Load Balancing:

- By:-Kirti Garg

Load Balancing

◆ Goal: All processors working all the time
- Efficiency of 1
- Distribute the load (work) to meet the goal

◆ Two types of load balancing
- Static
- Dynamic

Load Balancing

- The load balancing problem can be reduced to the bin-packing problem
  - NP-complete
- For simple cases, we can do well, but …
  - Heterogeneity
  - Different types of resources
    - Processor
    - Network, etc.

Evaluation of load balancing

- ◆ Efficiency
  - Are the processors always working?
  - How much processing overhead is associated with the load balance algorithm?
- ◆ Communication
  - Does load balance introduce or affect the communication pattern?
  - How much communication overhead is associated with the load balance algorithm?
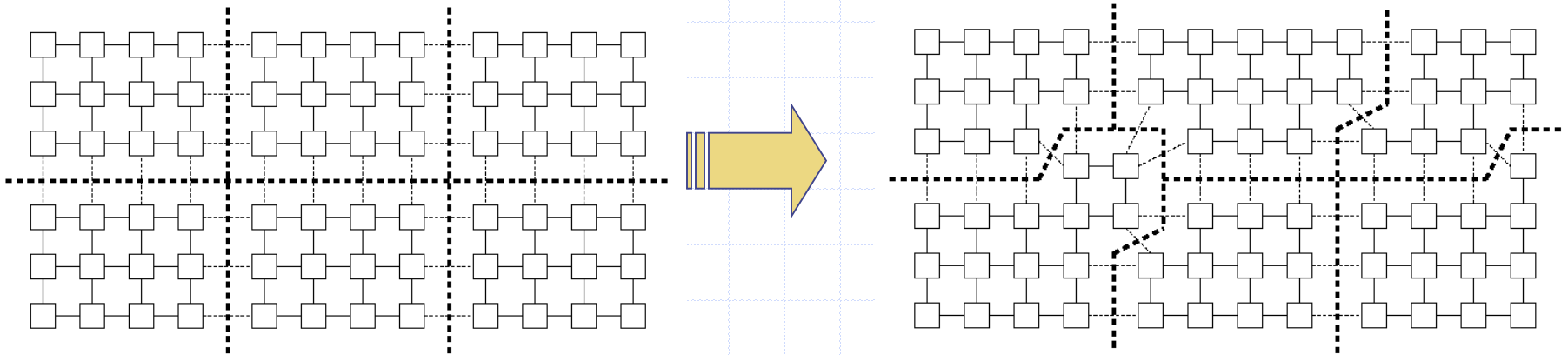 - How many edges are cut in communication graph?

Static Load Balancing

◆ This mode do not use node performance information to make distribution decisions.

◆ For example, **Round Robin** is a static load balancing mode

Static Load Balancing

- load balances connections between available nodes without measuring the relative performance of those nodes.

Dynamic Load Balancing

- Load is statically partitioned initially
- Adjust load when an imbalance is detected.
- Objectives
  - rebalance the load
  - keep edge cut minimized (communication)
  - avoid having too much overhead

Dynamic Load Balancing

- Consider adaptive algorithms
- After an interval of computation
  - mesh is adjusted according to an estimate of the discretization error
    - coarsened in areas
    - refined in others
- Mesh adjustment causes load imbalance