November 2013

IT Infrastructure Matters

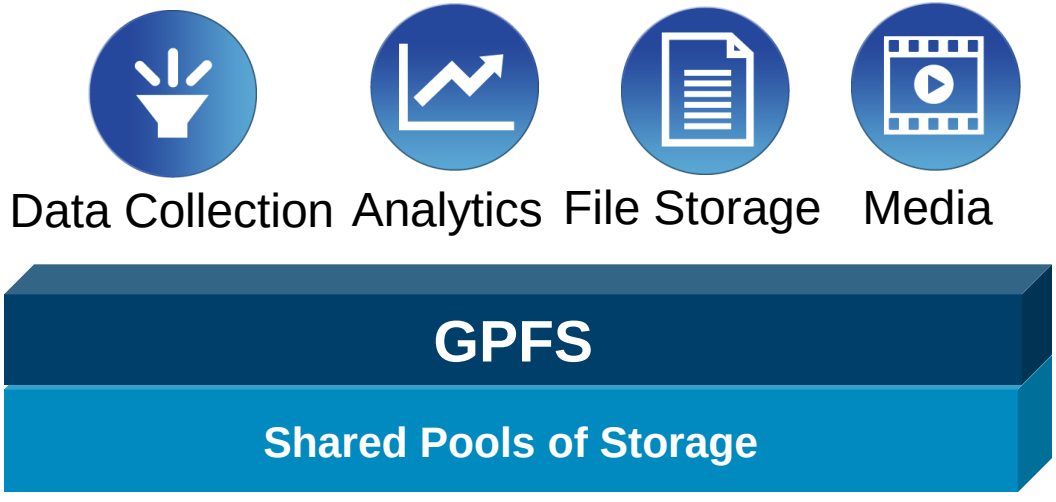
IBM General Parallel File System (GPFS™)

Status, what‘s new and what‘s coming

Agenda

- GPFS Updates – Status of new features
- Roadmap discussion
- Research Activities

High Performance Common Data Access



- Streamline Data access
- Centralize Storage Management
- Improve Data Availability

Latest news on recently released features

GPFS Updates

IBM

GPFS Scalability Updates

Extreme Scalability

File system

- Customer with 19 PB file system
- Others 10PB and larger
- 400 GB/sec IOR

Number of nodes

- 1 to 16384

GPFS for big data

World-wide global namespace

Solution:
Active File Management

Streaming Performance

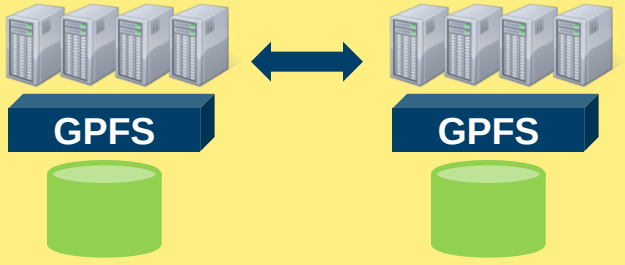
Solution:
GPFS Storage Server

Map Reduce Analytics

Solution:
File Placement Optimizer

Sharing Data Across Your Organization

IBM introduces GPFS providing concurrent file system access from multiple nodes



GPFS adds multi-clustering, expanding the global namespace by connecting multiple sites

AFM takes global namespace truly global by automatically managing asynchronous replication of data

1993 2005 2011

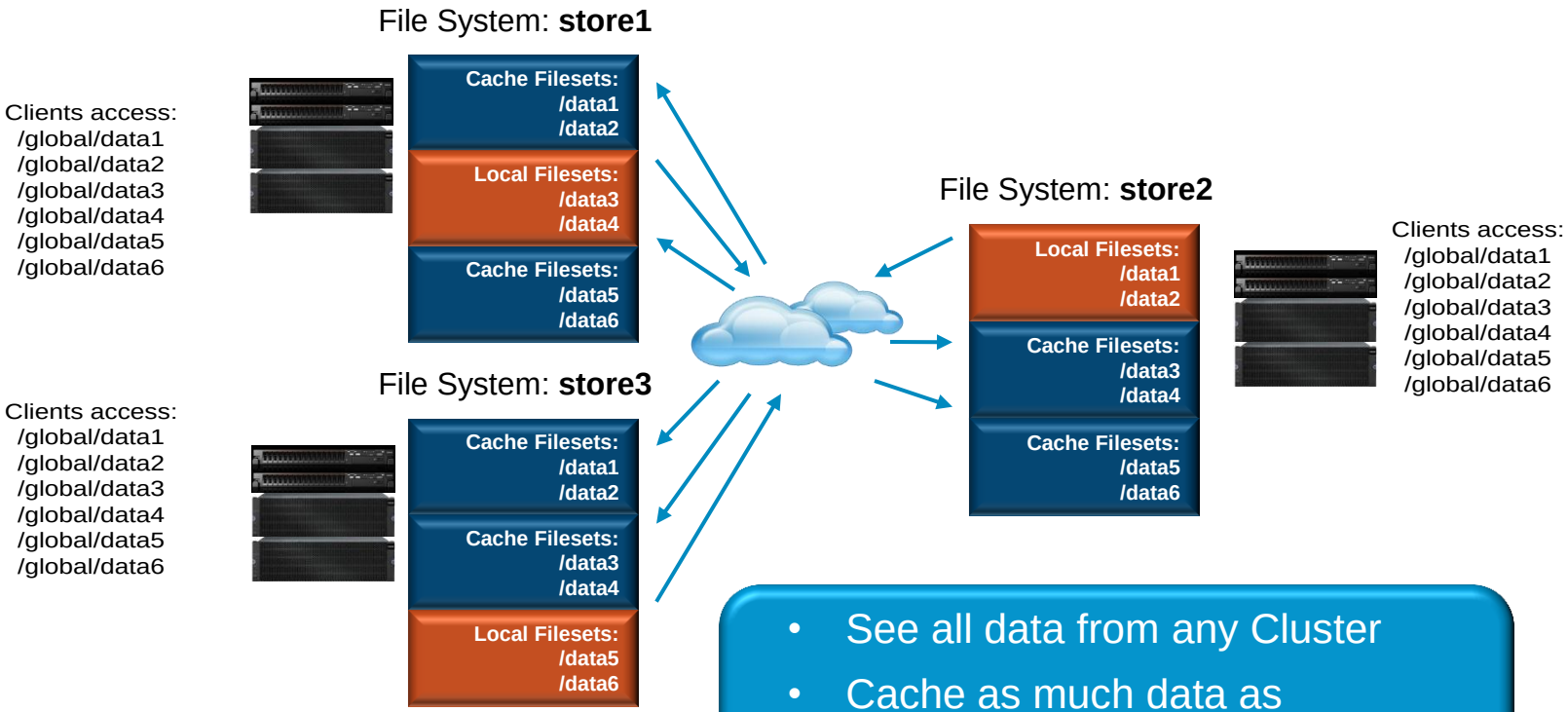
AFM Calendar

- GA with GPFS 3.5 (April 2012)
  - Read only
  - Single Writer
  - Local Update
- GPFS 3.5 TL3 (June 2013)
  - Independent writer

How is AFM being used?

- Read only
  - Acceleration of NFS stored data
- Local Update
  - Development systems
- Independent writer (GA planned June 2013)
  - Basic HA

Global Namespace



- See all data from any Cluster
- Cache as much data as required or fetch data on demand

LTFS Support in GPFS

- **Digital archives need and want the Value Proposition of Tape**:
  - $/GB – lowest cost storage
  - Watt/GB – green storage
  - Portability – ability to manage archive outside system
  - Investment protection – LTO has an 8 generation roadmap (up to a 32TB cartridge (compr.)
- **But - Inhibitors to use tape:**
  - Proprietary tape applications
  - Non-self describing data formats
  - Import/export & distribution of tapes in archive is difficult due to proprietary tape applications
- **SOLUTION: LTFS addresses the inhibitors and unlocks the value proposition of tape for digital archives**
  - Open, non-proprietary tape format
  - Self-describing data structure on cartridge
  - File system support on Linux, Mac, Windows provides:
  - Distribution and cross platform interchange
  - Enables transition to integrated file based tape/disk storage systems

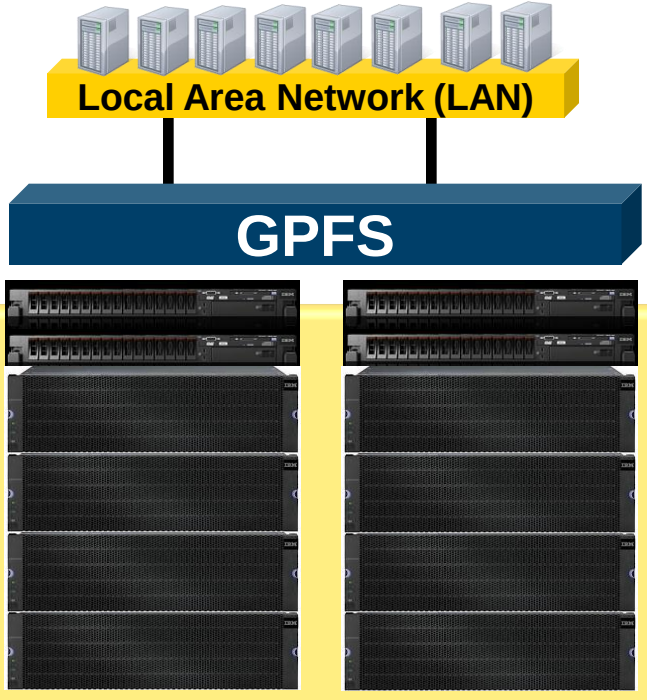
Why Build GPFS Storage Server?

- Disk rebuilding is a fact of life at petascale level
  - With 100,000 disks and an $MTBF_{disk}$ = 600 Khrs, rebuild is triggered about four times a day
  - 24-hour rebuild implies four concurrent, continuous rebuilds at all times
- Traditional, 1-fault-tolerant RAID-5 is a non-starter
  - Mean Time to Data loss (MDDTL) event every 6.5 days
- 2-fault-tolerant declustered RAID (8+2P) may not be sufficient
  - MTTDL ~ 7 years (simulated, $MTTF_{disk}$=600Khrs, Weibull, 100-PB usable)
- 3-fault-tolerant declustered RAID (8+3P) is 400,000x better
  - MTTDLL ~ $3x10^6$ years (simulated, $MTTF_{disk}$=600Khrs, Weibull, 100-PB usable)

Improve Reliability and Lower Costs

Benefits of GSS

- 3 years maintenance and support
- Improved storage affordability
- Delivers data integrity, end-to-end
- Faster rebuild and recovery times
- Reduces rebuild overhead by 3.5x



Features

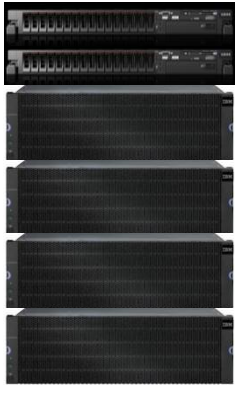
- De-clustered RAID (8+2p, 8+3p)
- 2- and 3-fault-tolerant erasure codes
- End-to-end checksum
- Protection against lost writes
- Off-the-shelf JBODs
- Standardized in-band SES management
- SSD Acceleration Built-in

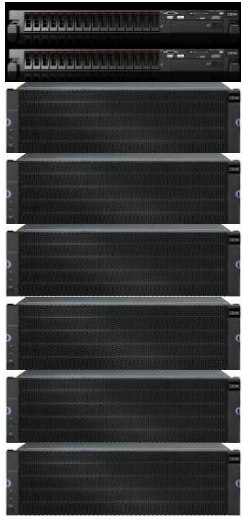
A Scalable Building Block Approach to Storage

Complete Storage Solution
Data Servers, Disk (SSD and NL-SAS), Software, Infiniband and Ethernet

x3650 M4

"Twin Tailed" JBOD
Disk Enclosure





Model 24: Light and Fast

- 4 Enclosures 20U
- 232 NL-SAS 6 SSD
- 10 GB/Second

Model 26: HPC Workhorse

- 6 Enclosures 28U
- 12 GB/Second
- 348 NL-SAS 6 SSD

High Density HPC Options

- 18 Enclosures
- 2 - 42u Standard Racks
- 1044 NL-SAS 18 SSD
- 36 GB/Second

Beyond GPFS 3.5

GPFS Roadmap

IBM

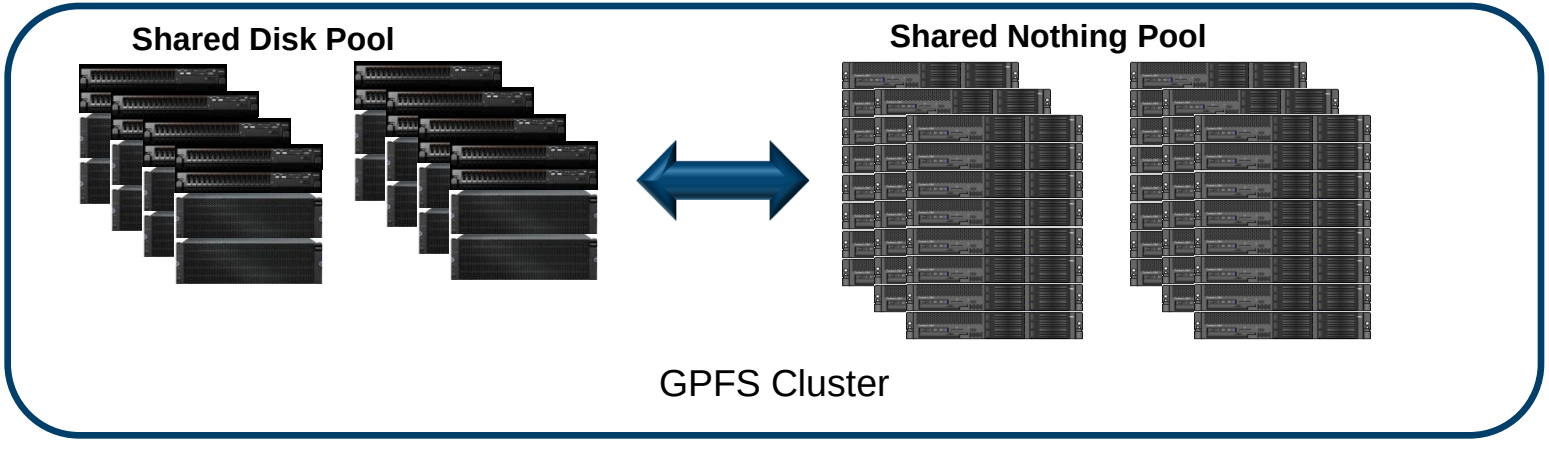
GPFS Storage Server

- GUI
  - Configuration
  - Performance
  - Monitoring
- Hardware Changes
  - New servers and cards
  - Smaller Trays
  - SSD, SAS and 4TB NL/SAS
- Software Enhancements
  - Enclosure protection

File Placement Optimizer (FPO) as a pool

- Shared Nothing Architecture
- Map Reduce Workloads (Hadoop)
- Archive Data

- Chunks
- Data Locality
- Efficient configurable replication
- Fast recovery



Local Read Only Cache

- Overflow pagepool to local storage
  - SSD
- Benefits
  - More memory for applications
  - Automatic management of local storage

Encryption

- Native encryption support for GPFS
- Addresses critical requirements
  - Encryption of data at rest
  - Multi-tenancy key enabler for cost efficient solutions
  - Secure deletion mandatory in a number of domains (FI, HR, healthcare, banking...)

Other new features

- RDMA over Ethernet
- Fileset snapshot restore
- 4k Disk sector support

Continuing innovation

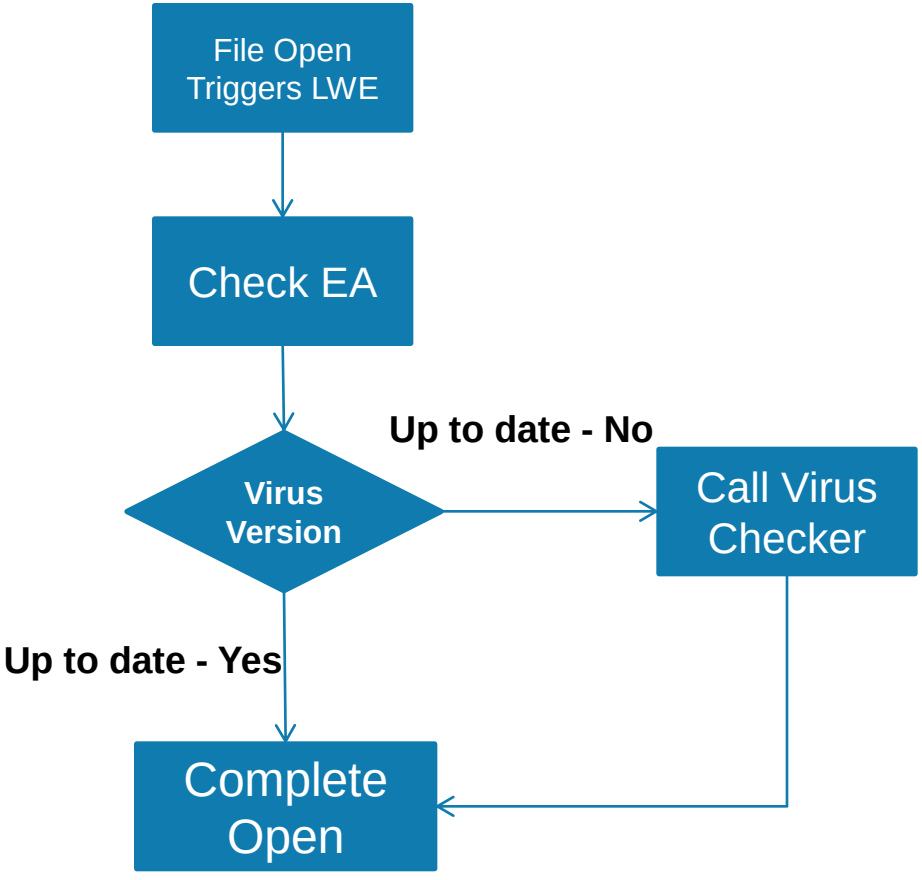
Research Involving GPFS

Burst Buffer

- Starting to research burst buffer implementations
- Multiple approaches
  - Information Lifecycle Management
  - Active File Management

Light Weight Events

- Local Action
- Triggers on different events
- Integrates with policies



Customer Examples

GPFS is used across many industries and data types

Digital media

- Broadcast and podcasts
  - Social networking
  - Streaming content
  - Media repositories
- Interactive and gaming

Analytics

- IBM Watson
- Financial analytics
- Inventory Control

Workflow

- Data Sharing
- Gene Sequencing
- ETL

And much more…

- Engineering design
- Gene Sequencing
- Seismic data processing
- Scientific research

Oil and Gas Customer Migrated from Lustre to GPFS

- **Customer goals**
  - Expand compute capacity by 10 times
  - Improve storage system stability
  - Scale storage to support additional compute nodes
  - Migrate to newer storage technology with no disruption
  - Better manage storage infrastructure
- **Customer had problems with Lustre installation**
  - Existing system was too **fragile**, requiring too much “firefighting”
  - Unpredictable and problematic upgrades
  - Poor understanding of actual space utilization
  - Lack of a consistent and efficient archive policy
- **New GPFS Solution**
  - Approximately 2.5 PB of active storage in place
  - Better understanding of size distribution and access
  - Migration of approximately 400 TB of Lustre data complete with disk hardware repurposed

Vestas Wind Systems

Maximize power generation and durability in its wind turbines with HPC

The Opportunity

This wind technology company relied on the World Research and Forecasting modeling system to run its **turbine location algorithms, in a process generally requiring weeks and posing inherent data capacity limitations.** Poised to begin the development of its own forecasts and adding actual historical data from existing customers to the mix of factors used in the model, Vestas needed a solution to its Big Data challenge that would be faster, more accurate, and better suited to the its expanding data set.

Solution Components

- IBM Technical Computing – General Parallel File System
- IBM InfoSphere® BigInsights Enterprise Edition
- IBM System x ®, iDataPlex ®

What Makes it Smarter

Precise placement of a wind turbine can make a significant difference in the turbine's performance–and its useful life. In the competitive new arena of sustainable energy, winning the business can depend on both value demonstrated in the proposal and the speed of RFP response. Vestas broke free of its dependency on the World Research and Forecasting model with a powerful solution that sliced weeks from the processing time and more than doubled the capacity needed to include all the factors it considers essential for accurately predicting turbine success. **Using a supercomputer that is one of the world's largest to-date and a modeling solution designed to harvest insights from both structured and unstructured data, the company can factor in temperature, barometric pressure, humidity, precipitation, wind direction and wind velocity at the ground level up to 300 feet, along with its own recorded data from customer turbine placements**. Other sources to be considered include global deforestation metrics, satellite images, geospatial data and data on phases of the moon and tides. The solution raises the bar for due diligence in determining effective turbine placement.

Real Business Results

- **Reduces from weeks to hours the response time for business user requests**
- Provides the capability to analyze ALL modeling and related data to improve the accuracy of turbine placement
- Reduces cost to customers per kilowatt hour produced and increases **the precision of customer ROI estimates**

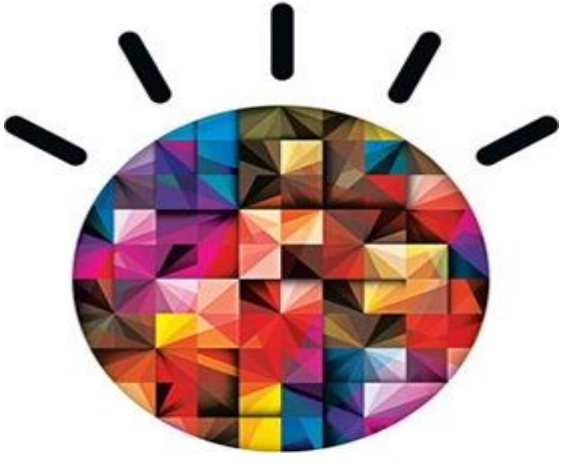
"Today, more and more sites are in complex terrain. Turbulence is a big factor at these sites, as the components in a turbine operating in turbulence are under more strain and consequently more likely to fail. Avoiding these pockets of turbulence means improved cost of energy for the customer."

IBM Technical Computing

- Anders Rhod Gregersen,
Senior Specialist, Plant Siting & Forecasting

Why GPFS?

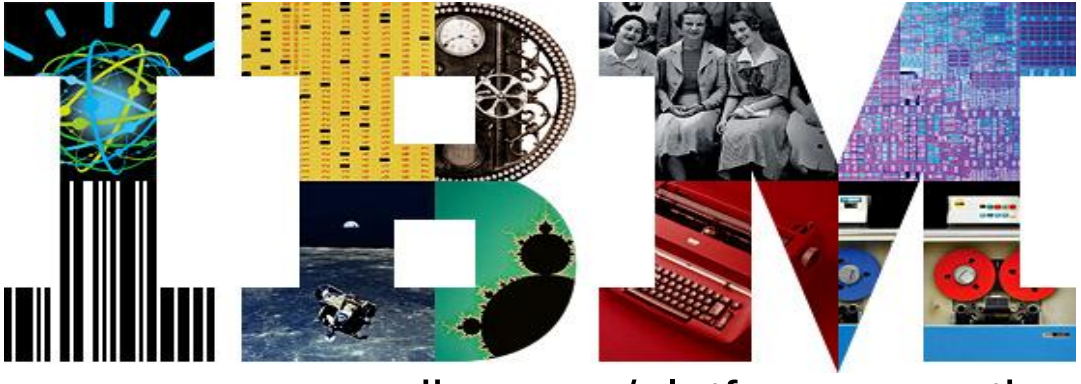
- High Performance
- Easy To Manage
- Advanced Features



Smart Storage

For more information

www.ibm.com/systems/software/gpfs



www.ibm.com/platformcomputing

Email gpfs@us.ibm.com or contact your IBM Representative