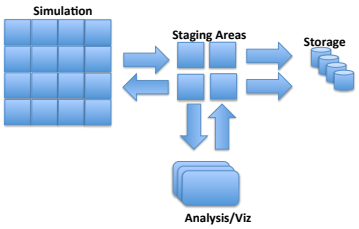
Exceptional service in the national interest

Sandia National Laboratories

D2T: Doubly Distributed Transactions for High Performance and Distributed Computing

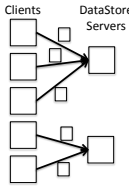
Jai Dayal, Jay Lofstead, Karsten Schwan, Ron Oldfield

jdayal3@gatech.edu, gflofst@sandia.gov, schwan@cc.gatech.edu raoldfi@sandia.gov



Example Architecture

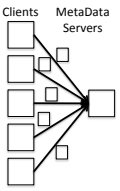
Datastore Service



Txn id	Obj ID	Object
1	123141	...
1	312412	...
1	123412	...
2	412312	...
2	412322	...

- Required Ops.
  - store
  - query
  - retrieve
  - activate objects
- Linking to Metadata
  - Object ID stored as a 'chunk' for a variable
  - All metadata (local array dims and offsets) stored in the metadata service
  - Datastore Service URL stored with each 'chunk'

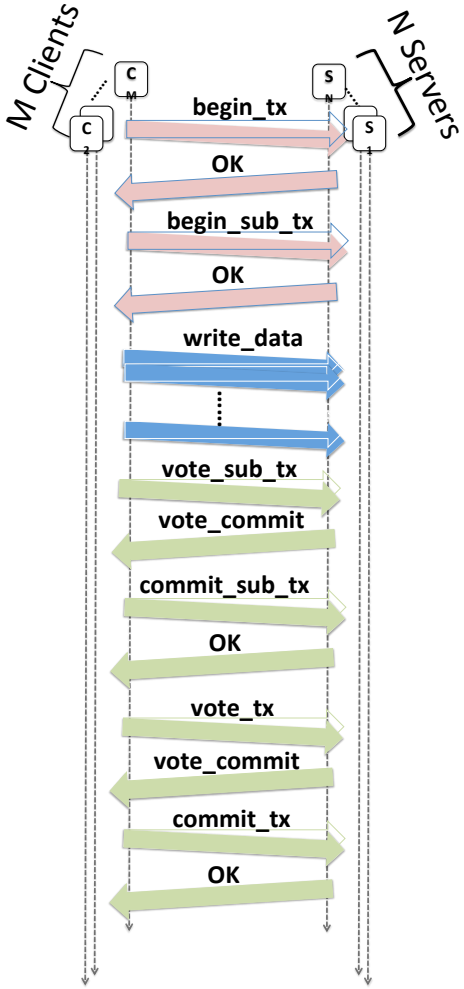
Metadata Service



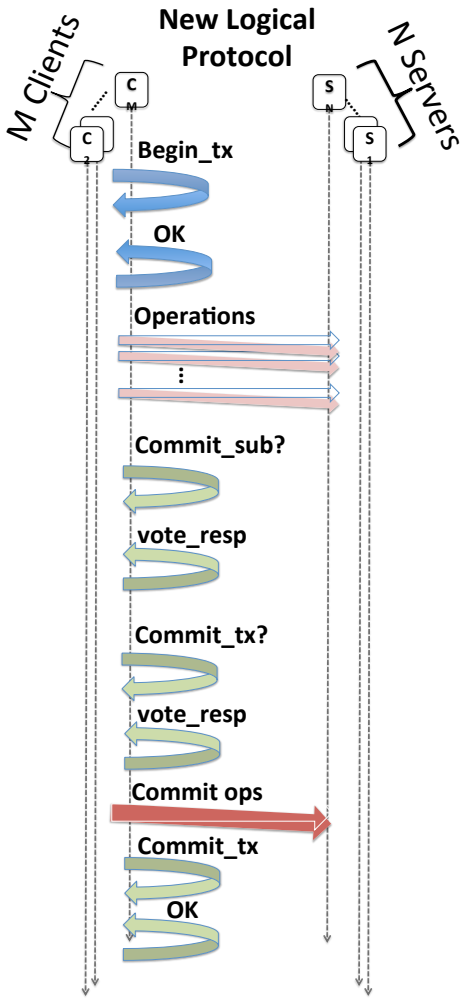
Txn id	Var ID	Obj ID	Chunk Attrs
1	5	123141	...
1	5	312412	...
1	5	123412	...
2	4	412312	...
2	4	412322	...

- Required Ops.
  - create var
  - insert chunk
  - catalog
  - activate var
  - get_chunk_list

Old Logical Protocol



New Logical Protocol



Implementation Changes

- Client-Only Coordinators
  - Eliminates 1-to-1 bottleneck
  - Avoids multiple transports
- Message Merging
  - Responses merged with next message start
  - Message Volumes:

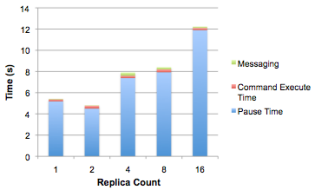
Old Protocol	New Protocol
20M + 12N + 12a	13M + 0N + 2Na

  M = Number of Clients
  N = Number of Servers
  a = messages across
- Sub-Transactions Now Categorized
  - single process now must be done early to detect failures
  - transaction knowledge now fully distributed

Simple API

```
trans = create_transaction(tx_id)
sub_tx1 = create_subtransaction(name, trans)
begin_transaction(trans)
sub_tx2 = create_subtransaction(name, trans)

        datastore_put(data1, id1, sub_tx2)
        datastore_put(data2, id2, sub_tx2)
        metadata_insert ('var', sub_tx1)

vote_subtransaction(sub_tx1)
vote_subtransaction(sub_tx2)
vote_transaction(trans)

        datastore_activate(sub_tx2)
        metadata_activate(sub_tx1)

commit_transaction(trans)
```

Use Case Result: I/O Containers



Time spent executing
Transaction protocol is negligible

U.S. DEPARTMENT OF ENERGY
NNSA National Nuclear Security Administration

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.
SAND No. 2011-XXXXP