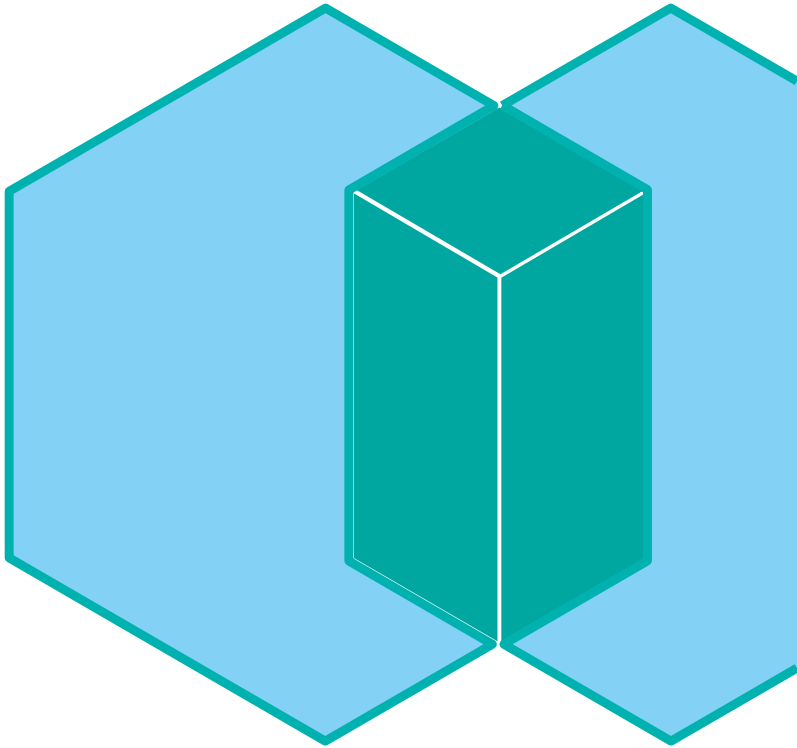
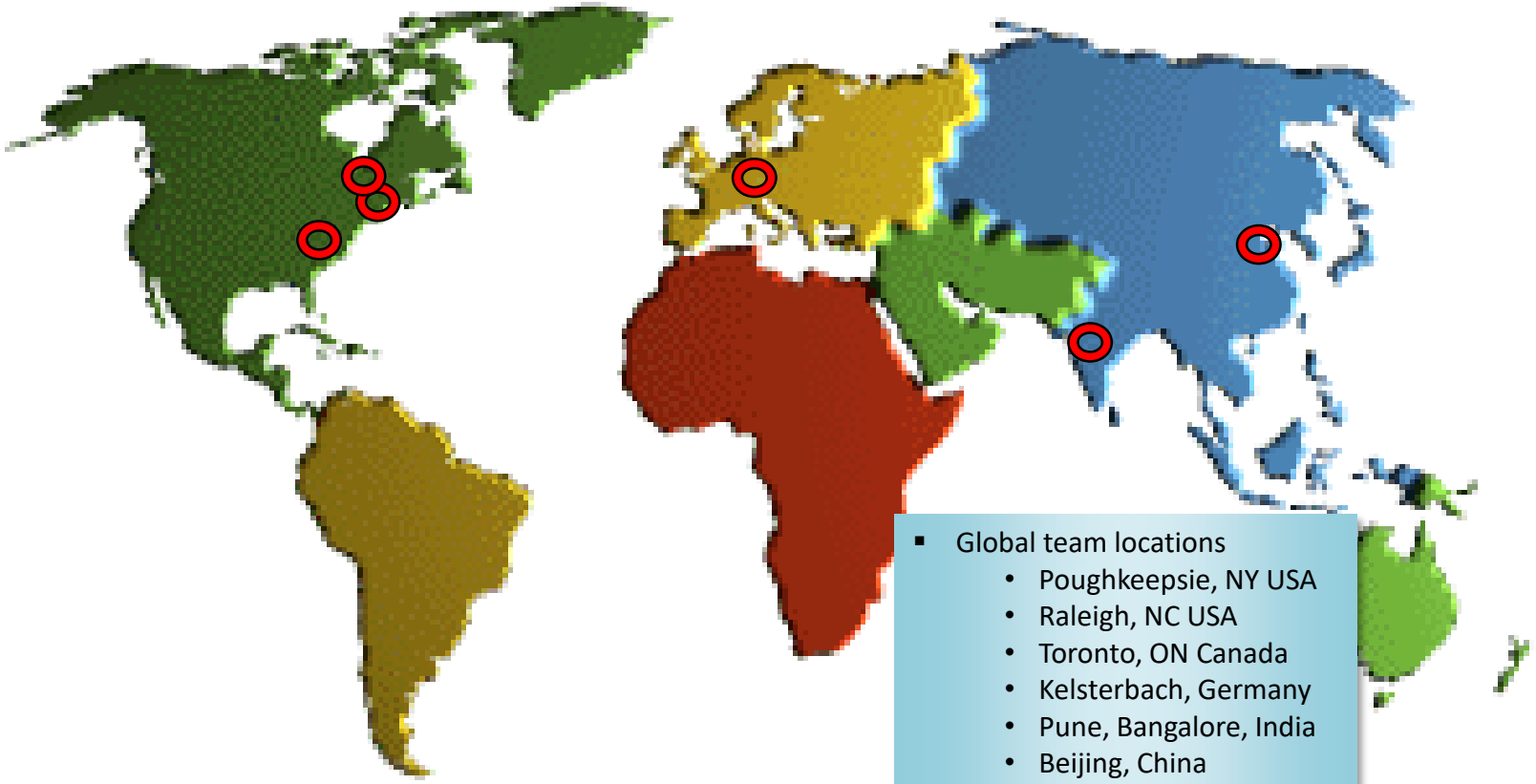


“Field Update”

Kai Korkow, Spectrum Scale and ESS Support EMEA
Kai.korkow@de.ibm.com

IBM Spectrum Scale Support Global Time Zone Coverage



- Global team locations
  - Poughkeepsie, NY USA
  - Raleigh, NC USA
  - Toronto, ON Canada
  - Kelsterbach, Germany
  - Pune, Bangalore, India
  - Beijing, China

Agenda

- 2 Interesting support cases
- Update Recommendations
- Helpful Links

Interesting case 1 : `No space left on device` with only 4% occupied?

Symptom

```
[root@pfecl1n1]/gpfs/cl1fs1/kai: dd if=/dev/zero of=./cantcreate bs=4K count=1
dd: error writing ‘./cantcreate’: No space left on device
1+0 records in
0+0 records out
0 bytes (0 B) copied, 0.276065 s, 0.0 kB/s
```

```
[root@pfecl1n1]/gpfs/cl1fs1/kai: ll cantcreate
-rw-r--r-- 1 root root 0 Apr 29 08:37 cantcreate
```

```
[root@pfecl1n1]/gpfs/cl1fs1/kai: df -h /gpfs/cl1fs1/
Filesystem      Size  Used Avail Use% Mounted on
cl1fs1           60G  2.3G   58G   4% /gpfs/cl1fs1
```

Interesting case 1 : `No space left on device` with only 4% occupied?

```
[root@pfecl1n1]/gpfs/cl1fs1/kai: mmlsfs cl1fs1
flag                value                    description
------------------- ------------------------ -----------------------------------
 -f                 8192                     Minimum fragment size in bytes
 -i                 4096                     Inode size in bytes
 -I                 32768                    Indirect block size in bytes
 -m                 1                        Default number of metadata replicas
 -r                 1                        Default number of data replicas
 -j                 scatter                  Block allocation type
 -B                 262144                   Block size
 -V                 17.00 (4.2.3.0)          File system version
 -L                 4194304                  Logfile size
 --inode-limit      8388608                  Maximum number of inodes
 --is4KAligned      Yes                      is4KAligned?
 --subblocks-per-full-block 32               Number of subblocks per full block
 -P                 system;data              Disk storage pools in file system
 -d                 SYS3;SYS4;SYS5;DAT3;DAT4;DAT5  Disks in file system
 -T                 /gpfs/cl1fs1             Default mount point
```

Creation loop

```
i=0
while true
do
  i=$[$i+1]
  dd if=/dev/zero of=/gpfs/cl1fs1/kai/$i bs=4K count=1 2>/dev/null
  echo $? $i written
done
```

Interesting case 1 : No space left on device with only 4% occupied?

```
[root@pfecl1n1]/gpfs/cl1fs1/kai: mmdf cl1fs1
disk                disk size  failure holds    holds              free KB             free KB
name                    in KB    group metadata data        in full blocks        in fragments
--------------- ------------- -------- -------- ----- -------------------- -------------------
Disks in storage pool: system (Maximum disk size allowed is 288 GB)
SYS3                 20971520        1 Yes      No           9494784 ( 45%)           472 ( 0%)
SYS4                 20971520        2 Yes      No           9499392 ( 45%)           696 ( 0%)
SYS5                 20971520        3 Yes      No           9497344 ( 45%)           632 ( 0%)
                -------------                         -------------------- -------------------
(pool total)         62914560                           28491520 ( 45%)          1800 ( 0%)

Disks in storage pool: data (Maximum disk size allowed is 288 GB)
DAT3                 20971520        1 No       Yes             0 (  0%)          7696 ( 0%)
DAT4                 20971520        2 No       Yes             0 (  0%)          6104 ( 0%)
DAT5                 20971520        3 No       Yes             0 (  0%)          7584 ( 0%)
                -------------                         -------------------- -------------------
(pool total)         62914560                                   0 (  0%)         21384 ( 0%)
                =============                         ==================== ===================
(data)               62914560                                   0 (  0%)         21384 ( 0%)
(metadata)           62914560                           28491520 ( 45%)          1800 ( 0%)
                =============                         ==================== ===================
(total)             125829120                           28491520 ( 23%)         23184 ( 0%)

Inode Information
-----------------
Number of used inodes:         7841056
Number of free inodes:          547552
Number of allocated inodes:    8388608
Maximum number of inodes:      8388608
```

```
[root@pfecl1n1]/gpfs/cl1fs1/kai: df /gpfs/cl1fs1/
Filesystem     1K-blocks     Used Available Use% Mounted on
cl1fs1          62914560 62893312     21248 100% /gpfs/cl1fs1

[root@pfecl1n1]/gpfs/cl1fs1/kai: df -h /gpfs/cl1fs1/
Filesystem      Size  Used Avail Use% Mounted on
cl1fs1           60G   60G   21M 100% /gpfs/cl1fs1

[root@pfecl1n1]/gpfs/cl1fs1/kai: ll 7836989
-rw-r--r-- 1 root root 4096 Apr 28 01:19 7836989
[root@pfecl1n1]/gpfs/cl1fs1/kai: ll 7836990

ls: cannot access 7836990: No such file or directory
```

Deletion loop

```
#!/bin/bash
# Script to delete files on subblocks 2-32

f=0
counter=1

while [ $f -le 7836989 ]; do
        counter=1
        f=$[$f+1]

                while [ $counter -lt 32 ]
                do
                ((f++))
#               echo will delete file: $f
                rm -rf $f
                ((counter++))
        done
done
```

Interesting case 1 : `No space left on device` with only 4% occupied?

After Deletion:

```
[root@pfecl1n1]/gpfs/cl1fs1/kai: mmdf cl1fs1
disk                disk size  failure holds    holds              free KB             free KB
name                    in KB    group metadata data        in full blocks        in fragments
--------------- ------------- -------- -------- ----- -------------------- -------------------
Disks in storage pool: system (Maximum disk size allowed is 288 GB)
SYS3                 20971520        1 Yes      No           9656064 ( 46%)            472 ( 0%)
SYS4                 20971520        2 Yes      No           9657088 ( 46%)            472 ( 0%)
SYS5                 20971520        3 Yes      No           9656320 ( 46%)            632 ( 0%)
                -------------                         -------------------- -------------------
(pool total)         62914560                              28969472 ( 46%)           1576 ( 0%)

Disks in storage pool: data (Maximum disk size allowed is 288 GB)
DAT3                 20971520        1 No       Yes            56576 (  0%)      20196056 (96%)
DAT4                 20971520        2 No       Yes            56576 (  0%)      20196152 (96%)
DAT5                 20971520        3 No       Yes            58624 (  0%)      20194312 (96%)
                -------------                         -------------------- -------------------
(pool total)         62914560                             171776 (  0%)      60586520 (96%)
                =============                         ==================== ===================
(data)               62914560                             171776 (  0%)      60586520 (96%)
(metadata)           62914560                           28969472 ( 46%)           1576 ( 0%)
                =============                         ==================== ===================
(total)             125829120                           29141248 ( 23%)      60588096 (48%)
Inode Information
-----------------
Number of used inodes:          248973
Number of free inodes:         8139635
Number of allocated inodes:    8388608
Maximum number of inodes:      8388608
```

```
171776 / 256
671
[root@pfecl1n1]/gpfs/cl1fs1/kai: dd if=/dev/zero of=./wastedspace bs=256K count=671
671+0 records in
671+0 records out
175898624 bytes (176 MB) copied, 2.26565 s, 77.6 MB/s
```

Interesting case 1 : `No space left on device` with only 4% occupied?

```
[root@pfecl1n1]/gpfs/cl1fs1/kai: mmdf cl1fs1
disk                disk size  failure holds    holds              free KB             free KB
name                    in KB    group metadata data        in full blocks        in fragments
--------------- ------------- -------- -------- ----- -------------------- -------------------
Disks in storage pool: system (Maximum disk size allowed is 288 GB)
SYS3                 20971520        1 Yes      No           9656064 ( 46%)            472 ( 0%)
SYS4                 20971520        2 Yes      No           9656832 ( 46%)            696 ( 0%)
SYS5                 20971520        3 Yes      No           9656320 ( 46%)            632 ( 0%)
               -------------                         -------------------- -------------------
(pool total)         62914560                              28969216 ( 46%)           1800 ( 0%)

Disks in storage pool: data (Maximum disk size allowed is 288 GB)
DAT3                 20971520        1 No       Yes                 0 (  0%)      20196056 (96%)
DAT4                 20971520        2 No       Yes                 0 (  0%)      20196152 (96%)
DAT5                 20971520        3 No       Yes                 0 (  0%)      20194312 (96%)
               -------------                         -------------------- -------------------
(pool total)         62914560                                    0 (  0%)      60586520 (96%)
               =============                         ==================== ====================
(data)               62914560                                    0 (  0%)      60586520 (96%)
(metadata)           62914560                              28969216 ( 46%)           1800 ( 0%)
               =============                         ==================== ====================
(total)             125829120                              28969216 ( 23%)      60588320 (48%)

Inode Information
-----------------
Number of used inodes:          248974
Number of free inodes:         8139634
Number of allocated inodes:    8388608
Maximum number of inodes:      8388608
```

```
[root@pfecl1n1]/gpfs/cl1fs1/kai: df -h /gpfs/cl1fs1/
Filesystem      Size  Used Avail Use% Mounted on
cl1fs1           60G  2.3G   58G   4% /gpfs/cl1fs1
```

Interesting case 2: ESS Power LE node hang due to oom

- **Symptom**
  - System hang. Unable to ssh into the node. No console login possible.
  - Will most likely still respond to ping over network.
  - Free memory is significantly higher than vm.min_free_kbytes.
- **Cause**
  - RHEL on Power LE systems is allocating 5% of the installed memory for kvm_cma by default.
  - This memory is reported as free memory but can't be used by normal applications.
- **Resolving The Problem**
  - kvm_cma allocation during boot can be disabled by adding the kernel option: kvm_cma_resv_ratio=0
  - `vi /etc/default/grub`
  - `GRUB_CMDLINE_LINUX="crashkernel=auto console=hvc0 kvm_cma_resv_ratio=0"`
  - `grub2-mkconfig -o /boot/grub2/grub.cfg`
  - `reboot`
- ESS Power LE node hang due to oom while free memory > vm.min_free_kbytes

Update Recommendations

http://www-01.ibm.com/support/docview.wss?uid=ssg1S1009703

IBM Spectrum Scale Software Version Recommendation Preventive Service Planning

Preventive Service Planning

Abstract

IBM Spectrum Scale Software Version Recommendation

Content

This generalized recommendation is made available to assist clients in implementing a code update strategy. It is a full field perspective, and as such, a customized recommendation which takes into account specifics such as business upgrade windows, length of time since last update, decommission plans, etc. may require assistance from local support teams. In general, recommendations assume planning updates annually.

IBM Spectrum Scale	Minimum Recommended Level	Field proven level	Latest Level
IBM Spectrum Scale	4.1.1.22 (**Alert**: Version 4.1.x is going End Of Support on 4/30/2019. Clients are advised to upgrade to one of the versions in the "field proven" column as soon as possible). [1]	4.x stream 4.2.3.12 [Nov 15, 2018] 5.x stream 5.0.2.2 [Dec 13, 2018]	4.x stream 4.2.3.13 [Jan 2019] 5.x stream 5.0.2.3 [Feb 2019]
IBM Spectrum Scale for ESS	ESS 4.6 [Mar 2017]	4.x stream: ESS 5.2.5 [Dec 2018] 5.x stream: ESS 5.3.2 [Nov 2018]	4.x stream ESS 5.2.5 [Dec 2018] 5.x stream: ESS 5.3.2.1 [Feb 2019]

[1]For information on upgrade, see the Upgrade section in the IBM Spectrum Scale Knowledge Center. The IBM Spectrum Scale supported upgrade

Useful Links

- Spectrum Scale FAQ : everything you may want to know about supported releases, Upgrade Paths, etc.
- Release currency, field approved level
- IBM Spectrum Scale Network Related Flows and Troubleshooting
- IBM Spectrum Scale Wiki

Questions?