A UT/ORNL PARTNERSHIP
NATIONAL INSTITUTE FOR COMPUTATIONAL SCIENCES
NICS

Automatic Library Tracking Database for High Performance Computing Systems

Bilel Hadri and Mark Fahey

University of Tennessee
NICS

Kraken

	Kraken
Compute processor type	AMD 2.6 GHz Istanbul -6
Compute cores	112,896
Compute nodes	9,408
Theoretical Peak Performance	1.17 PFLOPS
LinPACK Performance	919 TFLOPS
Total memory	147 TB
Compilers	**4 (PGI, Cray, GNU, Intel)**
Software/library installed by vendor & support staff	**136**

NICS

Software/library/applications

- Several categories of software/library
  - Linear algebra
  - I/O
  - Performance tools
  - Debugger
  - Chemistry
  - Molecular dynamic
  - Materials
  - Communications
  - Visualization
- Multiple versions
  - hdf5 (1.6.10 - 1.8.3 - 1.8.4 - 1.8.5 - 1.8.6)
  - netcdf (3.6.2 - 3.6.3 - 4.0 - 4.0.1 - 4.1 - 4.1.1)
  - libsci (10.3.9/ 10.4.1/ 10.4.2/ 10.4.4/ 10.4.5/ 10.4.9/ 10.5.0/ 10.5.01/ 10.5.02/ 11.0.00/ 11.0.01/)
- Multiple builds with different compiler:

  - example with HYPRE: cnl2.2_cray7.3.3/ cnl2.2_gnu4.4.4/ cnl2.2_gnu4.5.2/ cnl2.2_intel11.1.038/ cnl2.2_pgi10.6.0/ cnl2.2_pgi11.4.0/

→Kraken : Close to 800 combinations of software/library with different versions and builds with different compilers just in primary installation tree

NICS

Issue !

- How do HPC centers monitor/measure software usage and forecast needs?

- How do
  - we know when to change defaults? (to newer versions)
  - we know when we can get rid of old versions ?
  - we know which software is not used ? (reduce cost)
  - we find who is using
    - deprecated software?
    - non-optimal [math] libraries?
    - software with bugs?
    - software funded by NSF/DOE?

- As of today:
  - Rule of the thumbs from the staff:
    - not strictly accurate and reliable
  - Surveys:
    - Incomplete data

→ Solution ALTD: Automatic Library Tracking Database

NICS

Objectives and Goals

- A primary objective of ALTD :
  - track only libraries linked into the applications (not the function calls)
  - track parallel executables launched (how often are the libraries used?)

- Have as little impact on user as possible
  - Lightweight solution
    - No runtime increase
    - Only link time and job launch have marginal increase in time
  - Do not change user experience
    - Linker and job launcher work as expected

- Intercept the whole library path to retrieve valuable information on :
  - Package name
  - Version number
  - Build configuration

NICS

ALTD design

- Intercepting the GNU linker (ld) to get the linkage information

- Intercepting the job launcher (aprun)

→ Wrapping the linker and the job launcher through scripts is a simple and efficient way to obtain the information automatically and transparently.

- ld - Intercept link line
  - Update tags table
  - Call real linker (with tracemap option)
  - Use output from tracemap to find libraries linked into executable
  - Update linkline table

- aprun- Intercept job launcher
  - Pull information from ALTD section header in executable
  - Update jobs table
  - Call real job launcher

→ Storing information about compilation and execution into a database that can be mined to provide reports.

NICS

ALTD database results

linking_inc	linkline
14437	../bin/cg.B.4 /usr/lib/../lib64/crt1.o /usr/lib/../lib64/crti.o
	/opt/gcc/4.4.2/snos/lib/gcc/x86_64-suse-linux/4.4.2/crtbeginT.o
	/sw/xt/tau/2.19/cnl2.2_gnu4.4.1/tau-2.19/craycnl/lib/libTauMpi-gnu-mpi-pdt.a
	/sw/xt/tau/2.19/cnl2.2_gnu4.4.1/tau-2.19/craycnl/lib/libtau-gnu-mpi-pdt.a /usr/lib/../lib64/libpthread.a
	/opt/cray/mpt/4.0.1/xt/seastar/mpich2-gnu/lib/libmpich.a /opt/cray/pmi/1.0-1.0000.7628.10.2.ss/lib64/libpmi.a
	/usr/lib/alps/libalpslli.a /usr/lib/alps/libalpsutil.a /opt/xt-pe/2.2.41A/lib/snos64/libportals.a
	/opt/gcc/4.4.2/snos/lib/gcc/x86_64-suse-linux/4.4.2/libgfortranbegin.a
	/opt/gcc/4.4.2/snos/lib/gcc/x86_64-suse-linux/4.4.2/libgcc.a
	/opt/gcc/4.4.2/snos/lib/gcc/x86_64-suse-linux/4.4.2/libgcc_eh.a /usr/lib/../lib64/libc.a
	/opt/gcc/4.4.2/snos/lib/gcc/x86_64-suse-linux/4.4.2/crtend.o /usr/lib/../lib64/crtn.o
14438	highmass3d.Linux.CC.ex /usr/lib64/crt1.o /usr/lib64/crti.o /opt/pgi/9.0.4/linux86-64/9.0-4/lib/trace_init.o
	/usr/lib64/gcc/x86_64-suse-linux/4.1.2/crtbeginT.o /sw/xt/hypre/2.0.0/cnl2.2_pgi9.0.1/lib//libHYPRE.a
	/opt/cray/pmi/1.0-1.0000.7628.10.2.ss/lib64/libpmi.a /usr/lib/alps/libalpslli.a /usr/lib/alps/libalpsutil.a
	/opt/xt-pe/2.2.41A/lib/snos64/libportals.a /usr/lib64/libpthread.a /usr/lib64/libm.a
	/usr/local/lib/libmpich.a /opt/pgi/9.0.4/linux86-64/9.0-4/lib/libstd.a /opt/pgi/9.0.4/linux86-64/9.0-4/lib/libC.a
	/opt/pgi/9.0.4/linux86-64/9.0-4/lib/libpgf90.a /opt/pgi/9.0.4/linux86-64/9.0-4/lib/libpgc.a
	/usr/lib64/librt.a /usr/lib64/libpthread.a /usr/lib64/libm.a /usr/lib64/gcc/x86_64-suse-linux/4.1.2/libgcc_eh.a
	/usr/lib64/libc.a /usr/lib64/gcc/x86_64-suse-linux/4.1.2/crtend.o /usr/lib64/crtn.o
14439	probeTest /usr/lib/../lib64/crt1.o /usr/lib/../lib64/crti.o
	/opt/gcc/4.4.2/snos/lib/gcc/x86_64-suse-linux/4.4.2/crtbeginT.o
	/opt/cray/mpt/4.0.1/xt/seastar/mpich2-gnu/lib/libmpich.a /opt/cray/pmi/1.0-1.0000.7628.10.2.ss/lib64/libpmi.a
	/usr/lib/alps/libalpslli.a /usr/lib/alps/libalpsutil.a /opt/xt-pe/2.2.41A/lib/snos64/libportals.a
	/usr/lib/../lib64/libpthread.a /opt/gcc/4.4.2/snos/lib/gcc/x86_64-suse-linux/4.4.2/libgcc_eh.a
	/usr/lib/../lib64/libc.a /opt/gcc/4.4.2/snos/lib/gcc/x86_64-suse-linux/4.4.2/crtend.o /usr/lib/../lib64/crtn.o

a) linkline table

tag_id	linkline_id	username	exit_code	link_date
91126	14437	user1	0	2010-04-28
91127	0	user2	-1	2010-04-28
91128	14435	user3	0	2010-04-28
91129	6835	user2	0	2010-04-28
91130	14438	user4	0	2010-04-28
91131	14439	user1	0	2010-04-28
91132	14439	user1	0	2010-04-28

b) tag_id table

run_inc	tag_id	executable	username	run_date	job_launch_id	build_machine
144091	91126	/nics/b/home/user1/NPB3.3/bin/cg.B.4	user1	2010-04-28	548346	kraken
144099	91131	/nics/b/home/user1/probeTest	user1	2010-04-28	548357	kraken
144102	91132	/nics/b/home/user1/probeTest	user1	2010-04-28	548357	kraken
144179	91128	/lustre/scratch/user3/CH4/vasp_vtst.x	user3	2010-04-28	548444	kraken
144192	91128	/lustre/scratch/user3/CH4/vasp_vtst.x	user3	2010-04-28	548488	kraken
144356	91128	/lustre/scratch/user5/src/CH4/vasp_vtst.x	user5	2010-04-29	548638	kraken

c) job_id table

- **ALTD generates records into 3 tables:**
  - **Tags: entry for every link executed**
  - **Linkline: entry for each unique link line**
  - **Jobs: entry for each executable launched**

NICS

Linktable

linkline_id	linkline
14437	**../bin/cg.B.4** /usr/lib/../lib64/crt1.o /usr/lib/../lib64/crti.o /opt/gcc/4.4.2/snos/lib/gcc/x86_64-suse-linux/4.4.2/crtbeginT.o **/sw/xt/tau/2.19/cnl2.2_gnu4.4.1/tau-2.19/craycnl/lib/libTauMpi-gnu-mpi-pdt.a** **/sw/xt/tau/2.19/cnl2.2_gnu4.4.1/tau-2.19/craycnl/lib/libtau-gnu-mpi-pdt.a** /usr/lib/../lib64/libpthread.a /opt/cray/mpt/4.0.1/xt/seastar/mpich2-gnu/lib/libmpich.a /opt/cray/pmi/1.0-1.0000.7628.10.2.ss/lib64/libpmi.a /usr/lib/alps/libalpslli.a /usr/lib/alps/libalpsutil.a /opt/xt-pe/2.2.41A/lib/snos64/libportals.a **[…. gcc 4.4.2 libraries …]** /usr/lib/../lib64/libc.a /usr/lib/../lib64/crtn.o
14438	**highmass3d.Linux.CC.ex** /usr/lib64/crt1.o /usr/lib64/crti.o /opt/pgi/9.0.4/linux86-64/9.0-4/lib/trace_init.o /usr/lib64/gcc/x86_64-suse-linux/4.1.2/crtbeginT.o **/sw/xt/hypre/2.0.0/cnl2.2_pgi9.0.1/lib//libHYPRE.a** /opt/cray/pmi/1.0-1.0000.7628.10.2.ss/lib64/libpmi.a /usr/lib/alps/libalpslli.a /usr/lib/alps/libalpsutil.a /opt/xt-pe/2.2.41A/lib/snos64/libportals.a /usr/lib64/libpthread.a /usr/lib64/libm.a /usr/local/lib/libmpich.a **[… pgi 9.0.4 libraries …]** /usr/lib64/librt.a /usr/lib64/libpthread.a /usr/lib64/libm.a /usr/lib64/gcc/x86_64-suse-linux/4.1.2/libgcc_eh.a /usr/lib64/libc.a /usr/lib64/gcc/x86_64-suse-linux/4.1.2/crtend.o /usr/lib64/crtn.o
14439	**probeTest** /usr/lib/../lib64/crt1.o /usr/lib/../lib64/crti.o /opt/gcc/4.4.2/snos/lib/gcc/x86_64-suse-linux/4.4.2/crtbeginT.o /opt/cray/mpt/4.0.1/xt/seastar/mpich2-gnu/lib/libmpich.a /opt/cray/pmi/1.0-1.0000.7628.10.2.ss/lib64/libpmi.a /usr/lib/alps/libalpslli.a /usr/lib/alps/libalpsutil.a /opt/xt-pe/2.2.41A/lib/snos64/libportals.a /usr/lib/../lib64/libpthread.a **[… gcc 4.4.2 libraries …]** /usr/lib/../lib64/libc.a /usr/lib/../lib64/crtn.o

NICS

Job table

run_inc	tag_id	executable	username	run_date	Job_launch_id	build_machine
144091	91126	/nics/b/home/user1/NPB3.3/bin/cg.B.4	user1	2010-04-28	548346	kraken
144099	91131	/nics/b/home/user1/probeTest	user1	2010-04-28	548357	kraken
144102	91132	/nics/b/home/user1/probeTest	user1	2010-04-28	548357	kraken
144179	91128	/lustre/scratch/user3/CH4/vasp_vtst.x	user3	2010-04-28	548444	kraken
144192	91128	/lustre/scratch/user3/CH4/vasp_vtst.x	user3	2010-04-28	548488	kraken
144356	91128	/lustre/scratch/user5/CH4/vasp_vtst.x	user5	2010-04-28	548638	kraken

NICS

Reports : libraries used during linking

Libraries installed in /opt

Rank	Jaguar 2009	Jaguar 2010
1	hdf5	fftw
2	*craypat*	Hdf5
3	*papi*	petsc
4	petsc	*papi*
5	libsci	netcdf
6	netcdf	acml
7	acml	libsci
8	fftw	*craypat*
9	gromacs	*tau*
10	trilinos	trilinios

Libraries installed in /sw/xt/

Rank	Jaguar 2009	Jaguar 2010	Kraken 2010
1	szip/ 2.1	*tau/ 2.19*	sprng/ 2.0b
2	hdf5/ 1.6.8	szip/ 2.1	petsc/ 2.3.3
3	trilinos/ 9.0.2	hdf5/ 1.8.1	iobuf/ beta
4	pspline/ 1.0	hdf5/ 1.6.8	*tau/ 2.19*
5	netcdf/ 3.6.2	trilinos/1 0.0.4	szip/ 2.1
6	gromacs /4.0.5	*vampirtra ce/5.8*	p-netcdf/ 1.1.1
7	parmeti s/3.1	hdf5 /1.6.7	ncl/ 5.0.0
8	petsc/ 3.0.0	Adios /1.1.0	atlas/ 3.8.3
9	hdf5/ 1.6.7	*fpmpi /1.1*	upc/ 2.8.0
10	hdf5/ 1.8.2	p-netcdf/ 1.0.3	hdf5/1.8.3

Overall libraries

NICS

Reports : Usage during execution

Rank	Jaguar 2009	Jaguar 2010	Kraken 2010
1	nw_para	nw_para	interpo
2	ior	vbc1_7	namd
3	vbc1_4	vasp	chimera
4	vbc1_3	amber	amber
5	vasp	namd	mpiblast
6	visit	pltar	enzo
7	namd	chimera	espresso
8	espresso	espresso	lammps
9	spdcp	visit	vasp
10	lammps	gromacs	gromacs

Libraries installed in /sw/xt/

Rank	Jaguar 2009	Jaguar 2010	Kraken 2010
1	spdcp/ 0.3.6	vasp/ 4.6_r61	namd/ 2.7b1
2	namd/ 2.6	pltar/ 0.9.0	amber/ 10
3	vasp/ 4.6_r60	namd/ 2.6	lamps/ mar09
4	cpmd/ 3.13.2	namd/ 2.7b1	gromacs/4.0.7
5	namd/ 2.7b1	spdcp/ 0.3.9	lamps/ oct09

NICS

Conclusions

- ALTD tracks automatically and transparently library usage at compilation and at execution
  - Wrapping the linker and the job launcher
  - In production on several Cray XT machines at NICS and OLCF (ORNL)

- Track the most used libraries and it facilitates decisions for removing old/non-used libraries

- Data mining:
  - Usage at linking:
    - Linear algebra, I/O and Performance tools
  - Usage at execution
    - Molecular dynamic (NAMD and AMBER), climate modeling

- Alpha version is available if interested, please contact us !

NICS

Future Work

- Porting ALTD to additional HPC architectures
  - Top 500 (Cray is the third ranked vendor with 29 systems; however, Cray machines represent only 5.8% of all the TOP500 machines)
  - challenge: job launcher (mpirun, ibrun, …) and batch systems (SLURM, LSF…)
- Determining the usage of libraries and executables by a project
- Considering other metrics for the usage:
  - Rankings of "most used" executables based on CPU hours.
- Developing a web-interface to dynamically and easily show different reports.
- Building a HPC inventory

NICS

Contact

Bilel Hadri
University of Tennessee, NICS
bhadri@utk.edu

Mark Fahey
University of Tennessee, NICS
mfahey@utk.edu