With the advent of adaptive conformal radiotherapy, the development of very efficient methods of intensity modulated radiation therapy (IMRT) fluence map optimization (FMO) is required. The adaptive component of this approach requires the ability to generate high quality treatment plans "on-the-fly" to correct for changes in patient anatomy during the course of radiotherapy. We present a very efficient method of IMRT FMO that combines the well studied least-squares objective function with convex Conditional Value-at-Risk (CVaR) dose-volume histogram constraints. The resulting quadratic programming model is solved to global optimality using a commercial implementation of the primal-dual interior-point algorithm. We have applied this technique to 7 head-and-neck cases were definitive therapy and saliva gland sparing were desired. We also employed a multi-resolution grid-spacing model where decreasing resolutions were applied to targets, structures, and unspecified tissue without deterioration of the plan quality. Typical model sizes employed 16,000-28,000 variables and 22,000-39,000 constraints. Solution times varied between 6.5 and 15.5 seconds on a single 1.7 GHz Centrino processor with 1 GB of RAM. Dose distributions based on pre-computed beamlet doses could were recomputed in another 1-2 seconds for immediate review. The qualities of the solutions obtained were found to be comparable to a previously reported and potentially more flexible piece-wise linear-programming based model with CVaR constraint. The primal-dual interior-point algorithm employed can be parallelized to increase performance on a multiprocessor platform.