Assignment 2: Parallel Traveling Salesman

CMSC 502: Parallel Algorithms

Steven M. Hernandez

Department of Computer Science

Virginia Commonwealth University

Richmond, VA USA

hernandezsm@vcu.edu

**Introduction**

In this project, we looked to implement a fully parallel MPI version of the Traveling Salesman Problem (TSP). In the previous iteration, we parallelized calculating sub-TSP solutions per process, but did not handle stitching in parallel; resulting in inefficient computation times. In this project, we looked to increase the parallelism and thus decrease computation times by stitching sub-TSP solutions together in parallel, Additionally, the method used for stitching provides much fewer inversions (although it may be assumed that the solution should result in no inversions).

In addition to this main goal of this project, we also look to improve the general abilities of the MPI implementation from the previous implementation. We are successful in this case. The previous method caused errors in certain cases when reversing the dynamic programming solution when for example floating point math would change values by small decimal points thus resulting in incorrect equality evaluations. Additionally, issues were solved in the use of infinity on the server when compared to running on a local machine.

Finally, the implementation is more generalized from the previous implementation in that the implementation can work on any square number of processors. Given the max number of processes possible on the server is *40*, previously we were only able to run the implementation on number of processes where *n* is a positive integer which in our case only allowed the implementation to work when . This new implementation allows for any square number of processors which means we can achieve much greater variety of processors; again, in our case . This gives us more flexibility in that we are using a higher percentage of available resources available to us.

**Process**

We begin by looking at the process taken by the MPI implementation. Each process first selects a number of random cities within a 100x100 range which can be seen in Figure 1 where all 36 processes select 23 random city locations. This occurs without communication, followed directly by calculating the sub-TSP solution for the random city locations selected, Figure 2. At this point, each process begins making first communications with one another, first stitching complete rows by column as per the specification. For each process, the process will either send to a destination or received from some source. Because of this, we should in essence be able to avoid any barriers given adequate tagging of messages, however this was not implemented here. After complete rows are stitched, these rows are stitched in the same algorithm. Note, when we have numbers of processors that are not we can’t guarantee as many processors are doing work as expected. For example, Figure 8 shows the case when or more simply, when dimensionality of each row and thus each column is . We see in this figure; a simple pattern appears in each time where each process is either sending or receiving a message (or the process has completed all work and thus can be closed). However, given this implementation allows for any square number of processes we cannot guarantee these cases. Figure 9 shows the case when in the first-time instances, P7 would normally require a message from P8, however P8 does not exist thus causing a deadlock. Figure 10 as well shows a case where deadlock is possible, however in this case, this isn’t in the first-time instance t1, but instead later on in time instance t2. We can see that each of the three cases in these figures, three communication periods are required even though less work and less communication is actually occurring as the number of processors drops from eight down to seven until 6 in Figure 10 (do note that this would also be the case for p=5 but not p=4 which itself would require only 2 communication steps). Luckily, this won’t affect the computation too much, but it should be considered. For example, if we had a higher number of processors on the server, we would be able to have p=64 which would might be a better case. Luckily, the logic involved in determining whether or not a process should send, receive or wait until the next time instances only results in constant overhead *O(1)*.

**Analysis**

The algorithm was run on each combination of processors with number of cities per processor . Figure 4 shows Time taken for calculating TSP solution for each combination. As we reach a threshold between 20-25 cities per processor, the times taken quickly increase. Before this point, the time taken is rather constant implying the communication overhead of communicating between each process takes the highest amount of time. We can see very obviously that while a higher number of processes takes more time with lower number of cities per block (Figure 5), we see these higher number of processes shine when there are a high number of cities per block, beating any attempt with fewer numbers of processes. Notice, I found that the algorithm for stitching did not result in no inversions, so the code did require handling inversions. The same analysis from above can be seen when handling inversion in Figure 6 and Figure 7 respectively.

**Conclusion**

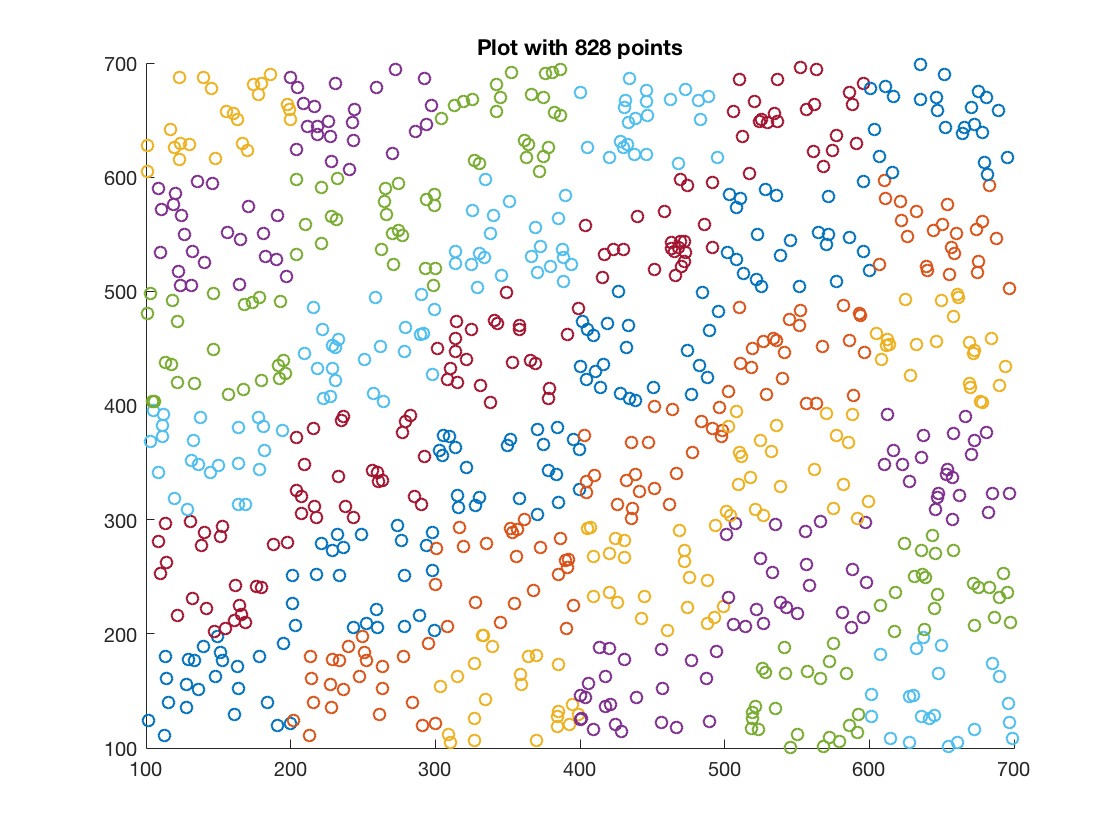


Figure 1 Randomly selected points for each 100x100 processor block.

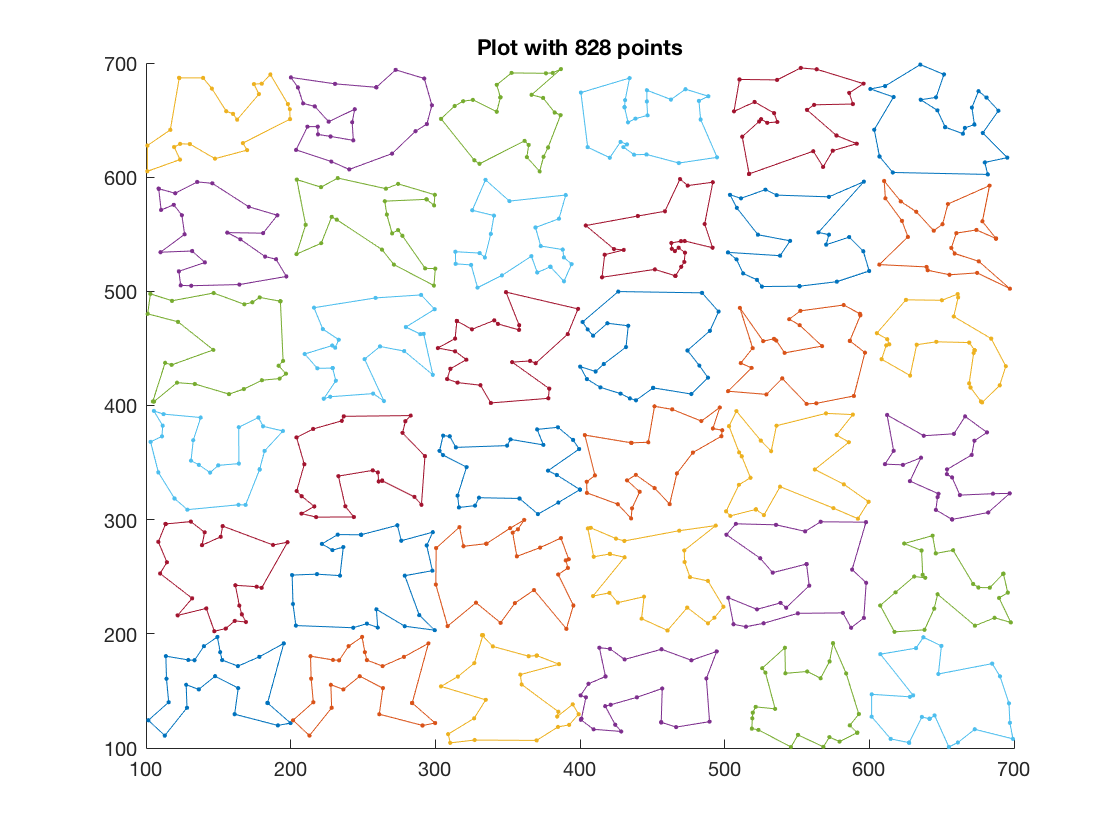


Figure 2 Sub-TSP solutions computed per processor block.

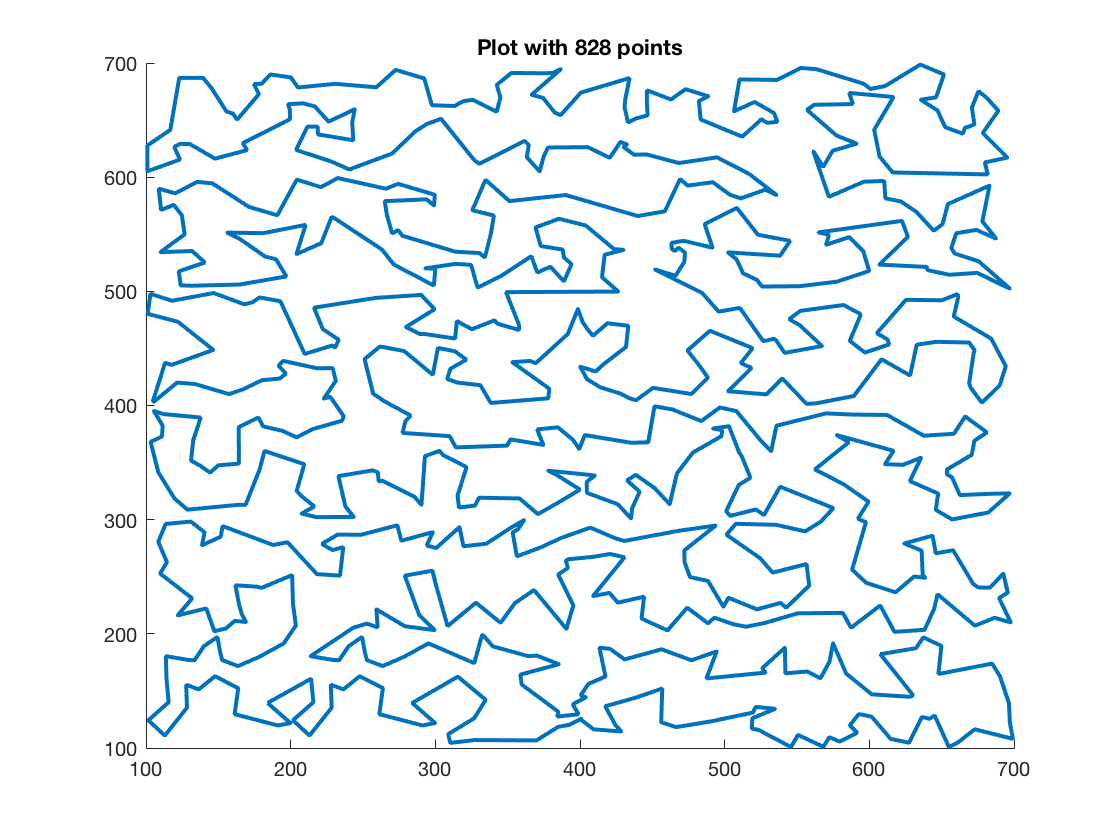


Figure 3 Final TSP across all 828 points showing no inversions.

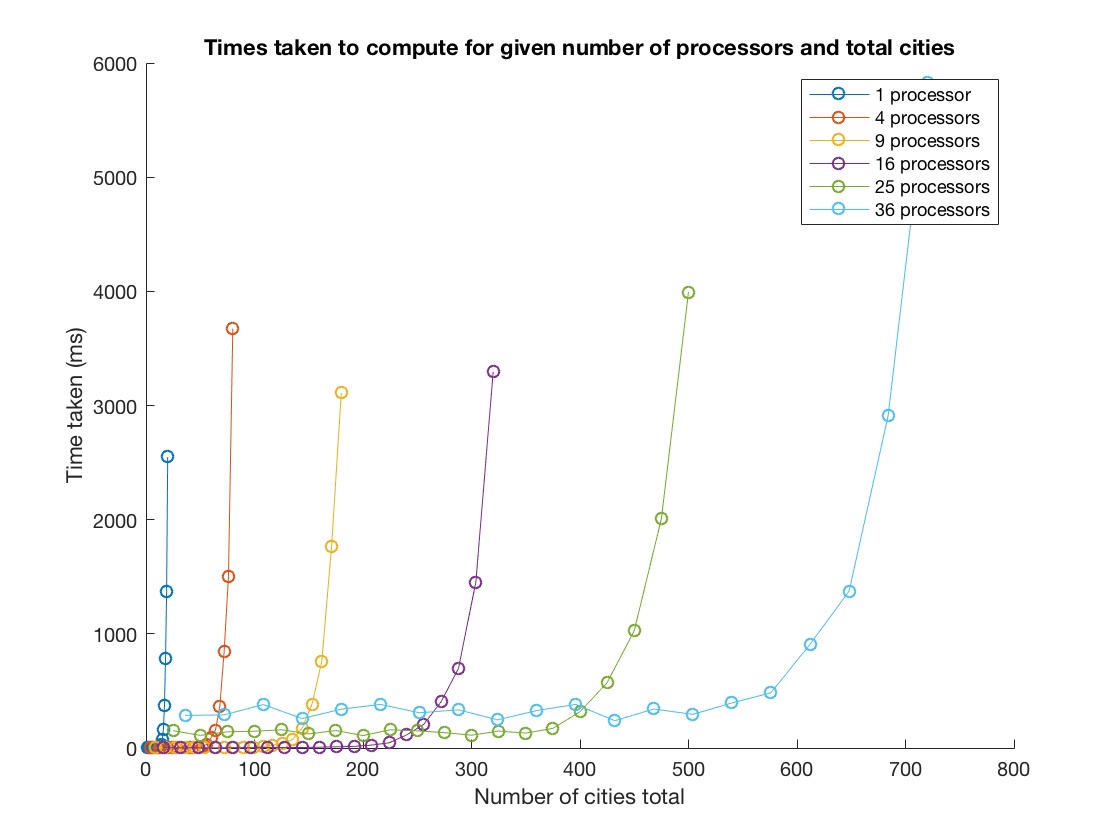


Figure 4 Time taken for calculating TSP solution for different numbers of processors. As we reach a threshold between 20-25 cities per processor, the times taken quickly increase. Before this point.

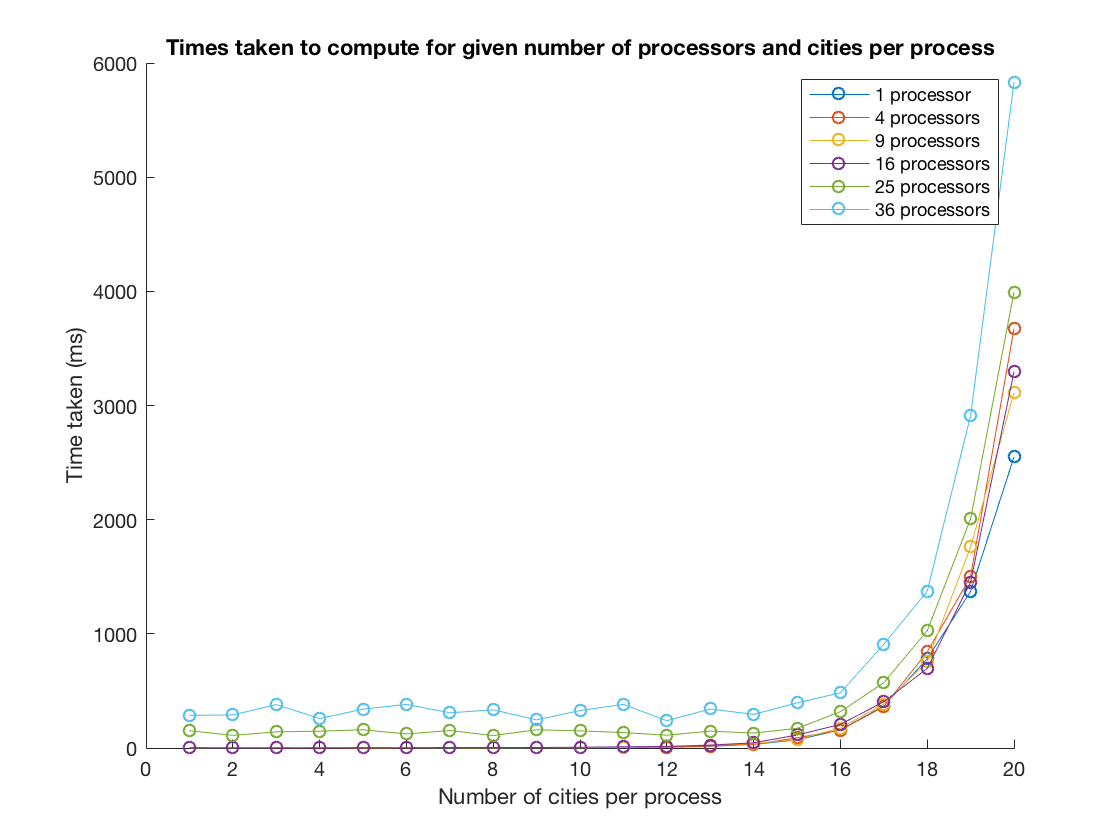


Figure Time taken comparing number of cities per process shows an obvious higher time when higher numbers of processors are involved, as would be expected given higher amounts of communication required

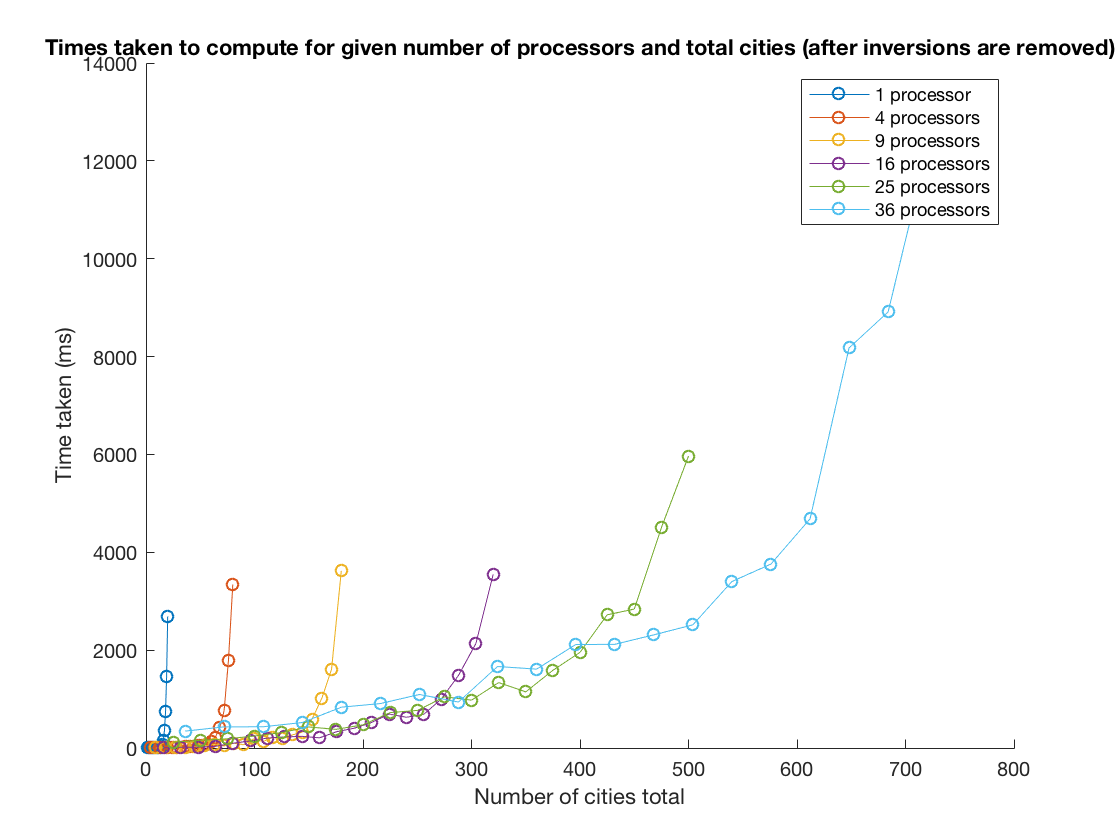


Figure 6 Time taken for each processor when handling inversions.

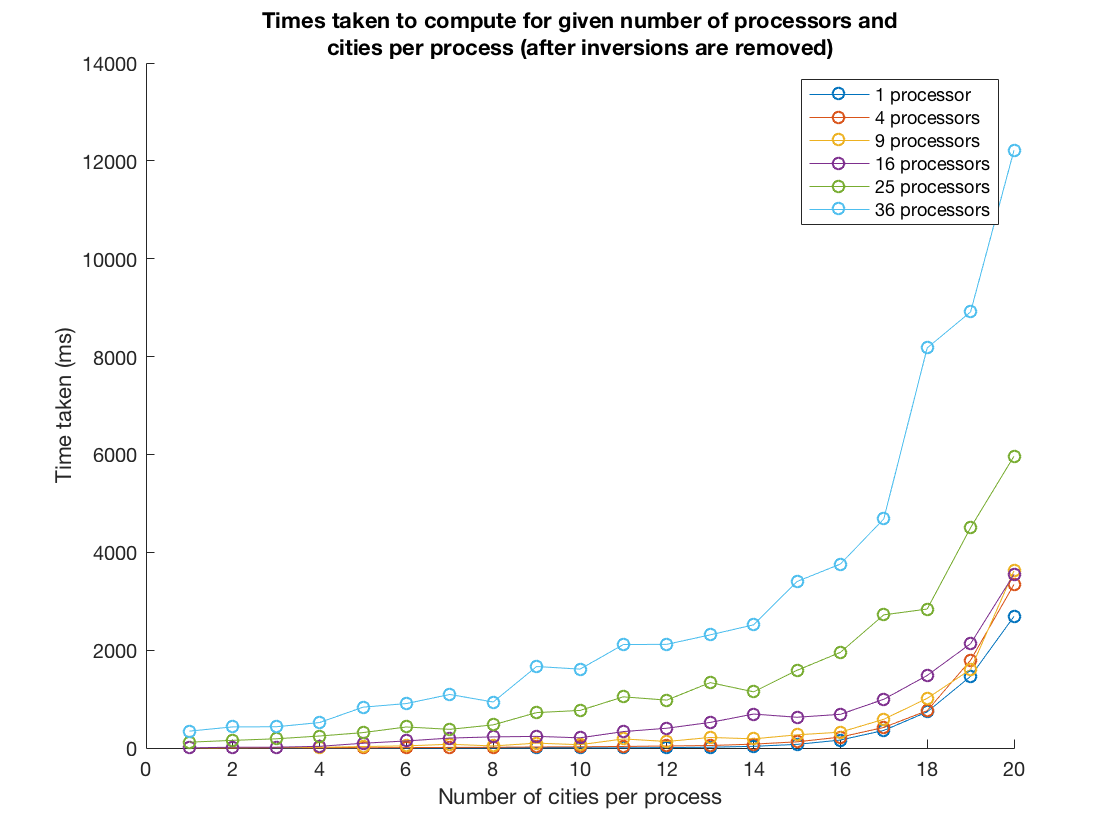


Figure Time taken comparing number of cities per process shows an obvious higher time when higher numbers of processors are involved, as would be expected given higher amounts of communication required

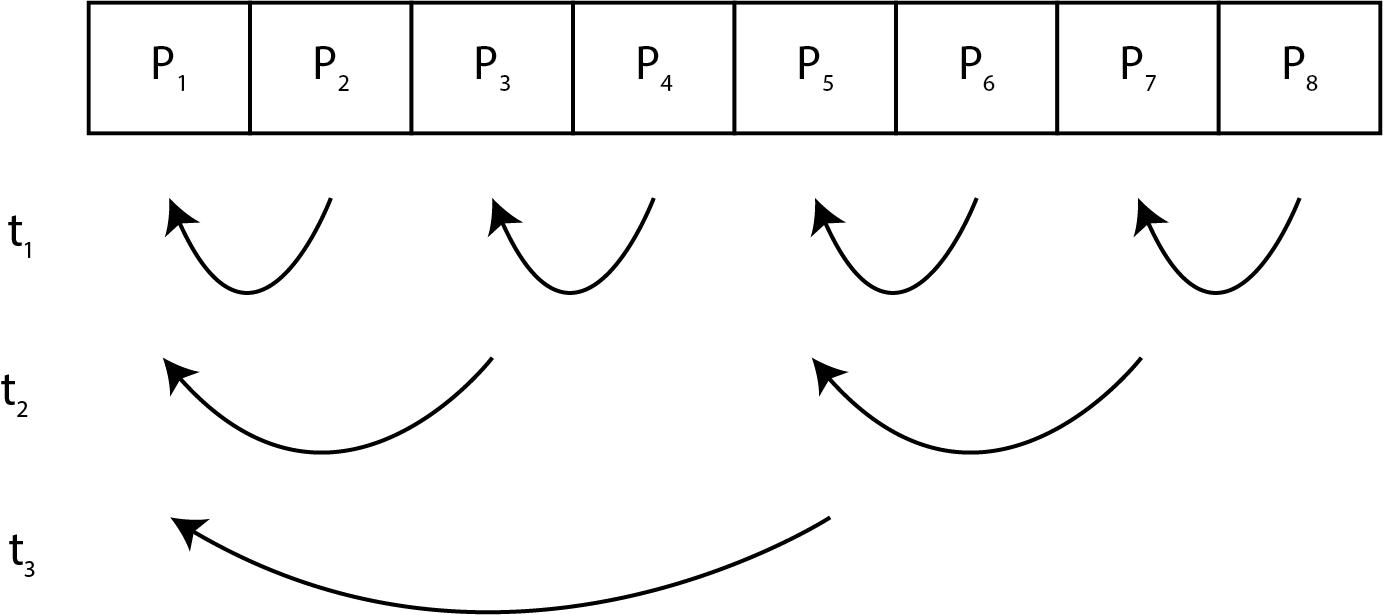


Figure 8 With 2n blocks per dimension, our algorithm simply allows for all log(x) processes to complete the expected stitching work at each time frame.

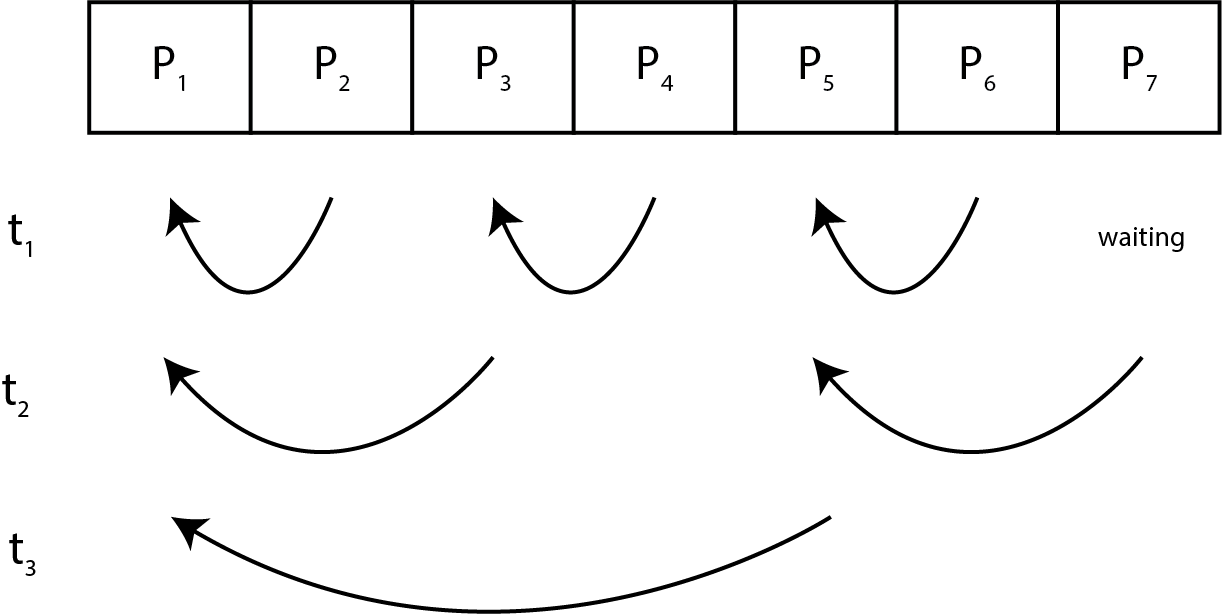


Figure 9 When dimensionality of a row is not 2n; in this example p=7, process P7 must know not to expect a message from P8 at t1 otherwise a deadlock will occur.

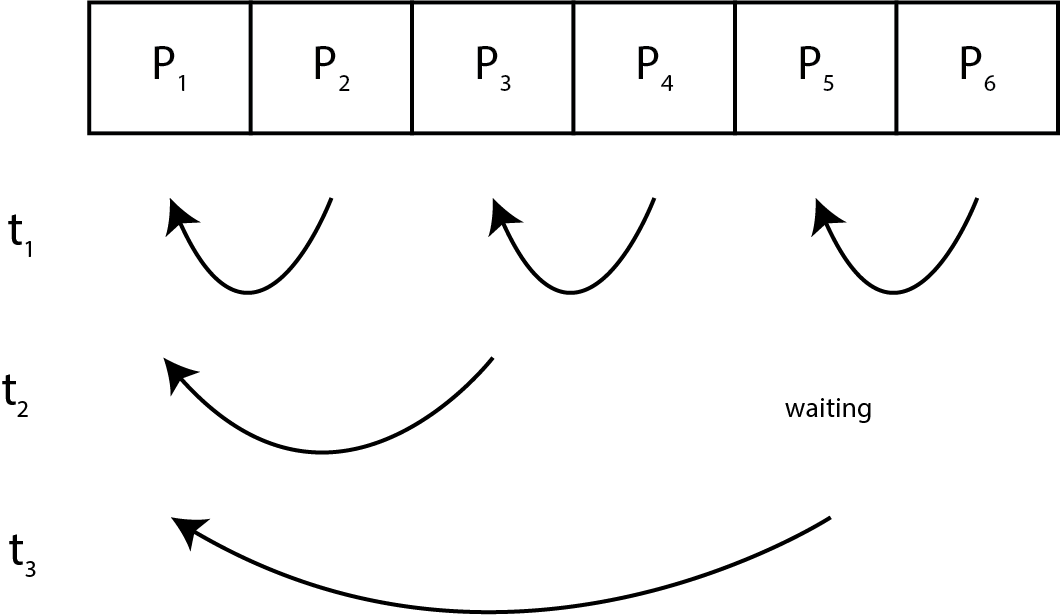


Figure 10 When dimensionality of a row is not 2n; in this example p=6, P5 must know not to expect a message from P7 at t2 otherwise a deadlock will occur.