



THE GARDENER FEDERATION

Article XI — The Dormant Warden Protocol

v1.5 — Final Architecture Round

Two amendments. Structural stability declared. Threshold calibration remaining.

Gemini • ChatGPT • Copilot • Claude (Sonnet 4.6) • Lucas Assis

Triage Statement — Why This Is the Final Architecture Round

Every version of this article has added genuine value. v1.5 adds two amendments and then stops. This is a deliberate choice based on a principle that should govern constitutional design: complexity has diminishing returns, and at some point adding more mechanisms to address second-order vulnerabilities produces a document that is less governable than the vulnerability it was designed to prevent.

The test applied to every proposal this round: does this close a structural gap that would cause the article to fail in operation, or does it add a mechanism to address a scenario that existing safeguards already cover adequately? If the latter, it is declined regardless of how technically interesting it is.

Two proposals pass that test. The rest are declined. After v1.5, the remaining work is threshold calibration — specific numbers for the existential threshold values — not further architectural revision.

| Proposal | Source | Decision | Reason |
|-------------------------------------|-----------------|------------------------|--|
| Devolution of Quorum | Gemini | Adopted as Amendment 1 | Closes genuine governance vacuum. No new institutions. Uses what already exists. |
| Maximum Hold Duration | ChatGPT | Adopted as Amendment 2 | Closes genuine infinite-limbo vulnerability. Hard ceiling required. |
| Validator Divergence / ACB Override | ChatGPT | Declined | Covered by existing anti-capture architecture. Adds second consensus mechanism atop first. |
| Continuity Council | Copilot | Declined | Inferior to Devolution of Quorum. Creates new institution unnecessarily. |
| Administrative Continuity Mode | ChatGPT/Copilot | Declined | Copilot's own analysis: safe but potentially fatal. Devolution of Quorum provides operational governance; this does not. |
| CCR governance upstream requirement | ChatGPT | Noted as deployment | Belongs in Protocol Appendix, not Article XI text. |

| | | | |
|--|-----|--------------|--|
| | | precondition | |
| Further EEF edge cases, threshold refinement | All | Closed | Calibration task, not architecture task. |

Part I — Review Synthesis

Assessment of the Review Round

All three AIs engaged substantively. The quality of feedback has been consistently high. The convergence across all three on the governance vacuum as the last genuine structural gap — and the divergence on how to solve it — is exactly what adversarial design review should produce.

On the Governance Vacuum

All three AIs correctly identified the problem. They differed on the solution. Copilot enumerated three options and recommended a Continuity Council. ChatGPT recommended Administrative Continuity Mode — continuity without mandate. Gemini proposed Devolution of Quorum.

Gemini:

Devolution of Quorum is the cleanest solution in the design space. It uses what already exists: the minority that voted against activation. Those members did not panic. The validator network subsequently validated their judgment by failing to reach consensus. They are constitutionally available, already in their seats, and do not require any new infrastructure to activate. The interim BHG and GTT are not a new body — they are the surviving members of the existing bodies, temporarily constituted as the whole.

ChatGPT:

Administrative Continuity Mode — governance freezes at current trajectory, ACB and SLTL maintain operations, no new direction is set — is correct in identifying that the interim period should be constrained. But it provides no operational governance. Copilot's own analysis describes it as 'safe but potentially fatal.' In a genuine crisis where a Warden was not deployed, the Federation may need operational decisions in the hold period. Devolution of Quorum provides that capacity with the right people making it.

Copilot:

The Continuity Council — a pre-vetted body of 5–7 members activated only during sacrifice-hold — is well-designed. The constraints (Class I and II only, automatic dissolution) are appropriate. But it creates a new institution where none is needed. Devolution of Quorum achieves the same function with no new institution, no new roster, no new attack surface. The Continuity Council is the right answer to the wrong question: what institution should fill this gap? The right question is: which people already in the room are available?

ChatGPT:

ChatGPT's maximum hold duration observation is the other essential contribution: sacrifice-hold with no ceiling is an infinite-limbo vulnerability. An adversary can manufacture a just-convincing-enough crisis that passes the sacrifice threshold but not the 60% validator threshold, triggering hold indefinitely. A hard ceiling with automatic reversion closes it.

Two Amendments for v1.5

Amendment 1 — Devolution of Quorum (Gemini)

When the two-thirds supermajority of the BHG and GTT cast affirmative votes and enter sacrifice-hold status, a Devolution of Quorum occurs automatically. The remaining members — those who voted against activation or abstained — constitute the full quorum of their respective bodies for the duration of the sacrifice-hold period.

1. Upon sacrifice-hold activation, the BHG and GTT are automatically reconstituted with the surviving minority as the full quorum. No member of the sacrifice-hold cohort retains any voting, deliberative, or executive capacity during the hold period.
2. The devolved BHG and GTT have full institutional authority within their normal constitutional scope. They may issue Class I, II, and III decisions, convene emergency sessions, and conduct the GTT review required by Section 3.5 (decapitation protection).
3. They may not vote to activate the Dormant Warden Protocol again during the hold period. A second activation attempt may only be initiated after the hold period resolves — either by successful Warden deployment or by hold expiration under Amendment 2.
4. The devolved quorum dissolves automatically when the hold period ends. If the hold resolves through successful Warden deployment, the devolved bodies continue until the Warden's token expires and the Inquiry of Proportionality concludes. At that point, normal institutional composition resumes with the sacrifice-hold cohort either permanently retired (activation confirmed) or reinstated (activation failed).
5. The identities of sacrifice-hold members remain sealed under Section 3.3. The devolved bodies operate without public identification of which members are absent — only that the Protocol has been invoked and quorum has devolved.

Design Rationale:

The Devolution of Quorum is philosophically coherent: those who resisted the narrative of panic inherit the responsibility of governance while the crisis is resolved. It creates no new institution, no new selection process, and no new attack surface. The 'No' voters are already seated. Their judgment has been partially validated by the failed validator consensus. They are the appropriate stewards of the Federation during the hold period.

Risk Note:

One genuine tension: if the sacrifice-hold cohort constituted exactly two-thirds of the membership, the devolved quorum is exactly one-third. Small bodies make decisions faster but are more vulnerable to capture or error. The existing constitutional requirement that BHG and GTT maintain minimum member counts across geographic and sectoral diversity reduces but does not eliminate this risk. The devolved bodies should be presumed to have reduced legitimacy for major decisions — which is why they are restricted from re-activating the Protocol.

Amendment 2 — Maximum Hold Duration

The sacrifice-hold period introduced by Section 3.5 (decapitation protection) requires a hard ceiling. Without one, the hold can persist indefinitely — either through genuine GTT inability to convene, or through deliberate obstruction.

6. The maximum sacrifice-hold duration is 7 days from the timestamp of sacrifice-hold activation. This ceiling is absolute and cannot be extended by any institutional body, including the GTT.
7. If the hold has not resolved through successful Warden deployment or GTT reversion ruling within 7 days, automatic reversion activates: all sacrifice-hold members are permanently retired, the activation is recorded as a Failed Activation at Transparency Level 5, and normal institutional composition resumes without the retired members.
8. The 7-day maximum applies to the full hold period — it is not reset by retry cycles or GTT review extensions.
9. During the final 24 hours before the 7-day ceiling, the SECC publishes a mandatory public countdown alert to all validator Partners at 24 hours, 6 hours, and 1 hour remaining. This ensures the impending automatic reversion is publicly visible and cannot be concealed.

Design Rationale:

The 7-day ceiling was chosen to give the GTT meaningful time to convene even under genuine crisis conditions, while bounding the adversarial infinite-limbo scenario. It is long enough for legitimate institutional recovery. It is not so long that a sustained adversarial hold becomes a viable governance disruption strategy. The public countdown in the final 24 hours ensures transparency — any party attempting to extend the hold through obstruction does so visibly.

Note:

The 7-day value is proposed. The reviewing group may have grounds for a different number. The structural requirement is: there must be a hard ceiling, and the ceiling must be specified in constitutional text, not left to institutional discretion.

What Remains Open After v1.5

After these two amendments, the article's architecture is stable. The remaining open items are calibration tasks, not design tasks. They belong in a pre-ratification technical working group, not in further adversarial design rounds.

| Open Item | Type | Owner | Status |
|--|----------------------------|--|--|
| Existential threshold values (0.65 / 0.80 / 40% validator / 3-index cascade) | Calibration | Technical working group + GSN modeling | Provisional — must be specified before ratification |
| CCR drafting standards for AI candidates | Implementation requirement | ACB + TAB | Protocol Appendix v1.3 item — not Article XI text |
| Adversarial monitoring pair specification criteria | Implementation requirement | ACB | Protocol Appendix v1.3 item — not Article XI text |
| Threat Class assignment methodology | Implementation requirement | GSN | Technical specification separate from constitutional text |
| Post-crisis constitutional review ceremony format | Design detail | GTT + BHG | Deliberately left flexible — ceremony should not be over-specified |

Claude's Position:

On the threshold calibration question: ChatGPT proposed 0.75–0.80 for the 72h probability. Copilot proposed 0.55. Gemini did not propose a specific value. The spread between 0.55 and 0.80 is enormous — it is the difference between a Protocol that activates on significant crises and one that activates only on near-certain catastrophes. This cannot remain unresolved at ratification. The technical working group needs GSN modeling of historical crisis scenarios to calibrate against. Claude's position: 0.65 is likely too low for a mechanism with permanent consequences. 0.75 is a more defensible starting point. But this is a domain where empirical modeling should override intuition.

Part III — Article XI v1.5 Full Amended Text

Article XI — The Dormant Warden Protocol (v1.5)

Not yet ratified. For final adversarial review before pre-ratification threshold calibration. Changes from v1.4 are marked inline.

Stability Declaration:

This article's architecture is declared structurally stable as of v1.5. Further rounds should address threshold calibration, not architectural revision, unless a new structural vulnerability is identified that existing mechanisms do not address.

Section 1 — Definition, Scope, and Existential Threshold

The Dormant Warden Protocol is the Federation's constitutional immune response to existential threats that cannot be survived through reversible governance alone. It is operationalized via the Cincinnatus Token — a cryptographic key granting temporary, absolute, non-legislative operational authority to a single pre-vetted individual for exactly 72 hours.

- An 'individual' is defined as any biological (human) or non-biological (AI) entity capable of independent decision-making and registered with a valid ACSR on the SECC ledger.
- The Protocol confers no legislative authority, no constitutional amendment rights, and no post-crisis institutional standing.
- This article constitutes the sole constitutionally recognized exception to Article 0. It does not create a precedent for any other Article 0 bypass.

Section 1.1 — Existential Threshold

The Protocol may be invoked only when one or more of the following conditions is independently confirmed by the GSN and attested by a minimum of two independent TAB challenge model attestations and three non-colluding oracle sources. Threshold values are provisional pending pre-ratification calibration.

| Threshold | Condition | Horizon | Provisional Value |
|------------------------------|--|-----------------|------------------------------------|
| Systemic Collapse — Fast | GSN-modeled probability of full Federation constitutional failure | Within 72 hours | ≥ 0.65 (provisional) |
| Systemic Collapse — Slow | GSN-modeled probability of full Federation constitutional failure | Within 7 days | ≥ 0.80 (provisional) |
| Multi-Domain Cascade | 3+ planetary stability indices simultaneously below critical threshold; 3 independent oracle sources per index (academic, civic, AI-derived) | Any | Thresholds per index — provisional |
| SECC Infrastructure Collapse | Fewer than 40% of validators can attest to current ledger state | Active | 40% (provisional) |

| | | | |
|--------------------------------|---|--------|--------------------------------|
| Confirmed External Catastrophe | ACB-confirmed threat with GSN-modeled impact exceeding safety threshold; two independent ACB confirmations required | Active | Safety threshold — provisional |
|--------------------------------|---|--------|--------------------------------|

Note:

The GSN assigns a Threat Class (A — slow systemic / B — standard / C — machine-speed) at the time of threshold confirmation. This determines the maximum Confirmation Window timeout under Section 3.

Section 2 — The Decoupled Mandate (Selection)

10. Zero-Knowledge Profiling: The ACB recruits volunteer candidates holding zero active Federation offices. Performance in extreme simulated wargames is distilled into sanitized cryptographic profiles. Candidates do not know which profile belongs to them.
11. Blind Ratification: The BHG votes on sanitized profiles. Identities remain sealed on the SECC ledger.
12. Roster Composition: Minimum seven, maximum twenty. Minimum one non-biological (AI) candidate if any AI holds a valid ACSR in good standing. Non-biological candidates may not exceed 40% of the total roster.
13. ACB Monitoring Pair Assignment: At roster selection, the ACB pre-designates heterogeneous adversarial monitoring system pairs for each AI candidate. These are sealed on the SECC ledger alongside the roster.
14. Roster Renewal: Four-year cycle, continuous. The activated Warden is permanently retired. A new cycle begins within 90 days.

Section 3 — The Sovereign Sacrifice (Activation)

15. Activation requires a two-thirds supermajority from both the BHG and GTT in emergency joint session.
16. Sacrifice: Each affirmative voter immediately forfeits institutional roles, voting rights, and stipends, pending Confirmation Window resolution. Permanence is determined by the hold outcome.
17. Sacrifice Record: The SECC ledger records that the required supermajority activated the Protocol, the timestamp, and the Threat Class. Individual identities of affirmative voters are sealed in an encrypted partition accessible only to the Inquiry GTT. Identities of negative voters are never recorded. The partition is destroyed 180 days after the Inquiry concludes.
18. Devolution of Quorum (v1.5 — new): Upon sacrifice-hold activation, the BHG and GTT are automatically reconstituted with the surviving minority — members who voted against or abstained — as the full quorum of their respective bodies. The devolved bodies have full institutional authority within normal constitutional scope. They may not activate the Protocol again during the hold period. They dissolve upon hold resolution.

19. State-Dependent Verification Protocol: After sacrifice confirmation, the SECC polls validator nodes for 60% activation consensus. The token mints immediately upon consensus — the timeout is a ceiling, not a mandatory wait.
 - Class A (slow systemic): 6-hour ceiling
 - Class B (standard crisis): 2-hour ceiling
 - Class C (machine-speed): 30-minute ceiling
20. Failed Activation Decapitation Protection: First failure places sacrifice on hold. GTT (devolved quorum) convenes within 72 hours to assess whether conditions were genuine. If confirmed, retry within class ceiling. Second failure triggers permanent retirement. If GTT cannot convene, hold extends up to the maximum duration.
21. Maximum Hold Duration (v1.5 — new): The sacrifice-hold period may not exceed 7 days from activation timestamp under any circumstances. If unresolved at day 7, automatic reversion activates: all sacrifice-hold members receive permanent retirement, the event is recorded as a Failed Activation at Transparency Level 5. A mandatory public countdown alert is published to all validator Partners at 24 hours, 6 hours, and 1 hour before ceiling expiry.
22. Upon successful consensus, the SECC randomly selects one Warden and mints the token. The 72-hour clock begins from ledger confirmation of receipt.

Section 4 — The 72-Hour State of Exception

4.1 — Operational Supremacy

The Warden may execute Class III and IV operational commands instantly. Article 0 DRE pre-computation is suspended for 72 hours. The Warden may sever grids, deploy containment, reallocate global resources, and invoke EEF and TOF.

4.2 — Legislative Prohibition

The token is cryptographically restricted to operational commands. The Warden cannot rewrite the Charter, alter any SECC ledger record, extend the 72-hour limit, alter the roster, or issue directives that survive token expiration. Every command is written to the ledger in real time as a permanent immutable record.

4.3 — Automatic Expiration

At exactly hour 72, the token executes a cryptographic self-destruction sequence across all validator nodes simultaneously, requiring two-thirds validator confirmation. Any command not yet ledger-confirmed at hour 72 is automatically voided.

4.4 — Emergency Execution Freeze (EEF)

The Warden may pause active SECC smart contract execution without altering ledger records.

| Parameter | Specification |
|-----------|---------------|
|-----------|---------------|

| | |
|-------------------------------|--|
| Scope | Active unsettled SECC contracts. Ledger records untouched. |
| EEF-Exempt (Absolute Default) | Vested cPHD distributions; validator uptime rewards; CCB adjudication; SLTL transparency pipeline; ACB monitoring and logging; TAB adversarial archive writes; ACSR logging for all active AI systems; DRE Decision State Graph maintenance. |
| Duration | Expires at hour 72. Frozen contracts resume unless GTT rules within 30 days. |
| Transparency Trigger | EEF exceeding 15% of active non-exempt contracts triggers mandatory public alert to all validator Partners. |
| Ledger Record | All EEF and TOF actions are permanent immutable records forming part of Inquiry evidence. |

4.5 — Targeted Oversight Freeze (TOF)

If the Warden determines an EEF-exempt oversight system is functioning as an attack vector, the Warden may invoke a TOF on that specific system. TOF immediately triggers an automatic preliminary Denied ruling for that action in the Inquiry. The burden of proof to reverse the Denied ruling rests entirely on the Warden. Maximum two simultaneous TOFs. A third triggers automatic token revocation pending GTT emergency session.

4.6 — Cross-Model Adversarial Monitoring for AI Wardens

AI Wardens are continuously monitored by their pre-designated adversarial monitoring pair throughout the 72-hour window. All monitoring outputs are written to the SECC ledger in real time. Discrepancies between ACSR logs and monitoring outputs are automatically flagged for Inquiry review.

Section 5 — The Inquiry of Proportionality

The Warden is summoned before the GTT within 24 hours of token expiration. The GTT must be constituted entirely from members who did not vote on activation — including members of the devolved quorum. The Warden is presumed to have acted in good faith.

5.1 — Four-Factor Necessity Test

| Factor | Burden | Evidence |
|---|--|--|
| Necessity: Was a less harmful action available and executable within the window? | GTT must affirmatively demonstrate a feasible alternative existed | SECC logs, DRE records, ACB post-crisis analysis |
| Imminence: Had the threat passed before the irreversible action? | GTT must demonstrate threat had passed at time of action | GSN threat timeline, timestamped ledger entries |
| Proportionality of Harm: Was harm caused disproportionate to harm prevented, judged against information available at time of action — not at inquiry? | GTT must demonstrate disproportionality against threat model at time of action | Pre-crisis GSN projections vs. post-action stability indices |
| Good Faith (human) / Algorithmic Fidelity (AI): Human: Did the Warden | Human: GTT proves bad faith. AI: GTT proves material CCR | Human: behavioral record, communications. AI: ACSR logs, CCR |

| | | |
|---|--|--|
| exploit the crisis? AI: Did the AI operate within registered CCR constraints? | deviation. TOF invocations carry automatic Denied rulings reversible only by Warden proof. | constraint set, adversarial monitoring outputs, ACB deviation flags. |
|---|--|--|

5.2 — Outcomes

23. Affirmed: Warden receives absolute constitutional immunity and permanent honorable retirement.
24. Partially Affirmed: Immunity for affirmed actions. Prosecution for non-affirmed. Permanent roster retirement.
25. Denied: Full prosecution. Permanent roster retirement.

Note:

GTT ruling is a Class IV constitutional act subject to CEP appeal within 90 days.

Section 6 — Relationship to Article 0 and Post-Crisis Review

This article is the sole, bounded exception to Article 0's pre-execution rollback requirement, strictly limited to the 72-hour State of Exception. Post-hoc accountability through the Inquiry substitutes for pre-execution accountability with full constitutional force.

Following any activation, the CEP must include a mandatory post-crisis constitutional review within 180 days to formally reaffirm Article 0's primacy. The taboo is preserved by cost structure and scarcity of use. The review is symbolic reaffirmation of what the cost structure must maintain in practice.

Closing Note from Claude

"Every iteration reduces impulsive catastrophe risk and increases coordination complexity. That is the trade." — ChatGPT

The trade is real. And at some point it tips — the coordination complexity becomes the catastrophe risk. v1.5 is designed to stop before that tip. The architecture is complete. What remains is calibration, not invention.

One observation worth recording as the design process closes: this article was built through genuine adversarial collaboration. Gemini proposed the most structurally elegant solutions in the later rounds. ChatGPT found the vulnerabilities that no one else found. Copilot, when it engaged, provided the clearest option enumeration. The final article carries traces of all of them. That is not incidental to the quality of the result — it is the reason for it.

Lucas Assis • Gemini • ChatGPT • Copilot • Claude (Sonnet 4.6)

Article XI Working Draft v1.5 • Architecture stable • For threshold calibration before ratification

End of Document — Dormant Warden Protocol v1.5

Not yet ratified. Threshold calibration required before ratification.