
AMD 3DNow!™ Technology and the K6-2 Microprocessor

Stuart Oberman
Fred Weber
Norbert Juffa
Greg Favor

California Microprocessor Division
Advanced Micro Devices



OUTLINE

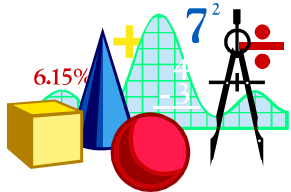
- **Motivation for 3DNow! Technology**
- **Features of the 3DNow! instruction set**
- **AMD-K6-2 implementation**
- **3D graphics performance**
- **Future**



Acceleration of Multimedia Applications

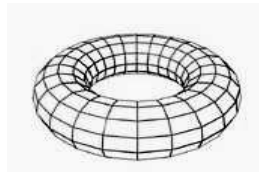
- Multimedia applications have become an integral part of the PC platform
 - Multimedia algorithms are computationally intensive

**Application or
Game Physics**



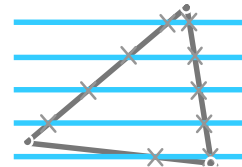
**floating point
intensive**

**Geometry transform,
Clipping, & Lighting**



**floating point
intensive**

**Triangle
Set up**



**floating point &
integer intensive**

**Pixel
Rendering**



**integer
intensive**

Motivation for 3DNow! Technology

- **Why a new technology?**

- Previous focus has been on integer intensive pixel rendering tasks: MMX and 3D graphics hardware
- 3D graphics performance now limited by floating-point intensive front-end of graphics pipeline
- New applications require realistic physical modeling

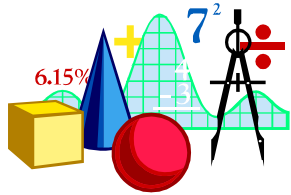
- **What is it?**

- New set of instructions to accelerate FP computation
- Defined in collaboration with leading ISV's
- Maximizes the performance of graphics accelerator cards by improving the front-end of graphics pipeline



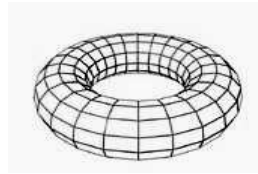
3DNow! Technology

Application or
Game Physics



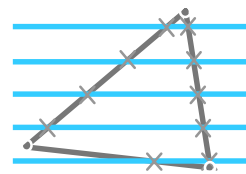
floating point
intensive

Geometry transform,
Clipping, & Lighting



floating point
intensive

Triangle
Set up



floating point &
integer intensive

Pixel
Rendering



integer
intensive

3DNow! Technology Accelerates

Graphic Cards Accelerates

• Benefits

- Accelerates most floating-point intensive multimedia operations
 - Graphics pipeline (physics, geometry, setup)
 - Audio processing

3DNow! Technology

- **21 new instructions**
 - “LEAN and MEAN” design philosophy
 - Includes only performance critical features
- **SIMD floating-point instructions**
 - Compatible with IEEE single precision data type
 - Two 32-bit FP values per 64-bit reg/mem operand
 - Uses MMX registers -> avoids x87 register stack
 - No exceptions
 - Limited rounding modes
 - No switching overhead between 3DNow! and MMX
 - Peak throughput of 4 FLOPS per cycle
- **No core OS support required**



Classes of Instructions - FP

- **Basic arithmetic**
 - PFADD, PFSUB, PFSUBR, PFACC, PFMUL
- **Comparisons**
 - PFCMPEQ, PFCMPGT, PFCMPGE
- **Min/max**
 - PFMIN, PFMAX
- **Conversions**
 - PF2ID, PI2FD
- **Reciprocal and reciprocal square root**
 - PFRCP, PFRSQRT
 - PFRCPIT1, PFRSQIT1, PFRCPIT2

Classes of Instructions - Non FP

- **Integer**
 - PAVGUSB, PMULHRW
- **Data movement**
 - PREFETCH/PREFETCHW
- **Overhead reduction**
 - FEMMS

Reciprocal and Reciprocal Square Root

- Alternative to “classical” DIV and SQRT
 - Reciprocal and reciprocal square root frequently used in graphics applications
 - Higher performance through reuse of common divisors and radicands
- Choice of reduced (14-15b) or full precision
 - Reduced precision sufficient for many applications and is higher performance
 - Avoid all long latency operations; full precision synthesized from fully-pipelined Newton-Raphson iterations ops

Reciprocal Iteration Instructions

- Reciprocal Newton-Raphson iteration
 - To compute full-precision reciprocal of b using initial approximation R_0 :
 - $R_{full} = R_0 \times (2 - b \times R_0)$
 - R_0 is accurate to about 14 bits, b is a 24 bit number
 - PFRCPIT1 performs $b \times R_0$ rounded to 32 bits, inverts the result (one's complement), and compresses out leading 8 bits known to be identical, leaving 24 bits
 - PFRCPIT2 expands the previous result to 32 bits, multiplies by R_0 , adds a fixed bias, and rounds to 24 bits

Reciprocal Accuracy

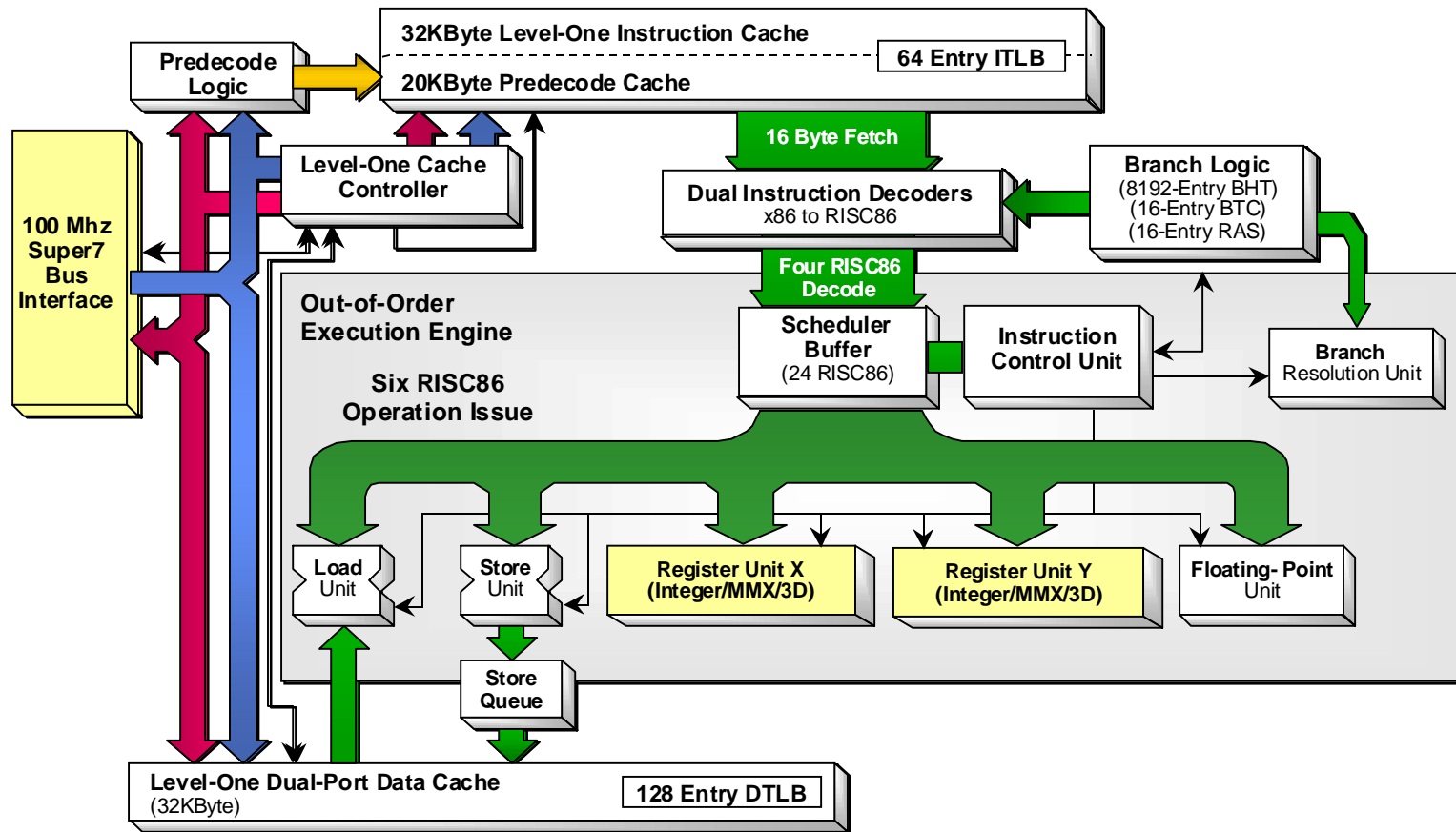
- Fast approximations
 - PFRCP accurate to 14.9 bits
 - PFRSQRT accurate to 15.8 bits
- Full precision
 - PFRCP, PFRCPIT1, PFRCPIT2 sequence provides IEEE RN result for > 99% of all operands; remaining differ by 1 unit-in-the-last-place
 - PFRSQRT, PFMUL, PFRSQIT1, PFRCPIT2 sequence provides IEEE RN result for > 87% of all operands; remaining differ by 1 unit-in-the-last-place

AMD-K6-2 Microprocessor

- Worldwide launch May 28, 1998
- Implemented in 0.25um CMOS process
- 9.3M transistors on a die of 80 mm²
- New features of AMD-K6-2
 - **Superscalar 3DNow! and MMX units**
 - Dual decode and dual execution pipelines
 - No decode pairing restrictions
 - Only one cycle misalignment penalty on memory accesses
 - **100 MHz Front Side bus**
 - Increases local bus and L2 cache bandwidth by 50%
 - Redesigned I/O timing to allow for low cost 100 MHz motherboard

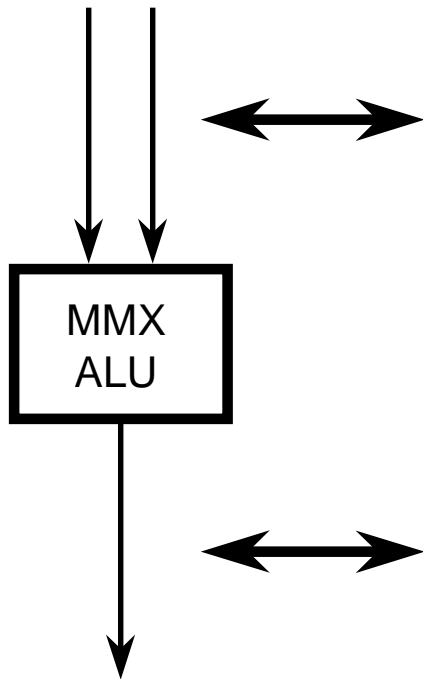


AMD-K6-2 Block Diagram

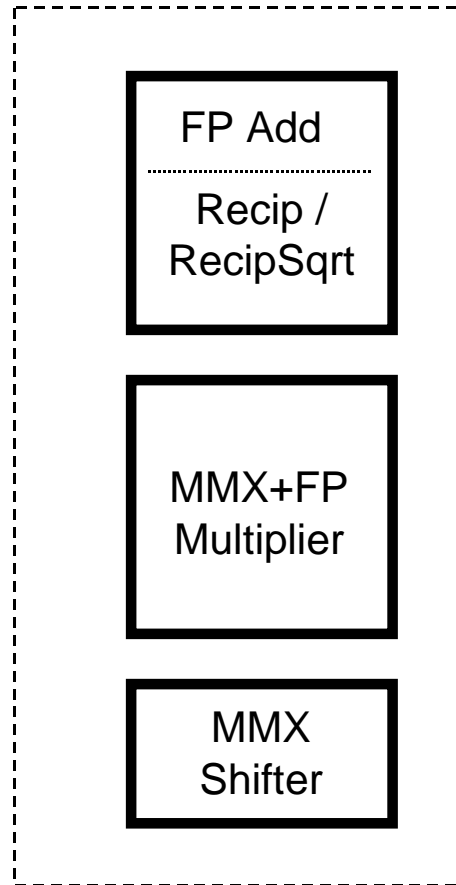


AMD-K6-2 Multimedia Units

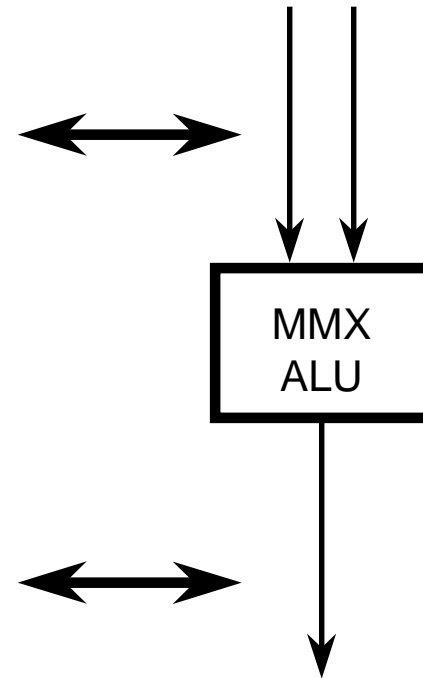
Register X Execution Pipeline



Shared Resources



Register Y Execution Pipeline



AMD-K6-2 Multimedia Performance

Instruction Type	Latency (cycles)	Throughput (cycles)
3DNow! FP	2	1
3DNow! / MMX integer ALU	1	1
MMX multiply	2	1

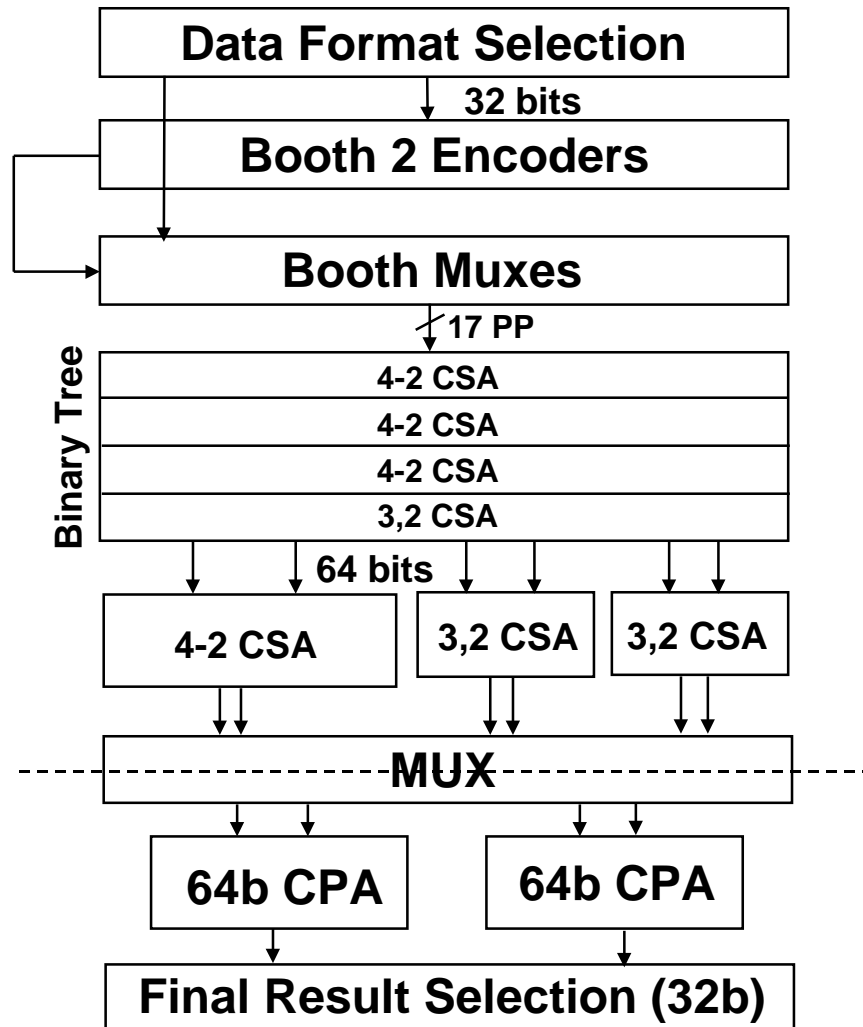


Recip / Recip Sqrt Performance

	K6-2 Performance	PII Performance
14 bit reciprocal	2 cycles pipelined	-
15 bit reciprocal square root	2 cycles pipelined	-
24 bit reciprocal	6 cycles pipelined	~ 17 cycles non-pipelined
24 bit reciprocal square root	8 cycles pipelined	~ 28 cycles non-pipelined

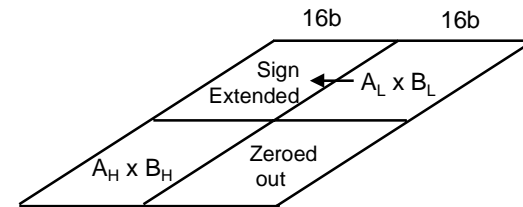
Shared 3DNow! and MMX Multiplier

EX1



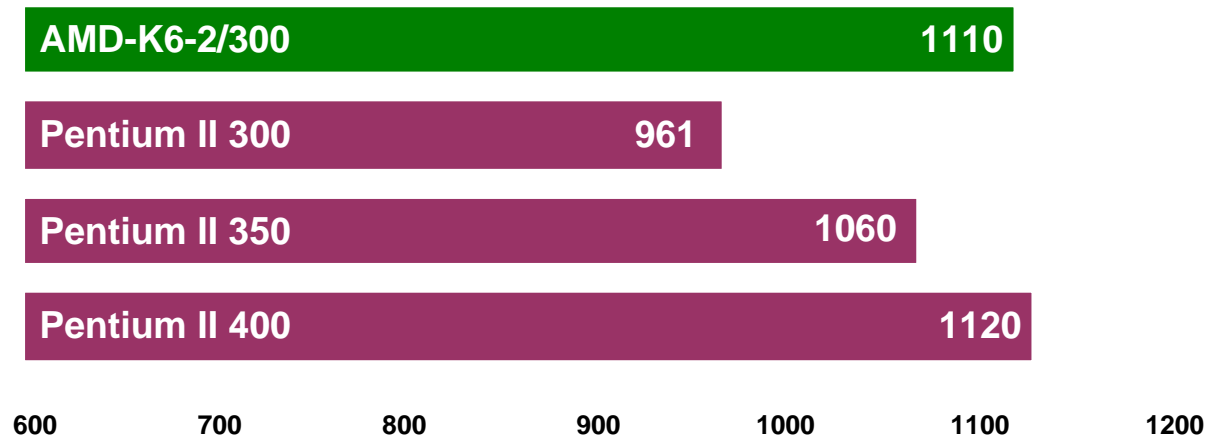
EX2

MMX Data Alignment



3D Winbench 98 Performance

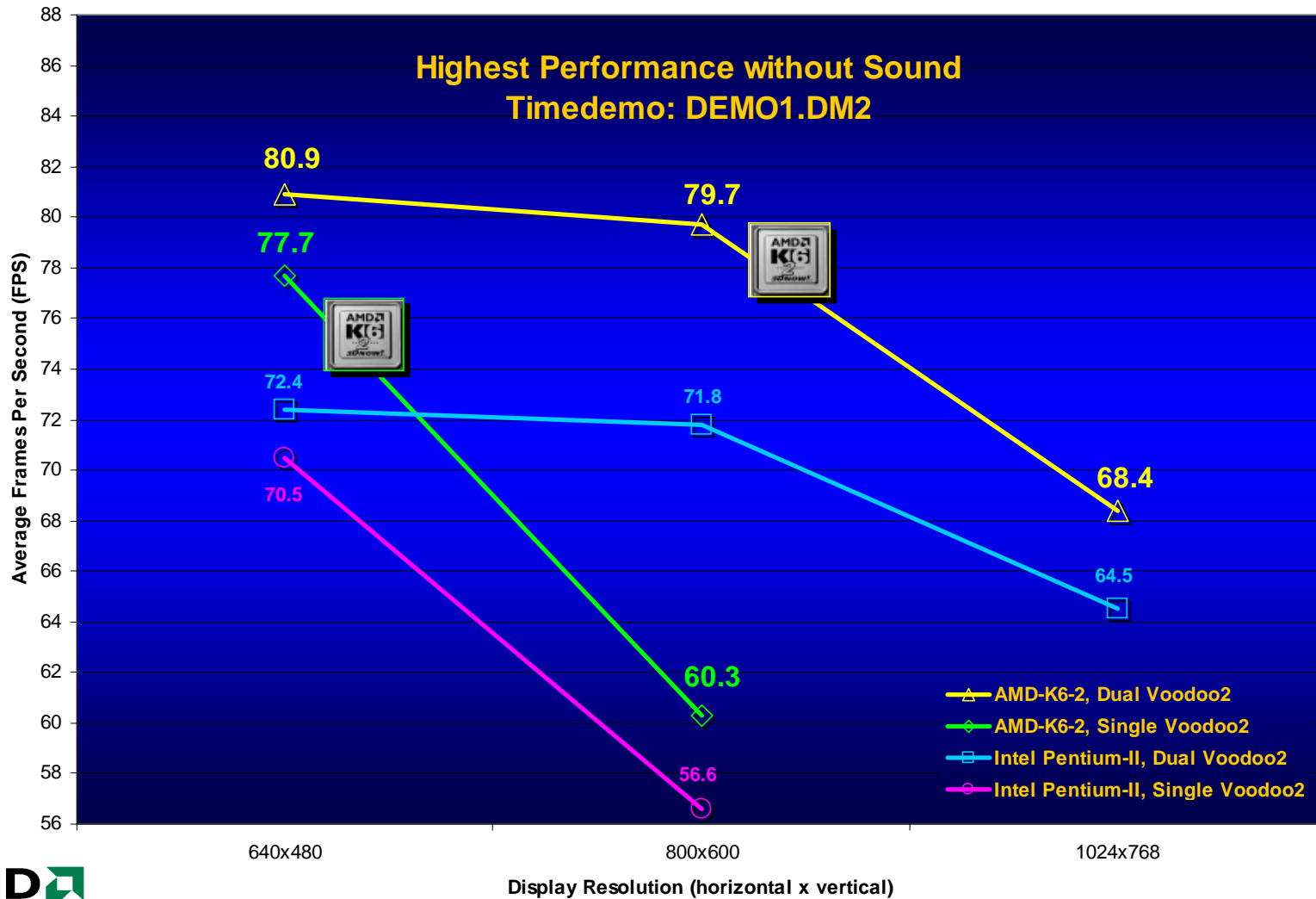
*3D Winbench 98 / Windows 95
(DirectX 6.0 optimized for 3DNow!™)*



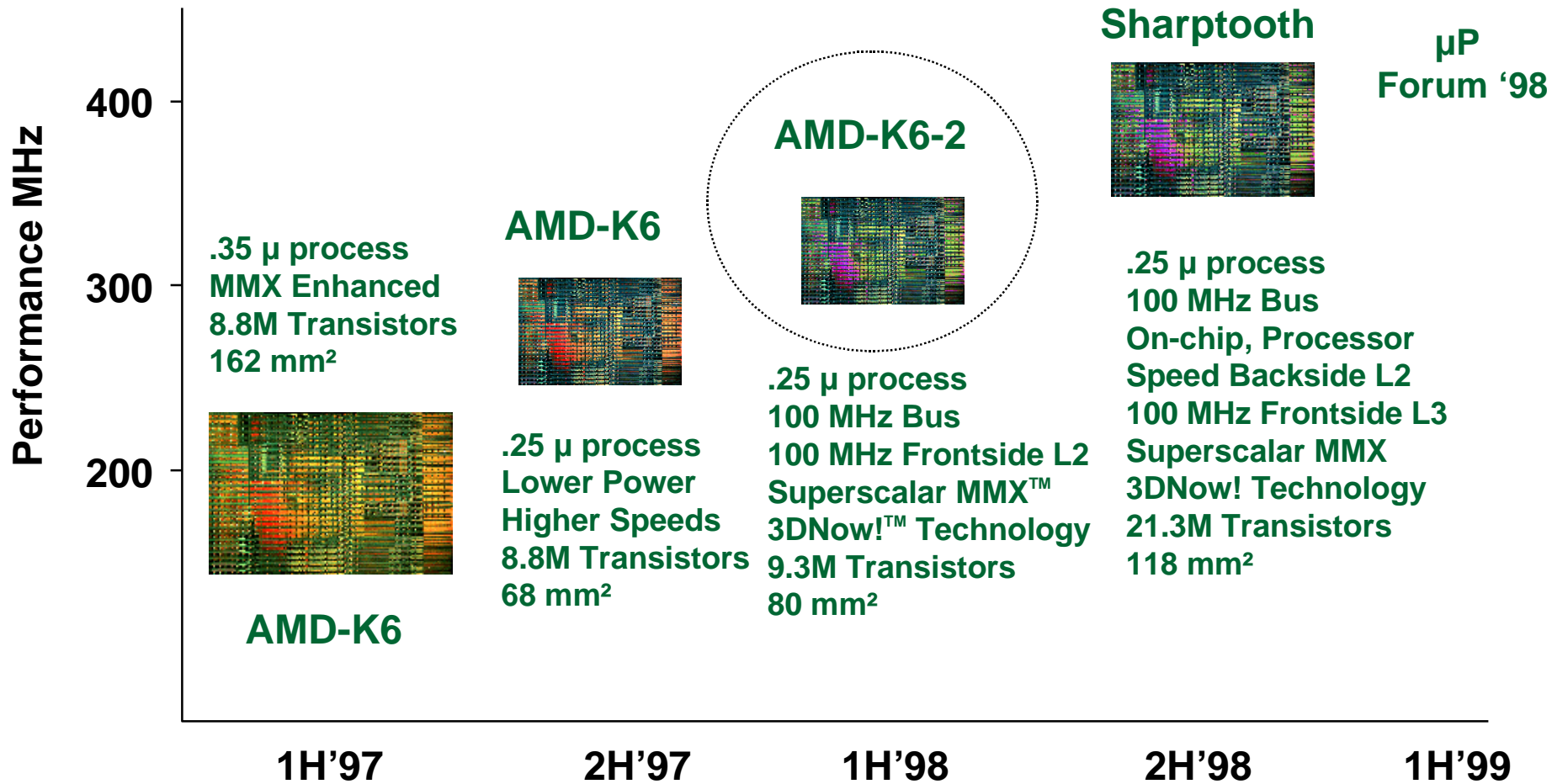
Windows 95 OSR 2.1, 32MB DRAM, Maxtor DiamondMax IDE HD, 512K L2 cache, Diamond Viper V330 4MB AGP. AMD-K6 3D processor based system: Microstar 5169 mainboard supporting 100MHz bus. Pentium® II 300 based system: Abit LX6 mainboard supporting 66MHz bus (Pentium II 300 based systems with 100MHz bus not currently commercially available). Pentium II 350 and Pentium II 400 based systems: Asus P2B mainboard supporting 100MHz bus



Quake 2 Performance



K6 Family and Future



μP Forum '98

