

Improving Auto-Negotiation Efficiency

Next Page Extension

Pat Thaler
September 2004



Agilent Technologies

Problem Statement

Clause 28 auto-negotiation was created at a time when only a few ability bits had to be transferred.

Usually a single negotiation page was needed.

Per page negotiation initialization time cost is high - about 1/4 s and next page payload is small - at most 10 bits per exchange plus an extra exchange for message code.

New physical layers require more bits to be exchanged.

Goal: Enhance auto-negotiation to complete negotiation in two page exchanges while maintaining legacy compatibility



Overview of proposal

- **For maximum backwards compatibility, we should leave the base page as it is.**
- **During base page exchange, determine whether both sides support extended next pages.**
- **Extended next page support should be mandatory for 10GBASE-T and optional for existing PHY layers.**
- **If extended next pages are supported, then next pages will be 48 bits rather than 16 bits per page.**
- **The existing state machines will support this with some updates to variable definitions.**



Extended next page format



Flag are the T, Ack2, MP, Ack and NP bits

- **Message code field and flag bit positions from legacy Next page message page encodings are retained to simplify transition.**
- **A 32 bit payload is supported. Allows space for**
 - **5 bits for 1000BASE-T configuration bits**
 - **4 bits for 10GBASE-T configuration bits**
 - **10 bits for random seed**
 - **13 bits for future technology**



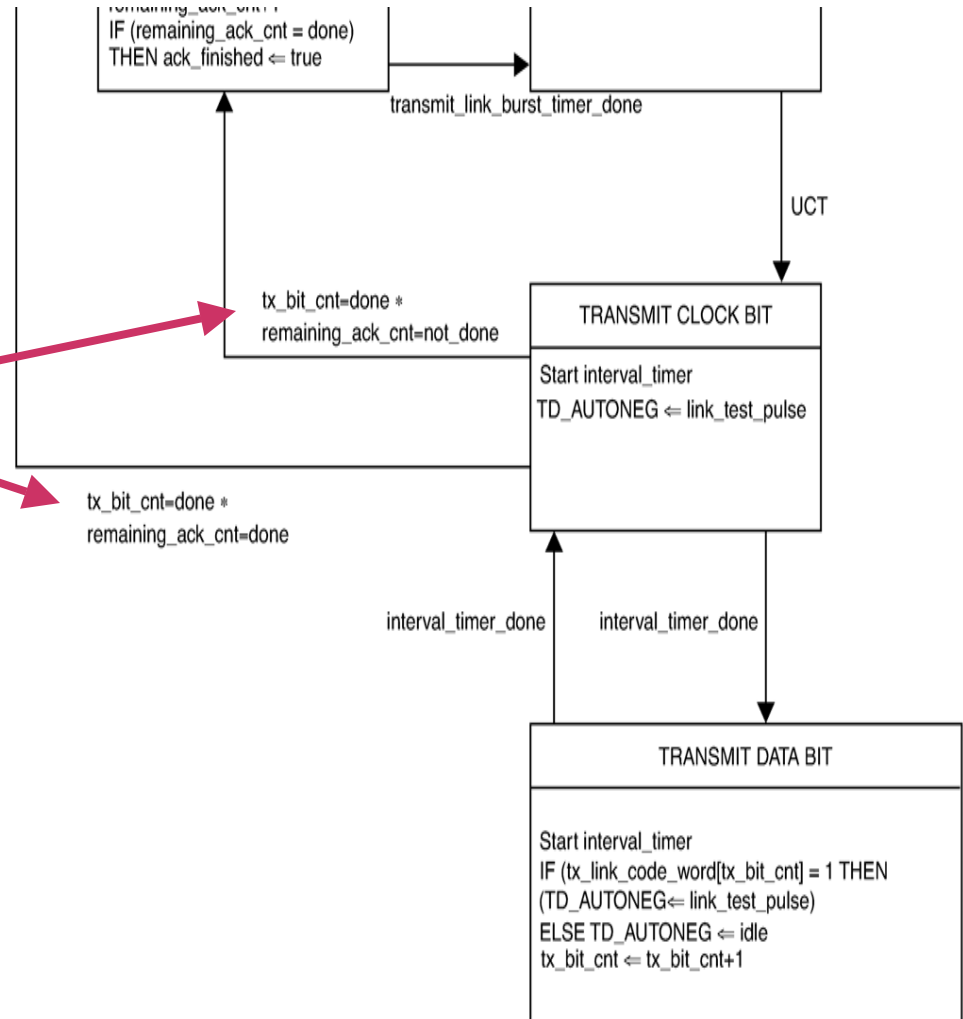
Negotiation of extended next page length

- **Bit A7 in the base page is currently reserved.**
- **Assign A7 to Extended next page ability**
 - **A value of 1 indicates that extended next pages are supported**
- **If the received and transmitted A7 bits are both 1, then extended next pages will be used for the remainder of the negotiation.**



No effect on transmit state diagram

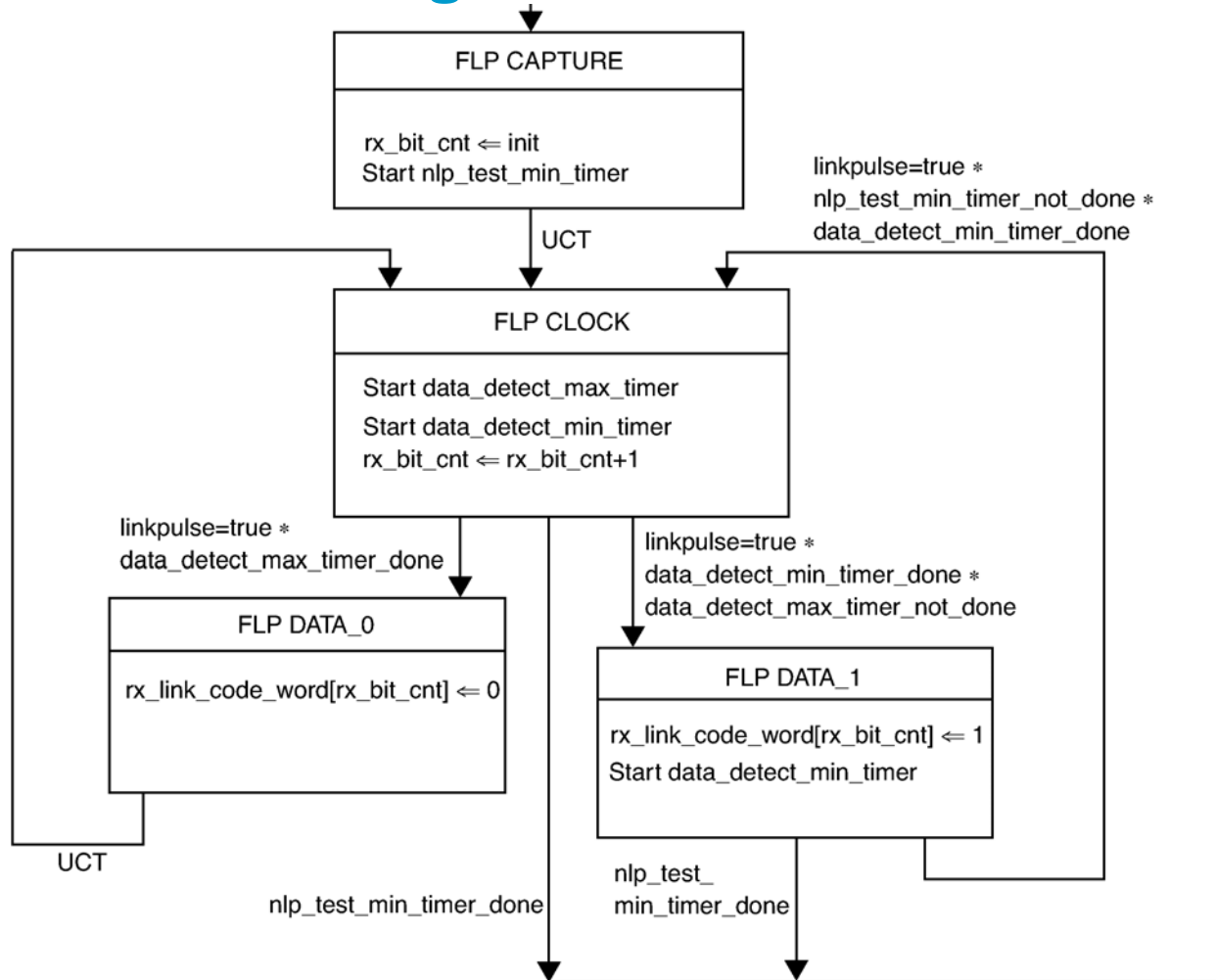
Note that **tx_bit_cnt** definition controls the number of bits sent.



No effect on receive state diagram

Note that rx_bit_cnt controls where bits are placed and the actual state diagram is independent of the number of bits received.

Total burst must be less than nlp_test_min_timer



Extended next page burst length effect

- **48 bits at 125 ns per bit (the time between clocks) require 6 us to complete.**
- **Existing nlp_test_min_timer range is 5 to 7 ms.**
- **The least change fix is to require a range of 6.5 to 7 ms for devices that support extended next page.**
- **An alternative would be to change the values of all the pulse separation timers to half their current value when in extended next page mode.**
 - **e.g. interval_timer goes from 62.5 us nominal to 31.2 us nominal for extended next pages.**
- **The simpler change is proposed here.**



Changes to state diagram variables.

- add a variable: `page_size` - page size at the beginning of negotiation is 16. When extended next page mode is on, `page_size` is increased to 48
- `rx_bit_cnt` - This counter takes values up to `page_size + 1` and does not increment beyond `page_size + 1`. It takes the value: `not_done` for 1 to `page_size-1` inclusive, `done` for values `page_size` or `page_size+1`
- `tx_bit_cnt` - This counter takes values up to `page_size + 1`. It takes the value `not_done` for values from 1 to `page_size`, `done` for `page_size+1`
- `nlp_test_min_timer` - The smallest change to the timers is to require this timer to be between 6.5 and 7 ms in implementations that support extended next page.



Management registers for extended next page

Registers could be read like the multi-register counters.

- **Reading the first register latches the other two registers**
- **The safest method is to allocate four additional clause 45 registers**
 - **two for the additional 32-bits of transmitted next page data**
 - **two for the additional 32-bits of received next page data**
- **An alternative would be to return the values in 3 reads of the same register but there is a risk of getting out of sync.**

The first method is recommended.



Other possible efficiency improvements.

- The original timers in auto-negotiation have very loose tolerances to allow for analog implementations (e.g. one-shots)
- FLP Burst to Burst time is allowed to vary from 8 to 24 ms.
- Each page negotiation requires enough time that:
 - Side A receives 3 matching FLP bursts and sets the Ack bit
 - Side B receives 3 matching FLP bursts with the Ack bit set
- Therefore, a single page exchange can be longer than 144 ms with the maximum burst to burst time.
- Negotiation time could be reduced by requiring burst to burst time for 10GBASE-T to more tightly controlled
 - Propose 8 us min to 8.5 us max
 - Improves time for 6 bursts to 51 ms.
- This change would be entirely backward compatible

Efficiency improvements - continued.

- Greater improvement can be made by shifting to a new set of timer values when extended page mode is operational.
- Pulse width is 100 ns
- Page exchange time would be improved from 144 ms to 1 ms but requires more to transition than the first proposal.

	Old nominal value	Extended page value
Clock pulse to clock pulse	125 us	1.6 us
Clock pulse to data pulse	62.5 us	0.8 us
clock pulses in burst	17	49
Burst width	2 ms	80 us
FLP burst to FLP burst	8-24 ms	160 us



Summary

Small changes can allow significant reduction in auto-negotiation time. At a minimum these changes should be made:

- **Add extended next page length**
- **Require tighter FLP burst to burst tolerance for 10GBASE-T**

Further improvement could be achieved by reducing the timer values in extended next page mode.

