Hochverfügbarkeitssysteme

Helmut Montsch

1 Hochverfügbarkeitssysteme

Ausfallzeit ist sehr teuer. Höchste Verfügbarkeit heißt: Ein System fängt unvorhergesehene Ausfälle von Komponenten unterbrechungsfrei ab oder steht nach einem Fehler in kürzester Zeit wieder zur Verfügung. Geplante Unterbrechungen des Systems werden durch den Einsatz von Ersatzsystemen für den Benutzer nicht spürbar und haben keine nachhaltige Auswirkung auf den Geschäftsbetrieb. Gerade in der Leistungsklasse der Netzwerk- und Datenbankserver besteht die Forderung nach Flexibilität, Leistungsfähigkeit, Skalier- und vor allem hoher Verfügbarkeit (HV). Studien der International Data Corp. (IDC), USA, ergaben durchschnittliche Kosten in Höhe von 78.000 $ pro Stunde Ausfallzeit für einen mittleren Betrieb. Man rechne nur die verlorene Arbeitsleistung der wartenden Mitarbeiter. In so genannten geschäftskritischen Anwendungen (mission critical applications), wie z. b. Geldtransaktionen bei Finanzinstituten oder Flugbuchungen bei Reiseanbietern kostet eine Stunde Arbeitsunterbrechung oft eine Million $ und mehr.

Eine Studie in Deutschland von Contingency Planning Research Inc. ergab folgende finanzielle Auswirkung im deutschen Geschäftsbereich:

Branche	Operation	Auswirkungen pro Stunde
Verkehr	Flugticket-Reservierung	ca. 254.000 DM
Finanzen	Börsenhandel	ca. 12.255.000 DM
Telekom	Aktivierung von Mobilfunktelefonen	ca. 114.000 DM
Unterhaltung	Ticketverkauf per Telefon	ca. 131.000 DM

Bei Ausfall der IT-Unterstützung ergaben sich folgende Werte für die Überlebensfähigkeit von Unternehmen:

Branche	Tage
Banken	2
Industrie	5
Versicherungen	5,5
Handel	2,5

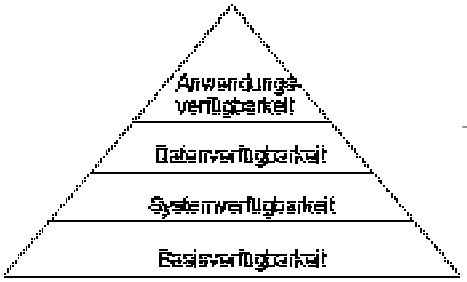
Ein kleines Praxisbeispiel soll zeigen, welche Auswirkungen Systemstillstände von EDV-Systemen haben können:

Je sensibler die Kunden-Lieferanten-Beziehung desto wichtiger wird die ständige Verfügbarkeit der Informationstechnik. Besonders deutlich wird dies am Beispiel der Automobilzulieferindustrie. Die Unternehmer dieser Branche unterhalten eine sehr enge Beziehung zu ihren Kunden. Außer attraktiven Preisen und hoher Funktionalität der Produkte erwarten die großen Automobilhersteller von ihren Zulieferbetrieben "Just-intime"-Lieferung und die laufende Qualitätssicherung.

Der Geschäftserfolg der Automobilzulieferer steht und fällt mit der Verfügbarkeit ihrer Informationssysteme, denn ihre Produktionsplanung und -steuerung ist Teil des Herstellprozesses ihrer Kunden.

Besonders wichtig für die Geschäftsbeziehung ist die Liefertreue. Oftmals bleibt zwischen Auftrag und Lieferung weniger als ein Tag. Im harten Wettbewerb der Zulieferindustrie kann ein Ausfall der Informationstechnik aufgrund hoher Konventionalstrafen und drohendem Auftragsverlust bei Verspätungen schnell zur Existenzfrage werden.

Die Hochverfügbarkeit baut auf vier Stufen auf:



- Die Basisverfügbarkeit wird durch den Qualitätsstandard der Serverprodukte gegeben. Die Fertigungsprozesse müssen die strengen ISO 9001-Richtliniern erfüllen. Das heißt für Hersteller, dass alle Komponenten auch die zugekauften Komponenten den Qualitätsansprüchen der Norm entsprechnen müssen.

  Daraus ergibt sich eine Standardverfügbarkeit von 99%, die normalerweise jedes System ohne zusätzliche Kosten liefern muss. Hochgerechnet bedeutet das, man muss mit 88 Stunden Stillstand im Jahr rechnen.
- Mittels zusätzlicher Komponenten und Mechanismen kann die Systemverfügbarkeit erhöht werden:
- Batteriegepufferte Speicher und unabhängige unterbrechungsfreie Stromversorgung,
- Redundante Prozessoren, Speicher und Controller,
- Auch im laufenden Betrieb austauschbare, redundante Netzteile, Lüfter und Magnetplatten.

  So erreicht man eine Verfügbarkeit des Gesamtsystems von 99,2 bis 99,6%.
- Durch weitere Maßnahmen kann man die Datenverfügbarkeit erhöhen. Zum Beispiel mit Spiegelplatten und RAID-Systemen (Redundant Array of independant Disks) sowie mit Hilfe von Mechanismen, die das Sichern und Restaurieren von Daten beschleunigen. So steht noch mehr Rechenzeit für die produktive Nutzung zur Verfügung (99,6 bis 99.9%).
- Höchste Verfügbarkeit der gesamten installierten Anwendung bedeutet, dass alle Komponenten (Hardware, Systemsoftware und Middleware) optimal aufeinander abgestimmt sein müssen.

  Die Sicherung von Cluster, also Server Teams, ist ein wichtige Option zu Sicherung der Verfügbarkeit. Die Server eines "Fail-over-Clusters" überwachen sich permanent gegenseitig. Bei einem Ausfall werden angeschlossene Arbeitsplätze und die Peripherie des betroffenen Servers automatisch auf einen anderen intakten Server umgeschaltet. Anwendungen und Netzverbindungen laufen sofort und ebenfalls automatisch wieder neu an. Datenzugriffe sind praktisch unterbrechungsfrei möglich. Die Anwendungsverfügbarkeit wird immer ein wichtigeres Kriterium für die EDV-Welt, dies gilt jetzt nicht mehr nur für das Pentagon sondern auch für Server die im Midrange-Bereich im Einsatz sind.

2 Basisverfügbarkeit

Die Basisverfügbarkeit muss - wie anfangs schon erwähnt - durch einen hohen Qualitätsstandard gesichert werden. Standardmaßnahme zur Fehlersicherung ist Cyclic Redundancy Check. Die für die Prüfsummenbildung erforderlichen Schieberegister und Vergleichsoperationen werden im allgemeinen in der Hardware implementiert.

3 Systemverfügbarkeit

Zur Grundausstattung für Server mit Systemverfügbarkeit sollten nachfolgende Leistungsmerkmale beinhalten:

3.1 ECC, EDC (Error Correction Code, Error Detection and Correction)

EDC ist ein Verfahren, um Speicherfehler in Speichermodulen zum Zeitpunkt des Auslesens der jeweiligen Speicherzelle zu korrigieren. ECC ist die dabei verwendete algorithmische Methode. Auftretende Speicherfehler werden korrigiert.

3.2 Memory Scrubbing

Der oben beschriebene ECC/EDC-Mechanismus korrigiert eventuelle Speicherfehler jeweils nur beim direkten Zugriff auf die betreffenden Speicherzellen. In selten benutzten Bereichen des Hauptspeichers kann dies dazu führen, dass sich mehrere Bitfehler anhäufen (Akkumulator-Effekt) und mit ECC/EDC bei einem späteren Speicherzugriff nicht mehr behoben werden können. Memory Scrubbing ist eine zusätzliche Firmeware-Funktion, die den gesamten Speicher

Helmut Montsch montsch.helmut@aon.at

zyklisch durchsucht und Fehler in den Speichermodulen direkt berichtigt, unabhängig davon, ob Zugriffe auf diesen Speicherbereich erfolgen, oder nicht. Die HW-basierte Methode vermeidet einerseits zusätzliche Systembelastung, andererseits ist die Konsistenz der Hauptspeicher-Dateninhalte auf ein nochmals höheres Sicherheitsniveau gehoben.

3.3 PDA (Prefailure Detection and Analysing)

Server sollten einen Fehlerfrüherkennungsmechanismus besitzen. Dies kann über eigene server-optimierte Firmware realisiert werden. Diese ist in einem speziellen elektronischen Baustein gekapselt (ASIC). Beispielsweise misst eine PDA-Funktion ständig die Leistung der Lüfter und der CMOS-Batterie. Sollten sich hier Veränderungen andeuten, so kann über eigene Software eine Vorwarnung erfolgen. Komponenten können vorsorglich ausgetauscht werden, bevor sie wirklich defekt sind. Zusätzlich sollte PDA über einen Speicher für die Ablage von Systemfehlermeldungen besitzen, den man mit spezielle Tools auslesen kann. Dies erleichtert für den Servicetechniker die Fehlersuche.

3.4 ASR&R (Automatic Server Reconfiguration & Restart)

Bei Ausfall einer Komponente (z.B. ein Prozessor im SMP-System, Speichermodule) wird der Server automatisch neu gestartet und das defekte Modul aus der Konfiguration ausgeblendet. Der Server ist nach Wiederstart weiterhin betriebsbereit und setzt seine Arbeit ohne die defekte Komponente fort.

3.5 Redundante Komponenten

Redundanz bedeutet, dass von einer Komponente in einem System mehr eingebaut sind, als im Normalbetrieb benötigt werden. Fällt eine aus, wird der Betrieb automatisch sofort mit den verbleibenden Elementen weitergeführt.

Redundant sind z.B. die Stromversorgung, das bedeutet, dass selbst in der Maximalkonfiguration optional eine Stromversorgung mehr in einen Server eingebaut werden kann als erforderlich und in Problemfällen den Serverbetrieb aufrecht erhält.

3.6 Hot Plug Funktionalitäten

Hot Plug-Komponenten können im laufenden Betrieb ausgetauscht werden. Dieses Verfahren kommt bei den Festplatten und bei den Stromversorgungen zum Einsatz. Fällt eine Festplatte aus, sind die Daten mit Hilfe der RAID-Technologie (Datenverfügbarkeit) vor Verlust geschützt. Über das Servermanagment z.B. wird der Systemverantwortliche informiert, welche Komponente defekt ist. Sie kann dann gegen eine neue ausgewechselt werden, ohne den Betrieb des Servers zu unterbrechen.

Ungeplante Serverstillstände werden durch den Einbau redundanter Komponenten verhindert, zusammen mit ihrer Hot Plug-Fähigkeit wird eine höhere Gesamtverfügbarkeit des Servers und somit eine merkliche Steigerung der Produktivität erreicht.

4 Datenverfügbarkeit

4.1 Datensicherung

Die Datensicherung wird meist bei kleineren Servern in den Hintergrund geschoben oder ganz außer Acht gelassen. Sie ist aber für einen sicheren Datenbestand unumgänglich Das Datensicherungskonzept sollte so ausgelegt werden, dass auch ein Kunde im Fehlerfall eine Rekonstruktion der Daten, es kann ja auch einmal ein Datenbankfehler auftreten, durchführen kann. Weiters sollte man sich überlegen, wie am schnellsten, wenn im Fehlerfall das ganze System ausfällt, das Operating System eingespielt werden kann. Entsprechende Backup-Software gibt es am Markt. Bei einen NT-Server könnte man z.B. ein Not-NT installieren und mit diesem Not-NT das Produktiv-NT sichern.

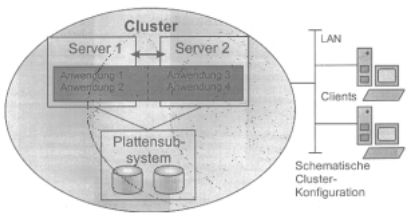
4.1 Raid-Systeme

Wer im Bereich Datensicherheit auf Leistung und Performance Wert legt, kommt heutzutage am Begriff RAID (Redundant Array of Independant Disks) nicht mehr vorbei. Wie so häufig wird man zunächst einmal mit einer Menge (vielleicht) neuer Begriffe konfrontiert. Es gibt im Grunde genommen vier klassische RAID-Level, die sich aus den Begriffen Stripe-Set und XOR-Verknüpfung herleiten lassen.

Zusätzlich kann bei Verwendung eines RAID-Controllers eine Hot Spare Festplatte (auch als Hot Standby bezeichnet) verwendet werden. Dies ist ein zusätzliches Laufwerk, das der Kontroller beim Ausfall einer anderen Platte automatisch in die RAID-Konfiguration einbindet, mit dem Datenset der ausgefallenen Festplatte beschreibt, zu einem späteren Zeitpunkt kann durch den IT-Service dann ein Austausch erfolgen.

5 Anwendungsverfügbarkeit

Gehen die Verfügbarkeitsanforderungen über diese Eigenschaften und deren Möglichkeiten hinaus, sind weitergehende Maßnahmen für die Serververfügbarkeit notwendig. Anwender unternehmenskritischer Applikationen, wie MS BackOffice Server, SAP R/3, fordern in vielen Fällen die Redundanz kompletter Server, um die IT-Verfügbarkeit auch bei Ausfall eines gesamten Servers gewährleisten zu können. Hier werden Cluster-Konfigurationen eingesetzt, bei denen neben höherer Verfügbarkeit als weiterer Vorteil die gesteigerte Verarbeitungsleistung durch den 2 Server entsteht.



Obige Grafik gibt den grundsätzlichen Aufbau eines Clusters wieder: Clustering stellt die Zusammenfassung meherer, voneinander unabhängiger Server zu einem logischen Gesamtsystem - dem Cluster dar. Nach außen hin verhält es sich wie ein einziger Server. Diese Struktur ist das Cluster, in der Grafik als großes Oval dargestellt. Ein Server innerhalb eines Clusters wird als Knoten bezeichnet. Dabei ist es unerheblich, ob es sich um Ein- oder Mehrprozessorsysteme handelt.

Auch müssen die Knoten innerhalb eines Clusters in Bezug auf Anzahl der CPUs und Speichergröße nicht identisch ausgestattet sein. Durch Server-Cluster wird - über redundante, online ersetzbare HW-Komponente hinaus - die System- und Applikationsredundanz erreicht. Clustering bildet somit die nächsthöhere Ebene der Maßnahmen für Server-HV.

Die Daten innerhalb des Cluster werden generell in einem gemeinsamen Plattensubsystem mit Zugang für alle Server gehalten. So ist bei einem Serverausfall die Übernahme der Applikationen sowie der erforderliche Zugriff zu den Daten durch den 2 Server möglich. Die erforderliche, gegenseitige "Lebendüberwachung" bei der Systeme erfolgt über den Server-Interconnect, der zusätzlich zum LAN-Schluss in beiden Servern eine vom allgemeinen Netzwerkverkehr unabhängige Kommunikation zwischen den Knoten ermöglicht (dargestellt durch den Doppelpfeil in obiger Grafik).

Clients im Netzwerk, die eine Applikation nutzen möchten, adressieren nicht mehr einen speziellen Server, sondern die logische Struktur "Cluster". Dadurch wird erreicht, dass die Anwendungen im Cluster unabhängig von den einzelnen Servern werden. Fällt ein Knoten aus, übernimmt ein anderer dessen Aufgaben, d.h. seine Applikationen und Festplatten. Da die Clients mit dem Cluster verbunden sind, laufen die Anwendungen aus Anwendersicht nach dem Neustart transparent auf dem übernehmenden Knoten weiter.

Clustering bietet vier wesentliche Vorteile:

- **Steigerung der Verfügbarkeit**

  Durch das Zusammenfassen mehrerer Server steigt die Verfügbarkeit der gesamten Konfiguration. Bei dem Ausfall eines Servers (ungeplante Ausfallzeit) übernimmt das verbleibende System dessen Aufgaben solange, bis ersterer wieder zur Verfügung steht. Diese Eigenschaft von Clustering ist auch dann von großem Nutzen, wenn wegen geplanten Ausfallzeiten ein Knoten für HW- oder SW-Wartung aus dem Betrieb genommen werden soll. In diesem Fall werden die relevanten Anwendungen zeitweilig auf den/die anderen Knoten ausgelagert. Das Gesamtsystem bleibt dabei weiter verfügbar.

- **Steigerung der Leistungsfähigkeit**

  Abhängig von der Anzahl der Knoten erhöht sich auch die gesamte Verarbeitungsleistung eines Clusters. Beide/alle Cluster-Knoten arbeiten

jeweils an unterschiedlichen Applikationen/Services. Nur ein geringer Teil der Verarbeitungsleistung ist für die Cluster-Koordination erforderlich.

- **Erhöht Flexibilität**

  Innerhalb des Clusters lassen sich auf den Knoten beliebig Anwendungen betreiben und Daten oder Services bereitstellen. Dies ist sogar dynamisch möglich, da je nach Auslastungsgrad des jeweiligen Knotens Anwendungen administratorgesteuert auf weniger belastete Server verlagert werden können.

- **Senkung der TCO** (Total Cost of Ownership)

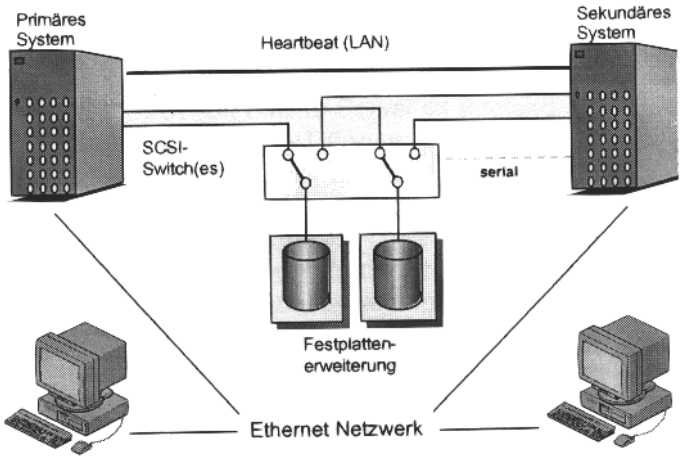
  Die Verwendung von Clustering wirkt sich signifikant auf die Reduktion von Kosten im Unternehmen aus: durch die erhöhte Verfügbarkeit des Gesamtsystems werden Ausfallzeiten vermieden. Es kommt nicht zu kostenintensiven Stillständen des ganzen Unternehmens oder von Teilbereichen, was wiederum auf für geplante Stillstandszeiten eines Servers zutrifft.

5.1 Server-Clustering

Für Clustering exisitiert heute eine Reihe von Lösungswegen. Sie decken nach unterschiedlichen Methoden hauptsächlich die Bedürfnisse nach Hochverfügbarkeit ab. Dabei wird die Unterscheidung getroffen nach:

5.1.1 Serverredundanz durch Failover Server

Grundsätzlich: Bei einer solchen Konfiguration ist von zwei Servern ein System als Primär-Server vorgesehen, der im Produktivbetrieb läuft und die kritischen Anwendungen abarbeitet. Der zweite ist der Sekundär-Server, er überwacht das Primärsystem mittels einer eigenen (LAN-)Verbindung. Wahlweise kann er zusätzlich auch - weniger kritische - Anwendungen übernehmen, zum Beispiel als Testsystem genutzt werden. Dies wird auch als "Warm Stand-by Betrieb" bezeichnet.



ServerShield von Siemens ist ein Failover Server Konzept für Server, das auf der Basis der zusätzlichen ServerShield Hard- und Software höhere Verfügbarkeit für Server mit MS Windows NT zur Verfügung stellt. Beide Rechner sind über die ServerShield SCSI-Umschaltbox (SCSI-Switch) mit dem Plattensubsystem verbunden, zu dem im Normalbetrieb jedoch nur das primäre System Zugang hat. Auf den externen Festplatten befinden sich neben dem Betriebssystem auch die Anwendungen und die Daten. Über ein separates LAN in jedem Server wird das Primärsystem durch das sekundäre überwacht. Hierüber erkennt der sekundäre Server das Auftreten signifikanter Störungen in dem Primärsystem, die dessen Betrieb unmöglich machen. Über den/die SCSI-Switch(es) wird in diesem Fall das gesamte Plattensubsystem dem Sekundärserver zugeordnet. Dessen anschließender Reboot erfolgt mit der SW-Installation auf den externen Festplatten (Betriebssystem, Anwendungen). Dadurch wird für die Anwender die vollständige Funktion des primären auf dem sekundären Server wiederhergestellt. Sind auf letzterem noch Anwendungen aktiv, werden sie vor dem Failover beendet und können später - nach der Instandsetzung des Primärservers - erneut auf dem Sekundärsystem gestartet werden.

Nutzen: Hohe Verfügbarkeit durch (aktive) Server-Redundanz. Das Konzept erfordert dabei absolut keinen Programmier- bzw. Anpassungsaufwand, da eine vollständige Unabhängigkeit zwischen Hochverfügbarkeitslösung und den Anwendungen besteht.

5.1.2 Server-/Applikations-Redundanz durch Server-Cluster

Nach ähnlicher Systematik wie redundante (Standby-) Server sorgen in einem Cluster die einzelnen Server durch mehrfaches Vorhandensein für Hochverfügbarkeit und ggf. auch Skalierung. Im Unterschied dazu steht bei Clusterlösungen mit der Integration der Clusterfähigkeit in das Betriebssystem selber jedoch ein weitergehender Lösungsansatz zur Verfügung: alle Clusterknoten bearbeiten gleichberechtigt ihre Applikationen und Services. Es gibt hier keine mit besonderer Rolle oder Hierarchie versehenen Primär- oder Sekundärserver. Parallel zum Produktiveinsatz erfolgt die gegenseitige Überwachung der Serversysteme untereinander, ebenfalls unter Nutzung zusätzlicher Kommunikationsverbindungen. Der Zugriff zu den Datenbeständen wird über ein gemeinsam genutztes Subsystem realisiert (Multi Hosted, Shared Disk).

5.1.2.1 Microsoft Cluster Server für Windows NT

Die Ankündigung einer Cluster-Option von Microsoft für Windows NT erfolgte im Mai 1997. Eine erweiterte Windows NT-Version, die sogenannte Microsoft Windows NT Server Enterprise Edition, beinhaltet im wesentlichen fünf Neuerungen, die zur Verbesserung der Hochverfügbarkeit sowie Skalierbarkeit des NT Servers beitragen. Eine dieser neuen Komponenten ist der Microsoft Cluster Server, kurz MSCS, der in der zurückliegenden Zeit unter dem Codenamen "Wolfpack" bekannt wurde.

Im Vergleich zu allen bislang im Windows NT Umfeld verfügbaren Clustervarianten ist nun bei MSCS die Clusterfähigkeit weitgehend im Betriebssystem integriert. Daraus resultieren folgende Vorteile:

- Die Integration in das weit verbreitete MS Windows NT mit seiner weiterhin rasch wachsenden Installationsbasis schafft einen Standard. Für Applikationen, die für den Einsatz in einem Clsuter optimiert sein sollen, steht somit eine standardisierte Schnittstelle zur Verfügung.

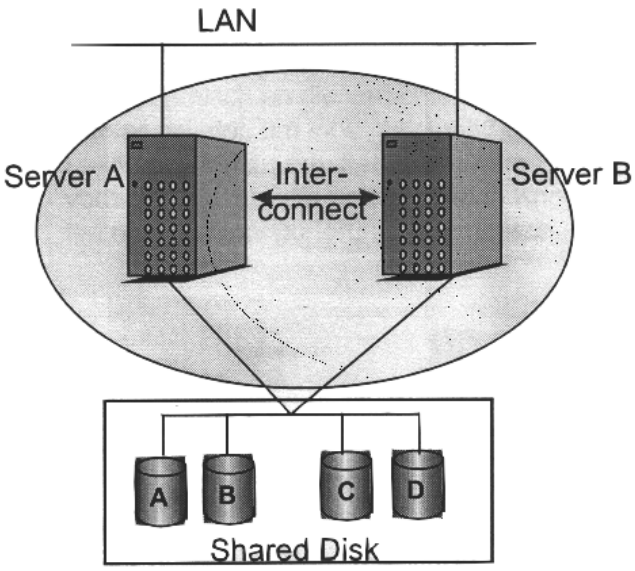
- Mit Microsoft als Hersteller im Hintergrund besteht die Sicherheit, dass auch in zukünftigen Versionen die Integrität und Kompatibilität des MSCS und der Anwendungen gewährleistet ist. Dies sorgt für zusätzlichen Investitionsschutz.

- Die Installation und Administration von MSCS erfolgen analog zur Bedienung anderer Microsoft Serverprodukte. Hierdurch sind einfache, zentrale Wartung und Konfiguration gewährleistet.

Voraussetzungen

Der Microsoft Cluster Server ist Bestandteil der Microsoft NT Server Enterprise Edition. Die erste Version unterstützt zwei Knoten pro Cluster. Auf beiden ist die Installation von Windows NT Server sowie MS Cluster Server erforderlich. Jeder Serverknoten muss verwaltungsseitig Mitglied der gleichen NT-Domäne sein und über je zwei Netzkarten verfügen. Jeweils eine davon ist für den Netzverkehr mit den Clients zuständig, das andere Board dient der internen Cluster-Kommunikation (Cluster-Management, Heartbeat). Grundsätzlich ist es sinnvoll, hierfür schnelle Verbindungen wie FDDI oder 100 Mbit-Ethernet zu verwenden.

MSCS arbeitet grundsätzlich nach dem Prinzip des "Shared Nothing": jeder Server im Cluster hat seine eigenen HW-Resourcen, wie Prozessoren, Hauptspeicher, Systemplatte und teilt diese Resourcen mit keinem weiteren Clusterknoten. Ausgenommen ist die gesamte Datenbasis: sie befindet sich auf einem gemeinsam genutzten externen Platten-Subsystem, das über einen gemeinsam genutzten SCSI-Bus im Shared Disk Modus von beiden Servern aus ansprechbar ist, was auch eine der Systemvoraussetzungen des MSS darstellt.



Eigenschaften

MSCS verfügt über zwei wesentliche Elemente:

- den NT-Systemdienst Cluster Server
- und das Administrationstool Cluster Administrator

Der Systemdienst beinhaltet die gesamte Steuerung der Cluster-Funktionalität. Zur Einrichtung, Überwachung und Konfiguration des Clusters und der Cluster-Applikationen steht der Cluster Administrator zur Verfügung.

Zur Definition von Cluster-Applikation stehen zwei Objekttypen bereit: Ressourcen und Gruppen. Ressourcen sind beispielsweise die IP-Adresse des Clusters, Festplatten, Netzwerk-Shares und Anwendungen. Alle zusammengehörigen Ressourcen werden in einer Gruppe zusammengefasst. Zusätzlich sind - soweit zutreffend - zwischen den Ressourcen bestehende Abhängigkeiten festzulegen, z.B. ein Knoten muss zuerst im Besitz einer Festplatte sein, bevor er eine darauf befindliche Applikation starten kann. Diese Unterteilung ist erforderlich, da die Applikationen zwar nach außen hin über das Cluster angesprochen werden, jedoch zu einem Zeitpunkt vollständig auf einem einzigen Knoten und dessen Peripherie ablaufen. All diese Informationen zum Cluster und seiner Konfiguration sind in den Servern (NT Registry) und auf einer für Knoten gemeinsamen Festplatte gespeichert, der sogenannten Quorum-Disk. Dazu stehen die Konfigurationsdaten beiden Servern jederzeit zur Verfügung und können nach einem Systemstart von dort gelesen werden.

Zu einem Zeitpunkt hat immer nur einer der Knoten einen exklusiven Zugriff auf eine Ressource, z.B. eine physikalische Festplatte. Alle Ressourcen des Clusters, die zu einer Gruppe zusammengefasst wurden, lassen sich zum einen aktiv durch den Administrator einem Knoten zuordnen bzw. per Kommando von diesem zum anderen Server bewegen. Das ist beispielsweise dann hilfreich, wenn ein Knoten für geplante HW- oder Software-Wartungsarbeiten aus dem Produktivbetrieb genommen werden soll. Zum anderen überträgt die Clusterverwaltung bei Ausfall eines Knoten automatisch die auf diesem Server laufenden Ressourcen auf den verbleibenden Knoten (Fail-over). Je nach Art der Applikation beträgt die Umschaltzeit zwischen wenigen Sekunden bei einfachen Anwendungen und bis in den Minutenbereich bei komplexen (wie SAP R/3). Hinzu addiert sich noch das Zurücksetzen nicht abgeschlossener Transaktionen der Datenbank (DB Recovery), dessen Zeit sehr stark von der Datenbankgröße abhängt.

Applikations-Unterstützung

Grundsätzlich sind alle heutigen Anwendungen für Windows NT auch auf dem MSCS ablauffähig. Es wird jedoch eine appliktionsspezifische Ressource-DLL (Dynamic Link Library) des jeweiligen Softwareherstellers benötigt, um die Applikation sinnvoll in den Cluster zu integrieren. Ist eine noch weitergehende Cluster-Integration der Anwendung bzw. Datenbank erwünscht, MSCS bereits heute das Cluster API an (Application Programming Interface). Auf dieser standardisierten Schnittstelle lassen sich Applikationen entwickeln, die noch optimaler mit MSCS ablaufen und zusammenarbeiten.

Eine Reihe bedeutender Software-Anbieter hat bereits die Unterstützung von MS Windows NT Server Enterprise Edition für ihre Anwendungs-Umgebungen angekündigt.

- SAP - R/3 Cluster für Windows NT
- MS SQL Server 7.0
- MS Exchange 5.5
- Oracle Fail Safe

5.1.2.2 Oracle Parallel Server (OPS)

Die Verteilung eines Datenbank - Managementsystems (DBMS) auf mehrere Server, den sog. Datenbank-Instanzen, erlaubt die für eine spezifische Datenbank oftmals geforderte Realisierung von Hochverfügbarkeit und auch Leistungsskalierung. Die Oracle Corp. bietet hierfür mit dem Oracle Parallel Server (OPS) für MS Windows NT eine Option für das Datenbank Serverclustering.

OPS stellt für Installation, Inbetriebnahme und Betrieb eine eigene Clusterverwaltung bereit und erlaubt die Konfiguration von bis zu vier Serverknoten zu einem OPS-Cluster. Ein weiteres Cluster - Managementsystem, wie MSCS wird dabei nicht benötigt. Die aus dem Netz an den Cluster eingehenden Anfragen werden dynamisch an die verschiedenen Server geleitet, um eine gleichmäßige Lastverteilung zu erhalten. Bei einem Serverausfall setzen die verbleibenden Knoten nach Rekonfiguration automatisch die Arbeit fort und führen die Abläufe des ausgefallenen Systems mit aus. Nach dessen Instandsetzung ist jederzeit das Wiedereinbringen in den Cluster möglich. Sämtliche Knoten haben Zugriff zu einem einzigen, gesamten Datenbestand auf einem Disk-Subsystem, das über Shared SCSI Fähigkeiten verfügen muss. Konkurrierende Schreibzugriffe werden von dem in OPS integrierten Distributed Lock Manager überwacht (DLM).