apply to new Cobol and PL/I workloads just like it can against Java and Java Script workloads.

Why Is IBM's Pricing Action Big News? First and foremost, it indicates that IBM's mainframe organization is thinking "outside the box" (literally) – looking at the major digital workload/services trend and asking "why aren't we participating in this marketplace?" IBM Z can run hundreds of thousands of containers extremely cost effectively – so why is the market not moving to z? Clearly, the conclusion that the IBM Z organization came to was that the right tools and interfaces were not in place to foster digital workload development on z – nor was the right pricing model in place. By fixing both of these, IBM can position IBM Z as modern, very cost competitive alternatives to the giant x86 server farms of cloud competitors.

To build out its new pricing model, IBM has started by focusing on three areas: 1) application development and test (DevTest) solutions; 2) new application solutions; and 3) payments pricing solutions. A closer look at each effort shows:

1. In DevTest, IBM will offer aggressive pricing for z/OS-based development and test workloads. Much effort has been spent to expand the tools and interfaces needed to build the wide range of new applications the company hopes to draw to IBM Z. Plus IBM is being flexible with DevTest, allowing customers to increase their DevTest capacity up to 3 times at no additional MLC cost (based on workload size). And DevTest tools can be added at "uniquely discounted prices."

2. In the new application solutions area, IBM will allow its customers to add new z/OS workloads and pay a highly competitive stand-alone price. The prices of existing workloads will not be impacted. IBM customers determine the size of their containers – and that sets in place the capacity-priced billing metric for the new application solution. For Reference-based software, licensing is only to the container size.

3. In the "payments pricing solutions" area, customers are given options for payment (on premise, Payments-as-a-Service, or other options). The new payment terms can apply to both software as well as hardware components. Usage will be measured by IBM's Financial Transaction Manager. And traditional software components (such as z/OS, MQ, DB2, ODM, Cognos BI) that are needed to run the new digital workloads will be offered under the traditional MLC payment arrangement – but discounted on a per payment basis. *The bottom line after all of these considerations is this: the new payments pricing solution provides predictable and consistent pricing regardless of peaks, and it has no impact to pricing of other workloads on a given IBM Z.*

Analytics and Machine Learning

In January, 2015, *Clabby Analytics* published a report entitled "The Case for Running Real-Time Transactional Analytics on IBM IBM Z" which opened with the following discussion:

> *"Most enterprises deploy their IBM IBM Z as high-volume transaction servers. But, when it comes to analytics processing, these same enterprises usually transfer their mainframe transactional data to other, distributed server platforms for analysis. This practice is known as the extract, transform and load (ETL) process. It is highly inefficient and extremely costly – and it has got to stop!"*

In this follow-on report, we argued that:

> *"In days gone by, the mainframe was essentially quarantined, locked into a position as a transaction processing engine. But with very significant improvements to mainframe hardware, with improvements to mainframe systems software, and with additional support coming from the independent software vendor (ISV) community – the mainframe has now become a powerful transaction and analytics serving environment that can deliver results faster than waiting for external servers. Accordingly, the need to move data away from the mainframe to other servers has dissipated. The ETL problem has now been solved."*

In short, the mainframe is already a powerful analytics server. And it makes a lot of sense to process data where it is – and the vast preponderance of enterprise-critical data can be found on mainframes.

With the arrival of z14, IBM's analytics/machine learning story just keeps getting better. With its greatly expanded memory, z14 customers can exploit much larger in-memory computing (32 TB of main memory) to achieve consistently rapid response times. Predictive behavioral models enable z14 customers to analyze and monitor large data sets and learn from that data. Analytics and machine learning on the z14 not only deliver insights more quickly – they are being used to greatly simplify systems management (see this report, this report and this report).

Machine learning techniques are also being used in infrastructure management, operational analytics, application performance management and security/fraud analysis. Often called "predict-tive analytics", machine "learners" (algorithms) are being used to examine machine data looking for the root causes of problems or suspicious behaviors – and are being used to trace application activity in order to troubleshoot application behavior as well as to tune application performance. In infrastructure management, some of the best examples of machine learning programs that we have found include IBM's zAware (a program that takes a snapshot of mainframe behavior when a mainframe is running in an optimized fashion, then compares machine data to that snapshot should performance degrade –isolating anomalies); IBM's Workload Automation product suite (report found here); CA Workload Automation iDash; IBM's Operational Analytics Log Analysis; and Virtual Instrument's infrastructure management portfolio. Each company simplifies management by using predictive analytics to improve workflow and tune infrastructure performance.

For those who still argue that mainframe management skills are hard to find, we suggest that new generation predictive analytics and other management tools should be examined. Over the past few years mainframe management has gotten significantly easier.

Modernization

For those who claim that mainframes are "old technology", we respond "balderdash!" Not only is IBM's mainframe hardware far more advanced than most Intel servers, it is also far more secure. Perhaps, however, those who argue that mainframes are old technology are referring to the mainframe software development environment. And, as part of the z14 announcement, IBM announced some major changes in mainframe software programming. First, consider the use of APIs. As part of its announcement, IBM showed how easy it is to extend and connect applications with one another using application program interfaces. Developers can quickly discover existing programs which are candidates to be APIs, and can make changes impacting those APIs. Developers can securely expose services as APIs at scale to provide access and personalized experiences across multiple channels. Developers can easily connect to external services such as IBM Blockchain High Security Business Network, Watson, Twitter, the Weather Channel and more

– giving IT operations staff can get an end-to-end enterprise view by integrating IBM Z data into operational analytic ecosystems like Splunk or the Elastic Stack.

Architects can choose a language of your choice to build microservices with over 20+ languages and database technologies. Developers can understand applications by visualizing them to shorten change efforts by 30%**.** Developers can leverage 17 (and counting) tools integrated into a single framework and extensible with plugins from an active community. Operations teams can reduce deployment times by 99% through the use of automation with both native IBM z and open source tools. Application owners can quickly identify performance bottlenecks of enterprise wide applications, including IBM z components, with application performance management tools like AppDynamics.

> ***For those who continue to argue that mainframes are old technology, the hardware and the new development environment prove otherwise. In fact, the skills being used to develop micro services and to integrate APIs are the exact same skills on distributed systems as they are on mainframes.***

Blockchain and Hyperledger

Blockchain is a transaction processing scheme originally designed to process cryptocurrency. The way it works is that maintains a list of records (blocks) that are linked (chained) to one another to create a transaction flow. Blocks in the transaction are timestamped, and can be viewed by all the intermediaries involved in the transaction. By linking and time-stamping these blocks a transaction can be initiated, recorded, viewed and approved – with money or value being exchanged at the end of the transactionData in blocks cannot be modified – and transacting parties have a fully transparent view of the progress of the transaction. When mutual consensus (using pre-arranged contracts) is reached by all parties, and all deliverables along the flow of the transaction have been met, the transaction is concluded. Money is exchanged at the completion of the transaction.

Hyperledger is an open source "umbrella project" that seeks to drive the runtime and augments the Blockchain ledger capabilities with a series of related open source tools. Hyperledger blockchains are being designed to serve financial, technical and supply chain markets – each market with its own consensus, storage and service models. Typically, these open source blockchains provide identity services, access control and contract agreement/administration facilities. Hyperledger is an open source project that seeks to enrich Blockchain with the kinds of tools and facilities needed to serve enterprise-class customers while further streamlining process flow

From our perspective, there are only two implementation choices for Blockchain/Hyperledger: 1) host them; or, 2) turn to a service provider. If your enterprise chooses to host Blockchain/-Hyperledger services, the biggest challenge it will face is security. Your other challenges will involve operations and performance tuning to handle large volume transaction processing, and improving governance. If your enterprises chooses the service provider route, it will need to ensure that your service provider can deal satisfactorily with security, performance and governance challenges.

For enterprises that choose to host their Blockchain environment, we strongly recommend deploying Blockchain solutions on highly secure mainframe servers. We also suggest that these enterprises evaluate IBM's Blockchain on Bluemix – an environment that helps developers build a blotchy network in hours instead of weeks; enables Blockchain to be run in a highly secure

environment (the mainframe); and that enables easy management of decentralized networks of members.

For enterprises that go the service provider route, we strongly recommend that the service provider use secure mainframes as back-end servers – or that IBM's Blockchain-as-a-Service offering (IBM Blockchain High Security Business Network) be evaluated. This service is hosted on secure mainframes… As for why, consider the security advantages articulated for z14 earlier in this Research Report – as well as the fact that Blockchain on IBM Z is 2.3 times faster than its "industry standard" Intel competitors; and that IBM has started to integrate its Watson cognitive computing technology with Blockchain and Hyperledger technologies – making it possible to automate and streamline certain transaction handling processes.

Summary Observations

There are three traditional obstacles to mainframe adoption: 1) the claim that mainframes are old technology; 2) the claim that mainframe skills are difficult to find; and, 3) the claim that mainframe costs are too high. x86-biased IT executives will cling to these antiquated objections – despite new fact that show otherwise.

For those willing to examine the facts, you will find that:

- The mainframe is a technological marvel especially when it comes to security. And IBM's new pervasive encryption further broadens the security gulf between mainframes and x86 architecture. Mainframes are clearly not old technology;
- With z14, IBM has modernized the z System programming environment – offering new tools and simpler ways to interface APIs to programs new and old (including COBOL programs). These new tools operate in much the same way that their distributed computing counterparts operate – and this helps close the skills gap;
- The mainframe management skills gap is being aggressively closed using analytics, machine learning and cognitive tools. IBM Z are starting to manage themselves – again addressing the alleged shortage of mainframe skills;
- IBM executive management is taking aggressive steps to lower mainframe costs and bring them more in line with cloud competitors such as Amazon Web Services and Microsoft Azure. This pricing throws the "mainframes cost too much" argument out the window; and,
- Blockchain runs faster and more securely on mainframes. This argument is quantifiably provable.

In recent years the mainframe has undergone some significant personality changes. It remains the traditional transaction processing powerhouse – but now, with the introduction of the z14, the mainframe is a potent analytics server, an enhanced machine learning environment, and the most securable Blockchain environment on the planet. Further, IBM has modernized the mainframe program environment – and has finally adjusted its pricing for new digital workloads. For those who contend mainframe are old, pricey and hard to staff – look again.

Clabby Analytics
http://www.clabbyanalytics.com
Telephone: 001 (207) 239-1211