COMPUTERWOCHE
VOICE OF DIGITAL

Link: https://www.computerwoche.de/a/wenn-der-server-abschmiert,1912848

Hochverfügbarkeitslösungen

Wenn der Server abschmiert

Datum: 09.12.2009
Autor(en): Klaus Manhart

Was können Sie sich im Hinblick auf Server-Ausfallzeiten leisten? Tage, Stunden, Minuten? Oder gar keine? Lesen Sie in diesem Kurzbeitrag mehr über Verfügbarkeitsklassen und die wichtigsten Hochverfügbarkeits-Lösungen.

Zumindest bei kritischen Unternehmensprozessen sind heute 24 Stunden Verfügbarkeit an jedem Tag der Woche unabdingbar. Im Idealfall dürfen Server nie stillstehen und Daten nie unerreichbar sein. In der Praxis ist perfekte Verfügbarkeit zwar so gut wie nicht gewährleistet, doch für viele Betriebe ist es wichtig, diesem Ideal möglichst nahe zu kommen.

Verfügbarkeiten werden als Verhältnis von Uptime (die Zeit, die der Server verfügbar ist) zur Gesamtzeit, also Uptime plus Downtime gemessen: Uptime / (Uptime + Downtime). Standardkomponenten erreichen heute eine Verfügbarkeit von 99,9 Prozent. Das klingt beeindruckend, reicht in der Praxis aber oft nicht. Denn immerhin lassen 99,9 Prozent einen Zugriffsverlust von 8,7 Stunden pro Jahr zu - sind es acht Stunden in der Geschäftszeit, kann auch dies zu lang sein.

Verfügbarkeitsklassen und maximale Ausfallzeiten

Verfügbarkeit	Verfügbarkeitsklasse	Max. Ausfall pro Monat	Max. Ausfall pro Jahr
99 Prozent	2	7,3 Stunden	87,66 Stunden
99,9 Prozent	3	43,8 Minuten	8,76 Stunden
99,99 Prozent	5	4,38 Minuten	52,6 Minuten
99,999 Prozen	5	26,3 Sekunden	5,26 Minuten

Eine der nächst höheren Stufen der Verfügbarkeit ist eine Ausfallsicherheit von 99,99 Prozent. Dies entspricht einer Downtime von etwa 53 Minuten pro Jahr. Selbst dies ist für manche Einsatzgebiete zu viel. Als magischer Wert gelten "Five-Nine": 99,999 Prozent Verfügbarkeit - das entspricht einer Ausfallzeit von maximal etwa fünf Minuten pro Jahr.

Hochverfügbarkeits-Cluster

Noch vor wenigen Jahren mussten Unternehmen bei **Hochverfügbarkeitslösungen**[1] auf proprietäre Systeme zurückgreifen. Die waren teuer und aufwändig und nur für größere Unternehmen geeignet. Mittlerweile gibt es jedoch hochverfügbare Systeme, die ganz auf Standard-Technologien aufbauen.

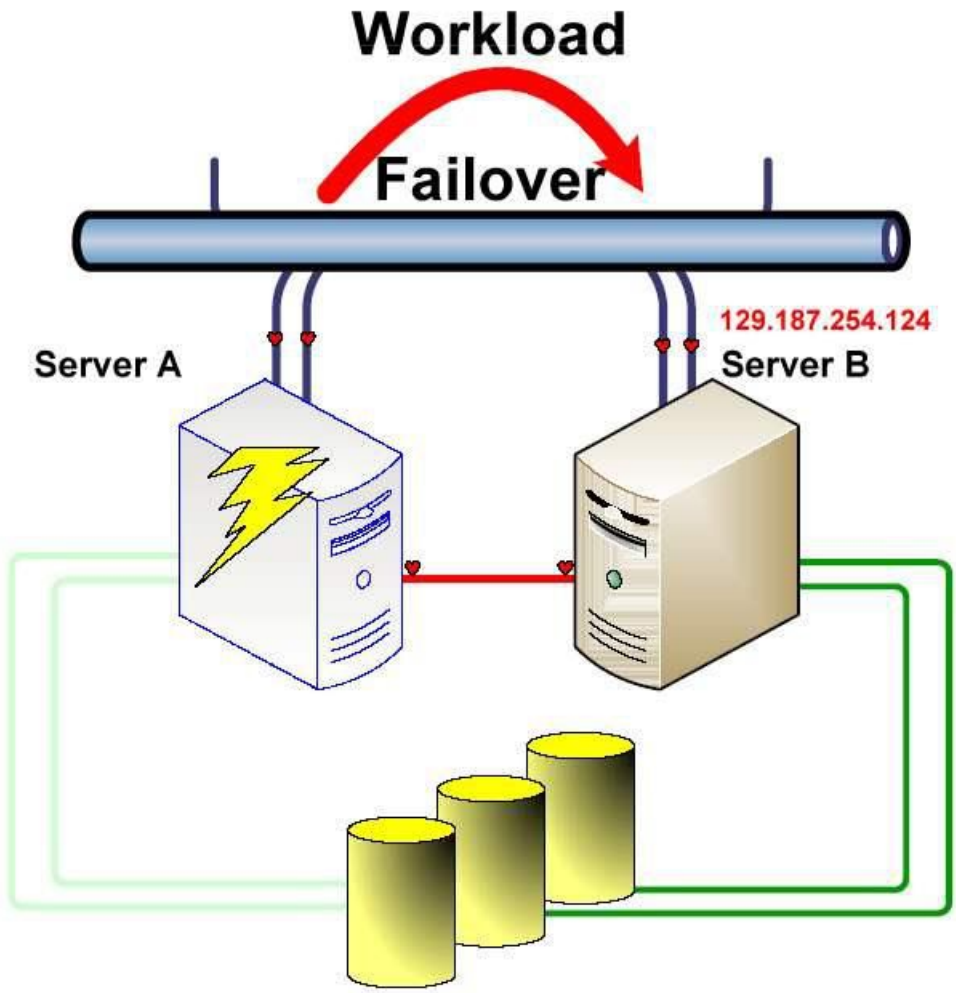
Hohe Ausfallsicherheit bieten vor allem redundante Server-Systeme wie **Hochverfügbarkeits-Cluster**[2]. Diese bestehen aus mindestens zwei identischen Servern, etwa zwei Datenbank-Servern, die miteinander verschaltet sind. Tritt auf dem produktiven Datenbank-Server ein Fehler auf, werden die auf diesem Rechner laufenden Dienste auf das andere System migriert.

Das Zweitsystem übernimmt im Fehlerfall die Aufgaben des Primärsystems. Alle Funktionen und Anwendungen laufen simultan ab und jedes System kann im Notfall autark weiterarbeiten. Ist das defekte System wieder betriebsbereit, synchronisieren sich die Server automatisch und arbeiten anschließend wieder parallel weiter.

Cluster lassen sich kompakt mit Blade-Servern realisieren.

Cluster werden vor allem über Rack- und **Blade-Server**[3] realisiert. Blades haben besonders in größeren Rechenzentren ihren Siegeszug angetreten - nicht zuletzt, weil Blade-Server höhere Packungsdichten erreichen als die üblichen horizontalen Einschübe für 19-Zoll-Racks. Ein einziges Rack kann hundert oder noch mehr Server-Blades aufnehmen. Auch beim Preis-Leistungs-Vergleich punkten Blades vor Rack-Servern, denn Blades nutzen Betriebsmittel wie Stromversorgung, Kühlung und Netz-Switches gemeinsam. Und schließlich ist die Verkabelung bei Blades auch weniger aufwändig als bei normalen Servern.

Alternative: Tandemrechner



Verstummt in einem HA-Cluster der Heartbeat werden die Dienste des Primärsystems (links) vom Standby-System (rechts) übernommen (Quelle: LRZ München).

Cluster arbeiten im Störungsfall allerdings **nicht unterbrechungsfrei**[4]. Die Übernahme der Prozesse durch das nicht gestörte System erfordert eine gewisse Failover-Zeit. Damit kommen Cluster-Server nicht über eine durchschnittliche Verfügbarkeit von rund 99,99 Prozent hinaus.

In manchen Fällen reichen Cluster-Server nicht mehr aus. Dann etwa, wenn Real-Time-Business betrieben wird, bei dem Daten ohne Zwischenspeicherung im flüchtigen RAM sofort verarbeitet werden. Geht beispielsweise bei Bestelleingang der Server kaputt, sind die Daten, die gerade in der Maschine waren, verloren - trotz des zweiten Rechners.

In solchen Fällen ist man bestrebt, Ausfälle ganz zu vermeiden. Eine Lösungsmöglichkeit sind Tandemrechner - zwei Server, auf denen zu jedem Zeitpunkt alle Prozesse parallel laufen. Fällt dann ein System des Duos aus, übernimmt einfach der Partnerrechner ohne Unterbrechung. Damit stehen die Applikationen die ganze Zeit über bereit, wobei Verfügbarkeiten von über 99,999 Prozent erreicht werden.

Links im Artikel:

[1] **https://www.computerwoche.de/hardware/data-center-server/2020347/index.html**
[2] **https://www.computerwoche.de/subnet/hp-intel/1894042/**
[3] **https://www.computerwoche.de/hardware/data-center-server/1906338/index2.html**
[4] **https://www.computerwoche.de/hardware/data-center-server/2020347/index4.html**