

# Something something something: A resource allocation framework to model cloud outage events<sup>☆</sup>

Jonathan Dunne<sup>a,\*</sup>, David Malone<sup>a</sup>

<sup>a</sup>*Hamilton Institute, Maynooth University, Kildare, Ireland*

---

## Abstract

As SME's adopt cloud technologies and rapid delivery models as a means to provide high value customer offers, there is a clear focus on uptime. Cloud outages represent a challenge to an SME's to deliver and maintain a services platform. If a Cloud platform suffers from downtime this can have a negative on business revenue. Additionally outages can divert resources from product development/delivery tasks to reactive remediation. These challenges are more immediate to an SME's with a small pool of resources at their disposal. Therefore it is necessary to develop a framework which can be used to predict the arrival of cloud outage events. A framework which can be used by cloud Operations teams to manage their scarce pool of resources to resolve outages, while minimising impact to service delivery. This article considers existing modelling techniques such as the M/M/1 queue, and proposes a more accurate approach. We first calculate the inter arrival and service distributions. Next we formally test for dependence between event arrivals. We then model a series of outage events in a G/G/1 queue with a Monte Carlo simulation to determine queue busy time. Finally we compare the precision of our framework against an M/M/1 simulation and real outage event data. The results demonstrated that our framework can improve the estimation of cloud outage events and aid DevOps resource planning.

*Keywords:* Outage simulation, Resource allocation model, Queuing Theory

---

## 1. Introduction

Cloud outage prediction and resolution is an important activity in the management of a cloud service. Recent studies and media reports have documented cases of cloud outages from high profile cloud service providers [ref]. During 2016 alone the CRN website has documented ten highest profile cloud outages to have occurred so far. Due to the increasing complex nature of the data centre infrastructure coupled with the rapid continuous delivery of incremental software updates it seems that cloud outages are with us for the time being.

For operations teams that maintain a cloud infrastructure, they rely on state of the art monitoring and alert systems to determine when an outage occurs [ref]. Once a new outage is observed, depending on the outage type (e.g. Software component, infrastructure, hardware etc) additional relevant experts may be called to remediate the issue. The time taken to resolve the issue may depend on a number of factors: Ability to find the relevant expert, swift problem diagnosis and velocity of the pushing a fix to production systems.

Both SME's and micro teams within large organisation's face a number of challenges when adopting a cloud

platform and a mechanism to deliver products and services. A number of recent studies have outlined that both frequency and duration of outage events are key challenges. Almost all European SME's (93%) employ less than ten people [ref]. Ensuring that adequate skills and resources are available to accommodate incoming outage events is highly desirable.

Downtime is bad for business. Whether the company provides a hosting platform, more commonly known as Platform as a Service (PaaS) or for a company that consumes such a platform to deliver their own services, more commonly known as Software as a Service (SaaS). The end result is the same: Business disruption, lost revenue, recovery/remediation costs etc. A recent US study which looked at the cost of data centre downtimes, calculated the mean cost to be \$5617 per minute of downtime[ref].

Another consideration is the idea of event dependence. Typical off the shelf single-server queue models such as M/M/1 and G/G/1 assume that the inter-arrival and service times between events are independent. However if some form of dependence is found between events how useful would a queuing model which assumes independence compare against that of a queuing model with dependence properties.

In this paper we propose a framework that a micro team or SME can leverage to best manage their existing resource pool. The core idea of this framework is for operations teams to use a special case of the G/G/1 queue

---

\*Corresponding author

Email address: [jonathan.dunne.2015@mumail.ie](mailto:jonathan.dunne.2015@mumail.ie) (Jonathan Dunne)

to model the inter arrival, service times of outage events. We show that the special case of G/G/1 out performs typical off the shelf queuing models like M/M/1 in terms of precision. Additionally our framework also considers independence between successive outage events. This article consists of a study of outage event data from a large enterprise dataset. By analysis of this outage event data this study shows the efficacy of the G/G/1 special case and how it can be used to reasonable model cloud outage event data. Additionally we highlight the shortcomings of the M/M/1 queue specifically in the realm of service times of repairable systems, such as cloud based platforms. Finally this study highlights how independence/dependence between cloud outage's plays a role in the frameworks precision.

To help researches reproduce and extend the work conducted as part of this study, pseudo-code of the queue modelling framework is provided . By utilising this queue framework, researches will have the ability to test their preferred case of the G/G/1 model against the M/M/1 model.

The rest of this article is divided up into the following sections: Section 2 introduces background and related work. Section 3 discusses the data set collected (and associated study terminology), outlines the research questions that are answered by this study and the limitations of the dataset. Section 4 outlines the experimental approach and associated results. Section 5 discusses the results of our experiments. Finally in Section 6, we conclude this paper and discuss future work.

## 2. Background and related work

The author names and affiliations could be formatted in two ways:

- (1) Group the authors per affiliation.
- (2) Use footnotes to indicate the affiliations.

See the front matter of this document for examples. You are recommended to conform your choice to the journal you are submitting to.

## 3. Data set and test methodology

## 4. Results

## 5. Discussion

## 6. Conclusion

## 7. Bibliography styles

There are various bibliography styles available. You can select the style of your choice in the preamble of this document. These styles are Elsevier styles based on standard styles like Harvard and Vancouver. Please use BibTEX to generate your bibliography and include DOIs whenever available.

Here are two sample references: [? ? ].

## References