Milestone Completion for Implementation Milestone 2 of the Distributed Namespace Project of contract SFS-DEV-001.

Revision History

Date	Revision	Author
11/07/12	Original	R. Henwood

Contents

Introduction

The following milestone completion document applies to Subproject 2.1 – Remote Directories subproject within the OpenSFS Lustre Development contract SFS-DEV-001 signed 7/30/2011.

Subproject Description

Per the contract, DNE 1: Remote Directories is described as follows:

> *This subproject distributes the Lustre namespace over multiple metadata targets (MDTs) under administrative control using a Lustre-specific mkdir command. Whereas normal users are only able to create child directories and files on the same MDT as the parent directory, administrators can use this command to create a directory on a different MDT. The contents of any directory remain limited to a single MDT. Rename and hardlink operations between files and directories on different MDTs return EXDEV, forcing applications and utilities to treat them as if they are on different file systems. This limits the complexity of the implementation of this subproject while delivering capacity and performance scaling benefits for the entire namespace in aggregate.*
>
> *Metadata update operations that span multiple MDTs are sequenced and synchronized to create and/or increment the link count on a MDT object before it is referenced by the remote directory entry and to update the remote directory entry before decrementing the link count and/or destroying the MDT object it referenced. Although this may result in an orphan MDT object under some failure conditions, it ensures that the Lustre namespace remains intact under any and all failure scenarios. All the other metadata operations avoid synchronous I/O and execute with full performance.*
>
> *This project includes the implementation of OST FIDs (File Identifiers). These are required to overcome a limitation in the current 2.x Lustre protocol that would otherwise prevent a single file system from having more than 8 MDTs. Addressing this technical debt in the first subproject of DNE avoids protocol compatibility issues that would arise if this feature were implemented after Remote Directories were used in production.*

Milestone Completion Criteria

Per the contract, three Implementation milestones have been defined by the Whamcloud. This document is concerned with completing the first Implementation milestone which is agreed as:

> *Demonstrate DNE recovery and failover. Suitable DNE-specific recovery and failover tests will be added and passed.*

These requirements are demonstrated below.

Demonstrate DNE recovery and fail-over

Recovery and fail-over testing is performed by the replay-dual, replay-single and recovery-small tests.

replay-dual

Verify recovery from two clients after server failure.
https://maloo.whamcloud.com/test_sets/7767e970-24cd-11e2-9e7c-52540035b04c

recovery-small

Verify RPC replay after communications failure
https://maloo.whamcloud.com/test_sets/76bb0e82-24cb-11e2-9e7c-52540035b04c

replay-single

https://maloo.whamcloud.com/test_sets/dd004406-24ca-11e2-9e7c-52540035b04c
The test platform is the OpenSFS Functional Test Cluster. The test run with these results is recorded in maloo at:
https://maloo.whamcloud.com/test_sessions/4f6f3d5a-24bf-11e2-9e7c-52540035b04c
A screenshot of the test session is recorded in Appendix A.

DNE-Specific recovery and failover tests will be added and passed.

Change	Commit Message
4318	OSF-69 tests: add parallel sanity tests to dne
4319	OSF-69 dne: add dne test into insanity.sh
4320	OSF-69 dne: add remote dir test to recovery-xxx-scale
4321	OSF-69 dne: add remote dir check in replay-vbr.
4367	OSF-133 tests: DNE fixes for conf sanity.
4366	OSF-69 tests: Add dne specific tests to sanityN
4365	OSF-69 tests: add create remote directory to racer
4364	OSF-69 tests: add DNE upgrade tests.
4363	OSF-110 tests: support multiple node fails
4362	OSF-110 tests: add dne test cases in replay-single
4361	OSF-110 tests: add dne tests cases in replay-dual
4360	OSF-110 tests: add DNE specific tests in recovery-small
4359	OSF-112 tests: Add test_mkdir in sanity for DNE
4358	OSF-69 tests: Add DNE test cases in sanity 230.

Conclusion

Implementation phase 2 has been completed according to the agreed criteria.

Appendix A: recovery and fail-over testing screenshot

Session for group regression (c21, Di Wang)

4 test sets passed out of 10.

Code review references

No code review references found.

Test sets

Name	Test group	Test host	Branch	Arch / Lustre Version	Run at (UTC)	Duration	Subtests passed	Bugs	Links	User	Status
racer	regression	c21	• {not reported}	• x86_64,el6,inkern	2012-11-01 11:14:46	0	2/2			Di Wang	ABORT
insanity	regression	c21	• {not reported}	• x86_64,el6,inkern	2012-11-01 10:19:51	3295	12/13			Di Wang	FAIL
replay-vbr	regression	c21	• {not reported}	• x86_64,el6,inkern	2012-11-01 05:20:17	17974	43/43			Di Wang	PASS
replay-dual	regression	c21	• {not reported}	• x86_64,el6,inkern	2012-11-01 03:51:45	5312	32/32			Di Wang	PASS
replay-ost-single	regression	c21	• {not reported}	• x86_64,el6,inkern	2012-11-01 03:41:17	628	7/8	LU-2245		Di Wang	FAIL
recovery-small	regression	c21	• {not reported}	• x86_64,el6,inkern	2012-11-01 02:25:23	4534	65/65			Di Wang	PASS
replay-single	regression	c21	• {not reported}	• x86_64,el6,inkern	2012-10-31 23:07:09	11874	125/125			Di Wang	PASS
conf-sanity	regression	c21	• {not reported}	• x86_64,el6,inkern	2012-10-31 21:29:24	5865	38/39			Di Wang	FAIL
sanityn	regression	c21	• {not reported}	• x86_64,el6,inkern	2012-10-31 20:40:15	2949	124/125	lu-2280		Di Wang	FAIL
sanity	regression	c21	• {not reported}	• x86_64,el6,inkern	2012-10-31 18:31:47	7708	431/446	LU-582, LU-2113		Di Wang	FAIL

Test nodes

c22	
Lustre Version:	jenkins-arch=x86_64,build_type=client,distro=el6,ib_stack=inkern
Architecture:	x86_64
Node Architecture:	x86_64
File System:	{not reported}
Name:	c22
Networks:	tcp
Distribution:	CentOS release 6.3
OS:	GNU/Linux

c21	
Lustre Version:	jenkins-arch=x86_64,build_type=client,distro=el6,ib_stack=inkern
Architecture:	x86_64
Node Architecture:	x86_64
File System:	{not reported}
Name:	c21
Networks:	tcp
Distribution:	CentOS release 6.3
OS:	GNU/Linux

NOTE: Tests unrelated to recovery and failover exhibit a small number of failures. These are a result of unresolved issues on Master branch. DNE is now based on Master so as the sources of failure are resolved they will be inherited by DNE.