

Triton

Voyager's Multimedia IO subsystem

Table of Contents

1. Introduction.....	2
2. Basic overview.....	2
2.1. Plugin List, Process Tree, Stream Groups.....	2
2.2. Nodes and node types in the process tree.....	4
3. The MMIO API for application programmers.....	5
3.1. Forewords.....	5
3.2. Quick usage.....	5
3.3. Advanced usage.....	6
3.3.1. Initializing MMIO.....	6
3.3.2. Managing the plugin list.....	6
3.3.3. Opening an URL.....	7
3.3.4. Creating stream groups.....	8
3.3.5. Setting direction and position of Stream Groups.....	9
3.3.6. Events of Stream Groups.....	9
3.3.7. Uninitializing the MMIO subsystem.....	10
4. Creating Plugins for the MMIO subsystem.....	10
5. Programming guidelines.....	10
6. API Reference.....	11

1. Introduction

The multimedia subsystem of Voyager (called Triton) is a plugin-based, extendable, flexible architecture, originally designed for stable playback of any kind of multimedia streams from any kind of source. The system is not yet capable of recording, it's planned for later versions.

The MMIO subsystem has a layered architecture. It means that in order to play back a given multimedia content, the subsystem builds a logical structure of at least five levels. This structure describes the inner structure of the given multimedia content, and makes it possible to divide the processing and playback of multimedia contents into different modular parts, making it easier to find and fix the bugs in plugins. The modular structure also helps so that once support for a new format is added, it will probably work from all the supported sources (like from local file or from an internet stream), and vice versa, a new plugin for a new source will make it possible to play back all the supported formats without any additional code.

The current version of MMIO has the following layers defined:

- Medium
- Channel
- Elementary Stream
- Raw Stream
- Terminator Nodes (Visualization)

These layers and the layered structure are handled internally by the system, and all these are built up from a URL. This effectively means that playing back a given file (or any kind of multimedia content) is as simple as opening a URL with one MMIO API call, and sending it a Play command.

These URLs look this simple:

file://c:/Musics/My_Favourite.mp3

http://local.video.store.com:8080/Horror/Killing_The_Killer.avi

rtsp://radio.com/live.ra

2. Basic overview

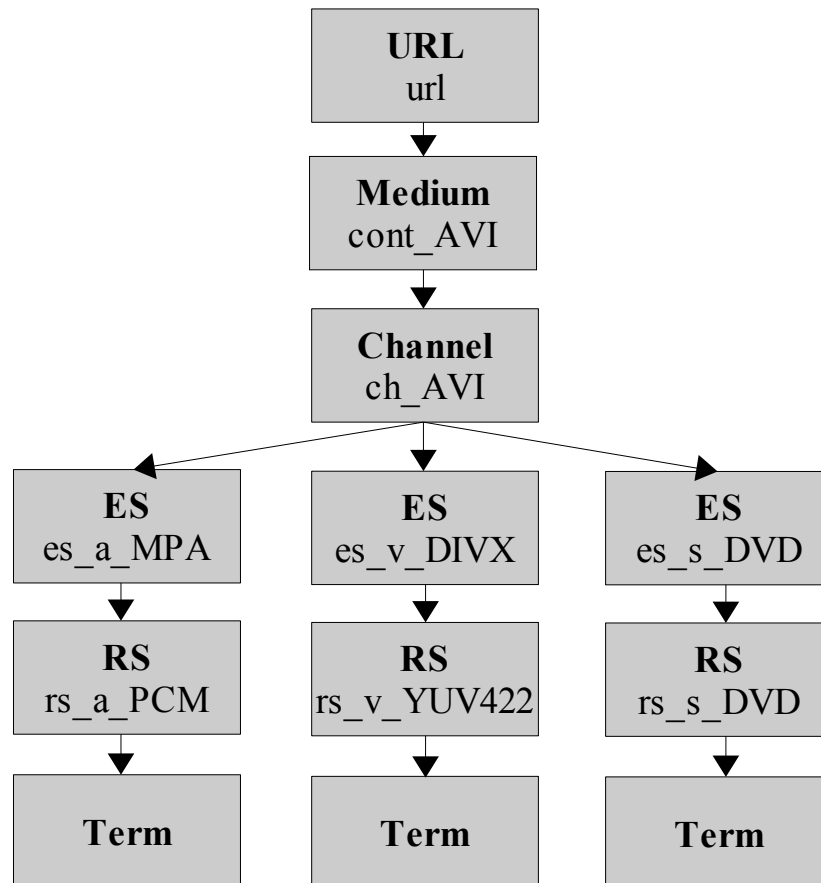
2.1. Plugin List, Process Tree, Stream Groups

There are three main expressions of MMIO with which one has to be familiar with.

The Plugin List means the list of available MMIO Plugins from where the MMIO subsystem can select plugins to be used for the playback of a given URL. This list can be queried and examined, so the application will get an idea about the capabilities of the system it runs on. Alternatively, it can decide to deregister old plugins and register new ones.

The Process Tree is probably the most important thing of the MMIO subsystem. It is a tree structure, having different kind of nodes connected together, describing the layered architecture of the multimedia content to be played back. This tree is built somewhat automatically by the MMIO subsystem, using the registered plugins of the Plugin List. The top of the tree has a URL node, containing only the URL string itself, while the bottom of the tree will have Terminator nodes, the "Visualization" nodes for every supported stream of the URL. Here is an example of an imaginary URL and the Process Tree created from it by the MMIO subsystem:

URL: http://www.server.com/Films/LOTR_Trailer.avi



As you can see, this example has an AVI file containing one video stream in DIVX format, one audio stream for the video in MPEG Audio format, and one Subtitle stream for all these, in DVD Subtitle format. All these elementary streams (ES) have a raw stream (RS) descendant, meaning that a decoder has been found for all of them, and all of them has a Terminator (Term) node, meaning that all the decoder outputs can be sent to the screen or speaker, so all of them can be played back. More about node types and their purpose later.

Having three Terminator nodes means that we have three independent streams we can play back. We could send a “Start Playback” command to all the three of these Terminator nodes, which would mean that the playback of all these streams would be started, and we would be able to watch and listen to the movie.

However, usually it's not this simple. Nothing guarantees that starting the playback this way, independently, stream-by-stream, will result in an experience where the video and the audio goes together. There must be “lipsync” as it's usually called, meaning that the audio and the video has to go together in order to have an enjoyable experience.

That's the reason why there is a third expression in the MMIO subsystem to get familiar with: the Stream Groups.

Stream Groups are logical entities, grouping Terminator nodes together. One can create a Stream Group from the three Terminator nodes of the example above, and send the “Start Playback” command to the Stream Group. The Stream Group itself will take care of the lipsync, and will make sure that all the streams grouped together in it will move and act together.

2.2. Nodes and node types in the process tree

The Process Tree is built from nodes. Every node in the tree has a type, and has a string describing the format which that node can provide. Using the Process Tree example we had before, there is a node which is a **URL node** (its type is URL), and can provide “urls” to any other nodes that are connecting to it.

The MMIO subsystem looked through its list of available plugins, and found a media handler plugin that can handle that URL node. Using that plugin, a new node was created, which is a **Media node**, describing the media contained in the URL. This new Media node was created by the media handler plugin (http handler plugin, in our example), and that plugin also checked what kind of stuff it can provide. It saw that the URL contains a file, which has the extension of “.avi”, so it reports that this node can provide an AVI container format to everyone who connects to it. As it's a Media node, everybody will be able to use the node's type-specific functions, like seeking, reading and other Media-specific functions.

The MMIO subsystem tried to “grow” this process tree, so again looked into its plugin list, and found a plugin which can handle AVI container formats (cont_AVI format strings). This is the AVI demuxer plugin. Using that plugin, new nodes of the Process Tree could be built. The first one is a **Channel node**, which describes logical channels in a given container. There is only one channel in an AVI file, but for example the MPEG Transport Stream containers, used by the digital television broadcasting systems have a lot of channels in one container, namely a lot of TV channels, and each of the channels have at least one Video elementary stream and one or more Audio elementary streams. So, the Channel node is to be able to group the available **Elementary Stream Nodes** of a given container.

It's still the AVI demuxer plugin which examined the AVI file format (using the Media node to read and seek in the file) and found out that there are three Elementary Streams (ES) in this AVI file, namely one video stream in DIVX format, one audio stream in MPEG Audio format, and a subtitle stream with DVD-format subtitles. So, the AVI demuxer plugin has created *four* new nodes into the Process Tree.

The ball is at the side of the MMIO subsystem again. It's still trying to grow the process tree, so it's looking for plugins to handle the three new Nodes of the tree. It's lucky again, and based on the plugin list, it finds decoder plugins for all the three formats, namely a DIVX decoder plugin for the es_v_DIVX format, an MPEG Audio decoder plugin for the es_a_MPA format, and a DVD Subtitle decoder/renderer plugin for the es_s_DVD format.

Using these decoder plugins for the corresponding nodes, each of the decoder plugins create one new Node into the tree, a **Raw Stream Node**. The Raw Stream (RS) terminology means decoded elementary streams, so while reading data from Elementary Streams would give encoded/compressed data, reading from Raw Streams gives decoded/uncompressed data, which can be directly shown on the screen, or sent to the audio card.

The MMIO subsystem does not give up, and still checks if it has plugins to handle the new nodes. It finds a video output plugin, which is capable of displaying YUV422 formatted images on the screen, and connects it to the video RS node. Also finds an audio output plugin which can handle the PCM audio format, and connects it to the audio RS node. The subtitle RS node will be handled by a subtitle visualizer plugin. All these plugins have one thing in common: they “eat” data from the upper layers, but do not provide any data, as they consume them by showing it. That's why they are called **Terminator Nodes**, because they terminate a given branch of the Process Tree.

Now the MMIO subsystem can see that all the leafs of the Process Tree are either Terminator nodes (so there is no task to do with them), or there was no plugin to handle them (none of the nodes are

such in our example). So, it gives up growing the tree, the Process Tree has been built successfully.

3. The MMIO API for application programmers

3.1. Forewords

Now that you've reached this chapter, you've either read through the hard part, or you've simply skipped the previous parts. If you've read through the previous parts, you know the basic idea behind the MMIO subsystem, and you'll be able to use it in a more advanced way, customize it for your needs, so the “Advanced usage” chapter will be interesting for you. However, if you haven't read through the previous parts, it's not a problem, the MMIO subsystem can be used very easily without that knowledge too, the “Quick usage” chapter is for you then.

3.2. Quick usage

We tried to create the MMIO API so that playback of a multimedia content can be done quickly, even with sparse knowledge about the MMIO API. So, if somebody does not want to do any special tricks with the multimedia content, just play it back, it's very easy to do with MMIO.

There are only some steps to follow:

- Initialize the MMIO library for the process by calling `MMIOInitialize()`
- Use the `MMIOOpen()` API to open the URL and get a Process Tree for it
- Create a Stream Group containing all the leaf nodes of that Process Tree using the `MMIOCreateEmptyStreamGroup()` and then the `MMIOAddStreamsToGroup()` APIs
- Start the playback of the Stream Group by setting its direction to “play” using the `MMIOSetDirection()` API
- Once the playback is done, destroy the Stream Group using the `MMIODestroyStreamGroup()` API
- Close the opened URL using the `MMIOClose()` API, which will destroy the Process Tree
- Uninitialize the MMIO library using the `MMIOUninitialize()` function

I know, it might seem to be a long list, but believe me, the code itself for it with comments is not longer either:

```
#include <stdlib.h>
#include "MMIO.h"

int main(int argc, char *argv[])
{
    mmioProcessTreeNode_p pURL;
    mmioStreamGroup_p pStreamGroup;

    /* Initialize MMIO subsystem */
    MMIOInitialize(".MMIO");

    /* Open the URL and create the Stream Group from it */
    MMIOOpen("file://example.avi", MMIO_NODETYPE_TERMINATOR, NULL, &pURL);
    MMIOCreateEmptyStreamGroup(&pStreamGroup);
    MMIOAddStreamsToGroup(pStreamGroup, pURL);

    /* Start the playback of the contents */
```

```

MMIOSetDirection(pStreamGroup, MMIO_DIRECTION_PLAY);

/* Now wait for the playback to stop... */
/* Sleep, or do whatever you want in here */
/* ... */

/* Destroy the Stream Group and the URL. */
/* It will eventually also stop the playback. :) */
MMIODestroyStreamGroup(pStreamGroup);
MMIOClose(pURL);

/* Uninitialize the MMIO subsystem, we don't use it anymore. */
/* The parameter means that if it detects a memory leak, it should print */
/* information about it to the screen. */
MMIOUninitialize(1);

return 0;
}

```

Of course, this code has no checks for failures anywhere, so I wouldn't call it the most fail-safe program written ever, but you can get the idea.

Also, you may have spotted that you should somehow solve what is not solved here, namely how to wait for the end of the playback. You can put there `Sleep()` calls, empty loops, and things like that. Sure, it can be done in a more professional way too, MMIO can notify you by sending you events like that, but it's discussed in the next chapter.

3.3. Advanced usage

This chapter will try to describe the possibilities of the MMIO subsystem by going through the steps required to play back a given URL. For greater understanding, it will be described what the given MMIO functions do in the background, what they are created for.

3.3.1. Initializing MMIO

The first step for any application using the MMIO API should be to initialize the MMIO subsystem by calling the `MMIOInitialize()` function.

This function allocates some internal MMIO resources for the calling process, and initializes the memory handling functions provided by the MMIO subsystem. These functions can be reached by including the `MMIOMem.h` header file and by using the `MMIOmalloc()`, `MMIOfree()` and other MMIO-prefixed memory handling functions. It's a simple memory allocation tracking set of functions, so memory leaks of plugins and the MMIO subsystem can be detected and fixed when these are used. Feel free to use these functions if you want to detect memory leaks in your own code too.

There is one optional parameter for this function, which tells where to find the MMIO plugins. If this parameter is not given (it is `NULL`), then the API will look for the MMIO plugin registry file and the MMIO plugins in the `$HOME\MMIO` directory. Otherwise it will use the path given there.

3.3.2. Managing the plugin list

Once the MMIO subsystem is initialized for the given process, it can work on its registry of plugins. The registry itself is a simple text file (at least in the current implementation), containing lines for every registered MMIO plugin. Still, it's not recommended to touch that file by hand, as it may change in the future, so better use the plugin-list handling functions provided by MMIO.

The list of registered MMIO plugins (for the current registry) can be queried by calling the `MMIOQueryRegisteredPluginList()` API, which will return the head of a simple linked list, where each of the elements of the linked list will describe one registered plugin.

As this list is allocated by the MMIO subsystem, there is an API to free this list. Use the `MMIOFreeRegisteredPluginList()` to free the memory allocated for the list.

If you want to register a new plugin or deregister an old one, there are functions for that too. The `MMIORegisterPlugin()` and `MMIODeregisterPlugin()` APIs can be used to modify the plugin-list of the current registry. For these functions, you have to know the filename of the plugin, the internal plugin-name of the plugin, the formats supported by the plugin and the importance factor of the plugin. The developer of the plugin should know these, so when he/she creates an installer for his plugin, he can use the `MMIORegisterPlugin()` function for that.

To make it even more easy to install new plugins, plugins (optionally) can contain information about what is their internal name, what formats they support, and what is their suggested (default) importance factor. If a given plugin contains this information, it can be queried using the `MMIOQueryPluginsOfBinary()` function. It will give back a linked list, containing the list of plugins implemented in the given binary plugin file. Please note that it can happen that one binary file contains more than one plugins, as the plugin internal names can differ, so there is nothing against putting more than one plugins with different internal names into one binary file.

The linked list that is created by the `MMIOQueryPluginsOfBinary()` function contains every information that is required for the plugin registration, so even unknown plugins can be registered this way. The list can be freed by the `MMIOFreePluginsOfBinary()` function, when it's not needed anymore.

3.3.3. Opening an URL

The most general descriptor for the MMIO subsystem is a URL. The URL is a simple text in the usual form: `protocol://resource_name`. The MMIO subsystem creates Process Trees from the URLs by creating first a URL node, and then looking for plugins that can handle the leaf nodes until every leaf node is a Terminator node (or no plugin can handle that node).

It should be noted here that every Process Tree has a parent node, because there is a Root node in MMIO. Every URL node must be the child of this Root node. This Root node can be got by calling the `MMIOGetRootNode()` API, and walking through the children of this node one can get the list of currently opened URLs and their Process Trees.

If one is masochist enough, the building of a Process Tree based on the URL can be done by hand, using the low-level functions, like `MMIOLoadPlugin()`, `MMIOUnloadPlugin()`, `MMIOInitializePlugin()`, `MMIOUninitializePlugin()`, `MMIOExamineNodeWithPlugin()`, `MMIOFreeExamineResult()`, `MMIOLinkPluginToNode()` and `MMIOUnlinkNode()`. However, I don't recommend doing this for anything else than debugging a newly created plugin, as this requires the deeper knowledge of how the MMIO plugin system works.

It's much easier to leave all this to the MMIO subsystem. Use the `MMIOOpen()` function to process a given URL and create a Process Tree from it, which is already linked to the Root node. This function uses the registry and all those functions mentioned above, and tries to build the Process Tree based on those informations as much as possible.

It is also possible to mix the two ways, to let the MMIO subsystem build the Process Tree to a given level, for example until it reaches the Terminator nodes, but put there a given Terminator node by hand using those low-level APIs mentioned above. It can be done by giving an `iOpenLevel` parameter smaller than `MMIO_NODETYPE_TERMINATOR`. This way, `MMIOOpen()` will return when reaches the given level, without further growing the Process Tree.

It's possible that the `MMIOOpen()` function will not know what to do when it reaches a given node. Imagine, if there are more than one plugins registered in the system that can handle a given format. In this case, one has to decide which one to use for the given node. The `MMIOOpen()` API creates a list of possible plugins for these nodes. This list is ordered by the “importance” factor of the plugins, which is a simple number in the registry, attached to the given plugin. The `MMIOOpen()` will try to use the plugins for the problematic node in the order of the list, and the first plugin which accepts the node will be used in there. To make this decision better controllable, the API has a parameter called `pfnSortList`, which (if not NULL) points to a function to be called by `MMIOOpen()` to re-sort the list of plugins to try for a given node. This way, one can set a callback in there to always move its own decoder plugin to the top of the list, no matter what “importance” is set to it in the registry, or one can program some kind of heuristics into his/her own callback to select the proper video renderer plugin for the current hardware configuration.

The very same callback is used when a plugin offers multiple output formats for a given node, so one would have to choose one from them. The callback function can reorder the list, this way influence the Process Tree creation process. Please check the description of the `MMIOOpen()` function for more information about this.

3.3.4. Walking through the process tree

Having built the Process Tree, we can examine the contents of the tree in more ways. The quickest and simplest way is to let MMIO print the structure of the process tree to the screen. The `MMIOShowProcessTree()` API is exactly for this purpose. It comes very handy for debugging purposes, to see how deeply the tree could be build and where it failed further growing.

Of course, one could walk the tree by hand, using the `pParent`, `pFirstChild` and `pNextBrother` fields of the `mmioProcessTreeNode_t` structure. However, there is a much easier method to do this. The `MMIOWalkProcessTree()` API can be used to walk through each of the nodes of the process tree and call a given callback function for them. This way, it's easy to collect information about the process tree, find given nodes in it, without knowing the internal structure of the nodes.

3.3.5. Creating stream groups

Once we have a Process Tree, we possibly have one or more Terminator nodes in that tree. Those nodes can be instructed to do the playback of their streams, to seek (if supported) in the stream, and things like that.

However, starting the playback of an audio stream and a video stream will not mean that they will

be played back in perfect sync. Most probably there will come some glitches in the playback, and the streams will simply go out of sync.

To solve this, MMIO has the concept of Stream Groups.

A Stream Group is like an intelligent container, where Terminator nodes can be put. The Stream Group will take care to keep these Terminator nodes in sync by collecting their position information all the time, and reporting them back if any of them are out of sync. The Stream Group always has one Terminator node which is the boss. This is called the Main Stream, and this is the one to which all the other streams will be synchronized.

By default, the Stream Group has a bit of intelligence, and as Streams (Terminator nodes) are added to the given Stream Group, it always selects the Main Stream based on an algorithm, so that if there is at least one audio stream then it will be the Main Stream (because it's always better to synchronize to an audio stream by speeding up or slowing down the video stream, than making hickups in the audio stream to catch up with the video).

So, to be able to do a perfect playback of a multimedia content, one has to create a Stream Group using the `MMIOCreateEmptyStreamGroup()` API, and then add the Terminator nodes of the streams to be played back into this group using the `MMIOAddStreamsToGroup()` function. Please note that this function is designed to be recursive, so if one wants to add every Terminator node of a given Process Tree to the Stream Group, it's possible to call this function only once, for the root of the Process Tree, and it will walk through the tree to collect all the Terminator nodes.

It's also possible to remove Streams from the Group by using the `MMIORemoveStreamsFromGroup()`, and if somebody is not satisfied with the Main Stream selected by the function, it can be changed by using the `MMIOSetMainStreamOfGroup()` function.

There are currently no functions to get the current Main Stream of a Group, or to check which Streams are part of a given Stream Group, but these informations can be easily got by directly accessing the `mmioStreamGroup_p` type (the Stream Group handle), and looking into the structure.

Once a Stream Group is not in use anymore, it must be destroyed by using the `MMIODestroyStreamGroup()` function. It will stop the playback if it's ongoing, and remove every Terminator node (every Stream) from the Stream Group, and free all the resources allocated for the Group.

3.3.6. Setting direction and position of Stream Groups

The playback of a given multimedia content can be started by setting the direction of a Stream Group to `MMIO_DIRECTION_PLAY`, using the `MMIOSetDirection()` API. This will result in setting the direction of every Terminator node of the Group.

The playback direction of a Stream is an integer number, where the value 1000 means the normal speed playback, 2000 the double speed playback and so on. It's also possible (theoretically) to have half speed playback (value 500), and even more, to play backwards (negative values).

However, not every multimedia format and codec supports all the directions. If an unsupported direction is requested, the `MMIOSetDirection()` API will return with the proper error code.

The current playback direction of a Stream Group can be queried by using the

`MMIOGetDirection()` API.

The same way, one can jump to a given position of the multimedia content. Some call it seeking, we call it “setting the position of the playback”. The position is in milliseconds, and can be set by the `MMIOSetPosition()` API. The current position can be queried by `MMIOGetPosition()`, and the length (the maximum position) of the current streams can be got by calling the `MMIOGetLength()` API.

It must be noted here again, that not every content can tell a length, or set position of the playback. For example, a streamed video through the internet might not be able to report its length, and also, it might not be able to seek in the audio or video stream. The functions return the proper result code in these cases.

3.3.7. Events of Stream Groups

The Stream Groups can report events to the owner, like when the direction of the playback changes (e.g. It suddenly stops because of a problem), when the stream reaches its end, when there is an error in the stream, and things like that (please check the `MMIO_STREAMGROUP_EVENT_` defines in `MMIO.h` header file).

One can subscribe for such events by using the `MMIOSubscribeEvents()` API, and giving a mask for the events in interest. Then the `MMIOGetEvent()` API can be used to wait for any of the subscribed events to arrive, with a given timeout, of course.

This way it's possible to write code that waits for the end of stream event, or which continuously prints the current position of the playback (by processing the position info event), and so on.

3.3.8. Uninitializing the MMIO subsystem

Once a process is done with using the MMIO subsystem, it can call the `MMIOUninitialize()` API. This will free all the resources allocated by the MMIO subsystem.

It has one parameter, which tells if the function should report the memory leaks it detects, if it detects any. If the parameter is `True`, and the function detects a memory leak, it will print the details of the allocated memory chunks to the screen, so one can debug the faulty plugin or application.

3.4. Simple Decoding with the MMIO subsystem

Even though the MMIO subsystem was designed and created for easy playback of media, one can always use it in other ways, thanks for its modular structure.

If all the application wants from MMIO is to decode a given media, it's easy to do.

First of all, the Process Tree has to be built for the URL. However, as we don't want the decoded audio or video frames to be sent to the outputs, it's enough to build the Process Tree until we reach the Raw Stream nodes. In other words, we can use the `MMIOOpen()` API as always, but give it an open level of `MMIO_NODETYPE_RAWSTREAM`. This way, the MMIO subsystem will not try to look for and use any output plugins, it will only build the Process Tree until it reaches the Raw Stream nodes.

To get the required Raw Stream node from the Process Tree, one can walk the tree to search for it using the `MMIOWalkProcessTree()` API, or choose the harder way to walk the tree by hand, using the tree link fields of the `mmioProcessTree_t` structure. No matter which method is chosen, the required Raw Stream node can be found easily.

Once we have it, we can ask the Raw Stream node directly to give us the raw (decoded) frame, just like the Terminator nodes would do it, but instead of being a Terminator node and displaying it, we can save it, or re-encode it into another format, or whatever is the task to do.

To ease this process, another utility function was created. It's called `MMIOGetOneSimpleRawFrame()`, and all it does is that it sets the playback speed in the given Raw Stream node to `MMIO_DIRECTION_PLAY` (this is needed so that if this API is called more than once, then the stream will give us back different frames, not the very same first frame all the time) and then gets one frame from the Raw Stream into the caller's buffer.

This API works well for decoding images or audio data, but it will most probably fail for the more sophisticated decoder plugins, for example for video decoder plugins that support B-Frames. The reason is that MMIO decoder plugins are allowed to “swallow” the buffers in which they should put the decoded data and ask for another buffer (temporarily), if needed. Now, for supporting B-Frames in video decoders, the most effective method is to use this feature, but this would make this API much more complex, so it wouldn't be called `MMIOGetOneSimpleRawFrame()` anymore.

4. Creating Plugins for the MMIO subsystem

TODO:

What is an MMIO Plugin

Binary file, has some well defined entry points, instantiable, mostly works on process tree, creates nodes, implements the node-specific functions for its nodes, can use Plugin Support functions from MMIO.

Public entry points of the plugin

Optional public entry points

Mandatory public entry points

Node types and the node-specific functions

Media nodes

Channel nodes

Elementary Stream nodes

Raw Stream nodes

Terminator nodes

5. Programming guidelines

This chapter is to be done.

Use `MMIOmem.h` and the Triton Porting Layer, where possible, so the plugins will be portable, memory leaks will be detectable and trackable.

6. API Reference

This chapter is to be done