Whitepaper

NetApp Cloud Insights: Monitoring für Ihre Cloud-Infrastruktur

Schnellere Innovation dank Einblick in den gesamten Applikationsinfrastruktur-Stack

Cliff Oberholtzer, NetApp

Juli 2023 | WP-7362

Zusammenfassung

Dieser Leitfaden erleichtert Ihnen den Einstieg in NetApp Cloud Insights und zeigt Ihnen in Kurzform, wie Sie Cloud Insights für die Überwachung, Fehlerbehebung und Optimierung Ihrer gesamten Infrastruktur einrichten und erste Schritte gehen.

INHALTSVERZEICHNIS

ABBILDUNGSVERZEICHNIS

Die Anforderungen beim Cloud-Betrieb

Ihre Infrastruktur wird zunehmend komplexer. Unternehmen schätzen die Flexibilität und die schnelle Bereitstellung von Tools und Services aus der Cloud. Allerdings handelt es sich in der Regel nicht nur um einen Cloud-Service, sondern um mehrere Clouds und eine zusätzlich vorhandene On-Premises-Infrastruktur. Trotz aller Komplexität sollen Sie aus immer weniger Ressourcen immer mehr herausholen. Die Cloud hat zwar die Bereitstellung vereinfacht und beschleunigt, dafür ist es umso schwerer geworden, die Kosten unter Kontrolle zu halten und die Auslastung zu optimieren. Als zuständige Abteilung halten Sie die Performance und Ausfallsicherheit am Laufen und werden bei Problemen zuallererst in die Verantwortung genommen.

Ihre Applikationsentwicklungs- und Implementierungsteams setzen immer häufiger auf die Vorteile neuer innovativer Cloud-Technologien, um Produkte schneller auf den Markt zu bringen und die Kundenbindung zu stärken. Den daraus resultierenden erhöhten Anforderungen an die Infrastruk-tur-Monitoring- und -Optimierungs-Tools nachzukommen, ist nicht so leicht. So erfordern neue Technologien wie Kubernetes, die Microservices und Distributed Tracing nutzen, neue Monitoring-Möglichkeiten. Diese Kluft zwischen Implementierungstechnologien und Monitoring-Tools kann den Ops-Teil Ihrer DevOps-Projekte einem erhöhten Risiko von Ausfällen und ausufernden Kosten aussetzen.

Als Site Reliability Engineer sind Sie Teil eines DevOps-Teams mit dem Auftrag zur Entwicklung neuer Cloud-Applikationen. Sie sind das einzige Bindeglied, das zwischen den Entwicklungs- und Betriebsteams vermittelt. Ihre Aufgabe ist es, den Betrieb effizienter zu gestalten und die Performance zu verbessern. Hierfür müssen Sie mögliche Risiken bei der Entwicklung und Implementierung dieser modernen Applikationen identifizieren und managen. Jeder Applikationsausfall wirkt sich nachteilig auf die Geschäfte Ihres Unternehmens aus.

Sie benötigen Echtzeit-Service-Level-Indikatoren (SLIs), um sicherzustellen, dass Ihre Systeme Ihre Service Level Objectives (SLOs) und Service Level Agreements (SLAs) erfüllen. Ihre Tools müssen sowohl Ihre Cloud- als auch Ihre On-Premises-Infrastruktur gleichzeitig überwachen können. Bei Infrastrukturausfällen müssen Sie das ursächliche Problem schnell identifizieren. Anders ausgedrückt: Sobald etwas schiefläuft, sind alle Augen auf Sie gerichtet.

Ergo: Sie brauchen ein einfach und intuitiv zu bedienendes Cloud-basiertes Infrastruktur-Monitoring-Tool, das die Fehlerbehebung beschleunigt, den Performance-Bedarf genau vorhersagt und Ihnen hilft, Kosten zu sparen. NetApp Cloud Insights wurde für genau diese Anforderungen entwickelt.

Cloud Insights ist ein SaaS-Monitoring-Tool (Software-as-a-Service), das Ihnen nützliche Einblicke in Ihre Infrastruktur gewährt. Cloud Insights ist äußerst benutzerfreundlich und durch das Hosting in der Cloud schnell einsatzbereit. Es ermöglicht die Echtzeit-Visualisierung der Verfügbarkeit, Performance und Auslastung Ihrer gesamten IT-Infrastruktur.

Dieser Leitfaden beleuchtet, welche Schwierigkeiten beim Monitoring von Cloud-Infrastrukturen bestehen und wie Sie mit Cloud Insights Zeit und Geld sparen können.

Wir zeigen Ihnen Schritt für Schritt, wie Sie:

- Daten zu Ihrer Infrastruktur sammeln
- Ihre Infrastruktur effektiv überwachen
- Fehlerbehebung betreiben, um Probleme zu finden und lösen
- Kosten optimieren und die betriebliche Effizienz steigern

Die wichtigen Funktionen von Cloud-Monitoring-Tools

Cloud-Applikationen sind so programmiert, dass sie die Infrastruktur als Code betrachten, d. h. sie provisionieren und deprovisionieren Infrastruktur dynamisch über APIs. Applikationen müssen daher während ihrer Ausführung den Zustand dieser Infrastrukturen kennen, um Anpassungen in Echtzeit vorzunehmen. Viele Cloud-Applikationen machen ausgiebig Gebrauch von Cloud-basierten Provisionierungs- und Kontroll-Services wie Puppet, Chef, Container und Kubernetes, um die von

ihnen verwendete Infrastruktur mit der Geschwindigkeit und Skalierbarkeit der Cloud zu erweitern oder zu schrumpfen.

Aufgrund dieser hohen Geschwindigkeit und Skalierbarkeit müssen Cloud-Monitoring-Tools Daten in Millisekunden erfassen können, nicht erst in Minuten oder gar Stunden. Zusätzlich müssen sie nicht nur den Zustand jeder Komponente, sondern auch die Beziehungen zwischen den Komponenten kennen.

Monitoring ist das forensische Mittel, um kurze Latenzspitzen oder flüchtige Ausfälle zu erkennen, die nur für einen kurzen Augenblick sichtbar sind. In einer dynamischen Infrastruktur mit Selbstreparatur-Eigenschaften können nicht behobene Probleme zu Überprovisionierung sowie unnötig hohen Kosten führen und im schlimmsten Fall Ihre Kunden und damit Ihr Geschäft beeinträchtigen.

NetApp Cloud Insights

NetApp Cloud Insights ist speziell auf Cloud-basierte Infrastrukturen und Bereitstellungstechnologien von heute ausgelegt. Mit seinen erweiterten Analytikfunktionen zu den Verbindungen zwischen Ressourcen innerhalb Ihrer Umgebung ermöglicht es Echtzeit-Visualisierung der Topologie, Verfügbarkeit, Performance und Auslastung Ihrer gesamten Infrastruktur, was Ihre Cloud- und Ihre On-Premises-Ressourcen gleichermaßen einschließt.

In Cloud Insights lassen sich unterschiedliche Infrastrukturschichten einbinden – herkömmliche servicebasierte und moderne softwarebasierte –, was Ihnen gleichermaßen Transparenz über Ihre herkömmlichen und Ihre modernen Applikationsarchitekturen erlaubt. In kurzer Zeit inventarisiert Cloud Insights Ihre Ressourcen, ermittelt gegenseitige Abhängigkeiten und erstellt eine Topologie Ihrer Umgebung. Sie erhalten End-to-End-Einblick in Ihre Infrastruktur und sehen genau, welche Ressourcen welche Applikationen unterstützen.

Cloud Insights wurde speziell für die transienten Cloud-Infrastrukturen von heute und ihre Verbindungen zu vielfältigen Services entwickelt und hilft Ihnen, die Nachfrage, Latenz, Fehler und Auslastungsgrenzen all Ihrer Services zu verstehen.

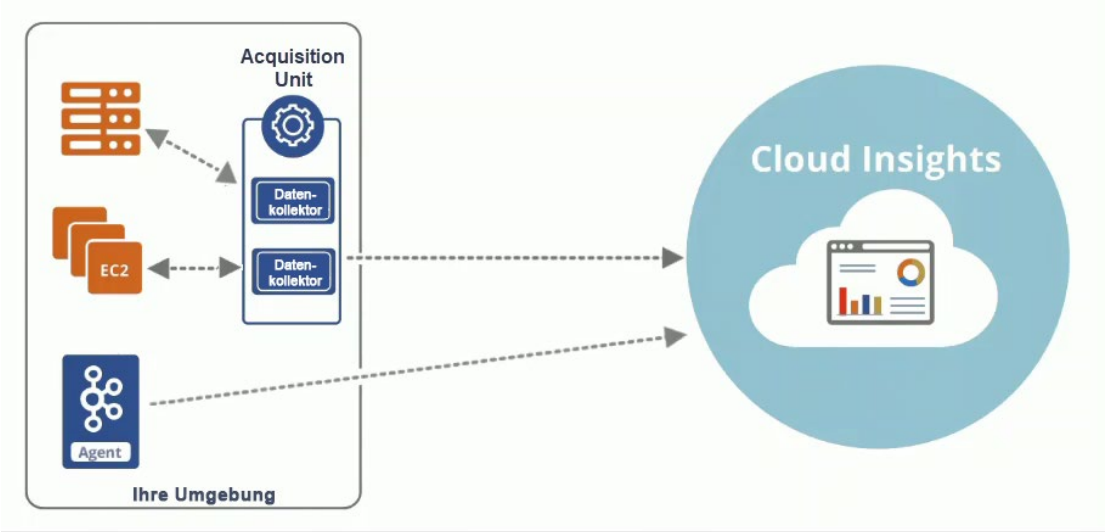
Mit Cloud Insights können Sie:

- **Die Kundenzufriedenheit steigern**, indem Sie bis zu 80 % aller Cloud-Infrastruktur-Probleme lösen, bevor sie von Anwendern bemerkt werden. Das proaktive Monitoring Ihrer gesamten Umgebung hilft Ihnen, die Anforderungen Ihrer Kunden besser zu erfüllen. Die automatische Erkennung visualisiert Ihre Topologie und stellt sämtliche Servicepfade transparent dar. So wissen Sie ganz genau, wie Ihre Systeme sich verhalten und wie sie genutzt werden. Sollte eine Performance-Level Violation auftreten, erhalten Sie alle erforderlichen Daten, um die Ursache so schnell wie möglich zu ermitteln. Diese Analyse hilft Ihnen, mit den Anforderungen Ihrer Kunden schrittzuhalten.
- **Ausfälle proaktiv vermeiden** und die Zeitspanne bis zur Problemlösung (MTTR) um bis zu 90 % verkürzen. Die erweiterte Analyse zeigt, welche Ressourcen übergriffig oder degradiert, also in der Leistung vermindert, sind. Mittels Korrelationsanalyse können Sie Services den von ihnen genutzten modernen, transienten Infrastrukturen zuordnen, um den Ursachen von Problemen schneller auf den Grund zu gehen. Sie können auch erweiterte bedingte Warnmeldungen konfigurieren, um die Anzahl der Falschmeldungen zu verringern. Moderne Machine-Learning-Algorithmen warnen Sie zudem vor möglichen Fehlern und empfehlen Lösungen, um zu verhindern, dass sie zu ausgewachsenen Problemen führen.
- **Cloud-Infrastruktur-Kosten optimieren und senken**, um durchschnittliche Einsparungen in Höhe von 33 % zu erzielen. Wenn unterschiedlichste Ressourcen bereitgestellt werden – im On-Premises-Datacenter bis hin zu mehreren Public Clouds –, lässt sich nur schwer sagen, welche Ressourcen davon wirklich genutzt werden und welche freigegeben werden können. Hierzu müssen Sie ungenutzte und verwaiste Ressourcen identifizieren können. Es ist außerdem hilfreich, die Performance-Anforderungen Ihrer Applikationen zu kennen, um feststellen zu können, auf welchen Ressourcen sie überprovisioniert sind. Mit diesem Wissen können Sie diese Applikationen auf günstigeren Infrastrukturen neu provisionieren.

1. Schritt: Ihre Daten erfassen und inventarisieren

Sobald Sie Cloud Insights gestartet und sich angemeldet haben, besteht der erste Schritt darin, die Infrastruktur zu erfassen, die Sie überwachen wollen. Hierzu müssen Sie zunächst eine Acquisition Unit installieren. Abbildung 1 zeigt die Rolle der Acquisition Unit innerhalb Ihrer Infrastrukturtopologie.

Abbildung 1) Acquisition Unit und Infrastrukturtopologie

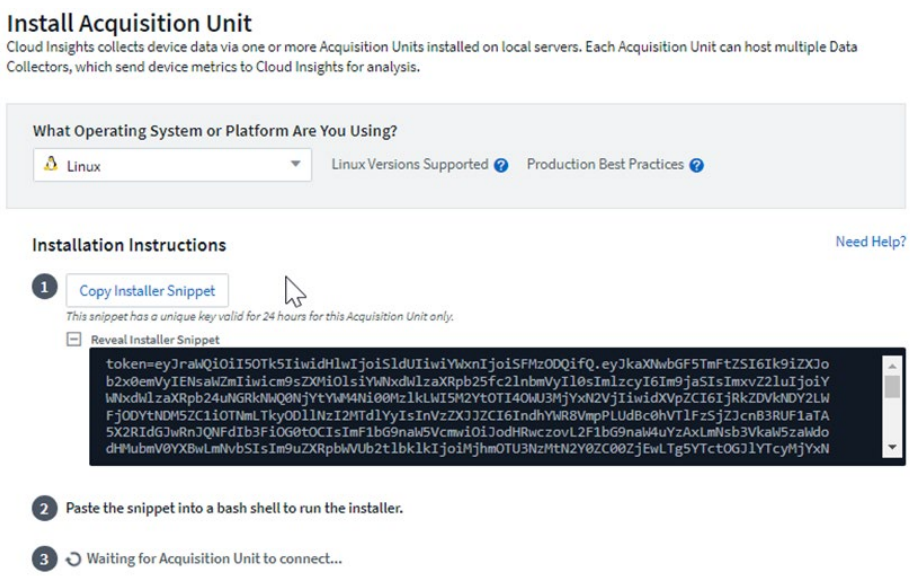


Die Acquisition Unit wird in Ihrer Infrastruktur bereitgestellt und befindet sich hinter Ihrer Firewall, Ihrer Virtual Private Cloud oder Ihrem VNet. Die Acquisition Unit wird auf einer Virtual Machine installiert, die durch die Firewall hindurch erreichbar ist und dafür sorgt, dass Daten aus der Cloud heraus und in die Cloud hinein übertragen werden können.

Cloud Insights unterstützt Acquisition Units für Windows und Linux. Die Acquisition Unit benötigt eine Virtual Machine mit lediglich 4 Kernen, 16 GB RAM und 4 GB Festplattenspeicher.

Für Linux müssen Sie nur den bereitgestellten Text in die Konsole kopieren (siehe Abbildung 2). Wenn Sie einen Proxyserver nutzen, müssen Sie auch die zusätzlich angezeigte Zeile mit einfügen, um die Proxyvariablen zu setzen.

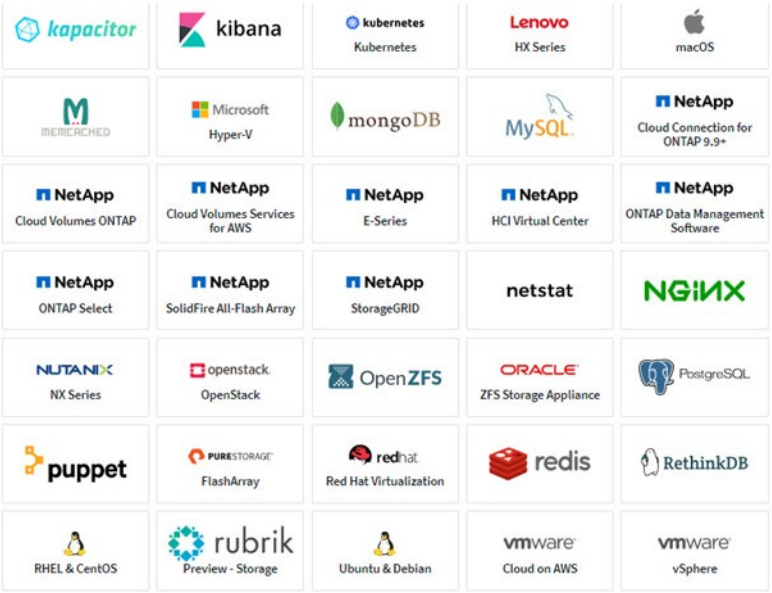
Abbildung 2) Einrichtung der Acquisition Unit



Für Windows müssen Sie den Microsoft Installer herunterladen und ausführen, um die Acquisition Unit einzurichten.

Mit Cloud Insights können Sie nicht nur NetApp Systeme überwachen. Es bietet auch volle Unterstützung für die Public Clouds der Anbieter AWS, Azure und Google Cloud sowie für Geräte von Drittanbietern wie Pure, Dell EMC, Fujitsu, IBM, Hitachi, Broadcom, Cisco und mehr. Abbildung 3 zeigt eine Auswahl der unterstützten Optionen.

Abbildung 3) Auswahl an verfügbaren Kollektoren



Die Einrichtung eines Kollektors ist sehr einfach: Sie geben dem Kollektor einen Namen innerhalb Ihrer Umgebung und wählen die zuvor eingerichtete Acquisition Unit aus. Im Anschluss geben Sie die für diesen Kollektor erforderlichen Zusatzinformationen ein. Abbildung 4 zeigt die Einrichtung eines AWS-Kollektors. In diesem Fall müssen Sie die Region, die Zugriffsschlüssel-ID und den geheimen Zugriffsschlüssel eingeben. Die Acquisition Unit hat nur Lese-Zugriff, damit Cloud Insights Daten der Ressource erfassen kann. Zu diesen Daten gehören Konfigurationsdetails, bestehende Verbindungen dieser Ressource zu anderen Ressourcen, Anbieter, Kapazität, Speicherauslastung und mehr. Cloud Insights nutzt diese Daten für wichtige Analysen, die Ihnen helfen, Ihre gesamte Umgebung zu überwachen, Fehler zu beheben und das Beste aus ihr herauszuholen.

Abbildung 4) Dialogfeld des AWS-Kollektors

2. Schritt: Überwachen und Probleme ausfindig machen

Sobald Daten ins System fließen, können Sie nach Problemen Ausschau halten. Cloud Insights stellt Ihnen eine Galerie mit Dashboards zur Auswahl. Die anfänglich angezeigte Galerie basiert auf den von Ihnen eingerichteten Kollektoren. Wenn Sie beispielsweise Ihre AWS-Umgebung erfassen wollen, wird Ihnen eine Auswahl an Berichten angezeigt, die Sie Ihrem personalisierten Dashboard hinzufügen können. Ihre Dashboards geben Ihnen Einsicht in Ihre Daten und ermöglichen Ihnen so, nach Problemen zu suchen.

Beim Monitoring sind vier SLI-Signale von höchster Relevanz:

- **Latenz** – Latenzspitzen beeinträchtigen Ihre Anwender. Dies führt zu Fehlertickets und erfordert aufwändige Korrekturmaßnahmen. Vermeiden Sie dies, indem Sie Latenzprobleme erkennen und beheben, bevor Ihre Anwender sie bemerken.
- **Sättigung/Auslastung** – Latenzprobleme entstehen häufig durch Überschreiten der Auslastungsgrenze der Geräte, die Ihre Workloads unterstützen.
  Überschreiten Ressourcen Ihre Performance-Grenzen? Speicherauslastung? CPU-Auslastung?

Festplattenkapazität? In diesen Fällen beeinflusst die Sättigung die Latenz. Zu wissen, was die Sättigung verursacht, ist wichtig, um das grundlegende Problem zu erkennen.

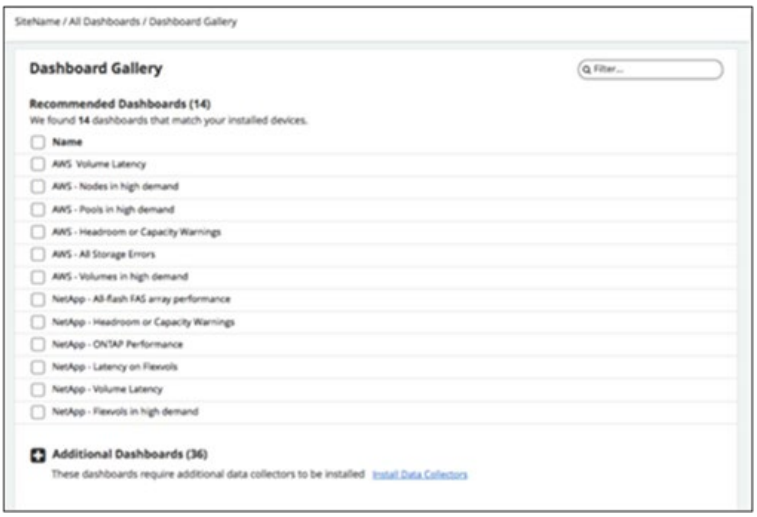
- **Traffic** – Erhöhter Netzwerkverkehr ist eine der häufigsten Ursachen für das Auftreten von Sättigung. Ein Anstieg der IOPS oder der Megabyte pro Sekunde kann dazu führen, dass bestimmte Geräte ihre Auslastungsgrenze überschreiten. Als für den Betrieb Zuständige müssen Sie wissen, wie die Werte für Latenz, Sättigung und Traffic aussehen. Je schneller Sie bemerken, dass SLI-Schwellenwerte überschritten werden, desto schneller können Sie notwendige Maßnahmen ergreifen, um Fehlertickets und Beschwerden zu vermeiden.
- **Fehler** – Zuguterletzt müssen Sie darüber informiert sein, ob in Ihrer Umgebung Fehler auftreten, und Sie müssen die Ursache für diese Fehler identifizieren können, um sie schnellst-möglich zu beheben.

Dashboards

Cloud Insights erstellt eine Standardauswahl an Dashboards anhand der von Ihnen gewählten Kollektoren. Jedes Dashboard ist so aufgesetzt, dass es die für Sie relevanten Fragen bei der Überwachung Ihrer Infrastruktur beantwortet.

- Bei welchen Systemen ist die Latenz sehr hoch?
- Wo wurden SLOs überschritten, wodurch es in Folge zu Fehlern kam?
- Welche VMs sind inaktiv oder heruntergefahren?

Abbildung 5) Beispiel für eine Dashboard-Galerie

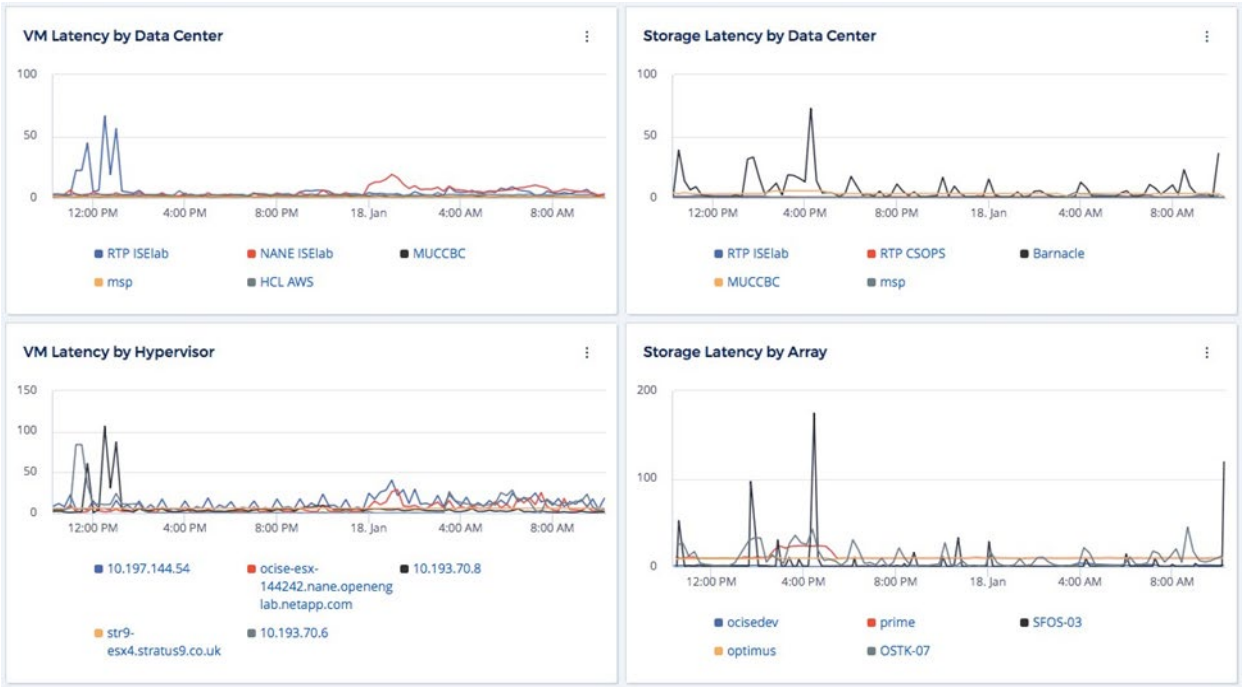


Ein typisches Monitoring-Szenario ist zum Beispiel die Fragestellung, welche Systeme unter hoher Latenz leiden, die ohne Korrektur zu verminderter Performance führt und ihre Anwender beeinträchtigt. Abbildung 6 zeigt ein Dashboard, das sämtliche Systeme innerhalb der gesamten Infrastruktur nach Latenz sortiert anzeigt. In diesem Beispiel gibt es mehrere Datacenter mit Ressourcen in AWS, Azure und Google Cloud sowie auf On-Premises-Systemen.

In einem solchen Dashboard können Sie sich einfach bis zu den einzelnen Ressourcen durchklicken, um sich anzusehen, welche davon für die hohe Latenz verantwortlich ist. Die Correlation Engine von Cloud Insights zeigt an, welche Ressourcen mit einer Latenzspitze in Bezug stehen und mit welcher Wahrscheinlichkeit sie beeinträchtigt werden.

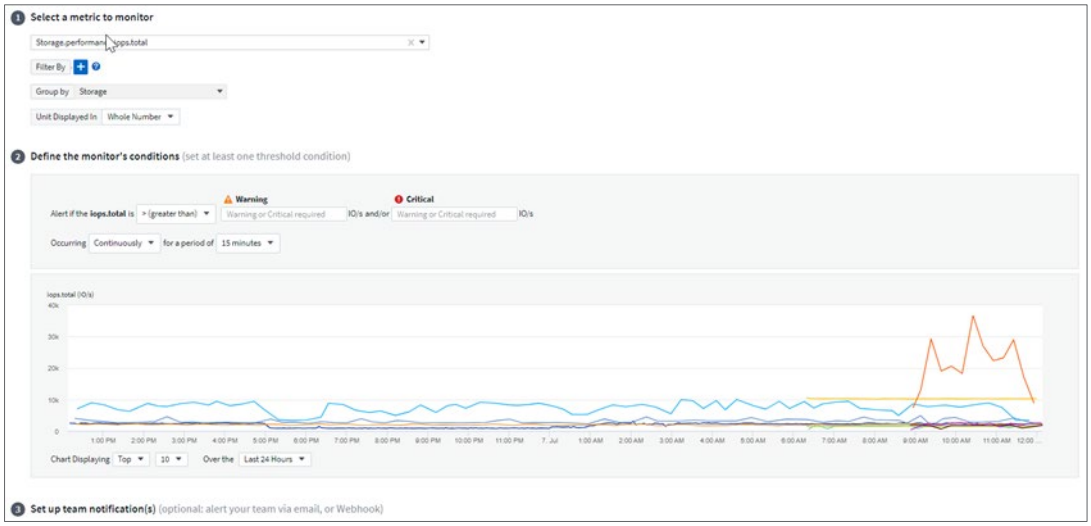
Wird eine beeinträchtigte Ressource gefunden, können Sie in der Topologie-Ansicht sehen, wie diese mit anderen Ressourcen verbunden ist.

Abbildung 6) Latenz-Dashboard nach Datacenter



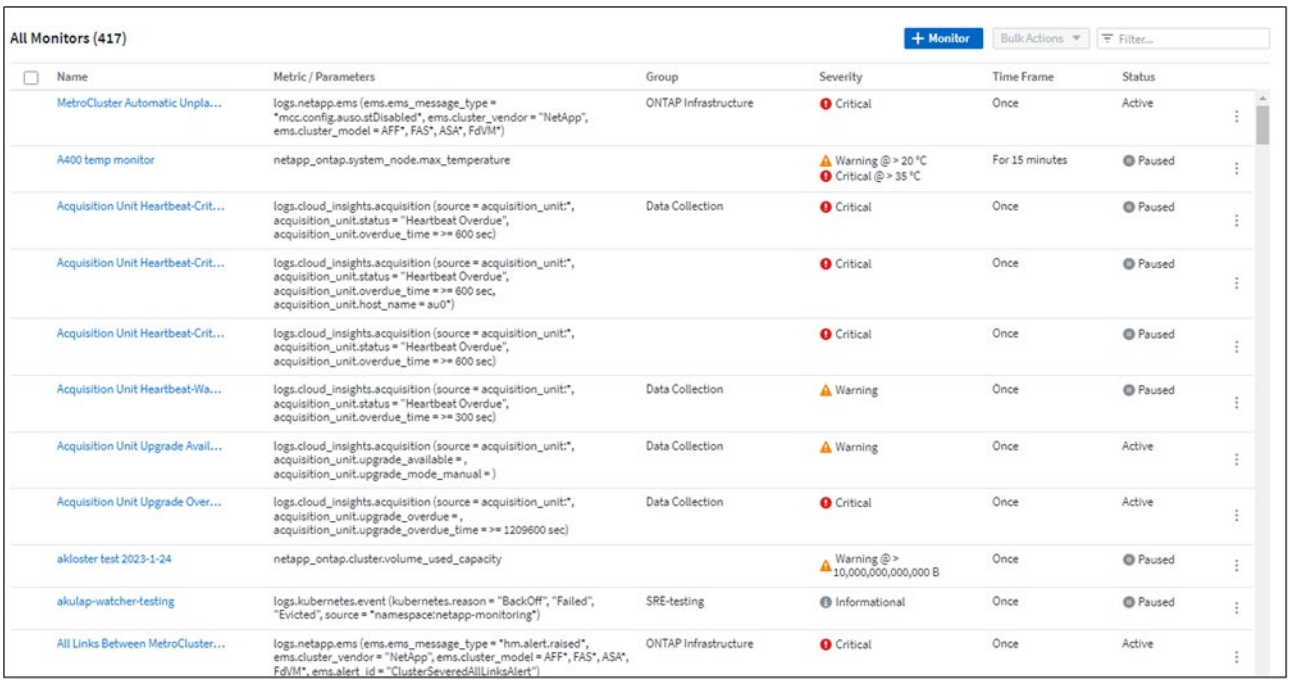
Cloud Insights ermöglicht die Einrichtung von Monitoring-Richtlinien, die Sie alarmieren, wenn eine Ressource ein bestimmtes Service Level oder einen festgelegten Schwellenwert überschreitet. Sie können Ihre Umgebung in Echtzeit überwachen und Probleme erkennen, bevor sie Ihre SLAs im eigentlichen Betrieb beeinträchtigen. Abbildung 7 zeigt, wie leicht Sie einen sogenannten Monitor zur Überwachung bestimmter Kriterien erstellen können. Hierzu geben Sie im Dialogfeld zum Hinzufügen eines Monitors einen Namen und die zu überwachenden Kriterien ein. Ihnen stehen zahlreiche Optionen zur Auswahl, darunter Datastores, VMDKs, Hypervisoren, Volumes und Virtual Machines. Sie können zudem den Typ des Alarms sowie die zu überwachenden Ressourcen festlegen. Alarmmonitore erfordern das Festlegen einer Ressource, einer Kennzahl und eines dazugehörigen Schwellenwerts oder Logs, wodurch Sie flexibel einstellen können, welche Bedingung genau den Alarm auslösen soll.

Abbildung 7) Dialogfeld zum Hinzufügen eines Monitors zur Überwachung nach Kriterien



Löst ein Monitor einen Alarm aus, erhalten Sie eine E-Mail mit einem Link, der Sie direkt zur Liste der Alarme führt, bei denen die Schwellenwerte überschritten wurden. Abbildung 8 zeigt beispielhaft eine Liste von Alarmen, die von überwachten Kennzahlen durch Überschreiten der festgelegten Schwellenwerte ausgelöst wurden. Sie können die Alarme in dieser Liste filtern und per Drill-Down die Ursache bestimmter Überschreitungen ermitteln.

Abbildung 8) Gemeldete Alarme



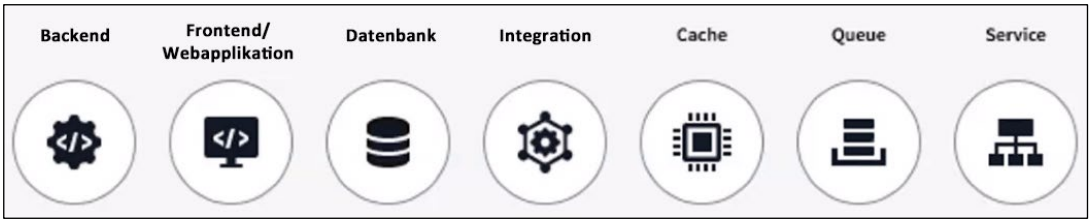
3. Schritt: Fehler schnell finden und beheben

Um Fehler effektiv beheben zu können, müssen Sie wissen, wie Ihre Ressourcen miteinander verbunden sind und wie sie interagieren. Die Abbildungen 9 bis 11 zeigen verschiedene Symbole, die von Cloud Insights genutzt werden, um einzelne Elemente Ihrer Infrastruktur darzustellen.

Abbildung 9) Infrastruktur-Symbole

Abbildung 10) Kubernetes-Symbole

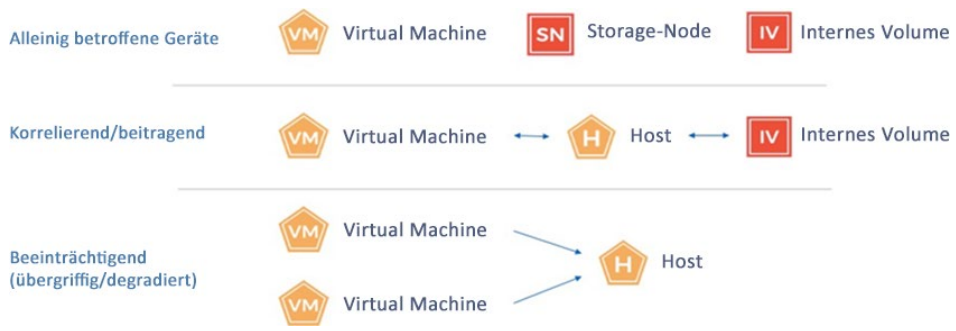
Abbildung 11) Symbole für Performance-Monitoring und Zuordnungen in Kubernetes-Netzwerken



Zur Behebung von Fehlern innerhalb Ihrer Infrastruktur suchen Sie in der Regel nach einem bestimmten Szenario, das Ihnen hilft, die Ursache des Problems zu erkennen. Abbildung 12 zeigt gängige Probleme:

- **Alleinig betroffene Geräte** (Individual device issues) – Das Problem wird möglicherweise durch eine einzelne Virtual Machine, einen Storage-Node oder ein Volume verursacht. In diesen Fällen sind verbundene Ressourcen nicht betroffen. So kann zum Beispiel eine einzelne VM Latenzüberschreitungen aufweisen, die nicht durch andere Ressourcen, sondern durch Programmierfehler (z. B. ein Memory Leak) verursacht werden.
- **Korrelierend oder beitragend** (Correlating / Contributing) – In diesem Fall betrifft das Problem einer Ressource weitere verbundene Ressourcen. Cloud Insights zeigt Ihnen sämtliche Zuordnungen zwischen den Ressourcen, damit Sie ermitteln können, ob der Ausfall einer Ressource andere verbundene Ressourcen beeinträchtigt.
- **Übergriffige Ressourcen** (Impacting: Greedy / Degraded) – Der dritte Fall in Abbildung 12 zeigt eine VM mit einem Problem, das eine andere VM innerhalb des Stacks beeinträchtigt. Cloud Insights erkennt diese Korrelationen und zeigt Ihnen, welche Ressourcen am wahrscheinlichsten von einer übergriffigen oder degradierten Ressource beeinträchtigt werden.

Abbildung 12) Mögliche Problemquellen im Infrastruktur-Stack



Wird Ihnen ein Problem gemeldet, klicken Sie einfach auf den Link in der Alarm-E-Mail und navigieren per Drill-Down zur Performance Violation. Abbildung 13 zeigt beispielhaft eine Violation, bei der eine VM um 4:00 Uhr morgens eine Latenzspitze aufwies. Sie sehen auch die anderen mit ihr verbundenen Ressourcen. Um deren Kennzahlen einzusehen, aktivieren Sie die zugehörigen Kontrollkästchen der verbundenen Ressourcen. In diesem Beispiel beeinträchtigt der Host „ocise-esx“ mit einer Wahrscheinlichkeit von 87 % die VM „bschoferW2k12S4“.

Die Kenntnis dieser Korrelationen hilft Ihnen, schnell zu verstehen, wo genau das Problem liegt und inwieweit es andere Ressourcen innerhalb Ihrer Infrastruktur beeinträchtigt.

Abbildung 13) Ursachenanalyse

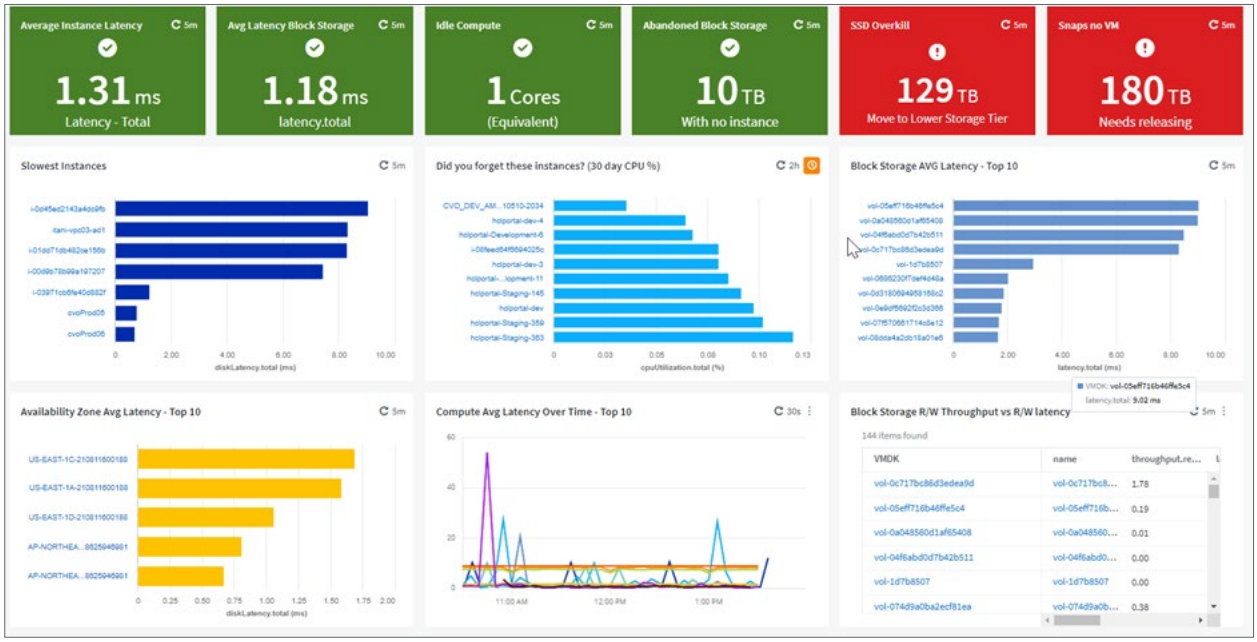
4. Schritt: Kosten optimieren und die betriebliche Effizienz steigern

Abseits der Überwachung Ihrer Umgebung und Fehlerbehebung wollen Sie mit Sicherheit auch Überprovisionierung vermeiden, um nicht mehr Geld auszugeben als nötig. Gerade in der Cloud kommt es häufig vor, dass Ressourcen provisioniert und dann vergessen werden.

Cloud Insights macht es Ihnen leicht, Kosten schnell einzusparen. Cloud Insights kennt den Zustand aller Virtual Machines in Ihrer Umgebung und kann Ihnen so schnell anzeigen, welche sich im Ruhezustand befinden und wie viel Kapazität von ihnen belegt wird. Sie haben dann die Möglichkeit, die Kosten dieser VMs ermitteln, um einschätzen zu können, wie viel Geld Sie durch ihre Deprovisionierung einsparen könnten.

Abbildung 14 zeigt einen Bericht zu allen heruntergefahrenen und pausierten VMs inklusive der von ihnen verursachten Kosten an. Sie können zu jeder einzelnen VM per Drill-Down Details anzeigen, um zu erfahren, wem sie zugewiesen und wie sie mit anderen Ressourcen innerhalb der Infrastruktur verbunden ist. Diese Informationen helfen Ihnen zu ermitteln, welche Ressourcen Sie freigeben können.

Abbildung 14) Einsparpotenzial ermitteln



Zusammenfassung und nächste Schritte

Cloud Insights hilft Ihnen bei der Überwachung, Fehlerbehebung, Optimierung und Sicherung Ihrer Infrastruktur.

Sie können Cloud Insights 30 Tage lang kostenlos testen und in Aktion erleben. Die Möglichkeit zur Registrierung und weitere Informationen finden Sie auf unserer Website zu Cloud Insights unter netapp.de. Folgen Sie dort einfach dem Link oder registrieren Sie sich hier für Ihre kostenlose Testphase.

NetApp