

A System Architecture Assisting User Trial-and-Error Process in *in-silico* Drug Design

Takashi Maeno, Susumu Date
*Graduate School of Information Science and
Technology, Osaka University*

Yoshiyuki Kido
Mitsui Knowledge Industry Co.,Ltd.

Ichiro Hasegawa
NEC System Technologies, Ltd.

Shinji Shimojo
Cybermedia Center, Osaka University

Outline

1. Background

- In-silico Screening
- DOCK as a tool for in-silico screening
- Related research of in-silico screening

2. Problem definition

- Our research target
- Problem of our research

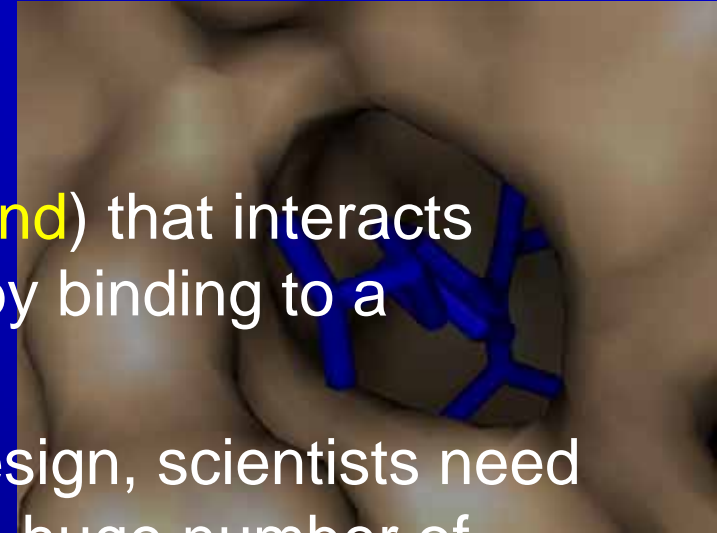
3. Design

- Modeling scientists' trial-and-error processes
- System image
- System Architecture
- Implementation overview

4. Conclusion and Future Works

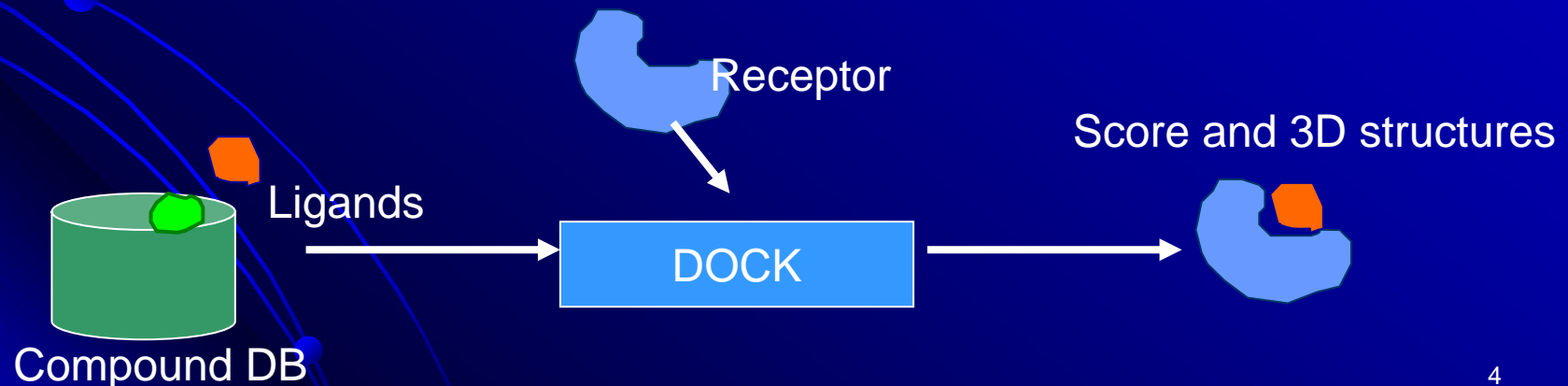
In-silico Screening

- Drug design and screening
 - A drug is a small compound (**ligand**) that interacts with a certain protein (**receptor**) by binding to a certain site of the protein
 - In the beginning stage of drug design, scientists need to screen drug candidates from a huge number of ligands (screening)
- In-silico screening attracts scientists
 - To reduce cost and time, scientists utilize computing resource for screening
- Scientists simulate receptor-ligand docking and get a score as a criterion for screening



DOCK

- DOCK is a tool for Receptor-Ligand Docking
 - Input: files of ligand and receptor
 - Output: score and 3D structure of docked ligand
- Scientists can screen ligands based on a score from DOCK



Current Researches of in-silico Screening

High throughput screening

- Each docking process can be executed independently.
- Many studies on Parallel processing of in-silico screening have been reported.
- e.g. [Buyya03] reports high throughput screening using DOCK and Nimrod/G

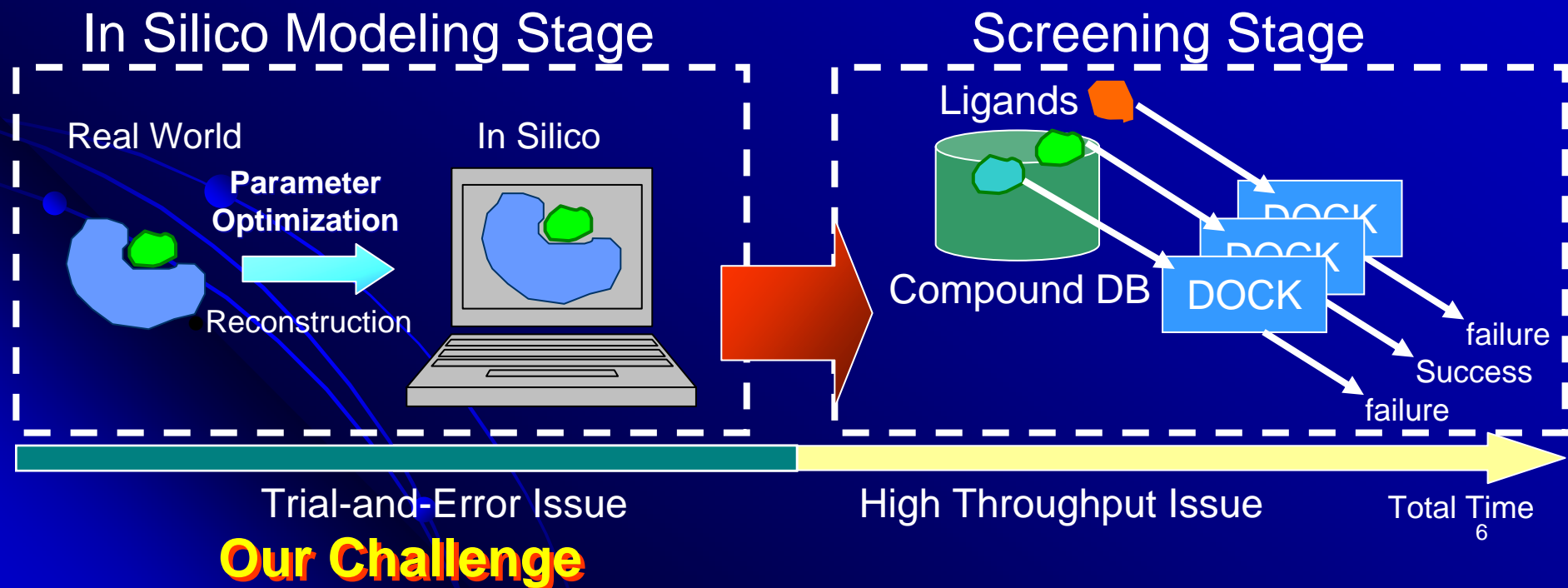
[Buyya03] R.Buyya, K.Branson, J.Giddy, and D.Abramson

“Virtual laboratory: Enabling molecular modeling for drug design on the world wide grid”

Our Research Target

In-silico modeling stage

- Parameter optimization for accurate in-silico modeling before screening



In-silico modeling stage

receptor
ligand

X-ray crystal structure of complex
(receptor with ligand binding it)
from a laboratory experiment

In-silico (in computers)

Scientist checks if the
parameter optimization is
proper or not, comparing
laboratory experiment result
and in-silico modeling

DOCK Suite Flow

For deciding position of ligand's atoms

atoms

Receptor

Scientists need to optimize all parameters properly for all tools

- They need to consider which tools to be combine and which parameters to be optimized
- They consider next coordination from the results of former coordination
- They need to consider which results to be used as criteria for parameter optimization
- If the coordination does not go well, they may have to coordinate former tools again (they also have to consider dependency between tools)

Mapping chemical properties on grid points

Complexity of Trial-and-Error Processes

- Three types of entities change for each consideration process
 - Which tools to use and how they combine these tools
 - How they coordinate parameters for each tool
 - Which results to be gathered and how they use the results as criteria of coordination process.
- The problem is these three types of entities differs in scientists knowledge and experience

necessary to assist trail-and-error processes

“Trial Set” Concept

To assisting scientists' trial-and-error processes, we propose unification of scientists' trial-and-error procedure as a trial set

1. Represents tools and their connection as a workflow template
2. Adding variation on workflow template represents a parameter optimization way

→ We aim a modeling of trial-and-error processes

System Image with Trial Set

Scientist A

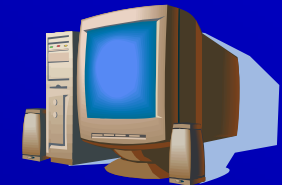


Creates and executes
A's trial set



Trial set of A

Obtains A's trial set



Scientist B

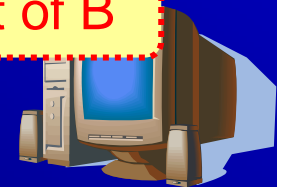


Modifies A's trial set and
executes it



Trial set of B

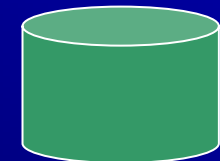
Obtains B's trial set



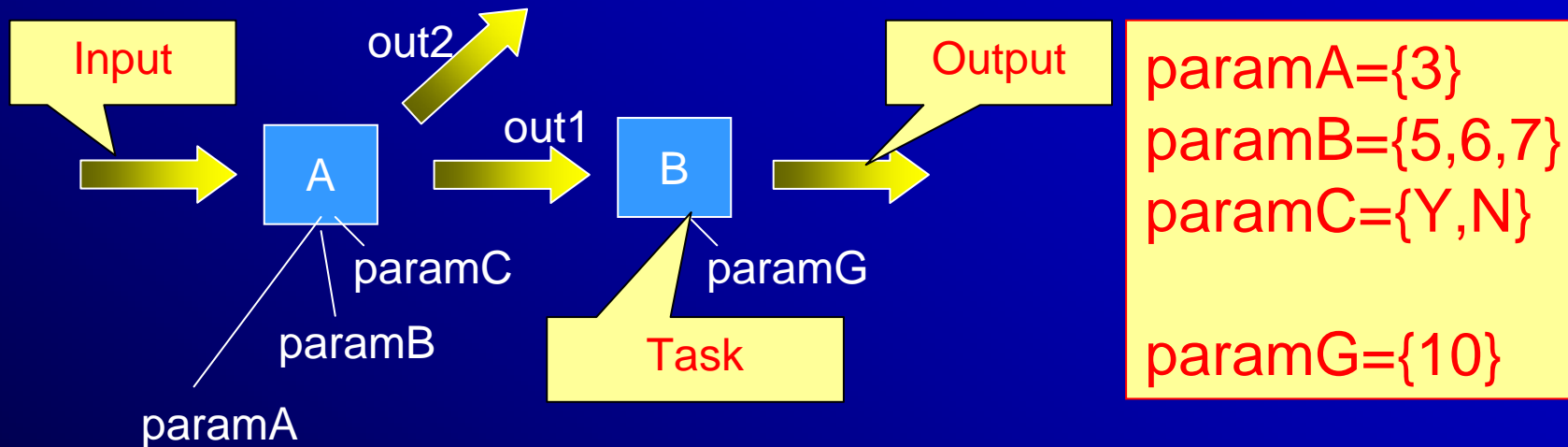
Scientist C



Executes B's trial set



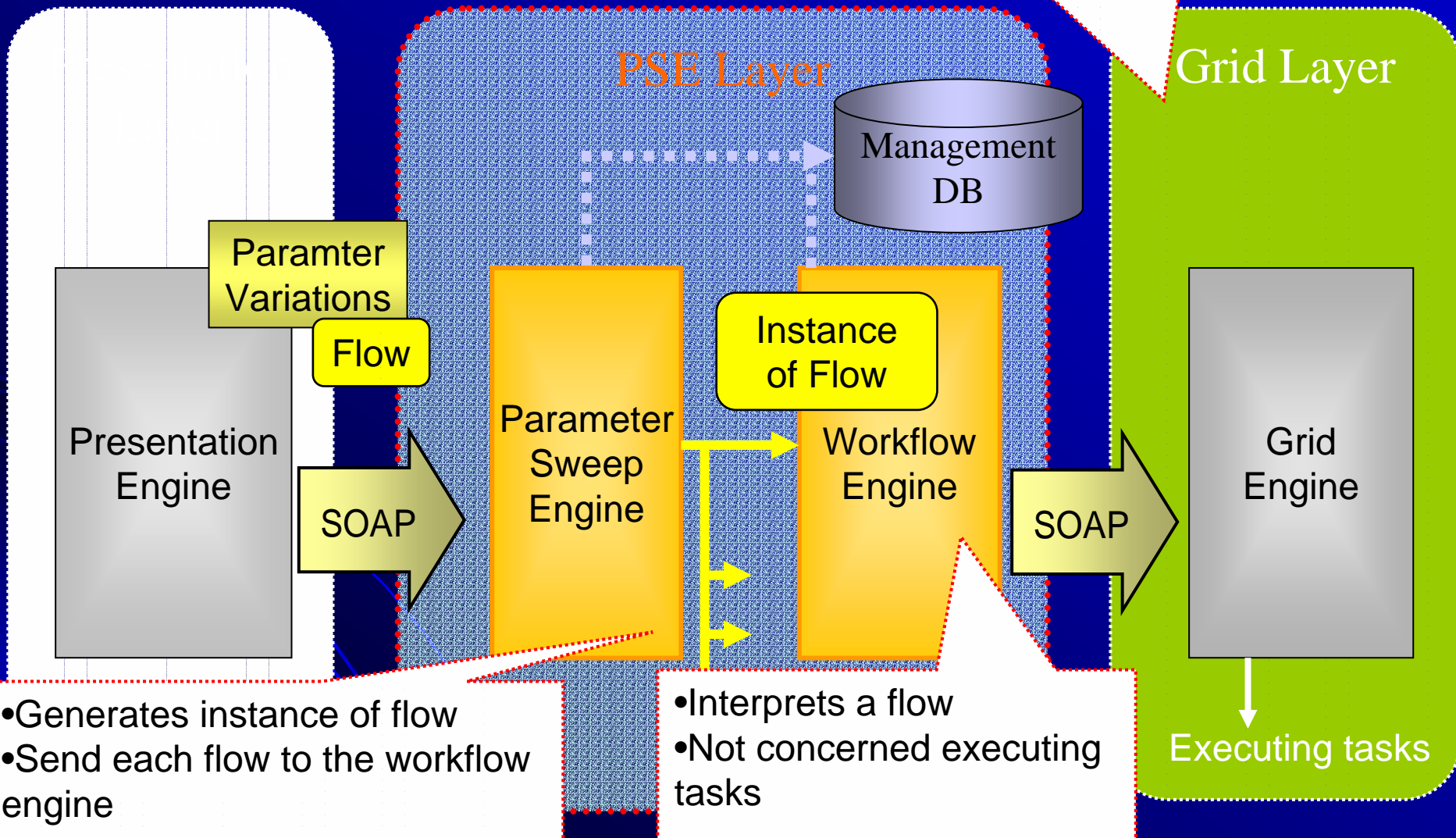
Example of Trial Set



	Task A			Task B
	paramA	paramB	paramC	paramG
P1	3	5	Y	10
P2	3	5	N	10
P3	3	6	Y	10
P4	3	6	N	10
P5	3	7	Y	10
P6	3	7	N	10

System Architecture for Executing Trial Set

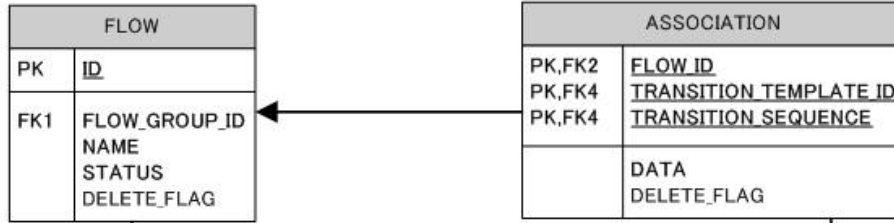
- Executes each task
- It doesn't know the flow



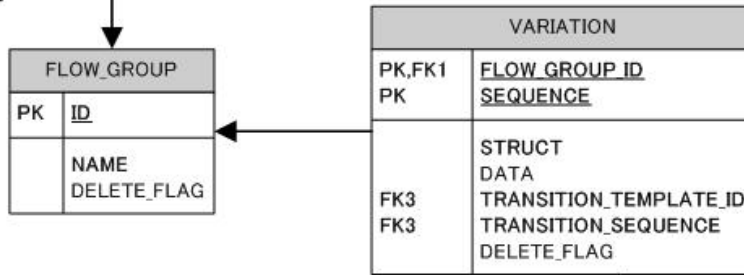
3.Design

Management DB

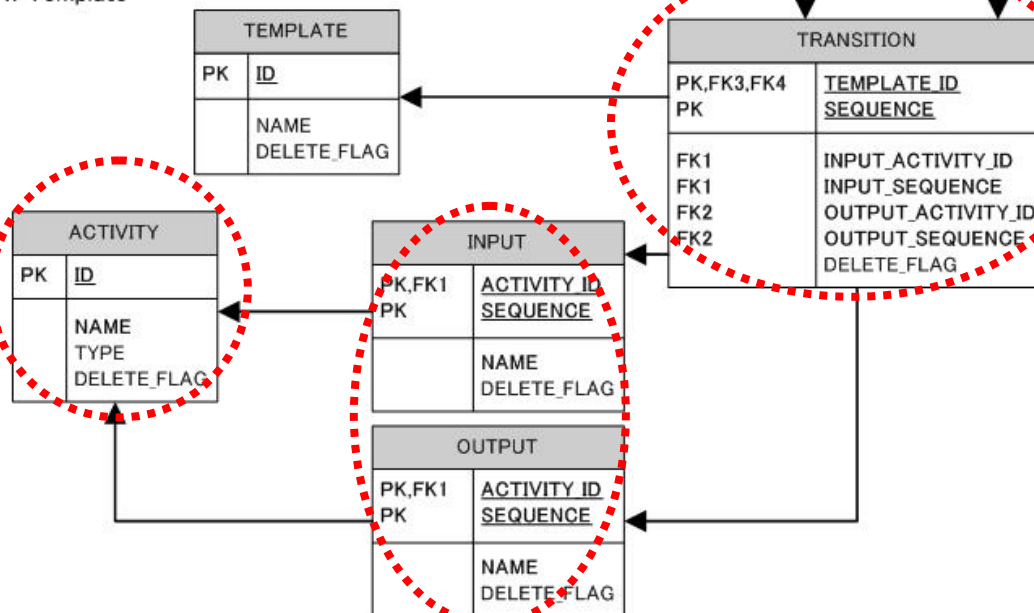
Flow Instance



Flow Instance Group



Flow Template



Trial set

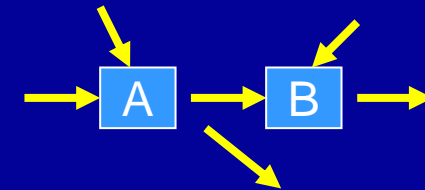
Variation

Yes, No

5,6,7

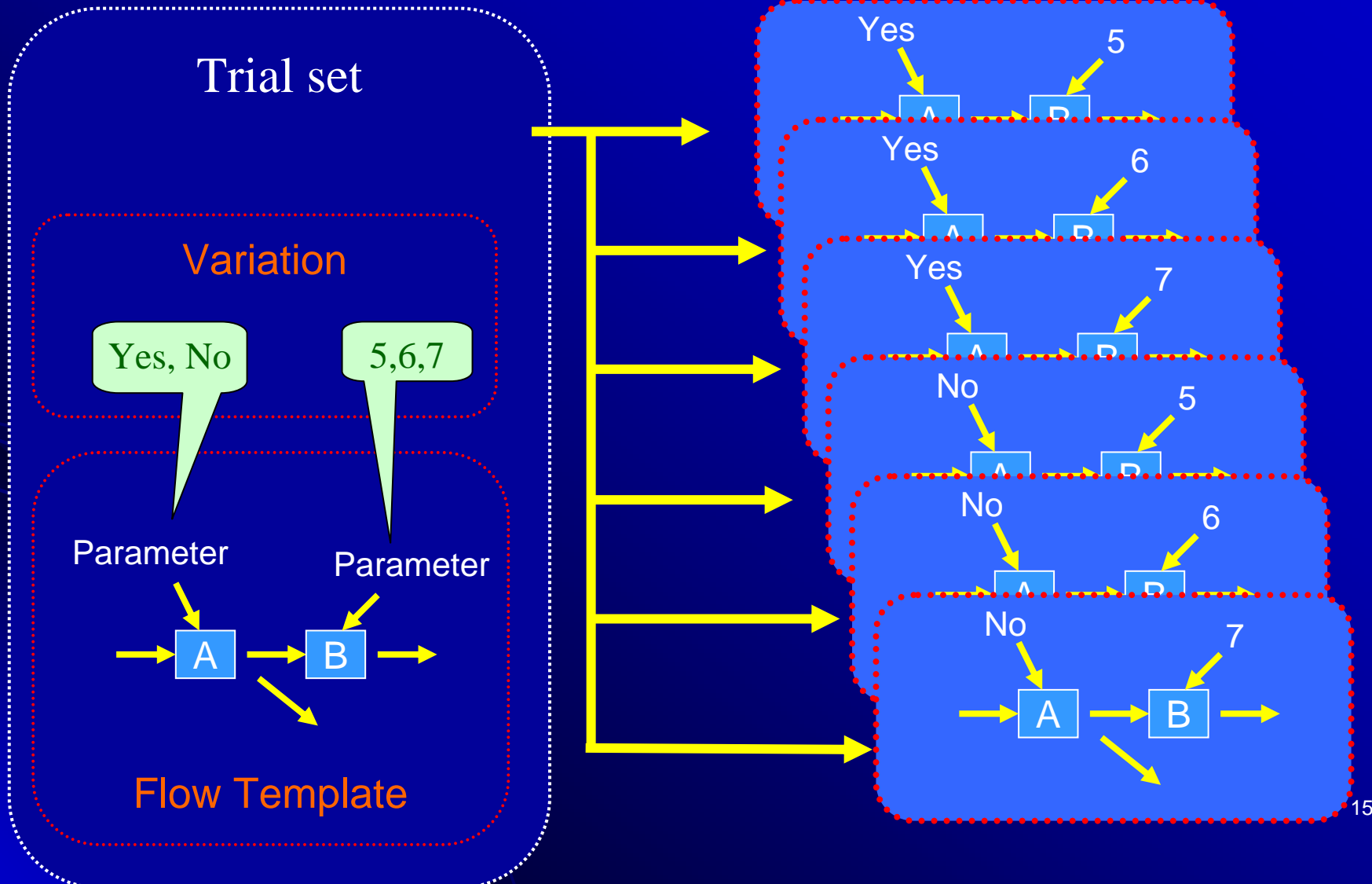
Parameter

Parameter

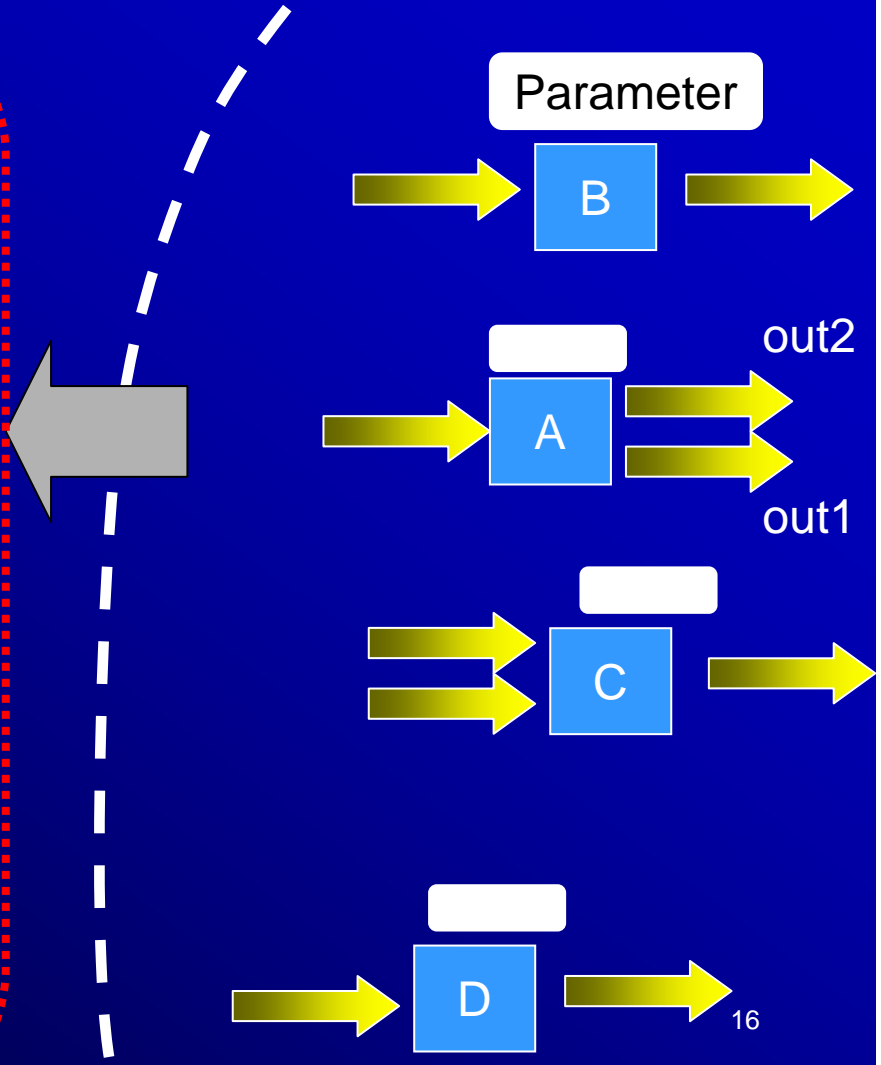
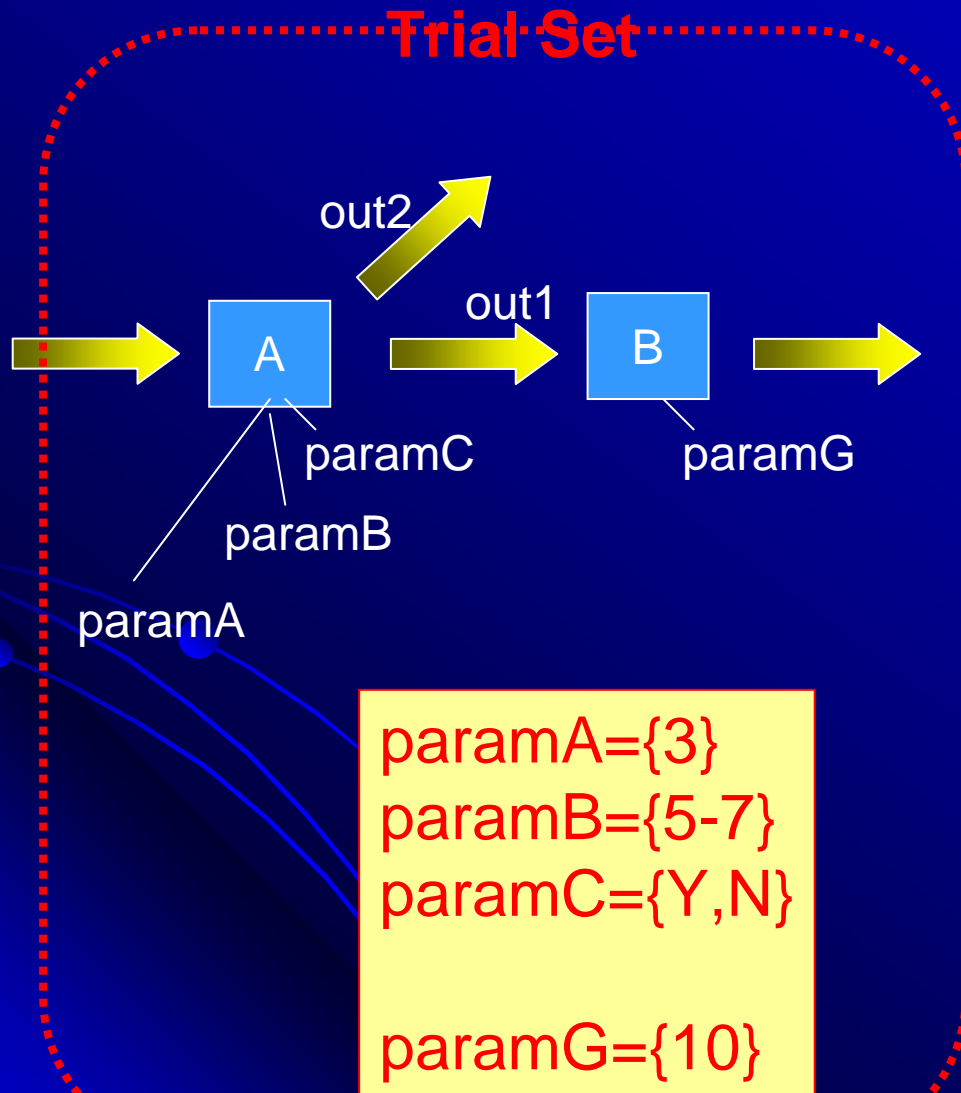


Flow Template

Execution of Trial Set

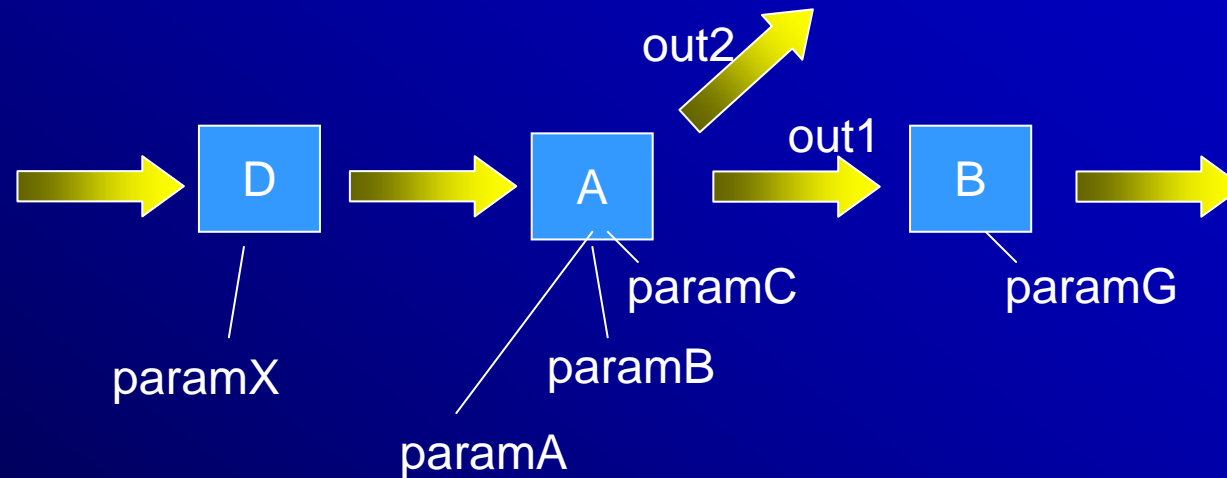


Creating Trial Set



Modifying Trial Set

~~Trial Set~~

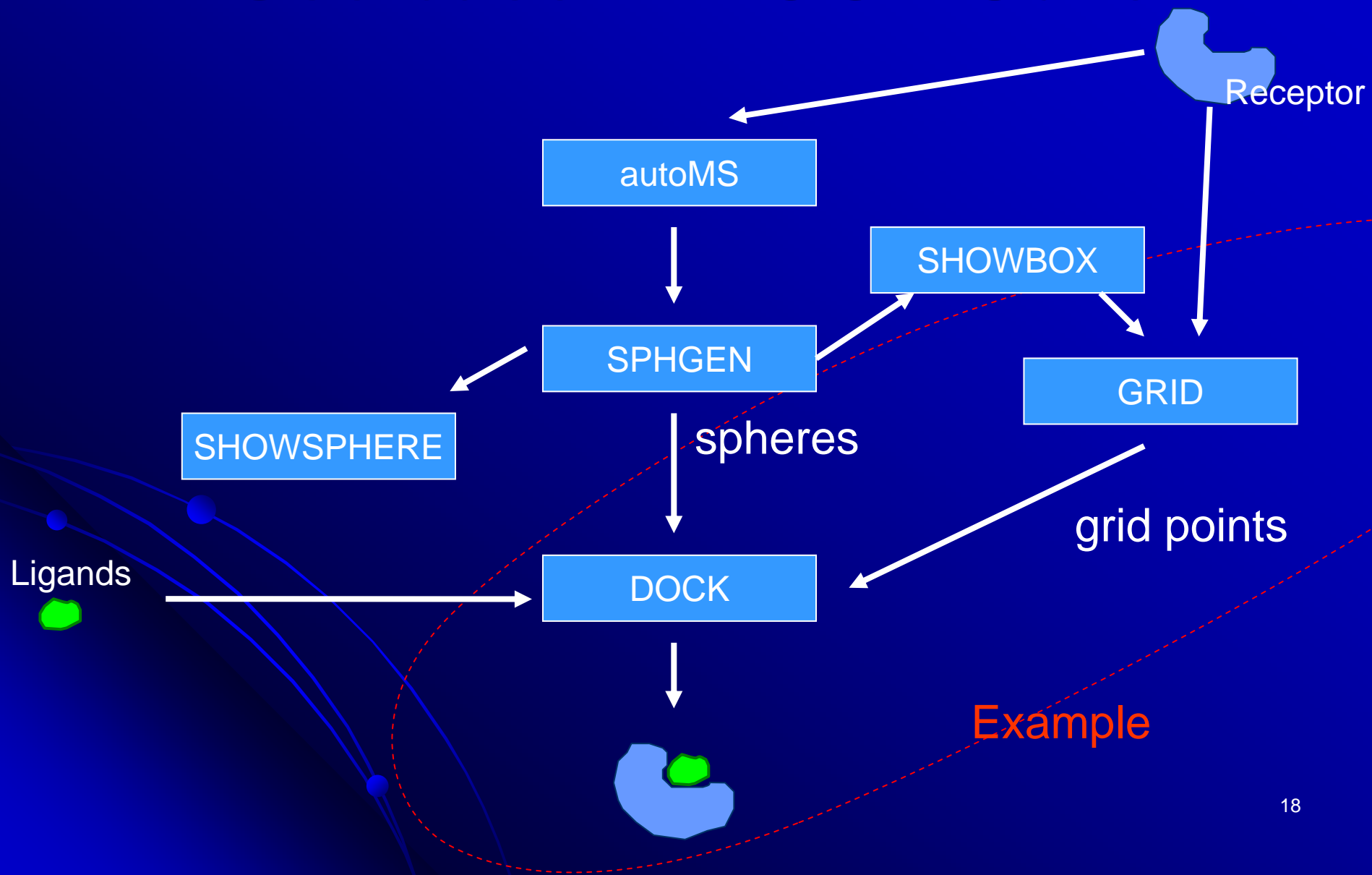


paramX={0.5, 0.25, 0.05}

paramA={3}
paramB={6}
paramC={Y,N}

paramG={10}

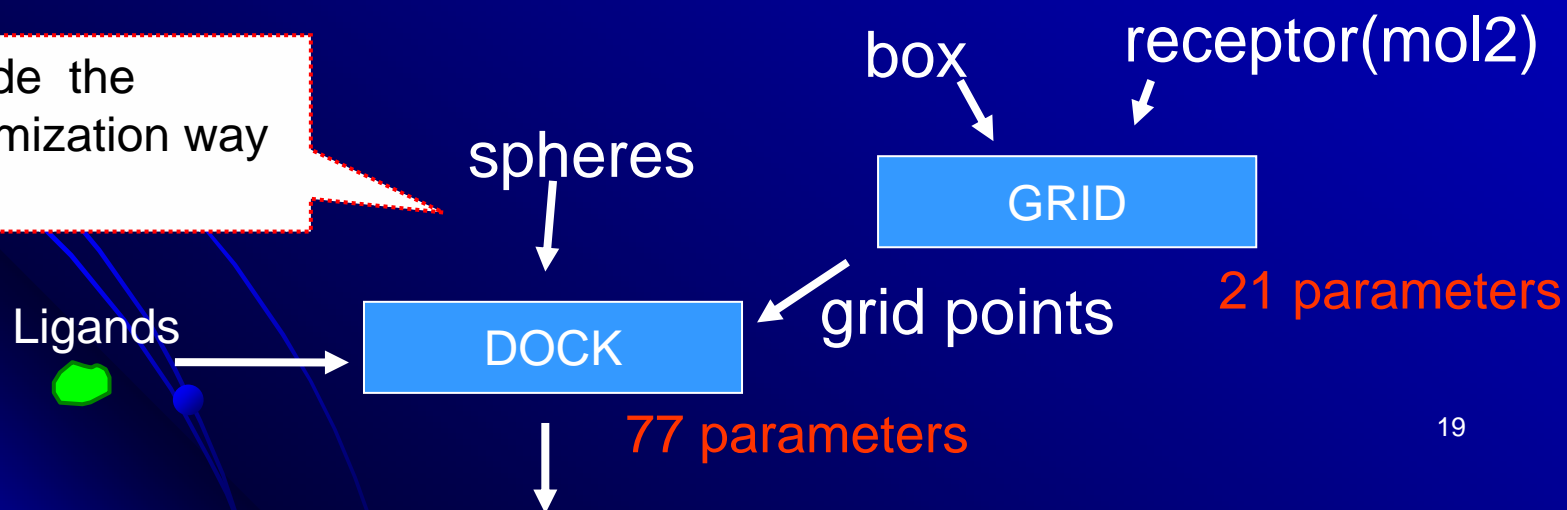
Use case in DOCK Suite



Example: DOCK Stage

- Scientists need to optimize DOCK's parameters (scoring function, ligand behavior in docking, etc)
- If the DOCK's parameter optimization does not go well, scientists try former procedure again such as GRID parameter optimization

Scientists decide the parameter optimization way again



Conclusions and Future Works

- We proposed,
 - modeling scientists' trial-and-error processes as a trial set
 - A system architecture to execute trial sets
- Future works
 - More efficient mechanism to support for scientists' trial-and-error processes
 - Sharing trial set among scientists
 - Reducing the analytic space of trial set
e.g. cutting the excessive flow dynamically