GMPLS-enabled, energy-efficient, self-organized network: MiDORi

Satoru Okamoto*, Yuki Nomura*, Haruka Yonezu*, Hidetoshi Takeshita*, Naoaki Yamanaka*
*Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa, Japan 223-8522

ABSTRACT

A network using minimum set of nodes and links that all communication data traffics among nodes can be carried while meeting the QoS requirements named MiDORi (Multi- (layer, path, and resources) Dynamically Optimized Routing) has been proposed. MiDORi network technologies that is composed of a low energy consumption network design tool (MiDORi path computation engine (PCE)), MiDORi Generalized Multi-Protocol Label Switching (GMPLS) protocol extensions for controlling network equipment, and a MiDORi router/switch controlled by MiDORi GMPLS, are now developing. This paper describes recent activities of the MiDORi network technologies.

Keywords: Energy efficient network, traffic engineering, GMPLS, path provisioning

1. INTRODUCTION

In 2007, the Information and Communication Technology (ICT) sector was responsible for 2% of global carbon emissions1 and rapidly increasing at a year rate of 6%. This means that the global carbon emissions will be doubled in 2020. The ICT sector is including personal computers (PCs) and peripherals, telecoms networks and devices, and datacenters. As focused in the telecoms core transport networks, energy consumption (i.e. carbon emission) of network equipment, such as Internet Protocol (IP)/Multi-Protocol Label Switching (MPLS) routers and Ethernet switches, are increasing rapidly, due to the exploding number of Internet users and increasing network transmission speed. It is estimated that the backbone network infrastructures, including IP/MPLS routers, transmission systems, and optical/electrical cross-connect systems (XCs), consume approximately 12% of total Internet energy consumption2, which is assumed to increase to 20% in 20202. Therefore, it is obvious that reducing energy consumption in the network is a crucial issue hereafter.

Previously, several techniques have been studied to reduce the network energy consumption3-8. These techniques can be categorized in to three approaches for simplification. The first approach is based on a concept of “sleeping components”3,4. It reduces the energy consumption of network components such as IP/MPLS routers and Ethernet switches by sleeping like a cellar phone. Reference 3 aims to save power consumption by putting network interfaces and other router and switch components to sleep. The timing to put these components to sleep is determined by estimating the expected inter-arrival time of traffic by monitoring traffic on all interfaces. The impact of the selection of which interface to put to sleep, is examined on the implementation of routing protocols, e.g. open shortest path first (OSPF) and internal border gateway protocol (iBGP). However, by alternately being turned on and off within short intervals, these components are concerned of being damaged. Reference 4 also aims to save power by switching the interface of components to sleep when no packets are buffered. However, the proposed method in 4 is a method to reduce the number of switching modes that result in quite short duration, by adopting “extra active mode”, which keeps the interface in active mode for an extra period after all packets are flushed out from the buffer. The second approach reduces the energy consumption of network links with concept of “adaptive link rate”5-7. The power consumption of links is constant regardless of the transmission rate. This approach focuses on adapting link rate and saving energy consumption. Reference 5 aims to reduce the energy consumption of Ethernet links by adaptively varying the link data rate in response to utilization. Some effective policies are proposed to determine when to change the link data rate. “IEEE802.3az Energy Efficient Ethernet (EEE)”8, which will be standardized a transition method to change the link rate of Ethernet interfaces from normal mode to energy-efficient mode. In Reference 7, a traffic optimum Ethernet physical media (PHY) (10Baste-T, 100Base-Tx, and 1000Base-T) is selected by medium access control (MAC) frame hand-shaking. The third approach, which this paper addresses, aims to reduce the energy consumption of the whole network by using network controlling techniques, i.e. traffic engineering (TE). This approach can be applied with first and second approaches.

*okamoto@ieee.org; phone 81 45 563-1151 ext. 48241; fax 81 45 560-1262; keio.ac.jp
*okamoto@ieee.org; phone 81 45 563-1151 ext. 48241; fax 81 45 560-1262; keio.ac.jp
We have been started a TE based network energy saving technology development project in early 2008\(^8\). The proposed concept is realizing a low energy consumption network by aggregating traffic on specific numbers of links and powering off link interfaces of routers/switches that are not used\(^8\)\(^-\)\(^10\). These operations are controlled by a path computation engine (PCE). A low energy consumption network design engine in the PCE creates optimal network topology\(^11\). Network equipment such as IP/MPLS routers, Ethernet switches, and XCs can be controlled by the PCE by Generalized Multi-Protocol Label Switching (GMPLS) based link power on/off control protocols\(^12\)\(^-\)\(^14\). This framework is named as “MiDORi (Multi- (layer, path, and resources) Dynamically Optimized Routing) Network Technologies”. “midori” is a Japanese word which means “green” in English. Under the MiDORi PCE, by monitoring the amount of traffic being transferred in the network\(^15\), the network topologies which include both a physical network topology and a path network topology can be dynamically reconfigured in an energy saving manner\(^17\). We already demonstrated the developed MiDORi Network Technologies in the Ethernet network\(^16\) and multilayer Ethernet and all optical network\(^17\). The third “traffic engineering” approach is now widely studied in other research institutes\(^8\)\(^-\)\(^21\).

This paper presents recent activities of the MiDORi networks technologies. The rest of the paper is organized as follows. In section 2, the MiDORi network architecture and technical components are described. In section 3, demonstrations of the MiDORi network testbed are presented. Finally in Section 4, summarize the paper.

## 2. MIDORI NETWORK ARCHITECTURE AND TECHNICAL COMPONENTS

### 2.1 TE based Low Energy Consumption Network Architecture

The amount of electrical energy consumption in one transmission link is constant regardless of the utilization of the link. This is because, the energy consumption of links during off-peak periods equals to that of busy periods. All packet gaps in layer 3 view or frame gaps in layer 2 view are fulfilled by the data. Therefore, in layer 1, a continuous bit rate stream is transmitted. From the viewpoint of the network operation, at off-peaks, all links in the network do not necessarily be powered on. All traffics in the network can be transferred through fewer links when the traffic is aggregated. Therefore, it is possible to save energy consumption with aggregating traffic into specific links and powering off other links.

![MiDORi Network System Architecture](image)

**Figure 1. MiDORi network system architecture.**

In the MiDORi network, as shown in Fig.1, the network management system monitors periodically the amount of traffic being transferred on each path. The MiDORi PCE receives a network topology and a traffic matrix of all paths from the network management system and computes energy optimum (or optimal) path layer accommodation design which leads
as much links to power off as possible. Calculated path layer accommodation design is reflected on the network via the MiDORi GMPLS protocol. The MiDORi GMPLS also reconfigures the physical layer network topology by controlling the ON/OFF state of all links.

The MiDORi network equipment which supports port power on/off and switching function power on/off is used for constructing the MiDORi networks. Candidates of the MiDORi network equipment are IP/MPLS routers, Ethernet switches, Synchronous Digital Hierarchy (SDH) XCs, Optical Transport Network (OTN) XCs, optical/photonic XCs, and transmission systems.

2.2 MiDORi PCE Energy Efficient Topology Design Algorithm

To get an optimum energy consumption network topology, selecting power off link procedure is very important. If we can examine all combinations of link ON/OFF patterns, we can get the optimum energy consumption network topology. However, this exhaustive search approach takes huge calculation time and is unsuitable for large network such as more than 20 nodes. If the optimum topology calculation can be written to use Integer Linea Programming (ILP), we can get the optimum topology. However, there is the same problem of the scaling issue. Thus, a practical calculation method is required to satisfy the calculation time. Additionally, link powering off may cause traffic congestion and adds extra hops. These may cause network performance deterioration such as traffic loss and propagation delay. Therefore, the least energy consumption network topology must be calculated under specific quality of service (QoS) requirements satisfaction. We have developed “depth d search algorithm” to overcome the calculation time problem and to satisfy the QoS restriction. The depth d search algorithm provides an optimal (not optimum) energy efficient topology.

2.2.1 Depth d Search Algorithm

The detailed operation of the depth d search algorithm is described in Fig.2. The algorithm iterates the process of determining one topology with the minimum value of r by comparing all patterns of topologies with arbitrary links powered off. Depth d can be set freely within the range of 1 to n-1. d can not exceed n-1 because the connectivity of nodes will be lost. d denotes the number of links powered off per iteration. r denotes the maximum link utilization. n denotes the number of nodes in the network. k denotes the number of on links in the network.

```
1: Input original topology.
2: Input traffic matrix.
3: Set d (d = 1, 2, ..., n - 1).
4: While d > 0 do
5: If k - d ≥ n - 1 then
6: Generate different patterns of topologies with d links powered off.
7: Check QoS of each topology.
8: If no topologies satisfy the QoS then
9: d ← d - 1
10: Else
11: Calculate r of each generated topology.
12: Choose one topology with the minimal value of r.
13: minr ← r
14: If minr > 1 then
15: d ← d - 1
16: Else
17: Replace the topology.
18: k ← k - d
19: End if
20: End if
21: Else
22: d ← d - 1
23: End if
24: While
```

Figure 2. Procedure of the depth d search algorithm.

The depth d algorithm can calculate 50 nodes network environment in 100 sec (d = 1), 500 sec (d = 2), and 30,000 sec (d = 3). Larger depth provides more optimal topology, however the calculation time is O(k^d+1). For the practical design algorithm depth 2 provides enough performance.

QoS restriction is evaluated in line 7 and 8 in Fig.2. Topologies that do not satisfy the QoS restriction are excluded in these steps. When none of the topologies satisfy the QoS restriction, d is decremented by 1. This is because a topology,
that satisfies the QoS restriction, might be found if the number of links to power off at a time is reduced. We have implemented QoS restrictions for end-to-end delay and usable capacity. Implementation of more strict QoS parameters, such as packet loss, delay jitter, etc., are for further study.

2.3 MiDORi GMPLS Protocol Extensions

The MiDORi PCE designs an energy optimal path route and physical topology. The designed path routes and link on/off information should be reflected to the network equipment. There are two methods for controlling path set up/tear down. One is direct setting from a management plane. Another is setting via a control plane. One of the features of the MiDORi network is to support “multi-layer”, “multi-path”, and “multi-resource”. To enhance the network scalability and multi-layer controllability, introducing the control plane and using the GMPLS protocol in the control plane is popular approach. The MiDORi network will have another feature called “self-organized network (SON)”. The self-organized covers many features such as plug-and-play, self-inventory, automatic neighbor discovery, automatic topology discovery, automatic traffic accommodation design, and automatic path provisioning. As a result, we can get an energy optimum network operation feature. The GMPLS protocol is useful to get self-organized feature. A routing protocol such as an open shortest path first (OSPF) and an intermediate system to intermediate system (IS-IS) supports the automatic topology discovery. A signaling protocol e.g. a resource reservation protocol (RSVP) supports the automatic path provisioning. A link management protocol (LMP) supports the self-inventory and the automatic neighbor discovery.

In GMPLS controlled networks, the network is described by label switch routers (LSRs) and TE links. A TE link is advertised as an adjunct to a "physical" link. When the link is up, both the regular internal gateway protocol (IGP) properties of the link and the TE properties of the link, such as bandwidth and switching capability, are then advertised. Therefore, basically, if the link is down then the TE link is also down. A TE link is not only defined between IGP neighbors but also defined on a forwarding adjacency (FA) label switched path (LSP). An LSP is composed with cross-connection of TE links. Therefore, if the composed TE link is down then the LSP is also down. This basic feature is not suitable for the energy efficient TE technique. Under the energy efficient TE, LSPs are rerouted to use least number of links, then some links are physically shutdown to reduce energy consumption of equipment. In traditional GMPLS networks, TE links associated in shutdown links are also down. Therefore, when emergency occurred, such as traffic explosion and link/equipment failure, downed TE links are not able to use for calculating protection LSP and LSP rerouting. To overcome this, MiDORi GMPLS extension is proposed. You can download the MiDORi GMPLS protocol software from “http://midori.yamanaka.ics.keio.ac.jp/?page_id=78”.

2.3.1 MiDORi GMPLS OSPF Extension

OSPF extension has a new logical TE link state for the link power up/down state. Figure 3 shows the basic OSPF extension images. In the current OSPF implementation, if a physical link becomes down then associated TE link is also down and removed from link database. As a result, PCE is not able to recognize the slept (power down) link. The slept link should be used when the traffic load becomes large and/or network failure occurred. The proposed OSPF extension keeps TE link in the link database. Therefore, slept TE links can be used when path network reconfiguration topology calculation is requested.

![Figure 3: OSPF extension image.](http://midori.yamanaka.ics.keio.ac.jp/?page_id=78)

2.3.2 MiDORi GMPLS LMP Extension

The energy efficient TE requires link power on/off control function. There are two possible implementations. One is using LMP, another is using RSVP. When using LMP, power on (or off) initiator LSR sends power on (or off) request to the neighbor LSR. The neighbor LSR sends Ack to the initiator LSR and power on (or off) the link and changes the TE link status. Then the initiator LSR receives Ack and power on (or off) the link and changes the TE link status. We have
implemented this procedure by using the LMP ChannelStatus message. To apply the power on procedure, IP control channel (IPCC) should be always up. Therefore, a dedicated IPCC is required to apply the LMP power on/off control.

2.3.3 MiDORi GMPLS RSVP Extension

When using RSVP, sequentially concatenated TE links can be controlled. There are two procedure candidates in the power off procedure shown in Fig.4.

[Power On] All TE links along with the LSP are power on.

[Power Off]
1. All TE links along with the LSP are power off. If other LSPs share the TE links then the LSPs should be rerouted.
2. All TE links but not shared by other LSPs are power off.

Both procedures are used according with the network operator's policy. Power control request has been implemented in the Admin_Status object for LSP status flag extension. The added flag indicates that the LSP is constructed on slept link(s) or not.

![Diagram showing RSVP extension power off procedure options. (a) option #1, (b) option #2.](image)

The power off procedure option #1 can be applicable not only to a single layer network but also to a multi-layer network. If the server layer TE-link becomes the “power off” state, upper layer LSP segment detects the status change and sends NOTIFY message to an LSP ingress node. The ingress node reroutes the LSP or changes the LSP status to “power off”.

2.4 Prototype MiDORi Network Equipment

To demonstrate the proposed energy optimal control system concept, we constructed a MiDORi network testbed. As a MiDORi network equipment, port power on/off controllable and switch function power on/off controllable Ethernet switches are developed. The required functions for constructing the MiDORi network equipment are following.

1. Power on/off controllable port. (Mandatory)
2. Port level traffic monitoring and path level traffic monitoring. (Mandatory)
3. Power on/off controllable switching module. (Optional)
4. Energy consumption visualization. (Optional)
5. Remote equipment control and management. (Mandatory)
6. Control plane protocol support, e.g. GMPLS, OpenFlow. (Optional)

The developed prototype equipment is an 8 ports Gigabit Ethernet (GbE) switch and a 16 ports Gigabit Ethernet switch. Both switch supports a traffic monitoring function of virtual local area networks (VLANs) assigned to each interface and switch module, which supports 4 GbE interfaces, power on/off control. The detailed specification is summarized in Table 1.

![Table 1. Prototype MiDORi Ethernet switch specification.](image)
Remote control and management | Management port 100Base-TX, Embedded Linux system
---|---
Switch implementation | FPGA (Virtex-5)
VLAN traffic monitoring | VLAN 1-127, all interfaces, send/receive packet counter and byte counter
Visualization | Front panel ammeter, Power consumption value via CLI
Unimplemented functions | Mac learning, Link aggregation, Spanning tree protocol
Size | W: 430 mm, D: 430 mm, H: 180 mm (8 port), 310 mm (16 port)
Maximum power consumption | 68.2 W (8 port), 106 W (16 port)

Power consumption reduction of 1.25 W/port and 17-21 W/module are achieved. This means that if 4 ports in the same switching module can be power off, 32 % energy saving for 8 ports switch and 21 % for 16 ports switch are possible. The embedded Linux system requires 3 to 7 W even if all ports and switch modules are power off.

### 3. MIDORI NETWORK DEMONSTRATIONS

#### 3.1 Cooperation Among MiDORi PCE, Network Management System, and MiDORi Network Equipment

To demonstrate the MiDORi network system concept, first, we constructed a small MiDORi network testbed. Figure 5 shows a sample test network structure. Five MiDORi Ethernet switches are connected via 7 GbE links. All switches are connected to the MiDORi PCE which includes network management system for traffic monitoring. Traffic information is reported to the network management system and the MiDORi PCE calculates the best network topology. The PCE control the traffic path route by controlling Ethernet VLANs. In this experiment, we can reduce total switch power consumption from 283.1 W to 276.5 W. The reduced power is 6.6W (2.3 % of the total power). However, this result shows that the MiDORi network technology is potentially feasible in the high speed interface and large scale network environment. This is because the rate that the power consumption of links occupies on the whole system becomes larger and the possibility of the switch module power off becomes high.

![Initial State, Total: 283.1 W](image1)

![After TE, Total: 276.5 W (-6.6W)](image2)

Figure 5. Five nodes MiDORi network trial. The MiDORi PCE calculates the energy optimal topology and sets the VLAN paths directly from the network management system.

Next, we examined the path reconfiguration time in the 6 nodes full mesh connection network environment. The required control time depends on the number of ports (control points) in the equipments. In this examination, the traffic amount ration is set as low traffic condition : high traffic condition = 1 : 5. The number of link reduction from the high traffic condition to the low traffic condition is 5. Used port number is reduced from 31 to 21. Traffic data collection time...
from all six switches to the network management system is 25.6 sec. The VLAN reconfiguration time is 29.0 sec. The optimal topology calculation time is 0.006 sec. Port On/Off control time is 7.2 sec. Total 61.8 sec is required. Our prototype switch does not support a “make before break” VLAN reconfiguration feature. Therefore, traffic data disruption is occurred during 29 sec in the network. The data disruption can be avoided by hard ware support of the “make before break” feature.

3.2 Cooperation Among MiDORi PCE, Network Management System, MiDORi GMPLS, and MiDORi Network Equipment

The MiDORi GMPLS is implemented into a small Linux box and the Linux box controls a equipment via telnet over Ethernet. We have developed MiDORi GMPLS enabled MiDORi Ethernet switch and optical switch. The MiDORi PCE is modified to enable energy efficient switching path selection\(^2\). If call set up requires larger bandwidth than preset threshold, the MiDORi PCE selects as much optical path switching route as possible. A multi-layer GMPLS signaling between a lambda switch capable (LSC) layer and a layer2 switch capable (L2SC) layer has been successfully demonstrated in iPOP2011\(^1\). Figure 6 shows the demonstrated network equipment and network topology. Traffic is increased and decreased with a period of 5 minutes. If the HD stream is requested the path which uses the optical switch is selected and the LSC LSP is set between 2 Ethernet switches.

![Figure 6. Multi-layer MiDORi demonstration network in iPOP2011. All equipment is controlled by the MiDORi GMPLS.](image)

4. CONCLUSIONS

This paper presented recent activities of the MiDORi networks technologies. The MiDORi PCE, the MiDORi GMPLS, and the MiDORi network equipment were successfully demonstrated. The MiDORi GMPLS is a key technology to realize SON feature in the MiDORi network. The prototype MiDORi Ethernet switch offers 1.25W/port power saving and 17 W/module power saving capabilities.

ACKNOELEGEMENT

This work is supported by the PREDICT program of the Ministry of Internal Affairs and Communications (MIC) of Japan.
REFERENCES