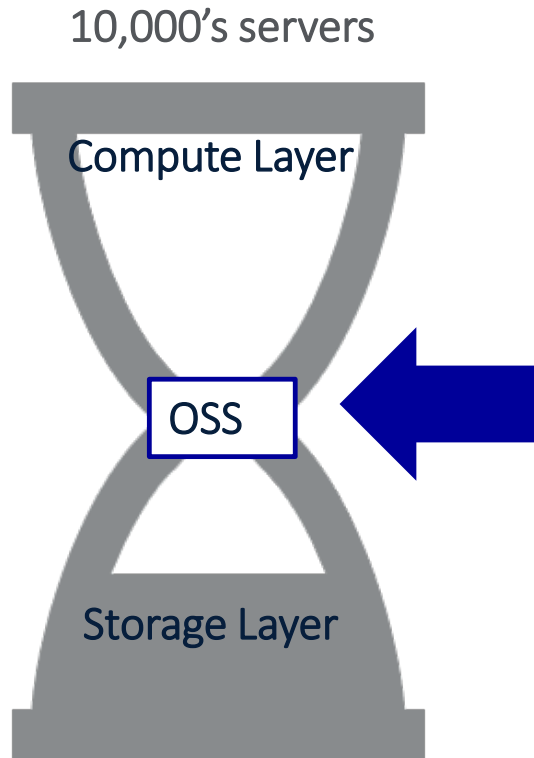


# Scalable Data Processing at Network transfer rates with nCorium Compute in Memory Modules

Suresh Devalapalli, Brett Neuman, Arvinth Lalam

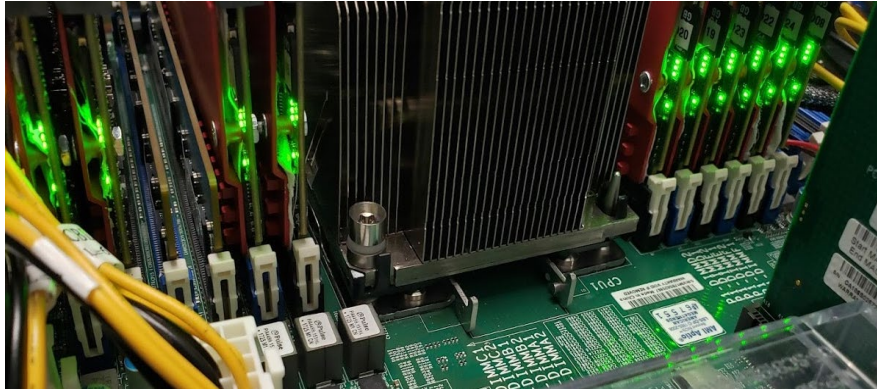


# OSS Layer Bottlenecks in HPC Facilities



- 10s of Peta Bytes of data moved in and out of Storage layer
- Challenges:
  1. Data processing on CPU unable to keep up with increase in network and storage speeds
  2. I/O throughput per node is limited

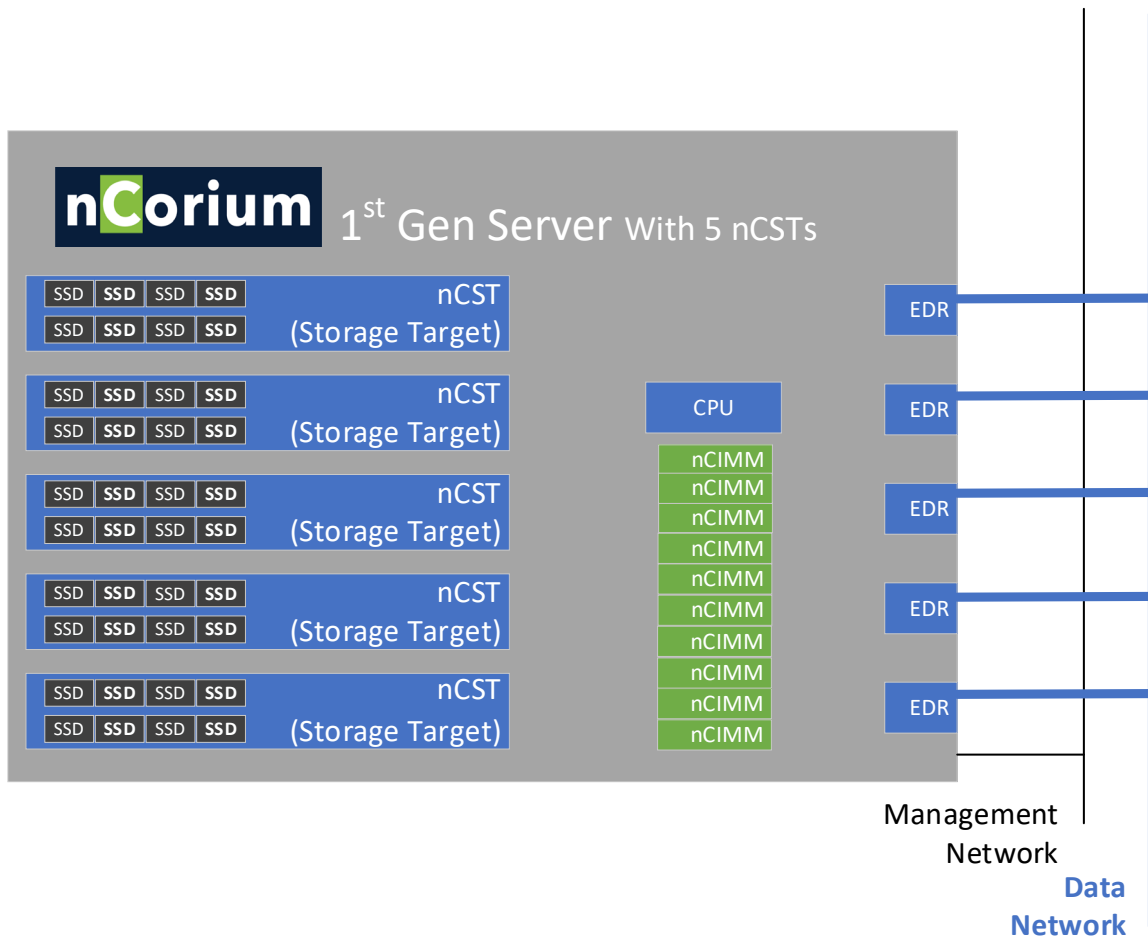
# nCorium Compute in Memory Module



nCIMM Array in the system

- DIMM form-factor
- Reconfigurable compute cores on the DIMM
- Off-load functions via libraries:
  - Compression
  - Erasure
  - Encryption
  - Video Transcoding
  - ....etc
- Data Stream based processing

# nCIMMs + nC-Grid = Scalable nCorium-OSS



- nCIMM to off-load data processing
- nC-Grid to open wider I/O
- Scalable Architecture:
  - Number of nCIMMs
  - Functions to support

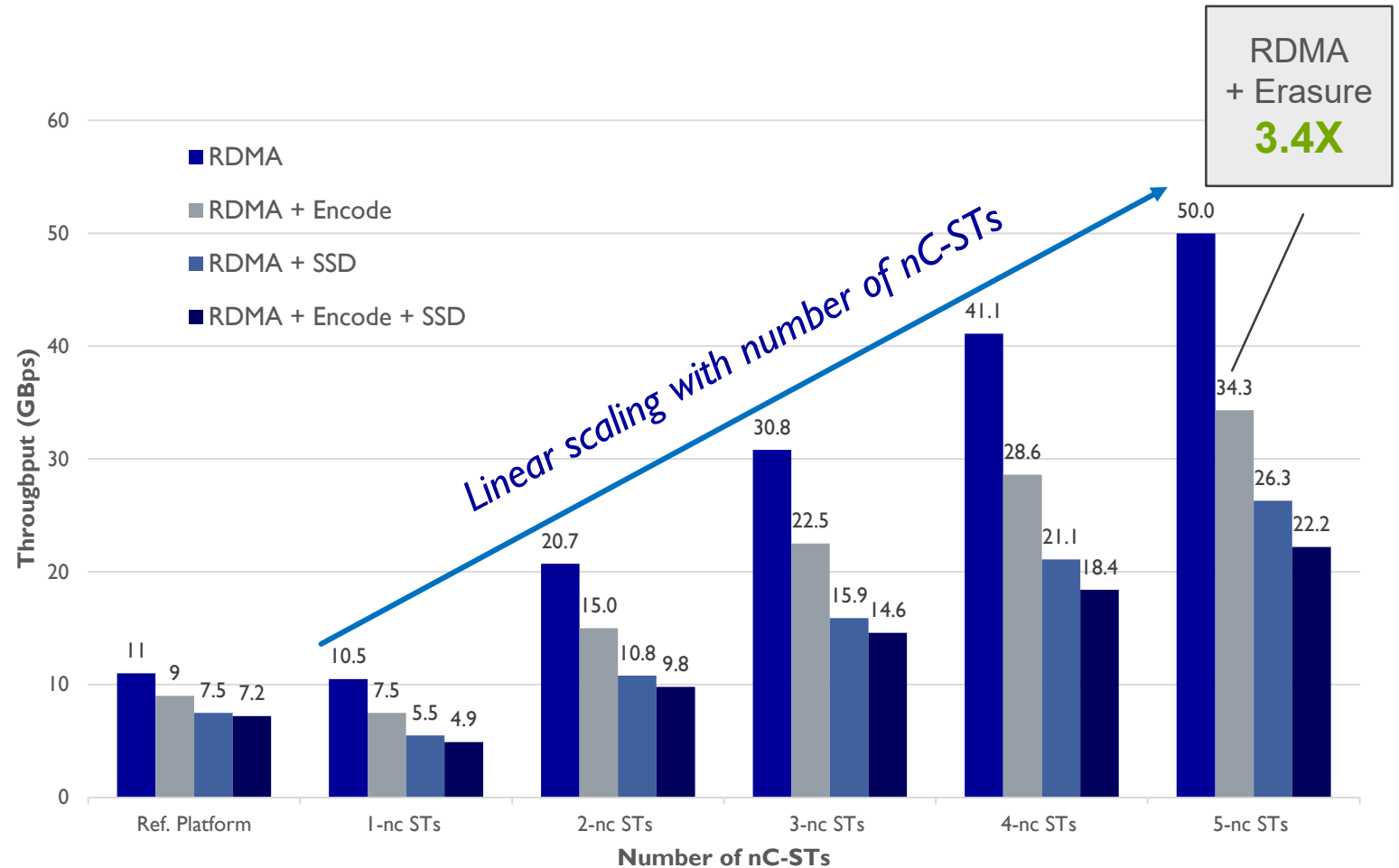
# Performance of Gen-1 nCorium OSS: Q'1 19

**5x** for RDMA

**3.4x** for RDMA+ Erasure

Performance compared with Intel Gold CPU based system

*Gen-1 released in Q1'19*



# Performance of Gen-2 nCorium OSS

- Improved reconfigurable cores
- More off-load functionality
- Higher Power efficiency

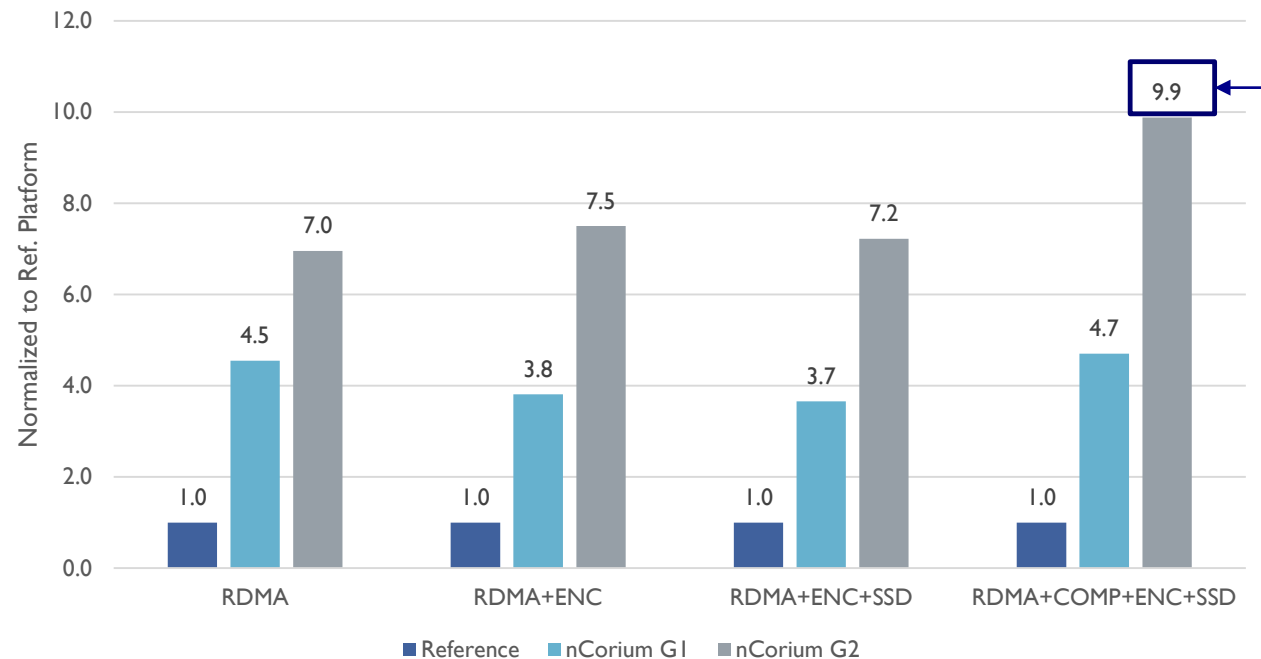
*Gen-2 now available*

|            | RDMA | RDMA+ENC | RDMA+ENC+SSD | RDMA+COMP+ENC+SSD |
|------------|------|----------|--------------|-------------------|
| Reference  | 11   | 9        | 7.2          | 5                 |
| nCorium G1 | 50   | 34.3     | 26.3         | 23.5              |
| nCorium G2 | 76.5 | 67.5     | 52           | 49.4              |

*Numbers extrapolated from initial data*

**~10x faster**

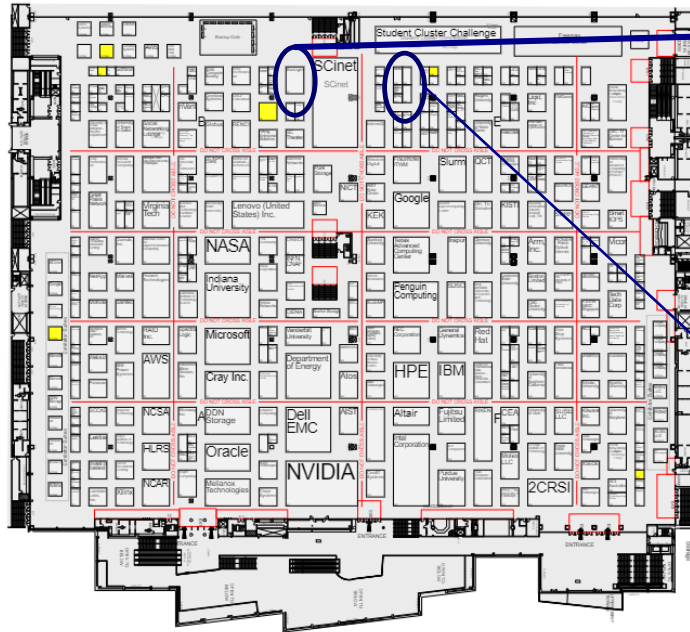
Ref. Platform vs nCorium G1 and nCorium G2



# Ongoing Work

- Parallel file system integration
  - Lustre
  - BeeGFS
- Other Applications
  - Key-Value Storage
  - Video Transcoding
  - Inferencing
  - Graph Analytics
  - Cyber Security
  - Content Distribution

# Demos / Exhibits



StarLight

993

StarLight: Booth 993

**Thank you!**

|  |             |
|--|-------------|
| Advantech Corporation  | OpenPOWER   |
| 1395   | 1494        |
| Interdisciplinary Centre for Mathematical and Computational Modeling, University of Warsaw<br>1393 | EchoStreams |
| Microchip Technology Inc.  | 1490        |
| 1391   |             |

EchoStreams: Booth 1490