

## АСИМПТОТИЧЕСКИЕ ОЦЕНКИ ЭФФЕКТИВНОСТИ СЕТЕЙ ХРАНЕНИЯ ДАННЫХ

Институт компьютерных технологий  
Национального авиационного университета

*Предложен метод расчета основных характеристик узлов сети на основании средних значений параметров входных потоков сообщений при проектировании систем, для которых потери системы обслуживания будут минимальны*

### Постановка задачи

В процессе развития информационных технологий, а также при современном развитии техники передача данных для хранения и архивирования снижается. Ситуация, когда та же база данных сохраняется в две или больше копии в разных подразделениях предприятия, а также копирование или архивирование данных происходит с использованием локальной сети, что приводит к торможению работы как сети, так и пользовательских приложений. Поэтому копирование происходит в после завершения рабочего дня. Это приводит к дополнительным издержкам организации (затраты на дополнительную оплату труда). В том случае если происходит внештатная ситуация и данные приходится восстанавливать это приводит к торможению всей работы подразделений предприятия, а на восстановление информации теряется время. Утраченные вычислительные ресурсы можно восстановить, а утраченные данные, при отсутствии грамотно спроектированной и внедренной системы резервирования, уже не подлежат восстановлению.

Ни один проект крупной сети со сложной топологией в настоящее время не обходится без моделирования будущей сети. Целью моделирования является определение оптимальной топологии, адекватный выбор сетевого оборудования, выбор массива для хранения данных, определение рабочих характеристик сети и возможных этапов будущего развития. Ведь сеть, слишком точно оптимизированная для решений задач текущего момента, может потребовать серьезных переделок в будущем. В процессе модели-

рования выясняются следующие параметры [1]:

Предельные пропускные способности различных фрагментов сети и зависимости потерь пакетов от загрузки отдельных станций и внешних каналов.

Время отклика основных серверов в самых разных режимах, в том числе таких, которые в реальной сети крайне нежелательны.

Влияние установки новых серверов на перераспределение информационных потоков (*Proxy, Firewall* и т.д.).

Решение оптимизации топологии при возникновении узких мест в сети (размещение серверов, службы централизованного разрешения имен *Domain Name System (DNS)*, внешних шлюзов, организация опорных каналов и пр.).

Выбор того или иного типа сетевого оборудования (например, *10BaseTX* или *100BaseFX*) или режима его работы (например, *cut-through, store-and-forward* для мостов и переключателей и т.д.).

Выбор внутреннего протокола маршрутизации и его параметров (например, метрики).

Определение предельно допустимого числа пользователей того или иного сервера.

Оценка необходимой полосы пропускания внешнего канала для обеспечения требуемого уровня *QoS*.

Однако, как бы ни были высоки скорости информационных потоков, это не исключает возможности образования так называемых "узких мест" в точках их концентрации, означающие нехватку ресурса для обслуживания заявок с заданным качеством.

Методи расчета требований к сетям новых поколений (пропускной способности каналов, емкости буферов и пр.) основанные на марковских моделях и формулах Эрланга, которые с успехом использовались при проектировании телефонных сетей, могут давать неоправданно оптимистические решения и приводить к недооценке нагрузки.

При исследовании любых систем управления необходимо решение одной из двух задач – анализа или синтеза. Анализ предусматривает выявление и изучение различных свойств данной сети при тех или иных условиях ее функционирования. Задача синтеза заключается в создании такой сети, которая обладает в определенных условиях заданными свойствами. Решение обеих задач (анализа в меньшей степени) предполагает наличие некоторой оценки качества работы сети, исходя из которой можно сказать, что одна система работает лучше, а другая – хуже и насколько. Коренная проблема количественной оценки предметов и процессов заключается в том, чтобы понятиям «лучше» и «хуже» поставить в соответствие понятия «больше» и «меньше». Для этой цели служат критерии качества, которые представляют собой количественные показатели, числовые значения которых являются мерой качества систем управления.

Решая задачу синтеза системы управления, мы должны начать с определения целей т.е. определить, какие функции в заданных условиях будет выполнять проектируемая система и ввести соответствующий набор количественных критериев качества ее работы.

При постановке и решении многих задач данной среде допускается предположение об экспоненциальном распределении исходных характеристик [2], что для современных систем не всегда является адекватной и удовлетворительной.

Рассмотрим многолинейную систему  $M/M/n/N$ , а также формулы каждого слагаемого из выражения по оптимизации величины расходов ( $C$ ) рассматриваемой системы. Эта система имеет  $n$  параллельных идентичных обслуживающих уст-

ройств и буфер для ожидания. На вход системы поступает простейший поток с параметром  $\lambda$ , а время обслуживания запроса имеет показательное распределение с параметром  $\mu$ .

Величина расходов данной системы определяется его техническими характеристиками, среди которых основное место занимают следующие параметры: интенсивность обслуживания, число каналов передачи и объема буфера.

Величину суммарного убытка ( $C$ ) многоканальной системы можно представить путем суммирования с некоторыми весами соответствующих средних значений характеристик:

$$C = b_1 \cdot t_{\text{простой}} + b_2 \cdot t_{\text{очередь}} + b_3 \cdot t_{\text{буфера}} + b_4 P_{\text{потерь}}$$

где  $t_{\text{простой}}$  – суммарное время простоя каналов узла;

$t_{\text{очередь}}$  – время задержки сообщений в очереди;

$t_{\text{буфера}}$  – простой буферных мест;

$P_{\text{потерь}}$  – число потерянных сообщений в единицу времени;

$b_1, b_2, b_3, b_4$  – весовые коэффициенты, соответственно, за простой канала связи, пребывания сообщения в очереди, за неиспользованное свободное место в буфере накопителя и штрафа за потерянное сообщение.

Состояние системы определяется случайным процессом  $V(t)$  – число запросов в рассматриваемой системе в момент времени  $t$ . Этот процесс является процессом гибели и размножения с параметрами [4]:

$$\lambda_i = \lambda, 0 \leq i \leq N$$

$$\mu_i = \begin{cases} i\mu, & \text{при } 1 \leq i \leq n. \\ n\mu, & \text{при } i > n \end{cases}$$

Важной характеристикой производительности многих реальных систем является распределение периода занятости системы. Период занятости является интервал времени с момента поступления запроса в пустую систему до того момен-

та, когда система впервые вновь окажется пустой [2].

Распределение состояний описанного процесса имеет вид:

$$M_k = \begin{cases} \frac{\alpha^k / k!}{\sum_{m=0}^n \frac{\alpha^m}{m!} + \sum_{m=n+1}^{n+N} \frac{\alpha^m}{n! n^{m-n}}}, & \text{при } k = 0, n \\ \frac{\alpha^k / (n! n^{k-n})}{\sum_{m=0}^n \frac{\alpha^m}{m!} + \sum_{m=n+1}^{n+N} \frac{\alpha^m}{n! n^{m-n}}}, & \text{при } k = n+1, n+N \end{cases}$$

Математическое ожидание величины очереди выражается формулой:

$$L_x = \frac{\frac{n^n}{n!} [(1-\rho^N) - N(1-\rho)\rho^N]}{\left( \sum_{m=0}^n \frac{\alpha^m}{m!} + \frac{1}{n!} \sum_{m=n+1}^{n+N} \frac{\alpha^m}{n^{m-n}} \right) (1-\rho)^2},$$

где  $\alpha = \lambda / \mu$ .

Параметр  $\rho$  – коэффициент загрузки системы, характеризующий соотношение интенсивности входящего потока и суммарной интенсивности обслуживания всеми приборами [4]:

$$\rho = \lambda / \mu n.$$

На значения параметров налагаются определенные ограничения, обусловленные как физическими характеристиками технологии построения узлов сети, так и условиями поддержки стационарного режима работы СМО. Ввиду того, что запросы источники и ответы узлов являются собой кадры *Ethernet*, интенсивности запросов и ответов ( $\lambda$ ) могут принимать любое значение в диапазоне  $\{0; \lambda_{max}\}$ , где  $\lambda_{max}$  – максимально возможная интенсивность передачи кадров в сети.

Параметры интенсивности обработки запросов и ответов ( $\mu$ ), в свою очередь, могут приобретать любое значение в диапазоне  $\{0; \mu_{max}\}$ , где  $\mu_{max}$  – максимально возможная интенсивность обработки кадров узлами сети.

Учитывая, что отмеченные ограничения, которые налагаются технологией

построения сети, а также условиями существования стационарного режима работы системы можно сформировать условия существования функции:

$$\alpha = (\lambda, \mu): \begin{cases} \lambda \leq \lambda_{max} \\ \lambda < \mu \\ \mu < \mu_{max} \end{cases}$$

Среднее время задержки пакета в системе ( $w$ ) вычисляется по формуле Литтла [3]:

$$L_x = \lambda \cdot w,$$

откуда получаем:

$$w = L_x / \lambda.$$

Для вычисления суммарного времени простоя  $t_{простой}$  воспользуемся формулой.

$M_k$  ( $k = 0, n$ ) означает, что заняты  $k$  каналов, а  $(n - k)$  простаивают. Отсюда

$$t_{простой} = \sum_{k=0}^{n-1} (n-k) \cdot M_k.$$

Количество неиспользованных мест в буфере накопителя вычисляется с использованием формулы:

$$t_{буфер} = N - M_k,$$

где  $N$  – количество мест в буфере.

Исходя из выше приведенных выражений, выражение по оптимизации величины расходов ( $C$ ) примет вид:

$$C = b_1 \sum_{k=0}^{n-1} (n-k) M_k + b_2 L_x + b_3 (N - L_x).$$

Полученное выражение целевой функции  $C$ , описывающей убытки узла системы трафика отображенные на рис. 1, зависит от трех параметров:  $V$  – скорости передачи сообщений,  $n$  – числа каналов и  $N$  – мест в буфере для ожидания в очереди. Другие характеристики системы передачи данных для определенности возьмем равными следующим часто принимающим значений для подобных узлов:  $\lambda = 2800$  пак/с.

Исходя из проведенных расчётов, в качестве минимального значения получены решения:

$$C_{min} = 0.412, V = 75 \cdot 10^6, n_{min} = 1, N_{min} = 9.$$

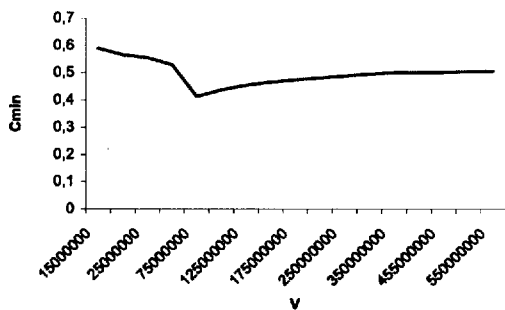


Рис. 1. Целевая функции ( $C$ ), описывающая убытки узла системы

Скорости передачи информации изменяются от  $15 \cdot 10^6$  бит/сек до  $700 \cdot 10^6$  бит/сек, число каналов  $n$  – от 1 до 10, объем буфера  $N$  – от 1 до 10. Для каждого значения скорости в качестве результата выводятся значения  $n$  и  $N$ , для которых величина убытка  $C_{min}$  является минимальной.

Следующим этапом рассмотрим модель  $M/M/n/M$  с использованием *RAID* массива. В данном случае размер буфера будет очень большой. Для получения её минимального значения рассчитаем значения, когда они будут изменяться. Характеристики системы для определенности возьмем равными следующих часто принимающих значений для подобных узлов:  $\lambda = 2800$  пак/с., средняя длина каждого пакета  $L = 127$  Мбит/с. В качестве величин убытка в системе с использованием внешнего массива в единицу времени предположим  $b_1 = 0.5$ ,  $b_2 = 0.01$ ,  $b_3 = 0.01$ ,  $b_4 = 0.1$  (рис. 2).

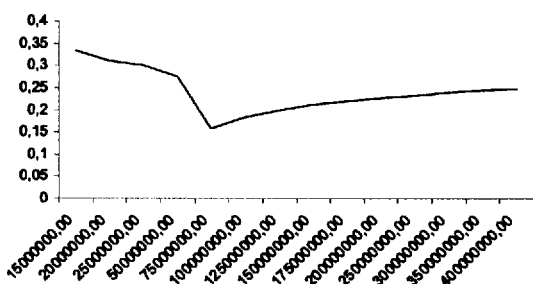


Рис. 2. Величина суммарного убытка в системе с использованием внешнего массива

Проведя расчёты величины суммарного убытка сети хранения данных, в качестве минимального значения получены решения

$$C_{min} = 0.1573, V = 250 \cdot 10^6, n_{min} = 1, N_{min} = 9.$$

## Выводы

Учитывая вышеизложенное, можно сделать вывод, что при проектировании систем на основании средних значений параметров входных потоков сообщений можно вычислить основные характеристики узлов: скорость передатчика ( $V$ ), количество необходимых каналов ( $n$ ) и объем буфера накопителя ( $N$ ), для которых потери системы обслуживания будут минимальны.

При выборе модели следует, учитывать параметры и характеристики основных видов сетевого трафика, поскольку при различных статистических моделях трафика характеристики сетей могут меняться в достаточно широких пределах.

## Список литературы

1. Столлингс В. Современные компьютерные сети. 2-е издание. – СПб.: Питер, 2003. – 783 с.
2. Клейнрок Л. Теория массового обслуживания. – М.: Машиностроение, 1979.
3. Олифер В.Г., Олифер Н.А. Компьютерные сети. Принципы, технологии, протоколы: Учебник для вузов. 2-е издание. – С.Пб.: Питер, 2003. – 894 с.
4. Вишневский В.М. Теоретические основы проектирования компьютерных сетей. – М.: Техносфера, 2003. – 512 с.