

# Diagonal Routing in High Performance Microprocessor Design

Noriyuki Ito, Hideaki Katagiri, Ryoichi Yamashita, Hiroshi Ikeda, Hiroyuki Sugiyama, Hiroaki Komatsu, Yoshiyasu Tanamura, Akihiko Yoshitake, Kazuhiro Nonomura, Kinya Ishizaka, Hiroaki Adachi, Yutaka Mori, Yutaka Isoda, Yaroku Sugiyama  
 Fujitsu Limited  
 4-1-1 Kamikodanaka, Nakahara-ku,  
 Kawasaki, 211-8588, Japan  
 ito.noriyuki@jp.fujitsu.com

## ABSTRACT

This paper presents a diagonal routing method which is applied to an actual microprocessor prototype chip. While including the layout functions for the conventional Manhattan routing with horizontal and vertical directions, a new diagonal routing capability is added as one of the routing functions. With this enhancement, diagonal routing becomes an additional strategy for improving delays of critical paths in the microprocessor design. This method was applied to the prototype chip of the Fujitsu SPARC64 microprocessor with two CPU cores using 90nm process technology. By applying the diagonal routing to long distance nets, net length is reduced by 36% per net on average. When the diagonal routing is applied to a critical path, path delay is improved by as much as about 14 pico-seconds per net on a path. This improvement is more than the delay of a gate with no load. This prototype chip proved that our method was effective in reducing the total net length and improving path delays.

**Keywords:** Microprocessor, diagonal routing, Manhattan routing

## 1. Introduction

For the performance improvement in microprocessor design, the continuous improvement of circuit design techniques and process technology is necessary. In this improvement, the CAD system also plays an important role. Designers achieve the targeted performance by tuning logical and physical designs to the last pico-second with various techniques [1]. For path delay optimization, there are several conventional techniques such as reduction of logical stages, custom macro design, buffer insertion, gate sizing, placement change, fan-out splitting, etc. Since interconnect delay is dominant in the deep sub-micron technology, some new techniques are used in routing to improve the delay. Timing-driven routing, which minimizes the delay from the driver pin of a net to the most critical receiver pin, is an effective routing technique. Another technique is using wider wires. As for the wire resistance, it is smaller for upper metal layers. This is because in general, a wire is wider and taller on an upper metal layer. Therefore, routing a critical net on an upper metal layer is another useful technique for improving the delay. After all these techniques are utilized, designers still improve delay by other techniques to achieve the targeted performance with the assistance of CAD tools. Although the delay improvement is small, pin permutation, which swaps nets connected to symmetric input pins of a cell so that a critical net is connected to a pin with

smaller gate delay, is applied. Also, a limited number of low  $V_{th}$  cells are used to improve delay on critical paths, even though the power dissipation may increase due to leakage current. These various techniques are used according to the progress of design in each stage. The technique that is used in the final phase of a design is very important even if the absolute delay improvement is small (e.g., a few pico-seconds). Therefore, it is very meaningful to provide designers of high performance microprocessors with new methods to improve delay.

This paper presents diagonal routing as a technique to improve delay of critical paths in microprocessor design. In order to add the new diagonal routing function to the current design methodology, several modifications are added to the existing in-house CAD tools. We made these modifications with a target completion deadline of at most six months. The rest of this paper is organized as follows. In Section 2, several diagonal routing techniques are surveyed, and the difficulty in the use of diagonal routing in LSI design is addressed. In Section 3, we estimate the contribution of improvement by our proposed diagonal routing scheme. Sections 4 and 5 describe the routing and capacitance extraction methods respectively. Experimental results are shown in Section 6, and Section 7 concludes the paper with directions for future work.

## 2. Related Work

Diagonal routing is not a new technique; it is routinely used in the printed circuit board design. In one of the Fujitsu mainframes, FACOM M-780 [2], one CPU is composed of 336 LSI's. These LSI's were mounted on both sides of a printed circuit board in a high density configuration [2]. In this printed circuit board design, diagonal routing was used to reduce delay between LSI's on the board [3]. In the context of LSI designs, several diagonal routing techniques have been proposed to reduce wire length and to improve density [4,5,6,7,8,9]. However, the use of diagonal routing has been very limited in LSI design. The reason is the difficulty to support diagonal routing in the related CAD tools such as routing, timing analysis, noise analysis, layout verification and manufacturing data generation. Recently, the pervasive use of diagonal routing with X-architecture was reported [10]. In the X-architecture, diagonal routing in 45 and 135 degree directions is used in conjunction with conventional Manhattan routing in horizontal and vertical directions. This architecture was applied to a RISC processor core. It was reported that the path delay improved by 19.8% and the area reduced by about 10%. As another routing model, Y-architecture based on 0, 60 and 120 degree directions is proposed [11]. With this

architecture, the wire length reduced by 13.4%. On average, the wire length was only 4.3% more than that of X-architecture. Since these routing architectures are different from the conventional Manhattan architecture, the development of CAD tools to support them is not easy. Even if we try to find some commercially available tools, it is very limited or none for LSI design. Although there are several recent research works on congestion reduction and Steiner tree generation assuming the use of diagonal routing [12, 13], actual use of diagonal routing in an industrial chip is not popular yet. We believe that to develop a high performance microprocessor, it is very important to enhance the existing CAD system with the new capability of diagonal routing. In this paper, we propose a new routing technique that uses diagonal wiring partially on a critical net. Such limited usage of diagonal wires minimizes the enhancement needed for the existing CAD system.

### 3. Preliminaries

Given two pins  $A$  and  $B$ , let  $L_{xy}$  denote the length of the net when it is routed with conventional Manhattan routing. Let  $L_{angle}(\alpha)$  denote the net length when the same net is routed diagonally in  $\alpha$  degree direction. As shown in Figure 1, let the elevation from  $A$  to  $B$  be  $\theta$  degree. Then,  $L_{xy}$  and  $L_{angle}(\alpha)$  are expressed by (1) and (2) respectively when  $\theta$  is within  $\alpha$  and when  $\theta$  is within 90 degree.

$$L_{xy} = L(1 + \tan \theta) \tag{1}$$

$$L_{angle}(\alpha) = L \tan \theta / \sin \alpha + L(1 - \tan \theta / \tan \alpha) \tag{2}$$

When diagonal routing is used in a chip,  $\alpha$  is fixed. The length of a net routed with a diagonal wire is the minimum when  $\theta$  is equal to  $\alpha$ . This is because a diagonal wire can connect two points directly with no Manhattan wire. Assuming that  $\theta$  is equal to  $\alpha$  and that  $\alpha$  varies from 0 to 90 degrees, the reduction ratio  $(L_{xy} - L_{angle}(\alpha)) / L_{xy}$  is the maximum when  $\alpha$  is equal to 45 degrees. As shown in Figure 2, the maximum reduction ratio is about 29.3%. Therefore, 45 degree direction is the most effective choice for diagonal routing to reduce the length of a net.

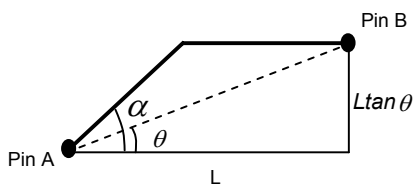


Figure 1.  $\alpha$ -degree diagonal routing

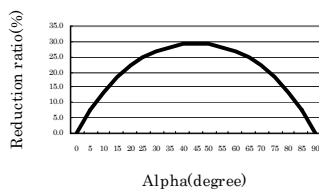


Figure 2. Length reduction ratio for  $\alpha$ -degree diagonal

Before actually applying our diagonal routing to the microprocessor design, we evaluated the net-length reduction when diagonal routing is applied to the actual design. When net length with a diagonal wire is estimated, the center location of receiver pins is calculated first. Then, for the connection between a driver pin and the center location, net length is estimated by both Manhattan and diagonal routing, as shown in Figure 3. The two lengths are compared for each net. We performed this evaluation for 6,000 long distance nets. The results are shown in Table 1. We see that the net length is reduced by about 12% on average by diagonal routing.

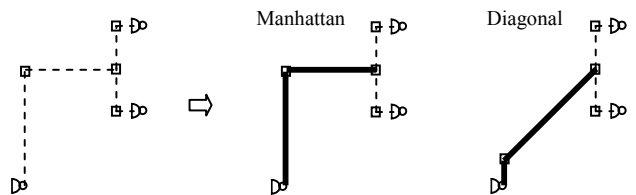


Figure 3. Manhattan routing and diagonal routing

Table 1. Length reduction by diagonal routing

Unit: grid

	Manhattan	Diagonal	difference
Total length	25,835,198	22,569,954	3,265,244
Average Net length	4,090	3,573	517

We also evaluated the improvement in delay for the typical model shown in Figure 4. In this model, the length of the net estimated using Manhattan routing is  $3L$ , and using diagonal routing is  $(1 + \sqrt{2})L$ . Diagonal routing can reduce the net length by about 20% in this case. For this model, delay is measured by HSPICE, and  $L$  is varied from 0 to 1500 grids. Figure 5 shows delay from the driver to a receiver for Manhattan and diagonal schemes. When  $L$  is 1500, the improvement of delay is about 7 pico-seconds. As it is clear from the graph, the improvement is larger when diagonal routing is applied to a longer net.

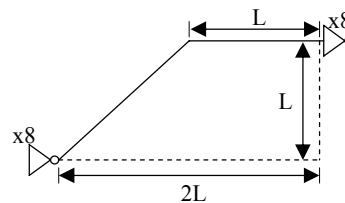


Figure 4. Model of a net with a diagonal wire

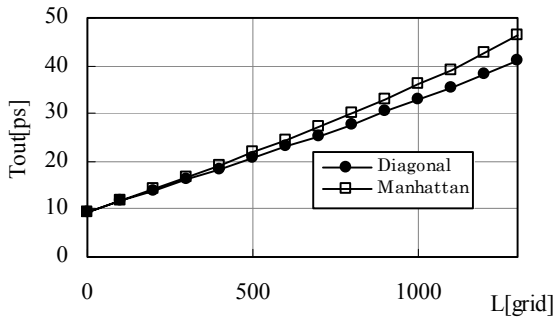


Figure 5. Reduction of delay by diagonal routing

#### 4. 45, 135 Degree Diagonal Routing

In order to apply our diagonal routing technique to a microprocessor prototype, it is desirable that we enhance our existing CAD tools with minimum effort. Therefore, we limited the use of diagonal wire only to critical nets. To meet these requirements, a diagonal wire is routed first as a straight-line trunk between driver and receivers. Then, the connection between the driver and the diagonal wire is routed using conventional Manhattan routing. Finally, the connection between the diagonal wire and receivers is routed using Manhattan routing. Figure 6 shows this idea.

An algorithm *Diagonal\_routing* to route a given net with diagonal wire is shown in Figure 7. A net is passed to *Diagonal\_routing* with a set of pins  $P$  of the net. *Routable\_check* checks whether diagonal routing can be applied or not. When a driver pin lies within the minimum bounding box of the receiver pins, the given net is routed by a conventional Manhattan router and the procedure *Diagonal\_routing* terminates. When the driver pin lies outside the bounding box of receiver pins, a further check is performed. A Steiner tree  $S$  is created for  $P$  by *Steiner\_generation*.  $S$  is passed to *Make\_receiver\_group*. This procedure finds a sub-tree  $S'$  of  $S$  that connects only receiver pins, and it returns a connection point  $R$  that connects  $S'$  and the rest of  $S$ . *Select\_driver* returns a driver pin  $D$  from  $P$ . Then, both  $R$  and  $D$  are passed to *Diagonal\_search*, which searches a diagonal route with no bend between  $R$  and  $D$  and returns the result as  $W_{diag}$ . The direction 45 or 135 degree is determined depending on the relative positions of driver and receivers. If no diagonal route is found, value *NONE* is returned. By passing  $W_{diag}$  and  $P$  to *Manhattan\_routing*, the connection between the diagonal wire  $W_{diag}$  and the driver pin  $D$ , and connection between  $W_{diag}$  and receiver pins are routed by the conventional Manhattan routing function. In *Diagonal\_search*, diagonal route search is not performed when the estimated length of the diagonal wire is less than some specific length. This is to apply diagonal routing only when its effect will be appreciable. When diagonal routing is applied between two points, the vertical and horizontal distances between these points for a diagonal route search are compared. When the horizontal distance is greater than the vertical distance, a diagonal route is searched in the horizontal direction by *Horizontal\_search*. Otherwise, a diagonal route is searched in the vertical direction by *Vertical\_search*. In *Horizontal\_search* or *Vertical\_search*, the angle of diagonal routing (45 or 135 degree) is determined based on the relative locations of  $D$  and  $R$ . Figure 8 shows these two types of diagonal route search. The length of the

diagonal wire is constant in the range from (b) to (c). Since the length of the diagonal wire varies in the range from (a) to (b) or from (c) to (d), search is performed within the range such that the length is larger than some specific value.

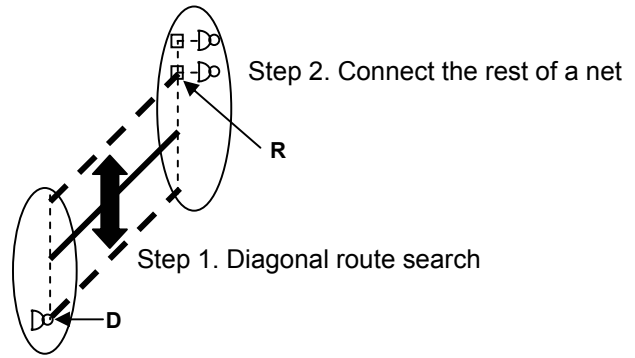


Figure 6. Basic procedure in diagonal routing

```

Diagonal_routing(P)
  flag = Routable_check(P)
  if flag == FALSE {
    W,V = Manhattan_routing(P)
  }
  else {
    S = Steiner_generation(P)
    R = Make_receiver_group(S)
    D = Select_driver(P)
    Wdiag = Diagonal_search(D,R)
    Wires,Vias = Manhattan_routing(P,Wdiag)
  }
  return(Wires,Vias)
end

Diagonal_search(D,R)
  L = Diagonal_length(D,R)
  if ( L < α ) then
    return(NONE)
  }
  else {
    x = Horizontal_distance(D,R)
    y = Vertical_distance(D,R)
    if ( x > y ) {
      diag_wire = Horizontal_search(D,R)
    }
    else {
      diag_wire = Vertical_search(D,R)
    }
  }
  return(diag_wire)
}

```

Figure 7. Diagonal routing algorithm

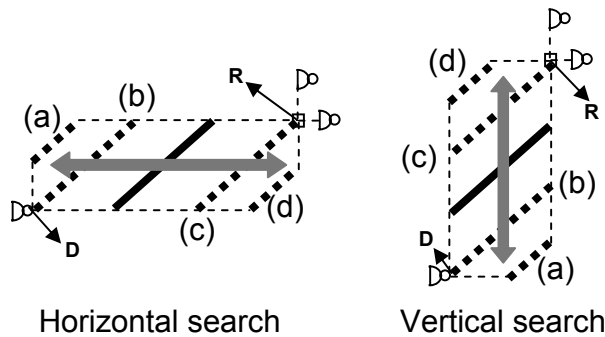


Figure 8. Search direction of diagonal route

Since diagonal routing is applied only to critical nets in our method, 45 and 135 degree layers are assigned to upper layers  $n$  and  $n+1$ , which have smaller wire resistance. After all nets with diagonal wiring are routed, some channels may be left unused on those layers. These available channels are used to route regular nets with the conventional Manhattan routing, as shown in Figure 9.

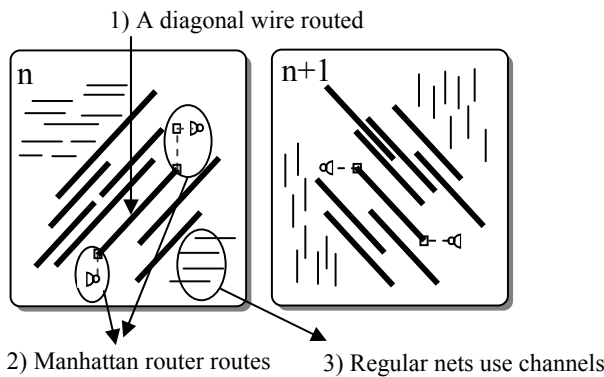


Figure 9. Manhattan and diagonal mixed layers

### 5. Capacitance Extraction

In order to run timing analysis for design data with diagonal wires, it is necessary to enhance the capacitance extraction tool to handle diagonal wires. Here, we will give an overview of the enhancement we made to our in-house capacitance extraction tool for microprocessor design. In this tool, capacitance for each wire is extracted grid by grid from a pre-computed table using the wire location and its adjacency with other wires. For a horizontal or vertical wire, a capacitance table CAPTBL\_M, which assumes the interval between two wires as  $N$  (1, 2, 3, ...) times of a regular grid, is used. For a diagonal wire with 45 or 135 degree, an additional grid with the  $1/\sqrt{2}$  grid length is introduced as shown in Figure 10. For the capacitance extraction of a diagonal wire, another capacitance table CAPTBL\_D, which assumes the interval between two wires as  $N$  (1, 2, 3, ...) times of a  $1/\sqrt{2}$  grid, is added.

With the addition of CAPTBL\_D, a procedure to search wires adjacent to a diagonal wire in the orthogonal direction is added to

the extraction tool. Figure 11 shows an example in which a wire adjacent to a given 45 degree diagonal wire is searched for in the orthogonal direction. As shown in this figure, whether the other adjacent wire exists or not in the  $1/\sqrt{2} * N$  grid is checked on the  $1/\sqrt{2}$  grid by  $1/\sqrt{2}$  grid. If  $N$  is even, the search point with the distance  $1/\sqrt{2} * N$  lies on the regular grid. If  $N$  is odd, the search point in the orthogonal direction does not lie on the regular grid. In this case, two regular grid points nearest to the search point are checked. According to this search procedure, regular grids are checked in the order of 1, 2, 3, 4, 5 which are the labels on a regular grid as shown in Figure 11. In this example, an adjacent wire is found when the search reaches the point with distance of  $1/\sqrt{2} * 5$  grids from the given wire.

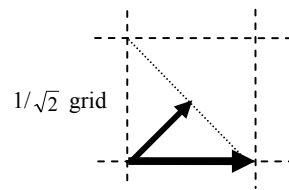


Figure 10. Basic grid for capacitance extraction

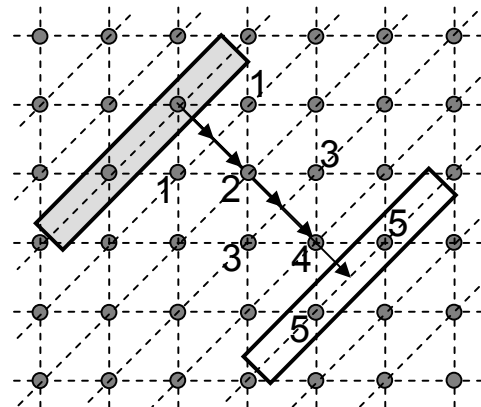


Figure 11. Capacitance extraction for a diagonal wire

### 6. Experimental Results

In order to apply the diagonal routing technique described in this paper, we enhanced the existing in-house CAD tools and database. The enhanced tools are automatic router, interactive placement and routing editor, RC extraction, timing analysis and layout verification. Calibre from Mentor Graphics was used for DRC and LVS sign-off. An actual prototype chip, which is called MTP (Multi Threaded Processor), is developed [14]. Figure 12 shows the micrograph with two CPU cores and six mega-byte second caches. This prototype chip is developed to verify MTP architecture, 90 nm process technology, diagonal routing, etc. In this chip, diagonal routing is applied to the most critical nets between an instruction unit and RAM's. The total number of nets to which diagonal routing is applied is about 1,500 nets per CPU core. The left figure in Figure 13 shows the 8<sup>th</sup> layer of an

instruction unit, and the right figure zooms in on the layout part containing diagonal wires.

We evaluated the improvement in net length and path delay with diagonal routing vis-à-vis the conventional Manhattan routing. Figure 14 shows this comparison for net length for about 1,500 nets to which diagonal routing is applied. Diagonal routing reduces the net length by about 36%. Since diagonal routing is applied only to a trunk of a net with no vias, the total detour of the net with diagonal routing is smaller than that with the Manhattan routing. Therefore, the maximum ratio of length reduction exceeds 29.3%, which was estimated in Section 3.

Figure 15 shows how the path delay is improved when diagonal routing is applied to critical nets. The number of nets with diagonal routing is 1,500, and the number of paths that pass them is about 27,000. In our prototype chip, one net on each path is routed by diagonal routing. The maximum improvement in path delay is about 14 pico-seconds, the average improvement being about 6 pico-seconds. There are some paths whose path delay increases by 2 pico-seconds. This is because the coupling capacitance with diagonal routing happens to be larger than that with the Manhattan routing. In general, coupling capacitance that can cause cross-talk noise is less when diagonal routing is applied. This is due to the associated reduction of net length. The graph in Figure 16 shows the ratio of the coupling capacitance per net with diagonal routing to that with the Manhattan routing. Although the coupling capacitance increases in some cases, overall coupling capacitance decreases by about 15%. This implies, in turn, that diagonal routing is effective from the point of view of reducing cross-talk noise.

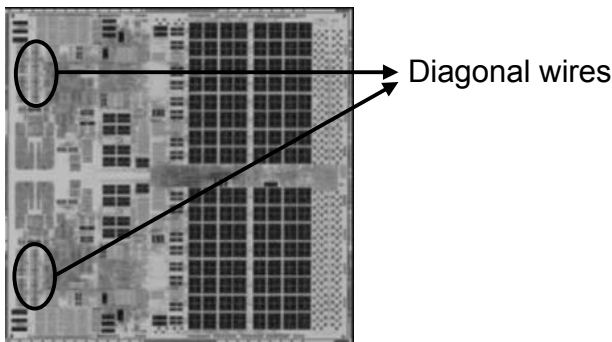


Figure 12. Prototype chip [14]

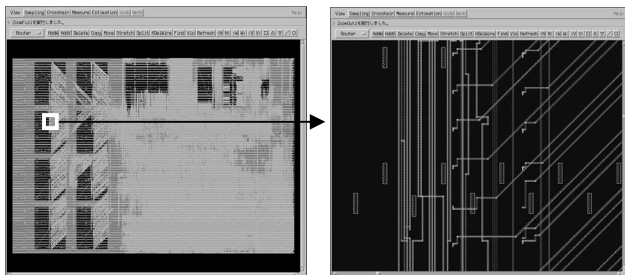


Figure 13. Actual diagonal wires

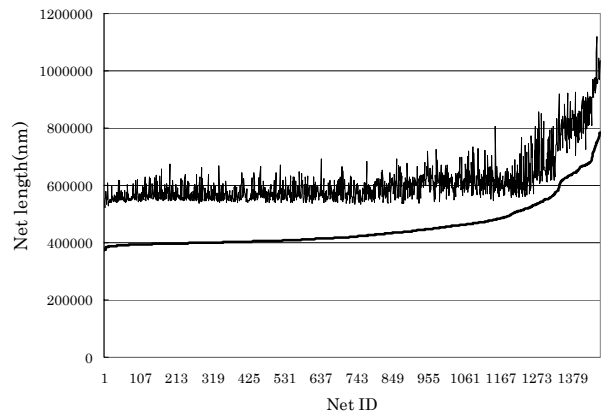


Figure 14. Reduction of net length

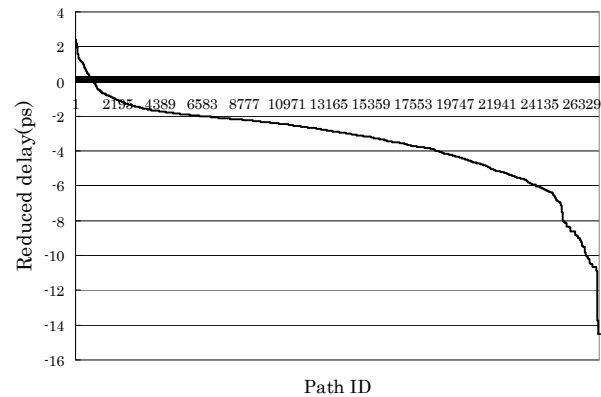


Figure 15. Reduction of path delay

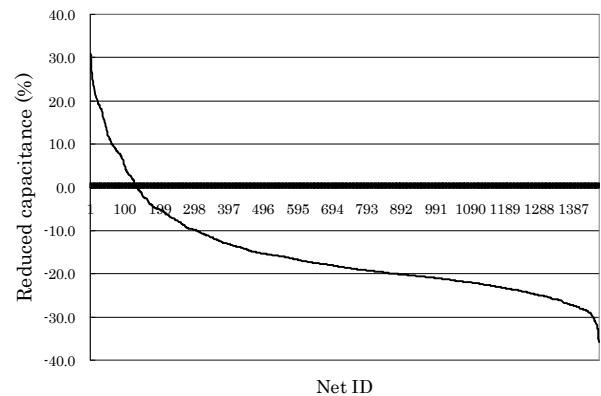


Figure 16. Reduction of net capacitance

## 7. Conclusion

With this work, we were able to provide microprocessor designers with diagonal routing capability as a new means to



improve path delay. This capability is built into our in-house CAD system for microprocessor design. We applied it to an actual prototype chip of a microprocessor with two CPU cores. We were able to reduce the net length by 36% per net (on average), and path delay by up to 14 pico-seconds. This delay improvement is more than the delay of a gate with no load. These results prove that the use of diagonal wires can improve the delay of critical paths in microprocessor design.

In the prototype processor design, diagonal routing was applied to about 3,000 nets in the chip. To increase the number of nets to which we can apply diagonal routing, further enhancements are needed, such as automatic net selection and floorplanning & cell placement that make diagonal routing more effective. We plan to address these issues in the near future and hopefully make diagonal routing an effective strategy for improving path delays in microprocessor designs.

## 8. ACKNOWLEDGMENTS

We would like to thank Toshiyuki Shibuya and Kaoru Kawamura and other researchers in the CAD group of Fujitsu Laboratories for their enhancement of the automatic routing engine. We also wish to thank Eizo Ninou and his technology group for the support they provide for this work.

## 9. REFERENCES

- [1] N. Ito, H. Komatsu, et al, "A Physical Design Methodology for 1.3GHz SPARC64 Microprocessor", Proc. ICCD, pp. 204-210, 2003.
- [2] M. Nishihara, T. Murase, et al, "Single-Board CPU Packaging for the FACOM M-780", Fujitsu Scientific and Technical Journal, 23, 4, pp.226-235, 1987.
- [3] M. Oda, T. Hamaguchi, "Oblique Routing System for FACOM M-780 SSC", Fujitsu Scientific and Technical Journal, 23, 4, pp.236-242, 1987.
- [4] H. H. Chen, "Routing L-Shaped Channels in Nonslicing-Structure Placement", Proc. DAC, pp. 152-158, 1987.
- [5] D. Marple, M. Smulders, et al, "An Efficient Compactor for 45° Layout", Proc. DAC, pp. 396-402, 1988.
- [6] C. Chiang, M. Sarrafzadeh, "Wirability of Knock-Knee Layouts with 45-degree wires", IEEE trans. On Circuits & Systems, vol. 38, No. 6, June 1991, pp. 613-624.
- [7] D. T. Lee, C. F. Shen, et at, "On Steiner Tree Problem with 45° Routing", Proc. ISCAS, pp. 1680-1682a, 1995.
- [8] S. Das, B. B. Bhattacharya, "Channel Routing in Manhattan-Diagonal Model", Proc. VLSI Design, pp. 43-48, 1996.
- [9] C. K. Koh, P. H. Madden, "Manhattan or Non-Manhattan? A Study of Alternative VLSI Routing Architectures", Proc. GLSVLSI, pp. 47-52, 2000.
- [10] I. Mutsunori, T. Mitsuhashi, et al, "A Diagonal-Interconnect Architecture and Its Application to RISC Core Design", Proc. ISSCC, pp. 684-689, 2002.
- [11] H. Chen, B. Yao, et al, "The Y-Architecture: Yet Another On-Chip Interconnect Solution", Proc. ASPDAC, pp. 840-846, 2003.
- [12] A. R. Agnihotri, P. H. Madden, "Congestion Reduction in Traditional and New Routing Architecture", Proc. GLSVLSI, pp. 211-214, 2003.
- [13] C. Chiang, Q. Su, "Wirelength Reduction by Using Diagonal Wire", Proc. GLSVLSI, pp. 104-107, 2003.
- [14] Takumi Maruyama, "SPARC64 VI: Fujitsu's Next Generation Processor", presented at the microprocessor forum, 2003..