Tuning the OpenAFS UNIX client cache manager

Michael Garrison
University of Michigan
ITCS/UMCE
University of Michigan Machine Configurations

- GigE Network
- 4Gb memory
- Dual 3.0Ghz Xeon
University of Michigan
Client/Server Setup

- 1.4.0 Clients and Servers
- Linux
- 13 AFS Servers:
  - 6 User
  - 3 Group
  - 4 Misc
University of Michigan Clients

- Major Clients
  - 5 Login Machines
  - 3 SFTP Machines
  - 4 Statistics and Computation Machines
  - 25 Web Machines
  - 15 Course tools

- All use disk caching
Cache manager flags

- **-fakestat**
  Fakes mount point stat information under /afs

- **-stat**
  Number of stat entries in memory

- **-dcache**
  Number of dcache entries in memory

- **-volumes**
  Number of memory structures for storing volume location information
Cache manager flags

- **-files**
  Number of files to create in the cache directory

- **-chunksize**
  Size of each cache chunk. Calculated as an exponent of 2

- **-rxpck**
  Sets rx_extraPackets to this value
What’s a good cache size?

- Single user: 128M
- Multi User Machine: 1-2G
- Very Big Machine: 3-4G
Monitoring the Clients

- Use rxdebug to check for noBuffers

bash-2.05b# rxdebug -rxstats -servers fala.web.itd.umich.edu -port 7001
Trying 141.211.144.202 (port 7001):
Free packets: 154, packet reclaims: 70, calls: 44641, used FDs: 64
not waiting for packets.
0 calls waiting for a thread
1 threads are idle
rx stats: free packets 154, allocs 228607986, alloc-failures(rcv 0/0, send 1/0, ack 0)
greedy 0, bogusReads 0 (last from host 0), noPackets 0, noBuffers 3, selects 0, sendSelects 0
packets read: data 62851053 ack 14069692 busy 0 abort 4494 ackall 0 challenge 6664 response 0 debug 61 params 0 unused 0
unused 0 unused 0 version 0
other read counters: data 62851051, ack 14065771, dup 18 spurious 3912 dally 11
packets sent: data 33492283 ack 101537117 busy 0 abort 16 ackall 0 challenge 0 response 6664 debug 0 params 0 unused 0
unused 0 unused 0 version 0
other send counters: ack 101537117, data 128446688 (not resends), resends 3509, pushed 0, acked&ignored 226843632
(these should be small) sendFailed 0, fatalErrors 0
Monitoring the Clients

- Use afsmonitor to look at your clients:
  `afsmonitor -cmhosts [hosts] -freq 5`

- The important stats
  - cache Blocks inUse
  - dcache Hits
  - vcache Hits
  - dcache Misses
  - vcache Misses
  - cache Flushes
2 Cache Managers monitored, 0 alerts on 0 machines

<table>
<thead>
<tr>
<th></th>
<th>Cache Num Entries</th>
<th>cache Blocks Total</th>
<th>cache Blocks InUse</th>
<th>cache Blocks Orig</th>
<th>cache MaxDirty Chunks</th>
<th>cache CurrDirty Chunks</th>
<th>dCache Hits</th>
<th>vCache Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>fala</td>
<td>50000</td>
<td>170000</td>
<td>1372806</td>
<td>170000</td>
<td>17708</td>
<td>0</td>
<td>177963169</td>
<td>25719971</td>
</tr>
<tr>
<td>gyruss</td>
<td>50000</td>
<td>170000</td>
<td>1591312</td>
<td>170000</td>
<td>2213</td>
<td>0</td>
<td>706047</td>
<td>95287</td>
</tr>
</tbody>
</table>

Command [overview, left, right]?
<table>
<thead>
<tr>
<th></th>
<th>dcache Misses</th>
<th>vcache Misses</th>
<th>cache Flushes</th>
<th>cacheFiles Reused</th>
<th>vcache XAllocs</th>
<th>dcache XAllocs</th>
<th>buf Allocated</th>
<th>buf Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>fala</td>
<td>8143088</td>
<td>22660077</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>213653764</td>
</tr>
<tr>
<td>gyruss</td>
<td>30992</td>
<td>84267</td>
<td>895</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>3914532</td>
</tr>
</tbody>
</table>
Flags to tweak

- **-chunksize**
  Default is 16 (64KB)
  19 (512KB) - 20 (1MB): 25% speed increase with large files

- **-rxpck**
  Increases the number of rx packets available. If rxdebug shows noBuffers, settings this will alleviate the problem
Flags to tweak

- **-stat**
  Increasing allows more stat entries in memory for caching stat information about cached AFS files.

- **-dcache**
  Increasing can result in a little better performance.

- **-files**
  Forces a set number of cache files to be created.
What Works Best (for us)

- stat 15000 - dcache 6000 - daemons 6 - volumes 256
- chunksize 19 - rxpck 2000 - files 50000
Misc

- Tuning requires a lot of time and patience
- Auto tuning
- avg.pl
  - Prints out average cache file size
  - Useful for chunksize
5568 files
1504381962 total bytes
270183 avg bytes

< 1k 1144 20%
< 2k 138 23%
< 3k 626 34%
< 4k 28 34%
< 5k 139 37%
< 6k 15 37%
< 7k 58 38%
< 8k 18 38%
< 9k 39 39%
< 10k 14 39%

[snip]
< 469k 1 49%
< 472k 1 49%
< 487k 1 50%
< 501k 2 50%
< 511k 1 50%
< 513k 2780 99%
Questions?

• Ask me and I’ll do my best to answer!

• Contact Information:
  Mike Garrison
  mcgarr@umich.edu