Optimal Mappings of m Dimensional FFT Communication to k Dimensional Mesh for Arbitrary m and k

Z. GEORGE MOU and XIAOJING WANG

Department of Computer Science and Center for Complex Systems
Brandeis University, Waltham, MA 02254, USA
{mou,wang}@cs.brandeis.edu

Abstract. The FFT communication patterns are important to not only FFT algorithms, but also many other algorithms over one or higher dimensional. The mapping of m dimensional FFT communication to k dimensional mesh has previously been considered only for the following special cases (a) m = 1 or 2, k = 1 or 2, (b) m = 1 or 2, k = log(n) where n is the size of the machine. In this paper, we present the optimal mappings of m dimensional FFT communication onto k dimensional mesh for arbitrary m and k. The mappings are optimal since the communication distances in the logarithmic steps sum to exactly the diameter of the mesh regardless of the dimension or the shape of the mesh. An m-k shuffle permutation, which subsumes perfect shuffle, is introduced and used to derive some of the optimal mappings. As a by-product, an optimal broadcast algorithm over any dimensional mesh, including binary hypercube as a special case, is presented.

1 Introduction

The butterfly communication patterns found in FFT algorithms, which we will refer to as FFT communications, are important to not only the FFT algorithms but also many other divide-and-conquer algorithms for a broad class of numerical problems including reduction, scan, polynomial evaluation, linear difference equations, and bitonic sort. On the other hand, mesh is an important underlying topology for a wide range of parallel architecture, ranging from the one dimensional linear array to the binary hypercube which is no more than a mesh of logarithmic dimensional mesh. The study of the mappings from FFT communication to mesh architectures however has been very limited in the past. Let m be the dimension of the FFT communications, and k the dimension of the mesh, then the previous studies only cover the following special cases:

- m = 1, 2, k = 1, 2 [3, 1, 2].
- m = 1, 2, k = log(n) [11, 13, 5, 12, 10].
- m = 1, 1 ≤ k ≤ log(n) [7, 8].

For the case of k = log(n), we have included Stone's work on perfect shuffle [12] and Preparata's work on cube-connected-shuffle [10] since the well-known isomorphism between the two architectures and a binary hypercube [14]. It thus can be observed that the mapping has not been studied for high dimensional FFT and the meshes
with the dimensionality between the two extremes \(- k = 1\) and \(2\) and \(k = \log(n)\). Moreover, there is a lack of general approach to the problem for different dimensions of the FFT and the mesh, and the inherent connections between the results for different dimensions cannot be easily seen.

In this paper, we study the mapping from \(m\) dimensional (\(m\)-d) FFT to \(k\) dimensional (\(k\)-d) mesh for arbitrary \(m\) and \(k\). The main results generalize the previous work in [1, 3, 2, 11, 13, 7, 8] and include

- A unified framework to study the mapping for arbitrary \(m\) and \(k\).
- A family of \((\log(n))!\) mappings for arbitrary dimensions of the FFT and the mesh with size \(n\).
- Establish the lower bound of the communication cost, and prove the optimality of the family of mappings.

As a by-product of the proof in Section 3, we present optimal broadcast algorithms which can be used to (1) broadcast over \(k\)-d mesh for any \(k\), including binary hypercube as a special case; (2) broadcast over a given dimension of an \(m\) dimensional array distributed over a \(k\)-d mesh for any \(m\) and \(k\). This result thus generalizes the well-known result on broadcasting in [4].

This paper is organized as follows. We begin by introducing some basic notions about mesh and FFT communication in Section 2. In Section 3, we establish the lower bound of the FFT communication on mesh. Section 4 and 5 introduces two specific mappings from FFT communications to mesh for arbitrary dimensions, present the implementations of the communications on mesh, and prove their optimality. In Section 6 we introduce a family of optimal mappings for which the previous two are members. In Section 7, we generalize the results in previous sections to allow arbitrary shape of the FFT array and the mesh. Some of the related results that cannot be reported here are mentioned in Section 8, where other comments are made as well.

2 Preliminary

2.1 \(k\) dimensional mesh

The topology of the mesh A \(k\) dimensional mesh with the shape \(P_{k-1} \times \ldots, P_0\) consists of processors in the set \(P^k\) of the form

\[
P = \{p = (p_{k-1}, \ldots, p_0) \mid 0 \leq p_i P_i - 1, \text{for } i = 0 \text{ to } k - 1\}
\]

where \(P_i\) is the size of the mesh along the \(i\)th dimension, \(p_i\) the coordinate of processors \(p\) along the \(i\)th dimension. Two processors \(u, v \in P^k\) are directly connected, denoted by \(uLv\), if only only if their coordinates differ along one dimension by one, namely

\[
uLv \iff |u_i - v_i| = 1, \text{ and } u_j - v_j = 0 \leq i, j < k - 1, j \neq i,
\]

The total number of processors \(s = |P| = \prod_{i=0}^{k-1} P_i\) is referred to as the size of the mesh. A metric function \(D : P^k \times P^k \rightarrow \mathbb{INT}\) is introduced to measure the distance between processors. Given two processors \(u = (u_0, u_1, \ldots, u_{k-1})\) and \(v = (v_0, v_1, \ldots, v_{k-1})\) in \(P^k\) the function \(D\) is

\[
D(u, v) = \sum_{i=0}^{k-1} |u_i - v_i|
\]