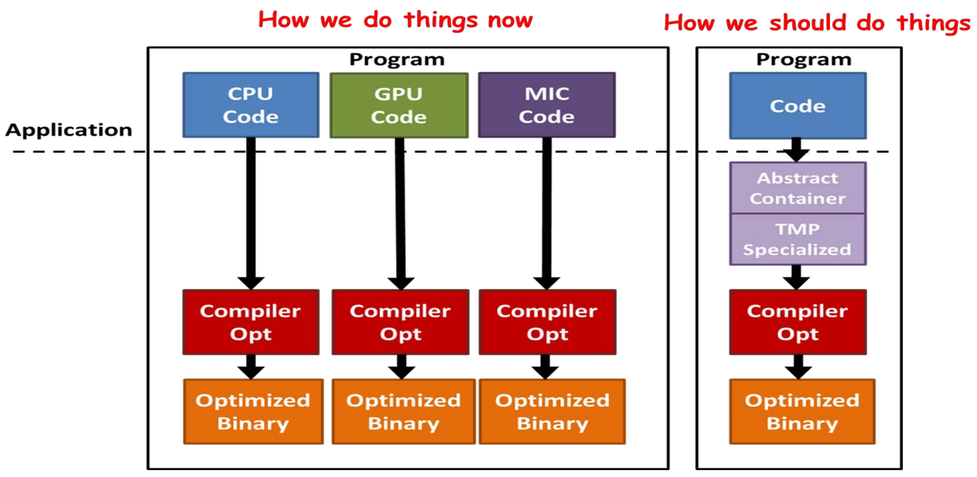
Programmability of Future Architectures

S&T Campaign: Computational Sciences
Advanced Computing Architectures

Jamie Infantolino and James Ross
(410) 278-7121
jamie.k.infantolino.civ@mail.mil
james.r.ross176.civ@mail.mil

Research Objective

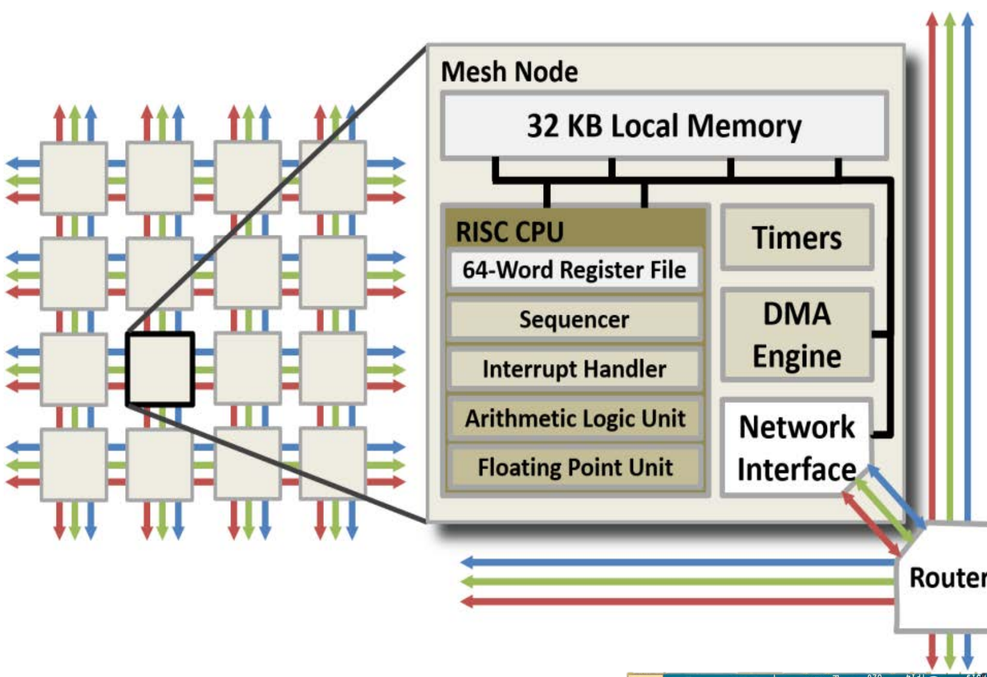
- **Enable performance-portable HPC software application for current and future architectures**
- **Lack of performance portability has been the single greatest barrier to effective use of rapidly emerging HPC architectures. Investigate future architectures from the prospective of programmability for high-performance applications**
- **Develop parallel programming models**
- **Define future architecture requirements based on sound principles for HPC software development**



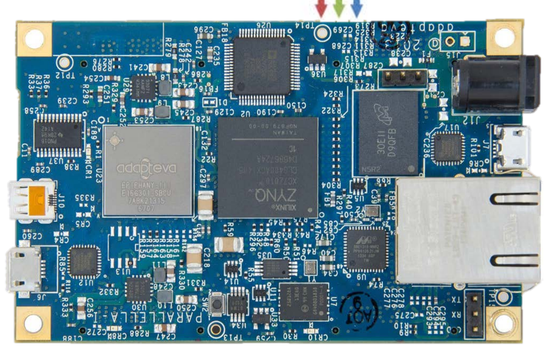
Example of high level design of Templated Meta-Programming

Challenges

- **One of the most long-standing challenges in HPC software design, performance portability for emerging hardware architectures.**
- **Proliferation of architectures requiring substantial re-factoring and porting efforts to achieve even functional if not optimized use**
- **Hardware-first model driven by industry ignores software requirements, drives ineffective utilization of HPC investment**

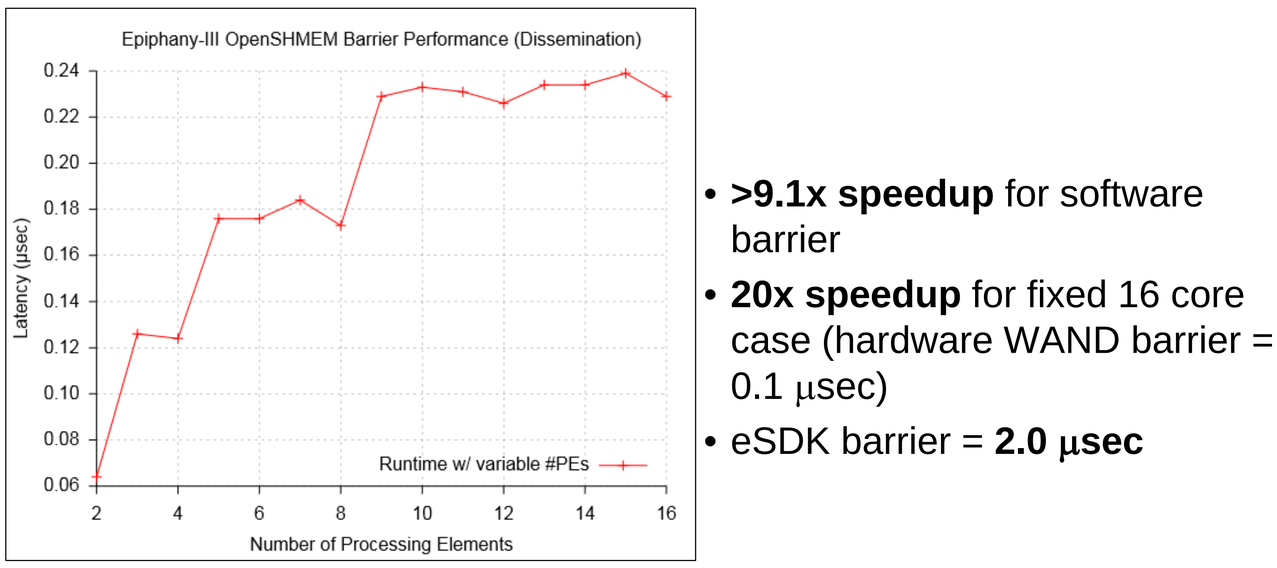


The Adapteva Epiphany RISC array processor architecture provides an example of a new and novel architecture with significant potential for large exascale systems and ultra-low-power tactical HPC platforms. ARL has led the development of parallel programming models for this architecture.



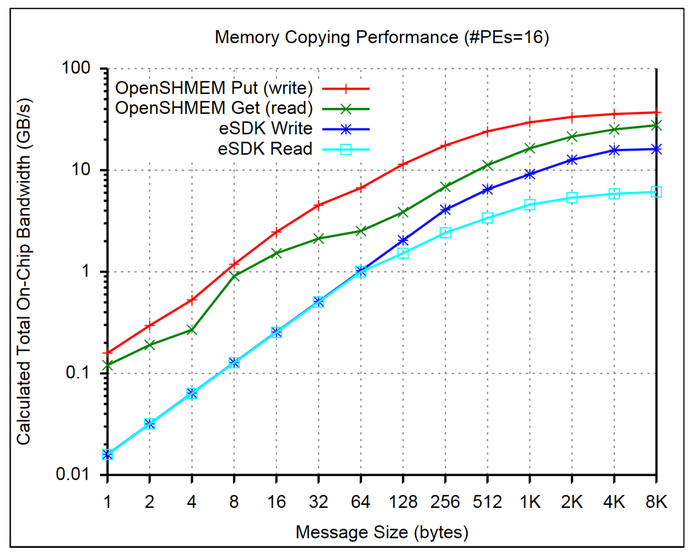
ARL Facilities and Capabilities Available to Support Collaborative Research

- **ARL leads the world in development for this architecture**
- **ARL continues to demonstrate technical leadership in this area (demonstrated first use of standard parallel programming model, OpenSHMEM implementation).**
- **First OpenSHMEM implementation for an embedded device**
- **Enables portability between clusters and embedded codes**
- **ARL now voting member of OpenSHMEM committee**



- **>9.1x speedup** for software barrier
- **20x speedup** for fixed 16 core case (hardware WAND barrier = 0.1 µsec)
- eSDK barrier = **2.0 µsec**

- **2.1-9.9x speedup** for all message sizes
- Peak Put bandwidth = **2.4 GB/s** per core
- Corresponds to alternating ldrd/strd



Complementary Expertise / Facilities / Capabilities Sought in Collaboration

- **Expertise in compiler optimizations**
- **Transition to DARPA-funded 1K-core Epiphany-V chip.**
- **Further areas to consider: expand programs, build out the supported architectures, add an auto-tuning mechanism**