Advancing the Life Cycle of iRODS for Data

DAVID SALLAK, VP PRODUCTS & SOLUTIONS

JUNE 2016

NGS NEEDS PERFORMANCE

Sequencing first human genome in 2003 (the original Human Genome Project) took ten years and cost $3B USD. Now it costs just over $1000 and takes three days. Next is HGP Write, producing synthetic genes on top of the read scanning performed today.

The impact?

- **Rapid adoption of high-throughput lab instruments creates massive amounts of new n-dimensional data**
  - DNA sequencers, mass spectrometers, confocal & electron microscopes, Nuclear Magnetic Resonance Imaging (NMRI), flow cytometry...

- **Science Changing Fast; IT Infrastructure must adapt easily**
  - Storage must be flexible and scalable

- **Now I/O Performance and Capacity Scaling are Both Critical**
  - Analysis of data on a massive scale leading to data becoming more active continually – often no longer archival in nature

STORAGE REQUIREMENTS

What does *Good – Fast – Easy* look like?

- **True Scale-out**
  - Storage scales easily when adding more devices
  - Seamless and non-disruptive
  - Highly available scalable architecture, single namespace
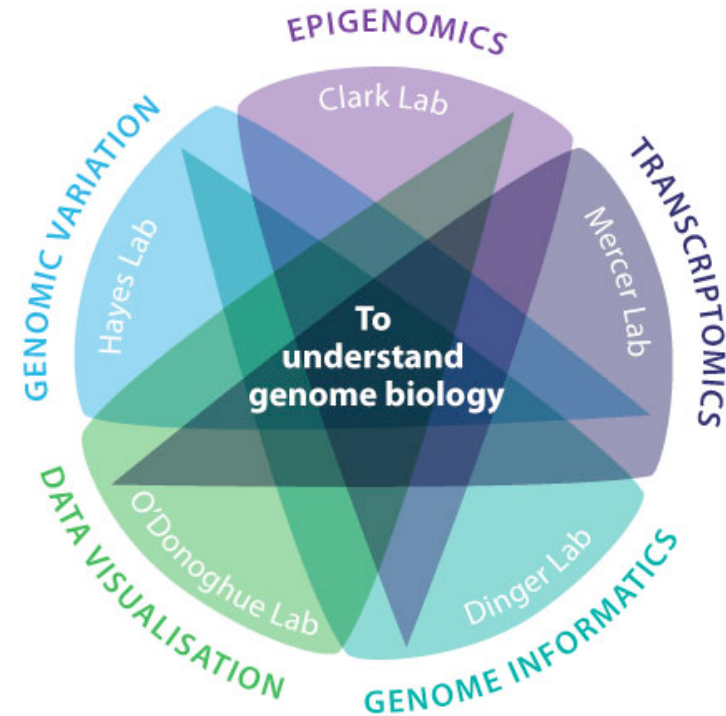 - Multi-protocol access to files from Linux, Windows, Mac clients

- **Performance via Direct Parallel Access**
  - Each compute client addresses the entire dimension of storage
  - No filer head or single node bottlenecks in the data path
  - Performance scales with capacity

- **Very Simple to Manage**
  - Data is automatically re-balanced and evenly distributed during expansion
  - Dynamic I/O load balancing
  - Less than one headcount manages terabytes or petabytes

INTRODUCING GARVAN INSTITUTE

- **Focused on translating technical advances into improvements in clinical practice**
  - Cancer risk assessment
  - Diagnosing children with intellectual disabilities

RESEARCH CHALLENGES AT GARVAN

This instrument changed everything

Illumina HiSeq X Ten

- **New Illumina performs faster**
  - Garvan was one of the first three organizations worldwide to deploy the Illumina HiSeq X Ten sequencing system. Each is capable of:
    - 50 human genomes per day = 1.8PB data per year

- **Growing Research Staff Size**
  - From 10 to 80 scientists
  - Existing storage incapable of sufficient scalability

- **Time to Results Matters**
  - Needed to integrate new sequencers, compute, and storage seamlessly into research workflows

- **Grants Cover Hardware, not Personnel**
  - Ease of installation and manageability vital for using scale-out NAS
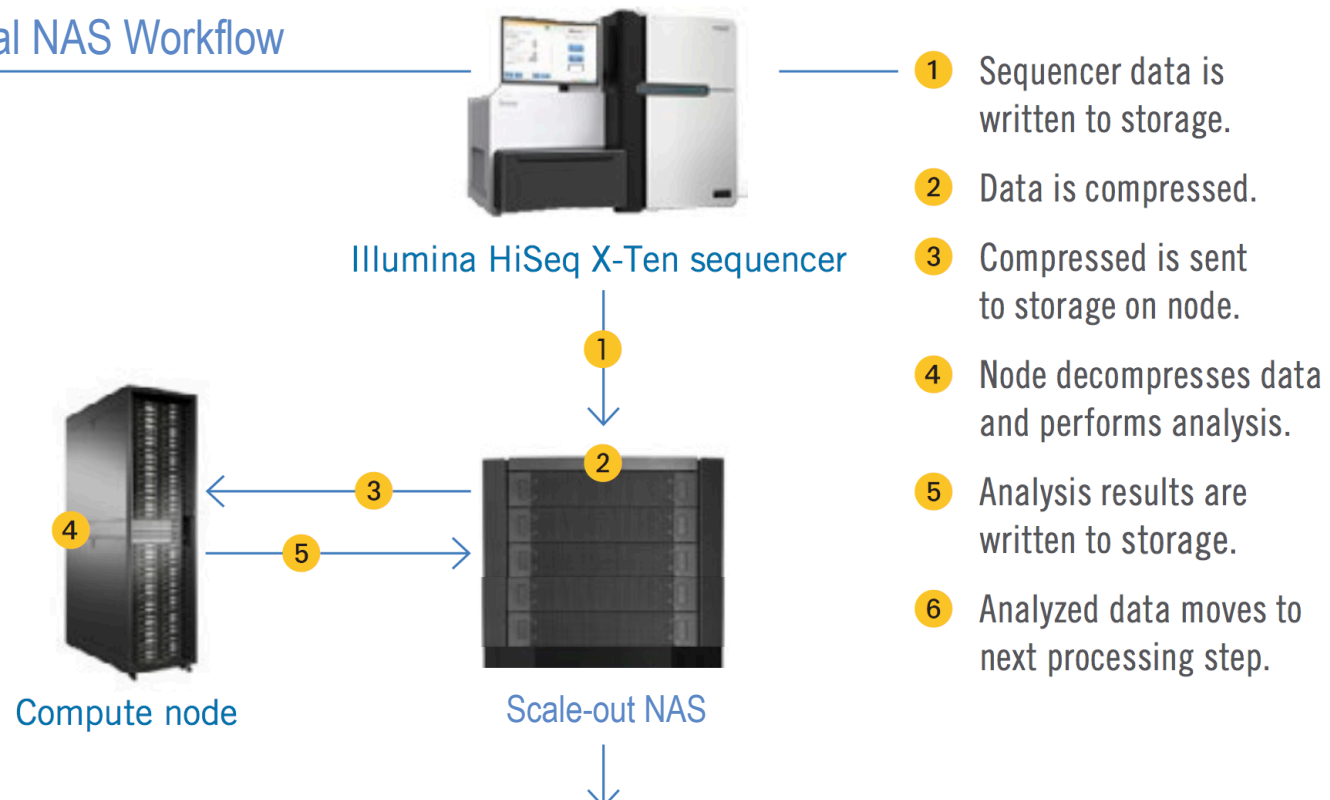
NGS WORKFLOW IMPROVEMENTS

Traditional NAS Workflow

Illumina HiSeq X-Ten sequencer

Compute node

Scale-out NAS

1. Sequencer data is written to storage.
2. Data is compressed.
3. Compressed is sent to storage on node.
4. Node decompresses data and performs analysis.
5. Analysis results are written to storage.
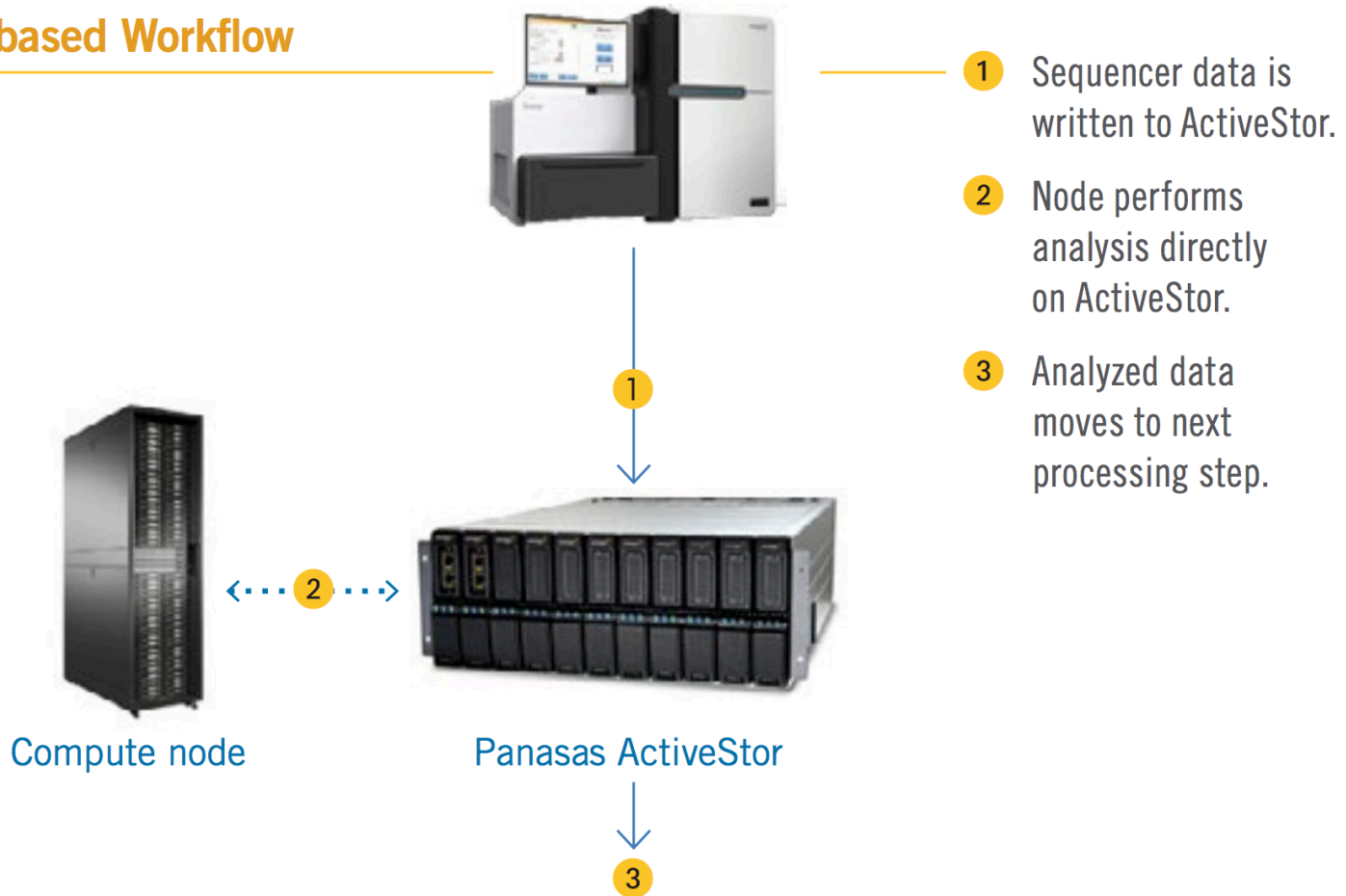6. Analyzed data moves to next processing step.

- **The traditional process of copying datasets from NAS to Compute introduces extra time, creating frustration and delay**

NGS WORKFLOW IMPROVEMENTS

panasas

Panasas based Workflow

1. Sequencer data is written to ActiveStor.

2. Node performs analysis directly on ActiveStor.

3. Analyzed data moves to next processing step.

Compute node

Panasas ActiveStor

- **The better workflow eliminates copying data to compute, reducing analysis time and frustration amongst research teams**

GARVAN'S RESULTS WITH PANASAS

- **Panasas Anchors Infrastructure Upgrade**
  - Half a petabyte of Panasas ActiveStor
  - One global namespace

- **Increased Sequencing Capacity by 50x**
  - Great fit with Illumina HiSeq X Ten and large Linux Compute Farm

- **Boosted Productivity of 80 Researchers**
  - Panasas "arrived just in time to quell a user mutiny over painfully slow application response"

- **Minimized IT Management Overhead**
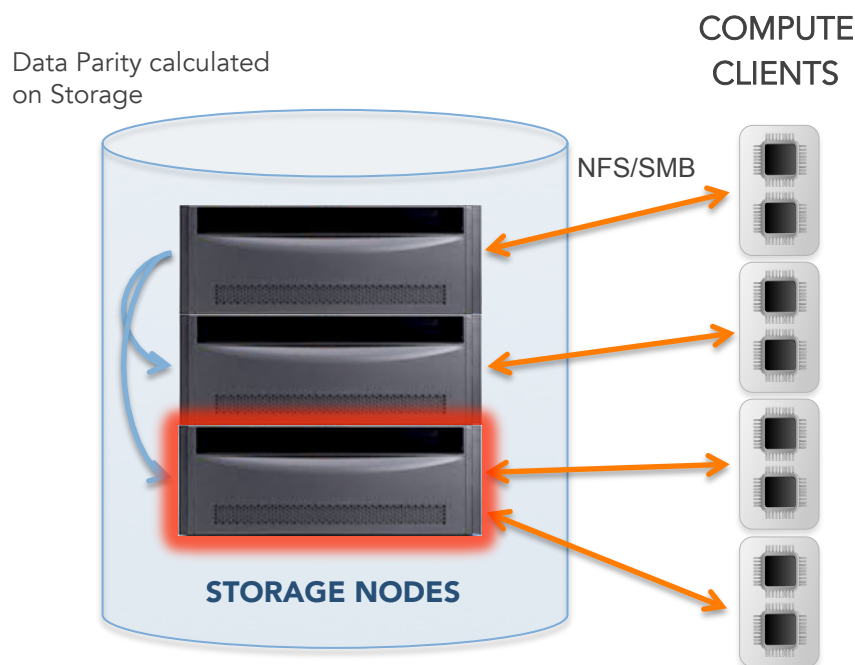 - Negligible maintenance and administration time

- **Simplified Workflows**
  - Sequencing data stayed in central repository throughout analysis
  - Saved time and bandwidth to achieve quicker results that are shared with researchers around the world
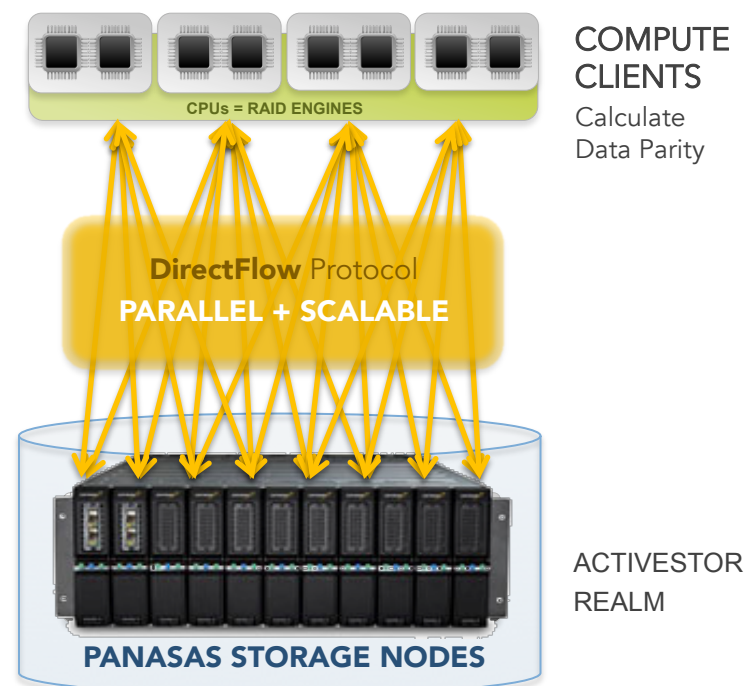
TRUE SCALE-OUT NAS ARCHITECTURE

panasas

Scale-out clustered NAS

Data Parity calculated
on Storage

COMPUTE
CLIENTS

NFS/SMB

STORAGE NODES

Performance Scale-out NAS

CPUs = RAID ENGINES

COMPUTE
CLIENTS

Calculate
Data Parity

DirectFlow Protocol
PARALLEL + SCALABLE
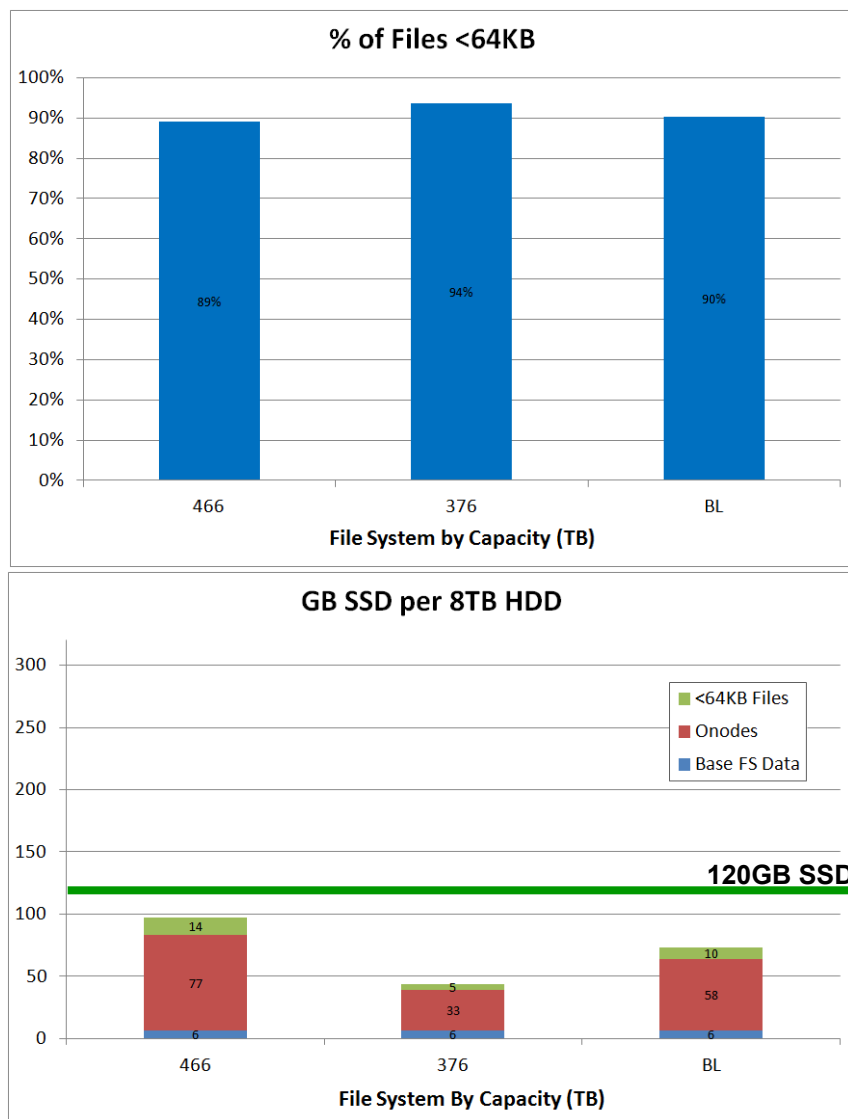
ACTIVESTOR
REALM

PANASAS STORAGE NODES

- Clients have point-to-point link to one cluster node
- I/O hotspots due to uneven distribution of clients
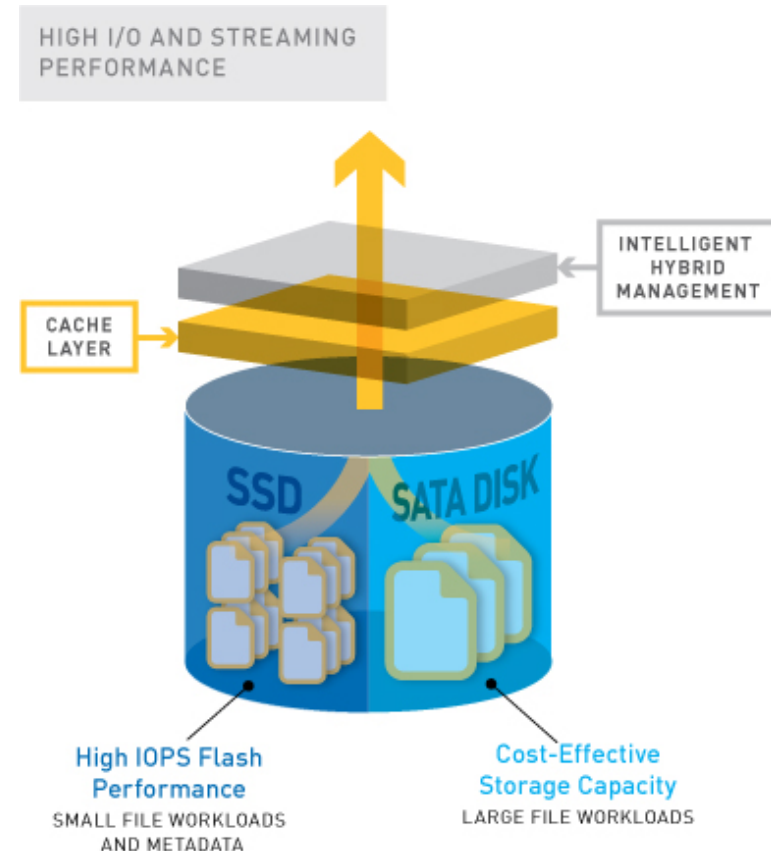- Data Parity calculated on storage – performance hit

- Each client has one-to-all connection to entire cluster
- Clients auto-balanced, wide front-end Ethernet profile
- Parity calculated by client – more cluster performance

A TYPICAL GENOMICS FILE SYSTEM

panasas

% of Files <64KB

GB SSD per 8TB HDD

- **Life Sciences researchers are typically not file systems experts**
  - You shouldn't be file system experts!

- **Instruments create a lot of tiny files, often for their own metadata**

- **On 376TB file system:**
  - 62% <=2KB file size
  - 83% <=12KB file size

- **Most scale-out NAS systems can struggle when handling small file IOPS and large sequential file throughput performance**

- **This is solved by the unique PanFS tiering algorithm optimized for Flash IOPS and SATA high-density storage**

10

HYBRID PERFORMANCE LEADERSHIP

- **Intelligent Flash-SATA Architecture**
  - Automated, multi-tier design
  - Cache + flash + SATA drives

- **Efficient Use of Storage Media**
  - Metadata & small files on flash
  - Large files on hard drives

- **Designed for the Real World**
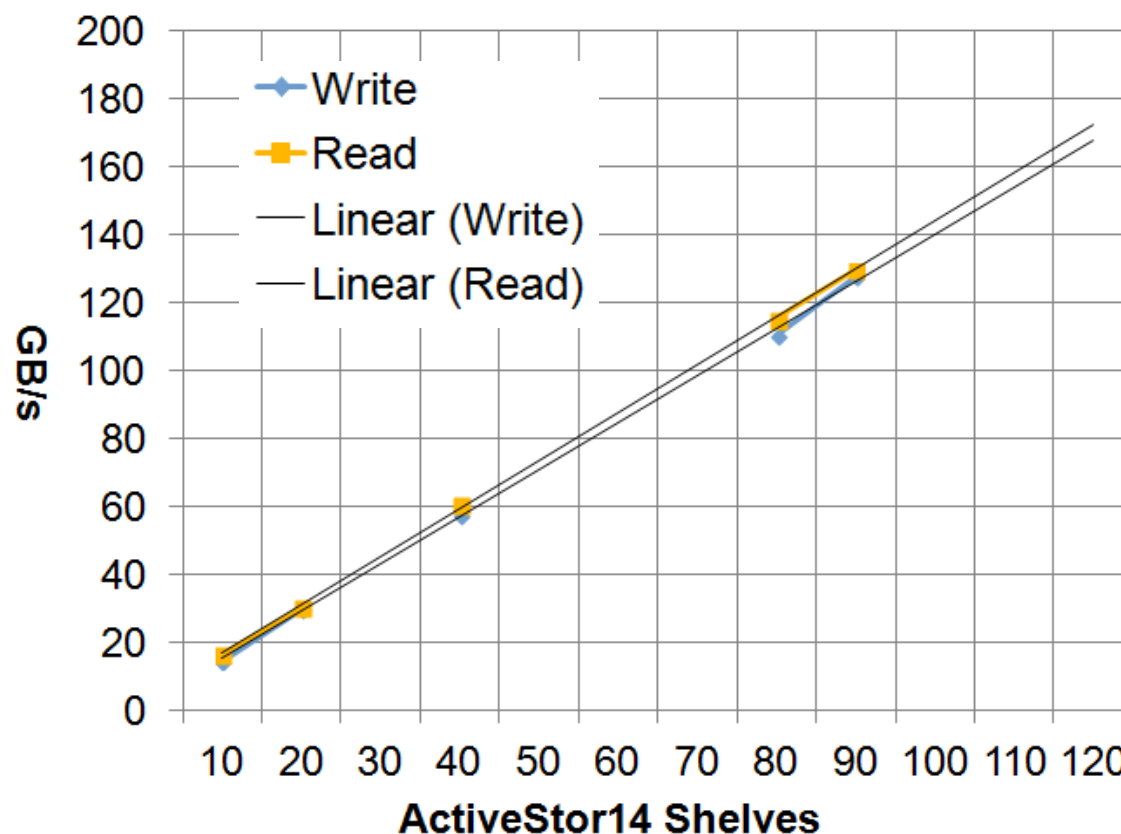 - High mixed-workload performance accelerates end-user applications, not just benchmarks

HIGH I/O AND STREAMING PERFORMANCE

INTELLIGENT HYBRID MANAGEMENT

CACHE LAYER

SSD   SATA DISK

High IOPS Flash Performance
SMALL FILE WORKLOADS AND METADATA

Cost-Effective Storage Capacity
LARGE FILE WORKLOADS

LINEAR SCALABILITY – CUSTOMER SITE

Configuration:

- 90 ActiveStor 14 shelves

- 27 director blades

- 180 compute nodes (3200 processes)

- 2 compute node ratio to ActiveStor

- **Linear Scalability as Expected!**

  - 128.6 GB/s write, 128.4 GB/s read (1.4GB/s per shelf) from 1926 hard drives

  - 1.6-1.7GB/s per shelf possible with current PanFS and 4 clients/shelf

  - Predicts up to 220GB/s from 130 shelves / 2750 drives

GARVAN INSTITUTE

Dr. Warren Kaplan
Chief of Informatics
Garvan Institute of Medical Research

"Thanks to Panasas' exceptional performance, our sequencing data stays in the central repository throughout the analysis. This streamlined workflow saves time and bandwidth, enabling us to deliver results quickly to researchers around the world."

"Panasas lives up to its promise of terrific performance with negligible maintenance and administration time."
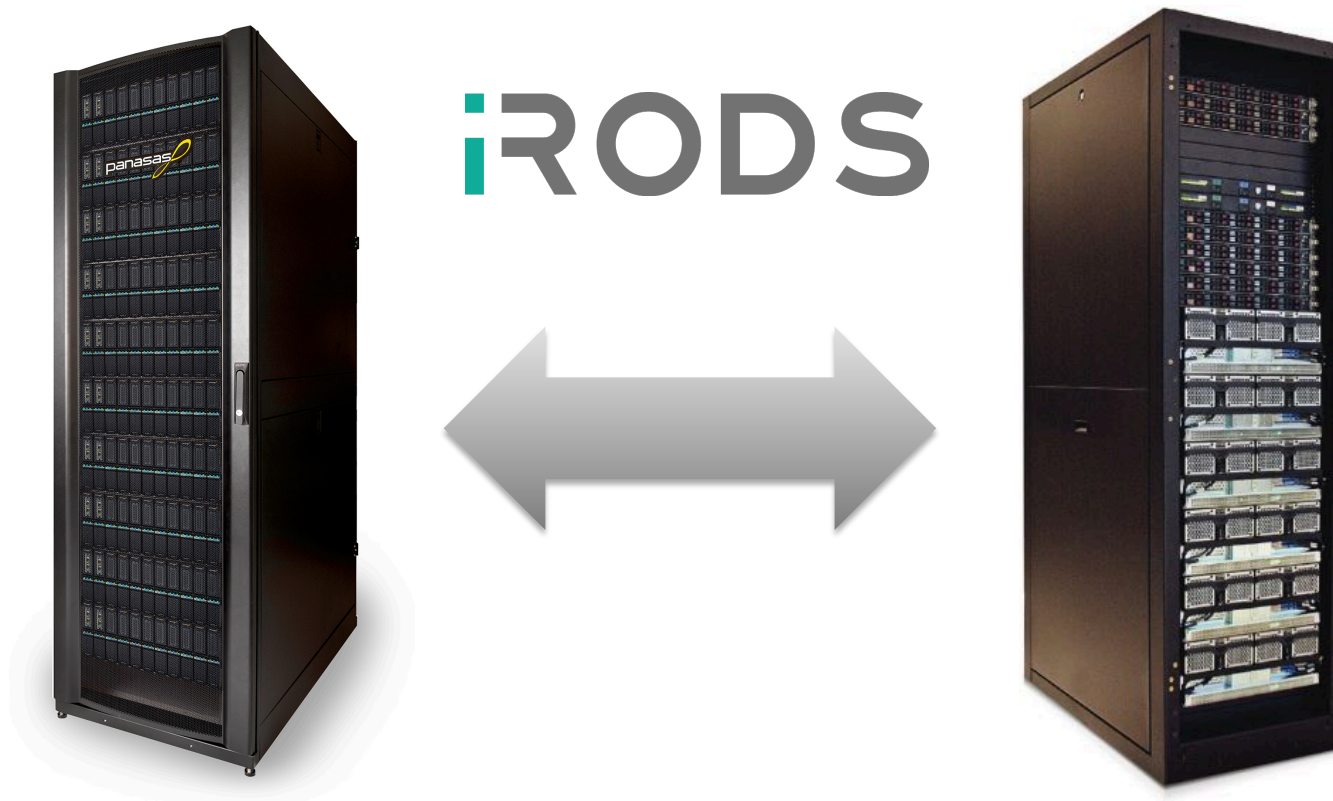
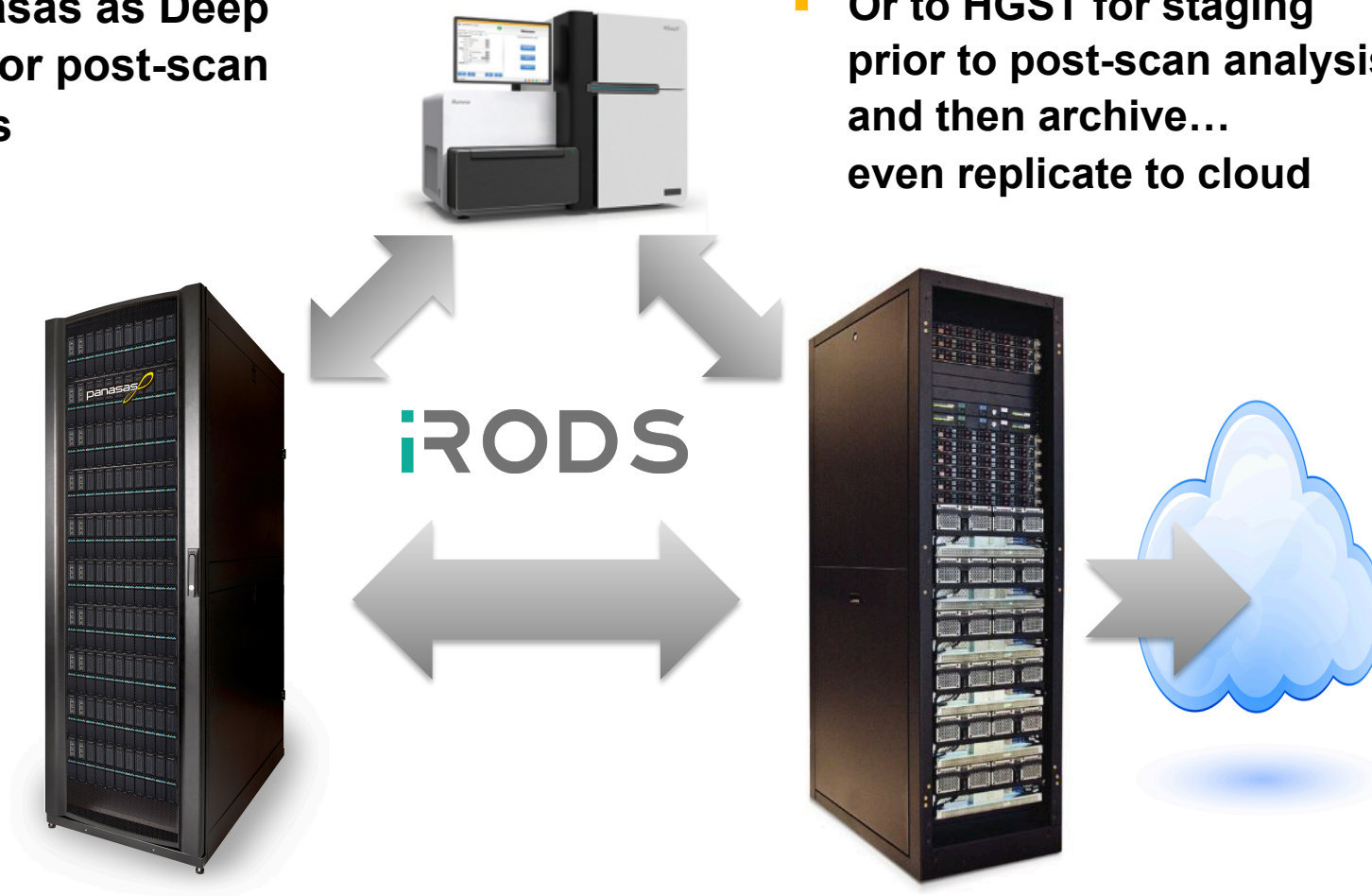# MANAGING LS DATA WITH IRODS

panasas

- **With iRODS, life science datasets are organized into *Collections***
- **iRODS Projects are managed as *Active* or *Passive***
- **iRODS Project Data is Tiered between Panasas (Active) and HGST Archive (Passive)**

iRODS

Copyright © 201   15

ILLUMINA X-TEN WITH IRODS

- **Illumina HiSeq outbound datasets managed by iRODS into collections**

- **To Panasas as Deep Cache for post-scan analysis**

- **Or to HGST for staging prior to post-scan analysis, and then archive… even replicate to cloud**

OUR GOALS WITH IRODS

- **Develop a Panasas DirectFlow Storage Resource plugin**
  - SDK exists
  - High-throughput data movement via iRODS servers

- **Coordinated Data Awareness with HGST via iRODS**

- **Support iRODS v4.3 efforts to step out of data path**

Panasas brings Good-Fast-Easy
to iRODS and HGST Active Archive

THANK YOU!

http://www.panasas.com

http://www.linkedin.com/company/panasas

http://twitter.com/#!/panasas

http://www.youtube.com/PanasasHPC