HubStor™

The Financial Business Case for Cloud Archiving

Executive Summary

Offloading inactive data to a cloud archive can greatly reduce in-house storage costs while providing other advantages. But storage economics in the cloud are quite different than in-house.

In this paper we'll compare consumption-based pricing of a cloud archive, based on public cloud economics, to the storage economics of a typical in-house storage array.

Storage is cheap, right? Not in the enterprise. Did you know many businesses pay **$1 million per 100TB** as a fully-burdened cost for in-house storage? [1]

In fact, dealing with data growth is the **#1 storage project** in 49% of organizations.[2] Maybe that's because 56% of organizations are saying **50% of data on primary storage is inactive**.[3]
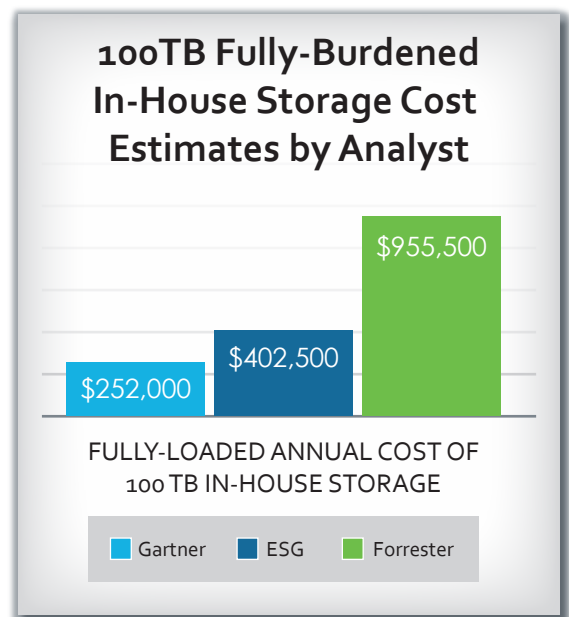
Archiving to cloud storage unlocks a smarter - and more budget friendly - storage strategy. Let's quantify it.

Wide Ranging Analyst Estimates on In-House Storage Costs

What is your fully-loaded annual cost per Terabyte of storage?

According to Gartner, "*the annual cost per raw TB of storage in 2014 was $2,520*"[4]. But that's raw storage (before things like RAID and vendor hold back, etc.) - not your actual usable capacity. Figures from ESG Global peg the annual cost per *usable* Terabyte at $4,025 [adjusting their 3 year cost of $603,778 for 50TB][5] Forrester Research Inc. places the fully-loaded annual cost per *usable* Terabyte at $9,555 [adjusting their annual cost of $955,500 for 100TB][6] Each analyst firm's estimate claims to account for hardware, software, personnel, facilities, power/cooling, and disaster recovery. Quoting on the raw versus usable storage explains some variance in these estimates, but certainly not to this degree.

$1M annually for every 100TB of storage seems to be an overestimate. But in our discussions with storage professionals at Fortune 1000 companies the $1M figure has been echoed to us again and again.

100TB Fully-Burdened In-House Storage Cost Estimates by Analyst

$252,000 — Gartner
$402,500 — ESG
$955,500 — Forrester

FULLY-LOADED ANNUAL COST OF 100 TB IN-HOUSE STORAGE

Unfortunately for some businesses it's even higher. Given the wide ranging cost figures from analysts we'll explore our own costing. But first let's understand what factors go into these models.

Cost Factors: In-House vs Cloud Storage

In-house storage typically involves large upfront expenditures. Cloud storage is usually pay-as-you-go. Here's a closer look at the cost factors and how they differ between cloud and in-house storage models:

	In-House Storage	Cloud Storage
Budget Impact	You're buying lots of equipment which is a Capital Expenditure (CAPEX).	It's a subscription to a service so you get to fund it with Operating Expenses (OPEX).
Capacity Management	You're always paying for more than you use. Consider overallocated and unused space, plus capacity for future growth.	You pay only for what you exactly use, and you don't worry about proactively buying additional buffer storage. It scales up as you need it.
Storage Overheads	Because of things like RAID, vendor hold backs, and how the OS measures capacity your usable storage is always less than what you purchased.	No storage overheads to worry about with cloud storage. Again, you pay for precisely what you use.
Redundancy Copies	You need to procure additional storage for local redundancy and offsite copies too. This may have additional costs such as backup software and 3rd service providers.	By default cloud storage makes 3 copies of your data – and you don't have to think about it. You can upgrade to have redundancy to a secondary datacenter.
Data Transfer	Continuously replicating to disk at a secondary facility requires bandwidth but otherwise bandwidth costs are not a factor with in-house storage.	Cloud storage models typically have a fee associated with data transfer out (retrieval from the cloud), and some may also have a data transfer fee on the way in.
Storage Transactions	Since you own the infrastructure there is no tangible cost associated with basic use of your storage.	The cloud storage provider will typically have a cost associated with storage transactions (each write, read, or modify touch to your data).
Facilities, Power, Cooling	Often overlooked or obscured, floor space and energy have real costs associated with in-house storage.	The cloud storage vendor bares this burden so you don't need to worry about it.
Vendor Support & Maintenance	Vendors often bundle in 3 years of maintenance and support, but after 3 years expect to pay 15% of the original purchase price annually.	Often bundled in to the subscription fee for the life of the subscription, but will vary depending on the vendor.
Data Migration	Storage equipment eventually ages, requiring data to be migrated and migration can be costly and painful.	The cloud storage vendor bares the burdened of refreshing hardware, with no migration costs or disruption to you.

So we see with in-house storage you have to buy, run, power, protect, maintain, and refresh your storage. It's why an organization's fully-burdened storage cost is substantially more than the upfront price tag you see from the storage vendor. As we'll see, cloud storage is a much simpler model.

Other Cost Factors

Not typically discussed in storage cost models – but indeed a reality for most enterprises and often owned by the IT team – are costs to do with the following:

Data Management Software & Associated Infrastructure	Initiatives in the areas of Sensitive Data Classification, File Analysis, Data Access Governance, and/or Chargeback/Showback Reporting will drive up storage costs with software licensing, machines, storage, professional services, and time. In some cases these costs might be more attributed to the security team's budget, but not always. Often before a migration, storage teams will clean house with such tools. According to Gartner the product costs alone for such initiatives can be "***around $1,000 per TB for metadata analysis only***"[7]. It's a significant added cost in the storage budget and it isn't factored in the referenced analyst cost estimates.
Content Indexing & Associated Infrastructure	Full-text indexing for content analysis and keyword search is another factor not in the analyst cost models. Often deployed for user productivity and/or this is also being used in storage management initiatives. It's another cost that's significant but often overlooked. For such initiatives Gartner says product costs from "***around $5,000 per TB or more for content analysis, including full text indexing***"[8], but don't forget that content indexing initiatives require large investments in compute and storage, as well as time and expertise.

Our Cost Model Example for In-House Storage

In our model we dropped personnel costs from the equation since these vary widely. And we added costs for data and storage management projects. That's because we believe projects relating to private/sensitive data tagging, chargeback/show-back reporting, full content indexing, access governance, etc. are very real and – in our view – should factor into the fully-burdened costs. **With these adjustments our model pegs the annual cost of each usable in-house Terabyte at $3,984.** Here's a breakdown of our estimates:

> *"One of the biggest differences between cloud and traditional storage[:] When you send 100 TB of data to the cloud, you pay for 100 TB. However, when you store 100 TB of data internally, you have to buy much more than that amount."*
>
> Andrew Reichman

Example 4 Year Cost Model for 100 TB In-House Storage			
	Amortized Monthly	**4 Yr Costs**	
Base Cost of Primary Storage	$7,813	$375,000	We estimate $3.75/GB per *usable* gigabyte of storage.
Redundancy & Backup	$19,531	$937,500	Conservatively, 2.5x the base storage costs for high availability copies and de-duped backup storage and infrastructure.
Facilities, Power, & Cooling	$1,563	$75,000	We apply Forrester's figure of 5% annually of the cost of the primary storage.
Vendor Support & Maintenance	$1,172	$56,250	Assumes the first 3 years are included but factors a 4th year subscription at 15%.
Data Migration	$1,042	$50,000	Assumes $500/TB cost amortized over 4 years.
Data Management	$2,083	$100,000	We conservatively use Gartner's product costs figure of $1,000/TB for metadata analysis of the full 100TB. *
Content Indexing	$3,646	$175,000	We assume 25% of the 100TB is to be full-text indexed. Our cost estimate conservatively adjusts Gartner's figure of $5,000/TB for product costs to $7,000/TB to rougly factor costs for indexing machines and storage and vendor services and support.
Total w/o Content Indexing	**$33,203**	**$1,593,750**	* We say this is conservative because it is cited strictly for "product costs" but we use it to cover licensing, services, machines, and storage costs across a variety of initiatives (tagging, analysis, reporting, retention, etc.) which may require multiple products in practice.
Total w/ Content Indexing	**$36,849**	**$1,768,750**	

Our Example Cost Model for Cloud Archiving with HubStor

The cost model for cloud archiving is drastically different. With HubStor – a software-as-a-service – it's a single line item for the monthly pay-as-you-go subscription. This includes the underlying cloud infrastructure needed to preserve, protect, manage, and access the 100TB. It also includes the cloud storage, redundancy copies, virtual machines, databases, vendor support and maintenance, and access to HubStor's on-premises cloud connector software.

For a detailed description of HubStor's pricing model mechanics see *Appendix A – How HubStor Pricing Works*.

Example 4 Year Cost Model for 100 TB Cloud Storage with HubStor

	Amortized Monthly	4 Yr Costs	
Cloud Storage	$8,219*	$394,512	An example HubStor monthly subscription fee for 100TB consumption, also covering the other "Included" items below.
Redundancy & Backup	Included*	Included	3 synchronous copies are made in the cloud by default.
User Access Integration	Included	Included	HubStor includes end-user access via a Web browser to the cloud as well as optional 'cloud stubs' in on-premises storage.
Vendor Support & Maintenance	Included	Included	Technical support with maintenance and upgrade services included.
Data Transfer	Included*	Included	We assume monthly retrieval of 1% (or 10TB) for data transfer costs, which are rolled into the monthly subscription fee.
Storage Transactions	Included*	Included	We assume 10 million storage transactions per month, which costs are rolled into the monthly subscription fee.
Cloud Storage Gateway (CSG)	$0	$0	No CSG. The HubStor subscription includes access to HubStor's on-premises connector service and suite of connectors.
Data Management	Included	Included	HubStor is a data-aware cloud archive with a built-in policy engine supporting retention, auditing, eDiscovery, and PII.
Content Indexing	$6,350*	$304,800	We've calculated 25% of the 100TB is indexed for the entire 4 years. This cost includes search cluster VMs for indexing and query as well as redundant index storage.
Total w/o Content Indexing	**$8,219**	**$394,512**	
Total w/ Content Indexing	**$14,569**	**$699,312**	

* This is example pricing only. HubStor pricing is based on a volume discounted markup schedule applied to the actual public cloud costs based on actual consumption. Our model makes estimates on consumption levels and configuration. Your actual production usage may yield higher or lower pricing. See Appendix A for a more detailed description.
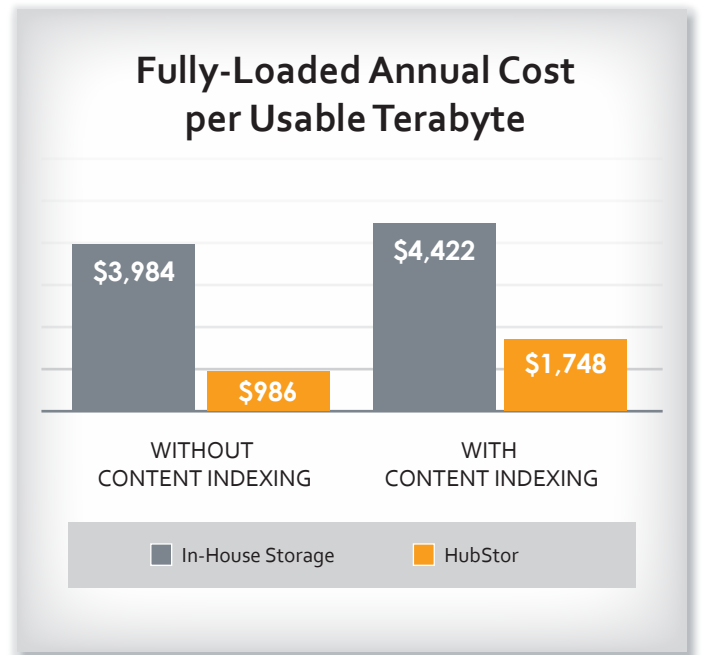
Normalizing the Models

It's easy to see cloud is a significantly more suitable place to keep inactive and low-touch data than in-house primary storage. Cloud also provides a list of business agility advantages.

Let's now look at a side-by-side comparison of the two models.

Fully-Loaded Monthly Cost / GB

IN-HOUSE	$0.32
HUBSTOR	$0.08

WITHOUT CONTENT INDEXING
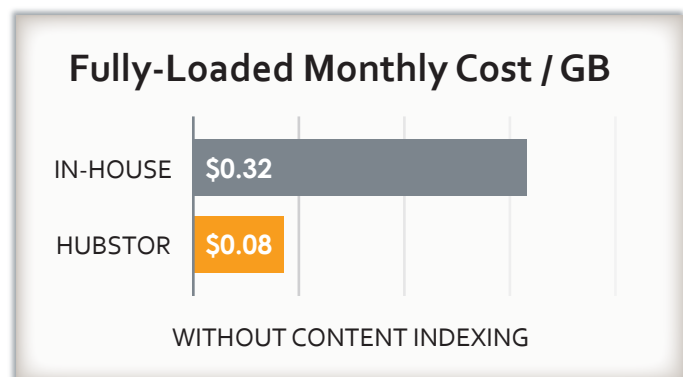
Cost Models - Monthly Comparison

100TB All-In	In-House	HubStor
Primary Storage	$7,813	N/A
Redundancy & Backup	$19,531	Included
Facilities, Power, & Cooling	$1,563	N/A
Vendor Support & Maintenance	$1,172	Included
Data Migration	$1,042	N/A
Cloud Storage	N/A	$8,219
User Access Integration	N/A	Included
Data Transfer	N/A	Included
Storage Transactions	N/A	Included
Cloud Storage Gateway (CSG)	N/A	$0
Data Management	$2,083	Included
Content Indexing	$3,646	$6,350
Total w/o Content Indexing	**$33,203**	**$8,219**
Total w/ Content Indexing	**$36,849**	**$14,569**

Cloud Consumption Estimate

Price estimates for HubStor's pay-as-you-go monthly subscription model are available upon request.

Fully-Loaded Annual Cost per Usable Terabyte

Without Content Indexing: In-House Storage $3,984, HubStor $986
With Content Indexing: In-House Storage $4,422, HubStor $1,748

In-House Storage ■ HubStor ■

Considerations

☁ **Measuring Fully-Burdened Costs** – Fully-burdened costs aren't something you think about in the cloud. But they are much more a factor in the in-house storage model. It's not only storage and backup but also floor space in the datacenter, staff to support, server infrastructure, electrical power and cooling, and a complete online duplicate infrastructure to support high availability. The all-in storage costs for your company may be drastically different than what we modeled or what analysts suggest. Forrester Research, Inc. provides a cost analysis tool which can help you assess your fully-burdened storage costs.[8] While the research is dated by a few years we believe the fundamentals for comparison remain unchanged.

☁ **Public Cloud Economics** – HubStor's model of passing through public cloud costs with a volume-discounted markup based on your consumption offers significant cost advantage that is impossible to match in-house, and a huge incentive to offload inactive/low-touch data from in-house primary storage.

☁ **Typical Cloud Storage Adoption Costs** – The HubStor software makes object storage in the cloud easily consumable, removing typical cloud storage cost variables such as development and integration costs for end-user access and procurement and deployment of a cloud storage gateway for extending existing infrastructure to the cloud.

☁ **Better Together** – HubStor is not a replacement storage solution for workloads requiring low-latency, high performance touch. However, HubStor is clearly an ideal storage tier for aged data. By offloading aged data you can see better performance from your disk arrays and reduce capacity demand on expensive in-house storage. HubStor's connector service provides policy mechanisms so that you can determine when (and what) data should be trickled up to the cloud.

☁ **Apples to Oranges** – For archive data we can see that HubStor provides significantly better economics than traditional in-house storage. Even when we compare HubStor with content indexing to in-house storage without content indexing we can see that HubStor is 40% the cost. This means you can reduce storage costs while gaining more functionality and control.

☁ **Apples to Apples on Content Indexing** – In our 4 year model for in-house storage we see amortized content indexing costs at $145/indexed TB monthly. We assume the in-house search infrastructure (machines and storage) isn't backed up, otherwise in-house search costs would be significantly higher. By comparison, content indexing in HubStor is more expensive, with a monthly cost per TB indexed at $254. Given the amortization, HubStor's pay-as-you-go search model is significantly less expensive in shorter time periods than implementing licensed search software in-house with its required compute and storage. HubStor also provides ways to scale down search during periods of lower consumption demand to reduce content indexing costs. In our model we made no consideration for scaling down HubStor's search during the 4 years.

☁ **Cloud Storage Gateway and HubStor** – With its on-premises connector service and suite of connectors, HubStor may negate the need for a CSG for the purpose of cloud-extending current storage infrastructure. But HubStor is not a CSG. CSGs commonly provide a global namespace and perform on-premises caching for high touch data. However, the two are not mutually exclusive. Keeping in mind that vendor integration would be necessary, a CSG could make a great front-end to HubStor's data-aware cloud archive.

HubStor Pricing Disclaimer

The pricing figures in this document are for example only. As described in this document, our model makes estimates on consumption levels and configuration. Your actual production usage may yield higher or lower costs per Terabyte in HubStor. See Appendix A for a more detailed explanation of HubStor pricing.

Appendix A - How HubStor Pricing Works

Overview

HubStor is a *monthly pay-as-you-go subscription* with no term commitments and no startup or cancellation fees. The subscription covers maintenance, support, and full-featured access to both HubStor's on-premises software and data-aware cloud archive.

Monthly subscription pricing is determined off the actual metered usage of public cloud infrastructure.
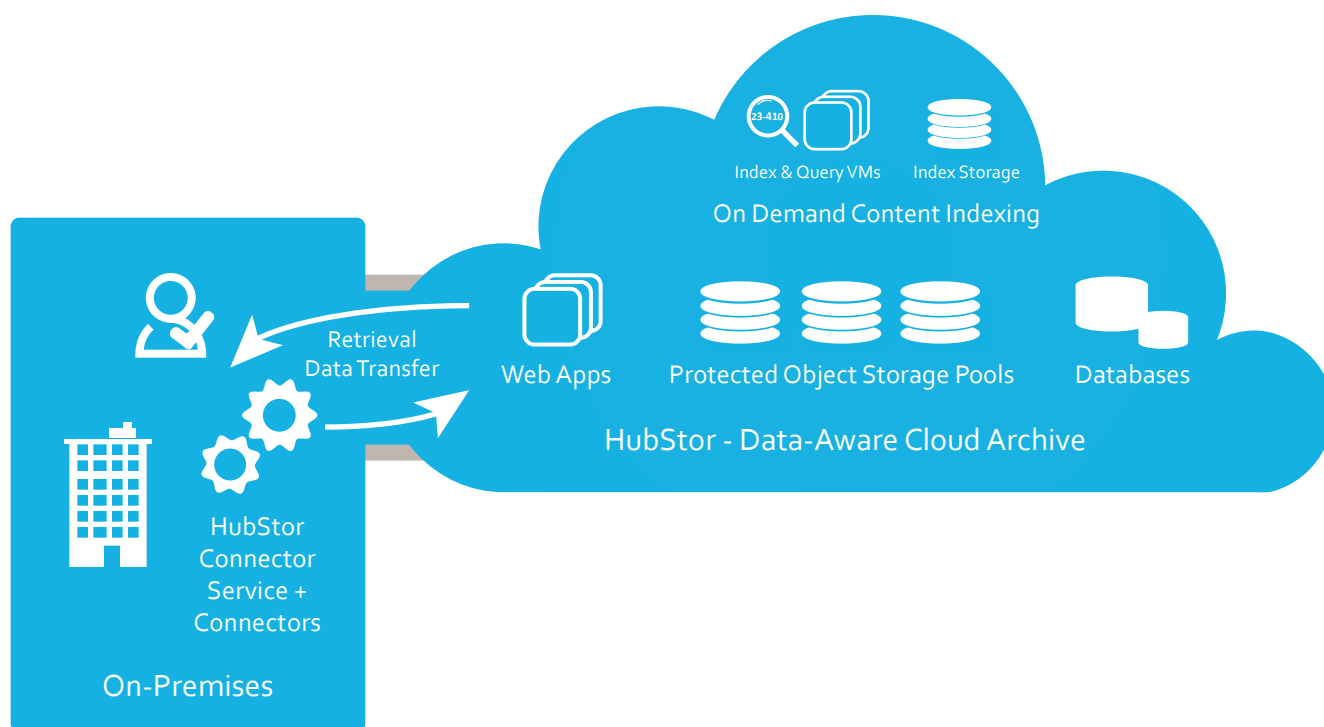
What is consumption?

When you subscribe to HubStor an archive environment is created for you in the cloud. This includes HubStor software deployed and configured with storage, redundancy, databases, Web apps, and an optional search cluster.

Then as you begin using HubStor - storing data in the cloud, accessing content, optionally using content indexing - your usage of the service incurs infrastructure costs in the cloud, which we call '*consumption*'.

Pay-as-you-go consumption pricing

HubStor pricing is determined by applying a volume discounted markup schedule to your public cloud costs. It's that simple.

Public Cloud Costs

+

HubStor Markup

=

Monthly
Subscription
Amount

Index & Query VMs   Index Storage

On Demand Content Indexing

Retrieval
Data Transfer   Web Apps   Protected Object Storage Pools   Databases

HubStor - Data-Aware Cloud Archive

HubStor
Connector
Service +
Connectors

On-Premises

ENDNOTES

1 http://media.amazonwebservices.com/Forrester_File_Storage_Costs_Less_In_The_Cloud.pdf

2 http://www.computerweekly.com/feature/Survey-Storage-budgets-shrink-while-capacity-balloons

3 TwinStrata, Industry Trends: State of Storage in 2014

4 http://www.gartner.com/technology/media-products/newsletters/active-navigation/1-2AYP85V/gartner.html

5 http://www.esg-global.com/lab-reports/hybrid-cloud-storage-from-microsoft/

6 http://media.amazonwebservices.com/Forrester_File_Storage_Costs_Less_In_The_Cloud.pdf

7 http://www.gartner.com/technology/media-products/newsletters/active-navigation/1-2AYP85V/gartner.html

8 http://www.gartner.com/technology/media-products/newsletters/active-navigation/1-2AYP85V/gartner.html

9 https://www.forrester.com/Cost+Analysis+Tool+Cloud+Versus+Internal+Storage+Deployment/fulltext/-/E-res58937

HubStor™

HubStor Inc.
Twitter: @hubstor

515 Legget Drive, Suite 800
Ottawa, ON K2K 3G4

T. 613.704.1737

Sales: sales@hubstor.net
Support: support@hubstor.net
www.hubstor.net