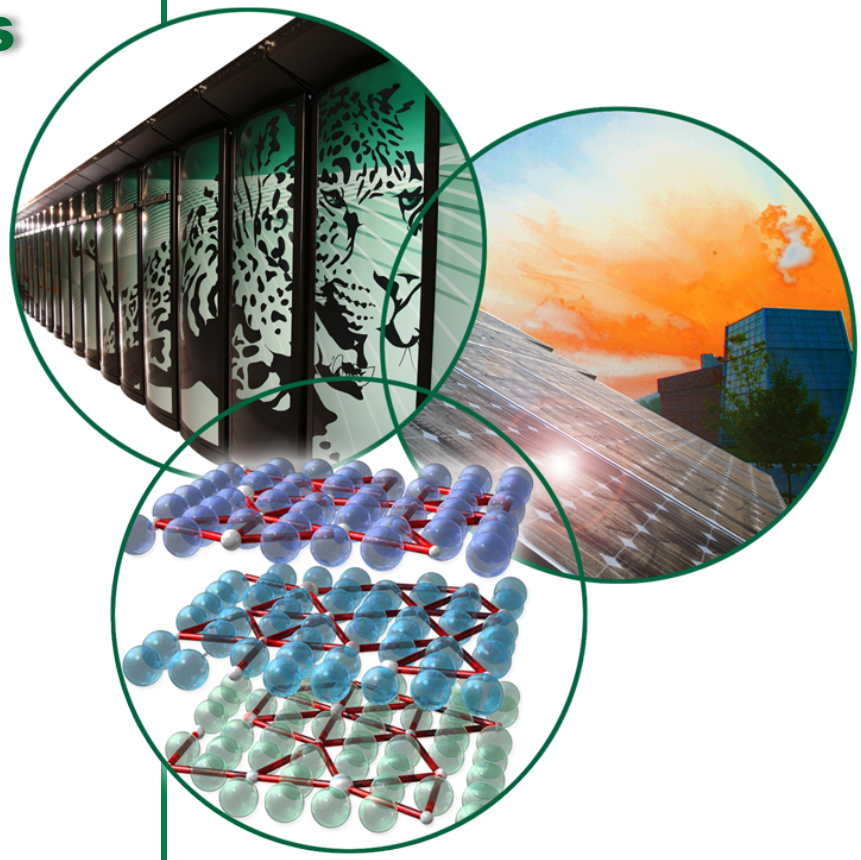
Scale+Fidelity+Speed+Integration = Necessary+Possible in Big Network Simulations

September 11, 2012



Kalyan S. Perumalla
Oak Ridge National Laboratory

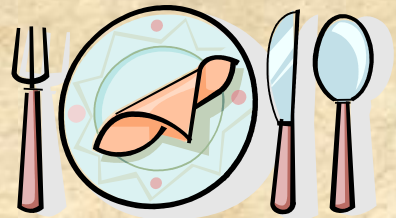
COMBINE Workshop

Washington, DC

U.S. DEPARTMENT OF ENERGY

OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

COMBINE Restaurant – A La Carte Menu



Scale*

Comes in filling portions
Specify **small**, **medium** or **large**

Fidelity*

Made with specialty ingredients
Specify **low**, **medium** or **high**

Speed

Recommended for to-go orders
Specify **slow**, **medium** or **fast**

Integration

Popular dish for large parties
Specify **ad-hoc** or **standard**

Note

- Limit any **two** items per customer
- Three or more available by **special order** only

* Please inform us if you are a pure theorist allergic to scale or fidelity ingredients

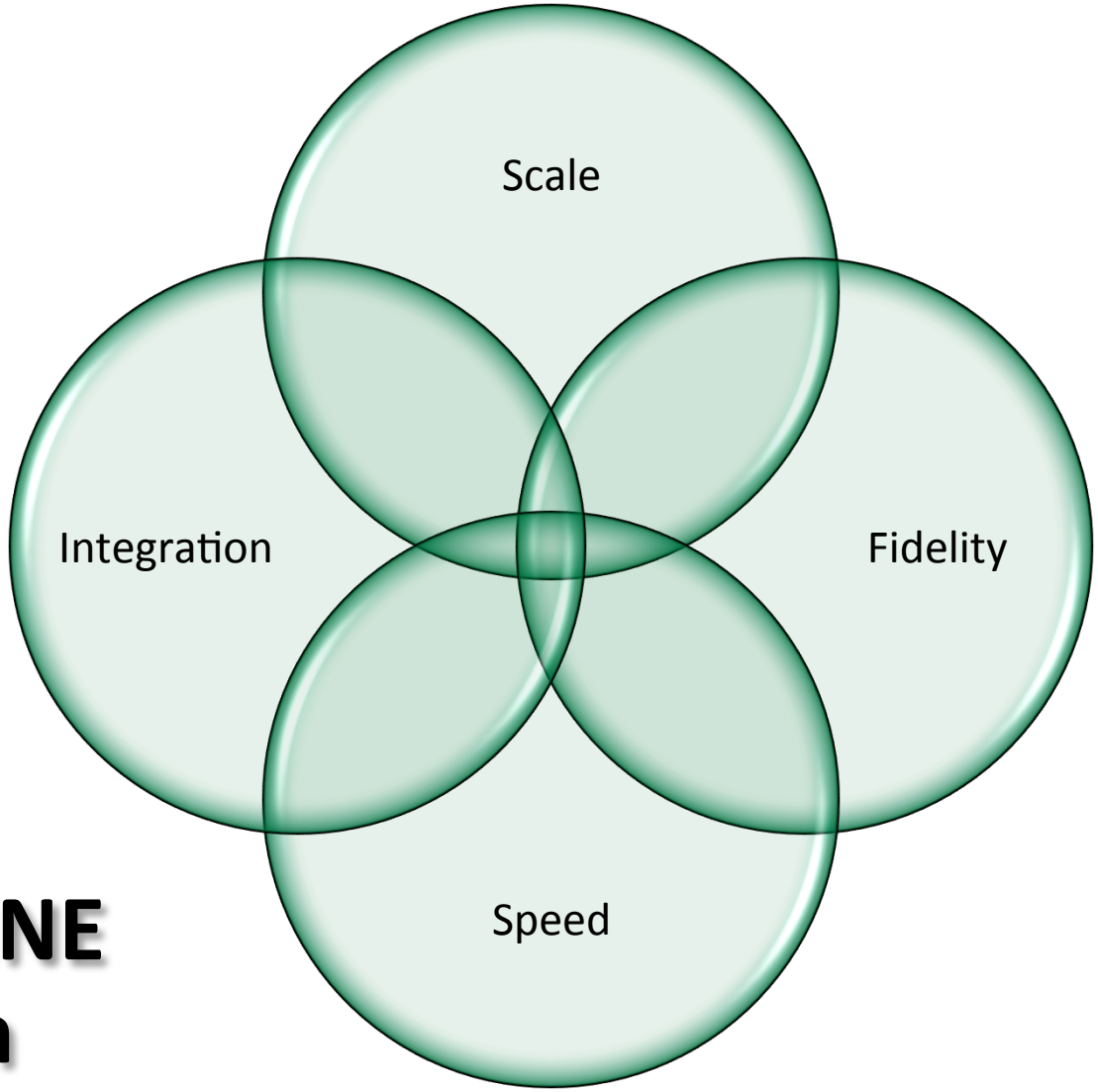
Speciality chefs available for one-of-a-kind events

OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

Lighter vein aside…

- These are independent factors

- Need high-end on all four *simultaneously* to meet COMBINE goals and vision



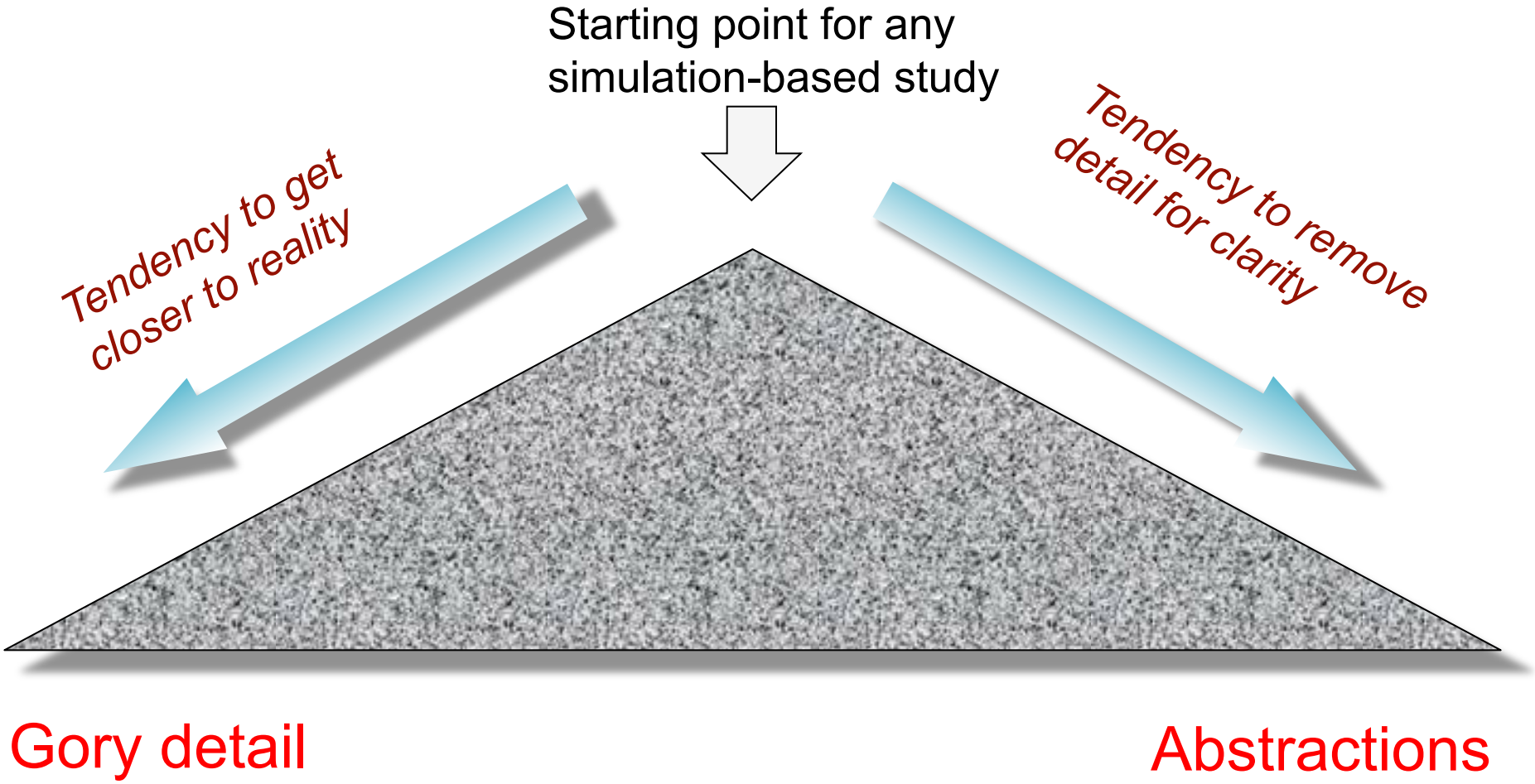
OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

Science and Technology Components

	Scale	Fidelity	Speed	Integration
Network Science	#Flows #Nodes #Packets #Apps Time horizon Global interconnects	Layers State Machines Interference Behavior network Overlays Authentication Malware	Scenarios/sec Parameter sweep Performance Design Provisioning Uncertainty Quant.	Traffic modeling Cyber security Social networks Vehicular networks Video models
Simulation Technology	Logical Processes Distributed memory Parallel synchronization	Virtual Machines Emulation Code Grafting Packets & Fluids	Discrete Event Supercomputing Multi-scale models	Multiscale interface Data-driven sims Standards Repositories Components

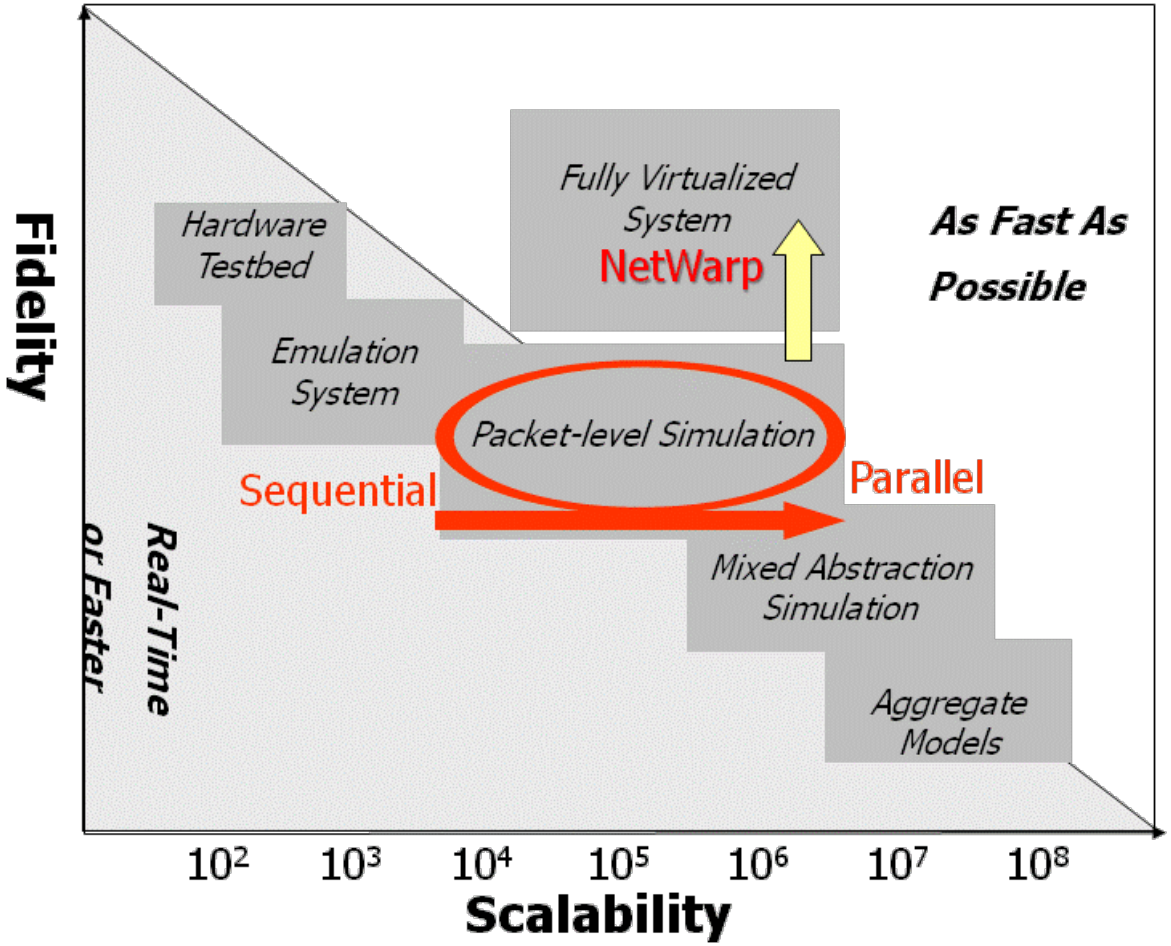
OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

It's very easy to slide down slippery slopes



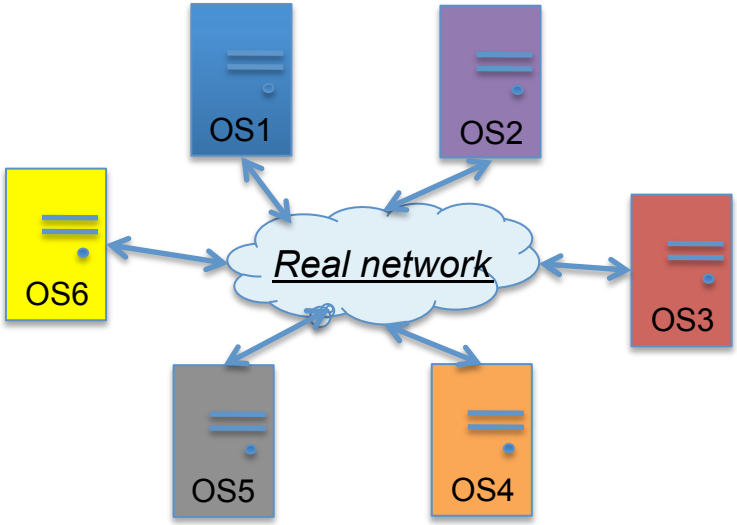
OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

NetWarp as a Fully Virtualized System*: Optimum of scalability vs. fidelity

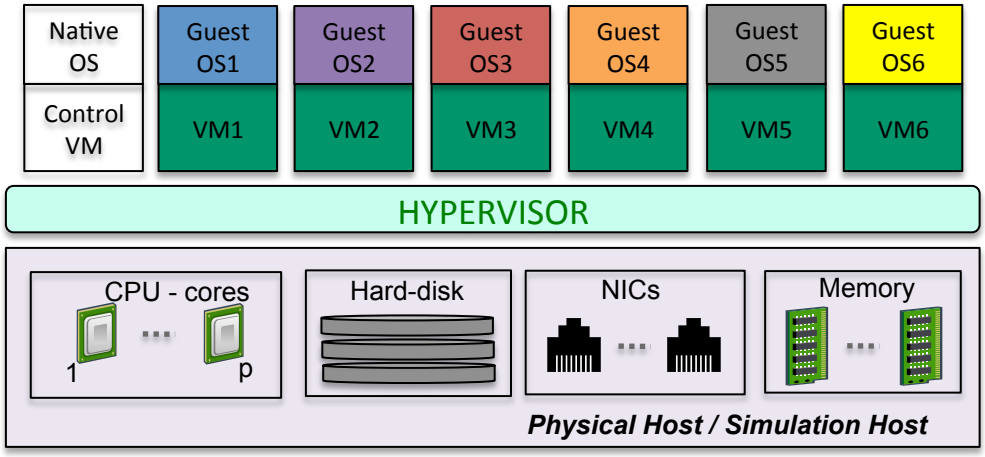


OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

NetWarp: Virtualization Applied to Network Emulations/Simulations



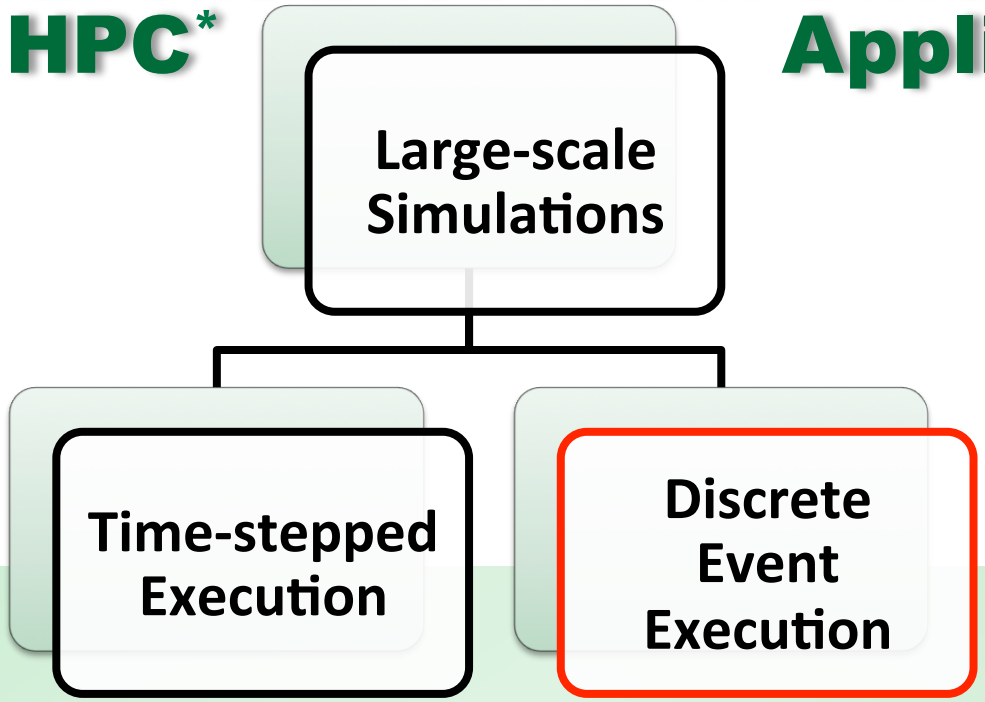
Real physical network to be simulated



Simulation/Emulation realization using VMs

OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

Discrete Event Execution: Different Class of HPC* Applications



Most-studied

- Coupled PDEs
- Implicit/explicit solutions
- Linear algebra, solvers
- Root-finding
- Communication computation overlap
- Asynchronous collectives
- Bulk synchrony
- Checkpoint-restart, …

Less-studied

- Fine-grained events
- Timestamp order
- Global Virtual Time computation
- Time-varying dependencies
- Lookahead extraction
- Primary/Secondary rollback
- Reverse computation, State-saving
- Fossil collection, …

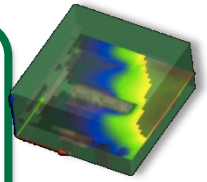
*HPC=High Performance Computing

OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

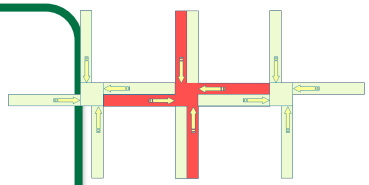
μsik Massively Parallel Discrete Event Engine

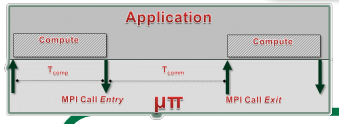
RCREDIF
World-scale Epidemics Simulator

TLM
Radio Signal Strength Simulator



SCATTER
Parallel Transportation Simulator





μπ
Exa-scale System Simulator

μsik
Discrete Event Engine

RCPHOLD
Standard Synthetic Benchmark

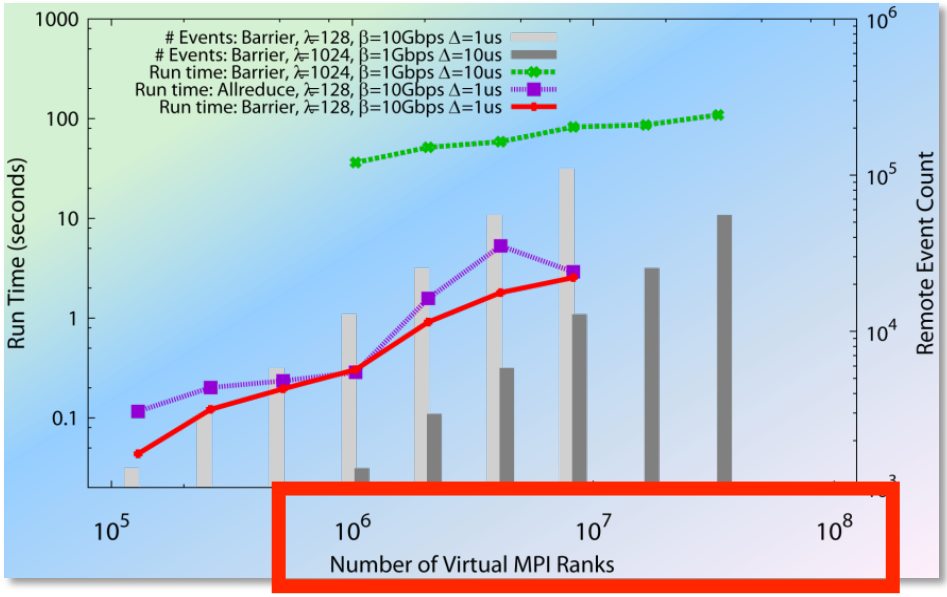
OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

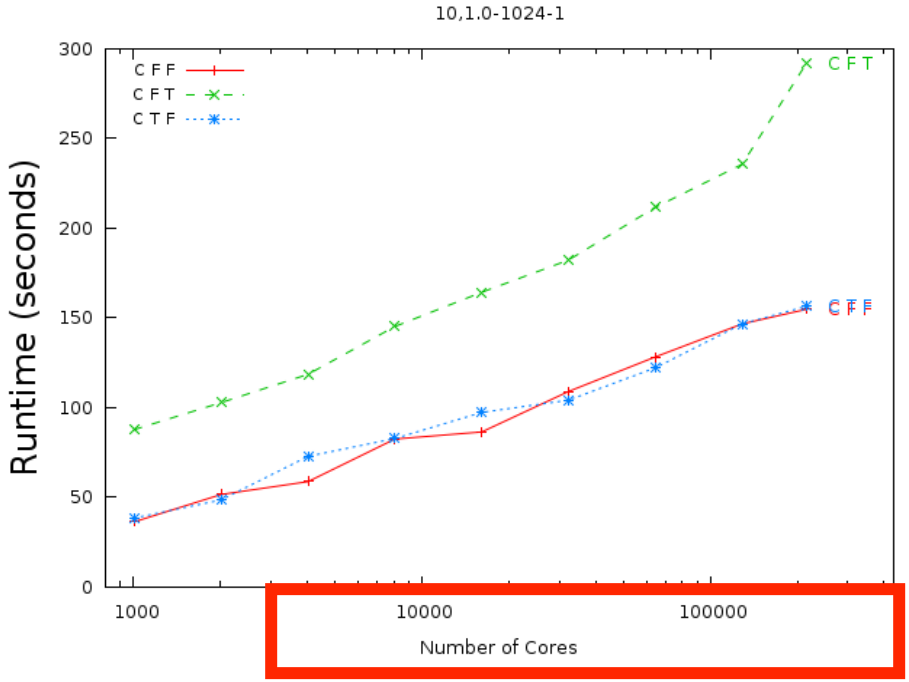
μπ Billion-way Concurrency Simulations over μsik

Scales to 216,000+ cores

0.25 billion control flows

Process-oriented flows



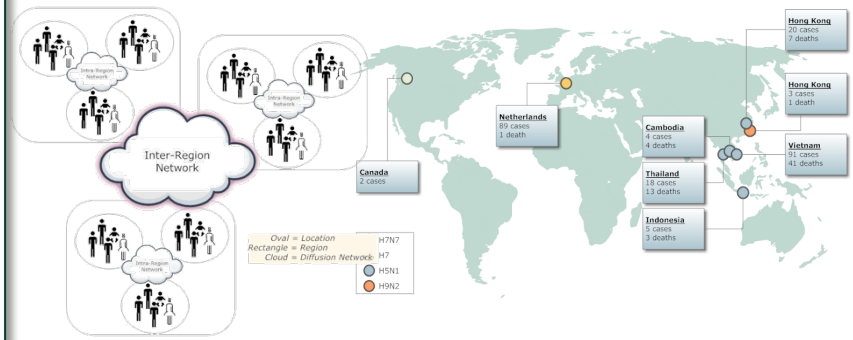


- **Articles in SimuTools'10, MASCOTS'11**
- **Applications tested over μπ**
  Tests: mpiping, allreduce, matmul, …
  Benchmarks: FFTE, GUPS, …
  Applications: Sweep3D, NWChem, …
  DOE Mini-Apps: pHPCG, CGPOP, …

OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

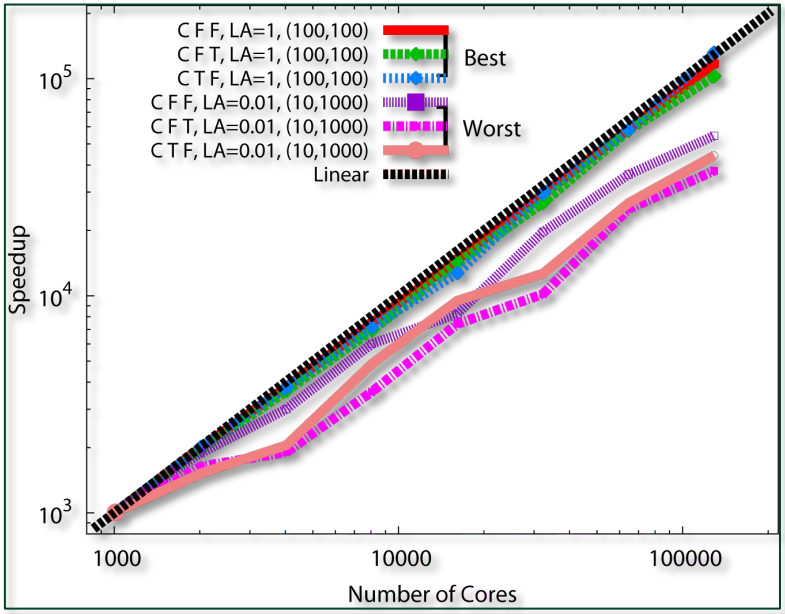
Billion-Individual Reaction-Diffusion Epidemiology Model over μsik

- Hierarchical network
- Multi-scale behavior
- Heterogeneous attributes
- Parameterized to specify many scenarios of interest



$$p_i = 1 - e^{t \sum_{r \in R} N_r ln(1 - rs_i \rho)}$$

System	Regions	Locations
Community	10	100
Small Town	100	100
Small City	1000	100
Large City	1000	1000
Rural State	10000	100



Articles in PADS'10, SIMULATION'12, HIPC'11

OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

Scale, Speed, Fidelity Possible

- **Very large-scale discrete event applications are indeed possible to execute efficiently**
- **Billions of events/sec implies large scale scenarios can be run in very short time. Examples:**
  - Large computer network simulations driven by complex behaviors
  - World-scale , individual-level epidemiological simulations
  - Transportation simulations with millions of vehicles, road intersections
- **Work is ongoing to further optimize, and test on larger number of processors, and newer architectures**
  - Cray XK6 (Titan) at ORNL, Blue Gene Q, …

OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

Issues and Challenges

A basic challenge

- Physical laws invariant during modeling and simulation effort

  vs.

  Network structure and behavior is extremely fast changing

What is the objective?

- Resilience?
- Security?
- Availability?
- Performance?
- Fairness?

Also, need

- **Cool graphics for PR**
  - Counterparts of images in astrophysics
- **Expectation management**
  - Reduced hype (contrast to "cyber range")

OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

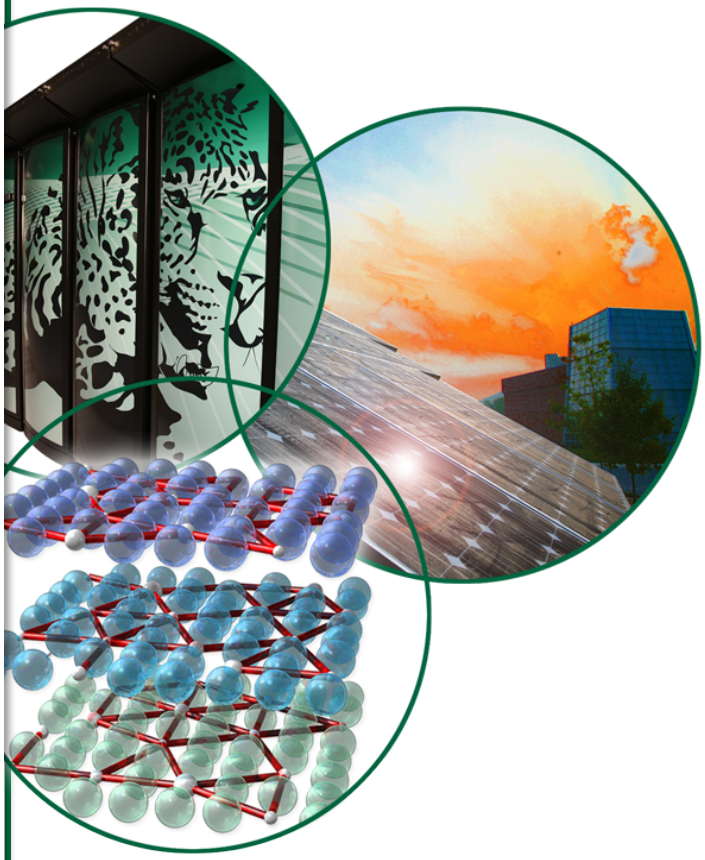
Fundamental Challenges

- **In the limit:**
  - The actual network is the only model of itself
- **Most frustrating:**
  - "We" don't control "the network"

OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

Thank you!
Questions/Comments?

Kalyan S. Perumalla
Senior R&D Staff & Manager, ORNL
Adjunct Professor, Georgia Tech

perumallaks@ornl.gov
www.ornl.gov/~2ip
+1 (865) 241-1315
+1 (865) 576-0003 (fax)

Oak Ridge National Laboratory
PO Box 2008, MS-6085
Oak Ridge, TN 37831-6085