# Event Monitoring Design

Authors: S. Kolos, I. Scholtes Keywords: monitoring, event, sampling

March 24, 2005

#### Abstract

This document presents the high-level design of the Event Monitoring framework within the ATLAS TDAQ system. The aim of the Event Monitoring component is to provide a framework in order to enable users to require samples of events or event fragments and distribute them to running monitoring tasks.

# 1 Introduction

The aim of the Event Monitoring component is to provide a framework in order to enable users to require samples of events or event fragments and distribute them to running event monitoring tasks. The monitoring system satisfies the requirements outlined in the user requirements document [1].

# 2 Event Monitoring Domain Decomposition

In order to simplify the design and implementation of the Event Monitoring system, different subcomponents have been seperated out, that have been implemented independently. The first subsystem provides a framework that enables users to implement applications, which sample events from the IO modules. In this document, this will be referred to as Event Sampling subsystem. The second subsystem is responsible for encapsulation of sampling applications, connection management and quality of service and will be referred to as Conductor. The third one provides a framework, enabling users to develop monitoring tasks that receive events from the event samplers and take care of the distribution of events to other monitoring tasks, being interested in the same events. This one will be referred to as Monitoring Tree subsystem.





### 2.1 Event Sampling subsystem

The Event Sampling subsystem provides a framework that enables users to develop event sampling applications, providing statistical samples of events flowing in the TDAQ system extended to the ROD crate. There is an independent Event Sampling subsystem per ROD, ROC or SFC crate, each providing the same functionality and using the same interface with the TDAQ Conductor subsystem. The Event Sampling subsystem provides the following facilities:

- handling request originating from the Conductor subsystem. Each valid request is composed of an address part specifying the source of the events, a sampling criteria and a CORBA callback reference to the monitor sending the request.
- for each selection criteria, the event sampling application starts a thread, sampling event data and pushing it to the appropriate event monitoring task.

### 2.2 Conductor subsystem

The Conductor subsystem separates the Event Sampling subsystem, which is directly connected to the TDAQ system, from the Event Monitoring subsystem and provides connection management as well as quality of service (QOS) facilities. Unlike as in former implementations, the Conductor system does not provide the distribution of event data itself, as this is now done following a peerto-peer paradigm. The Conductor subsystem provides the following facilities:

- it takes connection requests from monitoring tasks and passes them on to the appropriate event sampling application
- it provides error recovery strategies, in case event sampling applications or event monitoring tasks fail to respond due to network outage or program crashes
- it controls distribution performance of event sampling applications, adapting the event rate of the sampling application to the needs of monitoring tasks
- it is able to restart during the process of event sampling, without having influence on stability or performance of the event distribution process

### 2.3 Monitoring Tree subsystem

The Monitoring Tree subsystem provides a framework, that allows users to implement custom monitoring tasks receiving events or event fragments from the Event Sampling subsystem. Monitoring task may either receive event data directly from event sampling applications or from other monitoring tasks, that are arranged in a tree. The distribution of data in this tree is done transparently for the user, so users do not have to be aware whether they receive events from a sampler or a peer monitor. A monitoring task has the following functionality:

• it receives event data either from the Event Sampling subsystem or from other monitoring tasks in the tree

- it provides a buffer for event data being received from the Event Sampling subsystem or from other monitoring tasks in the tree
- it accepts a given number of child monitoring tasks, forwarding event data to them
- it recognizes malfunctioning child monitors and notifies the Conductor subsystem about it
- it detects buffer underruns or buffer overflows and notifies the Conductor subsystem about it



# 3 Event Monitoring architecture

This section describes the high level design of the Event Monitoring architecture and the interaction of the different subsystems. Before introducing the Event Monitoring architecture, some auxiliary definitions are given. They include the formal definition of the data types for event selection and event representation.

## 3.1 Event Sampling parameters

User monitoring tasks have the possibility to define the source of event sampling and certain characteristics of the events they are interested in. According to [1], the neccessary parameters for the sampling definition include:

- event sampling address
- event selection criteria

The sampling address defines the of the source event sampling and consists of a sequence of key, value pairs. Figure 3 shows the OMG IDL definition of the sampling address. The selection criteria is used to define the range of interesting events. This selection is based on serveral events' characteristics as defined in [1]. The OMG IDL declaration of the selection criteria is shown in figure 4. Users can

Figure 3: sampling address definition (OMG IDL)

```
typedef sequence<AddressComponent> SamplingAddress;
struct AddressComponent
{
   string key;
   string value;
};
```

select events according to different characteristics, depending on the level of selection.

- ROD-Level: L1 TriggerType and Detector Type
- ROS-Level: L1 TriggerType and Detector Type in ROB Header
- SFI-Level: L1 TriggerType and L2 TriggerInfo

So, at all levels of selection, users can define their selection using a two-word selection criteria. In each word, a selection can be made by specifying one masked value. In the masked value one can either specify that this parameter should be ignored (making events match, independent of that value) by setting ignored to true, or set ignored to false and specify an exact value that has to be matched.

The final two-word selection criteria will be an AND composition of two oneword selection criteria. Additionally, users may specify a long value statistics. A value of x for statistics shall cause every x-th event, in other words  $\frac{100}{x}\%$ of all events to be sampled. If used, this value must be greater than zero.

Figure 4: selection criteria definition (OMG IDL)

```
struct MaskedValue
{
    long value;
    boolean ignore;
};
struct SelectionCriteria
{
    MaskedValue lvl1_trigger_type;
    MaskedValue lvl2_trigger_info;
    MaskedValue detector_type;
    MaskedValue sc_status_word;
    long statistics;
};
```

# 3.2 Event definition

The OMG IDL declaration of the event is shown in figure 5.

Figure 5: The OMG IDL Event definition

typedef sequence<unsigned long>Event;

#### 3.3 Architecture definition

Figure 6 shows the interaction of the different subsystems of the Event Monitoring system. The EventConductor class implements the functions defined in



Figure 6: Distribution sequence diagram (UML notation)

the OMG IDL declaration shown in figure 7. It receives and stores registrations from event sampling applications and passes requests from monitoring task to the appropriate event sampling applications. In case of a crash of any monitoring task, it repairs the tree of monitoring tasks by rearranging it. If a sampling applications crashes, it notifies the appropriate tree of monitoring tasks that has been attached to this sampling application about the crash, so any monitoring task waiting for event data can cleanly exit. It also stores a local representation of the tree in order to repair it in case of the crash of a monitoring task. The class EventSampler implements functions that have been defined in the OMG IDL definition shown in 7. It handles subscriptions of monitoring tasks and manages a pool of threads, which will be used for event sampling for incoming subscriptions of monitoring tasks. The EventSampler receives one monitor task subscription per sampling thread and pushes events to this monitor task, which is called root monitoring task. Any consecutive monitoring tasks, willing to receive events from the same sampling thread will not be connected directly to EventSampler by the EventConductor, but to the root monitoring task. The root monitoring task can either accept connections to other monitoring tasks itself, registering them as child monitoring tasks locally, or pass them on to its child monitoring tasks. This will result in a tree of monitoring tasks, the type of the tree depending on the maximum amount of child monitoring tasks, each

monitoring task will accept. Figure 2 shows the example of a binary tree of monitoring tasks. The tree type is configurable by the user and can be passed as argument to the EventConductor application. If events are received from a sampling application, the MonitoringTree framework will take care of forwarding of events to child monitoring tasks. It also provides buffering facilities for the local user If the MonitoringTree framework recognizes a buffer underrun or a buffer overflow, it sends an appropriate adaptation message to EventConductor, which will cause EventSampler either to speed up or slow down the appropriate sampling thread, if possible.

## 3.4 IDL declaration

Figure 7 shows the OMG IDL declaration of the Event Monitoring System.

#### Figure 7: The Event Monitoring System interfaces (OMG IDL)

```
#include <ipc/ipc.idl>
module EventMonitoring {
  typedef sequence<unsigned long>Event;
     typedef sequencedumps/vent;
struct MaskedValue { long value }
struct SelectionCriteria {
MaskedValue detector_type;
MaskedValue lvll_trigger_type;
MaskedValue struc_vord;
long statistics;
      };
      typedef sequence<AddressComponent> SamplingAddress;
struct AddressComponent {
         string key;
string value;
       };
        interface EventMonitor;
       struct MonitorInfo {
SamplingAddress address;
SelectionCriteria criteria;
EventMonitor reference;
       }:
        typedef sequence<MonitorInfo>MonitorList;
        exception NoResources { };
       exception BadAddress { };
exception BadCriteria { };
exception NoMoreMonitors { };
       exception NoMoreKonitors { };
exception NotFound { };
exception NotFound { };
exception NotConnected { };
exception NotConnected { };
exception MaximumReached { };
exception AlreadyExit { };
exception AlreadyConnected { EventMonitor monitor; };
        interface EventMonitor {
              srface EventMonitor {
    void push_event( in Event e );
    void get_children( out MonitorList children );
    void add_child( in EventMonitor monitor );
    void remove_child( in EventMonitor monitor );
    void sampler_exit( );
       };
        enum Direction { SpeedUp, SlowDown };
       interface EventSampler : ipc::servant {
             in EventMonitor monitor )
raises ( BadAddress, BadCriteria, NotConnected );
               in EventMonitor monitor )
               raises ( NotConnected );
void adapt_sampling_rate( in SamplingAddress address,
in SelectionCriteria criteria,
              in Direction dir )
raises ( NotFound, MaximumReached );
void get_monitors( out MonitorList monitors );
       };
       interface Conductor : ipc::servant {
    const string name = "Conductor";
    void connect_sampler( in SamplingAddress address,
              raises (BadAddress, BadCriteria, NoResources);
void remove_monitor( in SamplingAddress address,
in SelectionCriteria criteria,
                                                   in EventMonitor monitor )
               raises ( NotFound );
void adapt_sampling_rate( in SamplingAddress address,
                                                            in SelectionCriteria criteria,
in EventMonitor monitor,
                         in Direction dir )
raises ( NotFound, NotAllowed, MaximumReached );
};
};
```

# 4 System APIs

### 4.1 Event Sampling API

In order to provide a simple API for the common user while at the same time delivering a powerful API to users who want to have full control about the sampling process, the API has been split into two parts. The Pull model API will be suitable for all users willing to develop a sampler application without having to deal with thread internals. Event Distribution is done by implementing a function sampleEvent, which will be used by an internal thread to pull events from the user. The Push Model API is suitable for all developers that want to have full control about the sampling process and any thread internals. Here, users may define a custom thread, using an EventChannel reference provided by the API to push events to the monitoring tasks. Both API will be discussed in detail in the following paragraphs.

#### 4.1.1 Pull model API

The user will have to provide the main entry point of the sampling application by instantiating a class, which is called EventSampler. The EventSampler constructor expects a minimum of three arguments as it is shown in Figure 8. The first parameter defines the TDAQ partition, to which the sampler will belong to. The second one provides the address, for which this sampler will be responsible. All the sampling requests with the same address will be forwarded by the Monitoring Conductor sub-system to this sampler. The third argument is a pointer to the user's sampling factory object, that will be used for event channel creation. This parameter is used to hide thread management aspects from application developers in order to simplify their work and to have sampler implementations less error prone. The optional fourth argument specifies the maximum number of channels this sampler might create. One so-called EventChannel will be created for every selection criteria being sampled by this sampling application. As one monitor connection will be allowed per EventChannel, this maximum number will influence the distribution performance of the sampling application. If not specified, the default value of 100 will be used. As the EventSampler class is common among the Push and Pull API, in this section we use the constructor suitable for the pull model.

A thread management technics, which is used by EventSampler class adheres

Figure 8: Event Sampler class (Pull API)

to the Factory pattern, which requires an application developer to implement

two abstract interfaces. In the case of Event Sampler these interfaces are called PullSampling and PullSamplingFactory, which are shown in Figure 9.

Figure 9: PullSampling and PullSamplingFactory interfaces

In order to implement the PullSamplingFactory interface an Event Sampler developer has to inherit his own class from the PullSamplingFactory and implement the startSampling virtual function. A possible implementation is shown in Figure 10.

An implementation of the startSampling method has to return an instance

Figure 10: A possible PullSamplingFactory implementation (with some pseudo code)

of the user specific class, which must implement the PullSampling abstract interface. An instance of the PullSampling class is responsible for performing the actual event sampling, i.e. for the interaction with the Data Flow system. The PullSamplingFactory is necessary to give to the developer some flexibility in defining signature for the PullSampling object constructor. Figure 11 shows a possible implementation of the PullSampling interface, which needs the SelectionCriteria parameter and also another one, which defines an initial buffer size. This is of course just an example and an Event Sampler developer is free to define any other parameters, which he may need to be provided for the PullSampling object. Please note, that there is a bijective relationship between instances of PullSampling and threads sampling events for a selection criteria. There are several things, which have to be done for a proper implementation

Figure 11: Example of the PullSampling interface implementation (with some pseudo code)

```
class MyPullSampling : public PullSampling
   public:
       MyPullSampling( const SelectionCriteria & criteria, int buffer_size )
       £
          unsigned long * buffer = new unsigned long[buffer_size];
          INITIALIZE DATAFLOW system WITH criteria
      }
       ~MyPullSampling( )
       {
          delete[] buffer;
       }
       void sampleEvent( EventChannel & ch )
       {
         long event_size = READ_EVENT_TO_THE_BUFFER( buffer );
         ch.pushEvent( buffer, event_size );
};
```

of PullSampling class. It is assumed that all the necessary initialisation, which has to be done to prepare for the event sampling has to be done in the constructor of the user defined MyPullSampling class. When the destructor of that class is called, this indicates that the sampling is not necessary anymore and MyPullSampling has to perform a proper clean up procedure. The main working method of the MyPullSampling class is the sampleEvent function, which is responsible for reading an appropriate event from the Data Flow system and pushing it to the EventChannel by using the pushEvent function. This technics is used to avoid complicated memory management in the user code, which would appear if the sampleEvent function was declared as returning event. Finally Figure 12 shows an example of the main function for the Event Sampler application. The function wait() of class EventSampler might be used to block the current thread until the event sampling process is stopped with function stop(). If initialization of the sampler fails for some reason, CannotInitialize is thrown from the constructor of class EventSampler. Figure 12: Example of the main function for Event Sampler application (with some pseudocode)

```
STOP_SAMPLING
{
    sampler->stop();
}
int main()
{
    IPCPartition partition = PARTITION;
    max_channels = 100;
    SamplingAddress address = ADDRESS;
    emon::EventSampler temp(partition, address, new MyPullSamplingFactory(), max_channels);
    sampler = &temp;
    sampler->wait;
    return 0;
}
```

#### 4.1.2 Push Model API

While the Pull model will be suitable for many users, some users might need more control of the event sampling process. Just like in the case of the Pull Model, the user will have to provide the main entry point of the sampling application by instantiating a class called EventSampler. The EventSampler constructor expects a minimum of three arguments as it is shown in Figure 13. The first parameter defines the TDAQ partition, to which the sampler will belong to. The second one provides the address, for which this sampler will be responsible. All the sampling requests with the same address will be forwarded by the Monitoring Conductor sub-system to this sampler. The third argument is a pointer to the user's sampling factory object, that will be used for event channel creation. The optional fourth argument specifies the maximum number of channels this sampler might create. One so-called EventChannel will be created for every selection criteria being sampled by this sampling application. As one monitor connection will be allowed per EventChannel, this maximum number will influence the distribution performance of the sampling application. If not specified, the default value of 100 will be used. As the EventSampler class is common among the Push and Pull API, in this section we use the constructor suitable for the push model.



Just like in the case of the Pull Model, the Push API uses the Factory pattern for thread creation, which requires an application developer to implement two abstract interfaces. In the case of Push Model these interfaces are called PushSampling and PushSamplingFactory, which are shown in Figure 14.

Figure 14: PushSampling and PushSamplingFactory interfaces

In order to implement the PushSamplingFactory interface an Event Sampler developer has to inherit his own class from PushSamplingFactory and implement the startSampling virtual function. A possible implementation is shown in Figure 15.

An implementation of the startSampling method has to return an instance

Figure 15: A possible PushSamplingFactory implementation (with some pseudo code)

of the user specific class, which must implement the PushSampling abstract interface. An instance of the PushSampling class is responsible for performing the actual event sampling, i.e. for the interaction with the Data Flow system. The PushSamplingFactory is necessary to give to the developer some flexibility in defining signature for the PushSampling object constructor. Figure 16 shows a possible implementation of the PushSampling interface, which needs the SelectionCriteria parameter and also another one, which defines an initial buffer size. This is of course just an example and an Event Sampler developer is free to define any other parameters, which he may need to be provided for the PullSampling object. Please note, that there is a bijective relationship between instances of PullSampling and threads sampling events for a selection criteria.

Again, there are several things, which have to be done for a proper implementation of PushSampling class. All necessary initialization which has to be Figure 16: Example of the PushSampling interface implementation (with some pseudo code)

```
class MyPushSampler : public emon::PushSampling,
                      -
public OWLThread
{
 public:
     emon::EventChannel * channel_;
    MyPushSampler( const emon::SelectionCriteria &sc,
                    long custom_arg1, long custom_arg2,
                    emon::EventChannel * channel )
             : channel_( channel )
     {
        INITIALIZE DATA FLOW
        // this will call run_undetached in a seperate thread
        start_undetached();
   };
   void * run_undetached( void * )
   ſ
        while( NOT TERMINATED )
        Ł
            RETRIEVE EVENT
            channel_->pushEvent(EVENT, SIZE);
       }
       return 0;
   7
    ~MyPushSampler()
    ſ
      TERMINATE AND CLEAN UP
   }
};
```

done to prepare for event sampling, as well as the initiation of the sampling thread, has to be done in the constructor of the user defined MyPushSampling class. When the destructor of that class is called, this indicates that the sampling is not necessary anymore and the sampling thread has to perform a proper clean up procedure. All thread issues and actual event distribution is left to the user. Event distribution can be done by calling the function pushEvent of the EventChannel object. Finally Figure 17 shows an example of the main function for the Event Sampler application. The function wait() of class EventSampler might be used to block the current thread until the event sampling process is stopped with function stop(). If initialization of the sampler fails for some reason, CannotInitialize is thrown from the constructor of class EventSampler. Figure 17: Example of the main function for Event Sampler application (with some pseudocode)

```
STOP_SAMPLING
{
    sampler->stop();
}
int main()
{
    IPCPartition partition = PARTITION;
    max_channels = 100;
    SamplingAddress address = ADDRESS;
    emon::EventSampler temp(partition, address, new MyPushSamplingFactory(), max_channels);
    sampler = &temp;
    sampler->wait;
    return 0;
}
```

### 4.2 Monitoring Tree API

The main entry point for a monitoring task application is the select function, which is defined in the emon name space as it is shown in Figure 13. The argument buffer\_limit specifies the maximum amount of events the buffer in this monitoring task can hold. If no buffer\_limit is specified, a maximum buffer-size of 1000 events is used. The parameter buffer\_limit will affect memory usage of the monitoring task.

This function allocates the instance of the EventIterator class, which a devel-

Figure 18: Monitoring task entry point

oper can use to retrieve events, which have been sampled from the sampling address address and satisfy the selection criteria criteria. This function may throw several exceptions in case of either address or criteria is invalid, conductor is not available, sampler does not accept any more monitoring tasks or this monitor cannot be added to the monitoring tree for some reason.

As shown in figure 19, the EventIterator class provides two functions to retrieve events. The nextEvent function returns a new event or suspends for the timeout milliseconds before throwing the NoMoreEvents exception in case there are no more events in the buffer. If no timeout is specified, it will wait block until a new event is available, waiting infinitely. The tryNextEvent function returns a new event or throws a NoMoreEvents exception in case the underlying event buffer is empty. The availableEvents function returns the number of events currently available in the monitoring tasks's buffer.

### Figure 19: Event Iterator interface

# List of Figures

1	Monitoring Collaboration Diagram (UML notation)	2
2	A binary example for a tree of monitoring tasks	4
3	sampling address definition (OMG IDL)	5
4	selection criteria definition (OMG IDL)	5
5	The OMG IDL Event definition	5
6	Distribution sequence diagram (UML notation)	6
7	The Event Monitoring System interfaces (OMG IDL)	8
8	Event Sampler class (Pull API)	9
9	PullSampling and PullSamplingFactory interfaces	10
10	A possible PullSamplingFactory implementation (with some	
	pseudo code)	10
11	Example of the PullSampling interface implementation (with	
	some pseudo code)	11
12	Example of the main function for Event Sampler application	
	(with some pseudocode)	12
13	Event Sampler class (Push API)	12
14	PushSampling and PushSamplingFactory interfaces	13
15	A possible PushSamplingFactory implementation (with some	
	pseudo code)	13
16	Example of the PushSampling interface implementation (with	
	some pseudo code)	14
17	Example of the main function for Event Sampler application	
	(with some pseudocode)	15
18	Monitoring task entry point	15
19	Event Iterator interface	16

# References

 M. Caprini, D. Francis, R. Jones, S. Kolos, J. Petersen, L. Tremblet User Requirements for the Online Monitoring of the ATLAS DAQ/EF Prototype -1, 2000