IBM Research

Dependability and Security Challenges in Tomorrow's Data Centers

Hari V. Ramasamy

IBM T.J. Watson Research Center, Hawthorne, NY

IFIP 10.4 WG on Dependable Computing and Fault Tolerance

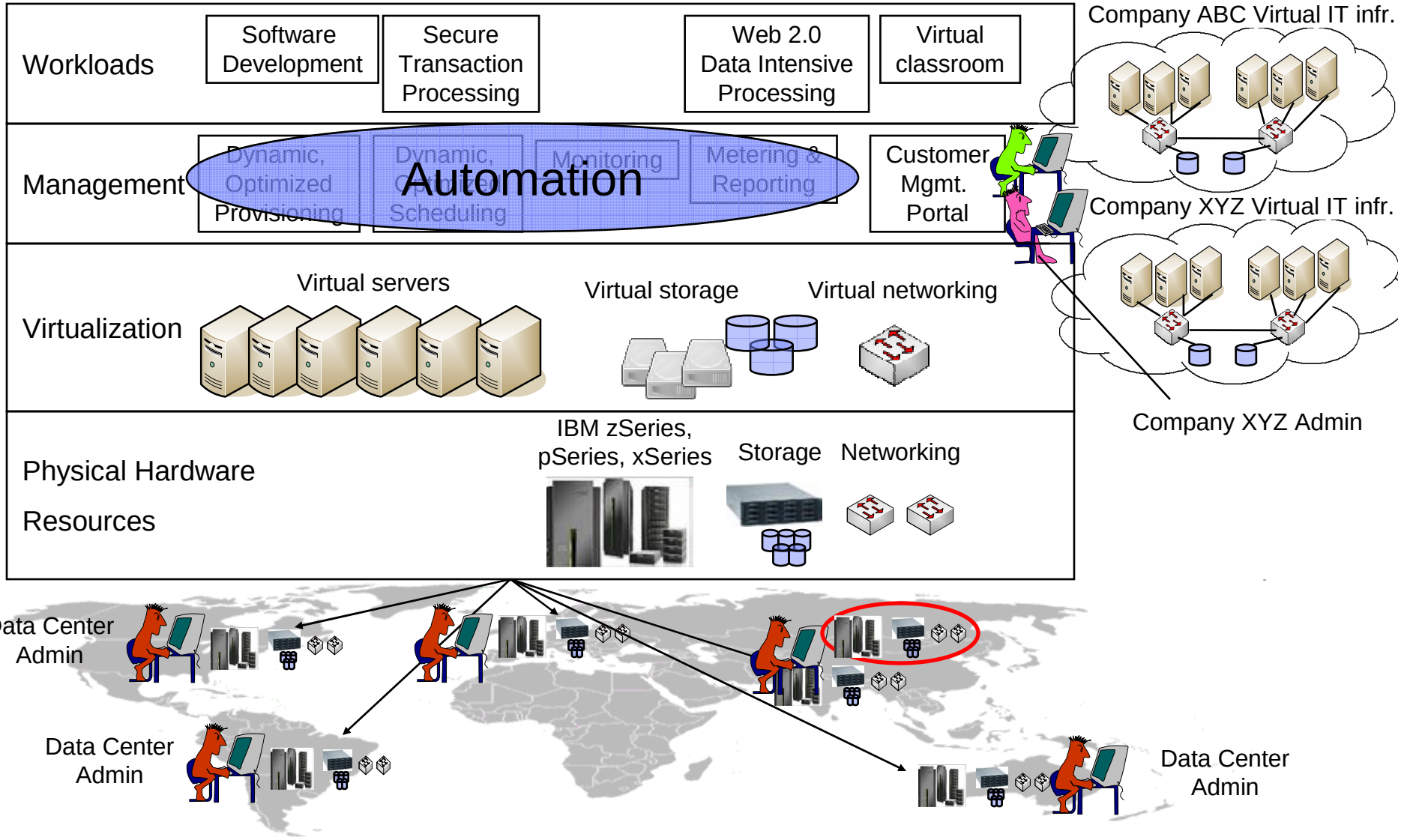
Today's Data Centers

- **Most Fortune 500 companies have their own large, dedicated data centers**
- **Smaller companies are increasingly outsourcing their IT infr., but still "physical cages" model at the data center provider**
- **Over-provisioning, over-engineering, under-utilization rampant**

What are the trends coming together? [Clabby Analytics]

- **At the operational level, problems with status quo have reached tipping point**
  - facilities cost, labor shortage, management costs, complexity that is beyond human capability
- **In terms of business needs, focus of IT management is moving away from a purely systems focus to a focus on satisfying business goals and service delivery**
- **In terms of technological trends,**
  - Consolidation
    - 30-50% enterprises have/are consolidating; a smaller % are doing some level of virtualization today [Source: IBM 2008]
  - Web-centric cloud computing ≈ Grid + virtualize everything + Web 2.0
    - Evolution from Grid computing, utility computing, Software-as-a-Service
    - *"By 2012, 80 percent of Fortune 1000 enterprises will pay for some cloud computing service and 30 percent of them will pay for cloud computing infrastructure."* – Gartner
  - SOA and XML standards

What might tomorrow's data center look like? [IBM NEDC White Paper]



What are the security challenges that might impede the move?

- **Securing the data in the cloud is a top concern**
  - data integrity
  - data confidentiality
  - customer isolation: familiar and intuitive notions of physical isolation need to be mirrored in the shared environment
- **Compliance Verification**
  - How to ensure that VM images all have necessary security updates?
  - If the data is moved from one country to another in the cloud, how to guarantee that the data respects the country's privacy laws?
- **Multiple levels of P*rovable Assurance***
- **Identity Management and Access Control**

Caveat: How do we address all these, while retaining the simplicity, efficiency, and usability that are the main driving factors behind cloud computing (read Amazon's EC2, Salesforce.com) taking off?

What are the dependability challenges that might impede the move?

- **Resilience at all levels**
  - facilities, infrastructure, applications, data
- **Failure Isolation**
  - between applications, virtual zones, customers sharing infrastructure
- **Continuous data availability and data preservation**
  - despite application, VM instance, or even data center failures
- **Better monitoring and fault tolerance at the virtualization layer level**
- **Ensuring that data among multiple sites is consistent**
  - we are talking really large volumes of data here (VM images, etc.), where network – despite improving bandwidths – could be a bottleneck
- **Really automated, tunable, real-time recovery solutions**
  - workload migration optimizing for cost (incl. energy-related) while satisfying performance, availability, compliance requirements simultaneously