

## **Downtime em Data Centers no Core e no Edge: Uma Pesquisa sobre Frequência, Duração e Ações**

---

**Patrocinada pela Vertiv**

Conduzida de forma independente pelo Ponemon Institute LLC

Data de Publicação: Janeiro 2021

# Downtime em Data Centers no Core e no Edge: Uma Pesquisa sobre Frequência, Duração e Ações

Ponemon Institute, Janeiro 2021

## Índice

	Seção	Página
Parte 1	Introdução	3
Parte 2	Principais Achados	4
	Tamanho da Instalação Frequência do Downtime no Core e no Edge Duração do Downtime no Core e no Edge Comparando Atributos dos Data Centers de Core e de Edge Causas do Downtime no Core e no Edge Ações para Evitar Eventos Downtime	
Parte 3	Perfil dos Participantes	10
Parte 4	Ressalvas	13

# **Downtime em Data Centers no Core e no Edge: Uma Pesquisa sobre Frequência, Duração e Ações**

Ponemon Institute, Janeiro 2021

## **Introdução**

O edge computing está se expandindo rapidamente e remodelando o ecossistema dos data centers na medida em que organizações de todos os setores levam a computação e o armazenamento para mais perto dos usuários para melhorar os tempos de resposta e reduzir as necessidades de largura de banda.

Embora algumas formas de computação distribuída tenham sido comuns em alguns setores por diversos anos, essa atual evolução é diferente no sentido em que ela está possibilitando uma ampla variedade de aplicações novas e emergentes e tem uma necessidade de criticidade mais alta do que sites de computação distribuída tradicionais.

Ao mesmo tempo, os gestores de data centers no core (centrais) estão lidando com uma maior complexidade e tendo que equilibrar diversas – e algumas vezes conflitantes – prioridades que podem comprometer a disponibilidade.

Como resultado, as redes dos data centers atuais estão mais vulneráveis ao downtime do que nunca. Em um esforço para quantificar esta vulnerabilidade, o Instituto Ponemon conduziu um estudo sobre a frequência, duração e ações relativas ao downtime no core e no edge, patrocinado pela Vertiv.

O estudo é baseado nas respostas de 425 participantes, representando 132 data centers e 1.667 locais de edge. Todos os data centers no core e no edge incluídos no estudo estão localizados nos Estados Unidos/Canadá e América Latina (LATAM).

O estudo encontrou redes de data centers vulneráveis a eventos de downtime por toda a rede. Data centers centrais passaram por uma média de 2,4 paradas totais das instalações por ano, com uma duração média de mais de duas horas (138 minutos). Isso é além dos quase 10 eventos de downtime anuais isolados em determinados racks ou servidores. No edge, a frequência das paradas completas da instalação foi ainda maior, embora a duração dessas indisponibilidades tenha sido menos da metade daquelas dos data centers centrais.

O estudo examinou também as ações que definem as decisões relativas aos data centers de core e de edge para ajudar a identificar fatores que poderiam contribuir para eventos de downtime. Mais da metade (54%) de todos os data centers de core não estão usando as melhores práticas em design do sistema nem redundância e 69% dizem que seu risco de uma indisponibilidade não planejada aumentou como decorrência da contenção de custos.

As principais causas de downtime não planejado no core e no edge incluem ciberataques, falhas em equipamentos de TI, erro humano, falha na bateria do UPS e falha do equipamento UPS.

Finalmente, a pesquisa pediu aos participantes que identificassem as medidas que suas organizações poderiam tomar para prevenir eventos futuros de downtime. Eles identificaram atividades que variavam desde investimentos em novos equipamentos e redundância até melhor treinamento e documentação.

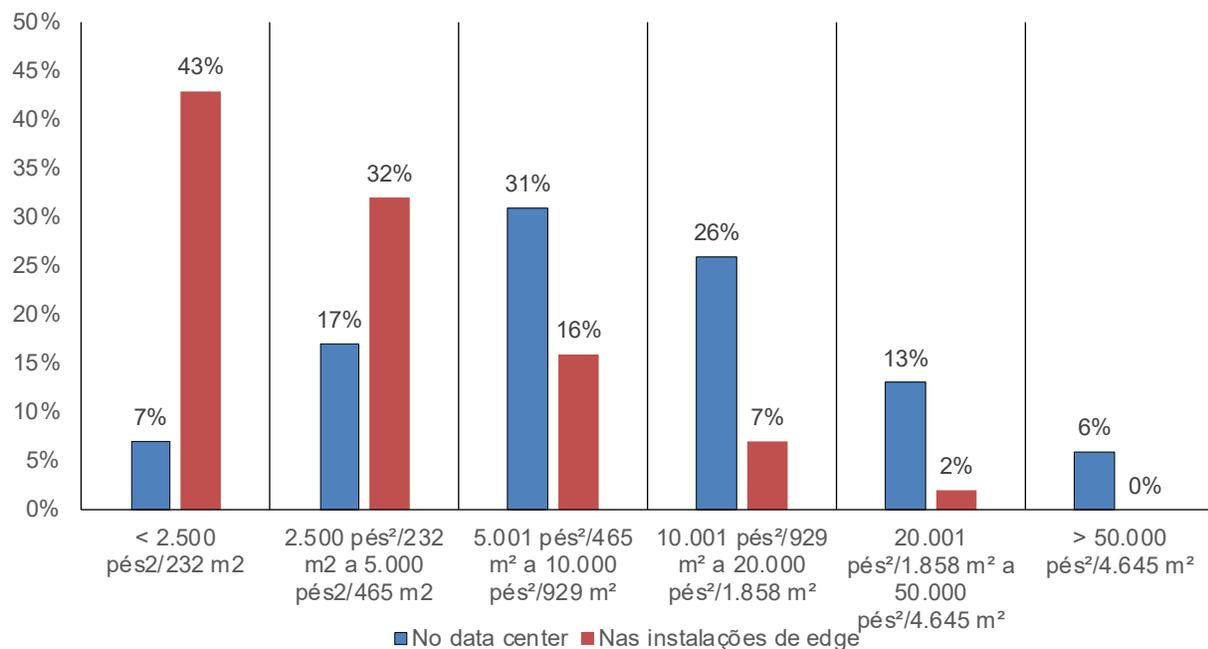
## Principais Achados

### Tamanho da Instalação

Data centers de edge não são necessariamente definidos pelo tamanho, mas sim por sua função. Para efeitos dessa pesquisa, data centers de edge são definidos como sendo instalações que trazem a computação e o armazenamento de dados para mais perto do local onde eles são necessários, para melhorar os tempos de resposta e economizar largura de banda. Entretanto, conforme visto na Figura 1, os data centers de edge tinham em média aproximadamente um terço do tamanho dos data centers centrais (core).

O tamanho extrapolado para data centers de core que participaram nesse estudo é de 1.408 metros quadrados/15.153 pés quadrados. Para as instalações no edge, o tamanho médio é de 465 metros quadrados/5.010 pés quadrados.

**Figura 1: Tamanho médio das instalações de core e edge computing**



A Figura 2 mostra o tamanho médio do data center, por quartil, nos EUA/Canadá e LATAM. Em todos os quartis, os data centers estudados nos EUA e no Canadá eram significativamente maiores do que os na América Latina.

**Figura 2: Tamanho do data center por quartil**

Quartil	Tamanho em geral	Tamanho nos EUA/Canadá	Tamanho na LATAM
Qtrl 1	4.001 pés²/372 m²	5.040 pés²/468 m²	2.962 pés²/275 m²
Qtrl 2	8.452 pés²/785 m²	10.732 pés²/997 m²	6.172 pés²/573 m²
Qtrl 3	15.898 pés²/1.477 m²	19.005 pés²/1.766 m²	12.791 pés²/1.188 m²
Qtrl 4	32.400 pés²/3.010 m²	40.500 pés²/3.763 m²	24.300 pés²/2.258 m²

## Frequência de Downtimes no Core e no Edge

A Figura 3 mostra a experiência de paradas nos data centers participantes, nos últimos 24 meses. Como pode ser visto, a parada total do data center teve a menor frequência (4,81). Entretanto, esses eventos também são os mais disruptivos, e as 4,81 paradas não planejadas das instalações em um período de 24 meses seria considerado inaceitável por muitas organizações.

Indisponibilidades parciais de determinados racks no data center têm a maior frequência, com 9,3; seguidas de indisponibilidades de servidores individuais, com 9,43.

Pode ser difícil comparar diretamente a quantidade total de eventos de downtime em instalações no core e no edge devido à maior complexidade geralmente encontrada nos data centers de core e a maior presença de pessoal nessas instalações. Entretanto, é possível comparar as paradas completas da instalação para data centers no core e no edge. Data centers no edge tiveram por uma frequência um pouco maior de paradas completas da instalação, com uma média de 5,39 em 24 meses. Conforme os sites no edge continuam a proliferar, reduzir a frequência das indisponibilidades no edge se tornará uma alta prioridade para muitas organizações.

**Figura 3: Frequência de downtimes em 24 meses**

Tipo de Evento	Frequência
<b>Data Center de Core</b>	
Falta de energia da rede elétrica primária	7,19
Parada completa da instalação	4,81
Indisponibilidades localizadas em determinados racks	9,93
Indisponibilidade limitada a servidores individuais	9,43
<b>Data Centers de Edge</b>	
Parada completa da instalação	5,39

A Figura 4 mostra que os data centers na América Latina tinham maior probabilidade de passar por todos os tipos de indisponibilidades, tanto no core quanto no edge, do que os data centers localizados nos Estados Unidos e Canadá.

**Figura 4: Frequência de indisponibilidades nos últimos 24 meses: EUA/Canadá versus LATAM**

Tipo de Indisponibilidade	Américas	EUA/Canadá	LATAM
<b>Data Center de Core</b>			
Falta de energia da rede elétrica primária	7,19	5,60	9,22
Parada completa do data center	4,81	4,00	5,84
Indisponibilidades localizadas em determinados racks	9,93	8,26	12,07
Indisponibilidade limitada a servidores individuais	9,43	7,68	11,66
<b>Data Center de Edge</b>			
Parada completa da Instalação de edge computing	5,39	4,98	5,92

## Duração dos Eventos de Downtime no Core e no Edge

A Figura 5 mostra a duração média de diversos tipos de eventos de indisponibilidade durante o período estudado. A parada completa de data centers no core teve a maior duração em 138 minutos. Por outro lado, indisponibilidades parciais de determinados racks no data center tiveram a maior frequência com 9,93, mas a menor duração, com um pouco menos de uma hora.

Os downtimes de instalações de edge tem uma duração consideravelmente menor do que eventos similares em data centers de core, apesar dos limitados recursos técnicos que, em geral, estão disponíveis nesses sites. Isso provavelmente se deve à funcionalidade mais focada desses sites, a qual limita complexidade.

**Figura 5. Duração do downtime por tipo de evento**

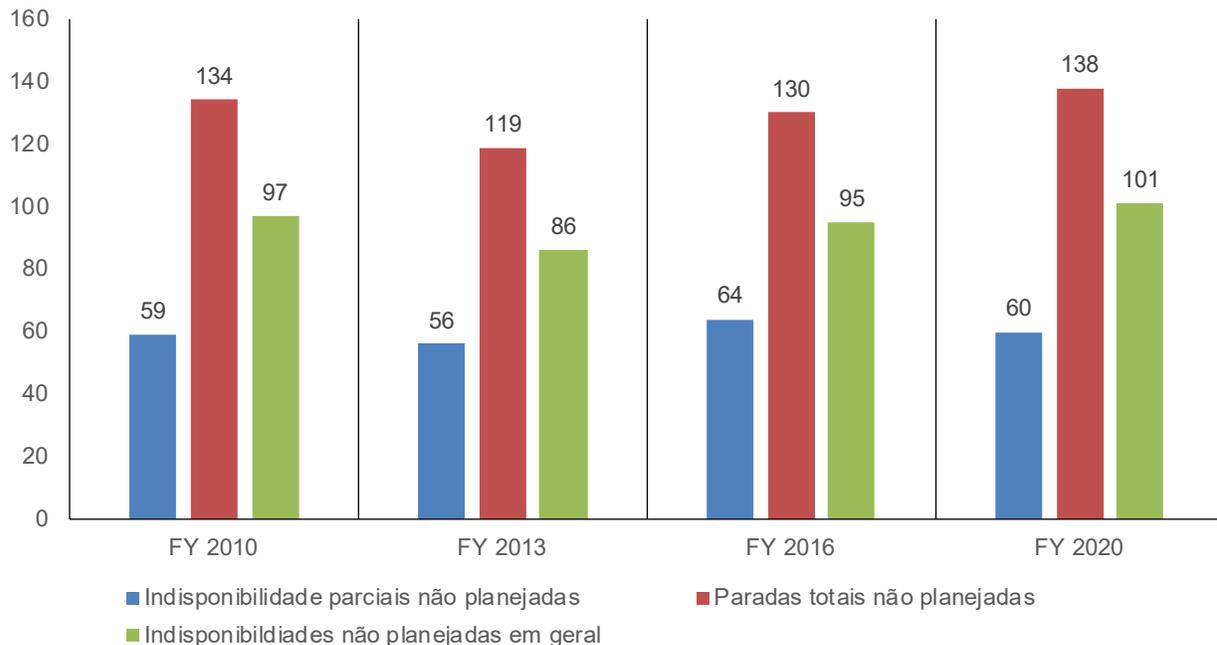
Tipo de Evento	Duração (min.)
<b>Data Center de Core</b>	
Falta de energia da rede elétrica primária	99,80
Parada completa do data center	137,87
Indisponibilidades localizadas em determinados racks	59,98
Indisponibilidade limitada a servidores individuais	60,05
<b>Data Center de Edge</b>	
Parada completa da instalação	45,40

A Figura 6 coloca os dados da duração dos downtimes atuais sob uma perspectiva histórica, comparando a média da duração das indisponibilidades parciais e totais para data centers de core com estudos anteriores que tenham usado a mesma metodologia. (A frequência e a duração para data centers de edge não foram incluídas em estudos anteriores.)

Todos os estudos mostram que, nos últimos dez anos, a duração das indisponibilidades completas não planejadas foi mais do que o dobro do tempo das indisponibilidades parciais. Por exemplo, em 2020 o total de paradas não planejadas foi de 138 minutos e o total das indisponibilidades parciais não planejadas foi de 60 minutos.

Os dados mostram também que a duração das indisponibilidades não planejadas aumentou regularmente nos três últimos estudos. A duração média de todas as indisponibilidades não planejadas aumentou de 86 minutos em 2013 para 101 minutos em 2020 e das paradas completas de instalações aumentou de 119 minutos em 2013 para 138 minutos em 2020.

**Figura 6: Tendência na duração de paradas parciais e totais (medidas em minutos)**  
Comparação entre resultados de 2010, 2013, 2016 e 2020



### Atributos dos Data Centers de Core e de Edge

Quando solicitados a comparar os atributos entre as instalações de core e de edge, os participantes da pesquisa mostraram diferenças relativamente pequenas nos riscos associados a cada uma delas (Figura 7). Curiosamente, a disponibilidade foi vista como uma prioridade mais alta para o edge do que para o data center de core. Sessenta e dois por cento dos participantes consideram a disponibilidade como a prioridade máxima, incluindo a minimização dos custos, comparado com 55% para os data centers de core. Isso pode ser devido aos limitados recursos técnicos disponíveis nesses sites para lidar com eventos de downtime e o número de sites de edge que as organizações acreditam terão nos próximos cinco anos.

Melhores práticas de design e redundância também foram citadas mais consistentemente para o edge. Entretanto, mesmo para o edge, apenas um pouco mais da metade dos participantes da pesquisa (54%) disse estar empregando melhores práticas. Isso indica que os participantes sabem das melhores práticas, mas as aplicam de forma limitada devido a contenções de custos, falha em priorizar a disponibilidade ou uma combinação dos dois motivos.

Uma maioria considerável de participantes citam as contenções de custo como um fator que aumenta o risco de indisponibilidades não planejadas tanto nos data centers de edge quanto de core, indicando que mesmo as instalações que priorizam a disponibilidade sobre a minimização de custos não estão fazendo os investimentos necessários para reduzir os riscos de downtime. Essa percepção é reforçada pela constatação de que apenas metade dos participantes diz que sua alta direção apoia plenamente os esforços para evitar indisponibilidades não planejadas.

A eficiência energética, que é motivada nos atuais data centers por preocupações relativas tanto aos custos quanto ambientais, foi citada entre as prioridades mais altas tanto para as instalações de core quanto de edge por aproximadamente a metade dos participantes - 51% para data centers de core e 49% para os de edge),

indicando que as organizações estão tendo uma abordagem similar em relação à eficiência energética no core e no edge.

Data centers de core têm uma tendência maior de relatar que os negócios para os quais dão suporte são dependentes dos sistemas para gerar receitas e conduzir o comércio eletrônico (67% dos data centers, contra 56% dos locais de edge).

**Figura 8: Atributos para data centers de core e de edge**



Cada percentual mostra a combinação das respostas Concordo e Concordo Firmemente (usando-se uma escala de concordância de cinco pontos)

### Causas do Downtime em Data Centers de Core e de Edge

Ciberataques, falhas em equipamentos de TI, falhas devidas a erros humanos, falhas nas baterias do UPS e falhas no UPS foram as principais causas de indisponibilidades citadas pelos participantes do estudo. Instalações no edge foram um pouco mais vulneráveis a ciberataques, falhas em equipamentos de TI e falhas devidas a erros humanos do que data centers de core.

Outras causas de indisponibilidades citadas incluíram falha em chave de transferência automática (ATS), falha no gerador, capacidade do UPS excedida, falhas relacionadas às condições climáticas, falhas relacionadas ao calor e à entrada de água.

## Ações para Evitar Eventos de Downtime

De acordo com a Pesquisa Global Data Center Survey de 2020, realizada pelo Uptime Institute, três entre cada quatro participantes disseram que seu evento de downtime mais recente era evitável. Embora não seja possível transferir aqueles resultados diretamente para esse estudo, ele efetivamente sugere que há uma oportunidade para reduzir a frequência dos eventos de downtime em muitos data centers.

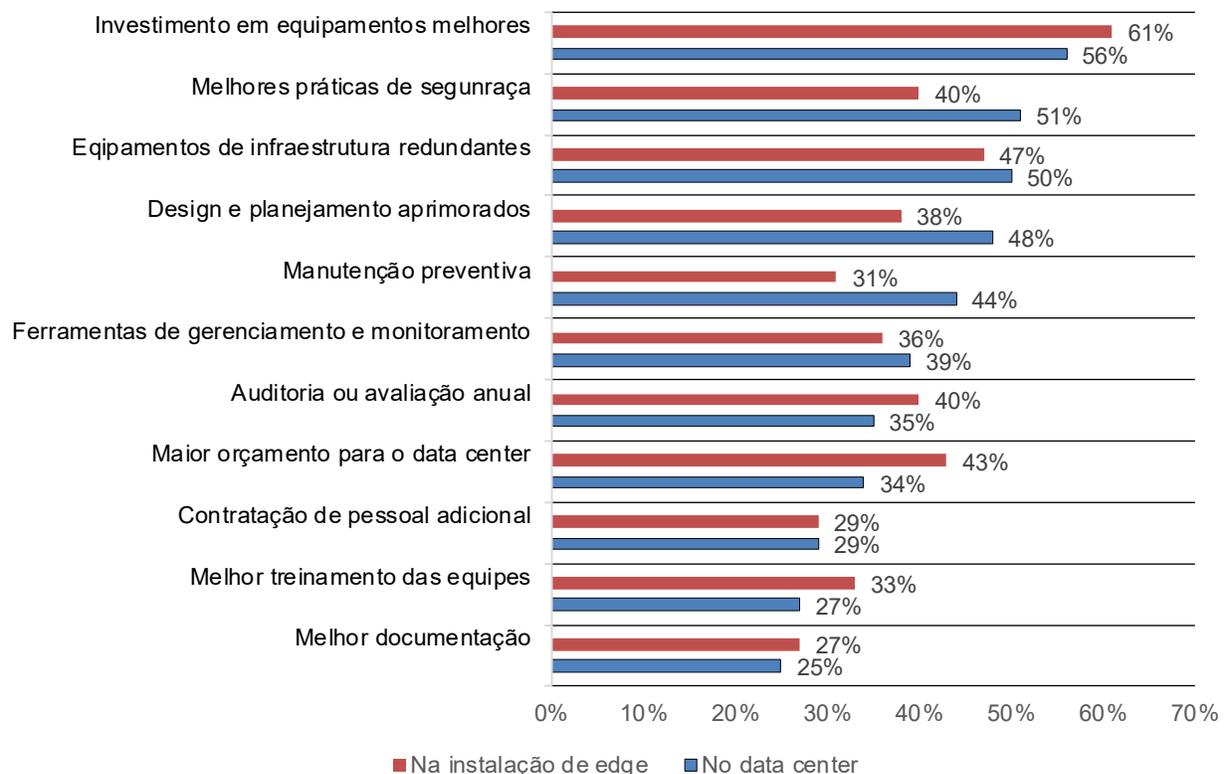
A Figura 9 mostra as ações que os participantes disseram poder ser tomadas para evitar indisponibilidades não planejadas no futuro. A ação mais citada, tanto para o core quanto para o edge, foi o investimento em equipamentos de TI. Isso provavelmente se deve à alta frequência de eventos de downtime relacionados a servidores individuais.

Da mesma forma, 51% dos data centers de core e 40% dos locais de edge selecionaram melhorar as práticas de segurança como uma das principais medidas para evitar indisponibilidades não planejadas, refletindo a preocupação constante com o impacto dos ataques de negação de serviço (DoS) sobre a disponibilidade do data center.

Outras medidas listadas pelos participantes poderiam ter um impacto na redução da longa duração de paradas totais das instalações, incluindo equipamentos de infraestrutura redundantes, design e planejamento aprimorados, manutenção preventiva, ferramentas de gerenciamento e monitoramento, auditorias ou avaliações, maior orçamento, contratação de mais colaboradores e melhoria no treinamento das equipes.

### Figura 8: O que pode ser feito para evitar as indisponibilidades não planejadas no futuro?

Comparação de data centers e locais de edge



## Perfil dos Participantes

O estudo é baseado nas respostas de 425 participantes, representando 132 data centers e 1.667 locais de edge. Todos os data centers de core e de edge incluídos no estudo estão localizados nos Estados Unidos e Canadá e América Latina incluindo o México (LATAM).

A tabela a seguir sumariza a distribuição das empresas e separa os data centers participando no estudo. Um total de 12 setores são representados na amostragem final. Nossa amostragem final inclui um total de 108 diferentes organizações, representando 132 data centers e 1.667 instalações de edge.

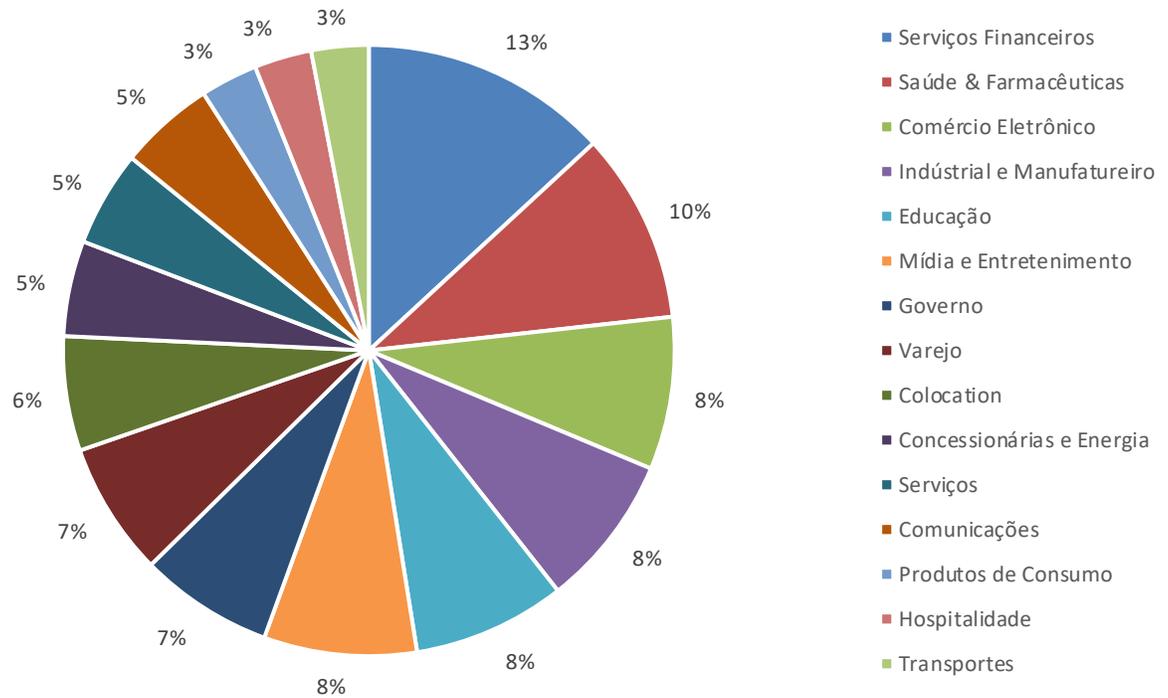
**Figura 9: Distribuição da amostra dos data centers localizados nos Estados Unidos, Canadá e LATAM**

Setor	Empresas	Data centers	EUA./Canadá	LATAM	Locais de Edge EUA/Canadá	Locais de Edge LATAM
Serviços financeiros	14	14	8	6	125	87
Saúde e Farmacêutico	11	10	7	3	78	31
Comércio Eletrônico	9	11	8	3	234	112
Industrial & manufatureiro	9	12	7	5	67	32
Educação	9	9	4	5	33	20
Mídia & entretenimento	9	9	3	6	27	13
Governo	8	13	5	8	89	56
Varejo	8	7	6	1	138	61
Colocation	7	9	6	3	24	5
Concessionárias & Energia	5	8	4	4	52	35
Serviços	5	9	5	4	53	42
Comunicações	5	7	3	4	63	52
Produtos de consumo	3	4	2	2	36	20
Hospitalidade	3	5	3	2	20	17
Transportes	3	5	3	2	24	21
<b>Total</b>	<b>108</b>	<b>132</b>	<b>74</b>	<b>58</b>	<b>1.063</b>	<b>604</b>

A Figura 10 resume a amostra dos data centers das empresas participantes, de acordo com 15 classificações de segmentos econômicos. Os serviços financeiros e o setor de saúde são os dois maiores segmentos, representando 13% e 10% da amostra respectivamente. Empresas de serviços financeiros incluem bancos de varejo, processadoras de pagamentos, seguradoras, corretoras e gestoras de investimentos.

**Figura 10: Distribuição das organizações participantes por segmento econômico**

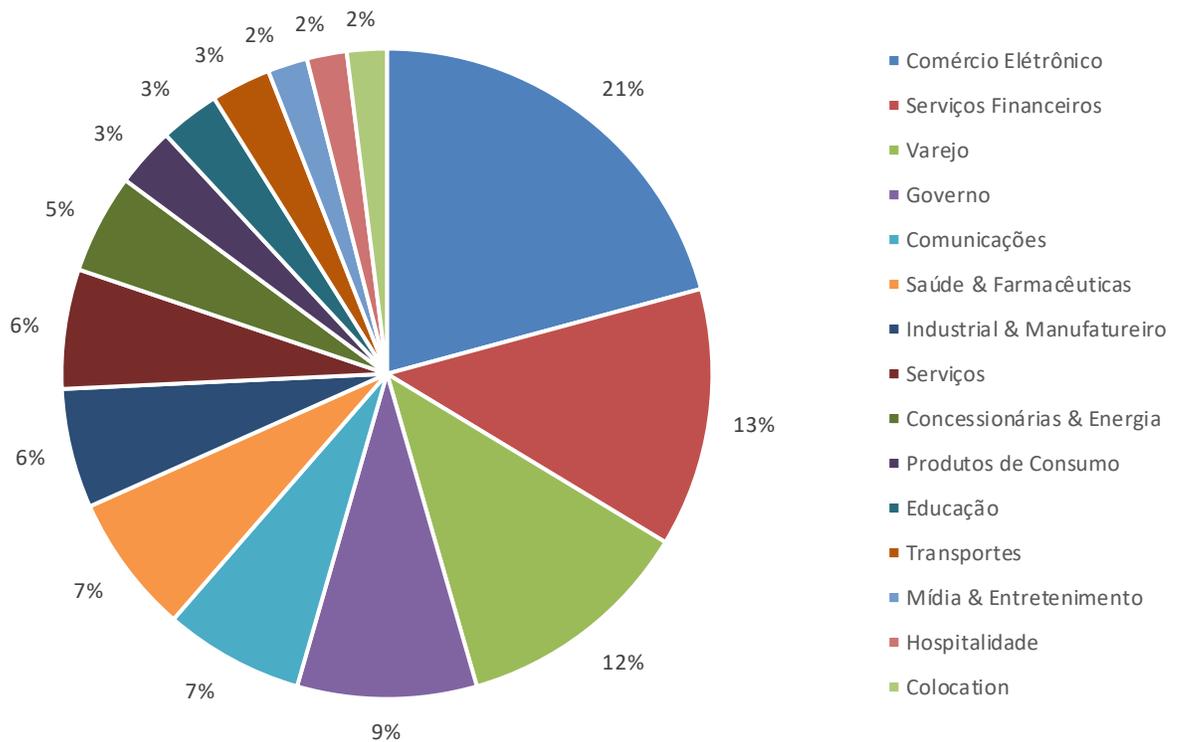
Computado a partir de 132 data centers avaliados



A Figura 11 mostra a frequência percentual de 1.667 locais de edge de acordo com o segmento econômico. Com 21%, o comércio eletrônico é o maior segmento, seguido pelos serviços financeiros com 13%.

**Figura 11: Distribuição das instalações de edge computing por segmento econômico**

Computado a partir de 1.667 locais de edge avaliados



Abaixo encontram-se os líderes funcionais dentro de cada organização que participou do processo de avaliação:

- Gerente de instalação
- Diretor-Executivo de TI (CIO)
- Gestor de data center
- Diretor-Executivo de segurança da informação
- Gerente de operações de TI
- Auditoria e compliance de TI
- Operações e engenharia
- Administrador de nuvem

## Ressalvas

Esse estudo usa um método de avaliação confidencial e proprietário, implementado com sucesso em pesquisa anterior do Ponemon Institute. Entretanto, há limitações inerentes a pesquisas de avaliação que precisam ser considerados com cuidado antes de estabelecer-se as conclusões a partir dos achados.

- **Resultados não estatísticos:** O objetivo desse estudo é descritivo ao invés de uma inferência normativa. O estudo atual é feito a partir de uma amostra representativa, não estatística, de data centers, todos tendo passado por pelo menos uma indisponibilidade não planejada nos últimos 12 meses. Inferências estatísticas, margens de erro e intervalos de confiança não podem ser aplicados a estes dados dada a natureza do nosso plano de amostragem.
- **Sem resposta:** Os achados atuais estão baseados em uma amostra representativa pequena de estudos de caso encerrados. Uma pesquisa de avaliação inicial foi enviada para um grupo de avaliação com mais de 600 organizações, todas as quais acreditava-se terem passado por uma ou mais indisponibilidade nos últimos 12 meses. Cento e trinta e dois data centers forneceram pesquisas de avaliação que puderam ser usadas. A pré-disposição para não responder não foi testada, então, é sempre possível que as empresas que não participaram sejam substancialmente diferentes em termos dos métodos usados para gerenciar os processos de detecção, contenção e recuperação.
- **Pré-disposição da base de amostragem:** Como nossa base de amostra é subjetiva, a qualidade dos resultados é influenciada pelo grau de representatividade do quadro da população de empresas e de data centers sendo estudados. Acreditamos que o atual quadro de amostras tem uma pré-disposição em relação a empresas com operações mais maduras de seus data centers.
- **Informações específicas da empresa:** As informações avaliadas são sensíveis e confidenciais. Portanto, o presente instrumento não captura informações que possam identificar as empresas. Ele também permite que indivíduos usem as variáveis de resposta por categoria para revelar informações demográficas sobre a empresa e segmento econômico. A classificação por segmento econômico baseia-se em informações fornecidas pelos participantes.
- **Fatores não medidos:** Para manter a pesquisa concisa e focada, decidimos omitir outras variáveis importantes de nossas análises, tais como principais tendências e características organizacionais. Não é possível estimar nesse momento até que ponto as variáveis omitidas poderiam explicar os resultados da avaliação.

Se tiver dúvidas comentários sobre esse relatório, por favor, entre em contato conosco por carta, telefone ou e-mail:

Ponemon Institute LLC  
Attn: Research Department  
2308 US 31 North  
Traverse City, Michigan 49686 EUA  
1.800.877.3118  
[research@ponemon.org](mailto:research@ponemon.org)

## **Ponemon Institute**

### ***Promoção da Gestão Responsável da Informação***

O Ponemon Institute dedica-se à pesquisa independente e educação que promovem práticas responsáveis de gestão da informação e privacidade junto a empresas e governo. Nossa missão é conduzir estudos empírico de alta qualidade sobre questões fundamentais que afetam a gestão e a segurança de informações sensíveis de pessoas e organizações.

Mantemos estrita confidencialidade de dados, privacidade e padrões éticos de pesquisa. Não coletamos quaisquer informações que possam identificar pessoalmente indivíduos (ou que identifiquem empresas) em nossas pesquisas. Além disso, temos padrões de qualidade rigorosos para garantir que os participantes não sejam submetidos a perguntas alheias ao assunto em questão, irrelevantes ou inapropriadas.