Azure Arc Systems Management

Governance and Administration of Multi-cloud and Hybrid IT Estates

Ramona Maxwell

Apress®

Azure Arc Systems Management

Governance and Administration of Multi-cloud and Hybrid IT Estates

Ramona Maxwell

Apress®

Azure Arc Systems Management: Governance and Administration of Multi-cloud and Hybrid IT Estates

Ramona Maxwell
Redwood City, CA, USA

ISBN-13 (pbk): 978-1-4842-9479-6 ISBN-13 (electronic): 978-1-4842-9480-2
https://doi.org/10.1007/978-1-4842-9480-2

Copyright © 2024 by Ramona Maxwell

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

Trademarked names, logos, and images may appear in this book. Rather than use a trademark symbol with every occurrence of a trademarked name, logo, or image we use the names, logos, and images only in an editorial fashion and to the benefit of the trademark owner, with no intention of infringement of the trademark.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Managing Director, Apress Media LLC: Welmoed Spahr
Acquisitions Editor: Ryan Byrnes
Development Editor: Laura Berendson
Editorial Assistant: Gryffin Winkler

Cover image designed by Arnon Thaneepoon at Dreamstime.com

Distributed to the book trade worldwide by Springer Science+Business Media New York, 1 New York Plaza, 1 FDR Dr, New York, NY 10004. Phone 1-800-SPRINGER, fax (201) 348-4505, e-mail orders-ny@springer-sbm.com, or visit www.springeronline.com. Apress Media, LLC is a California LLC and the sole member (owner) is Springer Science + Business Media Finance Inc (SSBM Finance Inc). SSBM Finance Inc is a **Delaware** corporation.

For information on translations, please e-mail booktranslations@springernature.com; for reprint, paperback, or audio rights, please e-mail bookpermissions@springernature.com.

Apress titles may be purchased in bulk for academic, corporate, or promotional use. eBook versions and licenses are also available for most titles. For more information, reference our Print and eBook Bulk Sales web page at http://www.apress.com/bulk-sales.

Any source code or other supplementary material referenced by the author in this book is available to readers on GitHub (https://www.apress.com/gp/services/source-code).

Paper in this product is recyclable

In honor of my bold and spirited adventurer
Charity "Cha Cha" Cassady (1974–2016)

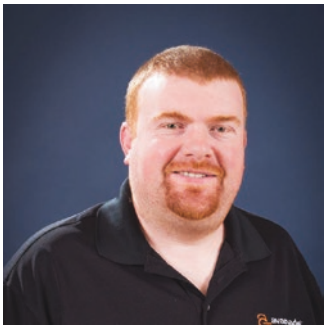
put your finger right in the middle of the cake
refuse to give up
refuse to let anyone you care about give up
pursue your dreams relentlessly

Table of Contents

About the Author

Ramona Maxwell is an experienced Enterprise Solution Architect who has worked with clients across a broad spectrum of the Fortune 100 such as utilities, finance, and healthcare organizations. She is certified by Microsoft as an Azure Solutions Architect Expert and is an AWS Certified Solutions Architect, Microsoft Certified Trainer (MCT), Associate as well as a 2020 graduate of VMware's Platform Acceleration Lab training for application architects. She has expertise in container technologies such as Kubernetes and Docker and extensive experience extending line-of-business systems across multiple technology stacks. Ramona is an occasional presenter at industry events on enterprise computing strategies and techniques. The author's full profile can be viewed at `https://ramonamaxwell.com`.

About the Technical Reviewer



Mike DeLuca is a recognized expert in public, private, and hybrid cloud. With seven patents, numerous cloud-related innovations, and a long career at Avanade and Microsoft, he has deep expertise helping the world's premier enterprises and online properties understand the journey to cloud. Mike currently serves as the leader for Avanade's suite of cloud-related tooling globally. He lives in the San Francisco Bay Area with his wife and four kids. He is also a contributor to the blog www.cryingcloud.com.

Introduction

The necessity of the methods and products covered in this book stands as a testament to the celerity with which enterprise IT estates have sprawled across multiple public clouds while never entirely leaving corporate-owned data centers. Exponential growth of data requires efficient access and processing methods, along with protection against corruption or exfiltration. Accomplishing these tasks manually is no longer possible; thus, tools and methodologies which are equal to these challenges are considered in this book. While Microsoft's Azure Arc is the primary focus, the principles of enterprise IT management discussed should serve to inform even those readers who are not currently in the market for an enterprise control plane. The book may be read topically if a particular area is of interest to you, with the caveat that earlier chapters are foundational to understanding more complex subjects that follow. Both business and IT leaders should find answers to their particular concerns around governing and securing technology assets in a fashion that promotes the objectives of the business.

CHAPTER 1

The Challenges of Enterprise-Scale Hybrid and Multi-cloud Architectures

The innovation of technology solutions available to enterprise-scale organizations continues to grow exponentially and at an astonishing pace. While adopting new technology is a vital strategy for profitability and growth, it also creates significant hurdles with transitions between product stacks, integration with existing solutions, and the critical need to observe and control the entire information technology [IT] estate of an organization. This book aims to investigate what the author believes to be the most comprehensive control plane for IT systems governance available on the market today, Azure Arc. It will cover the product's capabilities across management of database systems, Kubernetes, governance and policy, security, and its contributions to industry best practices such as DevOps. While focusing on Arc, some comparisons will be made to other products that have similar objectives. Finally, it will offer some historical context as to how the platforms impacted by Arc developed and what fostered the need for a product like Arc.

© Ramona Maxwell 2024
R. Maxwell, *Azure Arc Systems Management*,
https://doi.org/10.1007/978-1-4842-9480-2_1

In the world of enterprise architecture and the IT systems it creates, a use case or problem statement often generates a premise for the design. In the case of hybrid and multi-cloud system DevOps, governance, security, modernization, upgrades, monitoring, and management are all key areas of concern which we can examine before proposing any solution.

Over the last decade, the rush to the cloud has moved many enterprises to undertake large modernization projects. Motivating these behemoth efforts have been factors such as speed to market, adopting agile project methodologies in hopes of simplifying upgrades and expansion of technology stacks, shedding the large capital expense of running a data center, and more.

Commercial considerations may also influence which cloud an organization utilizes. The famous resistance of Walmart to use cloud services from Amazon, a chief competitor in the retail space, has led them to develop a hybrid solution combining their own infrastructure with services from Google Cloud Platform [GCP] and Microsoft Azure – but not Amazon Web Services [AWS].[1] Companies making acquisitions may find their new subsidiaries are already fully built out on a specific cloud platform, offering little justification to immediately rebuild on the parent company's provider platform. One of Azure Arc's key features, which will be covered extensively in this book, is its ability to manage IT assets on multiple clouds as well as in private data centers from one dashboard, thus enabling centralized governance and cost control.

Unfortunately, attempts to modernize and migrate IT workloads have a significant risk of failure. The UK consulting firm Advanced publishes an annual report surveying mainframe modernization initiatives in companies with more than 1B in revenue. In 2020, 74% of companies reported failed

[1] www.bloomberg.com/news/newsletters/2022-07-12/walmart-cloud-weans-itself-off-of-microsoft-azure-google-cloud

modernization projects,[2] and in 2021 that number had increased to 78%.[3] At the time of this writing their most recent update to the report was commissioned from Coleman Parkes early in 2022.[4] The summary noted progress as its subjects had gone beyond "embracing cloud hyperscalers as infrastructure providers" and begun "inviting them into the critical operations of their businesses" with multi-cloud identified as a strategy that would "predictably emerge as an organization's cloud maturity rises." Thirty-three percent of respondents to *Unisys'* detailed Cloud Success Barometer[5] report unsatisfactory outcomes in terms of organizational effectiveness following cloud migrations. As a result, vendors like Dell are seizing the opportunity to "repatriate" customers back to their own data centers,[6] an offering likely to entice the 54% of respondents to Advanced's latest survey that identified private cloud as part of their mainframe migration strategy.

Microsoft's answer to these throes of transition has not been not to abandon their cloud-first strategy and revert to selling server-based solutions. Instead, they looked at how the Azure cloud itself runs on Azure Resource Manager and began to extend this configuration-based approach to create a control plane capable of managing resources no matter where the customer hosted them including competitor's clouds, on-premise data centers, and edge locations. Over a period of experimenting with various control technologies (which will be detailed in the chapter covering Arc's history), Microsoft constructed and continues to expand upon Azure Arc.

---

[2] https://www.businesswire.com/news/home/20200528005186/en/74-Organizations-Fail-Complete-Legacy-System-Modernization

[3] https://modernsystems.oneadvanced.com/globalassets/modern-systems-assets/resources/reports/advanced_mainframe_report_2021.pdf

[4] https://modernsystems.oneadvanced.com/globalassets/modern-systems-assets/resources/reports/2022-mainframe-modernization-business-barometer-report.pdf

[5] https://www.unisys.com/siteassets/microsites/cloudbarometer/report_unisyscloudsuccessbarometer.pdf

[6] https://www.crn.com/news/data-center/michael-dell-it-s-prime-time-for-public-cloud-repatriation

The three primary cloud vendors (Google, Azure, and AWS) have each responded with offerings that allow customers to retain some critical workloads in-house while still benefiting from their platform's proprietary tools including deployment pipelines and administration consoles. Google Anthos offers an extension of Google Kubernetes Engine (GKE) to instances running on a competitor's clouds, edge locations, and bare metal running on the customer's hardware, while AWS Outposts use only AWS hardware that is shipped to the customer. While each of Google Anthos, AWS Outposts, and Azure Arc share Kubernetes management as a large part of their DNA, only Arc encapsulates the data center in its administration capabilities – a key differentiator which we will also devote a chapter to.

Microsoft does not even make a pretense of humility regarding Arc's capabilities and importance, with the page header in Arc's section of the Azure portal declaring, "*Govern and manage all your infrastructure, anywhere.*" Thus, while ordering an Outpost will ship your own AWS world with a subset of their services, and Google will let you run its industry-leading Kubernetes offering no matter where you park your servers, Microsoft's claim is that Arc can manage Kubernetes, SQL Servers, virtual machines, and more on any cloud or in your data center from a single control plane while additionally opening the doors to a plethora of Azure services to workloads not actually running on Azure.

The Challenges of Hybrid and Multi-cloud Architectures in a Modern Application Stack

What are some of the specific challenges posed in terms of implementing governance, DevOps, and security where systems are disparate and spread across large IT landscapes? How can existing workloads that are core to

business operations be modernized in order to be secure and competitive in their market? In the next few paragraphs, we will examine some of the barriers to success, while subsequent chapters will delve into how control plane technologies and specifically Azure Arc can offer steps to remediate them.

DevOps

Core principles of DevOps practice such as continuous integration, testing, deployment, and monitoring are made challenging by the very nature of distributed infrastructure. Reducing complexity and keeping work visible are extremely difficult as the number of systems and locations in large enterprise architectures continuously proliferate. A lack of compatible interfaces between disparate environments, product-specific management tools, and the need to have teams specializing in particular products or technology stacks raise up barriers against DevOps best practices such as minimizing handoffs and swarming in response to failures in order to prevent their impacting downstream systems (where the cause of failure may be difficult or impossible to diagnose).

The introduction of DevOps in the IT industry is generally attributed to examples of success in the automobile manufacturing industry where utter failures to consistently produce vehicles meeting safety and design specifications at a speed that was profitable led companies like Toyota to completely redesign their manufacturing process. Likewise, the software pipeline has been improved to prevent defective modules from entering the final product and speed time to market so that business can profit from software initiatives. Being able to see the product as a whole, eliminating barriers between teams, and shortening handoffs are only possible when infrastructure can be fully managed. In this way, the “Ops” side of the pipeline becomes a predictable target for “Dev” to deploy to.

Governance

Properly implemented enterprise IT governance does not start in the part of the organization being governed. Rather, it is a business initiative that protects assets such as personnel and financial information, trade secrets, client and distribution data, and many more adjuvant resources that information technology processes. Consequently, an organization generally has a steering committee and written policies within a governance plan that specify how all of the digital capital of a company will be managed.

When a mature IT organization implements a governance plan, they will choose technology and processes that allow them to implement its guidelines successfully. Monitoring of multiple facets of these IT systems such as security, performance, usability, and more must also be implemented and utilized to provide feedback to the steering committee on how the governing standards are being applied. This implies a mapping between a governing standard and all systems affected by the rule. For instance, a rule requiring all personally identifying (PI) data be encrypted would require monitoring of finance databases, front-end web applications, messaging systems, caching mechanisms, and more.

When an organization has assets distributed across multiple and diverse systems, capturing and analyzing the level of compliance may become so burdensome that if it cannot be "baked in" in the form of automated processes, its cost can weigh heavily against the revenue of the systems being governed.

Further, large IT landscapes may have legacy systems that cannot be successfully migrated or shadow IT initiatives that have sprung up to manage pain points or weak integrations. Accurate reporting around compliance and eDiscovery becomes excessively difficult not only due to disparate systems but also has origins in poor Enterprise Content Management [ECM].

Security

In terms of security, the difficulty of managing hybrid or multi-cloud infrastructures as a single unit cannot be overstated. One of the most valuable tools in terms of both governance and security is the ability to automatically apply consistent policies throughout an organization's IT assets, with the goal of 100% coverage.

For example, in August 2021, Microsoft took urgent action on reports of a critical flaw which allowed attackers to gain access to the primary account key for Azure Cosmos DB customers via vulnerable Jupyter notebooks with access to the database.[7] A classic conundrum of enterprise security is that some products, such as Jupyter notebooks, exist for the very purpose of open sharing and collaboration, while the enterprise data stores which they access are not intended to be open. It was estimated that 30% of Microsoft's customers utilizing Cosmos DB and Jupyter were forced into an emergency reset of account keys in order to reestablish database security, as well as forensic research to determine whether private data had been compromised. The ability to consistently monitor access patterns, enforce network security by requiring database instances live behind firewalls or within protected virtual networks, and regularly update account keys are just a few of the ramparts that must be continually built to ward off the onslaught of attacks targeting companies threatened by this and similar vulnerabilities.

The risk to a company of a data breach[8] can range from a loss of customer confidence and market share to failure and fines.[9] RiskIQ, a cybersecurity firm recently acquired by Microsoft, estimates in its 2021 Evil Internet Minute[10] report that cybercrime is currently costing organizations

[7] https://chaosdb.wiz.io

[8] https://www.upguard.com/blog/biggest-data-breaches

[9] https://www.linkedin.com/pulse/most-expensive-fines-companies-faced-due-security-failures-andre/

[10] https://www.riskiq.com/wp-content/uploads/2021/07/Evil-Internet-Minute-RiskIQ-Infographic-2021.pdf

roughly $1.8M per minute, while McAfee's faned 2020 whitepaper on the hidden cost of cybercrime estimated more than trillion dollars lost each year,[11] a figure some disputed as far too low.[12] Further, small and medium businesses may be forced to shutter entirely in the face of a single cybercrime attack.[13]

The onslaught of cybersecurity risk is contributing to a more severe regulatory environment. In July of 2023, the US Securities and Exchange Commission [SEC] imposed new reporting requirements on public companies,[14] as well as some private and foreign entities, requiring them to disclose certain cybersecurity incidents within just four days of identifying them as material to avoid being subject to fines and potential legal action by the SEC. While not requiring companies to engage board members with cybersecurity expertise, as was postulated by many while the rules were being formulated, the ruling calls out the necessity of competent security leadership and mandates that companies identify both their process for responding to incidents and list the potential impact of a security breach on the company. Are companies prepared? Forbes noted in February of 2023 that only 51% of Fortune 100 companies benefit from a qualified cybersecurity professional serving as a director on their boards, and the situation becomes increasingly grim as companies decrease in size and resources.[15] The SEC regulations are only one thread in a large net of compliance obligations being cast by governmental authorities as Forbes[16]

[11] https://www.businesswire.com/news/home/20201206005011/en/New-McAfee-Report-Estimates-Global-Cybercrime-Losses-to-Exceed-1-Trillion

[12] https://cybersecurityventures.com/mcafee-vastly-underestimates-the-cost-of-cybercrime/

[13] https://cybersecurityventures.com/60-percent-of-small-companies-close-within-6-months-of-being-hacked/

[14] www.sec.gov/news/press-release/2023-139

[15] www.forbes.com/sites/forbestechcouncil/2023/02/06/90-of-boards-are-not-ready-for-sec-cyber-regulations/?sh=665fcff588e7

[16] https://fortune.com/2023/12/20/quiet-cybersecurity-revolution-economy-us-allies-new-threats-regulation-politics-tech-eric-noonan/

also notes, "the federal government is quietly directing a seismic shift in the economy by mandating stringent cybersecurity compliance across all 16 critical infrastructure sectors,"[17] meaning that very few businesses will be exempt from these new requirements.

Modernization

As we head into the foreseeable future, enterprise modernization efforts that feature multi-cloud and hybrid systems have nearly a decade of experiential learning to draw on. Monumental failures in these attempts have been well documented, for instance, those in several of the State of California's governmental systems[18] and the national embarrassment suffered in the rollout of Healthcare.gov which was initially unable to enroll users or even remain accessible for them to attempt that process. Per the Brookings Institution,[19] "… the Centers for Medicare and Medicaid Services (CMS) eschewed four management practices recommended by the Software Engineering Institute and the GAO: scheduling, estimating the effort needed for project tasks, data management monitoring practices, and milestone project reviews." Despite that blunt assessment of CMS, the blame for the site's woes extended far beyond that agency. A comprehensive retrospective featured in the Harvard Business Review's Cold Call podcast[20] pointed out that support from key participants was subject to political whims and that base requirements for how the site should function were undefined. Robinson Meyer in *The Atlantic*[21] notes

[17] www.cisa.gov/topics/critical-infrastructure-security-and-resilience/critical-infrastructure-sectors

[18] https://www.linkedin.com/pulse/risk-mitigation-through-successful-consulting-ramona-maxwell/

[19] www.brookings.edu/blog/techtank/2015/04/09/a-look-back-at-technical-issues-with-healthcare-gov/

[20] https://bit.ly/4ON5UIx

[21] www.theatlantic.com/technology/archive/2015/07/the-secret-startup-saved-healthcare-gov-the-worst-website-in-america/397784/

that the project was ultimately rescued by "a team of young people" who applied agile software development principles in order to fix issues ranging from logins and security to the user interface in order to eventually deliver a working website.

The lack of publicity around enterprise project failures is not because they don't occur in a similarly dramatic fashion. It is largely because there has historically been less regulatory compulsion for them to disclose mistakes unless there has been a data breach or other injury to outside parties, but a review of industry survey statistics published in Computerworld[22] indicates that keeping pace the massive evolution of the computing industry is a struggle for businesses of all sorts.

Enterprise applications may be constructed as large monoliths, hundreds or thousands of individual applications, be using a service- oriented architecture, or even have graduated to a microservice model – but they all have one thing in common: *data*. Risks of data migration are risks to the heart of a running business, and those risks are multiple. Data quality, security, and the compatibility of data types between old and new systems are some of the first items that must be examined in the assessment phase of a data migration. The risk of corruption, undocumented dependencies, application downtime, and the impact of all of these issues on the flow of business illustrate why Ernst & Young Analytics partner Chris Michell states that data migration is not a technology project, but a "business-critical risk mitigation program."[23]

Finally, you cannot modernize what you cannot catalog. Some enterprise systems age gracefully due to consistent governance throughout their lifespan, while others degrade into byzantine morass of complexity. In these cases, documenting what the output of the system

[22] https://www.cio.com/article/221827/14-reasons-why-software-projects-fail.html

[23] https://www.ey.com/en_lu/consulting/why-data-migration-is-about-risk-mitigation-not-technology

is, replacing that functionality, and starting fresh may be the only way to provide a secure and performant platform which meets the business need. Whichever approach of reformation or replacement is chosen, the problem remains that many business-critical systems must remain operational while the transformation is accomplished.

Upgrades

Anyone with a history in enterprise IT has a war story or two about sacrificing a weekend for a marathon upgrade session that required downtime for production systems. The cadence of upgrades has increased with the need to meet ever-growing security threats and assure technical stacks for crucial applications remain inside support windows for specific product versions.

Because of that, a longtime industry goal has been zero-downtime upgrades, also known as Evergreen IT. IaaS, PaaS, and serverless computing models have come reasonably close to achieving that (with varying degrees of completeness). In the case of hybrid and multi-cloud systems though, some of the old paradigms of calling an outage still apply.

Regardless of where an upgrade is performed, certain protective steps must be taken to reach that moment of relief when all systems are back online, and high among these is monitoring. Similar to data migrations, application compatibility is a risk, and testing must be done before and after an upgrade is completed. This is particularly true in scenarios where an upgrade cannot be rolled back once it is performed. If incompatibility is discovered after a nonreversible action is taken, then remediation will likely be required that may not have been anticipated in terms of the operational costs of performing the upgrade.

Whether software is upgradeable can often depend on how well an application has been maintained and documented during its lifetime. I personally witnessed a core application at a major bank that became so prone to failure that updates were only made through stored procedures

in its database and deployed through feature flags in production. The risk of an application drifting into such a dysfunctional state cannot be overstated, and effective monitoring coupled with timely remediation is a proven strategy for preventing upgrade debacles – which is exactly the approach taken by Azure Arc.

Monitoring

As mentioned, monitoring is a cornerstone of successful modernization and upgrade efforts, but its value extends to the entire enterprise ecosystem from governance and implementation to application consumption. Without effective monitoring, it is impossible to create the continuous feedback loop between intent and outcome to gauge success. No matter the topic – be it performance of a web application, protection from DDOS, or whether traffic between internal applications is properly secured – what is not being monitored can be assumed to be at risk, if not in a state of failure.

Obstacles to effective monitoring are numerous. Middleware products are inclined to customize their output to fit their tool's design, and this output may not be consumable by a particular monitoring tool without extra transformative steps. Problems also occur with too much or too little information being outputted; there is a reason some log emitters are referred to as fire hoses! Logs have to be segregated and filtered for noise or sometimes composed into something that is meaningful for whomever is responsible for auditing their output. Transactional data, for instance, might go through extensive analytical processing after being shipped to a data warehouse. If these challenges are not met correctly, then organizations can be left "flying blind" in key areas of governance, security, and profitability of business objectives.

Management

A final and not insignificant pain point in having IT infrastructure spread across multi-cloud and hybrid systems is the inability to manage sprawling assets from a single point of view. While individual teams can and should manage systems nearest to them, a high-level view is still critical because of interdependencies throughout disparate systems. Even though an operationally mature organization will have strategies for maintaining high availability of critical systems, downstream systems should also be prepared to manage failure if their operation would be interrupted or perhaps fail also in the event of an outage.

System administrators with deep experience in topics such as networking, virtualization, or databases will still face entirely new paradigms when moving to cloud. Despite the fact that mega-cloud infrastructure is composed of roughly the same components, the parlance and the perspective from which they manage will be much different. Further complicating matters are attempted modernizations or upgrades that are stuck for long periods in an incomplete state, requiring management of "accidental hybrid" systems. When a multi-cloud approach is taken, the processes and tools will vary between vendors, adding yet another facet of management complexity.

Finally, management requires visibility and thus depends on accurate monitoring. Whether the metrics are financial, operational, or project based, there must be visibility into the entire corpus of an organization's IT assets if management is to successfully steer toward the business' objectives.

Throughout this book, we'll explore what Azure Arc is, its capabilities, and how it can lessen the hardships associated with development, governance, security, modernization, upgrades, monitoring, and management in multi-cloud and hybrid environments.

CHAPTER 2

What Is Azure Arc?

Arc – A Single Control Plane Across Multi-cloud and Hybrid Architectures

The unification of IT infrastructure management under a "single pane of glass" has been an end of the rainbow objective since the beginning of enterprise computing. The Jenga tower of interdependent resources that is created as organizations grow can be read like tree rings to examine what led to the current health of IT systems – the thin year when we didn't have sufficient staff and nothing was documented, the fat year when many new products were purchased but without sufficient time for integration testing, or the steady years when there was a consolidated stack working relatively well followed by the evidence of fire when a major security breach occurred.

In every case, both business and IT are seeking visibility not only for failure analysis but also to gather the data around what will support an optimal growth pattern. Classic IT Service Management [ITSM] processes and frameworks, notably Information Technology Infrastructure Library [ITIL], were initially seen as contrary to a modern DevOps approach as their elongated process flows and reduced agility hampered the speed of iteration despite the governance they provided being undeniably necessary. The principles for managing change processes across not only

© Ramona Maxwell 2024
R. Maxwell, *Azure Arc Systems Management*,
https://doi.org/10.1007/978-1-4842-9480-2_2

services but also infrastructure had to be blended with agile practice so that the ability to move faster would not jeopardize the outcome, and tools to accomplish this were not always readily available or mature.

Azure Arc is revolutionary in its ability to realize the goal of gathering assets into a single theater for observation and management. Its comprehensive approach to asset management uses the tools, processes, and approaches pioneered by Microsoft in the Azure cloud and allows you to leverage these same processes, tools, and approaches on any workload anywhere, even other cloud providers. For instance, along with Azure Stack Portfolio, Azure Stack IoT, and SQL Server, things like Kubernetes clusters or servers running in your own data centers or on competitor clouds are all viable targets.

What are some of the benefits? I recently gave a talk on container support for AWS Lambda (a compute service offering a runtime on which to execute functions to process data.[1] The function outputs can be consumed by various types of workloads), in which I noted I used a total of four command line interfaces (Docker, AWS, ECS, and SAM) to deploy my containerized function. Using Azure Arc as the management plane, you are able to use its CLI and APIs consistently across the entire catalog of managed assets (e.g., not switching to AWS CLI to manage a server located there). Not only that, but the view of asset inventory outside of the Azure portal is not segregated – instead, you can view all assets of a particular type together (see Figure 2-1) and manage them as a group. Further, you can apply policies and then monitor compliance since tools like Azure Policy and Azure Monitor are included.

[1] https://docs.aws.amazon.com/lambda/latest/dg/welcome.html

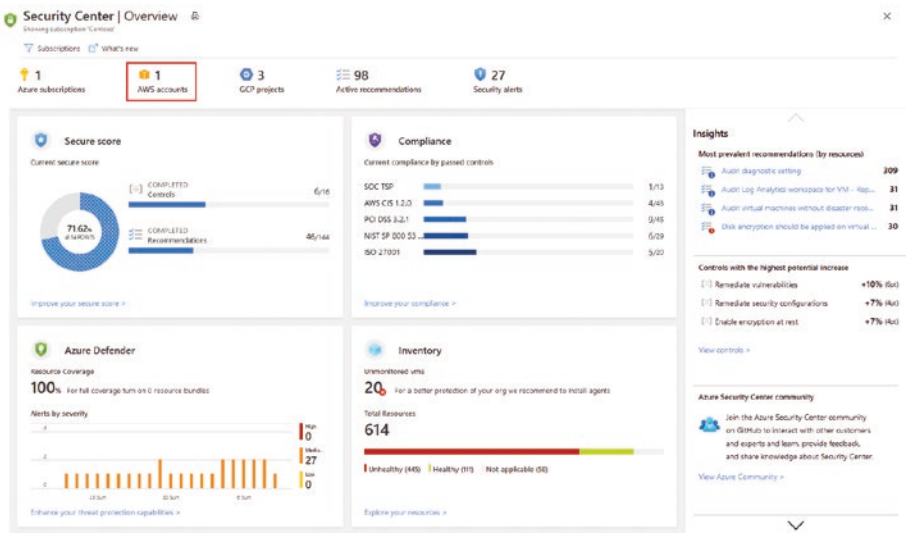


Figure 2-1. *Monitoring AWS policy compliance, including AWS CIS in Azure Security Center by means of Azure Arc analytics agent*[2]

Internal Architecture

The genesis of Arc is an interesting history that will be touched upon throughout this book. Here, though, I'd like to dive into the functional pieces of Arc and how the capabilities under discussion are achieved. To understand Arc, it's helpful to understand the core pieces of Azure itself. Long before operational visibility came along as icing, the Azure cake was baked into its current form with a couple of its key ingredients being Azure's Fabric Controller and Resource Manager.

The Fabric Controller reminds me most of the Diego Brain utilized by Cloud Foundry to spin up container instances on virtual machines based on requests from Cloud Foundry's Cloud Controller such as "give

[2] https://docs.microsoft.com/en-us/azure/security-center/quickstart-onboard-aws

me a Windows instance for this .Net Core app horizontally scalable by x factor" or an Alpine Linux instance for a Ruby app and many more combinations. Diego auctions off the app's request for a home to the server farm where the machine with the correct supporting OS and capacity can be selected. Next, Diego creates the container instance and runs the app, begins emitting its routing data to service requests, and sends its logs to an aggregator. If health checks are configured for an application, Diego also monitors whether the application is healthy and reports failures back to the Cloud Controller, so a new container instance to run the application can be created.

Initially, Azure also focused on cloud hosting of applications, and so its Fabric Controller, like Diego, serves as an intermediary mapping container requests to the hardware layer and continually monitors their health status in case any need to be replaced. Today, though, its purview is much broader than hosting applications since Azure provides not only PaaS but also IaaS services. Thus, the fabric controller also regulates other types of hardware, including networking and storage controllers.

The Fabric Controller itself sits on dedicated hardware in each cluster of approximately a thousand servers within an Azure Data Center.[3] The Fabric Controller servers are configured to be highly available, as are the clusters themselves organized according to what Microsoft refers to as "fault domains," which are at every crucial intersect where a hardware failure may occur (typically server racks which share a NIC and power supply, but since their definition is hierarchical, subsidiary components like disks may also belong to a fault domain). The Fabric Controller uses agent-based communication to manage guests, and this communication is always initiated by the host with responses from the guest assumed to be untrusted. This inability for a guest OS to initiate communications

[3] https://learn.microsoft.com/en-us/azure/security/fundamentals/infrastructure-components#azure-management-by-fabric-controllers

with the host is just one of the security boundaries that protects Azure infrastructure from the vulnerabilities that virtual machines running on nodes (physical servers) in the cluster may be exposed to. The Fabric Controller itself runs on VMs, allowing it to take advantage of security isolation provided by the virtual machine hypervisor including protection of the node's physical hardware and memory space.[4] Additionally, the hypervisor provides critical isolation of VMs, not only for the hardware layer but also limiting network communications to VMs within the same VNet so that "side channel" incursions where a malicious VM is used to attack other VMs in the same cluster are prevented.

The Fabric Controller manages three different operating systems, including Host OS (which is a minified and hardened version of Windows Server) on the node's root VM, Native OS which is the operating system for the Fabric Controller itself (as well as some Azure services that do not run under a hypervisor such as storage), and thirdly a guest OS which provides the runtime for the VM workload. The guest OS may be chosen from a variety of Linux distributions as well as recent versions of both Windows Server and consumer editions of Windows. The Fabric Controller also provides health monitoring and actively searches for a new host and restores a guest VM that goes down for any reason.

If you are one of those people who exercises the liberty to make fun of Hyper-V, you might reconsider when you understand how a modified version of the same hypervisor is foundational to Azure cloud infrastructure. The role of a hypervisor is of course to distribute the resources of a powerful server among as many virtualized server instances as the machine can comfortably host. The Azure hypervisor goes beyond this core task to provide security isolation so that no virtual machine can communicate with others on the same host, and additionally it is minimized to what is required to operate Azure so that its attack surface is shrunken.

[4] https://learn.microsoft.com/en-us/azure/azure-government/azure-secure-isolation-guidance#strongly-defined-security-boundaries