High Reliability Organization (HRO) in Practice

By Lionel Dyck

This article takes the concept of High Reliability Organizations (HRO) into the practical realm. Before reading this article it would be helpful, but not required, that you read the first article on HRO in the September 2006 issue of *Technical Support* magazine, "HRO: The New TLA to Solve All Our Problems."

To summarize, HRO is a way of looking at your job and your environment in a 'mindful' way, which basically means that you both anticipate potential problems and that you are also prepared for them should they occur. Problems rarely occur when you expect them or in the manner that you anticipate. Rather they typically occur when you least expect them and in ways that surprise you. This does not mean that you are unable to anticipate or prepare, although it does mean that you need to be creative.

The HRO concepts to keep in mind are:

1. Preoccupation with Failure
2. Reluctance to Simplify Interpretations
3. Sensitivity to Operations
4. Commitment to Resilience
5. Deference to Expertise

Some have asked whether or not Root Cause Analysis, or RCA, is the same thing. The short answer is a qualified no. The longer answer is that RCA is a process that should be used after every problem to learn what caused the problem and how to either prevent it in the future or mitigate its impact should it occur again. The qualification to the short answer is because RCA should be a part of your standard thought process to be used whenever you encounter anything out of the ordinary. One benefit of doing a true RCA is that the learning that you derive from it should lead to questions about how that learning applies elsewhere. For example, if problem 'x' is caused by actions 'a' and 'b,' what other potential problems could arise if those actions were to occur at a different time or environment?

Some examples to make the point (as you read each example, think which of the HRO concepts would have helped):

Scheduled downtime is something that applies to every system that has been developed. The schedule may be daily, weekly, monthly, quarterly, or less frequently and is the result of a need to perform some action on the system. What happens when that schedule conflicts with a mission-critical process? Think of an emergency room where there is a single workstation that is used to make requests for blood from the hospital's blood bank. Is there ever a good time to schedule an outage to that system? Probably no. There are better times than others, so the best way to schedule the downtime would be to coordinate with the emergency room management and staff. Another option would be to have a second workstation capable of remaining active while the other workstation is unavailable. In this example the HRO concepts that come into play would be *Deference to Expertise* and *Sensitivity to Operations* in scheduling the downtime and *Preoccupation with Failure* in making a second workstation available.

Patching workstations is something that we've gotten used to over the past few years with the advent of Microsoft Update being incorporated into the Windows operating system. The challenge is in configuring this update to occur without impacting the user, or at least minimizing the impact. Using the above example of a workstation in an emergency room, think about the impact if the patch is installed automatically and then immediately causes the workstation to reboot. What if that were to occur while the workstation was being used to review the results of a laboratory report indicating toxic levels of some substance and informing the physician that based on the levels there is less than five minutes before irreparable harm occurs to the patient? If the workstation were to reboot at the point before the five-minute warning were to be displayed, the patient could die or suffer some other major harm. Using the concept *Preoccupation with Failure*, it would make sense to either design the workstation so that it does not require the patches (unlikely) or to prompt the user to defer the reboot until a more convienent time.

Disaster Recovey (DR) and Business Continuity (BC) are evidence of two HRO concepts: *Preoccupation with Failure and Commitment to Resilience*, as they imply that thought has gone into what could happen and how to recover from disasters. A question for each of you: If you have either DR or BC, when was the last time you tested your DR or BC procedures? If your answer is never or more than a year ago, then you need to question how viable they still are. Things change frequenty and your procedures need to be validated after every change. This validation could be as simple as a desk check or as complex as senior management walking into the data center and declaring a mock disaster. If you have a DR and/or BC environment in place and haven't validated recently, then you really don't know if they will work.

Recovery from a problem is just as important as the prevention of the problem. This is the *Commitment to Resilience* HRO concept. How long does it take to return service to the user after a problem occurs? When a problem occurs you want to capture as much information as possible to be able to do a Root Cause Analysis (RCA), but what is the impact of taking that time? The answer

to how much time should be taken in analyzing and capturing data before taking recovery actions is one that has to be made by those who understand the environment. This is *Sensitivity to Operations*, yet the other HRO concept *Reluctance to Simplify Interpretations* also comes into play. You need to be sensitive to the impact of the problem (e.g. the workstation being down, preventing the ordering of blood for the emergency room patient who has lost a lot) while understanding that if you don't take the time to capture the information about the situation that the problem could, and probably will, occur again and again. In this case a *Preoccupation with Failure* would suggest that before implementing the workstation(s) that procedures (manual or better yet automated) be created to capture all critical pieces of information as quick as possible to allow service restoration to begin immediately.

In each of these examples the HRO concepts are used by those who are 'mindful' (a term used in the HRO literature to describe someone who is always prepared for any unexplained problem and who anticipates problems before they occur).

Those who practice the *Preoccupation with Failure* concept are also aware of the Law of Unintended Consequences. This 'law' effectively warns us that some simple action, such as turning on the hose to water the lawn, can have negative results, such as over-watering, if we are not careful.

Another example of *Preoccupation with Failure* is if a mechanic working on an aircraft carrier were to misplace a small tool, the consequences could be catastrophic if that tool were to be sucked into the intake of a jet engine on a fighter bomber taking off. The Navy has this situation covered in that should this occur, the mechanic is required to report it immediately, at which point flight operations will cease and everyone will begin a search for that tool. The mechanic is not reprimanded in this situation. Rather, they are commended for reporting the missing tool. This demonstrates another part of the HRO culture where individuals are encouraged to come forward to report problems and potential problems even if they are the cause. If the individual were to be reprimanded, they would be reluctant to come forward and the results could be significant (e.g. the loss of a flight crew and a multi-million dollar aircraft with possible damage to the ship). Once the missing tool is found, then an RCA is made to determine why the tool was lost and procedures or training will be implemented to prevent it in the future.

So how does all this relate to your job?

While experiencing the next changes that you make in your environment, ask yourself some questions:

- Is this a single point of failure and if so, how great is the risk? What can be done to mitigate this exposure? Have you informed management of this?
- Has the change been validated? Did you desk check it? Have you tested it in a non-production environment? Do you have a backout plan should the change fail? Have all documentation and procedures been updated to reflect any new operational characteristics?
- What additional exposures does this change introduce to your environment and are you prepared to deal with them?

While experiencing the next problems that you encounter in your environment, ask yourself some questions:

- Could this have been anticipated and prevented?
- If the problem cannot be prevented, then can the recovery be automated?
- Do you know what information to capture to be able to do a full Root Cause Analysis (RCA) and can that data capture be automated in the future?
- What can be done to mitigate the disruption this problem causes the next time it occurs?

To become a High Reliability Organization, or HRO, does not happen overnight and it does not require that you use the HRO terminology. It does require that the HRO concepts and philosophy become ingrained in the organization from the front line staff to senior management. If you have senior management calling for the head of someone who makes a mistake then mistakes will not be reported and when they occur, there will be finger pointing and excuses rather than remediating actions.

I encourage everyone to read the book *Managing the Unexpected: Assuring High Performance in an Age of Complexity* (1) to gain a better understanding of HRO. It does not matter what industry one works in because there are mission-critical applications running on mission-critical servers in every company. The loss of one of these applications or servers at the wrong time could have disasterous results from the loss of customers, the loss of revenue, to the loss of life (possibly yours).

As I closed the first article: Remember that you need to "be prepared" and that "if it can fail, it probably will fail and it will fail at the worst possible time."

Some useful resources are:

1. *Managing the Unexpected: Assuring High Performance in an Age of Complexity* by Karl E. Weick, Kathleen M. Sutcliffe. Published 2001.
2. *HRO Has Prominent History* by Karlene H. Roberts, PhD, http://www.apsf.org/resource_center/newsletter/2003/spring/hrohistory.htm
3. *Beyond Normal Accidents and High Reliability Organizations: The Need for an Alternative Approach to Safety in Complex Systems* by Karen Marais, Nicolas Dulac, and Nancy Leveson, MIT, March 24, 2004 http://esd.mit.edu/symposium/pdfs/papers/marais-b.pdf
4. *Safety, Reliability, Stewardship, and Regret: Contributions to Dependable System Design from the Study of Highly Reliable Organizations* by Andrew Koehler, PhD, Statistical Sciences, D-1, Los Alamos National Laboratory, 12/16/2005 http://ti.arc.nasa.gov/projects/ishem/Presentations/Koehler_High_Reliability.ppt
5. 5 Habits of Highly Reliable Organizations http://pf.fastcompany.com/magazine/58/chalktalk.html
6. High Reliability Organizations Conferences http://www.highreliability.org/

Lionel Dyck has been working with computers for over 34 years and has written many articles over the years for *Technical Support* magazine including several on his open source z/OS SMTP mailing utility XMITIP. You can learn more about his open source tools at http://www.lbdsoftware.com.