Optimising compute platforms for AI success

A strategic guide to aligning infrastructure with AI demands

vodafone business

Contents

The aim of this document is to help you understand why artificial intelligence (AI) is an emerging trend, how it could benefit your business, and how Vodafone can help you optimise your existing compute platforms for AI. It covers:

01
Powering AI potential

Artificial Intelligence (AI) presents unprecedented opportunities for businesses, from improving decision-making and automation to driving innovation and customer engagement. However, many organisations struggle to make the most out of AI, due to inefficient, unstructured, or outdated IT environments. To harness AI's full potential, businesses first need to organise and restructure their compute platforms to optimise both performance and cost efficiency.

This guide explores the importance of aligning IT infrastructure with AI workloads and the benefits it brings. This explainer provides a comprehensive, step-by-step guide to help organisations like yours seamlessly execute this transformation with Vodafone Business.

02

Overcoming infrastructure barriers to AI adoption

It's important to know that AI workloads demand significant computational power, storage space, and efficient data flow. Without a properly structured and optimised compute environment, your company could face several challenges, including:

Poor performance

AI models need powerful and scalable compute resources to function efficiently. Using outdated infrastructure can lead to slow processing times and bottlenecks.

High costs

Inefficient allocation of compute resources can result in unnecessary expenses, increasing the total cost of ownership (TCO) for AI deployments.

Scalability challenges

AI requires flexible infrastructure that can easily adjust to varying workload demands.

Security and compliance risks

It's crucial to process sensitive AI-driven data in a secure environment that meets industry regulations.

Inefficient data management

Achieving AI success depends on seamless data integration, processing, and storage capabilities, which require an optimised compute backbone.

Any one of these challenges is enough to knock your business's AI journey off course. However, by restructuring your compute environments, your organisation can eliminate inefficiencies, enhance AI performance, reduce costs, and accelerate time to value – unlocking the full potential of AI.

03 Business benefits of optimising your compute platforms for AI

Optimising compute platforms for AI isn't just about avoiding obstacles on your business's journey. It also offers numerous advantages, regardless of the industry your organisation is in, enabling you to discover the full potential of AI-driven applications. Below are the key benefits of optimised compute platforms for AI, with specific examples of AI applications across different sectors:

1. Enhanced AI model performance

For AI models to perform at their best, they need significant processing power. With fully optimised infrastructure, you can enjoy benefits such as:

- Faster training and inference times for AI models, leading to quicker insights.
- Reduced latency in real-time AI applications such as fraud detection in banking.
- Seamless deployment of AI-powered virtual assistants and chatbots in customer service.

Financial institutions using AI for fraud detection can process millions of transactions in real time, identifying anomalies faster and preventing fraudulent activities before they escalate.

2. Cost optimisation

Rightsizing compute resources ensures that your organisation does not over-provision or underutilise your infrastructure, and helps you achieve:

- Reduced cloud expenditure through use of reserved instances or spot pricing for AI workloads.
- Lower operational costs through use of serverless computing for on-demand AI processing.
- Efficient resource allocation that prevents unnecessary power consumption and cooling costs in data centres.

Retailers leveraging AI-driven recommendation engines, for instance, can optimise their cloud costs by using scalable graphics processing unit (GPU) instances only during peak shopping seasons.

3. Improved scalability and flexibility

An AI-ready compute platform allows your business to dynamically scale resources based on demand, enabling:

- Efficient handling of seasonal or event-driven AI workloads.
- Rapid expansion of AI applications without costly infrastructure overhauls.
- On-the-fly adjustments to computing power for AI-powered automation in logistics and supply chain management.

E-commerce platforms using AI for personalised recommendations can scale their compute power during sales events like Black Friday to manage increased traffic and AI-driven transactions.

4. Improved data management and integration

A structured compute environment streamlines data workflows, which is critical for AI applications such as:

- Predictive maintenance in manufacturing, where sensor data needs to be processed and analysed in real time.
- AI-powered analytics in healthcare, ensuring patient data is managed securely and processed efficiently to support diagnostic AI models.
- Real-time AI processing in smart cities, where traffic patterns and environmental data are analysed to improve urban planning.

In healthcare settings, high-performance computing can support hospitals and research institutions in processing large medical imaging datasets (such as X-rays, MRIs, and CT scans), enabling efficient data handling and workflow management while maintaining data security.

5. Increased security and compliance

AI models often process sensitive data, making security a top priority. An optimised compute platform helps ensure:

- Secure environments for AI training and inference, reducing risks of data breaches.
- Compliance with regulations such as GDPR and HIPAA, ensuring AI applications handle personal data responsibly.
- Robust identity and access management (IAM) to prevent unauthorised use of AI-powered analytics tools.

Financial services firms using AI for automated risk assessments must ensure their compute platforms adhere to stringent security and compliance standards to protect customer data.

6. Faster AI deployment and innovation

An AI-optimised infrastructure accelerates the development cycle, making it easier to:

- Train and fine-tune AI models in a fraction of the time.
- Deploy AI applications with minimal friction in production environments.
- Innovate with AI-powered solutions across different market segments, such as retail, healthcare, and automotive.

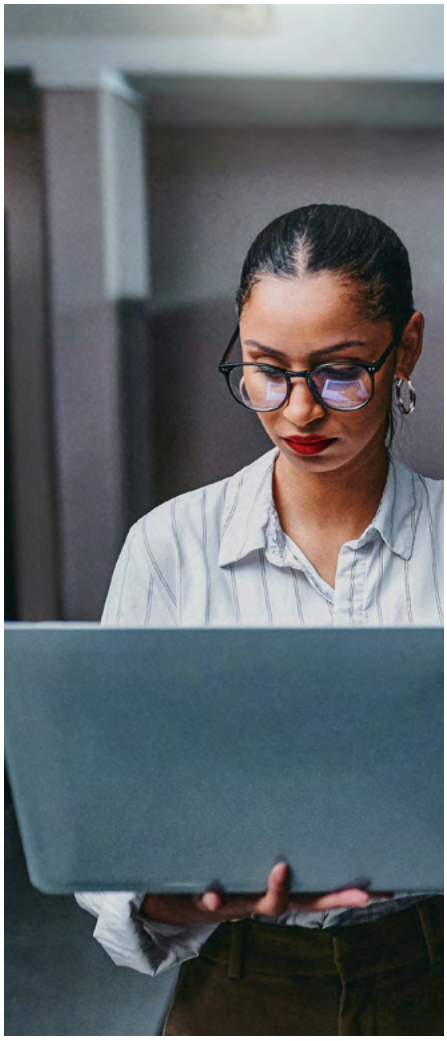
Autonomous vehicle manufacturers rely on highly optimised compute platforms to process vast amounts of sensor data in real time, improving vehicle safety and driving capabilities.

No matter what industry your business is in, investing in the right infrastructure can help you bring AI-powered solutions to the market faster. This not only allows you to stay ahead of competitors but also drives meaningful innovation.

04

Steps to reorganising and restructuring your compute platforms

Reorganising and restructuring compute platforms for AI is essential for businesses across all sectors. It's not just about overcoming challenges; it's about seizing new opportunities. At Vodafone Business, we recommend a structured approach to prepare your IT environment for AI. Read our step-by-step guide to help you through this transformation:



1. Assess current IT infrastructure and AI readiness

- Conduct an audit of your existing compute, storage, and networking resources.
- Identify bottlenecks and inefficiencies impacting AI performance.
- Evaluate your current cloud, on-premises, and hybrid architectures.
- Determine cost inefficiencies and underutilised resources.

2. Define AI workload requirements

- Analyse AI use cases and workload intensity (e.g. training vs. inference).
- Determine your compute needs (GPU, tensor processing units (TPU), central processing units) based on workload demands.
- Assess your data storage and management requirements.
- Define your security and compliance requirements.

3. Select the right compute platform

- Evaluate whether your workloads should run on public cloud, private cloud, or hybrid environments.
- Choose high-performance compute (HPC), GPU-optimised instances, or specialised AI chips for deep learning models.
- Assess cloud providers (AWS, Azure, Google Cloud) for AI-optimised services.
- Consider edge computing for latency-sensitive AI applications.

4. Optimise infrastructure for AI performance and cost

- Implement auto-scaling to dynamically adjust resources.
- Optimise workload placement for cost-effectiveness (e.g. spot instances, reserved instances, or serverless computing).
- Leverage containerisation (Docker, Kubernetes) for AI workload portability.
- Utilise AI-specific accelerators (e.g. NVIDIA A100 GPUs, Google TPU pods).

5. Streamline data workflows and storage

- Improve data ingestion, transformation, and integration processes.
- Optimise storage solutions (object storage, data lakes, high-speed SSDs) for AI efficiency.
- Implement data governance policies for security and compliance.
- Ensure real-time and batch data processing capabilities align with AI needs.

6. Implement security and compliance best practices

- Ensure AI models operate in a secure environment with proper IAM.
- Encrypt your data in transit and at rest.
- Adhere to regulatory frameworks, for example: GDPR, HIPAA, ISO 27001 for AI compliance.
- Deploy AI-specific security solutions to mitigate risks such as adversarial attacks and data poisoning.

7. Monitor, optimise, and iterate

- Continuously monitor AI workloads for performance, cost, and efficiency.
- Utilise AI-driven analytics to optimise compute resource allocation.
- Regularly assess your infrastructure needs as AI use cases evolve.
- Establish a feedback loop to refine AI-driven operations over time.

05 The importance of connectivity for powering AI

In an AI-driven landscape, robust connectivity is the vital foundation for moving data swiftly and reliably between compute resources – essential for experiencing the full potential of advanced analytics and machine learning. Seamless connectivity is a non-negotiable pillar of AI optimisation, and here are some of the key reasons why:

Reduced downtime

Reliable connectivity between clouds and end-users minimises service interruptions, ensuring AI workloads and applications stay available.

Scale with cloud requirements

Scalable networking infrastructure adapts to growing AI data volumes and computing needs. It enables smooth resource expansion and flexible bandwidth allocation to keep pace with evolving business demands.

Protect data in transit

Good connectivity helps ensure visibility and secure data flow whilst moving between cloud resources – helping to safeguard sensitive information and meet regulatory requirements.

Seamless interconnectivity

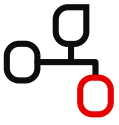
Fast connectivity enables smooth integration between cloud platforms and private datacentres. Workloads can communicate across clouds with minimal setup, using secure dedicated connectivity via express routes.

High availability and redundancy

Using multiple cloud providers alongside private infrastructure helps maintain uptime and resilience – crucial for continuity and disaster recovery.

Improved application performance

Solutions like SD-WAN dynamically direct traffic to the nearest cloud gateway, cutting latency and costs while boosting application speed and reliability.



Reduced networking complexity

Modern applications are built as microservices that run across hybrid and multi-cloud environments. Good connectivity helps manage the networking complexities.

Smooth application migration

Strong connectivity reduces the friction of moving apps across clouds, avoiding costly reconfigurations and keeping innovation on track.

Enhanced data redundancy and disaster recovery

Spreading data across clouds strengthens resilience. Reliable connectivity ensures effective data redundancy and disaster recovery.

06 How Vodafone Business can help

At Vodafone Business, we offer a variety of products and services designed to help your business harness the power of AI. Whether you're just starting out or considering how to optimise your existing infrastructure, we've got you covered. Our solutions include:

Data Location Optimisation (DLO)

DLO is an assessment we conduct to help businesses strategically place and manage their workloads. So, you can count on optimised cost and performance for your AI strategy today and tomorrow. For more information on DLO, please visit our website.

Dedicated Private Cloud (DPC)

DPC is a cloud infrastructure platform that helps businesses put AI workloads in the right environment. So, you can rely on secure, scalable infrastructure that's optimised for performance, data privacy, and compliance. With full control over compute and storage, your business can accelerate AI development and deployment. For more information on DPC, please click here.

Public Cloud

Our Public Cloud services help businesses scale their AI strategy with speed and flexibility. So, you can tap into vast compute resources, advanced AI services, and global reach – without the need for upfront infrastructure investment. It's ideal for experimenting, iterating, and deploying AI models quickly, while benefiting from continuous innovation and cost efficiency.

07
Next steps

AI is transforming every industry. But no business can reach its full potential without a well-structured and optimised compute environment.

By reorganising and aligning compute platforms with AI requirements, your organisation can achieve higher performance, lower costs, better scalability, and enhanced security. The key to success is a strategic, step-by-step approach and at Vodafone Business, we can offer you the solutions and services needed to ensure the right balance of compute power, data management, and operational efficiency, helping you to execute this transformation.

By working together with Vodafone Business, you can get the right infrastructure in place to harness the full potential of AI and drive innovation at scale.

Get started today

To discover more about optimising your compute platforms for AI success, visit our website or speak to your Vodafone account manager.

For other businesses, please contact us on the respective phone numbers below:

50 to 249 employees	0808 005 7404
250+ employees	0808 099 8877

The Nation's Network

Supporting the nation since 1984

Mobile | Fixed | Unified Comms | Security