Don't let your smart software suffer from poor platform choices

Mismatched AI and analytics can be massively inefficient

POWER PERSPECTIVE

FREEFORM DYNAMICS

IBM

If you and your organisation work with any of Artificial Intelligence (AI), Machine Learning (ML) and Advanced/Real Time Analytics, you know two things: first, they are all becoming much more important. And second, such workloads can have very different resource and security requirements at various stages of their life cycles. The question then is, do you need multiple different - and underutilised - systems to run different workloads, or could you depend on a single intelligent, optimised platform?

How high-scale, vertically integrated systems can help

Key system characteristic	Why this matters
Built using powerful CPUs with faster cores than standard server architectures, and with higher interconnect bandwidth between memory, I/O and storage while also requiring less floor space.	A system that delivers high performance in a small footprint reduces the chance of memory, processor and I/O bottlenecks, and therefore maximises the ability to run multiple demanding workloads.
Operating system and virtualisation hypervisor designed with the ability to logically partition the environment to deliver different service capabilities with minimal underutilisation of resources.	The elements of each AI/ML workload have individual needs that vary as models are created, applied and updated. Granular optimisation precisely targets and reallocates system resources as required.
A design able to support multiple different processor, storage and networking options in order to provide a range of system capabilities and meet the needs of diverse workload types.	The ability to include multiple processing units of various types, all tightly coupled via high-speed interconnections and fast memory, allows both learning and analysis to be run optimally.
Encryption through all elements of the system from the motherboard, CPU, networking elements and storage into memory and operations. Essentially, data encrypted at rest, in transit and in use.	High end-to-end security allows workloads to process and analyse large volumes of data, while still minimising the risk of accidental, or deliberate, disclosure of sensitive information.

Other things to think about

Organisations are making greater use of analytics as they seek to get more value from the expanding data resources they create and hold. The challenge is that analytics workloads vary hugely in terms of the mathematics they use, and this can require different IT resources to optimise each workload. One size doesn't fit all. Servicing these differing requirements can be achieved using x86 systems, but it is often the case that each system is optimised for particular types of analytics and AI/ML tasks, and can therefore only run a certain range of workloads. An alternative is to use a single highly-integrated system designed to cater efficiently and effectively for a broad range of high performance workloads.

Real-world solution example: IBM Power

In order to illustrate how high-scale, vertically-integrated systems can work in real-life scenarios, we will use a real-world example of a system which has been architected to suit a wide range of complex analytical workloads: the Power® platform from IBM®, the sponsor of this paper. Note that while nothing we say here should be taken as endorsing or recommending any particular product or service, it can be very useful to see how a specific example translates some of the key principles involved in running advanced analytics, say, into reality.

Flexibility is a key requirement here. As AI and analytics expand into more areas of business, organisations must run multiple disparate workloads, each of which has different platform or infrastructure needs. Some of these workloads must run frequently or continuously, while others come and go as business needs change over time.

A platform such as IBM® Power® can meet this complex requirement in a particularly interesting way: by incorporating a variety of different processor engines, and then logically and dynamically allocating those physical resources to different tasks using management software. Power has also been designed to minimise processing and memory interconnect bottlenecks in the backplane, allowing optimal performance at minimal energy consumption.

Some of these workloads may also process sensitive data, making the end-to-end encryption capabilities of the platform attractive as they ensure that information is protected at rest, in transit and in compute, considerably reducing the attack surface.

A final matter to bear in mind is that, given that AI/ML usage will continue to expand, using a platform like IBM Power that can be performance-upgraded with new, certified, components could also allow you to repurpose older kit to fit new and evolving business requirements.

About the Power Perspective series

This document is one of a series of similar pieces looking at how high-scale, vertically integrated systems can provide tangible business benefits in context for a range of different themes. Other Power Perspectives include:

Supporting the needs of highly unpredictable workloads
It's all about flexibility, scalability and cost effectiveness

Getting real about IT sustainability
From good intentions to tangible results through smart systems selection

Have you made the right platform choices to minimise risk?
Critical systems have to run and run

S/4HANA without compromise
A modern ERP architecture needs a powerful and future-proof platform

About Freeform Dynamics

Freeform Dynamics is an IT industry analyst firm. Through our research and insights, we help busy IT and business professionals get up to speed on the latest technology developments and make better-informed decisions.

For more information, visit www.freeformdynamics.com.

About IBM

Discover the next generation of IBM Power servers built to meet today's challenges with new levels of performance, core-to-cloud data protection, and streamlined automation and insights.

For more information, visit www.ibm.com/it-infrastructure/power.