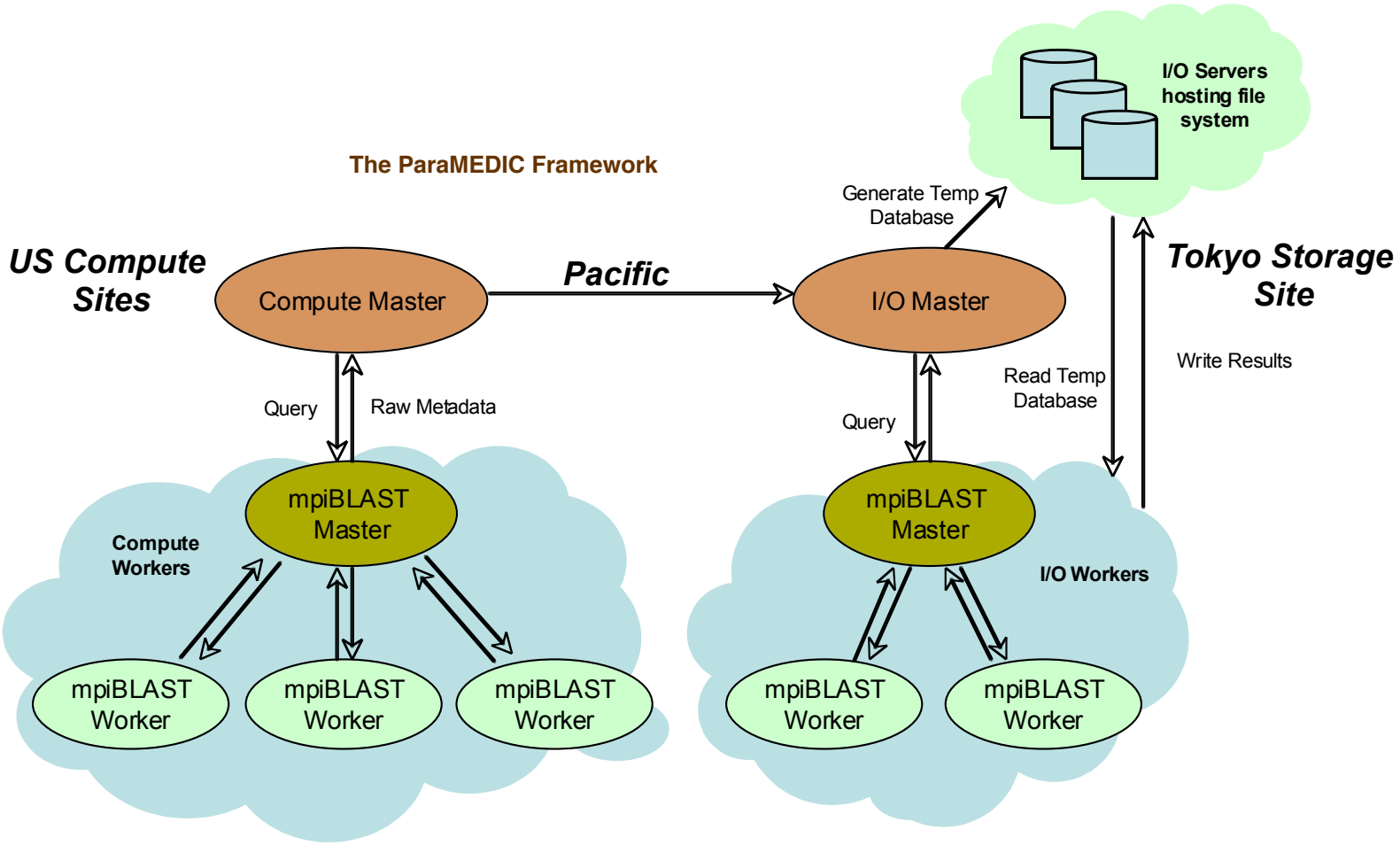
ParaMEDIC: Parallel Metadata Environment for Distributed I/O and Computing

P. Balaji, Argonne National Laboratory
W. Feng and J. Archuleta, Virginia Tech
H. Lin, North Carolina State University

Overview

- Biological Problems of Significance
  - Discover missing genes via sequence-similarity computations (i.e., mpiBLAST, http://www.mpiblast.org/)
  - Generate a complete genome sequence-similarity tree to speed-up future sequence searches
- Our Contributions
  - *Worldwide Supercomputer*
    - Compute: ~12,000 cores across six U.S. supercomputing centers
    - Storage: 0.5-petabyte at the Tokyo Institute of Technology
  - *ParaMEDIC: Parallel Metadata Environment for Distributed I/O and Computing*
    - Decouples computation and I/O and drastically reduces I/O overhead
    - Delivers 90% storage bandwidth utilization
      - A 100x improvement over (vanilla) mpiBLAST

Approach: ParaMEDIC Framework



The ParaMEDIC Paradigm

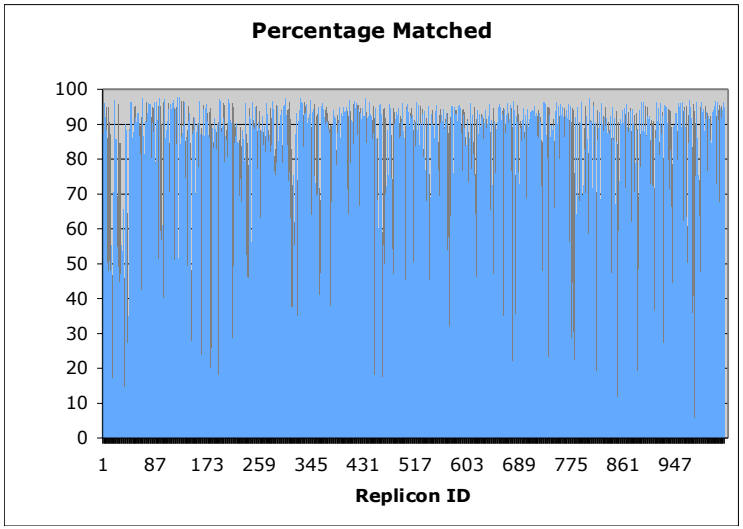
- A *semantics-based* data compression/decompression approach for distributed I/O
  - ParaMEDIC "understands" the output and transmits orders-of-magnitude smaller summarization (metadata) of the output
  - Application-level approach vs. System-level approach
    - Trading portability for performance
    - Serves a specific class of applications
- Example applications that benefit from ParaMEDIC
  - ParaMEDIC accelerated mpiBLAST by more than 25-fold
  - Remote visualization on TeraGrid optimized by 4 to 6-fold
  - Applicable for most applications that have semantic-aware data !
    - Bioinformatics (e.g. ClustalW, MUSCLE), Geophysics

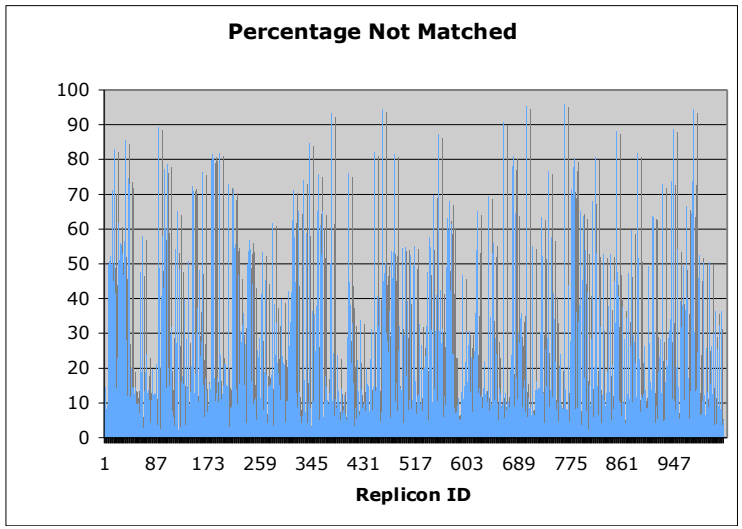
Addressing the Storage Challenge Criteria

- Performance and Scalability
  - Utilized 90% of the storage bandwidth
    - 100x improvement compared to (vanilla) mpiBLAST
  - Results show scalability up to 90% of available storage bandwidth
- Effective Use of Resources
  - Compute:
    - Earlier, compute resources had to wait idly for long, waiting for I/O
    - ParaMEDIC improved compute resource utilization to nearly 100%
  - Storage: Approximately 90% utilization of storage bandwidth
- Innovation
  - A semantics-based data compression/decompression approach for distributed I/O
    - A new approach to enable worldwide supercomputing !

Utility

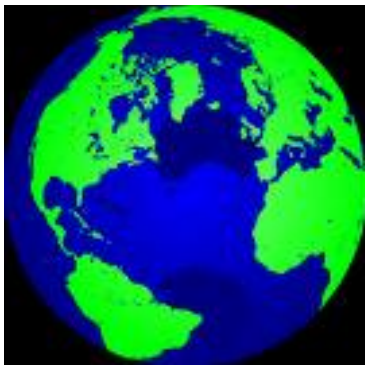
- Analysis of the Similarity Tree
  - Expect that replicons (i.e., chromosomes) will match other replicons reasonably well
  - But many replicons do not match many other replicons
    - 25% of all replicon-replicon searches do not match at all!





Utility

- Semantic-aware metadata gives scientists $2.5*10^{14}$ searches at their finger-tips
  - All metadata results from all searches can fit on iPod Nano
  - "Semantically compressed" 1 Petabyte into 4 Gigabytes (10^6X)
    - Usual compression results in 1 PB into 300 TB (3X)



Semantic Compression