/***
** Notes on Performance Application Programming Interface
**
** Intended audience: Those who would like to learn more about
** measuring program execution time in modern computer systems.
**
** Used: CPE 631 Advanced Computer Systems and Architectures
** CPE 619 Modeling and Analysis of Computer and Communication Systems
**
** ver 0.1, Spring 2005 PAPI-2.5.x
** ver 1.0, Spring 2006 PAPI-3.0.7
** ver 2.0, Spring 2007 PAPI-3.5.0 (changed interface)
** ver 3.0, Spring 2011 PAPI-4.1.0
** ver 4.0, Spring 2012 PAPI-5.3.0
** ver 5.0, Spring 2014 PAPI-5.3.0

** @Aleksandar Milenkovic, milenkovic@computer.org
***/

Performance Application Programming Interface

PAPI stands for Performance Application Programming Interface. It is a portable and efficient API (Application Programming Interface) to access hardware performance monitoring registers found on most modern microprocessors. The PAPI project is being developed by the ICL at the University of Tennessee.

1. To Learn More: http://icl.cs.utk.edu/papi/ (ultimate source of information)

2. Required reading:
User's Guide (http://icl.cs.utk.edu/projects/papi/files/documentation/PAPI_USER_GUIDE.pdf or http://icl.cs.utk.edu/projects/papi/files/documentation/PAPI_USER_GUIDE.htm)
This is a general overview of the PAPI library, with an introduction to the major feature areas, and comprehensive examples on usage.

3. Other documents (not required, but could be useful) can be found at:
http://icl.cs.utk.edu/projects/papi/files/documentation/

3.a. Programmer's Reference
(http://icl.cs.utk.edu/projects/papi/files/documentation/PAPI_Prog_Ref.pdf)
This is a technical reference (man pages) for PAPI programmers, providing detailed information on each call in the PAPI library, including calling sequences and examples.

4. To play with PAPI on Linux machine (Fedora 13), follow the steps below.

4.a. Login to your account.

4.b. Setup your account to have access to PAPI.

PAPI system variables should be in your path.
We will be using the papi 5.3.0 located in the /opt/papi-5.3.0 directory.

* Add path to papi utility programs and demo examples.
They are located in the papi bin directory and share directory.
To do that edit the .bashrc file in your home directory to include the following line.

```
export PATH=$PATH:./:/opt/papi-5.3.0/bin:/opt/papi-5.3.0/share/papi/ftests
```

* Add path to PAPI libraries.
To do that edit the .bashrc file in your home directory to include the following lines.

```
export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/opt/papi-5.3.0/lib
export LIBRARY_PATH=$LIBRARY_PATH:/opt/papi-5.3.0/lib
```

* Locate the PAPI include files in /opt/papi-5.3.0/include.

4.c. To see some predefined examples:

```
>cd /opt/papi-5.3.0/share/papi/ctests

>cd /opt/papi-5.3.0/share/papi/ftests
```

You can copy these directories and study examples that demonstrate the use of PAPI.

4.d. You can run some papi utilities

* To get the current version of the papi:

```
<<~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
[milenka@eb136i-nsf02 ~]$ papi_version
PAPI Version: 5.3.0.0
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~>>
```

* To get available events:

```
<<~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
[milenka@eb136i-nsf02 ~]$ papi_avail
Available events and hardware information.
--------------------------------------------------------------------------------
PAPI Version             : 5.3.0.0
Vendor string and code   : GenuineIntel (1)
Model string and code    : Intel(R) Xeon(R) CPU           X5570  @ 2.93GHz (26)
CPU Revision             : 5.000000
CPUID Info               : Family: 6  Model: 26  Stepping: 5
CPU Max Megahertz        : 2927
CPU Min Megahertz        : 1596
```

```
Hdw Threads per core     : 2
Cores per Socket         : 4
Sockets                  : 1
NUMA Nodes               : 1
CPUs per Node            : 8
Total CPUs               : 8
Running in a VM          : no
Number Hardware Counters : 7
Max Multiplex Counters   : 64
--------------------------------------------------------------------------------
```

The following correspond to fields in the PAPI_event_info_t structure.

```
    Name        Code    Avail Deriv Description (Note)
PAPI_L1_DCM  0x80000000  Yes   No   Level 1 data cache misses
PAPI_L1_ICM  0x80000001  Yes   No   Level 1 instruction cache misses
PAPI_L2_DCM  0x80000002  Yes   Yes  Level 2 data cache misses
PAPI_L2_ICM  0x80000003  Yes   No   Level 2 instruction cache misses
PAPI_L3_DCM  0x80000004  No    No   Level 3 data cache misses
PAPI_L3_ICM  0x80000005  No    No   Level 3 instruction cache misses
PAPI_L1_TCM  0x80000006  Yes   Yes  Level 1 cache misses
PAPI_L2_TCM  0x80000007  Yes   No   Level 2 cache misses
PAPI_L3_TCM  0x80000008  Yes   No   Level 3 cache misses
PAPI_CA_SNP  0x80000009  No    No   Requests for a snoop
PAPI_CA_SHR  0x8000000a  No    No   Requests for exclusive access to shared cache line
PAPI_CA_CLN  0x8000000b  No    No   Requests for exclusive access to clean cache line
PAPI_CA_INV  0x8000000c  No    No   Requests for cache line invalidation
PAPI_CA_ITV  0x8000000d  No    No   Requests for cache line intervention
PAPI_L3_LDM  0x8000000e  Yes   No   Level 3 load misses
PAPI_L3_STM  0x8000000f  No    No   Level 3 store misses
PAPI_BRU_IDL 0x80000010  No    No   Cycles branch units are idle
PAPI_FXU_IDL 0x80000011  No    No   Cycles integer units are idle
PAPI_FPU_IDL 0x80000012  No    No   Cycles floating point units are idle
PAPI_LSU_IDL 0x80000013  No    No   Cycles load/store units are idle
PAPI_TLB_DM  0x80000014  Yes   No   Data translation lookaside buffer misses
PAPI_TLB_IM  0x80000015  Yes   No   Instruction translation lookaside buffer misses
PAPI_TLB_TL  0x80000016  Yes   Yes  Total translation lookaside buffer misses
PAPI_L1_LDM  0x80000017  Yes   No   Level 1 load misses
PAPI_L1_STM  0x80000018  Yes   No   Level 1 store misses
PAPI_L2_LDM  0x80000019  Yes   No   Level 2 load misses
PAPI_L2_STM  0x8000001a  Yes   No   Level 2 store misses
PAPI_BTAC_M  0x8000001b  No    No   Branch target address cache misses
PAPI_PRF_DM  0x8000001c  No    No   Data prefetch cache misses
PAPI_L3_DCH  0x8000001d  No    No   Level 3 data cache hits
PAPI_TLB_SD  0x8000001e  No    No   Translation lookaside buffer shootdowns
PAPI_CSR_FAL 0x8000001f  No    No   Failed store conditional instructions
PAPI_CSR_SUC 0x80000020  No    No   Successful store conditional instructions
PAPI_CSR_TOT 0x80000021  No    No   Total store conditional instructions
PAPI_MEM_SCY 0x80000022  No    No   Cycles Stalled Waiting for memory accesses
PAPI_MEM_RCY 0x80000023  No    No   Cycles Stalled Waiting for memory Reads
PAPI_MEM_WCY 0x80000024  No    No   Cycles Stalled Waiting for memory writes
PAPI_STL_ICY 0x80000025  No    No   Cycles with no instruction issue
PAPI_FUL_ICY 0x80000026  No    No   Cycles with maximum instruction issue
PAPI_STL_CCY 0x80000027  No    No   Cycles with no instructions completed
PAPI_FUL_CCY 0x80000028  No    No   Cycles with maximum instructions completed
PAPI_HW_INT  0x80000029  No    No   Hardware interrupts
PAPI_BR_UCN  0x8000002a  Yes   No   Unconditional branch instructions
PAPI_BR_CN   0x8000002b  Yes   No   Conditional branch instructions
PAPI_BR_TKN  0x8000002c  Yes   No   Conditional branch instructions taken
PAPI_BR_NTK  0x8000002d  Yes   Yes  Conditional branch instructions not taken
PAPI_BR_MSP  0x8000002e  Yes   No   Conditional branch instructions mispredicted
```

```
PAPI_BR_PRC  0x8000002f  Yes   Yes  Conditional branch instructions correctly
predicted
PAPI_FMA_INS 0x80000030  No    No   FMA instructions completed
PAPI_TOT_IIS 0x80000031  Yes   No   Instructions issued
PAPI_TOT_INS 0x80000032  Yes   No   Instructions completed
PAPI_INT_INS 0x80000033  No    No   Integer instructions
PAPI_FP_INS  0x80000034  Yes   No   Floating point instructions
PAPI_LD_INS  0x80000035  Yes   No   Load instructions
PAPI_SR_INS  0x80000036  Yes   No   Store instructions
PAPI_BR_INS  0x80000037  Yes   No   Branch instructions
PAPI_VEC_INS 0x80000038  No    No   Vector/SIMD instructions (could include integer)
PAPI_RES_STL 0x80000039  Yes   No   Cycles stalled on any resource
PAPI_FP_STAL 0x8000003a  No    No   Cycles the FP unit(s) are stalled
PAPI_TOT_CYC 0x8000003b  Yes   No   Total cycles
PAPI_LST_INS 0x8000003c  Yes   Yes  Load/store instructions completed
PAPI_SYC_INS 0x8000003d  No    No   Synchronization instructions completed
PAPI_L1_DCH  0x8000003e  Yes   Yes  Level 1 data cache hits
PAPI_L2_DCH  0x8000003f  Yes   Yes  Level 2 data cache hits
PAPI_L1_DCA  0x80000040  Yes   No   Level 1 data cache accesses
PAPI_L2_DCA  0x80000041  Yes   No   Level 2 data cache accesses
PAPI_L3_DCA  0x80000042  Yes   Yes  Level 3 data cache accesses
PAPI_L1_DCR  0x80000043  Yes   No   Level 1 data cache reads
PAPI_L2_DCR  0x80000044  Yes   No   Level 2 data cache reads
PAPI_L3_DCR  0x80000045  Yes   No   Level 3 data cache reads
PAPI_L1_DCW  0x80000046  Yes   No   Level 1 data cache writes
PAPI_L2_DCW  0x80000047  Yes   No   Level 2 data cache writes
PAPI_L3_DCW  0x80000048  Yes   No   Level 3 data cache writes
PAPI_L1_ICH  0x80000049  Yes   No   Level 1 instruction cache hits
PAPI_L2_ICH  0x8000004a  Yes   No   Level 2 instruction cache hits
PAPI_L3_ICH  0x8000004b  No    No   Level 3 instruction cache hits
PAPI_L1_ICA  0x8000004c  Yes   No   Level 1 instruction cache accesses
PAPI_L2_ICA  0x8000004d  Yes   No   Level 2 instruction cache accesses
PAPI_L3_ICA  0x8000004e  Yes   No   Level 3 instruction cache accesses
PAPI_L1_ICR  0x8000004f  Yes   No   Level 1 instruction cache reads
PAPI_L2_ICR  0x80000050  Yes   No   Level 2 instruction cache reads
PAPI_L3_ICR  0x80000051  Yes   No   Level 3 instruction cache reads
PAPI_L1_ICW  0x80000052  No    No   Level 1 instruction cache writes
PAPI_L2_ICW  0x80000053  No    No   Level 2 instruction cache writes
PAPI_L3_ICW  0x80000054  No    No   Level 3 instruction cache writes
PAPI_L1_TCH  0x80000055  No    No   Level 1 total cache hits
PAPI_L2_TCH  0x80000056  Yes   Yes  Level 2 total cache hits
PAPI_L3_TCH  0x80000057  No    No   Level 3 total cache hits
PAPI_L1_TCA  0x80000058  Yes   Yes  Level 1 total cache accesses
PAPI_L2_TCA  0x80000059  Yes   No   Level 2 total cache accesses
PAPI_L3_TCA  0x8000005a  Yes   No   Level 3 total cache accesses
PAPI_L1_TCR  0x8000005b  Yes   Yes  Level 1 total cache reads
PAPI_L2_TCR  0x8000005c  Yes   Yes  Level 2 total cache reads
PAPI_L3_TCR  0x8000005d  Yes   Yes  Level 3 total cache reads
PAPI_L1_TCW  0x8000005e  No    No   Level 1 total cache writes
PAPI_L2_TCW  0x8000005f  Yes   No   Level 2 total cache writes
PAPI_L3_TCW  0x80000060  Yes   No   Level 3 total cache writes
PAPI_FML_INS 0x80000061  No    No   Floating point multiply instructions
PAPI_FAD_INS 0x80000062  No    No   Floating point add instructions
PAPI_FDV_INS 0x80000063  No    No   Floating point divide instructions
PAPI_FSQ_INS 0x80000064  No    No   Floating point square root instructions
PAPI_FNV_INS 0x80000065  No    No   Floating point inverse instructions
PAPI_FP_OPS  0x80000066  Yes   Yes  Floating point operations
PAPI_SP_OPS  0x80000067  Yes   Yes  Floating point operations; optimized to count
scaled single precision vector operations
PAPI_DP_OPS  0x80000068  Yes   Yes  Floating point operations; optimized to count
scaled double precision vector operations
PAPI_VEC_SP  0x80000069  Yes   No   Single precision vector/SIMD instructions
PAPI_VEC_DP  0x8000006a  Yes   No   Double precision vector/SIMD instructions
```

```
PAPI_REF_CYC 0x8000006b  Yes   No   Reference clock cycles
--------------------------------------------------------------------------
Of 108 possible events, 64 are available, of which 17 are derived.

avail.c                                     PASSED
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~>>
```

4.f. You can run papi tests from ctest and ftest directories. E.g.,

```
<<~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
[milenka@eb136i-nsf02 ctests]$ first
Test case 1: Non-overlapping start, stop, read.
-----------------------------------------------
Default domain is: 1 (PAPI_DOM_USER)
Default granularity is: 1 (PAPI_GRN_THR)
Using 20000000 iterations of c += a*b
--------------------------------------------------------------------------
Test type   :        1            2            3            4            5
PAPI_FP_INS:     40000211     40000180     80000362    120000590    120000590
PAPI_TOT_CYC:    200078129    203678021    405324079    606527023    606527023
--------------------------------------------------------------------------
Verification:
Row 1 Column 1 at least 20000000
% difference between PAPI_FP_INS: 1 & 2: 100.00
% difference between PAPI_TOT_CYC 1 & 2: 98.23
Column 1 approximately equals column 2
Column 3 approximately equals 2 * column 2
Column 4 approximately equals 3 * column 2
Column 4 exactly equals column 5
first.c                                     PASSED
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~>>
```

* To see the source code for this example read first.c, e.g.

```
<<~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
[milenka@eb136i-nsf02 ~]$ more /opt/papi-5.3.0/share/papi/ctests/first.c
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~>>
```

4.g. To see events supported by your substrate run.

```
<<~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
[milenka@eb136i-nsf02 ~]$ papi_native_avail | more
Available native events and hardware information.
---------------------------------------------------------------------------------
PAPI Version             : 5.3.0.0
Vendor string and code   : GenuineIntel (1)
Model string and code    : Intel(R) Xeon(R) CPU           X5570  @ 2.93GHz (26)
CPU Revision             : 5.000000
CPUID Info               : Family: 6  Model: 26  Stepping: 5
CPU Max Megahertz        : 2927
CPU Min Megahertz        : 1596
Hdw Threads per core     : 2
Cores per Socket         : 4
Sockets                  : 1
NUMA Nodes               : 1
CPUs per Node            : 8
Total CPUs               : 8
Running in a VM          : no
Number Hardware Counters : 7
Max Multiplex Counters   : 64
```

```
--------------------------------------------------------------------------------

===============================================================================
 Native Events in Component: perf_event
===============================================================================
| UNHALTED_CORE_CYCLES                                                         |
|            Count core clock cycles whenever the clock signal on the specific |
|            core is running (not halted)                                      |
|     :e=0                                                                     |
|            edge level (may require counter-mask >= 1)                        |
|     :i=0                                                                     |
|            invert                                                            |
|     :c=0                                                                     |
|            counter-mask in range [0-255]                                     |
|     :t=0                                                                     |
|            measure any thread                                                |
|     :u=0                                                                     |
|            monitor at user level                                             |
|     :k=0                                                                     |
|            monitor at kernel level                                           |
--------------------------------------------------------------------------------
| INSTRUCTION_RETIRED                                                          |
|            Count the number of instructions at retirement                    |
|     :e=0                                                                     |
|            edge level (may require counter-mask >= 1)                        |
|     :i=0                                                                     |
|            invert                                                            |
|     :c=0                                                                     |
|            counter-mask in range [0-255]                                     |
|     :t=0                                                                     |
|            measure any thread                                                |
|     :u=0                                                                     |
|            monitor at user level                                             |
|     :k=0                                                                     |
|            monitor at kernel level                                           |
--More--

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~>>
```

5. To help you get started I prepared a small example

5.a. Copy the example from /opt/arch.tut/test.papi

```
<<~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
[milenka@eb136i-nsf02 test.papi]$ cp -R /opt/arch.tut/test.papi .
[milenka@eb136i-nsf02 test.papi]$ cd test.papi
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~>>
```

5.b. Examine Makefile

```
<<~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
[milenka@eb136i-nsf02 test.papi]$ cat Makefile.test.papi
arrsum_papi: arrsum_papi.o
        gcc    -lpapi  arrsum_papi.o  -o arrsum_papi

arrsum_papi.o: arrsum_papi.c
        gcc -I/opt/papi-5.3.0/include -O0 -c arrsum_papi.c

clean:
        rm *.o
        rm arrsum_papi
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~>>
```

* If you want to learn more about GNU make visit
  http://www.gnu.org/software/make/manual/make.html

5.c. Examine the arrsum_papi.c (sums up elements of an integer array).
Observe PAPI events, use of PAPI functions, etc.
What metrics do we use in reporting performance?
How do we calculate them?

```
<<~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
[milenka@eb136i-nsf02 test.papi]$ joe arrsum_papi.c
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~>>
```

5.d. Compile the program.

```
<<~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
[milenka@eb136i-nsf02 test.papi]$ make -f Makefile.test.papi
gcc    -lpapi  arrsum_papi.o  -o arrsum_papi

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~>>
```

5.e. Run the program and study the result (play with different sizes)

```
<<~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
[milenka@eb136i-nsf02 test.papi]$ ./arrsum_papi 16384
17488290750180892l6.000000
TOT_INS for summing up:       163847
TOT_CYC for summing up:       164051
Caclulated CPI: 1.001247

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~>>
```

You may also calculate the total time needed to sum up elements of this array (164051/2.93exp(9))=55 us.

5.f. If you want to see the assembly code for this example type in the following:

```
<<~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
[milenka@eb136i-nsf02 test.papi]$ gcc -S -O0 arrsum_papi.c -lpapi
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~>>
```

arrsum_papi.s file will be generated.

5.g. To probe further:
* Add other events in test program (cycles, cache misses, etc).
* Change compiler optimization (in Makefile replace -O0 with O3)
* Play with other compilers (e.g., Intel's icc).
* Make your own examples

6. To learn more about
* gcc go to: http://gcc.gnu.org/onlinedocs/gcc/
* GNU make tools: http://www.gnu.org/software/make/manual/make.html