

# Google File System

## aus Wikipedia, der freien Enzyklopädie

Das **Google File System** (**GFS**) ist ein proprietäres, auf Linux basierendes verteiltes Dateisystem, das Google für seine Anwendungen benutzt. Es ist für Googles Internetsuche optimiert. Die Daten sind bleibend in teilweise mehrere Gigabyte großen Dateien gespeichert, die selten gelöscht, überschrieben oder verkleinert werden. Auch ist es für hohe Datendurchsätze optimiert.

### Inhaltsverzeichnis

- 1 Aufbau
- 2 Kritik
- 3 Quellen
- 4 Weblinks

## Aufbau

Das Google File System ist an die notwendigen Anforderungen der *Websuche* angepasst, die eine enorme Menge an zu speichernden Daten generiert. GFS entstand aus einem früheren Versuch Googles, welcher den Namen „*BigFiles*“ trägt und von Larry Page sowie Sergei Brin während ihrer Forschungstätigkeit an der Stanford-Universität entwickelt wurde.

Die Daten werden durchgehend in sehr großen, teilweise sogar mehrere Gigabyte großen Dateien gespeichert, welche nur in extrem seltenen Fällen gelöscht, überschrieben oder komprimiert werden; Daten werden üblicherweise angehängt oder ausgelesen. Das Dateisystem ist auch entworfen und optimiert worden, um auf Googles rechnenden Clustern laufen zu können, deren Netzknoten aus billigen Computern bestehen, was bedeutet, dass man der hohen Ausfallrate und dem damit verbundenen Datenverlust individueller Netzknoten vorbeugen muss. Andere Designentscheidungen setzen auf hohe Datendurchsatzraten, auch wenn dies auf Kosten der Latenzzeit geht.

Ein *GFS Cluster* besteht aus einem *Master* und hunderten oder tausenden *Chunkservern*. Die Chunkserver speichern die Dateien, wobei jede Datei in 64 MB große Stücke („*Chunks*“) gespalten ist, ähnlich Clustern oder Sektoren in gebräuchlichen Dateisystemen.

Um Datenverlust zu verhindern wird jede Datei beim GFS mindestens drei mal pro Cluster gespeichert, sodass bei Ausfall eines Chunkservers nur verschwindend geringe Verzögerungen auftreten, bis die Datei von einem anderen Chunkserver geliefert wird. Je nach Bedarf kann die Anzahl auch höher liegen, etwa bei ausführbaren Dateien. Jedem Chunk wird eine eindeutige, 64 Bit lange Kennzeichnung zugewiesen, logische Mappings der Dateien zu den einzelnen Chunks werden beibehalten.

Der Master dagegen speichert normalerweise keine Chunks, sondern vielmehr deren Metadaten, wie etwa Dateinamen, Dateigrößen, ihren Speicherort sowie den ihrer Kopien, welche Prozesse gerade auf welchen Chunk zugreifen etc. Die Master sind auch in der Lage, Kopien von Chunks zu erstellen und auf einen Chunkserver zu speichern, sollte ein solcher ausfallen und die Anzahl eines Chunks unter die Mindestanzahl fallen. Die Metadaten werden aktuell gehalten, indem die Chunkserver regelmäßig Updates an den Master senden („*heart-beat messages*“, auf deutsch etwa: „*Herzschlag-Mitteilungen*“).

## Kritik

Der Quelltext des GFS verbietet mehr als einen Master pro Cluster. Dies hat den Anschein, ein Fehler im System zu sein, die dessen Skalierbarkeit und Zuverlässigkeit begrenzt, da die maximale Größe und Uptime von der Leistungsfähigkeit und Uptime des Masters abhängt, da dieser die Metadaten katalogisiert und fast alle Anfragen durch ihn laufen; Googles Techniker behaupten allerdings, dies sei nicht der Fall und GFS sei sehr wohl skalierbar. Der Master ist im Normalfall der leistungsfähigste Netzknoten im Netzwerk. Um die Ausfallsicherheit zu gewähren gibt es gewöhnlich einen „*Schatten-Master*“, der den Hauptrechner spiegelt und notfalls, sollte der Master einmal ausfallen, einspringt. Engstellen gibt es nur selten, da Clients nur nach Metadaten fragen, die gewöhnlich in einem Cache gespeichert werden – sie sind sehr kompakt, pro Megabyte Daten fallen lediglich einige Kilobyte an. Durch den Einsatz nur eines Hauptrechners verringert sich die Softwarekomplexität drastisch, die erforderlich wäre, beispielsweise für Datenkonsistenz, Ausgleich der Netzbelastung und Systemsicherheit zu sorgen.

## Quellen

- <http://labs.google.com/papers/gfs.html>
- <http://www.baselinemag.com/article2/0,1540,1985047,00.asp>

## Weblinks

- How Google Works (<http://www.baselinemag.com/article2/0,1540,1985047,00.asp>)

Von „[http://de.wikipedia.org/wiki/Google\\_File\\_System](http://de.wikipedia.org/wiki/Google_File_System)“

Kategorien: Google | Dateisystem | Netzwerkarchitektur

---

- Diese Seite wurde zuletzt am 8. Mai 2008 um 15:34 Uhr geändert.
- Ihr Text steht unter der GNU-Lizenz für freie Dokumentation.  
Wikipedia® ist eine eingetragene Marke der Wikimedia Foundation Inc.