Improved Solutions for I/O Provisioning and Application Acceleration

August 11, 2015

Jeff Sisilli
Sr. Director Product Marketing
jsisilli@ddn.com

Why Burst Buffer?

The Supercomputing Tug-of-War

"A supercomputer is a device for turning compute-bound problems into I/O bound problems"

Ken Batcher,
Emeritus Professor of Computer Science, Kent State University

COMPUTE ACCELERATION

The Divide Driving Exascale Innovation

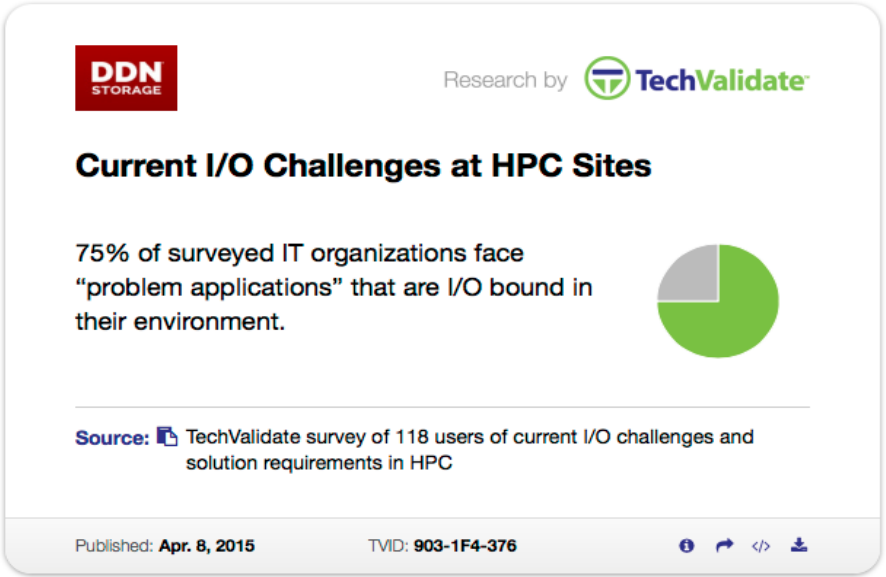
"DDN's storage mission is to eliminate I/O-bound problems and revert them back to compute-bound ones"

Alex Bouzari,
CEO & Founder, DDN

STORAGE ACCELERATION

PAIN: “Problem” I/O Bound Applications

Longstanding PFS I/O Bottlenecks Must Be Eliminated



TAKEAWAY

1. “Problem applications” are a huge source of known pain in HPC

HOW A BURST BUFFER HELPS

1. Accelerates applications and returns time available for computation by orders of magnitude

Polling HPC TOP500 Sites

RFP Mindshare of Various Flash-based I/O Acceleration Technologies

Which of the following are you including in upcoming RFP's to speed-up I/O and applications? Check all that Apply.

Choice	Responses	Percentage
Burst Buffer	40	44%
Flash and/or Edge Appliances	27	30%
All Flash Storage Arrays	23	25%
Hybrid (SSD + SAS) Arrays	63	69%

Current Motivation for I/O Accelerators

Faster Time to Results vs. Gaining New Efficiencies

For which purpose are you primarily considering these technologies?

Choice	Responses	Percentage
Faster time to discovery, insight or results	57	58%
Reducing hardware, cost and footprint of provisioning bandwidth in traditional spinning disk approaches	60	61%

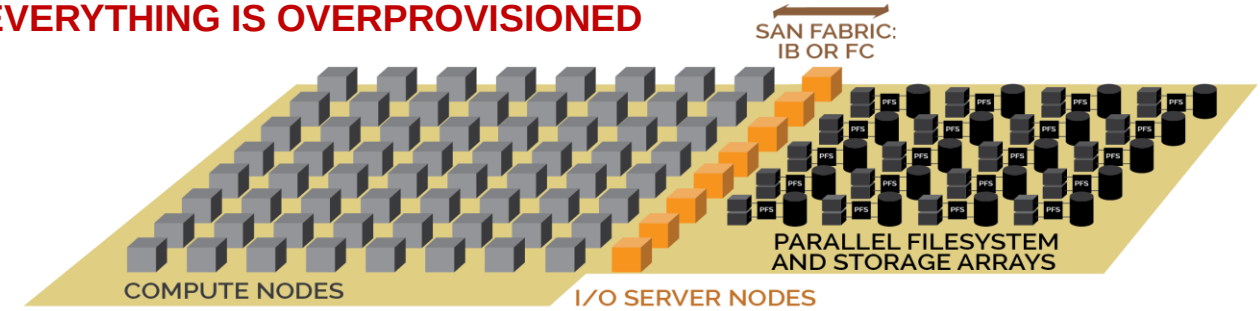
The Next I/O Provisioning Revolution:

Decoupling Physical Storage from Compute Resources!

BEFORE

Too Many: COMPUTE NODES DISKS, NETWORKING NODES, ARRAYS, ADMIN, H/W

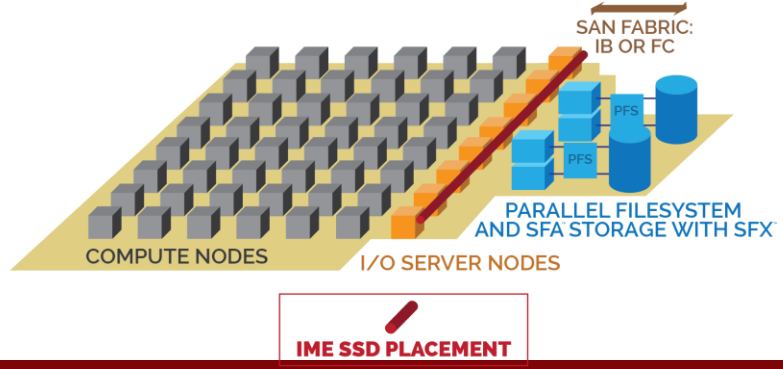
EVERYTHING IS OVERPROVISIONED



AFTER

Much Fewer: COMPUTE NODES DISKS, NETWORKING NODES, ARRAYS, ADMIN, H/W

A LOT MORE SPEED TO THE APPLICATION & A LOT LESS COMPONENTS



Even Building the World's Fastest PFS . . .

Will NOT Fix These I/O Challenges

PFS Locking

PFS are not designed for today's mixed I/O & ensembles

Storage Latency

HDD seek times & network traversing add latency

Fragmented I/O Patterns

Mal-aligned apps slow down the PFS & entire cluster

Out of Core Data

Many datasets are too big for expensive DRAM

Limited Power

Exascale or next scale needs more space & power

No matter how many HDDs you add to a PFS,
you can't break I/O bottlenecks without a burst buffer

DDN

Burst Buffer & Beyond: **IME®**

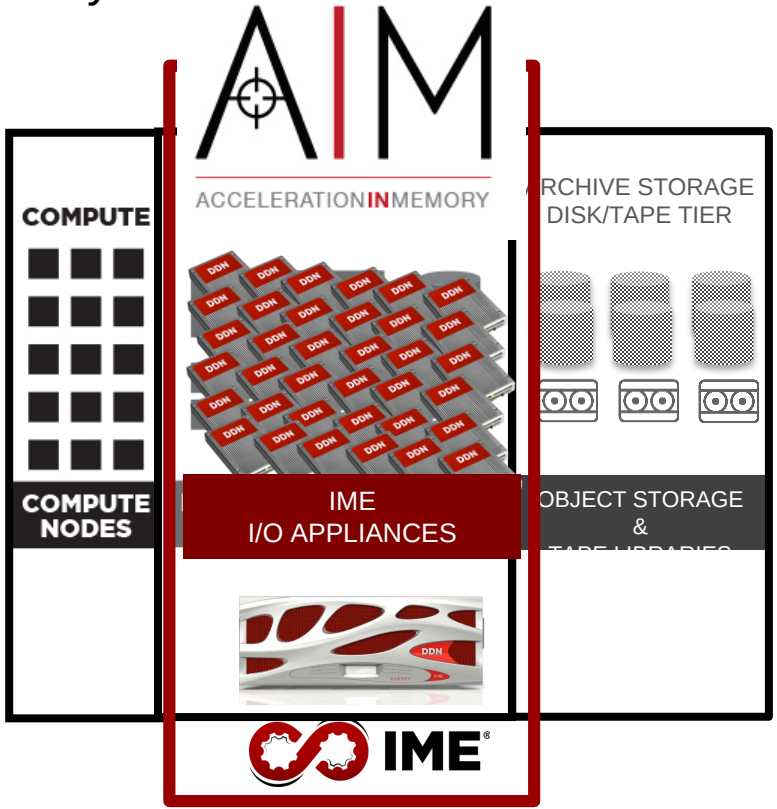
IME®

The New I/O Acceleration Architecture

AIM™ - Acceleration IN Memory

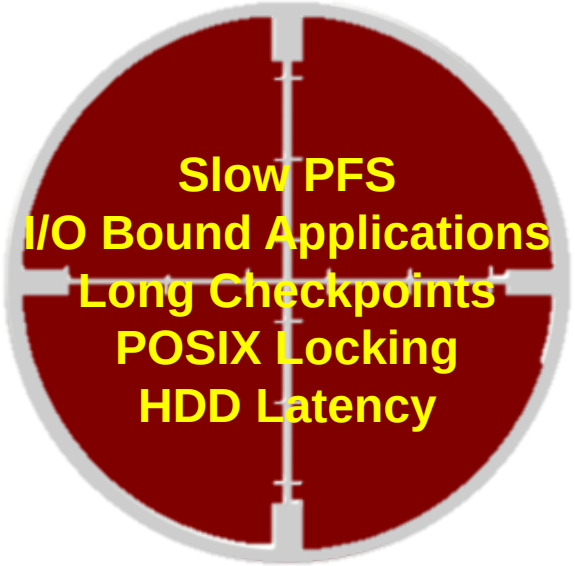
Introducing AIM, An *Active I/O Tier*, inserted right between compute and your PFS

Intelligent IME software virtualizes disparate NVMe SSDs into a single pool of shared memory that accelerates I/O, PFS & Applications



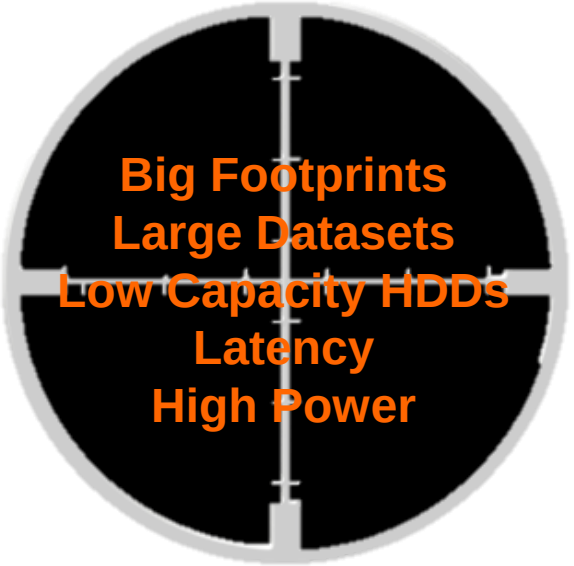
Take AIM & Target I/O Bottlenecks With IME

Eliminate Overprovisioning & Storage Sprawl



AIM

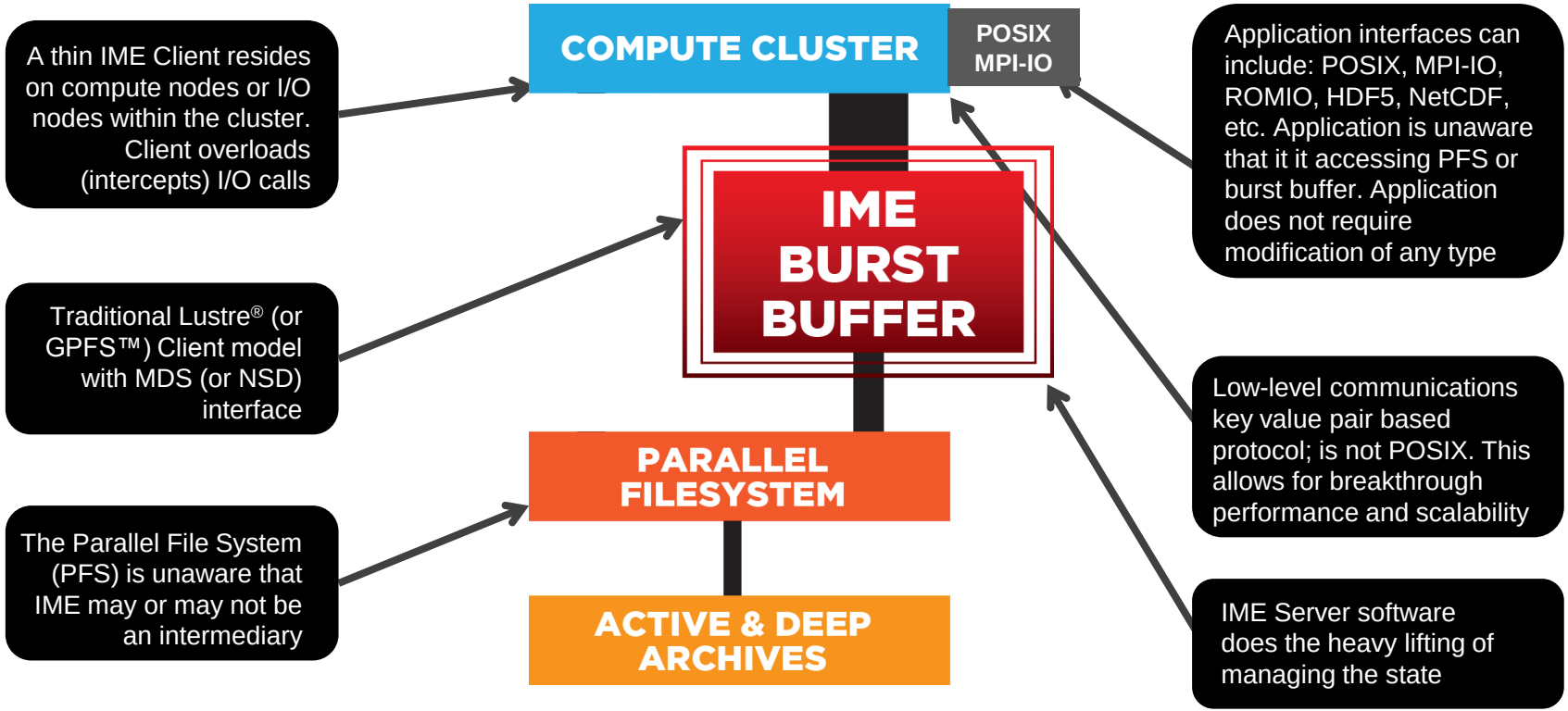
ACCELERATION IN MEMORY



Introducing IME®

Key Components and Operations

View Slide in Full Screen

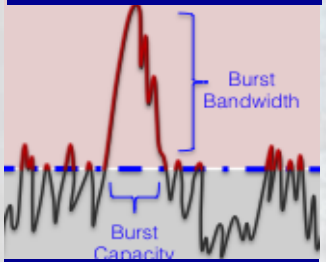


IME: A Burst Buffer & Way, Way, Way Beyond

Game Changing, Enabling Technology

Cache is only the beginning. Right out of the box, IME does so much more . . .

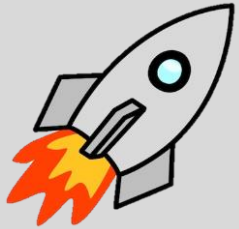
BURST BUFFER



Most cost & space efficient way to provision peak performance

+

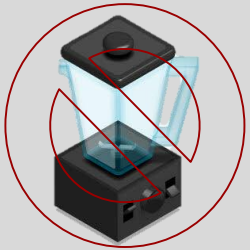
PFS ACCELERATOR



Finally breaks POSIX locking bottleneck with instant open/close

+

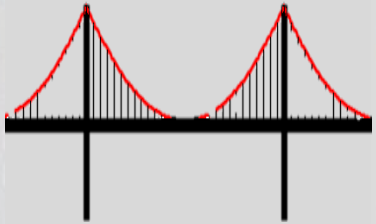
APP OPTIMIZER



Dynamically aligns mal-formed I/O into striped writes without code mods

+

DRAM EXTENDER



No dataset is too big with TBs or even PBs of fast, cost efficient NVMe

DDN

New Considerations for Architecting I/O Performance

IME

Game Changing Bandwidth

IME Disrupts How Performance is Provisioned

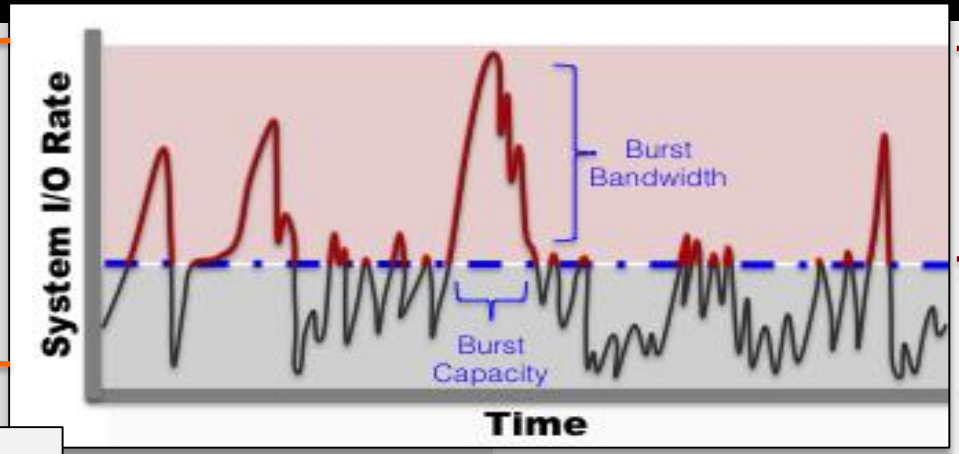
View Slide in Full Screen

IME introduces a more efficient way to provision performance than just storage arrays alone

BEFORE IME:

PFS Systems were designed to handle the entire performance load

This required lots of storage controllers, enclosures and drives to deliver full bandwidth –



TODAY, WITH IME:

IME's BURST BUFFER Absorbs the Peak Load

PARALLEL FILE SYSTEM Handles the Sustained Load

STORAGE BANDWIDTH UTILIZATION OF A MAJOR HPC PRODUCTION STORAGE SYSTEM

- **99%** of the time **< 33% of max**
- **70%** of the time **< 5% of max**

IME enables peak performance to be provisioned with much less hardware, power, space

IME Accelerates I/O in Several Ways

"Problem Application" Case Study: S3D

View Slide in Full Screen

10x

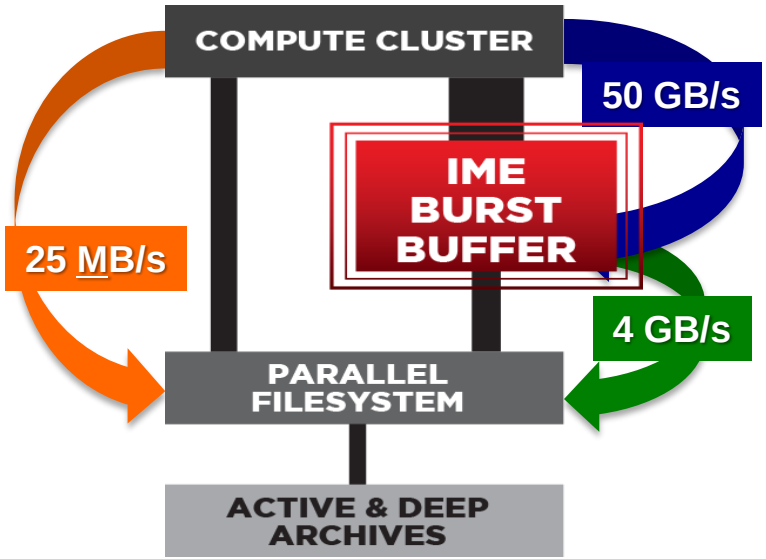
1) MITIGATES POOR PFS PERFORMANCE caused by PFS locking, small I/O, and mal-aligned, fragmented I/O patterns.

IME "**makes bad apps run well**" and also prevents a poor-behaving app from impacting the entire supercomputer.

This is especially valuable to **diverse workload** environments and **ISV** applications.

At SC14, we **demonstrated 1000x speed-up** on mal-formed I/O when using non-POSIX low-level communications.

1000x



S3D Turbulent Flow Model

2) PROVIDES HIGHER PERFORMANCE I/O (bandwidth and latency) to the application.

Providing additional bandwidth here is relatively inexpensive. Configuring **10x more bandwidth** compared to PFS is typical.

3) IME DRIVES I/O MORE EFFICIENTLY TO THE PFS by re-aligning and coalescing data within the non-volatile storage.

At SC14, we demonstrated **100x speed-up** due to this efficiency. IOR benchmarks show a **3x – 20x speedup on I/Os <32KB**.

100x

How Does IME Help?

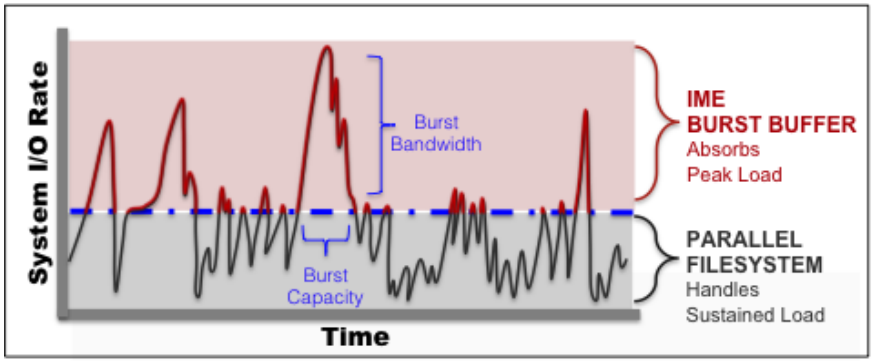
Disrupts the I/O Provisioning Paradigm & Reduces the Total Cost of Storage

1. IME enables organizations to separate the provisioning of peak & sustained performance requirements with greater operational efficiency and cost savings than utilizing exclusively disk-based parallel file systems

IME®

STORAGE BANDWIDTH UTILIZATION OF A MAJOR HPC PRODUCTION STORAGE SYSTEM

- **99%** of the time **< 33% of max**
- **70%** of the time**< 5% of max**



✓ **IME Reduces Storage Hardware up to 70%**

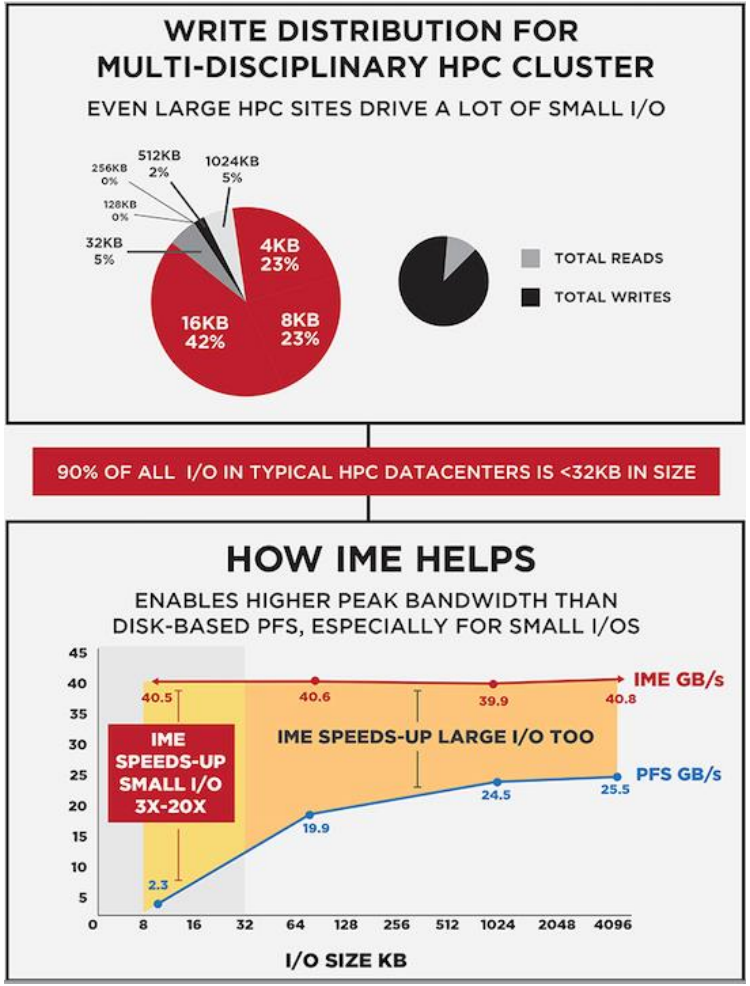
- Fewer systems to buy, power manage, maintain

How Does IME Help?

Increases I/O & Application Performance

2. IME Accelerates applications, especially those with small or mal-aligned I/O for faster time to results & insight

IME®



Thank You!

Keep in touch with us

sales@ddn.com

@ddn_limitless

company/datadirect-networks

2929 Patrick Henry Drive
Santa Clara, CA 95054

1.800.837.2298
1.818.700.4000