WHITE PAPER

The Missing Link in the Internet of Things Value Chain: Data Management

Pamela Kerman, Jim Lyons, and Ron Long, NetApp

July 2016 | WP-7236-0716

NetApp®

Section 1

Introduction

The Internet of Things (IoT), which extends compute and network capabilities to objects, sensors, and everyday items not normally considered computers, is reshaping our world. Today, the ever-expanding universe of connected devices includes everything from pipelines to electric meters, blood pressure gauges, automotive robots, cars, cities, weighing scales, cows, forks, and more. Every day the list of networked objects grows—along with the volume of data they generate.

According to Cisco, IoT was "born" between 2008 and 2009, when the number of "things or objects" connected to the Internet exceeded the number of people for the first time in history.[1] Cisco estimates that the IoT will consist of 50 billion devices connected to the Internet by 2020.[2]

However, at most companies, data from a variety of social, analytical, and mobile platforms already feeds into a variety of on-premises and cloud systems, resulting in siloes of data. The cumbersome process of collecting, managing, and analyzing data can quickly overwhelm the small IT teams that are tasked with making sense of it all. An inability to provide predictive analytics and predictive maintenance creates a lack of flexibility and business agility. This not only hinders future growth, but also creates inefficiencies, with more staff, more equipment, and more resources needed to manage disparate datasets across disparate environments.

How can you make sense of the tsunami of data generated by IoT devices? How should you architect your data management systems to convert data into value? The scale and scope of new IoT data call for new thinking and approaches. They call for advanced data management and the integration of IoT collection systems with traditional systems of record as part of a hybrid cloud infrastructure.

This paper focuses on how NetApp solves the challenges of managing IoT data across hybrid cloud environments with the Data Fabric—our vision for the future of data management and a solution to the unique challenges posed by IoT data. Data management is a critical link in the IoT value chain across all industries. In this paper, we examine examples and use cases specific to healthcare, manufacturing, and connected cities.

1. "The Internet of Things: How the Next Evolution of the Internet Is Changing Everything." April 2011. Cisco Internet Business Solutions Group (IBSG): www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf
2. "Internet of Things." Cisco website, June 2016. www.cisco.com/c/en/us/solutions/internet-of-things/overview.html

Section 2

Business Drivers and Challenges

Technology is reshaping our world, and connected devices are driving a new generation of data analytics applications. This connected, data-powered digital era—what IDC has called the "third platform" of IT—began circa 2010. The first platform was the mainframe computer system that started in the 1950s; the second platform was the client/server system that began in the 1980s.

In the digital era, enterprises must be faster, global, mobile, social, digital, and available on demand all the time. Enterprises must generate greater customer insight and engagement, create new business models, and improve productivity for their increasingly mobile workforces. They must move with speed and scale to capture new revenue opportunities while balancing cost and risk.

- **Data volume, velocity, and variety:** With IoT, data feeds into social, analytical, and mobile platforms across a variety of hybrid cloud environments. As sensors across these and other platforms transmit data, the volume becomes unmanageable and the different types of platforms create inconsistencies. Turning the volumes, velocity, and variety of data into real business value becomes a daunting task.
- **At the edge:** Large amounts of data need to be ingested, filtered, and classified at the network edge. At the same time, the data needs to be analyzed in real time for immediate action. Typically, the data needs to be aggregated and moved to the network core for analytics. Data needs to be protected at the edge and in flight when it is moved to the core.
- **At the core:** IoT produces large amounts of unstructured data with different data types. To anayze this data, IT needs a big-data infrastructure, including tools for analytics, data wrangling, and data management. The data needs to be made available to on-premises applications and cloud applications, depending on the IoT platform architecture and end-user applications.

3. "3rd Platform." IDC: www.idc.com/prodserv/3rd-platform/

NetApp®

In order to create value out of the raw data, CIOs need to process massive amounts of data being driven from a variety of sensors across connected devices and create actionable real-time analytics from large volumes of data in disparate locations versus simply collecting the data. They also must combine and integrate data into existing systems in innovative ways that can help reduce costs and improve visibility of market opportunities. Finally, they must improve productivity for mobile workers and drive new revenue streams by enhancing existing new products and developing additional services.

Section 3

Unlocking Value from Connected Devices with Data Management

NetApp believes that the hybrid cloud will become the dominant IT model across all industries. The hybrid cloud is a cloud computing environment that uses a mix of on-premises, private cloud, and third-party public cloud services with orchestration between all of the platforms. The hybrid cloud will also become the preferred model for IoT, enabling data and workloads to move across platforms to achievethe best results. NetApp solves the challenges of managing data across hybrid cloud environments with the Data Fabric.

The Data Fabric is a broad vision for enabling data to be in the right place at the right time with the right performance and cost. It is NetApp's vision for the future of data management, and data management is the key to unlocking value from connected devices.

At a high level, the Data Fabric helps you to do three things:

- Manage and secure data from connected devices across flash, disk, and hybrid clouds.
- Process large volumes of data from a variety of IoT sources with the high levels of visibility and performance you need to respond quickly.
- Choose from a global ecosystem of NetApp partners who can help you build a compliant IoT platform that connects and automates resources in the data center, near the cloud, and in the cloud.

For IoT, NetApp takes a layered approach—that is, NetApp uses a set of integrated functions and technologies to meet the challenges of IoT. There are four fundamental layers:

- **Partnerships and Solutions.** NetApp partnerships and solutions are focused on providing customers with complete, fully integrated approaches that solve common problems or needs.
- **Common APIs.** Common APIs are essential to make the layered approach work across a broad base of IT infrastructure, products, and platforms.
- **Data Fabric.** A Data Fabric enabled by NetApp provides a single, standardized data management and storage solution that can work across different architectures and platforms. The solution provides a common method for making data available without putting it in silos or limiting flexibility. The Data Fabric acts as the link between the data platform, data lakes, APIs, and key software products that will be used at a higher level to create IoT applications. The Data Fabric provides a consistent interface that makes it easy to provide IoT data to specific workloads or applications.
- **Data Platforms/Data Lakes.** The data platform layer is designed to provide an enhanced approach to platform as a service for IoT. The platform is made up of autonomous components, including capabilities such as compute, object store, networking, orchestration, and other functions. The key to IoT support is that these components are designed to support massive scale-out to provide necessary throughput. The term “data lake” refers to a set of global “object pools” that provide broad access to large amounts of data. Data lakes use a flat architecture to store vast amounts of raw data, precisely the type of requirement that will be common in IoT. The use of Hadoop with data lakes is already increasing, but data lakes are not specific or limited to Hadoop.

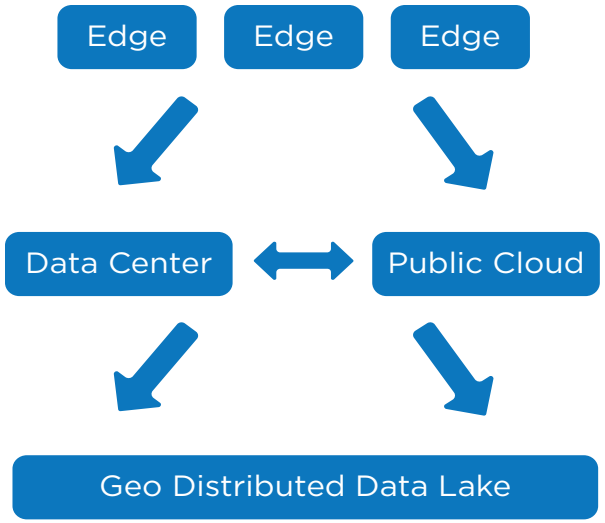


Figure 1) IoT data workflow.

Each of these four elements provides a critical part of a holistic solution that empowers the organization to use IoT data. These elements also provide IT with strategic infrastructure that isn't prohibitively expensive or too limited to support huge incoming data flows. Data is collected at the edge and analyzed in real time. Then it is aggregated and sent either to the data center or to the public cloud for analytics. Finally, data is moved back and forth between the data center and the public cloud for analytics and archival purposes.

Section 4

Industry Use Cases

What follows is a look at how data flows to create both personal and business impact across key use cases, including healthcare manufacturing, and "smart cities" (aka connected cities).

NetApp®

IoT Use Case #1: Healthcare

According to a report from MarketResearch.com, the healthcare IoT market segment is poised to hit $117 billion by 2020.[4] This growth in IoT is accelerating healthcare data management requirements that were already growing astronomically. From a data management perspective, one organization recently found that a single patient generated 3.67 terabytes of data. Consider what this means in the context of millions of healthcare consumers and the billions of everyday objects projected to be able to capture, receive, and share data by 2020.

Healthcare currently encompasses a range of IoT technologies, including:

- Traditional medical devices, including chemotherapy dispensers and insulin pumps
- Wireless devices, including scales, glucometers, and blood pressure monitors
- Consumer wearables such as the Apple Watch, Fitbit, and Nike FuelBand

The new generation of IoT sensors and patient devices means that healthcare providers can collect and analyze data from hospital devices, such as bedside sensors, and from patient devices outside the hospital. Combining these data sources can provide clinical specialists with a single platform for managing patient data.

IoT also provides real-time information to caregivers through the cloud, delivering anytime, anywhere patient status between in-person care visits and at the point of care. Clinicians can collect, record, and analyze new data streams faster and more accurately using data to enable better-informed decisions.

4. "$117 Billion Market for Internet of Things in Healthcare by 2020." April 2015. Forbes: www.forbes.com/sites/tjmccue/2015/04/22/117-billion-market-for-internet-of-things-in-healthcare-by-2020/#5124e1642471

NetApp®

With IoT, healthcare providers and patients can partner to collect and monitor health data, focusing on prevention and management to improve outcomes and save money. Next-generation predictive analytics will further support a focus on wellness versus illness and give healthcare practitioners real and ready tools for effective population health management.

IoT Data Management for Healthcare

With the Data Fabric, healthcare providers will be able to process large volumes of data from diverse IoT sources with visibility, performance, and efficiency. A hospital will be able to store data in lower-cost cloud or archival storage, then seamlessly move data to high-performance flash storage to enable predictive analytics to run on the data. NetApp's partners in the healthcare space include Epic, Cisco, and Splunk.

5. "HDB elderly alert system well received in test-bed." March 2015. The Straits Times, Singapore: www.straitstimes.com/singapore/housing/hdb-elderly-alert-system-well-received-in-test-bed

Customer Example

Singapore's Housing Development Board is piloting a smart monitoring alert system for elderly citizens living in apartments. The system uses motion sensors that track the citizens' living habits and alerts their caregivers through alarms and text messages in the event of irregular patterns—such as an unusually long period of inactivity. The system also comes with a portable panic button that the elderly person can press in times of distress. Most of those who used the system reported finding it easy to use and did not feel it invaded their privacy. The alert system might also become part of the smart-home infrastructure in other public housing. More information is available here: http://www.straitstimes.com/singapore/housing/hdb-elderly-alert-system-well-received-in-test-bed.

NetApp®

IoT Use Case #2: Manufacturing

IoT is changing the face of manufacturing. Factories and plants that are connected to the Internet are more productive, efficient, and smarter than their unconnected counterparts. And the stakes are massive; according to a 2015 report by Accenture,[6] the industrial Internet could contribute over $14 trillion of global output by 2030—with obvious appeal to manufacturers and buyers alike.

Driving this transformation are global competitive pressures that are challenging industrial and manufacturing companies to boost efficiencies in their systems, manage workforce skills gaps, and uncover new business opportunities. With a connected plant, manufacturers can use information to automate workflows to maintain and optimize production systems without human intervention. Company software that keeps a record of how different equipment performs can automatically adjust the machinery if it detects that a measurement—such as fan speed, temperature, or humidity—is outside the acceptable range. A combination of sensors, data, and analytics enables services such as predictive maintenance—a more proactive approach to maintenance than scheduled maintenance and the reactive "break-fix" mode of dealing with problems.

With connected cars, integrated telemetry—such as sensors that detect failures and call for help—collect and analyze valuable data on specific product performance for aftermarket services and proactive maintenance. The mobile nature of transportation systems has made traditional IT solutions and infrastructure difficult to use, while small IoT devices with connectivity are an excellent fit for transportation.

6. "Accenture estimates the Industrial Internet of Things (IIoT) could add $14.2 trillion to the global economy by 2030." January 2015. Accenture: www.accenture.com/us-en/insight-industrial-internet-of-things?c=glb_wef2015pr_10000001&n=otc_0115

NetApp®

IoT Data Management for Manufacturing

For manufacturing firms, creating and securing data across their entire data infrastructure is the key to unlocking value from their infrastructure and machinery—across the edge, in the data center, and in the public cloud.

For many manufacturers, the concept of predictive maintenance is a good use case starting point. Traditionally, scheduled maintenance for production machinery has been based on rotational intervals, triggered by parameters of time, usage, and so on. But although this is a step up from the break-fix approach, it is essentially a one-size-fits-all model that doesn't take into account how the machinery is actually operated—which is often different from what was expected or planned for. Despite strict maintenance frameworks and schedules, the result is often malfunctions, downtime, and production outages. With IoT, sensor data identifies not only a machine's actual usage and environment, but also predicts the respective root cause of potential problems.

NetApp and its ecosystem of partners in this space, including SAP, Cisco, MapR, Hadoop, Hortonworks, Siemens, Bosch, and Rockwell Automation, offer central infrastructure and application management.

Customer Example

Consider an enterprise in the automotive sector that produces thousands of cylinder heads in a 7 x 24 production process. For each cylinder head, a dataset is created comprising several hundred process parameters (measurements, times, temperatures, tools used, on so on) that are then crosschecked with target values. Doing this work manually would take days, even for the most experienced specialists. With predictive analytics, all relevant measurement data within the production flow is continuously collected, assigned time and content, and prepared for statistical analytics. Insights about deviations from the standard are now available within minutes, enabling corrective action to be taken.

Besides the production systems with sensors, actuators, software, and the secure transport of data, a high performance “data lake” incorporates the multiple formats of machine and environmental data. This process provides the right preconditions for a high-performing and flexible data analytics engine. As a result, manufacturing efficiency improved by 25% and production ramp-up time decreased by as much as 50%.

NetApp®

IoT Use Case #3: Smart Cities

From Rio de Janeiro to Barcelona, from Singapore to Tokyo, city administrators use technology to improve the quality of life for their citizens in collaboration with academia, business, and the technology industry. In 2025, there will be an estimated 26 global smart cities. Smart cities represent a $1.5 trillion market opportunity.[7]

Smart cities have multiple public and private services that need to be coordinated:

- **City administration—**to streamline management and deliver new services in an efficient way with smart meters and smart connection
- **Education—**to increase access, improve quality, and reduce costs
- Public safety—to use real-time information to anticipate and respond rapidly to emergencies and threats
- **Transportation—**to reduce traffic congestion while encouraging the use of public transportation by improving the customer experience and making travel more efficient, secure, and safe
- **Utilities—**to manage outages, control costs, and deliver only as much energy or water as is required

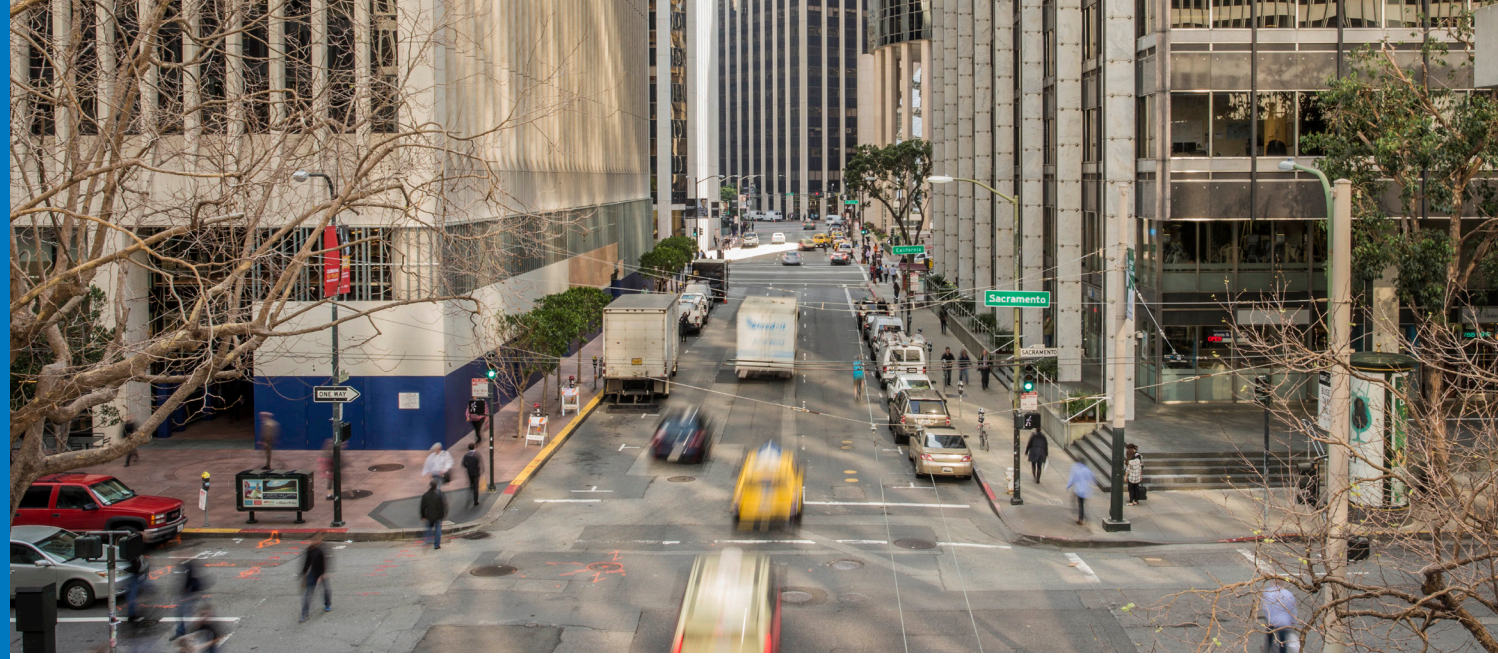
7. "Smart Cities—a $1.5 Trillion Market Opportunity." June 2014. Forbes: www.forbes.com/sites/sarwantsingh/2014/06/19/smart-cities-a-1-5-trillion-market-opportunity/#7e1cd26f7ef9

NetApp®

NetApp®

IoT Data Management for Smart Cities

The Data Fabric plays a key role in facilitating and enabling smart city initiatives. It gives organizations the freedom to move their data to where they need it most. The Data Fabric will transform how organizations manage, secure, protect, and move their data across disparate data environments, no matter where it lives. The Data Fabric enabled by NetApp helps organizations to seamlessly connect different data management environments into a cohesive, integrated whole. It allows organizations to manage their data outside the four walls of their business as a natural extension of what they do on the premises. NetApp has strong partnerships with leading technology companies worldwide, including Cisco, Samsung, and Siemens, and leading cloud service providers and hyperscalars such as Amazon Web Services. NetApp supports these partners with data management solutions in response to the requirements of the various smart city initiatives.



Partner Example

In China, NetApp partner Zhoushan Public Transportation Co., Ltd. (ZSPT), built a smart Bus Rapid Transit (BRT) system based on converged infrastructure. The system allows more people to join a "Green Travel Plan" to reduce carbon emissions. ZSPT is a state-owned enterprise mainly operating ground passenger transportation in Zhoushan City in Zhejiang province. The enterprise built a virtualized infrastructure with a multisite converged infrastructure solution in two data centers about 500 yards apart. The multisite solution is fully redundant and capable of multisite deployment, enabling both disaster avoidance and high availability for BRT services. The system allows ZSPT to monitor BRT traffic in real time, enabling the control center to provide new dispatching rules when there is unusual passenger flow volume. With the help of its always-on intelligent scheduling and management system, BRT delivers rapid, on time, convenient transportation services to local citizens, more and more of whom are choosing BRT as their preferred travel method.

NetApp®

Section 5

The Future of Data Management

"The future is already here—it's just not evenly distributed."
—William Gibson

NetApp®

It was writer William Gibson who famously declared, "The future is already here—it's just not evenly distributed." The same could be said of IoT. And with NetApp's vision of a Data Fabric, we might also say that the future of data management is already here. The Data Fabric gives you freedom—the choice to always make the best decisions for you and your business. It gives you data mobility—so that your data flows seamlessly to wherever you need it most, across flash, disk, and the cloud. And it gives you speed—because NetApp® technology, services, and partnerships allow you to innovate faster with fewer resources.

The Data Fabric lets you realize the full potential of the hybrid cloud, which is the IT foundation for IoT. In summary, data management across your entire data infrastructure is the key to unlocking value from connected devices. When data is free to flow where you need it most, you respond and innovate faster. That is the promise of the Data Fabric.

Next Steps

For more information, visit www.netapp.com/us/solutions/data-fabric/index.aspx.

NetApp®