

М.Серік, А.Е.Байғараева

Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Астана (E-mail: Serik_m@enu.kz)

Параллель есептеулер кластерін баптау

Кластер жоғары жылдамдықты байланыс арналарымен өзара біріккен компьютерлер (түйіндер) жиыны болып табылады, ал арнайы бағдарламаның қамтамасыз етілуі түйіндердің параллельді жұмысын ұйымдастыруға мүмкіндік береді. Қазіргі қуатты кластерлер бірнеше мыңдаған процессорлық элементтерден тұрады. Дегенмен қазіргі таңда кластер жүйелерін қолдану тиімділігі, бағдарламамен қамтамасыз етілуімен қатар (тілдері, технологиясы және параллель бағдарламалаудың аспапты ортасы), параллель есептеуді ұйымдастыру мен олардың тиімді орындалуы сияқты шешілмеген мәселелерге келіп тіреледі. Сондықтан кластерде параллель есептеулерді тиімді ұйымдастыру, олардың шекті мүмкіндіктерін мен нақты күрделі есепте осал жерлерін анықтау мәселесі өзекті болып тұр.

Кілт сөздер: кластер, супер компьютерлер, бұйрықтар, параллель есептеулер.

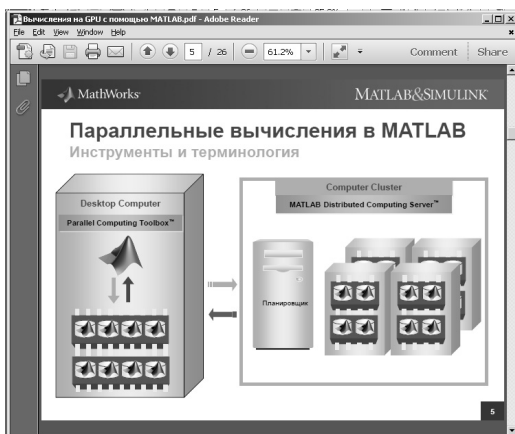
Кластер — біртұтас есептеуіш ресурс ретінде жұмыс істейтін жергілікті есептеуіш желіде біріктірілген компьютерлер тобы. Кластер тәсілдемесінің негізгі қасиеттері қолжетімділік және жеңіл кеңейтілімдік болып табылады.

Кластермен басқару жүйелеріне келесі негізін қалайтын талаптарды көрсетуге болады:

- «MATLAB Distributed Computing Server» қызметін әр компьютерге орнату;
- Windows операциялық жүйесінің ағылшын нұсқасы, яғни жұмыс кезінде сөздер тобымен қиындықтар туындамайды [1];
- бір типті компьютерлер, себебі кластер біртекті емес құрылымға ие болады да, мынадай жағдайлар ескерілуі керек болады:
- орталық процессордың қуаты, оперативтік жақтың көлемі, жергілікті желі бөлімшелерінің жылдамдығы;
- өнімділік, яғни өткізгіштік қабілеті;
- қолайлылық, яғни кластермен басқару жүйесі икемді және жақсы күйге келтірілген, сондай-ақ қолдануда қарапайым болуы қажет;
- сенімділік, яғни аппаратты платформа және бағдарламалық қамтамасыз ету жедел жұмыс жасауы қажет, істен шыққан жағдайда кластер тапсырмалардың орындалуын жалғастыруы қажет.

Егер орындалу тоқтаса, онда жүйеаралық ахуалды қалпына келтіруді басынан емес, тоқтаған жерінен жалғастыруы керек.

Кластер қарапайым жергілікті есептеуіш желіге қарағанда жоғарылау сенімділік пен тиімділікті жобалайды. Кластер типтік аппараттық және бағдарламалық шешімдерді қолданады, сондықтан параллель есептеуіш жүйелердің басқа осындай түрлерімен салыстырғанда айтарлықтай арзандау болады.



1-сурет. MATLAB ортасында параллельді есептеудің құралдары

Параллельді компьютерлер — бұл процессорлар тобы және жадтар мен олардың арасындағы коммуникацияның кейбір әдістері (1-сур.). Ол екі ядролы процессор болуы мүмкін, көп процессорлық сервер немесе, мысалы, кластер (суперкомпьютер) болуы мүмкін.

Дәстүр бойынша бағдарлама тізбекті алгоритмді түрде жазылды. Тізбекті бағдарлама бір орталық процессоры бар компьютерде орындалды. Бұйрықтар бағдарламада ретімен орындалды. Қазіргі бағдарламаларды орындауда мультипроцессорлы есептеулер қолданылады. Бағдарлама бір мезгілде орындалатын бөліктерге бөлінеді және әр түрлі процессорларда есептеледі. Ресурстар бір компьютерде немесе желіге қосылған бірнеше компьютерлерде тұруы мүмкін.

Параллельді бағдарламалау процессорларды көп ядролық архитектураларға ауысуы жағдайында өзекті болады. Осыған орай MathWork, Inc және басқалар өз пакеттеріне параллель және таралған есептеулерді енгізе бастады [2].

Параллельдеу — іздеу, операциялар нұсқауын немесе бір мезгілде орындалатын тізбекті алгоритм операцияларының жиындарының орындалу процесі.

Біз MATLAB жүйесінде пайдаланылатын кейбір қажетті ұғымдарды енгіземіз:

- *worker* (жұмыс процесі) — есептеуді орындайтын MATLAB жүйесіндегі процесс;
- *mdce* (MATLAB Distributed Computing Server) — *jobmanager* және *worker*-процесі қосылғанға дейін барлық машиналарды іске қосатын қызмет. Ол барлық процестерді іске қосу үшін негіз болатын орта болып табылады [3, 4];
- *jobmanager* — ішкі есептерді жоғары ретпен толық түрде басқару;
- *node* — кластер бөлігі болып табылатын компьютер.

Енді біз оқу процесінде қолдану үшін қолжетімді кластер құрамыз.

Параллель есептеулер кластерін баптау аппаратына қойылатын талаптар

Біздің жағдайда мынадай екі ядролы компьютер алынды: Pentium (R) Dual-Core CPU 2.8 GHz және Pentium(R) Dual-Core CPU 3.5 GHz. Әр компьютердің оперативтік жадысы 2 Gb тұрады.

Бұл екі компьютер төрт ядродан (екі-екіден) тұратын болды. Әр ядрода үш *worker* (жұмысшы) іске қосуға болатындықтан, 12 *worker*лер (жұмысшылар) құруға мүмкіндік аламыз, бұл оқу мақсаттарындағы параллель есептеулер үшін құрылған кластерді пайдалануға мүмкіндік береді.

Бағдарламалық қамтамасыз етуге қойылатын талаптар

1. Әр компьютерде MATLAB R2011b орнату және іске қосу.
2. Әр компьютерде MATLAB Distributed Computing Server-ді іске қосу.

Мынадай тапсырма орындайық:

Екі компьютерден тұратын кластерлерді баптауды қарастырайық, *mdce* баптау үшін *cmd line help* утилиттерін қарастырамыз: *mdce.bat*, *start jobmanager.bat*, *start worker.bat*, *nodestatus.bat*.

1. MATLABR2011b ортасын іске қосамыз.
2. Барлық компьютерлерде *mdce* баптауын орнатып, іске қосамыз.

Бұл үшін екі нұсқаны қолдануға болады:

1) Windows бұйрық жолы көмегімен.

1.1 cd “D:\Program Files\MATLAB\R2011b\toolbox\distcomp\bin” каталогына өтеміз.

1.2 Бұйрықты іске қосамыз:

– *mdce install*.

– *mdce start*.

1) MATLAB жүйесінің Command Window көмегімен:

Current Folder сол терезесінде келесі жолды ашамыз:

D:\Program Files\MATLAB\R2011b\toolbox\distcomp\bin

Command Window терезесінде бұйрықты іске қосамыз:

! *mdceinstall*

! *mdcestart*

«!» таңбасы жүйелік бұйрықтың іске қосылғанын білдіреді. MATLAB ортасында параллель есептеулерді басқару үшін жоспарлаушыны (*jobmanager*) іске қосамыз:

!*start jobmanager -name jm -v*

– *v* — бұйрық терезесінде іске қосудың процесінің толық бейнесін білдіреді.

- Тапсырмаларды жоспарлаушы деген;
- басқару жүйесінің орталық компоненті болып табылады;
- процессордың қажетті ядролары және есептеуіш түйіндердің қолдануын нақтылайды;
- тапсырмалардың орындауға қойылуын қадағалайды;
- жүйелік ресурстарды пайдалану сервисі;
- тапсырмалардың іске қосылуын қадағалайды;
- тапсырмалардың арасында ресурсты үлестіреді;
- тапсырмалар кезегін сүйемелдеу үшін жауап беретін сервис;
- тапсырмалар және есептеуіш түйіндердің жағдайын зерттеп отырады;
- істен шығу жағдайында қалпына келуді қамтамасыз етеді.

Ары қарай цикл көмегімен қойылған тапсырмаларды параллель орындайтын *worker*лерді іске қосамыз. Бұйрықтарды енгізу және келесі жолға өту үшін *shift+enter*, ал кодтың енгізуін аяқтау үшін *enter* пернелер комбинациясына басылсын.

Кластерді екі сервермен ұйымдастыру үшін мына бұйрықтар орындалады:

```
clientHost = 'Server1';
node = {'Server1','Server2'};
for i = 1:length(node)
for j=1:2
str = ['!startworker -name w_' num2str(j) '_' node{i} ' -jobmanagerhost ' clientHost ' -jobmanager jm -
remotehost ' node{i} ' -v'];
eval(str)
end
end
```

Host = 'Server1' — бұл жоспарлаушыны іске қосқан компьютер. Кластердің статусын бұйрықтың көмегімен тексеруге болады:

!nodestatus.bat –infolevel 3;

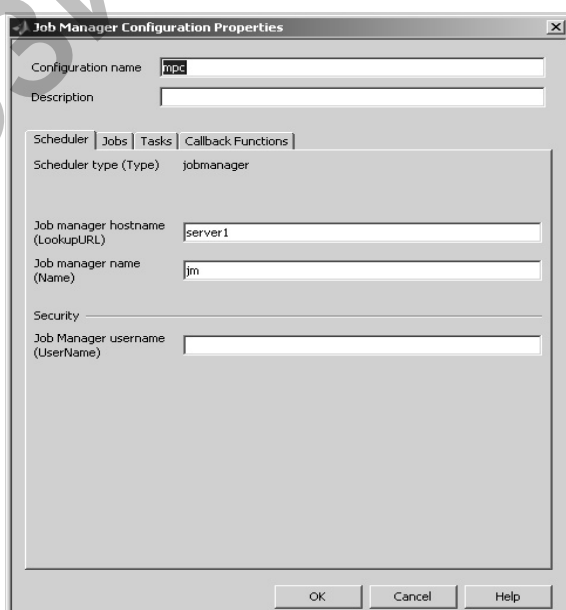
Infolevel – ақпарат алу деңгейі;

3 — «нодтар» туралы толық ақпаратты алу, егер 1 қойсақ, онда аз мөлшерде ақпарат аламыз.

Барлық процестерді іске қосу үшін сәтті баптаудан кейін, конфигурацияны баптауға кірісеміз (тек орталық серверде). Бұл үшін бұйрықтар тізбегін орындаймыз:

Parallel – Manage Configurations... – File – New – jobmanager.

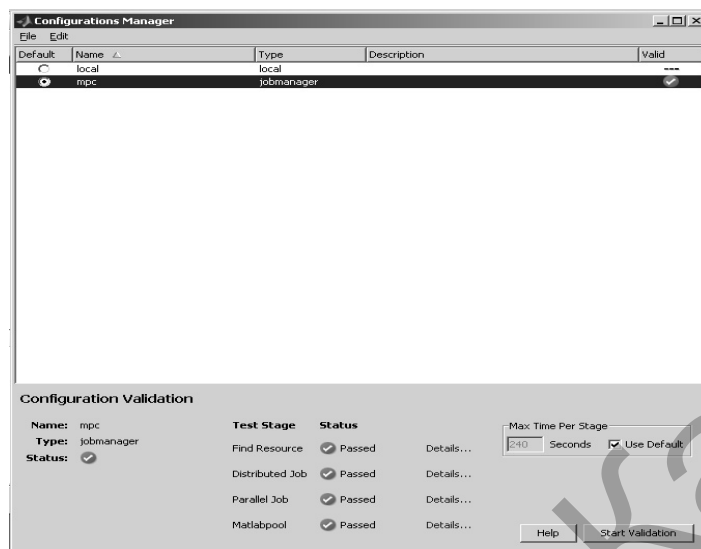
Осыдан соң құрылған *jobmanager* үшін қасиет таңдаймыз және 2-суреттегідей көрсетеміз:



2-сурет. Jobmanager терезесі

Конфигурация параметрлерін сақтаймыз және Start Validation батырмасы көмегімен тексереміз, ол үшін келесі бұйрықтарды орындау керек (3-сур.):

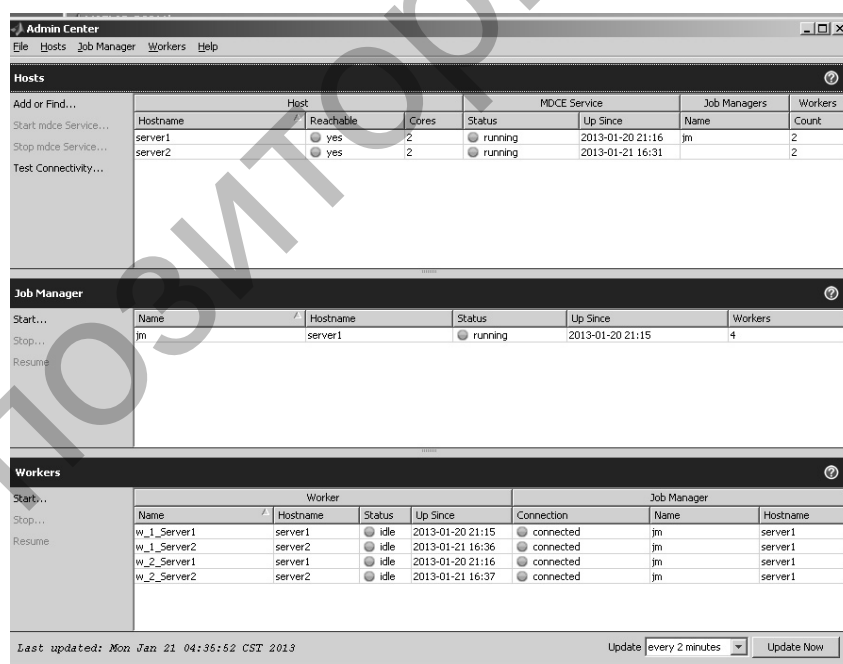
Parallel – Manage Configurations – mpc – контексті мәзір – Start Validation.



3-сурет. Start Validation бұйрығының орындалу нәтижесі

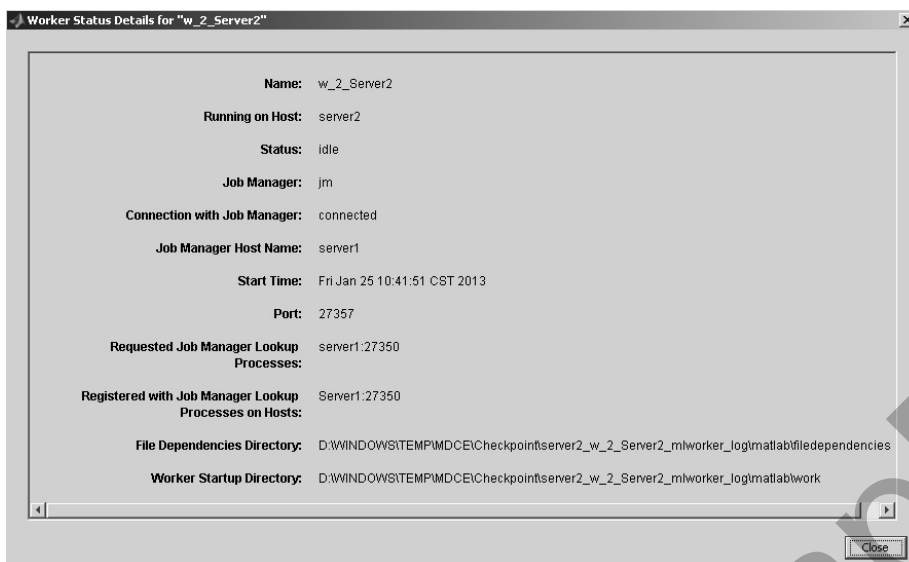
Кластердің параметрлерін қарап шығу және басқару үшін мына бұйрықты пайдаланамыз: !admincenter.bat

Келесі терезе пайда болады (4-сур.)



4-сурет. AdminCenter терезесі

Ұсынылған деректерді басқаруға болады, ол үшін тиісті жолда контекстік мәзір арқылы тиісті бұйрықты таңдау керек. Мысалы, w_2_Server2-де оң батырманы шертіп және пайда болған контекстік мәзірден Properties таңдап, келесі ақпаратты аламыз (5-сур.):



5-сурет. WorkerStatusDetailsfor “w_2_Server2” терезесі

Properties-тен басқа пайдаланылатын контекстік мәзірде тармақтар қатысады:

- Stop...
- Resume
- Destroy...
- Restart...

Мысалы, local1 конфигурациясының саны 8 workers-ке тең болғанда Admin Center келесі түрде болады (6-сур.):



6-сурет. 8-ге тең worker-лер санымен және таңдалған конфигурациясымен Admin Center терезесі кескінделген

Әдебиеттер тізімі

- 1 Специальные группы Windows и их перевод на русский язык. [ЭР]. Режим доступа: http://kaktusenok.blogspot.com/2012/08/windows_16.html
- 2 Кульгавая Е.А. Параллельные вычисления в MATLAB. — Минск: Белорус. гос. ун-т, 2010.

3 Гегель В.П., Стронгин Р.Г. Основы параллельных вычислений для многопроцессорных вычислительных систем. Нижний Новгород, 2003.

4 Оленев Н.Н., Печенкин Р.В., Чернецов А.М. Параллельное программирование в MATLAB и его приложения. — М.: ВЦ РАН, 2007. — 120 с.

М.Серик, А.Е.Байгараева

Настройка кластера параллельных вычислений

Кластер представляет собой множество компьютеров (узлов), соединенных между собой высокоскоростными каналами связи, а специальное программное обеспечение позволяет организовать параллельную работу узлов. Современные мощные кластеры уже содержат несколько сотен тысяч процессорных элементов. Однако эффективность применения кластерных систем в настоящее время наталкивается на целый ряд нерешённых проблем, касающихся как их программного обеспечения (языков, технологий и инструментальных сред параллельного программирования), так и собственно организации параллельных вычислений и эффективности их выполнения. Поэтому проблема комплексного анализа эффективности организации параллельных вычислений на кластерах, определения их предельных возможностей и их узких мест на реальных сложных задачах является актуальной.

M.Serik, A.E.Baygaraeva

Configuring a cluster of parallel computing

A cluster is a set of computers (nodes) connected by a high-speed links, and special software allows you to organize parallel operation nodes. Modern powerful clusters already contain hundreds of thousands of processing elements. However, the effectiveness of cluster systems currently encountering a number of unresolved issues related to their software (languages, technologies and tools for parallel programming environments), and the actual organization of parallel computation and the effectiveness of their implementation. Therefore, the problem of complex analysis of the effectiveness of parallel computing on clusters, defining their limits and possibilities of bottlenecks on the real challenge is urgent.

References

- 1 *Special Windows group and their translation into Russian*, [ER]. Access mode: http://kaktusenok.blogspot.com/2012/08/windows_16.html
- 2 Kulgavaya E.A. *Parallel computing in MATLAB*, Minsk: Belarusian State University, 2010.
- 3 Gergel V.P., Strongin R.G. *Introduction to Parallel Programming for multiprocessor computing systems*, Nizhny Novgorod, 2003.
- 4 Olenev N.N., Petchenkin R.V., Chernetcov A.M., *Parallel programming in MATLAB and its applications*, Moscow: CCAS, 2007, 120 p.