



**HPSS and SANs**  
**aka**  
**“SANitized HPSS”**  
**aka**  
**I/O Redirect Movers (IRMs)**

**Pam Hamilton**

**June 19, 2002**

# Overview



- ✦ **Goal**
- ✦ **How does a SAN help?**
- ✦ **Vision**
- ✦ **Ideas are not new...**
- ✦ **History Lesson**
- ✦ **Current HPSS Architecture**
- ✦ **“SANitized” HPSS Design Goals**
- ✦ **Data Movement for Stage and Migrate**
- ✦ **Data Movement between Workers and HPSS Disk**
- ✦ **Security Considerations**
- ✦ **“SANitized” HPSS Caveats**
- ✦ **Prototype Status**
- ✦ **Acknowledgements**

# Goal



- ✦ **To enhance the speed and efficiency of HPSS by optimizing device I/O.**

**If successful, will remove the HPSS mover as a limiting factor in data movement.**

- ✦ **To SAVE MONEY. IBM mover platforms are NOT cheap!**
- ✦ **To simplify HPSS administration. Mover tuning is an art form ... performance can vary dramatically.**

# How does a SAN help us reach this goal?



## ❖ First, what is a Storage Area Network, exactly?

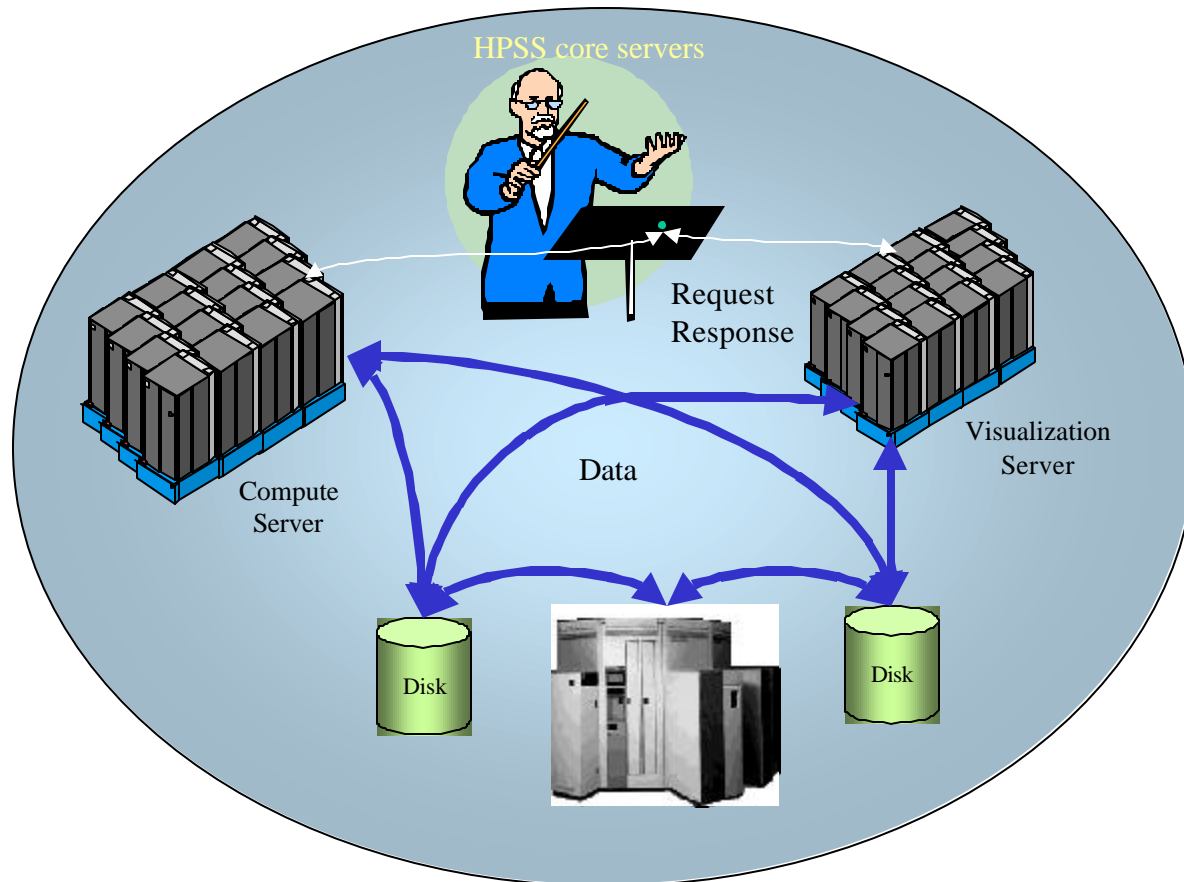
**“A storage system consisting of storage elements, storage devices, computer systems, and/or appliances, plus all control software, communicating over a network.”**

» Storage Networking Industry Association

## ❖ A SAN provides the common hardware interconnect infrastructure necessary to allow the storage subsystem to orchestrate data movement.



# Vision



## Ideas are not new...



- ✦ **Currently used by commercial backup software and called “LAN free backup” and “Serverless backup”.**
- ✦ **Using a combination of these two strategies, data moves from clients to/from disk and between disk and tape without flowing through a LAN or the storage server’s memory.**
- ✦ **Cuts mover bandwidth required to migrate or stage data in half by eliminating the disk mover from the data path.**
- ✦ **Early proof of concept established at the National Storage Laboratory.**

## **A History Lesson...**



- ❖ **National Storage Laboratory (NSL) founded in 1992**
- ❖ **Work at the NSL led to NSL-UniTree**
- ❖ **NSL-UniTree leveraged third-party data transfers almost 10 years ago.**
- ❖ **Third-party data transfer is a data transfer controlled by an agent.**
- ❖ **The agent controls the data transfer by communicating with both the data source and the data sink.**
- ❖ **The agent does not participate in the actual movement of the data.**

## History Lesson continued...



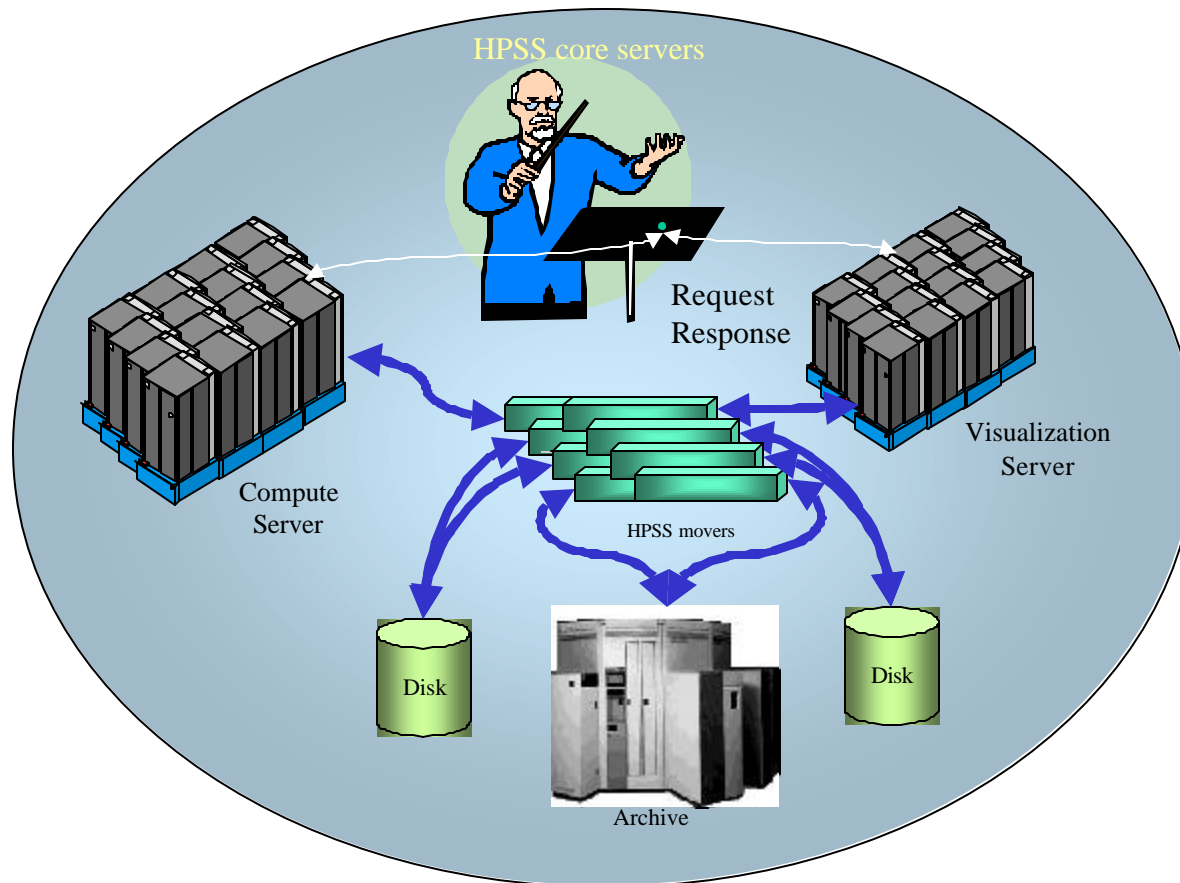
- ✦ **Lessons learned through:**
  - + MAXSTRAT RAID disk devices
  - + IPI3 – Intelligent Peripheral Interface
  - + HIPPI – High speed network
  
- ✦ **NSL work proved several points:**
  - + It established that data transfers between a client and network attached disks could give as good or better performance as native client disk.
  - + Third-party data transfer allowed the transformation of the NSL server to function as a metadata engine that could effectively control data sharing among clients while maintaining high data rates.
  - + Security is aided by separating control and data flow to separate networks.
  - + Hierarchical storage, with movement of data between disk and tape, could be implemented in the shared disk environment.
  
- ✦ **Support for intelligent third-party devices was carried forward in HPSS.**

# Current HPSS Architecture



- ✦ **Most of the storage devices attached to HPSS are, in fact, conventional disks, disk arrays, and tape libraries.**
- ✦ **HPSS utilizes the idea of a data “Mover”.**
- ✦ **SCSI and more recently Fibre-Channel disks and tape drives are attached to computers running a lightweight HPSS Mover protocol.**
- ✦ **A data Mover and the disks and tapes attached to it formed the equivalent of an intelligent third-party device.**
- ✦ **Use of multiple Movers allow many concurrent data transfers to provide very high aggregate data transfer rates.**
- ✦ **HPSS striping of devices and networks provides high bandwidth for individual transfers.**

# Current HPSS Architecture



# **“SANitized” HPSS Design Goals**



- ✦ **Development effort will not impact any current HPSS capability.**
- ✦ **Should be viewed as an enhancement to HPSS, which can be utilized if desired.**
- ✦ **Fundamental to the design is the concept of incremental deployment.**
- ✦ **Design supports the ability to run with a mixture of SAN based storage devices and conventional mover attached storage devices.**
- ✦ **SAN based I/O can be quickly and painlessly disabled on-the-fly for any SAN based device or client.**

## Data Movement for Stage and Migrate

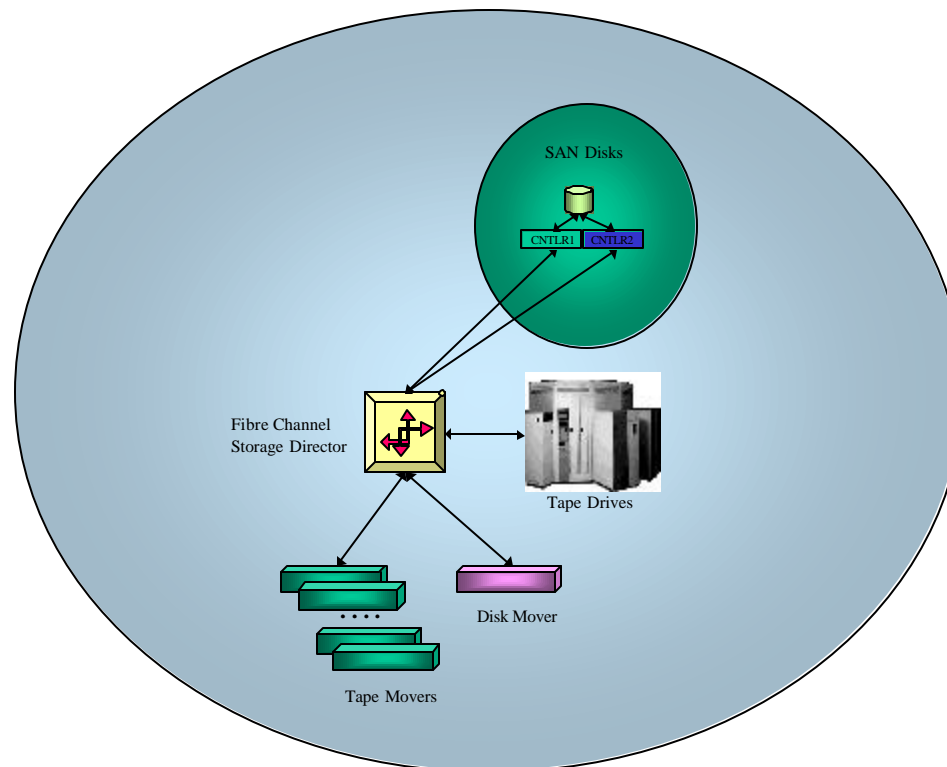


- ❖ **Eliminate the disk mover from the data path used for HPSS tape migrations and stages.**
- ❖ **Roughly equivalent to the “LAN free” backup capability.**
- ❖ **Most of the modifications required to produce “SANitized” HPSS limited to the HPSS mover code and the Client API.**
- ❖ **Modifications should fall in the same areas where IPI-3 transfers are currently supported.**
- ❖ **Plan to leverage the existing HPSS IPI-3 mechanism.**

# Stage and Migrate continued...



- ✦ SAN-enabled Mover will have the I/O descriptors for both the source and sink ends of the transfer.
- ✦ Will have the capability to migrate data from disk to tape or to stage data from tape to disk without moving data across the LAN.



# Data Movement between Workers and HPSS Disk



- ✦ **A worker is a compute platform running an HPSS application built using the HPSS Client API library.**
- ✦ **Currently, the Client API library has code to support the Mover protocol and communicates with HPSS Movers using TCP/IP.**
- ✦ **SAN-enabled HPSS applications can also access SAN-attached HPSS disks directly.**
- ✦ **Application would be passed an I/O descriptor that describes the I/O operation to be performed.**
- ✦ **The benefit of a SAN-enabled Client API library on a worker machine must be weighed against the security exposure.**

# Security Considerations



- ✦ **Currently, HPSS applications running on workers do not access HPSS devices directly. HPSS devices are effectively protected by a mover.**
- ✦ **Even if a mover is compromised, the amount of data at risk is limited to what's visible by that single mover.**
- ✦ **In current SAN-based environments where the goal is to share data across multiple workers, compromising a single SAN attached worker compromises the entire environment!**
- ✦ **A more robust security mechanism is sorely needed!**
- ✦ **Several promising efforts are currently under development.**

## **“SANitized” HPSS caveats...**



- ✦ **Initial proof of concept implementation relies upon the AIX logical volume manager, which restricts “SAN clients” to AIX based systems.**
  
- ✦ **Restrictions are due to:**
  - + **RAID providers limitations on the number of SCSI devices they can make visible on a channel**
  - + **Implies relatively small number of large disks will be visible to the SAN client**
  - + **HPSS currently has a limit of 16k segments on an HPSS volume**
  
- ✦ **Making use of the AIX Logical Volume manager is a way to bridge this gap.**
  
- ✦ **HPSS 5.1 removes the 16k segment restriction and will allow us to use the RAID vendor’s devices directly and remove our dependency upon AIX.**

## **“SANitized” HPSS caveats...**



- ✦ **In this strategy, clients (other than HPSS tape movers) do not access tape devices directly – all data is buffered on disk.**
- ✦ **HPSS tape based Classes of Service operate the way they always have – by transferring data through an HPSS mover.**
- ✦ **The “SCSI extended copy” feature was investigated to move data between disk devices and tape devices.**
- ✦ **It was found that this specialized SCSI command was only supported on a handful of fibre to SCSI bridge products and was not available native on either fibre RAID disk devices or any tape products.**
- ✦ **Conclusion was that performance through a specialized bridge would not be any better than using a “SANitized” HPSS mover and may actually be worse.**

# Prototype Status



- ✦ **Concept presented to HPSS Tech committee.**
- ✦ **An HPSS Change Request was submitted by IBM (I/O Redirect Movers).**
- ✦ **Coding is complete.**
- ✦ **Testing has begun in the I/O Testbed at LLNL.**
- ✦ **Formal process of integrating the modifications into HPSS is about to begin.**



# Acknowledgements



## ✦ Prototype Design and Development:

- + Keith Fitzgerald and Mike Gleicher

## ✦ References

- + K. Fitzgerald, M. Gleicher, “LLNL SANitized HPSS”, White paper.
- + K. Fitzgerald, O. Graf, H. Hulen, R. Watson, “Storage Area Networks and the High Performance Storage System,” *Proceedings Tenth Goddard Conference on Mass Storage Systems*, College Park, MD (Apr 2002).