Scalable and Coordinated Cross Layer Management

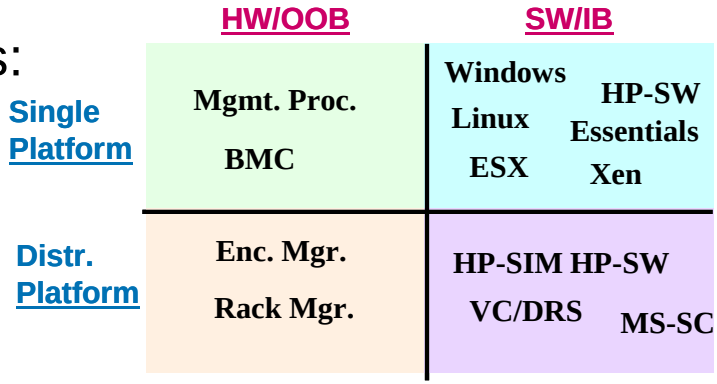
Vanish Talwar, Partha Ranganathan – HP Labs
Sanjay Kumar, Karsten Schwan – Georgia Tech

Hot Topics in Autonomic Computing

June 2, 2008

Context: Future Exascale Systems

- Key trends shaping future exascale technologies
  - move towards many-core chips
  - increased demand for even larger and reactive data centers
  - virtualization is inevitable in future systems
- Limitations applying current management solutions
  - partial, silo-ed subsystem level solns: leads to ineffectiveness & cost
  - centralized nature: limits scalability
  - virt. & many-cores stretch limitations



CHALLENGE: scalable and coordinated cross layer management for exascale systems

Research Required by the Challenge

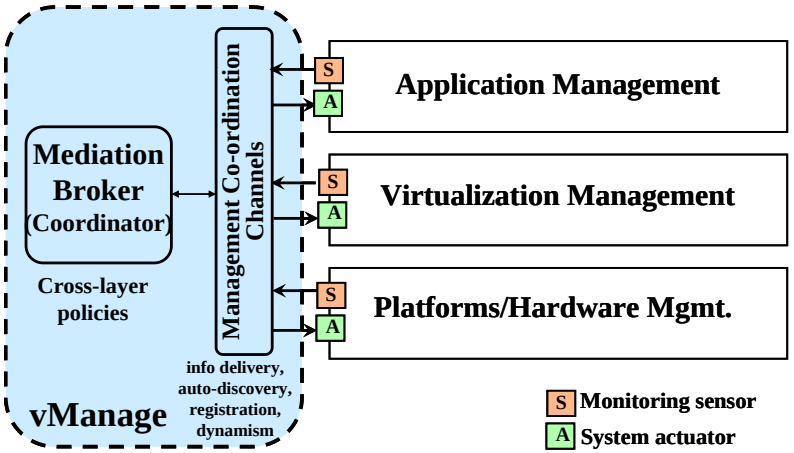
- Goals: low-cost, scalable performance and availability, reduced complexity, light weight, extensible
- Mgmt. coordination architecture and mechanisms
  - cross layer automated infrastructure
  - distributed virtual appliances overlay
  - scalable discovery & routing, directory services

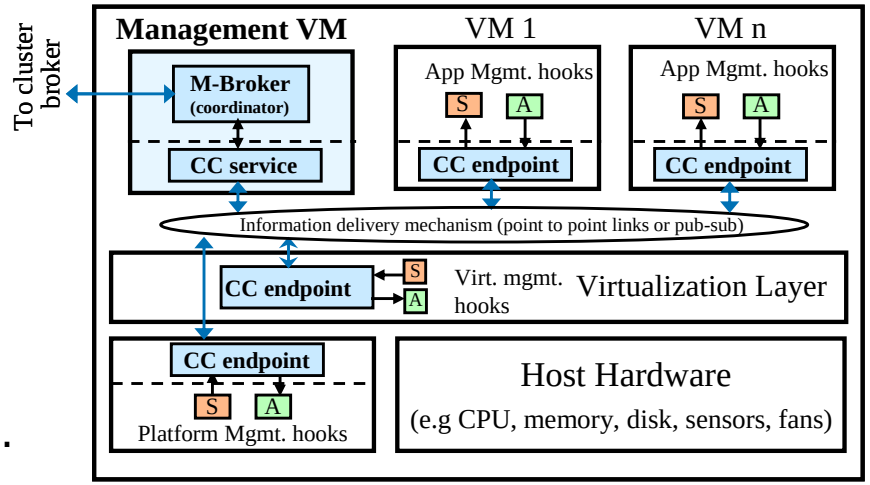
Manual	High latency, cost
Ad-hoc point solns	Non-scalable/portable, complex, costly
Automated	Structured, scalable, initial dev. costs

- Mgmt. principles & policies for improved efficiency
  - dynamic coordination/mgmt. assessment (for complex systems)
  - improved power-perf mgmt, fault mgmt., distr. aggregation
  - adaptive and predictive monitoring

Example: vManage

- VMM-level co-ord. arch. & approach
- Mgmt coord. channels (CC)
  - provide seamless discovery, meta-data unification, info delivery, dynamism across hw-sw-virt
- Mediation brokers (Coordinator)
  - execute in mgmt. VMs; interface to co-ord channels; implement policies
- Case Study: Coordinated VM placement & dynamic provisioning
  - considers VM and host metrics & actuators with stability analysis
  - example policy: SLA-based pwr reg. coordinated with pwr capping

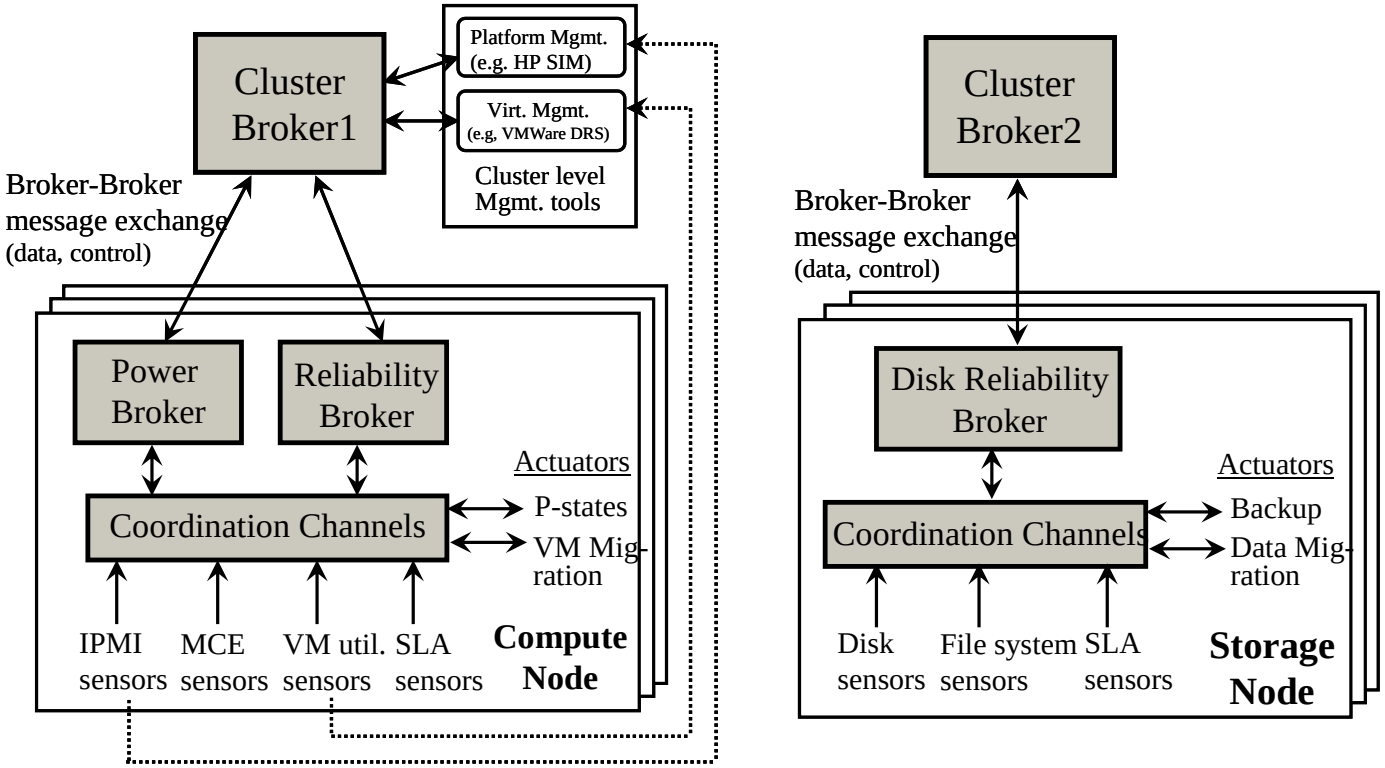




Backup

Case Study example

- Coordinated VM placement and dynamic provisioning



- Improved data center efficiency (power savings) and VM guarantees (SLA violations, availability) with automated infrastructure