

Datenbank Clustering / SQL Server Clustering

Der Microsoft SQL Server kommt in Unternehmen immer öfter zum Einsatz. Dank den Cluster Services dient er unterdessen auch vermehrt als hochverfügbare Unternehmensdatenbank. Doch gerade bei unternehmenskritischen Daten kommt man nicht umher, auch einen sehr sicheren Betrieb zu gewährleisten. Der nachfolgende Artikel befasst sich damit, wie dies sichergestellt werden kann.

Mit den stetig wachsenden Anforderungen an die Datenverfügbarkeit werden Microsoft SQL Server vermehrt clusterbasierend installiert und betrieben. Dabei stellt Microsoft mit dem Betriebssystem Windows 2000 Advanced Server alle benötigten Services integriert zur Verfügung. Nebst dem für die Datenhaltung benötigten Failover Microsoft Cluster Service (MSCS, max. 4 Nodes) sind auch Loadbalancing Cluster (Network Load Balancing / NLB und Component Load Balancing / CLB, max. 32 Nodes) Konfigurationen möglich. Der Hauptunterschied liegt in der Datenhaltung. In MSCS wird für die Datenspeicherung (Data Layer) ein Shared Storage System eingesetzt zu dem alle angeschlossenen Server verbunden sind. Jeder Server hat dabei exklusiven Zugriff auf dedizierte Diskressourcen. Diese Ressourcen sind entsprechend den geclusterten Services und Prozessen zugewiesen. In einem Fehlerfall werden diese Services und die abhängigen Diskressourcen auf einen verbleibenden Server übernommen und aktiviert.

Ein Zwei Node Cluster kann zudem in zwei Modi betrieben werden: aktiv/passiv oder aktiv/aktiv. Das bedeutet für den Aktiv/aktiv-Modus, dass beide Nodes aktive Prozesse bzw. Ressourcen gleichzeitig zur Verfügung stellen. In einem Failover kann jeder der beiden Nodes - mit daraus resultierender, reduzierter Performance - alle Ressourcen übernehmen. In der Konfiguration aktiv/passiv wäre nur ein Node mit allen Ressourcen aktiv und der andere Node sozusagen „stand-by“ für die Übernahme der Services im Fehlerfall.

Bei den Loadbalancing-Systemen werden die Daten (Webseiten, Applikationen) auf den jeweiligen Servern gespeichert und müssen mittels geeigneten Mechanismen synchronisiert werden. Daher sind Loadbalancing-Systeme nur für statische Daten (reine HTML-Seiten, ASP- und ASPX-Seiten, Bilder, ...) geeignet.

In der Architektur wird das gesamte System als Multi-Tier- bzw. Multi-Layer-Architektur bezeichnet. Abbildung 1 zeigt diese Architektur in .Net wobei .Net nicht zwingend ist. Dieses Schichtenmodell ist auch in Windows 2000 anzuwenden.

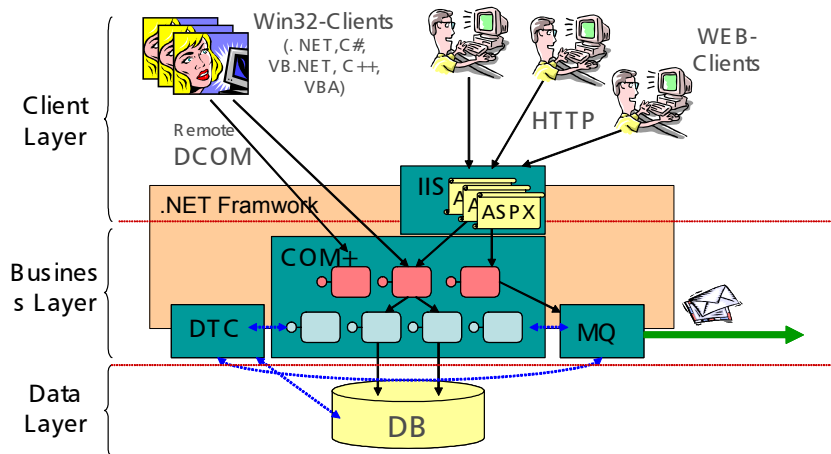


Abbildung 1

Data Layer

Für eine hohe System- und Datenverfügbarkeit ist nebst den Basistechnologien wie Stromversorgung, physikalischer Zugang, Elementarschutz (Wasser, Feuer) und Sicherheitskonzept auch die entsprechende Systemkonfiguration zwingend.

Microsoft bietet die Möglichkeit, bis zu vier Server (Nodes) zu einem MSCS zusammen zu führen. Dabei wird der Vier Node Cluster nur mit der Version Windows 2000 Datacenter unterstützt. Ein Zwei Node Cluster kann bereits mit Windows 2000 Advanced Server aufgesetzt werden.

Die Standard-Installation des Windows 2000 Clusters ist grundsätzlich problemlos. Es ist jedoch absolut notwendig, dass die Konfiguration exakt geplant und ausgeführt wird. So sind Systemwerte für Quorumdisk (Cluster-Konfiguration und -Zustand) und Diskcontroller-Einstellungen (write cache) optimal zu wählen.

Um die Data-Services auf Basis von Microsoft SQL Server 2000 implementieren zu können, ist ein genaues Design der Systemumgebung erforderlich. Dabei werden Parameter wie Wachstum der Datenbestände, Instanzierung, Anzahl Datenbanken oder Verwendungszweck analysiert. Somit hat diese Planung direkten Einfluss auf das Layout des Diskstorages. Die Diskressourcen sind selbstverständlich mit einem entsprechenden RAID-Level zu bilden.

Grundlegende Richtlinien für Datenbank-Implementierungen wie die physikalische Trennung von Data- und Log-Files müssen dabei genauso beachtet werden, wie auch Anforderungen an Performance und Sicherheit. Es spricht für sich, dass diese Konzeption von Spezialisten mit der nötigen Erfahrung durchgeführt werden muss. Im Speziellen sind Cluster-Installationen nur mit entsprechendem Know-how in der erwarteten hohen Qualität zu etablieren.

Ist das System nun entsprechend installiert und konfiguriert, muss eine Betriebsdokumentation alle wichtigen Parameter und Einstellungen aufzeigen.

Genauso, wie an ein Clustersystem hohe Erwartungen an die Verfügbarkeit gestellt werden, so werden auch an den Betrieb dieser Systeme hohe Anforderungen gesetzt. Ohne entsprechende Ausbildung und Unterstützung ist es leider oft der Fall, dass aus einem Fehlerfall schon fast ein Disaster wird. Erst dann die nötigen Massnahmen zu ergreifen und einen Schuldigen zu suchen, ist kaum richtig. Ein „gesundes“ System braucht Pflege und die Gewissheit der Verantwortlichen, einem hohen Qualitätsstand zu entsprechen.

Um diese Qualität sicher zu stellen, bietet die Trivadis AG einen SQL Server Health Check an. Dieses Package zeigt allfällige Mängel an der Konfiguration eines SQL Servers (2000 oder 7.0) auf. Dadurch kann Performance- und Systemsicherheitspotenzial verdeutlicht werden.

Glossar

Database-Engine > Datenbank System

Cluster > Computer System mit mehreren im Verbund arbeitenden Computern

Node > einzelner Computer aus dem Cluster

Failover > Wechsel eines Services von einem auf den anderen Node

Loadbalancing > Mechanismus um Anforderungen auf mehrere Server zu verteilen

Shared Storage > Festplatten-Speichersystem, welches von den Nodes aus dem Cluster verwendet wird

Write cache > Festplattenkontroller-Pufferspeicher für Schreibvorgänge

RAID > redundante Anordnung unabhängiger Festplatten