

Logs of Real Parallel Workloads from Production Systems
















































This page points to detailed workload logs collected from large scale parallel systems in production use in various places around the world.

The original logs come in different formats. Information about the individual format for each log is given to the degree available in the log's associated notes document. In addition to the original format, all logs are converted to the Standard Workload Format (SWF). Some of the logs also have cleaned versions; it is recommended that these be used when available.

To promote reproducibility, files are named according to the following convention. Each file name has two parts, identifying the dataset and the file version. The dataset identifier is constructed as `<site>-<machine>-<year>`. The version is a serial number and possible modifier. Number 0 is reserved for the original (unconverted) version. The number is increased by 1 for each new conversion (in case errors were found in a conversion and additional conversions were performed). Modified versions are indicated by a minor number. Thus a file name of CTC-SP2-1996-2.1-cln.swf is based on the dataset CTC-SP2-1996 (log of the SP2 machine at CTC starting from 1996), and represents a cleaned version of conversion 2 of this dataset. When using these logs, please specify exactly which file you are using, in order to enable others to reproduce your work.

Please send comments and additional information to feit@cs.huji.ac.il.


The Log Files:

| # | Name | From | To | Mon | CPUs | U | E | M | Standard Format: <i>CLEANED</i> | | | | Standard Format: <i>ORIGINAL</i> | | | | Original Log |
|----|-------------------------------|--|----------|-----|-------|---|---|---|------------------------------------|-------|-------|--|-------------------------------------|-------|-------|--|--|
| | | | | | | | | | Jobs | Users | Util% | File | Jobs | Users | Util% | File | |
| 1 | NASA iPSC |  Oct 1993 | Dec 1993 | 3 | 128 | | | | 18,239 | 69 | 46.6 |  204K | 42,264 | 69 | 46.7 |  400K |  290K |
| 2 | LANL CM5 |  Oct 1994 | Sep 1996 | 24 | 1,024 | ✓ | ✓ | | 122,060 | 213 | 74.4 |  2.1M | 201,387 | 213 | 75.2 |  2.8M |  14M  3.6M |
| 3 | SDSC Par95 |  Dec 1994 | Dec 1995 | 12 | 400 | | | | 53,970 | 98 | 71.6 |  761K | 76,872 | 98 | 71.7 |  1.1M |  1.4M |
| 4 | SDSC Par96 |  Dec 1995 | Dec 1996 | 12 | 400 | | | | 32,135 | 60 | 75.6 |  487K | 38,719 | 60 | 75.6 |  585K |  831K |
| 5 | Early CTC SP2 |  Jun 1995 | Apr 1996 | 10 | 430 | | ✓ | | use original | | | | 75,944 | 642 | 62.9 |  1.4M | |
| 6 | CTC SP2 |  Jun 1996 | May 1997 | 11 | 338 | ✓ | | | 77,222 | 679 | 85.2 |  1.5M | 79,302 | 679 | 85.2 |  1.5M |  3.5M |
| 7 | LLNL T3D |  Jun 1996 | Sep 1996 | 4 | 256 | | | | use original | | | | 22,779 | 153 | 60.1 |  243K |  569K  246K |
| 8 | KTH SP2 |  Sep 1996 | Aug 1997 | 11 | 100 | ✓ | | | 28,476 | 214 | 70.1 |  402K | 28,489 | 214 | 70.4 |  402K |  722K |
| 9 | SDSC SP2 |  Apr 1998 | Apr 2000 | 24 | 128 | ✓ | | | 59,715 | 437 | 83.4 |  1.2M | 73,496 | 437 | 83.7 |  1.4M |  3.1M |
| 10 | LANL O2K |  Nov 1999 | Apr 2000 | 5 | 2,048 | | ✓ | | use original | | | | 122,233 | 337 | 69.7 |  2.1M |  2.4M |
| 11 | OSC Cluster |  Jan 2000 | Nov 2001 | 22 | 178 | | | | 36,097 | 253 | 12.8 |  524K | 80,714 | 254 | 13.8 |  1.1M |  2.0M |
| 12 | SDSC BLUE |  Apr 2000 | Jan 2003 | 32 | 1,152 | ✓ | | | 243,306 | 468 | 76.7 |  3.9M | 250,440 | 468 | 76.8 |  4.0M |  6.0M |
| 13 | Sandia Ross |  Nov | Jan | 37 | 1,524 | ✓ | ✓ | | 57,882 | 203 | 49.9 | | 85,355 | 204 | 50.2 | | |

| | | | | | | | | | | | | | | | | | | | | |
|----|---------------------------------------|--|----------|----------|----|---------|---|---|--------------|-----|------|------|------------|---------|------|------|------|------|------|------|
| | | | 2001 | 2005 | | | | | | | | 1.1M | | | | | 1.5M | | | |
| 14 | HPC2N | | Jul 2002 | Jan 2006 | 42 | 240 | ✓ | ✓ | 202,871 | 257 | 60.1 | | 2.9M | 527,371 | 258 | 70.2 | | 6.7M | | 13M |
| 15 | DAS2 fs0 | | Jan 2003 | Jan 2004 | 12 | 144 | ✓ | ✓ | use original | | | | 225,711 | 102 | 14.9 | | 2.2M | | 14M | |
| 16 | DAS2 fs1 | | Jan 2003 | Dec 2003 | 12 | 64 | ✓ | ✓ | use original | | | | 40,315 | 36 | 12.0 | | 376K | | | |
| 17 | DAS2 fs2 | | Jan 2003 | Dec 2003 | 12 | 64 | ✓ | ✓ | use original | | | | 66,429 | 52 | 19.5 | | 641K | | | |
| 18 | DAS2 fs3 | | Jan 2003 | Dec 2003 | 12 | 64 | ✓ | ✓ | use original | | | | 66,737 | 64 | 10.7 | | 575K | | | |
| 19 | DAS2 fs4 | | Feb 2003 | Dec 2003 | 11 | 64 | ✓ | ✓ | use original | | | | 33,795 | 40 | 14.5 | | 311K | | | |
| 20 | SDSC DataStar | | Mar 2004 | Apr 2005 | 13 | 1,664 | ✓ | | 96,069 | 460 | 67.2 | | 1.6M | 96,089 | 460 | 63.1 | | 1.6M | | |
| 21 | LPC EGEE | | Aug 2004 | May 2005 | 9 | 140 | ✓ | ✓ | 234,889 | 56 | 24.4 | | 2.6M | 244,821 | 57 | 20.8 | | 2.7M | | 14M |
| 22 | LCG | | Nov 2005 | Dec 2005 | 1 | 24,515 | | | use original | | | | 188,041 | 216 | | | 1.9M | | 1.3M | |
| 23 | SHARCNET | | Dec 2005 | Jan 2007 | 13 | 6,828 | | ✓ | use original | | | | 1,195,242 | 412 | 43.6 | | 18M | | 31M | |
| 24 | SHARCNET Whale | | Jun 2006 | Jan 2007 | 7 | 3,072 | | ✓ | use original | | | | 589,251 | 154 | 72.1 | | 8.6M | | | |
| 25 | LLNL uBGL | | Nov 2006 | Jun 2007 | 7 | 2,048 | ✓ | | use original | | | | 112,611 | 62 | 56.1 | | 732K | | 1.2M | |
| 26 | LLNL Atlas | | Nov 2006 | Jun 2007 | 8 | 9,216 | ✓ | | 43,778 | 131 | 70.5 | | 492K | 60,332 | 132 | 64.1 | | 633K | | 1.2M |
| 27 | LLNL Thunder | | Jan 2007 | Jun 2007 | 5 | 4,008 | ✓ | | 121,039 | 283 | 86.7 | | 1.3M | 128,662 | 283 | 87.9 | | 1.4M | | 2.2M |
| 28 | ANL Intrepid | | Jan 2009 | Sep 2009 | 8 | 163,840 | ✓ | | use original | | | | 68,936 | 236 | 59.6 | | 893K | | | |
| 29 | MetaCentrum | | Dec 2008 | Jun 2009 | 7 | 806 | | ✓ | use original | | | | 103,656 | 147 | | | 1.5M | | 2.0M | |
| 30 | PIK IPLEX | | Apr 2009 | Jul 2012 | 40 | 2,560 | | ✓ | use original | | | | 742,965 | 225 | 38.0 | | 11M | | 150M | |
| 31 | RICC | | May 2010 | Sep 2010 | 5 | 8,192 | ✓ | ✓ | use original | | | | 447,794 | 176 | 87.2 | | 3.6M | | 4.0M | |
| 32 | CEA CURIE | | Feb 2011 | Oct 2012 | 20 | 93,312 | ✓ | | 312,826 | 582 | 62.0 | | 3.5M | 773,138 | 722 | 29.3 | | 6.4M | | 5.9M |
| 33 | Intel NetBatch pool A | | Nov 2012 | Dec 2012 | 1 | | | ✓ | use original | | | | 13,368,191 | 1,104 | | | 139M | | 576M | |
| 34 | Intel NetBatch pool B | | Oct 2012 | Nov 2012 | 1 | | | ✓ | use original | | | | 13,085,800 | 1,615 | | | 136M | | 462M | |
| 35 | Intel NetBatch pool C | | Oct 2012 | Nov 2012 | 1 | | | ✓ | use original | | | | 13,313,793 | 1,146 | | | 123M | | 586M | |
| 36 | Intel NetBatch pool D | | Oct 2012 | Nov 2012 | 1 | | | ✓ | use original | | | | 9,054,066 | 862 | | | 92M | | 411M | |
| 37 | UniLu Gaia | | May 2014 | Aug 2014 | 3 | 2,004 | ✓ | ✓ | use original | | | | 51,987 | 84 | 47.9 | | 837K | | 755K | |
| 38 | MetaCentrum2 | | Jan 2013 | Apr 2015 | 28 | 8,412 | ✓ | ✓ | use original | | | | 5,731,100 | 880 | | | 56M | | 69M | |

get [a list of the current \(cleaned\) versions of all logs](#), useful for downloading with wget -i

= Important notes associated with log

 = File to download

Mon = Duration of log in months

UE = Indicates whether log contains **user runtime estimates**

[You can add artificial estimates to logs that lack them by using [this utility](#)]

Mem = Indicates whether log contains data about memory (requested or used)

Util% = **Utilization** expressed as percentage

The ORIGINAL SWF files reflect the data "as is", with only minor modifications to make them self-consistent (for example, jobs that seem to start before they arrive are modified to equate the arrival and start times). Regrettably, the data often includes problematic and unrepresentative data, such as significant automated administrative activity or large-scale flurries of activity by single users. In order to ease the use of these logs in performance evaluations, we therefore also provide a CLEANED version of the logs, and recommend that this version be used (of course, you are urged to make your own decision regarding what cleaning if at all should be applied, but if you don't want to get into this, we suggest you use ours; this has the added benefit of ensuring your work is comparable with that of others). Specific information about the cleaning performed on each log is detailed in the "usage notes" section of the log's notes document. Further information about identifying and justifying the removal of problematic and unrepresentative data may be found in:

- D. G. Feitelson and D. Tsafir, "[Workload sanitation for performance evaluation](#)". In *IEEE Intl. Symp. Performance Analysis of Systems and Software*, pp. 221-230, Mar 2006.
- D. Tsafir and D. G. Feitelson, "[Instability in parallel job scheduling simulation: the role of workload flurries](#)". In *20th Intl. Parallel and Distributed Processing Symp.*, Apr 2006.

In addition, older versions of the converted files are sometimes available from the log's notes page. These should only be used for reproducing or verifying previous research that used those versions.

| | |
|--------------------------------|--|
| Last update | Mon Sep 7 11:06:50 IDT 2015 |
| Visit count since Aug 20, 2004 | 00067058 [see some access statistics here] |
| Automatically generated by | ./mk-logs.pl |

Change log:

| # | Date | Log File | Change |
|----|-------------|----------------------------|---|
| 83 | 07 Sep 2015 | METACENTRUM-2013-2.swf | Re-converted to adhere to SWF standard. |
| 82 | 07 Sep 2015 | METACENTRUM-2013-1.swf | Added to archive. |
| 81 | 31 Mar 2015 | UniLu-Gaia-2014-2.swf | Re-conversion with minor adjustments. |
| 80 | 31 Mar 2015 | UniLu-Gaia-2014-1.swf | Added to archive. |
| 79 | 08 Feb 2015 | HPC2N-2002-2.2-cln.swf | New cleaning (remove first 10 jobs) |
| 78 | 27 Jan 2015 | SDSC-DS-2004-2.1-cln.swf | Cleaned (remove first 20 jobs) |
| 77 | 27 Jan 2015 | SDSC-BLUE-2000-4.2-cln.swf | New cleaning (remove first 8 jobs) |
| 76 | 27 Jan 2015 | SDSC-SP2-1998-4.2-cln.swf | New cleaning (remove first 10 jobs) |
| 75 | 27 Jan | KTH-SP2-1996-2.1- | Cleaned (remove first 14 jobs) |

| | | | |
|----|-------------|-------------------------------|--|
| | 2015 | cln.swf | |
| 74 | 28 Oct 2013 | Intel-NetbatchD-2012-1.swf | Added to archive. |
| 73 | 28 Oct 2013 | Intel-NetbatchC-2012-1.swf | Added to archive. |
| 72 | 28 Oct 2013 | Intel-NetbatchB-2012-1.swf | Added to archive. |
| 71 | 28 Oct 2013 | Intel-NetbatchA-2012-1.swf | Added to archive. |
| 70 | 21 Aug 2013 | CEA-Curie-2011-2.1-cln.swf | Cleaned (removed initial period and 4 flurries) |
| 69 | 26 Jun 2013 | CEA-Curie-2011-2.swf | Re-conversion to correct sorting of jobs and numbering and remove bogus long runtimes. |
| 68 | 26 Jun 2013 | CEA-Curie-2011-1.swf | Added to archive. |
| 67 | 12 Sep 2012 | PIK-IPLEX-2009-1.swf | Added to archive. |
| 66 | 13 Dec 2011 | METACENTRUM-2009-2.swf | New conversion with clusters data. |
| 65 | 08 Dec 2011 | SHARCNET-Whale-2006-2.swf | Added to archive. |
| 64 | 30 Nov 2011 | CTC-SP2-1996-3.swf | Changed to 338 processors |
| 63 | 30 Nov 2011 | | New conversions providing improved checks and handling of missing data. |
| 62 | 13 Nov 2011 | OSC-Clust-2000-3.swf | New conversion with correct number of processors |
| 61 | 09 Nov 2011 | RICC-2010-2.swf | Re-conversion to correct sorting of jobs and numbering. |
| 60 | 08 Nov 2011 | Sandia-Ross-2001-1.1-cln.swf | Cleaned (removed multiple large flurries by 3 users) |
| 59 | 08 Nov 2011 | Sandia-Ross-2001-1.swf | Added to archive. |
| 58 | 26 Oct 2011 | RICC-2010-1.swf | Added to archive. |
| 57 | 25 Sep 2011 | METACENTRUM-2009-1.swf | Added to archive. |
| 56 | 23 Feb 2011 | ANL-Intrepid-2009-1.swf | Added to archive. |
| 55 | 08 Apr 2010 | SDSC-SP2-1998-0 | Added to archive. |
| 54 | 18 Mar 2008 | LLNL-thunder-2007-1.1-cln.swf | Cleaned (3 small flurries). |
| 53 | 18 Mar 2008 | LLNL-thunder-2007-1.swf | Added to archive. |
| 52 | 18 Mar 2008 | LLNL-atlas-2006-1.1-cln.swf | Cleaned (removed initial period and 6 flurries). |
| 51 | 18 Mar | LLNL-atlas-2006-1.swf | Added to archive. |

| | | | |
|----|-------------|---------------------------|--|
| | 2008 | | |
| 50 | 24 Feb 2008 | LLNL-uBGL-2006-1.swf | Added to archive. |
| 49 | 15 Jan 2008 | LPC-EGEE-2004-1.2-cln.swf | Corrected the deleted period. |
| 48 | 03 Jan 2008 | HPC2N-2002-1.1-cln.swf | Cleaned (removed all activity of user 2, and 2 flurries). |
| 47 | 03 Jan 2008 | HPC2N-2002-1.swf | Added to archive. |
| 46 | 24 Jan 2007 | LCG-2005-1.swf | Added to archive. |
| 45 | 21 Jan 2007 | SHARCNET-2005-1.swf | Added to archive. |
| 44 | 06 Aug 2006 | SDSC-DS-2004-1.swf | Added to archive. |
| 43 | 01 Aug 2006 | CTC-SP2-1995-1.swf | Added to archive. |
| 42 | 01 Aug 2006 | | Major overhaul with new conversions and renaming of files. |
| 41 | 13 Nov 2005 | LPC-EGEE-2004-1.1-cln.swf | Cleaned (removed initial testing period). |
| 40 | 13 Nov 2005 | LPC-EGEE-2004-1.swf | Added to archive. |
| 39 | 25 Sep 2005 | | Added TimeZoneString headers. |
| 38 | 16 Dec 2004 | SDSC-Par-1996-1.swf | Negative wait times changed to -1. |
| 37 | 16 Dec 2004 | SDSC-Par-1995-1.1-cln.swf | Removed 11 jobs with large negative wait times. |
| 36 | 16 Dec 2004 | SDSC-Par-1995-1.swf | Negative wait times and runtimes changed to -1. |
| 35 | 12 Dec 2004 | LANL-O2K-1999-1.swf | Changed the used-cpu field of 958 jobs from -2 to -1. |
| 34 | 09 Dec 2004 | DAS2-fs4-2003-1.swf | Added to archive. |
| 33 | 09 Dec 2004 | DAS2-fs3-2003-1.swf | Added to archive. |
| 32 | 09 Dec 2004 | DAS2-fs2-2003-1.swf | Added to archive. |
| 31 | 09 Dec 2004 | DAS2-fs1-2003-1.swf | Added to archive. |
| 30 | 09 Dec 2004 | DAS2-fs0-2003-1.swf | Added to archive. |
| 29 | 09 Dec 2004 | OSC-Clust-2000-1.swf | Added to archive. |
| 28 | 25 Nov 2004 | SDSC-Par-1995-1.1-cln.swf | Removed 15 jobs with large negative runtimes. |
| 27 | 04 Sep | LANL-CM5-1994-2.2- | Removed 3 more jobs. |

| | | | |
|----|-------------|----------------------------|--|
| | 2004 | cln.swf | |
| 26 | 07 Jul 2004 | SDSC-BLUE-2000-2.1-cln.swf | Cleaned (removed 3 flurries). |
| 25 | 07 Jul 2004 | SDSC-BLUE-2000-2.swf | New conversion with minor correction. |
| 24 | 07 Jul 2004 | SDSC-SP2-1998-2.1-cln.swf | Cleaned (4 flurries removed). |
| 23 | 06 Jul 2004 | CTC-SP2-1996-1.1-cln.swf | Cleaned (one flurry removed). |
| 22 | 06 Jul 2004 | LANL-CM5-1994-2.1-cln.swf | Cleaned (three large flurries removed). |
| 21 | 05 Jul 2004 | NASA-iPSC-1993-1.1-cln.swf | Cleaned (thousands of pwd jobs by sysadmin removed). |
| 20 | 09 Jan 2004 | SDSC-BLUE-2000-1.swf | Added to archive. |
| 19 | 08 Jan 2004 | LANL-O2K-1999-1.swf | Conversion to SWF added. |
| 18 | 08 Jan 2004 | SDSC-SP2-1998-2.swf | New conversion with corrections. |
| 17 | 23 May 2001 | LANL-CM5-1994-2.swf | Re-converted; substitute start for arrival if missing. |
| 16 | 18 Jun 2000 | LANL-O2K-1999-0.swf | Added to archive. |
| 15 | 29 May 2000 | SDSC-SP2-1998-1.swf | Added to archive. |
| 14 | 12 Nov 1999 | KTH-SP2-1996-1.swf | Conversion to SWF added. |
| 13 | 12 Nov 1999 | LANL-CM5-1994-1.swf | Conversion to SWF added. |
| 12 | 12 Nov 1999 | NASA-iPSC-1993-1.swf | Conversion to SWF added. |
| 11 | 08 Aug 1999 | CTC-SP2-1996-1.swf | Conversion to SWF added. |
| 10 | 01 Aug 1999 | SDSC-Par-1996-1.swf | Conversion to SWF added. |
| 9 | 01 Aug 1999 | SDSC-Par-1995-1.swf | Conversion to SWF added. |
| 8 | 11 Mar 1999 | LLNL-T3D-1996-1.swf | Added to archive. |
| 7 | 01 Dec 1998 | KTH-SP2-1996-0 | Added to archive. |
| 6 | 01 Dec 1998 | CTC-SP2-1996-0 | Added to archive. |
| 5 | 01 Dec 1998 | SDSC-Par-1996-0 | Added to archive. |
| 4 | 01 Dec 1998 | SDSC-Par-1995-0 | Added to archive. |
| 3 | 01 Dec | LANL-CM5-1994-0 | Added to archive. |

| | | | |
|---|-------------|------------------|----------------------|
| | 1998 | | |
| 2 | 01 Dec 1998 | NASA-iPSC-1993-0 | Added to archive. |
| 1 | 01 Dec 1998 | | Archive goes online. |