TSR: Topology Reduction from Tree to Star Data Grids

Ming-Chang Lee[#1], Fang-Yie Leu[*2], Ying-ping Chen[#3]

[#]Department of Computer Science, National Chiao Tung University, Taiwan

[*]Department of Computer Science, TungHai University, Taiwan

[1]mingchang1109@gmail.com; [2]leufy@thu.edu.tw; [3]ypchen@cs.nctu.edu.tw

Abstract—**To speed up data transmission of data grids, several co-allocation schemes have been proposed. However, data grids are often large in scale, heterogeneous in participating resources, and complicated in architecture and network topology, consequently increasing the analytical complexity of its data transmission behaviour. In other words, if we can reduce the data transmission topology for the grid, the analysis will be easier. Therefore, in this paper, we propose a topology reduction approach, called the *Tree-to-Star Reduction method* (TSR for short), which can reduce a packet delivery tree topology to a star for a data grid so that the data transmission of a co-allocation scheme can be more conveniently analyzed. Here, a delivery tree topology, as a tree topology rooted at the destination node, is a network topology for delivering all fragments of a file to the destination node.**

***Keywords*—data grid, delivery tree, delivery star, topology reduction, co-allocation scheme**

I. INTRODUCTION

To shorten file retrieval time and improve file replication performance for data grids, several co-allocation schemes have been proposed [1][2][3][4]. However, it is difficult for us to evaluate these schemes' data transmission behaviour since a data grid is often large in scale [5], heterogeneous in participating resources [6], and complicated in architecture and network topology. Coates et al. [7] logically transformed a complicated network topology into a simple tree to simplify their performance analyses. Levitin et al. [8] introduced a concept in which the network topology used to transmit the data generated by different service sites for a grid job can be treated as a star topology in which a logical link connects a service site and the destination node. With this star topology, users can more easily evaluate the execution time of a task and data transmission reliability. However, these studies neither mention how to reduce a complicated network topology to a simple one, nor discuss whether or not the performance of the two topologies is identical. Hence, in this paper, we propose a topology reduction approach, called the Tree-to-Star Reduction method (TSR for short), which can reduce a complicated packet delivery tree topology to a star in a data grid, where a delivery tree topology (or simply a delivery tree), as a tree topology rooted at the destination node, comprises a set of service sites and network components for cooperatively delivering a file or task results generated by the service sites to the destination node. Consequently, the complexity of evaluating a co-allocation scheme can be dramatically reduced.

II. BACKGROUND

A co-allocation scheme is an scheme invoked by the resource broker of a data grid to coordinate a set of service sites $X = \{X_1, X_2, \dots, X_m\}$ to cooperatively deliver all fragments of a file F to X_u where X_u, $X_u \notin X$, is the site at which a user U submits a request to access F [1]. In fact, from U's viewpoint, the routing paths that transmit all fragments of F to X_u form a delivery tree rooted at X_u. Each fragment may pass through several routers and links before arriving at X_u. Due to containing so many network components (i.e., links and routers) and parameters (e.g., link bandwidths and the departure and arrival rates of routers), it may be hard for users to analyse the file transmission time for the scheme in this delivery tree.

III. THE TSR METHOD

The process of reducing a delivery tree to a star is as follows. A delivery tree basically comprises two types of structures. The first is a series structure (SS for short) in which all network components form a chain topology. If the topology contains n nodes and $n-1$ links, $n > 2$, we call it an SS of length $n-1$. Given an SS of length n-1, if there does not exist a node that together with the SS forms an SS of length n, we call the given SS a maximum series structure (MSS for short). The chain topology shown in Fig. 1a, i.e., $X_k - l_{X_k} - Y_1 - l_{Y_1} - \dots - Y_{n-1} - l_{Y_{n-1}} - Y_n$, is an example in which X_k and Y_n are the head and tail nodes, respectively. We wish to reduce the MSS to a simple chain, which consists of the head node X_k and the tail node Y_n connected by a logical link l'_{X_k}, denoted by $X_k - l'_{X_k} - Y_n$ (see Fig. 1b), implying that $l_{X_k} - Y_1 - l_{Y_1} - \dots - Y_{n-1} - l_{Y_{n-1}}$ is simplified to l'_{X_k}. The reduction process is called the series structure reduction process (the SS-reduction process for short).

The other type is a parallel structure (PS for short), which consists of m upstream nodes $n_1, n_2, \dots$, and n_m, $m \geq 2$, and a conjunction node n_{conj}, denoted by $\{(n_1, n_2, \dots, n_m), n_{conj}\}$, in which n_k is connected to n_{conj} through a direct link l_{n_k}, i.e., $n_k - l_{n_k} - n_{conj}$ is a simple chain which is an SS of length 1, $1 \leq k \leq m$, where l_{n_k} may be a physical or logical link. The structure illustrated in the dashed rectangle shown in Fig. 2 is an example. Given a PS $\{(X_1, X_2, \dots, X_m), Y_{conj}\}$, if there exists at least one sub-tree rooted at Y_{conj} (see Fig. 2) with $X' = \{X'_1, X'_2, \dots, X'_p\}$ as its leaf nodes, and

$X' \cap \{X_1, X_2, \ldots, X_m\} = \emptyset$, then we call the PS an expandable PS. If no such sub-tree exists, we call the PS a maximum parallel structure (MPS for short), denoted by MAX $\{(X_1, X_2, \ldots, X_m), \Upsilon_{conj}\}$. The one contained in the dashed rectangle shown in Fig. 3a, i.e., $\{(X_1, X_2, \ldots, X_m), \Upsilon_{conj}\}$, is an MPS.

Given a MAX $\{(X_1, X_2, \ldots, X_m), \Upsilon_{conj}\}$, if there exists an MSS with Υ_{conj} as its head node, e.g., the one shown in Fig. 3a, i.e., $\Upsilon_{conj} - l_{\Upsilon_{conj}} - \Upsilon_1 - \ldots - \Upsilon_n - l_{\Upsilon_n} - X_S$, then after applying the SS-reduction process on the MSS, the resulting structure as depicted in Fig. 3b is an MPS followed by a simple chain $\Upsilon_{conj} - l_{\Upsilon'_{conj}} - X_S$. We call this structure a parallel-chain structure, which will be reduced to a PS $\{(X_1, X_2, \ldots, X_m), X_S\}$ (see Fig. 3c). The corresponding reduction process is called the parallel-chain reduction process (the PC-reduction process for short).

In the following, we assume that 1) a packet delivery tree is drop-free, implying that no packet will be lost and retransmitted during the packet transmission; 2) the queuing model of a node, either a router or a service site, for sending and receiving packets is M/M/1 [9], i.e., the arrival and departure rates follow a Poisson distribution; 3) due to applying the TSR to a data grid, each node n's packet service rate (service rate for short), denoted by μ_n, and packet departure rate (departure rate for short), denoted by μ'_n, are all larger than its packet arrival rate (arrival rate for short), named λ_n. Here, we separate service rate and departure rate since service rate means the maximum capacity in pkts/sec that a node n can transmit, whereas departure rate is the number of packets that n physically sends. For a node n in a drop free delivery tree, the capacity of n is either ∞ or $\mu_n \geq \mu'_n \geq \lambda_n$.

Let $F(l,t)$ be the function used to calculate the transmission time of a packet P through a link l at time point t. Then,

$$F(l,t) = \frac{|P|}{B_l(t)} \tag{1}$$

where $|P|$ is the size of P, and $B_l(t)$ is the bandwidth of l at t. Let $Q(n)$ be the function employed to calculate the expected queueing delay of P in a node n at t. Then,

$$Q(n) = \frac{1}{\mu_n - \lambda_n} - \frac{1}{\mu_n} \text{ [9]} \tag{2}$$

where λ_n is n's arrival rate and μ_n is n's service rate. Let $D(l_n, t)$ be a function used to calculate n's departure rate μ'_n. Then,

$$D(l_n, t) = min\left(\frac{B_{l_n}(t)}{|P|}, \lambda_n\right) \tag{3}$$

where $\frac{B_{l_n}(t)}{|P|}$ is the service rate μ_n of n at t. Let $A(S, n)$ be the function employed to calculate n's arrival rate λ_n where $S = \{s_1, s_2, \ldots, s_c\}$ is the set of n's immediate upstream nodes, in which s_v may be a service site or a router transmitting packets to n, $1 \leq v \leq c$. Then,

$$A(S, n) = A(\{s_1, s_2, \ldots, s_c\}, n) = \sum_{v=1}^{c} D(l_{s_v}, t) \tag{4}$$

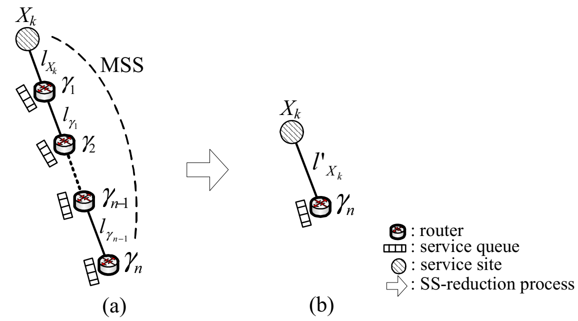


Fig. 1 The topology before and after the SS-reduction process

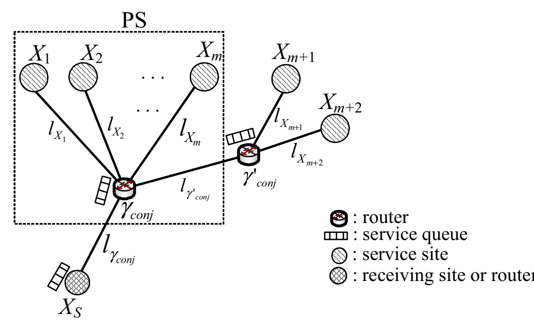


Fig. 2 An expandable parallel structure

A. The SS-reduction process

To describe the SS-reduction process, we assume that the packet generation rate, i.e., departure rate (rather than service rate), of a service site X_k, denoted by μ'_{X_k}, is known and deploy the topology shown in Fig. 1a as an example, in which X_k and Υ_n are respectively the head node and tail node of the MSS, and the set of routers $R = \{\Upsilon_1, \Upsilon_2, \ldots, \Upsilon_{n-1}\}$ are the intermediate nodes and the set of links $L = \{l_{X_k}, l_{\Upsilon_1}, l_{\Upsilon_2}, \ldots, l_{\Upsilon_{n-1}}\}$ are the intermediate links between X_k and Υ_n. The resulting topology is $X_k - l'_{X_k} - \Upsilon_n$. The SS-reduction process consisting of three phases is as follows.

1. The first phase uses the function shown in Eq. (1) to calculate the transmission time of P flowing through a link $l_{\Upsilon_{i-1}}$ at t, named $T_{l_{\Upsilon_{i-1}}}(t)$, where $T_{l_{\Upsilon_{i-1}}}(t) = F(l_{\Upsilon_{i-1}}, t)$, $i = 1, 2, \ldots, n$. When $i = 1$, $\Upsilon_0 = X_k$.
2. The second phase employs the functions shown in Eqs. (2)~(4) to calculate the expected queueing delay of P in router Υ_i at t, named W_{Υ_i}, where $W_{\Upsilon_i} = Q(\Upsilon_i) = \frac{1}{\mu_{\Upsilon_i} - \lambda_{\Upsilon_i}} - \frac{1}{\mu_{\Upsilon_i}}$, in which $\mu_{\Upsilon_i} = \frac{B_{l_{\Upsilon_i}}(t)}{|P|}$, $\mu'_{\Upsilon_i} = D(l_{\Upsilon_i}, t) = min\left(\frac{B_{l_{\Upsilon_i}}(t)}{|P|}, \lambda_{\Upsilon_i}\right)$, $i = 1, 2, \ldots, n$, and when $i = 1$, $\lambda_{\Upsilon_1} = A(\{X_k\}, \Upsilon_1) = D(l_{X_k}, t) = \mu'_{X_k}$, and when $2 \leq i \leq n$, $\lambda_{\Upsilon_i} = A(\{\Upsilon_{i-1}\}, \Upsilon_i) = D(l_{\Upsilon_{i-1}}, t)$.

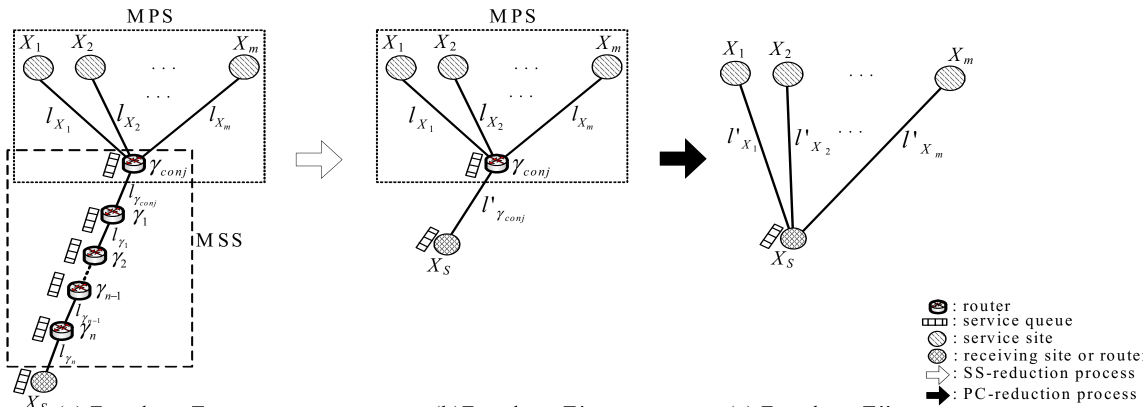


Fig. 3 The topology before and after the SS-reduction and PC-reduction processes. (a) A topology **T** in which an MPS is followed by an MSS; (b) The topology **T'** after reducing the MSS in **T** to a simple chain; (c) The topology **T''** is a PS structure, which is also an MPS after reducing the parallel-chain structure shown in **T'** with the PC-reduction process

3. The third phase calculates the accumulated transmission time of $l_{X_1} - \Upsilon_1 - l_{\Upsilon_1} - \ldots - \Upsilon_{n-1} - l_{\Upsilon_{n-1}}$ which is also the transmission time of P on l'_{X_k}, denoted by $T_{l'_{X_k}}(t)$, where

$$T_{l'_{X_k}}(t) = \sum_{i=1}^{n} F\left(l_{\Upsilon_{i-1}}, t\right) + \sum_{i=1}^{n-1} Q(\Upsilon_i) \quad (5)$$

and the equivalent bandwidth of l'_{X_k} at t, denoted by $B_{l'_{X_k}}(t)$, is

$$B_{l'_{X_k}}(t) = \frac{|P|}{T_{l'_{X_k}}(t)} \quad (6)$$

B. The PC-reduction process

In the following, we use the topology shown in Fig. 3b as an example, i.e., MAX $\{(X_1, X_2, \ldots, X_m), \Upsilon_{conj}\}$ followed by a simple chain $\Upsilon_{conj} - l'_{\Upsilon_{conj}} - X_S$, to describe the PC-reduction process which consists of four phases. Also assume that X_k's packet generation rate, i.e., μ'_{X_k}, is known, $k = 1, 2, \ldots, m$.

1. The first phase also uses Eq. (1) to calculate the transmission time of a link l_{X_k} at t, denoted by $T_{l_{X_k}}(t)$, where $T_{l_{X_k}}(t) = F\left(l_{X_k}, t\right)$, $k = 1, 2, \ldots, m$.
2. The second phase employs Eq. (2) to calculate the queueing delay of P in Υ_{conj} at t, denoted by $\mathrm{W}_{\Upsilon_{conj}}$, $\mathrm{W}_{\Upsilon_{conj}} = Q\left(\Upsilon_{conj}\right) = \frac{1}{\mu_{\Upsilon_{conj}} - \lambda_{\Upsilon_{conj}}} - \frac{1}{\mu_{\Upsilon_{conj}}}$, in which $\mu_{\Upsilon_{conj}} = \frac{B_{l_{\Upsilon'_{conj}}}(t)}{|P|}$ is the service rate of Υ_{conj}'s outlink $l_{\Upsilon'_{conj}}$ at t, $\mu'_{\Upsilon_{conj}} = D\left(l_{\Upsilon'_{conj}}, t\right)$, and $\lambda_{\Upsilon_{conj}} = A\left(\{X_1, X_2, \ldots, X_m\}, \lambda_{\Upsilon_{conj}}\right) = \sum_{k=1}^{m} D\left(l_{X_k}, t\right) = \sum_{k=1}^{m} \mu'_{X_k}$.
3. The third phase also invokes Eq. (1) to calculate the transmission time of P through the link $l_{\Upsilon'_{conj}}$ at t (see Fig. 3b), denoted by $T_{l_{\Upsilon'_{conj}}}(t)$, where $T_{l_{\Upsilon'_{conj}}}(t) = F\left(l_{\Upsilon'_{conj}}, t\right)$.
4. In the fourth phase, the accumulated delay of P on traveling through l_{X_k}, Υ_{conj}, and $l_{\Upsilon'_{conj}}$ at t, denoted by $T_{l'_{X_k}}(t)$, is calculated as

$$T_{l'_{X_k}}(t) = T_{l_{X_k}}(t) + \mathrm{W}_{\Upsilon_{conj}} + T_{l_{\Upsilon'_{conj}}}(t) \quad (7)$$

where $\mathrm{W}_{\Upsilon_{conj}} + T_{l_{\Upsilon'_{conj}}}(t)$ is a constant for all X_k, $1 \le k \le m$, since $\mathrm{W}_{\Upsilon_{conj}}$ and $T_{l_{\Upsilon'_{conj}}}(t)$ are the average values of all the packets flowing through Υ_{conj} and $l_{\Upsilon'_{conj}}$, respectively. The bandwidth $B_{l'_{X_k}}(t)$ of l'_{X_k} at t is also calculated by invoking Eq. (6), $k = 1, 2, \ldots, m$.

C. Tree-to-Star Reduction (TSR) algorithm

In the following, a delivery tree topology rooted at X_u with a set of service sites $X = \{X_1, X_2, \ldots, X_m\}$ shown in Fig. 4a is given as an example to describe the TSR algorithm illustrated in Fig. 5, in which between lines 1 and 3, the TSR first checks to see whether a service site X_k and its following MSS is a simple chain or not (assume that the conjunction node of the MSS is $\Upsilon_{conj,k}$), where $X_u \notin X$ and $X_k \in X$. If not, the TSR utilizes the SS-reduction process to reduce the MSS to a logical link l'_{X_k} connecting X_k to $\Upsilon_{conj,k}$. The result is shown in Fig. 4b. Now each service site is the head node of a simple chain.

Between lines 4 and 10, on identifying a new MPS, e.g., MAX $\{(X_1, X_2, \ldots, X_h), \Upsilon_{conj,k}\}$, the TSR checks to see whether the MSS with $\Upsilon_{conj,k}$ and n as its head and tail nodes, respectively, is a simple chain or not, $h \le m$. If yes, the MPS and the simple chain form a parallel-chain structure. If not, the TSR reduces the MSS to $\Upsilon_{conj,k} - l'_{\Upsilon_{conj,k}} - n$ by using the SS-reduction process. After that, the resulting topology is also a parallel-chain structure. Next, the TSR utilizes the PC-reduction process to reduce the parallel-chain structure to a PS $\{(X_1, X_2, \ldots, X_h), n\}$.

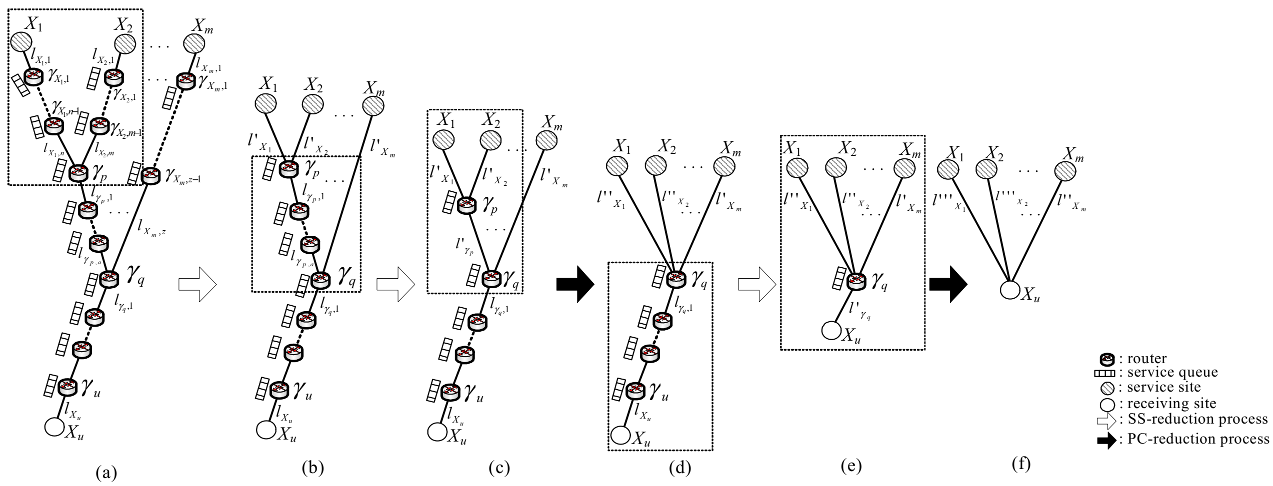


Fig. 4 A delivery tree is reduced to a delivery star by invoking the TSR algorithm

TSR algorithm:
Input: A delivery tree with X_u ($X = \{X_1, X_2, \dots, X_m\}$) as its root (leaf nodes);
Output: The equivalent delivery star;
Procedure:

For each service site X_k, $k = 1, 2, \dots, m${
  If there is an MSS with X_k and $n_{conj,k}$ as the head node and tail node, respectively{ /* $n_{conj,k}$ may be a router or X_u */
    Reduce the MSS by using the SS-reduction process; }} /* by invoking Eqs. (1)~(6) */
While there is a MAX $\{(X_1, X_2, \dots, X_h), \Upsilon_{conj,k}\}$ newly identified where $h \le m$ {
  If there is an MSS connecting $\Upsilon_{conj,k}$ and a node n, and its length is longer than 1{ /* n is a router or X_u*/
    Apply the SS-reduction process to reduce the MSS to a simple chain $\Upsilon_{conj,k} - l'_{\Upsilon_{conj,k}} - n$;}
  Apply the PC-reduction process to reduce the MAX $\{(X_1, X_2, \dots, X_h), \Upsilon_{conj,k}\}$ and $\Upsilon_{conj,k} - l'_{\Upsilon_{conj,k}} - n$ to a PS $\{(X_1, X_2, \dots, X_h), n\}$;
  /* by invoking Eqs. (1)~(7)*/
  Expand the $\{(X_1, X_2, \dots, X_h), n\}$ to a PS $\{(X_1, X_2, \dots, X_h, X_{h+1}, \dots, X_p), n\}$; } /* X_{h+1}, ..., X_p are those service sites currently and directly connected to n by a physical or logical link.*/

Fig. 5 The TSR algorithm

If the PS is an expandable one, the TSR expands the PS $\{(X_1, X_2, \dots, X_h), n\}$ to a MAX $\{(X_1, X_2, \dots, X_h, X_{h+1}, \dots, X_p), n\}$ by reducing each sub-tree rooted at n to an MPS by using the SS-reduction and PC-reduction processes interchangeably where $X_{h+1}, \dots, X_p$ are those service sites that are now directly connected to n by a physical or logical link. This procedure continues until n is X_u and no MPS can be newly identified. Now a delivery star topology (or simply a delivery star), in which all service sites are connected to the center node X_u, is derived. Fig. 4f is the example result. Figs. 4b, 4c, 4d, and 4e are the intermediate results of reducing Fig. 4a to Fig. 4f.

IV. A Demonstration Example

In this section, we would like to evaluate the transmission time of a packet P, 1 KB in length, in a delivery tree and its corresponding delivery star. We use the tree topology illustrated in Fig. 6a as an example, which has six service sites that send packets of a file F to X_u, to calculate P's delivery delay. The bandwidths of all the links are listed in TABLE I, and the service and departure rates of all service sites and routers are accordingly listed in TABLE II. We assume that service site X_k's packet generation rate μ'_{X_k} is 7574.4 pkts/sec, $k = 1 \sim 6$.

TABLE I
The network bandwidths of all the links shown in Fig. 6a

Network links	The bandwidth of a link (Mbps)
$l_{X_1}, l_{X_2}, l_{X_3}, l_{X_4}, l_{X_5}, l_{X_6}$	118.35
l_{Y_1}, l_{Y_2}	236.7
l_{Y_3}	473.4
$l_{Y_4}, l_{Y_5}, l_{Y_6}$	710.1

TABLE II
The service and departure rates of all service sites and routers shown in Fig. 6a

Network links	The bandwidth of a link (Mbps)
$l_{X_1}, l_{X_2}, l_{X_3}, l_{X_4}, l_{X_5}, l_{X_6}$	118.35
l_{Y_1}, l_{Y_2}	236.7
l_{Y_3}	473.4
$l_{Y_4}, l_{Y_5}, l_{Y_6}$	710.1

A. The transmission time in a delivery tree

The routing path of a packet P issued by X_1, denoted by P_1, consists of l_{X_1}, Y_1, l_{Y_1}, Y_3, l_{Y_3}, Y_4, l_{Y_4}, Y_5, l_{Y_5}, Y_6, and l_{Y_6}, and hence the total transmission time of P_1 at t, denoted by $T^{P_1}_{tree}(t)$, is

$$T^{P_1}_{tree}(t) = T_{l_{X_1}}(t) + \mathrm{W}_{Y_1} + T_{l_{Y_1}}(t) + \sum_{i=3}^{6} \left(\mathrm{W}_{Y_i} + T_{l_{Y_i}}(t) \right) \qquad (8)$$

where based on TABLE I, $T_{l_{X_1}}(t) = 66\left(\approx \frac{1\text{KB}}{118.35\text{ Mbps}}\right)$ μs, $T_{l_{Y_1}}(t) = 33\left(\approx \frac{1\text{KB}}{236.7\text{ Mbps}}\right)$ μs, $T_{l_{Y_3}}(t) = 16.5\left(\approx \frac{1\text{KB}}{473.4\text{ Mbps}}\right)$ μs, and $T_{l_{Y_4}}(t) = T_{l_{Y_5}}(t) = T_{l_{Y_6}}(t) = 11\left(\approx \frac{1\text{KB}}{710.1\text{ Mbps}}\right)$ μs. Based on Eq. (2), TABLE II, and Fig. 6a, $\mathrm{W}_{Y_1} = Q(Y_1) = \frac{1}{\mu_{Y_1}-\lambda_{Y_1}} - \frac{1}{\mu_{Y_1}} = \frac{1}{30297.6-7574.4\times 2} - \frac{1}{30297.6} =$ 33μs since $\mu_{Y_1} = 30297.6$ pkts/sec, and Y_1 has two incoming links l_{X_1} and l_{X_2}, implying $\lambda_{Y_1} = A(\{X_1, X_2\}, Y_1) = \sum_{k=1}^{2} D(l_{X_k}, t) = \mu'_{X_1} + \mu'_{X_2} = 15148.8\ (= 7574.4 \times 2)$ pkts/sec. Similarly, $\mathrm{W}_{Y_3} = Q(Y_3) = \frac{1}{60595.2-30297.6} - \frac{1}{60595.2} =$ 16.5μs in which $\mu_{Y_3} = 60595.2$ pkts/sec, $\lambda_{Y_3} = A(\{Y_1, Y_2\}, Y_3) = \sum_{i=1}^{2} D(l_{Y_i}, t) = min(30297.6, 15148.8) \times 2 = 30297.6$ pkts/sec, and $\mathrm{W}_{Y_4} = Q(Y_4) = \frac{1}{90892.8-45446.4} - \frac{1}{90892.8} = 11$μs because $\mu_{Y_4} = 90892.8$ pkts/sec and $\lambda_{Y_4} = A(\{Y_3, X_5, X_6\}, Y_4) = min(60595.2, 30297.6) + min(15148.8, 7574.4) \times 2 = 45446.4$ pkts/sec. Furthermore, $\mathrm{W}_{Y_5} = \frac{1}{\mu_{Y_5}-\lambda_{Y_5}} - \frac{1}{\mu_{Y_5}} = \mathrm{W}_{Y_6} = \frac{1}{\mu_{Y_6}-\lambda_{Y_6}} - \frac{1}{\mu_{Y_6}} = 11$μs since $\mu_{Y_5} = \mu_{Y_6} = 90892.8$ pkts/sec and $\lambda_{Y_5} = \lambda_{Y_6} = A(\{Y_4\}, Y_5) = A(\{Y_5\}, Y_6) = min(90892.8, 45446.4) = 45446.4$ pkts/sec. Based on Eq. (8) and the values calculated above, $T_{tree}^{P}(t) = 231$μs (= $66 + 33 + 33 + 16.5 + 16.5 + 11 + 11 + 11 + 11 + 11 + 11$). The delay of sending a packet P_k by X_k to X_u, denoted by $T_{tree}^{P_1}(t)$, can be calculated by the same method, $k = 2\sim 6$.

B. The transmission time in a delivery star

We first reduce the delivery tree shown in Fig. 6a to a delivery star by employing the TSR algorithm. All network components on the paths from X_1 and X_2 to Y_3, and from X_3 and X_4 to Y_3 form two parallel-chain structures. After applying the PC-reduction process to them, the result is shown in Fig. 6b, in which PS $\{(X_1, X_2, X_3, X_4), Y_3\}$ is an MPS and the path from X_k to Y_3 is reduced to a logical link l'_{X_k}, $k = 1\sim 4$. Based on Eq. (7),

$$T_{l'_{X_k}}(t) = \begin{cases} T_{l_{X_k}}(t) + \mathrm{W}_{Y_1} + T_{l_{Y_1}}(t), & if\ k = 1, 2 \quad (9) \\ T_{l_{X_k}}(t) + \mathrm{W}_{Y_2} + T_{l_{Y_2}}(t), & if\ k = 3, 4 \quad (10) \end{cases}$$

Hence, the bandwidth of l'_{X_k}, i.e., $B_{l'_{X_k}}(t)$, $k = 1\sim 4$, can be calculated by invoking Eq. (6). Now we can further reduce MAX $\{(X_1, X_2, X_3, X_4), Y_3\}$ and its following simple chain $Y_3 - l_{Y_3} - Y_4$, i.e., forming a parallel-chain structure, to a $\{(X_1, X_2, X_3, X_4), Y_4\}$, in which l'_{X_k} followed by Y_3 and l_{Y_3} is reduced to l''_{X_k}, $k = 1\sim 4$. The resulting structure is depicted in Fig. 6c. Based on Eq. (7),

$$T_{l''_{X_k}}(t) = T_{l'_{X_k}}(t) + \mathrm{W}_{Y_3} + T_{l_{Y_3}}(t), \qquad k = 1\sim 4 \qquad (11)$$

Therefore, the bandwidth of l''_{X_k} can be obtained by invoking Eq. (6), and then the TSR expands the PS $\{(X_1, X_2, X_3, X_4), Y_4\}$ to MAX $\{(X_1, X_2, X_3, X_4, X_5, X_6), Y_4\}$, in which X_5 and X_6 are individually connected to Y_4 by a physical link, and which as shown in Fig. 6c is a newly identified MPS, and $Y_4 - l_{Y_4} - Y_5 - l_{Y_5} - Y_6 - l_{Y_6} - X_u$ is an MSS which can be reduced by applying the SS-reduction process. The resulting structure is depicted in Fig. 6d, in which the path from Y_4 to X_u is reduced to a logical link l'_{Y_4}. Based on Eq. (5),

$$T_{l'_{Y_4}}(t) = T_{l_{Y_4}}(t) + \sum_{i=5}^{6}\left(\mathrm{W}_{Y_i} + T_{l_{Y_i}}(t)\right) \qquad (12)$$

where the bandwidths of l'_{Y_4} can be calculated by employing Eq. (6). Now MAX $\{(X_1, X_2, X_3, X_4, X_5, X_6), Y_4\}$ and $Y_4 - l'_{Y_4} - X_u$ is a parallel-chain structure which can be reduced to an MPS, i.e., MAX $\{(X_1, X_2, X_3, X_4, X_5, X_6), X_u\}$, in which l''_{X_k}, Y_4, and l'_{Y_4} become l'''_{X_k}, $k = 1\sim 4$, and l_{X_p}, Y_4, and l'_{Y_4} becomes l'_{X_p}, $p = 5, 6$. Up to this point, a delivery star centered at X_u as depicted in Fig. 6e has been derived. Each service site X_k is directly connected to X_u with either a logical link or physical link. Based on Eq. (7)

$$T_{l'''_{X_k}}(t) = T_{l''_{X_k}}(t) + \mathrm{W}_{Y_4} + T_{l'_{Y_4}}(t), \qquad k = 1\sim 4 \qquad (13)$$

and

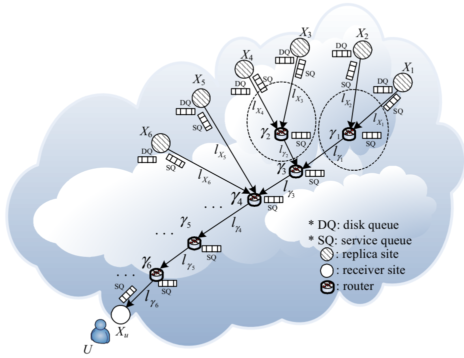
$$T_{l'_{X_p}}(t) = T_{l_{X_p}}(t) + \mathrm{W}_{Y_4} + T_{l'_{Y_4}}(t), \qquad p = 5, 6 \qquad (14)$$

Among the six service sites, the delay time of a packet sent by X_k, $k = 1, 2, 3$, and 4, to X_u is the longest. Hence, just as in calculating the transmission time of a packet in a delivery tree, we calculate the transmission time of a packet P_1 sent by X_1 to X_u as an example. Let $T_{star}^{P_1}(t)$ be the transmission delay of P_1 from X_1 to X_u. Then,

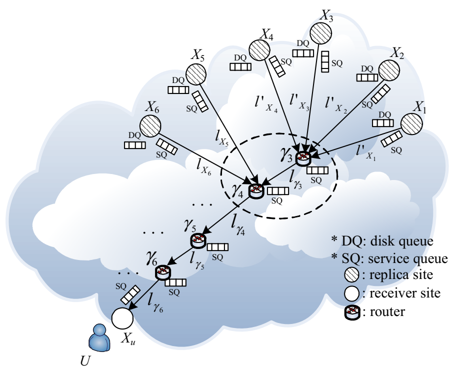
$$T_{star}^{P_1}(t) = T_{l'''_{X_1}}(t) \qquad (15)$$

Based on Eqs. (15), (13), (11), (9), and (12), $T_{star}^{P}(t) = T_{l''_{X_1}}(t) + \mathrm{W}_{Y_4} + T_{l'_{Y_4}}(t) = T_{l_{X_1}}(t) + \mathrm{W}_{Y_1} + T_{l_{Y_1}}(t) + \sum_{i=3}^{6}\left(\mathrm{W}_{Y_i} + T_{l_{Y_i}}(t)\right)$ which is equal to Eq. (8), implying that $T_{l'''_{X_1}}(t) = 231$μs and $B_{l'''_{X_1}}(t) = 35.46$Mbps $\left(= \frac{|P|}{T_{l'''_{X_1}}(t)} \approx \frac{1\text{KB}}{231\text{ μs}}\right)$.

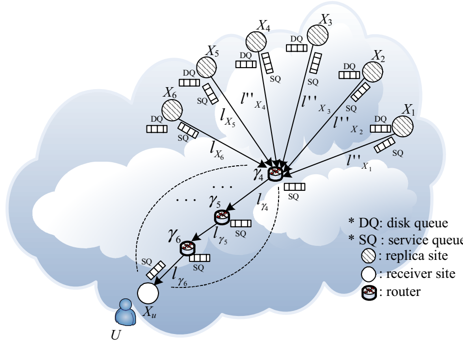
The delay of transmitting a packet from X_i to X_u, $i = 2\sim 6$, can be calculated by the same method. The delays are 231μs, 231μs, 231μs, 132μs, and 132μs, respectively, consequently demonstrating that the transmission delay of delivery star is equivalent to that of delivery tree. Additionally, if the delivery tree can be identified and the bandwidths of l'''_{X_k} and l'_{X_p}, $k = 1\sim 4$, $p = 5, 6$, are calculated beforehand, the calculation of $T_{star}^{P_1}(t)$, i.e., Eq. (15), is simpler than the calculation of $T_{tree}^{P_1}(t)$, i.e., Eq. (8). Therefore, it is easier for users to evaluate the data transmission delays of a co-allocation scheme.



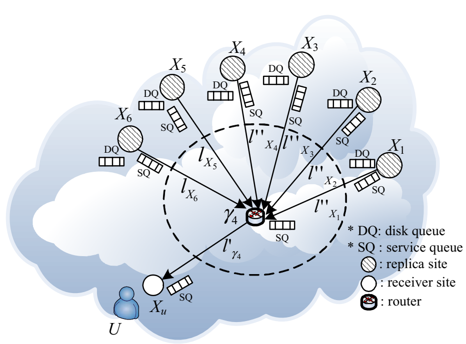
(a) A delivery tree data grid



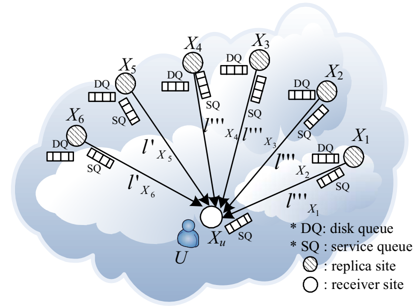
(b) A delivery tree reduced from the one shown in Fig. 6a



(c) A delivery tree reduced from the one shown in Fig. 6b



(d) A delivery tree reduced from the one shown in Fig. 6c



(e) A delivery tree reduced from the one shown in Fig. 6d

Fig. 6 An example of reducing a delivery tree data grid to a delivery star

V. Conclusions

In this article, we propose a topology reduction approach, called the TSR, to reduce a complicated delivery tree to a star in a data grid without impacting the evaluation accuracy and correctness of data transmission delays for a co-allocation scheme. The delivery star simplifies the evaluation and analytical complexities of the network delay time of a co-allocation scheme. Nevertheless, the TSR is also applicable to reducing the complexity of a delivery tree in other distribution and parallel systems, e.g., a grid computing, cloud computing, and P2P system.

ACKNOWLEDGMENT: The work was partially supported by TungHai University under the project GREENs and the National Science Council, Taiwan under Grants NSC 101-2221-E-029-003-MY3, NSC 100-2221-E-029-018 and NSC 101-2628-E-009-024-MY3.

References

[1] S. Vazhkudai, "Enabling the co-allocation of grid data transfers," *Proceedings of International Workshop on Grid Computing*, pp. 44–51, Phoenix, Arizona, USA, 17 November 2003.

[2] R.S. Chang, C.M. Wang, and P.H. Chen, "Replica selection on co-allocation data grids," *Proceedings of Second International Symposium on Parallel and Distributed Processing and Applications*, vol. 3358, pp. 584–593. Springer, Heidelberg, 2004.

[3] R.S. Bhuvaneswaran, Y. Katayama, and N. Takahashi, "Dynamic co-allocation scheme for parallel data transfer in grid environment," *Proceedings of First International Conference on Semantics, Knowledge, and Grid*, Beijing, China, vol. 17, pp. 28–29, Nov. 2005.

[4] C.T. Yang, I.H. Yang, S.Y. Wang, K.C. Li, and C.H. Hsu, "A recursively-adjusting co-allocation scheme with a cyber-transformer in data grids," *Future Generation Computer Systems*, vol. 25(7). Elsevier B.V: pp. 695–703, Amsterdam, July 2009.

[5] M.C. Lee, F.Y. Leu, and Y.-p. Chen, "PFRF: An Adaptive Data Replication Algorithm based on Star-Topology Data Grids," *Future Generation Computer systems*, vol. 28, issue 7, pp. 1045–1057, July 2012.

[6] I. Foster, C. Kesselman, and S. Tuecke, "The anatomy of the grid: enabling scalable virtual organizations," *International Journal of High Performance Computing Applications*, vol. 15, pp. 200–222, 2001.

[7] M. Coates, A. Hero, R. Nowak, and B. Yu, "Internet tomography," *IEEE Signal Processing Magazine*, vol. 19, issue 3, Jan. 2002.

[8] G. Levitin, Y.S. Dai, and B.H. Hanoch, "Reliability and performance of star topology grid service with precedence constraints on subtask execution," *IEEE Transactions on Reliability*, vol. 55, no. 3, pp. 507–515, Sept. 2006.

[9] L. Kleinrock, "Queuing Systems. In: Theory vol. 1," Wiley, New York (1975).