# Accelerating Large-scale Network Capacity Planning via Set-cover-enhanced Benders Decomposition

Eduardo Moreno[1], Babak Moazzez[2], Bruno de Backer[1], and Hugues Evrard[1]

[1]Google, Paris, France ✉ edomoreno@google.com
[2]Google, Sunnyvale, USA

Modern network capacity planning for large-scale infrastructures, such as Google's network, relies on complex Mixed-Integer Programming (MIP) models. These models determine cost-effective capacity allocations while guaranteeing Quality of Service (QoS). However, existing methods often struggle with long solving times due to the sheer size and complexity of the problem.

Consider a network capacity planning optimization model defined as:

$$\min_{cap, flow, x} \quad \text{Cost}(cap, x) \tag{1}$$

$$\text{s.t.} \quad g(cap, x) \leq 0 \tag{2}$$

$$\sum_{p \in \text{Paths}(d)} flow_p = D_d \qquad \forall d \in \text{Demands} \tag{3}$$

$$\sum_{d \in \text{Demands}} \sum_{p \in \text{Paths}(d): l \in p} flow_p \leq cap_l \qquad \forall l \in \text{Links} \tag{4}$$

Here, $cap_l$ represents the capacity of link $l$, and $x$ represents other decision variables required to set the optimal capacity, subject to arbitrary constraints represented by (2). We assume that all demand flow $D_d$ must be routed through allowed paths $p \in \text{Paths}(d)$ (constraints (3)), ensuring that the total flow over link $l$ does not exceed its capacity $cap_l$ (constraints (4)).

When the network size and the number of demand flows increase, this problem becomes computationally intractable. To address this, we propose a Benders decomposition framework. The core innovation is the dynamic generation of Benders cuts to replace the flow-conservation constraints on the main problem. We replace constraints (3) and (4) with Benders cuts constructed from the dual of a subproblem that minimizes unrouted flows, subject to the capacities provided by the main problem, resulting in:

$$\max_{\alpha, \beta} \quad \sum_{d \in \text{Demands}} \alpha_d - \sum_{l \in \text{Links}} \beta_l$$

$$\text{s.t.} \quad \alpha_d \leq \sum_{l \in p} \beta_l \quad \forall p \in \text{Paths}(d),\ d \in \text{Demands}$$

This subproblem can be viewed as a fractional weighted set-covering problem. By exploiting this structure, we apply specific set-cover heuristics to quickly generate violated Benders cuts for the main problem. This avoids the need for flow-dependent variables and significantly reduces complexity. Furthermore, we employ a novel hierarchical clustering technique to generate multiple violated cuts from a single solution of the set-cover problem, which dramatically improves decomposition efficiency. Finally, this approach allows us to solve multiple failure scenarios in parallel and incorporate specific SLO-aware constraints for demand flows.

Our preliminary results, using real data from Google's network infrastructure, validate that this approach significantly reduces computational time compared to existing methods, enabling rapid, high-quality network design decisions.