techUK
FOR WHAT COMES NEXT

Sector Readiness for Climate Change Risks: Data Centres

Report to DEFRA under the Adaptation Reporting Power: Third Round

December 2021

Contents

1 Preamble

1.1 Abstract

This voluntary submission to DEFRA under the Adaptation Reporting Power provides an update to, and should be read in conjunction with, our 2016 report. However, there is a material shift in scope: now that communications report separately, our 2021 submission is solely focused on data centres.

This submission explains what data centres are, what they do and why digital infrastructure is important to the UK economy. It provides a brief overview of the UK data centre market including approximate site numbers, distribution and business models. Climate change threats pertinent to the UK's data centre sector are identified and summarised. The report considers sector readiness for these risks and explores inherent resilience through redundancy, the application of industry standards and performance metrics and internal and external drivers. While data centres have not suffered climate change outages sufficient to provide an evidence base, proxy events including the pandemic provide invaluable insights and lessons. These are reviewed together with information from industry surveys on attitudes and preparedness.

Areas of concern include asymmetric interdependencies, potential gaps within existing industry standards in terms of scope and adoption, access to skills, supply chain bottlenecks and managing climate change risks when activity is offshored. The most pressing concern, however, relates to on-premise data centres and server rooms where resilience may not match the criticality of the activity. This is compounded by lack of transparency and reporting within this cohort and the data that does exist suggests that operational best practice lags far behind the commercial sector. Recommendations include greater scrutiny of on-premise data centres and a review of standards and practices to ensure climate change risks are accommodated.

1.2 Scope of this report

This is a voluntary report submitted under the Adaptation Reporting Power summarising the resilience of the UK's data centre sector to climate change risks. As an industry association we provide a collective voice for commercial operators in the UK so the scope and detail of our reporting is limited to a general summary of readiness at sector level and will not have the granularity of an individual corporate report because we are not in direct control of the way that risks are managed. Moreover, much of the UK's data centre estate is owned and operated by private sector organisations and site-specific information is generally not disclosed for reasons of operational security and resilience. We are also not party to, and are therefore unable to comment on, individual corporate risk plans.

The community that we represent comprises predominantly commercial operators, plus telcos, IT service providers and some financial institutions. This, however, does not provide a full picture of the data centre sector in the UK where a significant proportion of activity takes place in-house[i]: in dedicated purpose-built facilities, in small on-premise data centres, server rooms and distributed IT (servers in cupboards and closets within office buildings). While commercial operators compete on the basis of their ability to provide business continuity for customers and tend to work to international, peer-reviewed standards and KPIs, most enterprise operators (both public and private sector organisations that operate in-house data centres) are under no obligation to be candid about operational resilience. As a result, little is known about this cohort: while we must acknowledge their existence because of the associated risk, we cannot speak on behalf of these providers.

1.3 Relevant Adaptation Scenarios

This submission is based on the projections published by UKCP for 2018 for the period to 2050. This is because a data centre asset life is currently around 25-30 years, and because projections post 2050 are likely to be heavily influenced by global emissions and are therefore subject to much greater uncertainty. Broadly these anticipate hotter summers with the potential for significant reduction in precipitation, warmer wetter winters with significant potential for more precipitation, more storms and extreme weather events and sea level rise.

1.4 Quick review of our response under ARP in 2016

Our previous submission explored the climate change readiness of the UK's data centre sector, in addition to several informal observations on fixed line and mobile communications. The report explained the main features of our core digital infrastructure – data centres, fixed line telephony and mobile telephony - and how they fit together. It recorded the information sources that we were using to assess climate change risks at the time, and it highlighted the main threats to the operation of our digital infrastructure and to the delivery of the services that depend upon it.

We set out some of the approaches that were being deployed within the sector to identify, manage, and mitigate risks. The report then reviewed several recent climate change related incidents that had resulted in interruptions to service, considering what we, as a sector, had learned and what actions were being taken. Finally, it explored several areas that required further examination to ascertain whether they represent potential vulnerabilities, suggesting where there was scope for action.

1.5 What actions have we taken at sector level since the last round?

While the primary focus of this report is on operational and strategic measures undertaken within the UK's data centre operator community to address climate change risks and build adaptive capacity, here we take the opportunity to include a short summary of the steps we are taking as the relevant trade body to support the UK's commercial data centre service providers.

In 2016 we undertook to monitor incidents within the sector, to raise awareness of the nature of climate change risks, the information available, and how it should be used by data centre operators. We agreed to propose the addition of requirements for regular flood risk reviews to existing industry standards. We also undertook to continue to work with external stakeholders, including government and media.

We have monitored incidents through trade press, Uptime Institute surveys and DCIRN (the Data Centre Incident Reporting Network). We have held briefings with operators on climate change readiness, developed a library of relevant resources and engaged with standards bodies to upgrade existing requirements. We have engaged productively with other infrastructure operators through the IOAF (Infrastructure Operators Adaptation Forum) and held inter-sector dialogues on interdependencies. We have continued to work with government and have responded to relevant Inquiries. We have engaged with a wide range of stakeholders nationally and internationally and we have published multiple articles relating to climate change resilience.

1.6 Further Information and contacts

Links and reference relevant to the content of this report are included and listed separately at the end. All techUK publications relevant to data centres can also be found on our **Data Centres Programme Index**.

Contacts

Emma Fryer
Associate Director, techUK
Mob: 07595 410 653
emma.fryer@techuk.org

Adam Young
Programme Manager, Environment
Adam.young@techuk.org

Lucas Banach
Programme Assistant
Tel: 020 7331 2006
Lucas.banach@techuk.org

2 The UK Data Centre Sector: Overview

2.1 What are data centres?

Data centres[ii] are highly resilient facilities that underpin our modern economy by processing, managing, storing and transacting digital data and, with communications networks, form our core digital infrastructure[iii]. A data centre essentially consolidates organisational IT functions. It should provide a secure, resilient, and controlled environment for IT equipment (servers and networks) and supporting hardware and be equipped with guaranteed power supply and high bandwidth connectivity. Redundancy (duplication) of networks, power and other infrastructure ensures continuity. Building management controls such as air conditioning maintain environmental conditions for the equipment within a specified envelope of temperature and humidity, and advanced security systems ensure that the facility and its data remain secure. A data centre may be on-premises or remote. It may be purpose-built or be part of an existing building. It may be operated in-house to support organisational IT functions or by a commercial third-party provider. It may be dedicated to one organisation or service multiple customers.

2.2 And why do they matter?

Data centres support every conceivable part of our modern economy: business processes, Government services, telecommunications, transport infrastructures and social networks all depend on computers interacting in this way, exchanging digital information. Data centres enable retailers and banks to process financial payments, supermarkets to resupply, delivery companies to manage logistics and public authorities to deliver services and messaging. Some sites are officially deemed CNI (critical national infrastructure) to reflect the nature of the activity being managed therein.

Data centres underpin an internet economy that contributes over 16% of domestic output, 10% of employment and 24% of total UK exports[iv] and is growing faster than any other in the G-20. the UK sector is a real success story, is globally important and provides the technical infrastructure for financial services, aerospace, transport, healthcare, retail and utilities. Each new data centre contributes between £397 M and £436 M GVA per year to the UK economy[v] while that of each existing data centre is estimated to lie between £291 M and £320 M per annum.

Data centres are a *de facto* part of our infrastructure but UK Government has, until recently, demonstrated a strangely equivocal attitude to data infrastructure in the UK. Despite the fact that data centres are the point where our industrial strategy meets our digital strategy, they were not mentioned in either document. Data centres are also absent from the National Planning Policy Framework and the National Infrastructure Commission was established without any digital remit, and although communications were bolted on, the existence of a world class data centre sector in the UK does not appear to trouble this body. Officials regularly express surprise that the UK has a data centre sector and it has been exceptionally difficult to engage at strategic level with entities like National Grid on pressing issues like power provisioning. The establishment of a dedicated team within DCMS in 2020 as part of the UK's Covid response was a welcome development, looks likely to be permanent and is definitely a significant improvement.

2.3 How many data centres are there in the UK?

We estimate that there are around 500 recognisable facilities. We consider a recognisable facility to have a minimum power supply of 240KW, a floor area of over 200M^2, environmental controls and operational redundancy including emergency back-up power to allow continuous running.

Around 200, including most of our largest sites, are run by commercial operators who provide data centre services to third parties. These are called "colocation" as customers usually lease space in which to deploy their own IT hardware, "colocated" with the servers of other organisations. Reasons for outsourcing include security, resilience and cost: data centres are eye-wateringly expensive to build and maintain, especially for organisations where they are not part of the core business offering. Outsourcing moves this to a specialist provider and also moves the financing from a capex to an opex model. Perhaps as many as a hundred sites are run by telecoms operators and IT services providers. The remainder are dedicated

facilities, supporting corporate IT functions and customer services for organisations like banks, retailers and universities. These are known as "enterprise" facilities because they are dedicated to supporting the business; the enterprise. These may be on-premises or remote from the business. Many organisations mix and match – outsourcing mission critical activities but keeping more mundane functions in-house – or vice-versa.

Rapid expansion of the cloud services market is driving growth in UK data centre development because cloud services are delivered either directly or indirectly from data centres. Cloud computing is essentially the process of accessing IT functions via the internet – so applications and activity are held remotely in data centres, rather than on personal devices like laptops, PCs, tablets and phones. A cloud service provider may operate their own data centre or lease space from a colocation provider and there are many different business models adopted in the UK. See the business model diagram below.

Some organisations do not consolidate their IT into purpose built or dedicated data centre facilities. Instead they run smaller data centres or server rooms on premises. IT that is kept in server rooms, closets and cupboards is known as "distributed IT". There are no formal criteria that differentiate a data centre from a server room, but broadly speaking, a combination of power supply, resilience and server capacity are used. The fuzziness of the definition is a problem for reporting of this kind. Small data centres and server rooms may not provide the resilience to meet the definition above, but nevertheless might perform important functions. There are many thousands of server rooms and small data centres on premises throughout the UK, both in the public and private sector.

2.4 Where are data centres located in the UK?

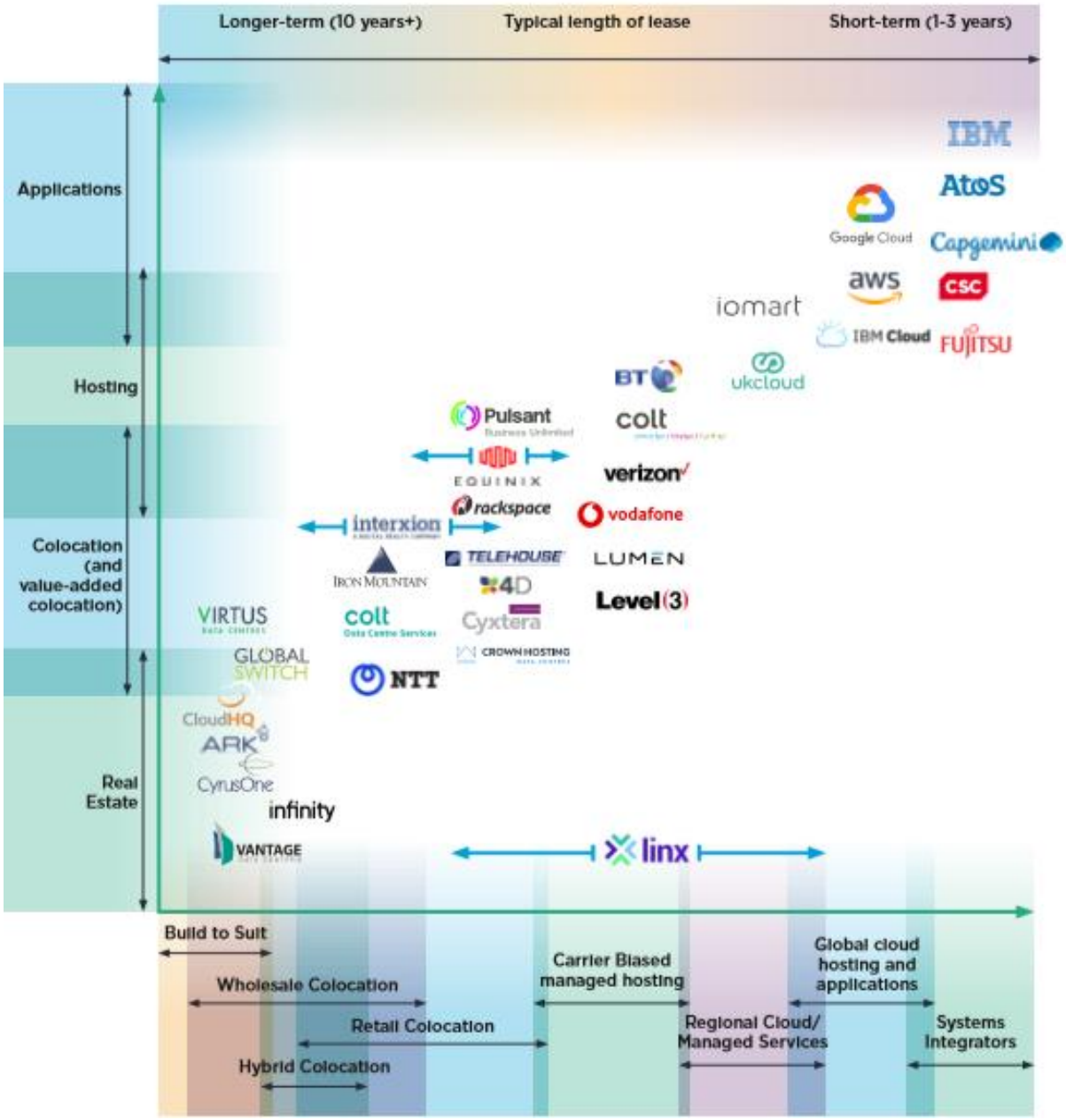
Within the UK, colocation (commercial) sites by and large are heavily clustered, with about 70% of the UK market in and around the M25. The second largest cluster is in Manchester. Smaller clusters and individual sites add further capacity and may also serve regional markets or provide disaster recovery. Enterprise sites, because they service a single business, tend to be located near head office. For example, sites were built in Norwich for Norwich Union and Halifax for the Halifax Building Society, though in many cases the original owners now outsource their data centre requirements and many such sites have been sold and now operate in the commercial market.

2.5 Data centres are a critical part of our core digital infrastructure

Our core digital infrastructure is not a single system but multiple systems and networks that interoperate. The three main constituents are fixed line telecommunications (made up of the high capacity and highly resilient core network plus the access network which runs from the exchanges to tens of millions of individual customer premises), mobile telecommunications (that interact with the core network but provide customer coverage through a cellular network) and data centres (that manage, transmit, process and store data for government, businesses, individuals and academia). Satellite and broadcast communications also play important roles in digital infrastructure.

2.6 UK data centre business models: Who does what?

This diagram roughly positions a subset of UK operators by primary service. The intention is just to present the range of activity and demonstrate that data centres don't all do the same thing. Many of the operators shown below provide services across multiple columns. The key deductions from this chart are a) that the data centre market is dynamic and b) that there are multiple providers at every level which indicates that a genuine competitive market exists for all type of service. This ensures continued focus on resilience and limits scope for complacency that might become a risk in a more monopolistic or oligarchic marketplace.



Notes

- Many data centre operators offer a broader range of services than implied here. There is a strong trend to broaden service offerings so in reality, boundaries between data centre services are blurred and likely to become less easy to define over time.
- Large (global) cloud service providers (CSPs) may operate their own data centres or take space in third parties, particularly within wholesale colocation. CSPs will often use colo facilities to scale initial operations and then look to build their own facilities over time or utilise a third party operator to enable a build to suit solution.
- Cloud, hosting and application service providers may be customers of the colocation providers.
- The importance of the global CSPs in our market is understated within this diagram. They are driving much of the growth we are seeing in the UK market at the moment and this is happening predominantly in wholesale colocation (the left hand side).
- This landscape tends to change rapidly so must only be regarded as a snapshot in time.
- This chart is not comprehensive and only represents a sample of operators. There are many providers not represented here.
- LINX (London Internet Exchange) provides interconnection services across multiple business models
- This overview will not reflect recent or ongoing M&A activity in the market.

3 Climate Change Threats and Impacts Relevant to Data Centres

Climate change risks relevant to data centres, and broader digital infrastructure, include fluvial and pluvial flooding, increased winter rainfall, increased severity and frequency of storms, lightning, high winds and heavy rain, increased summer temperature, higher humidity, increased speed of temperature and humidity change, prolonged periods of sustained high temperatures, and drought.

3.1 Potential physical impacts of climate change on data centres

Potential physical impacts include flooding of buildings, ducting and other assets; water, silt and salt damage, scour of cabling and foundations, subsidence to buildings and masts, access issues for engineers and staff, disruption to fleet operations, cable heave from uprooted trees, higher costs of cooling, shorter asset life, reduced reliability, fractured ducts, and higher operating costs. The table below summarises the first and second order impacts and the green and amber shading indicates where these are most likely to be manifest.

Impacts	First Order	Second order
Coastal flooding, (erosion, inundation by salt water, increase in salt spray)	Large scale data centres are unlikely to be built in coastal areas without major risk mitigation in the form of hard defences. Risks include flooding, scour damage to foundations, subsidence, salt damage to materials. Problems with emergency access for service engineers.	Interruptions or damage to power supply system. Flooding of exposed communications infrastructure, damage to cabling or cable landing stations, scour, corrosion. Slow response times if access is problematic.
Fluvial flooding (erosion, inundation, silt and sewage deposit)	Flooding, silt and sewage, water ingress and/or damage to heavy plant and switchgear, erosion and scour of cabling and buildings. Problems with emergency access for engineers.	Interruptions or damage to power supply system. Flooding of exposed communications infrastructure, damage to cabling or cable landing stations, scour. Slow response times if access is problematic.
Pluvial flooding (flash floods, inundation of localised area)	Flooding of facilities. Heavy plant and switchgear disabled, damage to cabling, water damage to other hardware. Problems with emergency access for engineers.	Interruptions or damage to power supply system. Flooding of exposed communications infrastructure, damage to cabling or cable landing stations, scour. Slow response times if access is problematic.
Sustained high summer temperatures	Poor working conditions for staff. Some legacy sites may struggle to maintain required temperature or avoid hot spots. May compromise some activity if cooling cannot be maintained. Cooling costs may increase for other facilities.	Power systems may need to meet temporary increase in demand
Increased rapidity of temp change	Higher HVAC (Heating, Ventilation, Air Conditioning) costs. Stress on components and hardware	Power supply and communications networks may be similarly affected.
increased humidity / humidity changes	More active humidity management required. higher risk of damage to hardware, may affect reliability and life expectancy.	Power supply and communications networks may be similarly affected.
increased storminess - wind and lightning	Most significant impacts likely to be second order	Interruptions in power supply, damage to overhead cables, substations or distribution system. Damage to exposed communication cables.
Drought	Access to cooling water for water cooled facilities. Subsidence.	Cable damage through subsidence

Note on Drought: Water use varies depending on the cooling technology deployed and while in the UK drought has not been reported as an issue for operators, lower reliance on water will improve resilience should such scenarios arise.

3.2 Non-physical impacts of climate change risks

Non-physical impacts include reputational damage, failure to meet customer SLAs (service level agreements), failure to meet regulatory objectives, high customer call volumes, impacts on staff wellbeing and unbudgeted costs.

4 Assessing Sector Readiness: How are Data Centres Preparing for Climate Change Risks and Building Adaptive Capacity?

Introduction

Climate change risks are handled as just one of a myriad of business risks facing data centre operators, and we consider this to be an appropriate approach. Risk planning for data centres happens at three stages – site selection, design, and operation.

Site selection: Data centre site selection is a specialist activity and all sites are subject to very rigorous assessment for a wide range of risks. Due diligence activity for data centre locations attracts larger premiums than any other sector due to the complexity of projects, the specialist expertise involved and the intense scrutiny required. Flood is high on the list of risk factors when choosing a location for data centres. Although there is no agreed risk threshold, industry practitioners generally seek a risk below 1 in 1000. This is, however balanced with other factors and emphasis is on managing and mitigating the risk rather than working to inflexible thresholds.

Design and build: Data centres are designed and built to be resilient and accommodate a wide range of risks, including those from severe weather, and extensive use is made of weather datasets to inform design specifications (see below). Peer reviewed, publicly available industry standards are widely adopted. Again, and as discussed in more detail below, the data centre sector has a well-developed range of bespoke operational standards focused on resilience in the form of the EN50600 series and the equivalent ISO/IEC 22237 series of Technical Specifications. Designing resilience is not limited to physical protection for individual sites: data centres, especially those developed for large cloud providers, may be mirrored in real time or be designed to move workloads between facilities and/or availability zones. Continuity of power, communications and cooling are major priorities (see below) with redundancy built in at every level.

Operation: Because climate change risks associated with a specific location may change over time, operational risk management is critical to ensure that a process of regular review is in place. Continuous application of resilience measures, regular reappraisal of risk, application and adherence to relevant industry standards are all relevant. While elements of the EN5600 series and parallel ISO/IEC 22237 technical specifications do cover data centre operations, the most commonly adopted operational resilience standard is ISO22301, governing business continuity. This is a generic standard unlike the EN5600 series, not bespoke to data centres.

Additional operational measures may become necessary to mitigate emerging risks - for instance if the flood risk associated with a specific location increases, or to upgrade or reconfigure infrastructure if new weather data suggests that the original design specification may not be sufficient to accommodate the scale of change in climate change risks. In more extreme cases it might be necessary to move workloads or, in time decide to downgrade or even decommission a site.

The following section explores how the data centre sector delivers resilience and is subdivided as follows:

- **Inherent or in-built resilience**
- **Resilience-related standards development**
- **Internal and external drivers: reputation, policy and financial drivers**

We then look at evidence to try and establish how resilient data centres are in practice

- **Proxy evidence from comparable events**
- **Survey responses**

4.1 Systemic resilience

Data centres, like broader ICT infrastructure, have systemic characteristics that make them relatively resilient to climate change.

Built-in redundancy

Data centres, because they compete on their ability to provide continuity of service, tend to build-in redundancy at every level of operation. These include but are by no means limited to:

- External power supply where a resilient facility will have separate, independent electricity feeds from the grid.
- Emergency generation, which will be predicated on the maximum theoretical power draw that the facility could impose, plus additional redundancy, denoted by configurations like 2N or N+2.w, where N is the maximum theoretical power demand of the site. Sites usually have priority arrangements for gasoil replenishment with fuel providers to accommodate longer outages.
- Communications: resilient sites will have dual or multiple independent connections.
- Cooling: again cooling will be configured to the maximum possible required, plus headroom, with additional redundancy denoted in a similar way (2N, N+2, 2N+1, etc) to ensure compliance with contractual obligations / SLAs.

Accommodating interdependency

As indicated by the tendency for built-in redundancy (see above), data centre design already accommodates some interdependency risks, primarily those from interruptions in power supply and communications. For instance, the data centre definition that we use for the Climate Change Agreement requires a back up electricity supply in the event of mains failure without which the facility is not considered to be a data centre. Interdependencies relating to data centres and ICT more generally seem to be asymmetric: data centres appear to be heavily dependent on a few sectors, but are critical to a wide range of sectors.

Nature of digital infrastructure

The fact that digital infrastructure comprises multiple systems interoperating confers some degree of natural redundancy. There is generally more scope to re-route data traffic than other utilities, even at scale – the way that cloud service providers build availability zones is an example (see more detail on this below). In terms of connectivity, there is capacity in the network to accommodate even quite a sudden increase in demand – the wholesale move to online activity during the pandemic was accommodated relatively seamlessly (see this useful LINX blog).

Asset life

Data centre asset life is relatively short. The design life expectation for the facility itself is around 30 years and although a well-positioned site will support data centre activity for longer, the data centre infrastructure would have to undergo a period of major refurbishment in order for it to remain fit for purpose and competitive in a world of rapidly evolving technology. Emergency generating plant is expected to serve for most or all of the lifetime of the facility but other critical elements, such as CRAC units (which provide cooling), are designed for around ten years of operation. This provides scope for regular upgrades of operational equipment in line with changing threat levels. The asset life of ICT hardware within the facility is significantly shorter than the supporting infrastructure. Computer servers in particular are refreshed every few years, so more resilient IT assets can be deployed as part of the natural replacement cycle within appropriate timeframes.

Technology development

Rapid pace of technological development and innovation gives the sector an inherent advantage in responding to change because of the potential to innovate around threats. This is particularly true of computing equipment (for instance improvements in chip design) but we are also seeing new developments in cooling technology. Currently, operators are trialling liquid or immersive cooling solutions which have much greater potential than air cooling to extract heat from computing equipment

which in turn may help reduce the risk of overheating. These approaches are also more likely to produce heat in a reusable form at the end of the process.

Backups, duplication and site replication

Data centres, especially those hosting critical functions, deploy a range of approaches to minimise data loss and speed up recovery from event-related outages or failures at site level, irrespective of cause. The Uptime Institute sums this up well: *"In many ways, data centers seem to be on top of the problem. The traditional method of attempting to guarantee a reasonably quick recovery from a catastrophic failure, regular backups to a secondary site, is deployed by 68 percent of the respondents, and 51 percent make near real-time replication to secondary sites (with 40 percent replicating to two or more sites). Newer methods are also gaining traction. Forty-two percent said they utilize some sort of disaster recovery as a service program, and 36 percent take advantage of cloud-based high availability service"*

Regions and Availability Zones: Ability to move workloads

Large cloud service providers like Google, Amazon and Microsoft have developed availability zones within the regions where they deliver services. Within each region, these availability zones are essentially separate locations from where customers access services. Workloads can be duplicated across different availability zones and activity can be moved, for instance if one site is compromised. This is important from a resilience perspective because it reduces reliance on individual sites. Redundancy is therefore delivered by the ability to move workloads and the provision of sufficient capacity to do so.

Competitive market around resilience

The sector competes on availability and continuity of service but not all data centres are built to the highest levels of resilience: the level of resilience should be dictated by the criticality of the activity that is hosted within the facility and this will be reflected contractually in service level agreements. The greater the level of redundancy the greater the cost and there is also an energy efficiency burden that includes the embedded energy associated with additional plant.

Genuinely competitive market for all service offerings

As mentioned earlier (see business model schematic) the UK is fortunate in the variety of data centre business models and service offerings available. While there is some consolidation in the market there is genuine competition at every level of provision which reduces the risks of complacency that might be seen in more oligarchic or monopolistic markets. It is also largely business rather than consumer facing. As a result the sector has not been burdened by an official regulator applying price controls. Evidence from other infrastructure sectors suggests that some regulators have focused too much on short-term considerations that fall within price review periods and have failed to accommodate, or allow for, longer term investment in resilience.

Outsourced business model

The tendency to consolidate and outsource IT function in purpose-built data centres improves resilience, transparency and accountability. Operational responsibility may remain in-house, with the data centre commissioned by an organisation as a dedicated facility to support its operations (for example a bank), it may be built and operated as a dedicated facility for a customer by a third party provider or activity may be migrated to a cloud or colocation service. Purpose built data centres are designed for exceptional levels of resilience that cannot be achieved on normal business premises or might impose very challenging levels of capex. Industry analysts concur that the need for resilience is a driver of the strong trend towards outsourcing.

Small and collaborative community

Data infrastructure may be critical, but the data centre industry is relatively small which means that news travels fast and nothing significant can happen without everyone knowing about it, even if outages are not reported formally. Lessons are learned quickly. Flood and wildfire incidents in the US during the last two years may be one of the reasons for Uptime surveys reporting increased focus on climate change among operators.

4.2 The role of industry standards and design specifications

Despite being a relatively immature industry, the data centre sector has developed an impressive range of peer reviewed, international standards, KPIs, metrics and toolkits covering resilient design and operation. Bespoke industry standards (specific to data centres) addressing resilience include the EN50600 series developed by CEN/CENELEC. They relate to the data centre facility itself (the infrastructure) and are described as "availability classes" and in this context availability is synonymous with resilience and essentially means "available to provide services". We explore these in more detail below. These standards are gradually being harmonised with global (ISO and ITU) standards to improve consistency.

Data centres also work to a range of generic (non data centre specific) risk standards such as ISO 31000 and ISO 22301. Operators frequently refer to metrics for uptime in terms like "nines". Working to "five nines" for instance means that they undertake to deliver at least 99.999% of continuous running (see the table below). There are also proprietary resilience standards available, such as the Uptime Institute's Tier series.

Operators are unlikely to adopt new resilience-specific standards in addition to, or in preference to, existing industry standards. The priority is therefore firstly to ensure that the standards that the data centre industry uses are fit for this purpose – i.e. they include appropriate climate change resilience requirements, and secondly to ensure that these standards are widely adopted. With respect to the first point, we understand anecdotally that the Uptime Institute is currently undertaking a major review of resilience standards and we will seek more information of the scope and timing of this important activity.

4.2.1 Generic reference data for building design specification

Design specifications for data centres are informed by granular, location-specific weather data. Common sources for weather data to inform design specifications for data centres (among other building types) are ASHRAE and CIBSE:

CIBSE (the Chartered Institute of Building Services Engineers) is a key reference for building design guidance and standards for the data centre sector. CIBSE publishes extensive weather data for use in building performance analysis and design. These figures are produced in collaboration with the UK Climate Impacts Programme, Arup and the University of Exeter, and include the latest climate change projections so that data centre design professionals, and others within the construction sector can future-proof buildings.

ASHRAE (the American Society of Heating, Refrigeration and Air Conditioning Engineers) publishes external ambient weather data for every country and major city in the world which informs the design specifications for new data centre builds as well as upgrades or refurbishment of existing facilities. ASHRAE weather data is retrospective but updated every four years. Designers choose to base specifications on 20 years, 50 years or 100 years of data. The ASHRAE retrospective 20-year dataset is recognised by Uptime Institute, an authority on resilience, as appropriate.

At present the relationship between the CIBSE weather data and ASHRAE data is unclear. Both datasets sit behind a paywall so we have not had the opportunity to compare them. The other concern that has been raised is that these datasets are generic and therefore not necessarily issued with data centre requirements in mind. We believe these points need further investigation and a comparative review of the datasets has been included on our action list.

Opinions differ within the industry as to whether retrospective data is adequate: some observers suggest that future projections should form the basis for current data centre design. However, most within the sector argue that, despite it being retrospective, the ASHRAE data provides more than enough headroom, that sites are designed to accommodate a full IT load although they never run above 85% of design capacity (more usually 50%), and that there is additional redundancy in cooling systems over and above the requirement specified against these criteria. They also note that the life of a cooling unit is around ten years, so there is scope to upgrade. Running a data centre with too much redundancy is economically and environmentally inefficient, so getting this balance right is important.

4.2.2 Bespoke design standards for data centres: EN 50600 Availability Classes

The EN 50600 series is a set of data centre standards developed by international standards body CENELEC. CENELEC is internationally recognised, vendor agnostic, not for profit and peer reviewed. The EN 50600 series is made up of multiple components. EN 50600-1 covers general requirements and availability classes. The following extract from "Data Centre Assessment vs Certification" explains this in more detail:

- EN 50600-2-2 defines four levels (1 to 4) for the design availability of the power supply and distribution system of the data centre;
- EN 50600-2-3 defines four levels (1 to 4) for the design availability of the environmental control system of the data centre;
- EN 50600-2-4 defines four levels (1 to 4) for the design availability of the telecommunications cabling systems of the data centre;
- EN 50600-1 defines the overall availability level of a data centre based on the lowest level of three infrastructures detailed above.
- EN 50600-2-5 defines requirements for the maintenance of physical security of data centre spaces independent of the infrastructure level.

In this way EN 50600-1, -2-2, -2-3, -2-4 and -2-5 provide a comprehensive framework for assessment of the design availability of a data centre. They are supported by EN 50600-2-1 and EN 50600-3-1 for building construction and operation respectively but these are currently treated as subsidiary. The EN 50600 series assesses the availability of the data centre infrastructure as opposed to the availability of the data centre function and this is an important distinction that has recently been clarified.

Availability classes

While not focused specifically on climate change risks, the resilience of a data centre facility is assessed against four classes of availability (see figures). A data centre with Class 1 availability would probably have a single power supply, a single communications connection and enough battery power to allow it to shut down safely in the event of a power cut. A class 4 facility would have three separate sources of power – most likely two separate grid supplies and one generator or vice versa. In terms of communications there would be diverse routed fibre backbones with multiples paths to devices enabled to receive multiple inputs (See chart).

Source[vi]: CEN/CENELEC/ETSI	**Availability of overall set of facilities and infrastructures**			
	Low	Medium	High	Very high
	AVAILABILITY CLASS			
Infrastructure	**1**	**2**	**3**	**4**
Power supply/ distribution EN 50600-2-2	Single-path (no redundancy of components)	Multi-path (resilience provided by redundancy of systems)	Multi-path (resilience provided by redundancy of systems)	Multi-path (fault tolerant even during maintenance)
Environmental control EN 50600-2-3	No specific requirements	Single-path (no redundancy of components)	Single-path (resilience provided by redundancy of components)	Multi-path (resilience provided by redundancy of systems), allows maintenance during operation
Telecommunications cabling EN 50600-2-4	Single-path using direct connections	Single-path using fixed infrastructure	Multi-path using fixed infrastructure	Multi-path using fixed infrastructure with diverse pathways

The objective of EN 50600 is not to be too prescriptive in terms of requirements. So it does not set a risk threshold for data centres or require, for instance, that no data centres are located in flood zones. There might be very good reasons for locating a data centre where the flood risk is slightly higher than is desirable. Instead, it focuses on mitigating and managing those risks. The standard also covers ongoing risk management and EN 50600-3-1 requires operators to have management procedures in place that ensure the risk assessment is ongoing. This is critical because flood zones change and need regular review.

The 50600 series is widely recognised within the industry and although the standards were developed relatively recently and are still being rolled out, adoption is growing rapidly. There is currently no certification process for this series but that does not mean that a data centre cannot be assessed against this standard or that the EN50600 standard cannot be used as part of the certification process against another standard such as ISO 9001, for which certification does exist.

Within the industry, data centre availability is sometimes described in terms of a percentage. So for example a data centre may have 99.999% (or "five nines") availability – see figure. What this means is that the data centre is designed and operated to ensure an absolute maximum of 5.3 minutes of down time in a year. Generally it has none. However; the nines approach has shortcomings – the 5.3 minutes could represent one outage or many.

Availability (A)	Common reference	Downtime (based on a 365 day year)
90 %	1-nine	36,5 days
99%	2-nines	3.65 days
99.9% (3-nines)	3-nines	8,76 hours
99.99% (4-nines)	4-nines	52,6 minutes
99.999% (5-nines)	5-nines	5,3 minutes
99.9999% (6-nines)	6-nines	31,5 seconds

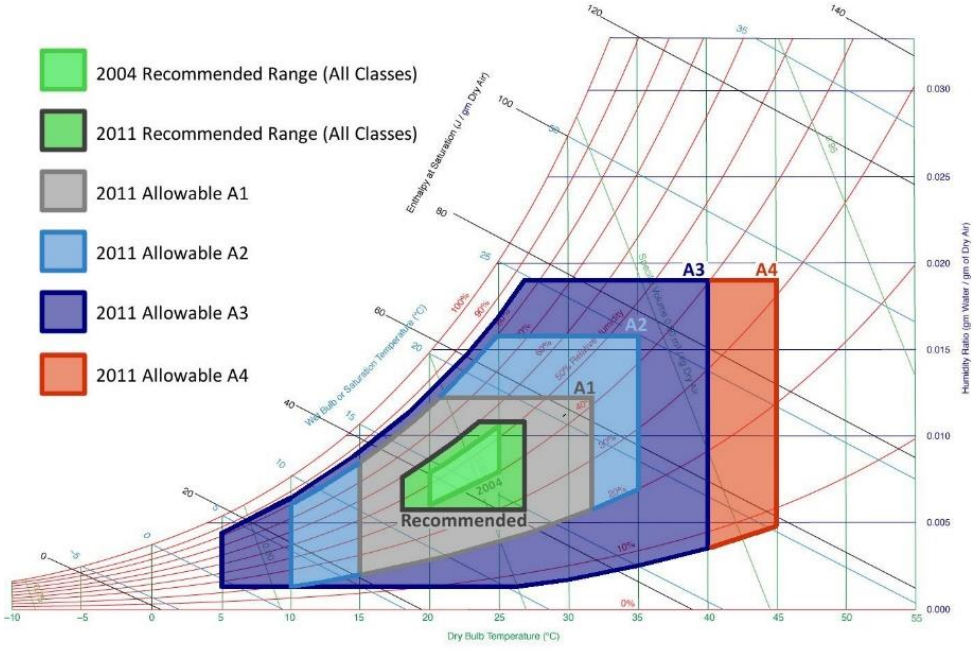
4.2.3 Standards relating to IT hardware relevant to an adaptation scenario

The fundamental purpose of a data centre is to provide a safe and secure operating environment for computer servers. Much work has been done to expand the temperature and humidity ranges within which computing equipment will work reliably. This has been driven in a bid to improve operational efficiency because a computer server that operates at higher temperatures needs less cooling. However, the ability of computing hardware to withstand fluctuations in temperature and humidity is clearly also relevant to climate change resilience.

ASHRAE has defined operating envelopes for servers in terms of temperature and humidity boundaries. Within these envelopes, manufacturers' warranties for reliable operation to remain valid. Some envelopes relate to continuous running and others to exceptional running. Although they relate to the IT equipment that the data centre houses, they are important in an adaptation and resilience context because in reality they set out operational temperature and humidity limits for the facility.

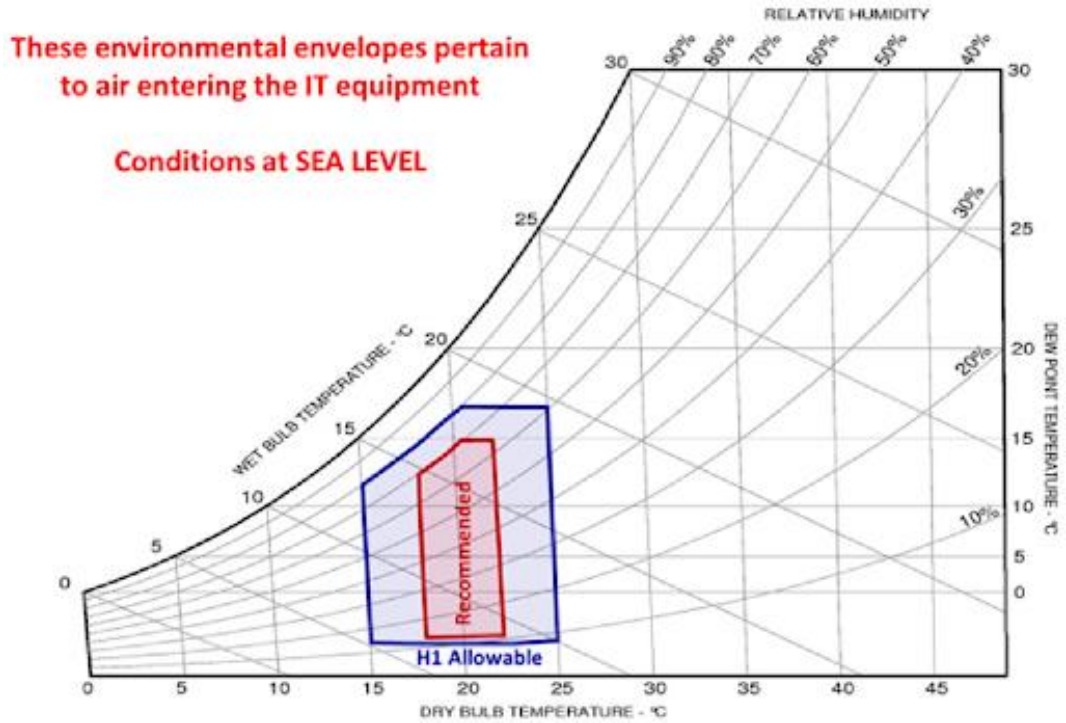
The ASHRAE envelopes have been expanded as modern server design has improved. The ability to operate at greater ranges of temperature and humidity delivers a material improvement in operational resilience at the IT hardware level. Legacy (old) servers, however, are unlikely to be warrantied to the same limits as more modern models. For this, and other operational reasons, data centres hosting old servers may be less able to cope with heatwaves, especially if high temperatures are sustained. See our input to The Environmental Audit Committee on heatwaves. These charts from ASHRAE illustrate these envelopes.

Server tolerances: ASHRAE psychrometric charts



ASHRAE psychrometric chart showing a common set of guidelines for operating conditions in data centres that would not invalidate the warranties provided by server manufacturers. In 2004 the original envelope proposed was 20°C to 25°C (68-77F). In 2008 the range expanded from 18°C to 27°C (64.4-80.6F). In 2011 the envelope was challenged again and allowable operating ranges as wide as 5° to 45°C (41 to 113 F) have been considered. Note that classes A1 and A2 include both new and legacy equipment. Classes A3 and A4 do not include legacy equipment. In other words, legacy (older) servers are not guaranteed to run as reliably as new ones at increased ranges of humidity and temperature.

Image source: ASHRAE/Don Beatty Associates



However, one size fits all approaches do not work in data centres. ASHRAE envelopes should be applied with care: allowable humidity levels are reduced if air pollutants are present, and for high density computing the envelopes are smaller to account for reduced airflow. This diagram shows the ranges suitable for high density computing. Source: ASHRAE, image accessed from DCD: https://www.datacenterdynamics.com/en/opinions/new-ashrae-guidelines-challenge-efficiency-drive/

4.2.4 Green Grid free cooling maps

The Green Grid produce free cooling maps which indicate areas in the world where free cooling (non mechanical cooling) of data centres is feasible. Originally produced in 2009 using ASHRAE data, they have since been reviewed but we believe that these maps will need to be updated regularly in line with changes in average and peak summer temperatures. The Uptime Institute makes a similar observation and notes in addition that a more granular approach may be needed in some locations. We concur with this.

4.3 Other drivers

4.3.1 Sharing best practice internationally

In recent years new international organisations and alliances have emerged within the data centre sphere of activity. These include, but are not limited to, The Climate Neutral Data Centre Pact, the Infrastructure Masons and the Coalition for Disaster Resilience Infrastructure (CDRI). Among these organisations, CDRI is particularly worthy of note because the scope of activity is specifically related to disaster recovery within infrastructure sectors, there is a clear focus and techUK is working actively and productively with CDRI on sharing best practices and knowledge relevant to data centre resilience.

There is also an excellent level of cooperation, information sharing and best practice related discussion at European levels through industry representative bodies like EUDCA and DIGITALEUROPE. Resilience is a regular subject of informal peer to peer activity among national trade associations focused on data centres and cloud services providers. There is an overview of associations and representative bodies on our website, although the rapid emergence of specialist data centre trade bodies means this requires regular updating.

4.3.2 Other internal drivers: Competition and reputation

Although we have covered a range of internal drivers above it is worth iterating the business drivers acting on the sector that encourage resilience here. The overriding purpose of a data centre is to consolidate IT function primarily to improve resilience. Commercial operators compete on resilience and adopt standards both to ensure adherence to best practice and provide third party evidence. A commercial data centre provider's reputation depends upon their ability to provide business continuity for customers and meet service level agreements and other contractual obligations.

4.3.3 External drivers: Policy and regulation

TCFD :

The Task Force on Climate-related Financial Disclosures (TCFD) was created in December 2015 by the Financial Stability Board to improve and increase the reporting of climate-related financial information. Over the past five years it has grown to be important to the data centre and digital infrastructure sectors. In 2017 the TCFD released several recommendations designed to help companies provide better information to investors. Governance, strategy, risk management, and metrics and targets underpin these recommendations. Although it began life as a voluntary scheme, there are movements to make the TCFD recommendations mandatory in several countries, the UK being the first to initiate this.

Set to come into force 6 April 2022, the UK will be the first G20 country to make it mandatory for firms to disclose high quality climate-related financial information. Although we would like to see a much more streamlined regulatory approach – operators are currently required to report the same carbon and energy multiple times through duplicative and often contradictory requirements - we see this largely as a positive development that will improve transparency and confidence. Data centres are capex intensive and are an exceptionally attractive asset class for a wide range of investors (See this guest blog from Vipa Digital on investor attitudes). Reporting obligations will encourage additional scrutiny of data centre assets with respect to emissions, net zero pathways and climate change risks. However, there are shortcomings; this approach will not cover assets within the public sector, which are most in need of scrutiny.

Banking Supply Chain Resilience:

In 2019 the European Banking Authority published guidance on outsourcing arrangements for financial institutions. This incorporated earlier guidance relating to cloud computing services that had been issued in 2017. Simplistically, the intention is to ensure that risk is adequately managed within the supply chain and that financial institutions have adequate control over, liability for, and insight into outsourced digital services, irrespective of where they are in the banking supply chain. Among the concerns was the potential for single points of failure to arise, for instance if core functions and disaster recovery provisions ended up in adjacent locations, or in some other way subject to a single point of failure, because one or both are outsourced without adequate oversight. The Bank of England issued a supervisory statement in March 2021 which will come into effect in March 2022 Again, in principle these measures should improve internal scrutiny and transparency and reduce the likelihood that both primary and back up operations could be compromised by the same event.

Other Policy Drivers

Climate change resilience is rising up the policy agenda and was a key discussion point of COP26. The Government's Net Zero Strategy mentions resilience repeatedly. The Downstream Oil Draft Resilience Bill published June 2021 aims to build resilience in critical infrastructure through the safeguarding of oil-based fuel supplies during crises. Government has also announced a research programme to help protect the country against the most serious weather-related impacts of climate change: heatwaves, flooding, and extreme storms. Extreme heat in buildings is also the target of the funding. These are cited as the areas of most concern in the CCRA3 briefing on ICT and telecoms, published by the UK climate risk group.

4.4 Evidence: What have we learned from recent events?

As we mentioned in our 2016 submission, data centres in the UK have been relatively resilient to severe weather with few, if any, significant outages attributable to climate change, so there is a very meagre evidence base to build a sectoral adaptation strategy upon – or to use as a basis for changing the current approach. In the absence of this, proxy events can be instructive. During 2020 and 2021 the sector, along with everyone else, had to deal with the existential threat of Covid-19: Commercial operators had to maintain adequate levels of staffing whilst managing strict infection control. At the same time demand for data centre services escalated rapidly as business, government and social activity moved online – almost overnight. In addition, organisations accelerated outsourcing programmes, expanded capacity in third party facilities or adopted cloud services to ensure their own business continuity. UK operators handled this escalation in demand and managed infection control without a single outage attributable to the pandemic.

Lessons Learned

In addition to demonstrating that the sector was able to handle unexpected threats robustly and without service interruption, useful lessons have proved instructive over the last two years. These include:

Early action, pre-empting requirements: One of the reasons that data centre operations were not generally compromised by infection was the speed with which operators imposed infection control measures. In many cases, especially for those with operations in multiple jurisdictions, these were implemented at the beginning of 2020, months ahead of the March 23rd lockdown. The sector does not wait for government instructions before rolling out its own risk management plans.

Information sharing: techUK set up a dedicated information source and established weekly operator calls where competitors and colleagues exchanged information and compared notes on their response strategies. There were many positive comments on the way that operators worked together cooperatively for the common good.

Rapid response from Government: As soon as techUK alerted DCMS (the Department for Digital, Culture, Media and Sport) to the need for data infrastructure to be included in the list of key workers, rapid action was taken to advocate internally within government. An emergency team was established in DCMS within 48 hours and has since been strengthened. A candid and productive dialogue was quickly established and has been maintained.

Reduced staffing: Data centres ran successfully with minimal staff numbers on site, and while the intention on this occasion was to reduce potential routes for infection this mimics a situation when not everyone can get to site, for instance during severe flooding. Under COVID, most restrictions were placed on visitors and data centre service providers performed operations on behalf of customers and rearranged maintenance routines to reflect staff availability. Under a climate change event scenario it is more likely that both staff and customers might be unable to access sites. There is also a parallel with staff absence due to illness during COVID: in a flood scenario local staff may be personally affected, potentially to a worse degree than the data centre itself, and be unavailable for work. In fact, exactly this scenario occurred due to flooding in 2017 in Houston, Texas. Some staff who could not return home remained on site in data centres for five days.

Dark site trials and automation: There were several examples of data centres being operated remotely from a sister site to avoid the need for staff onsite following potential contamination by an infected individual. Data centre facilities that are so highly automated that they do not need staff to run them do already exist and are known as Dark Sites, because lighting is unnecessary. Both anecdotal evidence and survey responses (see below) indicate that the pandemic has stimulated greater adoption of, and investment in, automation and remote management technologies by data centre operators. The obvious comparison to climate change is the ability to operate sites remotely, at least temporarily, if access is rendered impossible by flooding.

Increased investment in resilience measures

The Uptime Institute global survey on the impact of extreme weather on data centres (June 2021) reported that the pandemic had led to increased infrastructure investment and increased investment in monitoring and management systems. Uptime observed *"The data suggests that while the pandemic may subside during 2021 and 2022, the spending increase is likely to be sustained. Spending on protective equipment and extra staff may fall back, but capital technology investments, whether in increased automation/monitoring or in site resiliency, may take years to peak, and would then require ongoing operational support. As a result, data centers should be more resilient in the years ahead…"*

Supply chain: the combined impacts of the pandemic and Brexit have transformed data centre supply chains. The pandemic revealed geographical supply chain bottlenecks, for instance in cooling plant, and over the last two years the spot market for commodity supplies has all but disappeared, with delivery times against orders increasing often by multiples. While this reduces reliance on just-in-time business models, it may result in sub-optimal choices and lock-in as a result of market constraints. Inventory control and the supply chain are under greater scrutiny both by operators and government. Nevertheless we consider that the supply chain represents a second order risk that requires more investigation. We have already reported to DCMS and continue to work closely with Government on supply chain issues, including energy cost spikes.

Outsourced services

The data centre sector in the UK relies heavily on outsourcing, from environmental reporting to facilities management. During the pandemic there were concerns relating to contractors who moved between different customer premises, which was considered a risk. However, in an adaptation scenario, the ability to move staff to different sites is more likely to be an advantage.

Unintended consequences of emergency policy measures

Some of the biggest operational threats were external. In particular, the potential for denial of access to facilities due to lockdown. This was largely overcome by quick work by techUK and DCMS to ensure that data centres were included in the list of key workers but we had a very anxious wait before this was confirmed.

Even with key worker status conferred, during the first lockdown over-zealous policing went beyond government requirements and hampered access for operators: data centre workers on legitimate business were stopped by police and asked for documentation, despite that fact that there was no requirement for letters of authority at the time. Police had clearly misunderstood the guidance which required individuals to work at home unless they were unable to, but did not differentiate between essential and non-essential jobs. As a result a lot of urgent work had to be done by the sector, by DCMS, and by hard-pressed operators to provide workers with papers that were technically unnecessary but nevertheless demanded at every opportunity by ill-informed police. Later, the failure by BEIS to allow a test and release system to operate during summer 2021 proved to be a greater threat because data centres were not allowed to manage their own risk. A poor policy decision was compounded by government intransigence. See our comment here.

Incidents of this type add burdens and delay at the worst possible time and undermine confidence in government's ability to respond appropriately to threats. There is the obvious concern that illogical or poorly informed policy decisions may affect other operational activities, such as fuel oil deliveries for emergency generation in the event of a prolonged power outage. We reported in our 2016 submission that site access restrictions can hamper response in a climate change context and anecdotally we understand that response to one of the few flooding incidents in the UK relating to digital infrastructure was delayed because the operator's emergency engineering team were prevented from accessing their own site by emergency services. This may have been necessary for safety reasons but needs review.

The imposition of inappropriate constraints and clumsy mishandling of crises by government therefore has to be considered an operational risk, because it happens too often to ignore.

4.5 Intelligence from industry surveys

Industry analysts and resilience service providers like the Uptime Institute run regular surveys of the data centre sector, often on resilience-related topics. They also provide regular, anonymised information on outages. More recently the Uptime Institute has been surveying the sector on attitudes to climate change risk and trying to get a feel for the level of climate change related outages within the global digital infrastructure community. They have conducted global climate change resilience surveys in 2018, 2019, 2020 and 2021. In 2020 they also published a report on climate change and data centre resiliency. These and other surveys provide very useful insights into sector attitudes and readiness.

We highlight some of these findings below. **However, there are two important caveats to observe.** Firstly Uptime surveys are global and may not be representative of attitudes within the UK. Secondly the respondees are not limited to commercial data centre operators and represent the full suite of business models across the sector, including enterprise and on-premise, where practices may be different. With the exception of financial services, industry analysts concur that enterprise operations tend to lag behind commercial in terms of energy stewardship, resilience, investment and technology adoption, although obviously there are exceptions. The picture is therefore complex and it would be inappropriate to make simple deductions from these figures and apply them without qualification to the UK's commercial data centre market.

Uptime Institute survey responses

The most recent Uptime surveys confirm that awareness of climate change related risks is improving within the sector, which is a positive development, but from a relatively low base, which is less encouraging. In our 2016 submission under ARP we observed that awareness within the sector was variable but that some of this was down to terminology.

Uptime also reported in their 2020 report that in 2018 45% of operators were not adapting to climate change impacts because they felt that they were adequately covered under current risk assessments. Interestingly, after a spate of incidents including wildfires in California and flooding in Houston, this number dropped to 22% in 2019. This suggests that confidence has dropped and that more operators are proactively reviewing these risks.

Uptime welcomed this attitudinal change and summed up these and other improvements in their 2021 journal *"Fires, floods, big freezes and heat waves — coupled with investor activism and threatened legislation mandating greater resiliency to climate change impacts — have driven up awareness of the risks to critical infrastructure. More data center operators are now carrying out both internal and external assessments of their facilities' vulnerability to climate change-related events and long-term changes. A growing proportion is now reacting to what they perceive to be a dramatic increase in risk"*

The table below, reproduced from December 2021 Uptime Journal summarises some of the more recent changes within the sector in terms of recognition of climate change risks Awareness of the potential severity of climate change risks is growing. More operators had reviewed vulnerability and were confident that existing protection was in place. In addition Uptime reported that the number of data centre managers who had formally conducted climate change risk assessments of their data centre was up from 64% to 70%. This still suggests that nearly one third of operators have not conducted a formal risk assessment of the vulnerability of their data centres to climate change.

Operators respond to growing climate threats

Thinking about how your organization responds to the impacts of climate change and extreme weather events, which of these statements do you think is most accurate?	2020 (n=284)	2021 (n=312)	Change
Our management sees a dramatic increase in risk to our data centers due to climate change and is taking steps to improve the resiliency of our critical infrastructure	5%	9%	↑4%
Our management sees an increased risk to our data centers due to climate change and is taking steps to improve the resiliency of our critical infrastructure	28%	26%	↓2%
Our management has determined the vulnerability of our data centers to climate change and believes we have adequate protections in place for the foreseeable future	31%	35%	↑4%
Our management has yet to formally assess the vulnerability of our data centers to climate change	36%	30%	↓6%

UPTIME INSTITUTE CLIMATE CHANGE SURVEY 2020 & 2021 — UptimeInstitute | INTELLIGENCE 2021

Resilience may vary depending on business model

Uptime data also identified different approaches between colocation and enterprise facilities – in their 2020 climate change and data center resiliency report they state that 81% of colocation providers are actively preparing for climate change compared with a 55% industry average and that over half of colocation providers were willing to re-evaluate technology selection compared to one third of other providers. In the same report they also noted that

> *"colocation companies generally operate newer data centers than enterprises and are more likely to be planning new builds. These operators are therefore more likely to have lower risk profiles related to climate change."*

Outsourcing is a potential response to improve resilience

Uptime survey results referred to in the 2020 Data Center Resiliency report suggested that climate change risks are likely to drive organisations to outsource to cloud or colocation providers in the same way that the pandemic catalysed a significant increase in the adoption of third-party services by to improve resilience.

They suggested three reasons for this predicted effect:

- *"The ability to shift the growing risks and costs associated with defending data centers against extreme weather/climate change to the service provider.*
- *The ability of cloud providers to offer distributed resiliency, both within region and between regions — along with an overall high level of security/resiliency.*
- *The lower PUE and strong environmental commitment of large colocation and cloud operators."*

Water Use

Water was another issue raised by Uptime: in their 2020 global survey they reported that half of operators did not monitor water use. However, this is changing. Operators increasingly use standardised metrics like Water Use Effectiveness (WUE) and the Climate Neutral Data Centre Pact's signatories have committed to develop water use metrics and targets and meet them. The Pact is also making available a set of case studies on water conservation measures within data centres that will provide best practice for a number of different scenarios.

The important point to bear in mind, however, is that minimising water use imposes an operational energy burden: you can either optimise a data centre for water use or for energy efficiency, but not both. Therefore, optimising for water should usually take priority in areas that are water stressed.

Other industry sources

Data was also requested from DCIRN, the Data Centre Incident Reporting Network, and while they had concerns about the extent to which forward climate change projections were accommodated within existing industry standards (as we do - see above) they could not identify recent data centre outages specifically attributable to climate change in the UK.

5 Areas of Concern

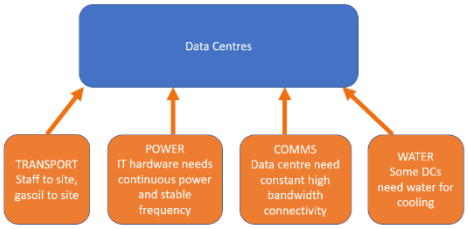
5.1 Growing dependence on digital infrastructure

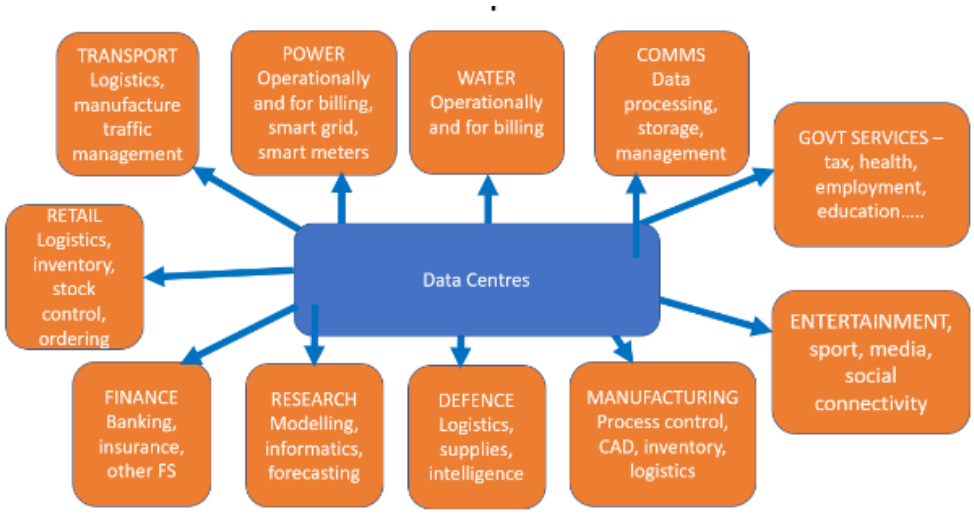
Although the data centre sector has not suffered the scale of problems encountered by other utilities as a result of climate change incidents and/or severe weather and has also demonstrated resilience to significant threats in recent months, there is no room for complacency. This is because, as the pandemic revealed, we are increasingly dependent on data infrastructure for every aspect of our daily lives. The corollary is that, as dependency grows, the consequences of a data centre service interruption become more serious and far-reaching. As referenced before, this familiar meme parodying Maslow's hierarchy of needs makes the point.

5.2 Interdependencies

Data infrastructure is not standalone: data centres, combined with telecommunications networks, comprise our digital infrastructure, which is a network of heavily interconnected and interdependent systems. Data centres are primarily dependent on connectivity – on which data flows rely. Without connectivity, data centres cannot function. Data centres are also heavily reliant on power supplies: computing equipment can only tolerate interruptions in supply of a few milliseconds, so resilient communications networks and an uninterrupted supply of power are critical. Power and connectivity interdependencies are, however, accommodated in data centre risk planning. Any facilities that underpin important functions should have duplicate power and connectivity feeds and embedded emergency generation.

Data centres are also dependent on transport (to get staff and if necessary, emergency generator fuel on site in the event of a power outage) and often water, although this depends on the cooling system of the facility. Data centres are also vulnerable to failures in physical "pinch points" like bridges that carry multiple utilities – communications, electricity and water. We would like to see greater scrutiny of the condition of such pinch points in terms of both identification and condition monitoring.





Interdependency works both ways and it is not just businesses and individuals who depend on data centres, but other infrastructure sectors are heavily reliant on digital infrastructure to function. Government has rightly recognised the potential for cascade failure and this should continue to be a very real concern.

5.3 Resilience risk within on-premise data centres and server rooms

While we provide a collective voice for commercial operators, telecommunications and IT services providers with significant data centre assets, and some financial institutions, these operators only represent part of the UK's data centre estate. Commercial operators have to meet service level agreements (SLAs) and demonstrate resilience capabilities in order to attract business, and this is usually confirmed contractually. As a result, commercial data centre operators are strongly motivated to prioritise business continuity and climate change risks are addressed as part of wider corporate risk assessment and management processes. Outside the commercial sector, the way that a data centre is run is up to the organisation it supports: in some cases it will be highly resilient and run as a business but this is not guaranteed. As Uptime's Annual Outage Report 2020 states:

"A particular issue that has affected financial services, as well as other industries such as air transport and retail, is "asymmetric criticality" or "creeping criticality." This refers to a situation in which the infrastructure and processes have not been upgraded or updated to reflect the growing criticality of the applications or business processes they support"

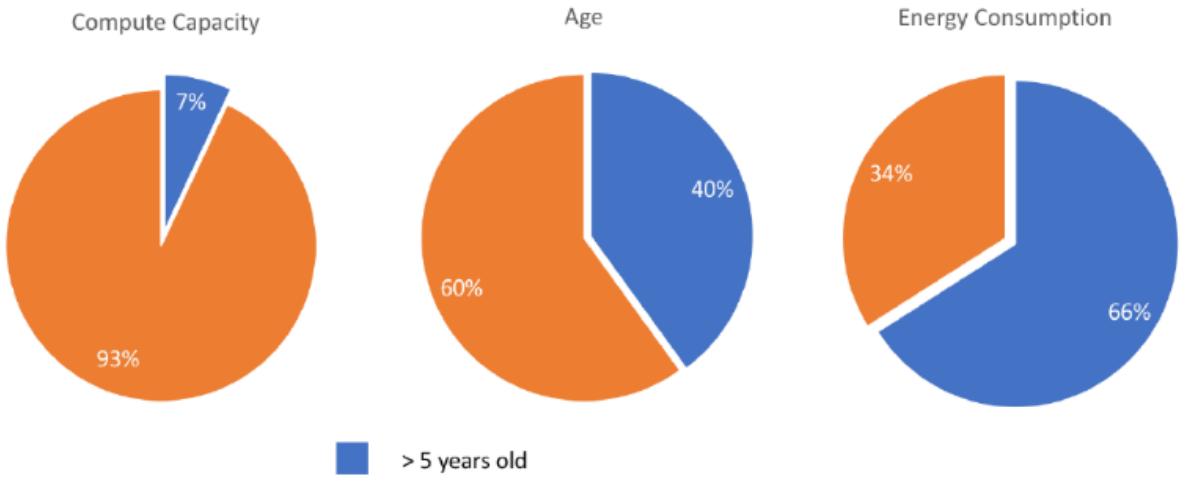
Moreover, excepting regulated sectors like financial services and telecommunications, the commercial drivers and regulatory obligations requiring enterprise data centres to demonstrate resilience are often greatly reduced, or even absent. This suggests that the adoption of industry standards is likely to be lower in enterprise data centres, especially smaller ones. This leads to a three-fold concern:

- Firstly, while we can estimate the number of larger enterprise data centres, it is impossible to judge the number of small on-premise data centres and server rooms. This is particularly problematic in the public sector because, although there are many such facilities, BEIS chose to exempt public sector bodies from the only regulatory instrument that could have shed light on how widespread they are and how they are managed – SECR (streamlined energy and carbon reporting).

- Secondly, the resilience status of these on-premise data centres and server rooms is unknown. Anecdotal evidence from the public sector is not encouraging. The only data we have relates to energy efficiency but this suggests that operational stewardship lags well behind the commercial sector: what is true for energy management may also apply to resilience planning (see below).

- Thirdly, we have no insight into what these on-premise data centres do, and whether the loss of such functions could have a material impact. For small on-premise data centres supporting a business, loss of functionality will affect that business, its customers and its customers' customers. For a local authority data centre the loss of functionality could impact the delivery of a wide range of public services to hundreds of thousands of individual citizens.

We mentioned above that existing data is not encouraging. In 2018 the Eureca project reviewed 337 on-premise public sector data centres in several European countries including the UK. The project revealed widespread shortcomings in energy stewardship: PUE[vii] averaged around 5, utilisation was low and servers were not being upgraded: 40% of servers were over 5 years old and this cohort was consuming 66% of electricity while only delivering 7% of compute. The use of old server assets is not only a very inefficient use of energy, but it has a resilience implication because older servers are unlikely to be designed to meet the latest ASHRAE envelopes of temperature and humidity. This means that the data centre must be kept cooler than for modern servers, which increases both cost and risk. See diagram below.



Ageing servers: source – Eureca project 2018. This illustrates the proportion of old servers in public sector data centres. While their energy consumption is disproportionate to their productivity, older servers are less likely to perform well in expanded ASHRAE temperature and humidity ranges.

Large public sector data centres were reviewed in 2010 by Lord Maude and significant concerns were raised about resilience –one report stated that public sector data centre resilience was "delivered through prayer". The outcomes included major consolidation and outsourcing through the Crown Hosting Joint Venture which has been one of the most successful government projects of all time, saving over £1Bn in costs and transforming transparency, accountability and resilience. While Crown Hosting is available to all public sector bodies, uptake by local authorities and agencies seems to lag behind central government.

In our view, the potential degree of exposure to climate change risks within on-premise data centres should be a matter of concern, especially those operating within the public sector. It looks likely that measures taken to ensure climate change resilience may be variable and in some cases inadequate, or even absent. Lord Maude has recently recommended the creation of a central register of all data centres operated by government and we believe this would be a very helpful step in improving transparency and should be extended as widely as possible across the public sector.

In the meantime we strongly recommend a spot check exercise to review resilience measures in a subset of on premise data centres, ideally across both public and private sectors to ensure that resilience measures are appropriate for the type of data that is managed by the facility.

5.4 Lack of clarity regarding climate change resilience requirements in recognised industry standards and common best practices

In our last submission we suggested the formal inclusion of regular review of flood risk in commonly used availability standards. Although operators do review flood risk we could not find a specific requirement in existing operational standards that would alert practitioners to conduct regular flood reviews. Although there were general requirements in place covering regular reassessment of risk they were not climate change specific: hence the proposal. While there is an important and carefully maintained balance to ensure that standards are flexible and avoid over-prescription while also being comprehensive, we would like greater confidence that climate change resilience is accommodated within the standards and common best practices used by the industry. We therefore recommend further review of design and operational standards and best practices to ensure that this is the case.

Standards adoption within the commercial data centre industry is high: standards provide evidence that best practices are being applied and allow customers to engage with confidence. Standards without certification processes like the EN50600 series are also widely adopted although operators are less likely to cite these. This means that we do not have good intelligence on the actual level of adherence to such

standards. The UK also leads in the implementation of best practices from the EU Code of Conduct which are highly regarded (and in fact form the basis of EN50600). However, due to administrative shortcomings of the Code by the JRC (on which we have provided vocal commentary in the past), operators tend to eschew formal participation through registration so the wide degree of deployment of these best practices within the industry is not recorded and is therefore under-reported. Outside the commercial sector standards adoption is likely to be more variable.
We think more transparency is needed on standards adoption.

5.5 Supply chain vulnerabilities

Potential supply chain issues were identified during the pandemic (see above) but are not restricted to the effects of COVID-19 which revealed a discomforting number of geographical bottlenecks in global data centre supply. We are particularly keen to avoid compounding this with supplier bottlenecks, which is one reason that we firmly reject current efforts by the Environment Agency to dictate which emergency generators operators purchase (see below).

Brexit, combined with COVID-19 have also had a significant impact on supply chain pricing and delivery timescales and as mentioned above, the spot market for many commodities that were previously readily available from stock has now largely evaporated. On the plus side the supply chain is now under much greater scrutiny and, as mentioned above, just-in-time business models are being displaced by forward bidding. However, supply scarcity may force operators into sub-optimal equipment and commodity choices, or supplier lock-in.

We consider that the supply chain represents a second order risk that requires more investigation and we have already reported separately to DCMS on supply chain issues that operators are facing in the UK.

5.6 Skills shortages

The data centre sector in the UK suffers from both short term and long term skills shortages. These are compounded by rapid growth, Brexit imposing constraints on free movement of labour, the lack of STEM qualified students in education and the relative obscurity of the sector. The shortage of appropriately skilled staff is therefore a continuous resilience concern for operators. At sector level a number of initiatives are underway: techUK has been working with other engineering rich sectors that are contracting or where, for other reasons, technical staff are seeking new career opportunities. In some cases this is working well, in others much more work is needed to smooth the technical transition pathway,

Looking further ahead, operators are now partnering with Heathrow University Technical College to develop a bespoke curriculum to prepare young people for careers in digital infrastructure. techUK has engaged operators in an outreach exercise to raise awareness of data centres among schoolchildren and encourage them to adopt STEM subjects.

Nevertheless, technical skills shortages are, and will continue to remain, problematic for data centres and comprise a continued threat that has the potential to undermine other resilience efforts.

5.7 Offshoring

While offshoring is technically outside a UK climate change resilience remit, we consider it a concern because modern computing processes may involve moving workloads to other regions, for instance to access renewable energy sources. Offshore activity supporting a UK customer or consumer base needs to be equally resilient to domestic functions despite being outside UK jurisdiction.

We believe that the resilience of offshored activity should be explored in more detail.

6 What are the barriers to improving resilience?

6.1 External constraints

Power supply: The data centre sector is growing and power supply is lagging well behind demand in data centre development hotspots. This supply issue appears to be largely because, as a heavily regulated entity, National Grid cannot expand capacity speculatively, irrespective of how much evidence there is for growth in demand. This is compromising the development of new data centre capacity across a wide swathe of land to the west of London. Inadequate power supply is a significant concern – adequate digital infrastructure capacity is not only a prerequisite for the UK to function as a digital economy but is also essential for sector resilience. Large cloud providers need to build out sufficient capacity within their availability zones to ensure business continuity in the event of outages or failures at single site level. Data centres of this type are location-specific and by constraining capacity, lack of power has an inevitable impact on resilience.

Unproductive regulatory burdens and distractions: Data centres are poorly understood by government policy makers and as a result tend to be, often accidentally, obliged under regulations that, because they were developed with other targets in mind, are either inappropriate (e.g. Heat Networks Metering and Billing Regulations), or are wholly disproportionate with respect to the compliance burdens compared to the policy outcome (E.g. the EU Emissions Trading Scheme – now UK ETS, and IED, the Industrial Emissions Directive, implemented through EPR). See our compliance checklist which captures just a few of the obligations placed on operators.

As mentioned above, the Environment Agency is currently trying to impose control over choice of emergency plant for the sector on the basis of air quality, irrespective of whether the facility is within an AQMZ (air quality management zone). We consider this inappropriate and an example of regulatory overreach. Generators are for resilience, and while rarely used, need to be selected on the basis of their ability to provide emergency power in the event of grid failure or instability (e.g. unacceptable fluctuation in frequency). At present, issues with generator permitting are challenging business viability and may force sites to operate well under design capacity. Proliferation and over-zealous implementation of burdensome and unproductive compliance requirements represent an abuse of process. They impose a significant toll on staffing, technical skills and divert resource that could be better deployed on core functions like resilience planning. See our statement here.

Regulatory requirements that undermine resilience efforts:
At present, the Government seems minded to "regulate out" diesel including standby generators. The result is that the installation of diesel standby capacity is extremely difficult and exceptionally costly for operators. While it makes sense to reduce reliance on diesel for sustainability and air quality reasons, reliable alternative sources of emergency power with the speed and power intensity of diesel are unavailable and/or tend to involve some degree of resilience compromise. In addition to the issues identified with IED above, operators are adopting biofuel alternatives to diesel for emergency in response to policy and reputational drivers towards net zero. However, there is a lack of long term trial data for many biofuels and some observers have raised concerns about the robustness of the supply chain for some biofuels. We anticipate that confidence will improve over time as more data becomes available.

Potential for government policy or agencies to hamper response
We mentioned above that one of the biggest threats during the pandemic was the potential for government policy, agencies or officials to prevent essential staff accessing sites.

We strongly recommend that government reviews regulatory instruments, including emergency interventions, on the basis of the potential to compromise, hamper or delay resilience related activity.

6.2 Internal constraints and barriers

Growth: the UK data centre sector is growing – fast, to meet demand, which means an increase in the number – or at least the size - of assets needing protection on the one hand, and an ever-growing dependency on digital services on the other - see our report. However, on the positive side, old on-premise activity is being consolidated and migrated to fewer, larger facilities and large enterprise facilities are being decommissioned or sold – sometimes for lease-back - as commercial data centres. This is effectively a process of modernisation that should improve resilience.

Clustering: the UK data centre sector is heavily clustered and a significant proportion of new development at scale is within or close to existing data centre clusters.

Complexity: The complexity of our digital infrastructure can sometimes make it difficult to understand and identify internal interdependencies. Data centres are not standalone but make up part of a network of networks.

Over-reliance on site selection standards

Some observers, including the Uptime Institute, are concerned that the robustness of site selection processes might be leading to a degree of complacency among operators who as a result are not reviewing flood risks of sites often enough. We raised this exact point in our 2016 report and requested that regular flood risk review be a requirement under EN50600.

Water based cooling systems

Data centres adopt different approaches to cooling. While we are seeing more closed-loop cooling systems and plenty of initiatives to reduce water consumption, especially the reliance on potable water, some approaches to data centre cooling, such as evaporative adiabatic systems, consume significant quantities of water and pose a risk if water stress becomes severe. There are already standardised metrics relating to water use (WUE) deployed within the industry and the Climate Neutral Data Centre Pact is currently developing more advanced water metrics where the requirements reflect local water availability. Data centres that are signatories to the pact will need to meet those requirements in future, which will eventually apply to existing facilities as well as new builds. However, this will not happen overnight and, as mentioned above, optimisation for water has an impact on energy, so there is a balance to be struck.

Terminology: Awareness has improved significantly within the sector and while we continue to work to help operators understand climate change risks and find relevant sources of information, it is clear that UK operators understand climate change and correctly view it as a business and operational risk. There is a residual issue relating to terminology, where operators are more likely to categorise climate change risks as severe weather risks, but this is more a question of semantics than of preparing for the wrong type of risk.

7 General Observations on the Process

Differentiating collective from individual responses

The current reporting process makes no formal distinction between collective responses and individual responses, and tends to ask for levels of evidence inappropriate for a collective position. The process needs to be more qualitative in approach. We have included this in our recommendations.

Segregating climate change risks

The data centre sector is resilient by design, with built-in operational redundancy. However, the sector does not segregate climate change risks and prepare for them in a different way: they are handled as part of the risk management process. We view this as appropriate.

Balancing climate change risks with other risks

Climate change risks must be balanced against an array of other, often more immediate, business and operational risks. Improving resilience usually adds a burden in terms of efficiency. Moreover, data centres have locational attributes and it may be that mitigating a certain degree of flood risk may be necessary in order to satisfy other criteria.

Power provisioning impact on interdependency

Many data centre businesses are exploring opportunities to operate at least partially off-grid and become energy prosumers, or source power through alternative means, such as PPAs. Moving to alternative energy sources in pursuit of net zero goals will change existing risks and may present operators with entirely new resilience challenges. This will also, inevitably, change the nature and level of inter-sector dependency.

Absence of Evidence is not Evidence of Absence

Lack of transparency and the absence of an evidence base from which to draw lessons and outline a resilience strategy are bug-bears of this process and as a result the ICT sector at large is repeatedly marked down in terms of its readiness for climate change. Within the data centre sector we would rather not have to rely on a long catalogue of catastrophes to inform our forward thinking. The lack of evidence relating to climate change related failures within the data centre sector should be regarded, at least to some extent, as an indication that climate change risks are being addressed successfully by operators rather than an indication that these risks are being ignored.

Terminology

We take the view that using ICT as a catch-all can be unhelpful and we were pleased to see greater clarity of terminology in recent outputs. Not all ICT is infrastructure. Instead, we think that CCRA and ARP processes should be focused on digital infrastructure. Communications and data centres comprise digital infrastructure. We recommend that the terminology applied to our sector be reviewed.

The transparency issue: balancing disclosure and resilience

Divulging too much detail about risk management within data centres could compromise or erode their resilience to other risks. We therefore do not disclose location details other than indicative distribution, or site-specific details. Data centres perform a critical role and for reasons of security, resilience and business continuity it is also inappropriate to disclose too much operational detail or information about the business functions that individual facilities or operators support. In any case, such information is usually protected under NDA.

We are extremely concerned by the level of granular operational detail already in the public domain due to over-zealous regulatory action by the Environment Agency, most notably relating to environmental permitting. We believe this could compromise resilience and should be removed, and that the practice of publishing detailed information about emergency standby configurations for individual sites should be immediately discontinued in the interests of business – and potentially national - security.

8 Actions for the Data Centre Sector

- Continue to review resilience related standards and metrics adopted by the industry to ensure they are fit for purpose with respect to climate change risks.
- Review levels of adoption of relevant standards and best practices within the sector.
- Continue to monitor relevant incidents, whether they are climate change related or could provide useful proxies.
- Escalate existing international outreach activities through organisations such as CDRI (Coalition for Disaster Resilient Infrastructure) and SDIA (Sustainable Digital Infrastructure Alliance). We work actively with both.
- Continue to work collaboratively with standards bodies and monitor and support the development of resilience-related standards and the inclusion of climate change resilience measures in existing standards.
- Continue to support the harmonisation process for relevant international standards.
- Encourage the adoption of resilience-related standards and other relevant metrics and KPIs within the industry.
- Continue to work with other infrastructure providers on interdependencies to reduce the potential for cascade failures.
- Continue to address short and long term skills shortages through outreach and inter-sector cooperation and information exchange.
- Continue to monitor supply chain constraints and risks in conjunction with DCMS.
- Explore the potential for climate change vulnerabilities in activity that is offshored.
- Continue to lobby for regulatory reform: many regulatory instruments are unfit for purpose and are hampering resilience by imposing operational constraints, or more usually, distracting attention and resource to compliance requirements that deliver no material policy outcome.

9 Our Recommendations

We were disappointed that none of the recommendations we made in our 2016 report appear to have been acted on. We believe that resilience reporting is not a one-way street but should be a consultative, two-way process. We suggest that those relating to data centre resilience are revisited

- We would like to see a more qualitative approach to the ARP assessment process that differentiates individual respondees from sector level respondees.

- We propose a revised definition of ICT as digital infrastructure, comprising telecommunications and data centres. Much of ICT has no infrastructural function.

- We believe that a "resilience impact" filter on proposed policy or regulatory measures is essential to ensure that new proposals will not compromise resilience

- We believe that a retrospective review of resilience impact is necessary for existing policy or regulatory measures to reform those that currently compromise or undermine resilience efforts.

- We recommend a spot check exercise to review resilience measures in a subset of on-premise data centres, ideally across both public and private sectors to ensure that resilience measures are appropriate for the type of data that is managed by the facility.

- We request that the Environment Agency practice of publishing detailed information about emergency standby configurations for individual data centre sites be immediately discontinued.

- We strongly recommend that government recognises the infrastructural importance of data centres when planning emergency responses in order to reduce the scope for protective measures to compromise, hamper or delay resilience related activity.

10 Useful Links and References

All techUK publications relevant to data centres can be found on our **Data Centres Programme Index** but should also be accessible via the direct links below.

Climate change adaptation

- 2016 Report to DEFRA: Core Digital Infrastructure – Climate Change Adaptation and Resilience
- 2020 Submission to CCC: High Winds and Lightning
- Template for Emergency Flood Plan
- Evidence to Environmental Audit Committee: Sector Resilience to Heatwaves, 2018
- Article for Telecoms Journal 2017: Digital Infrastructure – Impacts of Climate Change
- Data Centre Sector Position Statement on Self Isolation Rules, July 2021
- Data Centres and Covid-19: Dossier
- Guest Blog from LINX: Internet Traffic and the Coronavirus
- Changing relationship with Government

Data centres and digital infrastructure generally

- UK Data Centres Overview: the most important industry you've never heard of
- Securing our Digital Future: Data Centre Construction: challenges and opportunities:
- Ten Myths About Data Centres
- Input to Better Regulation Framework Consultation 2021
- Planning: Data Centres and Strategic Planning Policy: Position Statement
- techUK Response to Planning for the Future and techUK FAQ for Planners
- Guest blog on investor attitudes from VIPA Digital
- Compliance obligation shortlist: Compliance Healthcheck

Useful third party sources

- Uptime Institute 2020 Report
- Availability Zone – sample explanation from Microsoft Azure: https://docs.microsoft.com/en-us/azure/availability-zones/az-overview
- CIBSE weather data: https://www.cibse.org/weatherdata
- ASHRAE weather data: https://www.ashrae.org/technical-resources/bookstore/weather-data-center
- European Banking Authority: Revised guidance on outsourcing
- Bank of England Supervisory Statement 2021
- Lord Maude recommendations on cross cutting functions, 2021

Contacts

Emma Fryer
Associate Director
Mob: 07595 410 653
emma.fryer@techuk.org

Adam Young
Programme Manager, Environment
Adam.young@techuk.org

Lucas Banach
Programme Assistant
Tel: 020 7331 2006
Lucas.banach@techuk.org

About techUK

techUK is the UK's leading technology membership organisation, with more than 850 members spread across the UK. We are a network that enables our members to learn from each other and grow in a way which contributes to the country both socially and economically. www.techuk.org

Endnotes

[i] These are known as enterprise data centres because they are there solely to support the business enterprise, as opposed to commercial facilities, which provide data centre and/or cloud services to third parties.

[ii] **What is a data centre?**
A data centre is a building (or self-contained unit) used to house computing equipment such as servers along with associated components such as telecommunications, network and storage systems. A data centre is equipped with a guaranteed power supply and high bandwidth connectivity. Resilience is critical so redundancy (duplication) of networks, power and other infrastructure is common to ensure continuity. Building management controls such as air conditioning maintain the environmental conditions for the equipment within a specified envelope of temperature and humidity, and security systems ensure that the facility and its data remain secure.

[iii] **What is digital infrastructure?**
Our core digital infrastructure is not a single system but multiple systems and networks that interoperate. The three main constituents are fixed line telecommunications (made up of the high capacity and highly resilient core network plus the access network that runs from the exchanges to tens of millions of individual customer premises), mobile telecommunications (that interact with the core network but provide customer coverage through a cellular network) and data centres (that manage, transmit, process and store data for government, businesses, individuals and academia).

[iv] Frontier Economics 2017: The UK Digital Sectors After Brexit: https://www.techuk.org/insights/news/item/10086-the-uk-digital-sectors-after-brexit

[v] https://digitalrealty.box.com/s/bserfy44rne36jxupnnnirdcbwdcvp7f

[vi] Extracted from Review of Standardisation Activities: Energy Management and Viability of Data Centres basted on the edition 3 report of the CEN/CENELEC / ETSI coordination group on green data centres.

[vii] PUE or Power Usage Effectiveness, is the ratio of energy consumed by the facility compared to that consumed by the IT function within it. The greater the PUE, the larger the operational overhead and the less efficient the facility Modern operators strive for PUE as close to 1 as possible. Commercial average is currently around 1.7, which means that commercial facilities are running around six times more efficiently than the public sector data centres reviewed in the Eureca project.