

# GRID Enabled Biochip Analysis

2006. 12. 12.

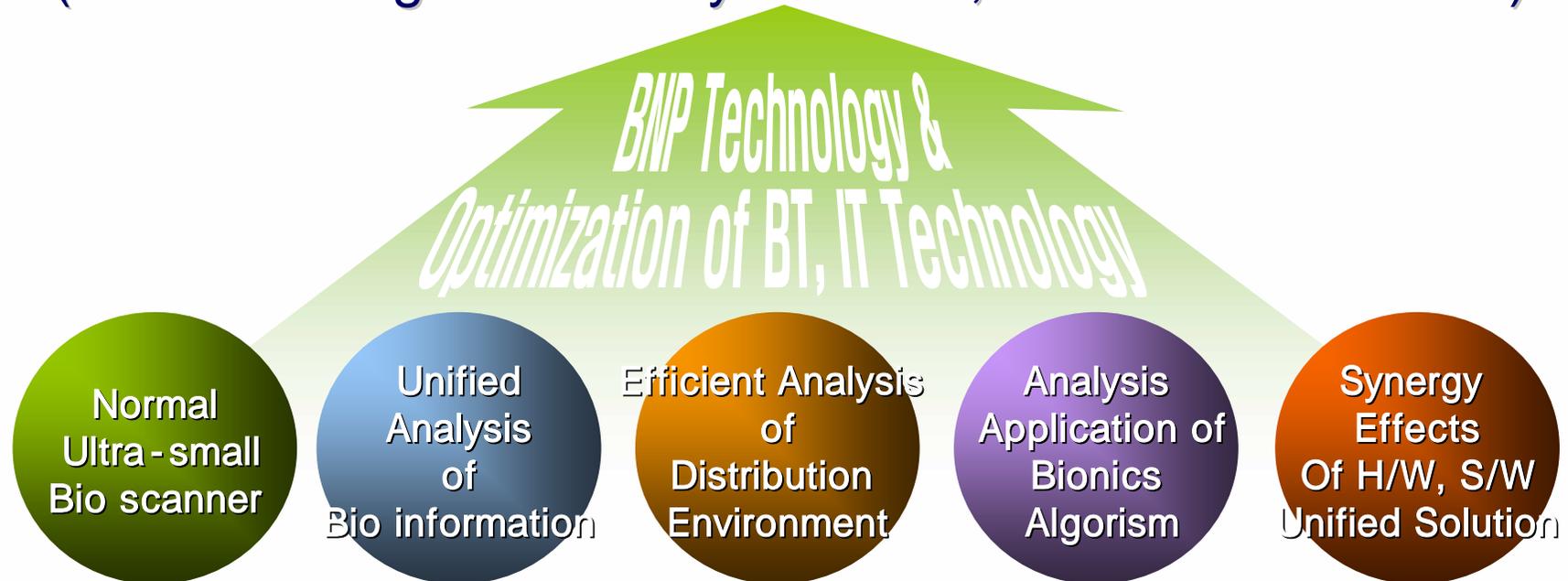
NationalGrid Tae Hoe Koo

# Contents

1. The Project Introduction
2. Biochip Analysis Process
3. The Application of Distribution Analysis

## Biochip Distribution Analysis System

(Medical Diagnosis Analysis Base, Commercialization)

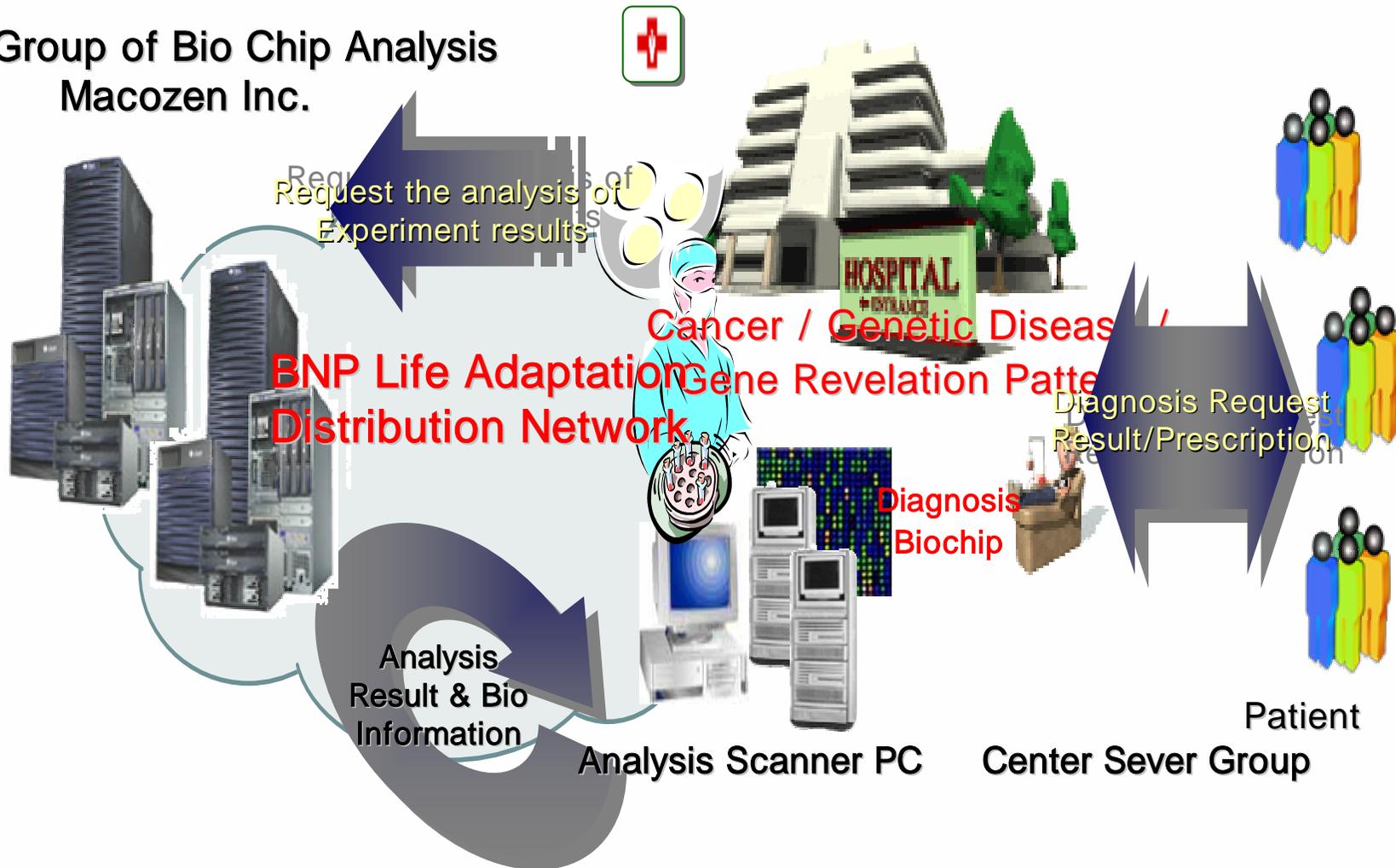


# 1-2. Concept Map of Development System

Project Introduction

Hospital/Diagnosis Center/Research Center

Sever Group of Bio Chip Analysis  
Macozen Inc.

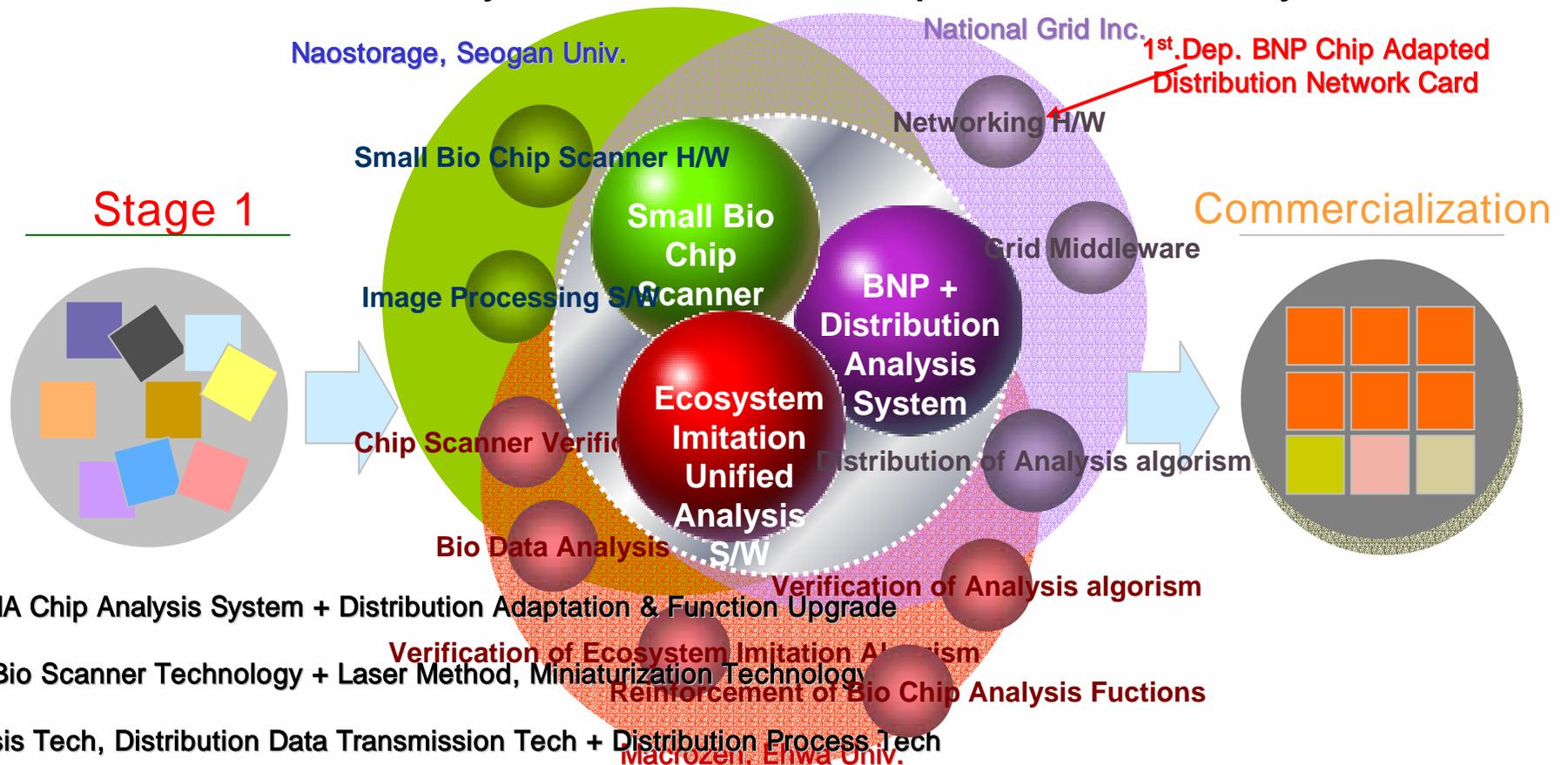


# 1-3. Stage & Related Institutions

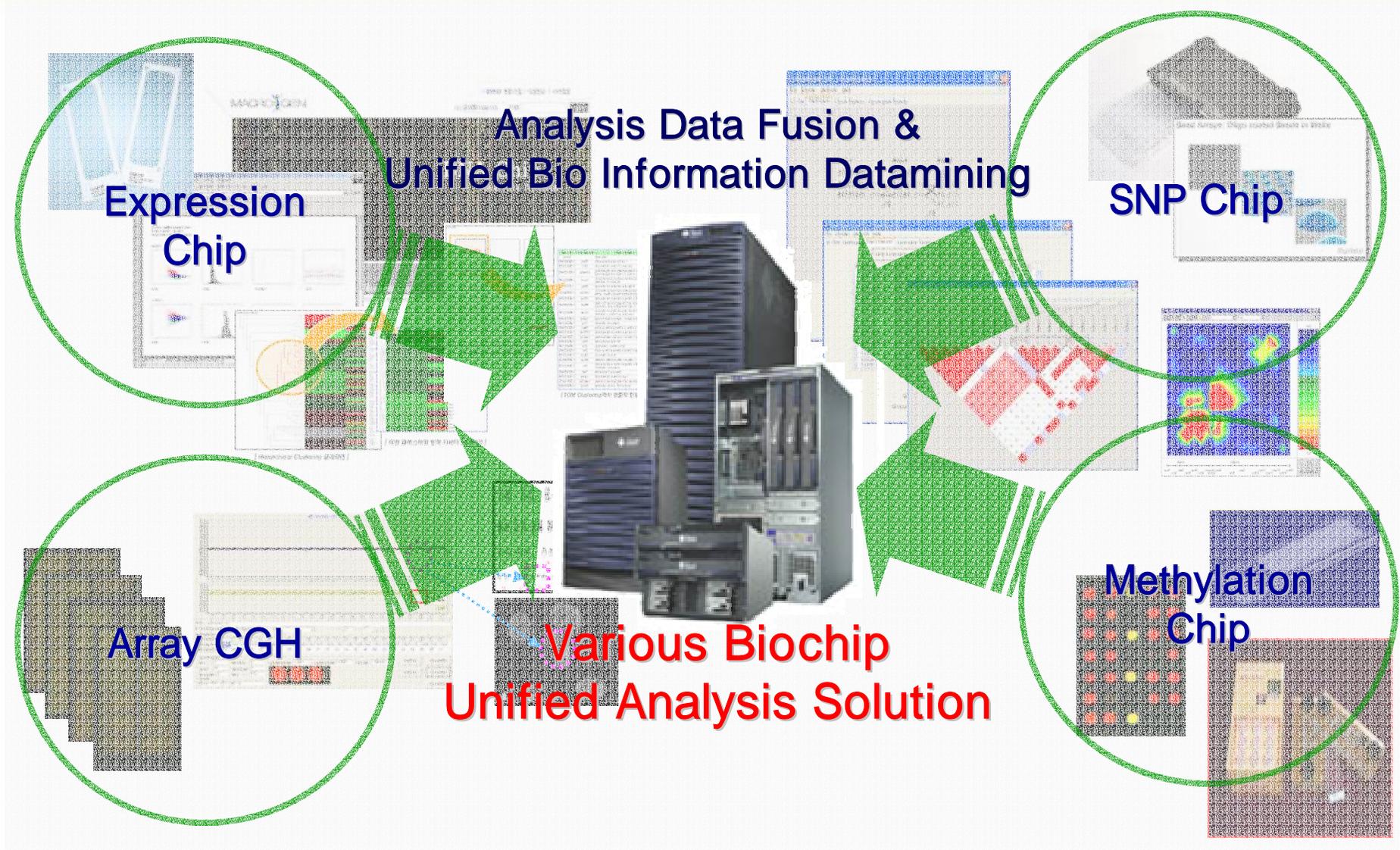
Project Introduction

## Stage 2

### Life Network System of Bio Chip Unified Analysis

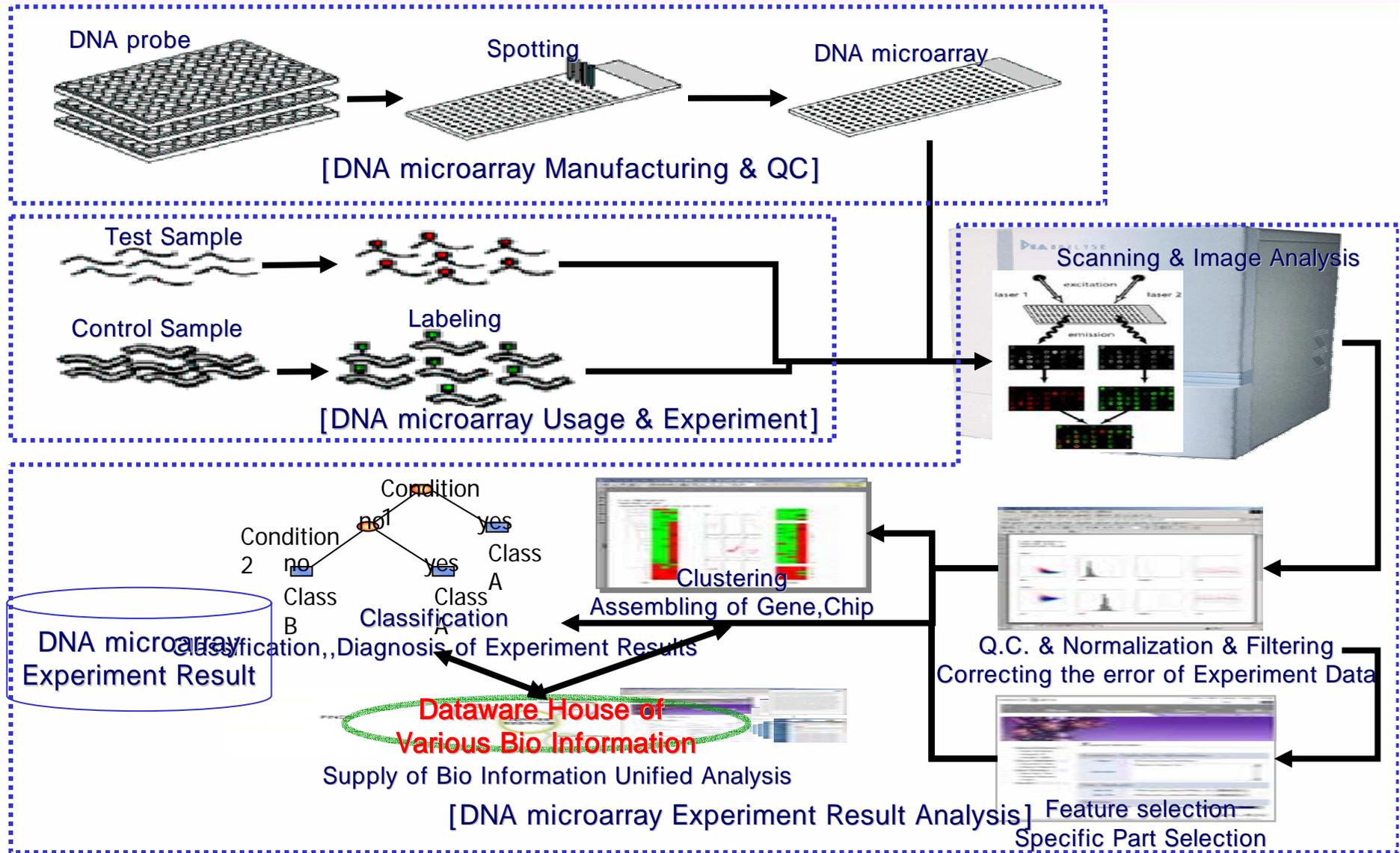


# 2-1. Biochip

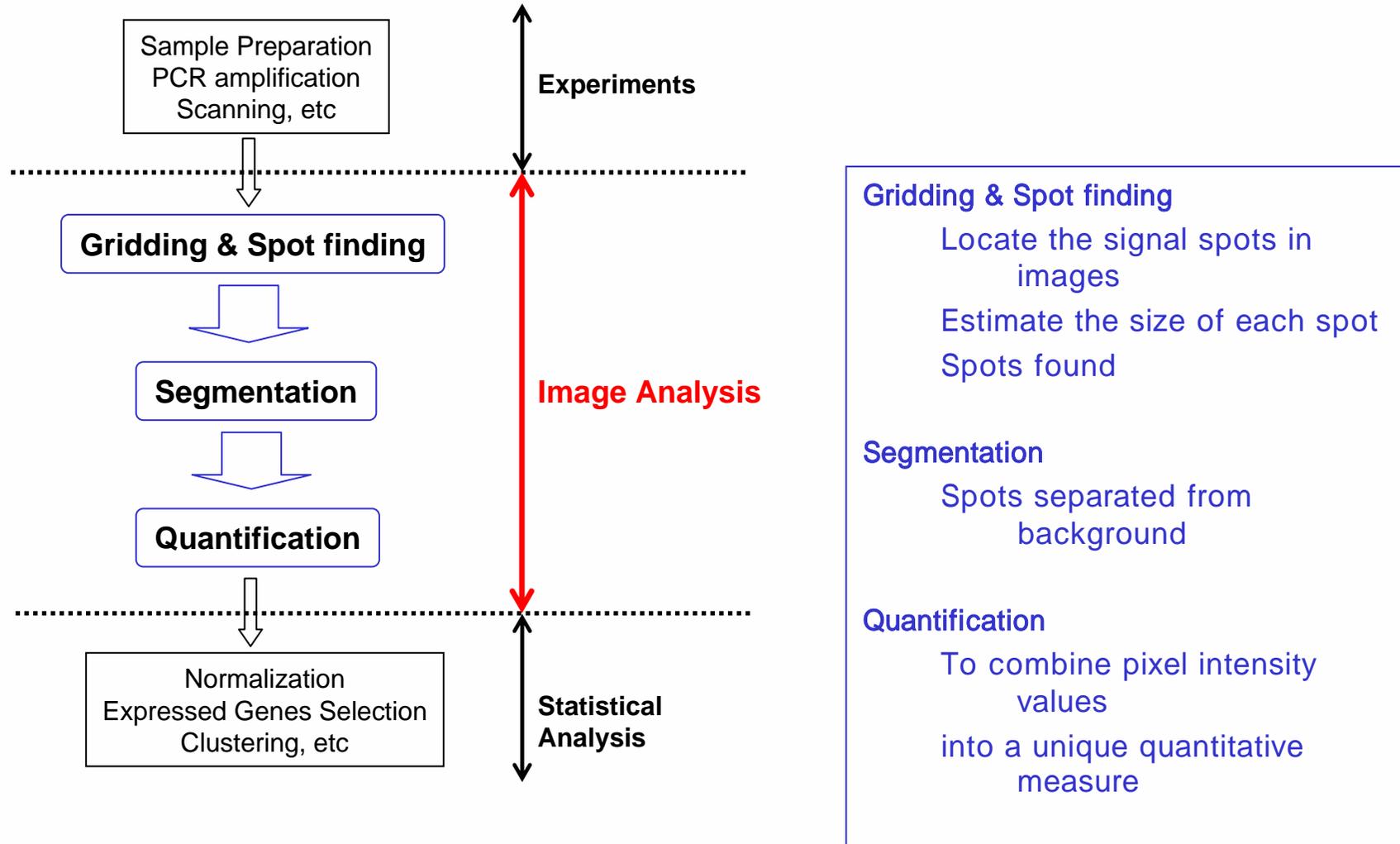


# 2-2. Total Process

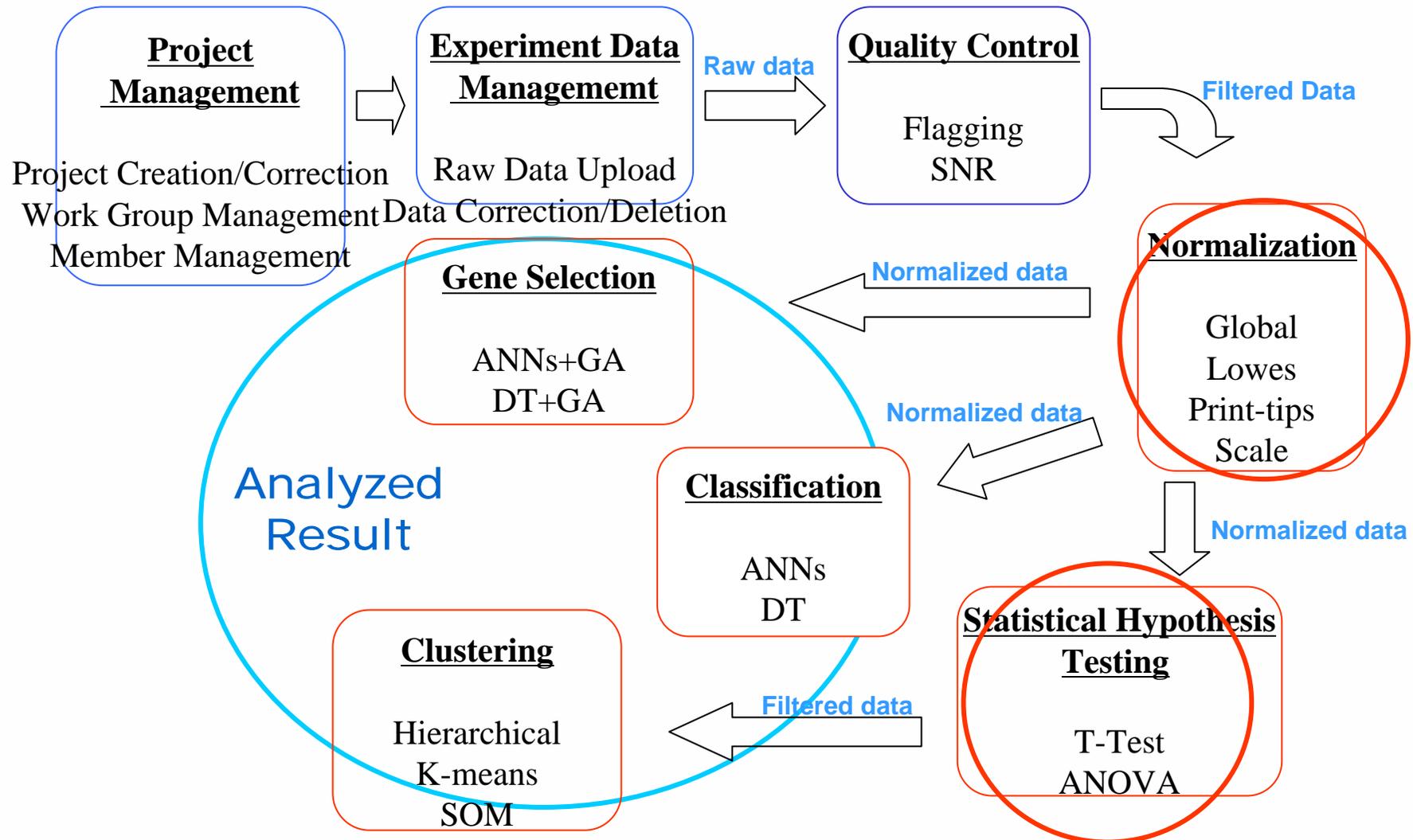
## Biochip Analysis Process



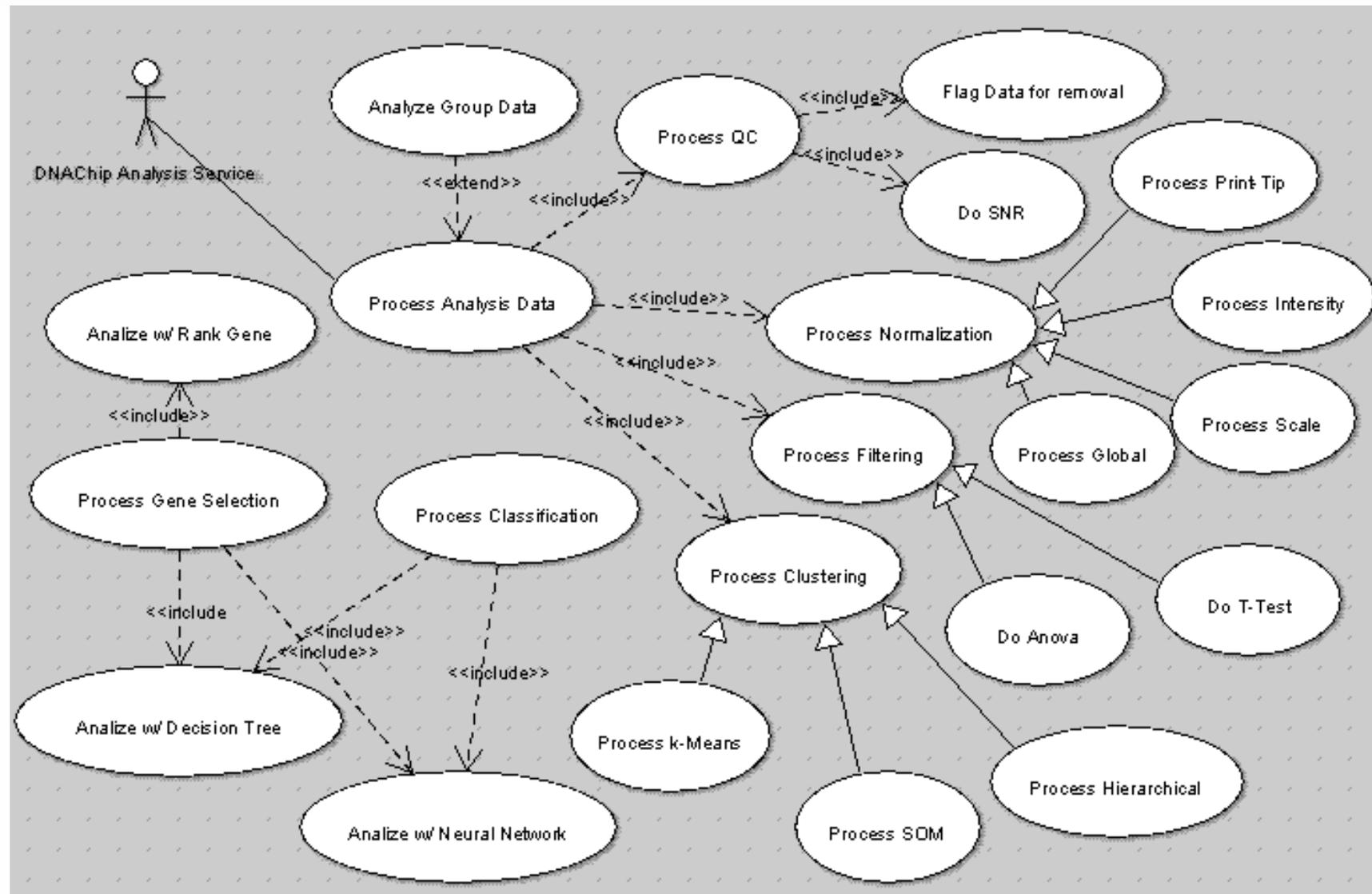
# 2-3. Image Processing



# 2-4. Statistical Analysis



# 2-5. Use case diagram



# 3-1. Summary

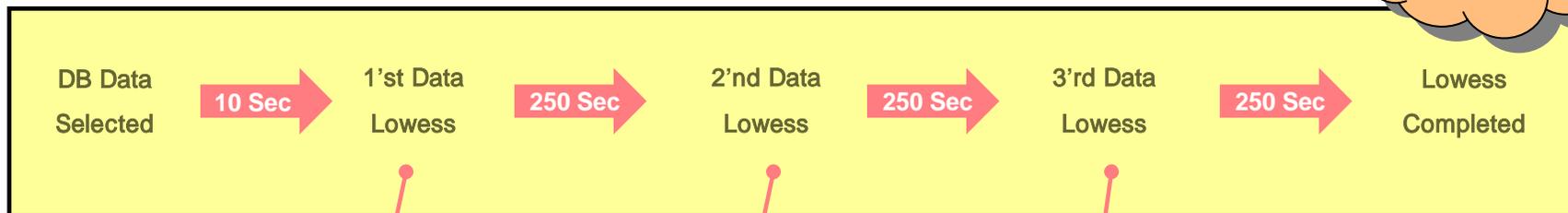
- **Lowess Normalization & Distribution Adaptation of T-Test**
- **Distribution Borker Server, Distribution Labor PC Environment**
  - 3.2GHz CPU, 2G Memory
  - Windows 2000 OS
  - Distribution Broker Server 1, Distribution Labor PC 3
- **Lowess Algorithm**
  - Sample Count : Reference 1, Test 3
  - Row Count : 15644
  - Option : Delta 0.0  
Smoothing Parameter 0.66  
Iteration 10
- **T-Test Algorithm**
  - Number of Sample per Group: 3 (Total 6 samples)
  - Number of Sample per row : 33,012
  - The part of Distribution Adaptation: The most time-consuming Data Reading Part
- **The effect of Distribution Analysis Adaptation**
  - Long analysis time is required to perform the analysis of large volume Bio Data.
  - Able to reduce the analysis time by the ratio of the number of work node
  - More effective for repetition work
  - The knowledge of analysis process is required to adapt distribution Algorithm

# 3-2. Lowess Normalization

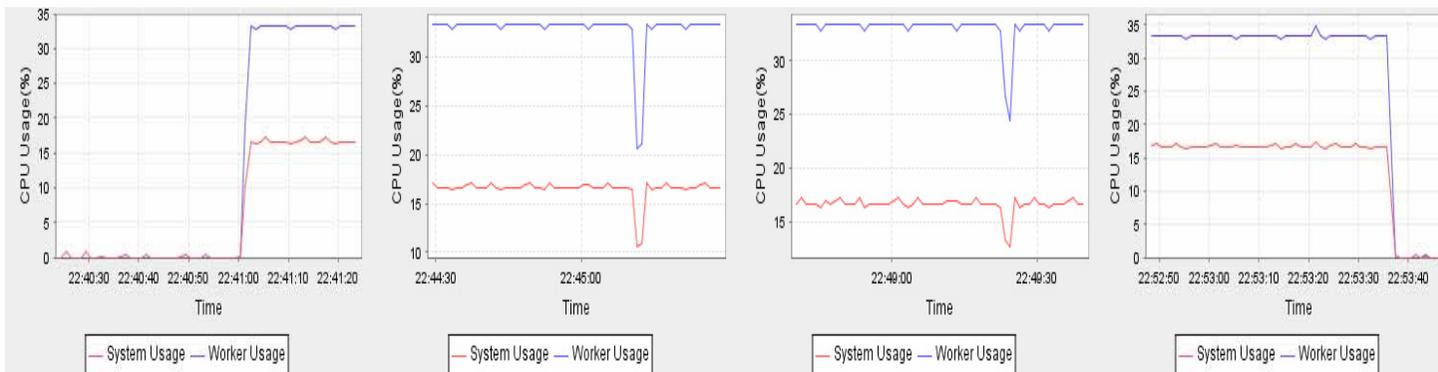
## The Adaptation of Distribution Analysis

- Distribution Adaptation of Lowess Normalization
  - Before Distribution Adaptation

Total  
760 Sec



Local PC  
CPU



# 3-2. Lowess Normalization

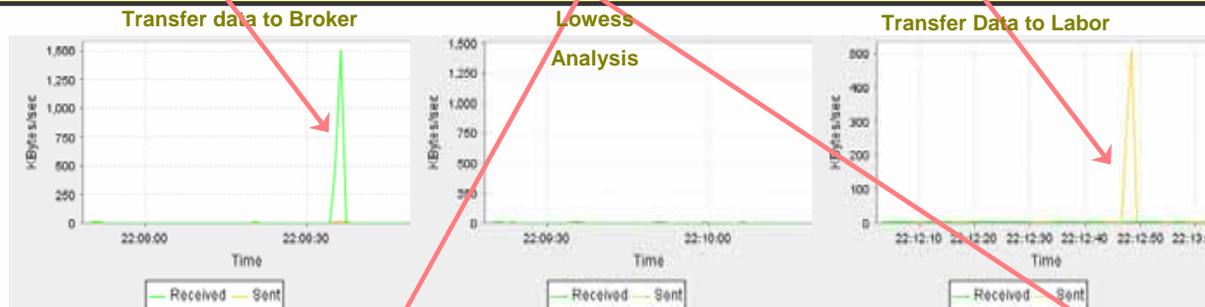
## The Adaptation of Distribution Analysis

- Distribution Adaptation of Lowess Normalization
  - After Distribution Adaptation

Total  
265 Sec



Network



Distribution Server  
CPU



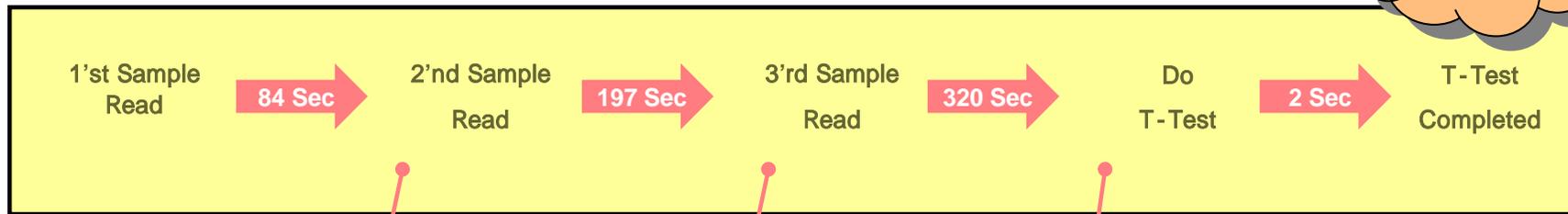
Use of 3 Nodes , Reduction  
of the analysis time by 1/3

# 3-3. T-Test

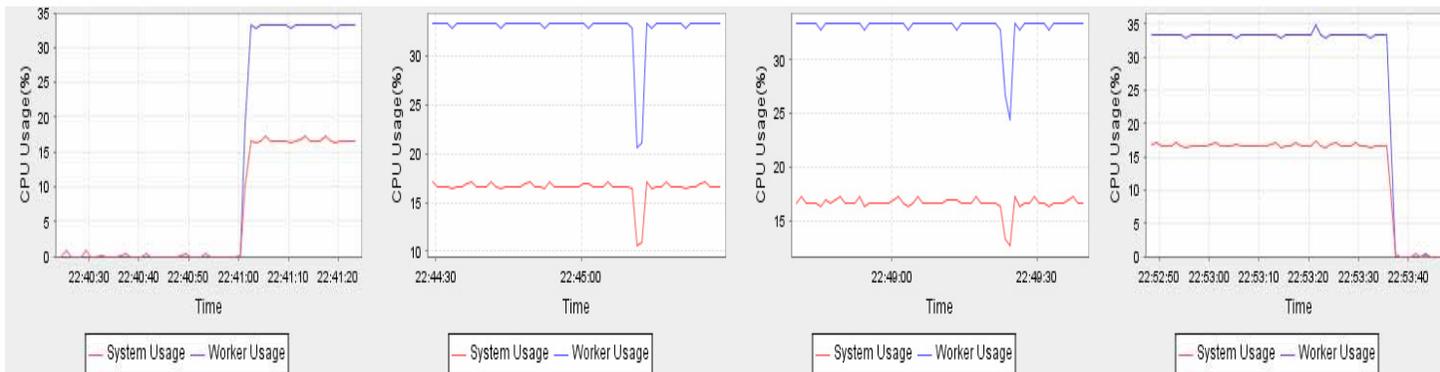
## The Adaptation of Distribution Analysis

- Distribution Adaptation of T-Test
  - Before Distribution Adaptation

Total  
603 Sec



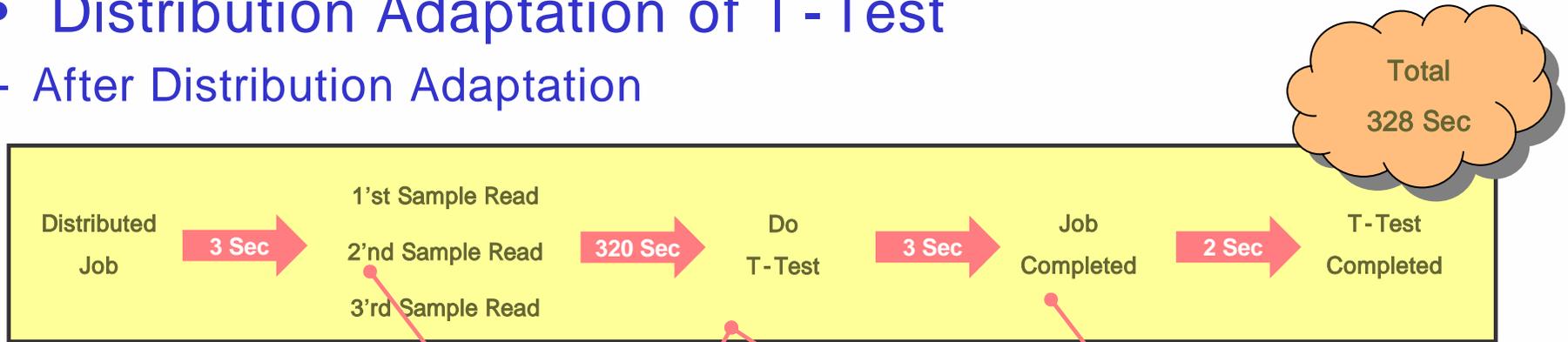
Local PC  
CPU



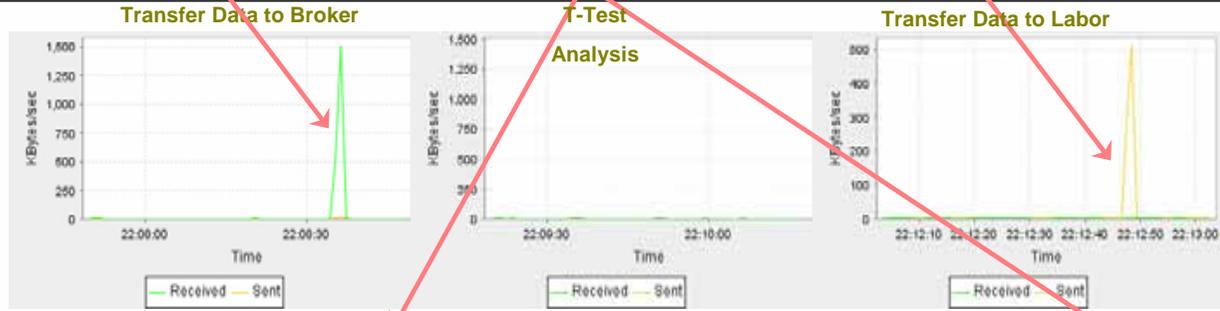
# 3-3. T-Test

## The Adaptation of Distribution Analysis

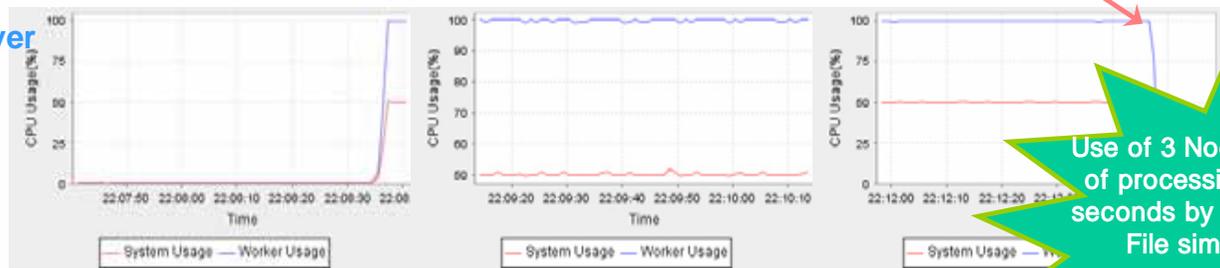
- Distribution Adaptation of T-Test
  - After Distribution Adaptation



Network



Distribution Server CPU



Use of 3 Nodes, Reduction of processing time to 275 seconds by reading Sample File simultaneously

## 4. Contributor

- Grid R&D Center, NationalGrid Inc.
  - Jong Hyun Youn
- Bioinformatics Research Center, Macrogen Inc.
  - Kyong Oh Yoon
  - Dae Hyun Kim
  - Keun Il Lee