

Google Performance Tools

Brad Chen

Cluster Performance Profiling Group

bradchen at google dot com

Background

- A big developer community: about 6000 developers
 - very bottom up, organic growth
 - most performance tools incidental to a larger project
- "Request-level parallelism" the predominant || model
- Diversity and sharing
 - Web search: crawl, index, and serve; ads, translation,...
 - External services: Gmail, Orkut, Maps, YouTube, Apps...
 - Internal infrastructure: Chubby, BigTable, MapReduce, ...
 - More dynamic than other systems I know about
- "Land-grab" mentality

Some Examples

- pprof, contentionz
- rgm, borgmon
- application structure
- Guppi

Some problems of interest

- I/O measurement and analysis
 - more generally, focus on tools around a bottleneck
- Coordination of resource sharing in a cluster
 - spotting sharing issues, bigger systems
 - utilization as an opportunity
- Using fine-grained parallelism to improve service latency
 - not a consensus opinion
- Accessibility of microarchitectural analysis and tuning
- Reducing diversity of performance data formats