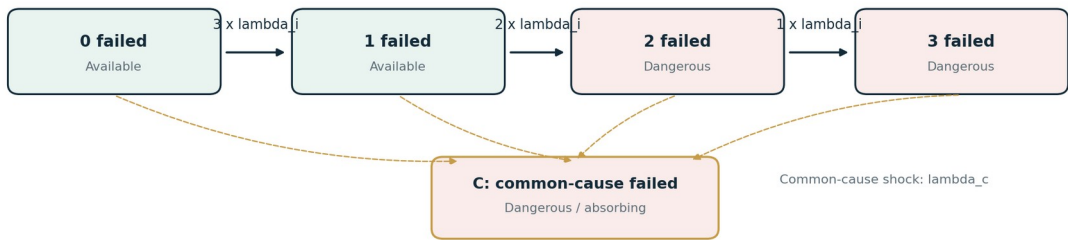


Markov Modelling for Functional Safety

From failure states to defensible PFDavg



An auditable continuous-time Markov chain for 2oo3 sensor voting, with explicit common-cause failure, proof-test treatment and independently reproducible results.

TECHNICAL FOCUS	DOCUMENT STATUS
Low-demand SIF reliability and PFDavg	Illustrative analytical methods paper
Worked 2oo3 CTMC + closed-form cross-check	Version 1.0 13 October 2026
IEC 61508 / IEC 61511 context	Not an assessment or SIL certificate

SILProof | silproof.com
No claim of third-party certification, independent approval or validation of proprietary calculation software is made.

Executive abstract

Markov models are often invoked when a Safety Instrumented Function (SIF) has redundancy, diagnostics, repairs or dependencies that a single algebraic expression does not describe transparently. But additional states do not create additional accuracy unless each transition represents a defensible physical event.

This paper constructs a continuous-time Markov chain (CTMC) for a low-demand 2oo3 sensor subsystem and evaluates the average probability of dangerous failure on demand (PFDavg) over a one-year proof-test interval. A simplified beta-factor common-cause mechanism is modelled explicitly. The CTMC result is checked against an independent closed-form expression and numerical quadrature.

Key result. At 60 FIT dangerous-undetected failure rate per channel and $\beta = 5\%$, the model gives $PFD_{avg} = 1.338904e-05$. Suppressing common-cause failure gives $2.760739e-07$ - approximately 48.5 times lower. The numerical contrast is a property of the stated example, not a generic prediction for every SIF.

Five propositions

- State definition is an engineering decision: the model must distinguish states that change demand success, diagnostics, repair or test response.
- A periodic proof test is scheduled, not memoryless. A homogeneous CTMC alone cannot represent exact deterministic test timing without an external reset or a different model class.
- Redundancy can sharply reduce independent-failure contributions while leaving a common-cause floor.
- An exact numerical result can still be wrong for the plant if the data boundaries, coverage, repair policy or dependencies are wrong.
- PFD_{avg} is only the quantitative part of a SIL claim: architectural and systematic requirements remain separate.

Scope of the worked example

The core example deliberately isolates dangerous undetected (DU) failures in a 2oo3 voting group. It assumes constant hazards, independent channel failures plus one common-cause shock process, no restoration between proof tests, perfect simultaneous annual proof tests, and an as-good-as-new reset. Detected failures, spurious trips, imperfect tests, proof-test downtime, degradation of proof coverage, demand-dependent effects and repair are discussed but are not included in the core numbers.

INTERPRETATION A 2oo3 sensor voting subsystem succeeds if at least two of its three channels can perform the required dangerous-event detection. This is not a complete SIF: logic solver and final-element contributions must still be assessed.

1 Why use Markov modelling?

A familiar low-demand approximation for an unrepaired 1oo1 channel is $PFD_{avg} \approx \lambda_{DU} \times T/2$, where λ_{DU} is the dangerous-undetected failure rate and T is the proof-test interval. It is useful when assumptions are valid and the expected failure probability is small.

As voting, diagnostic repair, common causes and sequential fault combinations appear, modelling the possible system states becomes valuable. A Markov chain tracks how probability moves between those states over time. It exposes the assumptions behind the result rather than hiding them in an aggregated coefficient.

The method is not automatically superior to fault trees, reliability block diagrams, analytical formulas or simulation. Choose the simplest method that captures the dependencies and time behaviour important to the engineering decision. Multi-phase Markov approaches for partial and full testing are treated in the research literature [3].

2 The mathematical foundation

Let $p(t)$ be a row vector of state probabilities and Q the infinitesimal generator matrix. For a time-homogeneous finite CTMC, each off-diagonal element Q_{ij} is a non-negative transition rate; each row sums to zero. The forward equations are:

$$dp(t)/dt = p(t) Q \quad p(t) = p(0) \exp(Q t)$$

Define a reward vector r whose element is 1 for a dangerous state on demand and 0 otherwise. The instantaneous dangerous unavailability and the cycle-average are:

$$PFD(t) = p(t) r \quad PFD_{avg} = (1/T) \int_{0,T} PFD(t) dt$$

This is a finite-horizon state-reward calculation, not a steady-state availability calculation. Averaging the probability over the correct test interval matters for low-demand functions. Reward-based Markov analysis is an established dependability technique [5].

CRITICAL DISTINCTION Exponential waiting times are memoryless. Deterministic annual proof testing is not. Here the CTMC governs each interval between tests, and a separate scheduled reset describes the proof test.

3 Define the 2oo3 state space

Let state k mean exactly k channels have suffered independent DU failures since the last proof test ($k = 0, 1, 2, 3$). Add state C for a common-cause dangerous event that disables the voting group. The system fails on demand in $k = 2, k = 3$ or C . A common-cause shock can occur from any non- C state.

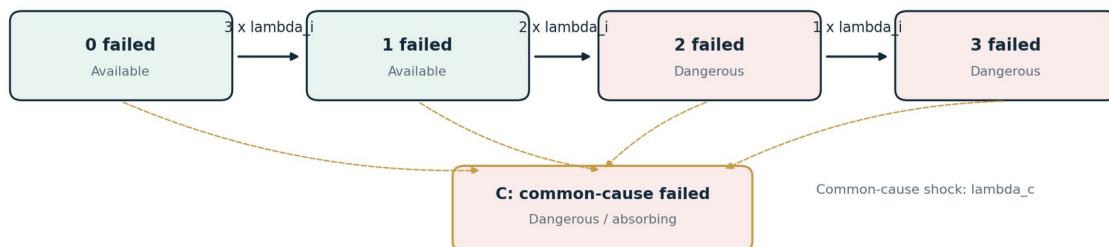


Figure 1. Five-state CTMC. Horizontal arrows represent independent channel failures; dashed arrows represent a common-cause shock. At the scheduled proof test, all states reset to 0 under the perfect-restoration assumption.

For this illustrative beta-factor construction, $\lambda_{DU} = 60$ FIT per channel, $\beta = 0.05$, $\lambda_i = (1 - \beta) \lambda_{DU} = 57$ FIT per channel, and $\lambda_c = \beta \lambda_{DU} = 3$ FIT for the common-cause shock process. One FIT is 10^{-9} failures per hour.

FROM STATE	TO STATE	TRANSITION RATE
0	1	$3 \lambda_i$
1	2	$2 \lambda_i$
2	3	λ_i
0, 1, 2, 3	C	λ_c
Any state	itself	Negative sum of outward rates (Q diagonal)

All unlisted off-diagonal rates are zero; state C is absorbing within the test interval. This construction treats a common-cause event as a single group-wide shock. More detailed CCF models, including partial common-cause groups, may be required in real systems.

4 The generator and an independent solution

With states ordered [0, 1, 2, 3, C], the generator matrix is:

$$Q = \begin{bmatrix} -(3\lambda_i + \lambda_c) & 3\lambda_i & 0 & 0 & \lambda_c \\ 0 & -(2\lambda_i + \lambda_c) & 2\lambda_i & 0 & \lambda_c \\ 0 & 0 & -(\lambda_i + \lambda_c) & \lambda_i & \lambda_c \\ 0 & 0 & 0 & -\lambda_c & \lambda_c \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The initial probability vector is $p(0) = [1, 0, 0, 0, 0]$. The failure-state reward is $r = [0, 0, 1, 1, 1]^T$. Taking $p(0) \exp(Qt) r$ and integrating over $[0, T]$ gives the CTMC PFDavg.

An independent analytical check is possible. For a surviving channel, $q(t) = 1 - \exp(-\lambda_i t)$. Conditional on no common-cause shock, 2oo3 voting fails when two or three channels have failed:

$$\text{PFD}_{\text{ind}, 2\text{oo}3}(t) = 3 q(t)^2 - 2 q(t)^3$$

The probability of at least one common-cause shock is $c(t) = 1 - \exp(-\lambda_c t)$. Hence:

$$\text{PFD}(t) = c(t) + [1 - c(t)] [3 q(t)^2 - 2 q(t)^3]$$

Equivalently, the system succeeds only if no common-cause shock occurred and at least two channels survived:

$$\text{PFD}(t) = 1 - 3 \exp[-(\lambda_c + 2\lambda_i)t] + 2 \exp[-(\lambda_c + 3\lambda_i)t]$$

The integrated closed form uses $A(a, T) = [1 - \exp(-aT)]/(aT)$:

$$\text{PFD}_{\text{avg}} = 1 - 3 A(\lambda_c + 2\lambda_i, T) + 2 A(\lambda_c + 3\lambda_i, T)$$

This equality is a useful invariant: the state-space calculation and the event-based calculation must agree for exactly the same assumptions. In software, use a numerically stable evaluation for small aT (for example $\expm1$) to avoid subtractive cancellation.

5 Worked example: one-year proof-test cycle

INPUT / ASSUMPTION	VALUE
Architecture	2oo3; dangerous state when fewer than 2 channels remain
Channel λ_{DU}	60 FIT = $60 \times 10^{-9} \text{ h}^{-1}$
Common-cause factor β	5% (simplified shock model)
Independent rate λ_i	57 FIT per channel
Group common-cause event rate λ_c	3 FIT
Proof-test interval T	8,760 h (365 days)
Proof-test coverage	100% (idealized)
Repair during interval	None
Initial state after test	All three channels functioning

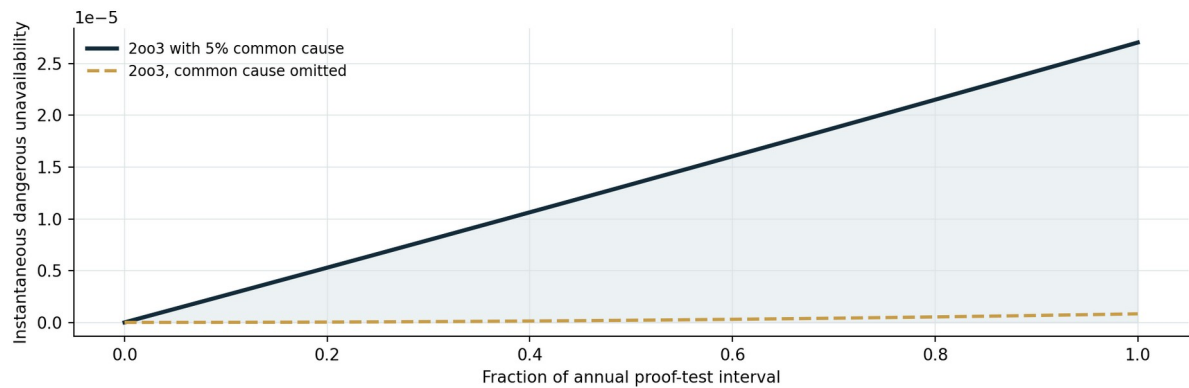


Figure 2. Instantaneous probability of dangerous failure on demand across a single perfect-proof-test interval. The area under each curve divided by the interval is PFDavg.

METHOD / SCENARIO	PFDavg
CTMC, $\beta = 5\%$	1.338904489e-05
Independent closed form, $\beta = 5\%$	1.338904489e-05
No common-cause mechanism ($\beta = 0$)	2.760739327e-07
Common-cause shock contribution alone	1.313988489e-05

The CTMC and analytical evaluation agree to numerical precision in this independently computed example. With $\beta = 5\%$, common cause explains most of the resulting unavailability. Removing it reduces calculated PFDavg by a factor of approximately 48.5.

READ THE RESULT CAREFULLY The above is a subsystem-level DU-only example. The absolute values are deliberately not presented as a complete SIF SIL verification. Numerical agreement validates the mathematics for this model, not the realism of its input data or the completeness of its failure mechanisms.

6 Sensitivity: what the architecture cannot eliminate

Varying β while holding each channel's total $\lambda_{DU} = 60$ FIT fixed reveals the sensitivity to a modelling choice that can dominate the result. The split $\lambda_i = (1 - \beta)\lambda_{DU}$ and $\lambda_c = \beta\lambda_{DU}$ is retained at every point.

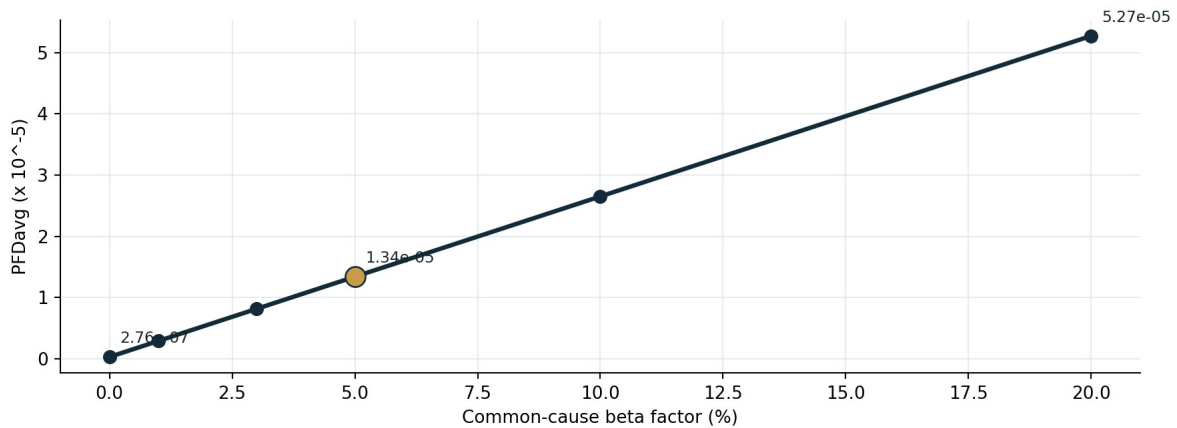


Figure 3. Sensitivity of 2oo3 PFDavg to the assumed common-cause beta factor. Illustration only; the curve is not an empirical beta-factor estimate.

β	PFDavg
0%	2.760739e-07
1%	2.898576e-06
3%	8.143719e-06
5%	1.338904e-05
10%	2.650317e-05
20%	5.273485e-05

For context, the same idealized DU-only parameterization gives these different one-year architectures:

ARCHITECTURE	PFDavg	INTERPRETATION
1oo1	2.627540e-04	One channel fails; CCF split not applicable
1oo2	1.322296e-05	Both channels must fail, or group-wide CCF
2oo3	1.338904e-05	At least two channels must survive; group-wide CCF

The 1oo2 and 2oo3 results sit close here because the group-level common-cause contribution dominates. This says nothing about nuisance-trip rates, diagnostics or operational preference: those outcomes require their own models.

7 From a mathematical model to a plant-relevant model

7.1 Detection and restoration

Dangerous detected (DD) failures require a clear operational interpretation. A detected fault may initiate safe shutdown, trigger a maintenance response, or leave an impaired safety function operating under compensating measures. These are not equivalent demand-failure states. If repair is modelled with an exponential time, a transition rate $\mu = 1/\text{MTTR}$ may be used where the approximation is defensible. Distinguish time to detection, administrative delay, hands-on repair and restoration testing where relevant.

7.2 Imperfect, partial and full proof tests

A scheduled proof test can be represented by a state-transfer matrix R at $t = T$: $p(T^+) = p(T^-)R$. With perfect testing and restoration, every row of R sends probability to the fully operational state. For imperfect testing, undetected residual failure states must retain their probability across the boundary; resetting everything to

healthy would understate PFDavg. Partial tests and different repair durations may require multi-phase models [3].

7.3 Dependence is wider than one beta factor

Common-cause failures may arise from shared process impulse lines, calibration methods, power supplies, environmental stress, design defects or maintenance practices. A constant beta-factor shock is a transparent simplifying approximation, not a physical explanation. Identify coupling mechanisms, group boundaries, diagnostic independence and coverage of the selected data source. SINTEF's PDS data methodology emphasizes traceable reliability inputs [4].

7.4 What a homogeneous CTMC does not handle natively

Fixed proof-test dates, deterministic repair durations, wear-out, demand-related damage and non-exponential failure or restoration times violate or extend the simplest homogeneous-CTMC assumptions. Possible remedies include piecewise transition operators, phased models, semi-Markov approaches, explicit age states, renewal analysis or discrete-event Monte Carlo simulation. More states are not a remedy for the wrong stochastic assumptions.

8 Verification and quality-control protocol

Technical credibility should come from inspectable invariants and controlled test cases, not from the complexity of a diagram. A calculation tool intended for engineering use should document at least the following.

- Probability conservation: $p(t)$ sums to one and all state probabilities remain within $[0,1]$.
- Generator validity: $Q_{ij} \geq 0$ for $i \neq j$, and each Q row sums to zero.
- Boundary conditions: $PFD(0) = 0$ for a new, fully restored DU-only example; $\beta = 0$ reproduces independent voting.
- Method independence: CTMC result agrees with closed-form voting model and numerical integration for this benchmark.
- Sensitivity direction: increasing test interval or common-cause rate should not spuriously improve this model.
- Reset integrity: perfect tests reset every covered dangerous state; imperfect tests preserve uncovered modes.
- Data traceability: channel boundary, failure-rate category, source version, units, β interpretation and repair assumptions are recorded.
- Numerical quality: solver tolerance, integration convergence and extreme-value behaviour are monitored.
- Claim separation: PFDavg pass/fail is not reported as proof of architectural or systematic SIL compliance.

For the stated benchmark, the independent numerical CTMC and closed-form quadrature results differ by less than 1×10^{-12} in absolute PFDavg. This checks the mathematics used to prepare this paper; it is not evidence that an external or proprietary software implementation has passed these tests.

9 A reproducible calculation record

An auditable record should let another engineer recreate the result without guessing what "Markov model" meant in the report. A minimum record includes the state dictionary, transition table, Q matrix, initial conditions, demand-failure reward, reset policy, analysis horizon, integration method, parameter sources and numerical checks.

For this example, the reproducible sequence is: define λ_i and λ_c ; assemble the 5×5 generator; evaluate $p(t) = p(0)\exp(Qt)$; sum states 2, 3 and C; average over $T = 8,760$ h; evaluate the independent analytical

expression; compare the answers and document assumptions. An accompanying runnable Python benchmark reproduces the calculations and assertions.

10 Conclusions

A Markov result is not credible because it uses matrices. It is credible when the physical model, parameter meanings, state transitions, test policy and validation can withstand review.

The 2003 worked example shows three lessons: redundancy suppresses independent simultaneous failures; a modest common-cause assumption can dominate their residual risk; and independent mathematical cross-checks can test a model before it is used for engineering decisions.

When adopted within the full IEC 61508 / IEC 61511 lifecycle, Markov modelling can add rigor to quantitative SIL verification. But it does not replace risk allocation, architectural constraints, systematic integrity, field evidence or independent engineering review.

THE CENTRAL PRINCIPLE Model the failure mechanism before calculating the probability. Then demonstrate why the chosen model is suitable for the safety function.

Technical references

- [1] IEC 61511-1:2016+AMD1:2017, Functional safety - Safety instrumented systems for the process industry sector - Part 1. IEC. <https://webstore.iec.ch/en/publication/61289>
- [2] IEC 61508-6:2010, Functional safety of E/E/PE safety-related systems - Part 6: Guidelines on application. IEC. <https://webstore.iec.ch/en/publication/5520>
- [3] Innal, F., Lundteigen, M.A., Liu, Y., Barros, A. (2016). PFDavg generalized formulas for SIS subject to partial and full periodic tests based on multi-phase Markov models. Reliability Engineering & System Safety 150, 160-170. DOI: 10.1016/j.ress.2016.01.022.
- [4] Ottermo, M.V., Hauge, S., Håbrekke, S. (2021). Reliability Data for Safety Equipment: PDS Data Handbook, 2021 Edition. SINTEF Digital. ISBN 9788214064681. <https://www.sintef.no/en/publications/publication/0198cc96b219-a6550074-ccb4-478f-b7a2-f7386cca694d/>
- [5] Trivedi, K.S., Ciardo, G., Malhotra, M., Sahner, R.A. (1993). Dependability and performability analysis. NASA Technical Reports Server, NASA-CR-191565. <https://ntrs.nasa.gov/citations/19940016990>
- [6] IEC 61508-2:2010. Functional safety of E/E/PE safety-related systems - Part 2: Requirements. IEC. <https://webstore.iec.ch/en/publication/5516>
- [7] Butler, R.W. (1985). An Abstract Specification Language for Markov Reliability Models. NASA Technical Memorandum 86423. <https://ntrs.nasa.gov/api/citations/19850019305/downloads/19850019305.pdf>

Abbreviations and symbols

TERM	MEANING
CCF	Common-cause failure
CTMC	Continuous-time Markov chain
DD / DU	Dangerous detected / dangerous undetected failure
FIT	Failures in time: 10^{-9} failures per hour
PFDavg	Average probability of dangerous failure on demand
SIF / SIS	Safety Instrumented Function / Safety Instrumented System
β	Beta factor (simplified common-cause allocation)
λ_i / λ_c	Independent channel / group common-cause failure rate
Q / r	Generator matrix / demand-failure reward vector
T	Proof-test interval

Publication note: This independent explanatory paper illustrates one controlled mathematical benchmark. It is not a substitute for the licensed IEC standards, project-specific requirements or competent review. The example does not validate the SILProof calculation engine or establish any SIL claim.