techskills

TEACHERS

Cloud and infrastructure in the classroom: the teacher pack

Activities, answers and the evidence behind every claim

Every figure was checked against its source on 5 November 2026. Free to use in your classroom.

About this pack

This is the printable companion to the Cloud and Infrastructure in the Classroom page on techskills.org. It holds everything you need at the front of a class: how to run each activity, the questions to ask, the numbers the games will show so you can check what students see, cards to print, and the evidence behind every claim.

Cloud and infrastructure has a desk side and a hands-on side, and the Career Matrix marks Data Centre Operations as a hard hat route. The balance in this pack is between activities that start at a desk with a sum and ones that start in the plant room, and between a study route, an apprenticeship at a desk and an apprenticeship on site.

There are thirteen activities. Three need no devices at all, and ten are interactive on this page. Five start at a desk with a sum, four start in the building where the real power, cooling and cabling are, and four suit both, so a teacher with a student who likes to design the system and a teacher with one who likes to keep the hardware running both have something to run on Monday. Timings are our own suggestions from how long each takes to run, not measurements. The teacher pack has the full notes.

Every shop, hall, server, price and failure rate in the games is made up for the lesson, and each game says so. The walk round the building and the role plays use things in your own building: choose the route in advance with your site manager or IT technician, keep students out of plant rooms, electrical cupboards, roofs and anywhere that is not open to them, and do not let students handle batteries, generators or live electrical equipment.

For	Teachers and lecturers in schools, colleges and training providers. Written for England; see the curriculum section
Activities	13
Sources	97, each read on 5 November 2026

The case for teaching it

We could just say that cloud and infrastructure matter and leave it there. But you are the one standing in front of students who will check, so here is what the evidence says, including the awkward parts, because the awkward parts are where the good lessons live.

It is a big, physical thing. The Department for Energy Security and Net Zero estimates that data centres used 4.5 TWh, which is 2% of the electricity consumed from the grid in Great Britain in 2024, and that around three quarters of it was in the South East and London. The same article says there is no universally agreed definition of a data centre, so counts and totals in other sources can differ, and they do: the National Energy System Operator estimates demand at 7.6 TWh from 2.4 GW of connected facilities, where DESNZ's meter-based estimate for 2023 is 4.1 TWh. The government classed UK data centres as critical national infrastructure in September 2024.

The demand for more is the story of the next few years, but be careful with the numbers. The International Energy Agency says data centres used around 1.5% of the world's electricity in 2024 and that consumption could reach around 945 TWh by 2030 in its base case. Ofgem says around 73 GW of data centre projects are in the grid connection queue, and NESO says that pipeline is likely to significantly overstate demand. A request for a connection is not a data centre.

It is also a market with a few very large names. The Competition and Markets Authority found that Microsoft and AWS each had between 30 and 40% of infrastructure as a service in 2024, across the UK and the European Economic Area, and that competition is not working well. Ofcom had said in 2023 that the two had a combined 70 to 80% in the UK in 2022. Those two figures measure different things in different years, so do not put them side by side.

Are there jobs? The honest answer is that we found no government figure for cloud or data centre vacancies. The Uptime Institute's 2025 survey of data centre operators worldwide found that 46% report difficulty finding qualified candidates, and that shortfalls in electrical and mechanical trades continue to be significant. That is an industry survey and not a UK one. The government's AI Growth Zones paper projects more than ten thousand jobs from its programme, which is a projection and not a count, and the Data Centre Covenant expects developers to invest in paid apprenticeships and training places without giving a number.

One more for the staff room. The Department for Science, Innovation and Technology's 2026 business data survey found that 31% of businesses that handle digitised data say they use a public cloud provider to store and process it. The question is about data storage and covers sole traders as well as big firms, so it is not the same as 31% of businesses using any cloud service. And the Technology Code of Practice tells government technology projects to consider public cloud first, which is guidance for government and not a rule for everyone.

The evidence

Fact	Condition	Source
DESNZ says there is no universally agreed definition of a data centre.	Applies to its own estimates, which cover only data centres serving external organisations or individuals (colocation, managed hosting, hyperscale) and exclude enterprise data centres.	Department for Energy Security and Net Zero (DESNZ). Official Statistics in Development, Energy Trends: June 2026, special feature article - Data centre electricity consumption in Great Britain, 2020 to 2024 (Published 30 June 2026)
In Great Britain, data centres accounted for an estimated 4.5 TWh (2%) of the electricity consumed from the grid in 2024.	Great Britain, not UK. Calendar year 2024. Official Statistics in Development (higher uncertainty). Meter-level data matched to a list of data centres; enterprise data centres excluded. The workbook (Table 3) gives 1.79 per cent, which the article rounds to 2%.	Department for Energy Security and Net Zero (DESNZ). Official Statistics in Development, Energy Trends: June 2026, special feature article - Data centre electricity consumption in Great Britain, 2020 to 2024 (Published 30 June 2026)
Around three-quarters (77%) of GB data centre electricity consumption in 2024 occurred in the South East (1.8 TWh) and London (1.7 TWh), and Slough alone accounted for 1.3 TWh (29%).	Same scope as above. The article adds that data centres were about 65% of all grid electricity used in Slough in 2024.	Department for Energy Security and Net Zero (DESNZ). Official Statistics in Development, Energy Trends: June 2026, special feature article - Data centre electricity consumption in Great Britain, 2020 to 2024 (Published 30 June 2026)
Between 2020 and 2024 total data centre electricity consumption from the grid in Great Britain increased by 1.3 TWh (41%).	Same scope as above. Most of the rise is in Slough and Hillingdon.	Department for Energy Security and Net Zero (DESNZ). Official Statistics in Development, Energy Trends: June 2026, special feature article - Data centre electricity consumption in Great Britain, 2020 to 2024 (Published 30 June 2026)
NESO's Future Energy Scenarios 2025 estimate Great Britain's data centre demand at 7.6 TWh from 2.4 GW of connected facilities.	NESO defines data centre demand as demand from dedicated buildings for computing, excluding servers within other commercial buildings. The page gives no year; DESNZ describes it as NESO's figure "for the total data centre electricity consumption across Great Britain in 2023". Based on capacities and assumed utilisation, not meter readings.	National Energy System Operator (NESO), Future Energy Scenarios: Pathways to Net Zero (V.5), section "Commercial | Data centres", page 110 (November 2025)

Fact	Condition	Source
DESNZ says NESO's published figure for 2023 is 7.6 TWh, while DESNZ's meter-based estimate for 2023 is a lower 4.1 TWh, and DESNZ's earlier ND-NEED figure for England and Wales in 2023 was 2.6 TWh.	These three numbers differ because of method and scope: capacity-and-utilisation assumptions (NESO), manual matching of meters to a data centre list (new DESNZ), automated matching to the business rates list (ND-NEED). DESNZ says its new estimate is a "higher 2023 total" than ND-NEED for England and Wales.	Department for Energy Security and Net Zero (DESNZ). Official Statistics in Development, Energy Trends: June 2026, special feature article - Data centre electricity consumption in Great Britain, 2020 to 2024 (Published 30 June 2026)
On 12 September 2024 the Technology Secretary announced that UK data centres are classed as Critical National Infrastructure, the first such designation in almost a decade.	Press release dated 12 September 2024. Source: "since the Space and Defence sectors gained the same status in 2015".	Department for Science, Innovation and Technology (DSIT), GOV.UK press release, Data centres to be given massive boost and protections from cyber criminals and IT blackouts (Published 12 September 2024)
The designation allows the government to support the sector in the event of critical incidents, and includes a dedicated CNI data infrastructure team of senior government officials.	Statement of what the designation "will" do, made at announcement in 2024. It describes intent, not an audit of what has happened since.	Department for Science, Innovation and Technology (DSIT), GOV.UK press release, Data centres to be given massive boost and protections from cyber criminals and IT blackouts (Published 12 September 2024)
The IEA says data centres accounted for around 1.5% of the world's electricity consumption in 2024, or 415 terawatt-hours.	Global, 2024. Energy and AI, Executive summary ("Energy demand from AI" section). Report published 10 April 2025.	International Energy Agency (IEA), Energy and AI, Executive summary (Published 10 April 2025)
The IEA says data centre electricity consumption is set to more than double to around 945 TWh by 2030, slightly more than Japan's total electricity consumption today.	Global. This is the IEA Base Case projection, not a forecast of certainty. "Today" is as of the April 2025 report.	International Energy Agency (IEA), Energy and AI, Executive summary (Published 10 April 2025)
The IEA's Base Case sees global data centre electricity consumption rising to around 1 200 TWh by 2035, and across its cases the 2035 range spans from 700 to 1 700 TWh.	Global. The IEA says uncertainties widen further after 2030. The range comes from its Lift-Off, Headwinds and High Efficiency sensitivity cases.	International Energy Agency (IEA), Energy and AI, Executive summary (Published 10 April 2025)
The IEA estimates that around 20% of planned data centre projects could be at risk of delays unless grid risks are addressed.	Global estimate, conditional on the risks not being addressed. Not a UK figure.	International Energy Agency (IEA), Energy and AI, Executive summary (Published 10 April 2025)

Fact	Condition	Source
Ofgem says that between November 2024 and June 2025 contracted demand in the grid connection queue rose from 41 GW to 125 GW, of which around 73 GW are data centres, compared with a Great Britain peak demand of 45 GW in 2025.	Connection requests, not data centres that exist. The consultation elsewhere puts the 73 GW at approximately 315 data centre projects ranging from 1 MW to 1,500 MW. Consultation published 29 July 2026.	Ofgem, Curate - Demand Connections Reform (consultation) (Publication date 29 July 2026; response deadline 16 September 2026)
NESO says the data centre connection pipeline is likely to significantly overstate peak data centre demand.	Slide deck of 22 September 2026. NESO also records stakeholder feedback that close to 50% of project capacity may be from "highly speculative" or duplicative applications. That is reported feedback, not a NESO measurement.	NESO, Electricity Markets Advisory Council, 22 September 2026 (slide deck, "Data centres" slide) (22 September 2026)
The January 2025 AI Opportunities Action Plan recommended establishing AI Growth Zones to facilitate the accelerated build out of AI data centres.	Plan published 13 January 2025. It is a recommendation to government by the plan's author.	DSIT, GOV.UK, AI Opportunities Action Plan (Published 13 January 2025)
The government's November 2025 paper says timely connections to the electricity grid for AI data centres is the single biggest blocker for establishing AI Growth Zones.	Published 13 November 2025. The same paper projects that its interventions will reduce time to power by up to 5 years. That is a government estimate, not an outcome.	DSIT, GOV.UK policy paper, Delivering AI Growth Zones (Published 13 November 2025)
The draft National Policy Statement says the Compute Roadmap forecasts the UK will need at least 6GW of AI-capable data centre capacity by 2030, and the Data Centre Covenant says the government estimates 6 GW of AI data centre capacity will be required by 2030.	Both published 28 October 2026. This is a government need estimate for AI-capable capacity, a different measure from the 1.6 to 2.4 GW current figures above.	Department for Business, Innovation, Science and Trade. Applies to England, Draft National Policy Statement for Data Centres (document title: Initial Proposal for a National Policy Statement (NPS) for Data Centres) (Published 28 October 2026; consultation closes 8 January 2027); HM Treasury, GOV.UK policy paper, The Data Centre Covenant (Published 28 October 2026)

Fact	Condition	Source
The CMA found that in infrastructure as a service (IaaS) the two largest providers, Microsoft and AWS, each had a share of supply of between 30 and 40% in 2024.	The CMA prints these as ranges in square brackets, [30-40]%. Geographic scope is UK+EEA, not UK only. IaaS only. Do not add the two to a single combined figure; the summary does not give one.	Competition and Markets Authority (CMA), Cloud services market investigation, Summary of final decision (31 July 2025)
In platform as a service (PaaS) the CMA found Microsoft's share had slightly increased to between 20 and 30% and AWS' share had remained stable at between 10 and 20% since 2020, with Google lower in both markets at 5 to 10%.	Ranges, 2024, UK+EEA scope. Google's 5 to 10% is stated for IaaS and PaaS together ("in both IaaS and PaaS markets").	Competition and Markets Authority (CMA), Cloud services market investigation, Summary of final decision (31 July 2025)
The CMA says customers spent £10.5 billion on cloud services in 2024, with spending growing by nearly 30% each year since 2020.	Public cloud infrastructure services (IaaS and PaaS), UK customers, 2024.	Competition and Markets Authority (CMA), Cloud services market investigation, Summary of final decision (31 July 2025)
The CMA found that less than 1% of customers switch cloud provider each year.	Customers of UK cloud services in the investigation. The summary says multi-cloud use is more common than switching but still uncommon for small and medium sized customers.	Competition and Markets Authority (CMA), Cloud services market investigation, Summary of final decision (31 July 2025)
The CMA's final decision, dated 31 July 2025, is that competition is not working well in cloud services, so it made recommendations to address this.	Final decision summary dated 31 July 2025; full report published 1 August 2025. Case page outcome: "Markets - phase 2 adverse effect on competition leading to remedies".	Competition and Markets Authority (CMA), Cloud services market investigation, Summary of final decision (31 July 2025)
The market investigation recommended that the CMA use its digital markets powers to prioritise commencing strategic market status (SMS) investigations into Microsoft and AWS.	Recommendation on the case page (last updated 31 July 2025). Recommendation only. The CMA's March 2026 statement, below, says what happened next.	CMA, GOV.UK, Cloud services market investigation (case page) (Published 5 October 2023, last updated 31 July 2025; case closed 31 July 2025)

Fact	Condition	Source
On 31 March 2026 the CMA said its 2025 investigation found that Amazon and Microsoft have positions of significant market power, and that it would launch an SMS investigation into Microsoft's business software ecosystem, commencing in May.	Press release 31 March 2026. The same release says Microsoft and Amazon took steps on cloud egress fees and interoperability. It does not say the CMA launched an SMS investigation into Microsoft or AWS cloud services themselves. Whether the May 2026 investigation actually started was not checked.	CMA, GOV.UK press release, CMA announces package of actions on business software and cloud services (Published 31 March 2026)
Ofcom said AWS and Microsoft had a combined market share of 70% to 80% in 2022 in UK cloud infrastructure services, and the UK market was worth £7.0 billion to £7.5 billion that year.	Final report "Published 5 October 2023". Combined share of IaaS and PaaS revenues, UK, 2022. Google, the closest competitor, had 5% to 10%. A different measure and year from the CMA's 2024 UK+EEA shares.	Ofcom, Cloud services market study, Final report (Published 5 October 2023)
Ofcom decided in October 2023 to refer the cloud infrastructure market to the Competition and Markets Authority for a market investigation.	Report dated 5 October 2023; the CMA case page shows the investigation opened 5 October 2023.	Ofcom, Cloud services market study, Final report (Published 5 October 2023)
Uptime Institute's 2025 survey found nearly half (46%) of operators report difficulties finding qualified candidates for vacant roles, and talent shortfalls in electrical and mechanical trades continue to be significant.	Global survey of data center operators, April to May 2025 (staffing question n=650). Primary industry research, not a UK government source, not UK specific.	Uptime Institute, Uptime Institute Global Data Center Survey 2025 (Keynote Report 180) (July 2025; survey run online April to May 2025)
The government's AI Growth Zones paper says its interventions will unlock up to £100 billion of additional investment and create more than ten thousand jobs.	Government projection for the AI Growth Zone programme as a whole, 13 November 2025. Not a count of current data centre or cloud jobs.	DSIT, GOV.UK policy paper, Delivering AI Growth Zones (Published 13 November 2025)
The Data Centre Covenant says data centres should create local jobs and invest in paid apprenticeships and training places.	An expectation of developers set by HM Treasury on 28 October 2026. It contains no job numbers.	HM Treasury, GOV.UK policy paper, The Data Centre Covenant (Published 28 October 2026)
techUK says it helped secure key worker status for the data centre workforce.	techUK's own statement about its own lobbying, undated page. Not checked against a government source.	techUK, Data Centres programme (Page carries no date)

Fact	Condition	Source
In the DSIT UK Business Data Survey 2026, 31% of businesses that handle digitised data said they use a public cloud provider such as Amazon Web Services, Microsoft Azure or Google Cloud to store and process data.	Fieldwork October 2025 to January 2026, 4,450 UK businesses in the main stage. This question's base: 2,050 UK businesses that handled digitised data and were randomly assigned to one module. It covers businesses of all sizes including sole traders, and measures data storage infrastructure, not "uses any cloud service". Other types in the same question: 27% on-premises servers, 22% a third party via software or web solution, 19% private cloud (businesses could pick more than one).	DSIT, Official Statistics (fieldwork by Ipsos), UK Business Data Survey 2026 (2025 to 2026 fieldwork) (Published 18 June 2026)
The Technology Code of Practice says to consider using public cloud solutions first, as stated in the Cloud First policy.	Point 5 "Use cloud first". It is guidance for government technology projects, not a rule for private business. Page last updated 29 October 2026.	Government Digital Service, GOV.UK, The Technology Code of Practice (Published 14 July 2021, last updated 29 October 2026)

Your school or college is a customer too

Every school and college depends on infrastructure that someone else may run: the network, the learning platform, the email, the timetable, the lunch payment system. A good first question for your IT technician is which of those run on computers in the building and which on someone else's. It makes a better teaching resource than any case study, because the answers are real and students can see who is responsible.

On who is responsible, the National Cyber Security Centre says there is a joint responsibility for the security and availability of data and workloads in a cloud service, shared between the provider and the customer. It says the provider will almost always be responsible for the availability and physical protections of its servers, and that the customer is always responsible for setting up the services it chooses securely and for deciding what data to store in them. When things go wrong it is not rare: the Department for Science, Innovation and Technology's survey found that 20% of businesses that store data away from their premises said they had been affected by a server or cloud outage in the last 12 months.

A student who likes the hands-on side should know it comes with rules. The Health and Safety Executive says people working on electrical equipment must be competent to do so, and those who cannot show it should not work unless supervised by someone who can. Regulation 14 of the Electricity at Work Regulations 1989 says nobody should work on or so near a live conductor that danger may arise unless, among other tests, it is unreasonable for it to be dead. The HSE also says falls from height are one of the biggest causes of workplace deaths and major injuries. None of that is specific to data centres, and the HSE has no page on data centres that we found.

The evidence

Fact	Condition	Source
The NCSC says there is a joint responsibility for the security and availability of data and workloads in a cloud service, shared between the cloud provider and the customer.	Guidance for cyber security professionals and organisations; page published 10 May 2022, reviewed 7 June 2023. The NCSC adds that the amount each party owns depends on the type of service and how the provider implements it.	National Cyber Security Centre (NCSC), Cloud security shared responsibility model (Published 10 May 2022, Reviewed 7 June 2023, Version 2.1)
The NCSC says the cloud provider will almost always be responsible for maintaining the availability and physical protections of their servers.	"Almost always" is the NCSC's wording. Responsibility for the application on those servers depends on the type of service.	National Cyber Security Centre (NCSC), Cloud security shared responsibility model (Published 10 May 2022, Reviewed 7 June 2023, Version 2.1)
The NCSC says the customer is always responsible for ensuring the service can meet their security needs, securely configuring the services they have chosen, and deciding which data they store in them.	NCSC list headed "You will always be responsible for". The NCSC also recommends handing as much security responsibility as possible to a trusted cloud provider.	National Cyber Security Centre (NCSC), Cloud security shared responsibility model (Published 10 May 2022, Reviewed 7 June 2023, Version 2.1)

Fact	Condition	Source
Of businesses that store data away from their premises, 20% said they were affected by a server or cloud outage or downtime in the previous 12 months.	Same survey; the 59% of businesses handling digitised data that store data away from their premises. Large businesses 58%, medium 28%, small 31%, micro 27%, sole traders 17%.	DSIT, Official Statistics (fieldwork by Ipsos), UK Business Data Survey 2026 (2025 to 2026 fieldwork) (Published 18 June 2026)
HSE says people working on electrical equipment, machinery or installations must be competent to do so, and that people who cannot demonstrate competence should not be allowed to work unless supervised by someone who is.	General HSE guidance on electrical work, not specific to data centres. Page updated 15 April 2025, as shown on the page.	Health and Safety Executive (HSE), Work on electrical equipment, machinery or installations (Updated 15 April 2025)
Regulation 14 of the Electricity at Work Regulations 1989 says no person shall be engaged in work on or so near any live conductor that danger may arise unless, among other tests, it is unreasonable in all the circumstances for it to be dead.	Regulation 14 has further conditions (reasonable to work live, suitable precautions). The page text is the regulation "as made" in 1989.	legislation.gov.uk, The Electricity at Work Regulations 1989, regulation 14 (Work on or near live conductors), as made (Statutory instrument 1989 No. 635, 1989)
HSE says falls from height are one of the biggest causes of workplace fatalities and major injuries.	HSE brief guide INDG401 on the Work at Height Regulations 2005, page updated 7 August 2025. General, not data centre specific. No HSE page on data centres specifically was found. HSE confined spaces guidance (INDG258, page updated 3 February 2026) exists but was not quoted.	HSE, Working at height (INDG401) (Updated 7 August 2025)

Starter: where is the cloud?

Before you say the word cloud, ask students where the last photo they saved is. Give them two minutes to answer. Most will say on their phone. Some will say in the cloud, and a few will say they have no idea. All three are fine answers.

That is the whole opportunity of this subject. The cloud is not a cloud. It is somebody's computers in a building that needs electricity, a way of getting rid of the heat, and cables in and out, and somebody's job is to keep all of that working. Some of those jobs are done at a desk and some are done on site with a hard hat on. This page is built to show both, and neither comes first.

Trace one saved photo from a phone to the building where a copy sits, then walk your own building to find the power, the back-up, the cooling and the cables that every data hall needs too.

What this is

A short trace and a short walk, with no devices. First you follow one saved photo from a phone to the building where a copy sits. Then you walk your own building with the site manager or the IT technician and find the things every data hall needs too.

What to do

- In pairs, think of the last photo you saved or sent. Write every place it passed through on the way, and every place a copy might be kept.
- Share your lists. Add anything missing: a mast or a router, cables, and a building full of computers called a data centre.
- Walk your own building with your teacher and the site manager or the IT technician. Find where electricity comes in, anything that gives back-up power, how rooms are heated and cooled, and where the network cables come in. Look, ask and write. Stay where you are shown.
- Back in the room, write who looks after each of those things, and what would stop working if one of them failed.

What you should notice

The cloud is other people's computers in real buildings, and those buildings need what yours needs: power that does not go off, air that stays the right temperature and cables in and out. Somebody looks after each one.

Time	25 to 35 minutes
You need	Paper and pens, and a short walk led by your site manager or IT technician.
Where	Can be done with no devices. Starts in the plant room and on site.

Students will be able to

- Say that the cloud is other people's computers in real buildings, and name what those buildings need.
- Name where electricity comes into your own building, what gives back-up power and how the rooms are kept comfortable.
- Say who looks after each of those things in your building.

Running it

- Ask students to think of the last photo they saved or sent. In pairs they list every place it passed through on the way, from the phone to the other person's phone, and every place a copy might be kept.
- Share the lists on the board. Add what is missing: the mast or the router, the cables, a building full of computers called a data centre, and the electricity and cooling that building needs.
- Show the page's figures for where UK data centres use their electricity. Ask why they might be in some places and not others.
- Walk your own building, led by the site manager or the IT technician. Find where electricity comes in, anything that gives back-up power, how rooms are heated and cooled, and where the network cables come in. Look, ask and write. Photograph only where you are allowed to.

Questions to ask

- Which parts of the photo's trip are buildings and cables, and which are software?
- If one of those failed, what would stop working?
- Who looks after each part in your school or college, and what happens when it goes wrong?

Watch for

- Students who think the cloud is not a place. It is computers in buildings that someone owns and powers.
- Plant rooms, electrical cupboards, roofs and anything locked. Choose the route in advance with the site team and keep students out of anywhere that is not open to them.
- Counts of UK data centres. The sources disagree because they count different things, and the page shows the range and says why.

The evidence

Fact	Condition	Source
DESNZ says there is no universally agreed definition of a data centre.	Applies to its own estimates, which cover only data centres serving external organisations or individuals (colocation, managed hosting, hyperscale) and exclude enterprise data centres.	Department for Energy Security and Net Zero (DESNZ). Official Statistics in Development, Energy Trends: June 2026, special feature article - Data centre electricity consumption in Great Britain, 2020 to 2024 (Published 30 June 2026)
DESNZ's workbook counts 239 data centres in operation in Great Britain in 2024, up from 221 in 2020.	Workbook Table 1, "Great Britain" row (221, 224, 230, 238, 239 for 2020 to 2024). A data centre counts as operational for a year if at least one electricity meter with consumption could be matched to it; a campus with several data centres counts each separately. Same scope as above (enterprise excluded).	Department for Energy Security and Net Zero (DESNZ). Official Statistics in Development, Energy Trends: June 2026, special feature article - Data centre electricity consumption in Great Britain, 2020 to 2024 (Published 30 June 2026)

Fact	Condition	Source
The draft National Policy Statement says the UK has an estimated more than 220 data centres and around 2 GW of capacity.	Draft policy paper, October 2026, applies to England. Footnote: covers colocation data centres including single-occupancy sites run by hyperscale cloud providers, excludes enterprise data centres, and misses some very small (under 100kW) sites.	Department for Business, Innovation, Science and Trade. Applies to England, Draft National Policy Statement for Data Centres (document title: Initial Proposal for a National Policy Statement (NPS) for Data Centres) (Published 28 October 2026; consultation closes 8 January 2027)
The government's estimate of Great Britain's data centre capacity, defined as the maximum rated IT load of colocation data centres, was a total IT power of 1.6 gigawatts, as of autumn 2024.	Colocation only; enterprise data centres excluded; mostly estimates, regional figures rounded. Answer to a parliamentary question first published 11 March 2025. Largest region: London 1,048 MW.	Department for Business, Innovation, Science and Trade and DSIT (GOV.UK, reproduces answer to Parliamentary Question HL5239), Estimate of Data Centre Capacity: Great Britain 2024 (Published 1 May 2025; the underlying answer was first published 11 March 2025; capacity is estimated as of autumn 2024)
Around three-quarters (77%) of GB data centre electricity consumption in 2024 occurred in the South East (1.8 TWh) and London (1.7 TWh), and Slough alone accounted for 1.3 TWh (29%).	Same scope as above. The article adds that data centres were about 65% of all grid electricity used in Slough in 2024.	Department for Energy Security and Net Zero (DESNZ). Official Statistics in Development, Energy Trends: June 2026, special feature article - Data centre electricity consumption in Great Britain, 2020 to 2024 (Published 30 June 2026)
On 12 September 2024 the Technology Secretary announced that UK data centres are classed as Critical National Infrastructure, the first such designation in almost a decade.	Press release dated 12 September 2024. Source: "since the Space and Defence sectors gained the same status in 2015".	Department for Science, Innovation and Technology (DSIT), GOV.UK press release, Data centres to be given massive boost and protections from cyber criminals and IT blackouts (Published 12 September 2024)
The designation allows the government to support the sector in the event of critical incidents, and includes a dedicated CNI data infrastructure team of senior government officials.	Statement of what the designation "will" do, made at announcement in 2024. It describes intent, not an audit of what has happened since.	Department for Science, Innovation and Technology (DSIT), GOV.UK press release, Data centres to be given massive boost and protections from cyber criminals and IT blackouts (Published 12 September 2024)
The NCSC says the cloud provider will almost always be responsible for maintaining the availability and physical protections of their servers.	"Almost always" is the NCSC's wording. Responsibility for the application on those servers depends on the type of service.	National Cyber Security Centre (NCSC), Cloud security shared responsibility model (Published 10 May 2022, Reviewed 7 June 2023, Version 2.1)

Your notes: where the power, back-up, cooling and cables are in your building

Lab: how many nines?

99% sounds like nearly perfect. Work out how long a year of 1% down really is, then keep the error budget of a made-up shop through a bad month.

What this is

A calculator for how much downtime a target allows, and then a month in the life of a made-up online shop. You see how a promise such as 99.9% turns into minutes, and how an error budget gets used up.

What to do

- Under Pick an availability target, choose 99%. Read how long the service may be down in a day, a month and a year. Then choose 99.9% and 99.99%.
- Before you choose 99.99%, guess its allowance for a year. The answer is 52 minutes 34 seconds.
- In part two, read the shop's budget. Tick A bad update goes out and is rolled back, The main database fails and the spare takes over and The shop's payment provider has an outage, and read how much is left.
- Now tick A security certificate expires and nobody notices. Read what happens to the bar.

What you should notice

At 99% a year of downtime is 3 days 15 hours. At 99.9% it is 8 hours 45 minutes. Each extra nine cuts the allowance to a tenth. With the first three incidents the shop has 2 minutes 42 seconds to spare in a month, and the fourth takes it 35 minutes 18 seconds over.

Time	20 to 30 minutes
You need	A projector for the whole class, or one device per pair.
Where	Interactive on the page. Needs a screen. Starts at a desk with a sum.

Students will be able to

- Work out how much downtime a target such as 99.9% allows in a day, a month and a year.
- Explain why each extra nine is ten times harder to promise.
- Say what an error budget is and what a team does when it has used it up.

Running it

- Show the page on a projector. Start at 90% and ask whether a service that works 90% of the time is good. Step up through 99%, 99.9% and 99.99%, reading out the time allowed in a year each time.
- Before you move from 99.9% to 99.99%, ask the class to predict the new allowance. Then click.
- Move to part two. Read out the shop's target and its budget. Have the class vote on each incident before you tick it.
- Ask what a team should do when the budget is used up, and what it can do while some is left.

Questions to ask

- Why would a company not promise 100%?

- Would a customer on a poor mobile signal notice the difference between 99.99% and 99.999%?
- Who decides how much downtime is acceptable, the engineers or the business?

Watch for

- Students who read 99% as nearly perfect. A year of 3 days 15 hours offline is a long time.
- Sources round these figures differently. Google's table gives 52.6 minutes a year for 99.99%, and the page works out 52 minutes 34 seconds from a 365 day year. Both are right to their own rounding.
- Google's error budgets run over a quarter, and the page uses a 30 day month so the sums are easy to check. A Tier from the Uptime Institute is not a promise of nines, because Uptime says Tiers do not certify an availability percentage.

What each target allows

Target	In a day	In a month of 30 days	In a year of 365 days
90%	2 hours 24 minutes	3 days	36 days 12 hours
99%	14 minutes 24 seconds	7 hours 12 minutes	3 days 15 hours
99.9%	1 minute 26 seconds	43 minutes 12 seconds	8 hours 45 minutes
99.99%	8.6 seconds	4 minutes 19 seconds	52 minutes 34 seconds
99.999%	0.9 seconds	25.9 seconds	5 minutes 15 seconds
99.9999%	0.1 seconds	2.6 seconds	31.5 seconds

The sums use a 365 day year and a 30 day month and ignore leap years. Sources round these figures differently, so a source's figure can differ from the one here by a small amount.

The made-up shop's month

Incident	Unavailable for
A bad update goes out and is rolled back	14 minutes
The main database fails and the spare takes over	1 minute 30 seconds
The shop's payment provider has an outage	25 minutes
A security certificate expires and nobody notices	38 minutes

The target is 99.9% over 30 days, which leaves a budget of 43 minutes 12 seconds. The first three incidents use 40 minutes 30 seconds and leave 2 minutes 42 seconds. With the fourth the shop has used 1 hour 18 minutes and is 35 minutes 18 seconds over. The shop and its incidents are made up.

The evidence

Fact	Condition	Source
Google's SRE book says unplanned downtime is usually expressed as a number of "nines" such as 99.9%, 99.99% or 99.999% availability, and each extra nine is an order of magnitude improvement toward 100%.	Google's own framing for its own services; written 2017.	Google (O'Reilly), sre.google, Embracing Risk (Chapter 3, Site Reliability Engineering book) (2017 (copyright Google, Inc.))
Google's availability table, assuming no planned downtime, allows this much unavailability per year: 99% is 3.65 days, 99.9% is 8.76 hours, 99.99% is 52.6 minutes and 99.999% is 5.26 minutes.	Assumes no planned downtime and a 365-day year. The chapter text rounds 99.99% to 52.56 minutes a year, and AWS and Microsoft pages round the same numbers slightly differently.	Google (O'Reilly), sre.google, Appendix A, Availability Table (Site Reliability Engineering book) (2017 book, with no separate page date)
The SRE book says a system with a 99.99% availability target can be down for up to 52.56 minutes in a year and still stay within its target.	Time-based availability; Google also says for its globally distributed services it usually measures the share of requests that succeed instead.	Google (O'Reilly), sre.google, Embracing Risk (Chapter 3, Site Reliability Engineering book) (2017 (copyright Google, Inc.))
Google's example of request-based availability is a system serving 2.5 million requests a day with a 99.99% daily target, which can serve up to 250 errors and still meet it.	Google's illustration for request success rate, not a time measure.	Google (O'Reilly), sre.google, Embracing Risk (Chapter 3, Site Reliability Engineering book) (2017 (copyright Google, Inc.))
Google says the error budget, based on the service level objective, gives a clear, objective measure of how unreliable the service is allowed to be within a single quarter.	Google's quarterly framing; other organisations use other periods.	Google (O'Reilly), sre.google, Embracing Risk (Chapter 3, Site Reliability Engineering book) (2017 (copyright Google, Inc.))
In Google's example, an SLO of serving 99.999% of queries per quarter means the error budget is a failure rate of 0.001% for that quarter.	Worked example in the book.	Google (O'Reilly), sre.google, Embracing Risk (Chapter 3, Site Reliability Engineering book) (2017 (copyright Google, Inc.))
Google says a user on a 99% reliable smartphone cannot tell the difference between 99.99% and 99.999% service reliability.	Google's argument for not chasing extreme reliability; it is part of why they treat targets as both a minimum and a maximum.	Google (O'Reilly), sre.google, Embracing Risk (Chapter 3, Site Reliability Engineering book) (2017 (copyright Google, Inc.))
Uptime Institute says its data center classifications are divided into four Tiers, which are progressive, and that a Tier IV data center is not automatically better than a Tier II because the levels fit differing business operations.	Uptime's description, undated page. "Progressive" is Uptime's word: each Tier incorporates the requirements of the lower Tiers.	Uptime Institute, Uptime Institute Tier Classification System (Page carries no date)

Fact	Condition	Source
Uptime describes Tier I as protecting against disruptions from human error but not unexpected failure or outage, Tier III as concurrently maintainable with redundant components, and Tier IV as adding fault tolerance to the Tier III topology.	Uptime's own wording. Tier II (redundant capacity components) is described on the same page but not quoted here.	Uptime Institute, Uptime Institute Tier Classification System (Page carries no date)
Uptime Institute says Tier awards do not certify, measure, or guarantee a specific performance outcome such as 99% uptime, five-nines reliability, or any other availability percentage.	From Uptime's publicity guidance to certified clients. Confirmed true on the raw page. The Tiers page itself prints no availability percentages.	Uptime Institute, Publicity Kit (Tier Certification) (Page carries no date)
Uptime Institute's 2025 survey found that half of operators report no outages in the past three years, but one in 10 outages still cause serious or severe disruption.	Global survey of operators, April to May 2025. Outage definitions vary between respondents (Uptime notes this).	Uptime Institute, Uptime Institute Global Data Center Survey 2025 (Keynote Report 180) (July 2025; survey run online April to May 2025)

Your notes

Lab: spread it across two

A made-up shop has three layers and one building. Add spare machines and a second building to reach 99.99%, and find out why more machines in one building stop helping.

What this is

A made-up shop with three layers, a web layer, an application layer and a database, and one building. You add spare machines and a second building and watch the availability, the downtime and the cost. All the numbers are made up.

What to do

- Read the starting figures: one machine in each layer, all in one building. The shop works 96.9329% of the time, which is about 11 days 4 hours down in a year.
- Use Add a copy of Web servers, Add a copy of Application servers and Add a copy of Database. Find the layer that is marked as the weakest and see what a second copy does.
- Give every layer two copies, still in one building, under Where does it all sit? Read the figure: 99.87%. Then add more copies and see that it hardly moves.
- Choose A full copy in a second building. Find the cheapest set-up that reaches 99.99%.

What you should notice

More machines in one building cannot beat the building: with every layer doubled the shop reaches 99.87%, about 11 hours 23 minutes down in a year. A second building changes the answer, and the cheapest way to 99.99% in this game is two copies of every layer in both buildings, which costs 42 units and gives 99.9998%.

Time	25 to 35 minutes
You need	A projector for the whole class, or one device per pair.
Where	Interactive on the page. Needs a screen. Starts at a desk with a sum.

Students will be able to

- Work out the availability of a chain of parts that must all work, and of copies where any one will do.
- Say why the weakest layer is the one to fix first.
- Explain that spare machines only help if they do not fail for the same reason.

Running it

- Start with one machine in each of the three layers in one building. Read out the availability and the downtime a year.
- Ask which single change would help most before anyone clicks. Then let pairs try. The goal is 99.99% for the least cost, and the page shows the weakest layer.
- Move to a second building. Ask why it changes the answer when more machines in the first building did not.
- Close with the note at the foot of the game: the sums assume the failures are independent. Ask the class to name a cause that two machines in one building would share.

Questions to ask

- What did the third copy of a layer buy compared with the second?
- What could take out both buildings at once?
- Is the cheapest way to 99.99% also the way you would choose for a hospital?

Watch for

- Students adding copies everywhere. The weakest layer is the one to fix first, and the page marks it.
- The numbers are made up. They are not real machines or real buildings.
- The formulas are the ones AWS prints, with its own condition: independent. The London heat story and the AWS story are both cases where the parts shared a cause.

Set-ups worked through

Set-up	Works	Down in a year	Cost in units
One machine in each layer, one building	96.9329%	11 days 4 hours	8
Two of each layer, one building	99.87%	11 hours 23 minutes	16
Four of each layer, one building	99.9%	8 hours 45 minutes	32
One of each layer in each of two buildings	99.9059%	8 hours 14 minutes	26
Two of each layer in each of two buildings	99.9998%	53.3 seconds	42

Every machine works 99% of the time and the building works 99.9% of the time. A copy costs 1 unit for the web layer, 2 for the application layer and 5 for the database, and a second building costs 10 units on top of its machines. The weakest layer in the starting set-up is the web servers layer. All of these figures are made up.

The evidence

Fact	Condition	Source
AWS says that where a system has hard dependencies, its availability is the product of the dependent systems' availabilities, so three independent parts each designed for 99.99% give 99.97%.	AWS says "theoretically" and "independent systems". It applies to hard dependencies only (a soft dependency, where the app can cope with the failure, is different).	Amazon Web Services, Availability, AWS Well-Architected Framework, Reliability Pillar (Page carries no date, read 5 November 2026)
AWS says that for independent, redundant components the theoretical availability is 100% minus the product of the component failure rates, so two independent components of 99.9% each give 99.9999%.	Independence is the stated condition (AWS gives "redundant resources in different Availability Zones" as the example). If the copies share a failure cause, the formula overstates availability.	Amazon Web Services, Availability, AWS Well-Architected Framework, Reliability Pillar (Page carries no date, read 5 November 2026)

Fact	Condition	Source
AWS gives a shortcut: if every availability is all nines, add up the count of nines, so two independent three-nines components give six nines.	Only for values made solely of the digit nine.	Amazon Web Services, Availability, AWS Well-Architected Framework, Reliability Pillar (Page carries no date, read 5 November 2026)
Microsoft says a composite SLO value is a product distribution of the contributing factors, and gives 99.95% x 99.99999% = about 99.95% as an example.	Microsoft's page does not state the redundant-copies formula; it says to include external services and shared platform components when they directly affect the customer experience.	Microsoft Learn, Define reliability and recovery targets (Azure Well-Architected Framework, Reliability) (Last updated 20 August 2024)
Microsoft says that for a headline SLA such as Azure Front Door's 99.99%, your design must adhere to specific criteria published in the agreement.	Microsoft's own caution that an SLA covers a narrow defined use case, not every feature of the product.	Microsoft Learn, Define reliability and recovery targets (Azure Well-Architected Framework, Reliability) (Last updated 20 August 2024)
AWS's Amazon Compute SLA (last updated May 25, 2022) says it will use commercially reasonable efforts to make EC2 available in each region with a Monthly Uptime Percentage of at least 99.99%, when all running instances are deployed across two or more Availability Zones.	Region-Level SLA only applies to instances deployed concurrently across two or more AZs in the same region. Check the live page before teaching a figure as current.	Amazon Web Services, Amazon Compute Service Level Agreement (Last updated 25 May 2022)
Under that Region-Level SLA the service credit is 10% of the monthly bill if uptime is below 99.99% but at least 99.0%, 30% if below 99.0% but at least 95.0%, and 100% if below 95.0%.	Credit is a percentage of the monthly bill for EC2 in the affected region, applied to future payments, not a refund; the customer must open a support claim.	Amazon Web Services, Amazon Compute Service Level Agreement (Last updated 25 May 2022)
AWS's Instance-Level SLA for a single EC2 instance is a Monthly (Instance-Level) Uptime Percentage of at least 99.5%, with the same 10%, 30% and 100% credit steps (at 99.0 and 95.0).	The tier boundaries for the instance-level SLA are 99.5%, 99.0% and 95.0%. "Unavailable" for an instance means no external connectivity.	Amazon Web Services, Amazon Compute Service Level Agreement (Last updated 25 May 2022)
AWS's SLA says service credits do not entitle the customer to any refund or other payment from AWS.	Part of the SLA terms.	Amazon Web Services, Amazon Compute Service Level Agreement (Last updated 25 May 2022)
Google Cloud's Compute Engine SLA says Instances in Multiple Zones have a Monthly Uptime Percentage of at least 99.99% (Premium Tier, cloud regions excluding Stockholm), and a single instance of most other machine families at least 99.9%.	Page says "Current", last modified October 14, 2026. The figure is lower for Stockholm and for Standard Tier (99.9% for multiple zones), and 99.95% for a single memory-optimized instance.	Google Cloud, Compute Engine Service Level Agreement (SLA) (Current version, last modified 14 October 2026)

Fact	Condition	Source
Google's Compute Engine SLA gives a Financial Credit of 10% of the monthly bill for uptime from 99.00% up to below 99.99% (Premium Tier), 25% from 95.00% to below 99.00%, and 100% below 95.00%.	For multi-zone instances and load balancing in regions other than Stockholm; credits go to future bills and the customer must notify Google support within 60 days with log files.	Google Cloud, Compute Engine Service Level Agreement (SLA) (Current version, last modified 14 October 2026)
For Instances in Multiple Zones, Google defines Downtime as loss of external connectivity or persistent disk access for all applicable running instances.	This is why one dead VM in a multi-zone set does not count as Downtime under the SLA.	Google Cloud, Compute Engine Service Level Agreement (SLA) (Current version, last modified 14 October 2026)
Microsoft says two or more VMs in an availability set are recommended to meet the 99.95% Azure SLA.	Learn page last updated 2025-12-05; this is Microsoft's summary, not the SLA document, which we could not read. It is not the Azure zones SLA.	Microsoft Learn, Availability options for Azure Virtual Machines (Last updated 5 December 2025)
The 2017 SRE book gave Google Compute Engine's published availability target as "three and a half nines", 99.95%.	Dated (2017) and superseded by the current multi-zone SLA above; use only as a history note.	Google (O'Reilly), sre.google, Service Level Objectives (Chapter 4, Site Reliability Engineering book) (2017 book)

Your notes

Game: lights out

The mains power fails in a made-up hall. A battery has minutes, a generator has to start and take over, and the fuel has to last. Choose the set-up and see whether the hall stays up.

What this is

A made-up data hall that draws 200 kW loses its mains power. A battery carries the load first, a generator has to start and take over, and its fuel has to last until the power comes back. You choose each and see whether the hall stays up.

What to do

- Leave the settings as they are and read the result for a cut of 1 hour. Under Battery size the battery lasts 4 minutes 3 seconds, the generator starts in 30 seconds and the hall stays up.
- Change The cut lasts to 3 minutes, then back to 1 hour, then to 6 hours and 24 hours. Read what ends each one.
- Change Generator takes to start to 10 minutes and read the result. Then change it back and tick The generator fails to start.
- Change Fuel in the tank lasts to 2 hours and set The cut lasts to 6 hours. Find the minute the hall goes dark.

What you should notice

A battery buys minutes and nothing more. A generator has to start before the battery is empty and must have fuel for the rest of the cut. Each one is a way for the hall to go dark.

Time	20 to 30 minutes
You need	A projector or one device per pair. Cards and a stopwatch for the role play version.
Where	Interactive on the page. Needs a screen. Starts in the plant room and on site.

Students will be able to

- Work out how long a battery lasts from its size and the load it carries.
- Say what a battery and a generator each do, and why a hall needs both.
- Explain that a back-up has to start in time and last long enough, and that it is only a plan if it has been tested.

Running it

- Ask students to list, in order, what keeps a data hall running when the mains power fails. Then show the page and the chain: mains, battery, generator.
- Set up the hall with a battery that lasts a few minutes and a generator that starts in seconds. Ask them to predict what happens in a 3 minute cut, a 1 hour cut and a 6 hour cut before you choose each.
- Tick the box for a generator that fails to start, and run the 1 hour cut again. Ask what the hall needed that it did not have.
- No devices? Run it as a role play with a stopwatch. One student is the battery and holds a card with the minutes it lasts, one is the generator and may only stand when its start time is up, and the rest hold cards

for the servers. Say when the mains fails and see who is still standing.

Questions to ask

- Why not just a very big battery?
- What would you test, and how often, to be sure the generator will start?
- What does a hall do about the fuel for a cut of two days?

Watch for

- The battery, generator and times are made up. Vendors quote typical ranges, and the page says the vendor said it.
- Students treating a generator as magic. It has to start, take the load and have fuel. Each is a way to fail.
- Do not let students handle real batteries, generators or live electrical equipment as part of this activity.

How long each battery lasts

Battery	Usable share	Load	Lasts
10 kWh	90%	200 kW	2 minutes 42 seconds
15 kWh	90%	200 kW	4 minutes 3 seconds
25 kWh	90%	200 kW	6 minutes 45 seconds
50 kWh	90%	200 kW	13 minutes 30 seconds

The hall, the batteries, the generator and the times are made up. The sum is the battery's usable energy divided by the load, in hours, turned into minutes.

Role play cards

- Mains power (announces the cut)
- Battery (holds a card with its minutes)
- Generator (may only stand when its start time is up)
- Fuel (holds a card with its hours)
- Servers (every other student)

The evidence

Fact	Condition	Source
The US Department of Energy says UPS systems provide backup power to data centers and can be based on battery banks, rotary machines, fuel cells or other technologies.	July 2024 revision; US government guide.	US Department of Energy (Federal Energy Management Program), Best Practices Guide for Energy-Efficient Data Center Design (Revised July 2024)
The Uptime Institute lists, among the requirements for a Tier I data center, a UPS for power sags, outages and spikes and an engine generator for power outages.	This describes the Tier I requirements; higher tiers add redundancy.	Uptime Institute, Tier Classification System (Page carries no date, read 5 November 2026)

Fact	Condition	Source
Schneider Electric says the generator set serves as long-term power backup (typically days) while the UPS serves as a bridge (typically minutes) until the generators come online.	Vendor source (Schneider Electric, 20 July 2020). Not a standards body or regulator. "Typically".	Schneider Electric (a UPS and generator vendor; vendor blog, not a standards body), Design Considerations of UPS Systems in Data Centers (20 July 2020)
Schneider Electric says generators start and the load is transferred to them in typically 10 to 20 seconds.	Vendor range, "typically", and "theoretically" you only need that much battery runtime. Not universal.	Schneider Electric (a UPS and generator vendor; vendor blog, not a standards body), Design Considerations of UPS Systems in Data Centers (20 July 2020)
Schneider Electric says hyperscale data centers are designed with 1 to 2 minutes of battery runtime, cloud and colocation data centers typically with 5 minutes, and the financial industry typically 10 to 15 minutes.	Vendor view of design practice, 2020; Schneider says the runtime is a customer business decision.	Schneider Electric (a UPS and generator vendor; vendor blog, not a standards body), Design Considerations of UPS Systems in Data Centers (20 July 2020)
Schneider Electric says that in a data center any power interruption of more than 20 milliseconds can cause IT systems to crash.	Vendor claim; used only to explain why a UPS exists.	Schneider Electric (a UPS and generator vendor; vendor blog, not a standards body), Design Considerations of UPS Systems in Data Centers (20 July 2020)
The BEA-RI explains that the site had two redundant 20 kV grid links, fuel generators and rooms of lead-acid batteries, and that the batteries feed the UPS in emergency before the generators are fully running.	Description of this OVH site. Translated from French by TechSkills; the source is in French.	BEA-RI, the French government's industrial-accident investigation bureau (CGEDD), Rapport d'enquete sur l'incendie au sein du centre de stockage de donnees OVH situe a Strasbourg (67) le 10 mars 2021 (report no. MTE-BEARI-2022-005, in French) (Report dated 24 May 2022)

Your notes

Game: Black Friday

A made-up shop has a quiet night, an evening peak and a sudden spike. Meet it with a fixed number of servers or let the number follow demand, and compare what turns away and what it costs.

What this is

A made-up shop has a day of demand: a quiet night, an evening peak and, if you switch it on, a sudden spike. You choose how many servers to run, either a fixed number or a number that follows demand, and see the share of requests turned away and the bill.

What to do

- With A fixed number of servers chosen and 8 servers, read the share turned away and the bill for the day.
- Tick Add a sudden spike at 10:00. Move Servers, all day up until the share turned away is 1% or less. Read the bill.
- Choose Add and remove servers as demand changes. Set Never more than to 30 and A new server takes to No delay, and read the result.
- Change A new server takes to 30 minutes, then 1 hour. Then lower Keep servers busy and read what a lower figure buys and costs.

What you should notice

A fixed fleet of 30 servers turns nothing away on the spike day and costs £72, because most of the servers sit idle for most of it. An elastic fleet with no delay costs £19 and turns nothing away. When a new server takes 30 minutes to start, the same plan turns away about 6.8% of requests: being late costs customers.

Time	25 to 35 minutes
You need	A projector for the whole class, or one device per pair.
Where	Interactive on the page. Needs a screen. Starts at a desk with a sum.

Students will be able to

- Say what elastic means, and what it costs to be slow to start.
- Compare a fixed fleet and an elastic one on requests turned away and on the bill.
- Explain why a fixed number of servers is either wasted for most of the day or too few at the peak.

Running it

- Show the day's demand and ask when the shop is busiest and when it is quiet. Ask how many servers they would buy for the whole day.
- Let pairs choose a fixed number of servers and read the share of requests turned away and the bill.
- Switch to the elastic plan. Ask what happens if a new server takes an hour to start, then add the spike.
- Close by asking what a team gives up for a lower bill, and who pays when a shop turns customers away.

Questions to ask

- When is a fixed number of servers the right choice?
- What would you want to measure before trusting an automatic plan?
- What does 'pay for what you use' mean for a business with an unpredictable day?

Watch for

- NIST says capabilities can be provisioned and released elastically 'in some cases automatically'. Do not tell students the cloud always scales itself.
- The traffic, the servers and the price are made up. A real service has other limits that this game leaves out.
- Students aiming for 0% turned away at any cost. The goal in the game is 1% or less for as little money as they can.

What a few plans give, on the day with the spike

Plan	Requests turned away	Most servers at once	Bill for the day
8 servers, all day	10.4%	8	£19
30 servers, all day	0%	30	£72
Elastic, no delay, at most 30	0%	30	£19
Elastic, 30 minute delay, at most 30	6.8%	30	£19
Elastic, 1 hour delay, at most 30	14.1%	30	£19

Each server answers 50 requests a second and costs £0.10 an hour, which is made up. The busiest ordinary moment is 400 requests a second, and the spike takes it to 1062. The quietest is 40.

The evidence

Fact	Condition	Source
NIST says the cloud model is composed of five essential characteristics, three service models and four deployment models.	SP 800-145, September 2011.	National Institute of Standards and Technology (US Department of Commerce), The NIST Definition of Cloud Computing, Special Publication 800-145 (September 2011)
NIST's five essential characteristics are on-demand self-service, broad network access, resource pooling, rapid elasticity and measured service.	September 2011 wording.	National Institute of Standards and Technology (US Department of Commerce), The NIST Definition of Cloud Computing, Special Publication 800-145 (September 2011)

Fact	Condition	Source
NIST says capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand.	"In some cases automatically": NIST does not say all elasticity is automatic. It adds that to the consumer the capabilities "often appear to be unlimited" (that is appearance, not a guarantee).	National Institute of Standards and Technology (US Department of Commerce), The NIST Definition of Cloud Computing, Special Publication 800-145 (September 2011)
NIST says cloud systems automatically control and optimize resource use with a metering capability, and its footnote says this is typically done on a pay-per-use or charge-per-use basis.	The pay-per-use wording is in a footnote with "typically".	National Institute of Standards and Technology (US Department of Commerce), The NIST Definition of Cloud Computing, Special Publication 800-145 (September 2011)
NIST defines three service models, Software as a Service, Platform as a Service and Infrastructure as a Service, and four deployment models, private, community, public and hybrid cloud.	September 2011; later vendor "as a service" terms are not NIST terms.	National Institute of Standards and Technology (US Department of Commerce), The NIST Definition of Cloud Computing, Special Publication 800-145 (September 2011)

Your notes

Game: keep it cool

Everything a data hall draws that does not reach the computers is a cost. Work out what a point of PUE is worth in energy, carbon and money, then find out how hot is too hot for the equipment.

What this is

Part one: how many more kilowatts a made-up data hall needs, on top of its computers, for cooling and everything else, and what that comes to in energy, carbon and money over a year. Part two: how hot is too hot for the equipment.

What to do

- Under How much do the computers in the hall draw? choose 1,000 kW. Press the button for the average in a 2025 survey of operators and read what everything else draws.
- Press the button for what the government expects of new builds. Read how the figures change, and the saving in the sentence below.
- Move the PUE slider to 1.00. Say why a real hall cannot get there.
- In part two, move Air temperature in the hall up from 22 degrees C. Find where it leaves the recommended range, and where it leaves the range every class of equipment is designed for.

What you should notice

At 1,000 kW of computing and a PUE of 1.54, everything else draws 540 kW, which is 4,730 MWh a year and about 619 tonnes of CO2e at the 2026 UK grid figure. At 1.3 the saving is 2,102 MWh a year. The hall and the price are made up, and PUE leaves out things such as the water a site uses.

Time	25 to 35 minutes
You need	A projector or one device per pair. Cards with the hall's figures for the paper version.
Where	Interactive on the page. Needs a screen. Starts in the plant room and on site.

Students will be able to

- Work out PUE from the power the whole site draws and the power the computers draw.
- Turn a power in kW into energy over a year in kWh, and into carbon using a stated grid figure.
- Say what the recommended temperature range for equipment is, and what happens beyond it.

Running it

- Ask what a building full of computers needs besides power. Take a few answers, then read out what the draft National Policy Statement says about the cooling systems that keep data centres from overheating.
- Show part one. Start at the average PUE from the survey. Ask what a kilowatt of computing needs besides, then move the slider to the new-build expectation and read out the saving.
- Move to part two and slide the temperature up through the recommended range and past it. Ask what the London heat story tells us about a hall built for cooler summers.
- No devices? Give pairs a card: 400 kW of computing and a whole site that draws 600 kW. They work out the PUE, then what the site would draw at 1.3, and the difference in a year.

Questions to ask

- Why is a PUE of 1.0 impossible?
- Who gains when a hall lowers its PUE, the operator, the customer or the grid?
- Why does the page say a 2026 grid figure is not a live reading?

Watch for

- The hall and the price are made up. The PUE figures, the grid figure and the temperature ranges are from sources, and each is listed under the game.
- PUE leaves out things, such as the water a site uses, so a low PUE is not the same as a green site. The page says so and so should you.
- The government's 1.3 is a draft expectation for new builds in England, not a rule for existing halls.

PUE worked through

Computers draw	PUE	Everything else draws	Energy a year	CO2e a year	Cost a year (made-up price)
500 kW	1.54	270 kW	2,365 MWh	310 tonnes	£591,300
500 kW	1.3	150 kW	1,314 MWh	172 tonnes	£328,500
1,000 kW	1.54	540 kW	4,730 MWh	619 tonnes	£1,182,600
1,000 kW	1.3	300 kW	2,628 MWh	344 tonnes	£657,000
2,000 kW	1.54	1,080 kW	9,461 MWh	1,239 tonnes	£2,365,200
2,000 kW	1.3	600 kW	5,256 MWh	688 tonnes	£1,314,000

A year is 8,760 hours. The carbon uses the government's 2026 factor for UK electricity, 130.96 grams of CO2e per kWh, which is a reporting average and not a live reading. The price is made up at 25p a kWh. The hall is made up.

The paper version

A hall's computers draw 400 kW and the whole site draws 600 kW. PUE is 600 divided by 400, which is 1.5. At 1.3 the site would draw 520 kW, which is 80 kW less, and over a year that is 701 MWh.

The evidence

Fact	Condition	Source
Uptime Institute says PUE is calculated by dividing total facility power by the power consumption of IT equipment.	Uptime says PUE was introduced by The Green Grid in 2007. A PUE of 1.0 would mean all power reaches IT equipment (that last point is arithmetic from the definition, not a quote). Uptime notes PUE excludes things such as facility water use or IT efficiency.	Uptime Institute, Uptime Institute Global Data Center Survey 2025 (Keynote Report 180) (July 2025; survey run online April to May 2025)
ISO/IEC 30134-2:2026 is the international standard that defines Power Usage Effectiveness (PUE) as a key performance indicator for data centres.	Edition 2, January 2026, replaces the 2016 edition (withdrawn). The catalogue pages give the scope but not the formula; the Green Grid white paper was not fetched.	ISO catalogue, ISO/IEC 30134-2:2026 Information technology - Data centres key performance indicators - Part 2: Power usage effectiveness (PUE) (Edition 2, January 2026)
In Uptime Institute's 2025 survey, respondents' annual PUE had a weighted average of 1.54, the sixth consecutive year that this headline figure has virtually stood still.	Global survey of data center owners and operators, online April to May 2025, more than 800 owner/operator responses overall; the PUE question had n=681. Self-reported by respondents. Uptime also reports 1.44 for facilities of 20 MW and above, and that many new facilities in high latitudes have PUEs of 1.3 or better.	Uptime Institute, Uptime Institute Global Data Center Survey 2025 (Keynote Report 180) (July 2025; survey run online April to May 2025)
The government expects new data centres to be designed to achieve a Power Usage Effectiveness of 1.3 or better.	Covenant (28 October 2026) says new facilities designed to achieve it; the draft NPS says applicants are expected to demonstrate it. Draft policy for new builds in England, not a rule for existing sites.	HM Treasury, GOV.UK policy paper, The Data Centre Covenant (Published 28 October 2026); Department for Business, Innovation, Science and Trade. Applies to England, Draft National Policy Statement for Data Centres (document title: Initial Proposal for a National Policy Statement (NPS) for Data Centres) (Published 28 October 2026; consultation closes 8 January 2027)
Ofgem says a kilowatt-hour equals the amount of energy generated by an installation with a power capacity of 1kW in an hour.	Ofgem's non-domestic RHI glossary; the definition is general (energy equals power times time).	Ofgem (the GB energy regulator), Key terms explained: Non-Domestic RHI (Page carries no date, read 5 November 2026)

Fact	Condition	Source
Ofgem says a kilowatt is a measure of power, the rate at which energy is transferred or converted.	Same page.	Ofgem (the GB energy regulator), Key terms explained: Non-Domestic RHI (Page carries no date, read 5 November 2026)
The UK government's 2026 conversion factor for UK electricity (electricity generated, supplied to the grid) is 0.13096 kg CO2e per kWh, about 131 grams CO2e per kWh.	DESNZ 2026 factors, published 11 June 2026, for calendar-year 2026 reporting. Generation-based (power station output, grid average, location-based), excludes transmission and distribution losses. It is an annual average factor, not a live reading. The 2026 method changed, as the fact on the method change says, so do not compare it with older years without saying so.	Department for Energy Security and Net Zero (DESNZ), GOV.UK, Greenhouse gas reporting: conversion factors 2026 (full set spreadsheet, UK electricity and Transmission and distribution sheets) (Published 11 June 2026, last updated 31 July 2026)
The 2026 transmission and distribution factor for UK electricity is 0.01299 kg CO2e per kWh, and DESNZ says an electricity consumption figure can be calculated by adding the generation and T&D values.	Consumption-based figure is therefore 0.14395 kg CO2e per kWh (0.13096 plus 0.01299, our addition following DESNZ's instruction, not a number DESNZ prints). DESNZ says company reporting should use the generation figure for Scope 2 and report T&D in Scope 3.	Department for Energy Security and Net Zero (DESNZ), GOV.UK, Greenhouse gas reporting: conversion factors 2026 (full set spreadsheet, UK electricity and Transmission and distribution sheets) (Published 11 June 2026, last updated 31 July 2026)
DESNZ says the 2026 electricity factor was updated to use more recent information on fuel used for power generation, quantities generated and imports and exports from Energy Trends, reducing the time lag.	The major changes report puts the change in the UK electricity CO2e factor at -26% against the 2025 publication and says that reflects two years of grid change rather than one.	DESNZ, GOV.UK, 2026 Government greenhouse gas conversion factors for company reporting: Methodology paper (part of the 11 June 2026 publication)
ASHRAE's TC 9.9 reference card gives 18 to 27 degrees C dry-bulb as the recommended range for equipment classes A1 to A4.	Card is from the Fifth Edition (2021), copyright "2021, 2024 ASHRAE". The card itself only dates 2021/2024; We did not read an older edition, so we do not say how long 18 to 27 has applied. Recommended is guidance for reliability and efficiency; allowable ranges are wider.	ASHRAE (Technical Committee 9.9), 2021 Equipment Thermal Guidelines for Data Processing Environments, ASHRAE TC 9.9 Reference Card (2021 and 2024, Fifth Edition of Thermal Guidelines for Data Processing Environments)

Fact	Condition	Source
ASHRAE's card gives wider allowable dry-bulb ranges of 15 to 32 degrees C for class A1, 10 to 35 for A2, 5 to 40 for A3 and 5 to 45 for A4.	"Allowable" is what the equipment is designed to function within; the card also has humidity, dew point and altitude conditions.	ASHRAE (Technical Committee 9.9), 2021 Equipment Thermal Guidelines for Data Processing Environments, ASHRAE TC 9.9 Reference Card (2021 and 2024, Fifth Edition of Thermal Guidelines for Data Processing Environments)
The draft National Policy Statement says data centres are often associated with using large amounts of water resources, usually consumed through the cooling systems required to keep data centres from overheating.	Draft policy for England, October 2026 (paragraph 417). The same paragraph says newer closed "waterless" cooling can reduce or remove the need for water.	Department for Business, Innovation, Science and Trade. Applies to England, Draft National Policy Statement for Data Centres (document title: Initial Proposal for a National Policy Statement (NPS) for Data Centres) (Published 28 October 2026; consultation closes 8 January 2027)
The National Framework for Water Resources says data centres can require a large amount of water for cooling processes.	England and Wales policy paper, page updated 27 August 2026. The text says large centres can consume millions of litres daily, with use rising in summer, but gives no sourced measurement.	Government policy paper on GOV.UK (Applies to England and Wales), National Framework for Water Resources 2025, chapter 9.4 "Water for data centres and artificial intelligence" (Updated 27 August 2026)
The government's water resources framework says it is experiencing barriers in gaining information about data centre water consumption.	Page updated 27 August 2026. In practice this means there is no official national figure for data centre water use.	Government policy paper on GOV.UK (Applies to England and Wales), National Framework for Water Resources 2025, chapter 9.4 "Water for data centres and artificial intelligence" (Updated 27 August 2026)
In techUK's survey run with the Environment Agency, 51% of surveyed data centres use waterless cooling systems and 64% use less than 10,000 m³ of water per year.	73 sites, mainly commercial colocation facilities in England; on-premise and enterprise data centres not included; submissions were anonymous and voluntary; techUK says it is "not a definitive snapshot". Industry body's report, 18 August 2025.	techUK, with a quote from the Environment Agency's Deputy Director for Water Resources, New report sheds light on commercial data centres water usage in England (press release for "Understanding Data Centre Water Use in England") (18 August 2025)
The government's estimate of Great Britain's data centre capacity, defined as the maximum rated IT load of colocation data centres, was a total IT power of 1.6 gigawatts, as of autumn 2024.	Colocation only; enterprise data centres excluded; mostly estimates, regional figures rounded. Answer to a parliamentary question first published 11 March 2025. Largest region: London 1,048 MW.	Department for Business, Innovation, Science and Trade and DSIT (GOV.UK, reproduces answer to Parliamentary Question HL5239), Estimate of Data Centre Capacity: Great Britain 2024 (Published 1 May 2025; the underlying answer was first published 11 March 2025; capacity is estimated as of autumn 2024)

Your notes

Lab: by hand or by pipeline

Twenty steps, each with a one in a hundred chance of a slip, is not a one in a hundred chance. See how small chances add up, and what a script changes and what it does not.

What this is

A made-up team releases its software over and over. By hand, a person follows a list of steps and any one of them can slip. A script does the same steps the same way, but a mistake in the script is in every release. You compare the two.

What to do

- Leave Steps in a release at 20 and Chance of slipping on one step, by hand at 1%. Read how often a release goes right by hand and by script.
- Move Steps in a release up to 40 and the chance of slipping to 2%. Read what happened to the figure by hand.
- Move Steps in a release down to 5 and the chance of slipping down to 0.5%. Then move Chance the script has a mistake that spoils a release up to 5%, and read what the page says about the two.
- Change Releases a week and read how many releases a year go wrong by hand and by script.

What you should notice

A 1% chance of a slip on each of 20 steps leaves only 81.8% of releases going right first time. At 40 steps and 2% it falls to 44.6%. A script is not automatically better: at 5 steps and 0.5%, doing it by hand goes right 97.5% of the time, and a script with a 5% chance of a mistake loses. At the starting figures a script with a 2% chance of a mistake means about 5 of 260 releases a year go wrong, against about 47 by hand.

Time	20 to 25 minutes
You need	A projector or one device per pair.
Where	Interactive on the page. Needs a screen. Starts at a desk with a sum.

Students will be able to

- Work out the chance that every step of a list goes right when each has a small chance of going wrong.
- Say what automation changes, and why a script can fail every release at once.
- Explain why a team tests its automation before it trusts it.

Running it

- Ask students for a task with a list of steps, such as making a sandwich or logging in to a game on a new device. How often do they miss a step?
- Show the page. Set 20 steps and a 1% chance of a slip, and ask what chance there is that all 20 go right. Take guesses, then read the answer.
- Raise the steps to 40 and the chance of a slip to 2%. Ask what changed.
- Now compare the script. Ask when a script is a better bet than a person, and when it is not, using the sliders.

Questions to ask

- Why is a list of 20 careful steps still likely to go wrong somewhere?
- If a script has a mistake in it, who finds it, and how quickly?
- What would you check before you trusted a script with the whole shop?

Watch for

- Students who say automation is always safer. The game lets the script lose, on purpose.
- The numbers are made up, and the sums assume each step goes wrong independently of the others.
- A student who has not met compound probability. Ask what the chance is of two heads in a row and build up from there.

How small chances add up

Steps	Chance of a slip on each	Chance every step goes right
5	0.5%	97.5%
20	1%	81.8%
20	2%	66.8%
40	1%	66.9%
40	2%	44.6%

At the starting figures (20 steps, 1% a step, 5 releases a week, which is 260 a year), about 47 releases a year go wrong by hand, against about 5 with a script that has a 2% chance of a mistake. Doing it by hand takes about 130 hours a year at 30 minutes a release. All the numbers are made up.

The evidence

Fact	Condition	Source
The standard describes the DevOps engineer role as enabling organisations to get valuable working software out in front of active users and improving the quality of digital services.	This is the page's overview line.	Skills England, DevOps engineer (ST0825) v1.1, Level 4 (Date updated 17/06/2026; approved for delivery 12 March 2020)
The framework says development operations engineers support the development and operation of software through tools, environments and practices.	Last updated 29 August 2025; nine role levels from apprentice DevOps engineer to principal DevOps engineer - management.	Government Digital Service, Development operations (DevOps) engineer, same framework (Published 7 January 2020, last updated 29 August 2025)

Your notes

Game: rack and stack

A rack has 42 units of space and 8 kW of power it may draw. Fill it from a made-up shelf to get as much done as you can, then see the best possible answer.

What this is

A made-up rack has 42 units of space and may draw 8 kW. You fill it from a made-up shelf to get as much work done as you can without going over either limit, then you see the best possible answer.

What to do

- Use Put a Large server in to add large servers. Watch the Space and Power and heat bars as you go.
- Keep going until you hit a limit. Say whether it was space or power, and what is left over.
- Take things out and try another mix. Use Put a GPU server in and Put a Storage shelf in.
- Press Show the best possible answer and compare it with yours.

What you should notice

In this game the best rack does 84 units of work a day using 2 GPU servers and 8 storage shelves. It uses 24 of the 42 units of space and all 8 kW of power. A rack can be full by power with space to spare, and a rack that is full of space may have power left over.

Time	25 to 35 minutes
You need	A projector or one device per pair. Squared paper and scissors for the paper version.
Where	Interactive on the page. Needs a screen. Starts in the plant room and on site.

Students will be able to

- Say that a rack is limited by its space and by its power at the same time.
- Add up the space and power of a set of equipment and say whether it fits.
- Compare their own answer with the best possible and say what made the difference.

Running it

- Show the made-up shelf and the two limits, space and power. Ask which they think will run out first.
- Let pairs fill the rack. They can see the space and the power gauge as they go. The goal is the most work done without going over either limit.
- Ask a pair that fills the space and has power left to say what they could do about it, and a pair that runs out of power with space left.
- Show the best possible answer and ask what the best rack does that most of the class did not.
- No devices? Give each pair a strip of squared paper 42 squares tall and cut-out blocks for each piece of kit, with its power written on it. They keep a running total of power on paper.

Questions to ask

- Why do real halls care about power per rack as much as space?

- What would you buy to use the space that is left when the power runs out?
- In the game, the power a rack draws counts as the heat the room must take away. Why would a hall care about the heat in each rack?

Watch for

- The rack limits, the kit and the work done are made up. A rack unit is a fixed height, but real racks and equipment vary.
- Students who treat the best possible answer as the only right answer. Several racks may tie, and the page says it keeps the first it finds.
- Heavy equipment. Weight is a real limit too, and the game leaves it out.

The made-up shelf

Equipment	Height in rack units	Power	Work a day (made up)	On the shelf
Network switch	1	0.15 kW	2	2
Small server	1	0.35 kW	2	14
Large server	2	0.8 kW	7	8
Storage shelf	2	0.25 kW	3	8
GPU server	4	3 kW	30	3

The rack has 42 units of space and may draw 8 kW. The best possible answer is 84 units of work a day, from 8 x Storage shelf and 2 x GPU server. It uses 24 of the 42 units and 8 of the 8 kW. The limits, the equipment and the work done are made up.

The evidence

Fact	Condition	Source
Wikipedia says a 19-inch rack module has a front panel that is 19 inches (482.6 mm) wide.	Large neutral source, not the standard itself. The page names EIA-310 and IEC 60297 as the formal standards (paywalled, not read).	Wikipedia (large neutral source; it lists EIA-310 and IEC 60297 as the formal standards), 19-inch rack (Last edited 12 October 2026)
Wikipedia says the height of rack-mountable equipment is standardized as multiples of 1.75 inches (44.45 mm), one rack unit or U.	The same page says equipment front panels are made 1/32 inch shorter than the full U count, so a 1U panel is about 1.719 inches (43.7 mm) tall.	Wikipedia (large neutral source; it lists EIA-310 and IEC 60297 as the formal standards), 19-inch rack (Last edited 12 October 2026)
Wikipedia says the two rack posts are separated by a gap of 17.75 inches (450.85 mm), giving an overall rack width of 19 inches.	Large neutral source.	Wikipedia (large neutral source; it lists EIA-310 and IEC 60297 as the formal standards), 19-inch rack (Last edited 12 October 2026)

Fact	Condition	Source
Wikipedia says 42U is a common rack configuration, and that some floor-mounted racks go up to 45U.	"Common", not a standard; the page's lead also says many data centers have taller racks. 42U stated as "common", never as "the" size.	Wikipedia (large neutral source; it lists EIA-310 and IEC 60297 as the formal standards), 19-inch rack (Last edited 12 October 2026)
HSE says falls from height are one of the biggest causes of workplace fatalities and major injuries.	HSE brief guide INDG401 on the Work at Height Regulations 2005, page updated 7 August 2025. General, not data centre specific. No HSE page on data centres specifically was found. HSE confined spaces guidance (INDG258, page updated 3 February 2026) exists but was not quoted.	HSE, Working at height (INDG401) (Updated 7 August 2025)

Your notes

Lab: rent by the hour or by the year

A made-up machine costs 10p an hour or a lower price for promising it all month. Which is cheaper depends on how much you use it, and the crossing point is a share of the month.

What this is

One made-up machine costs 10p an hour if you rent it by the hour, or a lower rate if you promise to keep it for the whole month. You pick what the machine is for and see both bills and the number of hours at which they are the same.

What to do

- Under What is the machine for? choose A test machine for developers. Read both bills.
- Choose An online shop that never closes. Read both bills again.
- Move Discount for promising to keep it all month between 10% and 60%, and watch the crossing point move.
- Find the workload where the answer changes as you move the discount.

What you should notice

At 40% off, promising the whole month costs £44, the same as renting for 438 hours, which is 60% of the month. The test machine used for 176 hours costs £18 by the hour, so renting wins. The shop that never closes costs £73 by the hour, so the promise wins. Which is cheaper depends on how much you use the machine.

Time	20 to 25 minutes
You need	A projector or one device per pair.
Where	Interactive on the page. Needs a screen. Starts at a desk with a sum.

Students will be able to

- Work out a bill for renting by the hour and for committing for a month.
- Find the number of hours at which the two cost the same.
- Say who in a company decides between the two, and what they need to know first.

Running it

- Show the four made-up workloads and ask students to guess which is cheaper to rent by the hour and which to commit to.
- Pick the test machine for developers and read both bills. Then pick the shop that never closes and read them again.
- Move the discount and watch the crossing point move. Ask what share of the month the machine would have to be busy for a 40% discount to pay off.
- Ask who in a company would decide this and what they would need to know before they did.

Questions to ask

- Why do providers offer a lower price for a promise?
- What is the risk of promising a month and then not needing the machine?
- Why might a company have someone whose job is to watch the cloud bill?

Watch for

- The price, the discount and the workloads are made up, and no real provider's prices are used. Real price lists change, so send students to the provider's own page for the day.
- A month is taken as 730 hours, which is 8,760 hours in a year divided by 12.
- The FinOps Foundation says FinOps is not just about saving money. Do not let the lesson become a race to the cheapest bill.

Both bills for each workload

Workload	Hours	By the hour	Promise the month (40% off)	Cheaper
A test machine for developers (8 hours a day, 22 working days)	176	£18	£44	By the hour
A night-time report that runs once a day (3 hours a day, 30 days)	90	£9.00	£44	By the hour
A results website, busy for one week (all day for 7 days)	168	£17	£44	By the hour
An online shop that never closes (every hour of the month)	730	£73	£44	The promise

The machine costs £0.10 an hour by the hour, which is made up. A month is about 730 hours. With 40% off, promising the month costs the same as renting for 438 hours, which is 60% of the month. No real provider's prices are used.

The evidence

Fact	Condition	Source
AWS says On-Demand Instances let you pay for compute capacity by the hour or second (minimum of 60 seconds) with no long-term commitments.	Page undated (footer 2026). Per-second billing applies to the instance types AWS lists; no price is used here.	Amazon Web Services, Amazon EC2 On-Demand Pricing (Page carries no date, read 5 November 2026)
AWS says EC2 Reserved Instances provide a significant discount compared with On-Demand pricing, are best suited to steady-state usage, and are offered on 1-year or 3-year terms.	Page undated. The page also states percentages (for example "up to 72%") that we cannot date; do not print them. Commitment is the trade: a lower rate for committing to steady use.	Amazon Web Services, Amazon EC2 Reserved Instances (Page carries no date, read 5 November 2026)

Fact	Condition	Source
The FinOps Foundation says FinOps is an operational framework and cultural practice which maximizes the business value of technology, enables timely data-driven decision making, and creates financial accountability through collaboration between engineering, finance and business teams.	Page "Updated: November 2026". The wording has changed over time (earlier versions said "cloud"), so check the page before quoting; the page also says FinOps is not just about saving money.	FinOps Foundation, What is FinOps? (Updated November 2026)
The CMA says customers spent £10.5 billion on cloud services in 2024, with spending growing by nearly 30% each year since 2020.	Public cloud infrastructure services (IaaS and PaaS), UK customers, 2024.	Competition and Markets Authority (CMA), Cloud services market investigation, Summary of final decision (31 July 2025)
The CMA found that less than 1% of customers switch cloud provider each year.	Customers of UK cloud services in the investigation. The summary says multi-cloud use is more common than switching but still uncommon for small and medium sized customers.	Competition and Markets Authority (CMA), Cloud services market investigation, Summary of final decision (31 July 2025)

Your notes

Cards: who looks after what?

When a company rents computers from a cloud provider, some jobs stay with the provider and some stay with the customer. Sort six jobs into the right pile, using what the National Cyber Security Centre says.

What this is

A card sort with no devices. A company that rents computers from a cloud provider shares jobs with the provider. You sort six jobs into the provider's, the customer's and those that depend on the kind of service, using what the National Cyber Security Centre says.

What to do

- Write the six jobs on cards: keeping the servers physically protected, keeping the servers available, setting up each service securely, deciding which data goes into it, owning your data and who can log in, and looking after the application that runs on the servers.
- Make three piles: the provider, the customer, and it depends. Sort the cards as a group and write one reason for each.
- Read the National Cyber Security Centre's wording on the page and move any card you would now place differently.
- Pick the one card that changed and say what convinced you.

What you should notice

The provider almost always looks after the availability and physical protection of its servers. The customer is always responsible for setting up the services it chooses securely and for deciding what data goes in them. Looking after the application depends on the type of service.

Time	15 to 20 minutes
You need	Paper cards or sticky notes. No devices.
Where	Can be done with no devices. Suits both a desk and a plant room.

Students will be able to

- Name two jobs the provider almost always does and two the customer always does.
- Say why the answer for some jobs depends on the kind of service.
- Explain why a customer cannot hand everything to the provider.

Running it

- Give each group six cards, one for each job on the page's list. Draw two piles on the board: the provider and the customer, and a third for 'it depends'.
- Groups sort the cards and write one reason for each. Then read out what the National Cyber Security Centre says and let them move cards.
- Ask what a customer would have to check if they hand the application to the provider.

Questions to ask

- Which card changed piles, and what changed your mind?
- Why does the answer depend on the kind of service?
- What does 'shared' not mean?

Watch for

- Students who think the provider looks after everything. The customer is always responsible for configuring the services and choosing what data to store.
- Statements from AWS and Microsoft describe their own models. The NCSC statement is the independent one.
- The page uses the NCSC's own 'almost always' and 'depends'. Keep those words.

The six cards, with the answer

Card	Pile	Why
Keeping the servers physically protected	The provider	The National Cyber Security Centre says the provider will almost always be responsible for the availability and physical protections of its servers.
Keeping the servers available	The provider	Same sentence: the provider will almost always be responsible for the availability of its servers.
Setting up each service you choose securely	The customer	The National Cyber Security Centre says the customer is always responsible for securely configuring the services they have chosen.
Deciding which data goes into it	The customer	The National Cyber Security Centre says the customer is always responsible for deciding which data they store in the services.
Owning your data and who can log in	The customer	Microsoft says that for all cloud deployment types you own your data and identities.
Looking after the application that runs on the servers	It depends	The National Cyber Security Centre says responsibility for the application on those servers depends on the type of service.

The evidence

Fact	Condition	Source
The NCSC says there is a joint responsibility for the security and availability of data and workloads in a cloud service, shared between the cloud provider and the customer.	Guidance for cyber security professionals and organisations; page published 10 May 2022, reviewed 7 June 2023. The NCSC adds that the amount each party owns depends on the type of service and how the provider implements it.	National Cyber Security Centre (NCSC), Cloud security shared responsibility model (Published 10 May 2022, Reviewed 7 June 2023, Version 2.1)

Fact	Condition	Source
The NCSC says the cloud provider will almost always be responsible for maintaining the availability and physical protections of their servers.	"Almost always" is the NCSC's wording. Responsibility for the application on those servers depends on the type of service.	National Cyber Security Centre (NCSC), Cloud security shared responsibility model (Published 10 May 2022, Reviewed 7 June 2023, Version 2.1)
The NCSC says the customer is always responsible for ensuring the service can meet their security needs, securely configuring the services they have chosen, and deciding which data they store in them.	NCSC list headed "You will always be responsible for". The NCSC also recommends handing as much security responsibility as possible to a trusted cloud provider.	National Cyber Security Centre (NCSC), Cloud security shared responsibility model (Published 10 May 2022, Reviewed 7 June 2023, Version 2.1)
AWS says it is responsible for protecting the infrastructure that runs all of the services offered in the AWS Cloud ("security of the cloud"), while customer responsibility ("security in the cloud") is determined by the AWS services the customer selects.	AWS's own marketing and compliance page, undated. AWS's description of its own model, not an independent view.	Amazon Web Services (vendor's own published model), Shared Responsibility Model (Page carries no date)
Microsoft says that for all cloud deployment types you own your data and identities, whereas in an on-premises datacenter you own the whole stack.	Microsoft Learn page "Last updated on 2026-08-24". Microsoft says its article is illustrative and does not modify any agreement.	Microsoft Learn (Azure), Shared responsibility in the cloud (Last updated 24 August 2026)

Your notes

Three true stories about infrastructure

Each story below comes from a document you can open, written by the company or the investigators, and each says only what that document says. Where the documents disagree, we show both. Read them first and decide how to handle them: these are real events that cost real people real work.

A fire in France. In the night of 9 to 10 March 2021 a fire started in the rooms that housed the batteries and the uninterruptible power supply at OVH's data centre in Strasbourg, says the French government's industrial accident investigation bureau, the BEA-RI. The building, called SBG2, was destroyed and nobody was hurt. The bureau says its detection systems worked perfectly, and that what went wrong came after: no automatic extinguishing system, one weak water supply, and generators that started automatically even though the building was on fire, so the site was only made electrically safe around 02h30. It says the cause of the failure in the power equipment was not established when it reported. OVH says in its reply that it is rolling out automatic extinguishing and putting battery rooms apart from the servers. Question for the class: which of the safety systems in this story was designed to help, and which got in the way?

A heatwave in London. On 19 July 2022, Google's own incident report says, a simultaneous failure of multiple, redundant cooling systems in one of the data centres hosting its London zone, combined with extraordinarily high outside temperatures, meant the site could not keep a safe operating temperature. Google powered down part of it on purpose to prevent an even longer outage. It counts 4 hours 8 minutes of cooling failure, 18 hours 23 minutes of impact on cloud services and 35 hours 15 minutes for the slowest recoveries, and says some regional services were hit because it inadvertently changed traffic routing. The Uptime Institute wrote that the heatwave approached or passed the design limits for data centres built before 2019, and that most data centres probably got through it without a major incident. Question for the class: the sums in the spare machines game assume failures are independent. Was this independent?

A bug in an automatic system. On 19 and 20 October 2025 Amazon Web Services' own account says a latent race condition in the automation that manages the DNS records for its DynamoDB database service in its North Virginia region left one record empty, and the automation could not repair it. Anything that depended on that database had trouble, and AWS says its engineers restored the records by 2:25 AM and that clearing the knock-on effects took until 2:20 PM that afternoon. It says servers that were already running stayed healthy, and what failed was starting new ones and reaching the database. It does not say who or what to blame, and it does not mention an attack. Question for the class: what depended on one empty record, and what would you have done to make it depend on less?

A fire in a French data centre, a cooling failure in a London heatwave and a bug that emptied one record at the heart of a huge cloud. Read what each company and each investigator says, and only that.

What this is

Three real events, each read from the document written by the company or the investigators, with a timeline and a plain list of what is safe to say and what is not. Nothing here has been added to what the documents say.

What to do

- Split into three groups, one for each story. Read the timeline and the lines on what is safe and not safe to say.

- In your group write two sentences on what happened, one on what changed afterwards, and one on something you would not say and why.
- Tell your story to the other groups. Listen for the parts that were physical and the parts that were software.
- As a class, decide which story shows spare parts that failed for a shared reason.

What you should notice

In France, the investigators say a fire started in the rooms with the batteries and the UPS, and that the cause was not established. In London, Google says several redundant cooling systems failed at the same time in a heatwave. At AWS, the company says an empty record in an automatic system caused the trouble. In the second and third, the spare parts failed for a shared reason.

Time	30 to 40 minutes
You need	Paper. A printout of the page, or a projector. No devices needed.
Where	Can be done with no devices. Suits both a desk and a plant room.

Students will be able to

- Say in two sentences what happened in each of the three stories, using only what the documents say.
- Name one thing in each story that was a physical failure and one that was a software failure.
- Say which of the three shows spare parts that failed for a shared reason.

Running it

- Split the class into three groups, one for each story. Give each group the story's timeline and the lines marked safe to say and not safe to say.
- Each group prepares two sentences on what happened, one on what changed afterwards and one on what they would not say, and why.
- Each group tells its story to the others. Ask the listeners to say which parts were physical and which were software.
- Finish by asking which story shows spare parts that failed for a shared reason.

Questions to ask

- What did each company or investigator say, and what did they not say?
- Which of these could happen in a school or college?
- What would you want to be true of a company's report on its own failure?

Watch for

- Blame. The documents say what they say, and the page attributes each statement. Do not add a cause the document does not give.
- The OVH fire has no official cause in the report, and its times differ between OVH and the investigators. Say both.
- The London cooling story names Google only because Google wrote the report. No other company is named.
- Treat these as real events that frightened real people and cost them real work.

The evidence: the fire in France

Fact	Condition	Source
The BEA-RI says a fire started in the night of 9 to 10 March 2021 at OVH's data centre in the Port du Rhin, Strasbourg, beginning at the ground floor of building SBG2 and spreading quickly through that whole building.	Official report. Alarm in the guard post at 00h35 per the BEA-RI; OVH's own page says "at 00:47". The 12 minute gap is unexplained in the documents. Translated from French by TechSkills; the source is in French.	BEA-RI, the French government's industrial-accident investigation bureau (CGEDD), Rapport d'enquete sur l'incendie au sein du centre de stockage de donnees OVH situe a Strasbourg (67) le 10 mars 2021 (report no. MTE-BEARI-2022-005, in French) (Report dated 24 May 2022)
The BEA-RI timeline: alarm at 00h35, building evacuated 00h39, fire service called 00h42, fire service on site 00h59, OVH cut the emergency power of SBG2 at 01h13, the grid operator cut power upstream at 01h50, the fire was under control about 06h45, and it was put out at 10h02 with the operation finished at 18h13.	Paris local time, from the BEA-RI. Translated from French by TechSkills; the source is in French.	BEA-RI, the French government's industrial-accident investigation bureau (CGEDD), Rapport d'enquete sur l'incendie au sein du centre de stockage de donnees OVH situe a Strasbourg (67) le 10 mars 2021 (report no. MTE-BEARI-2022-005, in French) (Report dated 24 May 2022)
The BEA-RI says the fire caused no victims and no injuries, but destroyed building SBG2 completely and partly destroyed SBG1 (4 of its 12 rooms).	"4 salles sur 12" for SBG1 comes from the report's synthesis and its text. Translated from French by TechSkills; the source is in French.	BEA-RI, the French government's industrial-accident investigation bureau (CGEDD), Rapport d'enquete sur l'incendie au sein du centre de stockage de donnees OVH situe a Strasbourg (67) le 10 mars 2021 (report no. MTE-BEARI-2022-005, in French) (Report dated 24 May 2022)
The BEA-RI says the fire began in the rooms that house the batteries and the UPS (the "energy rooms"), which had fire detection but no automatic extinguishing system.	Two start points almost at the same moment, one on batteries and one on a UPS (called ASI2 in the report). Translated from French by TechSkills; the source is in French.	BEA-RI, the French government's industrial-accident investigation bureau (CGEDD), Rapport d'enquete sur l'incendie au sein du centre de stockage de donnees OVH situe a Strasbourg (67) le 10 mars 2021 (report no. MTE-BEARI-2022-005, in French) (Report dated 24 May 2022)
The BEA-RI says the precise causes of the starts in those machines were not known at the date of the report because they were the subject of a court-ordered expert examination.	The BEA-RI floated hypotheses (liquid or humidity near the nearby cooling system, a maintenance operation done that morning, operation outside normal limits) and says it cannot say which. We did not find the later expert report. Translated from French by TechSkills; the source is in French.	BEA-RI, the French government's industrial-accident investigation bureau (CGEDD), Rapport d'enquete sur l'incendie au sein du centre de stockage de donnees OVH situe a Strasbourg (67) le 10 mars 2021 (report no. MTE-BEARI-2022-005, in French) (Report dated 24 May 2022)

Fact	Condition	Source
The BEA-RI says that, according to public sources, about 3.6 million websites (464,000 domain names) were unavailable at the peak of the crisis, close to 18% of the active IP addresses attributed to OVH in the previous two weeks.	The BEA-RI is citing "public sources", not its own count. Translated from French by TechSkills; the source is in French.	BEA-RI, the French government's industrial-accident investigation bureau (CGEDD), Rapport d'enquete sur l'incendie au sein du centre de stockage de donnees OVH situe a Strasbourg (67) le 10 mars 2021 (report no. MTE-BEARI-2022-005, in French) (Report dated 24 May 2022)
OVHcloud says it was producing servers to compensate for the destruction of the 14,046 servers that were in operation in the SBG2 data centre.	OVH's own figure, April 2021.	OVHcloud, Strasbourg datacentre: latest information (press release, updated) (Press release of 18 March 2021, updated to 23 April 2021)
OVHcloud says 120,000 services were fully or partially affected and, as of its 23 April 2021 update, 113,000 of them had been fully recovered.	OVH's own words in the 23 April 2021 update; "services" is OVH's unit, not "customers".	OVHcloud, Strasbourg datacentre: latest information (press release, updated) (Press release of 18 March 2021, updated to 23 April 2021)
The BEA-RI explains that the site had two redundant 20 kV grid links, fuel generators and rooms of lead-acid batteries, and that the batteries feed the UPS in emergency before the generators are fully running.	Description of this OVH site. Translated from French by TechSkills; the source is in French.	BEA-RI, the French government's industrial-accident investigation bureau (CGEDD), Rapport d'enquete sur l'incendie au sein du centre de stockage de donnees OVH situe a Strasbourg (67) le 10 mars 2021 (report no. MTE-BEARI-2022-005, in French) (Report dated 24 May 2022)
The BEA-RI says OVH chose not to fit any of the five Strasbourg buildings with an automatic fire protection system.	OVH's reply says the absence of an automatic extinguishing system was not a regulatory non-conformity or a fault of the operator, and that no guide for data centre fire safety existed when SBG2 was designed. Translated from French by TechSkills; the source is in French.	BEA-RI, the French government's industrial-accident investigation bureau (CGEDD), Rapport d'enquete sur l'incendie au sein du centre de stockage de donnees OVH situe a Strasbourg (67) le 10 mars 2021 (report no. MTE-BEARI-2022-005, in French) (Report dated 24 May 2022)
The BEA-RI says the public fire service had only one hydrant for this operation, with a flow below 60 cubic metres an hour.	The fire boat Europa arrived to supply water (the report gives about 3h00). Translated from French by TechSkills; the source is in French.	BEA-RI, the French government's industrial-accident investigation bureau (CGEDD), Rapport d'enquete sur l'incendie au sein du centre de stockage de donnees OVH situe a Strasbourg (67) le 10 mars 2021 (report no. MTE-BEARI-2022-005, in French) (Report dated 24 May 2022)

Fact	Condition	Source
The BEA-RI says the generators were tied only to supplying electricity and start automatically despite the fire, so OVH had to disable them so they would not start when power was cut.	Part of the report's explanation of why electrical isolation took so long. Translated from French by TechSkills; the source is in French.	BEA-RI, the French government's industrial-accident investigation bureau (CGEDD), Rapport d'enquete sur l'incendie au sein du centre de stockage de donnees OVH situe a Strasbourg (67) le 10 mars 2021 (report no. MTE-BEARI-2022-005, in French) (Report dated 24 May 2022)
The BEA-RI estimates the site was electrically secured only from 02h30, when water was put on the fire and SBG2 was already fully ablaze.	BEA-RI estimate. Reasons given include no easily reachable site-wide cut-off and the grid operator having to cut power upstream. Translated from French by TechSkills; the source is in French.	BEA-RI, the French government's industrial-accident investigation bureau (CGEDD), Rapport d'enquete sur l'incendie au sein du centre de stockage de donnees OVH situe a Strasbourg (67) le 10 mars 2021 (report no. MTE-BEARI-2022-005, in French) (Report dated 24 May 2022)
The BEA-RI says the detection systems (aspirating and optical) worked perfectly.	The failure was in what happened after detection (spread, water, power isolation), not in detection. Translated from French by TechSkills; the source is in French.	BEA-RI, the French government's industrial-accident investigation bureau (CGEDD), Rapport d'enquete sur l'incendie au sein du centre de stockage de donnees OVH situe a Strasbourg (67) le 10 mars 2021 (report no. MTE-BEARI-2022-005, in French) (Report dated 24 May 2022)
The BEA-RI says SBG2's design (modular, 6 levels, wooden floors treated with fire-retardant paint, open to outside air for cooling) played a role in how the fire spread.	The report says the structure and floors had one hour fire treatment; a design chosen to cool with outside air. Translated from French by TechSkills; the source is in French.	BEA-RI, the French government's industrial-accident investigation bureau (CGEDD), Rapport d'enquete sur l'incendie au sein du centre de stockage de donnees OVH situe a Strasbourg (67) le 10 mars 2021 (report no. MTE-BEARI-2022-005, in French) (Report dated 24 May 2022)
The BEA-RI recommended that OVH apply the lessons to rebuilding and new data centres, set up and rehearse emergency procedures for electrical shutdown at Strasbourg, and audit all its installations for fire vulnerability.	It also made a recommendation to the French ministry (DGPR) about rooms housing battery chargers. Translated from French by TechSkills; the source is in French.	BEA-RI, the French government's industrial-accident investigation bureau (CGEDD), Rapport d'enquete sur l'incendie au sein du centre de stockage de donnees OVH situe a Strasbourg (67) le 10 mars 2021 (report no. MTE-BEARI-2022-005, in French) (Report dated 24 May 2022)

Fact	Condition	Source
OVH says it launched an "Hyper Resilience" programme to generalise automatic extinguishing on sites without it, raise fire resistance from 60 to 120 minutes in zones with no specific rule, and put energy and battery rooms in separate rooms from the data centre.	OVH's own undertakings in its 25 July 2022 letter. The "separate rooms" rule is "systematic for new sites and when possible for sites already operating". The same letter says it will no longer build with that cooling model and is dismantling RBX4 at Roubaix. Translated from French by TechSkills; the source is in French.	OVH (published on the BEA-RI site), OVH letter to the BEA-RI with its observations and follow-up to the recommendations (in French) (Letter dated 25 July 2022)
OVHcloud lists, among the actions after the fire, offering support for customers in implementing their Disaster Recovery Plan mechanisms.	OVH's own list of priorities in its 2021 press release; it does not say here how many customers had working plans.	OVHcloud, Strasbourg datacentre: latest information (press release, updated) (Press release of 18 March 2021, updated to 23 April 2021)
OVHcloud said in April 2021 that it had decided to rebuild the SBG1, SBG2 and SBG4 data centres on the basis of new standards.	OVH statement, 2021. OVH's letter adds that its new SBG5 data centre opened on 6 July (2022).	OVHcloud, Strasbourg datacentre: latest information (press release, updated) (Press release of 18 March 2021, updated to 23 April 2021)

The evidence: the heatwave in London

Fact	Condition	Source
Google says that on Tuesday 19 July 2022 at 06:33 US/Pacific a simultaneous failure of multiple, redundant cooling systems in one of the data centers that hosts the zone europe-west2-a impacted multiple Google Cloud services.	Google's wording; all times US/Pacific (06:33 Pacific is 14:33 UK time, our conversion).	Google Cloud Service Health, Multiple Cloud products experiencing elevated error rates, latencies or service unavailability in europe-west2 (Incident report) (Incident report posted 29 July 2022; the incident ran from 19 July 2022 to 20 July 2022, US/Pacific time)
Google's root cause statement says one of the data centers that hosts europe-west2-a could not maintain a safe operating temperature due to a simultaneous failure of multiple, redundant cooling systems combined with extraordinarily high outside temperatures.	Google gives no temperature figure and does not explain why the cooling systems failed; it says its engineers are "actively conducting a detailed analysis of the cooling system failure".	Google Cloud Service Health, Multiple Cloud products experiencing elevated error rates, latencies or service unavailability in europe-west2 (Incident report) (Incident report posted 29 July 2022; the incident ran from 19 July 2022 to 20 July 2022, US/Pacific time)

Fact	Condition	Source
Google says engineers decided at 10:05 to power down servers in the impacted data center to prevent an even longer outage and further impact to infrastructure in the zone, and that the cooling system was repaired at 14:13.	Google's wording, US/Pacific.	Google Cloud Service Health, Multiple Cloud products experiencing elevated error rates, latencies or service unavailability in europe-west2 (Incident report) (Incident report posted 29 July 2022; the incident ran from 19 July 2022 to 20 July 2022, US/Pacific time)
Google gives three durations: 4 hours 8 minutes for the cooling system failure impact, 18 hours 23 minutes for impact on Cloud services (10:05 on 19 July to 04:28 on 20 July), and 35 hours 15 minutes for the long tail of some Compute Engine instances.	Google's own arithmetic. The incident closed at 21:20 on 20 July.	Google Cloud Service Health, Multiple Cloud products experiencing elevated error rates, latencies or service unavailability in europe-west2 (Incident report) (Incident report posted 29 July 2022; the incident ran from 19 July 2022 to 20 July 2022, US/Pacific time)
Google says Compute Engine terminated all VMs in the impacted data center, about 35% of the VMs in the europe-west2-a zone, and earlier terminated 42% of the Preemptible VMs across the europe-west2 region.	Google's figures. Different percentages refer to different groups (zone, region, preemptible only).	Google Cloud Service Health, Multiple Cloud products experiencing elevated error rates, latencies or service unavailability in europe-west2 (Incident report) (Incident report posted 29 July 2022; the incident ran from 19 July 2022 to 20 July 2022, US/Pacific time)
Google says approximately 38% of Persistent Disk volumes in zone europe-west2-a were unavailable from 09:13 on 19 July.	Zone-level figure.	Google Cloud Service Health, Multiple Cloud products experiencing elevated error rates, latencies or service unavailability in europe-west2 (Incident report) (Incident report posted 29 July 2022; the incident ran from 19 July 2022 to 20 July 2022, US/Pacific time)
Google says that although regional services are designed to survive the failure of a single zone, a number were affected because at the start it inadvertently modified traffic routing for internal services to avoid all three zones in the europe-west2 region.	Google attributes this to itself in its own report, corrected at 12:35 on 19 July. Do not add blame beyond what Google writes.	Google Cloud Service Health, Multiple Cloud products experiencing elevated error rates, latencies or service unavailability in europe-west2 (Incident report) (Incident report posted 29 July 2022; the incident ran from 19 July 2022 to 20 July 2022, US/Pacific time)
Google says between 07:18 and 08:54 on 19 July, ReadObject availability for buckets in europe-west2 dropped to approximately 86%, impacting 24% of customer projects in the region.	Google says a legacy placement algorithm let three storage clusters be affected by one data centre issue.	Google Cloud Service Health, Multiple Cloud products experiencing elevated error rates, latencies or service unavailability in europe-west2 (Incident report) (Incident report posted 29 July 2022; the incident ran from 19 July 2022 to 20 July 2022, US/Pacific time)

Fact	Condition	Source
Google says it will repair and carefully re-test its failover automation, develop more advanced methods to progressively decrease the thermal load within a single data center space, and audit cooling equipment and standards across the data centers that house Google Cloud globally.	Commitments in the report, not confirmed completed in anything we read.	Google Cloud Service Health, Multiple Cloud products experiencing elevated error rates, latencies or service unavailability in europe-west2 (Incident report) (Incident report posted 29 July 2022; the incident ran from 19 July 2022 to 20 July 2022, US/Pacific time)
Google's report says: "To our customers whose businesses were impacted during this outage, we sincerely apologize."	Google's own words (one sentence, quoted).	Google Cloud Service Health, Multiple Cloud products experiencing elevated error rates, latencies or service unavailability in europe-west2 (Incident report) (Incident report posted 29 July 2022; the incident ran from 19 July 2022 to 20 July 2022, US/Pacific time)
The Uptime Institute wrote on 27 July 2022 that temperatures at Heathrow Airport surpassed 40 degrees C to set an all-time record for the United Kingdom, and that this heatwave either approached or surpassed design specifications for ambient operating temperatures for data centers built before 2019.	Uptime's analysis; it also says most data centers likely made it through without a major incident, and notes some cloud providers "reportedly" saw thermal disruption. Uptime's piece does not name Google or Oracle.	Uptime Institute Journal (Daniel Bizo, Research Director), Extreme heat stress-tests European data centers, again (27 July 2022)
The Uptime Institute recommends reassessing climatic conditions for each site regularly, because the design conditions a data center was built for may be out of date.	Uptime's recommendation (27 July 2022), advice not a rule.	Uptime Institute Journal (Daniel Bizo, Research Director), Extreme heat stress-tests European data centers, again (27 July 2022)
UK Parliament's POSTnote 762 lists cooling failures at multiple London data centres during the July 2022 heatwave that disrupted services for UK businesses and hospitals and contributed to a cost of 1.4 million pounds for the NHS, and says the UK Government designated data centres as critical national infrastructure in September 2024.	16 March 2026 briefing, citing secondary sources for the cost; it does not name Google or Oracle in that passage.	UK Parliament, Parliamentary Office of Science and Technology, POSTnote 762, What are data centres and how sustainable are they? (16 March 2026)

The evidence: the bug at AWS

Fact	Condition	Source
AWS says the service disruption in the N. Virginia (us-east-1) Region began at 11:48 PM PDT on October 19, 2025 and ended at 2:20 PM PDT on October 20, 2025, with three distinct periods of impact to customer applications.	PDT. That is 14 hours 32 minutes by our subtraction. AWS also says Redshift cluster availability was only fully restored by 4:05 AM PDT on October 21 for some clusters.	Amazon Web Services (post-event summary), Summary of the Amazon DynamoDB Service Disruption in the Northern Virginia (US-EAST-1) Region (Event on 19 and 20 October 2025; the page carries no publication date)

Fact	Condition	Source
AWS's three periods were DynamoDB API errors from 11:48 PM on October 19 to 2:40 AM on October 20, Network Load Balancer connection errors from 5:30 AM to 2:09 PM, and failed new EC2 instance launches from 2:25 AM to 10:36 AM.	AWS adds that some newly launched instances had connectivity issues resolved by 1:50 PM.	Amazon Web Services (post-event summary), Summary of the Amazon DynamoDB Service Disruption in the Northern Virginia (US-EAST-1) Region (Event on 19 and 20 October 2025; the page carries no publication date)
AWS says the root cause was a latent race condition in the DynamoDB DNS management system that resulted in an incorrect empty DNS record for the service's regional endpoint that the automation failed to repair.	AWS's wording. It describes two independent components (a DNS Planner and DNS Enactors in three Availability Zones) and says the problem needed "manual operator intervention to correct".	Amazon Web Services (post-event summary), Summary of the Amazon DynamoDB Service Disruption in the Northern Virginia (US-EAST-1) Region (Event on 19 and 20 October 2025; the page carries no publication date)
AWS says that by 12:38 AM on October 20 its engineers had identified DynamoDB's DNS state as the source of the outage, and that by 2:25 AM all DNS information was restored.	PDT. Customers could connect again as cached DNS records expired between 2:25 AM and 2:40 AM, per AWS.	Amazon Web Services (post-event summary), Summary of the Amazon DynamoDB Service Disruption in the Northern Virginia (US-EAST-1) Region (Event on 19 and 20 October 2025; the page carries no publication date)
AWS says existing EC2 instances launched before the start of the event remained healthy and did not experience any impact.	The impact was on new launches and the control plane; running workloads were not stopped by this event, per AWS.	Amazon Web Services (post-event summary), Summary of the Amazon DynamoDB Service Disruption in the Northern Virginia (US-EAST-1) Region (Event on 19 and 20 October 2025; the page carries no publication date)
AWS says the EC2 launch subsystem DWFM entered a state of congestive collapse after DynamoDB recovered and that engineers throttled work and restarted DWFM hosts, with leases restored by 5:28 AM.	AWS wording; AWS says there was no established operational recovery procedure for this situation.	Amazon Web Services (post-event summary), Summary of the Amazon DynamoDB Service Disruption in the Northern Virginia (US-EAST-1) Region (Event on 19 and 20 October 2025; the page carries no publication date)
AWS says customers with DynamoDB global tables could connect to replica tables in other Regions but saw prolonged replication lag to and from us-east-1.	A real example of a multi-region design limiting impact.	Amazon Web Services (post-event summary), Summary of the Amazon DynamoDB Service Disruption in the Northern Virginia (US-EAST-1) Region (Event on 19 and 20 October 2025; the page carries no publication date)
AWS says it has already disabled the DynamoDB DNS Planner and DNS Enactor automation worldwide, will fix the race condition and add protections against incorrect DNS plans before re-enabling it, and will add NLB velocity controls, a new EC2 test suite and better throttling.	Commitments at the time of the summary; We did not check whether they are all complete.	Amazon Web Services (post-event summary), Summary of the Amazon DynamoDB Service Disruption in the Northern Virginia (US-EAST-1) Region (Event on 19 and 20 October 2025; the page carries no publication date)

Fact	Condition	Source
AWS's summary closes: "We apologize for the impact this event caused our customers."	AWS's own words.	Amazon Web Services (post-event summary), Summary of the Amazon DynamoDB Service Disruption in the Northern Virginia (US-EAST-1) Region (Event on 19 and 20 October 2025; the page carries no publication date)

Your notes

Eight kinds of cloud and infrastructure job

Cloud and infrastructure is eight areas of work, from designing systems to keeping the cables and cooling running. See the jobs in each, the pay with its source and the routes in.

What this is

Eight areas of cloud and infrastructure work from the TechSkills Career Matrix. For each one you can see the jobs, the pay with its own source, who it tends to suit and the routes in.

What to do

- Pick an area you know nothing about, such as Solutions Architecture. Read Jobs in this area and Best suited to.
- Open Pay, with its source for two areas, one that is done at a desk and one that is done on site, and compare.
- Find a route into an area that is not a degree, and note its level and length.
- Find one area that has no apprenticeship route of its own, and read what the page says about the job title cloud engineer.

What you should notice

Cloud and infrastructure is eight areas, some done at a desk and some on site. Pay ranges are a guide and come from the source shown. No apprenticeship standard is called cloud engineer, so the page names the standards where those job titles appear.

Time	20 to 25 minutes
You need	A projector or one device per pair.
Where	Interactive on the page. Needs a screen. Suits both a desk and a plant room.

Students will be able to

- Name at least three areas of cloud and infrastructure work and a job in each.
- Say which areas are done at a desk, which on site, and which both.
- Say what the sources say about pay, and what they do not.

Running it

- Show the eight areas and ask students to pick one they know nothing about and one they might try.
- Open the pay panel for two areas that differ in how they are done, one at a desk and one on site. Read out the pay and its source.
- Ask students to find a job in an area that does not sound like a computer job.

Questions to ask

- Which of these jobs would you not have known existed?
- Which needs the most hands-on work, and which the most analysis?

- Which would you try first, and what would you need to find out?

Watch for

- Students assuming cloud means only desk work. Data centre operations is on site and hands-on.
- Pay ranges are a guide from the sources shown, and they vary by employer and place.
- Any area can be reached by more than one route. Nothing here is a ceiling.

Area	Jobs	Pay, with its source	Best suited to
01. Cloud Engineering	Cloud Engineer, Cloud Architect, Platform Engineer, FinOps Analyst	£35,000 to £117,500+ (median junior Cloud Engineer to 90th-percentile Cloud Architect, IT Jobs Watch)	Analytical minds who enjoy building complex, logical systems in the cloud.
02. DevOps & Automation	DevOps Engineer, SRE (Reliability), Automation Lead, Build Engineer	£40,500 to £85,000+ (median, Build Engineer to Lead DevOps Engineer / Site Reliability Engineer, IT Jobs Watch)	Automation enthusiasts who hate doing the same manual task twice.
03. Systems & Virtualisation	Systems Administrator, Virtualisation Lead, Storage Engineer, Database Administrator	£35,750 to £75,000 (median, IT Systems Administrator to Senior Database Administrator, IT Jobs Watch)	Methodical thinkers who enjoy the 'engine room' of IT.
04. Data Centre Ops (hard hat route)	Data Centre Technician, Hardware Engineer, DC Ops Manager, Fibre Technician	£35,000 to £100,000 (median, Fibre Engineer to Data Centre Manager, IT Jobs Watch)	Practical people who enjoy working with physical technology and hardware.
05. Solutions Architecture	Solutions Architect, Enterprise Architect, Infrastructure Strategist	£54,481 to £100,000 (ONS median IT solutions architects/designers to IT Jobs Watch median Principal Solutions/Enterprise Architect)	Big-picture thinkers who can bridge the gap between CEO and Engineer.
06. Emerging Tech	Edge Architect, Sustainable Tech Lead, Quantum Infra Lead	£45,000 to £93,550+ (proxy: Sustainability Analyst median to 90th-percentile Edge Computing postings, IT Jobs Watch/Glassdoor, no dedicated listing for niche 'Quantum Infra Lead')	Early adopters who want to build the foundations of tomorrow.
07. Site Reliability & Observability	Site Reliability Engineer, Observability Engineer, Incident Manager, Performance Engineer	£52,500 to £85,000 (median, Observability-keyword postings to Site Reliability Engineer (London), IT Jobs Watch, no dedicated 'Observability Engineer' page)	Calm operators who thrive when things go wrong and need fixing fast.
08. Cloud Security & Compliance	Cloud Security Engineer, Compliance Analyst, FinOps & GRC Lead, Cloud Governance Manager	£50,000 to £84,750+ (median Compliance Analyst to 90th-percentile GRC/Cloud Security roles, IT Jobs Watch)	Detail-oriented engineers who enjoy translating regulation into technical controls.

A student who searches for a cloud engineer apprenticeship will not find one. Skills England has no current standard with cloud engineer, data centre technician or site reliability engineer in its title. Those titles appear inside other standards instead: the Information communications technician standard lists data centre support technician and cloud technician as typical job titles, and the DevOps engineer standard lists site reliability

engineer and platform engineer. The Career Matrix on this page shows each of them under the area of work it fits.

The government's own role framework is the same. The Digital and Data Profession Capability Framework has infrastructure engineer, infrastructure operations engineer, DevOps engineer and architect roles, and no site reliability engineer, cloud engineer or platform engineer role.

The evidence

Fact	Condition	Source
The Skills England listing shows no current apprenticeship standard titled cloud engineer, cloud specialist, data centre technician, data centre engineer or site reliability engineer.	Checked every non-retired entry by title and overview text across all 2086 listed items on 2026-11-05. "Cloud technician", "Data centre support technician" and "Site reliability engineer" appear only as job titles inside other standards.	Skills England, Apprenticeship search (full listing, 2086 entries) (Live listing, read 5 November 2026)
No current Skills England apprenticeship standard has data centre, cloud engineer or site reliability engineer in its title.	Checked by reading every non-retired entry in the Skills England listing of 2086 entries on 5 November 2026. Data centre support technician, cloud technician and site reliability engineer appear only as typical job titles inside other standards.	Skills England, Apprenticeship search (full listing, 2086 entries) (Live listing, read 5 November 2026)
Typical job titles in the ST0973 standard include data centre support technician and cloud technician.	Other listed titles include IT field technician, network field operative, first-line support, help desk support and telecommunications technician.	Skills England, Information communications technician (ST0973) v1.2, Level 3 (Date updated 30/06/2026; approved for delivery 4 May 2021)
Typical job titles in the DevOps engineer standard include infrastructure engineer, platform engineer and site reliability engineer.	The full list is: automation engineer, build and release engineer, deployment engineer, devops engineer, full stack developer, infrastructure engineer, platform engineer, reliability engineer, site reliability engineer.	Skills England, DevOps engineer (ST0825) v1.1, Level 4 (Date updated 17/06/2026; approved for delivery 12 March 2020)

Fact	Condition	Source
The framework role list has no site reliability engineer, cloud engineer or platform engineer role.	Checked the full role list on the home page on 2026-11-05. Infrastructure engineer, infrastructure operations engineer, DevOps engineer, technical architect and solution architect are the closest.	Government Digital Service, Government Digital and Data Profession Capability Framework home (Current page, retrieved 2026-11-05)
The framework says an infrastructure engineer designs, builds, manages and supports the infrastructure services that underpin all internal user services and services to the public.	Last updated 28 August 2026; five role levels from associate infrastructure engineer to principal infrastructure engineer.	Government Digital Service, GOV.UK, Infrastructure engineer, Government Digital and Data Profession Capability Framework (Published 7 January 2020, last updated 28 August 2026)
At the second role level (infrastructure engineer) the IT infrastructure skill is at "Working" level, which means you can "build, configure, administer and support infrastructure solutions and services".	Examples in the page include networking and physical infrastructure, COTS software, open source packages and "virtual and cloud computing". The associate level is "Awareness", then Working, Practitioner, Expert and Practitioner for levels 3 to 5. "Four ascending skill levels" is the page's wording.	Government Digital Service, GOV.UK, Infrastructure engineer, Government Digital and Data Profession Capability Framework (Published 7 January 2020, last updated 28 August 2026)
The associate infrastructure engineer is described as a trainee who works in an established team and supports infrastructure engineers in their daily work, under their direction.	Most often performed at Civil Service grade EO.	Government Digital Service, GOV.UK, Infrastructure engineer, Government Digital and Data Profession Capability Framework (Published 7 January 2020, last updated 28 August 2026)
The framework says an infrastructure operations engineer supports, manages and maintains the core infrastructure that underpins production services.	Last updated 30 November 2024; five levels.	Government Digital Service, Infrastructure operations engineer, same framework (Published 7 January 2020, last updated 30 November 2024)
The framework says development operations engineers support the development and operation of software through tools, environments and practices.	Last updated 29 August 2025; nine role levels from apprentice DevOps engineer to principal DevOps engineer - management.	Government Digital Service, Development operations (DevOps) engineer, same framework (Published 7 January 2020, last updated 29 August 2025)
The framework says a technical architect provides technical leadership and architectural design.	Last updated 28 August 2026; five levels.	Government Digital Service, Technical architect, same framework (Last updated 28 August 2026)

Fact	Condition	Source
The framework says a solution architect designs solutions for problems that affect the organisation.	Last updated 28 August 2026; five levels.	Government Digital Service, Solution architect, same framework (Last updated 28 August 2026)

Three ways in, side by side

Three kinds of route into cloud and infrastructure are worth putting next to each other, and none comes first on this page. One starts with study. Two are apprenticeships, which are jobs with training in them: one mostly at a desk and one mostly on site. Cloud and infrastructure is one of the families with a hard hat route in the Career Matrix, Data Centre Operations, so the columns are balanced between the grad cap and the hard hat.

A plain university degree route is missing from the first column on purpose: we did not find an official source that describes one for these jobs, and we would rather say so than guess. What we did find is sourced, and the first column shows it: a T Level, and a Level 7 apprenticeship that Skills England says is likely to need a degree. The T Level that fits best is Digital Support and Security, whose Digital infrastructure specialism is aligned to the Information communications technician standards.

The desk column is DevOps engineer at Level 4, IT solutions technician at Level 3, which started a new version on 2 November 2026, and the Level 6 integrated degree apprenticeship in digital and technology solutions, which is not specific to cloud work. The site column is the Level 2 foundation apprenticeship, the Information communications technician standard and the cooling standard, and it says plainly what is missing: no standard covers the back-up power a data centre depends on, and the cooling standard names data centres once, in a list of places the work is done.

Several of these are being replaced for new starts: DevOps engineer on 18 December 2026 and Information communications technician on 28 January 2027. Check the live Skills England page before giving a student a version. Pay for each area of work, with its source, is in the Career Matrix above.

A university degree, an apprenticeship at a desk and an apprenticeship on site, side by side, with the same information for each from the people who set the standards.

What this is

Three ways in side by side: a study route, an apprenticeship at a desk and an apprenticeship on site, with the same information for each, taken from Skills England and the government's T Level pages.

What to do

- Read the three headings and say what the columns have in common.
- Take one column each in a group. List three questions a student would ask, then find the answers on the page.
- Open a standard from your column. Read its level and length, then the line that says what the length leaves out.
- Find the sentence that says what none of the standards covers.

What you should notice

The three columns have equal weight and a student can move between them. No standard is called cloud engineer or data centre technician, and none covers a data centre's back-up power. Several standards are being replaced for new starts on dates the page gives.

Time	20 to 25 minutes
You need	A projector or one device per pair.

Where	Interactive on the page. Needs a screen. Suits both a desk and a plant room.

Students will be able to

- Compare three routes on level, length and who pays.
- Say what each route tends to suit, as a tendency and not a rule.
- Explain that no apprenticeship standard is called cloud engineer or data centre technician, and say where those jobs appear instead.

Running it

- Show the three columns together and read the headings. Ask students to say what the columns have in common before you compare them.
- Ask groups to take one column each and list three questions someone choosing it would want answered. Then check the page for the answers.
- Open a standard on the page and read the level and length together. Ask what the number of months leaves out.
- Close by asking what a student who picked one route and wanted to change would do.

Questions to ask

- What does each route cost a student, and what does it pay them?
- Who is each route likely to suit, and who might it leave out?
- Why might a job called cloud technician not have an apprenticeship with that name?

Watch for

- Treating any column as the better one. They have equal weight, and a student who picks one is not locked in.
- Entry requirements change each year and by employer. Send students to the standard, the employer and the course for the current ones.
- Several standards on the page are being replaced for new starts on dates the page gives. Check the live page before giving a student a version.
- No standard covers a data centre's backup power. The hands-on route here covers IT and communications field work, cooling and a foundation route, and the page says so.

Grad cap: the study route

Often suits someone who wants to study first and train after, or who already holds a degree and wants to move towards design and architecture.

Step	Level and length	What it is
T Level in Digital Support and Security	Level 3, 2 years	A course taken after GCSEs with an industry placement of at least 315 hours. Its Digital infrastructure specialism is aligned to the Information communications technician standards.

Step	Level and length	What it is
Digital and technology solutions specialist (integrated degree) (ST0482)	Level 7, typically 18 months	A specialist route with an enterprise systems architecture option. Skills England says entry is likely to include a degree at 2.1 or higher in a relevant subject, though some employers will accept other qualifications or experience.

Where it can lead: The T Level's Digital infrastructure specialism is aligned to the Information communications technician occupations, which is the Level 3 standard in the site column. The Level 7 standard is for people who already hold a degree, and government funding for it has been restricted since 1 January 2026.

The evidence

Fact	Condition	Source
A T Level is a 2-year course taken after GCSEs, broadly equivalent in size to 3 A levels.	Updated 26 January 2026; applies to the whole T Level programme.	Department for Education, GOV.UK, Introduction of T Levels (Updated 26 January 2026)
Every T Level includes an industry placement of a minimum of 315 hours, approximately 45 days.	gov.uk: "but can last longer". The tlevels.gov.uk Digital Support and Security page words it as "9 weeks minimum"; the home page says "at least 45 days".	Department for Education, GOV.UK, Introduction of T Levels (Updated 26 January 2026); Skills England, Occupational specialism: Digital infrastructure (TL0007a) (Approval date 17/12/2020; available from 01/09/2020)
DfE expects typical classroom time of between 1100 and 1300 hours across the 2 years, plus the mandatory industry placement of at least 315 hours.	Described as "typical time".	Department for Education, GOV.UK, Introduction of T Levels (Updated 26 January 2026)
The T Level is called Digital Support and Security on the T Levels student site, and the occupational specialisms are digital infrastructure, network cabling, digital support and cyber security.	The Skills England pages use the older title "Digital Support Services" and list the same four. Students choose one specialism.	T Levels (DfE), Digital Support and Security (Page carries no date)
The T Level core content covers problem solving, introduction to digital business support, data, business context, legislation and regulatory requirements, emerging issues, digital environments and security.	As listed on the student page.	T Levels (DfE), Digital Support and Security (Page carries no date)
The student page says the T Level is suitable for anyone wanting a career in digital infrastructure and support, with career options that might include infrastructure technician or a role in IT support.	"might include". The page also says entry requirements are set by the individual school or college.	T Levels (DfE), Digital Support and Security (Page carries no date)

Fact	Condition	Source
The Digital infrastructure specialism is a Level 3 T Level specialism aligned to the Information communications technician occupations (Digital Communications Technician, Network Technician and Support Technician).	Approval date 17/12/2020; available from 01/09/2020; awarding organisation NCFE. The Skills England page gives no description of the specialism's content, and we found no official description elsewhere.	Skills England, Occupational specialism: Digital infrastructure (TL0007a) (Approval date 17/12/2020; available from 01/09/2020)
The current T Levels subject list shows three digital T Levels: Digital software development, Digital data analytics and Digital support and security.	Updated 26 January 2026. We did not read any page on Digital Production, Design and Development or Digital Business Services, so nothing is claimed about them.	Department for Education, GOV.UK, Introduction of T Levels (Updated 26 January 2026)
The Digital and technology solutions specialist (integrated degree) apprenticeship (ST0482) is Level 7, 18 months, maximum funding £21,000, and includes a Digital Business and Enterprise Systems Architecture Specialist option.	The page says this standard is being revised and the current version remains approved; "this does not include assessment period"; entry is "likely to include a degree at 2.1 or higher in a relevant subject, although some employers will accept other relevant qualifications or experience".	Skills England, Digital and technology solutions specialist (integrated degree) (ST0482) v1.0, Level 7 (Date updated 14/01/2022; approved for delivery 7 August 2018; page notes it is being revised)
From 1 January 2026, Level 7 apprenticeships are government-funded only for apprentices aged 16 to 21 at the start, or under 25 with an Education, Health and Care plan or care experience.	Anyone who started a Level 7 apprenticeship before 1 January 2026 continues to be funded to completion.	Skills England, Digital and technology solutions specialist (integrated degree) (ST0482) v1.0, Level 7 (Date updated 14/01/2022; approved for delivery 7 August 2018; page notes it is being revised)

The desk-based apprenticeship route

Often suits someone who would rather learn the work inside an employer from the start, mostly at a screen, and who is comfortable with a longer route in exchange for being employed while they train.

Step	Level and length	What it is
IT solutions technician (ST0505)	Level 3, typically 18 months	Developing, implementing and maintaining complete IT solutions such as networks, operating systems and applications.
DevOps engineer (ST0825)	Level 4, typically 24 months	Enabling organisations to get valuable working software out in front of active users and improving the quality of digital services.

Step	Level and length	What it is
Digital and technology solutions professional (integrated degree) (ST0119)	Level 6, typically 48 months	A degree built into the job. It is not specific to cloud work: its options include network engineering professional, IT consultant professional and cyber security professional.

Where it can lead: The IT solutions technician page lists example progression routes including network engineer and DevOps engineer. The DevOps engineer standard lists infrastructure engineer, platform engineer and site reliability engineer among its typical job titles.

The evidence

Fact	Condition	Source
The IT solutions technician apprenticeship (ST0505) is at version 2.0, which started on 02/11/2026, and version 1.1 is marked Retired.	v1.1 had a latest start date of 01/11/2026.	Skills England, IT solutions technician (ST0505) v2.0, Level 3 (Date updated 02/11/2026; approved for delivery 18 February 2019; v2.0 earliest start 02/11/2026); Skills England, IT solutions technician (ST0505) v1.1 (Status Retired; Date updated 02/11/2026; latest start 01/11/2026)
The IT solutions technician apprenticeship is a Level 3 standard with a typical duration of 18 months.	Version 2.0. The summary tab says the end-point assessment period is typically 4 months, so assessment is in addition.	Skills England, IT solutions technician (ST0505) v2.0, Level 3 (Date updated 02/11/2026; approved for delivery 18 February 2019; v2.0 earliest start 02/11/2026)
The maximum funding for the IT solutions technician apprenticeship is £15,000.	Version 2.0; minimum hours for compliance 370.	Skills England, IT solutions technician (ST0505) v2.0, Level 3 (Date updated 02/11/2026; approved for delivery 18 February 2019; v2.0 earliest start 02/11/2026)
The standard describes the role as developing, implementing and maintaining complete IT solutions such as networks, operating systems and applications.	The page's overview line.	Skills England, IT solutions technician (ST0505) v2.0, Level 3 (Date updated 02/11/2026; approved for delivery 18 February 2019; v2.0 earliest start 02/11/2026)
Typical job titles in the IT solutions technician standard are junior applications support technician, junior database technician, junior devops engineer, junior IT support technician and junior IT systems support.	The page spells the third one "dev0ps" (a typo with a zero).	Skills England, IT solutions technician (ST0505) v2.0, Level 3 (Date updated 02/11/2026; approved for delivery 18 February 2019; v2.0 earliest start 02/11/2026)
The page lists example progression routes from ST0505 as software developer, software tester, network engineer, cyber security technologist (2021) and DevOps engineer.	These are "example progression routes", not entitlements.	Skills England, IT solutions technician (ST0505) v2.0, Level 3 (Date updated 02/11/2026; approved for delivery 18 February 2019; v2.0 earliest start 02/11/2026)

Fact	Condition	Source
The DevOps engineer apprenticeship (ST0825) is a Level 4 standard.	Version 1.1, status "Approved for delivery (available for starts)", page updated 17/06/2026.	Skills England, DevOps engineer (ST0825) v1.1, Level 4 (Date updated 17/06/2026; approved for delivery 12 March 2020)
The DevOps engineer apprenticeship has a typical duration of 24 months.	"this does not include assessment period".	Skills England, DevOps engineer (ST0825) v1.1, Level 4 (Date updated 17/06/2026; approved for delivery 12 March 2020)
The maximum funding for the DevOps engineer apprenticeship is £17,000.	Current version 1.1; minimum hours for compliance 466.	Skills England, DevOps engineer (ST0825) v1.1, Level 4 (Date updated 17/06/2026; approved for delivery 12 March 2020)
A revised DevOps engineer standard is approved and replaces version 1.1 for new starts on 18 December 2026.	Until then the current standard and end-point assessment remain in place; the revised version is "available for information only"; the revised version's page says the funding band is £17,000. Version 1.1 has a latest start date of 17/12/2026.	Skills England, DevOps engineer (ST0825) v1.1, Level 4 (Date updated 17/06/2026; approved for delivery 12 March 2020); Skills England, DevOps engineer (ST0825) revised version awaiting implementation (Page updated 17 June 2026; its status line reads Standard in development)
The standard describes the DevOps engineer role as enabling organisations to get valuable working software out in front of active users and improving the quality of digital services.	This is the page's overview line.	Skills England, DevOps engineer (ST0825) v1.1, Level 4 (Date updated 17/06/2026; approved for delivery 12 March 2020)
The standard says a DevOps engineer takes a cloud-infrastructure focused perspective and uses Cloud computing so that infrastructure can be defined in code.	Wording from the occupation summary; office based or remote working role.	Skills England, DevOps engineer (ST0825) v1.1, Level 4 (Date updated 17/06/2026; approved for delivery 12 March 2020)
The Digital and technology solutions professional (integrated degree) apprenticeship (ST0119) is Level 6 with a typical duration of 48 months and maximum funding of £27,000.	Version 1.2, updated 01/09/2023; options include network engineering professional, IT consultant professional and cyber security professional; the qualification is "Level: 6 (integrated degree)". The page states no entry requirement.	Skills England, Digital and technology solutions professional (integrated degree) (ST0119) v1.2, Level 6 (Date updated 01/09/2023; approved for delivery 26 March 2015)
The Network engineer apprenticeship (ST0127) is at version 2.0 and is Level 4, 24 months, with maximum funding of £19,000.	Date updated 02/11/2026; version 1.4 is shown as retired on 02 Nov 2026. Job titles include "Infrastructure engineer", "Network and cloud engineer" and "Systems engineer".	Skills England, Network engineer (ST0127) v2.0, Level 4 (Date updated 02/11/2026; approved for delivery 12 November 2014); Skills England, Apprenticeship search (full listing, 2086 entries) (Live listing, read 5 November 2026)

Fact	Condition	Source
The ST0127 standard says the demand for people who can manage, build and maintain on-site, cloud and hybrid networks is increasing.	Wording from the occupation summary (grammar as on the page).	Skills England, Network engineer (ST0127) v2.0, Level 4 (Date updated 02/11/2026; approved for delivery 12 November 2014)
The Digital support technician apprenticeship (ST0120) is at version 2.0, Level 3, 15 months, with maximum funding of £13,000.	Date updated 03/11/2026; v2.0 earliest start 29/10/2026; EPA typically 3 months; job titles include IT operations technician and database administrator.	Skills England, Digital support technician (ST0120) v2.0, Level 3 (Date updated 03/11/2026; v2.0 earliest start 29/10/2026)
The Cyber security technician apprenticeship (ST0865) is Level 3, 18 months, with maximum funding of £11,000.	Version 1.1, updated 08/09/2026; "this does not include assessment period"; a revised version replaces it for new starts on 9 March 2027. Its purpose is first line cyber security support.	Skills England, Cyber security technician (ST0865) v1.1, Level 3 (Date updated 08/09/2026; approved for delivery 13 May 2020)

Hard hat: the on-site apprenticeship route

Often suits someone who wants to keep real hardware running, who is comfortable with shift-based, site-based work and who would rather learn it hands-on from the start.

Step	Level and length	What it is
Hardware, network and infrastructure foundation apprenticeship (FA0004)	Level 2, typically 8 months to the gateway	A Level 2 start, normally for people aged 16 to 21 at the start, whose job titles include cloud technician and data centre support technician.
Information communications technician (ST0973)	Level 3, typically 18 months	Delivering efficient operation and control of IT and telecommunications infrastructure, on premises or provisioned as cloud services.
Refrigeration, air conditioning and heat pump engineering technician (ST0322)	Level 3, typically 36 months	Installing, maintaining, servicing, commissioning and decommissioning the cooling systems found in refrigeration, air conditioning and heat pumps. Data centres are one of many places the work is done.

Where it can lead: The Information communications technician standard lists data centre support technician and cloud technician among its job titles. No standard we found covers a data centre's backup power, such as batteries and generators, and the electrician and building services standards we read do not mention data centres.

The evidence

Fact	Condition	Source
The Hardware, network and infrastructure foundation apprenticeship (FA0004) is Level 2, with a typical duration to gateway of 8 months and maximum funding of £4,000.	Approved for delivery 29/04/2025; minimum 187 hours off-the-job learning; the apprentice "must normally be age 16 to 21 at the start".	Skills England, Hardware, network and infrastructure foundation apprenticeship (FA0004) v1.0, Level 2 (Approved for delivery 29/04/2025)
FA0004 lists cloud technician and data centre support technician among its job titles.	The list is shared across the occupations it covers (ST0505, ST0120, ST0865, ST0973 among them).	Skills England, Hardware, network and infrastructure foundation apprenticeship (FA0004) v1.0, Level 2 (Approved for delivery 29/04/2025)
The Information communications technician apprenticeship (ST0973) is a Level 3 standard with a typical duration of 18 months.	Version 1.2, updated 30/06/2026, "this does not include assessment period".	Skills England, Information communications technician (ST0973) v1.2, Level 3 (Date updated 30/06/2026; approved for delivery 4 May 2021)
The maximum funding for the Information communications technician apprenticeship is £15,000.	Minimum hours for compliance 348.	Skills England, Information communications technician (ST0973) v1.2, Level 3 (Date updated 30/06/2026; approved for delivery 4 May 2021)
ST0973 has three options: Support Technician, Network Technician and Digital Communications Technician.	The page says "This standard has options".	Skills England, Information communications technician (ST0973) v1.2, Level 3 (Date updated 30/06/2026; approved for delivery 4 May 2021)
The standard says the occupation is to deliver efficient operation and control of the IT and/or telecommunications infrastructure, comprising physical or virtual hardware, software, network services and data storage.	The page says this is "either on-premises or to end-users provisioned as cloud services".	Skills England, Information communications technician (ST0973) v1.2, Level 3 (Date updated 30/06/2026; approved for delivery 4 May 2021)
The standard's occupation summary includes the building and management of systems and components in virtualised and cloud computing environments.	Part of the occupation summary, and the knowledge list includes the principles of virtualisation of servers, applications and networks.	Skills England, Information communications technician (ST0973) v1.2, Level 3 (Date updated 30/06/2026; approved for delivery 4 May 2021)
Typical job titles in the ST0973 standard include data centre support technician and cloud technician.	Other listed titles include IT field technician, network field operative, first-line support, help desk support and telecommunications technician.	Skills England, Information communications technician (ST0973) v1.2, Level 3 (Date updated 30/06/2026; approved for delivery 4 May 2021)

Fact	Condition	Source
A revised ST0973 standard is approved and replaces version 1.2 for new starts on 28 January 2027.	Version 1.2 latest start date 27/01/2027; the revised version is available for information only until then.	Skills England, Information communications technician (ST0973) v1.2, Level 3 (Date updated 30/06/2026; approved for delivery 4 May 2021)
The ST0973 page lists no entry requirements and no mandatory qualification beyond English and maths.	The page gives no entry requirement or degree.	Skills England, Information communications technician (ST0973) v1.2, Level 3 (Date updated 30/06/2026; approved for delivery 4 May 2021)
The Refrigeration, air conditioning and heat pump engineering technician apprenticeship (ST0322) is a Level 3 standard with a typical duration of 36 months.	Version 1.3, updated 08/12/2025. The key information lists a typical assessment period of 6 months separately, so assessment is in addition.	Skills England, Refrigeration air conditioning and heat pump engineering technician (ST0322) v1.3, Level 3 (Date updated 08/12/2025; approved for delivery 10 September 2015)
The maximum funding for ST0322 is £20,000.	Minimum hours for compliance 835; route is "Construction and the built environment".	Skills England, Refrigeration air conditioning and heat pump engineering technician (ST0322) v1.3, Level 3 (Date updated 08/12/2025; approved for delivery 10 September 2015)
The standard describes the occupation as installation of components, routine maintenance, reactive service, commissioning and decommissioning of vapour compression systems, as found in refrigeration, air conditioning and heat pump systems.	"This is a specialist role".	Skills England, Refrigeration air conditioning and heat pump engineering technician (ST0322) v1.3, Level 3 (Date updated 08/12/2025; approved for delivery 10 September 2015)
The standard lists IT/data centres as one of several applications where this work is carried out.	It appears in a list that also includes food production, product distribution, retail storage and display, transport and office climate control, manufacturing processes and medical and healthcare services. It is the only mention of data centres in the standard.	Skills England, Refrigeration air conditioning and heat pump engineering technician (ST0322) v1.3, Level 3 (Date updated 08/12/2025; approved for delivery 10 September 2015)
Typical job titles in ST0322 include air conditioning engineer, commissioning engineer, heat pump engineer, maintenance engineer, refrigeration technician and service engineer.	"Typical job titles include".	Skills England, Refrigeration air conditioning and heat pump engineering technician (ST0322) v1.3, Level 3 (Date updated 08/12/2025; approved for delivery 10 September 2015)

Fact	Condition	Source
ST0322 requires three other qualifications before the end-point assessment: an Award in F Gas and ODS Regulations Category 1, a Level 2 Diploma in Refrigeration, Air Conditioning and Heat Pump Systems, and a Level 3 Certificate or Diploma in the same subject.	Listed as mandatory qualifications and gateway requirements; no degree is required.	Skills England, Refrigeration air conditioning and heat pump engineering technician (ST0322) v1.3, Level 3 (Date updated 08/12/2025; approved for delivery 10 September 2015)
The standard's duties are about planning tasks, installing, commissioning, servicing and decommissioning, and the word "design" does not describe the role.	Checked against the duty text; the page does not say the technician designs systems.	Skills England, Refrigeration air conditioning and heat pump engineering technician (ST0322) v1.3, Level 3 (Date updated 08/12/2025; approved for delivery 10 September 2015)
The former Infrastructure technician standard (ST0125, Level 3) is retired and cannot be named as a current route.	Listing: "Retired on 31 Jul 2021".	Skills England, Apprenticeship search (full listing, 2086 entries) (Live listing, read 5 November 2026)
Skills England's installation and maintenance electrician standard (ST0152) and its building services engineering service and maintenance engineer standard (ST0061) do not mention data centres.	Summary pages read on 5 November 2026. ST0152 describes low voltage electrical installation, testing and maintenance in industrial, commercial and residential settings, and ST0061 describes planned and corrective maintenance of building services such as ventilation, heating and water supply. Neither is a data centre standard.	Skills England, Installation and maintenance electrician (ST0152) (Summary page read 5 November 2026); Skills England, Building services engineering service and maintenance engineer (ST0061) (Summary page read 5 November 2026)
Skills England's maintenance and operations engineering technician standard (ST0154) does not mention data centres.	Page read on 5 November 2026, version 1.6, Level 3, typical duration 36 months, maximum funding £26,000, updated 05/03/2026.	Skills England, Maintenance and operations engineering technician (ST0154) (Summary page read 5 November 2026)

The evidence

Fact	Condition	Source
The Career Matrix maps six apprenticeship standards to this family: DevOps engineer (ST0825), IT solutions technician (ST0505), Information communications technician (ST0973), the Hardware, network and infrastructure foundation apprenticeship (FA0004), Refrigeration, air conditioning and heat pump engineering technician (ST0322) and the Level 7 Digital and technology solutions specialist (ST0482).	Read from the Matrix on 5 November 2026. DevOps engineer appears under two areas of work, DevOps and Automation and Site Reliability and Observability. A test checks that the pack and the Matrix agree.	TechSkills, Cloud and Infrastructure Career Matrix (Maintained by TechSkills, apprenticeship standards last checked 5 November 2026)

Fact	Condition	Source
Skills England shows the IT solutions technician standard at version 2.0 from 2 November 2026 and marks version 1.1 as retired.	Both pages read on 5 November 2026.	Skills England, IT solutions technician (ST0505) v2.0, Level 3 (Date updated 02/11/2026; approved for delivery 18 February 2019; v2.0 earliest start 02/11/2026); Skills England, IT solutions technician (ST0505) v1.1 (Status Retired; Date updated 02/11/2026; latest start 01/11/2026)

Students will hear about vendor certificates. The two beginner ones the Matrix names are AWS Certified Cloud Practitioner and Microsoft Azure Fundamentals. AWS says its exam is designed for people new to cloud who may not have an IT background. Both are issued by the companies that run the clouds, and the fees and exam details can change, so send students to the issuer's page.

The evidence

Fact	Condition	Source
AWS says the AWS Certified Cloud Practitioner validates a foundational, high-level understanding of AWS Cloud, services and terminology.	Category "Foundational". Page undated.	Amazon Web Services, AWS Certified Cloud Practitioner (Page carries no date, read 5 November 2026)
AWS says the exam is designed for candidates new to Cloud who may not have an information technology (IT) background, and that candidates could have up to 6 months of exposure to AWS Cloud, but this is not required.	Page undated.	Amazon Web Services, AWS Certified Cloud Practitioner (Page carries no date, read 5 November 2026)
The AWS Certified Cloud Practitioner exam is 90 minutes, with 65 questions that are either multiple choice or multiple response.	Test at a Pearson VUE centre or online proctored.	Amazon Web Services, AWS Certified Cloud Practitioner (Page carries no date, read 5 November 2026)
AWS lists the exam cost as 100 USD.	The page directs readers to "Exam pricing for additional cost information, including foreign exchange rates"; no sterling fee is stated.	Amazon Web Services, AWS Certified Cloud Practitioner (Page carries no date, read 5 November 2026)
Microsoft Learn currently lists the Microsoft Certified: Azure Fundamentals certification, at Beginner level, with the AZ-900 exam.	"Last Updated 07/20/2026". The page names the exam as AZ-900 in the study guide link.	Microsoft Learn, Microsoft Certified: Azure Fundamentals (Last Updated 07/20/2026)
The Azure Fundamentals exam page says "You will have 45 minutes to complete this assessment", and that the exam is proctored.	The page also says there may be interactive components.	Microsoft Learn, Microsoft Certified: Azure Fundamentals (Last Updated 07/20/2026)
Microsoft says the exam assesses: describe cloud concepts, describe Azure architecture and services, and describe Azure management and governance.	As listed under "Assessed on this exam".	Microsoft Learn, Microsoft Certified: Azure Fundamentals (Last Updated 07/20/2026)

Your notes

The Translator

Jargon in plain English, with a comparison from school life. Where the comparison stops working, it says so.

Term	In plain English	Think of it like this
Availability	The share of the time a service is working when people try to use it, usually written as a percentage such as 99.9%.	The share of lessons in a year in which the room was open and the projector worked. Where that stops working: A percentage hides how the time is spread. Ten minutes down every day feels very different from one long outage.
Availability zone	A separate location inside a cloud provider's region, designed so that the failure of one does not take down the others.	Two classrooms in different corridors, so that a problem in one does not close the other. Where that stops working: Two classrooms still share the same building and the same main supply, and zones can share more than they seem to.
Cloud	Computing services, such as storage, software and processing, that you use over the internet from someone else's computers and pay for by how much you use.	Borrowing a minibus from the council instead of buying one: you ask when you need it, use it and hand it back, and someone else keeps it running. Where that stops working: A minibus is one vehicle you can see. The cloud is a very large number of computers shared between many customers, and you never see which ones are yours.
Colocation	A data centre where a company rents space, power and cooling for its own computers, rather than building its own or renting computers from a cloud provider.	Renting a stall in a market hall: your goods, their building. Where that stops working: A stall holder can walk away with the stall. Moving a rack of servers between buildings is a project.
Data centre	A building, or part of one, full of computers, with the power, cooling and network connections they need. Sources count them differently, because there is no universally agreed definition.	The school kitchen. It is a room built round the things that make the meal happen: the cooker, the extraction, the water and the delivery door. Where that stops working: A kitchen serves one school. A data centre may hold computers for thousands of customers who never meet.
DevOps	A way of working in which the people who build software and the people who run it work as one team, using tools to release changes often and safely.	The same people who write the school play also set up the stage and run the lights on the night, instead of passing it over a wall. Where that stops working: A play is performed once. Software is changed and released again and again, so the tools matter more.
Elasticity	The ability to add computing capacity when demand rises and take it away when it falls, so you pay for what you need at the time.	Opening extra tills in the canteen at lunchtime and closing them again at two. Where that stops working: Opening a till takes a moment. In the cloud a new server also takes time to start, and a sudden rush can arrive before it is ready.
Error budget	The amount of failure a service is allowed under its target. If the target is 99.9% over a month, the budget is the 0.1% of that month that may go wrong.	A set number of late marks a student may have before a letter goes home. Where that stops working: A late mark is the student's fault. An error budget is spent by anything that breaks, including the team's own changes.

Term	In plain English	Think of it like this
FinOps	The practice of managing what the cloud costs, bringing engineering, finance and the business together to decide where spending is worth it.	A department deciding together how to spend a budget so that the money buys the best result and not just the cheapest. Where that stops working: A department budget is fixed once a year. Cloud spending changes by the hour, so the watching never stops.
Generator	An engine that makes electricity from fuel, started when the mains fails so the equipment can keep running for much longer than a battery can.	A caravan's portable power unit, used in a field when there is no plug. Where that stops working: A portable unit is switched on by a person. A data centre's generator has to start and take the load by itself, in seconds, and has to be tested to be trusted.
Infrastructure as code	Describing the computers, networks and settings a system needs in files that a tool can read, so the same set-up can be built again exactly.	A recipe, written down so that anyone can make the same dish, rather than one cook doing it from memory. Where that stops working: A recipe still needs a cook to check it. A file with a mistake builds the mistake exactly, every time.
kW and kWh	A kilowatt is a rate of using energy. A kilowatt-hour is the energy used by 1 kW running for one hour, which is what a bill counts.	Speed and distance. 60 miles an hour is a rate, and driving at it for an hour gives 60 miles. Where that stops working: A car's speed changes. A data hall's load is closer to steady, which is why the maths is simple.
Nines	A way of saying availability by counting the nines: 99.9% is three nines. Each extra nine cuts the time a service may be down to a tenth.	Attendance. Going from 90% to 99% attendance sounds small, and over a year it is the difference between missing weeks and missing days. Where that stops working: Attendance is counted in days. Availability is counted in minutes, and the sums get small very fast.
Pipeline	A series of automatic steps that takes a change to software, tests it, and, if it passes, releases it.	A homework routine that always runs in the same order: write it, check it, hand it in. Where that stops working: A routine you follow yourself can be skipped. A pipeline cannot, but a mistake in it is made every time.
PUE	The total power a data centre draws divided by the power its computers draw. A PUE of 1.5 means that for every 1 kW reaching the computers, 0.5 kW more goes on cooling and everything else.	The total of the school's electricity bill compared with what the classrooms' computers used. The rest goes on lights, heating and the kitchen. Where that stops working: PUE leaves things out, such as the water a site uses and how efficient the computers themselves are, so a low PUE is not the same as a green site.
Rack	A metal frame, a standard width, that holds computers and network equipment one above another, with power and cables at the back.	A bookcase for computers, where every shelf is the same width and each item takes a whole number of shelves. Where that stops working: A bookcase has no electricity to supply and no heat to take away. A rack has both, and they are often the real limit.

Term	In plain English	Think of it like this
Rack unit	The standard measure of height in a rack. One unit is 44.45 mm, and equipment is made in whole numbers of units.	A line on lined paper: everything has to take up a whole number of lines. Where that stops working: Lined paper has no weight limit. A rack has a limit on weight as well as space and power.
Redundancy	Having more than one of something that matters, so that when one fails another takes over.	A spare projector bulb, a second teacher who knows the register system and a back-up key to the music room. Where that stops working: A spare only helps if it does not break for the same reason. Two bulbs on one blown fuse are one fault.
Server	A computer whose job is to provide something to other computers, such as web pages, files or a database, when they ask.	The person in the school office who answers every phone call and hands out whatever was asked for: the timetable, the form, the key. Where that stops working: A real server answers many requests at the same moment and never needs a break, which the office cannot do.
Shared responsibility	The idea that a cloud provider looks after some things and the customer looks after others, and that the split depends on the kind of service.	A school trip: the coach company keeps the coach safe, the school decides who goes and the teacher counts heads. Where that stops working: Nobody on a trip can swap the jobs. In the cloud the customer can choose how much to hand over, but never everything.
Single point of failure	One part that, if it fails, stops the whole thing, because nothing else can do its job.	The one member of staff who knows the password to the timetable. Where that stops working: It is easy to spot in a drawing and hard to spot in real life, because many things share a hidden dependency such as one power feed or one DNS record.
SLA	A written promise from a provider to a customer about how well a service will work, and what the customer gets if it does not, which is usually a credit on a future bill.	The school's promise to parents that the office will answer within a day, and what it will do if it does not. Where that stops working: Real agreements have conditions. A headline figure applies only if you set things up the way the agreement says.
SLO	A target a team sets for itself for how well its service should work, such as 99.9% of requests answered. It is usually a little stricter than the promise it keeps.	A team's own target for the term, set a bit above what the teacher has asked for. Where that stops working: A target has no penalty attached. Missing it is a signal to the team, not a payment to a customer.
Tier	A grade from I to IV that the Uptime Institute gives to a data centre's design, based on how much redundancy it has and whether parts can be worked on without shutting it down.	A star rating for a restaurant's kitchen, covering the equipment and how it is run. Where that stops working: A star rating is a promise of quality. A Tier is not a promise of availability: Uptime says its Tier awards do not certify any availability percentage.
UPS	A battery-backed power supply that keeps equipment running for a short time when the mains fails, and cleans up dips and spikes.	The torch in the staff room that lasts long enough to get everybody to the exit. Where that stops working: A torch is for walking out. A UPS is meant to bridge the gap until the generator takes over, not to run the building.

Term	In plain English	Think of it like this
Virtual machine	A computer that exists as software, running on a real computer that can host several of them, each acting as if it were separate.	A hall that is used for assembly, an exam and a concert on different days: one building, several separate uses. Where that stops working: The three uses in the hall happen in turn. Virtual machines run at the same time and share the real computer's power between them.

The Curriculum Translator

Where it comes from	What the curriculum says	Activities that support it
Computing, national curriculum, secondary	"understand the hardware and software components that make up computer systems, and how they communicate with one another and with other systems"	Where is the cloud?, Who looks after what?
Computing, national curriculum, primary	"understand computer networks, including the internet"	Where is the cloud?
GCSE computer science, Department for Education subject content	"main and contemporary secondary storage and ways of storing data on devices"	Where is the cloud?, Who looks after what?
GCSE computer science, Department for Education subject content	"the ethical, legal and environmental impacts of digital technology on wider society"	Keep it cool, Three true stories
Department for Education response on the GCSE content, March 2025	"which may inhibit coverage of cloud computing within awarding organisations' specifications"	Where is the cloud?, Black Friday, Rent by the hour or by the year
AQA GCSE Computer Science 8525, section 3.4.5	"Explain the term cloud storage."	Where is the cloud?, Who looks after what?
AQA GCSE Computer Science 8525, section 3.8	AQA lists cloud storage among the areas that its exam questions on the ethical, legal and environmental impacts of digital technology will be taken from.	Keep it cool, Three true stories
OCR GCSE Computer Science J277, section 1.3.1	"The Cloud: remote service provision (e.g. storage, software, processing)"	Where is the cloud?, Black Friday, Rent by the hour or by the year
OCR GCSE Computer Science J277, section 1.3.1	"Concept of servers providing services"	Spread it across two, Who looks after what?
Mathematics, national curriculum, secondary	"calculate the probability of independent and dependent combined events, including using tree diagrams and other representations, and know the underlying assumptions"	How many nines?, Spread it across two, By hand or by pipeline
Science, national curriculum, secondary physics	"power as the rate of transfer of energy"	Lights out, Keep it cool, Rack and stack
Science, national curriculum, secondary physics	"comparing power ratings of appliances in watts (W, kW)"	Keep it cool, Rack and stack
Science, national curriculum, secondary physics	"comparing amounts of energy transferred (J, kJ, kW hour)"	Keep it cool, Lights out
Science, national curriculum, secondary physics	"domestic fuel bills, fuel use and costs"	Keep it cool

Where it comes from	What the curriculum says	Activities that support it
Science, national curriculum, secondary physics	"conservation of energy in a closed system, dissipation"	Rack and stack, Keep it cool, Lights out
Science, national curriculum, secondary physics	"calculating energy efficiency for any energy transfers"	Keep it cool
Curriculum for Wales, Science and Technology	"Computation is the foundation for our digital world."	Where is the cloud?, Who looks after what?
Curriculum for Excellence, Technologies	"business, computing science, digital literacy, food, textiles, craft, design, engineering, and graphics"	Eight kinds of job
Planned, first teaching from 2028	"publish a revised national curriculum in 2027 for first teaching in 2028"	Where is the cloud?
Planned, GCSEs updated for first teaching from 2029	"We will replace the computer science GCSE with a GCSE in computing"	Where is the cloud?

Be aware of the limits

- The words cloud, server, virtual and data centre appear nowhere in the national curriculum computing programmes of study or in the Department for Education's GCSE computer science content. The honest link at secondary level is the statement about how computer systems communicate with one another and with other systems, and the GCSE content on environmental impacts.
- The statement about computer networks and the internet sits in the primary programme of study. The secondary programmes do not name networks or the internet.
- The exam boards go further than the Department. AQA's specification names cloud storage, and OCR's names the cloud and servers. Neither names a data centre, and neither names virtualisation as content. Check your own board's current specification.
- The Department removed the phrase inside computers from the GCSE content because it may inhibit coverage of cloud computing. That change applies to GCSE exams from summer 2027.
- The sums in these games sit in mathematics and science, not in computing. The mathematics row is the wording on combined events and knowing the underlying assumptions, which is the assumption the spare machines game keeps coming back to: that the parts fail independently. The science rows are power, energy transfer, dissipation and efficiency. We have not mapped the electricity and thermal physics in more detail.
- The two rows from Wales and Scotland are one line each, to show that computing and technology are in those curricula too. In Scotland we could read only the general description of the technologies area and not the experiences and outcomes, and Northern Ireland is not covered.
- The last two rows are a plan, not a curriculum. The government's response to the curriculum and assessment review, published in November 2025, says it will replace the computer science GCSE with a GCSE in computing, publish a revised national curriculum in 2027 for first teaching in 2028, and update GCSEs for first teaching from 2029. Nothing has been published yet, and the response says only that it will explore a new level 3 qualification in data science and AI.
- We read the Department for Education's own documents and the current specifications from two exam boards. Check your own board's current specification, and for England the version your school currently follows.

Questions students, parents and colleagues ask

Do you need a degree to work in cloud and infrastructure?

For most of the apprenticeship standards the Career Matrix maps, no. The DevOps engineer standard, ST0825, names A levels, a Level 3 apprenticeship or another relevant qualification, relevant experience and an aptitude test as the likely entry, and lists no degree. The IT solutions technician and Information communications technician standards state no entry requirements of their own on the Skills England pages we read.

A degree matters in two places. The Level 6 integrated degree apprenticeship in digital and technology solutions, ST0119, has a degree built into it, and the Level 7 specialist standard, ST0482, says entry is likely to include a degree at 2.1 or higher in a relevant subject. We found no official source that says how many people in cloud and infrastructure jobs hold a degree, so we do not give a figure.

Is there a cloud engineer apprenticeship?

No. Skills England has no current standard with cloud engineer, data centre technician or site reliability engineer in its title. Those job titles appear inside other standards: data centre support technician and cloud technician in the Information communications technician standard, ST0973, and site reliability engineer and platform engineer in the DevOps engineer standard, ST0825. The Network engineer standard, ST0127, lists network and cloud engineer.

The government's own role framework also has no site reliability engineer, cloud engineer or platform engineer role. Its closest are infrastructure engineer, infrastructure operations engineer and DevOps engineer.

What is the hands-on, hard hat side of cloud and infrastructure?

It is the physical side: installing and keeping running the servers, cooling, cabling and backup power in a data centre. The Career Matrix calls the area Data Centre Operations and marks it as a site-based trade.

The sourced routes in are a Level 2 foundation apprenticeship, FA0004, normally for people aged 16 to 21 at the start, the Level 3 Information communications technician standard, ST0973, and the Level 3 refrigeration, air conditioning and heat pump standard, ST0322, which names data centres once in a list of places the work is done. We found no standard that covers a data centre's backup power, such as batteries and generators, and the electrician and building services standards we read do not mention data centres. Work on electrical equipment carries rules: the HSE says people must be competent to do it.

Is there a shortage of cloud and data centre workers?

We found no government figure for cloud or data centre vacancies, so we do not say there is a shortage. What we did find is partial. The Uptime Institute's 2025 survey of data centre operators worldwide found that 46% report difficulty finding qualified candidates, and that shortfalls in electrical and mechanical trades continue to be significant. It is an industry survey, not a UK one.

The government's AI Growth Zones paper projects more than ten thousand jobs from its programme. That is a projection, not a count. The Data Centre Covenant expects developers to invest in paid apprenticeships and training places, and gives no number.

What does 99.9% availability actually mean?

It means the service may be down for up to 0.1% of the time. Over a year that is about 8.76 hours, in Google's table, which assumes no planned downtime. Each extra nine cuts the allowance to a tenth: 99% allows 3.65 days a year, 99.99% allows 52.6 minutes and 99.999% allows 5.26 minutes.

It is a promise about the average, not about when. Ten minutes every day and one long outage can both come to the same figure. And the Uptime Institute says that a data centre's Tier does not certify an availability percentage, so a Tier is not a promise of nines.

Does the cloud use a lot of electricity?

It uses a lot in total, and the figures depend on who is counting. The Department for Energy Security and Net Zero estimates that data centres used 4.5 TWh, which is 2% of the electricity consumed from the grid in Great Britain in 2024, and that most of it was in the South East and London. The National Energy System Operator's estimate for 2023 is higher, 7.6 TWh, because it counts differently.

Globally, the International Energy Agency says data centres used around 1.5% of the world's electricity in 2024 and projects around 945 TWh by 2030 in its base case. For the future in Great Britain, be careful: around 73 GW of data centre projects are in the grid connection queue, and the National Energy System Operator says that pipeline is likely to significantly overstate demand.

How much water do data centres use?

Nobody can say for the whole country. The government's water resources framework says it is experiencing barriers in gaining information about data centre water consumption, and says data centres can require a large amount of water for cooling. The draft National Policy Statement says newer closed, waterless cooling can reduce or remove the need for water.

The only survey figures we found are from techUK's voluntary survey of 73 mainly commercial sites in England: 51% used waterless cooling systems and 64% used less than 10,000 cubic metres of water a year. techUK says it is not a definitive snapshot.

Is the cloud green?

It depends on what you measure. Power usage effectiveness, PUE, is the total power a site draws divided by what its computers draw. Uptime's 2025 survey found a weighted average of 1.54, and the government expects new data centres to be designed to achieve 1.3 or better.

PUE leaves out things such as the water a site uses and how efficient the computers are. And the carbon in each kilowatt-hour depends on the grid: the government's 2026 factor for UK electricity is 0.13096 kg of CO_2e per kWh, which is a reporting average and not a live reading.

Who owns the cloud?

A few very large companies supply most of it. The Competition and Markets Authority found that Microsoft and AWS each had between 30 and 40% of infrastructure as a service in 2024, across the UK and the European Economic Area, with Google lower, and that less than 1% of customers switch provider each year. Its final decision was that competition is not working well in cloud services.

Ofcom's earlier figure, a combined 70 to 80% for the two in the UK in 2022, measures a different thing in a different year, so do not set them side by side. In March 2026 the CMA said it would launch an investigation into Microsoft's business software ecosystem.

What happens when a data centre fails?

It depends on how, and the three stories on this page are three different ways. In France in 2021, the investigators say a fire started in the rooms with the batteries and the uninterruptible power supply and destroyed a building, and that the cause was not established when they reported. In London in 2022, Google says several cooling systems failed together in a heatwave. At AWS in 2025, the company says a bug in the automatic system that manages DNS records left one record empty.

In each, the company or the investigators say what changed afterwards. We show only what the documents say, and we do not add a cause that they do not give.

Who is responsible for security in the cloud?

Both sides, and the split depends on the kind of service. The National Cyber Security Centre says there is a joint responsibility for the security and availability of data and workloads in a cloud service, shared between the provider and the customer.

It says the provider will almost always be responsible for the availability and physical protections of its servers, and that the customer is always responsible for setting up the services it chooses securely and for deciding what data to store in them. AWS and Microsoft each describe their own model in their own words, and those are the companies' descriptions, not independent ones.

Do students need a vendor certificate?

Employers name them, and the two beginner ones are AWS Certified Cloud Practitioner and Microsoft Azure Fundamentals. AWS says its exam is designed for people new to cloud who may not have an IT background, and lists the exam as 90 minutes with 65 questions. Microsoft lists Azure Fundamentals at beginner level with a 45 minute exam.

Passing either one does not promise a job. Fees and exam details change, so send students to the issuer's own page.

Where does this fit in the curriculum?

Honestly, in a few places and thinly. The words cloud, server, virtual and data centre appear nowhere in the national curriculum computing programmes of study or in the Department for Education's GCSE computer science content. AQA's specification asks students to explain the term cloud storage, and OCR's lists the cloud and servers providing services. Neither names a data centre.

The government has said it will replace the computer science GCSE with a GCSE in computing and publish a revised national curriculum in 2027 for first teaching in 2028. That is a plan, and nothing has been published yet.

Are data centres part of the country's critical infrastructure?

In the UK, yes. On 12 September 2024 the Technology Secretary announced that UK data centres are classed as Critical National Infrastructure, the first such designation in almost a decade. The government said it allows

it to support the sector in the event of critical incidents.

That describes the government's intent at the time of the announcement, and it is not an audit of what has happened since.

Sources and how we checked them

Each source was opened and read on 5 November 2026. Where a source warns about its own figures, we pass the warning on.