



Messbericht

PHPProjekt Groupware

Inhaltsverzeichnis

1	Überblick	6
1.1	Einführung	6
1.2	Aufbau des Dokuments	6
2	Management Summary	7
2.1	Ziel der Messung	7
2.2	Kurze Beschreibung des Testablaufs	7
2.3	Interpretation der Messergebnisse	7
3	Beschreibung der Testumgebung	9
3.1	Testsystem	9
3.2	Testobjekt	10
3.2.1	Hardware	10
3.2.2	Software	11
4	Definition der Messungen	12
4.1	Beschreibung der Auftragsarten	12
4.2	Beschreibung der Kettenarten	12
4.3	Beschreibung der Benutzertypen (EAG)	13
4.4	Zeitvorgaben	13
4.4.1	Vorbereitungszeiten (AGVZ)	13
4.4.2	Durchlaufzeitforderungen (ADLZ)	14
4.5	Durchgeführte Messungen	14
4.6	Erläuterung zu den Messergebnissen	14
4.6.1	Ergebnisse der DIN-Messungen	14
4.6.2	Ergebnisse des System-Monitoring	15
5	Messreihe 1 Durchführung und Auswertung	16
5.1	Übersicht der Messläufe	16
5.2	Ergebnisse der Messungen	16
5.3	Auftragsart 'kontakte'	18
5.3.1	DIN/ISO Bewertungsgrößen	18
5.3.2	Antwortzeithistogramm 'kontakte'	18
5.3.3	Summenhistogramm 'kontakte'	19
5.4	Auftragsart 'notiz'	19
5.4.1	DIN/ISO Bewertungsgrößen	19
5.4.2	Antwortzeithistogramm 'notiz'	20
5.4.3	Summenhistogramm 'notiz'	20

5.5	Auftragsart 'forum'	21
5.5.1	DIN/ISO Bewertungsgrößen	21
5.5.2	Antwortzeithistogramm 'forum'	21
5.5.3	Summenhistogramm 'forum'	22
5.6	Auftragsart 'todo'	22
5.6.1	DIN/ISO Bewertungsgrößen	22
5.6.2	Antwortzeithistogramm 'todo'	23
5.6.3	Summenhistogramm 'todo'	23
5.7	System-Monitoring	24
5.7.1	Monitoring 'alpspitze'	24
5.7.2	Monitoring 'erongo'	24
5.7.3	Monitoring 'erongo vmw'	25
5.7.4	Monitoring 'tofana'	25
A	Anhang	26
A.1	Die DIN 66273 / ISO 14756 Methode	26
A.2	Durchführung einer Leistungsmessung nach DIN	26
A.2.1	Erstellung eines Anforderungsprofils	26
A.2.2	Durchführung einer Messung	27
A.2.3	Auswertung einer Messung	28
B	Glossar	30

Tabellenverzeichnis

/home/saturn/work/DEMOMESSB/result/s_reportde/env_s_aturn.csv	10
/home/saturn/work/DEMOMESSB/result/s_reportde/env_hardware.csv	11
/home/saturn/work/DEMOMESSB/result/s_reportde/env_software.csv	11
/home/saturn/work/DEMOMESSB/result/Series/iso_task.csv	12
/home/saturn/work/DEMOMESSB/result/Series/iso_chains.csv	13
/home/saturn/work/DEMOMESSB/result/Series/iso_utype.csv	13
/home/saturn/work/DEMOMESSB/result/Series/iso_ptime.csv	13
/home/saturn/work/DEMOMESSB/result/Series/iso_tclass.csv	14
Ausgewertete Messungen	16
Ergebnistabelle	17

Abbildungsverzeichnis

1	Beispiel Summenhistogramm 'todo'	8
2	Beispielmessung 'erongo vmw'	8
3	DIN/ISO Bewertungsgrößen 'kontakte'	18
4	Antwortzeithistogramm 'kontakte'	18
5	Summenhistogramm 'kontakte'	19
6	DIN/ISO Bewertungsgrößen 'notiz'	19
7	Antwortzeithistogramm 'notiz'	20
8	Summenhistogramm 'notiz'	20
9	DIN/ISO Bewertungsgrößen 'forum'	21
10	Antwortzeithistogramm 'forum'	21
11	Summenhistogramm 'forum'	22
12	DIN/ISO Bewertungsgrößen 'todo'	22
13	Antwortzeithistogramm 'todo'	23
14	Summenhistogramm 'todo'	23
15	Monitoring 'alpspitze'	24
16	Monitoring 'erongo'	24
17	Monitoring 'erongo vmw'	25
18	Monitoring 'tofana'	25
19	DIN-Bewertungsgrößen	29

1 Überblick

1.1 Einführung

Dieser Bericht dokumentiert die Ergebnisse der Lasttests an der beispielhaften Anwendung **Groupware PHProjekt**. Er zeigt die Leistungsfähigkeit des **s_atum** Systems sowie die Möglichkeit, aus den umfangreichen Daten einer Meßreihe weitgehend automatisiert einen Ergebnisbericht abzuleiten. Er beschreibt die Systemarchitektur und die beteiligten Komponenten zum Zeitpunkt der Durchführung der Messungen.

Die Hard- und Software des zu prüfenden Systems inklusive Betriebssystem und Netzwerkinfrastruktur wird dabei als **Testobjekt** bezeichnet. Der Begriff **Testsystem** umfasst analog Hard- und Software des eingesetzten Lasttesttreibers. Die im Text enthaltenen Abbildungen und Tabellen dokumentieren die Ergebnisse der Messungen. Ihre Datenbasis wurde während der Tests durch den Lasttesttreiber und die Monitoringwerkzeuge ermittelt. Während einer Testdurchführung fallen umfangreiche Datenmengen an, die automatisch verarbeitet und aufbereitet werden.

1.2 Aufbau des Dokuments

Den Vorbemerkungen folgt ein Management Summary und in Kapitel 3 eine Beschreibung des Testaufbaus. Kapitel 4 enthält eine Übersicht über den strukturellen Aufbau der Last durch Auftragsarten, Auftragsketten, Benutzertypen, Mengenangaben über Benutzer und Zeitvorgaben. Außerdem wird der Fluß dynamischer Daten den Auftragsarten zugeordnet und eine Übersicht über die durchgeführten Messungen aufgelistet. In Kapitel 2 findet sich ein Management Summary, das eine Zusammenfassung des Messberichts enthält. In Kapitel 4.6 findet eine allgemeine Erläuterung zu den Messergebnissen und DIN-Messungen statt. Kapitel 5 enthält die Auswertung und die Ergebnisse der Messung. Sie setzt sich aus einer, oder mehrerer Messreihen zusammen, die wiederum mehrere Messläufe enthalten kann. Anhang A liefert eine Erläuterung zur Messung nach DIN 66273 / ISO 14756 und Anhang B enthält ein Glossar über die wichtigsten Begriffe.

2 Management Summary

2.1 Ziel der Messung

Der Vergleich des Einsatzes der Beispiel-Anwendung **Groupware PHProjekt** auf vier verschiedenen Systemen soll die Leistungsfähigkeit der Systeme im ausgewählten Einsatzfall darstellen. Hierzu werden die Systeme durch die Simulation der Interaktion einer Vielzahl von Nutzern belastet. Randbedingung ist dabei die Einhaltung der Grenzwerte der vorgegebenen DIN-Bewertungsgrößen an Hand von vier festgelegten Anwender-Aktionen (Auftragsarten). Einzelne Vorgaben (Denk- und Durchlaufzeit) wurden hier mehr auf die demonstrative Aussagefähigkeit der Messergebnisse als auf eine realistische Nutzung der Anwendung ausgerichtet.

2.2 Kurze Beschreibung des Testablaufs

Die **Groupware PHProjekt** wurde als Web-Applikation auf vier Systemen mit abgestufter Leistungsfähigkeit eingerichtet, von denen eines in einer virtuellen Umgebung (VMware) abläuft. Als Webserver wird Apache, als Datenbank PostgreSQL eingesetzt. Die Simulation der Anwenderinteraktionen (Aufträge) erfolgt durch ein eigenes System mit dem Lasttesttreiber **s_aturn**, das die zu testenden Systeme über ein lokales Netz beschickt. Als Lastprofil wurde eine repräsentativer Mix aus vier ausgewählten Auftragsarten mit Vorgaben (z.B. Denkzeit, Systemreaktionszeit) entsprechend der Norm DIN 66273 / ISO 14756 definiert.

Zur Ermittlung der Leistungsfähigkeit wird bei der Durchführung der Messung die Zahl der simulierten Nutzer als ergebnisrelevante Größe variiert, bis die Normvorgaben im Rahmen der gewünschten Präzision erfüllt werden.

2.3 Interpretation der Messergebnisse

Die erreichte Nutzerzahl ist ein Maß für die Leistungsfähigkeit der getesteten Systeme. Mit den gewählten Vorgaben konnten die Interaktionen von 450 ('alpspitze'), 300 ('erongo'), 150 ('erongo vmw') und nochmals 150 ('tofana') Nutzern verarbeitet werden. 'alpspitze', 'erongo' und 'tofana' bezeichnen die unterschiedlichen Testobjekte.

Als Durchsatz wurden Werte zwischen ca. 2000 und 4200 Aufträgen pro Stunde gemessen. Hieraus abgeleitet ergeben sich für die Systeme Leistungsfaktoren von ca. 2 ('alpspitze'), 1.3 ('erongo'), 1 ('erongo vmw') und 1 ('tofana') bezogen auf das leistungsschwächste System ('erongo vmw'). Durch die Virtualisierung mit VMware wird hier ca. 75% der Leistung des realen Systems 'erongo' erreicht.

Die Auftragsarten unterscheiden sich geringfügig im Leistungsanspruch. Die Auftragsart 'kontakte' benötigt die geringsten Ressourcen, 'todo' die höchsten. Wegen der gleichwertigen Auftragsmischung (je 25% jeder Art) waren Schwerpunkte (CPU-Bedarf, Datenbanknutzung) nicht auszumachen.

Die Grenzwerte im Antwortzeitverhalten macht die Auftragsart 'todo' deutlich. Die treppenförmige Linie in Abbildung 1 stellt das Zeitklassenschema dar.

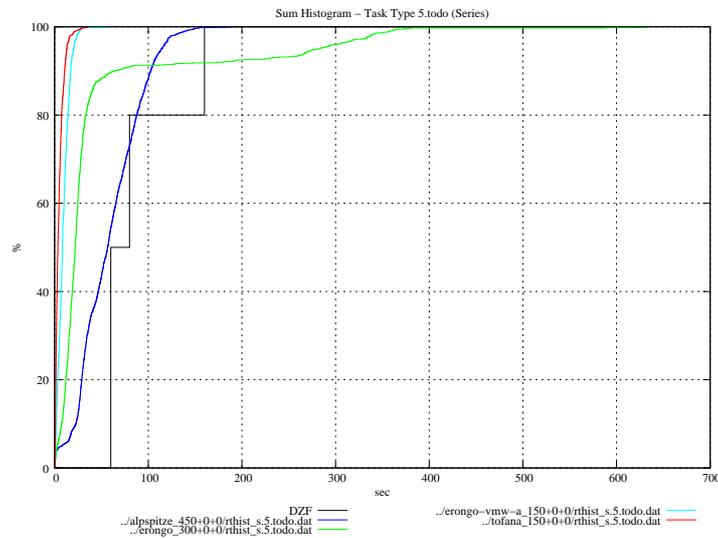


Abbildung 1: Beispiel Summenhistogramm 'todo'

Aus dem Systemmonitoring lässt sich hinsichtlich der Auslastung der Ressourcen in zwei Messläufen eine Korrelation der Werte für Webserver und Datenbank erkennen. Die CPU wird in allen Fällen während der Kernmesszeit zu 100% ausgelastet.

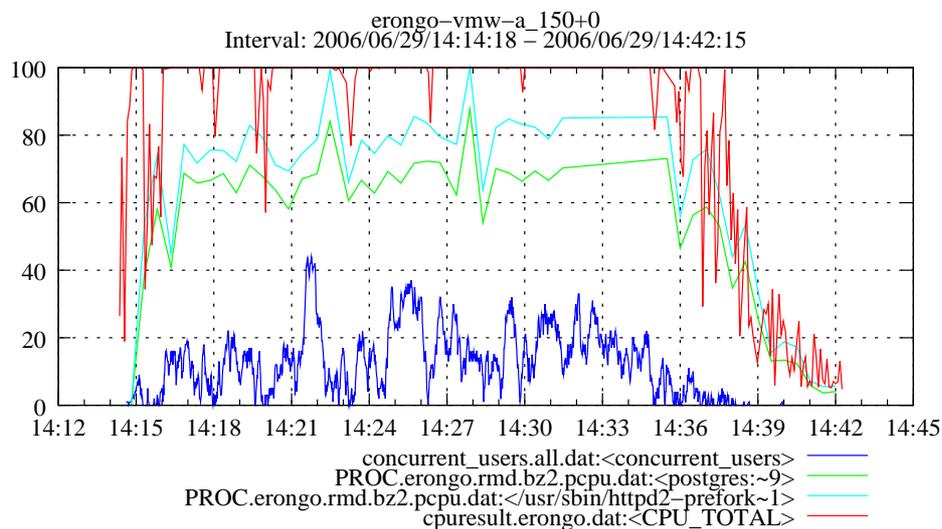
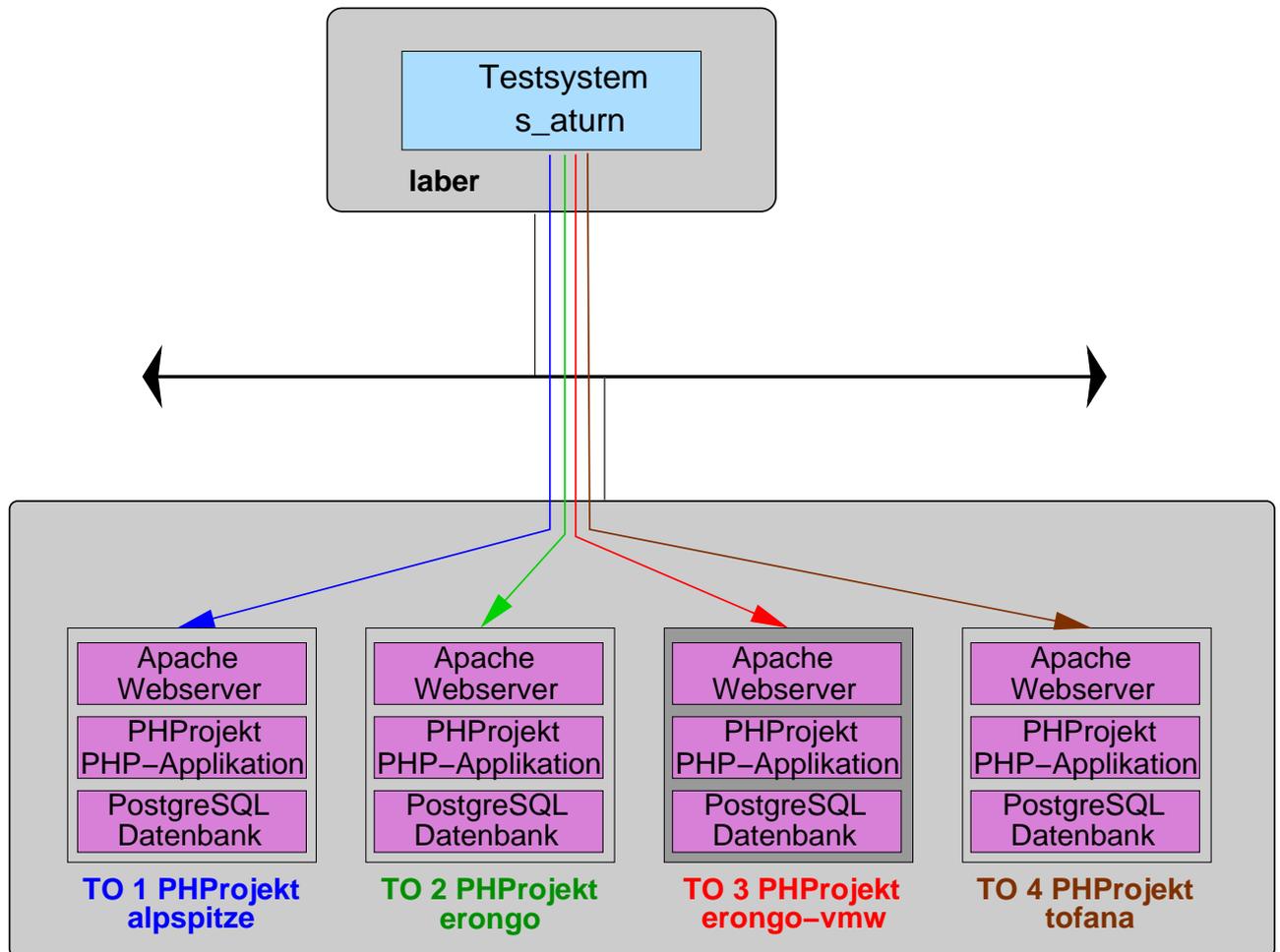


Abbildung 2: Beispielmessung 'erongo vmw'

Fazit: Mit den gewählten Nutzerzahlen befindet sich ein System an der Leistungsgrenze ('erongo'), ein System verfügt noch über Leistungsreserven ('tofana'), und die weiteren zwei Systeme erfüllen die Vorgaben ('alpspitze', 'erongo vmw').

3 Beschreibung der Testumgebung

Die Architektur der für den Test eingesetzten System- und Anwendungsumgebung wird durch folgendes Bild beschrieben:



3.1 Testsystem

Als Lasttreiber wird das System 'laber' eingesetzt. Die im folgenden beschriebene Konfiguration genügt hier den Anforderungen als **s_aturn** Testtreiber vollkommen. Im Bedarfsfall kann ein Monitoring der charakteristischen Auslastungsdaten des Lasttesttreibersystems vorgenommen werden.

Komponente	System 'Laber'
Prozessor CPU	AMD Athlon 1200 +
Hauptspeicher	1 GB
Festplattenkapazität	1*IDE 20 GB
Betriebssystem	Linux 10.0
Kernel	2.6.8-24-default
Virtualisierung	-
s_aturn-Version	s_aturn 2.14-1
<i>/home/saturn/work/DEMOMESSB/result/s_reportde/env_s_aturn.csv</i>	

3.2 Testobjekt

Das Testobjekt ist die **Groupware PHProjekt**, die sich als Beispiel für eine Web-Applikation anbietet. Sie eröffnet einen umfangreichen Funktionsumfang hinsichtlich Kontakt-, Termin- und Aufgabenverwaltung. Für den Test wurden die typischen Aktionen:

- Erfassung eines neuen Kontakts
- Erfassung einer neuen Notiz
- Neuer Beitrag für ein Forum
- Erfassung einer neuen Aufgabe (todo)

ausgewählt.

Der Zugang zur **Groupware PHProjekt** erfolgt über den Webserver Apache; das eingesetzte Database Management System (DBMS) ist PostgreSQL. Diese Konstellation wurde auf vier Systemkonfigurationen unterschiedlicher Leistungsfähigkeit implementiert, wovon für eine die virtuelle Systemplattform 'VMware' ausgewählt wurde. Die Bezeichnung der Systeme entspricht den Rechnernamen im Netz: 'alpspitze', 'erongo', 'erongo vmw' und 'tofana'.

3.2.1 Hardware

Die für die Tests eingesetzte Hardware wurde weitgehend entsprechend den Anforderungen für Groupware-Anwendungen in kleineren bis mittelständischen Unternehmen ausgewählt. Sie basiert durchweg auf Intel- bzw. AMD-Architektur. Die folgende Tabelle stellt die charakteristischen Daten dar.

Komponente	Systeme	Systeme
	M1 'alpspitze'	M2 'erongo'
Prozessor	4*Intel Xeon 2,4GHz	2*AMD Athlon 1800 +
Hauptspeicher	3 GB	3 GB
Festplattenkapazität	2*SCSI 34 GB	1*SCSI 18 GB,2*IDE 148 GB
	M3 'erongo vmw'	M4 'tofana'
Prozessor	2*AMD Athlon 1800 + (VMw)	1*Intel 2,26GHz
Hauptspeicher	1 GB (VMware)	1,5 GB
Festplattenkapazität	VMware 10 GB, 8 GB	1*IDE 80 GB
<i>/home/saturn/work/DEMOMESSB/result/s_reportde/env_hardware.csv</i>		

3.2.2 Software

Die eingesetzte Systemsoftware basiert auf der SUSE Linux Distribution 10.0 und den hier üblichen systemnahen Komponenten. Über einen Apache Webserver und die **Groupware PHProjekt** wird auf das DBMS PostgreSQL zugegriffen. Für die Virtualisierungsvariante bot sich der VMware-Server als etablierte Lösung an.

Komponente	Systeme	Systeme
	M1 'alpspitze'	M2 'erongo'
Betriebssystem	Linux 10.0	Linux 10.0
Kernel	2.6.13-15-smp	2.6.13-15
Virtualisierung	-	-
Web-Server	Apache2	Apache2
Datenbank	PostgreSQL	PostgreSQL
Testapplikation	PHProjekt 4.2	PHProjekt 4.2
	M3 'erongo vmw'	M4 'tofana'
Betriebssystem	Linux 10.0	Linux 10.0
Kernel	2.6.13-15-default	2.6.13-15.8
Virtualisierung	VMware Server	-
Web-Server	Apache2	Apache2
Datenbank	PostgreSQL	PostgreSQL
Testapplikation	PHProjekt 4.2	PHProjekt 4.2
<i>/home/saturn/work/DEMOMESSB/result/s_reportde/env_software.csv</i>		

4 Definition der Messungen

Die Messungen erfolgten gemäß DIN 66273 / ISO 14756 in Methodik, Durchführung und Auswertung. Im Folgenden findet eine Beschreibung von:

- Auftragsarten
- Auftragskettenarten
- Benutzertypen
- Denkzeiten (AGVZ)
- Zeitklassenschema

statt.

4.1 Beschreibung der Auftragsarten

Folgende Tabelle beschreibt die Auftragsarten als elementare Aktionen, die von einem Benutzer ausgeführt werden. Sie beschreiben jeweils einen Vorgang, der in verschiedenen Auftragsketten gleich ausgeführt wird.

Für die vorher beschriebenen Nutzer-Aktionen werden Denkzeiten im Bereich von 60 Sekunden und Antwortzeiten ab 60 Sekunden vorgegeben.

Diese Vorgaben erfüllen primär das Ziel der Demonstration. In der Realität werden deutlich niedrigere Werte gefordert.

AA	AAname	AGVZ	Zeitklasse	Kommentar
1	kontakte	pt60	tc60	Neuer Kontakt
2	notiz	pt60	tc60	Neue Notiz
4	forum	pt60	tc60	Eintrag Forum
5	todo	pt60	tc60	Neue Aufgabe

/home/saturn/work/DEMOMESSB/result/Series/iso_task.csv

4.2 Beschreibung der Kettenarten

Die Last wurde gemäss den Auftragsketten in folgender Tabelle und aus den elementaren Auftragsarten in der vorherigen Tabelle Auftragsarten definiert.

Die Auftragsketten bestehen hier nur aus gleichartigen Aufträgen.

Kette	Auftrag	Auftragsart	Kommentar
Ch0	kontakte	1	Nur neue Kontakte
Ch1	notiz	2	Nur neue Notizen
Ch2	todo	5	Nur neue Aufgaben
Ch3	forum	4	Nur Einträge Forum
<i>/home/saturn/work/DEMOMESSB/result/Series/iso_chains.csv</i>			

4.3 Beschreibung der Benutzertypen (EAG)

Für die Anwendung wurden verschiedene Benutzertypen definiert. In folgender Tabelle sind die Benutzertypen und die auszuführenden Auftragsketten nebst Kettenanteilen aufgelistet.

Hier führt ein Benutzer die vier Auftragsarten im Mittel gleichwertig aus (zu je 25%); es gibt keine besonderen Auftragsketten.

Die Definition der Last gegenüber Benutzertypen und die Anzahl der Benutzer eines Typs findet im Kapitel zur Messdurchführung der jeweiligen Messung statt.

Eagno	Name	Benutzer	Zyklen	C_CYC	Kette	Anteil(%)	Kommentar
0	Utype1	450	20	1	Ch0	25	25% Kette 0
	Utype1	450	20	1	Ch1	25	25% Kette 1
	Utype1	450	20	1	Ch2	25	25% Kette 2
	Utype1	450	20	1	Ch3	25	25% Kette 3
<i>/home/saturn/work/DEMOMESSB/result/Series/iso_utype.csv</i>							

4.4 Zeitvorgaben

4.4.1 Vorbereitungszeiten (AGVZ)

Die **Denkzeiten** — Zeiträume, die dazu dienen Pausen während der Bearbeitung durch einen Anwender zu simulieren — werden in untenstehender Tabelle definiert.

AGVZtyp	EAGtyp	MW(sec)	Stdabw.(sec)	Kommentar
pt5	all_types	5.000	1.000	‘5 Sekunden Denkzeit‘
pt30	all_types	30.000	15.000	‘30 Sekunden Denkzeit‘
pt45	all_types	45.000	20.000	‘45 Sekunden Denkzeit‘
pt60	all_types	60.000	30.000	‘60 Sekunden Denkzeit‘
pt80	all_types	80.000	40.000	‘80 Sekunden Denkzeit‘
<i>/home/saturn/work/DEMOMESSB/result/Series/iso_ptime.csv</i>				

4.4.2 Durchlaufzeitforderungen (ADLZ)

Die Durchlaufzeitforderungen wurden für alle Auftragsarten nach untenstehender Liste definiert. Die Einhaltung der Durchlaufzeitforderung schlägt sich bei der Auswertung im L3-Wert nieder.

Zeitklasse	Anteil(%)	Zeit(sec)	Kommentar
tc30	50	30.000	‘50% < 30 Sekunden‘
	80	60.000	‘80% < 60 Sekunden‘
	100	100.000	‘100% < 100 Sekunden‘
tc60	50	60.000	‘50% < 60 Sekunden‘
	80	80.000	‘80% < 80 Sekunden‘
	100	160.000	‘100% < 160 Sekunden‘
tc80	50	80.000	‘50% < 80 Sekunden‘
	80	100.000	‘80% < 100 Sekunden‘
	100	200.000	‘100% < 200 Sekunden‘
<i>/home/saturn/work/DEMOMESSB/result/Series/iso_tclass.csv</i>			

4.5 Durchgeführte Messungen

Nach Ermittlung der Lastgrenzen in Probemessungen wird je System ein Messlauf mit der festgelegten Nutzerzahl durchgeführt (M1 bis M4). In jeder Messung werden alle Auftragsarten bei statistischer Verteilung weitgehend gleichwertig ausgeführt.

4.6 Erläuterung zu den Messergebnissen

Eine Messung setzt sich aus einer Lastmessung nach DIN 66273 / ISO 14756 und optionalem Monitoring des Testobjekts zusammen.

4.6.1 Ergebnisse der DIN-Messungen

Der Durchführung der Messung und Auswertung der Messergebnisse geht die Definition eines theoretischen Referenzsystems voraus. Die darin aufgestellten Forderungen sind bei der Auswertung von besonderer Bedeutung. Eine "sinnvolle" Wahl der Forderungen nach DIN ist Voraussetzung für ein brauchbares Modell eines theoretischen Referenzsystems.

Falls die Messergebnisse sehr stark von den Forderungen abweichen, sollte immer auch in Betracht gezogen werden, das Referenzmodell zu bedenken und gegebenenfalls anzupassen.

Ein Wert erscheint in den Tabellen der Messergebnisse grün hinterlegt, falls bei der Messung die Forderungen eingehalten wurden. Ein gelber Eintrag in der Tabelle deutet auf eine Abweichung vom Referenzmodell von bis zu 25% hin, eine Abweichung die zwar bemerkt

werden sollte, aber nicht zwingend signifikant ist. Ob solche Ergebnisse toleriert werden können liegt im Ermessen des Betrachters. Rot hinterlegte Werte dagegen weichen bereits deutlich von den Vorgaben des Modells für die Referenzmaschine ab und zeigen Problemstellen auf, die unbedingt analysiert werden sollten. Basiert die Messung auf einem angemessenen Referenzmodell, geben die Daten des Monitoring Hinweise auf Ursachen im System, wie z.B. Engpässe bei den Systemressourcen, Prozessorlasten und Speicherplatz.

4.6.2 Ergebnisse des System-Monitoring

Während der Messungen wurden in 5-Sekunden-Intervallen die Auslastung der CPU, des Arbeitsspeichers und des Netzwerks mitprotokolliert. Diese Daten können mitunter helfen wichtige Zusammenhänge zwischen der Auslastung des Testobjekts und seinem Verhalten zu interpretieren. Sie sind allerdings, wie alle Daten aus Messreihen, immer vor einem entsprechenden Hintergrund zu deuten. Die Betrachtung der Daten alleine leistet keinen unmittelbaren Beweis, vielmehr liefert sie eine quantitative Basis, wie in der Statistik üblich.

5 Messreihe 1 Durchführung und Auswertung

5.1 Übersicht der Messläufe

Im Rahmen der Messreihe wurden je ein Messlauf pro System zur Betrachtung herangezogen. Die Auswahl erfolgte über eine Annäherung an die Erfüllung der DIN-Werte durch Variation der Zahl der simulierten Nutzer in Schritten von 50. Es ergaben sich so Nutzerzahlen von 150 bei den leistungsschwächeren Systemen bis zu 450 beim leistungsfähigsten System ('alpspitze').

M	n	MessID	Datum	Zeit	Start	Ende	Dauer	Ü
M1	450	alpspitze_450+0+0	07/18/06	17:38:54	2.4	2412.6	1959.9	0.93
M2	300	erongo_300+0+0	06/28/06	13:16:18	2.7	2778.5	2035.6	0.89
M3	150	erongo-vmw-a_150+0+0	06/29/06	14:14:19	2.5	1668.6	1336.9	0.92
M4	150	tofana_150+0+0	06/30/06	12:00:14	2.7	1655.7	1300.9	0.91
<i>Ausgewertete Messungen</i>								

5.2 Ergebnisse der Messungen

Die Ergebnisse der vier Messläufe werden in der folgenden Tabelle zusammengefasst. Die DIN-Werte L1 und L2 erfüllen in allen Fällen die Vorgaben. Die ungewöhnlich hohen Werte beim System 'tofana' (Messlauf M4) weisen darauf hin, dass hier mit 150 Nutzern die Leistungsreserven nicht ausgeschöpft werden. Das System 'erongo' hingegen befindet sich mit 300 Nutzern über seiner Leistungsgrenze, was durch die Nichteinhaltung der Termintreue L3 bei allen Auftragsarten dokumentiert wird.

In der Meßreihe M1 ('alpspitze') führt die hohe Nutzerzahl zu einer deutlicheren Gewichtung der vier Auftragsarten als bei den anderen Messreihen. Die Durchlaufzeitbewertung L2 bzw. die Antwortzeiten weisen hier signifikante Unterschiede auf. Dies führt dann bei der anspruchsvollsten Auftragsart 'todo' zur Verfehlung der Termintreue.

Das mit VMware virtualisierte System 'erongo vmw' (M3) verhält sich bei gleicher Nutzerzahl (150) in der Auftragsart 'kontakte' vergleichbar zum realen System 'tofana' (M4). Bei anspruchsvolleren Auftragsarten wird jedoch eine Verschlechterung der Antwortzeiten erkennbar. Die geringere Hauptspeicherausstattung könnte zu dieser Abweichung beitragen.

Die Durchsatzbewertung - der numerische Wert für die Anzahl abgearbeiteter Auftragsarten pro Stunde, Spalte "B" in der Tabelle - bietet eine Variante zum Leistungsvergleich der Systeme für die gemessene Anwendung: Das System 'erongo' erreicht ca. zwei Drittel der Leistung des stärksten Systems 'alpspitze'; sein virtualisiertes Gastsystem 'erongo vmw' ca. 75% seines realen Host-Systems.

Im Anschluss an die Übersichtstabelle werden die Ergebnisse der Messläufe (SoM - Series of Measurement) zu jeder Auftragsart dargestellt.

M	Typ	Auftragsart	N	t_AN	B	L1	L2	L3	Graph
M1	1	kontakte	2246	14.31	4186	1.51	6.01	1.00	
M2	1	kontakte	1482	38.11	2673	1.45	2.26	0.91	
M3	1	kontakte	753	3.21	2053	2.22	26.77	1.00	
M4	1	kontakte	753	5.50	2108	2.28	15.64	1.00	
SoM									3 - 5
M1	2	notiz	2269	30.85	4203	1.52	2.79	1.00	
M2	2	notiz	1510	40.57	2728	1.48	2.12	0.92	
M3	2	notiz	763	6.32	2072	2.24	13.60	1.00	
M4	2	notiz	763	4.91	2130	2.31	17.50	1.00	
SoM									6 - 8
M1	4	forum	2248	47.93	4128	1.49	1.79	1.00	
M2	4	forum	1517	43.00	2708	1.47	2.00	0.92	
M3	4	forum	756	8.13	2046	2.22	10.58	1.00	
M4	4	forum	756	4.23	2107	2.28	20.34	1.00	
SoM									9 - 11
M1	5	todo	2237	58.54	4102	1.48	1.47	0.93	
M2	5	todo	1491	45.09	2686	1.45	1.91	0.92	
M3	5	todo	728	9.17	1977	2.14	9.38	1.00	
M4	5	todo	728	4.87	2037	2.21	17.65	1.00	
SoM									12 - 14
<i>Ergebnistabelle</i>									

5.3 Auftragsart 'kontakte'

5.3.1 DIN/ISO Bewertungsgrößen

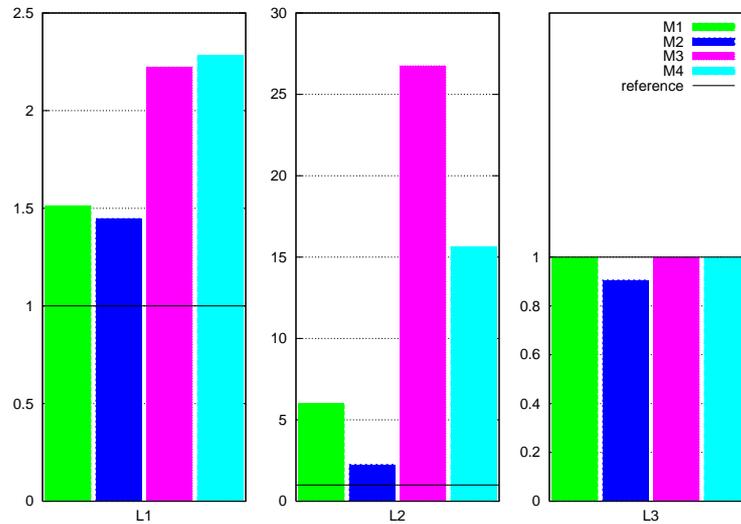


Abbildung 3: DIN/ISO Bewertungsgrößen 'kontakte'

5.3.2 Antwortzeithistogramm 'kontakte'

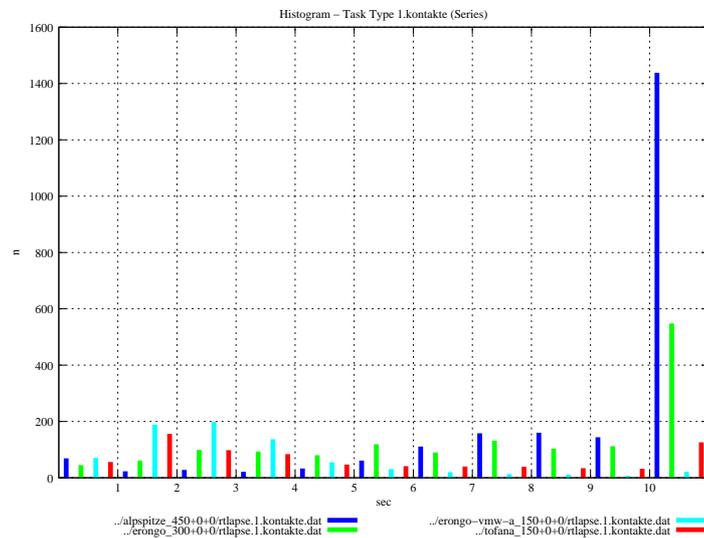


Abbildung 4: Antwortzeithistogramm 'kontakte'

5.3.3 Summenhistogramm 'kontakte'

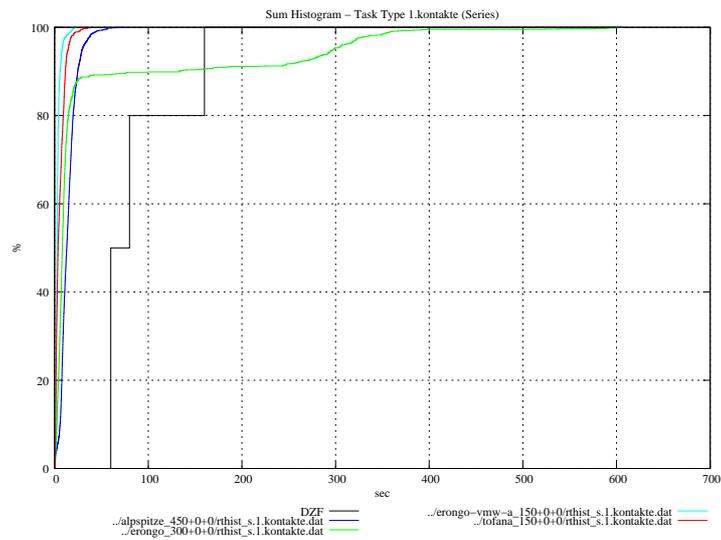


Abbildung 5: Summenhistogramm 'kontakte'

5.4 Auftragsart 'notiz'

5.4.1 DIN/ISO Bewertungsgrößen

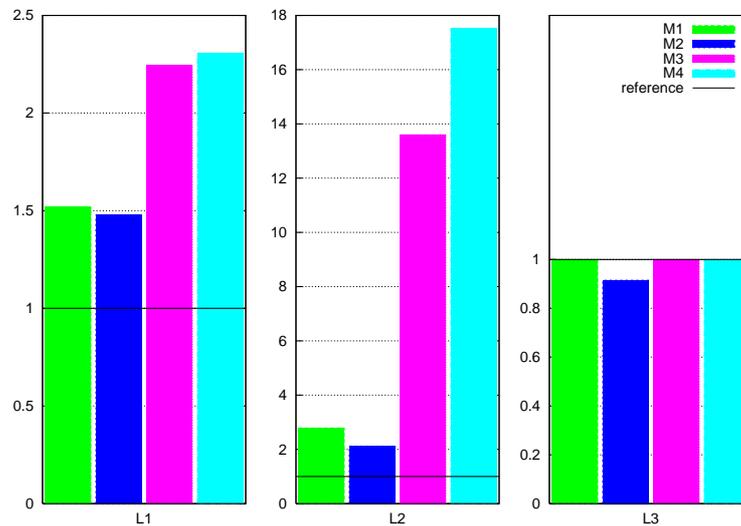


Abbildung 6: DIN/ISO Bewertungsgrößen 'notiz'

5.4.2 Antwortzeithistogramm 'notiz'

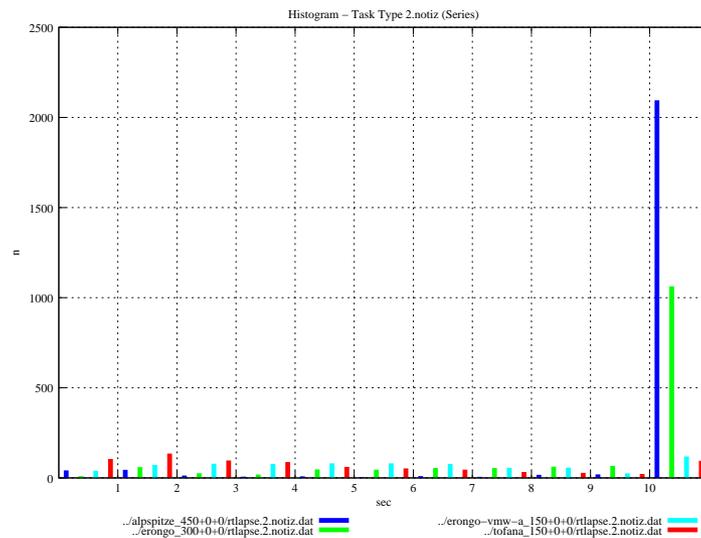


Abbildung 7: Antwortzeithistogramm 'notiz'

5.4.3 Summenhistogramm 'notiz'

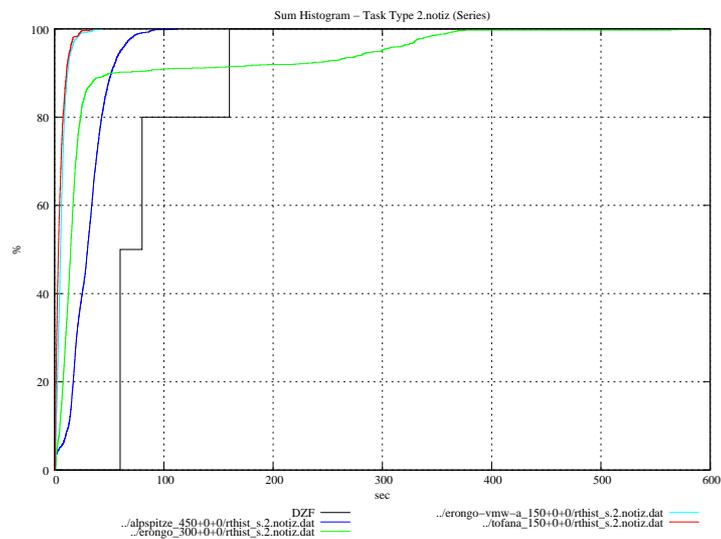


Abbildung 8: Summenhistogramm 'notiz'

5.5 Auftragsart 'forum'

5.5.1 DIN/ISO Bewertungsgrößen

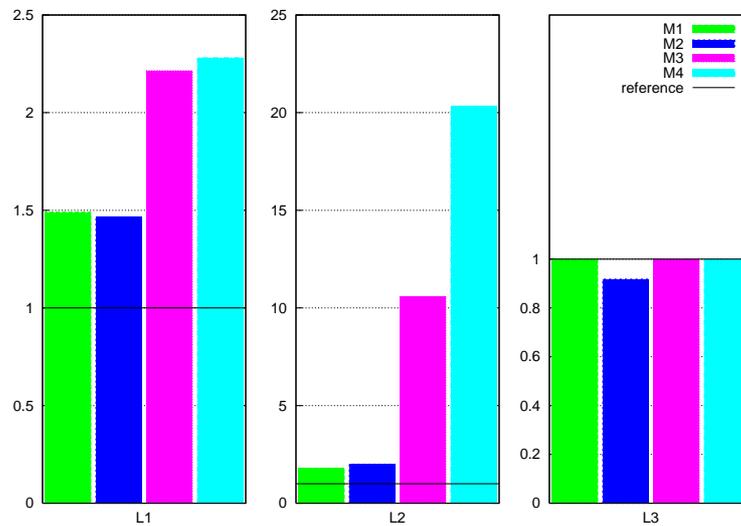


Abbildung 9: DIN/ISO Bewertungsgrößen 'forum'

5.5.2 Antwortzeithistogramm 'forum'

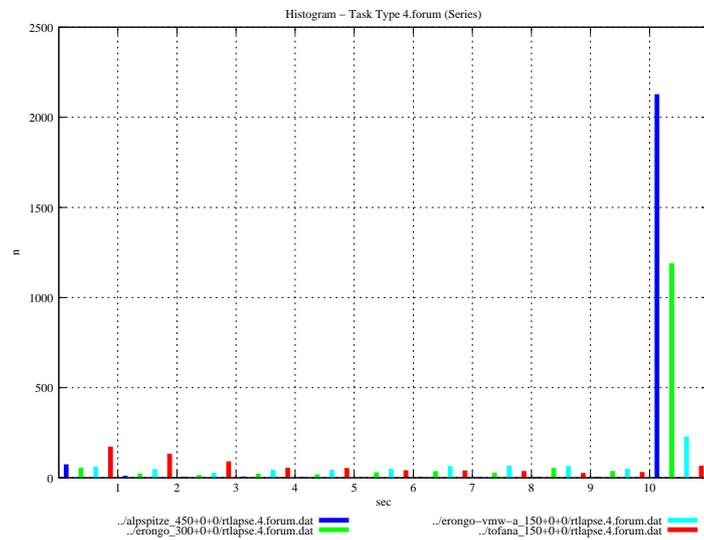


Abbildung 10: Antwortzeithistogramm 'forum'

5.5.3 Summenhistogramm 'forum'

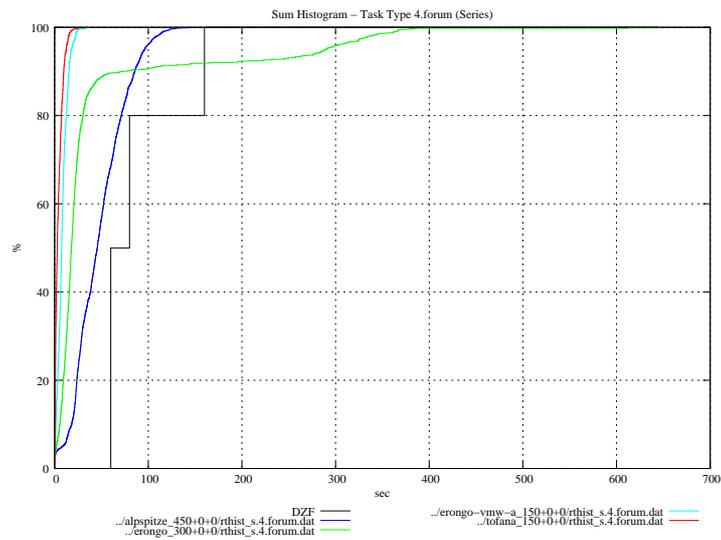


Abbildung 11: Summenhistogramm 'forum'

5.6 Auftragsart 'todo'

5.6.1 DIN/ISO Bewertungsgrößen

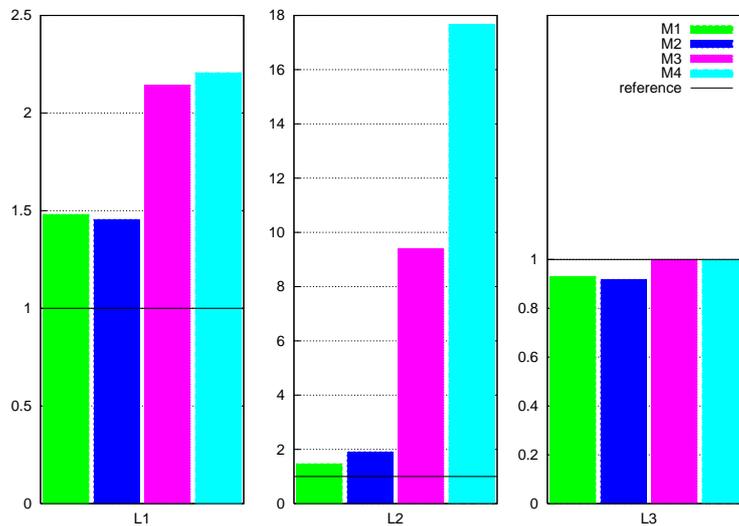


Abbildung 12: DIN/ISO Bewertungsgrößen 'todo'

5.6.2 Antwortzeithistogramm 'todo'

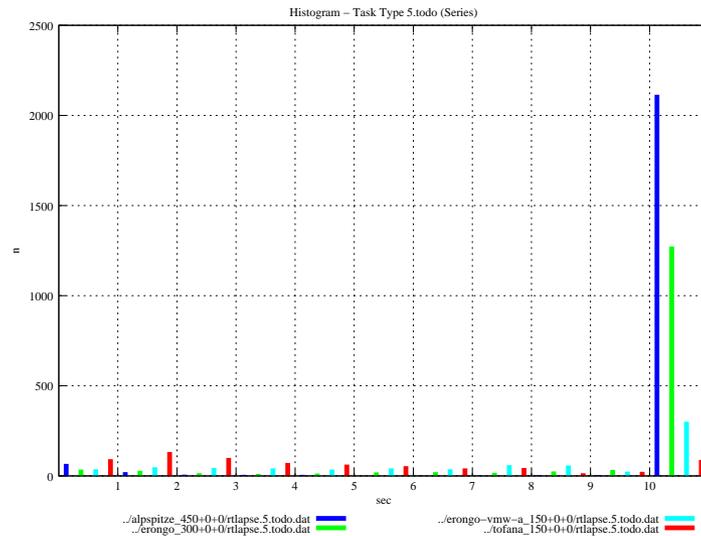


Abbildung 13: Antwortzeithistogramm 'todo'

5.6.3 Summenhistogramm 'todo'

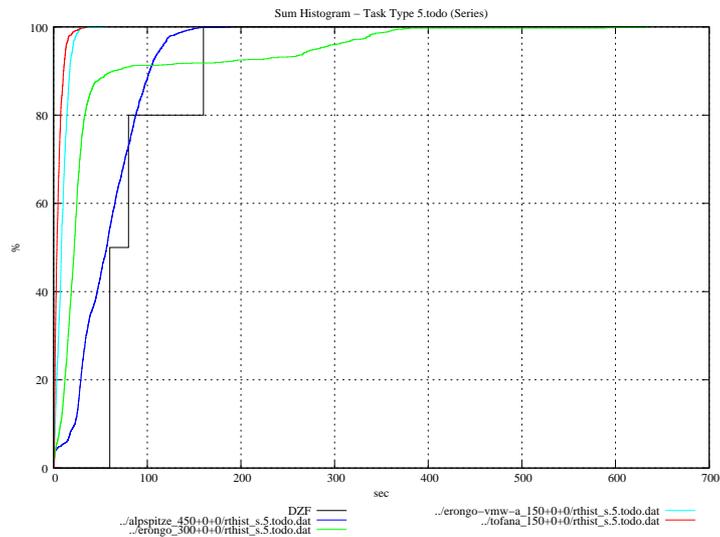


Abbildung 14: Summenhistogramm 'todo'

5.7 System-Monitoring

5.7.1 Monitoring 'alpspitze'

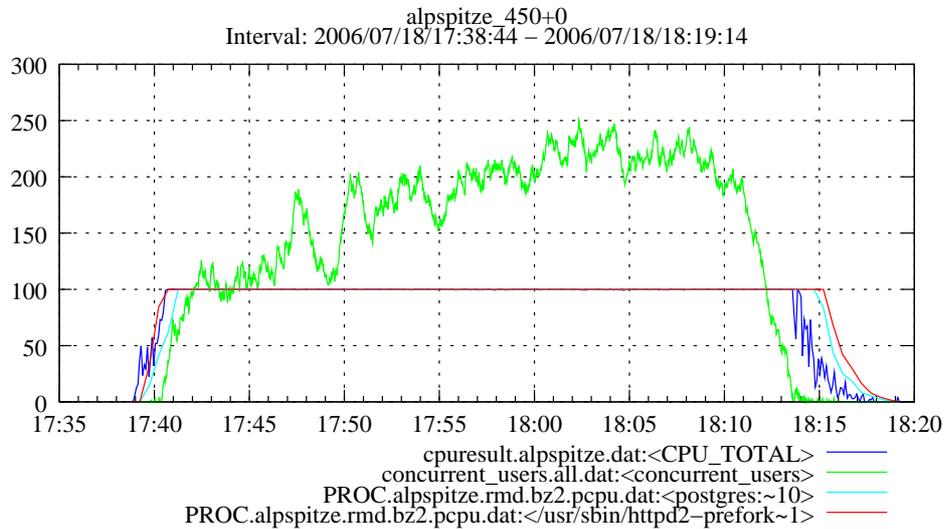


Abbildung 15: Monitoring 'alpspitze'

5.7.2 Monitoring 'erongo'

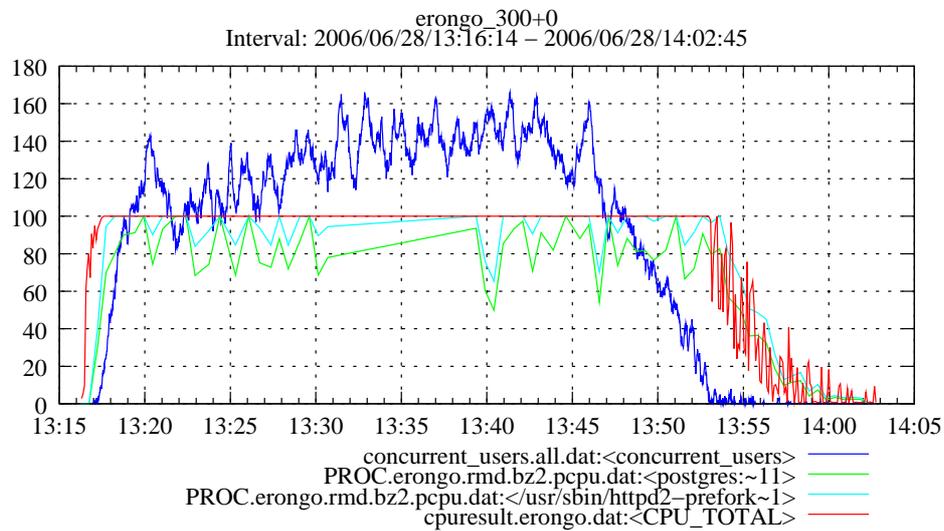


Abbildung 16: Monitoring 'erongo'

5.7.3 Monitoring 'erongo vmw'

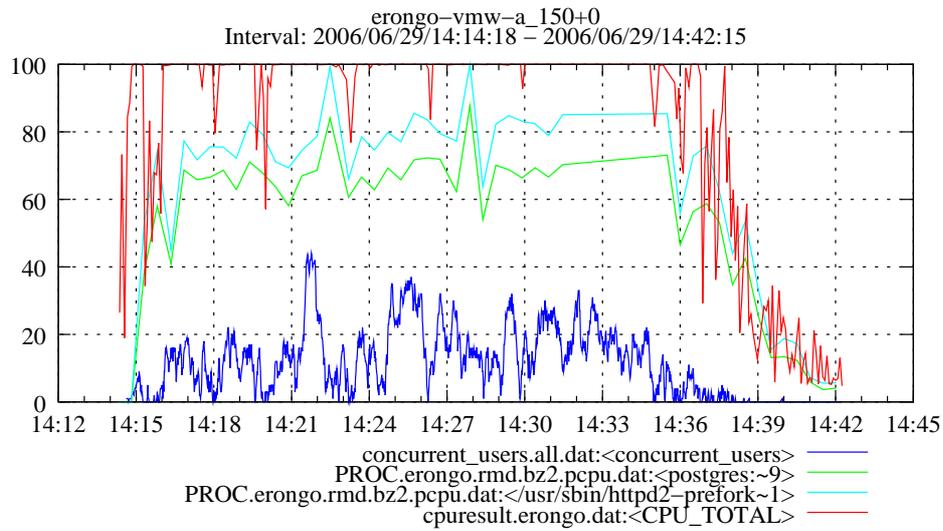


Abbildung 17: Monitoring 'erongo vmw'

5.7.4 Monitoring 'tofana'

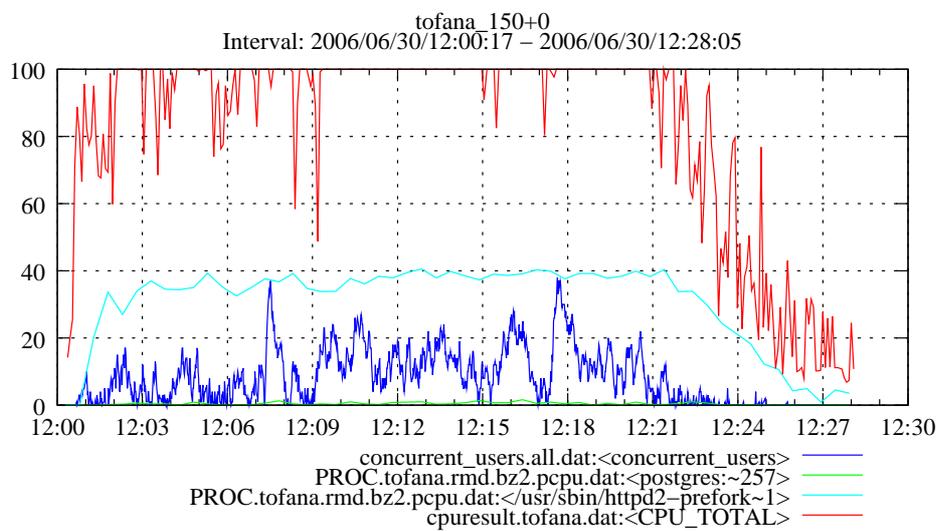


Abbildung 18: Monitoring 'tofana'

A Anhang

A.1 Die DIN 66273 / ISO 14756 Methode

Seit 1991 gibt es die nationale Norm DIN 66273, die inzwischen in die internationale Norm ISO 14756 überführt worden ist, Teil 1, mit dem Titel: "*Messung und Bewertung der Leistung von DV-Systemen*"¹ Sie beschreibt ein Begriffssystem sowie Mess- und Bewertungsverfahren, die den in der Einleitung aufgestellten Kriterien genügen. Die Norm wurde unter Beteiligung von Großanwendern und Herstellern entwickelt. Die DIN 66273 / ISO 14756 ist eine methodische Vorgehensweise der Leistungsbeurteilung von komplexen IT-Systemen. Die Bewertung erfolgt gegen ein theoretisches Referenzmodell. Die Methode DIN/ISO unterstützt bei: der Modellierung und Generierung von Lasttest-Szenarien der statistischen Auswertung von Lastmessungen. Durch die Auswertung gegen ein Zeitklassenschema für die erwarteten Antwortzeiten erhält man nicht nur statistische Informationen über das Antwortzeitverhalten sondern eine Bewertung der Messergebnisse, die eine eindeutige Beurteilung der Leistungsfähigkeit des vermessenen IT-Systems zulässt.

A.2 Durchführung einer Leistungsmessung nach DIN

Eine Leistungsmessung nach DIN wird in 3 Stufen durchgeführt:

- Erstellung des Anforderungsprofils
- Durchführung der Messung
- Auswertung und Berechnung der DIN-Bewertungsgrößen

A.2.1 Erstellung eines Anforderungsprofils

Das Anforderungsprofil umfasst die Beschreibung der auf dem Zielsystem arbeitenden Endbenutzer. Zunächst müssen die einzelnen Aufträge, die dem IT-System zu übergeben sind, genau beschrieben werden. Dies erfolgt in der Regel durch ein Manuskript mit den notwendigen Aktionen. Dazu müssen dem Benutzer Zeitvorgaben gemacht werden: wie lange soll nach einer Bildschirmausgabe gewartet werden, bis ein neuer Auftrag über die Tastatur oder die Maus eingegeben wird? Mit diesen Wartezeiten (auch Denkzeiten genannt) werden unvermeidliche Denkpausen, Arbeitspausen und andersartige Tätigkeiten (z.B. Gespräch mit einem Kunden am Bankschalter) nachgebildet. Schließlich soll festgelegt werden, wie lange die Bearbeitung eines Auftrags dauern darf. Dabei sollen sinnvolle Zeitforderungen vorgegeben werden: bei einem Kundengespräch in einem Reisebüro stört eine Bearbeitungszeit für eine Platzbuchung von 5 bis

¹Dirlewanger, Werner: Messung und Bewertung der DV-Leistung auf Basis der Norm DIN 66273/Werner Dirlewanger, Heidelberg, Hüthig-Verlag, 1994

10 Sekunden sicher niemand, wogegen eine Bearbeitungszeit von einer hundertstel Sekunde wenig praktischen Nutzen hat. Die Antwortzeitforderungen werden bei der **DIN 66273** in Form von Zeitklassenschemata vorgegeben, in welchen angegeben wird, welcher Anteil der Antwortzeiten welchen Grenzwert nicht überschreiten darf.

Beispiel: Antwortzeitforderung für Auftragsart Flugbuchung

40 % nicht länger als 5 sec

80 % nicht länger als 15 sec

100 % nicht länger als 30 sec

Die Anzahl der Zeitklassen, die in einer Antwortzeitforderung verwendet werden (im Beispiel sind es 3), kann der Anwender entsprechend seinen Bedürfnissen frei wählen. Nicht jeder Benutzer einer Benutzergruppe soll die gleiche Folge von Aufträgen mit jeweils gleichen Denkzeiten eingeben. Dies würde einem völlig praxisfremden Synchronarbeiten der Benutzer entsprechen. Vielmehr werden die Aufträge aus einem bestimmten Vorrat zufällig zusammengestellt und auch die Denkzeiten werden - unter Vorgabe eines Mittelwertes und einer Streuung - von einem Zufallsgenerator ermittelt. Dabei muss selbstverständlich die Tatsache berücksichtigt werden, dass bestimmte Aufträge nur in einer bestimmten Reihenfolge durchführbar sind (Bildung von Auftragsketten).

A.2.2 Durchführung einer Messung

Bei der Durchführung der Messung sind einige statistische Elemente zu beachten. Der Lasttreiber simuliert an 2 Stellen das statistische Verhalten einer realen Benutzerschaft.

- Anteil (Auftrittswahrscheinlichkeit) einer Auftragskette eines Benutzers
- Vorgabe der (Denkzeit) durch Mittelwert und Standardabweichung

Ein Lasttreiber kann die benötigten Zufallszahlen auf zwei Arten erzeugen:

- Freilaufender Zufallsgenerator mit entsprechender Parametrisierung

Vorteil: echter Zufall, sehr gute statistische Mischung der Benutzer

Nachteil: Einhaltung der gewünschten Vorgabewerte erst nach längerer Messzeit hinreichend genau

- Urnenmodell

Vorteil: genaue Einhaltung der Vorgabewerte auch bei kurzen Messzeiten

Nachteil: bei sehr kleinen Urnenfüllungen kein echtes Zufallsgeschehen mehr

Das Urnenmodell wird bei **s_atur** immer für die Kettenanteile benutzt, für die Denkzeiten wird alternativ auch der Zufallsgenerator eingesetzt. Als Folgerung dieser Tatsache ist es

notwendig Messzyklen zu definieren, die jede Auftragskette eines Benutzers *mindestens einmal* beinhaltet. Wird das Urnenmodell auch auf Denkzeiten angewendet, so muss jede Auftragsart *mindestens zweimal* auftauchen, um die Einhaltung von Mittelwert und Standardabweichung zu ermöglichen.

A.2.3 Auswertung einer Messung

Nach Durchführung der Messung mit einem der **s_atern** -Treibersystem steht für jeden simulierten Benutzer ein Recordfile zur Verfügung, welches die Datenströme zum und vom System, sowie die dazugehörigen Zeitstempel enthält. Mit der **s_atern** -Auswertesoftware können daraus sämtliche Bewertungsgrößen nach DIN 66273 berechnet werden. Die DIN 66273 kennt drei Bewertungsgrößen:

L1: Durchsatz

L2: Auftragsdurchlaufzeit (Antwortzeit)

L3: Termintreue

Alle drei Bewertungsgrößen sind Relativzahlen im Verhältnis zu Referenzwerten. Die Referenzwerte werden für eine theoretische Referenzmaschine ermittelt. Das ist ein System, welches in seinem Antwortzeitverhalten genau den aufgestellten Antwortzeitforderungen entspricht. Die Größe L1 beschreibt die Anzahl der bearbeiteten Aufträge in der Zeiteinheit (relativ zur Referenzmaschine), die Größe L2 gibt die mittlere Antwortzeit im Verhältnis zur Referenzmaschine wieder und die Größe L3 gibt den Anteil der termingerecht erledigten Aufträge im Verhältnis zur Anzahl der erteilten Aufträge an.

Die DIN-Bewertungsgrößen sollten nicht unter einem Wert von 1.0 liegen. Ein Wert unter 1.0 bedeutet, dass für die betrachtete Auftragsart das System nicht den vorher im Anforderungsprofil aufgestellten Forderungen entspricht.

Um die Grenzlast eines Systems zu ermitteln führt man eine Messreihe durch, bei der die Zahl der simulierten Benutzer so lange gesteigert wird, bis mindestens eine der DIN-Bewertungsgrößen unter 1.0 fällt. In Abbildung 19 ist in einem Beispiel der Verlauf der DIN-Bewertungsgrößen abhängig von der Benutzerzahl dargestellt. Die jeweils drei unterschiedlichen Kurven 1, 2 und 3 stellen den Verlauf für 3 verschiedene Auftragsarten dar. Diejenige DIN-Bewertungsgröße, die bei der niedrigsten Benutzerzahl unter 1 sinkt, ist die Antwortzeitbewertung L2 von Auftragsart drei (bei etwa 270 Benutzern).

Neben den eigentlichen Bewertungsgrößen enthält eine DIN-Auswertung noch eine Reihe von Kontrollwerten, welche die Funktion des Treibers und die statistische Qualität der Messung

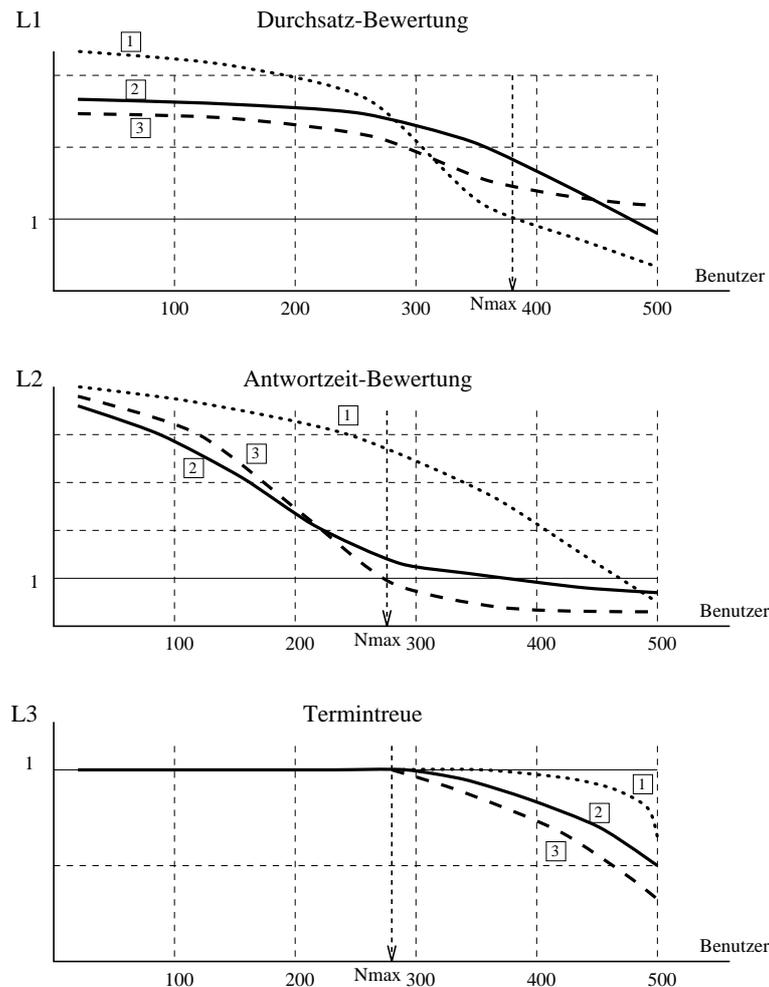


Abbildung 19: DIN-Bewertungsgrößen

beurteilen. So muss nach geprüft werden, ob die als Vorgabe gemachten Denkzeit-Verteilungen (Mittelwert und Standardabweichung) tatsächlich eingehalten worden sind. Außerdem müssen die Anteile der verschiedenen Auftragsketten, die jeder simulierte Benutzer in zufälliger Folge abgesetzt hat, den für die jeweilige Benutzergruppe vorgegebenen Mittelwerten entsprechen. Zur Beurteilung der statistischen Qualität der Messung wird das Konfidenzintervall des Mittelwertes der Antwortzeit berechnet. Über die Größe der noch hinnehmbaren Abweichungen des Treibers bei Denkzeiten und Auftragsanteilen, sowie über die maximale Breite des Konfidenzintervalls und die Höhe des Konfidenzniveaus macht die DIN 66273 keine Zahlenangaben. Die Zahlenwerte sollen aber vor der Messung festgelegt werden und bei Messreihen und unterschiedlichen Messobjekten immer gleich bleiben.

B Glossar

AA *siehe Auftragsart*

ADLZ Auftragsdurchlaufzeit:
Systemreaktionszeit + Netzlaufzeiten

AGVZ Auftraggebervorbereitungszeit:
Denkzeit + Eingabezeit eines Benutzers

AKA *siehe Auftragskettenart*

Auftragsketten Abfolge von Auftragsarten

Auftragsart Aktion eines Benutzers, mit Durchlaufzeitforderung und Auftraggebervorbereitungszeit. Die Auftragsart kann mit Transaktion gleichgesetzt werden

B Die tatsächliche Belastung B beschreibt das Verhältnis zwischen der Anzahl der erteilten Aufträge einer Auftragsart und der Dauer der Messung
Formel:

$$B = \frac{N}{\text{Messdauer}}$$

β vom Kunden definierte, theoretische Referenzbelastung (1/sec)

D(90.0)/% Messfehler in % vom Mittelwert

d(90.0)/s gibt den Messfehler in sec an. Die Ziffer (90.0) zeigt die Wahrscheinlichkeit an. Einfach ausgedrückt: Der Messfehler ist mit einer Wahrscheinlichkeit von 90 % nicht größer als der ermittelte Wert. Informationen können den DIN-Seminarunterlagen von Zott + Co entnommen werden.

DF Durchlaufzeitforderung:
Antwortzeitforderung

E E(j) (erbrachte Leistung) ist die Anzahl der zeitgerecht erledigten Aufträge einer Auftragsart je Zeiteinheit (Einheit: 1/sec).

Formel:

$$E = \frac{N * L_3}{\text{Messdauer}}$$

EAG *siehe Elementarer Auftraggeber*

EAG-Typ Benutzergruppe

Elementarer Auftraggeber Mitglied einer Benutzergruppe

L₁ Die Kennzahl L₁ (Durchsatzbewertung) errechnet sich aus dem Verhältnis von tatsächlicher Belastung (B) zu der errechneten Referenzbelastung (beta).

Formel:

$$L_1 = \frac{B}{\beta}$$

L₂ Die Kennzahl L₂ (Durchlaufzeit-Bewertung) errechnet sich, in dem t_{Ref} durch t_{AN} dividiert wird.

Formel:

$$L_2 = \frac{t_{Ref}}{t_{AN}}$$

L₃ Die Kennzahl L₃ beschreibt die Termintreue, d.h. Das Verhältnis der zeitgerecht erledigten Aufträge zur Anzahl der erteilten Aufträge in der Messzeit. Dabei bedeutet der Wert 1.0, dass die Durchlaufzeitforderungen zu 100 % erreicht worden sind (E ≤ 1).

Formel:

$$L_3 = \frac{E}{B}$$

Messlauf ist der Ablauf einer in sich geschlossenen Messung die in Form von DIN/ISO-Definitionen beschrieben ist.

Messreihe Bezeichnung für die Zusammenfassung mehrerer Messläufe

N Anzahl der Aufträge
Formel:

$$N = \text{Anzahl je Zyklus} \cdot \text{Zyklusanzahl} \cdot \text{Benutzer}$$

t_{AN} Mittelwert der gemessenen Durchlaufzeit (Einheit: sec)

t_{Ref} vor der Messung definierte Referenz-Durchlaufzeit (Einheit: sec);
ergibt sich als gewichtetes Mittel aus den Antwortzeitklassen