A synopsis of the paper:

Autonomic Energy Management in a Replicated Server System

Aeiman Gadafi, Laurent Broto, Noel Depalma
2010 Sixth International Conference on Autonomic and Autonomous Systems
Talk Agenda

• Introduction

• Context and motivation

• Autonomic energy management approach

• Experimentation and evaluation

• Contributions and conclusions
Introduction

- Resource management major challenge for server firms - most of time unused

- Autonomic management systems to dynamically adapt to runtime conditions - provide support for deployment, configuration etc.

- Reports an experiment for resource conservation in a clustered systems

- Dynamically adapts degree of replication to minimize energy consumption
• Emergence of ‘Cloud Computing’ brought the energy issue to forefront

• Energy consumption of server firms has doubled between 2000 and 2006

• QoS guarantee - naïve implementations over provision servers for peak periods

• Paper presents a way to dynamically allocate/de-allocate servers

• RAM suspension instead of machine on/off
Context and Motivation

• Web, Application and Database servers
• Increasing the number of users require more servers
• Load balancing among replicas – round-robin, random method
• How much replication is needed – tough problem to solve
• Paper talks about replicated database servers
• Main goal is to minimize energy consumption – maintaining response time
Autonomic energy management approach

- Autonomously increase/decrease replicated resource
- Decision based on maximum and minimum CPU utilization
- For more resource
  - Start a new node and a new server replica
  - Integrate the new replica into the load balancer
- For less resource
  - Remove a replica and stop it
  - Turn off the node
- Prevent oscillation – watch for stable condition before changing

Figure 3. Control loop

Figure 4. Autonomic replication of database servers
Autonomic energy management approach

- Performance scalability by adding database nodes and balancing the load among these nodes

- High availability of the database tier, i.e. C-JDBC tolerates database crashes and offers transparent failover using database replication techniques
Experimentation and evaluation

- 7 Machines running Linux and a programmable power meter
- Usual assortments of software products
- Tune management system and RUBiS - auction system from Rice University
- 1000 emulated clients - 45 minutes run
Experimentation and evaluation

- Requests processed with autonomic system is the same as it was for constant 3 servers scenarios
- Twice machines were added and twice they were RAM suspended
- Energy consumed was 21% less in autonomic system compared to the 3 server scenario
- Tune is providing best trade-off between QoS and energy consumption

Figure 8. Dynamic configuration with Tune

Figure 9. Comparison
Contributions and Conclusions

• Experiments show that autonomic approach can be used to do energy aware computing
• Reduced power consumption by 21%
• Promises to model network traffic
• Will integrate virtualization techniques for transparent process migration

• Paper does not delve into the details of Tune management system
Tag cloud for my assigned papers for the semester

ae (101) application (53) approach (55) architecture (141)
autonomic (62) availability (57) case (44) change (109) claim (46) class (73)
component (71) computing (55) controller (148) dynamic (40)
engineering (55) environment (54) example (62) execution (50) figure (64)
framework (53) function (74) global (42) goals (54) layer (59) level (51)
management (79) model (55) number (67) paper (60) patterns (69)
policy (53) problem (39) process (52) QOS (93) requests (51) response (55) sassy (42)
section (53) security (81) self-managed (44) servers (134) service (114)
software (118) state (44) system (380) used (59) utility (97)
va (52) values (66) work (56)