# ABCs of Grid Network Workload modeling

Author: Sandeep Sharma, Mercy Prasanna, Siva Prasad Angadi, Lakshmi Panat, Vasant Avaghade
Email: {ssharma,prasannam,asiva,lakshmip,vasant}@cdac.in

## Abstract:

***What differentiates grid networks from more traditional networks is the kind of applications and the traffic distribution profiles used by them. These applications and traffic distribution profiles make up the network workload model and grid network workload model is the heart of every grid network simulation experiment. This poster demystifies the methodology of deriving the workload model, by describing its components and then correlating them. This poster will help the readers in posing the right questions to different stakeholders of grid network (e.g. implementers/users) in order to gather inputs for creating a realistic load for grid network simulations.***

## 1. INTRODUCTION

The dictionary defines workload as "the amount of work assigned to, or done by, a worker or unit of workers in a given time period" (The *American Heritage Dictionary, 2nd Edition*). Within the confines of a network, workload is the amount of **work assigned to, or done** by, a client, workgroup, server, or inter network in a given **time period**. Therefore, workload characterization is the science that observes, identifies and explains the phenomena of work in a manner that simplifies your understanding of how the network is being used. Modeling workload plays a crucial role in simulation and can affect results and conclusions drawn from a simulation [12][15].

This poster write-up is divided into five sections. Section 2, describes the two components: Grid-Application-network traffic and profile/patterns.. It talks about the Workload-Traffic-Modeling techniques. It also elaborates the recommended distribution required for different application traffic models. It further discussed usage pattern/profile of various network applications. Section 3, explores the grid applications deployed in other grids and how those can be incorporated in Garuda traffic workload model. This document also incorporates a sample workload profile that is required for Network-Workload-modeling of grid users, so that it can be precisely modeled for simulation purposes. Conclusion and references are described in section 4 and section 5.

## 2. Grid-Network-Workload-Modeling

Grid-Network-traffic-workload modeling involves two main components viz. Grid-application-network-traffic and the pattern of traffic generation over a period of time. (Fig 1.)

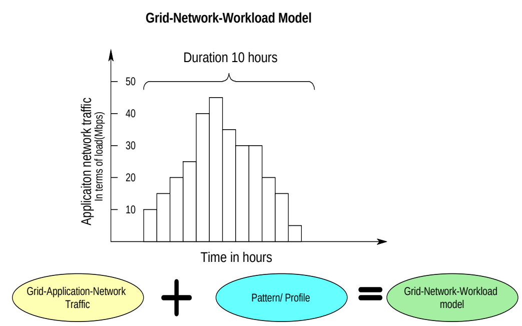


Fig.1.Componets of Grid network workload

### 2.1 Workload modeling techniques

Traffic can be created in a simulation model by importing from a live network (using traffic collection tools) or creating manually [9][10] by configuring various application and traffic attributes. Importing network traffic workload information from live network gives more realistic results as compared to manually configuring traffic. But in the absence of the real network or in the pre-deployment phase where the network is being designed, it may be required to predict the performance or answer what-if questions, in which case it becomes necessary to create the workload manually

Mathematical distributions are the basis for manually creating workload models. These distributions provide two options viz. Explicit traffic and Analytical traffic. Explicit traffic refers to a packet-by-packet data transfer with each transfer modeled as a discrete event. Analytically modeled traffic does not generate as many discrete events but it does impact explicit traffic in the form of network delays. Simulation studies may use explicit traffic, analytical traffic, or a combination of both [13][4].

### 2.1.1 Explicit traffic

Explicit traffic refers to packet-by-packet transfer where each transfer is modeled as a discrete event. It generates traffic explicitly; the system allocates memory for every packet, and queues up the packets and transmits them. At each layer of the protocol stack, it implements important protocol mechanisms, such as the following:

- Segmentation and reassembly
- Flow control
- Timeouts, back offs, and retransmissions
- Media access and quality of service prioritization

The application model is usually a simple model of client and server network applications [8]. Network simulators use a generic network application model to generate typical application traffic patterns known as applications model. Behavior of application models can be modified through parameters to represent wide variety of network applications. It does not, however, model in detail the behavior of any particular application. Fig 2. shows that application models along with the application attributes and mathematical probability distribution functions. These together make complete explicit application model.

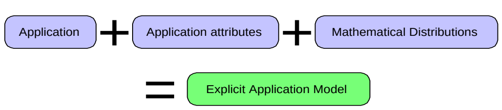


Fig 2. Explicit Application Model

The choice of a distribution for an application model is largely a function of what one is modeling and the input data points available. If measuring data in a production environment is possible, one can collect this data and fit it in a distribution using commercially available distribution fitting software. If data from a production environment is not available, one can use a predefined distribution as given in Table 1. Sometimes the distribution shown in table may not model the application behavior, so it may need some modifications to suit requirements.

| Applications | Attributes | Mathematical / Probability distribution functions | | | | |
|---|---|---|---|---|---|---|
| | | Uniform | Exponential | Constant | Normal | Poisson |
| *FTP* | Get/Total | | | x | | |
| | Inter-request time | | x | x | | |
| | File Size | | x | x | x | x |
| | | | | | | |
| | | | | | | |
| *Http* | Page Inter-arrival time | | x | | | |
| | Object size | | x | x | | |
| | Number of objects | x | | | x | |
| | | | | | | |
| *Video Conferencing* | Frame inter-arrival | | | x | x | |
| | Frame size | | x | x | x | |
| | | | | | | |
| *Email* | Send group size | x | | x | | |
| | Receive group size | x | | x | | |
| | Send inter-arrival time | | x | x | | |
| | Receive interarrival time | | x | x | | |
| | Size | | x | x | x | |
| *Failures* | | | | | | x |

Table 1. Recommended Traffic Model for applications

### 2.1.1 Analytical traffic

Analytical traffic refers to traffic that impacts performance of explicit traffic (packet flow) by inducing additional delays. The presence of analytical traffic results in queue build-up at intermediate devices and causes delays based on the queue length at any given time in the link. This model can be applied to routers, switches etc.

As an example consider an IP traffic flow used to model analytical load between a source and destination (Refer Fig. 3.) The important attributes for IP traffic flow are source, destination, protocol, traffic in bits per second, packets per second, type of service, traffic duration, traffic mix etc.

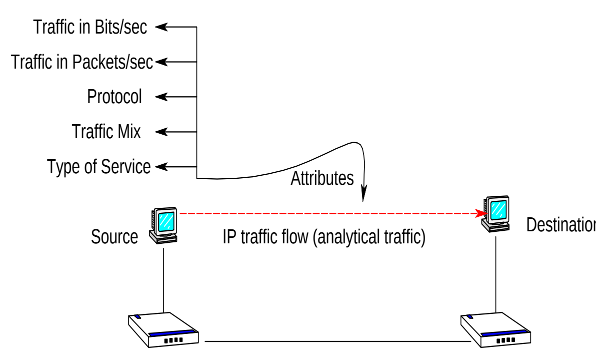


Fig. 3. IP traffic flow between source and destination

### 2.2. Usage pattern/ profile modeling

A Profile describes the activity patterns of a user or group of users in terms of the applications used over a period of time. There can be several different profiles running on a given network as per the usage pattern. Workload contains a list of profiles and for each the application can be executed or the pattern of the execution can be a) at the same time, b)one after the other in a specific order or c) in a random order. When describing the actions of a single user, generally the actions are serial since most people can only perform one activity at a given time (Fig. 4) However, when using applications that can perform non-blocking tasks, there can be more than one task running at a time (Fig. 5.) When describing the activities of a group of users, concurrency is common e.g. profile repetitions, application repetitions (Fig 6.)

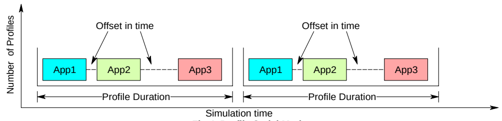


Fig.4. Profile Serial Mode

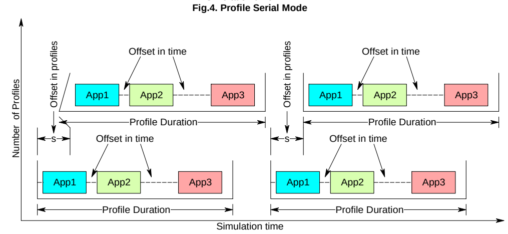


Fig.5. Profile Simultaneous Mode

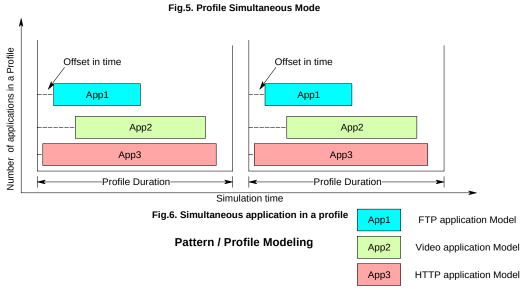


Fig.6. Simultaneous application in a profile

For capturing the effect of group of users a combination of simultaneous application in a profile with simultaneous mode is generally used. Profiles provide a mixture of different application running on the network. There can be different possible combinations of applications with different offsets.

# 3.GARUDA-Network-workload modeling methodology

Figure 7 shows the methodology we have followed for modeling GARUDA traffic. Some of the blocks are already discussed in section 2 in context of general simulators. The discussed features are already provided with OPNET Modeler, which is the most widely and commercially used simulator. We have chosen OPNET Simulator "Modeler" for modeling and rest of the document discusses the implementation of traffic modeling in context of OPNET modeler

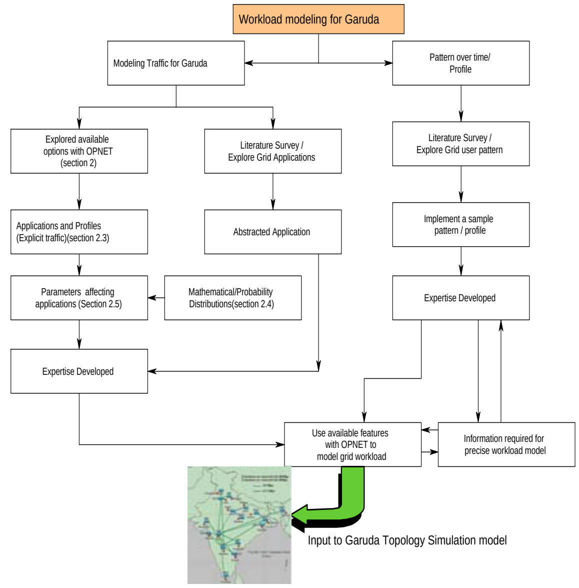


Fig.7. Workload modeling for Garuda Methodology

## 3.1 Traffic modeling techniques adopted for Garuda

As discussed in section 2., there are two types of modeling techniques: Explicit and Analytical, each of which can be modeled using OPNET.

For explicit end-to-end application traffic modeling OPNET provides applications and profiles whereas for analytical end-to-end application traffic modeling OPNET provides IP traffic flow. The following section describes how workload model was built using each of these techniques

## 3.2 Explicit traffic modeling using application and profiles

OPNET provides a facility to model an application behavior using tasks and phases. Exact application behavior can be designed for simulation purposes but it requires a deep understanding of a particular application. This approach is taken if a new application has to be deployed and its impact on the other application or network has to be studied. In such cases, instead of running the applications on the real network it is modeled in the simulator and its impact is studied.

However if the objective is not to study the performance for a particular application/profile but rather to study the overall network performance, then one can abstract out the application behavior using available applications in OPNET [1][2][3][8].

Following chart shows how various grid applications can be mapped to the available applications in OPNET Modeler.

| Grid Application | OPNET Specific applications | Transport Protocol | Port No. | Sensitivity |
|---|---|---|---|---|
| Video Conferencing | Video Conferencing / VoIP | UDP | 1719-1720, 1710-1715 | Delay, Delay-variation, Jitter |
| Remote visualization | | | | |
| Tele-immersion | | | | |
| Real time remote analysis | | | | |
| Data-oriented applications accessing large amounts of small data portions (remote file analysis) | Remote logins /FTP | TCP | 21 | Throughput |
| Remote hardware control | | | | |
| Client/server transactions in GRID middleware | Frequent but small TCP transfers | TCP | | Throughput |
| Database replication for load balancing and job locality | Custom Gridftp | TCP | 2811 | Throughput, Packet drop |
| Transfer of large data collections to one or more sites | | | | |
| Any grid-based application and middleware component under alpha and beta test | Random FTP | TCP | 21 | Throughput |
| Portal traffic | HTTP | TCP | 80 | Delay |

Table 2. Mapping Grid applications with OPNET specific application Model

### 3. Analytical traffic modeling using IP traffic flow

IP traffic flow is used to model analytical load between a source and destination (unidirectional). As discussed in section 2 this load affects explicit traffic flow. The important attributes for IP traffic flow are source, destination, protocol, traffic in bits per second and packets per second, type of service, traffic duration, traffic mix, etc.

As the underlying technology for access network in GARUDA is ethernet, the maximum transmission unit (MTU) for packet is 1500 bytes. Traffic in bits/second and traffic in packets per second is calculated using the formula:

Traffic in packets/sec=(traffic in bits/sec)/(8 * MTU)=(Traffic in bits/sec)/12000

E.g. suppose that traffic flow of 10Mbps has to be defined between two nodes. Then Packet/sec=10000000/12000=834 (approximate). Setting attributes for traffic in bits/second with start and stop time local/global attributes.

### 3.4 Sample Network-workload profile Modeling in Garuda

Once the application definition is done one has to define the usage patterns. This usage pattern is defined by the profile configuration. Multiple profiles can be defined as per the usage of different groups (refer section 2.2). A possible approach that can be taken is that one can develop a pattern for a client and then keep adding the number of clients with the same pattern with different offsets. Following table shows a workload profile, which was developed for 10 hours

| Application models | Duration (hour) | Average Bit rate (kbps) | Total data (Mbits) |
|---|---|---|---|
| | | | |
| Video conferencing | 2/0.5/0.4/0.5 | 512/2000/1000/2000 | 4000/4000/1600/4000 |
| Bulk transfers (FTP/ Custom Gridftp) | 4 | NA | 10000 |
| Remote login | 2 | NA | 4000 |
| HTTP | 2 | NA | 800 |
| E-mail | 2 | NA | 400 |

Workload model for 10 hours

Table. 3. Workload model for 10 hours

**This is very well based on the profile of a Physicist Circa 2005 given below.**

| Net Tasks of an average Physicist Circa 2005 | | | | |
|---|---|---|---|---|
| Task | Hours Per Day | KBits/sec | Total Mbits | |
| Conferencing | 2 | 512 | 4,000 | |
| "Coffee Room" | 0.5 | 2,000 | 4,000 | |
| Seminar | 0.4 | 1,000 | 1,600 | |
| Remote Sessions | 2 | 256 | 4,000 | |
| Analysis, Including Transfer in background | 4 | 700 | 10,000 | |
| Electronic Notebooks | 2 | 100 | 800 | |
| Papers: 20 Papers and Documents, Including Links | 2 | 100 | 800 | |
| E-Mail: 500 Multimedia Messages | 2 | 50 | 400 | |
| Interactive Virtual Reality | 0.5 | 2,000 | 4,000 | |
| TOTAL MBits Transferred Per Day | | | | 30,000 |
| Average Mbps During a Ten Hour Day | | | | 0.75 |

Table. 4. Sample Workload profile of a Physicist

### 4. CONCLUSION

Workload modeling involves two main components viz. traffic source and the pattern of traffic generation over a period of time. Each of these plays a crucial role in defining the network performance. Although many experts within the industry have come up with their own designs based upon their experience, they still have misconceptions and a daunting number of unanswered questions. Because of these unknowns in the network equation, workload characterization is a fundamental and essential element of network design.

There are very well defined traffic-modeling techniques (analytical and explicit) available in the simulators. One might not require to model an application from scratch, as most of the well-known application models are available within the simulators. Yet defining the right attributes with right choice of mathematical distribution is a complex task. Workload profile often follows some mathematical distribution, which is derived from the usage patterns of the groups/clients. So while designing workload model group/client usage is kept in mind and then mathematical distributions are applied. Workload modeling resulted into building certain level of expertise that has been listed below

- Traffic modeling techniques: Explicit and Analytical.
- Right attributes required for defining application behavior
- Mathematical distribution and how they could be used for modeling an application
- Workload profile modeling for a single or a group of users

**A crucial input required for Grid-Network-Workload-Modeling is the profile or usage pattern of the grid users as this will reflect the actual / desired user pattern. On availability of workload patterns new workload model can be designed and incorporated.**

## 5. References


1. Jun-Zhao Sun and Jaakko Sauvola "Modeling Techniques for simulation-based performance Evaluation of Mobile management Schemes"
2. Tiziana Ferrari,"Network Quality of Service for GRID-enabled Applications and Middleware"
3. James A. Rome, "Simulating large, high-speed networks on a supercomputer" SSFNet
4. Xavier Osó Alicart "Design and Implementation of a Traffic Engineering Research Platform based on OPNET "
5. Wu Feng" High-Performance Transport Protocols for Clusters and Grids"
6. Peter Key, Laurent Massoulié," A network flow model for mixtures of file transfers and streaming traffic "
7. OPNET Documentation 11.5:Application Model user Guide
8. OPNET Documentation 11.5:Modeling Concept reference Manual
9. OPNET Documentation 11.5:Representing network traffic
10. OPNET Documentation 11.5:Using distributions in simulation studies.
11. Ron Lee, An article on "An introduction to workload Characterization"
12. J. Potemans1, J. Theunis, B. Rodiers, B. Van den Broeck, P. Leys, E. Van Lil, A. Van de Capelle, "Simulation of a Campus Backbone Network, a case-study," Katholieke Universiteit Leuven (K.U.Leuven) Department of Electrical Engineering
13. David M. Nicol, Guanhua Yan, "Simulation of Network Traffic at Coarse Time-scales" University of Illinois, Urbana-Champaign
14. George Kola, Tev_k Kosar, and Miron Livny "Pro_ling Grid Data Transfer Protocols and Servers"