HP/ORNL RESEARCH COLLABORATION

HP at Supercomputing 2011

Dick Foster and Zarka Cvetanovic
HP Hyperscale BU / ISS

Research Project SOW

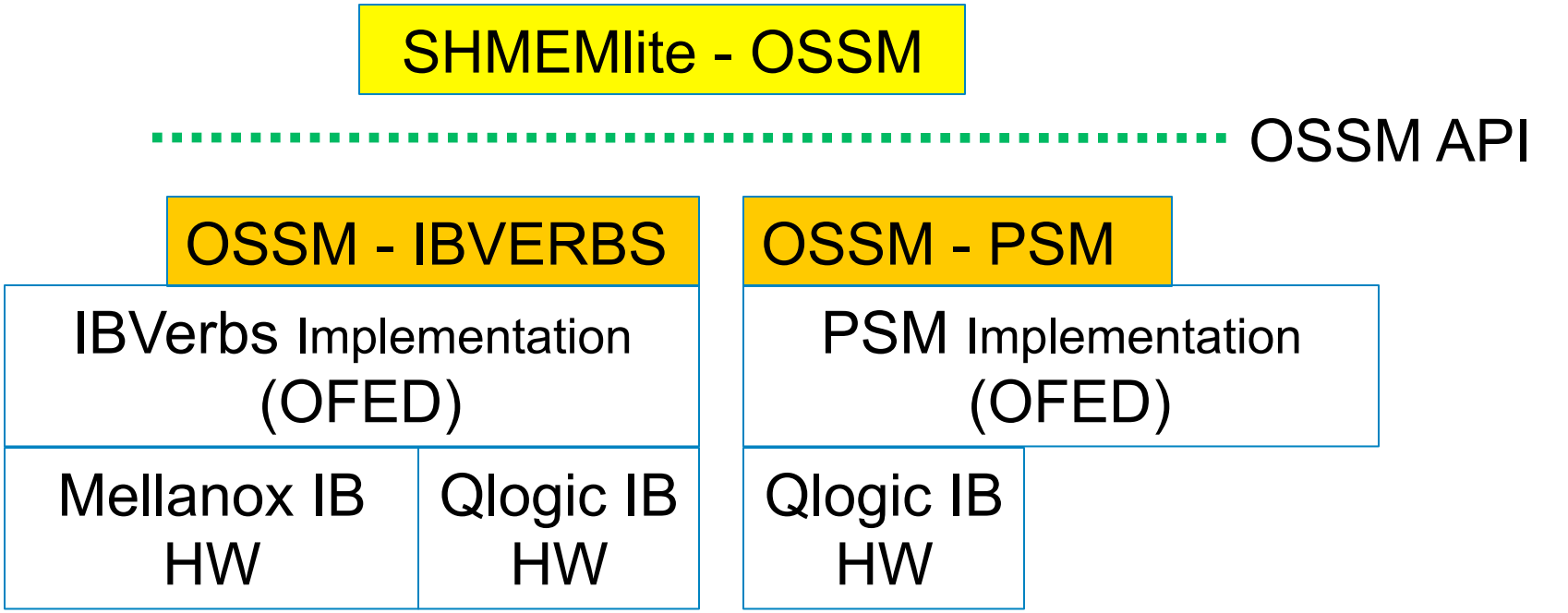
1. Develop software in support of PGAS languages on IB and contribute it to the Open Source Community
2. Investigations of scaling on IB
3. Build a standard suite of benchmarks

Task 2: Investigations of scaling on IB

Learning lessons: InfiniBand scaling studies

- More focus on performance and consistency of short messages
- More focus on workloads with random/high communication
  - A suite of workloads to address all aspects of interconnect
- Improvements in adaptive routing
  - Adaptive routing effective with irregular patterns (hot-spots)
- Improvements in collective operations
  - Hardware acceleration: improve performance and reduce system noise
- Further reductions in memory footprint at scale
- Hybrid approach for InfiniBand transports in hardware and software
- Improvements in fabric management/monitoring
  - Identify congestion and bottlenecks

Task 1: PGAS language support



- OSSM API =“one-sided symmetric memory” API
- SHMEMlite = OpenShmem subset (w/o collectives)
- OSSM-* = OSSM implementations

Supported by the ORNL Extreme Scale Systems Center

Futures

- Optimization work in OSSM API implementations
- Other OSSM API implementations
  - SMP version, Ethernet …
- Other “Symmetric Memory Object” types
  - GPU memories / partitioned global files / …
- Collectives routines for SHMEMlite
  - Complete implementation of Open SHMEM
- “Direct” UPC support layer over OSSM
  - Which UPC?
  - Other PGAS languages?

Supported by the ORNL Extreme Scale Systems Center

Task 3: Suite of Benchmarks

- Benchmarks
  - Scatter / gather (random puts/gets to symmetric memory)
  - RandomRing (like HPCC randomring BW)
  - BarrierHisto (distribution of global barrier timings)
  - LoadBalance (lock operation based load balancing)
  - P2P (point-to-point BW/latency tests)
  - ShmemSort (distributed sort)

- Functional Tests (not part of task)