

Medição do Desempenho

Benchmarks

Medição do Desempenho

IDEAL: obter um número que caracterize completamente o desempenho da máquina

IMPOSSÍVEL: complexidade dos sistemas
grande variedade de aplicações
grande variedade de condições de utilização

CUIDADO: os números fornecidos pelos vendedores podem ser enganadores.

- em que condições foram conseguidos?
- reflectem a carga real que o utilizador vai impor ao sistema?

Medição do Desempenho

Porquê medir o desempenho?

- escolha do *hardware* mais apropriado
- escolha do *software* (ex. compilador) mais apropriado
- escolha do algoritmo mais eficiente
- identificação de ineficiências num sistema (*hw*, *sw*)
- selecção de alternativas de desenho
 - qual o *instruction set* mais apropriado
 - selecção da organização do processador
 - selecção da organização da memória
 - ...

Definição de Desempenho

Avião	Capacidade	Alcance (milhas)	Velocidade (m.p.h)
Boeing 777	375	4630	610
Boeing 747	470	4150	610
Concorde	132	4000	1350

Qual o melhor avião?

Definição de Desempenho

Utilizador – minimização do **tempo de execução** (ou resposta)

Centro de Computação – maximização do **débito** (*throughput*)

$$Desempenho = \frac{1}{T_{exec}}$$

$$Desempenho \text{ relativo (ou ganho)} = \frac{Desempenho_A}{Desempenho_B}$$

Medição do Desempenho

PROBLEMA: Como garantir que a carga (*workload*) a que a máquina é sujeita durante a medição é a mesma carga a que o utilizador normalmente a sujeita?

São utilizados conjuntos de programas, designados por *benchmarks*, que se espera que reflectam a *workload* da maior parte dos utilizadores.

PROBLEMA: Como garantir que vendedores e fabricantes não manipulam os testes no sentido de melhorarem os resultados?

Os conjuntos de testes com maior credibilidade são definidos por comissões independentes, que especificam rigorosamente as condições em que estes devem ser realizados e a forma como os resultados devem ser documentados.

Tipos de *Benchmarks*

Synthetic Benchmarks – pequenos programas desenvolvidos especificamente para medir alguma característica específica da máquina. Normalmente não realizam nenhuma tarefa específica.

Desvantagens

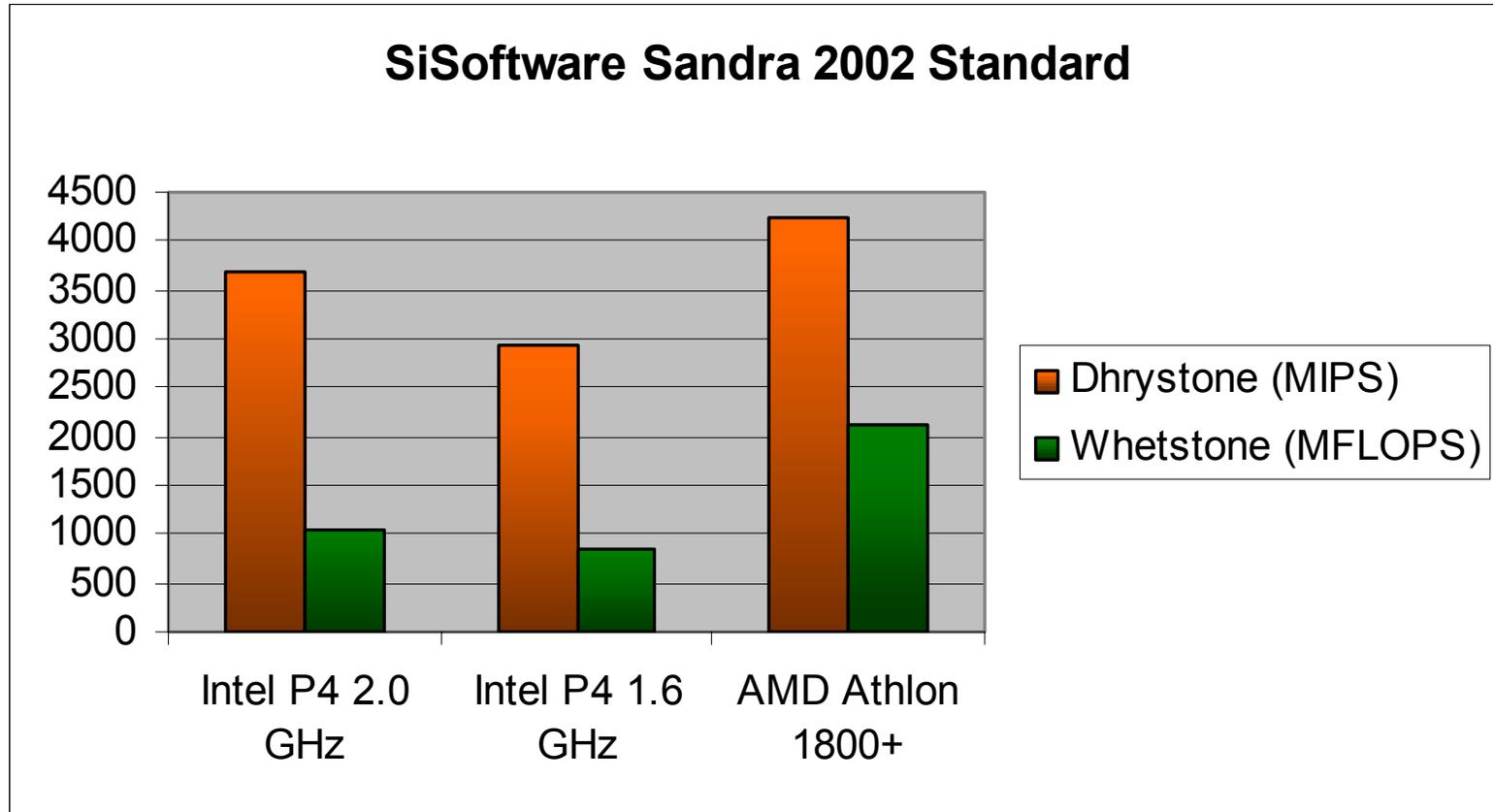
1. Não reflectem a carga que um utilizador aplica à sua máquina;
2. Programas pequenos que utilizam apenas a *cache*;
3. Alguns compiladores geram código otimizado para estes testes. Estas optimizações não podem depois ser usadas em aplicações reais.

Vantagens

1. Na fase inicial de desenho de um sistema estes testes são muito úteis, pois são fáceis de compilar e mesmo de simular.

Exemplos: Dhrystone (int) e Whetstone (FP)

Medição do Desempenho - Resultados



Tipos de *Benchmarks*

Conjuntos de aplicações reais – aplicações seleccionadas por representarem cargas típicas para um grande número de utilizadores

Desvantagens

1. Difícil seleccionar conjuntos de aplicações que garantidamente representem uma grande maioria de utilizadores;
2. Estes testes levam muito tempo a executar e as condições de teste e relatório de resultados são geralmente muito exigentes;

Vantagens

1. O utilizador pode geralmente confiar que os resultados reflectem com algum grau de precisão o desempenho a esperar da sua máquina;
2. Tratando-se de aplicações reais torna-se difícil aos fabricantes introduzirem características especiais no *hardware* ou nos compiladores para inflaccionar os resultados.

Exemplos: SPEC Benchmarks

SPEC - Standard Performance Evaluation Corporation

A SPEC (<http://www.spec.org/>) é uma associação criada em 1989 por um grupo de companhias para normalizar:

- o conjunto de testes a que as máquinas devem ser submetidas;
- as condições em que estes testes devem ser realizados;
- a forma como os resultados devem ser documentados

Benchmark	Descrição
SPEC WEB'99	Servidores WEB
SPEC HPC'2002	High Performance Computing
SPEC JVM'98	Java Virtual Machine
SPEC SFS'97_R1	Sistem File Server
SPEC MAIL'2001	Servidores de eMail
SPEC CPU'2000	Processador – memória - compilador

SPEC CPU'2000

Conjunto de programas cuidadosamente seleccionados para representarem a carga que um utilizador “regular” impõem à sua máquina.

Inclui vários testes dos quais se destacam:

Benchmark	Descrição
SPECint_base2000	Programas com operações em inteiros, compilados sem optimizações.
SPECint2000	Programas com operações em inteiros, compilados com optimizações.
SPECfp_base2000	Programas com operações em vírgula flutuante, compilados sem optimizações.
SPECfp2000	Programas com operações em vírgula flutuante, compilados com optimizações.

SPECint2000

Teste	Linguagem	Descrição
164.gzip	C	Compression
175.vpr	C	FPGA Circuit Placement and Routing
176.gcc	C	C Programming Language Compiler
181.mcf	C	Combinatorial Optimization
186.crafty	C	Game Playing: Chess
197.parser	C	Word Processing
252.eon	C++	Computer Visualization
253.perlbnk	C	PERL Programming Language
254.gap	C	Group Theory, Interpreter
255.vortex	C	Object-oriented Database
256.bzip2	C	Compression
300.twolf	C	Place and Route Simulator

12 programas

SPECfp2000

Teste	Linguagem	Descrição
168.wupwise	F77	Physics / Quantum Chromodynamics
171.swim	F77	Shallow Water Modeling
172.mgrid	F77	Multi-grid Solver: 3D Potential Field
173.aplu	F77	Parabolic / Elliptic Partial Differential Equations
177.mesa	C	3-D Graphics Library
178.galgel	F90	Computational Fluid Dynamics
179.art	C	Image Recognition / Neural Networks
183.quake	C	Seismic Wave Propagation Simulation
187.facerec	F90	Image Processing: Face Recognition
188.ammp	C	Computational Chemistry
189.lucas	F90	Number Theory / Primality Testing
191.fma3d	F90	Finite-element Crash Simulation
200.sixtrack	F77	High Energy Nuclear Physics Accelerator Design
301.apsi	F77	Meteorology: Pollutant Distribution

14 programas

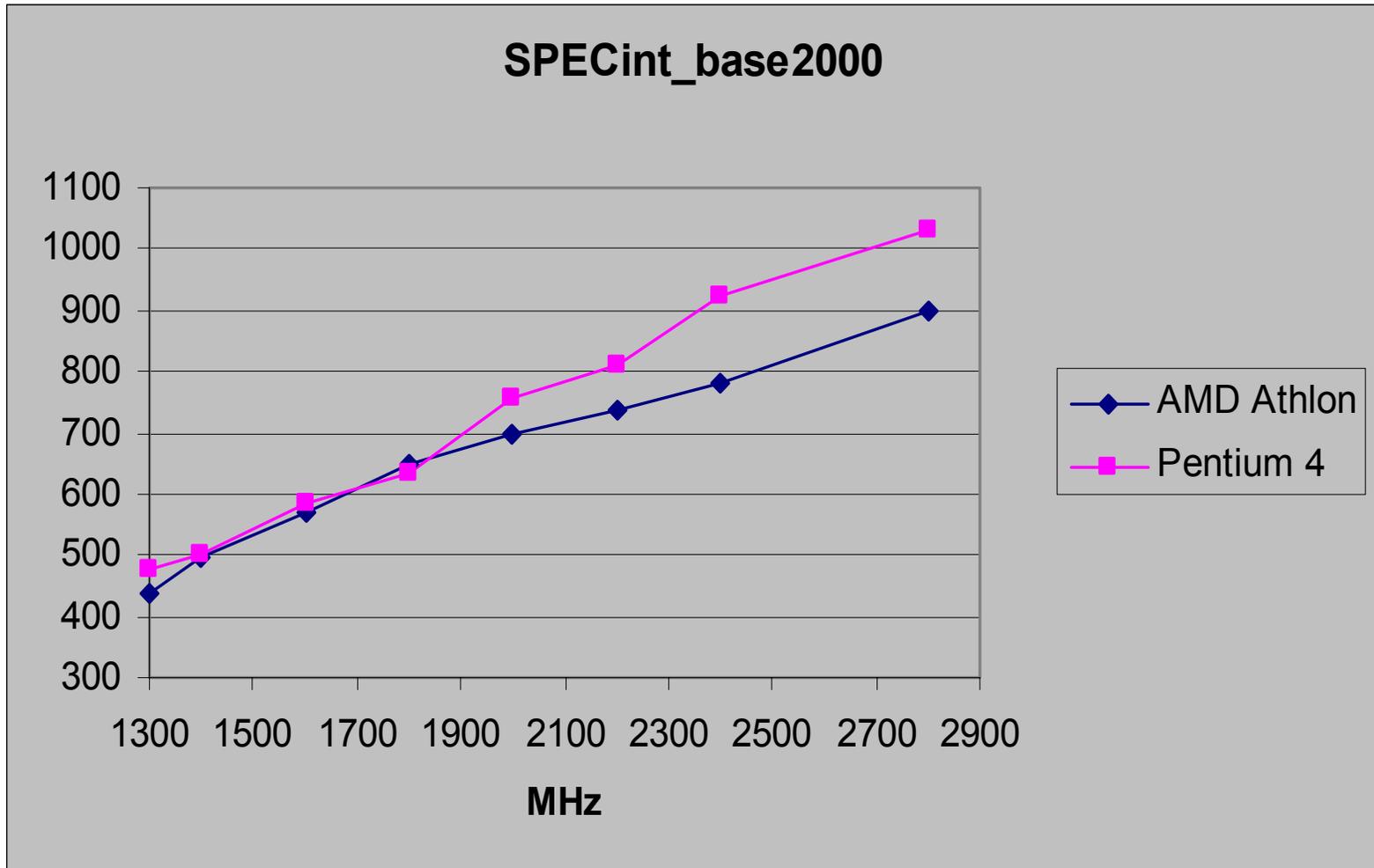
SPEC CPU2000

Como calcular os índices SPECint2000 e SPECfp2000?

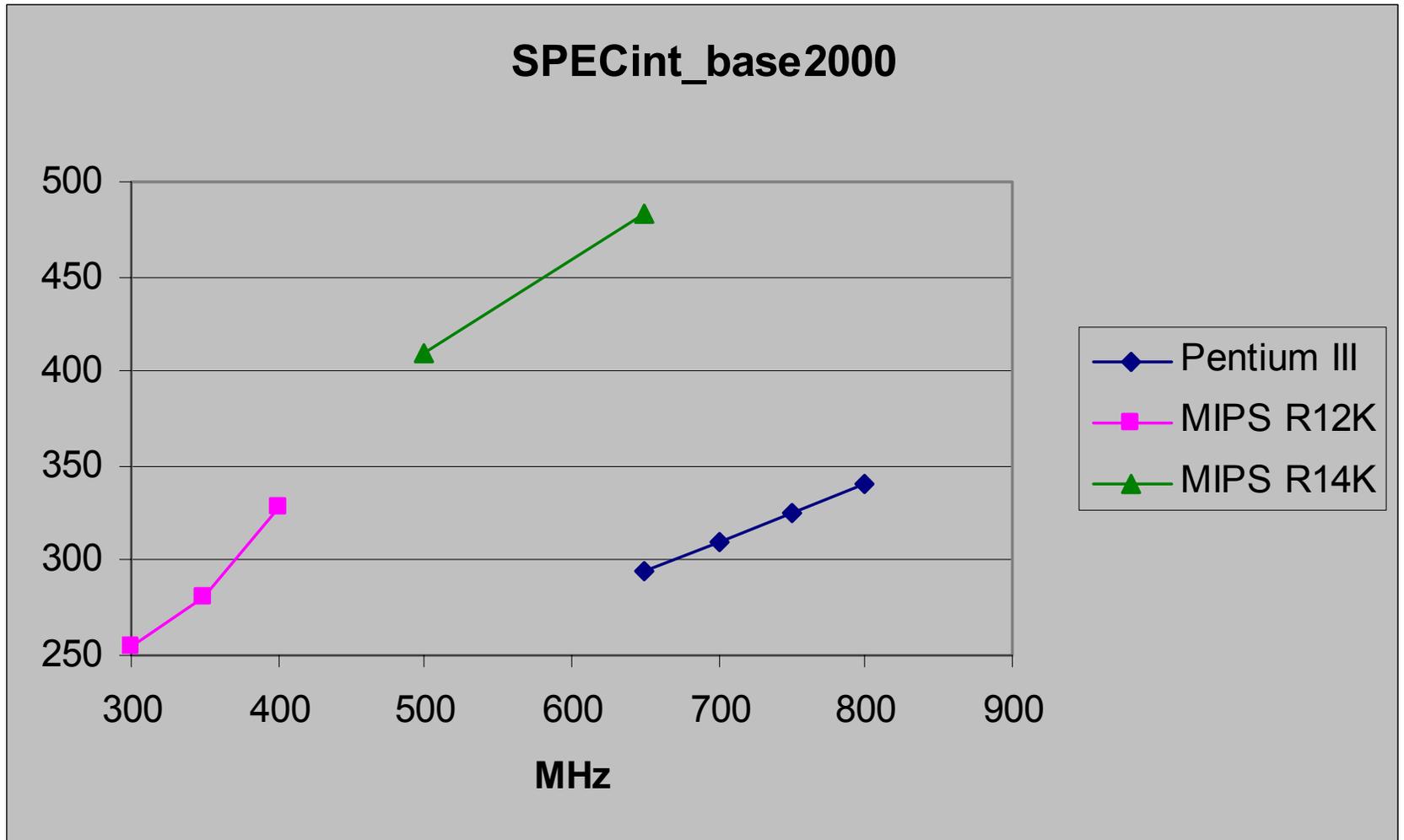
1. O tempo de execução de cada teste numa máquina de referência (SUN Ultra 5_10) é dividido pelo tempo de execução na máquina a testar. Chama-se a esta razão o **SPEC ratio**.
2. É calculada a média geométrica dos vários SPEC ratios.

$$SPEC = \sqrt[n]{\prod_i^n SPECratio_i}$$

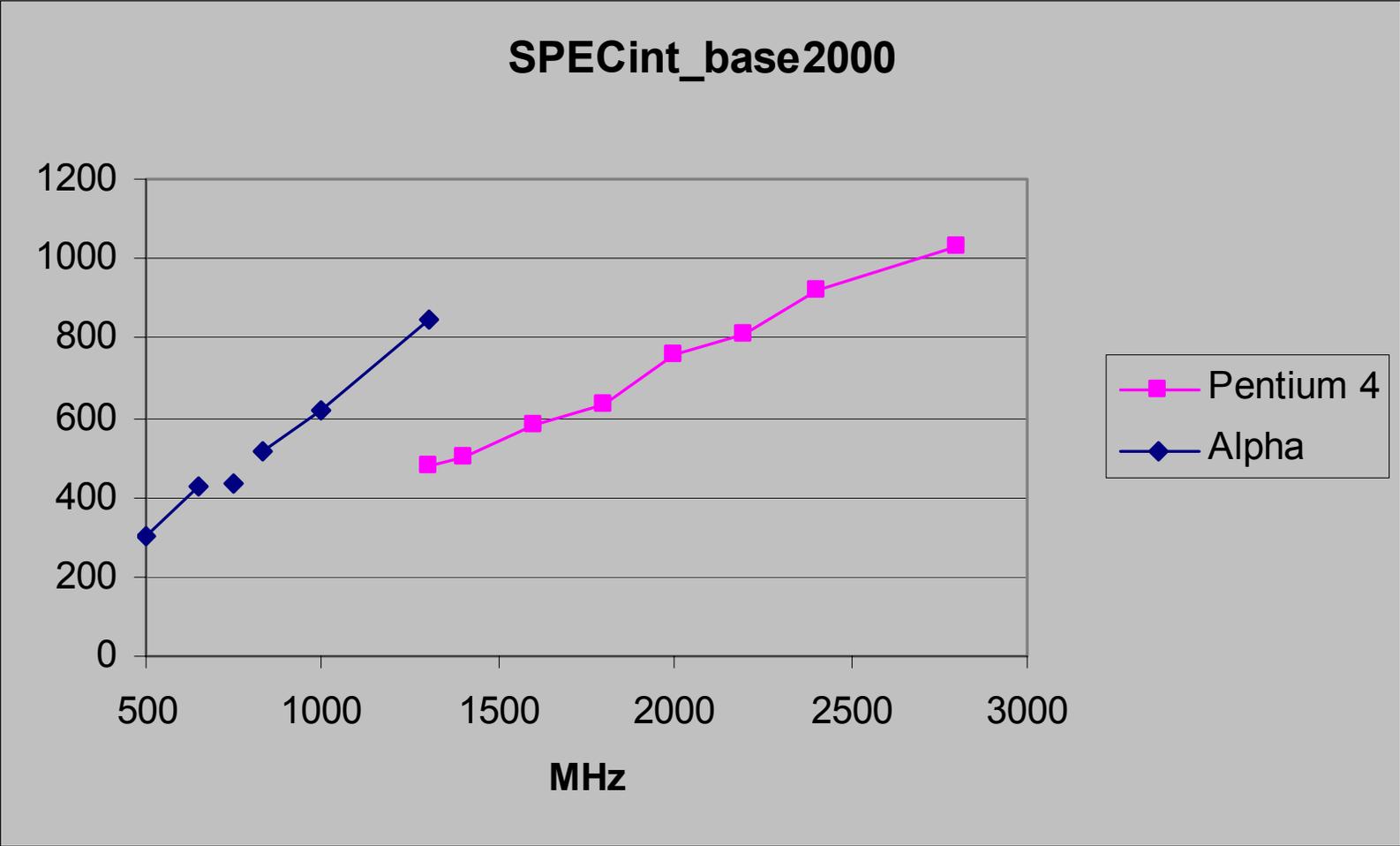
SPECint_base2000



SPECint_base2000



SPECint_base2000



Sumário

Tema	H & P
Benchmarks	Sec. 2.4
SPEC	Sec 2.6
Média Geométrica	Sec 2.7
Synthetic Benchmarks	Sec 2.7

Para mais informação ver: <http://www.spec.org/>