Assignment 2 of CISC 3018

ZHANG HUAKANG

April 9, 2021

1

1.1

$$\begin{aligned} N_{DAS} =& \frac{1Mb/s}{100kb/s} \\ =& 10.24 \\ =& 10, N_{DAS} \in \mathbb{N} \end{aligned}$$

1.2

1.2.1 How will the maximum number of the clients change

Maximum number will increase.

1.2.2 Reason

Because clients use TCP/IP to communicate with server. Based on Internet protocol, the data will be separated into small 'packets', then send the packets to client or server. This is why server can keep the communication with many clients without the limitation of connection speed between client and server.

1.2.3 Can the practical NAS support an infinite number of clients?

No, because the limitation of the machine's packet processing speed. Too many packets are received simultaneously will cause packet loss.

2

2.1

- DAS (Direct-Attached Storage)
- SAN (Storage Area Networks)
- NAS (Network-Attached Storage)

2.2

- All user connect the server directly, and connect the storage via the server in DAS. But in NAS, users can connect storage part directly and server just connect the network.
- DAS uses BLOCK I/O and SCSI protocols, while NAS uses TCP/IP and IP network.
- DAS uses dedicated link which can not be shared. So DAS is difficulty to extend stroge device. NAS uses the IP network and direct communication between storage and clients and thus it is easy to extend stroge device.
- DAS has the connection limitation but NAS does not.

3

3.1 Advantage

- Universal document access
- Enabling group sharing
- Increased flexibility and reliability
- Cost-efficient

3.2 Disadvantaged

- Internet connectivity dependent: poor preformance under a congested network.
- Security and privacy of the data

4

4.1 Map funtion

To generate a key-value pair for each subtask and obtain the intermediate files.

4.2 Reduce function

To combine the solution from the intermediate fils to generate the final solution.

4.3 Difference between MapReduce and Dryad

The main idea of MapReduce is process the subtask and organize the subtask and the result into a key-value pair while the Dryad tries to find the logical relationship among subtasks, connect them by a channel, and organize the subtasks and channels into a graph.

5

5.1 Yes

Beacuse PaaS is based on IaaS and PaaS providers will manage the servers, storage, data centers and networking resources.

5.2

The higher the abstraction-level is, the less control to the infrastructure and other basic things users has. Users have the deepest access to infrastructure configurations with low abstraction-level platforms. Users fource on middleware or softwave tasks and APIs abstracted from infrastructure with middle abstraction-level platforms. And Users get the entire technology stack with full abstraction of infrastructure.

6

6.1

- Amazon is fourcing on IaaS and PaaS mainly but Google is fourcing on SaaS.

- Amazon porvids VM from resources pool and let customer run its own OS on it while Google provides distributed stroge and caching pool, and the application based on this, such as Google Drive.

6.2

- IaaS: Cloud provider will provide infrastructure like server, networking, virtualization and storage in its physical data center. Users have to install their own OS and run application on the OS.
- PaaS: Cloud provider will provide not only infrastructure but also OS, middleware and runtime.
- SaaS: Cloud provide will provide data and applications in addition comparing to PaaS.