Understanding the Cost of Cloud Computing and Storage

ILLINOIS INSTITUTE OF TECHNOLOGY

DataSys
Data-Intensive Distributed Systems Laboratory

Iman Sadooghi
Computer Science Department
Illinois Institute of Technology
isadoogh@iit.edu

Dong Fang Zhao
Computer Science Department
Illinois Institute of Technology
dzhao8@iit.edu

Tonglin Li
Computer Science Department
Illinois Institute of Technology
tli13@hawk.iit.edu

Ioan Raicu
Computer Science Department
Illinois Institute of Technology
iraicu@cs.iit.edu

Abstract

Cloud computing is a new paradigm that enables everyone to use the remote resources in any demanded scale. Users are able to use computing and storage resources on-demand. Big companies as well as many startups have provided large scale cloud resources to different level of users. The users of cloud resources vary from huge tech and non-tech companies to the small start-ups and individual users. Unfortunately the pricing of these products have not yet adapted to the budget limit all of the potential users. There is a big need for lower cost storage system that can be used by the individual users. Also in the computing part, many small businesses have the potential to use these resources as their building blocks and avoid infrastructure costs.

In the storage, the gap between the raw storage price and the price that is offered by the current online storage providers determines that there is a big opportunity to provide a cloud storage system that focuses on normal end-users instead of big companies.

The compute part of the cloud also has potential to grow. Looking at the offered price by the cloud service providers brings the question in mind that whether there is opportunity to provide low cost compute resources in this area.

Cost of Storage
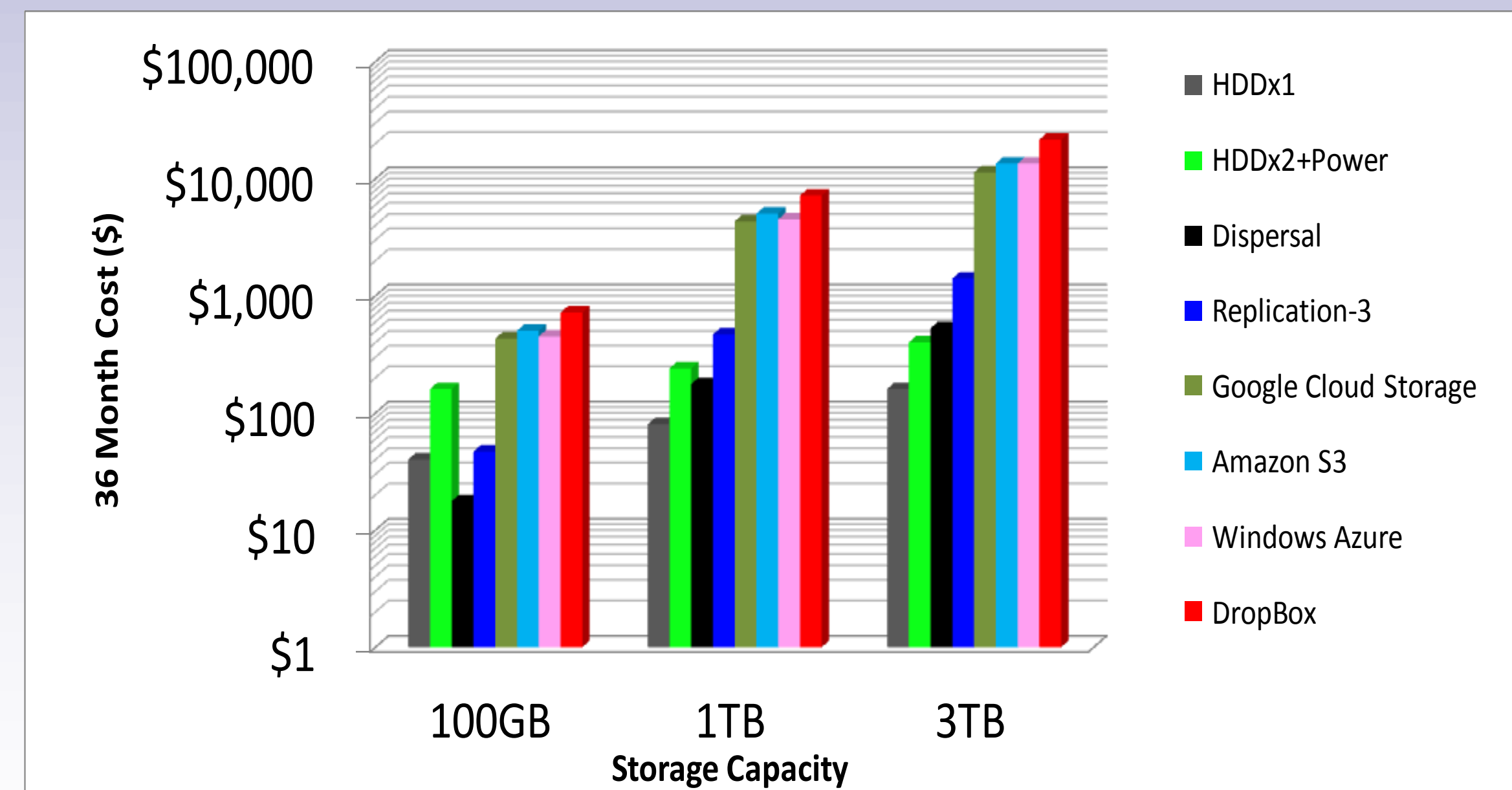
- The cost of hard drive it self for a three year period without considering other costs is still high.
- Setting a personal storage backup costs almost the same as the cost of keeping the data on cloud.
- Considering the maintenance cost, it is more cost effective to use cloud storage in small scale.

Storage Cost Comparison
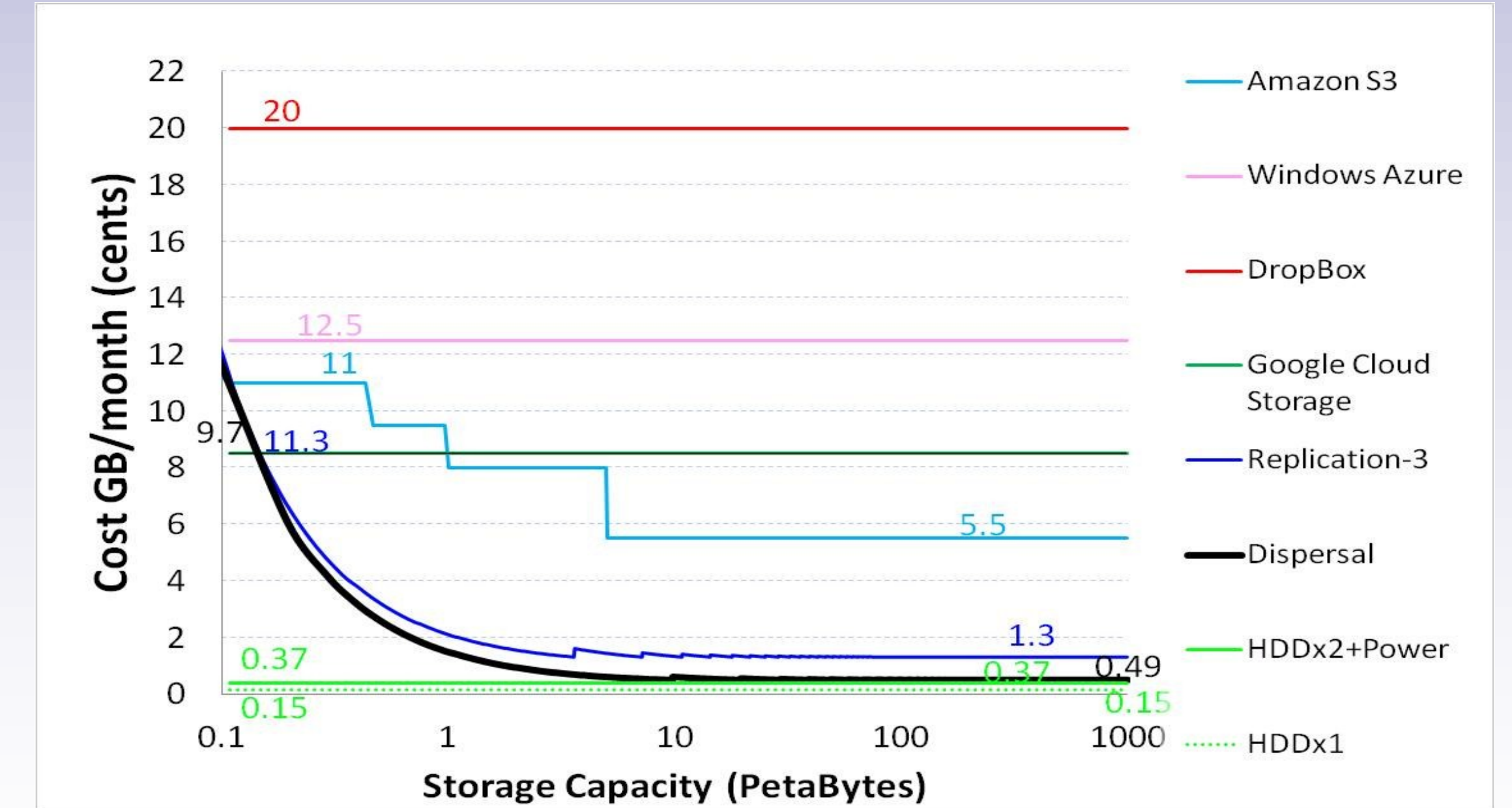
- The private cloud cost is slightly higher than the Amazon S3 and Google Cloud Storage pricing up to size of 100GB.
- At the size of 1 Peta Bytes the difference is up to 4x.
- At larger scales the private cloud is up to 11x cheaper than the cheapest provider cost.
- The dispersal storage technique price is 2.65x cheaper than the private cloud cost at large scale. Both techniques provide the same reliability.
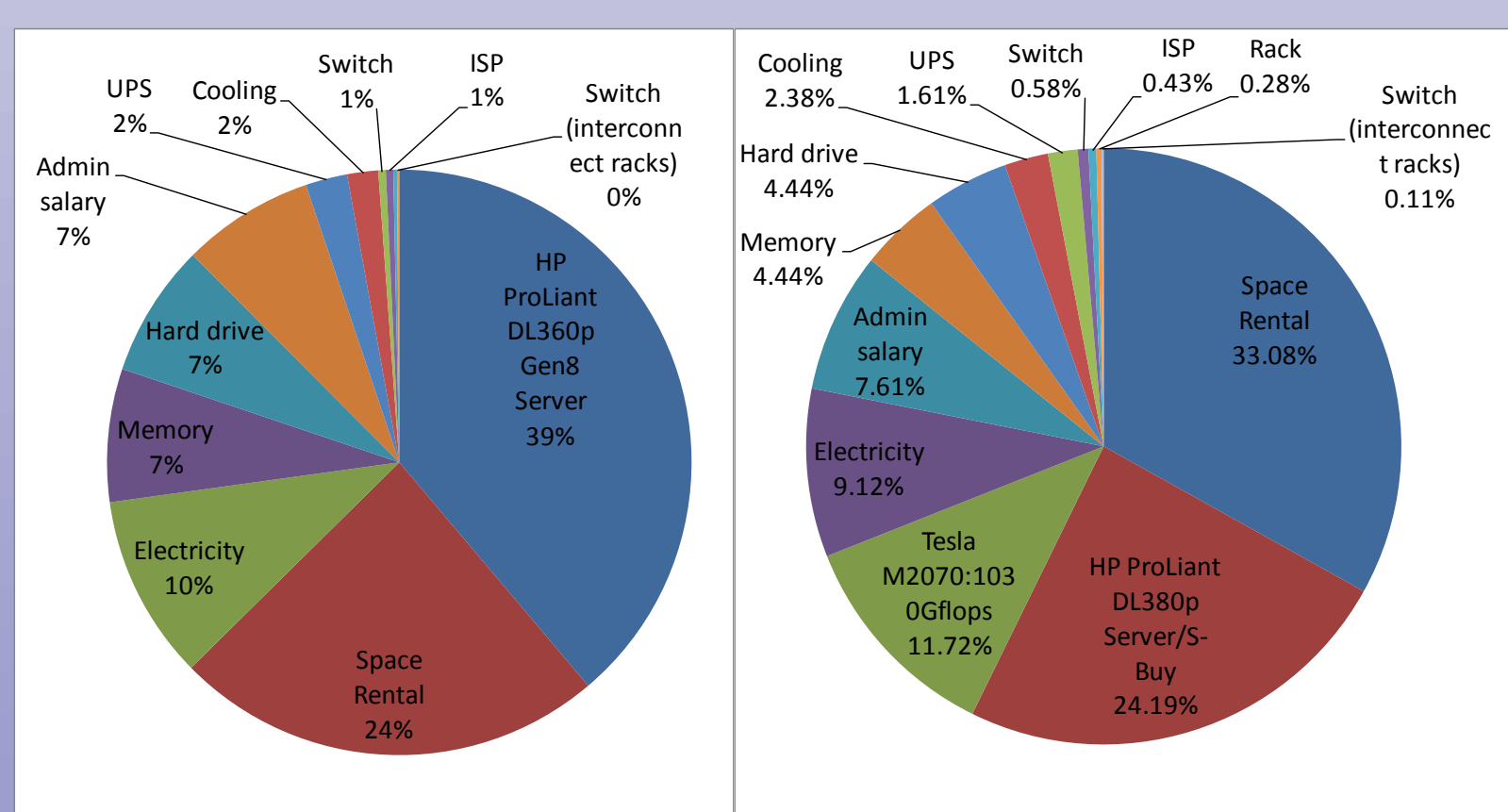
Compute Cost Comparison
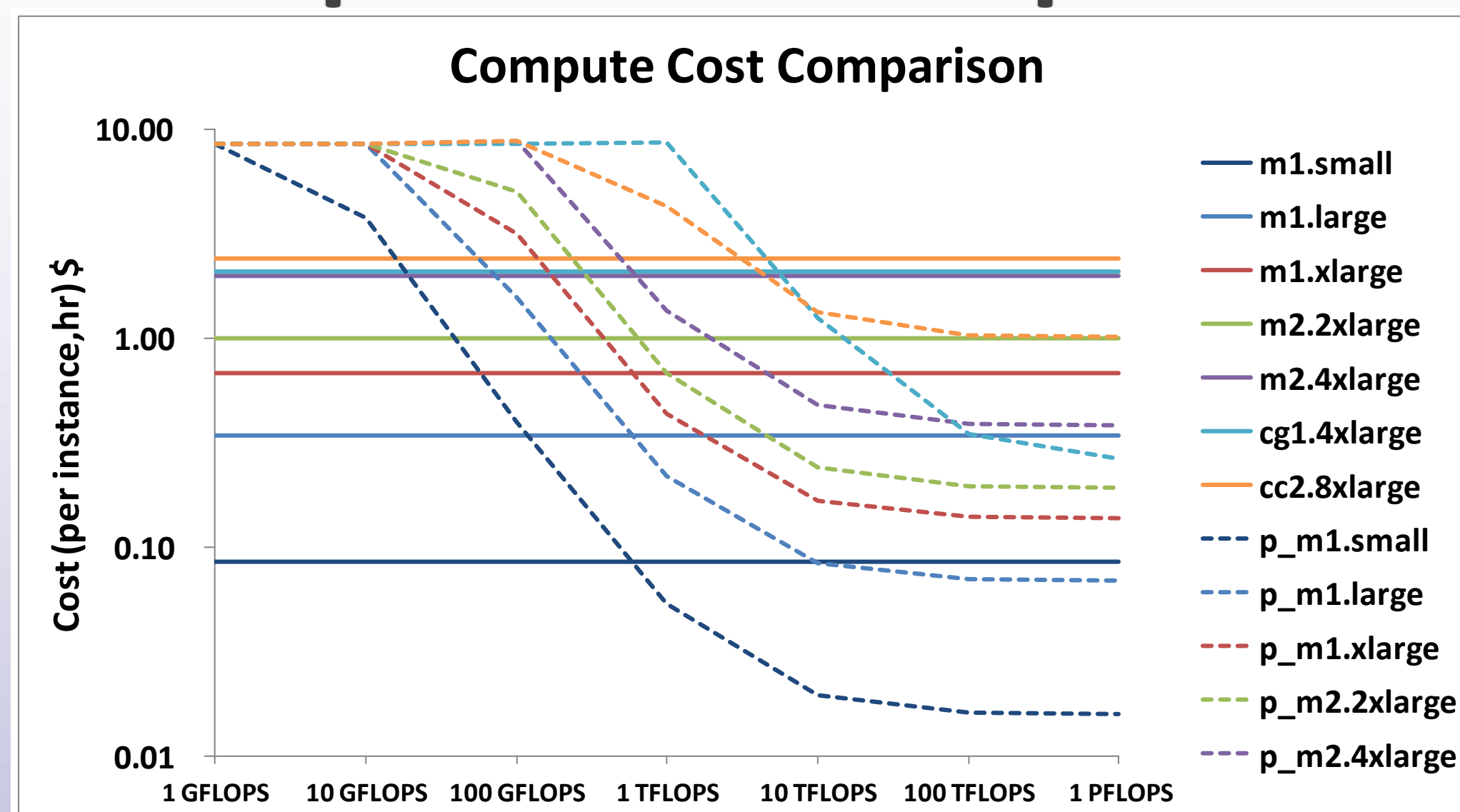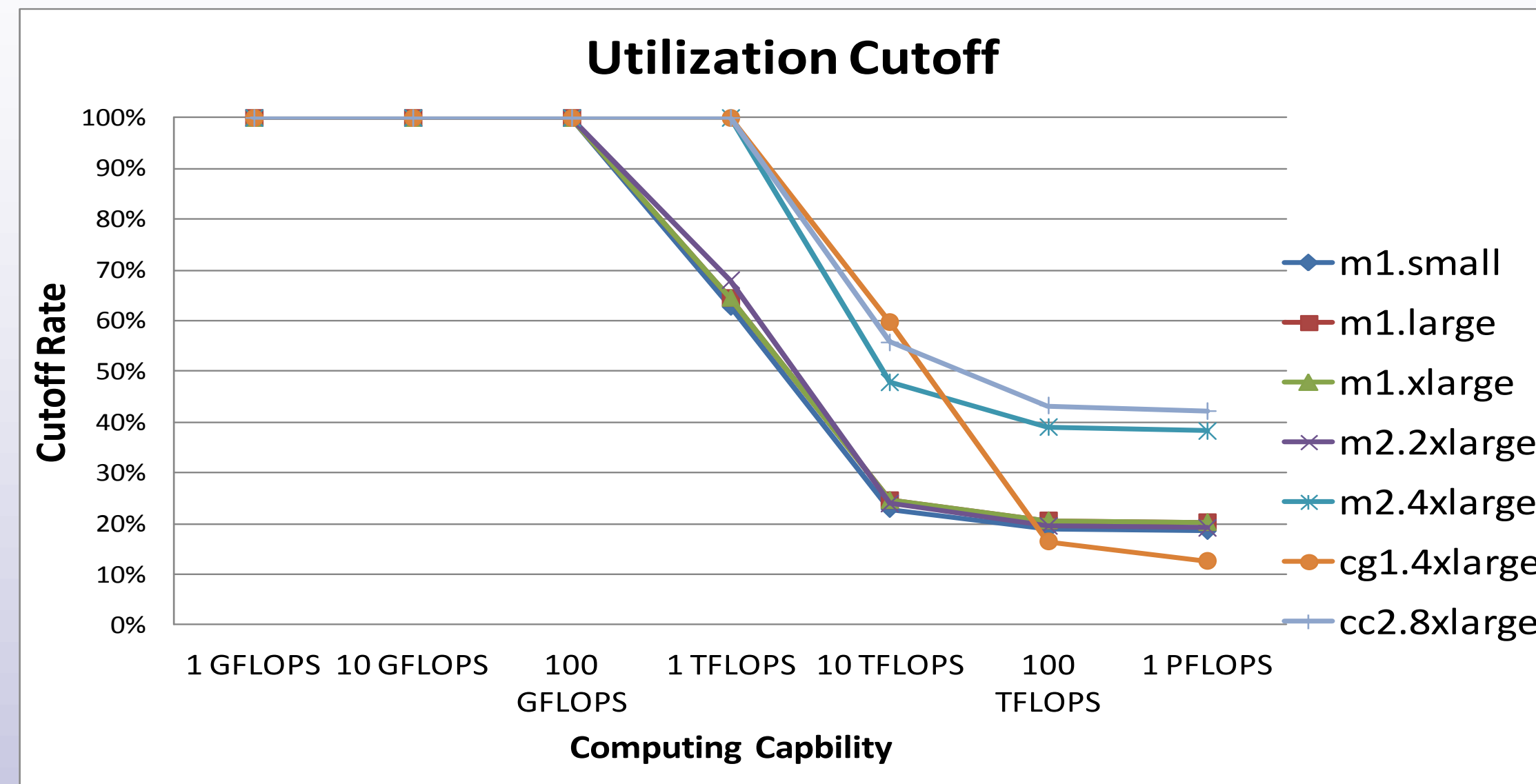
Cost Breakdown of Private Clouds Comparable to EC2 cg1.4xlarge and cc2.8xlarge

Utilization Cutoff

- EC2 is a better choice at smaller scales. At larger scales the administration costs are amortized to more compute nodes so private cloud is the right way to go.
- CPU-GPU hybrid instance (i.e. cg1.4xlarge) shows a very different trend from the CPU-only cases. It has a steeper downtrend before 100TFLOPS.
- CPU-GPU hybrid instance can almost reach 10% utilization rate for 1FLOPS.
- building a private cloud could easily beat the price of EC2 as long as he runs his tasks for more than (24 * 10%) = 2.4 hours per day.
- In GPU chart, the cost of machine space rental is the most costly part (33%), followed by the hardware purchase (24%).
- For the large instance (i.e. cc2.8xlarge), the most expensive potion is for the hardware (i.e. 39% for HP server), followed by rental as the 2nd (24%).

Conclusion

This work estimates the cost of the cloud resources. The comparison between the private and commercial cloud pricing in both storage and compute show promising results. Considering the number of potential users of such systems shows that there is a good opportunity in this area to provide such services in both compute and storage area. The results of the comparison in this project can reveal the potentials in the cloud area, and can be very useful to the new investors in this area.

Future work

The future work for this project is to perform a comprehensive performance evaluation of commercial clouds and compare them with the prototype private clouds.

The performance evaluation of the current cloud products would give a better view of the accuracy of the comparison. It also determines if the estimated bandwidth of the prototype private clouds is enough to manage large number of the users.

References

- Patel C, Shah A; "*Cost Model for Planning, Development and Operation of a Data Center*", HP Laboratories. http://www.hpl.hp.com/techreports/2005/HPL-2005-107R1.pdf
- T. Nufire. 2011 "Petabytes on a budget" from: http://blog.backblaze.com/2011/07/20/petabytes-on-a-budget-v2-0revealing-more-secrets/
- *Google Cloud Storage.* http://code.google.com/apis/storage/
- *Amazon Simple Storage Service (Amazon S3).* http://aws.amazon.com/s3/
- *Dropbox.* https://www.dropbox.com/features
- *Windows Azure Storage* http://www.microsoft.com/windowsazure/features/storage/