Ultra Reliable Closed Loop Life Support for Long Space Missions

Harry W. Jones¹
NASA Ames Research Center, Moffett Field, CA, 94035-0001

and

Michael K. Ewert²
NASA Johnson Space Center, Houston, TX, 77058

Spacecraft human life support systems can achieve ultra reliability by providing sufficient spares to replace all failed components. The additional mass of spares for ultra reliability is approximately equal to the original system mass, provided that the original system reliability is not too low. Acceptable reliability can be achieved for the Space Shuttle and Space Station by preventive maintenance and by replacing failed units. However, on-demand maintenance and repair requires a logistics supply chain in place to provide the needed spares. In contrast, a Mars or other long space mission must take along all the needed spares, since resupply is not possible. Long missions must achieve ultra reliability, a very low failure rate per hour, since they will take years rather than weeks and cannot be cut short if a failure occurs. Also, distant missions have a much higher mass launch cost per kilogram than near-Earth missions. Achieving ultra reliable spacecraft life support systems with acceptable mass will require a well-planned and extensive development effort. Analysis must determine the reliability requirement and allocate it to subsystems and components. Ultra reliability requires reducing the intrinsic failure causes, providing spares to replace failed components and having “graceful” failure modes. Technologies, components, and materials must be selected and designed for high reliability. Long duration testing is needed to confirm very low failure rates. Systems design should segregate the failure causes in the smallest, most easily replaceable parts. The system must be designed, developed, integrated, and tested with system reliability in mind. Maintenance and reparability of failed units must not add to the probability of failure. The overall system must be tested sufficiently to identify any design errors. A program to develop ultra reliable space life support systems with acceptable mass should start soon since it must be a long term effort.

Nomenclature

\[ C = \text{Cost} \]
\[ CDR = \text{Critical Design Review} \]
\[ CIL = \text{Critical Item List} \]
\[ F = \text{Failure rate} \]
\[ FMEA = \text{Failure Modes and Effects Analysis} \]
\[ FRACAS = \text{Failure Reporting And Corrective Action System} \]
\[ FRB = \text{Failure Review Board} \]
\[ FTA = \text{Fault Tree Analysis} \]
\[ LEO = \text{Low Earth Orbit} \]
\[ M = \text{Number of copies of each subsystem} \]
\[ MTBF = \text{Mean Time Between Failures} \]
\[ N = \text{Number of different subsystems} \]
\[ NRC = \text{National Research Council} \]

¹ Systems Engineer, Bioengineering Branch, Mail Stop N239-8, AIAA Senior Member.
\( PDR = \) Preliminary Design Review
\( \lambda = \) Failure rate per unit time

I. Introduction

This paper describes an approach to develop the ultra reliable, minimum mass life support systems needed for long duration human space missions. If the mission overall Loss of Crew requirement is 1 in 1,000, each of the major mission elements such as life support must have a 1 in 10,000 or 1 in 100,000 probability of failure over the mission duration. The cost of delivering mass to the Moon or Mars surface is roughly an order of magnitude greater than for Shuttle or Space Station, since the equipment must be launched to low Earth orbit, then on to the destination, and either landed on the surface or returned to Earth orbit.

Long, distant space mission operations and maintenance will require a new approach to provide ultra reliability. The Space Shuttle and Space Station are maintained using spares that are provided as needed. Long mission life support will be completely different, since all the spares must be provided in advance, launched either with or before the mission. There will be no opportunity to provide further spares.

Ultra reliability can be achieved with a reasonable mass of spare components. Recycling equipment for oxygen or water recovery can be designed to have many components so that the failure modes are localized, separated, and contained, and so that the failure rate of each system is small because the total failure rate is distributed over a large number of components. Design studies of life support recycling systems show that a 1 in 10 probability of failure can be decreased to 1 in 10,000 by providing spares, and that the mass of spares is approximately equal to the original mass. (Jones 2008, Jones 2010) For an acceptable risk, the system reliability must have low variation, as well as a very high expected value. In other words, if there is a significant risk that the actual reliability is low, many spares must be provided even if they are unlikely to be needed.

II. Planning for reliable life support

The initial system design choices largely determine reliability, just as they establish performance and cost. A plan should be established to achieve the needed reliability. The required system reliability should be analyzed and allocated to subsystems and components. Technologies, materials, and components should be selected to improve reliability. Planning must include life testing of components to confirm their reliability and maintainability and spares requirements must be defined. The design should isolate the predominant failure modes in small, low mass, easily replaceable components. Analysis and test data must confirm that the system is capable of the required ultra reliability. The subsystems and integrated systems should be tested sufficiently to discover design errors that may cause failures. Life testing at the systems level is usually unable to measure or improve reliability, as the number of systems is small and the test time is limited.

There is a time sequence of opportunities to improve the operational reliability of life support during the mission. System development efforts can assess reliability, reduce the intrinsic failure rate, design the spares logistics, and plan for maintenance and repair. The crew can be selected and trained for the needed repair skills. Additional spares and diagnosis and repair equipment can be provided. The crew can be allocated more time for maintenance, diagnosis, and repair.

These opportunities have associated costs; development cost, crew skills and training cost (e.g., for a repair expert versus a geologist), spares and equipment launch mass cost, and crew mission time cost. The costs occur in the successive mission phases, design, crew training, launch, and operations. The natural tendency to postpone costs may lead to higher total cost and lower performance than if more effort was expended in design and planning.

III. Ultra reliable, low mass life support is needed for long missions

It is well accepted that higher reliability and onboard repair will be required for deep space, the Moon, and Mars.

The factors of reliability and maintainability will assume immense importance as U.S. human spaceflight advances to extended operations in deep space, on the lunar surface, and on Mars. There will be no rapid return capability; resupply will be slow, difficult, and expensive; refurbishment now accomplished on the ground will have to be accomplished on site. (NRC, p. 77)

A. NASA needs a dedicated effort to achieve ultra reliable life support

Ultra reliable life support must be demonstrated to show that we can go to Mars or an asteroid safely. Former NASA Administrator Mike Griffin suggested using the Space Station and the Moon as testbeds for Mars. Send astronauts to the International Space Station for a six- or nine-month visit, after which they would be sent to the Moon for a similar amount of time, equipped with no additional supplies beyond those sent with them to the station.
Once they completed their Moon visit, this same group of astronauts would return directly to the Space Station for another six- to nine-month visit, again with no resupply. Only then would they return home. (de Selding)

This is an important test, but success is 99 percent likely even with an unacceptably high 1 in 100 probability of failure at the mission level. If lunar outpost life support fails, on-demand supply of spares or a quick return will prevent the possible Loss of Crew. But the Moon base or other acceptable analog test beds are required to be confident of the life support system for long missions.

B. Failure probability must be less than 1 in 10,000 over the mission

The ultra reliable life support design requirement for long missions is approximately, “Develop high reliability life support systems (atmosphere, water, waste, thermal, etc.) such that all the life support systems add less than 1 in 10,000 to the missions’ probability of Loss of Crew or Loss of Mission.” To achieve an overall Loss of Crew and Loss of Mission of 1 in 1,000, all the major mission elements, such as life support, must be have a failure probability of 1 in 10,000. The individual life support systems, such as water, must similarly be an order of magnitude more reliable, 1 in 100,000, and these systems’ subsystems and components, should be yet another order of magnitude more reliable, with a 1 in 1,000,000 chance of failure during the mission.

C. Current life support systems apparently have insufficient reliability for long missions

Current recycling life support systems do not appear to have the ultra reliability required for long missions, which is much higher than that needed for near Earth short duration missions.

Likens noted that actual life support failure rates have been significantly greater than predicted. The predicted failure rates ranged from $3 \times 10^{-4}$ to $3 \times 10^{-5}$, a one order of magnitude range. However, the actual failure rates ranged from $10^{-1}$ to $10^{-5}$, a four orders of magnitude range. Almost always, with only one exception in fifteen cases, the actual failure rates were higher than predicted. They were usually a full order of magnitude higher. Storage, resupply, and non-recycling technologies were found to be significantly more reliable than physical-chemical water processors. Only failure mitigation using emergency oxygen and water reserves or repair and work-arounds were able to prevent disaster. (Likens)

Russell and Klaus state “total ECLSS maintenance for 865 days was found to exceed the design estimate by a factor of 22.” A contributing factor was the oxygen generation system’s greater than expected failure rate. (Russell and Klaus) William Gerstenmaier, NASA’s associate administrator for space operations, expected this failure cause to continue on the International Space Station. “We know that oxygen generating systems in general have a lot of problems over the years during start-up. We think we'll have some problems with our oxygen generator system.” (Malik)

Although more reliable components are necessary, and significant effort must be devoted to obtaining them, using the most reliable available components is insufficient to achieve the needed life support reliability. Components will fail and have to be replaced. This means spares must be provided, thus putting more pressure on lowering equipment mass.

D. Lower mass life support is also needed for long missions

The cost to launch equipment to low Earth orbit (LEO) using the Space Shuttle is typically estimated at $20 k/kg. (London) Moon and Mars missions have a much higher mass launch cost per kilogram than the Shuttle or Station, since equipment must be sent from LEO on to the Moon or Mars, and either landed or returned to Earth orbit. From LEO to a Moon landing or return to LEO, or to a Mars landing, the rocket and propellant mass is roughly 20 times the payload mass in LEO. From LEO to Mars orbit and return to LEO, the required mass is roughly 50 times the LEO payload mass. (Jones, 2003-01-2635)

Significant effort and cost are justified to reduce life support system mass for Moon and Mars missions. Suppose a space qualified component costs $100k and weighs 1 kg. Its cost is $100k per kg. But the round trip cost to Mars and back is $1,000k per kg. The delivered equipment cost is $1,100k per kg. It is cost-effective to spend funds for mass reduction at a rate of up to one million dollars per kilogram saved.

IV. Planning to achieve ultra reliable life support

Delivering ultra reliable life support will be very difficult, and an intense effort should start soon. Reliability analysis assumes components that have known probabilities of failure and are used in ways that do not compromise their reliability. Reducing failure probability significantly requires a large investment and long development time. System problems are often discovered only during integration and test and may require redesign and retest. Rework is a major cause of schedule delay and cost escalation.
Ultimately all failures can be traced to specific engineering and management decisions. In hind sight required procedures or good practices were neglected and the planned institutional checks and oversight fail to detect the errors. The chance that serious problems occur is increased by pressure to meet budget and schedule and the unavailability of adequate resources (dollars, launch mass, time). Explicit risk analyses that define the cost and benefit of reliability enhancements will improve decision-making results and hopefully prevent in-flight failures.

A. The reliability program plan

The development of ultra reliable life support should be guided by a reliability program plan. A good approach to implementing a reliability program is to share the responsibility between a design and development team that knows the hardware well and a specialist reliability organization that is familiar with reliability methods and tools. The knowledgeable design and reliability organizations should be directed, coordinated, and funded to undertake an increased effort to develop very high reliability, minimum mass life support systems. Reliability problems can occur if the reliability program is not established before preliminary design, since important failure modes may not be identified and controlled early.

The reliability program plan should include objectives, budget, schedule, and responsibilities. The reliability program plan should answer the following questions. What are the reliability requirements? How will they be met? How will reliability be estimated, predicted, and achieved? How will conformance to the reliability requirements be demonstrated? What are the potential indicators of nonconformance? What are the potential remedies for nonconformance? How will a solution be selected?

The reliability program plan must be synchronized with the project plan. The usual project life cycle provides sequential phases for requirements, design, development, test, and operations. Reliability should be included in the standard end-of-phase reviews, such as Preliminary Design Review (PDR) and Critical Design Review (CDR). The key issues that should be discussed in reviews include the reliability requirements, reliability plan, failure rate information and test data, failure reporting, failure rate estimates, reliability analyses, redundancy planning, effects of space launch and environment, and effects of long term storage of primary hardware and spares. The reliability planning should be coordinated with logistics, maintainability, testing, safety, and quality assurance. The familiar waterfall schedule of sequential phases does not easily accommodate changes and sometimes contributes to escalating overruns. The reliability analysis and plan must be kept current with design changes and new information and not allowed to become obsolete in later phases.

B. Initial reliability analysis and allocation

The initial reliability analysis is the earliest, easiest, and most effective way to reduce failure rate. Generic reliability databases can be used to estimate failure rates by conventional “bottom-up” reliability prediction techniques, based on schematics, parts list, and reliability block diagram models. Preliminary analysis based on failure rate data does not give accurate estimates of the expected reliability. However, it can be used to compare alternative components and designs, for rough allocations of the reliability requirement to subsystems, and to estimate the achievable range of reliability. Significant preliminary design is needed before reasonably accurate reliability assessments and tradeoffs can be made. Using part failure rates to estimate system failure rates does not account for interactions, tolerances, and possible design errors.

The uncertainty of reliability estimates, the error bars, should be quantified. The component reliability uncertainty can be propagated from bottom to top using Monte Carlo simulations. (Heydorn and Railback) The exponential reliability function implies chi-square confidence levels. (Meaker, p. 506)

The standardized Failure Modes and Effects Analysis (FMEA) provides worksheets and reports that can be used to prioritize resources and remedies, to show if and how the reliability requirement will be met, and to plan spares and maintenance. This is where a so called “graceful” failure scenario can be devised. For example, if 4 pumps together provide the total flow for a wastewater processor and one has an electrical failure, only 25% capacity is lost. Further, if the remaining 3 pumps can each be turned up in speed to achieve 33% more flow, then the “cost” of the failure was just an increase in pump power. Using computer tools for FMEA can save engineering cost. FMEA problems are identified on the Critical Item List (CIL). Fault Tree Analysis (FTA) is a complementary top-down approach used for the most severe failure effects. These analytical techniques are limited, since they investigate only the known causes of unreliability. They do not detect potential operator error, unexpected environment, or poorly developed requirements. Their results should be validated against historical failures and lessons learned.
V. Reliability planning for design, development, testing, operations, and maintenance

The reliability planning and analysis done during design and development should answer the following questions:

- How is the single thread system reliability requirement allocated to subsystems and components? What is the estimated bottom-up reliability? Is there a comparison of the reliability of alternate components, technologies, and designs?
- What are the expected system failure modes? What is their estimated likelihood and severity? What are failure histories of existing or similar components and designs? What are the expected failure mechanisms for new components and designs? Does the system have any consumable, wear-out, or other end-of-life problems?
- What failure rate reduction methods will be used? What component margins or de-rating, inspection or selection, life testing, or design for reliability approaches will be used?
- How much of the overall system reliability requirement will be achieved by highly reliable components and subsystems and how much by redundancy and spares? Will the system be partitioned and spares provided for ultra reliability with minimum mass? (Jones 2008, Jones, 2010)
- What failure tracking methods will be used? At the end of development, will test results and failures be reviewed? Are unit operating times recorded? What Failure Reporting And Corrective Action System (FRACAS) will be used?
- What will be the response to problem reports? Will failures be processed by a failure review board? Is the FEMA updated as needed? Are unverified and intermittent failures treated as possible hard failure precursors? Are repairs and redesigns well documented?

A. Component selection and system design for ultra reliability

The basic single thread system probability of failure can be reduced by conservative design. Known parts should be used with wide operating margins in a limited stress environment. Reliability should be a prime consideration in technology selection trade-offs. Simplicity makes for reliability. In standard reliability estimation practice, the system reliability depends on the parts list and on the parts’ failure probabilities. Reliability tools and analysis are employed to estimate reliability and identify problem areas. Component tests are used to verify operating margins, endurable environments, and failure modes. Non-random failures caused by design oversights and incorrect processes are identified and reduced by testing, investigation of problems, and configuration control.

Previously used designs and components can be redesigned for the lowest possible failure rates, but new equipment may have unknown failure modes and rates. If fault avoidance by component selection and system design is insufficient to achieve the required overall system reliability, this goal must be achieved by increasing the fault tolerance using redundancy or spares.

B. Development tests for ultra reliability

Only early test results can be used in the initial systems design to improve reliability. Life testing to measure and bound very low component failure rates requires either a large number of test articles, or a long test duration, or both. In the usual reliability demonstration life test, many units are tested until they all fail. The test duration usually must be several times the Mean Time Between Failures (MTBF), where MTBF = 1/λ, and λ is the failure rate per unit time. Cutting reliability life tests short, before a failure occurs, underestimates the reliability. Life testing is more useful for parts and components than for full systems, since the system design is difficult to modify. Entire systems are not usually tested over a long term under stressful environments, while components are.

The standard development tests, although not conducted for reliability reasons, can provide useful failure mode and rate information. Breadboard, integration, qualification, and acceptance tests frequently result in failures. These should be captured in a Failure Reporting And Corrective Action System (FRACAS) and reviewed by a Failure Review Board (FRB). Newly identified failure modes should be added to the FMEA and design changes made as needed. Recordings of operational parameter values rather than scoring pass/fail can facilitate failure analysis and trend tracking. Many of the failures found in long term testing are traceable to assembly-induced damage. (Fragola and McFadden) Infant mortality and end-of-life replacements can be significant deviations from the usually assumed constant random failure rates. (Fragola and McFadden)

C. Common cause failures

Common cause failures can disable both the original component and its spares. They are not cured by using similar item redundancy. Common cause failure modes may be discovered by test. Another way to reduce common
cause failures is to use different technology alternates, but the cost is increased for duplicate design, test, operations, and maintenance. Nevertheless, this possibility should be considered as part of long duration, regenerative life support system development. Diverse technology is most appropriate when the environment is not well known or sufficient testing is not possible. The dissimilar redundant system could be an old design, or perhaps have only “life-boat quality.” If common cause failures are relatively few, redundancy is still an effective method to improve reliability.

D. Using redundancy or spares to achieve ultra reliability

Redundancy or spares are used to protect against random failures that cannot be eliminated by component selection or systems design. If active redundant units are used, the power and cooling loads increase, as well as the number of failures. Also, on-line redundant units require performance monitoring and automatic switching. But spares require installation time and may fail while dormant, although most likely at a much lower rate than active units.

Using redundancy or spares to increase fault tolerance does not depend on knowing the component failure modes, only the failure rates, so using spares for fault tolerance is more certain than designing for fault avoidance. Component redundancy or sparing is the last resort in improving reliability, and it is employed only after component reliability has been established. Redundancy and spares use valuable launch mass and crew time resources, so it is necessary to first reduce the failure rate as much as practical by design. A greater investment in improved reliability should be balanced against the cost of failure, but only if the failure probability is acceptable.

The time to repair can be reduced by built-in test systems, automated fault detection, design for repair, modular design, quick disconnects, and standardized tools and methods. Preventive maintenance should be used in addition to repair on failure, to reduce unscheduled repair. The system design should specifically consider low reliability and limited life items, such as filters, fans, and motors. It might be possible to produce replacement parts on-board, using stored materials and computer controlled manufacturing. However training mock-ups and ground labs are needed to develop these types of capabilities. Also, systems can be monitored and faults diagnosed on the ground as an alternative to crew time and on-board capabilities. (Grenouilleau)

VI. Cost and benefit analysis to guide reliability planning

The required ultra reliability must be achieved for long mission life support, regardless of cost. Generally, higher reliability can be obtained only with much higher costs, either by reducing the already small component failure rates or by adding still more redundancy or spares. Reducing the life support reliability requirement can cut costs, but the life support reliability requirement is determined by the mission need. Only an excess margin beyond the required reliability is tradable for other goals.

Making investments to reduce the failure rate should be guided by cost-benefit estimates. During development, reliability and cost estimates should be used to indicate the cost and benefit of the different potential investments to reduce failure rate. In providing spares, the failure probabilities and masses of the components indicate the cost and benefit of particular combinations of spares. Analysis can define the least costly approach to meet the failure rate goal. If resources are insufficient to meet the goal, cost-benefit optimization will produce the best achievable failure rate. Integer programming has been used to allocate spares to reduce the weight of a life support system while maintaining the required reliability. (Hwang et al.)

Design for failure prevention should minimize the total cost to meet the ultra reliability requirement. This is accomplished by increasing reliability in the most cost effective way. The fundamental trade off is between higher design and development cost to reduce the single string system failure rate, or more launch mass and cost to provide more spares. This trade off tends to favor providing spares for two reasons. Increasing component and system reliability is accomplished during the development phase, while launching spares and installing them as needed are part of the operations phase. Reducing early development costs results in larger later operations costs, but postponing costs is often necessary for budget reasons even if the total cost increases. The second reason favoring using spares is that the cost grows more rapidly for improving reliability by component development than by providing operational spares. The result is that operational phase cost tends to dominate the total mission life cycle cost.

A. Improve component reliability or add spares?

Developing intrinsically more reliable systems requires time and hardware experience. This means that keeping technology development and flight hardware design teams integrated and functioning for years, and even decades, is very important. It requires incurring costs for analysis, design, parts selection, process improvements, test, and
failure monitoring. Redundancy or spares are used against random failures remaining after the development of reasonably reliable single string systems. Spares are used at the component, assembly, or higher replaceable unit level. The mass cost of providing spares to improve reliability increases less rapidly than the development cost for improving component reliability, but it is very large since mass launch cost is a major factor in life cycle cost and may exceed the development cost.

Is it more cost effective to design more reliable components or to provide spares to replace failed components? A rule of thumb in reliability design is that cutting the probability of failure in half requires an investment equal to the original development cost. If the original cost is C for an original failure rate, F, of 1 percent, it costs 2 C to achieve 0.5 percent and 3 C to achieve 0.25 percent. (Rechtin, p. 165) The mathematical relation is Total Cost = C [1 + log2 (F/desired failure rate)]. An order of magnitude reduction in the failure rate, say from 1 percent to 0.1 percent, would cost 4.32 times the original cost. The cost increases as the logarithm of the failure rate decreases. More common estimates have the cost increasing much more rapidly, exponentially. Even the slow logarithmic rate of cost increase is much more rapid than the cost increase rate when spares are used to reduce the failure rate.

The cost of launching a system is proportional to the total number of units, M, including the original unit and spares. For one original unit and one spare, M = 2 and the launch cost doubles. Analysis shows that taking a system, dividing it into N series subsystems all with equal mass and failure rates, and then providing M copies of each subsystem reduces an initial failure probability of F to F^M/M! N^M-1. (Jones, 2008-01-2160) (Jones, 2010-01-dependability) Suppose that N = 100 and F = 0.01 for M = 1, corresponding to no spares. For M = 2, adding one set of spares, F = 0.01 is decreased to the failure probability F^2/[2 N] = 0.5 * 10^-4. Adding one spare per subsystem cuts the failure probability by a factor of F/[2 N] = 0.5 * 10^-4. Increasing the mass cost further to M = 3, for two sets of spares, will multiply the M = 2 failure probability by another slightly smaller factor of F/[3 N] = 0.33 * 10^-4, for a very small final failure rate of F^3/[3*2*N^2] = 0.17 * 10^-10. The cost increases as the exponent of the failure rate decreases.

The failure rate decreases at a dramatically more rapid rate for providing spares than for improving component reliability. If the total development cost is increased by an amount equal to the original development cost, the failure rate can be cut in half. If the launch cost is increased by an amount equal to the original launch cost, by adding one more spare for each subsystem, the original failure rate is reduced by a factor equal to the original failure rate divided by the number of subsystems, typically to only one percent of the original failure rate.

B. Systems design to improve the mass-reliability trade-off

Humphries et al. observe that mass savings are possible if heavier components are allocated lower reliability requirements and lighter components are allocated higher reliability requirements. (Humphries et al, p. 353) This would tend to occur automatically if mass is added in the most effective way to increase reliability, but it should be specifically implemented. But further, the observation suggests that systems should be designed so that the likely failure modes and much of the probability of failure are contained in low mass components. Again, this would tend to occur without specific attention to it in systems design, but it should be explicitly done. The spares mass versus reliability formula shows that providing spares at the unit level to achieve a certain failure rate requires more mass than at the sub-unit level, since the failure rate improvement is greater for a larger number of subsystems. However, lower level repair requires more different types of parts and higher skills.

VII. Conclusion

Human exploration of distant, long duration mission objectives in space requires the development of ultra reliable, minimum mass life support systems. Current life support system reliability is too low for missions that cannot be resupplied or cut short. Current life support system mass is too high for planetary missions.

The life support systems must be able to operate for long durations with ultra reliability using initially supplied spares for maintenance and repair. The first steps to providing such systems are improving basic life support system reliability and reducing mass during system design and development. Analysis and component testing can reduce failures caused by component selection, design, and process errors. If the life support systems have reasonably high reliability, ultra reliability can be achieved using spares with about two times the original mass. Therefore, a strong direct and integrated approach will be required to achieve ultra reliable, low mass life support systems for long missions.
VIII. References


