Dynamic Application of Slicing Technique in Grid Calling and Network Resources

Lixia Wang[1] , Aimin Yang[2] , Yaohui Li[1] , Liting Gao[1] ， Tieying Liu[3]
[1] Department of Computer Science
Hebei Institute of Architecture and Civil Engineering
No.33, Jianguo Road Zhang Jiakou City, Hebei Province
CHINA
http://www.hebiace.edu.cn
[2]Computer Science and Information Technology College
Zhejiang Wanli University
[2]No.8, South Qian Hu Road Ningbo, Zhejiang Province
CHINA
http://www.computer.zwu.edu.cn
Hebei CAIDA Securities Co Ltd
[3] Makers of the financial building-23, Self-Shijiazhuang City Road 35
CHINA

Abstract: The grid links all resources in different structures widespread geographically into one logical system via internet, characterized with complexity of network structure, dynamism of network resources and autonomy of each LAN. The high-efficient management of grid resources and high-efficient scheduling of tasks are the important contents in today's net grid research. Slices technology is the present quite popular popular technology, the application domain is widespread. The grid technology and the slice technology fuse in together, may create the more reasonable task scheduling and the scheduling of resources. In this article we divide the tasks into some small tasks assigned with scheduling priorities for each level, and then slice all resources in network into some time slices assigned with resource priorities according to trust mechanism. It promotes the effectiveness and high-efficiency by arranging task scheduling and resource scheduling in priority.

Key-Words: grid computing; Slicing Technique; Slice Min—min algorithm; Trust mechanism; Load Balancing; Distributional dispatch; Toughness; Correspondence detention

1 Introduction

The grid is one kind of integrated computing and resource sharing in network environment. [1]Web resources include computing resources, storage resources and bandwidth resources, resources, software, data resources, information resources, knowledge and resources.

Like one set of super computer the grid provides the user with integrated information and application service (such as compute store access and so on），makes up for computing power limitedness、resource limitedness and resource idling of independent computer , and realizes resource sharing in wide sense. The grid provides standards and supports between computer and network and allows the individual、corporation and organization on network to mutually make the best of shared resource，and also provide private source device to other remote users.[17] The difference in the

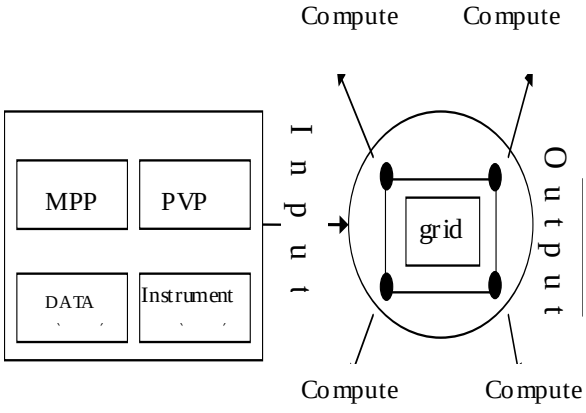

Fig.1 Grid constitution schematic drawing

structures of various resource、management mechanism、user and application program and dynamic change of network resources make resource management and task scheduling become one of the key technologies that affects the effectiveness and high-efficiency of grid computing.[13]

2 Grid Technology

The grid technology already obtained the success in the scientific research and educational fields' application, this point obtained the widespread approval at present. Obtains the success the primary cause is, the grid technology's user can enjoy the huge computing power and the data storage capacity. But, the people relatively are also strange to the grid technology in enterprise's application.

2.1 Present Situation of Grid

The grid computation is honored as after Internet and Web "the third generation information technology tide."[8]From the US, Europe, Japan and so on developed country to India and so on some developing countries, started the large-scale grid research project and obtains the industrial field to support vigorously. The British government has invested 100,000,000 pounds, the research and development "the British country grid", The US government uses in the grid technology the basic research funds amounting to 500,000,000 US dollars. The American Military is planning implements a giant grid to plan one "the global information grid", estimated that in 2020 completes.

Our country's grid computation research start soon. By the Chinese Academy of Science coordination "the national high performance computation environment (NHPCE)" the project and by the Tsinghua University coordination "the advanced computation infrastructure (ACI), Beijing Shanghai pilot project" two grid computational item has obtained the first fruits. Supported five national high performance data processing center which by the technical department organized already to move. At present our country's grid computation research mainly concentrates in Academia Sinica calculates, University of National Defense and Science, the south of Yangzi River to calculate, the Tsinghua University and so on several to have the strong strength research unit in the high performance computation aspect. These units have the very good technical accumulation in the high performance computation research aspect and very strong scientific research ability.

Until now, the grid computation did not have the official standard, but in the core technologies, the related organization achieved with the enterprise identically by government research and development organizations and so on United States Department of Energy and NASA impels for the center, the American Argonne country laboratory and University of Southern California information science institute (IS) the cooperation development's project Globus agreement has become the grid technology model to represent the other day with grid computation standard and the standard in fact. Globus Toolkit takes free software oneself after on the Internet public. Including Entropia, IBM, Microsoft, Compaq and so on 12 computers and the software manufacturers announced that will use Globus Toolkit, takes one kind of open construction and the opening standard infrastructure. Globus Toolkit has provided the primary service which the construction grid application needs, like security, resources discovery, resource management, data accessing and so on. At present possesses the significant grid project is the agreement which and the service construction provides based on Globus Tookit. In addition, including Global Grid Forum, the object management organizes (OMG), W3C, as well as standardized associations and so on Globus, org participated in the grid basis of calculation as well as the global big grid (GGG) standard formulation work.

2.2 The Basic Requirements of Grid

Computing power which provides regarding the grid, four basic requests, they respectively are the reliable request, the standardized request, easy visit request and low in price request

The grid reliability is refers to the computing power which the grid provides to guarantee that is continually, stable and safe, should not because of the grid interior individual resources change, but has the influence to the grid application, namely the grid interior partial resources' change to the grid application should be transparent, is similar to we use electric lamp's time daily should, because the individual power plant presents any breakdown to cause the entire electrical network electric power temporarily to make up the local electric power to be insufficient, the grid should also be able to guarantee that provides continually, the stable computing power. The grid should also satisfy the multiform safety requirements, for instance data transmission's encryption, the jurisdiction authentication, avoids

the illegal invasion and the illegal use and so on, if does not have the secure safeguard, this kind of advanced computing service cannot obtain the widespread promotion.

The grid standardized request is at the same time refers to between the grid resources to have the connection which or the protocol standard unified may visit mutually, because only then like this only then can realize the grid resources interoperability, thus realizes the full resource sharing, the standardization is the sharing premise; Standardized another meaning is refers to the grid the computing power which provides to the user to satisfy certain standard, one quite unified form, thus is advantageous by one unified way carries on the visit, regarding the visitor, cannot, because the time, the place, the concrete visit system and so on various requirement changes the visit form unceasingly, the visit form should have the uniformity, the certainly uniform premise is the grid must provide to the user a relative stabilization the standardized connection.
The grid easy visit request is refers to the user to be possible in any time, any place, visits and uses each kind of grid resources by the own custom's unified form. The grid computing power may assume the post what corner through the grid transportation, everywhere may result. That is, does not have the resources on the grid to occupy any position the concept, only then “on grid” or “not on grid” difference, regardless of you in any place, grid resources in yours side. People before when solution specific question can perhaps not but arrive at the specific place to carry on, for instance registers and the use special instrumentation equipment to some unit and so on, but when the grid solves in the problem, should not because of the visitor or the resources locus position different is restricted.

The grid expense's inexpensive request is the grid can accept generally with the promoted premise, no matter the grid has how many merits, if the majority users are unable to withstand its expense, the grid impossible to be popularized, its each kind of superiority is also unable on the basis to obtain manifests. The grid technology through shares fully the resources, maximum limit displays the resources the use value, may leave unused originally and the waste resources collects for the grid user use, moreover might avoid before, because the geographical position idle brought each kind of overhead, the grid has the very big insight expenses obviously to the user the potential.

These requests are the questions which the grid needs to solve, is also the grid technology display function place. The grid takes one kind new and the important infrastructure is not between one night can the miracle appear suddenly, needs each aspect to join up, joint effort only then to be possible to realize.

2.3 The characteristics of the grid

(1) distributivity: The distributivity is a grid most main characteristic. On the grid each kind of resources usual type is complex, the scale is big, the spanning geographical scope to be broad, under the distributed computing environment, needs to solve the resources and the duty assignment and the scheduling problem, safe transmission and correspondence question, timely safeguard question, between person and system as well as person and person's interactive question and so on.

(2) isomerism. The grid may contain many kinds of isomerism resources, including spanning geographic distribution many management territory. The constitution grid computing system's supercomputer has many kinds of types, the different type's supercomputer in the architecture, the operating system and the application software and so on many levels possibly has the different structure.

(3) extendibility: The grid may from contain the minority resources to develop at first has the tens of thousands of resources big grid. From this possibly brings a question is the performance drop which as well as the grid increases along with the grid resources causes retards, the grid must be able to adapt the scale change.

(4) sharing: The grid basic characteristic is the resource sharing, but is not its scale. Although the grid resources are distributed, but they are actually may share fully. The distribution is the grid hardware's in physics characteristic, but a share is in the logical characteristic which under the grid software support realizes.

(5) may the compatibility: In the grid, has many resources, the resources has the breakdown probability to be very high. The grid resource management or the application must be able dynamic adaptation these situations, to transfer in the grid the available resources and the service obtain the best performance. Is different with the common local area network system and single plane's structure, because the grid system the region distribution and the system cause its overall construction to change frequently complex, the grid system's application must be able to adapt this kind of uncertain structure.

Fig.2 grid system

(6) structure unpredictability: Dynamic and uncertain system behavior. In the traditional high performance computing system, the computing resource monopolizes, therefore system's behavior is may forecast. But in the grid system, creates the system behavior and the system performance by resources sharing changes frequently.

(7) multistage management territory: Because constitution grid system's resources and so on supercomputer usually belong to the different organization or organize and use the different safety mechanism, therefore needs each organization or the organization fellowship the solution multistage management territories the question.

2.4 Grid architecture

How is the grid architecture about constructs the grid the technology, including to grid basic building block and various part of function definition and description, grid various part of reciprocities and integrated method stipulation, grid effective operational mechanism portray. Obviously, the grid architecture is the grid skeleton and the soul, is the grid most core technology, only then establishes the reasonable grid architecture, only then can design and construct the good grid, only then can enable the grid to play the role effectively. At present the grid architecture's design had certain research, proposed the model includes: Five hour glass models, module model, opening grid architecture (OGSA) model, computation pond model, CPU model, neural network model, node model and so on. And five hour glasses are the classical models. These five respectively are:

(1) Fabric. Its function provides in the grid to be possible upwardly to supply sharing the resources, they are physics or the logic entity.

(2) Connectivity. It is in the grid network business processes corresponds and is authorized the control the core agreement. between the synthem submission's data exchange each resources' realizes under this control, authorization between various resources' examines the card, the safety control also to realize in here.

(3)Resource. Its function is to the single resources implementation control, carries on with the available resource grasps , to carry on the initialization, the monitor resources movement condition, the statistics and the payment related resources service data safely to the resources.

(4) Collective. This function is controlled the resources level submission the resources to collect in together, plans application procedure sharing which, the transfer for empty organizes. For to comes from the application sharing to carry on the management and the control, the collection level raises for the directory services, the resource distribution, the program arrangement, the resources proxy, the resources monitor diagnosis, the grid to start, the load to control the system, the account management and so on many kinds of functions.

(5) Applications. It is on the grid user's application procedure. The application procedure adjusts through each API with the corresponding service, completes the task again through the service transfer grid's on resources.

Five hour glass structures are an abstract hierarchical structure, its important characteristic is constitutes one “the hour glass” the shape. In the hour glass structure, the resources level and the connection level compose hour glass's bottleneck part together, provides the first floor for the grid computation the correspondence, the security as well as the partial resource management. Different high-level (hour glass's crown) the behavior maps their above, themselves can also map above the different basic strokes (hour glass's base), obviously bottleneck part's core agreement's quantity are few. The few core agreement is advantageous to the transplant, also realizes and has the support quite easily.

2.5 the key technology in Grid

To solve complicated technology computing and mass data service in different fields, the people construct different grids based on network interconnection, typically like computing grid,scavenging grid, data grid and so on. They vary in system structure and the problems to be solved, but they all common key technologies [9], mainly as

follow:

(1) high-performance scheduling technology. In grid system, a great number of applications share various resources of gird, so how to make these applications achieve the greatest performance is the problem that the scheduling need to solve. The grid scheduling system is more complicated than the one in traditional high-performance computing and mainly because the grid has some unique characteristics, like dynamic change 、 heterogeneity and diversity of grid resources and partial management of dispatcher, for grid scheduling, we need to create time-varying performance prediction model and make the best use of dynamic information of the grid to reflect the fluctuation of the grid performance. In grid scheduling, we also need to consider a series of matters, like portability 、 expandability 、 efficiency 、repeatability and combination of grid scheduling with local scheduling.

(2) resource management technology. The key matter for resource management is to allocate resource to the user effectively. High-efficient allocation involves resource allocation and scheduling and generally is reflected by a scheduling model containing a system model, while system model is abstraction of potential resources and in time provides visible resource information on all nodes to the allocator which will reasonably allocate resources to tasks after obtaining information, and thus optimize system performance.

(3) grid safety technology. The grid computing environment demands for safety more complicatedly than Internet for safety. In the grid computing environment, The users and resources are big in quantity and also dynamically changeable; Among many processes of one computing process exist different communication mechanism; resources support different authentication and authorization mechanism and may belong to multiple organizations. Those unique characteristics of the grid results in its much higher demand for safety, specifically like supporting safe communication between principals in grid computing environment to prevent principal impersonation and data leak; supporting; support the safety of crossing virtual organizations; supporting single sign-on of the use in grid computing environment, deferring trust and trust transfer of crossing multiple resources and sites, and so on.

The initial objective of the grid research is to hopefully link super computer into one remote-controllable computer system. Now, this objective has developed into constructing common base support structure for large-scale computing and data processing, and integrating various on-line high-performance computers 、 servers 、 PC 、 information systems 、 mass data storage and processing systems、application analog systems、 virtual reality systems、devices and information and information acquisition equipments(such as sensor), providing floor layer technology support for various applications and developments, to make Internet become one powerful and universal computer facility and finally realize resource sharing and distributed and coordinated working.[11]

2.6 task scheduling In grid computation

Briefly, the grid computation task scheduling's goal is must realize the most superior dispatch to the user submission duty, and tries to enhance the grid system's overall turnover rate. The concrete goal includes: Most superior span, grade of service, load equalization, economical principle and so on.

(1) most superior span. The span is one is most main, the most common goal, what refers to is the dispatch length, is also starts from the first duty to move the time which finished to the last task run experiences. The span shorter showing scheduling strategy is better. When user after grid system submission duty, the biggest desire is the grid system completes own task as soon as possible. Obviously, realizes the most superior span is the user and the grid system's common goal.

(2) grade of service QoS. The grid system must when provides the computation and the store service for the user, the user to the resource demand situation is through the QoS form reflects. The task management and the dispatching system when carries on the assignment dispatch task, safeguard grid application's QoS is must.

(3) load equalization. When development parallel and distributed computation application, the load stabilization is a key question. The grid system further expanded this question. The grid task scheduling is involves the overlapping territory and the large-scale application dispatch. Solves system's load equalization is a very important question.

(4) economical principle. In the grid environment's resources in the geography are the widespread distribution, moreover each resources belong to the different organization, has the respective resource management mechanism and the policy. According to real life's in market economy principle, the different resources' working costs should also be not same. The market economy actuation's resource management and the task

scheduling must make to expend both sides (resources user and resources tenderer) the mutually beneficial interaction, can cause the grid system to develop for a long time.

The objective of the grid task scheduling is to promote system throughput, and meet the demand of application program with existing limited computing resources.[8] For grid scheduling, at present, we mostly adopt heuristic scheduling algorithm. There are a lot of existing grid scheduling algorithms, like Min—min algorithm, Max—min algorithm, Suffrage algorithm, SA algorithm, etc. Min-min algorithm is to allocate resources to the task which is executed the earliest and has minimum execution time as many as possible, and is the algorithm based on minimum execution time. For this algorithm, in each time of task scheduling, what we consider is all non-allocated resources and its lack of system load balancing and lowness of resource utilization.[9] Max-Min algorithm refers when there are multiple parallel routes, they system gets maximum or minimum index, e.g. when there are multiple parallel execution work flow, it gets the maximum value for workflow execution time, and minimum for reliability, namely, For Max-min algorithm after getting minimum execution time of each task, it chooses the maximum execution time to conduct tasks and computer resource matching. For Suffrage algorithm it computes the difference value between minimum execution time and less maximum and chooses the task and computing resource with the minimum difference value in all tasks. For OTE Min-min algorithm it firstly allocates short tasks, and while when long tasks are running, other nodes remain idle and can't provide better load balancing. For POTE Min-min algorithm, it solves the problems existing in OTE Min-min algorithm by assigning priority based on OTE Min-min algorithm.[4] The task priority can be defined according to the demand of the tasks on data transmission capability of node computing，size of tasks、task consumption、task execution time limit and other different demands. After setting task priority, it divides all tasks into several subsets according to the level of priority. The subset task with higher priority is prior to the one with low priority in scheduling.

These algorithms study the scheduling at a different angle based on different scheduling model. In so many algorithms，Min-min algorithm is simple、fast and effective algorithm，but it can not guarantee load balancing because this algorithm always allocates small tasks. Aiming at the defect of Min-min algorithm, Segmented Min-min divides all tasks into N segments according to expected execution time and makes the scheduling starting from the segment where the task has longest execution time by Min-min algorithm. QoS Guided Min-min considers the demand of the tasks on data processing capability of node computing in scheduling process, and makes the scheduling firstly for the tasks which need high data transmission capability of node computing. The above methods all have improved Min-min algorithm at some angle, but none of them makes further analysis on the computing of task execution time, nor fully considers the influence of the demand of the task own itself on scheduling.

3 Slicing

Slices the technology in the medicine domain, the project processing domain and the computer domain application is widespread, is the current quite popular key technologies.

3.1 Slice definition

Slice, slide specimen one kind. Zoology and botany tissue slice which observe for the optical microscope or the electron microscope. Because requests differently, the available bit to carry on the unarmed slice, may also organize the block embedding in the paraffin wax either the collodion or by cryogenic freezing, with slicer slice. Slivers 5~10 microns thin slices, for optical microscope observation. Cuts the system with the epoxy resin or the methylacrylic acid embedding organization block the ultrathin section, its thickness in 20~50 nanometers, feeds specially under the electron microscope observes. Generally the teaching uses like the root point, stem's slice is generally called the paraffin section.
The slice, in the charting software or in the homepage manufacture software, slivers the image several parts, a piece of piece toward the upload, the speed which like this uploads is quite quick.

The slice, is carries in the slicer, slice thickness because of needs to decide, generally about 5-7 microns.

In tobacco's slice, in the piece smoke processing craft, after will open a box the volume big piece smoke block divides slivers certain thickness to favor loose and the resurgence technological process.

3.2 key technologies in Slice

In the slice technology application's major technique

is the feature selection and the feature Recognition.

Feature Selection

The feature selection is in the pattern recognition as well as an information theory important component. The feature selection is the expectation chooses a small space from the feature space, enables between the category to have the best separating capacity. How does feature selection's basic task discover these most effective characteristic from the multi-dimensional characteristic. Speaking of the validity, the feature selection divides into the following two kinds:

Chooses a dimension smaller characteristic subspace from the feature space by most effective expression some kind of object oneself.

Chooses a dimension smaller characteristic subspace from the feature space to use in the most effective discrimination different kind of object.

One of feature selection's key questions is the characteristic which how to appraise obtains, causes the characteristic which chooses to reflect classified well the information. Fuzzy mathematics' development appraised the characteristic for us the performance to provide the new method, they were the characteristic which appraised through the weight fuzzy set's fuzzy degree uses regarding the classified performance, namely used these characteristics the category separated difficulty degree.

In recent years, the feature selection domain's research has the new achievement, such as mutually entropy (or mutually information), multiple correlation/partial correlations, fuzzy theory measure, neural network as well as genetic algorithm in feature selection application and so on.

Feature Recognition

In pattern recognition process, the people always use the multi-dimensional data namely so-called characteristic, expresses the goal, moreover in the ordinary circumstances, always withdraws the numerous characteristics by the comprehensive expression goal. Uses the complete characteristic to carry on the target identification blindly, not only will bring the recognition operand to be big, recognition speed slow and so on questions, even possible, because used unreasonable has not been able to differentiate the different category goal the characteristic to cause the identification probability effectively to drop. Therefore, to enhance the classified recognition algorithm the reliability and the efficiency, needs to carry on the reasonable choice to the characteristic, chooses speaking of this sorter most can differentiate the different category effectively the characteristic.

3.3 Slice technology application in Computer

（1）Slice Technology Application in Image

In imagery processing, because image own data quantity is huge, uses the view picture chart to carry on the classified recognition directly not too to be feasible, requests us to carry on the transformation to the primitive image data, obtains most can reflect this image essence the characteristic. This is the feature extraction process. Usually, may withdraw its ensemble average to the image, the population variance, each different direction variance, traditional and so on 7 not bending moments characteristics; May also carry on it each kind of different transformation: Including the Fourier transformation, Gabor transformation, after each kind of wavelet transformation, obtains the data takes its characteristic; May also define other statistics even non-statistical data to take its characteristic voluntarily. Thus, speaking of an image, we may gain many characteristics.

In does not affect the page quality under the premise, to cause the homepage changes is small, only then reduces the picture the size, therefore cuts the chart to become manufactures page's fundamental mode now. Slices the tool is realizes the cutting picture.

Slices the technology mainly to use to the image carries on cutting to optimize, a big image cutting for many small image's forms, strengthens to the homepage the support, saves the upload, downloads and opens the homepage the time.

（2）Slices the technology in computer's application

As soon as the digitized hypothesized person is through to has certain characteristic corpse specimen to carry on the slice, cuts one piece every time, carries on the photograph, the scanning, then processing transforms as the computer data, in the computer the conformity reconstruction's human body three dimensional spatial structure, it is discipline intercrosses, the integrated development leading edge interdisciplinary studies and so on information technology and medicine.

4 Solutions

The Min_min algorithm is the classical grid task scheduling algorithm, it is one kind under isomerism environment heuristic dispatch algorithm. We study the task scheduling algorithm takes the basic philosophy by the Min_min algorithm, and performs in the inferior foundation to improve. Causes entire grid system's Makespan and the load stabilization has the enhancement.

4.1 Dynamic Task Scheduling Strategy Based on Slicing Technique

According to the demand of the tasks on data transmission capability of node computing，size of tasks、task consumption、task execution time limit and other different demands, we can divide one big task into several small time slices each of which is one scheduling on one resource, and assign to each one with one priority according to task consumption and task execution time limit, and integrate the scheduling tasks in one same resource into several big tasks, and make concentrative scheduling on grid resources according to relevant algorithms, which equals to splitting one big task into several small tasks, then integrating them into some big tasks, and meanwhile scheduling grid resources so it not only shortens scheduling time, but also makes network resources scheduled in balance.

4.2 Dynamic Task Scheduling Algorithm Based on Slicing Technique

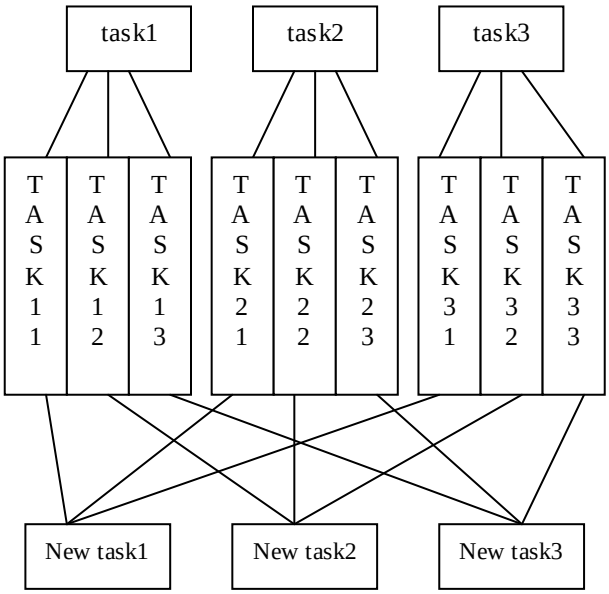

Fig.3 Task scheduling model

Description of the slice Min—min algorithm:
Slicing Min—min {
（1）split the task into some Slice by some reasons
（2）compute the priority of each Slice based on some requirement
（3）sort the tasks into a Collection in a order of their priority
（4）partition the tasks evenly into some parts and each part compose a sub aggregate Ti
（5）while task is not empty
（6）initialize scheduling list and initialize tasks in the beginning of each node of data transmission time
（7）compute time to complete the task
（8）Select the best task node to match
（9）Add it to the scheduling list
（10）Node amend all the tasks corresponding to the beginning of the data transmission time
（11）endwhile
}

The first line in algorithm is split into several slices according to the difference of the resources that tasks call firstly, then in the first line each time slice is assigned with one priority according to task consumption and task execution time limit; the third time will call the tasks in one same resource and integrate into one sets; the fourth line will call the tasks in one same resource and then split into several subsets according to time limit and other factors; the fifth to eleventh lines will make scheduling on the tasks in all subsets; the sixty line will initialize the scheduling list and the time that each task begins to transmit data in each node; the seventh line will compute task exeuction time; the eighth will select the optimal task/node to match; the ninth line will add it into scheduling list; the tenth line will modify the starting time of data tranmission of all tasks corresponding to nodes.

This algorithm is carries on the slice division differently according to the transfer resources. A dispatcher possibly only transfers one resources, also will possibly transfer several resources. Transfers the resources is possibly the serial structure, after transferring the preceding resources, carries on the next resources again the dispatch; Also possibly is the parallel structure, the transfer resources not successively order, may also carry on. We the dispatch which may simultaneously carry on, divides for the identical priority, has successively the smooth dispatch, according to dispatch time sequence division priority.

After slice division, carries on the combination the dispatch according to the transfer resources type. In this combination, carries on the dispatch according to the duty priority, if has the same priority, may defer to the smallest priority the method to carry on the dispatch.

In Slice Min-min algorithm, when the tasks and the nodes of scheudling resources are the same or close in quality, the performance of this algorithm is close to the one of tradional Min-minalgorithm, becasue the times of calling resources are not largely differnce and the time for calling each resource is near. When the tasks and the nodes of scheudling resources are far different in quality, Slice Min-min

algorithm will make total task exeuction time reduced becasue of less called resources and longer resource-occupying. Udner ideal condition, the execution of any task may be parallel to other tasks and does not use resources without releasing resources, and at that time the performance promotion rate is that the transmission time accounts 100% in task execution time.

4.3 Dynamic Task Scheduling Model Based on Slicing Technique

In grid, the users shares grid resources s by submitting computing tasks to grid system. The grid scheduling program schedules these tasks to appropriate computer resource nodes in some strategy according to the demands submitted by the users, like service quality, and positions the data needed by the tasks in grid system. The nodes of computing resources begin to execute after obtaining the data needed by the tasks.

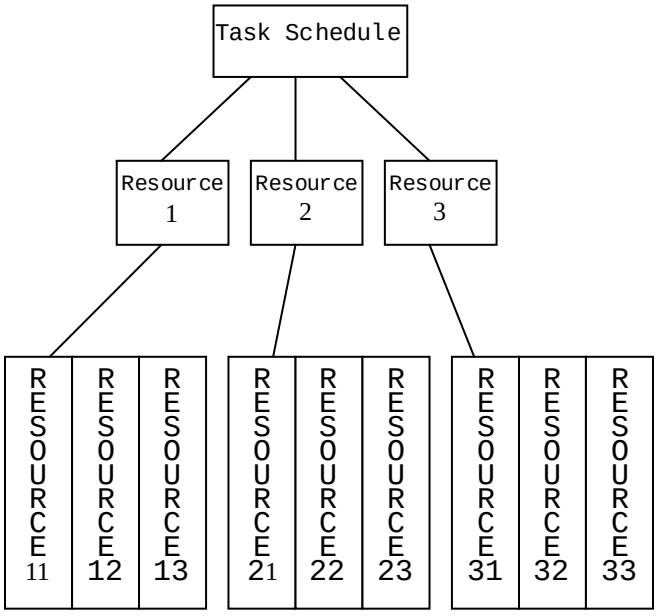

Fig.4 Resource scheduling model

In grid system, it statistics existing network usable resources, and assigns with relevant priority according to the execution situation and trustiness grade of each stituation when it affords tasks originally in grid, and make resource scheduling according to priority.

The dynamic characteristic of grid demands the matching relationship to adjust with the resourceand the change of task, and the scheduling event to repeatedly be called.Supposing when each scheduling event occurs,it is known:(1) node quantity of usable data resources for grid, node quantity of computing resources, network bandwidth, processing capability of computing resources; (2)size and position of imported file; (3)running tasks and data transmission list. The occurrence of scheduling event will be predicted by evaluating resource performance.

The basic model of Schedule（ ）is as follows：

Schedule（ ）
{
（1）Carries on the classification the existing grid resources according to the category
（2）May trust the degree each kind of resources according to the resources use to carry on the slice division
（3）compute the next scheduling event
（4）create aschedulelist
（5）foreach unscheduled tasks compute all estimate of its completion time
（6）select a task／host pair according to slice priority height rule and add it into the schedule list
（7）until each host has been assigned enough work and all task have been scheduled
（8）schedule the tasks according to the schedule list
}

The first line classify the grid resources by category; The second line divides each category of resources by the trustiness of resource usage; the third line computes next scheduling event. In computation, it will consider the change of grid resources as to increase or decrease the occurrence frequency of scheduling event. The occurrence frequency of scheduling event has great influence on the grid performance. High scheduling frequency can result in good adaptability and also is necessary for unstable grid environment, the higher the frequency, the more the scheduling consumption; the fourth line initializes scheduling list; the fifth line computes predicted execution time for each task, then in the sixth line, it selects task/node couple by the rule of dividing priority rank and adds the task and node into scheduling list. It repeats the sixth and seventh until all tasks have been allocated; the eighth makes task scheduling according to scheduling list.

The model makes layered scheduling according to task demand in the consideration of strategy of priority based on trustiness, which is beneficial for more reasonable scheduling on tasks. Because it has considered the trustiness grade among local areas and all nodes, it can avoid the tasks scheduled to unstable nodes as much as possible. If encountering node in error, it will transfer task to other trusty node, which largely increase the effectiveness of grid

computing.

5 Conclusions

At present the grid task scheduling algorithm faces the question is:

(1) design optimization (smallest execution time) the dispatch algorithm is a complete NP question. For example: The dispatch has the unit length duty which the unit correspondence retards to the resources core question which does not decide is a complete NP question. Optimized method dependence thumb principle, such as appraisal critical path question.

(2) accurate appraisal task execution time and the correspondence detention are difficult, for example in the compiling duration estimated that the correspondence extends is not feasible, because the movement time network struggles with the detention.

(3) static state method cannot solve the processor speed multiplicity, because their load variables.

The task scheduling and resource scheduling are the key matters for grid computing in distribute. The article puts forward task scheduling algorithm and resource scheduling model based on slicing technology. In order to differentiate each kind of dispatcher or segment a dispatcher certain sub-dispatches, needs according to the corresponding algorithm when carries on the slice division to the dispatch, the segmentation, entrusts with the corresponding priority, carries on the task scheduling according to the priority. The resources basis credible degree's difference, entrusts with the different priority, carries on the scheduling of resources again according to the priority. In comparison with traditional Min-min，it may make scheduling on tasks more reasonably and avoid the tasks scheduled to unstable LAN or nodes as much as possible , which largely increase the effectiveness of grid computing. Therefore, the task scheduling and resource scheduling based on slicing technology is one comparatively reasonable and effective solution to present grid computing.

But under the distributional system's task scheduling algorithm does not have the optimal solution, we study the task scheduling algorithm is also one heuristic algorithm, similarly is not the synergy. This topic discusses the task scheduling algorithm the goal is the hope through Slice Schedule model statement, provides a new research space for the grid task scheduling algorithm.

Moreover because the research level and the research time limit, but also many will wait for the question which will consummate and continues to discuss keeping in future work to give to solve. The grid computation is in the swift development the stage, makes many progresses from the first floor to the high-level application, but has many questions to need to solve, for instance in grid domain architecture question, security problem, cross territory resource management and union scheduling problem and so on. This article proposed Slice the Schedule model is also conducts the research theoretically, is merely regarding a model building simulation.

References:

[1] Du Zhihui.Chen Yu.Liu Peng,Grid Computing[M] Tsinghua University Press，2002

[2] FOSTER I.KESSEI MAN C,The Grid：Blueprint for a Future Computing Infrastructure [M] . Morgan Kaufmann Publishers，USA，1999

[3] BUYYA R.MURSHED M.A Deadline and Budget Constrained Cost-Time Optimization Algorithm for Scheduling Task Farming Applieations on Global Grids[R]. Technical Report，Monash University，March 2002

[4] HeYan.Li Kenli.Shi Kuiran.Liu Xiaoling.Wang Ying,Grid Scheduling Algorithm Based on Priority and Optimization Execution Time, Computer Application [M]，2006,Vol 26, No.1 ,PP.61-64

[5] Huang Baobian，Zen Wenhua. Dynamic Task Scheduling Based on Trust Mechanism in Grid Computing. Computer Application [M], 2006,Vol 26, No.1 ,PP.66-69

[6] Xu Wenqing,Yang Hongyu.Real Time Mixture Task Scheduling Algorithm ,2006,Vol 3, No.14 , PP.78-82

[7] Li Zaimin.Chen Ying,Optimal Scheduling. Systems Engineering and Electronics of Real-time Multi-task Stream System Based on Statstics and Anaylysis,2001,vol 10, PP. 721-727

[8] Liu huai,Shen jie.Fei Shumin,Use Double Priority Algorithm Scheduling to Control Rea-time Periodic Task ， 2003,vol 8, PP. 12- 19

[9] Wang Yongyan, Wang Qiang, Wang Hongan, Jin Hong,Dai Guozhong.Real-time Scheduling Algorithm and Its Realiazaiotn Based on Prioirty List ， 2004,vol 3,PP.360-371

[10] Liu Wen,Guo Li.Zhang Lichen,Li Shuguang.Deferring Algorithm of ATM Network Based on Static Priority ,2002,vol 3,PP.322-324

[11] Liu Zhengtao. Mao Yuguang, Real-time

Scheduling Algorithm of Data Stream and Database Matters and its Realization，Mini Microcomputer System，2006,vol 12, PP. 32-44

[12] Wu Xiongqi,Zeng Wenhua. Scheduling Algorithm of Resource Based on Improved Genetic Algorithm grid， Microelectronics &Computer，2006,vol 9,PP.26-28

[13] Dai Guozhong,Zhou Bing,Fang Ting ,Wang Hongan ,Wang Kun,Qiaoying.Real-time Research on Dynamic Scheduling Algorithm of Heterogeneous System, Computer Research and Development， 2002,vol 39, No.6, pp.725-732

[14] Zhang Huijuan,Cui Hongming ,Zhou Lihua. Research on Real-time Collaborative Scheduling Algorithm,Computer Engineering and Design，2004,9. 25(9). PP.1447-1450

[15] Cai Guoyong,Gu Tianlong, Dong Rongsheng, Zhao Lingzhong. Research on Predictable Scheduling of Non-periodic Task , Journal of Guilin University of Electronic Technology，2001,vol 3

[16] He Fugui，Wang Jiali. Dynamic Scheduling Algorithm Based on Limited Priority， Electronic Technology University Press， 2007,vol 3

[17] Jin Hong，Wang Qiang，Wang Hongan，Dai Guozhong. Real-time Scheduling Based on Dynamic Preemptive Threshold，Computer Research and Development， 2004,vol 3,PP.10-15

[18] Liu Yang，Gui Xiaolin，XuYuwen. Research on Scheduling Methods of Based on Priority in Grid Operation Stream，Journal of Xi'an Jiaotong University，2006,vol 4,PP.411-414

[19] TianYinhua，Fu You. One Adjustable Dynamic Task Scheduling Algorithm， Shandong：Journal of Shandong University of Sciences &Technology（Natural Sciences Edition）， 2007，vol 1,PP.74-78

[20] Zhou Bo，Qiu Weidong, Chen Weinan,Peng Chenglian. Quick Dynamic Scheduling Algorithm of Reconfigurable Real-time Task， Journal of CAD and Computer Gaphics，2006,vol 4,PP.53-59