X Window System Benchmarks

Nelson H. F. Beebe
Center for Scientific Computing
Department of Mathematics
University of Utah
Salt Lake City, UT 84112
USA
Tel: +1 801 581 5254
FAX: +1 801 581 4148
E-mail: `<beebe@math.utah.edu>`

21 June 1993
Version 1.08

1 Introduction

This report presents five benchmarks of workstation performance under the X Window System. Because a window system makes heavy use of integer, bitfield, and other graphical operations, and on some systems, receives graphics accelerator hardware assists, it should be expected that performance measurements may show quite different characteristics than would be seen with I/O benchmarks, or compute-intensive benchmarks like LINPACK, matrix multiplication, or SPEC. *Readers are therefore cautioned that the benchmark results presented here reflect only part of the performance characteristics that may be of interest in choosing a workstation.*

It must be recognized that the X Window System is a *generic* system developed by MIT's Project Athena in collaboration with Digital Equipment Corporation and IBM Corporation. The source code (over 6.8 million lines in X11R5) is publicly, and freely, available. That, together with the fact that it was the first non-proprietary network windowing system, has made it a *de facto* standard in the industry, and implementations can be found for virtually all UNIX systems on the market today, from low-cost workstations, up to the largest supercomputers. Public and commercial server implementations are also available for the IBM PC and Macintosh personal computers.

However, depending on the energy and commitment of the programmers who port the X Window System to a new environment, the quality of implementation should be expected to vary considerably. A great deal of attention needs to be given to the efficiency of low-level graphics primitives if high performance is to be achieved. I shall leave it to the reader to judge to what extent the systems studied here have received that attention.

It must also be noted that in the case of several vendors, including Evans and Sutherland, Silicon Graphics, and Stardent, the X Window System augments native high-performance graphics hardware and software support which the vendor has put much more development effort into. *This means that a system that performs poorly under the X Window System may still be a superb graphics system.*

Each of the systems studied here has its own particular strengths in hardware and software, and it is likely that specific benchmarks can be exhibited for each that place it above the others. If you cannot find such a benchmark, just ask your vendor's local sales person. For those users who tax the capabilities of a workstation's windowing environment for software development or graphics display, the window system performance results given here may be of considerable interest.

I hope that vendors will also find the comparisons here informative, and that the results will assist in revealing weak points of their systems. Computer architecture design is a very complex activity that is full of compromises and tradeoffs necessary to produce a marketable product. Only by actual use of the finished system under a wide variety of conditions is it possible to verify whether the design is satisfactory. Unlike floating-point benchmarks, which tend to be closely tied to CPU clock rates and instruction set characteristics, the results here show surprisingly wide variations that are unlikely to be humanly predictable, and can be exposed only by actual measurements.

2 Benchmark systems

The *HP Apollo 9000 Series 700 Performance Brief* contains graphics benchmark figures using the X Window System `x11perf` program for HP 720 (50 MHz PA-RISC) and 730 (66.7 MHz PA-RISC), DECstation 5000 (25 MHz MIPS R3000), and Sun SPARCstation 2 (40 MHz SPARC) systems. Detailed results for the DECstation 5000 can be found in an article in *Digital Review* magazine, 7 May 1990, p. 41.

The `x11perf` program was developed by Digital Equipment Corporation as part of the MIT Project Athena effort, carried out jointly with MIT and IBM Corporation. The program is about 5100 lines of C code in 20 files, and was originally used to tune the performance of the DECstation 3100. The authors of `x11perf` emphasize that *it is not possible to collapse the benchmark results into a single number*; instead, users are advised to select a subset of the tests that most closely resembles the operations required for their own work, and then to use those test subsets in making collective comparisons between different workstations.

what the numbers mean

`x11perf` results are reported in *operations per second*; that is, how many vectors of a given length, how many 80-character lines of text characters, how many dots, and so on, can be drawn in one second.

I have augmented the numerical results from the Hewlett-Packard report with data collected for some local systems of interest. These additional machines are:

- DEC Alpha AXP 3000/500 (150 MHz Alpha), OSF/1 T1.2-2 (Rev. 5), X11R5
- DEC Alpha AXP 3000/500X (200 MHz Alpha), OSF/1.2 (Rev. 10), X11R5
- DECstation 3100 (16.67 MHz MIPS R2000), ULTRIX V4.1 (Rev. 55), UWS V4.1 (Rev. 198), X11R4 (vendor delivers R3; upgraded locally to R4).
- Evans & Sutherland V50 (25 MHz MIPS R3000), BSD 4.3.
- Hewlett-Packard 9000/735 (99 MHz PA-RISC 1.1), HP-UX A.09.01, X11R5

- IBM RS/6000-320 (20 MHz RIOS), AIX 3.1 with IBM-supplied X11R3, AIX 3.1 with MIT X11R5, AIX 3.2 with MIT X11R5
- IBM RS/6000-520 (20 MHz RIOS) with High Performance 3D Color Graphics Processor, Color Graphics Video Card, 24-bit Color Graphics frame buffer card, and 24-bit Z-buffer option card, AIX 3.1.5, X11R3.
- Silicon Graphics 4D/20 (12 MHz MIPS R2000), GR1.1 Z-buffer option.
- Silicon Graphics 4D/240GTX (four 25 MHz MIPS R2000), GT option.
- Silicon Graphics 4D/320VGX (two 33 MHz MIPS R2000), VGX option.
- Silicon Graphics Indigo (33 MHz MIPS R3000), IRIX 4.0.1.
- Silicon Graphics Crimson (50 MHz MIPS R4000), VGX graphics accelerator option, IRIX 4.0.2.
- Stardent 1520 (dual-processor 16 MHz MIPS R2000), Stardent UNIX Release 2.2 with X11R4 libraries, but X11R3 server.
- Sun 3/110 (16.67 MHz Motorola 68020), SunOS 4.1.1, X11R4.
- Sun SPARCstation 2 (40 MHz SPARC CPU), SunOS 4.1.2, with X11R4, X11R5, OpenWindows 2.0, and OpenWindows 3.0.
- Sun SPARCstation LX (50 MHz MicroSPARC CPU and GX-Plus graphics chip), Solaris 2.1, with X11R5, OpenWindows 3.1
- Sun SPARCstation 10/41 (40 MHz SuperSPARC CPU), Solaris 2.1, with X11R5, OpenWindows 3.1

Note particularly that the DECstation 3100, Sun 3/110, HP systems run X11R4, while most of the others are at X11R3. Significant performance improvements were made in X11R4, at least for Sun color systems, and the results here for the Sun SPARCstation 2 show that X11R5 is even better.

X11R4 has been available from MIT since January 1990, and X11R5 was released in late August 1991. Vendor implementations lag far behind the MIT releases. DEC and HP first offered X11R4 in mid-1991, and IBM only released it in May 1992, nine months after R5 became publicly available. This tardiness is regrettable, because it compromises software portability, and deprives customers of sometimes substantial performance gains. I find it hard to understand why at least DEC and IBM are unable to offer prompt releases of the X Window System when they represent two-thirds of the development team! At least the new (1993) DEC Alpha systems are delivered with X11R5.

To provide a firm base for comparison, performance relative to my Sun 3/110 color workstation is given in each of the charts in this report. This system was retired in June 1992.

The Sun 3/110 was introduced about the beginning of 1987, and is the only non-RISC system included here. It has the same processor as mid-range models of the Apple Macintosh.

The RISC systems are all quite recent: Stardent 1520 and DECstations (1989), IBM RS/6000 (February 1990), Sun SPARCstation 2 (summer 1990), Sun SPARCstation LX (November 1992), Sun SPARCstation 10/41 (March 1993) HP Series 700 (March 1991), DEC Alpha (December 1992). I do not have dates of introduction of the first SGI systems; all are shown on a July 1990 price list. The Indigo was introduced in the summer of 1991, and the Crimson in January 1992.

List prices of these systems with 24MB of memory, 500MB of disk, and at least 8 color planes, are in the range of $10K to $30K, except for the Stardent 1520 (about $80K when it first appeared), the DEC Alpha 3000/500 (about $46K in 1993), the DEC Alpha 3000/500X ($70K), and the HP 9000/735 ($40K).

The older Sun 3/110 was $17K in 1987. The SGI systems contain a lot of specialized graphics hardware. The systems tested list from $11K (4D/20) to $135K (4D/240GTX and 4D/320VGX).

In matrix multiplication and Gaussian elimination (LINPACK) benchmarks, the Sun 3/110 performs at approximately 0.15 MFlop (floating-point) and 0.41 MFlop (32-bit integer). Most of the RISC systems included here have benchmark figures in the range of 2 to 32 MFlops. The DEC Alpha 3000/500 reaches 106 Mflops in single precision, and 89 Mflops in double precision, and the 3000/500X about 118 Mflops in each. The HP 9000/735 reaches 149 Mflops in each. Thus, one might expect their graphics performance to be five or more times larger than that of the Sun 3/110. *This turns out not to be the case for some systems*. In fact, on some benchmarks, the Sun 3/110 is up to 8 times faster than a RISC system.

3 Analysis of the benchmark results

The standard X11R3 and X11R4 distributions from MIT contain the `x11perf` benchmark program; the latest version, 1.2, was used here. It requires no toolkit libraries, only low-level X library routines, so it can be readily implemented on almost any system with a vendor-supplied X Window System implementation.

`x11perf` can run 146 different graphics performance benchmarks. Hewlett-Packard chose 5 of these for representative comparisons, following the practice in *Digital Review* magazine. These include single dots, short vectors, small filled rectangles, text with a variable-width font, and subwindow mapping, all operations which are expected to be typical of most X Window System applications.

The benchmarks are run on standalone systems whenever possible, with no other active jobs, or other windows that require updates (e.g. clocks and performance monitors). `x11perf` maps the cursor out of the benchmark window, because it otherwise seriously impacts performance. Any mouse movement would generate input interrupts to process, and all drawing operations with the cursor present must redraw the cursor in case it was overwritten. Some high-performance systems map the cursor to separate graphics memory planes, but current X implementations do not take advantage of that.

The results illustrated here in bar chart form in Figures 1–5 are rather surprising.

HP-9000 Series 700

The most obvious point is the exceedingly high performance of the new HP-9000 Series 700 machines, far beyond any of the competition. The differences are large enough that tripling or quadrupling the clock rates of the systems with slower CPUs will still *not* surpass the Series 700 workstations. Some of the benchmark results are 200 times faster than those for IBM RS/6000 models of similar price.

DECstation 5000

A second observation is that the relative performance of these machines changes drastically between the benchmarks. The DECstation 5000 graphics accelerator gives performance close to that of the HP systems only for lines and filled rectangles.

Since the `-line10` 10-pixel benchmark draws ten times as many pixels as the `-dot` benchmark, one might expect its operation count to be about ten times smaller. For the Sun 3/110, the ratio of the two is 9.17, but for most of the others, it is much smaller:

Machine	`-dot/-line10`
DEC Alpha 3000/500	1.60
DEC 3100	4.22
DEC 5000 PXGT	0.81
E&S V50	3.19
HP 720 GRX/CRX	1.44
HP 730 GRX/CRX	1.79
HP 735 GRX/CRX	2.87
IBM RS/6000-320	1.17
IBM RS/6000-520	1.14
SGI 4D/20	2.56
SGI 4D/240GTX	6.80
SGI 4D/320VGX	9.54
SGI Indigo	2.79
SGI Crimson VGX	2.42
Stardent 1520	0.37
Sun 3/110	9.17
Sun SPARCstation 2	6.87
Sun SPARCstation 2 GX (X11R4)	3.70
Sun SPARCstation 2 GX (X11R5)	1.75
Sun SPARCstation 2 GX (OW2)	0.86
Sun SPARCstation LX GX+ (OW3.1)	1.43
Sun SPARCstation 10/41 (OW3.1)	1.98

This means that the higher-performance systems here have optimized drawing of vectors compared to dots, or that the overhead in setting up to draw a single dot is very large compared to that for drawing one dot of many in a vector. It also suggests that the Sun systems have substandard vector drawing times, and could possibly be subjected to substantial optimizations.

DEC Alpha 3000/500

The DEC Alpha 3000/500 is comparable to the HP systems on three of the benchmarks, but substantially slower for dots and lines.

E&S V50

The E&S V50 is a very high performance graphics system in its native graphics mode. As with most of the other systems tested here, its X Window System performance is disappointing.

IBM RS/6000-320

The IBM RS/6000-320 outpaces the HP 9000-720 by about 50% in my floating-point matrix multiplication benchmarks, but has dismal graphics performance relative to the HP systems. In the subwindow mapping benchmark, its performance is only a small fraction of the Sun 3/110, and this seems to be independent of which version of X Window System is used. The IBM RS/6000-520 with hardware graphics assist is about the same performance as the model 320, except that window creation speed is doubled.

IBM RS/6000-520

SGI

The SGI systems show performance trends that match their price and date of introduction. There are no uniform trends with respect to products from the other vendors. In Silicon Graphics native GL mode, these systems are claimed to produce 90K (4D/20), 230K (Indigo), 400K (4D/240GTX), and 1M (4D/320VGX and Crimson) 10-pixel vectors/second. These figures are respectively 8.6, 1.3, 18.1, 45.4, and 8.5 times faster than the observed rates in the X Window System.

Sun SPARCstation 2

When version 1.00 of this report appeared in June 1990, the Sun SPARCstation 2 was at the top end of Sun's product line. While it is 1.8 to 25 times faster than the older Sun 3/110 on these benchmarks, its performance is definitely not competitive with comparably-priced systems from other vendors. The GX model has a graphics accelerator, and improves performance by a factor of two to four. There are sharp differences between software releases; results are presented for OpenWindows 2.0 and 3.0 (based on X11R4), X11R4, and X11R5.

Sun SPARCstation 10

In May 1992, the Sun SPARCstation 10 was announced, but shipments of the model 20 and 30 did not happen until the fall of 1992, and the model 41 was delayed until March 1993. The clock speeds are 33 MHz (model 10/30), 40 MHz (model 10/41), and 45 MHz (models 10/52 and 10/54). These are comparable to the 40MHz SPARCstation 2 clock speed; the extra performance comes from the superscalar implementation of the SPARC architecture. The 10/52 has two CPUs, and the 10/54, four CPUs. All but the model 10/30 have 1MB external cache for each CPU. Sun benchmark data suggests that these new systems may offer up to 3 times the performance of the SPARCstation 2.

Stardent 1520

The Stardent 1520 is billed as a parallel and vector graphics minisupercomputer, and indeed, its performance in native graphics mode is very impressive. However, graphics performance in X Windows is unacceptably slow, in 3 out of 5 cases far below the Sun 3/110.

According to Stardent technical specifications, this system will fill 20 $\times$ 20 rectangles at 60M pixels/sec, draw non-anti-aliased vectors at 11.6M pixels/sec, and anti-aliased vectors at 0.5M pixels/sec.

The `x11perf -rect10` results correspond to fill rates of about 334,000 pixels/sec, or about 180 times slower than in native mode. The `x11perf -line10` results are equivalent to 565,000 pixels/sec, or about 20 times slower than native mode.

We cannot make a similar comparison for anti-aliased vectors because the X Window System makes no provision for drawing them; that operation must be synthesized from line and point drawing operations, and intensity changes, at an enormous performance penalty.

It would be informative to make similar comparisons for the HP 9000 Series 700 machines, which also have a native graphics mode, called Starbase, but I do not yet have the necessary technical specification details to do so.

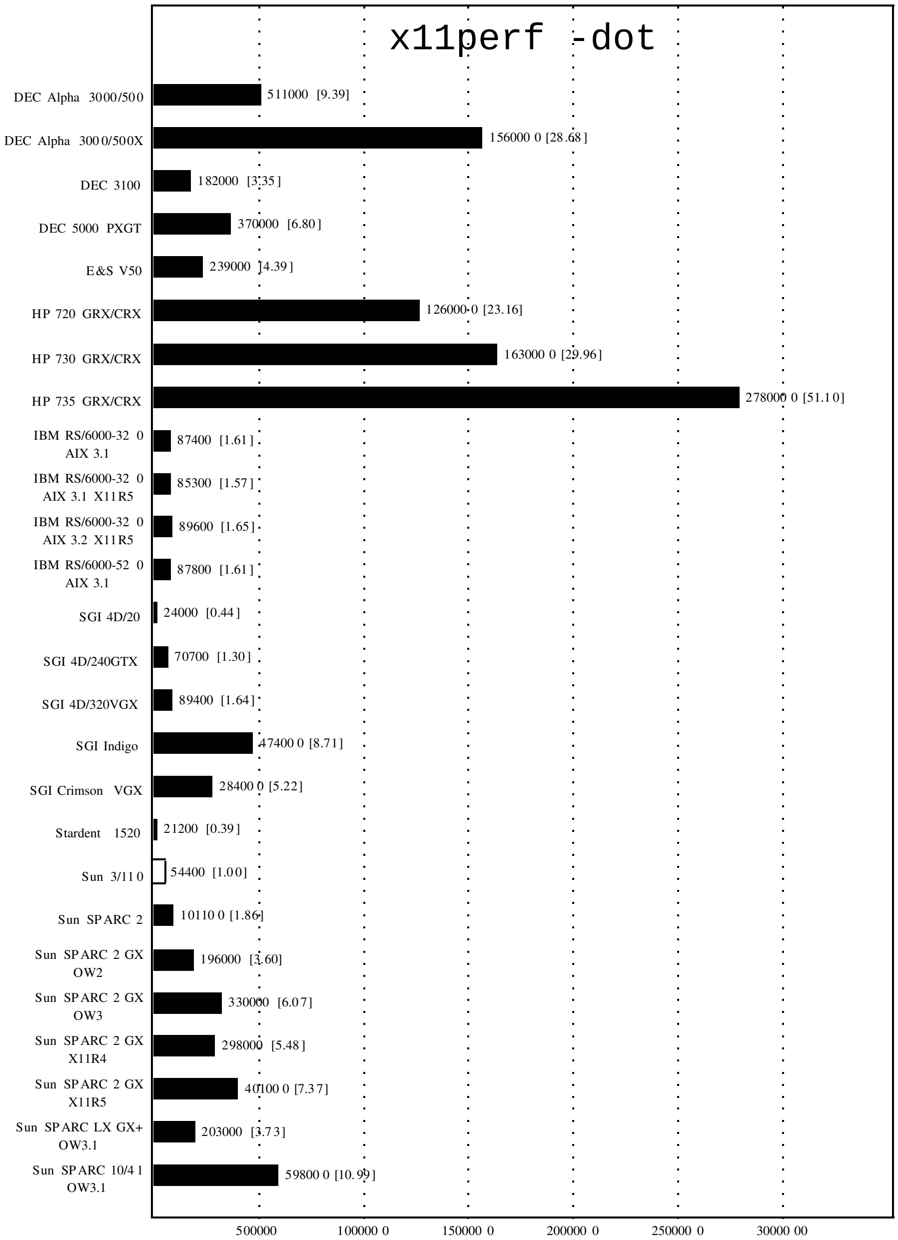

Figure 1: `x11perf -dot` benchmark in operations/second. In this, and following figures, the square-bracketed values are performance results measured relative to the Sun 3/110. The Sun 3/110 bars are always hollow, and systems are ordered alphabetically from top to bottom.

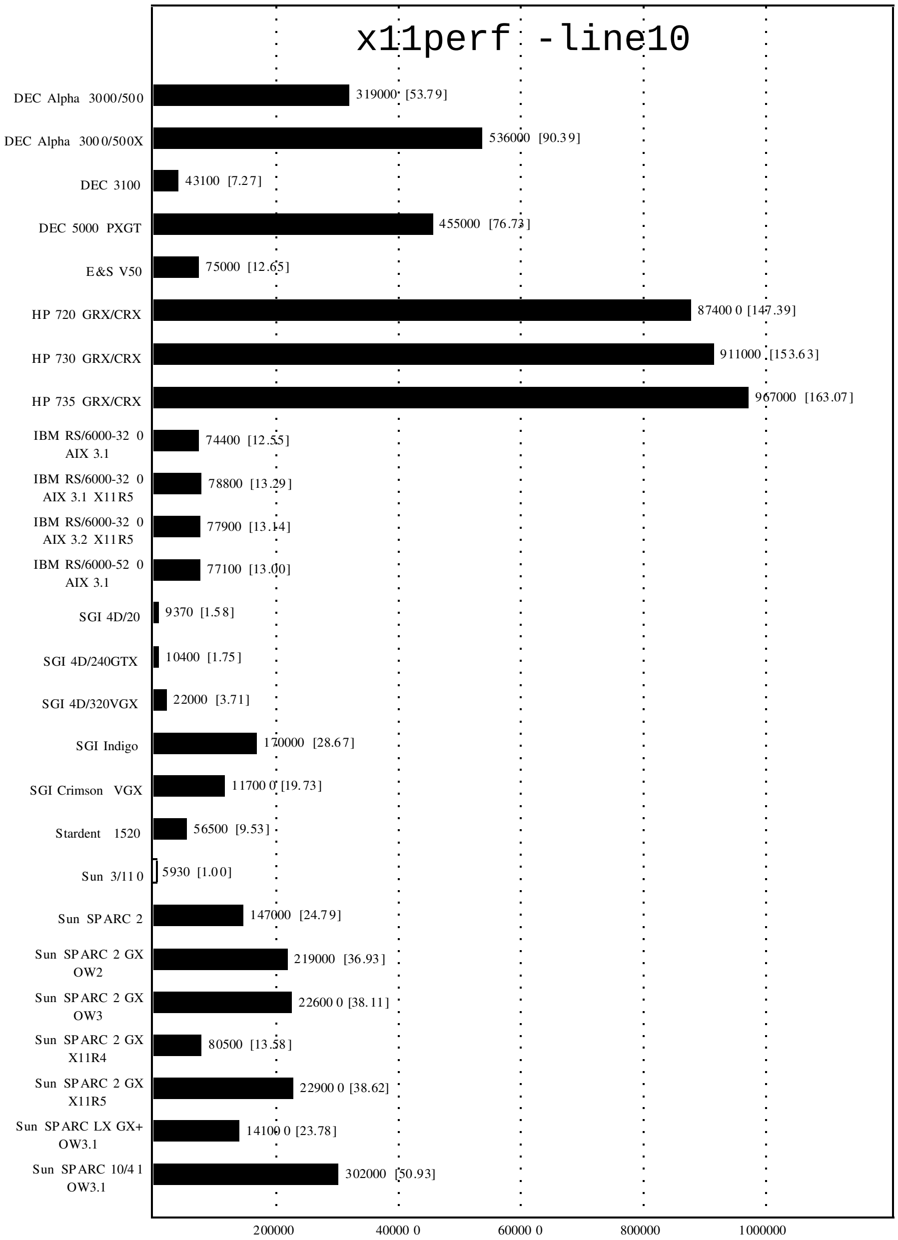

Figure 2: `x11perf -line10` (10-pixel lines) benchmark in operations/second, and relative to Sun 3/110.

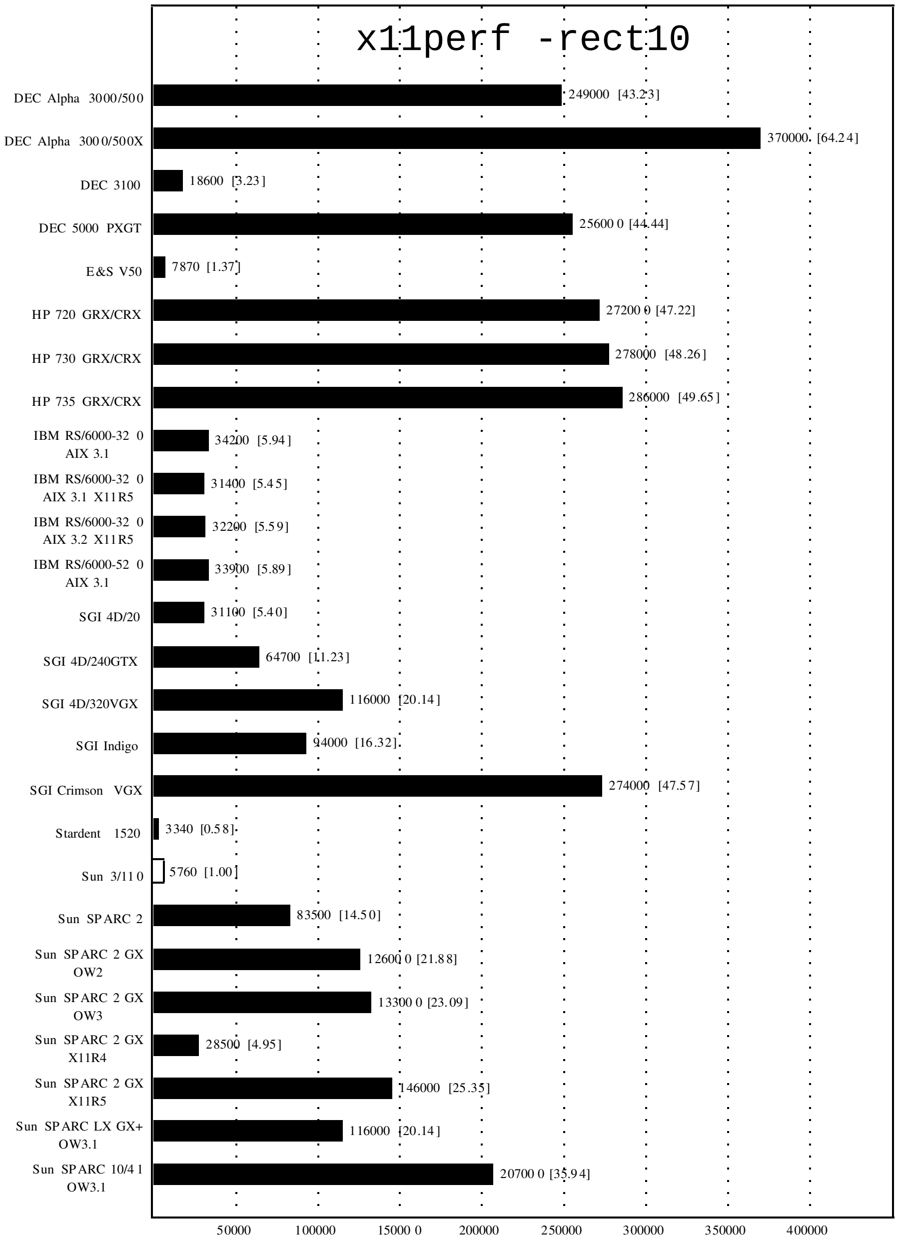

Figure 3: `x11perf -rect10` (10-pixel rectangles) benchmark in operations/second, and relative to Sun 3/110.

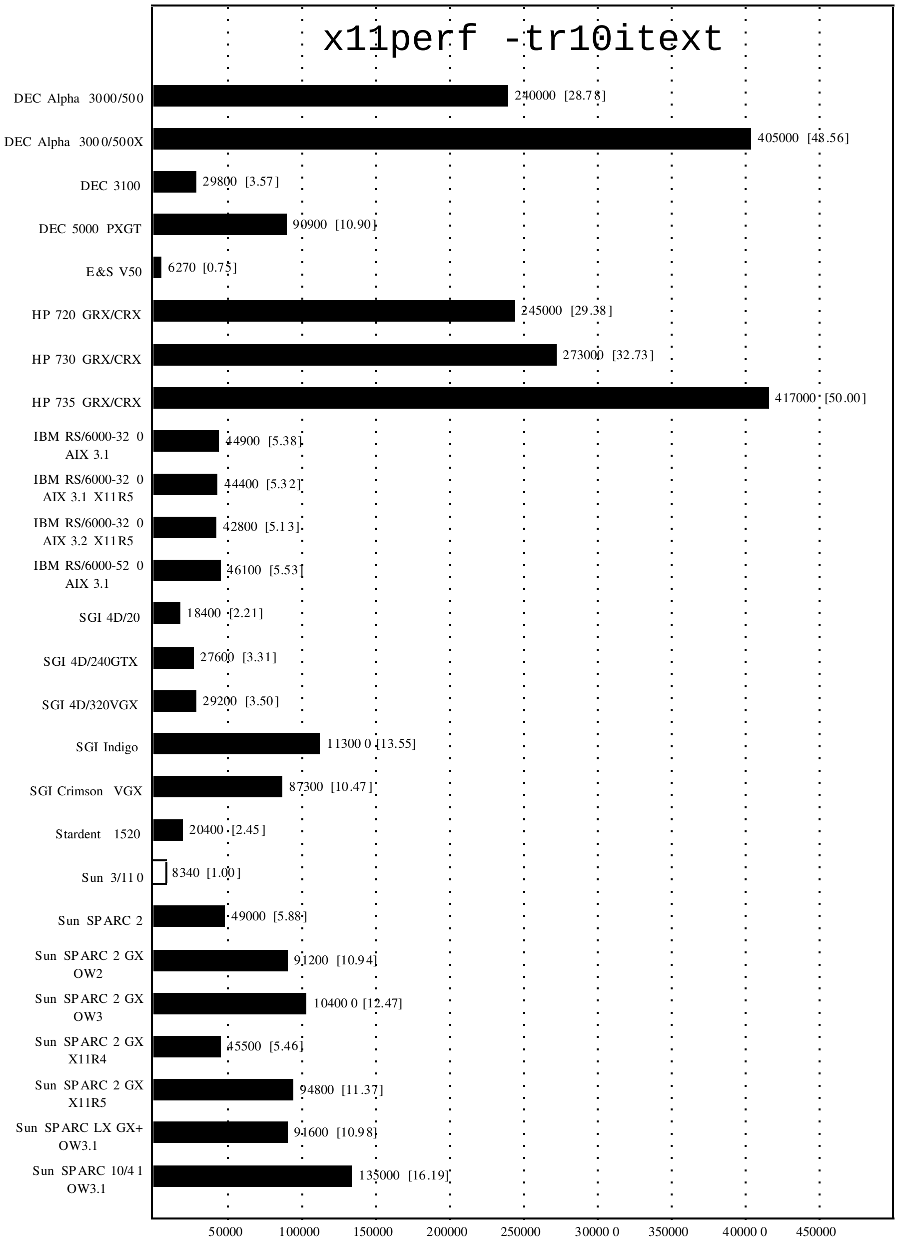

Figure 4: `x11perf -tr10itext` (80-character lines of 10pt variable-width Times Roman text) benchmark in characters/second, and relative to Sun 3/110.

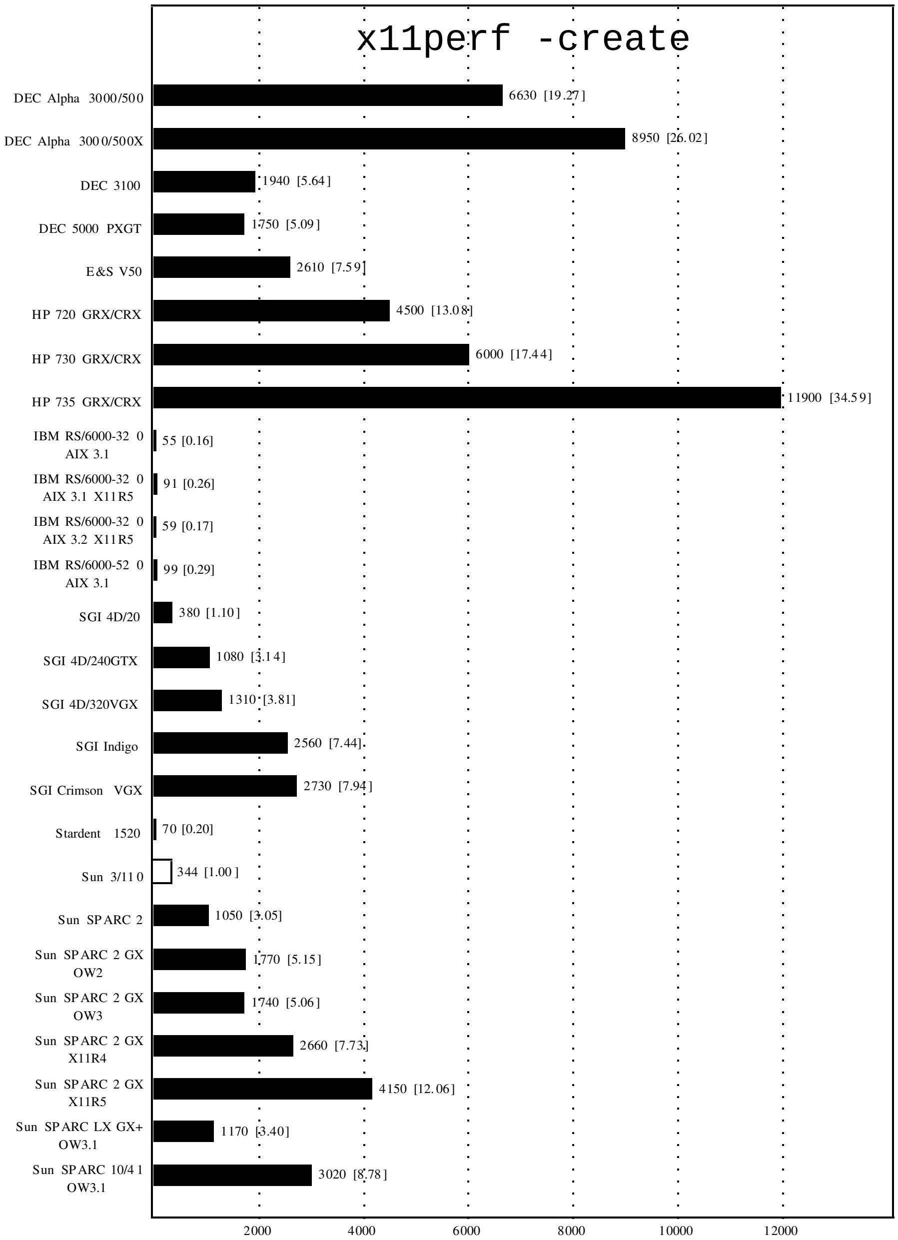

Figure 5: `x11perf -create` (50 kids) benchmark in operations/second, and relative to Sun 3/110.