А.В. Макаров, С.Ю. Скоробогатов, А.М. Чеповский

Учебный курс “CIL и системное программирование в Microsoft .NET”

Лекция 5.

Формат метаданных

Введение

- Метаданные служат для описания типов, находящихся в сборке .NET, и хранятся в исполняемых файлах
- Мы изучим формат метаданных на основе учебного примера

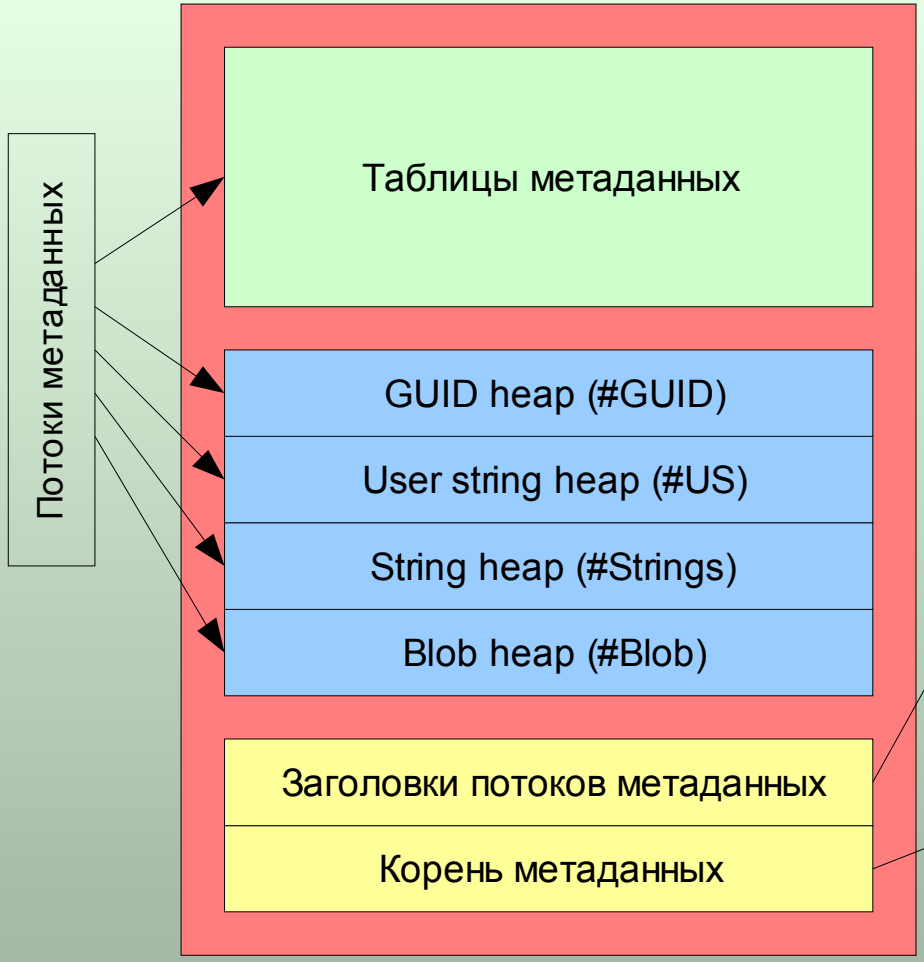
Учебный пример

```
.assembly HelloWorld
{
     .hash algorithm 0x00008004
     .ver 1:0:1:1
}

.module HelloWorld.exe

// hello() – единственный метод в нашей сборке
.method public static void hello() cil managed
{
     .entrypoint
     .maxstack 8

     // Загружаем строку "Hello, World!" на стек
     ldstr "Hello, World!"
     // Выводим строку на экран
     call void [mscorlib]System.Console::WriteLine(string)
     // Ожидаем, пока пользователь введет строку
     call string [mscorlib]System.Console::ReadLine()
     // Выводим введенную строку на экран
     call void [mscorlib]System.Console::WriteLine(string)
     // Завершаем выполнение
     ret
}
```

5.1. Расположение метаданных и кода внутри сборки

- В начале секции .text располагается тело метода hello, за которым следуют метаданные
- Мы можем выбрать для метаданных практически любое место внутри секции, и их расположение, изображенное на схеме, является лишь одним из многих возможных вариантов

5.2. Структура метаданных



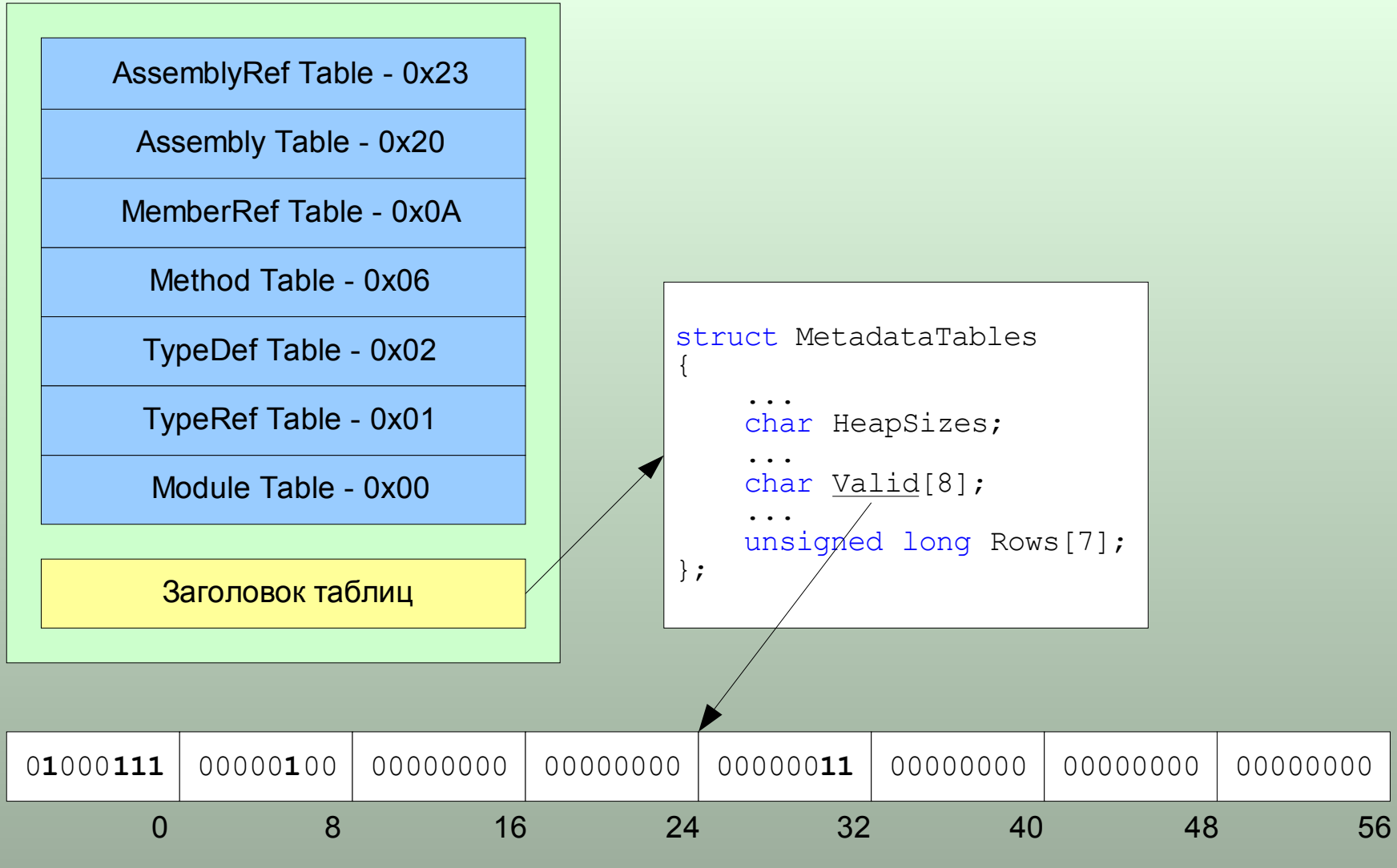

```
/* Для каждого потока: */
struct StreamHeader
{
    long Offset;
    long Size;
    char Name[x];
};
```

```
struct MetadataRoot
{
    long Signature;
        /* 0x424A5342 */
    ...
    unsigned short Streams;
};
```

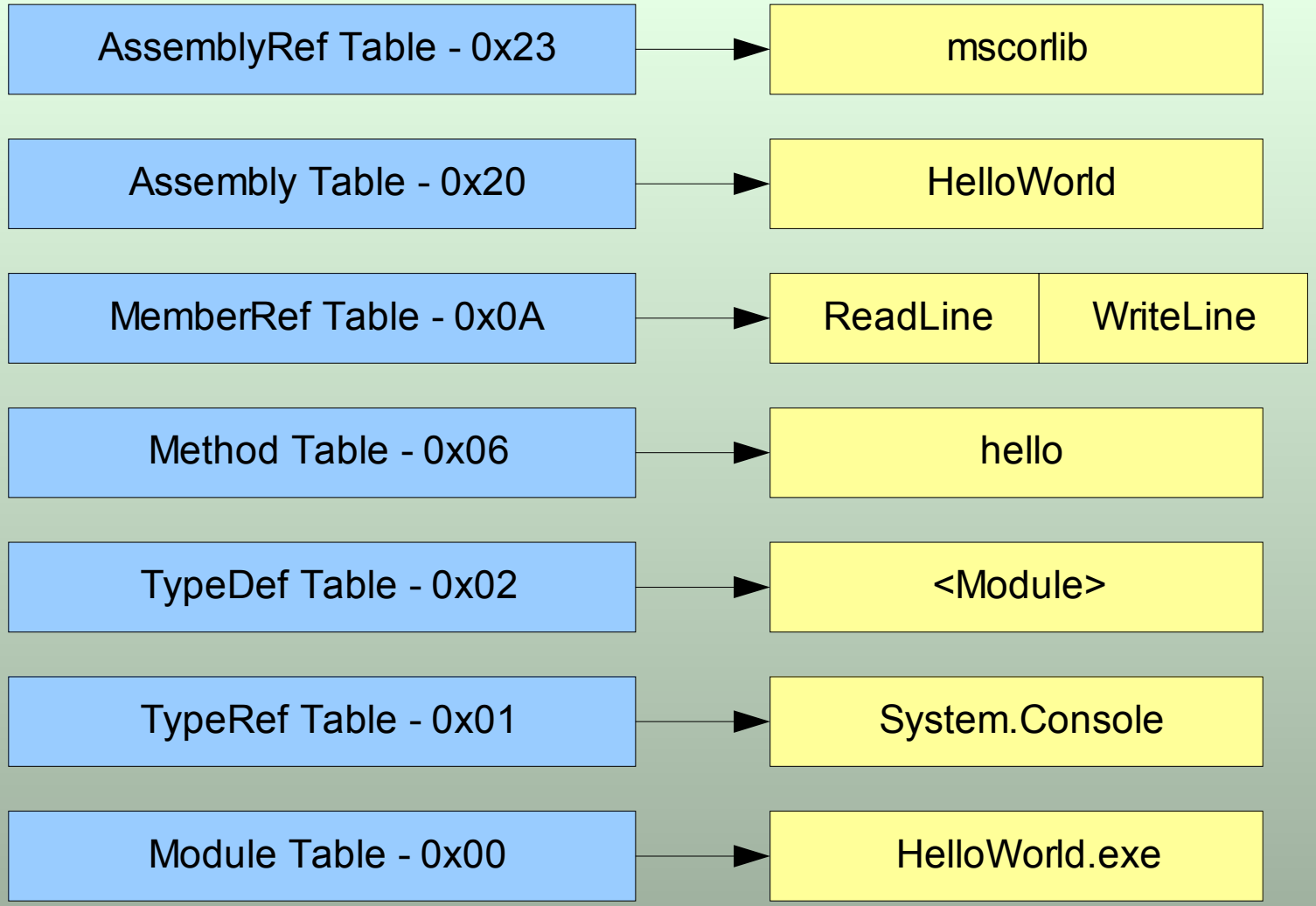
Потоки метаданных

Поток	Имя потока	Описание
Куча GUID'ов	`"#GUID"`	Представляет собой последовательность 128-битных глобальных уникальных идентификаторов.
Куча пользовательских строк	`"#US"`	Содержит строковые константы, определенные в программе.
Куча строк	`"#Strings"`	Содержит названия элементов метаданных (типов, методов, полей и т.п.).
Куча двоичных данных	`"#Blob"`	Содержит двоичные данные, описывающие метаданные (например, сигнатуры методов).
Таблицы метаданных	`"#~"`	Содержит физическое представление таблиц метаданных.

5.3. Таблицы метаданных



Распределение элементов метаданных учебного примера по таблицам



5.3.1. Таблица сборок (Assembly – 0x20)

- В этой таблице содержится ровно одна запись, описывающая нашу сборку. В этой записи для нас интерес представляют следующие поля:
  - `short MajorVersion;`
  - `short MinorVersion;`
  - `short BuildNumber;`
  - `short RevisionNumber;`
    - Эти четыре поля хранят информацию о версии сборки
  - `short Name;`
    - Это поле содержит индекс в куче строк, по которому хранится имя сборки ("HelloWorld")

5.3.2. Таблица модулей (Module – 0x00)

- Таблица модулей содержит ровно одну запись, описывающую модуль. При этом в этой записи существенными являются два поля:
  - `short Name;`
    - Это поле содержит индекс в куче строк, по которому хранится имя модуля ("HelloWorld.exe")
  - `short Mvid;`
    - Это поле содержит индекс в куче GUID'ов, по которому хранится глобальный уникальный идентификатор модуля

5.3.3. Таблица определенных в сборке типов (TypeDef – 0x02)

- В нашем учебном примере нет классов, но для того, чтобы объявить глобальную функцию, нам требуется специальный абстрактный тип <Module>
- Запись таблицы содержит следующие интересующие нас поля:
  - `short Name;`
    - Это поле содержит индекс в куче строк, по которому хранится имя типа ("<Module>")
  - `short FieldList;`
  - `short MethodList;`
    - Эти два поля содержат индексы в таблицах полей и методов, начиная с которых расположены описатели полей и методов типа

5.3.4. Таблица методов (Method – 0x06)

- Таблица методов описывает методы, объявленные в сборке. Каждая запись этой таблицы содержит информацию об одном методе, представленную в следующих полях:
  - `long RVA;`
    - RVA тела метода в исполняемом файле
  - `short Flags;`
    - Набор флагов, задающих область видимости метода и другие его атрибуты
  - `short Name;`
    - Это поле содержит индекс в куче строк, по которому хранится имя метода
  - `short Signature;`
    - Индекс в куче двоичных данных, по которому расположена сигнатура метода
  - `short ParamList;`
    - В этом поле хранится индекс в таблице описателей параметров метода

5.3.5. Таблица импортируемых сборок (AssemblyRef – 0x23)

- Все импортируемые сборки должны быть перечислены в таблице импортируемых сборок. Каждая запись этой таблицы содержит следующие поля:
  - `short MajorVersion;`
  - `short MinorVersion;`
  - `short BuildNumber;`
  - `short RevisionNumber;`
    - Эти четыре поля хранят информацию о версии импортируемой сборки
  - `short Name;`
    - Это поле содержит индекс в куче строк, по которому хранится имя импортируемой сборки (в нашем случае это "mscorlib")

5.3.6. Таблица импортируемых типов (TypeRef – 0x01)

- Каждая запись в этой таблице соответствует одному из импортируемых типов и содержит следующие поля:
  - `short ResolutionScope;`
    - Специальным образом закодированная информация о том, какой сборке или какому модулю принадлежит данный тип
  - `short Name;`
    - Это поле содержит индекс в куче строк, по которому хранится имя импортируемого типа (в нашем случае это "Console")
  - `short Namespace;`
    - Это поле содержит индекс в куче строк, по которому хранится имя пространства имен, которому принадлежит импортируемый тип (в нашем случае это "System")

5.3.7. Таблица членов импортируемых типов (MemberRef – 0x0A)

- В таблице членов импортируемых типов перечислены все методы, поля и свойства этих типов, которые используются в программе. Каждая запись этой таблицы содержит следующие поля:
  - `short Class;`
    - Специальным образом закодированная информация о об импортируемом типе
  - `short Name;`
    - Это поле содержит индекс в куче строк, по которому хранится имя члена импортируемого типа
  - `short Signature;`
    - Индекс в куче двоичных данных, по которому расположена сигнатура члена импортируемого типа