Ensuring Reliability in Tape Systems

Presented by:
David Cerf
EVP – Crossroads
+1 972 896 3164
dcerf@crossroads.com

History Of Peripheral Storage

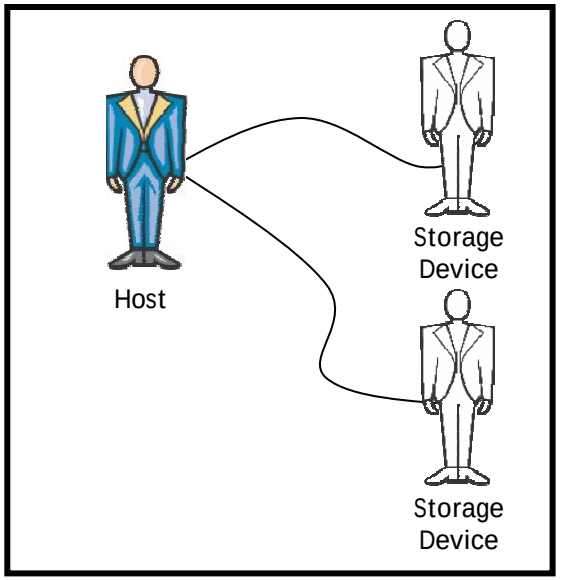
Peripheral Storage History

Prior to SCSI, most storage was "in the box"

- Controlled by one system (host)
- Completely isolated from any other devices

Why does that matter to storage?

- Only one Master – no sharing
- No clue about where it is, who else is in there with them, or what system they might be a part of
- Performance was solely dictated by the processor system – when it needs it, it gets it.



Computer System

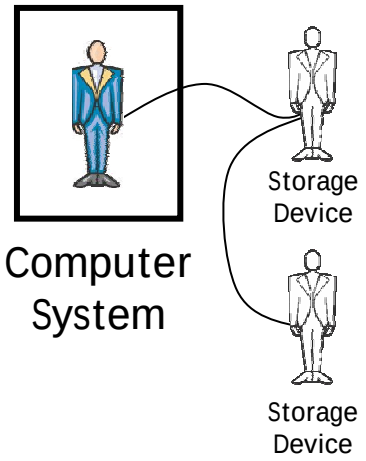
Peripheral Storage History

Post SCSI (Standard in 1986)

- Still controlled by one system (host)
- Shared bus with other devices

So Why Is This Significant?

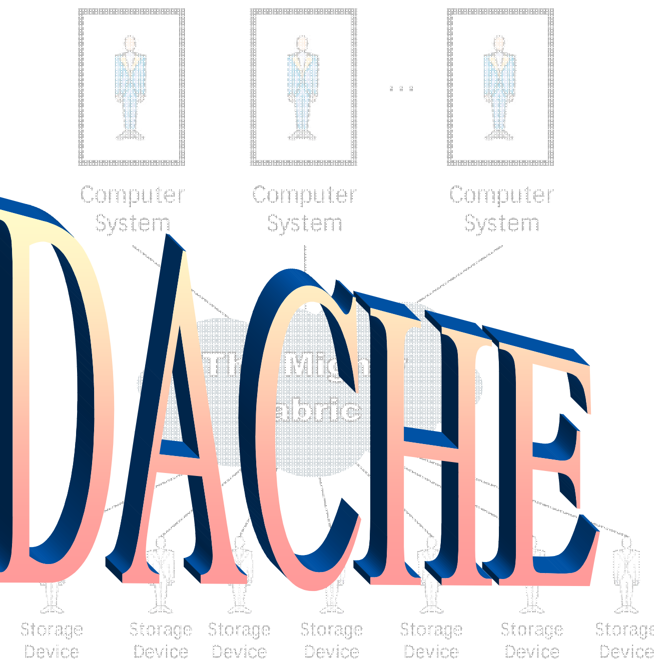
- Still only one Master – but am now sharing in the communication path with other people like me
- Still no clue about where it is, who else is with them, and if it belongs to a bigger system
- Performance is now affected by how many devices are on the bus and the overall performance of the bus – i.e. if one device is hogging the bus, then the others can't get their work done



Peripheral Storage History

Enter Fiber Channel (Standard in 1994)

- Multiple host systems
- Devices "shared" between hosts



So Now What?

WHAT A HEADACHE

- [illegible] multiple masters
- [illegible]
- Things like performance, utilization, failure rates, error handling, failover, load balancing all comes into the conversation

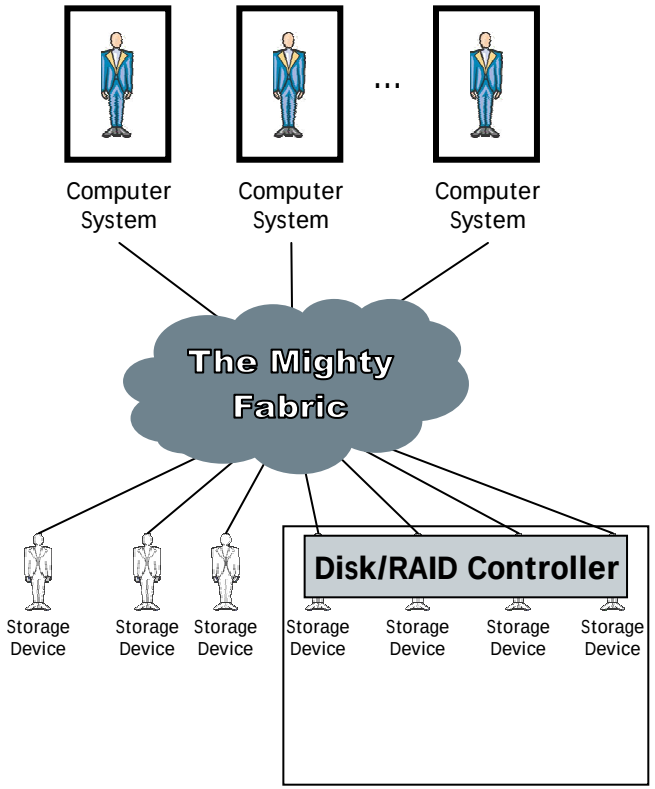
Peripheral Storage History

Disk & Tape Diverge in How to deal with this new connectivity

- Makes some sense as Fiber Channel was built by disk people for disk people
- Still no excuse for tape manufacturers however…

What did Disk do?

- Solve both the multiple host and multiple "devices" challenge by creating a super device – Disk Controller or RAID Controller
- Carve up their smaller (internal) devices by host, 3 for you, 2 for you, 4 for you, and "None for YOU" – Soup Nazi, Seinfeld
- Suddenly we are back to our comfort level with one Master and no sharing!



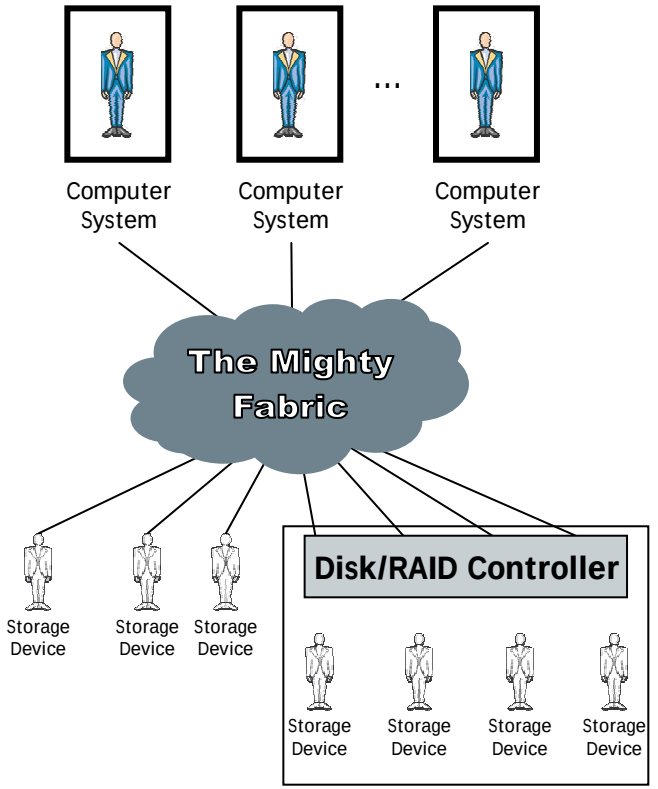
Peripheral Storage History

Disk & Tape Diverge: How to deal with this new connectivity

- Makes some sense as Fiber Channel was built by disk people for disk people
- Still no excuse for tape manufacturers however…

What did Tape do?

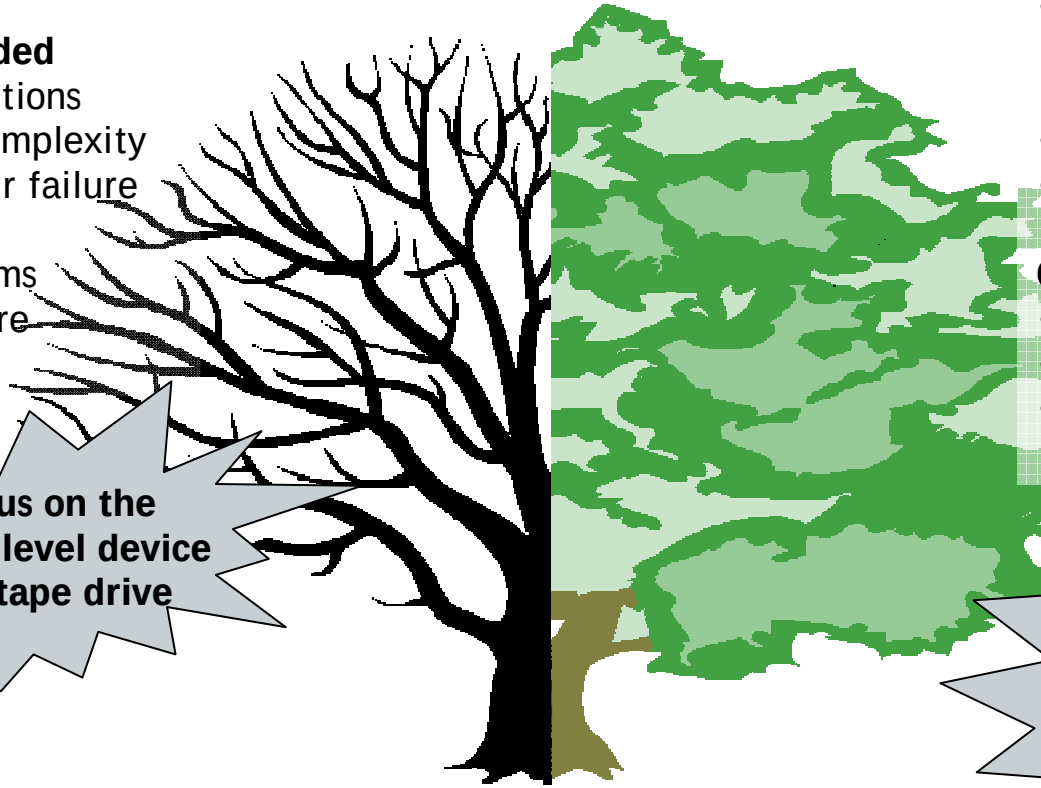
- "Hey we'll put a Fiber interface on our drives and attach them directly" – some genius tape guy
- Now the drives must grow in intelligence to deal with sharing, error handling, and switch fabric connectivity
- Suddenly we are in a whole new world with new requirements, new rules, new language, etc., etc., etc.



Tape & Disk – Diverging Paths

Manufacturers Focused On
- Double speed
- Double capacity

Customer's Needed
- Intelligent solutions
- Simplify the complexity
- Aspirin for their failure headaches
- Managed systems
- Proactive failure management



Focus on the lowest level device The tape drive

Tape

Manufacturers Focused On
- Double speed
- Double capacity
- Feature value add (replication, CDP, CAS, VTL, Deduplication, etc.
- Management
- Storage Virtualization

Customer's Needed
- Even more manageability
- Even more simplicity
- This stuff meant more $$$

Commoditized lowest level device, Drove value in system

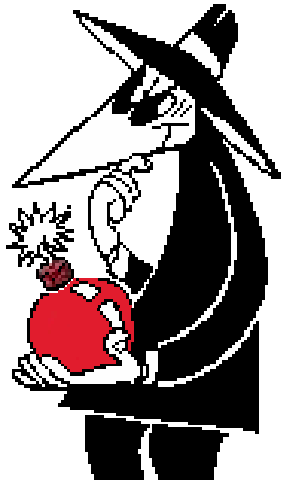
Disk

Disk vs. Tape

- **Marketing Driven**
  - EMC, DataDomain, others spend major $$$ telling everyone that tape isn't needed – If it in print, it must be true, right?
  - Paid Analysts have and will create "reports" showing Disk is better, cheaper, greener…
- **Customer Satisfaction driven**
  - Tape problems solicit finger pointing – not solution
  - Tape problems are typically dealt with by – replacing tape, and hoping the problem goes away
  - Managing tape library, drives, and media is a hassle, and with shrinking budgets – who has the time?

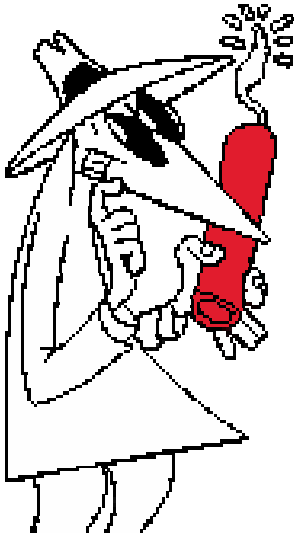
What can we learn from Disk

- **Embrace Failure – "Failure what failure?"**
  - Disks fail at a very high rate – Why don't customer's complain?
    - Built in recovery with RAID, and marketed that it is a good thing, a benefit!
    - Systems use their "call home feature" to call the manufacturer and support hustles out to replace bad disks – positioned as a "value-added service"
- **Sell a solution**
  - You don't buy disks, you don't even buy RAID – you buy a solution for "High Availability", or "Archive", etc.
  - You don't manage disk, you manage "volume"
- **Management**
  - Configuration, use, capacity, automatic alerting, and of course failure . . . I mean opportunities for service
  - Bottom line – provide the customer with less involvement, less hassle, makes them think it is a better solution



What can we do about Tape?

- **Embrace Tapes' Benefits**
  - Green technology
  - Data portability
  - Cost (if done right)
- **Improve the customer experience**
  - Don't just sell more drives, maximize their investment
  - Backup window problem could be caused by application, environment, tape drive defects, system performance, etc.
  - Solve defects – find root cause and resolve, don't just keep throwing tapes away
  - Do all of this for the customer, don't make them get a degree in backup physiology to protect their data
- **Management**
  - Provide mechanism to better manage complex tape infrastructure
  - Give customer ability to be alerted to issues in tape environment
  - Give customer ability to proactively solve potential problems before they become a major issue



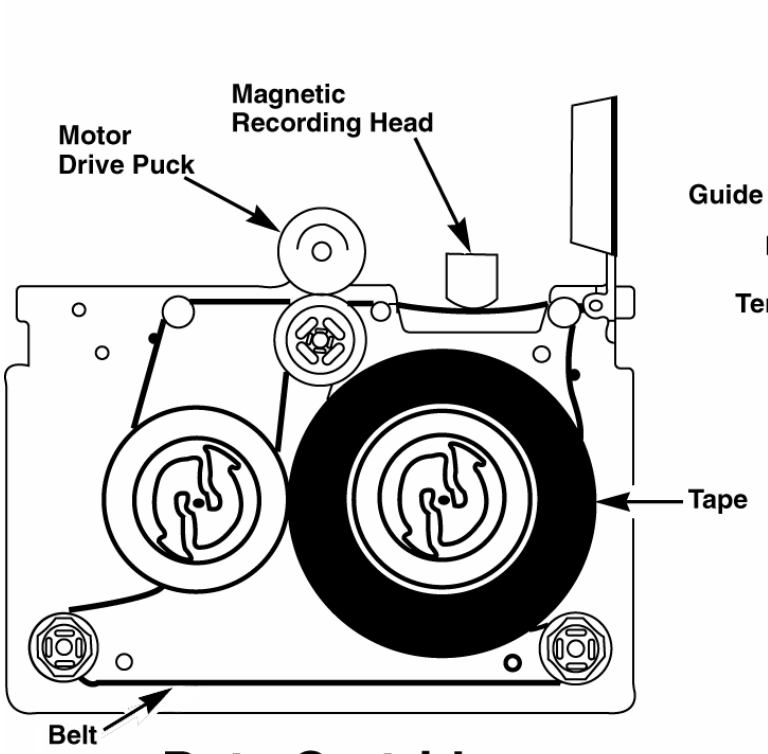

Tape 101

What makes Tape Tick?

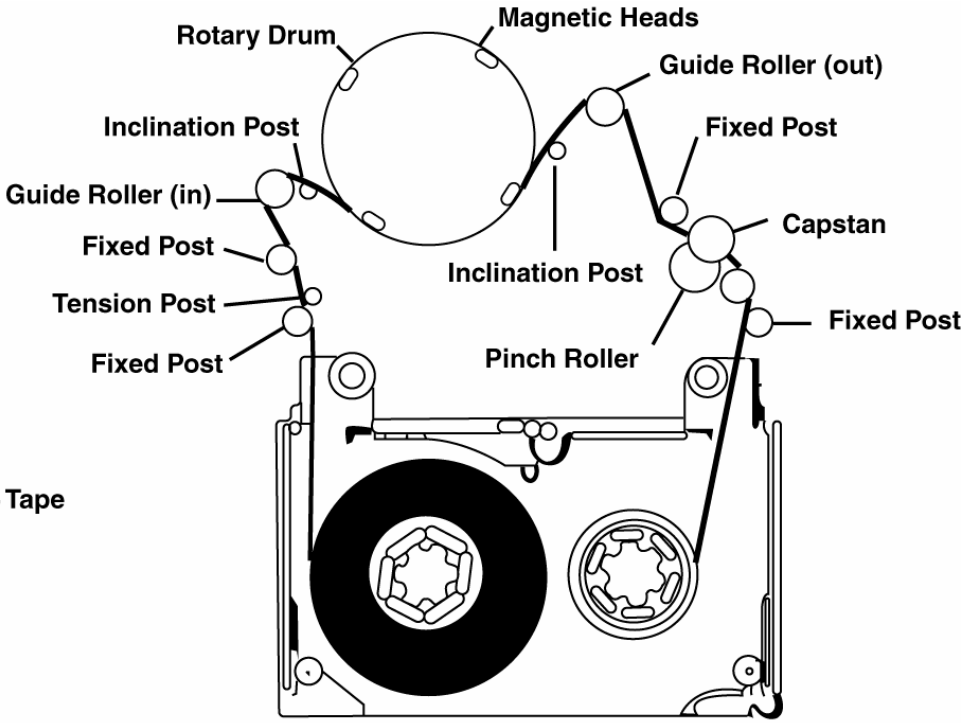
- **Lots and lots of data**
  - Tape wants to take large blocks of data in a continuous stream
- **Send large blocks quickly**
  - Tape drives consume data at very high data rates
- **Load media, fill media, unload media**
  - Tape load/unload process is the longest in data transfer process

What Makes Tape Inefficient or Ineffective?

- **Starting & Stopping**
  - **Media wear – expansion (stretching)**
  - **Media damage**



Data Cartridge



Helical Scan Cassette

What Makes Tape Inefficient or Ineffective?

- **Starting & Stopping**
  - Media wear – expansion (stretching)
  - Media damage
- **Tape drives solve this**
  - Always rewind before
  - Leader at fro
- **Problem s**
  - When drive is s
- Another side effect – Small data transfers with load/unloads will always require the tape to rewind before unload, and then fast forward on load to add new data – MAJOR TIME COST

IMPORTANT: Stay above streaming rate, write large chunks of data – maybe the whole tape if possible

What Makes Tape Inefficient or Ineffective?

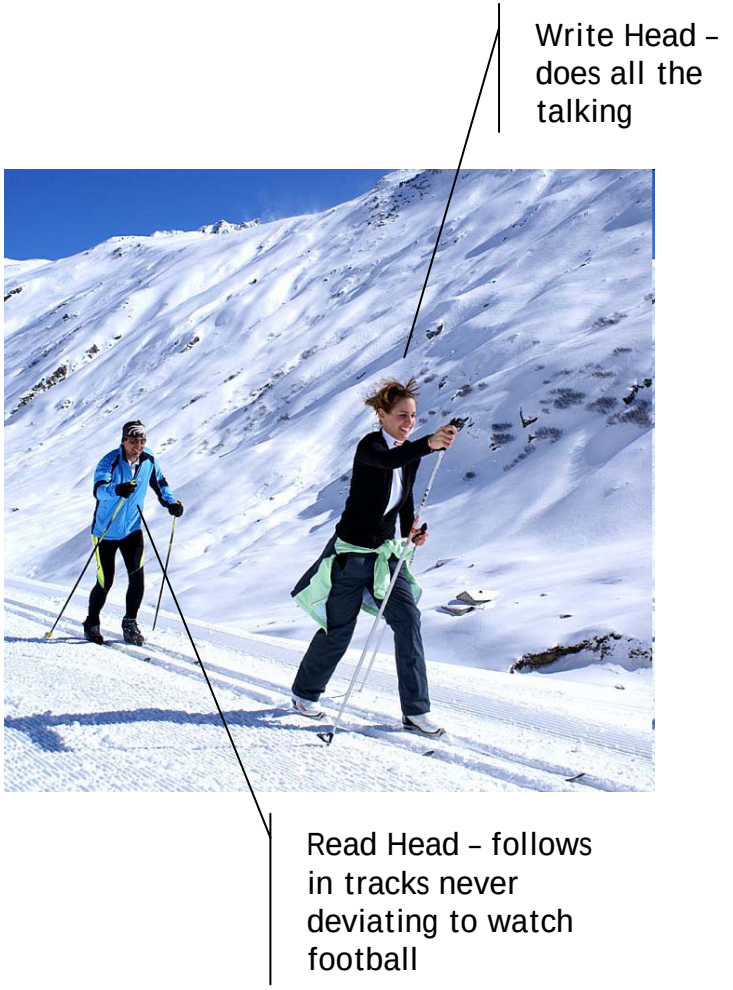
- **Loading & Unloading**
  - Mechanical system that loads cartridge and positions
    - Tape drives are rated for load/unload duty cycles that vary by drive type
    - Tape drives have a total load cycle as well
    - Library has load unload cycle rating as well
  - Time – Its not on our side
    - Library is occupied for both load & unload
    - Tape drives have different ratings for time to load and position tape
- **Occupied, but unused drives**
  - If a tape is loaded – the drive is unusable by any other system
  - If a tape is loaded but no data is going to or from it – then you have a very expensive door stop.

How does Tape deal with Defects?

- **Tape media – bunch of tiny little magnets all stuffed together**
  - Audio or even video recording having a bit or multiple bits in error is completely invisible to human brain
  - In data recording, a single bit in error means the whole data set is corrupt
- **Built in error recovery**
  - Strong Error Correction Codes (ECC) built into the data blocks and data tracks
    - Drive is able to recover from single bits in error within a block
    - Drive is able to recover from a multi-block error such as a scratch
  - Drive writes data to media and follows up by reading the data
    - Verifies data read matches data written – good idea
    - If ECC is invoked then data can be rewritten further down the tape
      - Different rules by drive type, by drive manufacturer
      - Causes slow down of transfer
- **Fire and forget**
  - Tape manufacturers assume if data was read after written, then that is good enough – lets call it a day and go on home
  - Sounds bullet proof – How could there every be a problem???

How does Tape deal with Defects?

- **Mechanical system – being more tolerant isn't a good thing in this case**
  - Drives vary and could have different tape paths such that a second drive would have difficulty reading a tape written on the first drive
  - Tape loading and seating could be different causing offsets
- **Read after write is the best it gets**
  - Read head immediately follows the write head – no chance to get off course
  - Subsequent read operations must use the servo tracking mechanism to stay on track (think of skiing by yourself after a light dusting of snow, the track is there, but you won't be exactly in the tracks all the time)



How does Tape deal with Defects?

Not all media is created equal

- Manufacturers use different processes
- Manufacturers have different acceptance criteria
- Manufacturers have different quality

What about Global Warming???

- The environment matters – not as much as disk, but it still matters
- Give me clean air – dust, and dirt can get onto media or heads and scratch the tape

Gosh its hot – Abominable Snowman to Bugs Bunny

Conclusion – Ideal Tape World

- Tapes loaded and fully written with large data blocks (minimize load/unload)
- Data transferred at compression rates (minimum requirement at streaming rate)
- Tapes unloaded immediately after written
- High quality media
- High quality tape drive
- Monitor the whole system to make sure all of the above is happening
- Verify the written media every so often to validate it is still viable

What does this mean for ME!

High performance backup environment

You thought you were Buying *this*…
But you got *this* instead

Current Challenges in HPC Tape Environments

Exploding Data Growth

Long-Term Data Protection

The tape environment is complex and there is limited visibility to understand the TRUE performance of your environment

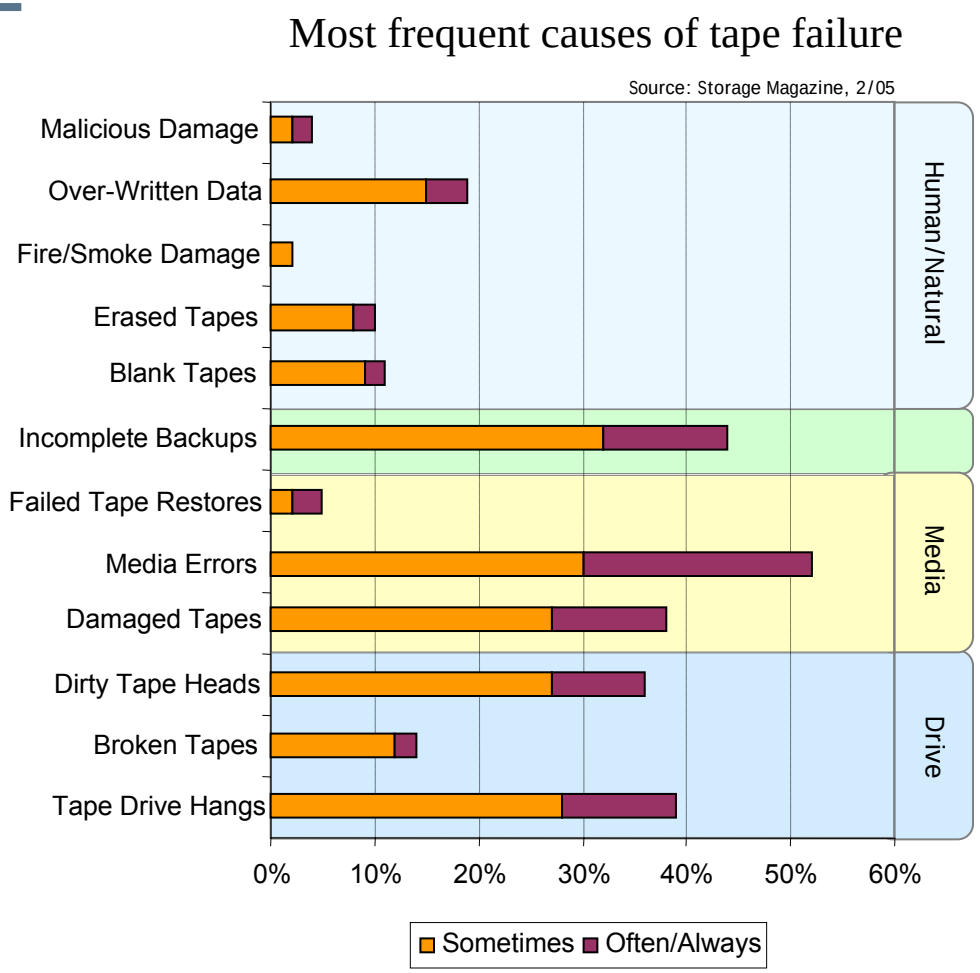
Tape Write Performance

Tape Write/Read Reliability

Reliability of Tape Backup

Causes of Backup/Recovery Failures

- **Human/Natural**
  - Lack of automation
  - Lack of repeatable process
- **Drives**
  - Mechanical errors
  - Dirty heads (write or read)
  - Firmware errors
- **System**
  - Drive utilization imbalance
  - Underperforming drives
  - Backup software errors/configuration
  - Network capacity
  - Network errors
- **Media**
  - Damaged (scratch, shoeshine, creased, etc.)
  - Degraded (worn, magnetic properties, etc.)



Tape Dyno

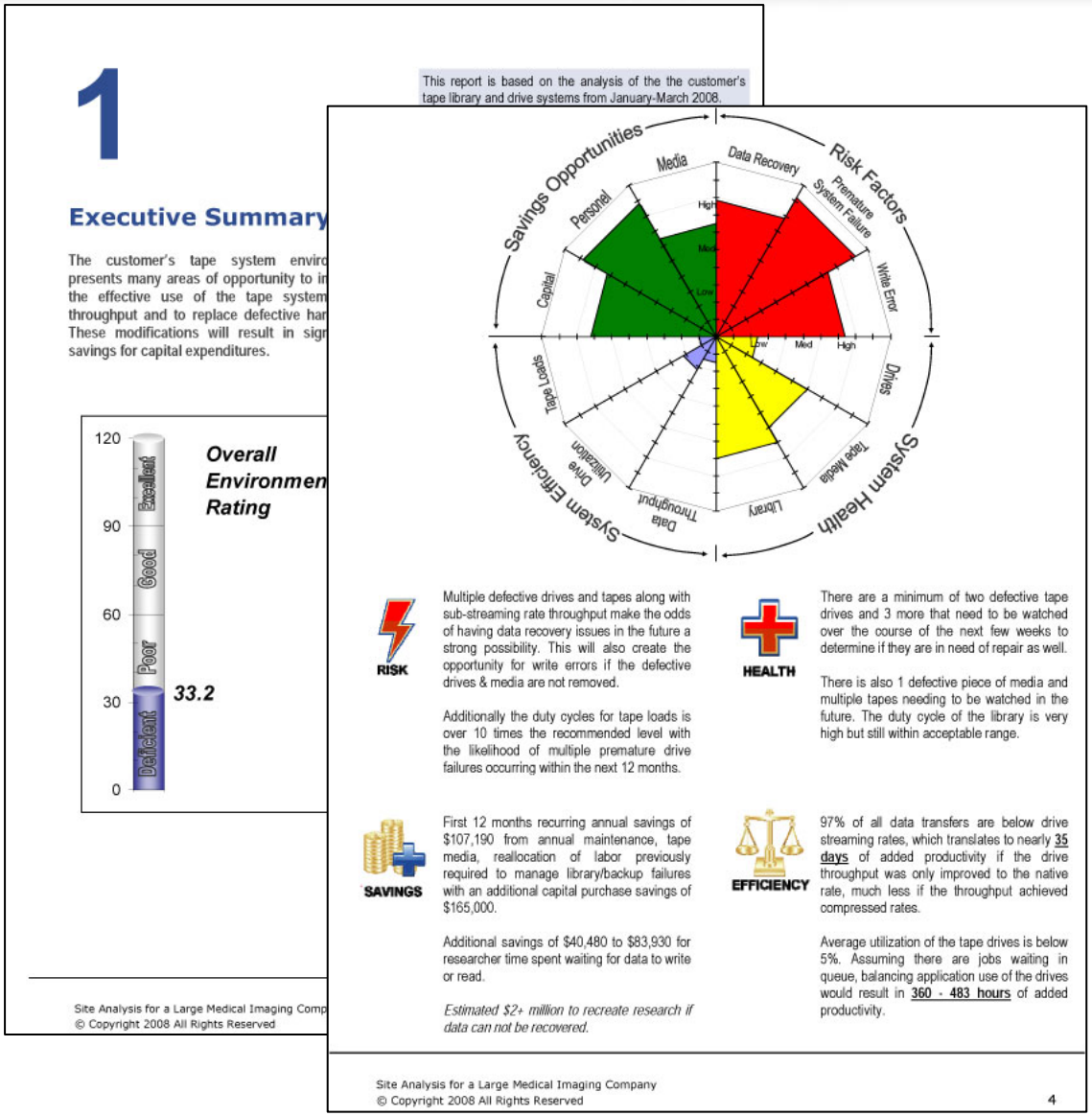

Horsepower
67.0
Vacuum
-0.7
129.71
ft-lb
2711.4
RPM
AFR
8.0
A/F
Oil Pressure
18.6
Pressure
Crankcase
-1.6
Psi
None
Play
Freeze
Rewind
Quick Print
Open Run
Replay Mode
Data Rate 100.000
Data Location
154
Degrees F.

CROSSROADS®
Assuring Global Information

Site Analysis Report

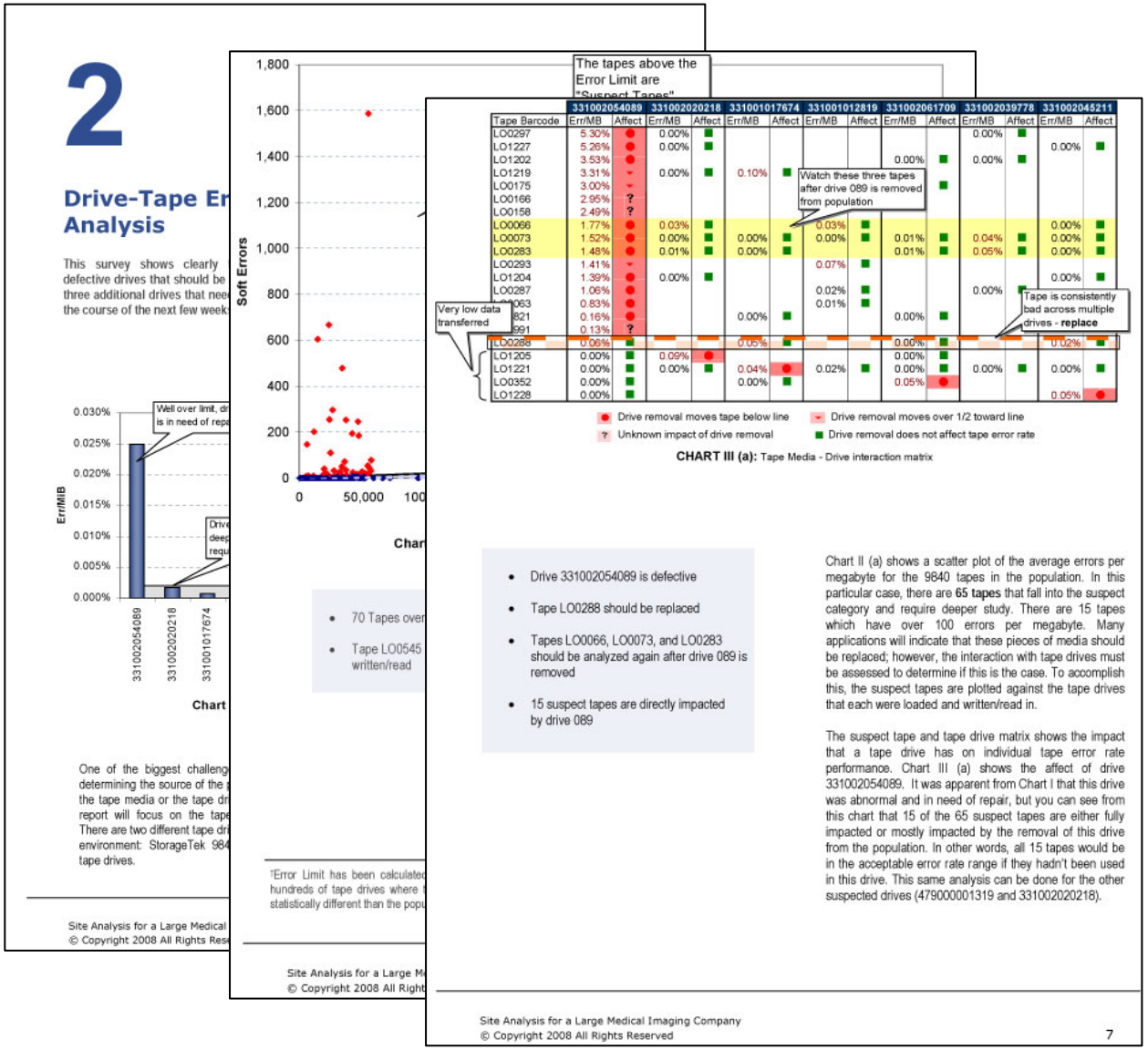
Executive Summary

- High level meant for customer to take to the CIO/CFO
- Overall rating useful for follow on reports to see progress
- System Wheel gives “at a glance” view of 12 areas within environment
- High level summary for
  - Risk to business
  - Health of environment
  - How efficient the systems are being used
  - Opportunity for savings



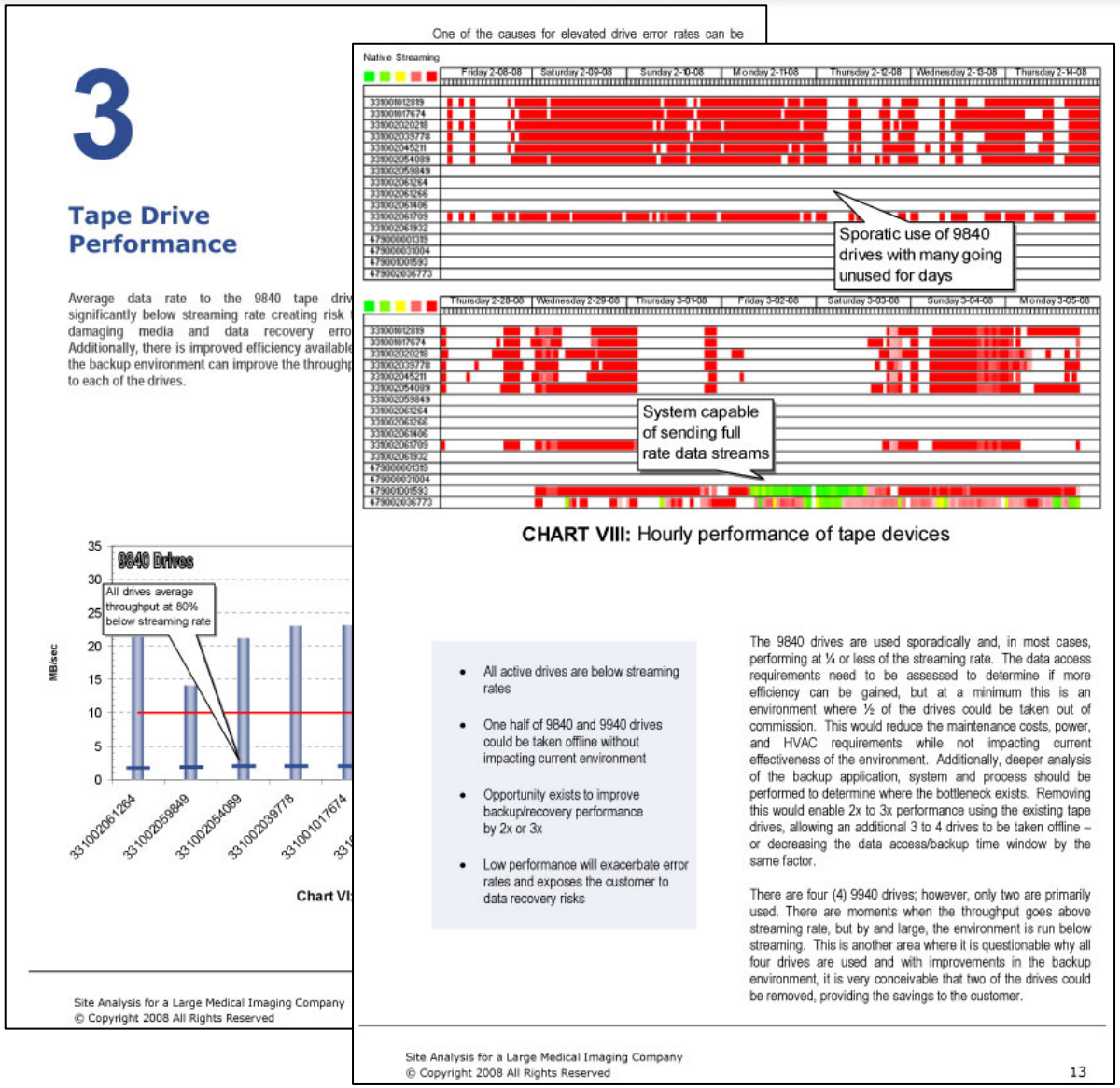
Drive-Tape Error Analysis

- Purpose is to isolate drives and media
- Plot “suspected” tapes – these would have been thrown out in normal operations
- A vertical red section indicates bad drives
- Horizontal red numbers indicates bad media



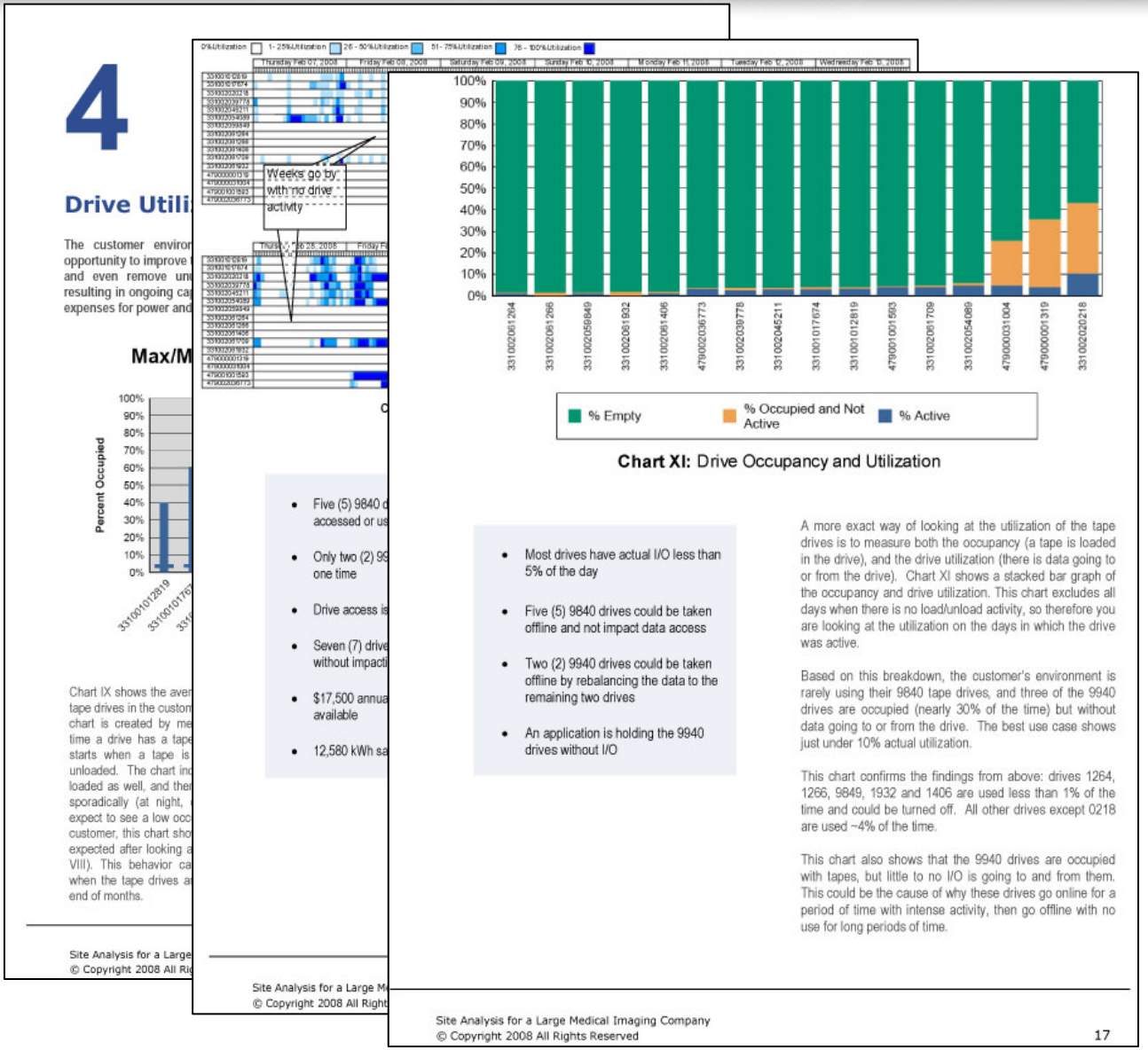
Tape Drive Performance

- Streaming rate is the bar that all is measured against
- The DNA chart give you the ability to determine if the customer has an isolated problem or a systemic one
- If yellow or green blocks, are shown then you know the environment is capable of supporting high data rates – the problem must be in the system or application



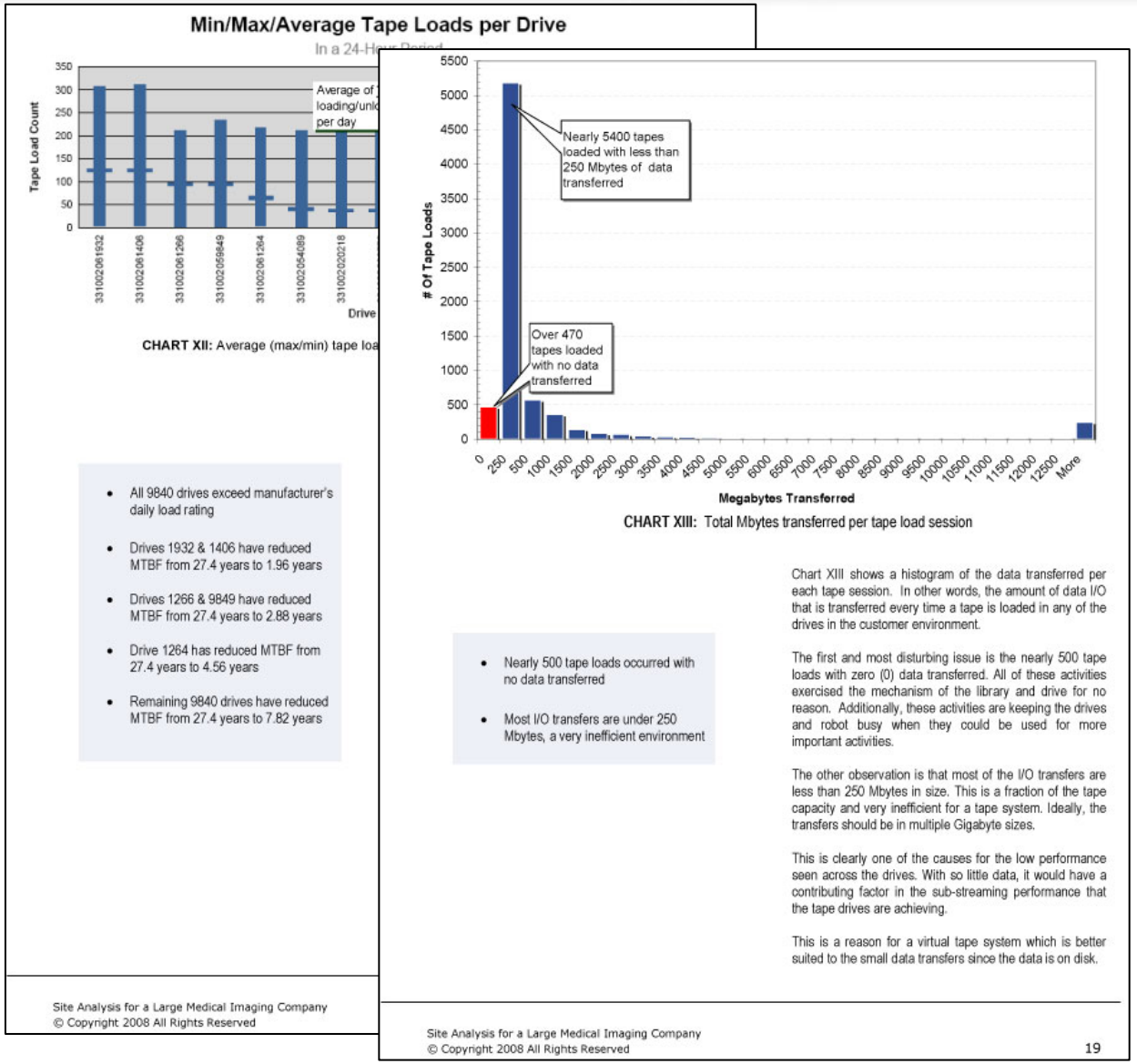
Drive Utilization

- This is all about maximizing the investment in the tape system
- DNA chart gives view into how the systems are being used on a daily/weekly/monthly basis
- Occupancy gives you a great view into the wasted time spent loaded, but nobody is home



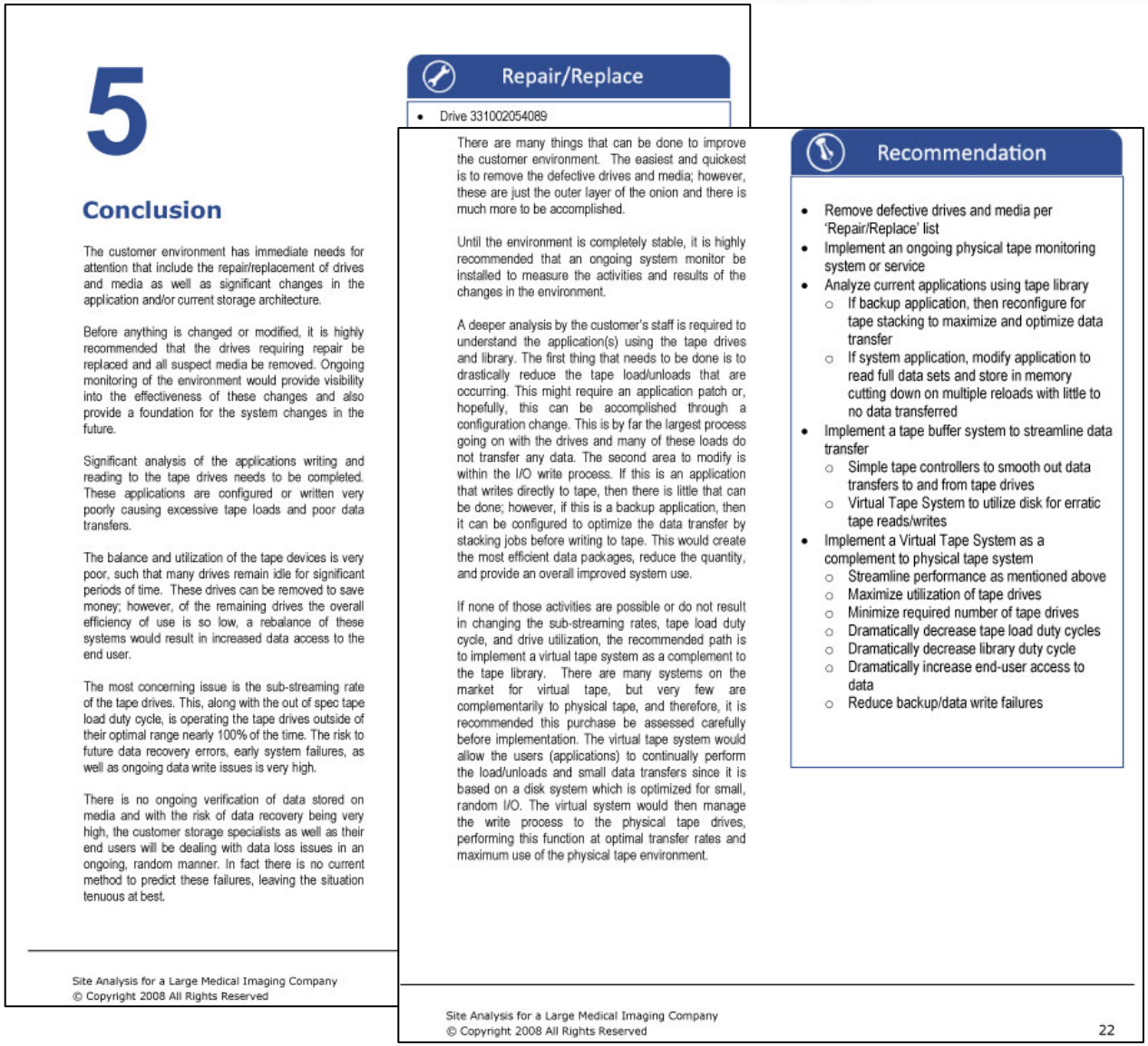
Drive Utilization

- Tape loads is all about time, and mechanical stress – low number is good for everyone
- Data transferred gives you insight into how efficient the system is being used – again we want large blocks of data



Conclusion

- Actionable results with recommended changes and follow-on steps
- Areas to improve, specific areas of risk – a tangible result that can drive follow on business or service



5

Conclusion

The customer environment has immediate needs for attention that include the repair/replacement of drives and media as well as significant changes in the application and/or current storage architecture.

Before anything is changed or modified, it is highly recommended that the drives requiring repair be replaced and all suspect media be removed. Ongoing monitoring of the environment would provide visibility into the effectiveness of these changes and also provide a foundation for the system changes in the future.

Significant analysis of the applications writing and reading to the tape drives needs to be completed. These applications are configured or written very poorly causing excessive tape loads and poor data transfers.

The balance and utilization of the tape devices is very poor, such that many drives remain idle for significant periods of time. These drives can be removed to save money; however, of the remaining drives the overall efficiency of use is so low, a rebalance of these systems would result in increased data access to the end user.

The most concerning issue is the sub-streaming rate of the tape drives. This, along with the out of spec tape load duty cycle, is operating the tape drives outside of their optimal range nearly 100% of the time. The risk to future data recovery errors, early system failures, as well as ongoing data write issues is very high.

There is no ongoing verification of data stored on media and with the risk of data recovery being very high, the customer storage specialists as well as their end users will be dealing with data loss issues in an ongoing, random manner. In fact there is no current method to predict these failures, leaving the situation tenuous at best.

Site Analysis for a Large Medical Imaging Company

Repair/Replace

- Drive 331002054089

There are many things that can be done to improve the customer environment. The easiest and quickest is to remove the defective drives and media; however, these are just the outer layer of the onion and there is much more to be accomplished.

Until the environment is completely stable, it is highly recommended that an ongoing system monitor be installed to measure the activities and results of the changes in the environment.

A deeper analysis by the customer's staff is required to understand the application(s) using the tape drives and library. The first thing that needs to be done is to drastically reduce the tape load/unloads that are occurring. This might require an application patch or, hopefully, this can be accomplished through a configuration change. This is by far the largest process going on with the drives and many of these loads do not transfer any data. The second area to modify is within the I/O write process. If this is an application that writes directly to tape, then there is little that can be done; however, if this is a backup application, then it can be configured to optimize the data transfer by stacking jobs before writing to tape. This would create the most efficient data packages, reduce the quantity, and provide an overall improved system use.

If none of those activities are possible or do not result in changing the sub-streaming rates, tape load duty cycle, and drive utilization, the recommended path is to implement a virtual tape system as a complement to the tape library. There are many systems on the market for virtual tape, but very few are complementarily to physical tape, and therefore, it is recommended this purchase be assessed carefully before implementation. The virtual tape system would allow the users (applications) to continually perform the load/unloads and small data transfers since it is based on a disk system which is optimized for small, random I/O. The virtual system would then manage the write process to the physical tape drives, performing this function at optimal transfer rates and maximum use of the physical tape environment.

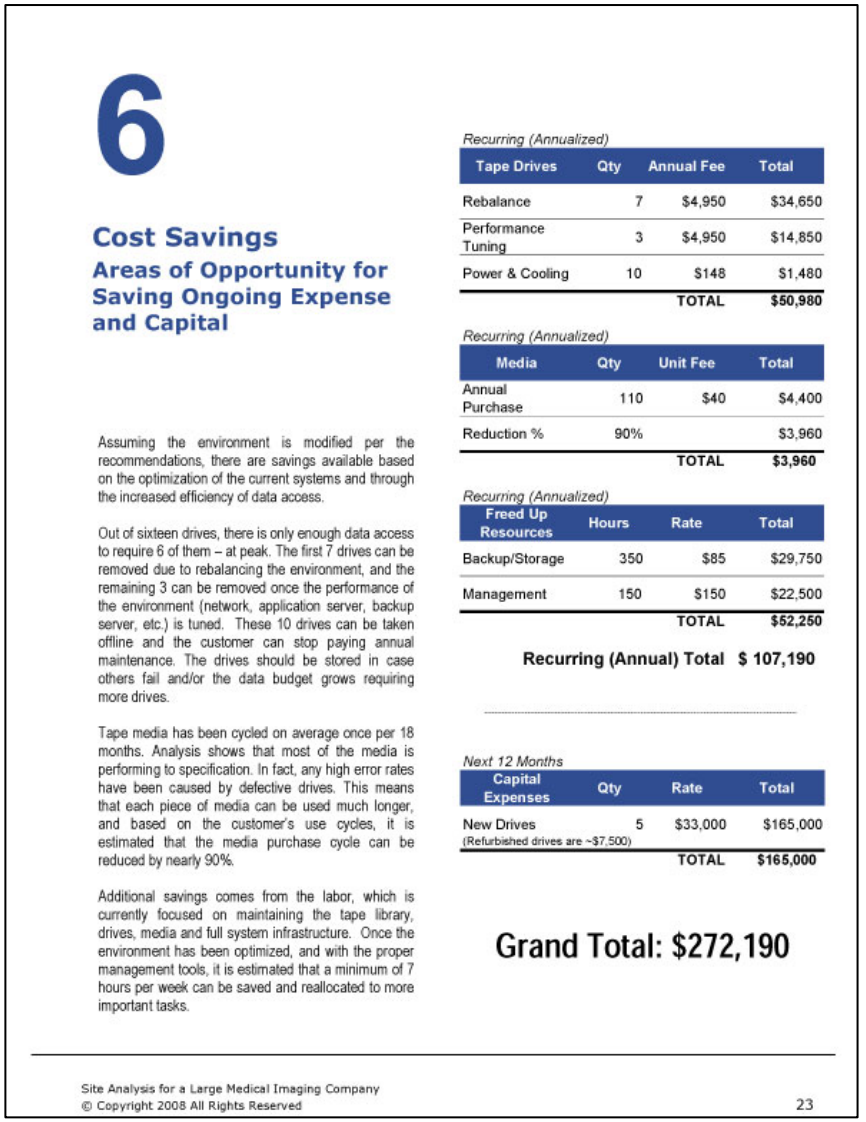
Recommendation

- Remove defective drives and media per 'Repair/Replace' list
- Implement an ongoing physical tape monitoring system or service
- Analyze current applications using tape library
  - If backup application, then reconfigure for tape stacking to maximize and optimize data transfer
  - If system application, modify application to read full data sets and store in memory cutting down on multiple reloads with little to no data transferred
- Implement a tape buffer system to streamline data transfer
  - Simple tape controllers to smooth out data transfers to and from tape drives
  - Virtual Tape System to utilize disk for erratic tape reads/writes
- Implement a Virtual Tape System as a complement to physical tape system
  - Streamline performance as mentioned above
  - Maximize utilization of tape drives
  - Minimize required number of tape drives
  - Dramatically decrease tape load duty cycles
  - Dramatically decrease library duty cycle
  - Dramatically increase end-user access to data
  - Reduce backup/data write failures

Site Analysis for a Large Medical Imaging Company

Cost Savings

- Opportunity for saving money for the customer if the recommendations are followed
- Both immediate as well on on-going yearly results
- Catastrophic failures are not monetized which only grows the value of the report



6

Cost Savings

Areas of Opportunity for Saving Ongoing Expense and Capital

Assuming the environment is modified per the recommendations, there are savings available based on the optimization of the current systems and through the increased efficiency of data access.

Out of sixteen drives, there is only enough data access to require 6 of them – at peak. The first 7 drives can be removed due to rebalancing the environment, and the remaining 3 can be removed once the performance of the environment (network, application server, backup server, etc.) is tuned. These 10 drives can be taken offline and the customer can stop paying annual maintenance. The drives should be stored in case others fail and/or the data budget grows requiring more drives.

Tape media has been cycled on average once per 18 months. Analysis shows that most of the media is performing to specification. In fact, any high error rates have been caused by defective drives. This means that each piece of media can be used much longer, and based on the customer's use cycles, it is estimated that the media purchase cycle can be reduced by nearly 90%.

Additional savings comes from the labor, which is currently focused on maintaining the tape library, drives, media and full system infrastructure. Once the environment has been optimized, and with the proper management tools, it is estimated that a minimum of 7 hours per week can be saved and reallocated to more important tasks.

Recurring (Annualized)

Tape Drives	Qty	Annual Fee	Total
Rebalance	7	$4,950	$34,650
Performance Tuning	3	$4,950	$14,850
Power & Cooling	10	$148	$1,480
		TOTAL	$50,980

Recurring (Annualized)

Media	Qty	Unit Fee	Total
Annual Purchase	110	$40	$4,400
Reduction %	90%		$3,960
		TOTAL	$3,960

Recurring (Annualized)

Freed Up Resources	Hours	Rate	Total
Backup/Storage	350	$85	$29,750
Management	150	$150	$22,500
		TOTAL	$52,250

Recurring (Annual) Total $ 107,190

Next 12 Months

Capital Expenses	Qty	Rate	Total
New Drives (Refurbished drives are ~$7,500)	5	$33,000	$165,000
		TOTAL	$165,000

Grand Total: $272,190

Site Analysis for a Large Medical Imaging Company