Enterprise Buyer's Guide for Data Protection

Pathfinder Report

July 2022

Commissioned by

Veeam

451 Research

S&P Global
Market Intelligence

About this paper

A Pathfinder paper navigates decision-makers through the issues surrounding a specific technology or business case, explores the business value of adoption, and recommends the range of considerations and concrete next steps in the decision-making process.

About the Author

Henry Baltazar

Research Director, Storage

Henry Baltazar is a Research Director for the storage practice at 451 Research, a part of S&P Global Market Intelligence. Henry returned to 451 Research after spending nearly three years at Forrester Research as a senior analyst serving Infrastructure & Operations Professionals and advising Forrester clients on datacenter infrastructure technologies. Henry has evaluated and tested storage hardware and software offerings for more than 15 years as an industry analyst and as a journalist.

Prior to 451 Research and Forrester, Henry spent nearly nine years working as a technical analyst for eWeek Labs, where he covered storage, server hardware and network operating systems. At eWeek Labs, he initiated the testing coverage of various technologies, including data replication, clustering, virtual tape libraries, storage virtualization, SAN management, NAS, iSCSI and email archiving. In addition, Henry was a member of eWeek's editorial board and provided content for the magazine's enterprise storage blog. Henry has been widely quoted in the press, including such media outlets as Silicon Valley Business Journal, Computerworld and SearchStorage.com.

Henry holds a BA in environmental sciences from the University of California, Berkeley.

Table of Contents

Executive Summary

Organizations have a broad range of choices when it comes to execution venues for workloads, from traditional on-premises infrastructures to hyperscaler public clouds and SaaS platforms. True hybrid IT, which can leverage on-premises and cloud resources interchangeably, is the desired state for organizations since it provides maximum flexibility. Adding to the burden is the emergence of newer technologies such as containers and cloud-native data services, as well as growing threats such as ransomware, which are forcing organizations to rethink their data protection and recovery strategies.

Meeting disaster recovery (DR) and business continuity requirements has long been a top pain point for organizations, and this goal has become more difficult to reach since customer expectations for uptime and data recovery include very little tolerance for lost time or lost data.

Modern data protection tools must evolve in multiple dimensions to keep pace:

- Provide comprehensive data protection
- Facilitate reliable recovery from ransomware
- Deliver enterprise-class scalability, performance and ease of use
- Provide support for modern applications and operational practices
- Have the ability to integrate cloud into a data protection strategy

Comprehensive Data Protection

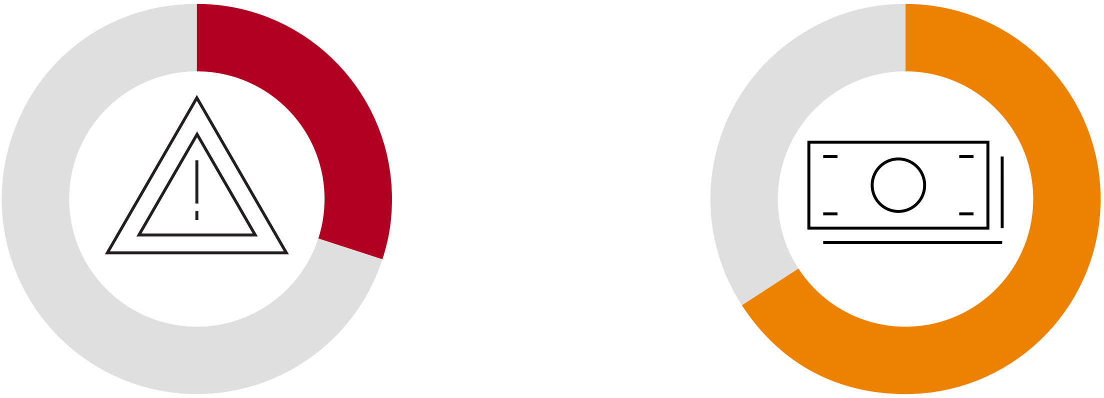
On the source side of the data protection equation, modern tools must not only be able to handle legacy applications running on physical servers and contemporary applications running on VMs, but they must also be able to protect assets such as network-attached storage and file servers, where a large portion of unstructured data resides.

Plug-ins and API support for applications are also important factors since individual applications have different recommended processes for quiescing a workload to capture data for a snapshot without impacting a production workload or accidentally corrupting data. Intelligent application integration can also facilitate key capabilities such as self-service restore for database or application administrators.

Another dramatic change that legacy backup tools were not designed for is the rise of public cloud, Kubernetes and SaaS workloads, where the data and workloads reside outside of a customer's on-premises estate. Recently, more organizations have begun using third-party backup providers, such as Salesforce and M365, to protect key SaaS platforms, with 34% of respondents now using them compared with just 15% in the 2021 study. Even though customers do not have to worry about the physical hardware and software running workloads in cloud and SaaS environments, data protection for workloads is still necessary since it can prevent data loss and data corruption due to intentional or unintentional file deletions, or from misbehaving applications. We note that public cloud and SaaS providers explicitly state that backup and recovery operations are part of a customer's ongoing responsibilities. Rising security threats such as ransomware are also highlighting the need for enhanced data protection.

Data Highlight

Figure 1: Outages (and Their Costs) Are Increasing



30%

of respondents experienced significant outages over the past two years

66%

of incidents cost the organization over $100,000

Q. When was the last time your organization experienced an outage that resulted in lost data or affected worker productivity?

Base: All respondents (n=372)

Q. Please estimate the total cost to your organization of its most recent cloud outage or downtime (from outage to full recovery, including direct costs, opportunity costs, etc.)

Base: Organizations with recent service outages/incidents and have estimated costs (n=192)

Source: 451 Research's Voice of the Enterprise: Storage, Data Management and Disaster Recovery 2022

Technology Considerations

Does your current backup platform protect physical servers, VMs and cloud services (SaaS & IaaS)?

Is your organization's data still accessible in the event of a cloud or SaaS provider outage?

How many multiple-backup platforms are you using to protect all of your workloads and data?

Does the organization prefer the convenience of integrated backup appliances or the flexibility of modular systems?

Do cloud or hardware offerings create a vendor lock-in situation?

Business Considerations

Where will data reside in your organization in the future, and is data protection available there?

Is software licensing complexity preventing your organization from transitioning data protection investments to new architectures such as containers?

Do business stakeholders and application owners (database admins) require self-service recovery?

Is the need for accelerated time to value increasing the usage of SaaS and public cloud?

Does your organization favor a perpetual license or a subscription-based model?

What to Look for

A comprehensive approach to data protection not only includes the workloads and data that are being protected, but it must also consider the application configuration, backup and storage resources that are responsible for running these processes and preserving data to facilitate recovery when needed. For some organizations, integrated appliances are a one-stop solution since they run data protection software while also handling storage capabilities. Modular systems are another option for customers that value flexibility. In either case, a modern data protection platform should be able to support either hardware format for on-premises deployments, not only to give customers freedom of deployment choice but also to reduce the threat of long-term vendor lock-in.

Flexible recovery options are also essential for a modern data protection platform to ensure a workload can failover or be rehydrated in a new execution venue if a large-scale disaster knocks out the primary site, or if the available resources are not sufficient to handle the production workload. Recovery flexibility also needs to provide options for customers to convert workloads to match resource availability, which is where conversion tools can transform workloads between physical, VM and container formats when necessary.

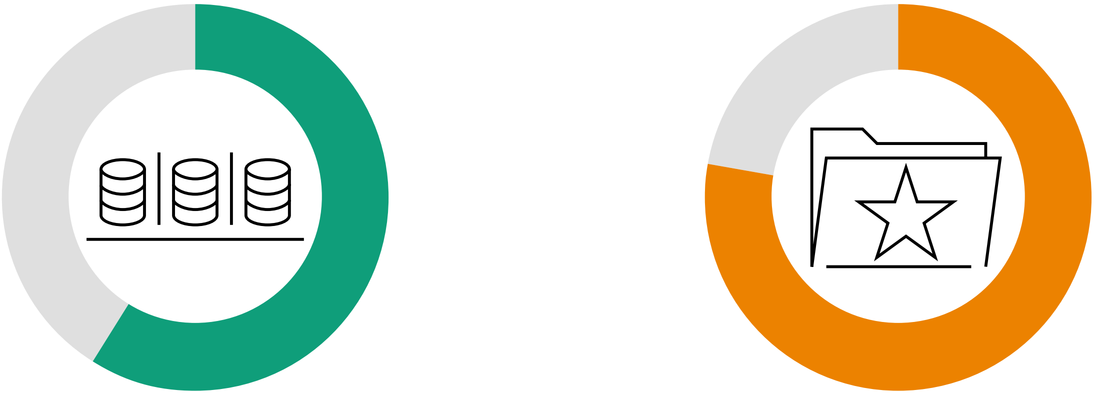
Ransomware Threat Drives the Need for Rapid Reliable Recovery

The "3-2-1 backup strategy," which calls for three backup copies, stored on at least two different media types, with one copy kept off-site, has been the recommended standard for many years. Ransomware had made it necessary to improve on that strategy since attackers are aggressively targeting the backup safety net to force organizations to cave in to their demands.

Ransomware deadlines require an accelerated response, and organizations need to have tools to facilitate rapid recovery and also to help create a secure testing environment in which to ensure that existing backups are free from ransomware before a recovery is attempted.

Data Highlight

Figure 2: The Importance of Being Prepared



59%

are increasing their spending on backup services and storage to ensure they can recover from a ransomware incident

78%

are maintaining a 'gold copy' of data for recovery after a ransomware incident

Q. Is your organization increasing spending for backup services and/or backup storage as a result of the potential threats of ransomware?
Base: All respondents (n=334)
Q. Does your organization maintain a 'gold copy' of data to recover after a ransomware incident?
Base: All respondents (n=315)
Source: 451 Research's Voice of the Enterprise: Storage, Data Management and Disaster Recovery 2022

Technology Considerations

Do you maintain an off-site and immutable data copy for recovery purposes?

Do you have a lab or secure recovery site to verify backups are free of ransomware before completing production recovery?

Do you have DR orchestration to provide recovery consistency and speed?

Business Considerations

Does your organization have comprehensive data protection to quickly recover from ransomware?

Do you have adequate reporting to prove compliance and data protection requirements are being met?

What to Look for

A secure backup infrastructure can provide additional protection even if cloud or on-premises accounts are compromised. This is where a "3-2-1-1-0 rule" comes into play and is considered best practice. The added "1" calls for at least one copy of data that is stored in an immutable format, and the "0" represents a backup with zero backup errors in order to ensure full recovery.

A few key capabilities should be included in a modern data protection solution, starting with immutable backup storage, which ensures that no delete operations or modifications are allowed within a set time frame to protect backup repositories. Automated backup validation ensures that backups are completed and that warnings are captured when backup jobs fail. Secure and validated restore is needed to ensure backups are ransomware-free and that a restore operation will run smoothly. A virtual labs capability can also be useful to provide security teams with a clean environment to ensure all traces of ransomware are eliminated and that no day zero vulnerabilities exist when workloads return to a production state. Disaster recovery orchestration is necessary to automate and accelerate recovery operations to minimize the length of an outage and the damage caused by the ransomware incident.

Scale, Speed and Ease of Use

The data protection and data management challenge will only get more difficult with time due to several factors. First and foremost is the rapid growth of data, which is the norm for nearly all organizations; many are seeing annual data growth approaching 30%. When taking into consideration the need to create multiple copies of data to fuel production, test/dev, backup and archive initiatives, organizations are maintaining an average of 6.5 copies of data.

In contrast, budgets for storage and data protection efforts are only growing at a modest 12% rate, which emphasizes that organizations will need to get more value out of their spending to keep the data growth challenge manageable. As workloads grow, performance requirements also increase for modern data protection platforms since they will need to retrieve and, in some cases, rehydrate backup data rapidly to restore production workloads in a timely fashion. Further, the gap between business-critical and normal workloads continues to shrink, suggesting that not only are data growth and tight budgets cause for concern, but that now nearly all data is deemed important.

Data Highlight

Figure 3: The Amount of Data Under Management at Organizations Is Enormous... and Growing

Data is expected to grow at a mean annual rate of **24%**

Organizations are managing an average of **6.5 copies of data** (production, backup, archive, etc.)

Q. Over the next 12 months, do you expect the amount of data your organization has under management to increase, decrease or not change?

Base: All respondents (n=432)

Q. How many copies of your business data exist across your organization, both on-premises and in the cloud (e.g., primary storage, backups, archives, etc.)?

Base: All respondents (n=232)

Source: 451 Research's Voice of the Enterprise: Storage, Data Management and Disaster Recovery 2022

Technology Considerations
Can your data protection infrastructure scale to match data growth?
Could a generalist IT professional run a restoration operation in the event of a disaster?
Are tools in place to validate backups from silent or intentional data corruption?

Business Considerations
How much damage would a lengthy outage create for a production workload?
Does your organization have standardized policies regarding recovery SLAs, the company's reputation and customer loyalty?
Is data protection automatically implemented when an application is provisioned?

What to Look for

To deliver adequate data protection capabilities on modest budgets, organizations must find ways to reduce their operational costs. Modern data protection tools need to be easy to use compared with legacy platforms, which were managed by dedicated specialists who are becoming more costly and difficult to find. Given that IT generalists are now handling much of the operational burden, data protection tools need to be intuitive so that generalists will be able to fulfill backup and restore requests without the aid of specialists.

Another way to greatly reduce the operational burden is through the implementation of automation, which can accelerate the provisioning of new resources when needed and can also reduce the burden of backup and disaster recovery testing. Automation can make the testing and recovery process far faster, which will give customers a higher success rate with their recovery operations. One-click recovery of complex workloads can also be facilitated with automation to reduce downtime. The documentation of recovery procedures and the creation of runbooks are important but time-consuming processes that could be automated. This automation can ensure that IT governance and operation control is preserved during the recovery process. Modern workloads should have data protection policies implemented when they are launched and provisioned to ensure no data protection gaps are exposed after an outage.

Modern data protection should provide a scale-out backup repository capability to manage backup storage across multiple sites while enforcing the data residency requirements, which will be extremely important for enterprises struggling with rapid data growth and stringent data retention requirements.

For many organizations, even non-business-critical workloads such as reporting and test/dev have recovery point objective and recovery time objective expectations that exceed the capabilities of daily backups and make technologies such as snapshots, continuous data protection and replication necessary since they can provide the required level of intraday protection.

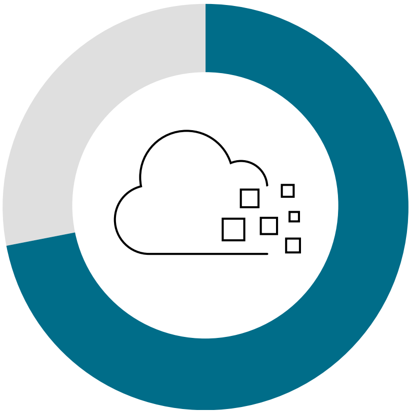
Integrating Cloud into Data Protection Strategy

Most organizations use public cloud storage services within their data protection strategies; only 28% of respondents in the study prefer to handle their data protection exclusively with on-premises resources. The hybrid cloud deployment, which runs backups locally on-premises and copies older backups to cloud storage, was the preferred model for 51% of respondents. SaaS-based offerings such as online backup and disaster recovery as a service (DRaaS) are also potential options for organizations that do not have on-premises data protection, though only about 20% prefer going this route.

The move toward cloud-based data protection is being driven by a variety of benefits for which public cloud services are known. Elasticity to immediately scale up or scale down resource consumption ensures that backup and disaster recovery operations do not run out of cloud compute or storage resources, even when demanding recovery requests come in. The ability to recover to a cloud site is another major benefit since it eliminates the need for an organization to set up and maintain costly failover sites at secondary datacenters.

Data Highlight

Figure 4: Cloud-Based Data Protection Is in Steady Use



72% of respondents are already using cloud-based data protection in the form of hybrid cloud, online backup or DRaaS

Q. Which of the following best describes your organization's current use of data protection (e.g., backup, disaster recovery)?
Base: All respondents (n=387)
Source: 451 Research's Voice of the Enterprise: Storage, Data Management and Disaster Recovery 2022

Technology Considerations
Does your backup platform have cloud cost optimization to eliminate unnecessary service consumption?
Is a pilot light DR site adequate, or is a warm standby or active/active site required?
Are physical transports available to retrieve data for large-scale recoveries?
Does the platform write data in a self-describing manner to facilitate data portability across various cloud environments?
Do the tools provide BIOS conversion to move between Windows and Linux clouds?

Business Considerations
Does your organization have recovery SLAs? What is the cost of outage?
Does your organization use multicloud storage now or plan to in the future?
Will the current infrastructure scale to meet requirements without leveraging public cloud?
Will re-platforming developer and licensing costs outweigh the cost benefits of moving to cloud?
Do you need the instant provisioning and elasticity of cloud to meet future time-to-value requirements?

What to Look for

The inefficient usage of public cloud resources can quickly eliminate the potential cost savings these services provide over traditional on-premises infrastructures. Modern data protection tools must be able to use a variety of cloud storage classes ranging from inexpensive archive storage for storing infrequently accessed data to high-performance object storage classes that can facilitate rapid recovery. Intelligent tiering capabilities should be included since they can cut down on long-term storage costs.

Egress charges and API access fees are additional elements that can bloat an organization's monthly cloud bill, and organizations should leverage cloud cost optimization tools to identify and warn them when consumption increases beyond typical usage.

Multicloud support is an important feature since a significant outage to a hyperscaler could prevent organizations from running recovery operations. The ability to use multicloud also provides protection against vendor lock-in and gives an organization leverage when it comes time to renegotiate contracts with a cloud vendor.

For large-scale recoveries and data migrations, physical shuttles have emerged as high-speed transport for moving data. Though networks are still the preferred means for transporting data to and from clouds, larger payloads in the hundreds of TB to PB range can often be moved faster using a provider-enabled physical transport.

Data Protection for Kubernetes and Next-Generation Workloads

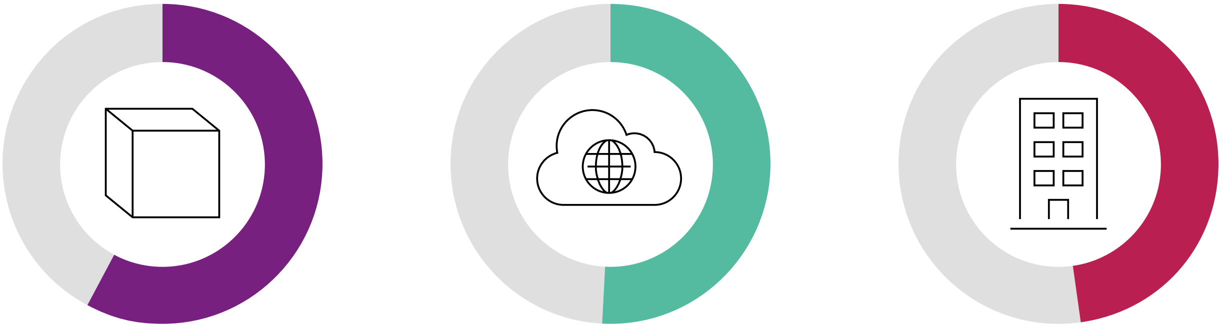
Cloud native is the future for infrastructures and workloads, but it creates challenges for data protection since technologies such as containers have the ability to deliver scalability and workload mobility at speeds that conventional infrastructures cannot match. While the advent of server virtualization and VMs greatly enhanced the utilization of servers and accelerated workload deployment beyond what was possible with physical infrastructures, Kubernetes can deliver provisioning and mobility superior to the standards set by VMs.

As cloud-native adoption, containerization and microservices deployment continue to explode, Kubernetes is becoming the enterprise infrastructure platform of choice. While Kubernetes eliminates the pain of ensuring high availability and scalability of your application services, these benefits do not extend to data. Delivering effective Kubernetes native data protection is one of the most critical imperatives faced by Kubernetes operators today.

Modern data protection tools must be able to keep up with the rapid changes and distributed nature of cloud-native environments. They must also have comparable scalability and agility to ensure recovery operations are comprehensive, with minimal data loss and downtime.

Data Highlight

Figure 5: Organizations Lean Toward Purpose-Built Data Management Tools



58% prefer backup and data management tools that were designed to support container platforms and orchestration

51% prefer to store data for stateful apps in public cloud environments,

while **48%** prefer on-premises storage

Q. What is your organization's primary data protection strategy for containerized applications and data volumes?

Base: Organizations that use containers (n=204)

Q. What is your organization's primary approach to storing data for stateful apps run on containers?

Base: Organizations that use containers, abbreviated fielding (n=213)

Source: 451 Research's Voice of the Enterprise: Storage, Data Management and Disaster Recovery 2022

Technology Considerations
Can your backup tool automatically discover and protect all applications and cluster data, states, configurations, secrets and relevant artifacts?
Can you protect different Kubernetes distributions (e.g., OpenShift, Rancher, etc.) and cloud vendors (Google, Azure, etc.)?
Can your tool protect leading applications and data sources (e.g., Cassandra, MongoDB, Kafka, etc.)
How security-minded is your backup tool in terms of adequate support for authentication, authorization, encryption and ransomware protection?

Business Considerations
Are your organization's development teams increasing their use of containers?
Are developers coding with resiliency in mind?
Is workload portability across hybrid and multicloud environments a key requirement?
Are your software development, DevOps and infrastructure teams capable of extracting the full value of the cloud-native model?

What to Look for

In 451 Research's VotE: Storage, Data Management and Data Protection 2022 study, 58% of respondents said they prefer backup and DR tools that were designed and optimized for cloud-native environments, up from 50% in the 2021 study. Conversely, those who prefer to rely on legacy tools dropped from 35% in the 2021 study to 29% in 2022. Organizations are now realizing that a new approach to data protection is required.

Modern data protection tools must be able to keep up with workload data, but they must also take into account the underlying application components such as the Kubernetes objects and cluster configurations. Next-generation data protection tools must understand the various snapshot procedures for protecting data services such as Cassandra, Kafka and Elasticsearch since they vary among the different data services. Using the wrong procedures for freezing workloads, taking snapshots and unfreezing can lead to data corruption and prevent organizations from doing a successful restoration when needed.

Reporting and advanced management tools with single-pane-of-glass management visibility are also important since they allow customers to quickly identify and fix backup issues. Integration with popular analysis tools such as Prometheus and Datadog are important since they tie data protection closer to cloud-native management and observability initiatives.

Support for modern operational practices is essential, and this includes support for self-service recovery and the ability to support DevSecOps. Modern data protection platforms must also be integrated into Kubernetes-native tooling and should be available for purchase on cloud marketplaces with flexible consumption models that can keep up with dynamic cloud-native environments.

CONTACTS

The Americas
+1 877 863 1306
market.intelligence@spglobal.com

Europe, Middle East & Africa
+44 20 7176 1234
market.intelligence@spglobal.com

Asia-Pacific
+852 2533 3565
market.intelligence@spglobal.com

www.spglobal.com/marketintelligence

These materials have been prepared solely for information purposes based upon information generally available to the public and from sources believed to be reliable. No content (including index data, ratings, credit-related analyses and data, research, model, software or other application or output therefrom) or any part thereof (Content) may be modified, reverse engineered, reproduced or distributed in any form by any means, or stored in a database or retrieval system, without the prior written permission of S&P Global Market Intelligence or its affiliates (collectively, S&P Global). The Content shall not be used for any unlawful or unauthorized purposes. S&P Global and any third-party providers, (collectively S&P Global Parties) do not guarantee the accuracy, completeness, timeliness or availability of the Content. S&P Global Parties are not responsible for any errors or omissions, regardless of the cause, for the results obtained from the use of the Content. THE CONTENT IS PROVIDED ON "AS IS" BASIS. S&P GLOBAL PARTIES DISCLAIM ANY AND ALL EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, ANY WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE OR USE, FREEDOM FROM BUGS, SOFTWARE ERRORS OR DEFECTS, THAT THE CONTENT'S FUNCTIONING WILL BE UNINTERRUPTED OR THAT THE CONTENT WILL OPERATE WITH ANY SOFTWARE OR HARDWARE CONFIGURATION. In no event shall S&P Global Parties be liable to any party for any direct, indirect, incidental, exemplary, compensatory, punitive, special or consequential damages, costs, expenses, legal fees, or losses (including, without limitation, lost income or lost profits and opportunity costs or losses caused by negligence) in connection with any use of the Content even if advised of the possibility of such damages.

S&P Global Market Intelligence's opinions, quotes and credit-related and other analyses are statements of opinion as of the date they are expressed and not statements of fact or recommendations to purchase, hold, or sell any securities or to make any investment decisions, and do not address the suitability of any security. S&P Global Market Intelligence may provide index data. Direct investment in an index is not possible. Exposure to an asset class represented by an index is available through investable instruments based on that index. S&P Global Market Intelligence assumes no obligation to update the Content following publication in any form or format. The Content should not be relied on and is not a substitute for the skill, judgment and experience of the user, its management, employees, advisors and/or clients when making investment and other business decisions. S&P Global Market Intelligence does not endorse companies, technologies, products, services, or solutions.

S&P Global keeps certain activities of its divisions separate from each other in order to preserve the independence and objectivity of their respective activities. As a result, certain divisions of S&P Global may have information that is not available to other S&P Global divisions. S&P Global has established policies and procedures to maintain the confidentiality of certain non-public information received in connection with each analytical process.

S&P Global may receive compensation for its ratings and certain analyses, normally from issuers or underwriters of securities or from obligors. S&P Global reserves the right to disseminate its opinions and analyses. S&P Global's public ratings and analyses are made available on its websites, www.standardandpoors.com (free of charge) and www.ratingsdirect.com (subscription), and may be distributed through other means, including via S&P Global publications and third-party redistributors. Additional information about our ratings fees is available at www.standardandpoors.com/usratingsfees.