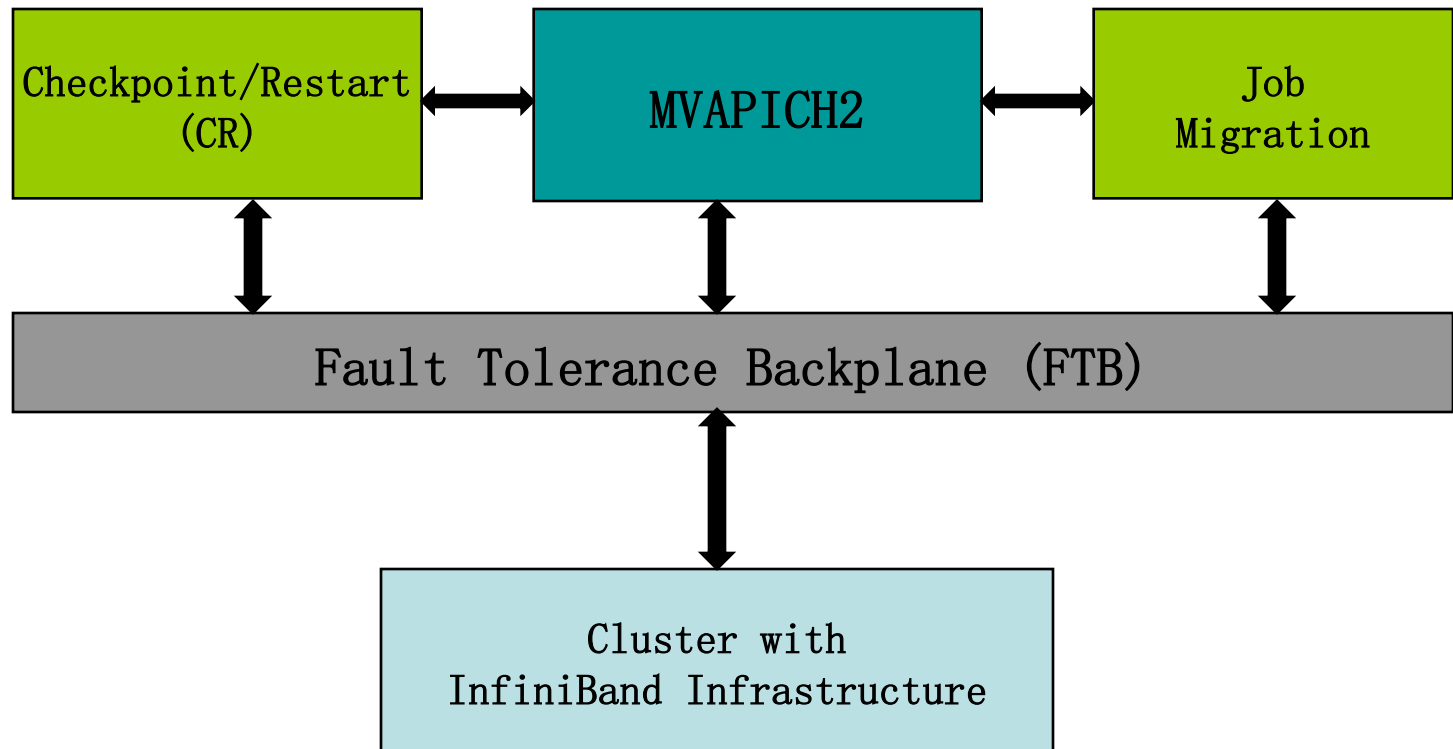# FTB-Enabled
Fault Tolerance in MVAPICH2

The Ohio State University

Fault Tolerance with MVAPICH2

FTB Workflows in MVAPICH2

- Checkpoint-Restart Mechanism

- Fast Check pointing by Write Aggregation with Dynamic Buffer Interleaving

- RDMA-Based Job Migration Framework for MPI over InfiniBand
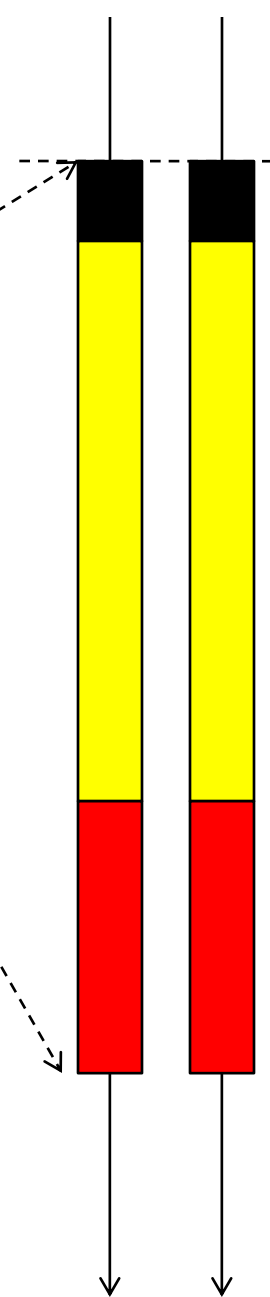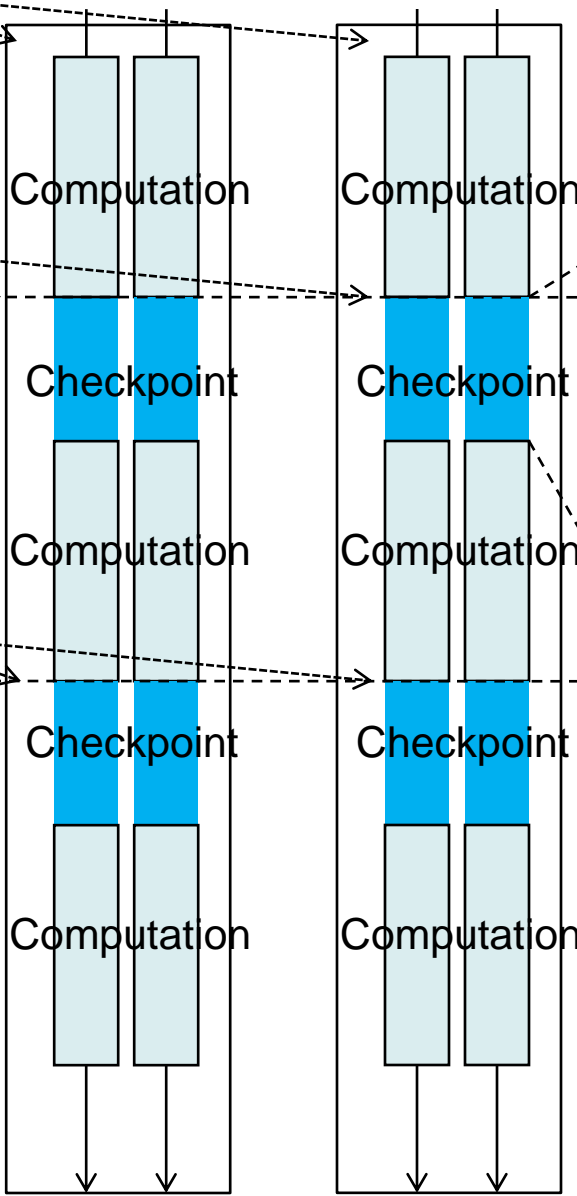
Checkpoint-Restart Workflow

Job Launcher

Start application

Compute Node   Compute Node

Ckpt Rqst

Computation   Computation

Checkpoint   Checkpoint

Computation   Computation

Ckpt Rqst

Checkpoint   Checkpoint

Computation   Computation

Phase 1: Suspend communication between all processes

Phase 2: Use the checkpoint library (BLCR) to checkpoint the individual processes

Phase 3: Re-establish connections between the processes, and continue execution

Write Aggregation Workflow

Application Processes   IO Processes

✓Aggregate many VFS writes into larger chunk writes
  •Reduce disk seeks, improve bandwidth utilization
✓Overlaps application progress with slow checkpoint file writing
  •Improve application execution time

Phase 1

Phase 2

Phase 3

•Application-perceived checkpoint time = *t2 - t1*

•*(t3 - t2)* is overlapping between Application-Progress and File-Write

•Vulnerable to failures in this window
•Provide an interface to poll for File-Write completion

t1

t2

t3
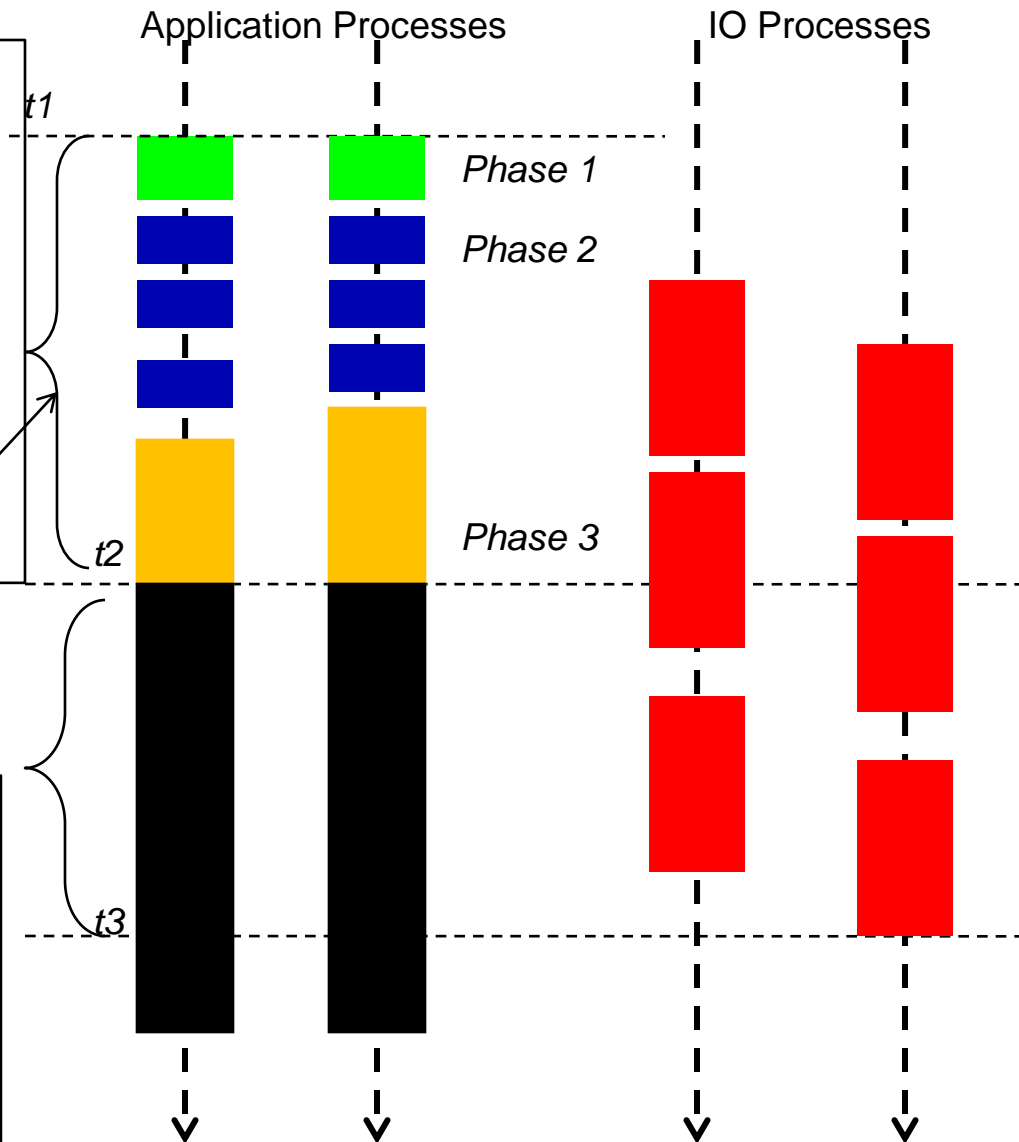
Job Migration Workflow
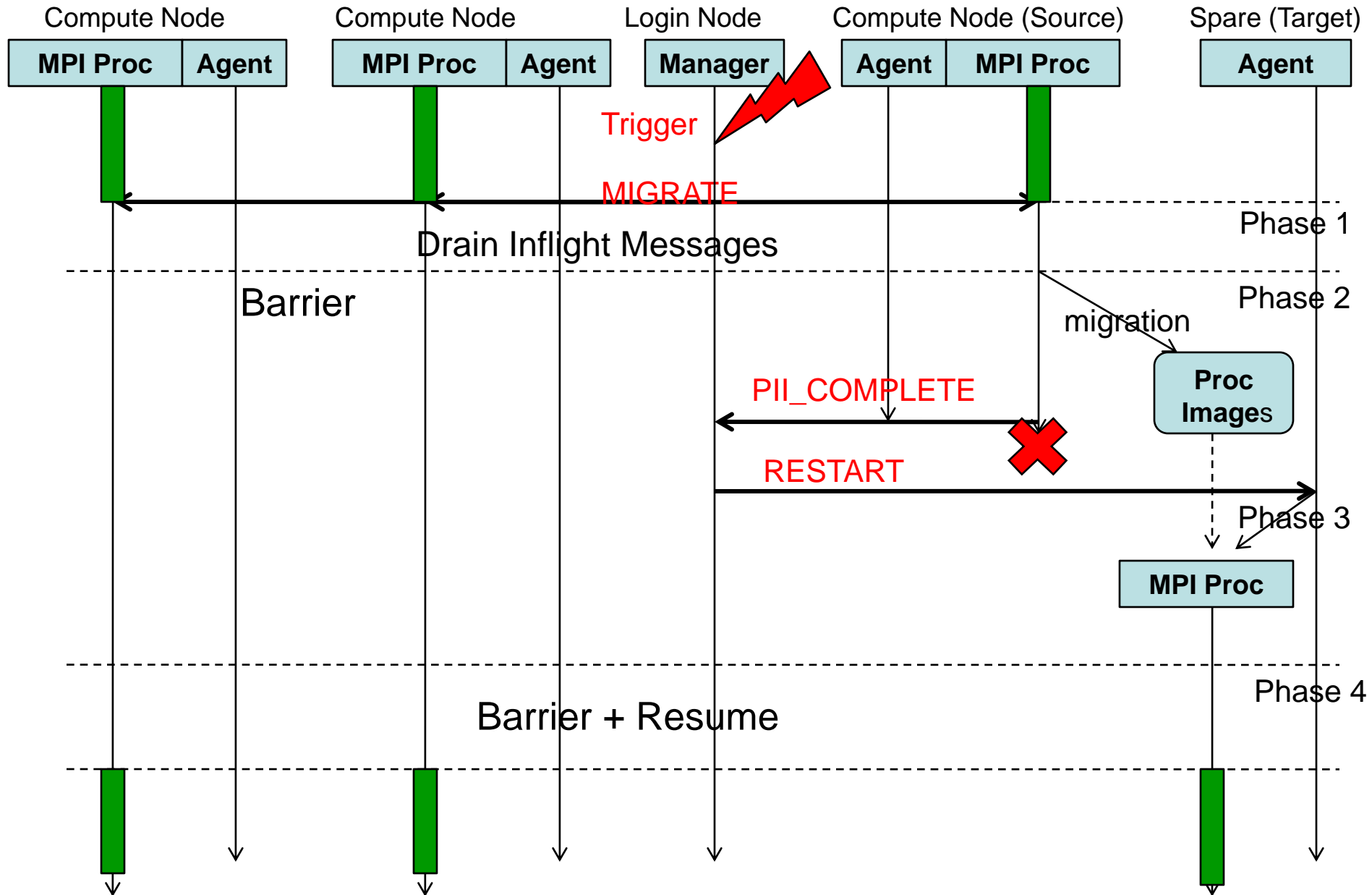
FTB Events in MVAPICH2 CR Framework

- CR_FTB_CHECKPOINT
  - Req. a Checkpoint of the MPI Job
  - Sent by mpirun_rsh

- CR_FTB_CKPT_DONE
  - Checkpoint completed successfully
  - Sent by MPI processes that were able to checkpoint

- CR_FTB_CKPT_FAIL
  - Checkpoint failed
  - Sent by MPI Processes that failed to take the checkpoint

- CR_FTB_RSRT_DONE
  - Restart completed successfully
  - Sent by MPI Processes that were able to restart

- CR_FTB_RSRT_FAIL
  - Restart failed
  - Sent by MPI Processes that failed to restart

- CR_FTB_CKPT_FINALIZE
  - CR module has shutdown
  - Sent by all MPI Processes doing MPI_Finalize()

- CR_FTB_APP_CKPT_REQ
  - Sent by MPI Process from where the user requested a checkpoint through MVAPICH2_Sync_Checkpoint command

FTB Events in MVAPICH2 Migration Framework

- PREDICTOR_NODES_FAILURE
  - An FTB-aware failure predictor informs the Migration Manager about a possible failure
  - Thrown by: Failure Predictor

- REQ_MIGRATE
  - The Migration manager arbitrates a spare node from a given set an provides it to mpispawn
  - Thrown by: Migration Manager

- MIGRATE_DONE
  - The mpispawn informs the mig. manager about successful migration
  - Thrown by: mpispawn on target node

- MIG_NODE_MIGRATED
  - Migration successfully completed
  - Thrown by: Migration Manager

- MIG_NODE_MIGRATE_FAILED
  - Migration failed
  - Thrown by: Migration Manager

Availability and Future Plans

- **Availability**

  - FTB-CR was incorporated in MVAPICH2 1.4

  - FTB-enabled CR with write aggregation and Job Migration has been incorporated in MVAPICH2 1.6

  - It is available at http://mvapich.cse.ohio-state.edu

  - Detailed guidelines for using FTB support in MVAPICH2 is available at https://mvapich.cse.ohio-state.edu/support/user_guide_mvapich2-1.5.1.html#x1-400006.6

- **Future Plans**

  - Optimize the design by analyzing the restart overhead at the target node