How green is my HSM?

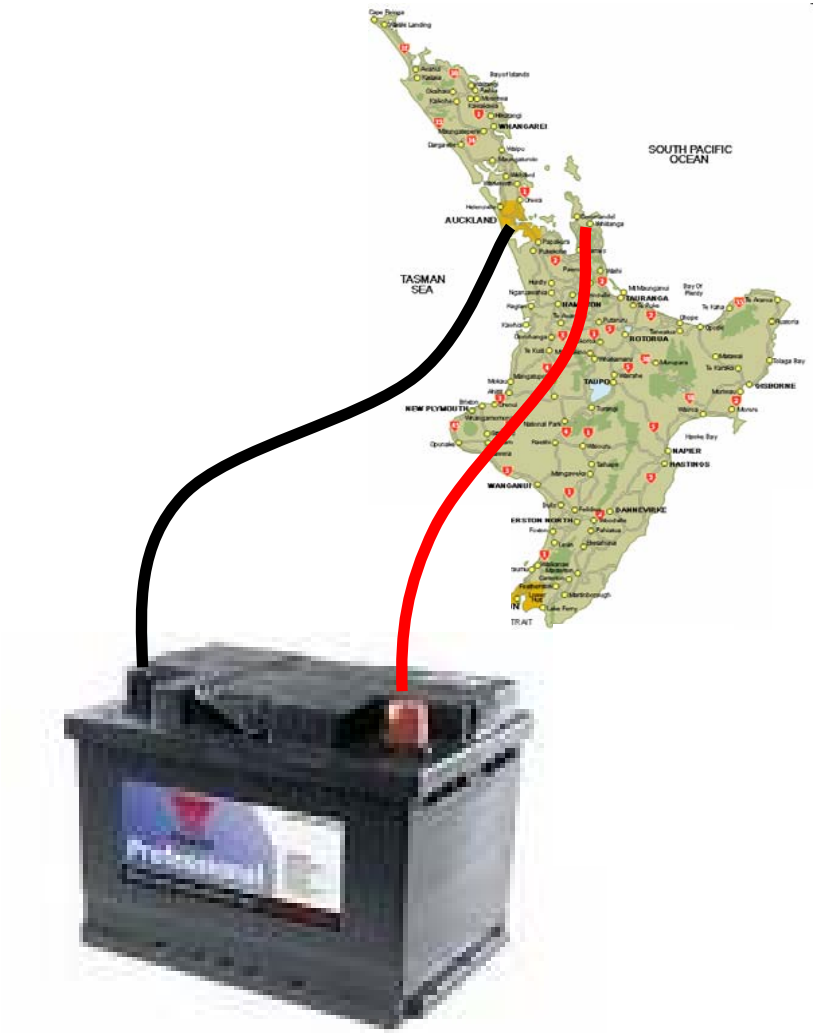
David Honey

SGI ANZ

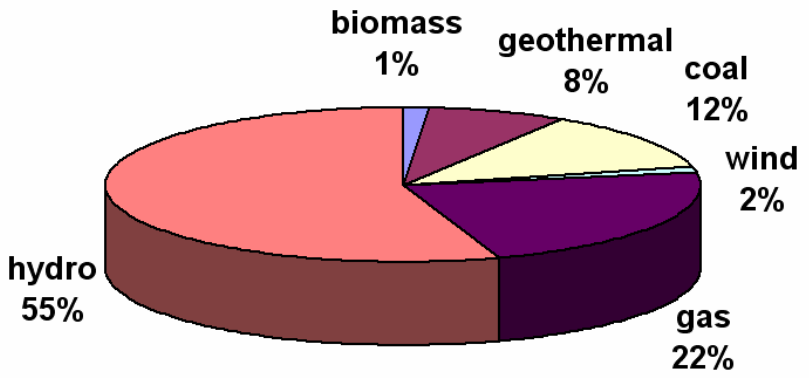
Agenda

- Where does energy come from and go to?
- Mass storage device power use; disk, SSD, tape, optical
- Multi tier vs disk only
- HSM is green
- What SGI is doing to help customers reach sustainability
- Q & A

Where does electricity come from ?



- **43 billion kWh of electricity generated in NZ in 2006**

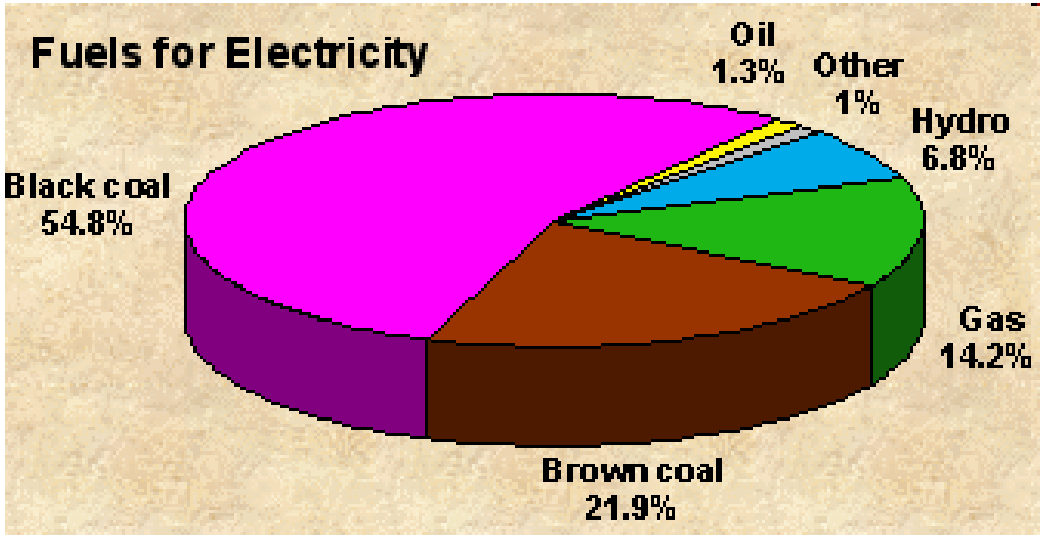


Source: World Nuclear Association

Australian electricity sources are different

- **In 2006 power stations produced 255 billion kilowatt hours (TWh) of electricity, 65% more than the 1990 level and growing at 3.3% pa.**
- **18 TWh is used by power stations, leaving 237 TWh net production. 17 TWh is lost or used in transmission and 9-10 more in energy sector consumption, leaving 210 TWh for final consumption (or 187 TWh apart from aluminium exports)**
- **Much of the energy exported is used for generating electricity overseas; three times as much black coal is exported as is used**

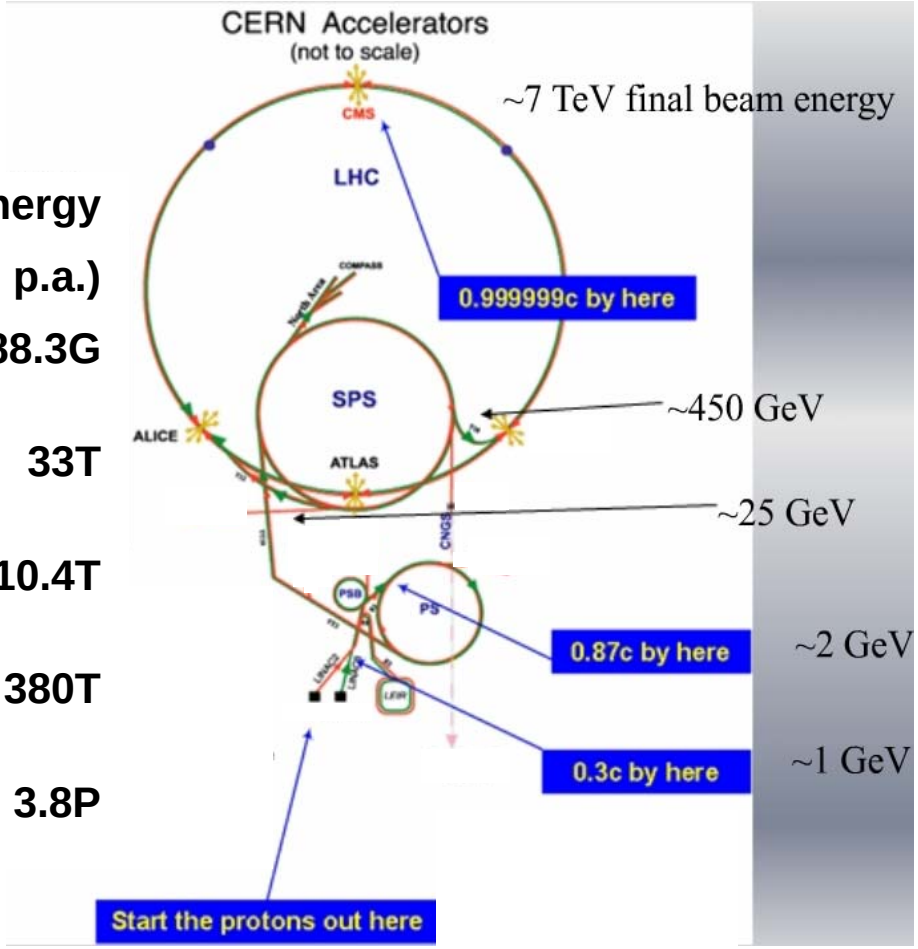
- **25% of electricity produced is lost in generation and transmission**
- **90% energy sources are fossil fuels**



Source: World Nuclear Association

Power: Energy used per unit of time

	Power (Watts)	Energy (Wh p.a.)	Energy (J p.a.)
Hubble telescope	2.8K	24.5M	88.3G
Space shuttle launch			33T
1920P Altix	330K	2.9G	10.4T
Synchrotron	12M	105G	380T
CERN LHC	120M	1T	3.8P

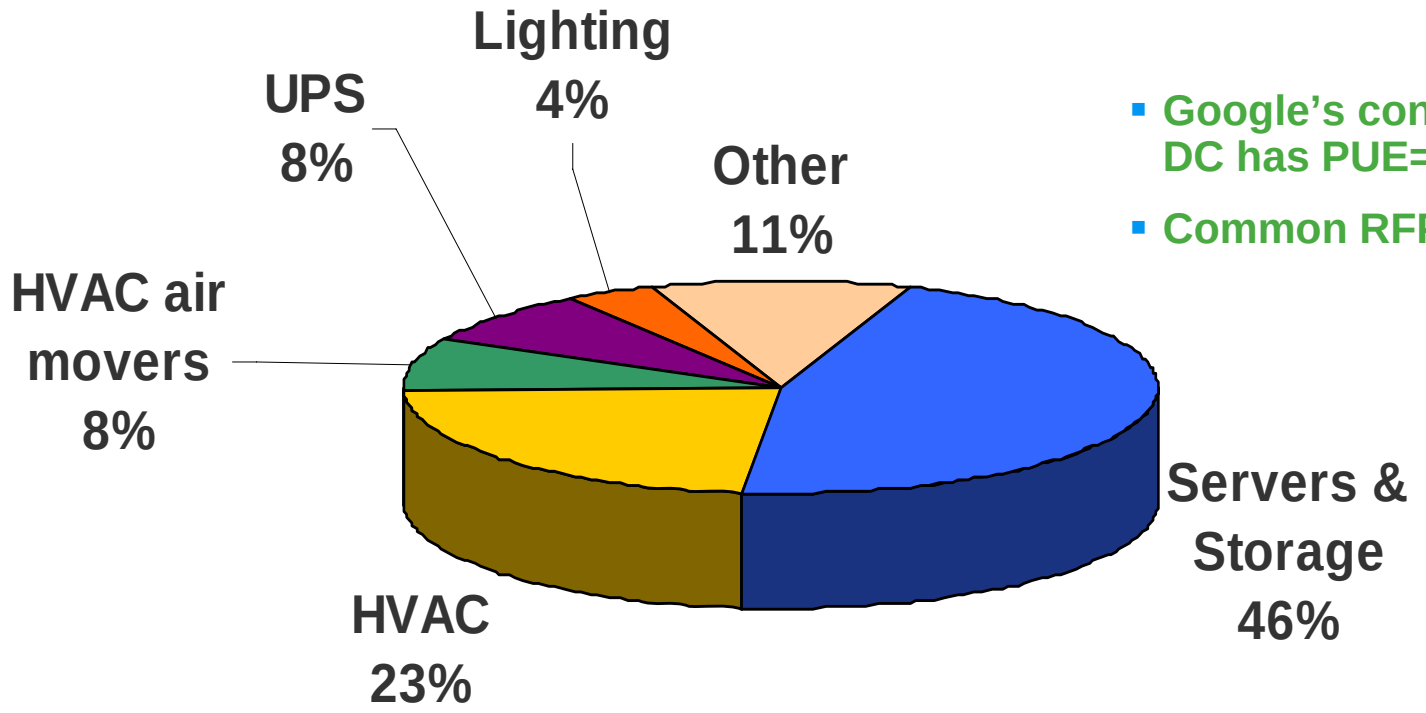


- Because the Data Centre is always on it consumes a lot of energy
- Conventional research uses energy too!

http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/faq/lhc-energy-consumption.htm

Where does energy go?

- **Average of 12 Data Centres in NY&CA: Co-Location, Hosting, Government, Financial Institution, Telecom, Scientific Computing, Data Storage**
- **PUE = Total datacenter power draw / IT load power draw ~ 2**



- **Google's containerised DC has PUE=1.2**
- **Common RFP reqmnt**

Source: Lawrence Berkeley National Laboratory, http://hightech.lbl.gov/benchmarking-dc.html

There's a Real Problem

- Energy consumption for US data centers and servers will nearly double by 2011… requiring an additional 10 power plants (EPA Study)
- Power consumption in data centers doubled from 2001 to 2005 (Lawrence Berkeley / Stanford study)
- Half the data centers will not have sufficient power for expansion by 2008 (Gartner survey '06)
- In the highest level of redundancy and reliability data centers, for every kilowatt used for processing, $22,000 is spent on power and cooling infrastructure (Uptime Institute)
- Through 2009, energy costs will emerge as the second highest operating cost in 70% of data centers worldwide (Gartner survey '06)

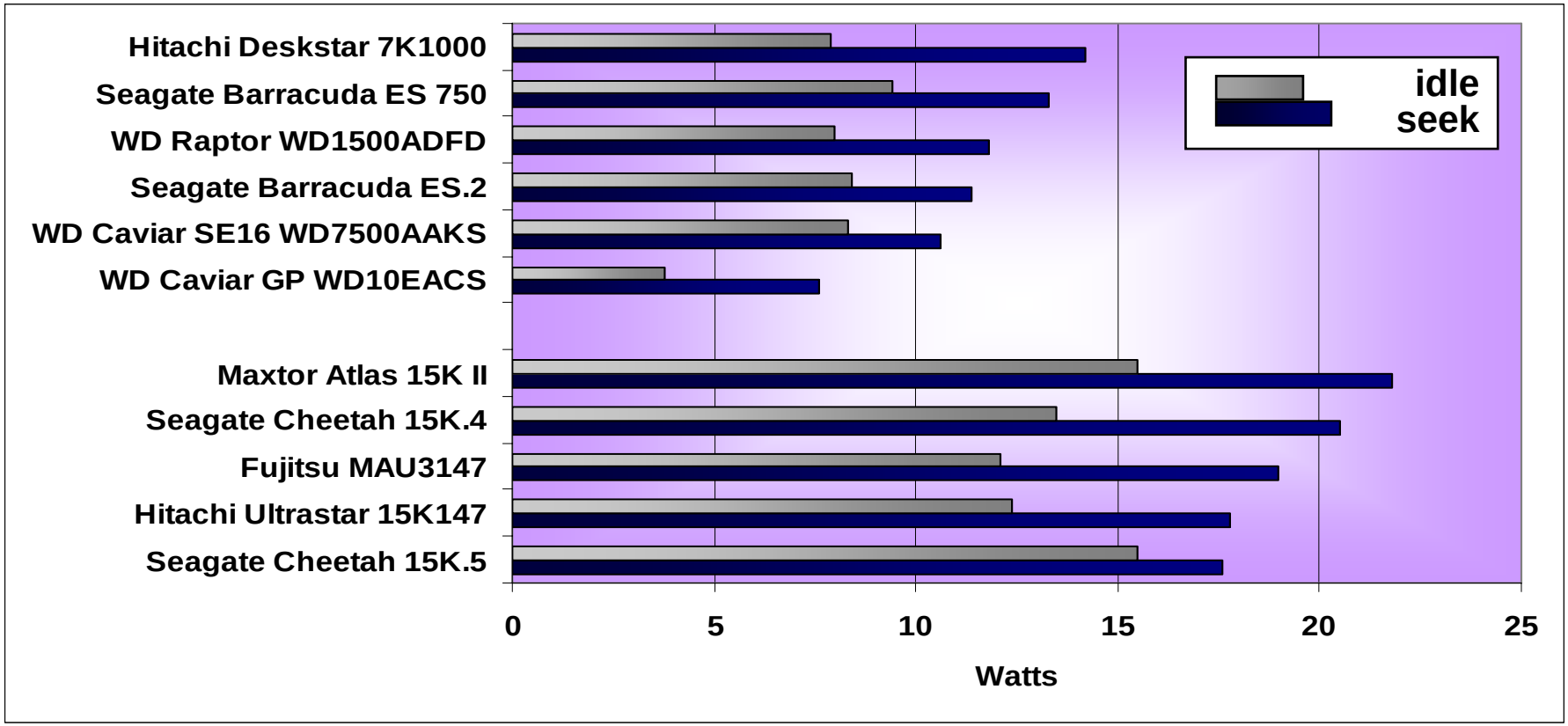
sgi

Concerns No Longer Focused on Servers

- Servers currently account for about 60% of data center power… storage will soon take that place (Glass House ’07)
- eWeek.com
  - Average number of storage terabytes maintained by Fortune 1000 organizations in 2004: 138
  - Average number of storage terabytes maintained by Fortune 1000 organizations in October 2006: 600

sgi

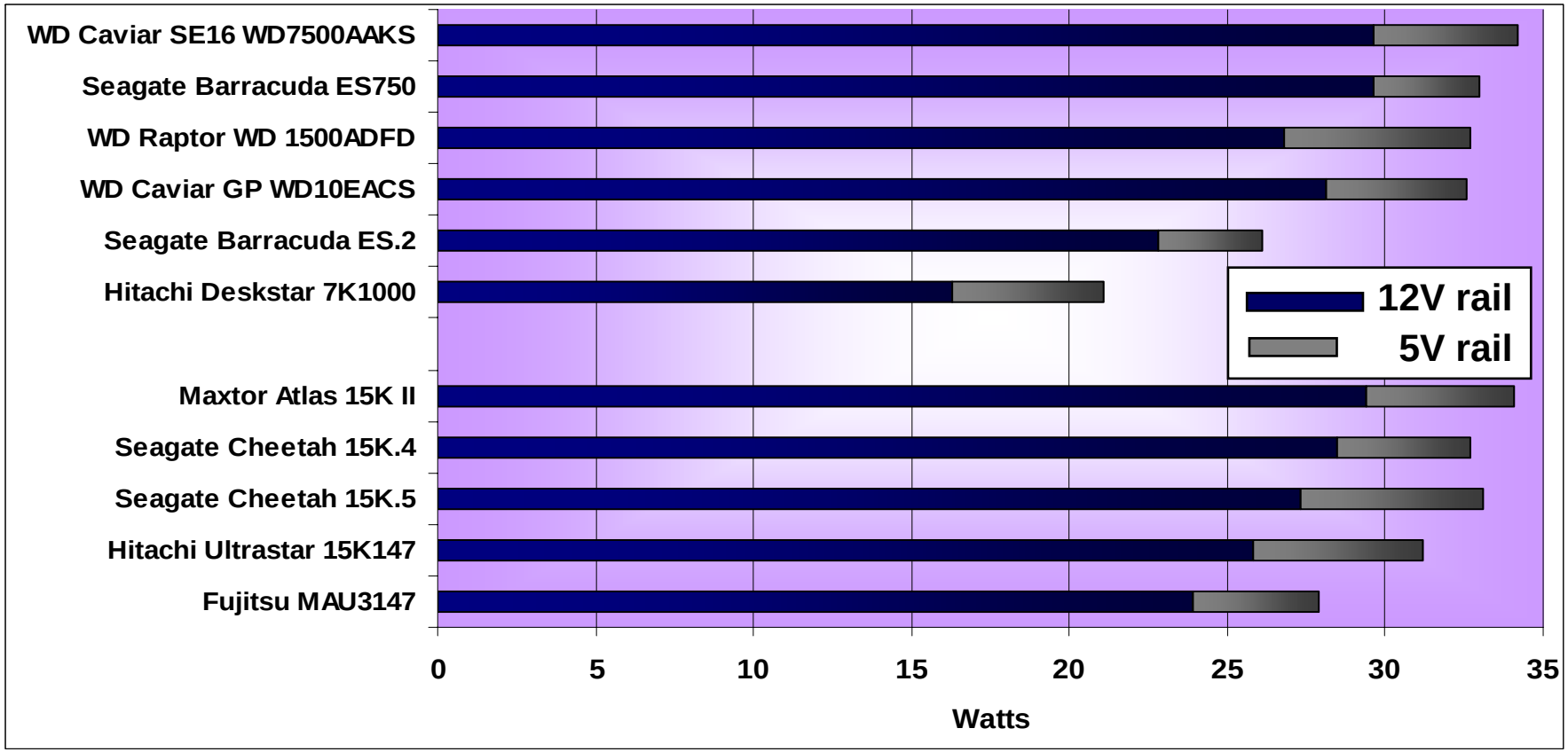
Disk drive power dissipation



- SATA drives use 8 – 10W
- SCSI drive use 12 – 20W
- SAS drives use ~1W more than SCSI
- FC drives use ~2W more than SCSI

Source: StorageReview.com

Disk drive startup power consumption



- **SATA similar to SCSI**
- **Drives use ~1.5x (scsi) – 2x (SATA) peak power to spin up**

Source: StorageReview.com

Energy required for 20TB of disk

	Capacity (GB)	Speed (rpm)	Type	Power (Watts)	Drives	Energy p.a. (KWh)	Joules (GJ)
Seagate Savvio 15K.1	73	15000	SAS	8.1	274	19,357	70
Seagate Cheetah 15K.4	146	15000	SCSI	19.1	137	22,936	83
Hitachi 10K300	300	10000	SCSI	15.8	67	9,257	33
Seagate Cheetah NS	400	10000	SAS	11.1	50	4,883	18
Seagate Barracuda ES	750	7200	SATA	12.5	27	2,927	11
Hitachi 7K1000	1000	7200	SATA	12.9	20	2,269	8
WD GP WD10EACS	1000	5400	SATA	6.8	20	1,199	4.3
Seagate Barracuda PR	1500	7200	SATA	10.9	14	1,274	4.5

- Power based on 80% seeks 20% idle
- 1TB drives use ~10x less power per TB than 146GB SCSI drives
- 1TB drives use ~2x less power per TB than 400GB SCSI drives
- SCSI/SAS/FC has 2x IOPs and less latency
- New drives from Seagate and Western Digital use 50% less power

Tape

SAIT / LTO4

800GB
45 / 120MB/s native
19sec load time
57sec file access time
30W (active)
5W (idle)

TS1130

1000GB
160MB/s native
13sec load time
49sec data access
46W (active)
17W (idle)

T10000B

Capacity cart	Sport cart
1000GB	75GB
120MB/s native	
16sec load time	
48sec data access	12sec
63W (N/A)	

- **No power used to retain data**
- **Very high data transfer rates per device – sequential access**
- **Tape Libraries 10TB - 8PB+**
- **30 year media life**

Sources: Manufacturers data sheets; Clipper Group

sgi

Optical Storage

Holographic (3D)

300GB
20MB/s data transfer
5 sec load time
50 year archive life
Blue laser
10,000,000 reads
WORM
80W

Ultra Dense Optical (UDO)

60GB (UDO2)
12MB/s data transfer
6.5sec media access time
50 year archive life
Blue laser
WORM
310W (19TB jbox)
20W (drive)

Blu-Ray / DVD

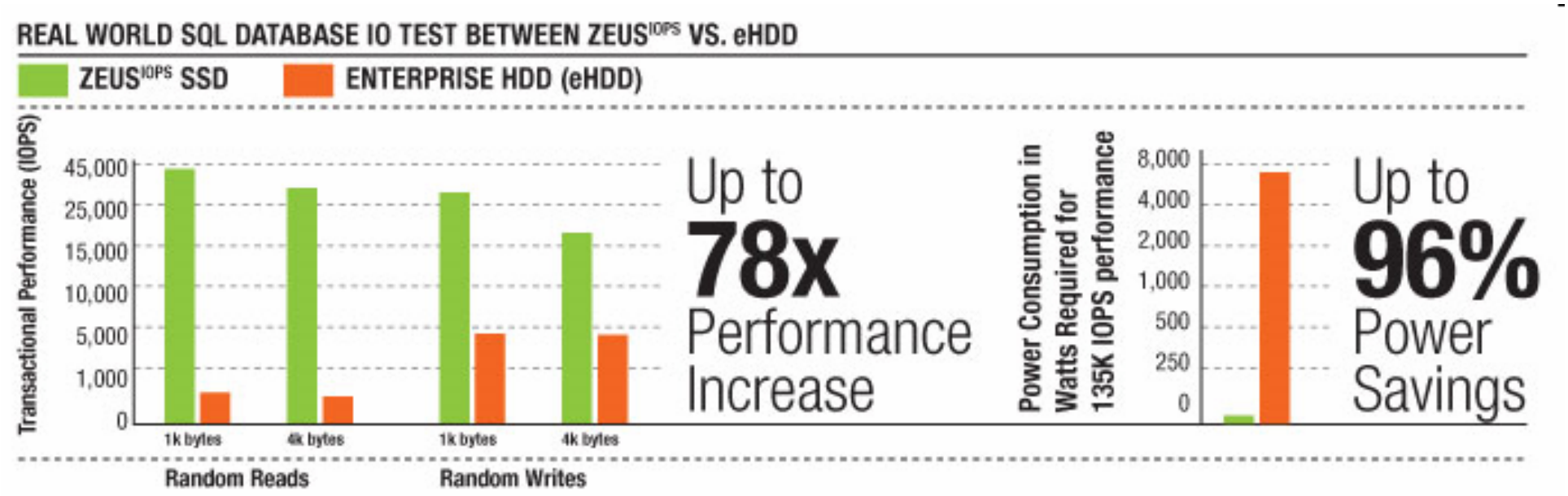
BD-R, BD-RE, DVD-R, DVD-RW

50GB / 5GB per disc
9MB/s data transfer
4.5sec media access time
Blue laser
Re-write 10,000 times
140W (jukebox)
20W (drive)

- **Like tape, optical media requires no electricity to store data**
- **Data transfer rates are low compared to tape and disk**
- **Random access once media loaded**
- **Largest jukebox capacity 30TB**

Sources: http://blogs.zdnet.com/storage/?p=313
http://www.incom.eu/jvc_mc-blx.html?L=1, http://www.kintronics.com/jukebox.html

Solid State Disk



- **SSDs use 5W idle and 8W during data transfers**
- **Excellent performance**
- **Very expensive for significant capacity**

Case Studies, tiered storage

Small

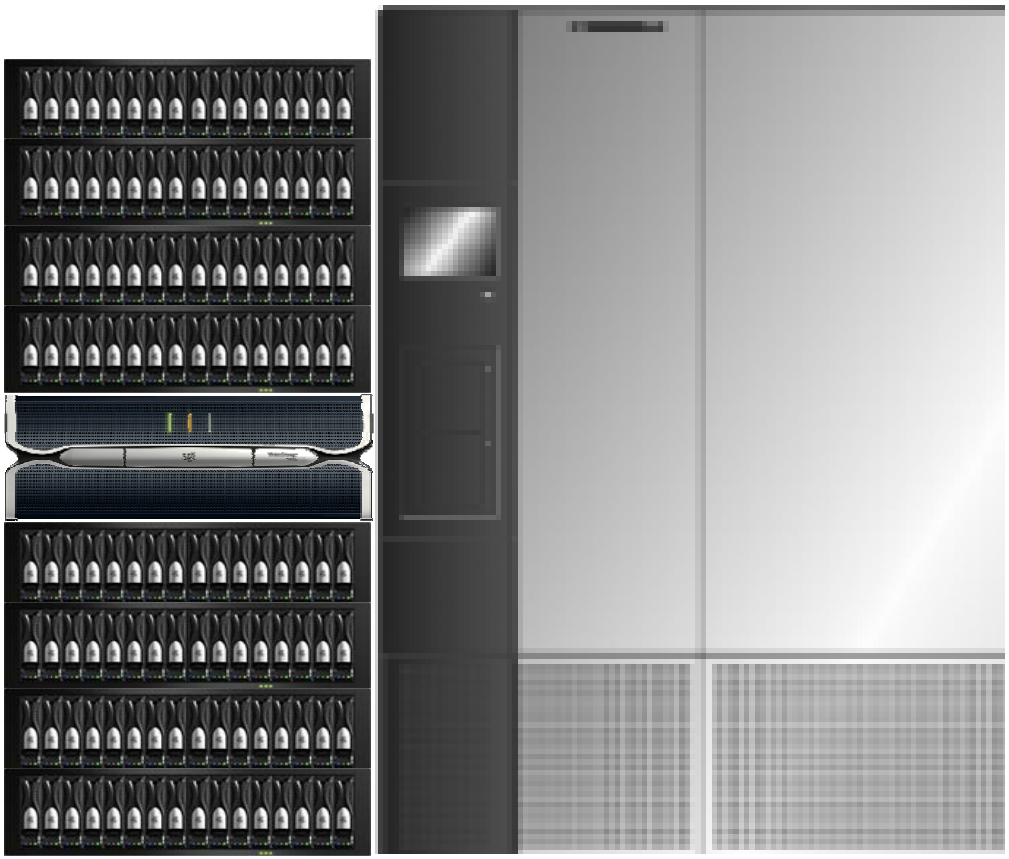
50TB

Medium

250TB

Large

1PB



50TB Storage Solution Power Comparison

Model	Watts	Type	Drives	Capacity (TB)
SGI IS220	510	SAS	30	10
+ SL T50	265	LTO4	2	40
	775			50
		2.3x		
Disk only				
SGI IS4000	1,793	SATA	64	51

- 8+2 RAID6
- 400GB FC drives for HSM
- 1TB SATA drives disk only

vs

250TB Storage Solution Power Comparison

Model	Watts	Type	Drives	Capacity (TB)
SGI IS4000	1,655	FC	64	20
STK SL500	420	LTO4	6	232
	2,079			252

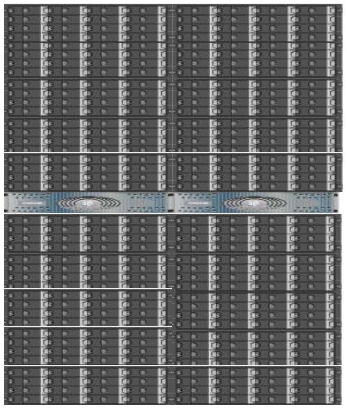
4.6x

Disk only

SGI IS6700	9,600	SATA	320	256

- 8+2 RAID6
- 400GB FC drives for HSM
- 1TB SATA drives disk only

VS



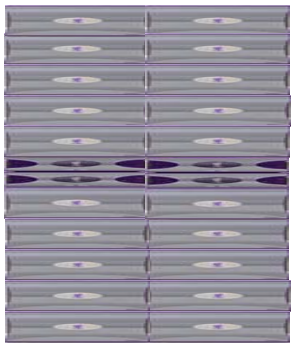
Source:http://www.sun.com/storagetek/tape_storage/tape_libraries/sl500/calc/calculator.jsp

1PB Storage Solution Power Comparison

Model	Watts	Type	Drives	Capacity (TB)
SGI IS4600	4,092	FC	128	41
SL T950	966	LTO4	12	1,154
	5,058			1,195
		7x		
Disk only				
SGI IS15000	35,800	SATA	1,200	960

- 8+2 RAID6
- 400GB FC drives for HSM
- 1TB SATA drives disk only

VS



Source: Manufacturers data sheets

Results Summary

	Multi tier	Disk only	Greener
Small	17W/TB	35W/TB	2.3x
Medium	7.6W/TB	37.5W/TB	4.6x
Large	4.4W/TB	37W/TB	7x

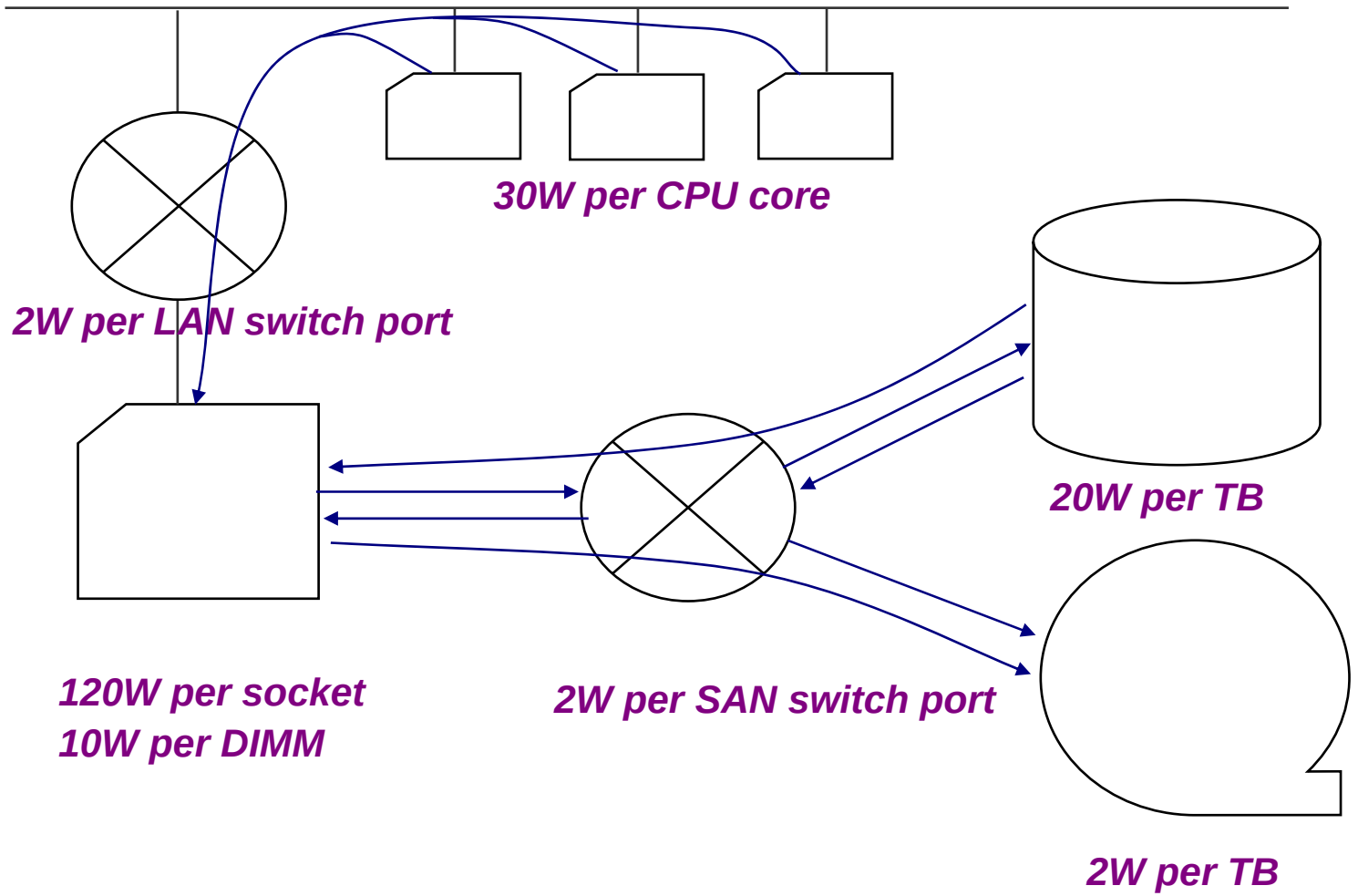
- In all cases multi tier RAIDs used SAS or FC disks to handle tape and network IO
- A fundamental difference in service is the higher latency involved in retrieving data from tape

sgi

HSM should be first choice for long term storage

- Disk and tape have a place in the data center
- With HSM, a relatively small volume of high performance disk satisfy user requests with good bandwidth and IOPS characteristics
- The bulk of data resides on tape when it's not active
- HSM provides the 'smarts' to move data to the appropriate storage medium in response to usage
- It is easy being green WRT storage – and capital investment is lower than a 'disk only' solution

Without HSM, Backups Consume More Energy



SGI Contribution to Energy Initiatives

tc99.ashraetcs.org/

www.thegreengrid.org/

www.climatesaverscomputing.org/

www.80plus.org

www.epa.gov

www.energystar.gov

www.eere.energy.gov/femp

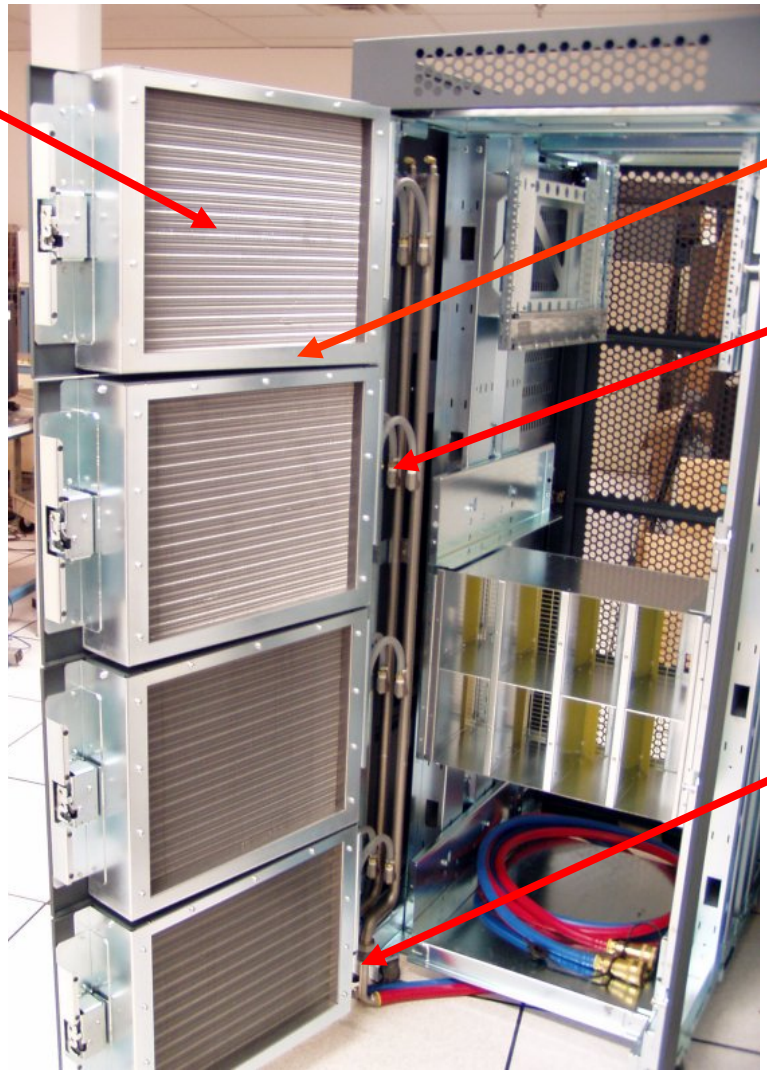
www.spec.org/power_ssj2008/

Building Water-Cooled Racks

(4) Individual Coils

Target Heat Rejection
95% water / 05% air

Chilled-Water Supply
45°F to 60°F (7.2°C to 15.6°C)
14.4 gpm (3.3 m^3/hr) Max.



Condensate Drain Pan

Branch Feed to Individual Coil

3/4” (1.91 cm)

Swivel Coupling to Supply Hose

Free Cooling Solution

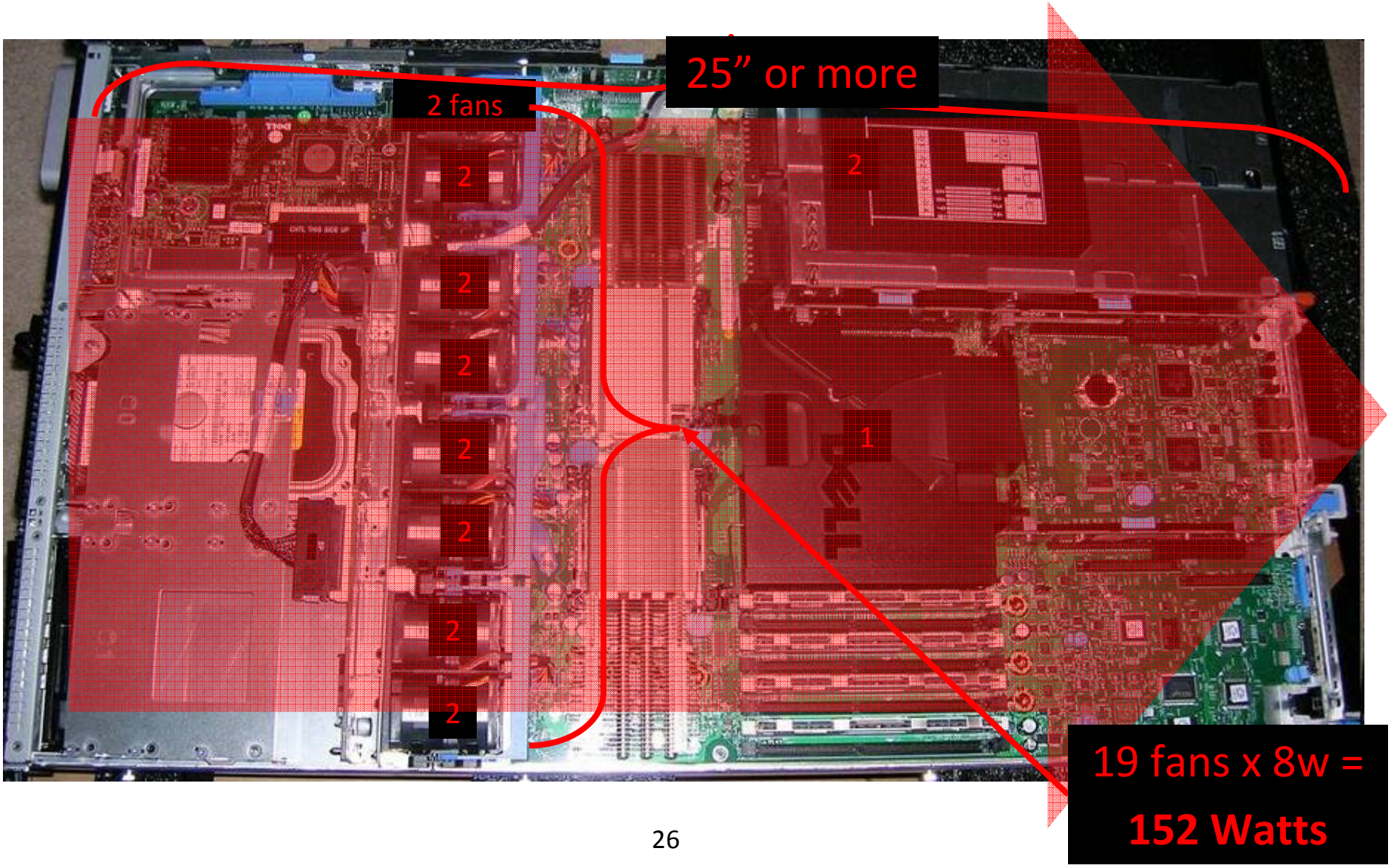
- Free Cooling technology saves on energy compared to the traditional refrigerated equipment, especially during autumn, winter and spring
  - Achieved by utilizing a low energy, dry air cooler and Free Cooling technology
  - Can partially or fully reduce the chiller depending on ambient air temperature.
  - Most effective when utilizing rear door and in-row water cooled solutions
- Return on investment,
  - payback in as little as six months
  - eligible for government grants (?)
  - Fully compliant with the latest EU directives for data centres.
- SGI PS development for easy installation on any chilled water systems. Existing installations can benefit from significant energy savings by retro-fitting.

Rack Centric Power Options

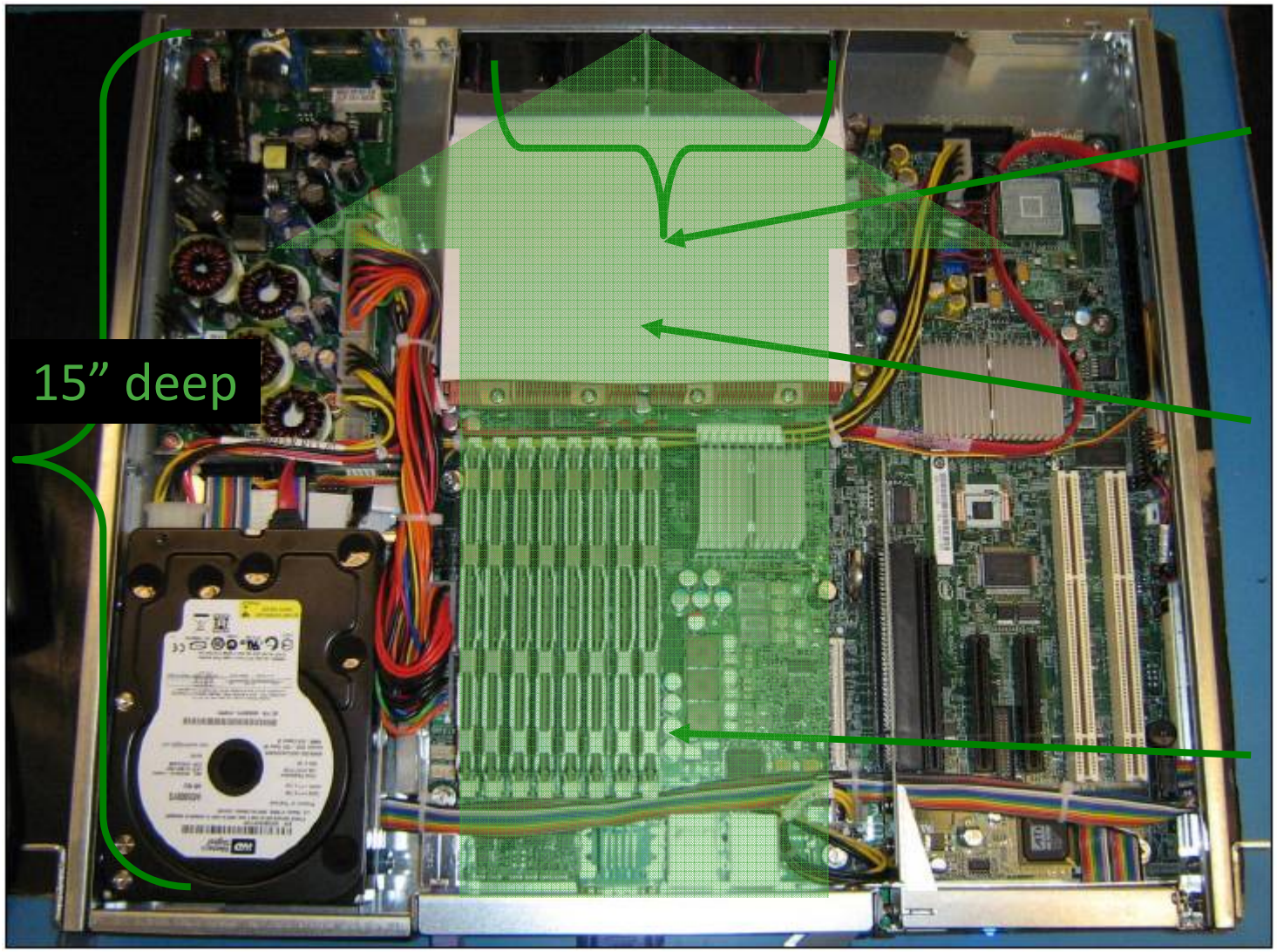
- AC
  - Plugs, cords, PDU's, ATS's
  - Rack-level
- Rectified DC
  - Rack level
- Direct DC

DC Card

Server Cooling Design - Dell 1950



Rackable Eco-Logical™ Chassis Design

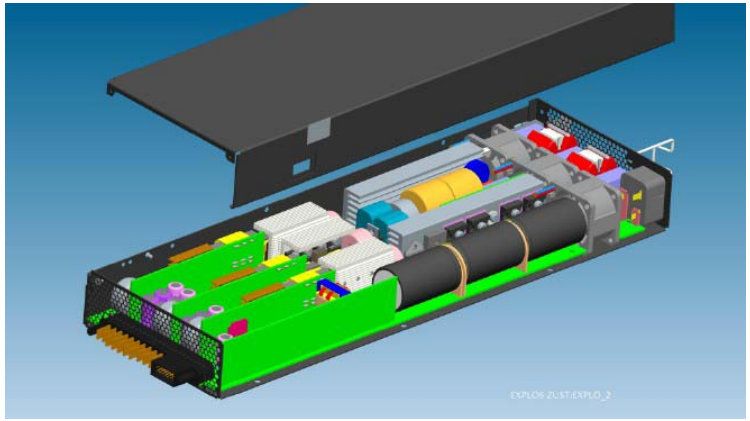


4 fans x 4.5w =
18 Watts

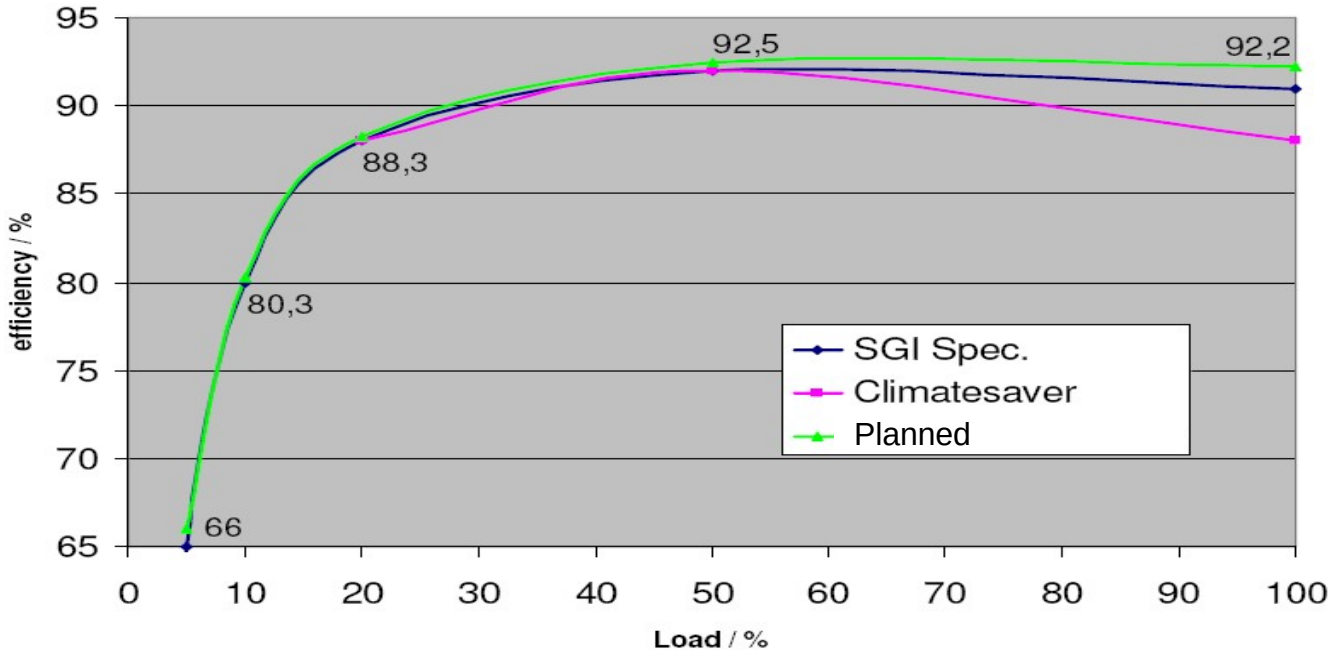
Cooling Shroud

Smart
Component
Placement

Building Award Winning Power Supplies

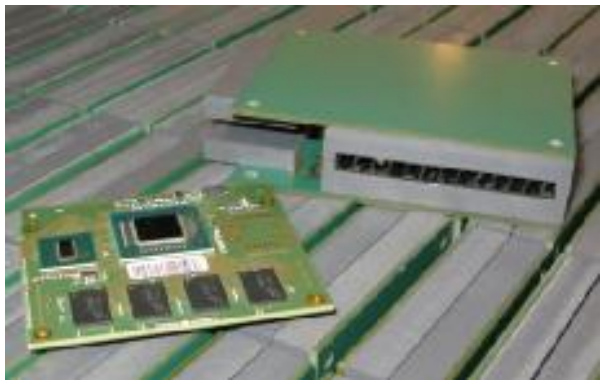


- SGI awarded 80plus silver and bronze certification for Altix systems shipping since January 2007.
- 2837W Power Supply exceeds ClimateSaver specifications for 1 and 2U servers.



Designing More Efficient HPC Systems

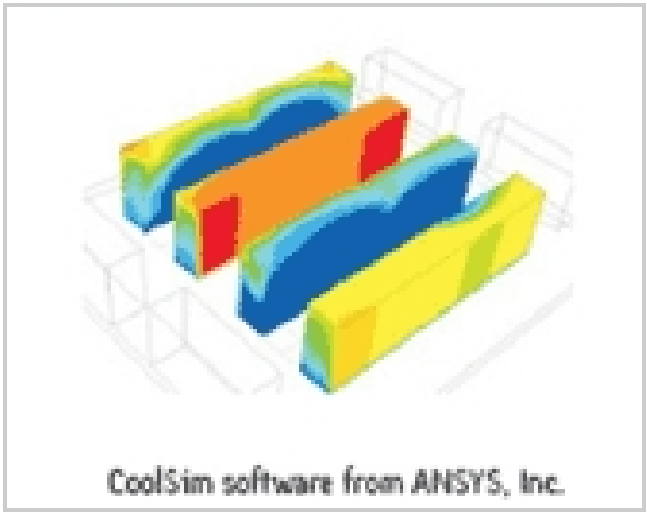
- SGI ‘Molecule’ concept computer
- 10x core density, ½ the power usage
- Commodity components (mobile device processor)
- 20x memory bw of a single rack x86 cluster
- Intel Atom N330 based; 8W TDP compare 120 W for XEON

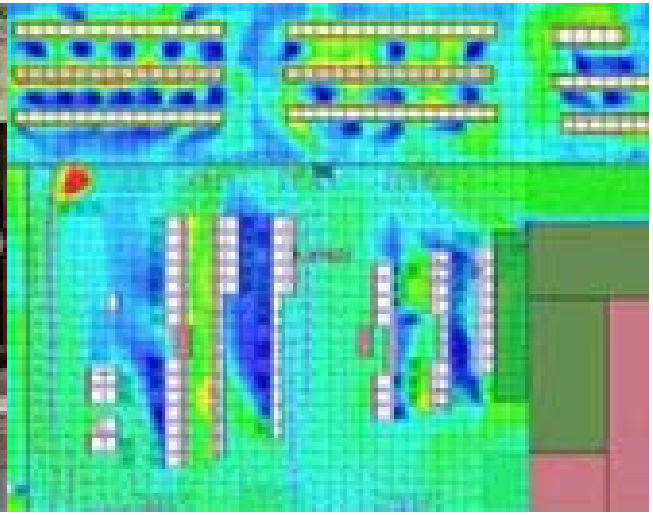


- ‘Kelvin’ liquid cooling system

sgi

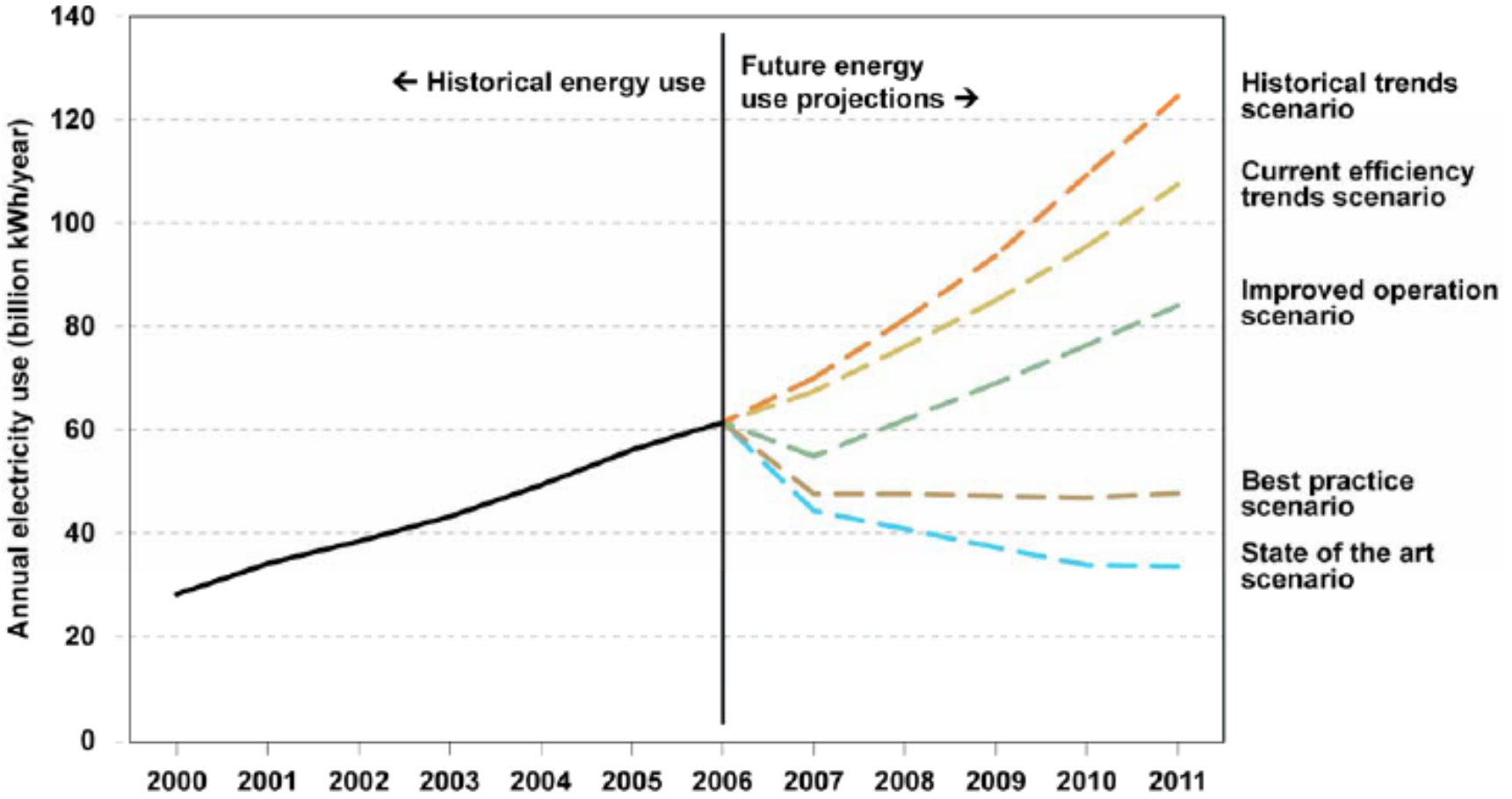
Delivering Green Consulting Services





- Data Center Assessment
- Dense Systems Assessment
- Energy Efficiency Assessment
- Power Distribution Assessment
- Storage Assessment

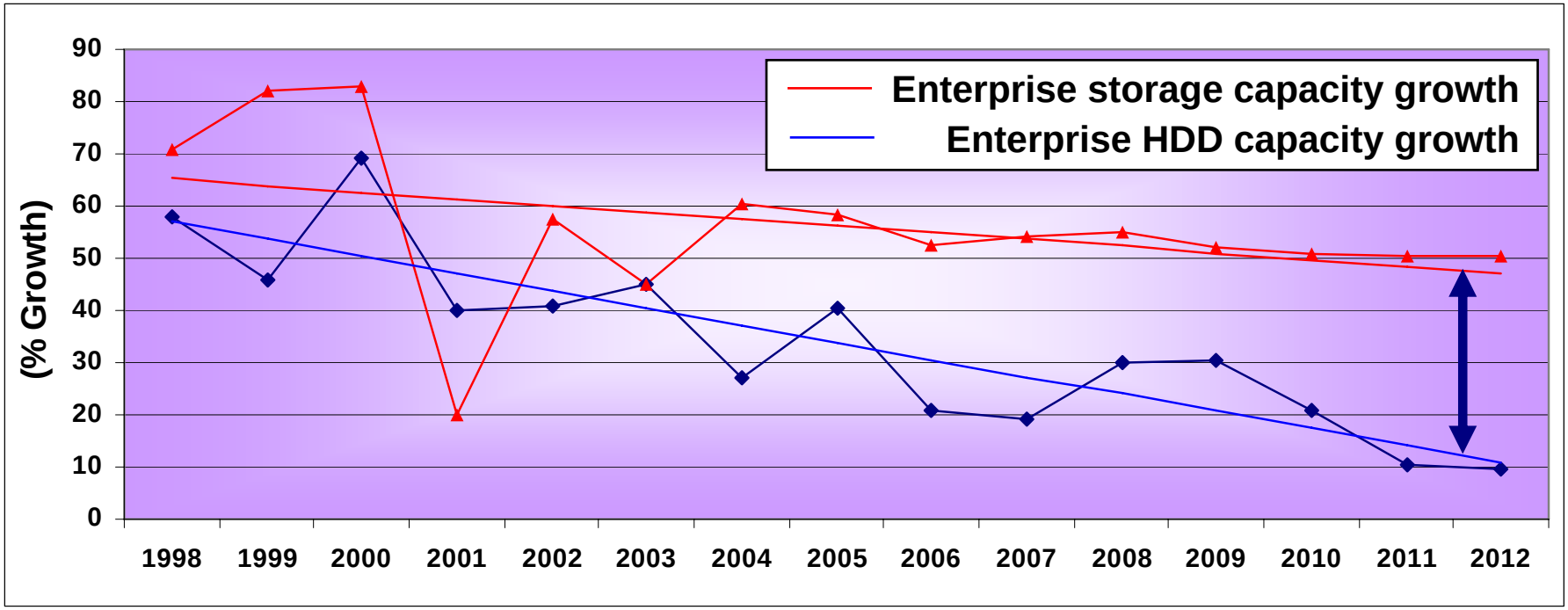
Projected Electricity Use, All Scenarios 2007 to 2011



Source: U.S. EPA Report to Congress on Server and Data Center Energy Efficiency, Aug 07

sgi

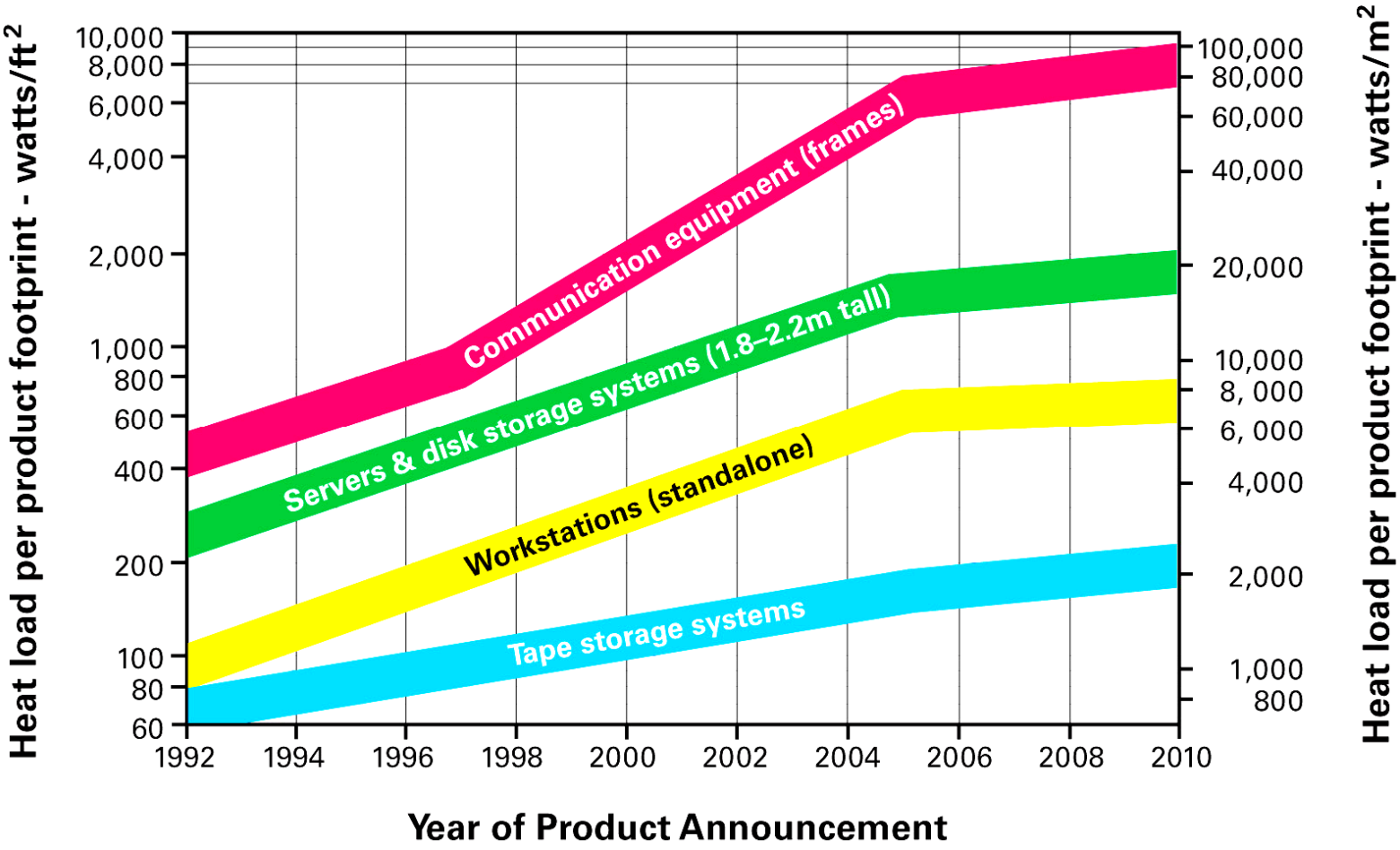
The problem with disk



- Drive capacity growing at diminishing rate ~30%
- Enterprise capacity increasing ~ constant 50%
- Drive capacity growth exceeds IOPs growth
- Forces tending to increase drive counts (and power consumption)

Source: IDC 2008

Power dissipation per unit area

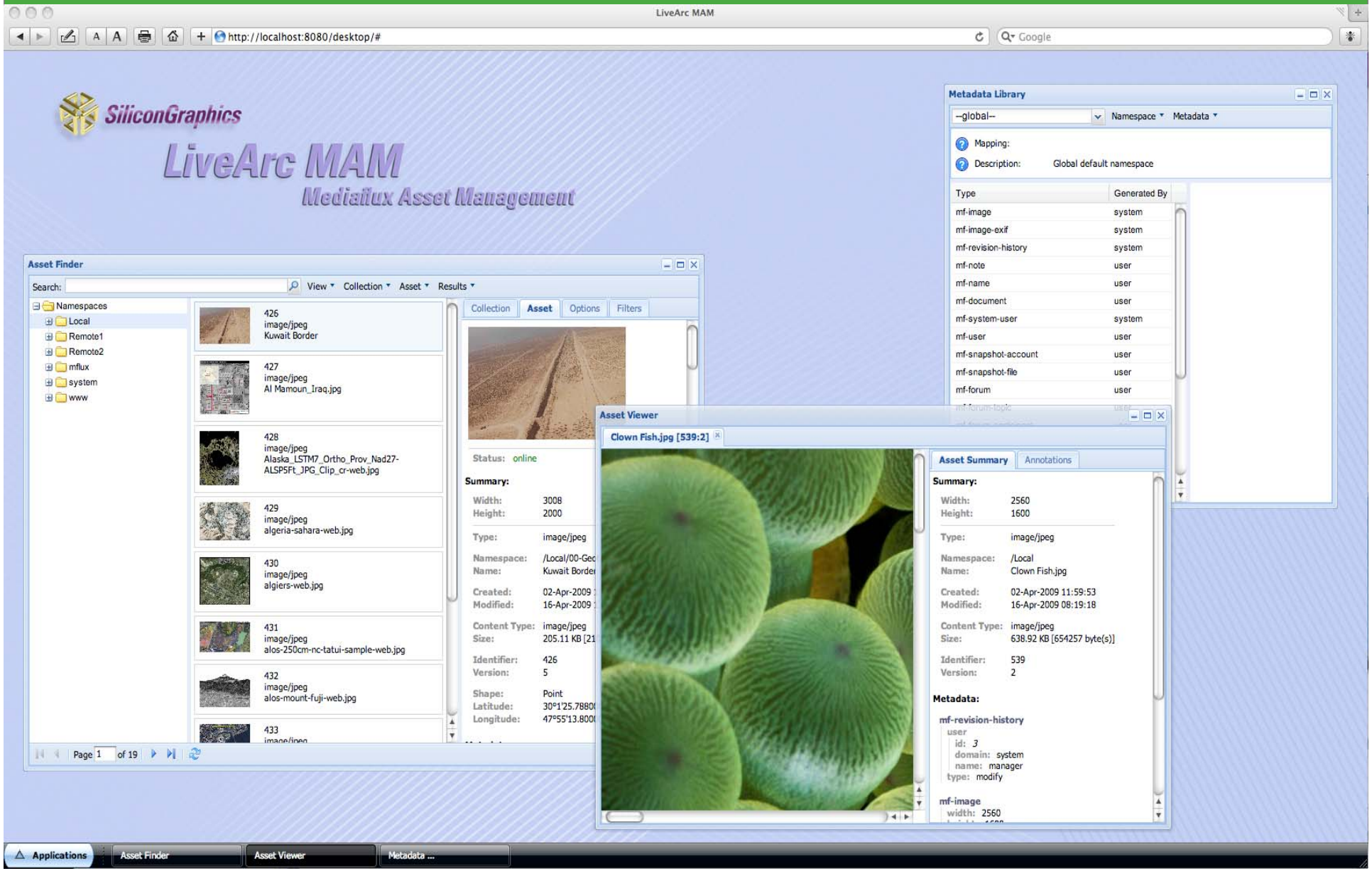


Source: Uptime Institute

DMF Saving the planet

- **Automated tape libraries use a fraction of the power of RAID storage**
- **SATA disk is greener than SCIS/SAS/FC but isn't appropriate for all workloads**
- **It is better for the planet for users to manage metadata better (MediaFlux) and wait a little longer for data**
- **"It's cheaper to push everything off to tape than have someone look at it"**
- **Centralised storage services improve data protection and data access to wider audiences**
- **Every Watt saved reduces green house gas emissions**

DMF integration with Mediaflux



designed. engineered. results.