

COSBench: A benchmark tool for Cloud Storage

Yaguang.wang@intel.com

Jiangang.duan@intel.com

Intel Corporation

Notices and Disclaimers

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY RELATING TO SALE AND/OR USE OF INTEL PRODUCTS, INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT, OR OTHER INTELLECTUAL PROPERTY RIGHT.

Intel may make changes to specifications, product descriptions, and plans at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined". Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product order.

All dates provided are subject to change without notice.

Intel and Intel logo are trademarks of Intel Corporation in the U.S. and other countries.

*Other names and brands may be claimed as the property of others.

Copyright © 2013, Intel Corporation. All rights reserved.

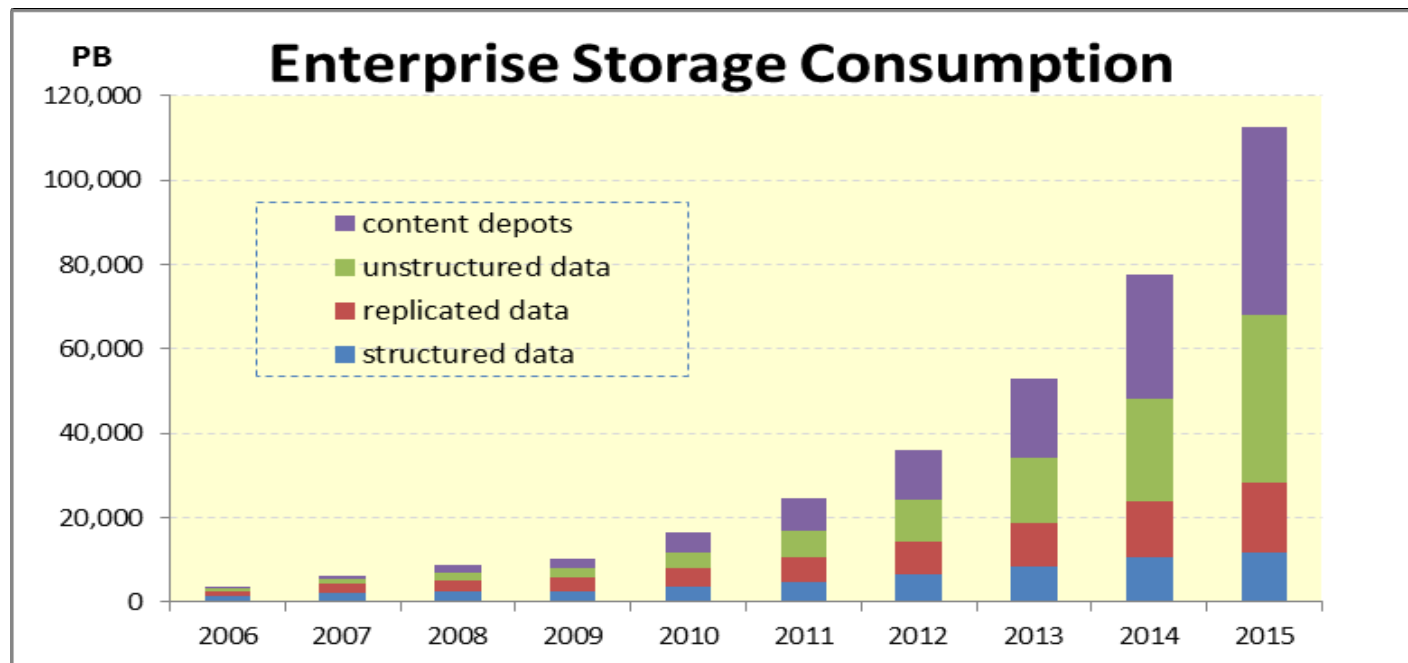
Agenda

- ❑ Why do we need COSBench?
- ❑ What is COSBench?
- ❑ How to test with COSBench?
- ❑ Call For Participation
- ❑ Q&A

Cloud Object Storage Overview

Amazon S3 goes exponential, now stores 2 trillion objects

It took six years for Amazon Web Services' Simple Storage Service, or S3, to grow to storing 1 trillion objects in June 2012. In a [blog post on Thursday](#), however — less than a year later — the company announced the service is now housing more than 2 trillion objects.

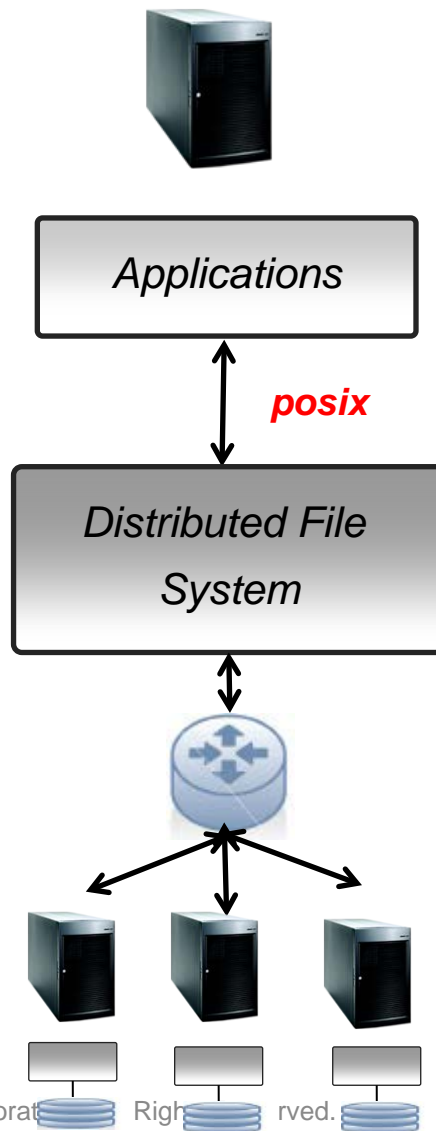
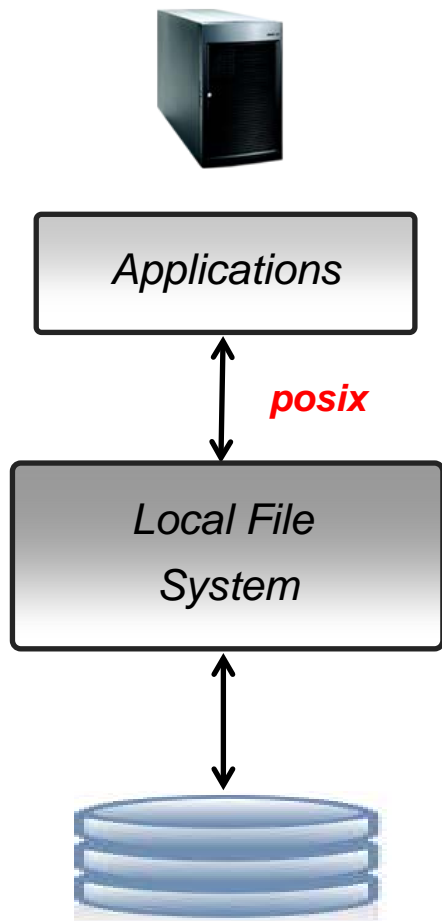


[†]Source: <http://gigaom.com/2013/04/18/amazon-s3-goes-exponential-now-stores-2-trillion-objects>

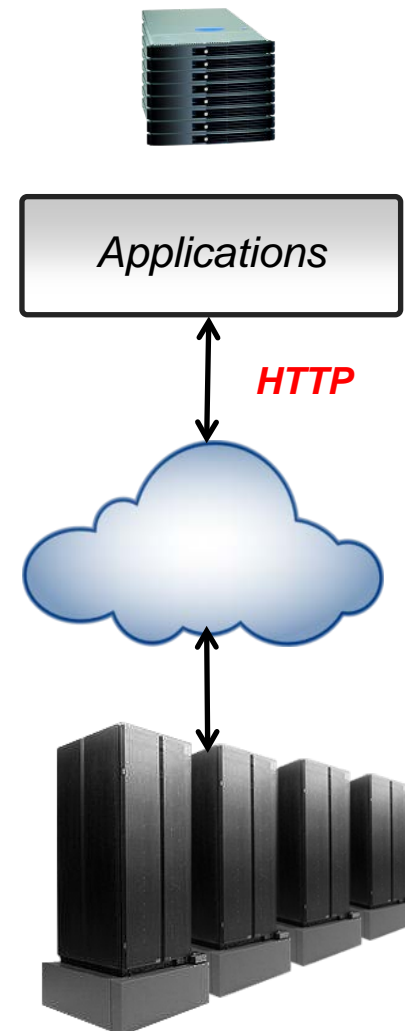
[†]Source: IDC, Worldwide Enterprise Storage Systems 2010–2014 Forecast: Recovery, Efficiency, and Digitization Shaping Customer Requirements for Storage Systems, Doc

Developer Perspective

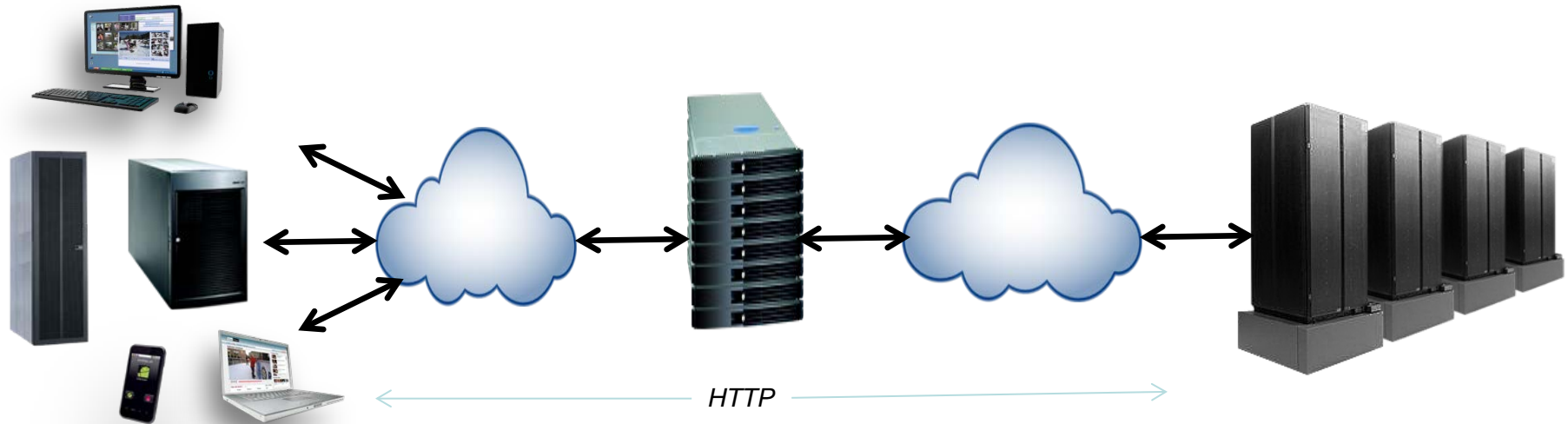
Traditional Compute Models



Service Oriented Models



Object Storage System Components



Application Interface

- RESTful
- API Variety
- PUT/GET/DELETE

Control/Proxy Nodes

- Location Mapping
- Encode/Decode
- Failure Masking

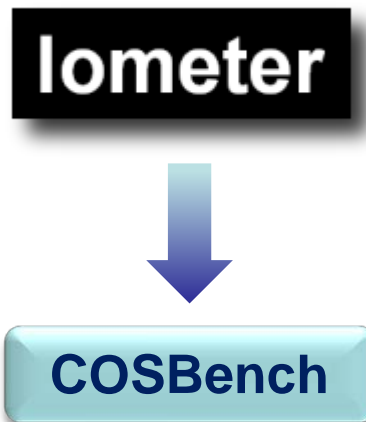
Storage Nodes

- House data
- Maintain data
- Scale Out

Object Storage: a Wide Variety of Usage Models

COSBench Overview

- ❑ COSBench is an open source **Benchmarking** tool developed by intel to measure **Cloud Object Storage** Service performance
 - ❑ For Object storage, like Amazon* S3, OpenStack* Swift.
 - ❑ Not for File system (NFS e.g) or Block Device system (EBS e.g.)
- ❑ Benefit:
 - ❑ For End User,
 - ❑ Compare public Cloud Object Storage services
 - ❑ For Cloud Service Provider,
 - ❑ Evaluate different Hardware/Software Stacks
 - ❑ Identify bottleneck and make optimization



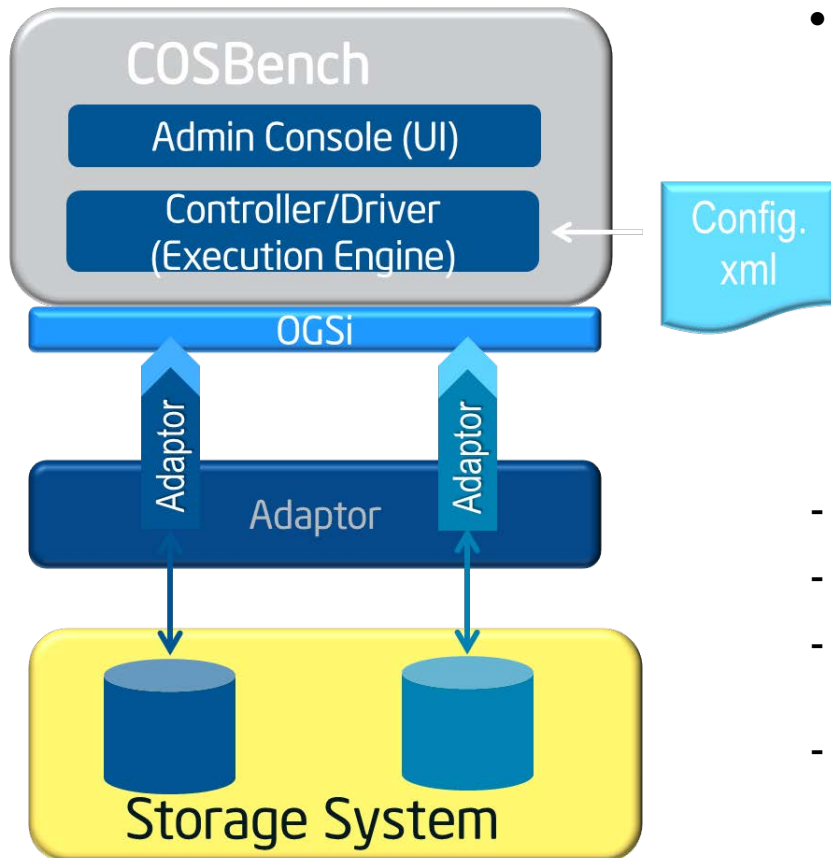
COSBench is open sourced with Apache V2 License, it is available on github: <https://github.com/intel-cloud/cosbench>.

Users



...

Major Features



- Cross Platform Deployment (Developed with java and OSGI-based)
 - Ubuntu 12.04 LTS /RedHat Enterprise Linux 6.1/Windows 7.
- Pluggable adaptors for different storage system.
 - OpenStack* Swift/Amplidata* Amplistor/Amazon* S3/Ceph
 - SNIA* CDMI (upcoming)
- Distributed load testing framework.
- Web-based real-time performance monitoring
- Flexible workload definition (Mixture of object sizes and operations, Configurable load balancing, Object integrity check)
- Rich performance metric reporting (Performance timeline, Response time histogram)
- Open Source (Apache License)

Workload Configuration

```
- <workflow>
- <workstage name="init">
  <work type="init" workers="8" config="containers=r(1,32)" />
</workstage>
- <workstage name="prepare">
  <work type="prepare" workers="8" config="containers=r(1,32);objects=r(1,50);sizes=c(64)KB" />
</workstage>
- <workstage name="main">
  <work name="main" workers="8" rampup="100" runtime="300">
    <operation type="read" ratio="80" config="containers=u(1,32);objects=u(1,50)" />
    <operation type="write" ratio="20" config="containers=u(1,32);objects=u(51,100);sizes=c(64)KB" />
  </work>
</workstage>
- <workstage name="cleanup">
  <work type="cleanup" workers="8" config="containers=r(1,32);objects=r(1,50)" />
</workstage>
- <workstage name="dispose">
  <work type="dispose" workers="8" config="containers=r(1,32)" />
</workstage>
</workflow>
</workload>
```

Load control

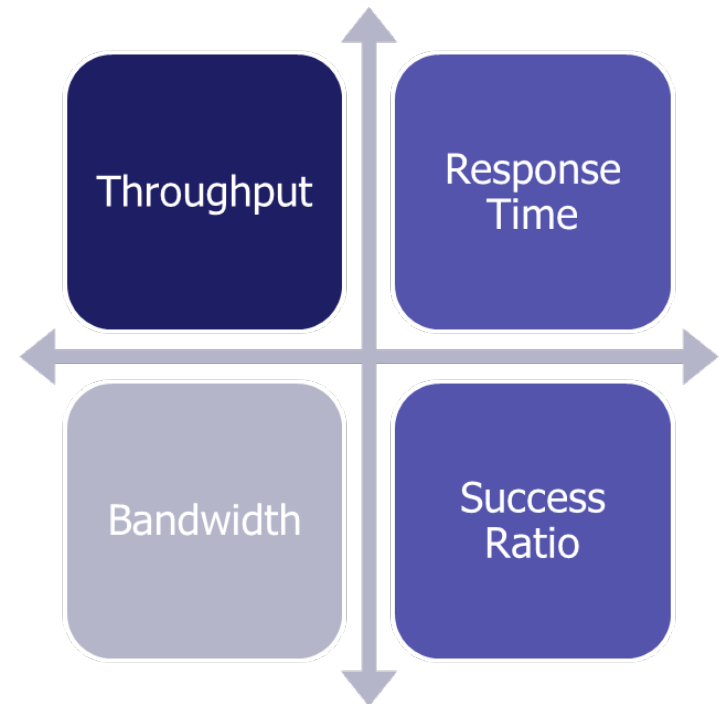
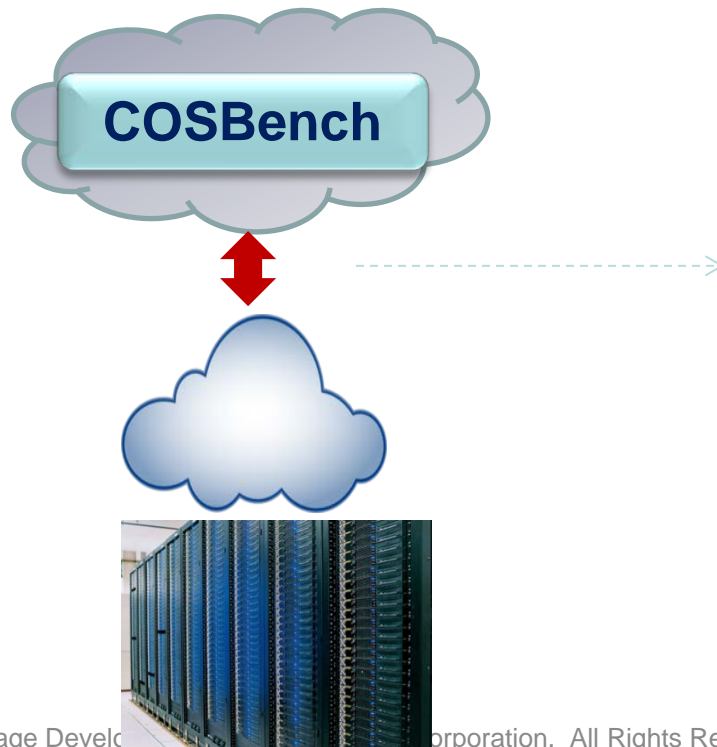
Mixture of operations

Mixture of object sizes

Workflow for stages

Performance Metrics

Op-Type	Op-Count	Byte-Count	Avg-ResTime	Throughput	Bandwidth	Succ-R
init-write	0 ops	0 B	N/A	0 op/s	0 B/S	N/A
prepare-write	10 kops	1.22 GiB	318.62 ms	321.32 op/s	40.16 MiB/S	100%
read	686.69 kops	83.82 GiB	43.67 ms	2289.17 op/s	286.15 MiB/S	100%
cleanup-delete	10 kops	0 B	41.96 ms	770.81 op/s	0 B/S	100%
dispose-delete	0 ops	0 B	N/A	0 op/s	0 B/S	N/A



Performance Reporting

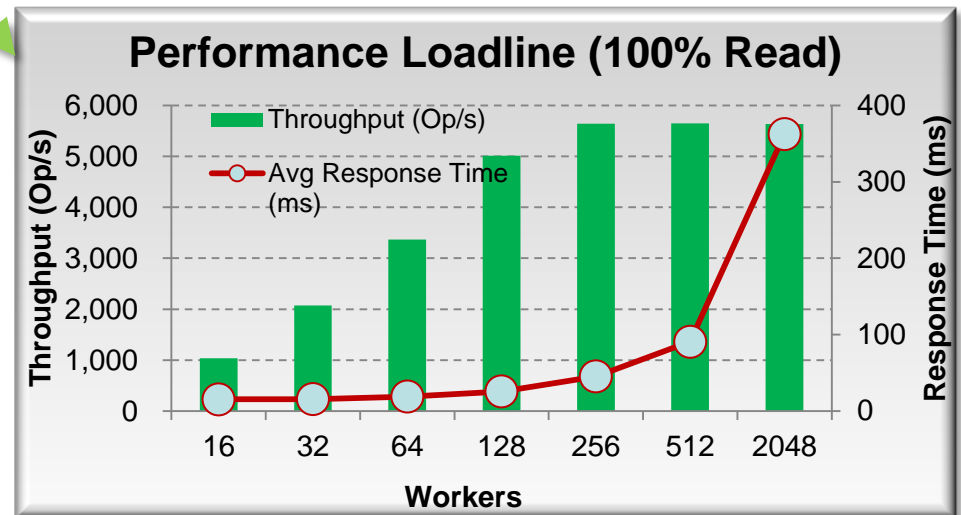
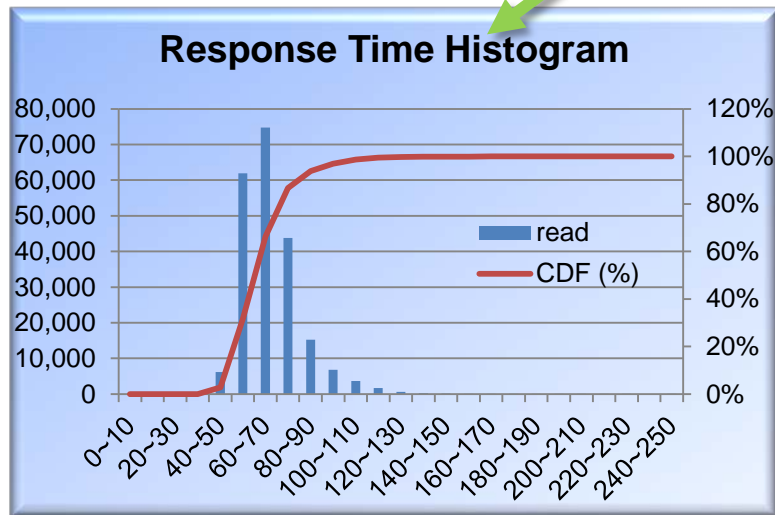
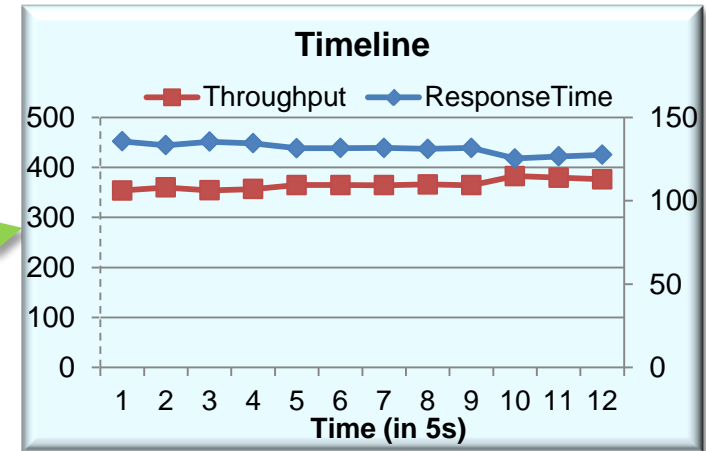
Op-Type	ResTime	Throughput	Bandwidth	Succ-Ratio
read	42.83 ms	2334.56 op/s	291.82 MiB/S	100%

summary

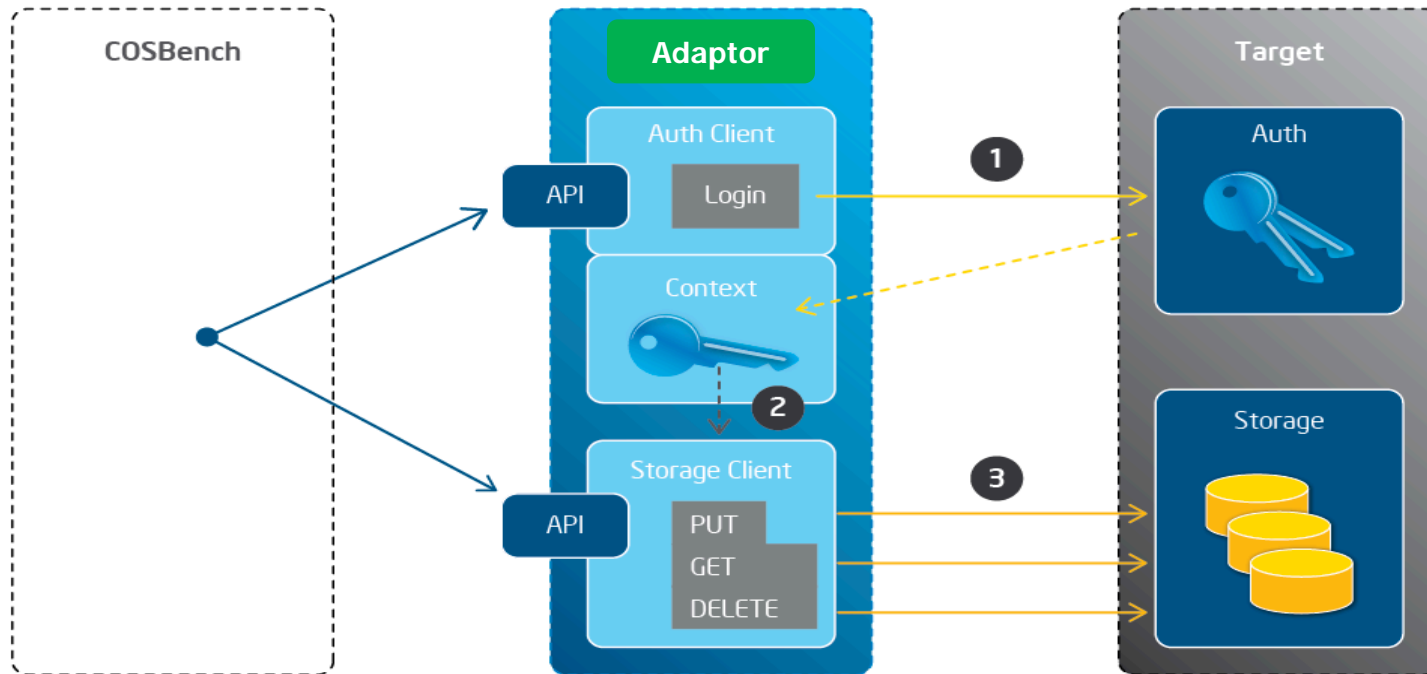
timeline

histogram

loadline



Add New Adaptor



- ❑ 1. COSBench calls Login() through an authentication client using credentials, which are passed to the target authentication server to authenticate the credentials passed in. If authentication succeeds, a corresponding token is returned and stored in authentication context.
- ❑ 2. COSBench acquires the authentication token from context.
- ❑ 3. COSBench calls PUT/GET/DELETE() through the storage client with acquired token.

Compatibility matrix (till the end of August'13)

Cloud Storage	Auth	Status
None	none	OK
Mock	mock	OK
OpenStack Swift	tempauth	OK
	swauth	OK
	keystone	OK
	direct	No support
	none	OK
Amplidata Amplistor	digest	No support
Ceph	librados	OK
	rados GW (swift)	OK
	rados GW (s3)	Not verified
Amazon S3	integrated	OK
CDMI	? (depends vendor's CDMI server)	No support

Live demo (5 min)

ID: w15 Name: demo Current State: **processing** Current Stage: prepare

Submitted At: Jun 13, 2013 2:06:04 PM Started At: Jun 13, 2013 2:06:04 PM Stopped At: N/A

[more info](#)

Snapshot

General Report

Op-Type	Op-Count	Byte-Count	Avg-ResTime	Throughput	Bandwidth	Succ-Ratio
prepare-write	494 ops	31.62 MB	10.11 ms	98.74 op/s	6.32 MB/S	100%

The snapshot was taken at 2:06:31 PM with version 2.

Stages

init --> **prepare** --> main --> cleanup --> dispose

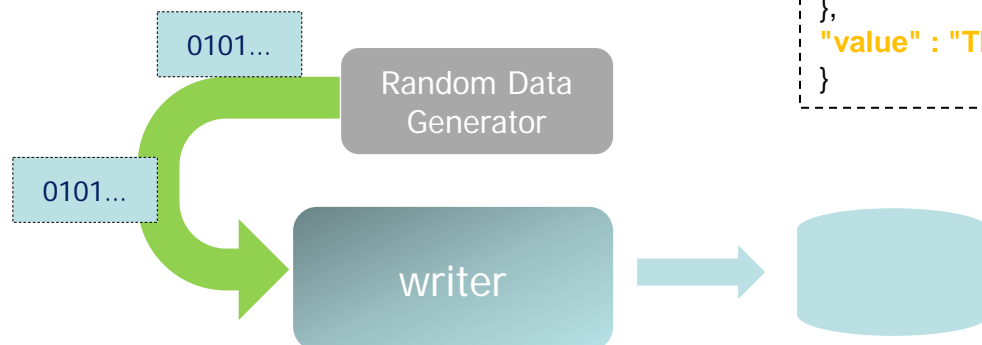
ID	Name	Works	Workers	Op-Info	State	Link
w15-s1	init	1 wks	1 wkrs	init	completed	view details
w15-s2	prepare	1 wks	1 wkrs	prepare	running	view details
w15-s3	main	1 wks	8 wkrs	read, write	waiting	view details

Intel Corporation

(Who wants to contribute the adaptor or work with us to deliver it?)

❑ Difficulties

- ❑ Json envelop requires new data generator.
- ❑ No authentication mechanism is included in CDMI standard.



† Source: <http://docs.aws.amazon.com/AmazonS3/latest/API/RESTObjectPUT.html>

* Source: cdmv1.0.2.pdf page 50

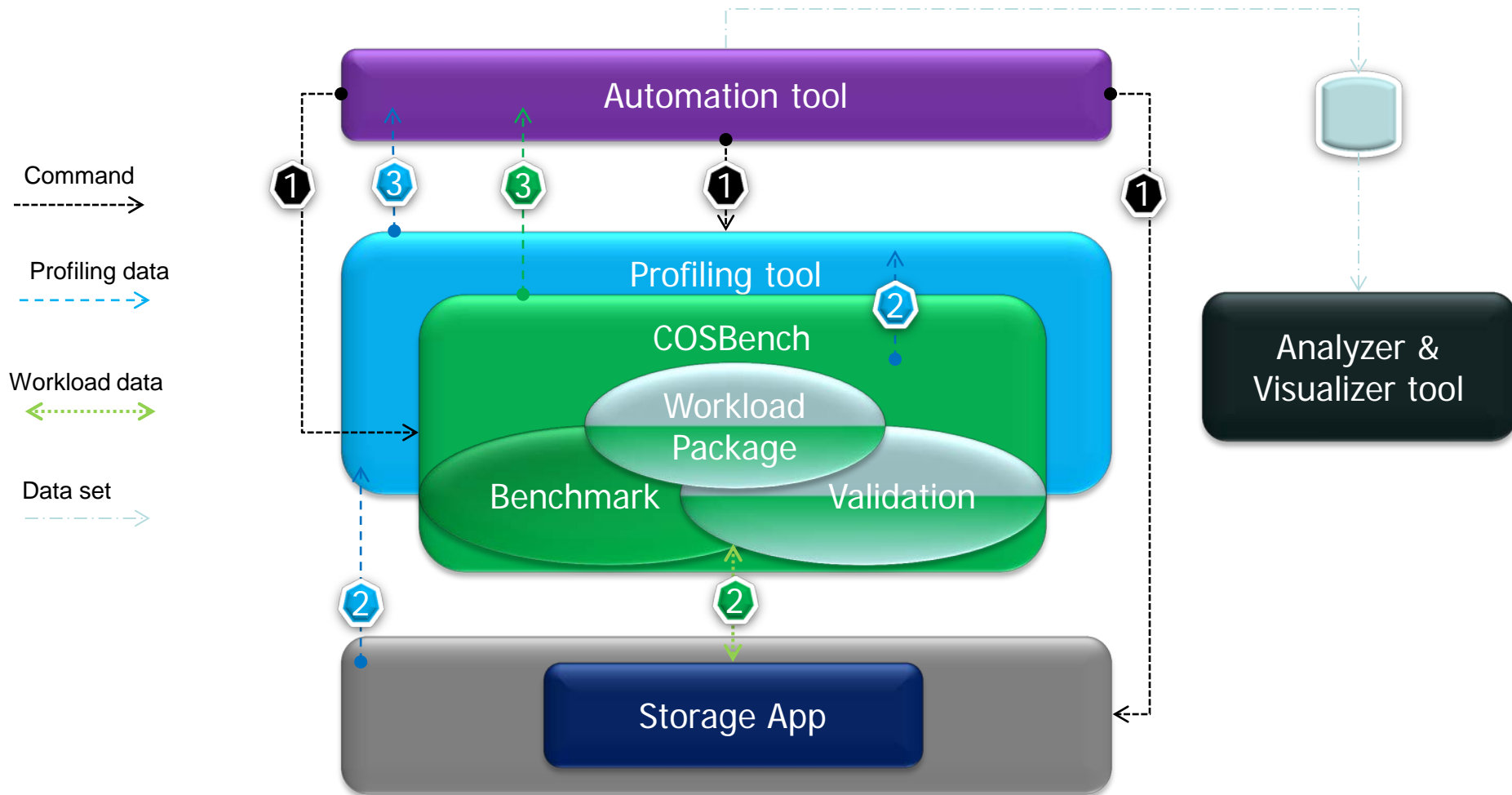
S3

```
PUT /my-image.jpg HTTP/1.1
Host: myBucket.s3.amazonaws.com
Date: Wed, 12 Oct 2009 17:50:00 GMT
Authorization: AWS
AKIAIOSFODNN7EXAMPLE:xQE0diMbLRpdf3YB+FIEXAMPLE=
Content-Type: text/plain
Content-Length: 11434
Expect: 100-continue
[11434 bytes of object data]
```

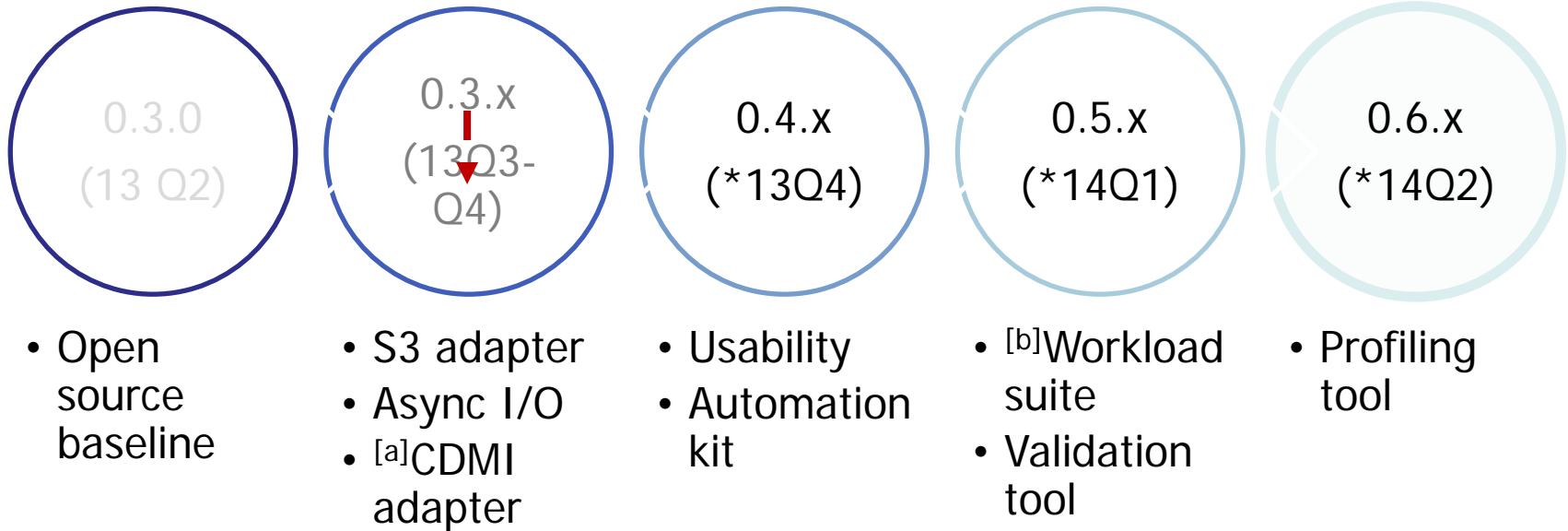
CDMI

```
PUT /MyContainer/MyDataObject.txt HTTP/1.1
Host: cloud.example.com
Accept: application/cdmi-object
Content-Type: application/cdmi-object
X-CDMI-Specification-Version: 1.0.2
{
  "mimetype" : "text/plain",
  "metadata" : {
  },
  "value" : "This is the Value of this Data Object"
}
```


More than Workload?



Roadmap



[a]: 1. No server for development and verification, who can help?

2. who wants to contribute the adaptor or work with us to deliver it?

[b]: 1. who want to contribute access pattern from different usage model?

* The time frame is roughly estimated, it may be changed if preconditions or resources change.

Call For Participation

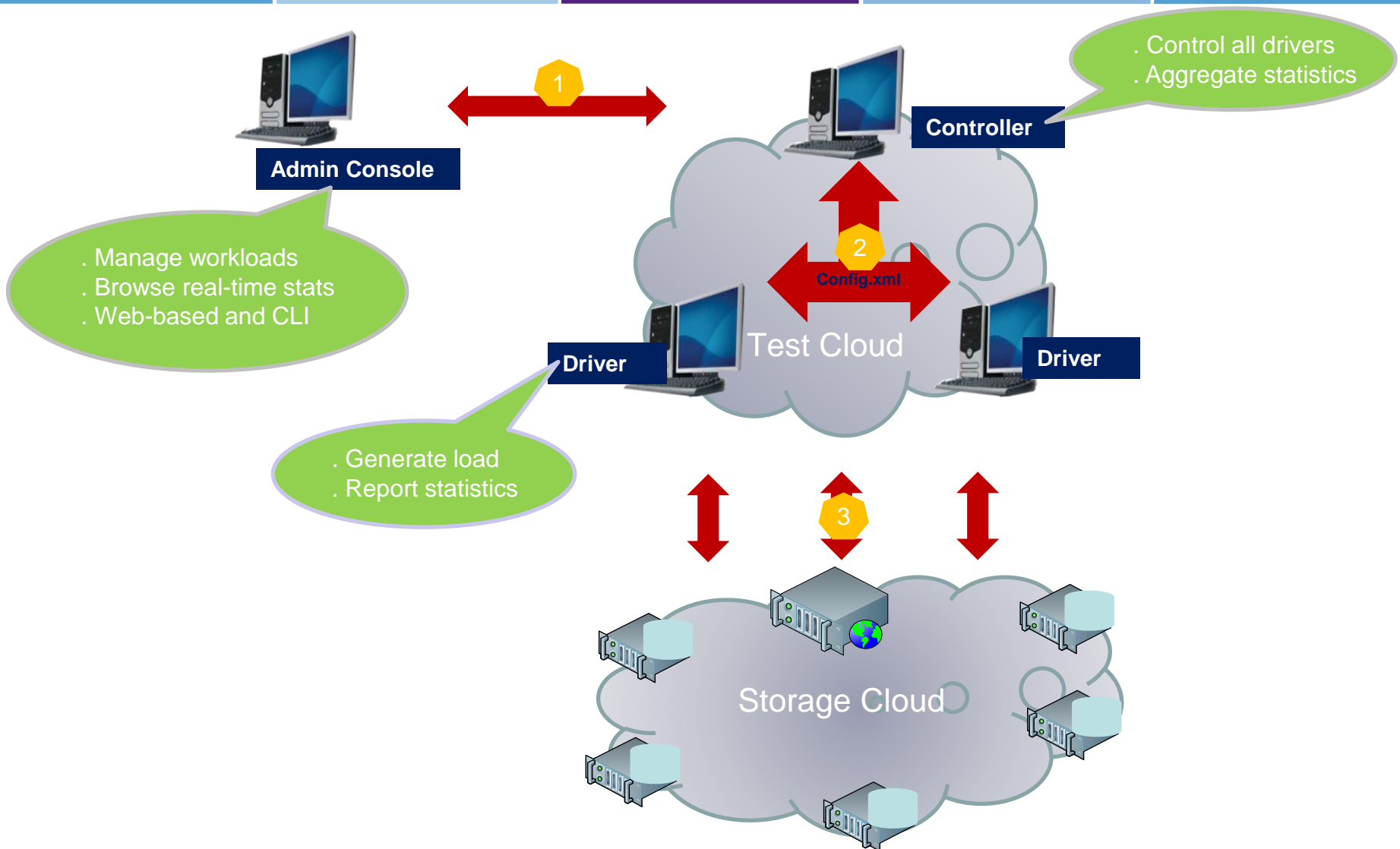
- ❑ Resource links:
 - ❑ Source code is hosted on github:
<https://github.com/intel-cloud/cosbench>,
 - ❑ One mailing list is hosted on Nabble:
<http://cosbench.1094679.n5.nabble.com/>
- ❑ We are continuing to improve it and welcome for contributions from you!
- ❑ Any questions or feedback, please contact me at: yaguang.wang@intel.com

Q&A



Backup

How to execute tests?



H/W Configuration

□ Proxy Node (1x)

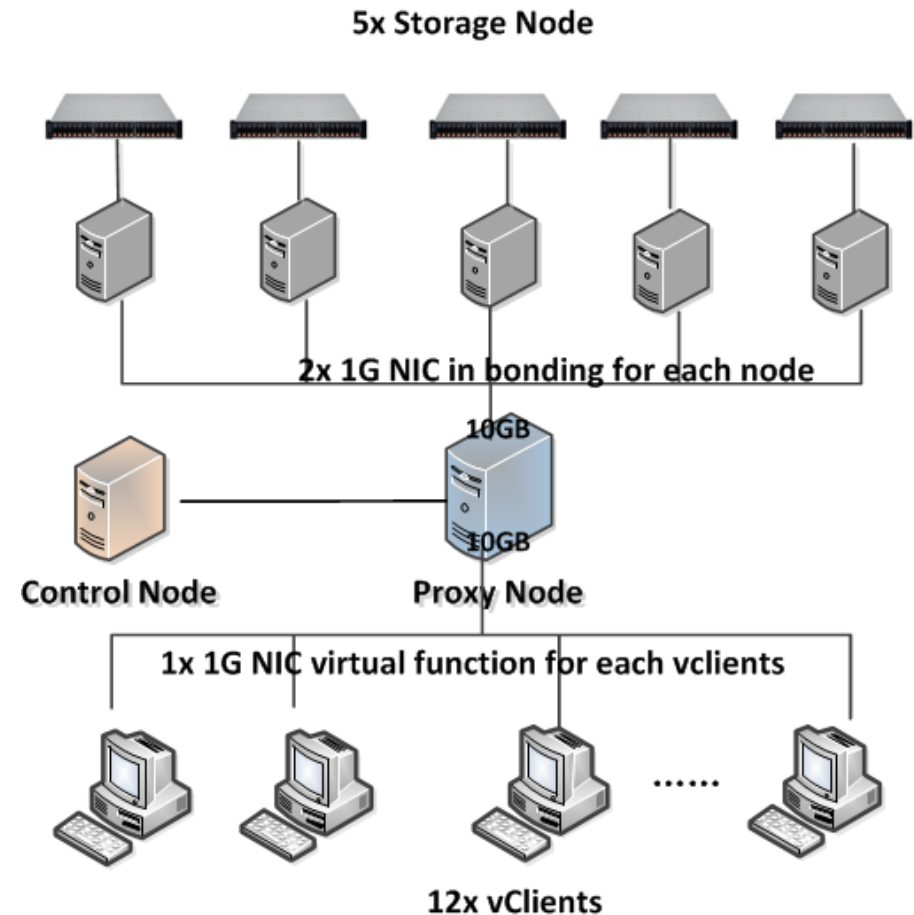
- CPU = 2 * Intel Xeon E5-2670 2.6GHz (8C/16T)
- Memory = 64GB
- NIC = Intel 82599 Dual port 10GbE
- OS = Ubuntu 12.04 (3.2.0 kernel)

□ Storage Node (5x)

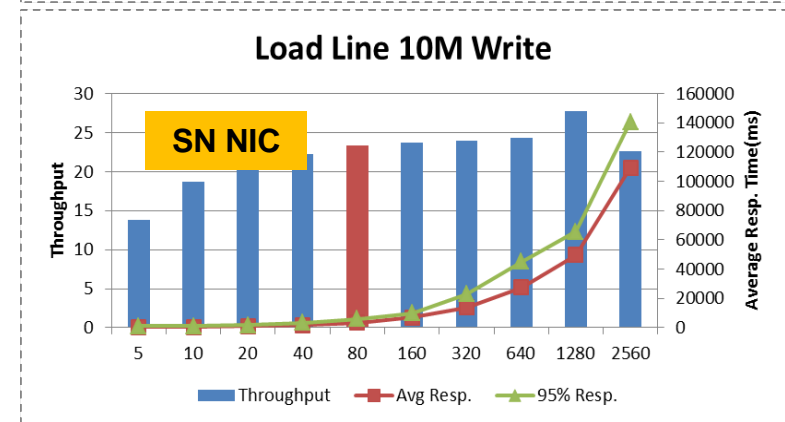
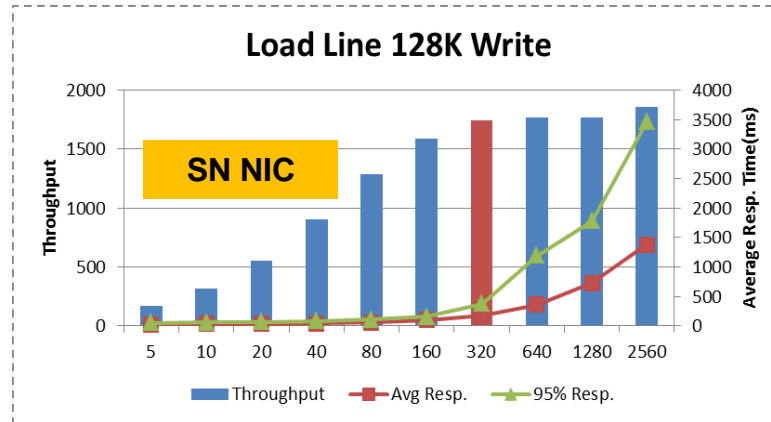
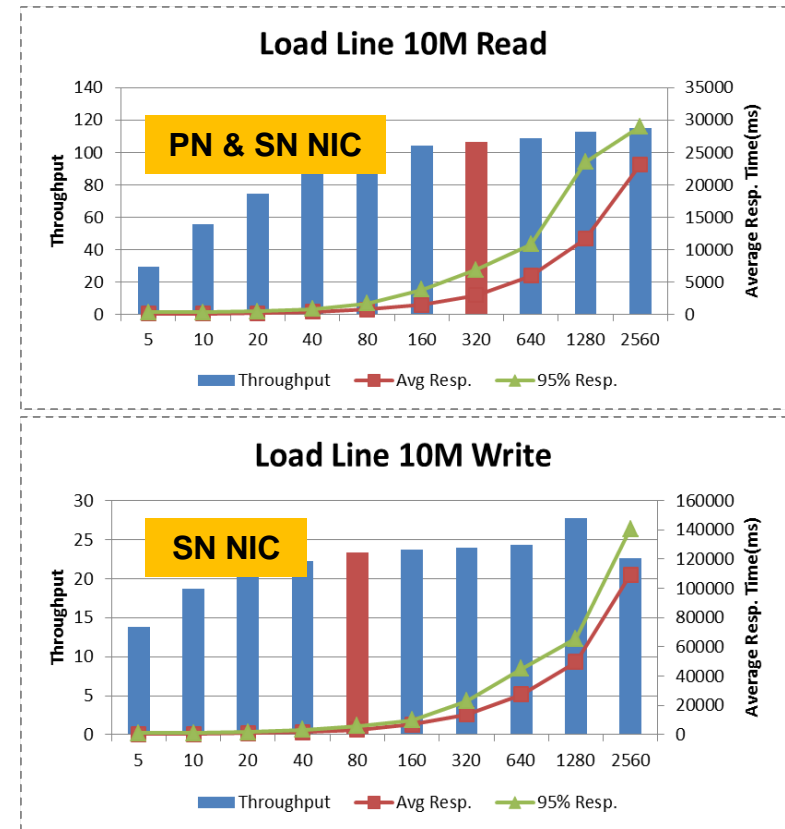
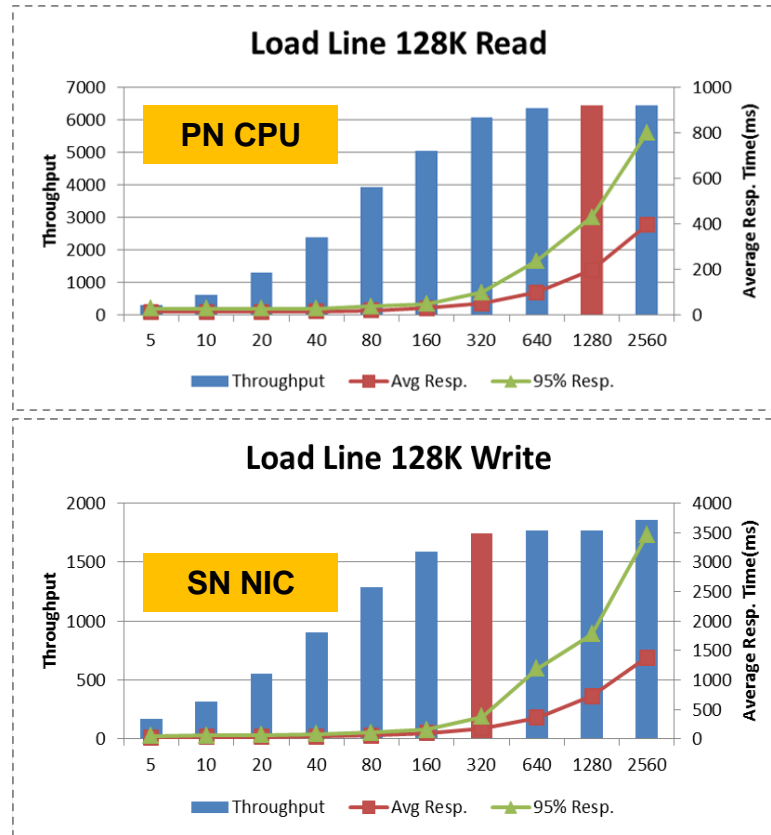
- CPU = 2 * Intel X5570 (4C/8T)
- Memory = 12 GB
- NIC = 2 * 1Gb (bonding; mode=rr)
- Disk = 12 * 70GB 15000 rpm SAS
- OS = Ubuntu 12.04 (3.2.0 kernel)

□ vClient (12x)

- CPU = 4 * Intel X5570 2.93GHz
- Memory = 10 GB
- 1 * Intel 82576 GbE virtual NIC



Performance results



□ Drive system to hit physical limitation at different cases.

* For more complete information about performance and benchmark results, visit www.intel.com/benchmarks