Timing Workshop Summary

Summary of the Timing Workshop held 15th February 2008

Present: J. Serrano, J. Lewis, P. Alvarez, F. Di Maio, K. Kostro, B. Puccio, U. Krause, R. Bär, K. Höppner, T. Hoffmann, C. Girerd, D. Autiero, B. Carlus, T. Rohlev, A. Borga, H. Weibel, J. Dedic, I. Verstovsek, P. Loschmidt, G. Gaderer, N. Kerö, J. Pietarinen, B. Todd.

Executive summary:

Several institutes and companies were invited to CERN in order to share ideas about future timing systems and launch a collaboration to benefit from common development. There were presentations by the different actors in the morning and discussion in the afternoon. The discussion was organized so as to converge on a certain technical solution and split the work to be done in work packages. The most promising technical solution turned out to be a combination of synchronous Ethernet and some kind of two-way scheme to evaluate delay, such as the Precision Time Protocol (PTP, IEEE1588). Work packages to do a preliminary study in this direction were created and distributed. Progress on these different fronts will be evaluated in a new meeting in about 3 months.

Workshop AM Agenda:

- Introduction to the Workshop [JS]
- CERN – Background and Requirements [JS]
- GSI – Background and Requirements [UK]
- Institut de Physique Nucleaire de Lyon (IN2P3) – Background and Requirements [CG]
- ITER – Scope and Schedule [FDM]
- FERMI LLRF – Sub-Nanosecond Timing [TR]
- Zurich University for Applied Sciences – Synchronous Ethernet and IEEE 1588 [HW]
- Cosylab – Supplier of Control Electronics for Particle Physics [JD]
- Austrian Academy of Sciences / Oregano Systems [PL]
- Micro-Research Finland [JP]

Workshop AM Summary

Introduction to the Workshop [JS]

There are two motivations of the timing workshop from CERN's point of view:

1. CERN will shortly start working on a new timing system to accommodate the requirements of the injector chain upgrade.
2. The old CERN model is no longer viable: collaboration with others to share experiences and present a common approach is in the interests of CERN.

CERN – Background and Requirements [JS]

CERN timing controls several interlinked machines and related experiments; this involves the fabrication of several different 'cycles'. A central beam and cycle manager creates these and distributes them to all of the front end equipment. For better efficiency, the system is multi-cycling. There are three different types of timing at CERN:

GMT – UTC synchronous 40MHz clock, with a granularity of 1 millisecond, potential for 8 messages per millisecond, all events get acknowledged at the next millisecond. The millisecond itself is an event. Overall jitter is less than one nanosecond; wander is about 10-15ns per year for long links, 2-3ns for shorter links. GMT is not required to be better than one microsecond.

TTC – This is the simple bunch crossing frequency transmission system.

BST – 40.079MHz, bunch crossing frequency with two channels A and B injected into this, receivers differentiate between the different elements.

GMT is the key system under question here. A Symmetricom system gives a pulse per second and a phase locked 10MHz, several further signals are generated from these using Phase Locked Loops. All of these frequencies are available in VME. To transmit the timing, event tables are first written into front ends, which receive UTC through NTP or GPS, this is synchronised to give the correct time of day in seconds. Sub-second counting is managed by the physical layer of the timing link. To achieve the best accuracy, cable delays are manually corrected. The resulting system is completely UTC synchronous, the CBCM cycle manager allows operators to draw a primary and backup cycle.

Flaws:

1. *Lack of bi-directionality:* forces a parallel data path to allow control and diagnostics, active cabling delay compensation schemes need bi-directional traffic.
2. *Lack of bandwidth:* 500 kbps means that different timing networks have to be used for different machines.

Next steps:

1. Identify commonalities, launch collaborative effort.
2. Promising Technologies: Ethernet PTP etc.
3. Be as Standard as Reasonably Possible
4. Be completely Open Source
5. Time Scheduled versus Event Based Systems
6. Unclear advantages of plain Ethernet and protocols stacks

GSI – Background and Requirements [UK]

GSI is a particle physics and medical institute having several different machines including LINAC, Storage Ring and Synchrotron. The control system and timing system were designed in the mid 80s. GSI intends to extend their facilities, introducing two new synchrotrons, each having ~1km circumference, four storage rings and some smaller machines. This will extend the research to higher energies, and anti-proton studies; the current LINAC and synchrotrons will become injectors to this mew complex.

A redesign of the timing system is needed for this expansion. GSI has similar multiplexing, multi-cycling as CERN, running several experiments in parallel. Experiments at GSI run from just one or two days up to a few weeks. In the enlarged complex the main ideas will remain the same.

The requirements on GSI are for an event based timing system, having ~50ns alignment between locations, compensation is already required to achieve this. Events are placed to microsecond granularity. Having these events separated by 10us is sufficient, 1us would be optimal.

Cycles are constructed from basic building blocks. Local delay generators are not used; in the front end equipment the arrival of an event directly triggers an activity. The Bunch Timing System (BuTIS) is the source of the timing for synchronisation, with GSI timing completely phase locked to the BuTIS. This source has a high frequency 200MHz clock and a lower 100kHz clock.

Institut de Physique Nucleaire de Lyon (IN2P3) – Background and Requirements [CG]
IN2P3 have recently developed distributed data acquisition systems, based on distributed smart sensors. This implements a common base having an embedded processor, a specific interface board is then designed for each experiment, all use the same dedicated bus for synchronisation. There are 1200 sensors connected on an Ethernet network. Sensors detect an event on arrival. Basically the events need to be time stamped to reconstruct any physics event.

IN2P3 has experimented with a PTP implementation based on a mixed hardware and software in a NIOS soft-core CPU. NIOS is implemented in conjunction with a Gb Mac IP, implementing the complete timing protocol in a single FPGA. A reconstruction of the received clock gives some *pico-second* accuracy, by recovering the 125MHz encoding clock. They are very interested in creating a Gb Ethernet Switch.

ITER – Scope and Schedule [FDM]
ITER prepared and closed the conceptual design in 2006/7. 2008 sees the start of the engineering specification for the realisation of the machine. The basic element of interest is the Control Data Access and Communications (CODAC): This CODAC is one of three core networks. The other two are the Interlock Network and the Safety Network.

The data which is required for transmission will include Events / Clock / Video and Synchronous Data. In addition there needs to be a predefined system which can be sent to the constructing factories to test the work that they do on building individual components.

The development of ITER is considerably hypothetical: Between now and 2010 the specifications will be completed, with the start of construction between 2010 and 2013. The initial work states that the plasma runs ought to start around 2016, after a commissioning period starting in 2014
Requirements:

1. Provide a set of general specifications
2. Provide some first ideas
3. Launch a R&D Contract in the near future (1-2 years) to start the work on the CODAC.

Note that machine Protection is completely isolated from the transmission of the timing.

FERMI LLRF – Sub-Nanosecond Timing [TR]
Fermilab has shown considerable effort in theoretical work to phase locking a distributed system to a reference clock. The absolute requirement is to be aligned to within 50fs. This is impossible, the reference cannot be preserved. A solution is proposed using a completely homemade solution based on Xilinx FPGAs with a complex feedback and correction mechanism.

Zurich University for Applied Sciences – Ethernet [HW]
Examples are shown for synchronisation and data transmission using the standard IEEE-1588, Precision Time Protocol (PTP). This is multi-cast, but a single cable point-to-point can allow an

accurate determinism of the delay. With commercial chips it is sometimes not possible to send a timestamp with the same command at the synch.

One Step: stamp and time in a single frame.
Two Step: stamp first then a second frame with time as a follow-up.

There are two sources of error, the delay of the line and the offset of the clock. The delay and offset can be calculated just with a pair of stamp operations, one in both directions.

If another network element is added into the link between the master and the slave, there is an additional boundary clock inferred between them. This then performs the synch between the left and the right equipments. The overall master clock which is the source of the time is called a 'grandmaster'. This clock is normally selected automatically based only relative measure qualities by the networks.

IEEE 1588 v2 allows for 'transparent clocks' with an *offset* field, allowing for more precise information regarding the delays. In addition, each of the links can be recalculated on-the-fly, this allows any changing topology due to transmission changes to be accommodated.

In addition Synchronous Ethernet can be applied to the gigabit Ethernet. In this case there is a single primary reference clock. Clocks are recovered at each stage, and fed forward. Each end node receives the same traceable clock. A Synchronous Ethernet switch will only use the best suited clock as the transmission clock for each transmission link. Operation and Administration Messages (OAM) can be used to give each node the information about the quality of the clocks being used.

Additionally Synchronous Ethernet is free of problems due to network traffic, PTP provides both a frequency and the time of day, but can have issues with the network traffic, as there must be a certain bandwidth to allow the synchronisation to be successfully carried out.

Several silicon vendors offer PHY chips to implement this, note that both the IEEE and the synchronous Ethernet use only the PHY, none of the higher layers of the OSI are used. In Principle, given a good design, as long as the link is established there is no worry about network overloading.

Cosylab – Supplier of Control Electronics for Particle Physics [JD]
Cosylab provide a lot of control electronics for several complexes around the world. They have about 75 FTE staff, with a division for accelerator and beam-line controls.
An example of their products is *MicroIOC*, this is a complete solution, hardware, software documentation, etc. All work has the potential to be open-source, they showed some systems working at 500MHz having just some 50ps error.

Austrian Academy of Sciences / Oregano Systems [PL]
Oregano provides design work from specification through design, with test and series production, including IP Cores and Embedded Systems. Clock distribution simulation is possible, in a simulated network based in OMNET ++, which is open source. This is a discrete event simulator; C code can be plugged in for each point, making some real determination of the way in which a system should work. An example of their work is SYN1588 implementing PTP (IEEE1588). A fault tolerant system which can allow for breaks and changes has also been considered. In this case security has an overhead, but

models have been developed to consider it. Generally the system runs at 1 GHz, one nanosecond resolution is feasible, standard Ethernet has shown sub 10ns accuracy.

Micro-Research Finland [JP]

A home grown company specialised in timing hardware. The principle is a simple clock source driving several slave devices.

In an example shown an event generator takes a high frequency N*125M RF clock, it's divided by the event generator to give a 125MHz reference which are used for Virtex-II Rocket-IO transceivers, following market progressions a system has been developed including bi-directional communications. This implementation has 255 event codes which can be used, events are sent at a maximum rate of 125MHz, they can be generated from external hardware sources, sequences of several thousand elements can be loaded into the generators.

Some examples of typical VME hardware demonstrate that the same ideas have been used from start to finish, some examples also exist in compact PCI. Resolution is ~8ns, jitter is measured at *<25ps* RMS, although the VME event receiver has a better jitter, *<5ps* for the CML outputs. Each receiver can synchronise to sub-ns accuracy. Measurements were shown that gave an indication of the accuracy, where fibre optic temperature effects were characterised using the basic hardware.

Common Conclusions and Questions from the Presentations

What data is being transferred (and in what directions)?
What are the basic requirements of accuracy, granularity, jitter, wander?
What are the final quantities of components that are required?
What is the timing system synchronised to?
What is the deadline for the work?
What are the implementations constraints? VME / Linux etc?
What will be the impact of obsolescence and enhancements on the availability of products in some years?

	Latency	Interval	Jitter	Align.	Sources	End Users	Distance	When?
CERN FIP	~ms	~ms	<ms	<1us			~2km	2010-2012
CERN GMT		~1 us	<1 ns		1 CBCM	~3000	~10km	2008-2010
GSI		1 – 10us		50ns	1 CBCM	2000	~5km?	2008-2013
IN2PL3				10ns	N/PTP		<1km	Soon
ITER SDN	1 ms		50 us		80	50	<1km	2008-2011
ITER TCN	10 us		<10 ns	10ns	50	50	<1km	2008-2011
FERMI			50 fs	<<1ns	1	21	200m	2008

Blue = Similar Requirements

Workshop PM Agenda:

1. Technical Issues
    a. Synchronous Ethernet plus PTP (IEEE 1588) (pro & con).
    b. Alternative Solutions in the Short Term.
    c. Form Factors
    d. Interest in Higher Level Protocols

2. Project Management Issues
    a. CERN driving project, remaining organisational structure.
    b. Open source definition.
    c. Timescale of work.
    d. Quantities of components / machine requirements.
    e. Financing.

3. Initial Thoughts Regarding Potential Work Packages.
    a. Synchronous Ethernet plus PTP (IEEE 1588) Background (theoretical)
        i. Manage project.
        ii. User needs gathering.
        iii. Fibre/copper choice.
        iv. Reliability studies.
        v. Is this constraining us in any way?. Can we mix commercial products with this? Verify this will stay standard.
        vi. Scalability system-level simulation for this solution.
        vii. Study Robustness of Cascaded Systems (includes analysis of switches).
        viii. Concrete Overall Design Specification.
    b. Synchronous Ethernet plus PTP (IEEE 1588) Prototype / Proof of Concept
        i. Drivers.
        ii. Test Programs.
        iii. Generator.
        iv. Switches, Study Combining of Ethernet and IEEE 1588 in Switch.
        v. Receiver.
    c. Synchronous Ethernet plus PTP (IEEE 1588)
        i. New Standardisation?
        ii. Dissemination.

Workshop PM Summary

ITER is interested in forming collaboration with CERN to study and implement some timing system. The direction of study seems appropriate for ITER.

IN2P3 is very much in favour of collaboration at a system level; perhaps the time scales are slightly different, in that some elements are perhaps needed slightly in advance of those that are required by others. IN2P3 is ready to make some tests with this kind of equipment; particular interest has been shown for the Switch.

Oregano Systems feels that the proposal falls in line with their expectations.

CosyLab feels that this is a good starting point, they would be happy to assist.

GSI welcomes the development, sharing singular time scale between this and the FAIR, Gb Ethernet has shown good promise, this approach has potential. Timing requirements are not so strict for this application, microsecond precision is enough. GSI must rationalise between a 'light' and 'heavy' system. The system must also be cost effective.

Micro-Research considers that the base is good; determinism may be a question, as is the potential obsolescence of the Gb Ethernet. It depends upon marked forces. PHY chip vendors may focus on high-end industrial products, these products should become more interesting. Latency wasn't deterministic in some examples for specific PHY; however, implementing a PHY within an FPGA has shown that it is possible to avoid this.

Zurich University agrees in principle with the approach, but needs to see more information in clarifying the benefit between Ethernet and IEEE-1588.

1. **Technical Issues**[1]

a. **Synchronous Ethernet plus PTP (IEEE 1588) (pro & con).**

It's more or less been agreed that the PTP plus Ethernet is an interesting proposal. However, it is unclear what the constraints apply when mixing different Ethernet types. It's known that the Master Slave concept is only an issue in the copper cabled Ethernet application, where a bi-directional link is shared. In the optical case this is a problem which doesn't arise. In this case there may be limits to scalability in the PTP network. A system simulation could be carried out to ensure that the system can be realised, this would ensure that *in principle* the specified alignment is possible. Work packages 3, 5, 6 and 7 are concerned with this.

Note that for PTP if a PHY can be manually set into Master Mode, the auto-negotiation will result in the other connected PHYs switching into slave mode. The selected Synchronous Ethernet master must also be the PTP master. But, the PTP master is normally selected automatically - If this was the case it would significantly degrade the accuracy of the system. As it stands now there is no alternative to be considered for CERN other than having a single node as the master for both.

The physical implementation needs to be a star-configuration for CERN, another type of hierarchy would be possible, but the number of cascaded PLLs for Synchronous Ethernet may be an issue.

A static network topology is employed at CERN and elsewhere, thus when switching between masters during a failure, the Synchronous Ethernet would likely become misaligned. PTP on the other hand can transfer automatically. Therefore there cannot be any automatic swapping between the two systems, this is an acceptable compromise.

Fibre optics allow for a passive fan-out, but copper is an alternative. It is widely known that high symmetry copper is a problem, as cables would need to be fabricated with a guaranteed pair length. Work package 3 will study this.

b. Alternative Solutions in the Short Term.

An alternative is to use Xilinx Virtex-II PRO, with a completely home-built solution. In this case it's no longer a standard solution. There are several pros and cons to using the non-standard approach. In an open standard we can in principle re-do everything as a manually built standard. This could become an argument both pro and con FPGA.
Using the standardised approach could save a lot of work as a standard solves many of the low level and complex issues that would all need to be considered with a completely home-built solution.

c. Form Factors

PCI-e, PCI, VME plus PMC, embedded receivers in PLCs are also a potential. It's possible to create a complete solution in VHDL which could be distributed as a solitary IP. A carrier board is also interested, MicroTCA is also of interest to some people.

d. Interest in Higher Level Protocols

None were expressed.

[1] Numbering in this section corresponds to that of the PM agenda shown before.

2. Project Management Issues

a. CERN driving project, remaining organisational structure.

This is generally considered as acceptable.

b. Open source definition.

Follow the SIM-PUTER open source agreement as a framework. Absolutely everything which is needed to manufacture the boards, understand their operation, and potentially modify them should be made open-source. Firstly we have to evaluate whether the Synchronous Ethernet and PTP combinations can operate correctly No objections were made.

c. Timescale of work.

Despite the time frame of work being defined as around two years, it is interesting to advance the work to meet everyone's requirements.

d. Quantities of components / machine requirements.

CERN requires around 1000 receivers only a few masters.
GSI requires around 2000 receivers only a few masters.
IN2P3 requires around 3000 receivers only a few masters.
ITER requires at least 100 receivers, increasing to several 1000s depending on final solution chosen.

e. Financing.

CERN will continue with their work as is, CERN is committed to developing this system. It will be developed as open-source, using the research budgets allocated for timing renovation. Payments for companies to work must be organised. Funds must be allocated to carry this out.
There are some possibilities for EU funding, but there may be a delay of 6-9 months to submit an application and hear for a response.
Equally some funds exist at National levels that may be tapped. In the short term there is no obvious alternative to CERN sponsoring the work to start.

Action: Contact CERN Legal Services to get this work organised as an official Collaboration.

3. Initial Thoughts Regarding Potential Work Packages[2].

a. Synchronous Ethernet plus PTP (IEEE 1588) Background (theoretical)
   i. Manage project (CERN).
   ii. User needs gathering (Cosylab, CERN).
   iii. Fibre/copper choice (FISS, InES).
   iv. Reliability studies (CERN, ITER).
   v. Is this constraining us in any way?. Can we mix commercial products with this? Verify this will stay standard (InES, ITER).
   vi. Scalability system-level simulation for this solution (FISS).
   vii. Study Robustness of Cascaded Systems (includes analysis of switches) (Oregano, IN2P3).
   viii. Concrete Overall Design Specification (CERN, IN2P3).

[2] This work package list is the final result after discussion. Attribution of work packages was only discussed for the first phase of the project (heading 3.a). In parenthesis, the Institutions responsible for each work package. The main responsible comes first. Names of responsible persons will be given shortly to have a main responsible per work package.

b. Synchronous Ethernet plus PTP (IEEE 1588) Prototype / Proof of Concept
    i. Drivers.
    ii. Test Programs.
    iii. Generator.
    iv. Switches, Study Combining of Ethernet and IEEE 1588 in Switch.
    v. Receiver.
c. Synchronous Ethernet plus PTP (IEEE 1588)
    i. New Standardisation?
    ii. Dissemination.

Benjamin Todd (minutes)
Javier Serrano (minor editing)