;

the Availability Digest

www.availabilitydigest.com

The State of Virginia – Down for Days

October 2010

The Commonwealth of Virginia lost dozens of its computer systems for over a week, bringing the activities of over two dozen of its agencies to a halt. Tens of thousands of the state's citizens were affected, some seriously. The web site of the Division of Emergency Management went dark just as Hurricane Earl was approaching.

How could so many major systems go down simultaneously? Why did it take so long to restore services? We look at this fiasco in this article.

The Outsourcing Decision

The story began in 2002 when then-Governor Mark Warner commissioned a legislative review of Virginia's computer systems. The review concluded that the state's computer systems were outdated and were becoming increasingly expensive to maintain.

Competitive bids were solicited for an overhaul of the entire computing network. In 2005, the contract was awarded to Northrup Grumman, a major defense contractor. The contract provided $2.4 billion over ten years to implement and manage a modern-day system to serve 76 state agencies. The contract, to be managed by the Virginia Information Technology Agency (VITA), was the largest contract of any kind ever awarded in Virginia's history.

The project was plagued from the beginning with poor performance and missed schedules. Under Governor Robert McDonnell, an October, 2009, legislative audit showed that the system had disrupted service at almost every state agency. There were an abundance of complaints about poor service and cost overruns. Rolling outages occurred because of the failure to install redundant communication networks. Northrup Grumman admitted it was nine months behind schedule. Yet the state's CIO was fired for criticizing Northrup Grumman's performance.

As a result of the audit, the contract was renegotiated in April of 2010. In return for increased penalties for poor performance, the contract was extended by three years and fattened with an additional $100 million.

The Debacle

The Cause

Around noontime on Wednesday, August 25, 2010, a large storage area network (SAN) in a data center in Chester, south of the state's capital, Richmond, began sending alert messages that something wasn't right. However, a redundant fault-tolerant EMC Symmetrix DMX SAN with automatic failover was in service, and the SAN continued operating properly.

;

It was determined by the maintenance staff that one of the SAN controllers needed replacing. A few hours later, a technician replaced the board; and pandemonium erupted. Though the root cause has yet to be announced, it was the initial belief that the technician pulled the good board of the fault-tolerant pair. At that point, the dual SAN crashed. 485 out of the state's 4,800 servers were knocked offline.

Maintenance staff tried through the night to fix the problem but to no avail. A decision the next day was made to shut down the entire system and to replace all suspected faulty components. The system was finally brought up two days later at 2:30 AM, Friday, August 27.

The Recovery

System testing then showed that the crash had caused severe database corruption. The only alternative was to rebuild the massive database from magnetic tape, a lengthy process. All data since the last tape backup was lost.

Database recovery from tape took almost a week to complete. Even then, only 97% of the data could be restored. A Minnesota company specializing in retrieving lost data was retained, but 3% of the data was irretrievably lost. For instance, thousands of photographs and signatures submitted by residents for drivers' licenses during the four days prior to the outage were lost.

EMC reacted quickly to try to return the system to operation. They said that they had never experienced such a disaster in one billion hours of operation of systems all over the world. They claimed that field experience showed five nines of SAN availability (an average of five minutes of downtime per year) and that internal checks ensured that dual failures did not corrupt data. It seems that you can protect data paths from sending corrupt data, but what do you do about fat fingers?

The Impact

The impact on state services was devastating. 26 of 89 state agencies were down, many for over a week. One such casualty was the Department of Motor Vehicles. All 74 DMV offices were closed, leaving an estimated 45,000 people unable to obtain or renew drivers' licenses.

The web site for the Department of Emergency Management went dark. It is used to communicate with Virginia residents during emergencies. This happened just as Hurricane Earl was approaching Virginia. In fact, Governor McDonnell declared a state of emergency for Virginia on September 1st, just as the system was being restored.

The State Department of Taxation was unable to access taxpayer accounts and to issue refunds for seven days. Local and state employees of the Department of Social Services worked through the weekend to issue food stamps and welfare checks to 380,000 people. The Department of Juvenile Justice could not release inmates. The Department of Veterans Services could not pay its bills. The Alcoholic Beverage Control, the Board of Elections, and the Virginia Employment Commission were down. The state could not process child-support payments or process claims for unemployment.

Following restoration of services, employees throughout the state government worked long hours to catch up on suspended services.

;

The Consequences

Financial

Northrup Grumman took responsibility for the outage. In a published announcement, it said:

> "We cannot afford to let any vulnerability in the infrastructure go unresolved. We have an obligation to [VITA] and the agencies and citizens they serve to learn everything we can from this occurrence. We will conduct a root cause analysis, carefully analyze and review the findings, develop lessons learned and make necessary changes."

Governor McDonnell, meanwhile, called for an independent third party to perform an "operational and performance review" of the outage and of the roles played by both VITA and Northrup Grumman. Northrup Grumman stated:

> "Northrup Grumman supports the Governor of Virginia Bob McDonnell's call for an independent analysis of the situation, from the cause of the initial failure through the process of restoration that has taken place. Our commitment to this partnership is absolute, and we are committed to making the Commonwealth's information technology infrastructure the best of its kind."

Per the contract, Northrup Grumman faces a $100,000 fine. In addition, it will pay $250,000 to cover the costs of the independent review. Should the review find negligence on Northrup Grumman's part, it could face significant additional fines.

Political

The costs of this outage went far beyond financial considerations. It became a serious political headache for Republican Governor McDonnell. As part of his election campaign, he had criticized his predecessor, Governor Timothy Kaine, for failing to properly manage the Northrup Grumman contract. He pledged to run such programs as a business. It was for this reason that he renegotiated the Northrup Grumman contract after he was elected to tighten the penalty clauses. In return, he increased the contract by three years and added $100 million to it.

The Democrats are using the incident to call into question Governor McDonnell's management style. They point out that Virginia cannot do without the computer systems but cannot afford the $400 million termination fee.

Lessons Learned

Fortunately, Northrup Grumman did one thing right. In designing the system, they elected to back up the database periodically to magnetic tape rather than to rely only on the replicated copy of the database on the fault-tolerant SAN. Otherwise, they would have found that both copies of the database were corrupted with little chance of recovery since any corruption to the online SAN would have been replicated to the backup SAN. Most if not all of the data for those 26 state agencies might have been irretrievably lost.

However, for a database this size, magnetic tape backup was perhaps not the right choice. There were too many tapes to manage, leading to lengthy restore times – in this case, several days. In addition, there was too much of a chance of finding a tape that could not be read. A better choice would have been virtual tape backup, in which tape images are replicated offsite to disk. (From there, they can be archived to tape if necessary.) Since virtual tape eliminates manual tape-handling procedures, it is simpler and can be scheduled more often.

;

Virtual tape can reduce both RPO and RTO. RPO is the *Recovery Point Objective*, or how much data loss can be tolerated. RTO is the *Recovery Time Objective*, or how long an outage can be tolerated. It appears that Northrup Grumman's tape backup schedule called for days between backups, as evidenced by the four days of data loss reported for the DMV. If virtual tape had been used, the database could perhaps be backed up one or more times per day, significantly reducing the amount of lost data. Furthermore, recovery time would have been significantly shortened since it is much faster to reconstruct a database from disk images than it is from magnetic tape.

Last, but certainly not least, it does not appear that recovery procedures had been thoroughly tested and practiced by the operations staff. It seems like the length of time it took to recover the database was a shocker to all. If recovery had been tested or practiced, staff would have known how long recovery should take.

The key lesson here is that if you are going to outsource your data center-operations, include in the contract the requirements to document, test, and periodically practice all recovery procedures and the right to witness and audit these procedures. This should be a fundamental tenant of any good business-recovery plan.

Acknowledgments

Virginia Data Center Experiences Massive Outage, *The Whir*; August 27, 2010.
Statewide computer meltdown in Virginia disrupts DMV, other government business, *The Washington Post*; August 28, 2010.
Massive computer failure hits Virginia systems, *SecurityInfoWatch.com*; August 30, 2010.
Virginia Governor Wants Investigation into Computer Failure, *myfoxdc.com*; August 30, 2010.
Northrup Grumman vows to fix Virginia computer problems, *The Washington Post*; August 31, 2010.
EMC system serving Va. breaks down, *Boston.com*; August 31, 2010.
Virginia Still Disrupted By IT Outage, *Information Week*; August 31, 2010.
Crash of Va. Computer network has implications for tech world, state politics, *The Washington Post*; September 2, 2010.
Governments, companies watch Virginia computer system failure, *newsleader.com*; September 2, 2010.
Contractor will pay $250K for Va. computer failure, *Bloomberg Businessweek*; September 13, 2010.
Audit Shows Failure of Virginia Computer Upgrade, *IT Business Edge*; October 14, 2009.
Tony Pearson, Inside System Storage, *IBM – My developerWorks: Blogs.*