Replica selection and downloading based on wavelength availability in λ-grid networks

Kouji Hirata[1], Khamisi Kalegele[2], Yoshinobu Higami[1], Shin-ya Kobayashi[1]
[1]Graduate School of Science and Engineering, Ehime University, Matsuyama 790-8577, Japan
[2]Graduate School of Information Sciences, Tohoku University, Sendai 980-8579, Japan
Email: [1]{hirata, higami, kob}@cs.ehime-u.ac.jp, [2]kalegs@ka.riec.tohoku.ac.jp

***Abstract*—This paper proposes a replica selection and downloading scheme in λ-grid networks. In λ-grid networks, in order to distribute loads and achieve high performance computing, a large amount of data is replicated on storage servers as files, and clients download these replicas. To download replicas efficiently, an appropriate replica selection scheme which avoids wavelength contention is required because λ-grid networks employ optical networking. The proposed scheme provides Replica Selection based on Wavelength Availability (RSWA). According to RSWA, a replica is selected based on wavelength availability collected by backward reservation. Furthermore, to suppress blocking probability of download attempts and improve downloading time of replicas, the proposed scheme introduces multi-wavelength downloading, which determines the number of wavelengths used for each downloading based on wavelength availability. Through simulation experiments, we show that the proposed scheme can reduce blocking probability of download attempts and improve average downloading time efficiently.**

***Index Terms*—λ-grid, replica selection, file replication, WDM, lightpath**

I. Introduction

Grid computing integrates geographically distributed computing resources, such as CPU and storage, through communication networks. A data grid, one of the grid computing techniques, manages distributed data and offers high performance computing [4], [6]. Fig. 1 shows a typical model of a data grid network, which consists of many sites. A single site comprises a master node, a storage node, and a computing node. The master node schedules jobs and exchanges information such as the amount of free space of data storage. The storage node stores data files required by jobs and the computing node executes jobs. When a user has a job to execute, the user submits it to a master node of a local site. The job is executed at the local site with data files stored in a storage node. If the data files required by the job do not exist in the local site, they are downloaded from a remote site. Then the job is executed. We call a local site "client" and a remote site "server" hereafter.

In data grid networks, data files are often huge. Thus, if the transmission rate of a data grid network is not sufficiently high, clients take longer to download data files. In such a case, computing resources are wasted

Manuscript received April 19, 2010; revised August 8, 2010; accepted September 1, 2010.

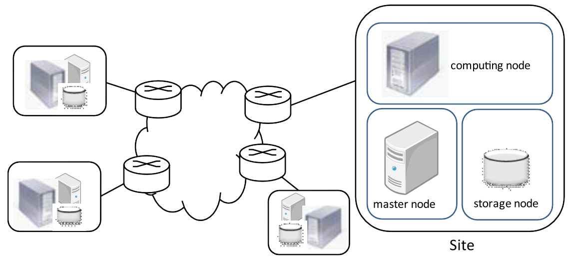

Fig. 1. A data grid network.

because a client cannot start a job during the download even though a computing node is available. To enhance the transmission rate, λ-grids which employ Wavelength Division Multiplexing (WDM) and lightpaths have been proposed [5], [9]. The WDM increases the capacity of a fiber optic link by simultaneously transmitting multiple signals with different wavelengths over a single fiber. A data file is transmitted via a lightpath which is a path assigned a dedicated wavelength. Because the lightpath is established between a client and a server before transmission, transmission bandwidth is guaranteed and thus reliable transmission is realized. Note that two or more lightpaths cannot be established with the same wavelength in the same link at the same time because wavelengths are dedicated. Wavelength conversion avoids such wavelength contention by converting one wavelength to another, and it improves performance of optical networks significantly [20]. However, the wavelength conversion technology is still immature and it requires extra hardware (e.g., wavelength converters and lasers for wavelength conversion) and control software. Without wavelength conversion, a lightpath must use a common wavelength in all links along the route (i.e., wavelength continuity constraint), so that lightpath establishments are often blocked. In λ-grid network, when a lightpath establishment is blocked, a corresponding file download attempt is also blocked and thus a client cannot execute a job. Therefore, wavelength contention is one of the significant issues to be resolved in λ-grid networks.

In order to distribute loads, data files are replicated on multiple sites in λ-grid networks. Each client can download those replicas. Note that blocking probability of download attempts of replicas depends on replica selection schemes such as [2], [10], [13], [14], [17]–[19],

doi:10.4304/jcm.5.9.692-702

[22] which select a replica to be downloaded based on certain information, e.g., RTT, server loads, and traffic status. Most existing replica selection schemes, however, do not consider wavelength resources because they are assumed to work in data grids without optical networking, so that they may not work well in λ-grid networks.

This paper proposes a replica selection and downloading scheme based on wavelength availability, which aims at improving blocking probability of download attempts and average downloading time. The proposed scheme provides Replica Selection based on Wavelength Availability (RSWA). The RSWA selects a replica based on wavelength availability collected by backward reservation [21], which is one of wavelength reservation protocols in optical networks. By doing so, RSWA utilizes wavelength resources efficiently even though wavelength usage in the network is unknown to the clients in advance. Furthermore, the proposed scheme introduces Multi-wavelength Downloading based on Wavelength Availability (MDWA). In general, the size of files shared in λ-grid networks is very huge. Thus files should be downloaded with multiple wavelengths (i.e., lightpaths) to reduce downloading time [8]. However, file downloading with multiple wavelengths often generates bottleneck links because it simultaneously uses more wavelength resources than that with a single wavelength. As a result, while the average downloading time decreases, blocking probability of download attempts increases. In order to suppress this trade-off, MDWA determines the number of wavelengths used for each downloading based on wavelength availability. The proposed scheme is expected to utilize the wavelength resources efficiently and thus improve blocking probability of download attempts and average downloading time.

The rest of this paper is organized as follows. In Section II, we describe existing replica selection schemes and their shortcomings. Section III explains wavelength reservation protocols in optical networks. Section IV discusses our proposed scheme. In Section V, the performance of the proposed scheme is discussed with the results of simulation experiments. Finally, we conclude the paper in Section VI.

II. File replication and replica selection

The size of data shared in data grid networks is generally huge. File downloading always consume large amounts of bandwidth. In order to resolve this problem, files are replicated on multiple sites in data grid networks. File replication reduces the bandwidth consumption and helps in load balancing. Replicas are dynamically created and deleted at each site according to replication schemes [1], [3], [12], [15], [16] or caching schemes [7], [11]. In replication schemes, a storage node decides when and where to create a replica of one of its files as necessary. On the other hand, in caching schemes, when a file is downloaded from a remote site, a storage node in a local site stores it. If there is no storage space, the storage node replaces stored files with the new file, based on information such as usage of files.

When a client downloads a file, it selects a replica from among replicas of the file stored in the network by means of replica selection schemes. In the past, many replica selection schemes, including schemes which use parallel downloading, have been proposed in data grid networks [2], [10], [13], [14], [18], [19], [22] . However, λ-grid networks differ from conventional data grid networks because in λ-grid networks, data files are transmitted via lightpaths. Those replica selection schemes do not consider wavelength availability, and thus they may not work well in λ-grid networks. We, therefore, need an appropriate replica selection scheme which avoids wavelength contention by considering wavelength availability.

As a replica selection scheme for λ-grid networks, [17] proposed a novel replica (server) selection scheme. In this scheme, replicas are selected while considering network resource availability. This scheme distributes loads for links and suppresses the generation of bottleneck links. As a result, this scheme reduces blocking probability of download attempts. However, in this scheme, every client needs to know wavelength availability of the whole network in advance in order to select replicas. On the other hand, RSWA selects a replica based on information on wavelength availability collected by backward reservation whenever a client tries to download a data file. Thus RSWA does not need to know wavelength availability in advance.

III. Wavelength reservation protocol

In optical networks, before data transmission, a lightpath is established by reserving a wavelength in all links along a route between a sender and a receiver. There are two types of wavelength reservation protocols: forward reservation and backward reservation [21]. In the forward reservation, a sender node decides which wavelength is used to transmit data. When a transmission request arrives, the sender node selects an available wavelength and transmits a RESV message to reserve the wavelength as shown in Fig. 2. The RESV message reserves the wavelength in all links along a route between the sender and the receiver. If wavelength contention occurs at an output port of an intermediate node, the RESV message is discarded at the node and the lightpath establishment is blocked. When the RESV message reaches the receiver node, an ACK message which will propagate from the receiver to the sender is generated. After receiving the ACK message, the sender node sends data. In general, the forward reservation has high blocking probability because sender nodes cannot get the wavelength information along routes.

On the other hand, in the backward reservation, a receiver node selects a wavelength. When a transmission request arrives, a sender node sends a PROB message as shown in Fig. 3. The PROB message collects information on available wavelengths in each link along a route. Note that it does not reserve wavelength resources. When the

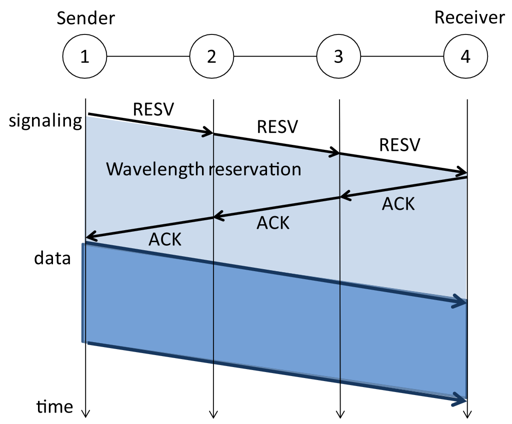

Fig. 2. Forward reservation.

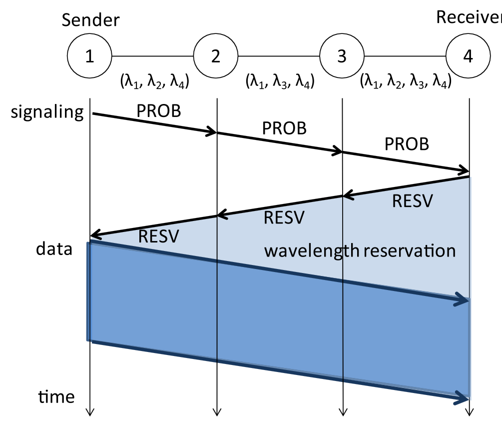

Fig. 3. Backward reservation.

PROB message reaches the receiver node, the receiver node selects a wavelength from a set of available wavelengths along the entire route based on certain criteria. Then the receiver node sends a RESV message to reserve the selected wavelength in each link along the route to the sender node. After the sender node receives the RESV massage, it sends data to the receiver node. The backward reservation can reduce blocking probability more efficiently than the forward reservation because wavelength usage in all links along a route is known before selection. Furthermore, duration of reservation in the backward reservation is smaller than that in the forward reservation. Note that when a PROB message detects that all wavelengths have already been reserved by other transmissions, the lightpath establishment is blocked. We call this blocking “forward blocking”. Also, when wavelength reservation by a RESV message collides with wavelength reservation by other RESV messages, the corresponding lightpath establishment is blocked. We call this blocking “backward blocking”.

We show an example of the backward reservation with Fig. 3. We assume that node 1 is a sender and node 4 is a receiver. At first, node 1 sends a PROB message. In this example, wavelengths λ_1, λ_2, and λ_4 are available between nodes 1 and 2. Similarly, wavelengths λ_1, λ_3, and λ_4 are available between nodes 2 and 3, and wavelengths λ_1, λ_2, λ_3, and λ_4 are available between nodes 3 and 4. Therefore, available wavelengths along the entire route are λ_1 and λ_4. After receiving the PROB message which includes these wavelength information, node 4 chooses a wavelength between those wavelengths. Then node 4 transmits a RESV message to reserve the wavelength along the route.

IV. Proposed Scheme

The proposed scheme consists of RSWA and MDWA. Note that these schemes can be applied to λ-grid networks independently. In this section, we explain the details of these schemes.

A. The proposed replica selection scheme

A replica is selected based on certain information such as server loads, the number of hops, RTT, traffic status, file size, and wavelength availability. Choice of a set of such information depends on requirements of systems. This paper aims at reducing blocking probability of download attempts. To efficiently avoid blocking of download attempts caused by blocking of lightpath establishments, it is important to suppress the generation of bottleneck links. Specifically, we expect to reduce blocking probability of download attempts by decreasing the number of wavelengths simultaneously used in each link. Furthermore, as we will see in Section IV-B, we assume that the selected replica is downloaded with multiple wavelengths. In order to efficiently download replicas with multiple wavelengths, we need to consider wavelength availability. Thus RSWA selects a replica based on not only the number of hops but also wavelength availability.

In what follows, we first explain the procedure of collecting information on wavelength availability. Then we explain how to select a replica. Note that we assume that the backward reservation is used as a wavelength reservation protocol. Note also that we assume that each client knows which servers have target replicas by using a replica catalog which has location information on replicas [4].

1) Collecting information with backward reservation: In conventional schemes, a client selects a replica first based on certain information such as wavelength availability, and then send a download request to a server with the selected replica. After receiving it, the server sends a PROB message to a client in order to establish a lightpath as shown in Fig. 4(a), where the server is a sender and the client is a receiver. In this case, the client must have information needed for replica selection in advance.

On the other hand, in RSWA, clients do not need to have such information before replica selection. RSWA collects information needed for replica selection using PROB messages of backward reservation every time a job arrives. The procedure is as follows. When a job arrives, a client sends download requests to multiple servers with a replica of a target file as shown in Fig. 4(b). After

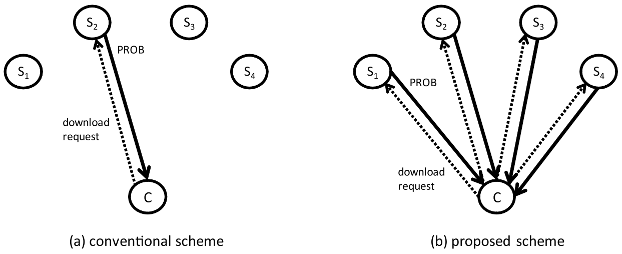

Fig. 4. Replica selection (C denotes a client and S_k $(k = 1, 2, 3, 4)$ denotes a server).

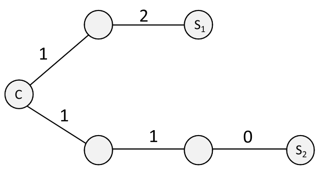

Fig. 5. Wavelength availability (C denotes a client and S_k $(k = 1, 2)$ denotes a server).

receiving the request, each server sends a PROB message to the client in order to establish a lightpath, based on the shortest path routing (in terms of the number of hops). The PROB message collects information on wavelength availability in links along the route, which is required for replica selection. The client waits until it receives all PROB messages transmitted from the requested servers. After receiving them, the client selects a replica based on information collected by these PROB messages, as we will discuss in Section IV-A2. Then the client sends a RESV message to only a server whose replica has been selected. The client also sends messages which cancel download requests to other servers and does not establish any lightpaths along routes to them. Because a lightpath is established along only a route between the client and the server with the selected replica, RSWA uses wavelength resources efficiently.

2) Replica selection based on wavelength availability: Let $\mathcal{L}_i$ denote a set of links along a route between a client and a server i. We denote the number of used wavelengths in link l by U_l. We define W_l as the total number of wavelengths in link l. We also define C_i as the cost of a replica in server i:

$$C_i = \sum_{l \in \mathcal{L}_i} \frac{U_l}{W_l}.$$

RSWA select a replica stored in a server i with the minimum cost C_i from among replicas stored in reachable servers each of which has at least one available wavelength along the entire route. If there are two or more replicas with the minimum cost, RSWA selects one randomly. Note that to obtain required information, i.e., U_l and W_l, RSWA have to collect such information in all links along routes by using PROB messages.

Fig. 5 shows an example of RSWA, where the numbers on links denote the numbers of used wavelengths. We assume that the total number W_l of wavelengths in all links is 4. The cost C_{S_1} of a replica in server S_1 is $1/4 + 2/4 = 3/4$. Similarly, cost C_{S_2} of a replica in server S_2 is $1/4 + 1/4 + 0/4 = 2/4$. Thus RSWA selects the replica in server S_2 even though its number of hops is larger.

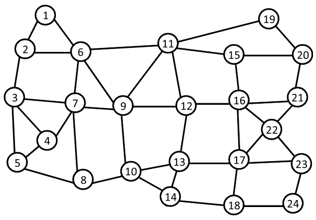

Fig. 6. Network model.

B. The proposed multi-wavelength downloading scheme

In order to reduce downloading time, this paper assumes that a selected replica is downloaded with multiple wavelengths. Thus, multiple wavelengths are reserved by a RESV message and multiple lightpaths are established between a client and a server with selected replica. Although file downloading with multiple wavelengths reduces downloading time, blocking probability of download attempts increases as the number of wavelengths (lightpaths) used for each downloading increases. This is because the increase in the number of wavelengths for each downloading causes the increase in wavelength resource utilization in each link, so that bottleneck links are often generated. Therefore, we should determine the number of wavelengths used for each file downloading carefully.

MDWA aims to suppress the generation of bottleneck links in order to reduce blocking probability of download attempts efficiently. To do so, MDWA uses fewer wavelengths for each downloading when a load along a route between a client and a server is high. On the other hand, when the load is low, MDWA uses many wavelengths in order to reduce downloading time. As a measure of the load, MDWA uses the number of available wavelengths along the entire route. Generally, the number of available wavelengths along the entire route tends to be small when the load is high. In contrast, the number of available wavelengths along the entire route tends to be large when the load is low. Based on this idea, MDWA determines the number N of wavelengths used for each downloading as follows:

$$N = \lceil aR \rceil, \tag{1}$$

where R denotes the number of available wavelengths along the entire route between a client and a server with a selected replica, and a $(0 < a \leq 1)$ denotes a parameter.

V. PERFORMANCE EVALUATION

A. Model

To evaluate performance of the proposed scheme, we conduct simulation experiments with a network model shown in Fig. 6. It consists of 24 nodes and 43 bi-directional links. Each node has a site. To establish lightpaths, we use wavelength routing with backward reservation and clients randomly select wavelengths from a set of available wavelengths. We assume that there are no wavelength converters at any nodes. For simplicity, we assume that the propagation delay of each link is equal to 1 [msec], and processing times of PROB and RESV messages at each node are 0.1 [msec]. We also assume that the bandwidth D of each wavelength is equal to 10 [Gbps]. The total number W of wavelengths in each link is set to be 32, unless stated otherwise. The number F of different files is set to be 24 and the storage size S of each site is set to be 10 [Tbyte], unless stated otherwise. Every original file is stored on one of sites. We define ρ as the offered load per wavelength:

$$\rho = \frac{8 \times \lambda \times L \times 10^3}{D \times W},$$

where λ [1/sec] denotes the average arrival rate of job execution requests at each site and L [Tbyte] denotes the average file size. In order to dynamically create and delete replicas, we use a simple caching scheme which replaces one or more stored replicas with the oldest access time with a new replica, namely LRU, every time the replica is downloaded. We collect 30 independent samples from simulation experiments, and 95% confidence intervals are shown in each figure (even though most of them are invisible).

B. Performance in a homogeneous model

In this section, we evaluate the performance of the proposed scheme in a homogeneous model. In this model, the size L_f of file f ($f = 1, 2, \cdots, F$) is fixed to 2 [Tbyte]. We assume that user requests of job execution at each site are generated according to a Poisson process with rate λ, and a target file is independently chosen equally likely among all possible files except original files stored at the site. We also assume that a client sends download requests to all servers with a target replica whenever a job arrives in the proposed scheme.

1) Basic performance of the proposed replica selection scheme: First, we evaluate the performance of RSWA without MDWA. In this section, we apply a simple multi-wavelength downloading scheme which determines the number N of wavelengths used for each downloading by

$$N = \min\{R, M\}, \tag{2}$$

where M ($M = 1, 2, \cdots W$) denotes maximum wavelength threshold. Specifically, a file is downloaded with M wavelengths when $R > M$. When $R \leq M$, a file is downloaded with R wavelengths, i.e., all available wavelengths.

For the sake of comparison, we use the following two replica selection schemes. In the first scheme, a client sends a download request to only one nearest server (in terms of the number of hops), based on the shortest path routing. The server then sends a PROB message to establish lightpaths. In the second scheme, a client sends download requests to all servers with a target replica based on the shortest path routing, similar to RSWA. The servers then send PROB messages to the client. The client selects a replica stored in a nearest server (in terms of the number of hops) from replicas stored in reachable servers. We call the first scheme "conventional-A" and the second scheme "conventional-B" hereafter. Note that if there are two or more shortest paths between a client and a server, we select one randomly whenever a job execution request arrives.

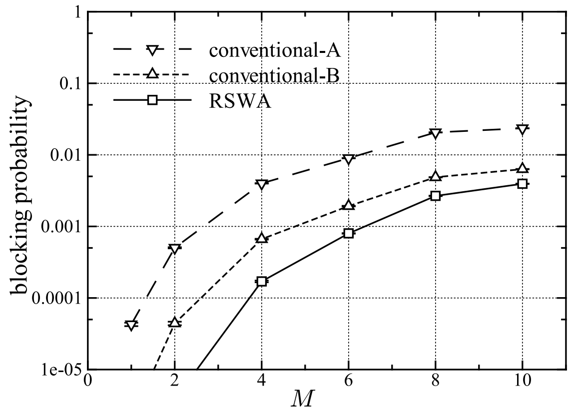

Fig. 7. Blocking probability ($\rho = 0.6$).

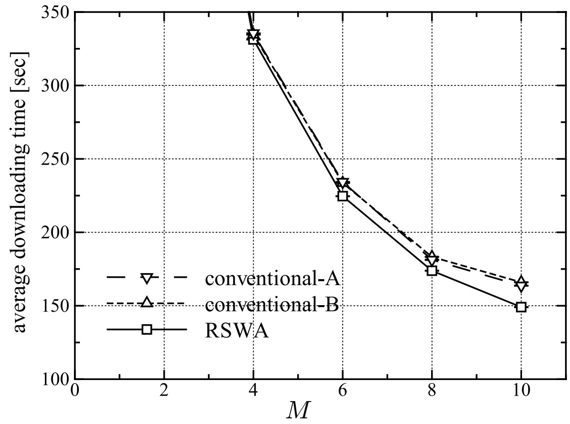

Fig. 8. Average downloading time ($\rho = 0.6$).

Fig. 7 shows the blocking probability of download attempts as a function of M for $\rho = 0.6$. Also, Fig. 8 shows the average downloading time as a function of M for $\rho = 0.6$. Note that the blocking probability BP of download attempts is defined as

$$\text{BP} = \frac{\text{\# of blocked file download attempts}}{\text{total \# of file download attempts}}.$$

File downloading is blocked when all PROB messages sent from servers with a target replica are discarded by forward blocking or when a RESV message cannot reserve any selected wavelengths along the entire route between a client and a server with a selected replica due to backward blocking. Downloading time is the time interval from when a client sends download requests to servers to the time when download of a selected replica finishes. We assume that downloading time is 0 when a replica is stored in a local site. As we can see in Fig. 7, the blocking probability increases with M, i.e., the maximum

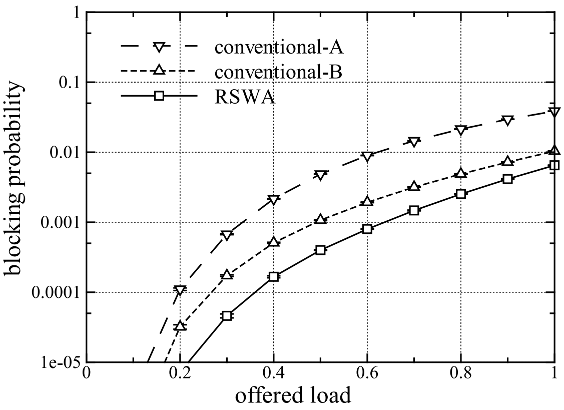

Fig. 9. Blocking probability ($M = 6$).

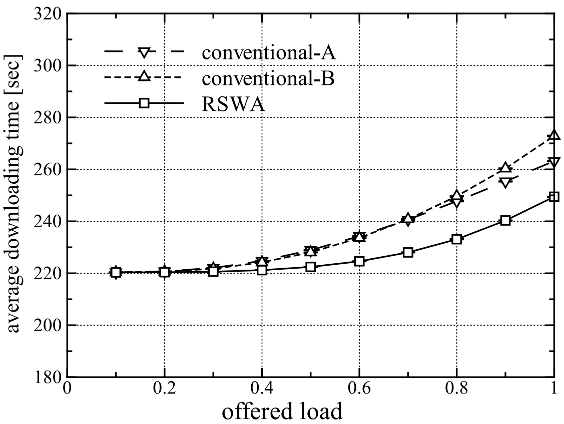

Fig. 10. Average downloading time ($M = 6$).

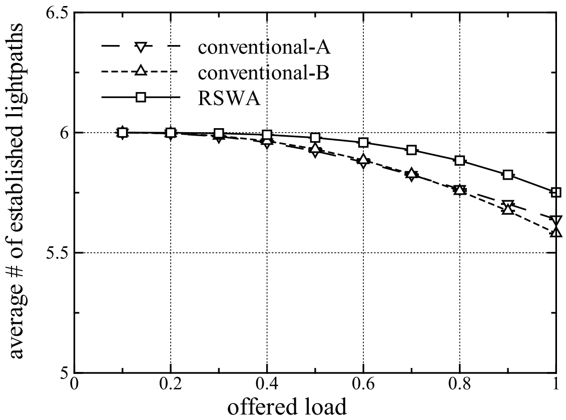

Fig. 11. Average number of lightpaths used for downloading ($M = 6$).

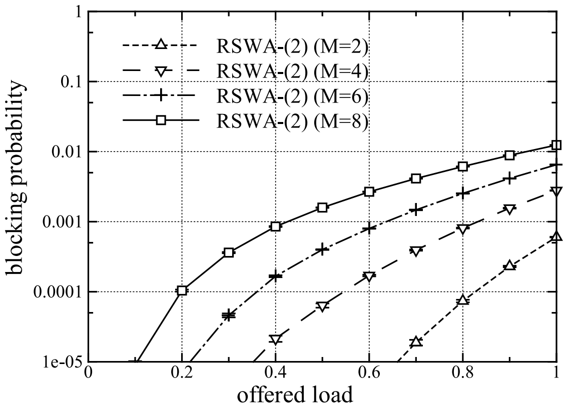

Fig. 12. Blocking probability.

number of wavelengths for each downloading, in all schemes. On the other hand, from Fig. 8, we observe that the average downloading time decreases as M increases. These results confirm that there is a trade-off between the blocking probability of download attempts and the average downloading time. We also observe that RSWA reduces the blocking probability of download attempts and improves the average downloading time because it uses wavelength resources more efficiently.

Fig. 9 shows the blocking probability of download attempts as a function of ρ, where $M = 6$. We observe that conventional-A shows high blocking probability. In conventional-A, a client sends a download request to only one server with a selected replica when a job arrives. The server sends a PROB message in order to establish lightpaths. If this PROB message is discarded by forward blocking, a corresponding file downloading is also blocked. On the other hand, RSWA and conventional-B send download requests to multiple servers. If at least one PROB message reaches the client, the client can continue to establish lightpaths. Therefore, RSWA and conventional-B reduce the blocking probability more efficiently than conventional-A. Furthermore, we observe that RSWA exhibits the excellent performance. This result implies that RSWA utilizes wavelength resources more effectively.

Fig. 10 shows the average downloading time as a function of ρ, where $M = 6$. We observe that RSWA decreases the average downloading time efficiently. We explain the cause of this phenomenon using Fig. 11, which shows the average number of lightpaths established in each downloading as a function of ρ, where $M = 6$. As we can see from this figure, RSWA establishes more lightpaths for each downloading than other schemes. Therefore, RSWA reduces the average downloading time efficiently.

2) Performance of the proposed multi-wavelength downloading: Next, we evaluate the performance of MDWA, which determines the number of wavelengths used for each downloading by (1). In this section, we call RSWA with MDWA "RSWA-(1)". We also call RSWA with multi-wavelength downloading determined by (2) "RSWA-(2)".

Figs. 12 and 13 show the blocking probability of download attempts and the average downloading time in RSWA-(2) respectively, as a function of the offered load. Also, Figs. 14 and 15 show the blocking probability of download attempts and the average downloading time in RSWA-(1) respectively, as a function of the offered load. Note that the increase in a means the increase in the number N of wavelengths used for each downloading. For example, when $a = 0.125$, the maximum number of wavelengths used for each downloading is 4. When $a = 0.375$, the maximum number of wavelengths used

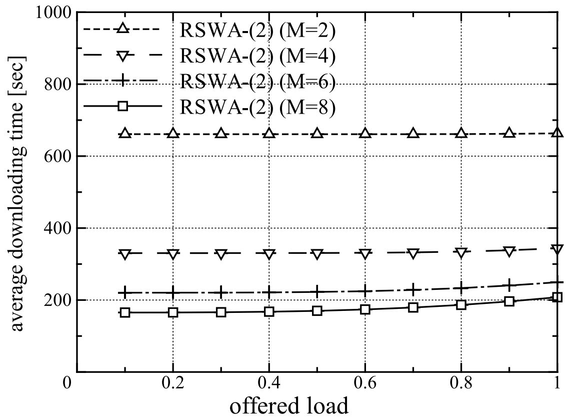

Fig. 13. Average downloading time.

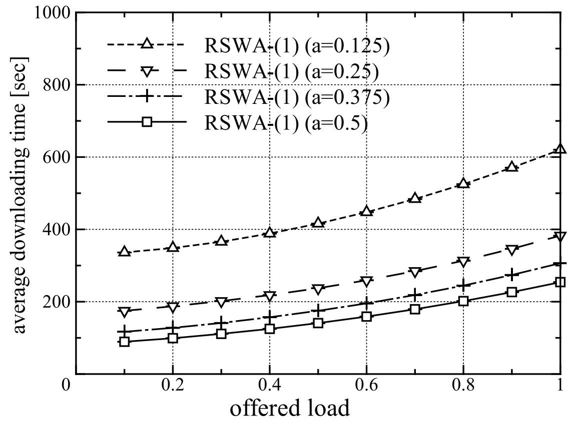

Fig. 15. Average downloading time.

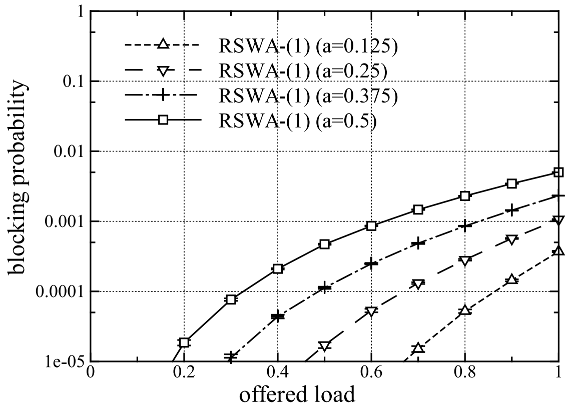

Fig. 14. Blocking probability.

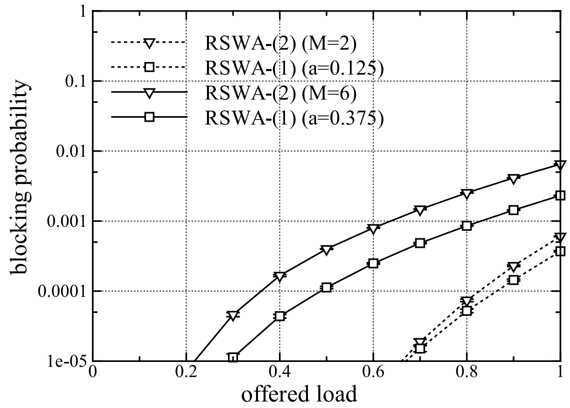

Fig. 16. Blocking probability.

for each downloading is 12. We observe that the blocking probabilities increase with the number N of wavelengths used for each downloading in both schemes. On the other hand, the average downloading time decreases with the increase in N in both schemes.

Although RSWA-(1) and RSWA-(2) show similar characteristics, RSWA-(1) has the superior performance. Figs. 16 and 17 show the blocking probability of download attempts and the average downloading time respectively, as a function of the offered load. In these figures, we compare RSWA-(1) with $a = 0.125$ to RSWA-(2) with $M = 2$. Also, we compare RSWA-(1) with $a = 0.375$ to RSWA-(2) with $M = 6$. We observe that both of the blocking probability and the average downloading time in RSWA-(1) with $a = 0.125$ are smaller than those in RSWA-(2) with $M = 2$. We also observe that RSWA-(1) with $a = 0.375$ reduce the blocking probability more efficiently than RSWA-(2) with $M = 6$. Furthermore, the average downloading time in RSWA-(1) with $a = 0.375$ is smaller than that in RSWA-(2) with $M = 6$, when $\rho < 0.8$. These results suggest that we can reduce both of the blocking probability and the average downloading time efficiently by strategically selecting a.

Fig. 18 shows the average number of lightpaths established in each downloading as a function of the offered load. We observe that RSWA-(1) establishes more lightpaths than RSWA-(2) especially in lightly loaded situations. Therefore, RSWA-(1) can reduce the average downloading time as shown in Fig. 17. Moreover, RSWA-(1) can reduce the blocking probability as shown in Fig. 16. These results imply that RSWA-(1) utilizes wavelength resources effectively.

3) The superiority of the proposed scheme: We examine the superiority of RSWA with MDWA over conventional-A with MDWA and conventional-B with MDWA. Fig. 19 shows the blocking probability of download attempts as a function of the offered load, where $a = 0.25$. Also, Fig. 20 shows the average downloading time as a function of the offered load, where $a = 0.25$. Note that in all schemes, the number of wavelengths used for each downloading is determined by (1), i.e., MDWA, in this section. We observe that RSWA improves both the blocking probability of download attempts and the average downloading time effectively. Furthermore, Figs. 21 and 22 show the blocking probability of download attempts and the average downloading time respectively, as a function of a, where $\rho = 0.6$. The blocking probability of download attempts in each scheme increases with a. On the other hand, the average downloading time decreases with the increase of a. These results coincide with the observation in Section V-B2. We also observe that RSWA decreases the blocking probability of download attempts and the average downloading time efficiently, regardless of a value of a. Thus we conclude that RSWA shows

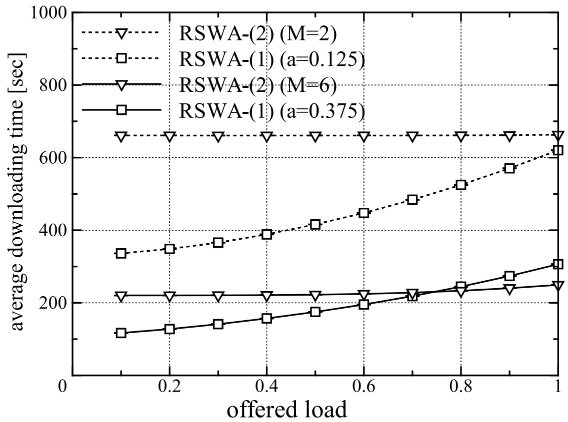

Fig. 17. Average downloading time.

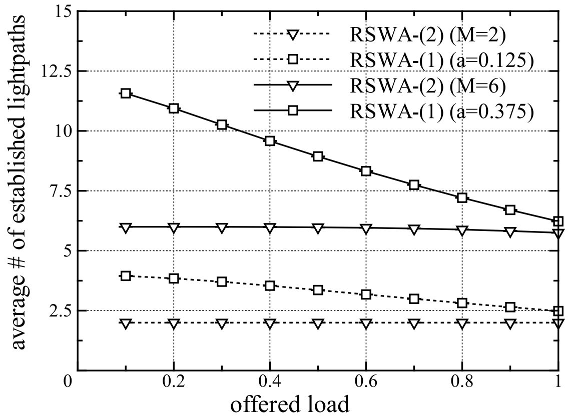

Fig. 18. Average number of lightpaths used for downloading.

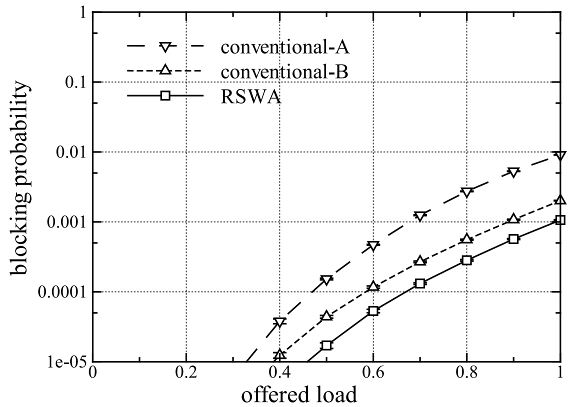

Fig. 19. Blocking probability ($a = 0.25$).

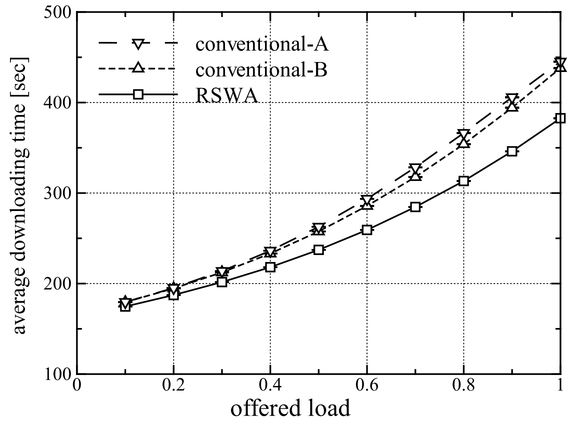

Fig. 20. Average downloading time ($a = 0.25$).

excellent performance when MDWA is applied.

Next, we examine the performance of RSWA against the total number W of wavelengths. Fig. 23 shows the blocking probability of download attempts as a function of W, where $\rho = 0.6$ and $a = 0.25$. The blocking probability of download attempts in each scheme decreases with the increase of W because of the large-scale effect. We also observe that RSWA reduces the blocking probability of download attempts more effectively than conventional schemes.

Fig. 24 shows the average downloading time as a function of W, where $\rho = 0.6$ and $a = 0.25$. The average downloading time decreases with the increase of W because the number of wavelengths used for each downloading increases with W. We also observe that RSWA reduces the average downloading time more effectively than conventional-A and conventional-B. We conclude that RSWA with MDWA keeps the superior performance, regardless of the total number W of wavelengths.

We examine the impact of the number F of different files. Fig. 25 shows the blocking probability of download attempts as a function of the number F of different files, where $\rho = 0.6$ and $a = 0.25$. Also, Fig. 26 shows the average downloading time against F, where $\rho = 0.6$ and $a = 0.25$. We observe that the blocking probability and the average downloading time of each scheme increase with F. This is because the number of replicas of each file stored in the network decreases and thus each downloading wastes wavelength resources. We also observe that RSWA exhibits an excellent performance in any case.

Finally, we examine the impact of the storage size S at each site. Figs. 27 and 28 show the blocking probability of download attempts and the average downloading time as a function of S, respectively, where $\rho = 0.6$ and $a = 0.25$. We observe that the blocking probability and the average downloading time decrease with the increase of the storage size because the increase of the storage size means that the number of replicas stored at each site increases. We also observe that RSWA improves the blocking probability and the average downloading time efficiently. We conclude that RSWA keeps the superior performance, regardless of storage size S.

C. Performance in a more realistic scenario

We evaluate the performance of our scheme in a more realistic scenario. For this purpose, we assume that user requests of job execution at site k are generated according to a Poisson process with rate λ_k which is proportional to the ratios in Table I. We also assume that a target file is chosen based on weights shown in Table II. Note that the values in these tables are randomly generated. Furthermore, we assume that the size L_f of file f follows

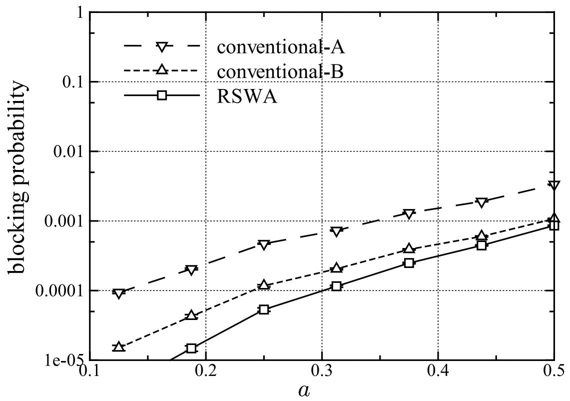

Fig. 21. Blocking probability ($\rho = 0.6$).

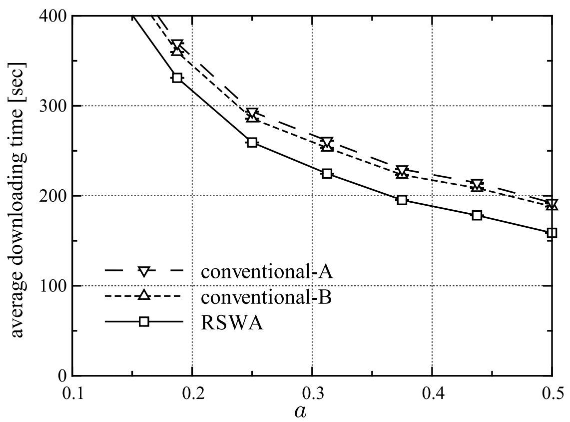

Fig. 22. Average downloading time ($\rho = 0.6$).

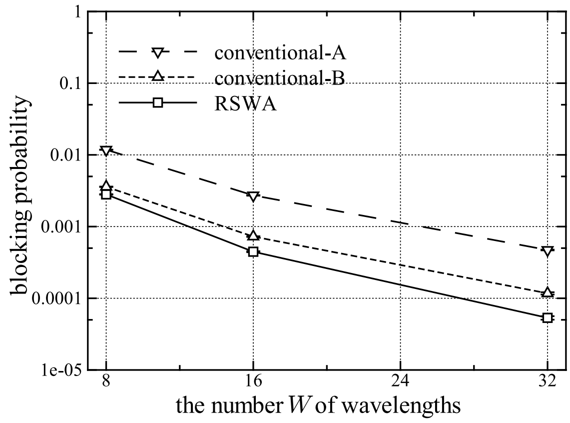

Fig. 23. Blocking probability ($\rho = 0.6$, $a = 0.25$).

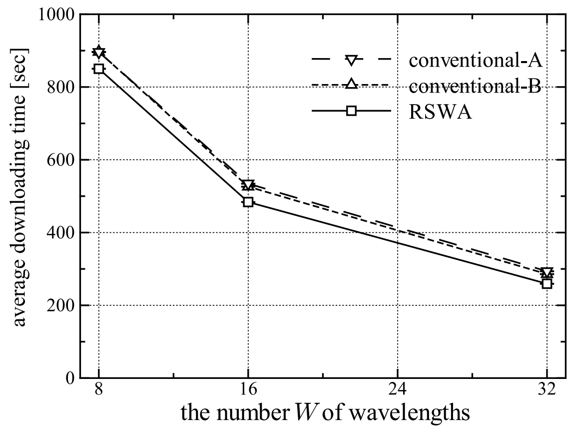

Fig. 24. Average downloading time ($\rho = 0.6$, $a = 0.25$).

an exponential distribution with mean $L = 2$ [Tbyte], where L_f is normalized in such a way that the total file size is equal to $\sum_{f=1}^{F} L_f = 48$.

Fig. 29 shows the blocking probability of download attempts as a function of the offered load, where $a = 0.25$. Also, Fig. 30 shows the average downloading time as a function of the offered load, where $a = 0.25$. Note that all schemes use MDWA. As shown in Fig. 29, the blocking probability of download attempts in RSWA is smaller than those in conventional schemes. Furthermore, as we can see from Fig. 30, the average downloading time in RSWA is smaller than those in conventional schemes.

Figs. 31 and 32 show the blocking probability of download attempts and the average downloading time as a function of a, respectively, where $\rho = 0.6$. We observe that the blocking probability of download attempts in each scheme increases with a because each scheme with large a uses many wavelengths resources. On the other hand, the average downloading time decreases with the increase of a. We also observe that RSWA decreases the blocking probability of download attempts and the average downloading time efficiently, regardless of a value of a. Generally, RSWA with MDWA shows the excellent performance in terms of both the blocking probability of download attempts and the average downloading time in this scenario, similar to the results in the homogeneous model.

VI. CONCLUSION

This paper proposed a replica selection and downloading scheme based on wavelength availability. The proposed scheme collects information on wavelength availability on routes between a client and multiple servers by means of backward reservation. The client then selects a replica based on collected information. Furthermore, the proposed scheme determines the number of wavelengths used for each downloading based on wavelength availability. Through the simulation experiments, we showed that the proposed scheme reduces the blocking probability of download attempts and improves the average downloading time efficiently.

ACKNOWLEDGMENT

The authors would like to thank Professor Tetsuya Takine of Graduate School of Engineering, Osaka University for his helpful advices.

REFERENCES

[1] W. H. Bell et al., "Evaluation of an economy-Based file replication strategy for a data grid," in *Proc. IEEE International Symposium on Cluster Computing and the Grid (CCGrid 2003)*, Tokyo, Japan, May 2003, pp. 661-668.

[2] R. Chang, M. Guo, and H. Lin, "A multiple parallel download scheme with server throughput and client bandwidth considerations for data grids," *Future Generation Computer Systems*, vol. 24, no. 8, pp. 798–805, 2008.

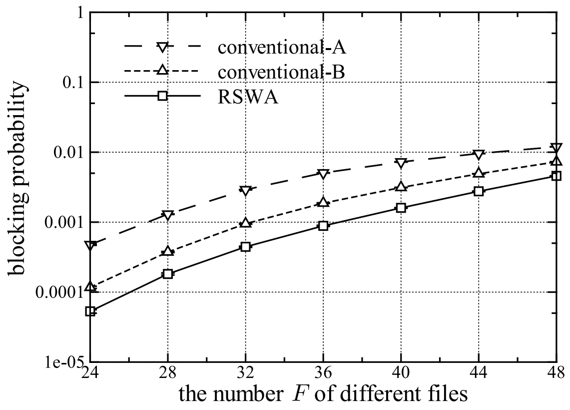

Fig. 25. Blocking probability ($\rho = 0.6$, $a = 0.25$).

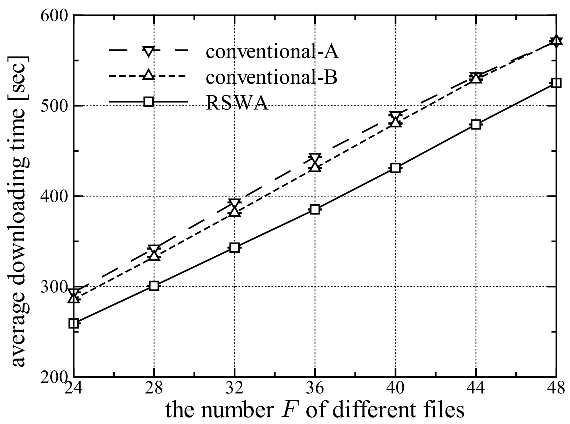

Fig. 26. Average downloading time ($\rho = 0.6$, $a = 0.25$).

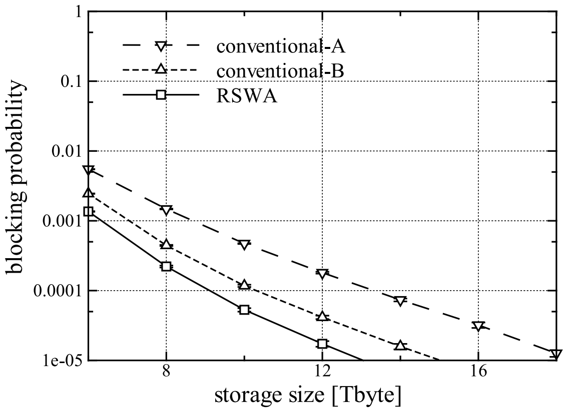

Fig. 27. Blocking probability ($\rho = 0.6$, $a = 0.25$).

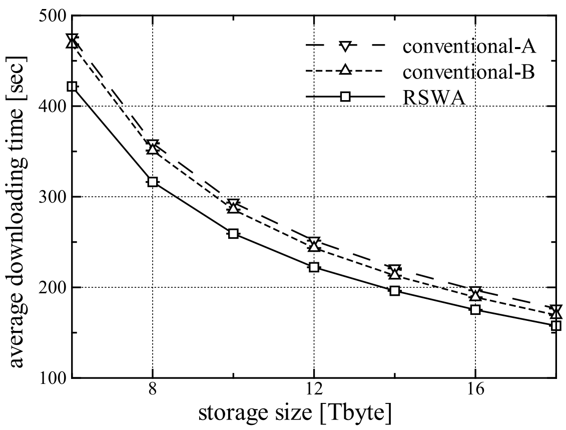

Fig. 28. Average downloading time ($\rho = 0.6$, $a = 0.25$).

[3] R. Chang, H. Chang, and Y. Wang, "A dynamic weighted data replication strategy in data grids," in *Proc. the 2008 IEEE/ACS International Conference on Computer Systems and Applications*, Doha, Qatar, Mar. 2008, pp. 414–421.

[4] A. Chervenak, I. Foster, C. Kesselman, C. Salisbury, and S. Tuecke, "The data grid: towards an architecture for the distributed management and analysis of large scientific datasets," *Journal of Network and Computer Application*, vol. 23, no. 3, pp. 187–200, 2000.

[5] S. Figueira et al., "DWDM-RAM: enabling grid services with dynamic optical networks," in *Proc. IEEE International Symposium on Cluster Computing and the Grid (CCGrid 2004)*, Chicago, IL, Apr. 2004, pp. 707–714.

[6] W. Hoschek, F. J. Jaen-Martinez, A. Samar, H. Stockinger, and K. Stockinger, "Data management in an international data grid project," in *Proc. the First IEEE/ACM International Workshop on Grid Computing*, Bangalore, India, Dec. 2000, pp. 77–90.

[7] S. Jiang and X. Zhang, "Efficient distributed disk caching in data grid management," in *Proc. IEEE International Conference on Cluster Computing*, Williamsburg, VA, Dec. 2003, pp. 446–451.

[8] N .R. Kaushik, S. M. Figueira, and S. A. Chiappari, "A hybrid algorithm for lightpath assignment," in *Journal of Networks*, vol. 4, no. 1, pp. 19–29, 2009.

[9] M. D. Leenheer, C. Develder, T. Stevens, B. Dhoedt, M. Pickavet, and P. Demeester, "Design and control of optical grid networks," in *Proc. Fourth International Conference on Broadband Communications, Networks and Systems (BROADNETS 2007)*, Raleigh, NC, Sep. 2007, pp. 107–115.

[10] H. H. E. AL-Mistarihi and C. H. Yong, "Response time optimization for replica selection service in data grids," *Journal of Computer Science*, vol. 4, no. 6, pp. 487–493, 2008.

[11] E. Otoo and A. Shoshani, "Accurate modeling of cache replacement policies in a data grid," in *Proc. the 20th IEEE/11th NASA Goddard Conference on Mass Storage Systems and Technologies*, San Diego, CA, Apr. 2003, pp. 10–19.

[12] S. Park, J. Kim, Y. Ko, and W. Yoon, "Dynamic data grid replication strategy based on Internet hierarchy," in *Proc. Second International Workshop on Grid and Cooperative Computing (GCC 2003)*, Shanghai, China, Dec. 2003, pp. 838–846.

[13] R. M. Rahman, K. Barker, and R. Alhajj, "Replica selection in grid environment: a data-mining approach," in *Proc. ACM Symposium on Applied Computing*, Santa Fe, New Mexico, Mar. 2005, pp. 695–700.

[14] R. M. Rahman, R. Alhajj, K. Barker, "Replica selection strategies in data grid," *Journal of Parallel and Distributed Computing*, vol. 68, no. 2, pp. 1561–1574, 2008.

[15] K. Ranganathan and I. Foster, "Identifying dynamic replication strategies for a high-performance data grid," in *Proc. the International Grid Computing Workshop*, Denver, CO, Nov. 2001, pp. 75–86.

[16] Q. Rasool, J. Li, G. S. Oreku, and E. U. Munir, "Fair-share replication in data grid," *Information Technology Journal*, vol. 7, no. 5, pp. 776–782, 2008.

[17] R. Usui, H. Miyagi, Y. Arakawa, S. Okamoto, and N. Yamanaka, "A novel distributed data access scheme considering with link resources and metric in lambda grid networks," in *Proc. the 14th Asia-Pacific Conference on Communications*, Tokyo, Japan, Oct. 2008.

[18] S. Vazhkudai, S. Tuecke, and I. Foster, "Replica selection in the Globus Data Grid," in *Proc. First IEEE/ACM International Symposium on Cluster Computing and The Grid*, Brisbane, Australia, May. 2001, pp. 106–113.

[19] C. Yang, I. Yang, C. Chen, and S. Wang, "Implementation of a dynamic adjustment mechanism with efficient replica selection in data grid environments," in *Proc. ACM Symposium on Applied Computing*, Dijon, France, Apr. 2006, pp. 797–804.

[20] S. Yao, B. Mukherjee, and S. Dixit, "A unified study of contention-resolution schemes in optical packet-switched networks," *Journal of Lightwave Technology*, vol. 21, no. 3, pp. 672–83, 2003.

[21] X. Yuan, R. Melham, and R. Gupta, "Distributed path reserva-

TABLE I
RATIO OF JOB EXECUTION REQUESTS FROM SITE k.

k	1	2	3	4	5	6	7	8	9	10	11	12
ratio	0.065	0.030	0.060	0.061	0.070	0.015	0.026	0.059	0.021	0.043	0.037	0.048
k	13	14	15	16	17	18	19	20	21	22	23	24
ratio	0.028	0.039	0.073	0.070	0.049	0.055	0.011	0.047	0.001	0.019	0.011	0.062

TABLE II
RATIO OF TARGET FILES.

no.	1	2	3	4	5	6	7	8	9	10	11	12
ratio	0.042	0.037	0.018	0.037	0.018	0.032	0.025	0.074	0.064	0.039	0.066	0.022
no.	13	14	15	16	17	18	19	20	21	22	23	24
ratio	0.015	0.076	0.025	0.031	0.047	0.076	0.040	0.052	0.047	0.002	0.064	0.051

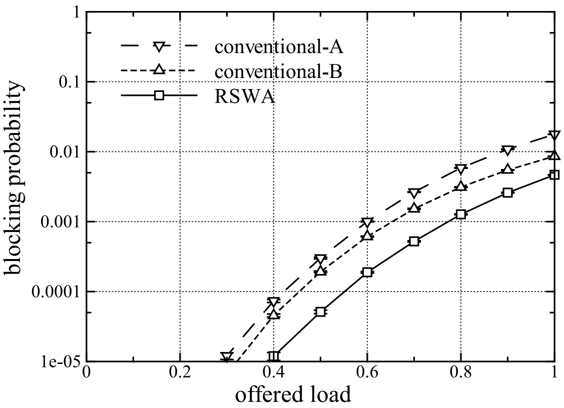

Fig. 29. Blocking probability ($a = 0.25$).

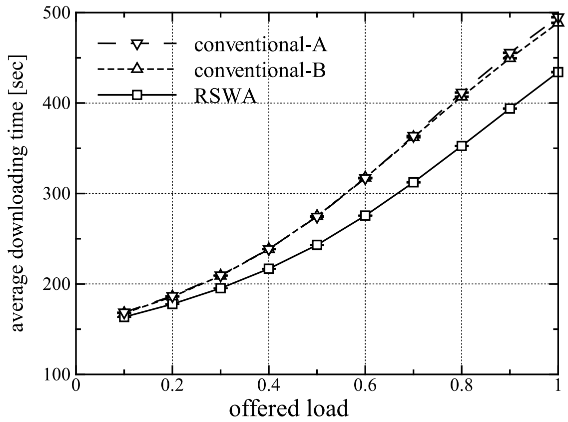

Fig. 30. Average downloading time ($a = 0.25$).

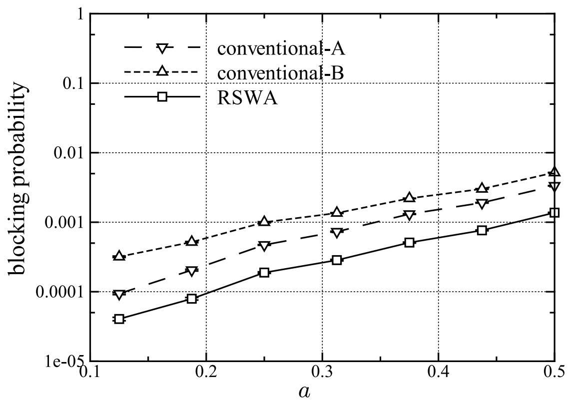

Fig. 31. Blocking probability ($\rho = 0.6$).

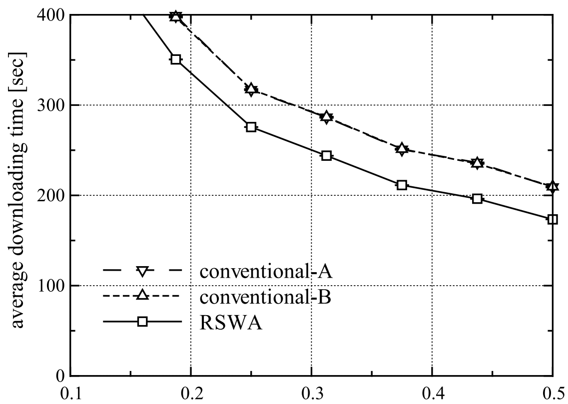

Fig. 32. Average downloading time ($\rho = 0.6$).

tion for multiplexed all-optical networks," *IEEE Transactions on Computers*, vol. 48, no. 12, pp. 1355–1363, 1999.

[22] Y. Zhao and Y. Hu, "GRESS - a grid replica selection service," in *Proc. the ISCA 16th International Conference on Parallel and Distributed Computing Systems*, Reno, NV, Aug. 2003, pp. 423–429.

Kouji Hirata received the B.E. and M.E. degrees in communications engineering from Osaka University in 2003 and 2005, respectively. He received the Ph.D. degree in electrical, electronic, and information engineering from Osaka University in 2008. At present, he is an Assistant Professor in the Department of Electrical and Electronic Engineering and Computer Science at the Graduate School of Science and Engineering of Ehime University. His research interests include optical networks and its performance evaluation. He is a member of IEEE, IPSJ and IEICE.

Khamisi Kalegele is a Ph.D. student at Graduate School of Information Sciences of Tohoku University. He received his M.E in Computer Sciences from Ehime University in 2010. He did his B.Sc in Computer Systems Engineering at the University of Dar es salaam in Tanzania from 1999 to 2003. He has 3 years of experience in GSM radio access network planning,dimensioning and optimization. His interests lie within things related to Information Systems. He is a member of ACM and IPSJ.

Yoshinobu Higami received his B.E., M.E., and D.E. degrees from Osaka University in 1991, 1993 and 1996, respectively. After serving as a research fellow of the Japan Society for the Promotion of Science, he joined Department of Computer Science, Ehime University in 1998. Currently he is an associate professor at Graduate School of Science and Engineering, Ehime University. In 1998 and 2006, he was also an honorary fellow at University of Wisconsin-Madison, U.S.A. He received the IEICE Best Paper Award in 2005. His research interests include test generation, design for testability and fault diagnosis of logic circuits. He is a senior member of the IEEE and a member of IEICE and IPSJ.

Shin-ya Kobayashi received the B.E. degree, M.E. degree, and Dr.E. degree in Communication Engineering from Osaka University in 1985, 1988, and 1991 respectively. He is a Professor at Graduate School of Science and Engineering, Ehime University. His research interests include distributed processing, and parallel processing. He is a member of the Information Processing Society of Japan, the Institute of Electrical Engineers of Japan, IEICE, IEEE, and ACM.