

Current Status and Issues of Visualization Research for Large-Scale Dataset

Kenji Ono RIKEN

AICS Cafe 6 July 2012

Introduction

— [大分県日田市 (1966)

— [日田高校 (1984卒業)

— [熊本大学工学部・大学院 (1984-1990)

— [日産自動車 中央研究所 (1990-2001)

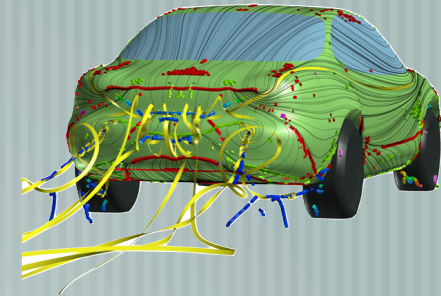
— [東京大学大学院工学系研究科 機械工学専攻 (2001-2004)

— [理化学研究所 (2004-)

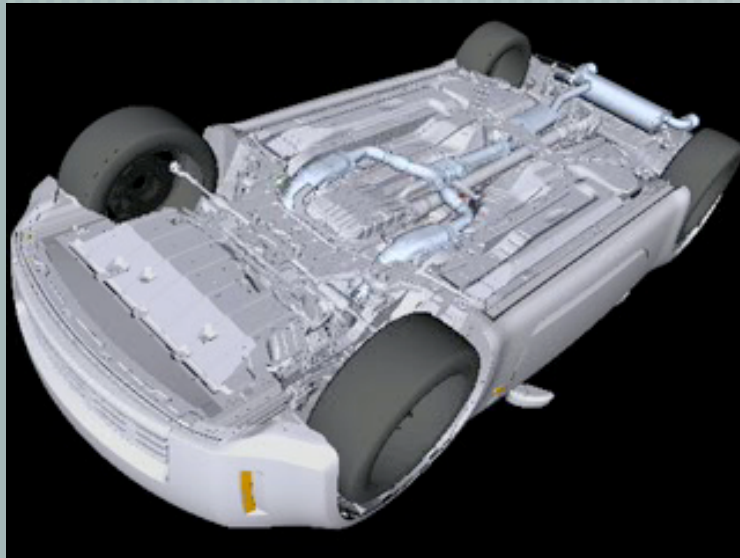
— [北海道大学大学院工学研究科 客員准教授 (2007-)

— [理研ベンチャー アイサイブ 代表取締役 (2007-)

— [東京大学 生産技術研究所 (2011-)



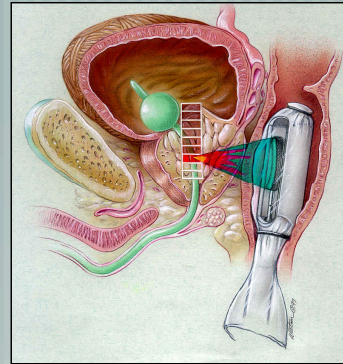
Automotive Engineering



東京大学大学院 機械工学専攻

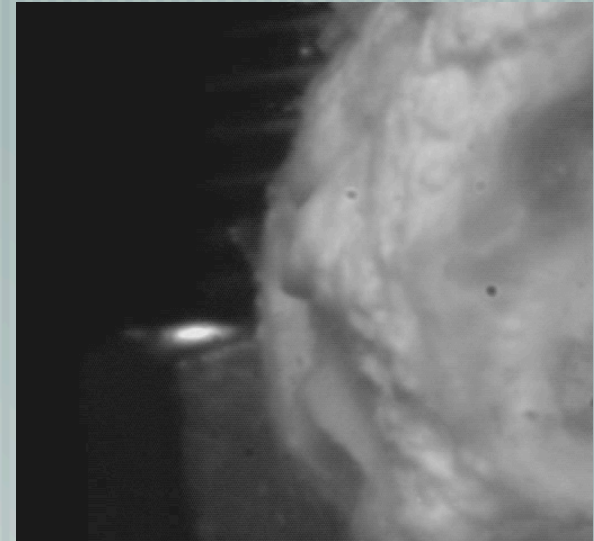
— [強力超音波による癌治療

— 脳腫瘍, 前立腺がん



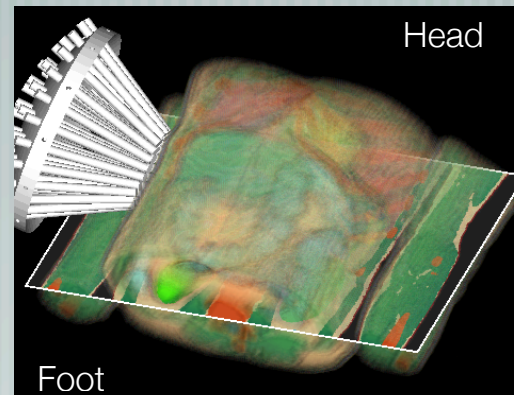
— [結石破砕

— 腎臓結石, 胆嚢結石



— [新しい研究へチャレンジ

— 可視化



— [文科省在外派遣プログラム

— 米国で半年の生活

— ワシントン大, カリフォルニア大デービス校

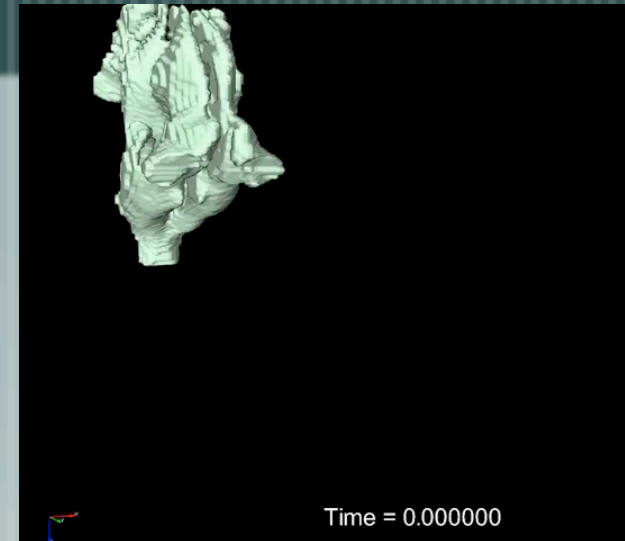
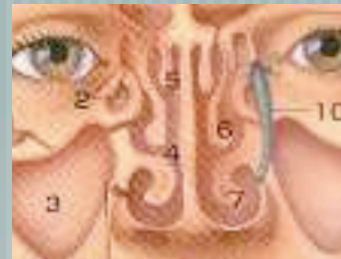


理化学研究所

VCADプロジェクト

熱流体シミュレーション

生体・工業製品



熱く燃えるザリガニ

小野謙二 1)、横田秀夫 2、3)、寛正信徳 3)、西村将臣 2)、姫野龍太郎 3)

- 1) (独) 理化学研究所 製品機能シミュレーション開発チーム
- 2) (独) 理化学研究所 VCAD 開発チーム
- 3) (独) 理化学研究所 生体力学シミュレーション特別研究ユニット



東京大学 生産技術研究所

- [神戸に移る
- [スパコンを使った研究
 - 日本ではここでしかできない研究
 - シミュレーションコードの開発
 - 結果の可視化研究

Large-Scale Computer Resource

— [for what?

— Science and Engineering, related to social activity

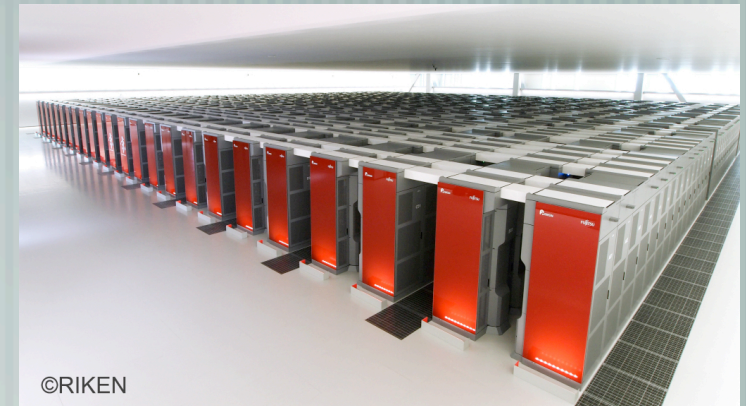
— [improve

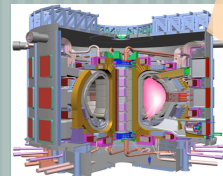
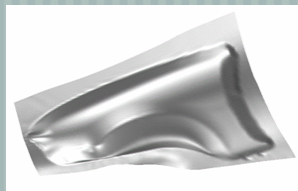
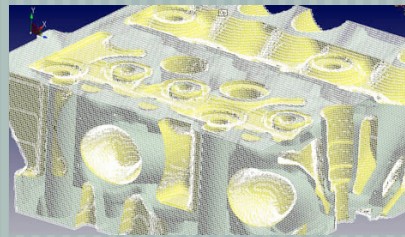
— Accuracy

— More complex physical model

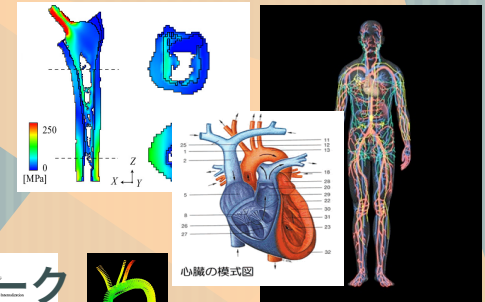
— Real world problem (complex geometry, BC,...)

— Time frame





人間系全体解析



デジタルエンジニアリング

核融合

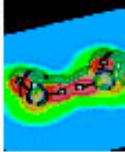
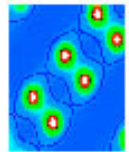
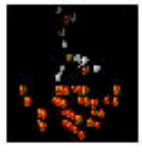
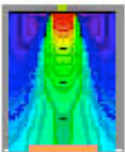
生体分子ネットワーク

Feature scale

Reactor scale

Atom, Molecule

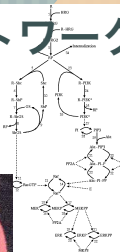
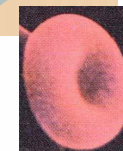
Electron, Nucleus



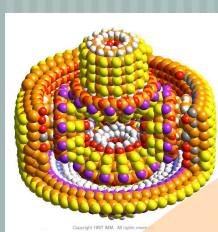
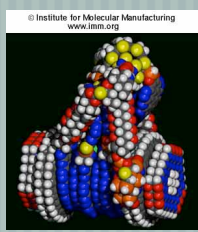
血流解析



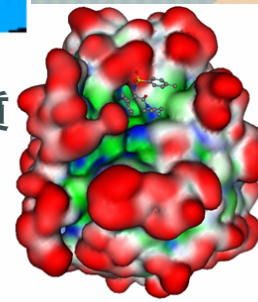
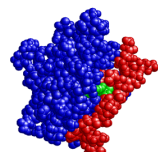
発病メカニズム解析



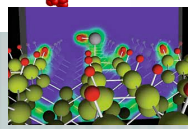
ナノマシン設計



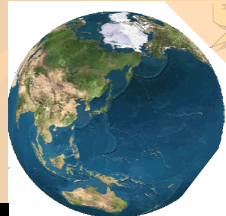
タンパク質



結合能解析 (創薬)

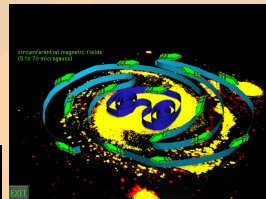


都市環境設計/地域防災

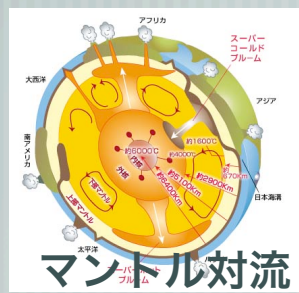


地球の誕生

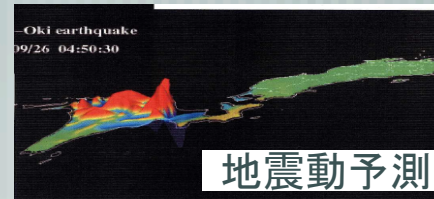
分子構造



銀河の形成



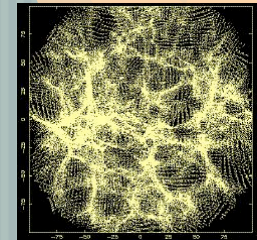
マントル対流



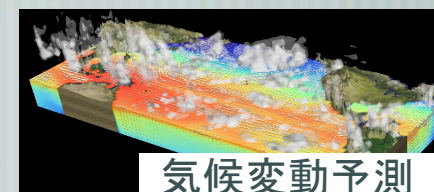
地震動予測



噴火予測



宇宙の誕生

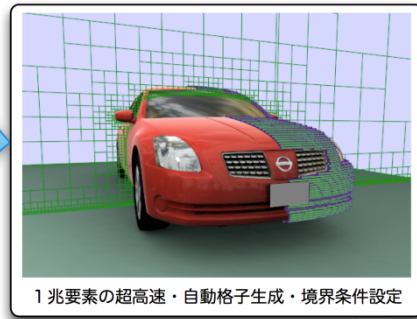
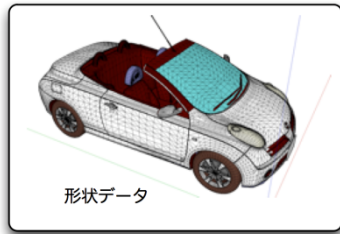


気候変動予測



溶岩流シミュレーション

低エネルギーインパクトなものづくり技術



プリプロセス

製造過程やランニングコストを抑制により、エネルギー問題に貢献
最適設計

実世界の現象を再現

複雑な現象・複雑な形状
構造，熱，流体，材料，分子などの連成シミュレーション

短時間シミュレーション

解析のターンアラウンド全体の短縮
複雑な境界条件を伴うソルバーの超並列解析技術

エンジニアが実問題の設計をシミュレーション
データに基づいて検討する

計算科学，計算機科学，情報技術，機械工学などの
分野融合技術を駆使した統合的なシミュレーション
技術の開発とシステム化が必要

製造・性能・構造設計

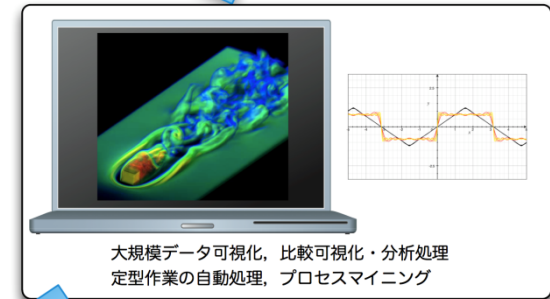
Pflops級
Strong scalingで1H

連成・マルチスケール解析

ポストプロセス



ワークフロー自動処理
最適化
ノンパラメトリック解析



知識共有



事例，テンプレート，分析手法



Issues of large-Scale Simulation

- [Pre-process
 - Large-size model
- [Simulation
 - Accuracy, time, complexity
- [Post-process
 - File size, # of files, IO, usability,...
 - Results -> 10~100TB
 - unsteady, multi-variate, parameter study...

Issues of Post-process

— [Many files -> one, two, three...many

— [File I/O -> Slow, slow...

— [Large dataset >> High cost of data move, copy

— [needs scaling visualization and data processing system

— [leads scientific discovery -> Important infrastructure

— Image and Analysis

— Concentrate on thinking

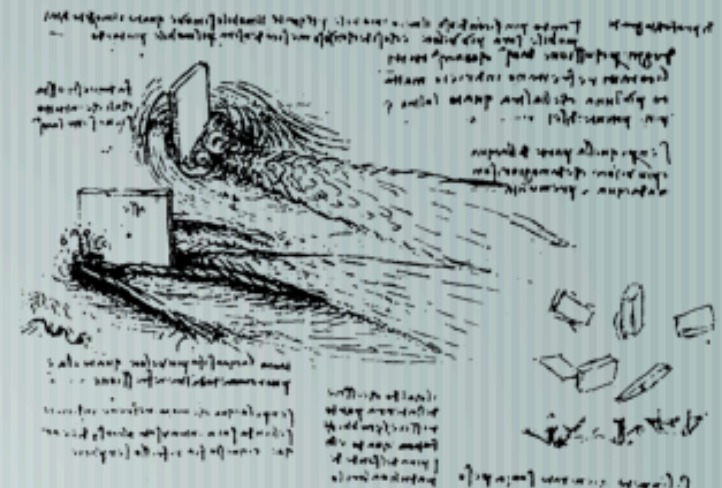


図1.14 障害物のまわりの流れ(レオナルド・ダ・ビンチのノートから)



Visualization Research in the Past

Current Techniques

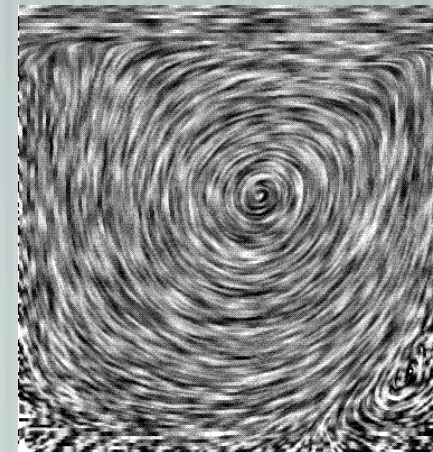
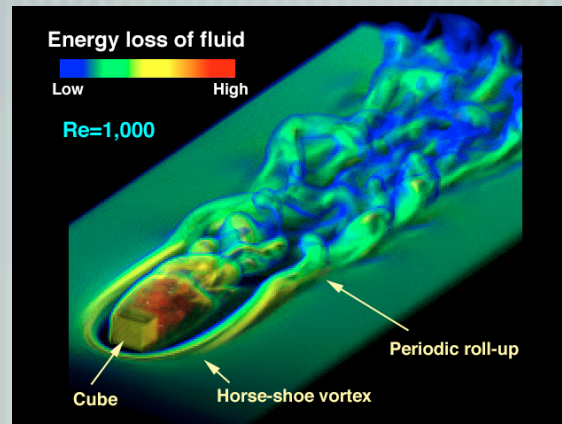
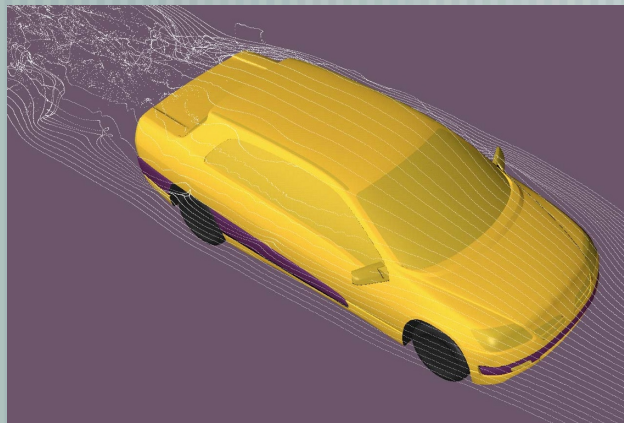
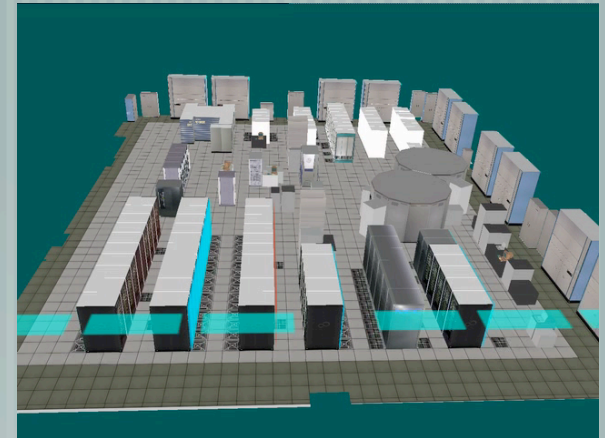
Invented in 80's

Increment

1990~ Parallel, GPU

2000~ Large-Scale Dataset

Scientific Data Management



ケーススタディ 可視化分析のストーリー

場の全体構造の把握

- 流線やボリュームレンダリングなど
- 非定常挙動, ROIの特定

写實的イメージによる観察

- 人間の認識力を利用したディテールの把握

簡素化した表現

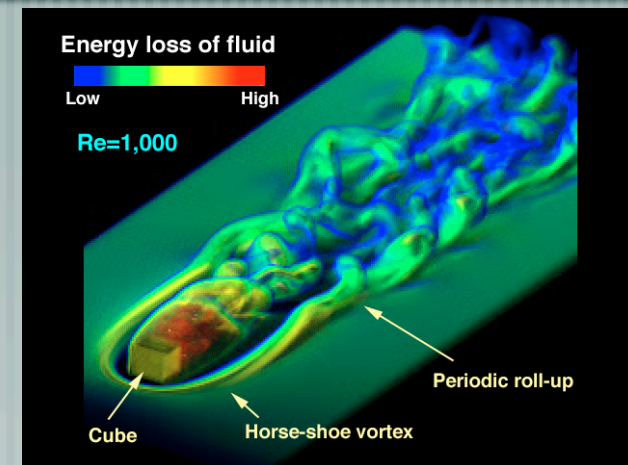
- 空間構造を記号的に把握
- 時系列の変化など複雑なものを抽象的に把握

詳細な分析

- 例えば,
 - ROIからの流線の逆積分
 - 温度勾配と渦との関係

知識への昇華

- 様々な分析結果を統合的に見て高次の知識へと昇華
- アノテーションによる知識表現



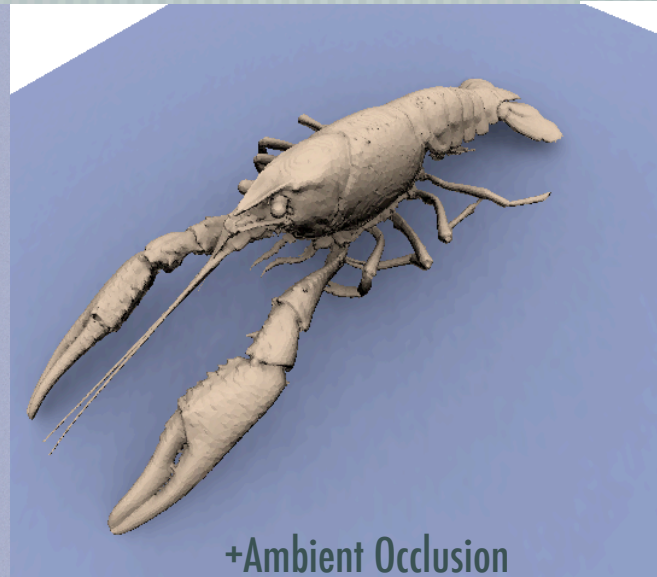
考える過程を支援
観察と分析を同時に

Technology Trend in Vis. Field

- [Vector/Tensor, Diffusion Tensor Imaging
- [Parallel Rendering
- [Visualization System
- [Large Data Visualization, Level Of Detail
- [Time-Varying, Out Of Core
- [Volume Visualization
- [Surface/Geometry
- [Feature Extraction
- [Applications: Flow Vis., Bio-Med.
- [Perception
- [NPR, Illustrative Rendering

Rendering (1)

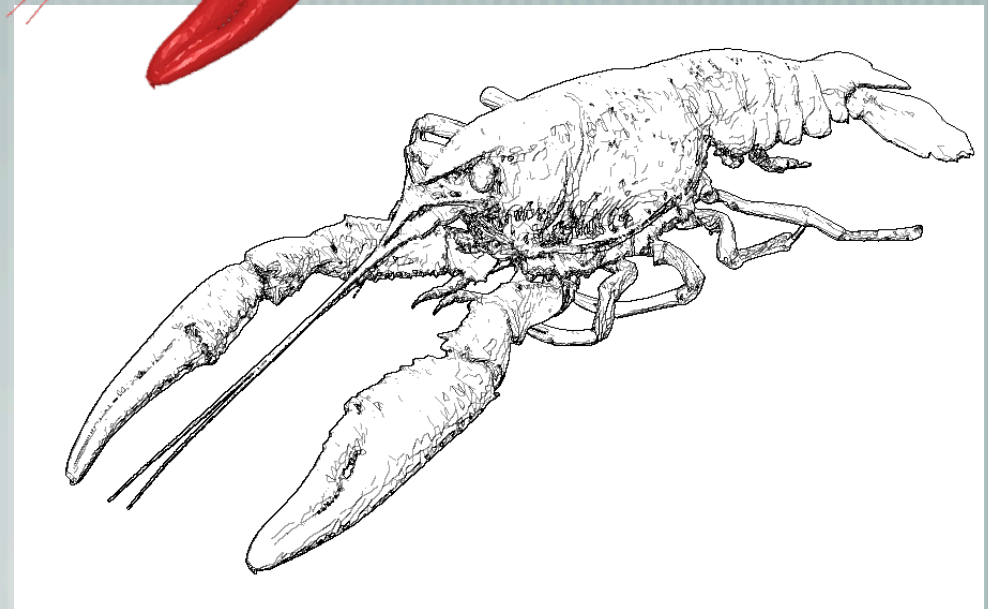
- [Photo Realistic Rendering
 - faithfully imitate the Optics
 - Ray Tracing
 - Path Tracing
 - Global Illumination Indirect
 - Ambient Occlusion
 - BDRF...



Rendering (2)

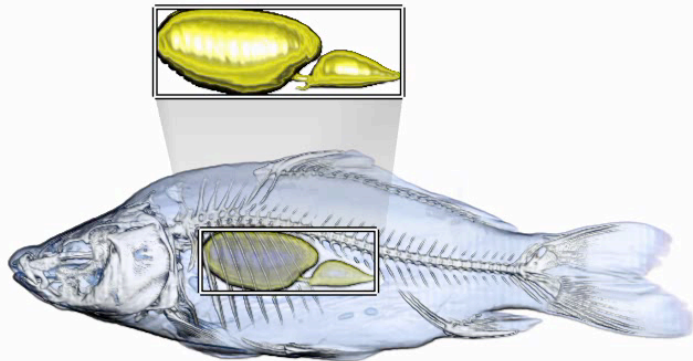
Non-Photo Realistic Rendering

- 線画, スケッチ風
- 特徴を際立たせる
 - 特徴の抽出
 - 本質を描画
 - S/N比の向上

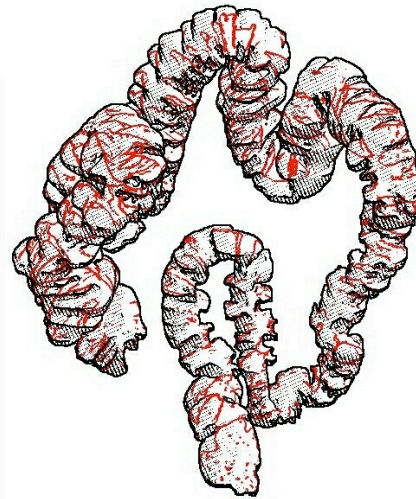


NPR, Illustrative Rendering

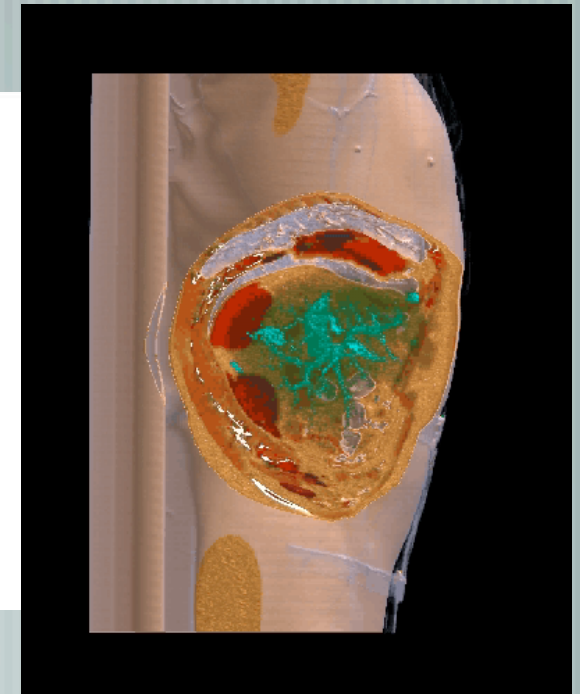
- [Fully dynamic illustration
- [Improve understanding
- [Importance Driven



Bruckner, Vis2005

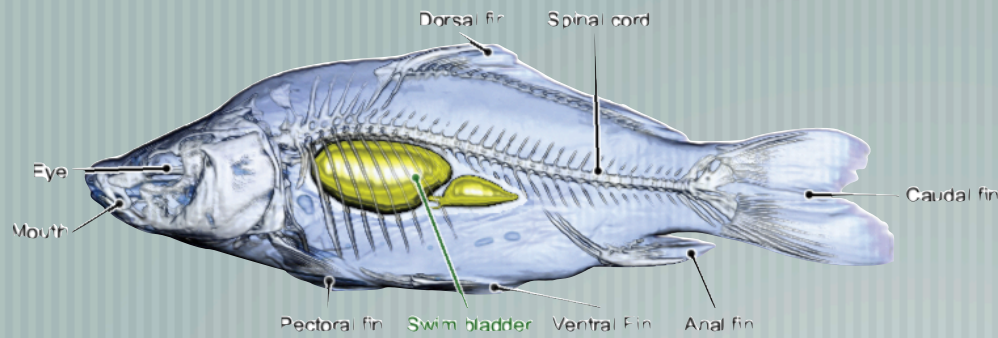


17

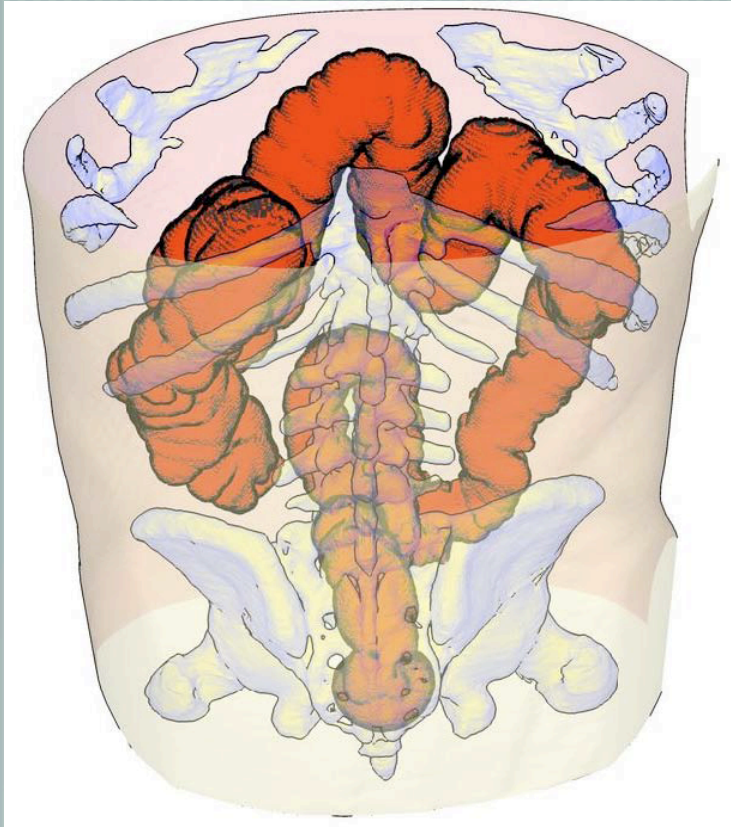


Viola, Vis2004

Illustrative Rendering & Annotation



Help to Comprehension



Simplification



Abstraction

GPU Technology (GPGPU)

Driving force

- フィルム産業, 医療分野, コンシューマーゲーム市場

データストリームを効率的に処理するHW

- 大量のデータに高密度な演算を施し, 低レイテンシで出力
- 並列処理向き (処理前後のデータ処理の依存関係がない)
- **有効活用のためには, CPUとは異なるアルゴリズムが必要**

Applications

- Linear algebra
- Sorting, Searching, Compression
- Physically-Based Simulations
- Rendering

Parallel Large Data Visualization

Query Driven Vis.

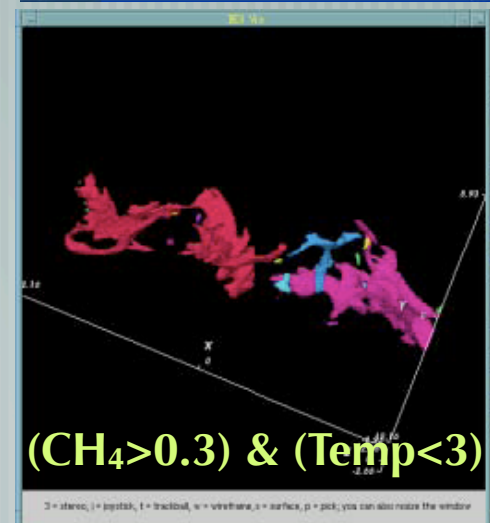
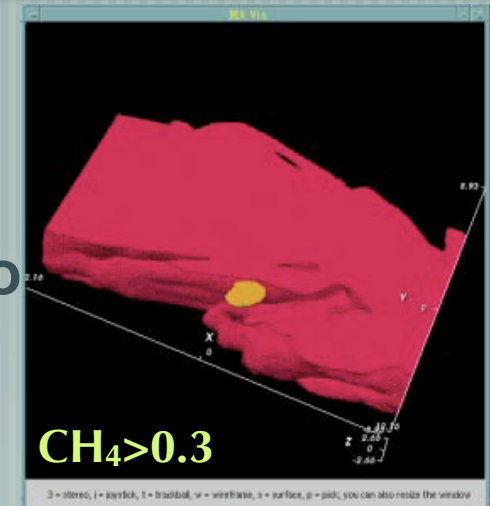
Distributed Data Management of Large Volumes

Level Of Detail

Roaming

COTS Cluster-based Sort-last Rendering

OpenGL Multi-pipe SDK



Time-Varying Large-Scale Data Set

— [並列分散データ

- シミュレーションのデータ分割方法 — 領域, オブジェクト
- データの移動はコストが高い (ディスクIOは低速)
- サーバクライアント方式
 - 要求 (コマンド) に対し, 結果 (イメージ, 情報) を返す

— [時系列データ

- 高速・並列IO
- プリフェッチ

— [圧縮, メタ化

Feature Extraction

— [特徴領域 (ROI) の抽出

— Segmentation

— [対象分野毎の個別技術

— 流体

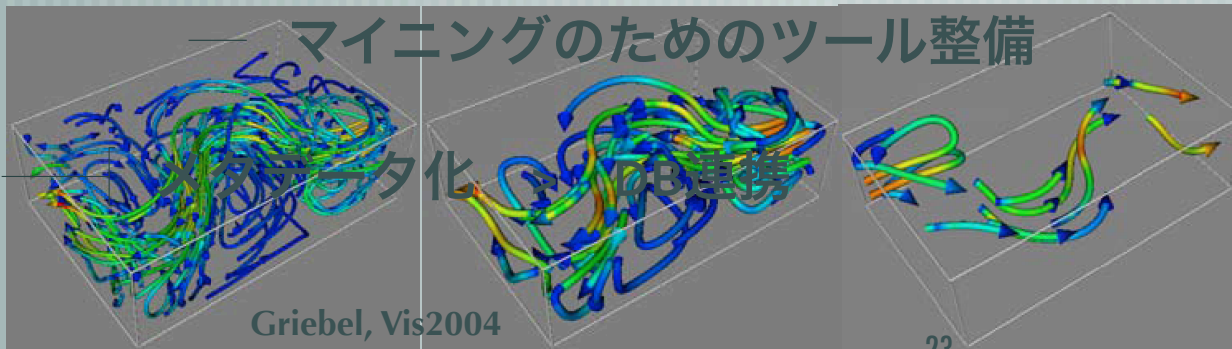
— 渦心の特定と抽出

— 渦構造の同定

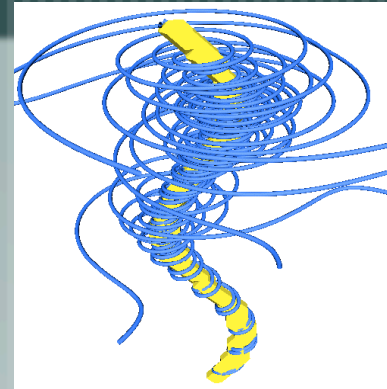
— ベクトル場

特異点解析

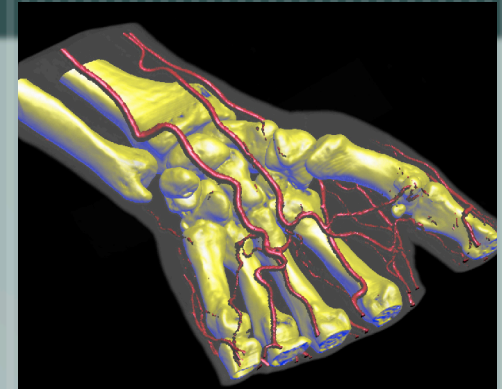
— マイニングのためのツール整備



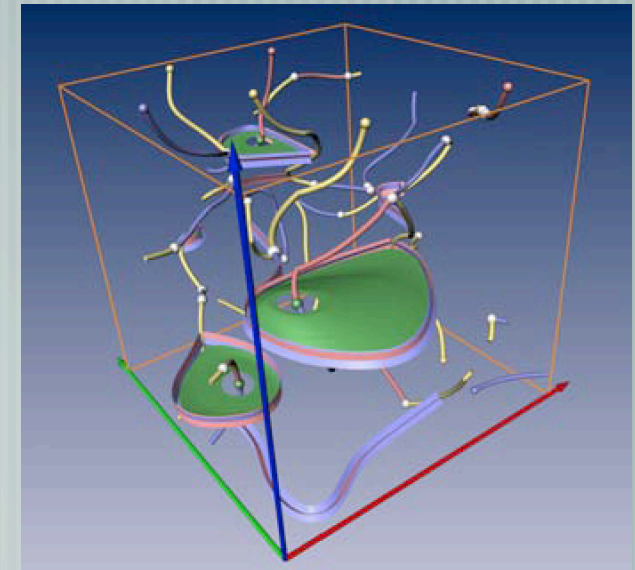
Griebel, Vis2004



Jiang, Vis2004



Hadwiger, Vis2003



Theisel, Vis2004

User Interface and Display Devices

- [Wall / Tiled Display
 - 大画面, 高解像度
 - プロジェクション技術
- [IPT
 - CAVE, CABIN
- [立体視
 - アクティブ方式, パッシブ方式
 - 裸眼立体視
- [ユーザーの利便性を考慮したインターフェイス
 - 使いやすさ, カスタマイズ性



CABIN (IML, U.Tokyo)

インテリジェントデータリダクション機能

— [データ圧縮技術

- 符号処理による圧縮・伸長
- 量子化
- 間引き
- LOD技術

— [観察対象となるデータの抽出

- 領域選択 (ROI, Region Of Interest) , ROIの移動 (Roaming)
- 特徴抽出 (Feature Extraction) により, 小規模のメタデータに変換

— [データベースの利用

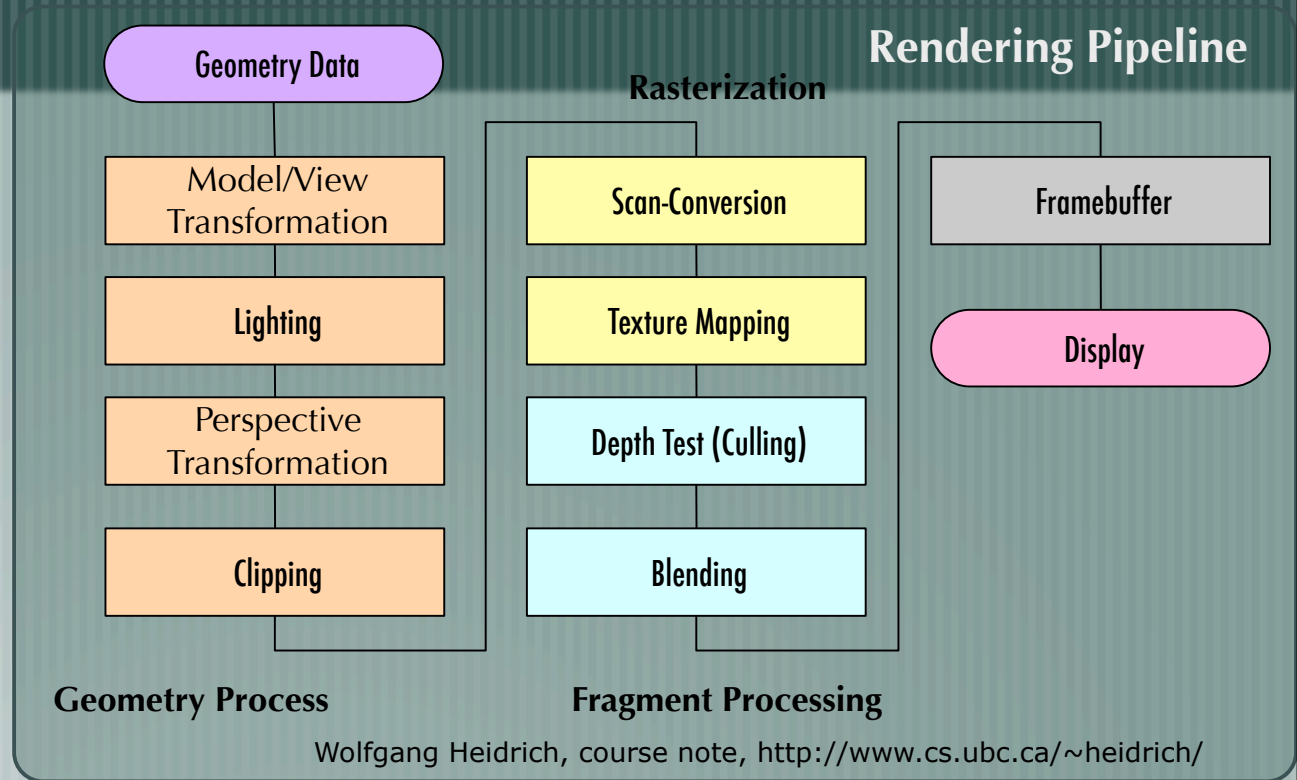
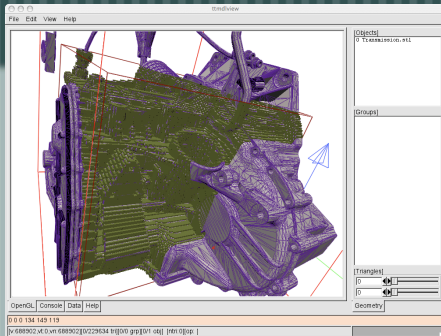
- パラメータ並列の分析作業時のデータ管理
- 可視化結果の管理
- 生データとメタデータとの関連づけ

— [インタラクティブ可視化の前処理の機能も提供

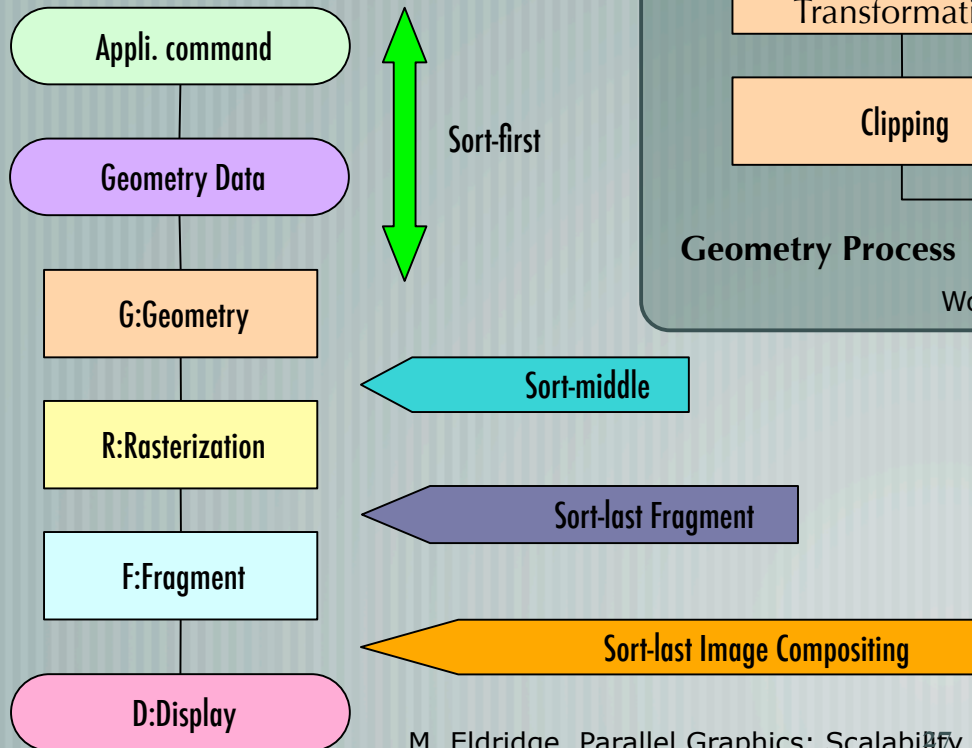
大規模可視化の要素技術

- [Sort-Last Image Compositing
- [Interactive Visualization
- [In Situ Visualization
- [(Intelligent) Data Reduction
- [Visual Analytic
- [Data Management
- [Workflow

Classification of Parallel Visualization



Sorting Taxonomy



描画処理のどのプロセスで並列化するか

- ・ 並列化の粒度
- ・ ネットワーク帯域
- ・ 描画性能

M. Eldridge, Parallel Graphics: Scalability and Communication, Course Note 37, SIGGRAPH 2001.

Various Scenario of Visualization

- サーバー/クライアントの**実行モード**

モード	クライアント	可視化プロセス
オンサイト	ユーザサイド	ユーザPC
	リモートサイド	サーバー
リモート	ユーザーサイド	サーバー

- **ユーザインタラクション**

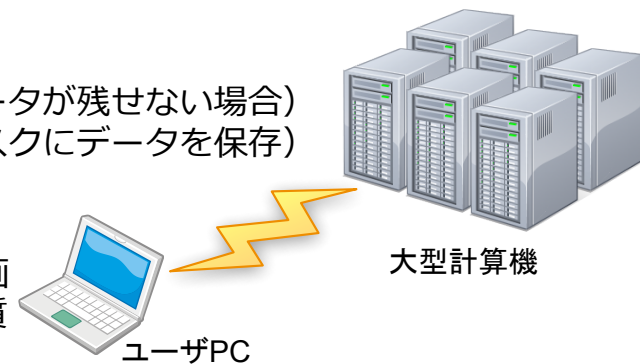
- インタラクティブ : ユーザの操作に応じて可視化
- バッチ : 定義されたシナリオに基づき可視化

- **実行タイミング**

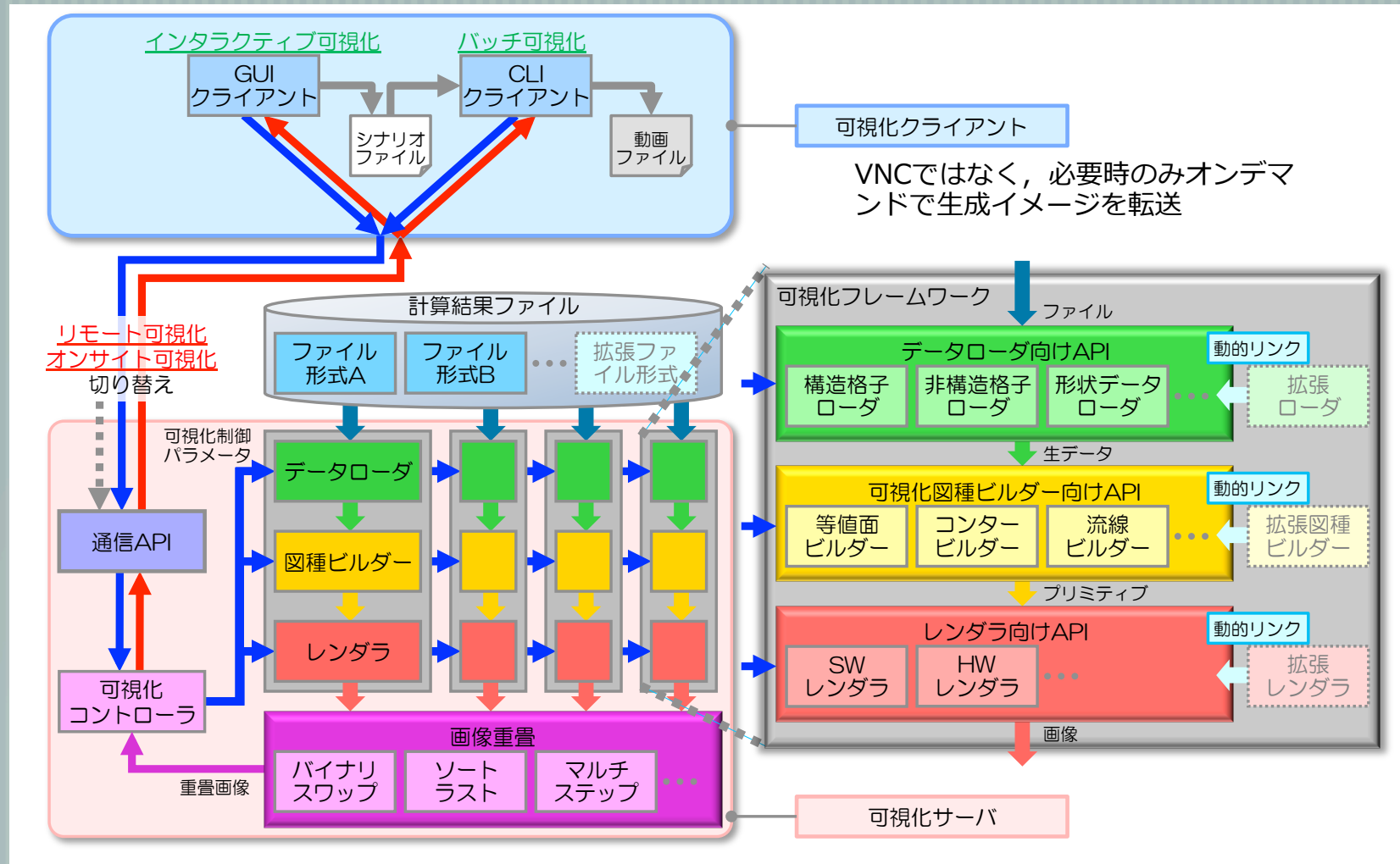
- リアルタイム : 計算途中に可視化 (データが残せない場合)
- ポスト : 終了後に可視化 (ディスクにデータを保存)

- **レンダラー**

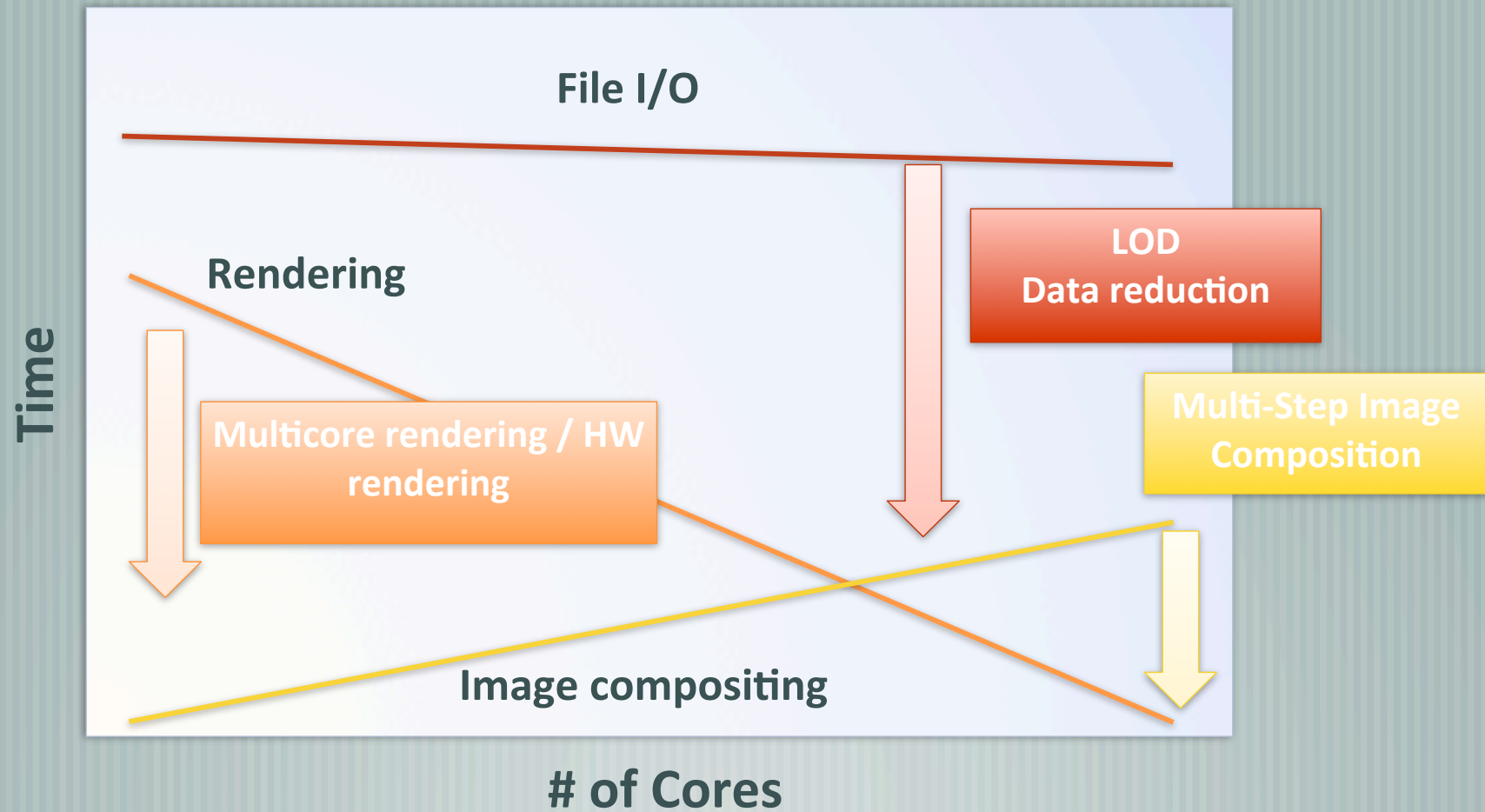
- ハードウェア : GPUを利用した高速描画
- ソフトウェア : CPUで描画可能, 高品質



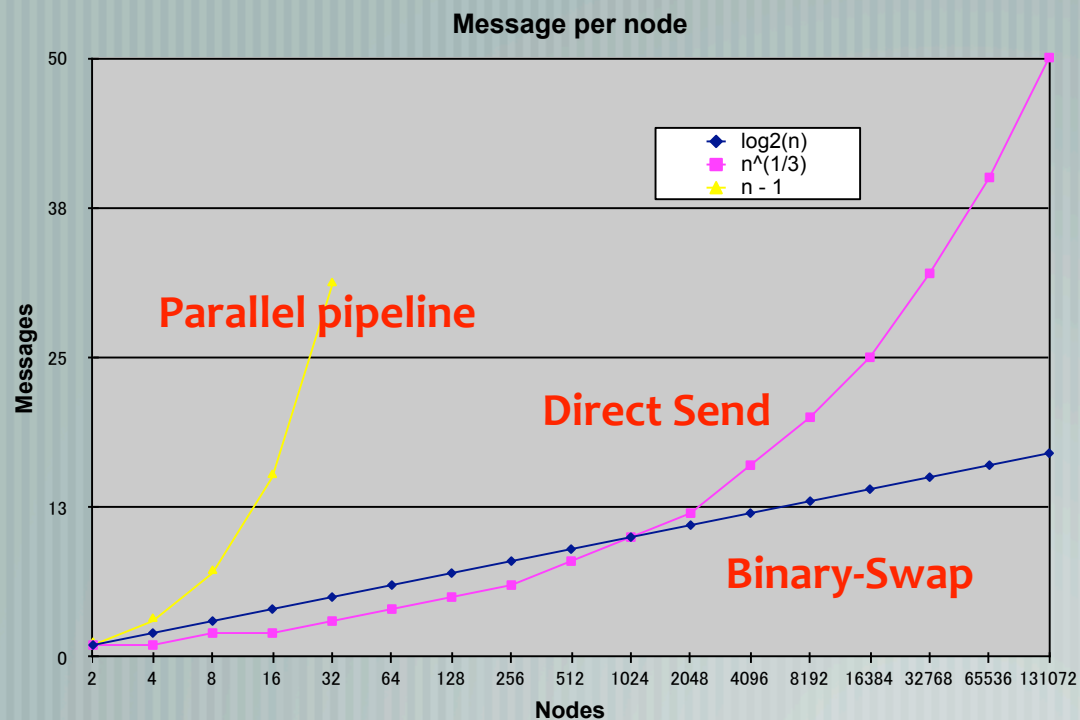
LSV (Large-Scale dataset Visualization)



Strategy: interactivity and scalability



Performance of Image Compositing



Binary-Swap

High scalability

Static scheduling

Several optimizations

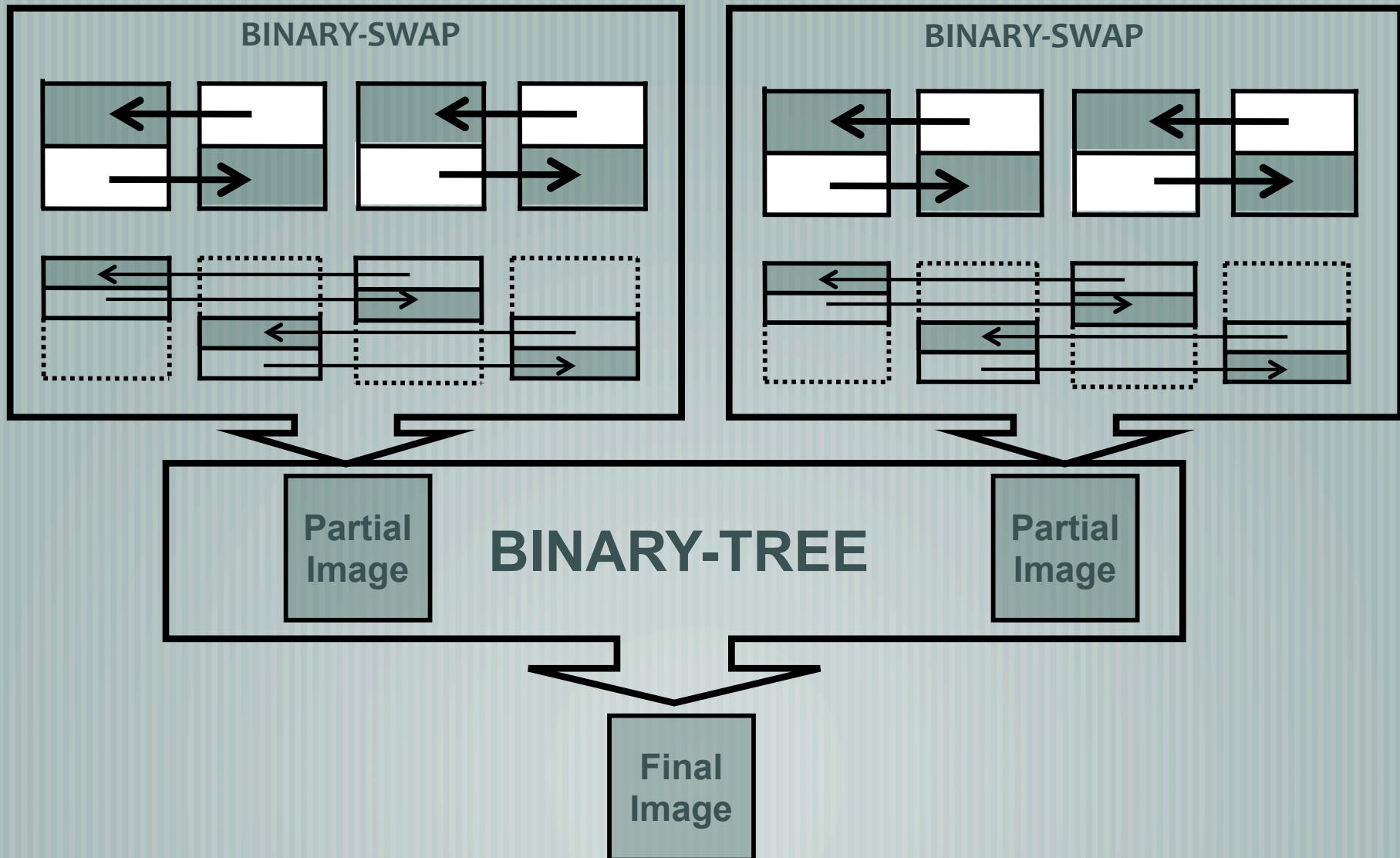
Easy combination

MPP

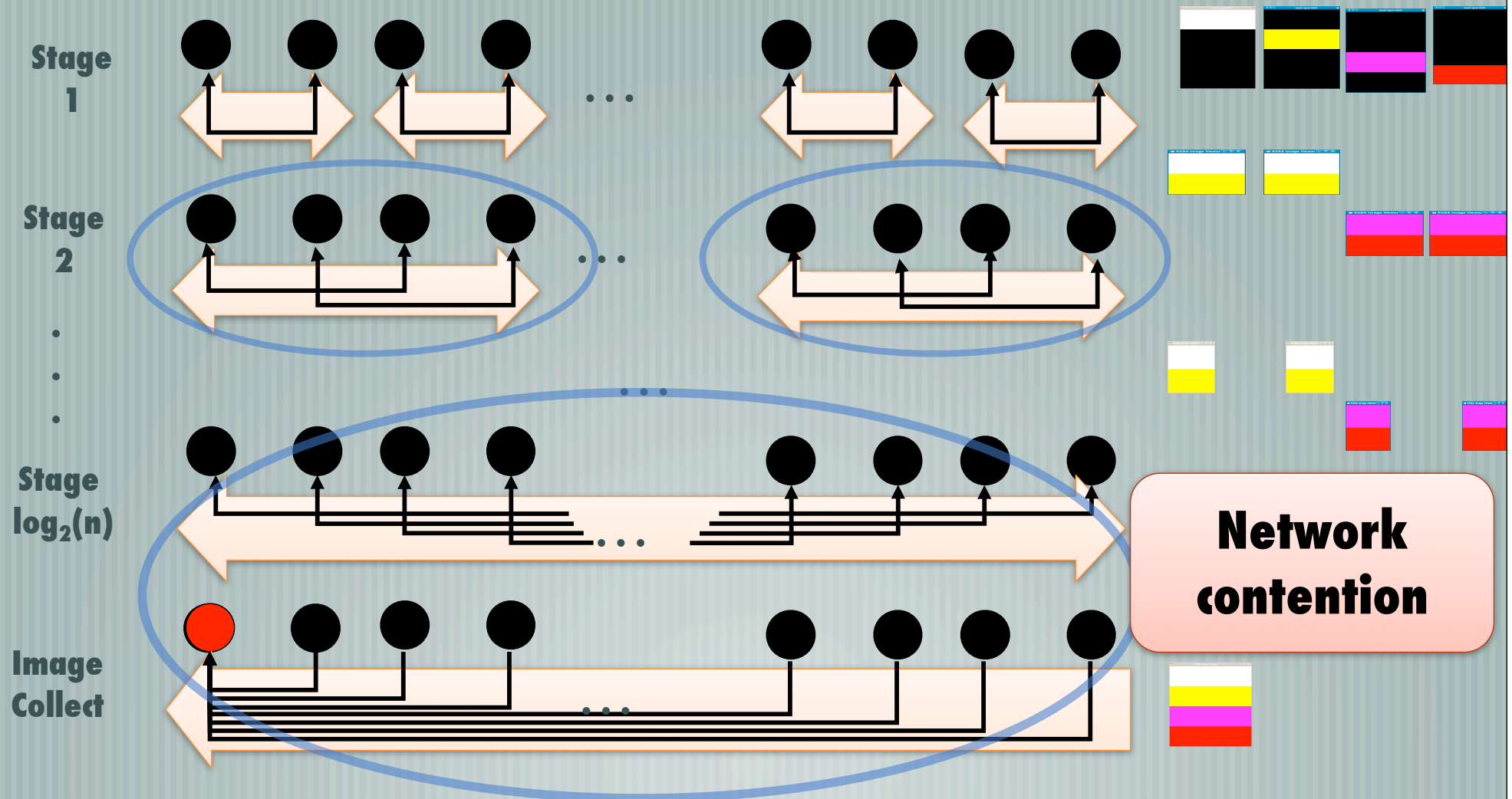
RIKEN One-Rack
BG/L

2048 processors

Hybrid Image Compositing

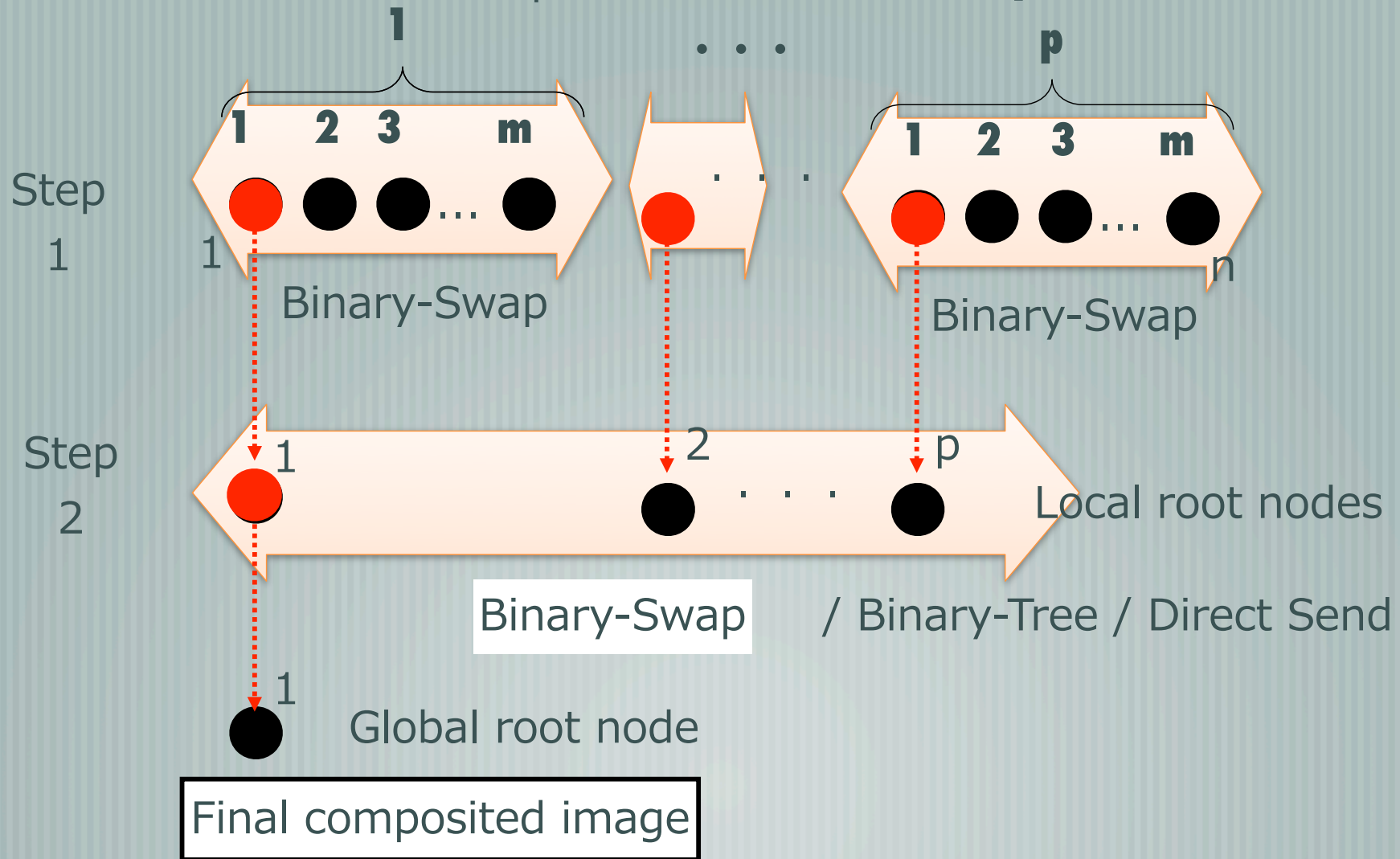


Binary-Swap Image Composition



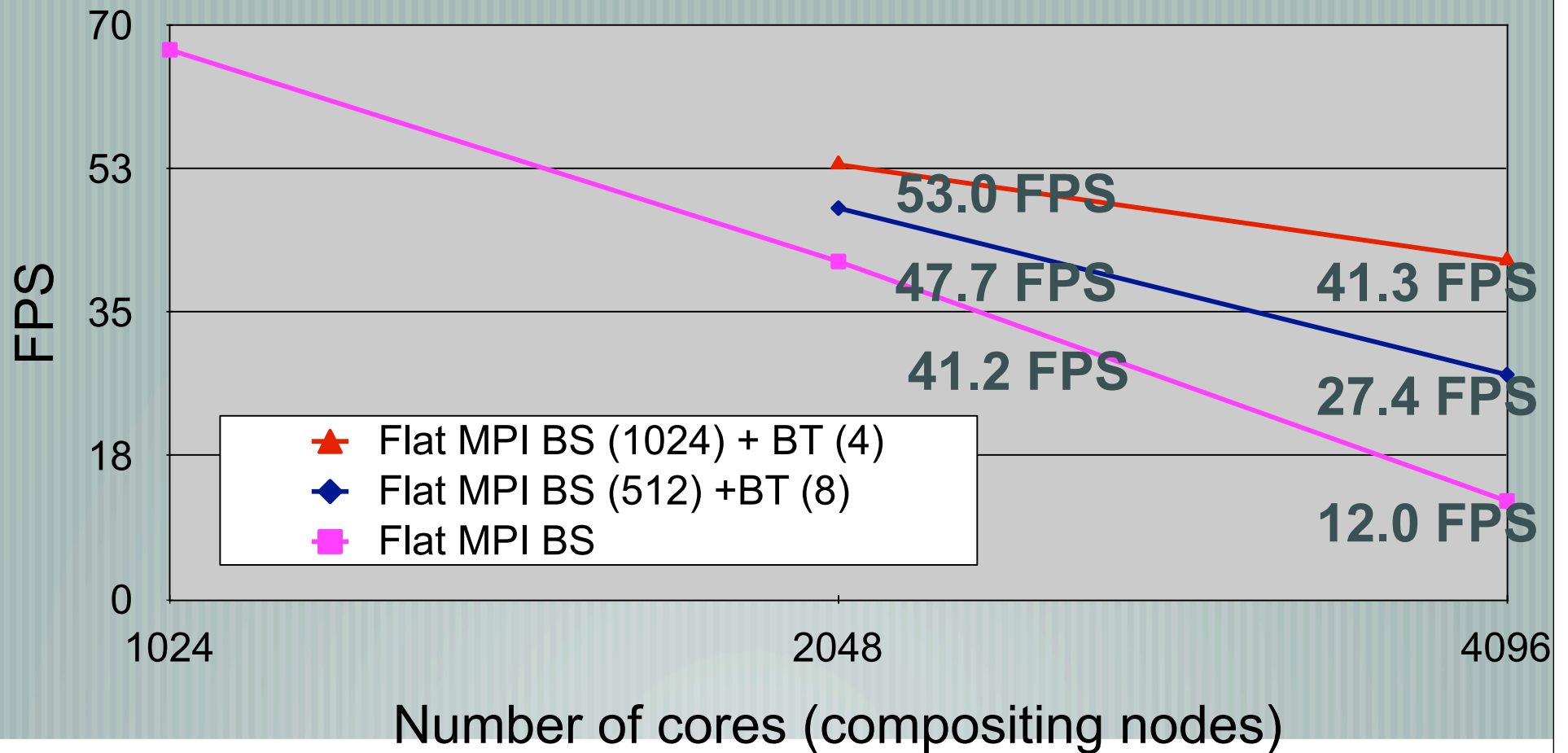
Multi-Step Image Composition

Composition Nodes: $n = p \cdot m$



Multi-Step Image Compositing

Flat MPI BS+BT Compositing (512x512 RGBA)



ポスト処理のコンセプト

— [シミュレーション結果から、有用な知識・情報を効率的に引き出す

- 様々な可視化手法による場の把握
- 生成された2次データの利用
- イメージ化
- データ分析
- 多様・大量なデータの管理のしくみ

— [データ管理

- シミュレーション開始から分析後までの全てのデータ
- 個々のデータの関連づけ

— [処理の効率化

- 定型作業の自動処理

— [ユーザ作業環境

- 単独で動作するコマンド，アプリの連携・緩やかな結合

シミュレーションの価値を増幅させる，知識を得るためのツール群