Technology Maturity for Adaptive Massively Parallel Computing

First Workshop 2009

March 2-3, 2009

Portland, OR, USA

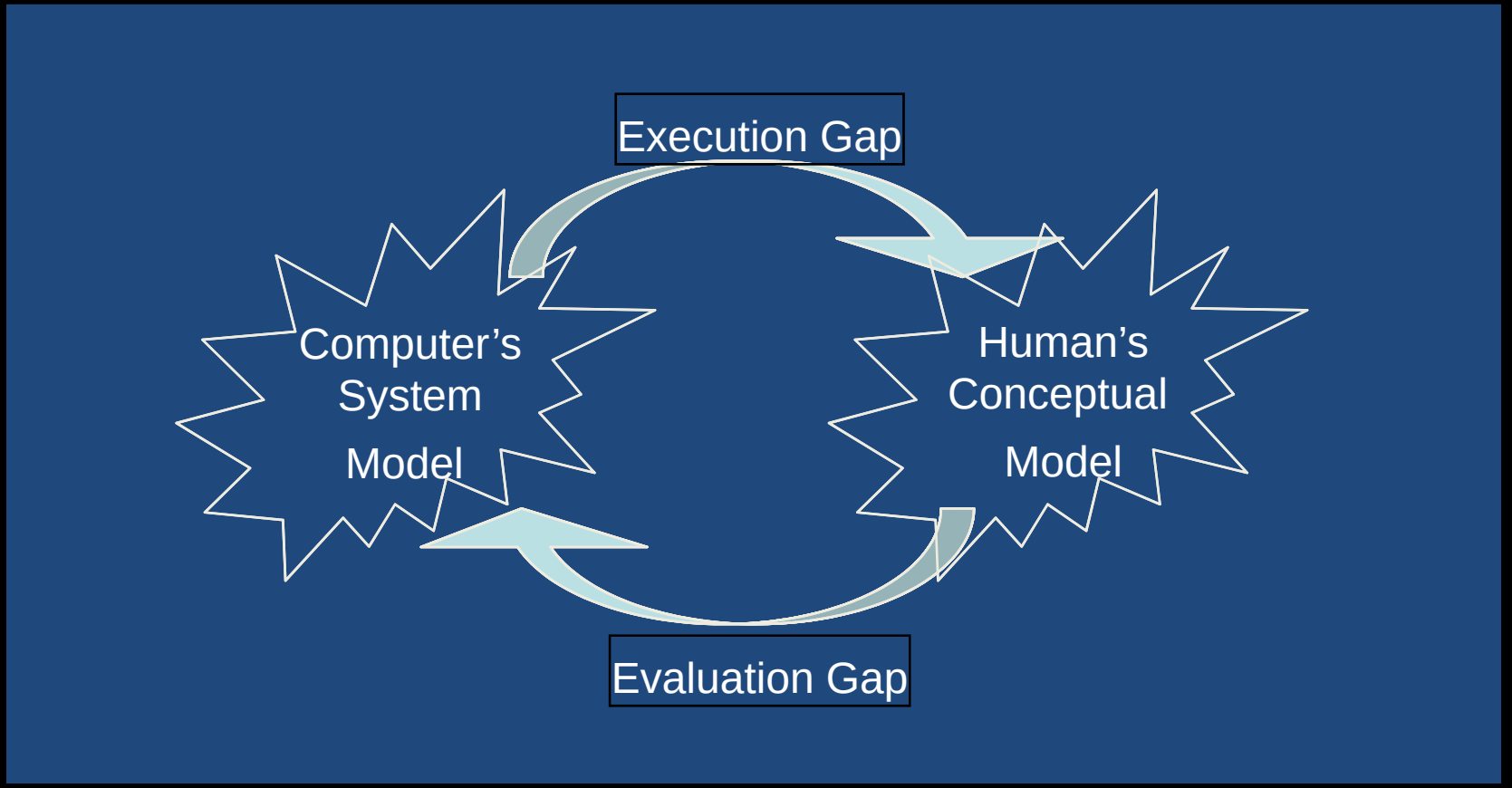
Massive Data Computing

Pradeep K. Dubey

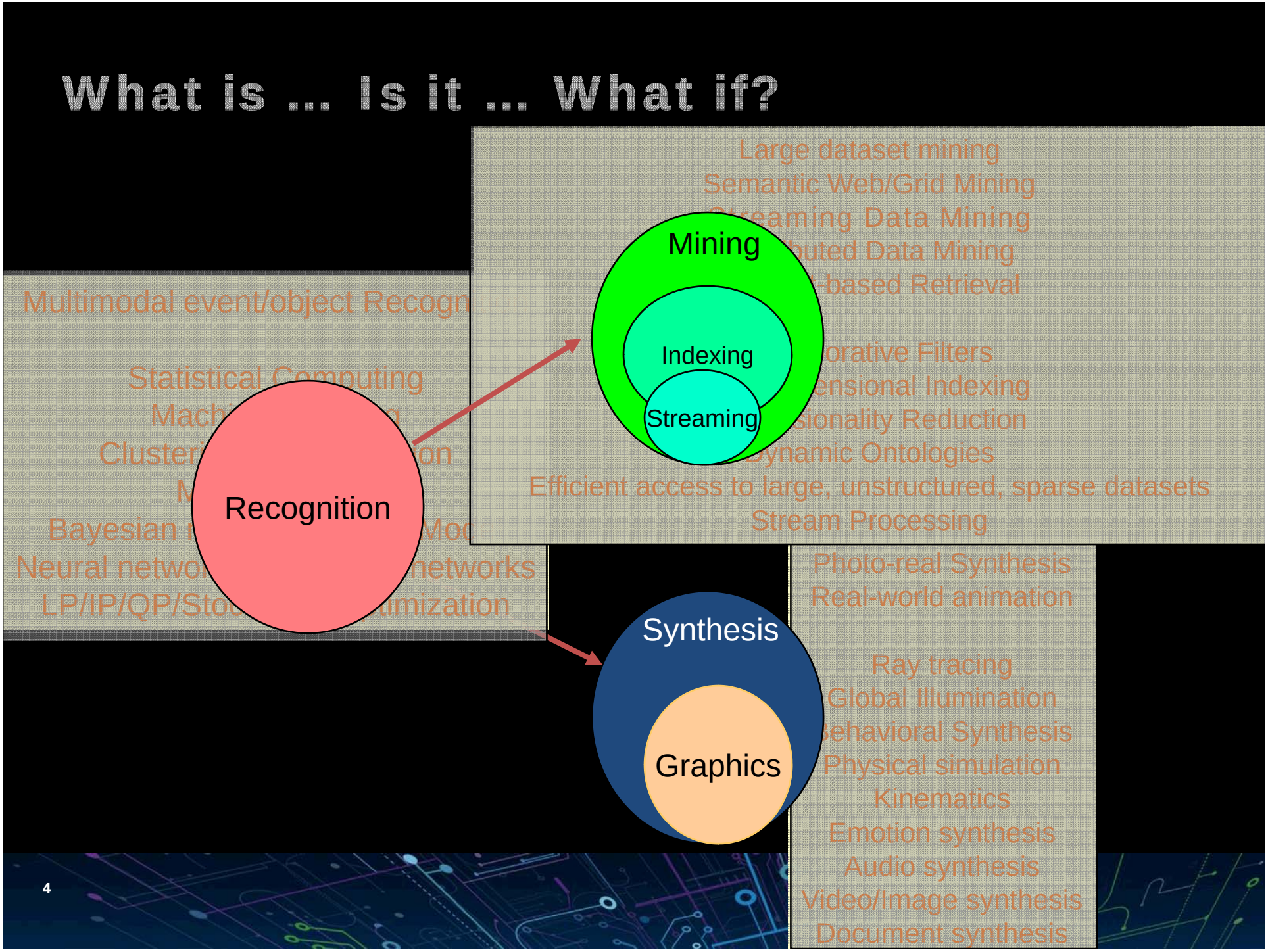
Senior Principal Engineer

Intel Corporation

Norman's Gulf

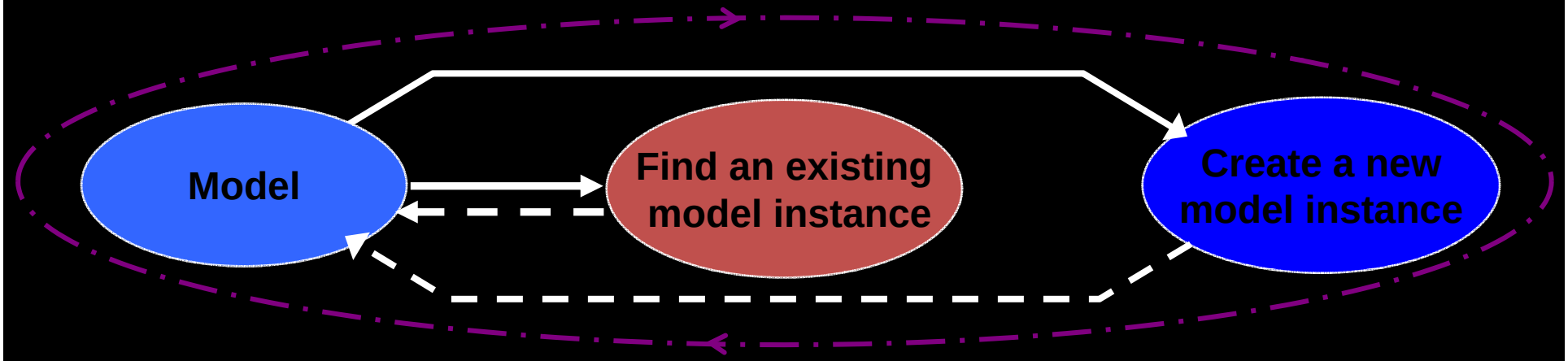


What is … Is it … What if?



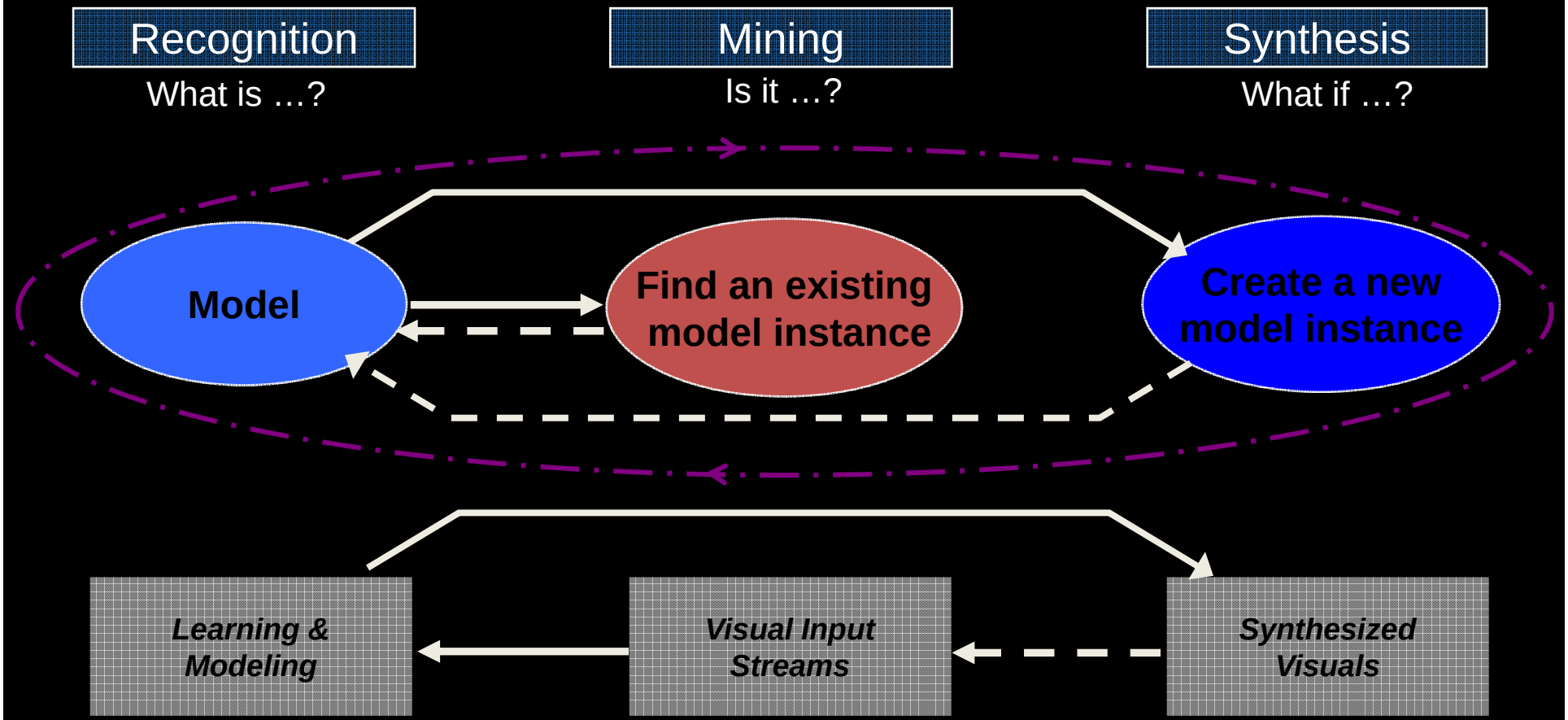
Interactive RMS

Recognition	Mining	Synthesis
What is …?	Is it …?	What if …?

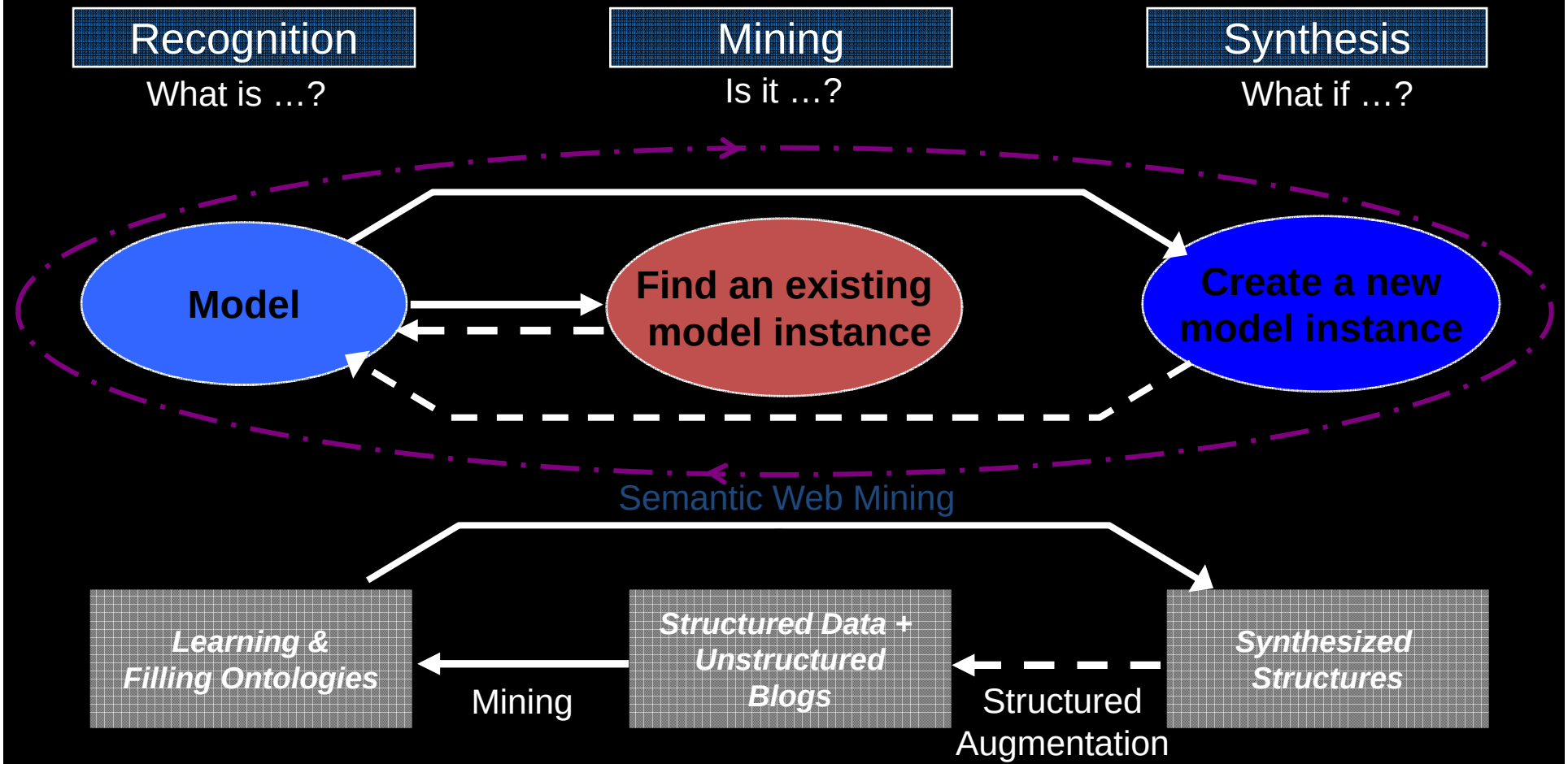


Most RMS apps are about enabling interactive (real-time) RMS Loop or iRMS

Visual Computing Loop

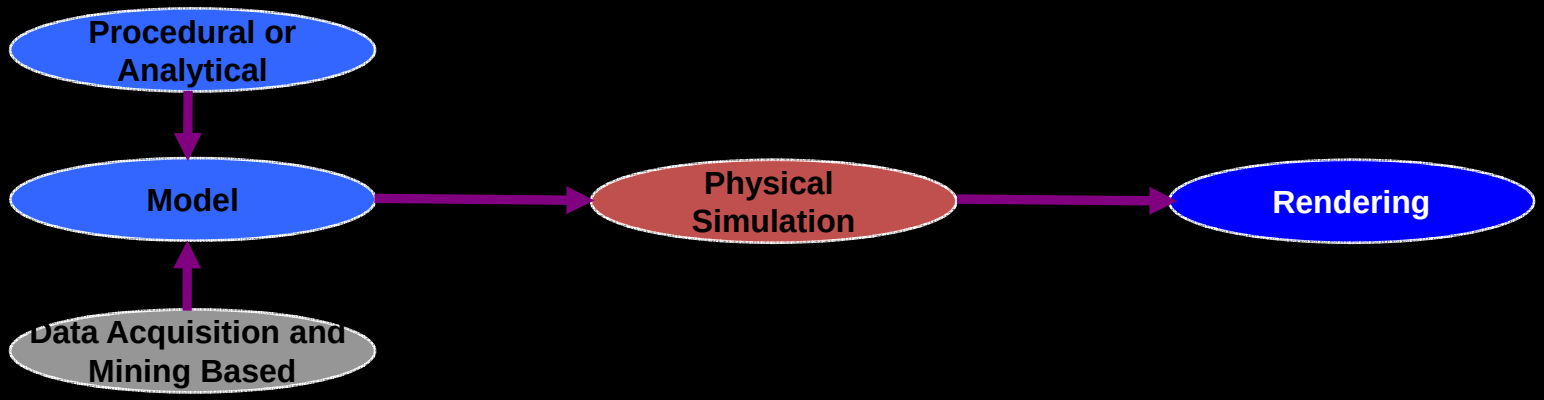


Analytics Loop

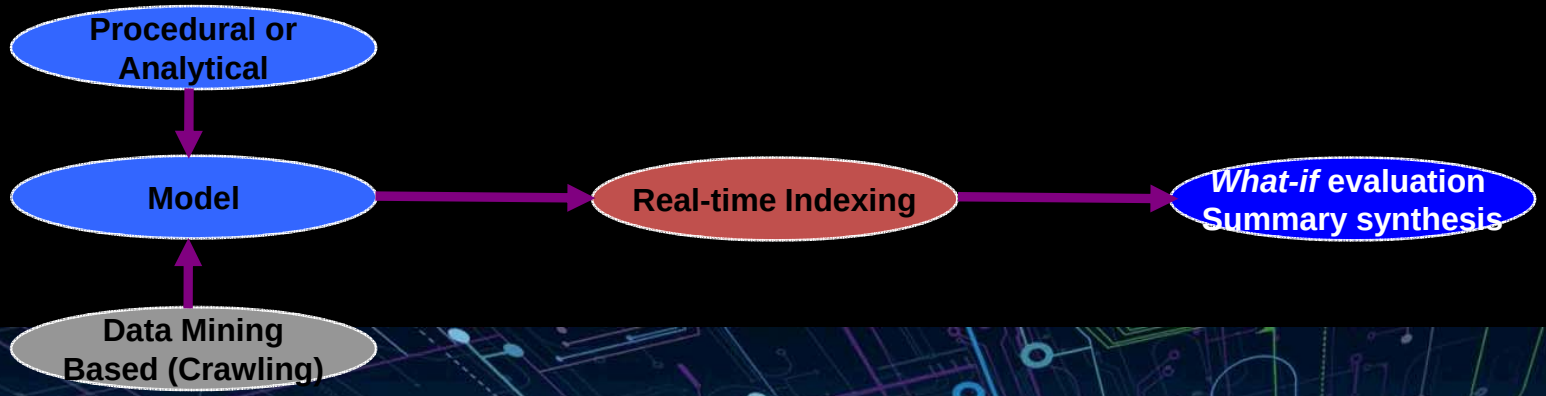


RMS Inner Loop Unrolled

Visual Computing (Graphics and Vision)



Non-Visual Computing (Search and Analytics)



Visual Computing Meets Analytics

Rendering
Simulation

collision detection
force solver
global illumination
...

Machine learning
Neural networks
Probabilistic reasoning

Fuzzy logic
Belief networks
Evolutionary computing
Chaos theory
...

Physics

Dynamics

Soft
Computing

Constraints

Soft
Physics?

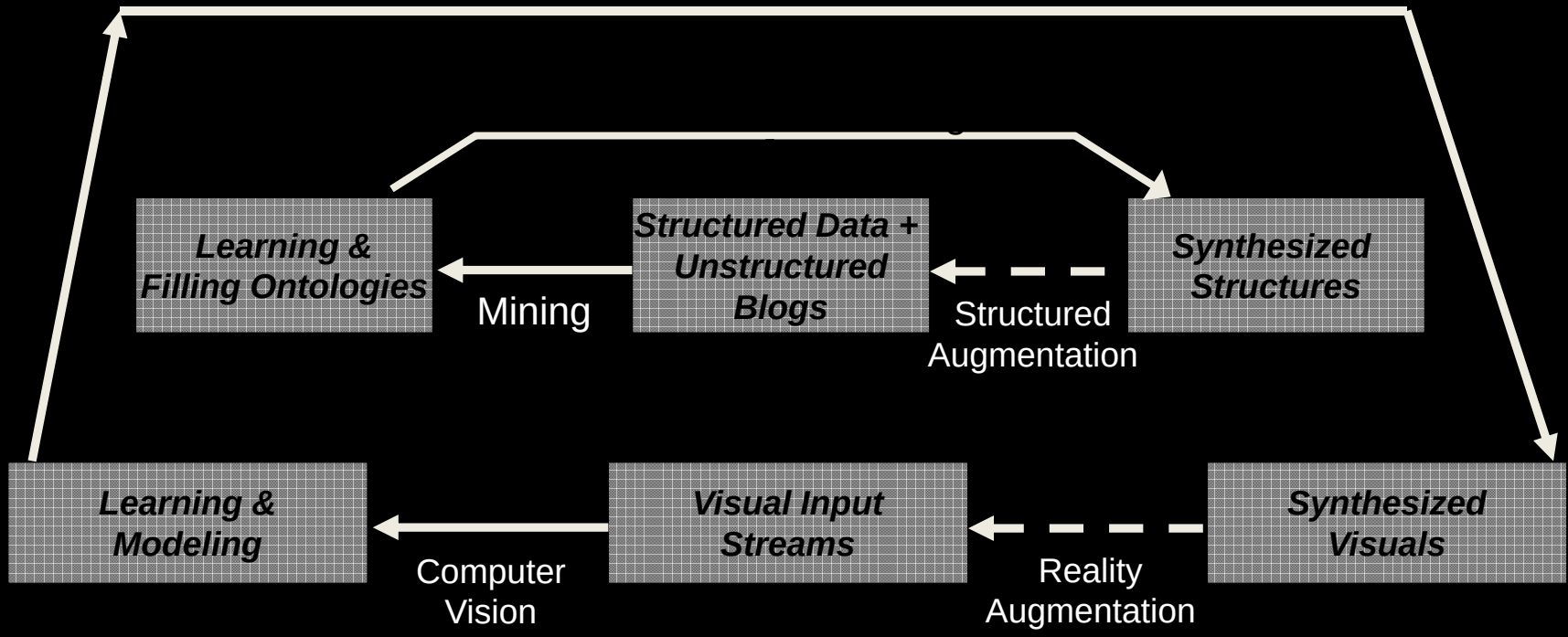
Constraint
Dynamics

Nested iRMS

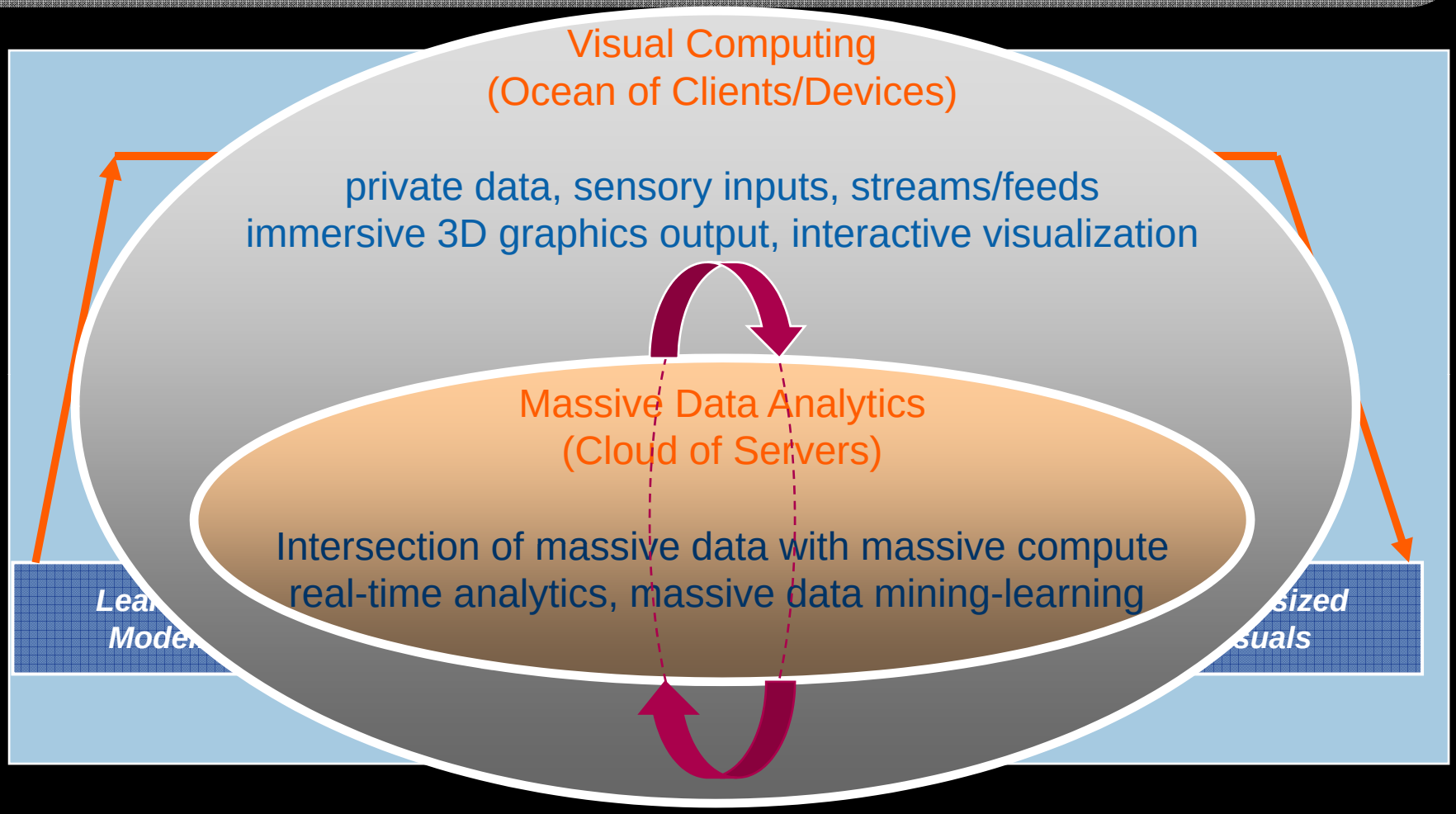
Recognition
What is …?

Mining
Is it …?

Synthesis
What if …?



Connected Computing

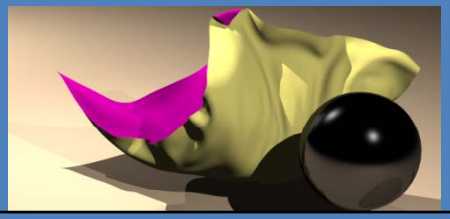


… Architectural implications are far more radical

Computational substrate must undergo a sea-change!

Immersive Computing

Sensory Immersion



Behavioral Immersion

Rendering
Simulation

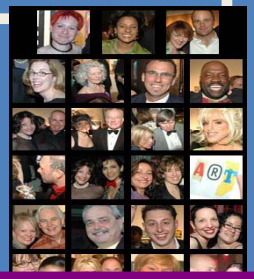
Machine learning
Neural networks
Probabilistic reasoning

Fuzzy logic
Belief networks

Super Immersion

Trade Bots

Outer loop:
iRMS Visual Loop:
One Per Real User



Chat Bots

Inner Loop:
iRMS Analytics Loop

Shop Bots

Performance Needs:
Can far exceed typical i/o limits of human perception

Limited by input/output limits of human perception

Computational requirements are huge, but …

Data changes the game

[Kasparov vs. Deep Blue]

Rule-based system exceeds human performance in a structured, deterministic domain

[Google MT wins NIST contest]

- Statistical inference (not rules)
- 100s of TB of training data
- Racks of computation

Newcomer Google beats decades of rule-based translation research

Opportunities Abound: Massive Data with Massive Compute

Machine Learning Algorithm Classes

- Model-based
  - Transparent model
    - Bayes nets
    - Regression
    - Discriminant Analysis
  - Opaque model
    - Neural networks
    - SVM
- Model-free
  - Supervised Classifiers
    - Decision trees
    - Ensembles of trees
  - Unsupervised
    - Clustering
    - Association Rules
    - Sequential Patterns
    - Principal Components
    - K-Nearest Neighbor

What are we doing?

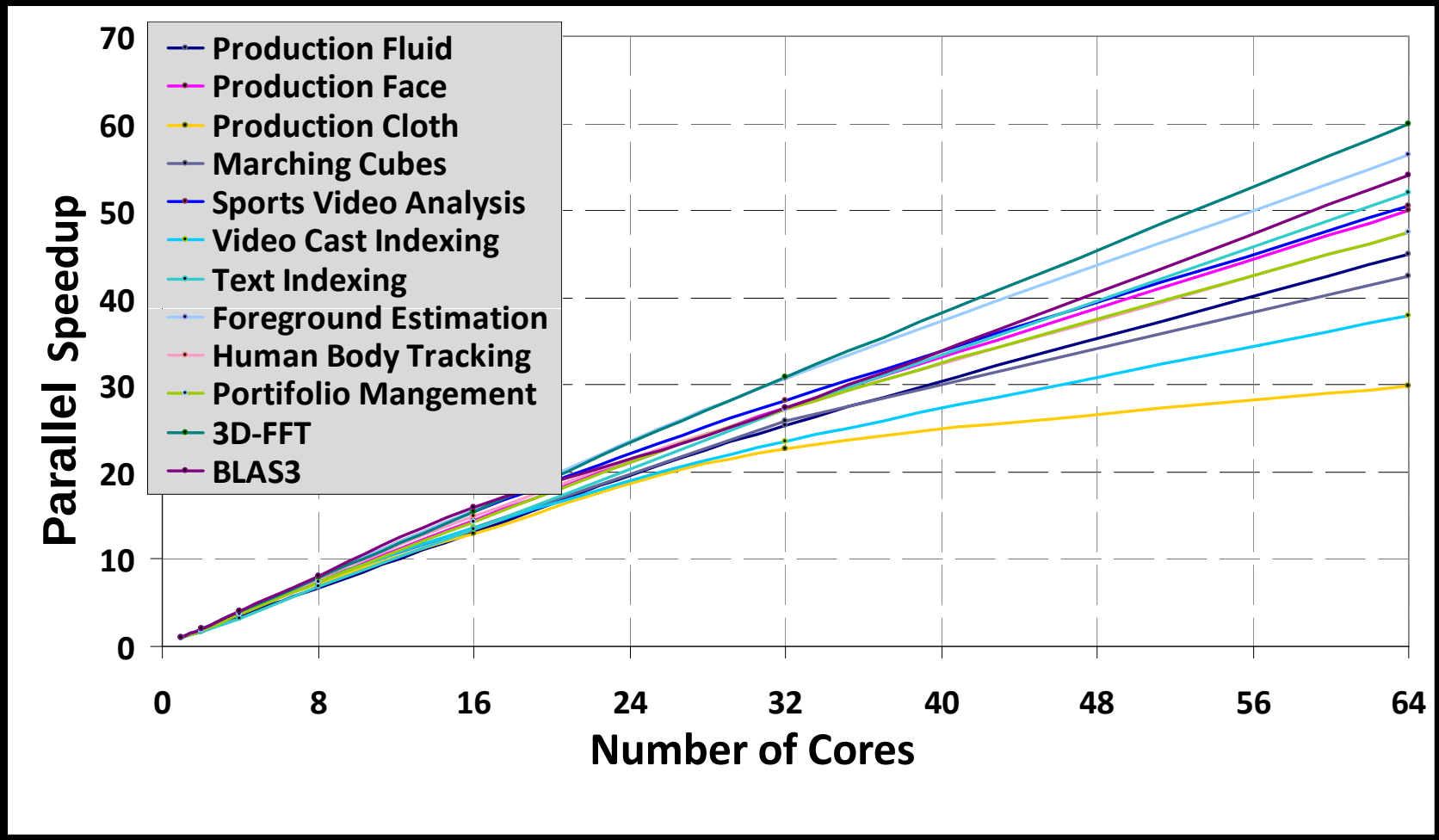


Level 1: Applications				
Ad-hoc search	...	Derivative Pricing	Ray-Tracing	Computer Vision
Semantic Search	Portfolio Selection	...	Physical Simulation	...
Level 2: Mathematical Models				
Partitioning Based	...	Diffusion Models	Level Sets	Tracking &Reconstr.
Generative non-linear	Quadratic Optimization	...	Particle Systems	...
Level 3: Mathematical Techniques				
SVD	...	Interior-Point Method	Collision Detection	Path Planning
K-means	Stochastic Simulation	...	Filtering&Anti-Aliasing	...
Level 4.1: Numerical Algorithms				
Direct Solvers	Iterative Solvers	Monte Carlo Simulation	Convex Collision (V-Clip, GJK)	
Level 4.2: Numerical Primitives and Data Structures				
Sparse BLAS123		Dense BLAS123	Structured matrix operat.	
Sparcity struct. (CRS, graphs, elimination tree		Basic geometry primitives (triangle, box, convex)	Partition structures (grids, kd-tree, BVH)	

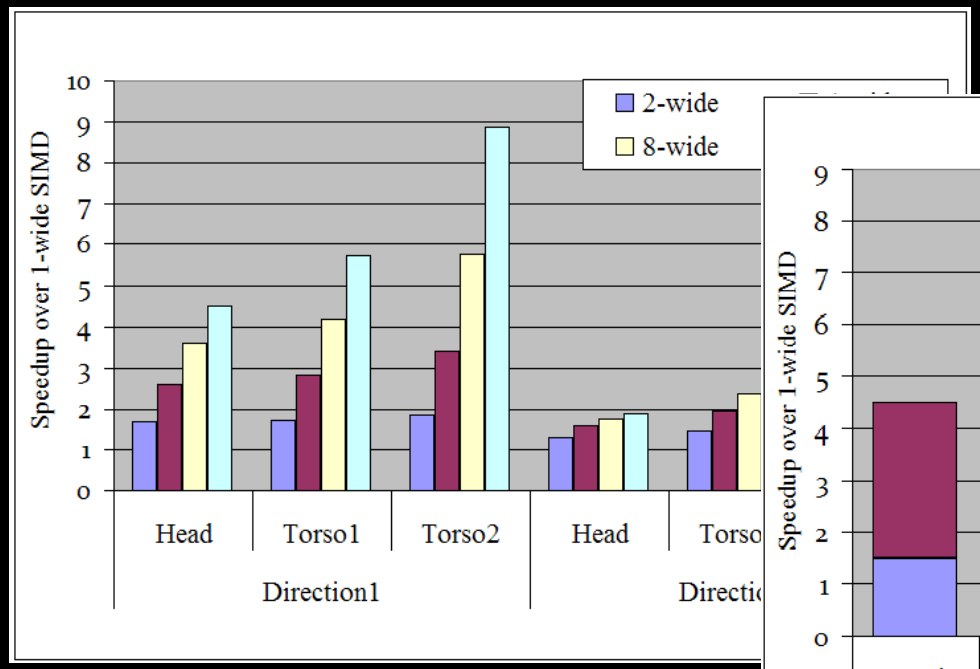
Workload-driven Architecture Research

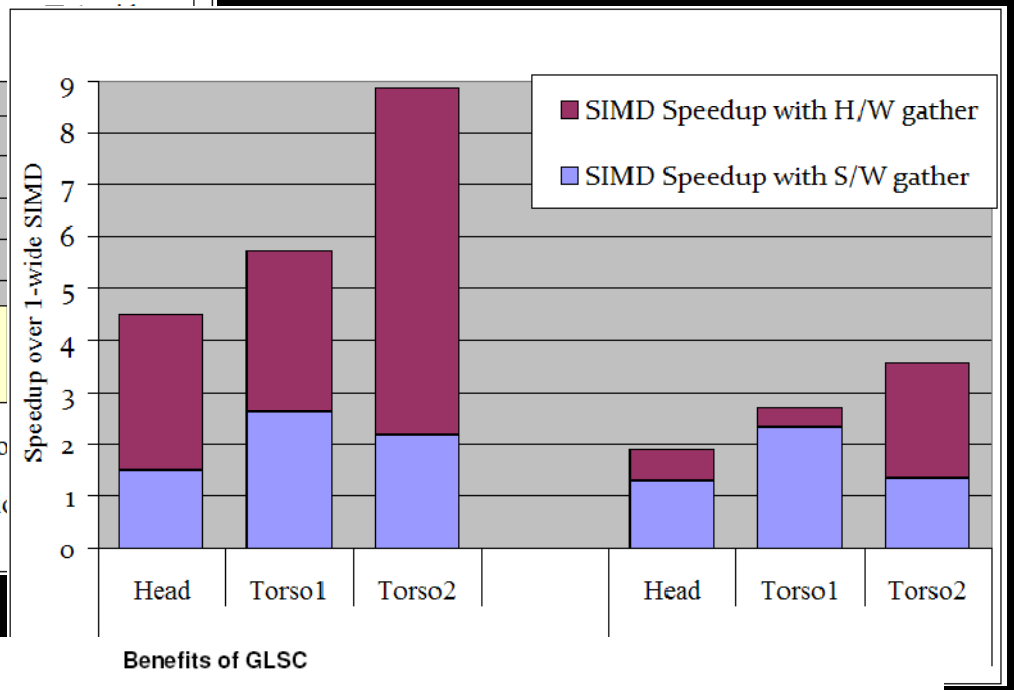
SAAR (Scalable Applications and Architecture Research)

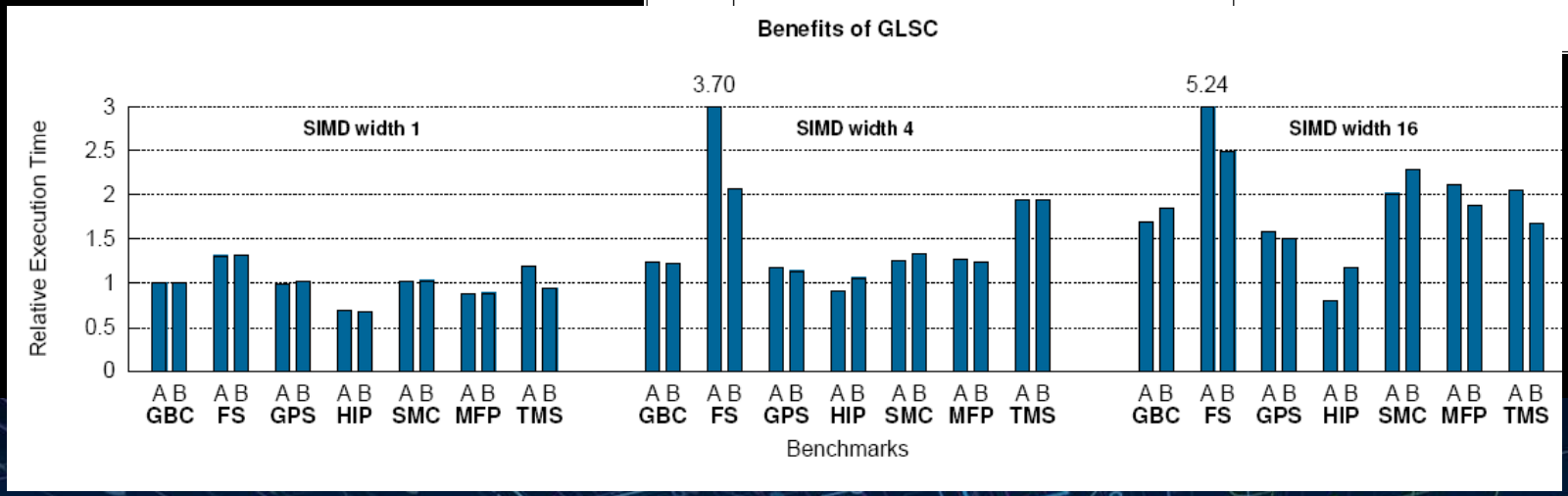
Scaling With Cores



Architectural Smarts

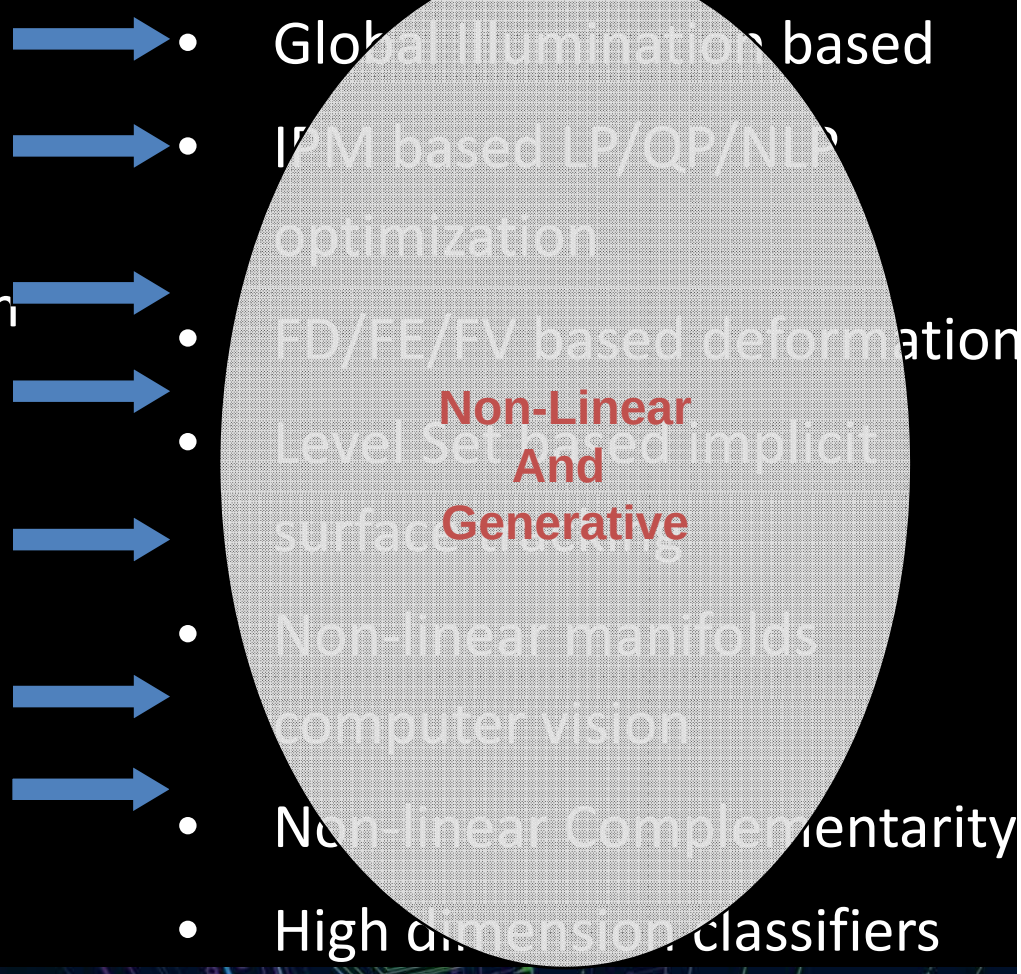






RMS Computing Core: Algorithmic Evolution

- Map-based shading
- SIMPLEX based linear optimization
- Mass-Spring based deformation
- Marker-based explicit surface tracking
- Linear manifold based recognition/modeling
- Linear Complementarity problems
- Low dimension classifiers



- Global Illumination based
- IPM based LP/QP/NLP optimization
- FD/FE/FV based deformation
- Level Set based implicit surface tracking
- Non-linear manifolds computer vision
- Non-linear Complementarity
- High dimension classifiers

Summary

- *Connected Computing*
  - It's all about three C's (above + content or data)
- Architectural Challenge
  - Moving the data real-time to where compute happens
- Algorithmic Opportunity
  - Massive data approach to traditional compute problems

BACKUP