SCALABLE COMPUTING SOFTWARE LABORATORY

ILLINOIS INSTITUTE OF TECHNOLOGY

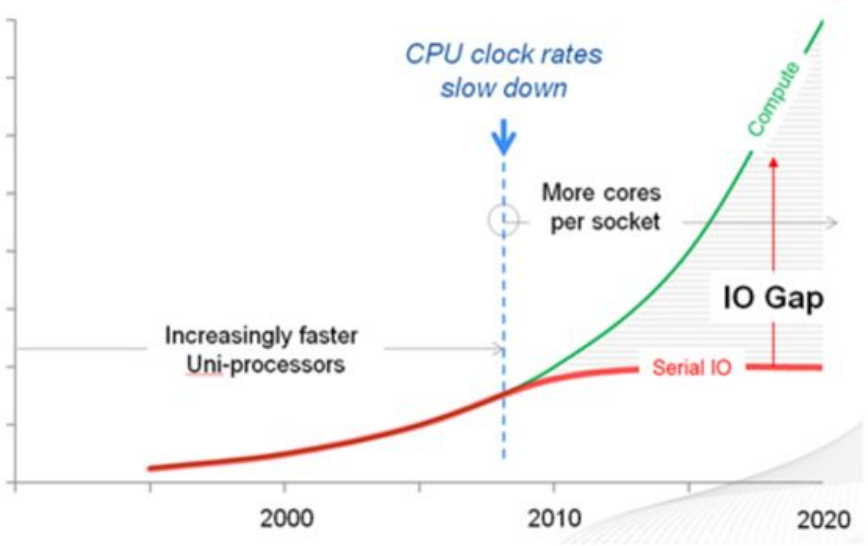
IEEE International Parallel and Distributed Processing Symposium (IPDPS)

Hierarchical Data Compression for Multi-Tiered Storage Environments

Hariharan Devarajan, Anthony Kougkas, Luke Logan, and Xian-He Sun
hdevarajan@hawk.iit.edu

I/O Bottleneck

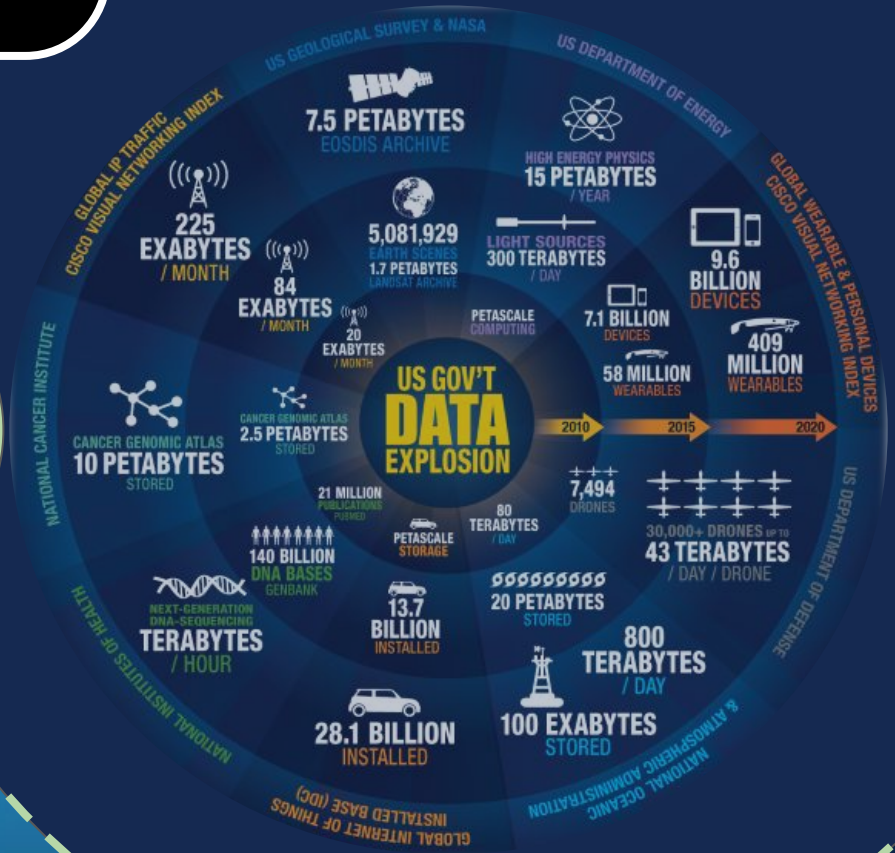
- In the data-intensive era, producing and consuming data is critical for scientific discovery.
- I/O subsystems struggle to match growing compute parallelism.
- System performance is bound by its slowest component. (Amdahl's “well-balanced” law)
- I/O performance is a concern in petascale, and would exaggerate even more as we ascend towards exascale.



Growing gap of CPU and I/O performance

Explosion of data

- Data is crucial to enable discovery.
- IDC reports predict that by 2025:
  - global data volume will grow to 163 ZB
  - 10x the data produced in 2016



Reduction of I/O bottleneck

- Several middleware solutions are proposed to reduce the I/O latency and increase application performance.
- In all approaches, the solutions utilize an Intermediate Temporary Scratch (ITS) space (e.g., Main Memory) to optimize I/O access.

I/O Optimization

<u>Increasing the space of ITS would greatly enhance the effectiveness of these solutions.</u>

Current approach: Increase ITS space

Tiered Hardware

- Add new intermediate resources to increase layers.
- E.g., HBM, NVRAM, NVMe SSD, etc.
- Increases space availability.

Software

- Reduce I/O footprint through data reduction techniques.
- E.g., Data Compression.
- Reduces Data Footprint.

Observation

Benefit of compression comes from trading CPU cycles to reduce I/O cost.

The new hardware reduces this I/O cost.

A combination of these two approaches can compound the increase of available ITS for I/O optimizations.

Hypothesis

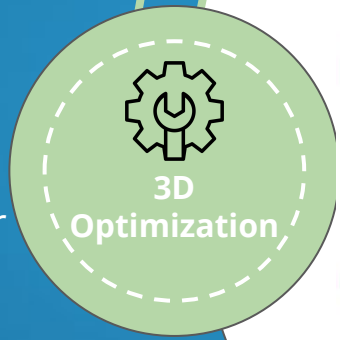
HCompress

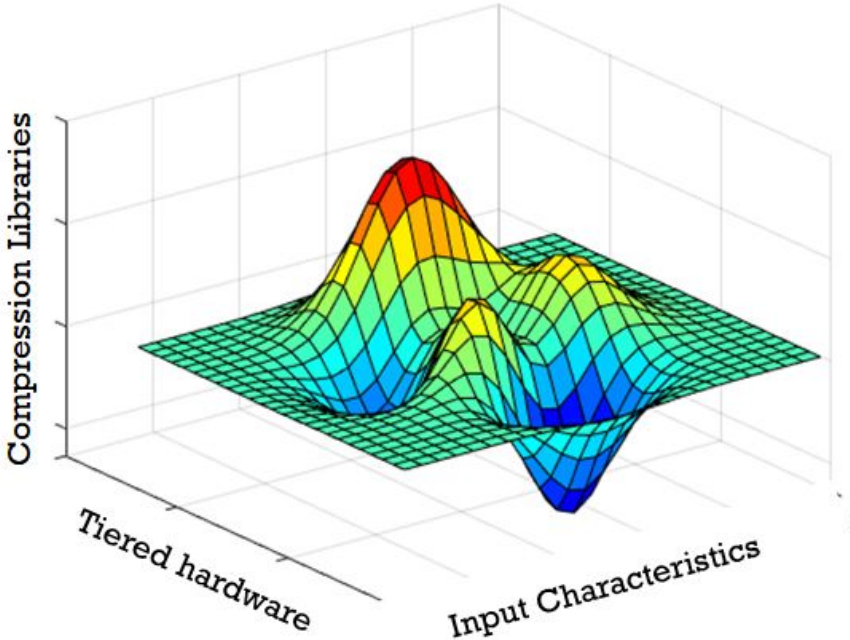
Hierarchical Data Compression for Multi-Tiered Storage Environments

Code: https://bitbucket.org/scs-io/hcompress

Problem Formulation

- Match three dimensions
  - Application Characteristics
  - Compression Characteristics
  - Hierarchical Tier Characteristics
- We can formulate it as a minimization of total time for executing an I/O task
- The constraints required
  - # sub-problems should be small.
  - Data compression is useful.
  - Compressed data fits in a tier.





Visual representation of 3D space.

HCompress Goals

Hierarchical

Utilize all storage hardware efficiently.

Match hardware speed of devices to ideal compression libraries

Dynamic

Dynamically switch compression libraries.

Cost of reconfiguration of compression engine should be low.

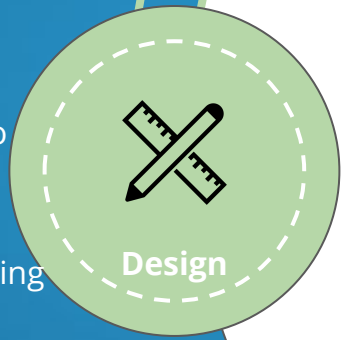
Flexible

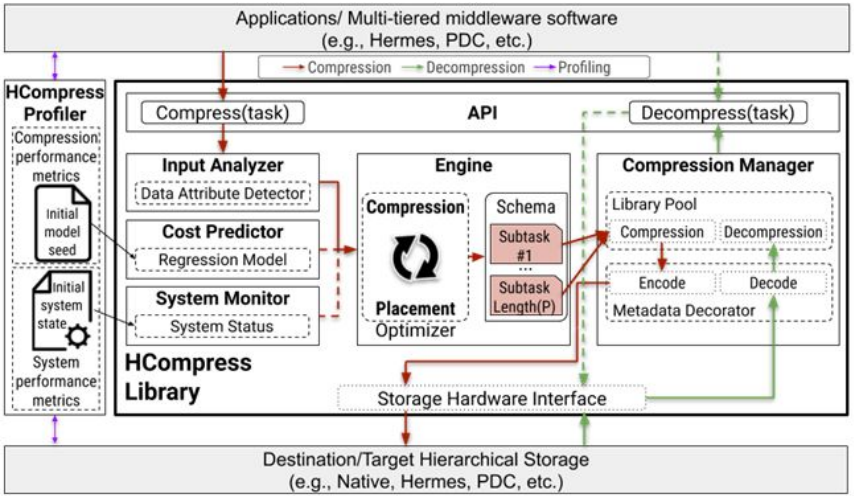
Unify the interface to compression libraries

Configure, add, and apply compression using a interface.

HCompress Design

- HCompress Profiler
  - Runs a exhaustive benchmark to capture system and compression characteristics.
- Compression Cost Predictor
  - Uses linear regression model
  - Uses reinforcement learning to improve accuracy.
- Engine
  - Employs a dynamic programming (DP)
    - Data characteristics, Compression libraries, and Storage tiers
- Compression Manager
  - Manages library pool
  - Performs metadata encoding





Evaluation

- **Cluster Configuration**
  - 64 compute nodes
  - 4 shared burst buffer nodes
  - 24 storage nodes
- **Node Configurations**
  - compute node
    - 64GB RAM and 512GB NVMe
    - Burst Buffer node
    - 64GB RAM and 2x512GB SSD
- **Storage node**
  - 64GB RAM and 2TB HDD

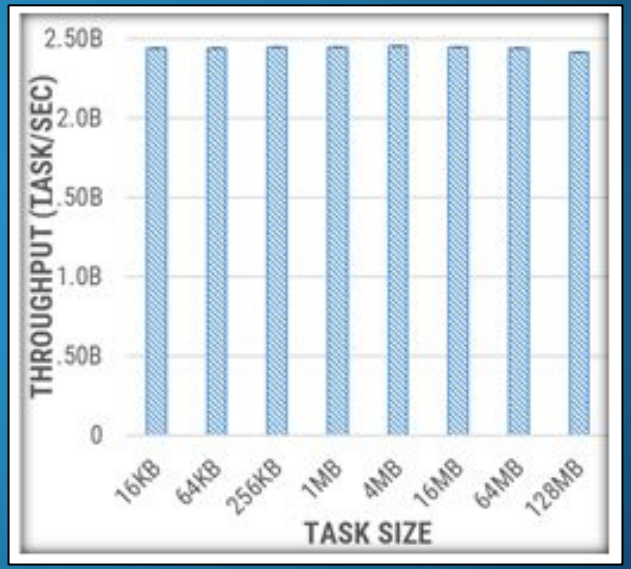
Tedbed

Config

- **Applications tested**
  - Synthetic Benchmarks,
  - VPIC, and
  - BD-CATS
- **Compared solutions**
  - Baseline vanilla PFS
  - Single-tier with compression
  - Multi-tiered without compression

HCompress Component: Engine

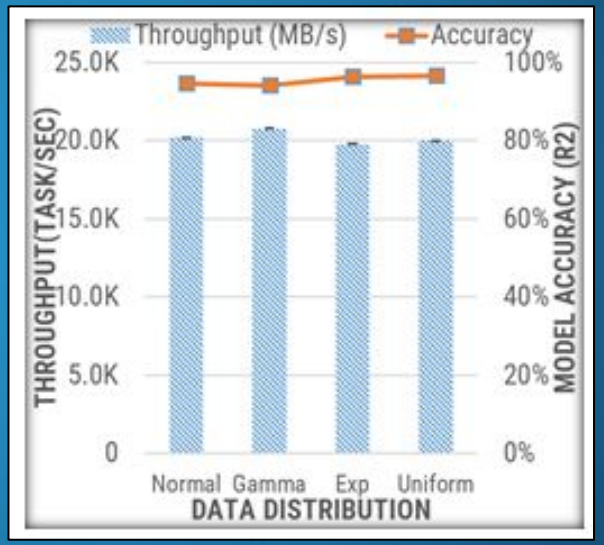
- Single Core test
- Test the Engine with different input sizes
  - As input size increases data is “potentially” split into more parts



- **<u>Observations</u>**
  - Constant throughput of **2.4B ops/sec**
  - Complexity **O(#layers)**.

HCompress Component: CCP

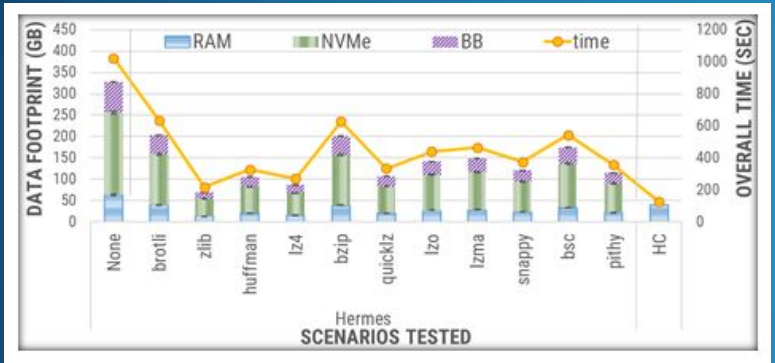
- Compression Cost Predictor(CCP)
- Test the CCP with different input distributions
  - As data distribution effects the prediction accuracy of compression cost



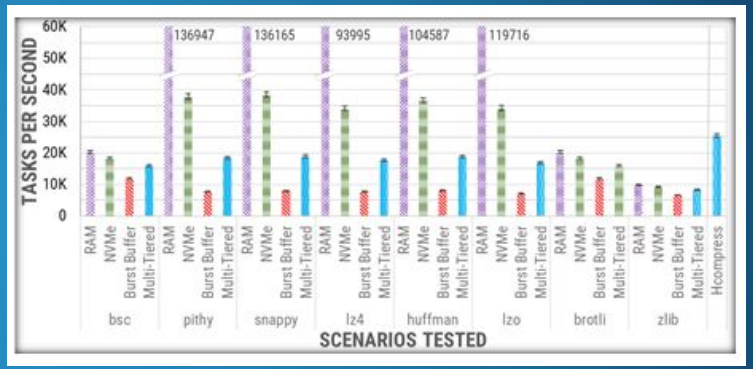
- **Observations**
  - Constant throughput of **20K ops/sec**
    - Bound by memory
  - Accuracy of **96%**

Impact of Tiered Compression

Compression on Tiered Storage



Tiered Storage on Compression



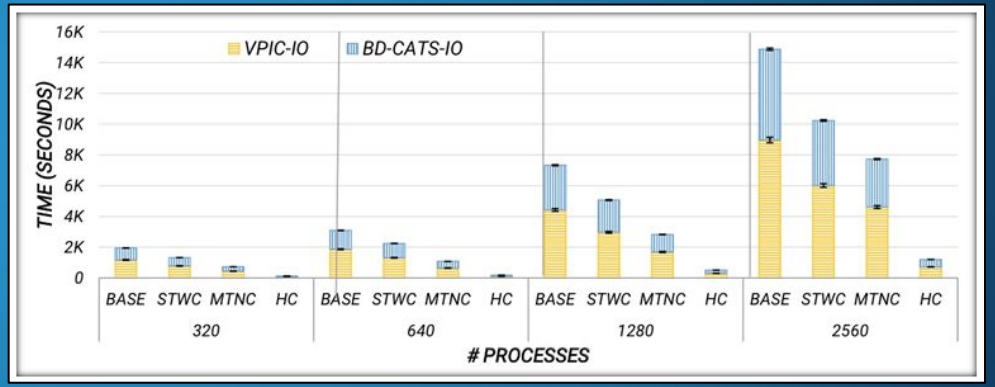
Observations:

- Performing multi-tiered buffering with single compression doesn't maximize the benefit.
  - data placement is not aware of compression.
- HCompress achieves a benefit of **2x**.

- Different tier effect differently for each compression
- HCompress balances trade-off dynamically and achieves the best multi-tiered throughput.

Scientific workflow

- Cluster test (64 nodes)
- **Workflow**: (HDF5 workload)
  - Simulation using VPIC
  - Analysis using BD-CATS
- **<u>Observations</u>**
  - Optimizes both write and read performance significantly
  - Optimizes all three parameters: compression time, decompression time and compression ratio equally
  - Achieves a performance boost of **7x.**



Conclusions

1. HCompress showcased how data characteristics and system characteristics affect data compression.
2. HCompress proposes a hierarchical compression engine for multi-tiered storage environments
3. Quantified the benefit of utilizing hierarchical hardware and data compression cohesively.
4. HCompress can optimize scientific workflows up to 7x compared to competitive solutions.

A list of all observations

Q&A

Thank you

hdevarajan@hawk.iit.edu

Video

SCAN ME

Code

SCAN ME