

# Cloud-based I/O burst buffer

October 10, 2014

## 1 Introduction

- introduce cloud computation and show the problem in data sharing through shared storage.
- compare supercomputer application execution time with cloud execution time, show how bad it is for data sensitive application.
- purpose a cloud based burst buffer to burst IO throughput between compute node and shared storage in cloud computation.

## 2 Background

### 2.1 a overview of cloud computation

introduce common cloud architecture, consist of compute node and shared storage, and shared storage system is globally distributed and connected via Internet causing low throughput.

### 2.2 a general view of burst buffer

### 2.3 introduction of some data sensitive applications

- montage
- supernoveas
- povray etc.

## **2.4 introduction of AWS (experiment environment)**

# **3 Architecture**

## **3.1 a overview of cloud based IO burstbuffer**

## **3.2 two IO patterns (just like SWoPP paper)**

# **4 Implementation**

## **4.1 a overview of implementation**

- a master manages all IOnodes info and maintain a namespace and file metadata, handle operation like INode addtion and deletion
- several IOnodes response to actually store data.
- a client connect with master to get file meta info and connect to IOnodes to transfer data

## **4.2 master**

function:

- manage IOnodes including addtion, deletion.
- manage data layout, including load balance, and data rebalance when IOnodes addtion and deletion.
- maintain namespace of buffered data.
- interact with client

## **4.3 INode**

function:

- buffer client output data.
- read data from storage.

# **5 Evaluation**

## **5.1 Evaluation of Implementation**

evaluation result published in SWoPP and first implementation evaluation result showed at meeting with amazon people.

## **5.2 Data Locality Evaluation**

montage etc. data locality results

## **5.3 Queuing model**

## **6 Related Work**

## **7 Conclusion**