

Technical Validation

# Dell EMC Isilon F810 with Inline Data Compression

## Efficient File Storage without Performance Penalty

By Kerry Dolan, Senior IT Validation Analyst

May 2019

This ESG Technical Validation was commissioned by Dell EMC and is distributed under license from ESG.



## Contents

- Introduction ..... 3
  - Background ..... 3
  - Storage Efficiency and Performance are Natural Enemies ..... 4
  - Dell EMC Isilon F810..... 4
  - Isilon F810 Compression Details ..... 5
- ESG Technical Validation..... 6
  - ESG Testing..... 6
- The Bigger Truth..... 9

### ESG Technical Validations

The goal of ESG Technical Validations is to educate IT professionals about information technology solutions for companies of all types and sizes. ESG Technical Validations are not meant to replace the evaluation process that should be conducted before making purchasing decisions, but rather to provide insight into these emerging technologies. Our objectives are to explore some of the more valuable features and functions of IT solutions, show how they can be used to solve real customer problems, and identify any areas needing improvement. The ESG Validation Team’s expert third-party perspective is based on our own hands-on testing as well as on interviews with customers who use these products in production environments.

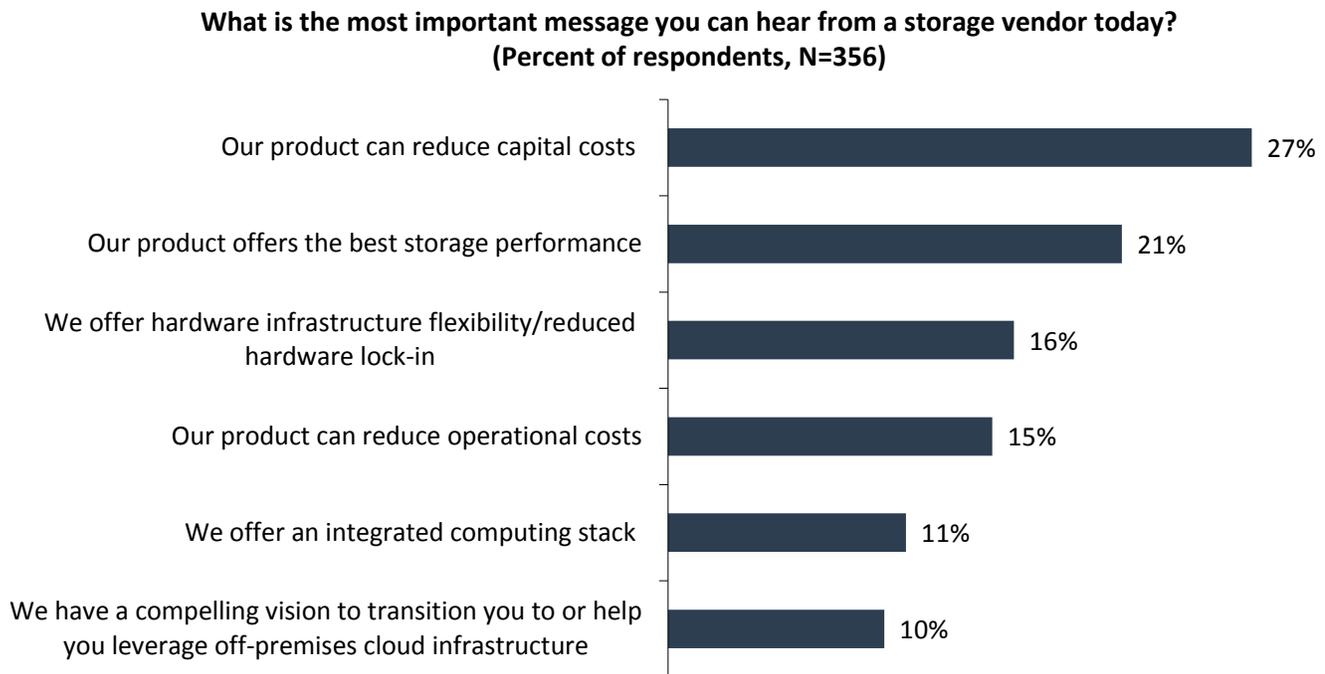
## Introduction

This report documents ESG’s audit of compression and performance testing of the Dell EMC Isilon F810 all-flash file storage array.

## Background

Data is a corporate asset of tremendous value, and organizations are collecting, storing, and analyzing masses of it to drive business decisions. Unlocking the potential of unstructured (file) data—the highest growth data type—can improve productivity, innovation, and corporate strategy. But what are the key features required to store all this data? When ESG asked survey respondents to name the most important messages they could hear from storage vendors, performance and reduction in operational costs were among the top four (Figure 1).<sup>1</sup>

**Figure 1. Storage Objectives**



*Source: Enterprise Strategy Group*

In past years, the only critical feature of file storage arrays was the ability to scale. “Big data” and analytics projects in particular have contributed to the huge escalation of file data, as organizations have realized that digital transformation can help them understand customers and trends to make better business decisions. Traditional file storage was built on the assumption that only a small percentage of data was generally active or “hot,” needing fast performance; the bulk of file data was inactive and could reside on slower media. In today’s data-driven economy, however, scale is necessary but not sufficient; speed is essential for delivering the real-time insights that analytics can generate. Solid-state drives can deliver the fast performance needed, but at higher cost, particularly for large amounts of data. But organizations need to use all of their file data, not just subsets of it, to understand their businesses and remain agile. These dual demands of massive file storage capacity and high performance can strain budgets, making file storage efficiency another essential characteristic. Data reduction technologies can enable more efficient storage for file data and reduce the costs of these essential resources.

<sup>1</sup> Source: ESG Master Survey Results, [2017 General Storage Trends](#), November 2017.

## Storage Efficiency and Performance are Natural Enemies

In practice, achieving storage efficiency and performance are normally conflicting objectives. The computations required to compress or deduplicate data sets pull CPU and memory resources away from data access activities. As a result, most IT organizations are well-acquainted with the tradeoffs between performance and data reduction. For example, electronic design automation (EDA) software, used to design silicon chips for use in mobile devices and laptops, often requires numerous concurrent tasks using many large files; as a result, these applications struggle with storage controller bottlenecks that impede performance. Typically, adding data reduction activities will decrease performance even more.

### Dell EMC Isilon F810

Dell EMC Isilon has been a leader in file storage solutions for more than a decade. The latest model, the Isilon F810 all-flash, scale-out NAS array, includes hardware-based, inline compression technology that is powered by the Isilon OneFS 8.1.3 operating system. This hardware-based compression enables storage efficiency with minimal performance impact, enabling organizations to gain efficiency without sacrificing performance.

The Isilon F810 fits into existing Isilon clusters and provides massive scalability plus performance. Dell EMC claims a single, four-node chassis can deliver up to 250K IOPS and 15GB/sec aggregate throughput, with a maximum of 144 nodes per cluster for 9 million IOPS. With the Isilon F810, storage efficiency is a hallmark. For example, the F810 provides a raw storage capacity of up to 33 PB in a 144-node cluster, which is enhanced by up to 3:1 data reduction (depending on the data set) for up to 79 PB effective capacity. Key efficiency features of the Isilon F810 are shown below

- Hardware-based, inline compression provides continuous compression of writes and decompression of reads. Compression is on by default, requires no configuration, and can be disabled with a single command.
- Isilon achieves 80%+ storage utilization due to its use of erasure coding and file-level protection instead of drive-level protection delivered through defined RAID sets and dedicated spares.
- OneFS distributes data evenly among nodes, continuously reallocating data for space conservation, and eliminating the need for reserve capacity overhead.
- Isilon SmartDedupe works in conjunction with compression to further reduce the storage used in some cases.
- As part of Dell EMC's Future-Proof Loyalty Program, the company guarantees that the Isilon F810 will provide logical usable capacity, including all data, equivalent to at least twice the usable physical capacity for one year from date of delivery.

Like all Isilon arrays, the F810 supports data access via NFS, SMB, HDFS, HTTP, and FTP, enabling a range of applications and workloads. Each node includes 256GB memory, and both 10GbE and 40GbE network interfaces. It offers N+1 through N+4 redundancy as well as backup and disaster recovery software options. It has no single point of failure, is self-healing, and includes intra-cluster failover.

All Isilon enterprise features remain available with inline compression. The OneFS operating system creates for each cluster a single file system and single global namespace. Security features include role-based access control, secure access zones, WORM data immutability, OneFS hardening, file system auditing support, and optional self-encrypting drives. In addition to SmartPools and CloudPools mentioned above, the Isilon F810 supports the full range of ecosystem software, including SmartDedupe deduplication; SyncIQ replication; SnapshotIQ data protection; SmartLock policy-based retention; SmartQuotas data management; InsightIQ performance management; and SmartConnect client connection load balancing and failover.

Figure 2. Dell EMC Isilon F810

### Isilon F810 with Hardware-based Inline Compression



- Up to 79 PB effective in a single cluster
- Retains all OneFS enterprise services



Up to 3:1 data compression



Negligible impact to performance



Reduces SSD wear rates



Savings on power, cooling, management

Source: Enterprise Strategy Group

### Isilon F810 Compression Details

With Isilon F810 inline compression, organizations can deliver high application performance while simultaneously enabling storage efficiency to keep costs down. Dell EMC chose hardware-based compression to maximize the compression capability and to reduce the performance impact of compression. Below is a high-level overview:

- The all-flash configuration delivers fast data access for better application performance.
- Lossless compression is performed inline as data is written to the Isilon F810.
  - Data is divided into 128KB chunks within node pools, and then compressed into 8KB blocks.
  - The compression activities are offloaded, with the calculations performed on a specialized, back-end NIC that leverages the OneFS 8.1.3 operating system to compress data. The compression algorithms identify and reduce or eliminate redundant bits so that less physical data is sent to the SSDs.
  - No separate license is required to enable compression on the F810 node.
- When compressed data is read, it is decompressed to its original form.

Offloading compression and decompression activities enables data reduction without involving node CPU and memory, resulting in minimal application latency and making compression/decompression transparent to applications and workflows. These activities reduce the costs of storage, power, cooling, and data management, while enabling high-performance data access. In addition, by writing less data to SSDs, compression reduces their wear rates, enabling improved SSD durability and life span.

Workloads that suffer similar challenges—and, therefore, may be good targets for Isilon F810 compression—include genomics, EDA, software development, and even general files. However, it is essential to understand that every data set is different, even in the same type of workload, with some data compressible and some not. Many factors will contribute to the ability of a workload to be compressed, including data types and amounts, file sizes, data layouts, protection levels, and random or sequential data access.

## ESG Technical Validation

ESG validated testing done by Dell EMC engineering teams in Shanghai, China. Testing included evaluating compression ratios in real-world data sets and performance with compression activated.

### ESG Testing

#### Compression

First, ESG reviewed Isilon F810 compression ratios for various data sets that Dell EMC had collected. Data included files from home directories; oil and gas exploration; electronic design automation; software build; media and entertainment, including images and video; and life sciences.

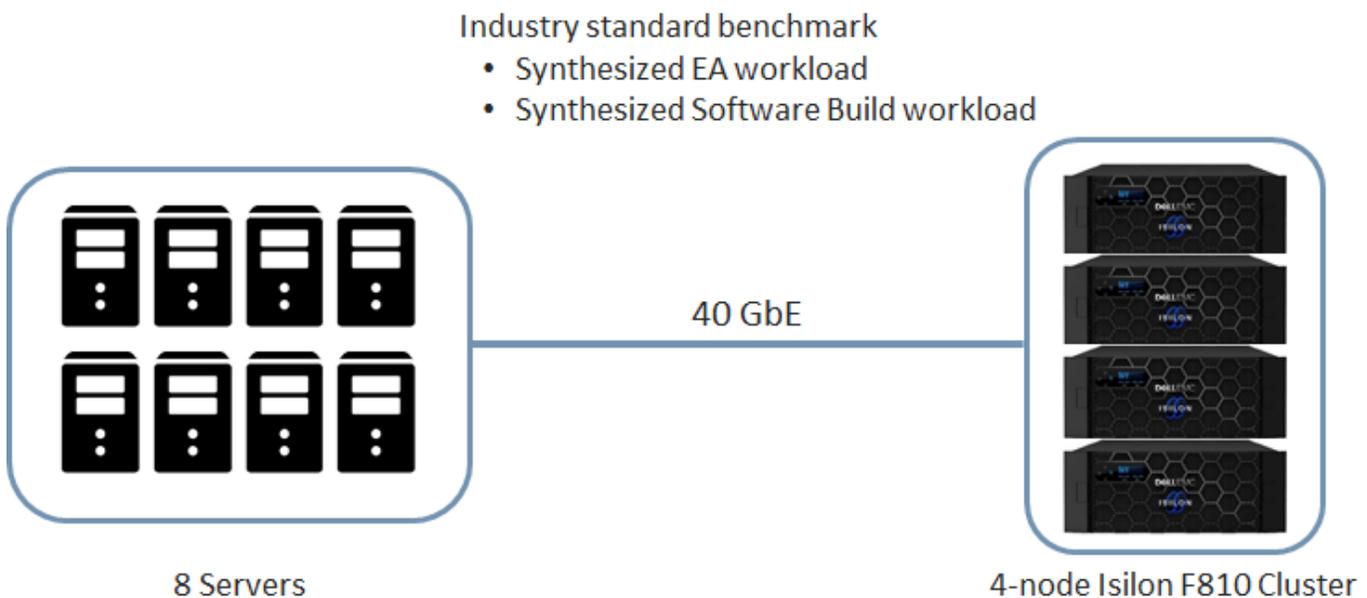
In our review, ESG noted that tested compression ratios ranged from 1:1 to 3.93:1 depending on the data set. The data exhibiting 1:1 compression will consume about the same amount of physical storage on the Isilon F810 as uncompressed data, while the 3.93:1 compressible data will take almost 4X less physical storage than uncompressed.

As we have noted, compression ratios vary by data set and within data sets. For example, some data in an EDA process may be very compressible, while data in other parts of the process may not be; most images and video are already compressed and so are not very compressible. The key takeaway is that both scenarios—compressible and uncompressible data—are abundant in real world data sets. The Isilon F810 can deliver significant savings on storage capacity, power, cooling, and management for the right workloads. Dell EMC’s guidance is that any data set exhibiting at least 1.5:1 average compressibility is a good fit for the F810. Less compressible data might be a better fit for the F800 platform, with no compression.

#### Performance

Next, ESG looked at how workloads on the Isilon F810 were impacted by turning on compression. The test bed included a four-node Isilon F810 cluster. Each Isilon node was configured with 16 Intel Xeon E5 CPU cores, 256GB RAM, 225TB of SSD, and 40GbE networks. On the front end, eight servers, each with 128 GiB memory, were used for load generation.

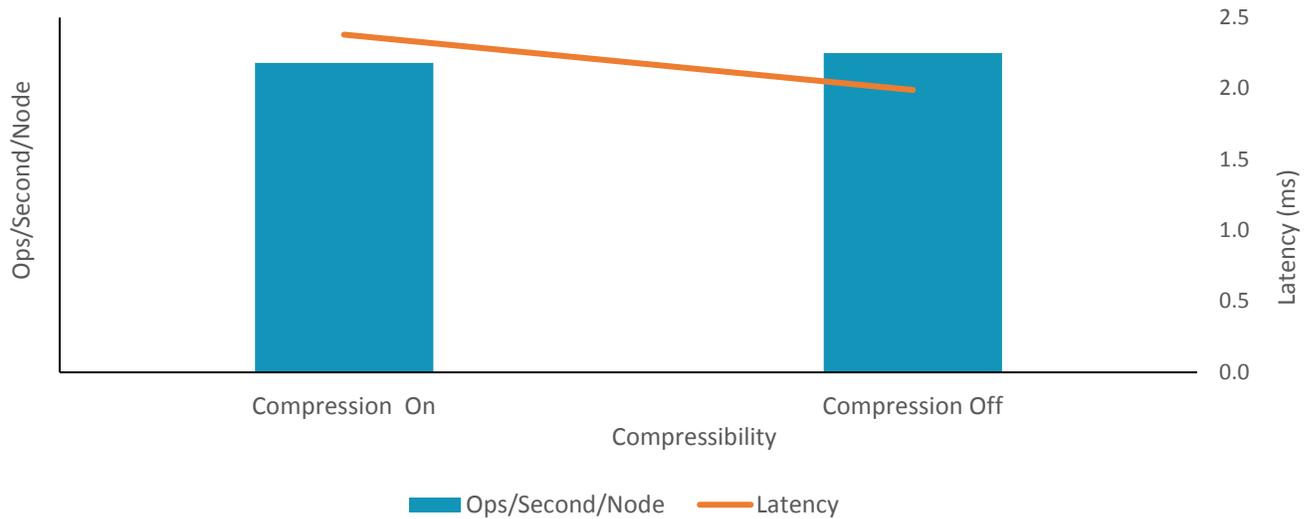
**Figure 3. Performance Test Bed**



Source: Enterprise Strategy Group

The performance test harness was an industry-standard benchmark tool that measures performance of various emulated workloads. The first test used EDA data that was 31% compressible. As Figure 4 shows, performance dropped only 3.13% when compression was enabled. Latency varied from just under 2 ms to just over 2 ms, which would not be noticeable to the application user.

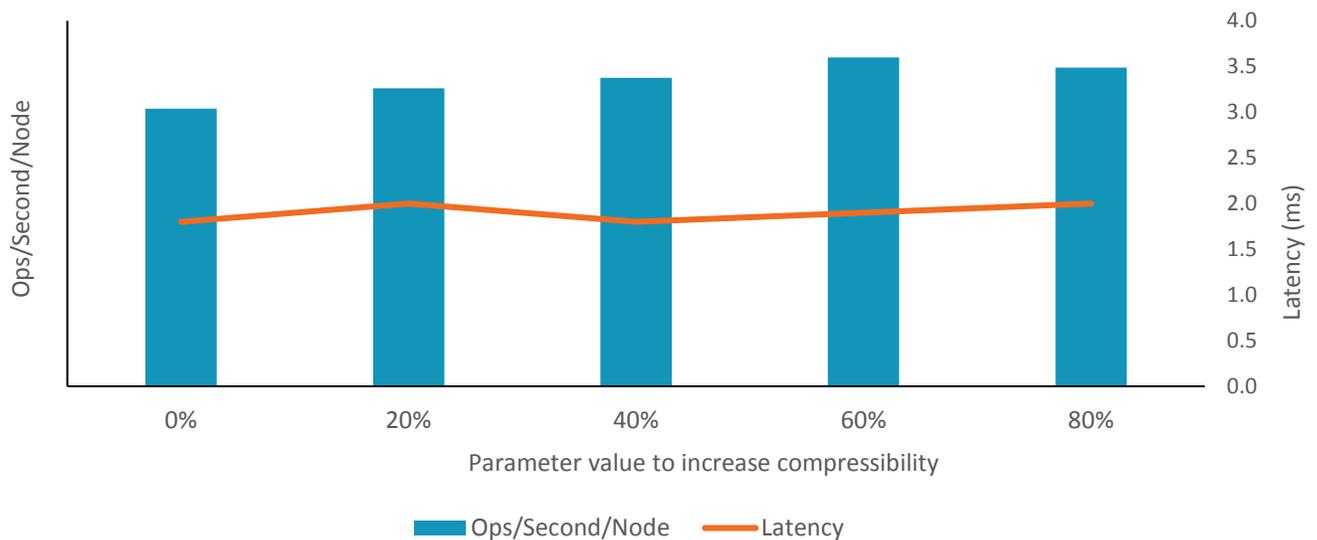
**Figure 4. EDA Performance with Compression**



Source: Enterprise Strategy Group

Next, ESG looked at the performance as EDA data compressibility increased. Figure 5 shows that performance increased as compressibility increased in 20% intervals; performance was 18% faster at 60% compressible than at 0% compressible. While performance dropped somewhat at 80% compressible, it was still 14% faster than the 0% compressible data. Latency remained virtually the same, hovering around 2 ms, regardless of data compressibility.

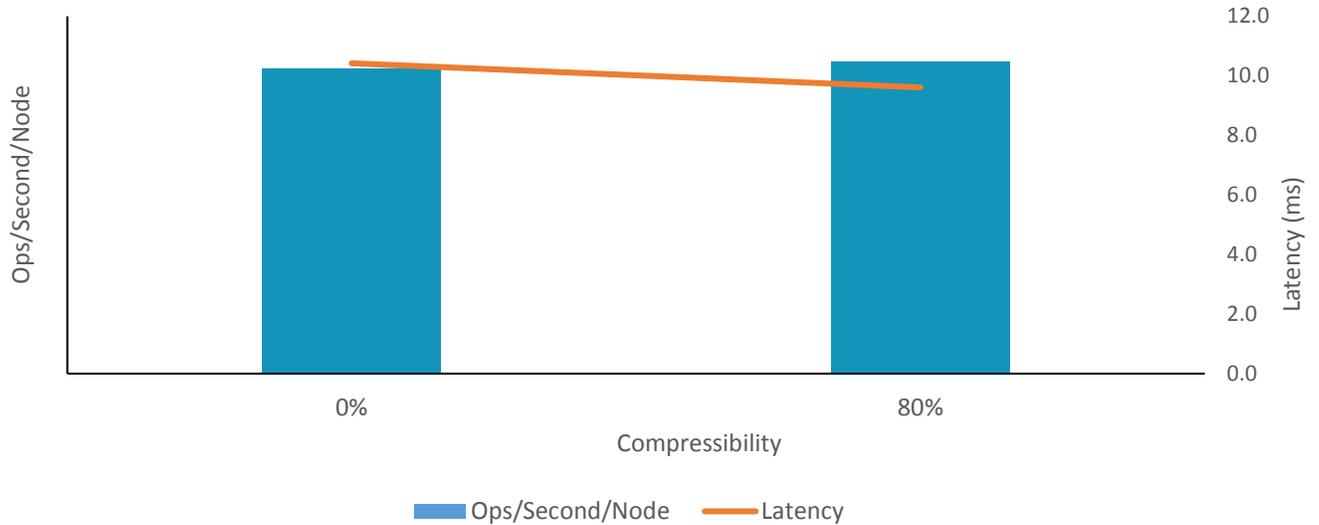
**Figure 5. Performance as Compressibility Increases**



Source: Enterprise Strategy Group

Finally, comparing 0% and 80% compressible software build data, the more compressible data showed 2% faster performance in ops/sec/node, with almost 8% lower latency.

**Figure 6. Software Build Data**



Source: Enterprise Strategy Group



### Why This Matters

File data continues to grow massively, and organizations want to leverage as much of it as possible for analysis. Scale is no longer the only requirement for storing file data; it must be quickly accessible, and efficiently stored to save money. However, data reduction technologies such as compression typically impede performance, forcing organizations to choose between cost efficiency and speed.

Dell EMC designed the Isilon F810 with a compression-offload NIC to enable scale and efficiency with minimal performance impact. ESG validated several data sets that demonstrated better performance with higher compression, without increasing latency. This enables organizations to gain the efficiency benefits of lower storage, power, and cooling costs, while enabling the performance that applications need.

## The Bigger Truth

File data is being used on a massive scale as organizations mine all the resources they have for analysis. But traditional file storage was designed for scale, and not for speed or efficiency. Today, all three are necessary—organizations need to access masses of data, with fast performance, and with efficiency to reduce the costs of storing and managing it. In fact, when asked about the most important objectives for their digital transformation initiatives, 55% of ESG survey respondents indicated that they wanted to “become more operationally efficient,” making it the most-cited objective.<sup>2</sup>

Dell EMC’s commitment to efficiency is demonstrated with the Isilon F810 all-flash array with hardware-based, inline compression. This feature adds to Isilon’s suite of data efficiency technologies, including post-process data deduplication; some data sets get more from inline compression, while others do better with deduplication. Dell EMC provides both to optimize efficiency. Additional efficiency technologies include high storage utilization based on using erasure coding for protection instead of defined RAID sets and sparing, and automatic data tiering to keep only the required data on fast storage and move cold data to less expensive media. Compression is easy to manage, and for the right data sets, enables cost savings without sacrificing performance.

Isilon F810 delivers all the same OneFS functionality and enterprise services as other Isilon arrays but adds inline compression. It remains easy to deploy, use, and expand for performance and capacity, and it retains all the advantages that flash drives deliver—density, performance, and power—while also providing more storage efficiency to reduce costs.

ESG Lab validated compression ratios of up to 3.93:1 for various workloads. We also witnessed increasing performance with more compressible data, with negligible latency impact. Organizations should feel confident that they can take advantage of potential savings without impacting application performance.

Every data set is different, and some will gain more from compression than others. This feature is not designed for the most heavily accessed data or the data needing the highest performance, but for most compressible applications, it reduces storage capacity and costs with minimal performance impact. Organizations looking for the benefits of flash, the storage efficiency of data compression, and all the enterprise services of OneFS should evaluate the Isilon F810.

---

<sup>2</sup> Source: ESG Master Survey Results, [2019 Technology Spending Intentions](#), March 2019.

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of The Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at 508.482.0188.



**Enterprise Strategy Group** is an IT analyst, research, validation, and strategy firm that provides market intelligence and actionable insight to the global IT community.

© 2019 by The Enterprise Strategy Group, Inc. All Rights Reserved.

