


УДК 519.248+519.872.8

 10.25209/2079-3316-2023-14-1-55-94

Исследование стационарных характеристик многосерверных моделей с избыточностью

Александр Сергеевич Румянцев^{1✉}, Диана Сергеевна Долгалева²,
Александр Станиславович Головин³

Институт прикладных математических исследований КарНЦ РАН, Петрозаводск, Россия

^{1✉}ar0@krc.karelia.ru

Аннотация. В работе предложены подходы к исследованию стохастических моделей вычислительных сетей из персональных компьютеров с помощью многосерверных систем обслуживания с избыточностью. Представлены перспективы и ограничения указанного подхода, а также возможные направления дальнейших исследований.

Ключевые слова и фразы: вычислительные сети из персональных компьютеров, многосерверные системы обслуживания, стохастическое моделирование, системы с избыточностью.

Благодарности: Авторы благодарны рецензентам за ценные замечания. Исследование выполнено за счет гранта РФФ 21-71-10135

Для цитирования: Румянцев А. С., Долгалева Д. С., Головин А. С. *Исследование стационарных характеристик многосерверных моделей с избыточностью* // Программные системы: теория и приложения. 2023. Т. 14. № 1(56). С. 55–94. https://psta.psisras.ru/read/psta2023_1_55-94.pdf

1. Введение

Системы распределенных вычислений, такие как вычислительные сети из персональных компьютеров типа Desktop Grid (ВСПК), ВСПК уровня предприятия (Enterprise Desktop Grid), являются актуальным инструментом для объединения простаивающих ресурсов неспециализированных вычислителей (компьютеров, серверов) с целью решения вычислительно-массивных задач [1]. Потенциально высокое быстродействие таких сетей компенсируется низкой надежностью, труднопрогнозируемым временем завершения расчетов, а также существенным ограничением на класс решаемых задач (решение которых можно разделить на множество слабо связанных подзадач, не требующих межпроцессорного обмена).

Для повышения надежности сети и достоверности получаемых результатов, в указанных сетях применяется *избыточность* на уровне подзаданий за счет *репликации* (создания копий подзадания с одними и теми же исходными данными) и *кворума* (числа копий, необходимых для успешного прохождения валидации и фиксации результата).

Избыточность информации является способом обеспечения надежной передачи данных между узлами вычислительных систем с помощью сетей передачи данных (в том числе беспроводных) [2]. Подобная задача возникает и в передовых системах хранения, основанных на технологии твердотельных накопителей, которые, в отличие от традиционных систем, обладают значительным внутренним резервом параллелизма, встроенной избыточностью, но существенно меньшим временем наработки на отказ отдельных модулей хранения [3]. В этой связи контролируемая избыточность, дублирование информации и высокая пропускная способность являются целевыми функциями для оптимизации параметров таких носителей [4].

Для анализа стационарных характеристик производительности телекоммуникационных и вычислительных систем целесообразно применять методы теории массового обслуживания в связи со случайным характером исследуемых процессов. Итогом применения указанных методов должно быть решение соответствующих задач по поиску оптимальных параметров конфигурации систем с избыточностью. Целью данной работы является анализ перспектив исследования класса многосерверных моделей с избыточностью, называемого системами разделения-слияния (split-merge), применяемых в качестве моделей вычислительных систем и систем передачи данных [5, 6]. Далее в работе кратко резюмированы основные результаты в области анализа систем разделения-слияния, получены новые результаты для стохастической упорядоченности ключевых характеристик и предложены направления дальнейших исследований указанного класса систем.

2. Обобщенные системы типа разделение-слияние

В данном разделе объектом исследования являются многосерверные системы обслуживания, в которых избыточность организована на уровне заявок (клиентов), поступающих в систему, с помощью репликации и кворума заявок. Такие системы являются моделями ВСПК (в том числе систем уровня предприятия) и ряда других систем (например, системы организации надежного хранения типа RAID [7]). Для исследования данного вида систем параллельного обслуживания можно применять многосерверные системы массового обслуживания с разделением и слиянием заявок. Именно, в момент поступления в систему заявка разделяется на несколько экземпляров, которые в дальнейшем обслуживаются в соответствии с выбранной дисциплиной. В соответствии с терминологией ВСПК, будем называть *подзаданием* экземпляр заявки на данном сервере. В зависимости от дисциплины обслуживания подзаданий, выделяют следующие виды систем:

- классические системы типа *разветвление-соединение* (Fork-Join, FJ), в которых подзадания попадают на независимые неограниченные очереди однотипных серверов, а обслуживание заявки завершается лишь в момент окончания обслуживания всех соответствующих подзаданий (в ряде случаев для целей накопления результатов выполнения подзаданий также рассматривается отдельная очередь, расположенная после серверов) [8–11];
- обобщенные системы типа *разветвление-соединение*, обозначаемые (r, q) -FJ, в которых для каждой заявки создаются r реплик (подзаданий), а окончание обслуживания соответствует кворуму из q результатов, при этом невостребованные $r - q$ подзаданий снимаются с расчетов/из очередей (в частности, классической FJ является система (r, r) -FJ) [12];
- обобщенные системы типа *разветвление-ранняя отмена* (Fork-Early-Cancel), обозначаемые (r, q) -FEC, в которых из r реплик, в момент начала обслуживания не менее q подзаданий, невостребованные $r - q$ подзаданий снимаются с расчетов/из очередей [12, 13];
- обобщенные системы типа *разделение-слияние* (Split-Merge), обозначаемые (c, r, q) -SM, в которых имеется одна общая очередь на обслуживание и c групп по r однотипных серверов каждая, подзадания каждого клиента ожидают освобождения (группы) серверов от подзаданий предшествующих клиентов и начинают

обслуживание на данной группе серверов одновременно, а окончанием обслуживания является достижение кворума q (невостребованные подзадания данной заявки снимаются с расчетов).

Отметим, что вероятностный анализ классических FJ-систем затруднен, поскольку входные потоки на сервера коррелированы, а размер очереди каждого сервера неограничен, что означает, потенциальную потребность в хранении неограниченного числа компонент для описания системы в виде цепи Маркова. В этой связи, как правило, для анализа классических FJ-систем прибегают к численному анализу или аппроксимации [7]. Однако, в ряде важных частных случаев удается свести анализ к известным классическим системам. В частности, в работах [12–14] доказаны следующие важные результаты:

- $(r, 1)$ -FJ система с типичным временем обслуживания подзадания S эквивалентна классической односерверной системе с типичным временем обслуживания $S_{1,r}$;
- $(r, 1)$ -FEC система эквивалентна классической r -серверной системе обслуживания;
- время пребывания (время отклика) в (r, q) -FJ системе ограничено сверху временем отклика в $(1, r, q)$ -SM системе;
- время пребывания (время отклика) в (c, r, q) -SM системе с типичным временем обслуживания подзадания S совпадает со временем пребывания в классической c -серверной системе с типичным временем обслуживания $S_{q;r}$,

где $S_{q;r}$ есть q -я порядковая статистика из r независимых одинаково распределенных случайных величин (с.в.), копий S . Это означает, в частности, что $(1, r, 1)$ -SM система эквивалентна $(r, 1)$ -FJ системе, а $(c, 1, 1)$ -SM система эквивалентна $(c, 1)$ -FEC системе. В этой связи далее приведем результаты для систем с разделяющимися клиентами типа (c, r, q) -SM.

Рассмотрим cr -серверную систему обслуживания, в которую с интенсивностью λ поступает поток заявок. Поступающая в систему, заявка ожидает одновременного освобождения группы из r серверов и создает r подзаданий, типичная (случайная) длительность которых, S , положительна. Подзадания данного клиента занимают освободившиеся r серверов одновременно, а при достижении кворума q (при окончании обслуживания q подзаданий), невостребованные $r - q$ подзаданий покидают сервера немедленно. Это означает, что группа из r серверов занимается только

подзданиями данного клиента, разделения данной группы серверов между различными клиентами не происходит. Таким образом, рассматриваемая система эквивалентна классической системе $G/G/c$, в которой типичное время обслуживания является порядковой статистикой $S_{q:r}$. Система стационарна в случае выполнения неравенства [15]

$$(1) \quad \lambda E S_{q:r} < c.$$

Целью анализа является, как правило, получение следующих ключевых характеристик производительности и эффективности систем в стационарном режиме в зависимости от числа серверов cr , параметров репликации r и кворума q (эти параметры использованы в обозначениях):

время отклика системы $V_{r,q}$ – время с момента попадания клиента в систему до момента окончания его обслуживания (получения результатов q подзаданий)

$$(2) \quad V_{r,q} = W_{r,q} + S_{q:r},$$

где $W_{r,q}$ есть стационарная задержка в системе;

стоимость вычислений $C_{r,q}$ – ценовая характеристика, пропорциональная реальному времени вычислений клиента,

$$(3) \quad C_{r,q} = r S_{q:r}.$$

Отметим, что для ВСПК под стоимостью может подразумеваться энергопотребление, в системах облачных вычислений (в которых ресурсы размещаются в распределенном разделяемом пуле) стоимость представляет затраты в денежном эквиваленте, в системах хранения данных на основе твердотельных накопителей стоимость определяется как физический износ накопителя.

2.1. Стохастическая упорядоченность характеристик

Свойства указанных характеристик находятся в сложной зависимости от параметров репликации r и кворума q и, в частности, это связано с особенностями распределения порядковых статистик. Пусть с.в. S имеет функцию распределения F . Известно [16], что функция распределения порядковой статистики $S_{q:r}$ имеет вид

$$(4) \quad F_{q:r}(x) = P\{S_{q:r} \leq x\} = \sum_{i=q}^r \binom{r}{i} F^i(x) \bar{F}^{r-i}(x),$$

где $\bar{F}(x) = 1 - F(x)$ есть хвост распределения. Равенство (4) означает, в частности, что имеет место следующее (очевидное) стохастическое неравенство:

$$S_{1:r} \leq_{st} S \leq_{st} S_{r:r}.$$

Кроме того, при фиксированном r для $q > 1$ имеет место следующее равенство [17]:

$$(5) \quad P(S_{q:r} \leq x) = P(S_{q-1:r} \leq x) - \binom{r}{q-1} F^{q-1}(x) \bar{F}^{r-q+1}(x).$$

Это означает, в частности, что

$$(6) \quad S_{q:r} \geq_{st} S_{q-1:r}.$$

Кроме того, при фиксированных $r, q > 1$, имеет место следующее равенство [17]

$$(7) \quad P(S_{q:r} \leq x) = P(S_{q-1:r-1} \leq x) - \binom{r-1}{q-1} F^{q-1}(x) \bar{F}^{r-q+1}(x).$$

В частности, это означает, что

$$(8) \quad S_{q:r} \geq_{st} S_{q-1:r-1}.$$

Наконец, при $q \geq 1$ выполнено [17]

$$(9) \quad P(S_{q:r} \leq x) = P(S_{q:r-1} \leq x) + \binom{r-1}{q-1} F^q(x) \bar{F}^{r-q}(x).$$

Это означает, что $S_{q:r} \leq_{st} S_{q:r-1}$, что с учетом (6) и (8) окончательно влечет

$$(10) \quad S_{q:r} \geq_{st} S_{q-1:r-1} \geq_{st} S_{q-1:r}.$$

Из (3) и (6) следует, что $C_{r,q}$ монотонно неубывает с возрастанием q (при фиксированном r), т.е.

$$(11) \quad C_{r,q} \geq_{st} C_{r,q-1}, \quad q > 1.$$

Из (8) следует, что $C_{r,q}$ монотонно неубывает при одновременном увеличении q и r , т.е.

$$(12) \quad C_{r,q} \geq_{st} C_{r-1,q-1}, \quad q, r > 1.$$

В то же время, в общем случае при фиксированном q с увеличением r величина $C_{r,q}$ немонотонна, что следует из (10).

Для исследования монотонности времени пребывания в системе в зависимости от параметров репликации и кворума, нам потребуется следующий результат, являющийся прямым следствием Теоремы 4 в работе [18]:

ТЕОРЕМА 1. Пусть в системах $\Sigma^{(1)}$ и $\Sigma^{(2)}$ (управляющие последовательности которых обозначены соответствующим верхним индексом) имеет место следующий стохастический порядок для типичных времен между приходами T и времен обслуживания S ,

$$(13) \quad T^{(1)} \geq_{st} T^{(2)}, \quad S^{(1)} \leq_{st} S^{(2)}.$$

Тогда стационарные задержки W в системах упорядочены следующим образом,

$$(14) \quad W^{(1)} \leq_{st} W^{(2)}.$$

Тогда из равенства (2), Теоремы 1 и неравенства (10) следует, что справедливо утверждение.

УТВЕРЖДЕНИЕ 1. Для $r, q > 1$ имеет место следующий стохастический порядок для стационарных времен отклика,

$$(15) \quad V_{r,q} \geq_{st} V_{r-1,q-1} \geq_{st} V_{r,q-1}.$$

Отметим, что стохастический порядок влечет, в частности, упорядоченность средних значений.

В общем случае выражение для среднего времени пребывания в системе $EV_{r,q}$ для системы типа $G/G/c$ неизвестно. Явный вид среднего времени отклика можно получить в случае $c = 1$, если входной поток является пуассоновским. Для этого можно воспользоваться формулой Поллачека–Хинчина

$$(16) \quad EW_{r,q} = \frac{\lambda ES_{q:r}^2}{2(1 - \lambda ES_{q:r})},$$

и получить из (3) выражение

$$(17) \quad EV_{r,q} = E(S_{q:r}) + \frac{\lambda E(S_{q:r}^2)}{2[1 - \lambda E(S_{q:r})]},$$

которое может быть использовано для расчета среднего времени отклика, если известны соответствующие выражения для моментов порядковых статистик.

Для иллюстрации стохастического порядка (15), построим графики функций распределения времени отклика в системе $M/G/c$ при количестве серверов $c = 3$, интенсивности входного потока $\lambda = 1$ и параметрах избыточности $q = 3$, $r = 5$ для случая, когда время обслуживания подзадачи имеет распределение Парето с плотностью

$$(18) \quad f(x) = \alpha x_0^\alpha x^{-\alpha-1}, \quad x > x_0, \quad (f(x) = 0, x \leq x_0).$$

Положим $\alpha = 5$, $x_0 = 1$ и заметим, что условие стационарности системы выполнено. Для указанных систем величина нагрузки $\rho_{r,q} = \lambda E S_{q;r}$ равна $\rho_{5,3} \approx 0.39$, $\rho_{4,2} \approx 0.38$ и $\rho_{5,2} = 0.37$, соответственно. На рисунке 1 видна упорядоченность функций распределения, соответствующая стохастическому порядку (15). Графики построены с помощью метода идеальной выборки [19].

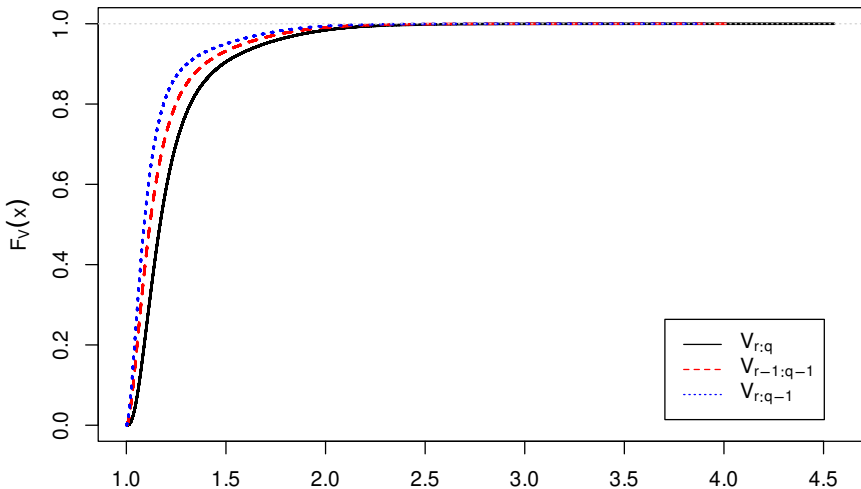


Рисунок 1. График функций распределения для системы $M/G/c$ при единичной интенсивности входного потока на 3 сервера

2.2. Фазовый тип распределения времени обслуживания

Стоимость вычислений (3) может быть получена в явном виде для рассматриваемой модели в случае, когда время обслуживания имеет распределение фазового типа $PH(\beta, \mathbf{B})$, т.е. время обслуживания соответствует времени до поглощения конечной цепи Маркова с начальным распределением β и (субгенераторной) матрицей интенсивностей переходов между m невозвратными состояниями \mathbf{B} (подробнее об этом плотном классе распределений см., напр. [20]). В этом случае можно показать (см. [20, 21]), что распределение порядковой статистики $S_{q:r}$ имеет вид $PH(\beta_{q:r}, \mathbf{B}_{q:r})$, где

$$(19) \quad \mathbf{B}_{q:r} = \begin{bmatrix} \mathbf{B}^{\oplus r} & \mathbf{b}_0^{\oplus r} & \mathbf{O} & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \mathbf{B}^{\oplus r-1} & \mathbf{b}_0^{\oplus r-1} & \mathbf{O} & \dots & \mathbf{O} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & \mathbf{O} & \dots & \mathbf{B}^{\oplus r-q+1} \end{bmatrix},$$

где неквадратные матрицы $\mathbf{b}_0^{\oplus k}$ содержат интенсивности перехода между состояниями в связи с поглощением одной компоненты (реализацией одной из k независимых одинаково распределенных с.в. с распределением $PH(\beta, \mathbf{B})$),

$$(20) \quad \mathbf{b}_0^{\oplus k} = \mathbf{b}_0 \otimes \mathbf{I}_{m^{k-1}} + \mathbf{I}_m \otimes \mathbf{b}_0 \otimes \mathbf{I}_{m^{k-2}} + \dots + \mathbf{I}_{m^{k-1}} \otimes \mathbf{b}_0,$$

\mathbf{I}_j есть единичная матрица размера $j \geq 1$, \otimes есть кронекерово произведение и \oplus есть кронекерова сумма матриц (соответственно, кратные произведение и сумма указаны в качестве верхнего индекса), а $\mathbf{b}_0 = -\mathbf{B}\mathbf{1}$ есть вектор интенсивностей поглощения цепи, соответствующей распределению $PH(\beta, \mathbf{B})$. Начальное распределение процесса можно задать в следующем виде

$$(21) \quad \beta_{q:r} = (\beta^{\otimes r}, \mathbf{0}),$$

Это означает, что класс распределений фазового типа замкнут относительно порядковых статистик, а следовательно, условие стационарности (1) и стоимость $EC_{r,q}$, определенная в (3), принимают явный вид. Действительно, известно, что для с.в. X с распределением $PH(\hat{\beta}, \hat{\mathbf{B}})$, моменты первого и

второго порядка имеют вид [22]

$$(22) \quad EX = -\hat{\beta}\hat{B}^{-1}\mathbf{1}, \quad EX^2 = 2\hat{\beta}\hat{B}^{-2}\mathbf{1}.$$

Тогда, с учетом (22),

$$(23) \quad EC_{r,q} = rES_{q:r} = -r\beta_{q:r}\mathbf{B}_{q:r}^{-1}\mathbf{1},$$

где $\beta_{q:r}$ и $\mathbf{B}_{q:r}$ определены в (21) и (19), соответственно.

Рассмотрим важный частный случай: пусть время обслуживания подзадания распределено экспоненциально с параметром μ . Тогда $\beta = 1$, $\mathbf{B} = -\mu$, $\mathbf{b}_0 = \mu$ и можно преобразовать (23) к следующему виду

$$(24) \quad EC_{r,q} = \frac{1}{\mu} \sum_{i=0}^{q-1} \frac{r}{r-i}.$$

В частности, при $q = 1$,

$$(25) \quad P\{S_{1:r} \leq x\} = 1 - e^{-r\mu},$$

поэтому выражение (24), примет вид

$$(26) \quad EC_{r,1} = \mu^{-1},$$

т.е. стоимость вычислений не зависит от параметра r при $q = 1$.

В общем случае,

$$(27) \quad EC_{r+1,q} - EC_{r,q} = -\frac{1}{\mu} \left(\sum_{i=1}^{q-1} \frac{1}{(r-i+1)(r-i)} \right) \leq 0.$$

Значит, стоимость монотонно убывает по r при фиксированном $q \geq 1$.

Если входной поток в систему является пуассоновским (либо так называемым марковским входным потоком), то время отклика системы может быть найдено в явном виде, но на практике это может быть выполнено лишь численно. В ряде частных случаев это удается сделать явно.

Рассмотрим важный частный случай $c = 1$, т.е. репликация происходит на все узлы системы, и пусть входной поток в систему пуассоновский. Тогда $(1, r, q)$ -SM система эквивалентна системе типа $M/PH/1$, для

которой время отклика (17), с учетом (22), примет вид

$$(28) \quad EV_{r,q} = \frac{\lambda \beta_{q:n} \mathbf{B}_{q:n}^{-2} \mathbf{1}}{1 + \lambda \beta_{q:n} \mathbf{B}_{q:n}^{-1} \mathbf{1}} - \beta_{q:r} \mathbf{B}_{q:r}^{-1} \mathbf{1}.$$

Если время обслуживания имеет экспоненциальное распределение с параметром μ , и при этом $q = 1$, то анализ упрощается за счет выражения (25). В частности, в системе $(c, r, 1)$ -SM для расчета времени отклика можно применить формулу Эрланга [23],

$$(29) \quad EV_{r,1} = \frac{1}{r\mu} + \frac{C(c, \lambda/(r\mu))}{c\mu - \lambda},$$

где так называемая С-формула Эрланга $C(n, \rho)$ задает вероятность попадания в очередь в момент прихода в системе M/M/n при нагрузке $\rho < n$,

$$(30) \quad C(n, \rho) = \left[1 + (1 - \rho) \frac{n!}{(n\rho)^n} \sum_{k=0}^{n-1} \frac{(n\rho)^k}{k!} \right]^{-1}.$$

В частности, время отклика системы $(c, 1, 1)$ -SM совпадает с классическим временем отклика системы M/M/c:

$$(31) \quad EV_{1,1} = \frac{1}{\mu} + \frac{C(c, \lambda/\mu)}{c\mu - \lambda}.$$

Из (30) видно, что $C(1, \rho) = \rho$, поэтому время отклика (29) в системе $(1, r, 1)$ -SM примет вид

$$EV_{r,1} = \frac{1}{r\mu} + \frac{\lambda}{r\mu(r\mu - \lambda)}.$$

2.3. Распределение времени обслуживания с тяжелым хвостом

С учетом (9), зависимость стоимостной характеристики $C_{r,q}$ от параметров r, q необходимо исследовать для отдельных классов распределений. В данном разделе величина средней стоимости $EC_{r,q}$ исследована для распределений Парето и Вейбулла, которые относятся (в последнем случае – при определенных параметрах) к классу распределений с тяжелым хвостом. Отметим, что такие распределения характерны для многих приложений в области вычислительных и телекоммуникационных систем [24].

Пусть типичное время обслуживания подзадания S имеет распределение Парето (18) с параметром $\alpha > 1$. Свойства порядковых статистик для с.в. с распределением Парето получены в работах [25, 26], именно,

$$(32) \quad \mathbb{E}(S_{q:r}^k) = x_0^k \frac{r!}{(r-q)!} \frac{\Gamma(r-q+1-k/\alpha)}{\Gamma(r+1-k/\alpha)}, \quad 0 < k < \alpha(r-q+1),$$

где Γ является гамма-функцией. Из (3) и (32) можно получить явный вид $EC_{r,q}$, а также следующее рекуррентное соотношение для величины $EC_{r,q}$ [6]:

$$(33) \quad EC_{r+1,q} = \frac{(r+1)^2(r-q+1-1/\alpha)}{r(r-q+1)(r+1-1/\alpha)} EC_{r,q}.$$

Преобразовывая выражение (33), можно получить, что $EC_{r,q}$ *возрастает* по r , если выполнено неравенство

$$(34) \quad \alpha r^2 + (2\alpha - \alpha q - q - 1)r - \alpha q + \alpha - 1 > 0,$$

в противном случае $EC_{r,q}$ *невозрастает* по r . При $\alpha > 1$ и фиксированном $q \geq 1$, квадратный трехчлен (относительно r) в левой части (34) имеет два вещественных корня, левый из которых отрицателен, а правый положителен и имеет вид [6]:

$$(35) \quad r_0(q) = \frac{1+q+\alpha q - 2\alpha + \sqrt{(\alpha q + q - 1)^2 + 4q}}{2\alpha}.$$

Таким образом, при заданном фиксированном q неравенство (34) выполнено при (целых) $r > r_0(q)$, а минимум стоимости достигается в одном из двух ближайших целых к $r_0(q)$.

Из (35), в частности, следует, что $r_0(1) < 1$ при $\alpha > 1.5$, т.е. стоимость $EC_{r,1}$ возрастает по r для всех $r \geq 1$. Действительно,

$$(36) \quad EC_{r,1} = x_0 \frac{r^2 \alpha}{r\alpha - 1} = x_0 \left(r + \frac{1}{\alpha} + \frac{1}{\alpha(r\alpha - 1)} \right), \quad r > 1/\alpha.$$

Кроме того, из (36) следует, что имеет место асимптотически линейный рост стоимости

$$\lim_{r \rightarrow \infty} r^{-1} EC_{r,1} = x_0.$$

Из (36) следует, что $EC_{1,1}$ может быть сколь угодно большим, если α близко к 1. Действительно, из (32) или из (36) видно, что

$$EC_{1,1} = ES = x_0 \frac{\alpha}{\alpha - 1}.$$

В то же время, из (36) следует, что

$$EC_{2,1} < x_0 \left(2 + \frac{2}{\alpha} \right),$$

поскольку $\alpha r - 1 > 1$ для $r \geq 2$. Это означает, в частности, что $EC_{2,1} < EC_{1,1}$ для $\alpha < \sqrt{2}$. Таким образом, с практической точки зрения репликация может быть полезна для существенного снижения стоимости вычислений при $r < r_0(q)$.

В отношении *времени отклика*, справедлив общий результат (15) о монотонности по параметрам избыточности, а также, с учетом (17) и (32), доступен явный вид в системе типа $M/G/1$. Однако более важной (в присутствии распределений с тяжелым хвостом) является проблема обеспечения конечности моментов (например, среднего или дисперсии) времени отклика системы. Конечность моментов, позволяет обеспечить надежное (доверительное) оценивание характеристик системы в стационарном режиме и (конечность в среднем) является необходимой для стационарности системы.

Обозначим $\kappa(X)$ *моментный индекс* с.в. X , т.е. [27, 28]

$$(37) \quad \kappa(X) = \sup\{k > 0 : E(X^k) < \infty\} \leq \infty.$$

Тогда из (18) следует, что с.в. S , имеющая распределение Парето, имеет моментный индекс α ,

$$(38) \quad \kappa(S) = \alpha,$$

при этом (32) означает, что

$$(39) \quad \kappa(S_{q;r}) = \alpha(r - q + 1) \geq \alpha.$$

Из (39) следует, что $\kappa(S_{q;r})$ *линейно растет* с ростом $r - q$, т.е. с помощью параметров избыточности q, r можно обеспечить конечность требуемого момента (например, дисперсии) времени обслуживания заявки даже при малых значениях моментных индексов подзаданий (например, в случае если подздания имеют неограниченную дисперсию длительности выполнения).

Известно, что моментный индекс времени ожидания в многосерверной системе типа $G/G/c$ определяется моментным индексом времени обслуживания [29]. В рассматриваемой системе для заданных λ, α, c, r, q

моментный индекс времени ожидания $W_{r,q}$ имеет вид [6]

$$(40) \quad \kappa(W_{r,q}) = (\alpha(r - q + 1) - 1)(c - \lfloor \rho \rfloor) \geq \alpha - 1,$$

где $\rho = \lambda \mathbb{E}S_{q:r}$, а $\mathbb{E}S_{q:r}$ определено в (32). Тогда из (3), используя неравенство Минковского [30]

$$(\mathbb{E}(W_{r,q} + S_{q:r})^k)^{1/k} \leq (\mathbb{E}W_{r,q}^k)^{1/k} + (\mathbb{E}S_{q:r}^k)^{1/k},$$

можно заключить, что

$$(41) \quad \kappa(V_{r,q}) \geq \min(\kappa(W_{r,q}), \kappa(S_{q:r})),$$

где моментные индексы $\kappa(W_{r,q})$ и $\kappa(S_{q:r})$ определены в (40) и (39), соответственно.

Равенство (40) при $c = 1$ и $r = q$ согласуется с формулой Поллачека–Хинчина (16), из которой следует, что для обеспечения конечности среднего времени ожидания необходима конечность *второго момента* времени обслуживания. В заключение заметим, что линейный рост моментного индекса $S_{q:r}$ по $r - q$, с учетом (41) и (40), может быть использован для подбора значений r, q таким образом, чтобы гарантировать требуемое значение моментного индекса (например, конечность дисперсии) времени пребывания клиента в системе даже если моментный индекс $\kappa(S) = \alpha$ подзадания достаточно мал (например, при α близком к единице).

Повторим анализ для системы, в которой типичное время выполнения подзадания S имеет распределение Вейбулла,

$$(42) \quad F(x) = 1 - e^{-x^\xi}, \quad \xi > 0.$$

Тогда моменты порядка $k > 0$ для $S_{q:r}$ могут быть получены в следующем виде [31, 32]

$$(43) \quad \mathbb{E}(S_{q:r}^k) = q \binom{r}{q} \Gamma(1 + k/\xi) \sum_{j=0}^{q-1} (-1)^j \binom{q-1}{j} (r - q + 1 + j)^{-1 - k/\xi}.$$

Таким образом, указанные моменты конечны для любого $k > 0$, т.е. $\kappa(S_{q:r}) = \infty$, а выражение (17) позволяет найти явный вид для $\mathbb{E}V_{r,q}$ для системы типа $M/G/1$. Кроме того, имеет место следующее рекуррентное соотношение [17, 33]

$$(44) \quad \mathbb{E}(S_{q:r}) = \mathbb{E}(S_{q-1:r-1}) + \binom{r-1}{q-1} J(r - q + 1, q - 1),$$

где $ES_{1:1} = J(1, 0) = \Gamma(1 + k/\xi)$, а $J(p, m)$ определяется рекуррентно:

$$J(p, m) = J(p, m - 1) + J(p + 1, m - 1), \quad J(p, 0) = p^{-1/\xi} \Gamma(1 + 1/\xi).$$

С учетом (43), имеем

$$(45) \quad EC_{r,q} = r \frac{r!}{(r-q)!} \frac{\Gamma(\xi^{-1})}{\xi} \sum_{j=0}^{q-1} (-1)^j \frac{(r-q+1+j)^{-1-\xi^{-1}}}{j!(q-1-j)!}.$$

Из (45) видно, что

$$(46) \quad EC_{r,1} = \Gamma(1 + 1/\xi) r^{1-1/\xi}, \quad r \geq 1.$$

Это означает, в частности, что $EC_{r,1}$ неубывает (невозрастает) по r при $\xi \geq 1$ ($\xi \leq 1$, соответственно). Кроме того, при $\xi = 1$ распределение Вейбулла совпадает с экспоненциальным распределением, и при этом (46) согласуется с (26).

Напрямую использовать (45) для установления монотонности $EC_{r,q}$ затруднительно. Однако, рассматривая разность $\delta(r) = EC_{2,r+1} - EC_{2,r}$ и воспользовавшись рекурсией (44), можно установить, что

$$\delta(r) = \Gamma(1 + 1/\xi) \left((2r^2 + r + 1)r^{-1/\xi} - r(r+1)^{1-1/\xi} - r^2(r-1)^{-1/\xi} \right).$$

В частности, при $\xi < 1$ величина $EC_{2,r}$ невозрастает по r . Действительно, исследуем функцию

$$\psi(r) = (2r^2 + r + 1)r^{-1/\xi} - r(r+1)^{1-1/\xi} - r^2(r-1)^{-1/\xi}.$$

Обозначим $a = \xi^{-1} > 1$. Преобразуем функцию к следующему виду

$$(47) \quad \psi(r) = r^2(2r^{-a} - (r+1)^{-a} - (r-1)^{-a}) + \phi(r),$$

где

$$\phi(r) = r(r+1)(r^{-a-1} - (r+1)^{-a-1}) = r^{1-a} + r^{-a} - (r+1)^{1-a} + (r+1)^{-a}.$$

Из выпуклости функции x^{-a} следует, что $2r^{-a} - (r+1)^{-a} - (r-1)^{-a} \leq 0$. Кроме того, можно убедиться, что $\partial\phi(r)/\partial r \leq 0$ для всех $a > 1$, т.е. функция $\phi(r)$ убывает по $r \geq 2$. При этом можно проверить, что $\psi(2) \leq 0$ при любом $\xi \in (0, 1)$. Таким образом, из (47) следует, что рассматриваемая разность $\delta(r)$ неположительна для всех $r \geq 2$, т.е. $EC_{2,r}$ монотонно невозрастает по r при любом $\xi \in (0, 1)$.

3. Итоги и перспективы исследований

Таблица 1 представляет результаты описанного выше исследования влияния параметров репликации r и кворума q на стоимость $C_{r,q}$ и время отклика $V_{r,q}$ в многосерверных моделях систем с избыточностью типа $(c, r, q) - SM$ для различных распределений времени обслуживания подзаданий S . Вид монотонности указан при увеличении r, q , в ряде случаев результат доступен для среднего значения, либо для соответствующей случайной величины (стохастический порядок). Результаты получены для случаев, когда время обслуживания имеет фазовый тип распределения, либо распределение Парето или Вейбулла (относящиеся к классу распределений с тяжелым хвостом).

Таблица 1. Результаты представленного исследования

Распределение	Характеристика	
	$C_{r,q}$	$V_{r,q}$
Общий случай	Немонотонна	Неубывает, неравенство (15)
Экспоненциальное	Невозрастает, неравенство (27); постоянна при $q = 1$, формула (26); явный вид (24)	Явный вид (29) в системе типа $M/M/c$ при $q = 1$
Фазовый тип	Явный вид (23)	Явный вид (28) в системе типа $M/PN/1$
Парето	Глобальный минимум, выражение (35)	Явный вид в системе типа $M/G/1$ с учетом (17) и (32)
Вейбулла с параметром ξ	Невозрастает, если $\xi \leq 1, q = 1, 2$; неубывает при $\xi \geq 1, q = 1$; выражение (46)	Явный вид в системе типа $M/G/1$ с учетом (17) и (43)

С учетом полученных результатов, можно выделить следующие направления их дальнейшего развития:

- монотонность средней стоимости (27) получена для частного случая экспоненциально распределенных времен обслуживания подзадач, при этом интерес представляет общий случай произвольного фазового типа распределений;

- было бы целесообразно рассмотреть также неоднородный случай, например, когда сервера, на которые осуществляется репликация, имеют разную производительность (при этом корректировки требует формула (19) и дальнейшие результаты);
- монотонность стоимости для системы с распределением Вейбулла показана в разделе 2.3 для $r = 1, 2$, случай произвольного r требует отдельного исследования;
- наконец, в модели предполагается, что число блоков серверов c неизменно, при этом r является изменяемым параметром, целесообразно исследовать случай, когда число блоков определяется как s/r (целочисленно), где s является (неизменным) числом серверов в исходной системе.

Последнее направление открывает возможности для использования стохастических сравнений для систем с различным числом серверов, однако интересна также параметрическая монотонность систем (например, по параметру α для распределения Парето или ξ для распределения Вейбулла).

Список литературы

- [1] Chernov I., Nikitina N., Ivashko E. *Task scheduling in desktop grids: open problems* // Open Engineering.– 2017.– Vol. **7**.– No. 1.– Pp. 343–357. doi ↑56
- [2] Verma N., Singh D. *Data redundancy implications in wireless sensor networks* // Procedia Computer Science.– 2018.– Vol. **132**.– Pp. 1210–1217. doi ↑56
- [3] Rummyantsev A., Ivashko E., Chernov I., Kositsyn D., Shabaev A., Ponomarev V. *Latency/wearout in a flash-based storage system with replication on write*, 2019 24th Conference of Open Innovations Association (FRUCT) (08–12 April 2019, Moscow, Russia).– Pp. 360–366. doi ↑56
- [4] Chernov I., Ivashko E., Kositsyn D., Ponomarev V., Rummyantsev A., Shabaev A. *Flash-based storage deduplication techniques: A survey* // International Journal Of Embedded And Real-Time Communication Systems (IJERTCS).– 2019.– Vol. **10**.– No. 3.– Pp. 32–48. doi ↑56
- [5] Osipov O., Rogachko E. *Fork-join queueing systems with heterogeneous servers threshold control policy*, SMARTY2020 (August 16–20, 2020, Petrozavodsk, Russia), CEUR Workshop Proc.– vol. **2792**.– 2020.– Pp. 76–88. URL ↑56
- [6] Rummyantsev A., Chakravarthy S., Morozov E., Remnev S. *Cost and effect of replication and quorum in desktop grid computing* // Information Technologies And Mathematical Modelling. Queueing Theory And Applications, Springer CCIS Series.– vol. **912**, Cham: Springer.– 2018.– ISBN 978-3-319-97594-8.– Pp. 143–156.

- [7] Thomasian A. *Analysis of fork/join and related queueing systems* // ACM Computing Surveys.– 2014.– Vol. **47**.– No. 2.– id. 17.– 71 pp. doi ↑^{57, 58}
- [8] Baccelli F., Makowski A., Shwartz A. *The fork-join queue and related systems with synchronization constraints: Stochastic ordering and computable bounds* // Advances in Applied Probability.– 1989.– Vol. **21**.– No. 3.– Pp. 629–660. doi ↑⁵⁷
- [9] Gardner K., Zbarsky S., Doroudi S., Harchol-Balter M., Hyytiä E., Scheller-Wolf A. *Queueing with redundant requests: Exact analysis* // Queueing Systems.– 2016.– Vol. **83**.– No. 3–4.– Pp. 227–259. doi ↑⁵⁷
- [10] Kim C., Agrawala A. K. *Analysis of the fork-join queue* // IEEE Transactions on Computers.– 1989.– Vol. **38**.– No. 2.– Pp. 250–255. doi ↑⁵⁷
- [11] Nelson R., Tantawi A. N. *Approximate analysis of fork/join synchronization in parallel queues* // IEEE Transactions on Computers.– 1988.– Vol. **37**.– No. 6.– Pp. 739–743. doi ↑⁵⁷
- [12] Joshi G., Soljanin E., Wornell G. *Queues with redundancy: Latency-cost analysis* // ACM SIGMETRICS Performance Evaluation Review.– 2015.– Vol. **43**.– No. 2.– Pp. 54–56. doi ↑^{57, 58}
- [13] Joshi G. *Efficient redundancy techniques to reduce delay in cloud systems*, PhD Thesis.– Massachusetts Institute of Technology.– 2016. ↑^{57, 58}
- [14] Joshi G., Soljanin E., Wornell G. *Efficient redundancy techniques for latency reduction in cloud systems* // ACM Transactions on Modeling and Performance Evaluation of Computing Systems.– 2017.– Vol. **2**.– No. 2.– id. 12.– 30 pp. doi ↑⁵⁸
- [15] Asmussen S. *Applied Probability and Queues*, SMAP.– Vol. **51**.– Springer.– 2003.– ISBN 978-0-387-00211-8.– 452 pp. doi ↑⁵⁹
- [16] Balakrishnan N. *Permanents, order statistics, outliers and robustness* // Revista Matemática Complutense.– 2007.– Vol. **20**.– No. 1.– Pp. 7–107. doi ↑⁵⁹
- [17] Balakrishnan N., Joshi P. C. *A note on order statistics from Weibull distribution* // Scandinavian Actuarial Journal.– 1981.– Vol. **1981**.– No. 2.– Pp. 121–122. doi ↑^{60, 68}
- [18] Whitt W. *Comparing counting processes and queues* // Advances in Applied Probability.– 1981.– Vol. **13**.– No. 1.– Pp. 207–220. doi ↑⁶¹
- [19] Blanchet J., Pei Y., Sigman K. *Exact sampling for some multi-dimensional queueing models with renewal input* // Advances in Applied Probability.– 2019.– Vol. **51**.– No. 4.– Pp. 1179–1208. doi ↑⁶²
- [20] Bladt M., Nielsen B. *Matrix-Exponential Distributions in Applied Probability*, PTSM.– Vol. **81**.– Boston, MA: Springer US.– 2017.– ISBN 978-1-4939-7047-6.– 749 pp. doi ↑⁶³
- [21] Rumyantsev A., Chakravarthy S. *Matrix-exponential distributions in applied probability* // Proceedings of the Third International Conference BOINC-based High Performance Computing: Fundamental Research and Development, BOINC:FAST 2017 (August 28–September 01, 2017, Petrozavodsk, Russia), CEUR Workshop Proc.– vol. **1973**.– 2017.– Pp. 27–34. URL ↑⁶³

- [22] Buchholz P., Kriege J., Felko I. *Input Modeling with Phase-Type Distributions and Markov Models. Theory and Applications*, SpringerBriefs in Mathematics.– Cham: Springer International Publishing.– 2014.– ISBN 978-3-319-06673-8.– xii+127 pp. [doi](#) ↑⁶⁴
- [23] Harchol-Balter M. *Performance Modeling and Design of Computer Systems: Queueing Theory in Action*.– Cambridge: Cambridge University Press.– 2013.– ISBN 9781107027503.– 574 pp. ↑⁶⁵
- [24] Feitelson D. *Workload Modeling for Computer Systems Performance Evaluation*.– Cambridge: Cambridge University Press.– 2015.– ISBN 9781139939690. [doi](#) ↑⁶⁵
- [25] Malik H. J. *Exact moments of order statistics from the Pareto distribution* // Scandinavian Actuarial Journal.– 1966.– Vol. **1966**.– No. 3–4.– Pp. 144–157. [doi](#) ↑⁶⁶
- [26] Huang J. S. *A note on order statistics from Pareto distribution* // Scandinavian Actuarial Journal.– 1975.– Vol. **1975**.– No. 3.– Pp. 187–190. [doi](#) ↑⁶⁶
- [27] Daley D. J., Goldie C. M. *The moment index of minima (II)* // Statistics & Probability Letters.– 2006.– Vol. **76**.– No. 8.– Pp. 831–837. [doi](#) ↑⁶⁷
- [28] Daley D. J. *The moment index of minima* // Journal of Applied Probability.– 2001.– Vol. **38**, Issue A: Probability, Statistics and Seismology.– Pp. 33–36. [doi](#) ↑⁶⁷
- [29] Scheller-Wolf A., Vesilo R. *Sink or swim together: necessary and sufficient conditions for finite moments of workload components in FIFO multiserver queues* // Queueing Systems.– 2011.– Vol. **67**.– No. 1.– Pp. 47–61. [doi](#) ↑⁶⁷
- [30] Hardy G. H., Littlewood J. E., Pólya G. *Inequalities*, 2nd ed.– Cambridge: Cambridge University Press.– 1991.– 340 pp. ↑⁶⁸
- [31] Lieblein J. *On moments of order statistics from the Weibull distribution* // Ann. Math. Statist..– 1955.– Vol. **26**.– No. 2.– Pp. 330–333. [doi](#) ↑⁶⁸
- [32] Sultan K. S., Moshref M. E. *Moments of order statistics from Weibull distribution in the presence of multiple outliers* // Communications in Statistics – Theory and Methods.– 2014.– Vol. **43**.– No. 10–12.– Pp. 2214–2226. [doi](#) ↑⁶⁸
- [33] Rinne H. *The Weibull Distribution. A Handbook*.– New York: Chapman and Hall/CRC.– 2009.– ISBN 9780367577469.– 808 pp. ↑⁶⁸

Поступила в редакцию	06.12.2022;
одобрена после рецензирования	27.01.2023;
принята к публикации	05.02.2023;
опубликована онлайн	15.02.2023.

Рекомендовал к публикации


д.т.н. А. М. Цирлин

Информация об авторах:



Александр Сергеевич Румянцев

с.н.с., к.ф.-м.н., научные интересы: теория массового обслуживания, многосерверные системы, стохастическое моделирование и анализ, высокопроизводительные и распределенные вычисления, матрично-аналитический метод, регенеративный подход, трехуровневое моделирование

 0000-0003-2364-5939

e-mail: ar0@krc.karelia.ru



Диана Сергеевна Долгалева

стажер-иссл., научные интересы: теория массового обслуживания, матрично-аналитический метод, биоинформатика, генные сети, распределенные вычисления


 0000-0002-0606-6583

e-mail: abcdi_do@mail.ru



Александр Станиславович Головин


вед.инж.-иссл., аспирант, научные интересы: теория массового обслуживания, матрично-аналитический метод, энергоэффективность, имитационное моделирование, распределенные вычисления, многосерверные системы обслуживания

 0000-0003-1325-3739

e-mail: golovin@krc.karelia.ru

*Все авторы сделали эквивалентный вклад в подготовку публикации.
Авторы заявляют об отсутствии конфликта интересов.*

UDC 519.248+519.872.8

 10.25209/2079-3316-2023-14-1-55-94


Steady-state performance analysis of multiserver queueing models with redundancy

Alexander Sergeevich **Rumyantsev**¹, Diana Sergeevna **Dolgaleva**²,
Alexander Stanislavovich **Golovin**³

Institute of Applied Mathematical Research KarRC of RAS, Petrozavodsk, Russia

¹*✉* ar0@krc.karelia.ru

Abstract. An approach to study of the stochastic models of distributed computing systems by means of the multiserver queueing models with redundancy is suggested. Perspectives and limitations of the approach, as well as possible future research directions are presented.

Key words and phrases: distributed computing systems, multiserver queueing systems, stochastic modeling, systems with redundancy

2020 *Mathematics Subject Classification:* 60K30; 90B22, 65C30

Acknowledgments: The authors are grateful to the referees for valuable comments. The study was supported by the Russian Science Foundation grant 21-71-10135

For citation: Alexander S. Rumyantsev, Diana S. Dolgaleva, Alexander S. Golovin. *Steady-state performance analysis of multiserver queueing models with redundancy*. Program Systems: Theory and Applications, 2023, **14**:1(56), pp. 55–94. https://psta.pspiras.ru/read/psta2023_1_55-94.pdf

1. Introduction

Distributed computing systems, such as Desktop Grid (DG), Enterprise Desktop Grid (EDG), are widely used for utilizing the idle resources of the generic computing resources (personal computers/servers) in order to solve computationally intensive tasks [1]. The potentially high performance of such networks coincides with low reliability, hardly predictable task completion time, as well as a significant limitation on the class of tasks to be solved (the solution of such tasks needs to be divided into many loosely coupled subtasks that do not require interprocessor exchange).

To increase the reliability of the network and the reliability of the results obtained, *redundancy* at the task level is used in the DG by means of *replication* (i.e. copying a task with the same input) and *quorum* (the number of replicas required for successful validation and result assimilation).

Redundancy of information is a way to ensure reliable data transmission between nodes of computing systems using data transmission networks (including wireless) [2]. A similar problem arises in advanced storage systems based on solid-state drive technology, which, unlike traditional systems, have a significant internal reserve of parallelism, built-in redundancy, but significantly less time to failure of individual storage modules [3]. In this regard, controlled redundancy, duplication of information and high throughput are the target functions for optimizing the parameters of such media [4].

To analyze the stationary performance characteristics of telecommunication and computing systems, the methods of queuing theory are used due to the random nature of the processes under study. As a result of such a study, an optimal configuration parameters of such systems with redundancy may arise. The purpose of this work is to analyze the perspectives for the study of a class of multi-server models with redundancy, called split-merge systems, used as models of computing systems and data transmission systems [5, 6]. Further, the paper briefly summarizes the main results in the field of analysis of the split-merge systems, new results for stochastic ordering of key characteristics are obtained and directions for further research of this class of systems are proposed.

2. Generalized Split-Merge-type Systems

In this section we study the multiserver queueing systems with task-level redundancy organized by means of replication and quorum mechanisms. Such a systems are the models of the DG (including EDG) and many other systems such as reliable storage systems, e.g. RAID [7]. To study such systems, multiserver queueing systems with task split and merge may be used. Specifically, upon arrival, each task is split into multiple copies which are then processed according to the selected service discipline. Using the terminology adopted in DG systems, we call such tasks *workunits*. According to the service discipline, such systems fall into one of the following categories:

- classical *Fork-Join* (FJ) where the workunits enter independent queues of (identical) independent servers, whereas the service of a specific task is complete only upon completion of all the corresponding workunits (in such a case usually a specific unbounded buffer accumulating the completed workunits is used) [8–11];
- generalized *Fork-Join* systems, denoted as (r, q) -FJ, in which r replicas of a task (workunits) are created and service completion is considered to be the time at which q earliest results of the workunits of a specific task are obtained, whereas the unnecessary $r - q$ workunits are immediately removed from the servers/queues (in particular, the classical FJ corresponds to the generalized (r, r) -FJ) [12];
- generalized *Fork-Early-Cancel* systems, denoted as (r, q) -FEC, where each task is copied into r replicas, and upon the service initialization of not less than q workunits, the unnecessary $r - q$ workunits are cancelled from the servers/queues [12, 13];
- generalized *Split-Merge* systems, denoted as (c, r, q) -SM, in which a single queue and a pool of c groups of r identical servers each are used, and the workunits of each task are dispatched to a specific group of servers. The group of servers dedicated to a task is seized by the workunits simultaneously (for random independent service times), whereas the service completion epoch is the time of obtaining q completed workunits (the unnecessary workunits are cancelled, and the group of servers is simultaneously released).

We note that stochastic analysis of the classical FJ systems is rather complicated, since the arrival processes of the servers in such a system are correlated, and the queue sizes are unbounded, requiring a potentially unbounded storage needed to keep the system description Markovian. As such, the classical FJ systems are generally studied by means of numerical analysis and approximation [7]. However, in some important special cases such an analysis may be performed by using the results for classical queueing systems. In particular, the following important properties were obtained in the papers [12–14]:

- $(r, 1)$ -FJ system with (generic) workunit service time S is equivalent to a classical single-server queue with generic service time $S_{1:r}$;
- $(r, 1)$ -FEC system is equivalent to a classical r -server system;
- sojourn (response) time in the (r, q) -FJ system is upper bounded by the sojourn time in $(1, r, q)$ -SM system;
- sojourn time in the (c, r, q) -SM system with generic service time of a workunit S coincides with the sojourn time in a c -server system with generic service time $S_{q:r}$,

where $S_{q:r}$ is the q -th order statistics from r independent identically distributed (iid) random variables (r.v.), stochastic copies of S . In particular this means that $(1, r, 1)$ -SM system is equivalent to $(r, 1)$ -FJ system, whereas $(c, 1, 1)$ -SM system is equivalent to $(c, 1)$ -FEC system. In this regard, we focus on studying (c, r, q) -SM systems in what follows.

We consider the cr -server queueing system in which the tasks arrive with input rate λ . An arriving task is waiting for (simultaneous) release of the group of r servers and is split into r workunits, each having iid service times, stochastic copies of the generic positive random service time S . All the workunits of a given task simultaneously seize the group of r servers, and upon receiving the quorum of q completed workunits, the remaining (busy) $r - q$ servers immediately are released. This means that the group of r servers is seized only by the workunits of a given task, and there is no split of such a group between different tasks. As such, the

system under consideration is equivalent to the classical $G/G/c$ system in which the typical service time is the order statistics $S_{q:r}$. The stability condition of such a systems is as follows, [15]

$$(1) \quad \lambda E S_{q:r} < c.$$

The main goal of the analysis is, in general, to obtain the dependence of the following key performance and efficiency characteristics on the number of servers, cr , replication r and quorum q (these parameters are used hereafter in the notation):

sojourn time $V_{r,q}$ – the time interval from the arrival of a (generic) task into the system until service completion (obtaining q completed workunits)

$$(2) \quad V_{r,q} = W_{r,q} + S_{q:r},$$

where $W_{r,q}$ is the (stationary) generic delay in the system;

computation cost $C_{r,q}$ – the characteristic proportional to the computation time of a task,

$$(3) \quad C_{r,q} = r S_{q:r}.$$

Note that in DG the cost may be thought of as the energy consumption of the system, whereas in the cloud computing such a cost directly coincides with the expenses for resources rent, in the solid state drive-based storage system such a cost may be considered as the wearout of the storage units.

2.1. Stochastic monotonicity of the performance

The properties of these characteristics are in a complex dependence on the replication parameters r and the quorum q and, in particular, this is due to the peculiarities of the order statistics distribution. Let r.v. S have a distribution function (cdf) F . It is known [16] that the order statistics cdf $S_{q:r}$ has the form

$$(4) \quad F_{q:r}(x) = P\{S_{q:r} \leq x\} = \sum_{i=q}^r \binom{r}{i} F^i(x) \bar{F}^{r-i}(x),$$

where $\bar{F}(x) = 1 - F(x)$ is the tail of distribution. The equality (4) means, in particular, that the following (obvious) stochastic inequality holds

$$S_{1:r} \leq_{st} S \leq_{st} S_{r:r}.$$

Moreover, for a fixed r and $q > 1$ the following equality holds [17]

$$(5) \quad P(S_{q:r} \leq x) = P(S_{q-1:r} \leq x) - \binom{r}{q-1} F^{q-1}(x) \bar{F}^{r-q+1}(x).$$

This means, in particular, that

$$(6) \quad S_{q:r} \geq_{st} S_{q-1:r}.$$

Moreover, for fixed $r, q > 1$ the following equality holds [17]

$$(7) \quad P(S_{q:r} \leq x) = P(S_{q-1:r-1} \leq x) - \binom{r-1}{q-1} F^{q-1}(x) \bar{F}^{r-q+1}(x).$$

In particular this means that

$$(8) \quad S_{q:r} \geq_{st} S_{q-1:r-1}.$$

Finally for $q \geq 1$ holds [17]

$$(9) \quad P(S_{q:r} \leq x) = P(S_{q:r-1} \leq x) + \binom{r-1}{q-1} F^q(x) \bar{F}^{r-q}(x).$$

This means that $S_{q:r} \leq_{st} S_{q:r-1}$ which taking account of (6) and (8) leads to

$$(10) \quad S_{q:r} \geq_{st} S_{q-1:r-1} \geq_{st} S_{q-1:r}.$$

It follows from (3) and (6) that $C_{r,q}$ is monotonically non-decreasing as q increases (for a fixed r), i.e.

$$(11) \quad C_{r,q} \geq_{st} C_{r,q-1}, \quad q > 1.$$

It follows from (8) that $C_{r,q}$ is monotonically non-decreasing as q and r increase simultaneously, i.e.

$$(12) \quad C_{r,q} \geq_{st} C_{r-1,q-1}, \quad q, r > 1.$$

At the same time, in the general case, for a fixed q , as r increases, the quantity $C_{r,q}$ is nonmonotonic, which follows from (10).

To study the monotonicity of the sojourn time in the system depending on the parameters of replication and quorum, we need the following result, which is a direct consequence of Theorem 4 in [18]:

THEOREM 1. *Let the systems $\Sigma^{(1)}$ and $\Sigma^{(2)}$ (whose escape sequences are denoted by the corresponding superscript) have the following stochastic order for typical times between arrivals T and service times S ,*

$$(13) \quad T^{(1)} \geq_{st} T^{(2)}, \quad S^{(1)} \leq_{st} S^{(2)}.$$

then the stationary delays W in the systems are ordered as follows,

$$(14) \quad W^{(1)} \leq_{st} W^{(2)}.$$

Then it follows from equality (2), Theorem 1 and inequality (10) that the assertion is true.

STATEMENT 1. *For $r, q > 1$, the following stochastic order for stationary sojourn times takes place,*

$$(15) \quad V_{r,q} \geq_{st} V_{r-1,q-1} \geq_{st} V_{r,q-1}.$$

Note that the stochastic order implies, in particular, the ordering of the mean values.

In general the expression for the average sojourn time $EV_{r,q}$ for a system of type $G/G/c$ is unknown. An explicit form of the average sojourn time can be obtained in the case of $c = 1$ if the arrival process is Poisson. To do this, you can use the Pollacek—Khinchin formula

$$(16) \quad EW_{r,q} = \frac{\lambda ES_{q:r}^2}{2(1 - \lambda ES_{q:r})},$$

and get from (3) the expression

$$(17) \quad EV_{r,q} = E(S_{q:r}) + \frac{\lambda E(S_{q:r}^2)}{2[1 - \lambda E(S_{q:r})]},$$

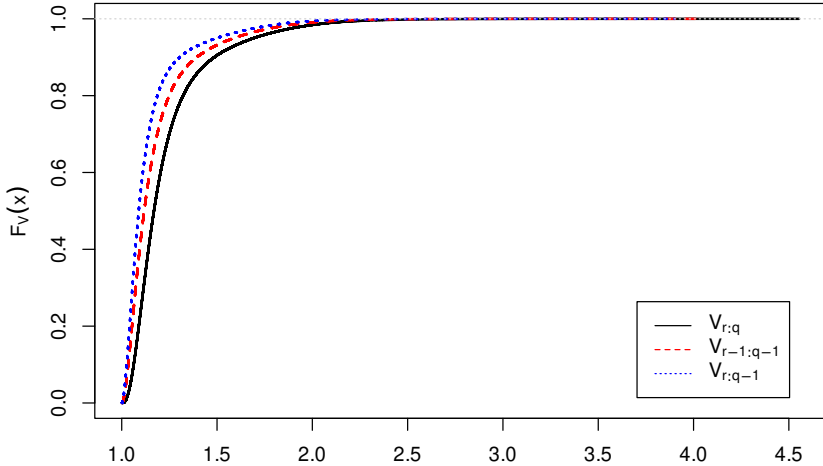


FIGURE 1. Cumulative distribution functions for the system $M/G/c$ with the input rate $\lambda = 1$ and the number of servers $c = 3$

which can be used to calculate the average sojourn time if the corresponding expressions for the moments of the order statistics are known.

To illustrate the stochastic order (15), let's plot the sojourn time distribution functions in the $M/G/c$ system for the case when the service time has Pareto distribution whose pdf has the form

$$(18) \quad f(x) = \alpha x_0^\alpha x^{-\alpha-1}, \quad x > x_0, \quad (f(x) = 0, x \leq x_0).$$

Let's set $\alpha = 5$, $x_0 = 1$ and redundancy parameters $r = 5$, $q = 3$ under input rate $\lambda = 1$ and the number of servers $c = 3$. For these systems, the load $\rho_{r,q} = \lambda ES_{q;r}$ is $\rho_{5,3} \approx 0.39$, $\rho_{4,2} \approx 0.38$ and $\rho_{5,2} = 0.37$, respectively. Note that the stationarity condition of the system is satisfied. Figure 1 shows the ordering of distribution functions corresponding to the stochastic order (15). Note that the graphics were built using the exact sampling method [19].

2.2. Phase-type service time distribution

The computation cost (3) may be obtained explicitly for the considered system in case the sojourn time has the so-called phase type distribution $PH(\boldsymbol{\beta}, \mathbf{B})$, i.e. the service time constitutes to the absorption time of a finite state space Markov chain with initial distribution $\boldsymbol{\beta}$ and (subgenerator) transition rate matrix for m transient states \mathbf{B} (for more details on this class of distributions see e.g. [20]). In such a case it can be shown (see [20, 21]) that the distribution of order statistics $S_{q:r}$ is given as $PH(\boldsymbol{\beta}_{q:r}, \mathbf{B}_{q:r})$, where

$$(19) \quad \mathbf{B}_{q:r} = \begin{bmatrix} \mathbf{B}^{\oplus r} & \mathbf{b}_0^{\oplus r} & \mathbf{O} & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \mathbf{B}^{\oplus r-1} & \mathbf{b}_0^{\oplus r-1} & \mathbf{O} & \dots & \mathbf{O} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & \mathbf{O} & \dots & \mathbf{B}^{\oplus r-q+1} \end{bmatrix},$$

where the (non-square) matrices $\mathbf{b}_0^{\oplus k}$ consist of such transition rates between the states that are related to absorption of one of the components of the chain (i.e. absorption of one of the k iid r.v.'s with distribution $PH(\boldsymbol{\beta}, \mathbf{B})$ each),

$$(20) \quad \mathbf{b}_0^{\oplus k} = \mathbf{b}_0 \otimes \mathbf{I}_{m^{k-1}} + \mathbf{I}_m \otimes \mathbf{b}_0 \otimes \mathbf{I}_{m^{k-2}} + \dots + \mathbf{I}_{m^{k-1}} \otimes \mathbf{b}_0,$$

the matrix \mathbf{I}_j is the identity matrix of size $j \geq 1$, \otimes is the Kronecker product and \oplus is the Kronecker sum of matrices (repeated sums and products of this type are given as upper indices), whereas $\mathbf{b}_0 = -\mathbf{B}\mathbf{1}$ is the absorption rate vector corresponding to the distribution $PH(\boldsymbol{\beta}, \mathbf{B})$. The initial distribution of such a chain may be given as

$$(21) \quad \boldsymbol{\beta}_{q:r} = (\boldsymbol{\beta}^{\otimes r}, \mathbf{0}),$$

This means that the phase-type distribution class is closed with respect to the order statistics, and hence the stability condition (1) and the cost $EC_{r,q}$, defined in (3), may be obtained in explicit form. Indeed, it is known that for the r.v. X having $PH(\hat{\boldsymbol{\beta}}, \hat{\mathbf{B}})$ distribution, the first and second

order moments are given as [22]

$$(22) \quad EX = -\hat{\beta}\hat{B}^{-1}\mathbf{1}, \quad EX^2 = 2\hat{\beta}\hat{B}^{-2}\mathbf{1}.$$

Hence, it follows from (22) that

$$(23) \quad EC_{r,q} = rES_{q:r} = -r\beta_{q:r}\mathbf{B}_{q:r}^{-1}\mathbf{1},$$

where $\beta_{q:r}$ and $\mathbf{B}_{q:r}$ are defined in (21) and (19), respectively.

Consider the following important special case, let the workunit service time be exponentially distributed with service rate μ . Then $\beta = 1$, $\mathbf{B} = -\mu$, $\mathbf{b}_0 = \mu$ and (23) may be transformed into

$$(24) \quad EC_{r,q} = \frac{1}{\mu} \sum_{i=0}^{q-1} \frac{r}{r-i}.$$

In particular, if $q = 1$,

$$(25) \quad P\{S_{1:r} \leq x\} = 1 - e^{-r\mu},$$

and hence (24) converts into

$$(26) \quad EC_{r,1} = \mu^{-1},$$

i.e. *the computation cost is independent on r when $q = 1$* . In general,

$$(27) \quad EC_{r+1,q} - EC_{r,q} = -\frac{1}{\mu} \left(\sum_{i=1}^{q-1} \frac{1}{(r-i+1)(r-i)} \right) \leq 0.$$

Hence the cost monotonically decreases with r for given fixed $q \geq 1$.

If the arrival process is Poisson (or the so-called Markovian arrival process), then the *sojourn time* in the system may be obtained explicitly, however, in practice it is possible only numerically. However, in some important special cases this may be done in explicit form.

Consider the case when $c = 1$, i.e. the replication is performed on all the servers of the system, and let the arrival process be Poisson. Then the $(1, r, q)$ -SM system is equivalent to the $M/PH/1$ system which, due to (22), has the following form of the sojourn time (17):

$$(28) \quad EV_{r,q} = \frac{\lambda\beta_{q:n}\mathbf{B}_{q:n}^{-2}\mathbf{1}}{1 + \lambda\beta_{q:n}\mathbf{B}_{q:n}^{-1}\mathbf{1}} - \beta_{q:r}\mathbf{B}_{q:r}^{-1}\mathbf{1}.$$

If the workunit service time is exponentially distributed with rate μ and if $q = 1$, then the analysis is simplified by using (25). In particular, for the $(c, r, 1)$ -SM, the sojourn time may be computed by using the Erlang formula [23],

$$(29) \quad EV_{r,1} = \frac{1}{r\mu} + \frac{C(c, \lambda/(r\mu))}{cr\mu - \lambda},$$

where the so-called Erlang C-formula $C(n, \rho)$ describes the probability of enqueueing in the M/M/n system having load $\rho < n$,

$$(30) \quad C(n, \rho) = \left[1 + (1 - \rho) \frac{n!}{(n\rho)^n} \sum_{k=0}^{n-1} \frac{(n\rho)^k}{k!} \right]^{-1}.$$

In particular, the sojourn time in $(c, 1, 1)$ -SM system coincides with the classical sojourn time in the $M/M/c$ system:

$$(31) \quad EV_{1,1} = \frac{1}{\mu} + \frac{C(c, \lambda/\mu)}{c\mu - \lambda}.$$

It follows from (30) that $C(1, \rho) = \rho$, and hence the sojourn time (29) in the $(1, r, 1)$ -SM system may be obtained as

$$EV_{r,1} = \frac{1}{r\mu} + \frac{\lambda}{r\mu(r\mu - \lambda)}.$$

2.3. Heavy-tailed service time distribution

It follows from (9) that the dependence of the cost $C_{r,q}$ on r, q is not monotone in general, and needs to be studied for particular classes of distributions. In this section we consider the dependence of the average cost $EC_{r,q}$ on the aforementioned parameters for the Pareto and Weibull service time distributions which (the latter - for specific parameter set) belong to the class of the heavy-tailed distributions. Note that the heavy-tailed distributions are widely used for studying the computing and communications systems [24].

Let the generic service time of the workunit S have the Pareto distribution (18) with parameter $\alpha > 1$. The properties of order statistics

from Pareto distribution are obtained in [25, 26], specifically,

$$(32) \quad \mathbb{E}(S_{q;r}^k) = x_0^k \frac{r!}{(r-q)!} \frac{\Gamma(r-q+1-k/\alpha)}{\Gamma(r+1-k/\alpha)}, \quad 0 < k < \alpha(r-q+1),$$

where Γ is the gamma function. It follows from (3) and (32) that the explicit expression for $EC_{r,q}$ may be obtained, and the following recurrence relation for $EC_{r,q}$ may be obtained as well [6]:

$$(33) \quad EC_{r+1,q} = \frac{(r+1)^2(r-q+1-1/\alpha)}{r(r-q+1)(r+1-1/\alpha)} EC_{r,q}.$$

It follows from (33), after some algebra, that $EC_{r,q}$ *increases* with r , if the following inequality holds good,

$$(34) \quad \alpha r^2 + (2\alpha - \alpha q - q - 1)r - \alpha q + \alpha - 1 > 0,$$

otherwise $EC_{r,q}$ *non-increases* with r . For $\alpha > 1$ and fixed $q \geq 1$, the quadratic polynomial (w.r.t. r) on the l.h.s. of (34) has two real roots, such that the smaller is negative and larger is positive and has the following form [6]:

$$(35) \quad r_0(q) = \frac{1 + q + \alpha q - 2\alpha + \sqrt{(\alpha q + q - 1)^2 + 4q}}{2\alpha}.$$

As such, for given fixed q the inequality (34) holds good for (integer) $r > r_0(q)$, and the minimal cost is attained at one of the integers nearest to $r_0(q)$.

It follows from (35), in particular, that $r_0(1) < 1$ for $\alpha > 1.5$, i.e. the cost $EC_{r,1}$ increases with r for all $r \geq 1$. Indeed,

$$(36) \quad EC_{r,1} = x_0 \frac{r^2 \alpha}{r\alpha - 1} = x_0 \left(r + \frac{1}{\alpha} + \frac{1}{\alpha(\alpha r - 1)} \right), \quad r > 1/\alpha.$$

Moreover, it follows from (36) that the cost is asymptotically linear

$$\lim_{r \rightarrow \infty} r^{-1} EC_{r,1} = x_0.$$

It can be seen from (36) that $EC_{1,1}$ may be arbitrary large if α is close to 1. Indeed, it can be derived from (32) or from (36) that

$$EC_{1,1} = ES = x_0 \frac{\alpha}{\alpha - 1}.$$

At the same time, it follows from (36) that

$$EC_{2,1} < x_0 \left(2 + \frac{2}{\alpha} \right),$$

since $\alpha r - 1 > 1$ for $r \geq 2$. This means that, in particular, $EC_{2,1} < EC_{1,1}$ for $\alpha < \sqrt{2}$. Thus, the replication may be used to significantly reduce the computation cost if $r < r_0(q)$.

With respect to the *sojourn time*, the general result (15) on the monotonicity w.r.t. replication/quorum parameters holds good, and, by taking into account (17) and (32), explicit expressions can be derived for the system of type $M/G/1$. However, an even more important problem to consider (in the presence of heavy tailed distributions) is to guarantee the finiteness of the moments (say, expectation or variance) of the sojourn time. This is especially important for obtaining reliable (interval) performance estimates in steady state and is important for the system stability.

Denote by $\kappa(X)$ the *moment index* of the r.v. X , i.e.. [27, 28]

$$(37) \quad \kappa(X) = \sup\{k > 0 : E(X^k) < \infty\} \leq \infty.$$

Then it follows from (18) that the r.v. S with Pareto distribution has the moment index equal to α ,

$$(38) \quad \kappa(S) = \alpha,$$

at the same time it follows from (32) that

$$(39) \quad \kappa(S_{q;r}) = \alpha(r - q + 1) \geq \alpha.$$

It follows from (39) that $\kappa(S_{q;r})$ is *linear* w.r.t. $r - q$, i.e. the given moment conditions (say, variance) may be configured by appropriate choice of the redundancy parameters q, r even in the case when the moment index of the workunits is relatively small (e.g. has infinite variance).

It is known that the moment index in the multiserver system $G/G/c$ may be obtained from the moment index of the service time [29]. In the system under consideration, for given fixed λ, α, c, r, q , the moment index of the waiting time $W_{r,q}$ may be obtained as follows [6]

$$(40) \quad \kappa(W_{r,q}) = (\alpha(r - q + 1) - 1)(c - [\rho]) \geq \alpha - 1,$$

where $\rho = \lambda ES_{q:r}$, and $ES_{q:r}$ is given in (32). Thus from (3), using the Minkowski inequality [30], we obtain

$$(E(W_{r,q} + S_{q:r})^k)^{1/k} \leq (EW_{r,q}^k)^{1/k} + (ES_{q:r}^k)^{1/k},$$

and hence

$$(41) \quad \kappa(V_{r,q}) \geq \min(\kappa(W_{r,q}), \kappa(S_{q:r})),$$

where the moment indices $\kappa(W_{r,q})$ and $\kappa(S_{q:r})$ are given in (40) and (39), respectively.

Note that equality in (40) holds good, in particular, for $c = 1$ and $r = q$, which agrees with the Pollacek–Khinchin formula (16), which states that finiteness of the *second moment* of the service time is necessary to ensure the finiteness of the average waiting time. Finally we note that, taking into account (41) and (40), the linear growth of the moment index $S_{q:r}$ in $r - q$ can be used to select the values of r, q in such way to guarantee the required value of the moment index (for instance, finiteness of the variance) of the task sojourn time even if the moment index $\kappa(S) = \alpha$ of the workunit is small enough (for example, when α is close to one).

Let's repeat the analysis for a system in which the typical workunit service time S has the Weibull distribution

$$(42) \quad F(x) = 1 - e^{-x^\xi}, \quad \xi > 0.$$

Then the moments of order $k > 0$ for $S_{q:r}$ can be obtained in the following form [31, 32]

$$(43) \quad E(S_{q:r}^k) = q \binom{r}{q} \Gamma(1 + k/\xi) \sum_{j=0}^{q-1} (-1)^j \binom{q-1}{j} (r - q + 1 + j)^{-1 - k/\xi}.$$

Thus, these moments are finite for any $k > 0$, i.e., $\kappa(S_{q:r}) = \infty$, while the expression (17) allows to find an explicit form for $EV_{r,q}$ for a system of type $M/G/1$. In addition, the following recurrence holds [17, 33]

$$(44) \quad E(S_{q:r}) = E(S_{q-1:r-1}) + \binom{r-1}{q-1} J(r - q + 1, q - 1),$$

where $ES_{1:1} = J(1, 0) = \Gamma(1 + k/\xi)$, and $J(p, m)$ is defined recursively:

$$J(p, m) = J(p, m - 1) + J(p + 1, m - 1), \quad J(p, 0) = p^{-1/\xi} \Gamma(1 + 1/\xi).$$

Taking account of (43), we have

$$(45) \quad EC_{r,q} = r \frac{r!}{(r-q)!} \frac{\Gamma(\xi^{-1})}{\xi} \sum_{j=0}^{q-1} (-1)^j \frac{(r-q+1+j)^{-1-\xi^{-1}}}{j!(q-1-j)!}.$$

It can be seen from (45) that

$$(46) \quad EC_{r,1} = \Gamma(1 + 1/\xi)r^{1-1/\xi}, \quad r \geq 1.$$

This means, in particular, that $EC_{r,1}$ is non-decreasing (non-increasing) in r for $\xi \geq 1$ ($\xi \leq 1$, respectively). Moreover, for $\xi = 1$ the Weibull distribution coincides with the exponential distribution, and at the same time (46) is consistent with (26).

It is difficult to use (45) directly to establish the monotonicity of $EC_{r,q}$. However, considering the difference $\delta(r) = EC_{2,r+1} - EC_{2,r}$ and using recursion (44), we obtain

$$\delta(r) = \Gamma(1 + 1/\xi) \left((2r^2 + r + 1)r^{-1/\xi} - r(r+1)^{1-1/\xi} - r^2(r-1)^{-1/\xi} \right).$$

In particular, for $\xi < 1$ the quantity $EC_{2,r}$ does not increase in r . Indeed, let us study the function

$$\psi(r) = (2r^2 + r + 1)r^{-1/\xi} - r(r+1)^{1-1/\xi} - r^2(r-1)^{-1/\xi}.$$

Denote $a = \xi^{-1} > 1$. Let's transform the function to the following form

$$(47) \quad \psi(r) = r^2(2r^{-a} - (r+1)^{-a} - (r-1)^{-a}) + \phi(r),$$

where

$$\phi(r) = r(r+1)(r^{-a-1} - (r+1)^{-a-1}) = r^{1-a} + r^{-a} - (r+1)^{1-a} + (r+1)^{-a}.$$

The convexity of the function x^{-a} implies that $2r^{-a} - (r+1)^{-a} - (r-1)^{-a} \leq 0$. Moreover, one can verify that $\partial\phi(r)/\partial r \leq 0$ for all $a > 1$, i.e. the function $\phi(r)$ is decreasing in $r \geq 2$. Moreover, one can check that $\psi(2) \leq 0$ for any $\xi \in (0, 1)$. Thus, it follows from (47) that the difference $\delta(r)$ is nonpositive for all $r \geq 2$, i.e., $EC_{2,r}$ is monotonically non-increasing in r for any $\xi \in (0, 1)$.

TABLE 1. A summary of the obtained results.

Distribution	Characteristic	
	$C_{r,q}$	$V_{r,q}$
In general	Non-monotonic	Non-decreasing, inequality (15)
Exponential	Non-increasing, inequality (27); is constant for $q = 1$, formula (26); explicit form (24)	Explicit form (29) in system $M/M/c$ for $q = 1$
Phase type	Explicit form (23)	Explicit form (28) in $M/PH/1$
Pareto	Global minimum, expression (35)	Explicit form in system $M/G/1$ with respect (17) and (32)
Weibull with parameter ξ	Not-increasing if $q = 1, 2$ and $\xi \leq 1$; non-decreasing for $q = 1, \xi \geq 1$, expression (46)	Explicit form in a system like $M/G/1$ taking account of (17) and (43)

3. Resume and perspectives of the research

Table 1 presents the results of studying the impact of replication parameter r and quorum q on the cost $C_{r,q}$ and the sojourn time $V_{r,q}$ in multi-server models with $(c, r, q) - SM$ redundancy for different distributions of workunit service time S . The type of monotonicity is indicated with an increase in r, q , in some cases the result is available for the average value or for the corresponding random variable (stochastic order). The results are obtained for cases where the service time has a phase-type distribution, Pareto or Weibull distribution (belonging to the class of heavy-tailed distributions).

Taking account of the obtained results, the following directions of their further development can be distinguished:















- the monotonicity of the average cost (27) was obtained for the special case of exponentially distributed workunit service times, while the general case of an arbitrary phase type of distributions is of interest;










- it would be interesting to consider also a heterogeneous case, for example, when the servers to which replication is carried out have different performance (in this case, the formula (19) and further results require appropriate adjustment);
- cost monotonicity for a system with Weibull distribution is shown in section 2.3 for $r = 1, 2$, the case of arbitrary r requires a separate study;
- finally, the model assumes that the number of server blocks c is unchanged, while r is a variable parameter, it is advisable to study the case when the number of blocks is defined as s/r (integer), where s is (unchanged) number of servers in source system.

The latter direction opens up opportunities for using stochastic comparisons for systems with a different number of servers, but the parametric monotonicity of systems is also interesting (for example, in the parameter α for the Pareto distribution or ξ for the Weibull distribution).

References

- [1] I. Chernov, N. Nikitina, E. Ivashko. “Task scheduling in desktop grids: open problems”, *Open Engineering*, **7**:1 (2017), pp. 343–357. [doi](#) ^{↑76}
- [2] N. Verma, D. Singh. “Data redundancy implications in wireless sensor networks”, *Procedia Computer Science*, **132** (2018), pp. 1210–1217. [doi](#) ^{↑76}
- [3] A. Rumyantsev, E. Ivashko, I. Chernov, D. Kositsyn, A. Shabaev, V. Ponomarev. “Latency/wearout in a flash-based storage system with replication on write”, 2019 24th Conference of Open Innovations Association (FRUCT) (08–12 April 2019, Moscow, Russia), pp. 360–366. [doi](#) ^{↑76}
- [4] I. Chernov, E. Ivashko, D. Kositsyn, V. Ponomarev, A. Rumyantsev, A. Shabaev. “Flash-based storage deduplication techniques: A survey”, *International Journal Of Embedded And Real-Time Communication Systems (IJERTCS)*, **10**:3 (2019), pp. 32–48. [doi](#) ^{↑76}
- [5] O. Osipov, E. Rogachko. “Fork-join queueing systems with heterogeneous servers threshold control policy”, SMARTY2020 (August 16–20, 2020, Petrozavodsk, Russia), CEUR Workshop Proc., vol. **2792**, 2020, pp. 76–88. [URL](#) ^{↑76}
- [6] A. Rumyantsev, S. Chakravarthy, E. Morozov, S. Remnev. “Cost and effect of replication and quorum in desktop grid computing”, *Information Technologies And Mathematical Modelling. Queueing Theory And Applications*, Springer CCIS Series, vol. **912**, Springer, Cham, 2018, ISBN 978-3-319-97594-8, pp. 143–156. [doi](#) ^{↑76, 86, 87}

- [7] A. Thomasian. “Analysis of fork/join and related queueing systems”, *ACM Computing Surveys*, **47**:2 (2014), id. 17, 71 pp.  [↑77, 78](#)
- [8] F. Baccelli, A. Makowski, A. Shwartz. “The fork-join queue and related systems with synchronization constraints: Stochastic ordering and computable bounds”, *Advances in Applied Probability*, **21**:3 (1989), pp. 629–660.  [↑77](#)
- [9] K. Gardner, S. Zbarsky, S. Doroudi, M. Harchol-Balter, E. Hyttiä, A. Scheller-Wolf. “Queueing with redundant requests: Exact analysis”, *Queueing Systems*, **83**:3–4 (2016), pp. 227–259.  [↑77](#)
- [10] C. Kim, A. K. Agrawala. “Analysis of the fork-join queue”, *IEEE Transactions on Computers*, **38**:2 (1989), pp. 250–255.  [↑77](#)
- [11] R. Nelson, A. N. Tantawi. “Approximate analysis of fork/join synchronization in parallel queues”, *IEEE Transactions on Computers*, **37**:6 (1988), pp. 739–743.  [↑77](#)
- [12] G. Joshi, E. Soljanin, G. Wornell. “Queues with redundancy: Latency-cost analysis”, *ACM SIGMETRICS Performance Evaluation Review*, **43**:2 (2015), pp. 54–56.  [↑77, 78](#)
- [13] G. Joshi. *Efficient redundancy techniques to reduce delay in cloud systems*, PhD Thesis, Massachusetts Institute of Technology, 2016. [↑77, 78](#)
- [14] G. Joshi, E. Soljanin, G. Wornell. “Efficient redundancy techniques for latency reduction in cloud systems”, *ACM Transactions on Modeling and Performance Evaluation of Computing Systems*, **2**:2 (2017), id. 12, 30 pp.  [↑78](#)
- [15] S. Asmussen. *Applied Probability and Queues*, SMAP, vol. **51**, Springer, 2003, ISBN 978-0-387-00211-8, 452 pp.  [↑79](#)
- [16] N. Balakrishnan. “Permanents, order statistics, outliers and robustness”, *Revista Matemática Complutense*, **20**:1 (2007), pp. 7–107.  [↑79](#)
- [17] N. Balakrishnan, P. C. Joshi. “A note on order statistics from Weibull distribution”, *Scandinavian Actuarial Journal*, **1981**:2 (1981), pp. 121–122.  [↑80, 88](#)
- [18] W. Whitt. “Comparing counting processes and queues”, *Advances in Applied Probability*, **13**:1 (1981), pp. 207–220.  [↑81](#)
- [19] J. Blanchet, Y. Pei, K. Sigman. “Exact sampling for some multi-dimensional queueing models with renewal input”, *Advances in Applied Probability*, **51**:4 (2019), pp. 1179–1208.  [↑82](#)
- [20] M. Bladt, B. Nielsen. *Matrix-Exponential Distributions in Applied Probability*, PTSM, vol. **81**, Springer US, Boston, MA, 2017, ISBN 978-1-4939-7047-6, 749 pp.  [↑83](#)
- [21] A. Rumyantsev, S. Chakravarthy. “Matrix-exponential distributions in applied probability”, *Proceedings of the Third International Conference BOINC-based High Performance Computing: Fundamental Research and Development*, BOINC:FAST 2017 (August 28–September 01, 2017, Petrozavodsk, Russia), CEUR Workshop Proc., vol. **1973**, 2017, pp. 27–34.  [↑83](#)

- [22] P. Buchholz, J. Krieger, I. Felko. *Input Modeling with Phase-Type Distributions and Markov Models. Theory and Applications*, SpringerBriefs in Mathematics, Springer International Publishing, Cham, 2014, ISBN 978-3-319-06673-8, xii+127 pp.  [↑84](#)
- [23] M. Harchol-Balder. *Performance Modeling and Design of Computer Systems: Queueing Theory in Action*, Cambridge University Press, Cambridge, 2013, ISBN 9781107027503, 574 pp. [↑85](#)
- [24] D. Fietelson. *Workload Modeling for Computer Systems Performance Evaluation*, Cambridge University Press, Cambridge, 2015, ISBN 9781139939690.  [↑85](#)
- [25] H. J. Malik. “Exact moments of order statistics from the Pareto distribution”, *Scandinavian Actuarial Journal*, **1966**:3–4 (1966), pp. 144–157.  [↑86](#)
- [26] J. S. Huang. “A note on order statistics from Pareto distribution”, *Scandinavian Actuarial Journal*, **1975**:3 (1975), pp. 187–190.  [↑86](#)
- [27] D. J. Daley, C. M. Goldie. “The moment index of minima (II)”, *Statistics & Probability Letters*, **76**:8 (2006), pp. 831–837.  [↑87](#)
- [28] D. J. Daley. “The moment index of minima”, *Journal of Applied Probability*, **38**, Issue A: Probability, Statistics and Seismology (2001), pp. 33–36.  [↑87](#)
- [29] A. Scheller-Wolf, R. Vesilo. “Sink or swim together: necessary and sufficient conditions for finite moments of workload components in FIFO multiserver queues”, *Queueing Systems*, **67**:1 (2011), pp. 47–61.  [↑87](#)
- [30] G. H. Hardy, J. E. Littlewood, G. Pólya. *Inequalities*, 2nd ed., Cambridge University Press, Cambridge, 1991, 340 pp. [↑88](#)
- [31] J. Lieblein. “On moments of order statistics from the Weibull distribution”, *Ann. Math. Statist.*, **26**:2 (1955), pp. 330–333.  [↑88](#)
- [32] K. S. Sultan, M. E. Moshref. “Moments of order statistics from Weibull distribution in the presence of multiple outliers”, *Communications in Statistics – Theory and Methods*, **43**:10–12 (2014), pp. 2214–2226.  [↑88](#)
- [33] H. Rinne. *The Weibull Distribution. A Handbook*, Chapman and Hall/CRC, New York, 2009, ISBN 9780367577469, 808 pp. [↑88](#)

Received	06.12.2022;
approved after reviewing	27.01.2023;
accepted for publication	05.02.2023;
published online	15.02.2023.

Recommended by


prof. A. M. Tsirlin

Information about the authors:



Alexander Sergeevich Rumyantsev

senior researcher, PhD, key interests: queueing theory, multiserver systems, stochastic modeling and analysis, high-performance and distributed computing, matrix-analytic method, regenerative approach, three-level modeling


 0000-0003-2364-5939

e-mail: ar0@krc.karelia.ru



Diana Sergeevna Dolgaleva

intern-researcher, key interests: queueing theory, matrix-analytic method, bioinformatics, gene networks, distributed computing


 0000-0002-0606-6583

e-mail: abcdi_do@mail.ru



Alexander Stanislavovich Golovin

leading engineer-researcher, PhD student, key interests: queueing theory, matrix-analytic method, energy efficiency, simulation, distributed computing, multiserver queues

 0000-0003-1325-3739

e-mail: golovin@krc.karelia.ru

The authors declare no conflicts of interests.