

RZ-Wiki

Resilienz von Rechenzentrums-
Infrastrukturen

Anwendungsbeispiele zur EN 50600

Inhalt

1	Abstrakt	3
2	Glossar	4
	2.1 Abkürzungen	4
	2.2 Definitionen	4
3	Resilienz von Rechenzentrums-Infrastrukturen	5
	3.1 Einführung	5
	3.1.1 Komplexität	5
	3.1.2 Begriffsbestimmung	6
	3.1.3 Aspekte der Resilienz	6
	3.1.4 Zuverlässigkeit	7
	3.1.5 Verfügbarkeit	7
	3.1.6 Fehlertoleranz	8
	3.1.8 Anwendungspraxis	9
4	Referenzen	10
	4.1 Quellenangaben	10
	4.2 Abbildungen	10
	4.3 Formeln	10
	4.4 Tabellen	10

1 Abstrakt

Resilienz im Kontext von Rechenzentren beschreibt verschiedene Eigenschaften der Infrastrukturen. Daraus folgt, dass Resilienz nicht durch eine einzige Kennzahl beziffert werden kann. Verschiedene Kennzahlen, genannt Key Performance Indikatoren (KPI), sind nötig, um jeweils verschiedene Aspekte der Resilienz zu beschreiben. Zu den Wesentlichen gehören die Verfügbarkeit, die Zuverlässigkeit und die Fehlertoleranz.

Die EN 50600 [1] respektive ISO/IEC 22237 [2] definiert Verfügbarkeitsklassen VK1 „niedrig“ bis VK4 „sehr hoch“ anhand von Eigenschaften. Über die Einteilung in Verfügbarkeitsklassen hinaus, liefert die ISO/IEC TS 22237-31 „Key performance indicators for resilience“ [3] ausgewählte Leistungskennzahlen zur quantitativen Bewertung von Rechenzentrums-Infrastrukturen. Mittels solchen KPIs ist es möglich, bestimmte Aspekte der Resilienz zu bestimmen. Damit ist die Grundlage zur quantitativen Bewertung des Rechenzentrums und fortführend die Möglichkeit zur Infrastruktur-Optimierung gegeben.

Parallel zur Analyse bzw. Optimierung der Effizienz der Rechenzentrums-Infrastruktur gilt die Resilienz-Analyse als Werkzeug um sicherzustellen, dass keine unentdeckten Kompromisse hinsichtlich der Betriebssicherheit eingegangen werden müssen. Die quantitative Bewertung der relevanten Aspekte der Resilienz gilt als unverzichtbarer Teil der ganzheitlichen Betrachtung eines Rechenzentrums, über dessen gesamten Lebenszyklus.

2 Glossar

2.1 Abkürzungen

Abkürzung	Bezeichnung	Beschreibung
DCI	Data Centre Infrastructure	Rechenzentrums-Infrastruktur
DPoF	Double Points of Failure	Zwei-Fehler-Kombination, die zum Gesamtausfall der Infrastruktur führen kann
EN	Europäische Norm	Regeln, welche durch ein europäisches Komitee für Standardisierung ratifiziert wurden
IEC	International Electrotechnical Commission	Internationale Normungsorganisation für Normen im Bereich der Elektrotechnik und Elektronik
IEEE	Institute of Electrical and Electronics Engineers	Weltweiter Verband von Ingenieuren, Technikern, Wissenschaftlern und angrenzender Berufe hauptsächlich aus den Bereichen Elektrotechnik und Informationstechnik
ISO	International Organization for Standardization	Internationale Vereinigung von Normungsorganisationen, erarbeitet Normen außerhalb der Zuständigkeit der IEC
KPI	Key Performance Indicator	Leistungskennzahl zum Messen bestimmter Eigenschaften
SLA	Service Level Agreement	Vereinbarung zur Güte einer Dienstleistung zwischen Auftraggeber und Dienstleister
SPoF	Single Points of Failure	Ein-Fehler-Stelle, die zum Gesamtausfall der Infrastruktur führen kann
VK	Verfügbarkeitsklasse	Ergebnis der Risikobewertung, siehe EN 50600

Tabelle 1: Abkürzungen

2.2 Definitionen

- A** Availability, Verfügbarkeit
- A_i** Inherent Availability, Inhärente Verfügbarkeit
- A_o** Operational Availability, Operationale Verfügbarkeit
- A_p** Past Availability, Zurückliegende Verfügbarkeit
- R(t)** Reliability, Zuverlässigkeit
- K** Anzahl gleichzeitig auftretender Fehler
- N_(K=1)** Anzahl der SPoF, auch N_{SPoF}
- N_(K=2)** Anzahl der DPoF, auch N_{DPoF}

3 Resilienz von Rechenzentrums-Infrastrukturen

Autor: Uwe Müller

3.1 Einführung

3.1.1 Komplexität

Rechenzentren und im Besonderen deren Infrastrukturen bestehen aus zahlreichen verknüpften technischen Teilsystemen verschiedener Gewerke. Jedes Teilsystem ist aus einer Menge von Komponenten konstruiert. Bereits die Anzahl der involvierten technischen Entitäten, sowie der Einfluss der Dimensionierung und deren Wartung, seien Beleg für die Komplexität der Aufgabe, Rechenzentren zu errichten und zu betreiben.

Durch den Kanon der Standards EN 50600 [1] bzw. ISO/IEC 22237 [2] ist es gelungen, die Phasen des Lebenszyklus von Rechenzentren international zu normieren. Zentraler Bestandteil ist die Ermittlung der Verfügbarkeitsklassen der verschiedenen Rechenzentrums-Infrastrukturen auf Grundlage von Risikobetrachtungen.

Rechenzentrums-Designer, -Planer oder -Betreiber, befinden sich in einem Spannungsfeld, welches einerseits die Minimierung der Investitions- und Betriebskosten fordert – auch in Notwendigkeit des Klima- und Umweltschutzes – sowie andererseits höchste Betriebssicherheit fordert. Im Allgemeinen gilt es vereinbarten Service-Level-Agreements (SLA) zu entsprechen. Fällt ein Rechenzentrum für eine bestimmte Zeitspanne aus, kann dies weitreichende Konsequenzen haben, nicht nur monetäre, sondern auch Leben gefährden und/oder überregionale Einschränkungen oder Schäden verursachen.

Dieses komplexe Feld, mit teils gegenteiligen Anforderungen zu bearbeiten, erfordert ganzheitliche, holistische Herangehensweisen. Ein solcher Bestandteil ist die Analyse der Aspekte der Resilienz oder kurz die Resilienz-Analyse.

3.1.2 Begriffsbestimmung

Die Verwendung des Resilienz-Begriffes folgt der Definition der IEEE Task Force PES-TR65 [4]. Mit dem Bezug auf technische Infrastrukturen beschreibt Resilienz Fähigkeiten zum:

- Widerstehen und Reduzieren des Ausmaßes oder der Dauer unterbrechender Ereignisse,
- Vorhersehen, Neutralisieren, Anpassen beim Eintritt unterbrechender Ereignisse,
- zügigen Wiederherstellen.

Das Widerstehen unterbrechender Ereignisse, deren Neutralisierung und Fähigkeit zur Anpassung ist durch die Ermittlung der Fehlertoleranz der Rechenzentrums-Infrastruktur quantifizierbar.

Die Dauer unterbrechender Ereignisse sowie deren Fähigkeit zum zügigen Wiederherstellen ist durch die Ermittlung verschiedener KPIs der Verfügbarkeit quantifizierbar.

Die Fähigkeit zum Wiederherstellen betrifft die Reparaturfähigkeit mit Hinblick auf materielle Ressourcen und adäquaten Personaleinsatz.

3.1.3 Aspekte der Resilienz

Die folgende Grafik illustriert verschiedene Aspekte der Resilienz von Rechenzentrums-Infrastrukturen:

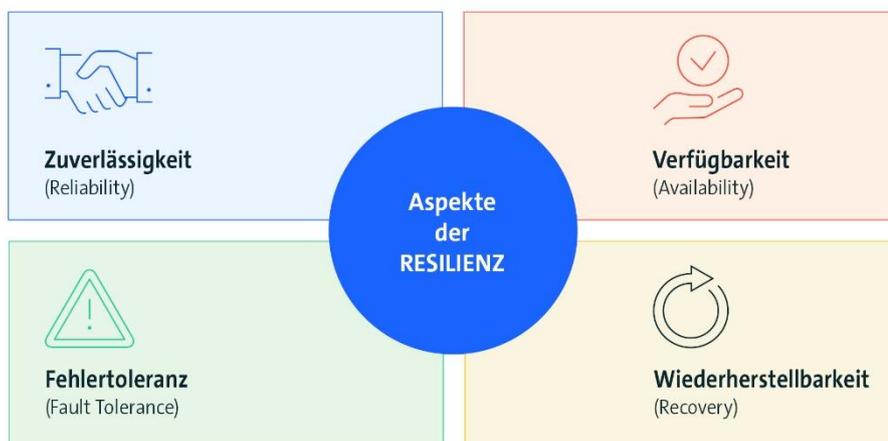


Abbildung 1: Aspekte der Resilienz

Jeder Aspekt der Resilienz hat seine spezifische Aussagekraft, die nicht zwangsläufig von einem anderen Aspekt abhängig sein muss. Auf Beispiele wird in Abschnitt 3.1.8 Anwendungspraxis eingegangen.

3.1.4 Zuverlässigkeit

Die Zuverlässigkeit $R(t)$, auch als Überlebenswahrscheinlichkeit bezeichnet, ist ein KPI für die Wahrscheinlichkeit der fehlerfreien Funktion. Sie kann für einzelne Komponenten, Teilsysteme als auch für das Gesamtsystem bestimmt werden. Gemäß IEEE Std. 493-2007 [5] ist die Zuverlässigkeit abhängig von der Fehlerrate sowie dem Alter. Sie unterliegt einer Verteilungsfunktion, im Allgemeinen der Exponentialfunktion.

Hohe Zuverlässigkeit ist ein Maß für geringe Fehlerwahrscheinlichkeit bzw. hohe Überlebenswahrscheinlichkeit.

3.1.5 Verfügbarkeit

Die Verfügbarkeit A ist Ausdruck für die Wahrscheinlichkeit, dass die Rechenzentrums-Infrastruktur funktionstüchtig ist. Gemäß IEEE Std. 493-2007 [5] haben sich verschiedene Verfahren zur Berechnung der Verfügbarkeit etabliert.

Inhärente Verfügbarkeit

Die Inhärente Verfügbarkeit A_i ist ein berechnetes Wahrscheinlichkeitsmaß der Verfügbarkeit auf Grundlage der eingesetzten Komponenten und Teilsysteme, unter Annahme „idealer“ Wartung und Instandsetzung.

Als Designziel für Rechenzentrums-Infrastrukturen der VK3 und VK4, unter Berücksichtigung der Stromversorgung und Klimatisierung, gilt oftmals das Erreichen der Inhärenten Verfügbarkeit von mindestens „Fünf-Neunen“. Damit ist gemeint $A_i \geq 0,99999$ bzw. 99,999 %.

Operationale Verfügbarkeit

Die Operationale Verfügbarkeit A_o ist ein berechnetes Wahrscheinlichkeitsmaß der Verfügbarkeit auf Grundlage der eingesetzten Komponenten und Teilsysteme, bei Inbezugnahme der Logistikzeiten sowie Ausfallzeiten für ungeplante und geplante Wartungen.

Im Vergleich zur Inhärenten Verfügbarkeit, fällt die Operationale Verfügbarkeit im Allgemeinen mindestens eine Größenordnung geringer aus. So kann eine Rechenzentrums-Infrastruktur der VK3, welche das Kriterium der inhärenten „Fünf-Neunen“ erfüllt, unter Berücksichtigung der Stromversorgung und Klimatisierung, durchaus eine Operationale Verfügbarkeit im Bereich $A_o \approx 0,9997$ bzw. 99,97 % aufweisen. Ein solcher Wert kann als Ausgangspunkt für nachfolgende Optimierungsschritte an der Infrastruktur dienen.

Zurückliegende Verfügbarkeit

Die Zurückliegende Verfügbarkeit A_p wird wie folgt ermittelt:

$$A_p = \frac{\text{Betrachtungszeitspanne} - \text{Gesamtausfallzeit}}{\text{Betrachtungszeitspanne}}$$

Formel 1: Zurückliegende Verfügbarkeit

Um die jährliche Verfügbarkeit zu ermitteln, hat sich die Betrachtungszeitspanne von 8760 h etabliert. Zur besseren Unterscheidbarkeit von A_i und A_o wird die Zurückliegende Verfügbarkeit oftmals in Prozent angegeben. Damit ergibt sich:

$$A_p = \frac{8760 \text{ h} - \text{Gesamtausfallzeit}}{8760 \text{ h}} * 100 \%$$

Formel 2: Jährliche Verfügbarkeit prozentual

3.1.6 Fehlertoleranz

Fehlertoleranz beschreibt die Eigenschaft der Rechenzentrums-Infrastruktur, trotz des Ausfalls von Teilsystemen funktionstüchtig zu bleiben, das heißt diesbezügliche SLAs nicht zu verletzen. Für ausgewählte Bereiche in hochverfügbaren Infrastrukturen fordert die EN 50600, dass nicht nur Ein-Fehler-Ereignisse, sondern auch Zwei-Fehler-Ereignisse ohne Gesamtausfall überstanden werden müssen.

Ein-Fehler-Toleranz

Ein-Fehler-Toleranz N_{SPoF} beziffert die Teilsysteme der Infrastruktur, welche im Fehlerfall zum Gesamtausfall der zugesicherten Dienstleistung führen. Dies kann zur Folge haben, dass das vereinbarte SLA verletzt wird.

Zwei-Fehler-Toleranz

Zwei-Fehler-Toleranz N_{DPoF} beziffert alle Kombinationen von zwei Teilsystemen der Infrastruktur, welche im Fehlerfall zum Gesamtausfall der zugesicherten Dienstleistung führen. Dies kann zur Folge haben, dass das vereinbarte SLA verletzt wird.

3.1.7 Wiederherstellbarkeit

Wiederherstellbarkeit drückt die Eigenschaft aus, dass Rechenzentrums-Infrastrukturen im Fehlerfall prinzipiell reparierbar sind. Voraussetzungen für Reparaturen sind:

- Hinreichend bereitstehende materielle Ressourcen bzw. Ersatzteile sowie

- Reparatur-Fachkräfte, welche im adäquaten Zeitrahmen wirksam sein können.

Zur Ermittlung der notwendigen Vorhaltungen und maximalen Reaktionszeiten können Verfügbarkeits-Untersuchungen der Teilsysteme der Infrastruktur durchgeführt werden.

Hinsichtlich betriebsorganisatorischer Aspekte sei auf die entsprechenden Normenteile der EN 50600 verwiesen.

3.1.8 Anwendungspraxis

In der praktischen Anwendung der KPIs zur Bewertung der Resilienz stellt sich heraus, dass sich deren Aufwand zur Berechnung recht unterschiedlich darstellt. Die Ermittlung des KPIs der Zurückliegenden Verfügbarkeit A_p sowie der Ein-Fehler-Toleranz der Rechenzentrums-Infrastruktur erfordert relativ geringe Anstrengungen.

Zur Berechnung der Inhärente Verfügbarkeit A_i , Operationalen Verfügbarkeit A_o sowie der Zwei-Fehler-Toleranz N_{DPoF} ist es im Allgemeinen nötig, die Rechenzentrums-Infrastruktur zu modellieren. Nach der Dateneinpfege der Komponenten in die Teilsysteme, erfolgt die Anwendung analytischer Verfahren [6]. Dabei kann sich die Unterstützung durch geeignete Analyse-Software als hilfreich erweisen. Insbesondere für Infrastrukturen der Verfügbarkeitsklassen VK3 und VK4 kann der Berechnungsaufwand auf Grund der Infrastruktur-Komplexität entsprechende Expertise erfordern.

Der Mehrwert von Resilienz-Analysen sei an drei Extrembeispielen aus der Praxis gezeigt:

1. Rechenzentrums-Infrastrukturen trotz dessen, dass sie nicht frei von Ein-Fehler-Stellen sind, die bisher nicht ausgefallen sind. Die Zurückliegende Verfügbarkeit wäre damit 100 %. Ein solches Ergebnis ist nicht gleichbedeutend mit ebensolcher Inhärenter oder Operationaler Verfügbarkeit, denn im Wartungsfall kann die Notwendigkeit bestehen, die Ein-Fehler-Stelle und damit das gesamte Rechenzentrum, abzuschalten.
2. Es sind Fälle bekannt, dass Rechenzentrums-Infrastrukturen ausgefallen sind, obwohl sie frei von Ein-Fehler-Stellen sind. Ursache können Verkettungen von Ereignissen, auch begünstigt durch Alterung sein.
3. Zusätzliche Redundanzen erhöhen die Komplexität der Infrastruktur und vermindern im Allgemeinen deren Effizienz. Ebenso können zusätzliche Redundanzen, trotz Verbesserung der Verfügbarkeit, die Zuverlässigkeit vermindern [6]. Die Frage, wieviel Redundanz nötig und sinnvoll ist, im Zusammenspiel mit Effizienzbetrachtungen, kann mittels Resilienz-Analysen beantwortet werden.

4 Referenzen

4.1 Quellenangaben

- [1] DIN EN 50600, Informationstechnik - Einrichtungen und Infrastrukturen von Rechenzentren, Normenfamilie.
- [2] ISO/IEC 22237, Information technology - Data centre facilities and infrastructures, Normenfamilie.
- [3] ISO/IEC TS 22237-31, Information technology - Data centre facilities and infrastructures, Part 31: Key performance indicators for resilience, 2023.
- [4] IEEE Task Force PES-TR65, The Definition and Quantification of Resilience, 2018.
- [5] IEEE Std. 493, Design of Reliable Industrial and Commercial Power Systems, 2007.
- [6] IET Journals, eISSN 2051-3305, joe.2018.5520, Resilience of data centre power system: modelling of sustained operation under outage, definition of metrics, and application, 2019.

4.2 Abbildungen

Abbildung 1: Aspekte der Resilienz 6

4.3 Formeln

Formel 1: Zurückliegende Verfügbarkeit 8

Formel 2: Jährliche Verfügbarkeit prozentual 8

4.4 Tabellen

Tabelle 1: Abkürzungen 4

Bitkom vertritt mehr als 2.200 Mitgliedsunternehmen aus der digitalen Wirtschaft. Sie generieren in Deutschland gut 200 Milliarden Euro Umsatz mit digitalen Technologien und Lösungen und beschäftigen mehr als 2 Millionen Menschen. 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, kreieren Content, bieten Plattformen an oder sind in anderer Weise Teil der digitalen Wirtschaft. 82 Prozent der im Bitkom engagierten Unternehmen haben ihren Hauptsitz in Deutschland, weitere 8 Prozent kommen aus dem restlichen Europa und 7 Prozent aus den USA. 3 Prozent stammen aus anderen Regionen der Welt. 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 leistungsfähigen und souveränen Digitalstandort zu machen.

Herausgeber

Bitkom e. V.
Albrechtstraße 10 | 10117 Berlin
Tel.: 030 27576-0 | bitkom@bitkom.org
www.bitkom.org

Autor und Ansprechpartner Inhalt

Uwe Müller | InfraOpt GmbH
uwe.mueller@infraopt.eu | T +49 172 8368 939

Ansprechpartner Bitkom

Kilian Wagner | Bereichsleiter für nachhaltige digitale Infrastrukturen
T +49 151 14824861 | k.wagner@bitkom.org

Verantwortliches Bitkom-Gremium

AK Rechenzentren

Copyright

Bitkom 2025

Diese Publikation stellt eine allgemeine unverbindliche Information dar. Die Inhalte spiegeln die Auffassung im Bitkom zum Zeitpunkt der Veröffentlichung wider. Obwohl die Informationen mit größtmöglicher Sorgfalt erstellt wurden, besteht kein Anspruch auf sachliche Richtigkeit, Vollständigkeit und/oder Aktualität, insbesondere kann diese Publikation nicht den besonderen Umständen des Einzelfalles Rechnung tragen. Eine Verwendung liegt daher in der eigenen Verantwortung des Lesers. Jegliche Haftung wird ausgeschlossen. Alle Rechte, auch der auszugsweisen Vervielfältigung, liegen beim Bitkom oder den jeweiligen Rechteinhabern.