Uptime Institute INTELLIGENCE

RISK & RESILIENCY

UI Intelligence report 45

Data center security: Reassessing physical, human and digital risks

Authors

Rhonda Ascierto, Vice President of Research, Uptime Institute
Todd Traver, Vice President for IT Optimization and Strategy, Uptime Institute

Data center operators routinely maintain multiple physical perimeters against intruders and regulate the activities of the people inside. However, cloud computing and increased remote monitoring and automation bring new security challenges. Human and digital vulnerabilities have expanded the attack surface for many data centers.

30-45 minutes to read

UII-45 v1.0PM published March 10, 2021 last updated March 10, 2021

This Uptime Institute Intelligence report includes:

ABOUT UPTIME INSTITUTE INTELLIGENCE

Uptime Institute Intelligence is an independent unit of Uptime Institute dedicated to identifying, analyzing and explaining the trends, technologies, operational practices and changing business models of the mission-critical infrastructure industry. For more about Uptime Institute Intelligence, visit uptimeinstitute.com/ui-intelligence or contact research@uptimeinstitute.com.

EXECUTIVE SUMMARY

The data center sector has made considerable effort to secure physical facilities. Infrastructure and access-control measures, as well as staff procedures, are routine in most, if not all, mission-critical data centers.

However, the likelihood of sabotage has grown, the surface area for attacks has expanded, and the methods used by intruders are increasingly sophisticated. Given the potential impact of a serious physical incursion, and in light of recent current events, data center managers should revisit their security approaches.

KEY FINDINGS

- Even with common tools and tactics, there is no singular approach or methodology for physical data center security. Every site is different.
- Data center owners and operators should continue to invest in strong physical security; a lack of incidents across the industry is a sign of success, and vigilance and investment should not be relaxed. The likelihood of physical breaches, unauthorized access to information, and the destruction of or tampering with data and interrupting services is higher than ever before.
- Cloud computing, internet connections and remote management technologies have introduced new threats for facility management teams. Data center management software and services, as well as all internet protocol (IP)-enabled equipment, should be closely assessed for security risks.
- The COVID-19 pandemic has increased the attack surface for many data centers globally due to on-site staff disruption and greater use of remote operations.
- Compliant does not mean secure. Cyberattacks are becoming more sophisticated. Intruders often use a combination of digital trails and social engineering to gain trust.
- Cybersecurity training, tools and processes are used to reduce security risks, as are threat assessment and penetration services.
- Threat models to identify vulnerabilities and prioritize mitigation efforts consider the unique risk surface of a data center by assessing all elements of access, including digital and human.
- A single undetected weak area in physical or digital security can compromise an entire data center.

Introduction

"Sounds like war! It would be a pity if someone with explosives training were to pay a visit to some AWS Data Centers — the locations of which are public knowledge."

User message posted on the social media service Parler
January 2021

While online threats against data centers are nothing new, the post above (and similar others) is a disquieting reminder of their increased frequency due to current events, as well as the growing criticality of IT. These recent threats are just one of the ways cloud computing and social media have changed the risk profile of physical data center security.

The growing use of remote monitoring and automation, as well as advances in cybersecurity attacks, are adding to the ways a threat might be executed. The attack surface of the data center — the physical, human and digital ways that security can be breached — is expanding. The challenge for data center operators is identifying new weaknesses, developing approaches to secure them, updating security processes and protocols, then continuously testing them.

In this report, Uptime Institute examines the growing scope of physical data center security, from infrastructure and procedural tactics for different data center types and risk profiles to digital vulnerabilities.

Risk perception vs. reality

Data center operators often maintain a low profile for their facilities out of necessity — security risks increase with greater public awareness. Yet as the critical role of data centers grows, so does the public's awareness.

High-profile outages of cloud, internet and other digital services have created news headlines for many data centers in recent years. Hostility to 5G has led to physical attacks on network infrastructure, including the suicide bombing near an AT&T facility in the US in late 2020. Governments are also drawing attention to data centers by investigating the resiliency of certain critical digital services.

In spite of the secrecy, information about data center locations can often be found on the internet. In some cases, the control systems of critical mechanical and electrical data center systems can even be directly accessed (see **Insecurity of digital systems**).

Terrorist groups, hackers, hostile states and even environmental campaigners could act against data centers. There have been examples where IT operations have been disrupted by remote access to critical infrastructure. In February 2021, for example, a hacker attempted to poison the water supply of a Florida (US) city by remotely accessing a software system and adjusting the sodium hydroxide concentration to a dangerous level. (The attack was detected, and harm avoided.) There are other examples, including some high-profile attacks at power plants.

In the 27 years Uptime Institute has been collecting AIRs (abnormal incident report) data, we have not recorded a single instance of a data center outage being caused by sabotage (although we have recorded many human error-related incidents, which could have been caused by unauthorized actions). Many physical security incidents (as opposed to general IT attacks, which are not routinely tracked by Uptime Institute) are accidental breaches of guidelines or policy, or a casual disregard of rules. However, this should not encourage complacency. The threat is real.

The likelihood of physical breaches, unauthorized access to information, and the destruction of or tampering with data and services is higher than ever before. In some geographies, the safety of people on-site is a concern.

Compliance vs. security mitigation

The security of IT systems used to manage and operate data centers, including digital control panels of certain infrastructure equipment, requires special attention. However, it is not uncommon for physical data center-related IT security to be viewed by executive management as a cost center and the focus is typically on compliance. While security and compliance are both essential components, they are not synonymous.

Effective security mitigation is an ongoing process that requires tools, technical systems and staff training. Compliance involves assessing security practices to ensure they meet specific legislation, regulation or standards based on best practices. Security is part of compliance but typically involves measures that go beyond it.

Operators of hyperscale cloud, internet, financial services and other data centers, for example, routinely employ specialist security penetration testers tasked with evading existing security measures to identify vulnerabilities. In some cases, these specialist firms are hired not by data center or IT management but by another department or by customers of the data center, effectively increasing the challenge of detecting them. These tests should not just be a box-checking exercise for compliance auditors; rather, they should be used to identify weak areas of security and, if necessary, as justification for investments in hardening a data center, either physically or digitally.

Some operators maintain relationships with security specialists to develop threat models that include both site-specific scenarios and threats resulting from potential changes to their business risk. Their plans are periodically reviewed to address new considerations that may evolve over the life of the data center (e.g., rising crime rate in the area) or any sudden developments that increase security requirements, ranging from nearby roadwork or construction to political instability in the region.

Managers should also consider compensating controls, which are alternative additional approaches to mitigating threats. For example, a biometric system may prevent unauthorized access to a computer room but can be thwarted by a brick thrown through a plain-glass window or by the use of a hostage as a "human key." Security approaches should ensure that all possible methods of unauthorized access are protected against, or at least assessed (in this example, mitigation would include windowless computer rooms and scanning all who enter for potential weapons).

It is also common for security to be implemented almost as an afterthought, following a data center's construction. Most security budgets are focused on the critical power and cooling infrastructure, though not necessarily on the operation and control of these systems. Physical security should be considered at an early stage of facility development; local site conditions can dictate tactics and technologies.

Risk profiles by data center type

Calculating the risk of a physical intrusion, or assessing how much should be spent on security, is not straightforward. The level of security required by a data center can depend on the business activities and workloads running in it. Some organizations are targets because of their business, customers or current affairs. The location of the data center and who knows of its existence also play a role.

Even with common tools and tactics, there is no singular approach to security. Organizations with multiple data centers typically create a unique threat/security matrix for each facility. They may employ similar strategies, technologies and approaches, but the underlying threats will differ.

The operators of different types of data centers also usually have varying abilities to isolate risk. Hyperscale cloud and internet data centers, for example, are often sited in remote locations where power is plentiful and inexpensive. Similar to government data centers that house sensitive information, security assessments and mitigation efforts are well-funded, ongoing and managed by specialist experts. Data centers of these types are more able to identify and isolate risk than other types of data centers.

A multi-tenant colocation (colo) facility is shared by multiple customers and often sited near a metropolitan area. This means a relatively high number of people are authorized to access the facility. While such data centers may sometimes employ security specialists, the number of visitors and their higher profile as commercial operators means their risks are greater.

The ability of privately owned enterprise data centers to isolate risk can vary considerably. These facilities are not shared and have greater freedom in location choice than a colo. Yet enterprise data centers can sometimes face budgetary pressures that stymie new or expanded security approaches. In the absence of a breach, managers often have to formalize a compelling business case for additional spending, which can be a deterrent.

Enterprise data centers housed in mixed-use facilities, rather than a dedicated purpose-built building, typically have the lowest ability to isolate risk. (In mixed-use facilities, data centers must share noncritical spaces and resources with other building occupants.) Depending on the existing infrastructure layout, for example, it may be inexpensive to create separate entrances to IT and building control functions to mitigate the risk from the greater number of visitors to the site, but the cost of separate and independent mechanical, electrical and control systems can be prohibitive. Typically, many of these systems are integrated into the existing infrastructure of the building, making them harder to isolate. Multipurpose facilities also often have separate teams responsible for the security of the data center versus the rest of the building. Table 1 compares the ability to isolate risk by data center type.

Table 1. **Ability to isolate risk by data center type**

Data center type	Ability to isolate risk	Typical focus areas of security
Enterprise: Mixed-use	Very low	Rack, computer room
Multi-tenant colocation	Low	Rack, computer room
Enterprise: Dedicated	Variable	Computer room, building
Hyperscale and high-security government	High	Computer room, building, site

Source: Uptime Institute Intelligence 2021

Edge data centers embedded in commercial buildings face some of the same challenges as mixed-use facilities. By their nature, edge data centers are near to users — in some cases, the general public — and may be in relatively unprotected physical locations (such as next to a busy street). This increases their vulnerability and therefore may require the use of safeguards in the IT itself (such as tamper-proof protective measures and extra encryption). At the same time, the small size and self-contained nature of edge data centers can make them easier to shield. Many are also likely to be hardened by distributed resiliency among multiple sites. However, this will not replace the need for physical hardening and protective processes.

Even among similar data center types, security approaches and implementations will vary, driven by circumstances and the risk tolerance of the business or customers. Ongoing security assessments should dictate the appropriate procedural, infrastructure, technology and training requirements for every data center.

Cloud resiliency as a security measure?

The use of multisite, cloud-based resiliency has opened up possibilities such as eliminating secondary backup data centers and moving to disaster recovery as a (cloud-based) service. In theory, this means protecting single sites and equipment may be less critical than it once was. However, our research shows that distributed or cloud-based resiliency does not necessarily eliminate the risks or consequences of single-site failure — and may, in fact, increase them (e.g., if infected or corrupted data is transmitted across multiple sites).

Cloud-based replication does make it easier and cheaper to store data and run applications in multiple locations. While this does not make it less important that each data center is adequately protected, it does provide a level of insurance in the event of a serious incursion.

Most operators, including those of public cloud facilities, still invest heavily in site-based resiliency — and, of course, in site security.

Physical perimeter and staff security

Many data centers are physically secured in two primary ways: by staff carrying out procedures and by physical infrastructure barriers, which may involve technology (such as biometric identification systems).

In some cases, such as government and military installations, or where the business risk from an incursion is atypically high, the physical security measures may include moats, razor wire-topped fences, armed security guards, and so on. The other end of the spectrum includes small colocation data centers in which security consists of a single receptionist sitting in a lobby area outside of the computer room.

Staff and procedural measures and physical infrastructure measures are discussed in the following sections. These approaches are separated into four areas: the data center site, the data center building, the computer room, and customer racks or cages. The security approaches for each area are organized to reflect escalating levels of risk mitigation, from basic measures to advanced capabilities.

Securing the site

The site of a data center encompasses all of the IT racks (and customer cages in colos), computer rooms, the building, and external critical infrastructure such as cooling towers, electrical substations and network feeds. The building and computer room can be fairly well-protected by preventing unwanted access to the site and limiting external risks.

SECURING THE SITE

BASIC

Staff and procedures

- Stationed security guards monitor CCTV cameras, ID/badge systems and a call box (direct special-purpose phone line).
- Security guards monitor utility services and equipment outside the site perimeter, including power, water and communications.
- Security guards conduct regular site patrols.
- Security guards receive advanced training on diffusing threats, using restraint techniques, etc. (similar to the training required for government security credentials).
- Security guards use an EVMS that includes software to track and manage site access.
- Security guards are armed.
- Security guards have active-shooter training (including lockdown procedures) and conduct rehearsals regularly.
- All vehicles are checked, including the undercarriage and trunk/boot.
- Aerial drones surveil the site, triggered by motion alarms or manual alerts.
- Internet social media threat monitoring and alerting procedures are followed.
- Visitor vehicles are parked at a remote guard house and security staff escort visitors.
- Visitors are scanned by a metal detector at the site perimeter.
- All visitor belongings are locked in their vehicle's trunk or a secure storage area.

Physical infrastructure

- Site entrance(s) are gated.
- Site grounds are monitored using CCTV cameras and infrared sensors.
- Multifactor authentication is required for access devices (badges, unique ID cards, biometrics, pin pads, mobile-phone text message, etc.).
- Underground network and electrical utility entrances are diverse and secured physically (often by gated, locked cabinets).
- Site has no signage indicating the name of the organization.
- Site includes perimeter security fences with razor wire.
- Site has a vehicular trap, including fences, moats and ditches.
- Concrete walls enclose cooling towers and critical electrical and mechanical equipment.
- Site incorporates earthen perimeter berms, trenches and landscaping.
- Retractable car-trap bollards secure the site entrance.

ADVANCED

EVMS, Electronic visitor management system
CCTV, Closed-circuit television
ID, Identification

Securing the building

The data center building contains IT racks, customer cages, computer rooms (including networks), and internal critical electrical and mechanical infrastructure. Securing the building protects computer rooms by restricting access through the use of advance approvals, escorting visitors, continuous monitoring and other measures.

SECURING THE BUILDING

BASIC → ADVANCED

Staff and procedures

- Facility has 24/7 staffing (building security guards, lobby reception staff, etc.).
- Security guards conduct regular foot patrols and monitor CCTV.
- Visitors are required to provide government-issued ID and to review and sign their acceptance of data center rules.
- Vendors and visitors must be sponsored and escorted.
- Visitors are scanned by a metal detector at the building entrance.
- Security guards record the serial number of computers entering and exiting the building.
- Multifactor authentication is required for access devices (badges, unique ID cards, biometrics, pin pads, mobile-phone text message, etc.).
- Pre-notification is required for all deliveries (unexpected/unscheduled deliveries are rejected).
- Security staff and loading dock personnel work together, logging approved packages, notifying the recipients and storing the packages in a secure place to prevent tampering (until collected).

Physical infrastructure

- Entrance to lobby area has reinforced or masonry walls and/or bulletproof glass.
- Lobby area is separate and secure from building (controlled by security/staff).
- Security guard station at entrance is protected by bulletproof glass.
- Site grounds are continuously monitored using CCTV and infrared cameras.
- All packages are scanned at the point of delivery (entrance or loading dock).
- Building entrance (past the lobby) has people-trap with volumetric sensing or security monitoring.
- Badge, pin and biometric readers are used to validate authorized access.
- Interior exit doors have alarms and security is notified if opened for more than a few seconds.
- Computer room and critical spaces have no windows.
- Site uses cell phone and global positioning system signal jammers.
- Exit doors have no handle on exterior side.
- Loading dock doors are secured, open only from inside and are monitored by security.
- Delivery storage areas are separate, isolated and recipient-specific.
- Exterior walls are lined with aramid (such as Kevlar) for additional physical protection.
- Exterior walls are lined with metal cladding for EMF protection.

EVMS, Electronic visitor management system
CCTV, Closed-circuit television
ID, Identification
EMF, Electromagnetic frequency

Uptime Institute | INTELLIGENCE

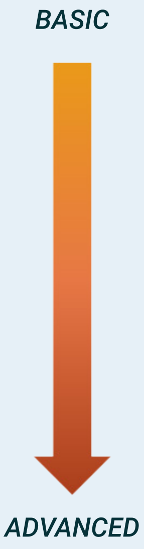
Securing the computer room

The computer room houses the physical IT assets, in racks and customer cages. It is the most vulnerable area, with unparalleled opportunities to damage IT operations and network connectivity, either intentionally or inadvertently. Access should be restricted to pre-approved, escorted and sponsored visitors or vendors only.

SECURING THE COMPUTER ROOM

Staff and procedures

- Security personnel authorize and maintain an access list to computer room.
- Multifactor authentication is required for access devices (badges, unique ID cards, biometrics, pin pads, mobile-phone text message, etc.).
- Access is physically monitored by security staff.
- Entrance has people-trap with volumetric sensing or security monitoring (security staff is alerted when the first door is opened, and the occupant is then monitored using a CCTV camera or via windows).



Physical infrastructure

- Site grounds are continuously monitored and recorded using CCTV and infrared cameras.
- Facility has double interlock EPO buttons, which are monitored by CCTV
- Exit doors have no handle on exterior side.
- Interior exit doors have alarms and security is notified if opened more than a few seconds.
- Site uses motion sensors, with additional video surveillance of entrances and exits.
- Raised flooring and ceiling tiles are secured with screws.
- Motion sensors are installed above the ceiling and below the raised floor.

CCTV, Closed-circuit television
ID, Identification
EPO, Emergency power off

Uptime Institute | INTELLIGENCE

Securing the rack or customer cage

Racks contain the physical IT assets (servers, storage, routers, etc.) and data. Dedicated enterprise and hyperscale/government facilities typically provide uniform security to all racks.

In colos, racks are often grouped into separate caged customer areas. The way customer spaces are secured will vary by the type of rack or cage. For example, some customer spaces can be free-standing, all-in-one micro data centers (which look like cabinets) with bespoke security. More commonly, wire-framed cages are used; some have double mesh enclosures that prevent passage of items such as a compact storage device (e.g., a thumb drive). Mesh cages typically include a ceiling; in raised floor environments, mesh should extend to the facility floor. In some data centers, cages are made of higher grades of steel at customer request. Access is restricted to the cage's customer only.

SECURING THE RACK OR CUSTOMER CAGE

BASIC → ADVANCED

Staff and procedures

- Motion sensor alarms notify security and customer.
- Access list to cage and racks is maintained and authorized by customer.
- Access tracking and reporting is provided to the customer.
- Equipment carries no customer identifiers, system server name or network IP information labels.

Physical infrastructure

- Multifactor authentication technology for access devices is installed on cage and racks.
- Raised flooring and ceiling tiles are secured with screws.
- Motion sensors are installed above the ceiling and below the raised floor.
- Motion sensors with additional video surveillance are installed inside customer cages.

IP, Internet protocol

UptimeInstitute | INTELLIGENCE

The human factor

Human error accounts for most data center incidents. Some even say all data center incidents are caused by human error, often the result of managers not providing adequate training. In the same way that someone may accidentally switch off a critical system, people can unintentionally act, or fail to act, in ways that allow security breaches.

There are increasingly sophisticated technologies to detect potential human breaches, ranging from aerial drone surveillance to motion software that can recognize the gait of authorized security guards. Yet most breaches are far more basic — staff reusing passwords or authorized people being duped into providing valuable information.

While fully isolating human risk is impossible, it can be minimized by training, tools and processes. Risk from humans can change with circumstances. Current events can elevate risk, as can local activity such as nearby construction; new vendors, customers or partners; and staff changes.

The insider threat

Uptime Institute Members say one of their most vexing security concerns is the insider threat — authorized staff, vendors or visitors acting with malicious intent.

Trusted individuals inside a facility can harm operations in a variety of ways. In extreme examples, they could power down servers and other equipment, damage network equipment, cut fiber paths, or steal data from servers or wipe the associated storage. Unfortunately, data centers cannot simply screen for trusted individuals with bad intent.

Most operators conduct background checks. Most have policies for different levels of access. Some may insist that all visitors have security escorts, and many have policies that prevent tailgating (physically following an authorized person through a door to gain access). Many have policies to limit the use of portable memory devices in computer rooms to only authorized work; some destroy them once the work is complete, and some insist that only specific computers assigned to specific worktables can be used.

Yet vulnerabilities exist. The use of single-source authentication of identification (ID), for example, can lead to the sharing of access cards and other unintended consequences. While some ID cards and badges have measures, such as encryption, to prevent them being copied, they can be cloned using specialist devices.

The ability for data centers to protect against insider threats can depend on the business, budget and other factors. It is easier and requires less effort, for example, for smaller organizations to focus on defense than for large corporations with multiple lines of business and many staff members to do so (although larger organizations have larger budgets).

The COVID-19 pandemic increased the risk for many data centers, at least temporarily. Some of the usual on-site staff were replaced by others, and routines were changed. When this happens, security and vetting procedures can be more successfully evaded.

Circumstances that are impossible to fully control, such as the insider threat, are typically mitigated against by adding layers of security. Multifactor authentication can significantly harden ingress and egress access.

Social engineering

Human psychology tactics are increasingly being used to trick authorized people into providing sensitive information. Social engineering, using deception to obtain unauthorized data or access, is becoming increasingly sophisticated. Tactics can include a mix of digital and physical reconnaissance.

The simplest approaches are often the most effective, such as manipulating people using phone or email, and using information available to the public (for example, on the internet). Social engineering is a concern for all businesses, particularly those with mission-critical infrastructure. Automated security systems can be used to detect anomalies in communications, such as email phishing campaigns on staff and visitors.

However, even routine communication can be exploited by hackers. For example, the host names derived from the headers of an email may contain information about the IP address of the computer that sent the email, such as its geographic location. Further information about, say, a data center employee can be obtained using online information (social media, typically), which can then be used for social manipulation — such as posing as a trusted source (spoofing caller IDs or creating unauthorized security certificates for a web domain, for example), tricking an employee into providing sensitive information. By surveilling employees, either physically or online, hackers can also obtain useful information at places they visit, such as credit card information used at a restaurant (by exploiting a vulnerability in the restaurant's digital system, for example). Hackers often gain trust by combining information gleaned from chasing digital trails with social engineering tactics.

Reviews of policies and procedures, including separation of duties, are recommended. There are also numerous cybersecurity software and training tools to minimize the scope for social engineering and unauthorized access. Some data center operations use automated open-source intelligence (OSInt) software to scan social media and the internet for mentions of keywords, such as their organization's name, associated with terror-related language. Some use automated cybersecurity tools to conduct open-source reconnaissance of exposed critical equipment and digital assets.

Insecurity of digital systems

While corporate IT networks and equipment are often protected by firewalls and credentials (if not always adequately), this may not be the case with data center infrastructure equipment, creating an open backdoor for access.

As data centers are increasingly automated and are monitored and managed using DCIM (data center infrastructure management) systems, and as suppliers seek to offer online diagnostics and control, more equipment can be accessed remotely. At least 90% of all uninterruptible power systems (UPSs) over 50 kVA (50 kilovolt-amperes or about 50 kilowatts, which is a small to medium-sized UPS) have IP addresses and can be managed remotely using the industry standard protocol, SNMP (simple network management protocol). Many power distribution units (PDUs) are IP addressable, as are many other items and types of equipment. At the same time, more facility-level data is being integrated and analyzed, sometimes using artificial intelligence (AI), in cloud environments.

All of this raises the possibility that this equipment can be discovered and hacked. The consequences could be dire. While manufacturers have put security into their devices (such as passwords), it can be lightweight, and very often the default codes are never changed. The ability to turn a UPS off remotely may be blocked, but that does not mean other critical settings that would trigger a power down are also blocked (a **system-of-systems attack**, for example).

Also, advances in hacking tools and techniques mean that some access controls for digital systems (that is, credentials for authorization, such as usernames and passwords) can be circumvented. Data center operators using biometrics, in particular, should not rely on vendors' claims that credential stores (whereby biometric information is linked to access privileges) are impenetrable.

While there is no simple solution and no single tool to achieving a sound cybersecurity approach, it is clear that data center facilities management needs to work closely with the organization's IT security team on an ongoing basis.

Malware and SNMP-based network management systems

Described by Microsoft as "the largest and most sophisticated attack the world has ever seen," the breach of the network and application monitoring software SolarWinds was possible because of a vulnerability in its software library. From there, hackers inserted malware that was used (via a SolarWinds software update to customers) to steal access credentials. (Malware is software designed to damage, disrupt or gain unauthorized access to computer systems.) The hackers then used this information to remotely access the systems of other SolarWinds customers, including nine US government entities and dozens of private companies, including Cisco, Intel, Microsoft, Nvidia, VMware and Belkin. Because it appeared to be fully authorized, the malware went undetected for three months before SolarWinds discovered it in late 2020.

The attack is relevant to data centers not least because the SolarWinds platform is used to provide many core network administration and configuration services in some large data centers. Also, SolarWinds is a broad-based SNMP (simple network management protocol) network management system — and SNMP is a protocol that is commonly used in most data centers. Conceivably, SNMP-enabled uninterruptible power supplies, power distribution units and other equipment of the kind commonly used by data center management and other control systems could also have malware hidden and ready to be activated. Other communication protocols commonly used in data center equipment and rack-level devices include BACnet (building automation and control networks) and Modbus — all of which could create malware pathways.

Common causes of insecurity

In what ways could a data center's management overlook the security of control system interfaces for power, cooling and other critical equipment?

Often there are multiple reasons for inadequate protection. The installer or user may be insufficiently trained in security best practices, and the system or equipment is often not properly integrated into the organization's larger, professionalized IT security structure.

Older equipment and technology that is still being used (often called legacy), in particular, can be left exposed online. Assignments of responsibility for older assets can fade, not least because the scenarios for the end of their life tend to be problematic. Upgrading, if it's even possible, costs time and requires process change, as well as migration to a newer system. Affordable extended vendor support can also be lacking.

The default (built-in) security of legacy control systems is also outdated; they were not designed with cybersecurity in mind. They are also not always routinely patched, either because they are overlooked, inadequately tested and assessed, or managers are unaware that they are connected to the corporate network or the internet.

Yet even new equipment, including servers, can be exposed because of a mundane mistake (e.g., a logging error or process error in an asset management or similar system).

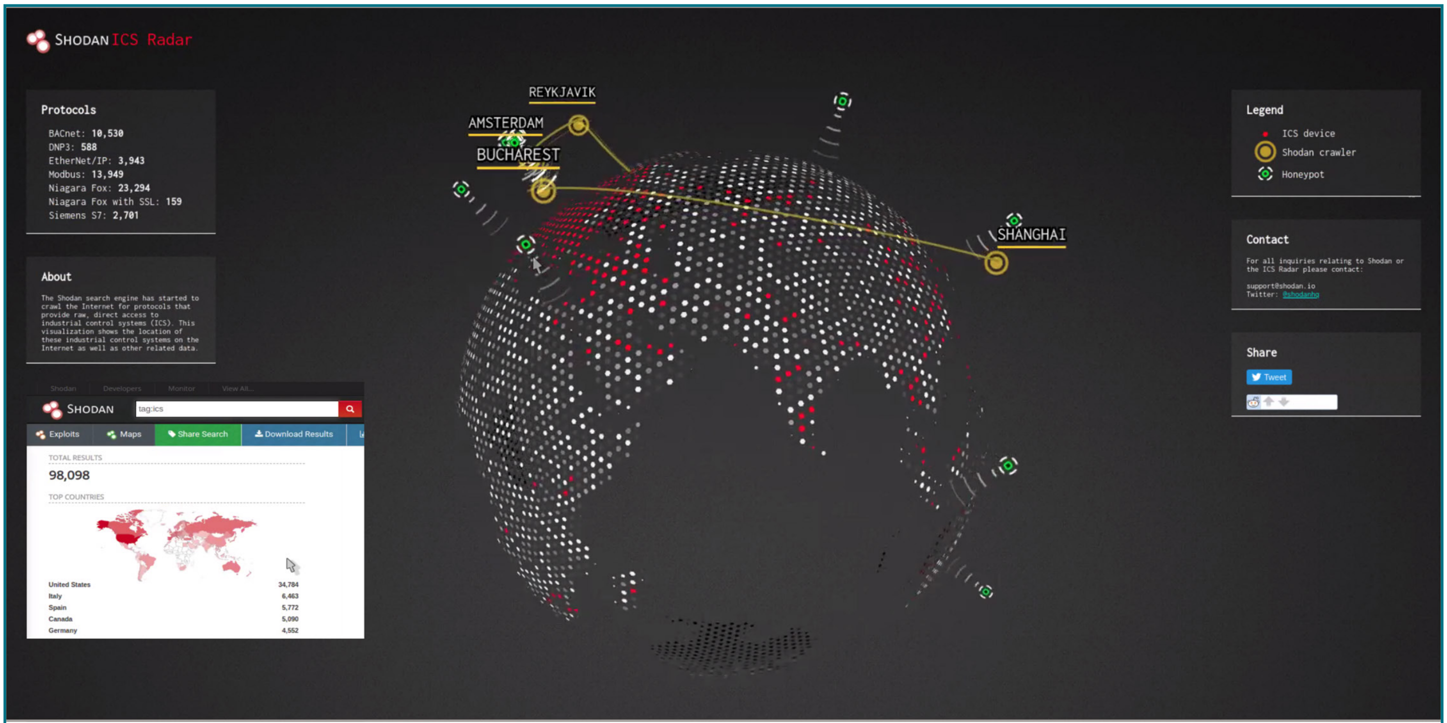
Online exposure of internal assets

How widespread is the problem of insecure facility assets? Our research of vulnerable systems on the open internet suggests it is not uncommon.

For close to a decade, the website Shodan has been used by hackers, benevolent and malevolent, to search for targets. Instead of fetching results that are webpages, Shodan crawls the internet for devices and industrial control systems (ICSs) that are connected to the internet but exposed.

Shodan and similar search engine websites (BinaryEdge, Censys and others) provide a compendium of port-scan data (locating open ports, which are a path to attack) on the internet. Expert users identify interesting characteristics about certain systems and set out to gain as much access as they can. Automation tools make the process more efficient, speeding up and also expanding what is possible for an exploit (e.g., by defeating login safeguards).

In a recent demonstration of Shodan, the cybersecurity firm Phobos Group showed more than 98,000 ICSs exposed globally, including data center equipment and devices. Phobos quickly discovered access to the login screens of control systems for most major data center equipment providers. In Figure 1 (as in all figures), screenshots of aggregate search results are shown to ensure privacy.



Protocols

BACnet: 10,530
DNP3: 588
EtherNet/IP: 3,943
Modbus: 13,949
Niagara Fox: 23,294
Niagara Fox with SSL: 159
Siemens S7: 2,701

About

The Shodan search engine has started to crawl the Internet for protocols that provide raw, direct access to industrial control systems (ICS). This visualization shows the location of these industrial control systems on the Internet as well as other related data.

TOTAL RESULTS
98,098

TOP COUNTRIES

United States	34,784
Italy	6,463
Spain	5,772
Canada	5,090
Germany	4,552

Legend	
●	ICS device
◉	Shodan crawler
◎	Honeypot

Contact

For all inquiries relating to Shodan or the ICS Radar please contact:

support@shodan.io
Twitter: @shodanhq

Source: Shodan screenshots courtesy of Phobos Group, February 2021 (composite image)

Figure 1. **Nearly 100,000 industrial control systems exposed**

The login process itself is highly problematic. Sometimes installers or users do not change the default credentials, which can be found online. During our demonstration, for example, Phobos used a default login to gain access to the control system for cooling units supplied by a widely used data center equipment vendor. If this exercise were carried out by a genuine intruder, they would be able to change setpoint temperatures and alarms.

Users' customized login credentials can be obtained from a data breach of one service and then used to try to log into another service, a type of cyberattack known as credential stuffing. The availability of lists of credentials has proliferated, and automated credential-stuffing tools have become more sophisticated, using bots to thwart traditional login protections. (Data breaches can happen without leaving any trace in corporate systems and can go undetected.)

Cybersecurity threat models

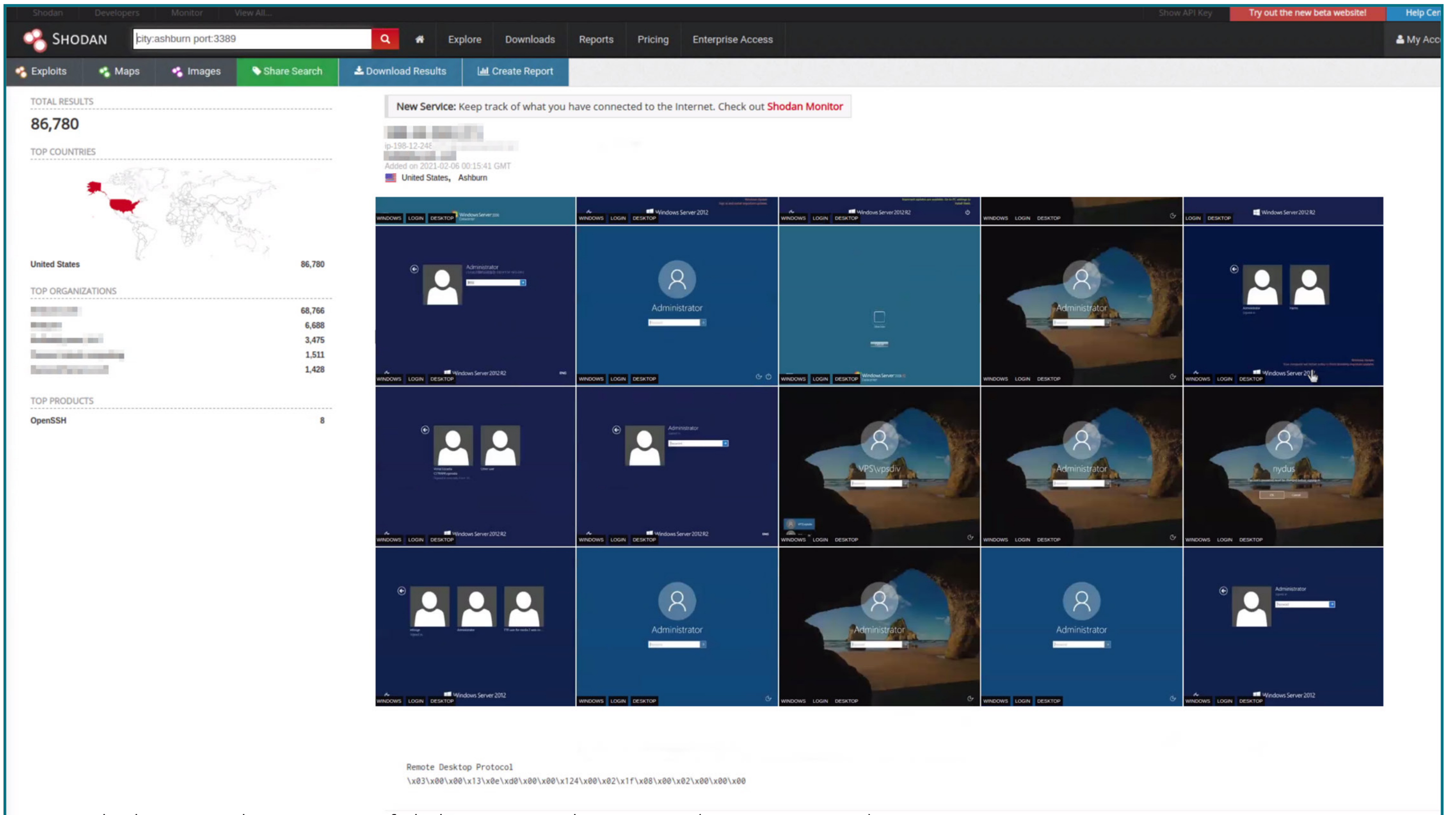
Cybersecurity of data center control systems and other internet protocol (IP)-enabled assets is multilayered and requires a combination of strategies. Specialists recommend creating a threat model using a structured process — presented in plain language — to identify potential threats and vulnerabilities and to prioritize mitigation efforts. Threat models should address broad questions such as, Who would break into here, and why? Who are our customers? Are they a target?

Threat models can determine a data center's risk surface by assessing all elements of how access is authorized. For example, in wholesale facilities, an operator and the lessor will need to communicate how access will be granted onto the site, into the building and so on.

This simple process of communicating — the host names derived from the headers of an email, for example — can introduce a risk surface. The IP address can mean that information discoverable online (using open-source intelligence or OSInt techniques) includes the lessee's or lessor's name and address and satellite-image "street" views (used by hackers for security reconnaissance). Information about the data center and its staff can be obtained using social engineering tactics, and/or by searching for exposed control and other systems online. As recent major attacks show, exploits can be disguised and go undiscovered for months.

As exploits of critical infrastructure in recent years have shown, control system interfaces may be the primary targets — but access to them is often through another system. Using the Shodan tool, the security company Phobos searched for exposed remote desktops, which can then provide access to multiple systems. This can be particularly troubling if a control system is accessible through a remote desktop and if the user employs the same or similar passwords across systems.

There are many remote desktops exposed online. As Figure 2 shows, in a recent Shodan search, over 86,700 remote desktops were exposed in the US city of Ashburn, Virginia, alone (a city known as the world's data center capital) — including a set of addresses for a major global data center wholesale capacity provider (not shown).



city:ashburn port:3389

TOTAL RESULTS

86,780

TOP COUNTRIES

United States	86,780

TOP ORGANIZATIONS

[redacted]	68,766
[redacted]	6,688
[redacted]	3,475
[redacted]	1,511
[redacted]	1,428

TOP PRODUCTS

OpenSSH	8

New Service: Keep track of what you have connected to the Internet. Check out Shodan Monitor

[redacted]
ip-198-12-248[redacted]
[redacted]
Added on 2021-02-06 00:15:41 GMT
United States, Ashburn

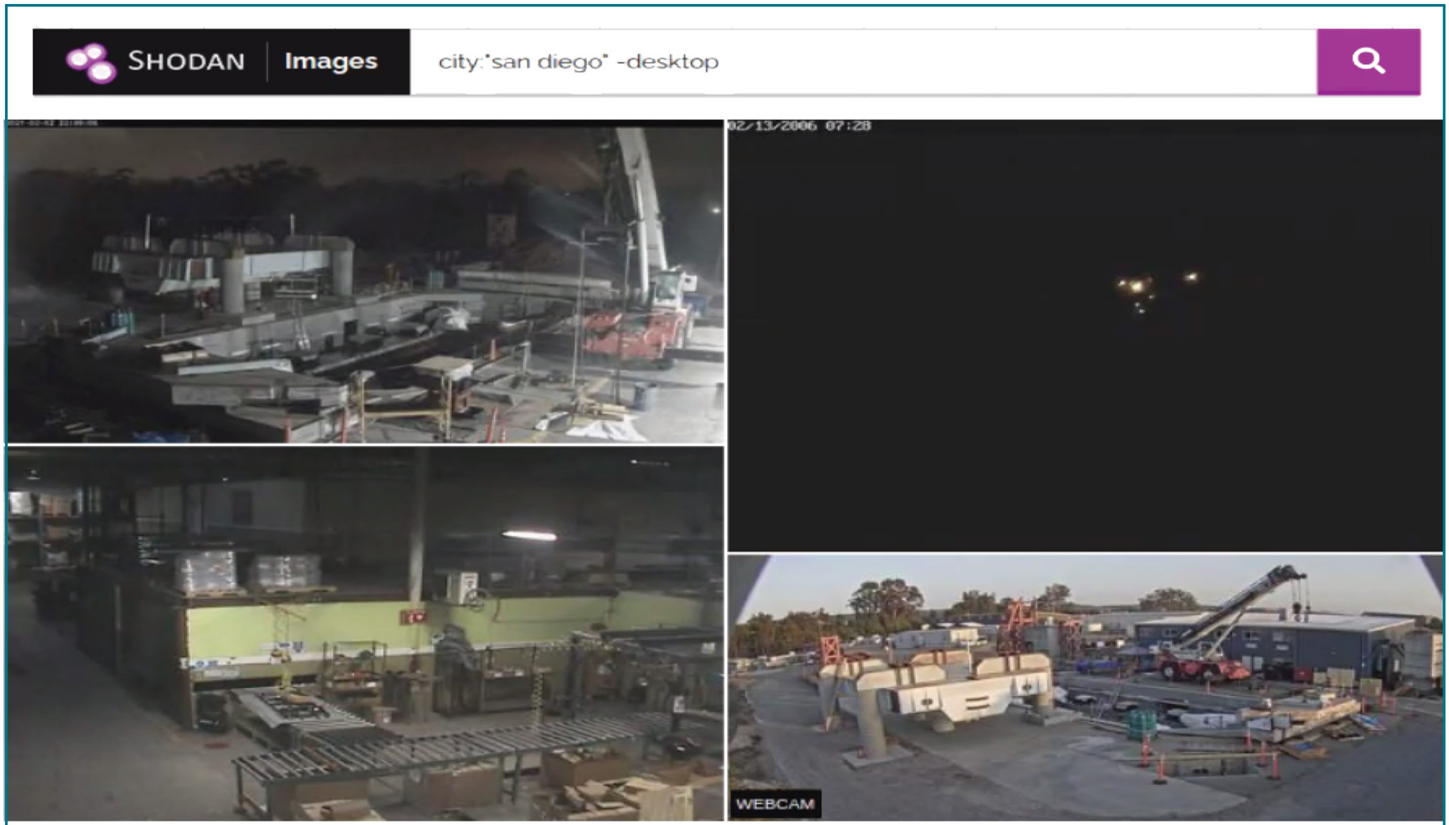
```
Remote Desktop Protocol
\x03\x00\x00\x13\x0e\xd0\x00\x00\x124\x00\x02\x1f\x08\x00\x02\x00\x00\x00
```

Source: Shodan screenshots courtesy of Phobos Group, February 2021 (composite image)

Figure 2. **Tens of thousands of remote desktops exposed in Ashburn, Virginia (US)**

Password reuse is one of the biggest security vulnerabilities humans introduce, but it can be minimized with training and tools, and by multifactor authentication where practicable. Installers and users should also be prevented from removing password protection controls (another vulnerability that Phobos demonstrated).

Surveillance cameras are also exposed; a casual search revealed open live footage of industrial settings (see Figure 3).



city:"san diego" -desktop

Source: Shodan screenshots courtesy of Phobos Group, February 2021 (composite image)

Figure 3. **Exposed live feeds of industrial surveillance cameras**

Password management systems can help maintain strong passwords and require frequent password updates. There are also cybersecurity tools to continuously scan for assets exposed online and to provide attack simulations. Services used at some facilities include threat intelligence and penetration tests on IP addresses and infrastructure. Low-tech approaches such as locked workstations and clean-desk policies also help protect sensitive information.

System-of-systems security issues

Industrial control systems, or ICSs, are often vulnerable because of a system-of-systems security issue. In other words, systems related to the way a control system is accessed (by authorized users) are exploited. In the February 2021 Oldsmar, Florida (US) city water plant attack, hackers took advantage of TeamViewer, a software tool for remote control, desktop and file sharing, and online meetings. The way TeamViewer is managed (its human-machine interface capabilities) provided a way into the control system, thwarting firewalls and virtual private networks. This is why security professionals recommend a zero-trust approach based on identity, rather than just a network-based perimeter security focus.

Securing DCIM software and services

Use of data center infrastructure management, or DCIM, systems is now mainstream, with more than half of respondents in a recent Uptime Institute survey having implemented some component, either commercial or homegrown. DCIM systems may include functions for security and compliance.

Ongoing and on-demand asset auditing features, for example, mean that requirements for security and regulatory standards can be monitored. And DCIM asset integrity monitoring helps prevent unauthorized physical devices being added to restricted areas.

DCIM software usually resides behind firewalls on dedicated (private) networks. However, even network segmentation or air gapping of software control systems is not foolproof; information is commonly exchanged, even if only periodically, including with systems that are connected to other networks. If a corporate IT network is compromised, DCIM systems can be vulnerable. If a DCIM system is penetrated, data values and alarm systems (e.g., threshold or shutdown settings) may be accessible. DCIM features that enable automation and control, such as for dynamic cooling optimization and IT power management, require a strong security approach. Vulnerability assessments of all internal and external (wide-area) network systems should be ongoing.

Modern DCIM products make use of underlying IT security systems. Virtualization management software is used to secure virtualized IT environments, and network management software secures the networking environment. However, these are not always adequate for critical systems, so a detailed security review should be considered.

Integration between applications, most likely using application programmable interfaces (APIs), may increase the security risk. The value of a DCIM investment multiplies when DCIM data is integrated and analyzed with other data. Bi-directional integrations of DCIM and third-party software are becoming common, usually with IT management systems. DCIM APIs are now standard in major products, including for the two foundational components of DCIM software:

- Asset management, which often connects to IT change management, such as work-order systems.
- Power and environmental monitoring, which is typically linked to IT service management (ITSM) and virtual machine management software.

APIs are used commonly across IT, including for cloud computing, and often are managed by automated tools. Yet all APIs broaden the attack surface by adding more avenues for intruders to exploit. Even APIs of noncritical systems can be a weakness; a compromised API can expose information about how an application is implemented and provide clues to underlying code architectures. Ongoing API security hygiene, including the patching of vulnerabilities, is critical. Many organizations are also using automated software to help ensure API security.

Security assessments are also needed when web access is made available to DCIM or other data center management systems for the first time or if the systems are linked to the internet for data transport — such as to consolidate DCIM data from different facilities or when using cloud-based analysis services, such as DMaaS (data center management as a service).

DMaaS: Cloudy data center management

DMaaS (data center management as a service) is a broad category of big data-driven cloud services that deliver customized analysis via a wide area network (paid for on a recurring, as-you-go basis). DMaaS aggregates and analyzes large sets of anonymized monitored data about equipment and operational environments from different facilities (customers). The data is analyzed using machine learning and other artificial intelligence and big data techniques. Results for individual customers may be tailored to their specific data center and delivered via dashboard (available online and/or in a mobile application), as well as in email, text and phone notifications.

Typically, each data center has gateway software (and sometimes also a physical gateway device) that gathers and sends data from monitored devices to the DMaaS supplier's cloud. Web-based mobile and/or desktop software acts as a personalized dashboard for managers to consume analysis, including for alarm notifications, for an overview of all sites under management, and for recommendations, assessments and reports.

Similar to all public cloud-delivered systems, the use of DMaaS requires security oversight, including adequate encryption of data, access controls to the system's interface and data, and API (application programmable interface) security hygiene.

Since the first service launched in late 2016, DMaaS adoption has grown steadily. However, some operators resist sending data (even encrypted) about their critical infrastructure outside their private firewall. Regulations may also be a barrier to DMaaS adoption for some.

Whether lack of certainty or clarity over data ownership and locality with DMaaS is a risk to data centers is vigorously debated. Some say that if hackers accessed the data, it would be of little use as the data is anonymized and, for example, does not include specific location details. Others say hackers could apply techniques, including AI, to piece together sensitive information and "de-anonymize" the data.

As with most areas of physical data center security, humans are a risk factor. For example, DCIM asset change and configuration software

Basic DCIM security

Questions to ask when deploying DCIM (data center infrastructure management) software include:

- How is user access managed, controlled, maintained, tracked, logged and reconciled?
- What is the process to remove individual user access from the DCIM system?
- Can user inputs into the DCIM system be audited?
- How is the system designed to prevent access from malevolent outsiders?
- What level of white box (source code analysis) and black box (website or binary code analysis) testing do you provide to ensure the integrity of the DCIM software?
- Does the system support any specific security standards? Please provide details.

typically includes an audit or trace function for saved changes. The people most frequently making changes are on-site operations and maintenance staff, which for many operators is likely to include at least some contractors.

DCIM and DMaaS security is extremely important. Documentation should outline security measures and practices, and systems should enable separation of user roles. Access history should also be logged, especially if customers are involved. In some environments, the software and services should be capable of meeting national or industry-specific security standards and should support practices such as encryption and multifactor authentication.

Conclusions

Securing the physical data center is an ongoing and multifaceted process. While data center risk profiles and strategies to counter threats can be similar across similar types of data centers, there is no singular approach. There is also no single process, product or service that protects against all the physical, human or digital threats.

Business risk and business requirements typically determine the level of physical data center security needed, and documented policies are important for all facilities. Approaches should be customized for individual sites and should be regularly updated and shared with stakeholders.

There are numerous cybersecurity tools to minimize the scope for social engineering and unauthorized access. Ongoing awareness training and reviews of policies and procedures, including separation of duties, are recommended.

Data security and cybersecurity will not reduce the need for physical security. Without good physical security, the entire virtual edifice is vulnerable.

Given the potential impact of a serious physical incursion, and in light of recent current events, additional vigilance and expense is likely to be justified for most management teams.

Uptime Institute® | INTELLIGENCE

ABOUT THE AUTHORS

Rhonda Ascierto is Uptime Institute's Vice President of Research. She has spent two decades at the crossroads of IT and business as an analyst, speaker, adviser, and editor covering the technology and competitive forces that shape the global IT industry. Contact: rascierto@uptimeinstitute.com

Todd Traver is Uptime Institute's Vice President for IT Optimization and Strategy and has over 30 years' experience in all aspects of data center and IT planning, reliability design and operational efficiency. Todd was IBM Global Technology Services' Chief Engineer and Senior Technical Staff Member responsible for developing global data center portfolio technology strategy, tools, and operational and process optimization plans. Todd's passion is to expand Uptime Institute's offerings based on these experiences, providing a pathway for others to innovate new approaches to integrate facilities and IT departments, which reduces total cost of operation and enhances end-user application and data reliability and resiliency experience. Contact: ttraver@uptimeinstitute.com

ABOUT UPTIME INSTITUTE

Uptime Institute is an advisory organization focused on improving the performance, efficiency and reliability of business critical infrastructure through innovation, collaboration and independent certifications. Uptime Institute serves all stakeholders responsible for IT service availability through industry leading standards, education, peer-to-peer networking, consulting and award programs delivered to enterprise organizations and third-party operators, manufacturers and providers. Uptime Institute is recognized globally for the creation and administration of the Tier Standards and Certifications for Data Center Design, Construction and Operations, along with its Management & Operations (M&O) Stamp of Approval, FORCSS® methodology and Efficient IT Stamp of Approval.

Uptime Institute – The Global Data Center Authority®, a division of The 451 Group, has office locations in the US, Mexico, Costa Rica, Brazil, UK, Spain, UAE, Russia, Taiwan, Singapore and Malaysia. Visit uptimeinstitute.com for more information.

All general queries:
Uptime Institute
5470 Shilshole Avenue NW, Suite 500
Seattle, WA 98107 USA
+1 206 783 0510
info@uptimeinstitute.com