

0.1. Павский К.В., Павский В.А. Математическая модель и оценки показателей надежности масштабируемых вычислительных систем с групповым восстановлением

Масштабируемые вычислительные системы (ВС) являются одним из естественных направлений развития вычислительных средств высокой производительности, обусловленных потребностями практики [1]. Например, кластерные ВС составляют 93% суперкомпьютеров списка Top500 (56я редакция). Количество элементарных машин (ЭМ, например, узлы) в современных суперкомпьютерах, масштабируемых вычислительных системах измеряется десятками тысяч и более, а элементная база близка к предельной. Несмотря на высокую надежность отдельного компонента, в силу своей большемасштабности, в ВС наличие отказов [2] становится проблемой. К таким системам предъявляются повышенные требования по надежности и живучести. Для масштабируемых вычислительных систем актуальным становится организация и анализ потенциальных возможностей.

Классическими средствами анализа стали методы теории массового обслуживания (ТМО), случайных процессов, имитационного моделирования [3]. В работе рассматриваем масштабируемую вычислительную систему с резервными машинами при групповом восстановлении отказавших ЭМ как стохастический объект. Обладая ограниченной надежностью, вычислительные системы должны выполнять свои функции по решению задач, несмотря на наличие отказов в системе. Считаем, что система переходит в состояние низкой производительности, если резерв исчерпан и нет возможности для замены отказавшей машины. В рамках ТМО рассматривается математическая модель, формализованная системой уравнений для вероятностей состояний системы отказавших машин, где под входящим потоком понимаются отказы с заданной интенсивностью, а восстановление группами описывается исходящим потоком.

В работе предложены аналитические решения для оценки надежности вычислительной системы, которая в любой момент времени может находиться в одном из двух состояний, высокой производительности или низкой производительности. В частности, предложена функция для нахождения вероятности пребывания масштабируемой вычислительной системы в состоянии низкой производительности в течение определенного времени при групповом восстановлении и заданном числе резервных машин.

Работа выполнена при финансовой поддержке РФФИ (грант № 20-07-00039).

Список литературы

[1] Хорошевский В. Г. Архитектура вычислительных систем / М: МГТУ им. Баумана, 2008. 520 с.

[2] GUPTA S., PATEL T., ENGELMANN C., TIWARI D. Failures in large scale systems: long-term measurement, analysis, and implications // SC '17: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis. Denver, Colorado — November 12 - 17, 2017. N. 44.

[3] MOR HARCHOL-BALTER Performance Modeling and Design of Computer Systems: Queueing Theory in Action / Cambridge University Press, 2013.