

Motivation

Addressing exascale challenges:

- ▶ Resiliency, power, memory access, concurrency and heterogeneous hardware.
- ▶ Codes/algorithms need to be flexible and extensible for unknown architectures.
- ▶ Errors and faults will become more frequent than on current platforms.

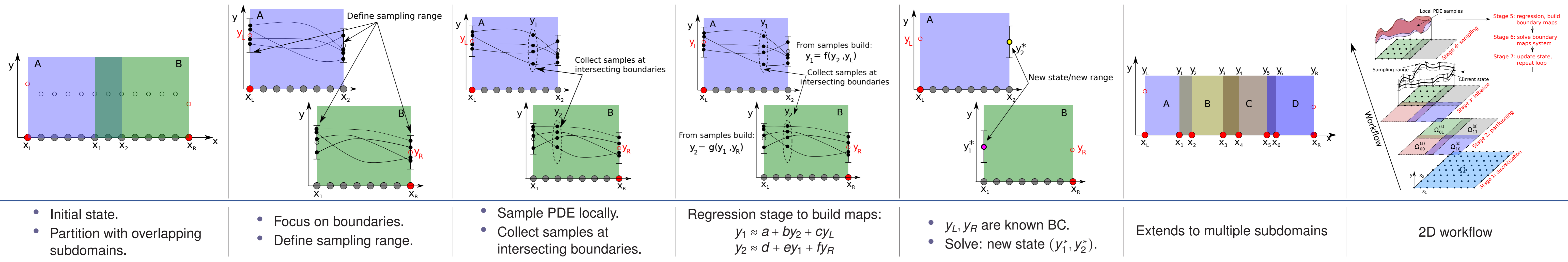
Objectives

- ▶ Well-defined formulations even in the presence of system faults.
- ▶ Assess predictive fidelity of extreme-scale scientific simulations.
- ▶ High performance on “uncertain” architectures.
- ▶ Portable, extensible and reusable codes.

Technical Approach

- ▶ What?
 - ▶ Resilient preconditioner for PDEs.
 - ▶ To achieve resiliency to both soft and hard faults.
- ▶ How?
 - ▶ Recast the original PDE as a sampling problem.
 - ▶ Focus solely on the information available.
 - ▶ State update through resilient data manipulation.

Resilient EXtreme Scale Scientific Simulation (REXSSS) Algorithm Overview

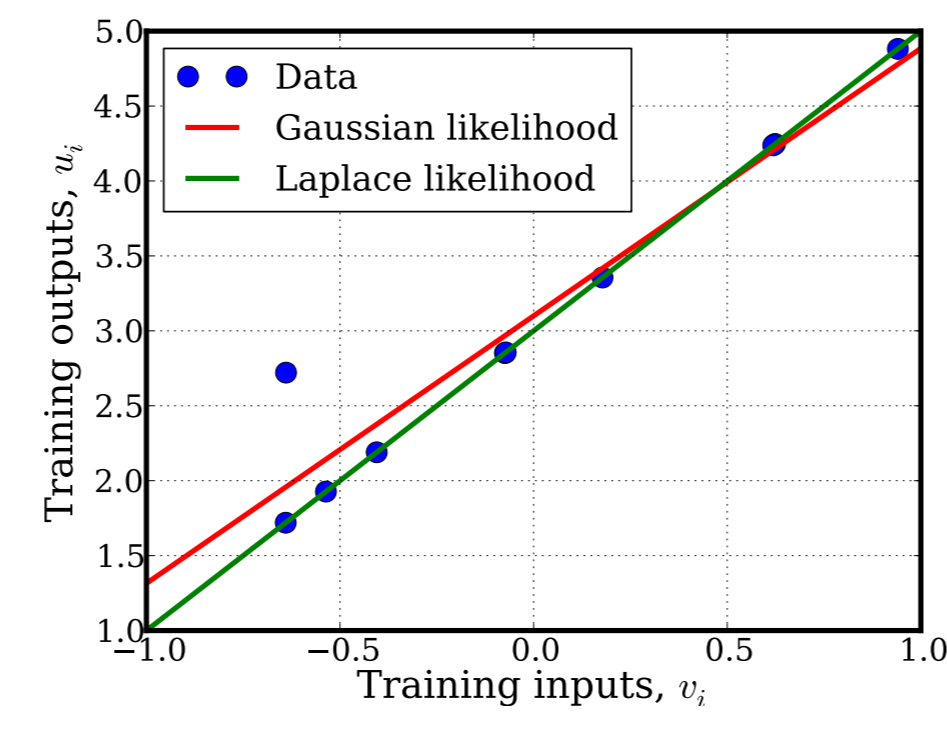


- Initial state.
- Partition with overlapping subdomains.
- Focus on boundaries.
- Define sampling range.
- Sample PDE locally.
- Collect samples at intersecting boundaries.
- Regression stage to build maps:
 - $y_1 \approx a + by_2 + cy_L$
 - $y_2 \approx d + ey_1 + fy_R$
- y_L, y_R are known BC.
- Solve: new state (y_1^*, y_2^*) .
- Extends to multiple subdomains
- 2D workflow

Regression stage: ℓ_1 vs ℓ_2

Regression Resilience

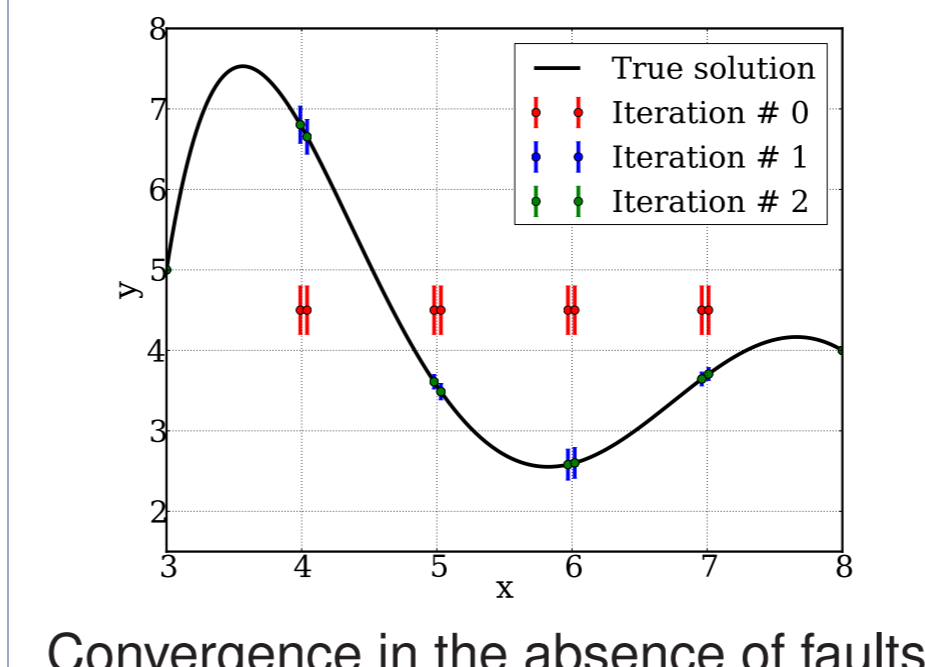
- Resilience to soft faults.
- Soft faults show up as corrupted data.
- Bayesian approach uses available information, incorporating uncertainties associated with missing or faulty data in a natural way.
- ℓ_2 : “balance” of all residual errors.
- ℓ_1 : solution with fewest possible non-zero residuals.



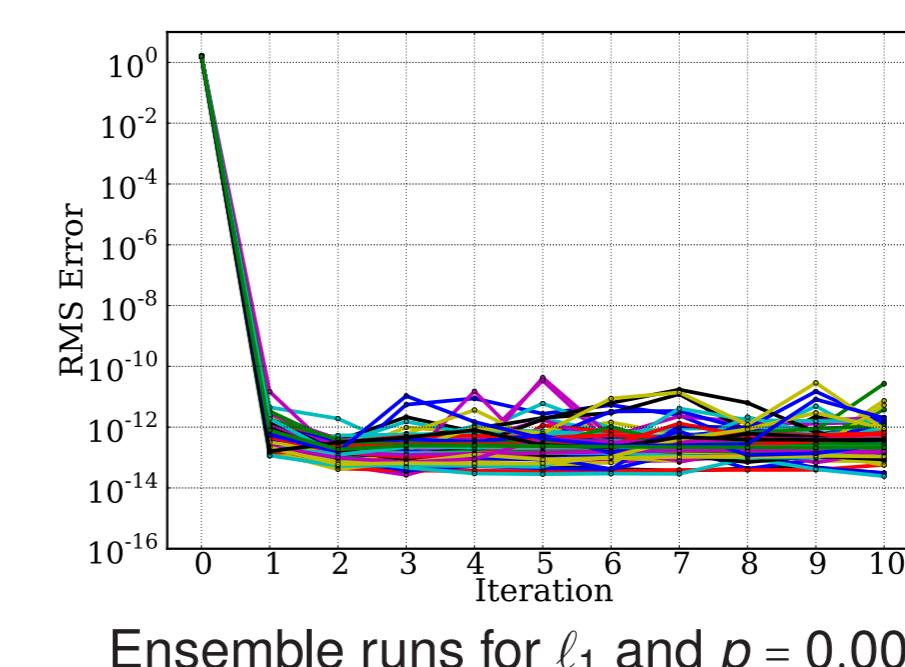
Regression test for single perturbed point.

1D Elliptic PDE Test

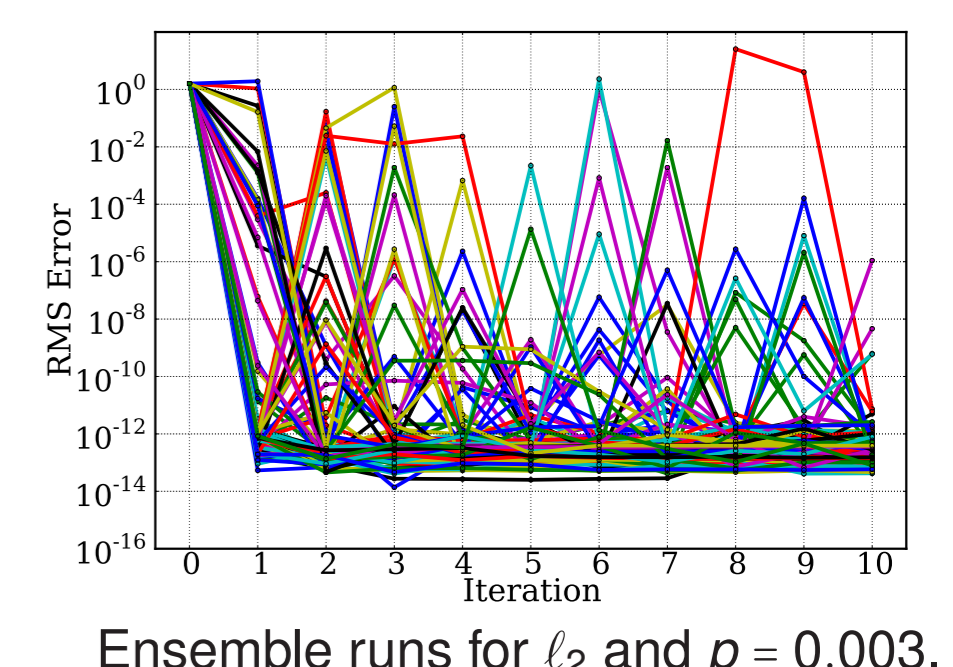
- Bit-flips are inserted during sampling with probability p .
- Laplace likelihood yields resiliency to soft faults.



Convergence in the absence of faults.



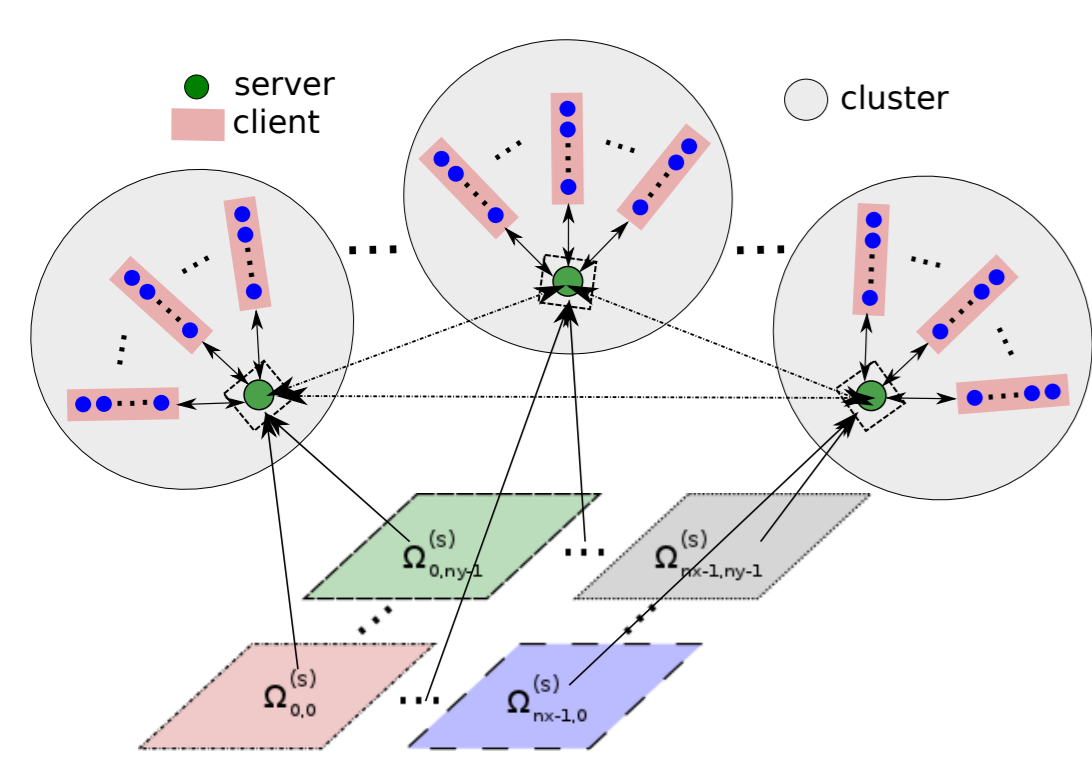
Ensemble runs for ℓ_1 and $p = 0.003$.



Ensemble runs for ℓ_2 and $p = 0.003$.

REXSSS Implementation

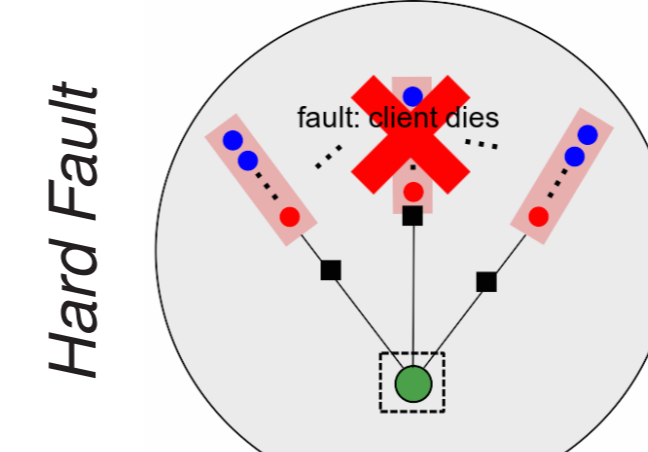
Server-Client (SC)



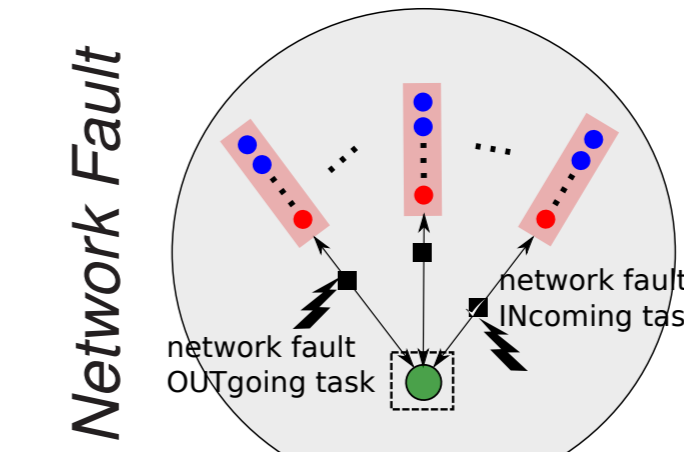
- Cluster: 1 server + n clients.
- Servers:
 - Communicate between each other.
 - Safe data/state storage (sandboxed).
- Clients:
 - Independent from one another.
 - Only serve as computing units.

- Separates state from computation.
- Confining the data to the server reduces the overall vulnerability.
- Resilient to clients crashing because even if tasks are lost, state is safe.
- It aligns with the vision of future exascale architectures involving heterogeneous and hierarchical hardware required to meet energy and cost constraints.

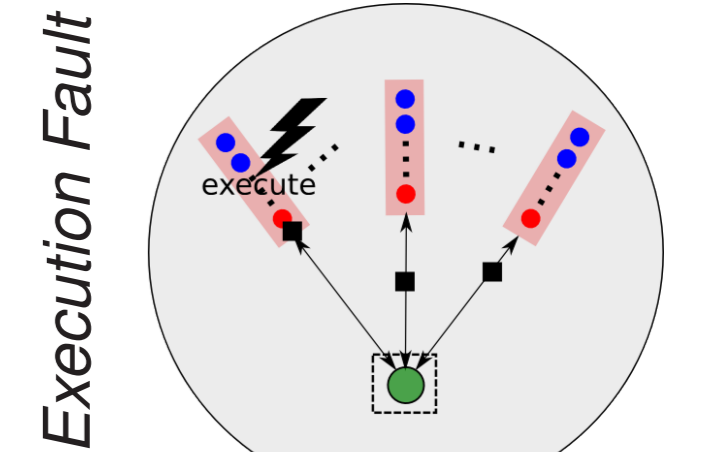
Currently Supported Fault Models



What? Client dies, lost task.
How? Client ranks set idle.



What? Data corruption during communication.
How? Random bit-flips on task's data.



What? Data corruption during computation.
How? Random bit-flips on task's data.

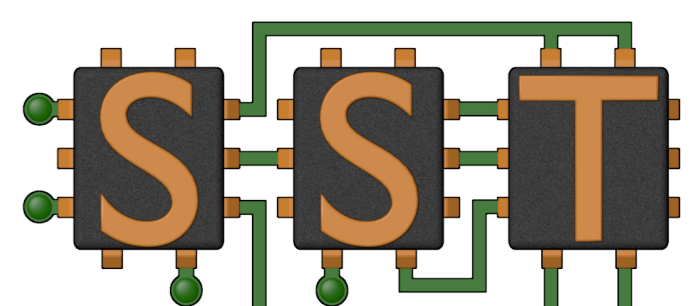
SST/macro Skeleton

SST/macro¹ overview:

- Coarse-grained simulator of distributed-memory applications.
- Uses machine models to estimate performance of processing and network components.

SST/macro skeleton of REXSSS:

- Skeleton of the SC implementation of the algorithm.
- Only need to reproduce basic workflow.



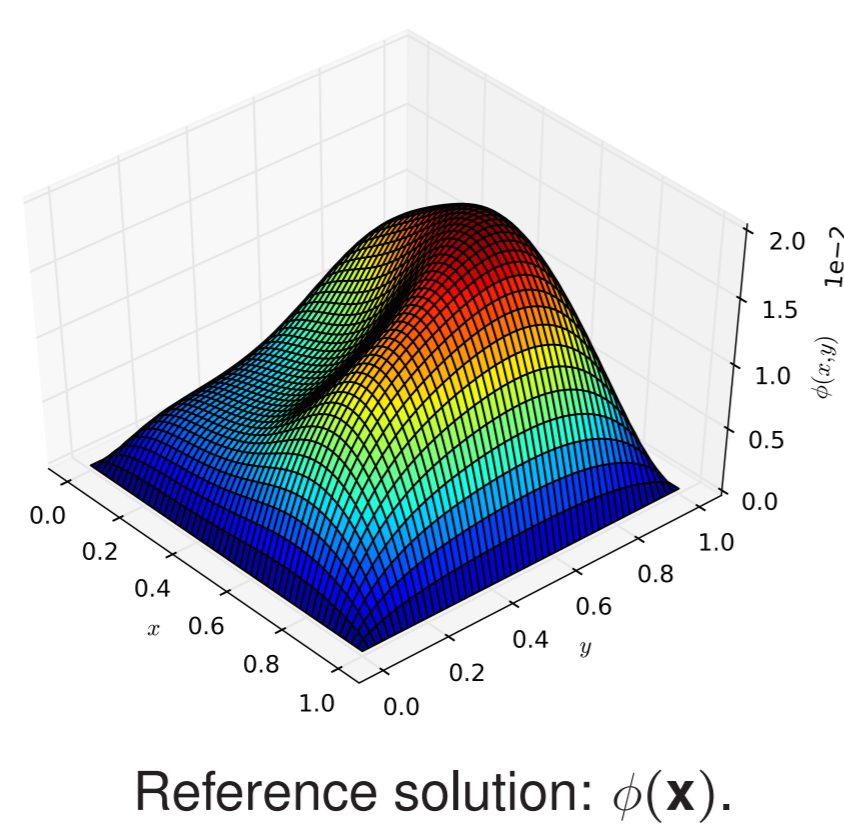
[1] <https://sst-simulator.org/>

Scalability

Test Problem

- Linear elliptic equation:

$$\nabla \cdot (k(\mathbf{x}) \nabla \phi(\mathbf{x})) = g(\mathbf{x})$$
- Unit square $(0, 1)^2$.
- Homogeneous Dirichlet boundary conditions.
- $k(\mathbf{x}), g(\mathbf{x})$ Gaussian form.
- 2nd-order finite-differences.
- Boundary maps are linear.
- For linear PDE, no faults, algorithm converges in one iteration.



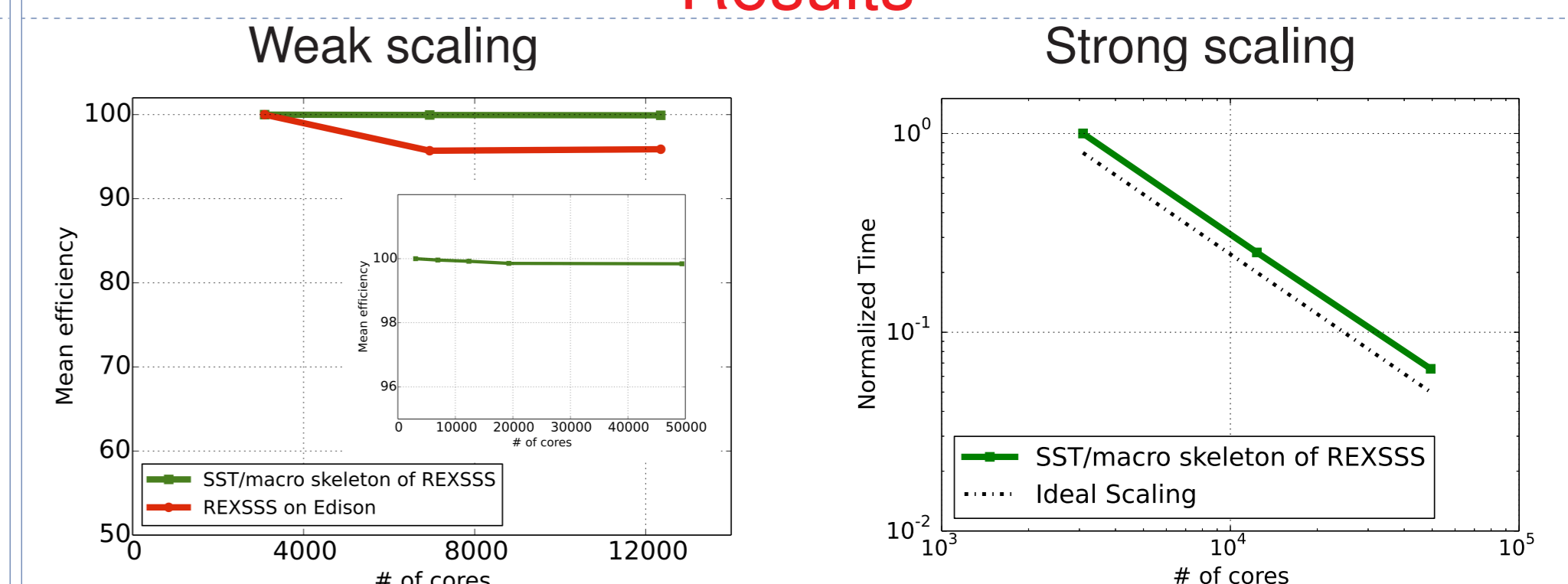
Runs

- Weak: number of subdomains per cluster is fixed.
- Strong: adding clusters for fixed problem size.

	Weak	Strong
Subdomains	12 ² , 18 ² , 24 ² , 30 ² , 48 ²	32 ²
Total Cores	3088, 6948, 12352, 19300, 49408	3088, 12352, 49408
Subdomain size	115 ²	
Servers	16, 36, 64, 100, 256	16, 64, 256
Subdomain Overlap (# grid cells)	12	12
Number of clients/server	48	48
Size of each client	4	4

- Focus on communication and computation intensive stages (sampling and regression) of the algorithm.

Results



- SST/macro skeleton:
 - Reproduces the REXSSS scaling behavior.
 - Allows extrapolation to larger configurations.
 - Will enable performance testing on future architectures.

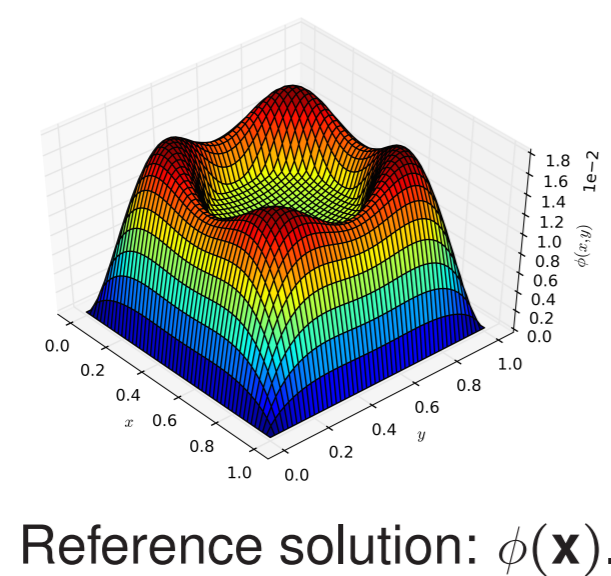
This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Resilience for 2D Elliptic PDE

Test Problem

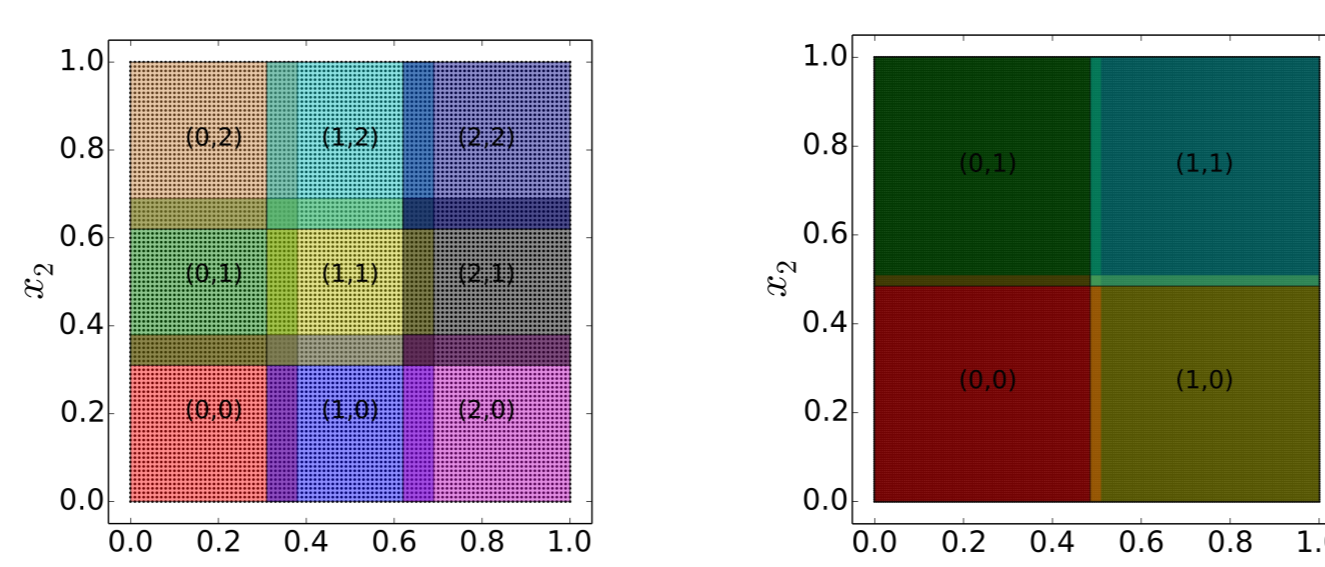
- Linear PDE:

$$\nabla \cdot (k(\mathbf{x}) \nabla \phi(\mathbf{x})) = g(\mathbf{x})$$
- Unit square $(0, 1)^2$.
- Homogeneous Dirichlet boundary conditions.
- $k(\mathbf{x}), g(\mathbf{x})$ hyperbolic tangent form.
- 2nd-order finite-differences.
- Boundary maps are linear.



Problem Size

- Explore the effect of problem size.
- 1 server, 62 clients of size 1.



$n_x = n_y = 101$
(tag=S9g_s37²)

$n_x = n_y = 201$
(tag=S4g_s103²)

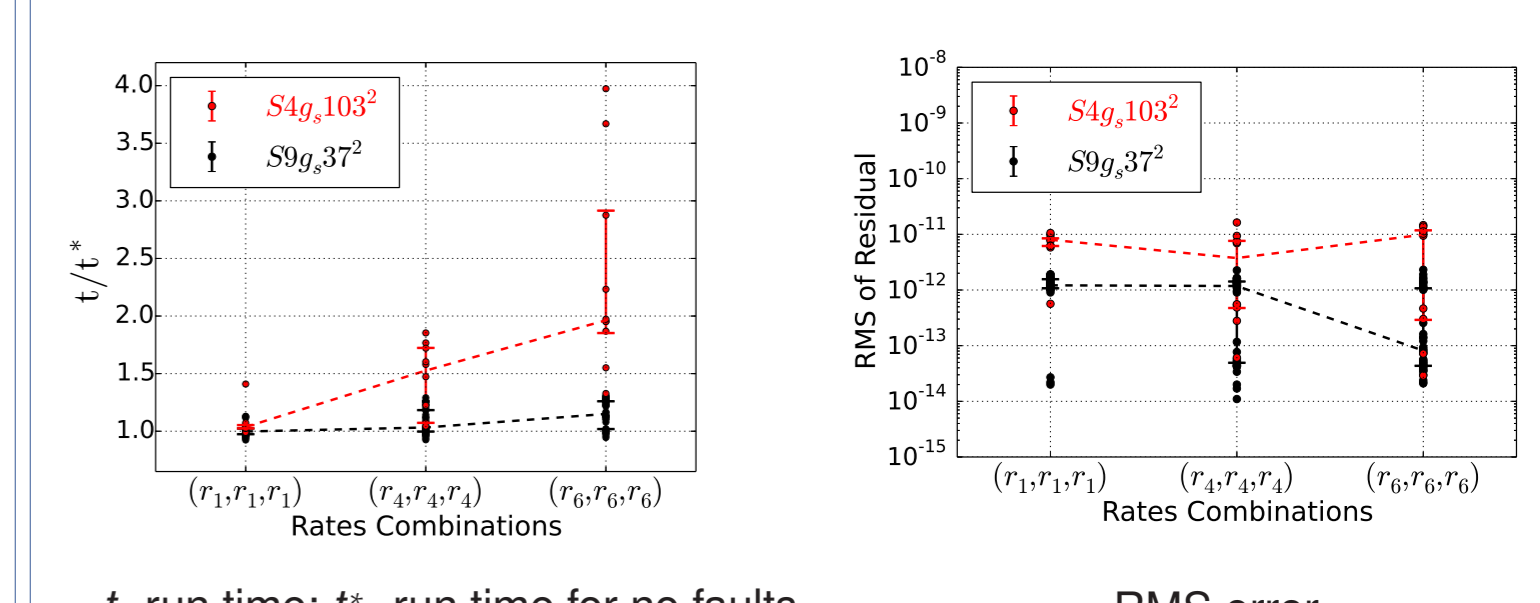
Failure Rates

Table: Failure Rates ($n_{\text{faults}}/\text{sec}$)

	Hard Faults	Computation Faults	Communication Faults
r_1	0.00005	0.00004	0.00069
r_4	0.00034	0.00035	0.00550
r_6	0.00090	0.00087	0.01400

- Faults modeled as a Poisson process.
- Test algorithm under extreme fault conditions.
- Faults extracted from literature by scaling up results assuming future architectures with 10^4 -way local concurrency and 10^5 nodes.

Results



t =run time; t^* =run time for no faults.

RMS error.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, under Award Number 13-016717.

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.