

Holistic Measurement Driven System Assessment (HMDSA)

Acknowledgements

Exascale-ready operations, performance, design and acquisition via comprehensive system monitoring and analysis

National Center for Supercomputing Applications and University of Illinois: Bill Kramer (PI)

Greg Bauer, Brett Bode, Mike Showerman, Jeremy Enos, Aaron Saxton, Saurabh Jha, Zbigniew Kalbarczyk, Ravi Iyer

Sandia National Laboratory: James Brandt, Ann Gentile

Institutional Collaborators: Boston University, New Mexico State University, University of Central Florida



Answering Key Operational Questions

Is my performance variation due to system conditions or code changes?

How can I know if the system is having problems?

User

Capability: Combined reasoning based on changes in application characteristics (e.g., MPI message size) and system characteristics (e.g., network congestion) to diagnose cause(s) of performance variation and predict application runtime.

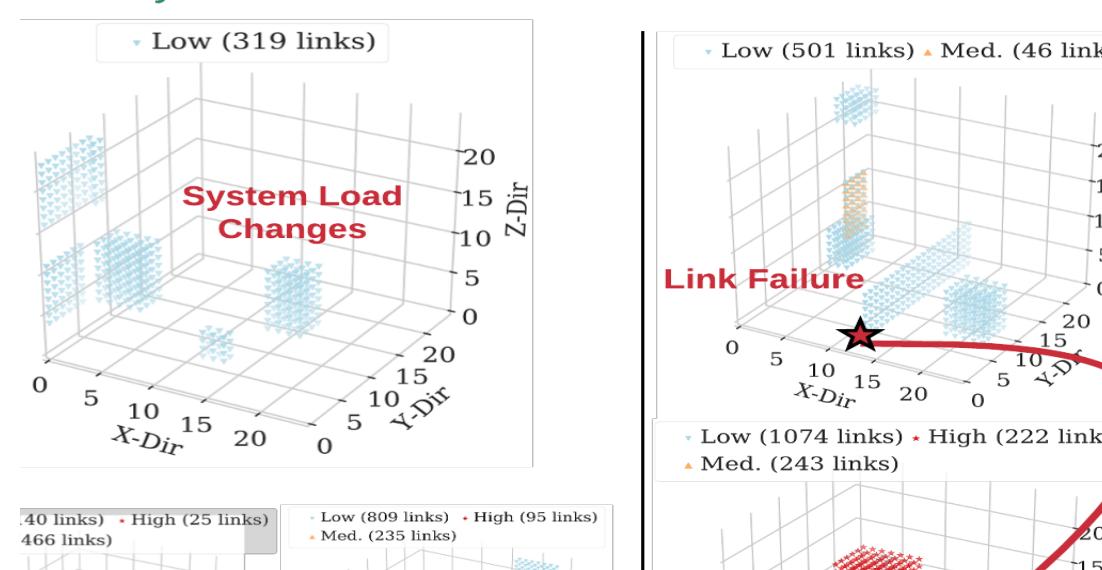
Figures of merit on Cray Aries interconnect system

Job ID	App ID	Node ID	Start Time	End Time	Treatment	HPages	NW Congest	Mem Score	Anomalies	PAPI Perf	App Perf	Comm Perf
10399	CoMD	nid000[53,59]	2018-09-21 00:28:40	2018-09-21 00:30:40	0	0.0	1.57	0.042	0x0	0.0	0.59	0.0
10398	Iulesh	nid000[60-61]	2018-09-21 00:19:26	2018-09-21 00:22:02	0	0.0	0.21	0.041	0x0	0.29	1.55	0.0
10397	CoMD	nid000[53,59]	2018-09-21 00:19:19	2018-09-21 00:21:28	0	0.0	0.06	0.042	0x0	0.1	0.59	0.0
10396	CoMD	nid000[21-22]	2018-09-21 00:19:16	2018-09-21 00:21:25	0	0.0	2.75	0.053	0x0	0.0	0.61	0.0

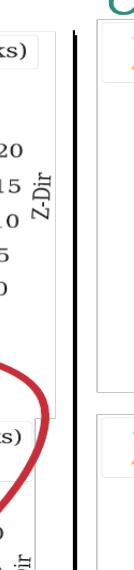
Different Cray Aries with competing application, different Aries, same Aries (self-contention)

Capability: Diagnose network issues using extracted congestion regions (e.g. Cray Gemini interconnect)

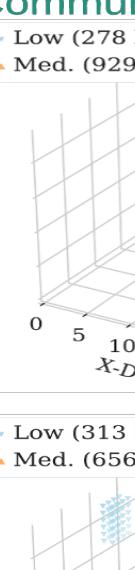
System Reasons



Failure Reasons



Communication Patt



System managers & users

Architects & acquisitions teams

Capability: Inform design decisions for future systems

Insight: Assessing dynamic range of memory needs for facilities workflows

What are the architectural requirements given the site's workload?

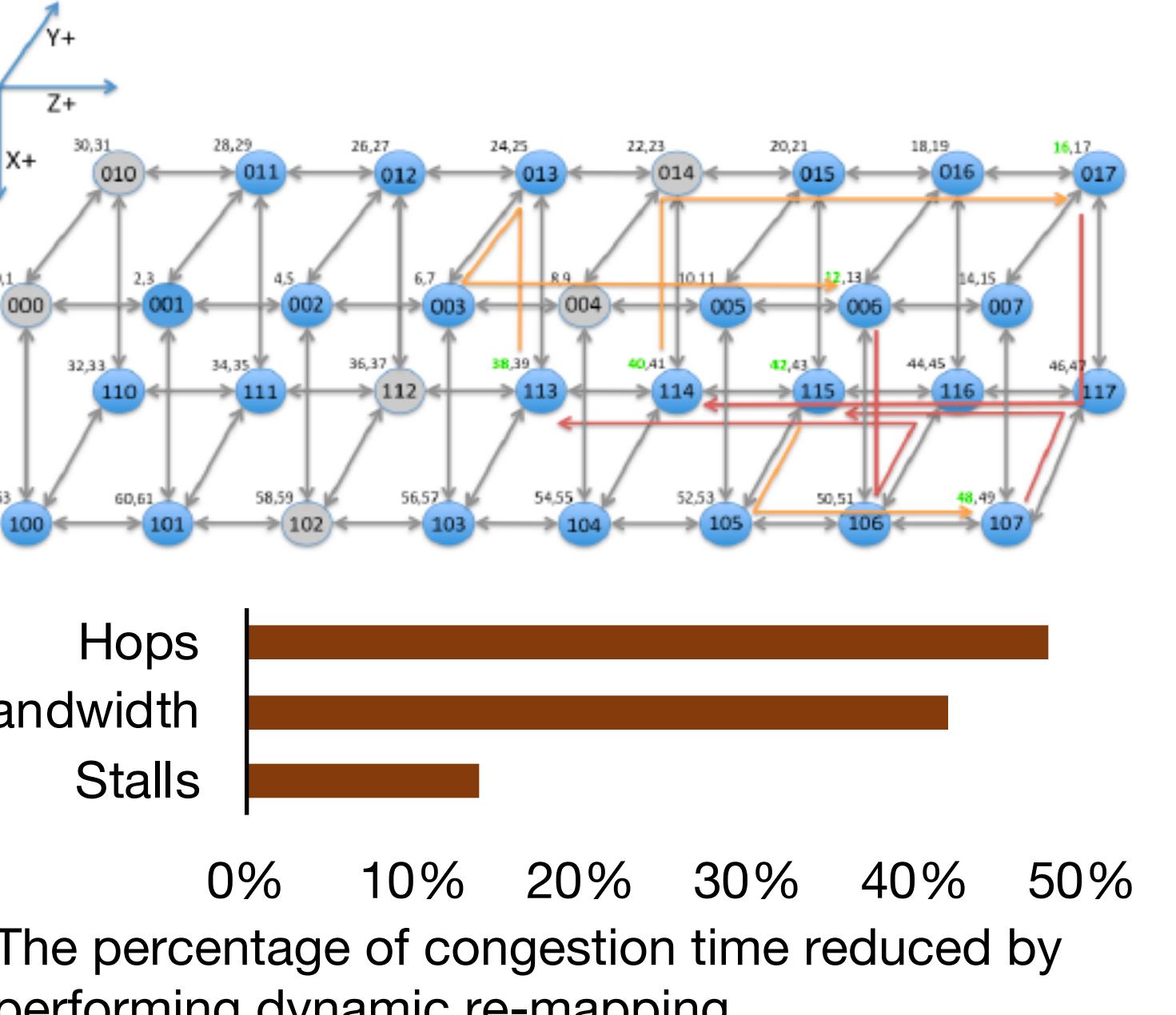
Architects & acquisitions teams

Capability: Dynamic task remapping based on congestion measurements

How can a system provide more effective and efficient services?

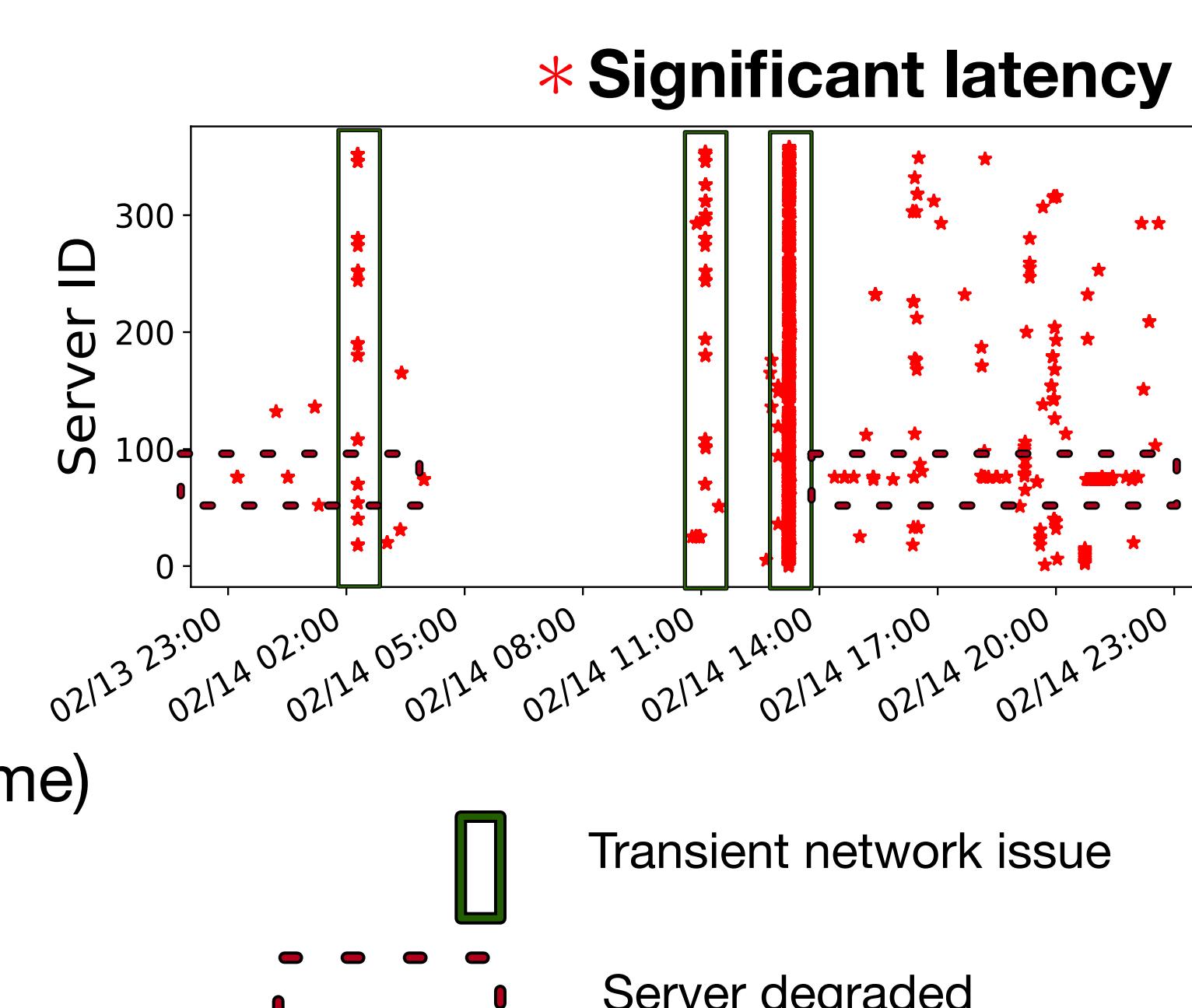
Architects, sys. managers & support

Capability: Dynamic task remapping based on congestion measurements



Capability: Diagnose file-system issues using periodic latency measurements of file system operations (e.g. write)

Insight: Identifying issues with adaptive scheduling and congestion mitigation procedures

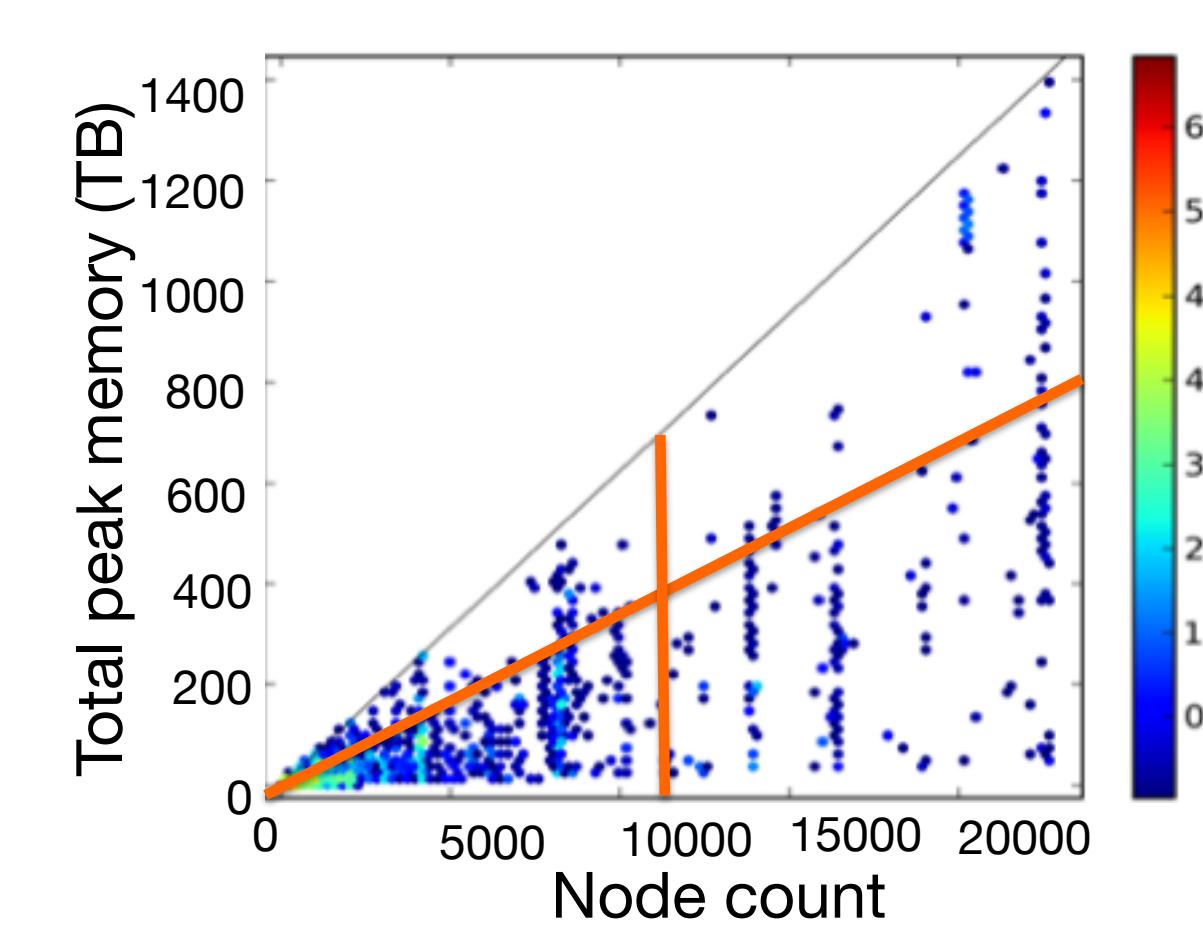


Capability: Diagnose code-related issues (e.g. impact of MPI message size on runtime)

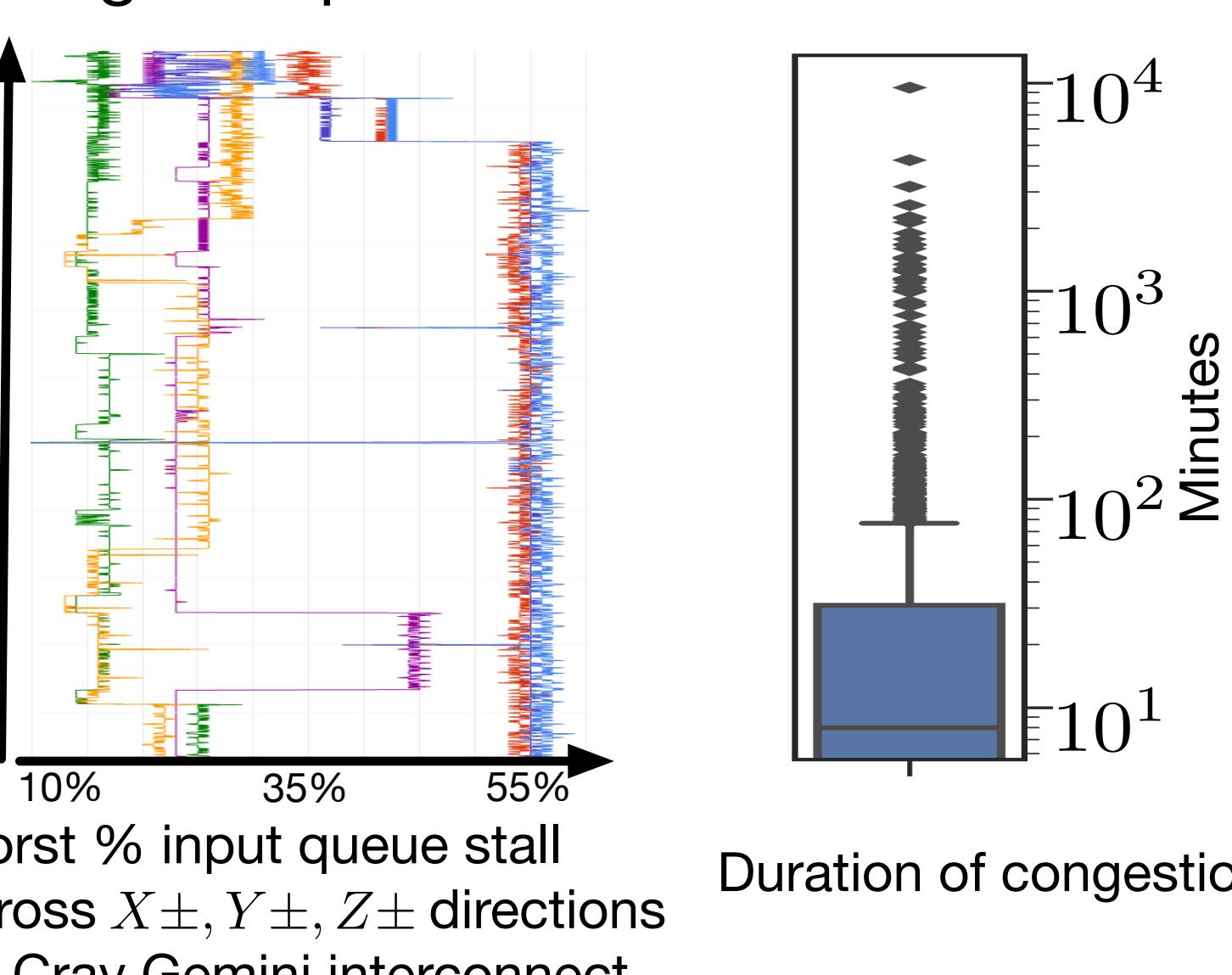


Transient network issue

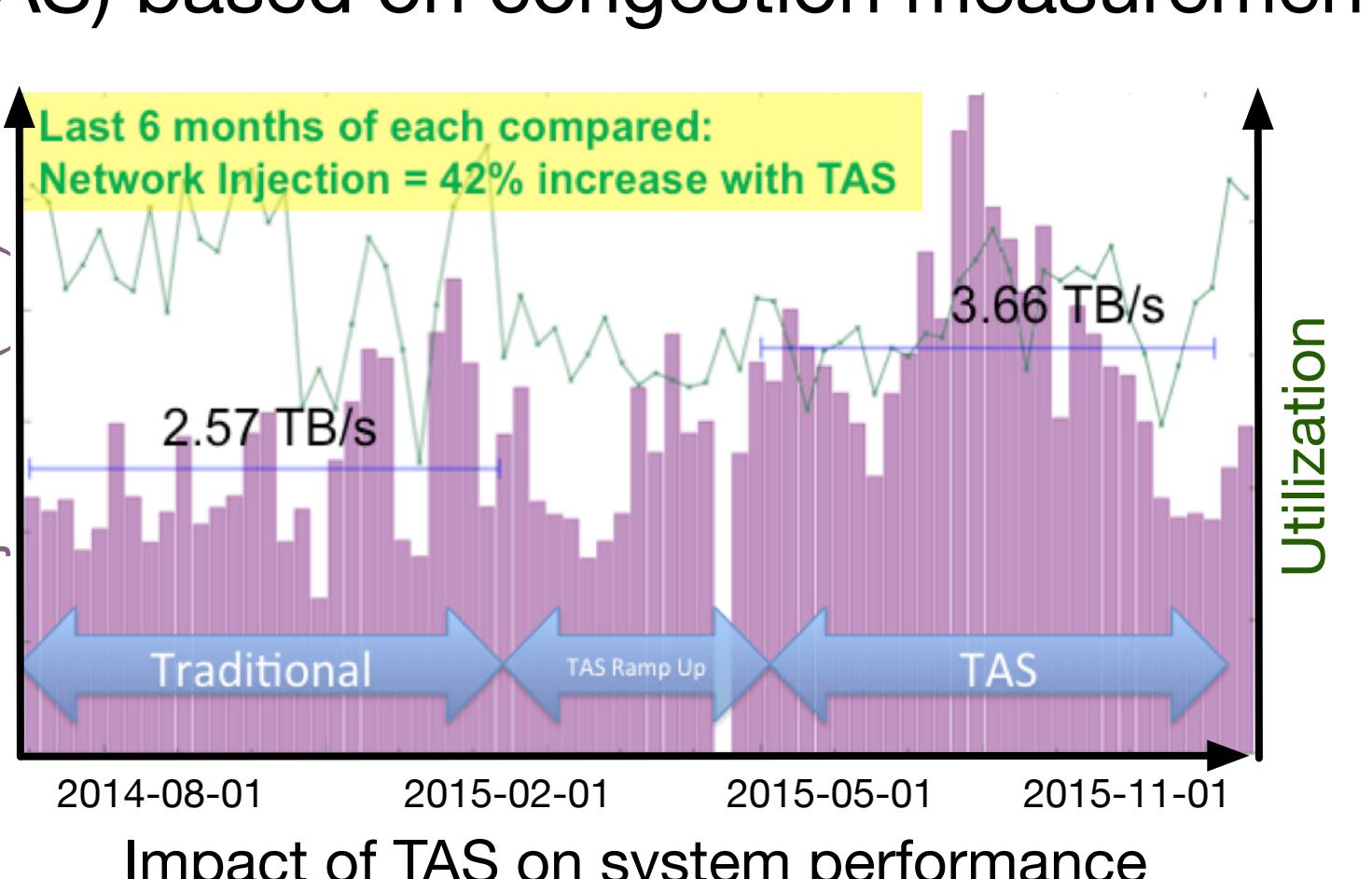
Server degraded



Insight: Identifying issues with adaptive scheduling and congestion mitigation procedures



Capability: Topology-aware scheduling (TAS) based on congestion measurements



Exascale-ready Features

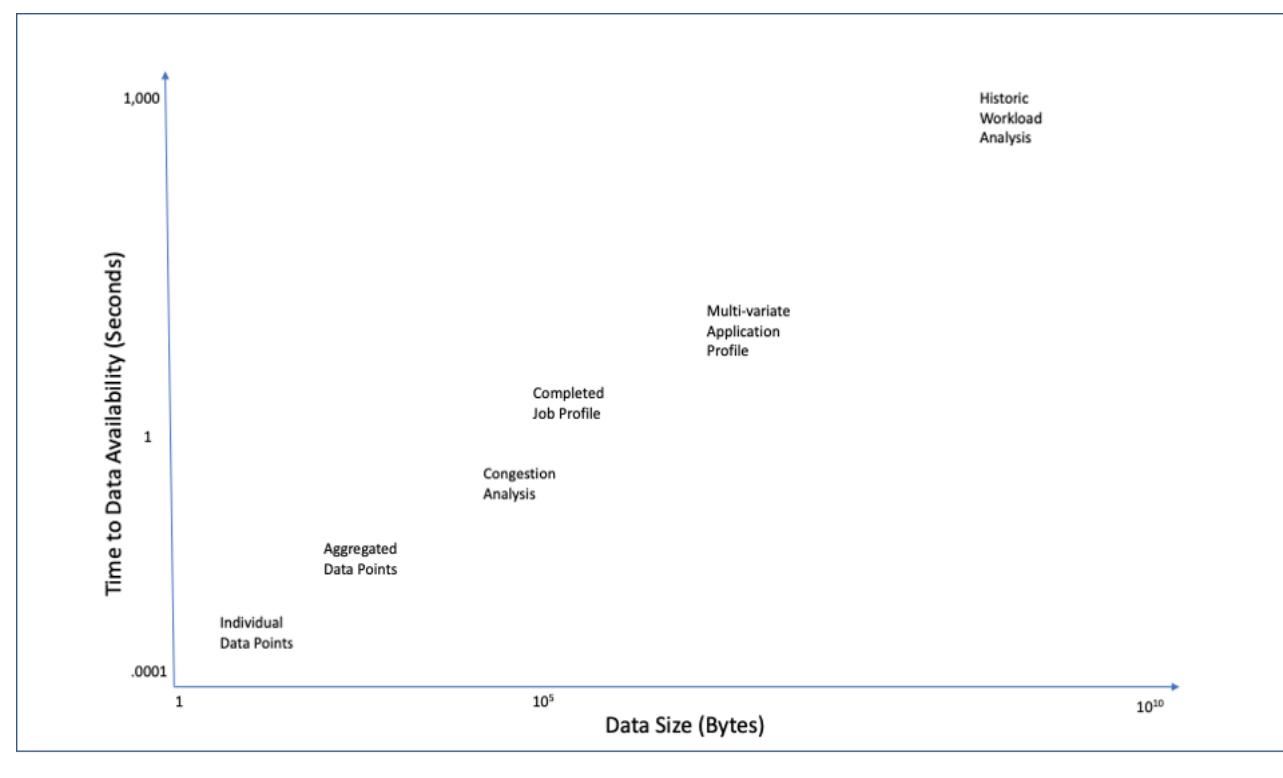
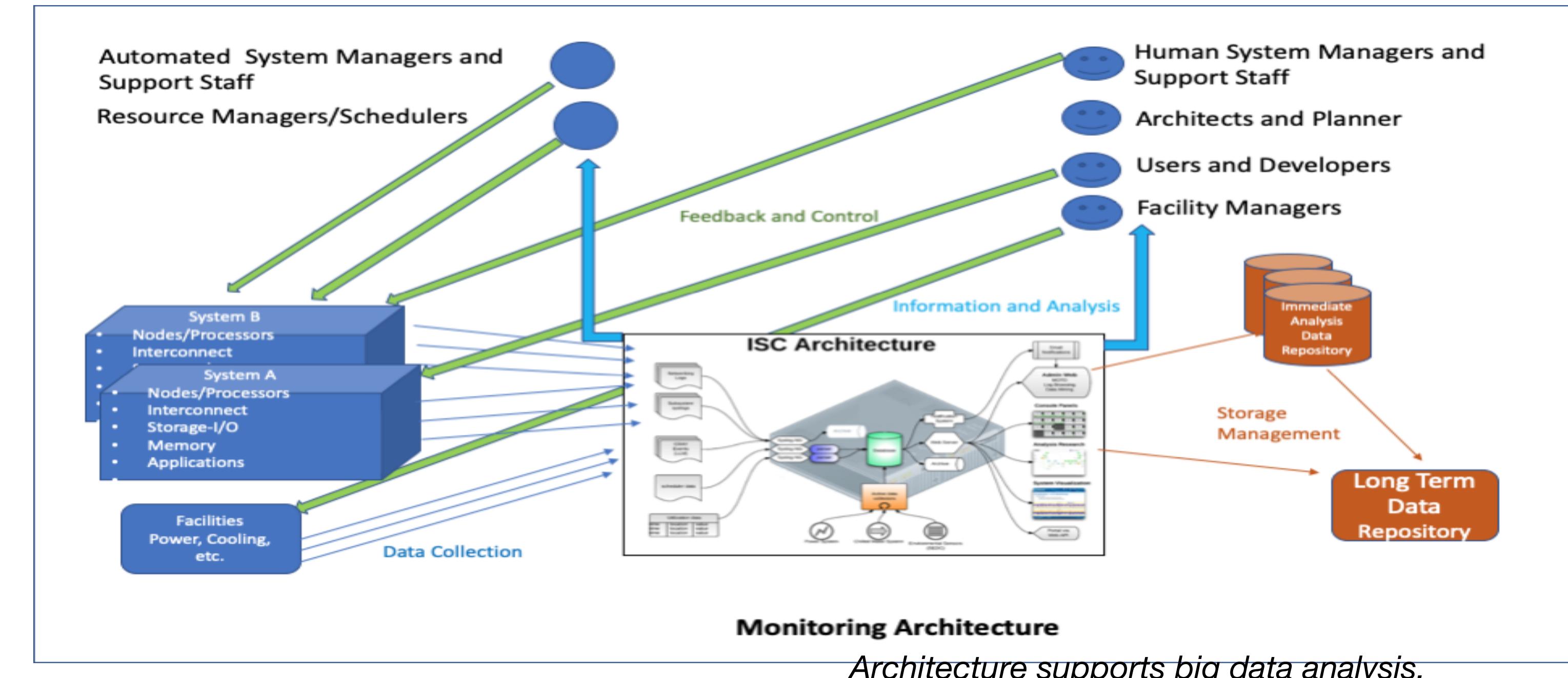
- Data-driven and Machine Learning (ML) based mechanisms for data collection, analysis, and automated runtime feedback and control
- Scalable APIs that enable flexible runtime interactivity (e.g., streaming analytics)
- Platform independent open-source solutions

Key Strengths

- Lightweight, low overhead mechanisms enable high fidelity (e.g. sub-second) synchronized, whole system numeric and event data capture with negligible impact on application runtime
- Discovery and analysis of performance and resilience related phenomena via integrated system logs and numeric data

- Enable highly efficient HPC system usage and inform future system improvements to produce more science
- Transform real-time data into actionable knowledge at runtime

Scalable System Design and Architecture



Expert and ML-Driven Large-scale Analysis

