

# **A Study on Application-aware Dynamic Interconnects**

Keichi Takahashi

August 6, 2018

# List of Publications



# Summary



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Background and Motivation . . . . .	1
1.2	Objective . . . . .	1
1.3	Organization of the Dissertation . . . . .	1
1.4	Current Trend of High-performance Computing Systems . . . . .	1
1.4.1	Cluster Architecture . . . . .	1
1.4.2	Interconnect . . . . .	1
1.4.3	Programming Model . . . . .	1
1.5	Related Work . . . . .	1
<b>2</b>	<b>Toolset for Analyzing Application-aware Dynamic Interconnects</b>	<b>3</b>
2.1	Introduction . . . . .	3
2.2	Problem and Objective . . . . .	3
2.3	Toolset for Analyzing Application-aware Dynamic Interconnects . . . . .	3
2.3.1	Fine-grained MPI Profiler . . . . .	3
2.3.2	Lightweight Interconnect Simulator . . . . .	3
2.4	Evaluation . . . . .	3
2.4.1	Experimental Environment . . . . .	3
2.4.2	Accuracy of Traffic Estimation . . . . .	3
2.4.3	Execution Time of Simulation . . . . .	3
2.4.4	Profiling Overhead . . . . .	3
2.5	Conclusion . . . . .	3
<b>3</b>	<b>SDN-enhanced MPI Collective</b>	<b>5</b>
3.1	Introduction . . . . .	5
3.2	Problem and Objective . . . . .	5
3.2.1	Related Work on Optimization of MPI Collectives . . . . .	5
3.2.2	Problem of Conventional Traffic Balancing Methods . . . . .	5

3.3	SDN-enhanced MPI_Allreduce . . . . .	5
3.3.1	Design . . . . .	5
3.3.2	Implementation . . . . .	5
3.4	Evaluation . . . . .	5
3.4.1	Experimental Environment . . . . .	5
3.4.2	Execution Time of SDN-enhanced MPI_Allreduce . . . . .	5
3.5	Conclusion . . . . .	5
<b>4</b>	<b>Coordination Mechanism of Communication and Computation</b>	<b>7</b>
4.1	Introduction . . . . .	7
4.2	Problem and Objective . . . . .	7
4.3	Coordination Mechanism of Communication and Computation . . . . .	7
4.3.1	Basic Idea . . . . .	7
4.3.2	Intra-node Architecture . . . . .	7
4.3.3	Inter-node Architecture . . . . .	7
4.4	Evaluation . . . . .	7
4.4.1	Experimental Environment . . . . .	7
4.4.2	Verification of the Coordination . . . . .	7
4.4.3	Overhead Incurred by the Coordination Mechanism . . . . .	7
4.5	Conclusion . . . . .	7
<b>5</b>	<b>Conclusion</b>	<b>9</b>
5.1	Concluding Remarks . . . . .	9
5.2	Future Directions . . . . .	9

# **1 Introduction**

## **1.1 Background and Motivation**

## **1.2 Objective**

## **1.3 Organization of the Dissertation**

## **1.4 Current Trend of High-performance Computing Systems**

### **1.4.1 Cluster Architecture**

### **1.4.2 Interconnect**

### **1.4.3 Programming Model**

## **1.5 Related Work**





## **2 Toolset for Analyzing Application-aware Dynamic Interconnects**

### **2.1 Introduction**

### **2.2 Problem and Objective**

### **2.3 Toolset for Analyzing Application-aware Dynamic Interconnects**

#### **2.3.1 Fine-grained MPI Profiler**

#### **2.3.2 Lightweight Interconnect Simulator**

### **2.4 Evaluation**

#### **2.4.1 Experimental Environment**

#### **2.4.2 Accuracy of Traffic Estimation**

#### **2.4.3 Execution Time of Simulation**

#### **2.4.4 Profiling Overhead**

### **2.5 Conclusion**



## **3 SDN-enhanced MPI Collective**

### **3.1 Introduction**

### **3.2 Problem and Objective**

#### **3.2.1 Related Work on Optimization of MPI Collectives**

#### **3.2.2 Problem of Conventional Traffic Balancing Methods**

### **3.3 SDN-enhanced MPI\_Allreduce**

#### **3.3.1 Design**

#### **3.3.2 Implementation**

### **3.4 Evaluation**

#### **3.4.1 Experimental Environment**

#### **3.4.2 Execution Time of SDN-enhanced MPI\_Allreduce**

### **3.5 Conclusion**



## **4 Coordination Mechanism of Communication and Computation**

### **4.1 Introduction**

### **4.2 Problem and Objective**

### **4.3 Coordination Mechanism of Communication and Computation**

#### **4.3.1 Basic Idea**

#### **4.3.2 Intra-node Architecture**

#### **4.3.3 Inter-node Architecture**

### **4.4 Evaluation**

#### **4.4.1 Experimental Environment**

#### **4.4.2 Verification of the Coordination**

#### **4.4.3 Overhead Incurred by the Coordination Mechanism**

### **4.5 Conclusion**



# **5 Conclusion**

## **5.1 Concluding Remarks**

## **5.2 Future Directions**





# Acknowledgements