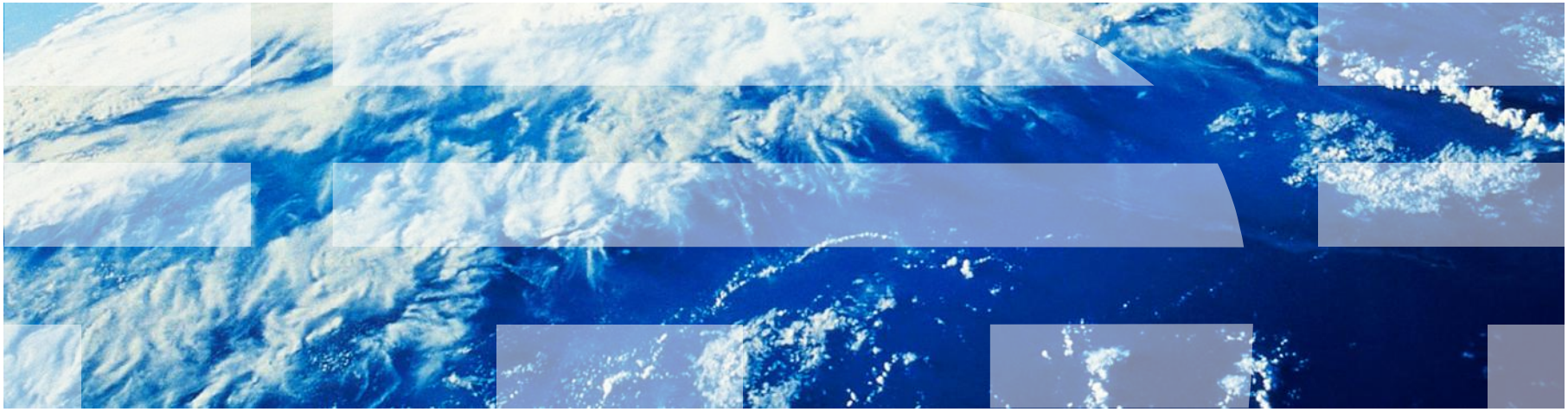
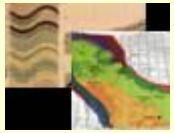
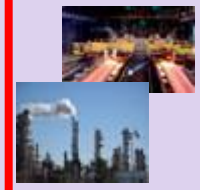
Hisashi Miyashita, Yoichi Hatsutori, Junya Shimizu
IBM Research - Tokyo

Model-based System Identification Cloud (MbSIC) for the integration of Engineering and Operation



Overview of Mining Industry

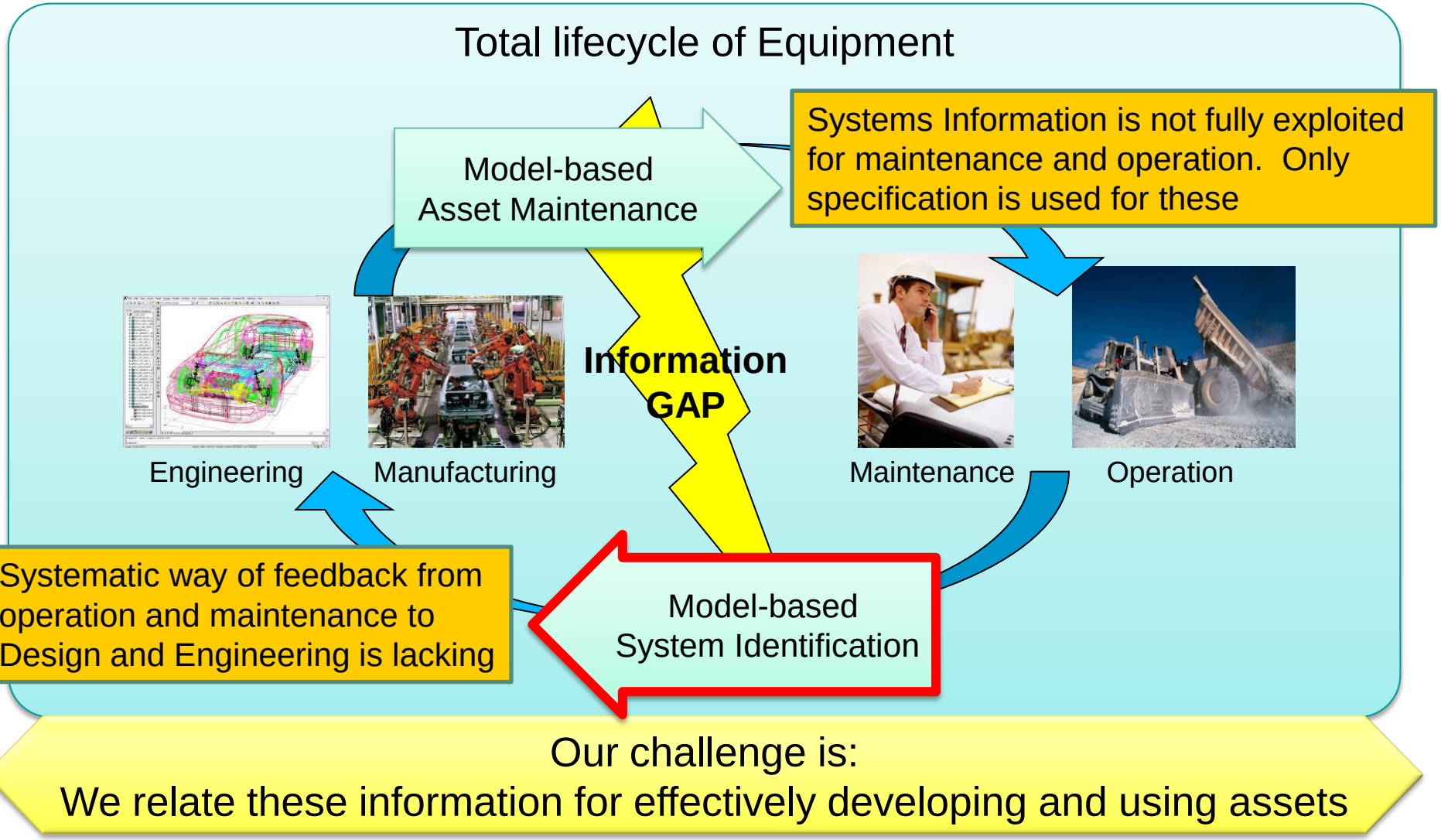
Plan for Supply	Exploration		Production		Process / Refine	Secondary Distribution
Understand reserves, policy, ... 	Prospecting 	Physical testing and assessment 	Extraction 	Primary distribution 	Refinery, smelter, ... 	Trucking, retail, .. 
Mining	**1% of cost**		**20% infrastructure set-up** **80% extraction and primary processing**			

Capital Investment in Mining is huge
- Total investment in WW is about $110B in 2011
- CODELCO announced $16.3B capital investment between 2011-2015

Effective use of such invested asset is critical

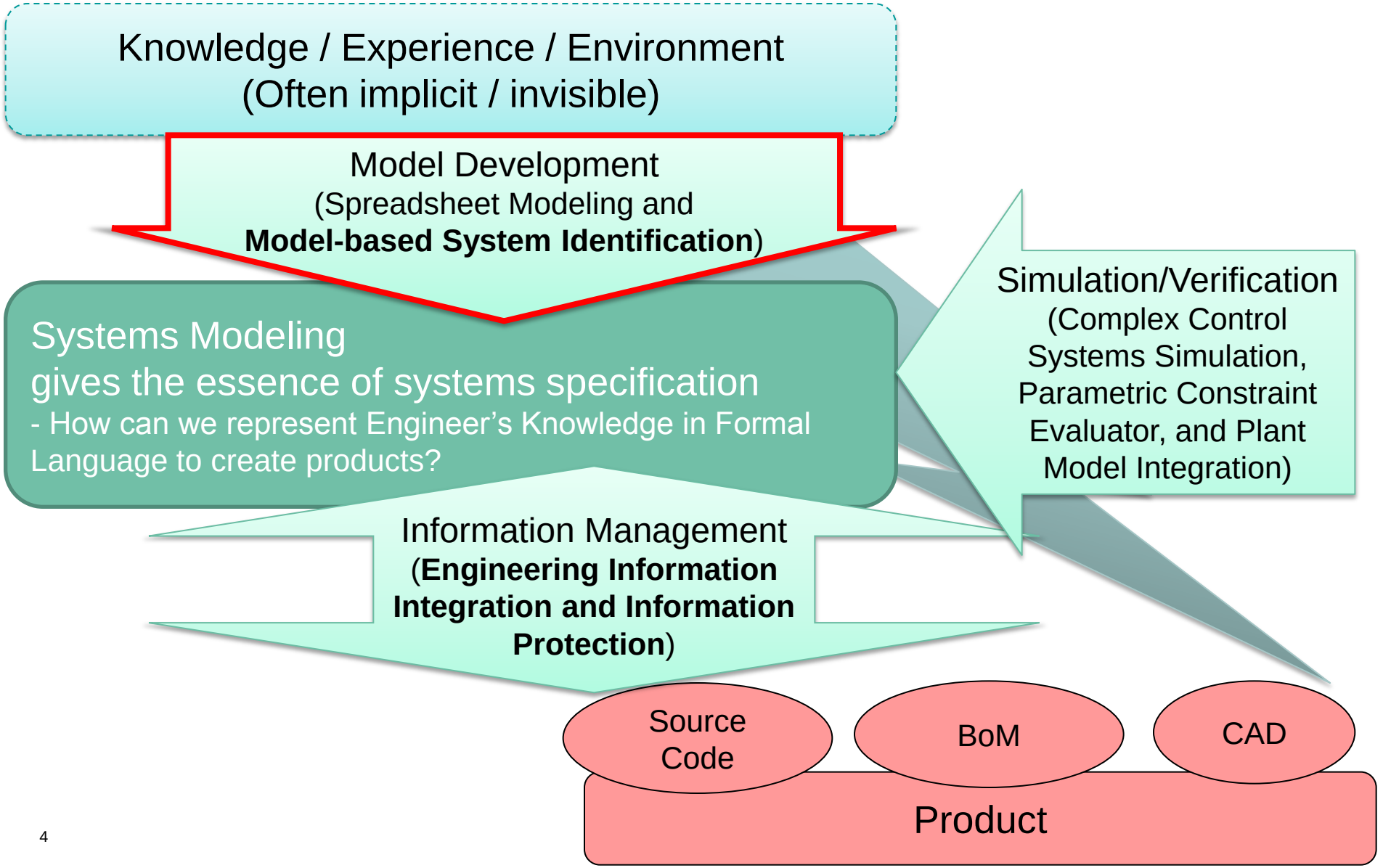
Challenges of Computational Engineering for Mining Industry

Our challenge is how we can leverage Engineering information for maintenance and operation and vice-versa



Our challenge is:
We relate these information for effectively developing and using assets

What are our values to bridge the gaps?

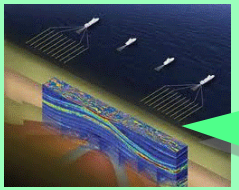


Model-based System Identification Cloud (MbSIC)

Since data analytics in natural resource industry is becoming enormous and complex, developing analysis system is getting hard. MbSIC addresses these issues by leveraging Model-driven Systems Engineering and Cloud Computing.

Issues & Challenges in System Identification

- **Enormous Amount of Data**

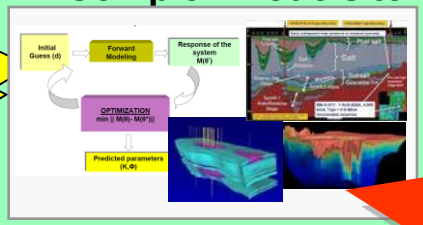


Seismic Analysis requires 100MB/sec to 0.5PB Tape

performance vs. flexibility

Optimized Code for Specific Models → General-purpose program for flexibility

- **Complex Models to be considered**



- Hard to develop it
- Difficult collaboration among many experts!

Solution

Model-based System Identification Cloud

massive number of observations

Map-Reduce based Cloud computing

Optimized code from Model

$$N=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1i} \\ & \ddots & & \\ a_{j1} & a_{12} & \dots & a_{ji} \end{pmatrix}$$

Code Generation

Efficiently identify systems by filling parameters from many observations

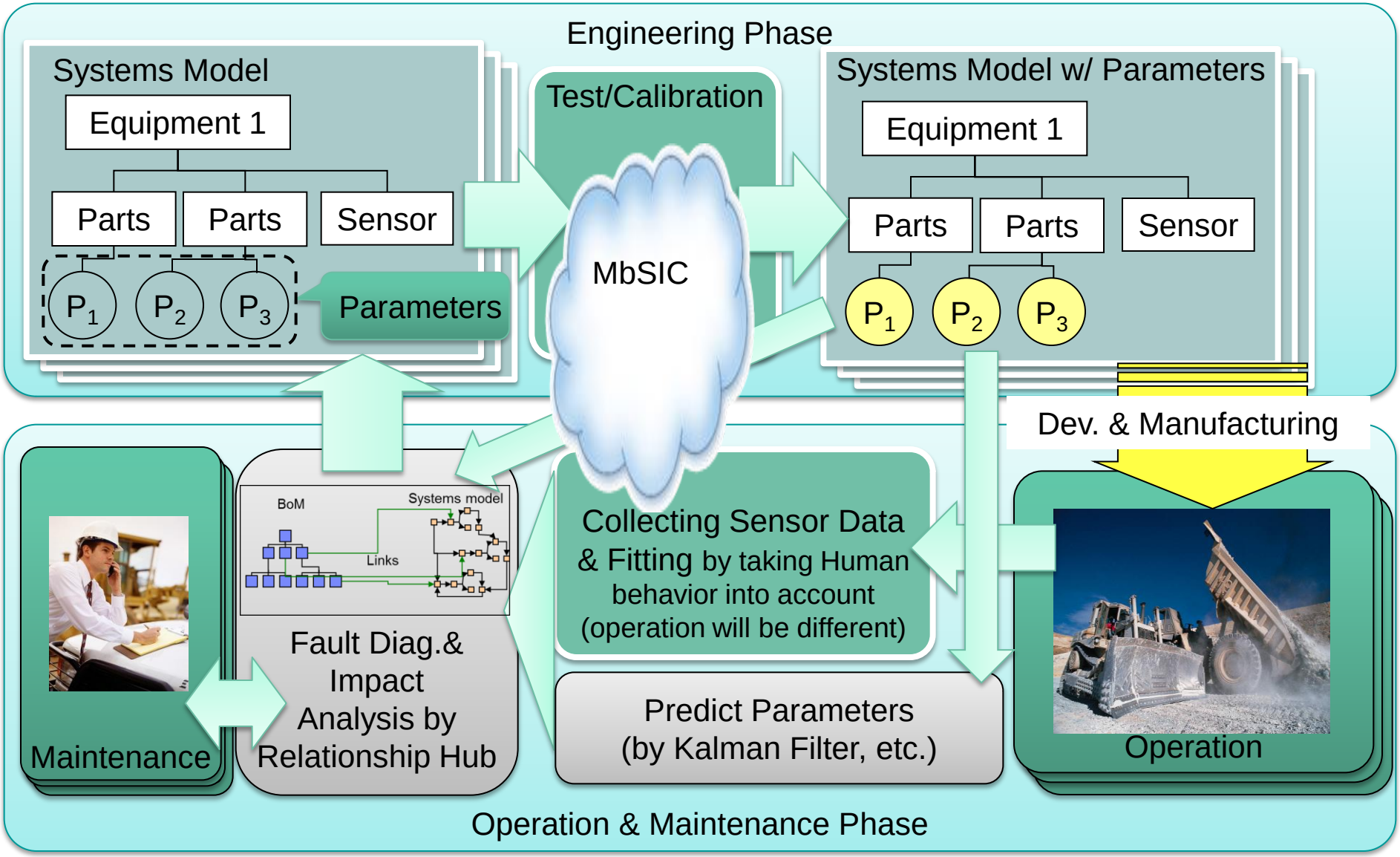
Scientists

Communication based on Systems Model and Mathematical Expressions

Engineers

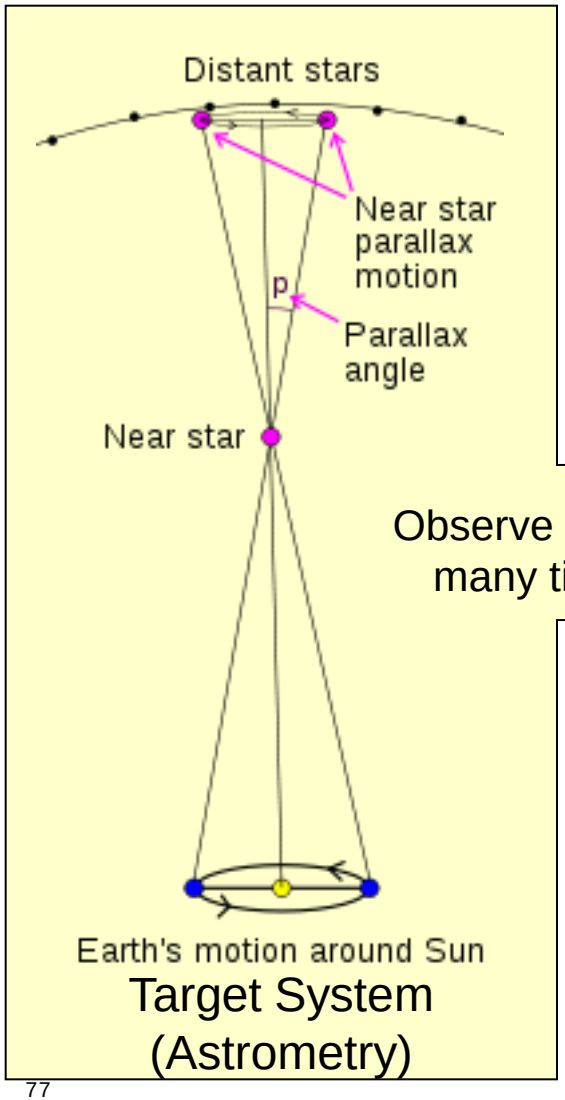
Systems and Geological Model development

Vision of Integration of Engineering and Operation by MbSIC



Target Problem by Example

JASMINE (Japan Astrometry Mission INfra-red Exploration) Project is a satellite-based high precision (1mas) astrometry project. MbSIC is challenging this data analysis.



Observe angles many times

Sensor noise, Fluctuations, Heat Effect, etc.

Enormous Amount of Experimental Data (TB to PB class)

Model-based System Identification Cloud

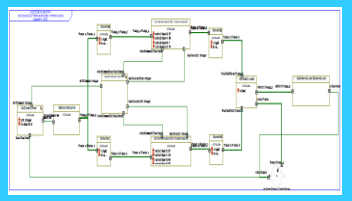
Nonlinear Least-square Method

Map of Gaia (Milky Way)

Satellite Plant Model with Estimated Parameters

Analysis Results

Systems Model of Stars, Sun System, Satellite
They need to fill many parameters

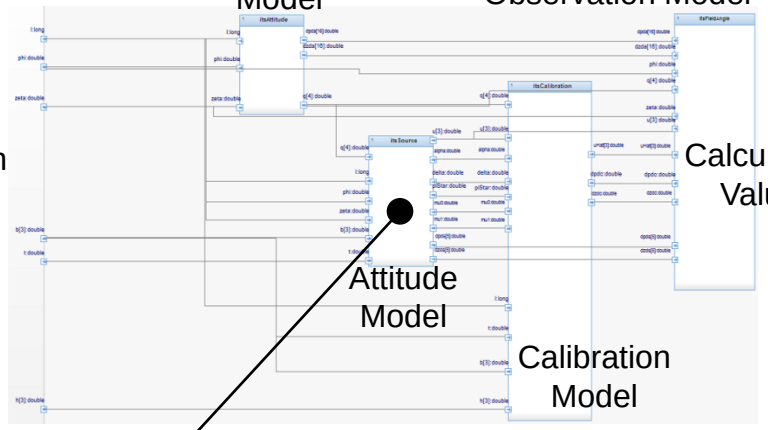


Quick look at Systems Model for JASMINE

Source Model

Observation Model

Observation
- Quaternion
- Position
- Velocity
- Time



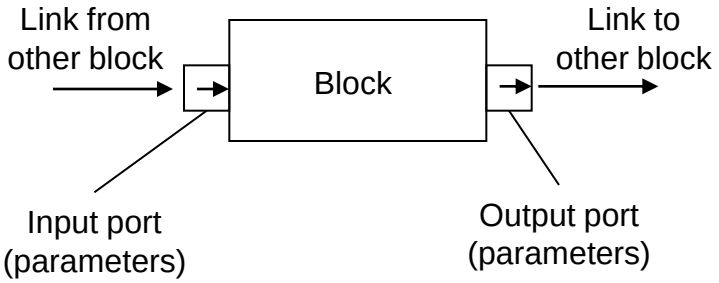
Calculated Value

Attitude Model

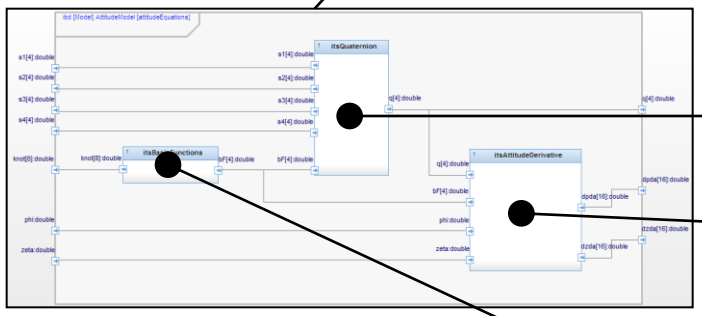
Calibration Model

Internal Block Diagram (IBD)

Each block defined in BDD is related to each other
Internal signal flows are defined as ports and links.



Attitude Model



$$\mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = \left\langle \sum_{n=l-M+1}^{l} \mathbf{a}_n B_n \right\rangle$$

$$-\frac{\partial f}{\partial \mathbf{q}} = -2\sec\zeta_l \mathbf{q}_l \left\langle \mathbf{n}_l, 0 \right\rangle$$

Inputs

Internal Signal flow

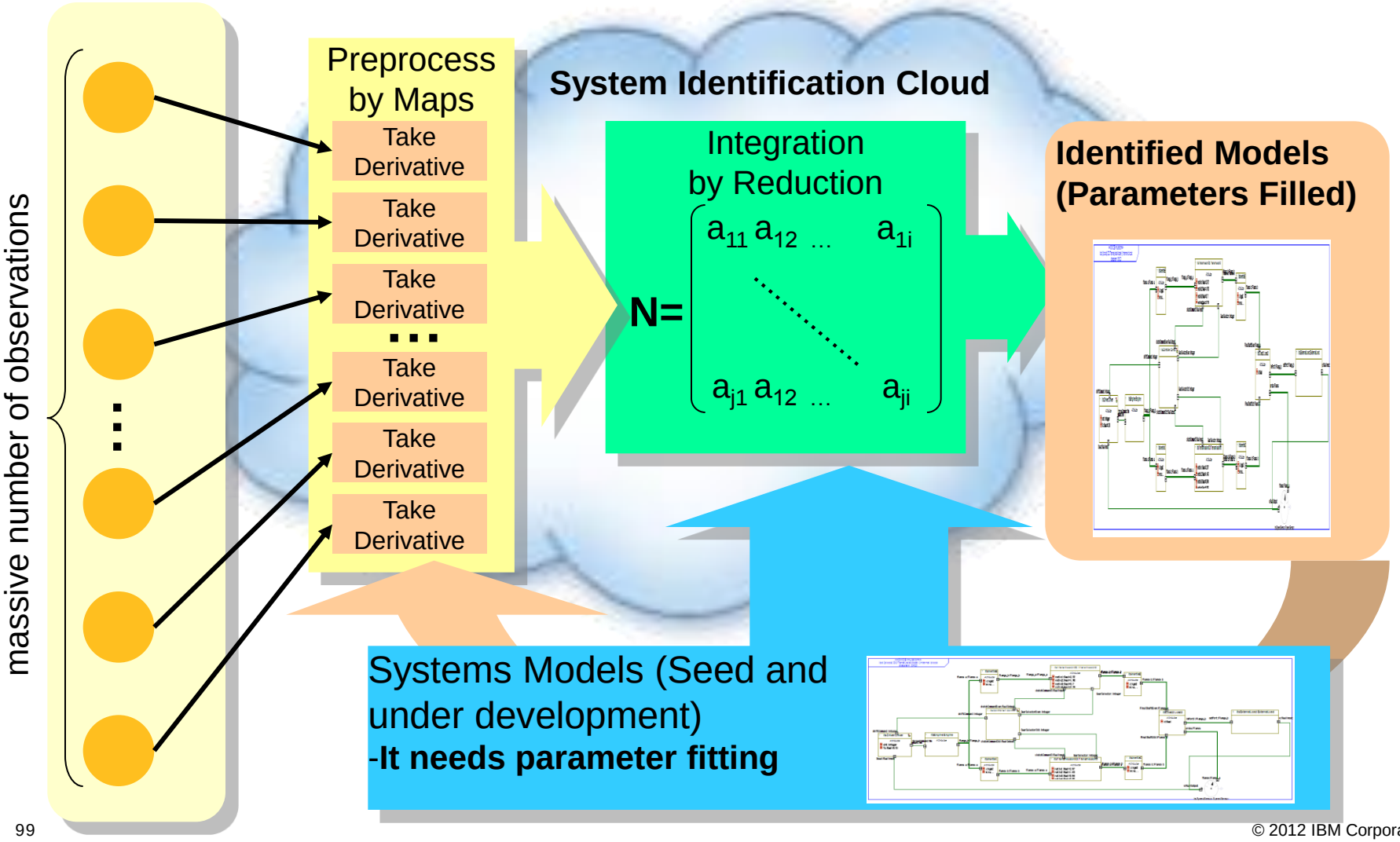
Outputs

$$B_{j,n} = \frac{t - t_j}{t_{j+n} - t_j} B_{j,n-1} + \frac{t_{j+n+1} - t}{t_{j+n+1} - t_{j+1}} B_{j+1,n-1}$$

$$B_{j,0} = \begin{cases} 1 & if\ t_j \le t < t_{j+1} \\ 0 & otherwise \end{cases}$$

The mechanism of MbSIC

MbSIC leverages MapReduce framework to process massive number of observations. It computes derivatives in Map Phase and then construct a matrix in reduce phase. And then solve that matrix in GPGPU computing, and repeats these until convergence

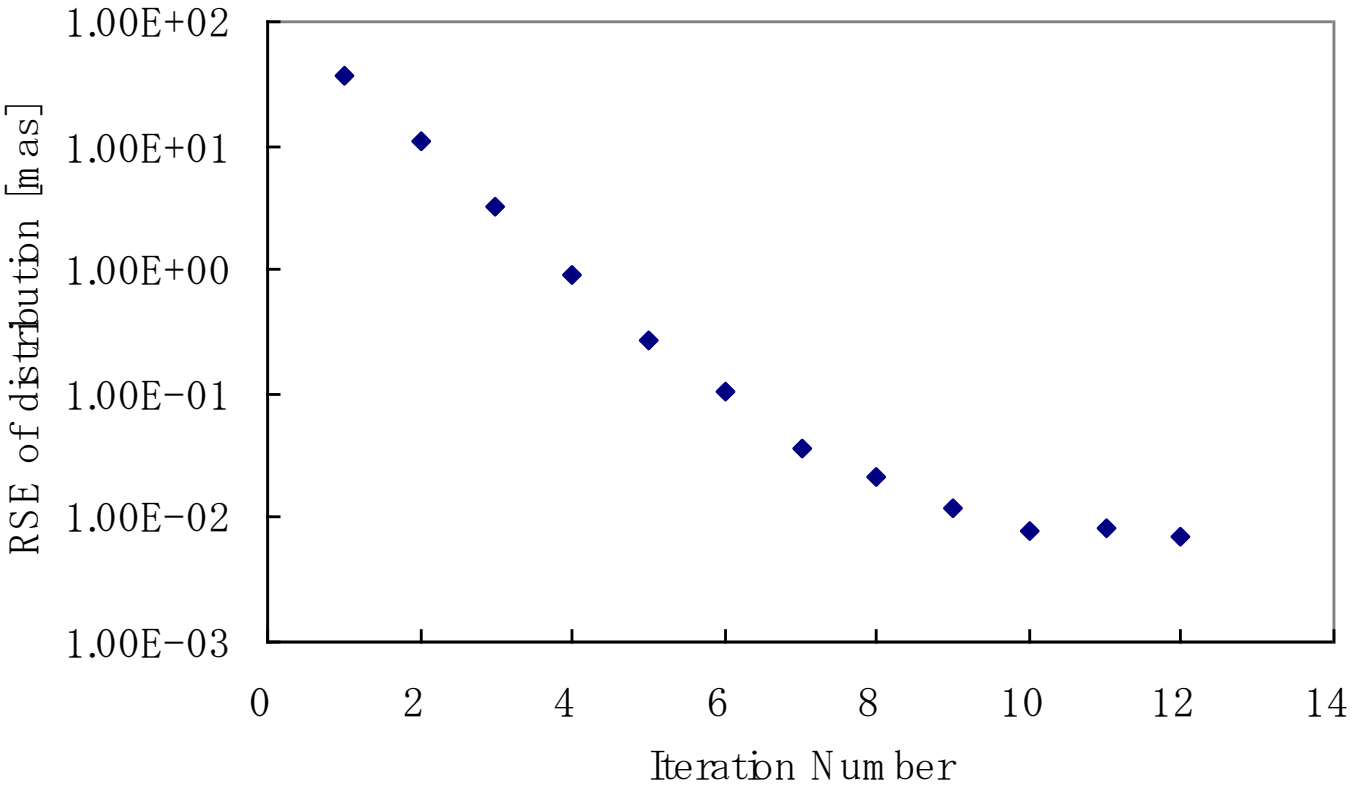


Current Status on Satellite Astrometry and Remaining Work

The input data were generated by using a simple model, that is:

The satellite is assumed as a single rigid body. Effects of disturbances, an attitude control and any degradation of equipment weren't considered. Only white Gaussian noise was added to the quaternion. A gravitational light deflection wasn't included in a source model.

Evolution of the parallax update vs. the iteration number for 10000 objects
RSE: Robust Scatter Estimate



Concluding Remarks

Model based System Identification will open up new era of the integration between Engineering and Operation

- Information integration between Engineering and Maintenance will enable
  - By exploiting the systems model information, we can improve maintenance, and then give proper feedback to engineering
- MbSIC is scalable, cloud-based technology for System Identification to capture huge amount of observations from sensors
  - It compiles systems model, and then execute least square method on top of Hadoop and GPGPU based cloud technology.
  - Currently, we could apply it to the astrometry, but we would like to apply it to mining industry
- Next Step:
  - We need to build target equipment and operation systems model
    - What are the difference and similarity between astrometry and equipment?
  - Support Recursive function (e.g. for Kalman Filter)
  - Hybrid System support for Transient (or Human) behavior