Stratified grid systems as a course of system engineering development in the oil production industry

Arsenevskiy Ivan Sergeevich
Graduate student,
National University of Oil and Gas «Gubkin University»
119296
arsenevsky@mail.ru

Kalinin Vasily Valerianovich
Professor., Dr.,
National University of Oil and Gas «Gubkin University»
119991
vm@gubkin.ru

Abstract: The paper issues the ways of system engineering development in the oil industry in the context of the digital economy of Russia. Oil fields are considered as control objects. The architecture of stratified hierarchical overlay grid systems is based on the hierarchical description of oil fields. The paper deals with several issues of introduction of grid systems into commercial operation. The criterion of applicability of partially distributed grid calculations is defined. The basement of issued grid systems for oil production solutions is developed.

Keywords: hierarchical analysis, oil production, digital economy, grid.

1 Introduction

Complexity of the modern economic system is reinforced by the crisis and the danger of recession. This determines necessity of transformations of industrial production. Technological improvements offer new opportunities for industrial development and economic growth. Modernization of information systems of Russian oil and gas complex is influenced by digital economy development. Intensification of information resources is becoming one of the main factors of competitiveness. Implementation of the principles of digital economy in industrial development denotes integration of breakthrough digital technologies. This leads to transformation of production infrastructure and system management.

Oil production is one of the most important industries in Russia. Advanced Russian oil and gas companies implement the concept of digital oilfields. This concept includes a set of engineering and information solutions for effective management and monitoring of oilfield development process. Information technology catalyzes progress of oil and gas complex.

Existing information systems of oil and gas industry have several disadvantages:

- dissociation of subsystems which provide actions of various corporative spheres;
- poor usage of advanced information technology;
- insufficient technical support of information systems;
- accumulation of technical debt.

There is a significant disproportion in real oil production objects between low level of information technologies implementation and their modern capabilities. System analysis revealed several problems and directions of development where these disproportions are most obvious:

- operative monitoring of efficiency factors of oil production;
- commercial exploitation of third-party software solutions whose manufacturers have stopped their support;
- implementation of interagency cooperation to solve planning and production problems;
- complicated interaction with SCADA-systems;
- usage of full-scaled analytical models to solve problems of oil production;

- low fault tolerance of systems leads to frequent failures;
- low level of information security.

These factors determine necessity of building a modern multi-level engineering and information system for automating solutions of oil production problems. This problem is interdisciplinary and includes various scientific fields: system analysis, field development, information technology, economics. A distinctive feature of the problem is that the solution should realize not only the current production requirements of oil and gas complex, but also accompany whole life cycle of oilfields. Therefore, the solution should be based on the principles and standards of system engineering (ISO/IEC 15288:2015, ISO 15926-1:2004, ISO/IEC 12207:2008, ISO/IEC 16326).

System engineering is an interdisciplinary approach to integrate technical and management efforts in a way to transform production tasks, specifications and operating conditions into solutions to ensure full life cycle of industrial facility management. Thus, the breakthrough development of information technologies in oil and gas companies requires development of system engineering in the oil industry.

2 Conception of stratified hierarchical grid systems

Hierarchical description can be used to describe development objects, information systems, management objects, economic processes and their integration at hierarchical levels. Thus, a hierarchical description is applicable to describe systems in oil and gas complex. There are the following production strata:

- economic forces;
- information processing and management;
- physical processes, objects of physical impact.

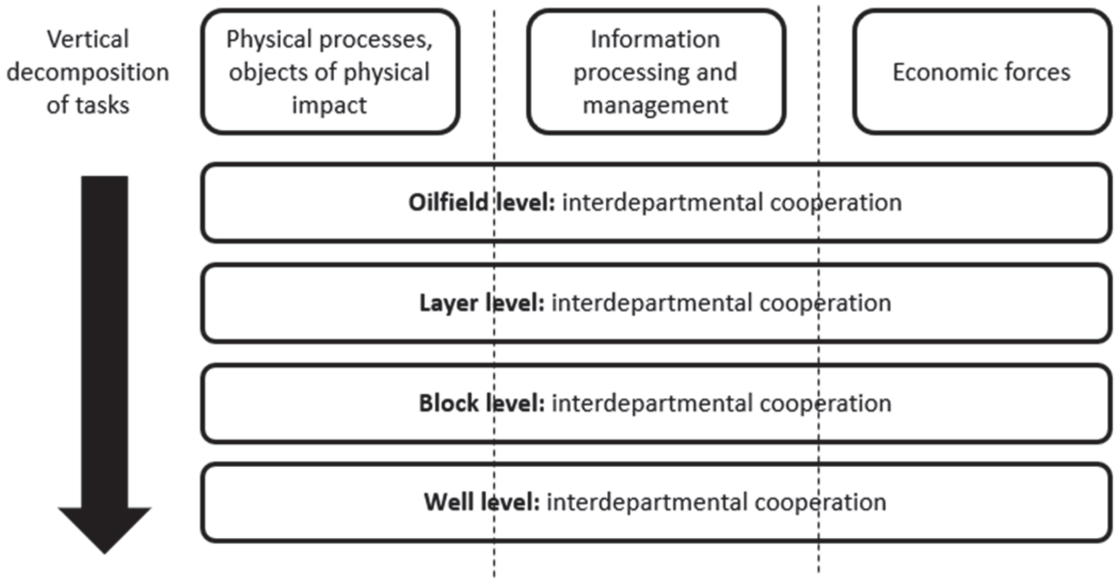

Figure 1 - Production strata and hierarchical levels in solving problems of block waterflooding of oilfields

Commercial exploitation of stratified hierarchical systems requires distributed hardware and software architecture. Conception of parallel computing has been developed in the middle of 20-th century. The first solutions were proposed in the 1960-th-1970-th: the model of collective computing (1962, E. V. Evreinov together with Yu. G. Kosarev), distributed computing in the local network PARC (1973, John Schoch and John Happ), distributed computing system "Minsk-32" (1977, head V. I. Zhiratkov). In 1967, Gene Amdahl formulated the law of limit acceleration of parallel computing:

$$S_p = \frac{1}{\alpha + \frac{1-\alpha}{p}}, \tag{1}$$

where

S_p - maximum speed-up of calculations in comparison with a single-processor system,

α - share of consecutive calculations in the total volume,

p - number of nodes involved in a multiprocessor system.

Development of modern information systems in the context of digital economy in Russia has a number of features: necessity for accelerated growth due to the lag of Russian digital economy from leading digital economies of the world, need for import substitution of software, financial constraints due to economic factors, high security requirements and fault tolerance due to the current competitive environment in the oil industry and geopolitical tensions.

Today there are several solutions for organization of high-performance systems: supercomputers, computing clusters and grid systems.

Supercomputers – specialized computing machines, which are complexes of high-performance servers united by a local high-speed data channel. Supercomputers have extremely high performance and compact resources. Implementation of supercomputers requires high financial costs for hardware: from 20-30 thousand dollars for basic models to several million for leading models. High centralization of supercomputer resources means a high risk of system failure. It also complicates the partial independent exploitation of resources which is important for building of complex distributed infrastructures.

Computing clusters are loosely coupled sets of several computing systems that work together to perform shared tasks. Clusters have a high total computing power and the lower (though rather high) cost compared to supercomputers. Dynamical load distribution among servers is available due to the less rigorous coupling between components of clusters. Taking into account a necessity to support high data rates between servers a cluster cannot be considered as a fully distributed system. Clusters often combine only the same type of computing devices. All these limits distributed usage of clusters in multi-layer information systems. Additional difficulties in working with clusters are associated with their lack of standardization and closed corporate developments that complicate their joint application.

Grid systems are dynamic (in composition and in structure) networks of different loosely coupled resources. Resources of grid systems are used to solve subtasks in the overall flow of information load. These grid systems provide a shared mode of access to resources across globally distributed virtual systems. Grid systems are based on standard and open protocols, services, interfaces and hardware and have relatively lower performance than supercomputers and computing clusters. Special software for combining all type of existing information resources of grid systems should be developed. This software is used for industrial exploitation of grid systems. This technological solution needs significantly lower financial investments in comparison with supercomputers and clusters. Grid systems have a high value of resources decentralization and flexibility of their integration into complex architectures. Existence of many decentralized loosely coupled resources ensures their extremely high fault tolerance. Grid system can be interpreted as geographically distributed infrastructure combined into a single computer network, or "mega cluster".

Due to the mentioned aspects, conception of stratified hierarchical overlay grid systems for development of information systems in the oil and gas complex was chosen.

The main components of described grid systems are:

- user interface for task management;
- cloud services for registration and decomposition of tasks and result output;
- queue servers and load balancers for distributed processing of decomposed tasks;
- data warehouse;
- distributed computing resources which work together to solve decomposed tasks.

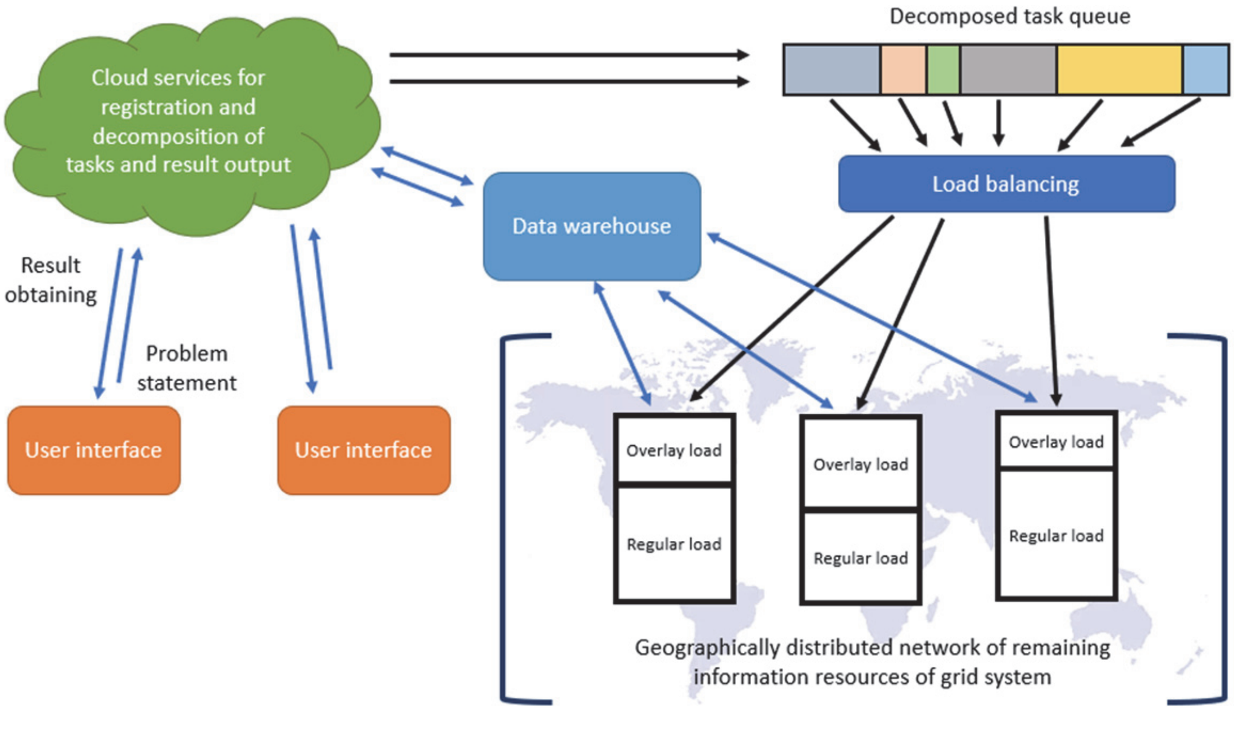

Figure 2 - Simple scheme of overlay grid system

Grid systems can include standard (and even outdated) devices: computers, databases, networks. Overlay grid systems use remaining computing resources of available hardware. All these facts offer conditions for budget introduction of breakthrough technologies in the context of development of digital economy (Fig. 2).

3 Design of stratified hierarchical grid system

There are several platforms for implementation of grid systems, including Globus Toolkit, UNICORE, and gLite. However, designing the architecture of grid systems for solving oil production problems stay an unresolved task. The grid system should have a hierarchical architecture according to overall hierarchical architecture of management which identifies developing oilfield. This architecture includes the following levels:

1. the level of oilfields;
2. the level of productive layers;
3. the level of waterflooding blocks (when contour waterflooding is implemented);
4. the level of wells.

Resources at each hierarchical level are combined into one information stratum, which is a network of distributed processing and storage devices (Fig. 3). Decomposed tasks from the upper hierarchical level set into shared queue of the lower level and are distributed between its devices through load balancing. Full failure of information stratum occurs only in case of failure of all located devices.

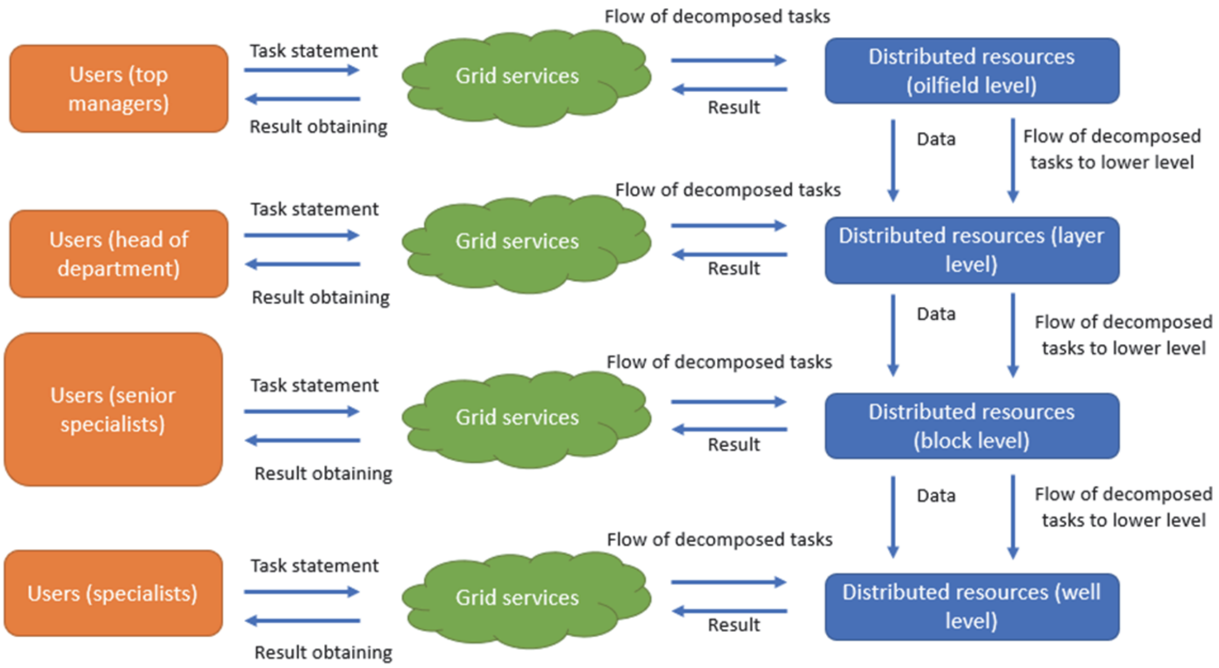

Figure 3 - General architecture of stratified hierarchical grid system

Methodology of construction of the grid systems is described with the two general aspects:

- development of a hierarchical information model (HIM) for decomposition and solution of oil production tasks;
- organization of grid system information resources according to HIM.

The main steps for developing HIM are: determination of production macro-tasks and decomposition of these macro-tasks at hierarchical levels.

The following steps are performed to organize grid system information resources according to HIM:

1. analysis of available information resources;
2. organization of data storage resources at hierarchical levels;
3. organization of resources for calculations at hierarchical levels;
4. integration of storage and calculation resources within levels;
5. realizing of resources interaction from different hierarchical levels.

Analysis of resources for overlay grid systems construction solve a number of problems:

- performance assessment of existing computing devices and their remaining resources for grid system tasks processing;
- assessment of network communication capacity;
- assessment of extensive growth capacity of information system.

Review of modern methods of hardware storage backup is carried out. Results of review used to solve the problem of storage resources organization. Implementation of permanent hardware reserving increases mean time between failures of reserved system according to the formula:

$$T_{avg\ r} = T_{av} \sum_{i=0}^{m} \frac{1}{(i+1)}, \tag{2}$$

where

$T_{avg\ r}$ – mean time between failures of reserved system,

T_{avg} – mean time between failures of non-reserved system,

m – reserving ratio.

Storage devices are organized into independent groups and distributed between hierarchical levels. Physical separation of storage devices will allow to realize different algorithms of resource balancing and data replication according to requirements of each hierarchical level. Data storage system at each hierarchical level will be a separated information space and will have its own hardware architecture. Thus, data storage systems will be stratified.

Distribution of computing resources for solving problems of oil production requires an assessment of resource intensity of proposed computational problems. Let grid calculations require execution time of one macro-task to be less than some target value $T_{tg}(sec)$. Also, let the resource intensity of each task be equivalent to a certain number of floating-point operations N_{eq}. Then, flow $Q \geq \frac{N_{eq}}{T_{tg}}$ should be carried out for successful execution of the task during $T < T_{tg}$ handling. Measurement in FLOPS, which equals to the number of floating-point operations per second, can be used to indicate value of Q. FLOPS measurement is widely used to determine the performance of information systems.

Evaluating the resource intensity of each task type should be provided. Then estimating the complexity of task groups of each hierarchical level becomes possible. Let N be the number of floating-point operations which equivalent to the computational complexity of the i-th hierarchical level tasks, then

$$N_{eq\ i} = \sum_j N_{eq\ ij}, \tag{3}$$

where $N_{eq\ ij}$ – equivalent of floating-point operations for the j-th task performed at the i-th hierarchical level.

Processing of all i-th level tasks within the target time should be guaranteed by enough summary performance of devices at the level:

$$Q_i \geq \frac{N_{eq\ i}}{T_{tg\ i}}, \tag{4}$$

where

$T_{tg\ i}$ – target processing time of the i-th hierarchical level tasks within the macro task execution,

Q_i – target performance of the i-th device group.

Distribution of remaining resources is performed according to resource intensive value of task groups at the system levels. Tasks are performed through progressing queues. So, the maximum performance of a computers group at the i-th hierarchical level can be described as

$$Q_{res\ i} = \sum_{j=1}^{M_i} Q_{res\ ij}, \tag{5}$$

where

$Q_{res\ i}$ – maximum performance of the i-th hierarchical computer group,

M_i – amount of i-th hierarchical level computers,

$Q_{res\ ij}$ – performance of the j-th computer at the i-th hierarchical level.

Probability of failure of a computer group greatly depends on devices amount:

$$P_i = 1 - \prod_{j=1}^{M_i} (1 - P_{ij}), \tag{6}$$

where

P_i – failure probability of computers group,

P_{ij} – failure probability of the j-th computer at the i-th hierarchical level.

Computational load is distributed among all devices in the formed groups. If any computer in a group will be lost task flow will be redistributed among active computers. A complete failure of computational group occurs when all its computers fail.

As a result, groups of computing resources are formed according to sets of tasks of hierarchical levels. This system of calculations is similar to the data storage system. The system can be extensively expanded. Also, every subsystem on different levels have its own dynamic balancing of resources. Therefore, each hierarchical level is a separate information space with its own hardware architecture.

All subsystems of the grid (data storage and calculations) should be integrated within hierarchical levels to successful solving problems of oil production. Efficiency factors of subsystems integration are:

- encapsulation of control tasks within hierarchical levels, it can be achieved through decomposition of original macro-tasks;
- balance of storage resources and resources for calculations at hierarchical levels;
- network speed.
- information security.

Integration of resources at hierarchical levels completes stratification process in the grid system.

Implementation of queues is assumed for interaction between stratified resources of the grid system (Fig. 4). Resources at one level should not make changes to information space at another level. This will preserve their integrity. Therefore, setting tasks between levels should be provided through a separated queue server (or a group of servers).

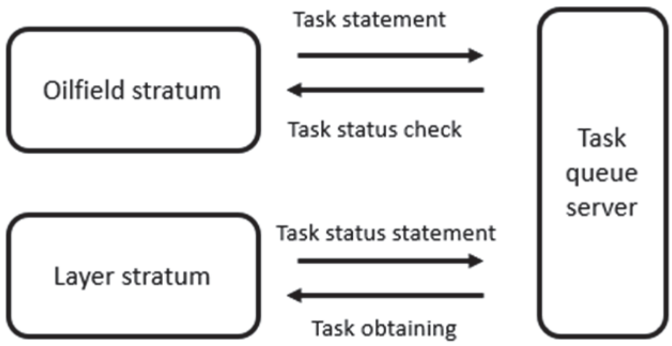

Figure 4 - Scheme of strata interaction through a queue server

A queue server controls information flows and provides active information protection against unauthorized access to the system.

Interaction between strata is necessary for computing operations with an expanded set of data (aggregation of data from lower levels, source data from higher levels). For this purpose, "read-only" access is implemented for requests from outer levels in storage devices (Fig. 5).

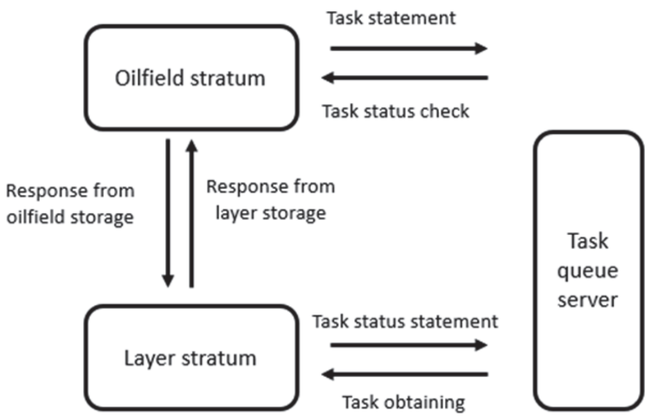

Figure 5 - Extended scheme of strata interaction

As a result, strata interaction is organized within safe expansion of information space of each stratum.

4 Test grid system for calculation of efficiency factors of oilfields development

Testing stand of the grid system was realized to check described conception of stratified hierarchical overlay systems for solving oil production problems. The problem of calculating efficiency indicators of contour waterflooding was solved within pilot exploitation of the system.

Waterflooding is one of the key processes carried out in oil production today. It is also one of the most common methods of influence on oil reservoirs to displace oil with water. Contour waterflooding is the most common type of waterflooding. It combines a number of modern methods. The general aim of contour waterflooding is a conditional separation of productive layers into individual blocks with subsequent localization of waterflooding processes.

According to the proposed methodology the main macro-tasks of the grid system were identified and decomposed:

- At the level of wells: calculation of fluid streamlines between wells based on the reservoir pressure map and Dupuit's formula of flow, calculation of efficiency indicators of waterflooding at the level.
- At the level of waterflooding blocks: identification of waterflooding blocks and calculation of participation rates of wells in blocks, calculation of efficiency indicators of waterflooding at the level.
- At the level of productive layers: analysis of fluid flows between blocks, calculation of efficiency indicators at the level.
- At the level of oilfields: final calculation of efficiency indicators for fields.

Table 1 - The main characteristics of test oilfields, their well stock and parameters of reservoir analytical models

Oilfield name	Layer name	Number of cells in models	Number of waterflooding blocks	Number of wells
Field 1	Layer 1	20000	42	205
	Layer 2	16250	35	158
	Layer 3	17550	36	180
Field 2	Layer 4	14762	28	148
	Layer 5	15867	30	155
Field 3	Layer 6	13800	23	137

The grid system was designed as hierarchical stratified overlay system including 4 hierarchical levels: oilfield, layer, waterflooding block, well.

Information resources of the grid system was based on 58 units with the following characteristics:

- 2.5 GHz Intel Core I5 7200 processor;
- 4GB RAM;
- 250GB SSD storage.

Average remaining resource of these devices in everyday usage was 39 %. 18 units was used as a storage warehouse and queue servers on different hierarchical levels. And 40 units were used for calculation performing.

Guided by the methodology of grid system resources distribution described in the article, an issue of calculation complexity of waterflooding efficiency factors of different strata was carried out. Distribution of tasks complexity between strata was provided. For this purpose, there were executed test calculation on a test device. Result of this distribution was described in seconds of processor time of the test device work:

- the level of oilfields – 10.1 seconds;
- the level of productive layers – 59.3 seconds;
- the level waterflooding blocks - 2343,5 seconds;
- the level of wells - 54226,5 seconds.

Average processor time per device of the whole system was 1 415,9 seconds.

At least at two devices were placed on each stratum to improve fault tolerance of its calculation resources. Because of this, there was an obvious excess of hardware resources on the two first hierarchical levels.

The final distribution of calculation resources between strata was as follows:

- the level of oilfields – 2 devices;
- the level of productive layers – 2 devices;
- the level waterflooding blocks – 2 devices;
- the level of wells - 34 devices.

At least at 1 backup device has been placed at each hierarchical level to provide enhanced fault tolerance for data storage. As a result, the following architecture was proposed:

- 2 queue servers with Memcached and PostgreSQL were placed at oilfield hierarchical level for interaction providing between oilfields and layers strata.
- 2 queue servers with Memcached and PostgreSQL were placed at layer hierarchical level for interaction providing between layers and blocks strata.
- 2 queue servers with Memcached and PostgreSQL were placed at blocks hierarchical level for interaction providing between blocks and wells strata.
- 2 storage servers for waterflooding data based on PostgreSQL and MongoDB was placed at oilfields hierarchical level;
- 2 storage servers for waterflooding data based on PostgreSQL and MongoDB was placed at layers hierarchical level;
- 3 storage servers for waterflooding data based on PostgreSQL and MongoDB was placed at blocks hierarchical level;

- 5 storage servers for waterflooding data based on PostgreSQL and MongoDB was placed at wells hierarchical level.

Target time to perform all the complex calculations (T_{tg}) was 30 minutes (1800 seconds). And actual time spent on the calculations was 1 820 seconds (there were no failures in the system operation).

The most productive device with the maximum remaining resource was chosen to analyze the performance increase. Programming code of software has been optimized to run calculations on a single device: decomposition of tasks and interaction with queue servers have been excluded, data processing was optimized. As a result, the execution time of computing took 41 192 seconds. The speed of calculations in the grid system exceeded the speed of calculations on one device by 22.6 times.

However, the grid system architecture should be corrected to perform the calculations for a time less than T_{cr}: new devices should be added to well and block strata or devices can be obtained from oilfield and layer strata with subsequent increase in the risk of system failure.

A web interface was developed to monitor status of the grid system. The interface can show dynamic of tasks completion, strata status and top performance devices.

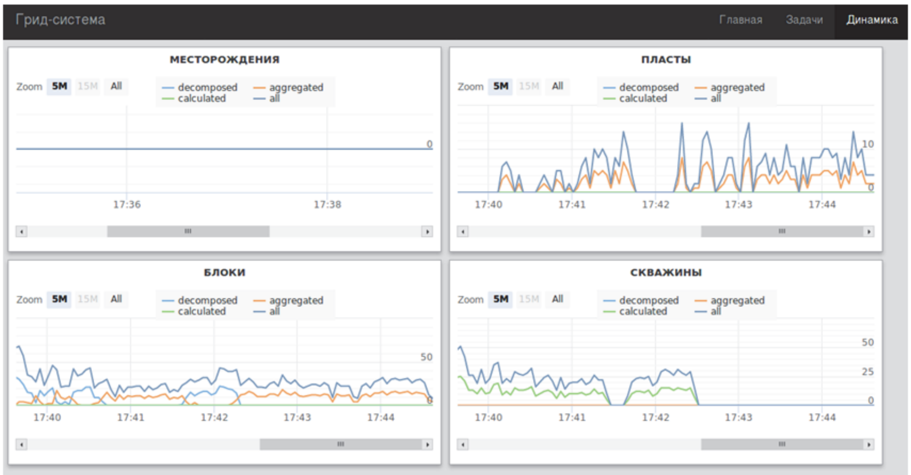

Figure 6 - Web interface for test stand of grid system for solving problems of contour waterflooding

5 Summary

The set of scientific results obtained in the paper can be a basis for solving problems of modernization of information systems in oil production industry. These results can be implemented at enterprises of the oil and gas complex of Russian Federation in the context of the digital economy. Conception of implementation of stratified hierarchical overlay grid systems as a guide for development of system engineering in this industry is proposed. These grid systems are able to provide a significant increase in productivity and fault tolerance at low financial costs and in a short time interval. This approach has a potential for breakthrough development of information technologies in Russian oil and gas complex.

The following results are obtained:

- Conception of stratified hierarchical overlay grid system for modernization of information systems of the oil and gas complex is proposed.
- The methodology of designing an architecture of grid systems for solving oil production problems is developed.
- Software implementation of the grid system for solving computational problems of oil production is realized. The functionality of grid systems: decomposition and task allocation in hierarchical levels, calculation of efficiency indicators of contour waterflooding on hierarchical levels, resources performance monitoring.
- Test exploitation of the developed grid system is carried out. Obtained results confirm the effectiveness of this system for solving problems of oil production.

List of references

[1] A. S. Tanenbaum, Maarten van Steen. Distributed systems. Principles and paradigms. Piter, 2003.
[2] L. A. Muravieva-Vitkovskaya. Osnovy raspredelennogo modelirovaniya. ITMO University, 2013.
[3] S. V. Vostokin. Graficheskaya ob"ektnaya model' parallel'nyh processov i ee primenenie v zadachah chislennogo modelirovaniya, SCN RAS, 2007.
[4] K. G. Skripkin. Ekonomicheskaya effektivnost' informacionnyh sistem. DMK Press, 2002.
[5] A. Borzenko. Servernye platformy Intel. Bestsellery IT-rynka, 4, April 2002.
[6] D. Schilyak. Decentralizovannoe upravlenie slozhnymi sistemami. Mir, 1994.
[7] B. Scott. Скотт Б. Razrabotka pravil informacionnoj bezopasnosti. Williams, 2002.

[8] I. Foster. What is the Grid? A Three Point Checklist. Argonne National Laboratory & University of Chicago, 2002.
[9] O. V. Sukhoruslov. Integraciya vychislitel'nyh prilozhenij i raspredelennyh resursov na baze oblachnoj programmnoj platformy. Programmnye sistemy: teoriya i prilozheniya, 4(22):171–182, 2014.
[10] I. Foster, C. Kesselman. The Grid: Blueprint for a New Computing Infrastructure. Morgan Kaufmann Publishers, 1998.
[11] A. Kh. Mirzadzhanzade, M. M. Hasanov, R. N. Bakhtizin. Etyudy o modelirovanii slozhnyh sistem neftedobychi. Neftegazovoe delo, 2009.
[12] E. Dubrova. Fault-Tolerant Design, Springer, 2013.
[13] Ch. Bell, M. Kindahl, L. Thalmann. MySQL High Availability. O`Reilly Media, 2010.
[14] D. G. Leonov. Postroenie geterogennyh raspredelennyh programmno-vychislitel'nyh kompleksov na osnove otkrytoj integracionnoj platformy. Proceedings of Russian State University of Oil and Gas, 2:125-135, 2017.
[15] V. V. Borduzhe, L. I. Grigorev, A. I. Kostogryzov, A. A. Nistratov. Importozameshchenie programmnogo obespecheniya v neftegazovom komplekse na principah sistemnoj inzhenerii. Cel', problemy i strategiya resheniya problem. Quality Management in Oil and Gas Industry, 4:121-137,2015.
[16] V. V. Borduzhe, L. I. Grigoryev, A. I. Kostogryzov, A. A. Nistratov. Importozameshchenie programmnogo obespecheniya v neftegazovom komplekse na principah sistemnoj inzhenerii. Cel', problemy i strategiya resheniya problem (Part 2). Quality Management in Oil and Gas Industry, 1:20-26,2016.
[17] A. I. Kostogryzov, O.I. Atakishchev, A. A. Nistratov, G. A. Nistratov, S. Klimov, L. I. Grigorev. The method of rational dispatching a sequence of heterogeneous repair works. Energetika, 4:154-162,2017.
[18] I. S. Arsenevskiy. Avtomatizirovannoe postroenie blokov v upravlenii vnutrikonturnym zavodneniem mestorozhdeniya nefti na osnove ierarhicheskogo analiza. Avtomatizaciya, telemekhanizaciya i svyaz' v neftyanoj promyshlennosti, 8:9-16, 2017.
[19] I. S. Arsenevskiy. Metodologiya postroeniya stratificirovannoj grid-sistemy po analizu effektivnosti i planirovaniyu vnutrikonturnogo zavodneniya neftyanyh mestorozhdenij. Avtomatizaciya, telemekhanizaciya i svyaz' v neftyanoj promyshlennosti, 6:21-31, 2018.
[20] Yu. A. Kovalchuk, I. M. Stepnov. Innovacii v menedzhmente. Cifrovaya ekonomika: transformaciya promyshlennyh predpriyatij, 11:32-43,2017.
[21] S. Abramov. Superkomp'yutery: obratnye rekordy. Nauka i zhizn', 1:42-45, 2019.
[22] V. A. Levin, V. V. Kalinin, E. V. Rybalka. Elementy linejnoj algebry i analiticheskoj geometrii na baze paketa «Mathematica». Physmathlit, 2007.
[23] A. A. Chien, S. Marlin, S. T. Elbert. Resource Management in the Entropia System. Kluwer Academic Publishers, 2004.
[24] V. V. Voevodin. Reshenie bol'shih zadach v raspredelennyh vychislitel'nyh sredah. Avtomatika i telemekhanika, 5:32-45, 2007.
[25] A. P. Demichev. V. A. Ilyin, A. P. Krukov. Vvedenie v grid-tekhnologii. NIIYAF MGU, 2007.
[26] I. Foster, G. Kesselman, S. Tuecke. The Anatomy of the Grid: Enabling Scalable Virtual Organizations. International Journal of High Performance Computing Applications, 3:200-222, 2001.
[27] V. N. Kovalenko, D. A. Koryagin. Organizaciya resursov GRID. IPM im. Keldysha RAS, 2004.