# Паралельне програмування з використанням бібліотеки TPL та PLINQ

# Лекція 7. Переваги і недоліки паралелізму

## 1. Переваги і недоліки паралельних обчислень

Паралелізм в програмах - це спосіб управляти розподілом спільних ресурсів, які використовуються одночасно. Внесення паралелізму до програм дозволяє витрачати системні ресурси найбільш ефективним чином.

Основна складність при проектуванні паралельних програм забезпечити правильну послідовність взаємодій між різними обчислювальними процесами, координацію a також ресурсів, ШО розділяються між процесами.

У цілому, головна мета паралельного програмування - прискорити виконання програми. Проте неправильне застосування методів паралельного програмування може призвести до зниження продуктивності, непередбачуваної поведінки програми і помилок, які важко усунути. Для того, щоб не такого сталося, програміст повинен розуміти механізми роботи потоків, як на апаратному, так і програмному рівні. Щоб зрозуміти, як реалізувати паралельну обробку у своїй програмі, необхідно знати:

- підхід до розробки і структуру програми;
- потоковий прикладний програмний інтерфейс (Application Programming Interface, API);
  - компілятор або середовище виконання програми;
  - цільові платформи, на яких працюватиме програма.

Виходячи з перерахованих позицій, важливо сформулювати стратегію паралельної обробки для конкретних частин вашої програми: що і коли розпаралелювати.

## Що розпаралелювати

Звичайно є сенс розпаралелювати такі конструкції як великі цикли, великі обсяги обробки об'єктів (масиви), велики SQL-запити, фонове введення-виведення.

## Колі є сенс розпаралелювати задачі

Розглянемо загальні міркування щодо застосування паралелізму у програмі з використанням засобів .Net.

Іноді при використанні паралелізму, можливі ситуації коли із збільшенням кількості задіяних процесорів не вдається добитися очікуваного зменшення часу виконання - в цьому випадку говорять про низьку продуктивність паралельної програми.

Причини таких ситуацій можуть бути різні:

- завдання (алгоритм) погано піддається розпаралелюванню;
- неправильне використання методів бібліотеки для розпаралелювання програми та ін.

Існує ряд загальних ситуацій, які призводять до низької ефективності паралельних програм, зокрема, що використовують методи паралельного

програмування .Net. Нижче розглянуто короткий опис деяких з таких ситуацій.

#### 1. Достатній об' єм обчислювальної роботи

Однією з основних вимог, яким повинна задовольняти програма для того, щоб була можливість її ефективного розпаралелювання, є наявність достатнього об' єму роботи, який вона повинна виконати. Припустимо, що тільки половина усіх обчислень в програмі може бути виконана паралельно, тоді згідно із законом Амдала множник продуктивності буде не більше двох. Це правило цілком очевидне, якщо уявити собі, що програма 90% часу виконання знаходиться в очікуванні відповіді на SQL -запит - паралельне виконання решти 10% коду не принесе істотного ефекту. У той самий час варто зазначити, що в цьому випадку ми можемо паралельно посилати SQL -запити різним серверам і використовувати асинхронну взаємодію з серверами.

Таким чином, для того, щоб отримати ефект від паралельного виконання завдань, треба мати достатній об'єм роботи для розпаралелювання, що як мінімум перекриває витрати самого процесу розпаралелювання.

# 2. Гранулярність завдань (розмір окремих завдань і їх загальна кількість)

Паралельна програма, що вирішує загальне завдання, зазвичай складається з окремих фрагментів - підзадач, які визначає в коді сам програміст. Якщо кількість підзадач буде занадто великою, то більшість з них простоюватимуть в черзі до ядер процесора, а об' єм витрат на запуск підзадач буде зростати. Якщо кількість підзадач буде занадто малою, то деякі ядра процесора будуть простоювати, знижуючи загальну продуктивність системи.

Деякі компоненти Parallel Extensions API, такі як, наприклад, Parallel.For() та PLINQ, здатні самі визначити відповідний для цієї системи об' єм і кількість завдань, що паралельно виконуються, тоді як при роботі з такими компонентами як TPL ( паралелізм задач - Task) відповідальність за ухвалення правильного рішення лежить на самому програмісті.

#### 3. Балансування навантаження

Навіть у тому випадку, коли гранулярність завдань підібрана вірно, виникає проблема розподілу цих завдань по ядрах (або по окремих процесорах). Ця проблема пов'язана з тим, що потенційно різні завдання можуть виконуватися різний час і, отже, в процесі виконання програми без балансування навантаження ядер можливі простої деяких ядер і простої завдань в чергах до інших ядер.

Так, наприклад, якщо Parallel.For() буде припускати, що усі ітерації розпаралелювання циклу виконуються однаковий час, то ми можемо просто розбити простір ітерацій на блоки по числу доступних ядер і паралельно виконати ці блоки. Проте насправді тривалість кожної ітерації може бути

різною, що веде до зниження продуктивності у відсутності балансування. Насправді, метод Parallel.For () реалізує автоматичне балансування, яке у багатьох випадках вирішує цю проблему.

## 4. Виділення пам' яті і збирання сміття

Деякі програми характеризуються інтенсивною взаємодією з пам'яттю, що викликає значні часові витрати як на виділення пам' яті, так і на збирання сміття. Зокрема, програма, яка обробляє текст, інтенсивно використовуватиме пам'ять, особливо, якщо програміст не звернув належної уваги на непрямі виділення пам' яті.

На жаль, виділення пам' яті на сучасній обчислювальній архітектурі це операція, яка може вимагати синхронізації між обчислювальними потоками, в яких виконується ця операція. Як мінімум ми маємо бути упевнені, що області пам' яті, що виділяються паралельно, не перекриваються.

Серйозні втрати часу виникають при збиранні сміття. Якщо збирач сміття знаходиться в постійній активності при виконанні програми, то це може стати вузьким місцем при її розпаралелюванні. Зрозуміло, збирач сміття в .NET може виконувати свою роботу паралельно з іншими завданнями програми, але для цього можуть знадобитися додаткові зусилля з боку програміста.

## 5. Кеш-промахи

Ця ситуація пов'язана з принципами кешування в сучасних комп'ютерах. Коли процесор робить вибірку значення з основної пам' яті, копія цього значення потрапляє в кеш процесора, що прискорює доступ до цього значення, якщо, звичайно, воно знадобиться в наступних обчисленнях. Насправді, процесор кэшує не лише одне вибране значення, але і деяку околицю цього значення в пам' яті. Значення, які переміщуються з основної пам' яті в кеш і зберігаються в ньому, називаються рядками кеша, і їх типовий розмір складає 64 або 128 байт.

Однією з проблем, пов'язаних з кешуванням на багатоядерних комп'ютерах, є ситуація при якій одне з ядер записує значення в певну область пам' яті, що знаходиться в кеші іншого ядра. У цьому випадку інше ядро при зверненні до відповідного рядка кеша отримає ситуацію промаху і буде вимушене відновити цей рядок з основної пам' яті. З деякою вірогідністю може скластися ситуація, при якій одне ядро постійно пише в основну пам'ять, порушуючи тим самим цілісність кеша іншого ядра, яке буде вимушене постійно оновлювати свій кеш, що призведе до різкого падіння продуктивності.

Тонша проблема виникає, коли ядра записують значення в різні області пам' яті, які, проте, розташовані так близько, що знаходяться в одному рядку кеша. Незважаючи на уявну малу вірогідність такої ситуації, вона досить часто зустрічається на практиці, оскільки, наприклад, поля об'єкту або проміжні результати обробки масивів об'єктів часто знаходяться в пам' яті відносно близько.

Є декілька способів уникнути подібних проблем. Можна, наприклад, проектувати структури даних спеціальним чином, знижуючи вірогідність появи проблем кешування, або виділяти пам'ять в різних потоках.

Крім того, проектуючи процеси паралельної обробки необхідно вже на алгоритмічному рівні зберігати локальність оброблюваних даних відносно потоку. Так якщо, наприклад, стоїть завдання обробити масив чисел, то очевидно, що локальність обробки кожним ядром своєї половини масиву буде набагато вища, ніж локальність обробки одним ядром елементів масиву з парними індексами, а іншим - з непарними. У останньому випадку тільки половина елементів рядка кеша ядра буде містити потрібні цьому ядру елементи.

#### 2. Закон Амдала для оцінки продуктивності

Найбільша складність в ефективному використанні багатоядерних процесорів полягає в законі Амдала, який стверджує, що максимальне підвищення ефективності за рахунок розпаралелювання обмежене кодом, який повинен обчислюватися послідовно. Наприклад, якщо тільки дві третини часу виконання алгоритму може бути розпаралелено, ви ніколи не отримаєте приріст продуктивності у 3 рази, навіть при нескінченній кількості ядер.

Отже, перш ніж продовжувати, варто перевірити, чи  $\epsilon$  код, який ми хочемо розпаралелити, вузьким місцем. Крім того, варто подумати про те, чи повинен код виконувати таку кількість обчислень; оптимізація зазвичай  $\epsilon$  найпростішим і найефективнішим підходом. Втім, тут може знадобитися компроміс, оскільки використання деякої техніки оптимізації може ускладнити розпаралелювання коду.

Найпростіше отримати вигоду при рішенні так званих завдань надзвичайного паралелізму (embarrassingly parallel problems) - коли уся задача може бути легко розбита на підзадачі, які можуть виконуватися незалежно. Прикладами можуть служити задачі обробки зображень, трасування променів, чисельні методи в математиці або криптографії.

#### Поняття продуктивності

Для того, щоб можна було оцінити виграш від розпаралелювання, необхідно вибрати показник ефективності. Одним з показників є порівняння часу виконання кращого послідовного алгоритму з часом виконання паралельної програми. Це відношення відоме як прискорення і характеризує те, наскільки швидше виконується програма при розпаралелюванні.

#### $S(p)=T\pi/T\pi$

Де S(p) - прискорення при використанні p потоків, Tn - час виконання послідовного алгоритму, Tp - час виконання паралельного алгоритму.

Тобто, прискорення визначається як функція від кількості фізичних потоків (р), використаних в паралельній реалізації.

#### Закон Амдала

3 урахуванням попереднього визначення прискорення, чи можна обчислити теоретичну межу підвищення продуктивності при збільшенні числа процесорних ядер (а отже, фізичних потоків) в програмі?

При вивченні цього питання зазвичай починають з роботи Джина Амдала (Gene Amdahl), виконаної в 1967 р. Його правило, відоме як закон Амдала, описує максимальний теоретичний виграш в продуктивності паралельного рішення по відношенню до кращого послідовного рішення.

Амдал почавши з інтуїтивно зрозумілого твердження, що прискорення програми  $\varepsilon$  функцією як від тієї частини програми, яка прискорюється, так і від того, наскільки ця частина програми прискорюється.

$$S=1/((1 - P1)+(P1/Sp))$$

Де S - прискорення, P1 - прискорена частина програми, Sp - прискорення прискореної частини програми.

## Приклад

Припустимо, ми змогли прискорити половину програми на 15 %. Тоді очікуване прискорення всієї програми може бути:

$$S = 1/((1 - 0.50) + (0.50/1.15)) = 1/(0.50 + 0.43) = 1.08.$$

Результат - прискорення на 8%, чого і слід було чекати. Якщо половина програми прискорилася на 15 %, то уся програма прискориться на половину цього значення.

Амдал потім пішов далі, щоб показати, в що виливається це рівняння після підстановок для тих частин програми, які розпаралелюються, і тих, які виконуються послідовно. Він вивів формулу, яку називають Законом Амдала.

Рівняння 1.1.

У цьому рівнянні S - прискорення, T - час, витрачений на виконання послідовної частини паралельної версії, p - це кількість процесорів (ядер процесора).

Зверніть увагу, що чисельник рівняння припускає, що програмі потрібно 1 одиницю часу для виконання кращого послідовного алгоритму.

Якщо ви підставите 1 замість кількості ядер процесора (p), то побачите, що прискорення немає. Якщо у вас двоядерна платформа робить половину роботи в паралельному режимі, то ви отримаєте:

$$1/(0.5+0.5/2)=1/0.75=1.33.$$

Тобто прискорення 33 %, оскільки час виконання (знаменник) складає 75% початкового часу виконання. Для восьмиядерного процесора прискорення складе:

$$1/(0.5+0.5/8)=1/0.56=1.78.$$

Підставивши  $p = \infty$  в рівняння 1.1 і припустивши, що кращий послідовний алгоритм виконується одну одиницю часу, отримаємо рівняння 1.2.

Рівняння 1.2. Верхня межа для додатка, що виконує послідовний код S часу

#### S=1/T (1.2)

Тобто Амдал вважає, що додавання процесорних ядер добре масштабується. Таке формулювання закону показує, що максимальний виграш від розпаралелювання деякої частини коду обмежений послідовною частиною коду. З цього результату можна побачити перший наслідок закону Амдала:

скорочення послідовної частини шляхом збільшення паралельної частини *важливіше*, ніж збільшення числа процесорних ядер. Наприклад, якщо у вас є програма, яка на 30 % распаралелена і виконується на двоядерній системі, то подвоєння числа процесорних ядер скорочує час виконання з 85 % часу послідовного алгоритму до 77,5 %, тоді як подвоєння кількості параллеленного коду зменшує час виконання з 85 до 70 %.

Збільшення кількості процесорів допомагає більше, ніж розпаралелювання коду, що залишився, тільки тоді, коли велика частина програми вже распаралелена.

Таким чином, Закон Амдала показує, що приріст ефективності обчислень залежить від алгоритму завдання і обмежений згори для будь-якого завдання. Не для всякого завдання має сенс нарощування числа процесорів в обчислювальній системі.

## 1. Огляд сучасних стандартів паралелізму

При створенні паралельних програм, які призначені для багатопроцесорних обчислювальних систем, використовуються різні види взаємодії між різними процесами, які можуть бути розділені на два типи:

- Взаємодія через спільну пам'ять (наприклад, в Java або С#). Цей вид паралельного програмування зазвичай вимагає якоїсь форми захоплення управління (мьютекси, семафори, монітори) для координації потоків між собою.
- Взаємодія с допомогою передачі повідомлень. Обмін повідомленнями може відбуватися асинхронно, або с використанням методу "рандеву", при якому відправник блокований до тих пір, поки його повідомлення не буде доставлено. Асинхронна передача повідомлень може бути надійною (з гарантією доставки) або ненадійною.

Методи паралельного програмування в .Net, які ми розглядали, використовують **спільну пам'ять одного процесу**. Крім цих методів існують інші стандарти паралельного програмування, які розроблені до появи .Net.

Основні з них:

- OpenMP стандарт інтерфейсу для паралельних систем із спільною пам'яттю.
  - POSIX Threads стандарт реалізації потоків виконання.
  - Windows API багатопотокові застосування для С++.
- PVM (Parallel Virtual Machine) дозволяє об'єднати різнорідний (але пов'язаний мережею) набір комп'ютерів в загальний обчислювальний ресурс.
- MPI (Message Passing Interface) стандарт систем передачі повідомлень між паралельно виконуваними процесами.
- В багатопроцесорних обчислювальних системах з **розподіленою** пам'яттю, процесори працюють незалежно один від одного. Для організації паралельних обчислень в таких умовах необхідно мати можливість розподіляти обчислювальне навантаження і організувати інформаційну взаємодію (передачу даних) між процесорами.

Вирішення усіх перерахованих питань забезпечує інтерфейс обміну повідомленнями (Message Passing Interface, MPI). При використанні цього інтерфейсу обмін даними між різними процесорами в програмі здійснюється за допомогою механізму передачі і приймання повідомлень. Окрім MPI при створенні паралельних програм можливі і інші методи обміну, наприклад, засновані на використанні функцій програмної системи PVM (Parallel Virtual Machine) або функцій бібліотеки SHMEM, розробленої компаніями Cray і Silicon Graphics.

При створенні паралельних програм, призначених для багатопроцесорних обчислювальних систем із спільною пам'яттю, нині для обміну даними між процесорами зазвичай використовуються або методи багатопотокового програмування за допомогою потоків (threads), або технологія **OpenMP**. Директиви OpenMP є спеціальними директивами для компіляторів. Вони створюють і організовують виконання паралельних процесів (потоків), а також обмін даними між процесами. Слід зазначити, що обмін даними між процесами в багатопроцесорних обчислювальних системах із спільною пам'яттю може здійснюватися і з застосуванням функцій MPI і SHMEM.

При використанні багатопроцесорних обчислювальних систем із спільною пам'яттю зазвичай передбачається, що наявні у складі системи процесори мають рівну продуктивність,  $\epsilon$  рівноправними при доступі до спільної пам'яті, і час доступу до пам'яті  $\epsilon$  однаковим (при одночасному доступі декількох процесорів до одного і того ж елементу пам'яті черговість і синхронізація доступу забезпечується на апаратному рівні). Багатопроцесорні системи подібного типу зазвичай іменуються симетричними

мультипроцесорами (symmetric multiprocessors, SMP).

вище набору припущень Перерахованому задовольняють багатоядерні процесори, в яких кожне ядро представляє практично незалежно функціонуючий обчислювальний пристрій. У рамках цієї технології директиви паралелізму використовуються для виділення в програмі паралельних фрагментів, в яких послідовний виконуваний код може бути розділений на декілька роздільних командних потоків (threads). Далі ці потоки можуть виконуватися на різних процесорах (процесорних ядрах) системи. В результаті такого підходу обчислювальної представляється у вигляді набору послідовних (однопотокових) і паралельних (багатопотокових) ділянок програмного коду. Подібний принцип організації паралелізму отримав найменування "вилкового" (fork - join) або пульсуючого паралелізму.

#### 4.1. Основні принципи МРІ

Найбільш поширеною технологією програмування для паралельних комп'ютерів з розподіленою пам'яттю нині є MPI. Основним способом взаємодії паралельних процесів в таких системах є передача повідомлень. Стандарт MPI фіксує інтерфейс, який повинен дотримуватися як системою програмування на кожній обчислювальній платформі, так і користувачем при створенні своїх програм. MPI підтримує роботу з мовами Фортран і C/C++.

МРІ -програма - це множина паралельних взаємодіючих процесів. Усі процеси породжуються один раз, утворюючи паралельну частину програми. В ході виконання МРІ -програми породження додаткових процесів або знищення існуючих не допускається. Кожен процес працює у своєму адресному просторі, ніяких спільних змінних або даних в МРІ немає. Основним способом взаємодії між процесами є явна посилка повідомлень.

## 4.2. Основні принципи ОрепМР

ОрепМР - це інтерфейс прикладного програмування для створення багатопотокових застосувань, призначених в основному для паралельних обчислювальних систем із спільною пам'яттю. ОрепМР складається з набору директив для компіляторів і бібліотек спеціальних функцій. Стандарти ОрепМР розроблялися протягом останніх 15 років для архітектури із загальною пам'яттю. Стандарт ОрепМР реалізований для мов Fortran і C/C++.

Модель програмування ОрепМР надає *незалежний від платформи* набір директив і прагм компілятора, викликів функцій і змінних середовища, які явно показують компілятору, як і де використовувати паралелізм в програмі.

Останніми роками дуже активно розробляється розширення стандартів ОрепМР для паралельних обчислювальних систем з *розподіленою* пам'яттю. У кінці 2005 року компанія Intel анонсувала продукт Cluster OpenMP, що реалізовує розширення OpenMP для обчислювальних систем з розподіленою

пам'яттю. Цей продукт дозволяє оголошувати області даних, доступні усім вузлам кластера, і здійснювати передачу даних між вузлами кластера неявно за допомогою протоколу Lazy Release Consistency.

ОрепМР дозволяє легко і швидко створювати багатопотокові застосування на алгоритмічних мовах Fortran і C/C++. При цьому директиви ОрепМР аналогічні директивам препроцесора для мови C/C++ і є аналогом коментарів в алгоритмічній мові Fortran. Це дозволяє у будь-який момент розробки паралельної реалізації програмного продукту при необхідності повернутися до послідовного варіанту програми.

Нині OpenMP підтримується більшістю розробників паралельних обчислювальних систем: компаніями Intel, Hewlett - Packard, Silicon Graphics, Sun, IBM, Fujitsu, Hitachi, Siemens, Bull і іншими. Багато відомих компаній в області розробки системного програмного забезпечення також приділяють значну увагу розробці системного програмного забезпечення з OpenMP. Серед цих компаній Intel, KAI, PGI, PSR, APR, Absoft і деякі інші. Значне число компаній і науково-дослідних організацій, розробляючих прикладне програмне забезпечення, нині використовує OpenMP при розробці своїх програмних продуктів.

#### Приклад.

Наступний цикл конвертує кожен 32-розрядний піксел масиву, відповідний колірній моделі RGB (Red, Green, Blue - червоний, зелений, синій), в монохромний 8-розрядний піксел. Єдина прагма, вставлена безпосередньо перед циклом, - це усе, що треба в ОрепМР для паралельного виконання.

```
#pragma omp parallel for for ( i - 0; i < numPixels; i++) {
  pGrayScaleBitmap[i] = (unsigned BYTE) ( pRGBB1tmap[i].red * 0.299 + pRGBBitmap[i].green * 0.587 + pRGBB1tmap[i].blue * 0.114 );
}
```

У прикладі має місце розподіл роботи - цим терміном в ОрепМР описують розподіл роботи між програмними потоками. Коли розподіл роботи відбувається за допомогою конструкції **for** (як в цьому прикладі), то ітерації циклу розподіляються між декількома потоками. При цьому реалізація ОрепМР сама визначає, скільки потоків треба створити і як краще ними управляти. Програміст просто повідомляє ОрепМР, який цикл треба обробляти за допомогою потоків. Компілятор ОрепМР і бібліотека часу виконання самі потурбуються про створення потоків і розподілі їх по процесорах.

#### Висновки

Паралелізм означає дійсно одночасне виконання декількох програмних потоків. Сучасні застосування часто складаються з декількох процесів або програмних потоків, які можуть виконуватися паралельно. - Більшість сучасних комп'ютерних платформ - це комп'ютери з декількома потоками команд і декількома потоками даних (МІМD). Ці комп'ютери дозволяють програмістам обробляти декілька потоків команд і даних одночасно.

На практиці закон Амдала неточно відбиває перевагу від збільшення кількості процесорних ядер на цій платформі. Шляхом збільшення кількості процесорних ядер можна досягти лінійного прискорення.

Крім засобів паралельного програмування, які реалізовані в платформі .Net Framework, існують інші технології організації паралельних обчислень.

Насамперед:

- OpenMP стандарт інтерфейсу для паралельних систем із спільною пам'яттю.
  - POSIX Threads стандарт реалізації потоків виконання.
  - Windows API багатопотокові застосування для С++.
- PVM (Parallel Virtual Machine) дозволяє об'єднати різнорідний (але пов'язаний мережею) набір комп'ютерів в загальний обчислювальний ресурс.
- MPI (Message Passing Interface) стандарт систем передачі повідомлень між паралельно виконуваними процесами.

Але засоби паралельного програмування, які ми вивчали в цьому курсу, особливо бібліотека TPL, дійсно спрощують розробку паралельних програм на .Net-сумісних мовах програмування. Це дозволяє програмісту зосередитися на бізнес-логіці програми, а не на технічних питаннях паралелізму.

## Контрольні запитання і завдання для самостійного виконання

- 1. У чому полягає складність внесення паралелізму в програму?
- 2. В яких випадках розпаралелювання може не принести очікуваних переваг, а навпаки, знизити продуктивність?
- 3. Яка головна мета паралельного програмування?
- 4. Які можуть бути причини низької продуктивності паралельної програми?
- 5. Що таке **гранулярність завдань** і як вона впливає на продуктивність?
- 6. Що означає поняття надзвичайного паралелізму і до чого воно застосовне?
- 7. Що означає наступне співвідношення  $S(p)=T\pi/T\pi$ ? Що таке прискорення виконання програми?
- 8. Закон Амдала.
- 9. Яка різниця між взаємодією процесів через спільну пам'ять і через обмін повідомленнями?
- 10. Технологія МРІ базується на обміні повідомленнями чи на спільній

#### пам'яті?

11. В яких мовах є реалізації ОрепМР?

#### Рекомендована література

#### Основна

1. Г. Шилдт. С# 4.0: полное руководство. М.: ООО "И.Д. Вильямс", 2011. -1056 с.

#### Додаткова

- 2. Эхтер Ш., Роберте Дж. Многоядерное программирование. СПб.: Питер, 2010. 316 с.
- 3. Эндрюс Г.Р. Основы многопоточного, параллельного и распределенного программирования. Пер. с англ. –М: Изд.дом Вильямс, 2003. -512 с.

#### Ресурси Інтернет

- 4. www.intuit.ru
- 5. www.parallel.ru
- 6. www.rsdn.ru/article/dotnet/
- 7. http://www.intuit.ru/studies/courses/5938/1074/info

Деякі приклади взяті з ресурсу msdn.microsoft.com/ru-ru:

http://msdn.microsoft.com/ru-ru/library/system.threading.tasks.task.aspx

http://msdn.microsoft.com/ru-ru/library/dd460705.aspx

http://msdn.microsoft.com/ru-ru/library/dd460713.aspx

http://msdn.microsoft.com/ru-ru/library/dd997425.aspx