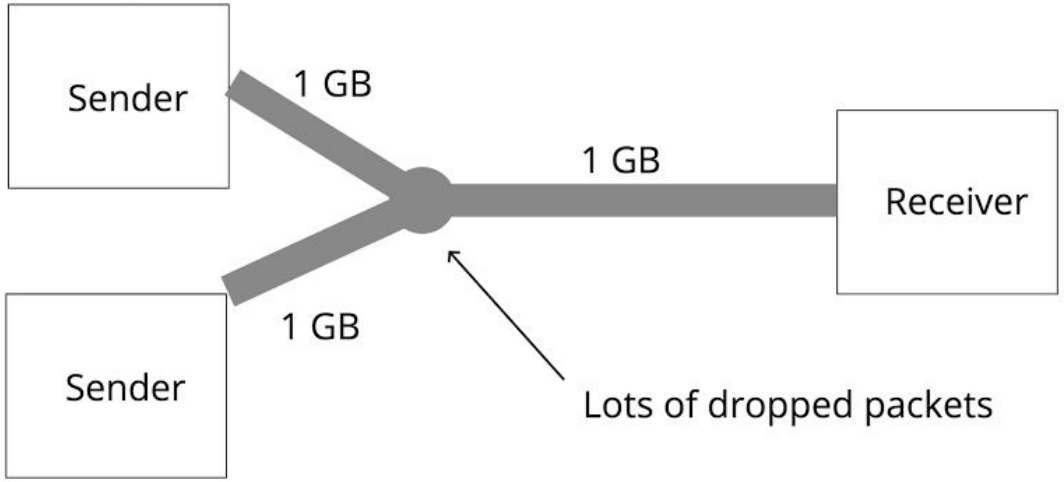
Congestion Control and Resorce Allocation

Name : Ansari M.Saeem
Uid : 2019430001

Congestion Control

- Efforts made by network nodes to prevent or respond to overload conditions



Resource Allocation

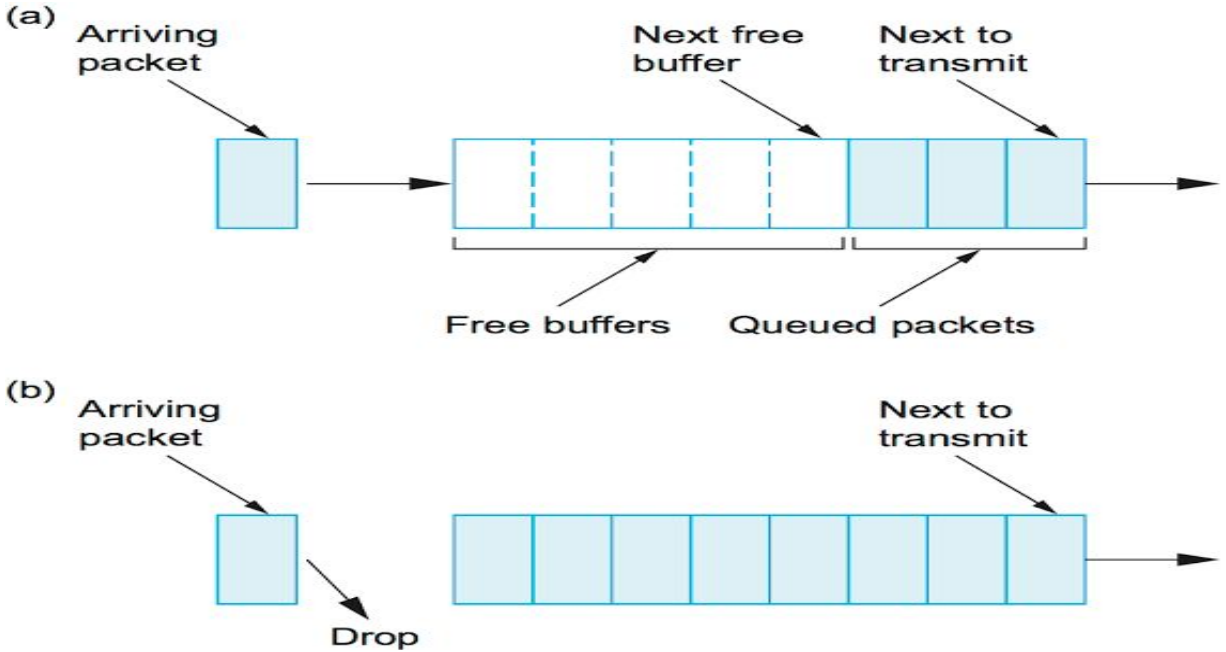
- Allocation or Distribution of resources to every node present in a network in a effective manner to avoid congestion.
- Resources are Bandwidth and Buffer memory.
- Congestion control and resource allocation are two sides of the same coin. If the network takes an active role in allocating resources, then congestion may be avoided.

Queuing Disciplines

- Each router must implement some queuing discipline that consists of a scheduling discipline and a drop policy.
- The scheduling discipline determines the order in which packets are transmitted (i.e., allocates link bandwidth).
- The drop policy determines which packets get discarded (i.e. allocates buffer space).
- Example :
  - FIFO Queue
  - Priority Queue
  - Fair Queue

FIFO (First-Come, First-Served) Queuing

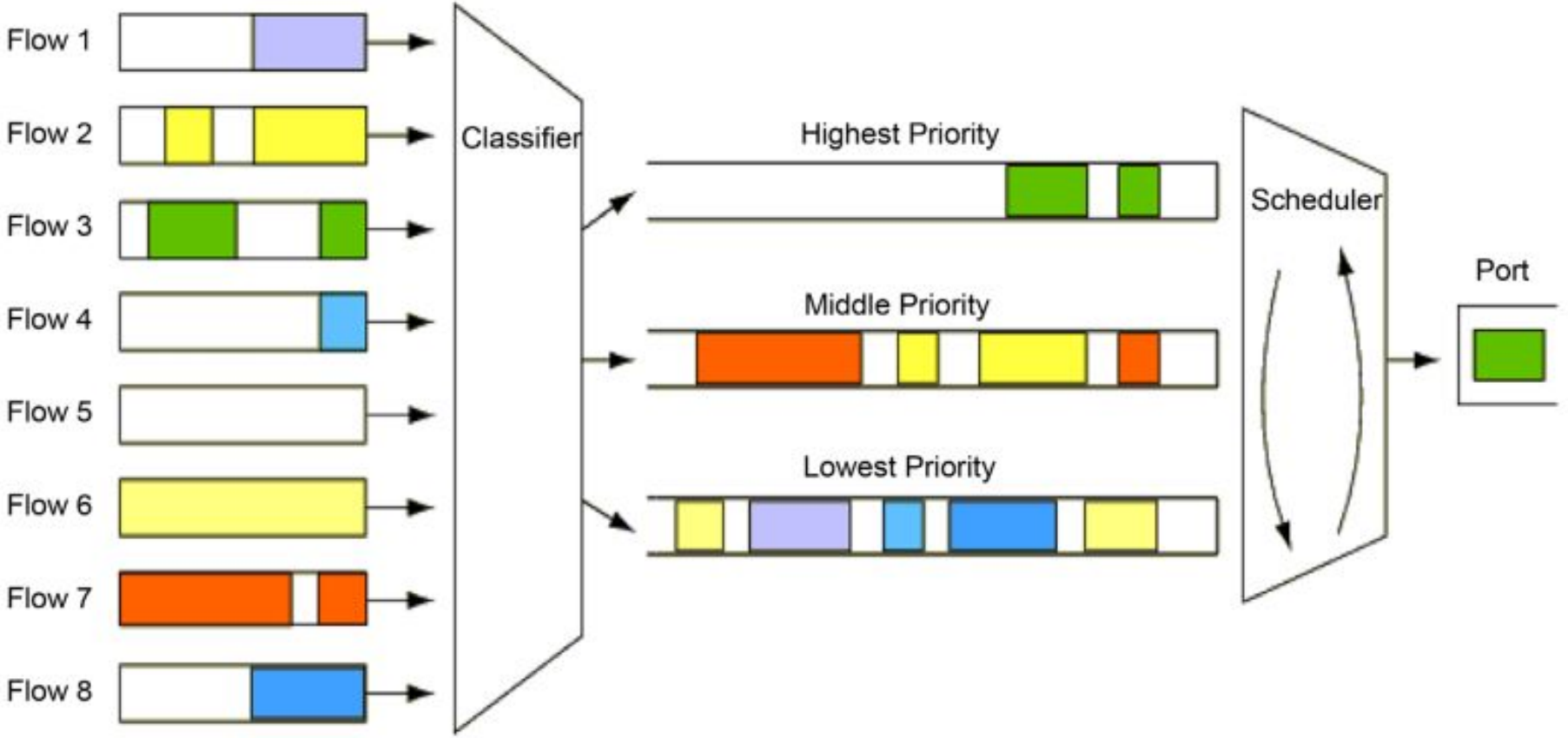
- The first packet that arrives at a router is the first packet to be transmitted.
- If a packet arrives and the queue (buffer space) is full, then the router discards that packet. This is called tail drop.



Priority Queuing

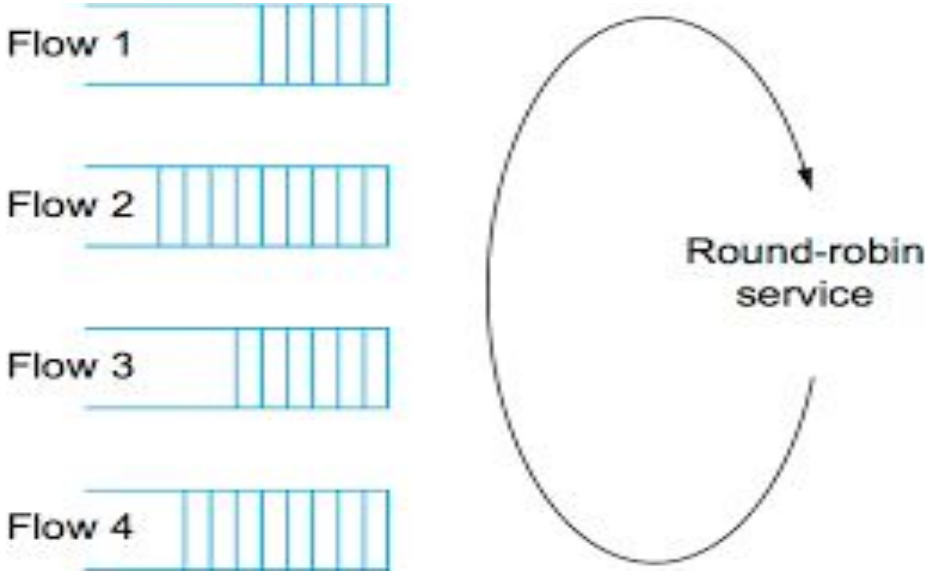
- Each packet is marked with a priority.
- The routers implement multiple FIFO queues, one for each
- priority class.
- Queues are serviced in strict order of queue priority. Within each priority, packets are managed in a FIFO manner.
- Problem: the high-priority queue can starve out all the other queues.

Priority Queuing (cont..)



Fair Queuing (FQ)

- Maintaining a separate queue for each flow currently being handled by the router. Queues are serviced in round-robin fashion.
- Packets being processed at a router are not necessarily the same length. To truly allocate the link bandwidth in a fair manner, it is necessary to take packet length into consideration.



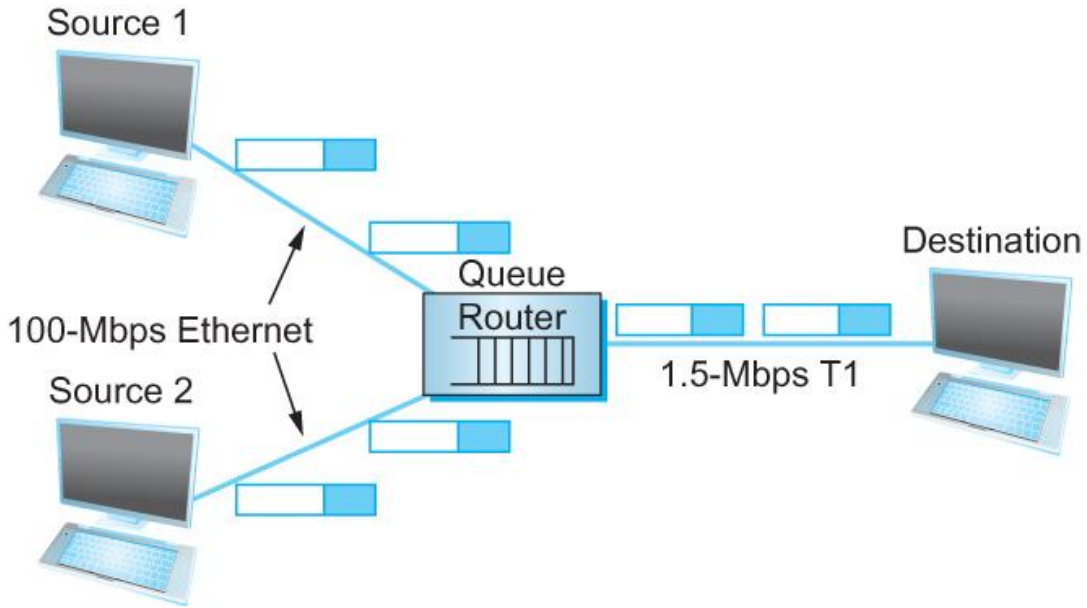
Issues in Resource Allocation

Packet switch network :

- We consider resource allocation in a packet-switched network (or internet) consisting of multiple links and switches (or routers).
- In such an environment, a given source may have more than enough capacity on the immediate outgoing link to send a packet, but somewhere in the middle of a network, its packets encounter a link that is being used by many different traffic sources

Issues in Resource Allocation

Packet Switch Network



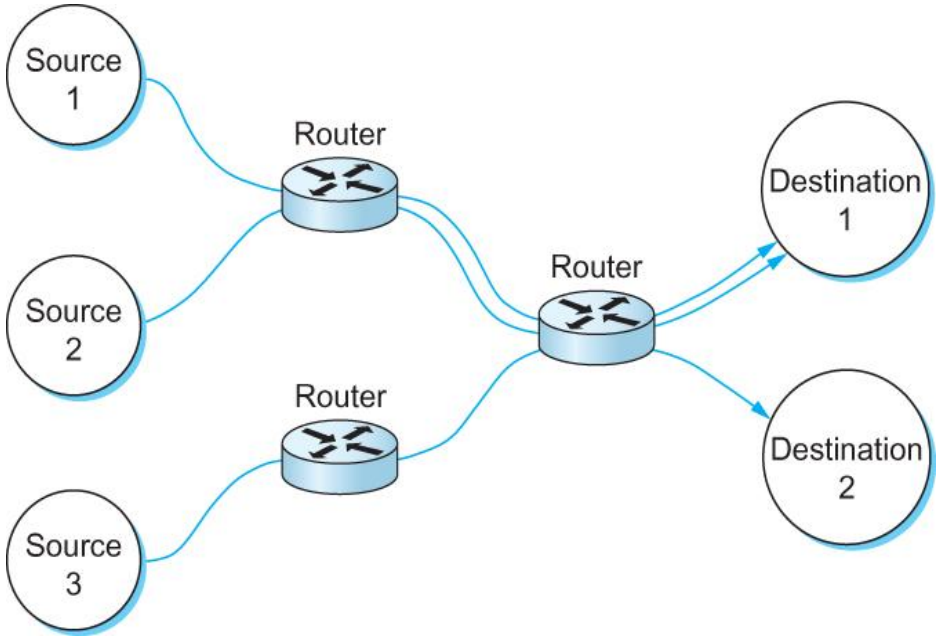
Issues in Resource Allocation

Connectionless Flows

- For much of our discussion, we assume that the network is essentially connectionless, with any connection-oriented service implemented in the transport protocol that is running on the end hosts.
- In particular, the assumption that all datagrams are completely independent in a connectionless network is too strong.
- The datagrams are certainly switched independently, but it is usually the case that a stream of datagrams between a particular pair of hosts flows through a particular set of routers
- Because multiple related packets flow through each router, it sometimes makes sense to maintain some state information for each flow, information that can be used to make resource allocation decisions about the packets that belong to the flow. This state is sometimes called soft state

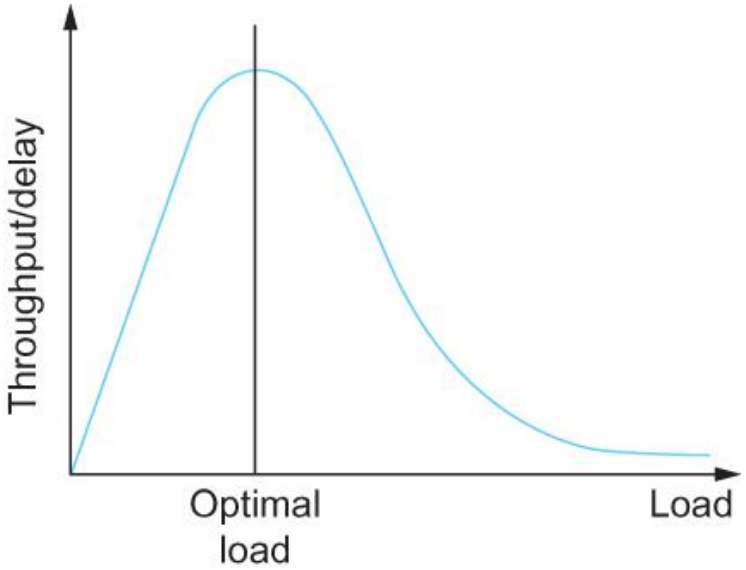
Issues in Resource Allocation

Connectionless Flow :



Issues in Resource Allocation

Evaluation criteria :

- A good starting point for evaluating the effectiveness of a resource allocation scheme is to consider the two principal metrics of networking: throughput and delay.
- Power = Throughput / Delay
- 

Paper : Task scheduling and resource allocation in cloud computing using a heuristic approach

Introduction :

- Cloud computing is an accelerating technology in the field of distributed computing.
- Cloud computing can be used in applications that include storing data, data analytics and IoT applications.

Input data : Cybershake scientific workflow

- User submit task request to cloud and it is responsibility of cloud to schedule and execute various requests and manage resources efficiently.
- The proposed system uses Cybershake scientific workflow data as input tasks which is used by the Southern California Earthquake Center (SCEC) to characterize earthquake hazards using the Probabilistic Seismic Hazard Analysis (PSHA) technique.

Paper : Task (Cont..)

Task-Table :

Table 1 Cybershake seismogram synthesis tasks

Tasks	Size of tasks	Time
Task 3	62,69,51,663	39.06
Task 5	69,47,76,323	38.49
Task 7	58,57,63,637	36.27
Task 9	53,68,97,326	32.29
Task 11	67,05,35,542	62.25
Task 14	40,67,28,38,798	96.91
Task 16	45,23,96,996	45.60
Task 18	50,27,64,231	28.67
Task 20	62,41,88,532	24.56
Task 22	42,65,77,006	31.05
Task 24	51,58,32,878	54.87
Task 26	68,14,99,417	23.99
Task 28	44,14,51,516	26.46

Paper : Architecture of Proposed System (Cont..)

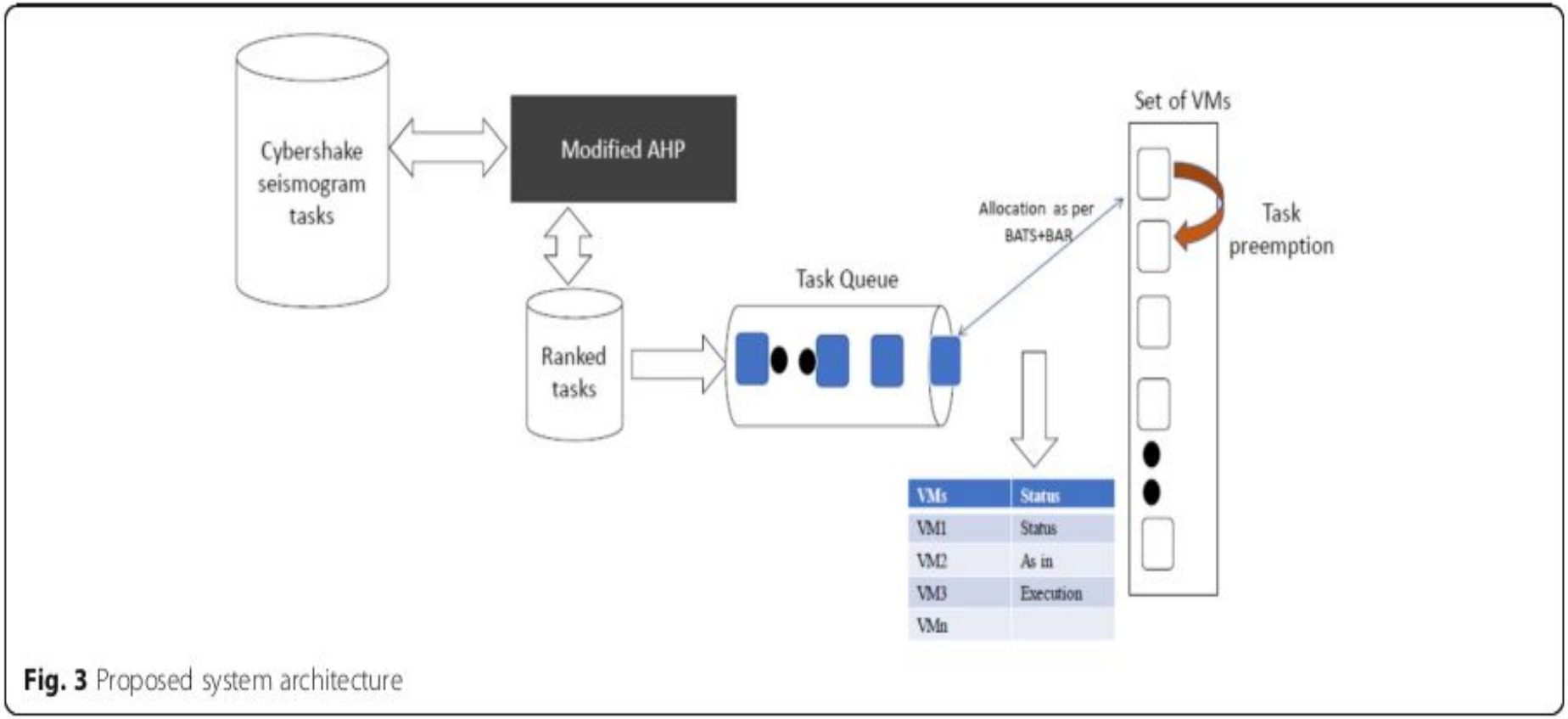


Fig. 3 Proposed system architecture

Paper : Methodology (cont..)

Analytic Hierarchy Process (AHP) :

- Design to solve complex problem with multiple criteria.
- AHP is used to rank the all the tasks given as input.
- To evaluate preferences the proposed system uses the saaty preference table with its numerical rating.

Numerical rating	Judgment preference
9	Extremely preferred
8	Very strongly to extremely preferred
7	Very strongly preferred to preferred
6	Strongly to very strongly
5	Strongly preferred
4	Moderately to strongly preferred
3	Moderately preferred
2	Equally to moderately preferred
1	Equally preferred

Paper : Methodology

BATS + BAR system :

- Independent tasks of equal size are considered in the design of this system.
- However, in allocating resources, the system does not consider the load on virtual machines because the waiting period for the tasks is long.
- In other cases, one virtual machine is busy while it executes a task, whereas others are occupied and waiting for jobs.
- Procedure :
  - Aggregate all of the task information that is ordered by rank.
  - Virtual machine (server) information is collected. This information includes the initial load on the virtual machine, its bandwidth and the time required to process the tasks on the server.
  - A bipartite graph is generated with the number of tasks. The ranking priorities, can be used to constructed a graph, by which each task is allocated to a virtual machine.

Paper: Flow Chart (cont..)

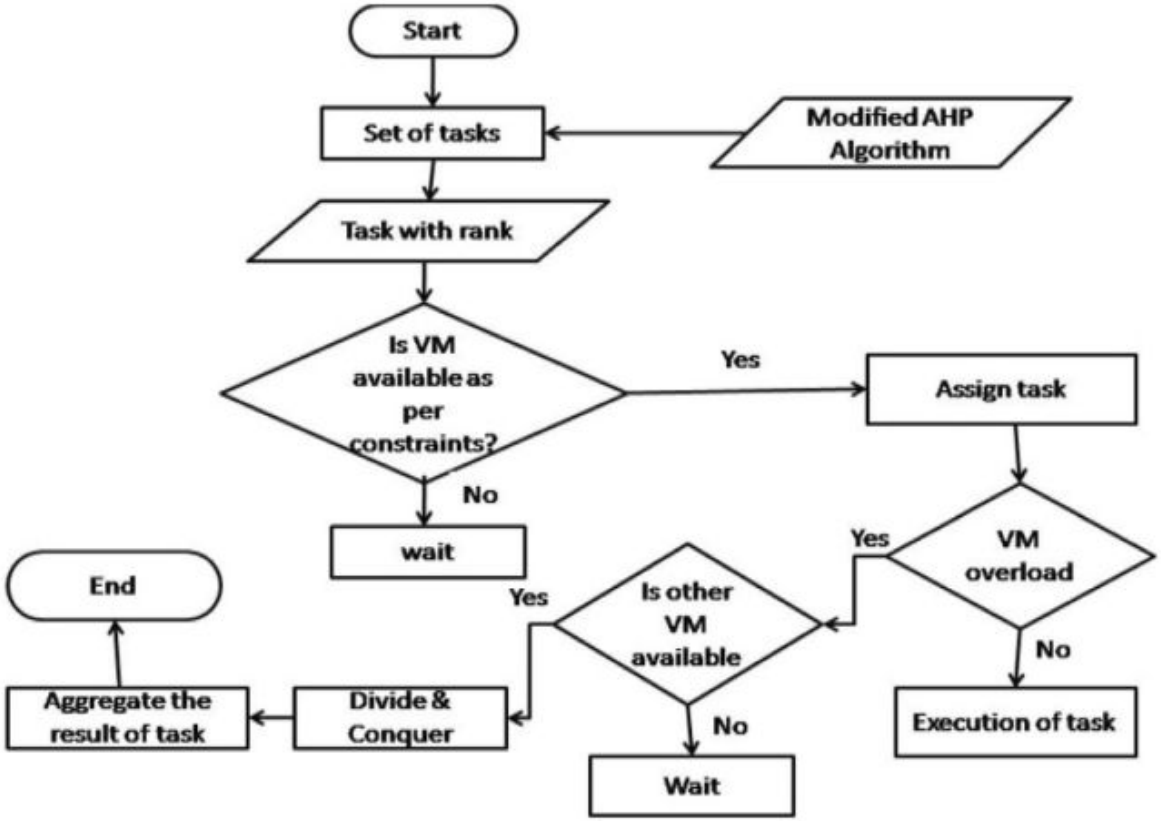


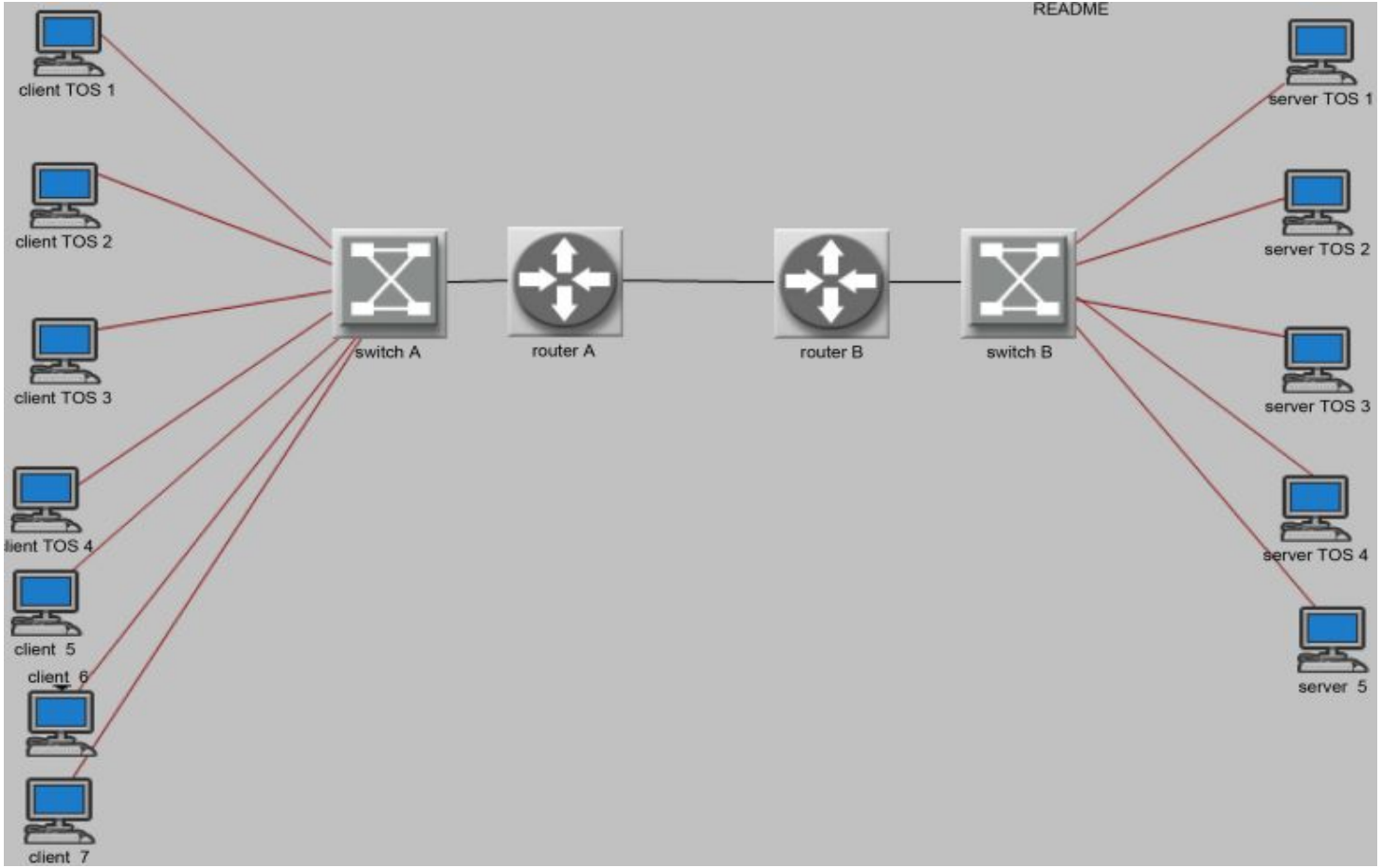
Fig. 4 Proposed system flowchart

Paper-2 : The Effect of Queuing Mechanisms First in First out (FIFO), Priority Queuing (PQ) and Weighted Fair Queuing (WFQ) on Network's Routers and Applications

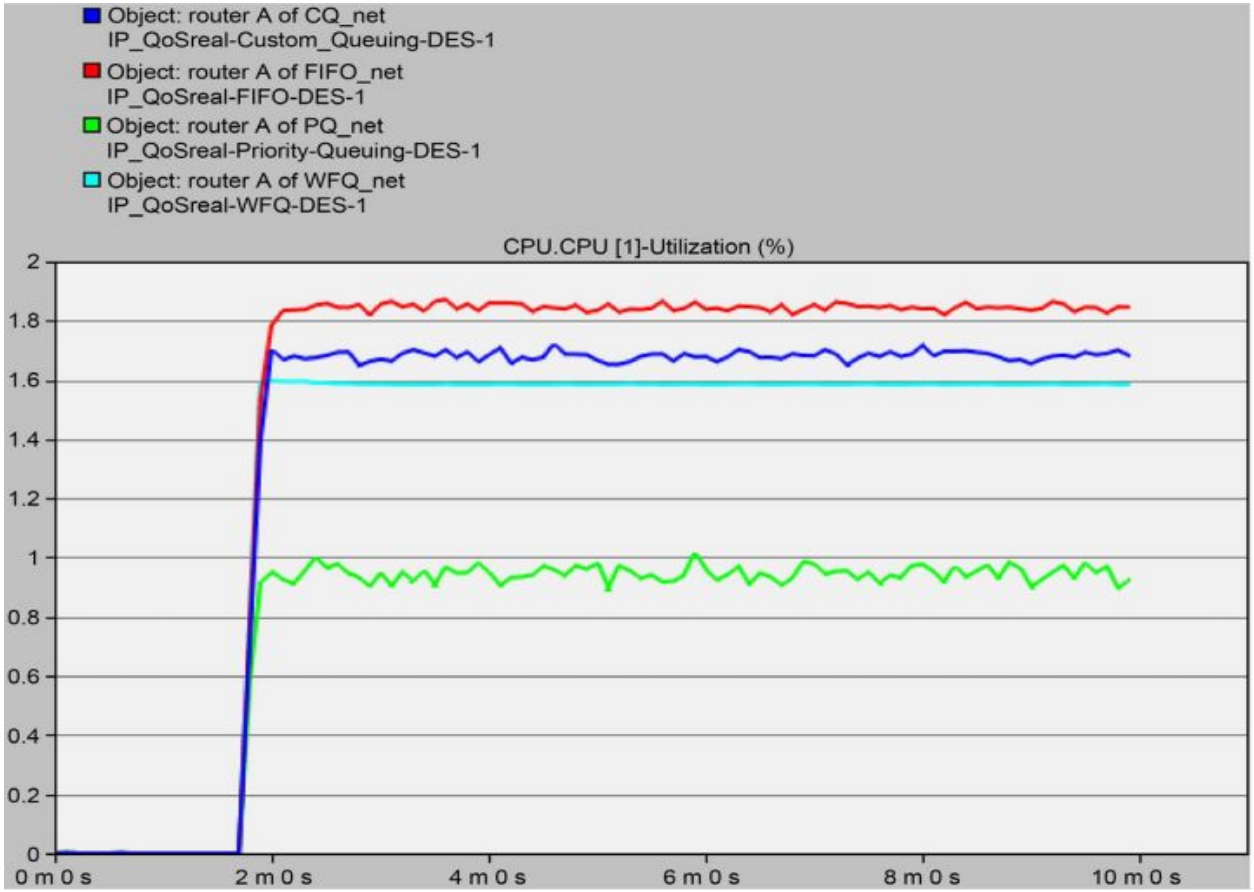
Introduction :

- The paper presents the simulation results of the comparison of three Queuing Mechanisms, First in First out (FIFO), Priority Queuing (PQ), and Weighted Fair Queuing (WFQ).

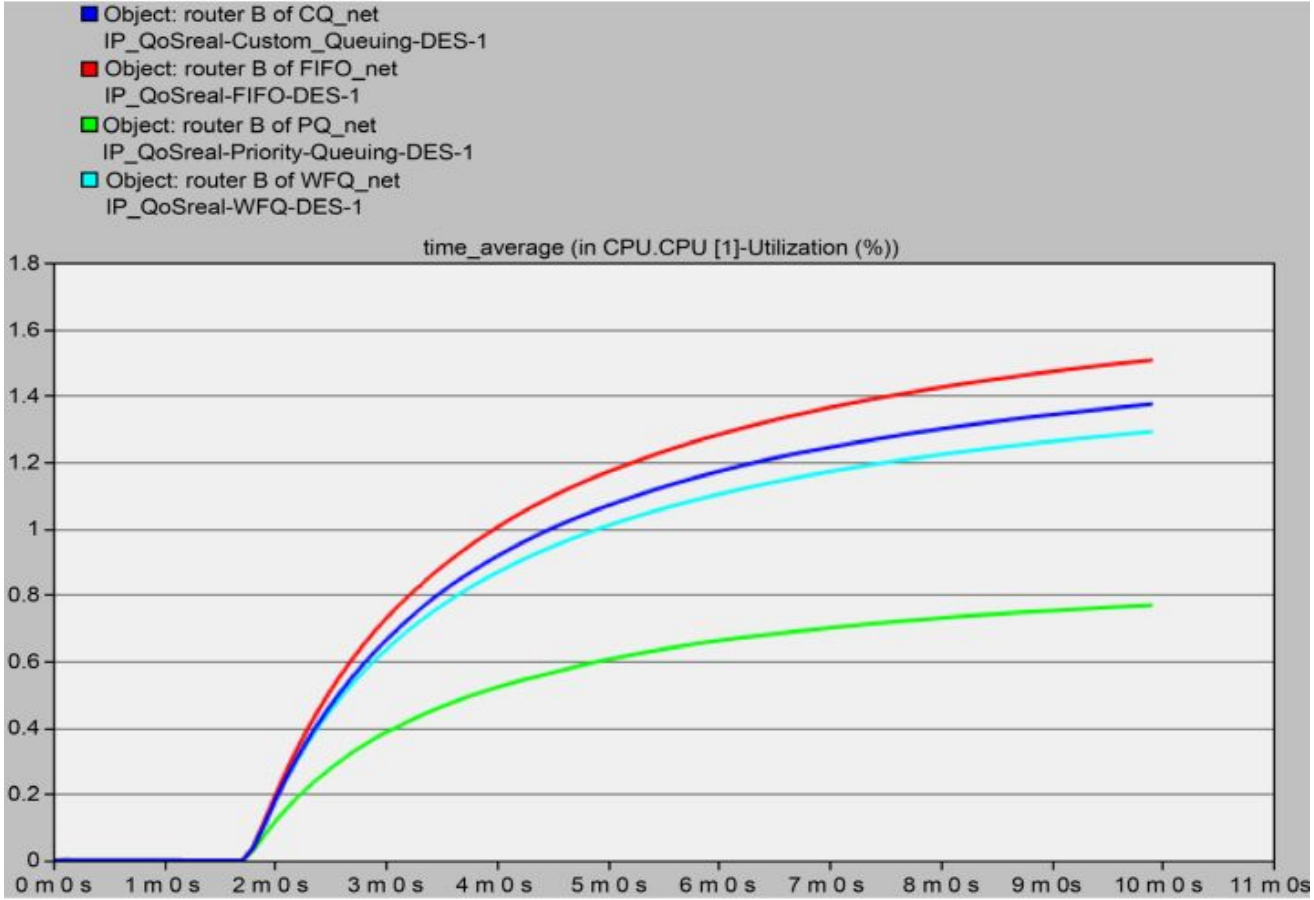
Paper-2: Network Topology (Cont..)



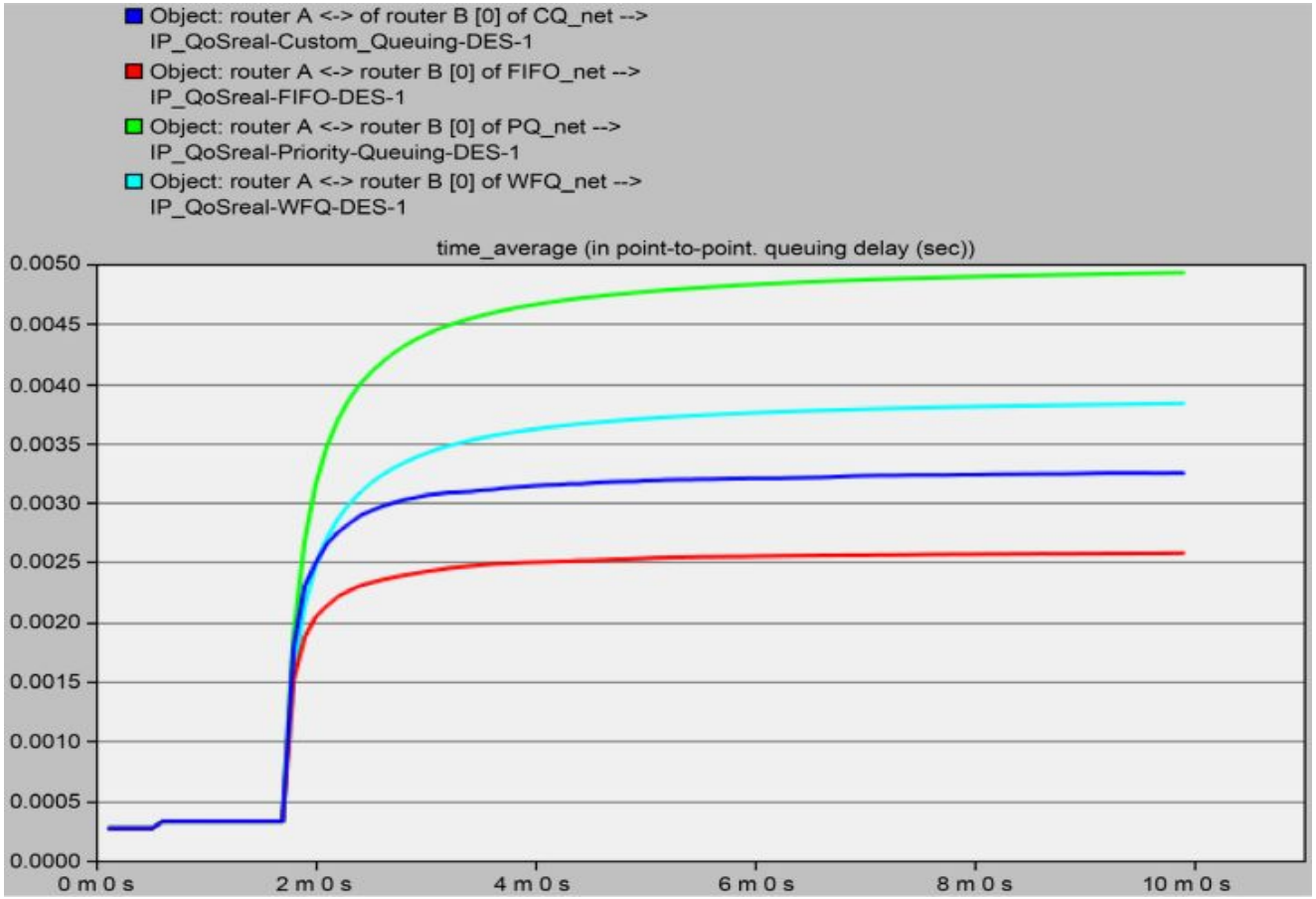
Paper-2 : CPU utilization of router-A (Cont..)



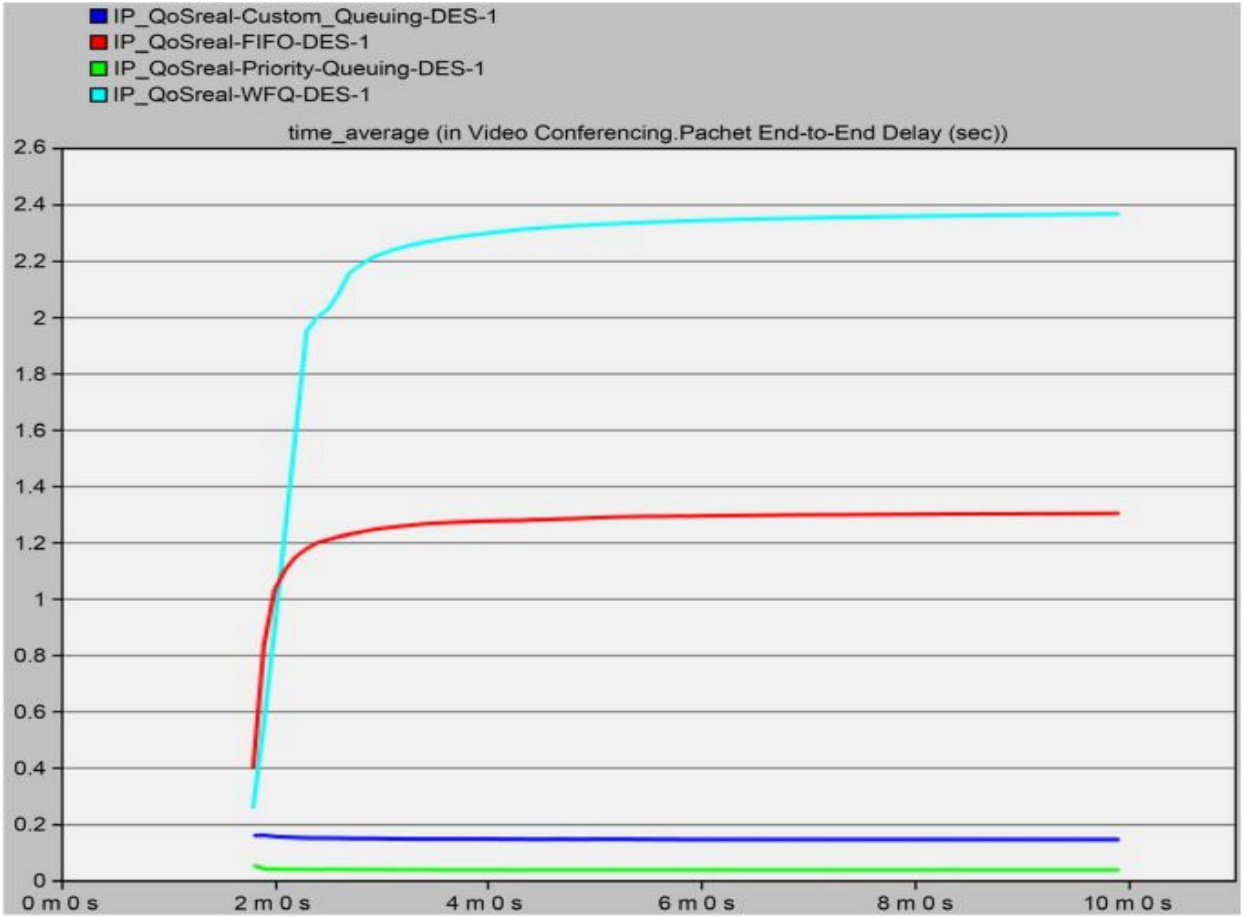
Paper-2 : CPU utilization of router-B (Cont..)



Paper-2 : Queuing delay from router A to B (Cont..)



Paper-2 : Overall Delay in Network (Cont..)



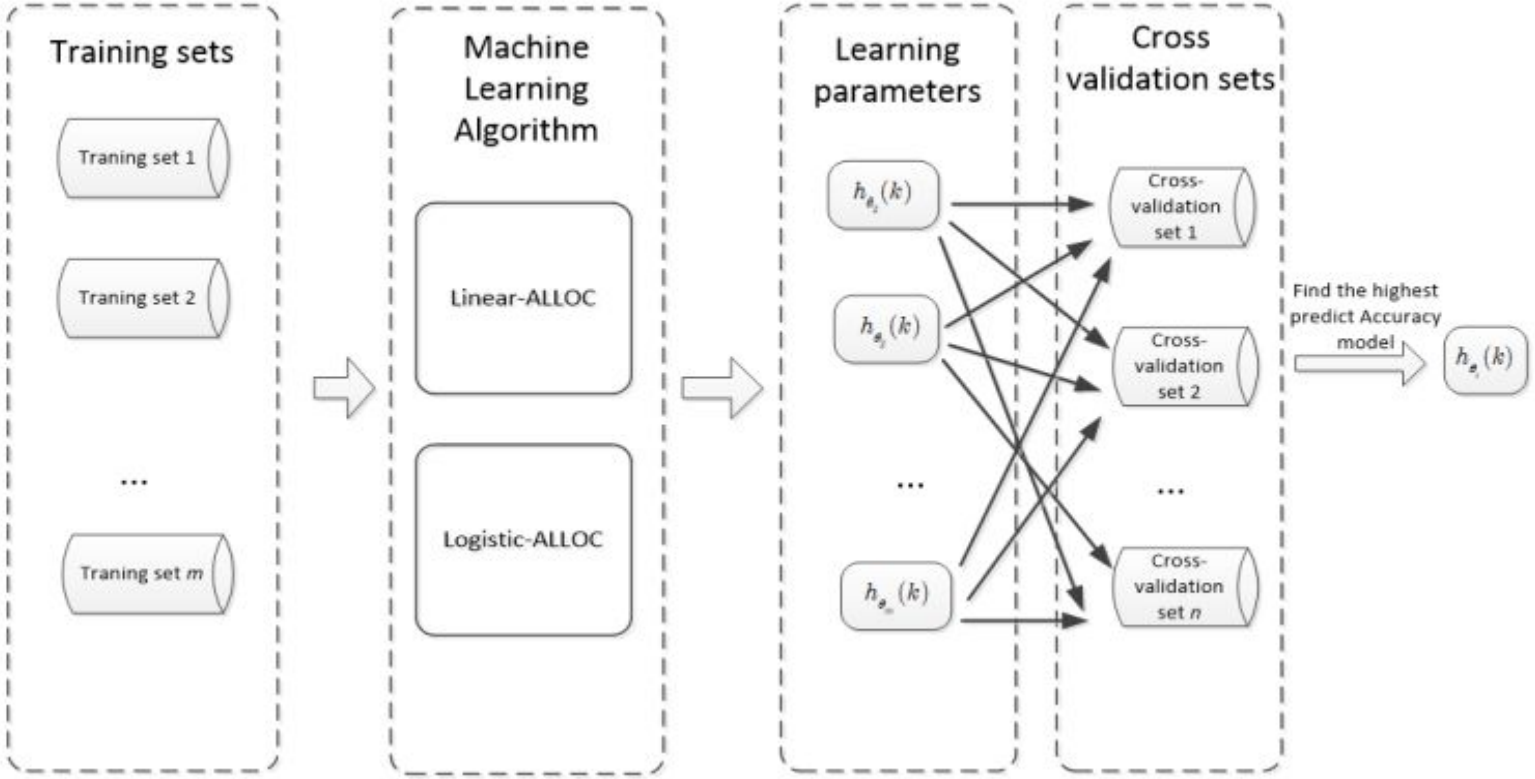
Paper-3 : Machine Learning Based Resource Allocation of Cloud Computing in Auction.

Introduction :

- It’s a challenge to determine how much amount of resource will be required for a specific node.
- A single-dimensional resource allocation problem has been proven to be NP-hard so it cannot be solved in polynomial time.
- In multi-dimensional resource allocation problems, different types of resources (e.g. CPU, memory, and storage) must be considered. Thus, multi-dimensional resource allocation problems are strongly NP-hard.
- The resource allocation problem can be solved by using an optimal algorithm such as integer programming and dynamic programming, or it can be solved by an approximate algorithm.

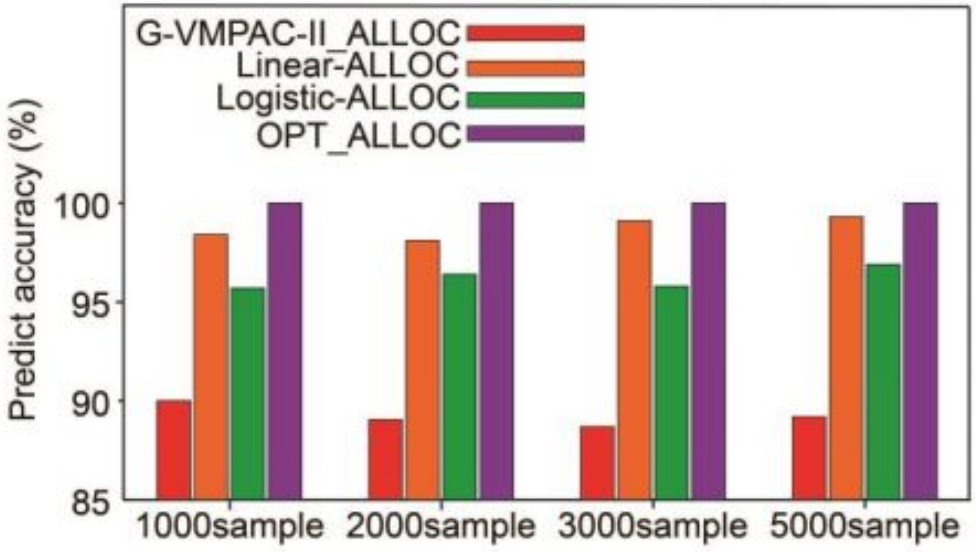
Paper-3 : Learning Process (Cont..)

Machine Learning Based Resource Allocation of Cloud Computing in Auction

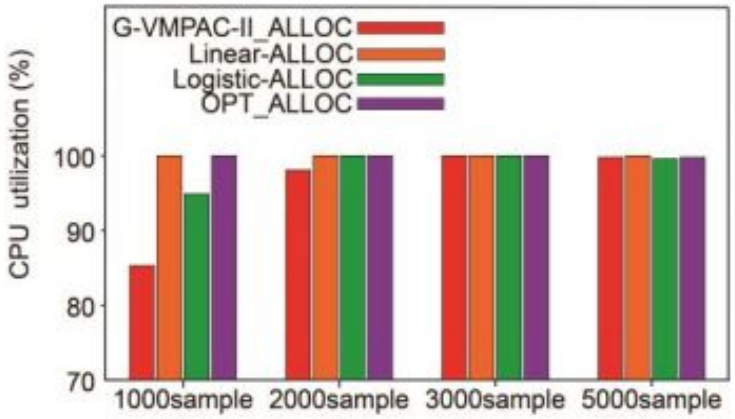


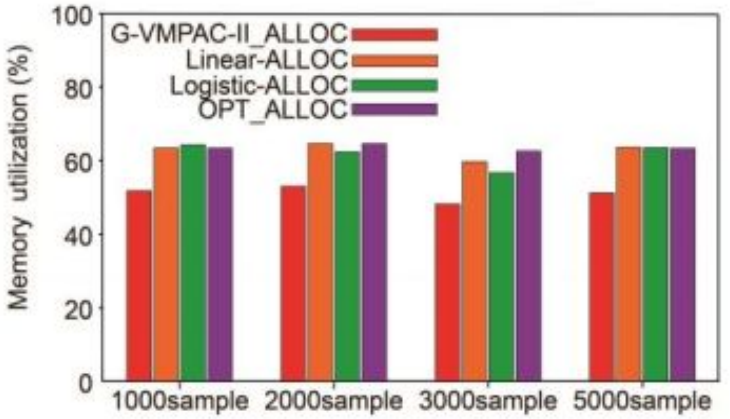
Paper-3 : Accuracy Comparison

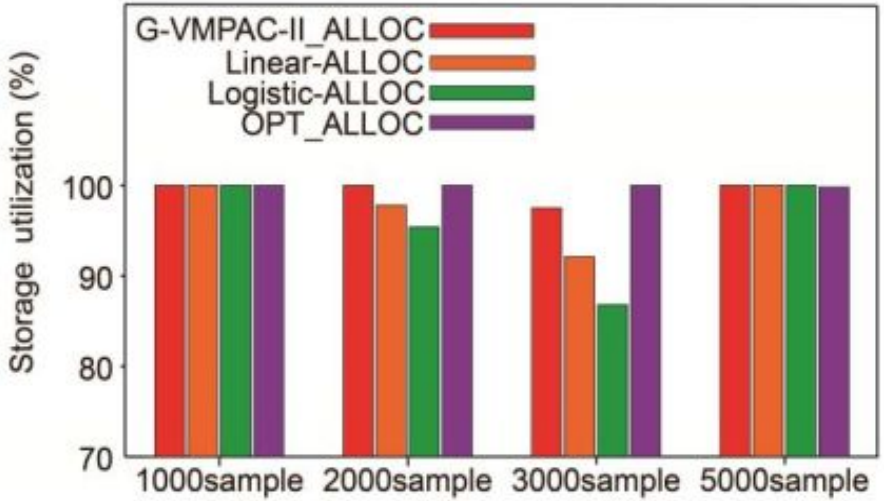
- Red - Approximation Allocation
- Orange - Linear Allocation
- Green - Logistic Allocation
- Purple - Optimal Allocation



Paper-3 : Comparison







Conclusion

- We studied congestion control and resource allocation and how they are related.
- We have seen various Queueing algorithm (FIFO, PQ, FQ)
- We have seen how to allocate resource in cloud computing environment.
- We did comparative studied of queueing algorithms (FIFO, PQ, FQ)

References

- Machine Learning Based Resource Allocation of Cloud Computing in Auction Jixian Zhang 1 , Ning Xie 1 , Xuejie Zhang 1 , Kun Yue 1 , Weidong Li 2, * and Deepesh Kumar 3
- The Effect of Queuing Mechanisms First in First out (FIFO), Priority Queuing (PQ) and Weighted Fair Queuing (WFQ) on Network's Routers and Applications Mustafa El Gili Mustafa 1,2 , Samani A. Talab 2
- Task scheduling and resource allocation in cloud computing using a heuristic approach Mahendra Bhatu Gawali 1* and Subhash K. Shinde 2