Research on the Grid-based Production Management System for Digital Oil Field

1, 2 Zhang Xian-wei, 1 Zhang Jing
1 School of Computer Science and Engineering, Xi’an University of Technology, xwzhang@xstu.edu.cn
2 School of Petroleum Resources, Xi’an Shiyou University, xwzhang@xstu.edu.cn

Abstract
With the development of Internet and information technology, many petroleum data management providers begin to research and develop their own data management platform in succession. Yet few of them pay much attention on the integrating of several services like automatically charging, resource description and searching, load balancing, which hinder the further application of data sharing and application. To overcome this shortcoming, Grid-based Production Management System for Digital Oil field is proposed and the architecture, components and workflow are described in detail. finally the classical application of data grid in digital oil field are proposed, which validate the validity of the system described in the paper..

Keywords: data grid, oil field, production management system.

1. Introduction
With the rapid development of information technology, there is gradual increasing on the automation level of cruel oil production and the concept of digital oil field [1-3] has also emerged. Digital oil field is to use a comprehensive information technology as a means to achieve corporate oilfield entities and digital, networked, intelligent and visualization. The full integration of Digital Earth Digital Oilfield technology, massive data integration and processing technology, multimedia and virtualization technologies, the process of trying to achieve the full crude oil production of information. As a result, the constructing of digital oil field relies on variety of services located on broad locations. So how to find the oil production network services they need to be a challenging problem. Regarding this problem, some people introduce the architecture of grid to the establishing of industry control network, which control the filed devices for oil production. The core idea of the grid is providing end users with an integrated network resource sharing and collaboration platform for the operation of services, allowing users to maximize the sharing of resources and coordination of resources to achieve targets. However, because of network connectivity in the form of diversity and heterogeneity of services, resources, oil field construction of the grid are still many problems, and QoS is the one of most prominent.

With the development of Internet and computation technique, data management becomes all computation areas’ bottleneck. How to manage multi-disciplinary distributed data effectively become interested issues for all computation areas. Some statistics show that the scientific researchers will spend about 80% time on preparing data needed and other only 20% time will be used to real computation. Especially, distributed data management issues become more important than centralized data management. Oil field is such an area, it involves all kinds of data, such as simple character data, media data and complicated bulk data volume and so on, and these data is distributed in different subsidiary companies. It is very difficult for subsidiary companies to manage these distributed data, further more, it is very difficult for scientific researchers to use these different type data to research geology, reservoir and other oil field features. In order to satisfy scientific researchers’ research, many oil fields begin to construct their own data center one after another. However, data center does not implement uniform data management, all the data can not implement independent data services’ functions, and current data management based on data center is bound with concrete applications. The final objective for oil field information construction is to implement digital oil.

To solve this problem, some key technologies like automatic charging, resource description, load balancing and service discovery are imported to construct the production management system for digital oil.
production, which provides great convenience on the production process control and producing situation displaying.

2. Type Style and Fonts

As the great significance of the application of data grid constructing in oil field production, lots of attentions are attracted in not only industry but also academy. Great amount of achievements are made in this field, some of them with great respective are listed as follows.

Davis, Brian[4-6] discussed the digital development concept Smart Fields by Shell, which is focused on the core management processes for production management processes for production operations, production optimization, well and reservoir surveillance, and field development planning and the application of advanced process control concepts to the integrated asset WHICH is treated as a single dynamic system from sub-surface well facilities, through the production system, and along the supply chain to the customer delivery point. Lvas, B. U.[7-9] characterized Digital oilfields by the extensive use of information technology. Based on new ICT tools, work processes are changed and functions are moved from offshore installations to onshore centers, where technology and people merge into new organizations, explores the main features of these new organizational changes, and assesses new threats, risks factors and vulnerabilities in an MTO perspective. Some uncertainties concerning the effect of digitalization on the overall level on new work processes, technology and environment are brought to light, as well as proposed measures to be taken, to improve risk management systems, methods and practices. Shen, Longbin[10] applies object relationship mapping (ORM) technology of iBatis in order to decouple the degree between program objects and the business rules, which is helpful to the integration of all kinds of oil-gas exploration and production applications in the oilfield. For bulk data, this paper takes a multitude of methods such as data compression, organization, cache and so forth, which focus on solving the performance of bulk data service. For GIS data, an adapter is designed by the standard protocol and commercial API packages, which supports the accessible performance and extensibility. The data service platform has been realized in the exploration decision support software system using coding language based on the technologies above.

The achievements mentioned above mainly focus on the architecture of data grid in the oil field production, few of them pay attention on the technologies like automatically charging, resource allocating etc. Regarding this problem, a data grid is imported into the oil field constructing and the architecture and workflow are described in detail.

3. Architecture of Grid based Oil Production Management System

3.1. Hierarchy of digital oil field

Petroleum area[11] integrates multi-disciplinary, this area involves all kinds of complicated knowledge, such as reservoir, geologic, logging, seismic, physical geography, drilling, management, plan and so on. It involves data from simple character string to complicated three-dimension bulk data volume, and the architecture of digital oil field can be described like figure1.

A typical oil field is composed of oil field, oil metering station, oil union station and oil field control center. To get service needed, a terminal like mobile phone, notebook and desktop should choose a template from service pub server and fill the parameter needed for the service first. Service pub server will then deal with the template and abstract data from it and search for the candidate service from database. In the second step, the monitor server will locate the service and produce the url according to the service description and the available computing and storing resources. In the third step, the service requestor will access the service from the url accepted and enjoy the service. Finally, the reply will be given to the service pub server for the QoS recommendation, which will be stored and analyzed for the preference model constructing.

Data from real time database and some other data source will be collected and categorized into exploration data, developing data, oil gathering data, ground engineering and research data. These data will be fused and mined to construct address model, reservoir model, workflow model and enterprise management model. The models provided are imported into several systems like production management, specialized software application, three-dimensional visualization system and GIS for trend predicting. The
top level enterprise information website is the entry of the service on the digital oil field, which provides variety of service online.

3.2. Architecture of Digital Grid

In order to facilitate resource sharing and service scheduling optimization, the digital grid architecture are introduced for digital oil field constructing. The structure can be described like figure 2, database, application, immediate communicating and storage service are provided online, which can be searched and accessed through service pub server and the state monitoring server are responsible for monitoring the state of clusters to acquiring the latest server state and publishing them on the service pub server, these critical resources are connected through oil field wan, which providing accessing point for mobile phone, notebook and terminals.

This service is focused on interactions with a single resource, it contains protocols and services (and APIs and SDKs) that are not associated with any specific resource but rather are global in nature and capture interactions across collections of resources. For this reason, we refer to this layer as collective layer. Because collective components build on the narrow resource and connectivity layer “neck” in the protocol hourglass, they can implement a wide variety of sharing behaviors without placing new requirements on the resources being shared.

Oil data grid core[12] is oil grid database, Oracle Company proposed a database Oracle 10g based on grid computation. Oracle Database 10g is the first database designed for enterprise grid computing, the most flexible and cost-effective way to manage enterprise information. It cuts costs of management while providing the highest possible quality of services. In addition to numerous quality and performance enhancements, Oracle Database 10g significantly reduces the costs of managing the IT environment, with a
simplified install, greatly reduces the configuration and management requirements, and automates the performance diagnosis and SQL tuning. These and other automated management capabilities help improve DBA and develop productivity and efficiency.

The storage cluster is responsible for mirroring image storage disks or local storage disks, these storage disks make up of storage pool. Database layer can manage the storage automatically.

4. Working flow description

A typical oil field is composed of oil field, oil metering station, oil union station and oil field control center. To get service needed, a terminal like mobile phone, notebook and desktop should choose a template from service pub server and fill the parameter needed for the service first. Service pub server will then deal with the template and abstract data from it and search for the candidate service from database. In the second step, the monitor server will locate the service and produce the url according to the service description and the available computing and storing resources. In the third step, the service requestor will access the service from the url accepted and enjoy the service. Finally, the reply will be given to the service pub server for the QoS recommendation, which will be stored and analyzed for the preference model constructing. The detail workflow can be described like figure3.

5. Key technology Description

5.1. Service charging mechanism

In the service providing process, the charging step is critical, which should be done collaborating with Pub Server, state monitor and service requestor with the help of globus, a famous service middleware for grid constructing. The charging step can be described like figure4

1. User forward a request to a grid
2. Grid execute the request with globus;
3. The state monitor server will log the request into database;
4. Service Pub server will calculate the resource used in the request responding and charge the default account;
5. Notify the terminal with email or short message
Each user has a grid system using a personal account, using the grid system for authentication; account can be divided into several levels, different levels have different permissions, Grid system to all resource nodes CPU time, main memory capacity, auxiliary storage capacity, I/O channel (such as network) traffic and bandwidth, software resources, use of time and other resources as may be involved in the valuation of resources indicators calculate the total price. Indicators of each resource you can use a static unit price, also a variety of resources based on demand, and user levels and permissions, dynamic generation unit price of each resource indicators. Grid usage charges can be used to calculate the real currency, you can also use e-currency and other virtual currency. Each time a user submitted a task (Job) is finished, the billing system on a user's account balance, billing is a mission that is conducted as a unit. Each task (Job) calculated the total cost of resources used by the way is: use a variety of resources (if R1, R2, R3 ...) are multiplied by the corresponding unit of resources used (such as R1, R2, R3 ... corresponding to the price is P1, P2, P3 ...), the use of all resources is the sum of the price the total price. That is, the total price of P = R1 * P1 + R2 * P2 + R3 * P3 + ..... 

Figure 3. Workflow of service providing

5.2. Replica service providing

This service can help users optimize access characteristics, such as grid structure and performance, storage system and file characteristics. In addition, it can provide intelligent scheduling to determine appropriate replica, site for (re)computation, etc. Data grid provides us with overall data management solution. We always explore how to manage those tremendous and complicated data, but there is not a good way to solve. Especially, in the oil exploration and production and oil management area, very day, lots of data be collected into a database, but how to effectively manage these valuable data asset and how to extract all kinds of data required by multi-discipline researches, these issues puzzled data administrators at all times. Data is the base, only data is managed effectively, scientific researchers can
focus their attentions on researches, and save much time to discover and collect those data interested. Data is managed effectively, can it implement multi-disciplinary cooperation researches of hard problem.

This service is focused on interactions with a single resource, it contains protocols and services (and APIs and SDKs) that are not associated with any specific resource but rather are global in nature and capture interactions across collections of resources. For this reason, we refer to this layer as collective layer. Because collective components build on the narrow resource and connectivity layer “neck” in the protocol hourglass, they can implement a wide variety of sharing behaviors without placing new requirements on the resources being shared. For example: directory services, co-allocation services, scheduling services, brokering services, monitoring and diagnostics services, data replication services, grid-enabled programming systems, workload management systems and collaboration frameworks, software discovery services and community authorization server, community accounting and payment services and collaboration services and so on.

![The workflow of service charging](image)

5.3. Trust value updating mechanism

The recommendation updating process of service is also a trust modeling workflow; the detail can be described as follows.

The definition of security has not been uniform, and there are several different definitions for the same time. The most popular one is that the behavior of entity can be trusted only if the entity's actions are consistent with the expectation. Regarding the definition given, the trust model and related security metrics can be described accordingly.

Trust is a binary relation, which can be one to one, one to many, many to one and many to many. In addition, trust is both subjective and objectivity, but not necessarily symmetric. Trust can be measured, and the degree of trust level can be divided. Trust can be passed, but the value will be partially lost during transmission. Trust is dynamic, and related to the context and time factors. Trust model is built out according to the characteristics mentioned above, which can be used to model the behavior of trust and calculate security metrics for trusted routing.
Trust itself is a judgment to the trustees which is made by the trustors according to their observations and knowledge. And trust relationship can be quantified as the trust degree.

Security metrics are used to calculate the trust value which quantify the extent of trust and can be used for routing path choosing. The security metrics are calculated according to the following rules.

If A and B had had prior contacts, The trustworthy of B can be determined by examining the past performance of B. We denotes this kind of trust value obtained through direct contact $DT_{AB}$

Let A and B has no contacts before, but now A has to judge the trustworthy of B. In this case, A can go to ask a more familiar entity C to obtain the trust value of B, like figure 1, provided that B and C have had direct contact. Because of the trust value is not obtained through the direct experience of A, we call the trust value obtained in this way indirect trust value, expressed by $IDT_{ACB}$

![Figure 5. Recommendation trust](image)

If C also did not experience direct contact with the B, C may have to ask D which is more familiar with the entity B to obtain the trust value of B. This is known as multi-level recommendation trust. In addition, A may ask both C and C for the trust value of B. In this case, calculating the trust value of B requires the synthesis of $IDT_{ACB}$ and $IDT_{ACB}$

![Figure 6. Composition of Trust](image)

The calculation process of $DT_{AB}$, $IDT_{ACB}$, $IDT_{ACB}$ and $IDT_{ACB}$ is determined by the features of specific application.

The composition rule of trust value can be described by formula (1)

$$T(i, j, t) = \prod((j, t) + \delta_1 DirT(i, j, t) + \delta_2 iDirT(i, j, t))$$

With $\prod((j, t) = 0$, the security state of $j$ can not be clarified or can not satisfy the requirement at $t$ denotes the trust value of service $j$ at time $t$, $DirT(i, j, t)$ and $iDirT(i, j, t)$ denotes the direct and indirect trust value separately. $[\delta_1, \delta_2]$, $\delta_1 + \delta_2 = 1$ is the weight vector of trust value composition.

The direct trust value can be described by formula (2)

$$DirT(i, j, t) = \alpha_0(j, t) \exp(-k_1 \cdot \phi(j \rightarrow i \mid 0..t-1)) - \beta_0(j, t) \exp(k_2 \cdot \phi(j \rightarrow i \mid 0..t-1))$$

where $\alpha_0(j, t) = C_2 \frac{\langle l \rangle_{N(t)} - \langle l \rangle_{N(t)}}{\langle l \rangle_{N(t)}} + 0.5$
if router \( j \) is admitted to access by \( i \), it may result in the reduction of the network characteristic path length. This feature is described by equation (3).

The static risk can be evaluated as the number of routing paths which would be cheated after router \( j \) has accessed to the network, like equation (4).

\[
\beta_0(j, i, t) = \frac{BTNS(e_j)}{C(sp)}
\]  

(4)

Trust model can only tolerate the risk to a certain extent. In order to better characterize this behavior, we use equation (5) to balance the risks and benefits

\[
k_1 + k_2 = \frac{\ln(\alpha_0(t, j)) - \ln(\beta_0(t, j))}{\epsilon_i}
\]  

(5)

The indirect trust value can be described by formula (3) \( DirT(i, j, t) \), a weighted mean recommendation, denotes the indirect trust value of \( i \) to \( j \) at time \( t \), which is composed from the recommendation from other services

\[
iDirT(i, j, t) = \frac{\sum BTNS(k) \cdot T(k, j, t-1)}{\sum BTNS(k)}
\]  

(6)

Where \( BTNS \) denotes the importance of the router

\[
TRE(p_\lambda) = T(i, k_1, t)^{-1} \ln T(i, k_1, t)^{-1} + T(k_1, k_2, t)^{-1} \ln T(k_1, k_2, t)^{-1} \\
+ \ldots + T(k_m, j, t)^{-1} \ln T(k_m, j, t)^{-1} \\
= \sum_{\epsilon \in P_i} T(n, m, t)^{-1} \ln T(n, m, t)^{-1}
\]  

(7)

6. Case Study

In order to validate the validity of the system designed, a drilling data calculating and displaying case is described. The request is to apply some resource including computing and storing needed by the request responding, and return back the dealing result with graphic interface. The software used in the case constructing includes (1) Globus Toolkit 2.X generic grid middleware platform that provides security, information, data, resources and task scheduling and other services. And to provide the above services programming interface API. This component is the core and backbone of the grid platform. (2) CoG Globus Toolkit provides the Java API. As the Globus Toolkit 2.X does not provide services or to provide a tool to generate the Portal user interface, CoG Portal can provide added functionality, and provides Web forms user interface. (3) Ganglia grid system performance monitoring components. (4) MPICH parallel job submission module. (5) MySQL provides a free database system. The result can be shown in figure 5.

![Figure 7. The result of the drilling data analyzing with data grid](image_url)

7. Conclusion and Future Works

With the rapid development of information technology, data sharing and service providing play a more and more important role in the production process of oil field production, traditional researchers
have made some achievements in grid constructing, yet few of them pay much attention on the variety of services can be integrated. Regarding this problem, a grid based oil production management system for oil field can be made and the structure, workflow and some key technologies are described in detail. In the future, our research will focus on the improvement of resource allocating algorithm which will greatly increase the efficiency of request responding.

8. Acknowledgement

The work was partly support by 863 key topics (2007AA010305).

9. References