Abstract: The progress of chip design technology faces two related challenges: power and on-chip communication. A recent study by Google shows that, as power-efficiency improves for server processors, the interconnection network is becoming a major power consumer in the datacenter. Likewise, on-chip communication now forms a power bottleneck in chip multiprocessors (CMPs) given the considerable progress on processor core power-efficiency. This talk focuses on employing dynamic voltage and frequency scaling (DVFS) and power gating policies for networks-on-chip (NoC) and shared, distributed last-level caches (LLC). In particular, I will present a practical system architecture where the distributed LLC and the NoC share a voltage/frequency domain which is separate from the core domain. This architecture enables controlling the relative speed between the cores and memory hierarchy without introducing synchronization delays within the NoC. DVFS for this architecture is more difficult than individual link/core-based DVFS since it involves spatially distributed monitoring and control. We propose an average memory access time (AMAT)-based monitoring technique and integrate it with DVFS based on PID control theory. Simulations on PARSEC benchmarks yield a 33% dynamic energy savings with a negligible impact on system performance.

I will also present preliminary results for a power gating policy which uses sampling to determine the active LLC footprint at run time, an power gates off portions of the cache when they do not contribute to performance. This technique achieves a power savings of 60% on workloads from the SPEC CPU2006 benchmark suite.

Biography: Paul V. Gratz is an Assistant Professor in the department of Electrical and Computer Engineering at Texas A&M University. His research interests include energy efficient and reliable design in the context of high performance computer architecture, processor memory systems and on-chip interconnection networks. He received his B.S. and M.S. degrees in Electrical Engineering from The University of Florida in 1994 and 1997 respectively. From 1997 to 2002 he was a design engineer with Intel Corporation. He received his Ph.D. degree in Electrical and Computer Engineering from the University of Texas at Austin in 2008. His paper "B-Fetch:Branch Prediction Directed Prefetching for In-Order Processors" was selected as one of four "Best Papers from IEEE Computer Architecture Letters in 2011". At ASPLOS '09, Dr. Gratz co-authored "An Evaluation of the TRIPS Computer System," receiving a best paper award. In Spring 2010, he received the "Teaching Excellence Award - Top 5%" award from the Texas A&M University System for a graduate-level Advanced Computer Architecture course he developed.