

# Performancetests für Web Map und Web Feature Services

Lydia GIETLER

*Dieser Beitrag wurde nach Begutachtung durch das Programmkomitee als „reviewed paper“ angenommen.*

## 1 Einleitung

Geodaten sind teuer, sowohl bei der Ersterfassung als auch bei der Fortführung. Werden die Geodaten über Netzwerke wie das Internet oder Intranets von großen Organisationen einem breiten Publikum zugänglich gemacht, können Mehrwerte geschaffen werden. Beispiele hierfür sind Dienste im Rahmen von E-Government oder E-Business für Bürger und Firmen sowie Zeit- und damit Geldersparnis aufgrund der Rationalisierung von Arbeitsprozessen mithilfe von Geodiensten innerhalb verschiedener Organisationen. Die zentralen Komponenten solcher Dienste stellen so genannte Geo-Webserver dar, für die Produkte von verschiedenen Herstellern angeboten werden. Um die Nutzer nicht an einen Hersteller zu binden, empfiehlt sich in vielen Fällen der Einsatz von Standards wie sie vom Open Geospatial Consortium (OGC) erarbeitet werden.

Zwei derartige OGC Standards sind das Web Map Service (WMS) und das Web Feature Service (WFS). Ein WMS dient der Verbreitung der visuellen Repräsentation von Geodaten, also der Verbreitung von Karten (OGC 2001). Ein WFS ermöglicht die standardisierte Verbreitung von Geobjekten (OGC 2002). Die Ausgabe der Geodaten erfolgt dabei in der Geography Markup Language (GML).

Neben der Funktionalität eines derartigen Dienstes trägt die Performance des Servers zu einem großen Teil zum Erfolg einer Applikation bei. Dementsprechend muss beim Aufbau einer Geodateninfrastruktur für die Verbreitung der Daten über ein Netzwerk die Kapazität der Hardware in Abhängigkeit der erwarteten Anzahl von Usern sowie der Größe der Datenbasis geplant werden. Die Ermittlung der erwarteten Anzahl von Usern ist Applikationsabhängig. Sie kann beispielsweise bei öffentlich zugänglichen Applikationen durch Vergleich mit Vorgänger- und/oder Konkurrenzapplikationen abgeschätzt werden, oder ist im Fall einer Intranetapplikation als Anzahl der Mitarbeiter, die mit dem System arbeiten sollen, bekannt. Um nicht erst nach der Implementierung eines derartigen Dienstes festzustellen, dass die Kapazität der Hardware nicht den Anforderungen der Applikation entspricht, können Performancetests bzw. deren Ergebnisse als Planungsgrundlage sowie zu Vergleichen von verschiedenen Produkten herangezogen werden.

Eine Reihe von Performancetests an verschiedenen Geo-Webserver-Produkten wurde bisher durchgeführt. Viele dieser Tests schmücken sich mit dem Titel Benchmark, jedoch kaum einer entspricht den Kriterien eines solchen (siehe GIETLER 2004a, 22f.). Einzige Ausnahme bildet hierzu der Benchmark des degree WMS (GIETLER 2004b). Der folgende Beitrag beschreibt detailliert diese Performancetests für Web Map Services, sowie Bench-

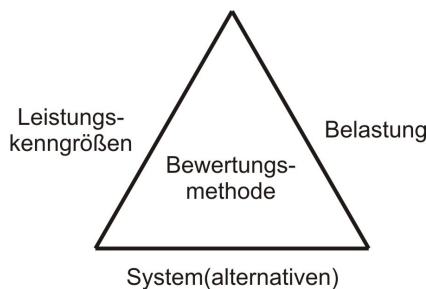
marks für Web Feature Services als Methode für die Leistungsbewertung solcher Produkte. Dabei wird nicht auf Ergebnisse von Durchführungen dieser Benchmarks an konkreten Produkten eingegangen (entsprechendes findet sich in GIETLER 2004a und GIETLER 2004b).

## 2 Was ist ein Benchmark?

Es gibt vier gleichwertige Aspekte, die bei der Leistungsbewertung von Computersystemen zu berücksichtigen sind (BODE & HELLWAGNER 1999). Diese Aspekte sind:

- Ein System bzw. eine Menge von Systemalternativen,
- zu ermittelnde Leistungskenngrößen,
- die Belastung der das System bzw. die Systemalternativen ausgesetzt sind sowie
- die angewandte Methode zur Ermittlung der Leistungskenngrößen.

Ein System bzw. eine Reihe von Systemalternativen sind mithilfe einer bestimmten Bewertungsmethode zu untersuchen. Das Ergebnis dieser Untersuchung sind Leistungskennzahlen. Bei der Ermittlung dieser Maßzahlen wird das System einer bestimmten Belastung, auch Workload genannt, ausgesetzt (vgl. Abb. 1).



**Abb. 1:** Vier gleichwertige Aspekte der Leistungsbewertung von Computersystemen (nach BODE & HELLWAGNER 1999)

Eine spezielle Form der Bewertungsmethode sind Benchmarks. Entsprechend der Definition von BODE & HELLWAGNER (1999) sind Benchmarks standardisierte Testprogramme zur Ermittlung der Performance von bestehenden Computersystemen. Hierbei wird die Leistungsfähigkeit eines Gesamtsystems, also der Kombination aus Hardware, Betriebssystem, Software und Daten, gemessen. Bei Netzwerk-basierten Systemen fließt auch die Übertragungsgeschwindigkeit des Netzwerks in diese Messungen ein.

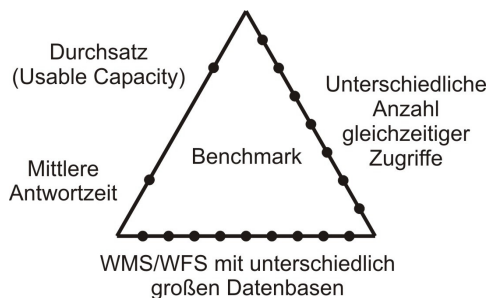
Leistungskenngrößen, die mithilfe von Benchmarks ermittelt werden können, sind die mittlere Antwortzeit sowie der Durchsatz eines Systems. Die Antwortzeit eines Webserver ist die Dauer zwischen Absenden der Anfrage, der Request, vom Client an den Server und Erhalt der Antwort, der Response, am Client. Der Durchsatz ist die Anzahl von derartigen Operationen, die vom System pro Zeiteinheit durchgeführt werden können und wird meist in Operationen pro Stunde angegeben. Eine spezielle Form des Durchsatzes ist die Usable Capacity. Hierbei darf die Antwortzeit der einzelnen Operationen einen vorgegeben Wert nicht überschreiten. Die Belastung, der ein Webserver bei solchen Tests ausgesetzt wird, ist die Anzahl der User, die gleichzeitig eine Webseite anfordern, also die Anzahl der gleichzeitig zu bearbeitenden Requests.

Die Standardisierung von Benchmarks bezieht sich auf ihre Wiederholbarkeit. Diese lässt sich erreichen durch Veröffentlichung der benutzten Hardware, des Betriebssystems, der Software, der Daten sowie einer detaillierten Beschreibung der benutzten Bewertungsmethode. Man unterscheidet zwischen so genannten *Paper and Pencil Tests*, einer textuellen Beschreibung der Tests inklusive der einzugebenden Daten, und implementierten Testprogrammen, die gemeinsam mit den zu benutzenden Daten z. B. im Internet zur Verfügung gestellt werden. In vielen Fällen wird der Ausdruck Benchmark fälschlicherweise für Performancetests, die nicht diesen Kriterien entsprechen, benutzt.

Zwei herstellerunabhängige Organisationen, die Benchmarks für verschiedene Zwecke entwickeln und freigeben, sind die *Standard Performance Evaluation Corporation (SPEC)* und das *Transaction Performance Council (TPC)*. Diese Organisationen stellen unter anderem eine Reihe von Performancetests für Webservices zur Verfügung, jedoch keine Tests für OGC-konforme Geo-Webservices.

### 3 Benchmarks für WMS und WFS

Entsprechend des allgemeinen Modells der Leistungsbewertung von Computersystemen wurde ein spezifisches Modell für Performancetests an OGC-konformen Web Map und Web Feature Services entwickelt (siehe Abb. 2). Die zu untersuchenden Systemalternativen sind jeweils Web Map bzw. Web Feature Services mit unterschiedlich großen Datenbasen. Die entwickelten Bewertungsmethoden entsprechen den Kriterien von Benchmarks nach BODE & HELLWAGNER (1999) und stehen gemeinsam mit den benutzten Daten auf [www.dotGIS.de/performancetests](http://www.dotGIS.de/performancetests) zur Verfügung. Die ermittelten Leistungskenngrößen sind die Usable Capacity sowie die mittleren Antwortzeiten der verschiedenen Systemalternativen. Die unterschiedlichen Belastungen, der die Systeme bei den Untersuchungen ausgesetzt sind, entsprechen einer unterschiedlichen Anzahl von gleichzeitigen Zugriffen, womit eine unterschiedliche Anzahl gleichzeitig mit dem System interagierender User simuliert wird.



**Abb. 2:** Modell der WMS- und WFS-Benchmarks (nach GIETLER 2004a)

Die Benchmarks sind ein Ergebnis der Diplomarbeit der Autorin (GIETLER 2004a). In dieser Arbeit waren aufgrund der Projektrahmenbedingungen die Geo-Webservice Produkte, die Hardware sowie die zu nutzenden Requests vorgegeben. Diese Rahmenbedingungen ergeben sich daraus, dass die Diplomarbeit ein Teilprojekt des vom Deutschen Bundesministerium für Bildung und Forschung geförderten Verbundprojektes *Weiterentwicklung*

von Geodaten war (vgl. GIETLER 2004a). Auch waren zum Zeitpunkt der Durchführung der Diplomarbeit die realen Daten noch nicht freigegeben, daher wurde auf Zufallsdaten zurückgegriffen. Im Folgenden werden die Daten, die Testsoftware sowie die Testpläne für Web Map sowie für Web Feature Services detailliert beschrieben.

### 3.1 Die Datenbasen

Die Testdaten wurden per Zufallsgenerator erzeugt (vgl. GIETLER 2004a, GIETLER 2004b). Jede Datenbasis setzt sich aus drei Objektklassen zusammen, wobei je eine Objektklasse punktförmige, eine weitere Objektklasse linienförmige und die dritte Objektklasse flächenförmige Objekte enthält. Als Attributdaten wurden jedem Objekt je ein Integer-, ein Float- und ein Stringwert beigefügt.

Ursprünglich wurden drei Objektklassen mit jeweils 1.000 punkt-, linien- bzw. flächenhaften Geoobjekten generiert. Diese wurden durch räumliche Spiegelung jeweils verdoppelt, sodass schließlich neun verschiedenen große Datenbasen zur Verfügung standen (siehe Tabelle 1). Dies ermöglicht es, aus den verschiedenen großen Datenbasen jeweils denselben Ausschnitt abzufragen.

**Tab. 1:** Die neun verschiedenen Datenbasen (nach GIETLER 2004b)

Name der Datenbasis	Anzahl der Punktobjekte	Anzahl der Linienobjekte	Anzahl der Flächenobjekte	Summe der Geoobjekte
DB1	1.000	1.000	1.000	3.000
DB2	2.000	2.000	2.000	6.000
DB3	4.000	4.000	4.000	12.000
DB4	8.000	8.000	8.000	24.000
DB5	16.000	16.000	16.000	48.000
DB6	32.000	32.000	32.000	96.000
DB7	64.000	64.000	64.000	192.000
DB8	128.000	128.000	128.000	384.000
DB9	256.000	256.000	256.000	768.000

### 3.2 Die Testsoftware

Als Basis für die Benchmarks wurde die Open Source Software *Apache JMeter 2.0* ausgewählt. Dieses frei zugängliche Tool ermöglicht unter anderem das Messen von Antwortzeiten und der Usable Capacity von Webapplikationen. Für die Validierung der Responses bietet *Apache JMeter* eine Reihe von Möglichkeiten wie beispielsweise Überprü-

fung der Größe der Response, Validierung der Syntax von XML-Dokumenten und Suche nach bestimmten Zeichensequenzen in der Response. Zur Aufbereitung der Ergebnisse stellt *Apache JMeter* eine Anzahl von statistischen Methoden wie zum Beispiel die Durchschnittsberechnung der Antwortzeiten oder die automatische Berechnung des Durchsatzes zur Verfügung.

Konkrete Benchmarks werden mithilfe von *Apache JMeter* als so genannte *JMeter* Testpläne realisiert. Das sind XML-basierte Konfigurationsdateien, die sämtliche notwendigen Informationen für das Senden der Requests, das Validieren der Responses sowie die Aufbereitung der Ergebnisse beinhalten. Zum Durchführen der Benchmarks wird *Apache JMeter* gestartet und ein *JMeter* Testplan geladen. Während des Tests kreiert *Apache JMeter* aus dem Testplan die entsprechenden Requests, sendet diese an den Server, misst die Zeit bis zum Erhalt der Response, validiert die Response entsprechend der Vorgaben des Testplans und bereitet die Ergebnisse, ebenfalls entsprechend der im Testplan enthaltenen Parameter, auf.

### 3.3 Die WMS-Testpläne

Die typischen Operationen eines Users bei Besuch einer Webapplikation die eine von einem WMS generierte Karte inkludiert sind:

- Das Anfordern einer Webseite mit der Karte sowie
- die Änderung des dargestellten Kartenausschnittes (Zoom, Pan).

Für den WMS bedeutet dies den Erhalt mehrerer GetMap-Requests mit unterschiedlichen Werten für die Bounding Boxes nacheinander. Aus diesem Grund wurden für die WMS Benchmarks drei GetMap-Requests mit den in Tabelle 2 dargestellten Parametern festgelegt. Die Responses zu diesen Requests sind jeweils eine Karte im Ausgabeformat JPEG mit einer Größe von 450 mal 450 Pixel. Die Inhalte dieser Karten zeigen jeweils einen Ausschnitt aus der Datenbasis DB1 mit immer größer werdendem Zoomfaktor. In Summe wurden neun Testpläne erstellt, die sich jeweils durch die Werte für den Parameter LAYER, der die abzufragende Datenbasis bestimmt, unterscheiden.

Der Durchlauf eines Testplans dauert jeweils drei Minuten, innerhalb derer die drei GetMap-Requests mehrfach, jeweils nach Erhalt der vorhergehenden Response, an den Server gesandt werden. Dabei wird jedes Mal die Antwortzeit gemessen und die Response wird entsprechend folgender Kriterien evaluiert:

- Die Response ist keine Karte,
- es wird nicht die angeforderte Karte übermittelt, oder
- die Antwortzeit beträgt mehr als 30 Sekunden.

Die Überprüfung, ob eine Karte bzw. die angeforderte Karte übermittelt würde, erfolgt durch Überprüfung der Response auf ihre Größe während der Laufzeit der Tests sowie durch Überprüfung der Zeichensequenzen des Headers der übertragenen Datei. Die maximale Antwortzeit von 30 Sekunden für die Usable Capacity ergibt sich aus Untersuchungen zum Thema Web Usability (NIELSON 2004 und ANDREWS 2004). Diese Untersuchungen besagen, dass nur fünf Prozent der User bereit sind, mehr als 30 Sekunden auf eine Webseite zu warten.

Jeder einzelne Testplan wird mehrfach durchgeführt, wobei jeweils eine unterschiedliche Anzahl von mit dem System interagierenden Usern simuliert wird. Dies beginnt bei einem Zugriff und verdoppelt sich jeweils bis zu 64 gleichzeitigen Zugriffen.

**Tab. 2:** Die drei GetMap-Requests der WMS-Benchmarks (nach GIETLER 2004a)

Parameter	GetMap 1	GetMap 2	GetMap 3
SERVICE	WMS	WMS	WMS
VERSION	1.1.1	1.1.1	1.1.1
REQUEST	GetMap	GetMap	GetMap
LAYERS	m:kat1,m:lin1,m:poi1	m:kat1,m:lin1,m:poi1	m:kat1,m:lin1,m:poi1
STYLES	default,default,default	default,default,default	default,default,default
SRS	EPSG:4326	EPSG:4326	EPSG:4326
WIDTH	450	450	450
HEIGHT	450	450	450
FORMAT	image/jpeg	image/jpeg	image/jpeg
BBOX	11.5437,48.1535, 11.5487,48.1573	11.5440,48.1540, 11.5465,48.1559	11.5450,48.1545, 11.5462,48.1554

### 3.4 Die WFS-Testpläne

Das Userverhalten bei der Nutzung eines WFS kann nicht so einfach identifiziert werden, da es in erster Linie von Zweck der Anforderung der Daten abhängt. Ist dieser Zweck der Download der Daten für eine spätere Nutzung, wie beispielsweise beim Online-Kauf von Geodaten, spielt die Performance des Servers keine ganz so große Rolle. Sollen die Daten hingegen sofort am Client weiterverarbeitet werden, ist die Antwortzeit wiederum ein wichtiges Kriterium für den Erfolg der WFS-Applikation.

Werden die Daten am Client dargestellt, liegt der Unterschied zur Nutzung eines WMS liegt in erster Linie darin, dass für die Änderung des Kartenausschnittes keine neuen Requests an den Server gesandt werden. Anstelle dessen werden anfangs Geodaten, die eine größere Fläche abdecken, heruntergeladen und die weitere Auswahl der darzustellenden Daten erfolgt am Client.

Dementsprechend wurde für die WFS-Testpläne nur ein GetFeature-Request festgelegt. Die Response zu diesem Request entspricht den in der Requests *GetMap 1* der WMS Benchmarks abgefragten Geodaten. Das Ausgabeformat ist in diesem Fall GML und neben der Geometrie sind auch sämtlichen Attributdaten in der Response enthalten. Des

Weiteren entsprechen die WFS-Testpläne jenen für einen WMS mit Ausnahme der Fehlerkriterien, die hier folgendermaßen aussehen:

- Die Response ist nicht in GML codiert,
- es werden nicht die angeforderten Geodaten übermittelt, oder
- die Antwortzeit beträgt mehr als 120 Sekunden.

Um zu überprüfen ob die angeforderten Daten in GML codiert übermittelt werden, wird die XML-Validierungsfunktion von *JMeter* mit der Suche nach Zeichenketten die einzelne GML-Objekte beschreiben kombiniert. Da es sich bei Nutzung eines WFS um den expliziten Download von Geodaten handelt, kann davon ausgegangen werden, dass User eine höhere Antwortzeit in Kauf nehmen als bei Besuch einer Webseite. Aus diesem Grund wurde die maximale Antwortzeit für die Usable Capacity auf 120 Sekunden erhöht.

Auch dieser Testplan wurde mehrfach durchgeführt um eine unterschiedliche Anzahl von mit dem System interagierenden Usern zu simulieren. In diesem Fall wurde begonnen bei einem Zugriff jeweils verdoppelt bis insgesamt 16 gleichzeitige Zugriffe abzuarbeiten waren.

## 4 Diskussion

In den vorgestellten Benchmarks werden immer wieder dieselben drei bzw. die eine Request an den Server gesandt. Dies entspricht nicht dem Verhalten von realen Usern, da jeder User an anderen Kartenausschnitten interessiert ist. Auch ein Zoom-Out wird durch diese Benchmarks nicht simuliert. Aus diesem Grund eignen sich diese Performancetests bzw. die entsprechenden Ergebnisse (GIETLER 2004a, GIETLER 2004b) nur als grobe Planungsgrundlage für neue Applikationen. Es wäre dementsprechend sinnvoll eine zusätzliche Anzahl von verschiedenen Requests, die auch in den abgefragten Objektklassen variieren, in die Tests inkludiert werden.

Auch sollten Untersuchungen mit anderen Datensätzen durchgeführt werden, um zu repräsentativeren Aussagen zu kommen. Beispielsweise ist es vorstellbar, dass bei Datensätzen, deren einzelne Objekte im Schnitt mehr Vertices besitzen als die hier benutzten, die Antwortzeiten ansteigen. Ähnliches gilt auch für die Anzahl der abgefragten Attribute bei den WFS-Benchmarks.

Auch die Beschränkung auf nur ein Ausgabeformat im Fall der WMS-Benchmarks ist nicht repräsentativ. Hier könnten Untersuchungen mit verschiedenen Ausgabeformaten zeigen, welche Formate bei welchen Produkten die beste Performance erzielen.

Wie erwähnt sind die Benchmarks ein Ergebnis der Diplomarbeit der Autorin und mussten daher den Projektrahmenbedingungen entsprechen. Eine Anpassung der Testpläne, etwa hinsichtlich der Parameter für die Requests, oder hinsichtlich der Anzahl der gleichzeitigen Zugriffe bei Nutzung leistungsfähigerer Hardware, ist aufgrund der entwickelten Struktur der Testpläne problemlos durchzuführen.

## 5 Fazit

Die beschriebenen Performancetests für OGC konforme Web Map und Web Feature Services sind gemeinsam mit den verwendeten Daten auf [www.dotGIS.de/performancetests](http://www.dotGIS.de/performancetests) frei zugänglich. Sie entsprechen den Kriterien von Benchmarks nach BODE & HELLWAGNER (1999). Aufgrund der standardisierten Interfaces können sie auf beliebige, diese Interfaces implementierende Produkte angewandt werden. Erste Ergebnisse der Durchführung dieser Performancetests (veröffentlicht in GIETLER 2004a und GIETLER 2004b) ermöglichen eine grobe Abschätzung der notwendigen Hardwarekapazität für neue Applikationen. Auch können diese Tests dazu benutzt werden, die Leistungsfähigkeit verschiedener Geo-Webservice-Produkte zu vergleichen, sofern die Tests unter denselben Bedingungen durchgeführt werden.

## Literatur

- ANDREWS, K. (2004): Information Architecture and Web Usability. <http://courses.iicm.edu/iaweb/iaweb.pdf>.
- BODE, A. & H. HELLWAGNER (1999): Leistungsbewertung und Fehlertoleranz. In: RECHENBERG, P. & G. POMBERGER (Hrsg.): Informatik-Handbuch, 409-422. Carl Hanser, München.
- GIETLER, L. (2004a): Verifikation der Auswahl von Open Source Produkten für OpenGIS Webservices. [www.dotgis.de/performancetests/download/gietler\\_verifikationOpenSourceOWS.pdf](http://www.dotgis.de/performancetests/download/gietler_verifikationOpenSourceOWS.pdf).
- GIETLER, L. (2004b): Benchmarking deegree WMS. In: Angewandte Geoinformatik 2004. STROBL, J., BLASCHKE, T. & G. GRIESEBNER (Hrsg.): Angewandte Geoinformatik 2004, 172-177. Wichmann, Heidelberg.
- NIELSEN, J. (2004): The Need For Speed – Update January 2004: Still Mostly Slow. <http://www.useit.com/alertbox/9703a.html>.
- OGC (2001): Web Map Service Implementation Specification. OGC 01-068r3, Open GIS Consortium Inc. <http://www.opengis.org/docs/01-068r2.pdf>.
- OGC (2002): Web Feature Service Implementation Specification. OGC 02-058, Open GIS Consortium Inc. <http://www.opengis.org/docs/02-058.pdf>.