Sensor Networks for Measurement and Data Reporting

Mequanint A. Moges
mmoges@uh.edu

COLLEGE OF TECHNOLOGY
University of Houston

Motivation

- Job request from various sources
  - Sensor generated data
  - Data intensive computing
  - Computational grids

- Challenge
  - Minimum processing time
  - Minimum energy use

- Solution
  - Efficient job scheduling strategies
    - Divisible Load Theory

Topics

- Research Purpose
- Sensor Networks
- Divisible Load Theory
- Scheduling Strategies
- Performance
- Conclusion

Research Purpose

- Optimal Job Scheduling
  - Measurement time
  - Computation time
  - Communication time
- Modeling
  - Divisible load theory
  - Markov chain models
- Application areas
  - Sensor networks
  - Grid networks

Sensor Networks

- ü Research Purpose
- q Sensor Networks
- q Divisible Load Theory
- q Scheduling Strategies
- q Performance
- q Conclusion

Sensor Networks

Sense <=> Process <=> Communicate

- Integrated collaboration
  - Gathering accurate information in a distributed manner from
    - Inaccessible geographic area
    - Disaster area
    - Industrial location
  - Perform task
    - Data gathering
    - Event detection and estimation
  - Relay Information
    - Network topologies
    - Access technology

Sensor Networks (cont)

ü Implications

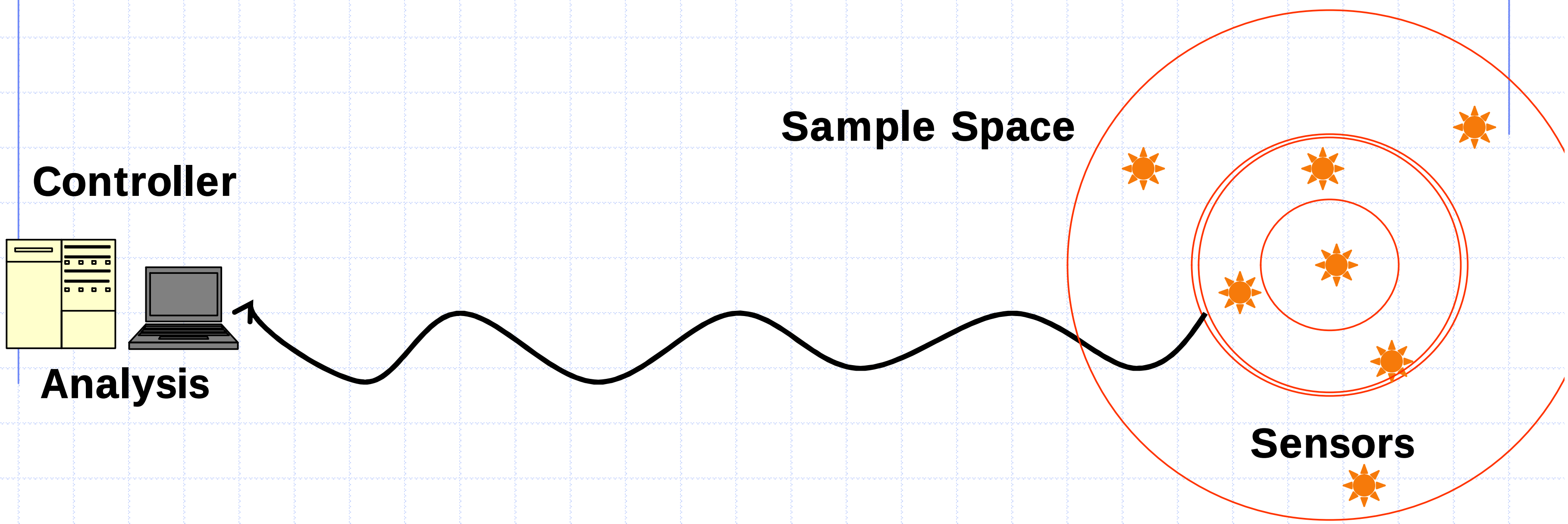
- Fixed wire-less network
- Avoid long distance communications
- No user attendance
- Deployed in large numbers
- Requires simple energy efficient routing

ü Objectives

- Minimum processing time
  - Fast response
- Energy use optimization
  - Extended life time

Our Scope

ü Sensing for measurement methods



- Sensors distributed on specific sample space
- Sensors measure stream of data (frequency)
- sensors report to the controller
- Sensors without computing power
- Sensors are power and energy limited

Our Scope

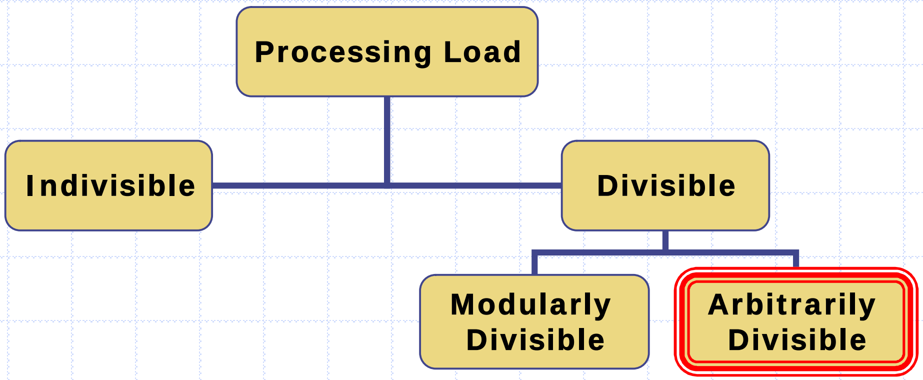
- Measurement Applications
  - Military Application
    - Space with large frequency range
    - Each sensor measures specific frequency and directional range
  - Very Large Array of Antennas (VLA)
    - Produce image of radio sky
      - Wide range of frequencies and directions
    - Extra-terrestrial signals

Divisible Load Theory

- ü Research Purpose
- ü Sensor Networks
- q Divisible Load Theory
- q Scheduling Strategies
- q Performance
- q Conclusion

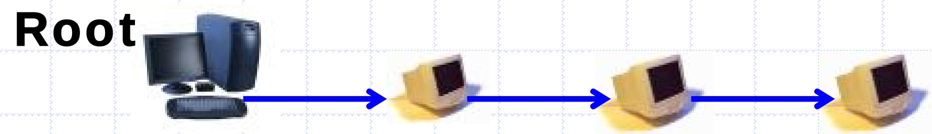
Divisible Load Theory

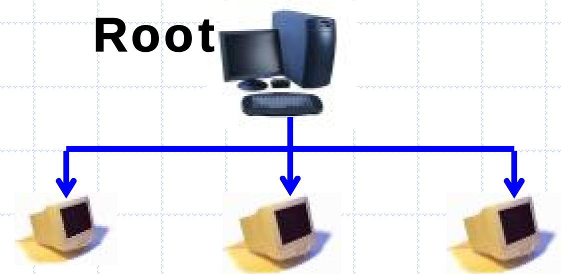
- Process of arbitrarily partitioning a large load
  - Shortest possible completion time
  - Lowest energy use
- Need for efficient partitioning
  - Prevalence of multiple processor systems
  - Data intensive computing
  - Ubiquity of sensor generated data

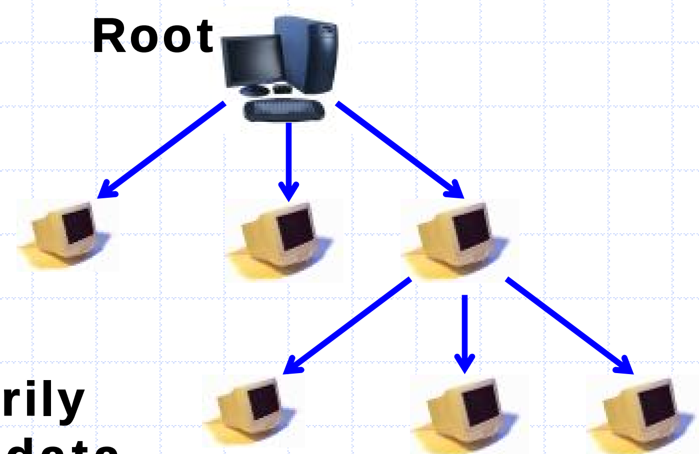


- Signal processing
- Image processing
- Sensor & Experimental data processing

System Model





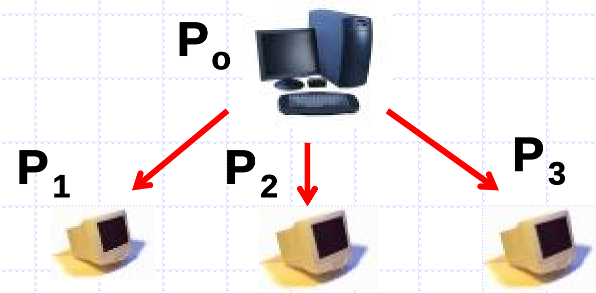


- **Topologies**
  - **Linear daisy chains**
  - **Bus networks**
  - **Tree Networks**

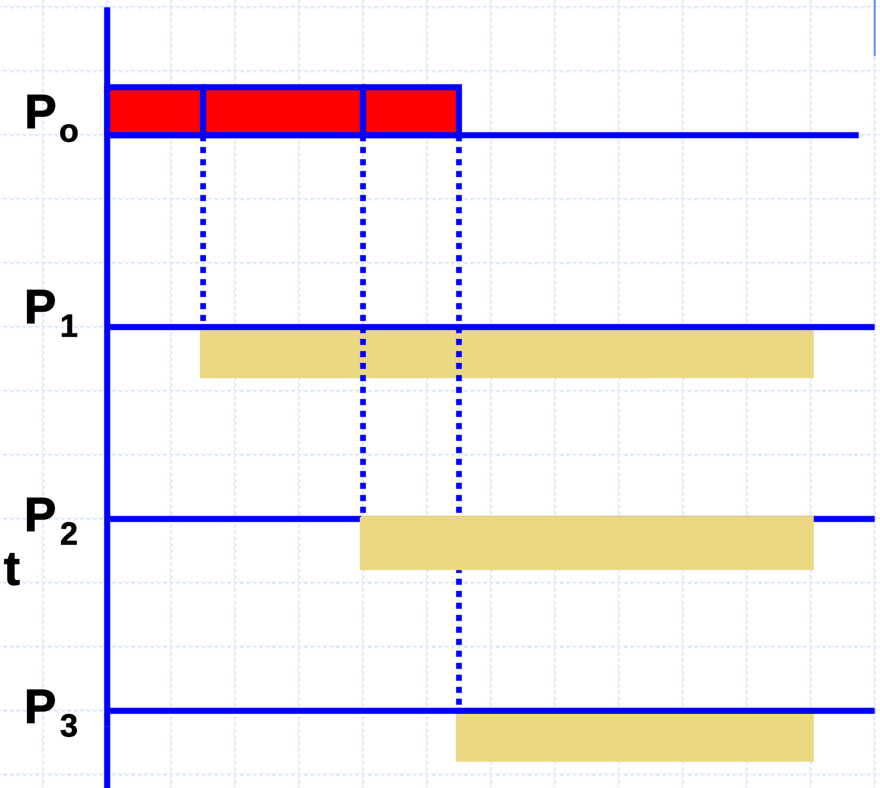
- **Assumptions**
  - **Load arrives at one particular processor (Root Processor)**
  - **Load can be partitioned arbitrarily**
  - **No precedence relation among data**
  - **Total load is normalized to be one**

System Model

DLT Model

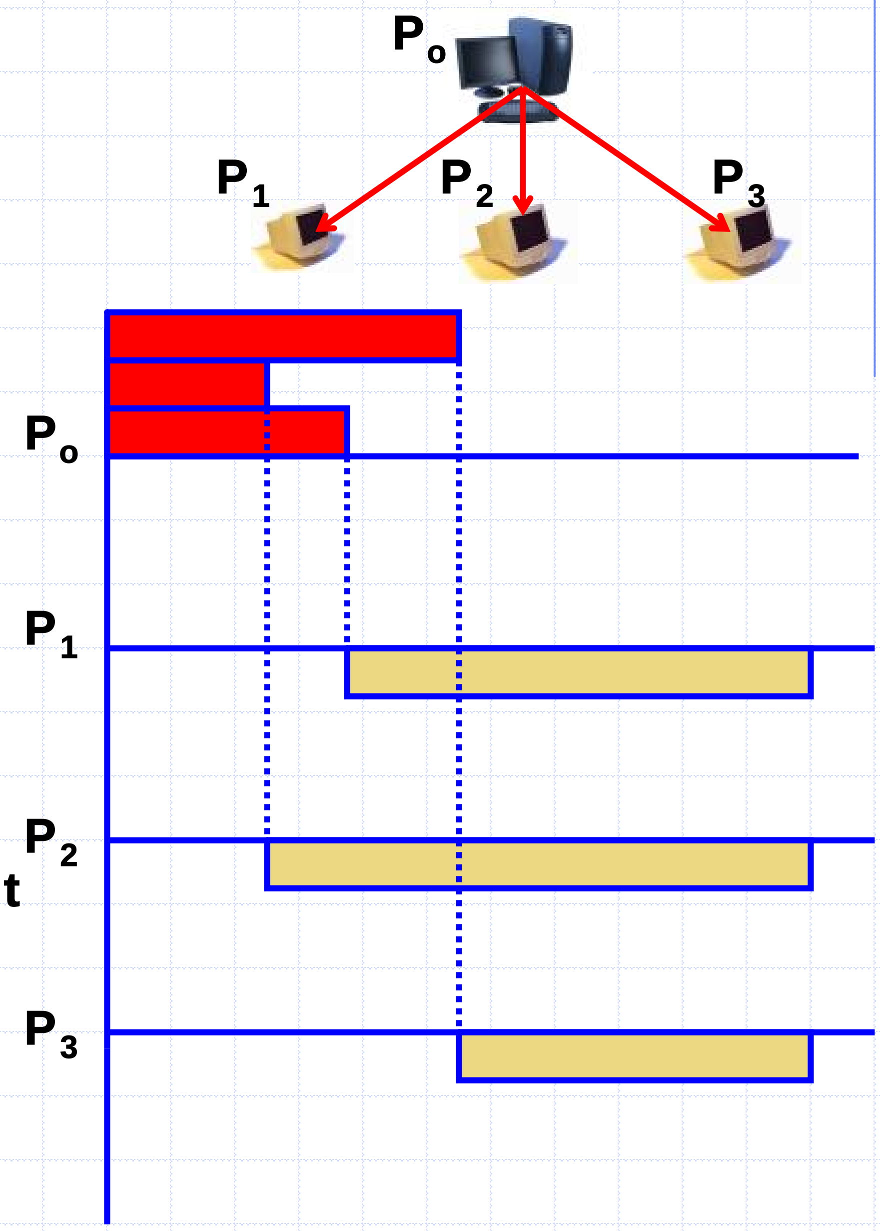


- Types of Loads
  - Communication loads
  - Computation loads
  - Measurement loads
- Scheduling policies
  - **Sequential**
  - Concurrent/Simultaneous
  - Single or multi-installment



Gantt-Chart Diagram

System Model



- Types of Loads
  - Communication loads
  - Computation loads
  - Measurement loads
- Scheduling policies
  - Sequential
  - Concurrent / Simultaneous
  - Single or multi-installment

System Model

- Optimality
  - Defined in the context of
    - Specific interconnection topology
    - Load scheduling policies
- Solution
  - Force **ALL** processors to finish at the same time.

Notations and Definitions

a_i – fraction of load processed by processor i

w_i – inverse computing speed of the i^{th} processor

y_i – inverse measurement speed of the i^{th} processor

Z_i - inverse transmission speed of the i^{th} link

T_{cp} – computation intensity

T_{cm} – communication intensity

T_{ms} – measurement intensity

$a_i\, w_i\, T_{cp}$ – time to process the i^{th} load fraction on the i^{th} processor

$a_i\, z_i\, T_{cm}$ – time to transmit the i^{th} load fraction on the i^{th} link

$a_i\, y_i\, T_{ms}$ – time to measure the i^{th} load fraction on the i^{th} processor

T_i – total time elapsed by the processor I

T_f – finish time of the process

Scheduling Strategies

- ü Research Purpose
- ü Sensor Networks
- ü Divisible Load Theory
- q Scheduling Strategies
- q Performance
- q Conclusion

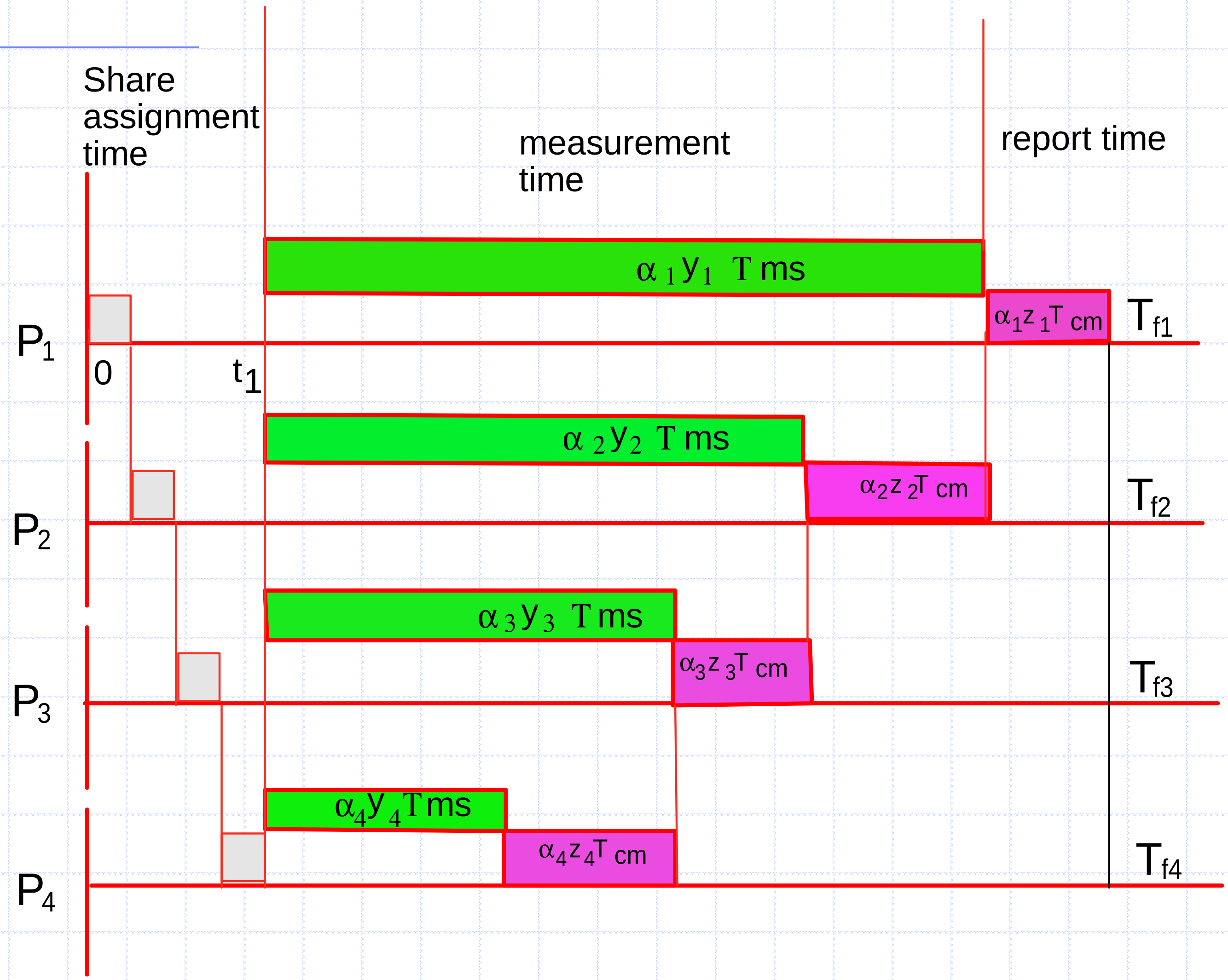
Data measurement and reporting strategies

- Three types
  - Simultaneous Measurement Start, Sequential Reporting
    - (SMS^{2}R)
  - Simultaneous Measurement Start, Simultaneous Reporting Time
    - (SMS^{2}RT)
  - Concurrent Measurement and Reporting
    - (CMR)

SMS²R

- ü **Single level tree network**
  - l **Three phases**
    - r **Job assignment phase**
      - r **Each processor will receive its measurement instruction**
    - r **Measurement phase**
      - r **Each processor starts measurement**
      - r **Measurement time è a¡y¡Tms**
    - r **Data Reporting phase**
      - r **Each processor starts reporting measured data**
      - r **Reporting is done sequentially**
      - r **Reporting time è a¡z¡Tcm**

SMS-SqR Timing Diagram



SMS²R

ü **Set of equations for the finish time**

- l P_1 è T_{f1} = a_1 y_1 Tms + a_1 z_1 Tcm
- l P_2 è T_{f2} = a_2 y_2 Tms + a_2 z_2 Tcm
- l P_3 è T_{f3} = a_3 y_3 Tms + a_3 z_3 Tcm
- l P_4 è T_{f4} = a_4 y_4 Tms + a_4 z_4 Tcm

$$T_f = t_1 + \frac{y_1 T_{ms} + z_1 T_{cm}}{1 + \sum_{i=2}^{N} \prod_{j=2}^{i} s_j}$$

ü **Normalization equation**

- l $a_1 + a_2 + a_2 + a_4 = 1$

$$s_j = \frac{y_{j-1} T_{ms}}{y_j T_{ms} + z_j Tcm}$$

ü **Different finish times**

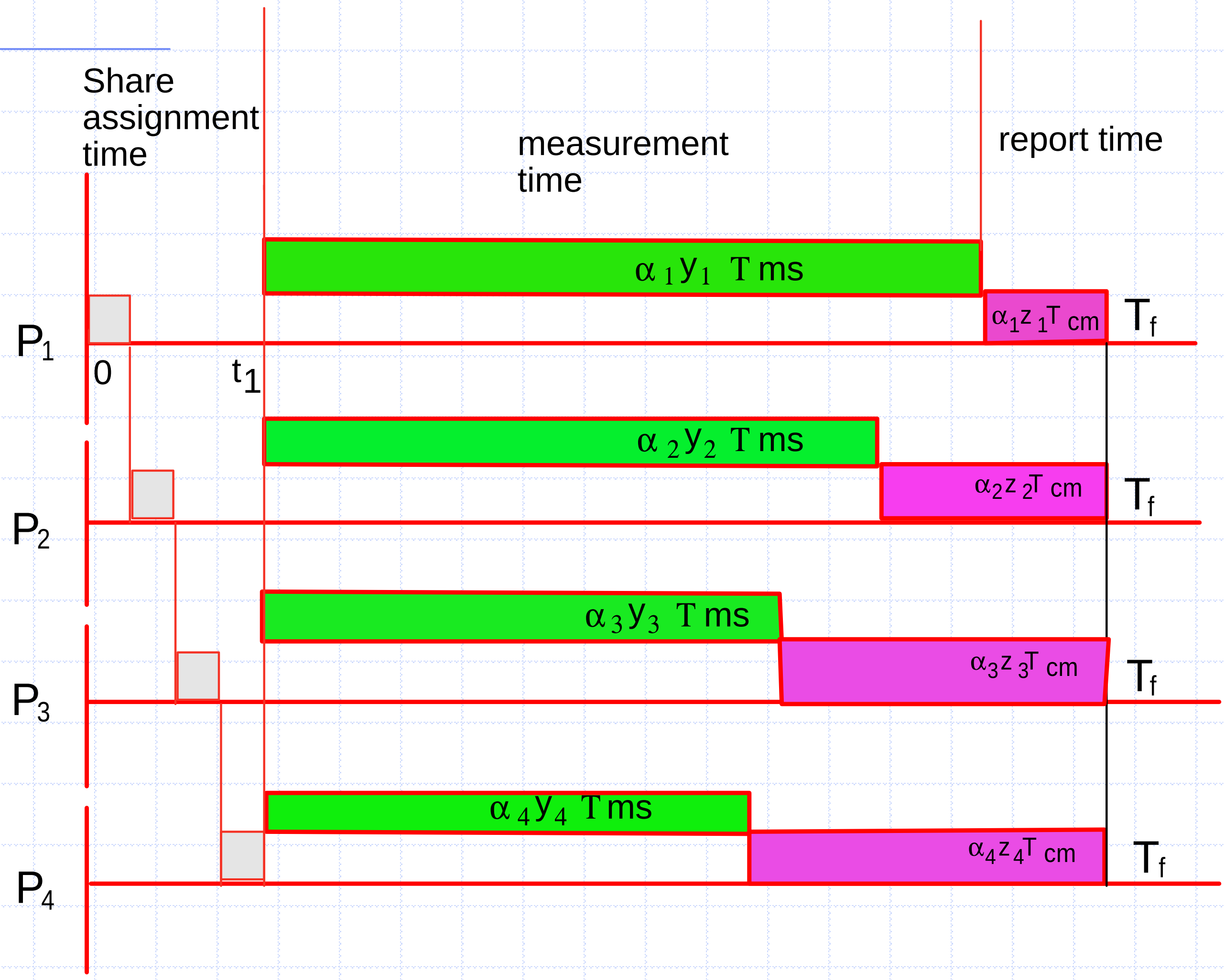
- l $T_{f(i+1)}$ = a_i y_i Tms

ü **Minimum finish time T_f**

SMS²RT

- Single level tree network
  - Job assignment phase
    - Each processor will receive its measurement instruction
  - Measurement phase
    - Each processor starts measurement
  - Data Reporting phase
    - Each processor starts reporting measured data
      - Simultaneously
      - Same finish time
- Each node has separate channel

SMS-SRT Timing Diagram



SMS²RT

- Set of equations for the finish time
  - P_1 è $T_1 = \alpha_1 y_1 Tms + \alpha_1 z_1 Tcm$
  - P_2 è $T_2 = \alpha_2 y_2 Tms + \alpha_2 z_2 Tcm$
  - P_3 è $T_3 = \alpha_3 y_3 Tms + \alpha_3 z_3 Tcm$
  - P_4 è $T_4 = \alpha_4 y_4 Tms + \alpha_4 z_4 Tcm$
- **In this case, we have:**
  - $T_1 = T_2 = T_3 = T_4 = T_f$
- **Normalization equation**
  - $\alpha_1 + \alpha_2 + \alpha_2 + \alpha_4 = 1$
- Minimum finish time

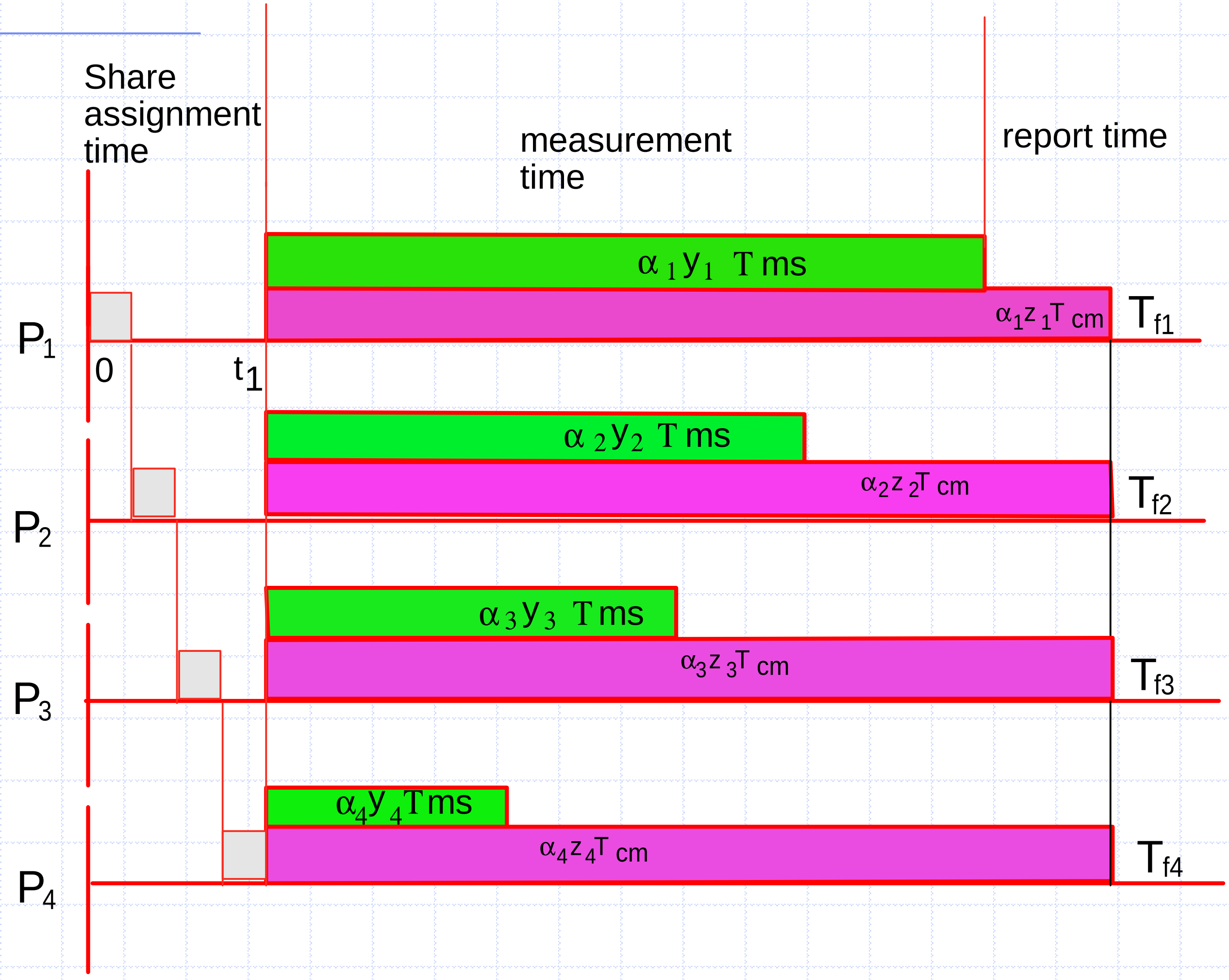
$$T_{fi} = t_1 + \frac{(y_i T_{ms} + z_i T_{cm})(\frac{1}{r_i})}{\sum_{i=1}^{N} (\frac{1}{r_i})},$$

$$r_i = y_i T_{ms} + z_i T_{cm}, i = 1,2,3\ldots$$

CMR

- ü Single level tree network
  - l Job assignment phase
    - r Each processor will receive its measurement instruction
  - l Measurement phase
    - r Each processor starts measurement
  - l Data Reporting phase
    - r Simultaneous data reporting
    - r Same finish time
- ü Each node has separate channel
- ü Concurrent measurement and data reporting

CMR Timing Diagram



CMR

- Set of equations for the finish time
  - P_1 → $T_1 = \alpha_1 z_1 Tcm$
  - P_2 → $T_2 = \alpha_2 z_2 Tcm$
  - P_3 → $T_3 = \alpha_3 z_3 Tcm$
  - P_4 → $T_4 = \alpha_4 z_4 Tcm$
- **Same time finish**
  - $T_1 = T_2 = T_3 = T_4 = T_f$
- **Normalization equation**
  - $\alpha_1 + \alpha_2 + \alpha_2 + \alpha_4 = 1$
- Minimum finish time

$$T_{fi} = t_1 + \frac{(z_i T_{cm})(\frac{1}{z_i})}{\sum_{i=1}^{N}(\frac{1}{z_i})},$$

Performance

- ü Research Purpose
- ü Sensor Networks
- ü Divisible Load Theory
- ü Scheduling Strategies
- q Performance
- q Conclusion

Performance

- Finish time Vs. Number of processors

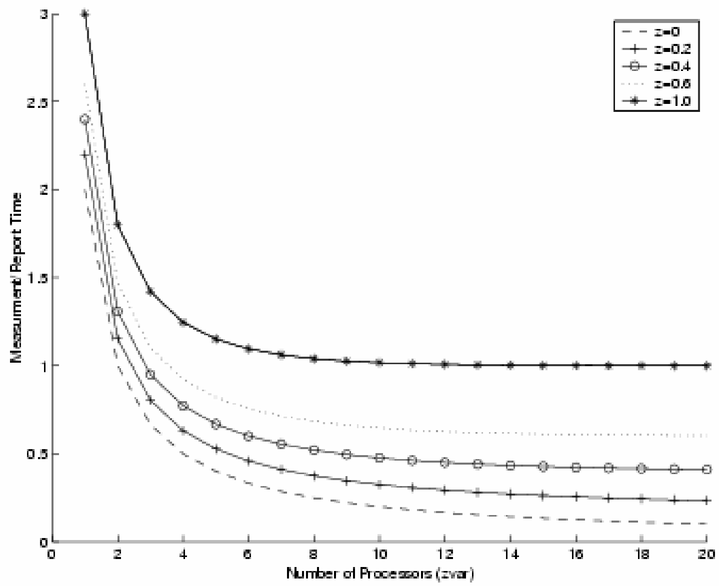
- Methodology
  - z_i is varied and y_i fixed
    - Impact of link speed
  - z_i is fixed and y_i varied
    - Impact of measurement speed

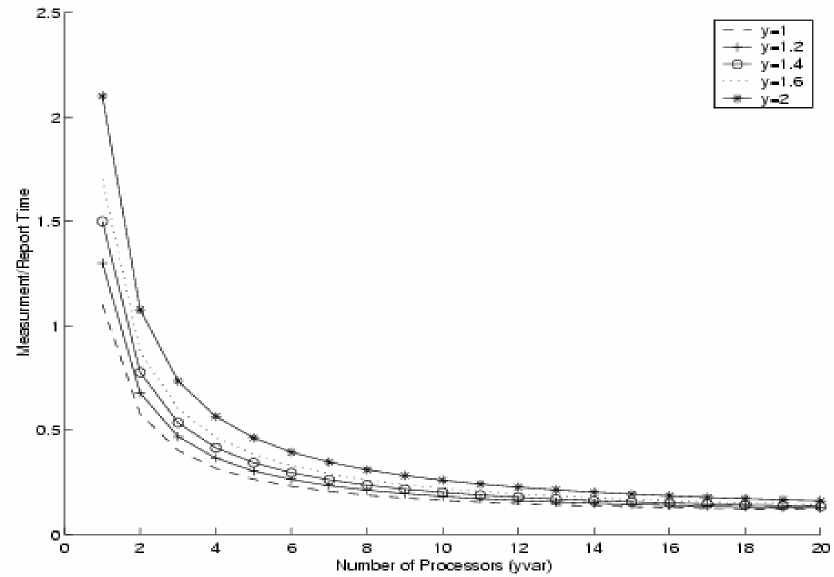
Performance

- SMS²R
  - z_i varied from 0 to 1
  - y_i fixed to be 2

z_i fixed to be 0.1

y_i varied from 1 -2



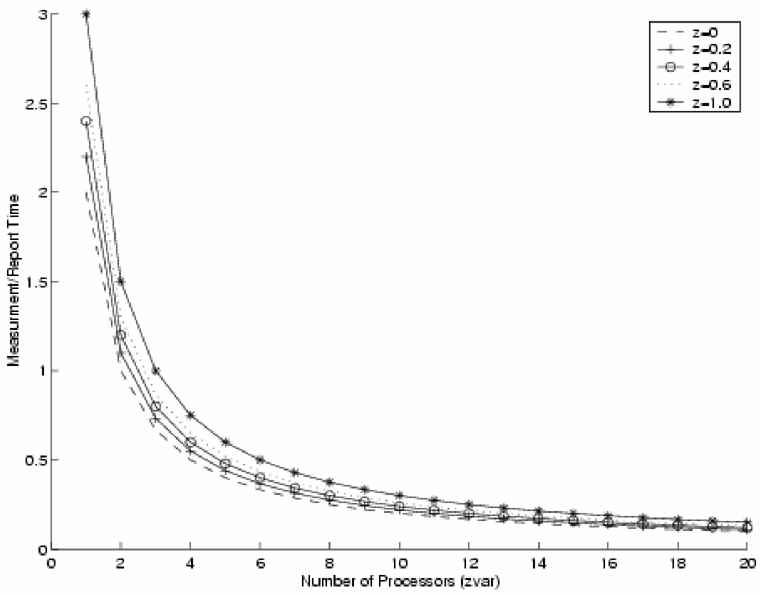


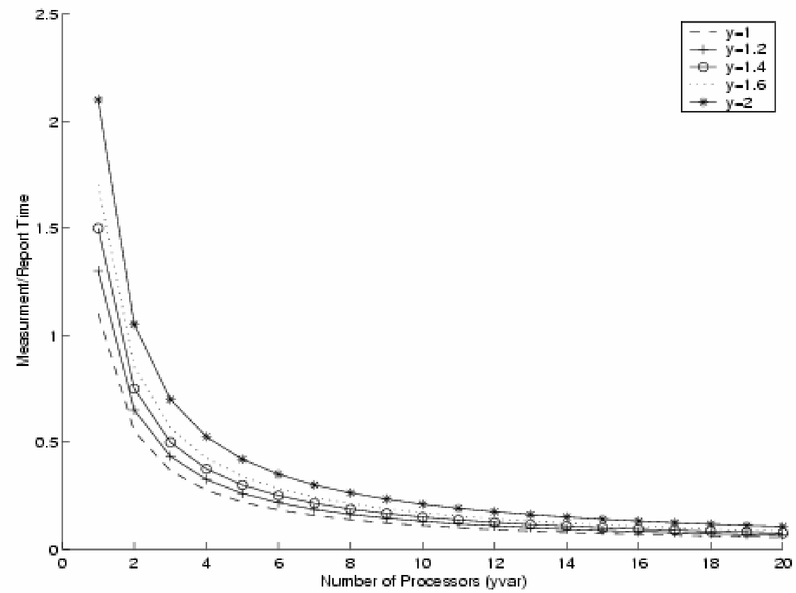
Performance

- **SMS²RT**
  - **z_i varied from 0 to 1**
  - **y_i fixed to be 2**

z_i fixed to be 0.1

y_i varied from 1 -2



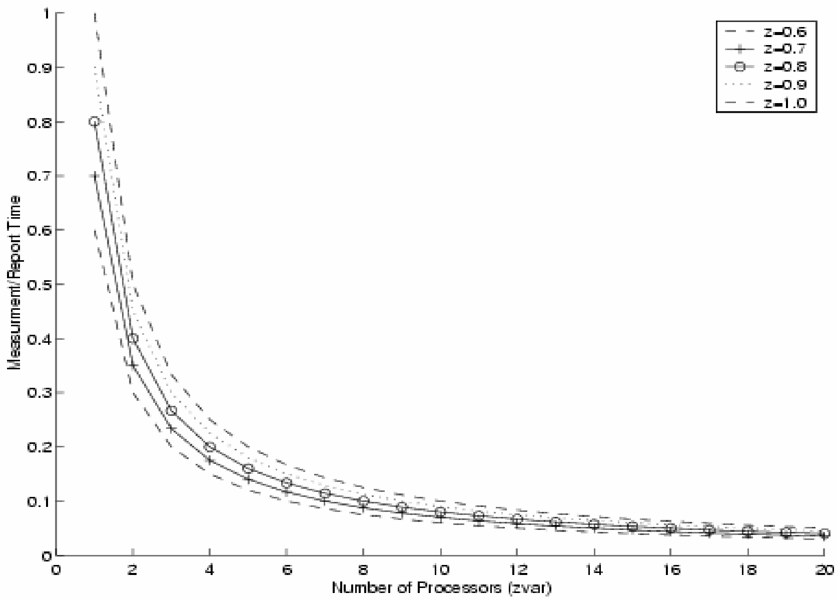


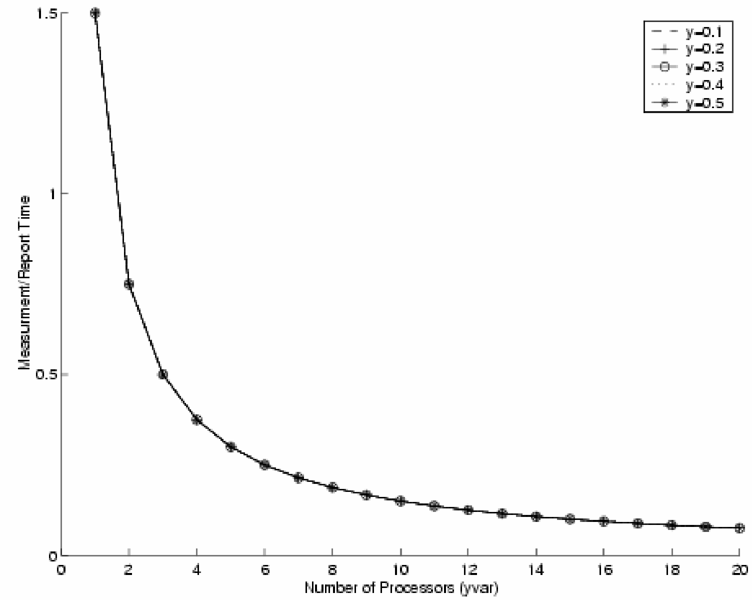
Performance

- CMR
  - z_i varied from 0 to 1
  - y_i fixed to be 2

z_i fixed to be 0.5

y_i varied from 0.1–0.5

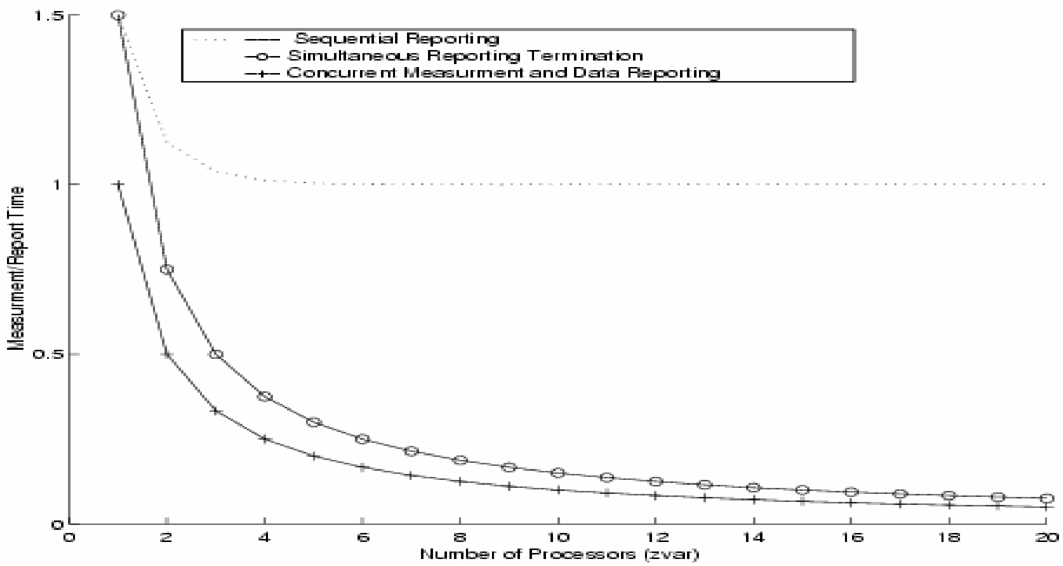




Performance

- **comparison**

  CMR performs WELL!



Conclusion

- ü Research Purpose
- ü Sensor Networks
- ü Divisible Load Theory
- ü Scheduling Strategies
- ü Performance
- q Conclusion

Conclusion

- Ongoing research
  - Monetary cost based energy use
  - Fault and rescheduling
  - STAR experiment in the RHIC project
    - Collaborative work with BNL
- Future research
  - Experimental level research
  - BlueGene/L , IBM
    - Massively-parallel supercomputing
  - Development of Integrated Scheduling Algorithm for Distributed Mobile Sensor Network
    - http://www.tech.uh.edu/isgrin/index.htm