# Adaptive MPI

## Dynamic Runtime Support for MPI Applications

Sam White
Charmworks Inc, UIUC

Charmworks

SC 17

# Adaptive MPI

- AMPI is an MPI implementation on top of Charm++. It provides dynamic runtime support for:
  - Process virtualization
  - Communication/computation overlap
  - Dynamic load balancing
  - Online fault tolerance

Charmworks

# Execution Model

- Ranks are User-Level Threads (ULTs)
  - Can have multiple per core
  - Fast to context switch
  - Scheduled based on message delivery
  - Migratable across cores and nodes

Charmworks

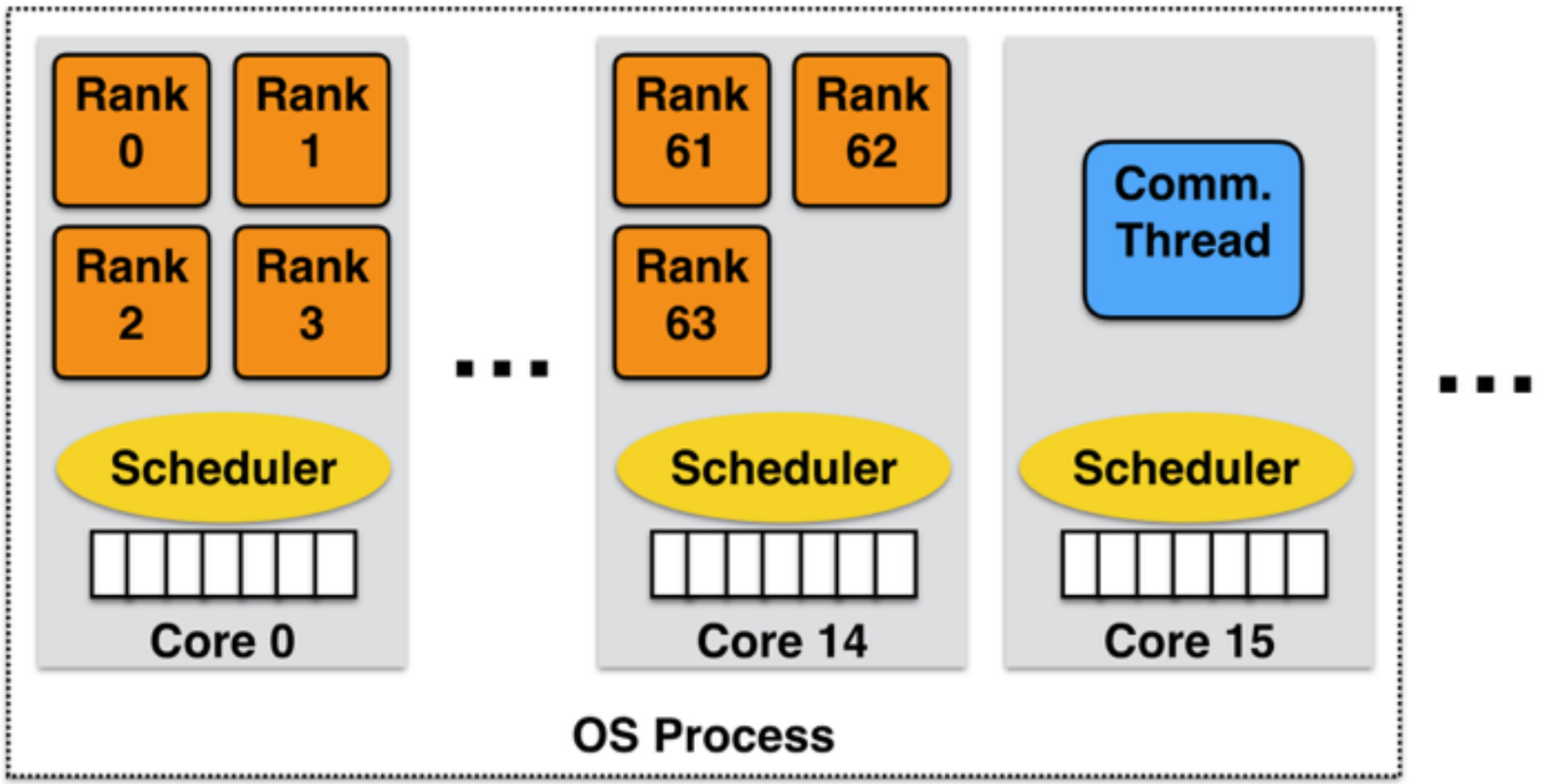
# Programming Model

- AMPI programs are MPI programs, with 2 caveats:
  1. Without mutable global/static variables
     - Or with them properly handled*
  2. Possibly with calls to AMPI's extensions
     - *AMPI_Migrate()*

Charmworks

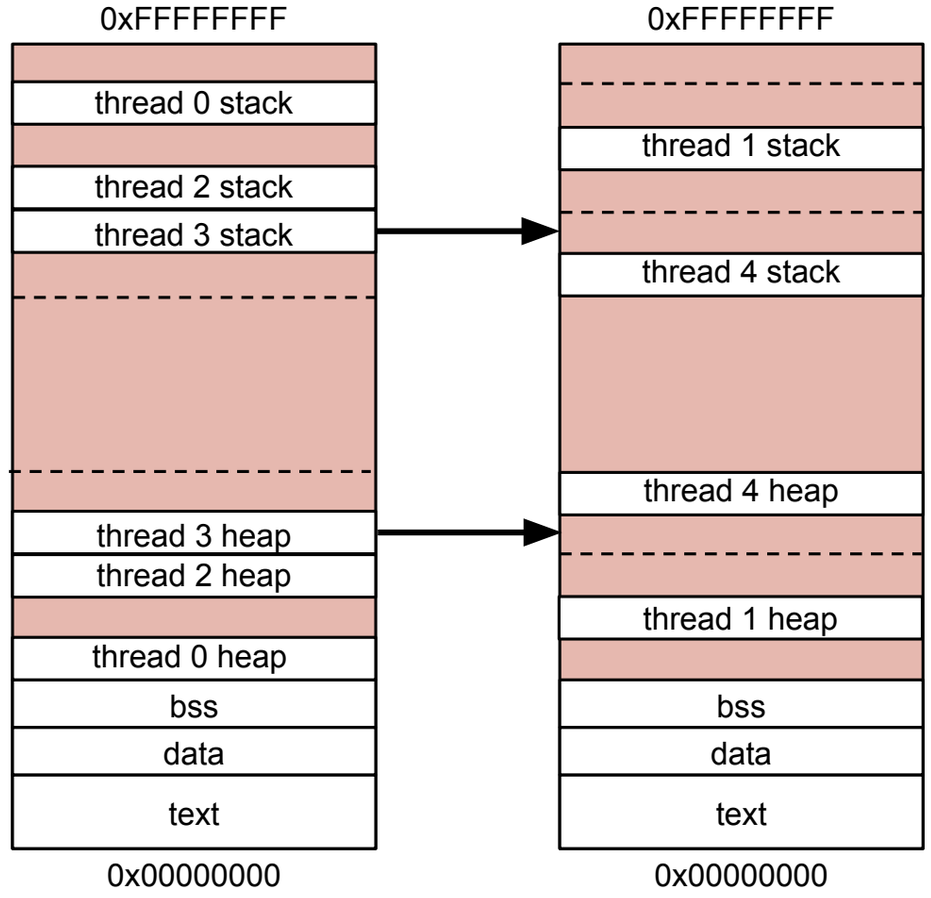
# Execution Model

- Many AMPI ranks share the same address space
- Ranks can migrate across address spaces



# Load Balancing

- AMPI ranks are migratable at runtime between address spaces
  - User-level thread stack + heap
- Isomalloc memory allocator
  - No application-specific serialization code needed




Charmworks

# Fault Tolerance

- Fault recovery is transparent to users:
  - A checkpoint is a migration to storage
    - Storage can be Disk, SSD, NVRAM, etc.
  - The runtime detects failures and restarts from last checkpoint automatically, online
    - No need to go thru the job scheduler to restart

Charmworks

# Recent Work

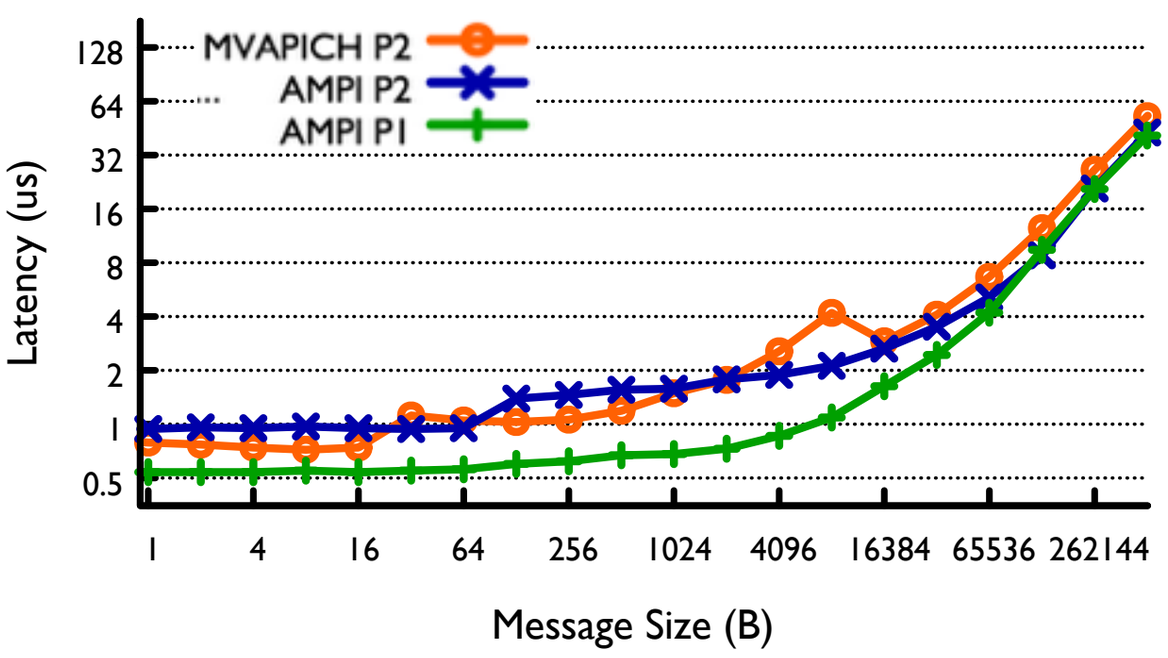
- Charmworks Inc was awarded a Phase I SBIR grant to improve AMPI's functionality and usability
  - Compiler & tools support for global/static variable privatization
  - MPI-3.1 standard conformance
  - Messaging performance: pt2pt & collectives

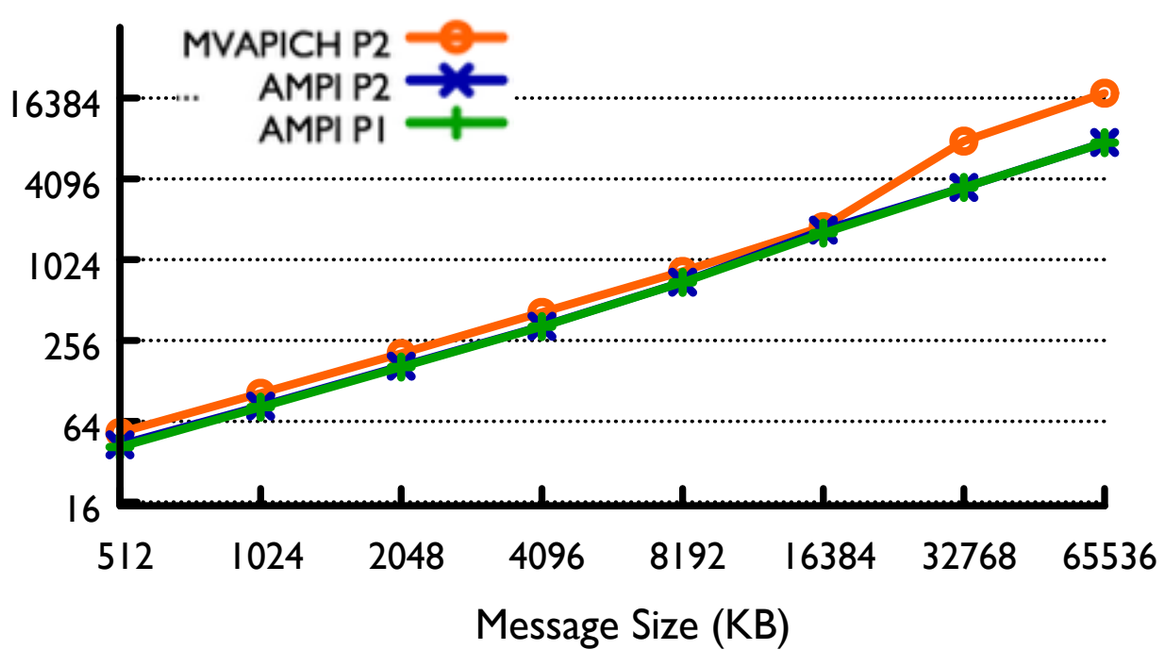
Charmworks

# Recent Work

- AMPI now exploits shared memory within a node
  - Lower latency for msgs ≥4 KB, up to 2.33x faster

OSU MPI Latency Benchmark on Quartz (LLNL)



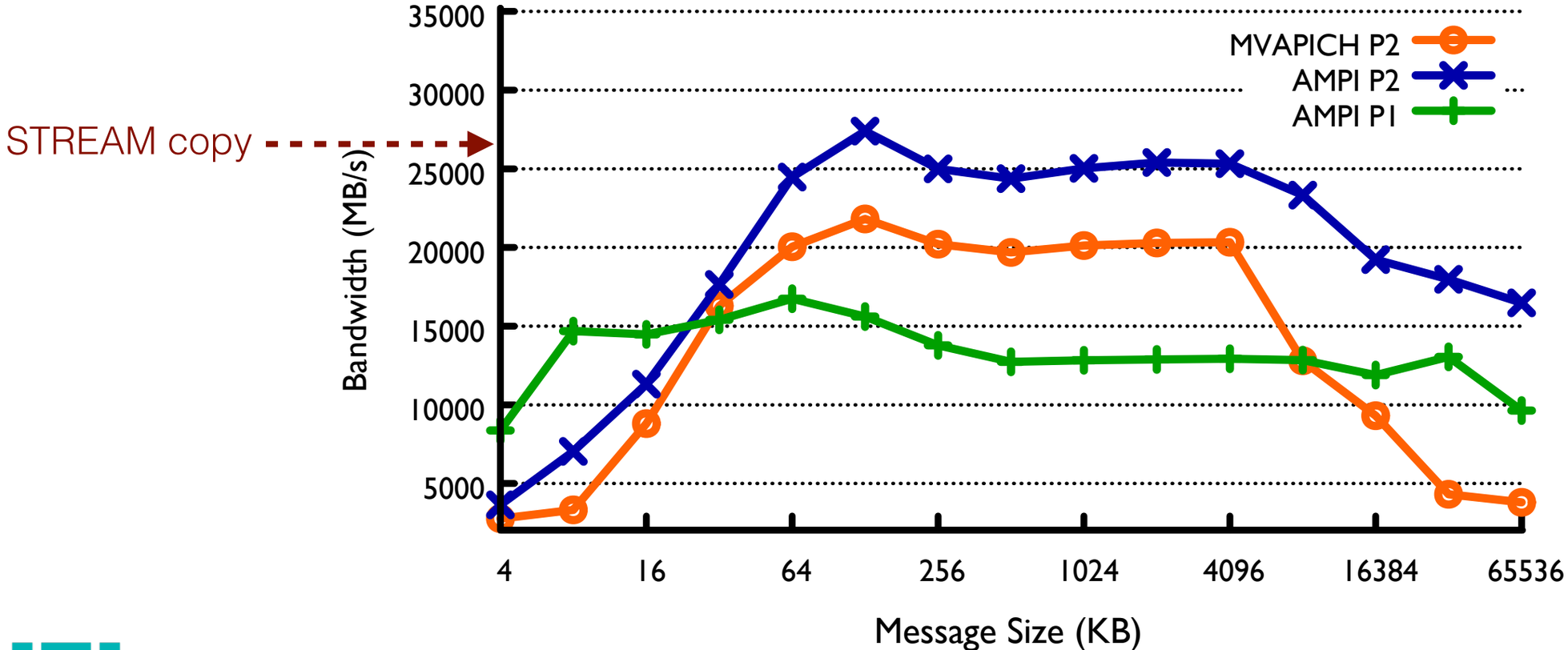




Charmworks

# Recent Work

- AMPI now exploits shared memory within a node
  - Up to 4x higher bandwidth

OSU MPI Bidirectional Bandwidth Benchmark on Quartz (LLNL)




Charmworks

# Summary

- AMPI provides application-independent runtime support for existing MPI applications:
  - Process virtualization
  - Latency tolerance
  - Load balancing
  - Fault tolerance
- See the AMPI manual for more info



# Questions?