

ABSTRACT

Software-defined networking (SDN) is the physical separation of the control and data planes in networking devices. A logically centralised controller plane which uses a network-wide view data structure to control several data plane devices is another defining attribute of SDN. The centralised controllers and the network-wide view data structure are difficult to scale as the network and the data it carries grow. Solutions which have been proposed to combat this challenge in SDN lack the use of the statistical properties of the workload or network traffic seen by SDN controllers. Hence, the objective of this research is twofold: Firstly, the statistical properties of the controller workload are investigated. Secondly, Autoregressive Integrated Moving Average Models (ARIMA) and Artificial Neural Network (ANN) models are investigated to establish the feasibility of forecasting the controller workload signal. Representations of the state of the controller plane in the network-wide view in the form of forecasts of the controller workload will enable control applications to detect dwindling controller resources and therefore alleviate controller congestion. On the other hand, realistic statistical traffic models of the controller workload variable are sought for the design and evaluation of SDN controllers. A data center network prototype is created by making use of an SDN network emulator called Mininet and an SDN controller called Onos. It was found that 1–2% of flows arrive within $10\mu\text{s}$ of each other and more than 80% have inter-arrival times in the range of $10\mu\text{s}$ –10ms. These inter-arrival times were found to follow a beta distribution, which is similar to findings made in Machine Type Communications (MTC). The use of ARIMA and ANN to forecast the controller workload established that it is feasible to forecast the workload seen by SDN controllers. The accuracy of these models was found to be comparable for continuously valued time series signals. The ANN model was found to be applicable even in discretely valued time series data.