**Background (motivation part)**

This is a technical description of ChartPoints MSVC 2015 extension.

All user information is here: <https://marketplace.visualstudio.com/items?itemName=aleksey.ChartPoints>

Source code:

***Postulate #1***: “*Tracing is helpful*”.

All we are using it from time to time. Me too.

To achieve this goal predefined IDE macros/classes, third-party libraries, own techniques are commonly used. One of the main problems is that we need to do it manually in code.

This leads to ***Postulate #2***: “All that can be automated must be automated.”

***Postulate #3:*** “Sometimes quick look from afar helps to find problematic areas easier than digging in details”

**Conclusions**

1. Give user the possibility to add trace points from IDE (inspired by ***Postulate #2***)
2. Provide interactive charts with traced data (inspired by ***Postulate #3***)
3. Minimize overhead

**Important notes (justification part #1)**

***Note #1***

It’s only the first experimental stage (cycle) with many limitations. Some parts need to be refactored/changed, e.g. choice of COM exe server as transport, but I didn’t refuse to it's usage in order to understand this more clearly, ☺

***Note #2***

My knowledge of C# is far away from professional level. There were 2 reasons to use it:

* It’s easier to develop MSVS extension using it than to develop in native C++
* It was interesting to me

*hence the conclusion*: please be lenient.

***Note #3***

This project contains many stand-alone parts, which will not be fully covered in this article. I will make common architecture overview and then concentrate on most significant aspects. If this is not enough, please mention this in comments. If will be interested I will gather all questions and write “Part 2”.

***Note #4***

I’ll not describe how to create tool windows, context menu handler, e.g. This is fully described in many samples. First of all in **VSSDK-Extensibility-Samples** (<https://github.com/Microsoft/VSSDK-Extensibility-Samples>). I’ll concentrate on most difficult (IMHO) parts.

***Note #5***

My English is not as good as want I’d like. Sorry..

**Limitations (justification part #2)**

* Supported languages: native C++ only
* Supported IDE: MSVC 2015
* Tracing variables: class members of C++ fundamental types only + some typedefs like std::uint32\_t (because they are useful)

**Used technologies/languages**

* C++ (injected code, COM out-of-proc server)
* C# (MSVS 2015 extension, MSBuild task)

**Used architecture patterns**

Two main common architecture patterns are used in this project: ***SOA (Service-Oriented Architecture)*** *&* ***EDA (Event-Driven Architecture)***. Both patterns are long known and well described, so I’ll focus only on details of current implementation.

***SOA (Service-Oriented Architecture)***

*Declaration*: …

// base interface for all ChartPoints services (maybe extended in the future)

public interface ICPService

{

}

//public delegate void OnCPServiceCreated<T>(T args);

// singleton service provider

public abstract partial class ICPServiceProvider

{

public abstract bool RegisterService<T>(T obj) where T : ICPService;

public abstract bool GetService<T>(out T obj) where T : class;

//public abstract bool GetService<T>(out T obj, OnCPServiceCreated<T> cb) where T : class;

public static ICPServiceProvider GetProvider()

{

return impl.ICPServiceProvider.GetProviderImpl();

}

}

*Implementation*: …

Service provider stores registered services on calling *RegisterService* for future return by *GetService*. Second (commented) version of *GetService* will allow to provide callback which will be used if service is not registered yet. There is no reasons to implement it now so I left it commented.

This approach allows to manage the order of service creation & query services from any place without worrying about their construction.

***EDA (Event-Driven Architecture)***

*Declaration*: …

public delegate void OnCPEvent<T>( T args );

public abstract class ICPEvent<T>

{

protected abstract ICPEvent<T> Add(OnCPEvent<T> cb);

public static ICPEvent<T> operator +(ICPEvent<T> me, OnCPEvent<T> cb)

{

return me.Add(cb);

}

protected abstract ICPEvent<T> Sub(OnCPEvent<T> cb);

public static ICPEvent<T> operator -(ICPEvent<T> me, OnCPEvent<T> cb)

{

return me.Sub(cb);

}

public abstract void Fire(T args);

}

*Implementation*: …

*ICPEvent* provides usual +/- operators. In current implementation all events are stored (there are not so many of them so we don’t have to worry). When new client subscribes it receives all earlier generated events. If it will be required, simple specialization without history will be added. Current implementation provides all the needs.

It’s done for two reasons:

1. Some objects are initialized not in predefined order (e.g. from event handlers of MSVS) and this guarantees that all events will be delivered to recipients.  
   *Example*: When tagger(object responsible for rendering glyphs in code editor) is created and subscribes to ChartPoints events, it successfully receives all fired earlier events and have all actual information for rendering glyphs.
2. It gives a possibility to add time marks to make them as part logging system(not implemented now, but can easily added).

**Class Factory**

Basic goals of current class factory implementation:

1. Allows to implement class factories based on any of existing ones (partial extension) and provide moqs & stubs for testing.
2. Strategy pattern via DI. Allows of implementing different strategies of same interface. As all objects query instances via factory methods, all that is need to do is to change appropriate class factory method for new object implementation.
3. Hide the possibility of explicitly creating objects intended to be created via class factory (in same assembly too). It was hard to understand how to achieve this. In C++ it is done easily, but in C# redundant entities need to be added to emulate ‘friend’ keyword. Maybe it can be done more easier..

***Declaration of class factory***

// class factory interface

public abstract partial class IClassFactory

{

// set custom class factory instance for DI purposes

public static void SetInstance(IClassFactory inst)

{

ClassFactory.SetInstanceImpl(inst);

}

// returns singleton instance

public static IClassFactory GetInstance()

{

return ClassFactory.GetInstanceImpl();

}

// factory methods

public abstract IChartPointsProcessor CreateCPProc();

**<..>**

}

***Classes intended to constructed via class factory declaration***

*Interface declaration*

public interface IChartPointsProcessor

{

**<..>**

}

*Implementation*

***Important:*** *marked abstract to hide possibility to explicitly construct it. Only derived classes can do that*

public abstract class ChartPointsProcessor : IChartPointsProcessor

{

**<..>**

}

***Implementation of class factory***

public abstract partial class IClassFactory

{

// implementation of ordinal class factory

// is accessible only by IClassFactory and class factory implementations (for DI purposes)

private partial class ClassFactory : IClassFactory

{

public ClassFactory()

{

**<..>**

}

private static IClassFactory Instance;

public static void SetInstanceImpl(IClassFactory inst)

{

Instance = inst;

}

public static IClassFactory GetInstanceImpl()

{

if (Instance == null)

Instance = new ClassFactory();

return Instance;

}

// IChartPointsProcessor factory

// Opens the back-door to construct ChartPointsProcessor object

// To access non-default constructors appropriate delegating ones need to be added

private class ChartPointsProcImpl : ChartPointsProcessor { }

// IChartPointsProcessor factory method implementation

public override IChartPointsProcessor CreateCPProc()

{

return new ChartPointsProcImpl();

}

**<..>**

}

***Extended class factory***

// IChartPointsProcessor implementation for dependency injection

namespace impl

{

namespace DI

{

// Can be fully implemented from IChartPointsProcessor or extend any existing implementation

public class DIChartPointsProcessor : *ChartPointsProcessor*

{

// methods to override

}

} // namespace DI

}

// dependency injection class factory implementation

public abstract partial class IClassFactory

{

// Can be fully implemented from IClassFactory or extend existing one

// In this case ordinal class factory used to override IChartPointsProcessor factory method only

class DIClassFactory\_01 : ClassFactory

{

private class ChartPointsProcImpl : DIChartPointsProcessor { }

// IChartPointsProcessor factory method implementation

public override *IChartPointsProcessor* CreateCPProc()

{

return new ChartPointsProcImpl();

}

}

// instantiate di class factory. Needed because IClassFactory.DIClassFactory\_01 declaration is inaccessible explicitly

public static IClassFactory GetInstanceDI\_01()

{

return new DIClassFactory\_01();

}

}

Somewhere (before construction of other class factory objects started) call

Utils.IClassFactory diCF = Utils.IClassFactory.GetInstanceDI\_01();

// after calling this all objects will be constructed via this class factory

Utils.IClassFactory.SetInstance(diCF);

**How it works**

***Quick view***

***ChartPoints***:

* User friendly interface for adding ChartPoints (aka breakpoints)
* Taggers in code editor indicating their placement
* List of ChartPoints in special tool window
* Simple code text changes listener (aka breakpoints)
* Save/Load defined ChartPoints
* Separate ChartPoints mode from ordinal builds
* ChartPoints validation before build, before/after save/load
* User interactive chart view
* Table view of ChartPoints values based on their generation time

***Code generation***

* MSBuild task for code tracing code injection
* Transport between MSVS extension & MSBuild task

***Trace library (publisher side – traced program)***

Minimize overhead between traced & untraced code execution. It is very important because if they differ much it will provide different behavior and tracing will become meaningless

***Trace library (consumer side - host)***

The requirements on this side are not as strict as on publisher’s. The main goal of this project is to perform post analysis. So some lag in run-time allowed.

***Trace transport***

As mentioned earlier I decided to use COM exe server as transport layer between the program being traced and the host. As it seems to me this wasn’t a good idea and need to be changed in the future. I’m planning to change transport layer later. So I will not describe it in detail.

***Description***

***Step #1 (Selection of ChartPoints)***

Visual Studio contains a set of interfaces for manipulating language code model. *CodeModel* – common for all supported languages. *VCCodeModel* – extended version for C/C++ projects. *FileCodeModel* – code model based on source file structure.

As I indicated in limitations section only class variables of C++ fundamental types are allowed to be traced. So the only one place where it can be done is class method definition.

Before showing context menu in code editor, checking availability of adding ChartPoint in performed. This is done in *CP.Code.Model.CheckCursorPos*() method. From *EnvDTE.ActiveDocument* current cursor position (*EnvDTE.ActiveDocument.Selection.ActivePoint*) and *FileCodeModel* (*EnvDTE.ActiveDocument.ProjectItem.FileCodeModel*) are acquired. Using *FileCodelModel.CodeElementFromPoint* method cursor position check is performed: being inside method body. If so, *Parent* property of returned *CodeElement* points to *VCCodeClass* object, which is used to get all class variables. The future injection point will be at the beginning of the line or immediately after open brace of method if cursor is on line containing it.

One ChartPoint can contain multiple traceable variables.

All set ChartPoints are added to “ChartPoints:design” tool window.

***Brief ChartPoints classes architecture***



All ChartPoints data is stored in tree structure. This objects provide events for subscribing on their Add/Move/Remove/Status changes. This objects composition gives the possibility to easily operate them in forward (from root to leafs) and backward (based on events notifications) order. Both approaches will be actively used further.

**Step #2 (Taggers)**

**VSSDK-Extensibility-Samples** contains samples showing basic usage of taggers. Also MSDN has several articles describing it. The beginning point is here: *Inside the Editor* (<https://msdn.microsoft.com/en-us/library/dd885240.aspx>).

But I want more:

1. Force taggers appearance/change
2. Optimize performance (exclude redundant updates)

**Short taggers overview**

Every time new document is opened/changed MSVS calls custom *IViewTaggerProvider* implementation (if any) method CreateTagger to create *ITagger* object. It’s method *GetTags* will be later called from MSVS environment in order to determine if (and where) tags are present.

***Problem:***

*IViewTaggerProvider*.CreateTagger is called multiple times for the same document. It looks like it called for each window that can contain tags: code editor, find results (???). As I found the last one is called for code editor window. Yes, it’s works but I don’t have full understanding.. So this needs to be researched more clearly.

**Custom taggers implementation**



All created tagger are stored in association array with file names as keys. ChartPoints tagger provider subscribes on *IFileChartPoints* Add/Remove events with subsequent providing *IFileChartPoints* object to stored taggers. This gives them possibility to subscribe on *ILineChartPoint* events notifications.

When document is opened for the first time or changed, *GetTags* method of ChartPoints tagger is called. In this method intersection of lines in SnapshotSpan and stored containing ChartPoints numbers calculated.

If needed to update tag manually from outside, *IChartPointsTagger.*RaiseTagChangedEvent is fired with parameter containing line number. *ITagger<ChartPointTag>.TagsChanged* event with *SnapshotSpan* containing only 1 line, where ChartPoint is placed, fired. This helps to exclude redundant checks on tags creation and provides the possibility to force (re-)draw tags.

***Important***: all indexes (line/character numbers) used here are 0-based. *EnvDTE* indexes, which are used to calculate ChartPoints positions, starts from 1. And this is the cause of constant headache.

**Step #3 (Save/Load ChartPoints)**

All information about contained ChartPoints is saved per solution basis in \*.suo (Solution User Options) file.

In order to do it I use implementation of *IVsPersistSolutionOpts* interface which provides overloaded methods and a reference to *Microsoft*.*VisualStudio*.*OLE*.*Interop*.*IStream* object. On load this object is cloned & stored for use after solution loaded.

**Step #4 (Text changed tracker)**

Code changes are tracked only by simple text changes now. Perhaps this is enough. I slightly experimented with VCCodeModel but decided that it is too complicated & expensive.

Tracking system divided into 2 parts: UI (MSVS side) and Model (ChartPoints). It was done so when I thought that I will use both text change listeners & code model changes. Maybe someday I will return back to this.

***UI***

MSVS services for listen text changes are: *IWpfTextViewCreationListener* and *IWpfTextView*. Implementation of the first one provides handle TextViewCreated(IWpfTextView) event. The second gives the possibility to subscribe to *TextBuffer*.*Changed* event.

***Model***

*ICPTrackService* tracks ChartPoints Add/Remove events and provides small wrapper objects which hides ChartPoints objects references. This service and several events bind UI & Model.

*ChartPoints track service sequence diagram*

**

*IWpfTextViewCreationListener.*TextViewCreated(IWpfTextView) is called for each opened document. If no FileTracker objects for this file registered in *ICPTrackService,* *TextChangeListener* will store IWpfTextView object within filename. Later, If Model.FileTracker create event will be received, *FileChangeTracker* object with references to IWpfTextView & FileTracker will be created. It will subscribe to buffer changed event and query validation from *FileTracker*.

**Step #5 (Code instrumentation)**

Code instrumentation is performed via MSBuild task placed in CPInstBuildTask.dll.

When start building (Globals.dte.*Events*.*BuildEvents*.*OnBuildProjConfigBegin* event handler), following actions performed:

1. Check the existence of ChartPoints in current project
2. Disable debug information generation (it is not needed because after instrumentation executing code will differ from original one)

Here is the code doing it:

*EnvDTE*.*Project* proj = ..

...

*VCProject* vcProj = (*VCProject*)proj.*Object*;

*VCConfiguration* vcConfig = vcProj.*Configurations*.*Item*(projConfig);

*IVCCollection* tools = vcConfig.*Tools* as *IVCCollection*;

*VCLinkerTool* tool = tools.*Item*("VCLinkerTool") as *VCLinkerTool*;

tool.*GenerateDebugInformation* = false;

1. Validate ChartPoints
2. Transport for communication between MSVS host and MSBuild task is opened. ServiceHost with *NetNamedPipeBinding* is used for this. Why? It was the first that I saw starting to dig C# capabilities, :). As an address project file full name is used. It looks ugly but gives unique address. Maybe someone someday will decide to synchronously build the same project from several instances of MSVS and problems will be guarantied. But I believe in the power of reason.

MSVS host provides IPCChartPoint interface (and few others placed in the same file) which contains method for calculating ChartPoints injection layout for injection points:

1. Trace variables initialization
2. Trace points
3. Additional include file injection
4. .. and so on

***MSBuild task***

1. Open *ServiceHost* with same address
2. Acquire IPCChartPoint object
3. Calculate injection points layout IPCChartPoint.GetInjectionData(<project name>)
4. Copy required source files to %TEMP% directory
5. Instrument them (see detailed information in next section **Step #6**).
6. Pass instrumented files to MSBuild (add them to build and remove ordinal ones from build)

**Step #6 (C++ tracing library)**

To organize the correct variables tracking, the following data is required:

1. Identifier of traced variable.   
   The variable address casted to unsigned 64-bit value used for this purpose (\*)
2. Variable name
3. Type id  
   For further usage. Not used now.
4. Variable value
5. Timestamp  
   It’s taken at moment of tracing to provide reliable information

(\*) It is guarantied that in current moment address value is unique. But the key phrase is “in current moment”. Same address can be used for multiple times. It depends on variable lifetime. Workaround of this issue will discussed late in “***Partially unique identifier***” section.

Predefined entities used for code instrumentation:

1. cpti(64).dll libraries containing tracing logic. Are explicitly loaded by injected code.
2. \_\_cp\_\_.test\_01.h  
   Declares type id wrapper class type\_id and class tracer which methods are used by instrumented code. It is implemented in *cpti(64).dll* libraries.
3. \_\_cp\_\_.test\_01.cpp  
   Template specializations of type\_id class for all supported types used in tracing  
   Implements tracer\_ptr tracer::instance() method which loads cpti(64).dll depending on Platform (x86/64) used by instrumented module.

*Instrumentation details:*

#include "\_\_cp\_\_.test\_01.h" is added to beginning of each instrumented file. As instrumented files are the copies of ordinal ones, all their include occurrences are changed to use new ones.

\_\_cp\_\_.test\_01.cpp is added to project.

*Registration:*

tracer::pub\_reg\_elem("test\_01::d\_01",d\_01);

Type and variable identifiers are generated on the fly. First – using specialized version of type\_id class. Second by taking and casting variable address

*Tracing:*

tracer::pub\_trace(d\_01);

*Under the hood:*

As I mentioned earlier, one of the main requirements is to minimize overhead.

Following class is used for this purpose:

template<typename TData>

class data\_queue

{

public:

typedef *std*::*queue*< TData > data\_cont;

private:

data\_cont data\_1;

data\_cont data\_2;

public:

data\_cont \*in\_ptr;

data\_cont \*out\_ptr;

data\_queue()

{

in\_ptr = &data\_1;

out\_ptr = &data\_2;

}

void swap()

{

*std*::*swap*( in\_ptr, out\_ptr );

}

};

When tracer::pub\_trace is called, access to data\_queue object is locked and new value with id and timestamp is added to the queue::in\_ptr queue.

Passing of stored values to transport is performed in separate thread, which locks access to data\_queue object only for data\_queue::swap() which swaps queue::in\_ptr & queue::out\_ptr pointers. Then this thread selects all the data from the queue pointed to by queue::out\_ptr and passes it to transport.

This gives the possibility to block calling (trace source) thread for minimal time.

***Partially unique identifier***

As I said “Same address can be used for multiple times” and it complicates it’s usage as identifier. But statement “in current moment” helps.

It resolved in the following way:

One more thread(tracer\_impl::reg\_proc) and data\_queue are created. It works in common the same way as data tracing except one important moment: **before any registration is performed all accumulated data must be sent.**

For this purposes utility class notifier is used. It works like boost::barrier (http://www.boost.org/doc/libs/1\_33\_1/doc/html/barrier.html )/*semaphore* (both are based on waiting on desired counter value). Each time tracer\_impl::reg\_elem is called, registration info (***sic***, containing registration timestamp) is stored, notifier object counter is increased and waiting registration thread is notified. Registration thread publish all accumulated tracing data preceding current registration entity timestamp. After this is done it decreases notifier object counter. All this time data sending thread is sleeping waiting for notifier zero counter value. It provides the current order of reg/trace messages delivery. **Note:** The host (MSVS ChartPoints extension) knows about it, which helps it to correctly handle received messages.

**Step #7 (COM exe server, CPTracer.exe)**

Choice of COM out-of-proc server as a transport was a temporary experiment in order to estimate it’s capabilities. But as is well known "there is nothing more permanent than a temporary"©. So it accompanied me all the time developing the first version, :). And I will refuse using it at the first opportunity. Thant why I’ll say about it usage only few words.

The same data\_queue class instance used for sending data to customer. It’s done to decrease COM events delivery calls. Sending thread sleeps for 500ms on each iteration and then compose all data to arrays that are send to customer in one call.

**Future plans (not in order of priority)**

* Refuse COM server usage. Move to some net protocol.
* Move logic described in **Step #6 (C++ tracing library)** to customer.
* Add tracing capabilities of local variables.
* Remove “<..> [ChartPoints]” configuration. Use MSBuild manually without changing origin \*.sln & \*.vcxproj files.
* Move ChartPoints storage to separate configuration files.
* Relax.. Looks like a candidate for number one in the list of priorities.