Scalable Communication Infrastructure through Resource Provisioning Approach for Reliable Social Networking System

Masnida Hussin, Asma Zubaida Mohamed Ibrahim, Azizol Abdullah

Abstract: ***Nowadays social networking system serves as an important information technology (IT) infrastructure of growth for community. Demands for scalable resource sharing in the system have become extremely important due to strong requirements imposed by the dynamic behaviour of users and resource communities. With the increasing number of entities in the system, inadequate information and unsuccessful accessibility of resources have becoming critical factors that degraded system performance. We present a hybrid resource provision scheme for improving resource communication in terms of availability and reliability while aiming for cost-effective processing. Specifically, our resource provision scheme incorporates an automated analytic for achieving accurate resource information while minimizing processing overheads and inter-communication latency. The provision scheme then incorporates trust-based scheduling policy to deal with diverse processing requirements and heterogeneous resources. Simulation experiments proved the efficacy of our scheme in achieving better trade-off between system scalability and performance; and helps sustain cost-effective computing.***

Keywords: ***Resource management, Communication infrastructure, Social computing, Social networking system.***

I. INTRODUCTION

Generally, continuous growth of community computing in social networking system has led to complex systems of behavior from handling large users' communications to resource capacities [1-5]. The resources in regards processing and communication capacities forms the core of support infrastructure for performance optimization in such system. The computing paradigm of the social networking system poses new challenges to the scalability problem, because they have to process the large volumes of datasets and might have loosely coupled tasks. The wide range of resource selection and high degree of heterogeneity leads to problem of selection. In response to scalable computing, processing and communication services are necessary to be operated and maintained in efficient and cost-effective management. In the social networking systems, every users' applications are completely different, dynamic and independent.

Revised Manuscript Received on May 22, 2019.

Masnida Hussin, Department of Communication Technology and Network, Faculty of Computer Science and Information Technology, University of Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

Asma Zubaida Mohamed Ibrahim, Department of Communication Technology and Network, Faculty of Computer Science and Information Technology, University of Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

Azizol Abdullah, Department of Communication Technology and Network, Faculty of Computer Science and Information Technology, University of Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

They are generally allocated to the high-performance computing resources. However, resources' information in regards their capabilities are hard to determine prior in dynamic computing environment. Meanwhile, the information service in network management systems provides only the most elementary information to guide on resource selection process [6, 7]. In Cloud, for example, not all information is available for the resource manager/scheduler to assist in the selection process for resources [8]. In many cases, if not all, the information returned by the service is costly to obtain, inaccurate or outdated. Hence, it is necessary to develop resource provision scheme of accurately evaluating resource capabilities (i.e., capacity and availability). An effective provision scheme also needs for reducing expense of processing and communication during intersection of social behavior and computational systems. Further, it significantly increases successful task execution.

This motivated us to design a scalable communication infrastructure that able to adapt heterogeneous and dynamic social computing environment while integrating the notion of cost-effective and trust processing. We define scalable as the capability of resources for maintaining to function well in social computing and adapting to processing demands.

Specifically, our communication infrastructure utilizes the resource provision scheme for reconciling the dynamic mapping rules with trust-based scheduling. With the diverse processing requirements from the systems' users, we employ resource trust factor for the available resources. The trust factor aims to ensure the continuous processing reputation of inter-communication among resources in the social network management. In this work, we also focus on interoperability among resource sites where the communication happens through the network links that are able to provide effective transmission within the social entities (i.e., end-users, public Cloud, organization etc.). Due to unpredictable changes in resource availability, our communication infrastructure incorporates automated analytics for determining the accurate resources' information. Such analytic required the resource manager/scheduler to continuously monitor the valuation of resources' behaviors in the social computing. The analytic approach intents for minimizing computing cost in terms of overhead and latency. This work evaluates in our extensive simulations (i.e., comparison with other provision schemes) through varying values in processing and communication capacities, and a diverse set of data communication.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The results obtained from our comparative evaluation study clearly show that our communication infrastructure improves communication time while experiencing low in processing cost. The remainder of this paper is organized as follows. A review of related work is presented in Section 2. In Section 3, we described the models used in the paper. Section 4 details our resource allocation approach. Experimental settings and the experimental results are presented in Section 5. Finally, conclusions are made in Section 6.

II. RELATED WORK

Despite of decades of research advances, social networking system keeps posing challenging research questions [4, 9-11]. It is due to ever-increasing data communication in variety and scale, and increasing diversity of resources and network domains. There is new finding in [12] where to consider the use of variety resource management infrastructure. Such ideas raises due to it will be anticipated to impact areas, such as connecting people and devices, data-intensive computing, the service space and self-learning systems. Due to uncertainties in the resource availability and cost of using resources (e.g., in Clouds), resource management mechanisms such as resource provisioning, rescheduling, migration have been extensively studied in the literature (e.g, [13-15]). There has been increasing interest in addressing dynamic resource provisioning decisions for performance optimization.

Due to the resources are geographically located over many locations in the network, resource provision scheme should aim for scheduling resources that closer to users to increase availability while reducing the costs. The existing resource management mechanisms are either has considered computational power and communication link as one perspective, or only considered one of those. The game theoretical framework is proposed in [15] focus on distributed resource management and formulates the bandwidth allocation problem. Their cloud-based social networking is designed to efficiently deal with quality of service provision in multimedia sharing and distribution. The scope of the model is for the mobile users that watch multimedia live-streaming where their cheat-proof mechanism aims to motivate mobile users to share bandwidth. Meanwhile, the authors in [9] developed the resource allocation framework for Device-to-device (D2D) communication that called as the social-community-aware D2D. They proposed the two-step coalition game for allowing the cellular users to share their channels with D2D communications within the same community group of people. The study is considered both processing and bandwidth capacities. The factors of processing and communication in resource management are significant for the system performance. It is been claimed by the authors in [16] that efficient algorithms for resource allocation is reliant on processing latency and communication costs by minimizing these metrics. Hence, they developed data-center selection algorithms for virtual machines (VM) placement that minimize the maximum distance between the selected data centers in order to get the maximum benefit from the distributed cloud system. Meanwhile, with the heuristic algorithms, their allocation scheme performs well for assigning VMs to processing resources in the chosen data center locations.

Note that the success of the overall resource management mechanisms lies with the scalability, efficiency, manageability, reliability of the systems in providing and managing the resources (i.e., workers, computation time, memory, energy). It is vital for the social system to manage the resources efficiently and ensure the computing and communication demands are able to access the appropriate information; regardless of the place and time. It is because failing to manage the resources would not only impairing the system functionality but also affecting the overall cost and performance. The automated resource management system for Cloud computing environment that proposed in [17] addressed the issue of data centers operation where always underutilized due to over provisioning for the peak demand. The work achieves good balance between the overload avoidance and green computing. The authors in [3] propose a failure-aware resource allocation strategy by calculating resources' reliability states. They defined the capacity-reliability metric and used best-fit algorithm to find the best qualified resources on which to instantiate virtual machines (VMs) to compute users' requests. The authors in [18] proposed dynamic resource provision that considered the user-specified level of accuracy in order to transmit the communication from one resource to another. Their work focuses on the traffic characterization for determining the right set of resources assigned to each communication at each switch. While considering the high accuracy for data communication, the provisioning approach also taking network-wide resource constraints as well as traffic and data dynamics into account. In the wide-networks such as social networking system, if there is multiple communication relays exist, the selection process, wherein scheduler chooses the "best" relay, is expected for minimizing overhead [18, 19]. The authors in [4] proposed the assessment mechanism for social resource allocation that considered welfare, allocation fairness, and algorithmic runtime. They utilized the social network model of virtualized containers from the users' personal computers or smart devices. The assessment mechanism investigated on how the social networks able to leverage in the construction of Cloud computing infrastructures and how resources are allocated in the presence of user preferences. Our communication infrastructure explicitly takes into account both computational power and link communication capacity based on the user preferences (i.e., priority) in selection and mapping strategies to efficiently meet fluctuating demands.

Note that the success of the overall resource management mechanisms lies with the scalability, efficiency, manageability, reliability of the systems in providing and managing the resources (i.e., workers, computation time, memory, energy). It is vital for the social system to manage the resources efficiently and ensure the computing and communication demands are able to access the appropriate information; regardless of the place and time.

It is because failing to manage the resources would not only impairing the system functionality but also affecting the overall cost and performance. The automated resource management system for Cloud computing environment that proposed in [17] addressed the issue of data centers operation where always underutilized due to over provisioning for the peak demand. The work achieves good balance between the overload avoidance and green computing. The authors in [3] propose a failure-aware resource allocation strategy by calculating resources' reliability states. They defined the capacity-reliability metric and used best-fit algorithm to find the best qualified resources on which to instantiate virtual machines (VMs) to compute users' requests. The authors in [18] proposed dynamic resource provision that considered the user-specified level of accuracy in order to transmit the communication from one resource to another. Their work focuses on the traffic characterization for determining the right set of resources assigned to each communication at each switch. While considering the high accuracy for data communication, the provisioning approach also taking network-wide resource constraints as well as traffic and data dynamics into account. In the wide-networks such as social networking system, if there is multiple communication relays exist, the selection process, wherein scheduler chooses the "best" relay, is expected for minimizing overhead [18, 19]. The authors in [4] proposed the assessment mechanism for social resource allocation that considered welfare, allocation fairness, and algorithmic runtime. They utilized the social network model of virtualized containers from the users' personal computers or smart devices. The assessment mechanism investigated on how the social networks able to leverage in the construction of Cloud computing infrastructures and how resources are allocated in the presence of user preferences. Our communication infrastructure explicitly takes into account both computational power and link communication capacity based on the user preferences (i.e., priority) in selection and mapping strategies to efficiently meet fluctuating demands.

$$CT(i,,j) = (wait\_t + exe\_t) \quad (1)$$

where wait_t is the elapsed time between a submission, and the start of execution and exe_t is actual execution time of task, respectively. Workloads arrive in a Poisson process. We assume that the workload's profile is available and can be provided by the user using job profiling, analytical models or historical information.

Social Networking Model

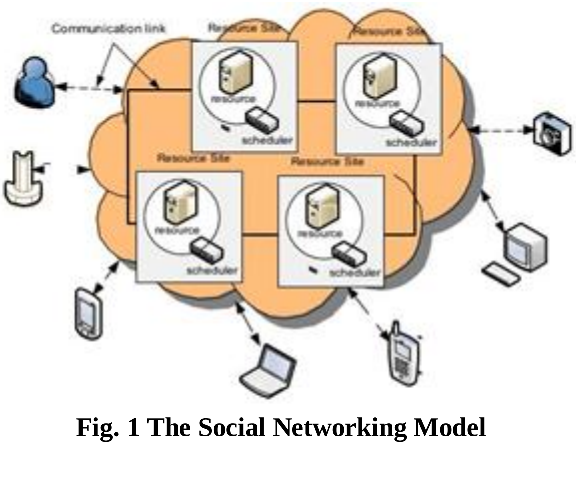

Fig. 1 The Social Networking Model

The target system used in this work (Fig. 1) consists of several distributed resource sites or nodes given as Sx where x = [1, 2…, s] which are loosely-coupled sites connected by communication links. In each site a scheduler resides which it handles workloads from system users that comes from variety network devices where it can be considered as data centers, mobile devices, wireless access point (WAP) etc. that is illustrated in Fig. 1. The scheduler then maps the workloads/demands onto compute resource. Each scheduler is given the authority to keep track of its resources' details; where different scheduler deals with its workloads in parallel. Each resource j is associated with processing capacity to complete the scheduled workloads. More formally, the processing capacity of resource j is defined as:

$$PCj = \sum speed + (L/\sum exe\_t) \quad (2)$$

where speed is relative speed of resources and L is the total number of workloads completed within some observation period, respectively. For a given resource, its processing and availability fluctuated. Therefore, the actual completion time of a workload on a particular resource is difficult, if not impossible, to determine a prior. The number of resources in this study is assumed to be relatively less than the number of users' demands to be processed.

The resource scheduler may communicate with each other to share and exchange resources' information, in the sense that there is a communication link between them. The link represents connection between sites and it might be an actual link of cable or a virtual link of the Internet. There are two type of connections considered in this study; direct connection and indirect connection. If the direct connection is not operational or no direct connection exists between the two sites, the schedulers can communicate through other intermediate peers bandwidth between any two individual sites varies corresponding to realistic network; thus, the inter-site communications are heterogeneous. The inter-site communications are assumed to dispute with some delays.

The performance of communication link also affects resource performance to a certain degree due to heterogeneous links. It counts all gains and losses in transferring the workloads between the schedulers. For a given workload, start time of processing might be affected by the fluctuation in the link capacity; more specifically, delays in communication time and the average rate of successful communication. Hence, data transfer rate between resource schedulers (i.e., scheduler a and b) through the communication link l can be expressed as in Eq. 3.

$$lc_l = (1 - dy_l) + CD\,(scheduler\ a, b) \quad (3)$$

where dyl is average of transaction delay in communication link l and CD(scheduler a, b) can be derived from the size of workload divided by the mean speed of the current link between the scheduler a and b. We believe that minimization of the processing cost in terms of delay and latency robustly contributes for enhancing system scalability. By referring Eq. (1) and (3), we define the processing cost for workload of user i as;

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

$$cost_i = lc_i + \frac{}{CT(i,j)} \qquad (4)$$

In this study, cost component such as local disk I/O and cooling systems are neglected, given that the cost-effective decisions focus on timing factor.

III. HYBRID PROVISION SCHEME

This section begins by describing our automated analytics strategy for identifying the resource information that based on availability and capacity. It is an effective means for effectively handling fluctuating in resource performance. Later, we describe the trust-based scheduling that take into account processing and communication link availabilities in order to meet the ever- growing processing requirements.

Automated Analytics for Resource Discovery

Note that the network communication infrastructure takes into account both computation and communication factors to determine resource capability. In particular, the system performance could be lower due to involvement of unpredictable changes in processing capacity hence, increases execution time during peak workload (or heavily loaded). That makes the communication links becomes more intensive due to heavy traffic that implicitly increases communication latency. In response to the issue, it is significant to identify accurate information of resources in the communication infrastructure. Our automated analytic for resource discovery strategy concerns two procedures, (i) local and (ii) global explanatories. These two explanatories are different in terms of information scale that is collected by the resource scheduler. For the local explanatory, only the computing availability is taken into account while the global explanatory tracking both computing power and communication link availabilities. The local explanatory determines resource availability by evaluating the total processing capacity in the site. The resource is considered available if its average waiting time smaller than average of waiting time in the site. It aims to avoid lengthy waiting time in incapable resources (i.e., *average wait_t@j > average wait_t site*).

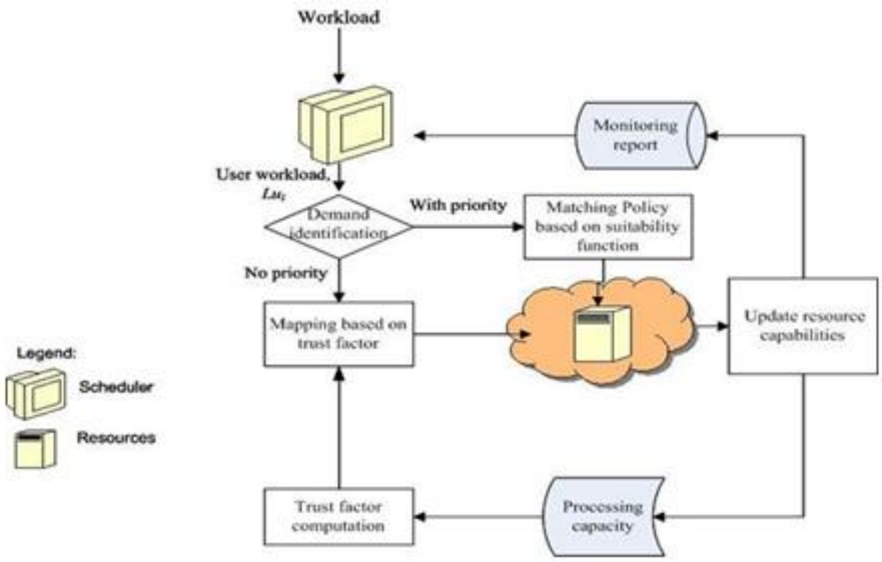

Fig. 2 Overview of the Trust-based Scheduling

The global explanatory involved the inter-communication between sites. This resource discovery is needed and applied when there are workloads need to be rescheduled or migrate between sites. Note that the sites connected through heterogeneous communication links; hence the availability of communication link comes into a picture. In initial phase, the resource scheduler traced all the available links that are connected to them. Basically, we have two communicating sites that refer as sender site and receiver site where the sender site might connect to more than one receiver sites. The link availability is defined as the probability that there is link exists between sites either direct connection or indirect connection at a particular given time (describe in Section 3.2). In the second phase of the global explanatory, the resource scheduler calculated and compared all data transfer rate of the available links. It only selects the link with the smallest data transfer rate to transfer the workload to the available compute resources. When the workload reaches to the receiver site, it then scheduled into the site's queue as a new arrival workload.

Trust-based Scheduling Formation

We focus on large-scale resource sharing that resides in the social networking system. Hence, the information of both processing power and communication link has significant on performance optimization. Basically, the system is able to improve execution time if the resources are available and able to meet certain level of demands. However, the available resources might not always be suitable for the users' workloads due to resource heterogeneity. In response to this, we formulate suitability function between resource and workload for looking resource alternatives (Fig. 2).

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Due to the workloads come with different priorities, there is also increasing interest from the users to schedule their workloads into the most dependable resource for reliable performance. In some cases, the workload might be required additional processing requirement; such as the higher priority workload is essentially to be executed first. Hence, it is significant to effectively match the processing capacity with users' demands and aim for better completion time. The matching policy for both resource and workload is determined based on a suitability/fitness function between compute resource j, data rate of communication link l and workload weight wi as defined to be:

$$fit(j,i)l = \frac{PC_j + w_i}{lc_j} \quad (5)$$

where PCj and lc are taken from the Eqs. (2) and (3). The workload is assigned to the resource that gives the highest suitability/fitness value fit(j,i)l where it is connected through communication link with low data transfer rate. Note that, it allows the workloads to meet the processing requirement with low communication cost. In some cases, for instance, if there are calculated with same fitness value for a given workload, it then assigned to the resource that gives the highest processing capacity. The workloads without any additional processing requirement are assigned into the available resources based on the preferences of resource trust factor. The resource trust factor aims to express trustworthiness of the underlying resources based on their capacities (i.e., processing and communication). By utilizing the Eq. (3), the trust factor of resource j and link i, it is formally defined as follows:

1, if average wait_t@j< average wait_t site
A(j,l) = aj + lcl where aj = :
0 otherwise (6)

As the resource availability is also affected by the waiting time of the workload in the scheduler, we further introduced an adaptive ranking that based on workload priority. Initially, the workloads at resource site are mapped according to their priority value pri. However, the resource scheduler is permitted to modify the priority of the workload from lower to higher value when the waiting time of the workload is 50% higher than the average waiting time in the queue (i.e., 0.5 wait_ti > average wait_t@q). The scheduler regulated a new priority of a given workload by increased 5% of initial priority value. Note that, the value satisfies with our workload priority ranges that being set in experiments. The workload rank is frequently updated and it repeats until no further processing requirements to be executed.

IV. PERFORMANCE EVALUATION

In this section, we first describe the experiment configuration. Then, experimental results are presented. We study the performance of our network communication infrastructure; name Adaptive Social Networking Communication (AdaSoN) by using the resource provision scheme for system scalability that is compared with three other resource provision algorithms, which are Suit-fit, Max-max and Rand. In the Suit-fit, the suitable resource for a particular workload determined merely according to the highest suitability value fit(j,i). Max-max heuristic algorithm maps workload to the resource, which gives the highest suitability value fit(j,i) from its maximum list.

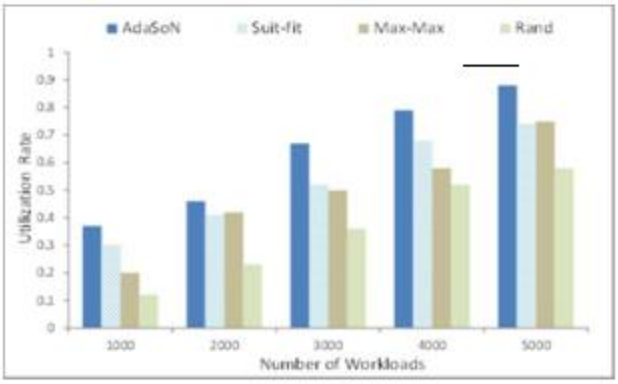

Fig. 3 Resource Utilization with Different Provision Approaches

It allows the workload that assigned into the most suitable resource to be executed first in order to achieve better execution time. In Rand, workloads are randomly mapped to available resources. Performance metrics used for the experiment are resource utilization rate, processing cost and success rate. The utilization rate which is defined as the percentage of time the resource is busy servicing users' workloads, given by:

$$RU = \frac{L}{\sum exe\_t} \quad (7)$$

where L denotes is the total number of workloads completed within some observation period. In addition to utilization rate RU, the processing cost denotes the cost to fulfil users' processing requirements, as in Eq. 8.

$$\beta = (\text{average } cost_i) * \sum exe\_t \quad (8)$$

where the cost value is given by calculating Eq. (3).

The resource with a higher trust factor is more likely to improve user satisfaction in terms of minimizing execution time and processing cost compared than the resource with a low contribution. The scheduler of each site constantly checks the resource trust factor to indicate the resource that is sufficiently qualified to process the workloads.

To better understand on how trust-based scheduling helps in providing reliable communication, we invent the success rate denotes the number of priority workloads that are completed within users' satisfaction obtained. It is generated by dividing the total number of high priority workload that successfully completed over the total number of workloads that been executed, specifically by using Eq. (9).

$$\delta = \frac{\sum pri\_w}{\sum L} \quad (9)$$

The users' satisfaction is considered that happen when there are no complained or resubmission of the workload in the social computing. These metrics are used to measure the degree of system performance and to identify how well our communication infrastructure deals with heterogeneous processing and communication requirements.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

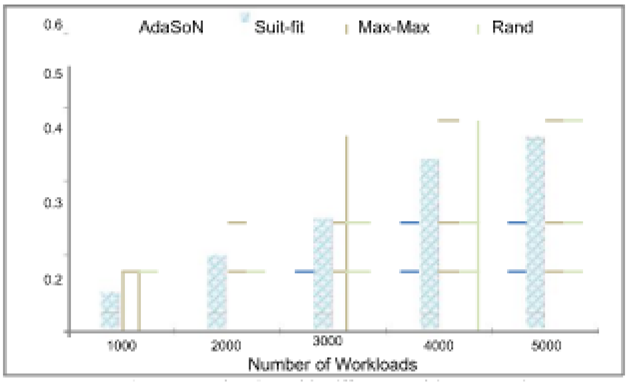

Fig. 4 Processing Cost with Different Provision Approaches

Experimental Settings

We evaluated our resource provision scheme via simulation with number of resource sites ranging from five to twenty in each of which its own scheduler resides. Each resource site contains a varying number of core ranging from 4 to 8 units. The relative processing power (speed) in each core is selected within the range of 1 and 7.5. The number of workloads in a particular simulation is set between 1000 for indicating non-peak communication and 5000, for the peak communications happen in the simulation system. Inter-arrival times (iat) of the workload follow a Poisson distribution with a mean of five time units. Note that, iat satisfies with the mapping technique without explicitly increasing significant time delay in the system. For the given workload, the communication weight is randomly generated from a uniform distribution ranging from 103 to 106. The workload priority values are between 0 and 1, where 0 indicates the lowest priority and 1 indicates the highest priority. The speed of a communication link is uniformly distributed within the range of 10 and 100. We vary the percentage of the transaction delay in communication link from 10% to 50%. Note that the smallest delay has a success workload transaction close to 100%.

V. RESULTS

Experimental results are presented in two different ways based on system performance and communication effectiveness.

Experiment 1: The impact of resource provision scheme on system performance

As shown in Fig. 3, the proposed social networking communication outperformed others in terms of utilization rate. The superior performance of our communication strategy is primarily achieved by scheduling approach that take into account both processing and link availabilities. Specifically, the incorporation of two different mapping strategies (i.e., suitability value and resource trust factor) into our provision scheme leads to better completion time of users' workloads. It also observes that Suit-fit is comparable with our approach. It is due to the fact that both strategies considered suitability value during their matching processes for optimal scheduling. However, it indicates that AdaSoN works 30% better in the case of more workloads coming or being processed.

Fig. 4 shows the average processing cost that is plotted against number of workloads, respectively. Our communication strategy obtains appealing processing cost by using adaptive provision scheme compared to three other schemes, about 60% on average. Although all resource provision schemes in the simulation considered communication links during the scheduling process, our AdaSoN benefits for handling various workload priorities and heterogeneous resources by introducing the analytic resource discovery.

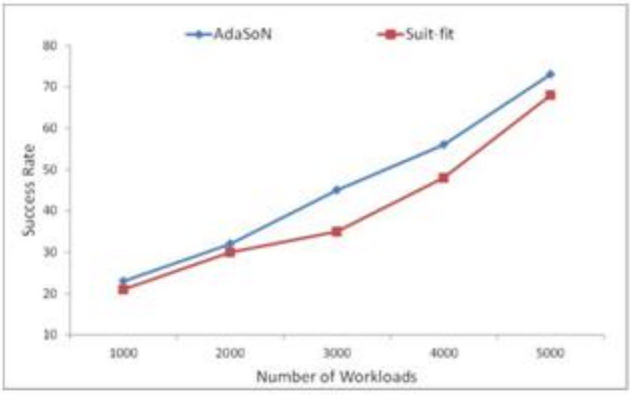

Fig. 5 Success Rate with Different Provision Approaches

Fig. 5 shows the results of successful workload execution between AdaSoN and Suit-fit. Although the experiments are conducted with other resource provision schemes (i.e., Max-max and Rand), they were not explicitly stated. It is due to such algorithms does not significantly differentiate results in terms of success rate. In Max-max and Rand algorithms, most of the workloads met their priorities (i.e., good in success rate), however they are noticeably increasing the processing cost. Therefore, in Fig. 5 we merely plotted the results with reliable in both success rate and processing cost. The figure clearly signifies the superior performance of our approach over Suit-fit, regardless of higher successful workload. It indicates that our scalable communication strategy with direct and indirect connections is able to capture better execution time without excessive processing cost.

Experiment 2: The impact of communication strategy on system variability

In response to appealing result in the Experiment 1, we enhance the experiment of AdaSoN in order to measure on how the variation in resource availability affects the computing and communication cost (denoted as processing cost). In this experiment, we vary the heterogeneity of resources availability according to four different states of availabilities between communication link and processing power, as given in Table 1. The state of resource availabilities is derived based on normalization values from the first experiment.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Table. 1 Different States in Resource Availabilities

State of availability	Processing node	Communication link
1st state	High	High
2nd state	High	Low
3rd state	Low	High
4th state	Low	Low

Fig. 6 shows efficient results in terms of processing cost in the first three states of resource availability in AdaSoN (less than 50% on average). This is can be explained by the high availability of resources either in processing powers or communication links is much preferable helps the network systems to continuously exhibit improvement in cost-effective processing.

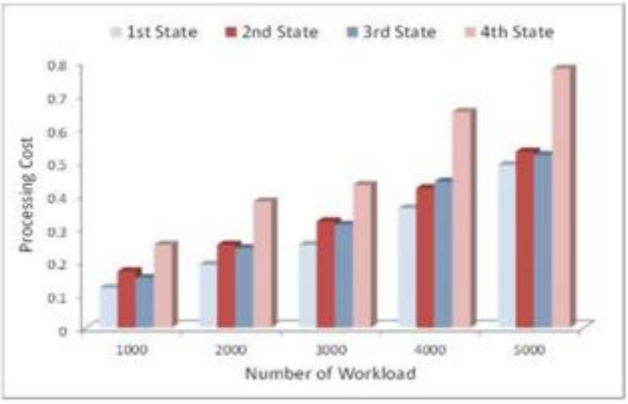

Fig. 6 Processing Cost with Different State of Resource Availabilities

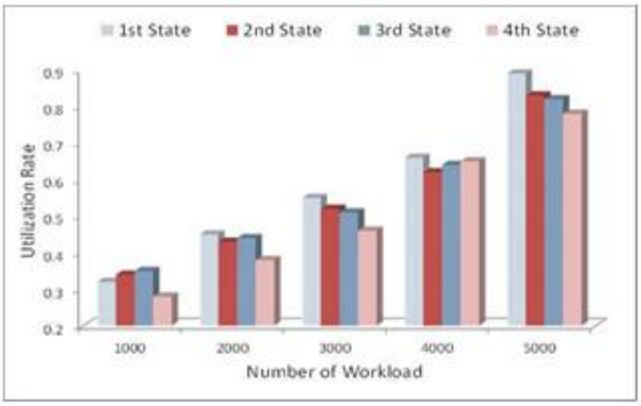

Fig. 7 Resource Utilization with Different State of Resource Availabilities

In Fig. 7, we observe that once again, the AdaSoN performs extremely well by achieves resource utilization, up to 80%, in all available states. The results provide further evidence that, in fact, our social communication strategy capable of adapting to transient workload condition without impacting system performance.

VI. CONCLUSION

Advances in the state of art in social networking system offer an ideal environment for advancing management procedure of resources including the resource communication. The resource provision scheme that take into account processing and communication complexities need to be designed on top of such large-scale systems for conserving reliable communication. In this work, we addressed the scalability issue for social networking system in the context of inter-communication in dynamic and heterogeneous computing. We have effectively modelled communication infrastructure by utilizing the automated analytic for encompassing resource information that helps in identifying accurate resource state. In addition, the scheduling approach with trust factor is introduced to benefit in handling diversity in users' processing requirements, with minimal possible performance degradation. Our communication strategy for social computing also able to yield interesting insights into the cost-effective processing in regards minimizing overhead and latency. Based on our extensive simulation results, the communication strategy demonstrates robust decisions in the sense that system performance remains at a good level regardless of different workload patterns. We highlight that the scalable communication infrastructure contributes for improving system reliability, which will result in a proliferation in the performance of social networking system.

REFERENCES

1. H. Chen and M. Attia, "Fair-sharing Resource Management Scheme for Multi-Service Wireless and Mobile Networks," IEEE, pp. 700-704, 2008.
2. I. Llorente, R. Moreno-Vozmediano, and R. Montero, "Cloud Computing for On-Demand Grid Resource Provisioning," in Advances in Parallel Computing vol. 18, ed: IOS Press, 2009, pp. 177-191.
3. S. Fu, "Failure-aware resource management for high-availability computing clusters with distributed virtual machines," Journal Parallel and Distributed Computing, vol. 70, pp. 384-393, 2010.
4. S. Caton, C. Haas, K. Chard, K. Bubendorfer, and O. F. Rana, "A Social Compute Cloud: Allocating and Sharing Infrastructure Resources via Social Networks " IEEE Transactions on Services Computing vol. 7, pp. 359 - 372, 2014.
5. T. Paul, A. Famulari, and T. Strufe, "A survey on decentralized online social networks," Computer Networks, vol. 75, pp. 437-452, 2014.
6. J. M. Ferris, "Adjusting resource usage for cloud-based networks," ed: Google Patents, 2017.
7. C. Wu and R. Buyya, Cloud Data Centers and Cost Modeling: A complete guide to planning, designing and building a cloud data center,: Morgan Kaufmann, 2015.
8. S. F. Jalal and M. Hussin, "Multi-Level Priority-based Scheduling Model in Heterogenous Cloud," Journal of Computer Science, vol. 10, p. 2628, 2014.
9. F. Wang, Y. Li, Z. Wang, and Z. Yang, "Social-Community-Aware Resource Allocation for D2D Communications Underlaying Cellular Networks," IEEE Transactions on Vehicular Technology, vol. 65, pp. 3628 - 3640 2016.
10. D. Serrano, S. Bouchenak, Y. Kouki, F. A. de Oliveira Jr, T. Ledoux, J. Lejeune, J. Sopena, L. Arantes, and P. Sens, "SLA guarantees for cloud services," Future Generation Computer Systems, vol. 54, pp. 233-246, 2016
11. S. Son, G. Jung, and S. C. Jun, "An SLA-based cloud computing that facilitates resource allocation in the distributed data centers of a cloud provider," Journal Supercomputing, vol. 64, pp. 606-637, 2013.
12. B. Varghese and R. Buyya, "Next generation cloud computing: New trends and research directions," Future Generation Computer Systems, vol. 79, pp. 849-861, 2018.
13. O. Adam, Y. C. Lee, and A. Y. Zomaya, "Stochastic resource provisioning for containerized multi-tier web services in clouds," IEEE Transactions on Parallel and Distributed Systems, vol. 28, pp. 2060-2073, 2017.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

14. M.-A. Vasile, F. Popa, R.-I. Tutueanua, V. Cristea, and J. Kołodziej, "Resource-aware hybrid scheduling algorithm in heterogeneous distributed computing," Future Generation Computer Systems, vol. 51 pp. 61-71, 2015.
15. G. Nan, Z. Mao, M. Li, Y. Zhang, S. Gjessing, H. Wang, and M. Guizani, "Distributed resource allocation in cloud-based wireless multimedia social networks " IEEE Network, vol. 28, pp. 74 - 80, 2014
16. M. Alicherry and T. V. Lakshman, "Network Aware Resource Allocation in Distributed Clouds," in IEEE INFOCOM, Orlando, FL, 2012, pp. 963-971.
17. Z. Xiao, W. Song, and Q. Chen, "Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment," IEEE Transaction on Parallel and Distributed Systems, vol. 24, pp. 1107-1117, 2013.
18. M. Moshref, M. Yu, R. Govindan, and A. Vahdat, "DREAM: Dynamic Resource Allocation for Software-defined Measurement," presented at the SIGCOMM, Illinois USA, 2014.
19. M. S. Alam, J. W. Mark, and X. S. Shen, "Relay Selection and Resource Allocation for Multi-User Cooperative OFDMA Networks," IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, vol. 12, pp. 2193-2205, 2013.