Benchmarking parallel DHT storage Techniques (Phase 2)

Abha Chaudhary, Chenfei Lin, Adil Ansari

December 8, 2013

1 Objective

Our objective is to study and compare different implementations of Distributed Hash Table and benchmark the space/time performance of these across different combination of loads and sizes.

2 APIs

2.1 C-MPI

- We are able to setup DHT and do a basic communication across nodes.

```
sleeping...
[D 00:17:52.635313 R:4] client: 4
[4] cmpi_params_get(key=nodes,result=4)...
[4] contact: 0
[4] contact: 1
[4] contact: 2
[4] contact: 3
[4] cmpi_client_code()...
sleeping...
[D 00:17:52.635531 R:0] node: 0
[0] cmpi_params_get(key=nodes,result=4)...
[0] DENSE_Init()...
[0] rpc_bootping(1)...
[0] MPIRPC_Block(bootping)...
[0] MPIRPC_Block_blob(bootping)...
[0] MPIRPC_Call_blob(target.rank=1,name=bootping,args=(null),rpc->unique=1)...
[3] cmpi_params_get(key=tag,result=(null))...
[D 00:17:52.635968 R:1] node: 1
[1] cmpi_params_get(key=nodes,result=4)...
[1] DENSE_Init()...
[1] listen_loop()...
[3] I am node
node: 3 @ roxy
[1] MPIRPC_Request(name=bootping,caller.rank=0)...
[1] MPIRPC_Return(caller.rank=0,unique=1,length=-1)...
[1] rpc_bootping(2)...
[1] MPIRPC_Block(bootping)...
[1] MPIRPC_Block_blob(bootping)...
[1] MPIRPC_Call_blob(target.rank=2,name=bootping,args=(null),rpc->unique=1)...
[0] MPIRPC_Retrieve(sender.rank=1,unique=1)...
[0] Unblocked: bootping
[0] MPIRPC_Free(rpc->name=bootping,rpc->unique=1)...
[0] listen_loop()...
```

Figure 1: Startup test

- Were able to set up the DHT on the cluster upto a 70 nodes. So total of 140 processes running on 70 nodes wth 70 other communicating clients. Following is the time taken to start, set up the dht, do the initial communication and shutdown the system:

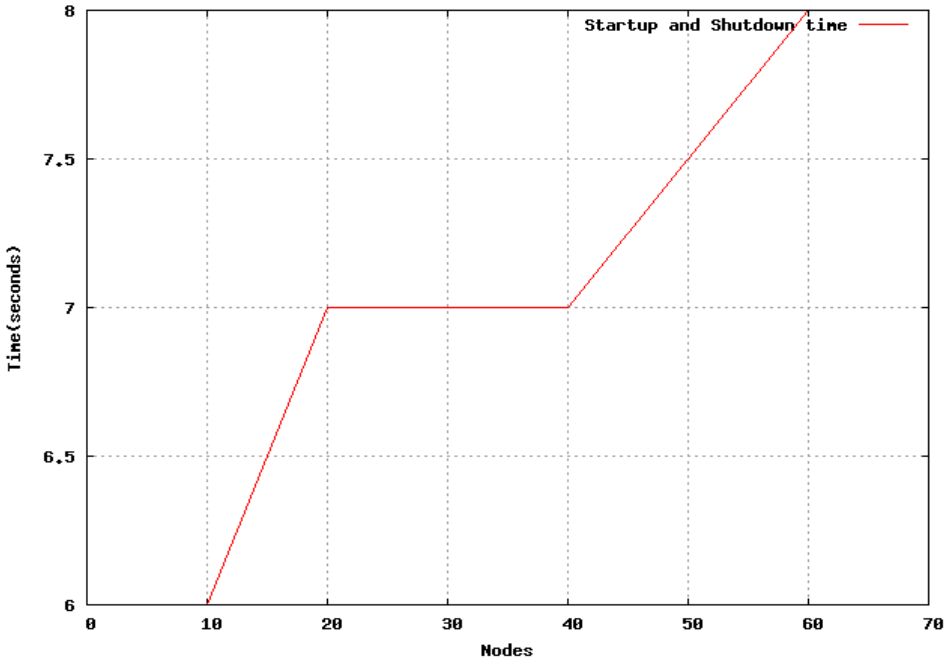

Figure 2: BambooDHT menu

2.2 BamboDHT

- Use a 32-bit JDK version and set JAVAHOME
- 'ls' into the directory
- 'make'
- 'ls' into bamboo/test/
- 'perl -w location-test-menu.pl'
- menu-[4] start a node

```
Menu:
    1.  Check node status
    2.  Check object pointers
    3.  Check for exceptions
    4.  Start a node
    5.  Stop a node
    6.  Quit
Your choice: [2]  1

/tmp/experiment-12550-localhost-3630.log:Tapestry: ready
/tmp/experiment-12550-localhost-3633.log:Tapestry: ready

Menu:
    1.  Check node status
    2.  Check object pointers
    3.  Check for exceptions
    4.  Start a node
    5.  Stop a node
    6.  Quit
Your choice: [1]
```

Figure 3: BambooDHT menu

- *Error* compiling with higher level of JDK, and running with lower version of JRE
- Some **major issues** are encountered because the software is built on **Staged Event-Driven Architecture** which uses 32-bit word length and the program has major compatility problems with 64-bit architectures.

```
Your choice: [2]  2

127.0.0.1:3630, 0x19d5e887 is storing
  0x35e9f1c7 0x35e9f1c7
  0x09eec5f6 0x09eec5f6
  0x2f747891 0x2f747891
  0x1030f3d2 0x1030f3d2
  0x1a1e9f2a 0x1a1e9f2a
  0x5189c6e1 0x5189c6e1
  0x1a36c07d 0x1a36c07d
  0xd99fdf9c 0xd99fdf9c
  0x7d6da37d 0x7d6da37d

127.0.0.1:3633, 0xc7663e70 is storing
  0x35e9f1c7 0x35e9f1c7
  0x09eec5f6 0x09eec5f6
  0x2f747891 0x2f747891
  0x1030f3d2 0x1030f3d2
  0x1a1e9f2a 0x1a1e9f2a
  0x5189c6e1 0x5189c6e1
  0x1a36c07d 0x1a36c07d
  0xd99fdf9c 0xd99fdf9c
  0x7d6da37d 0x7d6da37d

0x09eec5f6 0x09eec5f6 2
0x1030f3d2 0x1030f3d2 2
0x1a1e9f2a 0x1a1e9f2a 2
0x1a36c07d 0x1a36c07d 2
0x2f747891 0x2f747891 2
0x35e9f1c7 0x35e9f1c7 2
0x5189c6e1 0x5189c6e1 2
0x7d6da37d 0x7d6da37d 2
0xd99fdf9c 0xd99fdf9c 2

All counts okay.
```

Figure 4: Communication in bamboo DHT

2.3 Chimera

2.4 Steps to install

- **URL** *http://current.cs.ucsb.edu/projects/chimera*
- ls into chimera directory
- ./configure –prefix='curent directory'
- make
- ls into /test
- sha1_keygen [*number_of_keys*] [*hostname*] will generate a set of pairs of ports and keys
- ./test [*port*] [*key*], where post is the port and key is from pair generated above. You can then send messages.

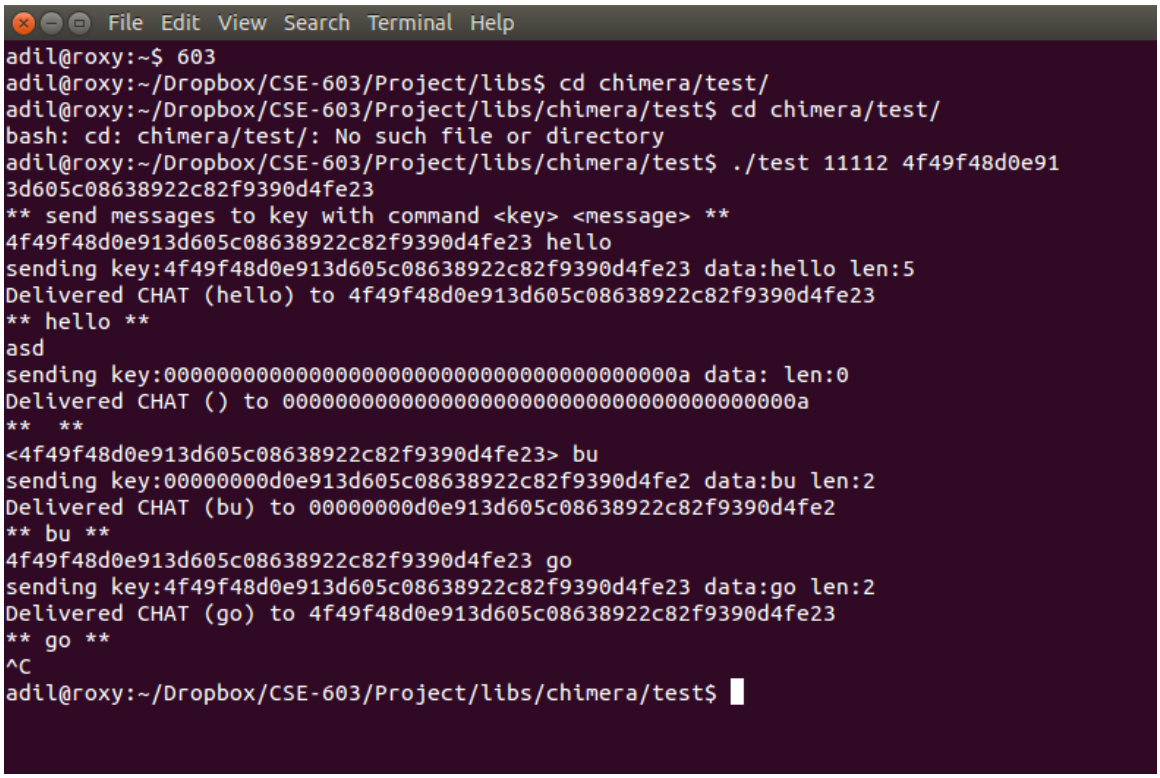

Figure 5: Chimera communication model

3 Summary

- **C-MPI :** We are able to execute it on a cluster of 80 nodes and get the statistics for starting all the nodes, database setup, basic communication and shutdown.
- **Bamboo :** Build and run simple test.
- **Chimera :** Executed it on a single node but were not able to do on cluster.

4 Future work

- Scale c-mpi to hundreds of nodes.
- Try to get some more put and get test cases of c-mpi running.
- Deploy bamboo DHT on cluster