Business white paper

Hewlett Packard Enterprise

A new era of intelligence for the university data center

Smart moves: HPE InfoSight transforms data center storage

Table of contents

Introduction

The bar for success is rising in higher education. Students and parents expect instruction enhanced with the latest technologies, affordable learning options, opportunities to connect to job prospects, and experiences that improve student quality of life. These expectations place ever-growing pressures on colleges and universities like yours: Attract and retain the best students and staff, provide cutting-edge, technology-enabled teaching methods and student experiences, maintain consistent funding levels—and make the most efficient use of the many applications that make this all possible.

Impacts are severe when the infrastructure powering these applications is hard to administer (or worse, fails). The digital era has disrupted IT's role in this high-stakes scenario. Traditionally, IT carried much of the responsibility to ensure infrastructure availability and performance. But today, educational institutions must evolve to rely more on intelligence to self-heal, self-manage, and automate their data centers.

Hewlett Packard Enterprise (HPE) has broken away from convention, transforming how storage is managed and supported for institutions of higher education. The transformational key is HPE InfoSight, a cloud-based, predictive analytics and machine-learning engine that watches over the data center 24x7 so IT administrators don't have to. HPE InfoSight ensures 99.9999% guaranteed availability and predicts and resolves 86% of problems before they can disrupt the business of learning, research, and administration. And because HPE InfoSight analyzes millions of sensors a second, everyone benefits as the university's systems grow continuously smarter and more reliable.

This paper explores how **HPE InfoSight** and its recommendation engine can pave your institution's path toward an automated, self-healing and self-managing data center that frees IT to focus on creating value and allows your institution to turn exponentially growing data into valuable, actionable insights.

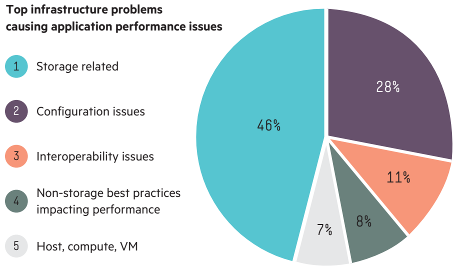
The case for a smarter data center

University leaders and IT administrators are aware of the compelling benefits of digital transformation overall—and artificial intelligence (AI) in particular. AI can amplify human capabilities by using machine learning, or deep learning, to convert the fast-growing and plentiful sources of data about all aspects of a university into actionable insights that drive better decisions. But when planning a transformational strategy, these leaders must prioritize operational continuity. It's critical to protect the everyday activities of learning, research, and administration that rely on the IT infrastructure to consistently deliver data to its applications.

Next-generation platforms such as **HPE Storage enterprise flash arrays** continue to drive up storage and application performance specs. But fast storage alone can't guarantee that you'll have reliable, non-disruptive access to data or eliminate the need for manual attention from IT. The unsustainable break-fix-repeat demands of the data center's increasingly complex infrastructure ultimately hold back both IT operations and everyone (students, staff, and administrative leaders) who relies on it.

Traditional monitoring and support no longer make the grade.
In the past, university IT organizations could only rely on monitoring tools to troubleshoot their environment. Unfortunately, this required IT staff to spend dozens of hours mining log files and interpreting graphs, trying to find the cause of a disruption so they could resolve it. When IT couldn't do it alone, they had to bring in external consultants for support—resulting in expensive, time-consuming, and multi-tiered escalations.

Today, IT organizations in higher education have reached a flashpoint: The historical approach to managing the IT infrastructure no longer makes the grade. Manual monitoring and reactive troubleshooting can't keep pace with fast-growing student and staff demands for future-facing technology to enhance all aspects of university experience and operations. And it's no longer acceptable to discover disruption-causing issues only after the fact: The IT infrastructure has become vital to the educational institution's bottom line.



The costs of manual infrastructure management are too high.
Your university needs IT resources to support your most strategic initiatives. This means that IT can no longer afford the one-off, manual care and feeding required to constantly ensure optimal performance for every application. Institutional workloads change regularly, requiring ongoing allocation of specialized resources and time-consuming trial and error to fine-tune the infrastructure. Overprovisioning has been a go-to work-around in this scenario—but at the cost of paying for more than what's needed. And even if business requirements stay constant, an IT organization caught in the guesswork of the reactive break-fix cycle misses opportunities to improve performance: Maybe moving an application from all-flash array (AFA) to hybrid storage or resizing a volume will improve a specific issue. Or maybe not. The inherent uncertainty in this approach carries a tremendous opportunity cost.

Intelligent, predictive data center management is the future of higher education.

Forward-looking universities seek to set themselves up for successful digital transformation that can both drive successful student outcomes and contribute to sustainable, long-term institutional growth and profit. This requires a data center management solution that transforms how the infrastructure is managed and supported. IT needs a way to predict problems before they occur and receive in-depth recommendations for the actions needed to optimize performance and available resources.

AI sees beyond limits

Humans can see the present and remember a little bit about the past. This isn't so different from how an IT administrator's tools have managed data centers in the past. Removing the burden of managing the infrastructure in this way requires foresight to predict problems before they occur. Determining how to optimize the environment requires deep intelligence on the underlining workloads and resources.

Unfortunately, traditional tools get an "F" in this situation for four key reasons:

1. Inability to learn from others: Analytics that simply report on local system metrics offer limited value: They can't tap into data gathered about the behavior of thousands of peer systems across the university that could help detect and diagnose developing issues. In contrast, a global approach to data collection and analysis can pool observations from an immense variety of workloads. This allows for rare events identified at one site to be preemptively avoided at another and for more common events to be detected earlier and with greater accuracy.
2. Analytics stuck in infrastructure silos: Problems that disrupt an application can occur anywhere in an institution's infrastructure stack. Tools that only provide status on a per-device basis tell just a part of the story. In contrast, cross-stack analytics that correlate across multiple layers (applications, compute, virtualization, databases, networks, storage) can tell the full story.
3. Missing domain expertise: Predictive modeling demands deep domain experience—understanding all the operating, environmental, and telemetry parameters in each system of the infrastructure stack. General-purpose analytics only can go so deep. To bridge the gap, pairing domain experts with AI can enable machine-learning algorithms to identify causation from historical events to predict the most complex and damaging problems before they occur.
4. Inability to act: The ideal state of the university data center is a self-managing operation that's free of the need for human intervention. This requires knowing what changes are needed to avoid a problem or improve the data center environment—and being able to act on behalf of the system administrator. Achieving this level of automation demands a proven history of automated recommendations that provide the needed level of trust and confidence.

Infrastructure powered by artificial intelligence can overcome these limitations through the following framework:

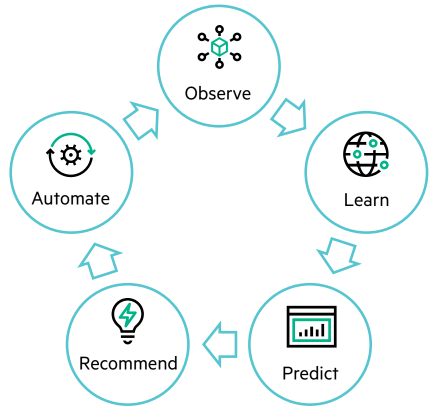


Figure 1. Artificial intelligence for the infrastructure framework

1. **Observe:** AI simultaneously monitors all the institution's systems to develop a steady-state understanding of the ideal operating environment for every workload and application. AI then identifies abnormal behavior by recognizing the underlining I/O patterns and configurations of each environment.
2. **Learn:** Deep system telemetry coupled with global connectivity creates a foundation of data that exploits the experiences of every system connected throughout the institution. Machine learning in the cloud rapidly accelerates the AI knowledge and global learning.
3. **Predict:** For any new problem detected, AI can learn to predict the issue and use pattern-matching algorithms to determine whether any other system in the installed base is susceptible. Application performance can also be modeled and tuned for new infrastructure based on historical configurations and workload patterns.
4. **Recommend:** Based on the predictive analytics, the AI determines the appropriate recommendation what's needed to improve and ensure the ideal environment. Recommendations are system operational decisions that free the institution's IT organization for more strategic work and eliminate the guesswork in managing the infrastructure.
5. **Automate:** Through mutual trust between the infrastructure and the AI, recommendations can be applied automatically on behalf of the IT administrators. When automation is unavailable, specific recommendations can be delivered through support case automation.

AI can watch over the infrastructure; continuously learn from a global installed base; and apply what it learns to predict and prevent issues, eliminating the guesswork in managing the infrastructure. The result is a smarter and more reliable infrastructure for your institution.

Benefits of HPE InfoSight:
86% of issues automatically predicted and resolved[1]

99.9999% of proven availability[2]

79% reduction in IT storage OPEX[3]

85% less time spent on storage issues[4]

Mesa Community College
Objective: Sought to standardize all IT infrastructure to ease management and enable easy scalability

Approach: Chose to standardize on HPE Nimble, including HPE InfoSight, for storage

Outcome: Cut infrastructure management time by 3:1, saving hundreds of thousands of dollars over 5 years

Global learning

AI and machine learning requires massive amounts of data beyond the limited logs and metrics of traditional hardware platforms. The HPE storage platforms, which use **Intel® Xeon® processors and SSDs**, were architected with deep diagnostic sensors. HPE InfoSight has been collecting this data since 2010, giving it a breadth of telemetry that creates an architectural advantage.

HPE InfoSight: A smart move for colleges and universities

HPE InfoSight was founded on the belief that infrastructure management and support must evolve. Instead of dealing with unexpected problems and reactive vendor support, the IT organization can rely on AI to make the infrastructure smart enough to anticipate issues before they occur and resolve them without human intervention. Only this self-healing model can enable institutions like yours to most efficiently use your resources to drive the needed outcomes.

Through HPE InfoSight, HPE successfully enables the university to exploit data assets across the full spectrum of AI machine learning, providing context and guidance on both current situations to ideal future states:

- **Insight:** What's happening and why? (For instance, how many students are enrolled, what are they doing, and what's driving these actions?)
- **Foresight:** What will happen next and why? (For instance, how many students will return next year and why others aren't being retained? Or can a university save money by retiring certain equipment?)
- **Forecast:** Given what will happen next and why, what's the best way to proceed? (For instance, what specific actions can the institution take to retain a certain student? Or when is the ideal time to replace a specific air conditioning system to obtain the best outcomes?)
- **Deep learning:** Create and exploit vast data sets to take insight to the next level. (For instance, use computer vision to answer questions such as whether a particular student should be in a certain location right now; whether a piece of equipment should be moved, whether people smoking in a designated non-smoking location, etc.) Further, deep learning can use voice-to-text analytics to not only transcribe a full lecture in real-time and automatically send transcripts to all participants, but also analyze what was said by both lecturers and students for sentiment or content.

Built on a unique approach to data collection and analysis, HPE InfoSight collects and analyzes millions of sensor data points every second from our globally connected installed base. This sensor data provides comprehensive measurements of the operation and state of each system, subsystem, and surrounding IT infrastructure. It learns from this data to drive its **predictive analytics** and **recommendation engines**, resulting in significant impact for all our customers.

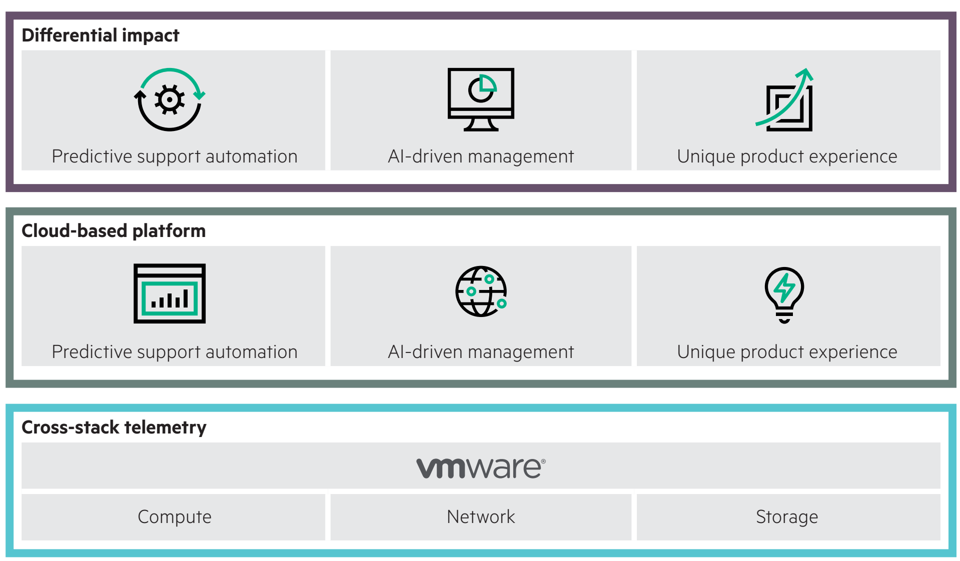


Figure 2. HPE InfoSight platform

[1] **"Redefining the standard for system availability,"** 2017

[2] **"HPE Get 6-Nines Guarantee,"** 2017

[3, 4] **"Assessing the financial impact of HPE InfoSight predictive analytics,"** 2017

Predictive analytics engine

Seeing ahead to eliminate disruptions and to get IT ahead

HPE InfoSight offers predictive analytics that extend across the university infrastructure lifecycle—from planning to expanding.

- **For planning:** HPE InfoSight “right-sizes” new infrastructure by anticipating performance and resources needed based on different applications seen in our installed base. Through telemetry from deployed systems, HPE InfoSight continuously refines its machine-learned models for better sizing accuracy.
- **Once arrays are deployed:** Predictive analytics transform the product and support experience. HPE InfoSight constantly looks for leading indicators of problems and automatically resolves them before IT organizations that rely on HPE InfoSight even realize that there is an issue. If HPE InfoSight detects a new issue, it learns to predict the issue and prevent other systems in the installed base from seeing the same problem.
- **Completing the lifecycle:** HPE InfoSight accurately predicts future capacity, performance, and bandwidth needs based on historical use and autoregressive and Monte Carlo simulations.

Predictive analytics go beyond storage
The predictive capabilities of HPE InfoSight go beyond storage.

For example, HPE InfoSight predicted and prevented a catastrophic All-Paths-Down situation for HPE Nimble Storage customers due to a potential issue with a network VIC card in the host. Leveraging HPE InfoSight, HPE Nimble Storage support engineers determined that the Fibre Channel recovery mechanism might fail due to a double abort issue within the card.

HPE InfoSight used a signature pattern-matching algorithm to identify 100 customers susceptible to this issue and applied a workaround that prevented the issue.

As demonstrated with **HPE Nimble Storage**, HPE InfoSight automatically predicts and resolves 86% of issues. This translates to a 79% reduction in IT operational expenses, 85% less time spent on storage issues, and over 99.9999% of proven availability across the HPE Nimble Storage installed base.

Recommendation engine

Making infrastructure effortless to manage

To make a university's infrastructure truly autonomous, HPE InfoSight must not only see what's ahead to predict issues, but also dynamically make intelligent recommendations and decisions that improve and optimize each environment proactively. It must be application-aware to serve the right recommendation at the right time—without impacting other applications.

Through its proprietary recommendation engine, HPE InfoSight builds off of its predictive capabilities to automatically tell IT how to prevent issues, proactively improve performance, and optimize resources. The engine advises based on the experience learned from its knowledge base.

Figure 3. Benefits of HPE InfoSight recommendation engine

Many of the more difficult responsibilities in a university's infrastructure management are related to system performance. Next, we'll take a closer look at what the recommendation engine does for performance management.

AI performance recommendations

Without an automated approach, ensuring optimal performance is a time-consuming and costly ordeal for your institution's IT organization:

- Application-impacting problems come up unexpectedly, so this process is reactive.
- The analytics from other tools aren't smart enough to understand why a problem came up, so they can't determine how to resolve it.
- This process requires heavy manual tuning and guesswork, which creates pain and inefficiency in your IT organization.

Through advanced machine learning, HPE InfoSight's recommendation engine removes guesswork and optimizes performance and resources. It:

- Identifies opportunities to improve performance based on I/O workload patterns
- Accurately determines variables that have the highest impact
- Proactively provides the right recommendation to improve performance

Architecting the recommendation engine

Let's look more closely into the recommendation engine: We'll show the design methodology as well as the architecture.

Figure 4, on the next page, represents the problem space of potential infrastructure problems on a bar chart with the type of problems and frequency labeled respectively. Problems generally fall within two categories—those that are **simple and common** (marked in grey), and those that are **complex and unique** (marked in blue)—forming a Pareto distribution. It's important to note the pain curve correlated with the type of problems.

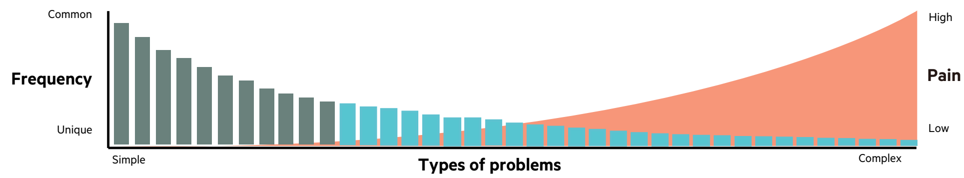


Figure 4. Problem space correlated with the relative pain index

Simple and common problems (for example, failed disk drives) are more frequent but only account for a small percentage of the pain that your IT administrators must handle. The frequency of these problems makes them easier to predict and resolve with an automated solution. But in reality, problems in IT environments can vary widely, so the complex and unique issues—the ones that come up unexpectedly and require numerous people and resources to resolve— cause the most pain and suffering.

The higher education institution needs prediction and auto-resolution for the entire range of problems—from the most basic to the most complex and multifaceted. Simple problems can be identified by looking at only a few pieces of data qualitatively with hard-coded rules to trigger events and alarms. Predictive analytics solutions from other vendors may claim they provide recommendations. But unlike HPE InfoSight, much of their proficiency is generally limited to addressing issues that lie in the simple and common side of the distribution.

For complex and unique problems, the number of variables and the level of quantitative precision required to make a diagnostic determination increases almost exponentially. As problems become more complex, hard-coding rules involving numerous quantitative variables become error-prone and inefficient. Even the most talented human experts struggle with challenges that go beyond the simple threshold behavior for quantitative problems (for example, this problem should trigger when sensor X climbs above threshold Y). And all too often, even those solutions are derived from anecdotal experience rather than rigorous analysis.

HPE InfoSight's recommendation engine goes beyond the simple and common problems to identify and prevent the complex and unique issues. AI and machine learning make it possible to address the long tail and provide recommendations to avoid disruptions to institutional operations and to the daily experience of students and staff.

Design methodology for the AI performance recommendations
Constructing a robust recommendation engine for performance requires answering three key questions.

Question 1: Are performance metrics actually an accurate indicator of an unoptimized system or potential problem?

Sensors collect real-time measurements of their environment with the purpose of detecting events or changes. Typically, IT administrators rely on the value of these sensors (that is, read latency, write latency, IOPS, throughput, and others) to determine if a behavior is problematic. But this approach is flawed: sensors alone lack the full context to determine whether their values are truly indicative of impact on the application and end user experience with a particular system or application.

Different workloads and applications have differing performance characteristics and sensitivity to the end user experience. For instance, large block operations such as backup jobs are naturally more latent—but less response-time sensitive as a transactional workload. Naively assuming higher latency means there are problems resulting in false positives and wasted time chasing down the wrong events—a fundamental problem to event management.

Sensors Alone Lack Context
Universities are increasingly collecting and analyzing student data for a variety of reasons. For instance, a university could collect data from the swipes of student ID cards at locations across campus. The data points are of limited use on their own, but with the right context and analysis, the university could uncover what they reveal about student routines, behaviors, and relationships, potentially identifying opportunities to improve the campus experience for students.

Design approach

Determining how much latency is truly impactful depends on the sensitivity of the underlying application. Leveraging global system telemetry in HPE InfoSight, we have developed machine-learned models of typical performance to more precisely identify events that actually matter to users. We have validated these models using customer case data that reflect the **potential impact** also referred to as the **latency severity score** that can negatively affect performance.

Outcome

As shown in Figure 5, HPE InfoSight understands the true impact of latency and provides a severity index within a defined time frame as color coded in orange and accompanying numerical value (1 to 10). Darker shades of orange indicate higher potential impact to latency.

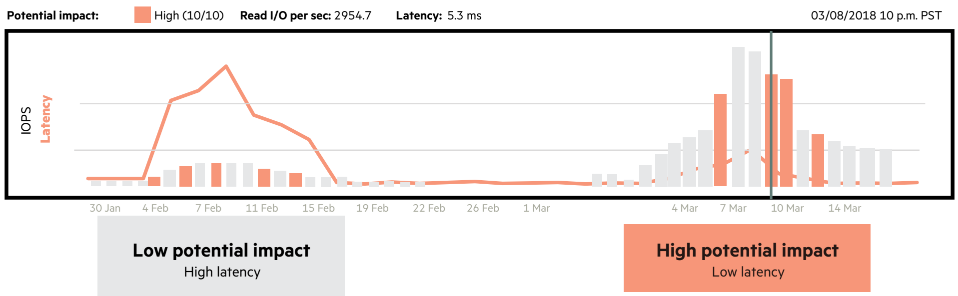


Figure 5. Historical IOPS activity with potential impact reflected in orange color and a numerical latency severity score

This visualization filters the noise and allows the college or university's IT administrators to focus only on the events that matter. The net result is the elimination of false positives and the ability to discern when performance can be improved.

Question 2: Based on workloads running on the system, what factors can be affecting the application performance and to what degree?

Now that we know if and when sensor measurements are indicative of an unoptimized system, we're ready to determine the cause of it.

Traditionally, IT administrators have gone through a series of trial-and-error exercises to resolve a performance problem, hoping to find some way to keep the problem away. But this guesswork is time consuming and often resolves the problem only temporarily—if at all.

Design approach

To ensure that our system can recognize issues across the problem space with high precision, HPE integrates the analysis from the models described in Question 1 with two types of machine-learning models: **expert trained** and **globally trained**. Expert-trained models are trained and validated against specific examples of rare events that have been labeled by HPE's support engineers. Globally trained models are trained and validated against our installed-base telemetry to recognize uncommon problems by looking for expected correlations with latency or when a system is underperforming relative to expectations.

This hybrid approach ensures that HPE InfoSight can address the long tail of complex and unique issues.

Why machine learning?
Machine learning is ideal for problems that require the examination of multiple quantitative variables simultaneously and require signatures that lack a concise qualitative description. Human-curated rules are poorly suited for solving these problems in the same way that it would be implausible for a human to handwrite code that determine if a matrix of pixels matches a specific person's face.

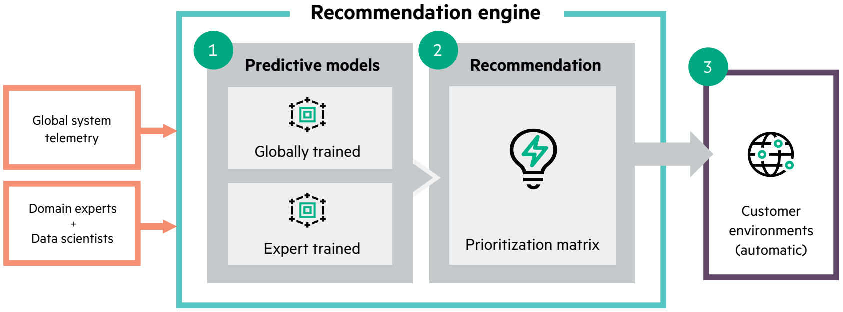


Figure 6. Architecture for the recommendation engine

Multivariate analysis

Expert trained classifiers have been useful, for example, for identifying instances of **SSD bandwidth saturation:** an uncommon event in which high degree of I/O throughput is being directed to SSDs. This scenario is interesting because we have determined that looking at any one SSD metric (for example, latency, queue depth, IOPS, MB/s, proportion of recent milliseconds spent active, and such) is wholly insufficient to accurately determine whether the SSD is causing upstream performance issues. Instead, multiple samples of these metrics need to be examined simultaneously. Looking at one metric alone produces a heuristic that either produces many false positives (low precision) or fails to identify a large proportion of the problematic events (low recall). To produce a model that could simultaneously achieve both high precision and recall: a multivariate machine-learned model was needed.

Because of the quantitative complexity of the problem, our machine-learned classifier was able to recognize this issue much more effectively than any of the human-written heuristics that preceded it.

Expert-trained models

Our expert-trained models are classifiers that use human-labeled instances of a problem that has been observed in the field. Through the support process, our data scientists train the classifiers to recognize new instances of those events in the field without human intervention and with high level of accuracy. Augmenting telemetry with human labels ensures that the system will make correct diagnosis and recommendations in the face of uncommon events.

Globally trained models

Expert-trained models work well for identifying conditions that are discretely true or false but are not well suited for problems that can have multiple root causes co-occurring to varying degrees. If several distinct contributors to latency are detected on a system, it is important to have a consistent way of determining which of these is most responsible for the detected problem. In this situation, it is implausible for a human expert to produce training examples of sufficient quantity. Instead, we train models against our global install-based telemetry to quantify how different sources of latency contribute (often nonlinearly) to produce any observed latency. With this model, we can identify which issues are most important to resolve first. The breadth and richness of our telemetry allows us to generate very comprehensive diagnostic models that would be impossible to train otherwise.

Outcome

Our hybrid machine-learning approach continuously improves the accuracy of the fault monitoring system and its breadth of coverage, minimizing unknown problems. The net result is accurate, root cause diagnostics for every system in our installed base.

Question 3: What is the right recommendation to improve performance?

From the output of Questions 1 and 2, HPE InfoSight can determine whether there is opportunity to improve the institution's environment. The approach to Question 3 results in HPE InfoSight automatically telling the IT administrators what they should do to improve the situation.

Design approach
The easiest (but most inefficient) recommendation would be to advise the need to upgrade hardware: In other words, simply tell an institution's IT administrators that resources are beyond physical limits and larger hardware is required.

This is as far as some vendors go. In contrast, HPE InfoSight provides a much richer set of recommendations. These include but are not limited to QoS limits, software updates, workload changes, configuration changes, and hardware upgrades. HPE InfoSight is aware of the applications, resources, and preferences (for example, time of day and days of week critical, sensitivity to latency) for every system. And it uses this understanding to prioritize the recommendations.

Details provided with the recommendation inform those portions of the IT organization's workload consuming a saturated resource (for example, the volumes using the most storage CPU when the array is CPU bound). These details are critical because IT needs them to decide between workload-based remediation (throttling volume activity or otherwise attenuating the volumes' requirements) or hardware remediation (adding hardware to the system to extend the system's capabilities and alleviate the resource bottleneck).

Outcome
Before HPE InfoSight was available, IT administrators had to deal with the pain of manually and reactively managing storage performance. HPE InfoSight's recommendation engine has removed this worry. HPE InfoSight informs IT when there's an opportunity to improve performance and tells them what they need to do. Confident that their systems are running optimally, your IT administrators can run their storage systems hard and consolidate multiple applications—without worrying about the threat of applications slowing down due to infrastructure problems. This frees up more IT resources to support strategic institutional objectives. Further, colleges and universities can use HPE InfoSight to access to rich sets of AI-generated insights that make it possible to realize the true promise of AI—enhancing the quality of the educational experience for all.

In summary, the recommendations generated by HPE InfoSight are:

- **Automatic:** Available at any time to every authorized institutional user
- **Preemptive:** Providing foresight to bottlenecks before they can impact institutional operations
- **Extensive:** Using machine learning to predict the long tail of complex and unique problems
- **Prescriptive:** Reaching beyond hardware upgrades and including specific operational changes

See how colleges, universities, and other public institutions are becoming smarter with an intelligent architecture that includes HPE InfoSight: **hpe.com/info/public-sector.**

A lesson plan for the predictive data center

It's time to transform your data center into an asset that drives your college or university's success. HPE's vision is a data center that no longer needs constant attention, manual tuning, or reactive troubleshooting. This means a data center whose infrastructure truly self-manages, self-heals, and self-optimizes—and that can provide you with the competitive advantage of access to sophisticated, institutional-wide data and analytics that help drive better decisions. Colleges, universities, and other public institutions with infrastructure powered by AI from HPE InfoSight are coming increasingly close to making this vision a reality.

As the industry's most experienced provider of artificial intelligence, HPE InfoSight has fundamentally changed how infrastructure is managed and supported. Through cloud-based machine learning, it predicts and prevents problems while delivering the optimal performance and availability from the infrastructure it supports. And with almost a decade of experience and learning, HPE InfoSight continues to grow more sophisticated and proficient.

Regardless of where your institution is in the digital transformation journey, it's important to tailor the right solution to meet your needs based on your institution's individual charter and objectives. From strategic advising to purpose-built solutions, HPE Infosight and HPE's deep learning expertise can go to work to help you create your ideal, end-to-end AI solution—helping to enable an agile IT foundation, a smarter university, and a campus experience that's safe, efficient, innovative, and fun.

Learn more at **hpe.com/us/en/storage/infosight.html**

Make the right purchase decision. Click here to chat with our presales specialists.

Share now

Get updates

a00059047enw, December 2018