Cost and Time Trade-off of Scheduling Grid Tasks over Grooming Capable Networks

Norbert Szabó, József Gál, Ákos Ladányi, Tibor Cinkler

Outline

- Motivation
- Network architecture
- GRID Model
- Results
- Summary

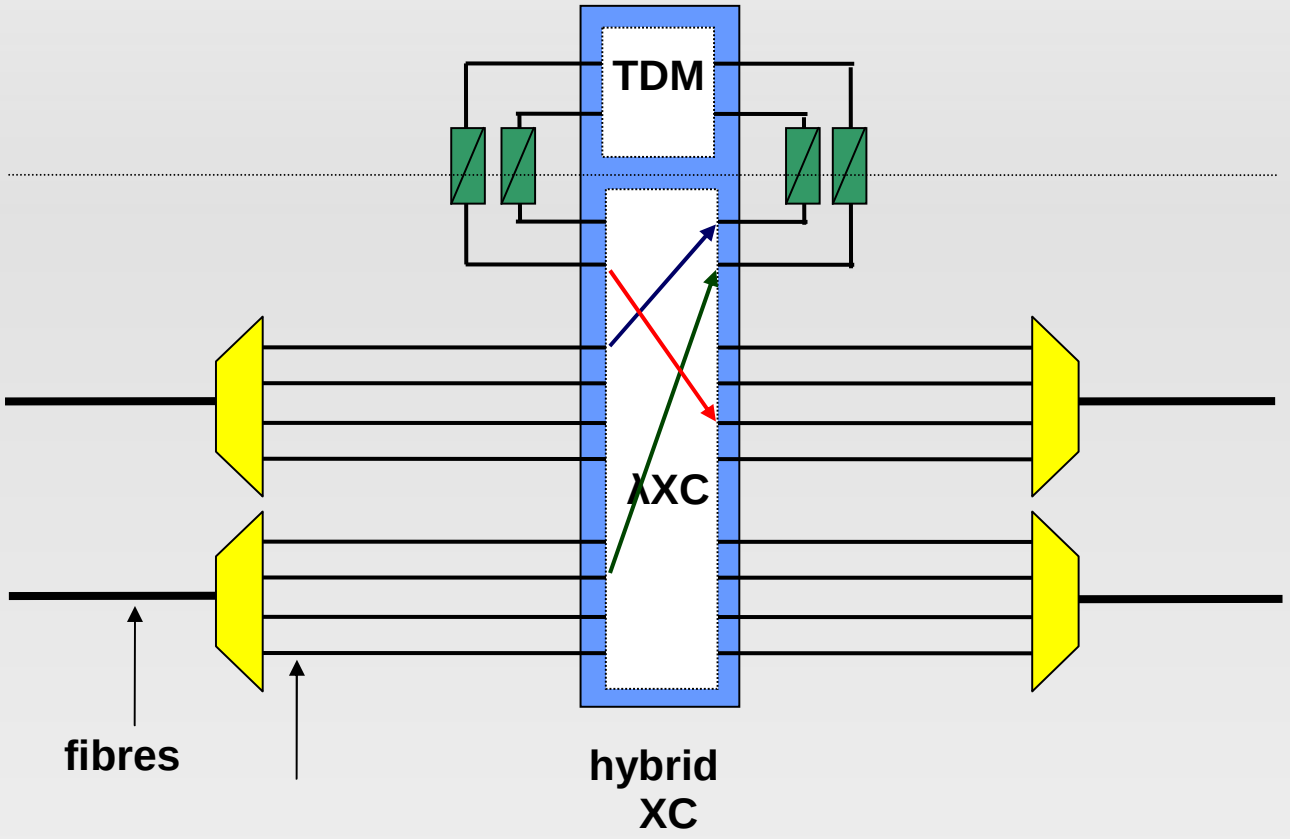
Motivation

-
  - Loosely-coupled resources forming a virtual computer
- Resources
  - Computation elements
  - Storage elements
  - ...
- Distributed
  - When a task requires data not residing on a local computation element then it needs to be transferred there
- Data-Intensive Applications
  - Video transcoding
  - Certain scientific computations

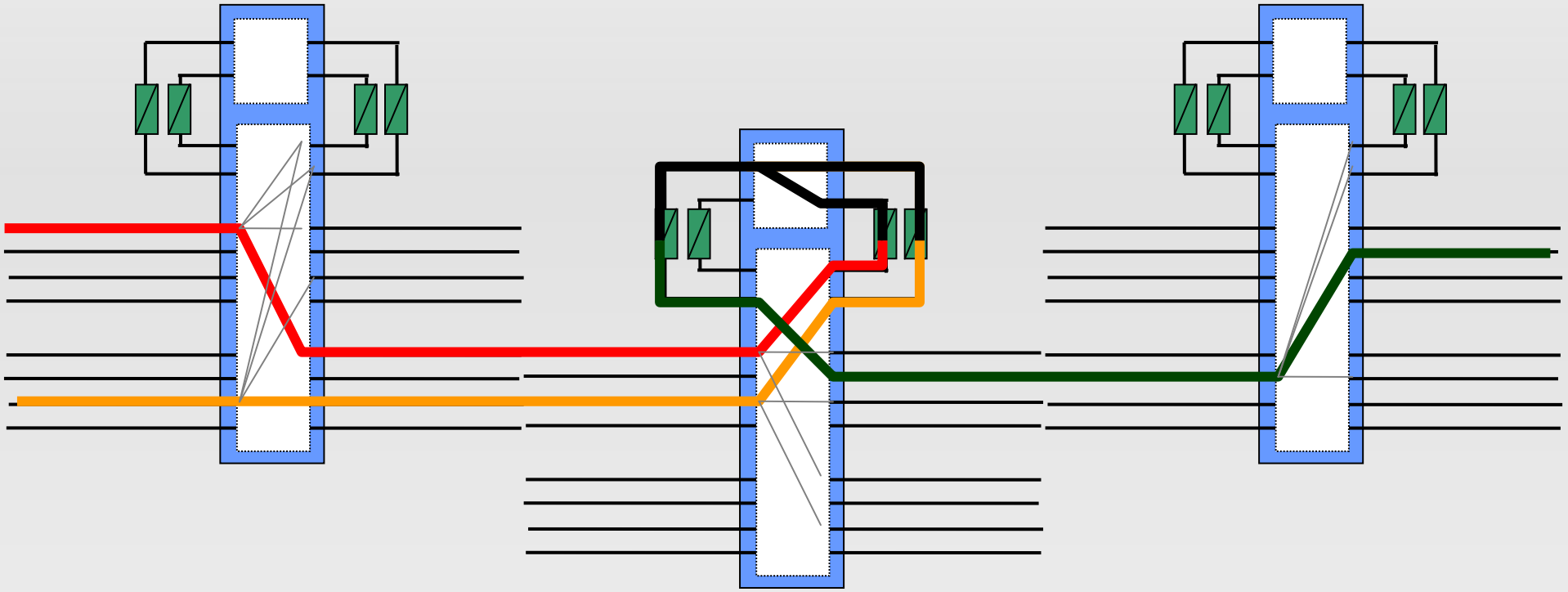
Network architecture

- Two-layer optical network
  - The lower one is the λ-path system (a WR-DWDM network)
  - The upper one is an ”electronic” layer (SDH, ATM, GbE)
- Can be the part of a multi-layer GMPLS/ASTN architecture

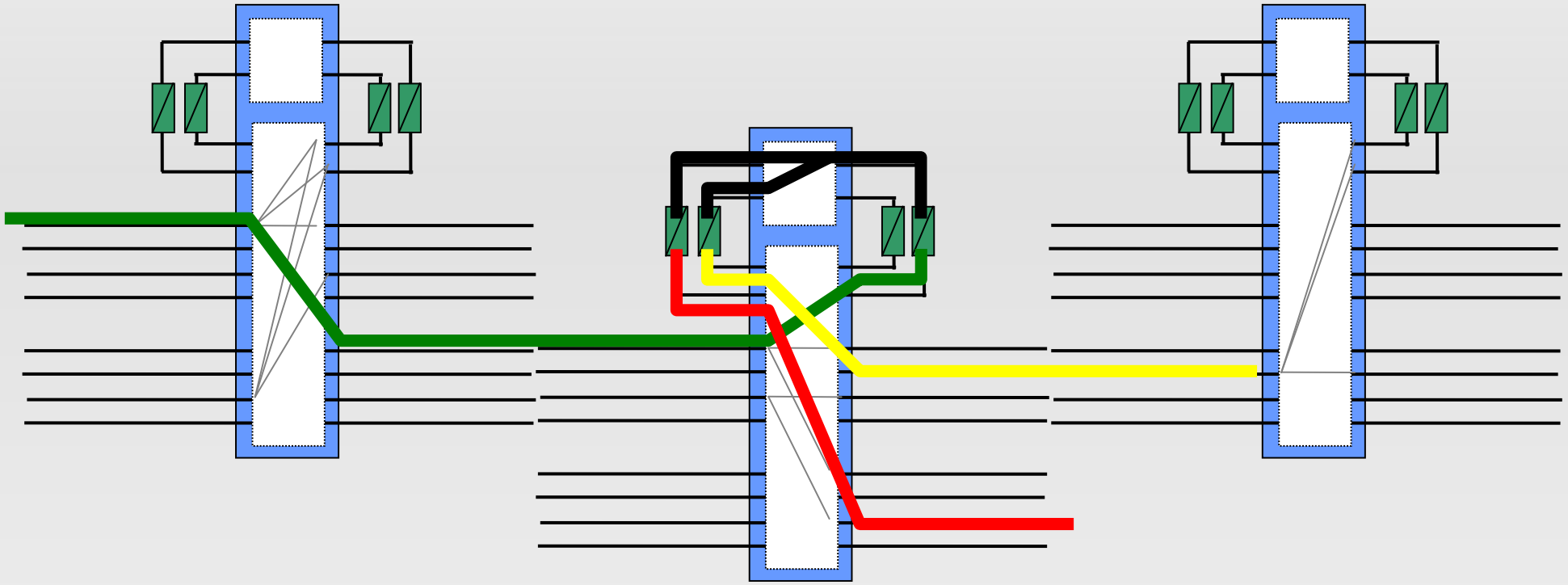
Traffic Grooming Node Architecture



Grooming - an example (1)



Grooming - an example (2)



GRID model

- Demand
  - Data size
  - Computation size
  - Number of subtasks the task can be divided to
- Steps of processing a demand
  - Determine which clusters and processors we assign the subtasks
  - Set up connections from the data location to the chosen clusters
  - Transfer the data
  - Tear down connections
  - Start computations

ILP Formulation: variables and parameters

- $Y_i \in \{0,1\}$ 1 if cluster *i* is used
- d_{ij} computation capacity of processor *j* of cluster *i*
- D_i transmission delay to cluster *i*
- c_{ij} calculation cost for a unit of time of proc. *j* of cluster *i*
- C_i transmission cost to cluster *i*
- (MI, T, K, P, D) → a GRID demand, given as
  - **MI: Computational size of the task to execute**
  - **T: Total Time limit**
  - **K: Total Cost limit**
  - **P: Number of subtasks the task can be divided to**
  - **D: Amount of data that has to be transferred to the processor**

ILP Formulation

$$\min\left\{\alpha t_{max} + (1-\alpha)\, k_{max}\right\}, 0 \le \alpha \le 1$$

cost

time

$$\sum_{\forall i,j} X_{i,j,k} = 1 \quad \forall k$$

one processor per subtask

$$\sum_{\forall k,j} X_{i,j,k} \ge Y_i \quad \forall i$$

if non of processors of cluster i is used, Y=0

if any of processors of cluster i is used, Y=1

$$X_{i,j,k} \le Y_i \quad \forall i,j,k$$

Total of transmission and execution time to be limited by tmax

$$Y_i D_i + \frac{\frac{MI}{P}\sum_{\forall k} X_{i,j,k}}{d_{i,j}} \le t_{max}$$

$$t_{max} \le T$$

Total cost limit

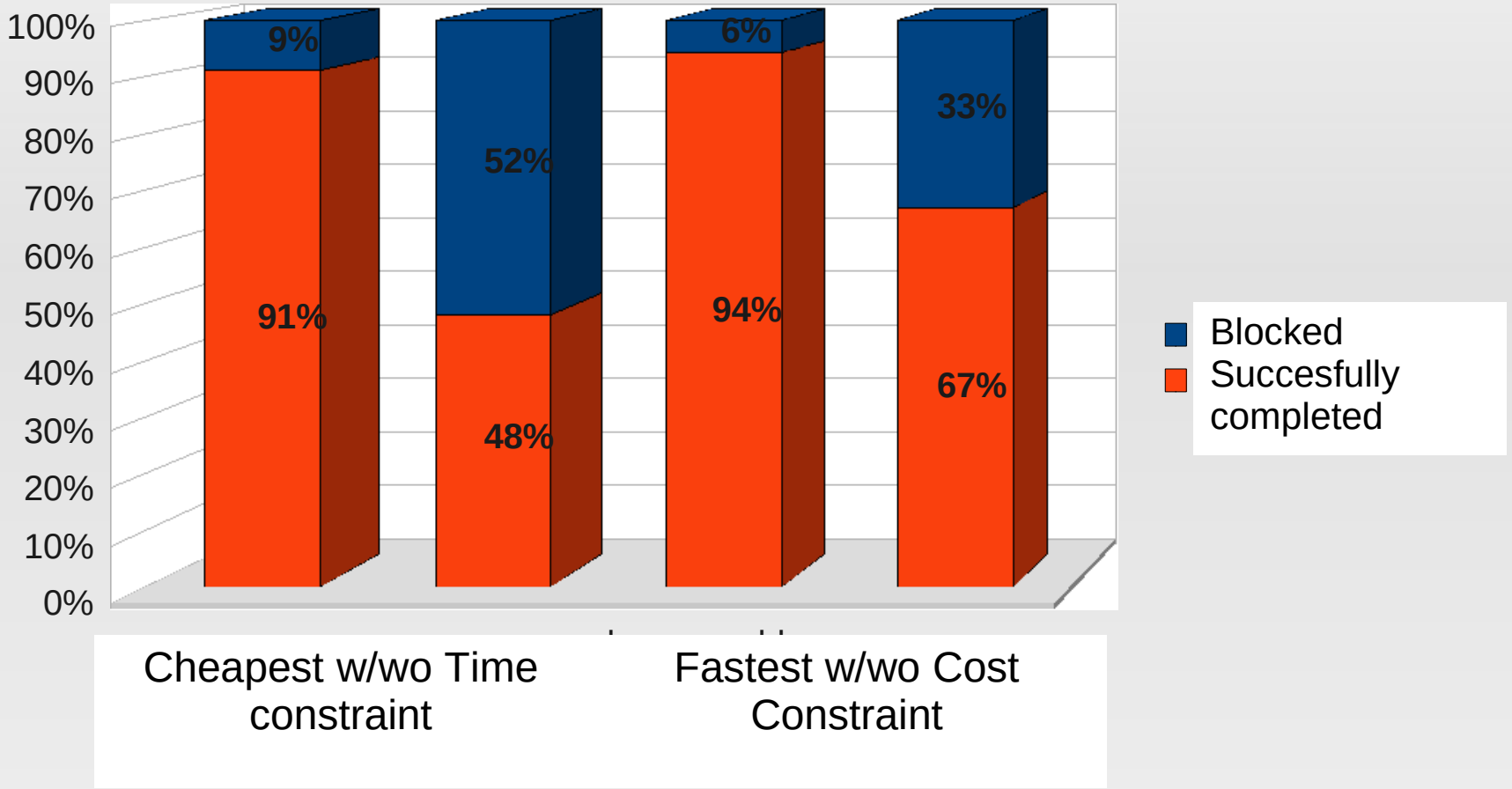
$$\sum_{\forall i} Y_i C_i + \sum_{\forall i,j} \frac{MI}{P} c_{i,j} \sum_{\forall k} X_{i,j,k} \le k_{max}$$

$$k_{max} \le K$$

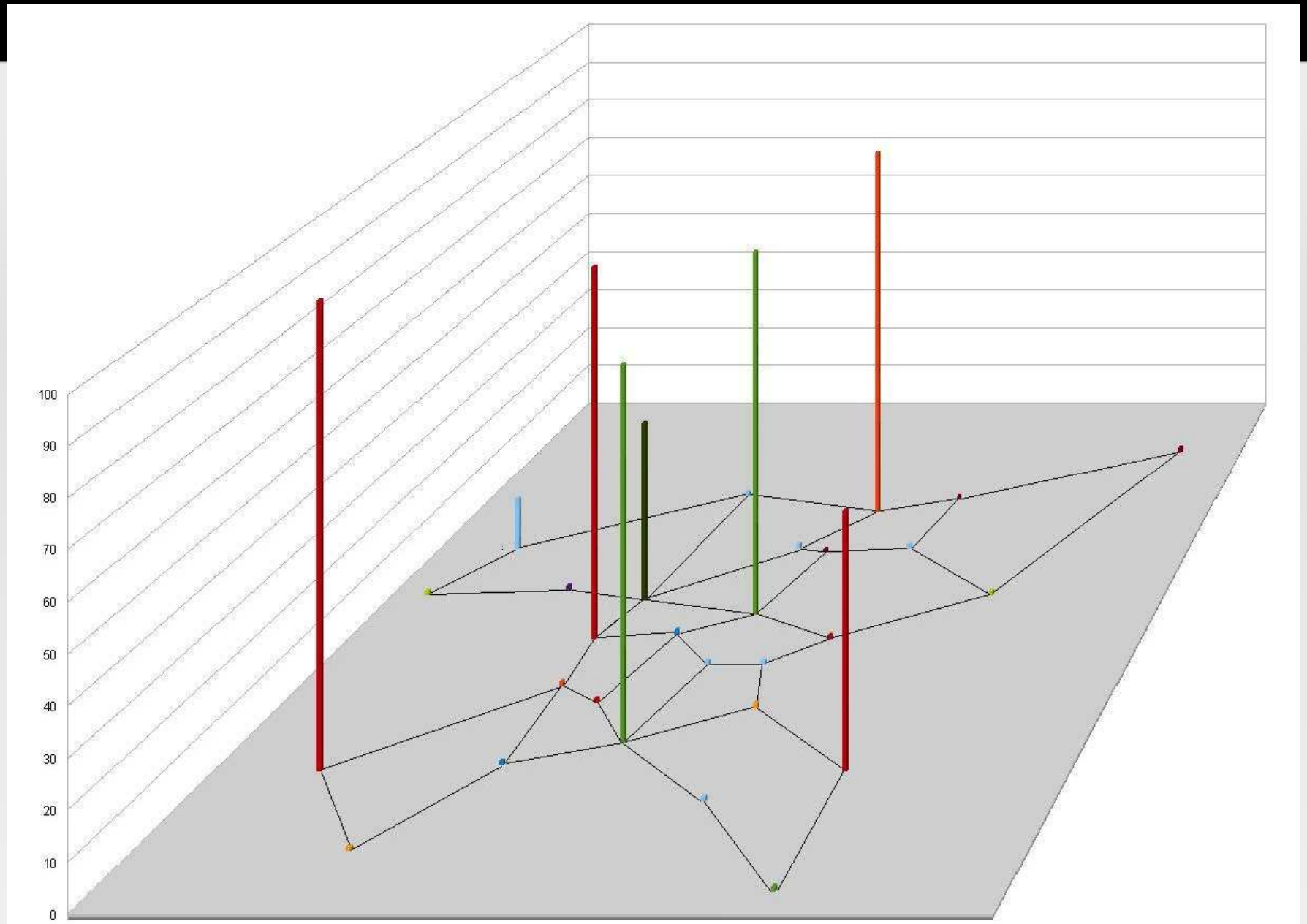
Objectives and Constraints

Variants	Constraints	Objectives
Cheapest	-	Cost
Fastest	-	Time
Cheapest within constrained time	Time-Limit: $t_{max}<T$	Cost
Fastest within limited cost budget	Cost-Limit: $k_{max}<K$	Time

Effect of cost and time constraints on the blocking



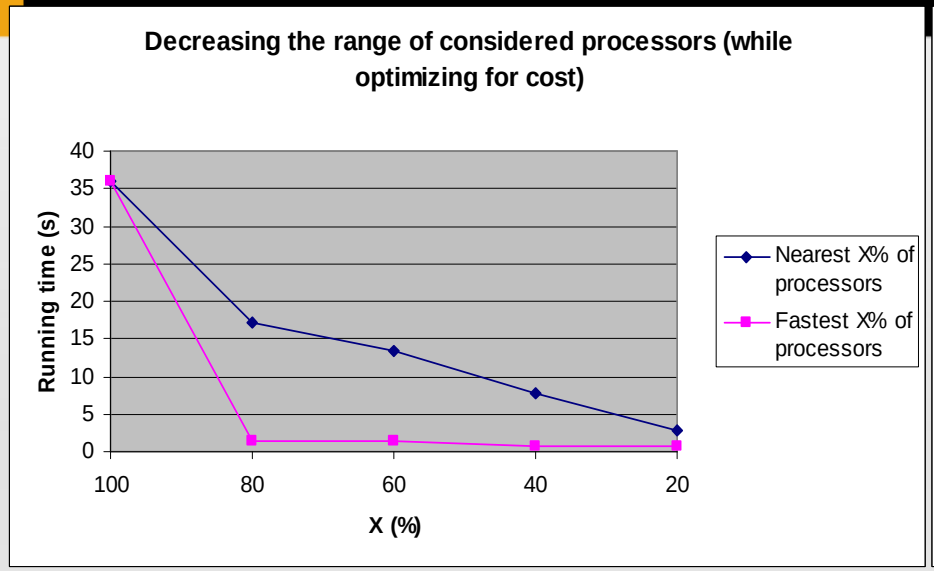
Geographical distribution of processing

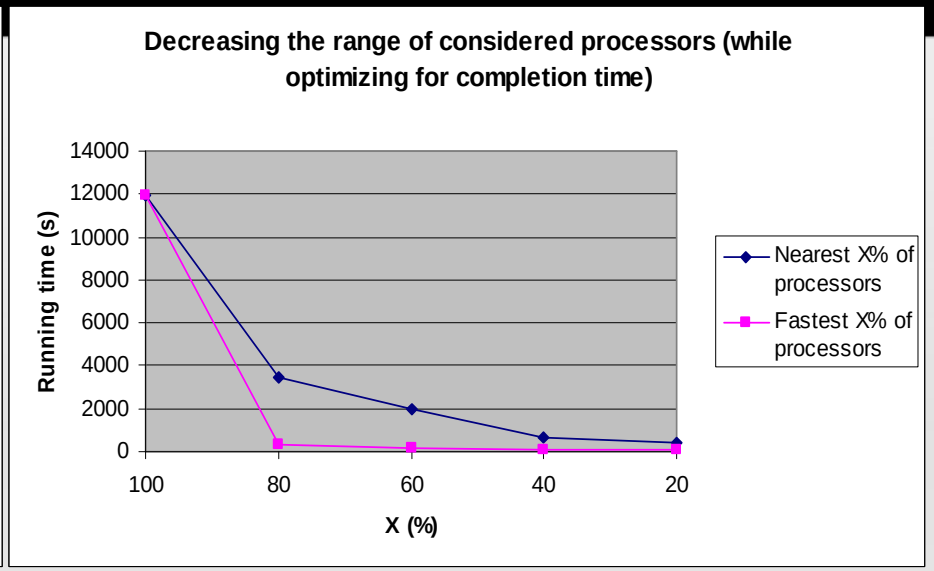


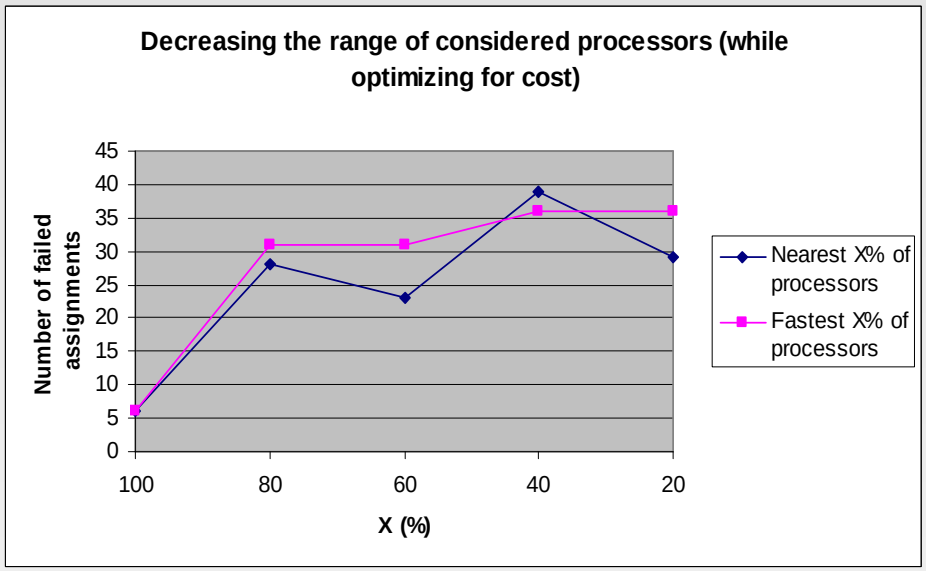
Speeding up the algorithm

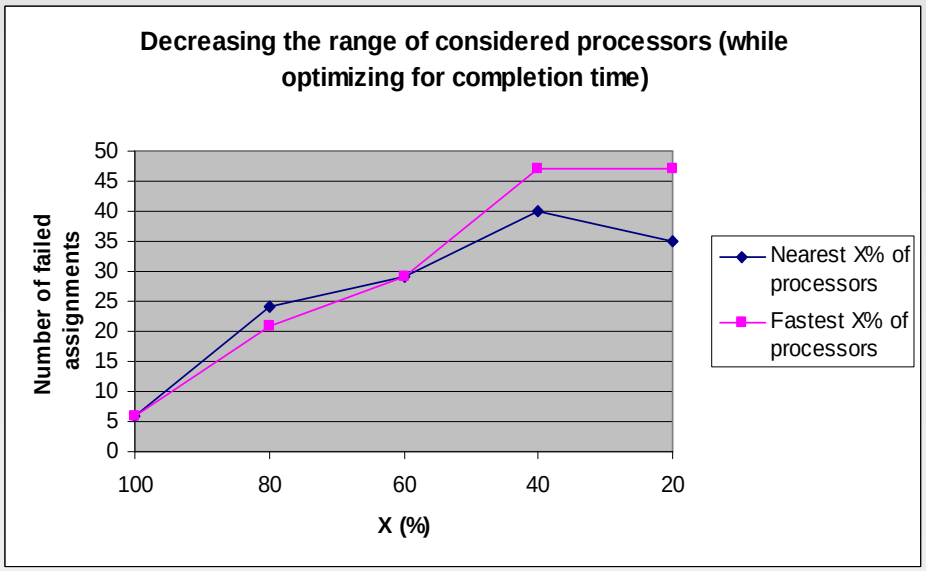
- Do not consider all the clusters, but only a certain ratio of fastest and closest
- Simulations
  - Cost 266 reference network (28 nodes, 41 links)
  - 34 processors in 7 random nodes
  - 8 wavelengths, each 2.5 Gbps
  - Each fourth demand is a GRID one (the rest is background traffic)

Running time and blocking while decreasing the range of clusters

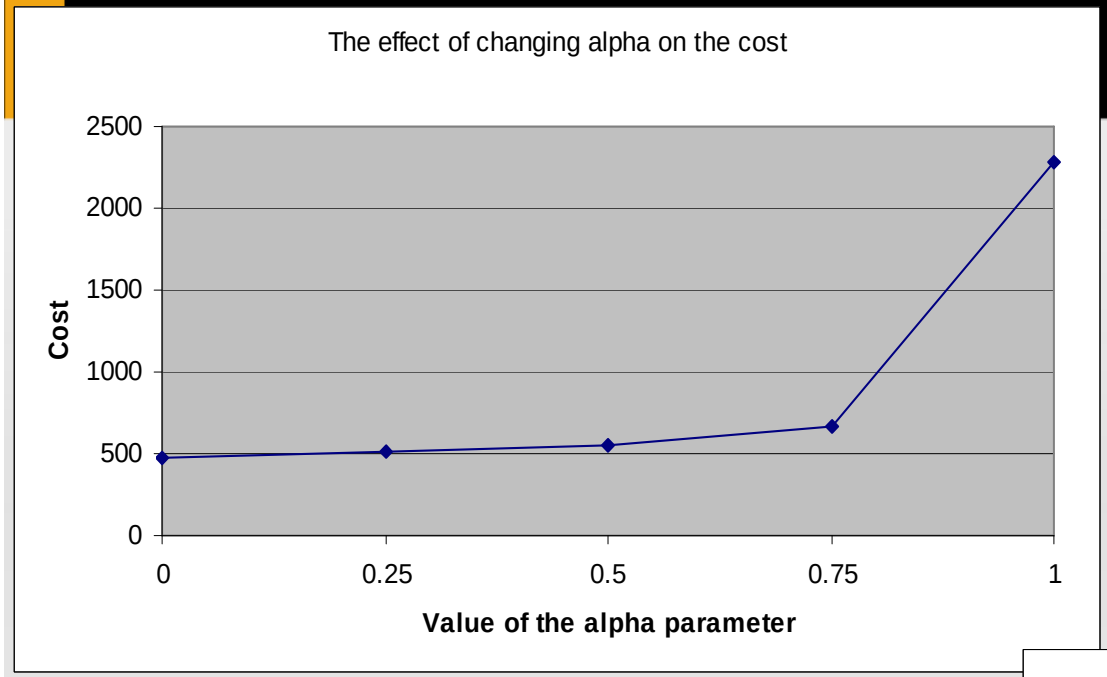


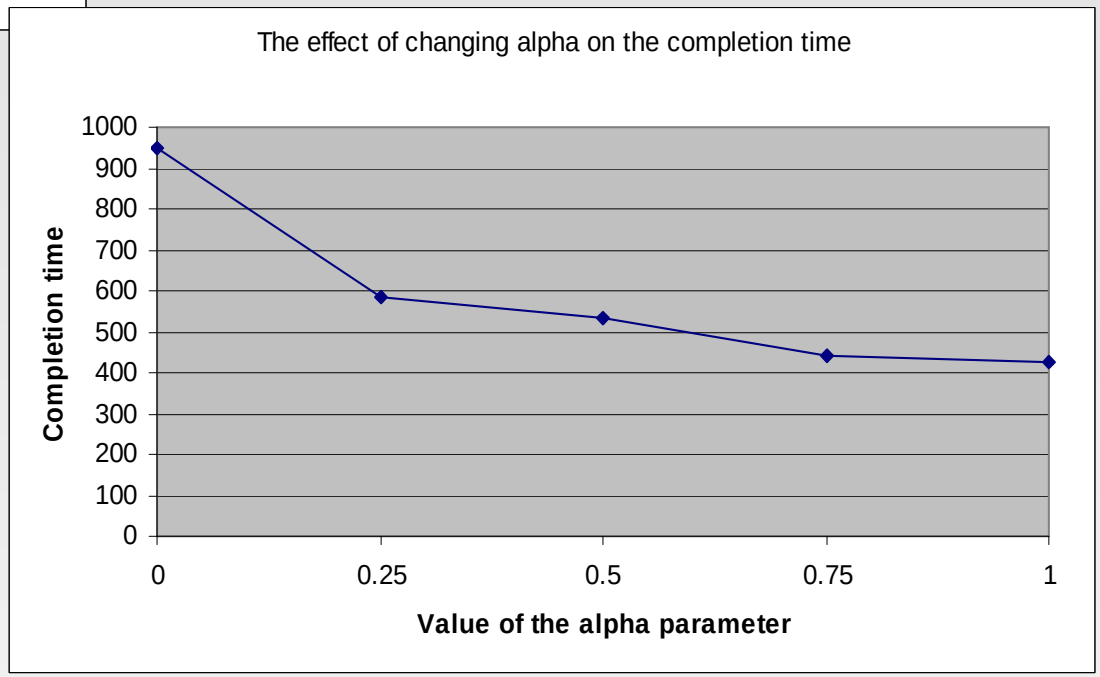






The effect of changing the α parameter





Summary

- Optimization framework
- Heuristics to speed-up the computations
- Mutual impact of networking and processing