



0 🗰 🗣 🔺

2024 Interconnection Report

The impact of **generative Al** on networks_

Contents

Executive Summary		
Part 1: The unstoppable rise of generative AI	3	
Part 2: The transformative power of generative AI for businesses	5	

Part 3: Understanding the network		
generative Al		
Part 4: Where is generative AI		
heading	15	
Appendix	20	

"66% of CTOs said their current network infrastructure did not have the capacity to fully embrace gen AI."

Executive Summary

- The generative AI revolution is advancing at an astonishing pace, and business leaders predict it will have a seismic impact on their operations, but most IT executives say their networks are unprepared for this groundbreaking technology.
- In January 2024, <u>Console Connect</u> conducted a global survey with Chief Technical Officers (CTOs) and senior IT Leaders to understand their thinking on generative AI, including both the opportunities and challenges this emerging technology has created for their networks.
- Our survey involved 1,000 CTOs and senior IT Leaders in the technology, financial services, communications, manufacturing, and retail sectors in the UK, the USA, Australia, Hong Kong, and Singapore, and provided a comprehensive perspective on the network demands associated with generative AI adoption.
- Alarmingly, over two-thirds (69%) of senior IT leaders and 66% of CTOs said their current network infrastructure did not have the capacity to embrace generative AI to its full potential.

- Of the respondents that do not currently use gen Al, 88% said they plan to deploy the technology in the near future. Furthermore, almost four in five leaders (76%) agreed the rush to adopt generative Al may have long term repercussions on their technology infrastructure planning. Additionally, 76% believed their IT teams were under increasing pressure to embrace generative Al within their organisation.
- Network security is another big concern when it comes to the generative AI adoption. Seven-in-ten (70%) feared the use of generative AI is going to put their network at risk of cyberattacks or data breaches

 a figure which rises to 90% in Australia.
 Cybersecurity risks (41%) and a lack of IT skills or expertise (39%) were also seen as major barriers to immediate adoption.
- While these findings underscore a growing awareness among industry leaders regarding the potential dangers of embracing generative AI without strategic planning, they equally draw attention to the fact that careful groundwork is essential before adopting this rapidly evolving technology into their current infrastructure.



Part

The unstoppable rise of generative AI



The dawn of generative AI has gripped the business world like no other technology since the dawn of the internet.

This transformational technology is now being rapidly adopted by almost every industry and is poised to revolutionise our jobs, our daily routines, and even how we communicate with each other.

Indeed, as we approach the first quarter of the 21st century, generative AI has the potential to unlock billions of dollars in business value and radically transform the way we work. In the process, it will drive innovation through its ability to make intelligent decisions at unprecedented speeds.

But what exactly is generative AI? At its core, generative AI is a type of artificial intelligence technology involving complex algorithms that can create new and sophisticated content such as text, images, audio, code, or other media which mimic human ability.

For example, if you asked a generative AI tool to write a four-volume book on the history of England, it could, in theory, take the entire contents of the British Library's English history department and generate a sequence of words, paragraphs, and chapters to create something that looked like – but not a reproduction of – Winston Churchill's A History of the English-Speaking Peoples. This phenomenal ability to process vast amounts of raw data to generate new content is a significant turning point in our understanding of what intelligent automated systems can do for our world.

Furthermore, as these AI systems become increasingly more sophisticated, they have quickly become an invaluable business asset in their facility to transform operations, from managing daily mundane tasks and predicting customer behaviours through to writing code and co-creating innovative prototypes.

In January 2024, Console Connect commissioned Arlington Research to conduct a global research survey about the challenges and demands placed on business infrastructure when embracing and deploying generative AI. Our survey involved 1,000 Chief Technical Officers (CTOs) and senior IT Leaders in the technology, financial services, communications, manufacturing, and retail sectors in the UK, the US, Australia, Hong Kong, and Singapore.

Awareness and understanding of the technology is relatively high, with 87% of global respondents claiming 'some knowledge' of the technology. So too is the appetite for adoption. Of the respondents who haven't already adopted AI within their organisation, 88% said that they are likely or extremely likely to deploy generative AI in the near future (see fig.1).



Figure 1: Respondents are asked the likelihood of deploying generative AI in the near future



That said, 29% of respondents who 'know a little' about generative AI said there would be no barriers to adopting its use within their organisation.

This suggests that interest seems to be focused on the usages and benefits of generative AI, rather than how to implement it effectively into their network infrastructures, not to mention the very specific connectivity requirements needed for generative AI to bloom.

However, as business leaders confront the awesome potential of generative AI technology, it is becoming apparent that many companies will be forced to rethink their network management systems.

The reason for this is simple. Generative AI requires staggering amounts of computing

power and data-crunching to perform effectively, and it demands fast and flexible network solutions that can keep up with its rapid advancements.

On this matter, 76% (see fig.2) of global CTOs and IT Leaders agreed that the rush to adopt generative AI will have long term repercussions on technology infrastructure planning for their organisation, but an equal 76% said their IT teams are being pressured to support generative AI adoption.

Therefore, for companies to truly harness the potential of generative AI, it is essential to first understand the network connectivity requirements – and the risks – before deploying this revolutionary technology safely, securely, and responsibly.



Figure 2: Respondents are asked to agree with the following statements about gen AI





The transformative power of generative AI in business networks



Several industries have already gained a first-mover advantage in the realm of generative AI by switching up their network management systems to accommodate its connectivity requirements.



Leading the charge is the banking sector, where generative AI has been widely deployed to create not just responsive chatbots but intelligent virtual assistants.

These innovative systems have streamlined customer-facing operations by providing seamless interactions with personalised and conversational responses. They can also pour through masses of customer data and then execute automated tasks in the bank's network infrastructure such as fund transfers, monthly payments, financial history tracking, and even new account onboarding – all without human assistance.

This evolution in customer service is the reason banking giant NatWest is now collaborating with Amazon Web Services to accelerate its uptake of new generative AI technology from the cloud provider.



Similarly, the rapid adoption of generative AI platforms in the retail sector has led to stronger customer relationships. These platforms sift through customer data in nearreal time to create compelling individualised recommendations based on user preferences and search behaviour patterns.

One example is multinational furniture company lkea's Kreativ app which uses generative AI to offer customers a virtual replica of their homes that can be fitted out with products for each room.



Healthcare

In the medical research sector, generative Al's ability to simulate predictive outcomes has accelerated the clinical drug trial development process as well as helped researchers to reduce the time and costs in bringing life-saving medications to patients.

Hong Kong biotechnology company Insilico Medicine, for instance, uses a revolutionary drug discovery tool called Pharma.AI that applies millions of data samples and multiple data types to identify the most promising targets for cancer and other signature diseases.



Finance

Likewise, finance and fintech companies are deploying generative AI technology across their networks in the guise of predictive trading algorithms. These tools ingest gigantic volumes of financial data and then forecast nuanced investment decisions, allowing portfolio managers to anticipate shifts in the market before they happen.

The world's largest hedge fund firm, Bridgewater Associates, even has its own Al Lab where investors and Al experts can explore cutting-edge use cases.



Examining your existing network

However, for all the benefits that generative Al offers, they will remain beyond reach if your current network infrastructure is unable to store, process, and retrieve the massive datasets which this technology thrives on.

So, before you consider which generative Al tool or platform to use, you should conduct a thorough analysis of your current network and determine whether it has the capabilities to ensure seamless and augmented Al workflows.

For instance, does your current network have the edge computing capabilities to process IoT

data and deliver real-time quality insights? Also, can you monitor and audit how your generative Al tools are interacting with your network? Another essential factor is making sure your network is secure.

Our global study found that network security is a major concern when it comes to the adoption of generative AI. 70% of respondents (see fig.3) feared the use of generative AI is going to put their network at risk of cyberattacks or data breaches – a figure which rises to 90% in Australia. Cybersecurity risks (41%) and a lack of IT skills or expertise (39%) were seen as major barriers to immediate adoption (see fig.4).



Figure 3: Respondents select how likely gen Al will put their network at risk of a cyberattack or data breach

Figure 4: The main barriers for respondents to adopting gen AI within their organisation



A lack of IT skills or expertise (39%) were also seen as major barriers to immediate adoption. It's true that most forward-thinking businesses are now operating in a hybrid or multi-cloud environment where they are pulling in data from a variety of public and private clouds.

This certainly improves operational efficiency, but it can also compromise sensitive data if the underlying network is insecure.

A third factor is the rapid progression of generative AI technology. Your network needs to be quick and agile enough to support future AI advancements, which means having the flexibility to scale or upgrade on demand to support not only your business growth, but also prevent the onslaught of cyberattacks that will inevitably follow as AI evolves. Then there is the human factor. Introducing generative AI platforms into your business operations requires training your IT team to make sure they can harness its full potential and deploy it effectively across your network.

But is this training happening? Our survey suggests otherwise. 43% of respondents (see fig.5) put training for IT staff as their top priority.

Not only that, but with generative AI being so dynamic, your team will need regular upskilling to ensure your network capabilities remain equally adaptive and dynamic.

Figure 5: Respondents rate the top priorities for their IT team in adopting gen AI

Training for IT staff Increasing the IT budget Increasing cybersecurity Increasing the use of cloud services Need for more IT staff Upgrading the network infrastructure Enhancing network connectivity Ensure it complies with regulatory requirements Moving to cloud services None of these



Businesses are still educating themselves on the infrastructure requirements of gen Al. Not only do CTOs have long-term concerns about the repurcussions of gen Al on technology infrastructure planning, but they also have more immediate concerns about the capabilities of their network. Just over twothirds of respondents (69%) agreed or strongly agreed that their network infrastructure does not have the capacity to embrace Gen AI to its full potential (see fig.6).

It's clear, then, that as generative AI moves from a desirable concept to an actionable solution, a deeper understanding is needed as to how it can perform effectively with business networks.

Figure 6: Respondents are asked if they believe their existing network infrastructure doesn't have the capacity to embrace gen AI to its full potential







Understanding the connectivity requirements of generative AI



Our global survey identifies major concerns from the IT community around data privacy and the accuracy and reliability of data (see fig.7).

Over one third of organisations said they are concerned that the adoption of gen AI will compromise data privacy and security, which was also ranked as the highest concern by 16% of respondents.

So, whether you're a small start-up company or a multinational corporation, it is beneficial to understand the following connectivity requirements before deploying generative AI safely and effectively across your network:

Data processing

One of the key connectivity provisions for training generative AI tools in cloud environments is affordable access to the scale of datasets. One example is so-called large language models (LLMs) such as OpenAI ChatGPT, Microsoft Bing Chat, and Google Bard. By their very definition, LLMs are extremely large, and training them requires vast amounts of data and hyper-fast computations.

However, the enormous processing power required to train LLMs is only one part of the jigsaw. You also need to manage the sovereignty, security, and privacy requirements of the data transiting in your public cloud. This is reflected in our survey, with over one-third saying their organisation would be concerned about the risk to data privacy and security.

Data protection

Data security and privacy are paramount network requirements when deploying generative AI technology. Simply put, if you are connecting over the public internet then you are exposed. It is no longer a question of whether a data breach will happen, but when. These attacks will certainly increase over the next few years,



Figure 7: Fears and concerns around using gen AI within the enterprise network - highest concerns versus most selected concerns



especially with hackers deploying Al-powered tactics to get to your data.

Indeed, the protection of personal data is perhaps the major concern of the generative AI revolution, and businesses will need to closely examine the security and resilience of their cloud network infrastructure and decide if they need to move to a private network environment.

Latency and network congestion

While risk to data privacy (16%) and the accuracy and reliability of responses (14%) were the highest concerns in the global survey (see fig.8), network congestion and latency were seen as less of a worry. This is worrying, and points to either a lack of awareness or misplaced confidence, or both. Latency is a critical factor in terms of interactions with your workforce and your customers. We have all become latency sensitive, especially with the volume of voice and video calls that we experience daily, but the massive datasets used for training AI models can lead to serious latency issues on the public cloud.

Network congestion, meanwhile, could impact your ability to build models on time. If you have significant congestion in getting your fresh data into your LLMs it's going to start to backlog, and you won't be able to achieve the learning outcomes that you are hoping for.

Figure 8: Largest fears and concerns amongst respondents for using gen AI within their networks







Regulatory frameworks

Business leaders should pay close attention to the key public policies and regulation trends which are rapidly emerging around the AI landscape. Just 28% of respondents who 'know a little' about generative AI see them as a concern (see fig.9).

In truth, the maze of regulatory frameworks globally is very complex and subject to change.

Keeping abreast of this regulatory landscape is going to be challenging to say the least, and the understanding of localisation of data within the boundaries of jurisdictions is becoming a

top priority. With the explosion of data growing exponentially, and with data sources becoming more fragmented and sitting in lots of separate places, companies will need to manage how their data is being moved around.

If their networks are not secure, then cracks will soon appear.

Data sovereignty

As the world becomes more digitally interconnected, the widespread acceptance of generative AI technology will likely create longlasting challenges around the thorny issue of data sovereignty.

Figure 9: Comparing barriers to adoption between respondents who know a lot about generative Al versus those who know little





This has already prompted some nations to define and regulate their own legislation regarding where data can be stored, and where the LLMs processing that data can be housed.

The only way your company will have assurance of maintaining your sovereign border may be to use a form of private connectivity while the data is in transit.

The same applies to AI training models on the public cloud; companies will need some type of secure connectivity from their private cloud to their public cloud where they do their AI training models, and then use that private connectivity to bring their inference models back.

Al governance

A final requirement to consider is governance. In other words, who gets access to the data and where is the traceability of the approval of that data available? Without proper AI governance, there could be high consequences for companies that may result in commercial and reputational damage.

The EU's new AI Act – the world's first comprehensive AI law – is the most significant in terms of responsible implementation. Specifically designed to address the risks of AI governance, the act outlines a series of requirements that intend to safeguard "the health, safety, and fundamental rights of EU citizens and beyond," and is expected to have a considerable impact worldwide when it is formally adopted by the European Parliament in 2024.

Businesses using or planning to use AI systems should therefore start to map their processes and assess the level of compliance with the new rules.





Where is generative AI heading?

The rise of generative AI is one of the most significant transitions in computing history and will transform our industries beyond recognition.

Given that machine learning and specifically generative AI technology is advancing at such breakneck speed, the question arises – where is all this heading?

7 emerging AI trends to watch:



Mass adoption of LLMs

As we have discussed, LLMs operate by analysing massive pools of data to generate language text. These LLM programmes are trained on millions of gigabytes of data that are pulled from the internet and then fine-tuned to interpret questions, generate responses, or analyse research.

The most well-known examples are OpenAl's ChatGPT, Google's Gemini, and Microsoft's Bing Chat.

To capitalise on LLM models, businesses will need to create centralised locations for these data sets to exist, such as cloud storage or on-premises storage. Naturally, this requires significant investment not only in infrastructure costs but networking costs such as egress fees that come with moving this data.



Unified data and AI platforms

We are already seeing the emergence of fully managed data and AI analytics platforms that can support businesses through their generative AI led transformation.

These unified 'AI ecosystems' are designed to be multi-engine, multi-format, and multi-cloud, making it easier to store, analyse, govern, and transform your business data. One example is Google's BigQuery where you can connect your data to Al and then tune, train, and deploy multiple LLMs without copying or moving data.

This allows businesses to both reduce costs and minimise the risk of migrating data workloads through a one-stop portal. Companies will also be able to build data pipelines that blend structured data, unstructured data, and multimodel generative AI tools to create a new class of analytical application.



Virtual GPUs

Virtual graphics processing units (or GPUs) are one of the hottest AI products at present. GPUs were originally developed for rendering graphics in video games but have recently become entwined with generative AI because both technologies align with neural networks.

When technology firm Nvidia released its virtual GPU a few years back, the virtualisation of data centre GPUs allowed it to be shared across multiple devices through a back-end server rather than on devices such as PCs or laptops.

This vastly improved availability and allowed users to perform thousands of operations simultaneously. That said, the configuration of virtual GPUs can be extremely technical and requires a deep understanding of neural network technology.



Al-as-a-Service

Over the past 18 months, the rise of Al-as-a-Service (AlaaS) has allowed businesses to incorporate generative Al solutions into their operations without heavily investing in resources or development.

Instead, they can simply subscribe to an AlaaS



provider and choose the relevant AI tools for their customer service needs, so that could be things like chatbots, virtual assistants, machine learning frameworks, or application programming interfaces (APIs) such as Amazon, Rekognition or IBM Watson.



The move towards Edge AI

Edge AI allows data to be processed where the data is collected, rather than in the cloud or at offsite data centres. This means devices like smartphones or laptops can make smarter and quicker decisions.

As more AI models become configured inside these devices, they will be able to process data within milliseconds and deliver results in nearreal time. What's more, Edge AI will lower latency, reduce bandwidth, and provide greater data privacy and security.

Given these game-changing benefits, it's no surprise that some analysts predict a 50% rise in AI data processing moving towards the edge over the coming years.

Indeed, such is the unquenchable demand for Edge AI processing that data centres are beginning to feel the pressure, especially when you factor in the lack of available urban spaces and rising rental costs. This may require edge computing to find its own solutions in responding to market demand.



Managing larger volumes of data with network automation

The large volumes of additional data being generated by Gen AI and the requirement to move this data to and between private and public clouds is already starting to introduce greater cost and complexity to enterprise networks. As enterprises build hybrid and multi-cloud architectures to deliver and support Gen AI, they need to reexamine how they access the cloud.

"The rapid development of Gen Al creates a demand on networks that we have not seen before," says Paul Gampe, CTO of Console Connect. "These survey results demonstrate that when it comes to deploying Al tools, businesses are growing increasingly concerned about the need to be securely connected and that the public internet is no longer suitable for handling many of these applications and workloads."

"Automation and the move to a Network-asa-Service (NaaS) model presents a way for businesses to break away from traditional network infrastructure and the public Internet, and instead utilise automated, private and secure network connections to the cloud that can be dynamically adapted to meet the needs of generative AI."

"These survey results demonstrate that when it comes to deploying AI tools, businesses are growing increasingly concerned about the need to be securely connected and that the public internet is no longer suitable for handling many of these applications and workloads."

Paul Gampe, CTO of Console Connect



Automating at the edge

It is likely that as developers build LLM applications, they will need easier access to multiple data sets stored in the cloud and onpremise data centres.

NaaS platforms such as Console Connect are an effective tool for businesses to build networks between private and public clouds to facilitate



18

the transfer of data – without making significant investment in additional infrastructure.

If, as anticipated, more data is pushed to the network edge for processing, NaaS can also play a critical role in linking clouds and data centres with edge devices.

Like generative AI, the technology behind edge networking is evolving all the time. A recent advancement is the automated edge. In simplified terms, the automated edge is intelligence that is embedded into edge devices which takes care of networking tasks like configuration, responsiveness, provisioning, and even monitoring all by itself.

As daily operations shift and change, it will automatically optimise network performance by intelligently steering traffic through different pathways to reduce latency and interruptions while providing greater reliability and security.

"The automated edge can work hand-in glove with generative AI models to improve connectivity and real-time decision making from different devices, locations, and applications. This not only accelerates the analysis of your data close to the source, but also speeds up customer interactions and provides further opportunities for growth and innovation," says Paul.

How Console Connect can help with the move to Al

The rise of Gen Al and the migration of more workloads to the cloud means networks are handling larger volumes of more sensitive and mission-critical data than ever before. With this comes the needs for businesses to reassess their network model and ensure they have the immediate and long-term capabilities to support this new technology.

Console Connect is a Network-as-a-Service (NaaS) platform that makes it easy to provision and scale your connectivity globally. The platform runs on Console Connect's own underlying global network infrastructure, helping businesses connect securely and on-demand to clouds, applications, devices, data centres and office locations - using a network they can trust. With increased network complexity and cost on the horizon, Console Connect can help businesses meet the requirements of generative Al projects by:

- Increasing agility: Al technology is improving at a rapid pace. Peaks and subsequent troughs from multimodal LLMs are anticipated over the coming years. Switching to a NaaS model gives businesses the power to change and adapt their connectivity requirements to meet the unpredictable traffic patterns of Al. Users of our platform can order, change and manage network connections in real-time, with point and click provisioning. They can also dynamically adjust bandwidth to meet the needs of workloads - and only pay for the bandwidth they use.
- Enhancing security: Most generative Al systems store in the data cloud, presenting a new threat of data leakage. Businesses have been urged to integrate operational processes, architectures and tools that can ensure data privacy. The network plays an important role in ensuring data remains protected. With the public internet becoming increasingly unsuitable for carrying sensitive data, leaving it vulnerable to cyberattacks, businesses should explore alternative ways to access the cloud.

Console Connect improves data reliability by taking your traffic away from the public internet. The platform meets stringent privacy and security requirements by leveraging our private network infrastructure, which is backed by carriergrade SLAs. Users also have the ability to choose from different Class of Service (CoS) options so they can easily prioritise sensitive traffic.

• Reducing cost: Large volumes of data transferring to and from the cloud can present significant additional costs for your



business. Egress charges have been in the news a lot recently as enterprises find themselves battling spiralling cloud costs. Private dedicated connections from an enterprise network to a cloud provider can significantly reduce egress fees, which cloud providers charge for moving data out of the cloud.

Console Connect can help deliver egress savings of up to 50%. Recent research suggests that if your company is exchanging more than 25Mbps with a public cloud provider, or in price terms, spending more than \$700 per month on internet egress fees, then using a private network can already start saving you money.

Saving time: Transferring large volumes of data between public clouds or between clouds and on-premises isn't just costly. It can be time consuming too.

As well as significantly reducing service provisioning times from months to minutes, the high-speed connections delivered by Console Connect can also reduce the time it takes to transfer large volumes of data for AI applications.

For example, recently Console Connect helped an ecommerce business complete data migration between two public cloud platforms in just three weeks - an estimated ten times faster than if completed over a standard internet access service.

Improving access to the cloud: As highlighted in this report, access to the cloud is a critical consideration for IT decision makers as they move forward with gen AI projects.

For businesses that that have adopted a hybrid or multicloud strategy, Console Connect can help:

- Connect between different cloud platforms and cloud regions

- Connect on-premises data centres to the public cloud

- Automate data backups and recovery between clouds

The platform offers global reach to all the hyperscalers, such as AWS, Google Cloud and Microsoft Azure, as well as specialist providers, such as OVHcloud, Vultr and Wasabi. You can view our full cloud ecosystem <u>here</u>.

To find out more about how Console Connect and automation can meet the needs of your generative AI project, visit www.consoleconnect.com

Or talk with one of our NaaS experts by contacting sales@consoleconnect.com

19

This study was conducted by Arlington Research and was answered by 1,000 respondents in the UK, the US, Australia, Hong Kong, and Singapore between 3-16 January 2024. Please note, that graphs may not add up to 100% as each figure has been rounded to the closest decimal.

The global study focused on CTOs and senior IT leaders at director level, or above, in organisations with at least 50 employees. The sample demographics are as follows:

	IT	C-Suite
USA	80%	20%
UK	74%	26%
Australia	94%	6%
Hong Kong	76%	24%
Singapore	24%	76%

IT Leaders and C-Suite by country:

Size of organisation:

50 - 249 employees - 21% 250 - 999 employees - 40% 1,000 - 4,999 employees - 31% 5,000 or more employees - 8%

Employees by country:

	50-249 employees	250 - 999 employees	1,000 - 4,999 employees	5,000 or more employees
USA	16%	36%	38%	10%
UK	18%	34%	36%	12%
Australia	28%	48%	20%	4%
Hong Kong	10%	47%	38%	6%
Singapore	33%	36%	26%	6%

NET employees by country:

	NET: 50-999 employees	NET: 1,000+ employees
USA	52%	48%
UK	52%	48%
Australia	77%	23%
Hong Kong	57%	43%
Singapore	69%	31%

Level of seniority and expertise:

Owner / Partner – 20% Senior Management (e.g. CEO / VP/ Managing Director) – 42% CTO / CIO – 38%

Level of seniority and expertise by country:

	Owner/Partner	Senior Management	сто/сіо
USA	20%	40%	40%
UK	20%	45%	35%
Australia	20%	40%	40%
Hong Kong	21%	42%	37%
Singapore	22%	41%	37%

Level of decision making:

I am the sole decision-maker – 74% I make the final decision with input from staff/management – 19% I help reach the final decision as part of a group/committee – 7%

Sole decision-makers by country:

Total	74%
USA	78%
UK	69%
Australia	68%
Hong Kong	92%
Singapore	70%

Industry sector:

IT/Technology/Communications – 78% Finance/Banking/Insurance – 7% Manufacturing – 5% Retail – 3% Business Services – 3% Education – 2% Energy/Utilities – 1% Travel/Transportation/Logistics – 1% Hospitality – 1% Other – 1%

IT/Technology/Communications by country:

Total	78%
USA	86%
UK	80%
Australia	94%
Hong Kong	62%
Singapore	39%



www.consoleconnect.com sales@consoleconnect.com

