IBM Storage for a Smarter Planet …

Jon Laycock, XIV Technical Lead
IBM XIV
laycockj@uk.ibm.com 07764 354230

AGENDA

- Introduction

- XIV - Background

- XIV And Oracle

- Wrap-up | Q&A

The Role of IBM Storage in Building a Smarter Planet

IBM's Smarter Planet vision

The world is flatter and smaller. Now it's becoming smarter.

The world is generating 15 petabytes of data daily

Two major storage initiatives

Storage Optimization

Improve performance up to **300%**

Retention and Archive

Reduce physical storage up to **50%**

IBM Storage
Grow Capacity without Complexity

IBM Storage Optimization solutions help clients meet new performance, scalability, and reliability requirements.

- Grid and Multi-tier Storage Systems
- Storage Virtualization
- Smarter Movement and Management of Information
- Remote Mirroring

IBM Information Retention and Archiving solutions help clients reduce the cost of storing and protecting data throughout its useful life

- Data Reduction
- Archive
- Backup
- Tape and Virtual Tape Systems

IBM XIV Storage System – Key Business Value

High-End System:

- ✓ Performance
- ✓ Reliability
- ✓ Scalability
- ✓ Features

Unmatched TCO:

- ✓ High capacity utilization
- ✓ Low power consumption
- ✓ Minimal administration
- ✓ High responsiveness

Single Tier:

- ✓ High-end system, mid-range TCO
- ✓ Avoids ILM and HSM overhead

5 Key Attributes for Enterprise Storage

Performance

- Consistent performance under all conditions; eliminating hotspots and delivering consistent high performance even during rebuilds

Reliability

- Business data is more critical than ever; no tolerance for down time or disruption of service; greater than "5 nines" is the new standard

Functionality

- Tier 1 functions (e.g., replication, thin provisioning, snapshots) that scale with no performance penalty are built-in, out of the box

Manageability

- Total system virtualization; emphasis on extreme, industry-leading ease of use and management

Cost

- Superior cost containment and control; best in class TCO with lower acquisition cost and considerably lower on-going management cost

XIV is all of these key attributes with <u>massive scalability</u>

Where did XIV come from??

- Moshe Yanai, XIV CEO (former EMC CTO & designer of Symmetrix) designed XIV

    - While at EMC identified that the dual controller architecture had limitations

        - Controllers, disks, interconnects

        - Scale by more and more powerful controllers

            - Faster CPUs, bigger cache

        - Disks are getting bigger, not faster

            - Year 2000: 18.4GB 15k rpm, 2010: 3TB (largest), 15k fastest.

- New architecture design required

    - MY proposed grid based architecture

    - Rejected by EMC

- MY considered future architecture should be grid, rather than dual controller approach

    - Departed EMC to develop grid based approach

    - XIV (founded 2002)

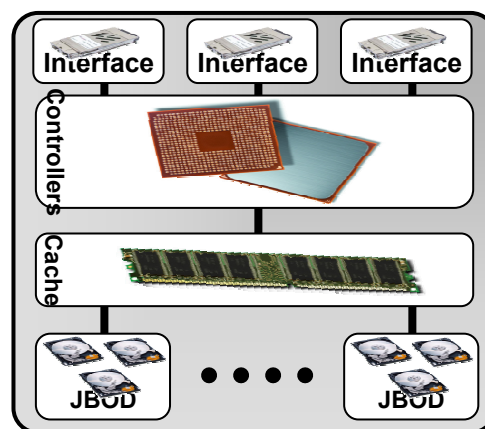# Storage Architecture Revolutions

1970

- Mainframe
- Monolithic
- Gates design
- Very expensive

- Complex
- Downtimes
- On site technician
- Manual

1990

- Cluster architecture
- Tightly coupled
- Custom HW design
- Expensive components

- Long, complex development cycles
- System exposed on failures (the human factor)
- Complex reactive service
- Requires tuning for optimal performance

2010

- Scalable grid architecture
- Node independent
- Commodity H/W building blocks
- Off-the-shelf low cost components
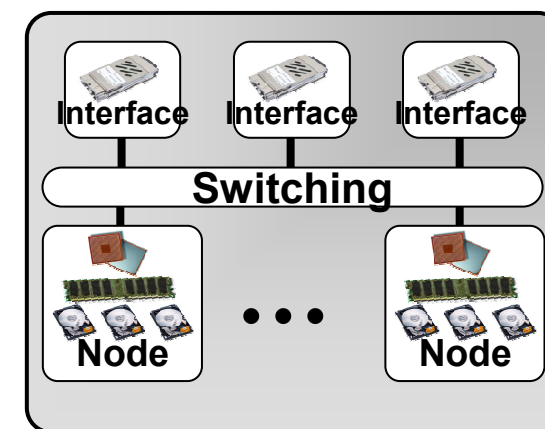
- Fast, efficient development cycles
- WEB resiliency
- Self healing
- Data layout eliminate hotspots

Who is XIV and why did IBM Acquire XIV?

- Part of a $1 billion strategy to displace EMC/HP/HDS and lead the market

- IBM identified that the current disk architecture has limitations

    - A new approach was needed

    - Carried out an evaluation of several "new" technologies

    - XIV was one of the technologies and organisations evaluated

- XIV (founded 2002)

    - Acquired by IBM on 31/12/2007

- Moshe Yanai, XIV CEO (former EMC CTO & designer of Symmetrix) designed XIV

Escalating Complexity of Traditional Storage

- **Disk Configurations to Manage**
  - 1-6 Types and Sizes of Disk
  - Flash/SSD?
  - Replica Pool
  - Sparing & rebuild time
  - Many RAID groups / Different RAID Levels -

- **Data Center Efficiency**
  - Power
  - Space
  - Utilization

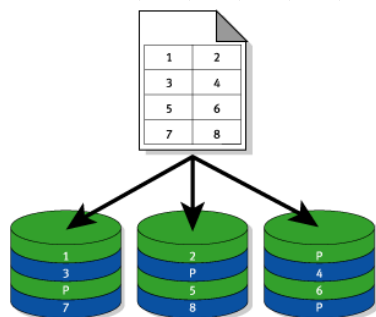- **Responsiveness / Cost to Business Units**

This results in THOUSANDS of components to manage in a traditional array!
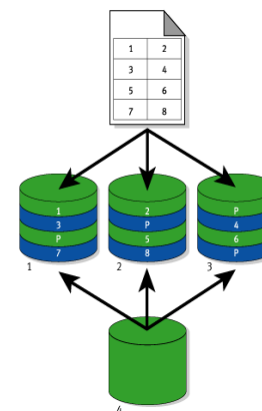
Complexity = COST = *Diminished Service*

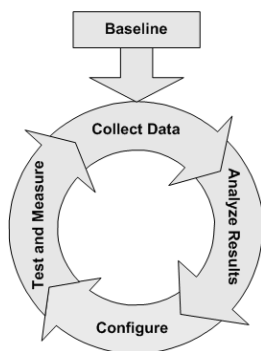## With XIV, manage storage capacity - NOT technology!

What's not required with XIV….

- ## RAID… 0,1,3,4,5,6,xx

- ## Performance Tuning

- ## Hot-Spares… and how many!

- ## Multi-Tier Storage with XIV

What is XIV storage?

- **Fully virtual SAN based on grid design**
  - Data distributed over all grid modules/drives
  - Self-healing/self-tuning – no hotspots
  - Non-disruptive hardware & software service
  - Supports all major open systems & Linux on System z
  - Sold as <u>usable</u> capacity
  - Each grid node comprised of CPU + Cache + 12 disks

- **Ultra fast rebuild times:**
  - Recover from a 1 TB drive failure in 40 minutes or less for most workloads

- **Dynamically add new capacity; autonomic load re-balancing**

- **New capacity added seamlessly**

- **Provision new storage in < 1 minute**

AVAILABILITY

Five 9s+

Multiple levels of redundancy

Autonomous self-healing

Ultra-fast fault recovery
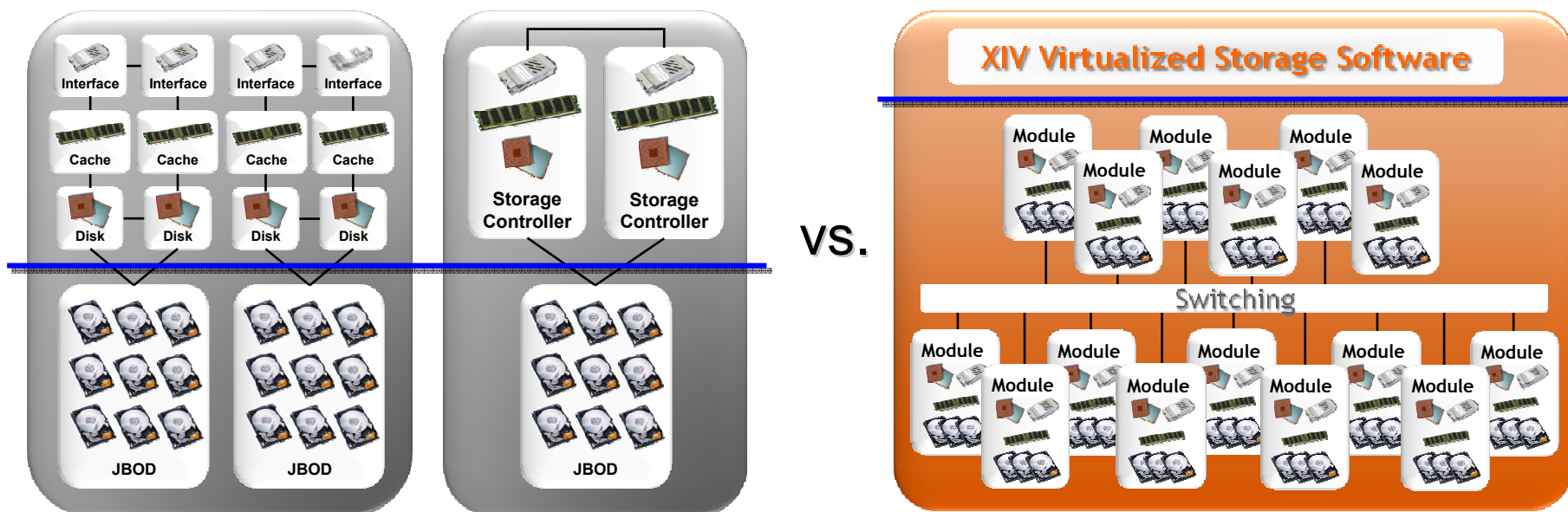
PERFORMANCE

Self-tuning

Automatic load balancing eliminates bottlenecks

Performance scales with capacity

EASE OF USE

Provision new storage in <1 minute

Dynamically resize/adapt

Easily managed by junior staff

Drive out cost while exceeding existing SLAs.

How is XIV different?

VS.

XIV Virtualized Storage Software

Capacity Only Vs Every Dimension

Monolithic Vs Grid

Custom Vs Commodity Components

Centralized Processing Vs Massive Parallelism

Grid - Explained

1941: The "Allegheny" 2-6-6-6

110,200 lbs of work power

1946: A Better Approach – The Alco FA-2

61,000 lbs **x 5 = 305,000 lbs of work power**

1:41.09

World record 800m

David Rushida, 2010
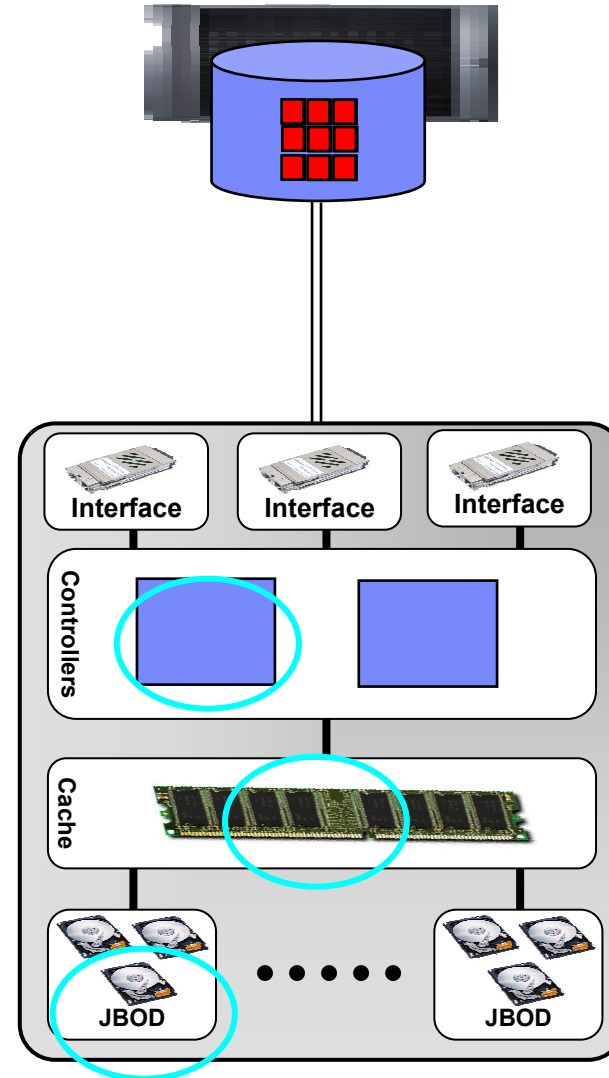
(101.09s)

… or 8 x 100m?

10.03s

Last place, 100m

Darvis Patton, 2008

9.69s (WR) Usain Bolt, 2008

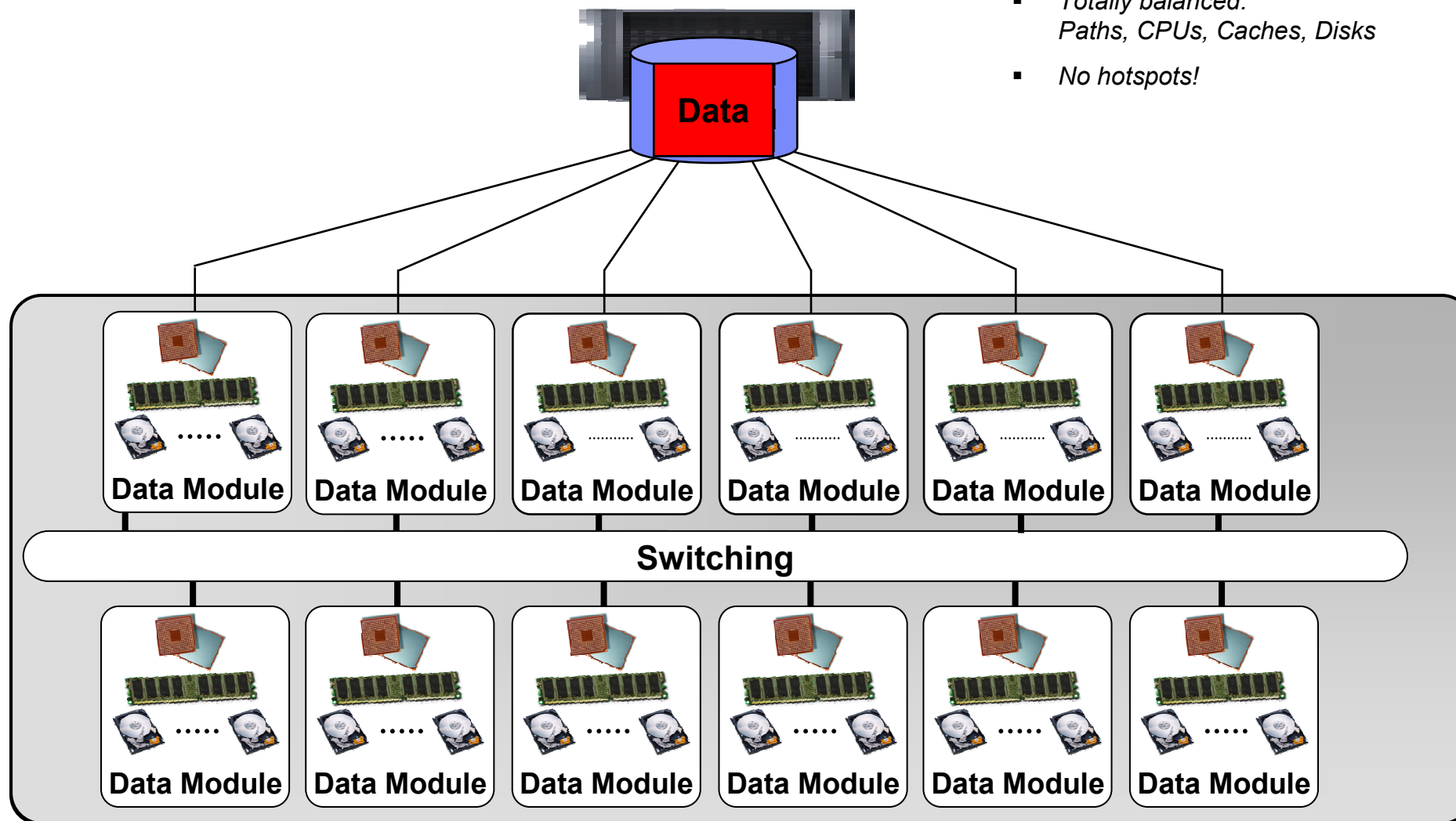# Traditional Dual Controller I/O

- IOs must pass single or dual controller

- No parallel IO -> needs fast Disks!

- Host spots most likely to occur

Interface **Interface** **Interface**

Controllers

Cache

JBOD • • • • • **JBOD**

XIV's Parallel Architecture

- All IO *parallel !*

- *Totally balanced:*
  Paths, CPUs, Caches, Disks

- *No hotspots!*

Data

| Data Module | Data Module | Data Module | Data Module | Data Module | Data Module |

Switching

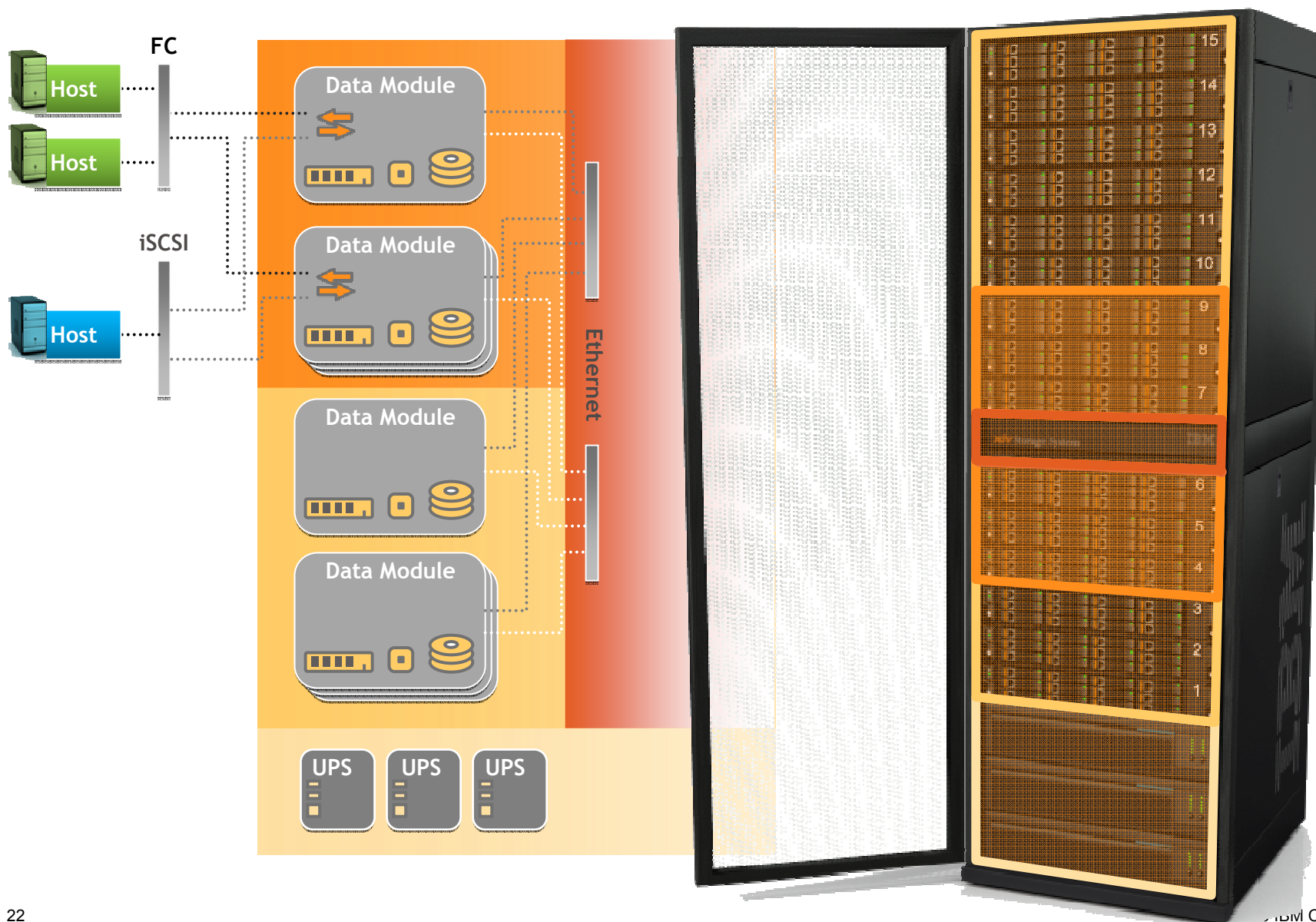| Data Module | Data Module | Data Module | Data Module | Data Module | Data Module |

IBM XIV Architecture: Scale Out

Design principles:

- Massive parallelism
- Granular distribution
- Off-the-shelf components
- Coupled disk, RAM and CPU
- User simplicity

Interface	Interface	Interface		Interface		Interface

Switching

| Data Module	Data Module	Data Module	Data Module	Data Module	Data Module	Data Module
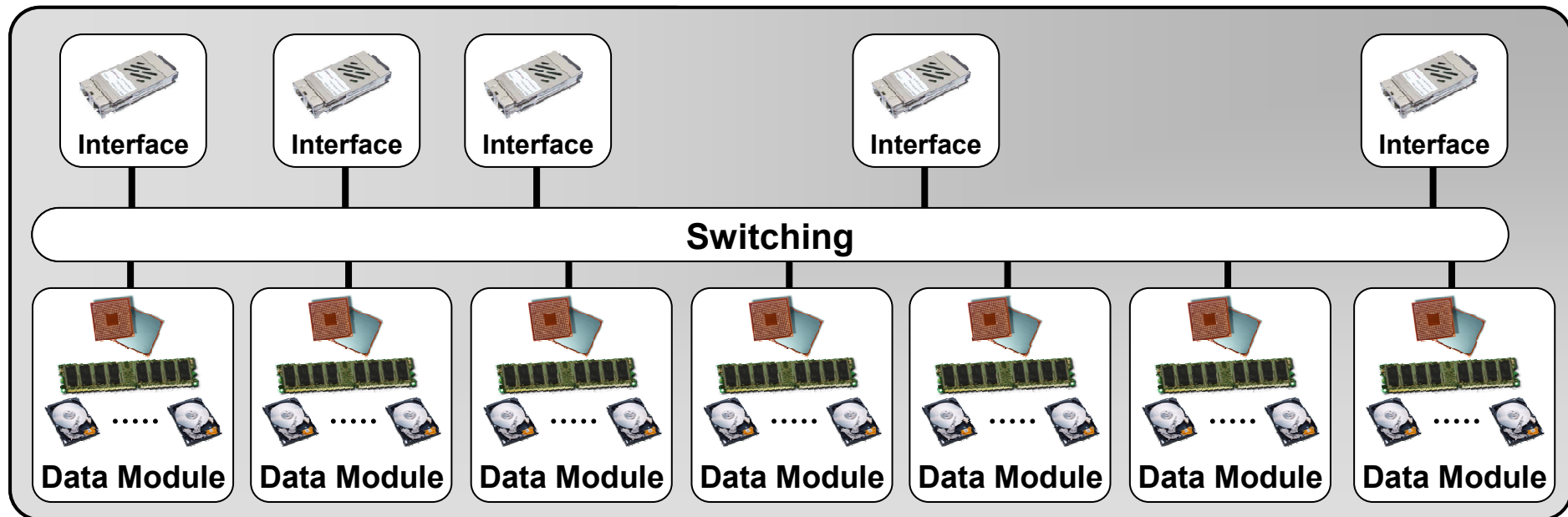
XIV - Explained

IBM XIV Storage System Distribution Algorithm

- Each volume is spread across all drives

- Data is "cut" into 1MB "partitions" and stored on the disks

- XIV algorithm ... in the system pseudo-rando...

> **XIV disks behave like <u>connected</u> <u>vessels</u>, as the distribution algorithm aims for <u>constant disk equilibrium</u>.**
>
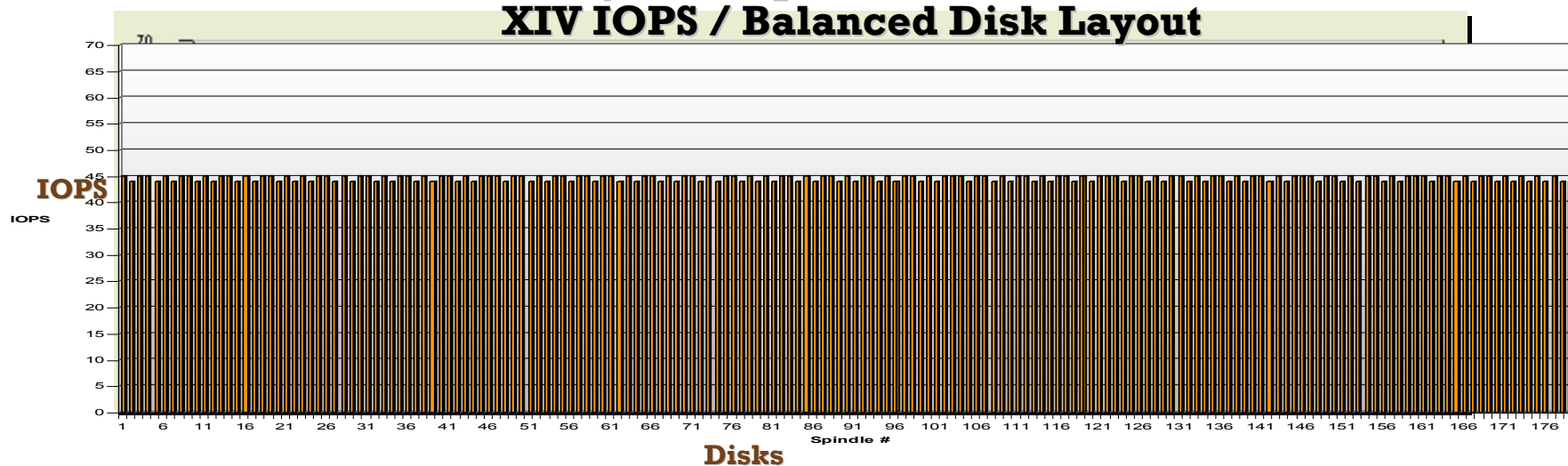> **Thus, IBM XIV's Storage overall <u>disk usage</u> could approach 100% utilization when loaded**

Interface   Interface

Data Module   Data Module   Data Module
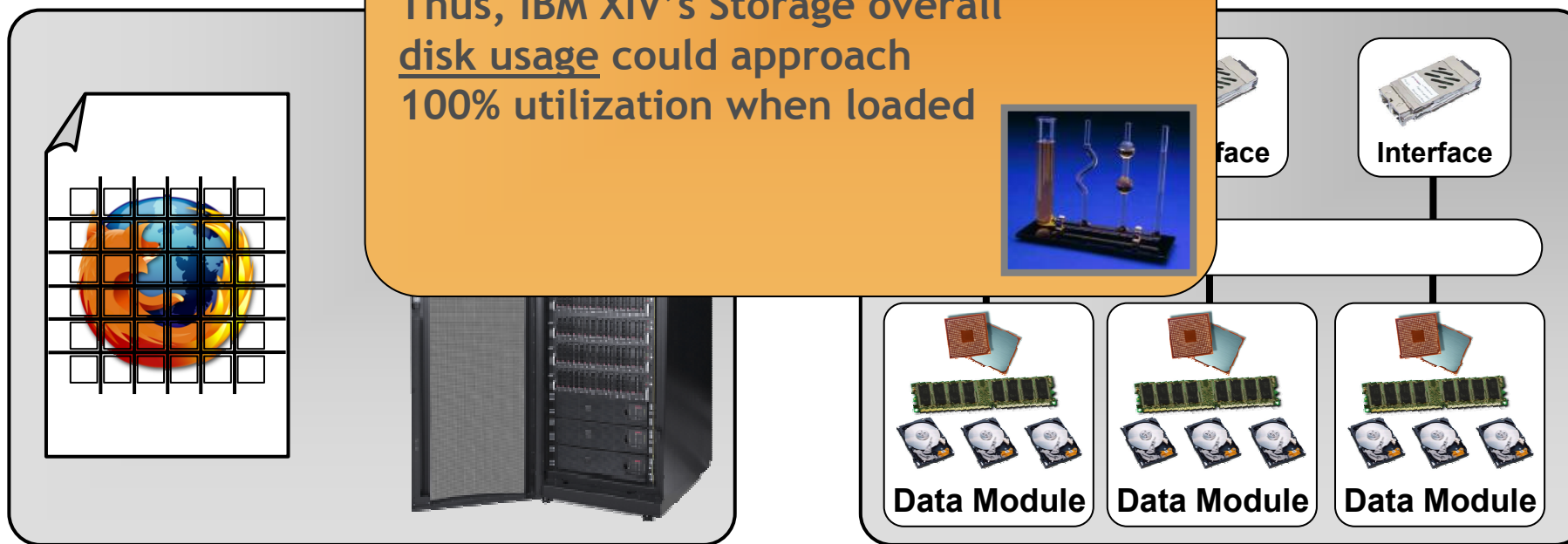
- The result?   We go from this…

Traditional Arrays IOPS profile: Unbalanced Disk
XIV IOPS / Balanced Disk Layout

IOPS

Disks

IBM XIV Distribution Algorithm on System Changes

- ## Data distribution only changes when the system changes
  - Equilibrium is kept when new hardware is added
  - Equilibrium is kept when old hardware is removed
  - Equilibrium is kept after a hardware failure

Data Module 1
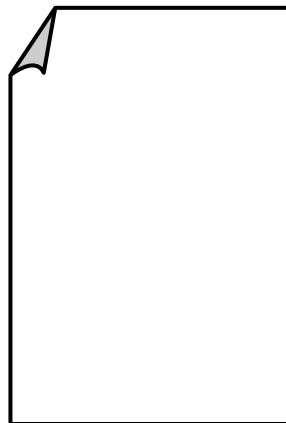
Data Module 2

Data Module 3

IBM XIV Distribution Algorithm on System Changes

- **Data distribution only changes when the system changes**
  - Equilibrium is kept when new hardware is added
  - Equilibrium is kept when old hardware is removed
  - Equilibrium is kept after a hardware failure

Data Module 1

Data Module 2

Data Module 3

Data Module 4
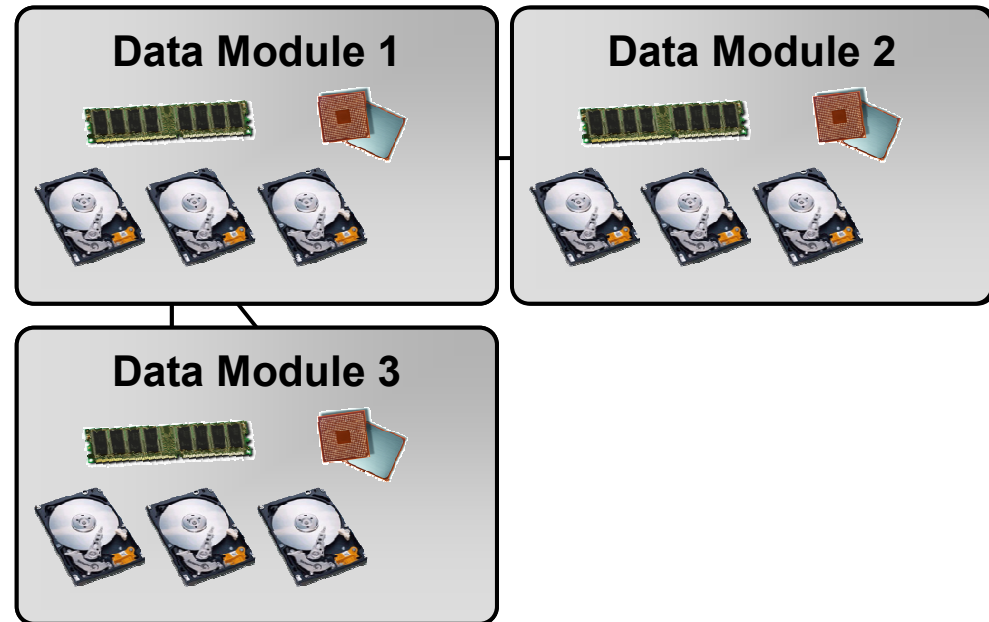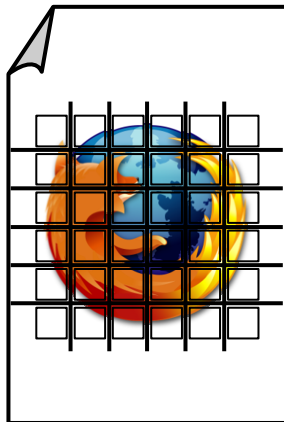
[hardware upgrade]

IBM XIV Distribution Algorithm on System Changes

- Data distribution only changes when the system changes
  - Equilibrium i
  - Equilibrium i
  - Equilibrium i

> **The fact that distribution is _full_ and _automatic_ ensures that all spindles join the effort of data re-distribution after configuration change.**
>
> **Tremendous performance gains are seen in recovery/optimization times thanks to this fact.**

Module 2

Data Module 3

Data Module 4
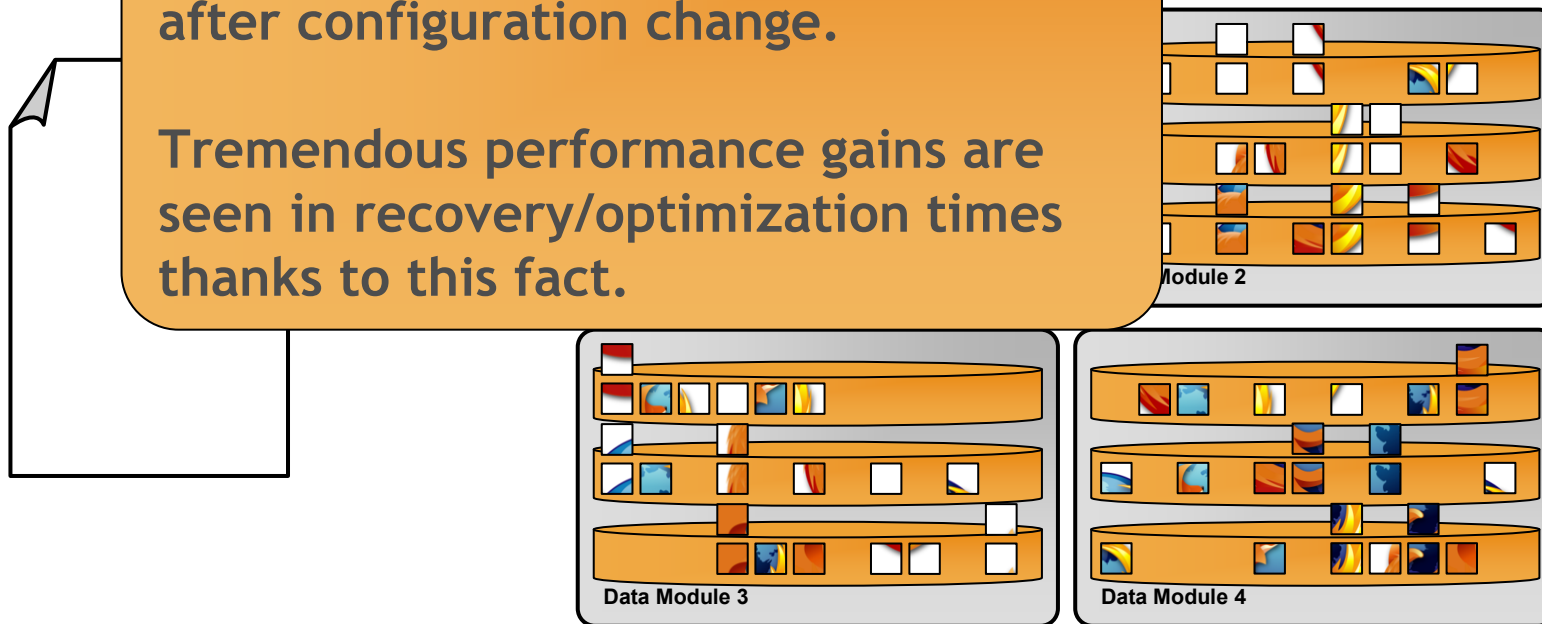
XIV – Delivers: Agile, Responsive, Flexible Applications

- **Provision Storage in < 1 Minute**

- **Resize Volumes Dynamically**

- **High Performance Snapshots**

  – Replicate data for backups or other apps

  – Empower DBAs to manage Snaps

- **Add New Capacity Quickly, Transparently**

  – Added capacity available near real-time

  – Existing data automatically re-balanced across all storage resources

- **All Advanced Functionality Included**
  – **Standard XIV System Software**
  – **No Extra Charge**

| Feature	
Async Replication	✓
Sync Replication	✓
Thin Provisioning	✓
Snapshots	✓
Clones	✓
Read/Writeable Snapshots Clones	✓
Consistency Groups	✓
Migration Tool	✓
Management Tools - GUI/CLI/Top	✓
Real Time Performance Monitoring	✓
Historical Performance Collection	✓
RAID X Protection	✓
Sparing	✓
FC and iSCSI	✓
Host Based MPIO and MPIO GUI	✓
Automatic Performance Tuning	✓
Microsoft VSS integration	✓

Improved Capacity Utilization = TCO Control

- **XIV sold as USABLE capacity**
  - NO lost capacity due to : spares, special system areas, volume set asides for replication, etc.

- **Capacity usage easy to monitor**
  - Complete system, storage pool, or volume

- **XIV all virtual**
  - Configured with a single disk type and no RAID groups to minimize islands of capacity
  - No physical disk binding
  - THIN provisioning standard

- **Designed to Perform Well at >90% Capacity Utilization**

+50%

Superior Effective Capacity

- **Effective capacity**
  - Net space used by application data

- **Thin provisioning**
  - Only stores the data, not the zeros

- **Snapshots**
  - Hi perfomance, pointer based, 1000's capable

- **No Orphaned Space**
  - Capacity can be used.

- **Space Reclamation**
  - Reclamation of capacity that is only storing zeros
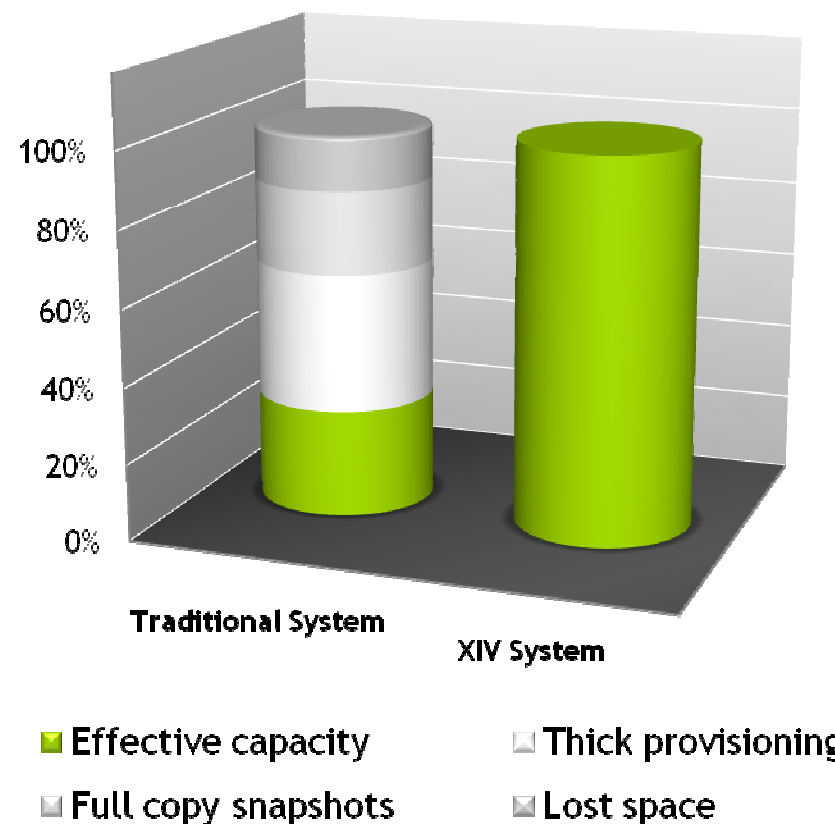
☑ **Fewer TBs address the same applicative needs**

Administration Made Easy

- Use your time to provision storage:
  - Define volumes in seconds
  - Resize volumes painlessly
  - Create instant snapshots
  - Create test environments with writable snapshots

- Don't spend time optimizing:
  - No need for performance tuning
  - No need for complex layout patterns

- Role-based management allows safe delegation of tasks
  - Application administrators manage their own snapshots
  - Integrate with the organization's LDAP

- TPC 4.1 Support through SMI-S

☑ **Simplicity saves time and money**

☑ **Better service to the organization**

Where does XIV fit ?

Requirements	Apps/DBs Supported	Platforms Supported

Requirements

- Minimum: 20TB
- SAN/iSCSI

Apps/DBs Supported

- Exchange
- MS-SQL
- Oracle
- DB2
- SAP
- VDI
- Many more...

Platforms Supported

- Open Systems
  - Linux
  - Microsoft
  - AIX
  - Solaris
  - HP-UX
  - VMware
  - System i with VIO (FC)
  - Mac OS (FC)
  - zLinux (FC)
    - Suse, Redhat
  - z/VM (FC)

- Excellent fit for virtualisation incl: VMware, Power etc.
- Shrinks the storage management for complex environments such as Oracle, SAP and Exchange
- Ideal technology for anyone consolidating storage infrastructures
- Excellent vehicle for dramatically reducing storage running costs
- Compared with the rest …
  - It's functionally richer than Entry/Midrange
  - It's more cost effective than High End (£/TB) but with same functionality
  - All functionality is included within the £/TB price, as opposed to other platforms

XIV + Oracle

Oracle plus XIV – Combined Benefits

EASY
- Ground-breaking ease of use
- Rapid deployment
- Storage management dramatically simplified

RELIABLE
- Highly fault-tolerant, grid
- Self-healing
- Five 9's+
- NO recorded incidents of data loss

FAST
- Massively parallel
- Extreme disk-to-cache bandwidth
- Self-tuning

COST-EFFECTIVE
- All-inclusive software model
- Avg. 63% TCO reduction over 3-year period

GREEN
- Higher effective utilization = lower power consumption
- Low power components

SCALABLE
- Scales linearly in all dimensions
- Adapts next gen HW seamlessly

XIV + Oracle = Optimized DBMS Storage

- Storage configuration for RDBMS's has traditionally been a complex process that requires:
  - Detailed knowledge of the application
  - Careful mapping of expected I/Os to storage subsystem
  - Tradeoffs among applications sharing the storage
  - Ongoing (often manual) load balancing

- XIV mitigates these challenges at the architectural level – <u>SAME* implemented in hardware</u>
  (*Stripe All-over, Mirror Everything)

- XIV's grid architecture and <u>massive parallelism</u> removes the need for most end user planning, deployment and ongoing tuning

- Oracle will aggressively parallelize many operations (scan, sorts, joins, hash, load, create index, etc.) on its own, which makes it a perfect candidate for XIV

- XIV architecture does not require the use of:
  - Multiple logical volumes
  - Host based striping

Why XIV for consolidation?

- **Decrease the Total Cost of Ownership (TCO) versus disparate arrays**

    - XIV efficiencies drive down costs, overheads and improve ROI

    - Power, cooling, floorspace and utility savings

- **Deliver improved service and management efficiencies**

    - Manage more in a single unit, because XIV is less complex – as single point of management for all capacity.

    - Eliminates the traditional performance monitoring and tunining overhead

    - Delivers advanced functionality, improves service levels – PiT, DR and cloning

- **Storage efficiencies**

    - Thin provisioning, snapshots, replication and no orphaned space drive usable utilisation realistically over 90%

- **Forget a complex tiering model and introduce a predictable service and financial storage solution**

    - Based on a usable TB model – "All in": all hardware, software functionality and maintenance

    - No additional capacity, connectivity, or software royalty based pricing

    - A predictable model so customers can plan for future growth without concern

Why XIV plus Oracle and/or Databases?

- **Reduce the planning and deployment of complex environments**
    - No complex data placement, no RAID, no disk size/speed considerations
    - XIV architecture removes the need for multiple volumes and host based striping
    - Speedy deployment and reproduction of environments, locally or remotely

- **Remove performance management and tuning overheads**
    - Eradicate IO bottlenecks that can occur in tradional RAID implementations
    - Consistently high performance delivered with little storage admin resource input
    - Maximise DB parallelism techniques with a solution based on a parallelized grid architecture
    - All modules and disks deliver performance in "parallel"
    - DBA's see a reduction in the amount of effort required to manage & monitor storage resources

- **Storage efficiency improvements**
    - Drive real useable capacity utilization over 90%+
    - Thin Provisioning – drives up utilization & reduces required hardware
    - Pointer-based snaphots enable capacity efficient environment duplication

- **Efficient environment duplication**
    - Snapshots and replication of XIV

Why XIV and Oracle? A real life UK example…

- **Major UK private healthcare organisation**

  - Offices on three continents and more than ten million customers in over 200 countries.

  - After a period of consistent growth, the company found itself in need of immediate additional capacity, as well as a plan to incrementally expand storage as needed in the future.

  - Implemented 2 XIV frames – one for each of its UK-based datacenters. The XIV frames feature 55TB and 27TB of storage capacity,

- **Benefits**

  - Leveraged enterprise-class storage performance for less than the price of EMC Clariion technology.

  - Reduced overhead for operational staff to manage & monitor

  - Able to take advantage of capacity efficiencies and CoD capability

- **Comments:**

  - "XXXXX is extremely satisfied with the new solution from IBM, which has effectively solved its present and future storage needs."

  - "With IBM XIV, we make far more efficient use of the available capacity, which means less frequent upgrades and lower costs for us…"

Grid made simple

Background

CONCLUSION – XIV and Oracle

- **XIV**
  - **Innovative, Enterprise Class Storage Solution**
    - **Efficient, easy, reliable and fast!**
  - **Key Component in New IBM Storage Strategy**
  - **Invented by Moshe Yanai, visionary behind modern enterprise storage**
  - **Field Proven Since 2003 – Aggressive, Cross-Industry Customer and Market Adoption**

- **… with Oracle**

  - **Game-Changing Solution – Measurable, Real World Benefits, Unrivaled Efficiency, TCO and ROI Benefits**

"IBM XIV has reduced our total cost of operations while improving our service levels to our worldwide SAP user community. The XIV architecture has allowed Gerber Scientific to improve performance for operational tasks: one function that used to take eight hours is now completed in just 15 minutes, giving our IT staff time to focus on other projects."

Raf Cohen,
CIO of Gerber Scientific

Q&A

Disclaimer

Copyright © 2011 by International Business Machines Corporation.

This publication is provided "AS IS." IBM product information is subject to change without notice.

No part of this document may be reproduced or transmitted in any form without written permission from IBM Corporation.

Product data has been reviewed for accuracy as of the date of initial publication. Product data is subject to change without notice. This information could include technical inaccuracies or typographical errors. IBM may make improvements and/or changes in the product(s) and/or programs(s) at any time without notice. The information provided in this document is distributed "AS IS" without any warranty, either express or implied. IBM EXPRESSLY DISCLAIMS any warranties of merchantability, fitness for a particular purpose OR INFRINGEMENT. IBM shall have no responsibility to update this information. IBM products are warranted according to the terms and conditions of the agreements (e.g., IBM Customer Agreement, Statement of Limited Warranty, International Program License Agreement, etc.) under which they are provided.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products in connection with this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products.

IBM makes no representations or warranties, expressed or implied, regarding non-IBM products and services, including those designated as ServerProven.

IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. Contact your local IBM office or IBM authorized reseller for the full text of the specific Statement of Direction.