



Research®
Now a Part of

BLACK & WHITE PAPER

S&P Global Market Intelligence

Designing for Edge and IoT Success

Comparing edge and cloud-centric compute models
in several use cases

COMMISSIONED BY



AUGUST 2020

©COPYRIGHT 2020 451 RESEARCH.
ALL RIGHTS RESERVED.

About this paper

A Black & White paper is a study based on primary research survey data that assesses the market dynamics of a key enterprise technology segment through the lens of the “on the ground” experience and opinions of real practitioners — what they are doing, and why they are doing it.

This paper is commissioned by Dell Technologies.

Welcome to the digital future where every organization needs to be a digital organization – powered by data, running in a multi-cloud world.

The digital future demands a data-first perspective. Rather than start by trying to determine what any particular cloud can do, begin by asking what your data can do for you – then ask what architecture (edge to core to cloud) will get you there, in the most efficient and effective way. In this study you will see that Edge and IoT data should not be framed as an ‘edge vs. cloud’ topology; rather, the two should work in tandem where it makes the most sense for the purposes of performance, security, reliability and cost.

ABOUT THE AUTHOR



CHRISTIAN RENAUD

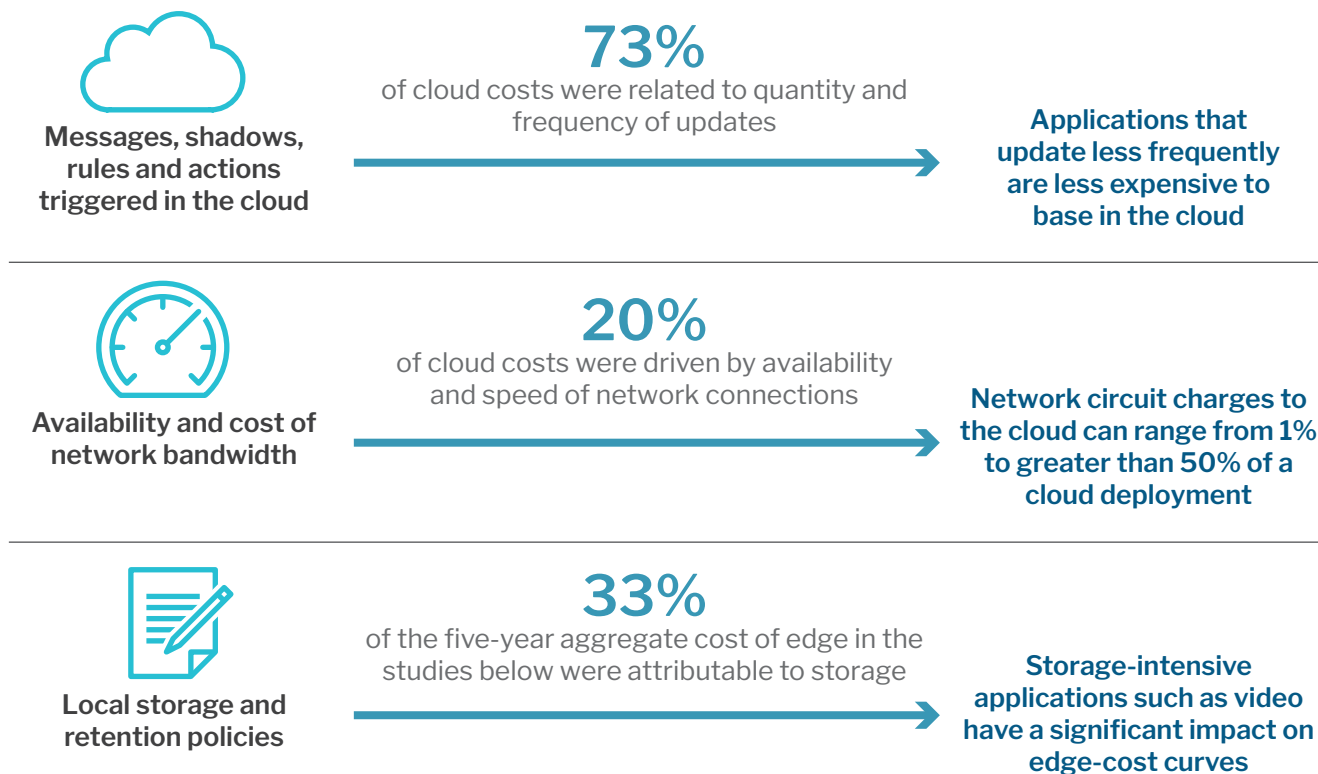
RESEARCH VICE PRESIDENT,
INTERNET OF THINGS

As Research Vice President of 451 Research’s Internet of Things practice, Christian Renaud covers the ongoing virtualization and digitization of the physical world around us.

For nearly one year, 451 Research worked with end-user organizations in manufacturing, oil and gas, and smart cities to identify the key drivers in determining the optimal location to execute IoT workloads. These interactions yielded extensive data on application demands for bandwidth, latency and storage, which in turn allowed 451 to model the five-year cost of these applications at the edge and in the cloud.

Figure 1: Key cost considerations across use cases

Source: 451 Research Custom Research



Key levers impacting workload location included:

- The requirements of real-time applications, such as oil and gas production. These requirements favored edge computing over cloud computing given the latency inherent in wide area networking (WAN) connections, if available, to the cloud.
- A key cost driver for cloud compute IoT deployments was the frequency of messages between the cloud and the end asset. If these are simple keepalive messages, triggers or rules, they add up to increased traffic, compute and cloud storage, which was the primary driver in cloud compute scenarios.
- Availability and speed of network connections for many IoT applications, specifically in oil and gas, is inconsistent. These network costs were a key lever in many of the use cases detailed below.

- The volume of data generated from video and radar sensors is considerable (ranging from tens of terabytes to petabytes). The cost of WAN links, cloud compute, and long-term storage of this volume of data heavily impacted the five-year costs involved in cloud-centric scenarios, with costs of 1.7x to 61x over comparable edge computing scenarios.

Of the six use cases detailed below, edge computing was less expensive over a five-year period than comparable cloud-based implementations. These industries and applications were heavily skewed toward real-time and other production applications, which naturally demand more than casual, low-frequency applications in less real-time industries.

Digital Transformation and Emerging Applications

An integral step in any organization's digital transformation journey is the adoption of emerging technologies that drive transformation. We often hear organizations claim that emerging trends are hype; we're here to prove that the emerging technologies are reality, and early adopters of these trends are driving their organizations ahead of the competition. When we refer to 'emerging technologies,' we are referring to the data-driven trends such as the Internet of Things (IoT), artificial intelligence (AI), augmented and virtual reality (AR/VR), and blockchain.

This report focuses on the important step in the digital transformation journey that includes edge computing – specifically IoT – or connecting new or existing equipment in order to extract data for analysis, insight and action.

Edge computing consists of pushing networking, compute and storage as close to the IoT device generating data as possible; 78% of all IoT data is first analyzed at one or more edge compute locations.

According to respondents to 451 Research's Voice of the Enterprise (VoE): Digital Pulse 2019 survey, 61% of organizations are underway in their own digital transformation journey, and an additional 37% anticipated they would be within the next 12-24 months. IoT project deployments are equally strong: 68% of respondents to 451 Research's VoE: Internet of Things, Workloads and Key Projects 2019 survey indicated that they are in production or proof of concept with one or more IoT projects, and an additional 22% anticipated they would be within the next 12-24 months.

While use cases for deploying IoT technologies vary from industry to industry, mature industries such as manufacturing, oil and gas (O&G), and transportation primarily focus on taking costs out of the business, reducing overhead and worker safety. In contrast, more recent adopters of sensors and analytics have begun with growing the top line, increasing customer targeting (retail), or increasing the quality of care (healthcare). Running alongside these key business drivers are IT organizations, which are looking at the potential deluge of operational (machine, environmental and even biological) data and trying to fit it into existing architectures and tools.

This has resulted in a large amount of experimentation by IT teams to grow in-house capabilities in advance of or in response to these projects. While some projects have clear objectives and outcomes, others are less well defined, resulting in only 53% of IT-driven IoT projects moving from the drawing board to production. Two key drivers to success in IoT projects are a clear understanding of the intended outcome(s), which requires understanding of the applications and workloads themselves, and a firm grasp of the costs involved in delivering against these outcomes. This serves not only IT teams that want to grow internal skills, but gives business units and operational leaders the data they need to use in refining and modeling business operations.

Location, Location, Location – IoT Workload Execution Venues

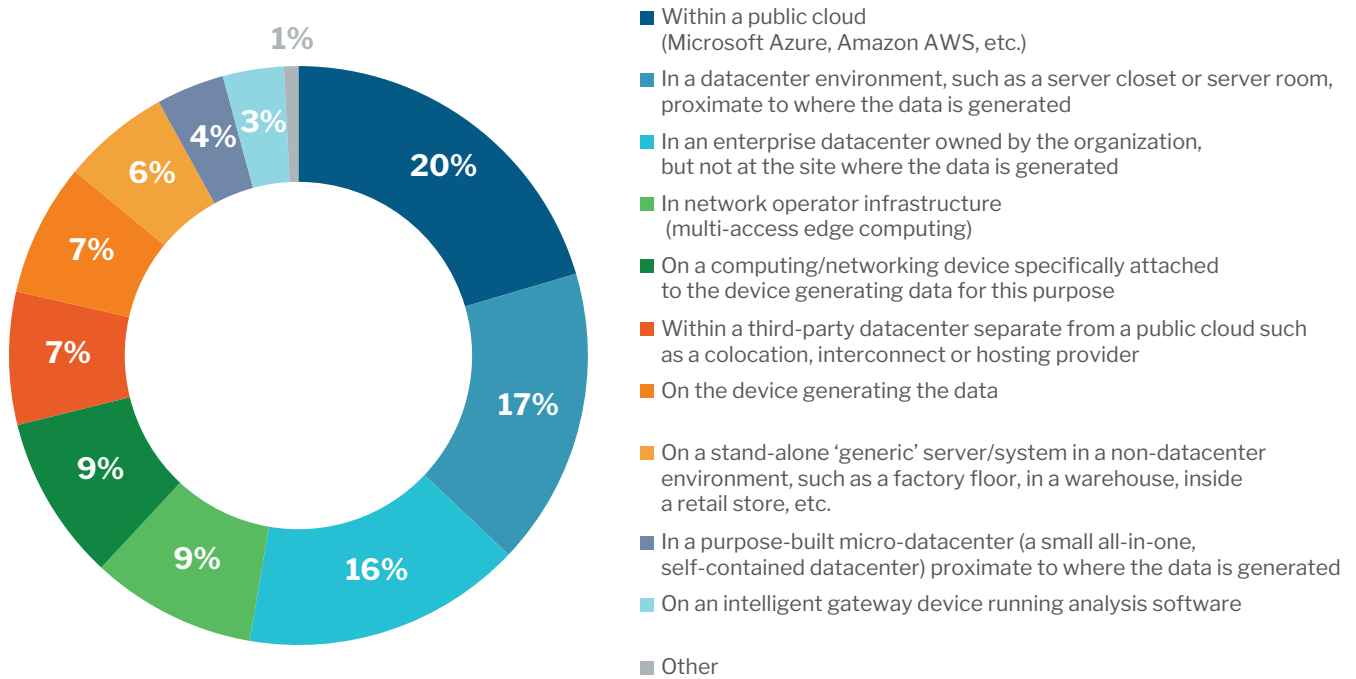
The first of these two drivers – understanding the applications being run to achieve the business outcome, such as lower cost of operation or less unscheduled downtime – requires that you first inventory the requirements of the workload itself.

- Does the application require data and analytics occur at a particular speed (i.e., low latency), such as a manufacturing assembly line or electric grid?
- How much bandwidth does it require to reach any related data sources or analysis tools?
- Is the data sensitive in nature and require a heightened level of security?
- Is the data subject to regulation or data protection laws?
- Does your application require a specialized device to acquire the data from the originating equipment?
- Are there AI models in the cloud that require training data for local inferencing and action?

The answers to these questions dictate where application analytics occur. It should come as no surprise, therefore, that when enterprises were asked where they are performing these analytics today, survey respondents indicated more than 10 'landing pads' for IoT analytics processes, ranging from on-device analytics at the edge to analytics in the cloud. While the largest by sheer percentage is analytics in the public cloud at 20%, a total of 46% of the responses indicated locations on-premises close to where the data originates, and an additional 32% were 'near edge' locations that may serve as the first execution venue of IoT analytics, such as third-party datacenters or a network operator network (multi-access edge).

Figure 2: Location of first IoT analysis

Source: 451 Research's Voice of the Enterprise: Workloads and Key Projects 2019



Sample edge and near-edge use cases range from production monitoring and condition-based maintenance in the manufacturing sector due to time sensitivity, wellhead monitoring and O&G exploration in the utilities/O&G sector due to connectivity limitations, and video surveillance and environmental monitoring in cities and government due to data volumes. Future workloads will drive an increase in this trend of edge computing and analytics due to the sheer volume of data generated by video and other sensors such as LiDAR (light detection and ranging) and Radar.

Key Considerations

451 Research has documented six case studies in detail. They are based on nearly a year of intensive quantitative analysis and qualitative in-depth interviews with enterprises and government agencies currently deploying IoT projects. Through this work, we identified a number of key variables that play a role in how IoT efforts are architected and the impact of those decisions on the total project cost over a five-year period.

The top three 'levers' that impacted the cost of these deployments were:

- The frequency of updates, rules and triggers required for cloud computing and digital twins.
- The availability and cost of inexpensive, high-speed WAN broadband connections for cloud connectivity.
- The amount of persistent storage required for high-traffic applications such as video analytics in manufacturing quality or video surveillance in cities.

Ultimately, we discovered that the more frequently an application updates – for example, the collection of time-series data such as vibration or temperature – the more this can drive up the cost of cloud computing, especially when paired with digital twins and shadows. Twins and shadows that require near-real-time updates require a constant stream of device data, which impacts the costs of cloud compute, network bandwidth and storage.

IoT data should not be framed as an 'edge vs. cloud' topology; rather, the two should work in tandem where it makes the most sense for the purposes of performance, security, reliability and cost.

These findings point to use cases that naturally align with edge computing topologies, and others that are more ideal for cloud architectures. We recognize that much of the post-processed data from edge applications will ultimately be incorporated into cloud models and historians, so IoT data should not be framed as an 'edge vs. cloud' topology; rather, the two should work in tandem where it makes the most sense for the purposes of performance, security, reliability and cost.

To model the costs of your own deployments, you need to collect several key parameters that will ultimately inform your placement of IoT workloads. These include:

- Volume of data generated by the application
- Frequency of analysis
- Availability and cost of a high-speed WAN
- Data storage and retention policies for the applications

Interoperability and Standards

A hidden variable in the cost of any IoT deployment can be the intensive development frequently required to connect to and reconcile data siloes or inoperable sources of data. This stems from the legacy equipment that needs to be connected, often via proprietary interfaces and protocols. This proliferation of vendor-specific connections negates much of the value and promise of the Internet of Things in the unlocking of previously inaccessible data to analytics, as well as enrichment and correlation with other data sources, within an organization. Freeing the data from these walled gardens opens up the data to multiple analysis tools and allows the opportunity to combine it with adjacent data tools and business systems such as manufacturing resource planning, asset performance management and customer relationship management (CRM). Early cloud solutions for IoT were equally 'walled,' and while this hasn't yet been entirely resolved, barriers to interoperability and data accessibility are coming down, reducing the potential for high pricing or lock-in.

One notable organization working on interoperability at the edge of the IoT network is the EdgeX Foundry open source project within the Linux Foundation. EdgeX represents an open source community of more than 100 contributors working together on interoperability and compatibility among vendor edge solutions in the market. In order for IoT to be cost-effective and reach its true potential around the monetization of data, resources and IP, standards-based edge devices will need to work in tandem with any cloud collaboratively, utilizing standards-based interfaces.

Case Studies

Methodology

The case studies that follow are the result of quantitative and qualitative research performed by the 451 Research IoT and Cloud Economics analyst teams between July 2018 and March 2019. This research included quantitative surveys of enterprise IoT practitioners – both IT decision-makers and operational technology (OT) professionals – as well as in-depth interviews with multiple end users in each industry vertical. The six workloads described below span three industries and are the result of 16 telephone and in-person interviews with subject-matter experts and implementors. These subject matter experts were distributed across the US, Western Europe and Japan.

In addition to the survey and interviews conducted, 451 Research developed a custom application to query the pricing tools of leading cloud providers. This work extensively leveraged the Cloud Price Index (CPI) service of 451 Research. The Cloud Price Index is 451 Research's benchmarking service covering the costs of public, private and managed clouds, and the foundation of 451 Research's Cloud Economics practice, alongside the Cloud Price Codex, our global survey of cloud pricing methods and mechanisms. The CPI was used extensively to benchmark cloud pricing variables for each of the workloads described below across the three leading public cloud providers of IoT services: Microsoft Azure, Amazon Web Services and Google Cloud IoT.

The deployment costs of these use cases were compared when deploying these workloads in the cloud and at the edge. Numerous factors, including hardware costs, support, network equipment requirements, circuit costs, labor costs and electricity were considered in the economic analysis. For the cloud computing scenarios, the variables contributing to pricing included keepalive ratios, messages per day (and size of messages), shadows and twins, rules, actions, and frequency of new device registrations. These were determined by use case and ended up being a key lever driving overall cloud pricing.

Staffing and labor costs in both edge and cloud scenarios in all six use cases were determined to be too similar in both scenarios to be relevant to the analysis and were ultimately excluded. This was due to similar head count requirements to provision and deploy an on-premises or cloud-based deployment. In each case, two to three full-time engineers for three months were modeled, followed by fractional support of either on-premises staff to maintain the servers, network equipment and storage while operational, or the ongoing maintenance of on-premises network routers to the cloud, and the part-time services of a cloud architect/systems administrator. These costs were nearly identical, and ultimately served no useful purpose in comparing venues.

Workload Requirements

In each of the cases, the applications were characterized for data generation rates, data retention percentages (storage), data collection intervals, digital shadows and connected uptime, triggers, rules and bandwidth. These workloads were then scoped for both edge implementations, requiring on-premises compute and storage, and cloud execution, which requires edge routing, a WAN connection, and cloud compute and storage resources. These two execution venues were then compared as far as total cost of ownership over a 12-, 36-, and 60-month period.

For annual support, estimates were based on current vendor pricing for five-year support contracts on appropriately specified hardware (low end and high end) and, where unavailable, were estimated at approximately 15% of the original purchase price of the equipment over a five-year period for ongoing software and maintenance support. For cloud computing, these support costs were included in the annual cloud price calculations.

For electricity, the average power consumption of each edge device was factored into the model on a device-by-device basis. This is true for both the edge use case (on-premises edge server(s) and storage) and for the cloud model (network router).

Real estate considerations were not included given the diverse nature of these environments. This is a factor that needs to be considered for the more intensive edge use cases, such as video analytics, where considerable storage may be required on-premises. Also not included were any 'salvage' costs of on-premises hardware, compute or networking, either for resale or redeployment.

Manufacturing Case Studies

The North American Industry Classification System describes the manufacturing sector, which comprises nearly 16% of gross world product according to the World Bank (2016), as the following:

The Manufacturing sector comprises establishments engaged in the mechanical, physical, or chemical transformation of materials, substances, or components into new products. Establishments in the Manufacturing sector are often described as plants, factories, or mills and characteristically use power-driven machines and materials-handling equipment. Manufacturing establishments may process materials or may contract with other establishments to process their materials for them. Both types of establishments are included in manufacturing.

This range of classification means that manufacturing data requirements and instrumentation vary greatly across the board from single status byte messages to a high throughput of time series data. For the use cases within this section, one semiconductor manufacturing quality application was characterized using machine instrumentation, and a second application was characterized for discrete filter manufacturing with production data obtained from programmable logic controllers (PLCs).

Discrete Semiconductor Manufacturing Quality and Yield Monitoring

A semiconductor manufacturing firm assesses the quality and yield rates of the manufacturing process by evaluating the overall equipment effectiveness and the quality of product output. It tracks the location of production line items and evaluates test results. Historical data is important to the process. In the flow of the manufacturing process, anomalies need to be detected early before they can compound losses during secondary assembly or processing. While there is no regulatory requirement to retain data, 'customer track back' and warranty policies require long-term storage to demonstrate manufacturing quality. The company identified that a single day's outage or production failure in one fabrication line could cause a loss of about \$155m.

The semiconductor manufacturing process utilized by the firm is made up of as many as 60 discrete steps with four or five sub-steps within each. Several hundred tests – visual, electrical and mechanical – are statistically run against historical data and normal parameters.

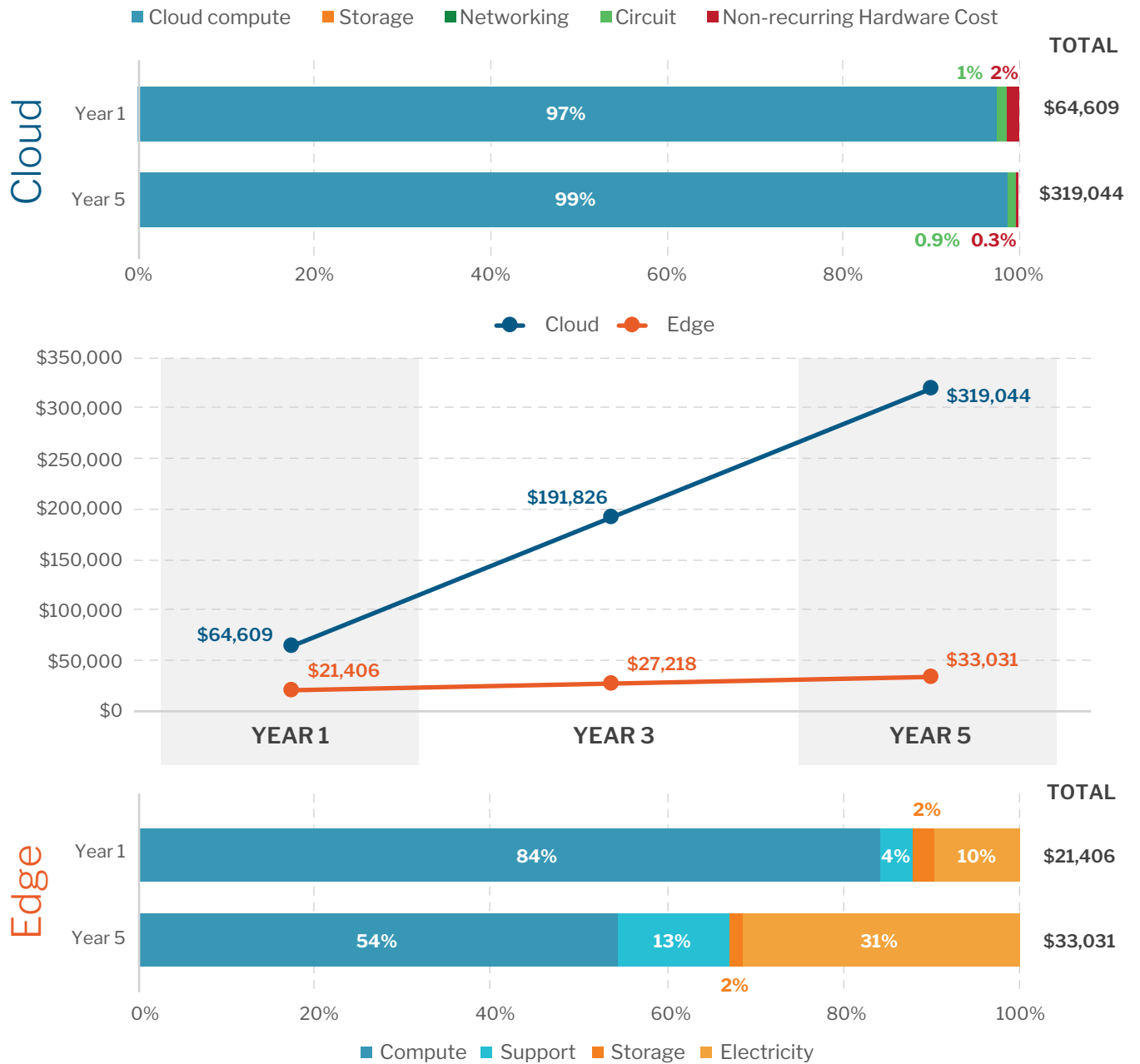
One differentiator in semiconductor manufacturing is the ability to provide higher levels of efficiency and/or quality based on the manufacturing process 'path' utilized. This is an area where semiconductor manufacturers are able to offer different levels of service depending on customer needs and budget.

Key Takeaways

- This use case is a good example of one that we see frequently started on cloud services because it is fast and easy to do so, but we see from the analysis that over the five-year period, a cloud-only model would be more than 9.5x more expensive. The frequency of updates and keepalive messages, as noted above, were a key driver for the cloud computing price component of the cloud model.
- The price of the edge computing appliance was the primary cost driver for the edge computing scenario, followed by the cost of electricity for the appliance. This is a robust edge computing device that can be leveraged, via virtualization, to run additional workloads without incremental investment.
- The cloud deployment scenario was heavily influenced by the high quantity of endpoints, as well as the frequency and size of messages from the manufacturing equipment, causing the price of cloud compute resources to far outweigh any other factor over a five-year period.

Figure 3: Year 1 and year 5 cost breakdowns for manufacturing quality/yield monitoring

Source: 451 Research custom research



BLACK & WHITE | DESIGNING FOR EDGE AND IOT SUCCESS

Discrete Manufacturing of Filter Components – Production Monitoring

The firm is a manufacturer of filter components used in other industrial processes. Production monitoring data is primarily gathered from the PLCs in the plant. A proprietary architecture supports the production process. Data is monitored at a high rate and kept for a 30-day window then averaged down into a 60-day window for overall trends. The majority of machinery is pressing and folding. Plant distribution covers 51 countries.

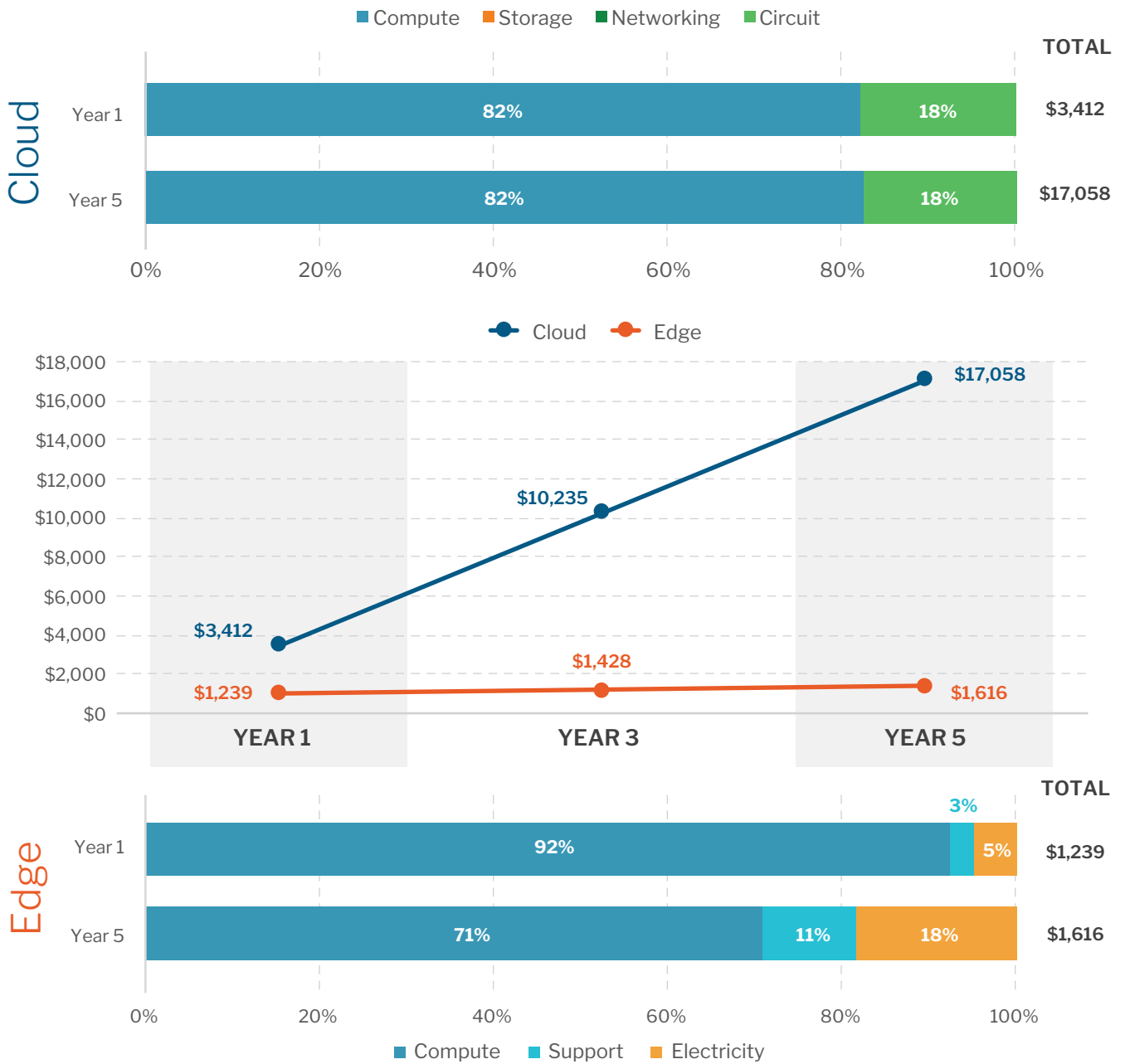
Data retention is critical for analysis of historic data trends (manufacturing historians), as well as training of models and, in some countries and manufacturing sectors, regulatory reasons. The volume of data retained can impact both on-premises and cloud-based storage costs. Equipment manufacturers may require access to machine data to ensure that the equipment is being properly operated within warranty parameters and may seek secondary monetization of the data for additional services.

Key Takeaways

- Both cloud and edge computing scenarios were inexpensive for this low-traffic use case; however, an emphasis on edge computing proved to be less expensive over a five-year period, given the frequency of updates to the cloud and its follow-on impact to overall cloud compute pricing.
- Lowering the frequency of updates to the cloud could bring cloud pricing more in line with the edge scenario; however, wide-area circuit costs to the cloud, unless the cost is absorbed as part of a broader enterprise cloud initiative, are still a recurring fee that, alone, exceed the cost of the edge hardware, support and electricity.

Figure 4: Year 1 and year 5 cost breakdowns for manufacturing production monitoring

Source: 451 Research custom research



BLACK & WHITE | DESIGNING FOR EDGE AND IOT SUCCESS

Oil and Gas Case Studies

The multi-trillion-dollar O&G industry is integral to not only the global transportation sector, but also to chemical feedstocks and other industries. The fluctuations in oil prices of previous years have given way to more sustainable price and production agreements between non-OPEC and OPEC countries. Although well below historical high prices, this stability has led to O&G operators making strategic investments in solutions such as IoT.

The main segments of the O&G market are upstream (exploration and production), midstream (transportation and storage) and downstream (refining and processing). For this section, two upstream applications were characterized, one in seismic exploration in remote sites, and one in wellhead production monitoring. In both cases, the remote environments complicated WAN connectivity and raised the price, both of which tilted the scales in favor of local edge computing.

Seismic Exploration

The exploration for oil and gas reserves has been evolving since the watershed modern discovery of oil in Pennsylvania in 1859. Beginning in the 1930s, exploration was performed with cabled geophones, or distributed microphones that were connected via copper wires to a central recording device to detect the signal injected by a seismic source, similar to 'pinging' a radar signal at an object from a submarine in the ocean. This cabled/wired approach faced challenges in the uncertain terrains encountered in oil and gas exploration.

Modern surveying and exploration approaches utilize a nodal system with each independent node integrating a geophone and recording device, as well as localized storage. These devices are distributed at consistent intervals throughout an exploration site and record at half millisecond or millisecond intervals. Today, there are a limited number of implementations connecting these nodes wirelessly to a centralized location for fusion of the geophone readings of each node.

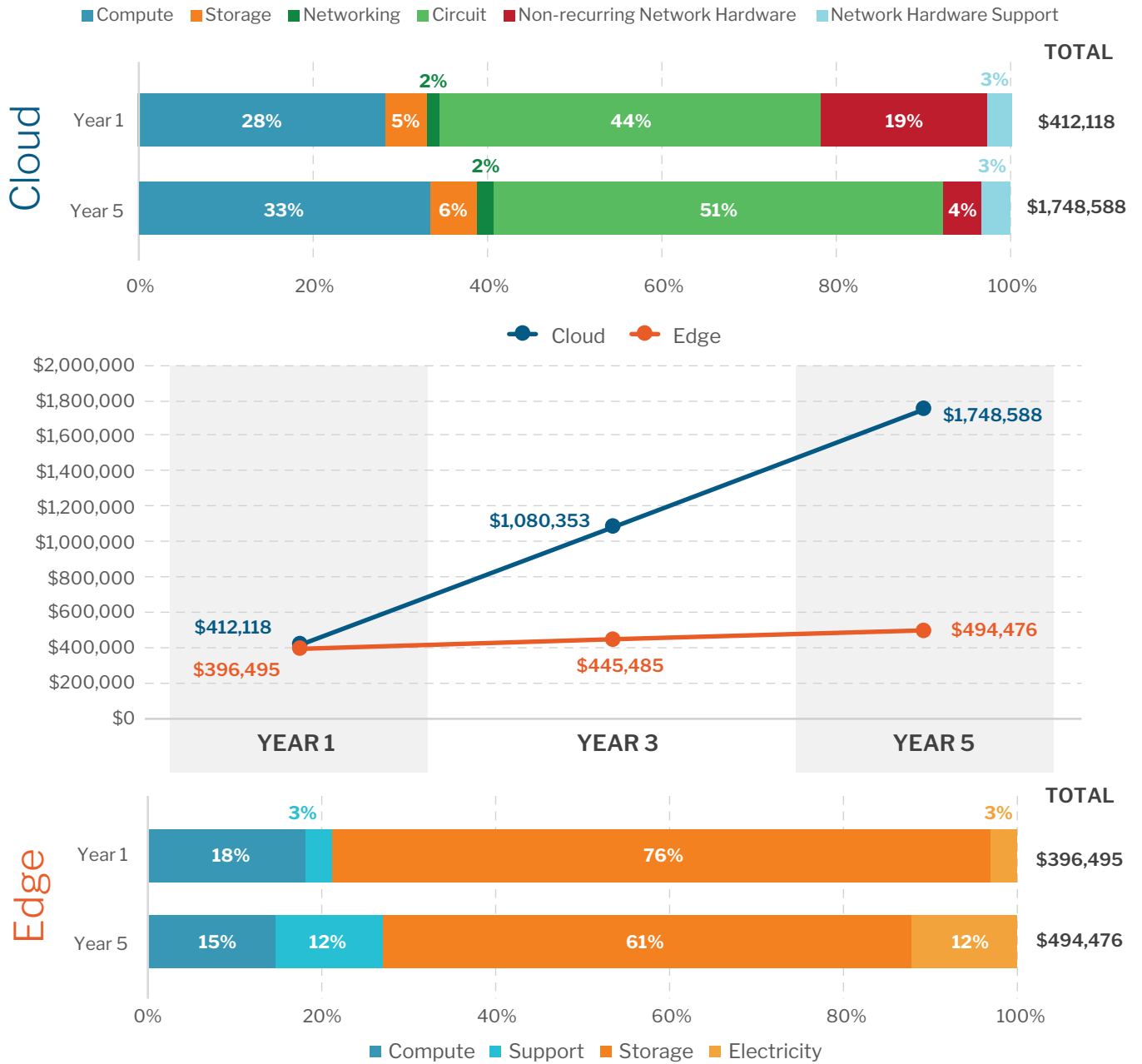
Currently, and in this use case, hard disk drives are integrated into each node to record readings throughout the signal injection process. These hard drives are then manually retrieved, and the data from these devices is integrated into a single model from the distributed recordings/readings. This model presents two primary challenges: battery life of the individual nodes, and the potential for errors in configuration or deployment of the nodes themselves, that are not discovered until the data is 'harvested' days later. In addition, manual retrieval of these hard drives is a labor- and time-intensive process as the data is extracted and integrated, and the scale of the nodal deployment can reach numbers as high as 30,000 nodes per site.

Key Takeaways

- The large amount of data generated in this use case drove higher prices in both cloud and edge scenarios, but for different reasons. The size of the samples require a petabyte of storage, either on-premises or in the cloud. In the edge scenario, the price of the storage devices (and support) proved substantial, and was the key lever of pricing, constituting 61% of the total five-year deployment.
- The cloud scenario pricing was equally impacted by storage, albeit indirectly in the cost of the robust WAN circuits to transport the sample data from the remote location to the cloud. The price of the circuit over time proved to be the largest driver at 52% of the five-year cumulative price.

Figure 5: Year 1 and year 5 cost breakdowns for seismic exploration

Source: 451 Research custom research



BLACK & WHITE | DESIGNING FOR EDGE AND IOT SUCCESS

Wellhead Production Monitoring

Once a promising site has been identified by seismic exploration, the next step is to continually refine the efficacy of the drilling by performing wellhead analytics at each drilling rig. Leveraging data from previous wells drilled in a field or site can yield valuable insights about where to drill another well and what geological formations are likely to be encountered during the drilling process before reaching oil. These wells are instrumented with a number of sensors and increasingly robust analytics to inform the best depth and direction of drilling to avoid a dry hole.

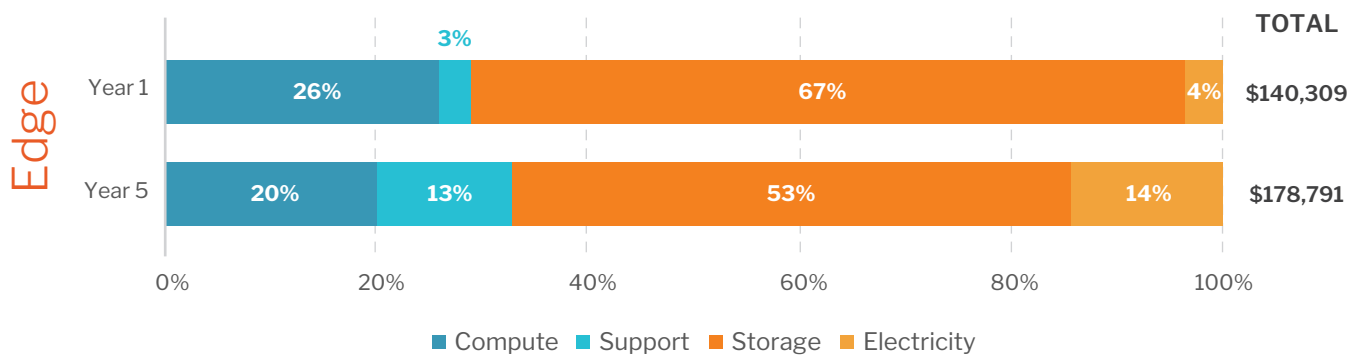
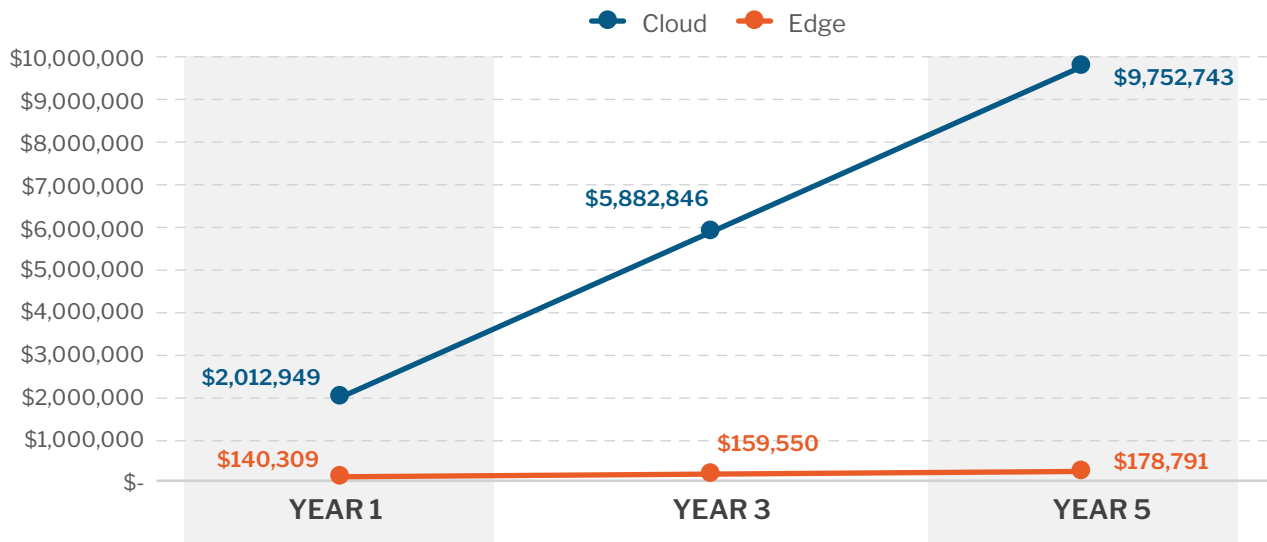
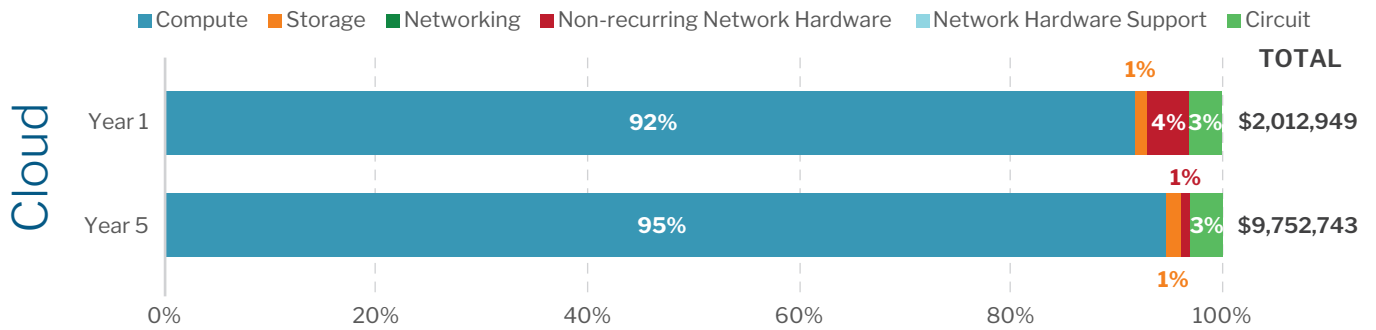
Located in South Texas, one of the more complex drill sites has 800 wells with 70,000 data sensors generating data. The site is constantly operating, with each well returning a constant stream of telemetry data, which is then used to inform models for other wells operating in the immediate vicinity. Data from the aggregate wells is stored, and new settings are determined based on a three-well lookback. This allows the wellhead operator to adjust location, direction and depth of the hole for optimal results in near-real-time.

Key Takeaways

- The sheer volume of data generated at each wellsite and the need for rapid adjustments based on real-time telemetry coming from each wellhead make oil and gas production a likely edge case from the start – if not entirely edge-driven, then at least doing pre-processing and summarization of production data prior to sending to the cloud.
- The frequency of updates from the wellheads and the requirement to constantly monitor pressure and adjust the direction of the drill results in a deluge of telemetry data from the wellsite. Although this requires considerable bandwidth to transport to the cloud, the annual cost of the custom fiber line was quickly surpassed by the price of cloud computing updates in the cumulative price calculations.

Figure 6: Year 1 and year 5 cost breakdowns for wellhead monitoring

Source: 451 Research custom research



BLACK & WHITE | DESIGNING FOR EDGE AND IOT SUCCESS

Smart City Case Studies

According to United Nations' estimates, 55% of the world's population, or 4.2 billion people, live in urban areas. That number will grow to 68%, or 6.7 billion by 2050. The challenges facing those cities are vast – maintaining city infrastructure, managing traffic and transportation, and ensuring public safety are just a few. Just as technology has helped automate and improve business and industrial operations, so too can information and communications technology help to improve the economic outlook and overall quality of city life. 'Smart cities' apply technology to address pressing urban issues, improve the quality of life for citizens and minimize the impact on the planet.

For the use cases in this section, we address two of the most pressing issues facing the public sector: environmental monitoring, which is driven by both environmental and public safety concerns, and public transportation management, which leverages IoT to aid both the city (capacity and route management) and the citizenry (real-time notifications and scheduling information).

Environmental Monitoring – Seismic Activity – Public Safety/Notification

A US state government (state population: three million) took on the task of tracking seismic activity throughout the state (and sharing it across its region) to understand the volume, degree and ultimately the cause of seismic incidents. To support the use case, seismic data is collected via ruggedized and buried environmental motion sensors (measuring x,y,z seismic movements) at 12 locations throughout the state (six are permanent; six are mobile and on the move). Given that a seismic incident can occur at any moment, with little to no warning, the system runs 24/7/365, collecting data every 30 seconds on average (and more frequently during an incident, as often as every microsecond) – an average of 2,880 collections per day, or 86,400 collections per month. Each endpoint sensor location collects and transmits an average of 1TB of data per month per device.

Depending on the location, the data is transported via wired (permanent) or cellular 3G/4G (mobile) connection to a state-owned enterprise datacenter location for initial analysis if needed (i.e., in case of seismic emergency) and archival raw data storage. The data is also forwarded to a second enterprise datacenter at a regional university that is better equipped with both software and staff for primary (and deeper) analysis. At the regional university location, the data is processed and 'noise' stripped out. The degree of earth movement is analyzed and correlated with data and movement analysis at other sensor locations across the region to determine the epicenter of movement; 100% of data is analyzed for event detection, with 20-30 events per day registering as at least somewhat significant.

Data at the regional university datacenter is deleted after three months. The state, meanwhile, retains data for 10 years in its enterprise datacenter. Data is retained for research and legal purposes (i.e., if there is ever an issue that could be resolved with the help of the data, a fracking accident, for example). Although the state has discussed using public cloud for archival data storage, it has opted to keep the data stored in its own datacenter for now, based on both cost and availability reasons.

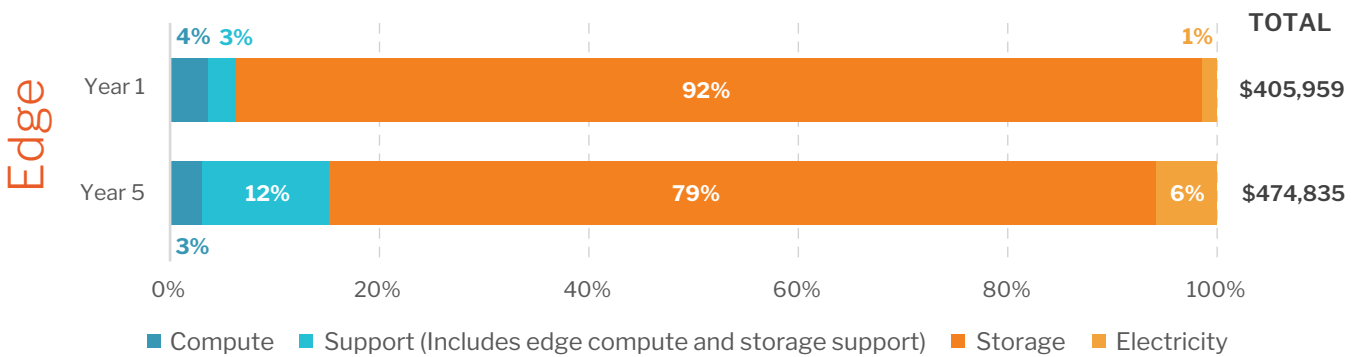
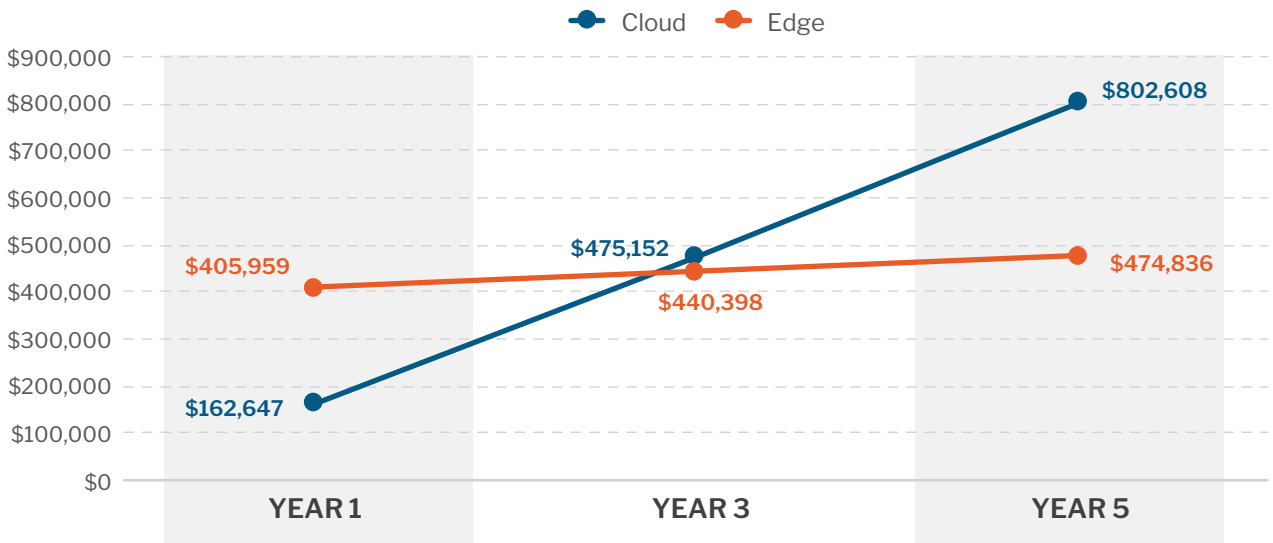
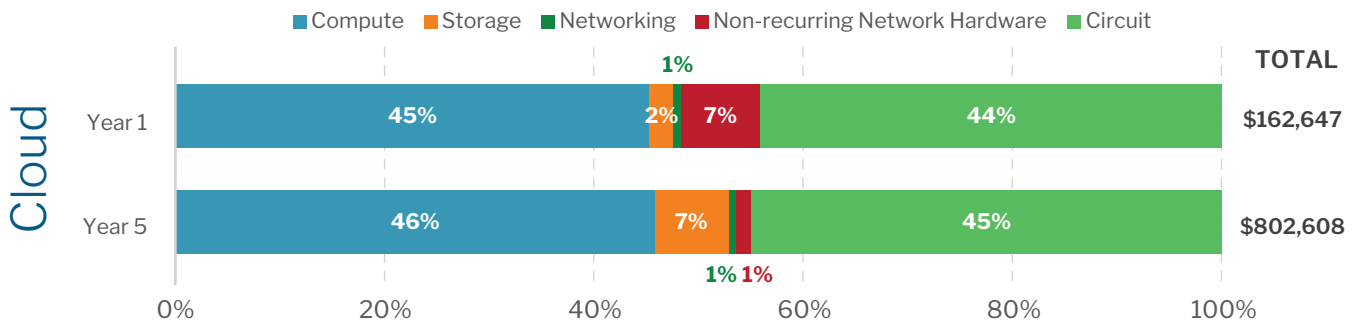
The state's IoT/analytics application has yielded positive results. The wider data collection and near- and long-term analysis has made it easier to distinguish between true seismic events and other activities, such as O&G fracking operations. And while there hasn't been an emergency-level seismic event since the project was launched, scientists – including students at universities across the wider region – have been able to study the historical data to better understand seismic triggers and impacts, which will help them anticipate and manage events.

Key Takeaways

- Cloud and edge computing scenarios in this use case were similar in five-year cumulative cost, with cloud computing only becoming more expensive than edge in the fifth year of deployment, driven by ongoing cloud computing and wide-area circuit costs.
- The requirement for data retention of seismic data drove increased storage requirements in the edge scenario, making it the key price lever (79%) over five years. In the cloud scenario, this is reflected in the ongoing circuit costs (45% over five years) combined with the processing of the seismic data (cloud compute, 46%), however with minimal cloud storage fees (7%) as an overall percentage of five-year cumulative price.

Figure 7: Year 1 and year 5 cost breakdowns for environmental monitoring

Source: 451 Research custom research



BLACK & WHITE | DESIGNING FOR EDGE AND IOT SUCCESS

Intelligent Public Transportation Management – Capacity Management/Citizen Route Notifications

A midsized US county (population: 500,000) deployed a series of interlinked IoT applications and use cases to help it better manage its regional public transport bus system. Computer vision using video camera endpoint-enabled deployment helps the county government manage its public transportation bus system. Capabilities include management and optimization of bus route capacity/utilization and dispatch; delivery of real-time location data and schedule information to citizens via smartphone app; and on-bus incident-based video surveillance. Primary measurements include the number of people on/off buses (counting people via facial recognition); real-time bus location vs. schedule data (via real-time cellular location data and enterprise application); and real-time surveillance video feed access and on-demand access to limited-time video archive as needed.

The on-bus setup includes 10 video cameras per bus, each with its own storage on each vehicle; one CPU per bus for compute/analytics; and one cellular router per bus to transport people-counting and location data). Beyond the on-bus endpoints, the system has deployed digital signage at bus stops (at 100 out of 5,000 total stops in the system). Video is collected in real time on each bus while it is in service and is stored and analyzed at the edge – i.e., on the bus. People-counting metadata is sent to a near-edge compute running in a county datacenter. Location data is also collected every six seconds for each bus in service and sent to compute resources running in the city datacenter and smartphone/scheduling applications running on AWS.

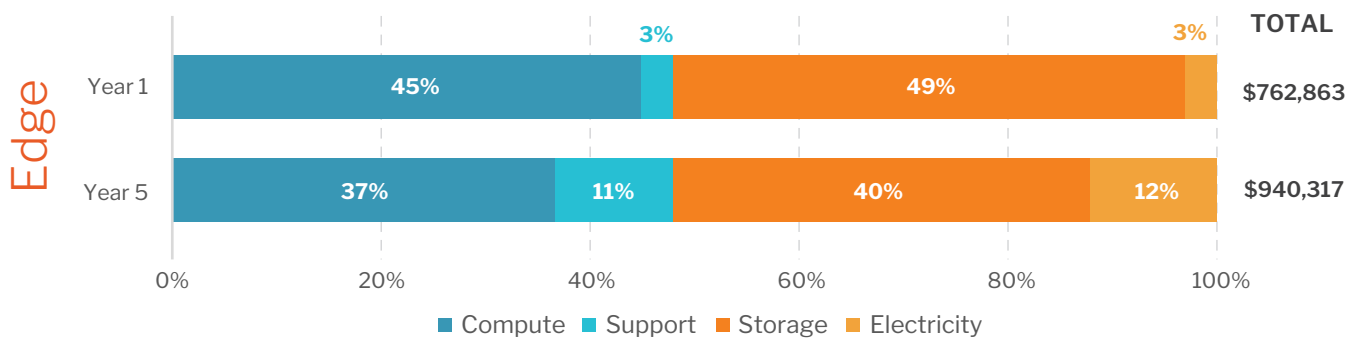
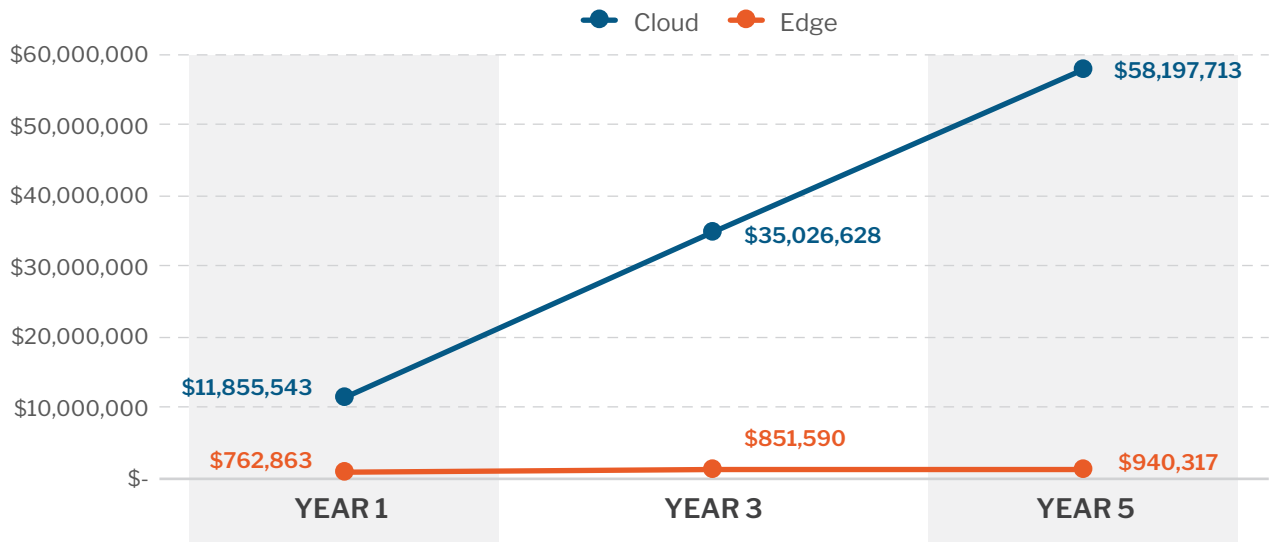
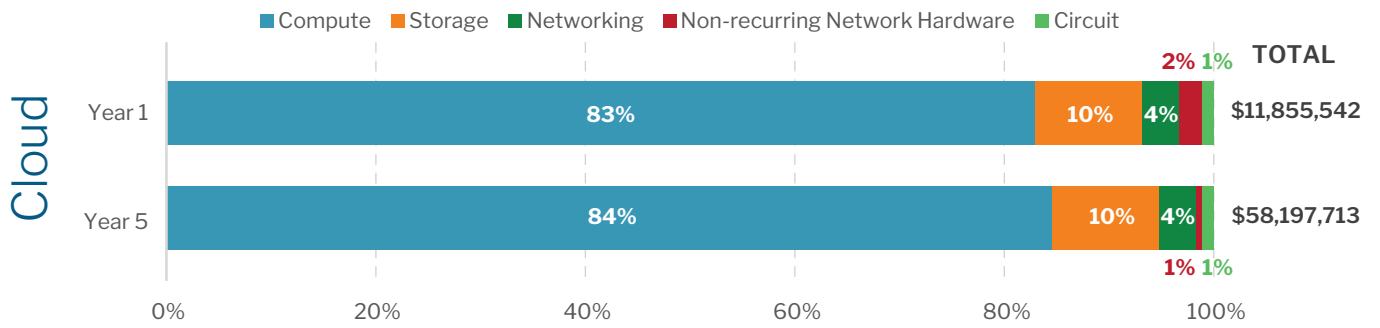
The county's intelligent transportation application has yielded positive results. Access to the locally stored data has helped local law enforcement improve public safety across the bus system, which averages 110,000 riders per day. Beyond rider safety, county officials cite a range of benefits, including improved transit operations (such as system capacity management and information sharing with other city departments), reduced fuel consumption, and improved rider satisfaction due to fewer delays.

Key Takeaways

- Video analytics is both traffic- and compute-intensive, and distributed video analytics across moving vehicles is one of the most rigorous use cases for IoT. The compute and retention (storage) of the video proved to be key cost levers over a five-year period in the edge scenario, with 77% of the total price being absorbed by these two categories.
- Video analytics in the cloud is equally challenging, and complicated by the difficulty in getting streaming video to a cloud service for analysis. Cloud compute and storage constituted 94% of the five-year cumulative price of cloud in this use case, driven in part by requirements for data retention.

Figure 8: Year 1 and year 5 cost breakdowns for intelligent public transportation management

Source: 451 Research custom research



BLACK & WHITE | DESIGNING FOR EDGE AND IOT SUCCESS

Conclusions

While there is no 'one size fits all' solution or design for an IoT deployment, the use cases presented in this paper heavily favor edge computing over a five-year period from a cumulative cost basis. Other use cases that are less data-intensive or cost-prohibitive for storage or bandwidth could favor deployment of edge computing. Ultimately, network designs will have edge compute and storage paired with cloud-based AI models and integration with cloud-hosted applications such as manufacturing historians or CRM. It will be the combination of these two, the edge and the cloud, leveraging each when best suited to a particular use case or environment, that will prove the IoT topology of the future.

In the future, topologies will atomize into IoT analytics and control functions delivered as microservices, leveraging cloud-native technologies such as containers. These networks will decouple the applications from the infrastructure components, abstracting them via open APIs, and enabling them to be located at the best location for optimal latency, cost and manageability. This abstraction layer itself functions as a platform for IoT applications and enables flexibility of deployment both to the cloud and the myriad edge locations across industries.



Now a Part of

S&P Global Market Intelligence

About 451 Research

451 Research is a leading information technology research and advisory company focusing on technology innovation and market disruption. More than 100 analysts and consultants provide essential insight to more than 1,000 client organizations globally through a combination of syndicated research and data, advisory and go-to-market services, and live events. Founded in 2000, 451 Research is a part of S&P Global Market Intelligence.

© 2020 S&P Global Market Intelligence. All Rights Reserved. Reproduction and distribution of this publication, in whole or in part, in any form without prior written permission from S&P Global Market Intelligence is forbidden. The terms of use regarding distribution, both internally and externally, shall be governed by the terms laid out in your Service Agreement with 451 Research and/or its Affiliates. The information contained herein has been obtained from sources believed to be reliable. 451 Research and S&P Global Market Intelligence disclaim all warranties as to the accuracy, completeness or adequacy of such information. Although 451 Research may discuss legal issues related to the information technology business, 451 Research does not provide legal advice or services and their research should not be construed or used as such.

The content of this artifact is for educational purposes only. S&P Global Market Intelligence does not endorse any companies, technologies, products, services, or solutions. S&P Global Market Intelligence shall have no liability for errors, omissions or inadequacies in the information contained herein or for interpretations thereof. The reader assumes sole re-sponsibility for the selection of these materials to achieve its intended results. The opinions expressed herein are subject to change without notice.



NEW YORK

55 Water Street
New York, NY 10041
+1 212 505 3030



SAN FRANCISCO

One California Street, 31st Floor
San Francisco, CA 94111
+1 212 505 3030



LONDON

20 Canada Square
Canary Wharf
London E14 5LH, UK
+44 (0) 203 929 5700



BOSTON

75-101 Federal Street
Boston, MA 02110
+1 617 598 7200

