

Comparison of Various Packing Algorithms for Usage with the **DXF Viewer**

Rammi^{*}

October 27, 2005

1 Introduction

A customer told me that it took at least twice the time to read a file packed with `bzip2` than the naked file. So I made some simple tests and took a 30 MByte DXF file, packed it into all formats the **DXF Viewer** understands, and loaded all files into a newly started viewer to avoid problems with garbage collection.

The testing environment was a Athlon 1600+ machine with 1 GByte of memory under Linux, using a Sun Java 1.4.2-b28, and a Mobile Pentium 2000 with 512 MByte under Windows 2000, with Java 1.5.0_02-b09. In both case I used the `-mx400m` switch to increase the memory available to the virtual machine. I reran some of the tests to see whether the results are reliable, and the measured durations did not differ more than a second, good enough for our needs. The files were on the local hard drive.

2 Results

Type	Size	Time/Linux	Time/Windows	Pack Ratio	Rate Limit
.dxf	31022919	15466	11487	1.0	<i>n.a.</i>
.bz2	1617418	79922	56230	19.2	5 MBit/s
.gz	2200634	76278	49982	14.1	7 MBit/s
.zip	2200761	16709	12958	14.1	180 MBit/s

The last column displays the speed of the connection which would be necessary to load the file as fast as it is unpacked by the viewer. That means if your connection is slower you will gain by packing the file before transmission: it is displayed faster.

The times to compute this value are taken from the Windows results (because computers are always getting faster), and having a machine which is twice as fast would push up the rates by a factor of two. Because of taking differences the last column has a huge error span, especially for the `.zip` result where the time difference is comparable to the timing error.

3 Conclusions

The winner is perfectly clear: unless storage space is very limited use `zip` to pack your DXF files for the **DXF Viewer**. Only if you have severe space restrictions or the files are accessed over very slow internet lines¹ you can think about packing with `bzip2`. From the **DXF Viewer**'s point of view packing with `gzip` has no advantage: `bzip2` packs a lot better with comparable unpack times, and `zip` packs comparable, but has nearly no unpacking time overhead.

¹To make `bzip2` win over `zip` your access line has to be slower than 13.5 kBit/s.