Service Recovery & Availability

Robert Dickerson

June 2010

- Started in 1971 with $3,000, 40 clients and 1 employee. 2009: over $2B revenue, 500,000+ clients, 13,000 employees.
- Payroll / Tax Services / 401(k) / Employee Benefits / HR Admin / Time & Labor / Employee Leasing / 1-50 Employees / 50+ Employees

Disaster Recovery & Business Continuity

- The technical solutions in place to recover IT Computer processing functions and restore service to internal and external clients. The ability to restore critical business functions and protect the corporate assets.

Paychex Assets

- 13,000 employees across 100 locations, 9 within Monroe County, 2 in Germany.
- Data centers running critical applications to support the movement of $1B per day in support of 500,000 businesses. 1,600 windows servers. 2 megawatts of power. 1,400 unix servers. 1,000 terabytes storage. 147tb backed up weekly. 75,000 inbound / 65,000 outbound phone calls each day.
- Technical infrastructure to support branch offices, data centers and connections to banking and trading partners.

Definitions

High Availability (HA):

Minimized *Unplanned* Downtime

Unplanned = Hardware failures, network outages, data corruptions, data center failures, etc.

Continuous Operations (CO):

Minimize *Planned* Downtime

Planned = Patching, application upgrades, code releases, data center maintenance, etc.

High Availability + Continuous Operations = Continuous Availability

Service Availability Percentages

Best In Class Platform Availability (per Gartner)

- High Availability – approx. 5 hours of unplanned downtime per year or 99.95%
- Continuous Operations – 12 Hours of planned downtime per Year or 99.86%
- Continuous Availability – 17 Hours of downtime per Year or 99.81%

Availability Percentage	Downtime per Year	Downtime per Month
90%	876 Hours	72 Hours
99%	87.6 Hours	7.20 Hours
99.9%	8.76 Hours	43.0 Min
99.95%	4.38 Hours	21.9 Min
99.99%	52 min	4.32 Min
99.999%	5 min 15 sec	25.9 Sec
99.9999%	31.5 sec	2.59 Sec
99.99999%	3 sec	0.25 Sec

Cost of Availability Levels (Gartner)

- The cost of continuous availability increases by multiples as the target level increases.

  If X represents the cost of 99% availability then

  99.9% has a cost equal to 3X, and

  99.999% has a cost equal to 8X

- Achieving Continuous Availability in the range of 99.999% (just minutes of downtime a year) is very rare and expensive.

Downtime Root Cause

Outages affecting Availability

Industry Breakdown

Unplanned
20% - Hardware, OS, Environmental Factors
40% - Applications
40% - Operations

Planned
65% - Application and Database Updates
35% - HW/SW/DC Maintenance, Systems and Applications Management

Paychex

Unplanned
2% - Hardware, OS, Environmental Factor
22% - Operations (network, configurations, memory, data integrity)
52% - Applications
24% - Unknown (efforts underway to minimize this classification)

Contributing factors to downtime: component failure, rate of change, expertise, process & procedures.

High Level Standards for Achieving Increased Availability

Infrastructure and Data Resiliency

Design and Implement infrastructure that prevents unplanned downtime. Use fault-tolerant hardware and operating systems. Provide redundancy to account for component failures as well as disasters. Limit complexity. Reduce MTTF (Mean Time to Failure).

Automation

Avoid human intervention for failovers as well as maintenance, test and release procedures. Use hardware load balancers or clustering software to automate failovers. Use automated release and testing tools and procedures. Reduce MTTR (Mean Time to Recovery).

IT management processes

Improve the maturity level of standard IT management processes (proactive monitoring, change, release, configuration and incident/problem). Reduce amount of change and/or probability that change will cause downtime. Monitor failures and capacity proactively vs. reactively. Understand integrations and maintain consistent configurations. Formally report and investigate incidents affecting availability. Per Gartner, IT Maturity Levels of 3 or above are required in order to be highly available.

Application Development

Design Applications with availability in mind (non-disruptive installation, no hard-coded limits, dynamic initialization, backwards compatibility) to support minimizing planned downtime (i.e. rolling upgrades. Consider availability requirements early in the design process. Once 1/3 of code has changed, Gartner recommends a complete application rewrite.

Testing

Test all possible use cases in representative test environments. Complete integrated, performance, and post-release testing.

Strategic HA and CO Technologies (Currently Available)

Load Balancing	F5 Devices for Load Balancing active/active, redundant configurations and automated failovers
Session State Management	Coherence*Web for session state management to support active/active configurations
Hardware Clustering	Veritas and Microsoft Hardware clustering software to mitigate hardware and OS failures
Database/Application Clustering	Oracle RAC for multi-node database clustering to mitigate hardware or software failures on database server nodes
Fault-Tolerant Hardware/Self-Healing OS	We use hardware models and Operating Systems that have features to prevent downtime from occurring due to hardware/OS failures.
Data Replication	Dataguard for database replication and SRDF for storage-based date replication to allow for hardware recovery on standby systems
Data Recovery	Backups to disk and tape for data recovery. Snapshot and snapback technologies.
Proactive Monitoring	RUM, Aternity and Edgesight for End User Experience Monitoring. OEM for DB monitoring. SCOM and OVO for hardware/OS/App monitoring.
Incident Tracking and Reporting	Service Center is currently in used for Problem Management.
Automation	Tidal for automated releases. F5/SM & WL plugins for failovers.
Configuration Management	Service Center is currently in used for Configuration Management.
Development Lifecycle	Quality Center for development lifecycle quality and consistency
Design Principles and Standards	Troux for architecture principles and standards for quality and consistency of designs.
Continuous Operations	Hardware Clustering allows for rolling HW/OS patching. Some hardware has hot-swappable components.

🟩 = Relatively New to Paychex

Strategic HA and CO Technologies (Investigative Stage)

Load Balancing	All Feature/functions of the F5 devices (LTMs/GTMs) that can enhance availability
Session State Management	No new technology investigations at this time.
Hardware Clustering	No new technology investigations at this time.
Database/Application Clustering	Oracle SOA Suite 11g to allow for implementation across datacenters
Fault-Tolerant HW/Self-Healing OS	Use of Superdome technologies.
Data Replication	Active Dataguard for data replication and read-only access to standby. Coherence for data caching and replication.
Data Recovery	Volume snapshots for data recovery.
Proactive Monitoring	Central monitoring console, extend use of existing tools for proactive vs. reactive monitoring
Incident Tracking and Reporting	Event Correlation tools, enhanced service level reporting tools
Automation	Technologies to: provide automated post-release test to decrease planned downtime windows, further automate release processes and build/turnover processes
Configuration Management	Is Service Center adequate for mapping application and system interdependencies?
Development Lifecycle	Technologies to deliver self-healing applications with rolling upgrade capabilites.
Design Principles and Standards	No new technology investigations at this time.
Virtualization	Investigate role in providing HA and Continuous Operations solutions.
Continuous Operations	Transient Logical Standby procedures in Oracle 11g to apply database patches in a rolling fashion

HA and CO Design Considerations and Procedures

Redundant Infrastructure	Storage, Network, Security, Server infrastructure must have redundant components.
Load Balancing	Load Balance Across data centers when appropriate. Load Balance based on performance metrics of target vs. using round robin approach.
Session State Management	Maintain active/active configurations by utilizing Session State tools.
Proactive Monitoring	Build health monitors into applications to be utilized by F5 devices. Continue process toward making Proactive Monitoring part of the design process not just implementation.
Incident Tracking and Reporting	Capture all Availability affecting incidents in Service Center while reducing the number incidents classified with unknown root cause.
Automation	Further Automate failover and recovery procedures, post-release testing and build processes to reduce time and errors.
Configuration Management	Ensure standby and primary systems have matching configurations. Thoroughly map application and system interdependencies/integrations.
Development Lifecycle	Matching availability of Build/Release automation tools with Test Engineering requirements Analyze Agile Development Model for impacts across all areas (support, release management, testing) Architect applications with high availability features to allow for more seamless failovers and rolling upgrades as well as self-healing features. Consider rewriting applications that have been “over-updated”. Determine best balance between frequency and complexity of release.
Design Principles and Standards	Develop Availability Design standards
Service Level Agreements/BIAs	Define accurate and comprehensive SLAs and availability requirements early in the project lifecycle process in order to affect approaches and designs. Define conditional SLAs for impaired vs. unavailable.
Testing Environments	Assess impacts for downtime on build automation and testing tools on testing timelines and completeness. Enhance integrated testing to make it more comprehensive. Design production-like test environments. Refine resource scheduling processes to make the most of the available test environments.

Predictive Model

- **What is it?**

  Model used to predict the overall platform availability based on probabilities of unplanned downtime.

- **What does it provide?**

  Allows for calculating the impact of component changes to overall availability. Can be used to help create Availability roadmap. Provides consistent way to assign availability expectations to given design as well as identify gaps.

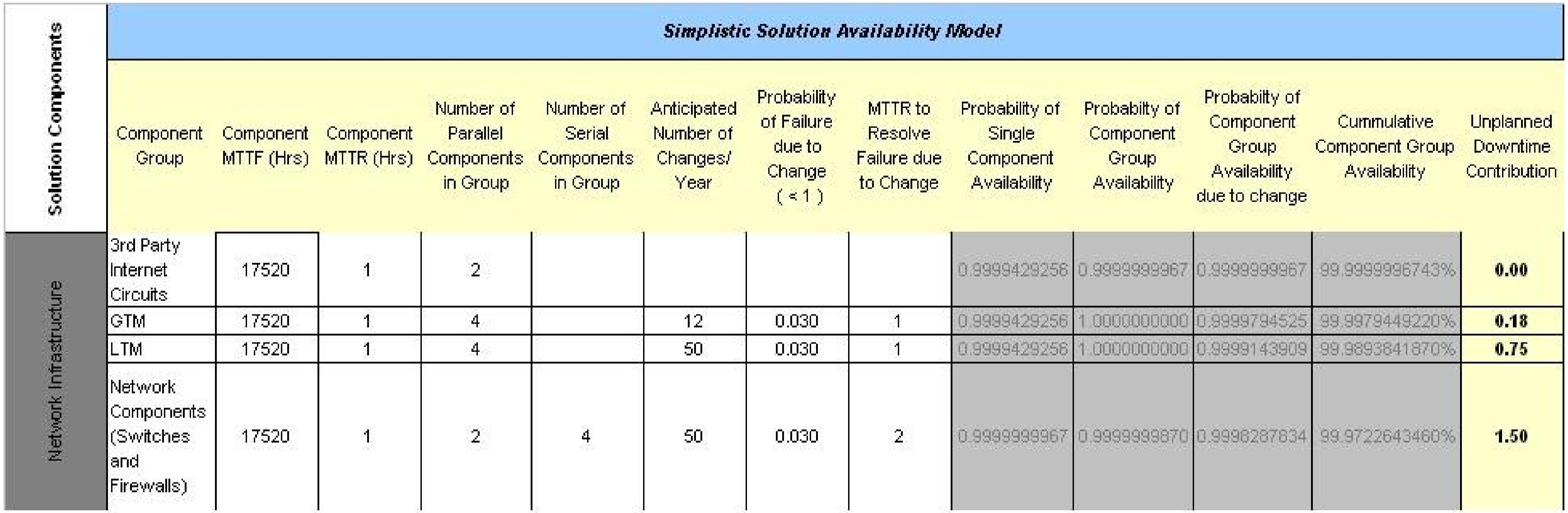
- **How is it used?**

  Prediction is attained through industry failure rates and Paychex history. Continuous feedback from actual experience used to update predictions.

- **When is it used?**

  This model can be used during the Solutions Approach and Design Phases to assess the predicted availability percentage of given Solution/Design. It can also be used as part of an assessment of an existing Product. Work is being done to add it to the project lifecycle as a deliverable.

Predictive Model Sample



Solution Components	Simplistic Solution Availability Model												
	Component Group	Component MTTF (Hrs)	Component MTTR (Hrs)	Number of Parallel Components in Group	Number of Serial Components in Group	Anticipated Number of Changes/ Year	Probability of Failure due to Change (< 1)	MTTR to Resolve Failure due to Change	Probability of Single Component Availability	Probabilty of Component Group Availability	Probabilty of Component Group Availability due to change	Cummulative Component Group Availability	Unplanned Downtime Contribution
Network Infrastructure	3rd Party Internet Circuits	17520	1	2					0.9999429256	0.9999999967	0.9999999967	99.9999996743%	**0.00**
	GTM	17520	1	4		12	0.030	1	0.9999429256	1.0000000000	0.9999794525	99.9979449220%	**0.18**
	LTM	17520	1	4		50	0.030	1	0.9999429256	1.0000000000	0.9999143909	99.9893841870%	**0.75**
	Network Components (Switches and Firewalls)	17520	1	2	4	50	0.030	2	0.9999999967	0.9999999870	0.9996287834	99.9722643460%	**1.50**