SAND2015-10001C

Exceptional service in the national interest

Sandia National Laboratories

Infrastructure for In Situ System Monitoring and Application Data Analysis

J. Brandt, K. Devine, A. Gentile

U.S. DEPARTMENT OF ENERGY

National Nuclear Security Administration

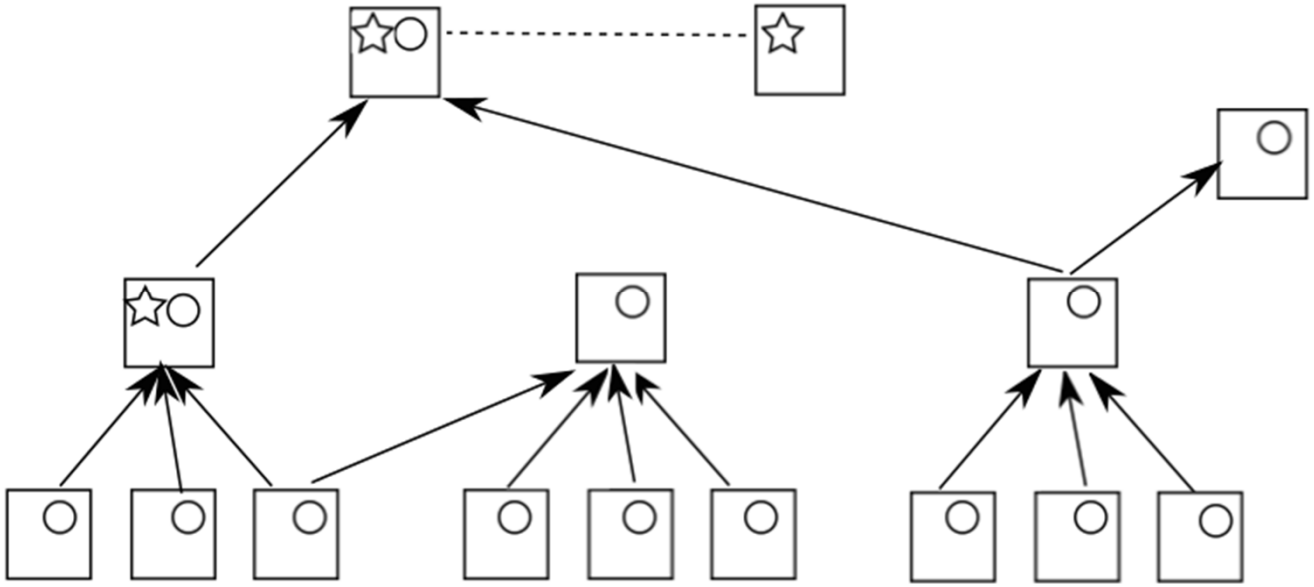
Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2011-XXXXP

Motivation

Sandia National Laboratories

- HPC system monitoring and application execution typically separate
  - Applications have no insight into resource, particularly global, state
  - Monitoring systems send to centralized server, low collection frequency
- Expose system state data for run-time analysis within the monitoring system and to the running application:
  - Application gains exposure to privileged monitoring data
  - Lightweight monitoring system provides high frequency data
  - Integration provides low latency
- Use cases:
  - Power, CPU utilization etc.
  - Response to discovered contention for shared resources and dynamic events
- Goal: Provide a system service for applications and system software to improve execution and system performance

Architecture Overview



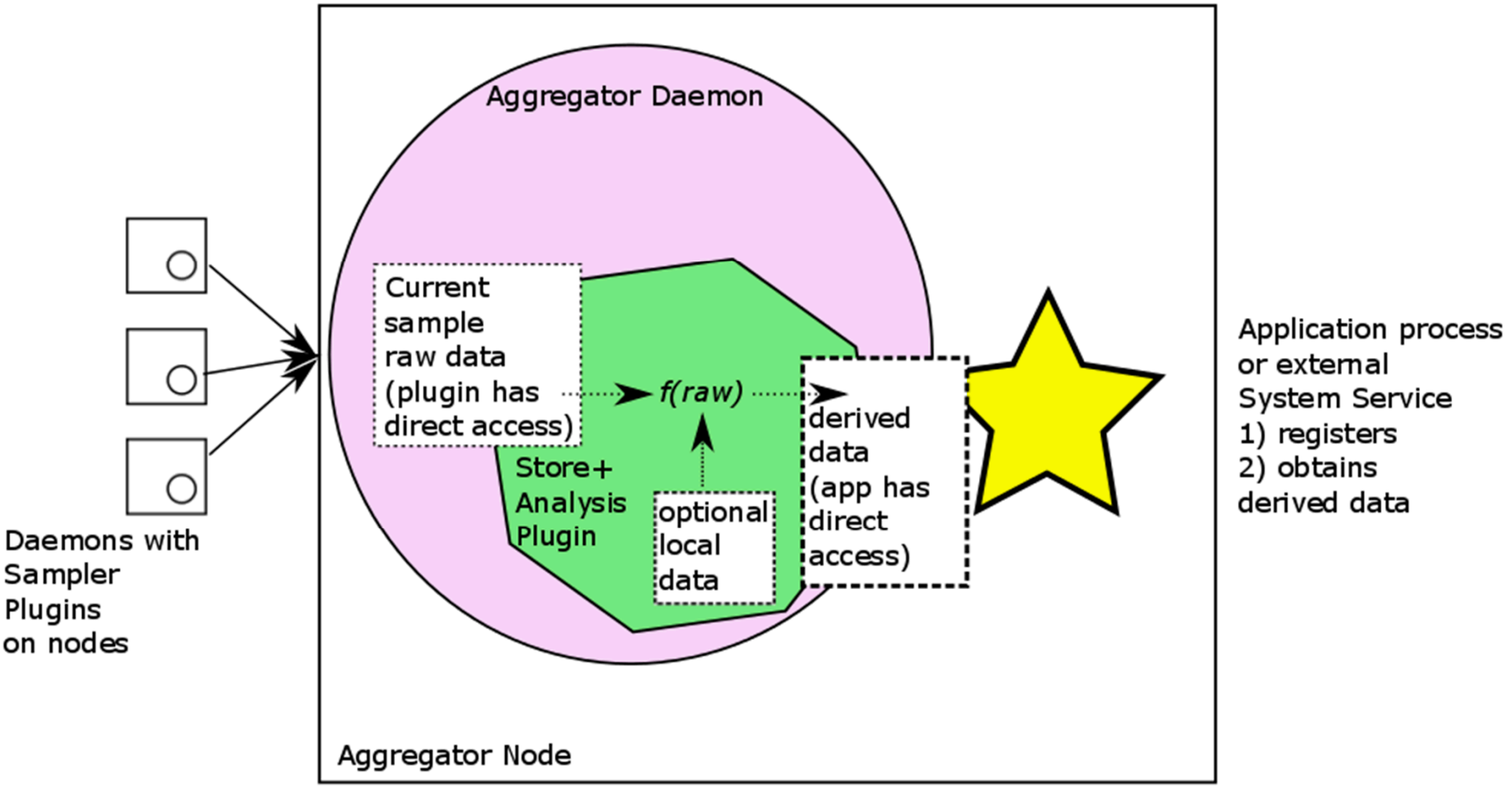
Star = application. No global data access

LDMS – lightweight monitoring with arbitrary aggregation topologies

- Circle = monitoring daemons
- *Continuous, full-resolution in situ data analysis via plugins as the data is resident in the aggregators*
  - Supports parallel analysis, use of run-time and historical data
  - Low-latency on-node access to results

Monitoring Analysis Interface

- Data consumer registers the intent to gather a defined set of data from the analysis plugin. Continuous analysis reduces latency.
- Memory mapped region can reduce latency
- Co-location of resource data, analysis, and application response



Dynamic Analysis and Response

Sandia National Laboratories

- Data Requirements:
  - Relevant but not jittery data
  - Reasonable expectation of future performance based on recent past
  - Fast in-situ analysis: redistribution vs analysis cost

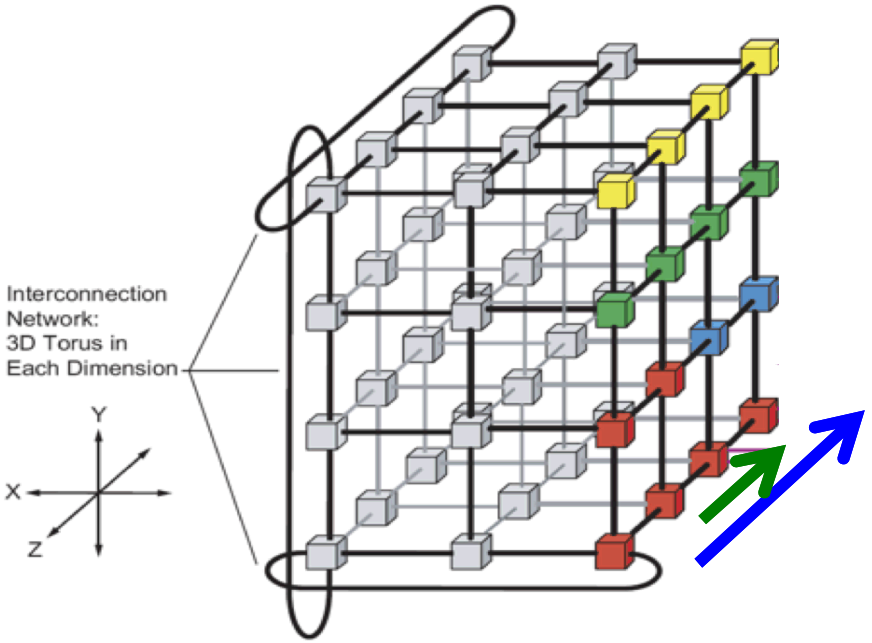
Use case: Analysis of network congestion and placement of MPI tasks on core to avoid detected congestion.

- Prior work: Can recover up to 49% of the execution time lost to congestion
- Goal: increase scale, decrease analysis time, reduce latency.
- New: Continuous analysis can proactively discover contention and trigger dynamic response

Network Congestion Analysis and Response

- Gemini network: 3D torus
- Traffic between source and dest is sent via a deterministic multihop route (X/Y/Z) (+/-)
- Congestion on intermediate and shared hops can affect performance. Info on those would not be accessible to an application.
- Congestion Measures:
  - Max % BW used
  - Max % time spend in stalls

  along the route



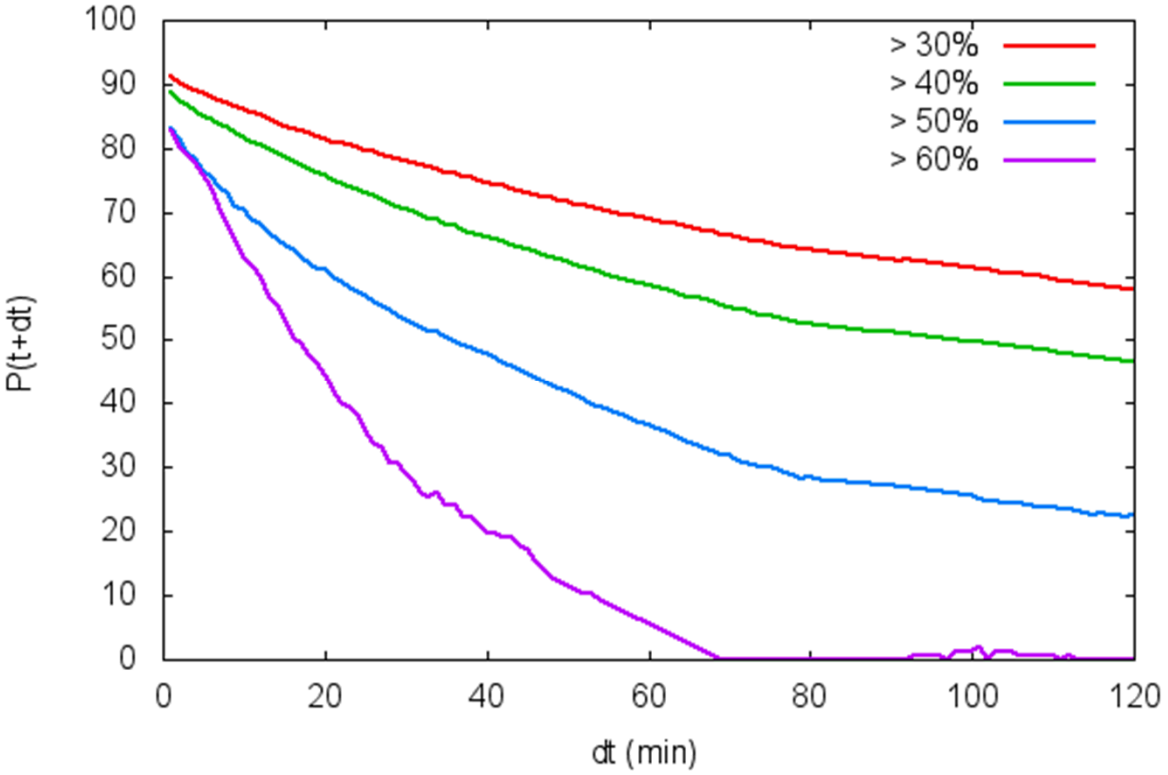
Analysis for Large Scale Systems

Sandia National Laboratories

- Analyze data at run-time as it streams thru the aggregator
- Calculation reduction:
  - Brute force: 27648*27648 routes for NCSA Blue Waters
  - Register only for the routes of interest
  - Perform all routes' calculation in parallel
  - Calculation of max along a route can be done across multiple aggregators, each with a subset of the data, and then combined
  - Include only the most important components:
    - Significantly less congestion in Z+/-
- Privileged data -- More relevant interpretation:
  - Network quiesce events cause the values to drop although congestion is still high
  - Rerouting
  - Stale counter values

Change Detection and Response Triggering

- Probability of congestion duration on a given link
- Congestion infrequent (val >= 30% only ~0.6% of the entire day)
- Congestion endures (val >=50% has over 60% probability of being so 20 min later)



Discover congestion to advise other jobs to redistribute data

Characterizations can guide data collection/recomputation frequency

- X+ congestion (Max % time spend in credit stalls) duration from 1 day
- Data collected at 1 min intervals

Network Congestion Responses

Sandia National Laboratories

- Goal: Minimize cost of communications and/or data movement.
- Tools: Map geometric or graph-based representations onto each other to improve associations (e.g., reduce costs, minimize edge cuts)
- Congestion-aware Task Placement: Map tasks to nodes. Inputs:
  - Task graph (derived from the application) - Vertices represent MPI tasks. Weighted edges represent # bytes communicated between tasks.
  - Architecture graph (uses system information) - Vertices represent allocated nodes. Weighted edges represent cost of communication between nodes. More heavily congested paths between processors have higher weights.
- Congestion-aware Data Partitioning: Balance workloads and maintain locality. Inputs:
  - Application data graph (derived from the application) - Weighted vertices represent data and computational cost. Weighted edges represent strength of dependence.
  - Network graph (uses system information) - Dynamically weighted based on congestion information
- Continuous Analysis can trigger response

Conclusion

Sandia National Laboratories

- Infrastructure for analyzing run-time system monitoring data and providing it to applications and system services
  - Improved task placement and work distribution
- In situ analysis
  - Higher frequency data for analysis
    - Need not be stored by the monitoring system
  - Low latency on-node access to results
  - Trigger events based on continuous analysis
  - Global data for decision making
  - Privileged data
- Analysis questions:
  - Characterize the impact of contention measures on performance to refine the weighting
  - Determine the effectiveness of response
  - Aries adaptive routing