



Evaluierung und Monitoring von Dienstqualität (Quality of Service) dargestellt am Beispiel der Marinen Dateninfrastruktur Deutschland (MDI-DE)

CHRISTIAN SEIP, Rostock

Keywords: spatial web services, INSPIRE, evaluation, monitoring, service quality

Zusammenfassung: Service-orientierte Architekturen (SOA) sind heute weit verbreitet. Deutschland entwickelte eine marinespezifische Service-orientierte Dateninfrastruktur (MDI-DE, Marine Dateninfrastruktur Deutschland) von 2010 bis 2013. Die Dienste in der MDI-DE tragen zur Erfüllung von Berichtspflichten für verschiedene europäische und nationale Regularien bei. Die Dienste der MDI-DE ebenso wie andere Dienste, die z. B. von INSPIRE betroffen sind, müssen spezifische Anforderungen an Performanz, Verfügbarkeit und Konformität (Quality of Service, QoS) erfüllen. Obwohl SOA ein wichtiger Bereich in der Forschung ist, gibt es nur sehr wenige Publikationen und Studien über QoS, insbesondere hinsichtlich der INSPIRE-Anforderungen.

Die Dienste der MDI-DE wurden mit verschiedenen Tools analysiert. Die Analyse erstreckte sich auf die Kontrolle der Qualität von Konformität und Leistung (Performanz). Aufgrund teils widersprüchlicher Ergebnisse und einiger Unzulänglichkeiten der unterschiedlichen Werkzeuge kann vereinfacht festgestellt werden, dass je mehr Werkzeuge verwandt werden, desto aussagekräftiger das Ergebnis ist. Die Servicequalitäten waren nicht kohärent, wenn sie mit verschiedenen Werkzeugen gemessen wurden. Die Studie zeigt, dass für das Erreichen objektiver und nachvollziehbarer Messungen von QoS-Parametern als Teil von INSPIRE noch einige Herausforderungen bestehen.

Summary: *Evaluation and Monitoring of Service Quality illustrated by the Example of the German Marine Spatial Data Infrastructure (MDI-DE).* Service oriented architectures (SOA) are widely used nowadays. As the name implies such architectures rely on services. Germany developed a marine specific service oriented data infrastructure (MDI-DE, Marine Dateninfrastruktur Deutschland) from 2010 to 2013. The services in MDI-DE can contribute to fulfilling reporting commitments for various European and national legislation. The services of MDI-DE (just like other services affected, for instance, by INSPIRE) have to meet specific requirements regarding performance, availability and conformity (quality of service, QoS). Although SOA is an important field in scientific research there are only very few publications and studies available on QoS, especially regarding INSPIRE requirements.

The services of MDI-DE were analysed regarding their QoS using various existing tools. The paper discusses the tools' usefulness to reflect whether the services performance or conformity need improvement. Due to partly contradicting results and the failure of some tools concerning specific aspects, it can at least be stated that the more tools are applied the more conclusive the outcome. Furthermore, the service quality results were not coherent when measured with different tools. The study indicates that achieving objective and comparable measurements of QoS parameters as part of the INSPIRE implementation will still be challenging.

1 Einführung

Der Aufbau von regional begrenzten Geodateninfrastrukturen (GDI) wie die des Bundes, der Länder oder der Landkreise ist im vollen

Gange. Um hingegen Fachdaten aus dem Küsten- und Meeresbereich für Wissenschaft, Planung, Öffentlichkeit, Politik und Verwaltung gemäß den Anforderungen an eine integrierte europäische Meerespolitik bereitzustellen, ist

es notwendig, thematisch begrenzte GDI-Regionen übergreifend aufzubauen. Das BMBF förderte von 2010 bis 2013 den Aufbau einer solchen marinen Dateninfrastruktur (MDI) mit dem Projekt *MDI-DE*. Das Ziel der MDI-DE war es, sowohl vorhandene technische Lösungen wie das Nordsee-Ostsee-Küsteninformationssystem (NOKIS) und die Geodateninfrastruktur des Bundesamtes für Seeschifffahrt und Hydrographie (GDI-BSH) als Grundstock zu integrieren, als auch Informationen aus allen Bereichen des Küsteningenieurwesens und Küstengewässerschutzes, des Meeresumweltschutzes, des marinen Naturschutzes, der Raumordnung und der Küstenforschung zusammenzuführen.

Eine GDI wie die MDI-DE besteht aus einer Vielzahl von verteilten Diensten und ist somit von deren Verfügbarkeit, Leistung und Einhaltung von Normen und Standards abhängig. Die MDI-DE-Umgebung ist insofern ein gutes Testbeispiel zur Prüfung von Dienstqualitäten. Nicht verfügbare Dienste sind negativ für alle aufsetzenden Anwendungen in einer GDI, z. B. Portale, und hinderlich für die Nutzer. Aber auch wenn Dienste verfügbar sind, ist ihre Leistung für eine zufriedenstellende Nutzung sowie die Einhaltung von regulatorischen Anforderungen, z. B. INSPIRE, von entscheidender Bedeutung. Darüber hinaus müssen Dienste-Anfragen und Dienste-Antworten abgestimmten Anforderungen genügen, um sicherzustellen, dass beispielsweise ein Dienst in das lokale GIS eines Anwenders integriert werden kann. Daher ist Konformität ein weiterer wichtiger Faktor, der betrachtet werden sollte und Grundlage für Interoperabilität ist. LUPP (2008) unterstreicht, dass Standardisierung von Diensten Interoperabilität ermöglichen.

Interoperabilität wiederum lässt sich unter anderem in die zwei Hauptbereiche organisatorische und technische Interoperabilität unterteilen. Innerhalb der technischen Interoperabilität werden die beiden Merkmale semantische und syntaktische Interoperabilität unterschieden (RÜH 2014). Im Kontext der EU Richtlinie INSPIRE werden die Ebenen der Interoperabilität weitergehend definiert. Das European Interoperability Framework (EIF) definiert vier Ebenen der Interoperabilität (JAKOBSSON 2012): rechtlich, organisato-

risch, semantisch und technisch. In Bezug auf Konformität hinsichtlich INSPIRE gilt, dass Konformität mit INSPIRE auf der einen Seite Konformität mit dessen Daten- bzw. Metadaten-Spezifikationen und auf der anderen Seite die Übereinstimmung mit den Spezifikationen bezüglich der Servicequalität bedeutet.

1.1 *Quality of Service*

Die Evaluierung von Diensten lässt sich in zwei Kategorien einteilen: Performanz und Verfügbarkeit auf der einen Seite und Konformität auf der anderen Seite mit vielen Aspekten auf beiden Seiten. Diese Kategorien werden für die fünf verschiedenen INSPIRE Diensttypen Discovery, View, Download, Transformation und Invoke Spatial Data Services (INSPIRE 2007) konkretisiert.

Dienstqualität (Quality of Service, QoS) wird durch PARASURAMAN et al. (1985) als Diskrepanz zwischen den Erwartungen der Konsumenten und den tatsächlich für einen Dienst festgelegten Spezifikationen definiert. Ist die Spezifikation nutzerorientiert, sollte der Dienst den Erwartungen des Nutzers entsprechen, also das, was er erwartet, sollte mit dem, was er wahrnimmt, übereinstimmen. Obwohl diese Definition aus dem Marketingbereich stammt, gilt sie auch für Web-Dienste (und wahrscheinlich viele andere Dienstypen). FRANKEN (1996) unterstreicht mit einem technischen Hintergrund dies durch seine Definition von Dienstqualität als vom Benutzer wahrgenommene Performanz. Dadurch stellt er Performanz bzw. Leistung auf die gleiche Ebene wie QoS und setzt diese beiden Begriffe praktisch gleich. INSPIRE (2007) setzt QoS zunächst auch mit Performanz gleich, indem nach Artikel 16 die Durchführungsbestimmungen unter anderem Mindestleistungskriterien für Dienste festlegen. Wie jedoch bereits erwähnt, ist im technischen Sinne die Performanz nur ein Aspekt der Dienstqualität.

Ein Nutzer erwartet zuerst, dass ein Dienst verfügbar ist und genutzt werden kann. Eine Analyse zur Auffindbarkeit von WMS über Google ergab eine Nichtauffindbarkeit von 30% (MÜLLER & MANDERY 2010). Ein weiteres Qualitätsmerkmal eines Dienstes sowie einer GDI ist Performanz bzw. Leistung, weshalb dies auch eine INSPIRE-Anforderung

ist. YANG & EVANS (2008) prägten den Begriff „Network GIS Performance“ (NGP), der sowohl die effiziente Nutzung von Netzwerk-GIS-Ressourcen als auch die Geschwindigkeit eines Netzwerk-GIS umfasst. INSPIRE beschreibt Performanz als Geschwindigkeit, mit der eine Anfrage an einen Dienst abgeschlossen werden kann (INSPIRE CONSOLIDATION TEAM 2007). YANG & EVANS (2008) sagen, dass ein durchschnittlicher Benutzer ca. 8 Sekunden auf eine Web-Antwort zu warten bereit ist (8-Sekunden-Regel). Nach MÜLLER & MANDERY (2010) sind 86% der verfügbaren Dienste in der Lage, die Grafiken mit einer Auflösung von 400 x 400 Pixeln in weniger als einer Sekunde zu liefern. Allerdings ist eine Performanz-Messung nur sinnvoll, wenn sie über einen längeren Zeitraum durchgeführt wird (MÜLLER & MANDERY 2010).

CIBULKA (2013) und das INSPIRE CONSOLIDATION TEAM (2007) erwähnen darüber hinaus Interoperabilität als Attribut für Dienstqualität. Ein interoperabler Dienst kann mit den anderen Diensten zusammenarbeiten, die die gleichen Normen und Standards einhalten. Somit ist die Konformität der zweite Hauptbereich für die Dienstevaluierung.

CIBULKA (2013) listet eine Vielzahl weiterer Attribute auf, die Leistung beeinflussen:

- Zuverlässigkeit*: Fähigkeit eines Service, eine bestimmte Qualität einzuhalten.
 - Skalierbarkeit: Fähigkeit eines Systems, die Rechenkapazität nach aktuellen Anforderungen zu erhöhen.
 - Kapazität*: Obergrenze der Anzahl von gleichzeitigen Anfragen mit einer garantierten Performanz.
 - Genauigkeit: Vom Dienst generierte Fehlerrate.
 - Zugänglichkeit*: Fähigkeit eines Dienstes, der Anfrage eines Clients nachzukommen.
- Mit einem Stern sind Attribute gekennzeichnet, die als Mindestleistungskriterien für die INSPIRE Network Services verwendet werden (INSPIRE CONSOLIDATION TEAM 2007).

1.2 Messung von QoS-Parametern

Zur Messung von QoS-Parametern, speziell der relevanten Kriterien Performanz, Verfügbarkeit und Konformität, finden sich einige

Ansätze. CIBULKA (2013) entwickelte ein eigenes Werkzeug für seine Messungen.

GIULIANI et al. (2013) verwendeten *JMeter Skripte*, die vom FOSS4GWMS (Grid-enabled Web Map Service) Benchmark im Jahr 2009 stammten und ausgebaut und angepasst wurden (AIME & MCKENNA 2009). Allerdings spiegelt dies nur die Messungen von einem einzigen Werkzeug (*JMeter*) wider, so dass es keinen Vergleich gibt, der die Genauigkeit der Ergebnisse überprüft. Das gleiche gilt für HORÁK et al. (2009), die ausschließlich *WAPT (Web Application Load, Stress and Performance Testing)* von Softlogica LLC für ihre Messungen verwendeten. Im Zusammenhang mit INSPIRE weist DRERUP (2010) jedoch darauf hin, dass *WAPT* nicht nutzbar ist, um die INSPIRE-Anforderungen zu erfüllen, da es nicht in der Lage sei, die Zeit zum Download einzeln zu messen. Es kann nur die Gesamtantwortzeit messen, die auch die Zeit beinhaltet, die der Server braucht, um eine Anfrage zu bearbeiten. INSPIRE verlangt jedoch, dass jede Zeitspanne separat gemessen wird.

Obwohl sich zahlreiche Arbeiten zu SOA finden, stellen GIULIANI et al. (2013) in Bezug auf Servicequalität fest, dass trotz der Wichtigkeit wenig Forschung zu diesem Thema veröffentlicht worden sei.

HECKEL & MARIANI (2005) schlugen vor, dass ein Suchdienst nur andere Dienste registriert bzw. aufnimmt, wenn diese einen Konformitäts-Test bestehen. Aber dieser Vorschlag bezieht sich nur auf Suchdienste und Konformität und somit nicht auf Verfügbarkeit und Performanz. BOZKURT et al. (2010) zeigen neun verschiedene Funktionstesttechniken und -ansätze, um Web-Dienste zu testen. Diese Techniken und Ansätze zielen jedoch generell auf Web-Dienste ab und nicht speziell auf Dienste einer GDI.

2 Evaluierung und Testwerkzeuge

Bei der Evaluierung stellen Testwerkzeuge eine Anfrage an einen Dienst und messen dann entweder die Zeit bis zur Antwort (Performanz) oder überprüfen, ob die Antwort bestimmten Vorgaben entspricht (Konformität). Insbesondere bei der Leistung gibt

es jedoch einiges zu beachten. Beispielsweise können zeitgleiche Anfragen anderer Nutzer die Antwortzeit beeinflussen. Aus diesem Grund beschäftigt sich zunächst Abschnitt 2.1 mit solchen Überlegungen. Das dieser Studie zugrunde liegende Test-Setup wird zu Beginn von Abschnitt 3 beschrieben.

In Abschnitt 2.2 werden Werkzeuge vorgestellt, die Konformität und/oder Leistung und Verfügbarkeit überprüfen. Dabei wird hauptsächlich freie Software (Free and Open-Source-Software, FOSS) beschrieben. Eine Ausnahme bildet der *serviceMonitor*, der in der *sdi.suite* enthalten ist, die im MDI-DE-Projekt verwendet wird. Nach FRÖLUND & KOISTINEN (1998) erwartet der Benutzer vor allem einen Dienst, der verfügbar ist und schnell Daten liefert. Abschnitt 2.2.2 stellt Testwerkzeuge zur Überwachung der Leistung und Verfügbarkeit und Abschnitt 2.2.1 einen Überblick über Werkzeuge zur Überprüfung der Konformität als Grundlage der Interoperabilität vor. Eines der Werkzeuge, die GDI-DE *Testsuite*, erlaubt den Test beider Eigenschaften und wird separat in Abschnitt 2.2.3 vorgestellt.

2.1 Vorüberlegungen und mögliche Messprobleme

Gemäß Abschnitt 1.1 sind unter anderem Performanz und Verfügbarkeit Merkmale für die Dienstqualität. CIBULKA (2013) weist darauf hin, dass diese Eigenschaften von der Anzahl der Benutzer, der Art der Operation (GetMap, GetCapabilities, GetFeature etc.), dem Volumen der verarbeiteten Daten, und einer Reihe von anderen Parametern (Hardware, Software, Netzwerk, etc.) abhängig sind. Daraus ergeben sich einige Fragen an die Messungen:

- 1) Wo wird gemessen?
- 2) Welcher Anfragetyp wird verwendet, z. B. GetMap?
- 3) Wie groß ist das Datenvolumen der Antwort?
- 4) Unter welchen Umständen wird gemessen, z. B. Tageszeit?
- 5) Welche Eigenschaften hat der Server, z. B. bezüglich Netzanbindung?
- 6) Welche Eigenschaften hat der Rechner, von dem aus getestet wird, z. B. Hardware?

Die INSPIRE-Anforderungen an Performanz und Verfügbarkeit wurden unter anderem im INSPIRE NETWORK SERVICES DRAFTING TEAM (2013) definiert. Für eine 470 KB große Grafik sollte die Reaktionszeit bis zur initialen Antwort auf eine GetMap-Anfrage an einen Darstellungsdienst maximal 5 Sekunden in einer normalen Situation betragen, d. h. einem Zeitraum außerhalb der Spitzenlast, also in 90% der Gesamtzeit. Damit legt INSPIRE für die Fragen 1) bis 3) konkrete Anforderungen fest.

Hinsichtlich Frage 1) muss sich das Testwerkzeug für INSPIRE-Konformität auf dem zu testenden Server installieren lassen. Abgesehen von *MapMatters* und *serviceMonitor* ist dies bei allen vorgestellten Werkzeugen möglich.

Für die Frage 2) legt INSPIRE eine GetMap-Anfrage fest. Zwei der verwendeten Werkzeuge, nämlich die GDI-DE *Testsuite* und der *serviceMonitor*, bieten die Möglichkeit, eine Anfrage selbst zu formulieren. Der *Service Status Checker (SSC)* bietet eine solche Möglichkeit nicht und erzeugt selbst eine nicht veränderbare GetMap-Anfrage.

Zum Datenvolumen (Frage 4) hat CIBULKA (2013) festgestellt, dass die Performanz auch vom Kartenmaßstab sowie von der Lage abhängt. Dies bedeutet, dass das Anfragen von Informationen mit wenigen Daten, z. B. einige wenige Punkte statt vieler komplexer Polygone, zu geringeren Reaktionszeiten führt. Mit etwas Aufwand wäre es möglich, dieses Verhalten von räumlichen Webdiensten mit dem *serviceMonitor* und mit der GDI-DE *Testsuite* zu untersuchen und somit die Aussage von CIBULKA (2013) zu überprüfen. Beide Werkzeuge ermöglichen es, mehrere Tests auf einem Dienst mit unterschiedlichen Anfragen anzuwenden. So können verschiedene Maßstäbe und Raumausschnitte verwendet werden, um den Dienst zu testen. *MapMatters* und der *SSC* sind dazu nicht in der Lage, da Anfragen nicht vom Benutzer definiert werden können.

Die Anzahl der Benutzer beeinflusst die Dienstqualität, da Nutzer Last auf einem Server erzeugen. Die Anzahl der Benutzer variiert mit der Tageszeit, wie DRERUP (2010) ausführt (Frage 4). Auch die Verfügbarkeit ist zeitabhängig, zum Beispiel eingeschränkt

von typischerweise nachts durchgeführten Wartungsarbeiten. Während der Zeitpunkt der Messungen von keinem der Werkzeuge spezifiziert werden kann, gibt es Werkzeuge, die messen und auch Last erzeugen können (2.2.2).

Die Fragen 5 und 6 haben zwar sicherlich den größten Einfluss auf die Leistung, können aber beim Messen nicht beeinflusst werden und müssen daher als gegebene Größen angesehen werden, die es zu dokumentieren gilt.

2.2 Testwerkzeuge

Abschnitt 2.2.1 gibt eine Übersicht zu Werkzeugen zur Überprüfung der Konformität. Abschnitt 2.2.2 stellt Testwerkzeuge zur Überwachung der Leistung und Verfügbarkeit vor. Das Testwerkzeug, das beide Parameter testen kann, wird in Abschnitt 2.2.3 erläutert (GDI-DE *Testsuite*).

2.2.1 Überprüfung der Konformität

Das Open Geospatial Consortium (OGC) bietet eine Testumgebung zur Kontrolle von Software auf Konformität mit OGC-Spezifikationen (Compliance and Interoperability Testing Initiative, CITE) (TEAMENGINE 2015).

Während der *Neogeo WMS INSPIRE-Tester* (NEOGEO 2015) sich nur auf INSPIRE-Darstellungsdienste konzentriert, bietet der offizielle INSPIRE *Geoportal Metadata Validator* (EUROPEAN COMMISSION 2013) Tests für Metadaten, Suchdienste und Download-Dienste.

2.2.2 Überwachung der Leistung und Verfügbarkeit

MapMatters (MAPMATTERS 2015) ist ein von geOps entwickeltes Portal, dessen Hauptaufgabe die Katalogisierung von Kartendiensten (WMS) ist. Es bietet dafür eine Textsuche sowie eine Suche nach deren geographischer Ausdehnung. *MapMatters* verwendet sowohl Robots als auch Benutzereingaben, um die Datenbank mit Kartendiensten zu füllen. Zusätzlich wird eine Monitoring-Komponente zur Aufzeichnung der Verfügbarkeit und der Antwortzeiten angeboten.

Der *serviceMonitor* von con terra überwacht verschiedene Dienstarten wie WMS, Web Feature Service (WFS), Web Coverage Service (WCS) sowie ArcIMS und ArcGIS und INSPIRE Network Services. Die Zeitintervalle sowie die Schwellenwerte in Bezug auf Verfügbarkeit und Performanz, die nicht unterschritten werden sollen, können konfiguriert werden. Wenn ein Dienst unter die vorgeschriebenen Qualitätsanforderungen fällt, kann ein Alarm per E-Mail oder SMS ausgelöst werden. Für INSPIRE-Netzwerkdienste sind die INSPIRE Quality-of-Service-Parameter bereits vorkonfiguriert. Außerdem können Daten über die Verfügbarkeit und Performanz der Dienste als Berichte mit Diagrammen oder Statistiken angezeigt und im MS Excel-Format exportiert werden.

Die Ziele des SSC vom US-amerikanischen Federal Geographic Data Committee (FGDC) werden von ANTHONY & NEBERT (2012) wie folgt zusammenfasst:

- Überwachung und Bewertung räumlicher Web-Dienste,
- Benachrichtigung der Besitzer bei Problemen mit Diensten,
- Entdeckung von Performanz-Problemen,
- Ermittlung von Funktionsstörungen und
- Austausch von Testergebnissen.

Listen mit Diensten können von Katalogen oder von einem registrierten Benutzer über Atom-Feeds bereitgestellt werden. Danach können die Dienste dann in einem bestimmten Zeitintervall geprüft (zweimal pro Tag) und die Ergebnisse in einer Datenbank archiviert werden. Die Ergebnisse sind über einen Web-Dienst, über eine API mit GET-Anfragen, die Ausgaben im JSON-Format erzeugt, und über HTML-Service-Berichte verfügbar.

Das INSPIRE CONSOLIDATION TEAM (2007) definiert Kapazität als die Grenze der Anzahl von gleichzeitigen Anfragen, die mit garantierter Leistung zur Verfügung gestellt werden können. Das bedeutet, dass es eine starke Korrelation zwischen der Performanz und der Menge an gleichzeitigen Benutzern, also Anfragen, gibt. BARANSKI et al. (2011) zeigten, dass ihr Server-Aufbau höchstens fünf gleichzeitige Anfragen erlaubte, um der INSPIRE-Anforderung einer Reaktionszeit von weniger als fünf Sekunden zu entsprechen (3.1). Lei-

der ist keines der bisher dargestellten Werkzeuge in der Lage, eine bestimmte Anzahl von gleichzeitigen Anfragen zu simulieren.

Die Website SoftwareQATest.com von RICK HOWER beschäftigt sich mit Qualitätssicherung und Tests. Sie listet 66 Werkzeuge in der Kategorie *Load and Performance Test Tool*, von denen *Proxy Sniffer*, *WAPT (Web Application Load Stress und Leistungstests)*, *Apache JMeter* und *ApacheBench* Beispiele sind, die bereits auf räumliche Web-Dienste angewandt wurden. Auch wenn es eine Vielzahl solcher Last-Tools für spezielle Anforderungen wie das in BARANSKI et al. (2011) vorgestellte Hybrid Cloud-Konzept gibt, sind weitere Anpassungen oder neue Implementierungen für den Einsatz in GDI erforderlich.

2.2.3 Die GDI-DE Testsuite

Die GDI-DE ist die nationale Geodateninfrastruktur Deutschlands. Eine ihrer Komponenten ist die GDI-DE *Testsuite* (TESTSUITE 2015), die von der Koordinierungsstelle GDI-DE beim Bundesamt für Kartographie und Geodäsie (BKG) in Abstimmung mit dem Lenkungsausschuss GDI-DE zwischen Ende des Jahres 2010 und Mitte 2011 entwickelt wurde. Die *Testsuite* kann als Online-Web-Anwendung oder über eine API genutzt werden oder lokal installiert und verwendet werden. Ho-

GREBE (2012) fasst die Ziele der GDI-DE *Testsuite* wie folgt zusammen:

- Unterstützung der Interoperabilität innerhalb der GDI-DE (gemeinsames Werkzeug).
- Unterstützung des Umsetzungsprozesses der GDI-DE und INSPIRE (gemeinsames Verständnis der einschlägigen Normen und Spezifikationen).
- Unterstützung des INSPIRE Monitoring (Konformitätsindikatoren)

3 Evaluierung der MDI-DE-Dienste

Dieser Abschnitt analysiert die im MDI-DE-Portal verfügbaren Dienste hinsichtlich ihrer Performanz (3.2) und Konformität (3.1). Alle in Abschnitt 2 vorgestellten Werkzeuge wurden mit Ausnahme von *MapMatters* (keine Dienstleistungsstatistiken) benutzt. Die meisten Werkzeuge wurden von ihren Herstellern gehostet. Eine Ausnahme bildete der *service-Monitor* der *sdi.suite*, der auf einem Server beim BSH gehostet wurde. Damit musste bei den anderen Servern gegen die INSPIRE-Anforderung verstoßen werden, auf dem Server selbst zu messen. Jedoch war dies nicht möglich, da ein Zugriff auf die Server nicht möglich war. Allerdings entsprechen diese Tests

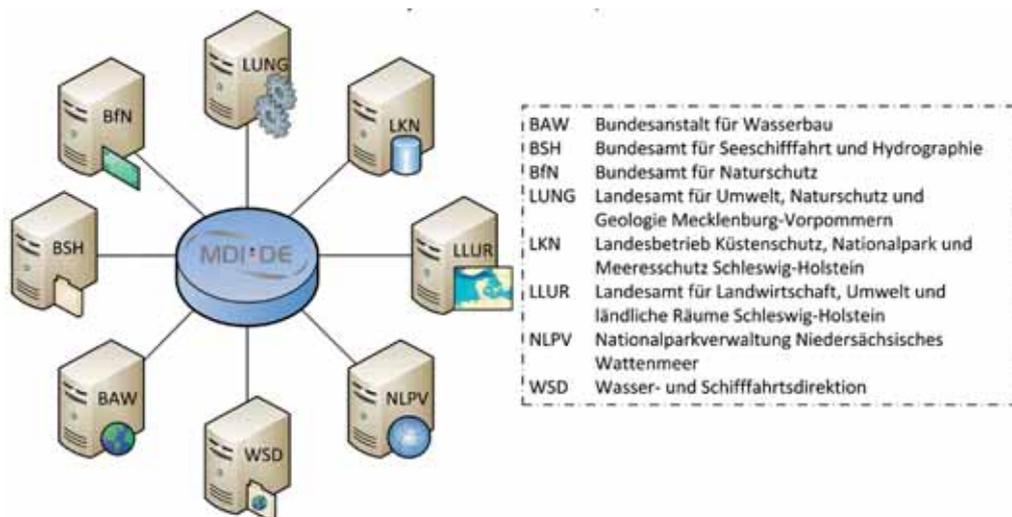


Abb. 1: Für Überwachung und Testen genutzte MDI-DE-Infrastrukturknoten.

damit eher der Nutzersicht: Es interessiert weniger, wie schnell ein Dienst auf dem Server reagiert, sondern eher wie schnell eine Antwort beim Nutzer eintrifft.

Last-Tools wurden nicht verwendet. Zum einen ist nicht bekannt, wie viele Benutzer den/die Dienst/e gerade während eines Testlaufs benutzen. Bei der Erzeugung gleichzeitiger Anfragen kann daher nicht bestimmt werden, wie groß die Gesamtsumme aller gleichzeitigen Anfragen ist. Zum anderen bedeutet künstlich erzeugte Last und erhöhte Bandbreite ggf. höhere Kosten für die beteiligten Institutionen. Der Testaufbau (Abb. 1) beinhaltet acht Institutionen und deren Infrastrukturlösungen mit neun Serverinstanzen und drei verschiedenen Arten von Karten-Server-Software. Insgesamt bieten sie 27 Dienste mit 882 Layern, von einem bis zu 397 Layern pro Dienst. Von den 27 Diensten sind 23 WMS und vier WFS. Die verwendeten WMS-Versionen waren annähernd gleich verteilt: elf mit der Version 1.3.0 und zwölf mit Version 1.1.1. Von den WFS nutzen drei die Version 2.0.0 und einer die Version 1.0.0.

Die GetMap-Anfragen wurden von der GDI-DE *Testsuite* generiert. Diese wurden auch für den *serviceMonitor* genutzt, so dass beide die gleichen Anfragen nutzen und somit vergleichbar sind. Zusätzlich wurden GetCapabilities-Anfragen verwendet, um in der

Lage zu sein, die Anfragetypen zu vergleichen. Leider bietet SSC eine solche Möglichkeit nicht, denn es erzeugt, wie oben erwähnt, selbst eine GetMap-Anfrage, die nicht verändert werden kann.

Über weitere die Messung beeinflussende Charakteristika, z. B. Netzanbindungen und die unterschiedlichen Latenzen und Hardware der Server, ist nichts bekannt.

3.1 Konformität mit INSPIRE und OGC

Detaillierte Fehleranalysen für jeden Dienst und jedes Werkzeug sind aufgrund des Umfangs der Dienste nicht möglich. Daher werden in diesem Abschnitt nur Zusammenfassungen für die Konformität mit den Spezifikationen des OGC (3.1.1) und von INSPIRE (3.1.2) vorgestellt.

3.1.1 OGC

Die OGC *TEAM Engine* produzierte mehrere falsch-negative Ergebnisse, die die Anzahl der tatsächlichen Fehler beeinflussten. Dienste bestanden die Tests auf vorgeschriebene Ausgaben bei Exceptions nicht, obwohl eine manuelle Analyse ergab, dass die Ausgaben den Anforderungen entsprachen.

Tab. 1: Vergleich der Fehlermeldungen: OGC *TEAM Engine* und GDI-DE *Testsuite*.

No	Error	<i>TEAM Engine</i>	<i>Testsuite</i>
01	Wrong format when exception is raised	+	-
02	GetFeatureInfo request to non-queryable services	+	-
03	GetMap request with TRANSPARENT=TRUE and „empty” bounding box	+	-
04	Use of exponential notation values for the bounding box	+	-
05	LegendGraphic is not exactly pixels	20 x 20	100 x 512
06	Invalid response when requesting a layer with a certain CRS	-	+
07	Invalid response when requesting a layer with the LAYERS parameter set to a specific layer	-	+
08	No service exception when requesting a layer with the CRS parameter set to an invalid CRS	-	+
09	MIME type of the response matches the format image/png when the FORMAT parameter is set to image/png	-	+

Für beide Testwerkzeuge (OGC *TEAM Engine* und die GDI-DE *Testsuite*) müssen die Ergebnisse relativiert werden, weil die Anzahl der Tests, die ein Dienst zu durchlaufen hat, stark variiert (55 bis 625 Tests pro Dienst; i.d.R. abhängig von der Anzahl der angebotenen Layer). Die Anzahl der fehlgeschlagenen Tests, also die Anzahl der Fehlermeldungen, dass etwas nicht konform zur Spezifikation ist, ist eine zentrale Aussage über einen Dienst, hilft in diesem Fall aber nicht die Dienste zu vergleichen. Denn diese Anzahl wird wesentlich durch die Anzahl der Layer eines Dienstes beeinflusst. In Verbindung mit kleineren Fehlern wie „die Legendengrafik misst nicht genau 20 × 20 Pixel“ (größere Legendengrafiken waren eine Anforderung im MDI-DE-Projekt) führt dies bei vielen Layern zu einer großen Anzahl von Fehlermeldungen. Somit ist die Aussagekraft der Ergebnisse recht klein und eine Prüfung des Einzelfalls

für jeden Dienst erforderlich. Interessanterweise haben die Ergebnisse der OGC *TEAM Engine* und der GDI-DE *Testsuite* nur einen Fehler gemeinsam, wie Tab. 1 zeigt. Dies bedeutet, dass beide Werkzeuge benötigt werden, um alle Fehler zu finden.

3.1.2 INSPIRE

Tab. 2 zeigt, dass der *Neogeo WMS INSPIRE Tester* die wenigsten Fehler der drei INSPIRE-Konformitätstest-Tools fand. Der *INSPIRE Geoportal Metadata Validator* findet fast so viele Fehler wie die GDI-DE *Testsuite*. Die GDI-DE *Testsuite* erkennt aber nicht nur mehr Fehler, sondern beschreibt die Fehler auch detaillierter. Da leider nicht alle Dienste mit der GDI-DE *Testsuite* getestet werden können, kann nicht auf den *INSPIRE Geoportal Metadata Validator* verzichtet werden.

Tab. 2: Vergleich der INSPIRE-Fehlermeldungen (N = *Neogeo WMS INSPIRE tester*, I = *INSPIRE Geoportal Metadata Validator*, G = *GDI-DE Testsuite*).

No	Error	N	I	G
01	Unexpected value for AccessConstraints	+	+	+
02	BoundingBox missing for some CRS	+	-	+
03	MetadataUrl not found in ExtendedCapabilities	+	+	+
04	Fees element mandatory but not found	+	+	-
05	DefaultLanguage, ResponseLanguage, ...	+	+	+
06	One or more layers failed the INSPIRE validation	-	+	+
07	Metadata element „Layers” is missing, empty or incomplete but it is required	-	+	+
08	Metadata element „Responsible Organisation” is missing, ...	-	+	-
09	Metadata element „Temporal Reference” is missing, ...	-	+	+
10	Metadata element „Resource Type” is missing, ...	-	+	+
11	Metadata element „Metadata Point Of Contact” missing, ...	-	+	+
12	Metadata element „Metadata Date” is missing, ...	-	+	+
13	Metadata element „Metadata Language” is missing, ...	-	+	+
14	Metadata element „Spatial Data Service Type” is missing, ...	-	+	+
15	Metadata element „Mandatory Keyword” is missing, ...	-	+	+
16	Metadata element „Resource Abstract” is missing, ...	-	+	+
17	Not every layer supports EPSG:4326 and EPSG:4258	-	-	+
18	Not every layer has an AuthorityURL and a correctly formatted identifier	-	-	+
19	Service title for identification not specified	-	-	+
20	A conformity statement with a result of conformance evaluation is not given	-	-	+

3.2 Performanz und Verfügbarkeit

Die INSPIRE Anforderungen wurden zusammen mit weiteren zu beachtenden Aspekten bereits in Abschnitt 2 erläutert. Darauf aufbauend wurden die MDI-DE Dienste zunächst mit dem *Service Status Checker* (SSC, 3.2.1) getestet. Hierbei war es nur möglich, die 23 WMS Dienste zu testen, da für WFS Dienste

nur Fehlermeldungen zurückgeliefert wurden. Jedoch gab es für drei der Dienste generell keine Ergebnisse und für weitere drei keine Ergebnisse für die Reaktionszeit. Anschließend folgt in Abschnitt 3.2.2 die Auswertung der Ergebnisse des *serviceMonitors*, der sowohl alle 23 WMS als auch alle vier WFS Dienste untersuchen konnte. Mit der GDI-DE *Testsuite* (3.2.3) können nur WMS Dienste ge-

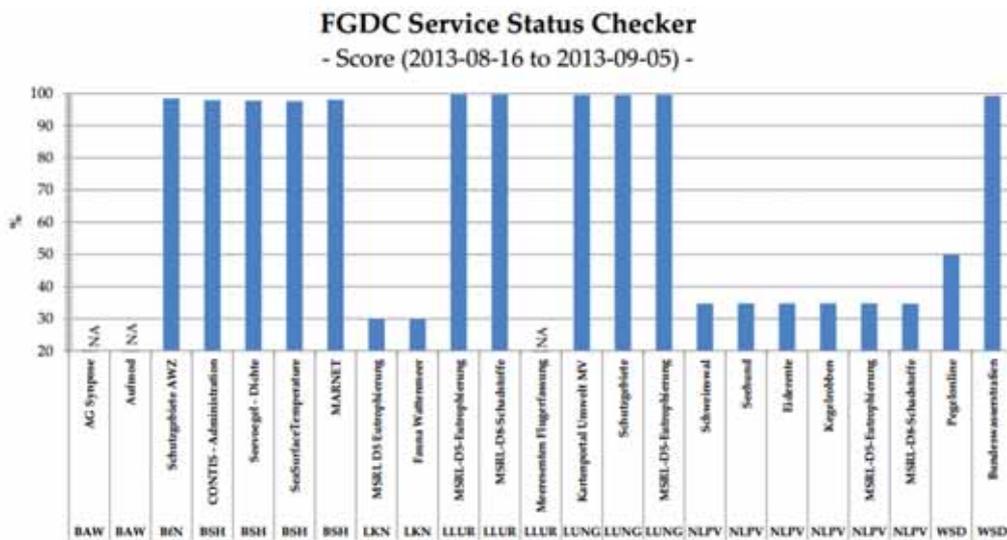


Abb. 2: Resultate für score (Punktzahl) (Service Status Checker).

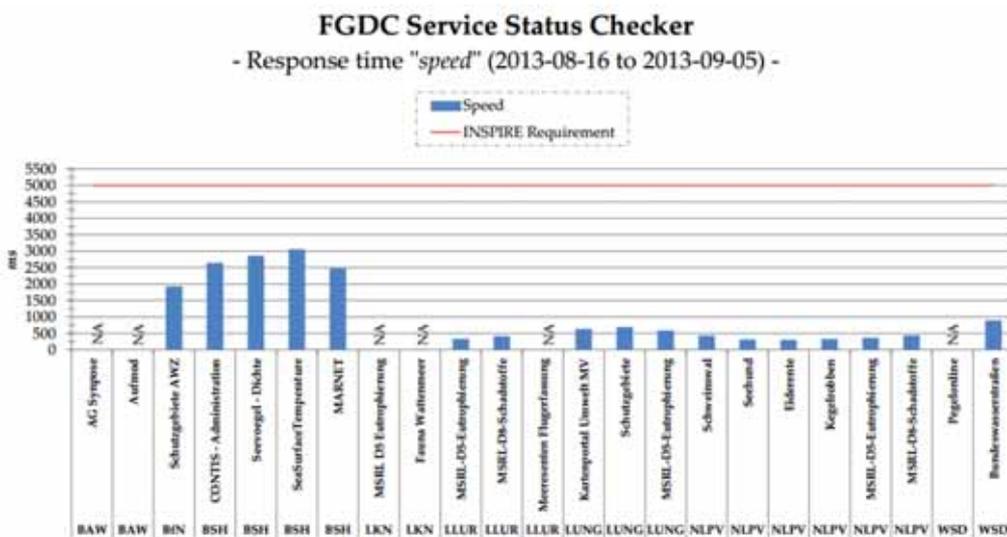


Abb. 3: Resultate für speed (Performanz, Reaktionszeit) (Service Status Checker).

testet werden, wobei einer der 24 Dienste aus nicht nachvollziehbaren Gründen nicht getestet werden konnte.

3.2.1 FGDC Service Status Checker (SSC)

Der SSC gibt eine Punktzahl (*score*) und die Performanz in Form der Reaktionszeit (*speed*)

aus. Die Methode der Berechnung der Punktzahl (*score*) ist unklar, weil eine Anfrage beim FGDC unbeantwortet blieb. Abb. 2 und 3 zeigen die Ergebnisse jeweils für *score* und *speed* im Zeitraum Mitte August bis Anfang September 2013. Die Zeitspanne ist so kurz, weil die API des FGDC SSC maximal 1.000 Datensätze zurückliefern kann. Abb. 2 zeigt, dass die Ergebnisse der Bewertungen in zwei Grup-

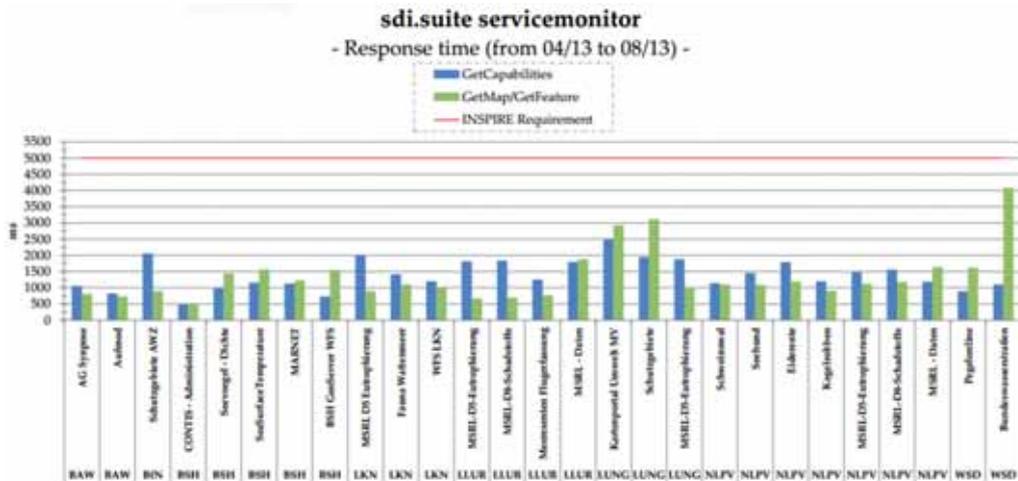


Abb. 4: Resultate für *speed* (Performanz, Reaktionszeit) (sdi.suite serviceMonitor).

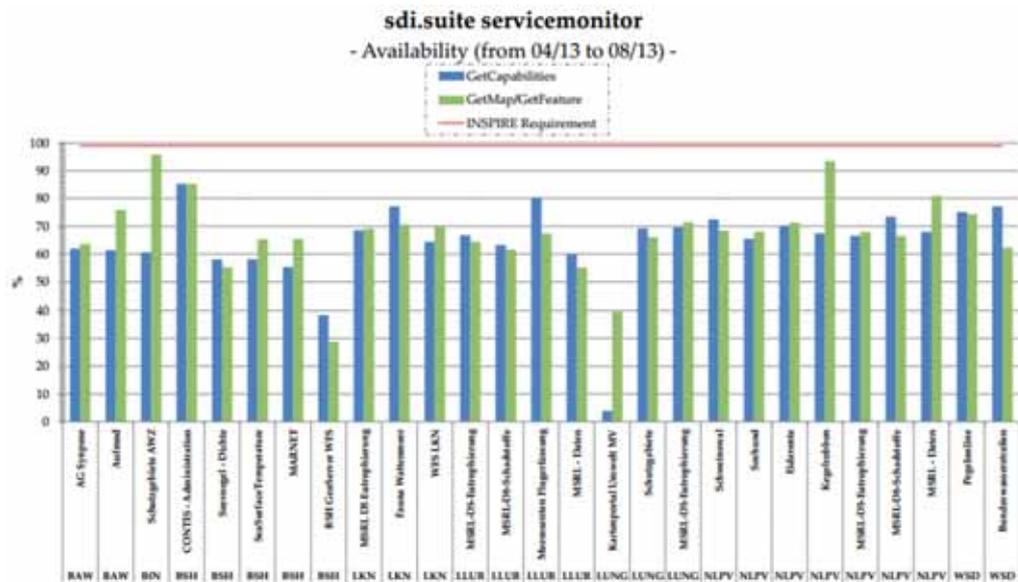


Abb. 5: Resultate für Verfügbarkeit (sdi.suite serviceMonitor).

pen geteilt sind. Die eine Gruppe (elf Dienste) erreicht Werte über 95%, während die andere Gruppe (neun Dienste, mit Ausnahme der „NA“ (nicht verfügbaren) Dienste) Werte unter 50% erreichte. Abb. 3 zeigt auch, dass alle Dienste die Anforderungen von INSPIRE erheblich unterboten. Leider hilft auch die Betrachtung von *speed* nicht, die Ergebnisse besser nachzuvollziehen, da Dienste mit hohen (= schlechten) Reaktionszeiten eine sehr hohe (= gute) Punktzahl bekamen, während andere Dienste mit sehr niedrigen (= guten) Reaktionszeiten sehr niedrige (= schlechte) Werte erreichten. Weiterhin ist nicht klar, warum einige Tests nicht für bestimmte Dienste durchgeführt werden konnten.

und GetMap-/GetFeature-Anfragen, aber es gibt keinen Hinweis darauf, dass ein Anfragetyp in der Regel langsamer oder schneller ist. Die Abbildung zeigt auch, dass die Reaktionszeiten die INSPIRE Anforderung von 5.000 ms unterschreiten. Wie Abb. 5 zeigt, erreicht jedoch keiner der Dienste die von INSPIRE geforderte Verfügbarkeit von 99%. Interessanterweise ist die Verfügbarkeit für die verschiedenen Anfragetypen nicht gleich. Sowohl die Ergebnisse der Reaktionszeiten sowie der Verfügbarkeit sind für die verschiedenen Anfragen ungleich verteilt. Dies bedeutet, dass die Antwortzeit bzw. Verfügbarkeit für eine GetCapabilities-Anfrage kürzer bzw. geringer als für eine GetMap-Anfrage ist.

3.2.2 sdi.suite serviceMonitor

Die Messungen mit dem sdi.suite *serviceMonitor* wurden über einen Zeitraum von etwa vier Monaten durchgeführt (April bis August 2013). Wie Abb. 4 zeigt, unterscheiden sich die Antwortzeiten für GetCapabilities-

3.2.3 GDI-DE Testsuite

Die Messungen mit der GDI-DE *Testsuite* wurden über einen Zeitraum von etwa vier Monaten (von April bis August 2013) vorgenommen. Im Gegensatz zum sdi.suite *serviceMonitor* weisen fast alle Dienste eine Verfüg-

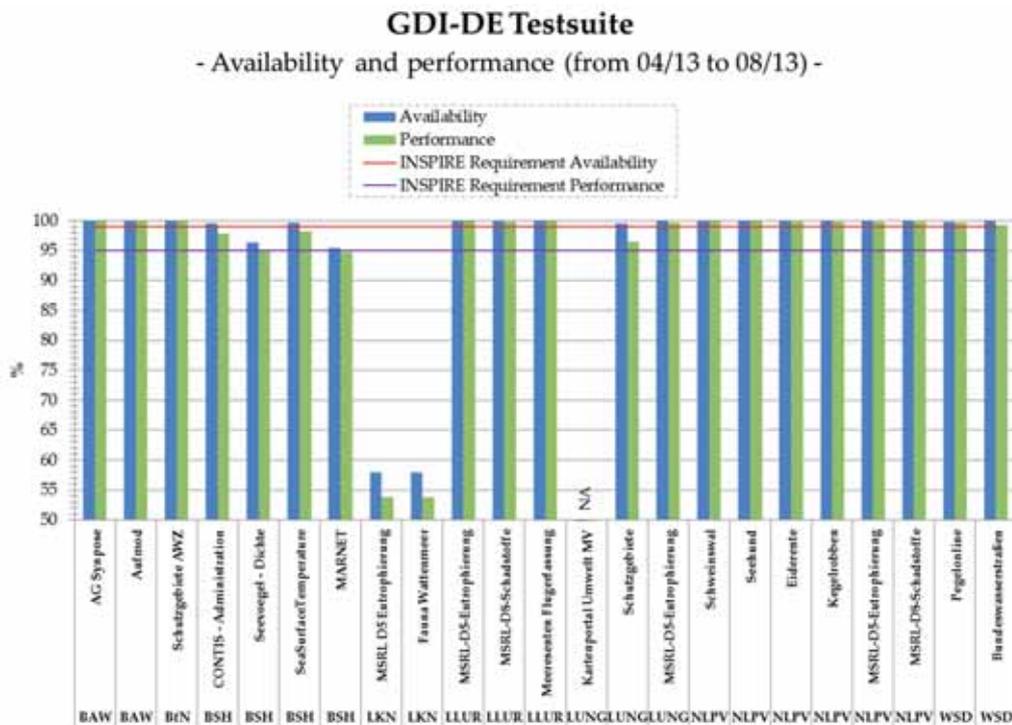


Abb. 6: Resultate für *speed* (Perfanz, Reaktionszeit) und Verfügbarkeit (GDI-DE *Testsuite*).

Ein ähnliches Bild ergibt sich hinsichtlich der Reaktionszeit. Obwohl der Korrelationskoeffizient hier im Vergleich zur Verfügbarkeit ($r = 0,064$) höher ist, bleibt er immer noch sehr gering. Dies wird durch die Dienste „Schutzgebiete“ und „MSRL-D5-Eutrophierung“ (LUNG), illustriert. Während beide eine ähnliche Reaktionszeit von 688 ms bzw. 590 ms beim SSC aufweisen, gibt der *serviceMonitor* einen Wert von 3.117 ms für „Schutzgebiete“ und 980 ms für „MSRL-D5-Eutrophierung“ aus. Abb. 8 zeigt, dass die Messungen des *serviceMonitor* und *SSC* zumindest darin übereinstimmen, dass kein Dienst die INSPIRE-Anforderung einer Reaktionszeit von 5.000 ms erfüllt.

Insbesondere für die Messungen der Verfügbarkeit wären die Ergebnisse besser vergleichbar gewesen, wenn genau die gleichen Zeiträume verwendet worden wären. Allerdings ist das Hauptproblem hier, dass die Zeiträume nicht gleich lang waren (etwa vier Monate gegenüber fast einem Monat) aufgrund der Beschränkung des *SSC* auf maximal 1.000 Datensätze.

Da auch die jeweiligen Messzeitpunkte nicht die gleichen für die verschiedenen Werk-

zeuge sind, sind die Ergebnisse auch diesbezüglich schwer vergleichbar. DRERUP (2010) kommt zu dem Schluss, dass diese Variationen einen Unterschied von bis zu 4,25% ausmachen können. Daher ist anzunehmen, dass ungleiche Messzeiten nicht allein für die ungleichen Ergebnisse verantwortlich sind.

Der von DRERUP (2010) beschriebene Einfluss unterschiedlicher Anfragetypen auf die Reaktionszeit konnte hier jedoch nicht nachgewiesen werden. Die Ergebnisse für GetCapabilities- und GetMap-Anfragen gemessen mit dem *serviceMonitor* weisen eine sehr geringe Korrelation ($r = 0,19$, siehe Abb. 4) auf. Da jedoch in den INSPIRE-Anforderungen GetMap- beziehungsweise GetFeature-Anfragen vorgeschrieben sind, ist die Bedeutung solcher Unterschiede zumindest im INSPIRE-Kontext nicht groß.

3.2.5 Aktuelle Entwicklung der Ergebnisse

Für die beiden Werkzeuge GDI-DE *Testsuite* und *SSC* wurden die Testreihen ein Jahr lang fortgeführt (August 2013 bis August 2014). Durch das Ende der Entwicklungsarbeiten am

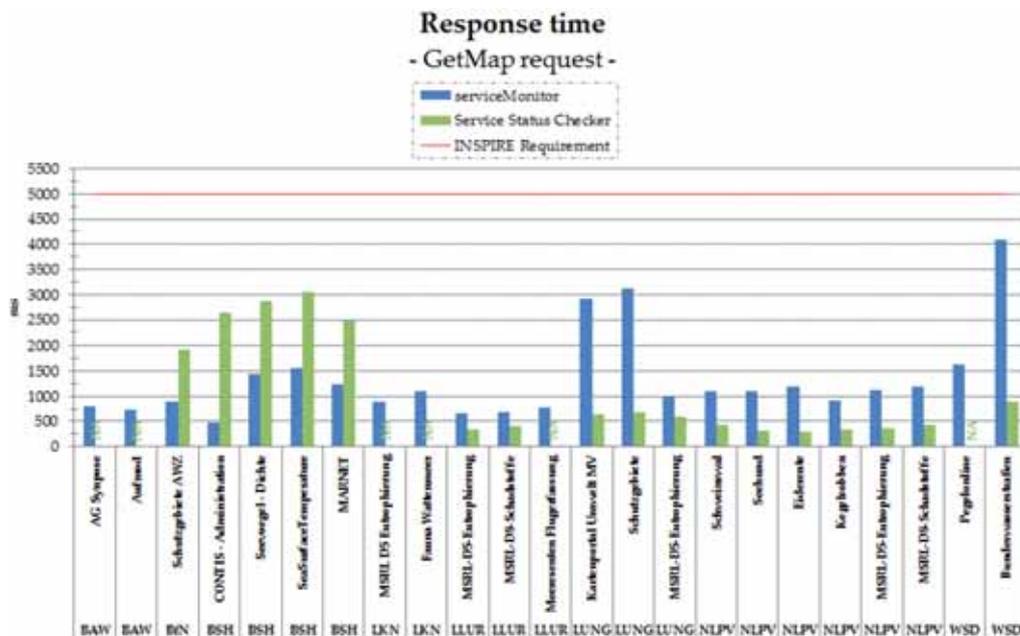


Abb. 8: Vergleich aller Monitoring-Ergebnisse bezüglich Antwortzeit.

MDI-DE-Projekt wurde der *sdi.suite service-Monitor* abgeschaltet, so dass es von diesem keine aktuellen Resultate gibt. Aus diesem Grund kann nur partiell überprüft werden, wie die Dienste zu einem späteren Zeitpunkt bzw. über einen anderen (längeren) Zeitraum zu bewerten sind.

Der Zeitraum, über den die Ergebnisse der GDI-DE *Testsuite* aggregiert wurden, hat sich vergrößert. Er gilt nun für das Jahr 2014 bis zum 26. August (also ungefähr 8 Monate). Abb. 9 zeigt die Resultate der Messungen und es fällt auf, dass zum einen nun Testergebnisse für alle Dienste verfügbar sind und dass die Reaktionszeit nun nur noch 90% statt vorher 95% der Zeit unter 5.000 ms liegen muss. Dies ändert das Resultat hinsichtlich des Scheiterns an dieser Grenze jedoch nicht, da alle Dienste entweder über 95% oder unter 90% liegen. Ansonsten lässt sich sagen, dass sich auf der einen Seite die Ergebnisse der Dienste des

LKN stark verbessert, auf der anderen Seite sich die Resultate der Dienste des BAW stark verschlechtert haben. Im Durchschnitt vermindert sich die Verfügbarkeit jedoch nur um ca. ein Sechstel Prozentpunkt während sich die Reaktionszeit um fast fünf Prozentpunkte erhöht. Doch trotz dieser marginalen Verminderung und der Erhöhung scheitern nun sieben (vorher vier) Dienste an den Vorgaben bezüglich Verfügbarkeit und wiederum sieben (vorher zwei) an den Vorgaben bezüglich der Reaktionszeit.

Der Zeitraum der mit *SSC* durchgeführten Testreihe ist wiederum kurz, da nach wie vor die Beschränkung von 1.000 Datensätzen besteht. Bei der in Abb. 10 dargestellten Testreihe bezüglich *score* wird der Zeitraum vom 16. bis zum 25. August 2014 abgebildet. Auf die Darstellung der Reaktionszeiten wurde verzichtet, da diese wieder die 5.000 ms Marke weit unterschreiten und die Veränderungen

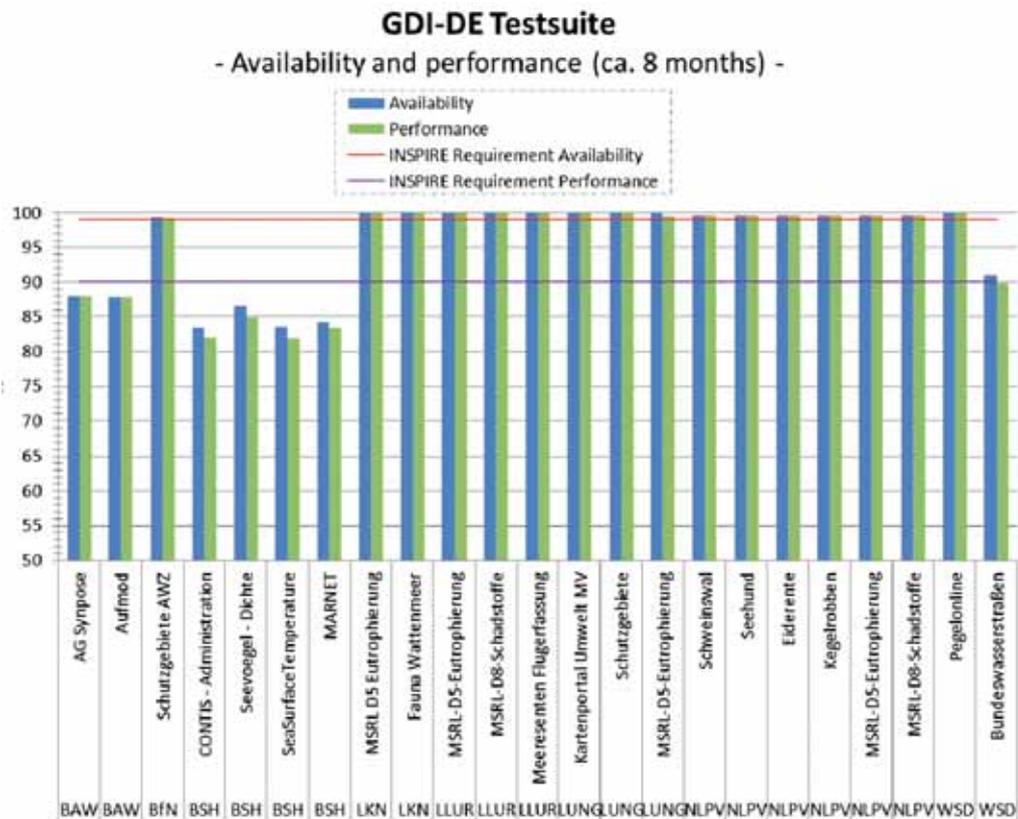


Abb. 9: Resultate für 2014 zur Reaktionszeit und Verfügbarkeit (GDI-DE *Testsuite*).

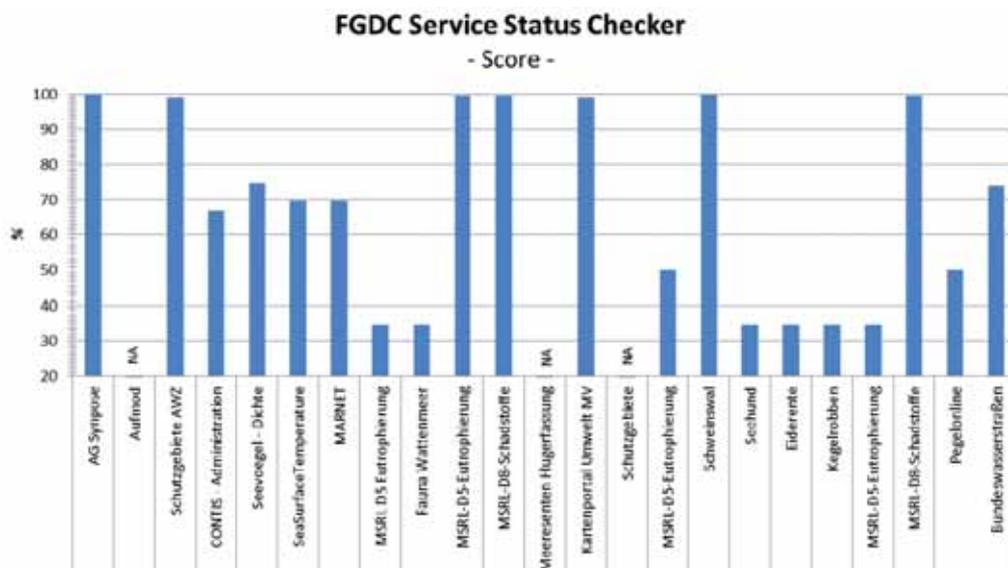


Abb. 10: Resultate von 2014 für *score* (FGDC Service Status Checker).

gegenüber 2013 recht gering sind. Allerdings sind für die Reaktionszeit deutlich weniger Ergebnisse verfügbar als bei der Auswertung der Testreihe aus dem Jahr 2013. So ist beispielsweise für keinen der Dienste des BSHs die Reaktionszeit verfügbar. Die BSH-Dienste sind auch diejenigen Dienste, bei denen die Verminderung des *score* am auffälligsten ist. Kurioserweise erhöhte sich bei zwei von sechs Diensten der NLPV der *score* um ca. 200% (bzw. 65 Prozentpunkte), obwohl die sechs Dienste alle auf dem gleichen Server laufen, was in den Resultaten der GDI-DE *Testsuite* auch abgebildet wird. Hier sind die Abweichungen der Resultate der NLPV-Dienste sehr gering.

3.3 Auswirkungen der Serverwahl auf die Dienstqualität

Es wäre interessant zu sehen, ob es einen Zusammenhang zwischen den Ergebnissen und

der verwendeten Server-Software gibt. Die verwendete Server-Software ist für sechs Hosts, die 15 Dienste umfassen, bekannt, von denen vier GeoServer, einer ArcGIS Server und einer MapServer einsetzen. Es liegt jedoch für jeden Server jeweils nur eine Stichprobe vor, so dass statistische Aussagen nicht möglich sind. Dennoch zeigt Tab. 3 die durchschnittlichen Verfügbarkeiten und Antwortzeiten pro Server-Software. Die Verfügbarkeit ist vor allem vom Host abhängig, kann jedoch auch durch die verwendete Server-Software verursacht sein. Wenn vereinfacht angenommen wird, dass die Beeinflussung der Reaktionszeit durch die Internetverbindung für alle Server gleich ist, dann hängt die Antwortzeit für eine GetMap-Anfrage im Wesentlichen von der Server-Software ab. Tab. 3 zeigt, dass die Ergebnisse für die Reaktionszeit für GeoServer und MapServer fast gleich sind. Diese Erkenntnis wird von MÜLLER & MANDERY (2009) und AIME & MCKENNA (2009) unterstützt, die angeben, dass beide eine ähnliche

Tab. 3: Durchschnittliche Antwortzeiten und Verfügbarkeiten der verschiedenen Server-Software.

	GeoServer	MapServer	ArcGIS Server
Verfügbarkeit (%)	77,3	85,2	89,1
Antwortzeit (ms)	775	785	1.412

Leistung liefern. Die Tabelle zeigt auch, dass die Reaktionszeit des ArcGIS Server Hosts deutlich höher ist. Ähnliches berichtet NASSER (2009), als er schrieb, dass sich viele Benutzer über die Leistung des ESRI ArcGIS Servers beschwert haben.

4 Zusammenfassung und Ausblick

Die deutsche marine Geodateninfrastruktur MDI-DE ist stark von Berichtspflichten für verschiedene europäische und nationale Gesetzgebungsinitiativen beeinflusst, die teilweise unter Zuhilfenahme von Diensten erfüllt werden können. Diese Dienste müssen spezifische durch die EU vorgegebene Anforderungen zu Konformität und Servicequalität erfüllen. Die Dienste der MDI-DE wurden mit verschiedenen bestehenden Tools analysiert, um sicherzustellen, dass die Dienste wie gefordert laufen.

Insbesondere Konformität kann recht gut bewertet werden. Jedoch kann aufgrund der zum Teil widersprüchlichen Ergebnisse und der Nichtprüfbarkeit einiger Dienste gesagt werden, dass stets mehrere Werkzeuge eingesetzt werden müssen, um schlüssige Ergebnisse zu erhalten. Auf der anderen Seite war die Servicequalität nicht kohärent, wenn sie mit verschiedenen Werkzeugen gemessen wurden. Die ermittelten Servicequalitäten zeigten sich als nicht vergleichbar. Die Ergebnisse sind nicht miteinander korreliert. Dies wirft die Frage auf, wie die Monitoring-Werkzeuge ihre Ergebnisse ermitteln, wie sie sie bewerten und analysieren und ob ihre Verwendung wirklich verantwortbar ist, zumal jeder Dienst in Zukunft spezifische Anforderungen von INSPIRE erfüllen muss. Außerdem besteht die Gefahr, dass die Werkzeuge gewählt werden, die die besten Ergebnisse liefern. Ansonsten lassen sich folgende Anforderungen an Testwerkzeuge besonders hinsichtlich INSPIRE-Konformität formulieren:

- Installierbarkeit auf dem Server,
- Spezifizierbarkeit der GetMap-Anfrage,
- Simulierbarkeit der Last,
- Spezifizierbarkeit der Messungszeitpunkte.

In Zukunft wird es darüber hinaus wichtig sein, einen Weg zu finden, solche GetMap-

Anfragen zu erzeugen, die Grafiken mit genau 470 KB zurückliefern. Nur dann werden die Ergebnisse vergleichbar und die Dienste wirklich die Anforderungen einhalten.

Außerdem bieten nicht alle Überwachungstools (Performanz und Verfügbarkeit) Visualisierungen oder eine strukturierte Sicht auf ihre Ergebnisse, so dass sie schwer zu interpretieren sind. Deshalb ist eine Webanwendung für die Ergebnisse des *Service Status Checker* vom FGDC vom Verfasser (RÜH 2014) auf Grundlage der JavaScript Object Notation (JSON) Ausgaben seiner Application Programming Interface (API) implementiert worden.

Literatur

- AIME, A. & MCKENNA, J., 2009: WMS Performance Shootout. – Presentation at FOSS4G 2009. – <http://download.osgeo.org/osgeo/foss4g/2009/SPREP/3Fri/Parkside%20Auditorium/1500/fri%20paud%201500%20mckenna-aime.pdf> (30.8.2014).
- ANTHONY, M. & NEBERT, D., 2012: Monitoring the Performance and Reliability of Geospatial Web Services – Service Status Checker (SSC) System Overview. – Global Geospatial Conference (GSDI13). – <http://www.gsdi.org/gsdiconf/gsd13/slides/1.4c.pdf> (30.8.2014).
- BARANSKI, B., FOERSTER, T., SCHÄFFER, B. & LANGE, K., 2011: Matching INSPIRE Quality of Service Requirements with Hybrid Clouds. – *Transactions in GIS* **15**: 125–142.
- BOZKURT, M., HARMAN, M. & HASSOUN, Y., 2010: Testing Web Services: A Survey. Technical report, King's College London. – <http://www.dcs.kcl.ac.uk/technical-reports/papers/TR-10-01.pdf> (30.8.2014).
- CIBULKA, D., 2013: Performance Testing of Web Map Services in three Dimensions – X, Y, Scale. – *Slovak Journal of Civil Engineering*, volume **XXI** (1): 31–36. – <http://www.degruyter.com/view/j/sjce.2013.21.issue-1/sjce-2013-0005/sjce-2013-0005.xml> (30.8.2014).
- DRERUP, K., 2010: Setup and test of a WFS for NatureSDI+ according to INSPIRE. – Bachelor Thesis. – <http://hig.diva-portal.org/smash/record.jsf?pid=diva2:354500> (30.8.2014).
- EUROPEAN COMMISSION, 2013: INSPIRE Geoportal Metadata Validator. – <http://inspire-geoportal.ec.europa.eu/validator2/> (30.8.2014).
- FRANKEN, L., 1996: Quality of service management – a model-based approach. – PhD Thesis, KPN Research.

- FRÖLUND, S. & KOISTINEN, J., 1998: Quality-of-Service Specification in Distributed Object Systems. – *Distributed System Engineering Journal*, December, **5** (4). – <http://www.hpl.hp.com/techreports/98/HPL-98-159.pdf> (18.2.2015).
- GIULIANI, G., DUBOIS, A. & LACROIX, P., 2013: Testing OGC Web Feature and Coverage Service performance: Towards efficient delivery of geospatial data. – *Journal of Spatial Information Science* **7**: 1–23. – <http://www.josis.org/index.php/josis/article/view/112> (30.8.2014).
- HECKEL, R. & MARIANI, L., 2005: Automatic Conformance Testing of Web Services. – *Fundamental Approaches to Software Engineering*. – <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.95.6051&rep=rep1&type=pdf> (30.8.2014).
- HOGREBE, D., 2012: GDI-DE Testsuite. Improving interoperability. – *INSPIRE conference 2012*. – http://inspire.jrc.ec.europa.eu/events/conferences/inspire_2012/presentations/90.pdf (30.8.2014).
- HORÁK, J., ARDIELLI, J. & HORÁKOVÁ, B., 2009: Testing of Web Map Services. – *International Journal of Spatial Data Infrastructures Research, Special Issue GSDI-11*. – <http://www.gsdi.org/gsdiconf/gsdil1/papers/pdf/330.pdf> (30.8.2014).
- INSPIRE, 2007: Directive 2007/2/EC of the European Parliament and of the Council. – <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:0014:en:PDF> (30.8.2014).
- INSPIRE CONSOLIDATION TEAM, 2007: INSPIRE Network Services Performance Guidelines. – *Technical report*. – http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/network/Network_Services_Performance_Guidelines_%20v1.0.pdf (30.8.2014).
- INSPIRE NETWORK SERVICES DRAFTING TEAM, 2009: Draft Implementing Rules for Download Services (Version 3.0). – *Technical report*. – http://inspire.ec.europa.eu/documents/Network_Services/Technical_Guidance_Download_Services_3.0.pdf (18.6.2015).
- INSPIRE NETWORK SERVICES DRAFTING TEAM, 2013: Technical Guidance for the implementation of INSPIRE View Services (Version 3.11). – *Technical report*. – http://inspire.ec.europa.eu/documents/Network_Services/TechnicalGuidance_ViewServices_v3.11.pdf (18.6.2015).
- JAKOBSSON, A., 2012: European Location Framework (White Paper v1.0). – *Technical report*. – <http://www.elfproject.eu/sites/default/files/ELF%20White%20Paper.pdf> (18.2.2015).
- LUPP, M., 2008: OGC's Open Standards for Geospatial Interoperability. – *Encyclopedia of GIS*, Springer US: p. 800.
- MAPMATTERS, 2015: <http://www.geops.de> (23.5.2015).
- MÜLLER, U. & MANDERY, N., 2009: Leistungsfähigkeit aktueller WMS-Dienste. – *Intergeo 2009*. – http://www.fossgis.de/w/images/9/9d/Geops_wmsperformance_090922.pdf (30.8.2014).
- MÜLLER, U. & MANDERY, N., 2010: Qualitätsaspekte aktueller Web Map Services (WMS) als Grundlage einer dienstbasierten Geodateninfrastruktur. – *STROBL, J., BLASCHKE, T. & GRIESEBNER, G. (Hrsg.): Angewandte Geoinformatik 2010: Beiträge zum 22. AGIT-Symposium Salzburg, XVIII: 813–816*, Wichmann, Berlin. – <https://www.vde-verlag.de/proceedings-en/537495104.html> (30.8.2014).
- NASSER, H., 2009: ArcGIS Server Performance. – *Blog post*. – <http://www.husseinasser.com/2009/02/esri-arcgis-server-performance.html> (30.8.2014).
- NEOGEO, 2015: http://inspire_tester.neogeo-online.net (23.5.2015).
- PARASURAMAN, A., ZEITHHAML, V.A. & BERRY, L.L., 1985: A Conceptual Model of Service Quality and Its Implications for Future Research. – *Journal of Marketing* **49** (4): 41–50, ISSN 00222429. – <http://www.jstor.org/stable/1251430> (30.8.2014).
- RÜH, C., 2014: Marine spatial data infrastructures – Approaches on evaluation, design and implementation. – *DGK Reihe C Heft 732*, Verlag der Bayerischen Akademie der Wissenschaften, München 2014.
- TEAMENGINE, 2015: <http://cite.opengeospatial.org/teamengine/>, <http://sourceforge.net/projects/teamengine/> (23.5.2015).
- TESTSUITE, 2015: <http://testsuite.gdi-de.org/> (23.5.2015).
- YANG, C.P. & EVANS, J., 2008: Network GIS Performance. – *SHEKHAR, S. & XIONG, H. (Hrsg.): Encyclopedia of GIS: 787–790*, Springer. – <http://dblp.uni-trier.de/db/reference/gis/gis2008.html#YangE08> (30.8.2014).

Adresse des Autors:

Christian Seip, Universität Rostock, Agrar- und Umweltwissenschaftliche Fakultät (AUF), Professur für Geodäsie und Geoinformatik (GG), Justus-von-Liebig-Weg 6, D-18059 Rostock, Tel.: +49-381-498-3212, Fax: +49-381-498-3002, e-mail: christian.seip@uni-rostock.de

Manuskript eingereicht: September 2014
Angenommen: Mai 2015