

# Benchmarking Projekt – Einführung



Projekt im Rahmen der Vorlesung:

**Computer Systems Performance**

**Analysis and Benchmarking (37-235)**

Dipl. Inf. Ing. Christian Kurmann

# Überblick

- **Projekt Performance Evaluation - Zielsetzungen**
- **SPEC - Wissenschaftliche Applikationsbenchmark**
- **Imbench - Synthetische Komponenten-Benchmark**
- **Winstone - Scriptgesteuerte Applikationsbenchmark**
- **WinBench - Synthetische Subsystem Benchmark**
  
- **Labors**
- **Beispieluntersuchungen**

# Projekt Performance Evaluation

- **Ziel**

**”Vorlesungstheorie” in die Praxis umsetzen, indem verschiedene PC-Systeme bzw. deren Komponenten sowie Softwareprodukte mit entsprechenden Benchmarks ausgemessen und die Resultate herausgearbeitet und dargestellt werden.**

- **Praktikumsbericht**

- Festlegung eines Evaluations-Ziels/Systems/Metriken
- Beschreibung der gewählten Evaluationstechnik/Workload
- Messresultate auf möglichst *verschiedenen* Rechnern bzw. verschiedenen Komponenten (CPU, Memory, Graphik, Disk, Netzwerk)
- Interpretation und Analyse von aufgetretenen Effekten und Resultaten

- **Arbeitsaufwand ca. 20-30 h (1P, ½U)**

# SPEC

- **Standard Performance Evaluation Corporation: Non-Profit Organisation** verschiedener Computerhersteller, Verleger und Consultants
- **Standardbenchmark zum Vergleichen von Computern über verschiedene Hardware- und OS-Plattformen**
- **SPEC CPU2000** enthält zwei Messfolgen: **CINT, CFP**
- **misst Prozessor, Speicher-Architektur und Compiler Performance**
- **wechselt ca. alle paar Jahre, da z.B. Compiler für SPEC Benchmarks optimieren werden, Lasten ändern**
- **grosse Problemstellungen, lange Laufzeiten, grosse Applikationsdiversität**
- **ausgelegt für rechenintensive Applikationen, die auf verschiedenen Plattformen zur Verfügung stehen**

# SPEC 95

- **CINT95**
- **go** - An internationally ranked go-playing program. INT-Performance, small Cache activity
- **m88ksim** - A chip simulator for the Motorola 88100 microprocessor.
- **gcc** - Based on the GNU C compiler version 2.5.3. 50 Input files result in the highest number of fork/exec/open system calls.
- **compress** - An in-memory version of the common UNIX utility. Memory and INT-Benchmark
- **li** - Xlisp interpreter. Inlining Optimization test.
- **jpeg** - Image compression/decompression on in-memory images. Shows of superscaler integer capabilities.
- **perl** - An interpreter for the Perl language. Word lookups and basic math. Tests libc routines: malloc, free, memcpy
- **vortex** - An object oriented database. Largest Benchmark in the suite, tests System TLB handler.

## SPEC Resultate:

<http://www.specbench.org/osg/cpu95/results/>

<http://www.intel.com/procs/perf/>

# SPEC 95

- **CFP95**
- **tomcatv** - Vectorized mesh generation. Sensitive to the speed of memory.
- **swim** - Shallow water equations. Can be parallelized well.
- **su2cor** - Monte-Carlo method. Can be parallelized well.
- **hydro2d** - Navier Stokes equations. Can be parallelized well.
- **mgrid** - 3d potential field.
- **applu** - Partial differential equations.
- **turbo3d** - Turbulence modeling. Large 1D FFT
- **apsi** - Weather prediction. Can not be easily parallelized.
- **fpppp** - From Gaussian series of quantum chemistry benchmarks. Can not be parallelized. Tests register allocation/performance.
- **wave5** - Maxwell's equations.

# SPEC 95 Resultate

- Ausführungszeiten für die einzelnen CINT Benchmarks sind gegeben
- Berechnung des Verhältnisses der Ausführungszeiten zur Basismaschine (SPARCstation 10 Model 40).
- Die einzelnen Verhältnisse entsprechen dem Median von mindestens drei Durchgängen.
- SPEC<int/fp>\_base95 wird berechnet aus dem Geometrischen Mittel der individuellen Verhältnisse.

SPEC CINT95 Results							
DELL OptiPlex GXpro				SPECint95 =	--		
				SPECint_base95 =	7.07		
SPEC license #	1326	Tested By:	CK, CS, ETH Zuerich	Test Date:	March '98	Hardware Avail: Now	Software Avail: Now
Hardware/Software Configuration for: OptiPlex GXpro							
	Hardware	Benchmark # and Name	Reference Time	Base Run Time	Base SPEC Ratio	Run Time	SPEC Ratio
Model Name:	Dell Optiplex GXpro	099.go	4600	569	8.08	--	--
CPU:	Intel Pentium Pro 200 MHz	124.m88ksim	1900	316	6.01	--	--
FPU:	Integrated	126.gcc	1700	264	6.44	--	--
Number of CPU(s):	2	129.compress	1800	294	6.12	--	--
Primary Cache:	8 KB (I) + 8 KB (D)	130.li	1900	249	7.63	--	--
Secondary Cache:	256 KB (I+D)	132.jpeg	2400	331	7.25	--	--
Other Cache:	none	134.perl	1900	233	8.15	--	--
Memory:	64 MB	147.vortex	2700	374	7.22	--	--
Disk Subsystem:	3 GB, 2.5 GB EIDE	SPECint_base95 (G. Mean)		7.07			
Other Hardware:	Ethernet Fast-Ethernet			SPECint95 (G. Mean)		--	
	Software						
Operating System:	Windows NT 4.0 Service Pack 3						
Compiler:	Microsoft Visual C++ 5.0						
File System:	FAT						
System State:	single user						

# SPEC 95 : HowTo

- **Starten von SPEC Benchmarks mit `runspec`:** übernimmt Compilation, Bereitstellen der Eingabedaten, Starten der Programme, Generierung der Resultate
- **Beispielkommando:**  
**`runspec -a validate -c my_config -i ref -n 3 -o ps int`**
  - **`a validate`:** action validate, führt alle notwendigen Aktionen aus wie Compilieren, Vorbereiten der Daten, Starten der Benchmarks, Vergleichen und Ausgeben der Resultate.
  - **`c my_config`:** wählt die zu verwendende Konfiguration. Diese enthält Angaben zum System und Compiler.
  - **`i ref`:** bestimmt die zu verwendenden Eingabedaten, test (Testdaten), train (zum Optimieren der Compiler args) und ref (Referenz-Daten).
  - **`n 3`:** wieviel mal der Benchmarks durchgeführt werden soll (für gültige Resultate mindestens 3 mal).
  - **`o ps`:** wählt das Ausgabeformat der Testresultate z.B. PostScript.
  - **`int`:** wählt den oder die auszuführenden Benchmarks. INTEGER-Benchmarks (SPEC CINT 95)



# Imbench

- **Synthetische Benchmark**: misst spezifisch die Performanz individueller Komponenten mit simplen, portablen Workloads
- **Nachteil**: Resultate repräsentieren nicht die Performanz eines Computersystems in real-life Situationen
- **Gnu GPL**, Larry McVoy (SGI), Carl Staelin (HP)
- **Resultatdatenbank als Vergleichsbasis**
- **Low-Level Benchmark**

# Imbench

- **Bandwidth benchmarks**

- Cached file read
- Memory copy (bcopy)
- Memory read

- **Latency benchmarks**

- Context switching
- Networking: connection establishment, pipe, TCP, UDP, and RPC hot potato
- File system creates and deletes

- **Miscellaneous**

- Processor clock rate calculation

- Memory write
- Pipe
- TCP

- Process creation
- Signal handling
- System call overhead
- Memory read latency

# Winstone

- Skriptgesteuerte System-Level Applikationsbenchmark
- allumfassendes Mass der Performance eines PC's unter Windows 2k / NT / 98, normalisiert zu einer Basismaschine
- **Business Applikationen** (populäre Applikationen, alltäglich benutzt durch die meisten Benutzer) und **High-End Applikationen** (bezieht sich auf spezialisierte Benutzer wie Applikationsentwickler, Photo Editor)
  - **Business WinMark, HighEnd WinMark**
- Installiert Demoverversionen der Applikationen und verändert Systemeinstellungen, heikel in produktiver Umgebung

# Winstone 99 : Applikationen

Suiten	Kategorie	Applikation
Business	Business Browser	Netscape Navigator® 4.04
	Business Corel WordPerfect®	Quattro® Pro 8, WordPerfect® 8
	Business Lotus SmartSuite®	1-2-3® 97, Word Pro® 97
	Business Microsoft Office®	Access 97, Excel 97, PowerPoint 97, Word 97
High End	High End Applications (Development, Graphic, CAD)	Adobe® Photoshop® 4.01, Adobe Premiere® 4.2, AVS/Express® 3.4, Bentley System's MicroStation® SE, Microsoft® FrontPage® 98, Microsoft® Visual C++® 5.0, Sonic Foundry® Sound Forge® 4.0
Dual-Processor Inspection	High End Applications	Adobe® Photoshop® 4.01, MicroStation® SE, Microsoft® Visual C++® 5.0

# Winstone 99 : Basismaschine

- Computer: Dell Dimension™ XPS P5-233 MMX
- Floating Point: Integriert
- Display Adapter: Matrox® Millennium II™
- Display Adapter Memory: 8192KB WRAM
- Display Driver: MGA-2164W-2064W-1064SG
- Display Mode: 1024 X 768 16 bits/pixel
- Display Refresh Rate: 60 Hz
- Hard Disk: IBM DHEA 26480 with 6.4GB
- Hard Disk Controller: Primary EIDE
- Off-Chip Processor Cache: 512KB Pipeline Burst
- Processor: intel Pentium x86 Family 5 Model 4, Step. 3
- Processor RAM: 32MB SDRAM
- Operating System: WindowsNT 4.0 Service Pack 3, Build 1381
- Bus Type: PCI
- **erreicht 10 Winstones für Business/HighEnd Test, bzw. 1 Winstone für jede Kategorie, Dauer 1 h für sämtliche Tests**

# WinBench 99

- **Subsystem-Level Benchmark (Graphik-, Disk, Prozessor/Memory, Video, CD-ROM Performance) für Windows 2k / NT / 98 Plattformen**
- **Synthetischer Nachvollzug der mit Winstone aufgezeichneten Grafik- und Diskoperationen für Business- und High-End Applikationen**
  - Business Graphics WinMark**
  - High-End Graphics WinMark**
  - Business Disk WinMark**
  - High-End Disk WinMark**

# WinBench 99 : Benchmarks

- **Synthetische Benchmarks**
  - **GDI/USER**: Graphics-Device- und USER-Interface Tests
  - **DISK WinMarks**: Disk Read/Write Sequential/Random
  - **DISK Transfer Rate**: Transferrate der Disk/Controller/Treiber-Kette
  - **DISK ReadCPU Utilisation**: CPU Beanspruchung während Disk Read
  - **CPUMark 99**: Processorleistung für 32 Bit Programme, Applikationsmix (Kompression, Wortanalyse, Textformatierung)
  - **FPU Winmark**: Processor-FloatingPoint Leistungsmessung, Applikationsmix (FFT, Poisson Gleichung, Lin. Gleichungssyst.)
  - **CD-Rom WinMark**: CD Rom Read Sequential/Random, Access Time, Applikationsmix (Business, Games und Unterhaltung, Reference und Lehre)
  - **Video WinMark, DirectDraw WinMark**

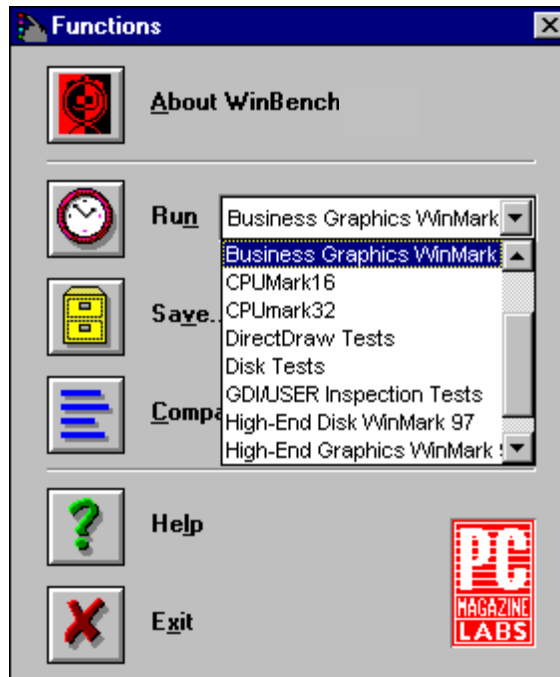
# WinBench 99: Einheiten

Test	Einheiten der Resultate:
<b>Disk</b>	
WinMark und Inspektion Tests	1000 Bytes/Sek
CPU Benutzung	Prozent benutzt (kleiner ist besser)
<b>Graphics</b>	
WinMark Tests	Keine Einheit
Inspektion Tests	Millionen Pixels/Sek
<b>Processor</b>	
CPUmark	Keine Einheit
<b>CD-ROM</b>	
WinMark und Inspektion Tests	1000 Bytes/Sek
Zugriffszeit	Millisekunden
CPU Benutzung	Prozent benutzt (kleiner ist besser)
<b>Video</b>	
Framerate/Qualität	Frames/Sek, fallengelassene Frames



# WinBench 99: Oberfläche

## Arbeitsoberfläche von WinBench99 / WinStone99



**Testauswahl / Teststart**

**Speichern von Resultaten**

**Laden von Vergleichstests in Excel**

**Hilfe / Erklärungen**

# Beispiel-Untersuchungen: SPEC

- **Compiler auf UltraSparc:** Vergleichen Sie den Sun WorkShop Compilers mit dem GNU gcc auf zwei Sun UltraSparc-Rechnern.
- **Betriebssysteme und Compiler für PC:** Vergleichen Sie die zwei typischen PC-Entwicklungsoberflächen Linux mit GNU gcc und Windows 2000 mit Microsoft Visual C++ auf einem Pentium III PC. Wie sind die Unterschiede zu erklären?
- **Gänzlich unterschiedliche Rechner:** Vergleichen Sie eine Sun Blade mit Solaris 8 und Sun WorkShop Compiler mit einem Dell Precision 420 mit Linux und GNU gcc.
- Vergleichen Sie die Resultate mit der SPEC Datenbank:  
<http://www.specbench.org/osg/cpu95/results>

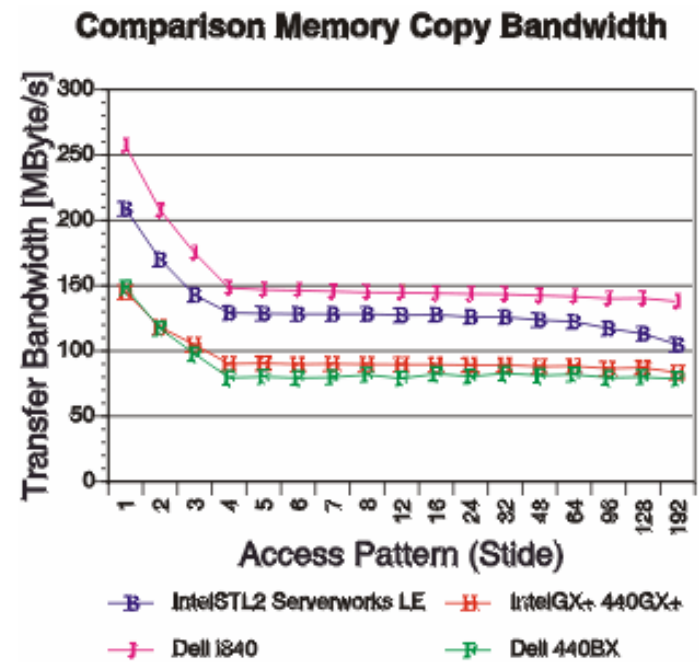
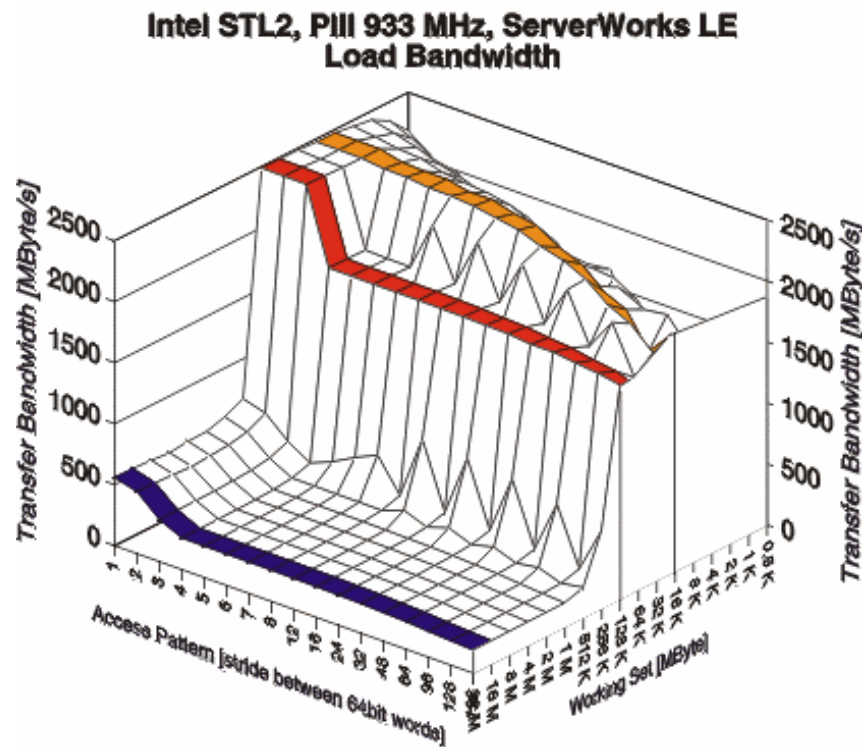
# Beispiele: WinBench/WinStone

- Vergleichen Sie die **Business Graphics WinMarks99** bzw. die **Business Winstone99** eines Pentium III Rechners mit verschiedenen Graphik Adaptern.
- Wieviel % der CPU Zeit wird für einen **Disk/Read Access** benötigt für ein SCSI und ein EIDE-System (mit und ohne DMA). Gibts Unterschiede?
- Messen Sie die **Disk/Read Transfer Rate** und **Disk/Read Access Random** einer durchschnittlichen EIDE Festplatte und einer **High-End Seagate Cheetah SCSI** mit 10'200 RPM. Was macht die richtige Konfiguration (PIO, DMA) aus.
- Vergleichen Sie **Winstone** Tests-Resultate mit den Beispielresultaten und den Resultaten der synthetischen **WinBench**.

# Beispiele: Imbench / ECT-memperf

- Kompilieren Sie **Imbench** unter LINUX, unter Windows und Solaris und untersuchen Sie die Resultate unter den verschiedenen Betriebssystemen
- Vergleichen Sie die Resultate von Imbench mit denen von WinBench (wo möglich)
- **ECT memperf** ist eine Methode um die Leistung eines Memorysystems zu charakterisieren. Sie erfasst zwei Aspekte der Memoryhierarchie:
  - *temporal locality*: variieren der Workingset-Grösse
  - *spacial locality*: variieren des Zugriffsmusters
- Anhand der resultierenden Bandbreiten-Werte können die Memorysysteme sehr genau verglichen bzw. Applikationen mit Hilfe der Resultate optimiert werden.
- <http://www.cs.inf.ethz.ch/cops/ECT>

# Beispiele: ECT-memperf



# Weitere Benchmarks

## Weitere Möglichkeiten:

- **Wireless Lan Performance Evaluation**
- **Multimedia: Videoencoding mit Flask (MPEG4, DivX)**
- **Netzwerkkomponenten (netpipe, netperf, ttcp)**
- **Speichersysteme (ECT-memperf, STREAM Benchmark)**
- **Graphikbeschleuniger**
- **ganze Softwaresysteme, Compiler, JVMs, Webserver usw.**

**... es sind keine Grenzen gesetzt ...**

- **Vorjahresprojekte:**  
**<http://www.cs.inf.ethz.ch/37-235/projects.html>**

# Weitere Benchmarks / Infos

- SPEC Bench: <http://www.specbench.org>
- Ziff-Davis Benchmarks Winstone, WinBench, 3D, CD, Battery, Web, Net Bench: <http://www.etestinglabs.com/benchmarks/>
- Imbench – Analysis Tools: <http://www.bitmover.com/Im/Imbench>
- ECT-memperf - Memory System Performance Characterization <http://www.cs.inf.ethz.ch/cops/ECT>
- STREAM Memory Bandwidth: <http://www.cs.virginia.edu/stream/>
- Netpipe: <http://www.scl.ameslab.gov/netpipe/>
- Netperf: <http://www.netperf.org/netperf/NetperfPage.html>
- SiSoft Sandra (System ANalyser, Diagnostic and Reporting Assistant), CPU, Memory, Drives, CD-Rom/DVD, Network: <http://www.sisoftware.demon.co.uk/sandra/>
- Weitere: <http://www.wintotaldb.de/yad/softw.php?urubrik=28>