KTH Royal Institute of Technology

Highly Available and Robust Network Services in Under-served Areas

Jiannan Guo

Master's thesis

Supervisor: Prof. Björn Pehrson, KTH
Examiner: Prof. Markus Hidell, KTH

Stockholm, 15th September, 2014

Part I

Robust Network Infrastructure in Rural Areas of Tanzania

Part II

Highly Available Web Services

Chapter 1

Introduction

1.1 Background

Affordable, yet stable web services are highly desired in rural area, as a mean to alleviate digital divide and improve life quality. When it comes to under-developed regions in Africa, requirements and conditions need to be carefully assessed and analyzed, for that challenges could be unique and dramatically different than metropolitan.

1.2 E-learning for Open University of Tanzania

Open University of Tanzania (OUT) [1] is the first university of East Africa Region to provide open and distance learning programmes. To distribute course content through the whole country, Moodle has been chosen as underlying digital resource management platform. Moodle[**?**] is an open-source industrial-level online learning platform and resource management system. As a typical data-driven web service, Moodle runs over an underlying database and assemble its webpages on-the-fly based on user requests. It is written in PHP and heavily tested against Apache, Nginx and MySQL. At present, the platform is running as a standalone web service in a central server and mainly serve static content such as PDF, Text and Slides. Although OUT has the vision to introduce multi-media materials to enhance education quality. OUT also establishes learning centers in major cities and towns all over the whole country and is ambitious to extend to a larger scale. An emerging obstacle is to provide services in remote areas with poor network connection and bandwidth.

1.3 Problem Identification

As part of the project, we investigated local conditions and needs within the scale of Serengeti Broadband Network, especially in areas with evident

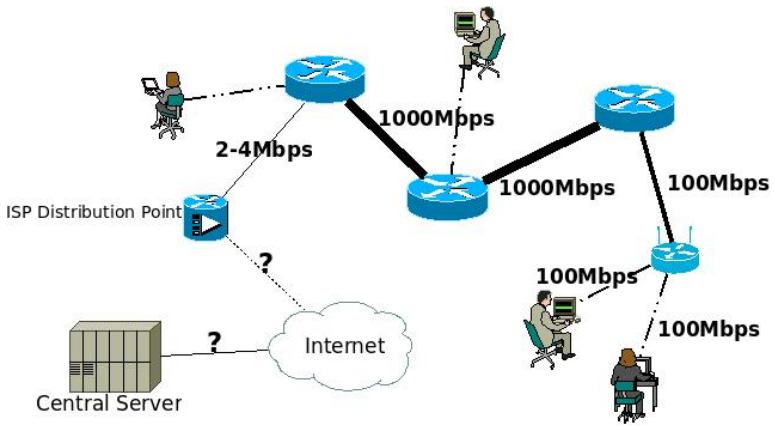

Figure 1.1: A Typical Setting of Rural Local Access Network (LAN)

demands of services and lack of infrastructures. And we were able to identify following challenges:

1.3.1 Power Outages

Power grid in rural areas of Tanzania is so unreliable that UPS for critical device is almost a must. While people are gradually adapting to mobile platforms, such as smart phones and tablets, backbone infrastructures are also required to be more persistent. Equipments powered up by solar and battery are highly desired due to cost-efficiency. Although power consumption need to be optimized in this circumstance in order to prolong battery life and improve reliability.

1.3.2 Poor network quality and frequent failure

Although local network is operated by ICT4RD project and can be fairly reliable, uplink is still depending on national-wide ISP and is somewhat unpredictable according to our observation. Network failure could occur anytime and can last for random period (several minutes to several days). Those web services that depend on a central server are apparently not accessible during the failure. On the other hand, the uplink can be very narrow due to poor infrastructure and limited budget. It could be difficult to squeeze multimedia services into such bandwidth.

To better illustrate this problem, suppose a typical setting in Figure 1.1. Major backbone components in this LAN are interconnected through fiber-optical lines, and network is distributed to users through WiFi or Ethernet. The LAN is linked to the Internet through ISP distribution link and central server resides on the otherside of the Internet.

Due to limited budget and ISP capacity, the upper link is equipped

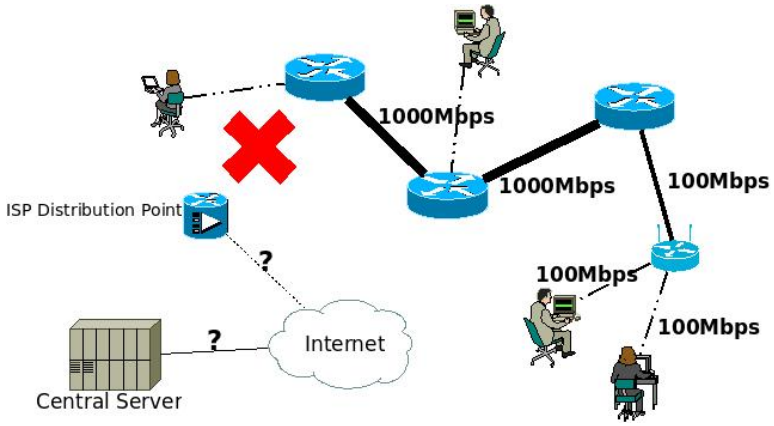

Figure 1.2: Uplink Failure leads to the isolation of LAN

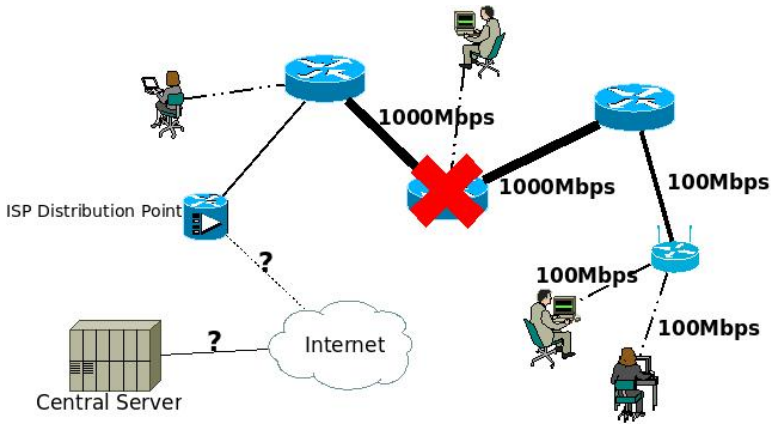

Figure 1.3: Network Separation due to Component Failure

with an average bandwidth of 2 4Mbps which is shared among all users in LAN. While a minimum bandwidth of 1.5Mbps is recommended for video streaming, it is difficult for users to get decent service from central server. To worsen the situation, the upper link is somewhat unpreditable, which leads to the isolation of LAN, as shown in 1.2.

On the other hand, components in the LAN can also break down which leads to network separation, see Figure 1.3. In the first case, multi-media content can hardly reach end users. And in other two cases, users cannot get service at all.

1.3.3 Limited budget

Cost is an essential factor during rural ICT development. Given relatively smaller user base and weaker demand, equipments need to be chosen wisely. Although future maintainence and development also need to be considered.

Chapter 2

Adapt to Frequent Network Failure and Limited Bandwidth

It is important to formalize the problem and clarify what we attempt to achieve.

Overall goals are:

- To reduce user-perceived latency and bandwidth usage within the context of rural LAN with a narrow upper link.
- To serve up-to-date content.
- To achieve partition tolerance by continuing service when network failure occurs.

It has been proven that consistency, high availability and partition tolerence are impossible to be achieved at same time[**?**][**?**], however weak consistency is acceptable in our case.

Since OUT online learning platform is already running in production phase, it is desired to impose minimum modifications to existing software stack.

2.1 A closer look at Moodle

As introduced in section 1.2, Moodle is deployed as underlying course management system for OUT E-learning platform. Moodle is an open sourse project written in PHP and well-documented[**?**][**?**]. Similiar to other web applications, it can be deployed in a typical LAMP or LNMP stack. In this chapter, we mainly focus on possible solutions for two problems stated previously, and leave the choice of actual server to chapter 3

Moodle is a typical database-driven web application where all the pages are generated on-the-fly based on user request. The whole application is composed of three main components:

- PHP source code, typically in `/var/www/moodle/`
- A database to store data or metadata including site configuration, student information, course details, events, etc.
- A directory to store materials and resources, as well as cache and temparory files. Typically it is named as `moodledata/`

As an online learning platform, interactive sessions cannot be neglected such as forum, quiz and user-generated blogs. Hence, modifications need to be preserved.

2.2 Web Caching and Prefetching

An intuitive and common solution for the problem of limited bandwidth is to cache popular web content locally, as illustrated in Figure 2.1. Web caching has been proven as an effective approach to reduce bandwidth usage, user-perceived latency and loads on original server[**?**].

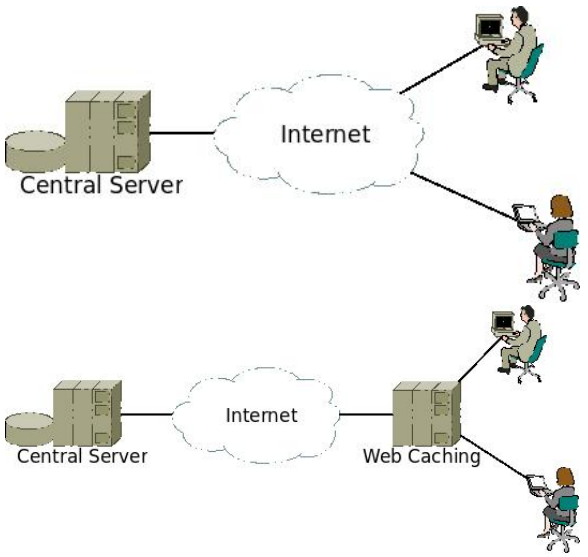

Figure 2.1: Serve users without and with a Gateway Cache

A proxy cache is normally a shared cache deployed in user network by administrator. It fetches Internet objects on the behalf of user and stores them locally. When next request for the same object arrives at proxy server, it responds with local copy and in turn reduces latency and saves bandwidth.

Although, two disadvantages are found when investigated against our case:

- Read-only Cache Web cache proxy processes READ requests from users, whereas WRITE requests still go to original servers. Even though WRITE request can be evidently less than READ, they should still be satisfied during disconnections.
- Cache Miss Web cache proxy aims at accelarating web service, rather than replicating it. When cache miss occurs, users simply cannot access the service.

2.3 Content Delivery Network

A Content Delivery Network is a collaborative set of surrogate servers spanning the network, where web contents are mirrored[**?**]. Users will perceive a smaller latency while fetching content from a nearby CDN surrogate server rather than original web server. The essence of CDN is illustrated in figure X

Since more and more web services are evolving to provide dynamic content, CDN also takes advantages of cachebility hints when dealing with dynamic contents[**?**].

2.4 Network File System

A more brute-force approach is to replicate all three main components in section 2.1 at local server.

2.4.1 Database Cluster

2.4.2 Database Replication and Synchronization

2.5 Multi-Master Database Synchronization

2.5.1 Concurrency Control Protocol

2.5.2 Operational Transformation

Chapter 3

Low Power, yet Powerful

3.1 Power Reduction on Servers

3.2 Benchmark of Server Ability

3.2.1 Nginx or Apache

3.2.2 Processor

3.2.3 Find the bottleneck

3.3 Small-scale Cluster

Chapter 4

Conclusion and Future Work

4.1 Conclusion

4.2 Future Work

Bibliography

[1] http://www.out.ac.tz/. [Online]. Available: http://www.out.ac.tz/