ethernet alliance

Improving the Energy Efficiency of Ethernet: Adaptive Link Rate Proposal

Version 1.0, July 15, 2006

Authors: Mike Bennett
Lawrence Berkeley National Laboratory
Ken Christensen
University of South Florida
Bruce Nordman
Lawrence Berkeley National Laboratory

ethernet alliance | p.o. box 200757 | austin, tx | 78720-0757 | usa
www.ethernetalliance.org

executive summary

There is a major push for chip, component, and system vendors to improve the energy efficiency of their products. Reasons for this push include reducing heat dissipation, improving battery lifetime, and reducing environmental impact. This includes complying with regulatory specifications on energy use (such as the EPA Energy Star specification). Power management capabilities exist for desktop computers where the power consumption of disk drives, processors, displays, and even in some cases Ethernet links can be scaled-down or entirely powered-off to achieve energy savings. Enabling Ethernet to quickly and adaptively set its link rate in response to link utilization during full-power operation can yield even greater energy savings, and achieve an unbeatable energy efficiency of the most popular networking link layer protocol in the world. Energy savings of hundreds of millions of dollars per year in the US alone are possible by implementing and adopting an Adaptive Link Rate (ALR) method described in this white paper.

potential for energy savings

There are roughly one billion Ethernet connections worldwide. In the US alone, there are about 400 million connections, many which will soon be running at 1 Gb/s. The majority of Ethernet controllers sold today support a 1 Gb/s data rate. As bandwidth demands continue to increase, the likelihood of 10 Gb/s becoming a commonly used data rate is very real. With an increasing data rate comes an increasing energy use. For a typical PC if an Ethernet link is operating at 1 Gb/s versus 100 Mb/s, a difference of about 2 W (AC) can be measured. This same 2 W difference can also be measured at the other end of the link at the LAN switch. This energy use is independent of traffic being carried on the link.

Most of the time, Ethernet links connecting a desktop PC to a LAN switch are idle, or very nearly idle. Indeed, measurements show that link utilization generally remains in the 1% to 5% range. High link data rates, such as 1 Gb/s, are desirable when downloading large files. Thus, high data rates to the desktop are primarily to support burst bandwidth. But, what about the "most of the time" when only a very small fraction of the available bandwidth in a desktop to LAN switch link is utilized? If the link data rate were reduced from 1 Gb/s to 100 Mb/s the user would never know it **if it could be increased back up to 1 Gb/s instantaneously and automatically when needed.**

What are the possible energy savings if existing and future 1 Gb/s (and 10 Gb/s) desktop to switch Ethernet links could operate at 1 Gb/s (or 10 Gb/s) when needed, but at 100 Mb/s when idle or lightly utilized? That is, what savings could be achieved if an Ethernet link data rate was adaptive as a function of link utilization? The following assumptions were made:

- Two-thirds of commercial sector desktop PCs are fully powered-on 24/7 as are many residential desktops;
- A full 1 Gb/s data rate is required only a cumulative one hour per workday (and the rest of the time operation at 100 Mb/s is possible with no perceivable difference to the user);
- The power difference between 1 Gb/s and 100 Mb/s is 4 W for a desktop to switch link;
- By 2012, there will be 80 million Ethernet-connected commercial desktop PCs in the US half of which could be enabled and operating at 1 Gb/s and support ALR; and
- ALR can also apply to residential desktops, notebooks, and other IT and CE products.

At this 50% penetration rate, ALR in commercial desktop PCs could save about US $80 million in energy per year in the US alone — with considerable additional savings from residential and other IT and CE products. Table 1 shows the calculations for the commercial desktop PC savings estimate.

Parameter	Value
Power savings per Ethernet link	4 W
1 Gb/s traffic time per day	1hr
ALR low-traffic time per day	18 hr 75 %
ALR low-traffic time per year	70 %
Energy savings per PC per year	24.5 kWh/yr
Energy savings for 40 million PCs per year	0.98 TWh/yr
Savings at US$ 0.08 per kW/h per year	$78 million/yr

Table 1. Estimated energy savings from the use of ALR in commercial desktop PCs

The potential for energy savings is so great that the revised EPA Energy Star Program Requirements for Computers: Draft 2 Version 4.0 states, "All computers shall reduce their network link speeds during times of low data traffic levels in accordance with any industry standards that provides for quick transitions among link rates. … With such capabilities in place, reduced link rates are expected to be heavily used on Tier 2 computers, and have the potential for significant savings."

existing power management for non-Ethernet technologies

Asynchronous digital subscriber loop (ADSL) standards and implementations now support power management in ADSL2 and ADSL2+ (ITU G.992.3/4/5). ADSL2+ is a last mile technology with a data rate of up to 24 Mb/s with a power consumption of about 2 W per port at full data rate. ADSL2 (and ADSL2+) support three power states corresponding to link data rates. The states are L0 (full on), L2 (low power), and L3 (off). By reducing power consumption, heat is also reduced at remote transceiver cabinets. Transitions between the L0 and L2 states are near-instantaneous and user-imperceptible. Transitions from the L3 state will take a maximum of 3 seconds. The targeted power consumption per port in states L2 and L3 are 0.75W and 0.3W, respectively.

The European Commission's Joint Research Center under the European Union (EU) Stand-by Initiative for improving the energy efficiency of electrical equipment is proposing to standardize on ADSL2 and ADSL2+ for home broadband delivery due to its power management capabilities.

existing power management for Ethernet

A new generation of Ethernet controllers is beginning to support enhanced power management capabilities and enable reduced power usage. For example, many desktop and laptop computers will drop their Ethernet data rate from 1 Gb/s to 100 or 10 Mb/s if the system enters a low power sleep state (such as the Microsoft Windows Standby state). The Ethernet controllers can also automatically drop their internal data rate and power-down their PHY circuitry if no signal is detected on a link (that is, when the other end of the link has been powered down or unplugged).

This existing power management makes use of Ethernet auto-negotiation to force a lower data rate (for example, when entering a system sleep state) or higher data rate (for example, when leaving a system sleep state to resume full power operation). Existing auto-negotiation requires several 100s of milliseconds to force a data rate change. The use of auto-negotiation works well if the time to change the data rate is otherwise masked by system power-up time such as in the case of a system resuming full-power operation from a low-power sleep state. The use of auto-negotiation cannot be acceptable during periods of user operation where a data rate transition time of 100s of milliseconds could introduce a delay glitch noticeable by a user and might also cause data loss due to packets from a high-speed burst overflowing buffers in a LAN switch.

To be transparent to a user, a link data rate change has to take place in much less time than hundreds of milliseconds. For this, a new mechanism is needed. A policy is also needed to determine when to change the data rate from low to high, and high to low. This is addressed in the proposed Adaptive Link Rate (ALR) for Ethernet.

adaptive link rate (ALR) for Ethernet

The ALR proposal is provided to the Ethernet community and the IEEE 802.3 standards community as a means to foster discuss and initiate the study of new mechanisms and policies for fast transition between existing link data rates. The ALR proposal uses a two-way MAC frame handshake that would be faster than existing auto-negotiation. Such a MAC frame handshake could be implemented either in the device driver or within the Ethernet controller itself. The proposal also contains a policy based on buffer thresholds and utilization measurements.

ALR mechanisms and policies must be implemented in both the desktop Ethernet controller and LAN switch Ethernet controller. Thus, ALR is a two-way protocol where either end (for example, desktop computer or LAN switch port) can initiate a link data rate change. The existence of an ALR capability would need to be agreed-upon at link start-up via the existing auto-negotiation mechanism. ALR would only run within the bounds of an advertised capability of both ends of a link. For example, a request to transition from low to high data rate at a desktop computer would only be made if it is known that the other end (for example, the LAN switch port) supports the higher data rate.

ALR mechanism

A fast signaling mechanism is needed for ALR to be transparent to users. One possible mechanism is a two-way MAC frame handshake. Figure 1 shows a possible MAC frame definition. The opcode field is used to indicate the required action. The control field is the requested data rate (which could be defined similarly to the Technology Ability Field in the auto-negotiation link code word). Three new opcodes are needed, one each for the ALR transition request, ACK, and NACK MAC frames. These opcode values are to be determined. An ACK response would indicate that a data rate change is acceptable to the other end and would be followed by a link resynchronization at the new data rate. A NACK response would indicate that a data rate change is not acceptable to the other end.

DA	SA	Type	Opcode	Data rate	Padding	FCS
TBD	Source MAC addr	Control (88-08)	TBD	TBD	Reserved (sent as zeroes)	CRC32
6 octets	6 octets	2 octets	2 octets	2 octets	42 octets	4 octets

Figure 1. ALR MAC frame

Figure 2 shows a two-way MAC frame handshake with a link resynchronization immediately following the ACK MAC frame.

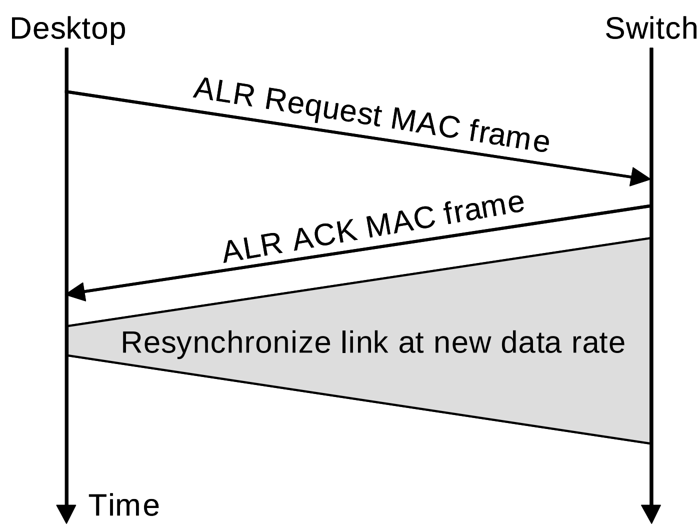

Figure 2. ALR MAC frame handshake

ALR control policy

When should a low data rate be changed to a high data rate? When should a high data rate be dropped to a low data rate? The ALR control policy determines when to change the data rate. The policy must be as simple as possible to allow for direct implementation within an Ethernet controller. An effective policy is one that allows the link to operate in a low data rate for as much time as possible without measurably affecting application response-time as perceived by a user. Significant increases to packet delay would increase application response time.

Buffer thresholds and utilization monitoring are two key ingredients for an effective ALR policy.

Transition from low to high data rate: It is proposed that a high queue threshold should trigger a transition from low to high data rate. That is, when the transmit buffer in an Ethernet controller have more than a specified large (high threshold) number of bytes or packets queued, a rate transition MAC frame requesting a transition to a high data rate should be sent to the other end. This request to a higher data rate must never be denied by the other end and thus would always be acknowledged with an ACK MAC frame.

Transition from high to low data rate: It is proposed that link utilization should be monitored and when below a threshold value, a rate transition MAC frame requesting a transition to a low data rate is sent to the other end. If the other end has conditions – such a high measured utilization or a transmit buffer that is above threshold – it would deny the request with a NACK MAC frame. Thus, link operation at a high data rate always takes precedence over operation at a low data rate.

Utilization monitoring should be used for triggering transitions from high to low data rate to prevent data rate oscillation of a link. Utilization monitoring can be implemented by counting bytes sent during a sampling period. Simulation-based experiments have shown that 5% utilization threshold at a high data rate and 10 milliseconds for a link utilization sampling period are reasonable values. ALR has been studied at the University of Florida and University of South Florida (see http://www.csee.usf.edu/~christen/energy/main.html). The effects of ALR on packet delay versus percentage of time at low data rate was studied using simulation models of Ethernet links with both traced traffic and synthetic traffic as inputs. Figures 3 and 4 show the results from an experiment using synthetically generated bursty traffic. The experiment assumed a high data rate of 1 Gb/s, low rate of 100 Mb/s, and 1 millisecond delay to handshake and resynchronize the link between the two data rates. A utilization threshold of 5% was used. Figure 3 shows mean frame delay (this is the delay seen by a frame spent queueing in the Ethernet controller and transmitting on the link) and Figure 4 show the percentage of time spent in low data rate. The three traces in each graph are for the utilization sample period (of 1, 10, and 100 milliseconds). It can be observed that at low utilization levels (normal for desktop Ethernet links) most of the time can be spent in the low power 100 Mb/s data rate at the cost of a non-user perceptible increase in delay. For example, at 5% link utilization, for a utilization sample period of 10 milliseconds, the mean response time is less than 0.5 milliseconds and about 80% of the time is spent at 100 Mb/s. At higher link utilizations, the link operates only at its high data rate as would be desired.

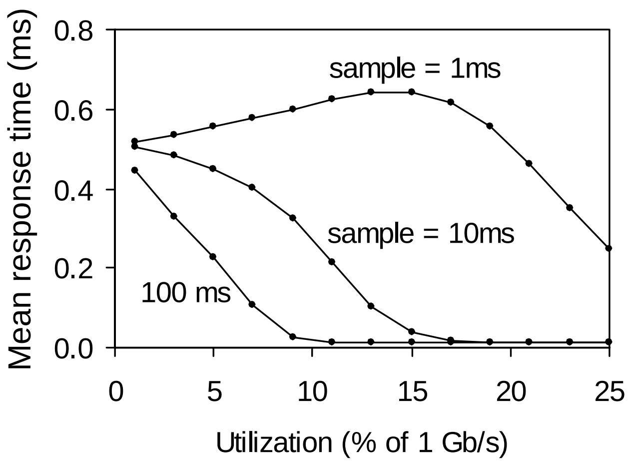

Figure 3. Mean response time

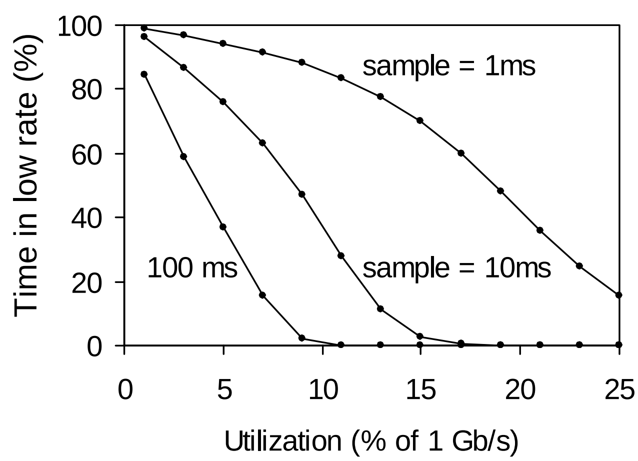

Figure 4. Percentage of time in low rate

open challenges

Open challenges remain. **An IEEE 802.3 study group is needed to fully identity and address these open challenges.** Open challenges include:

- Link resynchronization without re-training for 1 Gb/s to 10 Gb/s operation needs to be investigated;
- Start-up conditions and requirements needed to be considered (e.g., in terms of high layer protocols waking-up and for link failover/redundancy configurations);
- Master/slave roles between the two ends of the link need to be addressed; and
- Possible interactions with Power-over-Ethernet (POE) need to be investigated.

summary and next steps

Ethernet has the potential to be the most energy efficient wired link layer protocol in the world. Adaptive Link Rate (ALR) is one step that can be taken to achieve this necessary goal. There is a substantial energy savings potential from implementation of a standard method for quickly changing link data rates in response to traffic levels or link utilization. ALR is in some respects an extension of Ethernet power management capabilities that use auto-negotiation to force link data rate changes. Developments in ADSL set a precedent for Ethernet to follow and improve upon in this important area of power management.

For ALR to be successfully implemented and widely adopted it must first become a standard. The IEEE 802.3 standards community needs to address the open challenges outlined above.

The appendix contains a FAQ for ALR.

acknowledgments

The authors thank Bob Grow of Intel Corporation and Brad Booth of Quake Technologies for their helpful comments to this whitepaper and support for ALR.

author biographies

Mike Bennett is a Senior Network Engineer at Lawrence Berkeley National Laboratory with over 20 years of combined experience in electronics and enterprise networking. He co-manages the LBLnet Services Group which is responsible for operating and maintaining the Lab's campus area network, a LAN with over 13,000 attached devices. Mike is an IEEE standards association member, and a contributor of end-user's perspective to a number of Ethernet standards.

Ken Christensen is an Associate Professor in the Department of Computer Science and Engineering at the University of South Florida. His interest in energy efficiency in computer networks goes back to the mid-1990s with work on Wake-On-LAN at IBM. His current research in energy efficiency of networks is funded by the National Science Foundation.

Bruce Nordman is a researcher with the Environmental Energy Technologies Department of Lawrence Berkeley National Laboratory. He has worked on energy efficiency in electronics at LBNL since the early 1990s, doing research for the EPA Energy Star program, California Energy Commission, and U.S. Department of Energy.

appendix – ALR FAQ

This appendix contains an FAQ for Ethernet Adaptive Link Rate (ALR).

1) What is ALR?

Adaptive Link Rate (ALR) is a proposal that Ethernet links dynamically change their data rates in response to traffic levels. For high levels of traffic, a high link data rate is necessary and must be used. For low levels of traffic, a low data rate can be sufficient and should be used. It is well known that most Ethernet links – especially PC to LAN switch links – have very low traffic levels most of the time. ALR could apply to 10 Gb/s in the future or to 1 Gb/s currently. ALR does not propose that new data rates be invented, but rather that existing 10, 100, 1000, and 10000 Mb/s data rates be used.

2) What is the need for ALR?

The need for ALR is part of a growing need to reduce the energy costs associated with operating residential and commercial computing equipment. ALR can be a first step to address direct energy use of network links and equipment. ALR may be incorporated into future EPA Energy Star specifications.

3) What is the primary application for ALR?

The primary application for ALR is expected to be for the Ethernet link from the PC to first-level LAN switch. PC links are idle the great majority of the time (see question #16). ALR is not intended for the data center where link utilization is often high and requirements for very low latencies (and high throughputs) are far greater than for PC links. The number of relevant Ethernet controllers that could use ALR is well over 100 million in the US alone.

4) What is the expected operating cost savings achievable from ALR?

For the individual user, ALR could achieve energy savings of several dollars per year (more than recovering the expected additional cost of adding ALR to an Ethernet controller). On a larger scale, ALR has the potential of saving hundreds of millions of dollars per year in energy costs in the USA alone. Companies with thousands of PCs would see a measurable reduction in their yearly energy bill. Most savings come from desktop PCs and the network equipment they link to, both in the commercial and residential sectors. Other savings come from the increasing number of IT and CE products that use Ethernet technology. Additional savings come from reduced air conditioning costs — these are not estimated here.

5) Is ALR technically feasible for 1 Gb/s?

The answer to this question can best be determined as part of a multi-company study work though it is believed that ALR is easily feasible for 1 Gb/s. Evidence for this is that existing Ethernet controllers already have the capability to change their link data rate when a laptop or desktop computer enters a low-power sleep state (e.g., Microsoft standby state).

6) Is ALR technically feasible for 10 Gb/s on copper medium?

ALR is almost certainly feasible for 1 Gb/s. ALR for 10 Gb/s is a more challenging technical problem. The answer to this question can best be determined as part of a multi-company study.

7) How is a link data rate change initiated?

A link data rate change could be initiated by a handshake mechanism (e.g., via newly defined MAC frames). Use of existing auto-negotiation mechanisms can/should also be investigated. When to initiate a link data rate change is a question of policy. Queue thresholds and/or utilization measurements can be used to determine when to increase and decrease data rate such that packet delay is not perceivably affected.

8) Can existing auto-negotiation be used to change data rates?

The existing auto-negotiation scheme in 802.3 takes 100's of milliseconds to change data rates at 1 Gb/s link data rate. This is too slow for ALR to be most effective. See also question #10.

9) Do both ends of a link have to be ALR compatible?

Yes. Both the Ethernet controller in the PC and the port in the switch must support ALR. Auto-negotiation would need to be used to detect and agree on the use of ALR between two ends of a link.

10) Can there be different levels of implementation of ALR?

Yes. This is based on time scales of interest. For fast switching (e.g., 1 millisecond) a new handshake mechanism is needed. For slow switching (say, 100s of milliseconds to few seconds) one could use existing mechanisms and this is then primarily a "software problem".

11) Will ALR result in a user-perceptible performance impact?

No. A link data rate switching time of roughly 1 millisecond would not be perceivable to a user. A goal of ALR must be to have no impact to user performance.

12) Does ALR have to be a standard?

For fast switching of data rates, standardization of new mechanisms would ensure compatibility and interoperability at all levels, and greatly increase the chances that the installed functionality is actually used.

13) Is there any precedent or "competition" for ALR in other link layer technologies?

ADSL2+ supports multiple power states and data rates.

14) What are the key reference documents for ALR?

The key reference for ALR is the pre-proposal document: Adaptive Link Rate (ALR) for Full-Duplex Ethernet Pre-Proposal for an IEEE 802.3 Study Group (http://www.csee.usf.edu/~christen/energy/alr_proposal.pdf).

15) Why support an activity to standardize ALR?

ALR offers an opportunity for product differentiation – there is a market for low-power and low-energy devices. ALR also offers an opportunity to improve energy efficiency in a studied manner, rather than possibly being told how to do it by regulatory entities. ALR could be part of future Energy Star and similar guidelines either in the US and/or abroad.

16) What makes a link relatively "idle"?

For purposes of ALR, an idle link in not one with no data being transmitted. Rather, it is one in which the amount of data being transmitted is very low compared to peak capacity. PC to switch links are typically very lightly utilized over the long run as they are primarily intended for high-speed burst capability (e.g., occasional large file transfers). Studies have shown that Ethernet links are typically utilized in the range of a few percent.

17) What key assumptions drive savings estimates from ALR?

The 4 W figure reflects 2 W of savings at each end of the link. This is the savings seen "at the wall", at the AC supply. This reflects a smaller amount of savings in DC power, due to power supply losses. Field research by LBNL has shown that about two-thirds of desktop PCs in commercial buildings are fully powered-on 24/7. Residential surveys by TIAX have shown that residential desktop PCs are on over half of the time on average. Stock figures for 5 years in the future are inherently speculative. The figures for this estimate are a combination of those from various LBNL studies and from TIAX.

18) Who are the key contacts for ALR?

The primary contact for ALR in its pre-proposal stage is:

Mike Bennett
Lawrence Berkeley National Laboratory
MJBennett@lbl.gov
(510) 486-7913