











- [2] Angius, A., Horváth, A. and Wolf, V. 2013. Approximate transient analysis of queuing networks by quasi product forms. *Analytical and Stochastic Modeling Techniques and Applications*. Springer. 22–36.
- [3] Ardagna, D., Casale, G., Ciavotta, M., Pérez, J.F. and Wang, W. 2014. Quality-of-service in cloud computing: modeling techniques and their applications. *Journal of Internet Services and Applications*. 5, 1 (2014), 1–17.
- [4] Au-Yeung, S.W., Dingle, N.J. and Knottenbelt, W.J. 2004. Efficient approximation of response time densities and quantiles in stochastic models. *ACM SIGSOFT Software Engineering Notes* (2004), 151–155.
- [5] Broadwell, P.M. 2004. Response time as a performability metric for online services. Report No. UCB//CSD-04-1324. Computer Science Division (EECS), University of California, Berkeley.
- [6] Grottko, M., Apte, V., Trivedi, K.S. and Woollet, S. 2011. Response time distributions in networks of queues. *Queueing Networks*. Springer. 587–641.
- [7] Grozev, N. and Buyya, R. 2014. Multi-cloud provisioning and load distribution for three-tier applications. *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*. 9, 3 (2014), 13.
- [8] Gullhav, A.N., Nygreen, B. and Heegaard, P.E. 2013. Approximating the response time distribution of fault-tolerant multi-tier cloud services. *Proceedings of the 2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing* (2013), 287–291.
- [9] Harchol-Balter, M. 2013. *Performance Modeling and Design of Computer Systems: Queueing Theory in Action*. Cambridge University Press.
- [10] Van Houdt, B. and Blondia, C. 2005. Approximated transient queue length and waiting time distributions via steady state analysis. *Stochastic Models*. 21, 2-3 (2005), 725–744.
- [11] Van Velthoven, J., Van Houdt, B. and Blondia, C. 2005. Response time distribution in a D-MAP/PH/1 queue with general customer impatience. *Stochastic Models*. 21, 2-3 (2005), 745–765.