Ceph - Bug #19200

RHEL 7.3 Selinux denials at OSD start

03/06/2017 07:51 PM - Ben Meekhof

Status:	Resolved	**Start date:**	03/06/2017
Priority:	Low	**Due date:**	
Assignee:	Boris Ranto	**% Done:**	0%
Category:		**Estimated time:**	0.00 hour
Target version:		**Spent time:**	0.00 hour
Source:		**Reviewed:**	
Tags:		**Affected Versions:**	v11.2.1
Backport:	kraken, jewel, luminous	**ceph-qa-suite:**	
Regression:	No	**Pull request ID:**	
Severity:	4 - irritation	**Crash signature:**	

Description

I get a batch of SElinux denials when starting Kraken OSD. However there does not seem to be any impairment of the function of the OSD.
SElinux package is ceph-selinux-11.2.0-0.el7.x86_64. OS is RHEL-derived Scientific Linux 7.3, kernel 3.10.0-514.10.2.el7.x86_64, selinux-policy-targeted-3.13.1-102.el7_3.15.noarch.

Messages:
type=AVC msg=audit(1488828549.766:5687): avc: denied { write } for pid=1501374 comm="journal_write" path="/dev/nvme2n1p9" dev="devtmpfs" ino=37915 scontext=system_u:system_r:ceph_t:s0 tcontext=system_u:object_r:nvme_device_t:s0 tclass=blk_file
type=SYSCALL msg=audit(1488828549.766:5687): arch=c000003e syscall=209 success=yes exit=1 a0=7f53c72ee000 a1=1 a2=7f53bcd9d5e8 a3=0 items=0 ppid=1 pid=1501374 auid=4294967295 uid=167 gid=167 euid=167 suid=167 fsuid=167 egid=167 sgid=167 fsgid=167 tty=(none) ses=4294967295 comm="journal_write" exe="/usr/bin/ceph-osd" subj=system_u:system_r:ceph_t:s0 key=(null)
type=AVC msg=audit(1488828550.243:5688): avc: denied { name_connect } for pid=1502719 comm="ceph-osd" dest=7001 scontext=system_u:system_r:ceph_t:s0 tcontext=system_u:object_r:afs3_callback_port_t:s0 tclass=tcp_socket
type=SYSCALL msg=audit(1488828550.243:5688): arch=c000003e syscall=42 success=no exit=-115 a0=3e a1=7fad0b55a88c a2=10 a3=7facf4117f5c items=0 ppid=1 pid=1502719 auid=4294967295 uid=167 gid=167 euid=167 suid=167 fsuid=167 egid=167 sgid=167 fsgid=167 tty=(none) ses=4294967295 comm="ceph-osd" exe="/usr/bin/ceph-osd" subj=system_u:system_r:ceph_t:s0 key=(null)
type=SERVICE_START msg=audit(1488828551.920:5689): pid=1 uid=0 auid=4294967295 ses=4294967295 subj=system_u:system_r:init_t:s0 msg='unit=ceph-osd@550 comm="systemd" exe="/usr/lib/systemd/systemd" hostname=? addr=? terminal=? res=success'
type=SERVICE_STOP msg=audit(1488828551.921:5690): pid=1 uid=0 auid=4294967295 ses=4294967295 subj=system_u:system_r:init_t:s0 msg='unit=ceph-osd@550 comm="systemd" exe="/usr/lib/systemd/systemd" hostname=? addr=? terminal=? res=success'
type=SERVICE_START msg=audit(1488828551.976:5691): pid=1 uid=0 auid=4294967295 ses=4294967295 subj=system_u:system_r:init_t:s0 msg='unit=ceph-osd@550 comm="systemd" exe="/usr/lib/systemd/systemd" hostname=? addr=? terminal=? res=success'
type=AVC msg=audit(1488828552.079:5692): avc: denied { read } for pid=1535965 comm="ceph-osd" name="nvme0n1p10" dev="devtmpfs" ino=37933 scontext=system_u:system_r:ceph_t:s0 tcontext=system_u:object_r:nvme_device_t:s0 tclass=blk_file
type=AVC msg=audit(1488828552.079:5692): avc: denied { open } for pid=1535965 comm="ceph-osd" path="/dev/nvme0n1p10" dev="devtmpfs" ino=37933 scontext=system_u:system_r:ceph_t:s0 tcontext=system_u:object_r:nvme_device_t:s0 tclass=blk_file
type=SYSCALL msg=audit(1488828552.079:5692): arch=c000003e syscall=2 success=yes exit=21 a0=7fd5a87f7898 a1=0 a2=1a4 a3=1 items=0 ppid=1 pid=1535965 auid=4294967295 uid=167 gid=167 euid=167 suid=167 fsuid=167 egid=167 sgid=167 fsgid=167 tty=(none) ses=4294967295 comm="ceph-osd" exe="/usr/bin/ceph-osd" subj=system_u:system_r:ceph_t:s0 key=(null)
type=AVC msg=audit(1488828552.080:5693): avc: denied { getattr } for pid=1535965 comm="ceph-osd" path="/dev/nvme0n1p10" dev="devtmpfs" ino=37933 scontext=system_u:system_r:ceph_t:s0 tcontext=system_u:object_r:nvme_device_t:s0 tclass=blk_file
type=SYSCALL msg=audit(1488828552.080:5693): arch=c000003e syscall=5 success=yes exit=0 a0=15 a1=7ffd15772ac0 a2=7ffd15772ac0 a3=1 items=0 ppid=1 pid=1535965 auid=4294967295 uid=167 gid=167 euid=167 suid=167 fsuid=167 egid=167 sgid=167 fsgid=167 tty=(none) ses=4294967295 comm="ceph-osd" exe="/usr/bin/ceph-osd" subj=system_u:system_r:ceph_t:s0 key=(null)
type=AVC msg=audit(1488828552.080:5694): avc: denied { ioctl } for pid=1535965 comm="ceph-osd" path="/dev/nvme0n1p10" dev="devtmpfs" ino=37933 scontext=system_u:system_r:ceph_t:s0 tcontext=system_u:object_r:nvme_device_t:s0 tclass=blk_file
type=SYSCALL msg=audit(1488828552.080:5694): arch=c000003e syscall=16 success=yes exit=0 a0=15 a1=80081272 a2=7ffd15772918 a3=7ffd15772680 items=0 ppid=1 pid=1535965 auid=4294967295 uid=167 gid=167 euid=167 suid=167 fsuid=167

egid=167 sgid=167 fsgid=167 tty=(none) ses=4294967295 comm="ceph-osd" exe="/usr/bin/ceph-osd" subj=system_u:system_r:ceph_t:s0 key=(null)
type=AVC msg=audit(1488828554.749:5695): avc: denied { name_connect } for pid=1535967 comm="ceph-osd" dest=6969 scontext=system_u:system_r:ceph_t:s0 tcontext=system_u:object_r:tor_port_t:s0 tclass=tcp_socket
type=SYSCALL msg=audit(1488828554.749:5695): arch=c000003e syscall=42 success=no exit=-115 a0=84 a1=7fd5ace7188c a2=10 a3=7fd597f05f5c items=0 ppid=1 pid=1535967 auid=4294967295 uid=167 gid=167 euid=167 suid=167 fsuid=167 egid=167 sgid=167 fsgid=167 tty=(none) ses=4294967295 comm="ceph-osd" exe="/usr/bin/ceph-osd" subj=system_u:system_r:ceph_t:s0 key=(null)
type=AVC msg=audit(1488828554.758:5696): avc: denied { name_connect } for pid=1535967 comm="ceph-osd" dest=7002 scontext=system_u:system_r:ceph_t:s0 tcontext=system_u:object_r:afs_pt_port_t:s0 tclass=tcp_socket
type=SYSCALL msg=audit(1488828554.758:5696): arch=c000003e syscall=42 success=no exit=-115 a0=113 a1=7fd5ad17388c a2=10 a3=7fd597f05f5c items=0 ppid=1 pid=1535967 auid=4294967295 uid=167 gid=167 euid=167 suid=167 fsuid=167 egid=167 sgid=167 fsgid=167 tty=(none) ses=4294967295 comm="ceph-osd" exe="/usr/bin/ceph-osd" subj=system_u:system_r:ceph_t:s0 key=(null)
type=AVC msg=audit(1488828554.935:5697): avc: denied { name_connect } for pid=1505236 comm="ceph-osd" dest=7000 scontext=system_u:system_r:ceph_t:s0 tcontext=system_u:object_r:gatekeeper_port_t:s0 tclass=tcp_socket
type=SYSCALL msg=audit(1488828554.935:5697): arch=c000003e syscall=42 success=no exit=-115 a0=7b a1=7f2183c1308c a2=10 a3=7f216e482f5c items=0 ppid=1 pid=1505236 auid=4294967295 uid=167 gid=167 euid=167 suid=167 fsuid=167 egid=167 sgid=167 fsgid=167 tty=(none) ses=4294967295 comm="ceph-osd" exe="/usr/bin/ceph-osd" subj=system_u:system_r:ceph_t:s0 key=(null)

This policy is what audit2allow generates from the messages:

```
require {
type tor_port_t;
type gatekeeper_port_t;
type nvme_device_t;
type ceph_t;
type afs3_callback_port_t;
type afs_pt_port_t;
class tcp_socket name_connect;
class blk_file { getattr ioctl open read write };
}

#============= ceph_t ==============
allow ceph_t afs3_callback_port_t:tcp_socket name_connect;
allow ceph_t afs_pt_port_t:tcp_socket name_connect;
allow ceph_t gatekeeper_port_t:tcp_socket name_connect;
allow ceph_t nvme_device_t:blk_file { getattr ioctl open read write };
allow ceph_t tor_port_t:tcp_socket name_connect;
```

Related issues:

Copied to Ceph - Backport #21037: kraken: RHEL 7.3 Selinux denials at OSD start	**Rejected**
Copied to Ceph - Backport #21052: luminous: RHEL 7.3 Selinux denials at OSD s...	**Resolved**
Copied to Ceph - Backport #21053: jewel: RHEL 7.3 Selinux denials at OSD start	**Resolved**

History

#1 - 04/26/2017 12:07 PM - Ruben Kerkhof

I see these on Jewel (10.2.7) too:

type=AVC msg=audit(1493207893.757:66): avc: denied { read } for pid=2448 comm="ceph-osd" name="nvme0n1p10" dev="devtmpfs" ino=14211 scontext=system_u:system_r:ceph_t:s0 tcontext=system_u:object_r:nvme_device_t:s0 tclass=blk_file
type=AVC msg=audit(1493207893.757:67): avc: denied { open } for pid=2448 comm="ceph-osd" path="/dev/nvme0n1p10" dev="devtmpfs" ino=14211 scontext=system_u:system_r:ceph_t:s0 tcontext=system_u:object_r:nvme_device_t:s0 tclass=blk_file
type=AVC msg=audit(1493207893.760:68): avc: denied { getattr } for pid=2448 comm="ceph-osd" path="/dev/nvme0n1p10" dev="devtmpfs" ino=14211 scontext=system_u:system_r:ceph_t:s0 tcontext=system_u:object_r:nvme_device_t:s0 tclass=blk_file
type=AVC msg=audit(1493207893.760:69): avc: denied { ioctl } for pid=2448 comm="ceph-osd" path="/dev/nvme0n1p10" dev="devtmpfs" ino=14211 scontext=system_u:system_r:ceph_t:s0 tcontext=system_u:object_r:nvme_device_t:s0 tclass=blk_file
type=AVC msg=audit(1493207894.620:71): avc: denied { write } for pid=2444 comm="ceph-osd" name="nvme0n1p12" dev="devtmpfs" ino=14213 scontext=system_u:system_r:ceph_t:s0 tcontext=system_u:object_r:nvme_device_t:s0 tclass=blk_file

#2 - 04/26/2017 12:19 PM - Ruben Kerkhof

And some more:
type=AVC msg=audit(1493209002.775:222): avc: denied { read } for pid=5831 comm="fn_odsk_fstore" name="nvme0n1p5" dev="devtmpfs" ino=10436 scontext=system_u:system_r:ceph_t:s0 tcontext=system_u:object_r:nvme_device_t:s0 tclass=blk_file
type=AVC msg=audit(1493209002.775:223): avc: denied { open } for pid=5831 comm="fn_odsk_fstore" path="/dev/nvme0n1p5" dev="devtmpfs" ino=10436 scontext=system_u:system_r:ceph_t:s0 tcontext=system_u:object_r:nvme_device_t:s0 tclass=blk_file
type=AVC msg=audit(1493209002.775:224): avc: denied { ioctl } for pid=5831 comm="fn_odsk_fstore" path="/dev/nvme0n1p5" dev="devtmpfs" ino=10436 scontext=system_u:system_r:ceph_t:s0 tcontext=system_u:object_r:nvme_device_t:s0 tclass=blk_file

#3 - 06/07/2017 09:03 PM - Boris Ranto

- Assignee set to Boris Ranto

Hi Ruben,

what do you use as a storage? It looks like ceph-osd is trying to read/write /dev/nvme0n1p10. Do you use an nvme device for a journal? Are you running in permissive or enforcing mode?

Unfortunately, we currently do not allow nvme devices in the policy but we can add it.

Regards,
Boris

#4 - 06/08/2017 11:04 AM - Ruben Kerkhof

Hi Boris,

Indeed, I use a single NVME device for both the OS filesystems, and the 12 journal partitions for my OSDS. My OSDs are on HDD.
I'm currently running in permissive mode, mainly because of this.

If you could add support in the policy that would be great.

#5 - 06/08/2017 02:25 PM - Ben Meekhof

I originally opened this issue on Kraken and my configuration consists of HDD OSD and multiple journal partitions on NVMe devices.

I see the same denied errors as Ruben on our Jewel OSD as well.

My systems are all in enforcing mode - we haven't seen any noticeable effect from this and only happened to notice the logs when tracking down a different SElinux problem.

#6 - 06/09/2017 12:40 PM - Boris Ranto

- Status changed from New to In Progress

- Backport set to kraken, jewel

Upstream PR:

https://github.com/ceph/ceph/pull/15597

#7 - 06/12/2017 11:58 AM - Kefu Chai

Ruben, are your ceph-osd daemons are using the default settings?

as ceph-osd is not supposed to connect to these ports 6969,7000,7001,7002 by default.

#8 - 06/12/2017 12:01 PM - Ruben Kerkhof

Kefu Chai wrote:

> Ruben, are your ceph-osd daemons are using the default settings?
>
> as ceph-osd is not supposed to connect to these ports 6969,7000,7001,7002 by default.

Hi Kefu,

Where in my AVC's are you seeing the OSDs connecting to these ports?
Or was this meant for Ben?

#9 - 06/12/2017 02:45 PM - Ben Meekhof

We are using the default settings though we do set some debug levels, crush location, etc unrelated to network ports. Just because it is a network related setting I'll note that we do use separate cluster_network and public_network subnets in the global section.

#10 - 06/12/2017 03:40 PM - Kefu Chai

Ruben, very sorry. my queries were meant for Ben.

#11 - 08/16/2017 12:31 PM - Kefu Chai

- Status changed from In Progress to Pending Backport

- Backport changed from kraken, jewel to kraken, jewel, luminous

#12 - 08/18/2017 09:44 AM - Nathan Cutler

- Copied to Backport #21037: kraken: RHEL 7.3 Selinux denials at OSD start added

#13 - 08/21/2017 04:15 PM - Nathan Cutler

- Copied to Backport #21052: luminous: RHEL 7.3 Selinux denials at OSD start added

#14 - 08/21/2017 04:15 PM - Nathan Cutler

- Copied to Backport #21053: jewel: RHEL 7.3 Selinux denials at OSD start added

#15 - 10/26/2018 10:51 AM - Nathan Cutler

- Status changed from Pending Backport to Resolved