The definitive guide to practical AIOps

AI software that supports IT operations

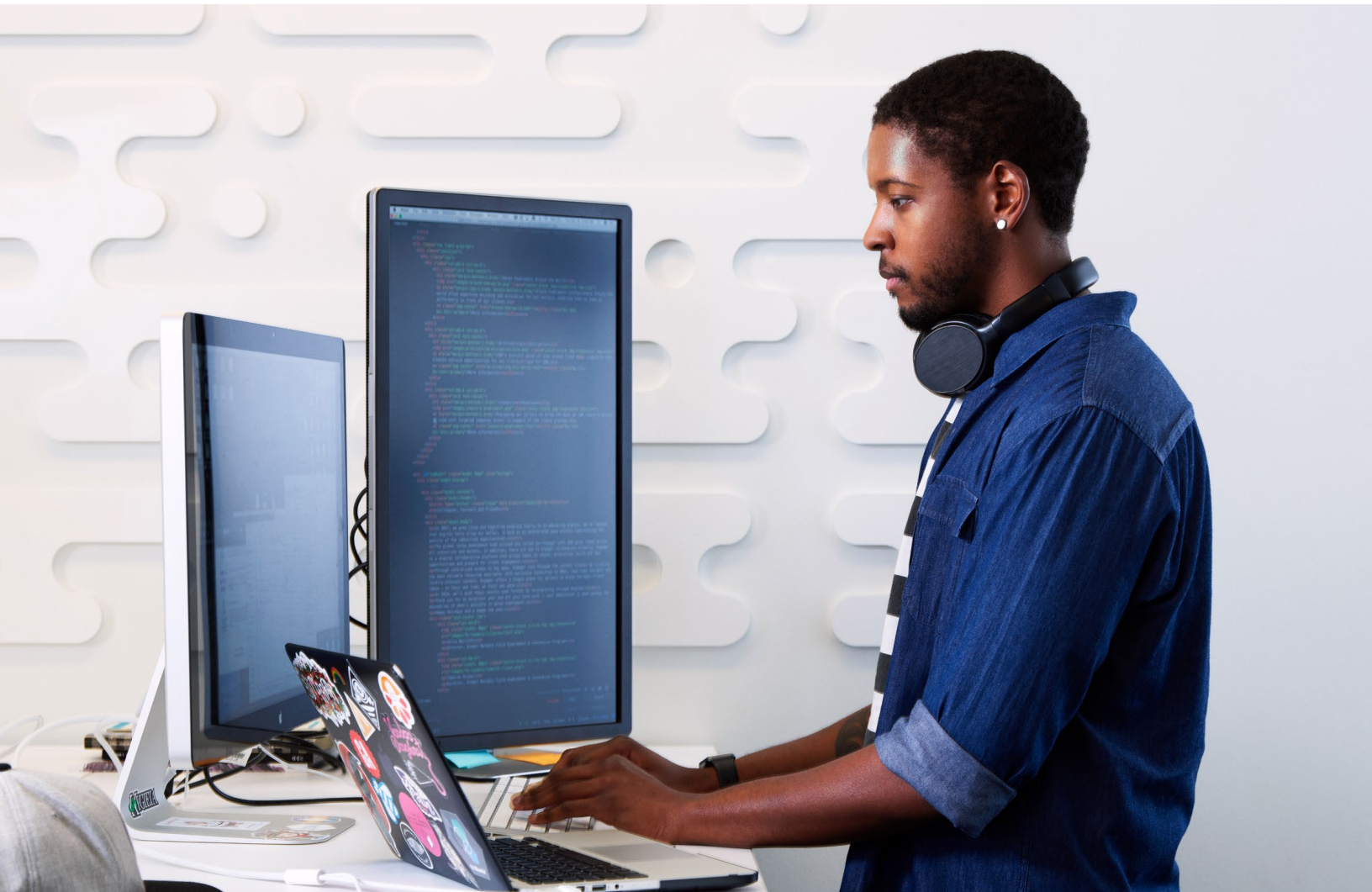

Table of contents

Introduction

Today's digital transformation initiatives are aimed at providing better user experiences that drive increased customer engagement and loyalty. Customers are increasingly interacting with enterprises through digital and mobile experiences, and, in turn, business outcomes are determined by how quickly an enterprise delivers new services to market that surpass the services provided by its competitors.

Applications underpin all experiences, and new generations of cloud-native applications are designed to enable rapid introduction of new services and personalized customer experiences. Highly distributed applications now employ microservice-based and container-based building blocks, hosted across complex environments, spanning data centers, public cloud and the edge. The complexity of managing these environments has given rise to software that can operate beyond human scale and comprehension to tackle the decision-making of where and how applications can be supported.

In this ebook, we introduce the current practical applications of artificial intelligence (AI) for IT operations (AIOps), provide guidance on how and why enterprises should be using AIOps today, and offer insight into what the future holds. The goal is to help maximize the business value derived from AIOps while preparing for a future where AIOps is an imperative to provide resilient application performance and allow employees to focus on driving business innovation.

What is AIOps and why does it matter?

AIOps is the application of AI to IT operations. According to Gartner's definition, AI "applies advanced analysis and logic-based techniques—including machine learning—to interpret events, support and automate decisions, and to take actions."[1]

It's important to note that the objective of AIOps is not to mimic human intelligence. Rather, AIOps seeks to apply algorithms to solve specific problems, often much faster, much more accurately and at much higher scale than a human. As applications become more distributed and complex, and as the infrastructure those applications run on gets more distributed and complex—often spanning from data centers to public cloud to edge computing—it becomes untenable for applications to perform reliably and efficiently at scale without AIOps. And as enterprises that adopt AIOps are discovering, their employees can be more productive and spend more time on innovation when AIOps frees them from troubleshooting problems, performing root cause analysis, or conducting routine maintenance and other "keeping the lights on" activities.

A practical way to think about IT operations

IT organizations have embraced a traditional approach to operations for several decades. Resources are intentionally overallocated to accommodate for future growth and potential unforeseen circumstances. Monitoring tools dedicated to constituent parts of IT, for example, applications, servers, storage, networks elements and cloud instances, are pervasively used to track access, utilization, latency, adherence to rules and other key performance indicators (KPIs) of interest. KPIs and threshold-based alerts indicate when actions need to be taken to avoid or mitigate issues or optimize efficiencies.

Most IT organizations spend significant resources tending to their IT environments. In many enterprises, there are so many alerts being generated by different monitoring tools that a class of monitoring has emerged to filter out all but the most severe performance issues or risks, which are then surfaced for IT staff to investigate and remediate. These tools incorporate advanced analytics and logic-based capabilities to classify alerts that can in all probability be ignored and suppressed from view, so staff can more quickly identify the root cause of a problem when a significant issue exists or, better yet, address a risk before it becomes a big problem.

At its core, AIOps can be thought of as managing two types of challenges:

– Artifacts and attributes that aren't supposed to change, for example, static, or may change in predictable ways, for example, periodic. AIOps addresses these scenarios through machine learning (ML) programs that establish baselines and identify a deviation from normal.

– Artifacts and attributes that may change in unpredictable ways, for example, dynamic. Dynamic optimization establishes rules and analytics intended to appropriately respond to situations that have never been seen before.

Let's explore these management challenges using common examples in IT.

AIOps use case: Domain-specific deviation from baseline
Application performance management (APM) represents a class of tools that are used by a majority of Fortune 100 companies and large global enterprises. The outcomes sought through APM include facilitating root cause analysis (RCA) and remediation of application performance degradation issues, reduction in the frequency of application performance degradation, and reduction in the number of IT tickets associated with application performance.

APM platforms ingest millions of disparate records every second and embed AIOps capabilities to establish a baseline of what normal looks like and, therefore, can identify deviation from normal behavior as potential root cause of performance issues. ML algorithms in APM platforms can recognize seasonal and periodic patterns, adjusting the baseline accordingly.

Here, AIOps works in concert with IT staff responsible for resolving issues and mitigating risks that may lead to issues. The leading APM platforms provide visibility into a rich set of KPIs and the analytics that provide actionable insight into what's happening and what to do about it.

AIOps use case: Predictive IT operations
IBM Cloud Pak® for Watson AIOps can incorporate structured and unstructured data from a variety of IT systems, including monitoring tools and workflow platforms. Logs and KPIs, alerts, trouble tickets, and topology information can be ingested and analyzed by ML to establish baselines and norms over time. IBM Watson® includes AI and natural language processing (NLP) technologies to correlate unstructured and structured data to provide as much context as possible about an event to speed RCA and remediation. Watson can integrate with collaboration tools to speed resolution and can also compare situations to events that have occurred in the past to learn from past resolution efforts.

IBM Cloud Pak for Watson AIOps helps customers achieve more comprehensive incident analyses, faster incident resolution and improved incident avoidance. Cross-platform integration serves to bring siloed teams together, aligned around a common understand of the entire environment and circumstances, resulting in better end-user experiences and the avoidance of downtime that can inflict both economic and reputational damage.

AIOps use case: Dynamic optimization
Application environments consist of multiple stacks or layers; at the top is the most important entity, the application. The purpose of IT operations (ITOps) is to make applications run well. Therefore, everything below the application layer should support the application. Insight into the application's fluctuating demand and configuration can be used to dynamically match demand to required resources.

Every resource in the stack can experience contention or issues that will impact the higher layers, all the way up to the application. For example, an Amazon Elastic Compute Cloud (Amazon EC2) instance assigned with an Amazon Elastic Block Store (EBS) volume not suitable for the application input/output operations per second (IOPS) demand may negatively impact application performance despite Amazon EC2 having ample compute capacity.

IBM® Turbonomic® offers full-stack visibility and control into every environment it manages. IBM Turbonomic software determines the relationships between the layers, the available capacity of resources and fluctuating demand of the application. Application resource management (ARM) is accomplished by abstracting all layers of the IT stack, from discrete physical resources up through every layer to the application components, into a market economy supply chain. IBM Turbonomic applies economic principles to manage IT resources and can execute actions at each layer to assure application performance. Analytics are employed to compare the available supply in the underlying layer to the demand being requested by the higher layers, providing the ability to continuously make informed decisions about how to meet demand. By making each layer application aware, IT resources are continually optimized to assure performance. It's no longer necessary for IT staff to adjust resources manually after an application performance issue is identified.

AIOps use case: Service management automation
Traditional IT resourcing has remained reactive and manual due to siloed resource alerts that arise with limited context in relation to application performance. The burden of resource resolution falls to IT staff. Process automation is useful but insufficient. To become agile, companies must embrace automation, from provisioning infrastructure, for example, infrastructure as code, to workload and application deployment.

The potential of AIOps is to avoid issues rather than waiting for them to occur and requiring IT staff to resolve them. Preventive measures must be determined and automated in application-aware software because only software can prevent performance issues, which are beyond human scale. But before clients begin to accept AI, they need to build trust that the decisions made by the software are safe to implement. As customers gain confidence in the trustworthiness of AI decisions, a natural step is to automate service management processes that previously required manual review, approval and execution of corrective action. IBM Turbonomic software is authorized to raise a service management ticket identifying a congestion issue, propose an action to resolve the issue, approve the workflow action based upon full analysis of the action's implications, validate that taking the action has eliminated the congestion, and close the service management ticket—all without IT staff involvement.

IBM Turbonomic is deployed in many of the world's largest and most complex environments, including more than half of the Fortune 100 companies. In large enterprises, multiple teams and stakeholders initially review the actions generated by IBM Turbonomic and manually execute actions prior to enabling scheduled or fully automated actions. In complex organizational structures, automation can be implemented to span company culture, organizational processes and IT operations workflow systems, fostering cross-team collaboration through a "single source of truth." IBM Turbonomic integrates with major IT platforms and environments, enabling teams formerly in distinct silos to see their familiar data in the context of the larger environment. Consequently, they see and understand the relevance of initial automation decisions. Through a process of starting with manual actions, followed by actions scheduled during change windows, and moving to immediate automated execution, IT organizations can move at their own pace and according to their own policies.

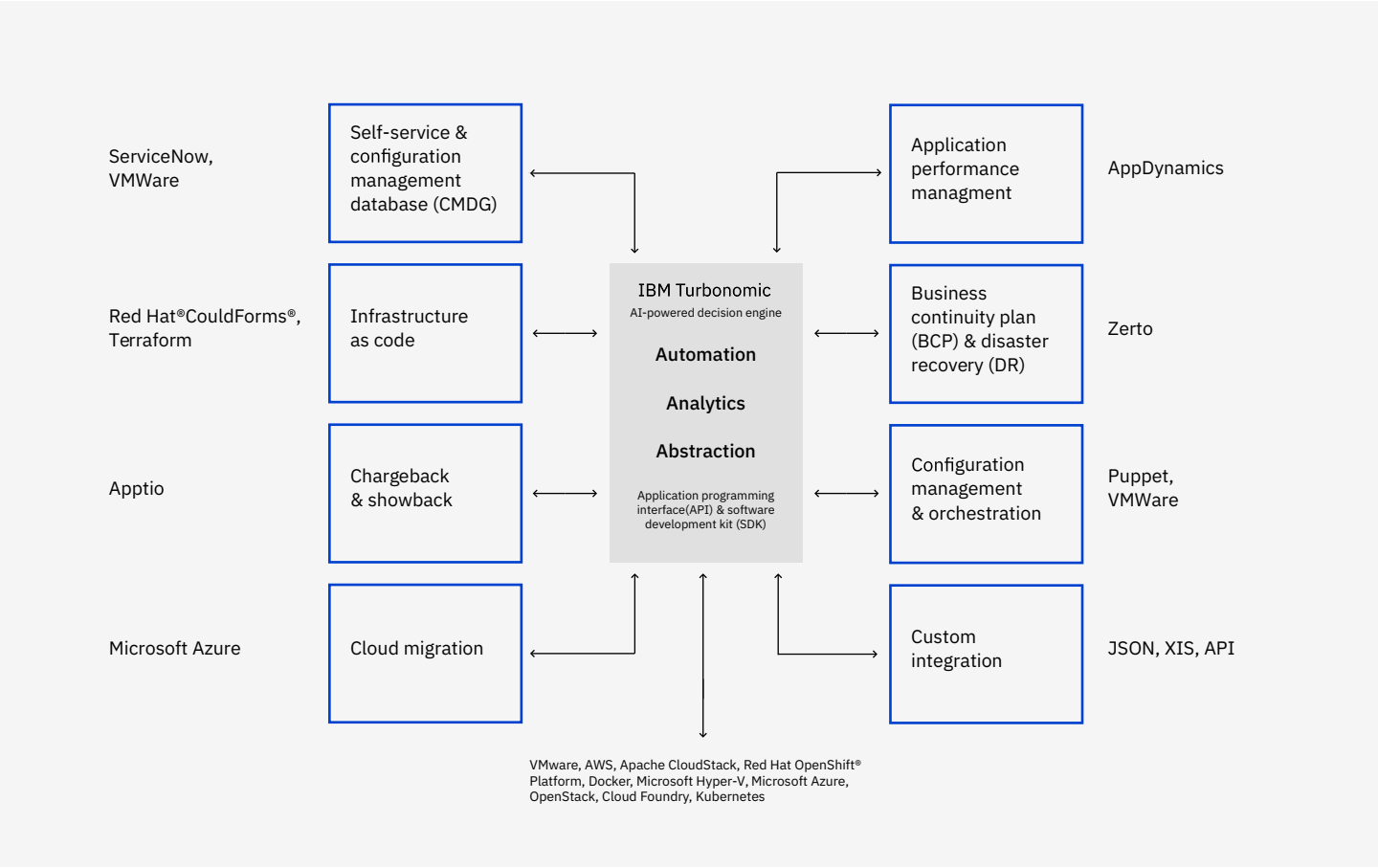

Figure 1. Application resource management lifecycle ecosystem

What problem are you trying to solve?
Most enterprises ignore thousands of alerts each day, manually addressing a subset of those that pose a risk of a serious performance issue or threat, often not preventing issues but reacting to them. AIOps tools and platforms that incorporate ML techniques to establish baselines can detect meaningful deviations from baseline. The value they provide translates into faster resolution of issues and the ability to prevent issues that arise from periodic fluctuations in application demand or IT resource consumption. Unfortunately, they're not able to take customers very far on the path to self-driving ITOps because they can't address unpredictable demand in dynamic environments.

As applications get more complex and the IT resource options introduce new complexities, such as the rapid ascent of dynamic and ephemeral container environments, it becomes more difficult to understand the connection between full-stack relationships and dependencies. More importantly, a primary reason that applications are getting more complex and distributed is because digital engagement with end users continually evolves and becomes more dynamic, which leads to unpredictable application demand situations. Applications composed hundreds of microservices that can be independently modified to continually update services for thousands or millions of consumers who are also targets of new types of advertising campaigns means that unpredictable events will become commonplace. Never-before-seen events may be a regular occurrence.

There's a clear and immediate need for a solution that can address the challenges related to multicloud adoption and containerization of applications that support highly dynamic demand scenarios. Simply put, the problem to be solved is giving applications the resources they need—when they need them. To help assure application performance, consideration should also be given to the relationships between each layer of the IT stack, from the application components to the physical resources that are assembled to build containers and virtual elements. This understanding of the full-stack relationships is what's missing in traditional IT toolsets and approaches.

IBM Turbonomic AIOps

The most important objective of IT infrastructure is to provide applications with the resources they need to deliver their service levels. A companion objective is to do it as cost-efficiently as possible and adapt to changing environment and application demand scenarios by dynamically adjusting resources over time. Key capabilities include:


Application-aware optimization


Support for on-premises, hybrid and multicloud deployments


Full-stack visibility and control across the full environment


Trustworthy and automatable actions


Enforcement of business policy compliance

"Turbonomic is very focused on the mission of providing automated application resource management.For Turbonomic, cost optimization is just a benefit of the perfect alignment between infrastructure resource supply and performance requirements."[1]

Abstraction, analytics, automation
Thus far, we've touched upon several use cases for AIOps, including:

- Real-time anomaly detection for risk mitigation and problem avoidance
- Faster root cause analysis through event correlation when problems do occur
- Low-priority alert suppression, so high-priority alerts get better visibility
- Capacity planning and management, based upon predictive analytics
- IT service management automation

A primary challenge for large enterprises is exponential complexity inherent in modern applications built on microservices and deployed on a modern containerized multicloud infrastructure.

Enterprises are rapidly adopting AIOps with ML embedded in many of their monitoring and management systems, including APM, service management, infrastructure as code and configuration management. IBM Turbonomic serves as a control plane, a modern application hosting platform, to tie those systems together and scale to millions of managed elements in a single instance. As AIOps technology continues to evolve in independent tools, IBM Turbonomic learns about the changes through its integrations and incorporates that data into the IBM Turbonomic AI engine's decision process, enabling large enterprise development and IT organizations to grow along with their technology investments.

Conclusion

In today's competitive business climate, applications drive revenue, profits, customer experience and customer retention. Application success is determined by an enterprise's ability to innovate through development teams and IT operations, but application and infrastructure complexity, scale and scope are rising exponentially. AIOps represents a wave of opportunity for IT organizations in the coming years. Business leaders and ITOps teams will decouple themselves from reactive tools and embrace technologies that accelerate business innovation, prevent issues and maximize cost efficiencies. Rather than responding to issues when they occur, AIOps offers probabilistic outcomes based upon advanced analytics.

Future planning must contemplate the speed at which AIOps technologies evolve and the speed at which organizations trust AIOps-generated insights and actions. IT staff will be increasingly called upon to focus on business innovation, and it will be up to intelligent AI software to manage the complexity of IT environments and create a control plane capable of self-driving operations that ensure continuous health versus merely striving to speed mitigation and self-healing.

Why IBM?

IBM is one of the few companies with AI-powered automation capabilities that span business and IT. This technology, when used by clients of IBM, can help assure application performance and governance by dynamically resourcing applications across hybrid and multicloud environments. With IBM, gain fully automated actions to ensure applications get what they need to perform while adhering to your business policies.

For further information, visit ibm.com/cloud/turbonomic.

IBM Corporation
New Orchard Road
Armonk, NY 10504

Produced in the United States of America
April 2022

[1] Solution Comparison for Public Cloud Third-Party Cost Optimization Tools, Gartner Research, 10 December 2019.