POWER7 System RAS

Schlüsselaspekte für die Zuverlässigkeit, Verfügbarkeit und Wartungsfreundlichkeit von Power Systems

1. November 2010

IBM Server and Technology Group
Daniel Henderson, Jim Mitchell und George Ahrens

Inhalt

Einführung

Das Jahrzehnt von Smarter Planet

Im Jahr 2008 trug das IBM Führungsteam dem überaus großen Einfluss der Technik auf die Gesellschaft Rechnung und eröffnete einen Dialog über die besten Möglichkeiten, im Rahmen von ***Smarter Planet*** einen globalen Beitrag zu leisten. Smarter Planet bedeutet, dass die Systeme und Prozesse, die weltweit wirken, intelligenter gemacht werden sollen. Diese Technologien umfassen sowohl herkömmliche IT-Infrastrukturen als auch die zunehmende Anzahl von Einsatzbereichen für intelligente Einheiten: Smart Phones, GPS-Systeme, Kraftfahrzeuge, Geräte, Straßen, Stromnetze, Wasserversorgungssysteme.

Nach unserer Einschätzung werden bereits 2011 nahezu eine Billion Geräte an das Internet angeschlossen sein. Diese mit dem Internet verbundenen Geräte ermöglichen neue Arten sozialer Interaktionen ebenso wie neue Möglichkeiten für Unternehmen, mit ihren Kunden, Mitarbeitern und Zulieferern in Kontakt zu treten. Zugleich entstehen bei diesen Interaktionen riesige Datenmengen – Rohdaten über die Art und Weise, wie die Menschen ihre Ressourcen nutzen und wie auf den Märkten Daten fließen. Diese können genutzt werden, um die Funktionsweise von Gesellschaften zu verstehen. Server in einer solchen Umgebung müssen aus dem Internet empfangene Transaktionen schnell verarbeiten und große Datenmengen in Back-End-Datenbanken speichern und verarbeiten können. Darüber hinaus müssen sie Analysetools unterstützen: Mit diesen können Führungskräfte von Unternehmen und Communitys die erforderlichen Einblicke erhalten, um innerhalb ihrer Zuständigkeitsbereiche fundierte und rechtzeitige Entscheidungen treffen zu können.

Intelligentere Systeme im Rahmen von Smarter Planet

Im Februar 2010 kündigte IBM die ersten Modelle einer neuen Generation von Power-Servern an, die auf dem Mikroprozessor POWER7™ basieren. Obwohl diese neuen Server so konzipiert sind, dass sie herkömmliche IT-Ansprüche an höhere Leistung und Kapazität erfüllen, erfordert der Umstieg auf „intelligentere“ Lösungen entsprechende Server, die schnell und effizient skaliert werden können, die Auslastung automatisch optimieren und Ressourcen je nach Bedarf der Anwendung flexibel verwalten, Ausfallzeiten minimieren, Energie sparen und Management-Tasks automatisieren.

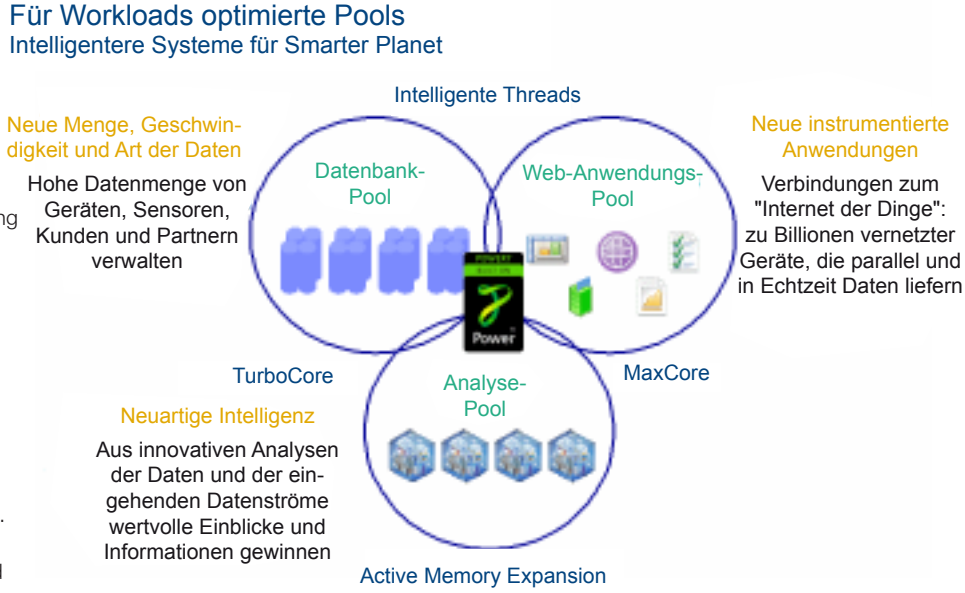
Deshalb erhöhte IBM in großem Umfang die Fähigkeiten von POWER7-Systemen zur Parallelverarbeitung (mit Hardware- und Softwareintegration). Dies ist wichtig, damit Millionen gleichzeitig ablaufender Transaktionen verwaltet werden können. Die neuen Power Systems treten, wie erwartet, das stolze Erbe der IBM Power-Server an – sie bieten branchenführende Geschwindigkeiten für die Transaktionsverarbeitung, damit auch die höchsten Datenbankauslastungen effizient verarbeitet werden können. Darüber hinaus bedeuten diese neuen, für die Ausführung massiver Internet-Workloads optimierten Produkte einen Entwicklungssprung hin zur Datenverarbeitung mit hohen Durchsätzen.

Intelligentere Systeme im Rahmen von Smarter Planet

Die schnelle Implementierung von Technologie stellt neue Anforderungen an IT-Systeme. Server innerhalb dieser Umgebung müssen Transaktionen aus dem Intranet schnell verarbeiten und große Datenmengen in Back-End-Datenbanken speichern und verarbeiten können.

Darüber hinaus müssen sie innovative Analysetools verwenden können, mit denen wichtige Einblicke in diese wertvollen Daten gewonnen werden können.

Server müssen heute schnell und effizient skaliert werden können, die Auslastung automatisch optimieren und Ressourcen je nach Bedarf der Anwendung flexibel umverteilen, Ausfallzeiten minimieren, Energie sparen und Verwaltungsaufgaben automatisieren.



Im Rahmen von neu entstehenden Geschäftsmodellen werden große Datenmengen gesammelt (zum Beispiel aus dem Internet, von Sensoren in Stromnetzen, Straßen oder aus der Lieferkette). Diese Daten werden in großen Datenbanken abgelegt und mithilfe von intelligenten Analysetools untersucht. Dabei werden Informationen gefunden, die einen Wettbewerbsvorsprung bedeuten können. Pools schneller Server mit Multithread-POWER7-Prozessoren können in optimierten Pools implementiert werden. Diese Pools ermöglichen es, mit Internet-Workloads effizient umzugehen, große Datenmengen in Datenbanken zu speichern und zu verarbeiten und spezialisierte Analysetools einzusetzen (wie zum Beispiel IBM Smart Analytics System 7700). Dies dient dem Zweck, geschäftlich nutzbare Informationen abzuleiten. Die drei Modi der Datenverarbeitung – massive Parallelverarbeitung, Datenverarbeitung mit hohen Durchsätzen und Analysefähigkeiten – werden mithilfe der IBM Systems Director-Software konsistent integriert und verwaltet.

Zuverlässigkeit, Verfügbarkeit und Wartungsfreundlichkeit (RAS) von Servern

Seit den frühen 1990er Jahren trieb das Power-Entwicklerteam in Austin die Integration bewährter Technologien für die Zuverlässigkeit von Mainframes in Power-Servern energisch voran. Wohl eines der wichtigsten Leistungsmerkmale, das im Jahr 1997 in allen IBM Power System-Servern eingeführt wurde, ist eine Methodik für die Konzeption von Hardware, die als First Failure Data Capture (FFDC) bezeichnet wird. Bei dieser Methodik werden hardwarebasierte Fehlerdetektoren eingesetzt, um interne Systemkomponenten umfassend zu instrumentieren. Die einzelnen Detektoren sind Diagnosetests, die in der Lage sind, einem dedizierten Serviceprozessor Fehlerdetails zu melden. Wenn FFDC mit einer automatischen Firmwareanalyse gekoppelt ist, wird diese Methodik dazu verwendet, um die einem Fehler zugrunde liegende Ursache bei seinem ersten Auftreten schnell und präzise zu ermitteln, unabhängig von der Phase des Systembetriebs und ohne eine Diagnose mit Fehlerreproduktion ausführen zu müssen. Dabei ist es entscheidend, zu bestimmen, ***welche Komponente*** einen Fehler – bei ***seinem ersten Auftreten*** – verursacht hat, und dessen erneutes Auftreten zu verhindern. Dieses Produktmerkmal wurde in einer Reihe technischer RAS-Artikel und „White Papers“ detailliert beschrieben, die technische Einzelheiten zum IBM Power-Systemaufbau enthalten.

Der Artikel „Fault-tolerant design of the IBM pSeries 690 system using POWER4™ tolerant design of the IBM pSerie processor technology“ [1] beleuchtet, wie POWER-Systeme konzipiert wurden – vom ersten RAS-Konzept bis zur vollständigen Implementierung. In den neun Jahren seit der Einführung von POWER4 sind von IBM einige aufeinanderfolgende Generationen von POWER-Prozessoren eingeführt worden, die jeweils neue RAS-Funktionen enthalten. In nachfolgenden White Papers[2,3] wurde beschrieben, wie RAS-Attribute so definiert und gemessen wurden, dass die Ziele von RAS auch tatsächlich erfüllt werden. Außerdem wurde genau angegeben, wie die einzelnen neuen Funktionen zur Zuverlässigkeit von Systemoperationen beitragen. Allgemein umrissen diese Dokumente die wesentlichen Prinzipien, an denen sich der IBM Systemaufbau bei der Implementierung der RAS-Architektur orientierte. Ein Anwender kann zu Recht erwarten, dass ein Server physische Sicherheit, Systemintegrität und automatisierte Fehlererkennung und -bestimmung bietet.

1. Systeme müssen in einem umfassenden Sinne zuverlässig sein. Sie müssen Folgendes leisten:
   - Korrekte Datenverarbeitungsergebnisse generieren
   - Keinerlei physische Risiken darstellen
   - Frei von Konstruktionsfehlern sein, die die Zuverlässigkeit oder die Leistung beeinträchtigen
   - Solange sie gewartet werden können, wenige Hardwarefehler zulassen
2. Systeme müssen für das Erreichen der erforderlichen Verfügbarkeitsebenen konfiguriert werden können und dabei Folgendes einhalten:
   - Nicht die Leistung, die Auslastung oder die Virtualisierung beeinträchtigen
   - Bewährten Aufbau und bewährte Verfahren für die Verfügbarkeit verwenden
   - Clustering-Verfahren nutzen können, um die höchsten Verfügbarkeitsebenen zu erreichen
3. Systeme müssen Fehler automatisch und proaktiv diagnostizieren und dabei:
   - Wenn möglich, die automatische Fehlerbehebung und die selektive Aufhebung der Konfiguration nutzen, um Anwendungen aktiv zu halten
   - Weder auf Diagnoseprogramme zurückgreifen, um Fehler zu reproduzieren, noch auf eine manuelle Analyse von Speicherauszügen
   - Sich nicht auf die Unterstützung des Betriebssystems oder auf Anwendungen verlassen

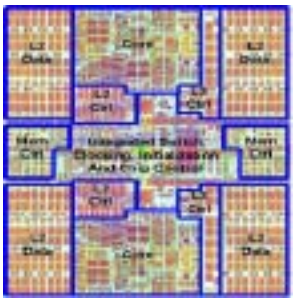
[1] D. C.Bossen A. Kitamorn, K. F. Reick und M. S. Floyd, „Fault-tolerant design of the IBM pSeries 690 system using POWER4 processor technology“, IBM Journal of Research and Development, VOL.46 NO.1, Januar 2002.

[2] J. Mitchell, D. Henderson, G. Ahrens, „IBM ###[[Zeichen für eServer einfügen!]] p5: A Highly Available Design for Business-Critical Applications“, p5tiPOWER5RASwp121504.doc, Dezember 2004.

[3] J. Mitchell, D. Henderson, G. Ahrens und J. Villarreal, „Highly Available IBM Power Systems Servers for Business-Critical Applications “, PSW03003-USEN-00, April 2008.

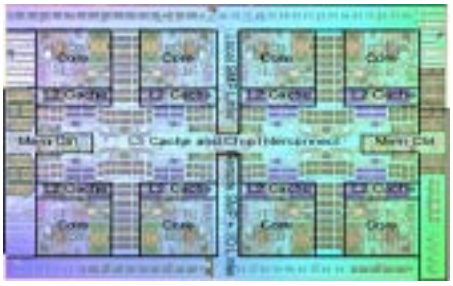
Die Absicht dieses White Paper besteht dann darin, Produktmerkmale von POWER7-Servern zu beschreiben, die die inhärenten Power Architecture-RAS-Funktionen vorheriger Servergenerationen erweitern.

Kurzer Vergleich: POWER6 und POWER7



POWER6 (2007)

- 65-nm-Technologie – 341 mm^2
- 0,79 Milliarden Transistoren
- 2 Kerne
  - 2 SMT-Threads/Kern
- 9 Ausführungseinheiten/Kern
  - 2 Ganzzahleinheiten und 2 binäre Gleitkommaeinheiten
  - 1 Vektoreinheit und 1 dezimale Gleitkommaeinheit
  - 2 Load/Store-Einheiten, 1 Verzweigungseinheit
- Integrierter L2-Cache
- L3-Verzeichnis und -Controller (Off-Chip-L3-Cache)
- 2 Speichercontroller



POWER7 (2010)

- 45-nm-Technologie – 567 mm2
- 1,2 Milliarden Transistoren
- 8 Kerne
  - 4 SMT-Threads/Kern
- 12 Ausführungseinheiten/Kern
  - 2 Ganzzahleinheiten und 4 binäre Gleitkommaeinheiten
  - 1 Vektoreinheit und 1 dezimale Gleitkommaeinheit
  - 2 Load/Store-Einheiten, 1 Verzweigungseinheit, 1 Bedingungsregistereinheit (Condition Register Unit)
- Integrierter L2-Cache
- Integrierter L3-Cache
- 2 Speichercontroller

Das POWER7-Modul ist ein weitaus dichterer Chip als sein Vorgänger POWER6 mit nur zwei statt acht Kernen. Außerdem verfügt das POWER7-Modul über 32 MB integrierten (nicht externen) L3-Cache. Im Vergleich zu POWER6 umfasst es auch wesentlich mehr Funktionen mit höheren Ebenen des simultanen Multithreading (SMT) pro Kern.

Zusätzlich zu den Fortschritten bei der Dichte, der Virtualisierung und der Leistung enthält das Prozessormodul wesentliche neue Features für die Zuverlässigkeit und die Verfügbarkeit, die auf dem reichen Erbe früherer Prozessorgenerationen aufbauen.

Die Zuverlässigkeit eines Systems und die Verfügbarkeit der Anwendungen, die es unterstützt, hängen ohne jeden Zweifel von mehr ab als nur von der Zuverlässigkeit der Prozessoren oder auch von der gesamten Systemhardware. Eine vollständige Beschreibung eines Systemaufbaus für RAS muss die gesamte Hardware, die Firmware, das Betriebssystem, die Middleware, die Anwendungen, die Betriebsumgebung, den Arbeitszyklus usw. einschließen.

Eine Definition von RAS

Da die Zuverlässigkeit, die Verfügbarkeit und die Wartungsfreundlichkeit auf viele Arten definiert werden können, ist es ein sinnvoller Ansatz, zunächst zu erläutern, was das IBM Entwicklungsteam unter RAS versteht. Hier also eine technisch etwas ungenaue, jedoch für Server-Hardware nützliche Definition:

Zuverlässigkeit

- Wie selten ein Defekt oder Fehler bei einem Server auftritt

Verfügbarkeit

- Wie selten die Funktionalität eines Systems oder einer Anwendung durch einen Defekt oder einen Fehler beeinträchtigt wird

Wartungsfreundlichkeit

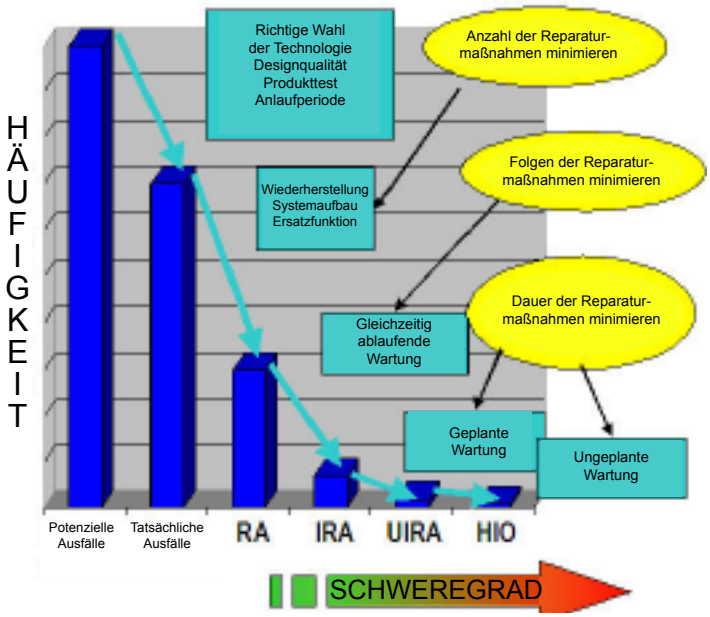
- Wie gut Fehler und deren Auswirkungen gegenüber Endanwendern und Kundendienstmitarbeitern kommuniziert werden und wie effizient sie behoben werden, ohne den Betrieb zu unterbrechen

Laut dieser Definition geht es bei der Hardwarezuverlässigkeit darum, wie oft ein Hardwarefehler eine Systemwartung erforderlich macht – je seltener die Fehler, desto größer die Zuverlässigkeit. Verfügbarkeit bedeutet demnach, wie selten ein solcher Fehler den Betrieb des Systems oder der Anwendung beeinträchtigt. Damit höhere Verfügbarkeitsebenen erreicht werden können, dürfen Hardwarefehler den korrekten Systembetrieb nicht negativ beeinflussen. Anders ausgedrückt: Ein hoch verfügbarer Systemaufbau stellt sicher, dass die meisten Hardwarefehler nicht zu einem Ausfall der Anwendung führen. Eine hohe Wartungsfreundlichkeit ist gegeben, wenn die Fehlerursache schnell erkannt wird und der Fehler (bei einer bestimmten Komponente, bei der Firmware oder bei der Software) effizient behoben wird.

RAS minimiert Ausfallzeiten

Hardware fällt letztlich irgendwann immer einmal aus, unabhängig davon, wie gut sie aufgebaut ist. Die Absicht hinter jedem guten Systemaufbau für Verfügbarkeit besteht darin, den Einfluss dieser Fehler auf den Systembetrieb dramatisch zu reduzieren.

IBM verwendet die voraussichtlichen Auswirkungenn des Ausfalls einer Komponente auf den Kunden als Maßstab für den Erfolg des Verfügbarkeitskonzepts. Diese Messgröße wird als Ausfallzeit der Anwendung, der Partition oder des Systems definiert. IBM klassifiziert herkömmlicherweise Hardwarefehlerereignisse auf mehrere unterschiedliche Arten:



1. **Repair Action (RA)** steht in Beziehung zur Definition von MTBF (Mean Time Between Failures) gemäß Industriestandard. Eine RA ist jedes Hardwareereignis, das bei einem System eine Wartung erfordert. RAs umfassen Störungen, die die Systemverfügbarkeit beeinträchtigen, und Störungen, die ohne Unterbrechung des Systembetriebs behoben werden können.
2. **Interrupt Repair Action (IRA)**. Eine IRA ist ein Hardwareereignis, das eine zu einem geplanten Zeitpunkt stattfindende Systemausfallzeit erfordert, damit es behoben werden kann.
3. **Unscheduled Incident Repair Action (UIRA)**. Eine UIRA ist ein Hardwareereignis, das dazu führt, dass ein System vollständig oder im herabgesetzten Modus neu gestartet werden muss. Das System fällt zu einer ungeplanten Zeit aus. Beim Neustart kann in gewissem Umfang Funktionalität verloren gehen, die verbleibenden Ressourcen werden jedoch für den produktiven Betrieb verfügbar gemacht.
4. **High Impact Outage (HIO)**. Ein HIO ist ein Hardwarefehler, der einen Systemabsturz auslöst, der nicht durch einen sofortigen Neustart behoben werden kann. Dieser Fehler wird normalerweise durch den Ausfall einer Komponente verursacht, die für den Systembetrieb kritisch ist. Dies ist in gewissem Sinne ein Maß für Single Points of Failure des Systems. HIOs wirken sich am stärksten auf das System aus, da Fehlerbehebungen nicht ohne einen Serviceaufruf durchgeführt werden können.

Das klare Ziel für den Systemaufbau von Power Systems besteht darin, zu verhindern, dass Hardwarefehler den Ausfall einer Plattform oder einer Partition verursachen. Die Teileauswahl im Hinblick auf Zuverlässigkeit, Redundanz, Fehlerbehebung und selbstheilende Verfahren sowie herabgesetzte Betriebsmodi werden im Rahmen einer kohärenten, methodischen Strategie eingesetzt, um Anwendungsausfallzeiten zu verhindern.

POWER7-Zuverlässigkeit
Übersicht

Die Zuverlässigkeit oder die Verfügbarkeit eines Systems oder einer Komponente kann anhand verschiedener Messgrößen beschrieben werden. Eine der häufigsten Größen ist die mittlere Zeit zwischen auftretenden Fehlern (Mean Time Between Failures – MTBF), die herkömmlicherweise als die durchschnittliche Nutzungsdauer gilt, bevor eine Komponente nicht mehr ordnungsgemäß funktioniert.

Wenn eine Komponentenbeschreibung besagt, dass die Komponente eine MTBF von 50 Jahren aufweist, ist hiermit nicht garantiert, dass sie 50 Jahre lang betrieben werden kann, bevor ein Fehler auftritt. Diese statistische Größe gibt jedoch an, dass innerhalb einer großen Menge von Komponenten bei einer geringen Ausfallwahrscheinlichkeit einer einzelnen Komponente davon ausgegangen werden kann, dass zwei Prozent der Komponenten pro Jahr ausfallen. (Beispiel: Wenn 100 Server mit einer spezifizierten MBTF von jeweils 50 Jahren ein Jahr lang betrieben werden, fallen voraussichtlich im Durchschnitt zwei Server aus.)

Dies wird am deutlichsten, wenn betrachtet wird, wie Hardwarefehler in einem System bauartbedingt auftreten. Erfahrungsgemäß weisen neu hergestellte Teile eine relativ hohe Fehlerrate auf. Auf eine hohe Fehlerrate in der Anfangslaufzeit folgt ein längerer Zeitraum, in dem Teile gleichmäßig oft, aber selten ausfallen. Anschließend folgt ein Zeitraum, in dem Komponenten zu verschleißen beginnen und in dem vermehrt Fehler auftreten.

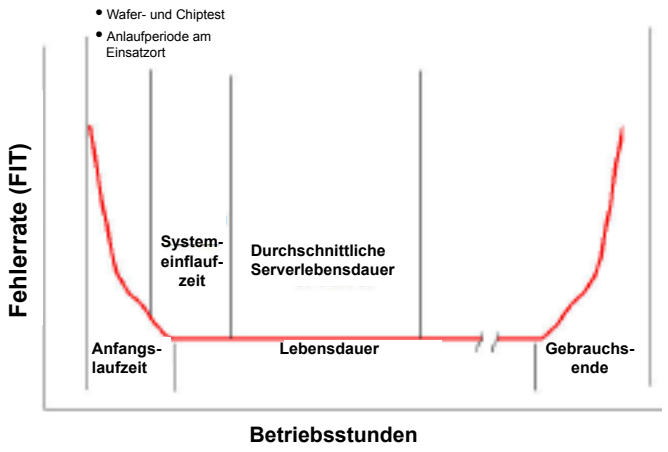
Dies kann veranschaulicht werden, indem in einem Diagramm die Fehler im zeitlichen Verlauf dargestellt werden. Eine Kurve der typischen Fehlerraten ist an ihrem wannenförmigen Verlauf zu erkennen. Ziel des RAS-Systemaufbaus für POWER-Server ist es, dass Systeme getestet, gesiebt und, wenn nötig, „eingefahren“ werden, sodass Komponenten, die einen Ausfall verursachen können, bei der Lieferung an Kunden für die Installation bereits keine „Kinderkrankheiten“ aus der Anfangslaufzeit mehr mitbringen, wie sie in der wannenförmigen Kurve zu sehen sind.

Ausfälle verursachende Teile erst auftritt, nachdem die Systeme außer Betrieb genommen worden sind.

Zu diesem Zweck kommen mehrere unterschiedliche Elemente ins Spiel. Im Folgenden werden einige davon genannt:

- Komponenten mit einem hohen erwarteten Verschleiß wie Lüfter, Stromversorgungskomponenten usw. werden redundant und im laufenden Betrieb reparierbar ausgelegt oder im Systemaufbau anderweitig so ausgelegt, dass ein Defekt bei solchen Komponenten weder zu geplanten noch zu ungeplanten Ausfallzeiten führt.
- Die thermische Systemumgebung wird reguliert und überwacht, damit Komponenten aufgrund thermischer Belastung nicht schnell verschleißen.
- Zu Anschlüssen, die dazu neigen, bei häufigem Gebrauch auszufallen (aufgrund von Korrosion oder aufgrund anderer Faktoren), werden Ersatzsteckerpins hinzugefügt.

Bauartbedingte Zuverlässigkeit als Zeitfunktion



Im Allgemeinen tritt bei Teilen nach ihrer Herstellung eine relativ hohe Fehlerrate auf. Auf eine hohe Fehlerrate in der Anfangslaufzeit folgt ein längerer Zeitraum, in dem Teile gleichmäßig, aber selten ausfallen. Anschließend folgt ein Zeitraum, in dem Komponenten zu verschleißen beginnen und in dem vermehrt Fehler auftreten. Im Fertigungsprozess werden verschiedene Verfahren eingesetzt, um Teilefehler der Anfangslaufzeit festzustellen, bevor Server an Kunden geliefert werden.

Transiente Fehler

Komponentenfehler („dauerhafte“ Ausfälle, die zur bauartbedingten Zuverlässigkeit in Beziehung stehen) sind nicht die einzigen Verursacher von Hardwareausfallzeiten. Verschiedene transiente oder temporäre Fehler können auftreten, wenn die zugrunde liegende Hardware nicht defekt ist. Dazu gehören Fehler aufgrund vorübergehender externer Störungen, zum Beispiel Stromausfälle, elektrostatische Entladungen, durch kosmische Strahlung in die Elektronik abgegebene Energie sowie gelegentliche elektronische Störungen aufgrund sonstiger zufällig auftretender Störquellen.

Beim Aufbau zuverlässiger Systeme müssen die Quellen dieser potenziellen transienten Störungen so analysiert werden, dass Hardware auf geeignete Weise vor ihnen geschützt werden kann.

Zu diesem Zweck verwenden Entwickler Komponenten und entwickeln Verfahren, die auf inhärente Weise Komponenten und Subsysteme vor transienten Fehlern schützen, indem sie zum Beispiel einen Aufbau für Prozessorflipflops verwenden, der allgemein die Energie aus Höhenstrahlungsspitzen verträgt. Außerdem werden Verfahren eingesetzt, mit denen Fehler bei ihrem Auftreten behoben werden können. Beispielsweise kann im Arbeitsspeicher ein DED/SEC-ECC-Code (DED/SEC – Double Error Detect/Single Error Correction) verwendet werden. Dadurch kann ein transientes Störsignal in einem Einzelbit automatisch korrigiert werden.

Systemaufbau und Test

Ausfallsichere Verfahren für den Systemaufbau, eine Komponentenauswahl im Hinblick auf die Zuverlässigkeit, das Wärmemanagement und die Überwachung der Spannungsregulierung sind einige der Verfahren, die das IBM Systementwicklerteam einsetzt, um transiente Fehler zu bekämpfen.

Test zur Überprüfung der automatischen Fehlerbehebung[4]

Damit die Wirksamkeit der RAS-Verfahren im POWER6-Prozessor überprüft werden kann, hat ein IBM Entwicklerteam ein Testszenario erstellt, bei dem zufällige Fehler in die Kerne „injiziert“ werden können.

Entwickler richteten den Strahl eines Protonengenerators auf einen POWER6-Chip. Dadurch wurden energiereiche Protonen mit weitaus höherem Fluss in den Chip injiziert, als dies normalerweise in einem System in einer typischen Anwendung vorkommt. Das Team wandte ein methodisches Vorgehen an, um ein Fehlerabdeckungsmodell mit der unter Testbedingungen gemessenen Systemantwort zu korrelieren.

Das Testteam schloss daraus, dass der Mikroprozessor POWER6 beträchtliche Verbesserungen bei der Behebung von Soft Errors gegenüber den zuvor veröffentlichten Ergebnissen zeigte.

Ausweitung der Tests auf POWER7

Dieses Jahr führte ein IBM Entwicklungsteam eine Reihe von Protoneninjektions- und „Hot Underfill“-Tests mit dem Mikroprozessor POWER7 durch, um den Systemaufbau und die systemeigenen Fehlerbehebungsprozeduren zu prüfen.

POWER7-Testsystem, das im Strahlungsweg befestigt wurde

POWER7 schließt Verfahren zur extensiven Soft-Error-Behebung ein. Dazu gehören Verfahren wie PIR (Processor Instruction Retry) und APR (Alternate Processor Recovery). Diese verleihen eine „Immunität“ gegen Fehler, die bei anderen Prozessordesigns Ausfallzeiten verursachen können.

In einem der Tests versahen Entwickler sechs POWER7-Module mit radioaktivem Underfill-Material und betrieben diese in Power 780-Systemen. In extremem Zeitraffer entsprach dies Tausenden von Betriebsstunden. Bei diesem Verfahren wird das Testen dadurch beschleunigt, dass die Wahrscheinlichkeit von Soft Errors im POWER7-Chip beträchtlich erhöht wird. Dieser Test ist äquivalent zu einem ungefähr eine Million Jahre dauernden Betrieb innerhalb einer normalen Betriebsumgebung. Dabei wurden nahezu 26.600 transiente Fehler künstlich generiert. Lediglich 3 Fehler führten zu einer UIRA. Dies entspricht einer MTBF-UIRA für Soft Errors von über 330.000 Jahren.

Diese Art der Systemprüfung ermöglicht es dem Entwicklerteam, Problemstellungen im Zusammenhang mit der Wiederherstellungsstrategie und -implementierung gründlich zu untersuchen und den Geltungsbereich der Prüfung, die Prioritäten der Prüfung sowie das Design der Flipflops und Arrays zu prüfen.

Protoneninjektion

- Bestrahlen eines POWER7-Moduls mit Protonenstrahlen
  - Injektion von 720 Milliarden energiereichen Protonen
  - Um viele Größenordnungen höherer Fluss als im Normalfall
  - Prüfen der Erkennung/Ausfallsicherheit und der Wiederherstellungsmechanismen im Modul

Testen mit Hot Underfill

- Strahlungsfreies Underfill-Material (zur Abschirmung gegen Alphateilchen) wird ersetzt durch ein verstrahltes Underfill-Material, das die Module der Energie von Alphateilchen aussetzen soll.
- Der Betrieb von Modulen in Systemen wird über einen langen Zeitraum getestet.

[4] Jeffrey W. Kellington, Ryan McBeth, Pia Sanda und Ronald N. Kalla, „IBM POWER6 Processor Soft Error Tolerance Analysis Using Proton Irradiation“, SELSE III (2007).

Er reicht jedoch nicht aus, lediglich den Systemaufbau auf Zuverlässigkeit auszulegen, um einen robusten Schutz vor transienten Fehlern sicherzustellen. IBM erkannte bereits vor langer Zeit, dass Systeme getestet werden müssen, um ihre Ausfallsicherheit für Soft-Error-Ereignisse zu prüfen. Konformitätstests von Systemen bei IBM umfassen das Testen von Stromausfällen mit unterschiedlichen Intensitäten und unterschiedlicher Dauer. Auch Tests von Störungen aufgrund von Blitzen, die in der Nähe einschlagen, werden durchgeführt, die Systeme werden systematisch elektrostatischen Entladungen und Wärme ausgesetzt und es werden Profile erstellt usw.

Außerdem hat IBM in neuerer Zeit beträchtliche Anstrengungen unternommen, Elektronik hohen Strahlungsmengen auszusetzen (mit weitaus höherer Intensität, als sie im praktischen Einsatz vorkommt), um sicherzustellen, dass der Systemaufbau den beabsichtigten hohen Schutz gegen Höhenstrahlung und gegen Emissionen von Alphateilchen bietet.

Mängel im Systemaufbau

Bisweilen wird ein Mangel im Systemaufbau (ein „Bug“ in der Hardware oder der Firmware) als ungeplante Betriebsunterbrechung sichtbar. In aller Regel treten solche Fehler für die seltenen Mängel im Systemaufbau auf, die nur unter sehr besonderen Betriebsbedingungen vorkommen. Diese selten auftretenden Probleme sind oft nur schwer zu erkennen und zu beheben. In anderen Fällen können die für die Fehlerbestimmung oder -behebung eingesetzten Methoden selbst Mängel aufweisen, was zu einer Folge von weiteren Fehlern und zu Systemausfällen führen kann.

Die Vorhersage von Systemausfällen (Bestimmung der Systemzuverlässigkeit), die durch bauartbedingte Fehler verursacht werden, ist kompliziert, aber mit wenig Unsicherheit behaftet. Bei Zuverlässigkeitsberechnungen können auch die vorhergesagten Folgen transienter Fehler berücksichtigt werden. Es ist jedoch sehr schwierig, in diese Berechnungen die Auswirkungen von Mängeln im Systemaufbau einzubeziehen.

Ein gutes Zuverlässigkeitskonzept zielt auf die Beseitigung von Mängeln im Systemaufbau ab, statt diese vorherzusagen. Daher werden Prozessoren und andere von IBM entwickelte angepasste Module vor der Fertigung eingehenden Simulationen unterzogen. Bei der Simulation sollen sowohl die Funktion der Einheiten als auch die Fähigkeit des Systems geprüft werden, Fehler bei ihrem Auftreten zu erkennen und auf diese Fehler zu reagieren. Die Simulation und der Test auf Komponentenebene sind der Ausgangspunkt innerhalb einer mehrstufigen Strategie zur Systemprüfung vor der Freigabe eines Produkts. Entwickler erstellen vollständige IT-Systemumgebungen, in denen die Funktionen ganzer ineinandergreifender Komponentenstacks – Hardware, Firmware, Betriebssystem und Anwendungen – so getestet werden, wie Anwender sie wahrscheinlich einsetzen.

IBM hat viel in den Systemaufbau von Hardware investiert, die Fehlererkennung und -behebung im normalen Betrieb unterstützt. Zu diesem Konzept gehört auch die Fähigkeit, während der Start- und Validierungsphase für den Systemaufbau nach Bedarf Fehler injizieren zu können, damit die Funktionen eines Systems und dessen Fehlerbehandlungsfähigkeiten effizient und gründlich geprüft werden können.

In Power Systems wird eine FFDC-Methodik angewendet, bei der eine umfassende Gruppe von Fehlerprüfprogrammen und FIRs (Fault Isolation Registers) eingesetzt wird, um Fehlerbedingungen in einem Server zu erkennen, einzugrenzen und zu bestimmen. Insbesondere können mit diesem Typ einer automatischen Fehlererfassung und -bestimmung ungeplante Hardwareausfälle schnell behoben werden. Obwohl diese Daten eine Basis für die Analyse von Komponentenfehlern bieten, können sie auch zur Verbesserung der Zuverlässigkeit der Komponente genutzt werden und als Ausgangspunkt für Verbesserungen des Systemaufbaus in zukünftigen Systemen dienen.

IBM RAS-Entwickler verwenden speziell entworfene logische Schaltungen, um Fehler zu generieren, die erkannt und in FIR-Bits gespeichert werden können. Dadurch werden interne Chipfehler simuliert. Dieses Verfahren, das als Fehlerinjektion bezeichnet wird, wird eingesetzt, um Server-RAS-Funktionen und Diagnosefunktionen bei unterschiedlichen Betriebsbedingungen zu prüfen (beim Einschalten, beim Start und im laufenden Betrieb). Fehler werden injiziert, um sowohl die Ausführung geeigneter Analyseroutinen als auch die ordnungsgemäße Funktion von Fehlereingrenzungsprozeduren zu prüfen, die Meldungen an vorgelagerte Anwendungen übermitteln (an POWER Hypervisor, an das Betriebssystem und an Service Focal Point sowie an Service Agent-Anwendungen). Außerdem prüft diese Testmethode, ob Wiederherstellungsalgorithmen aktiviert sind und ob Aktionen zur Systemwiederherstellung ausgeführt werden. Fehlerberichtspfade für Clientbenachrichtigung, Pageraufrufe und Verwendung der Call-Home-Funktion für Service an IBM werden geprüft und RAS-Entwickler sorgen dafür, dass grundlegende und erweiterte Fehlerinformationen aufgezeichnet werden. Ein für Tests zuständiger Kundendienstmitarbeiter „durchläuft“ anschließend unter Verwendung des Wartungspakets Reparaturszenarien, die dem Systemfehler zugeordnet sind. Dadurch soll sichergestellt werden, dass alle Bestandteile des Wartungspakets zusammenarbeiten und dass das System wieder voll funktionsfähig gemacht werden kann. Auf diese Weise wird geprüft, ob RAS-Funktionen und -Komponenten einschließlich des Wartungspakets im Betrieb gemäß den Spezifikationen für den Systemaufbau funktionieren.

Fortlaufende Verbesserung der Zuverlässigkeit

Selbstverständlich tragen das Setzen von Zuverlässigkeitszielen für die Fehlerrate von Komponenten, der Entwurf von Systemen nach Maßgabe bestimmter Verfügbarkeitsmessgrößen, die gründliche Validierung und das gründliche Testen des Verfügbarkeitsentwurfs sowie die Vorhersage der erwarteten Leistung zu einem zuverlässigen Serversystemaufbau bei. Es reicht jedoch nicht aus, einfach Ziele zu setzen und den Systemaufbau zu simulieren.

Darüber hinaus ist es für eine langfristige Systemzuverlässigkeit entscheidend, dass die Leistung implementierter Systeme überwacht, Mängel aufgedeckt und Korrekturmaßnahmen ergriffen werden, sodass die Systemzuverlässigkeit fortlaufend sichergestellt wird.

IBM Entwicklerteams im Außendienst verfolgen und protokollieren Reparaturen von Systemkomponenten, die von der Gewährleistung oder vom Wartungsvertrag abgedeckt werden. IBM Commodity-Manager, die überwachen, wie häufig Teile ausgetauscht werden, stellen für jedes Teil Informationen zur Fehlerrate zusammen und analysieren diese. Falls eine Komponente ihre Zuverlässigkeitsziele verfehlt, erstellt der verantwortliche Commodity-Manager einen Aktionsplan und ergreift geeignete Korrekturmaßnahmen, um das Problem zu lösen.

Commodity-Manager stützen sich auf die FFDC-Methodik von IBM (FFDC - First-Failure Data Capture) und auf die zugeordnete Fehlerprotokollierungsstrategie, erstellen ein genaues Profil der Typen der in der Praxis auftretenden Fehler und leiten Programme ein, um Korrekturmaßnahmen zu ermöglichen. In vielen Fällen können diese Korrekturmaßnahmen eingeleitet werden, ohne darauf zu warten, dass Teile für eine Fehleranalyse zurückgesendet werden.

Das Team des IBM Technischer Außendienstes analysiert kritische Systemfehler fortlaufend. Dazu werden Tests durchgeführt, um zu bestimmen, ob Systemfirmware, Wartungsverfahren und Tools Fehler effektiv beheben und aufzeichnen. Diese Struktur zur fortlaufenden Überwachung und Verbesserung beim Kunden ermöglicht es IBM Entwicklern, mit einer gewissen Bestimmtheit festzustellen, wie Systeme in Kundenumgebungen arbeiten, statt lediglich von Vorhersagen abhängig zu sein. Bei Bedarf nutzen IBM Entwickler diese Informationen, um während des Betriebs Korrekturen vorzunehmen, mit denen aktuelle Produkte während der Implementierung verbessert werden. Diese wertvollen Daten aus Kundenumgebungen werden auch genutzt, um zukünftige Server zu planen und zu entwerfen.

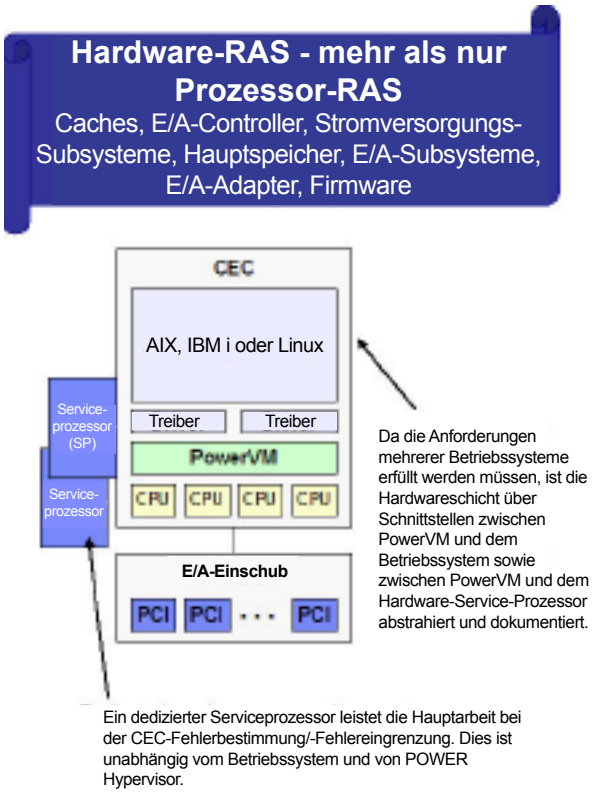
IBM stellt Systeme und einen großen Teil des Stacks her…

Da IBM Rechtsinhaber des Server-Designs ist, Komponenten herstellt, testet und im Rahmen der Gewährleistung und der Wartung für die Systeme Services bereitstellt (einschließlich Prozessoren, Speicherpuffern, E-A-Hub-Controllern, Serviceprozessoren, Firmware usw.), verfügen unsere Systemarchitekten über einen umfassenden Einblick in das Design zuverlässiger Server und über eine umfassende Fähigkeit, alle Teile des Designs zu beeinflussen.

Die Bereitstellung einer hohen Systemzuverlässigkeit erfordert es, den Schwerpunkt entschlossen auf die Optimierung des Systemaufbaus zu legen.

IBM entwickelt, testet, integriert Folgendes:

- E/A-Einschübe
- Prozessoren
- Hauptspeicherpuffer
- E/A-Hubs und andere Hardwarekomponenten
- POWER Hypervisor und VIOS (virtueller E/A-Server)
- Einheitentreiber (AIX™)
- Betriebssystem
- Middleware
- Cluster-Software (PowerHA SystemMirror™, Systempools)
- Management-Software (IBM Systems Director™)



Da IBM Eigner des Server-Designs ist, Komponenten herstellt, testet und im Rahmen der Gewährleistung und der Wartung Services für die Systeme bereitstellt (einschließlich Prozessoren, Speicherpuffern, E-A-Hub-Controllern, Serviceprozessoren, Firmware usw.), verfügen IBM Systemarchitekten über einen umfassenden Einblick in das Design zuverlässiger Server und über eine umfassende Fähigkeit, alle Teile des Systemaufbaus zu beeinflussen. Beim Entwurf einer bestimmten Komponente

gibt es die normale Tendenz, die Komponente möglichst „gut“ oder möglichst „schnell“ zu konzipieren. Das Risiko besteht darin, dass eine Komponente zwar nach bestimmten Kriterien optimiert sein kann, aber möglicherweise innerhalb der Architektur eines Systems nicht optimal funktioniert. Ein solcher Designprozess wird als Suboptimierung bezeichnet. Obwohl Komponenten für sich alleine gut funktionieren, wird die Gesamtleistung oder die Verfügbarkeit des Servers beeinträchtigt. Diese Art des Server-Designs kann sich ergeben, wenn ein Hersteller Komponenten aus verschiedenen Quellen erhält und daraus Systeme „zusammenzubasteln“ versucht.

Ein vorrangiges Ziel des IBM Power-Entwicklungsteams ist es, „intelligentere“ Systeme bereitzustellen, die eine aufeinander abgestimmte Leistung aufweisen (Kopplung von Prozessorgeschwindigkeit und Hauptspeicherkapazität sowie E-A-Bandbreite) und gut mit der Systemsoftware (Betriebssystem, Middleware, Anwendungen) harmonieren.

IBM ist in der IT-Branche insofern einzigartig positioniert, als IBM Systemschlüsselkomponenten konzipiert, baut (oder IBM Entwickler den Aufbau vorgeben und die Fertigung überwachen und kontrollieren), integriert und testet. IBM verwandelt also rohes Silizium in Server und integriert Betriebssystem, Middleware und Management-Software. Diese umfassende Kontrolle gibt darüber hinaus IBM Entwicklern ein sehr effizientes Mittel an die Hand, Lösungen zu optimieren und zu entscheiden, ob eine Designänderung der Hardware oder Firmware erforderlich ist bzw. wie diese hergestellt, getestet oder gewartet wird.

Verbesserungen am Systemaufbau für POWER7-Zuverlässigkeit

Aus dem Blickwinkel von RAS beginnt der Systemaufbauprozess mit der Definition eines Konzepts für die Zuverlässigkeit. Eine unternehmensweite IBM Richtlinie legt fest, dass neue Produkte eine mindestens so hohe Zuverlässigkeit bieten müssen wie ihre Vorgänger. Dieses Ergebnis wird erwartet, obwohl jede neue Servergeneration normalerweise eine höhere Leistung und Kapazität bereitstellt – obwohl neue Komponenten mehr Ausfälle bedeuten können (da weitere Kerne, Hauptspeicher oder E/A-Kanäle hinzukommen).

Der Übergang von POWER6 zu POWER7 stellte für das RAS-Team eine interessante Herausforderung dar. IBM kann einerseits aufgrund der dramatischen Erhöhung der Anzahl Kerne pro Chip Server mit gleich vielen Kernen bereitstellen, die jedoch auf viel weniger Prozessorchips passen (weniger Fehlerquellen). Darüber hinaus ist in POWER7 ein L3-Cache integriert, wodurch sich die erforderliche Chipanzahl in umfangreicheren Server-Designs reduziert. Andererseits liefert IBM Server mit weitaus mehr Kernen (bis zu 256 Kerne bei Power 795) – und es kommen beträchtlich mehr Hauptspeicher- und E/A-Subsysteme hinzu, damit diese Kerne in optimalen Leistungsbereichen arbeiten können.

Als zusätzliche Herausforderung für RAS-Entwickler führte das Herabsetzen der Größe der Prozessorkomponente auf die 45-nm-Technologie zu einer potenziellen Vermehrung transienter Fehler. IBM führte in Servern, die auf POWER6 basieren, die Verfahren Processor Instruction Retry und Alternate Processor Recovery ein, um diese Arten von Soft Errors zu verhindern. Bei POWER7 wird diese Technologie weiterhin eingesetzt. Darüber hinaus wird jedoch deren gesamte Widerstandsfähigkeit gegen diese Art von Ereignis erweitert: Dazu dient der „Stacked Latch“-Systemaufbau (mit übereinander angeordneten Flipflops), der mit bauartbedingt geringerer Wahrscheinlichkeit den Zustand ändert, falls eine energetische Entladung aufgrund von Höhenstrahlung oder Alphateilchen auftritt.

Während weiterhin der Schwerpunkt auf der Verhinderung von Soft Errors liegt, ist der Speicherbus in Servern, die auf dem Prozessor POWER7 basieren, anders aufgebaut: Er ist vor Soft Errors dadurch geschützt, dass Fehler mithilfe einer CRC-Prüfung erkannt und fehlerhafte Operationen wiederholt werden. Der Bus ist, wie bei POWER6 eingeführt, dazu in der Lage, eine fehlerhafte Datenbitzeile auf dem Bus dynamisch durch eine Ersatzbitzeile zu ersetzen.

Wegen der von Kunden erwarteten dramatischen Vergrößerung der Gesamtspeicherkapazität, die im System verfügbar ist, hat sich für größere Server auch bei POWER7 die Kapazität des Zusatzspeichers auf angepassten DIMMs beträchtlich erhöht. Die Verwendung von Zusatz-DRAM-Chips bietet für ausgewählte POWER7-Server im oberen Leistungsbereich eine „Selbstheilungsfähigkeit“ im Speicher, bei der Daten aus ausgefallenen DRAM-Chips automatisch in verfügbare Zusatzspeicher verschoben werden. Mit diesem Design kann ein DIMM-Paar, pro Speicherbank einen und manchmal zwei DRAM-Fehler verkraften, ohne die DIMMs zu ersetzen und dabei die Chipkill-Erkennung und -Fehlerbehebungsfunktionalität zu erhalten. (Es können potenziell drei DRAM-Chipfehler für jedes Speicherbankpaar auf einem DIMM-Paar toleriert werden.) Weitere Einzelheiten hierzu finden Sie im Abschnitt „Speicher-DIMMs und ECC-Wörter“ auf Seite 27.

Fehlererkennung und Fehlereingrenzung

Einführung

Damit Systeme in einem umfassenderen Sinne zuverlässig sind, müssen Fehler, die sich auf Rechenoperationen auswirken können, erkannt und behoben werden, sodass falsche Ergebnisse nicht bis in Kundenanwendungen gelangen. Systeme, die für diese Aufgabe ausgelegt sind, erfordern eine umfassende Architektur für die Fehlererkennung, -behebung und -eingrenzung.

Eine derartige Infrastruktur muss auch dazu in der Lage sein, Fehler in den Daten zu erkennen und, wo auch immer sie gespeichert sind, möglichst zu korrigieren. Die Beibehaltung zumindest der ECC-Codes für die Erkennung von Doppelbitfehlern und für die Korrektur von Einzelbitfehlern, um gespeicherte Daten zu prüfen, kann dazu beitragen, dieses Ziel zu erreichen.

Obwohl der Umfang der Elektronik, die für derartige ECC-Korrekturcodes reserviert ist, hoch sein kann, muss die Fehlererkennung weit mehr als nur die ECC-Datenprüfung einschließen. Damit die Fehlererkennung wirksam ist, muss darüber hinaus Folgendes erfüllt sein:

- Es müssen Fälle erkannt werden können, in denen Daten an einer falschen Position gespeichert sind oder von einer falschen Position abgerufen werden, oder Fälle, in denen eine Anforderung zum Speichern oder Abrufen von Daten nicht beantwortet wurde.
- Es müssen Mechanismen vorhanden sein, mit denen geprüft werden kann, ob die Logikfunktionen von verschiedenen Verarbeitungskomponenten und von Komponenten, die zum Verschieben von Daten dienen, ebenfalls fehlerfrei sind. Beispiele:
  - Fehlererkennung und Datenkorrektur in Flipflops und Arrays, in denen der logische Status gespeichert wird
  - Prozesse, die ungültige Zustandsmaschinenübergänge markieren, die durch Komponentenfehler und transiente Bedingungen verursacht wurden
  - Mathematische Mechanismen zur statistischen Validierung der Ergebnisse, die Rechenwerke liefern (zum Beispiel mithilfe von Verfahren wie der Modulo-Prüfung)

Damit Systeme mit nicht behebbaren Fehlern effizient gewartet werden, muss eine umfassende Designmethodik nicht nur einen Fehler erkennen. Sie muss auch über ausreichend Daten verfügen, um die Fehlerquelle zu ermitteln, damit eine bestimmte Komponente (bzw. Komponenten) korrekt identifiziert werden kann (bzw. können). Diese Teile müssen möglicherweise dekonfiguriert werden, damit ein Fehler nicht erneut auftritt, und bei Bedarf letztendlich ausgetauscht werden.

Diese Architektur erfordert, dass Fehlerprüfprogramme überall im System verteilt sein müssen, damit die eigentliche Fehlerursache ermittelt werden kann. Zum Beispiel reicht es nicht aus, einfach die Gültigkeit von Daten zu prüfen (beispielsweise mithilfe von ECC), wenn sie in einen Datencache eines Prozessors oder in den Hauptspeicher geschrieben werden. Die Daten müssen an jedem Punkt geprüft werden, an dem sie die Grenze von einer Komponente zu einer anderen überschreiten (also überall, wo sie beschädigt werden können).

Damit ein Server das höchste Niveau der Fehlererkennung und -behebung erreicht, muss er einen Fehler ***sowohl*** bei seinem Auftreten identifizieren können ***als auch*** über einen Mechanismus verfügen, auf den Fehler automatisch im laufenden Betrieb zu ***reagieren***.

Bei einem derartigen Aufbau kann es zum Beispiel erforderlich sein, dass eine Gleitkommaeinheit die Modulo-Prüfung der Ausführung einer Gleitkommaanweisung beendet haben muss, bevor die nächste Anweisung beginnt. Dadurch soll eine fehlgeschlagene Anweisung unmittelbar nach dem eventuellen Auftreten eines Fehlers wiederholt werden können.

Schließlich gilt für die Fehlererkennung und Fehlereingrenzung Folgendes:

- Sie muss in der Hardware und in der Firmware automatisch ausgeführt werden, ohne dass ein Diagnoseprogramm einen Fehler reproduzieren muss und ohne dass ein Mensch in irgendeiner Weise Fehlerindikatoren interpretieren muss.
- Sie muss ***unabhängig vom Betriebssystem*** in der Hardware/Firmware ausgeführt werden und idealerweise auch unabhängig von den Systemverarbeitungsressourcen sein (sodass die Server-Hardware mehrere Betriebssysteme effizient unterstützen kann und die dem System zur Verfügung stehenden IT-Ressourcen nicht reduziert werden).

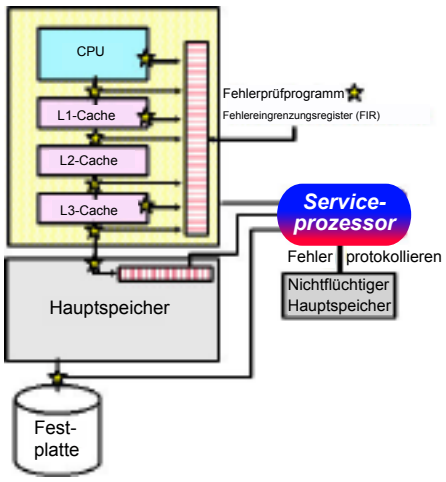
First Failure Data Capture (FFDC) mit dedizierten Serviceprozessoren

Komponenten in der zentralen Elektronik (Prozessoren, Hauptspeicherpuffer und E/A-Hub-Controller) verwenden hardwarebasierte Fehlerdetektoren, um interne Systemkomponenten umfassend zu instrumentieren.

Diese Architektur wird als First Failure Data Capture (FFDC) bezeichnet. Sie erfordert es, dass ein Server über „eingebaute“ Stationen zur Hardwarefehlerprüfung verfügt, die Fehlerbedingungen erfassen und identifizieren. Ein Power 795-Server mit 256 Kernen umfasst zum Beispiel über 598.000 Prüfprogramme, die Fehlerbedingungen erfassen und identifizieren. Diese werden in über 96.000 Bits des Fehlereingrenzungsregisters (Fault Isolation Register) gespeichert. Die einzelnen Detektoren sind Diagnosetests, die in der Lage sind, einem dedizierten Serviceprozessor Fehlerdetails zu melden. Relevante Fehlerdaten, die mit dem Fehler in Beziehung stehen, werden in Fehlereingrenzungsregistern aufgezeichnet und erfasst sowie mit spezialisierter Analyselogik für die Fehlerrelevanz verarbeitet.

Wenn bei Anwendung dieses Verfahrens in einem Server ein Fehler erkannt wird, wird die Fehlerursache normalerweise erfasst, ohne den Fehler reproduzieren und ohne eine umfangreiche Zurückverfolgung oder Diagnoseprogramme ausführen zu müssen. Für die überwältigende Mehrheit der Fehler bedeutet ein gutes FFDC-Design, dass die Fehlerursache auch automatisch und ohne Kundendienstmitarbeiter erkannt werden kann.

Der Serviceprozessor ist ein dedizierter Mikroprozessor, der vom POWER7-Mikroprozessor unabhängig ist. Er wird eingesetzt, um Fehlerinformationen zu korrelieren und Fehlerbehebungsmechanismen in Verbindung mit der Hardware und mit POWER Hypervisor (Firmware) zu koordinieren. Der Serviceprozessor empfängt immer dann Signale, wenn die Status der Fehlereingrenzungsregister sich ändern, und er kann jederzeit auf Registerdaten zugreifen, ohne die Leistung dieser Module zu beeinflussen.



Die Erfassung von Fehlerdaten beim ersten Auftreten (FFDC - First Failure Data Capture) spielt eine entscheidende Rolle, wenn Server bereitgestellt werden sollen, die eine Selbstdiagnose und eine Selbstheilung durchführen können. Mithilfe von Tausenden von Prüfprogrammen (Diagnosetests), die an kritischen Nahtstellen im gesamten Server implementiert werden, „fängt“ das System Hardwarefehler im laufenden Betrieb „ab“.

Mit dem separat mit Strom versorgten Serviceprozessor werden anschließend die Prüfprogramme analysiert und die Fehler bestimmt. Mithilfe dieser Methode umgeht IBM die üblicherweise eingesetzte nicht unterbrechungsfreie Fehlererkennungsstrategie „Neustart und neuer Versuch“, da das System mit gewisser Sicherheit das fehlerhafte Teil kennt. Bei dieser automatisierten Lösung kann die Laufzeitfehlerdiagnose deterministisch sein, sodass für jede Prüfstation der eindeutige Fehlerbereich für dieses Prüfprogramm definiert und dokumentiert ist. Schließlich wird aus dem Fehlerbereich der FRU-Aufruf (FRU - Field-Replacable Unit) und normalerweise müssen die Daten nicht manuell interpretiert werden. E/A-Adapter und -Einheiten, die Partitionen physisch zugeordnet sind, verwalten normalerweise die Fehlererkennung und -meldung von Partitionen an Einheitentreiber des Betriebssystems.

Diese Architektur ist auch die Basis für die Analyse vorhersehbarer Fehler (Predictive Failure Analysis) von IBM, da der Serviceprozessor sporadisch auftretende Komponentenfehler nun zählen und protokollieren sowie Zuordnungen aufheben und andere Korrekturmaßnahmen ergreifen kann, sobald eine Fehlerschwelle erreicht wird.

Zentrale Elektronik (CEC – Central Electronics Complex)

- Integrierte Hardwarefehlerprüfprogramme
  - Erkennen, wo und wann Fehler auftreten I
  - Fehler auf die ausgefallene Komponente oder Schnittstelle eingrenzen
    - √ Fehler auf die ausgefallene Komponente oder Schnittstelle eingrenzen
  - Informationen in besonderen Fehlereingrenzungsregistern (FIR – Fault Isolation Registers) aufzeichnen
- Hardware/Firmware führt geeignete interne Maßnahmen mit dem folgendem Ziel aus
  - Fehler möglichst automatisch beheben
  - Serviceprozessor oder Hypervisor über behebbare Fehler benachrichtigen
  - Erforderliche Maßnahmen ausführen, um den Schaden infolge nicht behebbarer Fehler zu minimieren
- Keine Notwendigkeit, Fehler zu reproduzieren, um Fehler zu erkennen/einzugrenzen

Serviceprozessor

- Empfängt Fehlerberichte und liest Fehlereingrenzungsregister (FIRs - Fault Isolation Registers).
- Ergreift Maßnahmen bei Fehlern:
  - Implementiert Wiederherstellungsaktionen.I
  - Ermöglicht Selbstheilungsfunktionen.
  - Gibt Ressourcen frei.
    - √ Als Prognose, die auf der Häufigkeit von Ereignissen basiert (Schwellenwerte)
    - √ Als Antwort auf Fehlerberichte
- Reagiert auf nicht behebbare Fehler.
  - Erfasst zusätzliche erweiterte Fehlerinformationen.
  - Leitet Wiederherstellungsmaßnahmen für die Partition oder das System gemäß Konfiguration ein.
- Einige Typen von Fehlern behandelt möglicherweise POWER Hypervisor.

Betriebssysteme und Anwendungen

- Nicht an der Fehlererkennung/Fehlereingrenzung oder -wiederherstellung von CEC-Komponentenfehlern beteiligt
  - Fehlererkennung und Fehlereingrenzung hängen bei diesen Komponenten stark von den Betriebssystemen ab. (Das Betriebssystem muss nicht gewechselt werden, damit die meisten RAS-Funktionen genutzt werden können.)
- Unter bestimmten Umständen möglicherweise nützlich bei der dynamischen Aufhebung der Prozessorzuordnung (Dynamic Processor Deallocation)

Die RAS-Architektur koppelt dann FFDC mit ***automatisierter*** Firmwareanalyse, um die einem Fehler zugrunde liegende Ursache bei seinem ersten Auftreten schnell und präzise zu ermitteln, unabhängig von der Phase des Systembetriebs und ohne eine Diagnose mit Fehlerreproduktion ausführen zu müssen. Dabei ist es entscheidend, zu bestimmen, ***welche Komponente*** einen Fehler – ***bei seinem ersten Auftreten*** – verursacht hat, und dessen erneutes Auftreten zu verhindern.

IBM Entwickler steuern das gesamte CEC-Design (vom Silizium bis hin zur Architektur) und die RAS-Architekten korrelieren Fehlerereignisse im gesamten System. Damit eine korrekte Fehlerabdeckung und -erkennung sichergestellt wird, enthalten Power-Server eine ***zusätzliche*** Fehlerinfrastruktur, damit die ursprüngliche Fehlerquelle einfacher erkannt werden kann, auch wenn sich der Fehler auf mehrere Komponenten verbreitet hat (zum Beispiel Daten aus dem Hauptspeicher, die nicht behebbare Fehler aufweisen).

Zusätzlich zum Zugriff auf Fehlereingrenzungsregister ist der er Serviceprozessor beim Auftreten eines schwerwiegenden Fehlers in der Lage, den Status in Modulen enthaltener funktioneller Flipflops zu überprüfen, um beim Auftreten des Fehlers eine Momentaufnahme des Status der Prozessoren und anderer Komponenten zu erstellen. Diese zusätzlichen Informationen, die als erweiterte Fehlerdaten bezeichnet werden, können beim Auftreten eines Fehlers an IBM übertragen werden. Dazu werden automatisierte Call-Home-Funktionen verwendet (sofern der Systemadministrator diese aktiviert hat), sodass IBM Kundendienstmitarbeiter die Begleitumstände des Fehlerereignisses feststellen können. Dieser Prozess unterstützt eine detaillierte Fehleranalyse und stellt somit sicher, dass systematisch auftretende Probleme im Rahmen eines fortlaufenden Qualitätsverbesserungsprozesses schnell erkannt und behandelt werden [siehe „Fortlaufende Verbesserung der Zuverlässigkeit“ auf Seite 12].

Verfügbarkeit der zentralen Elektronik (CEC – Central Electronics Complex)

Systemaufbau für Verfügbarkeit

Die Verfügbarkeitsstrategie für IBM Power ist tief verwurzelt in einer jahrzehntelangen Geschichte der Entwicklung von Mainframes. Durch die Nutzung eines umfassenden Hintergrunds in Predictive Failure Analysis™ (Analyse vorhersehbarer Fehler) und in der dynamischen Systemanpassung hat das IBM Verfügbarkeitsteam dazu beigetragen, einen einzigartigen Prozessor zu entwickeln, der dem Kunden einen hohen Mehrwert bietet.

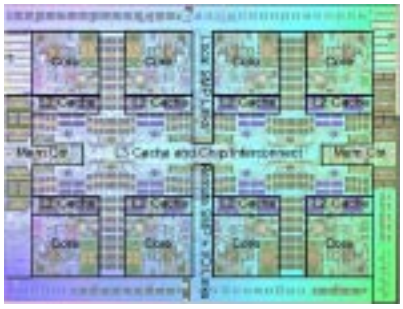
Zusammen ergeben diese Merkmale den Eckpfeiler für die POWER7-Verfügbarkeitsstrategie, wie sie im Lieferumfang der Power-Serverfamilie enthalten ist. Diese Leistungsmerkmale ermöglichen einen völlig neuen Grad der Eigenintelligenz in Systemen, die auf dem Prozessor POWER7 basieren. Das FFDC-Verfahren ermöglicht es dem System, Situationen proaktiv zu analysieren, die auf einen bevorstehenden Fehler hinweisen, und setzt die betreffende Komponente außer Betrieb, bevor diese die Systemstabilität gefährdet.

Zum qualitativen Verständnis des Systemaufbaus für die Verfügbarkeit von Power Systems ist es erforderlich, die Verfügbarkeitsmerkmale der einzelnen Komponenten des Systems zu bestimmen. Dies gilt für Hardware – Prozessoren, Hauptspeicher, E/A-Adapter, Stromversorgungs- und Kühlungssubsysteme sowie Gehäuse und Firmware, soweit diese zur Funktion der Hardware in Beziehung steht – und für die einzelnen im Server eingesetzten Softwarekomponenten. Selbst wenn eine RAS-Beschreibung sich auf Hardwareelemente beschränkt, muss diese daher im Kontext des Gesamtsystems untersucht werden: Sie darf sich nicht nur auf ein einzelnes Element konzentrieren, zum Beispiel nur auf einen Prozessorkern oder auf eine Hauptspeicherkomponente.

In den folgenden Abschnitten werden die Leistungsmerkmale von RAS insbesondere im Hinblick auf die Funktionalität in Systemen für Unternehmen im oberen Leistungsbereich beleuchtet. Die Methodik für das Server-Design konzentriert sich üblicherweise zunächst einmal auf die sehr strengen Anforderungen an die Verfügbarkeit großer, hochgradig virtualisierter Hochleistungssysteme wie Power 795. RAS-Funktionen, die für diese Klasse von Systemen entwickelt werden, fließen anschließend häufig auch in die leistungsschwächeren Modelle der Produktlinie mit ein. Dadurch sollen die Investitionen in die Prozessor- oder Firmwarewiederherstellung genutzt und auch Einstiegsserver mit Verfügbarkeitsmerkmalen des oberen Leistungsbereichs ausgestattet werden. Eine Tabelle mit RAS-Funktionen nach Modellen finden Sie in Anhang A: Systemunterstützung für ausgewählte RAS-Funktionen [Seite 51].

POWER7-Prozessor

POWER7-Prozessorchip



Der POWER7-Chip weist acht Kerne auf, die jeweils Single Threading und simultanes Multithreading (SMT) beherrschen (mit der Fähigkeit, vier Threads auf demselben Kern zur gleichen Zeit auszuführen). POWER7 bleibt kompatibel mit vorhandenen POWER6-basierten Systemen, damit sichergestellt wird, dass Binärdateien auf den neueren Systemen weiterhin ordnungsgemäß ausgeführt werden. POWER7-Technologie unterstützt Virtualisierungstechnologien wie auch ihr Vorgänger POWER6 und weist sowohl auf Chipebene als auch auf Systemebene eine verbesserte Zuverlässigkeit und Wartungsfreundlichkeit auf. Damit die Datenbandbreite eines Prozessors mit acht Kernen und 4 GHz Taktfrequenz (bzw. im TurboCore-Modus 4,25 GHz) unterstützt wird, verwendet der POWER7-Chip eingebetteten dynamischen Arbeitsspeicher (eDRAM), um einen intelligenten 32-MB-L3-Cache auf dem Prozessorchip zu integrieren. Dieser Cache reduziert die Latenzzeit des Datenzugriffs im Vergleich zur Implementierung eines externen Cache beträchtlich, nimmt jedoch ungefähr ein Drittel des Platzes ein und weist ein Fünftel des Standby-Strombedarfs auf. Zugleich schützt er im Vergleich zum konventionellen statischen Arbeitsspeicher (SRAM) wesentlich besser vor Soft Errors.

Jeder Vier-Wege-SMT-Kern im POWER7-Mikroprozessor basiert auf einem Superskalardesign und schließt zwölf Ausführungseinheiten für Anweisungen ein. Der POWER7-Chip stellt Fehlerbehebungshardware für Processor Instruction Retry für den automatischen Neustart von Workloads auf demselben oder einem alternativen Kern im selben Server bereit und verwendet darüber hinaus einen „Stacked Latch"-Systemaufbau (mit übereinander angeordneten Flipflops), der die Störunempfindlichkeit gegenüber Soft Errors verbessert, und er umfasst viele verschiedene Features für die Energieverwaltung, einschließlich eines „Schlaf-" und eines „Schlummermodus" sowie einer dynamischen Verwaltung der Taktfrequenz und der Spannung.

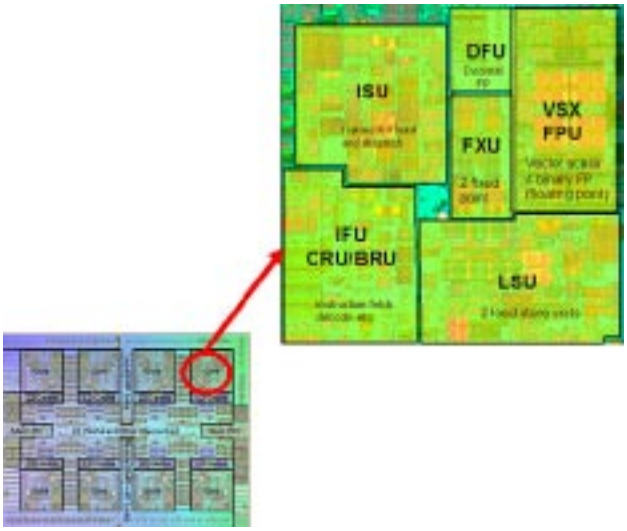
Der POWER7-Prozessorchip besteht aus acht Prozessorkernen, die jeweils einen 32-KB-8-Wege-Assoziativdatencache, einen 32-KB-2-Wege-Assoziativinstruktionscache und einen eng gekoppelten 256–KB-8-Wege-L2-Assoziativcache umfassen. Darüber hinaus verwendet der Chip eingebettete dynamische Arbeitsspeichertechnologie (eDRAM), um 32 MB internen L3-Cache zur Verfügung zu stellen. Dieser Onboard-L3-Cache ist sehr wichtig, um einen Systemaufbau mit ausgeglichener Auslastung bereitzustellen, da dieser Folgendes bietet:

- Gegenüber externem L3-Cache eine Verbesserung der Latenzzeit im Verhältnis von bis zu 6 zu 1 für L3-Zugriffe, da keine Treiber oder Empfänger im L3-Zugriffspfad außerhalb des Chips vorhanden sind
- Eine doppelt so große Bandbreite aufgrund der Verbindungen auf dem Chip selbst
- Eine Reduzierung des Platzbedarfs gegenüber herkömmlichem SRAM-Speicher
  - √ 1/3 des Platzbedarfs einer herkömmlichen SRAM-Implementierung
  - √ 1/5 des Standby-Stromverbrauchs
- Ein innovatives, intelligentes Cache-Design mit einer schnellen 4-MB-L3-Region pro Prozessorkern, in der das System Folgendes ausführt:
  - √ Automatisches Migrieren von privatem Speicherbedarf (bis zu 4 MB) in diese schnelle lokale Region (pro Kern) bei ca. fünfmal geringerer Latenzzeit im Vergleich zu vollständigem L3-Cache
  - √ Möglichkeit des automatischen Klonens gemeinsamer Daten auf mehrere private Regionen.

POWER7 schließt darüber hinaus zwei Speichercontroller (jeweils mit vier Hochgeschwindigkeitsspeicherkanälen) und eine Reihe von Buscontrollern ein, die Schnittstellen zu anderen Prozessormodulen, zu E/A-Hub-Controllern und zu einem Serviceprozessor verwalten.

POWER7-Prozessorkern

Jeder 64-Bit-Prozessorkern kann Anweisungen in anderer Reihenfolge ausführen (OoOE – Out-of-Order Execution) und ist dazu in der Lage, 6 Anweisungen pro Zyklus zuzuteilen und 8 Anweisungen pro Zyklus abzusetzen. Die einzelnen Kerne unterstützen 12 Ausführungseinheiten für Anweisungen. Dazu gehören 2 für Festkommaverarbeitung (FXU), 2 Load/Store-Einheiten (LSU – Load/Store Unit), 4 Einheiten für Gleitkommaoperationen mit doppelter Genauigkeit (FPU – Floating Point Unit), 1 Vektorprozessoreinheit (VMX), 1 Verzweigungseinheit (BRU - Branch Unit), 1 Bedingungsregistereinheit (CRU – Condition Register Unit) und 1 Einheit für dezimale Gleitkommaoperationen (DFU – Decimal Floating Point Unit).



Ausgewählte wichtige RAS-Funktionen

- „retry“ und „set delete“ für Instruktions- und Datencache
- Processor Instruction Retry (PIR)
- Alternate Processor Recovery (APR)
- Wiederherstellung bei Blockierung
- GPR-ECC mit Fehlerbehebung
- Core-Contained-Checkstop-Fehler)

POWER7-Prozessorkern

Jeder 64-Bit-Prozessorkern kann Anweisungen in anderer Reihenfolge ausführen (OoOE – Out-of-Order Execution) und ist dazu in der Lage, 6 Anweisungen pro Zyklus zuzuteilen und 8 Anweisungen pro Zyklus abzusetzen. Die einzelnen Kerne unterstützen 12 Ausführungseinheiten für Anweisungen. Dazu gehören 2 für Festkommaverarbeitung, 2 Load/Store-Einheiten, 4 Einheiten für Gleitkommaoperationen mit doppelter Genauigkeit, 1 Vektorprozessoreinheit, 1 Verzweigungseinheit, 1 Bedingungsregistereinheit und 1 Einheit für dezimale Gleitkommaoperationen.

Die einzelnen Kerne mit variabler Taktfrequenz (bis zu 4,25 GHz) können Instruktionsströme im Einzelthreadmodus (Single Threading) oder im Modus für simultanes 2- oder 4-Wege-Multithreading (SMT) verarbeiten. SMT-Modi können für jede virtuelle Maschine separat dynamisch festgelegt oder vom Betriebssystem zur Laufzeit für eine optimale Leistung geändert werden.

Damit die höchste Serververfügbarkeit und -integrität erreicht wird, müssen FFDC und Sicherheitseinrichtungen für die Fehlerbehebung die Gültigkeit von Benutzerdaten überall im Server schützen. Dazu gehören auch alle internen Speicherbereiche sowie die Busse für die Datenübertragung. Genauso wichtig ist es, den ordnungsgemäßen Betrieb interner Flipflops (Register), Arrays und Logik innerhalb eines Prozessorkerns zu authentifizieren, die die Ausführungssteuerung des Systems umfassen, und geeignete Maßnahmen zu ergreifen, wenn ein Fehler erkannt wird.

Fehlerbehebung beim Kern – Dynamic Processor Deallocation, Processor Instruction Retry und Alternate Processor Recovery

In den POWER6-basierten Servern wurde ein Prozessorkerndesign eingeführt, das umfassende Einrichtungen zur Erkennung und Eingrenzung vieler verschiedener Fehler während der Anweisungsverarbeitung einschließt. Da eine genaue Kopie des CPU-Status gespeichert ist, kann der Kern eine fehlgeschlagene Anweisung wiederholen und dadurch zahlreiche transiente Fehler verhindern.

Bei diesen Systemen wurden einige Leistungsmerkmale eingeführt, die dynamische Anpassungen unterstützen, sobald Probleme auftreten, die die Verfügbarkeit gefährden. Insbesondere wird die PIR-Tool-Suite (PIR - Processor Instruction Retry) implementiert, die die Leistungsmerkmale Processor Instruction Retry, Alternate Processor Recovery, Partition Availability Prioritization und Single Processor Checkstop umfasst. Zusammengenommen ermöglichen es diese Leistungsmerkmale in vielen Fehlerszenarios einem Power-Server (POWER6 oder POWER7), den Fehler transparent zu beheben, ohne eine Partition zu beeinflussen, die den Kern verwendet. Dieses Bündel von Leistungsmerkmalen wird in einem White Paper zu RAS-Funktionen von POWER6 detailliert beschrieben[5].

Processor Instruction Retry (Wiederholung einer aufgrund eines transienten Fehlers fehlgeschlagenen Operation) ist ein wichtiges Mittel, mit dem ein POWER6- oder POWER7-System transiente Fehler im Prozessorkern behebt. Diese Prozessorchips enthalten darüber hinaus Mechanismen zur Behandlung von Soft Errors, die mit POWER Hypervisor zusammenarbeiten. Diese Verfahren zur Behebung transienter Fehler werden ohne Beteiligung des Betriebssystems ausgeführt.

Obwohl jedoch der Schutz vor transienten Fehlern oder Soft Errors ein wichtiger Verfügbarkeitsaspekt ist, reicht dies für einen vollständigen RAS-Systemaufbau nicht aus. Entwickler müssen darüber hinaus die Rate für „permanente“ Fehler von Komponenten berücksichtigen. Aus diesem Grund ist in POWER6 ein Mechanismus für die Erstellung eines Prüfpunkts integriert, der den Betriebszustand des Kerns vor der Ausführung einer Anweisung beschreibt. Somit kann POWER Hypervisor in vielen Fällen die Statusinformationen mit dem Betriebszustand eines permanent ausgefallenen Prozessorkerns in einen alternativen Prozessorkern laden – und dabei die Threadausführung ohne Anwendungsunterbrechung dynamisch verlagern.

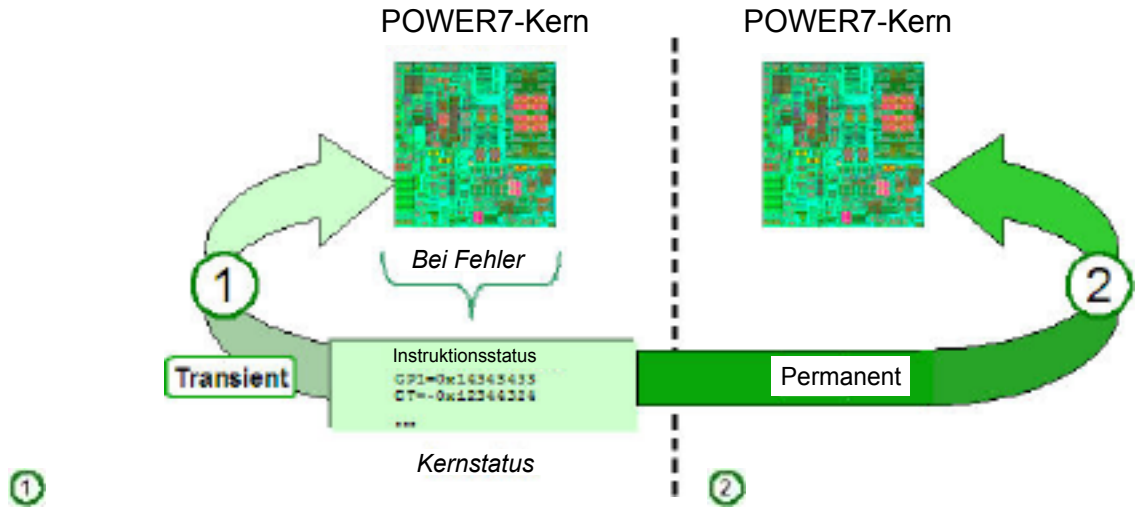
[5] Jim Mitchell, Daniel Henderson, George Ahrens und Julissa Villarreal: „IBM Power Platform Reliability, Availability, and Serviceability (RAS). Highly Available IBM Power Systems Servers for Business-Critical Applications“, POW03003-USEN-03 (5. Juni 2009).

Dieses Verfahren wird als Alternate Processor Recovery bezeichnet. Mit ***Alternate Processor Recovery*** können Systeme einige permanente Fehler im Prozessor beheben, die andernfalls zu Ausfällen von Partitionen oder Gesamtsystemen führen. Dieses Produktmerkmal wird nun von POWER7-Servern unterstützt.

> *Ein wichtiger Aspekt dieses Systemaufbaus besteht darin, dass die gesamte Funktionalität für die Fehlerbehebung sowohl für permanente als auch für transiente Fehler in der Hardware und in POWER Hypervisor enthalten ist. Die Fehlerbehebung soll für Anwendungen transparent sein und nicht von den Betriebssystemen abhängen, die in Partitionen implementiert sind.*

Wenn ein vom Prozessor erkannter Fehler nicht durch PIR (Processor Instruction Retry) oder APR (Alternate Processor Recovery) behoben werden kann, beendet POWER Hypervisor die Partition, auf die der Prozessorkern zum Zeitpunkt des Fehlers zugegriffen hat (Checkstop-Fehler). Dies wird als ***Core-Contained-Checkstop-Fehler*** bezeichnet. Im Allgemeinen wird dadurch der Ausfall auf eine einzelne Partition begrenzt. Wenn jedoch ein Fehler nicht auf eine einzelne Anweisung begrenzt werden kann oder wenn ein Fehler in einer POWER Hypervisor-Anweisung aufgetreten ist, wird der Server neu gestartet.

Fehlerbehebungsverfahren für den Kern



PIR (Processor Instruction Retry)

POWER7-Kerne unterstützen PIR (Processor Instruction Retry), eine Methode zur Behebung von Fehlern des Prozessorkerns. Da ein POWER7-Kern Architekturstatusinformationen erfasst, kann er neu starten und viele transiente Fehler automatisch beheben. Dieses Verfahren leistet Folgendes:

- Schutz vor sporadisch auftretenden Fehlern/Soft Errors
- Behebung von Cache- und Logikfehlern
- *Unabhängigkeit vom Betriebssystem und vom Hypervisor*

POWER7-Kerne enthalten verschiedene weitere Wiederherstellungsfunktionen, zum Beispiel die Erkennung/Behebung von Blockierungen. Eine Blockierung eines Kerns ist selten und wird in der Regel durch einen Soft Error im Prozessor verursacht, der die Anweisungsverarbeitung stoppt. Dieser Zustand wird möglicherweise nicht unmittelbar als „Standardlogikfehler" oder als Datenfehler erkannt. Hardware mit POWER7-Kernen kann diese Bedingung automatisch erkennen und beheben.

APR (Alternate Processor Recovery

Bei permanenten Fehlern versucht POWER Hypervisor, den Instruktionsstrom auf einen Ersatzkern zu verlagern. Diese Verfahren (PIR und APR) funktionieren sowohl für dedizierte Kerne als auch für gemeinsame Pools von Kernen.

Wenn die Partitionen ordnungsgemäß konfiguriert sind, findet die Fehlerbehebung statt, ohne Ausfallzeiten zu verursachen.

In einer virtualisierten Umgebung kann ein Systemadministrator mithilfe einer Partition Availability Priority-Einstufung eine Richtlinie festlegen, mit der ein Ersatzkern ermittelt werden kann, und dabei Folgendes durchführen:

- Leistungszuteilungen zu ausgewählten Kernen reduzieren, um die Kapazität eines Kerns wiederherzustellen
- Falls keine andere Möglichkeit besteht, eine vom Benutzer vordefinierte Partition mit der niedrigsten Priorität beenden, um eine Partition mit einer höheren Priorität weiterhin ausführen zu können

Predictive Processor Deallocation

Vor der Einführung der zukunftsweisenden Fehlerbehebungsfunktionen für den Kern in POWER6, war Predictive Processor Deallocation die wichtigste Möglichkeit, Fehler des Prozessorkerns zu behandeln. Dieses Feature betrachtet Muster behebbarer Fehler in der Prozessorfunktion als Grundlage für die Vorhersage eines zukünftigen nicht behebbaren Prozessorfehlers.

Dynamic Processor Deallocation (dynamische Aufhebung der Prozessorzuordnung) unterstützt die automatische Aufhebung der Konfiguration eines fehlerträchtigen Prozessorkerns, bevor dieser einen nicht behebbaren Systemfehler verursacht (ungeplanter Serverausfall). Dynamic Processor Deallocation basiert auf der Fähigkeit des Serviceprozessors, von FFDC generierte Informationen zu behebbaren Fehlern zu verwenden und POWER Hypervisor darüber zu benachrichtigen,

wenn der Prozessorkern seine vordefinierte Fehlergrenze erreicht. Anschließend „leert“ POWER Hypervisor zusammen mit dem Betriebssystem die Ausführungswarteschlange für den problematischen Kern, verteilt die Arbeit auf die verbleibenden CPUs um, gibt die problematische CPU frei und setzt den normalen Betrieb fort – allerdings möglicherweise mit geringerer Systemleistung.

Dieses Produktmerkmal kann mit Dynamic Processor Sparing gekoppelt werden, damit verfügbare zusätzliche Verarbeitungskapazität dynamisch ersetzt wird, bevor der ausgefallene Kern freigegeben wird.

Obwohl dieses Verfahren in POWER7-Servern weiterhin eingesetzt wird, weist es im Vergleich zu PIR und APR einige wesentliche Nachteile auf:

- Wenn Prozessoren Partitionen zugeordnet sind, ist es bei der Aufhebung der Zuordnung erforderlich, dass das Betriebssystem und/oder die Anwendungen zusammenarbeiten, damit der aktuell auf dem Prozessor ausgeführte Thread beendet wird, bevor die Aufhebung der Zuordnung stattfindet.
- Dieser Mechanismus behebt vorhersehbare, behebbare Fehler. Unvorhersehbare, sporadisch auftretende Fehler oder permanente Fehler können nicht behandelt werden.

Dynamic Processor Sparing

Die auf PowerVM basierende Power-Virtualisierungsumgebung von IBM ist sicher und kann eine ungenügende Auslastung von Servern dadurch verhindern. Dazu werden Ressourcen in Pools zusammengefasst und ihre Nutzung wird in mehreren Anwendungsumgebungen und Betriebssystemen übergreifend optimiert. Über intelligente Funktionalität für die dynamische logische Partitionierung (LPAR) kann eine einzelne Partition als vollständig separate AIX-, IBM i- oder Linux™-Betriebsumgebung fungieren. Partitionen können über dedizierte oder gemeinsam genutzte Ressourcen verfügen. Bei gemeinsam genutzten Ressourcen kann PowerVM in Pools gestellte Prozessorressourcen automatisch betriebssystemübergreifend anpassen. Dabei wird Rechenleistung von inaktiven Partitionen ausgeliehen, um hohe Transaktionsvolumen in anderen Partitionen zu verarbeiten.

Dieses sehr leistungsfähige Konzept für die Partitionierung maximiert die Partitionierungsflexibilität und -wartung. Diese PowerVM-Funktionalität stellt allen Server in der Power Systems-Familie neben der Unterstützung für genau abgestimmte Ressourcenzuordnung die Möglichkeit zur Verfügung, beliebige Ressourcen (Prozessorkerne, Speichersegmente, E/A-Steckplätze) jeder beliebigen Partition in jeder beliebigen Kombination zuzuordnen oder deren Zuordnung aufzuheben – oder Ressourcen mit mehreren virtuellen Maschinen gemeinsam dynamisch zu nutzen.

In einer logisch partitionierten Architektur steht der gesamte Serverhauptspeicher allen Prozessorkernen und allen E/A-Einheiten im System physisch zum Zugriff zur Verfügung, unabhängig von der physischen Position des Hauptspeichers oder davon, wo die logische Partition betrieben wird. POWER Hypervisor ist so konzipiert, dass beliebiger Code, der in einer Partition (Betriebssystem und Firmware) aktiv ist, nur Zugriff auf den physischen Hauptspeicher hat, der der dynamischen logischen Partition zugeordnet ist. Power Systems-Modelle verfügen darüber hinaus über von IBM entworfene PCI-zu-PCI-Bridges, damit POWER Hypervisor den Zugriff über DMA (Direct Memory Access), der von E/A-Einheiten ausgeht, auf Hauptspeicher begrenzen kann, dessen Eigner die die Einheit verwendende Partition ist. Der Aufbau mit einer Speichercachekoheränzdomäne ist eine wichtige Anforderung für die Bereitstellung der höchsten SMP-Leistung. Da IBM die Strategie verfolgt, Hunderte dynamisch konfigurierbarer logischer Partitionen bereitzustellen, die eine verbesserte Systemauslastung ermöglichen und die Gesamtkosten für die Datenverarbeitung reduzieren, müssen diese Server auf bestimmte Weise ausgelegt sein: Bedingungen müssen verhindert oder minimiert werden, die zu einem vollständigen Serverausfall führen.

Die Verfügbarkeitsarchitektur von IBM stellt einen hohen Schutz für die einzelnen Komponenten bereit, aus denen die Speicherkohärenzdomäne besteht und zu denen Speicher, Caches und Fabric-Bus gehören. Sie bietet darüber hinaus innovative Verfahren, mit denen Fehler innerhalb der Kohärenzdomäne auf einen Teil des Servers beschränkt werden können. Durch einen sorgfältigen Systemaufbau werden Fehler trotz der gemeinsamen Nutzung der Hardware in vielen Fällen auf eine Komponente oder auf eine Partition begrenzt. Viele dieser Verfahren werden im vorliegenden Dokument beschrieben.

Diese Architektur ermöglicht nicht nur eine außerordentliche Konfigurationsflexibilität, sondern weist auch Merkmale auf, die in anderen, weniger flexiblen Architekturen nicht vorhanden sind. Da Prozessoren und Prozessorzyklen zwischen logischen Partitionen dynamisch neu zugeordnet werden können, gehört zum ***Processor Sparing*** (Prozessorersatz) nicht nur die Neuzuordnung vollständiger Prozessorkerne, sondern auch die Neuzuordnung von Zyklen (Leistung) zwischen Partitionen. Dadurch werden Ausfälle verhindert, wenn zusätzliche Kerne (Ersatzkerne) nicht verfügbar sind.

Naturgemäß erfordert Alternate Processor Recovery Ersatzressourcen (die einem Kern entsprechen), wenn die Auslastung aus einem ausgefallenen Prozessor verlagert werden muss. Das Verfahren Predictive Processor Deallocation schließt auch die Möglichkeit ein, einen Ersatzkern zu verwenden, damit die Partitionen ohne Leistungsverluste weiterarbeiten können.

Aufgrund der Virtualisierung betrachtet das System einen „Ersatzkern“ in einem weit gefassten Sinn. Mit PowerVM verfügt ein Power-Server über eine beträchtliche Flexibilität beim Zusammenstellen freier Ressourcen. Selbst Bruchteile der Verarbeitungskapazität können verwendet werden, um den „Ersatz“ einzurichten.

Ersatzkapazität wird nach Priorität in der folgenden Reihenfolge gesucht:

- *Ein unlizenzierter CoD-Kern* (CoD – Capacity on Demand). Der am deutlichsten erkennbare Ersatzkern in einem System ist ein noch unlizenzierter Kern im System, also ein Kern, der für den zukünftigen Gebrauch in Reserve gehalten wird. Dieser unlizenzierte Kern kann verwendet werden, um den freigegebenen, fehlerhaften Prozessorkern automatisch zu ersetzen. In den meisten Fällen ist diese Operation für den Systemadministrator und für Endbenutzer transparent und der Server setzt den normalen Betrieb mit voller Funktionalität und voller Leistung fort. CoD-Zusatzressourcen können ohne zusätzliche Lizenzen eingesetzt werden und beeinflussen eine spätere Lizenzierung des Kerns nicht, wenn das System repariert wird.

  Dieser Ansatz weist einige wichtige Vorteile auf:

  1. Der Ersatz kann von einer beliebigen Position im System stammen. Er muss sich nicht im selben physischen Einschub oder in derselben Partition wie der ausgefallene Kern befinden.
  2. Eine einzelne Ersatzressource reicht für eine Aufhebung der Zuordnung an beliebiger Position innerhalb des gesamten Servers aus.
- *Falls kein CoD-Prozessorkern verfügbar ist*, versucht POWER Hypervisor, im System irgendwo einen nicht zugeordneten lizenzierten Kern zu erkennen. Dieser lizenzierte Kern ist der Ersatz und die ausgefallene CPU wird freigegeben.
- *Falls kein Ersatzprozessorkern verfügbar ist*, versucht POWER Hypervisor, die Kapazität eines vollständigen Prozessorkerns aus einem gemeinsamen Prozessorpool (sofern unterstützt) zu erkennen, und verteilt die Auslastung um. (Pools sind eine Einrichtung, die von der logischen Partitionierung unterstützt wird, da Ressourcen an beliebiger Stelle gemeinsam genutzt werden können.) In diesem Fall reduziert POWER Hypervisor die Zuordnung von Verarbeitungsressourcen, die Partitionen zugeteilt werden, um den Verlust eines Prozessors auszugleichen. Dies kann mit der Granularität eines Bruchteils eines Prozessors erfolgen.
- *Falls die erforderliche Ersatzkapazität nicht verfügbar ist*, zum Beispiel wenn die Partitionen zusammen mehr Verarbeitungsressourcen benötigen, als durch die Aufhebung der Zuordnung eines Prozessors verfügbar wird, oder wenn kein gemeinsamer Prozessorpool verwendet wird, wird Folgendes durchgeführt:

  1. Predictive Processor Deallocation:
     – Gemeinsame Poolpartitionen: Ein Kern wird erst beim nächsten Serverneustart freigegeben.
     – Dedizierte Prozessorpartitionen: Die Aufhebung einer Kernzuordnung wird durchgeführt (mehrere Kerne in der Partition), jedoch kein Ersatz. In einer Einzelprozessorpartition wird ein Kern erst nach einem Serverneustart freigegeben.
  2. Alternate Processor Recovery: Mindestens eine Partition wird beendet – jedoch nicht notwendigerweise diejenige, die den unzuverlässigen Kern verwendet. Kerne werden von POWER Hypervisor unter Verwendung von Partition Availability Priority ausgewählt.

Partition Availability Priority

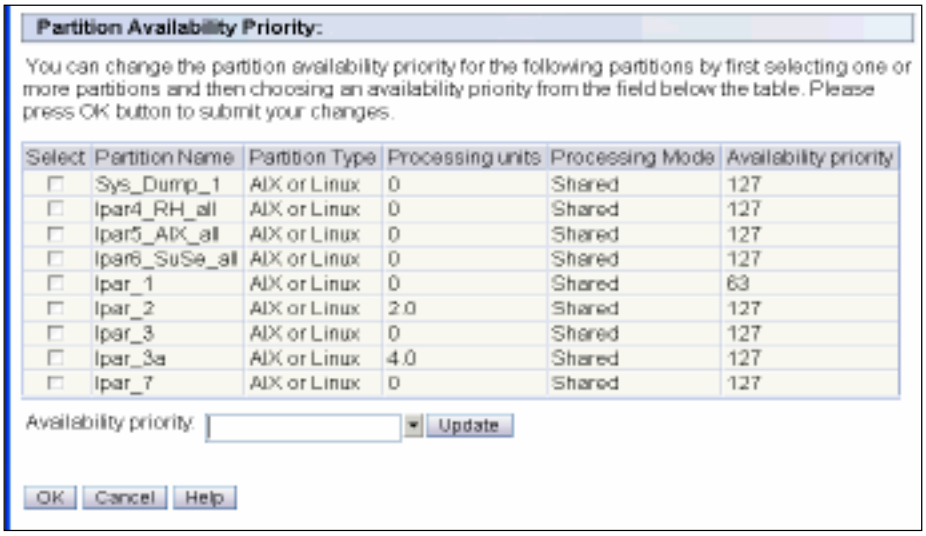
Power Systems geben Benutzern die Möglichkeit, Partitionen vorab zu priorisieren. Wenn Prioritäten für die Verfügbarkeit von Partitionen angegeben werden, kann POWER Hypervisor eine Partition (falls erforderlich auch mehrere Partitionen) mit geringerer Priorität beenden und die nun verfügbare Ressource nutzen, um Partitionen mit höherer Priorität aktiv zu halten.

Wenn Partitionen priorisiert werden, steuern die Prioritäten selbst in den Fällen, in denen keine Partition beendet werden muss, die Reihenfolge, in der für Teilressourcen des Prozessors die Zuordnung aufgehoben wird.

Weitere Designmerkmale der Wiederherstellung von POWER7-Prozessorkernen

Level-1-Caches (Instruktions- und Datencachebehandlung)

Instruktions- und Datencaches, die normalerweise als Teil eines Prozessorkerns betrachtet werden, sind so ausgelegt, dass erkannte Fehler, wenn nötig, korrigiert werden, indem fehlerhafte Daten oder Anweisungen von einer anderen Position in der Speicherhierarchie abgerufen werden. Die Caches sind in mehrere Datasets strukturiert. Falls in einem Dataset ein permanenter Fehler erkannt wird, kann der Kern die Verwendung dieser Daten beenden, wodurch das Problem effektiv gelöst ist. Mehrere set delete-Ereignisse können ebenfalls eine Predictive Processor Deallocation auslösen, bei der der fehlerhafte Cache (und der Kern) entfernt wird, bevor ein katastrophaler Fehler auftritt.



Partition Availability Priority

Wenn mit Processor Instruction Retry (PIR) ein Prozessorkernfehler nicht behoben werden kann, greift POWER Hypervisor auf das Verfahren Alternate Processor Recovery (APR) zurück, bei dem Ersatzkapazität (CoD oder nicht zugeordnete Prozessorkernressourcen) verwendet wird, um Workloads dynamisch zu verlagern. Dieses Verfahren kann auf einem POWER6-basierten Server die unterbrechungsfreie Anwendungsverfügbarkeit aufrechterhalten.

Falls kein Ersatzkern verfügbar ist, können Administratoren die Auswirkungen von Alternate Processor Recovery verwalten, indem sie Partition Availability Priority (Priorität der Partitionsverfügbarkeit) einrichten. Dieses Verfahren wird über HMC-Konfigurationsanzeigen eingerichtet und stellt eine numerische Rangfolge (0 bis 255) für die einzelnen Partitionen dar.

Anhand dieser Einstufung verteilt POWER Hypervisor Kapazitäten aus Partitionen mit niedrigerer Priorität um (und reduziert deren normale Leistung) oder stoppt, falls erforderlich, Partitionen mit niedrigerer Priorität, sodass Anwendungen mit hoher Priorität weiterhin normal betrieben werden können.

Level-2- und Level-3-Caches

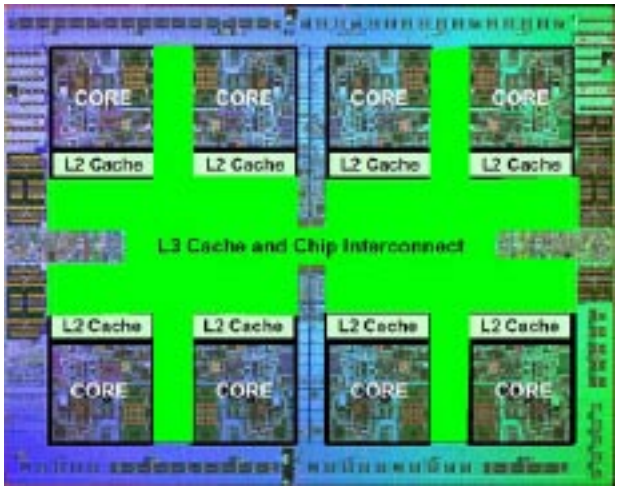
Die Level-2- und Level-3-Caches im POWER7-Prozessor verwenden mindestens DED/SEC-ECC-Codes (DED/SEC - Double Error Detect/ Single Error Correct), um sicherzustellen, dass ein Doppelbitfehler in einem beliebigen ECC-Datenwort erkannt werden und ein Einzelbitfehler in einem beliebigen ECC-Wort behoben werden kann. Dieser Mechanismus bietet in Kombination mit dem Layout der verschiedenen Datenbits innerhalb des ECC-Worts den wesentlichen Schutz vor einem Soft Error in einem der beiden Caches.

Wenn Daten in einem Cache gelesen werden, werden Fehler vom Cache-Controller überwacht. Falls eine Gruppe von Cachedaten – eine Cachezeile – einen permanenten, behebbaren Fehler aufzuweisen scheint, gibt der Cache-Controller auf der Basis einer Vorhersage die Cachezeile über einen Vorgang frei, der als *Cache Line Delete* (Cachezeilenlöschung) bezeichnet wird.

Bei diesem Prozess wird die Cachezeile im Cachezeilenverzeichnis als nicht verfügbar gekennzeichnet. Für den Fall, dass Daten aus dem Cache wegen einer Änderung im Cache in den Hauptspeicher geschrieben werden müssen, werden vor der Zeilenlöschung die Daten per ECC korrigiert und in den Hauptspeicher geschrieben.

Wenn trotz dieser Vorsichtsmaßnahmen in Daten aus einer Cachezeile ein Fehler gefunden wird, der nicht per ECC behoben werden kann, wird die Cachezeile dennoch gelöscht. Falls die Daten im Cache nicht geändert wurden, werden sie „wiederhergestellt“, indem eine weitere Kopie aus dem Hauptspeicher abgerufen wird. Wenn jedoch die Daten in der Cachezeile im Cache geändert wurden, werden die Daten in den Hauptspeicher geschrieben. Alle Fehler, die per ECC behoben werden können, werden beim Schreiben behoben. Alle Fehler, die nicht behoben werden können, werden beim Speichern identifiziert, indem ein besonderer Code in die ECC-Prüfbits der nicht korrigierbaren Wörter geschrieben wird.

Die Firmware beendet keinerlei Prozesse oder Partitionen allein aufgrund eines nicht behebbaren Fehlers. Stattdessen wird die Verwendung der markierten fehlerhaften Daten im System durch einen Prozess überwacht, der als S*pecial Uncorrectable Error-Behandlung* (SUE) bezeichnet wird. Wenn die Daten nie verwendet werden, wird die Datenverarbeitung ohne Unterbrechung fortgesetzt. Wenn die Daten jedoch verwendet werden, können die Folgen in vielen Fällen auf die Beendigung nur einer Anwendung oder einer Partition begrenzt werden.

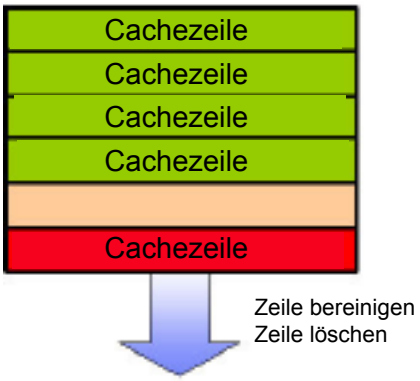


Behandlung von L2- und L3-Cachedaten

Der POWER7-Prozessorchip besteht aus acht Prozessorkernen, die jeweils einen 32-KB-8-Wege-Assoziativdatencache, einen 32-KB-2-Wege-Assoziativinstruktionscache und einen eng gekoppelten 256–KB-8-Wege-L2-Assoziativcache umfassen.

Darüber hinaus verwendet der Chip eingebettete dynamische Arbeitsspeichertechnologie (eDRAM), um 32 MB internen L3-Cache zur Verfügung zu stellen. Dieser Onboard-L3-Cache ist sehr wichtig, um einen Systemaufbau mit ausgeglichener Auslastung bereitzustellen.

Die Level-2- und Level-3-Caches im POWER7-Prozessor verwenden mindestens DED/SEC-ECC-Codes (DED/SEC – Double Error Detect/Single Error Correct), um sicherzustellen, dass ein Doppelbitfehler in einem beliebigen ECC-Datenwort erkannt werden und ein Einzelbitfehler in einem beliebigen ECC-Wort behoben werden kann. Dieser Mechanismus bietet in Kombination mit dem Layout der verschiedenen Datenbits innerhalb des ECC-Worts den wesentlichen Schutz vor einem Soft Error in beiden Caches. Zum Schutz der Daten werden verschiedene weitere Mechanismen eingesetzt.



Einzelbit-Soft-Errors

- Mit ECC behoben

Permanente Einzelbitfehler

- Cachezeile bereinigt
  - – Im Cache nicht geänderte Daten werden ungültig gemacht.
  - – Im Cache geänderte Daten werden in den Hauptspeicher geschrieben.
    - √ Von Hardware als nicht behoben markiert (SUE-Markierung)
- Cachezeile bereinigt
  - – Cachezeile nicht mehr verwendet; keine künftigen Fehler

Mehrbitfehler

- Cachezeile dynamisch gelöscht
  - – Im Cache nicht geänderte Daten werden ungültig gemacht.
  - – Im Cache geänderte Daten werden in den Hauptspeicher geschrieben.
    - √ Alle nicht korrigierbaren Daten werden mit einem SUE-Code markiert und behandelt, wenn die Daten verwendet werden.
- Cachezeile dynamisch gelöscht
  - – Cachezeile nicht mehr verwendet; keine künftigen Fehler

Ein nicht behebbarer Datenfehler (UE) kann schwerwiegende Folgen haben. Wenn ein UE auftritt, gibt die Systemfirmware die Anforderung aus, dass das Teil, das den Cache enthält, ausgetauscht werden muss. Wenn der Fehler jedoch auf eine einzelne Cachezeile eingegrenzt wird, reicht die Löschfunktion für Cachezeilen aus, um künftige Probleme im Zusammenhang mit diesem Fehler zu verhindern.

Die L2- und L3-Caches werden von viel kleineren Arrays gesteuert, die als Verzeichnisse bezeichnet werden. Diese enthalten Informationen zum Inhalt bestimmter Cachezeilen. Diese Verzeichnisse werden von einem DED/SEC-ECC-Code abgedeckt.

Special Uncorrectable Error Handling (spezielle Behandlung nicht behebbarer Fehler)

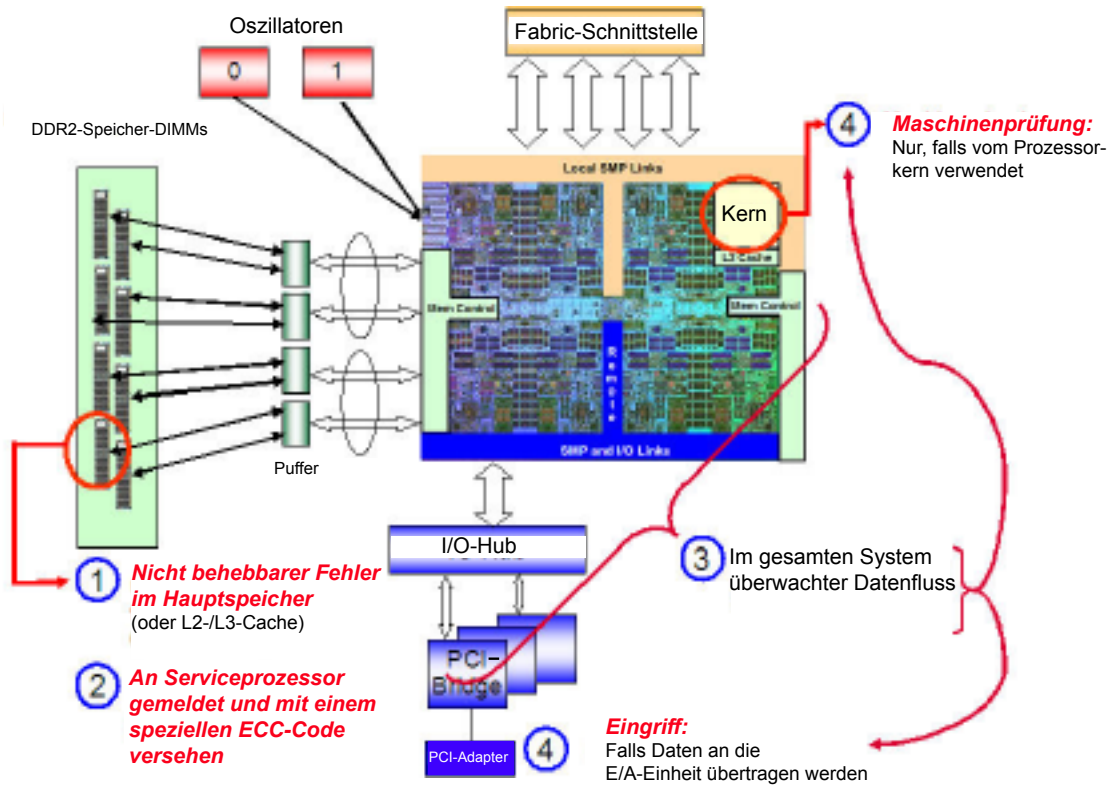
Obwohl es selten vorkommt, dass ein nicht behebbarer Datenfehler im Hauptspeicher oder in einem Cache auftritt, kann dies trotz aller Vorsichtsmaßnahmen geschehen. Das Ziel aller Power Systems ist es, die Auswirkungen eines nicht behebbaren Fehlers auf so wenige Betriebsunterbrechungen wie möglich zu begrenzen, indem eine klar strukturierte Strategie angewendet wird, die bereits mit der Berücksichtigung der Datenquelle beginnt.

Manchmal ist ein nicht behebbarer Fehler seinem Wesen nach transient und tritt in Daten auf, die aus einem anderen Repository wiederhergestellt werden können. Beispiel: In einem L3-Cache ist möglicherweise eine nicht geänderte Kopie von Daten gespeichert, die sich in einem Teil des Hauptspeichers befinden. In diesem Fall löst ein nicht behebbarer Fehler im L3-Cache einfach das „erneute Laden“ einer Cachezeile aus dem Hauptspeicher aus. Diese Möglichkeit ist auch im L2-Cache verfügbar.

Für Fälle, in denen die Daten nicht aus einer anderen Quelle wiederhergestellt werden können, wird ein Verfahren eingesetzt, das als SUE-Behandlung (SUE - Special Uncorrectable Error) bezeichnet wird. Ein SUE kann auf mehrere Arten erkannt werden:

1. ECC-Wörter, die nicht behebbare Fehler in L2- und L3-Caches enthalten, werden schließlich mit einem SUE-Code (SUE – Special Uncorrectable Error) in den Hauptspeicher geschrieben. Dieser Code gibt an, dass die Daten unzuverlässig sind.
2. Fehler in anderen Teilen des Systems, zum Beispiel im Hauptspeicher, können ebenfalls dazu führen, dass Datenwörter mit dem SUE-Code markiert werden.

Dieser Code gibt an, dass ein nicht behebbarer Fehler in der Datenquelle gefunden wurde und dass die Daten im „Standard-ECC-Wort“ nicht mehr gültig sind. Außerdem sendet die Prüfhardware ein Signal an den Serviceprozessor und gibt die Quelle des Fehlers an. Der Serviceprozessor leitet Aktionen ein, um die Komponente (bzw. die Komponenten) zu identifizieren, die den Fehler enthalten, sodass sie für die Fehlerbehebung markiert sind. Der Serviceprozessor führt ebenfalls geeignete Maßnahmen durch, um den Fehler zu behandeln.



Obwohl es selten vorkommt, dass ein nicht behebbarer Datenfehler im Hauptspeicher oder in einem Cache auftritt, kann dies trotz aller Vorsichtsmaßnahmen geschehen. Für Fälle, in denen die Daten nicht aus einer anderen Quelle wiederhergestellt werden können, wird ein Verfahren eingesetzt, das als SUE-Behandlung (SUE - Special Uncorrectable Error) bezeichnet wird. Dieser Code gibt an, dass ein nicht behebbarer Fehler in der Datenquelle gefunden wurde und dass die Daten im „Standard-ECC-Wort“ nicht mehr gültig sind.

Wenn ein Fehler lediglich erkannt wird, führt dies nicht automatisch zum Systemabschluss oder zur Beendigung der betreffenden Partition. Stattdessen wird die Verwendung der Daten im gesamten System überwacht. Bisweilen kann es vorkommen, dass die Daten nie verwendet, jedoch einfach gelöscht werden, bevor sie in einen Prozessorkern geladen oder an eine E/A-Einheit gesendet werden. In diesem Fall wird die Fehlerbedingung ungültig gemacht und das System setzt den normalen Betrieb fort.

Wenn ein Prozess versucht, die Daten zu verwenden, wird das Betriebssystem (AIX oder Linux) über den Fehler informiert und beendet nur das betroffene Benutzerprogramm. Falls die Daten über DMA an einen E/A-Controller gesendet werden, „friert“ das System den GX-Adapter ein. Dies führt zum Verlust aller E/A-Operationen über diesen Adapter. Ein ordnungsgemäß konfigurierter Server (mit E/A-Redundanz) kann einen Ausfall verhindern. Nur wenn beschädigte Daten von POWER Hypervisor verwendet werden und kein gespiegelter Speicher für POWER Hypervisor (POWER Hypervisor Mirrored Memory) eingesetzt wird, muss das gesamte System neu gestartet werden, um die Systemintegrität aufrechtzuerhalten.

Eine Architektur zum Eingrenzen eines nicht behebbaren Datenfehlers und zur Minimierung der Folgen des Fehlers für das System muss Folgendes leisten:

- Sie muss für nicht behebbare Fehler im Hauptspeicher und im Cache vorhanden sein.
- Sie darf sich nicht auf das System auswirken, wenn die Daten nicht verwendet werden.
- Wenn sie von einem Prozessor verwendet wird, hängt die Auswirkung davon ab, was der Prozessor ausführt:
  - Im Fall von Partitionsdaten:
    - √ Höchstens ein Fehler auf einer Partition
  - Im Fall von POWER Hypervisor:
    - √ Alle Partitionen beendet
    - √ Kein Fehler, falls Mirrored Hypervisor Memory eingesetzt wird
- Wenn die Daten eine E/A-Übertragung (DMA) sind, werden am GX-Controller angeschlossene E/A-Einheiten eingefroren.
  - Dies kann zum Verlust von Partitionen führen, die die Ein-/Ausgabe verwenden, falls die Ein-/Ausgabe kritisch ist und keine redundanten Ressourcen verfügbar sind.

Wenn ein Fehler lediglich erkannt wird, führt dies nicht automatisch zum Systemabschluss oder zur Beendigung der betreffenden Partition.

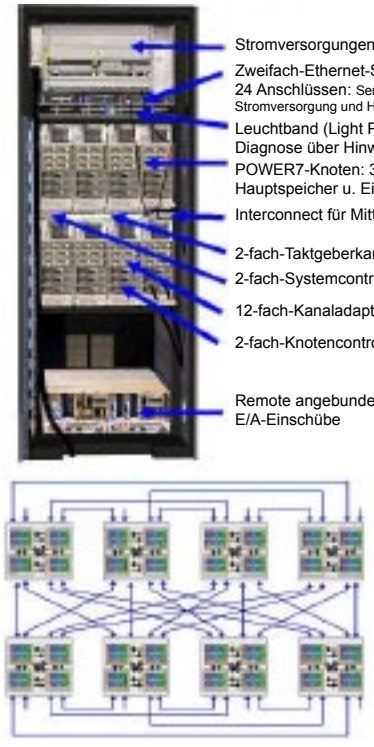
Stattdessen wird die Verwendung der Daten im gesamten System überwacht. Bisweilen kann es vorkommen, dass die Daten nie verwendet, jedoch einfach gelöscht werden, bevor sie in einen Prozessorkern geladen oder an eine E/A-Einheit gesendet werden. In diesem Fal kannl die Fehlerbedingung ohne Bedenken ungültig gemacht werden und das System setzt den normalen Betrieb fort.

Wenn der Prozessorkern versucht, die fehlerhaften Daten zu laden, bewirkt diese Anforderung, dass eine synchrone Maschinenprüfungsunterbrechung generiert wird. Die Firmware stellt einen Zeiger auf die Anweisung bereit, die auf die beschädigten Daten verweist. Diese Daten sind häufig nützlich, um das Problem abzumildern:

1. 1Wenn eine Benutzeranwendung Eigner der referenzierten Daten ist (Code des Benutzeradressbereichs) verfügt das Betriebssystem (derzeit unterstützte AIX- und Linux-Versionen) über die Möglichkeit, lediglich die Anwendung zu beenden, die Eigner der Daten ist. Wenn die Daten Kerneldaten des Betriebssystems sind, wird die Partition normalerweise beendet, das übrige System kann jedoch den Betrieb fortsetzen.
2. Nur falls POWER Hypervisor Eigner der beschädigten Daten ist oder die Daten sich in einem kritischen Bereich des POWER Hypervisor-Hauptspeichers befinden, wird das gesamte System beendet und automatisch neu gestartet. Dadurch bleibt die Gesamtintegrität des Systems erhalten.

Falls die Daten über DMA an einen E/A-Controller gesendet werden, „friert“ das System den GX-Adapter ein. Dies führt zum Verlust aller E/A-Operationen über diesen Adapter. Ein ordnungsgemäß konfigurierter Server (mit E/A-Redundanz) kann einen Ausfall verhindern.

Dieses Design zur Behandlung eines zufälligen UE ist so optimiert, dass die Verfügbarkeit von Anwendungen und Partitionen maximiert wird, da nichts beendet wird, ***bis*** und ***erst wenn*** es erforderlich ist, um zu verhindern, dass das System falsche Daten verwendet. Wenn eine Beendigung ausgeführt wird, *muss nur das Programmelement beendet werden, das Eigner der fehlerhaften Daten ist.*



Über lokale und remote angebundene SMP-Links werden bis zu 32 POWER7-Chips verbunden.

Power 795: A Mehrknotenserver mit Fabric-Interconnect

Der Server IBM Power 795 ist so konzipiert, dass er die Verarbeitung großer Transaktionsvolumen und Datenbankanwendungen innerhalb einer hochgradig virtualisierten Systeminfrastruktur unterstützt. Dies ermöglicht einen neuartigen Umfang des Lastausgleichs, der Ressourcenauslastung und der Effizienz. In seiner Eigenschaft als leistungsfähigstes Mitglied der IBM Power Systems-Familie bietet dieser Server eine sehr hohe Leistung, massive Skalierbarkeit und entsprechende Bandbreite, das volle Spektrum komplexer, unternehmenskritischer Anwendungen zu unterstützen.

Der Server Power 795 ist mit bis zu 256 POWER7-Prozessorkernen bestückt und kann schnell und nahtlos so skaliert werden, dass er sich dem wechselnden Bedarf des heutigen Geschäftslebens anpasst. Der Server Power 795 ist so ausgelegt, dass er zusammen mit dem Unternehmen wächst. Zu diesem Zweck ist er ausgezeichnet skalierbar und flexibel konfigurierbar. Prozessorblöcke, Hauptspeicher, E/A-Einschübe, Adapter und Plattenpositionen können mühelos hinzugefügt werden, um das Potenzial der Leistungsfähigkeit und der Kapazität des Systems zu nutzen.

Der große Umfang der von Großsystemen abgeleiteten Funktionen für Zuverlässigkeit, Verfügbarkeit und Wartungsfreundlichkeit (RAS) stellen im Power 795 sicher, dass unternehmenskritische Anwendungen zuverlässig und rund um die Uhr laufen. Alle POWER7-Prozessormodule sind miteinander über Fabric-Busse zwischen Prozessoren verbunden. Mehrere Fabric-Bus-Schnittstellen in den einzelnen Prozessormodulen ermöglichen Punkt-zu-Punkt-Verbindungen zwischen Modulen innerhalb eines Knotens und Punkt-zu-Punkt-Verbindungen zwischen Knoten – und dies alles ohne Umschalten und ohne Kreuzschienenmodule. In einem Mehrknotenserver kann diese Architektur innerhalb eines Systems von einem Prozessorblock dekonfiguriert oder neu konfiguriert werden, während die übrigen Prozessorblöcke weiterhin verbunden sind. Diese Operation kann durchgeführt werden, ohne Leistung oder Konnektivität zu verlieren und ohne den Datenverkehr auf den Fabric-Bussen zwischen den übrigen Knoten ausgleichen zu müssen.

CEC-Knoten und Prozessor-zu-Prozessor-Schnittstellen (Fabric-Busse)

Die Server Power System 770, 780 und 795 verwenden in der zentralen Elektronik (CEC - Central Electronics Complex) eine Hardwarearchitektur aus „Bausteinen“ oder „Knoten“. In diesen Servern kann ein System mindestens einen Knoten einschließen, der jeweils mindestens einen POWER7-Chip, ein Speicher-DIMM, einen E/A-Hub-Chip und zugeordnete Komponenten enthält.

In Power 770 oder 780 ist ein Knoten ein CEC-Einschub. Bis zu vier CEC-Einschübe sind miteinander durch Fabric-Bus-Kabel zu einem hoch skalierbaren System verbunden.

Power 795 enthält bis zu acht Prozessor/Speicher-Blöcke. Jeder Block wird in ein CEC-Rack eingeschoben (wie Bücher in ein Bücherregal) und über eine passive (also keine aktiven elektronischen Komponenten enthaltende) Platine namens „Mittelplatine“ elektrisch angeschlossen.

Alle POWER7-Prozessormodule sind miteinander über Fabric-Busse zwischen Prozessoren verbunden. Mehrere Fabric-Bus-Schnittstellen in den einzelnen Prozessormodulen ermöglichen Punkt-zu-Punkt-Verbindungen zwischen Modulen innerhalb eines Knotens und Punkt-zu-Punkt-Verbindungen zwischen Knoten – und dies alles ohne Umschalten und ohne Kreuzschienenmodule.

In einem Mehrknotenserver kann diese Architektur innerhalb eines Systems von einem Prozessorblock dekonfiguriert oder neu konfiguriert werden, während die übrigen Prozessorblöcke weiterhin verbunden sind. Diese Operation kann durchgeführt werden, ohne Leistung oder Konnektivität zu verlieren und ohne den Datenverkehr auf den Fabric-Bussen zwischen den übrigen Knoten ausgleichen zu müssen.

Auf diese Weise unterstützt die als Knoten strukturierte Hardware ***das dynamische Entfernen eines Knotens*** innerhalb eines Systems (nachdem die Fabric-Aktivität ausgesetzt worden ist). Außerdem unterstützt sie die ***dynamische Integration eines reparierten Knotens*** oder ***das Hinzufügen eines neuen Knotens*** während eines Upgradevorgangs[6].

Zu den Architekturmerkmalen des Fabrics (Konnektivität, Leistung) kommen die guten Eigenschaften dieses Aufbaus der Übertragung hinzu. Der Fabric-Bus ist hinsichtlich der Signalcharakteristik sehr stabil: Es gibt keine erwartete Soft-Error-Rate und keine Fehlerrate für sporadisch auftretende Fehler des Busses aufgrund elektrischer Störungen oder Taktproblemen. Dies gilt auch für ein vollständig konfiguriertes System. Teilweise liegt dies daran, dass IBM Rechtsinhaber des Aufbaus aller Komponenten ist, die die Fabric-Bus-Signale übertragen. IBM Entwickler führen den Entwurf, die Optimierung und die Anpassung der Komponenten für die einzelnen Busse und Systemimplementierungen durch.

Ein im Grundkonzept angelegtes Ziel des Fabric-Bus-RAS-Aufbaus besteht darin, einen Fehler einer einzelnen Datenbitzeile zu verkraften, ohne einen nicht behebbaren Fehler oder eine Verschlechterung der Signalstärke zu verursachen, die ihrerseits zu Leistungsproblemen führen können. Dies wird dadurch erreicht, dass auf dem Fabric-Bus ein DED/SEC-ECC-Code verwendet wird, der sowohl transiente als auch permanente Fehler in einer Datenbitzeile behebt.

Speichersubsystem

Das Speichersubsystem eines POWER7-Systems enthält drei wesentliche Teile: DRAM-Chips, in denen die Daten gespeichert werden und die den „Hauptspeicher“ des Systems umfassen, Speichercontroller (zwei pro Prozessormodul), die die Prozessorschnittstelle zum Speicher verwalten, und Speicherpufferchips, die die Schnittstelle zwischen den Speichercontrollern und den DRAM-Chips bilden.

Es werden zwei grundlegende DIMM-Aufbaus (DIMM - Dual In-line Memory Module) verwendet:

1. Ein angepasster Aufbau, bei dem Hauptspeicherpuffer und die DRAM-Chips auf einem DIMM gemäß IBM Spezifikationen eingebettet sind
2. Ein Aufbau, bei dem das Speicher-DIMMs gemäß Industriestandard verwendet und eine separate Platine mit Pufferchips bestückt werden

Busschnittstellen des Speichers

Im Unterschied zu den Fabric-Bus-Schnittstellen zwischen den Prozessoren ist vom Bus zwischen Prozessorspeichercontrollern und den Pufferchips zu erwarten, dass aufgrund elektrischer Störungen, Abweichungen in der Taktung und verschiedener anderer Faktoren gelegentliche transiente Fehler auftreten.

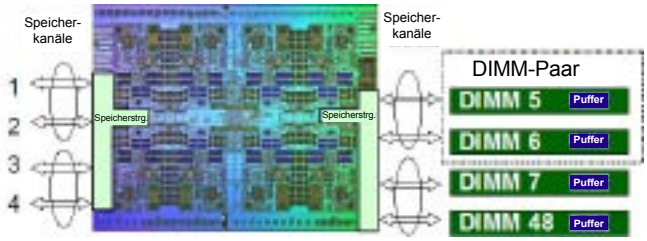
Daher unterscheidet sich der Speicherbusaufbau stark von der Fabric-Bus-Schnittstelle. In diesem Fall wird bei diesem Konzept ein CRC-Code (CRC - Cyclic Redundancy Check) zur Fehlererkennung verwendet. Diese Methode ist eine leistungsfähige Möglichkeit, die Typen mehrerer Bitfehler zu erkennen, die voraussichtlich wegen transienter Busfehler auftreten.

[6] Verfügbarkeitstermine finden Sie im IBM United States Hardware Announcement 110-158, vom 17. August 2010 und im IBM United States Hardware Announcement 110-025 vom 9. Februar 2010.

Mit solchen Codes können Fehler in der Regel zwar erkannt, aber nicht behoben werden. Auf dem Speicherbus können Daten dadurch korrigiert werden, dass der Speichercontroller ***eine fehlerhafte Operation wiederholt.***

Bei jedem Einschalten eines Systems durchläuft der Fabric-Bus des Speichers eine Reihe von Tests, mit denen die Datenübertragungsleistung optimiert werden soll. Wenn bei einem Bus mehrere CRC-Fehler auftreten, die durch eine Wiederholung behoben werden müssen, kann der Speichercontroller ***dynamisch angepasst werden***, um die optimale Busleistung wiederherzustellen.

Wenn ein persistenter Soft Error durch die dynamische Anpassung eines Busses nicht behoben werden kann, wurde der Fehler möglicherweise von einer fehlerhaften Bitzeile im Bus selbst verursacht. Der Speicherbus enthält eine Funktion für ***dynamische Ersatzbitzeile*** (Dynamic Memory Channel Repair). Mit dieser Funktion kann der Speichercontroller eine persistent fehlerhafte Bitzeile erkennen und durch eine Ersatzbitzeile ersetzen.



POWER7-Speicherlayout

Der POWER7-Chip enthält ein Paar von Speichercontrollern mit jeweils vier Anschlüssen. In einigen Power-Server-Modellen (Power 770, 780 und 795) werden alle acht verfügbaren Speicherkanäle und angepasste DIMMs sowie ein auf dem DIMM integrierter Pufferchip eingesetzt, um eine höhere Leistung des Speichers zu erzielen. In anderen Modellen befinden sich Standard-DIMMs und ein separat eingebauter Pufferchip, der für eine ausgeglichene Leistung bei Verwendung nur eines einzigen Speichercontrollers sorgt. In POWER7-Servern wird folgendes eingesetzt:

- Mit einem auf ein DIMM-Paar verteilten 72-Byte-ECC-Wort

Ein oder zwei Speichercontroller mit vier Anschlüssen. Die einzelnen Anschlüsse werden immer mit einem Pufferchip verbunden.

- Der Puffer befindet sich direkt auf dem DIMM oder auf einer Platine für DIMMs gemäß Industriestandard.

Speicherbusse mit CRC-Wiederholung („retry“) und dynamischer Anpassung des Busses

- Diese können eine dynamische Reparatur einer Zeile durchführen, falls eine Bitzeile fehlerhaft ist.
- Die Fehlerbehebung kann von Hardware eingeleitet oder von Firmware ausgelöst werden, falls zu viele behebbare Fehler verarbeitet werden.

Pufferchips, die von IBM entworfen und hergestellt werden. Ein Pufferchip muss Folgendes leisten:

- Er ist gegenüber permanenten Fehlern und Soft Errors sehr tolerant.
- Er setzt die UE-Fehlerbehandlung und den CRC-Wiederholung für Soft Errors ein.
- Die DRAM-Chipschnittstelle ist ECC-geschützt und enthält Funktionalität für Wiederholungen („retry“).

Hauptspeicherpuffer

Auf die RAS-Merkmale des Chips für den Hauptspeicherpuffer wurde großes Augenmerk gelegt. Dieser von IBM entworfene Chip weist dieselbe Art von Fehlererkennungs- und Fehlereingrenzungsverfahren auf wie im POWER7-Prozessormodul. In beiden Fällen war es ein Hauptziel, Soft Errors zu erkennen und zu beheben.

Nötigenfalls kann der Speicherpufferchip, während er auf die Beendigung einer vorherigen Fehlerbehebungsoperation wartet, temporär ausstehende Datenübertragungen zwischen ihm selbst und dem Speichercontroller „fernhalten“, indem er immer dann einen ***CRC-Fehler*** signalisiert, wenn Daten übertragen werden sollen. Dies bedeutet, dass ein POWER7-Speichersubsystem Soft Errors in den Pufferchips beheben kann, die bei einem anderen Speicheraufbau zu einem Absturz führen. Der Datenbus zwischen dem Hauptspeicherpuffer und den DRAM-Chips auf einem DIMM verwendet den ECC-Schutz: Ein Fehler führt zur Wiederholung von Operationen, mit denen transiente Busfehler behoben werden.

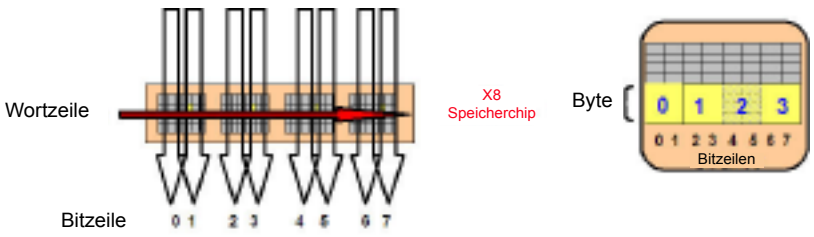
Speicher-DIMMs und ECC-Wörter

Der Hauptspeicher in einem POWER7-System besteht aus Speicher-DIMMs, die wiederum aus mehreren DRAM-Chips bestehen. In den meisten POWER7-Systemen befinden sich DRAM-Chips mit 8-Bit-Übertragung, d. h., dass ein einzelner DRAM-Chip 8 Bits Daten auf einmal liefert. Die einzelnen Bits werden aus einem DRAM-Chip als sogenannte ***Bitzeile*** ausgelesen.

Daten werden in ECC-Wörtern aus dem Speicher ausgelesen und in ihm gespeichert. Ein ECC-Wort besteht aus Daten und zugehörigen Prüfbits, mit denen bestimmt wird, ob die Daten korrekt gespeichert und gelesen wurden.

Bei POWER7 besteht ein ECC-Wort aus 72 Byte Daten. Davon werden 64 Byte zum Speichern von Nutzdaten verwendet. Die übrigen 8 Byte enthalten Prüfbits und zusätzliche Informationen zum ECC-Wort.

POWER7-Systeme greifen auf Speicher-DIMMs paarweise zu, denn die 72 Byte eines ECC-Worts sind gleichmäßig auf zwei unterschiedliche DIMMs verteilt. Also werden aus jedem DIMM 36 Byte eines ECC-Worts abgerufen. Die einzelnen DIMMs sind weiter unterteilt in mehrere Speicherbänke. Eine ***Speicherbank*** ist eine Einheit von Daten, die aus einigen oder aus allen Speicherchips auf einem einzelnen Speicher-DIMM gebildet wird. Dieser Aufbau wird vor allem eingesetzt, um die Leistung zu steigern, er weist jedoch auch bestimmte Vorteile für RAS auf; er vereinfacht zum Beispiel die Erkennung defekter Speicherpufferchips oder DIMMs in einem ECC-Wort.



DRAM-Chipstruktur

Dieses logische Diagramm veranschaulicht eine der Möglichkeiten, wie ein Speicher-DRAM-Chip strukturiert sein kann. Der DRAM-Chip enthält einen Array aus Tausenden von Speicherzellen. Der Zugriff auf diese Zellen erfolgt auf viele Arten. Wie zugegriffen wird, hängt von der Struktur des Chips ab. In diesem Beispiel handelt es sich um einen 8-Bit-Speicherchip. Auf Daten wird in acht Bitzeilen gleichzeitig zugegriffen. Aus mehreren DRAM-Chips werden Bitzeilen so zusammengestellt, dass sie Speicherwörter bilden.

Das POWER7-ECC-Schema verwendet jeweils vier Bits aus zwei Bitzeilen, um ein Byte aus acht Bits zu bilden. Die einzelnen DRAM-Chips liefern vier Bytes (oder vier Zeichen) für den ECC-Code. In einem POWER7-Hauptspeichersubsystem wird auf 18 DRAM-Chips gleichzeitig zugegriffen, um die 72 Byte zusammenzustellen, die für das ECC-Schema erforderlich sind. Diese Bytes werden aus zwei DIMMs (einem DIMM-Paar) abgerufen.

Das ECC-Design basiert darauf, dass der Speicherchip bei einem Ausfall eines ganzen Chips auf einem DIMM dauerhaft identifiziert und als defekt markiert werden kann. Anschließend kann der ECC-Algorithmus, dem nun der Defekt eines DRAM-Chips bekannt ist, die Datenverarbeitung fortsetzen. Zugleich wird sichergestellt, dass die übrigen Daten im ECC-Wort geprüft, gegebenenfalls korrigiert und ordnungsgemäß übertragen werden. Diese Ebene der Fehlerkorrektur wird als ***Chipkill-Fehlerkorrektur*** bezeichnet. Chipkill ist die Grundlage für Hauptspeicher-RAS, da damit sichergestellt wird, dass ein System den Betrieb auch dann ordnungsgemäß fortsetzen kann, wenn ein einzelner permanenter DRAM-Chipfehler auftritt.

Diese Ebene des DRAM-Schutzes wurde in POWER6-Systemen bereitgestellt. Während allerdings POWER6-Systeme ein komplettes Chipkill-Ereignis korrigieren konnten, war es in einigen Konfigurationen möglich, dass ein einzelner Fehler an einer anderen Stelle im ECC-Wort zu einem nicht behebbaren Speicherfehler führen konnte. Zum Schutz vor diesem Ereignis stellten POWER6-Systeme redundanten Speicher bereit, sodass eine fehlerhafte Bitzeile auf einem DRAM-Chip durch eine Ersatzbitzeile aus einem anderen DRAM-Chip ersetzt werden konnte (bis zur vom redundanten Speicher gesetzten Grenze).

Für POWER7-Server wurde für alle Konfigurationen der grundlegende ECC-Aufbau selbst so verbessert, dass der ECC-Algorithmus, nachdem ein DRAM-Chip als fehlerhaft markiert worden ist, einen Fehler auch dann beheben kann, wenn bei einem *weiteren Halbbyte (Bitzeile)* ein Fehler auftritt.

POWER7-DIMMs mit integrierten Pufferchips – „selbstheilender" Hauptspeicheraufbau

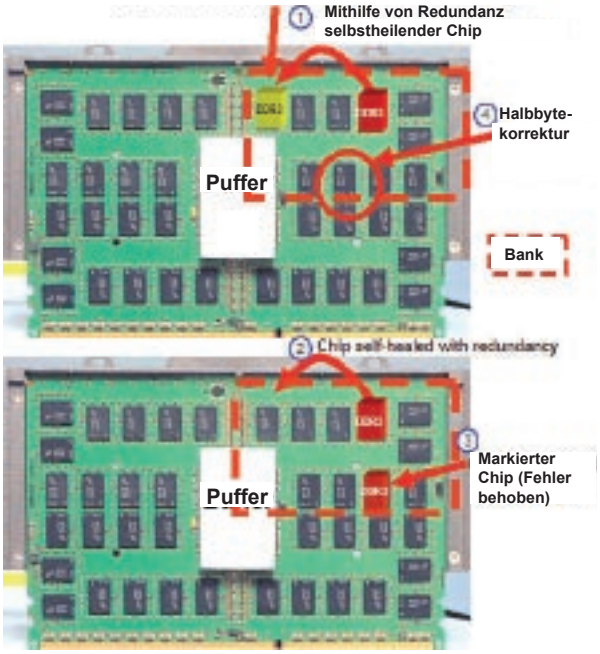
Für POWER7-Server wurde für alle Konfigurationen der grundlegende ECC-Aufbau selbst so verbessert, dass der ECC-Algorithmus, nachdem ein DRAM-Chip als fehlerhaft markiert worden ist, einen Fehler auch dann beheben kann, wenn bei einem weiteren Halbbyte (Bitzeile) ein Fehler auftritt.

Aufgrund dieser Fehlerkorrekturfunktionalität kann ein DIMM-Chip auch mit einem fehlerhaften DRAM-Chip in einem System verbleiben, denn ein zufälliger Fehler in einem weiteren Halbbyte kann dennoch behoben werden.

Dieses Foto zeigt Entwicklungsmodelle der DIMM-Typen, die in den Servern Power 770, 780 und 795 eingesetzt werden. Die Anordnung der Bänke kann von der hier abgebildeten Anordnung abweichen. In diesen Modellen werden angepasste DIMMs mit Ersatzkapazität eingesetzt.

- Diese umfassen pro DIMM einen zusätzlichen DIMM-Chip pro Bank, der einen defekten DRAM-Chip ersetzen kann.
- Nach dem Ersetzen weisen alle die Fehlererkennung und die Fehlerbehebung der Basiskonfiguration auf.
- Ein einzelnes DIMM kann aus mehreren Bänken bestehen, die jeweils über Ersatzkapazität verfügen. Die Anzahl der Bänke und der Umfang der Ersatzkapazität hängen von der Speicher-DIMM-Größe ab.

Zusätzlich: 1 DRAM-Chip pro Bank, der einen beliebigen ausgefallenen DRAM-Chip ersetzen soll



Aufgrund dieser Fehlerkorrekturfunktionalität kann ein DIMM-Chip auch mit einem fehlerhaften DRAM-Chip in einem System verbleiben, denn ein zufälliger Fehler in einem weiteren Halbbyte kann dennoch behoben werden.

Da jedoch sowohl geplante als auch ungeplante Ausfallzeiten in Systemen mit angepassten DIMMs (Power 770, 780 und 795) minimiert werden müssen, wurden die RAS-Funktionen von DIMMs über diesen Grundschutz hinaus erweitert: Es wurde weiterer Speicher hinzugefügt: ***Zusatzspeicherchips***. In diesen Servern besteht eine Speicherbank aus zehn DRAM-Chips. Die einzelnen DRAM-Chips stellen acht Datenbitzeilen bereit. Für ECC sind Bitzeilen in Paaren angeordnet und auf die Daten wird über ECC-Bytes aus den einzelnen Bitzeilenpaaren zugegriffen. Ein Byte besteht in diesem Fall aus vier Bits pro Bitzeile (einem Halbbyte) oder aus einem Byte pro Bitzeilenpaar. Auf diese Weise steuert jeder DRAM-Chip vier Bytes zum ECC-Wort bei. Neun DRAM-Chips liefern 36 Byte und das DIMM-Paar die vollen 72 Byte, die vom Fehlerkorrekturschema verwendet werden. Der „zusätzliche“ DRAM-Chip (der zehnte) stellt einen Zusatz-DRAM-Chip für jede Speicherbank auf den einzelnen DIMMs bereit. Wenn in einem DRAM-Chip ein permanenter Fehler auftritt, wird der Chip im ECC-Schema als fehlerhaft markiert. Die Daten aus dem fehlerhaften DRAM-Chip werden anschließend auf den Zusatz-DRAM-Speicher verlagert und das System wird in seinen ursprünglichen ECC-Status zurückversetzt (d. h. mit vollständigem ECC-Schutz). Dies ist ein ausgezeichnetes Beispiel für eine „selbstheilende“ Architektur, in der der Server automatisch fehlerhafte Komponenten durch funktionsfähige ersetzt, ohne dass sich die Leistung oder die Funktionalität verringert.

Da ein ECC-Wort auf ein Bankpaar verteilt wird, sind zwei Zusatz-DRAM-Chips pro Bankpaar vorhanden. Jeder dieser zusätzlichen DRAM-Chips kann als Ersatz für einen fehlerhaften DRAM-Chip auf demselben DIMM verwendet werden. Dadurch ist es möglich, dass ein einzelner DRAM-Chip in einem DIMM-Paar fehlerhaft ist und auf dem einen DIMM über Zusatzspeicher verfügt und später ein ähnlicher DRAM-Chipfehler auftritt, der jedoch auf dem anderen DIMM des Paares über Zusatzspeicher verfügt. Bei diesem Aufbau kann das System noch weitere DRAM-Chipfehler verkraften und zudem noch einen Bitzeilenfehler auf einem der DRAM-Chips im Bankpaar beheben.

Bei einem solchen Umfang an Zusatzspeicherkapazität wird erwartet, dass selbst geplante Ausfallzeiten zum Austauschen von DIMMs, die aufgrund von DRAM-Chipfehlern ausfallen, beträchtlich reduziert werden.

Speichertest und Aufhebung der Zuordnung von Seiten und logischen Speicherblöcken

POWER7-Systemspeichercontroller erkennen und beheben Fehler bei ihrem Auftreten nicht nur. Sie testen (bereinigen) den Speicher auch proaktiv. Dies wird durch den Serviceprozessor gesteuert. Der Speichertest umfasst das periodische Lesen des gesamten Speicherinhalts im System und die Suche nach Datenfehlern. Der Test umfasst den gesamten Hauptspeicher des Systems, unabhängig davon, ob er lizenziert ist oder nicht.

Wenn bei diesem Vorgang festgestellt wird, dass eine einzelne Zelle (Einzelbit an einer Einzeladresse) persistent fehlerhaft ist, kann der Serviceprozessor Folgendes ausführen: Er kann über POWER Hypervisor anfordern, dass die Speicherseite mit der Adresse keiner Partition und keinem Hypervisor-Code zugeordnet wird. Falls sie bereits zugeordnet ist, kann er anfordern, dass sie freigegeben wird. Diese ***dynamische Aufhebung der Speicherzuordnung*** ist so konzipiert, dass ein persistenter Fehler in einer einzelnen Speicherzelle auf einem DRAM-Chip nicht parallel zu einem Fehler in einem anderen DRAM-Chip im selben Speicherwort an derselben Adresse auftritt (um dadurch Doppelbitfehler zu verhindern).

Aus der Sicht der Hardware wird der Hauptspeicher aus ECC-Wörtern aus jeweils 72 Byte zusammengesetzt, die auf zwei DIMMs verteilt sind. Auf Betriebssystemebene erfolgt der Speicherzugriff über Speicherseiten. Speicherseiten sind Datensammlungen in der Größenordnung von einigen Kilobyte. Die Größe der Speicherseiten wird vom Betriebssystem festgelegt. Hauptspeicher wird Betriebssystemen in logischen Speicherblöcken (LMBs) zugeordnet. Ein logischer Speicherblock ist eine Gruppe von Speicherseiten in der Größenordnung von 16 bis 256 MB.

Für eine optimale Leistung enthält ein einzelner LMB in der Regel Daten aus allen DIMM-Paaren, die an einem Prozessormodul angeschlossen sind. Bei einem Unternehmenssystem wie Power 795 können die einzelnen LMBs in einem System Daten aus allen acht DIMMs enthalten, die an einem POWER7-Prozessorchip angeschlossen sind.

Wenn während einer Speichertestoperation ein nicht behebbarer Speicherfehler erkannt wird, fordert POWER Hypervisor an, dass der zugeordnete LMB im System ***freigegeben*** und nicht wiederverwendet werden soll. Wenn der LMB einer aktiven Partition zugeordnet wurde, wird die Partition beendet, bevor der LMB freigegeben ist. Der Serviceprozessor fordert darüber hinaus an, dass das Speicher-DIMM-Paar mit dem nicht behebbaren Fehler ersetzt wird.

Wenn analog dazu POWER Hypervisor oder eine Anwendung im normalen Systembetrieb einen nicht behebbaren Fehler im Speicher findet, wird die SUE-Fehlerbehandlung (SUE - Special Uncorrectable Error) eingesetzt. Diese dient dem Zweck, das mögliche Ausmaß einer UE-bedingten Ausfallzeit zu minimieren, und der LMB mit dem nicht behebbaren Fehler wird, wie oben erwähnt, freigegeben.

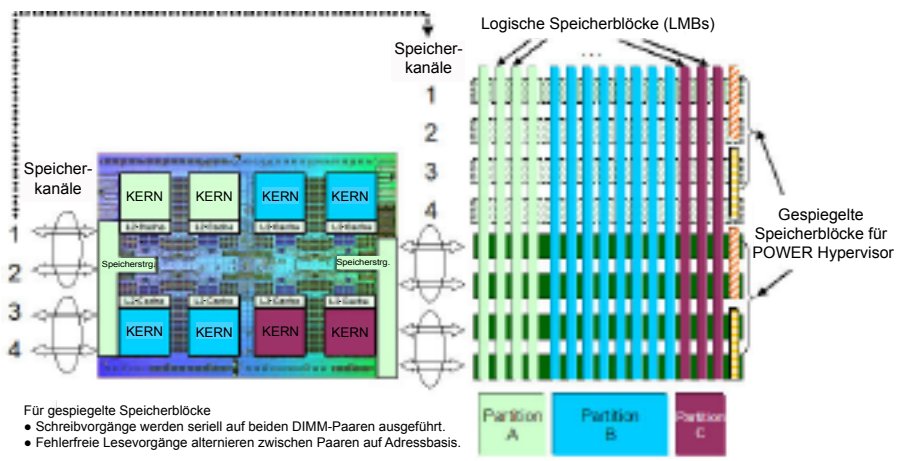
Active Memory Mirroring for Hypervisor

Für Systeme mit angepassten DIMMs ist die Wahrscheinlichkeit sehr gering, dass aufgrund defekter DRAM-Chips im Speicher ein nicht behebbarer Fehler auftritt. Der Grund dafür ist der selbstheilende Aufbau, bei dem ein innovatives ECC-Schema mit Zusatz-DRAM-Chips gekoppelt wird. Dadurch sollen Speicherfehler automatisch erkannt und behoben und Daten nötigenfalls auf Backup-Chips verlagert werden. Wie bereits beschrieben, sind mit diesen DIMMs ausgestattete Server dazu in der Lage, nacheinander Fehler bei zwei DRAM-Chips zu beheben. Nun müssen gegebenenfalls in einigen Fällen fehlerhafte Daten aus drei DRAM-Chips in einem DIMM-Paar und zudem mindestens eine zusätzliche Bitzeile korrigiert werden (eine für ***jedes*** Bankpaar im DRAM-Paar). Die Wahrscheinlichkeit eines nicht behebbaren Fehlers (UE) wird noch weiter reduziert, wenn der Serviceprozessor angibt, dass ein DIMM ausgetauscht werden muss und dies rechtzeitig durchgeführt wird.

Eine weitere mögliche Fehlerquelle für einen Speicher-UE ist ein Fehler in der DIMM-Logik, die die DRAM-Chips verwaltet oder die an der Verlagerung von Daten beteiligt ist. Fehler in dieser „sonstigen" Elektronik können auch ein katastrophaler Ausfall des Pufferchips oder verschiedene sonstige Probleme beim DIMM sein, wie zum Beispiel beim Spannungsregler. Ein derartiger Fehler kann mehrere nicht behebbare Fehler verursachen, die nicht auf einen einzelnen LMB begrenzt sind.

Der Hauptspeicherpufferchip (Seite ###27) ist so ausgelegt, dass er Soft Errors erkennt und behebt. IBM steuert den Fertigungsprozess der Pufferchips so, dass die Kontrolle der Chipqualität gewährleistet ist. Wenn dennoch ein katastrophaler Fehler auf DIMM-Ebene auftritt, können alle Partitionen im System ausfallen, die einem LMB auf dem fehlerhaften DIMM zugordnet sind. Wenn darüber hinaus ein DIMM einen LMB mit kritischen POWER Hypervisor-Daten enthält, kann dies zu einem systemweiten Ausfall führen.

Damit diese letztgenannte Möglichkeit ausgeschlossen wird, wurde der Speichercontroller von Power 795 so konzipiert, dass er das Spiegeln von ECC-Wörtern eines DIMM-Paars in ECC-Wörter eines anderen DIMM-Paars ermöglicht. Die Funktion ist insofern selektiv, als dafür nicht der gesamte Speicher und nicht einmal ein wesentlicher Teil davon gespiegelt werden muss. In der Praxis wird diese Funktion verwendet, um lediglich diejenigen LMBs zu spiegeln, die von POWER Hypervisor verwendet werden.



Ausgewählte RAS-Funktionen für den Speichercontroller

Hardwarespeichertest

- Korrigiert Einzelbit-Soft-Errors.
- Nutzt die Aufhebung der Seitenzuordnung des Betriebssystems für permanente Einzelbitfehler.
- Chipkill-Ereignisse können zur Nutzung von Zusatz-DRAM-Chips führen.

Für nicht behebbare Fehler im Speicher:

- Kann logische Speicherblöcke freigeben.

Active Memory Mirroring for Hypervisor

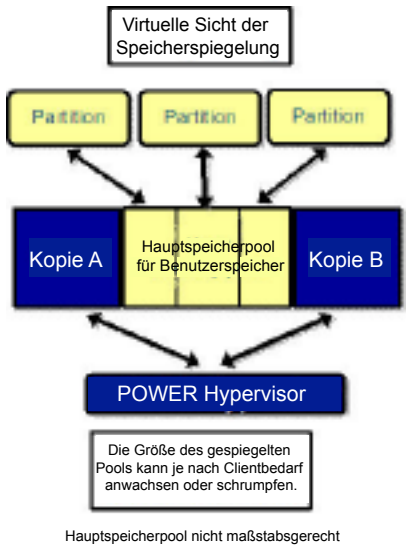
Active Memory Mirroring for Hypervisor

In dieser Zeichnung ist ein POWER7-Prozessormodul dargestellt, das an mehrere logische Speicherblöcke (LMBs) angeschlossen ist. Ein tatsächliches System verfügt über viel mehr Blöcke als in der Zeichnung. LMBs, die Partitionen zugeordnet sind, werden nicht gespiegelt. LMBs, die von POWER Hypervisor verwendet werden, können jedoch optional gespiegelt werden. Diese werden in der Abbildung durch die zwei kleineren LMBs ganz rechts dargestellt. Beachten Sie, dass sie auf DIMM-Paare verteilt gespiegelt werden.

Die meisten LMBs werden auf normale Weise gelesen und geschrieben. Für die LMBs, die POWER Hypervisor zugeordnet sind, werden alle Speicherschreibvorgänge an beide Speicherpaare gesendet, sodass die Daten gespiegelt werden. Wenn kein nicht behebbarer Fehler auftritt, alterniert der Speichercontroller die Lesevorgänge: Er liest Daten aus dem einen DIMM-Paar aus und anschließend aus dem nächsten. Falls beim Lesen ein nicht behebbarer Fehler auftritt und der Fehler nicht dadurch behoben wird, dass der Puffer auf dem Bus zwischen Prozessor und Speicher dynamisch angepasst wird, werden die Daten aus dem spiegelgleichen Paar abgerufen.

Dieser Vorgang wird als Active Memory Mirroring for Hypervisor oder als Systemfirmwarespiegelung bezeichnet. Dieses Produktmerkmal ist in einem Power 795-System standardmäßig aktiviert, kann jedoch bei Bedarf inaktiviert werden. Active Memory Mirroring setzt voraus, dass in den einzelnen Knoten eines Power 795-Systems mindestens ein Prozessormodul mit acht DIMMs vollständig konfiguriert sein muss. Falls Speicher nach einem Systemneustart so dekonfiguriert wird, dass dies nicht mehr zutrifft, inaktiviert das System möglicherweise vollständig oder teilweise das Spiegeln, bis eine erforderliche Mindestkonfiguration wiederhergestellt ist.

Die meisten LMBs werden auf normale Weise gelesen und geschrieben. Für die LMBs, die POWER Hypervisor zugeordnet sind, werden alle Speicherschreibvorgänge an beide Speicherpaare gesendet, sodass die Daten gespiegelt werden. Wenn kein nicht behebbarer Fehler auftritt, alterniert der Speichercontroller die Lesevorgänge: Er liest Daten aus einem DIMM-Paar aus und anschließend aus dem nächsten. Falls beim Lesen ein nicht behebbarer Fehler auftritt und der Fehler nicht dadurch behoben wird, dass der Puffer auf dem Bus zwischen Prozessor und Speicher dynamisch angepasst wird, werden die Daten aus dem spiegelgleichen Paar abgerufen.



Vereinfachte Ansicht:

Active Memory Mirroring

Beseitigt Systemausfälle aufgrund nicht behebbarer Fehler im POWER Hypervisor-Speicher.

Verwaltet zu jedem Zeitpunkt zwei identische Kopien von POWER Hypervisor im Speicher.

Beide Kopien werden bei allen Änderungen gleichzeitig aktualisiert.

Falls in der primären Kopie ein Speicherfehler auftritt, wird automatisch die zweite Kopie aufgerufen und über Electronic Service Agent eine Benachrichtigung an IBM gesendet (wennn der Administrator Electronic Service Agent aktiviert hat).

Dieser Vorgang wird als *Active Memory Mirroring for Hypervisor* oder als Systemfirmwarespiegelung bezeichnet. Dieses Produktmerkmal ist in einem Power 795-System standardmäßig aktiviert, kann jedoch bei Bedarf inaktiviert werden. Active Memory Mirroring setzt voraus, dass in den einzelnen Knoten eines Power 795-Systems mindestens ein Prozessormodul mit acht DIMMs vollständig konfiguriert sein muss. Falls Speicher nach einem Systemneustart so dekonfiguriert wird, dass dies nicht mehr zutrifft, inaktiviert das System möglicherweise vollständig oder teilweise das Spiegeln, bis eine erforderliche Mindestkonfiguration wiederhergestellt ist.

Obwohl dieses Produktmerkmal vor systemweiten Ausfällen aufgrund eines nicht behebbaren Fehlers im POWER Hypervisor-Speicher schützt, können Partitionen weiterhin ausfallen. Welche Partitionen genau davon betroffen sind und auf welche Weise, hängt von den jeweiligen LMB-Zuordnungen für Partitionen beim Auftreten des DIMM-Fehlers ab.

Darüber hinaus ist es möglicherweise nicht möglich und nicht einmal ratsam nach einem Partitionsausfall aufgrund eines DIMM-Fehlers, betroffene LMBs neu zu starten. Falls der nicht behebbare Fehler (UE) von einem katastrophalen DIMM-Ausfall verursacht wurde, treten bei allen dem DIMM zugeordneten LMBs möglicherweise schließlich UE-Ereignisse auf. Falls ein UE während des Neustarts einer Partition auftritt, verhindert er möglicherweise den Neustart dieser Partition.

Persistente Aufhebung der Zuordnung von Komponenten und System-IPL

Erkennung und Aufhebung der Zuordnung fehlerhafter Komponenten

Zur Laufzeit behebbare/korrigierbare Fehler werden überwacht, um zu bestimmen, ob ein Fehlermuster oder ein „Trend in Richtung Nichtbehebbarkeit“ vorliegt. Wenn eine Komponente eine vordefinierte Fehlergrenze erreicht, leitet der Serviceprozessor eine Maßnahme ein, um die „fehlerhafte“ Hardware zu dekonfigurieren. Hierdurch wird ein potenzieller Systemausfall verhindert und die Systemverfügbarkeit erhöht. Dieser Vorgang wurde anhand mehrerer Beispiele veranschaulicht (zum Beispiel anhand der dynamischen Aufhebung der Prozessorzuordnung, der Cachezeilenlöschung, der dynamischen Aufhebung der Seitenzuordnung und der Speicherblocklöschung). Fehlergrenzen werden von IBM Entwicklern voreingestellt und basieren auf protokollierten Mustern des Komponentenverhaltens in unterschiedlichen Betriebsumgebungen. Fehlerschwellen werden in der Regel von Algorithmen unterstützt, die einen zeitgesteuerten Zähler für behebbare Fehler einschließen. Dies bedeutet, dass der Serviceprozessor auf eine Bedingung antwortet, bei der innerhalb eines definierten Zeitraums zu viele Fehler auftreten.

Persistente Aufhebung der Zuordnung

Wenn ein Fehler erkannt und eine Komponente zum Reparieren markiert wird, markiert das System zur Verbesserung der Systemverfügbarkeit das ausgefallene Element, um es bei nachfolgenden Systemneustarts zu dekonfigurieren. Dies wird als persistente Aufhebung der Zuordnung (Persistent Deallocation) bezeichnet und soll sicherstellen, dass eine ausgefallene Komponente nicht wieder in Betrieb genommen wird, wenn das System vor der Reparatur der Komponente neu gestartet wird.

Das Entfernen von Komponenten kann entweder dynamisch erfolgen (während das System aktiv ist) oder beim Booten (IPL). Dies hängt jeweils vom Typ des Fehlers und vom Zeitpunkt ab, wann der Fehler erkannt wurde.

Darüber hinaus können Komponenten, bei denen nicht behebbare Hardwarefehler während des Betriebs auftreten, nach dem ersten Auftreten im System dekonfiguriert werden. Das System kann unmittelbar nach dem Fehler neu gestartet werden und den Betrieb mit der verbleibenden fehlerfreien Hardware fortsetzen. Dadurch wird verhindert, dass dieselbe „fehlerhafte" Hardware erneut den Systembetrieb beeinträchtigt. So kann die Reparaturaktion auf einen passenderen, unkritischeren Zeitpunkt verschoben werden.

Aus verschiedenen Gründen, unter anderem aus Leistungsgründen und wegen der Vorbereitung auf die Wartung, müssen die während des Systembetriebs dekonfigurierten Komponenten nicht notwendigerweise dieselben sein wie beim erneuten Booten des Systems. Wenn zum Beispiel während des Betriebs ein nicht behebbarer Fehler in einem Level-2-Prozessorcache auftritt, wird die fehlerhafte Cachezeile gelöscht. Falls eine große Anzahl Cachezeilen gelöscht werden, kann es beim erneuten Booten effizienter und aus Leistungsgründen besser sein, den gesamten Cache und den ihm zugeordneten Prozessorkern zu dekonfigurieren.

Ein anderes Beispiel: Obwohl im Systembetrieb einzelne LMBs des Hauptspeichers möglicherweise dekonfiguriert werden, wird bei einem erneuten Booten ein vollständiges DIMM dekonfiguriert. Da ECC-Wörter immer auf zwei DIMMs verteilt werden, wird aufgrund der Zugehörigkeit das jeweils andere DIMM im DIMM-Paar ebenfalls dekonfiguriert. Obwohl dadurch zwei vollständige DIMMs aus dem System entfernt werden, können auf den übrigen DIMMs logische Speicherblöcke freigegeben werden. Dies ist häufig die effizienteste Methode, das System für optimale Leistung und Speicherkapazität zu konfigurieren, wenn mehrere LMBs an einem Ausfall beteiligt sind.

Außerdem kann es in einigen Fällen nützlich sein, einen vollständigen Knoten beim Booten zu dekonfigurieren, wenn eine Komponente im Knoten ausgefallen ist. (Dies ist jedoch nicht das Standardverfahren.) Dadurch können die übrigen Knoten im System betrieben werden, während der dekonfigurierte Knoten repariert und später wieder integriert wird.

Da Systeme in vielen unterschiedlichen Umgebungen implementiert und verwaltet werden, legt der Anwender eine Richtlinie fest. Diese steuert, ob bestimmte Fehler während des Betriebs und des Neustarts das Dekonfigurieren bewirken und ob beim Neustart ein vollständiger Knoten aufgrund eines Fehlers im Knoten dekonfiguriert wird.

Schließlich gibt es den Fall, dass eine festgelegte Richtlinie eine Bedingung verursacht, bei der keine ausreichenden Ressourcen verfügbar werden, um einen beliebigen Teil eines Systems bei einem Neustart aktivieren zu können. In diesem Fall können einige vorher dekonfigurierte Komponenten wieder in Betrieb genommen werden, damit nach einem Ausfall der Systembetrieb fortgesetzt werden kann.

Verfügbarkeit des E/A-Subsystems

Basisaufbau

Zweifellos umfasst ein System mehr als die zentrale Elektronik (CEC - Central Electronics Complex). Dies muss also auch für den RAS-Systemaufbau gelten. Das Erstellen eines soliden, zuverlässigen, verfügbaren E/A-Subsystems ist entscheidend, um eine höhere System-RAS-Ebene zu erreichen. Leider erschwert es die große Vielfalt und Komplexität verfügbarer E/A-Einheiten, die RAS-Merkmale einzelner Einheiten ausführlich zu behandeln. Eine Beschreibung dazu, wie die Systemverfügbarkeit durch das E/A-Subsystem beeinflusst wird, ist jedoch wichtig für das Verständnis der Anwendungsverfügbarkeit und kann in diesem Dokument behandelt werden.

POWER7-Server wurden so entworfen, dass sie eine fortlaufende Verfügbarkeit von Systemen, Partitionen und Adaptern ermöglichen. Dies wird vor allem durch Folgendes erreicht:

1. Herstellen von Einheiten, die das System mit den E/A-Einheiten verbinden und dabei zuverlässige, über eine bestimmte Ausfallsicherheit verfügende Komponenten verwenden (Komponenten, die einige zu erwartende Fehlertypen verkraften)

2. Bereitstellen von Redundanzen in E/A-Adaptern und Einheiten

3. Sicherstellen, dass Einheiten und Einheitentreiber gemäß strenger Spezifikationen aufgebaut und vor der Implementierung gründlich getestet wurden

Als das erste POWER4-System eingeführt wurde, hießen die bereitgestellten E/A-Adapter PCI-Adapter, da sie die Spezifikation Peripheral Connect Interchange unterstützten. Diese Spezifikation war damals Standard.

Diese PCI-Adapter waren dafür ausgelegt, an Systeme angeschlossen zu werden, die so skaliert werden konnten, dass sie mehrere E/A-Adapter in mehreren E/A-Einschüben unterstützten. IBM entwickelte daher eine hierarchische E/A-Struktur, um sie zu unterstützen.

In dieser Architektur kommunizieren die Systemprozessormodule über einen Hochgeschwindigkeitsbus (den GX-Bus) mit einer E/A-Hubkarte, die als GX-Busadapter bezeichnet wird.

E/A-Hubkarten verfügten über redundante Verbindungen zu E/A-Einschüben, in denen sich mehrere PCI-Adapter befanden. Der Bus zwischen dem E/A-Hub und den E/A-Einschüben war eine Hochgeschwindigkeitsverbindung (namens Remote Input/Output oder einfach RIO), die erweiterte RAS-Merkmale unterstützte (einschließlich der Wiederholung fehlgeschlagener Operationen und der Redundanz mit Failover-Fähigkeit). Die RAS-Merkmale ermöglichten es, E/A-Einschübe gleichzeitig zum übrigen Systembetrieb einzustecken und zu entfernen. Die einzelnen E/A-Einschübe verfügten über Module, die an diesen Bus angeschlossen wurden und ihrerseits die Funktion einer PCI-Hostbrücke erfüllten.

Für dieses System bestand ein wichtiges Ziel von RAS darin, das *dynamische Entfernen und Austauschen von PCI-Adaptern* zu unterstützen. Dazu war eine elektrische PCI-Adapterisolation erforderlich (Adapter in jeweils separaten Stromversorgungsdomänen), sodass das übrige System vor möglichen transienten elektrischen Störsignalen während des Einsteckvorgangs geschützt war. Aus diesem Grund stellte IBM einen Controller-Chip her (den EADS-Controller), der Daten von der PCI-Hostbrücke empfing, und *isolierte den PCI-Bus* für jeden von ihm gesteuerten E/A-Steckplatz elektrisch.

Die damalige PCI-Spezifikation ermöglichte die Paritätsprüfung auf dem PCI-Bus. Dadurch konnten Busfehler erkannt werden. Allerdings konnten sie über den Prüfcode nicht behoben werden. Da ein Unternehmenssystem zahlreiche PCI-Adapter enthalten kann, war es wichtig, Datenbusfehler im E/A-System nicht nur zu erkennen, sondern auch zu beheben. Daher integrierten IBM Entwickler auf dem EADS-Chip eine entsprechende Unterstützung, mit der an einen Einheitentreiber signalisiert werden konnte, dass ein PCI-Busfehler aufgetreten war. Dadurch konnte der Einheitentreiber den Steckplatz zurücksetzen und die fehlgeschlagene Operation wiederholen. Falls Wiederholungen nicht durchgeführt wurden oder erfolglos waren, konnte der Einheitentreiber die Nutzung des fehlerhaften PCI-Adapters beenden. Eine Partition oder ein System musste dazu nicht direkt beendet werden.

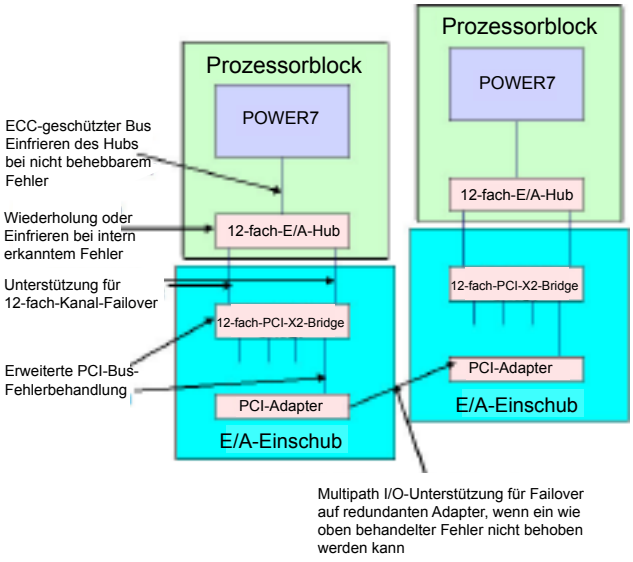
Das Basiskonzept für die Behandlung von Fehlern auf dem PCI-Bus wurde als *Enhanced Error Handling* (EEH) für E/A-Adapter bezeichnet.

E/A-Konnektivität eines für optimale Verfügbarkeit konfigurierten E/A-Adapterpaars

IBM hat bereits lange Zeit Server mit innovativen RAS-Funktionen gebaut, um die Integrität der E/A-Datenübertragung zu schützen. Jede nachfolgende Power-Generation enthält neue Funktionsmerkmale, die die Serververbindungen zwischen CEC und E/A-Einheiten schützen.

POWER7 verwendet weiterhin hoch entwickelte Funktionsmerkmale für angeschlossene E/A-Einheiten wie den über 12X-DDR-InfiniBand angeschlossenen E/A-Hub mit Einfriermodus. Wenn Systeme für eine maximale Verfügbarkeit konfiguriert sind, können sie nahezu völlig vor Ausfällen (geplanten und ungeplanten) geschützt werden, die von Fehlern im E/A-System verursacht werden.

In dieser Abbildung ist eine Konfiguration für die hohe Verfügbarkeit bei E/A-Operationen dargestellt. In diesem Fall müssen die einzelnen kritischen E/A-Adapterfunktionen von einem der beiden E/A-Adapter ausgeführt werden können. Hierbei befindet sich der erste Adapter in einem E/A-Einschub, der über einen auf 12X-DDR-InfiniBand basierenden E/A-Hub-Controller an einen CEC-Knoten angeschlossen ist. Der zweite Adapter befindet sich in einem anderen E/A-Einschub, der an einen anderen auf 12X-DDR-InfiniBand basierenden E/A-Hub-Controller in einem separaten CEC-Knoten angeschlossen ist. In dieser Konfiguration gibt es eine vollständige Redundanz für Fehler bei einem Adapter, bei der E/A-Platine oder beim E/A-Hub selbst. Außerdem kann eine ständige Verbindung zu E/A-Einheiten aufrechterhalten werden, wenn ein CEC-Knoten während des Betriebs entfernt wird, um ihn zu reparieren.



Spätere Entwicklungen

Dieses Basiskonzept für die E/A-Fehlerbehandlung wurde in nachfolgenden Adaptergenerationen auf mehrere Arten erweitert. Die Fähigkeit eines Betriebssystems, die Funktion auf einen redundanten Adapter in einem anderen E/A-Einschub eines Systems zu übertragen und dabei Verfahren wie ***Multi-Path I/O*** einzusetzen, wurde für Fälle hinzugefügt, in denen die Wiederholung von Operationen das Problem nicht behebt.

Dieser Mechanismus wurde primär entwickelt, damit der Betrieb fortgesetzt werden kann, wenn bei einem einzelnen PCI-Adapter ein nicht behebbarer Fehler auftritt. Im Laufe der Zeit wurde das EEH-Verhalten für E/A-Adapter so erweitert, dass Fehler in der Prozessorhostbrücke (PHB – Processor Host Bridge) abgedeckt wurden. Dies war so konzipiert, dass beim ***Einfrieren einer PHB*** alle der PHB untergeordneten Adapter entweder zurückgesetzt werden und E/A-Operationen wiederholt werden mussten oder die Verwendung der Adapter beendet werden musste.

Übergänge von PCI zu PCI-X und PCI-E und von RIO zu RIO-G und InfiniBand

Während diese Grundverbesserungen der Verfügbarkeit im E/A-Subsystem durchgeführt wurden, gab es teilweise auch bei der zugrunde liegenden Struktur des E/A-Subsystems Fortschritte. Beispiele:

- Es wurden Zusatzeinrichtungen zur Unterstützung PCI-X- und PCI-E-basierter E/A-Steckplätze hinzugefügt. Dazu gehörte die Unterstützung für neue RAS-Merkmale, die in den Busarchitekturen selbst begründet waren.
- Der E/A-Bus, über den ein CEC mit seinen E/A-Einschüben verbunden wurde, wurde mit einer höheren Leistung ausgestattet, als er von einem RIO-Bus zu einem RIO-G-Bus und schließlich zu einem 12-fach-Kanal migriert wurde (der InfiniBand-Busprotokolle verwendet).

SUE-Behandlung für E/A-Hub-RAS und Ein-/Ausgabe

Der Bus, der einen Prozessorchip mit einer E/A-Hubkarte verbindet (GX-Busadapter) wurde mehrmals verbessert: Die Geschwindigkeit wurde erhöht und es wurde Funktionalität hinzugefügt. Die Busversionen waren daran erkennbar, dass an den Busnamen einfach „+“ oder „++“ angehängt wurde. Ein GX-Adapter in POWER7-Systemen behält jedoch, unabhängig von der Version, einige wichtige RAS-Funktionen bei. Der Bus ist mit einem ECC-Schutz ausgestattet. Dadurch wird sichergestellt, dass gegen Einzelbitfehler ein durchgehender Schutz vorhanden ist.

Wenn Daten, die mit einem SUE-Code (SUE - Special Uncorrectable Error) markiert sind, über den GX-Bus übertragen werden, wird der SUE im E/A-Hub-Controller verarbeitet. Der Hub-Controller stellt sicher, dass die fehlerhaften Daten gelöscht werden, bevor sie einen E/A-Adapter erreichen. Nach dem Löschen werden alle nachfolgenden E/A-Operationen an alle dem Hub untergeordneten Einheiten (auf PCI-Busseite) „eingefroren“. Dadurch haben Einheitentreiber keinen Zugriff mehr auf ihre Einheiten. Diese Fehler werden als persistente EEH-Ereignisse behandelt (und der normale Systembetrieb kann fortgesetzt werden, wenn geeignete Redundanzen vorhanden sind).

Fehler auf der Schnittstelle zwischen dem E/A-Hubadapter und einem E/A-Einschub werden allgemein dadurch behandelt, dass im Busprotokoll redundante Pfade zu den E/A-Einschüben sowie Wiederholungsmechanismen verfügbar sind. Aufgrund der Eigenschaften des GX-Busses selbst kann jedoch ein älterer E/A-Hubadapter (der für nicht mehr verfügbare RIO-E/A-Einschübe ausgelegt ist) Ausfälle verursachen, falls ein nicht behebbarer Fehler beim Bus oder bestimmte andere interne Fehler im Hub auftreten. Dieser Systemaufbau soll die Datenintegrität sicherstellen, denn falsche Daten dieser Art können nicht nur für die E/A-Hubeinheit, sondern auch für die übrigen Prozessorkomponenten sichtbar sein, die zur Cachekohärenzdomäne gehören.

Über 12x-DDR-InfiniBand angeschlossene E/A-Hubadapter und „Einfriermodus“.

Innerhalb des POWER6-Zeitrahmens wurde ein neuer E/A-Hubadapter für den Anschluss an E/A-Einschübe eingeführt. Dabei wurde ein optionaler doppelter 12x-DDR-InfiniBand-Anschluss verwendet (12x Channel Adapter). Dieser Adapter unterstützte eine weitere wichtige Verbesserung des E/A-Subsystemschutzes.

Der E/A-Hub-Controller wurde so überarbeitet, dass in Zusammenarbeit mit dem POWER-Prozessor bei Fehlern im Hub, die entweder mit dem GX-Bus oder mit dem Prozessor in Beziehung stehen, in einigen Fällen eine Wiederholung durchgeführt werden kann. Dadurch wird bei transienten Ereignissen ein Ausfall verhindert. Wenn für einen Hubfehler eine Wiederholung nicht geeignet ist oder wenn dadurch das Problem nicht gelöst werden kann, kann der E/A-Hub auf ähnliche Weise wie im oben erwähnten SUE-Fall alle angeschlossenen E/A-Adapter einfrieren. Dieser ***Einfriermodus*** kann mit E/A-Redundanz kombiniert werden, um Ausfälle zu verhindern.

POWER7-E/A-Gehäuse und integrierte Ein-/Ausgabe

Verfügbarkeit angeschlossener E/A-Einheiten

POWER7 übernimmt die in früheren Prozessorgenerationen entwickelten RAS-Merkmale für angeschlossene E/A-Einheiten. Dazu gehören Zusatzeinrichtungen wie die Verwendung des über 12X-DDR-InfiniBand angeschlossenen E/A-Hubs mit Einfriermodus. Wenn Systeme für eine maximale Verfügbarkeit konfiguriert sind, können sie nahezu völlig vor Ausfällen (geplanten und ungeplanten) geschützt werden, die von Fehlern im E/A-System verursacht werden.

Zum Erreichen dieser Verfügbarkeitsebene müssen die einzelnen kritischen E/A-Adapterfunktionen von einem der beiden E/A-Adapter ausgeführt werden können. Hierbei befindet sich der erste E/A-Adapter eines Systems in einem E/A-Einschub, der über einen auf 12X-DDR-InfiniBand basierenden E/A-Hub-Controller an einen CEC-Knoten angeschlossen ist. Der zweite Adapter befindet sich in einem anderen E/A-Einschub, der an einen anderen auf 12X-DDR-InfiniBand basierenden E/A-Hub-Controller in einem separaten CEC-Knoten angeschlossen ist. In dieser Konfiguration gibt es eine vollständige Redundanz für Fehler bei einem Adapter, bei der E/A-Platine oder beim E/A-Hub selbst. Außerdem kann eine ständige Verbindung zu E/A-Einheiten aufrechterhalten werden, wenn ein CEC-Knoten während des Betriebs zur Reparatur entfernt wird.

Darüber hinaus ist es wichtig, die RAS-Merkmale der Einheiten zu berücksichtigen, die an diesen E/A-Adaptern angeschlossen sind. Damit der Failover funktioniert, ist es selbstverständlich erforderlich, dass das Netz/die Einheiten, mit dem/denen die Adapter verbunden ist/sind, von mehreren E/A-Controllern aus gesteuert werden können. Es wird vorausgesetzt, dass jeder kritische Speicher auf irgendeine Weise redundant ist, zum Beispiel mithilfe der DASD-Spiegelungstechnologie oder mithilfe von RAID.

Außerdem muss beachtet werden, dass die einzelnen angeschlossenen E/A-Einschübe redundante Stromversorgungen und Kühlungen umfassen. Dadurch erhöht sich die Verfügbarkeitsebene für die Einschübe.

Integrierte Ein-/Ausgabe

In der bisherigen Beschreibung wurde weitgehend davon ausgegangen, dass sich alle E/A-Adapter in einem System in separaten E/A-Einschüben befinden. In anderen Systemmodellen können sich E/A-Adapter und -Einheiten jedoch in den CEC-Knoten selbst befinden.

Die einzelnen CEC-Einschübe von Power 780 unterstützen zum Beispiel intern sechs PCI-E-8x-Steckplätze, vier 1-Gb-Ethernet-Anschlüsse und optional zwei optische/twinaxiale 10-Gb-Steckplätze mit zwei 1-Gb-Ethernet-Anschlüssen, zwei SAS-DASD-Controller mit internem optionalem RAID, einen SATA-Datenträgercontroller und sonstige Unterstützung für serielle Anschlüsse sowie USB- und HMC-Anschlüsse (HMC - Hardware Management Console).

Obwohl diese „integrierte Ein-/Ausgabe“ zahlreiche RAS-Merkmale umfasst, die für separate E/A-Einschübe beschrieben wurden, bestehen gewisse RAS-Einschränkungen. Beispiele:

- Wenn ein CEC-Knoten aus irgendeinem Grund nicht verfügbar wird, sind alle integrierten E/A-Einheiten im CEC-Knoten ebenfalls nicht verfügbar.
- Power 770- und 780-Systeme verwenden einen integrierten E/A-Hub-Controller, der den E/A-Hubeinfriermodus nicht unterstützt. Der integrierte E/A-Hub-Controller unterstützt das Einfrieren auf Steckplatzebene.
- Obwohl in jedem Gehäuse eine interne RAID-Zusatzeinrichtung unterstützt wird, schließt RAID nicht die Möglichkeit aus, dass auf die integrierten DASD-Einheiten nicht mehr zugegriffen werden kann, falls während des Betriebs bestimmte CEC-Ausfälle auftreten, zu denen auch geplante Wartungen mit Reparatur des Knotens gehören.

Zuverlässigkeit und Verfügbarkeit der Systeminfrastruktur

Allgemeine Systemumgebungen

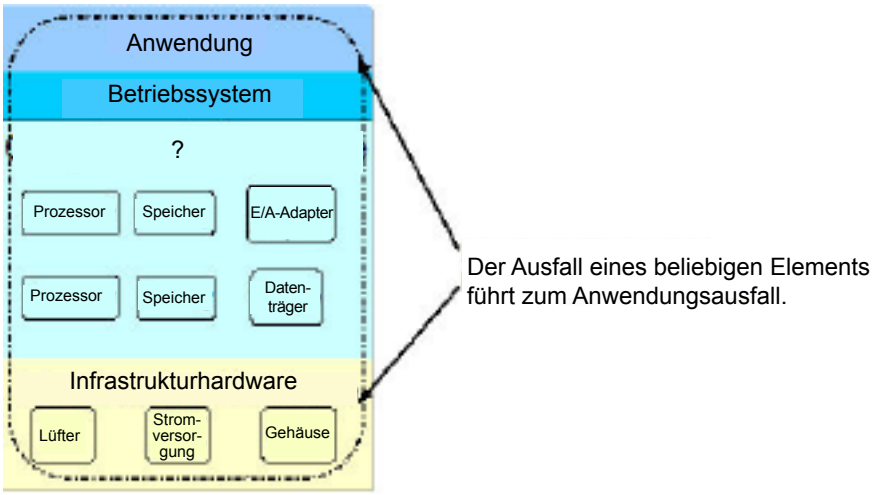
Die vorherigen Abschnitte konzentrierten sich auf RAS-Merkmale der IT-Ressourcen eines Computersystems – Prozessoren, Hauptspeicher und E/A-Einheiten. Zur Systemhardware gehören jedoch auch Infrastrukturressourcen – Stromversorgungen, Spannungsregler, Lüfter, Kühlkomponenten, Taktgeber und die Infrastruktur zur Unterstützung von Serviceprozessoren.

Die Verfügbarkeit von Systemanwendungen hängt von der Zuverlässigkeit und Verfügbarkeit all dieser Ressourcen ab, auch wenn die Folgen eines Ausfalls einer solchen Ressource davon abhängen, wie Systemressourcen implementiert sind und welche Hardwarevirtualisierung eingesetzt wird.

Bevor wir uns der Infrastrukturverfügbarkeit zuwenden, ist es nützlich, verschiedene Serverumgebungen zu untersuchen, um zu sehen, wie die Infrastrukturverfügbarkeit von Rechenzentren sich auf den Systemaufbau für Verfügbarkeit auswirkt.

Einzelserver/Betriebssystemumgebung

In einem einfachen Standalone-System kann der Ausfall praktisch jedes beliebigen Systemelements einen Anwendungsausfall verursachen. Daher hängt die Verfügbarkeit einer Anwendung sehr unmittelbar von der Verfügbarkeit der Computerkomponenten, der Infrastruktur, dem Betriebssystem usw. ab.

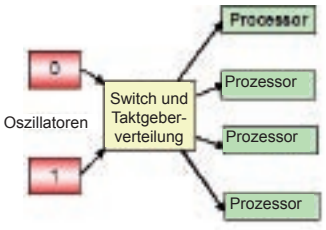


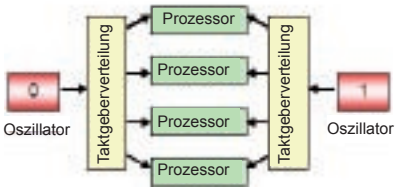
Verdeckte „Single Points of Failure" im Standalone-Server-Design

Das Ziel jedes Servers für „geschäftskritische" Anwendungen ist einfach: Die Anwendungen müssen aktiv bleiben. Die Verfügbarkeit auf Systemebene ist umso höher, je zuverlässiger die zugrunde liegende Hardware und die Verfahren sind, mit denen die Folgen auftretender Fehler minimiert werden. Zu diesem Zweck setzen Entwickler sehr zuverlässige Komponenten in einem hoch verfügbaren Systemaufbau ein, der sicherstellt, dass die meisten Hardwarefehler zu keinem Ausfall einer Anwendung führen. Die Entwickler streben Folgendes an:

1. Sie bauen eine Systemumgebung auf, die zu einem ordnungsgemäßen Betrieb von Komponenten beiträgt.
2. Sie ermitteln Komponenten, die höhere Fehlerraten aufweisen. Sie konzipieren den Server so, dass sporadisch auftretende Fehler in diesen Komponenten behoben werden und/oder ein Failover zu redundanten Komponenten ausgeführt wird. Sie wenden zum Beispiel Strategien wie die Fehlerbehebung durch automatische Wiederholung und Zusatzressourcen (Redundanz) an, um Fehler zu beheben.

In einer Einzelsystemumgebung sind diese Typen von Bedingungen nicht immer leicht erkennbar, obwohl möglicherweise dafür gesorgt wurde, dass die Anzahl der SPOFs (Single Points of Failure) reduziert ist. Im Beispiel unten sehen Sie zwei Server-Designs, die „redundante Taktgeber" bereitstellen. Obwohl der Server vor einem Ausfall des Oszillators geschützt ist, kann ein Fehler bei einer Schaltkomponente eines Taktgebers ohne Zweifel einen Serverausfall aufgrund einer einzigen Fehlerstelle verursachen. In Server-Designs gibt es viele derartige Beispiele (siehe E/A-Konnektivität eines für optimale Verfügbarkeit konfigurierten E/A-Adapterpaars auf Seite 33).





Aufbau der Taktgeberverteilung

Redundanter Oszillatoraufbau 1

- Bei einem Oszillatorfehler gibt es keinen Ausfall.
- Ein Ausfall der Verteilungskomponente für Schalter/Taktgeber führt zu einem Ausfall aller Prozessoren.
- Wenn die Taktgeberschnittstelle im Prozessor ausfällt, ist der Ausfall je nach Systemaufbau möglicherweise auf den Prozessor beschränkt.
- Ein permanenter Schalterfehler führt dazu, dass das System auch nach einem Neustart ausgefallen bleibt.

Redundanter Oszillatoraufbau 2

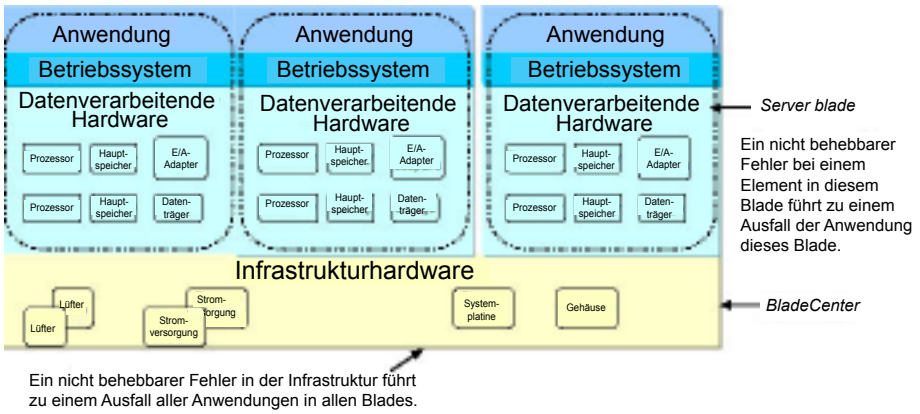
- Bei einem Oszillatorfehler oder bei einem Fehler einer Verteilungskomponente für Taktgeber gibt es keinen Ausfall.
- Wenn die Taktgeberschnittstelle im Prozessor ausfällt, ist der Ausfall je nach Systemaufbau möglicherweise auf den Prozessor beschränkt.
- Kein Single Point of Failure verhindert den Neustart.

[6] Verfügbarkeitstermine finden Sie im IBM United States Hardware Announcement 110-158, vom 17. August 2010 und im IBM United States Hardware Announcement 110-025 vom 9. Februar 2010.

BladeCenter und Bladeumgebung

Als Alternative zur Implementierung mehrerer kleiner Standalone-Systeme, implementieren einige Clients mit vielen, relativ kleinen Workloads sogenannte „Blade-Server“, die in einem einzigen Gehäuse zusammengefasst werden können. Aufgrund ihres allgemeinen Formfaktors (häufig befindet sich ein Blade auf einer einzigen Schaltkreiskarte) können diese Server auf einfache Weise in ein Standardgehäuse wie das Gehäuse einer IBM BladeCenter-Umgebung eingebaut werden.

In einem BladeCenter nutzen mehrere unabhängige Server-Blades eine gemeinsame Infrastruktur aus gemeinsamen Stromversorgungen, zentralen Kühlelementen und Zusatzeinrichtungen zum Herstellen von Verbindungen zwischen Blades, einem gemeinsamen Gehäuse usw.



Verfügbarkeit in einer BladeCenter-Umgebung

Da ein Fehler in der Infrastrukturhardware eines BladeCenter jedes Blade im System beeinträchtigen kann, sind Bladegehäuse so ausgelegt, dass diese Fehlerrate sehr gering ist. Das BladeCenter-Entwicklerteam betrachtete es als entscheidend, eine extrem robuste Hardwareinfrastruktur zu bauen.

Im Jahr 2008 wurde in einer Studie vorhergesagt, dass die BladeCenter-Infrastruktur im Durchschnitt ungefähr 0,8 Minuten der Gesamtausfallzeit eines Einzelblade pro Jahr beiträgt und eine Gesamtausfallzeit pro Jahr für diesen Blade von über 14 Minuten pro Jahr zu erwarten ist. Dieses BladeCenter-Beispiel verdeutlicht den sehr zuverlässigen Aufbau einer gemeinsamen Infrastruktur mit vielen Redundanzen. Dennoch werden Ausfallzeiten infolge von Fehlern bei der Hardwareinfrastruktur oder bei einzelnen Blades nicht vollständig beseitigt.

Da ein Fehler in der Infrastrukturhardware jedes Blade im System beeinträchtigen kann, sind Bladegehäuse so ausgelegt, dass die Fehlerrate der Infrastrukturhardware sehr gering ist. Dies wird normalerweise erreicht, indem Komponenten mit sehr niedrigen bauartbedingten Fehlerraten eingesetzt werden (zum Beispiel passive Systemplatinen ohne aktive Komponenten) oder indem redundante Ressourcen verwendet werden (zum Beispiel redundante Stromversorgungen und Lüfter).

Selbst bei Einsatz all dieser Verfahren werden üblicherweise einige Fehler oder Kombinationen von Fehlern innerhalb der Infrastrukturhardware erwartet, die sowohl zu geplanten als auch zu ungeplanten Ausfallzeiten führen können.

Bewertung der Fehlerrate einer „redundanten“ und „passiven“ Infrastruktur

Das BladeCenter-Entwicklerteam betrachtete es als entscheidend, eine extrem robuste Hardwareinfrastruktur zu bauen. Es erwies sich als durchaus erfolgreich und ein beiläufiger Blick auf die Beschreibung des Systemaufbaus könnte zum Schluss führen, dass Infrastrukturfehler nie Ausfälle von Anwendungen verursacht hätten, die in Blades ausgeführt wurden.

Im Jahr 2008 veröffentlichte IBM jedoch einen Artikel in IBM Systems Journal mit dem Titel „Availability analysis of blade server systems“[7]. Darin wurden der Aufbau und die Zuverlässigkeit der Infrastruktur einer Blade-/BladeCenter-Umgebung, die auf BladeCenter HS20-Servern basierte und mit Xeon-Prozessoren von Intel™ konfiguriert war, gründlich behandelt. Der Artikel zeigte, wie mit beträchtlichem Entwicklungsaufwand Ausfälle, die mit einer BladeCenter-Infrastruktur im Zusammenhang standen, auf lediglich vorhergesagte 0,8 Minuten pro Jahr reduziert werden konnten. Dies gilt sowohl für geplante als auch für ungeplante Ausfallzeiten.

Obwohl die Infrastruktur als sehr robust angesehen werden konnte, beschrieb der Artikel einige Ursachen dafür, weshalb von einer gemeinsamen Infrastruktur nicht erwartet werden kann, dass bei ihr keine Fehler auftreten.

[7] W.E. Smith, K.S. Trivedi, L.A. Tomek, J. Ackaret, „Availability analysis of blade server systems“, IBM Systems Journal, Band 47, Nr. 4, 2008.

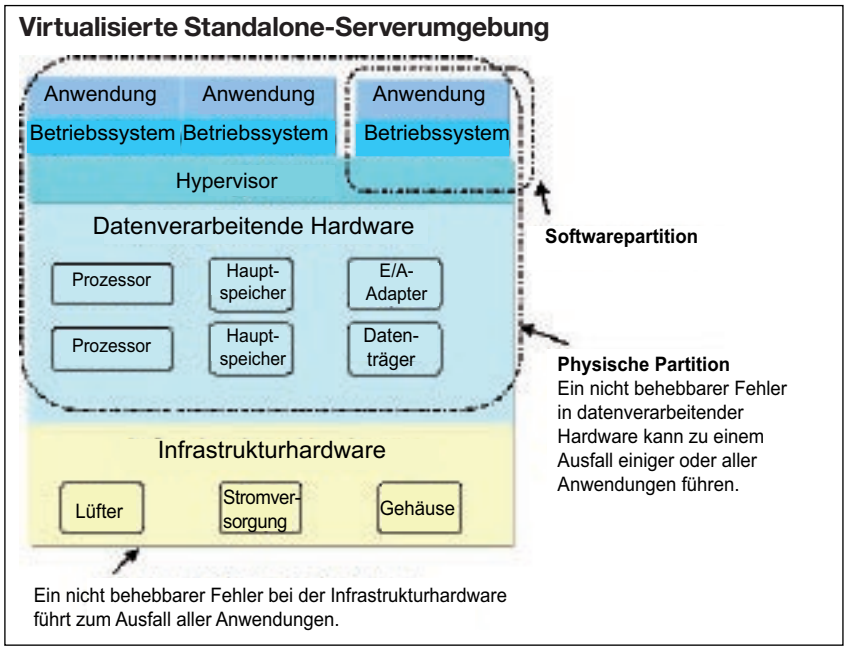
Dafür gibt es verschiedene Gründe. Beispiele:

1. Eine Komponente eines Paars redundanter Komponenten fällt möglicherweise aus. Noch bevor die erste ausgetauscht worden ist, fällt die andere Komponente des Paars möglicherweise ebenfalls aus.
2. Der Ausfall der ersten Komponente in einem redundanten Paar deckt möglicherweise einen verborgenen Defekt im redundanten Paar auf, besonders wenn in einem aktiven System unzureichende Mittel vorhanden sind, den fehlerfreien Zustand des Paares sicherzustellen.
3. Der Failover zu einer redundanten Einheit erfordert möglicherweise Code und der Code ist möglicherweise beschädigt.
4. Und schließlich gibt es selbst in Fällen mit zwei Komponenten, die redundant sein sollen, häufig eine gemeinsame Komponente oder Verbindung zwischen den redundanten Teilen. Der Ausfall dieser gemeinsamen Komponente kann zu einem Ausfall führen. Ein mögliches Beispiel dafür: Zwei Lüfter in einem System werden von beiden Stromversorgungen in einem System mit Strom versorgt, damit beide Lüfter nach dem Ausfall einer Stromversorgung weiterhin betrieben werden können. Falls einer der Lüfter in einem System so ausfällt, dass ein Überstrom (oder ein elektrischer Kurzschluss) auftritt, sind möglicherweise beide Stromversorgungen davon betroffen und versorgen beide Lüfter nicht mehr mit Strom. Obwohl Möglichkeiten entwickelt wurden, diese normalen Ausfälle zu verhindern oder zu umgehen, gibt es weiterhin einige derartige Bedingungen.

Es muss jedoch noch einmal Folgendes betont werden: Obwohl die Auswirkungen der Infrastruktur auf Bladeausfälle minimiert werden können, ist die Datenverarbeitungshardware innerhalb der Gruppe der von Hardware verursachten Ausfälle die Hauptursache von Bladeausfällen.

Während die Infrastruktur voraussichtlich ungefähr 0,8 Minuten der jährlichen Gesamtausfallzeit im Bezug auf den Durchschnitt für einen einzelnen Blade ausmacht, beträgt die jährliche Gesamtausfallzeit eines einzelnen Blade voraussichtlich ungefähr 14 Minuten im Jahr. Dieses BladeCenter-Beispiel verdeutlicht den ***sehr zuverlässigen Aufbau*** einer gemeinsamen Infrastruktur mit vielen Redundanzen. Dennoch werden Ausfallzeiten infolge von Fehlern bei der Hardwareinfrastruktur oder bei einzelnen Blades nicht vollständig beseitigt.

Größere virtualisierte Standalone-SMP-Server-Umgebung



Wenn auf einem Standalone-Server genügend Prozessor- und Hauptspeicherressourcen vorhanden sind, können darauf mehrere Instanzen eines Betriebssystems aktiv sein. Dies wird normalerweise erreicht, indem eine Codeschicht mit der Bezeichnung „Hypervisor“ verwendet wird, die sich im Systemspeicher befindet und auf den Systemprozessoren ausgeführt wird. Ein Hypervisor verwaltet die Hardwareressourcen direkt und stellt Betriebssystemen eine Standardschnittstelle sowie eine Gruppe wichtiger Services bereit. Der Hypervisor, der sich zwischen einem Betriebssystem und der Hardware befindet, kann die Hardware „virtualisieren“. Dies bedeutet, dass er den Betriebssystemen eine Gruppe von Hardwareressourcen zur Verfügung stellen kann (Prozessoren, Hauptspeicher und E/A-Einheiten). Diese Ressourcen können entweder dediziert sein(zum Beispiel ein Prozessorkern) oder gemeinsam genutzt werden. (Beispiel: Bei einem Kern mit Time-Sharing kann jedem „gemeinsam nutzenden“ Betriebssystem ein Prozentsatz der auf einem Kern verfügbaren Zyklen zugeteilt werden.)

Ein direkter Ansatz zum Erstellen eines Hypervisors besteht darin, ein vorhandenes Betriebssystem (dem bereits bekannt ist, wie Hardwareressourcen zeitlich gesteuert und verwaltet werden) um Code zu erweitern, mit dem zusätzliche Betriebssystemimages gebootet werden können. Dieser Hypervisortyp umfasst allgemein Funktionsmerkmale, die eine feinkörnige Zuordnung von Ressourcen ermöglichen, um Betriebssystemimages (Partitionen) voneinander zu trennen und die Sicherheit und die Isolierung zwischen Images sicherzustellen. Diese Typen von Partitionen heißen normalerweise ***Softwarepartitionen*** oder *Soft Partitions.*

Obwohl diese Struktur allgemein eine Softwareisolierung zwischen Betriebssystemimages (und Anwendungen) bereitstellt, verursacht ein katastrophaler Fehler in der Hypervisorsoftware oder der Infrastrukturhardware den Ausfall aller Softwarepartitionen in einem System. Außerdem kann der Ausfall einer beliebigen IT-Ressource zu einem vollständigen Systemausfall führen.

Ein ausgefeilterer Hypervisoraufbau ähnelt nicht so sehr einem Betriebssystem, sondern ist kleiner und weist daher tendenziell weniger Programmierfehler („Bugs“) auf. Außerdem verwendet er spezialisierte Hardwarezusatzeinrichtungen, um Hardwarefehler besser auf einzelne Partitionen einzugrenzen. Selbst in diesem Fall wird das erreichbare Ausmaß der Isolierung jedoch vom Gesamtaufbau der Hardware beschränkt.

Wie im BladeCenter-Beispiel angegeben wurde, sollten mithilfe von Verfahren für das Hinzufügen redundanter Ressourcen, passiver Systemplatinen usw. die Ausfälle minimiert werden können, die mit der gemeinsamen Infrastrukturhardware im Zusammenhang stehen. Es sollte eine sehr hohe MTBF für die Hardwareinfrastruktur möglich sein.

Dies ist jedoch nicht als Erwartung zu verstehen, dass die MTBF der Anwendungen ebenfalls im selben Bereich liegt, da dieser Wert sich lediglich auf die Infrastruktur bezieht. Alle Ausfälle von Datenverarbeitungskomponenten (Prozessoren, Hauptspeicher usw.) können auch zu Ausfällen mindestens einer Partition führen, je nachdem, welcher Hypervisor und welcher Hardwaresystemaufbau verwendet werden. Störungen dieser Art werden am ehesten durch Anwendungsausfälle verursacht (obwohl Benutzer-, Administrator- und Softwarefehler häufig die wichtigsten Ursachen für Anwendungsausfälle sind).

Mehrere große Standalone-Systeme oder ein physisch partitioniertes Einzelsystem

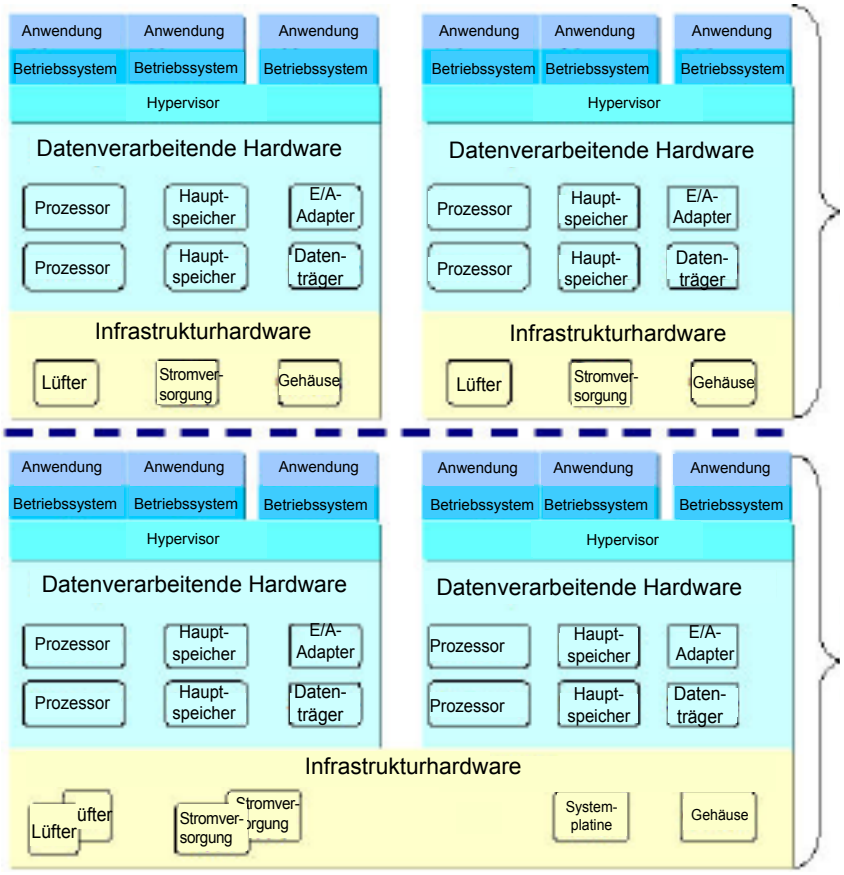
Ein großes Unternehmen mit weitreichenden Anforderungen an die Anwendungsverfügbarkeit implementiert möglicherweise mehrere große Standalone-Systeme, um die erforderlichen IT-Ressourcen bereitzustellen und die erforderliche Verfügbarkeit kritischer Anwendungen zu gewährleisten.

Alternativ dazu kann ein Systemaufbau wie bei einer BladeCenter-Umgebung, bei der zu Verwaltungszwecken mehrere Blades in einem einzigen Gehäuse aktiv sein können, es auch ermöglichen, mehrere große SMP-Server in einem einzigen Gehäuse zu konsolidieren und eine gemeinsame Hardwareinfrastruktur zu nutzen. In einer solchen Umgebung können die einzelnen konsolidierten Hardwareelemente als „physische“ oder „hardwaregebundene“ Partitionen betrachtet werden. Dies gilt analog zu den einzelnen Partitionen, die einer physischen Partition übergeordnet sind und die als logische Partitionen betrachtet werden können.

In beiden Umgebungen gilt ebenfalls: Je zuverlässiger die Datenverarbeitungshardware und die Infrastrukturhardware ist, desto zuverlässiger ist eine beliebige einzelne Anwendung.

Wenn die Datenverarbeitungshardware in einer physischen Partition dieselbe ist wie die Hardware eines Standalone-Servers, ist die Anwendungsverfügbarkeit aufgrund von Hardwarefehlern für beide Konfigurationen gleich. Wenn in der Infrastruktur eines konsolidierten Servers ein Fehler auftritt, fallen zweifellos alle Anwendungen aus. Ein Infrastrukturfehler auf einem partitionierten Server kann jedoch einen größeren Einfluss auf den Betrieb haben, da bei jeder Anwendung auf jeder Partition ein Ausfall auftritt.

Zur Veranschaulichung kann ein Rechenzentrum dienen, das drei Server implementiert, die jeweils eine Infrastruktur-MTBF von fünf Jahren aufweisen. In diesem Beispiel wird ein großer Server in zwei gleiche Partitionen aufgeteilt. Dabei stellt jede Partition die Leistung und Kapazität zweier kleiner Server bereit. Die kleineren Server werden verwendet, um eine einzige geschäftskritische Anwendung auszuführen, während die partitionierten Server eine Anwendung pro Partition ausführen. In einem Zeitraum von fünf Jahren tritt auf jedem dieser Server voraussichtlich ein Ausfall auf. Daher ist innerhalb dieses Zeitrahmens die Gesamtzahl der Anwendungsausfälle (zwei Server gegenüber partitioniertem Server) genau gleich. (Der größere Server fällt einmal aus, wobei zwei Anwendungsausfälle auftreten, und auf den kleineren Servern tritt zusammen ebenfalls je ein Anwendungsausfall auf.) Möglicherweise wirkt sich dies jedoch auf den Betrieb eines Rechenzentrums stärker aus, wenn der größere Server ausfällt und alle kritischen Anwendungen ausfallen.



Implementierung auf mehreren Standalone-Servern gegenüber physischen Partitionen auf einem einzigen großen Server

Einige Hersteller bieten „fest partitionierte" Server an, auf denen Untersegmente des Servers in isolierten, relativ unflexiblen „physischen Partitionen" gruppiert werden. Physische Partitionen sind im Allgemeinen an den Kern und an die Grenzen der Hauptspeicherplatine gebunden. Diese Art der Partitionierung ist nicht so flexibel und kann nicht so gut ausgelastet werden, „verspricht" jedoch eine höhere Verfügbarkeit, da ein Hardwarefehler in der einen Partitionen normalerweise keine Fehler in einer anderen Partition verursacht. Dadurch sieht der Anwender lediglich den Ausfall einer einzigen Partition, jedoch nicht eines vollständigen Systems.

Wenn jedoch ein System über die physische Partitionierung vor allem die weitaus meisten Systeminfrastrukturausfälle beseitigt (beseitigt tatsächlich ***alle*** SPOFs in ***allen*** Komponenten, die mehrere Partitionen miteinander verbinden, und in der gesamten Software, die den Server betreibt) ist dennoch Folgendes möglich: Trotz sehr geringer Rate der Systemabstürze kann eine hohe Ausfallrate für einzelne Partitionen auftreten. Dadurch kann sich trotz physischer Partitionen eine hohe Rate der Anwendungsausfälle ergeben. Obwohl einige Konzepte möglicherweise mit hohen MTBFs für Infrastrukturausfälle werben, erhöhen diese Ausfälle einfach die Gesamtrate der Anwendungsausfälle: Häufig führen weitere Faktoren (wie geplante Ausfallzeiten für die Reparatur von Komponenten) zu Anwendungsausfällen. Viele Kunden zögern, „geschäftskritische" Anwendungen in einer derartigen Umgebung zu implementieren. In vielen Fällen ist es für sie die geeignetere Lösung, einen hoch verfügbaren Server-Cluster zu implementieren, der keine gemeinsam genutzten Komponenten aufweist.

Wenn in einer konsolidierten Umgebung in Frage steht, ob der Ausfall einer einzelnen Anwendung bedeutungslos ist und ob der einzig relevante Ausfall geschieht, wenn jede Anwendung in einem Komplex ausfällt, ist nur die MTBF der Infrastruktur von Bedeutung (neben anderen Fehlern, die ein Gesamtsystem zum Absturz bringen). Bestenfalls kann die konsolidierte Umgebung sich lediglich der MTBF zweier Standalone-Server annähern. Eine 300 Jahre betragende MTBF für eine konsolidierte Systeminfrastruktur erreicht beispielsweise niemals die durchschnittliche Zeit zwischen dem gleichzeitigen Ausfall zweier Standalone-Systeme.

Allgemein sind Anwendungsausfälle relevant und die Anwendungsverfügbarkeit in einem gut aufgebauten System ist vor allem dann umso höher, je verfügbarer die datenverarbeitenden Elemente sind. Dazu gehört nicht nur, wie häufig Datenverarbeitungsausfälle sich auswirken, jedoch auch in welchem Umfang diese in einer einzigen Anwendung isoliert werden können.

Infrastrukturansatz für POWER7-Server

Power Systems teilen Hardware nicht so in mehrere physische Partitionen auf, dass ein einzelnes konsolidiertes System annähernd die Verfügbarkeit aufweist, die durch den Betrieb von zwei, vier oder acht separaten Standalone-Systemen erreicht werden kann. Stattdessen wird davon ausgegangen, dass Unternehmen am meisten von ihren Investitionen profitieren, wenn sie Server in hochgradig virtualisierten Umgebungen implementieren. In diesen Umgebungen können sie dynamisch und flexibel auf Systemressourcen zugreifen und die Serverleistung dadurch nutzen, dass sie die Serverauslastung maximieren. Ziel ist es dabei, ihre Serverinvestitionen rentabel zu machen.

Server müssen schnell und effizient skaliert werden können, die Auslastung automatisch optimieren und Ressourcen je nach Bedarf der jeweiligen Anwendung flexibel umverteilen, Ausfallzeiten minimieren, Energie sparen und Management-Tasks automatisieren. Diese Anforderungen definieren die Merkmale von Servern im Rahmen von Smarter Planet. Außerdem stehen diese Anforderungen hinter einer Strategie für den RAS-Systemaufbau von Power-Systemen, die die Verfügbarkeit unabhängig von der Systemkonfiguration optimieren soll.

Im vorliegenden Dokument wurde relativ ausführlich behandelt, wie der einzigartige Systemaufbau für Zuverlässigkeit und Verfügbarkeit von Power-IT-Komponenten zur Anwendungsverfügbarkeit beiträgt. Außerdem konnten die IBM Entwickler Hardwarefunktionen integrieren, die einen zuverlässigeren Betrieb der Systemfirmware unterstützen, da sie gleichzeitig die Hardware und die Firmware entwickelt haben. Beispielsweise verbessert das Verfahren Active Memory Mirroring for Hypervisor [siehe Seite 30] die Systemverfügbarkeit dadurch, dass ein seltener, jedoch relevanter, potenzieller System-SPOF beseitigt wird. Auch POWER Hypervisor ist eine kritische Komponente zum Erreichen der Virtualisierungsziele und einer besseren Serververfügbarkeit.

Systemaufbau von PowerVM/POWER Hypervisor

Für die innovativen Virtualisierungsverfahren, die bei der POWER-Technologie verfügbar sind, sind leistungsfähige Tools erforderlich, mit denen ein System in mehrere Partitionen aufgeteilt werden kann, auf denen jeweils eine separate Instanz eines separaten Betriebssystemimage ausgeführt wird. Dazu wird Firmware namens POWER Hypervisor eingesetzt. POWER Hypervisor stellt für alle Partitionen die Softwareisolierung und die Sicherheit zur Verfügung. POWER Hypervisor ist eine „dünne“ Codeschicht, eine Abstraktionsebene (Codeschnittstelle), die sich zwischen dem Betriebssystem und der Hardware befindet. POWER Hypervisor ist im Vergleich zu einer vollständigen Betriebssystemimplementierung weniger komplex und kleiner und kann daher unter dem Aspekt des Systemaufbaus und der Qualitätssicherung besser gehandhabt werden.

Der Code von POWER Hypervisor ist auf diskreten Flashspeicherbereichen gespeichert und in allen Systemen aktiv – auch in denjenigen, die lediglich eine einzige Partition enthalten. In vielerlei Hinsicht wird POWER Hypervisor einfach als eine weitere Infrastrukturkomponente behandelt. Da diese Komponente auf Systemhardware ausgeführt wird, hängt deren Zuverlässigkeit teilweise von der Zuverlässigkeit des Kernsystems ab.

PowerVM unterstützt nicht nur feinkörnige Ressourcenzuordnungen, die POWER Hypervisor-Firmware ermöglicht es auch, dass nahezu alle Systemressourcen (Prozessorkerne und -zyklen, Hauptspeichersegmente, E/A-Steckplätze) einer beliebigen Partition in beliebiger Kombination dynamisch zugeordnet werden können (dediziert oder gemeinsam genutzt). Dadurch wird eine außerordentliche hohe Konfigurationsflexibilität erreicht und viele Hochverfügbarkeitsfunktionen werden unterstützt.

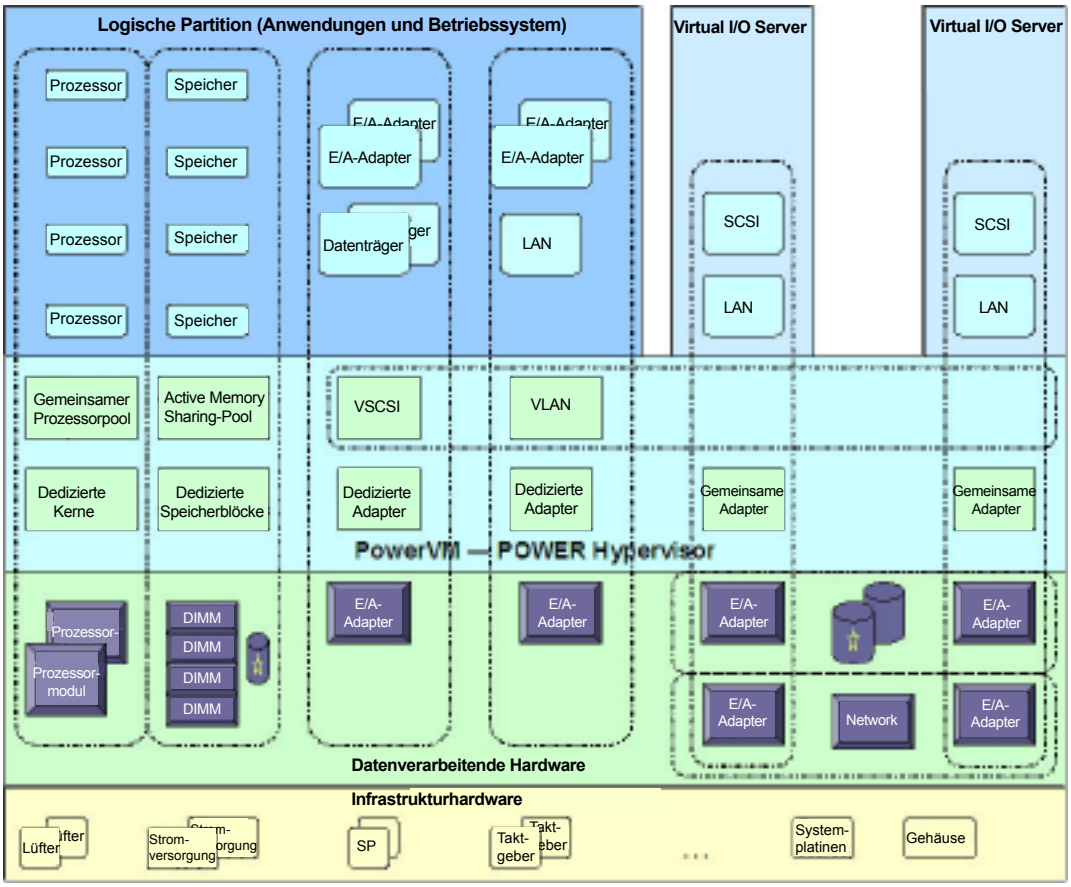
Virtualisierung von Prozessorressourcen

Es sind unterschiedliche PowerVM Editions verfügbar, die Funktionalität wie zum Beispiel die Mikropartitionierung (Micro-Partitioning™), den virtuellen E/A-Server (Virtual I/O Server), gemeinsame Prozessorpools (Multiple Shared Processor Pools) und gemeinsame dedizierte Kapazität (Shared Dedicated Capacity) zur Verfügung stellen. Diese Funktionalität soll für Unternehmen die Systemauslastung erhöhen und zugleich sicherstellen, dass Anwendungen die von ihnen benötigten Ressourcen erhalten.

PowerVM Micro-Partitioning unterstützt eine Reihe logischer Partitionen (virtueller Maschinen):

1. Dediziert: Zugeordnete Prozessorkerne werden ausschließlich vom Betriebssystem verwendet, das in der Partition ausgeführt wird. Kein anderes Betriebssystem kann diese Kerne verwenden.

2. Gemeinsam genutzt: Einem gemeinsamen Prozessorpool zugeordnete Prozessorkerne sind gemeinsame Ressourcen. POWER Hypervisor ordnet den Partitionen im Zeitscheibenverfahren Prozessorzyklen zu, die diese Kerne verwenden. Eine virtuelle Maschine erhält einen Teil der verfügbaren Zyklen (im gesamten Pool). Die Zuteilung basiert auf einer Richtlinie, die von einer HMC (Hardware Management Console) oder vom Integrated Virtualization Manager festgelegt wird. Die Ressourcenzuordnung ist dynamisch und automatisch.
3. Gemeinsam und dediziert (Shared Dedicated): Gemeinsame und dedizierte Kapazität (Shared Dedicated Capacity) ermöglicht die „Spende“ von Zusatz-CPU-Zyklen aus dedizierten Prozessorpartitionen an einen gemeinsamen Prozessorpool. Die dedizierte Partition behält für dedizierte CPU-Zyklen die absolute Priorität. Das Aktivieren dieser Funktion kann dazu beitragen, die Systemauslastung zu verbessern, ohne die Verarbeitungskapazität für kritische Workloads in einem dedizierten Prozessor zu beeinträchtigen.



Vereinfachte Ansicht der Systemvirtualisierung mit PowerVM

Diese Abbildung veranschaulicht grob die unterschiedlichen Hardware- und Softwareschichten, die auf Power 770-, 780- oder 795-Servern mit Unterstützung der Partitionierung und Virtualisierung eines Power-Systems unter Verwendung von POWER Hypervisor verfügbar sind.

In allen Power Systems-Servern werden Benutzeranwendungen in dynamischen logischen Partitionen ausgeführt, wobei die einzelnen Partitionen ihre jeweils eigenen Betriebssysteminstanzen ausführen. Diese Architektur wird als Micro-Partitioning™ bezeichnet und ermöglicht es einzelnen Kernen, bis zu zehn Kopien des Betriebssystems auszuführen. Die Prozessorkerne unterstützen sowohl logische dedizierte Prozessorenpartitionen als auch dynamische LPARs mit gemeinsam genutzten Prozessoren. In einer dedizierten Prozessorpartition ist der Partition mindestens ein physischer Kern zugeordnet. In Partitionen mit gemeinsam genutzten Prozessoren wird ein „gemeinsamer Pool“ physischer Prozessorkerne definiert. Dieser gemeinsame Pool besteht aus mindestens einem physischen Prozessorkern. Für jeden physischen Prozessorkern im Pool können bis zu zehn logische Partitionen definiert werden. Diese Technologie wird als PowerVM bezeichnet.

IBM PowerVM ist die Familie der Power-Technologien, -Funktionen und -Produktangebote, die eine sichere, bewährte Virtualisierung bieten. PowerVM umfasst Basiskomponenten, die mit Power Systems-Firmware ausgestattet sind, sowie optionale Komponenten – PowerVM Editions – zur Erstellung von Pools aus Ressourcen und zur Optimierung ihrer anwendungsumgebungs- und betriebssystemübergreifenden Nutzung mit dem Ziel, die Leistung und die Flexibilität bereitzustellen, um in einem einzigen System mehrere Systemanforderungen zu erfüllen. Andere Ausgaben (Standard und Advanced Edition) bieten eine Reihe von Funktionen wie Micro-Partitioning™, Virtual I/O Server, Multiple Shared Processor Pools und Shared Dedicated Capacity sowie Live Partition Mobility (LPM) und Active Memory Sharing.

Active Memory Expansion (Erweiterung des aktiven Hauptspeichers) ist eine neue POWER7-Technologie, mit der die tatsächliche Hauptspeicherkapazität des Systems viel größer sein kann als der tatsächliche physische Hauptspeicher. Mit der innovativen Komprimierung/Dekomprimierung von Speicherinhalten kann bis zu doppelt so viel Speicher erzielt werden. Active Memory Expansion ist für Partitionen unter AIX ab Version 6.1 verfügbar.

Serverkonfigurationen können darüber hinaus gemeinsame Prozessorpools (Multiple Shared Processor Pools) umfassen. Mit diesen Pools kann ein automatischer unterbrechungsfreier Lastausgleich der Verarbeitungskapazität zwischen Partitionen erreicht werden, die gemeinsamen Pools zugeordnet sind. Dadurch wird ein höherer Durchsatz erreicht. Außerdem besteht die Möglichkeit, die Prozessorkernressourcen, die von einer Gruppe von Partitionen genutzt werden, zu begrenzen, um Kosten für prozessorbasierte Softwarelizenzen zu reduzieren.

Wenn in all diesen Fällen ein Prozessor wegen permanenter Fehler oder wegen einer Reihe von Soft Errors dekonfiguriert (entfernt) werden muss, stehen Mechanismen zur Verfügung, um aktive Programme vom Prozessorkern in verfügbare Zusatzkapazität auf dem Server zu verlagern. Diese Zusatzkapazität kann von einer beliebigen Stelle im System abgerufen werden, ohne zum Beispiel für die einzelnen Partitionen einen Zusatzprozessorkern zu dedizieren. Für die Nutzung gemeinsamer Prozessorpools ist es nicht erforderlich, diese Funktionen in Kooperation mit dem Betriebssystem oder mit einer Anwendung auszuführen [siehe Dynamic Processor Sparing auf Seite 20].

Außerdem ist zu beachten, dass in einem POWER7-System jeder Prozessorkern bis zu vier gleichzeitige Threads mithilfe des simultanen Multithreading ausführen kann. Wenn Multithreading unterstützt wird, werden die einzelnen Threads ebenfalls virtualisiert.

Virtualisierung von Hauptspeicherressourcen

Speicherzuordnung zu virtuellen Maschinen

Hauptspeicher wird Partitionen in Form von logischen Speicherblöcken mit vordefinierter Größe zugeordnet (zum Beispiel als logischer Speicherblock mit 256 MB). Nachdem ein logischer Speicherblock einer Partition zugeordnet wurde, ist er dediziert. Keine andere Partition hat Zugriff auf Daten, die in diesem logischen Speicherblock gespeichert sind.

Alternativ dazu kann mithilfe der Funktion Active Memory Sharing Speicher in einem gemeinsamen Hauptspeicherpool zusammengefasst und dynamisch Partitionen zugeordnet werden. Mit Active Memory Sharing können logische Partitionen einen Pool aus physischem Hauptspeicher auf einem Einzelserver gemeinsam nutzen. Dadurch wird die Speicherauslastung verbessert und die Systemkosten werden reduziert. POWER Hypervisor verwaltet die Beziehung zwischen den logischen Speicherblöcken und der physischen Speicherhardware. Dabei wird ein Virtual I/O Server verwendet, um selten genutzte Speicherblöcke auf DASD-Speicher zu schreiben und das Zuordnen weiterer logischer Speicherblöcke zu ermöglichen. Hauptspeicher wird nach Bedarf dynamisch zwischen den Partitionen zugeordnet, um den physischen Gesamtspeicher im Pool optimal zu nutzen.

Active Memory Expansion für Power 795

Active Memory Expansion (Erweiterung des aktiven Hauptspeichers) ist eine neue POWER7-Technologie, mit der die tatsächliche Hauptspeicherkapazität des Systems viel größer werden kann als der tatsächliche physische Hauptspeicher. Mit der innovativen Komprimierung/Dekomprimierung von Speicherinhalten kann bis zu doppelt so viel Speicherkapazität erzielt werden. Dadurch kann eine Partition im Verhältnis zum Umfang ihres physischen Speichers beträchtlich mehr Arbeit übernehmen oder ein Server mehr Partitionen ausführen und für denselben Umfang an physischem Hauptspeicher mehr Daten verarbeiten. Active Memory Expansion ist für Partitionen unter AIX ab Version 6.1 verfügbar.

Gespiegelter Hauptspeicher als Schutz für POWER Hypervisor

Active Memory Mirroring for Hypervisor [siehe Seite30] wurde konzipiert, um einen Systemausfall selbst dann zu verhindern, wenn im Hauptspeicher ein nicht behebbarer Fehler auftritt.

Aufhebung der Speicherkonfiguration und Zusatzspeicher

Wenn der Serviceprozessor beim Booten einen Hauptspeicherfehler erkennt, wird der betroffene Speicher als fehlerhaft markiert und beim aktuellen IPL oder bei nachfolgenden IPLs nicht mehr verwendet (Memory Persistent Deallocation).

Wenn eine Partition wegen eines nicht behebbaren Fehlers (UE - Uncorrectable Error) abstürzt, belegt POWER Hypervisor beim nächsten Neustart nicht zugeordneten Hauptspeicher (aus einer beliebigen Position im System) einschließlich CoD-Ressourcen (CoD - Capacity on Demand), damit die Partition neu gestartet werden kann. Falls nicht genügend freier Speicher vorhanden ist, um die fehlerhaften LMBs zu ersetzen, startet POWER Hypervisor eine Partition, sofern der verfügbare Speicher dem vom Benutzer angegebenen Minimum für die Partition entspricht.

In beiden Fällen generiert POWER Hypervisor eine Fehlernachricht, die einen Serviceaufruf zum Austauschen des fehlerhaften Speichers auslöst.

Virtualisierung von E/A-Adaptern

Dedizierte Ein-/Ausgabe

POWER Hypervisor kann entweder PCI-x- oder PCIe-E/A-Adapter zu Partitionen zuordnen. Diese Adapter werden ausschließlich von einer einzelnen Partition verwendet und als dedizierte E/A-Adapter bezeichnet. Für Power-Server ist eine große Vielfalt von Adaptern verfügbar, mit denen der Server Verbindungen zu Festplatten, LANs oder Speichereinheiten herstellen kann. Die Adapterverfügbarkeit kann mit Betriebssystemunterstützung verbessert werden: Es können redundante Adapter und verschiedene Wiederherstellungsschemen wie Multipath I/O mit automatischem Failover verwendet werden.

Gemeinsam genutzte (virtuelle) Ein-/Ausgabe

Power-Server unterstützen die gemeinsame Nutzung von E/A-Einheiten. Dazu definieren sie eine spezielle Partition, die als Virtual I/O Server-Partition (VIOS-Partition) bezeichnet wird. Obwohl E/A-Adapter (und die daran angeschlossenen Ressourcen) für eine VIOS-Partition dediziert sind, gehört zu Virtual I/O Server auch Code, mit dem PCI-x- oder PCIe-Adapter von Einheitentreibern gemeinsam genutzt werden können, die auf unterschiedlichen virtuellen Maschinen (logischen Partitionen) aktiv sind. VIOS ermöglicht die gemeinsame Nutzung von Festplatten und optischen Einheiten sowie Datenübertragungs- und Fibre-Channel-Adaptern. Dadurch sollen die Komplexität und die Kosten für Systeme/Verwaltung reduziert werden.

Aus Gründen der besseren Verfügbarkeit können die einzelnen Virtual I/O Server redundante E/A-Adapter mit dynamischer Failover-Fähigkeit nutzen. Der dynamische Failover ist unabhängig von den Betriebssystemen und von Benutzeranwendungen. Diese Konfiguration kann eingesetzt werden, um Ausfallzeiten zu verhindern, die von E/A-Adaptern verursacht werden.

Mit Betriebssystemunterstützung kann eine Partition unter Verwendung redundanter virtueller E/A-Server (Virtual I/O Server) auf dasselbe E/A-Subsystem zugreifen. Durch die Nutzung redundanter E/A-Kapazitäten, die auf implementierten virtuellen E/A-Servern verfügbar sind, können Kundenanwendungen fortlaufend auf E/A-Adapter zugreifen, auch wenn in einer VIOS-Partition ein Ausfall auftritt.

Live Partition Mobility

PowerVM unterstützt die Verlagerung einer aktiven AIX- oder Linux-Partition von einem physischen Server auf einen anderen kompatiblen Server, ohne dass Anwendungen ausfallen. Dadurch werden Anwendungsausfälle für die geplante Systemwartung, die Einrichtung und das Auslastungsmanagement verhindert. Dieses Produktmerkmal heißt Live Partition Mobility (LPM). Es nutzt viele eigene Funktionen der Power-Virtualisierungsarchitektur (virtuelle Prozessoren, VIOS, dynamische Zuordnung von Ressourcen usw.), um zwei Server zu koppeln, die E/A-Ressourcen gemeinsam nutzen. Dadurch ist eine dynamische Verlagerung virtueller Maschinen zwischen Servern ohne Anwendungsausfälle möglich. Live Partition Mobility wird zwischen POWER6-Servern, zwischen POWER7-Servern und zwischen POWER6- und POWER7-Servern unterstützt. Mit LPM kann die Migration von Anwendungen zwischen Servergenerationen (POWER6 zu POWER7) im Betrieb von Rechenzentren unterbrechungsfrei durchgeführt werden: Die Anwendungen werden einfach auf den neuen Server verlagert.

Diese Funktionalität kann verwendet werden, um Workloads systemübergreifend gleichmäßig zu verteilen oder um Energie zu sparen, indem Server bei geringem Bedarf konsolidiert und im Leerlauf ausgeschaltet werden. Darüber hinaus können Leiter von Rechenzentren geplante Ausfallzeiten mithilfe der Funktion Live Partition Mobility von PowerVM Enterprise Edition besser steuern. Workloads können von Server zu Server verlagert werden, wodurch eine Routinewartung zu den am besten geeigneten Zeiten durchgeführt werden kann.

Sonstige Elemente des Systemaufbaus für Zuverlässigkeit/Verfügbarkeit der Infrastruktur

Stromversorgung/Kühlung

Power 770-, Power 780- und Power 795-Server umfassen redundante Stromversorgungskomponenten, Kühlkomponenten, wie Gebläse oder Lüfter, und Ausgaben von Spannungsreglermodulen. Darüber hinaus werden doppelte Wechselstromeingänge unterstützt. Andere Systeme unterstützen Zusatzeinrichtungen für redundante Stromversorgung und Kühlung [siehe Anhang A auf Seite 51].

TPMD

Alle POWER7-Systeme verfügen über mindestens ein TPMD (Thermal Power Management Device). Diese Energiemanagementeinheit überwacht die Wärmebelastung und den Stromverbrauch von Knoten. TPMD ist eine kritische Komponente der Technologie EnergyScale™[8]von IBM. Sie wird von IBM Systems Director Active Energy Manager verwendet, um es Systemadministratoren zu ermöglichen, Richtlinien für den Leistungsausgleich, für die Kühlung und für den Stromverbrauch in einem System festzulegen. Diese Richtlinien schließen die Unterstützung verschiedener Stromsparmodi ein.

Der RAS-Systemaufbau soll einen vollständigen Ausfall einer TPMD verkraften können. Dazu wird ein „abgesicherter" Modus verwendet, in dem der Systembetrieb fortgesetzt werden kann. Die Power 795-Server enthalten redundante TPMD-Module, die die Wahrscheinlichkeit eines Ausfalls von Energiemanagementfunktionen aufgrund von TPMD-Ausfällen noch weiter reduzieren.

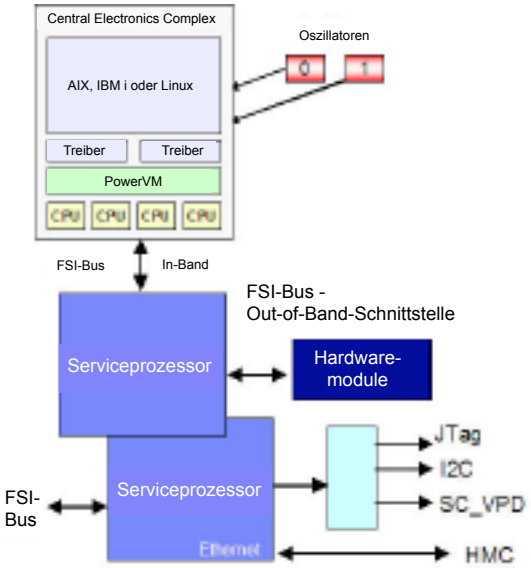
Taktgeber

Power 795-Systeme und Power 770- und 780-Mehrknotensysteme stellen ***redundante Prozessorbezugstaktgeber*** für jeden Prozessor bereit, die den dynamischen Failover des Taktgebers unterstützen (ohne eine Systemausfallzeit zu verursachen), wenn ein Ausfall in den Taktgeberkomponenten erkannt wird. Da der Wechsel innerhalb der Prozessormodule erfolgt (ohne externes Umschaltelement), wird der Bezugstaktgeber als Datenverarbeitungskomponente, nicht als Infrastrukturelement, betrachtet.

Serviceprozessoren

POWER7-Systeme benötigen einen einzigen Serviceprozessor, um das Einschalten des Systems und Fehlerereignisse während des Betriebs überwachen zu können. Der Serviceprozessor ist ein separat mit Strom versorgter Mikroprozessor, der von der Hauptelektronik für die Instruktionsverarbeitung getrennt ist. Der Serviceprozessor unterstützt die Überwachung mit POWER Hypervisor und mit der HMC (Hardware Management Console), eine selektive remote angebundene

Taktgeber- und Serviceprozessoranschlüsse



Der Serviceprozessor ist ein separat mit Strom versorgter Mikroprozessor, der von der Hauptelektronik für die Instruktionsverarbeitung getrennt ist. Der Serviceprozessor unterstützt die Überwachung mit POWER Hypervisor und mit der HMC (Hardware Management Console), eine selektive remote angebundene Stromversorgungssteuerung, Umgebungsüberwachung, Funktionen zum Zurücksetzen und Booten, Fernwartung und Diagnoseaktivitäten sowie Konsolspiegelung.

Die einzelnen Serviceprozessoren kommunizieren mit POWER Hypervisor mithilfe einer PSI-Verbindung (PSI - Processor Service Interface) zu einem POWER7-Prozessor. Dieser Bus ist ECC-geschützt und viele Serviceprozessoren in einem System können mit mehreren POWER7-Prozessoren über redundante PSI-Busse kommunizieren.

Server mit redundanten Serviceprozessoren oder Taktgebern unterstützen den dynamischen Failover von der primären Einheit zur Ausweicheinheit, ohne den normalen Systembetrieb zu unterbrechen.

[8] Martha Broyles, Chris Francois, Andrew Geissler, Michael Hollinger, Todd Rosedahl, Guillermo Silva, Jeff Van Heuklon, Brian Veale, „IBM EnergyScale for POWER7 Processor-Based Systems", http://www-03.ibm.com/systems/power/hardware/whitepapers/energyscale7.html, POW03039-USEN-03, August 2010.

Stromversorgungssteuerung, Umgebungsüberwachung, Funktionen zum Zurücksetzen und Booten, Fernwartung und Diagnoseaktivitäten sowie Konsolspiegelung. In Systemen ohne Hardware Management Console kann der Serviceprozessor Aufrufe absetzen, um Überwachungsfehler bei POWER Hypervisor, kritische umgebungsbedingte Fehler und kritische Verarbeitungsfehler zu melden, selbst wenn die Hauptverarbeitungsarbeit nicht funktionsfähig ist. Der Serviceprozessor stellt Services bereit, die für zeitgemäße Computer üblich sind: zum Beispiel die Umgebungsüberwachung, die gegenseitige Überwachung und eine Reihe von Verfügbarkeitsfunktionen.

Wenn bei einem Serviceprozessor ein transienter Laufzeitfehler in der Hardware oder in der Firmware auftritt, kann er automatisch neu gebootet werden, während das System aktiv bleibt. Transiente Fehler beim Serviceprozessor wirken sich nicht auf Serveranwendungen aus. Wenn der Serviceprozessor die Bedingung einer „Codeblockierung" findet, kann POWER Hypervisor den Fehler erkennen und den Serviceprozessor anweisen, neu zu booten, um einen weiteren Ausfall zu verhindern. Dies wird als Fähigkeit zum ***reset/reload (***zurücksetzen/neu laden) bezeichnet.

Damit eine verbesserte Verfügbarkeit unterstützt wird, werden separate Kopien des Serviceprozessor-Microcodes und des POWER Hypervisor-Codes in diskreten Flashspeicherbereichen gespeichert. Der Codezugriff ist CRC-geschützt. Außerdem unterstützen die Server dynamische Firmware-Updates, bei denen Anwendungen funktionsfähig bleiben, während IBM Systemfirmware für bestimmte Funktionalitäten aktualisiert wird. Dadurch dass zwei Kopien gespeichert sind, kann ein neues Firmware-Image in den Serviceprozessor geladen werden, während dieser weiterhin mit dem älteren Image betrieben wird. Außerdem wird dadurch sichergestellt, dass der Serviceprozessor selbst dann aktiv sein kann, wenn eine Flashspeicherkopie beschädigt wird.

Power 770- und 780-Systeme stellen redundante Serviceprozessoren in Mehrknotensystemen mit der Fähigkeit bereit, bei permanenten Serviceprozessorfehlern, einen dynamischen Failover vom primären Serviceprozessor zu einem sekundären Serviceprozessor durchzuführen.

In einem Power 795-Server wird die Serviceprozessorfunktion zwischen Systemcontrollern und Knotencontrollern verteilt. Die Systemcontroller (ein aktiver Controller und ein Ausweichcontroller) fungieren als Überwachungsinstanz, die einen zentralen Steuerungspunkt bereitstellen und den Hauptteil der herkömmlichen Serviceprozessorfunktionen ausführen. Die Knotencontroller, von denen pro Knoten einer aktiv und einer ein Ausweichcontroller ist, haben Servicezugriff auf die Knotenhardware. Alle Befehle vom primären Systemcontroller werden sowohl an den primären als auch an den redundanten Knotencontroller weitergeleitet. Falls ein primärer Knotencontroller ausfällt, übernimmt der redundante Controller automatisch alle Aufgaben des primären.

Systemcontroller kommunizieren über redundante LANs, die Verbindungen zwischen den Stromversorgungscontrollern (in den Hauptnetzteilen), den Knotencontrollern und einer HMC oder zwei HMCs herstellen. Dieser Systemaufbau ermöglicht den automatischen Failover und einen durchgehenden Serverbetrieb, falls bei einer bestimmten Komponente ein Fehler auftritt.

Firmware-Updates

Aufgrund des robusten Konzepts für die Verwaltung von Hardwarefehlern im Serviceprozessor ist es wichtig, dass Defekte in der Serviceprozessorfirmware, in POWER Hypervisor und in sonstiger Systemfirmware behoben werden können, ohne für die Programmkorrektur eine Ausfallzeit zu verursachen.

Die Firmwarestruktur für den Serviceprozessor, für das Stromversorgungssubsystem und für POWER Hypervisor unterstützt eine Strategie für ***gleichzeitig ablaufende Programmkorrektur,*** die es Administratoren erlaubt, viele Firmware-Updates zu installieren und zu aktivieren, ohne die Stromversorgung des Systems aus- und wieder einzuschalten oder den Server neu zu starten. Da Änderungen an einigen Serverfunktionen (zum Beispiel die Änderung der Initialisierungswerte für Chipsteuerungen) nicht während des Betriebs stattfinden können, erfordert eine Programmkorrektur in diesem Bereich zur Aktivierung einen Neustart des Systems.

Das Aktivieren neuer Firmwarefunktionen setzt die Installation eines Firmware-Release-Level voraus. Dieser Prozess unterbricht den Serverbetrieb und setzt eine geplante Ausfallzeit sowie einen vollständigen Neustart des Servers voraus.

Hardware Management Console

Die Hardware Management Console (HMC) unterstützt die IBM Virtualisierungsstrategie und schließt zahlreiche Verbesserungen für den Service und für die Unterstützung ein: automatische Installation und automatisches Upgrade sowie Parallelwartung und Upgrade für Hardware und Firmware. Die HMC stellt darüber hinaus einen zentralen Punkt für den Empfang von Services, die Protokollierung, die Verfolgung von Systemfehlern und (falls aktiviert) die Weiterleitung von Problemberichten

an den IBM Service und an den IBM Support bereit. Obwohl die HMC für einige Konfigurationen ein optionales Angebot ist, kann mit ihr ein beliebiger Server[9] in der Produktfamilie der IBM Power-Server unterstützt werden.

Hardware Management Consoles sind separate Einheiten, die über ein Ethernet-Netz an ein System angeschlossen werden. Power-Server unterstützen aus Gründen der Redundanz Verbindungen zu zwei HMCs[9].

Über die Hardware hinausgehende Verfügbarkeit und Virtualisierung

Der Schwerpunkt des vorliegenden Dokuments ist die Beschreibung von RAS-Merkmalen in der Power-Hardware, die für die Verfügbarkeit der Systemhardware selbst sorgen. Betriebssysteme, Middleware und Anwendungen erweitern dies um zusätzliche Schlüsselfunktionen, die ihre eigene Verfügbarkeit betreffen. Diese Verfügbarkeit ist nicht Thema der vorliegenden Hardwarebeschreibung.

Es lohnt sich jedoch, zu beachten, dass Hardware- und Firmware-RAS-Funktionen eine wichtige Unterstützung für ausgewählte Funktionen der Softwareverfügbarkeit bieten können.

Funktionsmerkmale von Betriebssystemen und Anwendungen

Es gibt viele in Betriebssysteme und Anwendungen integrierte Fehlererkennungs-/Fehlereingrenzungs- und Verfügbarkeitsfunktionen, mit denen im Code Fehler erkannt und diagnostiziert werden sollen, welche keine Beziehung zu zugrunde liegenden Hardwarefehlern haben.

Obwohl die meisten dieser Funktionen hardwareunabhängig sind, nutzen manche von ihnen Funktionen, die mit Hardware oder Firmware in Beziehung stehen. Beispielsweise sammelt AIX bei der Beendigung einer Partition in einem sogenannten „Speicherauszug“ Daten, die mit dem Problem zusammenhängen, das die Beendigung verursacht hat. POWER Hypervisor trägt zum Sammeln dieser Daten bei.

Storage Protection Keys

Unbeabsichtigte Hauptspeicherüberlagerungen – bei denen Code versehentlich Speicher überschreibt, dessen Eigner anderer Code ist – sind eine verbreitete Ursache für Anwendungs- und Betriebssystemausfälle.

Ein Betriebssystem kann gemäß seinem Systemaufbau verhindern, dass eine Anwendung direkt den Code oder Daten einer anderen Anwendung überschreibt. Voraussetzung dafür ist es, dass die Anwendung keinen Hauptspeicher gemeinsam nutzt. Es besteht jedoch die Möglichkeit, dass ein Teil einer Anwendung Hauptspeicher eines nicht zu ihr gehörigen Teils dieser Anwendung überlagert.

Power Storage Protection Keys (Power-Speicherschutzschlüssel) bieten einen hardwaregestützten Zugriffsmechanismus für Hauptspeicherbereiche. Nur Programme, die den richtigen Schlüssel verwenden, dürfen Daten aus geschützten Speicherpositionen auslesen oder in diese schreiben. Unter AIX[10] kann jede virtuelle Speicherseite einem Schlüssel zugeordnet werden, der die Art des Zugriffs, zu der die Seite berechtigt ist (zum Beispiel nur Lesen oder Lesen und Schreiben) nach Codesegment definiert. POWER7-Server ermöglichen eine direkte Zuordnung von virtuellen Schlüsseln zu Hardwareschlüsseln. Dadurch kann die Prozessorhardware selbst einen Schutz vor unberechtigten Zugriffen bieten.

Diese neuartige Hardware ermöglicht es Programmierern, den Speicherzugriff innerhalb klar strukturierter, hardwaregestützter Grenzen zu beschränken. Dadurch werden kritische Teile von AIX[10] und von Anwendungssoftware vor unbeabsichtigten Speicherüberlagerungen geschützt. Speicherschutzschlüssel können die Anzahl der sporadisch auftretenden Ausfälle im Zusammenhang mit nicht erkannten Speicherüberlagerungen innerhalb des AIX-Kernels reduzieren. Außerdem können Programmierer die Power-Funktion für Speicherschutzschlüssel verwenden, um die Zuverlässigkeit großer, komplexer Anwendungen zu steigern, die unter AIX ab Version 5.3 ausgeführt werden.

Hochverfügbarkeitslösungen

Im gesamten White Paper wurde ein Hauptaugenmerk auf die Anwendungsverfügbarkeit innerhalb der Umgebung eines Einzelsystems gelegt. Obwohl die Wahrscheinlichkeit von Anwendungsausfällen aufgrund von Hardwarefehlern in dieser Umgebung verringert werden kann, können sie nicht vollständig beseitigt werden. Ausfälle können jedoch auch durch Fehler im Software-Stack verursacht werden. Im vorliegenden Dokument wurden viele Verfügbarkeitsfunktionen behandelt, die bei Bedarf automatisch aufgerufen werden.

Eine geeignete Planung der Serverkonfigurationen kann dazu beitragen, die Systemverfügbarkeit zu maximieren. Sie können die gesamte Anwendungsverfügbarkeit verbessern, indem Sie E/A-Einheiten für die Redundanz ordnungsgemäß konfigurieren und Partitionsdefinitionen erstellen, bei denen der Verlust von Kern- oder Hauptspeicherressourcen nicht zum Systemausfall führt.

[9] Außer Power-Blades 10 AIX ab Version 6.1
[10] AIX ab Version 6.1

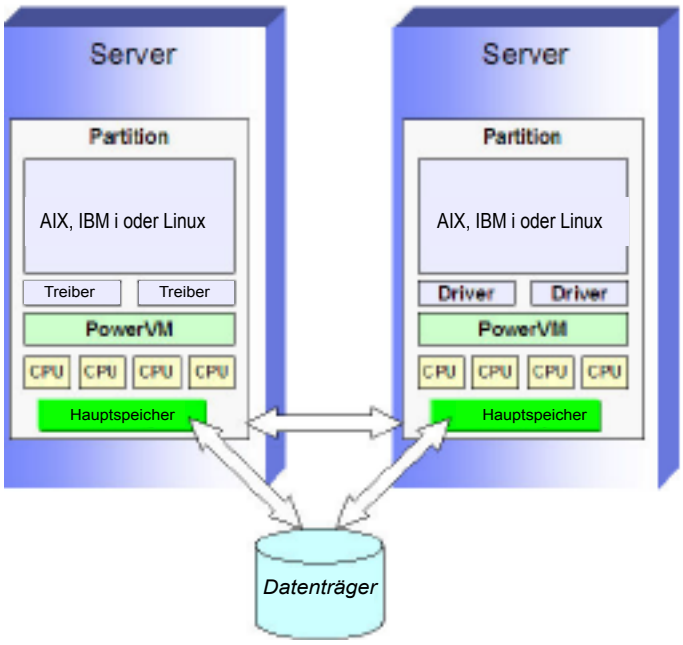
Wenn die höchste Ebene der Anwendungsverfügbarkeit und der Fehlerbehebung erreicht werden soll, ist es nützlich, über Mechanismen zu verfügen, die ausgefallene Anwendungen schnell wiederherstellen: Idealerweise erfolgt dies auf anderen Hardware- und Softwareimages als denjenigen, die beim Auftreten des Fehlers im Gebrauch waren. In der Regel wird dies mit Clustering-Lösungen erreicht, die eine vollständige Redundanz von Hardwaresystemen (und häufig auch Software-Stacks) ermöglichen. Bei diesen Konzepten ist eine Methode für die Fehlerüberwachung integriert, sodass bei einem Systemausfall ein automatisierter Neustart von Anwendungen aufgerufen werden kann.

Für die Verfügbarkeit mit vollständiger Redundanz stellen IBM und andere Softwareanbieter einige Hochverfügbarkeitsclusterlösungen bereit, zum Beispiel IBM PowerHA SystemMirror. Diese Lösung bietet viele Optionen für den Failover von Partitionen. Ein weiteres Beispiel ist IBM DB2 Purescale, ein Produkt, das eine hohe Verfügbarkeit für die Datenbank IBM DB2 unterstützt.

Verschiedene Clustering-Methoden, die redundante Hardware, Firmware und Anwendungen einbeziehen, sind ebenfalls vorhanden oder wurden angeboten. In allen Fällen ist jedoch die Zuverlässigkeit von auf Unternehmen abgestimmten Einzelservern die Grundlage für eine tatsächlich zuverlässige IT-Lösung.

1. Ein offensichtlicher Grund dafür ist, dass die Implementierung komplexer Anwendungen mit Redundanz mehrerer Systeme eine kostenintensive Alternative sein kann und möglicherweise kein optimales Mittel darstellt, um die gewünschte Verfügbarkeitsebene zu erreichen.
2. Einige Clustering-Lösungen stellen möglicherweise eine geringe Anwendungsleistung bereit und häufig ist für Failover-Verfahren eine Unterbrechung des Systembetriebs erforderlich. Schlecht geplante Konfigurationen weisen möglicherweise

Cluster aus zwei Systemen für eine höhere Verfügbarkeit



Wenn die höchste Ebene der Anwendungsverfügbarkeit und der Fehlerbehebung erreicht werden soll, ist es nützlich, über Mechanismen zu verfügen, die ausgefallene Anwendungen schnell wiederherstellen: Dies erfolgt auf anderen Hardware- und Softwareimages als denjenigen, die beim Auftreten des Fehlers im Gebrauch waren. In der Regel wird dies mit Clustering-Lösungen erreicht, die eine vollständige Redundanz von Hardwaresystemen (und häufig auch Software-Stacks) ermöglichen.

Für die Verfügbarkeit mit vollständiger Redundanz stellen IBM und andere Softwareanbieter einige Hochverfügbarkeitsclusterlösungen bereit, zum Beispiel IBM PowerHA SystemMirror. Diese Lösung bietet viele Optionen für den Failover von Partitionen. Ein weiteres Beispiel ist IBM DB2 Purescale mit seiner Unterstützung der hohen Verfügbarkeit für die Datenbank IBM DB2.

Verfügbarkeit über Clustering

Clustering bedeutet, dass bei einem Anwendungsausfall in einer Partition auf einem Server die Anwendung auf einem anderen Server neu gestartet und fortgesetzt wird.

Die Fehlererkennung einer primären Partition kann auf viele Arten erfolgen:

- Heartbeat
- Aktive Signalisierung von Fehlern, …

Das sekundäre System kann Folgendes sein:

- „Hot Standby“ (Bereitschaftsmodus)
- System, das andere Daten verarbeitet, jedoch bei Bedarf Kapazität für zusätzliche Workload übernimmt …

Die Konnektivität kann Folgendes sein:

- Einfach die Übergabe von Heartbeats
- Gemeinsam genutzte einzelne DASD-Images
- Remote angebundene DASD-Spiegelung, …

Speicher kann sich im DASD oder im Hauptspeicher befinden:

- Von den einzelnen Servern gespiegelt
- Oder von den einzelnen Servern separat gespeichert, kopiert und gemeinsam genutzt

Anwendungen müssen möglicherweise den Failover und den Neustart unterstützen:

- Anpassung der Anwendungsumgebung normalerweise erforderlich
- Einige Anwendungen eventuell mit integrierter Redundanzunterstützung (z. B. Datenbanken)

Normalerweise für Softwarefehler und für erkannte Hardwarefehler geeignet:

- Nicht erkannte Fehler auf einem primären System werden möglicherweise nicht mit diesem Clustering-Ansatz behoben.

Die Kosten der Lösung hängen von der Methode und von der auf dem sekundären Server gewünschten Verfügbarkeit ab:

- Ein Einzelserver ist möglicherweise dazu in der Lage, als sekundärer Server für mehrere primäre Server zu fungieren.

kritische Verzögerungen bei der Fehlererkennung sowie beim aufgerufenen Failover auf und der nachfolgende Anwendungsneustart dauert möglicherweise eine beträchtliche Zeit. Allerdings ist es möglich, dass ein Systemaufbau aus einem Einzelserver einen Fehler erkennt, eine Ausfallzeit beansprucht, die ausgefallene Komponente dekonfiguriert und die Anwendungsverfügbarkeit schneller wiederherstellt als eine Clustering-Lösung.

3. Selbst in einer Umgebung, die Redundanz bereitstellt, ist es wichtig, die ausgefallene Hardware nach einem Failover vollständig wiederherzustellen. Dies kann nur Erfolg haben, wenn bei der Fehlererkennung und -eingrenzung eine richtige Investitionsentscheidung getroffen wird.
4. Erfahrungsgemäß funktionieren Failover-Verfahren nur ordnungsgemäß, wenn sie bei der Implementierung auf geeignete Weise getestet werden und wenn sie regelmäßig erneut getestet werden. Häufig schrecken Unternehmen davor zurück, in Produktionssystemen eine Ausfallzeit zu verursachen, um sich für ein zukünftiges Katastrophenszenario vorzubereiten.

Außerdem ist es wichtig, zu beachten, dass Failover-Verfahren nur zuverlässig sind, wenn ein Fehler richtig ermittelt wird. Zuverlässige Hardware mit einer hohen Qualität der Fehlererkennung und -eingrenzung ist entscheidend, um Fehler so zu bestimmen und einzugrenzen, dass der Failover den Ausfall ordnungsgemäß behebt, ohne falsche Ergebnisse zu verbreiten.

Smarter Planet – instrumentiert, miteinander verbunden und intelligent

Das Mantra der Elektronikbranche in den vergangenen Jahrzehnten lautete: „kompakter, schneller, dichter, kostengünstiger". Als Ergebnis dieser Entwicklung gibt es in diesem Jahr weltweit pro Kopf schätzungsweise über eine Milliarde Transistoren. Jeder Transistor kostet weniger als ein Millionstel eines US-Pennys. Mit dieser breit verfügbaren Technologie kann nahezu alles instrumentiert werden. Dadurch können entscheidende Informationen zur Umgebung, zur Verpackung und zum Zustand von Artikeln, zum Nutzungsverhalten, zu warmen Aufwinden, zum Verkehrsfluss oder zu beliebigen anderen Faktoren erfasst werden. Millionen intelligenter Einheiten werden implementiert und miteinander verbunden. Das Internet brachte eine Umgestaltung der Kommunikation und der Interaktion der Menschen untereinander und mit der Umwelt mit sich. Einzelpersonen und Unternehmen schaffen mit ihren Transaktionen über viele Kanäle, Online-Communitys, Registrierungen und Ortsänderungen Millionen neuer digitaler Spuren und daraus entstehen buchstäblich Gebirge aus Daten. Diese Daten enthalten Perlen der Intelligenz: Informationen, die gefiltert werden können, um Einblicke zu erhalten, die einen globalen Fortschritt bedeuten können.

Unternehmen gehen mit mehr Informationen um als je zuvor. Diese gewaltige Explosion der Datenmenge bringt jedoch Datenmüll, Ungenauigkeit und verpasste Chancen mit sich. Unternehmen gehen dieses Problem an, indem sie intelligentere IT-Modelle mit neuen, flexiblen Infrastrukturen implementieren, die große, hoch skalierbare, mit Lastausgleich ausgestattete Computersysteme mit offener, dynamischer Software koppeln. Dieser Übergang zu „intelligenteren" Lösungen stellt neue Anforderungen an die Entwickler von Servern. Heute müssen Server schnell und effizient skaliert werden können. Zugleich müssen sie die Auslastung automatisch optimieren und Ressourcen je nach Bedarf der Anwendungen flexibel neu zuordnen. Dies alles soll zudem mit einem hohen Grad der Verfügbarkeit von Anwendungen stattfinden.

Als Antwort auf diese Anforderungen hat IBM die POWER7-Server mit ihren ausgezeichneten Leistungsmerkmalen entwickelt, damit drei Modi der Datenverarbeitung auf elegante Weise umgesetzt werden: die massive Parallelverarbeitung, die Datenverarbeitung mit hohen Durchsätzen sowie Analysefähigkeiten. Diese Server mit Lastausgleich setzen dynamische Virtualisierungsfunktionalität ein, die eine bessere Nutzung der Rechenleistung ermöglicht (über eine hohe Auslastung und über die gemeinsame Nutzung von Hauptspeicher und E/A-Ressourcen). In Power Systems ist die Hardware so mit Systemsoftware verbunden, dass für viele virtuelle Maschinen, die eine große Vielfalt von Anwendungen unterstützen, eine optimierte Leistung bereitgestellt wird. In dieser Umgebung stellen die Systemzuverlässigkeit und -verfügbarkeit grundlegende Anforderungen an den Systemaufbau dar. Server müssen für Zuverlässigkeit, Verfügbarkeit und Wartungsfreundlichkeit ausgelegt sein – und das gilt nicht nur für die Infrastruktur, sondern auch für die datenverarbeitenden Elemente.

Alle Power Systems-Server folgen einer auf der Architektur basierenden Strategie, die darauf ausgelegt ist, geplante und ungeplante Ausfallzeiten zu verhindern. Diese Server umfassen eine breite Vielfalt von Funktionen, um fehlerhafte Komponenten automatisch zu analysieren, zu identifizieren und einzugrenzen, sodass Reparaturen so schnell und so effizient wie möglich durchgeführt werden können. Die Entwickler des Systemaufbaus integrieren technologisch ausgereifte Komponenten und intelligente Verfahren zum Zusammenstellen der Komponenten, wählen Teile mit geringen bauartbedingten Fehlerraten aus und bündeln diese mit einem Serverpaket, mit dem diese Komponenten zuverlässig betrieben werden können. Es wurde sorgfältig darauf geachtet, robuste und zuverlässige Verbindungen bereitzustellen und Zusatzeinrichtungen aufzunehmen, die den Service erleichtern: zum Beispiel Kartenführungen, PCI-Adapterträger, Kabelhaltebänder und „selbstsichernde" Anschlüsse. Falls tatsächlich ein Hardwarefehler auftritt, wurden diese Server so konzipiert, dass sie ausfallsicher sind und trotz des Fehlers weiter in Betrieb bleiben. Jeder POWER7-Server enthält intelligente Verfügbarkeitsfunktionalität: Processor Instruction Retry (PIR), Alternate Processor Recovery (APR), Dynamic Processor Deallocation, Behebung von PCI-Busfehlern, Chipkill-Speicher, L2- und L3-Cache-Line-Delete, dynamisches Firmware-Update, redundante Hot-Plug-Kühlungslüfter sowie Hot-Plug-N+1-Stromversorgung und -Netzkabel (in einigen Konfigurationen optional).

Viele dieser Funktionen basieren auf der IBM Technologie FFDC (First Failure Data Capture), die es dem Server ermöglicht, Hardwarefehler bei ihrem ersten Auftreten effizient zu erfassen, zu diagnostizieren und darauf zu reagieren.

Diese Verfügbarkeitsverfahren werden durch Serviceangebote gestützt: zum Beispiel durch die automatisierte Installation, das automatisierte Upgrade und die automatisierte Wartung, die von IBM Kundendienstmitarbeitern oder IBM Kunden (für ausgewählte Modelle) angewendet werden können und mit denen Servicekräfte beider Organisationen neue Systeme oder Funktionsmerkmale installieren und Fehler in diesen Systemen effektiv diagnostizieren und beheben können. Das Herzstück jedes POWER7-Servers ist POWER Hypervisor. Dieser stellt nicht nur eine feinkörnige Zuordnung von Systemressourcen bereit, die intelligente Virtualisierungsfunktionalität unterstützen, sondern bietet darüber hinaus viele Verbesserungen bei der Verfügbarkeit. POWER Hypervisor unterstützt Folgendes: den Ersatz von Ressourcen (CPU und Hauptspeicherpools), die automatische Weitergabe von Kapazität in N+1-Konfigurationen, redundante LPAR-konfigurationenübergreifende Ein-/Ausgabe, die Fähigkeit, ein System während des Betriebs neu zu konfigurieren, die automatische Skalierung hoch verfügbarer Ausweichserver, die serialisierte gemeinsame Nutzung von Einheiten, die gemeinsame Nutzung von E/A-Einheiten über E/A-Serverpartitionen sowie das Verlagern „aktiver“ Partitionen zwischen Power-Servern.

Power Systems profitieren von der langen Geschichte ihrer Vorgängersysteme und der Integration vieler intelligenter Verfahren, die in IBM Mainframes erstmals eingesetzt wurden. Sie sind dafür konzipiert, Ihnen eine zukunftsweisende Zuverlässigkeit, Verfügbarkeit und Wartungsfreundlichkeit zur Verfügung zu stellen.

Anhang A: Systemunterstützung für ausgewählte RAS-Funktionen [● = verfügbar, ○ = optional]

RAS-Funktion	BladeCenter PS700, 701, 702 Express	Power 710 und 730 Express	Power 720 und 740 Express	Power 750 Express und 755	Power 770 und 780	Power 795
Prozessor						
Prozessor-Fabric-Bus-Schutz	●	●	●	●	●	●
Dynamic Processor Deallocation	●	●	●	●	●	●
Dynamic Processor Sparing						
◆ Einsatz von CoD-Kernen					○	○
◆ Nutzung von Kapazität aus Zusatzpools	○	○	○	○	○	○
Fehlerbehebung beim Prozessorkern						
Processor Instruction Retry (PIR)	●	●	●	●	●	●
◆ Alternate Processor Recovery (APR)	●	●	●	●	●	●
◆ Core-Contained-Checkstop-Fehler für Partition	●	●	●	●[1]	●	●
Persistente Aufhebung der Prozessorzuordnung	●	●	●	●	●	●
E/A-Subsystem						
Persistente Aufhebung der GX+-Buszuordnung[2]		○	○	○	○	○
Optionaler ECC-E/A-Hub mit Einfrierverhalten (Freeze)		○	○	○	○	○
Erweiterte Fehlererkennung für PCI-Bus	●	●	●	●	●	●
Erweiterte Fehlerbehebung für PCI-Bus	●	●	●	●	●	●
Erweiterte Fehlerbehandlung (EEH - Enhanced Error Handling) für PCI-PCI-Bridge		●	●	●	●	●
Redundante 12-fach-Kanalverbindung			○[3]	○[4]	○	○
Taktgeber und Serviceprozessor (SP)						
Dynamischer SP-Failover während des Betriebs					●[5]	●[6]
Failover für Taktgeber während des Betriebs					●[15]	●
Verfügbarkeit des Hauptspeichers						
ECC-Speicher, L2-Cache, L3-Cache	●	●	●	●	●	●
Fehlererkennung/-behebung						
◆ Chipkill-Speicher mit zusätzlicher Halbbytekorrektur	●	●	●	●	●	●
◆ DRAM-Ersatzspeicherfunktion					●	●
Ersatzspeicherfunktion mit CoD bei IPL					○	○
CRC (Cyclic Redundancy Check) mit Wiederholung auf dem Hauptspeicherdatenbus (CPU-to-Buffer)	●	●	●	●[11]	●	●
ECC für Datenbus (Hauptspeicherpuffer zu DRAM) mit Wiederholung	●	●	●	●	●	●
Dynamische Speicherkanalreparatur	●	●	●	●	●	●
Speichertest (Memory Scrubbing) für Prozessorspeicher-controller	●	●	●	●	●	●
Aufhebung der Zuordnung von Hauptspeicherseiten	●	●	●	●	●	●
L1-Paritätsprüfung mit „retry“/„set delete“	●	●	●	●	●	●
L2-Cache-Line-Delete (L2-Cachezeilenlöschung)	●	●	●	●	●	●
L3-Cache-Line-Delete (L3-Cachezeilenlöschung)	●	●	●	●	●	●
Special Uncorrectable Error Handling (spezielle Behandlung nicht behebbarer Fehler)	●	●	●	●	●	●
Active Memory Mirroring for Hypervisor						○[7]
Fehlererkennung und -eingrenzung						
FFDC für Fehlererkennung und -eingrenzung	●	●	●	●	●	●
Storage Protection Keys	●	●	●	●	●	●
Fehlerprotokollanalyse	●	●	●	●	●	●

[1] Die Verfügbarkeit kann von Systemprozessoren und Firmware-Level abhängen.
[2] Mehrere GX-Busse
[3] Nur Power 740
[4] Nur Power 750 mit mindestens zwei Stecksockeln
[5] Erfordert mehrere Knoten.
[6] Auch: redundante Knotencontroller mit dynamischem Failover während des Betriebs
[7] Verfügbarkeitstermine siehe IBM United States Hardware Announcement 110-158 vom 17. August 2010

RAS-Funktion	BladeCenter PS700, 701, 702 Express	Power 710 und 730 Express	Power 720 und 740 Express	Power 750 Express und 755	Power 770 und 780	Power 795
Wartungsfreundlichkeit						
Fortschrittsanzeiger beim Booten	●	●	●	●	●	●
Firmware-Fehler-Codes	●	●	●	●	●	●
Betriebssystemfehlercodes	●	●	●	●	●	●
Bestandserfassung	●	●	●	●	●	●
Umgebungs- und Netzüberwachung	●	●	●	●	●	●
PCI-Kartenaustausch während des Betriebs (Hot Swap)			○[8]	○[19]	●[9]	○[18]
Austausch von DASD-Einheiten/Datenträgern während des Betriebs (Hot Swap)	●[10]	●	●	●	●	●[11]
Zweifache Plattencontroller/geteilte Rückwandplatine (Split Backplane)			○	○	●	○[18]
Erweiterte Fehlerdatensammlung	●	●	●	●[1]	●	●
SP-„Call-Home-Funktion“ in Nicht-HMC-Konfigurationen		●	●	●	NZ[12]	NZ
Redundante Anschlüsse für E/A-Einschübe			●	●	●	●
Einbau von E/A-Einschüben während des Betriebs und Reparatur bei eingeschalteter Einheit			●	●	●	●
GX-Adaptereinbau während des Betriebs und Reparatur bei eingeschalteter Einheit				●	●	●[13]
Gleichzeitig ablaufender Einbau eines E/A-Gehäuses mit Stromversorgung			●	●	●	●
Gegenseitige SP-Überwachung mit POWER Hypervisor	●	●	●	●	●	●
Dynamische Firmwareaktualisierung mit HMC (Hardware Management Console)		○	○	○	●	●
Service Agent-Call-Home-Funktion	●	●	●	●	●	●
Serviceanzeigen – Hinweisanzeige oder Anzeigen im Diagnosefeld “Light Path Diagnostics”	Light Path	Light Path	Light Path	Light Path	Hinweis	Hinweis
Serviceprozessorunterstützung für integrierten Selbsttest (BIST) für Logik/Arrays, Verbindungstests, Komponenteninitialisierung	●	●	●	●	●	●
Systemspeicherauszug für Hauptspeicher, POWER Hypervisor, Serviceprozessor	●	●	●	●	●	●
Veröffentlichungen im InfoCenter und auf der Service-Site für Systemunterstützung (Systems Support Site)	●	●	●	●	●	●
Wartungshilfe mit Repair & Verify	NZ	○	○	○	●	●
Meldung von Betriebssystemfehlern an HMC SFP-Anwendung	NZ	○	○	○	●	●
RMC-Subsystem für sichere Fehlerübertragung	NZ	○	○	○	●	●
Statusprüfung für geplante Operationen mit HMC	NZ	○	○	○	●	●
Bedienerkonsole (real oder virtuell)	●	●	●	●	●	●
Gleichzeitig ablaufende Bedienerkonsolenwartung	NZ				●	NZ
Redundante HMCs	NZ	○	○	○	○	○
Automatisierte Serverwiederherstellung/automatisierter Serverneustart	●	●	●	●	●	●
Knoten während des Betrieb hinzufügen/Knoten außerhalb des Betriebs reparieren	Blades während des Betriebs austauschen				●[14]	●[23]
Knoten während des Betriebs reparieren/Hauptspeicherupgrade während des Betriebs	Blades während des Betriebs austauschen				●[24]	●[23]
PowerVM Live Partition Mobility / Live Application Mobility	○	○	○	○	○	○
Stromversorgung und Kühlung						
Redundante Hot-Swap-Lüfter und -Gebläse für CEC	●[15]	●	●	●	●	●
Redundante Hot-Swap-Stromversorgungen für CEC	●[25]	710 ○730 ●	720 ○740 ●	●	●	●
Redundante Spannungsreglerausgaben					●	●
TPMD für Stromversorgung und Wärmemanagement des Systems	●	●	●	●	●	●
CEC-Stromversorgungs-/Wärmesensoren (CPU und Hauptspeicher)	●	●	●	●	●	●
Redundante Stromversorgung für E/A-Einschübe			●[18]	●[18]	●[18]	●[18]

[8] Nur in angeschlossenen E/A-Einschüben unterstützt
[9] In Basiseinheit und in E/A-Einschüben unterstützt (Anmerkung: E/A-Einschübe werden in Power 755 nicht unterstützt.)
[10] Im BladeCenter S-Gehäuse
[11] Funktionalität in E/A-Einschüben
[12] NZ = Nicht zutreffend
[13] Verfügbarkeitstermine siehe IBM United States Hardware Announcement 110-158 vom 17. August 2010
[14] Verfügbarkeitstermine siehe IBM United States Hardware Announcement 110-025 vom 9. Februar 2010
[15] Im BladeCenter-Gehäuse
[16] Weitere Einzelheiten finden Sie im Dokument zu Fakten und Leistungsmerkmalen von Power Systems (http://www-03.ibm.com/systems/power/hardware/reports/factsfeatures.html) für POWER7-Server (POB03022-USEN-05).

Anhang B: Betriebssystemunterstützung für ausgewählte RAS-Funktionen [● = verfügbar, ○ = optional]

RAS-Funktion	AIX V5.3	AIX V6	AIX V7	IBM i V6	IBM i V7	RHEL 5.5 ,	SLES 10	SLES 11
Prozessor								
Prozessor-Fabric-Bus-Schutz	●	●	●	●	●	●	●	●
Dynamic Processor Deallocation	●	●	●	●	●	●	●	●
Dynamic Processor Sparing	●	●	●	●	●	●	●	●
◆ Einsatz von CoD-Kernen	●	●	●	●	●	●	●	●
◆ Nutzung von Kapazität aus Zusatzpools	●	●	●	●	●	●	●	●
Fehlerbehebung beim Prozessorkern								
Processor Instruction Retry (PIR)	●	●	●	●	●	●	●	●
◆ Alternate Processor Recovery (APR)	●	●	●	●	●	●	●	●
◆ Core-Contained-Checkstop-Fehler für Partition	●	●	●	●	●	●	●	●
Persistente Aufhebung der Prozessorzuordnung	●	●	●	●	●	●	●	●
E/A-Subsystem								
Persistente Aufhebung der GX+-Buszuordnung	●	●	●	●	●			
Optionaler E/A-Hub mit Einfrierverhalten (Freeze)	●	●	●	●	●	●	●	●
Erweiterte Fehlererkennung für PCI-Bus	●	●	●	●	●	●	●	●
Erweiterte Fehlerbehebung für PCI-Bus	●	●	●	●	●	Einge-schränkt	Einge-schränkt	Einge-schränkt
Erweiterte Fehlerbehandlung (EEH - Enhanced Error Handling) für PCI-PCI-Bridge	●	●	●	●	●			
Redundante 12-fach-Kanalverbindung	●	●	●	●	●	●	●	●
Taktgeber und Serviceprozessor (SP)								
Dynamischer SP-Failover während des Betriebs	●	●	●	●	●	●	●	●
Redundante SP- und Knotencontroller mit dynamischem Failover während des Betriebs	●	●	●	●	●	●	●	●
Failover für Taktgeber während des Betriebs	●	●	●	●	●	●	●	●
Verfügbarkeit des Hauptspeichers								
ECC-Speicher, L2-Cache, L3-Cache	●	●	●	●	●	●	●	●
Fehlererkennung/-behebung	●	●	●	●	●	●	●	●
◆ Chipkill-Speicher mit zusätzlicher Halbbytekorrektur	●	●	●	●	●	●	●	●
◆ DRAM-Ersatzspeicherfunktion	●	●	●	●	●	●	●	●
Ersatzspeicherfunktion mit CoD bei IPL	●	●	●	●	●	●	●	●
CRC (Cyclic Redundancy Check) mit Wiederholung auf dem Hauptspeicher-datenbus (CPU-to-Buffer)	●	●	●	●	●	●	●	●
ECC für Datenbus (Hauptspeicherpuffer zu DRAM) mit Wiederholung	●	●	●	●	●	●	●	●
Dynamische Speicherkanalreparatur	●	●	●	●	●	●	●	●
Speichertest (Memory Scrubbing) für Prozessorspeichercontroller	●	●	●	●	●	●	●	●
Aufhebung der Zuordnung von Hauptspeicherseiten	●	●	●	●	●	●	●	●
L1-Paritätsprüfung mit „retry“/„set delete“	●	●	●	●	●	●	●	●
L2-Cache-Line-Delete (L2-Cachezeilenlöschung)	●	●	●	●	●	●	●	●
L3-Cache-Line-Delete (L3-Cachezeilenlöschung)	●	●	●	●	●	●	●	●
Special Uncorrectable Error Handling (spezielle Behandlung nicht behebbarer Fehler)	●	●	●	●[21]	●[31]	●	●	●
Active Memory Mirroring for Hypervisor	●	●	●	●	●	●	●	●
Fehlererkennung und -eingrenzung								
FFDC für Fehlererkennung/Fehlereingrenzung	●	●	●	●	●	●	●	●
Storage Protection Keys	●	●	●	●	●			
Fehlerprotokollanalyse	●	●	●	●	●	●	●	●

[17] RHEL 5.5 = Red Hat Enterprise Linux 5.5
[18] IBM Absichtserklärung für die Unterstützung von POWER7-basierten Power Systems-Servern und Blade-Systemen in der künftigen Version Red Hat Enterprise Linux 6 von Red Hat. Bei weiteren Fragen zur Verfügbarkeit dieses Release wenden Sie sich bitte an Red Hat.
[19] SLES 10 = Novell SUSE LINUX Enterprise Server 10
[20] SLES 11 = Novell SUSE LINUX Enterprise Server 11
[21] Checkstop-Fehler für Partition
[14] Verfügbarkeitstermine siehe IBM United States Hardware Announcement 110-025 vom 9. Februar 2010
[15] Im BladeCenter-Gehäuse

RAS-Funktion	AIX V5.3	AIX V6	AIX V7	IBM i V6	IBM i V7	RHEL 5.5 ,	SLES 10	SLES 11
Wartungsfreundlichkeit								
Fortschrittsanzeiger beim Booten	●	●	●	●	●	Einge-schränkt	Einge-schränkt	Einge-schränkt
Firmware-Fehler-Codes	●	●	●	●	●	●	●	●
Betriebssystemfehlercodes	●	●	●	●	●	Einge-schränkt	Einge-schränkt	Einge-schränkt
Umgebungs- und Netzüberwachung	●	●	●	●	●	●	●	●
Austausch von DASD-Einheiten/Datenträgern/PCI-Adaptern während des Betriebs (Hot Swap)	●	●	●	●	●	●[32]	●[32]	●[22]
PCI-Kartenaustausch während des Betriebs (Hot Swap)	●	●	●	●	●	●	●	●
Zweifache Plattencontroller/geteilte Rückwandplatine (Split Backplane)	●	●	●	●	●	●	●	●
Erweiterte Fehlerdatensammlung	●	●	●	●	●	●	●	●
SP-„Call-Home-Funktion“ in Nicht-HMC-Konfigurationen	●	●	●	●	●	●	●	●
Redundante Anschlüsse für E/A-Einschübe	●	●	●	●	●	●	●	●
Einbau von E/A-Einschüben während des Betriebs und Reparatur bei eingeschalteter Einheit	●	●	●	●	●	●	●	●
GX-Adaptereinbau während des Betriebs und Reparatur bei eingeschalteter Einheit	●	●	●	●	●	●	●	●
Gleichzeitig ablaufender Einbau eines E/A-Gehäuses mit Stromversorgung	●	●	●	●	●	●	●	●
Gegenseitige SP-Überwachung mit POWER Hypervisor	●	●	●	●	●	●	●	●
Dynamische Firmwareaktualisierung mit HMC (Hardware Management Console)	●	●	●	●	●	●	●	●
Service Agent-Call-Home-Funktion	●	●	●	●	●	●	●	●
Serviceanzeigen – Hinweisanzeige oder Anzeigen im Diagnosefeld „Light Path Diagnostics“	●	●	●	●	●	●[32]	●[32]	●[32]
Serviceprozessorunterstützung für integrierten Selbsttest (BIST) für Logik/ Arrays, Verbindungstests, Komponenteninitialisierung	●	●	●	●	●	●	●	●
Systemspeicherauszug für Hauptspeicher, POWER Hypervisor, Serviceprozessor	●	●	●	●	●	●[32]	●[32]	●[32]
Veröffentlichungen im InfoCenter und auf der Service-Site für Systemunterstützung (Systems Support Site)	●	●	●	●	●	●	●	●
Wartungshilfe mit Repair & Verify	●	●	●	●	●	Einge-schränkt	Einge-schränkt	Einge-schränkt
Vor-Ort-Schulung der Systemunterstützung	●	●	●	●	●	●	●	●
Meldung von Betriebssystemfehlern an HMC SFP-Anwendung	●	●	●	●	●	●[32]	●[32]	●[32]
RMC-Subsystem für sichere Fehlerübertragung	●	●	●	●	●	●	●	●
Statusprüfung für geplante Operationen mit HMC	●	●	●	●	●	●	●	●
Bedienerkonsole (real oder virtuell)	●	●	●	●	●	●	●	●
Gleichzeitig ablaufende Bedienerkonsolenwartung	●	●	●	●	●	●	●	●
Redundante HMCs	●	●	●	●	●	●	●	●
Automatisierte Serverwiederherstellung/automatisierter Serverneustart	●	●	●	●	●	●	●	●
Unterstützung für Hochverfügbarkeitsclustering	●	●	●	●	●	●	●	●
Gleichzeitig ablaufende Kernelaktualisierung		●	●	●	●			
Knoten während des Betrieb hinzufügen/Knoten außerhalb des Betriebs reparieren	●	●	●	●	●	●	●	●
Knoten während des Betriebs reparieren/Hauptspeicherupgrade während des Betriebs	●	●	●	●	●	●	●	●
PowerVM Live Partition Mobility / Live Application Mobility	○	○	○			○	○	○
Stromversorgung und Kühlung								
Redundante Hot-Swap-Lüfter und -Gebläse für CEC	●	●	●	●	●	●	●	●
Redundante Hot-Swap-Stromversorgungen für CEC	●	●	●	●	●	●	●	●
Redundante Spannungsreglerausgaben	●	●	●	●	●	●	●	●
TPMD für Stromversorgung und Wärmemanagement des Systems	●	●	●	●	●	●	●	●
CEC-Stromversorgungs-/Wärmesensoren (CPU und Hauptspeicher)	●	●	●	●	●	●	●	●
Redundante Stromversorgung für E/A-Einschübe	●	●	●	●	●	●	●	●

[22] Alle Systeme außer Blade-Systeme

Zu den Autoren:

Jim Mitchell ist leitender IBM Entwickler und Certified IT Infrastructure Consultant (zertifizierter IT-Infrastrukturberater). Er arbeitete im Bereich Mikroprozessordesign und leitete ein Entwicklerteam für Betriebssysteme. Als Inhaber eines IBM Patents veröffentlichte Jim Mitchell zahlreiche Artikel über das Design von Gleitkommaprozessoren, über die Systemsimulation und -modellierung und über Serversystemarchitekturen. Jim Mitchell gehört derzeit dem Austin Executive Briefing Center an.

Daniel Henderson ist ein IBM Senior Technical Staff Member. Er war seit den frühesten Tagen RISC-basierter Produkte Mitglied des Entwicklerteams in Austin und ist derzeit der maßgebende Entwickler, der sich mit der Verfügbarkeit von Systemen für IBM Power Systems-Plattformen beschäftigt.

George Ahrens ist ein IBM Senior Technical Staff Member. Er ist für die Servicestrategie und für die Architektur von Power Systems zuständig. George Ahrens veröffentlichte mehrere Artikel über die RAS-Modellierung sowie einige White Paper über den RAS-Systemaufbau und über bewährte Verfahren für die Verfügbarkeit. Er ist Inhaber zahlreicher Patente im Zusammenhang mit den RAS-Funktionen und mit dem RAS-Systemaufbau auf partitionierten Servern. George Ahrens leitet derzeit eine Gruppe von Servicearchitekten, die an der Festlegung der Servicestrategie und -architektur für Produkte der IBM Systems and Technology Group arbeitet.

IBM Deutschland GmbH
IBM-Allee 1
71139 Ehningen
ibm.com/de

IBM Österreich
Obere Donaustrasse 95
1020 Wien
ibm.com/at

IBM Schweiz
Vulkanstrasse 106
8010 Zürich
ibm.com/ch

Die IBM Homepage finden Sie unter:
ibm.com

Die IBM Power Systems-Homepage finden Sie unter:
ibm.com/systems/power/

Bei IBM heißt Dienst am Kunden zugleich auch Dienst an unserer Umwelt: Wir nehmen Ihre IBM Altgeräte und Zubehörteile zurück und stellen deren umweltfreundliche Entsorgung zum Selbstkostenpreis sicher. IBM Hardwareprodukte sind fabrikneu hergestellt.

Alle Informationen zu Produkten anderer Anbieter stammen von den Anbietern der aufgeführten Produkte oder aus anderen allgemein verfügbaren Quellen. Fragen zu den Leistungsmerkmalen von Produkten anderer Anbieter sind an den jeweiligen Anbieter zu richten.

Alle in diesem Dokument enthaltenen Leistungsdaten stammen aus einer kontrollierten Umgebung. Die tatsächlichen Ergebnisse können davon abweichen. Alle Leistungsdaten werden von IBM auf der Grundlage des gegenwärtigen Zustands (auf „as-is“-Basis) ohne jegliche Gewährleistung zur Verfügung gestellt. Käufer sollten andere Informationsquellen, einschließlich Systembenchmarks, zu Rate ziehen, um die Leistung eines Systems zu bewerten, dessen Kauf sie in Erwägung ziehen.

POW03056-USEN-04