You Survived the Great Recession, Now What?

Toward a Sustainable Strategy for Cost-Containment, Compliance, Continuity and Carbon Footprint Reduction

Presented by
Jon William Toigo
CEO Toigo Partners International
Founder The Data Management Institute

TECHNOLOGY STRATEGY

TOIGO PARTNERS INTERNATIONAL LLC

Is It Over? Really?

- Depending on how you cook the books
  - IT Budgets continue to be lean
  - Staff levels remain frozen
  - “Do more with less” remains the dominant mantra
  - But, economists say that two quarters of consecutive growth translates to a turnaround

In Every Company I Visit Today…

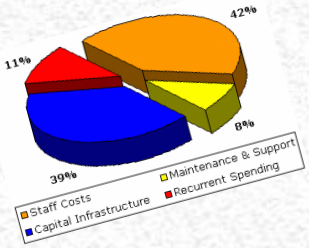
- The four issues that are front of mind for business planners:

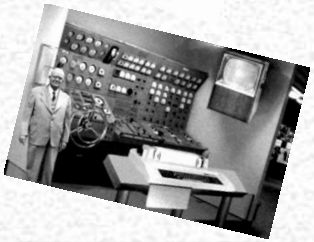
One Significant Change

Decreasing Headcount
Equals
Increasing Automation Dependency

Coming After a Decade of Questioning About the Worth of IT

- Management view of IT (and vice versa) has been poor since 2001…
- Exacerbated by current pain…
  - Budget/credit issues
  - Staffing issues
  - Cost of maintenance/downtime
- Showing up in interesting ways
  - "Risk aversion" equals "retrenchment" – Some seek safety in "love brands"
  - Frantic adoption of "methodologies"
  - Keeping gear longer, dropping OEM maintenance
  - Interest in Outsourcing/Clouds



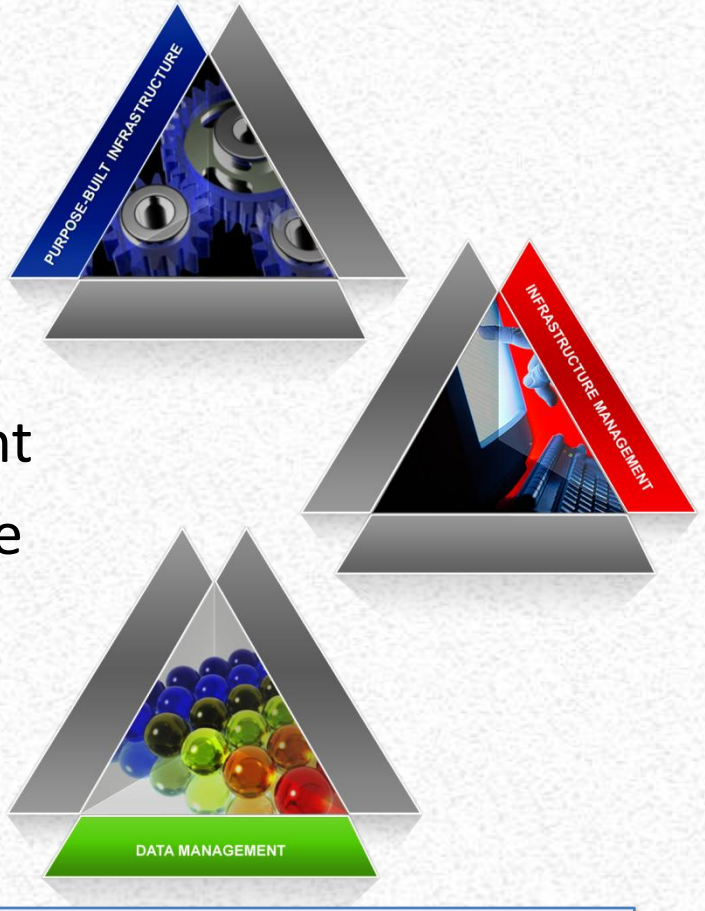


Today's C-4 Challenge:
A Possibility to Heal the "Rift"

- The Front Office
  - Top Line Growth
  - Cost-Savings
  - Governance, Risk & Compliance
  - Return on Investment
- The Back Office
  - Infrastructure Planning
  - Performance Management
  - Service Level Improvement
  - New Information Products

Current Situation Not Just a Reflection of the Economy…

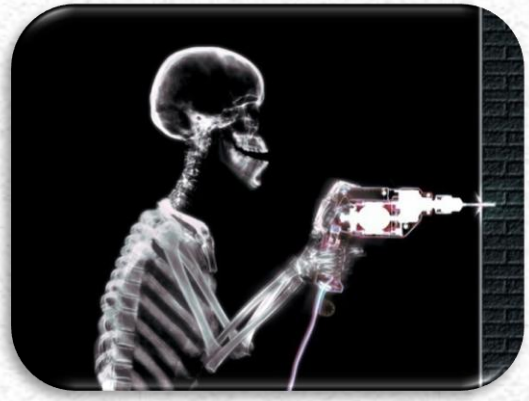
- The C-4 issues have been building steadily, a function of…
  1. Failure to "purpose-build" infrastructure…especially in the distributed computing environment
  2. Failure to instrument infrastructure for monitoring, measurement and management
  3. Failure to manage data effectively



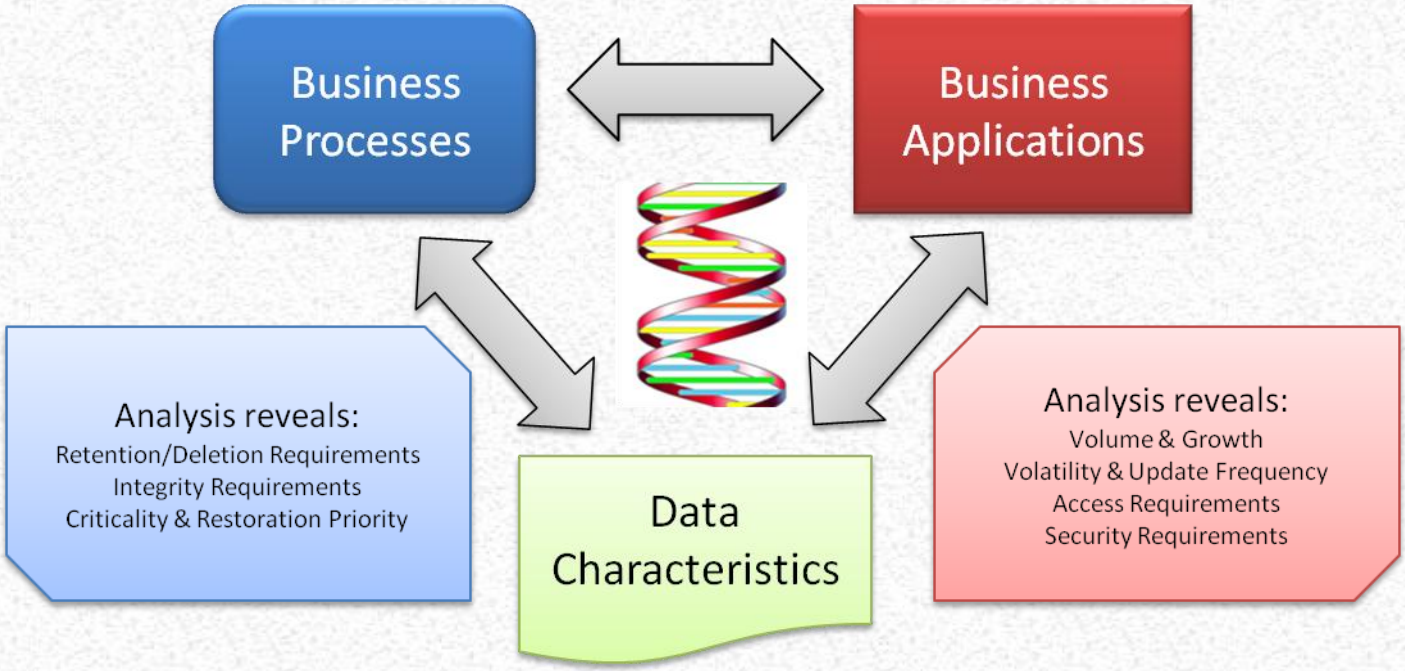
The current recession has only underscored the problems and, perhaps, given us a "pause" to think **strategically** about ways for solving them…

Let's Drill Down

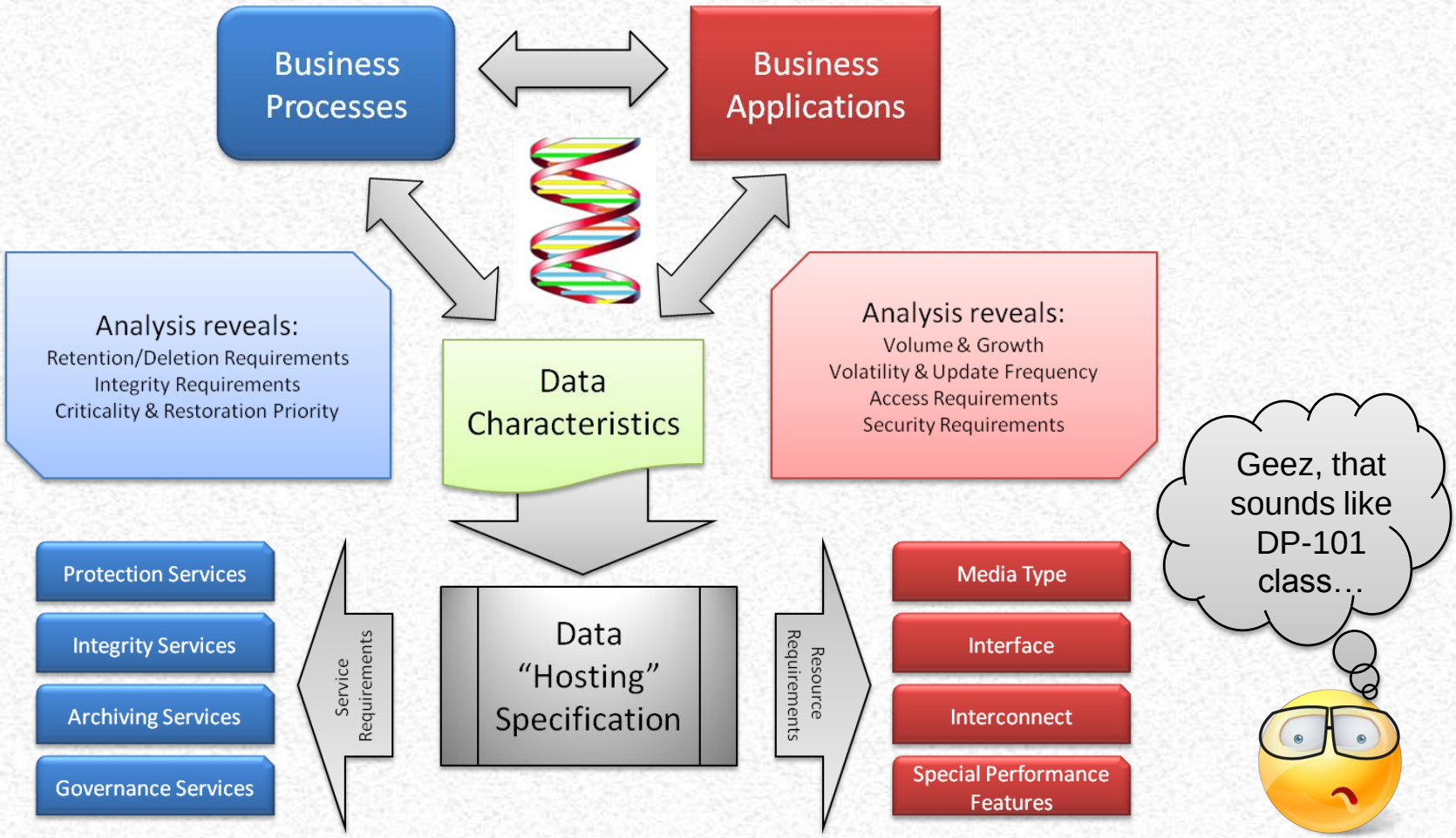
- What is "purpose-building" infrastructure?
  - n. a kind of "intelligent design" – a deliberate and well conceived matching of infrastructure "services" (resources and functions) to meet well-defined business needs…
  - v. the act of buying or developing and deploying technology that serves a business need and not simply outsourcing your thinking process to vendors whose top line growth is more important than your top line growth



Business Process and Business Apps are Supposed to Define Workload, Which Defines Infrastructure…

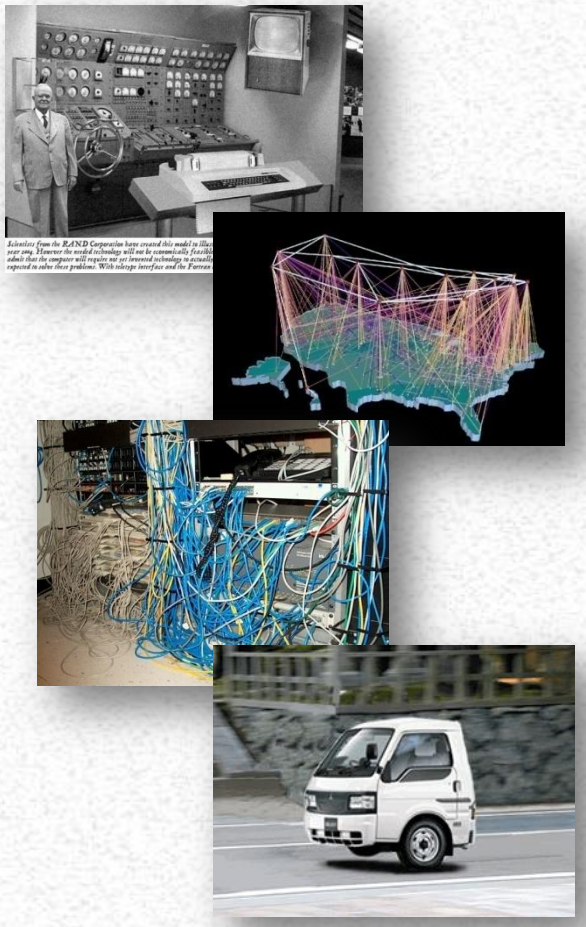


Workload (Data Characteristics) Identifies Necessary Resources and Services



Chances are Good That Your Infrastructure Was Originally Built That Way…

- Might have started with a data-centric, business process-focused view, but it rarely stayed that way over the years
- Why?
  - Management changed: Different masters bring different views…
  - Technology changed: Distributed systems, TCP/IP, the Internet, the "V" word, *et al*…
  - Business changed: Mergers and acquisitions, new products, new go to market strategies…
  - Attitudes changed: "IT is our strategic differentiator," "Bricks and Sticks are Dead," "Does IT Matter?," "Legal is running IT now," "Web 2.0," "Enterprise 2.0," "Mash-ups will replace IT applications," Clouds will replace data centers, *blah blah blah*…



Biggest Problems Attributable to Triumph of Marketecture

- Marketecture: n. that which otherwise good architecture becomes when subjected to vendor marketing departments; see *kluge*
  - The rise of distributed systems: mainframe deconstruction gone awry…
  - The network is the computer: distributed computing gone awry…
  - SANs: storage infrastructure gone awry
  - x86 server virtualization: LPARs gone awry
  - "Cloud computing:" multi-tenancy gone awry

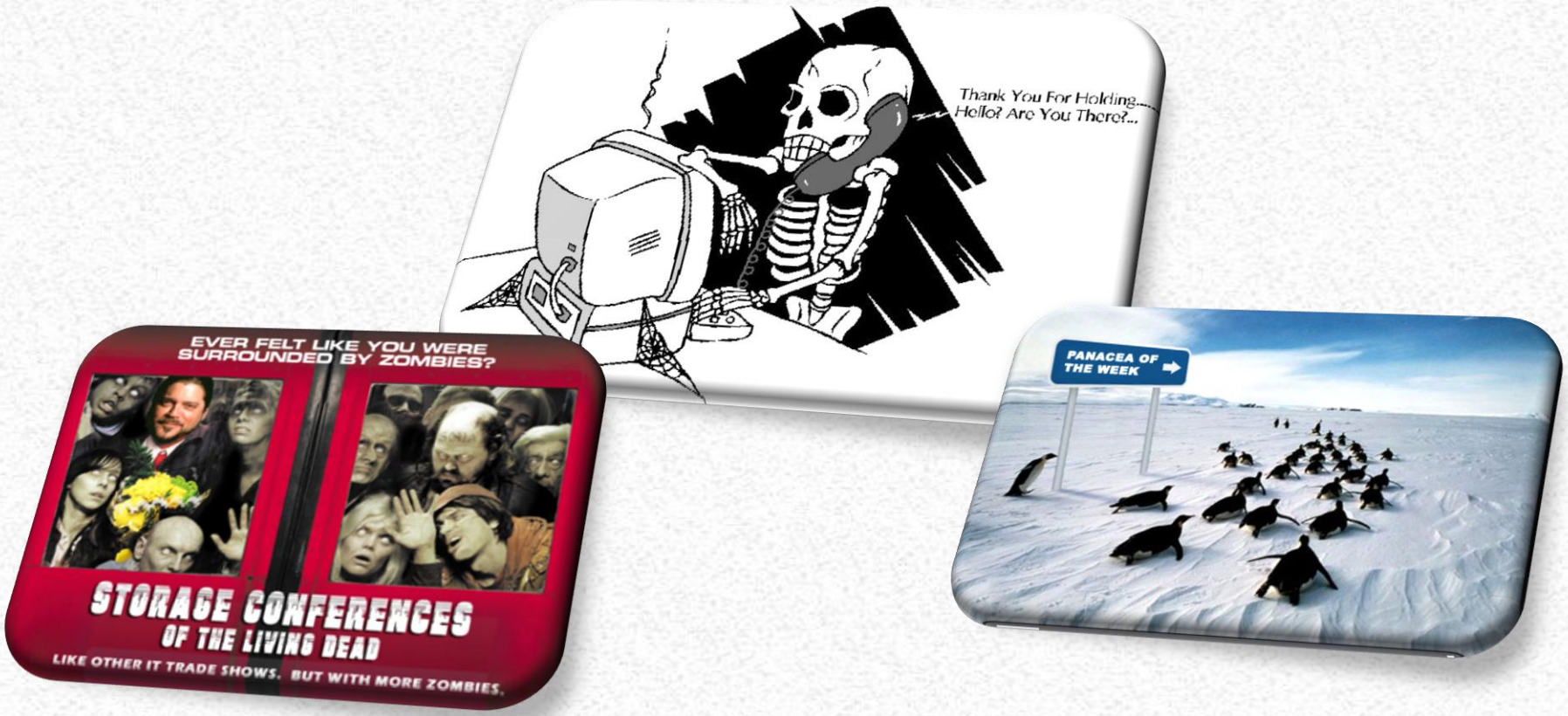
Marketecture Myths Abound

- “IT isn’t about building and managing infrastructure anymore, it’s about managing a few trusted vendors…”
- “Smarter systems reduce management burden (and its associated labor costs) and fit the needs of all businesses well…”
- “Resource utilization efficiency is best achieved through server virtualization…”
- “Why deal with problems? Source your services from a cloud…”

Marketecture Succeeds Because…

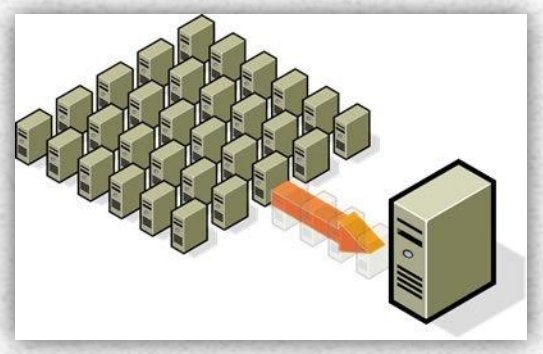
- Leading vendors tend to sell to the Front Office, not to the Back Office (where IT lives)…
  - Eschewing "tech speak" and framing pitch in terms of business problem solving
  - Purchasing mind share with PPV analyst reports
  - Casting aspersions on internal IT competency
  - Plus, the usual schmoozing…

Not to Suggest that IT is Blameless…



No Silver Bullets

- Server virtualization isn't the big fix everyone is hoping for…
- Nothing wrong in concept: mainframe LPARs and PRISM around for a couple of decades
- But x86 extent code for multi-tenancy not fully ripened
- And hypervisors confront a two-front war…



A War on Two Fronts…

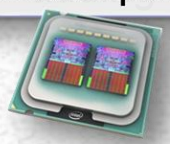
App software vendors don't write code to support VM hosting…

- Good hypervisors can do a great job with x86 extent code…
- The challenge to st
  - From abo
    making "i
  - From belo
    functiona
    vendors (t
    one) can co
    resource r

Virtual Machine Layer
- Application (and OS) containerized in Virtual Machine
- Issue resource calls to OS

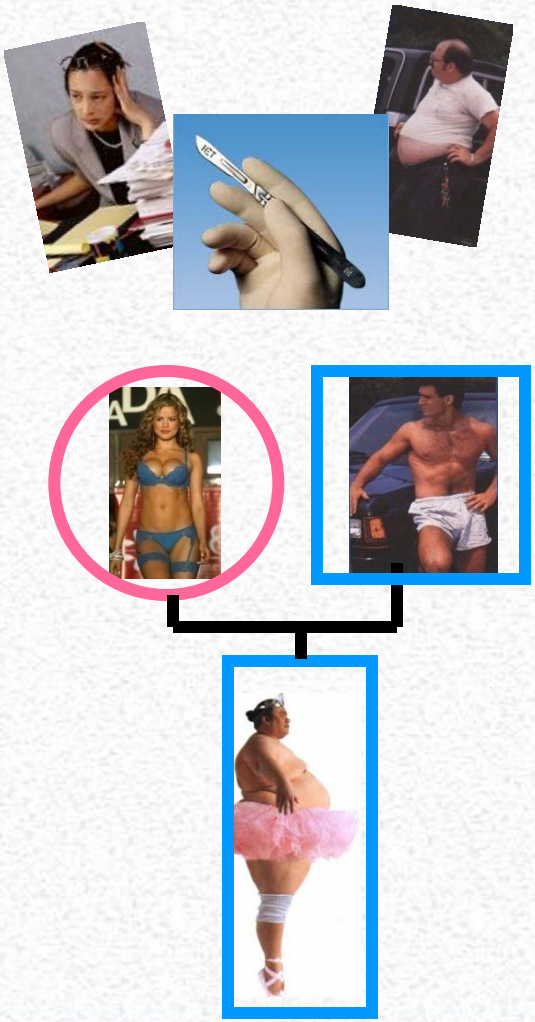
Hypervisor Layer
- Manager of VMs
- Intercepts resource calls and allocate resources based on requests

Operating System Layer
- Host to hypervisor server software
- Passes resources/services to applications per hypervisor-brokered requests



The Simple Truth

- Good stuff: x86 Hypervisor Hype has focused attention on
  - ✓ The commoditization of server, network and storage components
  - ✓ The need to consider resource allocation and utilization inefficiencies
  - ✓ The need to understand the relationship between data I/O and infrastructure
- But Don't Forget: Virtualization changes nothing, it just masks problems (and makes them more difficult to fix in too many cases)



Getting Back to Purpose-Built

1. Requires common sense review of the status quo…
2. A vision for how things need to be…
3. A plan for getting there…incrementally and in a business savvy way…

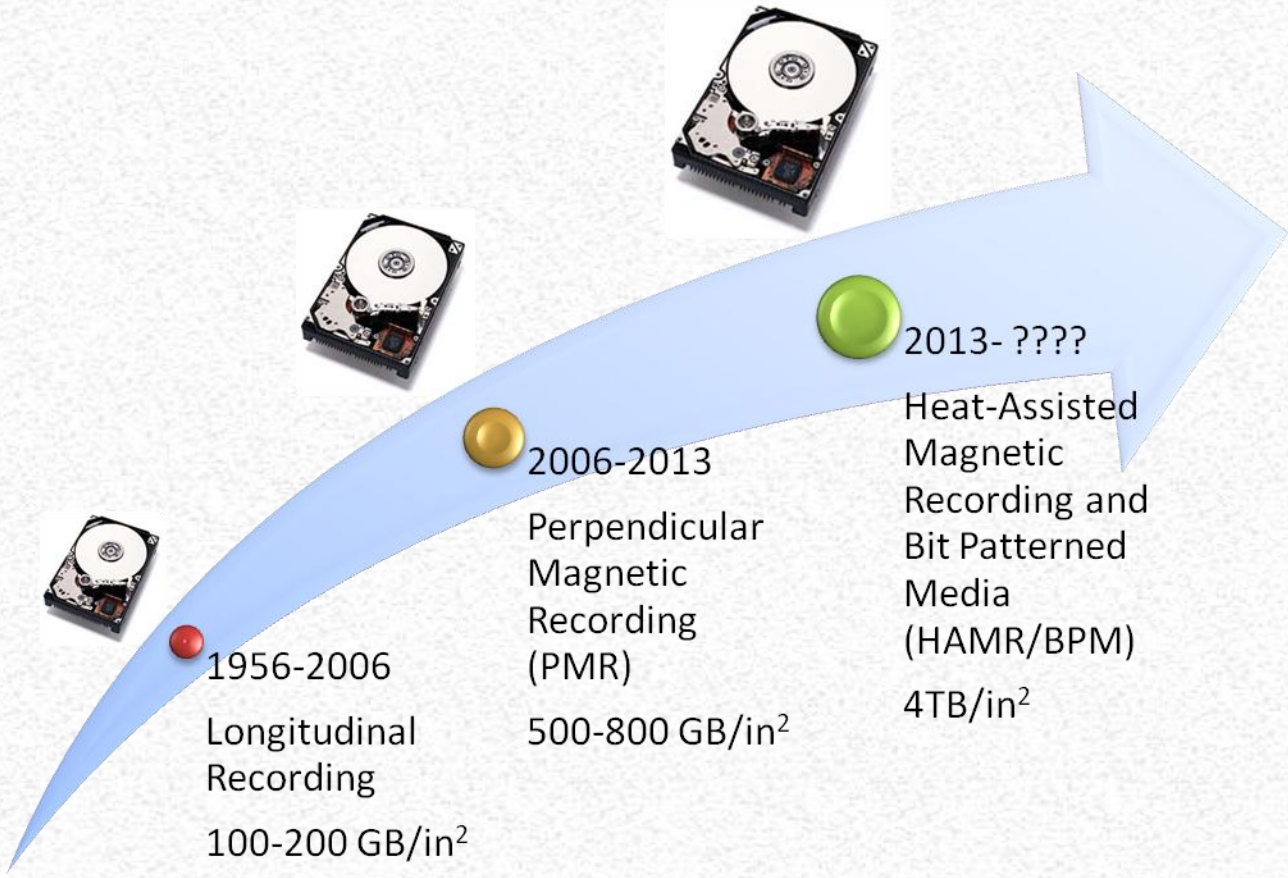
(A set of noise cancelling headphones is also a good investment.)

Storage is Key to Purpose-Built Design

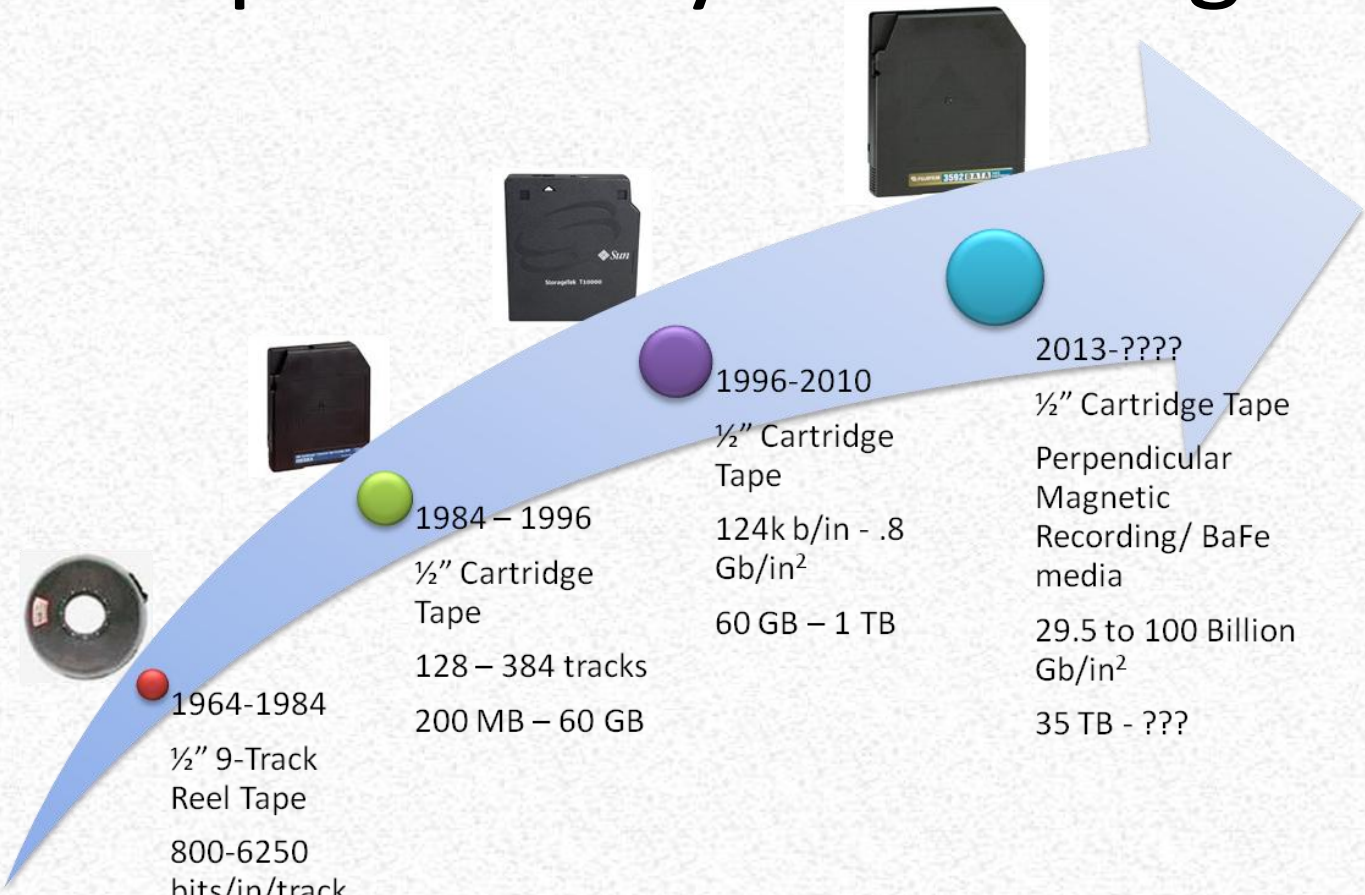
- 33 to 70 cents of every dollar spent on IT hardware today
- 300% capacity growth expected between 2007 and 2011
- Biggest power pig in the data center

- Where data [illegible] …*er, goes to sleep*

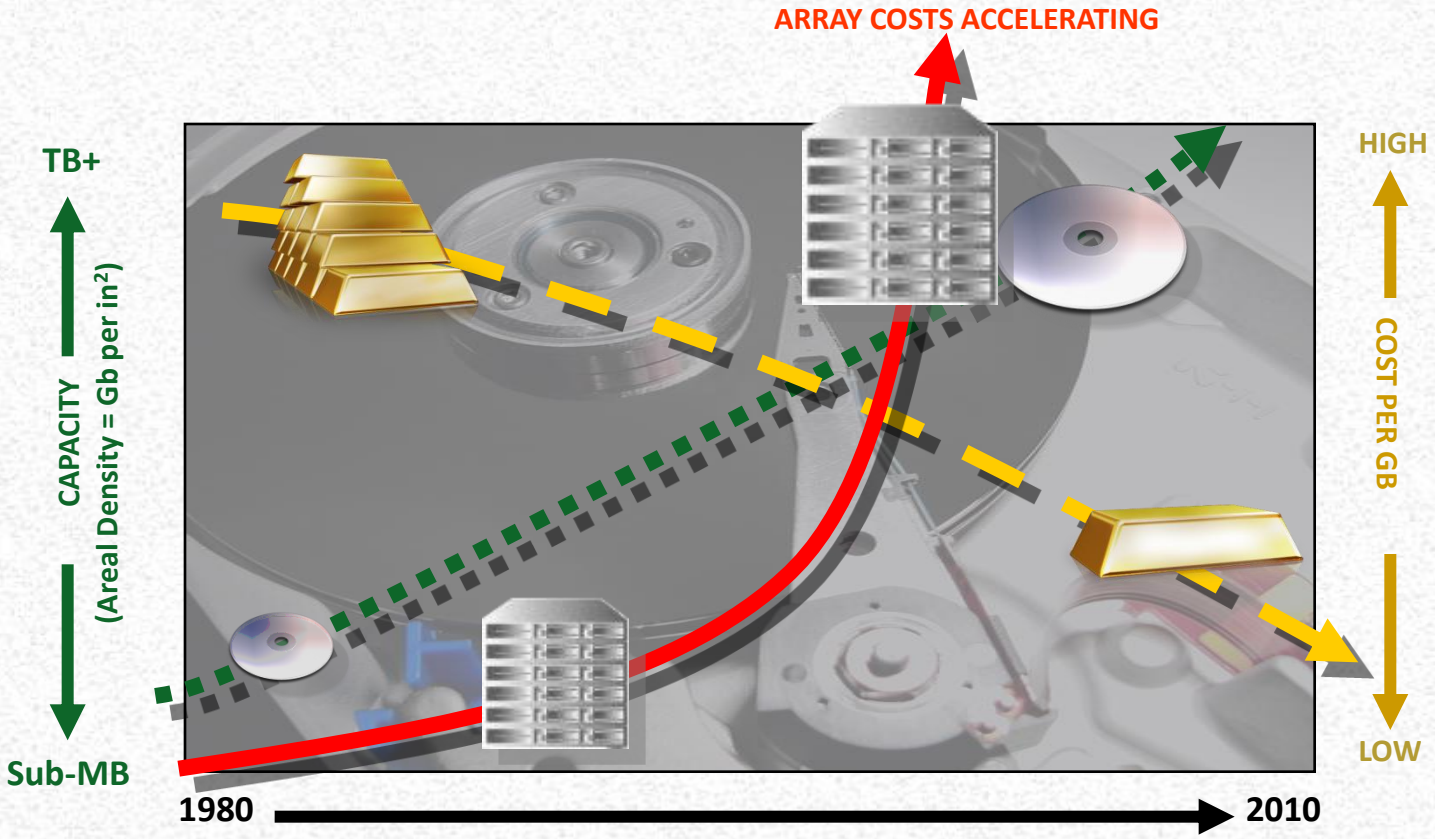
Good News:
Disk Areal Density Increasing



Also Good News: Tape Density Increasing

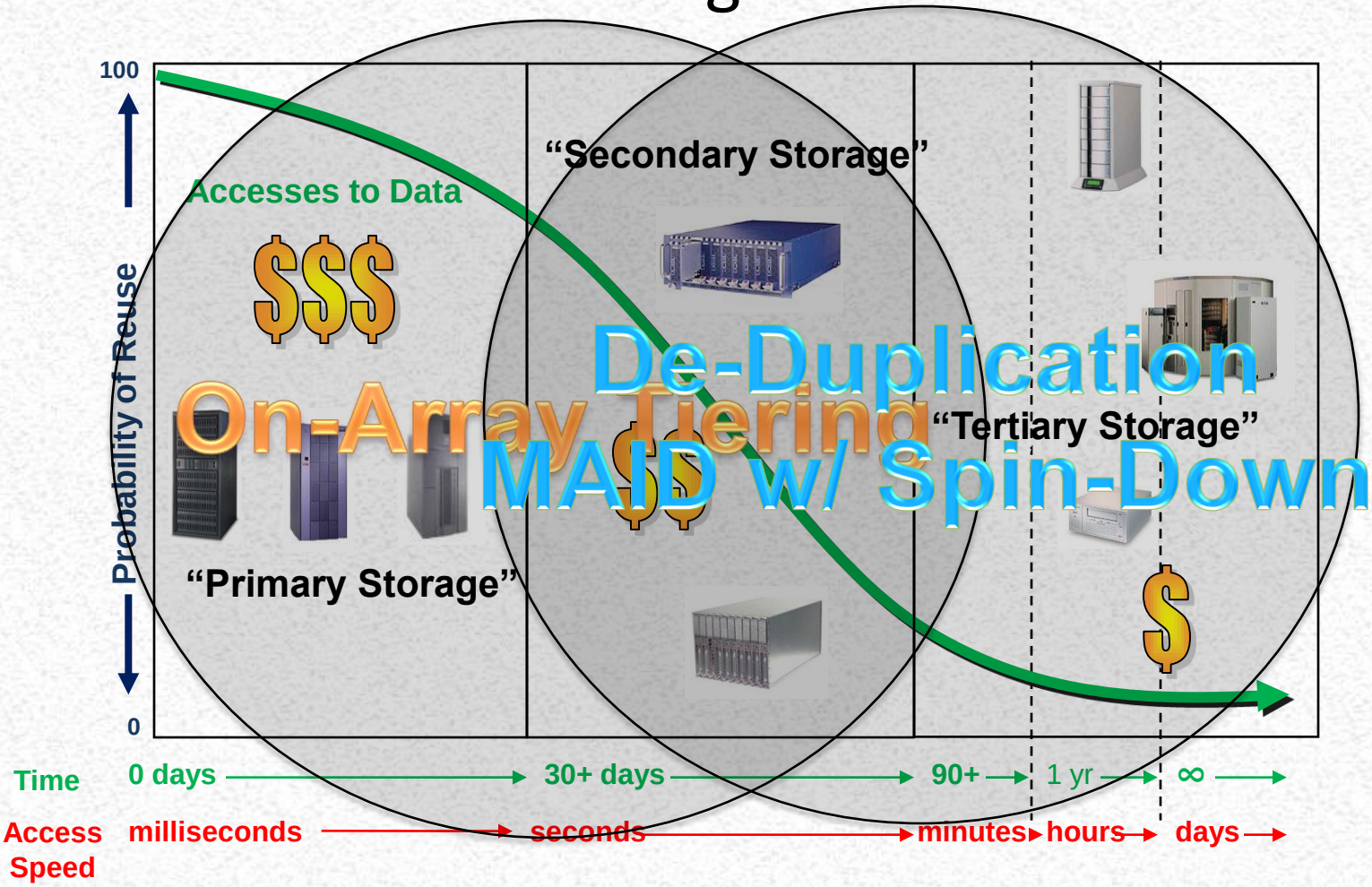


Should Bode Well For Containing IT Costs, But…



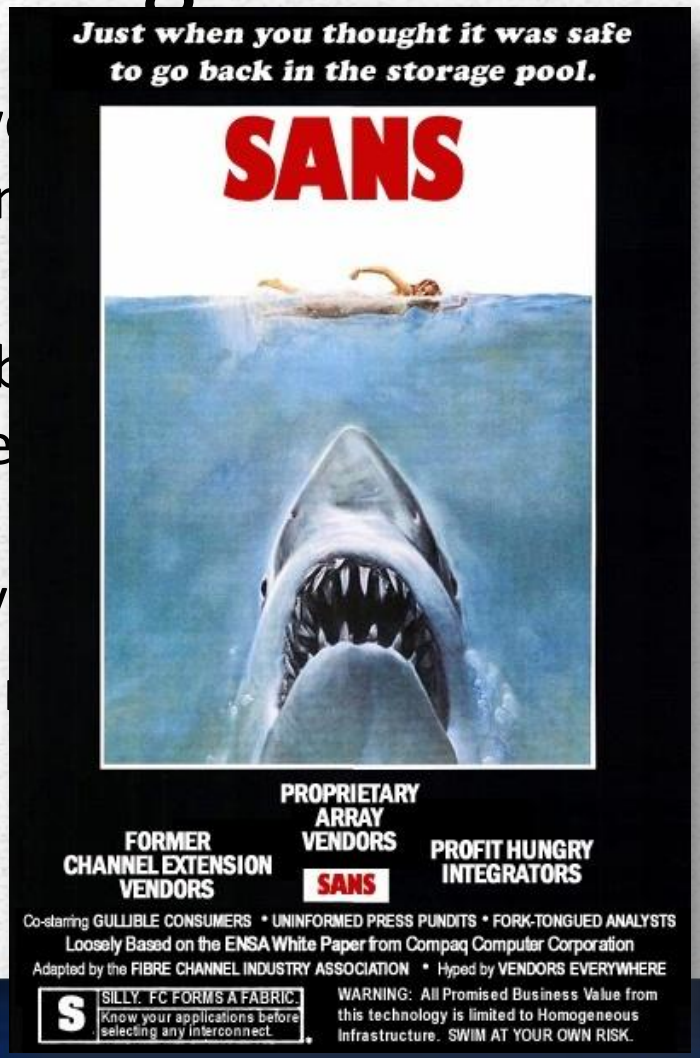
Source: "Avoiding the Storage Crunch," Jon Toigo, *Scientific American*

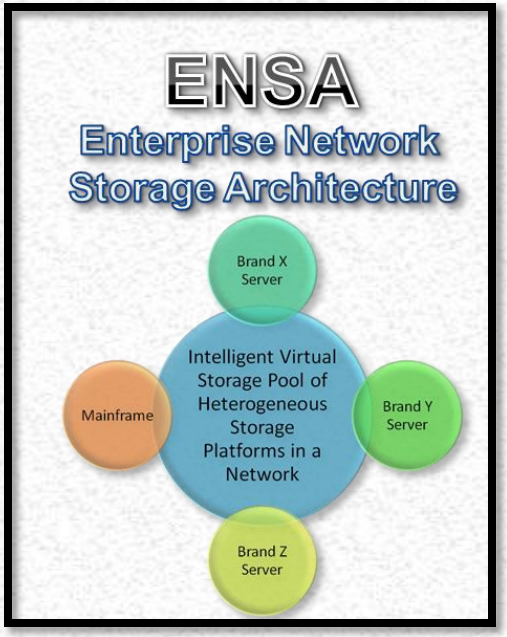
Data Re-Reference Rates Should Dictate Data Hosting Choices



Placed in a SAN, Now You're Talking Real Money

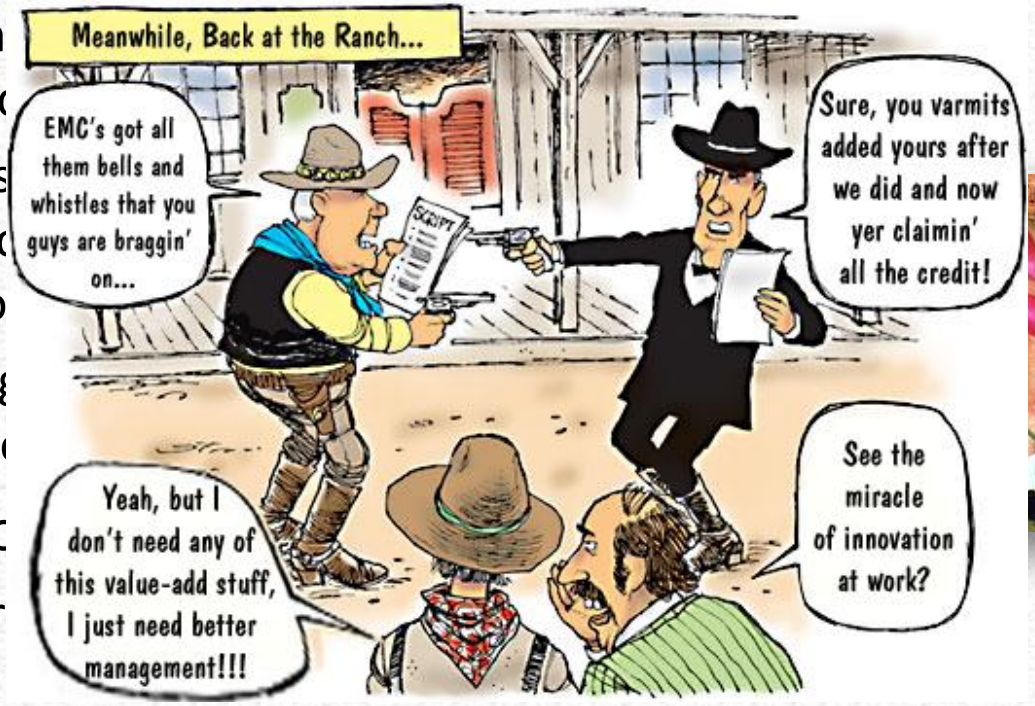
- SANs aren't netw
  been brought to r
- SAN switch ports
  underutilized (sub
  accelerating price
  high as $150/GB,
  $110/GB, tertiary
- At those prices, a
  growth…





Embedded “Value-Add” Software: A Good Idea?

- The array controller is getting bloated
  - Certain fun
    to the spin
  - Dynamic es
    “value” to
    look like yo
  - Vendors ar
    discussion
- Driving up co
  - You pay for
    or not
  - Management is obfuscated: need to hire more monks when you buy a new array



A Visionary C-4 Infrastructure Should Work Something Like This…

Data from this application…

…In a highly manageable way…

Storage Devices

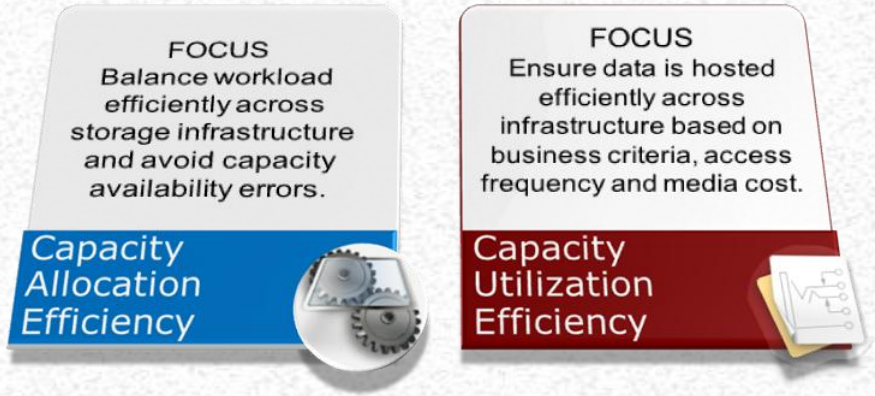
But Not Before We Tackle the Second Issue: Management

Realizing a Purpose-Built Infrastructure Requires…

- "Deconstructionalist" thinking
  - Willingness to let go of stovepipe vendors
  - Willingness to cobble together required services from best of breed suppliers rather than one-stop-shop providers
  - Desire to achieve best efficiency at lowest cost
- Embrace of an effective, preferably standards-based, approach to managing infrastructure
- Understanding of data itself and the services it requires for secure access, integrity, retention and protection

Compared to Mainframes, Open Systems Infrastructure Management…

- Er…is something out of the Stone Age
  - Mainframe storage allocation efficiency is 4x that of distributed storage allocation efficiency
  - Mainframe storage utilization efficiency is 3-4x that of distributed storage
- x86 Server Virtualization woo underscores this problem



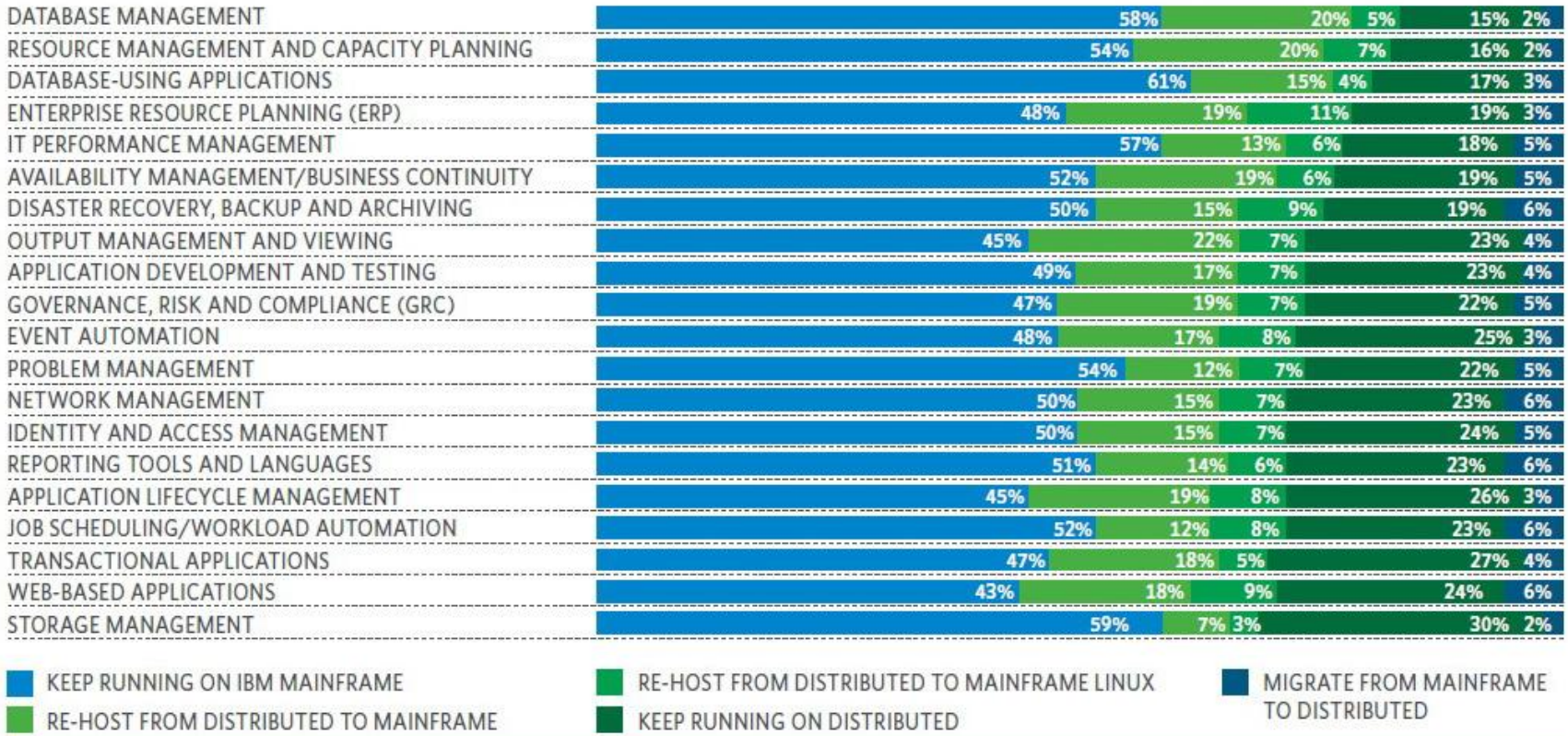
Mainframe Storage CAE: ~80%*
Distributed Storage CAE: ~20%

Mainframe Storage CUE: ~70%**
Distributed Storage CUE: ~17-20%***

*Assumes SMS
**Assumes HSM
***Depends on Server OS

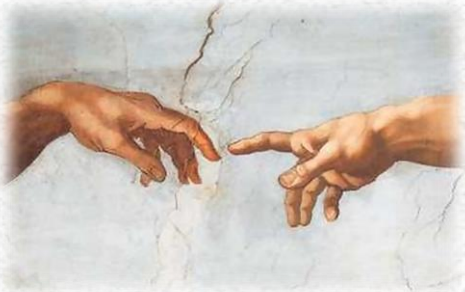
Mainframes and Mini-Me's

- CHART 8: MAINFRAME SOFTWARE SOLUTIONS HOSTING — DETAILS



IBM's Contribution Was a Blessing

- But mainframe management was also limited to one vendor's rig…
  - *IBM Cosmology*: Mainframe at the center of the IT universe
    - Leveraged *de facto* standards and well-defined architectural model
    - Homogeneity: all storage must conform to mainframe rules
    - Sets up a rigid workflow model
    - Doesn't work and play well with off-brand kit
  - *Actual goal:* sell more DASD, while maintaining sparse labor force



Open Systems Corollaries

- "Trust your vendors to solve your management problems."
- "Smarter storage arrays do it all."
  - Capacity management through embedded tiering algorithms
  - Capacity management through embedded thin provisioning algorithms
  - Capacity management through embedded compression and block de-duplication algorithms
- Hypervisors on the server allocate storage resources to apps-in-VMs more efficiently: the rise of "clouds"

We Have Tended to be Lackadaisical About Unified Management

- Primed to see things this way from early on…
  - Common view in small shops with only one storage platform from one vendor
  - Indicator lights, self-articulating web pages with status and configuration data, or a utility CD provides a solid "storage management" story
  - Meets the standard requirements of storage management
    - ✓ Provides configuration control
    - ✓ Monitors device status and capacity
    - ✓ Reports when problems occur so they can be addressed
    - ✓ Easy to use, ready out of the box: no muss, no fuss

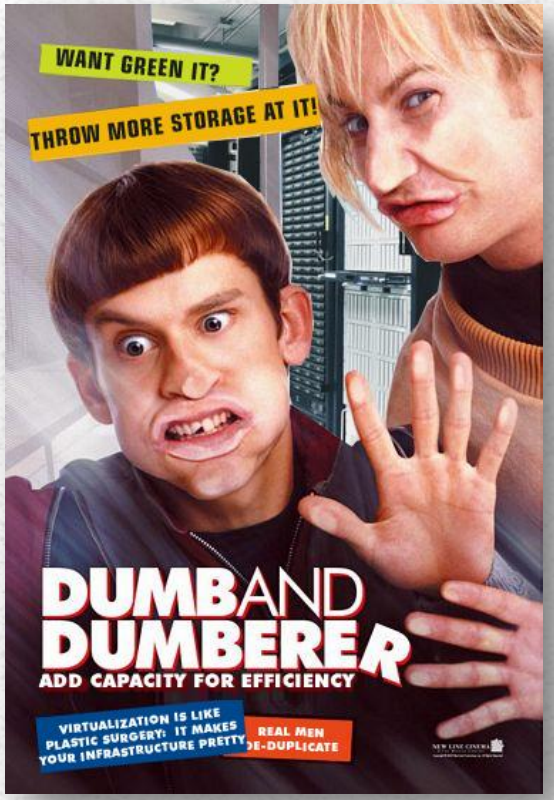
"See that green light? I know I've got a problem if it turns red. That's all the management I need."

But Element Managers Become a Problem When…

- Volume of data grows…
- The number of storage products increases…
- The storage becomes heterogeneous…
  - Products deployed from different vendors
  - Different products deployed from the same vendor
- Storage becomes “smart”…
  - Vendors build “spoofs” into gear to improve performance (example: NVRAM caching)
  - Array controller-embedded “value-add” software obfuscates real monitoring of capacity

Industry Response

- Make the Storage Array "Smarter" – More Value-Add Functionality
  - Automated Tiering
  - On-Array Thin Provisioning
  - Compression and De-duplication
- Storage management is really just about capacity management



“Tears of Storage”

- Quick definition: On-Array Storage Tiering is the movement of data between different types of disk within the same rig
  - Moves based on “watermarks” and FIFO, or date last accessed/date last modified attributes
  - Eliminating manual migration processes
  - All in one array: one throat to choke
- Purported value: simplification of capacity management, lower labor costs, less gear to manage, greater efficiency, elimination of tape from hardware mix

- Issues:
  - Non-granular data moves ignore business context of data: simplistic Hierarchical Storage Management
  - Increased price of all disks “enhanced with tiering controls” offsets purported time savings
  - Hard to manage stovepipe

On-Array Thin Provisioning

- Simple definition: Oversubscribe disk for greater allocation efficiency, reduced capacity management costs
- Done at a storage virtualization layer, maybe a valid strategy. But on an array controller? Hmm…
- Reality: A storage capacity "shell game"
  - No standards, only as effective as proprietary disk space demand forecasting algorithm
  - Data at risk of a "margin call"
  - Stovepipes are hard to manage

Thin is In, or Is It?

Survey of 249 Companies

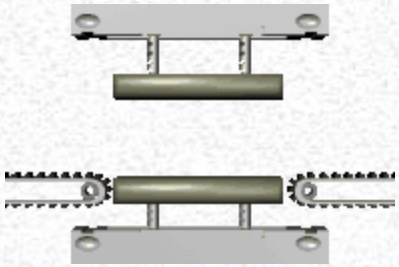
- 79% expect TP to deliver improved capacity allocation efficiency
- 49% expect easier provisioning with less disruption

Conversely,

- 44% concerned that TP will result in greater risk of running out of storage
- 43% worry that TP adds complexity
- 42% worry about a lack of capacity management tools

What About Compression and De-Dupe?

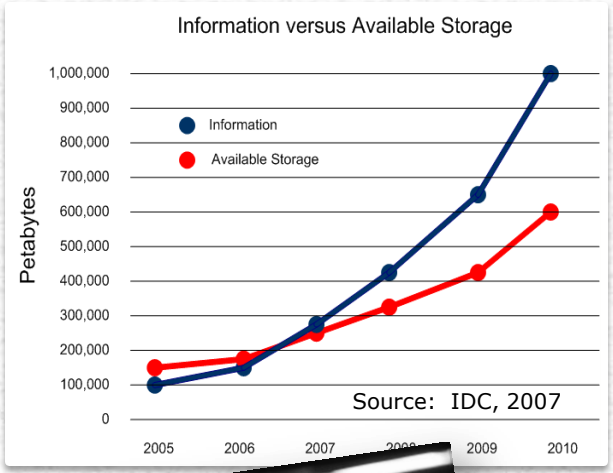
- Quick definition: reduction in the number of bits to represent data; can also mean:
  - Removing/stubbing bit patterns
  - Storing only change bytes or bits in versioning system
  - Elimination of file duplicates
- Ask a vendor about the differences and make some coffee…



- Purported Management value:
  - Squeeze more data onto media
  - Defer need to buy more capacity
- Issues:
  - Your mileage may vary
  - Possible compliance issues
  - More equipment to power
  - The Problem Dark Storage worsens…

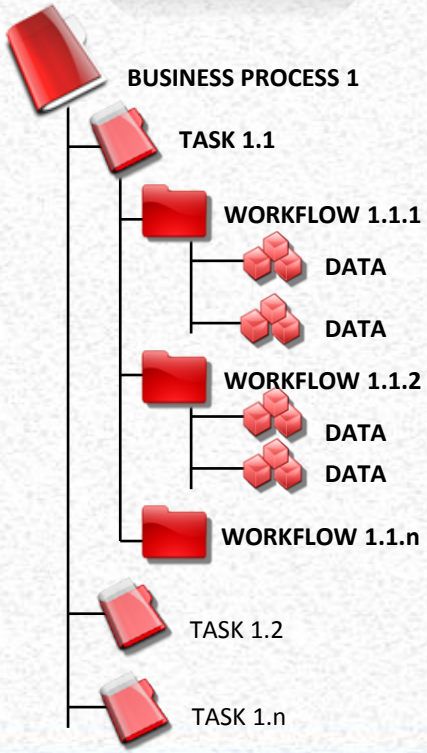
Infrastructure Management is More than Configuration and Capacity Management

- A view inherited from the mainframe world
- It is more than a knee-jerk response to the “Data Explosion”
- It is about aligning infrastructure to business requirements to provide
  - Status monitoring to facilitate planning and proactive troubleshooting
  - Modeling of infrastructure costs and performance to business goals to track ROI on changes



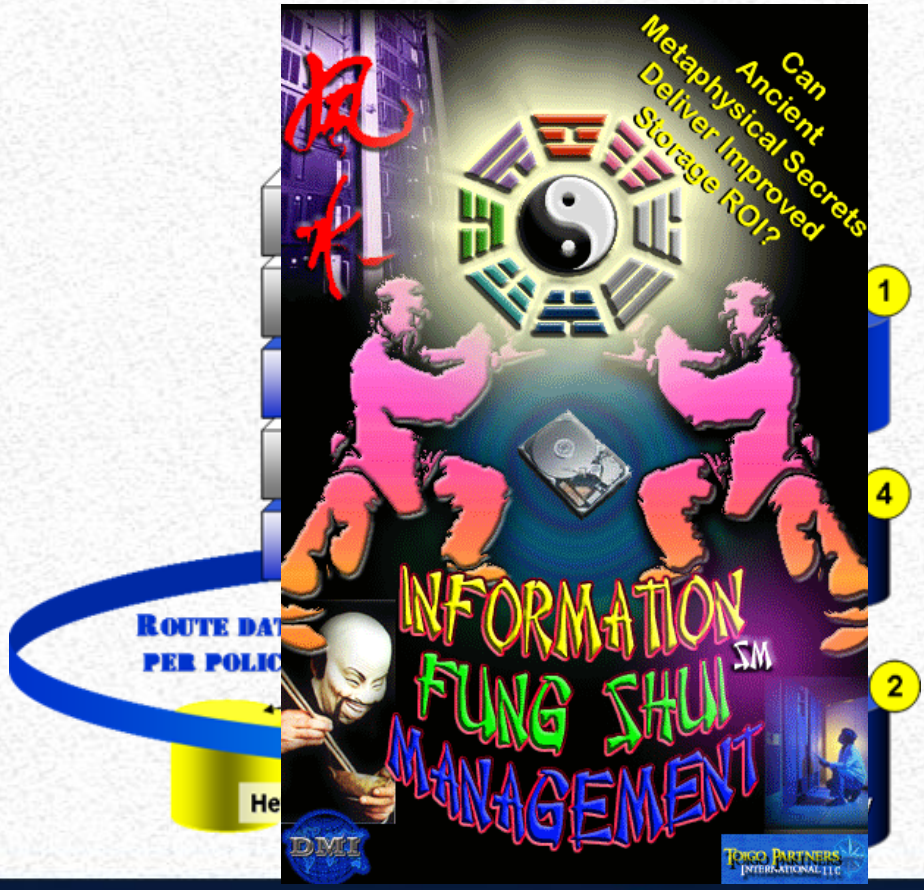
Ultimately, Infrastructure Management is Closely Tied to Data Management

- To manage infrastructure effectively, you need to understand the workload it supports
- You must understand data itself…
  - Business context of data (business value of data, regulatory requirements, criticality, etc.)
  - Volume and volatility characteristics
  - Access requirements (how many folks need access, what security must be provided, etc.)
  - Application-imposed parameters (ex 300GB SQL DB constraints in MS SharePoint)
  - Continuity requirements



No, I'm Not Pitching ILM…

1. A data classification scheme and a means to apply it (WHAT TO MOVE)
2. A storage classification scheme to identify target devices to host data based on the mix of services, performance and cost they deliver (WHERE TO MOVE IT TO)
3. An access frequency counter to help identify data usage patterns (WHEN TO MOVE IT)
4. A policy-driven data mover to move data from one place to another (HOW TO MOVE IT)

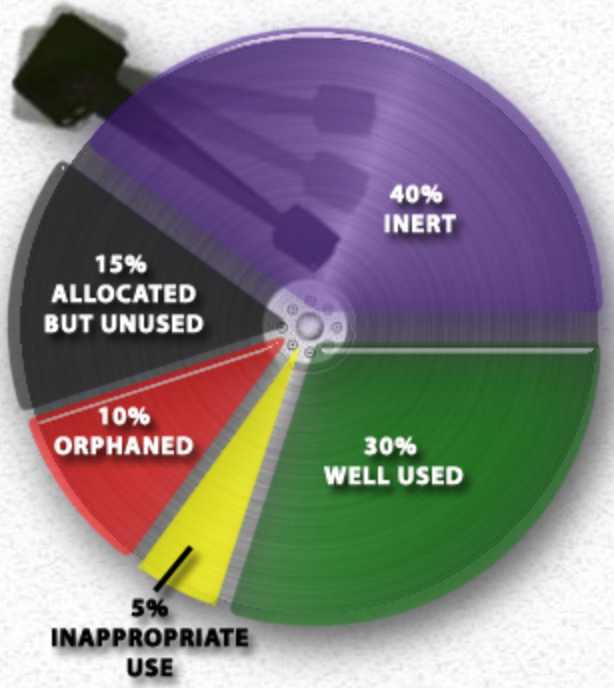


Lack of Data Management is Preventing Companies from Deriving Business Value from IT Investments!

- Think about it
  - Currently, the job of the IT administrator is largely confined to capacity management and break/fix
  - But, capacity allocation efficiency and uptime are not the sole determinants of infrastructure business value
  - Real cost-savings, risk reduction and process improvement are a function of resource utilization efficiency
- Reaching any sort of utilization efficiency requires data management
- Everybody knows this, but no one is saying it…

In the Absence of Data Management

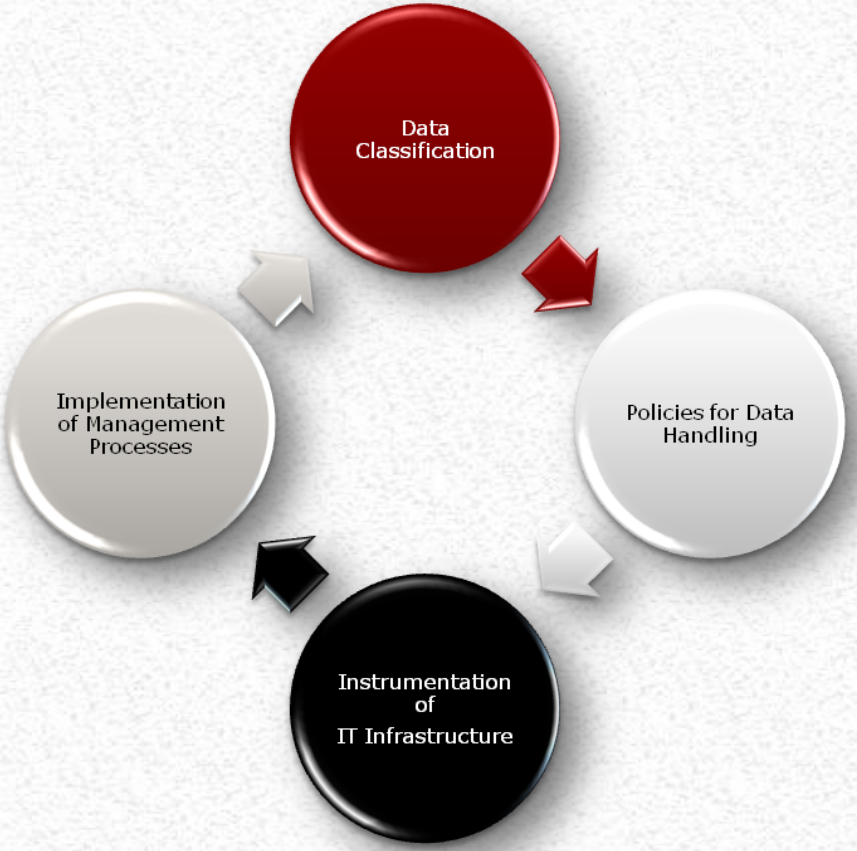
- Storage is a junk drawer – as much as 70% of capacity is wasted
- After normalizing over 10,000 storage assessments, we find
  - 30% of the data occupying spindles is useful
  - 40% is inert (needed for retention, but never accessed)
  - 15% of capacity is allocated but unused (often “dark storage”)
  - 10% is orphan data whose owner/server no longer exist
  - 5% (low estimate) is contraband



Source: *Making IT Matter* (Toigo, 2010)

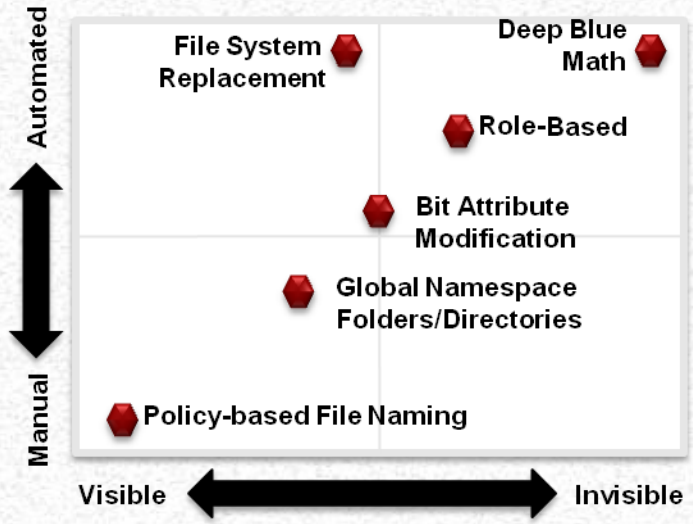
We Need to Start Sorting Out the Storage Junk Drawer

- Conceptually, a straightforward undertaking
  1. Classify your data
  2. Create policies for data handling by class
  3. Instrument your infrastructure for policy-based data movement
  4. Move your data around per policy
- Then, as in writing, just open a vein and bleed on the page…



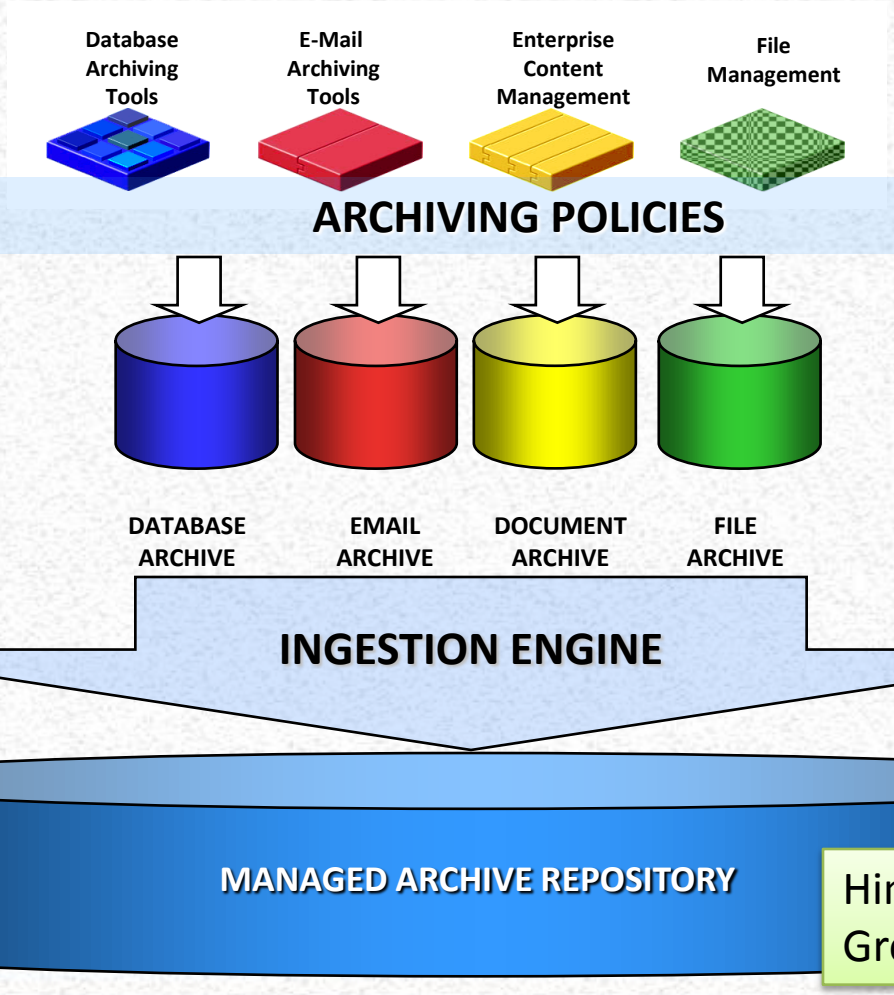
Classify Files at Creation: Many Approaches Exist

- But none are perfect.
  - Strive for a level of granularity you can live with (but note that simple metadata is rarely sufficient)
  - User invisibility is key – users generally bristle at having to participate in a classification-at-point-of-creation scheme
    - Most of the search-and-auto-classify tools we have tested produce many "false positives"
    - Some applications avail their data output with ready classification tools
  - In general, homegrown schemes are as effective as anything you can buy shrink-wrapped
- There are no silver bullets, no "ILM in a box"



User File Classification
"Visibility" to users is a gating factor for success

Build Your Own “Unified Model”



Hint: Tape is a good Green Archive Media.

Which Takes Us Back to the Beginning

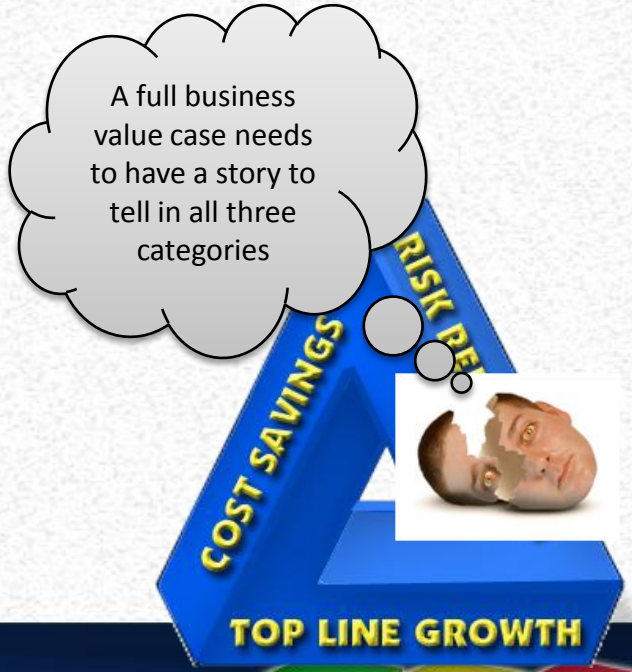
- The C-4 issues are real…
- Management sees them and wants a strategy to address them…
- Your job is to give them one.

Plus, They Want a Full Business Case to Support Your Proposal…

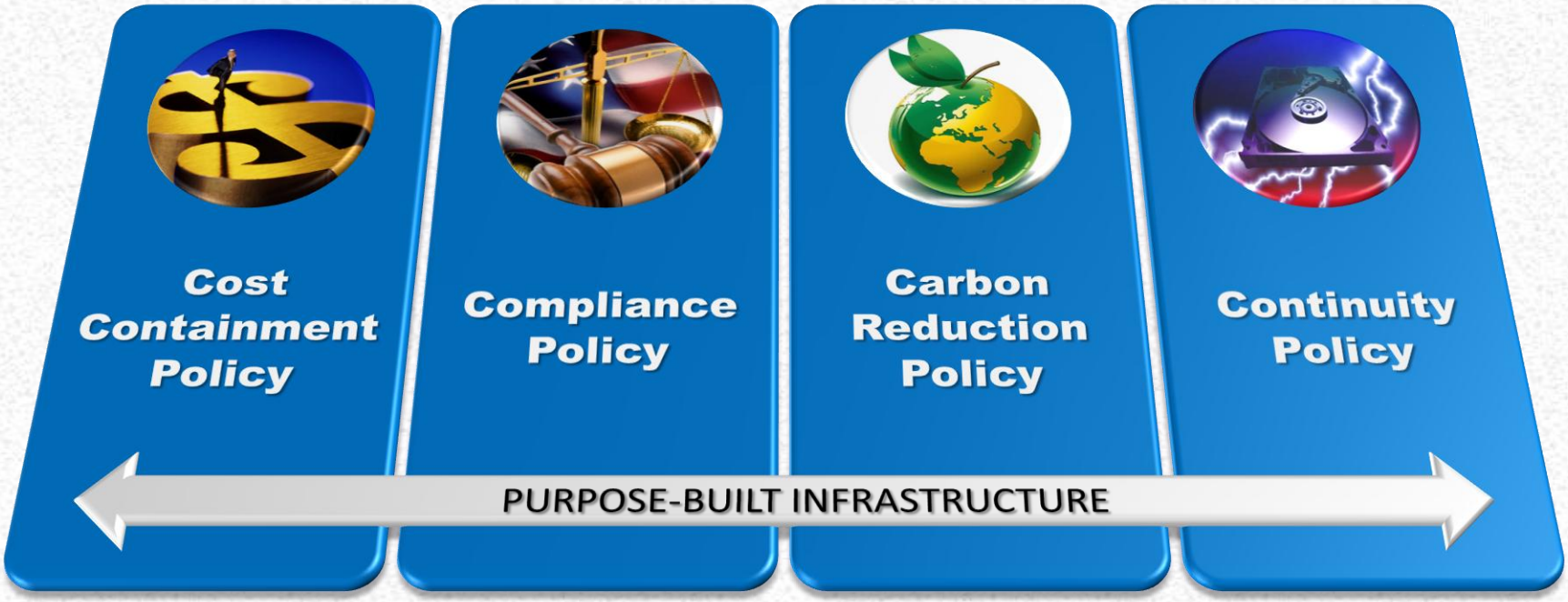
How will what you are proposing defer CAPEX, defer new investments, or trim OPEX?

How will what you are proposing protect existing investments, reduce downtime, mitigate operational or legal risk?

How will what you are proposing improve productivity, help expand into new markets, facilitate intelligent and timely decisions?



First Ingredient: Purpose-Built Infrastructure



- Map infrastructure to application and business process requirements
- Capture economic trends in technology

- Eliminate stovepipe technologies that impair visibility of data
- Leverage service based architecture that enhances process auditability and adapts to change

- Power only what equipment you need to do work
- Platform data according to access requirements

- Simplify recovery hosting platform requirements
- Enhance visibility and control of replication and failover processes

Second Ingredient:
Intelligent Infrastructure Management

- To identify resource allocation efficiency and opportunities for improvement
- To minimize OPEX and downtime costs
- To facilitate trend analysis for better CAPEX planning

- To monitor data asset service provisioning (security, DR, HSM, archive, etc.)
- To identify all infrastructure associated with a business process

- To identify power consumption trends in infrastructure and opportunities for consolidation
- To guide selection and deployment of greener technologies

- To identify burgeoning disaster potentials so they can be eliminated
- To augment testing capabilities
- To facilitate uptime and failover

Third Ingredient: Data Management

- To defer CAPEX spend
- To achieve hosting efficiency
- To protect IT investment

- To identify data assets for retention
- To apply appropriate privacy/integrity measures
- To discover appropriate data within timeframes

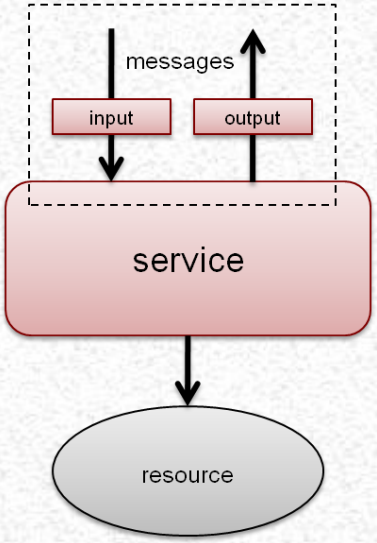
- To improve storage utilization efficiency
- To green IT infrastructure
- To slow demand for more energy

- To identify data assets requiring replication
- To right size data protection services
- To expedite recovery timeframes

Together, These Deliver the Full Business Value Case

My Preferred Paradigm: Business-Centric, Data-Focused, and **Standards-Based**

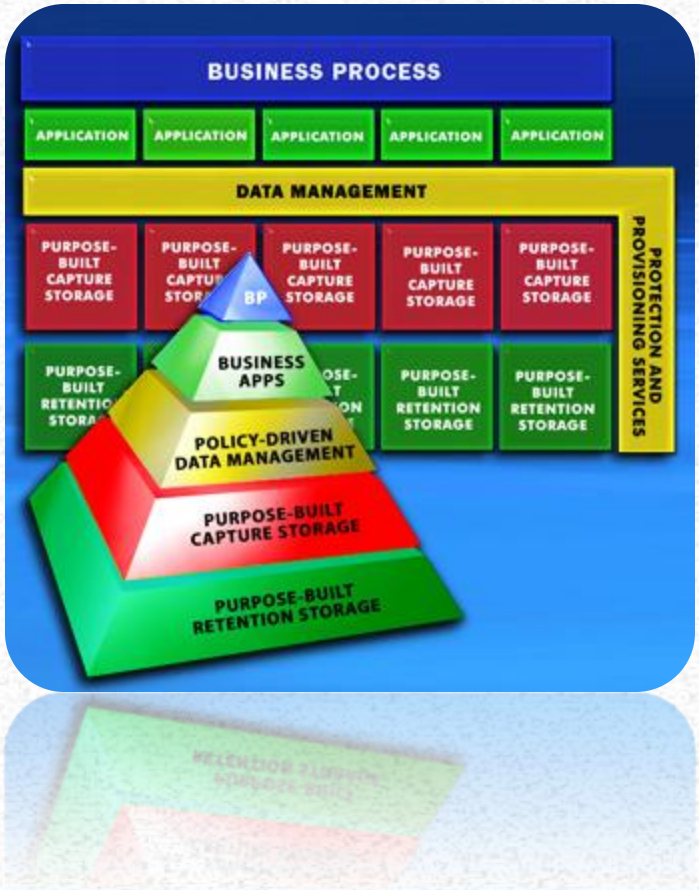
- Effective infrastructure management is required to get to the Holy Grail -- Data Lifecycle Management -- across purpose-built storage infrastructure
  - For real D/ILM to happen, storage classes need to be defined and operational status monitored continuously
  - Selecting or building storage using heterogeneous components and fielding services on networked platforms (aka appliances, routers, gateways, etc.) makes management a must-have
- W3C Web Services-based management support by gear vendors would make for an easily deployed and integrated **enterprise** management solution (covering servers, networks and storage)



SOAP/XML could provide the "universal glue" required to auto-provision and auto-protect applications with infrastructure-based resources and services.

Already supported by a broad range of applications, OS software, network hardware…and now storage.

C-4 Nirvana: Automated Data Service Provisioning



App Seeks to Write Data

Services Requested for Data Handling

Services Status Inventoried

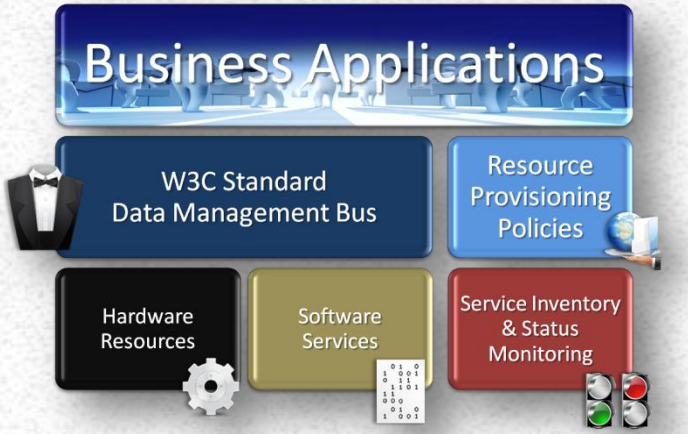
Provisioning and Protection Policies Applied

"Route" Established for Services

Requested Services Applied to Data

Going Forward

- Toigo Partners International and the Data Management Institute have launched a C-4 Community Portal to aid in W3C Web Services storage reference architecture design, development and education: You Are Invited!
- Working groups will span technical and nontechnical subjects and will combine consumers, vendors and reseller/integrator perspectives…
- Leading to the establishment of a C-4 Microfactory – a "farmer's market" web site where Web Services-enabled hardware and software products can be discovered and sourced for integrated solutions…



That's It: A C-4 Strategy In a Nutshell

- Questions? jtoigo@toigopartners.com
- Thanks! Keep me apprised of your progress and maybe we can publish your tips and tricks…

www.archivemgt.org
www.enterprisemgt.org
www.drplanning.org
www.compliancemgt.org
www.it-sense.org
www.greendataproject.org
www.storagemgt.org
www.securitymgt.org

For a Deeper Dive…
Join Us at the C-4 Summit in Cyberspace

C-4 Summit

Cost Containment • Compliance • Continuity • Carbon Footprint

WWW.C4PROJECT.ORG