Active Archival Storage

A Cost of Ownership Analysis

By Brian Garrett
Analyst, Enterprise Strategy Group

October 2006

Executive Summary

The technology choice for implementing deep and active digital archives has historically been between tape and optical. Content addressed systems that are based on commodity disk drives have recently become available as an alternative. The Enterprise Strategy Group (ESG) was contracted by Plasmon, the leading optical archiving solutions provider, to perform a third party analysis of the cost of acquisition and ownership of each of these archival technologies – tape, optical, and disk.

This analysis is based on the actual case study of a financial services firm shopping for a 12 TB archive. The financial services firm needed a solution to archive 8 GB of new data and handle 2,500 queries daily. The cost of acquisition and ownership was measured over 3 years of operation. Only clearly quantifiable list prices in effect in May 2006 were included in this study to avoid the effects of subjective interpretation.

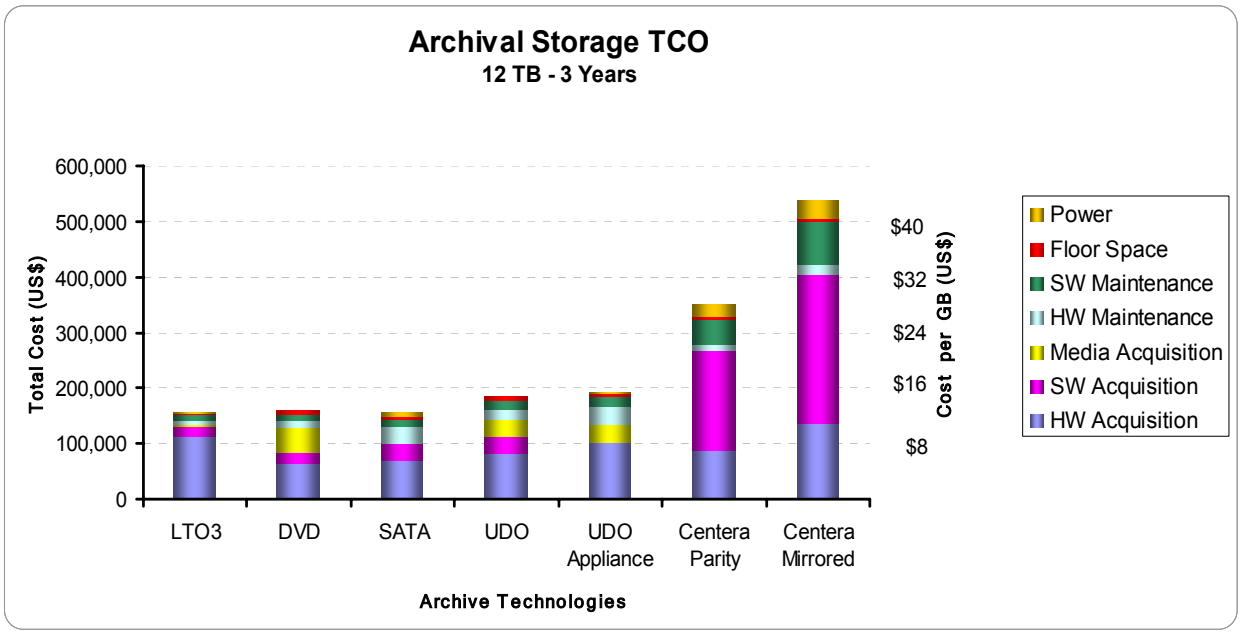

Figure One: ESG Archival Cost Analysis

The results, as shown in Figure One, *ESG Archival Cost Analysis,* clearly indicate that the cost of a Plasmon library populated with Ultra Dense Optical (UDO) technology is competitive with LTO-3 tape, SATA disk, and DVD optical technology and represents a fraction of the cost of a Centera disk solution[1]. In fact, the UDO Archive Appliance which contains a NAS device for easy integration with any IP network, a RAID box for caching to enable rapid reads and writes and a full compliment of UDO cartridges is still significantly cheaper than a Centera. **As a matter of fact, our analysis indicates that 12 TB of parity protected Centera capacity is 81% more expensive than a comparably configured automated UDO appliance, and a Centera Mirrored configuration is three times more expensive.** A closer look at the results reveals that the cost of a Centera solution is inflated significantly due to the price of software acquisition and maintenance. A more careful examination uncovers noticeable power consumption costs for the disk-based Centera system and the elimination of third party software acquisition costs for the UDO appliance due its built-in Hierarchical Storage Migration (HSM) capability.

[1] The methodology and data behind this diagram are documented later in this report and in the Appendix.

Proponents of the Centera architecture might object to the conclusion that optical technologies like UDO are significantly more cost effective than a disk-based Centera system. They would argue that the self-healing and scalable Centera architecture reduces the cost of administration when compared to multiple tape or optical libraries. That argument is explored in Figure Two, *Scalable UDO vs. Centera Analysis*.

The diagram compares a pair of 16 node parity protected Centera clusters to a pair of fully loaded Plasmon model UDO Archive Appliance 638. Pairs of Centera clusters and UDO Archive Appliances are then added to scale near-line capacity to 150 TB. The total cost per Gigabyte presented above was used to calculate the costs depicted on the Y-axis in millions of US dollars.

Let's focus first on the first set of data points at the bottom left of the diagram, where we've already seen that an automated UDO Archive Appliance is 81% more cost effective than parity protected Centera. Since all of the media for the automated UDO solution is inside the library, there is no need for a system administrator to handle removable media. The EMC Centera is known for its ease of administration. However, customers may find it difficult to justify the $310,309 difference due to reduced administration costs.[2]

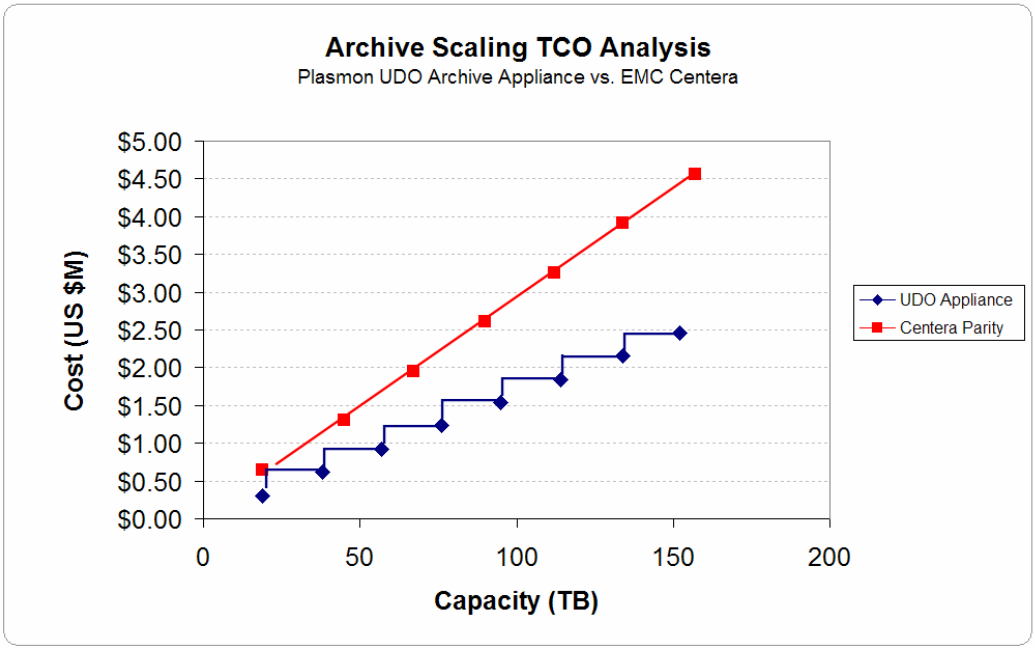

Figure Two: Scalable UDO vs. Centera Analysis

Now consider the cost of a 150 TB capacity point as shown toward the upper right of the diagram, where the cost of eight optical libraries is compared to a 224 node Centera cluster. The Centera upgrade path from 22 TB to 157 TB is depicted as a smooth red-line to depict the fact that nodes can be added to a singly managed pool of storage as needed. Although removable capacity can be added within an optical library for additional capacity as needed, we depict the optical solution as a step-wise series of fully populated library additions. The cost difference at 150 TB is $2.1M. **Conceding the fact that managing sixteen optical libraries is more difficult than managing a single Centera system, ESG finds it difficult to believe that the majority of end users would incur $2.1M of additional system administration costs over three years for a multi-library UDO optical archiving solution.**

[2] It should be noted that while this type of analysis holds true for near-line archiving applications with all media maintained within an automated solution, it does not apply when some portion of the archive resides outside of a library on removable media.

ESG Research

End user research performed by ESG indicates that cost is a major concern of storage professionals considering a move to disk-based archival. ESG conducted a survey of 163 North American storage professionals and IT managers in November 2004. The majority (58%) of respondents responsible for 10 TB or more of primary storage fell into a "fast follower" camp, indicating that they would consider replacing some or all of their enterprise-class tape libraries with a disk-based archive, but have yet to take any action to do so.

Nearly half of the fast followers (44%) say they would consider a disk-based solution to support on-line or near-line archiving. These fast followers who are motivated by compliance and corporate governance initiatives are looking for a new cost-effective tier of near-line storage between on-line disk and off-line tape. As shown in the following diagram, the vast majority of fast followers (77%) indicate that the cost of new disk–based solutions is the primary objection. It should also be noted that fast followers are also concerned about the reliability of ATA drives, the lack of media portability and the ability to ensure regulatory compliance due to the absence of Write Once Read Many (WORM) media – all of which are not a concern when UDO optical technology are used for near-line archival.

"We looked at a variety of archiving technologies to replace DVD jukeboxes that were multiplying like rabbits. We found the disk-based EMC Centera to be too expensive. Our fully automated Plasmon UDO library costs three to four times less than a Centera from EMC and is eight times faster than our old DVD jukeboxes."

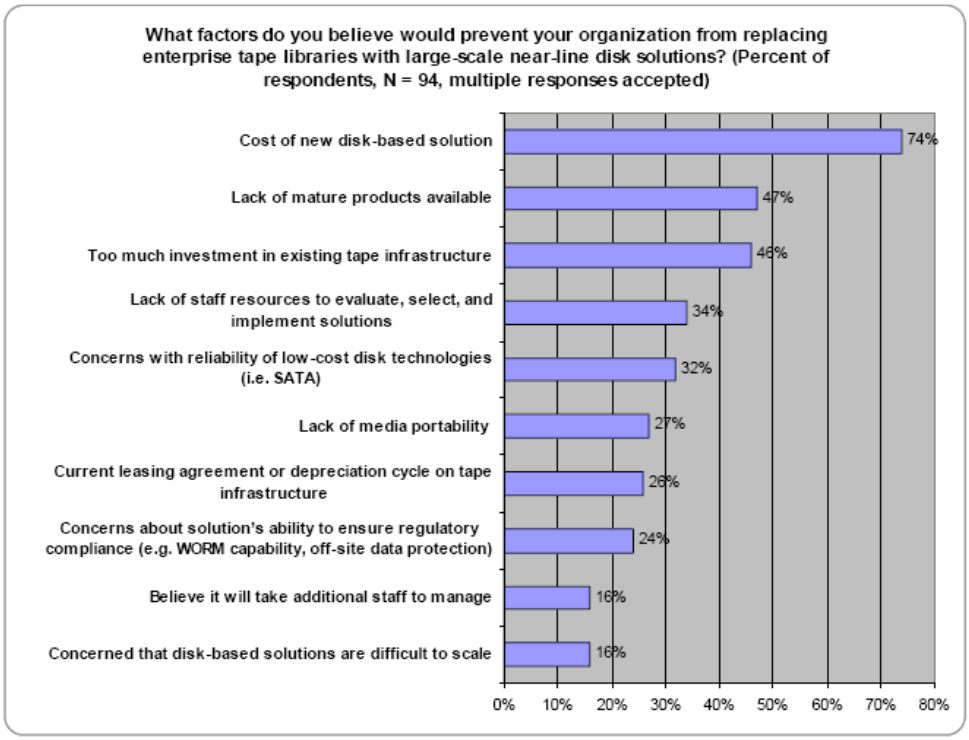

Figure Three: ESG Research

Technology and Cost

There are significant differences between the CD and DVD technology we use in our homes and the professional UDO optical technology analyzed in this report. The first and most obvious difference is automation. Professional optical libraries can be packed full of robotically controlled media, support multiple drives which act in parallel for improved performance, and come in sizes as big as a refrigerator. For example, Plasmon optical libraries (Figure Four, *Plasmon UDO Archive Appliance and G-series Libraries*), with capacities in excess of 19 TB, support advanced features including dual pickers, barcode readers, hot/warm swappable drives, redundant power supplies, and very high duty cycles. Also, with the introduction of the UDO archive Appliance there is now an integrated solution available that enables not only ease of use but fast access along with the authenticity and reliability of UDO.

Figure Four: Plasmon UDO Archive Appliance and G-series Libraries

Figure Five: UDO Cartridges

Another major difference between professional optical and the technology we use at home is how the media is packaged and handled. UDO disks are completely enclosed in a rugged, shock resistant cartridge and are only exposed to the elements when the media is loaded into the drive for read and write operations.. (Figure Five, *UDO Cartridges*). The cartridge packaging of UDO media virtually eliminates the risk of scratches and contamination experienced with traditional optical discs.

UDO uses the latest blue laser technology, which delivers 30 GB of capacity per disc, compared to DVD with 9.4 GB of capacity per disc. Legacy optical discs can be significantly more expensive than UDO solutions. This is due to the high cost of lower capacity legacy media, which increases the number of libraries needed for a comparable UDO configuration.

UDO drives use Write-Once-Read-Many (WORM) media that cannot be erased or overwritten. Although software and firmware can be used to make non-WORM technologies such as hard drives behave like WORM media, the native WORM capability of UDO is a better choice due to the absolute data authenticity it provides and reduced system overhead. UDO uses a high quality phase change recording surface that is manufactured to much higher standards than consumer DVDs.. UDO media is also available in a Rewritable and Compliant WORM formats, which are based upon physically different media materials and structures., UDO Compliant Write Once media provides WORM authenticity with the ability to physically destroy individual records when they reach the end of their lifecycle. Multifunction UDO drives and G-Series libraries support all three UDO media types allowing different data sets to be archived on appropriate media.

As mentioned earlier, in addition to the family of UDO libraries, Plasmon offers a UDO Archive Appliance, which combines RAID protected disk and automated UDO into a single network attached storage system. The tiered storage design of a Plasmon UDO appliance combines the ease of use and access speed of a network attached storage system with the cost effective long term assurance of automated, UDO optical media. A four drive G438 library and a similarly configured four drive UDO Archive Appliance 438 with 2 TB of usable RAID protected hard drive capacity were used for this analysis. The UDO Archive Appliance includes integrated archive management software so no additional interconnect software is needed.

LTO tape technology has price/performance characteristics and WORM emulation support which are suitable for moderately active archive applications. Un-compressed 400 GB LTO-3 tape cartridge capacity was used for this analysis, based on the assumption that data and images stored by archiving applications is typically compressed before it is written to an archive[3].

The financial services application examined in this study needs to be able to handle 8 GB of new data and 2,500 accesses for existing data daily. 2.500 requests daily equates to 313 requests per hour. The performance characteristics of LTO-3 tape drives indicated that 12 tape drives are needed to meet that requirement[4]. Given the larger capacity of LTO tape cartridges compared to UDO optical cartridges and the likelihood that tape exchanges will be less frequent, an ADIC Scalar i5000 with only six LTO-3 drives was used for this analysis. It should be noted that an archiving application which requires less read activity could be serviced with a tape library configured with fewer LTO drives and would cost less than the LTO library presented.

Two types of disk technology were chosen for this analysis – an EMC Centera and an EMC CLARiiON CX300. The Centera was chosen because it is a market leading disk-based archive system that is specifically designed to meet the WORM functionality, management, and authenticity requirements of compliance-mandated archive applications. EMC Centera software costs include mandatory CentraStor software and the CentraStor CompliancePlus Option, which is needed in mandated environments that require auditable WORM attributes. Two Centera configurations were included in the analysis, one with parity for protection from a hard disk failure (CPP) and the other with mirroring (CPM). The parity protected configuration, which is the most popular, requires fewer disk drives and therefore costs less than a mirrored configuration with less chance of data loss due to a hard drive failure.

Fibre Channel attached disk arrays full of affordable dense SATA hard drives are increasingly being adopted by a large number of users for backup and archival. At first blush the low cost and high density of SATA disk arrays seems an attractive option for archiving, yet SATA arrays lack the WORM capabilities of Plasmon UDO systems. It should also be noted that SATA arrays and the Centera lack the media portability of a Plasmon professional optical solution. With these considerations in mind, an EMC CX500 with 500 GB SATA drives was included in this analysis. The cost of a pair of FC host bus adapters, an eight port FC switch and PowerPath failover software for a single server is also included.

Since each technology option could not be configured to exactly meet the 12 TB requirement, fully populated configurations that came closest to 12 TB were used. A price per GB of capacity was then calculated for each configuration and multiplied up to 12 TB for a rational cost comparison. The number of tape and optical cartridges were determined by matching publicly available performance specifications with the need to sustain 8 GB of write data and 2,500 requests daily[5].

The cost of QStar HSM software was included in this analysis. QStar HSM is an enterprise-class hierarchical storage management software package that supports tape, optical, and Centera hardware.

[3] A compression rate of 2:1, which would cut the media cost in half, would not noticeably change the results of this cost analysis, due to the fact that LTO media costs contribute relatively little to the overall cost of a 12 TB automated LTO-3 solution.
[4] Details are available in the Appendix
[5] Details are available in the Appendix.

Although there are a large number of software packages that can be used for a near-line archival cost analysis, the choice of archival software matters little in this analysis since the cost of archival software, and the server(s) it runs on, should be the same regardless of the hardware technology chosen for archival. Archiving software solutions that have not been ported to the Centera programming interface require the use of CUA servers and software so that a Centera can be accessed using standard network file system protocols. Because QStar HSM supports the Centera programming interface, the cost of Centera Universal Access (CUA) hardware and software was not included in this analysis.

The cost of commercial power at $0.0919 per Kilowatt hour[6], floor space at $3,235 per square meter, and cooling at 60% of the cost of power were included in this analysis. These rates, which are typical for a large US or European city, would need to be adjusted higher for areas like California and New York City. The cost of power for the Centera system, which ranged between $6,183 and $9,274 a year, was noticeable compared to the automated libraries, which averaged $451 a year. The relative cost of floor space for all configurations was negligible, although a 12TB Centera configuration approaches a ton in weight, which is an order of magnitude higher than a UDO library.

Although the cost of disaster avoidance was not included in this analysis, media survivability and remote vaulting are issues that should be considered when implementing a long term digital archive. Data residing on optical media has a better chance of surviving a disaster (e.g. an earthquake or flood) than tape or disk. Removable tape and optical media are easy to replicate and transport to a remote site for safe keeping. Replicating a disk-based Centera or CLARiiON system is also easy, but can be quite expensive due to the cost of another set of disk drives, EMC software at the remote site, and the recurring cost of WAN bandwidth.

And finally, the cost of media maintenance should be considered when deploying a digital archive in which data will reside for many years. Tape and disk technology have inherent media maintenance costs (e.g. tape re-tensioning and disk migration) that can inflate the cost of digital archiving as compared to optical media, with a typical specified shelf-life of 50 years.

[6] The rolling 12 month average cost of electricity in the United States for the period ending July, 2006 as reported by the US department of energy was used for this analysis (http://www.eia.doe.gov/cneaf/electricity/epa/epat7p4.html)

The Bottom Line

Compliance regulations, corporate governance initiatives, and the explosive growth of unstructured digital content are driving the adoption of archiving strategies and technologies. Organizations supporting business needs with on-line disk technologies are evaluating the deployment of a new tier of cost effective near-line capacity for long term data archival. ESG Research has shown that the high cost of emerging near-line disk-based systems is a major concern for storage professionals and IT managers.

The true cost of acquisition and ownership of tape and optical solutions, compared to systems based on commodity disk drives, has been widely debated in the industry recently. Common mistakes made during such a comparison include a myopic focus on the cost of the raw media and an inflated estimate of system administration costs. Despite the fact that some vendors suggest that tape and library configurations require dramatically higher administration costs, the subjective costs associated with system administration were not included in this study. The Enterprise Strategy Group stands behind this assumption based on customer feedback that confirms the fact that modern tape and optical libraries can be configured with massive amounts of near-line archival capacity, which eliminates the system administration costs associated with removable media handling.

ESG analysis indicates that the cost of a Plasmon UDO library or UDO Archive Appliance populated with professional optical technology is competitive with LTO-3 tape, SATA disk, and DVD optical technology and represents a fraction of the cost of a Centera disk solution. The next generation of professional optical technology (UDO2), scheduled for availability in early 2007, will double the capacity to 60 GB and will be offered at roughly the same price as existing 30 GB cartridges. Plasmon UDO libraries shipping today will support the newer, more cost effective UDO2 media when it becomes available. As a result, automated UDO technology will become an even more attractive and cost competitive solution for digital archiving.

The intent of this analysis is to provide readers with a reasonable starting point for the comparison of the cost of archival technologies available on the market today. Although a variety of hard and soft costs were not addressed in this analysis, ESG believes that the methodology and results presented in this report form a valid relative comparison of the most significant costs of archival storage ownership. ESG encourages readers of this report who are considering an active archiving solution to perform their own cost of ownership analysis. We are confident that such an analysis will make a compelling case for the consideration of automated professional optical technology.

Appendix

Table One: 12 TB Adjusted Cost Summary

Archive Type	System Capacity	Actual System Cost	$/GB	Adjusted 12 TB Cost
LTO-3 Library	20.0 TB	$156,909	$13.08	$156,909
DVD Library	13.8 TB	$184,585	$13.38	$160,509
CLARiiON CX300	12.0 TB	$167,189	$13.93	$167,189
UDO Library	13.1 TB	$203,432	$15.53	$186,350
UDO Appliance	15.1 TB	$212,063	$16.19	$194,256
Centera CPP	11.6 TB	$337,764	$29.12	$349,411
Centera CPM	10.1 TB	$452,920	$44.84	$538,123

Table Two: System Cost Breakdown

Archive Type	HW$	SW$	Media$	HW Maint$	SW Maint$[7]	Floor Space$	Power$	Total$
LTO-3 Library	113,185	18,100	2,700	8,160	9,774	3,386	1,604	$156,909
DVD Library	64,743	20,957	44,144	11,605	11,317	6,646	1,097	$160,509
CLARiiON CX300	69,560	30,740	0	29,922	13,650	5,805	6,715	$155,852
UDO Library	83,472	30,321	30,813	17,125	16,379	7,001	1,239	$186,350
UDO Archive Appliance	103,878	0	30,813	33,683	0	7,001	2,503	$194,256
Centera CPP	87,724	179,690	0	11,346	45,859	5,605	19,188	$349,411
Centera CPM	136,634	268,634	0	17,337	76,025	6,437	33,056	$538,123

Media Type	Vendor	Product	Drive Count	Media/ Drive Capacity	Media Count	Usable System Capacity
Tape (LTO-3)	ADIC	Scalar i500	6	400GB	50	20 TB
Disk	EMC	Centera Parity	32	500GB	-	11.6 TB
Disk	EMC	Centera Mirrored	48	500 GB	-	10.1 TB
DVD	Plasmon	D1525	8	9.4GB	1,475	13.8TB
UDO	Plasmon	G438	4	30GB	438	13.1TB
Disk	EMC	CLARiiON	29	500 GB	-	12 TB

[3]This column indicates the cost of QSTAR HSM software maintenance plus the cost of storage system software maintenance if applicable.

Table Four: Optical and Tape Drive Specifications

Drive / Library Specs.	DVD	UDO	LTO-3
Load Time	15 sec	5 sec	19 sec
Unload Time	3 sec	3 sec	19 sec
Average Seek Time	200 msec	35 - 50 msec	46 sec
Average Rewind Time	0 sec	0 sec	44 sec
Media Exchange Time	6 sec	6 sec	6 sec
Average Data Access	5 sec	5 sec	5 sec
Average Access Cycle	29 sec	19 sec	139 sec
Access Cycles per Hour	124 cycles	189 cycles	26 cycles
Drive Count for Read	3 drives	2 drives	12 drives

Table Five: Environmental Factors

Archive Type	Power Watts	BTU/hr	Power $/hr	Cooling $/hr	Total $/hr	Total $/yr	Total 3 years	Adjusted 3 years
LTO-3	415	1,418	0.0381	0.0229	0.0610	535	1,604	$1,745
DVD	300	1,025	0.0276	0.0165	0.0441	386	1159	$1,097
CLARiiON CX300	1598	5374	0.1469	0.0881	0.2350	2058	6175	$17,512
UDO Library	350	1,196	0.0322	0.0193	0.0515	451	1352	$1,348
UDO Archive Appliance	707	2,414	0.0650	0.0390	0.1040	911	2732	$2,225
Centera CPP	4,800	16,400	0.4411	0.2647	0.7058	6183	18548	$20,879
Centera CPM	7,200	24,600	0.6617	0.3970	1.0587	9274	27,822	$35,970