

P1394.1 WORKING GROUP
March 19 - 20, 1997
San Jose, CA

Dick Scheel, chair of the working group, called the meeting to order in the vicinity of 9:00 AM. As is customary, the attendees introduced themselves after which regular business started with an acceptance of the agenda.

Dick explained that IEEE procedures recognize individual contributors not company representatives.

The draft minutes of the October 7, 1996 meeting were accepted unanimously. The status report filed with the MSC was distributed.

The draft agenda accepted was:

- * Draft 0.02 of the standard
- * P1394.2 Liaison Report
- * Progress on Sony net simulation
- * Document Registry
- * Work Plan
- * Deadlock and starvation avoidance
- * IP/1394 Issues
- * Reset notification
- * Speed map
- * Open issues
- * Review of action items
- * Meeting schedule

DRAFT 0.02 REVIEW

Calto Wong raised a question about the rules in 8.1 that prevent a bridge from being the net cycle master. NB: The new name "net cycle master" replaces the informal "cycle monster" appellation for the singular clock source for the entire Serial Bus net. The rules are incorrectly stated; it should be possible for a bridge to have no portals that are cycle slaves.

Additional discussion on clock distribution throughout the net lead to the conclusion that isochronous data packets need to flow along the same paths used for clock distribution. The reason for this is that isochronous exchanges between talker and listener(s) rely upon a shared clock value: if isochronous packets are permitted to short-circuit the path used for clock distribution they are no longer guaranteed to be in phase with the clock. The reasons for this conclusion need to be stated in the document.

Dave James suggested that the "remote transaction" capabilities need to be reconsidered for the sake of security. The kinds of asynchronous requests that may be forwarded could vary according to whether or not the new

domain on the other side of a bridge portal is unsecured (uninitialized) or secured (already initialized by another bridge manager). This suggests that bridges may be manufactured in two styles: secure or open.

P1394.2 LIAISON REPORT

Dave James presented an extract from Draft 0.775 of the P1394.2 draft, which extract concerns bridging in the Serial Express environment.

Some 1394.2 concepts have evolved over the last year and may be useful to the Serial Bus bridging work:

- a) Simple bus ID routing tables that offer the possibility of more complex routing than base / bounds but occupy only slightly more storage space (e.g., four times as many bits as the current P1394.1 scheme);
- b) Remote transaction capabilities to access portal-specific registers (contrast this with the "window" method described in the current P1394.1 draft);
- c) Constrained remote transaction capabilities into domains that have already been "secured" by other bridge managers. This permits secure bridges to be implemented;
- d) In the same vein, remote "ring" notification to a bridge manager (or other node) that does not require knowledge of the destination---the intended recipient of the "ring" had previously configured the bridge to transmit an alert.

The group decided to investigate the usefulness of these ideas to P1394.1 and create a proposal couched in terms familiar to Serial Bus users.

SONY NET SIMULATION

Du Hung Hou presented inception-to-date results from Sony's efforts to simulate bridge behavior. The project was started to provide proof-of-concept and experimental test-bed for Sony's exploration of bridge architectures.

The simulation may be scripted to express node, bridge and isochronous behavior. Events occur sequentially and time delays may also be expressed within the script.

Simulation has produced recommendations for changes to the P1394.1 draft:

- a) Relocate PORTAL_SELECT register outside of the portal-specific window to avoid window "nesting" effect.

The document that describes the work is BR001r00.pdf.

DOCUMENT REGISTRY

The working group selected the following nomenclature for documents:

BRnnnRrr.pdf

The nnn is a document number assigned by the Secretary, Peter Johansson. Contact the secretary by EMail at pjohansson@aol.com for a document number; when the document is available, forward it to the Secretary who will see that it is placed on the FTP site. The revision level of the document, rr, is assigned by the document author and is customarily expected to start at zero and increment monotonically.

Working group minutes are labeled as follows:

Mddmmmyy.pdf

In the above, dd is the numeric day of the month, mmm is the alphabetic abbreviation for the month and yy is the two least significant digits of the year.

The working drafts of the standard are identified as Dvv_rr.pdf, where vv is the version level and rr is the revision level.

Working group participants are requested to post new documents to the FTP site at least three (3) full business days before the next meeting. If an author cannot meet this requirement, the onus is on the author to bring an adequate number of printed copies of the document to the meeting.

WORK PLAN

After a detailed discussion of IEEE procedures and terminology (sponsor ballot, recirculation ballot, RevCom, MSC, etc.), the working group concluded that December, 1997 is a realistic target date for completion of the working draft. The next MSC meeting is in January, 1998, at which time the request for sponsor ballot would be made. If the sponsor ballot, preparation of ballot responses, recirculation ballot and final review by RevCom follow patterns typical of the IEEE in the past, January, 1999, is projected for final approval of the standard.

The first significant milestone before December will be the publication of a Draft 1.00 when the group deems the document to be substantially complete. No accurate estimate of when this might be was offered.

DEADLOCK / STARVATION AVOIDANCE

Dave James gave a short tutorial on the nature of deadlock and starvation on a split-transaction bus. The design principals that avoid these problems include:

- a) Separate request and response queues; and
- b) A preference to prioritize responses over responses.

Note that although b) is generally desirable, if carried to extreme can result in starvation.

A common misunderstanding is that the separation and independence of request and response handling requires additional hardware. This is not necessarily so: careful design can permit independent, non-blocking behavior while sharing hardware resources between requests and responses.

The deadlock and starvation problems are not unique to bridges---it is just that the higher likelihood of congestion within bridges exacerbates the problem.

P1394a intends to add an informative annex to illustrate the nature of the problem and educate implementers. It is difficult to write exacting rules to prevent deadlock and starvation because particular strategies may be robust for some applications / implementations but not for others.

IP/1394 and GLOBAL BROADCAST

The anticipated standardization of IP traffic over 1394 by the IETF raises an issue for the bridge working group: global broadcast. "Global broadcast" is defined to mean an unconfirmed write request that is propagated throughout the Serial Bus net.

At present there are two proposals for global broadcast:

- a) Global broadcast is identified by destination_bus_ID 0x3FE; and
- b) Global broadcast is identified by a combination, when destination_ID equals 0xFFFF and source_bus_ID is not equal to 0x3FF.

Taka Fujimori and Peter Johansson explained the respective proposals to the group. The ideas are the subject of controversy at present and are explained in more detail in documents to be submitted to the FTP site. Please consult the document index.

There was agreement in the working group that a facility for global broadcast of asynchronous write transactions is needed for P1394.1 but further study and debate are required before a choice of either (or both) methods can be made.

RESET NOTIFICATION

Dick Scheel presented a short paper on problems that may occur when 1394 errors occur in the transmission of broadcast write packets to the RESET_NOTIFICATION register. These documents are available as BR002r00.pdf and BR003r00.pdf at the FTP site.

The simplifying assumption is that subsequent to a reset notification, the notified node receives no information as to which bus experienced the reset. This requires the affected node to assume that ALL buses may have experienced a reset and to interrogate the configuration ROM of all remote nodes to which it had directed requests. The principal advantage is the idempotent nature of the write to

RESET_NOTIFICATION; there is no requirement to latch the first write until software retrieves the value.

If this is workable, it can lead to additional simplifications in the bridges and in the reset notification / acknowledgment protocol.

Dick then presented a proposal to eliminate the quarantine bit in bridges. It is based upon the use of a spanning tree topology for the routing of asynchronous packets and envisions an expanding "wave front" of reset notifications that place bridges into states where they discard all remote request and response packets. In some ways this creates a global reset (although it does not affect local traffic).

SPEED MAP

There are a number of different issues for the management of speed in a Serial Bus net:

- a)The maximum speed that may be used to transmit a packet between two nodes on different buses;
- b)In order to maximize bus utilization, an "in transit" transmission, i.e., between an outbound portal of one bridge and the corresponding inbound portal of another bridge on the same bus, should occur at the fastest possible speed between those portals; and
- c)The recreation of an appropriate speed code for the transmission of the packet to the ultimate recipient (the speed at which the packet was first transmitted reflected the fastest speed to the first bridge portal).

This is a continuation of the same discussion from the October meeting in Redmond, WA.

With respect to a), a net-wide speed map is not useful. On the other hand, the slowest point-to-point PHY connection between any two nodes in the net determines the maximum size packet that may be transmitted. Possible solutions are:

- a)Trial and error. After discovering the existence of an interesting remote node at S100, the requester attempts other packet sizes to probe the maximum supported;
- b)The bridge manager maintains a very large array that enables table-lookup of the maximum size packet permitted between any two of the 65,535 possible Serial Bus nodes. Note that although this array might appear very large in terms of the Serial Bus address space it consumed the actual implementation in a bridge manager might be very sparse; or
- c)The bridge manager provides a service where the requester may supply arguments of source_ID and destination_ID and the bus manager returns the largest packet size. An advantage of this approach is that

other information could be returned as well, such as the speed to use when the requester first transmits the packet (i.e., the maximum speed between the requester and the first bridge portal).

With respect to transmission between one portal and an adjacent portal while in transit, three solutions were suggested:

- a) There is one speed value that represents the lowest common denominator of speeds between the portal and all other portals on the same bus;
- b) There is a 1,023 entry table of 3-bit entries that provides complete speed information based on bus ID; and
- c) If a "multiplexed" routing scheme (such as the one proposed in Annex C of the P1394.2 draft) is used, it would require 64 3-bit entries to provide the necessary information.

After considerable discussion, the working group selected a) as the most workable. If this is inadequate for future, sophisticated bridges it is never the less sufficient for environments where there only two bridge portals on any one bus. Even in topologies where there are more than two portals on a bus it may often be possible to group bridge portals close to each other so that they may all transmit at a single high speed.

As far as the speed used for the final transmission to the destination node, there was agreement that the bridge needs to keep a copy of part of the local bus's SPEED_MAP (only the row that corresponds to the portal's PHY) in order to obtain the speed code.

OPEN ISSUES

Does the jitter behavior of cycle timers within bridges have to be constrained beyond the requirements of 1394-1995 in order to make bridged systems work?

What is the security domain of a Serial Bus net? Are bridges supposed to provide security between enumerated buses? Or is it an acceptable model if the entire net is a single security domain? Raise the security issue at the 1394 TA.

Define a minimum set of transaction capabilities for bridges. For instance, are read4, write4, lock4 and lock8 adequate? Perhaps block transactions with lengths up to 64 bytes would be useful, even if not necessary.

Do bridges have different requirements for fair and priority arbitration than other 1394 nodes? How does the P1394a proposal to waive fairness requirements for responses affect bridges?

The standard needs to have a separate section or annex that enumerates the hardware changes needed in link designs to specifically enable bridges. Some of these requirements may be in conflict with compliance requirements of other standards, such as P1394a, and it may be necessary to make explicit that they are exemptions for bridges and bridges only.

ACTION ITEMS

Peter Johansson to work with David James to prepare a proposal based upon concepts from P1394.2 Annex C.

Bob Gugel to investigate within TI and the backplane community whether or not the write behavior of NODE_IDS could be redefined so that the physical ID is not altered.

Dick Scheel to post the new document registry requirements to the reflector.

Peter Johansson to edit the next document revision and post it to the FTP site.

MEETING SCHEDULE

The next scheduled meeting of P1394.1 is Tuesday and Wednesday, May 13 - 14, 1997, in the Bay area (most likely the South Bay). Dick Scheel will confirm the meeting location.

ADJOURNMENT

The meeting was adjourned at 1:00 PM on Thursday, March 20.

ATTENDANCE

Nancy Chang	(408) 955-5743	nancyc@lsi.sel.sony.com
Taka Fujimori	81-3-5488-6353	fujimori@cv.sony.co.jp
Bob Gugel	(972) 480-3715	r-gugel@ti.com
Du Hung Hou	(408) 955-5742	duh@lsi.sel.sony.com
Dave James	(415) 494-0926	dvj@scizzl.com
Peter Johansson	(510) 531-5472	pjohansson@aol.com
Michael Nguyen		michael.nguyen@fcpa.fujitsu.com
Dick Scheel	(408) 955-4295	dicks@lsi.sel.sony.com
Ming L. So	(408) 486-8126	mso@opti.com
Calto Wong	(914) 945-6382	cxw@philabs.research.philips.com
Benjamin Pan	(408) 588-8379	bpans3.com

POINTS of CONTACT

FTP site: ftp.symbios.com:/pub/standards/io/1394/p1394.1

Reflector: p1394.bridges@sun.com
(maintained manually by bob.snively@sun.com)

Chair: dicks@lsi.sel.sony.com

Editor and Secretary: pjohansson@aol.com