Power Distribution Topology Design

Ashok Vittal and Malgorzata Marek-Sadowska

Department of Electrical and Computer Engineering,
University of California, Santa Barbara, CA 93106

Abstract

We propose topology design of power distribution nets using a novel method for capturing the temporal characteristics of sink currents - the current compatibility graph. This graph carries information necessary for net area optimization. We propose a new algorithm for simultaneous topology design and wire sizing that can handle large designs. Our techniques result in significant area improvements on benchmark instances.

I. Introduction *

The power and ground nets of an electronic system supply reference voltages to the circuit elements. The resistance of the interconnect causes voltage drops between the pad and the sinks. These voltage drops are undesirable as they result in speed degradation due to smaller current drive. They also introduce timing noise which may invalidate timing verification. Besides, the interconnect metal can carry a prescribed maximum current density if the reliability of the system is not to be compromised. The power distribution synthesis therefore affects the reliability of the system. As the power supply net connects a large number of sinks, its area is a significant fraction of the total available layout area. With technologies scaling down, interconnect resistance increases, total current drawn increases and noise margins are reduced, resulting in reduced signal-to-noise ratios [14] and making power distribution synthesis more critical. Besides, portable applications have motivated low power design. The primary means of achieving low power is by scaling down the supply voltage, which further reduces the noise margin and good power distribution is essential for correct system operation. The size of the problem also gets larger as the number of sinks increases. The problem has therefore received considerable attention.

The first generation power distribution synthesis tools concentrated on single layer routing of the power and ground nets [13], [16]. The emphasis was on connectivity and electrical effects were not considered. Subsequent research focussed on changing wire widths of a *given topology* to satisfy electromigration and voltage drop constraints, while minimizing area requirements [3], [6], [10]. The characteristics of the problem allow the sizing problem to be solved by optimization techniques. While the significant influence that the topology has on the final layout area was recognized, good topology design remained open.

Topology optimization was proposed in [12]. This approach was specifically meant for standard cell layouts. Enhancement buses of minimum area were added to pre-existing power buses so that performance constraints were met. However, high performance systems are usually full custom, so that unrestricted topologies can be better than highly constrained topologies in terms of area. Nevertheless, this work clearly demonstrated that topology design was an important degree of freedom.

Simultaneous sizing and topology optimization is considered in [15]. The entire problem is solved using simulated annealing. However, all the topologies in the search space are sub-sets of a “general grid”. This is restrictive and misses a large class of optimum solutions. For instance, star routing to pins in the same channel is not allowed by the formulation. This is a shortcoming due to the use of a simulator in the optimization loop. In order to keep the simulation tractable, parallel wires in the same channel are not allowed, losing optimality. Besides, the time required is tens of hours for instances with tens of sinks. While the power supply routing problem is usually solved only once during a chip design, power supply routing using this approach for LSI designs with tens of thousands of sinks is ruled out. The approach is meant mainly for analog power distribution where noise coupling is critical and the problem scale is small. The small scale of the problem (tens of modules) allows simulated annealing to be used for optimization; this does not exploit the special structure of the problem and is therefore computationally expensive. The philosophy of using a simulator in a general optimization loop is inherently computationally expensive. Optimization techniques that exploit problem structure can return much better solutions.

In [8], a topology optimization procedure was proposed that starts with a sized power distribution mesh and returns a sized power distribution tree which satisfies the same electromigration and voltage drop constraints, while using considerably less area. This was extended in [7] to show that area-minimal topologies are trees. A characterization of those trees which are area minimal was also given.

Previous work on power net routing has not used information regarding the times at which sinks draw currents. We show that this misses out on a significant source of optimization. In particular, we show that star routes, shown to be optimal in [7], can be *arbitrarily bad* if the problem formulation allows the use of such temporal information. The peak currents drawn by the sinks have to be used for voltage drop

*. This work was supported in part by the Advanced Research Projects Agency under contract DABT63-93-C-0039, NSF Grant MIP 9117328 and by AT&T Bell Labs and Xilinx through the California MICRO program.

32nd ACM/IEEE Design Automation Conference ®

computation. Usually, however, all sinks do not require their respective peak currents at the same time instant. For instance, in domino logic a stage does not evaluate till its predecessor is done evaluating. This disjoint temporal nature of sink current requirements can be used to optimize the power supply route. Proponents of self-timed systems cite this as a major incentive[5]. However, no previous power distribution tool takes advantage of this. We introduce the notion of a current compatibility graph in Section II to capture the temporal sink current information necessary for area optimization. This graph enables the quick evaluation of tighter worst case scenarios so that optimizing routes become possible.

The use of the current compatibility graph affects the structure of optimal solutions. For instance, the merging of two sub-trees which never draw current at the same time would be preferred. Section III explores the relationship between performance driven interconnect design, clock routing and power distribution design to enable the reuse of intuition gained on the other problems. For instance, the need for bounding sink voltage differences is analogous to bounding the skew during clock routing. However, we will see that traditional clock or signal net routers are inadequate for power net routing, motivating our greedy heuristic. Our heuristic designs the topology and sizes the wires simultaneously.

Section IV reports our results and Section V concludes with a recap of the major contributions and directions for future work.

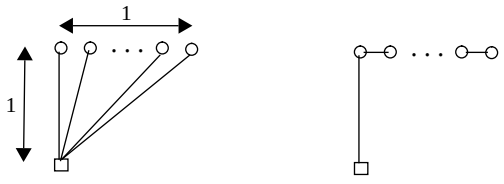
II. Problem Formulation

We wish to design minimal area, sized power supply nets given only the positions of the sinks, currents drawn at sinks, temporal sink current information and technology information. The temporal sink current information is in the form of a current compatibility graph. This is explained below.

Electromigration in power supply routes is avoided by limiting the maximum current density [14]. We therefore need to estimate the maximum current density in each branch of the route over all input patterns of interest. This is inherently a simulation problem. As there is also a limit on the current a pad can carry, estimation of the minimum number of supply pads also requires such information and techniques for finding the maximum current drawn by a chip have been proposed in [2], [11], [18]. Thus, given a power supply route, it is possible to estimate the maximum current drawn by any branch. These techniques solve the *analysis* problem. However, we are addressing the *design* issue. If we are simultaneously designing the route and the wire sizes, optimization of routes using time-domain sink current characteristics requires a new concept - the current compatibility graph.

We say that two sinks i and j are *current-compatible* if and only if i and j never draw current at the same time. This information can be obtained either by simulation or by structural analysis of the circuit netlist. The *current compatibility graph* G(V,E) is given by the set of sinks V and the set of current-compatible sink pairs E. The current compatibility graph captures the information essential for computing better bounds on wire width so that routing area can be saved. Note that it is possible to get such a graph from structural analysis rather than explicit simulation. For instance, the levels of a Boolean network could give the times at which gates switch during a clock period. In this case, there is an edge between two sinks if the corresponding gates are at different levels. This is, of course, only an existence proof. More sophisticated methods for doing this are possible. A high level synthesis system, following scheduling and allocation, may give us information about disjoint temporal usage of functional units.

Figure 1 shows that such temporal information can be used for net area optimization. The star route, which is optimum if no temporal information is available, uses 1.5n wire length while the optimum route uses only 2 units of wire length. It is assumed that all sinks require currents at disjoint intervals of time.



(i) star route (ii) optimum route

Figure 1. The star route can be arbitrarily bad. i) the star route uses O(n) wirelength. ii) the optimum route uses O(1) wirelength as no two sinks draw current at the same time.

Thus we state the problem as

Given P = {p_1, p_2,..., p_n} a set of sinks on the Manhattan plane, the interconnect resistance per unit length R_0, the maximum allowed vertical voltage drop V_{vmax}, the maximum horizontal voltage drop V_{hmax}, the peak sink currents {I_i}, the current compatibility graph G (the novel feature of our formulation) and the maximum allowed current density J_{max},

$$Minimize \sum_{j \in B} w_j \bullet l_j$$

where B is the set of branches, w_j and l_j are the width and length of branch j. The cost function is therefore the net area.

The electromigration constraints are

$$(w_j \geq \frac{I_j}{J_{max}}) \; \forall j \qquad (1)$$

This means that the width of a branch should be large enough to keep the estimated maximum current density within limits.

The vertical voltage drop constraints are

$$\left(\sum_{j \in P_i} I_{jmax} R_0 \frac{l_j}{w_j} \leq V_{Vmax} \right) \forall i \qquad (2)$$

where the constraints are for all leaf nodes (sinks), P_i is the unique path from the root to the i^{th} sink, R_0 is the resistance per unit length and I_{jmax} is the current in the sub-tree downstream of the branch j.

The horizontal voltage drop constraints are

$$(|V_i - V_j| \le V_{Hmax})\ \forall(i,j) \qquad (3)$$

The horizontal voltage drop constraints impose the need for bounding the difference between reference voltages seen by communicating circuit elements. For CMOS, they are implied by the vertical voltage drop constraints [3], as the minimum non-zero current drawn by a sub-tree is very small. However, this may not be the case for custom ECL circuits which essentially draw large, constant currents.

The decision version of the problem is NP-complete by restriction to the minimum Steiner tree problem. Without electromigration and voltage drop constraints, the problem is one of finding a Steiner tree of minimum cost. This means that efficient exact solution in polynomial time is unlikely. This leads to the quest for good polynomial time approximation algorithms.

III. A Heuristic

In this section, we show that power net routing can be formulated as signal routing with delay and skew bounds, with an added dimension - temporal information. Conventional signal routers are inadequate for power net routing as they do not design topologies and size them simultaneously. We therefore propose a new algorithm for simultaneous topology design and wire sizing.

The expression for the voltage drop at sink i is the LHS of inequality (2). It is isomorphic to the expression for calculating the Elmore delay in the performance-driven interconnect design problem [4]. The expression for the delay to the j^{th} sink with resistance per unit length R_0, driver resistance R_d, capacitance per unit length C_0 and load capacitances C_{Li} is given by

$$D_i = \sum_{k \in P_i} C_k R_0 \frac{l_k}{w_k} + R_d C_0 L(T) + R_d \sum_{\forall i} C_{Li}$$

where C_k is the downstream capacitance seen by branch k (this is the sum of sub-tree load capacitances and sub-tree interconnect capacitance), l_k and w_k are the length and width of branch k.

The expressions for voltage drop and sink delay are isomorphic if C_0=0 and R_d=0 in the sink delay equation. The delay in performance-driven interconnect design corresponds to the voltage drop in our problem and the sink capacitances correspond to the sink currents. Wire sizing behavior is the same too - wire widening decreases the interconnect resistance in both cases. The horizontal constraints are exactly the same as bounded skew constraints in clock routing. There are differences however - the load capacitances are constant while the load currents are time-varying. Besides, the temporal information adds a dimension that is not present in signal routing. The sub-tree current is a function of the entire set of sinks it contains.

Figure 2 shows that this new dimension renders conventional methods ineffective. The problem instance is shown in (2i). The sink and pad positions are shown on the left and the current compatibility graph is shown on the right - sinks A and B do not draw currents at the same time. Star routing, minimum Steiner routing and bounded-radius bounded-cost trees are all sub-optimum. The added dimension calls for algorithms that see the critical temporal information.

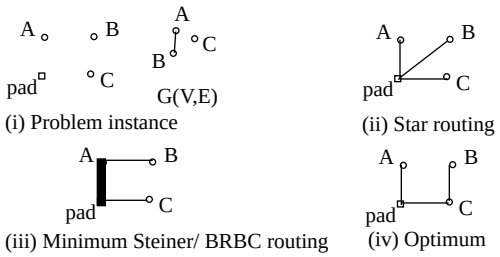


Figure 2. New dimension added by temporal information

Our problem appears to be related to the performance-driven interconnect design problem and to the clock routing problem - problems which have seen considerable research. Greedy methods have been effective for performance driven interconnect design [1], [4], [9]. It is therefore natural to expect a greedy algorithm to do well for power supply net routing too. We propose to build the power net in a bottom up fashion, simultaneously designing the topology and sizing the wires. The basic “move” in a greedy algorithm for our problem consists of a merge of sub-trees. We discuss what a greedy merge should do.

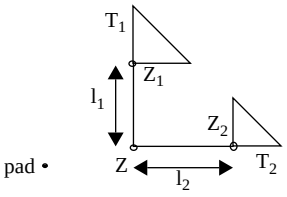


Figure 3. Merging sub-trees T_1 and T_2 to get a new sub-tree T

Consider two sub-trees T_1 and T_2, rooted at positions Z_1 and Z_2 on the Manhattan plane. Let the maximum voltage drops from the sub-tree roots to any of the sinks be V_{M1} and V_{M2} respectively. Let their sub-tree currents be I_1 and I_2. We need to decide a merge point as the position of the root. This is chosen to be the point on the bounding box of the two sub-tree roots that is closest to the pad. The widths of the segments w_1 and w_2 are determined next. This is done using the wire sizing techniques of [10] for homogeneous vertical voltage drop constraints.

We have the electromigration constraints

$$(w_i \ge \frac{I_i}{J_{max}})\ i = 1, 2$$

The voltage drop constraints are

$$(V_{max} \ge V_{Mi} + I_i R_0 \frac{l_i}{w_i})\ i = 1, 2$$

If the minimum widths given by the electromigration constraints are insufficient for satisfying the voltage drop

constraints, we size up the entire sub-tree by a factor α. This decreases the sub-tree voltage drop by a factor of α but increases the sub-tree area A_i, so that there is an optimum α for minimum area increase, while satisfying constraints [10].The optimum α and w_i can be computed in O(1) time for any pair of sub-trees. The new largest voltage drop to any sink can also be found in O(1) time.

The new sub-tree current is the largest current that the sub-tree current can possibly draw. Such a bound can be found using the current compatibility graph induced by the vertex sub-set consisting of sinks from the sub-tree. Let the edge weighted current compatibility graph be the current compatibility graph with the weights being the smaller of the two sink currents for that edge. Then a sub-tree current bound can be inferred using a maximal matching on the edge weighted current compatibility graph. We illustrate this with an example.

Consider the current compatibility graph shown in Figure 4. There are four sinks drawing currents of 3, 1, 4 and 8 units in a sub-tree. The weight of an edge in the graph is the gain that result by using the edge. Using conventional methods, the bound that would be obtained for the sub-tree current is 16 (the sum of sink currents). Using a maximal matching, however, we get 11 as the bound on the maximum sub-tree current. Each edge contributes a gain equal to its weight. The maximal matching gives us a better bound.

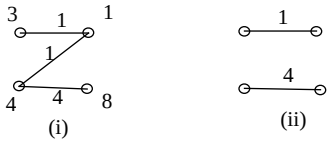


Figure 4. Using matching to compute sub-tree currents.
i) The edge weighted current compatibility graph
ii) The maximum matching

Our heuristic chooses the sub-trees to merge, using a minimum area increase criterion as described above. In other words, we find the cost increase for the merge of each pair of current sub-trees and choose the merge that gives the smallest area increase. The algorithm is outlined below.

```
Algorithm GREEDY
Input: Sink positions {Z_i}, voltage drop & electromigration constraints,
       technology
Output: Sized topology of minimal area

begin
    initialize list of sub-trees to the set of sinks
    repeat{
        minimum_cost_merge()
    } n-1 times
end

minimum_cost_merge()
begin
    find min{cost of sub-tree merges for all pairs of sub-trees}
    merge the sub-trees
end
```

The time complexity of the algorithm is $O(n^3)$ as we have n merges and each merge decision takes at most $O(n^2)$ time. Each sub-tree is represented by its area, its maximum voltage drop to any sink and its current, so that a merge of sub-trees involves computing these quantities for the new sub-tree. This can be done in O(1) time.

IV. Results

We implemented our algorithm in C on a DEC 5000 workstation. We compare our algorithm with the previous approaches of wire sizing of star routes. This represents the conventional solution to the problem [7]. We check how useful time domain current information as represented by the current compatibility graph is in our experiments.

There are no benchmarks for power supply routing. We therefore introduce our own benchmarks, based on the widely available clock routing benchmarks from [17]. We derive 8 benchmarks from the benchmarks R1 and R2. A typical high performance design has tens of thousands of gates and close to a hundred power supply pads. The number of sinks per pad is therefore a few hundred. We therefore choose each of R1 and R2 to have four supply pads at the four corners of the die. The assignments of pins to pads is done by a closest point heuristic, i.e., a pin is assigned to its closest pad. The total current is chosen to be 1A. The distribution of sink currents I_i is chosen to be proportional to the corresponding load capacitances C_{Li}. The die size for R1 and R2 are chosen to be 7.5mmX7.5mm and 1cmX1cm respectively. The number of sinks in each of our eight benchmarks is shown in Table 1. The suffix of each name indicates which corner the pad is at. E.g., R1.LL represents the benchmark obtained from R1 with sinks closest to the lower left corner.

Table 1: Benchmark sizes

Name	Number of sinks	Sum of sink currents (A)
R1.LL	72	0.274
R1.LR	76	0.282
R1.UL	87	0.325
R1.UR	32	0.119
R2.LL	116	0.193
R2.LR	160	0.277
R2.UL	142	0.237
R2.UR	180	0.294

In our experiments we compare sized star routes with our algorithm. The results are shown in Table 2 and 3. The resistance per unit grid is 1mΩ and the maximum current through a minimum width wire is 1mA. The vertical voltage drop constraint is 0.2 V. All lengths are in terms of 10^6 grid units(0.1 μm). Areas are in units of 10^6 square grids (0.1μm X 3μm). We compare the areas obtained by the greedy algorithm for three different kinds of compatibility graphs - one with no edges, second with a complete graph and third with a

graph with O(n) edges. The compatibility graph with no edges corresponds to the conventional methods. The complete graph corresponds to a low parallelism situation and gives us a bound on the savings possible. The current after the merge of sub-trees is now the largest sub-tree current. The graph with O(n) edges has much weaker time-domain information. We get these results by setting the current after the merge of two sink nodes to be the larger of the sink currents. The current compatibility graph therefore has at most n/2 edges. The area in the complete graph case is typically

Table 2: Utility of temporal information

Bench mark	Empty edge set	Complete graph	Graph with O(n) edges		Star route
	Area	Area	Area	# edges	Area
R1.LL	4.280	0.743	3.128	27	5.772
R1.LR	3.784	0.695	2.832	27	5.716
R1.UL	4.208	0.735	2.982	33	6.243
R1.UR	1.631	0.423	1.197	11	2.481
R2.LL	8.298	2.093	5.860	44	7.928
R2.LR	9.242	2.466	7.372	61	10.19
R2.UL	7.371	2.489	5.541	56	8.061
R2.UR	10.41	2.619	7.812	71	11.63

five times less than the empty edge set case. Even with O(n) edges we get significantly smaller areas - about 30% smaller.

The error caused by using real domain solutions when sizing wires is bounded by the net length. This is typically only about 10% of the total area for our heuristic, as Table 3 shows, so that our experimental framework is sound.

TABLE 3. Net lengths

Benchmark	Star route	Greedy
R1.LL	3.371	0.455
R1.LR	3.483	0.436
R1.UL	3.809	0.447
R1.UR	1.601	0.242
R2.LL	7.739	2.075
R2.LR	10.01	2.367
R2.UL	7.871	2.259
R2.UR	11.35	2.460

It is easily possible for the current compatibility graph to have $O(n^2)$ edges. Typically, very few nodes in a Boolean network switch during a clock period. A Boolean network in which levels give the order of circuit activity will have $O(n^2)$ edges. Similarly, in a single write port register file, only one of the registers is written into at any given time and we have $O(n^2)$ edges. During a DRAM refresh, a row is refreshed at a time. This means that different rows draw current at different times, giving $O(n^2)$ edges again.

V. Conclusions

We have seen that the current compatibility graph is a useful model for capturing time-domain current information in power net area optimization. We have seen that our algorithm is a viable approach for simultaneous topology design and wire sizing. Our formulation does not sacrifice feasible regions of the design space. Extensions for use in standard cell and channel-based layout styles are straightforward - merges are defined over the routing graph.

Future work will include characterization of optimum solutions, better methods for representing temporal information and approximation algorithms with performance guarantees.

References

[1] K.D. Boese, A.B. Kahng, B.A. McCoy and G. Robins, “Rectilinear Steiner trees with minimum Elmore delay”, Proceedings of the Design Automation Conference, 1994, pp. 381-386.

[2] S. Chowdhury and J.S. Barkatullah, “Estimation of maximum currents in MOS IC logic circuits”, IEEE Transactions on CAD, Vol. 9, No. 6, June 1990, pp. 642-654.

[3] S. Chowdhury and M.A. Breuer, “Optimum design of IC power/ground nets subject to reliability constraints”, IEEE Transactions on CAD, Vol. 7, No.7, July 1988, pp. 787-796.

[4] J. Cong, K.-S. Leung and D. Zhou, “Performance-driven interconnect design based on a distributed RC delay model”, Proceedings of the Design Automation Conference, 1993, pp. 606-611.

[5] L. Conway and C. Mead, *Introduction to VLSI Systems*, Addison-Wesley, 1980, Chapter7.

[6] R. Dutta and M. Marek-Sadowska, “Algorithm for wire sizing of power and ground networks in VLSI designs”, Journal of Circuits, Systems and Computers, Vol. 2, No. 2, 1992, pp. 141-157

[7] K.-H Erhard and F.M. Johannes, “Power/ground networks in VLSI: are general graphs better than trees?”, Integration, the VLSI journal, Vol. 14, 1992, pp. 91-109.

[8] K.-H. Erhard, F.M. Johannes and R. Dachauer, “Topology optimization techniques for power/ground networks in VLSI”, Proceeding of the European Design Automation Conference, 1992, pp. 362-367.

[9] T.D. Hodes, B.A. McCoy and G. Robins, “Dynamically-wiresized Elmore-based routing constructions”, Proceedings of the International Symposium on Circuits and Systems, 1994, pp.463-466.

[10] R. Kolla, “A dynamic programming approach to the power supply net sizing problem”, Proceedings of the European Design Automation Conference, March 1990, pp. 600-604.

[11] H. Kriplani, F. Najm and I. Hajj, “Maximum current estimation in CMOS circuits”, Proceedings of the Design Automation Conference, 1992, pp. 2-7.

[12] T. Mitsuhashi and E.S. Kuh, “Power and ground network topology optimization for cell-based VLSIs”, Proceedings of the Design Automation Conference, 1992, pp. 524-529.

[13] A.S. Moulton, “Laying the power and ground wires on a VLSI chip”, Proceedings of the Design Automation Conference, 1983, pp. 754-755.

[14] W.S. Song and L.A. Glasser, “Power distribution techniques for VLSI circuits”, IEEE Journal of Solid State Circuits, Vol. 21, No. 1, February 1986, pp. 150-156.

[15] B.R. Stanisic, “Automatic analog power distribution synthesis in RAIL”, Ph.D. thesis, Carnegie Mellon University, 1993.

[16] Z.A. Syed and A. El Gamal, “Single layer routing of power and ground networks in integrated circuits”, Journal of Digital Systems, Vol. 6, No. 1, 1982, pp. 53-63.

[17] R.-S. Tsay, An exact zero-skew clock routing algorithm”, IEEE Transactions on CAD, Vol. 12, 1993, pp. 242-249.

[18] P. Vanoostende, P. Six and H.J. de Man, “PRITI: estimation of maximal currents and current derivatives in complex CMOS circuits using activity waveforms”, Proceedings of the European Design Automation Conference, 1993, pp. 347-353.