Modeling halo exchange on the BlueGene/Q

Yadu N. Babuji & Timothy G. Armstrong

The University of Chicago

December 9, 2014

Overview

Why is task placements in Halo exchange interesting ?

- Halo exchange is a very common nearest neighbor communication pattern
- Solving PDEs require halo exchange
- Task placements can affect halo exchange performance by upto 7.5x
- Cheap optimisations with no code change

What did we do ?

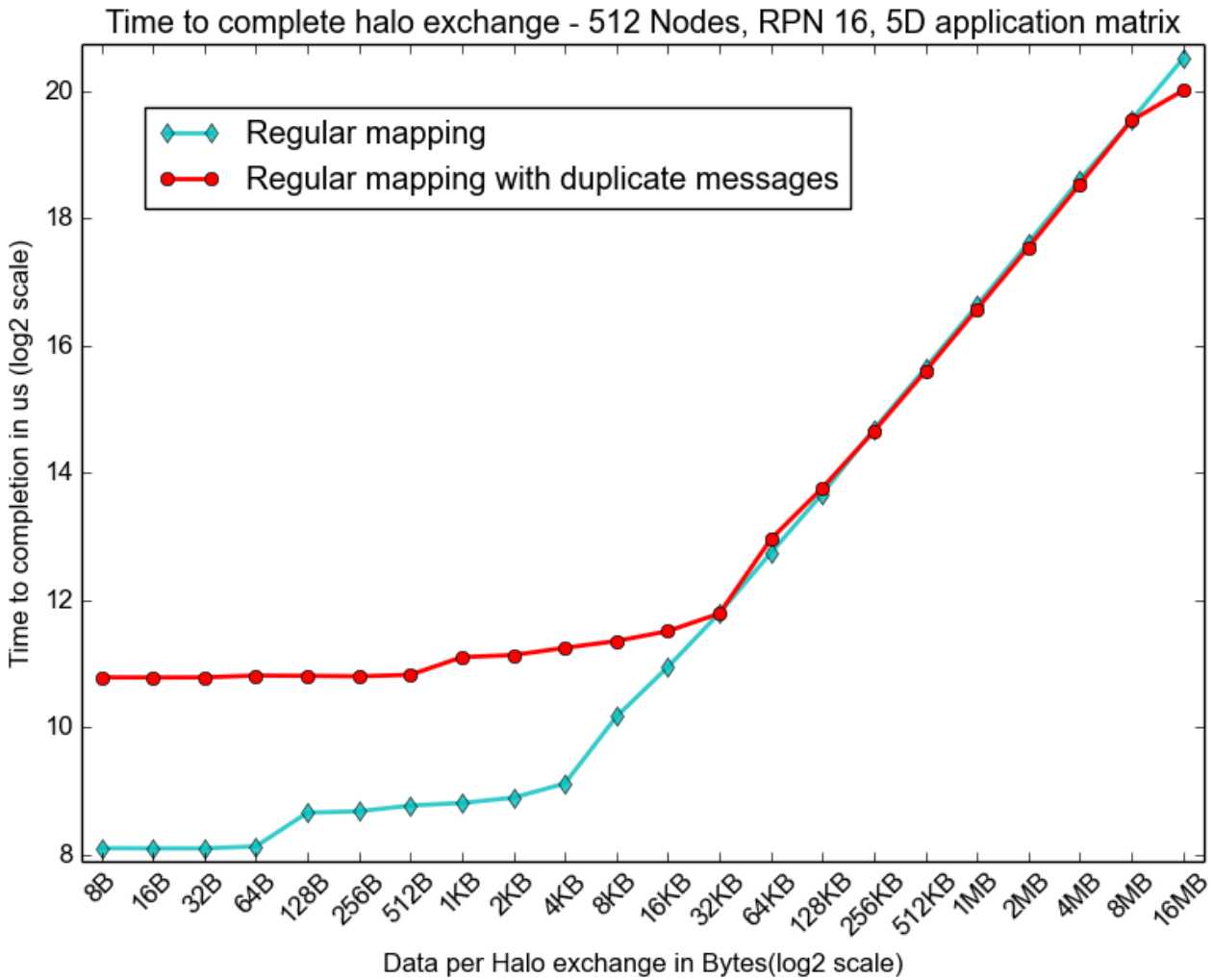
- We made an Analytical model from first principles to model the performance.
- Introduced a reasonably effective metric for mappings
- Experiments to study what factors affect performance
- Made nearly optimal and pessimal mappings
- Analysis and Plots!

What affects performance ?

- Caching effects when message sizes do not fit in L3 cache ? (No)
- Does longer distances result in higher latency ? (Surprisingly No!)
- Higher overall traffic ?

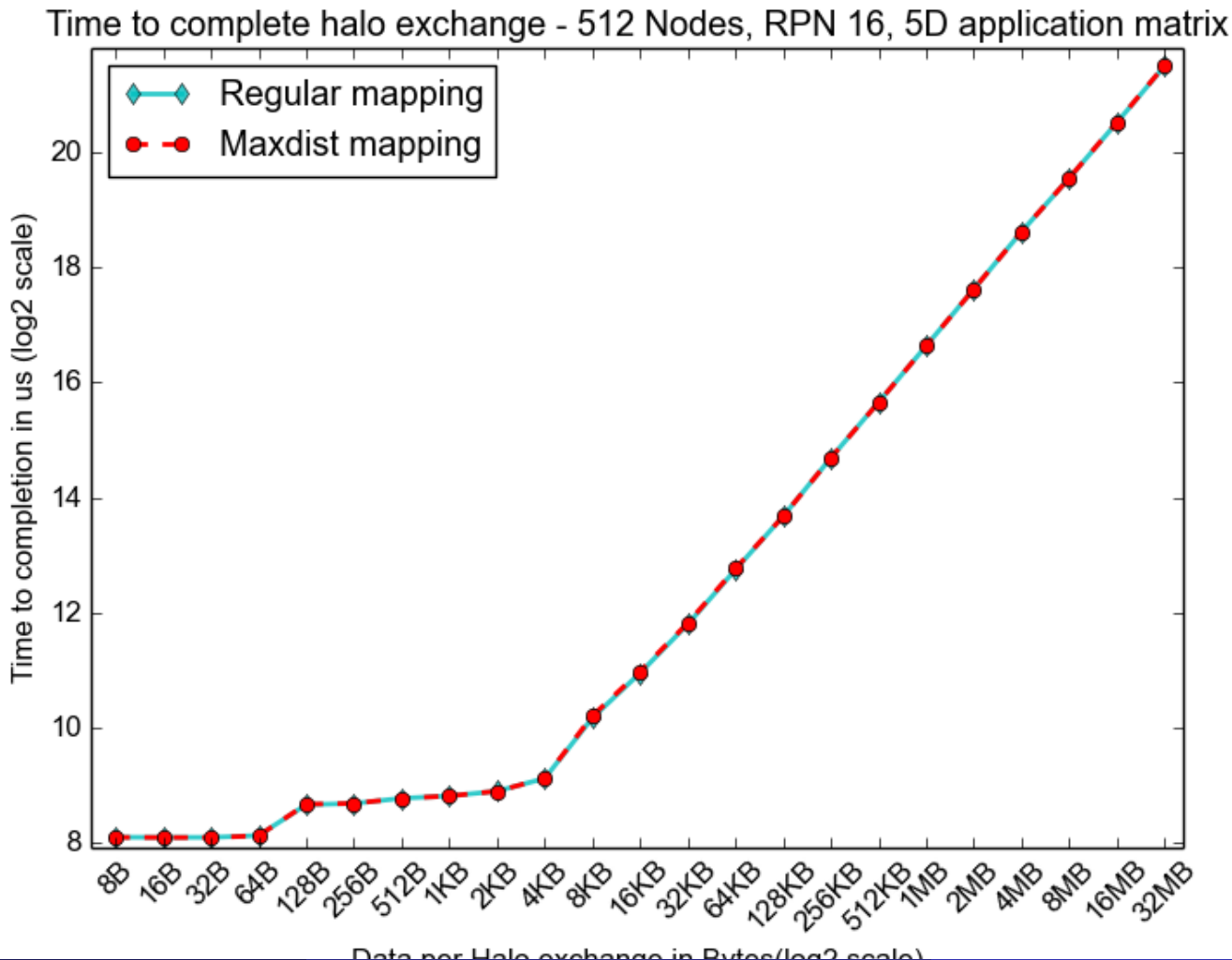
Caching plots

Figure : Caching effects



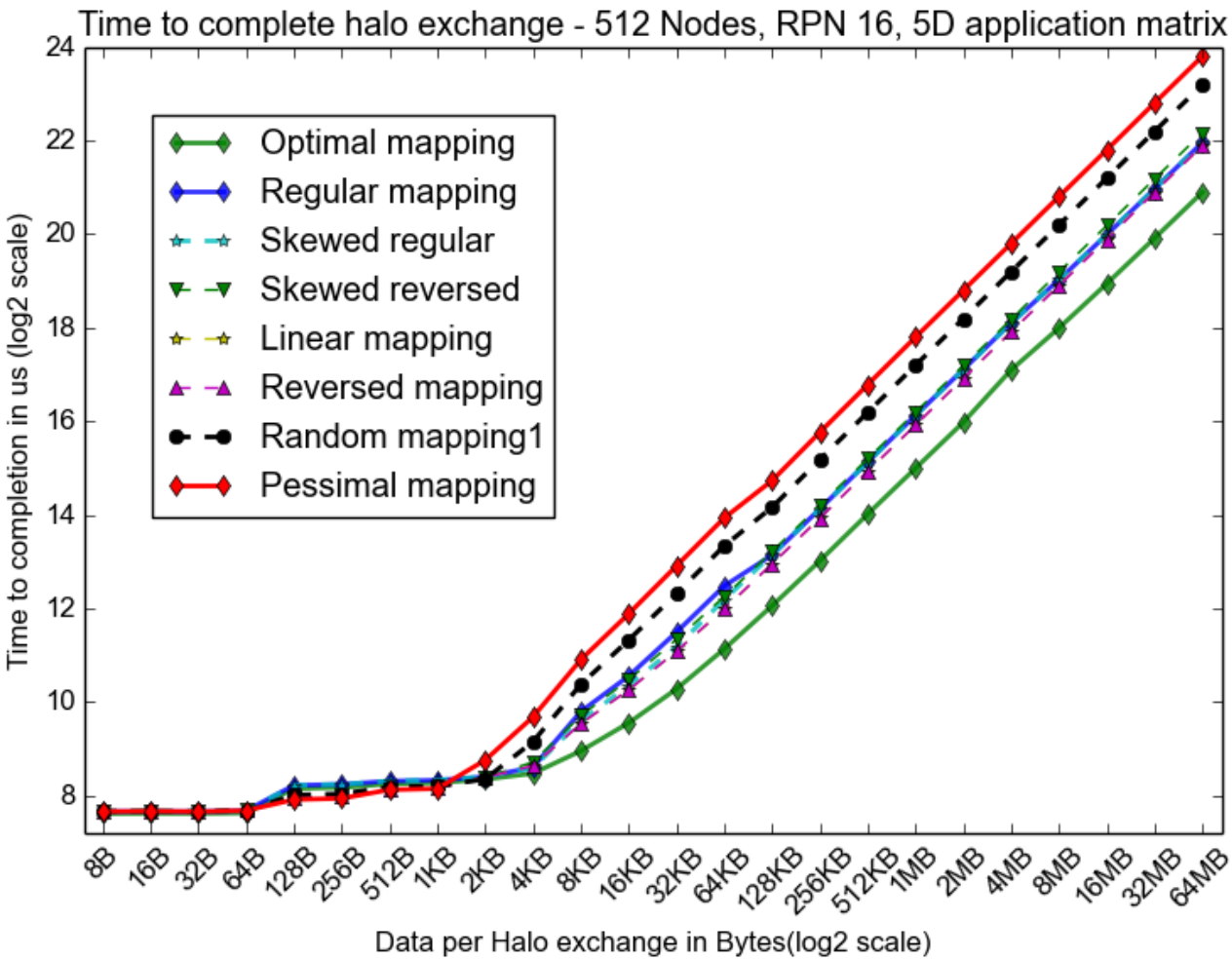
Latency plots

Figure : Latency effects



Overall traffic plots

Figure : Increasing traffic



Analytical model for Halo exchange performance

- Total number of neighbors $T_{neighbors}$, D dimensionality of application.

$$T_{neighbors} = N_{ranks} * 2 * D \tag{1}$$

- Average steps a message travels N_{steps}

$$N_{steps} = \frac{\sum_{u,v} dist_{u,v}}{T_{neighbors}} \tag{2}$$

- Time to complete a halo exchange:

$$T = t_c + (N_{steps} * N_{procs} * N * t_b * \alpha) \tag{3}$$

What mapping strategies did we try?

- Regular/Default
- Skewed regular & Skewed reversed
- Random
- Linear & Reversed
- **Pessimal mapping generated by Simulated Annealing**
- **Optimal mapping by partitioning Application domains**

Figure : 5D Linear mapping

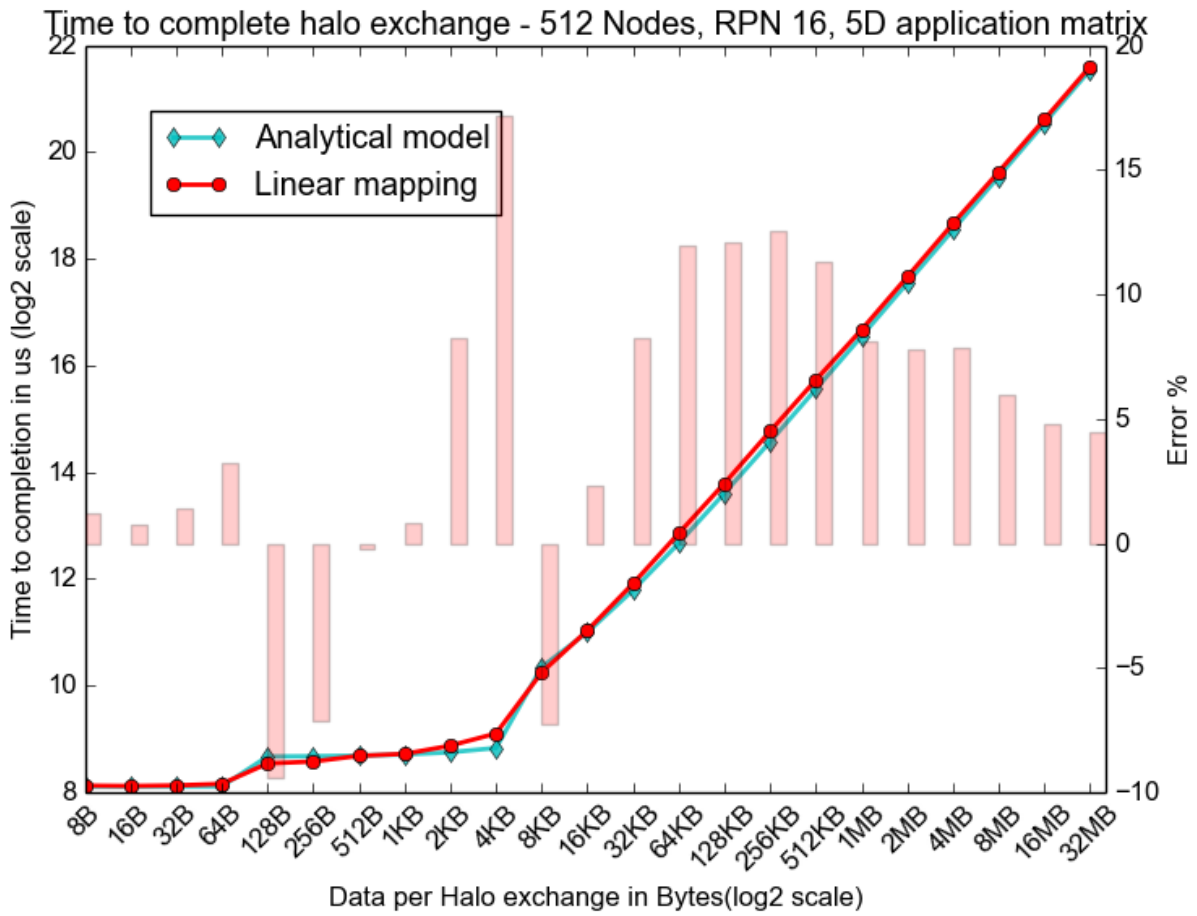


Figure : 5D Optimal mapping

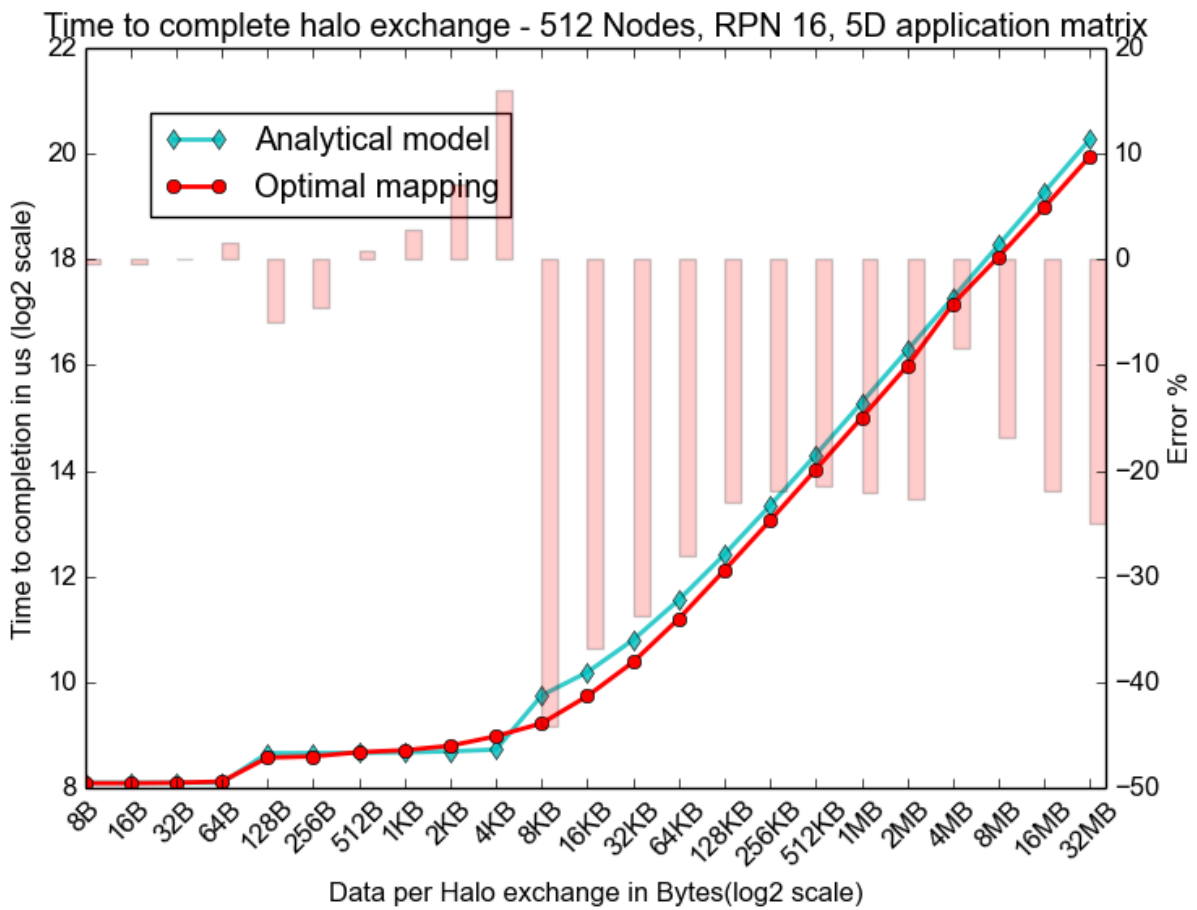


Figure : 5D Random mapping

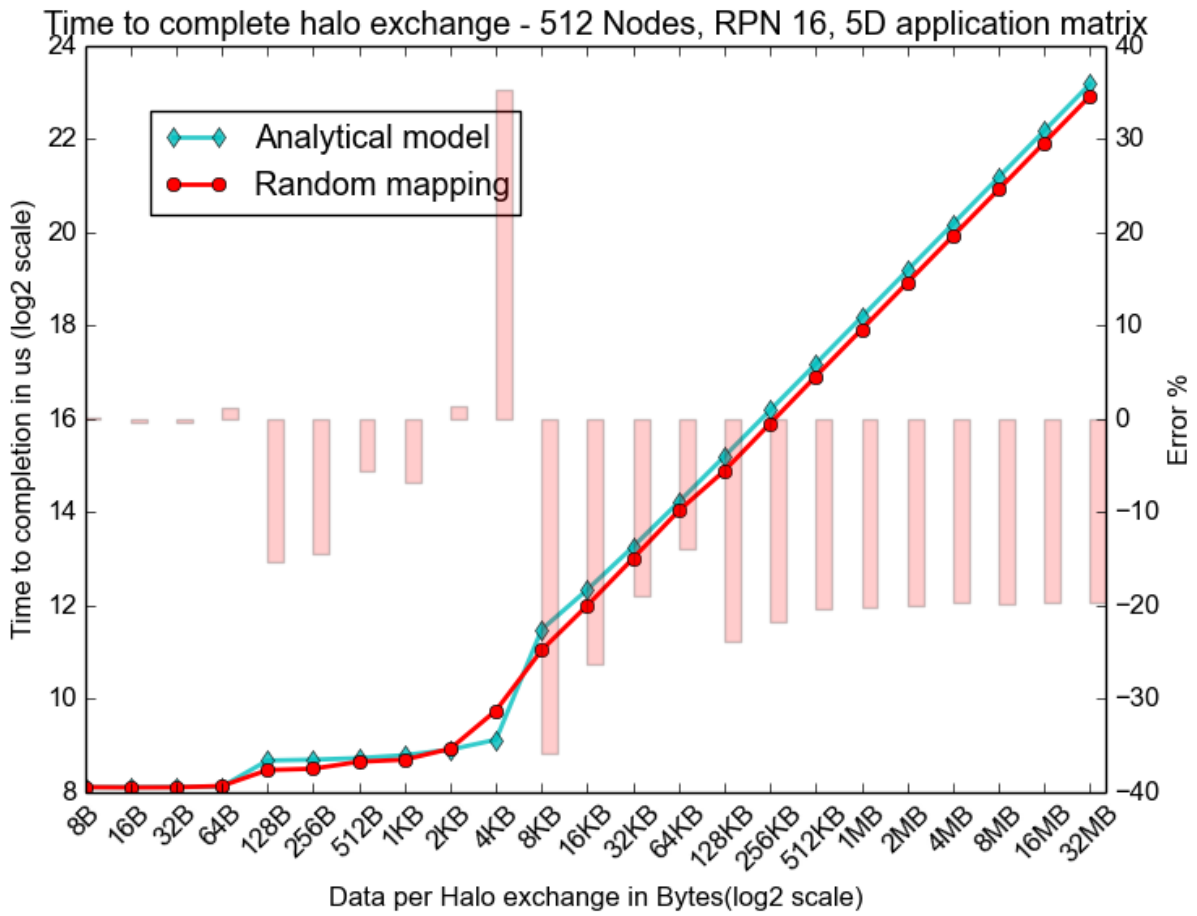


Figure : 5D Regular mapping

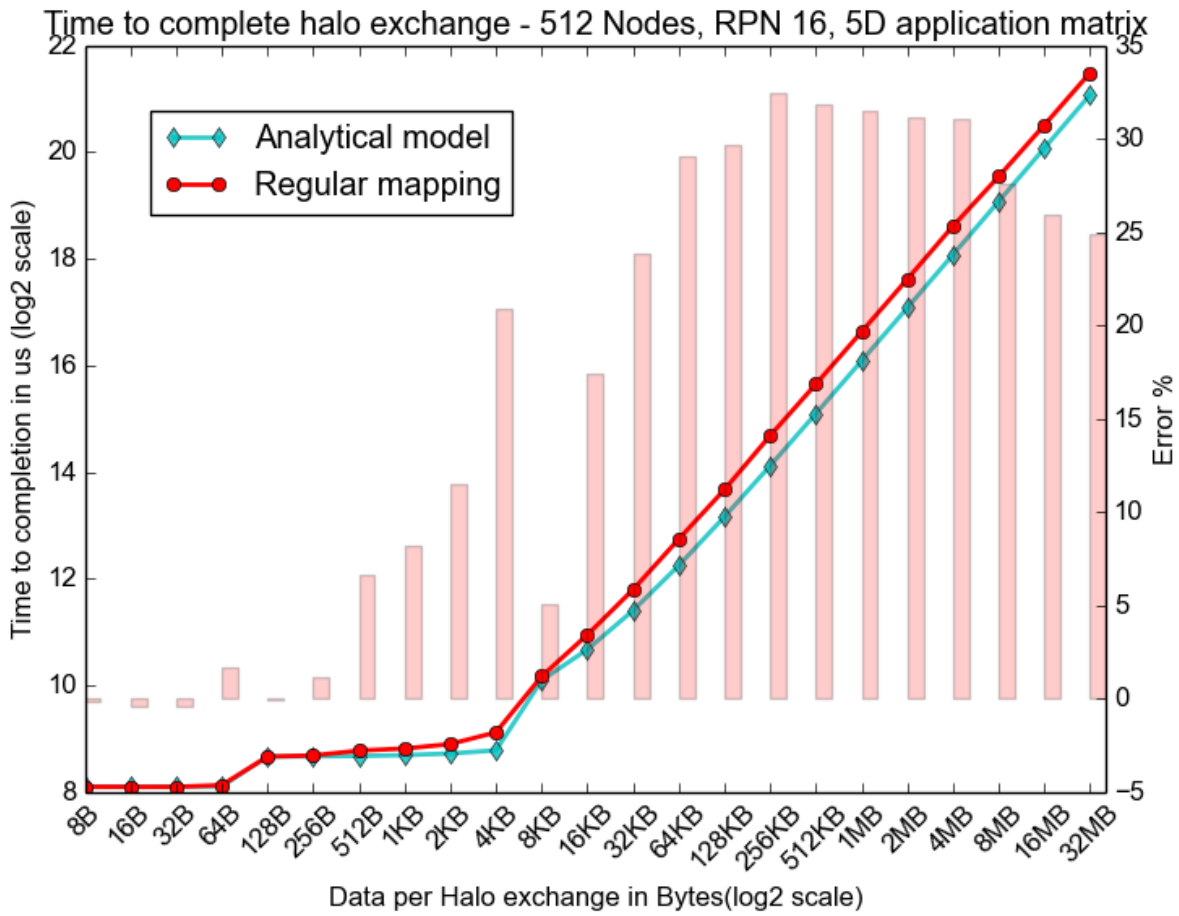


Figure : 5D Reversed mapping

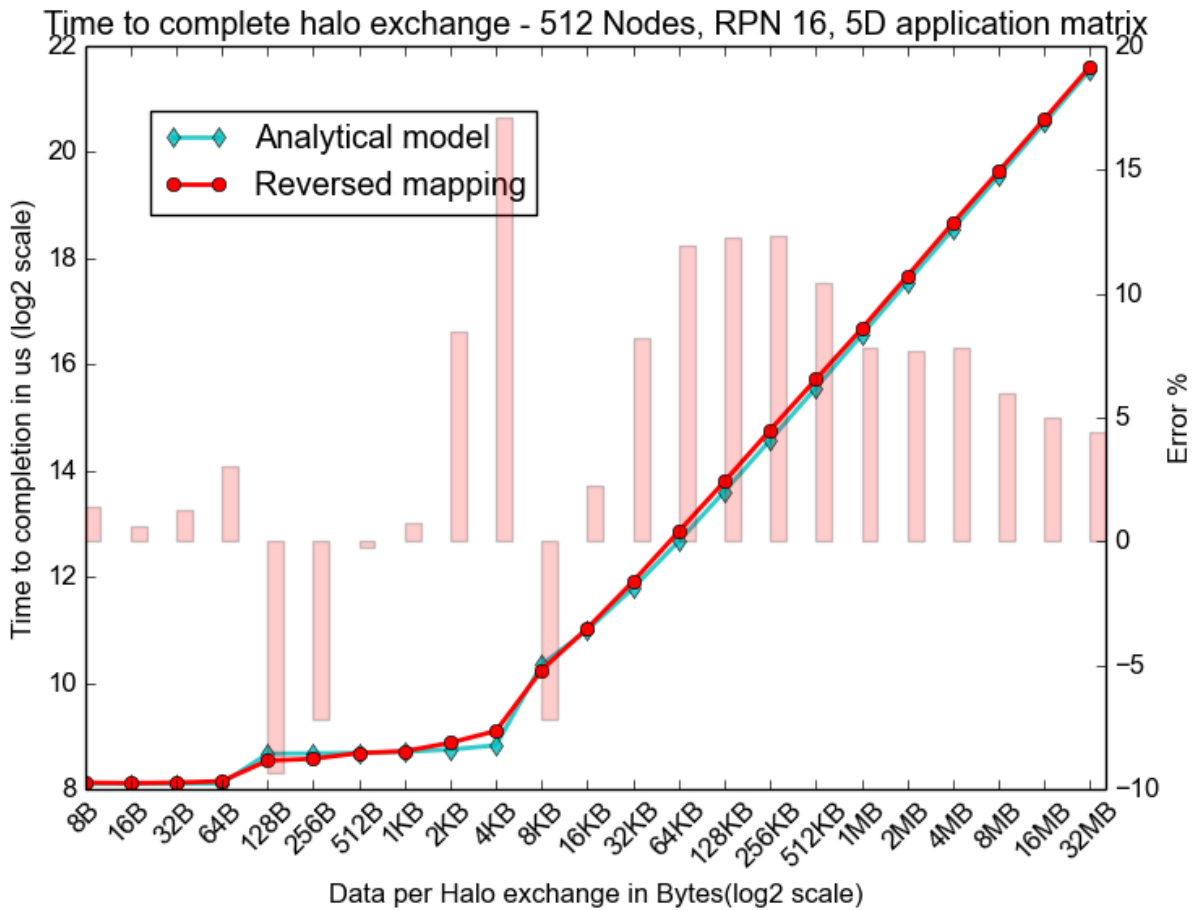


Figure : 5D Skewed Regular

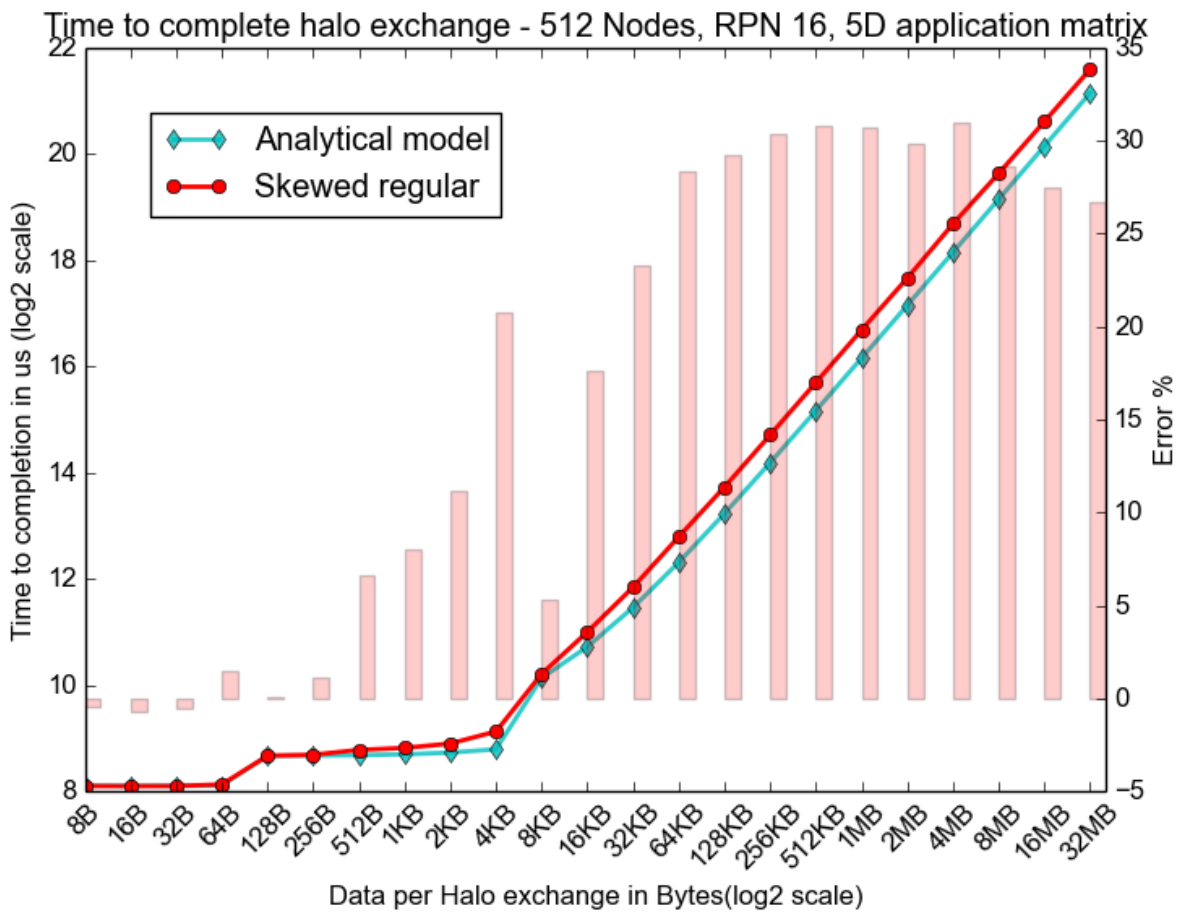


Figure : 3D Optimal mapping

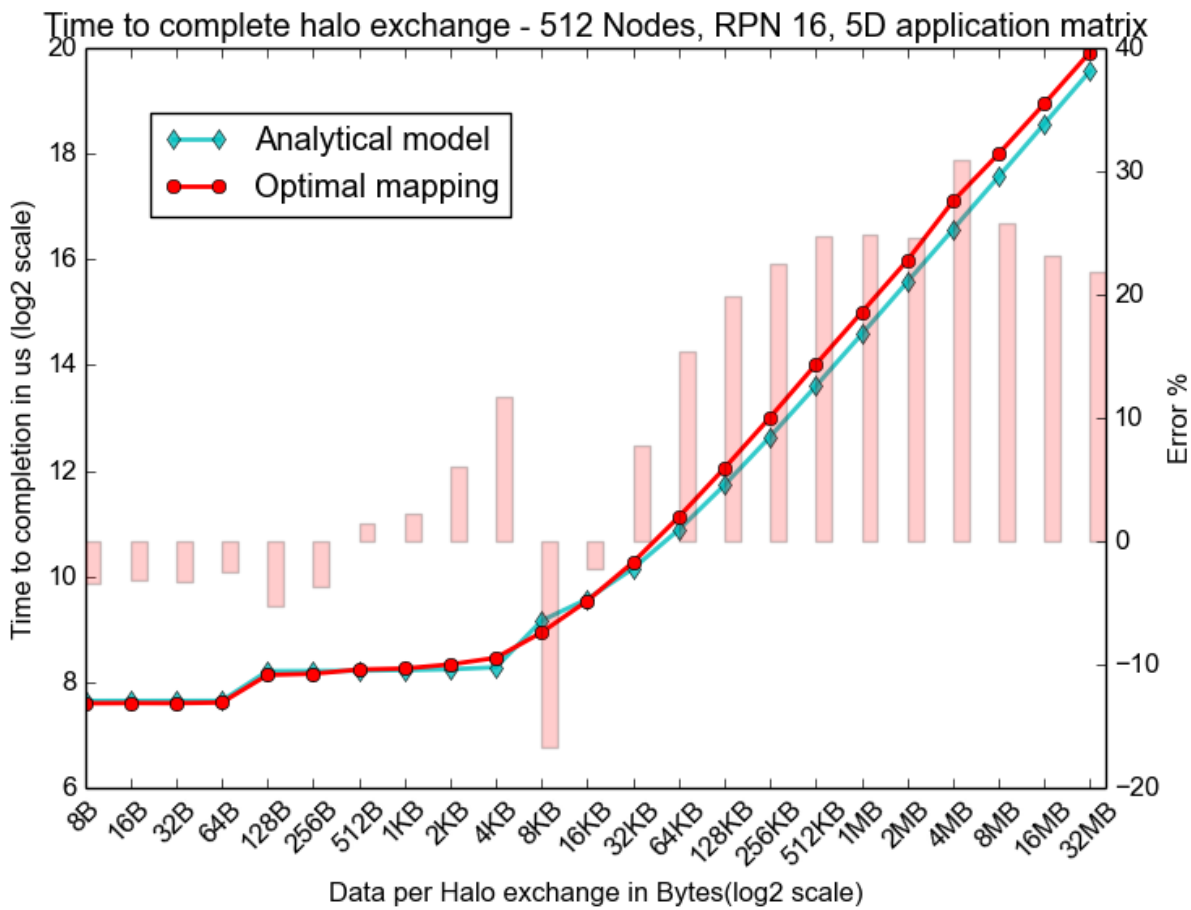


Figure : 3D Pessimal mapping

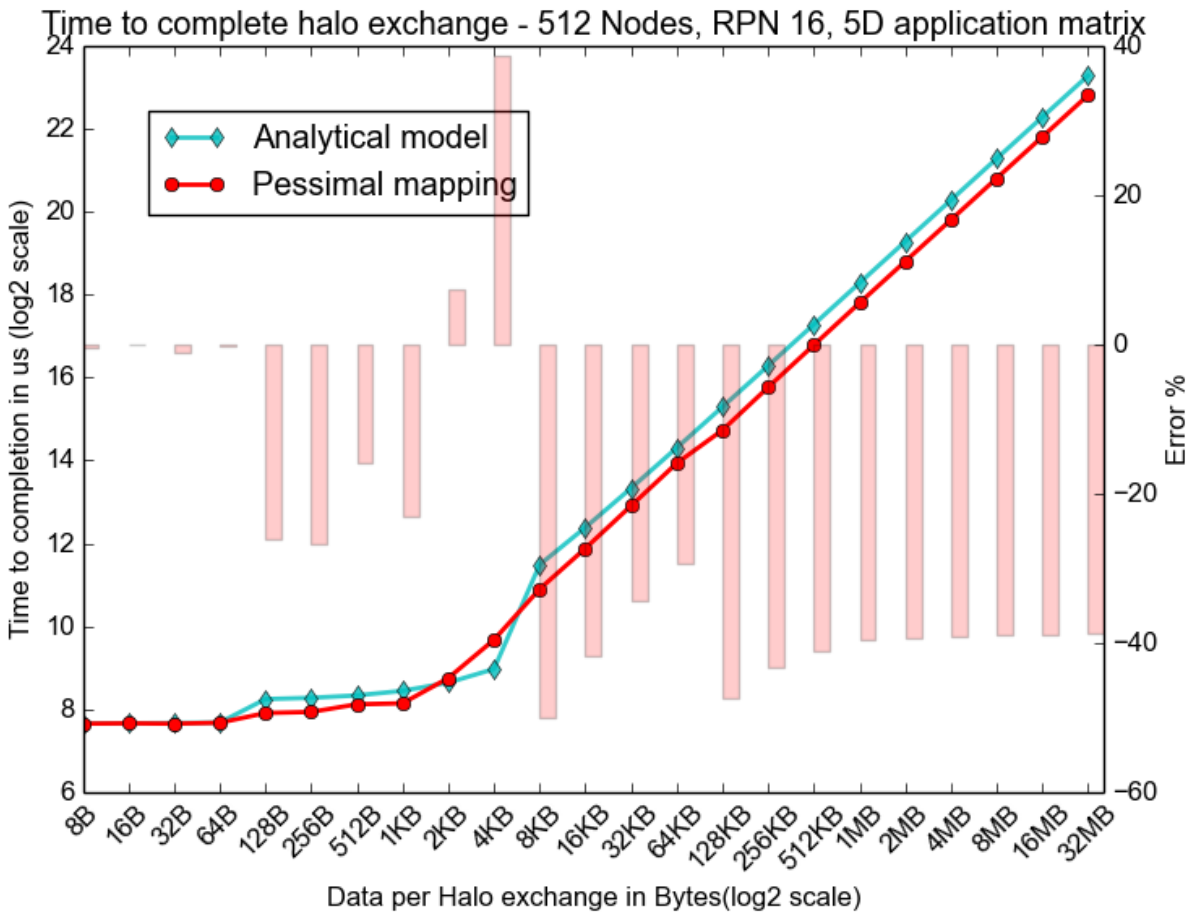


Figure : 3D Random mapping

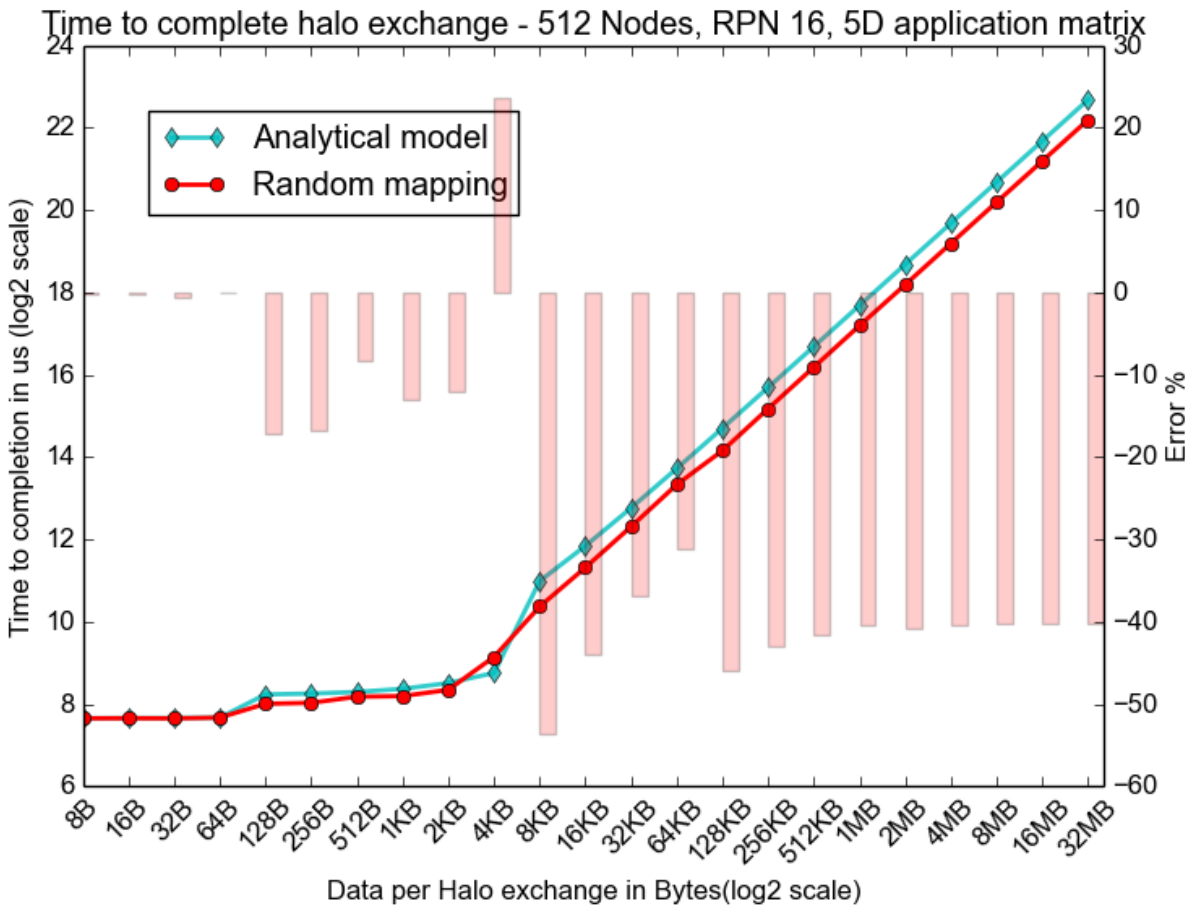


Figure : 3D Regular mapping

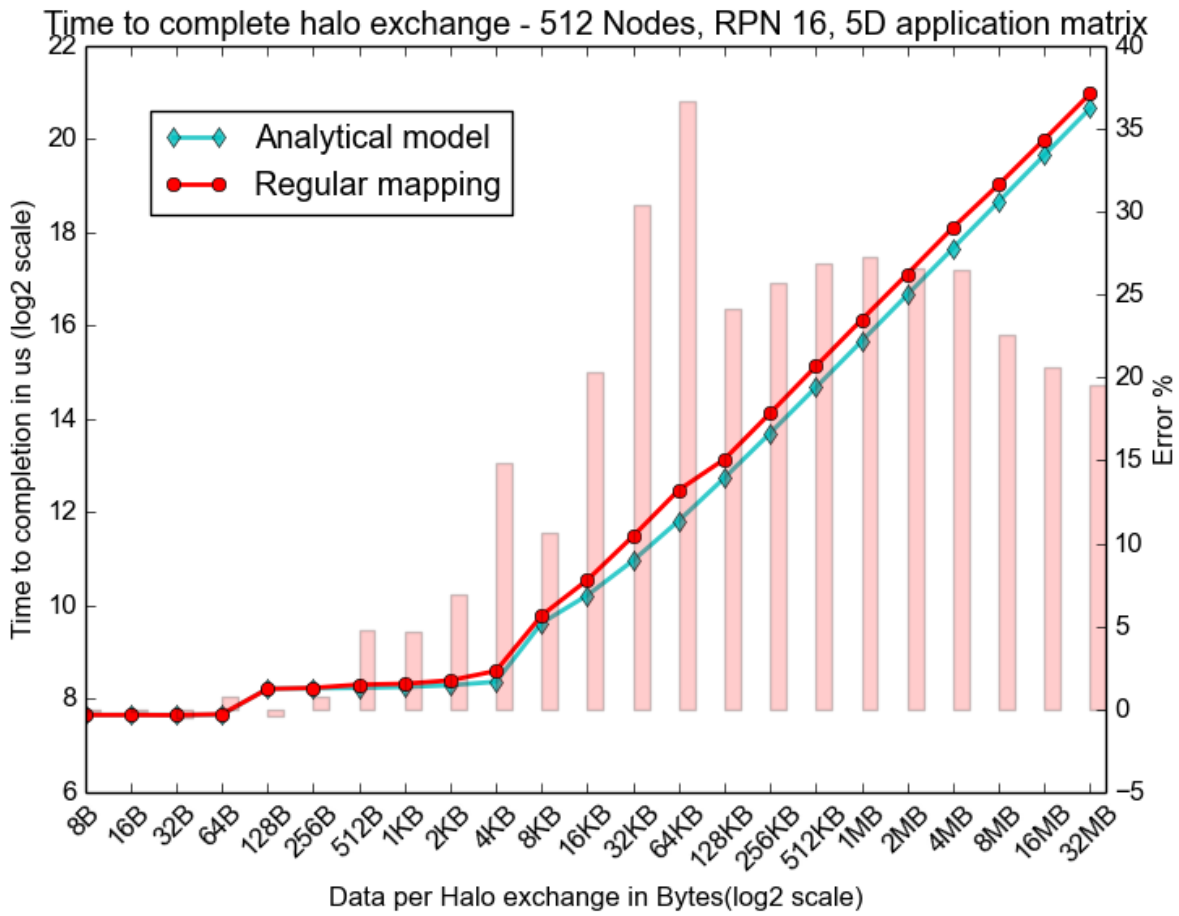


Figure : 3D Reversed mapping

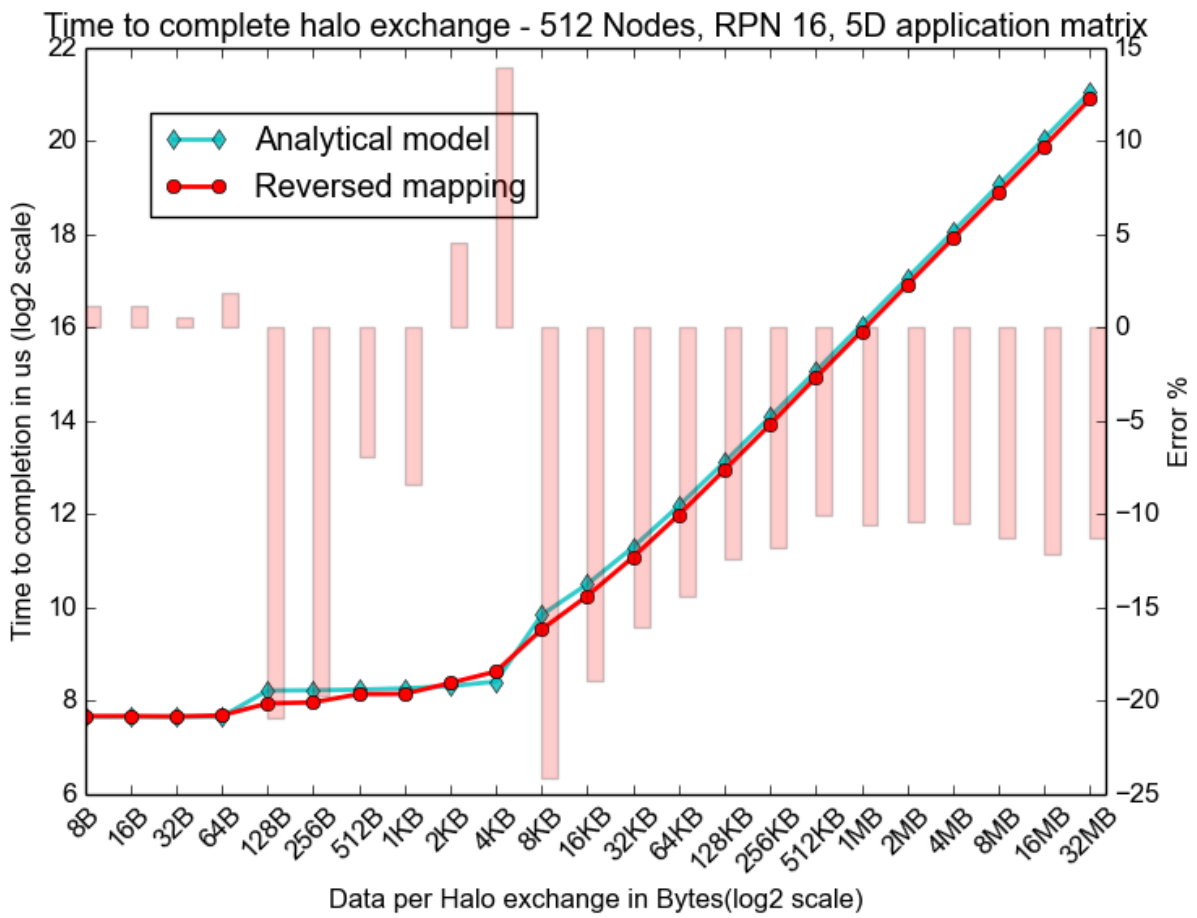


Figure : 3D Skewed regular mapping