Monitoring in hybriden IT-Umgebungen

Whitepaper

www.bitkom.org

bitkom

Herausgeber

Bitkom e. V.
Albrechtstraße 10 | 10117 Berlin

Projektleitung

Nick Kriegeskotte | Leiter Infrastruktur & Regulierung
+49 30 27576-224 | n.kriegeskotte@bitkom.org

Autoren:

Thore Bahr (SUSE)
Roman Bansen
Frank Beckereit (NTT)
Martin Beuse (HPE)
Stefan Böhler (Microsoft)
Axel Frentzen (NetApp)
Andreas Jahnke (DATEV)
Michael Kambeck (SVA)
Sven Kaminski (SVA)
Fred Klein (IBM)
Andreas Landenberger (DATEV)
Jürgen Lang (IBM)
Holger Nicolay (Interxion)
Rui Manuel Tavares (Fujitsu)

Satz & Layout

Anna Stolz | Bitkom e. V.

Titelbild

© Benjamin Suter – www.pexels.com

Bitkom vertritt mehr als 2.000 Mitgliedsunternehmen aus der digitalen Wirtschaft. Sie erzielen allein mit IT- und Telekommunikationsleistungen jährlich Umsätze von 190 Milliarden Euro, darunter Exporte in Höhe von 50 Milliarden Euro. Die Bitkom-Mitglieder beschäftigen in Deutschland mehr als 2 Millionen Mitarbeiterinnen und Mitarbeiter. Zu den Mitgliedern zählen mehr als 1.000 Mittelständler, über 500 Startups und nahezu alle Global Player. Sie bieten Software, IT-Services, Telekommunikations- oder Internetdienste an, stellen Geräte und Bauteile her, sind im Bereich der digitalen Medien tätig oder in anderer Weise Teil der digitalen Wirtschaft. 80 Prozent der Unternehmen haben ihren Hauptsitz in Deutschland, jeweils 8 Prozent kommen aus Europa und den USA, 4 Prozent aus anderen Regionen. Bitkom fördert und treibt die digitale Transformation der deutschen Wirtschaft und setzt sich für eine breite gesellschaftliche Teilhabe an den digitalen Entwicklungen ein. Ziel ist es, Deutschland zu einem weltweit führenden Digitalstandort zu machen.

Inhaltsverzeichnis

Abbildungsverzeichnis

Über dieses Dokument

Mit dem ersten **↗ Whitepaper, welches der AK Hybride-IT in 2021 veröffentlicht hat,** wurde anhand eines detaillierten Schaubildes erklärt, was der Begriff »Hybride-IT« bedeutet. Zudem wurden die unterschiedlichen Aspekte verdeutlicht, die bei der Auswahl und Planung von Anwendungen und Services für unterschiedliche Use-cases zu beachten sind. Das vorliegende aktuelle Whitepaper zeigt nun speziell die Besonderheiten des Monitorings in hybriden IT-Umgebungen auf.

Das nun vorliegende Whitepaper geht speziell auf die Unterschiede und auch Gemeinsamkeiten bei verschiedenen Monitoring Dimensionen im Fall einer hybriden IT ein. Generell lässt sich sagen, dass alle Monitoring Aktivitäten (Performance, Verfügbarkeit, Kapazität, Compliance, Security, Kosten, Anomalien und Betrieb) auch in klassischen IT-Umgebungen vorhanden sein sollten. In einer hybriden Umgebung kommen aber zusätzliche Anforderungen und auch Herausforderung hinzu.

Im Kapitel **↗ Performance-Monitoring** ist eine der Herausforderungen die Key Performance Indicators aus den verschiedenen Bereichen einer hybriden IT in einheitlicher Art zu monitoren und zu berichten.

Bei den **↗ SLAs Verfügbarkeiten** ist einer der wesentliche Faktoren, dass sich die Gesamtverfügbarkeit in einer hybriden Umgebung im Regelfall aus dem Produkt der Einzel-Verfügbarkeiten von deutlich mehr/granularen Services ergibt.

↗ Kapazitäts-Monitoring muss in einer hybriden IT-Umgebung damit umgehen, dass ein Teil der IT-Ressourcen quasi unlimitiert (aus der Cloud), andere dagegen nur limitiert (On-Premises/ Hosting) zur Verfügung stehen.

Im **↗ Compliance-Monitoring** ist zu berücksichtigen, dass es zu verteilten Verantwortlichkeiten und damit zu einem Kontrollverlust des Unternehmens kommt. Dies muss durch ein einheitliches Set von Compliance-Kontrollpunkten und ein automatisiertes Monitoring mitigiert werden.

Auch dem **↗ Security-Monitoring** kommt in einer hybriden IT-Umgebung eine deutlich höhere Bedeutung zu, da neben den »klassischen« Bedrohungen auch die Besonderheiten einer verteilten/ cloud basierten Umgebung hinzukommen. Dies gilt in ähnlicher Weise auch für das **↗ Anomalie-Monitoring**.

Das **↗ Kosten-Monitoring** steht in hybriden Umgebungen vor der Herausforderung, zwei Bereiche – die festen/kalkulierbaren Kosten der traditionellen IT-Umgebung und die variablen Kosten einer Cloud-Umgebung - gleichermaßen transparent zu machen.

Das **↗ Monitoring im operativen Betrieb** beschreibt die Voraussetzungen, die für das Monitoring der Infrastruktur und Services notwendig sind, sowie welche Informationen zum Aufrechterhalten der operativen Systeme hilfreich sind.

In den folgenden Kapiteln sind die Erkenntnisse und Empfehlungen weiter im Detail ausgeführt.

1 Einleitung

1 Einleitung

Mit der Cloud als Treiber für Hybride IT erscheint zunächst alles einfach. Doch in der Regel ist das die Sicht der Anwender und Anwenderinnen, die als Konsumenten oder Konsumentinnen ganz leicht die angebotenen Service in der Cloud nutzen können um ihren jeweiligen Bedarf zu stillen. Die neue und bunte Welt der Cloud mit all ihren Services, Apps, Endgeräten und Webseiten sorgt für Wirtschaftswachstum im Zeitalter der Digitalisierung durch neue, intelligente, stets erreichbare und verfügbare Dienste.

Doch aus Sicht der Unternehmens-IT ist die Cloud anders zu betrachten. Denn auch die interne IT ist – genauso wie die Cloud-Anbieter und -Anbieterinnen – Dienstleister oder Dienstleisterinnen für Mitarbeiter oder Mitarbeiterinnen und Kunden oder Kundinnen des Unternehmens und muss sich nun mit den neuen Standards, die die Cloud Services etabliert haben, messen. Denn wie heißt es so schön: »Die Konkurrenz ist nur ein Klick entfernt« und Kunden/Kundinnen sind heutzutage bereit, für einen besseren Service schnell den Anbieter oder die Anbeiterin zu wechseln. Somit ist für jedes Unternehmen, das sich aus den unterschiedlichsten Gründen für den Einsatz von Cloud-Services entscheidet, ein wesentlicher Punkt, die Gesamtzufriedenheit mit den IT-Services für Mitarbeiter oder Mitarbeiterinnnen und Kunden oder Kundinnen sicher zu stellen.

Dieser Leitfaden soll Erläuterungen und Hilfestellung geben, wie Unternehmen, die ihre IT mit Mehrwertdiensten der Cloud ergänzen, bestehende Dienste in die Cloud verlagern oder ersetzen. Somit werden sie sich automatisch mit einer hybriden IT auseinandersetzen. Diese einschneidende Veränderung im Bereich der IT-Architekturen muss weiterhin den Überblick, die Kontrolle und die Anwenderzufriedenheit im Blick behalten und weitere neue Aspekte beim Monitoring hybrider IT-Umgebungen berücksichtigen. Dabei gehen wir sowohl auf die Einzelkomponenten des hybriden Monitorings ein, als auch auch auf die jeweiligen Personas, die im Bereich Monitoring ihre unterschiedlichen Ansprüche an ein Monitoring- und Alarmierungssystem erheben.

Wir wünschen uns, dass dieser Leitfaden den Leserinnen und Leser eine Grundlage vermittelt, mit deren Hilfe eine ggf. notwendige Neuausrichtung oder Erweiterung eines bestehenden Monitoring- und Alarmierungssystems erfolgreich gelingt und neu aufkommende Aspekte nicht übersehen werden.

1.1 Allgemeine Einleitung

In den heutigen IT Umgebungen existieren bereits schon lange Monitoring- und Alarmierungssysteme. Warum also muss dieses Thema mit dem Einzug der Cloud neu überdacht werden? Erweitert man nicht einfach sein bestehendes Monitoring-Tool um die hinzugekommenen Infrastrukturkomponenten und Services und das Thema kann abgehakt werden? Wird mein existierendes Monitoringsystem nicht sowieso beim nächsten Update eine erweiterte Funktionspalette mitliefern, die eine Änderung zu einer hybriden IT-Umgebung berücksichtigt? Oft ist es leider nicht so einfach. Denn durch die Kopplung von Cloud-Services an eine bisher eher abgeschirmte

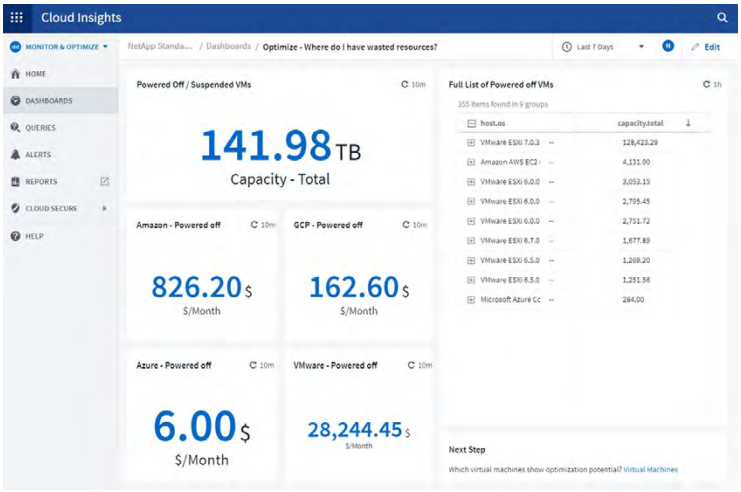
IT-Umgebung und damit die Entstehung einer hybriden IT-Umgebung, kommen wesentliche neue Aspekte auf, die in den bisherigen Architekturen nicht oder weniger priorisiert berücksichtigt werden mussten.

Sehr oft ändert sich beispielsweise mit der Cloud-Nutzung auch das Bezahlmodell, denn in der Regel erfolgt damit auch ein Wechsel von einem Upfront-CAPEX-Modell zu einem »Pay-as-you-go«-OPEX-Modell, wodurch die Überwachung und Kontrolle der Ressourcen und deren erzeugter Kosten ein neuer, zu berücksichtigender Aspekt ist:

- Welche Kosten erzeuge ich für welche Ressourcen?
- Liegen Ressourcen brach, für die ich dennoch Geld bezahle?
- Können Ressourcen minimiert werden?
- Können Services kostengünstiger abgebildet werden?

Situation:

- Welche Kosten erzeuge ich wofür?
- Liegen Ressourcen brach?



Situation:

- Können Ressourcen minimiert werden?
- Gibt es andere Möglichkeiten der Abbildung von Services?

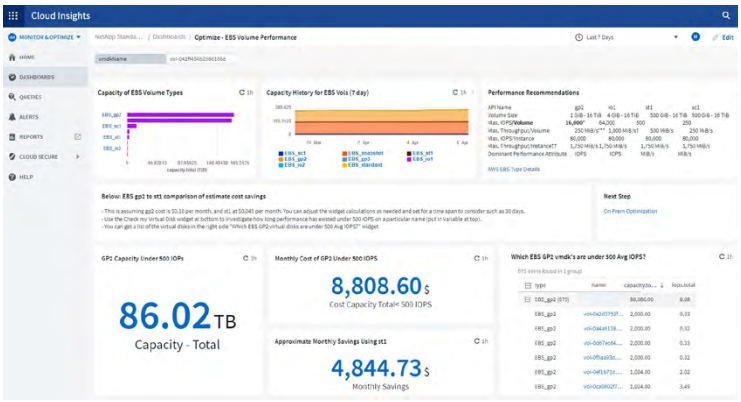


Abbildung 1 und 2: Ressourcen

Auch das Thema Security spielt bei der Nutzung der Cloud eine weitaus wichtigere Rolle als bisher. Denn die neuen Architekturen öffnen weitere Zugangsmöglichkeiten zu den Unternehmens- und Kunden- oder Kundinnendaten , die es intensiv zu schützen gilt. Zumal die Daten eines Unternehmens nicht nur über neue Kanäle zugreifbar sind, sondern je nach Anforderung und Architektur an unterschiedlichsten Stellen abgelegt werden. Auf folgende Fragen sollte man stets eine Antwort haben:

- Wo liegen die Unternehmens- und Kunden- oder Kundinnendaten?
- Ist stets eine Verschlüsselung - beim Transfer und bei der Speicherung - gegeben?
- Treten Anomalien beim Zugriff auf die Daten auf, die auf ein Sicherheitsleck schließen lassen?

Situation:

- Gibt es Anomalien?

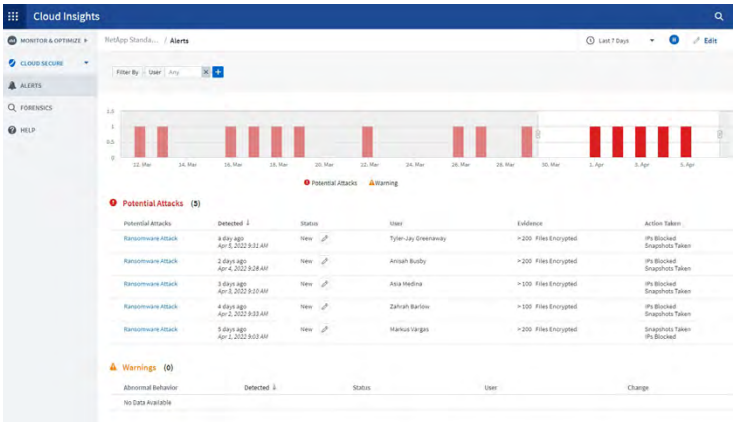


Abbildung 3: Anomalien bei Zugriff auf Daten

Und natürlich ist der Compliance-Überblick über alle Unternehmens- und Kunden- oder Kundinnendaten eine extrem ernst zu nehmende Disziplin. Hier drohen bei Nichteinhaltung, beispielsweise gesetzlicher Vorgaben, empfindliche Strafen und die Schädigung des Rufs eines Unternehmens.

Unter Compliance fällt nicht nur die datenbezogene Compliance, sondern auch die Einhaltung der für die jeweilige Branche gesetzlich, oder in Form von Branchen- bzw. Unternehmensrichtlinien, vorgegebenen Regeln. Mehr Details dazu sind im Kapitel: ↗ **Compliance-Monitoring** zu finden.

Die Ernennung eines Informationssicherheitsbeauftragten im Unternehmen sollte bereits geschehen sein und dieser sollte über die folgenden Punkte (je nach Branche und deren gesetzlichen Vorgaben) auskunftsfähig sein:

Situation:

- Haben wir Compliance-relevante Daten in der Cloud und ist dies mit den Regularien vereinbar?
- Sind die Daten nach Vorgaben geschützt?
- haben wir jederzeit die Möglichkeit, dies zu prüfen und zu reporten?

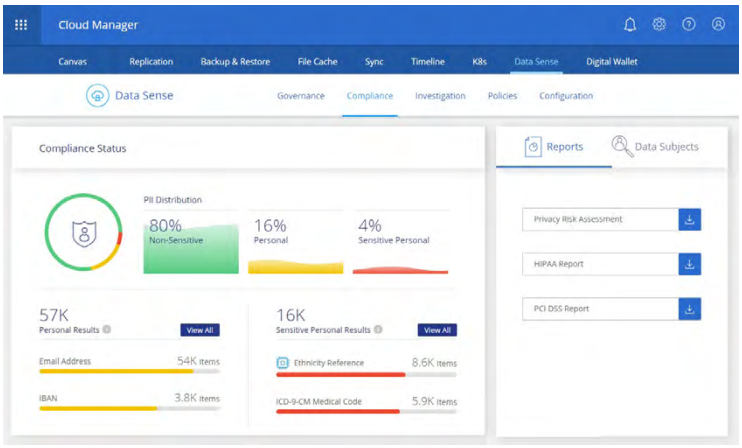


Abbildung 4: Compliance Status

Die drei angesprochenen Punkte - Kosten, Security und Compliance - sind wahrscheinlich die wesentlichsten Aspekte, die bei einem hybriden Ansatz von lokalen, eigenen Datacentern mit Cloud-Umgebungen Dritter eine zusätzliche Rolle spielen. Da diese Aspekte in den meisten Fällen bei den bestehenden Monitoring- und Alarmierungssystemen noch keine ausgeprägte Rolle gespielt haben, empfehlen wir, diese Aspekte individuell auf die eigene IT-Strategie anzuwenden, zu beleuchten und entsprechende Maßnahmen zu planen. Ein gesamtheitliches Monitoring- und Alarmierungssystem sorgt für Sicherheit im Betrieb und bei zukünftigen Erweiterungen und ist damit eine wesentliche Komponente der sich ständig ändernden IT-Landschaft (Build → Monitor → Optimize).

Hinzu kommt eine weitere Ebene: Das Monitoring der Applikationen bzw. der Services, die hinter einer Applikation stehen. Dieser Bereich des Monitorings ist eine logische Klammer, die den traditionellen Aspekt des Monitorings, ob eine Applikation funktioniert oder nicht, mit den neuen Bereichen zusammenführt. Das ermöglicht nicht nur neue Betrachtungsweisen hinsichtlich des Services, sondern macht sie auch notwendig.

Durch die neuen Faktoren und Bereiche, die im Monitoring von hybriden Umgebungen betrachtet werden müssen, kommt natürlich auch die Frage auf, wer sich nun aus welchem Bereich mit dem Monitoring auseinander setzen muss und wer in den unterschiedlichen Bereichen welche Aufgaben und Verantwortungen hat. Oder, ob das Monitoring überhaupt noch im eigenen Bereich, wenn nicht noch im eigenen Unternehmen liegt. Diese Sichtweisen wollen wir explizit beleuchten, um den Leserinnen und Lesern die Möglichkeit zu geben, sich in die Rollen hineinzuversetzen und Vorteile, aber auch Herausforderungen zu verstehen und daraus die eigenen nächsten Schritte abzuleiten.

1.2 Personas / Zielgruppen:

Personas veranschaulichen typische Vertreter einer Zielgruppe. Die Personas, von denen hier gesprochen wird, sind auch beim Einsatz von hybriden IT-Umgebungen von der Disziplin Monitoring tangiert. Hierzu existieren Erwartungen, Wünsche und Ziele, was Monitoring im gesamten Life-Cycle eines Services/Produkts leisten soll. Die Bandbreite der genutzten IT-Umgebung beinhaltet Kombinationen aus On-Premises-Umgebungen und Provider-Services.

Die nachfolgenden Personas orientieren sich an den Mitarbeitenden in den DevOps-Teams und anderen wichtigen Stakeholdern.

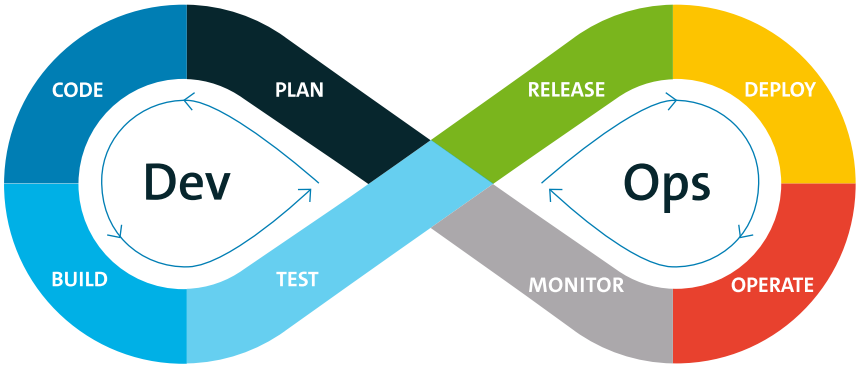


Abbildung 5: Funktionen im Product-Life-Cycle für DevOps

Persona Product Owner: Diese Persona ist ähnlich der Persona Product Manager. Im Wesentlichen geht es um das Verantworten des wirtschaftlichen Erfolgs eines Produkts/Services mit starkem Kunden oder Kundinnen/Nutzer- oder Nutzerinnen-Fokus und darum, dessen Qualität im gesamten benötigten Ökosystem unter optimalem Einsatz des/der zur Verfügung stehenden DevOps-Team(s) sicher zu stellen. Der unterstützende Beitrag des Monitorings liegt in der automatisierten Auswertung adäquater Metriken und Logdateien (Fullstack) nach verschiedensten Kriterien, einer aggregierten Darstellung mit Drill-Down und beim Über- oder Unterschreiten von definierten Schwellwerten (statisch/dynamisch) einer idealerweise präventiven Alarmierung. Monitoring dient über den gesamten Produkt/Service-Life-Cycle der transparenten Qualitätssicherung, Nutzung und Darstellung von Kosten und Umsätzen zum Produkt/Service.

Persona DevOps-Engineer: In dieser Persona sind die drei Personas Legacy-, Fullstack- und Cloud-Native-Entwickler subsumiert. Im Wesentlichen geht es um die Entwicklung und den Betrieb eigener Produkte/Services. Dies können auch Basisprodukte/-services wie zum Beispiel APIs sein. Das Monitoring unterstützt hierbei mit tiefgehenden, automatisierten Auswertungen. Die Basis bilden adäquate Metriken und Logdateien (Fullstack) nach verschiedensten Kriterien, die Darstellung mit Drill-Down und beim Über- oder Unterschreiten von definierten Schwellwerten (statisch/dynamisch), einer idealerweise präventiven Alarmierung und liefert im Falle einer Störung Hinweise auf relevante Ereignisse (Events) zur schnellstmöglichen Entstörung. Monitoring dient über den gesamten Produkt-/Service-Lifecycle mit den benötigten Systemplattformen einer transparenten Qualitätssicherung (Sicherstellung von Funktionen, Performance, Ressourcenbedarf, Security, Stabilität und Verfügbarkeit) und Nutzung der Produkte/Services.

Persona Operations (IT-Betrieb): Diese Persona stellt als Mitarbeiter und Mitarbeiterinnen im klassischen IT-Operations-Center den Betrieb von Produkten/Services auf der Basis von Service Level Agreements (SLAs) sicher. Das Monitoring unterstützt dabei, die Arbeitsfähigkeit des eigenen Unternehmens sicherzustellen und die Ausfallzeiten möglichst gering zu halten. Dazu werden aus dem Monitoring schnellstmöglich Events über sich abzeichnende Tendenzen, die zu Performance-Engpässen oder Fehlern führen können, oder über bereits aufgetretene Fehler erzeugt. Anhand dieser Events ist es möglich, die Fehlerursache (Root Cause) sowie die dadurch erzeugten Auswirkungen abzuleiten und u. a. über Automatismen die Fehler zu beheben. Zusätzlich werden die externen Kunden oder Kundinnen, internen Nutzer oder Nutzerinnen sowie internen Stakeholder über die Dauer der Beseitigung der Störung(en) über geeignete Mechanismen informiert. Dazu gehört unter anderem das Security Operations Center (SOC), dessen spezielle Aufgabe es ist, im Operativen die Security des Unternehmens mit der Abwehr von potentiellen Angriffen zu gewährleisten.

Persona Service: Diese Persona bedient als Mitarbeiter und Mitarbeiterinnen die Schnittstelle zu den internen und externen Kunden oder Kundinnen für Service-Requests und Auskünfte bezüglich der vom Unternehmen mit seinem Ökosystem verantworteten IT-Services. Das Monitoring liefert hierzu Kennzahlen und die vom Service benötigten Detailinformationen für eine optimale Auskunftsfähigkeit gegenüber den Kunden oder der Kundin.

Persona Informationssicherheitsbeauftragter (CISO): Diese Persona umfasst die Wahrnehmung aller Belange der Informationssicherheit innerhalb des Unternehmens und gegenüber externen Erbringern von IT-Services. Sie stellt sicher, dass die in der IT-Strategie, der Informationssicherheitsleitlinie und dem Informationssicherheitskonzept des Unternehmens niedergelegten Ziele und Maßnahmen sowohl intern als auch bei den beauftragten IT-Dienstleistern oder -Dienstleisterinnen eingehalten werden. Monitoring liefert hierzu Kennzahlen und angeforderte Detailinformationen zu den verantworteten IT-Services.

Persona Datenschutz: Diese Persona ist direkt der Geschäftsleitung unterstellt und nimmt im Wesentlichen eine Kontroll- und Organisationsfunktion wahr. Sie sichert als Datenschutzbeauftragte und Datenschutzbeauftragter die Einhaltung der EU-DSGVO anhand von unternehmensweit geltenden Policies ab und berät dazu die Bereiche im Unternehmen. Sie legt den Umgang mit personenbezogenen Daten fest, kontrolliert u. a. auch Arbeitsabläufe und verhindert damit Verstöße gegen gesetzliche und innerbetriebliche Regelungen. Monitoring-Daten sind davon in den Punkten gespeicherte Daten mit deren Inhalten (auch Anonymisierung, Pseudonymisierung etc.), deren Vorhaltezeiten, Löschverfahren und der Zulässigkeit an Verarbeitung betroffen. Verstöße gegen gesetzliche Regelungen sind als Datenschutzverletzung an die zuständige Aufsichtsbehörde zu melden.

Persona Betriebsrat: Diese Persona nimmt ähnlich der Persona Datenschutz im Wesentlichen eine Kontroll- und Beratungsfunktion wahr. Sie sichert zusätzlich zur EU-DSGVO die Rechte und Belange der Arbeitnehmerinnen und Arbeitnehmer gemäß Betriebsverfassungsgesetz und interner Betriebsvereinbarungen ab. Monitoring-Daten sind davon wie bei der Persona Datenschutz in den Punkten gespeicherte Daten mit deren Inhalten (auch Anonymisierung, Pseudonymisierung etc.), deren Vorhaltezeiten, Löschverfahren und der Zulässigkeit an Verarbeitung insbesondere zur Leistungskontrolle betroffen. Dabei kann es durchaus zu Kollisionen aus den unterschiedlichen Interessenlagen der beteiligten Personas kommen.

Persona IT-Revision: Diese Persona ist eine unabhängige und objektive Einheit zur systematischen, risikoorientierten und zielgerichteten Prüfung und Beratung zu allen informationsverarbeitenden Funktionen - damit auch dem Bereich Monitoring - im Unternehmen. Wie bei der Persona des Datenschutz besteht hier eine Kontroll- und Organisationsfunktion.

Persona Product Owner Monitoring: Diese Persona verantwortet die Strategie und die Ausrichtung des Monitorings innerhalb des eigenen Unternehmens im Zusammenspiel mit dem gesamten Ökosystem unter Betrachtung der Wertschöpfungskette (End-to-End-Verantwortlichkeit). Dazu gehört, dass der/die Product Owner mit seinem/ihrem Team als Single-Point-of-Contact (SPoC) alle Anforderungen zum Monitoring erhält und diese in aktuelle Konzepte einbaut, sowie in einer Strategie münden lässt. Er entscheidet über Art und Einsatz der Tool-Landschaft. In der Konsequenz existiert im Unternehmen ein internes Produkt »Monitoring-as-a-Service«.

2 Performance Monitoring

2 Performance Monitoring

Wird man das erste Mal mit dem Begriff Performance Monitoring konfrontiert, stellt man sich die Frage, was überhaupt mit Performance gemeint ist. Je nach Sichtweise, also aus der jeweiligen Rolle der Mitarbeitender und Mitarbeitenden oder aus der Sichtweise der Abteilung bzw. Technologiegruppe, kann dies stark variieren. Trotz des sehr starken Fokus auf Technologie, kann Performance auch hinsichtlich Erfolg und Monetarisierung verstanden werden, was den technischen und nichttechnischen Bereich zusammenfügt und eins werden lässt.

Betrachtet man das Monitoring nun zusätzlich aus der Sichtweise der hybriden IT, so müssen zudem die Performance-Indikatoren aus unterschiedlichen Bereichen zusammengebracht werden, so dass diese als kumulierte Kennzahlen von allen verschiedenen Rollen, egal ob technisch oder nicht, verstanden werden können.

In welchen Bereichen haben wir nun also eine Performance zu betrachten und wie können wir diese überwachen? Aufgrund der Vielzahl der unterschiedlichen Sichtweisen und zu gewinnenden Informationen müssen wir uns auf die Grundpfeiler des Performance Monitoring fokussieren.

Qualitätssicherung in den unterschiedlichen Bereichen

Es kann unter Umständen verwundern, dass man bei der Betrachtung des Monitoring-Aspektes in hybriden Umgebungen mit der Softwareentwicklung beginnt. Die Betrachtungsweise hat sich hier in den letzten Jahren merklich geändert und die Grenzen zwischen den unterschiedlichen Technologie-Bereichen verschwimmen.

Wir fangen nun also mit dem Performance Monitoring zu einem Zeitpunkt an, an dem noch überhaupt keine Software existiert bzw. zu einem Zeitpunkt, in dem sich die Software noch in einem sehr frühen Zustand befindet. Durch die Betrachtung in der hybriden IT ist zudem der Begriff Software mit Service gleichzusetzen und ähnlich wie in den anderen Bereichen zu betrachten. Um so früher der Entwicklungsprozess und die daraus iterativ entstehenden Minimal Viable Products überwacht und ausgewertet werden, desto eher ist es möglich, einzugreifen und Fehler zu vermeiden, die später zu einer Herausforderung im finalen Service werden können. Der Quellcode, der kompilierte Code und im gleichen Zuge der ausgeführte Code in den unterschiedlichsten Hosting-Varianten wird überwacht und alle daraus gewonnenen Informationen in Relation gestellt. Somit gewinnt die Softwareentwicklung Einblicke in den Zeitpunkt nach Auslieferung des Codes und kann auch aus Entwicklungssicht besser Einfluss auf die Performance der Applikation nehmen.

Aus der Sicht von Operations wird die Performance hinsichtlich der hardwarenahen Faktoren betrachtet. Es wird bewertet, wie viel Arbeitsspeicher eine Applikation benötigt, wie viel CPU durch eine Applikation beansprucht wird, oder wie viel Last durch Datenbankabfragen generiert wird. Operations beschreiben den direkten Übergabepunkt zum Support oder ggf. sogar bereits zu den Anwender und Anwenderinnen des Services selbst. Sie sind darauf bedacht, die eigene Performance soweit hochzuhalten, dass die Nutzer oder Nutzerinnen zufrieden sind und ein hohes Akzeptanz-Level herrscht. Zudem wird bereits im Voraus bewertet, wie die Performance in der

Zukunft entweder besser oder aber auch schlechter werden kann. Dies geschieht zum einem im On-Premises-Bereich bzw. in der eigenen Private Cloud, mit den zugehörigen Einschränkungen hinsichtlich Skalierbarkeit, aber auch im Bereich Public Cloud mit dem Management von On-demand-Ressourcen.

Auch im Performance Monitoring oder allgemein im Monitoring ist die Verschmelzung der Bereiche Development und Operations zu erkennen, die aus den beiden Rollen den DevOps-Mitarbeiter oder -Mitarbeiterinnen werden lässt. Dadurch entsteht ein gemeinsames Verständnis von Performance und damit verbundener Qualität und Qualitätssicherung, die in die darauf aufbauenden Bereiche weitergegeben werden können.

Als letzte Ebene vor den Endanwender und Endanwenderinnen ist der Support zu betrachten. Besteht von der Entwicklung über den Betrieb bereits ein gemeinsames Verständnis der Performance von Services, so ist es auch dem Support möglich, bereits direkt nach einer Anfrage die notwendigen Informationen weiterzuleiten oder durch Automatisierung bereits selber die Lösung des Problems durchzuführen.

Key Performance Indicators

Um die Auswertung der Monitoring-Ergebnisse und daraus abzuleitenden Aktionen – ob manuell oder auch automatisch – erstellen zu können, benötigen wir entsprechende Key Performance Indicators, die einen maßgeblichen Einfluss auf die Gesamtbewertung der Performance haben. Alle Indikatoren können für sich selbst betrachtet werden, haben aber in den meisten Fällen direkte Auswirkungen auf andere Indikatoren und müssen gemeinsam ausgewertet werden.

Durch die Verschmelzung der unterschiedlichen Bereiche nehmen wir hier keine Trennung mehr in Private Cloud bzw. On-Premises oder Public Cloud, sowie keine Trennung in unterschiedliche Rollen vor. Key Performance Indicators im Bereich Performance Monitoring sind z. B.

- Ausfälle von Festplatten im Storage-Bereich
- Voraussage von Ausfällen im Bereich Storage
- Erreichen von Storage Limits
- Erreichen von Memory Exhaustion
- Erreichen von CPU Exhaustion
- Antwortzeiten von Backend Services und deren Frameworks
- Antwortzeiten von Frontend Services und deren Frameworks
- Code-Hotspots aus Auslöser von Memory und CPU Exhaustions

- High Consumption Datenbankabfragen
- Akzeptanz von Services

Da wir mit dem Monitoring und dem Sammeln der Key Performance Indicators das Ziel haben, die negativen Einflüsse auf unsere Umgebung zu minimieren, liegt der Fokus auf den negativen Aspekten bzw. der negativen Bewertung der Key Performance Indicators. Dies sollte aber nicht allgemeingültig angewandt werden. Alle Indikatoren können negativ sowie positiv betrachtet werden und müssen somit regelmäßig neu angewandt und bewertet werden. Hierfür müssen in den einzelnen Dimensionen, aber auch auf multidimensionaler Ebene Baselines gebildet werden, um negative Auswirkungen zu minimieren und durch daraus abgeleiteten Aktionen positive Effekte zu haben und die jeweiligen Bereiche zu optimieren.

Business Monitoring

Der Fokus lag bisher auf den Key Performance Indicators des technischen Bereiches, die Performance kann und muss aber auch in anderen Bereichen bzw. Richtungen betrachtet werden. Die Betrachtungsweisen können bei der Nutzung von externen Services wie auch bei der Entwicklung von eigenen Services nahezu gleichermaßen angewandt werden.

Werden Services selbst entwickelt, dann wird z. B. gemessen, wie lange es braucht einen Service nach außen hin zu veröffentlichen und zur Verfügung zu stellen. Hierbei ist auch die Vermeidung von Fehlern bei der Entwicklung wichtig, damit Services nicht neu entwickelt oder mit viel Aufwand umgeschrieben werden müssen. Bei der Konsumierung von anderen Services ist es wichtig, diese möglichst schnell zu platzieren und die externen Quellen einzubinden. Dies hat Einfluss auf die Akzeptanz der Anwenderinnen und Anwender sowie auf die Bewertung hinsichtlich Time-to-Market.

Sind die Services zur Verfügung gestellt und werden diese genutzt, so wird im Business Monitoring die Performance hinsichtlich der Nutzung wichtig. Hierzu gehören diejenigen Aspekte, die direkten Einfluss auf das Business haben, ob durch die Nutzung von Services oder das Anbieten dieser.

Hierzu gehören Kennzahlen wie z. B.:

- Welche Lizenzen sind für den eigenen Anwendungsfall relevant und werden genutzt?
- Welche Lizenzen werden durch Konsumentinnen und Konsumenten des Services genutzt bzw. präferiert?
- Welche Versionen von Services und Komponenten werden genutzt?
- Welche Features werden in den Anwendungen besonders stark genutzt und welche Features eher weniger?

- Der sich aus Kombinationen der genannten Kennzahlen ableitende effektive Umsatz bzw. Gewinn durch angebotene und verwendete Services
- Conversion Rate bei der Nutzung spezifischer Dienste
- Nutzung der Anwendungen bezogen auf geographische Regionen und Zeitzonen

Application Performance Monitoring & User Experience

Um die Gesamtperformance zu messen und die Relationen zwischen allen Komponenten herzustellen, kann das Application Performance Monitoring genutzt werden. Je komplexer die Anwendungslandschaft sowie die Anwendungen und deren Konstruktion werden, desto wichtiger wird es, diese Verbindungen sichtbar zu machen. Obwohl es nicht von Beginn an ersichtlich ist, kann jede Komponente Einfluss auf jede andere Komponente haben.

Der Begriff der Anwendung bzw. Lösung wird im Application Performance Monitoring anders angewendet, als man dies traditionell verwendet. Anstelle der einzelnen Anwendungen als Software, wird nun von einem logischen Zusammenschluss aller Elemente gesprochen.

Beispiele für den Einfluss der jeweiligen Komponenten aufeinander sind:

- Ein nicht optimierter Query, der an eine Datenbank geschickt wird, hat Einfluss auf die Performance einer einzelnen Funktionalität auf einer Homepage und somit auch auf die Conversion Rate
- Ein Klick auf einen Button in einer Web Application triggert eine Funktion, die einen Code Hotspot enthält, der dafür sorgt, dass der Arbeitsspeicher des Host-Systems so stark belastet wird, dass es ausfällt
- Der Ausfall eines einzelnen Services sorgt dafür, dass davon abhängige Services bzw. Service-Verbunde nicht mehr funktionen

Werden diese Abhängigkeiten schnell und automatisiert erkannt, so ist es direkt möglich bei der Root-Ursache anzusetzen. Die Root-Ursache und die für die Lösung abgeleiteten Aktionen werden in den meisten Fällen über Funktionalitäten Künstlicher Intelligenz (KI) errechnet bzw. bestimmt. Die Lösungsfindung hat also ggf. Einfluss auf die Verbesserung von mehreren Performance-Kategorien, sowie auf die Gesamtperformance.

Die letzte Ebene, die durch alle zu überwachenden Komponenten bestimmt und maßgeblich beeinflusst wird, ist die User Experience. Ob die User Experience für die eigenen Anwender und Anwenderinnen gemessen wird oder für die Anwender und Anwenderinnen einer extern bereitgestellten Software, macht grundsätzlich keinen Unterschied. Ist ein Anwender oder eine Anwenderin nicht zufrieden, hat dies Einfluss auf alle Performance-Bewertungen und die Performance-Kennzahlen – in Summe bestimmen sie die User Experience.

3 SLAs, Verfügbarkeiten

3 SLAs, Verfügbarkeiten

Spätestens seit dem Siegeszug von ITIL sind Service Levels und Verfügbarkeiten in der IT seit Jahren etablierte Messgrößen für die Qualität von IT-Services. In hybriden Umgebungen besteht die besondere Herausforderung darin, dass sich ein spezifischer IT-Service aus mehreren Teilservices speisen kann, die wiederum bei unterschiedlichen Anbieter und Anbieterinnen gehostet werden. Jeder dieser Teilservices wird dabei sehr wahrscheinlich anderen Metriken und Maßgaben folgen, so dass sich ein SLA bzw. die Bestimmung von Verfügbarkeiten für den gesamten Service aufwendig gestalten kann. Beispielsweise kann die Datenbank auf internen Servern liegen, die Middleware durch einen Dienstleister oder eine Dienstleisterin bereitgestellt werden und die virtuellen Loadbalancer sowie die Webfrontends in der Cloud skalieren.

Service Level Agreements (SLA)

Für die Erbringung bzw. den Konsum der Services werden zwischen Lieferanten und Konsumenten sogenannte Service Level Agreements (SLA) definiert, welche Kenngrößen, KPI, Reaktions- und Wiederanlaufzeiten oder auch Verrechnungsmodelle definieren.

SLA definieren aber auch Reaktions- oder Lösungszeiten im Störungsfall, Kapazitäten, Change-Prozesse, Tools und Methoden. Hier ist sicherzustellen, dass die betreffenden Leistungen der Teilservices zusammenpassen.

Benötigt beispielsweise ein Service eine zugesicherte Wiederanlaufzeit von maximal acht Stunden, so darf die Summe der Wiederanlaufzeiten aller Teilservices dies nicht überschreiten. Zusätzlich müssen Puffer eingerechnet werden, wenn das Incident Handling ggf. sequenziell durch die Teilservices läuft und so zusätzliche Zeitaufwände bedeutet.

Verfügbarkeit

Unter Verfügbarkeit versteht man die Zeit, die ein (Teil-) System pro Zeiteinheit (meistens pro Jahr) effektiv zur Nutzung bereitsteht, bzw. wie lange pro Zeiteinheit das System maximal ausfallen darf. Dabei werden Verfügbarkeiten in Prozenten angegeben.

So bedeuten beispielsweise 98 Prozent bei 7x24x365 (Tage pro Woche x Stunden pro Tag x Tage pro Jahr) zugesagter Verfügbarkeit, dass ein solches System im Jahr maximal 175,2 Stunden ausfallen darf, und dabei immer noch die zugesicherte Verfügbarkeit einhalten würde. Das entspräche dann 7,3 Tagen erlaubter Ausfall. Je nach Service kann das ausreichend sein. Eine Verfügbarkeit von 99,95 Prozent bei 7x24x365 entspricht dagegen nur noch 4,38 Stunden maximaler Ausfallzeit pro Jahr, was aber bei unternehmenskritischen Systemen immer noch zu lange sein kann.

Bei vielen Cloud-Services fällt diese dabei relativ niedrig aus im Vergleich zu Legacy-Services. Hintergrund ist dabei meist, dass höhere Verfügbarkeiten entweder technologisch anders gelöst

werden müssen, als das beispielsweise bei virtuellen Maschinen oder typischen Clustern der Fall ist (z. B. bei Containern). Oder die Cloud-Betreiber bieten höhere Verfügbarkeiten erst in den teureren Tarifen an, so dass hieraus zusätzliches Geschäft generiert werden kann.

Mit steigender Verfügbarkeit wird der technische Aufwand immer größer und die Kosten somit höher. Daher ist bei der Planung von IT-Services die Frage nach der tatsächlich erforderlichen Verfügbarkeit zu prüfen, um die Aufwände und Kosten in einer gesunden Relation zum Nutzen des konkreten Services zu halten.

Auch wird die Verfügbarkeit bei jeden Anbieter und Anbieterin auf eine andere Weise angegeben. Häufig sind Cloud-Services nicht auf 24x7x365, sondern auf »typische« Bürozeiten wie 8x5x365 berechnet. Bei der Kombination von Teilservices muss daher auf eine einheitliche Berechnungsbasis normalisiert werden, damit die Gesamtverfügbarkeit des Systems sauber ermittelt werden kann.

Zur Bestimmung der Gesamtverfügbarkeit werden die Teilverfügbarkeiten multipliziert. Besteht also ein spezifischer, unternehmenskritischer Service aus 3 Teilservices zu je 99,99 % (extrem teuer), so hat der Gesamtservice allerdings nur

99,99 % x 99,99 % x 99,99 % = 99,97 % Verfügbarkeit.

Anstatt einer maximal erlaubten Ausfallzeit bei 99,99 % Verfügbarkeit einzelner Services von (365 Tage x 24 Stunden/Tag) - (365 Tage x 24 Stunden/Tag x 99,99 %) = 0,876 Stunden pro Jahr, wären dann für den Gesamtservice bei 99,97 % Gesamtverfügbarkeit bereits

(365 Tage x 24 Stunden/Tag x 99,97 %) = 2,6 Stunden (!) pro Jahr Ausfall erlaubt.

Hier ist also immer darauf zu achten, welche Metriken die Teilservices aufweisen, diese zu normalisieren und einen Gesamt-SLA sowie die Gesamtverfügbarkeit mit Blick auf den Gesamtservice zu ermitteln.

Fazit

Zusammenfassend kann daher gesagt werden, dass das Monitoring von SLA und zugehörigen Verfügbarkeiten besonderes Augenmerk erfordert und sowohl vertraglich als auch technisch aufwendiger sein kann als bei selbst erbrachten Services.

4 Kapazitäts-Monitoring

4 Kapazitäts-Monitoring

Im Kapitel ↗ **Performance Monitoring** wurden die Begriffe Business- und Performance Monitoring eingeführt, welche einen starken Bezug zu Applikationen haben. Kapazitäts-Monitoring beschreibt die Größenordnung und die Auslastung von Ressourcen bezüglich solcher Applikationen, also die Auslastung der zugeordneten Hardware. Hardware in diesem Sinne sind dann CPU Cores, RAM aber auch Netzwerk und Massenspeicher.

Diese Kapazität bezieht sich heute nicht mehr nur auf dedizierte Server (man spricht von Bare Metal), sondern eher auf eine oder mehrere Virtualisierungsschichten, z. B. auf virtualisierte Server und den darauf laufenden Umgebungen für containerisierten Code, sogenannte »Micro Service Architekturen«. Diese Virtualisierungsschichten werden auch Hypervisor oder Virtual Machine Monitor (VMM) genannt. Auch in der Netzwerktechnik setzen sich Virtualisierungen immer mehr durch: Software Defined Networks (SDN) und Netzwerkvirtualisierung (NFV, VNF).

Eingeführt wurde dieses Schichtenmodell um Server besser auslasten zu können. Die Container waren dann eine Weiterentwicklung, um diese virtuellen Instanzen als solches schlanker gestalten zu können. Nicht jeder Container benötigt das volle Betriebssystem, sondern nur Teile davon. Nimmt man diese nicht benötigten Teile weg, landet man automatisch beim Containermodell. Durch die Art, wie Container genutzt werden, ergibt sich ein ganz anderes, potenzell wesentlich dynamischeres Lastprofil als bei bisherigen Client/Server-Architekturen. Teilweise können Container mit hoher Frequenz provisioniert und auch wieder abgebaut werden. Für das Monitoring der Hardware-Kapazitäten ist daher eine engere und schnellere Grenzwertüberwachung erforderlich, um Engpässe oder Überlastung zu vermeiden. Darüber hinaus ergeben sich für die künftige Planung von Kapazitäten neue Herausforderungen, da die effektive Nutzung der Plattform sehr schnell sehr stark schwanken kann.

Kapazitäts-Monitoring ist das Messen der Auslastung der verwendeten Hardware. Bei dedizierten oder virtualisierten Servern wird in der Regel mit Agenten basierten Systemen gearbeitet. Ein Agent sammelt Informationen über die zu überwachenden Ressourcen und übermittelt diese dann an ein zentrales System. Die Container-Laufzeitumgebungen stellen oft mittels API Messdaten zur Verfügung. Dadurch reduziert sich der Aufwand, den Agenten aufbringen, weil über die APIs zentral eine ganze Umgebung gemessen werden kann.

In cloudbasierten Umgebungen hat man als Dienstnehmer oder -nehmerin nicht zwangsläufig Zugriff auf die zugeordnete Hardware, somit ist auch kein eigener Agent installierbar. Cloud-Anbieter und -Anbieterinnen stellen in der Regel über APIs Messdaten zu Fehlersituationen und Ressourcen-Auslastung zur Verfügung. REST APIs und Webhooks sind gängige Integrationen.

Die gewonnenen Messdaten werden beim Kapazitäts-Monitoring dazu genutzt, die Planung von neuen Ressourcen zu unterstützen. In der »traditionellen Welt« sind durch die Prozesslaufzeiten für Bestellung, Lieferung, Einbau und Installation Voraussagen nötig, um immer genügend Hardware-Ressourcen zur Verfügung zu haben. Das Kapazitäts-Monitoring kann hier Trends aufzeigen und somit die Optimierung von freien Ressourcen unterstützen. Das Kapazitäts-Monitoring spielt also sowohl in der Service-Qualität (immer genügend Ressourcen vorhanden,

um den Service auch zu erbringen) seine Rolle, als auch in der Kostenplanung (umso weniger freie, ungenutzte Ressourcen bereitgestellt sind, umso günstiger ist der Betrieb).

In Cloud-Service-Umgebungen stehen den Nutzerinnen und Nutzern theoretisch unendlich viele Ressourcen zur Verfügung – der Cloud-Service-Provider ist in der Pflicht, immer für die notwendigen Ressourcen zu sorgen. Er nutzt das Kapazitäts-Monitoring intern für die Qualitätssicherung seiner Dienstleistung. Als Kundin oder Kunde bzw. User der Cloud kann ich diese vom Provider bereitgestellten Kapazitäts Informationen aber weiterhin zur Kostenanalyse und für das Trending nutzen - die Prozesszeit für die Bereitstellung einer Ressource ist im Gegensatz zu einer traditionellen On-Premises-Umgebung sehr kurz. Diese unterschiedliche Sichtweise und die verschiedenen Schnittstellen (Agent vs. API) auf das Kapazitäts-Monitoring muss in einer hybriden Umgebung beachtet werden.

Ein Vorteil der hybriden Umgebung kann dabei in der schnellen Überbrückung von Ressourcen-Engpässen durch Cloud-Ressourcen gesehen werden. Wenn durch das Monitoring in der On-Premises-Umgebung Engpässe erkannt werden, die durch eine Bestellung von interner Hardware nicht zeitnah ausgeglichen werden kann, dann können kurzfristig Cloud-Ressourcen eingebunden werden. Somit kann sehr flexibel auf Änderungen reagiert werden, ohne die Servicequalität zu gefährden.

Ein neuer Trend, der gerade in hoch agilen und flexiblen Cloud-Architekturen zu finden ist, ist »Application-driven Infrastructure« (kurz: ADI). Unter ADI versteht man die stetige Analyse des konkreten Ressourcen-Bedarfs einer Applikation oder eines Services und der automatischen Bereitstellung der genau an deren Bedarf angepassten Ressourcen, wodurch wiederum kaum Überprovisionierungen entstehen und damit einer Verschwendung vorgebeugt werden kann. Die flexible Nutzung und verbrauchsorientierte Abrechnung von Cloud-Anbieter und -Anbieterinnen haben diesen Trend befeuert. Damit ADI überhaupt umgesetzt werden kann, ist das Kapazitäts-Monitoring als grundlegende Disziplin eine Voraussetzung. Denn wenn ich keinen konkreten Ressourcenbedarf messen kann, kann ich auch keine darauf abgestimmten Ressourcen zur Verfügung stellen. Das Thema ADI wird in diesem Whitepaper nicht weiter betrachtet. Wir empfehlen aber, sich mit diesem Thema auseinander zu setzen, um die Konzepte zu verstehen, dadurch auch Cloud-Ressourcen automatisch optimal bereitstellen und Cloud-Kosten ggf. senken zu können.

5 Compliance Monitoring

5 Compliance Monitoring

Compliance und deren Einhaltung kann unter verschiedenen Blickwinkeln erfolgen. Im Folgenden liegt der Fokus auf der Einhaltung von gesetzlichen, regulatorischen und branchen- bzw. kunden- oder kundinnenspezifischen Compliance Vorgaben.

Dimensionen der Compliance

Anforderungen an die Compliance die in einer (hybriden) IT-Umgebung monitored werden (IT-Compliance), lassen sich sich (ohne Anspruch auf Vollständigkeit) in folgende Gruppen aufteilen:

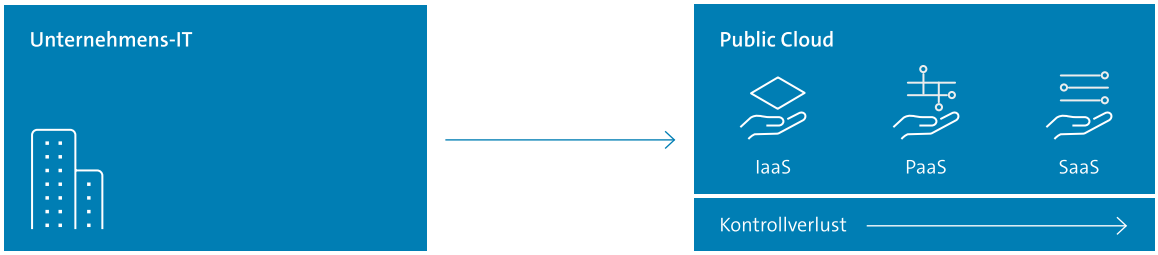
- Direkte, gesetzliche Anforderungen wie:
  - Die ↗ **Datenschutz-Grundverordnung**
  - Das ↗ **zehnte Sozialgesetzbuch – § 80** Verarbeitung von Sozialdaten im Auftrag
  - Das ↗ **IT-Sicherheitsgesetz**
  - Der Sarbanes-Oxley Act (SOX) für an der US-Börse notierte Unternehmen
- Anforderungen und Kriterienkataloge, die sich direkt oder abgeleitet aus Normen und Prüfschemata ergeben, bzw. auf die Anforderungen aus den beiden oben genannten Gruppe referenziert sind:
  - Der Kriterienkatalog ↗ **C5 des BSI**
  - Die Normen der ↗ **ISO 270xx** Reihe
  - Das insbesondere in angelsächsischen Länder verbreitete ↗ **NIST 800-53 Framework**
  - Die ↗ **Cloud Controls Matrix (CCM)** bzw. auch die ↗ **CIS Controlls** des Center for Internet Security
  - ↗ **Die TISAX** Vorgabe in der Automobilbranche
- Industrie spezifische Compliance Anforderungen (hier aufgezeigt am Beispiel der Finanzindustrie) durch:
  - Regulatoren wie zum Beispiel die BAFIN, ↗ **EBA** EIOPA, ect. Bankaufsichtliche Einzelaspekte, die Compliance-relevant sind, basieren i.d.R. auf EU- oder nationaler Gesetzgebung, und präzisieren genannte Anforderungen. Beispielsweise leitet die BaFin aus dem ↗ **Kreditwesengesetz** die ↗ **Mindestanforderungen an das Risikomanagement** (MaRisk) und daraus wiederum die ↗ **Bankaufsichtlichen Anforderungen an die IT von Banken** ab.
  - die EU Regulatory Technical Standards wie dem ↗ **Digital Operational Resilience Act** (DORA).
  - Community basierte Vorgaben wie der ↗ **PCI-DSS** Standard oder Vorgaben von Nutzer- oder Nutzerinnengruppen wie z. B. der ↗ **European Cloud User Coalition**
  - Standards die durch Anbieter und Anbieterinnen wie z. B. ↗ **S.W.I.F.T.** gesetzt werden.

- Kunden- oder Kundinnenspezifische Kontrollpunkte, die eine Ergänzung/Abwandlung der oben aufgeführten Anforderungen darstellen.

Der Großteil der regulatorischen/unternehmensinternen Anforderungen gibt vor, was zu erfüllen ist, lässt dabei jedoch offen, wie die Regelkonformität (=Compliance) im Detail zu erreichen ist. Handlungsanweisungen sind z. B. in den BSI Grundschutz-Katalogen aufgeführt.

Verteilte Zuständigkeiten

In hybriden IT-Umgebungen kommt es regelmäßig zu einem Kontrollverlust des Unternehmens, der durch die verteilten Zuständigkeiten zwischen Unternehmens-IT und IT-Dienstleistern oder - Dienstleisterinnen begründet ist.



Zuständigkeit und Verantwortlichkeit des Unternehmens

Verteilte Zuständigkeit (Verantwortlichkeit des Unternehmers)

Abbildung 6: Zuständigkeit versus Verantwortlichkeit

Zu Sicherstellung der erforderlichen Transparenz über die gesamte IT-Umgebung muss das Compliance-Monitoring ein gemeinsames Set von Kontrollpunkten nutzen und sich über alle Bereiche, in denen im hybriden Modell IT-Services angesiedelt sind, erstrecken.

End to End Compliance

Alle unternehmensweit zu erfüllenden Compliance Anforderungen sind bestenfalls in einem Global Risk and Compliance System unternehmensweit definiert.

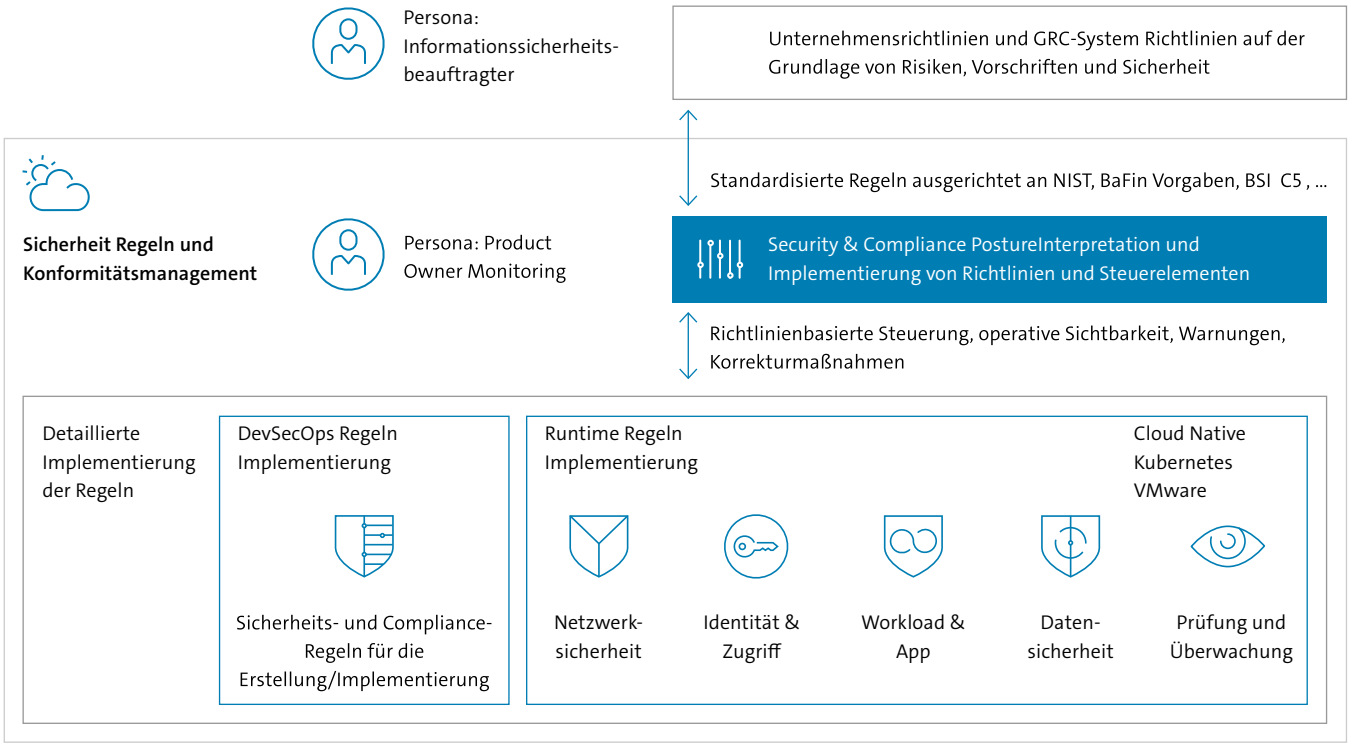


Abbildung 7: Einordnung Compliance Posture Management

Ein Untermenge von Kontrollpunkten daraus wird auf die hybride IT-Umgebung angewendet und ist damit Teil des für die Umgebung anzuwendenden Monitorings.

Automatisierung von Compliance Monitoring und Reporting

Neben dem gemeinsamen Set von Kontrollpunkten ist es in einer hybriden IT-Umgebung erforderlich, die unterschiedlichen Dimensionen und Granularitäten/Abstraktionsebenen von Compliance-Anforderungen normiert technisch abzubilden, um diese dann automatisiert zu überwachen. Dabei lässt sich unterscheiden in:

- Security/Compliance Posture Management (Verwaltung des Sicherheits- und Compliance-Niveaus)
  - Ergebnis ist die Abbildung des Ist-Status (im Vergleich zum Soll-Status)
- Configuration Governance (Konfigurations-Steuerung) zur Unterstützung bei der Durchsetzung von Compliance-Regeln
  - Präventive Kontrolle ergänzend zur Überprüfung (Stichwort DevOps / CI/CD-Integration)
- Sicherheitsanalysen übernehmen das Logging und Monitoring aller sicherheitsrelevanten Aktivitäten. Diese sind im Abschnitt ↗ **Security-Monitoring** und ↗ **Anomalie-Monitoring und Alarmierung** näher erläutert.

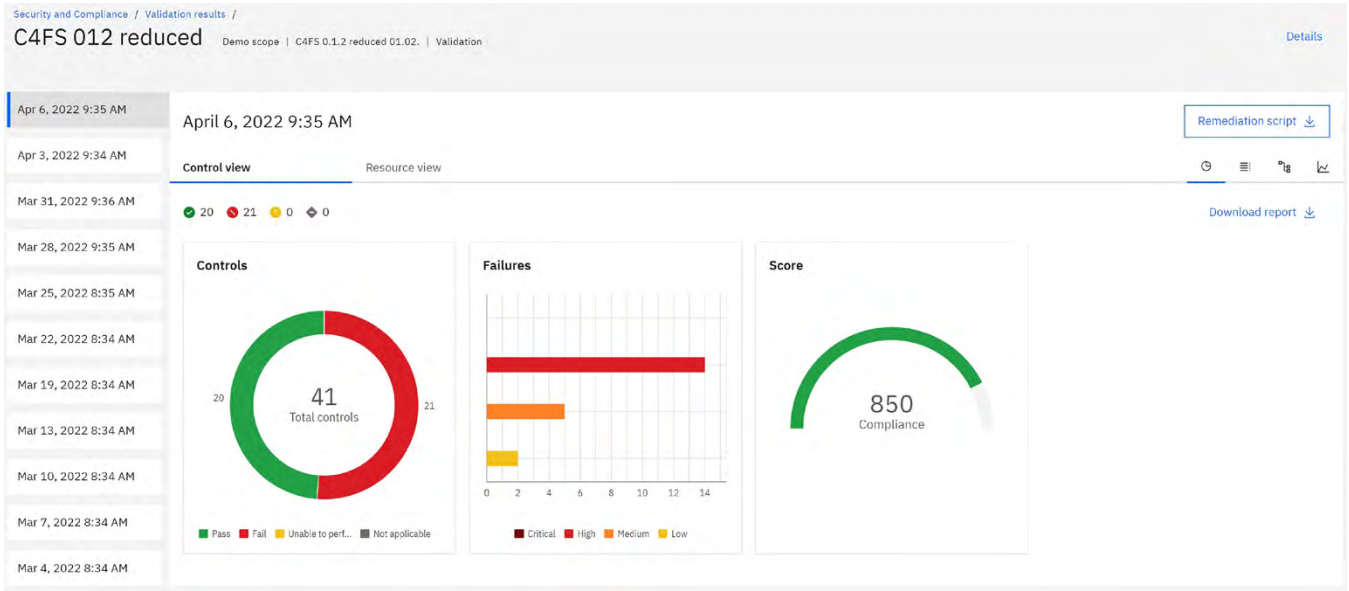


Abbildung 8: Beispielhafte Monitoring Ergebnisse im Rahmen des Posture Management

Security Monitoring

Der Ursprung des Monitorings liegt in der Überwachung von Prozessorleistung, Speicher und Netzwerk. Prozesse und IT-Vorgänge werden dabei in der klassischen IT überwacht und protokolliert. Mit wachsenden Anforderungen wurde das Monitoring immer weiter ausgedehnt. Die heutigen Standards umfassen auch noch die Überwachung von:

- Application
- End-to-end-Monitoring
- SLA Monitoring (SLA = Service Level Agreement)
- Performance and Growth
- Disaster Recovery and Business Continuity
- Identity und Access

Neben diesen klassischen, schon lange vorhandenen Monitor Lösungen, kommen jetzt neue Monitoring-Anforderungen durch die Hybride IT auf die Verantwortlichen zu.

Die weltweite Akzeptanz von Containern als Laufzeitumgebung nimmt drastisch zu, da eine stetig wachsende Anzahl von Firmen neue Anwendungen damit entwickelt und einsetzt. Es wird prognostiziert, dass 75 Prozent aller Organisationen bis 2022 Container-Anwendungen eingeführt haben werden. Dies ermöglicht es der Anwendung, sehr flexibel Ressourcen, entweder von On-Premises oder XaaS, zu nutzen.

Während das »Lift and Shift« traditioneller virtueller Maschinen (VMs) in Container möglich ist, lässt sich das volle Potenzial und die Vorteile von Containern erst dann nutzen, wenn neue, cloud-native Anwendungen unter Verwendung von Containern und Microservices entwickelt werden. Die Konvertierung ineffizienter VMs in Container wird eine temporäre Maßnahme dieser Transformation werden.

In diesem Szenario kommen hybride, IT-spezifische Anforderungen in Bezug auf die Überwachung von persistenter Speicherung, Datenmanagement, Sicherheit und Multi-Cloud- oder Datenzentrumsübergreifende Unterstützung, anwendungsgesteuerte Überwachung, Kostenüberwachung hinzu.

Beim Monitoring-Betrieb in hybriden IT-Umgebungen hat man es auch mit unterschiedlichen Verantwortlichkeiten zu tun. Beim eigenen Monitoring für die eigene IT spricht man von »above the line«, während man bei dem Monitoring der als XaaS bezogenen Dienste von »below the line« spricht. Letztere liegen in der Verantwortung des Cloud Providers (vgl. auch Kapitel ↗ **Compliance-Monitoring**).

Ein traditionelles Monitoring setzt oft auf die Verwendung von Agenten auf. Der Agent überwacht die entsprechende Ressource und die Daten werden an die Monitoring-Instanz geschickt. Die Hardware-Ressourcen einer Cloud lassen sich so nicht überwachen – allenfalls ist es möglich, im Rahmen von »Lift and Shift« auch einen Agenten in der VM in die Überwachung zu integrieren. Service-Provider stellen aber eigene Methoden zur Leistungsmessung bereit, die in das eigene Monitoring übertragen werden können. Somit wird das Monitoring »agentless« betrieben. Wie im Bereich Anwendungsbasiertes Monitoring beschrieben, kann dies auch auf die Anwendungsschicht erweitert werden.

Beispiele für Monitoring relevante Bereiche

Persistent Storage

Beim persistenten Storage handelt es sich um die fest zugeordneten Speicher-Bereiche die einer VM, einem Container oder auch Applikation spezifisch zugeordnet sind. Die bekannten Messattribute wie Verfügbarkeit, Kapazität, Auslastung und Performance gelten auch hier. Neu ist die Agilität der Applikation und damit auch das häufige Verschieben von persistentem Storage. Das muss in einem Monitoring-System berücksichtigt werden.

Multi-Cloud- oder Cross-Datacenter-Support

In der hybriden IT kann es durchaus vorkommen, dass Applikationsteile eines Lösungsstacks verteilt laufen. So kann eine Datenbank im eigenen Rechenzentrum, also On-Premises laufen. Ein Applikationsteil läuft bei einem Cloud-Anbieter oder -Anbieterin und ein zweiter Teil bei einem zweitem Cloud-Anbieter oder -Anbieterin. Auch Technologie wie »Bursting«, also die temporäre Zuhilfenahme von Cloud-Ressourcen zur Absicherung von Lastspitzen, gilt es im Monitoring mit aufzunehmen. Hier muss das Monitoring aufzeichnen, wo und wann die verschiedenen Teile der Applikation laufen. Das Monitoring muss dabei der Applikation automatisch folgen.

Anwendungsbasiertes Monitoring

In der traditionellen IT-Welt dominiert das Monitoring der Ressourcen (CPU / Speicher / Netzwerk). Container-basierte Umgebungen lassen sich dort nur bedingt integrieren, da die Agilität und fehlende Hoheit über die Infrastruktur nicht mit einer statischen Monitoring-Lösung zu erfassen ist. Hier zeigt sich allerdings auch der Trend, dass Anwendungen selbst geeignete Monitoring-Schnittstellen per API bereitstellen, die dann ein direktes Monitoring erlauben (Beispiel Prometheus als OpenSource Projekt). In einer hybriden Umgebung ist zu beachten, dass für das Monitoring alle Komponenten (lokal / dezentral) integriert werden müssen, um somit einen gesamtheitlichen Blick auf den Zustand zu ermöglichen. Existierende Monitoring Lösungen sollten beim Einsatz einer hybriden Infrastruktur entsprechend erweitert oder komplett überarbeitet werden.

Netzwerk- und Firewall-Monitoring

Hybrid IT lebt auf der Basis eines sicheren und alle Services miteinander verbindenden Netzwerkes. Klare, vertraglich vereinbarte Trennungen regeln auch hier die Zuständigkeit des Netzwerk-Monitorings. Den Umfang von Netzwerk-Monitoring bilden dabei altbekannte IT-Aufgaben: Verfügbarkeit der Netze, Verschlüsselungen, Firewall Einstellungen, Intrusion Detection, Auslastungs- und Bandbreiten-Monitoring, End-to-End-Funktion sowie Latenzzeiten, um nur einige zu nennen.

Security

Das häufigere Verschieben zwischen On-Premises- und XaaS-Bereichen erfordert es, ein Security-Framework aufzubauen, das die Sicherheitsanforderungen der genutzten Bereiche definiert. Das Monitoring-Tool sollte automatisiert diesen Rahmen auf Einhaltung prüfen. Die neu entstehenden Container-Applikationen müssen ebenfalls den Security-Richtlinien der DevOps- (DevSecOPS-) Entwicklung gerecht werden und diesen folgen. Siehe hierzu auch den Punkt Anwendungsbasiertes Monitoring.

Weit verbreitete Security Mechanismen sind:

- **SIEM**
  »Security Information and Event Management« kombiniert die zwei Konzepte »Security Information Management« und »Security Event Management« für die Echtzeitanalyse von Sicherheitsalarmen aus den Quellen Anwendungen und Netzwerkkomponenten

- **FRAUD Detection**
  Fraud Detection und Fraud Prevention sind wesentliche Elemente für jedes strategisches Betrugsrisiko-Management, die Kunden- oder Kundinnen- und Unternehmensinformationen schützen müssen. Dies geschieht durch die Echtzeit-, Fast-Echtzeit- oder Stapel-Analyse von Aktivitäten der Benutzerinnen und Benutzern sowie anderen definierten Entitäten, wie zum Beispiel Bots. Im Hintergrund laufen serverbasierte Prozesse, die das Zugriffs- und Verhaltensmuster der Benutzerinnen und Benutzer und Entitäten untersuchen. Die Informationen werden mit einem erwarteten Profil verglichen. Die Benutzerinnen und Benutzer selbst bemerken diese Prüfung nicht. Sie ist nicht störend, außer es liegt eine verdächtige Aktivität vor.

- **DDoS-Detection**
  Denial of Service bezeichnet in der Informationstechnik die Nichtverfügbarkeit eines Internetdienstes, der eigentlich verfügbar sein sollte. Häufigster Grund ist die Überlastung des Datennetzes oder auch Hacker Angriffe, die gezielt Services lahm legen sollen.

Erweiterte Monitoring Parameter

Cloud-basierte Lösungen ermöglichen in der kommerziellen Betrachtung einen Wechsel von Investitionskosten zu Betriebskosten (CAPEX → OPEX). Bei einem »Pay-as-you-use«-Modell sind die Kosten von der genutzten Leistung abhängig und müssen daher auch in ein Monitoring mit einbezogen werden. Service-Provider bieten hier eigene Schnittstellen und Services an. Bei dem Einsatz einer hybriden Lösung müssen daher die Kosten der lokalen (On-Premises) und der Remote-Anteile kombiniert werden. Dies ist sicherlich keine Kernaufgabe des Monitorings selbst – sie dient allerdings als Datenquelle und sollte daher in die Konzeption der Gesamtlösung mit einbezogen werden, um eine servicebasierte Verrechnung intern wie extern zu ermöglichen.

Dokumentation und Monitoring

Das Thema Dokumentation und Monitoring Report ist dabei nicht zu vernachlässigen. Die Dokumentation des Monitorings sollte einmal jährlich überprüft werden, ggf. sind Anpassungen notwendig. Monitoring-Reports sind in die regelmäßig stattfindenden IT-Reviews mit zu integrieren.

6 Cloud-Kosten-Monitoring

6 Cloud-Kosten-Monitoring

In der Public Cloud (Hyperscaler) und auch teilweise bei Service-Providern werden Kosten entweder verbrauchs- bzw. nutzungsabhängig berechnet (man spricht von »metered«, »pay per use« oder »on demand« Modellen) oder verbindlich auf eine Laufzeit (man spricht von »reserved«, »committed term«) kalkuliert. Bei Compute-Leistung kann man meist zwischen den Modellen wählen, Netzwerkverkehr ist dagegen selten zum Festpreis (capped) möglich.

Dies sind für sich schon große Unterschiede zu On-Premises-Kosten, da hier Systeme fast ausschließlich gekauft und auf 36 Monate oder länger abgeschrieben werden.

Im eigenen Rechenzentrum vor Ort, ist man es gewohnt, dass die meisten Kosten in der Anschaffung von IT-Ressourcen inklusive der Software bekannt sind oder man diese als einen festen Kostenblock im Blick hat. Die IT-Fachbereiche können ganz frei auf der Infrastruktur-Umgebung agieren, ohne sich Gedanken machen zu müssen, dass so schnell neue, unbekannte Infrastrukturkosten entstehen könnten.

Anders aber funktioniert es in der Cloud. Hier fallen beim Buchen neuer Ressourcen üblicherweise keine Einmalkosten an. Bei XaaS entstehen jeweils variable, laufende Service-Kosten, die nicht immer sofort transparent sind. Beispielsweise verursacht bei Auswahl einer festen oder dynamischen IP-Adresse nur die feste IP-Adresse zusätzliche Kosten. Wenn zudem unterschiedliche Abteilungen, interne oder externe Kundinnen und Kunden frei und unkontrolliert Cloud-Ressourcen buchen, verliert man sehr schnell den Überblick und erlebt Überraschungen, wenn die Abrechnung kommt. Verantwortliche sollten immer wissen, durch wen, wofür, wann, wie und an welchen Stellen Ihr Cloud-Budget genutzt wird.

Auch wenn zum Beispiel virtuelle Maschinen für Tests deployed und anschließend wieder gelöscht werden, bleiben meistens Restkomponenten, bzw. feste gebuchte Ressourcen in der Cloud bestehen. Oft sind es die Disk und/oder Netzwerk-Ressourcen. Diese werden nie wieder genutzt, verursachen aber weiterhin versteckte Kosten. Diese können sich bei größeren Unternehmen sehr schnell summieren und man wundert sich, warum die Cloud-Kosten unkontrolliert schleichend und stetig ansteigen.

Zudem können Entwickler und Entwicklerinnen schnell agieren und nach Bedarf Cloud-Ressourcen variabel nutzen. Meist verwenden sie dazu S. 30 cloud-native Technologien wie Microservices und Container und haben ggf. keinen vollständigen Überblick über die erzeugten Kosten.

Um daher die Kosten für die Nutzung einer (Multi-) Cloud im Blick zu behalten, sind neue Ansätze zur Verwaltung und Optimierung dieser Kosten erforderlich. Die manuelle Verwaltung der Cloud-Kosten ist weder einfach noch skalierbar, weshalb ein entsprechendes Tool genutzt werden sollte.

Folgende Punkte sind dabei besonders wichtig und bei Cloud Monitoring zu empfehlen:

- Identifizieren unnötiger Kosten, Echtzeit-Kostenüberwachung in öffentlichen, privaten, hybriden und Multi-Cloud-Umgebungen, einschließlich der Kosten für Container
- Einsparungsempfehlungen für Workloads und die dahinterstehende Datennutzung, einschließlich Rightsizing/Größen-Anpassungen
- Anomalieerkennungs-, Budgetierungs- und Prognosefunktionen, mit denen Teams Überraschungen bei Cloud-Ausgaben vermeiden können
- Festlegen von Schwellwerten und Budget-Benachrichtigungen
- Unterstützung von Multi-Cloud-Kostentransparenz in allen genutzten Cloud Umgebungen
- Granularität bei Einblick in die Kosten von Cloud-Ressourcen nach »genutzt«, »ungenutzt« oder »nicht zugewiesen«
- Genaue Zuordnung von Kosten an Fachbereiche, Projekte oder Organisationen um diesen die Möglichkeit zu geben, die Kostenauswirkungen ihrer Arbeit zu erkennen. Bei manchen Tools kann es vorkommen, dass die gebuchten Services, Ressourcen mit Tags bezeichnet werden müssen, damit sie besser den Fachbereichen zugeordnet werden können.

Die Gesamtkosten sind natürlich sichtbar, aber es ist selten offensichtlich, wie genau sie sich zusammensetzen. Und damit ist es auch nicht gut nachvollziehbar, welche Bereiche kosteneffizient arbeiten und welche ungeahnt zu hohe Kosten verursachen oder sogar unterfinanziert sind.

In jedem Fall ist Cloud Monitoring der erste Schritt auf dem Weg, Transparenz darüber zu erhalten, wo Geld ausgegeben wird. Die intelligente Cloud-Kostenoptimierung ist das zweite Standbein bei solchen Vorhaben.

Um überhöhte Kosten schon vorab in der Cloud zu vermeiden, sollte jedes Projekt oder jede Ressourcen-Buchung vorher genau in einem geplanten Konzept durchdacht worden sein. Hierbei ist es wichtig, die Ressourcen optimiert und im Vorfeld richtig zu dimensionieren, entlang der Fragen »Welchen Anspruch an die Auslastung habe ich genau mit dieser Anwendung?« und »Wie sollen die Berechtigungen, das Netzwerk, etc. sein?«. Dies vermeidet Korrekturen im Nachhinein.

Ein Beispiel für erfolgreiches Kosten-Monitoring könnte das Aufzeigen von Einsparpotenzialen sein. Ggf. kann dann die Nutzung von »reserved Modellen« in Betracht gezogen werden. Diese sind aufgrund der verbindlichen Laufzeit deutlich günstiger als die on-demand-Verwendung bei gleicher Laufzeit. Wenn man also genau weiß, dass bestimmte Anwendungen für eine lange Laufzeit geplant sind, lassen sich 40-70 Prozent der Kosten einsparen. Die Bedingung hierbei ist, dass man sich mit dieser Ressourcen-Kategorie auf drei oder fünf Jahre festlegt.

7 Anomalie-Monitoring und Alarmierung

7 Anomalie-Monitoring und Alarmierung

Was sind Anomalien?

Bei Wikipedia findet man als Erklärung einer Anomalie den Begriff der Unregelmäßigkeit. Weiterführend, im Themenbereich der Softwaretests beschreibt eine Anomalie einen »Zustand, der vom Erwarteten abweicht«. Auch im Bereich komplexer IT-Umgebungen kommen solche Abweichungen vor. Dabei ist aber nicht gemeint, dass sich die Auslastungen bestimmter Systeme zu definierten Tages- oder Nachtzeiten verändern (Beispiele: »Boot-Storm am Montagmorgen«, »Hohe Netzwerklast zu Backup-Zeiten«, »Compute-Last am Monatsende wegen diverser Rechnungsläufe«; dies sind eher typische regelmäßige Lasten), sondern tatsächlich der erwartete Zustand des Systems in einem typischen Nutzungszyklus (möglicherweise innerhalb eines Monats) vom tatsächlichen Zustand stark abweicht.

Beispiele dafür könnten sein:

- Einmalige Lösch-Aktivitäten großer Datenmengen
- Extrem hohe Anzahl von wiederkehrenden Fehlermeldungen
- System-Peaks zu untypischen Zeiten

Solche Situationen können oft auch eine Indikation von Cyber-Attacken sein, bei denen es sich lohnt, schnell genauer hinzuschauen, um die tatsächliche Ursache dieser Indikation festzustellen und dementsprechend sinnvoll zu reagieren.

Die drei oben genannten Beispiele können tatsächlich als erste Anzeichen von verschiedenen Bedrohungen verstanden werden:

Anzeichen	Risiko
Einmalige Lösch-Aktivitäten großer Datenmengen	Sabotage von Innen
Extrem hohe Anzahl von wiederkehrenden Fehlermeldungen	Unerwünschter Datenzugriff von außen
System-Peaks zu untypischen Zeiten	Unerwünschte Zweckentfremdung von IT-Ressourcen oder unerwünschter Datentransfer

Die Auswirkungen solcher Bedrohungen können, gerade in der heutigen Zeit bei vermehrter Öffnung der IT-Ressourcen nach außen, sehr hoch sein und kommen immer häufiger vor. Die Auswirkungen können erheblich sein: Sie reichen vom Systemdefekt über Datenverlust bis hin zu Reputationsverlust und Schadensersatzzahlungen. Dies ist der Grund, dass moderne Monitoring- und Alarmierungssysteme mit einer Anomalieerkennung ausgestattet sind. Denn wenn solche Anomalien schnell vom Monitoringsystem erkannt werden, kann auch schnell und ggf. automa-

tisiert darauf reagiert werden und ein möglicher Schaden abgewendet werden. Der Aufwand lohnt sich im Vergleich zum Risiko.

Die automatisierte Reaktion spielt dabei eine wichtige Rolle, denn Cyber-Attacken zum Beispiel passieren oft zu Zeiten, die nicht die typischen Arbeitszeiten der IT-Mannschaft (Persona »Operations«, Siehe Kapitel 1.2) widerspiegeln. Dennoch wollen der verantwortliche IT-Betriebsmanager oder Datenschutzbeauftragte (Persona »Datenschutz«, Siehe Kapitel 1.2) die Gewissheit haben, dass die Unternehmensdaten jederzeit sicher und geschützt sind.

Anomalien lassen sich in unterschiedliche Typen klassifizieren, die sich im Groben entlang des OSI-Schichtenmodells wiederfinden.

Zunächst gibt es auf der Ebene der IT-Komponenten (Server, Storage, Switche, Router, Firewall, Loadbalancer, ...), klassische Lastanaomalien. Hierbei wird ein typisches Lastprofil mit einem aktuellen Lastprofil verglichen und stellt sich ggf. so als Unregelmäßigkeit heraus. Beispiel: Die CPU-Last der Unternehmensfirewall steigt an einem untypischen Zeitpunkt von 20 Prozent auf 80 Prozent.

Daneben gibt es die Ebenen der Applikationen, die auf Basis von Zugriffen (zum Beispiel mittels Protokolle) Anomalien erfahren können. Beispiel: Die durchschnittlichen Datenbankabfragen pro Minute erhöhen sich um 300 Prozent.

Darüber hinaus können Anomalien auch auf Basis eines Benutzer- oder einer Benutzerinverhalten auftreten. Beispiel: Das Benutzerkonto 471100 initiierte in den letzten 60 Minuten Datei-Leseoperationen, die um das 10-fache so hoch waren wie im Durchschnitt der letzten Woche.

Erkennen von Anomalien

Wichtig bei der Anomalie-Erkennung ist, dass alle Anomalien für sich gesehen auch sogenannte False-Positives sein können, das heisst keine tatsächliche Bedrohung darstellen müssen. Oft wird die Bedrohung erst offensichtlich, wenn mehrere solcher Anomalien gleichzeitig auftreten.

Zunächst müssen solche Anomalien aber einzeln erkannt und bewertet werden. Wie wird dies bewerkstelligt?

In der Vergangenheit wurde die Definition einer Anomalie »von Hand« durchgeführt. Das bedeutet, dass einzelne Indikatoren der jeweiligen Komponenten herangezogenund drei individuelle Werte definiert wurden: Normalwert, Schwellwert-Maximum, Schwellwert-Minimum. Mittels des Monitorings dieser Werte könnten dann Alarmierungen definiert werden, die bei Überschreiten des Schwellwert-Maximums beziehungsweise das Unterschreiten des Schwellwert-Minimums ausgelöst werden. Dieses Vorgehen war und ist offensichtlich extrem zeitaufwändig, sehr starr und bedarf großer Erfahrung mit den entsprechenden Systemen und deren Lastverhalten.

Heutige Verfahren setzen intelligentere Mechanismen ein. Der Ansatz der modernen Lösungen ist es, dass die verantwortlichen Administratoren und Administratorinnen die Definition einer Anomalie nicht mehr selbst vornehmen, sondern durch ein Werkzeug, das selbstständig die Umgebung »lernt«. In der Regel benötigen sie für diesen Lernvorgang, der durch Künstliche Intelligenz (KI; besser noch: ML = Machine Learning) unterstützt wird, einige Wochen, um die tatsächlichen Verhaltens- und Auslastungsmuster kennen zu lernen.

Anschließend wird ein jeweiliger Bewegungs-»Korridor« erstellt, in dessen Rahmen sich das individuelle Normalverhalten üblicherweise bewegen wird. Dieser Automatismus ist heutzutage genauer und schneller, als ein manuelles Erstellen eines Umgebungsprofils. Außerdem ist es frei von eventuellen »Tunnelblicken« der IT-Betriebsmannschaft und berücksichtigt die entsprechenden Best-Practices pro Instanz. Alle so definierten Korridore können auch individuell angepasst werden. Zusammengefasst lernen die neuen Monitoring- und Alamierungswerkzeuge das »normale« Verhalten einer IT Umgebung, stellen dieses als Nutzungsbasis dar und erkennen darauf basierend Anomalien und alarmieren entsprechend.

Darüber hinaus sind sie in der Lage, Korrelationen einzelner Korridore zu bilden um damit Zusammenhänge aufzuzeigen, die bisher noch nicht bekannt waren. Das ist nachvollziehbarerweise nur mit eine KI-Integration möglich. Denn die Erkenntnis »etwas ist anders, als es gestern war«, ohne dass vordefinierte Schwellenwerte überschritten worden sind, lässt sich über die bisherigen Ansätze nicht realisieren.

Reagieren

Die übliche, ausgelöste Aktion eines Monitoring- und Alarmierungssystems ist (wie der Name schon ausdrückt) die Absetzung eines Alarms. Aber was nützt einem IT-Verantwortlichen oder seinen Vorgesetzten ein Alarm, wenn die notwendige Reaktion ausbleibt oder zu spät erfolgt. Zum Beispiel erfolgt eine Ransomware-Attacke innerhalb von wenigen Minuten bis Stunden. Und da kann ein reiner Alarmierungsmechanismus nur wenig Besserung versprechen. Deshalb bieten einige Monitoring-Werkzeuge direkt die Option, tatsächliche Aktivitäten auszulösen, welche die Gefahr direkt abschalten oder eindämmen. Beispiel: Sobald eine Ransomware Attacke erkannt wird, wird automatisiert ein Storage-Snapshot ausgelöst und der entsprechenden Benutzer oder Benutzerinnen Zugriff gesperrt. Der alarmierte Systemadministrator oder -administratorin kann sich dann im Nachgang schneller mit dem Wiederherstellen des Ursprungszustandes befassen (z. B. Endgeräte »säubern«, Snapshot-Wiederherstellung).

Relevanz

Die Relevanz von modernen Monitoring- und Alamierungswerkzeugen ist in den letzten Jahren stark gestiegen. Das hat mit den verschiedenen Trends rund um die IT zu tun.

Die Rechenzentrums-Infrastrukturen werden tendenziell im Backend mit Cloud-Diensten ergänzt oder ersetzt. Beispiele: Cloud Anbindung für Backup, Disaster Recovery, Web Services, Cloud-Bursting. Dies hat zur Konsequenz, dass die Gesamt-Infrastrukturen durch den Einzug von hybriden Cloud-Architekturen immer komplexer werden.

Eine Anomalieerkennung im Kontext hybrider Architekturen wird schon allein dadurch notwendig, dass die entstandene Komplexität kaum noch mit den bisherigen Werkzeugen und Mechanismen abzudecken ist. Zumal für den Aufbau eines solchen Regelwerks im tagtäglichen IT-Betrieb schlicht die Zeit fehlt. Spätestens, wenn es um Web-Services in der Cloud geht, die einen direkten Endkunden- oder Endkundinnen-Zugriff anbieten, sollte die Verfügbarkeit und die Stabilität (kurz: das SLA) dieser Services sichergestellt werden, wozu eben auch die Anomalie-Erkennung, aber auch die Korrelationen mit anderen Monitoring-Daten einen wesentlichen Teil beiträgt.

Auch im Frontend öffnen sich die IT-Infrastrukturen immer mehr. Beispiele: Bring-your-own-Device, Home-Office, Online-Apps. Sie erhöhen damit das Angriffsrisiko für Angriffe von innen heraus. Beispiele hierfür sind Rough-Admin und Ransomware. Die negativen Auswirkungen eines Ransomware Angriffs sind hoch: Diese umfassen neben den reinen IT-Ausfall- und Bereinigungskosten auch die Kosten für eventuelle Strafen, Imageschäden, Umsatzeinbußen oder Erpressungskosten[1].

Alleine die durchschnittlichen, IT-bezogenen Kosten sind erschreckend[2]:

- Zunahme von Ransomware-Vorfällen pro Jahr: 41 Prozent
- Durchschnittliche Ausfallzeit pro Vorfall: 16 Tage
- Anteil der Unternehmen, die von einem Angriff betroffen waren und alle oder einen Teil ihrer Daten dauerhaft verloren haben: 67 Prozent
- Durchschnittliche Kosten der Ausfallzeit: 300.000 US-Dollar pro Stunde
- Gesamtkosten für Unternehmen im Jahr 2019 aufgrund von bösartigen Aktivitäten: 170 Milliarden US-Dollar

1 Siehe hierzu auch Bitkom-Studie »Spionage, Sabotage und Datendiebstahl – Wirtschaftsschutz in der vernetzten Welt« Studienbericht 2020, abrufbar unter; ↗ **https://www.bitkom.org/sites/default/files/2020-02/200211_bitkom_studie_wirtschaftsschutz_2020_final.pdf**

2 Quelle: ↗ **https://cloud.netapp.com/ci-sde-plp-cloud-secure-info-trial**

Empfehlung

Der Trend zu einer, wie auch immer im Detail ausgearbeiteten, hybriden IT-Infrastruktur ist real und wird auf absehbare Zeit auch nicht enden. Dadurch steigt unmittelbar die Komplexität individueller IT-Umgebungen, auch wenn dies im Rahmen einer 100 Prozent-Cloud-Strategie eventuell nur eine Übergangssituation darstellt. Die entsprechenden Veränderungen verschlingen außerdem die Zeit der Administratoren und Administratorinnen, wodurch die Pflege der bestehenden Monitoring- und Alarmierungssysteme ebenfalls leidet.

Beides führt quasi zwangsläufig zu der Empfehlung, sein Bestandssystem um die hier erläuterte Anomalieerkennung zu aktualisieren beziehungsweise zu ergänzen, da hier erhebliches Potential zur Risikovermeidung und damit ggf. Kostenvermeidung besteht.

Neben der automatischen Erkennung von Anomalien empfehlen wir zusätzlich die Abwägung des Einsatzes von Schutz-Automatismen wie zum Beispiel die Erstellung von Snapshots, das Sperren von Zugriffen und die Anpassung von Ressourcen.

Je nach Branche gelten neben den hier erläuterten Aspekten aber auch spezifische IT-Regularien (siehe Kapitel Compliance-Monitoring), die im jeweiligen Rahmen die Befolgung von Auditing-, Reporting-, Security- und Compliance-Anforderungen voraussetzen. Auch aus diesem Aspekt heraus kann sich der Einsatz eines modernen hybriden IT-Monitoring- und Alarmierungssystems (inklusive Anomalieerkennung) lohnen.

8 Monitoring im operativen Betrieb

8 Monitoring im operativen Betrieb

Die unterschiedlichen Monitoring-Technologien und -Vorgehensweisen im operativen Betrieb vor allem im Bereich von hybriden IT-Umgebungen – umzusetzen, stellt Betreiber vor verschiedene Herausforderungen. Neben der Frage, welche Technologie für welchen Anwendungsfall die Richtige ist, geht es auch um die Frage, wie diese nach der Auswahl implementiert und möglichst effektiv genutzt werden kann sowie welche Informationen diese zur Verfügung stellt.

Im operativen Betrieb liegt der Fokus ganz klar darauf, die in der eigenen IT-Landschaft, ob im eigenen Rechenzentrum, in der Public Cloud oder in einem hybriden Format entstehenden Störungen beziehungsweise Incidents, möglichst schnell und zielgenau zu erkennen und die für die Lösung geeigneten Maßnahmen, entsprechend zeiteffizient, einzuleiten. Hierbei sollten die entsprechenden Performance-Metriken mit einbezogen werden, wobei das Sicherstellen der Arbeitsfähigkeit der eigenen Mitarbeiterinnen und Mitarbeiter die höchste Priorität hat. Dies ist nur möglich, wenn entsprechender Einblick in die Umgebung und die Applikation gegeben ist sowie eine Recherche der notwendigen Informationen selbst nicht bereits zu viel Zeit in Anspruch nimmt.

Application Performance Management

Der zentrale Unterschied zwischen dem Application Performance Monitoring und der nächsten Stufe, dem Application Performance Management, ist der Scope der beim Monitoring gewonnenen Informationen und Erkenntnisse und der daraus abgeleiteten Aktionen. Diese Aktionen können Auto-Remediation-Tasks sein, also automatisch durchgeführte Aktionen wie Scaling, das Neustarten von Containern und Services, Re-Deployments oder das Triggern von fachspezifischen APIs.

Betrieb des Monitorings

Die perspektivische Entwicklung und Nutzung der verschiedenen Monitoring-Technologien ist klar darauf ausgerichtet, in einem Software-as-a-Service-Betriebsmodell verwendet zu werden. Hierbei muss für die zu überwachenden Systeme sichergestellt werden, dass diese eine sichere Verbindung zu den Schnittstellen der Monitoring-Systeme herstellen können. Dies kann entweder über eine direkte Verbindung zwischen dem auf dem System installierten Agent und der Schnittstelle erfolgen oder, sofern diese Verbindung auf direktem Weg nicht möglich ist, über Proxies. Abgesehen von der initialen Installation der Agents, dem Sicherstellen der Kommunikation und der Konfiguration der für die Monitoring-Lösung spezifischen Informationen zur Prozessierung und Anzeige der Daten, muss hier nichts weiter betrachtet werden.

Der Großteil der Monitoring-Lösungen kann alternativ auch On-Premises betrieben und genutzt werden. Aufgrund der für das eigentliche Monitoring und der Auswertung der gesammelten Informationen notwendigen Ressourcen, der eingeschränkten Skalierungsmöglichkeiten und

den damit verbundenen Kosten ist es notwendig, diese Variante kritisch zu betrachten und zu bewerten. Die Nutzung der Systeme erfolgt nach der Inbetriebnahme nach dem gleichen Prinzip wie die Nutzung der Software-as-a-Service-Variante.

Funktionsweise

Das Monitoring funktioniert entweder aktiv oder passiv und muss in entsprechenden Sinks beziehungsweise zentralen Systemen zur Analyse zusammengeführt werden. Dies passiert in Form von Agents auf den zu überwachenden Systemen, auf dem Orchestrator beziehungsweise Hypervisor der Services und Applikationen oder Agents beziehungsweise Workern auf den Monitoring-Systemen.

Ist es nicht möglich, Systeme direkt zu überwachen, werden die zur Verfügung gestellten APIs abgefragt, um Metriken auszulesen. Diese können entweder Schnittstellen in den Betriebssystemen direkt sein, oder standardisiert zur Verfügung gestellte Schnittstellen, die auf das unterliegende Betriebssystem zugreifen. Für Web-Technologien können neben den Hosting-Systemen auch die Clients, zum Beispiel mit Hilfe von JavaScript/TypeScript, in das Monitoring aufgenommen werden.

Die hiermit gesammelten Informationen werden kontinuierlich verwendet, um Baselines zu bilden und zu lernen, wie die Umgebungen standardmäßig funktionieren. Abweichungen von diesen Baselines, KI-gestützt oder manuell beziehungsweise automatisiert bestimmt, erstellen dann Incidents, die in den hierfür verantwortlichen Bereich weitergeleitet werden. Durch die regelmäßig wechselnden Anforderungen ist es notwendig, diese Baselines beziehungsweise die damit verbundenen Schwellwerte und Metriken regelmäßig an die eigenen Bedürfnisse anzupassen.

Full Stack Observability

Der Begriff Full Stack Observability sagt aus, dass in alle unterschiedlichen Bereiche, ob organisatorisch oder technisch, Einblick genommen werden kann. Die daraus entstehenden Informationen werden miteinander verknüpft und für alle Beteiligten, individuell oder aus Sicht eines Teams, zur Verfügung gestellt, damit ein gemeinsames Verständnis gebildet werden kann.

Die zu Anfang definierten Personas bzw. Rollen sind elementar wichtig dafür. Die Einsicht kann kann zum Beispiel aus folgenden Rollen, welche für die Bereiche des operativen Monitoring am wichtigsten sind, entstehen:

- Development
- Operations (IT-Betrieb)
- DevOps

Um alle Vorzüge der Full-Stack-Observability nutzen zu können, ist es für die unterschiedlichen Rollen notwendig, entsprechendes Wissen in den anderen Bereichen zu haben.

Somit sind alle Metriken, die durch das Monitoring aufgezeichnet werden wichtig, da diese in eine Gesamt-Metrik bzw. die Performance einfließen und so die Ziele der jeweiligen Bereichen verschmelzen beziehungsweise geteilt werden. Für das gemeinsame Verständnis können einfache Hilfsmittel wie Dashboards, gefiltert für den eigenen Bereich, verwendet werden. Diese greifen aber grundsätzlich auf die große Informationsbasis zu, die allen über die jeweiligen Tools zur Verfügung steht.

Root Cause Analysis

Das wichtigste Werkzeug für das moderne Monitoring ist die Root-Cause-Analysis. Im Gegensatz zum traditionellen Ansatz im Bereich des Monitorings und der nachgelagerten Auswertung der Informationen wird hier mit neuen Technologien ein weitestgehend automatisiertes Verfahren etabliert. Das Ziel ist nicht mehr, den Fehler selbst zu finden, sondern den Fehler beziehungsweise den Auslöser des Fehlers direkt aufgezeigt zu bekommen.

Verwendet man den traditionellen Ansatz, so ist es erforderlich, diejenigen Informationen aus den eigenen Applikationen zu extrahieren, die für den Betrieb und die Fehlersuche notwendig sind. Dies muss für das überwachte System sowie alle angrenzenden, in Wechselwirkungen stehenden Systeme durchgeführt werden. Es wird also ein eigenes Framework aus Informationen aufgebaut, die für sich selber stehen und verknüpft werden müssen. Tritt hier nun ein Fehler auf und die Informationen sind nicht bereits in Relation zueinander gestellt, kann es im schlechtesten Fall notwendig sein, die maximale Anzahl von Systemen zu durchsuchen und bei der Problemlösung selbst erst die Relationen herzustellen. In diesem Fall wird kostbare Zeit verloren.

Verwendet man den modernen Ansatz, wird das Monitoring der Systeme durch einen KI-Ansatz durchgeführt. Im Idealfall werden alle relevanten Daten geloggt und die KI wertet aus, was wichtig ist und welche Systeme in Relation stehen. Über einen längeren Zeitraum wird zudem automatisch bestimmt, welchen Zustand die Systeme im Normalfall haben. Tritt nun ein Fehler auf, kann die Kette der Fehler direkt bestimmt und an die Anwender und Anwenderinnen des Monitoring-Tools weitergegeben werden. Die Zeit für die Analyse der Log-Dateien oder Systeme fällt somit weg oder wird auf ein Minimum beschränkt. Durch das Lernen der Zustände der Systeme können Probleme zusätzlich proaktiv erkannt werden. Hierzu gehören Situationen wie zyklisch wiederkehrende Probleme oder ggf. in der Zukunft wahrscheinliche Ausfälle von Hardware oder Services.

Begriffserläuterungen

1. **API**
   Die Abkürzung API steht für Application Programming Interface und bezeichnet eine Programmierschnittstelle, die einen Dienst für andere Software anbietet. Sie ermöglicht es Anwendungen und Programmteilen untereinander standardisiert zu kommunizieren, indem sie die Zugangspunkte regelt.

2. **Virtuelle Maschine (VM)**
   Ein virtueller Computer, welcher denselben Funktionsumfang bietet wie ein physischer Rechner. Die Abstraktion erfolgt auf Betriebssystem-Ebene und mit Hilfe eines Hypervisors, welcher die Hardware-Ressourcen des zugrundeliegenden Systems – dem sogenannten Host, aufteilt, isoliert und verwaltet.

3. **Container**
   Ein isolierter Anwendungsbehälter. Die Abstraktion erfolgt im Gegensatz zur virtuellen Maschine (siehe oben) auf Anwendungsebene, weshalb Container über die geteilte Hardware hinaus auch auf ein geteiltes Betriebssystem und Kernel zugreifen. Der Container selbst umfasst nur die Anwendung, Abhängigkeiten und Konfiguration.

4. **DevOps**
   Eine Sammlung unterschiedlicher technischer Methoden und eine Kultur zur Zusammenarbeit zwischen Software-Entwicklung und IT-Betrieb. Das Ziel dieses Prozess-Verbesserungs-Ansatzes ist es, historisches Silodenken zu überwinden. Der Begriff ist dabei eine Kombination der Wörter Development und Operations.

5. **Micro Service**
   Ein Architekturmuster, welches komplexe Software modular aus voneinander entkoppelten Diensten aufbaut. Diese kommunizieren untereinander mit Programmierschnittstellen.

Bitkom vertritt mehr als 2.000 Mitgliedsunternehmen aus der digitalen Wirtschaft. Sie erzielen allein mit IT- und Telekommunikationsleistungen jährlich Umsätze von 190 Milliarden Euro, darunter Exporte in Höhe von 50 Milliarden Euro. Die Bitkom-Mitglieder beschäftigen in Deutschland mehr als 2 Millionen Mitarbeiterinnen und Mitarbeiter. Zu den Mitgliedern zählen mehr als 1.000 Mittelständler, über 500 Startups und nahezu alle Global Player. Sie bieten Software, IT-Services, Telekommunikations- oder Internetdienste an, stellen Geräte und Bauteile her, sind im Bereich der digitalen Medien tätig oder in anderer Weise Teil der digitalen Wirtschaft. 80 Prozent der Unternehmen haben ihren Hauptsitz in Deutschland, jeweils 8 Prozent kommen aus Europa und den USA, 4 Prozent aus anderen Regionen. Bitkom fördert und treibt die digitale Transformation der deutschen Wirtschaft und setzt sich für eine breite gesellschaftliche Teilhabe an den digitalen Entwicklungen ein. Ziel ist es, Deutschland zu einem weltweit führenden Digitalstandort zu machen.

Bitkom e.V.

Albrechtstraße 10
10117 Berlin
T 030 27576-0
F 030 27576-400
bitkom@bitkom.org
www.bitkom.org