The benefits of colocating workloads on IBM Z®

Colocation refers to the presentation, business logic, and data serving layers of a multitier workload that is deployed onto a single physical server.

Key benefits

- Reductions in latency and improvements in throughput
- Improved security and availability
- Cost savings

Does colocation matter?

Colocation can make a significant difference. Businesses and IT organizations must provide fast access to data across the web, application, and database tiers that comprise most enterprise workloads.

The physical location or proximity of the workloads can make a difference when these multitiered workloads have communication patterns that are network intensive. This means, they either frequently communicate, exchange many messages to complete a single transaction, or exchange large amounts of data. For example, big data type solutions such as analytics related workloads.

IBM Z technologies supporting co-location

IBM HiperSockets™ and Shared Memory Communication (SMC) are the IBM Z networking technologies that provide differentiated benefit to customers who colocate workloads on the platform.

HiperSockets provide a very efficient memory to memory transfer of standard packets without requiring physical networking hardware.

SMC eliminates all packets along with all the TCP/IP protocol and packet-related processing, thus providing significant savings in host network processing.

Benefits of colocation on IBM Z

Reductions in latency and improvements in throughput

Workloads that access data on IBM Z can do so at memory speeds with the lowest possible latency, bypassing network traffic and delays. Network latency can affect the overall throughput capabilities for a workload.

The result of an IBM internal study demonstrates the reduced latency:

An OLTP workload on Red Hat® OpenShift® Container Platform 4.4 that is colocated to a database on an IBM z15™ T01 using a HiperSockets connection, runs with up to 4.7x lower latency than the compared x86 platform, which uses a 10 Gb TCP/IP connection to the same database.[1]

Improved security and availability

External components are vulnerable to physical attacks. Because HiperSockets and SMC do not have any physical interfaces, these technologies are protected against any attacks from the outside. And since there are no network switches, routers, adapters, or wires that can break or that have to be maintained, the availability is improved.

Cost savings

Colocated workloads run on one server. Therefore, hardware boxes can be eliminated, and floor space and power consumption can be reduced. In addition, the central processing unit (CPU) overhead for the networking software stack is significantly reduced. With HiperSockets and SMC, there are zero external components or cables to pay for, to replace, to maintain, or to wear out. Also, the complexity that is associated with procuring and managing physical network equipment is reduced and simplified.

Bottom line

Colocating workloads running on IBM z/OS®, Linux, IBM z/VSE®, or IBM z/TPF workloads on a physical IBM Z server, benefits not only from great performance and operational efficiency, but also leverages investments in existing assets.

Cloud-native applications can be located close to existing workloads to improve throughput and reduce latency, empowering organizations to integrate and modernize without disrupting current services as they take their cloud-native journey.

[1] *IBM internal study designed to replicate banking OLTP workload usage in the marketplace deployed on Red Hat OpenShift Container Platform (OCP) 4.4.12 on IBM z15 T01 using IBM z/VM® versus on compared x86 platform using KVM accessing the same PostgreSQL 12 database running in a z15 T01 LPAR. 3 OLTP workload instances were run in parallel driven remotely from JMeter 5.2.1 with 16 parallel threads. Results may vary. z15 T01 configuration: The PostgreSQL database ran in a LPAR with 12 dedicated IFLs, 256 GB memory, 1TB FlashSystem 900 storage, Red Hat Enterprise Linux 7.7 (SMT mode). The OCP control plane and compute nodes ran on z/VM 7.1 in a LPAR with 30 dedicated IFLs, 448 GB memory, DASD storage, and HiperSockets connection to the PostgreSQL LPAR. x86 configuration: The OCP control plane and compute nodes ran on KVM on Red Hat Enterprise Linux 8.2 on 30 Skylake Intel® Xeon® Gold CPU @ 2.30GHz with Hyperthreading turned on, 448 GB memory, RAID5 local SSD storage, and 10Gbit Ethernet connection to the PostgreSQL LPAR.*