Accounting system for heterogeneous IP-networks (IPNA) implemented at Kaiserslautern University

Brian Worden[1], Claudia Baltes[2], Inga Scheler[3], Paul Müller[4] and Hans Hagen[5]

[1] University of Kaiserslautern, RHRK
G.34 R. 224, P.O. Box 3049, 67653 Kaiserslautern
worden@rhrk.uni-kl.de
[2] fgn GmbH
Gottlieb-Daimler-Strasse, G.38, 67663 Kaiserslautern
baltes@fg-networking.de
[3] University of Kaiserslautern, RHRK
G.34 R. 227, P.O. Box 3049, 67653 Kaiserslautern
scheler@rhrk.uni-kl.de
[4] University of Kaiserslautern, Department of integrated communications systems
G.34 R. 324, P.O. Box 3049, 67653 Kaiserslautern
pmueller@informatik.uni-kl.de
[5] University of Kaiserslautern, Department of Computer Graphics
G.36 R. 226, P.O. Box 3049, 67653 Kaiserslautern
hagen@informatik.uni-kl.de

Abstract. This paper describes an accounting system (IPNA) for heterogenous IP-networks with arbitrary topologies implemented at the university of Kaiserslautern. The produced data volume per unit is numerated. The collected data is stored in a database and offers different analysis possibilities. The results can be visualized and adapted to the users requirements.
The main effort was to build a data traffic quota system for single units as well as groups of devices that also report exceeded quotas. The system itself only observes the network traffic. Interfaces offer tools to interact with the network. The IPNA consists of a back-end for the data-acquisition and -preparation and a front-end for configuration and visualization tasks including quality control.

Keywords. Accounting system, IP-network, Communication, information visualization, online quality control

1 Introduction

An overview of the utilization within a communications network is necessary to display and control user behaviour and to analyze and maintain the network structure. Additionally, the increasing number of real-time measured data on the central network systems requires the need of a new accounting system. Accounting systems are used to protocol network traffic between internal and external

Dagstuhl Seminar Proceedings 09211
Visualization and Monitoring of Network Traffic
http://drops.dagstuhl.de/opus/volltexte/2009/2155

network connections. Goal of our work was to develop an accounting system that is now utilized to monitor the complex network in the RHRK at the university of Kaiserslautern. Contrary to the old system we can define the amount of data independent of the topology of the heterogenous network.
Primarily we define the processes and the census in arbitrary networks. The collected data is stored in a database. This data is then analyzed to provide different kinds of overviews. Within these overviews the IP-address is allocated to the DNS or user-name to allow a network view as well as a user-specific view. The main focus within our work was the configuration and analysis of a user-quota with multiple rules that occurs overlapping and conflicts that can be recognized and solved by using the system. Different interfaces allow the use of software that provides more tools to administrate the network [**?**]. The developed accounting system serves as a online quality control system of the network at Kaiserslautern university.

2 Data collection

In this section we describe the data and the way it is collected to build the accounting system. To identify the data we first extract the existing network model. As already mentioned, the main goal of the accounting system is to measure data-connection volumes between internal and external networks. Therefore we have to define the data-volume of each connection and send this information to the main node, the kernel of the system. We distinguish between internal connections and all other connections. Internal connections are set to zero because they are not taken into account. The connections of external nodes are also out of interest because the kernel node does not reach the single nodes of an external network.

2.1 Network model

To start the data acquisition the internal network and the way single connections are allocated is defined. For that purpose all active network components like computers, switches, routers and so on have to be arranged in a graph. The edges between the nodes result from the physical connections between the nodes, i.e. network cable. An internal network arises from all components within a group and describes a partial graph. All edges of the internal network compose the internal route. nodes of the partial graph that are linked by an edge to a node that is not part of the internal network are boundary nodes and the edges boundary edges. All data connections between nodes of the partial graph and external nodes are of interest for counting. For every internal network in the whole network we define a single logical counting point called gate. This point represents, independent of the number of boundary nodes, the point carrying all external connections of a network. To catch all external connections it is sufficient to setup gates on boundary nodes, so called gate-devices. These nodes must be able to read and handle IP-packets and send this information to the kernel [**?**].

2.2 Configuration of the gate devices

The gate device needs the information about all internal and external routes (see Fig. **??**).

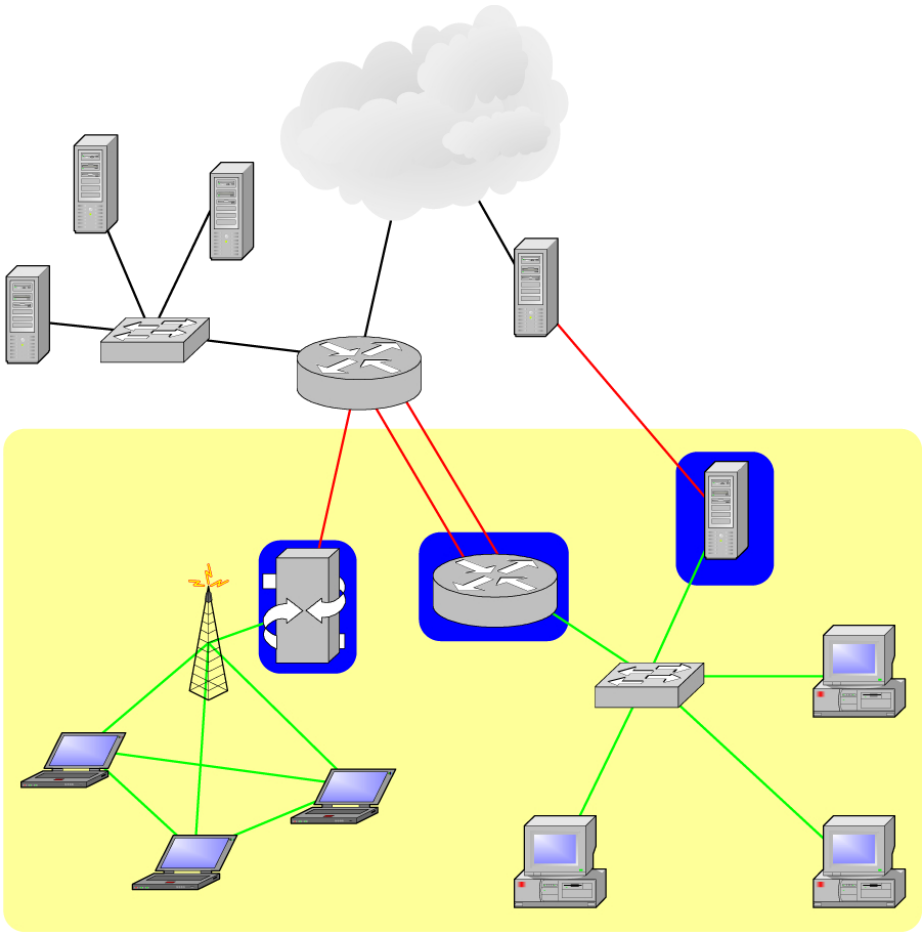

Fig. 1. network graph: internal network (yellow), internal nodes, boundary nodes (blue), internal routes (green) and external routes (red)

With this definition the decision about internal and external connections can be handled with regard to the following points:

- A packet coming in by an internal interface and going out by an external one is an **outbound external connection**
- A packet coming in by an external interface and going out by an internal one is an **inbound external connection**

– All other cases can be ignored.

There exists one exclusion. Proxyservers are usually used to reduce the system load of an internal network. They handle as Cache identical requests. Only the first request provokes a data load. This data is buffered for additional requests. This cache has to be taken into account as additional gates. They send the information about external traffic on behalf of an internal node to the kernel.

– Cache-Miss: The inquiry of an internal node occurs a data load of an external node. The incoming data volume together with the information about the query is send to the kernel.
– Cache-Hit: The inquiry of an internal node is stored in the cache. The is no additional data traffic and the gate does not inform the kernel about the event.
– All other cases don't count.

2.3 Data format

The kernel must get all information by the gate devices. Therefore we need a consistent data format that is used by all units to store the data. We spread text files for every device with the information about the monitored connection within a specific period of time. After this period the file is closed and enables the work with data. The main information regarding a connection is:

– timestamp
– source-interface
– source-IP-address: interface, the packet entered the gate-device
– end-interface: interface, the packet left the gate-device
– end-IP-address: receives packets
– Octets: Number of Bytes accumulated in the flow

2.4 Measurement method

Measurement methods for the system provide the above described information. The measurement technique and the transfer to the kernel are not predefined. Measurement methods are combinations of Network sniffers like tcpdump or wireshark. and scripts that analyze the data and send the relevant data to the kernel. The text files are sorted in the kernel and stored in different directories per device.

2.5 Data processing

The storage of the text files in the kernel allows the direct data processing. As opposed to existing accounting systems our system stores the data in a database in order to support the analysis and to reduce memory. For all gates the connections and the users are stored in a table to control the gates. It is essential to have a unique allocation of data volume and IP-address. Therefore it is not possible to change the IP-address because it is linked to the mac-address and entered in the ACL-lists on the gates.

3 Data analysis

Subsequent to the storage of the data in a database different analysis can be arranged. The main aspect of the analysis is the extraction of data volumes per user over the time as statistical analysis. Additionally we filter network users producing too much data in one hour. They exceed the predefined limit. This event is called Netburst and the users get an e-mail information about this event. Besides this we periodically produce overviews, i.e. top-N-lists and invoices. These lists can be visualized as shown in Fig. **??**. The raw data as well as the database entry are shown in Fig. **??** and Fig. **??**. The raw NetFlow data is useless to the user.

```
1251707399    517   2    <srcip> 21    7    <dstip>
0    6    51134    80    5    901

the fields:
time engine input srcip srcmask snmp dstip dstmask proto sport dport inp oct
```

Fig. 2. raw data: network quota

database entry:

client	addr	date	incoming	cache	outgoing
<hostname_or_AAA-User>	<ip>	2009-08-31 10:00:00	1051702	0	815538

Fig. 3. database: network quota

For the moment it is a trivial visualization but sufficient to give the result to the users. The visualization work is still in progress towards intuitive interfaces by using different information visualization methods and a quality control system.
To control the efficiency of a network we installed a so called quota on the system. The users get a limited access within a specific time frame and get informed about the violation of the limit. All processes running on the system are permanently controlled to secure proper network usage. At the University of Kaiserslautern we use the program Nagios to monitor the condition of our network [**?**].

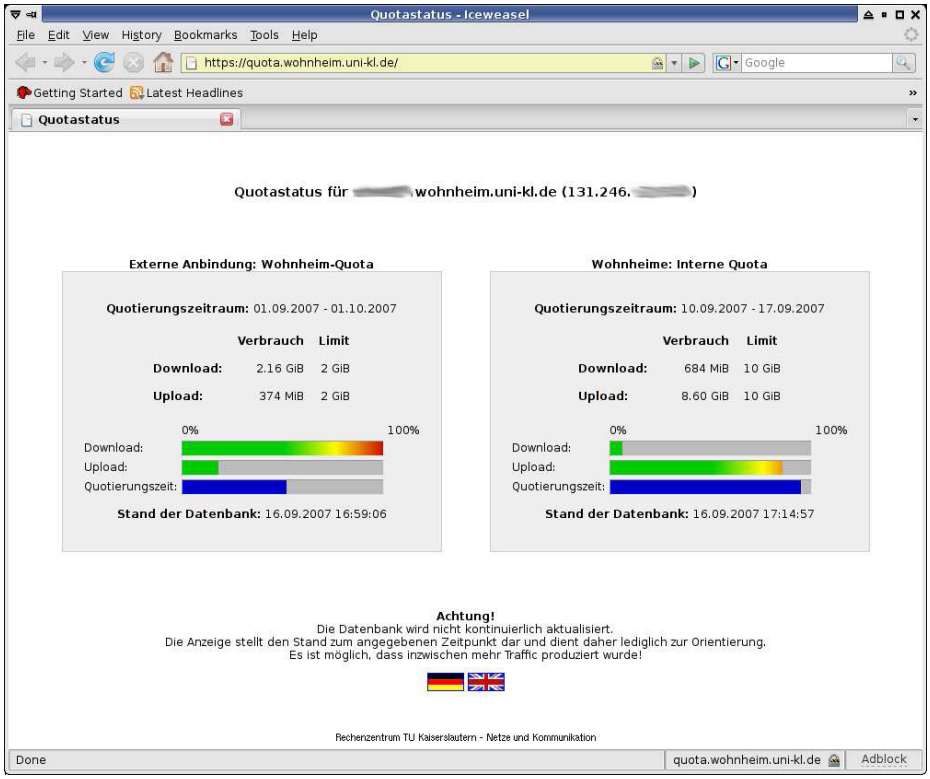

Fig. 4. visualization: overview of a network quota

4 Conclusion

We developed an accounting system that allows to monitor the usage of a network per FQDN (Fully Qualified Domain Name). One goal for the accounting system was to include devices of different manufacturers in a heterogeneous network independent of their topology to define the data volume per user. The monitored data had to be prepared in a database to ensure the analysis. The analysis leads to an overview of the network traffic per user including a monthly bill about used traffic. Additionally a user quota Is implemented as part of the system.

5 Acknowledgement

The authors wish to thank Tobias Föhst for all his work within the project

References

1. Paul Mueller, C.B., Koppen, P.: Self management in heterogeneous networks using a service-oriented architecture. In: IEEE CCNC 2007

4th IEEE Consumer Communications and Networking Conference. (2006) http://dspace.icsy.de/handle/123456789/185
2. for Standardization, I.O.: Information Processing Systems Open Systems Interconnection (OSI) Basic Reference Model. http://standards.iso.org/ittf/PubliclyAvailableStandards/s020269 ISO IEC 7498-1 1994(E).zip (1994)
3. 802.1q:, I.: Vlan. http://www.ieee802.org/1/pages/802.1Q.html (2004)