Disaster Prevention and Recovery Architecture

A Presentation for Disaster Recovery Planning Professionals

Steven J. Ross
Executive Principal
Risk Masters, Inc.
New York, NY
Abstract

Many Disaster Recovery plans were written with the assumption that all that was required was an alternate facility in which to run production in an emergency. Businesses now have less tolerance for downtime and data loss and a combination of technology and economics has made prevention of downtime just as realistic a goal for many organizations as is recovery. Server virtualization, cloud computing, data center consolidation, SAN-based replication, virtual tape libraries, ultra-high bandwidth and distributed staffing models have made it possible to design enterprise data centers with continuous operations as a realistic goal. Putting all the pieces together calls for more than a planner; you need an architect. This session presents techniques for the development of an IT Disaster Prevention and Recovery Architecture that is driven by the business needs for continuity in data centers as well as the offices and factories that depend on them. It distinguishes HA from DR, showing how closely aligned the two concepts are. It will enable attendees to begin the lengthy road to true protection of the enterprise’s IT assets.
The Drivers for Disaster Recovery Architecture
It Ain’t Your Daddy’s Disaster Recovery Plan

• Most Disaster Recovery Plans (DRP) were developed to recover the operation of a single site
  – Alternate data center, often at a commercial recovery service
  – Recovery of the infrastructure, applications and data maintained in that data center
  – Often recovery of data and software from backup tapes
  – Recovery time measured in hours to days
• Each data center served the business needs of either
  – Functions in geographic proximity
  – Widespread locations in the same business unit
• Rarely did DRPs address the needs of all the data centers taken as a whole, serving the entire enterprise
Consolidation Forces DRP Change

• As the decision to consolidate data centers is made and implemented, it becomes apparent that Disaster Recovery Planning needs to be re-thought
  – Applications and data spread across locations
  – Resulting concentration of key applications (e.g., ERP, CRM)
    • Or costly duplication of application instances
  – Inability to back up and restore immense amounts of data
  – Vastly increased risk from the failure of a single data center
• How many data centers is too many?
• How few data centers is too few?
Key Indicators of the Need for a Disaster Recovery Architecture

- Disaster Recovery infrastructure and practices
- Critical enterprise-wide applications
- The number of data centers
- Backbone routing
- Dependency on the Extended Enterprise
In order to recover the applications supporting the business, it is first necessary to recover the infrastructure:

- The physical data center
- The network
- System software
- Storage

Only then can applications and external connections be recovered.
Disaster Recovery Infrastructure and Practices, continued

• With an alternate data center, it is necessary to port infrastructure, network, applications and data
  – Usually less capacity
Disaster Recovery Infrastructure and Practices, continued

- This structure breaks down when applications and infrastructure are spread over multiple data centers
  - What goes to the alternate site and what is retained locally?
- How do shared applications operate when one site is down?
Critical Enterprise-wide Applications

• In some cases, local data centers serve only local needs
  – In these cases, localized Disaster Recovery is sufficient
• Increasingly we see common applications serving broad-based geographies and user populations
  – Email (Exchange)
  – ERP (SAP, Oracle)
  – Customer Relationship Management (Seibel)
  – Procurement (Ariba)
  – Warehouse Management (often industry specialized)
  – Human Resources (if not in an ERP)
• These may cross data centers and cannot be recovered easily on a site basis
The Number of Data Centers

• You know you have too many data centers if...
  – Support and control staffs are duplicated at each site
  – Overall server utilization is less than 10%
  – Data is replicated from site to site
  – Total cost of ownership (TCO) of information resources is growing rapidly
    • Software licenses
    • Operating systems
    • Real estate
  – High cost to replicate security from site to site
  – General inefficiencies of scale
Backbone Routing

• In some cases, the intra-company backbone network runs through one site, usually the major acquirer's original data center
  – This makes the core data center a giant single point of failure

• Reducing the risk requires major network re-architecture, including attention to diversity and redundancy (i.e., Disaster Recovery)
  – Network protocols (MPLS)
  – Carrier diversity or guarantees of route diversity
  – Bandwidth on demand
Backbone Routing, continued

• In other cases, companies have a diverse backbone network

• This reduces the dependence on one site, but also creates a condition in which regional clusters add cost for network routing but does not accomplish savings by increasing efficiencies of scale

• The risk of a disaster at a concentrator site, while not as great as a star-shaped backbone, is still significant
Dependency on the Extended Enterprise

• The problems of recovering multiple data centers are exacerbated when one or more of those data centers are owned and operated by third parties
  – Joint ventures/coopetition
  – ASPs
  – Data sources
• The more critical these third party relationships are to the enterprise, the more difficult it is to manage recoverability on a site-by-site basis
  – Overlapping and contradictory contracts
  – Possible prohibitions on relocating services or connectivity
• Consideration must be given to a disruption caused by an outage at another company’s data center
Risk Masters, Inc.

Disaster Recovery and High Availability
A Subtle Difference...But a Real One

• Disaster Recovery focuses on the *restoration* of services after physical destruction of facilities, equipment and/or data.
  – Policies, processes and plans
  – Accomplished remotely from the primary site
  – The timeframe for doing so is dictated by the needs of the business

• High Availability deals with *ensuring a prearranged level of operational performance*
  – Redundancy and sound operational methods
  – Applies to all level applications and infrastructure
  – Aimed at preventing downtime and ensuring data center availability (i.e. in response to a component failure)
  – Both within and among data centers

• Highly available data centers are also recoverable
# Availability Tiers (Uptime Institute)

<table>
<thead>
<tr>
<th></th>
<th>Tier 1</th>
<th>Tier 2</th>
<th>Tier 3</th>
<th>Tier 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Active Capacity</strong></td>
<td></td>
<td></td>
<td></td>
<td>N after any failure (or 2N)</td>
</tr>
<tr>
<td><strong>Components to Support the IT Load</strong></td>
<td>N</td>
<td>N+1</td>
<td>N+1</td>
<td></td>
</tr>
<tr>
<td><strong>Distribution Paths</strong></td>
<td>1</td>
<td>1</td>
<td>1 active+1 alternate</td>
<td>2 simultaneously active</td>
</tr>
<tr>
<td><strong>Concurrently Maintainable</strong></td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Fault Tolerance</strong></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Compartmentalization</strong></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Continuous Cooling</strong></td>
<td>Load Density Dependent</td>
<td>Load Density Dependent</td>
<td>Load Density Dependent</td>
<td>Class A</td>
</tr>
</tbody>
</table>
Different Drivers

• Disaster Recovery is driven by the needs of the business for information systems
  – Business Impact Analysis

• High Availability is driven by the degree of risk minimization that management is prepared to pay for
  – Risk Analysis

• If some downtime is acceptable, is high availability necessary?
  – Especially if there is adequate recoverability
  – As usual, cost is a critical consideration
Disaster Recovery Architecture Methodology
Disaster Recovery Architecture Model

- **Document Current State**
  - Primary Data Centers
  - Network

- **Consider Business Drivers**
  - Corporate
  - Market
  - Recovery Time Objectives
  - Recovery Point Objectives

**Data Center Implications and Considerations**

- Technology Trends
- Growth

**Infrastructure Strategy Analysis**

- Location Analysis
- Data Center Strategy Options
- Sourcing & Vendor Analysis
- Cost Analysis

**Filtering Process**

**Final Strategy**

- Strategy Optimization
- Cost Optimization
- Present to Executive Management
- Management Determination of Acceptable Risk Level
- Solution Recommendations
- Transition Guidelines

**Implementation and Testing**

- Implementation Work Plans
- Document Procedures
- Proof of Concept Testing
Assessing the Current State of Resilience

• The ideal
  – Recoverability and resilience are considerations in consolidation, placement and construction decisions
  – All data centers are highly available in accord with management’s perception of risk
  – A consolidated Disaster Recovery Plan is developed in parallel with the consolidation of data centers, created by
    • Business Continuity Management
    • Disaster Recovery Planning
    • Technical Engineering
    • Network Engineering
    • Operations
Assessing the Current State of Resilience, continued

• The reality
  – Recoverability and resilience are *afterthoughts* in equipment acquisition, consolidation, placement and construction decisions
  – A consolidated Disaster Recovery Plan is developed *after the fact* of the consolidation of data centers, created by
    • Disaster Recovery Planning, with the aid of
      – Consultants
      – Auditors
      – Anyone who will listen and help out
  – Both recoverability and availability are constrained by budgets, not needs
The needs of the business set the parameters for recoverability, resilience and availability. Those needs are expressed in different ways that overlap and sometimes contradict one another.
Corporate Strategy and Disaster Recovery

While business needs drive disaster recovery requirements, it is also true that achievement of corporate strategic goals is supported by and dependent upon information systems availability.

- **Corporate Strategy**
  - Organic growth
  - Acquisition
  - Customer service
  - Cost containment
  - New products and service
  - Shareholder value

- **Disaster Recovery**
  - Maximum uptime
  - Minimum data loss
  - Service restoration based on criticality
  - Flexibility
  - Balanced cost
I.T. Strategy and Disaster Recovery

Similarly, the strategic goals of the Information Technology function (which may include resilience and recoverability) are also dependent on Disaster Recovery.

- **Information Technology Strategy**
  - Error-free processing
  - Cost containment
  - Service orientation
  - Standardization (ITIL)

- **Disaster Recovery**
  - Maximum uptime
  - Minimum data loss
  - Service restoration based on criticality
  - Flexibility
  - Balanced cost
Technology Trends Affecting Disaster Recovery
Technology Trends

**IT Operations**
- ITIL framework adoption
- ITIL process improvement
- CMM service improvement
- Data center consolidation
- Service-centricity
- Asset management
- Configuration change management
- Inventory validation with maintenance contracts
- Server provisioning and coordination
- CPU, channel, memory and OS resource management
- Virtualization
- Remote operations
- Cloud computing and recoverability
- Network-based failover

**Facilities**
- Increased server density ratios
- Green IT and sustainability requirements
- Reduced raised floor area
- Increased power and HVAC requirements
- Voice and data network convergence
- Active-active data centers
- Thin storage provisioning on the desktop
- Network based failover
- Virtualized bare metal restore
- Virtual tape libraries

**Applications**
- Standardization, restacking and rightsizing
- Workload consolidation
- Storage tiering
- Hierarchical Storage Management
- Active-Active applications
- Service and data dependency mapping
- Applications as a service
- Legacy application dependencies

**Infrastructure**
- Heightened platform performance
- Infrastructure as a service
- Commodity servers
- Operating system maturation
- Network security
- Infrastructure as a service
- Storage consolidation
- Storage as a service
- Storage virtualization
- Resilient network protocols
- Cost of bandwidth
It Ain’t Your Daddy’s Data Center Anymore

• In the past decade, a number of trends have combined to make the data centers of previous times obsolete
  – Blade servers
  – Power cost
  – Business growth by acquisition
  – Voice and data network convergence
  – Virtualization
Blade Servers

• Blades are modular, stripped down computer systems arrayed on a common hardware backbone for power cooling, network interface, etc.

• Although each blade server draws less power and runs cooler than a rack-mounted PC, their dense concentration results in much higher demands for electricity and air conditioning
  – Since the power feed is to the entire blade enclosure, has led to the use of additional UPS
  – Beyond the need for cooling, it is important to have adequate air flow within a data center to avoid hot spots

• There is far greater need for underfloor wiring and overhead space to dissipate heat
Power Cost

• At the same time, the cost of electricity has skyrocketed, in part due to increased demand for power and HVAC

• This, more than any other factor, has been the force behind the drive towards Green IT

Source: US Energy Information Administration
Business Growth by Acquisition

• Mergers and acquisitions commonly result in the acquiring companies retention – at least for a while – of the other companies’ data centers
  – As a result, some very large organizations have data centers spread nationally and globally
  – While this diminishes the risk to the parent corporation of the failure of any one data center, the resulting costs are unsustainable
    • Bandwidth
    • Labor
• The result has been a wave of data center consolidations, made possible in part by blade servers and virtualization
Voice and Data Network Convergence

• The advent of voice over IP (VoIP) has increased the demand for bandwidth passing through the data center
• Previously, telephone service was relatively immune to power failures
  – The possibility of losing both the voice and data networks simultaneously has raised the necessity for backup power in the data center
  – Similarly, the need for multiple demarcas (entry points) for the network has been transformed to a need for multiple computing sites to provide assured service
• As a result, the proportion of network terminating equipment has increased relative to servers and storage
Virtualization

- There are many rationales for server and storage virtualization

**Servers**
- Capacity optimization
- Rapid server provisioning
- Server portability
- Reduced hardware expense
- Improved disaster recovery

**Storage**
- Increased utilization of the existing storage environment
- Improved ROI of existing data storage assets
- Reduction in downtime due to data management issues
- Improvement of backup and recovery procedures
- Improved “quality of service” offerings
- Masking of data storage management complexity

---

**“Please rate the importance of the following factors in considering or choosing to use x86 server virtualization:”**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Very important</th>
<th>Important</th>
<th>Not very important</th>
<th>Not at all important</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increasing flexibility and speed of making changes to the server environment</td>
<td>34%</td>
<td>58%</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>Consolidating servers to save on hardware costs</td>
<td>32%</td>
<td>46%</td>
<td>22%</td>
<td></td>
</tr>
<tr>
<td>Helping with disaster recovery and/or business continuity</td>
<td>26%</td>
<td>46%</td>
<td>24%</td>
<td>4%</td>
</tr>
<tr>
<td>Migrating legacy OSes like Win NT to newer, unsupported hardware</td>
<td>14%</td>
<td>26%</td>
<td>38%</td>
<td>22%</td>
</tr>
</tbody>
</table>

Base: 50 IT decision-makers at North American firms with 500 or more employees

Source: Forrester Research, Inc.
The Growth of Colocation

• After abortive growth in the early part of the century, colocation firms are well established

• The economies of scale of shared MEP with dedicated computing and networking space alter the build vs. rent decision

• Colocation offers the possibility of geographic diversity without the need for large-scale dispersed staffing
The Data Center of the Past

- Many data centers today are in the same space that housed their legacy mainframe systems
  - Some even built before the advent of CMOS computers
  - Raised floor predominated
  - Provisioning was a major event, planned years in advance
Data Center of the Future (and Now)

- There is less requirement for raised floor due to miniaturization and virtualization
- Mechanical, electrical, and plumbing equipment takes more space, both relatively and absolutely
- More room is needed for wiring, cooling, and heat dissipation
Consider Capacity and Performance

- There is a difference between recovering an application and recovering it to the same service level:
  - Is the recovery site configuration equivalent to the primary one?
  - Is the network connectivity to the recovery site the same as to the primary data center?
- Are there a sufficient number of data center personnel to recover all applications at each level of criticality within the required timeframes?
  - Operators cannot work continuously for days at a time
- Will virtualized applications work the same way on a different configuration?
  - What will be the effect of a different degree of compression?
  - Will there be the same mix of applications on recovery site servers as in the primary data center?
Disaster Recovery Architecture Selections
Alternate Locations

- Dedicated DR Site
- Shared DR Sites
- Cloud Based Recovery
Servers

Hot or Warm Servers

Repurposed Servers

At Time of Disaster (ATOD)

April 16-18, 2012 • Talking Stick Resort • Scottsdale, Arizona
Storage

• Copying
  – Tape
  – Virtual Tape

• Replication
  – Software
  – SAN
Networks – Inter-Data Center

Point to Point (DR dedicated)

MPLS (shared with users)

April 16-18, 2012 • Talking Stick Resort • Scottsdale, Arizona
Networks – Local Loops

- **Direct Connection**
  - Local Central Office
  - DR site

- **Metropolitan Area Network**
  - Local Central Office
  - DR site
Networks – Carriers

Carrier Diversity

Route Diversity

Central Office Carrier 2

Central Office Carrier 1

Carrier Diversity

Local Central Office

Route Diversity

DR site

DR site
Next Steps

• To achieve the objectives of a Disaster Prevention and Recovery Architecture, you need
  - Skill and experience
  - Fortitude and determination
  - Technology and business management support
  - Lawyers, guns and money
  - All the above
  - Did I mention money?
Feedback?
Final Thoughts

- An architect's most useful tools are an eraser at the drafting board, and a wrecking bar at the site.
  – Frank Lloyd Wright

Risk Masters, Inc.
377 W. 11th Street
Suite 1E
New York, NY 10014
(917) 837-2484
stross@riskmastersinc.com

Steven J. Ross
Executive Principal
Risk Masters™

“Experience Matters”