

Craypat Automatic Performance Analysis A(APA)

Automatic Profiling Analysis (APA)

1. Load CrayPat & Cray Apprentice² module files
 - % module load xt-craypat/4.4.0.4 apprentice2/4.4.0.4
2. Build application
 - % make clean
 - % make
3. Instrument application for automatic profiling analysis
 - % pat_build **-O apa** a.out
 - You should get an instrumented program a.out+pat
4. Run application to get top time consuming routines
 - Remember to modify <script> to **run a.out+pat**
 - Remember to **run on Lustre**
 - % aprun ... **a.out+pat** (or qsub <pat script>)
 - You should get a performance file (“<sdatafile>.xf”) or multiple files in a directory <sdatadir>
5. **Generate .apa file (4.2 only)**
 - % pat_report -o *my_sampling_report* [<sdatafile>.xf | <sdatadir>]
 - creates a report file & **an automatic profile analysis file <apafilename>.apa**

APA File Example

```
# You can edit this file, if desired, and use it
# to reinstrument the program for tracing like this:
#
# pat_build -O mhd3d.Oapa.x+4125-401sdt.apa
#
# These suggested trace options are based on data from:
#
# /home/crayadm/ldr/mhd3d/run/mhd3d.Oapa.x+4125-401sdt.ap2, /home/
#   crayadm/ldr/mhd3d/run/mhd3d.Oapa.x+4125-401sdt.xf
#
# -----
#
# HWPC group to collect by default.
#
# -Drtenv=PAT_RT_HWPC=1 # Summary with instructions metrics.
#
# -----
#
# Libraries to trace.
#
# -g mpi
#
# -----
#
# User-defined functions to trace, sorted by % of samples.
# Limited to top 200. A function is commented out if it has < 1%
# of samples, or if a cumulative threshold of 90% has been reached,
# or if it has size < 200 bytes.
#
# -w # Enable tracing of user-defined functions.
# Note: -u should NOT be specified as an additional option.
```

```
# 43.37% 99659 bytes
#   -T mlwxyz_
#
# 16.09% 17615 bytes
#   -T half_
#
# 7.98% 12666 bytes
#   -T full_
#
# 6.82% 6846 bytes
#   -T artv_
#
# ...
#
# 1.29% 5352 bytes
#   -T currenh_
#
# 1.03% 25294 bytes
#   -T bndbo_
#
# Functions below this point account for less than 10% of samples.
#
# 1.03% 31240 bytes
#   -T bndto_
#
# 0.51% 11169 bytes
#   -T bncto_
#
# ...
#
# -----
#
# -o mhd3d.x+apa # New instrumented program.
#
# /work/crayadm/ldr/mhd3d/mhd3d.x # Original program.
```

Steps to collecting performance data

6. Look at <apafilename>.apa file
 - Verify if additional instrumentation is wanted
7. Instrument application for further analysis (a.out+apa)
 - % pat_build -O <apafilename>.apa
 - You should get an instrumented program a.out+apa
8. Run application
 - Remember to modify <script> to run a.out+apa
 - % aprun ... a.out+apa (or qsub <apa script>)
 - You should get a performance file (“<datafile>.xf”) or multiple files in a directory <datadir>
9. Create text report
 - % pat_report -o my_text_report.txt [<datafile>.xf | <datadir>]
 - Will generate a compressed performance file (<datafile>.ap2)
10. View results in text (my_text_report.txt) and/or with Cray Apprentice²
 - % app2 <datafile>.ap2

Using Craypat on large numbers of processors

- Pat_report can use an inordinate amount of time on the front-end system
 - Try submitting the pat_report as a batch job
 - Only give Pat_report a subset of the .xf files
 - Pat_report fms_cs_test13.x+apa+25430-12755tdt/*3.xf

Using Craypat MPI statistics

```

MPI Msg Bytes | MPI Msg | MsgSz | 16B<= | 256B<= | 4KB<= | Experiment=1
              | Count  |       | MsgSz  | MsgSz  | MsgSz  | Function
              |       | Count | <16B   | <256B  | <4KB   | Caller
              |       |       | Count  | Count  | Count  | PE[mmm]
3062457144.0 | 144952.0 | 15022.0 | 39.0 | 64522.0 | 65369.0 | Total
-----
| 3059984152.0 | 129926.0 | -- | 36.0 | 64522.0 | 65368.0 | mpi_isend_
|-----
|| 1727628971.0 | 63645.1 | -- | 4.0 | 31817.1 | 31824.0 | MPP_DO_UPDATE_R8_3DV.in.MPP_DOMAINS_MOD
3| | | | | | | | | MPP_UPDATE_DOMAIN2D_R8_3DV.in.MPP_DOMAINS_MOD
| | | |-----
4| | | 1680716892.0 | 61909.4 | -- | -- | 30949.4 | 30960.0 | DYN_CORE.in.DYN_CORE_MOD
5| | | | | | | | | FV_DYNAMICS.in.FV_DYNAMICS_MOD
6| | | | | | | | | ATMOSPHERE.in.ATMOSPHERE_MOD
7| | | | | | | | | MAIN__
8| | | | | | | | | main
| | | | | | | |-----
9| | | | | | | 1680756480.0 | 61920.0 | -- | -- | 30960.0 | 30960.0 | pe.13666
9| | | | | | | 1680756480.0 | 61920.0 | -- | -- | 30960.0 | 30960.0 | pe.8949
9| | | | | | | 1651777920.0 | 54180.0 | -- | -- | 23220.0 | 30960.0 | pe.12549
| | | | | | |=====

```

Memory allocation data from Craypat

Table 7: Heap Leaks during Main Program

Tracked MBytes Not Freed %	Tracked MBytes Not Freed	Tracked Objects Not Freed	Experiment=1 Caller PE[mmm]
100.0%	593.479	43673	Total
97.7%	579.580	43493	_F90_ALLOCATE
61.4%	364.394	106	SET_DOMAIN2D.in.MPP_DOMAINS_MOD
3			MPP_DEFINE_DOMAINS2D.in.MPP_DOMAINS_MOD
4			MPP_DEFINE_MOSAIC.in.MPP_DOMAINS_MOD
5			DOMAIN_DECOMP.in.FV_MP_MOD
6			RUN_SETUP.in.FV_CONTROL_MOD
7			FV_INIT.in.FV_CONTROL_MOD
8			ATMOSPHERE_INIT.in.ATMOSPHERE_MOD
9			ATMOS_MODEL_INIT.in.ATMOS_MODEL
10			MAIN__
11			main
12	0.0%	364.395	110 pe.43
12	0.0%	364.394	107 pe.8181
12	0.0%	364.391	88 pe.1047

Craypat load-imbalance data

Table 1: Profile by Function Group and Function

Time %	Time	Imb. Time	Imb. Time %	Calls	Experiment=1
					Group
					Function
					PE='HIDE'
100.0%	1061.141647	--	--	3454195.8	Total

70.7%	750.564025	--	--	280169.0	MPI_SYNC

45.3%	480.828018	163.575446	25.4%	14653.0	mpi_barrier_(sync)
18.4%	195.548030	33.071062	14.5%	257546.0	mpi_allreduce_(sync)
7.0%	74.187977	5.261545	6.6%	7970.0	mpi_bcast_(sync)
=====					
15.2%	161.166842	--	--	3174022.8	MPI

10.1%	106.808182	8.237162	7.2%	257546.0	mpi_allreduce_
3.2%	33.841961	342.085777	91.0%	755495.8	mpi_waitall_
=====					
14.1%	149.410781	--	--	4.0	USER

14.0%	148.048597	446.124165	75.1%	1.0	main
=====					

Hardware Counters

USER / MPP_DO_UPDATE_R8_3DV.in.MPP_DOMAINS_MOD

Time%		10.2%	
Time		49.386043	secs
Imb.Time		1.359548	secs
Imb.Time%		2.7%	
Calls	167.1 /sec	8176.0	calls
PAPI_L1_DCM	10.512M/sec	514376509	misses
PAPI_TLB_DM	2.104M/sec	102970863	misses
PAPI_L1_DCA	155.710M/sec	7619492785	refs
PAPI_FP_OPS		0	ops
User time (approx)	48.934 secs	112547914072 cycles	99.1%Time
Average Time per Call		0.006040	sec
CrayPat Overhead : Time	0.0%		
HW FP Ops / User time		0 ops	0.0%peak (DP)
HW FP Ops / WCT			
Computational intensity	0.00 ops/cycle	0.00	ops/ref
MFLOPS (aggregate)	0.00M/sec		
TLB utilization	74.00 refs/miss	0.145	avg uses
D1 cache hit,miss ratios	93.2% hits	6.8%	misses
D1 cache utilization (M)	14.81 refs/miss	1.852	avg uses

Table 2: Profile by Group, Function, and Line

Samp %	Samp	Imb. Samp	Imb. Samp %	Experiment=1
				Group
				Function
				Source
				Line
				PE='HIDE'
100.0%	103828	--	--	Total

48.9%	50784	--	--	USER

11.0%	11468	--	--	MPP_DO_UPDATE_R8_3DV.in.MPP_DOMAINS_MO
3				shared/mpp/include/mpp_do_updateV.h

2.9%	3056	238.53	7.2%	line.380
2.8%	2875	231.97	7.5%	line.967
2.0%	2071	310.19	13.0%	line.1028
=====				