Lecture 14 Summary

This lecture talks about the dragonfly network. Previous interconnection networks were built with low-radix routers, which had a small number of ports. Examples were Cray T3D, T3E and XT3. The case existed because for the packaging and technology constraints and the relatively low pin bandwidth available at the time, low-radix networks provided optimal latency for a given cost. However, the pin bandwidth of router chips has increased by a rate like Moore’s law over the past 20 years, which results in increasing in the signaling rate and the number of signals available to a router chip. Therefore, the best way to exploit the increasing off-chip bandwidth is to build high-radix routers with thin channel, which leads to a reduce of hop count and lower latency and lower cost. However, a change in conventional network topologies is required to better exploit the benefits. So, the dragonfly network topology was proposed, which uses a group of routers as a virtual router to increase the effective radix of the network and hence reduce network diameter, cost and latency. Existing topologies like folded-Clos and fat-tree, consume costly bandwidth to load-balance traffic that is already balanced. And a conventional butterfly network, incurs significant lower cost on balanced traffic though, its performance is severely limited on adversarial traffic patterns because of no path diversity.

The dragonfly network is a hierarchical network with three levels: router, group, and system. At the bottom level, each router has connections to p terminals, a-1 local channels, and h global channels. And the radix of each router is k = p + a + h -1. With adaptive routing, the dragonfly topology’s benefits can be fully exploited. Although the topology provides high path diversity, it needs nonminimal global adaptive routing to properly exploit the diverse paths. UGAL-L is used as the baseline routing algorithm where the routing decision is based on local queue information at the current router node. With tow modifications, it can overcome its limitation about the dragonfly topology, and yield performance results approaching an ideal implementation using global information. By reducing global channels, a dragonfly reduces cost by 20 percent compared to a flattened butterfly and by 52 percent compared to a folded-Clos network in configuration with more than 16K nodes.