

5 Steps to getting started with Open Source Software Defined Storage – and why you should take them

EXECUTIVE SUMMARY

Back in 2013, analyst group IDC calculated that the total amount of data created and replicated in the world had edged beyond 4.4 zettabytes – a staggering number. The statement made the headlines and was widely repeated across media websites dealing with Big Data and the related storage issues. At the time, IDC attributed the enormous growth to approximately 11bn connected devices, – all generating and transmitting data, many containing sensors which also generate data.

IDC also predicted that the number of connected devices would triple to 30bn by 2020, before near tripling again to 80bn a few years later. If you've ever wondered what analysts mean by 'exponential' data growth this is what they are talking about, and the growth keeps on coming, even the forecasts for data growth are growing: three years later in 2016, IDC revised their predictions upwards, forecasting that by 2025 the total volume of data stored globally would hit 180 zettabytes. Divide 180 by 4.4 and you have a staggering growth rate of 40 x in just nine years.

Of course not all of that data is made by enterprises, but IDC say they are responsible for 85% of it at some point in its lifecycle. So, whilst enterprises might not make all the data, and might not drive all its growth, they still have to architect and manage storage systems that can cope with the multiple challenges it brings.

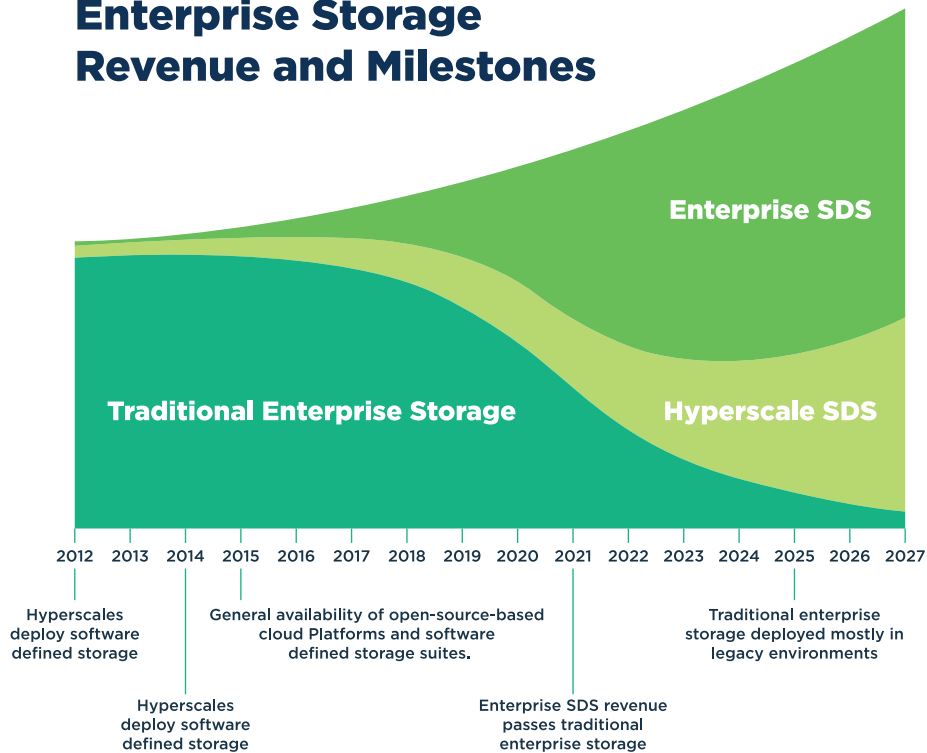
OPERATIONAL CHALLENGES: VOLUME GROWTH, DIGITAL TRANSFORMATION AND ANALYTICS

Storage costs may have come down a lot in recent years, but the operational issues associated with managing it keep on piling up. Systems reach capacity and must be replaced. The surrounding architecture is shifting as organisations undergo digital transformation and migrate to hybrid and public cloud environments. Decisions must be made about what data should be kept and what should be deleted – decisions which must be kept on the right side of the law, and which revolve not only around data itself, but on the value of that data to the enterprise; a bigger challenge than some might think as the financial potential in data is not always clear to the IT team, who are after all better placed to understand volume than value: a shortcoming which can lead to the enterprise equivalent of assessing the complete works of Shakespeare based on the number of pages in the book.

There are also substantial problems that come from moving large data sets over limited cabling: the backup routines that have decreasing windows, the challenges with replication and recovery that increase with the related increase in disk failure, the volume of unstructured data that comes with data like video, the security and compliance challenges, making data available for analytics, and for many, the ongoing cost of skilled technical staff for management.

These challenges aren't going away: like your data, they are only going to get bigger. Unsurprisingly, enterprises are turning to software defined storage as the solution, indeed IT Brand Pulse predict that not only will SDS overtake traditional storage by 2020, but that 70 to 80% of storage will be managed on less expensive or commodity hardware managed by software in the same timeframe. If software defined is the answer to this challenge, why SUSE?

Enterprise Storage Revenue and Milestones



SEVEN REASONS WHY YOU SHOULD CHOOSE OPEN SOURCE SDS from SUSE.

Open source software defined storage on Ceph platform offers several key advantages:

- Cost reduction through elimination of proprietary software licensing costs
- Avoidance of proprietary vendor software lock-down
- Reduction of hardware costs by moving to commodity hardware
- Support for Object, Block and File and key protocols on a single platform
- Scale out infrastructure – simply add new servers and nodes as capacity increases
- Service, support and management to mitigate risks and control operational cost
- Consistent innovation and first-to-market roadmap improvements

GETTING STARTED WITH OPEN SOURCE SOFTWARE DEFINED STORAGE

- 1** Start small. Storage administrators are rightly risk averse – so choose your first deployment where you can prove the value in terms of cost reduction without putting mission critical data or processes at risk.
- 2** Find the right use cases. Good applications for Ceph Jewell include unstructured data like video footage, where sheer volume of data presents challenges in costs,

volumes, back-up and retention – simply being able to keep video files into the mid-term. Another good example is the cold store – where Ceph can be cheaper than services like Amazon Glacier in terms of dollars per GB, yet remain on premise and avoid hidden costs for retrieval should you need your data back quickly.

- 3** Scale your usage with your skillset. As with any new technology, it takes time to become familiar with Ceph and build skills and confidence – both your own and your organisations'. Up your deployment in line with your knowledge and capability.
- 4** Align your strategy for storage with your strategy for the data centre – its not only storage that is moving to software defined. Consider what your infrastructure will look in the future as enterprises moved towards software defined everything. How will your data centre look in five years' time?
- 5** Seek expert help when and where you need it. As you move from the periphery to the centre, complexity and risk increase – manage that risk and maximise the benefits by working with skilled third parties.