CEng 471 Parallel Computing Grid

Erdal ERDAL
200511019

Outline

1. History of Grid
2. What is Grid?
3. Development of Grid.
4. Why Grid?
5. Grid Application Fields.
6. Sample Application (LHC).
7. Grid Approaches.
8. Grid in World.
9. EGEE
10. Ulakbilim and Grid.
11. Türkiye Grid with Photos
12. Türkiye Grid Applications.
13. Travelling Salesman Problem (TSP)
14. Travelling Salesman Problem with Parallel Computing

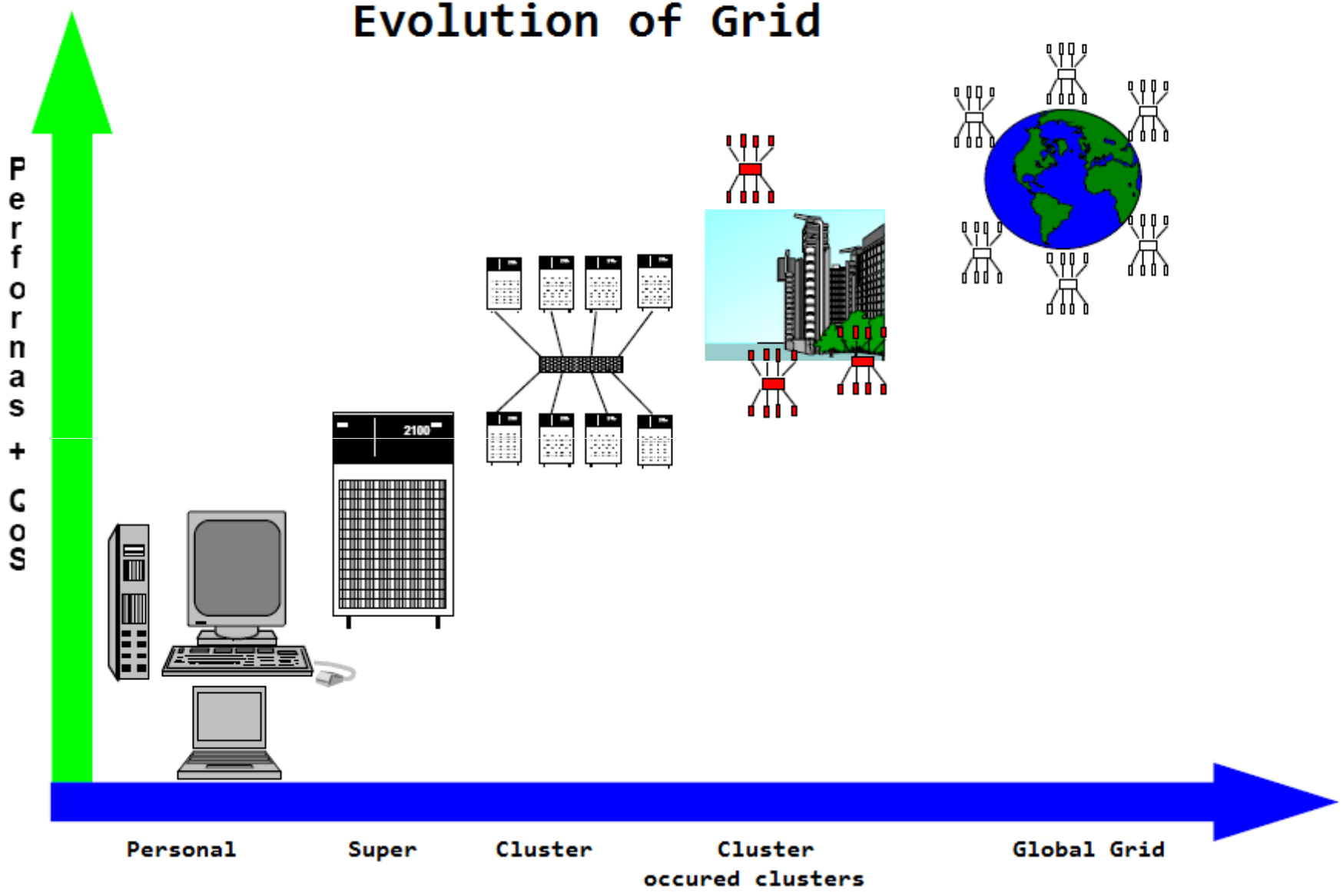
History of Grid

- 1970s - 80s: Internet
  - Email, ftp, etc
- 1980s - 90s: Remote-, Distributed-, Meta-computing
  - Beginning to think of ways to exploit distributed resources
  - US SC Centers Programs
    - Had to access remotely
    - Think of harnessing together
  - SC95: I-Way
  - First large scale experiments
  - Most failed, but I-Way a success!
- Post I-Way: The Grid
  - Explosion of activity, but still immature. Get prepared for future

What is Grid?

- Grid is a service that the resources of the computer sharing via internet. Such as CPU and memory.
- The aim of the Grid technology is convert all the global computer network to a big calculation source.
- Grid aims to use the resources efficiently that are using inefficiently.

Development of Grid.



Why Grid?

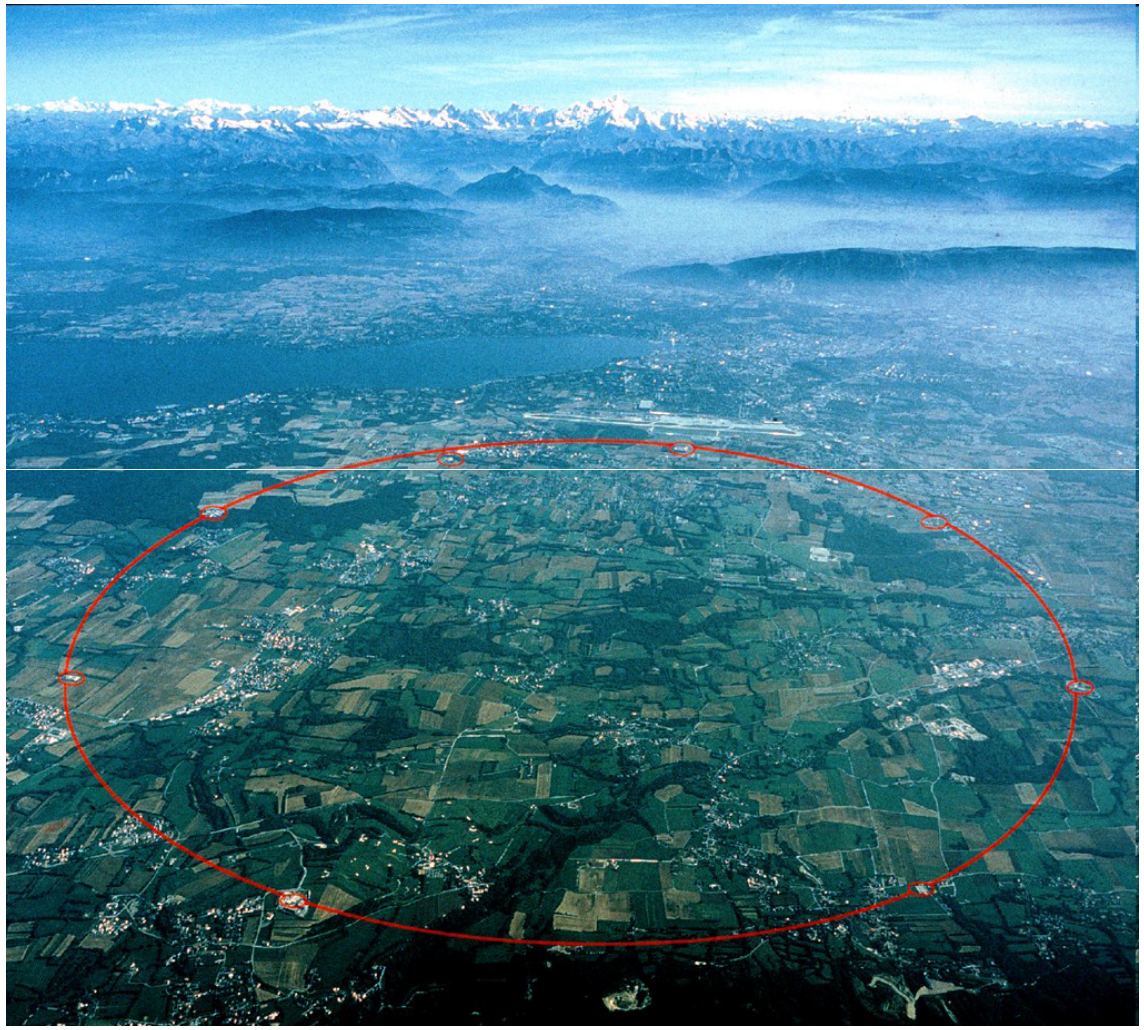
- Decrease the calculation time.
- Access the big data.
- Access the special equipment.
- Study with the other users.
- Resources can be sharing.
- Decrease the costs.

Grid Application Fields

- Fundamental Science (Mathematics, Physics, Chemistry)
- High Enegy Physics
- Medicine Science
- Genetic Science
- Brain Research
- Meteorology
- Electronic and Nano Technology
- Computer Science
- ...

LHC (Large Hadron Collider)

- Grid is used in Cern.
- 10 PetaByte in a year.
  (40 million event in a second)
- All the data must be processed and saved.



3 Approaches in Development Process

- Grid with Dedicated Clusters
  - EEGE
- Grid with Super Computers
  - Deisa Grid Infrastructure
- Grid with Desktop
  - Home Projects

Grid in World

Europe

- EGEE (II) (Enabling Grids for E-science)
- SEE-Grid (South-Eastern European Grid),
- DEISA (“Distributed European Infrastructure for Supercomputer Applications”)
- EUMEDGRID (Mediterranean Countries),
- EU-ChinaGrid, BalticGrid, NorduGrid, EELA (Latin America Countries), …

America

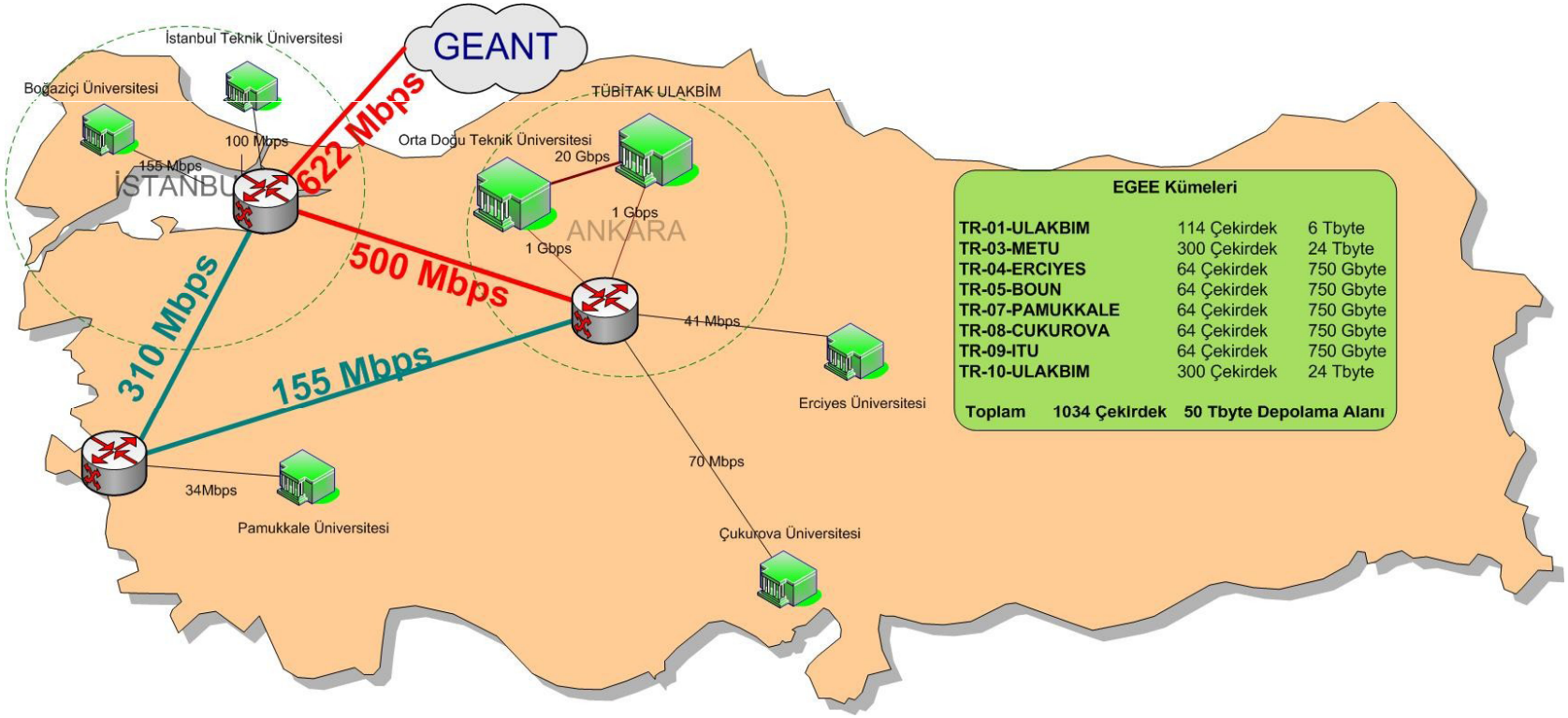
- TeraGrid

- The biggest Grid Creation.
- 224 sites in 49 countries.
- 38.000 CPU.
- 15 PB storage.

Tübitak-ULAKBİM has some roles in the Europe Grid Projects.

- Network Activies
- Special Service Activities.

- Türkiye National Grid Infrastructure (TUGA)
- TUGA started to work under the ULAKBİLİM in 2003.



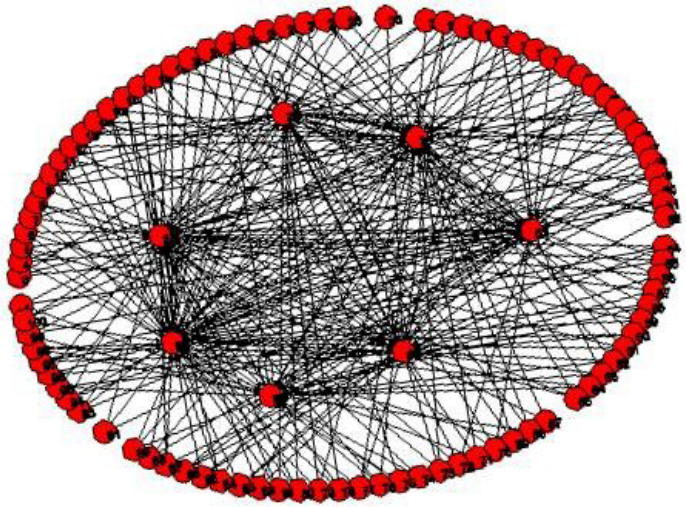
TR-Grid Infrastructure

SİTE	İSİM	Çekirdek	Veri Alanı	TİPİ
TÜBİTAK ULAKBİM	TR-01-ULAKBIM	114	5 Tbyte	EGEE, SEE-GRID, EUMEDGRID
ODTÜ	TR-03-METU	300	20 Tbyte	EGEE
Erciyes Üniversitesi	TR-04-ERCIYES	64	1 Tbyte	EGEE
Boğaziçi Üniversitesi	TR-05-BOUN	64	1 Tbyte	EGEE
Pamukkale Üniversitesi	TR-07-PAMUKKALE	64	1 Tbyte	EGEE
Çukurova Üniversitesi	**TR-08-CUKUROVA**	**64**	**1 Tbyte**	**EGEE**
İstanbul Teknik Üniversitesi	TR-09-ITU	64	1 Tbyte	EGEE
TÜBİTAK ULAKBİM	TR-10-ULAKBIM	300	20 Tbyte	EGEE
TOPLAM		**1034**	**50 Tbyte**	

First Cluster Project of Ulakbilim – Deniz - Levrek

Grid System of Ulakbim

Modeling the human memory simulation project



- Modelling the human memory
- Simulating human interaction
- Examining the affect of memory dimesion.

- 105 users in ULAKBİLİM and ODTU.
- 76 users have TR-Grid Certificate to work on Grid.
- 40 articles.
- TR-Grid is being used in 23 different universities.

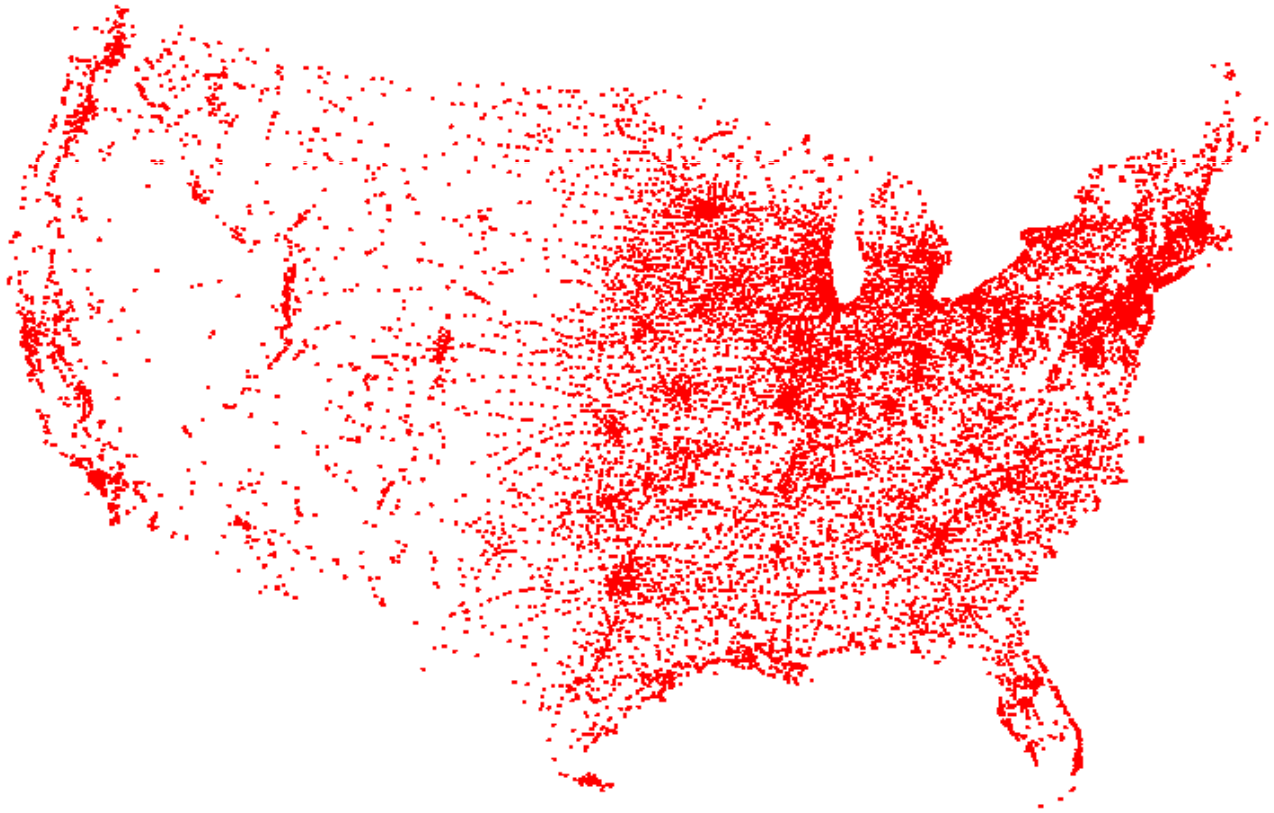
Sample Application of Parallel Computing

Given a collection of cities and the cost of travel between each pair of them, the **travelling salesman problem,** or **TSP** for short, is to find the cheapest way of visiting all of the cities and returning to your starting point.

The TSP is one of the most intensely studied problems in computational mathematics and yet no effective solution method is known for the general case.

TSP Sample

USA 13.509 is one of the larger TSP instances in TSPLIB. It contains 13,509 cities in the United States (all cities having a population of at least 500 at the time the instance was contributed to TSPLIB).



TSP Algorithms

1- Exact Algorithms

2- Heuristics

a) Constructive Heuristics

b) Iterative Improvement

c) Randomized Improvement

3- Greedy Algorithm

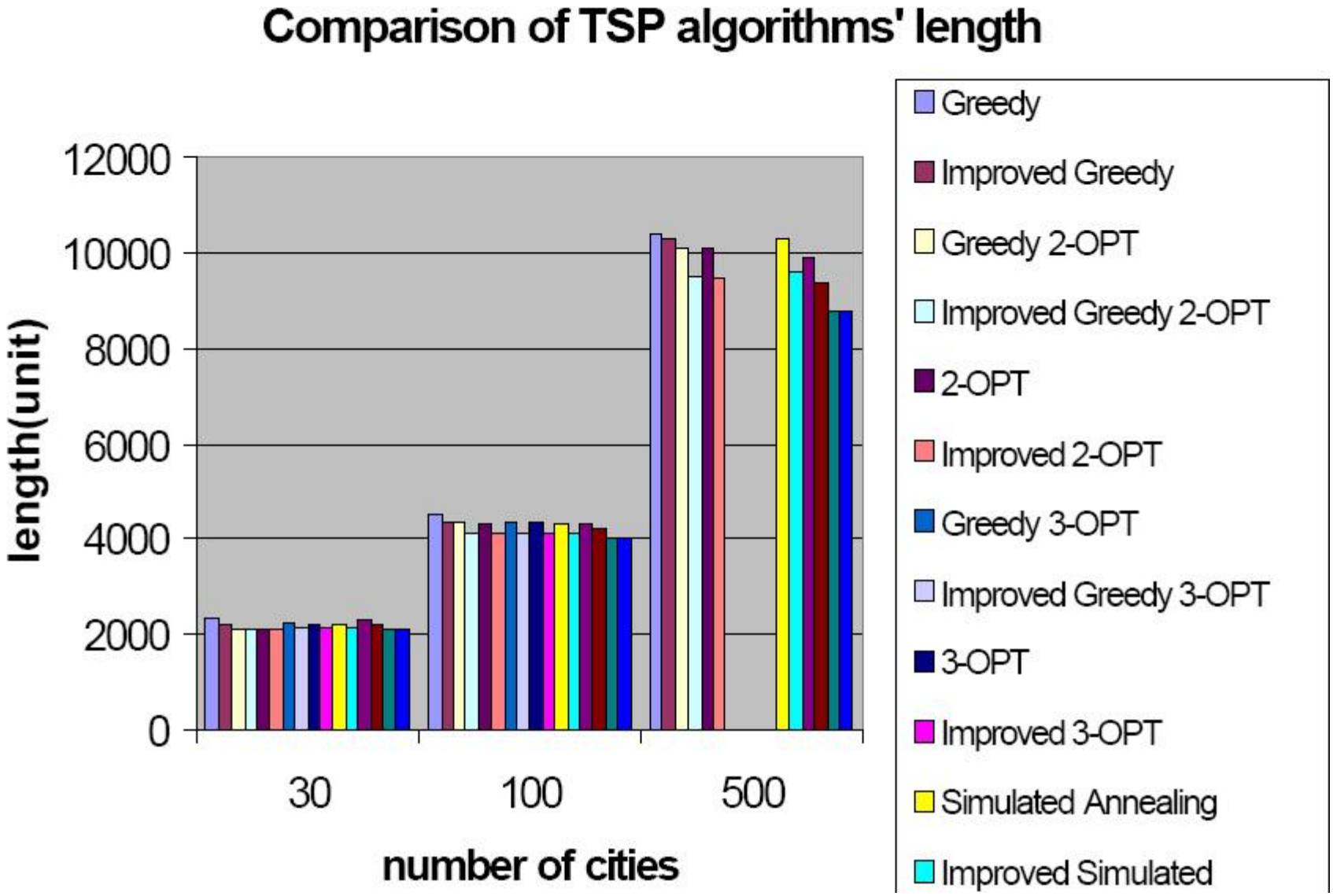
4- 2-Opt Algorithm

5- Greedy 2-Opt Algorithm

6- 3-Opt Algorithm

7- Genetic Algorithm

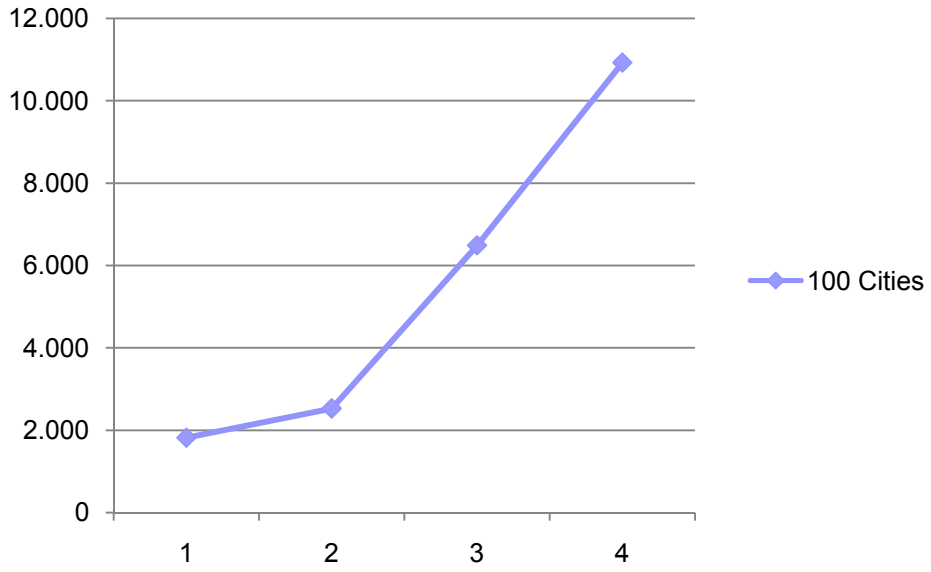
8- Simulated Annealing


Comparison of TSP algorithms' length
length(unit)
12000
10000
8000
6000
4000
2000
0
30
100
500
number of cities
Greedy
Improved Greedy
Greedy 2-OPT
Improved Greedy 2-OPT
2-OPT
Improved 2-OPT
Greedy 3-OPT
Improved Greedy 3-OPT
3-OPT
Improved 3-OPT
Simulated Annealing
Improved Simulated

In this application of TSP with parallel computing, all the cities are selected randomly. And the efficiency of paralel computing tested.

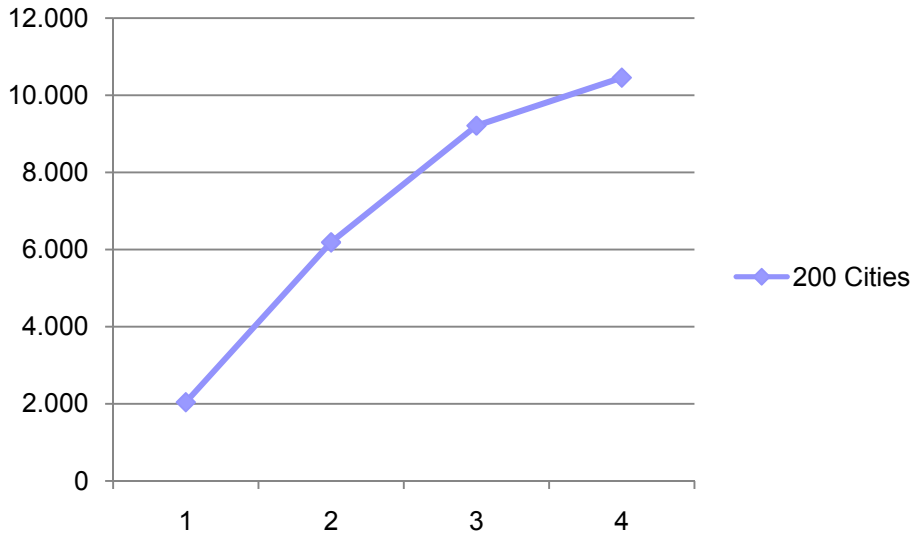
100 Cities

Number	Number of City	Number of CPU	Time
1	100	1	1.819
2	100	2	2.529
3	100	5	6.488
4	100	10	10.924



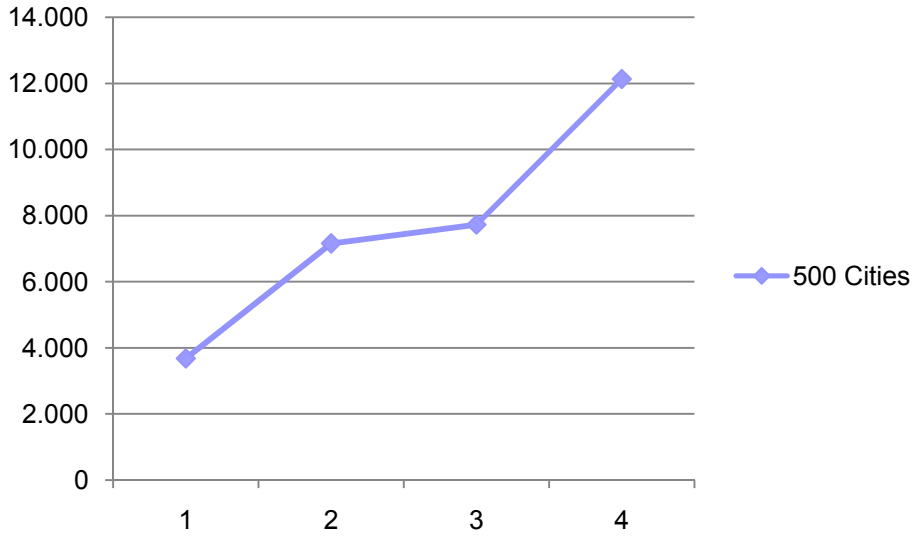
200 Cities

Number	Number of City	Number of CPU	Time
1	200	1	2.040
2	200	2	6.185
3	200	5	9.210
4	200	10	10.454



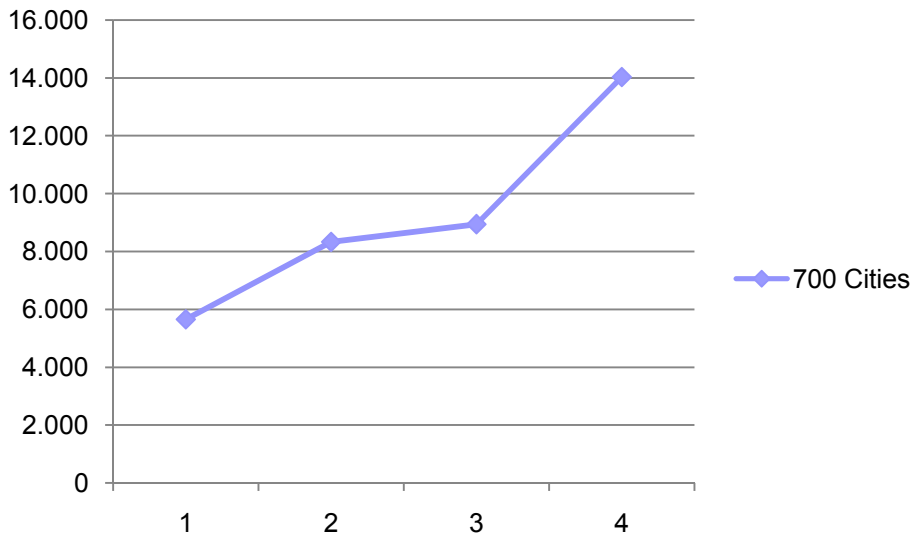
500 Cities

Number	Number of City	Number of CPU	Time
1	500	1	3.682
2	500	2	7.159
3	500	5	7.726
4	500	10	12.128



700 Cities

Number	Number of City	Number of CPU	Time
1	700	1	5.654
2	700	2	8.332
3	700	5	8.942
4	700	10	14.027



Conclusion :

In this application of TSP in parallel computing, we use random city selection algorithm so there is no any difficulty for processors.

As a result, it’s an unefficient way to use parallel computing for the random city selection algorithm in TSP.

CEng 471 Parallel Computing Grid

Thanks

Erdal ERDAL
200511019