



# The Green Genes of CAStor

By Paul Carpentier, CTO and Founder, Caringo.

March 2008

---

"Caringo is committed to advancing CAStor such that it is not only proclaimed be "the Porsche of read/write functions" but also "the Prius of power consumption".

## Introduction

The rapidly growing environmental awareness in all segments of society is causing a landslide in the positioning and marketing of IT technologies around the concept of "green." Aspects that were previously ignored have made it to the top of the short list of decisive criteria. Small wonder that all IT vendors – regardless of the nature of their wares – are trying to find whatever green arguments they can find to secure a rank in this new pecking order. The resulting spectacle is often painful to watch, as most often their technologies are strongly rooted in an era where power consumption, or "bang for the watt", simply wasn't a factor. Bang for the buck was what counted, and power was cheap.

Caringo designed CAStor for this new era where power and cooling are now major defining characteristics of any storage solution. This paper demonstrates how the same architectural foundation that makes CAStor economically cost effective also makes it naturally "green". This results in total solutions that have a substantially smaller footprint not only in capital expenditures and management labor but also in energy consumption, cooling, data center space and associated costs. Being environmentally friendly or "green" is part of the genetic makeup of CAStor software.

## The Information Power Cycle

When considering power economics in a storage context, it is vital to look at the total picture over the entire lifetime of the data being stored.

## Unlimited hardware choices —

### Let your business and energy requirements drive the decision

The decision to develop and offer CASTor as a pure software offering wasn't just a marketing or distribution choice. It was made to empower customers to select the perfectly tuned solution among the widest possible range of hardware. This stands in stark contrast with the one-size-fits-all bundled storage appliances in the market today. These are a poor match for the very diverse use cases of active and archive content storage, and invariably result in real power-guzzling implementations.

CASTor runs on standard, commodity server hardware (X86 architecture) and as a result, many choices are available as to the selection of CPU speed and power consumption, disk to CPU ratio, RAM to disk ratio, and number of nodes in a cluster. CASTor performance remains outstanding, even on low power processors. The key here is proper planning. Consider the requirements of the planned content storage solution relative to client throughput and capacity, and select CPU speed/power profiles accordingly. Selecting large capacity SATA disk drives will allow savings not just on the dollars per Gigabyte, but also on the watts per Gigabyte. Since CASTor can be scaled in real time, without ever taking down the cluster, a moderate amount of over-provisioning will take care of any eventuality in case of failing hardware. This safety factor does not adversely impact the power effectiveness of the total solution. In other words, there is no reason you should ever run more hardware than strictly necessary for the capacity at any given time. CASTor leaves you in total control.

Because CASTor can use the most power-efficient hardware available today, you can take advantage of the 'greenest' servers available. There is no need to wait for long product cycles typical at hardware vendors today. So, if you consider that next year's hardware is going to be even more cost-effective *and* more power-efficient, there's no reason to pre-finance tomorrow's capacity at today's prices while needlessly consuming power in the process. CASTor allows for the inter-mixing of hardware brands, types, speeds and capacities. When a more power or cost-effective server becomes available, it, will seamlessly plug right in. CASTor's market-based bidding mechanisms in the cluster automatically load-balance between the older and the newer nodes based on characteristics such as performance and RAM-to-disk ratio. Hardware or capacity upgrades may be done in small increments on a quarterly basis, or in larger increments on a yearly or longer basis. With CASTor, your business requirements dictate your choices, not your vendor. CASTor's pricing model based on installed capacity encourages and amplifies the returns of such a practice.

#### **A final thought**

CASTor perfectly supports asymmetrical sub-clusters where the "primary" replica of an object resides on fairly fast nodes with not too many disks per CPU, while the "spare" replica is stored on slower but lower power nodes with many disks per node. Read and write access times will hardly be influenced since they are defined by the behavior of primary replicas, but the power savings in such a configuration can be as much as 25%.

## Better Data Management is Naturally Greener

Nearly all industrial-strength storage solutions today provide some form of protection against hardware failure. Most often some form of hardware or software RAID or a block- or object-based equivalent thereof is used where the level of safety provided can be varied by selecting one or the other RAID schema. Unfortunately, this always applies equally to the whole infrastructure, independent of actual business value of the provided protection level for any specific piece of information. If you are considering a storage platform storing hundreds of Terabytes where only 25% of the information is crucial, this 'wasted' RAID usage can result in significant wasted power and money.

CAStor allows for the very precise and granular specification of the number of replicas of any individual stored object (file). The driving application may provide metadata with each object, or class of objects that tells CAStor how many replicas of the object to maintain. It can also specify to change the replica count over time such as when the business value of the protection decreases. Automatic deletion after a given period may also be programmed in this fashion. Excessive replicas will be automatically trimmed by the system resulting in space recovery. For detailed storage management tasks such as content trimming to re-capture and reuse storage capacity, the full automation offered by CAStor provides the only reasonable guarantee they will actually happen.

Even more critically, rather than blind "remote mirroring", CAStor's Wide Area Replication facility designed for disaster recovery and content distribution and collection purposes allows for similar, very selective and thus space and power saving policies.

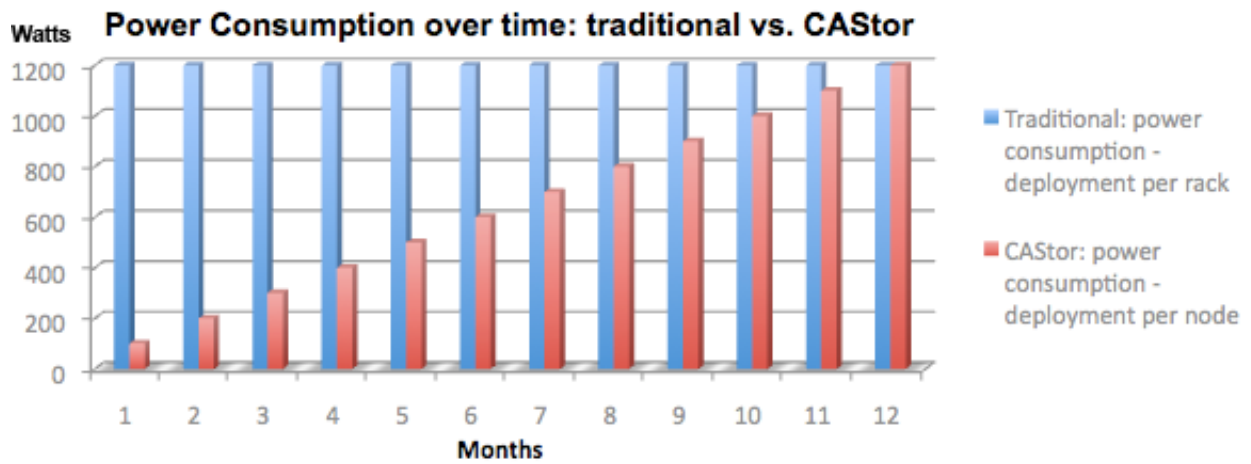
When optimally designed and implemented, Caringo's exclusive Single Tier ILM (Information Lifecycle Management) obsoletes the practice of separating primary and secondary content storage, archives, backups and disaster recovery. Indeed, CAStor's performance is fast enough to assume the role of primary content storage, yet provides the economics that eliminates the need to move content to any other tier of storage before the end of its useful life. When designing corporate storage strategies, IT architects should compare this scenario with the full array of hardware, software and services needed for any multi-tiered alternative, including the data overlap between tiers and other unavoidable inefficiencies when cobbling together separate and diverse technologies. Again, economy and ecology will prove to go hand in hand.

### **The bottom line**

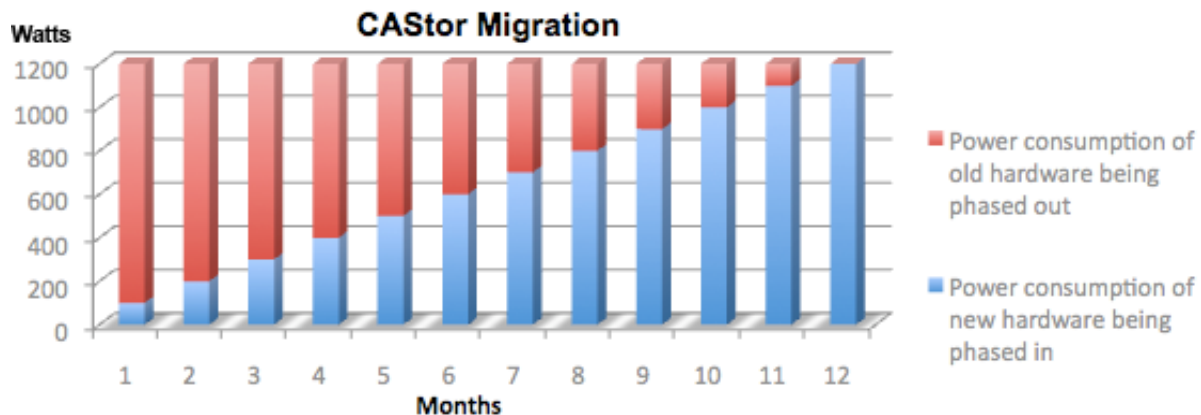
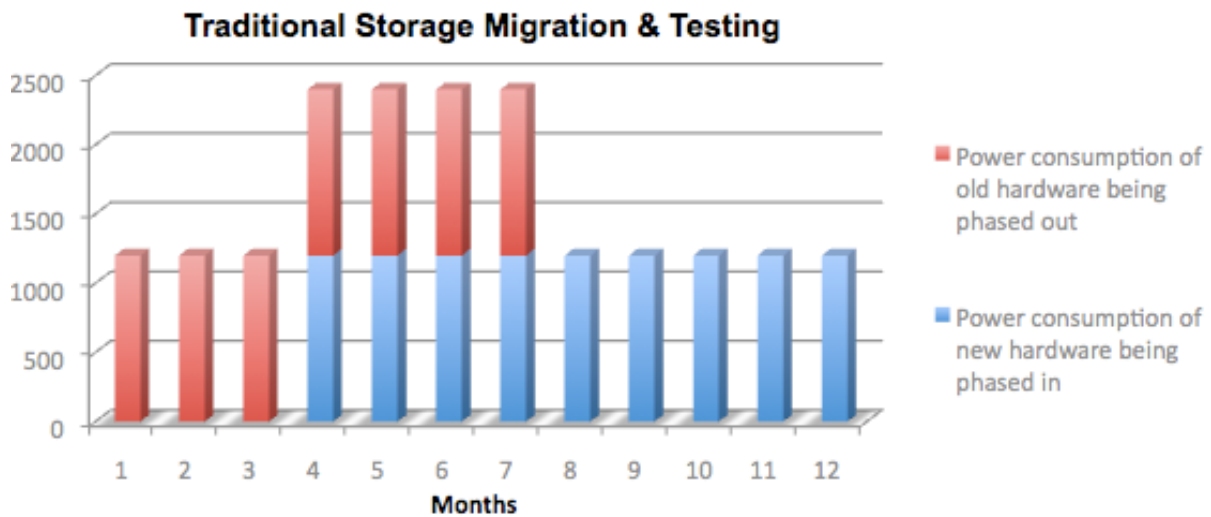
When using CAStor technology, the actual data and the amount of space needed to store it closely aligns to the business value more so than traditional storage offerings. In a storage infrastructure, less unneeded data simply means less power consumed.

### **Thrifty by design**

As is often the case with many things, the devil is in the details, and it's no different with the economics of content storage. Many platforms have trouble operating at high disk utilization levels, either reliability or performance wise; consequently, it is discouraged by standard operating procedures or field support personnel. Very often that fact is conveniently forgotten when comparing solution economics. Since CAStor has no problem whatsoever to run at very high utilization rates, Caringo recommends that all unused, unusable or spare capacity of any kind be considered when doing the math.



It was mentioned earlier in this document that economic and power savings are possible because you can always run with the most recent and advanced server hardware. This means the CASTor cluster can be built or expanded with the most affordable and energy efficient hardware available. But when comparing to other platforms, the upgrade granularity itself needs to be taken into account. The CASTor method is continuous, smooth expansion without disruption along a gradual slope rather than large coarse increments. Assuming comparable hardware, adding a node a month will consume about half the power as compared to a single expansion with a 12-node cabinet (appliance model), as illustrated above.



Moreover, refresh of hardware in a CASstor storage cluster works in the same fashion by adding new and removing old nodes in a running cluster. Compare that to traditional scenarios above where all old and new hardware need to be running full steam side-by-side for months and where you can consider yourself lucky if the doubled power consumption is your primary concern at that point.

## Architectural perspectives for saving more power

Some companies would probably be quite satisfied with this level of achievement in an area as important as power consumption and cooling. Not Caringo. We continue to work hard to leverage the unique advantages of CASstor's underlying CAS architecture to improve energy efficiency. Looking forward, Caringo has identified opportunities in the near future to further optimize CASstor's use of hardware in areas such as:

- Support for automatic spin-down of disks and stepping down CPUs based on usage patterns.
- Systematically spinning down half of the disks (or more) in a cluster, for as much as 70% to 90% of the time. These disks will contain only "spare" replicas and need to be awakened simply for housekeeping and content recovery.

### Longer Term:

- Content compression, potentially using specific content-optimized compressors to reduce capacity and related power consumption
- More controls for throttling health processing tasks to meet, not exceed requirements, resulting in less CPU/disk use, hence less power consumption
- Metadata-driven archiving of known-silent data to largely powered-down nodes

## Conclusion

When comparing Caringo's CASstor with traditional content storage offerings, and evaluating the complete picture in terms of actual watts consumed per GB stored, it is evident that CASstor was designed for an era where power and heat considerations are critical. Ongoing development will leverage further potential in the near future to realize power savings of 50% to 80% as compared to today. Caringo is committed to advancing CASstor such that it is not only proclaimed be "the Porsche of read/write functions"<sup>1</sup> but also "the Prius of power consumption".

### About Caringo

With CASstor content storage software, Caringo has re-imagined content addressed storage (CAS) from the ground up. Its flagship CASstor product is third-generation technology that leverages standard commodity hardware that dramatically changes the economics of corporate content storage.

For more info, visit [www.caringo.com](http://www.caringo.com).

---

<sup>1</sup> Preimesberger, Chris. "10 Disruptive New Storage Technologies." eWeek. December 2007