

SLOs – ein Leitfaden für Führungskräfte

Neue Paradigmen im Bereich Kundenerlebnis erfordern von Führungskräften bei Service-Level-Zielen ein Umdenken



Zusammenfassung

Moderne Anwendungen sind komplexer und stärker verteilt als je zuvor. Und ihre Performance ist für Unternehmen kritischer als je zuvor. Daher stehen Führungskräfte unter Druck, den positiven Einfluss ihrer Abteilungen bzw. Teams auf den Geschäftserfolg nachzuweisen. Glücklicherweise schaffen die gleichen Technologien, die für die Modernisierung und Verwaltung von Anwendungen genutzt werden, auch Möglichkeiten zur direkten Verknüpfung von IT-Maßnahmen mit dem geschäftlichen Nutzen. Die für die IT und Plattformen verantwortlichen Teams haben neue Wege gefunden, wie sie den Zustand ihrer Umgebungen messen, Erwartungen an ihre Anwendungen definieren und ihre Arbeit mit dem Geschäftskontext verknüpfen können – genannt Service-Level-Ziele (Service Level Objectives, SLOs). Angesichts des geschäftskritischen Charakters heutiger moderner Anwendungen müssen Führungskräfte die Aufmerksamkeit ihrer Teams auf die Definition von SLOs in Bezug auf die geschäftlichen Auswirkungen und das Kundenerlebnis richten.

In diesem Leitfaden werden häufige Strategien und Probleme bei der Definition und Implementierung von SLOs untersucht und klare Schritte hin zu Best Practices bei der Nutzung von SLOs aufgezeigt, um einen höheren geschäftlichen Nutzen zu erzielen.

Kontext

Heute ist „die Zufriedenheit der Endbenutzer mit der Anwendungsleistung und -zuverlässigkeit für den erfolgreichen digitalen Geschäftsbetrieb entscheidend“.¹ Die Bereitstellung hervorragender Endnutzererlebnisse ist jedoch schwierig, da moderne Anwendungen sowie die Plattformen und Infrastrukturen, auf denen sie ausgeführt werden, komplexer und stärker verteilt sind als je zuvor. So nannten Unternehmen auch in unserem Bericht „Multicloud – der aktuelle Stand 2021“ Komplexität als eine der wichtigsten Herausforderungen bei der Erreichung ihrer Geschäftsziele.²

Neben der Entwicklung moderner Anwendungen bereitet der „Umstieg auf eine agile Infrastruktur – inklusive Hybrid-IT, Multicloud und Containern“ den IT-Teams Schwierigkeiten: Enorme Datenmengen müssen sinnvoll verarbeitet werden und „die Brauchbarkeit herkömmlicher Tools zur Überwachung der Infrastruktur wird infrage gestellt“.³

Als Reaktion auf diese Herausforderungen haben IT- und Plattform-Teams neue Wege gefunden, den Zustand ihrer Umgebungen zu messen, Erwartungen für ihre Anwendungen zu definieren und ihre Arbeit mit einem geschäftlichen Nutzen zu verknüpfen. Und angesichts der immer höheren Erwartungen der Endnutzer ist es unerlässlich, dass Führungskräfte bei der Beurteilung ihrer Anwendungen und deren geschäftlichen Nutzen in ihren Abteilungen aussagekräftige Kennzahlen festlegen.

1 Quelle: Worldwide Application Performance Management Software Market Shares, 2020, Juni 2021, IDC #US47989021

2 Quelle: 2021 Turbonomic State of Multicloud Report, 2020 CNCF Survey

3 Quelle: Monitoring and Observability for Modern Services and Infrastructure, Juni 2020, Gartner G00720854

Die aktuelle Strategie: Arbeitsaufwendig und reaktiv

Ein häufiger Ansatz, den IT- und Plattform-Teams zur Messung des Zustands ihrer Umgebungen wählen, besteht in der Identifizierung und Konfigurierung von SLIs (Service-Level-Indikatoren), SLAs (Service-Level-Agreements) und SLOs (Service-Level-Zielen).

- **Ein Service-Level-Indikator (SLI) ist eine Messgröße, mit der Teams die Leistung eines Service in Bezug auf das Ziel messen.**
- **Ein Service-Level-Agreement (SLA) ist ein Leistungsversprechen zur Verfügbarkeit, das IT- und Plattform-Teams gegenüber Kunden oder Endbenutzern geben.**
- **Und ein Service-Level-Ziel (SLO) ist eine Vorgabe bzw. Zielsetzung auf Serviceebene, gemessen anhand eines SLI, zu deren Erreichung sich Teams verpflichten.**

SLOs sind von besonderer Bedeutung, weil sie die Kriterien und klare Erwartungen an die Leistung der Anwendungen eines Unternehmens definieren.

Heute durchlaufen die meisten Unternehmen einen zeit- und arbeitsaufwendigen manuellen Prozess der SLO-Konfiguration. IT-Teams müssen zunächst festlegen, welche Services einer Anwendung sich direkt auf das Geschäft und das Erlebnis des Endbenutzers auswirken. Daraufhin müssen sie identifizieren, welche Messgrößen sie als Service-Level-Indikatoren nutzen möchten. Häufige SLI-Messgrößen sind zum Beispiel Verfügbarkeit, Latenz oder Transaktionsdurchsatz. Sobald die geeigneten SLIs ausgewählt sind, müssen die SLO-Zielvorgabe für die jeweilige Messgröße sowie ein konkreter Messzeitraum festgelegt werden. Nach Abschluss dieses Prozesses müssen die IT-Teams wiederum Fehlerbudgets für jedes SLO erstellen und diese mit einem schwellenwertbasierten Warnmeldesystem verknüpfen. Viele Unternehmen nutzen diesen schwellenwertbasierten Ansatz, weil es unmöglich ist, die Leistung ihrer Anwendungen täglich und rund um die Uhr manuell zu überwachen.

Leider löst die Festlegung von SLO-Schwellenwerten nicht die Frage der Leistung. Diese IT-Strategie ist ineffektiv, weil sie für heutige agile Infrastrukturen, in denen Anwendungen in Multicloud- oder containerisierten Umgebungen ausgeführt werden, zu reaktiv ist. Wenn etwas bei einem Service fehlschlägt und ein Schwellenwert oder eine Warnmeldung ausgelöst wird, ist es bereits zu einer Leistungseinbuße für diesen Service mit einem schlechten Endbenutzererlebnis gekommen.

IT-Teams haben in der Vergangenheit versucht, ihr System aus Schwellenwerten/Warnmeldungen zu verbessern, indem sie Horizontal Pod Autoscaler (HPAs) für Anwendungen in containerisierten Umgebungen implementiert haben. Jedoch scheitern auch HPAs dabei, ein erstklassiges Endbenutzererlebnis zu gewährleisten und Leistungseinbußen zu verhindern. Ähnlich wie beim SLO-Konfigurationsprozess müssen IT-Teams zur Abstimmung der automatischen Skalierung auf den Ressourcenbedarf Messgrößen identifizieren, die diesen Bedarf am besten repräsentieren, sie müssen Ziele konfigurieren, Schwellenwerte festlegen und alles testen. Dieser Prozess muss für jeden Service einer Anwendung wiederholt werden. Da einige Anwendungen Hunderte von verschiedenen Services beinhalten, ist die Implementierung von HPA im richtigen Maß sehr schwierig. Darüber hinaus verlassen sich verschiedene HPA-Strategien nach wie vor auf Schwellenwerte und sind nicht miteinander korreliert oder kompatibel. Dadurch kann sich die Skalierung von einem Service negativ auf einen anderen auswirken. Abschließend handelt es sich hierbei nicht um einen einmaligen Vorgang, da wirksame HPA-Skalierungsstrategien kontinuierlich neu konfiguriert und überwacht werden müssen.



Die richtige Perspektive für SLOs

Definieren Sie SLIs und SLOs in Bezug auf das Kundenerlebnis.

Laut Branchenexperten „genießt die Konzentration auf das Management von Systemen und Anwendungen zur Optimierung der Endnutzenerfahrung eine hohe Priorität, weil **die schnelle Performance und 100-prozentige Uptime Mindestanforderungen für den Erfolg eines digitalen Geschäfts darstellen**“.⁴

Weil eine schnelle Performance und die Uptime heute Mindestanforderungen darstellen, können Unternehmen keine Zeit mit der Erfassung von Daten vergeuden, die bei der Bewertung ihrer Anwendungsleistung keine direkten Schlüsse auf das Endbenutzererlebnis zulassen.

Die richtigen Daten zu erfassen, ist keine einfache Aufgabe. So ist zum Beispiel die Verfügbarkeit eine häufig genutzte Messgröße, doch stellt sie keinen direkten Ausdruck der Leistung dar: Eine Anwendung kann hochverfügbar sein und dennoch unter Ressourcenengpässen und Leistungseinbußen leiden. Es gibt zahlreiche verschiedene Messgrößen, die Unternehmen für die Leistungsmessung nutzen können. Daher ist es wichtig, dass IT-Abteilung und die jeweilige Fachabteilung gemeinsam entscheiden, welche Daten erfasst und in Berichte aufgenommen werden sollen. Wie Gartner schreibt: „Die Auswahl repräsentativer und aussagekräftiger SLIs ist von kritischer Bedeutung. In den meisten Fällen ist eine Messgröße auf Infrastrukturbasis (wie „verfügbarer Speicher“ oder „Anteil freier Workerknoten“) nicht aussagekräftig, weil sie für die Benutzer eines Service nicht interessant ist. Wählen Sie SLIs, die eine direkte Messung des Service-Erlebnisses Ihrer Benutzer darstellen.“⁵

Einfach gesagt sollten SLOs als Werkzeug genutzt werden, das das Unternehmen über die geschäftlich relevante korrekte Funktionsweise seiner Anwendungen informiert. Unternehmen, die ein aussagekräftiges SLO auf der richtigen Ebene ihrer Anwendung definiert haben, wissen bei Leistungsproblemen ihrer Anwendung genau, welche Maßnahmen am effektivsten für eine Wiederherstellung der Leistung sind. Unterschiedliche Unternehmen müssen je nach ihren Geschäftsbereichen auch unterschiedliche Messgrößen messen. Mit der Abkehr von monolithischen Anwendungsarchitekturen zugunsten von modernen und stärker verteilten Architekturen verlieren herkömmliche Leistungsindikatoren wie hohe Speicherkapazitäten und CPU-Auslastung ihren Nutzen. Stattdessen sollten Unternehmen versuchen, SLIs/SLOs in Bezug auf Messgrößen zu definieren, die typischerweise für das Geschäft relevant sind – beispielsweise Antwortzeiten und Transaktionsdurchsatz. Diese Messgrößen stellen eine direktere Messung der Leistung dar. Beispielsweise können IT-Teams durch die Definition von SLOs für den Transaktionsdurchsatz genau bestimmen, wie Anfragen an den einzelnen Pod bzw. die einzelne VM bearbeitet werden. Diese Messgröße misst die Leistung direkter als SLO-Definitionen auf Basis von Auslastungsmessgrößen, die dann lose mit Messgrößen wie Antwortzeiten und Durchsatz assoziiert werden.

⁴ Quelle: Worldwide Application Performance Management Software Market Shares, 2020, Juni 2021, IDC #US47989021

⁵ Solution Path for Modern Infrastructure and Application Monitoring, Juni 2019, Gartner

Auch wenn Antwortzeiten und Transaktionsdurchsatz zu den häufigsten Bewertungsarten für das Kundenerlebnis und die Anwendungsleistung gehören, funktionieren diese Messgrößen doch nicht für alle Unternehmen. So würde zum Beispiel ein VDI-Serviceanbieter (Virtual Desktop Infrastructure) keine SLOs für den Transaktionsdurchsatz definieren, sondern stattdessen für Support-Tickets. Support-Tickets wären hier eine effektivere Messgröße für ein SLO, denn dies bedeutet eine direkte Messung der Leistung virtueller Desktops. Schließlich müssen SLIs/SLOs auch spezifisch für das jeweilige Unternehmen und die jeweilige Fachabteilung definiert werden.

Eine kontinuierliche Analyse der sich verändernden Anwendungsanforderungen ist erforderlich.

Moderne Anwendungen und Infrastrukturen sind elastisch und der Ressourcenbedarf ist dynamisch. Aufgrund dieser ständigen Veränderungen ist es sehr schwierig, die Beziehungen zwischen verschiedenen Quellen und Datentypen zu verstehen. Unternehmen nutzen häufig verschiedene Tools zur Überwachung verschiedener Schichten ihres Stacks über verschiedene Teams hinweg – und all das, während sie versuchen, verschiedene Seiten desselben Problems zu lösen. Diese mangelnde Koordination führt bei der Suche nach Problemursachen zu zahlreichen falschen Hinweisen. Zudem müssen die Daten kontinuierlich erfasst werden, denn andernfalls müssen die IT- und Plattformteams raten, welche Daten sie zur Lösung eines Problems benötigen. Wie Gartner in einem Bericht zur Überwachung und Beobachtbarkeit von 2020 anmerkt, „muss die Erfassung der möglichen Ursachendaten kontinuierlich stattfinden, denn bei einer Aktivierung nur in Reaktion auf ein Symptom kann die Ursache völlig falsch bewertet werden“.⁶

Um dieses Dilemma zu vermeiden, benötigen Unternehmen ein zentrales Repository, das Daten aus jeder Schicht eines Anwendungsstacks aggregieren und korrelieren kann. Mit diesem System können Unternehmen ihre kontinuierlich erfassten Daten zusammenbringen und unverändert an ihre SLOs weiterleiten. Durch diese kontinuierliche Analyse können Unternehmen anschließend alles kontextualisieren und wissen, welches Ziel sie für eine erfolgreiche Anwendungsleistung erreichen müssen. Und wenn dieses Ziel nicht erfüllt wird, können sie feststellen, welche Probleme in den zugrunde liegenden Schichten des Anwendungsstacks zu beheben sind. Letztlich ist die Implementierung von Systemen, die diese kontinuierliche Analyse leisten können, die einzige Möglichkeit, um mit den dynamischen Ressourcenanforderungen moderner Anwendungen und Infrastrukturen Schritt zu halten.



Automatisieren Sie die dynamische Ressourcenbeschaffung.

Da der dynamische Charakter moderner Anwendungen und die Infrastruktur, auf der sie ausgeführt werden, eine kontinuierliche Analyse erfordern, sollten Führungskräfte und Unternehmen auch die Implementierung einer kontinuierlichen Automatisierung in Betracht ziehen. So können Sie SLOs vollständig nutzen und ein präventives System für das Management ihrer Anwendung entwickeln. Laut eines aktuellen Analystenberichts von IDC müssen Unternehmen „die Rolle von Automatisierung bei der Produktfunktionalität berücksichtigen“, um in Zukunft konkurrenzfähig zu bleiben.⁷

⁶ Quelle: Monitoring and Observability for Modern Services and Infrastructure, Juni 2020, Gartner G00720854

⁷ Quelle: Worldwide Application Performance Management Software Market Shares, 2020, Juni 2021, IDC #US47989021

Die Entwicklung einer modernen Infrastruktur versprach eine künftige Resilienz und Elastizität für Anwendungen, doch viele Unternehmen haben heute damit zu kämpfen, die Leistung ihrer Anwendungen aufrechtzuerhalten. Die Implementierung von Automatisierung für das Management moderner Anwendungen ist deswegen essenziell, weil sich damit die Vorteile der Resilienz und Elastizität ernten lassen. Ohne Automatisierung sind die Lösungen, die durch die Konfiguration der richtigen SLOs und der kontinuierlichen Erfassung von Daten möglich werden, nicht effektiv. IT- und Plattform-Teams können ohne Automatisierung der dynamischen Ressourcenbeschaffung keine Anwendungsleistung zusichern: Zu dem Zeitpunkt, an dem eine Warnmeldung ausgelöst und manuell eine Ressourcenentscheidung ausgeführt wird, ist es bereits zur LeistungseinbuÙe gekommen.

Dieser Ansatz ist deshalb mit Schwierigkeiten verbunden, da er erst greift, wenn sich ein Problem bereits ereignet hat.

Automatisierung kann mehr bedeuten als nur eine Reaktion auf ein Ereignis, das durch einen Schwellenwert ausgelöst wird. Durch effektive kontinuierliche Datenerfassung und Identifizierung entscheidender SLIs und SLOs für die jeweilige Fachabteilung können Unternehmen Software umsetzbare Entscheidungen generieren lassen, die automatisiert werden sollten. Unternehmen, die die notwendigen Maßnahmen ergreifen und automatisieren, können ein wirklich elastisches Umfeld schaffen, das proaktiv steuert, wie Anwendungen mit Ressourcen ausgestattet werden, und so eine kontinuierliche Leistung sicherstellen. Diese Art von Automatisierung erfordert ein intelligentes System, das eine dynamische, sich verändernde Umgebung analysiert und die notwendige Entscheidungskette automatisiert, sodass Probleme gelöst werden, *bevor* es zu LeistungseinbuÙen kommt. Diese Art von Leistung und Elastizität lässt sich unmöglich mit einem Prozess aus Spreadsheets und Warnmeldungen erreichen.

Um von einer reaktiven Strategie abzurücken, ist die vollständige Unterstützung der Stakeholder nötig, sowohl aus den Teams, die die Anwendungen verwalten, als auch aus denen, die die Infrastruktur verwalten. Anwendungs- und Produktverantwortliche zögern häufig, Kontrolle über ihre Anwendungen an die Automatisierung abzugeben. Diese Vorbehalte beruhen auf mangelndem Vertrauen in die Automatisierung. Doch dies lässt sich überwinden. Automatisierung erfordert eine kulturelle Transformation Ihrer IT-Organisation. Für echte Elastizität und Resilienz müssen Unternehmen den zu automatisierenden Aktionen vertrauen können. Mit aussagekräftigen geschäfts- und anwendungsorientierten SLOs, die direkt mit der dynamischen Ressourcenbeschaffung in der Plattform und Infrastruktur verknüpft sind, fällt es Anwendungs- und Produktverantwortlichen leichter, der Automatisierung zu vertrauen und sich mit der Erwartung wohl zu fühlen, dass sie vollständig in Geschäftsabläufe implementiert wird.

Was Sie beachten sollten

Anwendungen werden immer komplexer, da Unternehmen zunehmend agile Infrastrukturen wie Hybrid-IT, Containerplattformen und Multicloud einführen. Angesichts dieser künftigen Entwicklungen werden sich Unternehmen, die die erhöhten Erwartungen an das Kundenerlebnis und die Anwendungsleistung erfüllen können, erfolgreich entwickeln. Zu den Best Practices gehören die Festlegung von SLOs in Bezug auf das Kundenerlebnis, die kontinuierliche Analyse der sich verändernden Anwendungsanforderungen und die Automatisierung der dynamischen Ressourcenbeschaffung. Für ein erfolgreiches digitales Geschäft müssen Unternehmen die notwendigen Maßnahmen zur Implementierung einer Automatisierung ergreifen, die Anwendungen dynamisch mit Ressourcen versorgt und dadurch eine sich verändernde Nachfrage und Unternehmens-SLOs erfüllen kann.

Automatisierte Anwendungs-SLOs mit Turbonomic

Turbonomic verwandelt Daten in Aktionen und liefert eine Automatisierung, die Risiken für die Anwendungsleistung vorbeugt und die Elastizität maximiert.

Die Modernisierung geschäftskritischer Anwendungen und Infrastruktur ist eine Investition mit zahlreichen Vorteilen. Doch um die Vorteile von Elastizität, Resilienz und schnellen Markteinführungszeiten nutzen zu können, benötigen Sie Software, die kontinuierlich Ihre Umgebung analysiert und die richtigen Ressourcenentscheidungen zur richtigen Zeit ausführt und dadurch die Anwendungsleistung sicherstellt. Mit Turbonomic können Sie Antwortzeiten, Transaktionsdurchsatz oder andere SLIs/SLOs für Anwendungen mit der dynamischen Ressourcenschaffung korrelieren. Bei schwankender Nachfrage sorgt die dynamische Ressourcenbeschaffung von Turbonomic für eine kontinuierliche Anwendungsleistung.


HPA reicht nicht mehr aus. Turbonomic stellt Ihre SLOs durch eine Top-Down-, Full-Stack-Analyse dynamisch sicher. Sie legen Ihre SLOs fest und unsere KI-gestützte Software sorgt dafür, dass die Plattform und die zugrunde liegende Infrastruktur die für die Erfüllung dieser SLOs benötigten Ressourcen liefert, wo auch immer Ihre Anwendungen laufen.

Nahtlose Integration in Geschäftsabläufe. Durch die Integration von Turbonomic mit Webhooks können Sie Turbonomic Aktionen leicht in Anwendungs-Lebenszyklen, DevOps- und Infrastruktur-Pipelines, Autorisierungs- und Audit-Workflows sowie Kommunikationsprozesse einspeisen.

Minimieren Sie manuelle Tätigkeiten: Entwickler, DevOps und SREs müssen keine Schwellenwerte, Beschränkungen oder Richtlinien für die automatische Skalierung festlegen. Die Software trifft die richtigen Ressourcenentscheidungen für Sie und liefert Aktionen, die sich tatsächlich automatisieren lassen.

Vermeiden Sie zu hohe Ausgaben für Kapazitäten: Sie sind für Ressourcenentscheidungen nicht mehr auf die Entwicklung angewiesen (die häufig Überkapazitäten bereitstellt, um ganz sicher zu gehen, richtig?). Unsere Software bestimmt genau, welche Ressourcen die einzelnen Services benötigen – alles auf Basis des Anwendungsbedarfs.

Planen Sie Wachstum schnell und einfach: Simulieren Sie das Onboarding neuer Services mit unserer Software. Bestimmen Sie exakt, was Sie benötigen, um neues Wachstum zu unterstützen.



Testen Sie
Turbonomic
noch heute

turbonomic.com/try-SLO

© Copyright IBM Corporation 2022

IBM Deutschland GmbH

IBM-Allee 1
71139 Ehningen
[ibm.com/de](https://www.ibm.com/de)

IBM Österreich

Obere Donaustraße 95
1020 Wien
[ibm.com/at](https://www.ibm.com/at)

IBM Schweiz

Vulkanstrasse 106
8010 Zürich
[ibm.com/ch](https://www.ibm.com/ch)

Hergestellt in den Vereinigten Staaten von Amerika
Januar 2022

IBM und das IBM-Logo sind Marken oder eingetragene Marken der International Business Machines Corporation, eingetragen in den USA und/oder anderen Ländern. Weitere Produkt- und Servicennamen können Marken von IBM oder anderen Unternehmen sein. Eine aktuelle Liste der Marken von IBM finden Sie auf [ibm.com/trademark](https://www.ibm.com/trademark).

Das vorliegende Dokument ist ab dem Datum der Erstveröffentlichung aktuell und kann jederzeit von IBM geändert werden. Nicht alle Angebote sind in allen Ländern verfügbar, in denen IBM tätig ist.

Die genannten Performance-Daten und Kundenbeispiele dienen ausschließlich zur Veranschaulichung. Tatsächliche Leistungsergebnisse hängen von den jeweiligen Konfigurationen und Betriebsbedingungen ab.

DIE INFORMATIONEN IN DIESEM DOKUMENT WERDEN OHNE JEGLICHE AUSDRÜCKLICHE ODER STILLSCHWEIGENDE GARANTIE ZUR VERFÜGUNG GESTELLT, EINSCHLIESSLICH DER GARANTIE DER MARKTGÄNGIGKEIT, DER EIGNUNG FÜR EINEN BESTIMMTEN ZWECK UND DER GARANTIE ODER BEDINGUNG DER NICHTVERLETZUNG VON RECHTEN.

Die Garantie für Produkte von IBM richtet sich nach den Bestimmungen und Bedingungen der Vereinbarungen, unter denen sie bereitgestellt werden.

