Status of SELinux support in Lustre

DataDirect Networks, Inc.

Sebastien Buisson sbuisson@ddn.com

2017/05

Status of SELinux support in Lustre

- **LU-8956: security context at create**
  - Performance & correctness
- **LU-9193: security context at lookup**
  - Atomicity
- **LU-8955: Send SELinux policy info to server**
  - Kernel side SELinux enhancements
  - Lustre side security awareness

© 2017 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others.
Any statements or representations around future events are subject to change.

LU-8956: security context at create

- **SELinux-wise, atomicity of create addressed with patch:**

```
LU-5560 security: send file security context
                  for creates
```

  - Send file security context to MDT along with create RPC
  - Close the insecure window between creation and setting of the security context

© 2017 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others.
Any statements or representations around future events are subject to change.

LU-8956: security context at create

- **But patch needs to be improved:**
  - security context must be set on inode after create
    - Do not leave inode with uninitialized sec context
    - Avoid getxattr to fetch on-disk sec context

- **New patch at:**

  `https://review.whamcloud.com/24426`

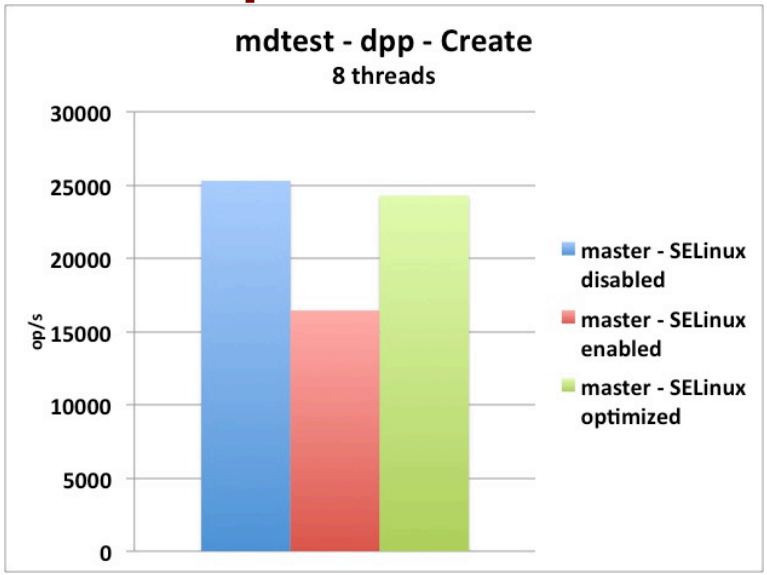
  Calls security_inode_notifysecctx() before d_instantiate().

  Landed on Monday!

© 2017 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others.
Any statements or representations around future events are subject to change.

LU-8956: security context at create

- **Performance improvement:**



- **Reduction of MDS-side contention…**

© 2017 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others.
Any statements or representations around future events are subject to change.

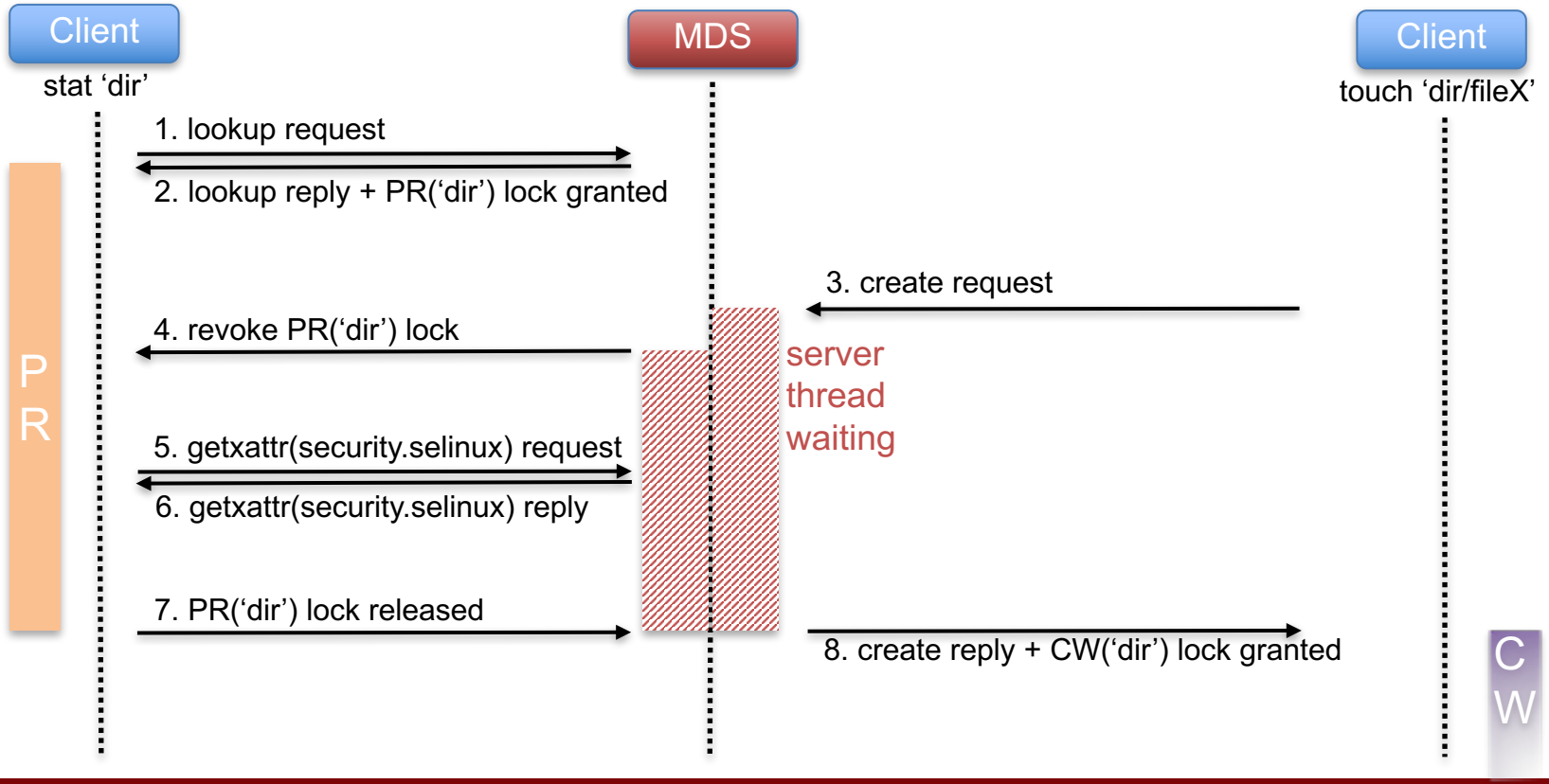
LU-9193: security context at lookup

- **SELinux-wise, lookup operation not atomic**

ll_lookup_it
→ ***sends request to MDS***
→ ***receives reply from MDS with PR lock granted***
ll_lookup_it_finish
ll_splice_alias
d_instantiate
security_d_instantiate
[…]
ll_getxattr(*security.selinux*)
→ ***sends request to MDS***
→ ***now able to release PR lock***

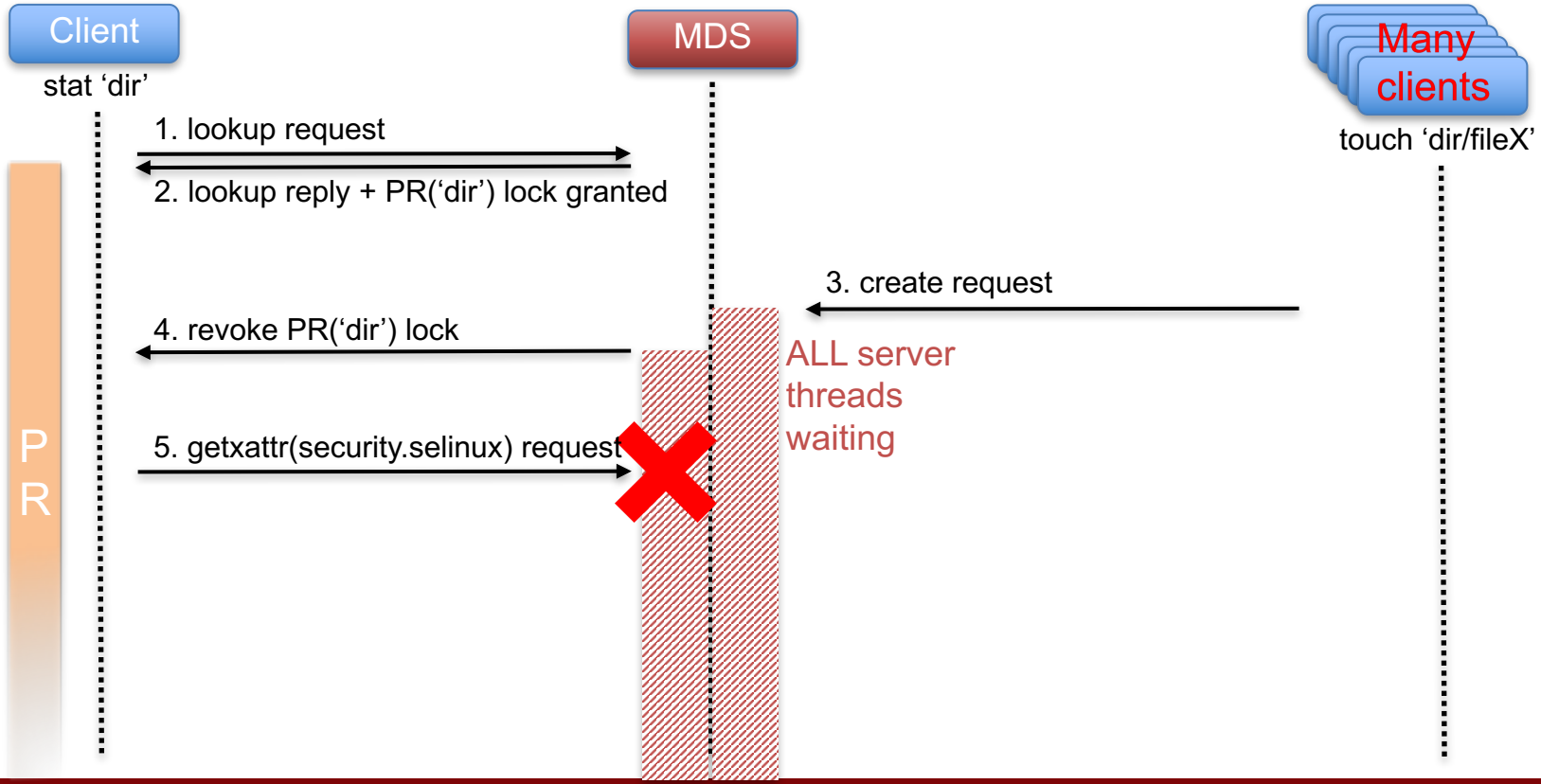
© 2017 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others.
Any statements or representations around future events are subject to change.

LU-9193: security context at lookup



© 2017 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others.
Any statements or representations around future events are subject to change.

LU-9193: security context at lookup



© 2017 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others.
Any statements or representations around future events are subject to change.

LU-9193: security context at lookup

- **New patch, owner Bruno Faccini (Intel):**
  `https://review.whamcloud.com/26831`

ll_lookup_it

→ ***sends request to MDS***

→ ***receives reply from MDS with PR lock granted + sec ctx***

ll_lookup_it_finish

security_inode_notifysecctx

ll_splice_alias

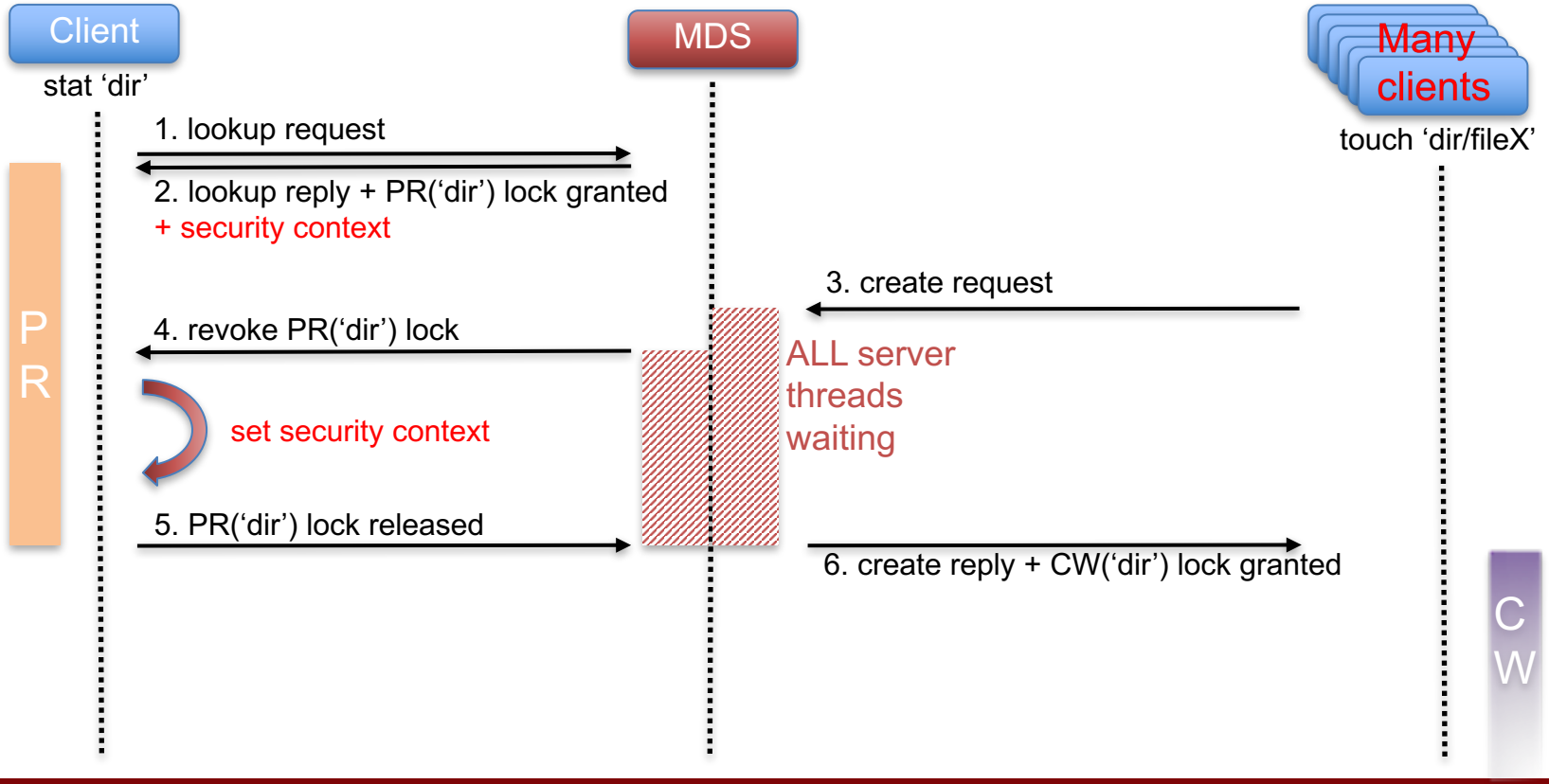
d_instantiate

security_d_instantiate

→ ***now able to release PR lock***

© 2017 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others.
Any statements or representations around future events are subject to change.

LU-9193: security context at lookup



© 2017 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others.
Any statements or representations around future events are subject to change.

LU-8955: Send SELinux policy info to server

- **Distributed file systems specificity:**
  - Really need to make sure data is always accessed by nodes with SELinux policy **properly enforced**
    - Otherwise data is not protected, especially with SELinux MLS
- **Retrieve SELinux status on client nodes**
  - Build a representation of policy's characteristics
- **Send clients' SELinux status to servers along with requests**
  - connect
  - ldlm locks
  - create - open - unlink - rename - getxattr - setxattr
- **On servers, compare info received from clients with reference status stored into nodemap**
  - If they differ => Permission Denied

© 2017 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others.
Any statements or representations around future events are subject to change.

LU-8955: Send SELinux policy info to server

- **Currently, kernel does not expose necessary information**
  - Most is only available via userspace commands
  - or via /sys/fs/selinux/
- **Create a new Lustre usermode helper**
  - l_getsepol, writing back SELinux status to /proc/fs/lustre/<target>/srpc_sepol
- **Usermode helper drawbacks**
  - Performance penalty!
    - Call only when policy changed
  - Security flaw?

© 2017 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others.
Any statements or representations around future events are subject to change.

LU-8955: Send SELinux policy info to server

Kernel side SELinux enhancements

- Compute a “policy brief” in SELinux kernel code:

```
selinux(enforce=<0 or 1>:checkreqprot=<0 or 1>:<hashalg>=<checksum>)
```

- Update every time policy is modified.
- Add hook to expose policy brief to the rest of the kernel.
- Implement notifications to have hook called only when policy brief is updated.

Usage on Lustre side

- “policy brief” used as SELinux status info sent from client to server.

© 2017 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others.
Any statements or representations around future events are subject to change.

LU-8955: Send SELinux policy info to server

Kernel side SELinux enhancements

- Work on kernel patches is actively in progress.
  - Many different ideas from different maintainers…

- Need Lustre Community push for
  - landing in upstream kernel
    - Lustre being in staging is an obstacle
    - Maintainers want to see Lustre code using the new SELinux hook
  - merge in CentOS/RH
    - Lustre client's kernel cannot easily be patched at customer sites!

© 2017 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others.
Any statements or representations around future events are subject to change.

LU-8955: Send SELinux policy info to server

- **Lustre side security awareness**
  - A lot of stuff happens on Lustre client after server's reply, eg:

.atomic_open = ll_atomic_open
ll_lookup_it → ***sends req and receives reply from MDS***
ll_lookup_it_finish
ll_splice_alias
d_instantiate
security_d_instantiate → ***sets inode sec ctx***
finish_open
do_dentry_open
security_file_open → ***applies policy on inode sec ctx***
.open = ll_file_open

© 2017 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others.
Any statements or representations around future events are subject to change.

LU-8955: Send SELinux policy info to server

- **Lustre side security awareness**
  - A lot of stuff happens on Lustre client after server's reply.

.atomic_open = ll_atomic_open
ll_lookup_it → ***check SELinux status on server side***
ll_lookup_it_finish
ll_splice_alias
d_instantiate
security_d_instantiate → ***sets inode sec ctx***
finish_open
do_dentry_open
security_file_open → ***applied policy is not the one checked before***
.open = ll_file_open

policy modified in the meantime

© 2017 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others.
Any statements or representations around future events are subject to change.

LU-8955: Send SELinux policy info to server

- **Lustre side security awareness**

.atomic_open = ll_atomic_open
ll_lookup_it → ***sends req and receives reply from MDS***
ll_lookup_it_finish
ll_splice_alias
d_instantiate
security_d_instantiate → ***sets inode sec ctx***
finish_open
do_dentry_open
security_file_open → ***applies policy on inode sec ctx***
.open = ll_file_open
compare sequence of policy used in req with current
if they differ call ll_invalidate_aliases() and return -EACCES

© 2017 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others.
Any statements or representations around future events are subject to change.

Thank You!

Keep in touch with us

Team-jpsales@ddn.com

@ddn_limitless

company/datadirect-networks

102-0081
東京都千代田区四番町6-2
東急番町ビル 8F

TEL:03-3261-9101
FAX：03-3261-9140

© 2017 DataDirect Networks, Inc. * Other names and brands may be claimed as the property of others.
Any statements or representations around future events are subject to change.