

Übung 6. – Leistungsmodellierung, Visuelle Datenanalyse

1 Leistung eines Servers

Wir haben an einem Server die folgenden Leistungsparameter gemessen:

Zeitpunkt der Probe [ms]	500	600	700	800	900
Anzahl der in den letzten 100ms bedienten Anfragen	11	12	21	18	20
Durchschnittsbedienzeit der letzten 100ms [ms]	15	20	21	25	27
Prozessorauslastung der letzten 100ms [%]	12	13	16	17	19
Festplatten-I/O-Auslastung der letzten 100ms [%]	55	63	87	61	73

- Anhand der zur Verfügung stehenden Daten, welche Ressource des Servers kann der Engpass im System sein? Warum?
- Wie hoch ist der Durchsatz zu der Zeit der ersten Probe? Anhand der Ergebnisse der 5 Abtastungen, wie hoch ist der gemessene Durchschnitt und Median des Durchsatzes?
- Anhand dieser 5 Messungen, wie würden Sie die durchschnittliche Anzahl der zu gleicher Zeit in Bearbeitung stehender Anfragen abschätzen?

2 Gemeinschaftsseite

Wir betreiben eine Gemeinschaftsseite am Internet. In der letzten Zeit ist die Seite viel populärer geworden, deswegen ist aber die Antwortzeit auch gewachsen. Unser Geschäftsziel ist es, dass die Homepage in der Spitzenzeit auf einmal 1500 Benutzer mit 4 Sekunden durchschnittlicher Antwortzeit bedienen kann.

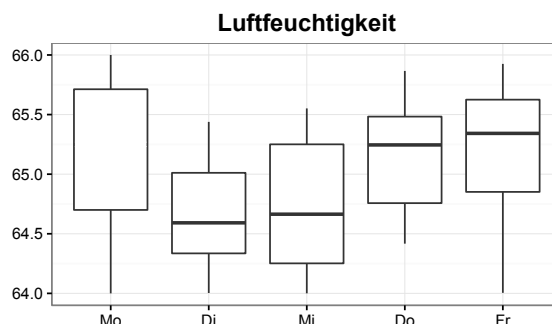
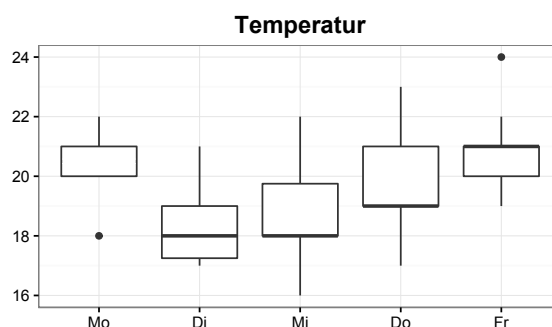
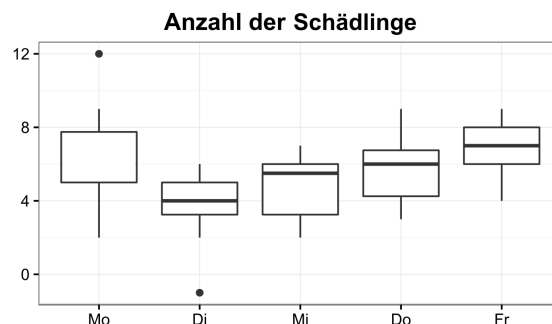
- Mindestens wie groß muss man den Durchsatz der Bedienungsinfrastruktur planen, falls die Zeit aller anderen Verzögerungen (Netzverkehr, HTML-Erscheinung beim Klienten) als eine Sekunde betrachtet werden kann?
- Entsprechend den Messungen an der wiederentworfenen Webseite braucht die Bedienung einer Anfrage durchschnittlich 20 ms CPU-Zeit am Webserver, und es besetzt einen Datenbankserver für 12,5 ms. Zurzeit bedienen 15 Webserver die Anfragen, und die Datenbank ist für 5 Bediener repliziert. Welche Art und mindestens wie viele Computer müssen wir noch kaufen, um das obige Ziel zu erfüllen, falls wir eine lineare Skalierbarkeit annehmen?
- (*) Wie groß wird die Auslastung im erweiterten System? Wie fern müsste das System erweitert werden, falls wir auch in der Spitzenzeit nur höchstens 50% Auslastung erlauben wollen?
- Betrachten wir nur die 2 Webserver und die 3 Datenbanken. Konstruieren wir zustandsbasierte Modelle, die die Ressourcen der Infrastruktur nach Verfügbarkeit (frei/besetzt) modellieren! Welche Designerentscheidungen treffen wir? Was sind die Vor- und Nachteile der verschiedenen Möglichkeiten?

3 Netzwerk von Sensoren (frühere Klausuraufgabe) – Datenanalyse

Sei ein Netzwerk von Landwirtschaftssensoren gegeben, mit dem Freilandstücke, Glashäuser und Folien überwacht werden. Die Sensoren messen Daten wie Temperatur, Luftfeuchtigkeit, Helligkeit, Windstärke, Anzahl der Schädlinge, usw.

Datum	Temp. [°C]	Luftf. [%]	Schädl. [Stück]
4. Mai 2015 08:00	18	66,00	3
4. Mai 2015 09:00	20	65,75	6
4. Mai 2015 10:00	20	65,75	8
4. Mai 2015 11:00	20	65,50	9
4. Mai 2015 12:00	20	65,50	5
4. Mai 2015 13:00	21	65,00	12
4. Mai 2015 14:00	21	64,70	5
4. Mai 2015 15:00	21	64,70	6
4. Mai 2015 16:00	21	64,60	7
4. Mai 2015 17:00	22	64,00	2

- Leider sind die Mittelwerte (Median) von Montag, dem 4. Mai von den Bildern verschwunden, zeichnen Sie sie anhand der Daten in der Tabelle ein.
- Interpretieren Sie die Diagramme: Von welchen Variablen weisen die 1. Quartile eine streng monotone Änderung mit der Zeit auf?
- (Ergänzende Aufgabe.) Die Temperaturwerte und die Anzahldaten der Schädlinge von Montag, dem 4. Mai sollen an einem Diagramm mit parallelen Koordinaten verglichen werden.



4 Netzwerk von Sensoren (frühere Klausuraufgabe)

– Leistungsanalyse

(Zu Aufgabe 3. gehörende Aufgaben für Analyse von Leistungen.)

Die verschiedenen Sensoren liefern entsprechende Daten von ihrer 100 m-Umgebung, die sie samt Zeitstempel über dem drahtlosen Netzwerk zu der Zentrale schicken. Der Prozessor des zentralen Rechners verarbeitet diese Nachrichten, dann schreibt er sie auf eine Platte zur Langzeit-speicherung. Unser Betrieb hat insgesamt 4500 Sensoren installiert, die je eine Nachricht pro Minute schicken. Unsere Infrastruktur bedient die aktuelle Belastung erfolgreich. Das drahtlose Netzwerk kann 100 Messdatenpakete pro Sekunde übertragen. Der Prozessor des zentralen Rechners läuft zu 75% leer. Die Platte braucht 8 ms für das Speichern einer Nachricht.

- Wie viele Nachrichten pro Sekunde ist der aktuelle Durchsatz der Infrastruktur?
- Wie hoch sind der aktuelle Durchsatz, der Grenzdurchsatz und die Auslastung des Netzwerks, des Prozessors, bzw. der Platte?
- Um die Messgenauigkeit zu erhöhen, wollen wir weitere Sensoren installieren. Wie viele dürfen wir ohne Verstärkung der Infrastruktur? Nehmen Sie eine lineare Skalierbarkeit an.
- Mit geschickter Kodierung unterstützt das Netzwerk, dass mehrere Sensoren gleichzeitig Messdaten schicken können. Im Durchschnitt wie viele Sensoren senden gleichzeitig im Netzwerk (wie hoch ist die Überlappung im Netzwerk) aktuell bzw. bei der maximalen Auslastung des Netzwerks, wenn das Abschicken einer Nachricht 40 ms dauert?