Applied Mathematics & Information Sciences 5 (2) (2011), 113S-121S
– An International Journal

Research on Data Service Reliability Based upon Limited Buffer for Data Grid

QU Ming-Cheng, WU Xiang-Hu and Yang Xiao-Zong
School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China
Email Address: qumingcheng@126.com

Received May 17, 2010; Revised December 05, 2010

In order to build reliable data service for a data grid based upon an unreliable grid node, a service failure model based upon a parallel transmission mechanism is proposed by introducing a limited buffer on the consumption side to describe the quantitative relationship among the size of buffer, the probability of data service failure, the probability of node failure, the transmission speed of each node, the number of nodes, the size of the replica and the probability of data failure. The correctness of the model is proved by comparing the theoretical values with experimental results, analyzing the difference between theoretical values and experimental results and detecting the impact on service failure of a variety of parameters.

Keywords: Data grid; data service failure model; limited buffer.

1 Introduction

It is a current challenge to provide reliable data services based upon unreliable grid nodes to ensure the reliability of the service of a data grid [1-2]. In all kinds of applications based upon network buffer technology is always used to overcome the uncertainty of network and grid node [3-4].

Currently there are interrelated researches in the field of streaming media [5]. In streaming media applications the data that cannot be consumed is cached on the consumer node so that, when transmission is interrupted, the cached data can support the service to continue for a long time. However, the data in a streaming media application is not very large so that the size of buffer is always not to be considered and all the data which cannot be consumed in time is cached. Most streaming media applications transmit data from single server based upon multithreading technology [6].

Many parallel transmission algorithms based on multi-opy and gridFTP from multi-servers in the grid community enhance download speed greatly [7]. There is less research based upon the two strategies of parallel transmission and limited buffer. Qu [8] proposed a service failure model based upon parallel transmission mode and buffer. The model can be used to calculate the approximate probability of service failure under given conditions.

However, the size of buffer is not be considered in the deduction process. So the feasibility of its practical application is poor.

Now it is urgently needed to use a model based upon parallel transmission mode effectively to express the quantitative relation among the size of buffer, service failure probability and other important parameters.

2 Buffer model and definition

2.1 Buffer Model Based on Parallel Transmission

As is shown in Figure 2.1, it is assumed that in the model there are 3 identical replicas in three nodes, each of the three nodes transmits data at rate of V_1, V_2, V_3 to the data consumption node and the consumption node consumes data at rate of V_0. The others continue to transmit data when one node is not available. During this period of transmission, the data that has not been consumed by the data consumer is cached into its buffer. Obviously the buffer size is a key parameter and affects the data consumer directly. If the size is too small, only a little data can be cached so that, when some grid nodes are unavailable, the data service of consumer fails. On the other hand,if a large buffer is allocated, maybe waste occurs.

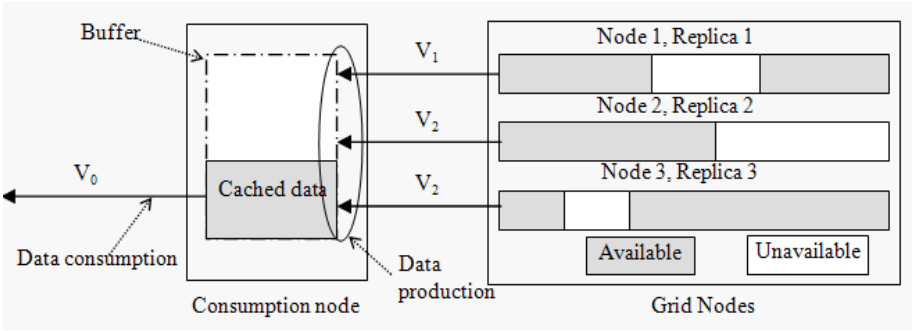


Figure 2-1 Buffer model based upon parallel transmission

2.2 Basic Definitions

Let the buffer size be m, data size of the replica be M and the average data consumption rate of the consumer be V_0. Based upon these parameters we can deduce the service model.

The probability of failure of grid node N_i is defined as P_i. Let node-i transmit data to the consumer node at rate of V_i, $1 \le i \le k$, and k nodes are used to store the same copy, while the consumption node download in parallel from k storage nodes and the sum of average download speed is $\sum_{i=1}^{k}(1-P_i)V_i$.

Definition 2.1. (service failure probability, *f*): When the cached data on the consumer side is 0, then a service failure happens so that the probability of failure occurrence is called service failure probability.

Definition 2.2. (data failure probability, *F*): k copies of data are stored separately in

k nodes and the failure probability of each node is P_i. Let $F = \prod_{i=1}^{k} P_i$. F is called the data failure probability.

Definition 2.3. (production to consumption ratio, $\theta(k, P, V)$): The ratio of the sum of the average transmission (production) rate for each copy of the node to the speed of consumption is defined as $\theta = \left(\sum_{i=1}^{k}(1-P_i)V_i\right)/V_0$ and θ>1. It is called the production to consumption ratio. It can also be expressed as: $\theta V_0 = \sum_{i=1}^{k}(1-P_i)V_i$.

The purpose of Definitions 2.2 and 2.3 is equivalently to make multiple nodes into a whole which has the probability of failure F and average data transfer rate θV_0.

Definition 2.4. (effective buffer limit, m_0): The data that has not been consumed in the data production process in time is cached on the consumption-side. The maximum amount of cached data in a data request is called as the effective buffer limit (for effective use) and expressed as m_0.

It can be seen from the definition of production to consumption ratio that, when the cached data grows to the maximum amount, the data produced upon completion of data production is M and the amount of data consumed at this time is M-m_0 (i.e., the amount of data produced minus the amount of data consumed), which can be expressed as shown below:

$$\theta = \frac{M}{M - m_0} \text{ or } m_0 = M(1 - \frac{1}{\theta}) \tag{2-1}$$

Definition 2.5. (cache speed): The production rate minus the consumption rate is the net growth rate of the data per unit time in the buffer. See Eq. (2-2)

$$(\theta - 1)V_0 \tag{2-2}$$

Condition (1) The time of data failure for each grid node during the production of all the data is continuous. Due to the cached data an empty buffer caused by continuous failure is more likely to occur and result in a greater service failure probability. We can thus reach the upper bound of service failure probability.

Condition (2) When the buffer is not full in the whole period of data consumption, the production rate is θV_0 (see definition 2.3), i.e. the total transmission speed of all the nodes; when the buffer is full, the speeds of production and consumption are equal, i.e. V_0. As long as the cached data is not 0, service failure never occurs.

2.3. Average Data Production Time

The data production rate differs whether the buffer is full or not and so the corresponding data production time is not the same. The average data production time is derived by comparing buffer size m and size of effective buffer limit m_0.

(1) If $0 \le m < m_0$, $t_1 = \frac{m}{(\theta - 1)V_0}$ can be obtained using Eq. (2-1). It occurs when the buffer is full and the data is not completely downloaded. If one assumes that the time when the buffer is just full is t_1, the remaining production time is t_2. The following Eqs. (2-3) and (2-4) can be derived:

$$t_1 = \frac{m}{(\theta - 1)V_0} \tag{2-3}$$

where $(\theta - 1)V_0 = \theta V_0 - V_0 = \sum_{i=1}^{k}(1 - P_i)V_i - V_0$ (as defined in 3), namely, the production rate minus the consumption rate (cache speed: Eq. (2-2)). The meaning of the formula is: the buffer size divided by cache speed, derived t_1 the time is required by the buffer changes from the time when it is empty to when it is full and

$$t_2 = \frac{M - V_0\theta t_1}{V_0} = \frac{M - V_0\theta\dfrac{m}{(\theta - 1)V_0}}{V_0} = \frac{M}{V_0} - \frac{\theta m}{(\theta - 1)V_0} \tag{2-4}$$

where $(V_0\theta)t_1$ represents the amount of data downloaded at time t_1, and $M - (V_0\theta)t_1$ represents the remaining amount of data. To t_1 the buffer is full. By condition ② the cache speed is V_0 at this time, which is equal to consumption rate while $\dfrac{M - V_0\theta t_1}{V_0}$ indicates the time required for the remaining data production. Eq. (2-4) is derived by substituting t_1 for simplification.

Then the total expected completion time for data transmission is:

$$T_1 = t_1 + t_2 = \frac{m}{(\theta - 1)V_0} + \frac{M}{V_0} - \frac{\theta m}{(\theta - 1)V_0} = \frac{M - m}{V_0} \tag{2-5}$$

(1)If $m \geq m_0$: If the buffer overflow situation which causes the production rate to be equal to the consumption speed does not occur in the data production process, then the overall data production time expected can be expressed as shown below, i.e. the total requested data divided by the whole data production rate.

$$T_2 = \frac{M}{\theta V_0} \tag{2-6}$$

The average data production time can be derived by combining formulas (2-5) and (2-6) as shown below.

$$T(m) = \begin{cases} \dfrac{M - m}{V_0} & 0 \leq m < m_0 \\ \dfrac{M}{\theta V_0} & m \geq m_0 \end{cases} \tag{2-7}$$

Theorem 2.1. *T(m) is a continuous function when $0 \leq m < +\infty$.*

Proof: *T(m)* is a piecewise function in $0 \ m<+\infty$. The demarcation point is m_0. Substitute $m=m_0$ into Eq. (2-6): $T = \dfrac{M - m_0}{V_0}$. Then substitute formula (1) into it: $T = \dfrac{M - M(1 - 1/\theta)}{V_0} = \dfrac{M}{\theta V_0}$, where $T^+(m_0) = T^-(m_0)$. Hence the piecewise function *T(m)* is a continuous function. The proof is completed.

Formula (2-6) can be expressed as:

$$T(m)=\frac{(M-m)}{V_0}\quad\begin{cases}0\le m\le m_0\\ m=m_0(m>m_0)\end{cases} \tag{2-8}$$

Conditions ③ From Definition 2.4 the effect buffer is still m_0 when $m>m_0$ and so just $m\le m_0$ is discussed in following sections.

According to condition ① and Theorem 1 in the whole data production period the average time length of data failure can be expressed as:

$$\overline{T}=F\cdot T(m)=F\frac{(M-m)}{V_0}\quad\begin{cases}0\le m\le m_0\\ m=m_0(m>m_0)\end{cases} \tag{2-9}$$

3 Service Failure Model

According to the size of $\overline{T}V_0$ and m, two types of model should be deduced.

3.1 Service Failure Model When $\overline{T}V_0 < m$

As $\overline{T}V_0 < m$, i.e., $0<\frac{FM}{1+F}<m$ (with Eq. (2-9) substituted into), a service failure happens in the given interval [0, t_1]. Once the buffer is full, even if data failure occurs, the cached data can continue to provide data services to the consumer side so that no service failure arises.

Let the maximum time when service failure possibly occurs in (0, t_1) be t_p. It satisfies Eq. (3-1), namely, when the amount of cached data $(\theta-1)V_0t_p$ (we can obtain this from formula (2-2)) is less than the amount of data consumed during the data failure period $\overline{T}V_0$, a service failure occurs.

$$(\theta-1)V_0t_p<\overline{T}V_0\Rightarrow t_p<\frac{\overline{T}}{\theta-1}\qquad\left(\frac{t_p}{\overline{T}}<1\right) \tag{3-1}$$

From formula (3-1) the maximum t_p can be

$$t_p=\frac{\overline{T}}{\theta-1}\qquad\left(\frac{t_p}{\overline{T}}<1\right) \tag{3-2}$$

Then the length of time that the service failure occurs in [0, t_p] is shown in Eq. (3-3), where $(\theta-1)V_0t$ is the amount of cached data up to t, divided by V_0, which indicates the length of time when the data is consumed. The data failure time (Eq. (2-9)) minus this time derives the length of service failure time at t.

$$\bar{t}_p=\overline{T}-\frac{(\theta-1)V_0t}{V_0}=\overline{T}-(\theta-1)t\quad(0<t<t_p) \tag{3-3}$$

From Eq (3-3) we can see $\bar{t}_p$ is a function of t. If we want to get an average service failure time, then a integral should be calculated. By solving the integral of $\bar{t}_p$ in range [0, t_p] and letting the result be divided by total length $\delta=\frac{F}{f}=F\frac{(F(M-m)+2M)}{F(M-m)}\cdot\frac{(\theta-1)}{F}=\frac{(\theta-1)}{F}\left(F+\frac{2M}{M-m}\right)$, an average time of service failure (t_f) is derived in the range, see Eq

(3-4).

Now we know the average service failure time is t_f. The valid data service time is M/V_0. So the total data service time is t_f+M/V_0. Let the ratio of service failure time to the total data service time be r, see Eq. (3-5). So r is the data service failure probability when data failure occurs in $[0,t_p]$.

$$t_f = \frac{\int_0^{t_p} \bar{t}_p dt}{t_p} = \frac{\int_0^{t_p} \left(\overline{T} - (\theta-1)t\right) dt}{t_p} = FT - \frac{(\theta-1)t_p}{2} = FT - \frac{(\theta-1)}{2}\frac{FT}{\theta-1} = \frac{FT}{2} = F\frac{(M-m)}{2V_0} \quad (3\text{-}4)$$

$$r = \frac{t_f}{t_f + \frac{M}{V_0}} = \frac{F\frac{(M-m)}{2V_0}}{F\frac{(M-m)}{2V_0} + \frac{M}{V_0}} = \frac{F(M-m)}{F(M-m)+2M} \quad (3\text{-}5)$$

The distribution of data failure time is arbitrary whereas the probability P_p that failure occurs in $[0, t_p]$ is Eq. (3-6), i.e., t_p divided by the average data production time.

$$p_p = \frac{t_p}{T} = \frac{F}{\theta-1} < 1 \Rightarrow \frac{\theta-1}{F} > 1 \quad (3\text{-}6)$$

We can achieve a condition: if P_p , then r, i.e., if data failure occurs in $[0,t_p]$, then the probability of service failure probability is r.

Now a service failure model can be derived as shown in formula (3-7).

$$f = rp_p = \frac{F(M-m)}{\left(F(M-m)+2M\right)} \times \frac{F}{(\theta-1)} \quad (3\text{-}7)$$

In order to compare service failure probability with the average data failure probability, their ratio is defined as δ, as shown below.

$$\delta = \frac{F}{f} = F\frac{\left(F(M-m)+2M\right)}{F(M-m)} \cdot \frac{(\theta-1)}{F} = \frac{(\theta-1)}{F}\left(F + \frac{2M}{M-m}\right) \quad (3\text{-}8)$$

3.2 Service Failure Model When $\overline{TV}_0 \geq m$

When $\overline{TV}_0 \geq m$, i.e., $F(M-m) \geq m > 0 \Rightarrow 0 < m \leq \frac{MF}{1+F}$, the amount of data consumed during the data failure period is greater than or equal to the buffer size. There are two possible locations for the time when data failure occurs depending upon whether the buffer is full or not full.

We can know from Eq. (2-3) that the time when the buffer is full is t_1 and the time from t_1 to transmission completion is t_2 (Eq. (2-4)).

Data failure occurs in $[0, t_1]$: The deduction process is the same with Eq. (3-3) except that the interval of t should become $0<t<t_1$. According to the calculation process of Eq. (3-4) the following formula can be derived by replacing the upper limit of the integral with t_1 and doing the following calculations.

$$t_f = \frac{\int_0^{t_1} \bar{t}_p dt}{t_p} = \frac{\int_0^{t_1} \left(\overline{T} - (\theta-1)t\right) dt}{t_1} = FT - \frac{(\theta-1)t_1}{2} = FT - \frac{(\theta-1)}{2}\frac{m}{(\theta-1)V_0} = FT - \frac{m}{2V_0}$$

We get the service failure time in $[0, t_1]$ as shown in formula (14).

$$t_f = FT - \frac{m}{2V_0} \tag{3-9}$$

Data failure occurs in [t_1, t_2]: If data failure occurs during t_2, as the buffer is full, the cached data is m. The length of time that the cached data can continue to support data service is m/V_0.

So in the interval [t_1, t_2] the time span of service failure (t_f) is the total duration of data failure (Eq. (2-9)) minus the time when the cached data has been completely consumed (m/V_0). The time, t_f', can be derived using Eq. (3-10)

$$t_f' = FT - \frac{m}{V_0} \tag{3-10}$$

The complete service failure model can be expressed as:

$$f = f_1\frac{t_1}{T} + f_2\frac{t_2}{T} = \frac{t_f}{t_f + M/V_0}\times\frac{t_1}{T} + \frac{t_f'}{t_f' + M/V_0}\times\frac{t_2}{T} \tag{3-11}$$

When $m \neq 0$, a detailed model of service failure can be obtained by substituting Eqs. (2-3), (2-4), (2-8), (3-9) and (3-10) into formula (3-11) as shown below.

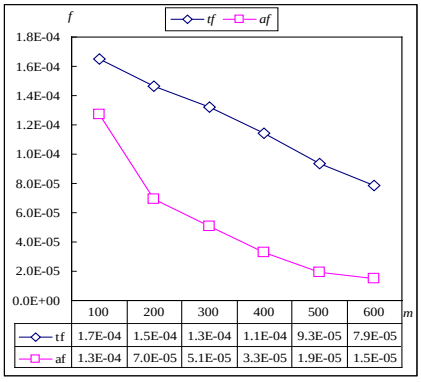
$$f = \frac{2F(M-m)-m}{2F(M-m)-m+2M}\cdot\frac{m}{(\theta-1)(M-m)} + \frac{F(M-m)-m}{F(M-m)-m+M}\cdot\frac{(\theta-1)M-\theta m}{(\theta-1)(M-m)} \tag{3-12}$$

4 Experiments

Physical environment: The experiments were performed in a simulated environment. We configure five nodes to act as replica nodes and we also simulate a consumer node to get data from the replica nodes. Meanwhile the network transmission speed is simulated.

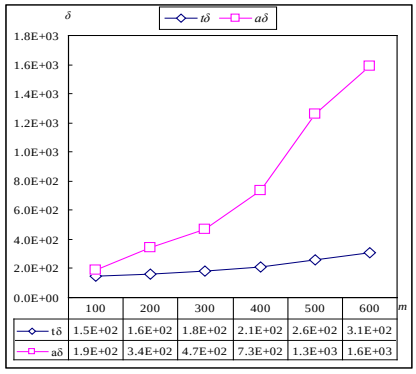
4.1 Experiment 1: condition (1) of model (18) is met

Parameters: Selected three grid nodes, configure $V_0=100$, $V_1=100$, $V_2=120$, $V_3=140$, $P_1=0.3$, $P_2=0.4$, $P_3=0.2$, size of replica is 1*G*. We configured different size of buffer by (100,200,300,400,500,600) Mbytes in consumer side separately.



	100	200	300	400	500	600
tf	1.7E-04	1.5E-04	1.3E-04	1.1E-04	9.3E-05	7.9E-05
af	1.3E-04	7.0E-05	5.1E-05	3.3E-05	1.9E-05	1.5E-05

Figure 4.1 Impact of m on f when $\overline{TV}_0 < m$



	100	200	300	400	500	600
tδ	1.5E+02	1.6E+02	1.8E+02	2.1E+02	2.6E+02	3.1E+02
aδ	1.9E+02	3.4E+02	4.7E+02	7.3E+02	1.3E+03	1.6E+03

Figure 4.2 Impact of m on δ when $\overline{TV}_0 < m$

Experimental Analysis

Based upon the values we obtained Figures 4.1 and 4.2 are drawn. Here from Figure

4.1 we can see clearly that, as m increases, tf and af show a downward trend, and af is always better than tf.

Figure 4.2 shows the trend of δ from the theoretical value $t\delta$ calculated from the model and the experimental value $a\delta$. It can be seen that service failure probability is far less than data failure probability and also $a\delta$ is always better than $t\delta$.

Here the difference between experimental results and theoretical values is from one to five times and the experimental results are all better than theoretical values. Why?

Analysis: F, used in the model derivation, is a whole approximate equivalence of the failure of each node (assuming the failure of each node occurs at the same time) and service failure depends upon the buffer state (empty, nonempty). It makes a increase of probability of empty buffer so that the value calculated by the service failure model should be the upper bound of the actual service failures.

4.2 Experiment 2: condition (2) of model (18) is met

Here we configured different size of buffer by (6,9,12,15,18,21) Mbytes. Based upon the values we obtained Figures 4.3 and 4.4 are drawn.

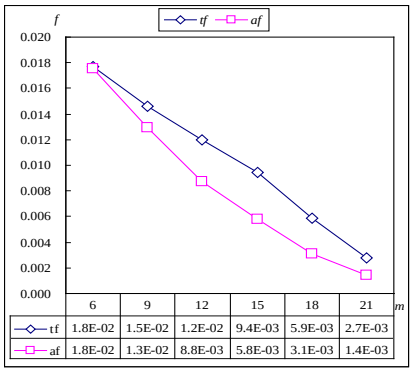


Figure 4.3 Impact of m on f when $\overline{TV}_0 \geq m$

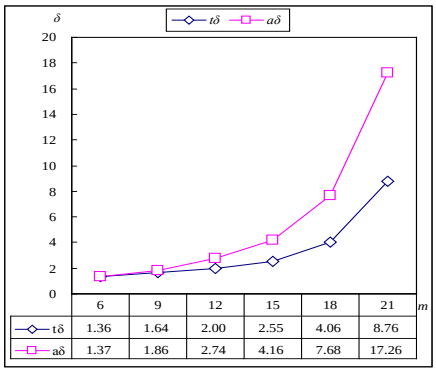


Figure 4.4 Impact of m on δ when $\overline{TV}_0 \geq m$

From Sections 4.1 and 4.2: The experimental service failure probability is approximately consistent with theoretical ones and the experimental value takes the theoretical value as its upper bound.

5. Conclusion

The model effectively describes the relation between the size of buffer and service failure probability and the impact of various parameters (P\V\k\M\F) on buffer size and data service failure. Experimental results indicate that the theoretical results are approximately consistent with the experimental results. Both theoretical proof and experimental results indicate that by the introduction of buffer model based on parallel transmission the probability of service failure becomes far less than the probability of data failure, that is, with fewer copies a greater reliability of data services can be got. The

theoretical service failure probability is the upper bound of actual service failure probability so that models can be used to determine whether the determinate reliability of services can be met under given conditions.

References

[1] C. Lei and L. S. Li, A calking dynamic replication distribution algorithm in data grid. *Acta Electronica Sinica*, 11, 2006:1-4

[2] T. Dong, C. S. Yu and C. Feng, A dynamic fault detection algorithm under grid environments, *Journal of Computer: Research and Development*, 11,2006,: 1870～1875

[3] K. Shi. A replication and cache based distributed metadata management system for data grid, *SNPD2007*, IEEE Computer Society:20-25

[4] M. Mitzenmacher, Digital fountains: A survey and look forward, *Proceedings of IEEE Information Theory Workshop*, (2004) October 1-2; San Antonio, USA: 271-276

[5] P. K. Prasad and G. G. Das, Congestion controlling for streaming media through buffer management and jitter control. *IJCSNS International Journal of Computer Science and Network Security*. 2, 2009:1-10

[6] R. Rangaswami and Z. Dimitrijevic, E. Chang. MEMS-based disk buffer for streaming media servers. *ACM Transactions on Storage*, 2, 2007:1-31

[7] Sudharshan and Vazhkudai, Enabling the Co-Allocation of Grid Data Transfers *Journal of Grid Computing*, 2, (2004):1-8

[8] Q. M. Cheng, W. X. Hu and L. M. Hong, Research of service reliability for data grid based on production & consumption model and slack time. *Computer Integrated Manufacturing Systems*, 11, 2009:2166-2172

Ming-Cheng Qu is a Ph.D. candidate in the School of Computer Science and Technology of Harbin Institute of Technology (HIT). He received his BS and MS degree from HIT. His research interests include Grid Computing etc.

Xiang-Hu Wu is a Professor in the School of Computer Science and Technology of Harbin Institute of Technology (HIT). He is an advanced member of CCF. His research interests include Grid Computing and Embedded Computing.