

## Introduction

This document will guide you through the process of using Integrated Performance Monitoring on HPC Wales.

### What is IPM?

*Integrated Performance Monitoring (IPM) is a portable profiling infrastructure that provides a low-overhead framework for measuring the performance and resource utilization of parallel codes. It collects metrics about communication, computation and IO.*

Further details on accessing HPC Wales systems can be found in the User Guide and help can be obtained through the Support Desk:

- E-mail [support@hpcwales.co.uk](mailto:support@hpcwales.co.uk)
- Telephone 08452 572 207
- Website <https://hpcwprod.service-now.com/>

## Table of Contents

Profiling With IPM	2
IPM Output	2
Controlling the Level of Output	5
Online Documentation	5

## Profiling With IPM

You do not need to recompile your code to use IPM. Simply load the ipm module in your batch script and replace your line that calls mpirun with the following:

```
LD_PRELOAD=$IPM_LD_PRELOAD mpirun [args] ./myexec
```

where [args] are the arguments you pass to mpirun and myexec is your executable.

## IPM Output

IPM prints a report at the end of the output file from your job:

```
##IPMv0.983#####
#
# command : unknown (completed)
# host      : cf-htc-103/x86_64_Linux          mpi_tasks : 12 on 1 nodes
# start     : 05/09/13/11:46:03              wallclock  : 0.509226 sec
# stop      : 05/09/13/11:46:04              %comm     : 27.33
# gbytes    : 0.00000e+00 total              gflop/sec  : 0.00000e+00 total
#
#####
# region   : *          [ntasks] =      12
#
#          [total]          <avg>          min          max
# entries          12          1          1          1
# wallclock        6.10535      0.508779    0.508311    0.509226
# user             5.73912      0.47826    0.469928    0.486925
# system           0.237957     0.0198297  0.014997    0.023996
# mpi              1.67033      0.139194   0.132735    0.14891
# %comm            27.3344      26.0661    29.2481
# gflop/sec        0           0           0           0
# gbytes           0           0           0           0
#
#          [time]          [calls]          <%mpi>          <%wall>
# MPI_Allreduce    1.15443      231900      69.11         18.91
# MPI_Waitall      0.281504     664938     16.85         4.61
# MPI_Isend        0.110112     430254      6.59          1.80
# MPI_Irecv        0.0655586    430254      3.92          1.07
# MPI_Barrier      0.0587168     24          3.52          0.96
#####
```

The report has three sections. The header provides general information such as total run-time and percentage of time spent on communication. The second section shows the variation across processes by reporting the average, minimum and maximum values of the observed timing data. The third section reports the number of calls for every MPI function used in the code and the total time spent in each.

There will also be an XML file in the same directory in which your job ran that will have a name of the form <user.name><user.name>.<numbers>.<numbers>.<numbers>. This can be post-processed with the ipm\_parse utility to give a detailed web-page view of IPM's output:

## IPM Output

`ipm_parse -html <filename>`

This will produce a directory containing some html pages that you can view in the browser of your choice. To do this, copy the files from HPC Wales to your local machine. The report will look like this:

The screenshot shows a Mozilla Firefox browser window displaying an IPM profile for process 263844. The browser's address bar shows 'IPM profile for 263844'. The main content area is divided into several sections:

- Process Information:** A table showing details for process 263844.
 

command: unknown			
codename:	unknown	state:	running
username:		group:	
host:	cf-htc-023 (x86_64_Linux)	mpi_tasks:	24 on 2 hosts
start:	12/20/13/14:52:37	wallclock:	6.61298e-01 sec
stop:	12/20/13/14:52:38	%comm:	26.8019800957108
total memory:	0 gbytes	total gflop/sec:	-36.2922615825241
switch(send):	0 gbytes	switch(recv):	0 gbytes
- Computation:** A table with columns 'Event', 'Count', and 'Pop'.
 

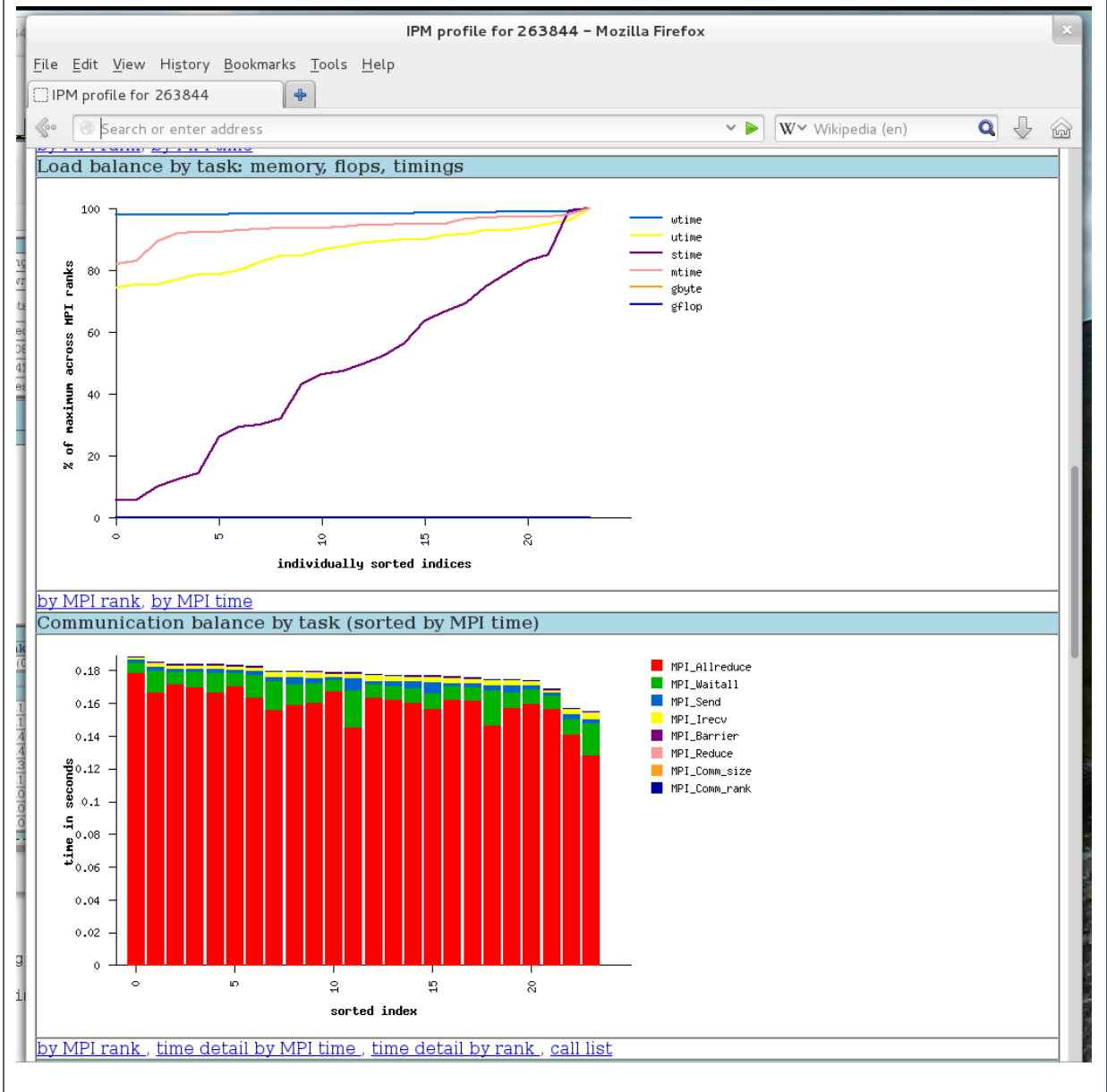
Event	Count	Pop
NULL	0	*
- Communication:** A pie chart showing the distribution of MPI operations. The legend includes: MPI\_Allreduce (red), MPI\_Waitall (green), MPI\_Send (blue), MPI\_Irecv (yellow), MPI\_Barrier (purple), MPI\_Reduce (orange), MPI\_Comm\_size (light blue), and MPI\_Comm\_rank (dark blue). MPI\_Allreduce is the dominant operation.
- HPM Counter Statistics:** A table with columns 'Event', 'Ntasks', 'Avg', 'Min(rank)', and 'Max(rank)'.
 

Event	Ntasks	Avg	Min(rank)	Max(rank)
NULL	*	0.00	0 (0)	0 (0)
- Communication Event Statistics (100.00% detail, 2.6153e-05 error):** A table with columns 'Event', 'Buffer Size', 'Ncalls', 'Total Time', 'Min Time', 'Max Time', '%MPI', and '%Wall'.
 

Event	Buffer Size	Ncalls	Total Time	Min Time	Max Time	%MPI	%Wall
MPI_Allreduce	8	377112	3.840	4.897e-06	4.631e-03	90.27	24.19
MPI_Waitall	16	172832	0.182	3.914e-07	8.975e-05	4.27	1.14
MPI_Send	8	361376	0.073	1.739e-07	2.226e-05	1.72	0.46
MPI_Waitall	92	24	0.070	3.264e-03	3.277e-03	1.65	0.44
MPI_Irecv	8	361376	0.059	1.243e-07	2.791e-05	1.38	0.37
MPI_Waitall	8	15734	0.018	1.320e-07	3.113e-03	0.42	0.11
MPI_Barrier	0	48	0.006	4.648e-05	1.824e-04	0.14	0.04
MPI_Irecv	4	598	0.004	1.080e-07	2.059e-04	0.09	0.02
MPI_Reduce	8	168	0.002	1.194e-06	6.401e-05	0.05	0.01

Scroll down to find lots of informative graphs about message sizes, communication patterns and so forth. These provide a per-process view of the collected readings:

# IPM Output



## Controlling the Level of Output

The level of detail that IPM reports is controlled by the `IPM_REPORT` environment variable.

Value	Description
<code>terse</code>	(default) Aggregate wallclock time, memory usage and flops are reported along with the percentage of wallclock time spent in MPI calls.
<code>full</code>	Each hardware counter is reported (where available) as are all of wallclock, user, system, and MPI time. The contribution of each MPI call to the communication time is given.
<code>none</code>	No report

## Online Documentation

The [full IPM user guide](#) is available online.