Chemistry CQoS:
Automated Tuning of SCF Procedures

Joseph P. Kenny

August 25th, 2007

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company,
for the United States Department of Energy's National Nuclear Security Administration
under contract DE-AC04-94AL85000.

CCA
Common Component Architecture

NNSA
National Nuclear Security Administration

Sandia National Laboratories

Goals for this phase

- Integrate chem cqos efforts with core efforts
- Design prototype application encompassing important features but which is simple enough to demonstrate near-term progress

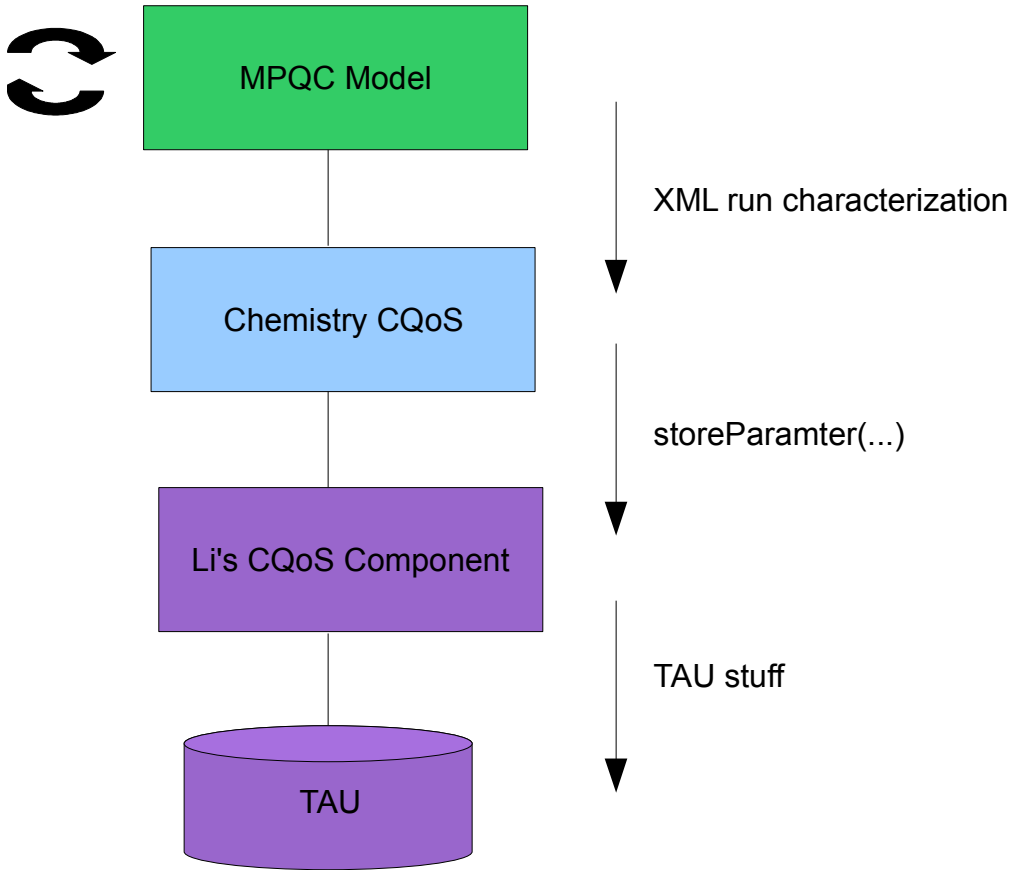
Scope

- Global database intriguing, stick with per machine for now
  - Seed database with large number of performance runs, grow with future runs

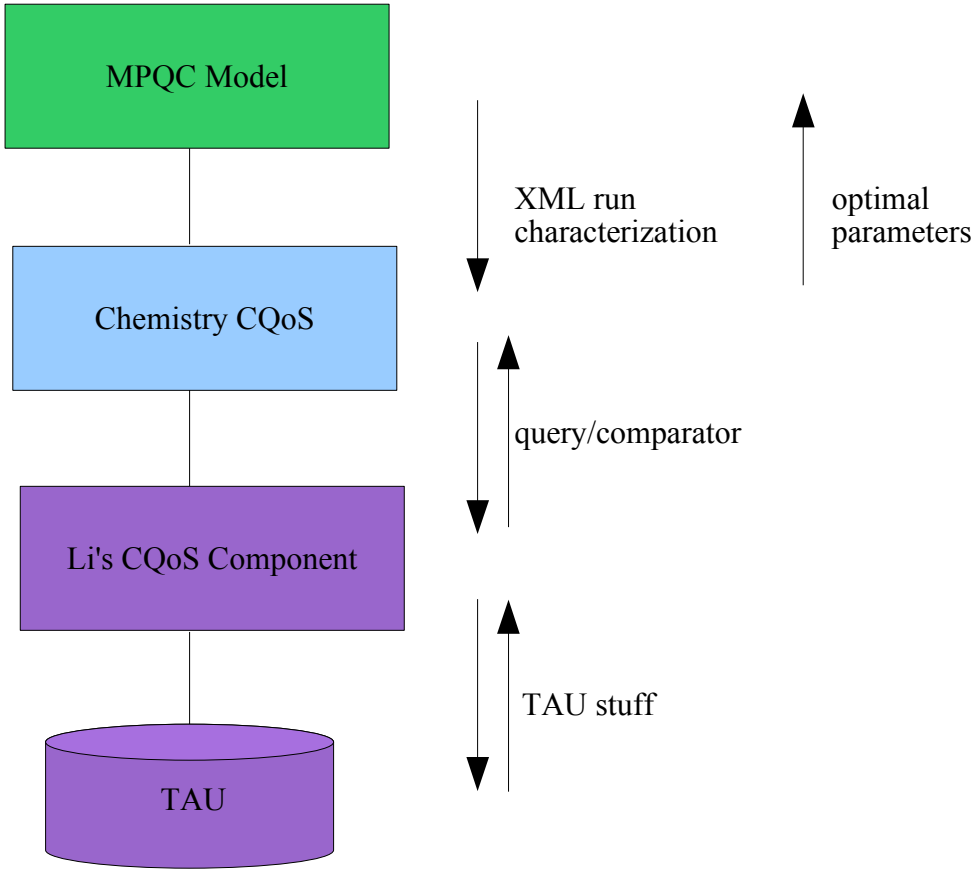
Target: SCF tuning

- Many parameters to tune, down-select for now
  - Number of nodes, multiple MPI tasks or threads
  - Use integral storage (disk-based for GAMESS, memory-resident integral cache for MPQC)

Chemistry CQoS Component: Training Phase



Chemistry CQoS Component: Production Phase



Tuning Algorithm

- Parse input to determine molecule characterization
  - Obtain runs with similar molecule characterization from database
    - same spin multiplicity
    - Choose percentage (10%?) of runs with least difference in atom type numbers (average difference, max difference)?
- Determine best performing (10%?) calcs with similar molecules based on execution time
- How many nodes?
  - Use average nodes for best calcs

Tuning Algorithm

- Use integral cache?
  - Best calcs each get a vote
- If using integral cache, how much memory
  - Average of best calcs (that used cache)

Database Entities

SCF Configuration

- Use integral cache/disk storage {on,off}

Hardware Configuration

- Number of nodes {*integer*}
- MPI tasks per node {*integer*}
- Threads per node {*integer*}

Database Entities

Molecule Characterization

- Number of H,C,N,O {*integer*}
- Numbers of other atom types (first row, third row transition, etc) {*integer*}
- Spin multiplicity {*integer*}

Performance Metric

- Execution time {*integer*}