

С.Ю. Скрупский, ассистент,  
Р.К. Кудерметов, канд. техн. наук, доц.,  
Запорожский национальный технический университет  
88sts88@mail.ru, krk@zntu.edu.ua

## Имитационные модели распределенных систем компрессии видеоинформации

*Представлены имитационные модели распределенных систем компрессии видеоинформации, позволяющие планировать выбор вычислительных ресурсов для достижения желаемой производительности в процессе сжатия видеоинформации. Экспериментально проверена адекватность предложенных моделей.*

**Ключевые слова:** видеоинформация, компрессия, распределенная система, имитационная модель, система массового обслуживания.

### 1. Введение

Для сжатия видеоинформации применяют современные методы устранения интракадровой и интеркадровой избыточности видеоинформации [1-4]. Высокая вычислительная сложность таких методов не позволяет сжимать видеоинформацию стандарта HDTV [5] в реальном времени на однопроцессорных компьютерах. Поэтому для сокращения времени сжатия видеоинформации (не за счет снижения вычислительной сложности методов сжатия и, как следствие, снижения качества результата сжатия) используют специальные вычислительные средства и в том числе параллельные и распределенные компьютерные системы [6-8].

### 2. Постановка задачи

Применение распределенных систем для компрессии видеоинформации предполагает предварительное планирование необходимых ресурсов, например, ресурсов Грид, запрашиваемых в виртуальной организации. При этом преследуется цель уменьшить время сжатия видеоинформации путем использования нескольких вычислительных ресурсов. Рациональный выбор вычислительных ресурсов можно осуществить при помощи моделирования процесса сжатия видеоинформации в распределенной системе.

В данной работе ставится задача разработать имитационные модели распределенных систем компрессии видеоинформации, позволяющие планировать выбор вычислительных ресурсов для достижения желаемой производительности в процессе сжатия видеоинформации при сохранении приемлемого качества результата сжатия.

### 3. Модель распределенной системы сжатия видеоинформации как системы массового обслуживания

Рассмотрим процесс сжатия видеоинформации вычислительными узлами распределенной системы как процесс обслуживания требований в системе массового обслуживания (СМО), состоящей из входящего потока требований, очереди требований, обслуживающих приборов и выходящего потока требований. Пусть в распределенную систему поступает видеоинформация для сжатия. Специальный вычислительный узел, например сервер системы, принимает ее и выполняет разделение видеопоследовательности на части предложенным в [9] методом. Суть метода заключается в вычислении коэффициентов корреляции между всеми смежными кадрами видеопоследовательности. Затем для коэффициента корреляции текущей пары кадров вычисляется значение функции чувствительности к смене сцен анализатора видеопоследовательности. Если значение функции чувствительности превышает значение модуля коэффициента корреляции, то видеопоследовательность разделяется на две части. Первая часть содержит кадры исходной видеопоследовательности до определенного с помощью функции чувствительности места слабой межкадровой корреляции, а вторая видеопоследовательность – с кадрами после места слабой межкадровой корреляции. Если же значение функции чувствительности меньше значения модуля коэффициента корреляции, тогда вычисляется значение функции чувствительности для следующей пары смежных кадров, которое сравнивается с соответствующим значением модуля коэффициента корреляции.

Полученные в результате разделения части видеопоследовательности становятся в очередь на обслуживание (сжатие) вычислительными узлами и образуют входящий в СМО поток требований. Если какой-либо вычислительный

узел свободен, то он изымает требование из очереди и обслуживает его. Обслуженное требование (сжатая часть видеоинформации) поступает в очередь выходящего потока требований, который направлен в хранилище сжатой видеоинформации. Например, в Грид имеется планировщик, который в каждый момент времени располагает информацией о свободных и занятых вычислительных ресурсах, очередях требований. Если все ресурсы заняты, требование ожидает в очереди, пока не освободится необходимый ресурс. Планировщик извлекает задачу из очереди и отправляет ее на вычислительный узел – обслуживающий прибор.

Наиболее подходящим типом распределенной системы сжатия видеоинформации является вычислительный кластер, поскольку вычислительные узлы на базе однотипных (равных по производительности) процессоров в таком кластере объединены высокоскоростной средой пе-

редачи данных с низкой латентностью. Высокоскоростная среда передачи данных важна для уменьшения накладных расходов вычислительного процесса, а однотипность процессоров обеспечивает универсальность и простоту использования ресурсов.

Для конечного результата сжатия видеоинформации важно, чтобы все ее части были сжаты, т.е. чтобы ни одно из требований не получило отказ на обслуживание. Следовательно, отнесем рассматриваемую СМО к классу систем с неограниченным временем ожидания требований в очередях.

Таким образом, распределенную систему сжатия видеоинформации, объединяющую вычислительные узлы, можно представить в виде многофазной СМО с неограниченным временем ожидания требований в очередях и с однотипными обслуживающими приборами (рис. 1).

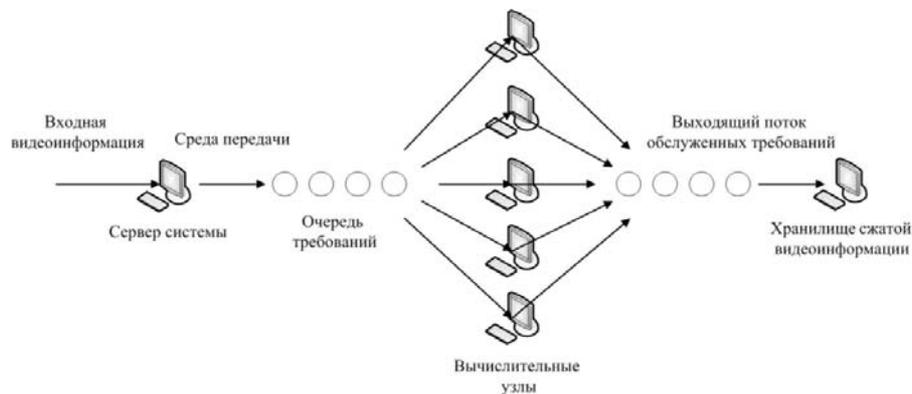


Рисунок 1 – Структурная схема распределенной системы сжатия видеоинформации

В качестве характеристик эффективности распределенной системы сжатия видеоинформации, учитываемых в модели, были выбраны следующие:

- общее время вычислительного процесса;
- коэффициент загрузки вычислительных узлов;
- среднее время нахождения требований в очередях.

#### 4. Разработка и исследование имитационной модели распределенной системы компрессии видеоинформации

Для описания имитационной модели в терминах СМО введем следующие обозначения:

- $\lambda$  – интенсивность входящего потока требований;
- $\mu$  – интенсивность потока обслуживания;
- $n$  – количество приборов в многоканальной системе;
- $k$  – количество требований, поступающих в систему;

$t_{обс}$  – среднее время обслуживания требования в системе;

$t_{пер1}$  – среднее время, затрачиваемое на доставку требования от источника требований к любому обслуживающему прибору;

$t_{сжат}$  – среднее время сжатия требования;

$t_{пер2}$  – среднее время, затрачиваемое на доставку требования от любого обслуживающего прибора в хранилище.

Среднее время обслуживания требования в системе  $t_{обс}$  является суммой среднего времени доставки требования от сервера к вычислительным узлам, среднего времени собственно сжатия требования и среднего времени доставки требования от вычислительных узлов в хранилище:

$$t_{обс} = t_{пер1} + t_{сжат} + t_{пер2} \cdot \quad (1)$$

Соответственно, интенсивность потока обслуживания в системе можно определить по формуле:

$$\mu = \frac{1}{t_{\text{обс}}} = \frac{1}{t_{\text{неп1}} + t_{\text{сжат}} + t_{\text{неп2}}}. \quad (2)$$

Поскольку несжатая видеоинформация занимает больший объем памяти чем сжатая, то для доставки сжатой видеоинформации в хранилище требуется меньше времени, чем для доставки несжатой видеоинформации к вычислительным узлам при постоянной пропускной способности среды передачи, т.е. выполняется следующее условие:

$$t_{\text{неп2}} < t_{\text{неп1}}. \quad (3)$$

Интенсивность  $\lambda$  зависит от стандарта представления видеоинформации и от производительности анализатора видеопоследовательности, выполняющего разделение ее на части. Чем больше разрешение кадра видеоинформации (в соответствии со стандартом ее представления), тем меньше интенсивность входящего потока требований, формируемых анализатором видеопоследовательностей при его постоянной производительности. Вместе с тем, чем больше производительность анализатора видеопоследовательностей, тем больше интенсивность  $\lambda$ . Интенсивность  $\mu$  прямо пропорциональна пропускной способности среды передачи, производительности вычислительных узлов распределенной системы и обратно пропорциональна вычислительной сложности алгоритмов стандарта сжатия видеоинформации.

Предложенная имитационная модель распределенной системы сжатия видеоинформации включает в себя модели следующих систем:

- анализатор видеопоследовательности – генератор требований, который создает поток из  $k$  требований с интенсивностью  $\lambda$ ;
- среда передачи – одноканальный обслуживающий прибор со средним временем обслуживания  $t_{\text{неп1}}$ , который доставляет требования от генератора требований в многоканальную систему;
- вычислительные узлы –  $n$ -канальная система со средним временем обслуживания требований  $t_{\text{сжат}}$ ;
- среда передачи – одноканальный обслуживающий прибор со средним временем обслуживания  $t_{\text{неп2}}$ , доставляющий требования от многоканальной системы в уничтожитель требований;
- хранилище – уничтожитель требований.

Таким образом, предложенная имитационная модель учитывает характеристики входящего потока требований (интенсивность потока требований, количество требований, т.е. частоту разделенной предложенным в [9] методом видеопоследовательности), характеристики среды передачи (латентность и пропускную способ-

ность) и характеристики обслуживания (количество вычислительных узлов и среднее время обслуживания требования каждым узлом). Структура предложенной имитационной модели приведена на рис. 2.

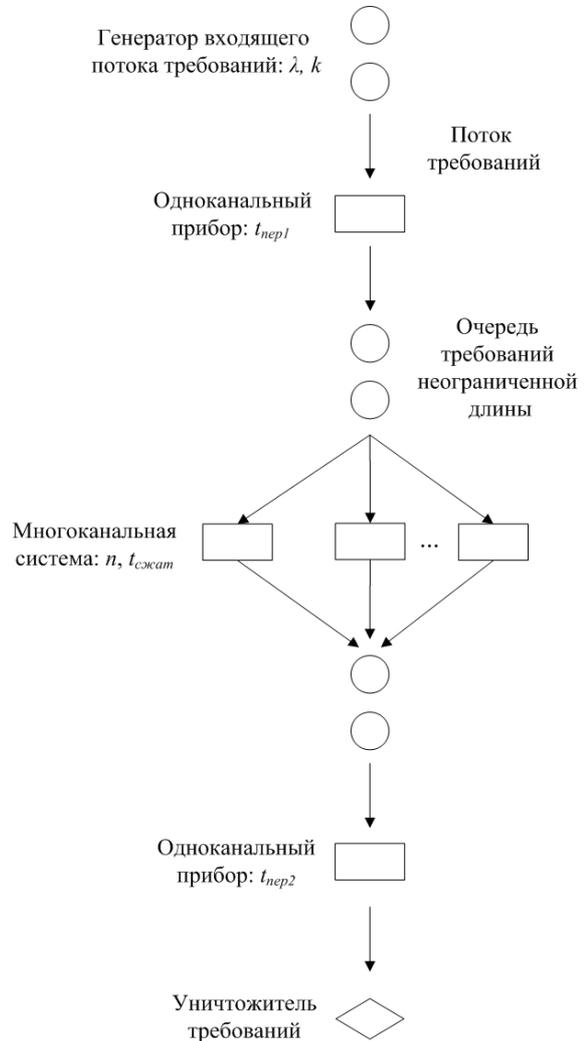


Рисунок 2 – Структура предложенной имитационной модели

#### 4.1 Модель распределенной системы сжатия видеоинформации для эксклюзивного случая использования ресурсов

Рассмотрим случай эксклюзивного использования ресурсов распределенной системы. При этом мы предполагаем, что ресурсы такой системы всегда доступны, обслуживание каждого требования начинается сразу после его поступления в систему, реальная производительность выделенных для сжатия видеоинформации ресурсов соответствует номинальной, т.е. все процессорное время выделенных вычислительных узлов задействовано для сжатия видеоинформации. Среда передачи данных выделенного сегмента распределенной системы используется

только для передачи видеоинформации. Такое предположение необходимо для анализа простейшего случая, например, когда пользователь монополюбно владеет ресурсами распределенной системы.

Эксклюзивное использование пропускной способности сети превращает условие (3) в следующее выражение:

$$t_{пер2} = t_{пер1} \cdot K_{сж}, \quad (4)$$

где  $K_{сж}$  – коэффициент сжатия видеопоследовательности.

Поскольку производительность анализатора видеопоследовательности постоянна в ходе всего процесса и видеопоследовательности подаются в анализатор последовательно одна за другой, т.е. вероятность поступления в СМО определенного числа требований в течение отрезка времени  $\tau$  зависит от длины этого отрезка, а не от начала его отсчета на оси времени, то такой поток требований можно считать стационарным. В любой момент времени анализатор выдает в СМО только одно требование, т.е. поток ординарный. Условия разделения видеопоследовательностей анализатором постоянны и видеопоследовательности подаются в анализатор последовательно одна за другой, т.е. вероятность поступления в СМО за отрезок времени  $\tau$  определенного числа требований не зависит от того, сколько требований уже поступило в систему. Следовательно, такой поток можно считать потоком без последствия.

Учитывая эти свойства входящего в СМО потока требований, есть основания выдвинуть

гипотезу о том, что поток является пуассоновским.

Данная гипотеза была подтверждена статистически по критерию хи-квадрат в ходе эксперимента, описанного ниже. На вход модели анализатора видеоинформации, реализованной в пакете Matlab, в произвольном порядке последовательно одна за другой подавались тестовые видеопоследовательности, полученные конкатенацией файлов, содержащих общепринятые для тестирования методов сжатия видеопоследовательности [10, 11] и охватывающие все доступные типы цифровой видеоинформации. Конкатенация обеспечивала дополнительные смены сцен [9]. Весь период времени наблюдения в 100 минут был разделен на 100 одноминутных интервалов, в каждом из которых определялось количество  $a_i$  поступивших от анализатора видеоинформации частей видеопоследовательности. Интервалы с одинаковым количеством поступивших требований были сгруппированы. Затем было вычислено математическое ожидание  $\bar{a}$  количества требований в течение принятого интервала времени – 1 минута. При помощи этого математического ожидания вычислялись вероятности  $P_i$  количества требований в интервале времени в соответствии с пуассоновским распределением. Также вычислялись математические ожидания  $n_m$  числа интервалов с заданным количеством требований. Результаты эксперимента приведены в табл. 1.

Таблица 1. Характеристики входящего потока требований

Число требований $a_i$ в интервале времени $t=1$ мин	Число интервалов $n_i$ с одинаковым числом требований	Значение вероятности $P_i$ числа требований в интервале	Математическое ожидание $n_m$ числа интервалов с заданным числом требований
0	0	0.0003	0.0304
1	0	0.0025	0.2459
2	1	0.0100	0.9958
3	3	0.0269	2.6886
4	6	0.0544	5.4443
5	4	0.0882	8.8198
6	16	0.1191	11.9067
7	16	0.1378	13.7778
8	12	0.1395	13.9500
9	11	0.1256	12.5550
10	9	0.1017	10.1696
11	10	0.0749	7.4885
12	6	0.0505	5.0547
13	4	0.0315	3.1495
14	2	0.0182	1.8222
15	0	0.0098	0.9840
16	0	0.0050	0.4981
17	0	0.0024	0.2374
18	0	0.0011	0.1068
19	0	0.0005	0.0455

Вычислена величина хи-квадрат по формуле:

$$\chi^2 = \sum \frac{(n_i - n_m)^2}{n_m}. \quad (5)$$

Получено  $\chi^2 = 8.51$ . По таблице значений  $\chi_{кр}^2$  [12] для числа степеней свободы  $w = 20 - 2 = 18$  и уровня значимости  $\alpha = 0.05$  определена критическая точка правосторонней критической области  $\chi_{кр}^2(0.05, 18) = 9.3905$ . Поскольку  $\chi^2 < \chi_{кр}^2$  – нет оснований отвергнуть гипотезу о распределении входящего в СМО потока требований по закону Пуассона.

Поскольку производительность вычислительных узлов в эксклюзивном режиме постоянна и видеопоследовательности, поступающие в систему от анализатора, однотипны, то время сжатия  $t_{сжат}$  для каждого узла детерминировано для видеопоследовательностей каждого стандарта представления видеoinформации. Среднее значение  $t_{сжат}$  для вычислительных узлов всей распределенной системы так же детерминировано для видеопоследовательностей каждого стандарта представления видеoinформации, например, для видеопоследовательностей стандарта CIF величина  $t_{сжат}$  меньше, а для видеопоследовательностей стандарта HDTV – больше. Средние значения характеристик среды передачи данных  $t_{пер1}$ ,  $t_{пер2}$  – детерминированные. Такой вывод следует из того, что производительность распределенной системы определяется ее техническими характеристиками и неизменна при использовании в эксклюзивном режиме.

#### 4.2 Модель распределенной системы сжатия видеoinформации для неэксклюзивного случая использования ресурсов

Рассмотрим случай неэксклюзивного использования ресурсов распределенной системы. В этом случае требования сжатия видеoinформации могут обслуживаться не сразу, а по мере высвобождения запрошенных ресурсов. Анализатор формирует пуассоновский поток требований. Требования становятся в очередь, ожидая выделения запрошенных ресурсов распределенной системы. В такой имитационной модели (в отличие от эксклюзивного случая использования ресурсов) требования после их появления накапливаются в очереди и ожидают на протяжении времени  $t_{ож}$ , т.е. времени от момента появления первого требования до момента выделения запрошенных ресурсов распределенной системы.

Время  $t_{ож}$  прямо пропорционально количеству запрошенных вычислительных узлов и обратно пропорционально количеству свободных вычислительных узлов и их производительности.

Среда передачи данных распределенной системы используется для передачи не только видеoinформации, но и других данных. Следовательно,  $t_{пер1}$  и  $t_{пер2}$  могут изменяться во времени. В этом случае условие (3) не превращается в выражение (4).

Будем считать, что производительность выделенных вычислительных узлов, как и для эксклюзивного случая, постоянная, т.к. если вычислительный ресурс уже выделен, то все его процессорное время используется для обслуживания текущего требования. Усредненное значение  $t_{сжат}$  для всей распределенной системы детерминировано для каждого стандарта представления видеoinформации.

#### 5. Организация экспериментальной проверки адекватности предложенных моделей

Для проверки адекватности предложенных имитационных моделей были проведены натурные эксперименты по сжатию видеoinформации в распределенных системах в эксклюзивном и в неэксклюзивном режимах. Анализатор видеoinформации распределенной системы генерировал входящий поток требований – отправлял части видеопоследовательностей, полученные после разделения видеoinформации в соответствии с предложенным в [9] методом, по среде передачи на вычислительные узлы, которые выполняли сжатие поступивших видеопоследовательностей по стандарту MPEG-4/H.264 с помощью программы FFmpeg. Настройки кодера видеoinформации были следующими:

- количество В-кадров в группе: 3;
- количество восстановленных кадров для операции оценки движения: 5;
- четвертьпиксельная точность для операции оценки движения блоков полным перебором с блоками переменного размера и областью поиска 128 пикселей;
- битрейт переменный с сохранением исходного качества.

Стандарт MPEG-4/H.264 был выбран для сжатия потому, что его алгоритмы обладают значительно большей вычислительной сложностью по сравнению с алгоритмами других стандартов, обеспечивают наилучшее качество сжатия, наименьший коэффициент сжатия и востребованы чаще алгоритмов других стандартов сжатия. В качестве тестовых использованы QCIF- и CIF-видеопоследовательности, описанные в [9]. Набор тестовых видеопоследовательностей пода-

вался на вход анализатора в произвольном порядке 6 раз для каждого эксперимента. Так обеспечивалось увеличение объема выборки для исследования. Все требования поступали в очередь распределенной системы с равным приоритетом, поэтому изымались из очереди по принципу "FIFO". Сжатые видеопоследовательности от вычислительных узлов транспортировались по среде передачи в хранилище. Регистрировалось время передачи видеоинформации от сервера к вычислительным узлам  $t_1$ , время сжатия видеопоследовательности  $t_2$ , время передачи сжатой видеоинформации от вычислительных узлов в хранилище  $t_3$  и общее время выполнения процесса сжатия всех тестовых видеопоследовательностей в распределенной системе  $T$ .

Для повышения достоверности результатов эксперименты проводились на трех различных кластерах:

1. Кластер Запорожского национального технического университета (ЗНТУ). Было задействовано 8 узлов следующей конфигурации:

- процессор Intel E3200;
- оперативная память 1 ГБ DDR-2 на каждый вычислительный узел;
- коммуникационная среда Fast Ethernet 100 Мб/с.

На узлах кластера установлено промежуточное программное обеспечение (middleware): gLite middleware 3.1 [13] с планировщиком Torque на операционной системе Scientific Linux 4.8.

2. Кластер НТУУ "КПИ" г. Киев. Было задействовано 8 логических узлов следующей конфигурации:

- процессор Intel Xeon 5440;
- оперативная память 4 ГБ DDR-2 на каждый вычислительный узел;
- коммуникационная среда InfiniBand 10Гб/с.

На узлах кластера установлено middleware: Torque [14] на операционной системе CentOS 5.2.

3. Кластер Института проблем моделирования в энергетике (ИПМЭ) им. Г.Е. Пухова г. Киев. Было задействовано 8 логических узлов следующей конфигурации:

- процессор Intel Xeon 5405;
- оперативная память 4×2 ГБ DDR-2 на каждый вычислительный узел;
- коммуникационная среда InfiniBand 20Гб/с.

На узлах кластера установлено middleware: пакет ARC [15] с планировщиком Torque на операционной системе CentOS 5.2.

Под логическим вычислительным узлом подразумевается полнофункциональный независимый процессор, который расположен в одном физическом корпусе с другими логическими вычислительными узлами. Так в современных кластерах объединяют до 8 логических вычисли-

тельных узлов в одном физическом [16]. Влияние пакетов gLite и ARC на производительность распределенных систем практически одинаковое [17].

### 5.1 Проверка адекватности имитационной модели для эксклюзивного случая

Для проверки адекватности предложенной модели кластер ЗНТУ использовался в эксклюзивном режиме. Анализатор видеоинформации выдавал в кластер пуассоновский поток требований. Для оценки величин  $t_{неп1}$ ,  $t_{сжат}$ ,  $t_{неп2}$ , которые вводились в имитационную модель, был задействован только один вычислительный узел кластера. Эксперимент повторялся 10 раз. Усредненные данные сведены в табл. 2. Результаты экспериментов сравнивались с результатами имитационного моделирования в среде GPSS.

Для имитационного моделирования распределенной системы сжатия видеоинформации была использована среда GPSS на базе программы GPSS World [18], которая хорошо зарекомендовала себя в научном обществе и часто используется для моделирования СМО [19].

Особенности реализации предложенной модели в среде GPSS:

- единица модельного времени – 1 секунда;
- интенсивность  $\lambda$  входящего в СМО потока требований вводится в модель путем преобразования интенсивности в соответствующее ей среднее время между последовательными поступлениями требований;
- пуассоновский входящий поток требований задается в блоке *GENERATE* с параметрами  $\lambda$ , *FN\$EXPON*, где *FN\$EXPON* определена следующим образом:  
*EXPON FUNCTION RNI,C24*  
*0,0/.1,.104/.2,.222/.3,.355/.4,.509/.5,.69/.6,.915/.7,1.2/.75,1.38/.8,1.6/.84,1.83/.88,2.12/.9,2.3/.92,2.52/.94,2.81/.95,2.81/.96,3.2/.97,3.5/.98,3.9/.99,4.6/.995,5.3/.998,6.2/.999,7/.9998,8;*

–  $t_{неп1}$ ,  $t_{неп2}$  являются параметрами одноканальных обслуживающих устройств (среда передачи) и используются в соответствующих блоках *ADVANCE*, т.е. указывают модельное время, в течение которого среда передачи транспортирует требование к вычислительным узлам или от вычислительных узлов в хранилище соответственно;

– вычислительные узлы имитируются в модели с помощью оператора *STORAGE* с параметром  $n$ , соответствующим емкости многоканального прибора;

– одно требование при поступлении занимает, а после обслуживания – освобождает один канал многоканального прибора;

– если обслуживающий прибор занят, то

требование ожидает неограниченное время в очереди до освобождения прибора;

- $t_{сжат}$  является параметром многоканального прибора и используется в соответствующем ему блоке *ADVANCE*;
- хранилище – блок, уничтожающий по-

ступающее в него требование.

Предложенная имитационная модель использовалась для моделирования аналогичного эксперимента с характеристиками кластера, полученными в ходе натуральных экспериментов.

Таблица 2. Результаты эксперимента для кластера ЗНТУ на одном узле

Тестовая видеопоследовательность	$t_1, c$	$t_2, c$	$t_3, c$
Akiyo Carphone Claire part1	0.72	8.21	0.27
Akiyo Carphone Claire part2	0.83	25.22	0.36
Akiyo Carphone Claire part3	1.02	21.05	0.31
Coastguard Waterfall Container part1	0.93	90.72	0.34
Coastguard Waterfall Container part2	0.89	44.82	0.29
Coastguard Waterfall Container part3	0.99	40.9	0.31
Silent Mother and Daughter Bus part1	0.88	30.6	0.29
Silent Mother and Daughter Bus part2	0.93	46.41	0.35
Silent Mother and Daughter Bus part3	0.89	70.11	0.33
Mobile Bridge Paris part1	0.86	62.52	0.33
Mobile Bridge Paris part2	1.19	47.16	0.35
Mobile Bridge Paris part3	0.86	33.45	0.29
News Tempete Hall monitor part1	1.13	36.74	0.31
News Tempete Hall monitor part2	0.54	11.26	0.27
News Tempete Hall monitor part3	1.19	82.86	0.36
News Tempete Hall monitor part4	0.86	45.52	0.33

Из табл. 2 путем усреднения соответствующих параметров по всем тестовым видеопоследовательностям получены следующие данные:

- $t_{пер1} = 0.919$  с;
- $t_{сжат} = 43.597$  с;
- $t_{пер2} = 0.318$  с.

Натурный эксперимент повторялся на кластере для двух, четырех и восьми запрошенных узлов. В модели изменялось только количество вычислительных узлов и моделирование повторялось. Экспериментальные данные, полученные на кластере ЗНТУ, и результаты имитационного моделирования приведены на рис. 3, где по оси  $X$  на графике отложено количество задействованных вычислительных узлов распределенной системы, а по оси  $Y$  – суммарное время сжатия всех тестовых видеопоследовательно-

стей.

Аналогичные эксперименты были проведены на кластере НТУУ "КПИ" и на кластере ИПМЭ им. Г.Е. Пухова. Экспериментальные данные, полученные на кластерах, и результаты имитационного моделирования показаны на рис. 4 и 5. В табл. 3 приведены максимальные значения относительных отклонений ( $\delta$ ) результатов моделирования от результатов экспериментов.

Таким образом, максимальное относительное отклонение результатов моделирования от экспериментальных данных для эксклюзивного случая использования распределенных систем составило 6.45%, что подтверждает адекватность предложенной модели распределенной системы сжатия видеоинформации для эксклюзивного случая.

Таблица 3. Максимальные значения относительных отклонений результатов моделирования от экспериментальных данных для эксклюзивного случая

	Кластер ЗНТУ	Кластер НТУУ "КПИ"	Кластер ИПМЭ им. Г.Е. Пухова
$\delta, \%$	6.45	6.17	4.13

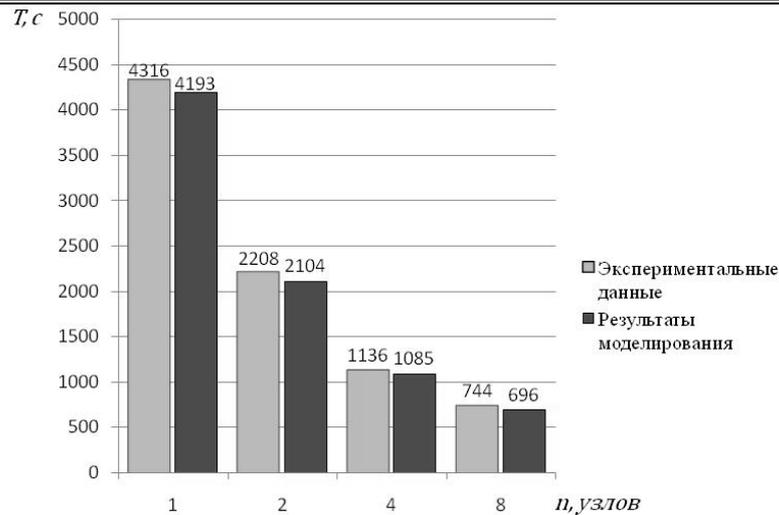


Рисунок 3 – Экспериментальные данные, полученные на кластере ЗНТУ, и результаты имитационного моделирования

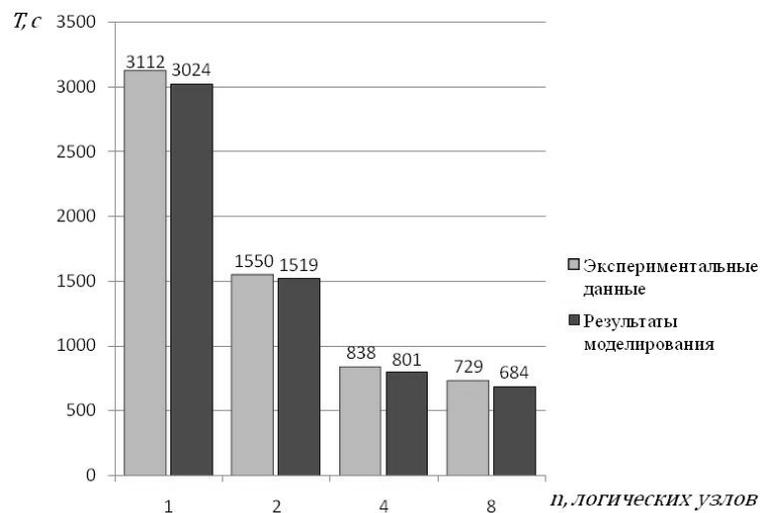


Рисунок 4 – Экспериментальные данные, полученные на кластере НТУУ "КПИ", и результаты имитационного моделирования

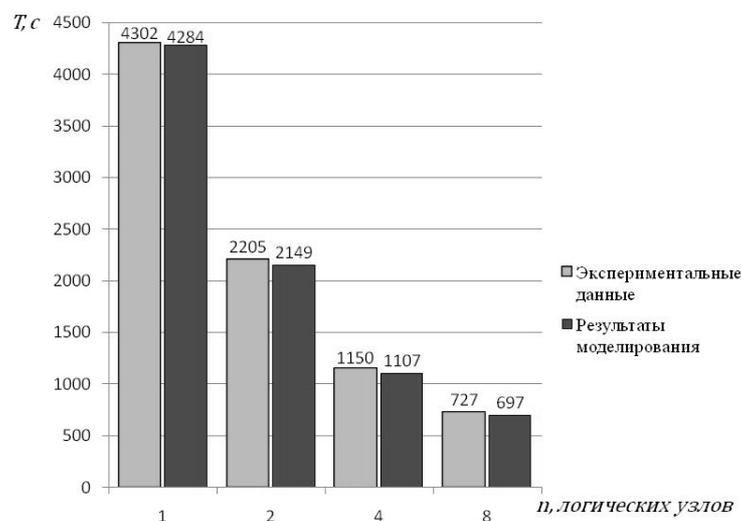


Рисунок 5 – Экспериментальные данные, полученные на кластере ИПМЭ им. Г.Е. Пухова, и результаты имитационного моделирования

## 5.2 Проверка адекватности модели для неэксклюзивного случая

Для неэксклюзивного случая использования распределенной системы вышеупомянутые кластеры были задействованы в произвольные моменты загрузки (количество свободных узлов в каждый момент времени случайно, пропускная способность среды передачи данных варьируется). Эксперименты выполнены аналогично, как описано в п. 5.1, но количество повторений экспериментов – 100 раз для каждого кластера. Экспериментально установлено, что величины  $t_{пер1}$  и  $t_{пер2}$  могут изменяться от 80% до 100% от соответствующих величин, полученных для эксклюзивного случая.

В отличие от эксклюзивного случая, для моделирования в GPSS неэксклюзивного случая использования распределенной системы между генератором входящего потока требований и средой передачи требований введен блок, имитирующий задержку перед началом обслуживания требований в распределенной системе:

```
GENERATE 0,,1
FUNAVAIL TTRANSFERI
ADVANCE TWAIT
FAVAIL TTRANSFERI
TERMINATE
```

Величина *TWAIT* соответствует  $t_{ож}$ ; *TTRANSFERI* – среда передачи данных.

Значения параметра  $t_{ож}$  для кластеров,

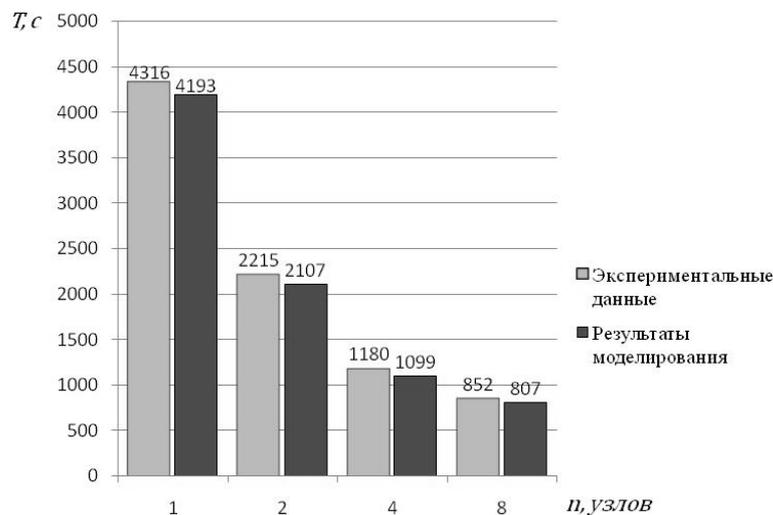


Рисунок 6 – Экспериментальные данные, полученные на кластере ЗНТУ в неэксклюзивном режиме, и результаты имитационного моделирования

задействованных в экспериментах, получены экспериментально и приведены в табл. 4.

Таблица 4. Значения  $t_{ож}$  для кластеров

Узлы, шт	Величина $t_{ож}$ (с) для кластера		
	ЗНТУ	НТУУ "КПИ"	ИПМЭ им. Г.Е. Пухова
1	2	5	2
2	12	8	6
4	21	15	10
8	31	30	25

Изменение пропускной способности среды передачи данных на GPSS имитируется в блоках *ADVANCE*, соответствующих  $t_{пер1}$  и  $t_{пер2}$ . Экспериментальные данные, полученные на кластерах в неэксклюзивном режиме, и результаты имитационного моделирования приведены на рис. 6, 7, 8. В табл. 5 приведены максимальные значения относительных отклонений результатов моделирования от экспериментальных данных ( $\delta$ ), полученных для распределенных систем, задействованных в экспериментах.

Таким образом, максимальное относительное отклонение результатов моделирования от экспериментальных данных для неэксклюзивного случая использования распределенных систем составило 7.65%, что подтверждает адекватность предложенной модели распределенной системы сжатия видеoinформации для неэксклюзивного случая использования ресурсов.

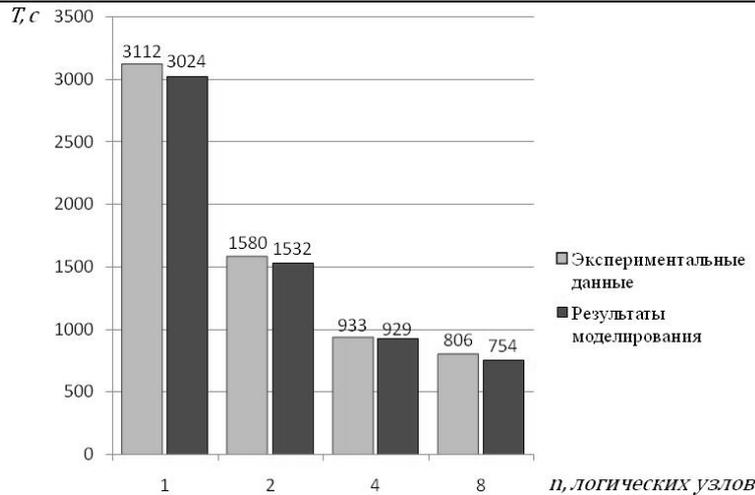


Рисунок 7 – Экспериментальные данные, полученные на кластере НТУУ "КПИ" в неэксклюзивном режиме, и результаты имитационного моделирования

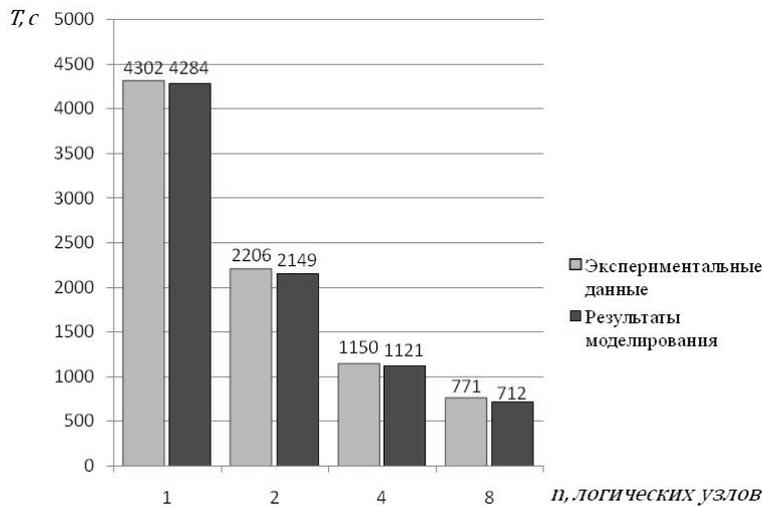


Рисунок 8 – Экспериментальные данные, полученные на кластере ИПМЭ им. Г.Е. Пухова в неэксклюзивном режиме, и результаты имитационного моделирования

Таблица 5. Максимальные значения относительных отклонений результатов моделирования от экспериментальных данных для неэксклюзивного случая

	Кластер ЗНТУ	Кластер НТУУ "КПИ"	Кластер ИПМЭ им. Г.Е. Пухова
$\delta, \%$	6.86	6.45	7.65

**6. Имитационное моделирование процесса сжатия видеоинформации в распределенных системах**

Предложенная имитационная модель была использована для моделирования различных конфигураций распределенной системы с целью достижения желаемой производительности в сжатии видеоинформации, представленной по стандарту HDTV с разрешением кадра 1920×1080 пикселей, в эксклюзивном режиме. Производительность вычислительных узлов, введенная в модель, предварительно оценена экспериментальным путем для процессоров Intel E3200, сжимающих видеоинформацию, представленную по стандарту HDTV. Параметры

среды передачи данных, которые были введены в модель, соответствуют параметрам Infiniband 20 Гб/с, полученным экспериментально.

В модели изменялось количество вычислительных узлов и моделирование повторялось. Результаты имитационного моделирования приведены в табл. 6, где  $k$  – количество частей видеопоследовательности, полученных от анализатора после разделения видеоинформации по предложенному в [9] методу;  $t_{np1}$  – среднее время нахождения требований в очереди в среде передачи от сервера к приборам;  $t_{np2}$  – среднее время нахождения требований в очереди на сжатие;  $t_{np3}$  – среднее время нахождения требований в очереди в среде передачи от приборов в

хранилище;  $k_3$  – коэффициент загрузки вычислительных узлов;  $T$  – общее время сжатия;  $S$  – ускорение вычислительного процесса по сравнению с одним процессором;  $E$  – эффективность распределенной системы. Результаты в табл. 6 получены для параметров модели  $k=128$ ,  $\lambda=8.1$  требований в минуту.

Как видно из табл. 6, эффективность распределенной системы для эксклюзивного случая тем выше, чем больше количество задач  $k$  по сравнению с количеством запрошенных вычислительных узлов  $n$ . Среднее время нахождения требований в очереди в среде передачи от при-

боров в хранилище равно нулю, следовательно, любая сжатая видеопоследовательность доставляется в хранилище быстрее, чем появляется другая сжатая видеопоследовательность.

В предложенную модель введена производительность вычислительных узлов, оцененная для процессоров Intel Xeon 5440, идентичная производительности процессоров кластера НТУУ "КПИ". Остальные параметры модели – такие же, как и в предыдущем эксперименте. Результаты имитационного моделирования приведены в табл. 7.

Таблица 6. Результаты имитационного моделирования

$n$	$t_{nep1}, c$	$t_{сжат}, c$	$t_{nep2}, c$	$k_3$	$t_{np1}, c$	$t_{np2}, c$	$t_{np3}, c$	$T, c$	$S$	$E$
1	4.55	665	0.21	1	3.63	42087	0	85132	1	1
2				1	3.55	21037		42576	1.99	0.99
4				0.99	3.79	10239		21322	3.99	0.99
8				0.99	3.67	5000		10711	7.95	0.99
16				0.98	3.84	2228		5439	15.65	0.98
32				0.95	3.51	881		2910	29.25	0.91
64				0.86	2.67	214		1805	47.16	0.74
128				0.56	2.69	0		1584	53.74	0.42

Таблица 7. Результаты имитационного моделирования

$n$	$t_{nep1}, c$	$t_{сжат}, c$	$t_{nep2}, c$	$k_3$	$t_{np1}, c$	$t_{np2}, c$	$t_{np3}, c$	$T, c$	$S$	$E$
1	4.55	442	0.21	1	3.57	27927	0	56588	1	1
2				1	3.47	13897		28304	1.99	0.99
4				0.99	3.74	6693		14186	3.99	0.99
8				0.99	3.78	3102		7143	7.92	0.99
16				0.98	3.54	1301		3654	15.49	0.97
32				0.93	2.97	432		2018	28.04	0.88
64				0.8	2.71	5.27		1362	41.55	0.65
128				0.4	2.71	0		1362	41.55	0.32

Применение более производительных процессоров типа Intel Xeon E5440 в кластере по сравнению с процессорами типа Intel E3200 позволило уменьшить общее время сжатия 128 видеопоследовательностей, представленных по стандарту HDTV, на 64 узлах с 1805 с до 1362 с. Однако, этой производительности не достаточно для того, чтобы сжимать видеoinформацию, представленную по стандарту HDTV, в реальном времени. Экспериментально установлено, что при постоянном количестве видеопоследовательностей, поступивших в распределенную систему и постоянной производительности анализатора видеoinформации, увеличение количества узлов системы с некоторого момента не повышает ускорение вычислительного процесса. Это объясняется тем, что вычислительные узлы, обслуживающие требования, освобождаются и обслуживают очередные требования из очереди.

Следовательно, запрошенные сверх необходимости ресурсы распределенной системы никогда не будут задействованы. Поэтому, планирование выбора ресурсов распределенной системы, которое можно осуществить при помощи предложенных моделей, позволит избежать такой ситуации.

### Выводы

Предложены имитационные модели распределенных систем сжатия видеoinформации, позволяющие планировать выбор ресурсов с целью уменьшения времени сжатия видеoinформации. Экспериментально подтверждена гипотеза о том, что части видеопоследовательности, получаемые путем разделения видеoinформации предложенным в [9] методом, образуют входящий в СМО пуассоновский поток требований. Адекватность предложенных моделей проверена экспериментальным путем на кластерах ЗНТУ,

НТУУ "КПИ", ИПМЭ им. Г.Е. Пухова для эксклюзивного и неэксклюзивного случаев их использования.

Результаты имитационного моделирования позволяют судить о том, что производительности современных кластеров (на примере кластера НТУУ "КПИ") не достаточно для того, чтобы сжимать видеoinформацию, представленную по стандарту HDTV, в реальном времени с сохранением приемлемого качества результата

сжатия. Для достижения такой производительности без значительной деградации качества результата сжатия требуется модификация методов оценки движения блоков в видеопоследовательностях или, при наличии такой возможности, распараллеливание процесса компрессии видеoinформации на каждом узле распределенной системы, которое можно осуществить путем использования специальных вычислительных средств.

### Список использованной литературы

1. Bhaskaran V. Image and video compression standards: algorithms and architectures / V. Bhaskaran, K. Konstantinides. – MA, USA : Kluwer Academic Publishers Norwell, 1997. – 454 p.
2. Сэломон Д. Сжатие данных, изображений и звука / Сэломон Д.; перевод с англ. В.В. Чепыжова. – М. : Техносфера, 2004. – 368 с.
3. Hardware Architecture Design of Video Compression for Multimedia Communication Systems / S-Y. Chien, Y-W. Huang, C-Y. Chen et al. // IEEE Communications Magazine. – 2005. – P. 123 – 131.
4. Обзор блочных методов оценки движения в цифровых видео сигналах / С.В. Гришин, Д.С. Ватолин, А.С. Лукин и др. // Программные системы и инструменты. Тематический сборник. – 2008. – №9. – С. 50 – 62.
5. Parameter values for the HDTV standards for production and international programme exchange : ITU-R Recommendation BT.709, 2008.
6. Rodriguez A. Performance Evaluation of Parallel MPEG-4 Video Coding Algorithms on Clusters of Workstations : PARELEC '04 / A. Rodriguez, A. Gonzalez, M. P. Malumbres // Proceedings of the international conference on Parallel Computing in Electrical Engineering. – Washington DC, 2004. – P. 354 – 357.
7. Parallel Scalability of H.264 / C.H. Meenderinck, A. Azevedo, M. Alvarez et al. / Proceedings of the first Workshop on Programmability Issues for Multi-Core Computers – Goteborg, Sweden, 2008.
8. ON-DEMAND HD VIDEO USING JINI BASED GRID : ICME 2008 / S. Kent, P. Broadbent, N. Warren et al. – 2008. – P. 1045–1048.
9. Скрупский С.Ю. Повышение эффективности сжатия видеoinформации в распределенных компьютерных системах / С.Ю. Скрупский // Электронное моделирование. – 2011. – №6 (33). – С. 57–72.
10. YUV Video Sequences. Arizona State University. – Режим доступа: \www/ URL: <http://trace.eas.asu.edu/yuv/> — 1.11.2010. — Загл. с экрана.
11. Xiph.org Test Media. – Режим доступа: \www/ URL: <http://media.xiph.org/video/derf/> — 1.11.2010. — Загл. с экрана.
12. Новиков О.А. Прикладные вопросы теории массового обслуживания / О.А. Новиков, С.И. Петухов. – М. : Советское радио, 1969. – 400 с.
13. Навчально-дослідницький Grid-кластер ЗНТУ / Кудерметов Р.К., Шкарупило В.В., Скрупський С.Ю. и др. // V міжнародна науково-технічна конференція "Сучасні проблеми і досягнення в галузі радіотехніки, телекомунікацій та інформаційних технологій РТІТ-2010", Запоріжжя, 22-24 вересня 2010. – Запоріжжя, 2010. – С. 41 – 42.
14. Система пакетной обработки заданий torque: руководство пользователя. – М. : Т-Платформы, 2008. – 31 с.
15. Advanced Resource Connector middleware for lightweight computational Grids / M. Ellert, M. Grønager, A. Konstantinov et al. // Future Generation Computer Systems. – 2007. – Vol. 23. – Iss. 2. – P. 219 – 240.
16. Platform Brief Intel® Xeon® Processor Technology. – Режим доступа: \www/ URL: <http://www.intel.com/assets/PDF/prodbrief/xeon-5400-animated.pdf> — 20.01.2011. — Загл. с экрана.
17. A comparison between ARC and gLite for medical image processing on Grids : in 2nd workshop on Medical imaging on GRID, HPC and GPU infrastructures: achievements and perspectives (MICCAI-Grid) / T. Glatard, X. Zhou, S. Camarasu-Pop et al. – London, UK, September 2009. – P. 29 – 39.
18. GPSS World Products. – Режим доступа: \www/ URL: <http://www.minutemansoftware.com/product.htm> — 1.02.2011. — Загл. с экрана.
19. Власов С.А. Язык моделирования GPSS World и системы автоматизации имитационных исследований: опыт применения и перспективы использования / С.А. Власов, В.В. Девятков, Т.В. Девятков // Четвертая всероссийская научно-практическая конференция по имитационному моделированию и его применению в науке и промышленности "Имитационное моделирование. Теория и практика". – СПб, 2009. – Т 1. – С. 11 – 18.

Надійшла до редколегії 20.01.2012

**С.Ю.СКРУПСЬКИЙ, Р.К. КУДЕРМЕТОВ**  
Запорізький національний технічний університет

**S.Y.SKRUPSKY, R.K. KUDERMETOV**  
Zaporizhzhya National Technical University

**ІМІТАЦІЙНІ МОДЕЛІ РОЗПОДІЛЕНИХ СИСТЕМ КОМПРЕСІЇ ВІДЕОІНФОРМАЦІЇ**

**THE SIMULATION MODELS OF VIDEO INFORMATION COMPRESSION DISTRIBUTED SYSTEMS**

Представлені імітаційні моделі розподілених систем компресії відеоінформації, що дозволяють планувати вибір обчислювальних ресурсів для досягнення бажаної продуктивності в процесі компресії відеоінформації. Експериментально перевірена адекватність запропонованих моделей.

The simulation models of video information compression distributed systems are presented. They allow planning the selection of the computing resources to achieve the desired performance in the process of video information compression. The adequacy of the proposed models is verified experimentally.

**Ключові слова:** *відеоінформація, компресія, розподілена система, імітаційна модель, система масового обслуговування.*

**Keywords:** *video information, compression, distributed system, simulation model, queueing system.*