



Lustre DNE update

Lai Siyao

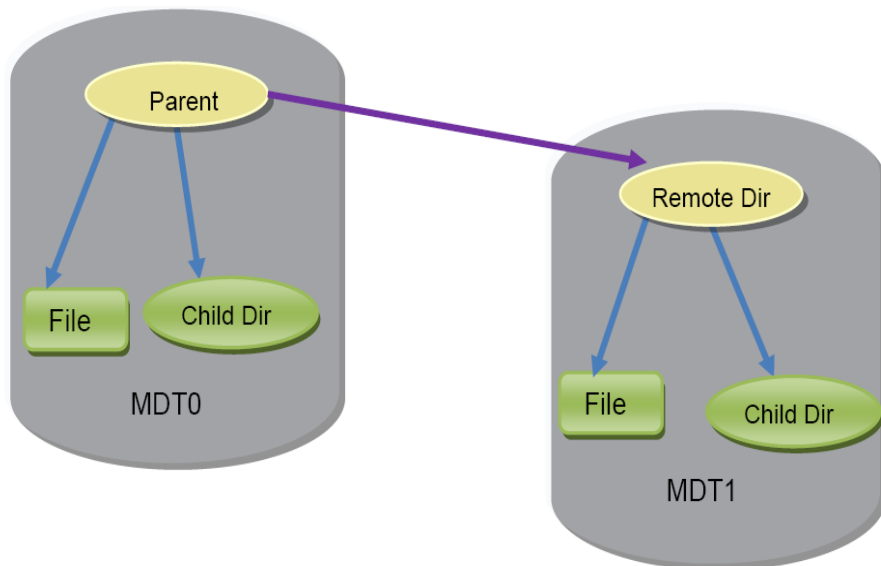
Lustre DNE 是什么

Distributed NamespacE

- 不同于以前的CMD (Clustered MetaData), 采用了全新的设计和实现
- DNE 1, 远端目录, 2.6正式支持
- DNE 2, 条带化目录, 2.8正式支持

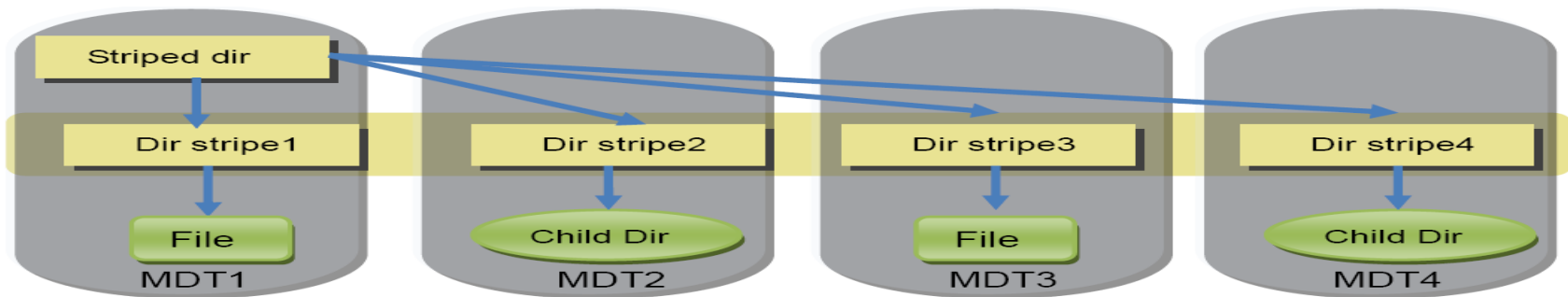
远端目录 remote directory

- 创建目录时指定元数据服务器：`lfs mkdir -i <mdt_index>`
- 普通文件创建在父目录所在元数据服务器
- 涉及多个元数据服务器的修改操作使用同步事务机制
- 不支持跨元数据服务器rename/link



条带化目录 striped directory

- 目录项跨越多个元数据服务器，提高大目录操作并行度
- 在元数据服务器之间迁移文件或目录，避免数据拷贝
- 支持跨元数据服务器rename/link



异步跨元数据服务器操作

- operation (客户端请求) 被分割成一系列updates
- master MDT (接收请求的元数据服务器) 负责分割operation , 并且发送请求到slave MDT (远端元数据服务器)
- 所有的元数据操作都允许跨元数据服务器 , 如果不依赖于其它的操作则异步执行 , 否则使用Commit-On-Share机制保证一致性
- 在不同元数据服务器之间迁移文件或目录 : `lfs migrate -i <mdt_index>`
- Failover需要所有元数据服务器共同参与 , 确定哪些updates在本地重新执行 , 哪些在远端元数据服务器重新执行。

Source Parent

Source child

Target parent

Target child

