Efficient Network-on-Chip based Multi-core Platform for Hierarchical Parallel Genetic Algorithm

Ming Hsieh Department of Electrical Engineering

Yuankun Xue and Paul Bogdan
CPS Group

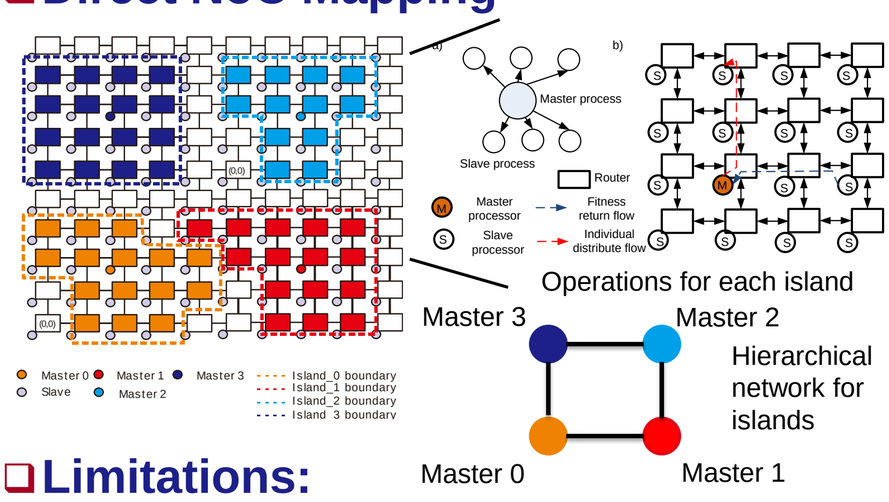
USC Viterbi School of Engineering

Abstract

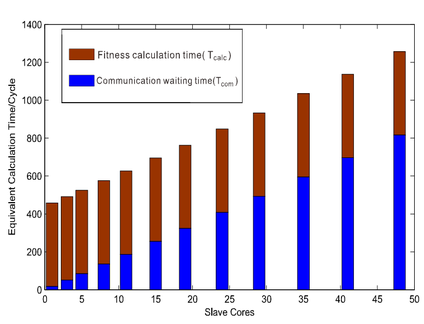
Direct mapping of Hierarchical Parallel Genetic Algorithm to Network-on-Chip for massive parallelism is less favored by performance bottleneck contributed by fixed node bandwidth and low core-utilization which are predicted formally and observed in our experiments. Motivated by the theoretical analysis, we propose a new architecture with dynamic injection bandwidth multiplexing (DIBM) and time-division based island multiplexing (TDIM), to improve the speedup and reduce the hardware overhead. A task-aware adaptive routing algorithm is designed for the proposed architecture to take advantage of multiplexing schemes to further reduce the hardware overhead. We demonstrate the benefits of our approach using the problem of protein folding prediction. Our experimental results show that the proposed NoC architecture achieves up to 240X speedup compared to a single island design. The hardware cost is also reduced by 50% compared to a direct NoC-based HPGA implementation.

Motivation

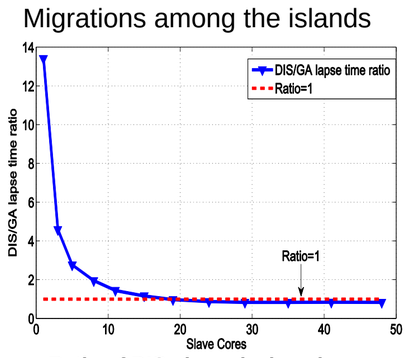
Direct NoC Mapping



Limitations:



Breakdown of distribution and fitness calculation times [Y.-k.Xue, DAC,2014]

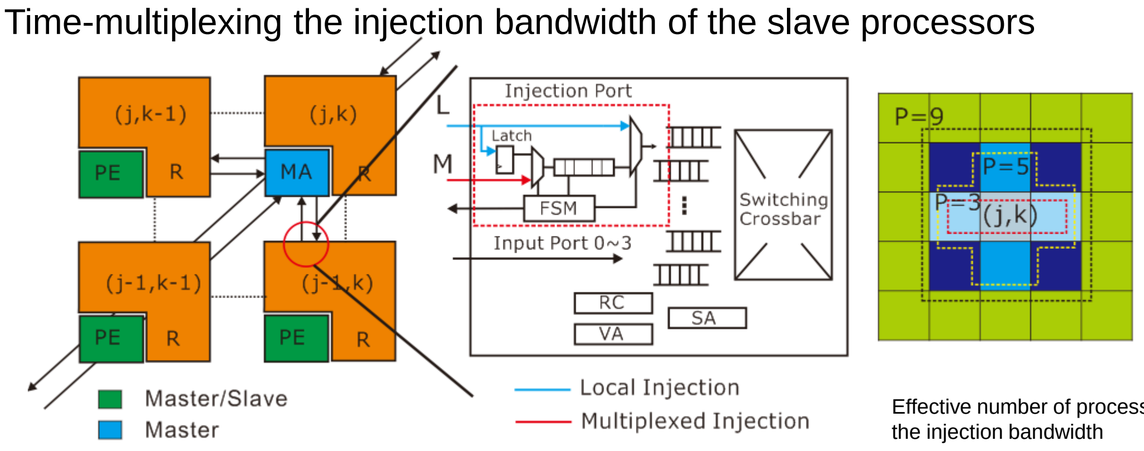


Ratio of DIS phase (using slave processors for calculation) to GA (slave processors are idle) phase

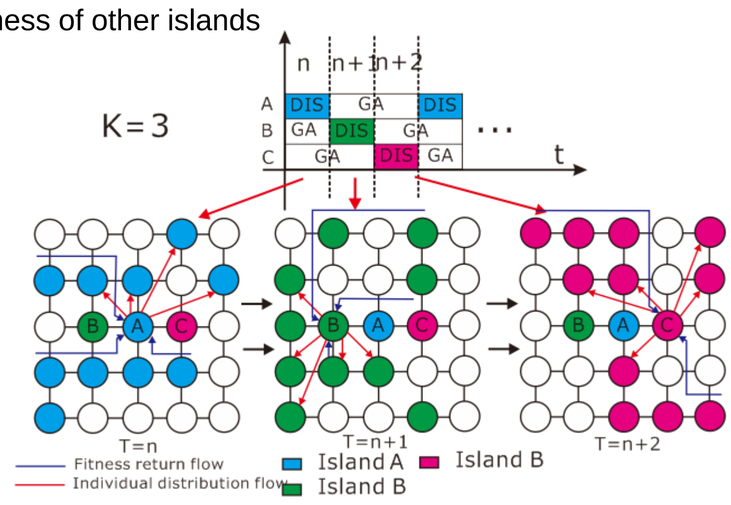
1.Communication cost dominates 2.Low utilization

Contributions

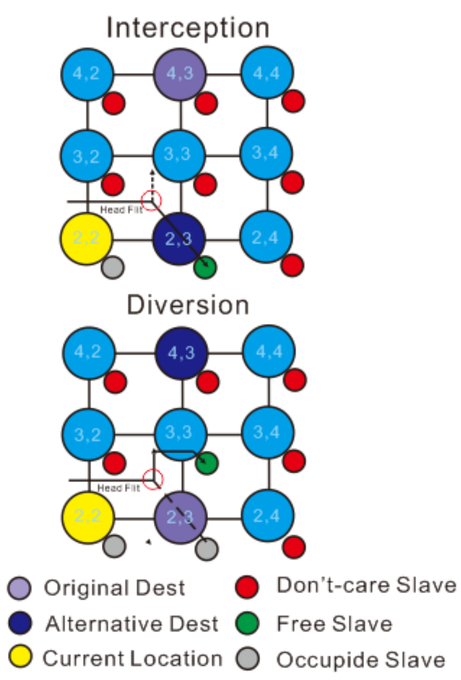
- **Dynamic injection bandwidth multiplexing (DIBM) scheme**
  - **Address limitation 1 by improving master injection bandwidth**
    - Unbalanced utilization of master and slave injection bandwidth
    - Time-multiplexing the injection bandwidth of the slave processors



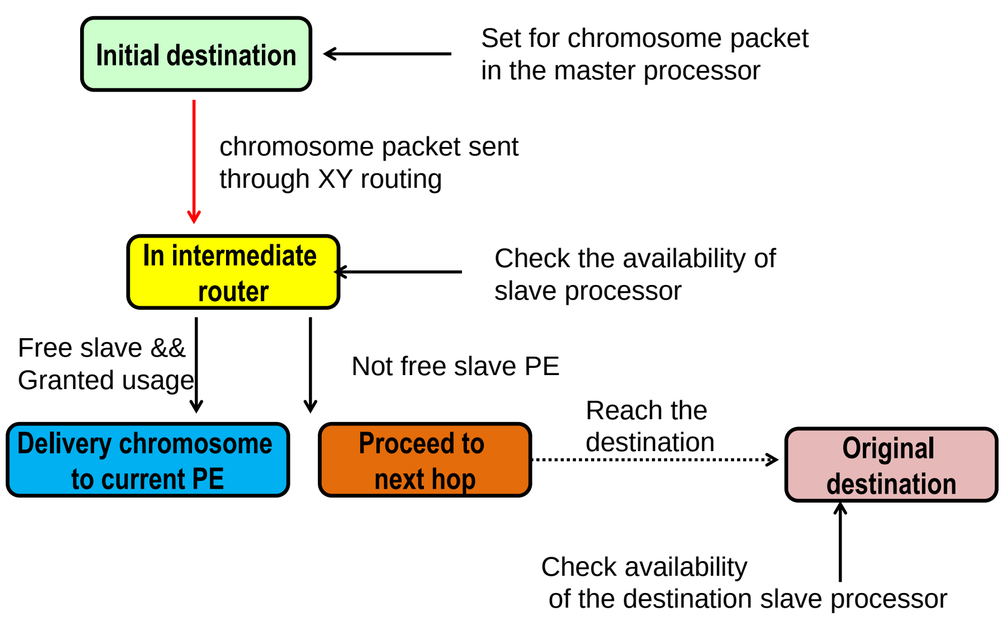
- **Time-division island multiplexing (TIDM) scheme**
  - **Address limitation 2 by improving the slave processor's utilization**
    - Time-sharing the GA phase
      - The slave idle time in one island can be used to calculate the individual fitness of other islands



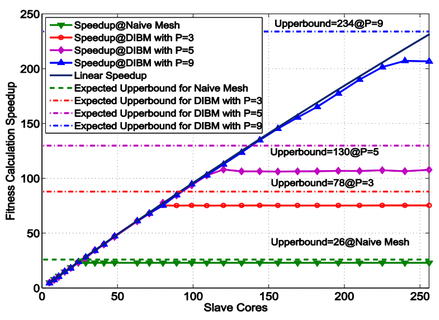
- **Task-aware adaptive routing**
  - **Avoid extra delays in TIDM scheme when two masters distribute individuals simultaneously**
  - **In the routing, packets (individuals) change the destinations adaptively**



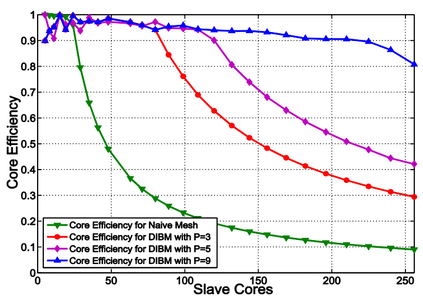
- **Adaptive routing flow:**



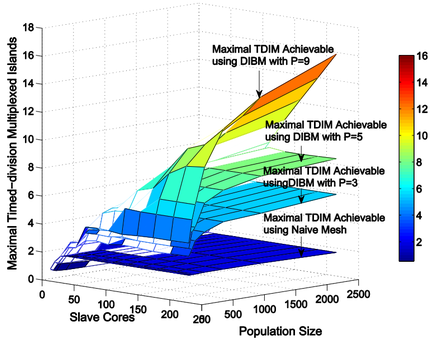
Experimental Results



- Upperbound is obtained by theoretical derivation considering level of multiplexing
- For baseline design (naïve mesh), the speedup tends to saturate early as two types of limitations exist
- For NoC with DIBM, a 75X-206X speedup can be obtained

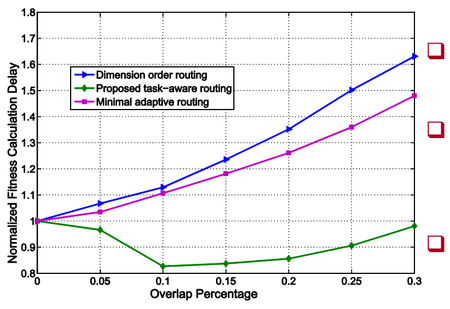


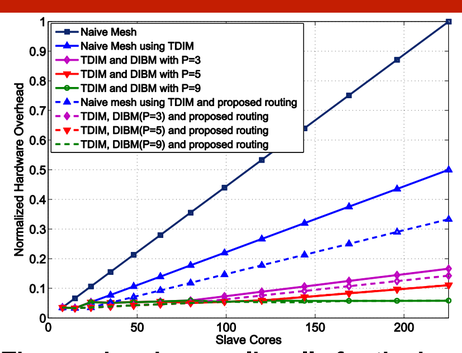
- Core utilization in baseline (naïve mesh) drops significantly
- TDIM schemes efficiently improves the slave cores' utilization as the number of slave processor increases



- Maximal number of islands that can be multiplexed on a physical island
- Up to 14 islands could be multiplexed over the same network under experimental settings

- Overlap ratio is the percentage of time DIS phase of two logic islands overlap
- The proposed routing algorithm achieves 10-15% reduction in the fitness calculation time





- The overhead grows lineally for the baseline design
- The proposed TDIM scheme greatly reduced the number of PEs and routers required by sharing the same resources in the island
- Combining DIBM, the proposed routing with TDIM further reduces the hardware requirement

Yuankun Xue, Zhiliang Qian, Guopeng Wei, Paul Bogdan, Chi-Ying Tsui, Radu Marculescu, "An Efficient Network-on-Chip (NoC) based Multicore Platform for Hierarchical Parallel Genetic Algorithms", NOCS 2014, {yuankunx, pbogdan}@usc.edu

Ming Hsieh Institute
Ming Hsieh Department of Electrical Engineering