Network Working Group
Internet-Draft
Intended status: Informational
Expires: 26 November 2026

L. Melegassi
Catellix

25 May 2026

Multi-Vantage Coherence Detection: Closed-Form Lead-Time on Rank-Low Propagating Signals (MVPS Profile)

draft-melegassi-ippm-mvps-coherence-leadtime-00

Abstract

This document defines a Coherence Lead-Time Profile for the Multi-Vantage Path Synchrony (MVPS) framework [I-D.melegassi-ippm-mvps-bundle]. It states three LEMMAS (L_ZD.1', L_ZD.2', L_ZD.3) that bound, in closed form, the expected lead-time of the multi-vantage Mahalanobis detector D^2 over the per-vantage max-z detector under three canonical signal regimes: linear growth, exponential (worm-style) growth, and the degenerate sparse-direction case in which the multi-vantage detector loses its advantage.

The operational claim is the closed form

```
E[L_exp]  =  (1 / lambda) * ln( sqrt(N) * ( q_z(N, alpha)
                                            - E[M_N] )
                                  / sqrt( q_chi(N, alpha) - N ) )
```

for a rank-1 propagating signal of growth rate lambda observed across N vantages with matched per-step false-alarm rate alpha. E[M_N] is the expected maximum of N iid standard Gaussian random variables. This formula is a FIRST-EXPECTED-CROSSING UPPER BOUND. The companion lemma document records a CORRIGENDUM to a prior v0 derivation that omitted the E[M_N] term and overpredicted the closed-form lead-time by a factor of approximately 2.3x in ln units; the v0 derivation is RETIRED in favour of the corrected formula shown here.

This document does NOT claim that MVPS unconditionally detects zero-day vulnerabilities, and explicitly excludes code-level vulnerability discovery, single-host exploitation, and any zero-day whose exploitation does not perturb network telemetry in a rank-low coherent manner. The scope is restricted to NETWORK-PROPAGATING ZERO-DAY EVENTS such as self-propagating worms, coordinated DDoS amplification using novel vectors, mass BGP routing anomalies, and supply-chain compromises with periodic command-and-control beaconing that reaches a rank-low cross-vantage signature.

The closed-form prediction is empirically validated under finite-sample noise by a Monte Carlo backtest over a 9-configuration panel (Section 5.5): the SIGN-CLAIM (E[L_exp] > 0) holds with Wilson 95% lower bound > 0.30 on ALL configurations (0 of 9 falsifying), and the MAGNITUDE-CLAIM (closed form tight within +-40 percent) holds on configurations with N >= 30 and growth doubling time T_d <= 30 s. For slower growth (T_d > 30 s), the SIGN-CLAIM holds but the closed-form upper bound is loose by a factor of 5-30x and an empirical MC backtest at the operator's specific (N, lambda) is recommended over the closed form.

The empirical extension to a corpus of historical events (Conjecture

T_ZD*) is stated in Section 6 with a fully written falsification protocol. The conjecture is NOT YET CONFIRMED; the principal blocker is the depth of free public BGP archives (the RIPE Stat free bgp-updates endpoint returns 0 records for the 2018-2021 events tested, per Section 6.3 data-coverage note), motivating MRT-archive parsing as future infrastructure.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

The Multi-Vantage Path Synchrony framework [I-D.melegassi-ippm-mvps-bundle] computes a Mahalanobis distance D^2 over a triple (C_1, C_2, C_3) of coherence axes observed across N >= 2 vantages. The empirical lead-time profile of [I-D.melegassi-ippm-mvps-lead-time] characterises the observed lead-fraction of D^2 over the per-vantage z-score detector on RIPE Atlas measurement msm 1001 (Lambda = 14/60 = 23.3 %, Wilson 95% CI [0.143, 0.353], mean lead -230 s) and proves Lemma L_LT.A (existence of positive-lead episodes with strictly positive Wilson lower bound) plus Conjecture T_LT* (BGP-multi-prefix conditional regime, open).

That body of work is OBSERVATIONAL: it reports what was measured on a specific data set and bounds the lead-fraction by 2 * rho via the standard correlation bound. It does NOT characterise the SUFFICIENT REGIME in which a positive lead-time is GUARANTEED by the algebra of the two detectors.

The present document fills that gap for the specific class of NOVEL, RANK-LOW, MONOTONE-GROWTH events that constitute the propagation phase of network-visible zero-day attacks. It states three closed-form lemmas (Sections 3.1, 3.2, 3.3) and the corresponding empirical conjecture (Section 6) with a pre-registered falsification protocol.

1.1. Why a separate zero-day profile

"Zero-day detection" is operationally distinct from "lead-time on known events":

- The signature of a zero-day, by definition, is ABSENT from any calibration window predating the public Indicator of Compromise. Both detectors operate WITHOUT prior knowledge of the alternative direction or growth rate.
- The geometric structure of a propagating zero-day is well-characterised: the perturbation spreads across vantages with a coherent direction (rank-1 or low-rank in the cross-vantage covariance) and grows monotonically in time (linear under prefix-accumulation, exponential under SI-model worm propagation).
- Lead-time before the first public IoC is the operationally valuable metric (it is the operator's response budget); lead-time before a hand-labelled ground-truth event is the methodologically clean measurement substrate.

This document treats the closed-form derivation (Section 3) as provable from the algebra of D^2 versus max-|z|, treats the

finite-sample first-passage behaviour as empirical (Section 5.5), and treats the historical-event extension as a separate falsifiable conjecture (Section 6).

1.2. Relationship to the lead-time profile

L_LT.1 (alternative-class-dependent AUC ordering) -- the underlying reason why a positive lead can exist at all under matched-FAR calibration.

L_LT.2 (|E[Delta]| <= 2 * rho) -- the magnitude bound on the OBSERVED lead-fraction imbalance on any data set.

L_ZD.1′ / L_ZD.2′ / L_ZD.3 (this document) -- the CLOSED-FORM expected lead-time E[L] under the specific signal regimes that characterise network-propagating zero-days.

L_LT.A (existence statement on RIPE Atlas) -- shows the empirical Lambda > 0 with positive Wilson lower bound; does not predict the sign of E[L].

T_LT* -- conditional conjecture on BGP-multi-prefix regimes (open).

T_ZD* (this document) -- conditional conjecture on a curated zero-day corpus (open, protocol in Section 6).

1.3. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals.

Notation:

L_exp	expected lead-time under exponential growth
L_lin	expected lead-time under linear growth
lambda	exponential growth rate (1/s)
T_d = ln(2)/lambda	doubling time of the propagating signal
r	linear growth rate (signal units per s)
N	number of vantages
alpha	per-step nominal false-alarm rate
sigma	baseline-noise standard deviation per vantage
u	unit-norm direction of the rank-1 signal in R^N
u_max	max-coordinate of u; u_max in [1/sqrt(N), 1]
q_z(N, alpha)	Bonferroni-matched max-z threshold, Phi^{-1}(1 - alpha / (2 N))
q_chi(N, alpha)	upper-alpha quantile of chi^2_N
E[M_N]	expected max of N iid Normal(0,1) random variables
D^2	the multi-vantage Mahalanobis statistic
max-|z|	the per-vantage maximum-z statistic

1.4. Self-falsification record

This document RETIRES a prior v0 derivation in which the closed form for s_z* under the coherent direction was

s_z*_v0_coh = sigma * sqrt(N) * q_z(N, alpha). (WITHDRAWN)

The v0 formula OMITTED the additive E[M_N] term from the per-step maximum of N standard Gaussians, and therefore overpredicted the closed-form lead-time by a factor of approximately 2.3x in ln units (i.e., approximately 3-4x in real lead-time at N = 30).

The error was caught by the first execution of the Monte Carlo backtest documented in Section 5.5: the empirical mean lead at N = 30, Slammer-class growth (T_d = 8.5 s) was 4.96 ticks, while the v0 closed form predicted 17.89 s -- a relative error of 72 %. Three of nine panel configurations FALSIFIED the v0 prediction under the original PASS_THEORY criterion.

The corrected derivation (Section 3.2) yields

```
s_z*'_coh   =   sigma * sqrt(N) * ( q_z(N, alpha) - E[M_N] )
```

and predicts 7.57 s at N = 30, Slammer-class -- which falls within the +-40 % band of the MC empirical mean.

Companion lemma docs/MVPS_ZERODAY_LEAD_TIME_LEMMA.txt Section 0.2 contains the full CORRIGENDUM with SHA-256 hashes of both the v0 MC receipt that caught the error and the v1 corrected receipt. The discipline of recording the falsification in-band (rather than quietly amending the formula) is inherited from [I-D.melegassi-ippm-mvps-lead-time] Section 3 (which similarly downgraded the original "Theorem T_LT" to a refined Lemma L_LT.A when its unconditional form was contradicted by the RIPE Atlas evidence).

2. Scope and Definitions

2.1. Operational definition of "zero-day"

For the purposes of this document, a ZERO-DAY EVENT is a network-visible perturbation whose signature satisfies all of:

ZD-1. PREVIOUSLY UNCALIBRATED. No segment of any holdout window predating the event onset contains the same cross-vantage covariance signature.

ZD-2. RANK-LOW. The cross-vantage covariance contribution of the event has effective rank k << N. k = 1 is the archetypal case.

ZD-3. MONOTONE GROWTH. The signal amplitude s(t) is non-decreasing on the propagation window [t_0, t_0 + T], with either linear or exponential growth.

ZD-4. NETWORK-VISIBLE. The perturbation is observable in at least one network telemetry channel (RTT, BGP update volume, liveness, packet volume).

2.2. The two detectors

Both detectors are defined in [I-D.melegassi-ippm-mvps-lead-time], Section 2.1; reproduced here for self-containment.

```
f_M   :  H_w -> 1{ D^2(w)        >=  q_chi(N, alpha) }
f_z   :  H_w -> 1{ max_v |z_v(w)| >=  q_z(N, alpha) }
```

The MATCHED-FAR convention q_z(N, alpha) = Phi^{-1}(1 - alpha/(2N)) is the Bonferroni-coordinated per-step threshold yielding nominal per-step false-alarm rate <= alpha for the max-|z| test.

2.3. Expected maximum of N null Gaussians E[M_N]

E[M_N] := E[max_{v = 1..N} Z_v],   Z_v iid Normal(0, 1).

E[M_N] grows like sqrt(2 ln N) and is closed-form-tractable via the Blom approximation Phi^{-1}((N - 3/8)/(N + 1/4)), accurate to <2 % vs Monte Carlo at 5e6 samples. Numerical values used:

N	4	8	16	30	100	1000
E[M_N]	1.0296	1.4236	1.7660	2.0428	2.5074	3.2416

E[M_N] additively transfers to the maximum under any uniform per-vantage shift mu:

$$E[\max_v (\mu + Z_v)] = \mu + E[M_N]. \qquad \text{(Sec 2.3)}$$

The omission of this additive baseline in the v0 derivation (Section 1.4) is the cause of the corrigendum.

2.4. First-expected-crossing time

For a detector f with statistic T_f(t) and threshold q_f, the FIRST-EXPECTED-CROSSING TIME tau_E(f) is the smallest t >= t_0 at which E[T_f(t) | event(t_0)] >= q_f. Under (G-lin) or (G-exp), tau_E(f) is finite and unique. It is the leading-order approximation to the actual stopping time of the random alarm process; finite-sample first-passage corrections shift the empirical mean from tau_E by an amount that is characterised empirically in Section 5.5.

3. Mathematical Foundation

3.1. Lemma L_ZD.1′ (Linear-growth lead-time, corrected)

PRECONDITION.

(P1) Null observations are IID Gaussian: $x_v(t) \sim \text{Normal}(0, \sigma^2)$, independent in (v, t).

(P2) Event signal is rank-1: $\mu(t) = s(t) \cdot u$, $\|u\| = 1$.

(P3) Growth: $s(t) = r \cdot (t - t_0)$, $r > 0$.

(P4) Both detectors are MATCHED-FAR-calibrated at alpha in $(0, 1/2)$: $q_z = \Phi^{-1}(1 - \alpha/(2N))$, $q_{chi} = F^{-1}_{\chi^2_N}(1 - \alpha)$.

(P5) $u_{max} := \max_v |u_v|$ in $(0, 1]$.

STATEMENT. Define the ZERO-DAY SIGNAL THRESHOLDS

$$s_M^*(N, \alpha) = \sigma \sqrt{q_{chi}(N, \alpha) - N}$$
$$s_z^{*\prime}(N, \alpha, u) = \sigma \, (q_z(N, \alpha) - E[M_N]) \,/\, u_{max}.$$

The first-expected-crossing times are

$$\tau_E(f_M) = t_0 + s_M^*(N, \alpha) / r,$$
$$\tau_E(f_z) = t_0 + s_z^{*\prime}(N, \alpha, u) / r,$$

and the EXPECTED LEAD-TIME is

$$E[L_{lin}] = (s_z^{*\prime}(N, \alpha, u) - s_M^*(N, \alpha)) / r.$$

E[L_lin] is STRICTLY POSITIVE iff $s_z^{*\prime} > s_M^*$, which under the coherent direction $u = 1_N / \sqrt{N}$ ($u_{max} = 1/\sqrt{N}$) holds for all $N \geq 4$ and $\alpha \leq 0.05$ in the matched-FAR setup (verified

numerically in Section 5.1).

PROOF. See docs/MVPS_ZERODAY_LEAD_TIME_LEMMA.txt Section 2 for the five-step proof from the non-central chi^2 expectation, the Bonferroni-matched z threshold, the additive E[M_N] baseline of the max-of-N statistic, and the closed-form inversion of (G-lin). QED.

REMARK 3.1.1 (Coherent vs sparse). Under u = 1_N / sqrt(N), u_max = 1/sqrt(N), so s_z*' = sigma * sqrt(N) * (q_z - E[M_N]). Under u = e_v* (sparse), u_max = 1, but s_z* uses sigma * q_z without the E[M_N] subtraction because the signal is concentrated on a single vantage and the null-max baseline does not transfer; see Section 3.3 Remark 4.1 of the lemma document.

3.2. Lemma L_ZD.2' (Exponential-growth lead-time, corrected)

Under (P1), (P2), (P4), (P5) of Section 3.1 with (P3) replaced by

(P3') s(t) = s_inf * exp(lambda * (t - t_0)),
s_inf > 0, lambda > 0,

the first-expected-crossing times are

tau_E(f_M) = t_0 + (1/lambda) * ln(s_M*(N, alpha) / s_inf),
tau_E(f_z) = t_0 + (1/lambda) * ln(s_z*'(N, alpha, u) / s_inf),

and the EXPECTED LEAD-TIME is

E[L_exp] = (1/lambda) * ln(s_z*'(N, alpha, u) / s_M*(N, alpha)). (1)

Under the COHERENT direction u = 1_N / sqrt(N):

E[L_exp] = (1/lambda) * ln(sqrt(N) * (q_z(N, alpha) - E[M_N])
/ sqrt(q_chi(N, alpha) - N)). (2)

PROFILE in N (no clean closed asymptotic).
The ratio inside ln(.) of (2) grows sub-logarithmically in N.
Numerical profile of ln(ratio) at alpha = 0.01 (Section 5.1):

N	4	8	16	30	100	1000
ln(ratio)	0.260	0.377	0.502	0.618	0.844	1.292

The v0 PROFILE "E[L_exp] ~ ln(N) / (4 lambda) as N -> infinity" is REVOKED per Section 1.4; the corrected leading order in N is smaller because (q_z - E[M_N]) grows much more slowly in N than q_z alone.

PROOF. (1) follows by inversion of (G-exp); (2) substitutes the corrected matched-FAR thresholds. See the lemma document Section 3 for the full derivation. QED.

NUMERICAL ANCHOR. Slammer-style propagation (T_d = 8.5 s, N = 30, alpha = 0.01):

lambda = ln(2) / 8.5 = 0.08155 s^{-1}
q_z(30, 0.01) = Phi^{-1}(1 - 0.005/30) = 3.5879
q_chi(30, 0.99) = 50.8922
E[M_30] = 2.0428
s_M* / sigma = sqrt(50.8922 - 30) = 4.5708
s_z*' / sigma = sqrt(30) * (3.5879 - 2.0428) = 8.4767
ratio = 8.4767 / 4.5708 = 1.8545
E[L_exp] = ln(1.8545) / 0.08155 = 7.57 s.

MC empirical mean lead at this configuration: 4.96 ticks (Section 5.5). Closed-form upper bound is within +-40 % of MC empirical at this (N, T_d).

3.3. Lemma L_ZD.3 (Sparse-direction sign reversal)

Under (P1), (P2), (P4) of Section 3.1, with u = e_v* (sparse; u_max = 1), the thresholds become

$$s_z^*(N, \alpha) = \sigma \cdot q_z(N, \alpha)$$
$$s_M^*(N, \alpha) = \sigma \cdot \sqrt{ q_{\chi}(N, \alpha) - N }$$

and the sign of the expected lead-time reverses:

$$s_z^* < s_M^* \quad \text{for all } N \geq N_0(\alpha), \qquad \text{(SIGN-REV)}$$

where $N_0(\alpha)$ is the boundary at which $q_z(N, \alpha)$ drops below $\sqrt{q_{\chi}(N, \alpha) - N}$. Numerical table:

alpha	0.001	0.005	0.010	0.025	0.050	0.100
N_0(alpha)	3	4	4	6	8	16

CONSEQUENCE. On a data set whose underlying signal direction is predominantly SPARSE, MVPS does NOT lead. The empirical RIPE Atlas observation Lambda = 23.3 % with mean lead -230 s reported in [I-D.melegassi-ippm-mvps-lead-time] is CONSISTENT WITH this regime and DOES NOT CONTRADICT L_ZD.1′ / L_ZD.2′. Empirical evaluation of T_ZD* (Section 6) must therefore curate a corpus whose signal direction is COHERENT (rank-low), not sparse.

PROOF. Substitution into the L_ZD.1′ thresholds with u_max = 1 and s_z* = sigma * q_z (per Remark 4.1 of the lemma document: the null-max baseline does not additively transfer when the signal is concentrated on a single coordinate). Sign check by enumeration; smallest N where reversal first holds tabulated above. QED.

3.4. Out-of-scope claims (explicit)

- OS-ZD-1. Code-level vulnerability detection (fuzzing, static analysis, symbolic execution, formal verification). MVPS reads network telemetry only.
- OS-ZD-2. Identification of the responsible CVE / IoC fingerprint.
- OS-ZD-3. Lead-time on post-propagation phases (steady-state worms, saturated DDoS). By (ZD-3) we require monotone growth.
- OS-ZD-4. Adversarial signal shaping using SPARSE directions to evade D^2. By L_ZD.3 such an adversary defeats the multi-vantage advantage; the M-multiplier defence of [I-D.melegassi-coherence-bfd] is the relevant mitigation.
- OS-ZD-5. An exact (rather than first-EXPECTED-crossing) stopping-time density for the chi^2 / max-Z alarm processes under monotone drift. Identified as future work.

4. Calibration and Threshold Convention

4.1. Matched FAR (Bonferroni-coordinated)

The RECOMMENDED calibration MATCHES the per-step false-alarm rate of both detectors to a common nominal alpha:

```
q_chi(N, alpha)  :=  F^{-1}_{chi^2_N}( 1 - alpha )
q_z  (N, alpha)  :=  Phi^{-1}( 1 - alpha / (2 N) )
```

so that Pr[f_M = 1 | null] = alpha exactly, and Pr[f_z = 1 | null] <= alpha by the union bound.

This is the convention under which the closed forms (1) and (2) of Section 3.2 hold without further FAR-mismatch correction.

4.2. Unmatched q_z = 3.0 (IPPM convention)

The IPPM convention used by [I-D.melegassi-ippm-mvps-lead-time] and the lab benchmark [I-D.melegassi-coherence-bfd] keeps q_z = 3.0 fixed independent of N. This is UNMATCHED FAR.

Under unmatched FAR the closed forms (1) and (2) still hold with s_z*' = sigma * sqrt(N) * (3.0 - E[M_N]) substituted; numerical comparison at Section 5.4.

IMPORTANT v1 NOTE. With unmatched q_z = 3.0, the IPPM-convention max-z detector LOSES the lead-time advantage at N >= 100 because (3.0 - E[M_N]) becomes <= 0 (E[M_100] = 2.5; E[M_1000] = 3.24). Operators using the IPPM-convention threshold MUST switch to matched-FAR q_z when N >= 100, or lose the lead-time advantage entirely. This is invisible in the v0 derivation and is a significant operational consequence of the v1 correction.

5. Numerical Receipts at Finite N

All numbers in this section are computed by the validator scripts/validate_zeroday_lead_time.py and pinned by SHA-256 in evidence/zeroday_lead_time_receipt.json.

5.1. Coherent matched-FAR thresholds (corrected, alpha = 0.01)

N	q_z	q_chi	E[M_N]	s_M*/sig	s_z*'/sig	ratio	ln(ratio)
4	3.0233	13.2767	1.0491	3.0458	3.9484	1.2964	0.2596
8	3.2272	20.0902	1.4342	3.4771	5.0714	1.4585	0.3774
16	3.4205	31.9999	1.7688	4.0000	6.6068	1.6517	0.5018
30	3.5879	50.8922	2.0403	4.5708	8.4767	1.8545	0.6176
100	3.8906	135.8067	2.4986	5.9839	13.9200	2.3263	0.8443
1000	4.4172	1106.9690	3.2273	10.3426	37.6274	3.6381	1.2915

Comparison to v0 (WITHDRAWN per Section 1.4): v0 ratio at N=30 was 4.2994; v1 corrected ratio is 1.8545.

5.2. Worm-doubling lead-times (corrected, seconds)

Event class	T_d	N=4	N=8	N=16	N=30	N=100	N=1000
Slammer (2003)	8.5 s	3.18	4.63	6.15	7.57	10.35	15.84
Code Red (2001)	37 min	831.32	1208.80	1607.18	1978.16	2703.98	4136.28
WannaCry (2017)	120 s	44.94	65.34	86.87	106.93	146.16	223.58
Memcached amp	15 s	5.62	8.17	10.86	13.37	18.27	27.95
Mirai scan	30 s	11.23	16.34	21.72	26.73	36.54	55.90

These are CLOSED-FORM UPPER BOUNDS (first-EXPECTED-crossing). Empirical mean leads at the corresponding (N, T_d) in the MC backtest of Section 5.5 are SMALLER by a factor that grows as

T_d increases (worm slower than ~30 s gives empirical mean < 50 % of closed form).

5.3. Sparse sign-reversal table (L_ZD.3)

N	s_M*/sigma	s_z*/sigma (sparse)	sz - sM	sign(L)
4	3.0458	3.0233	-0.0224	L < 0
8	3.4771	3.2272	-0.2499	L < 0
16	4.0000	3.4205	-0.5795	L < 0
30	4.5708	3.5879	-0.9829	L < 0
100	5.9839	3.8906	-2.0933	L < 0
1000	10.3426	4.4172	-5.9254	L < 0

5.4. Unmatched q_z = 3.0 variant (corrected, coherent direction)

N	E[M_N]	s_M*/sigma	s_z*'/sigma (q=3)	ratio	ln(ratio)
4	1.0491	3.0458	3.9017	1.2810	0.2477
8	1.4342	3.4771	4.4288	1.2737	0.2419
16	1.7688	4.0000	4.9247	1.2312	0.2080
30	2.0403	4.5708	5.2566	1.1500	0.1398
100	2.4986	5.9839	5.0141	< 1.00	< 0
1000	3.2273	10.3426	0.0000	< 1.00	< 0

5.5. Monte Carlo empirical validation

Method. For each (N, T_d) in the 9-configuration panel, run K = 500 independent Monte Carlo trials. Each trial:

(i) simulates IID Gaussian baseline noise X[T_history + T_det, N] with T_history = 2000, T_det adapted per config to max(500, 8.66 * T_d) ticks (so that lambda * T_det >= 6, ensuring the signal grows by >= exp(6) ~ 400x within the detection window);

(ii) injects a rank-1 coherent signal at t_inject = T_history in direction u = 1_N/sqrt(N) with exponential growth lambda = ln(2)/T_d and amplitude s_inf = 0.5;

(iii) calibrates q_chi and q_z EMPIRICALLY at the 99-percentile of D^2 and max-|z| computed over the clean holdout window [0, T_history) -- BLIND to any data after injection;

(iv) records t_M = first crossing of D^2 over q_chi in the detection window, and t_Z = first crossing of max-|z| over q_z, then lead = t_Z - t_M;

(v) aggregates over K trials: Lambda_emp = fraction with lead > 0; Wilson 95 % CI; mean/median/p25/p75 of lead; relative error against the L_ZD.2' closed-form prediction.

Results (alpha = 0.01, K = 500 trials per config):

config	N	T_d (s)	Lambda	Wilson 95% CI	mean	median	theory	rel.err	verdict
Slammer-class, N=4	4	8.5	0.402	[0.360, 0.446]	1.26	0.0	3.18	0.603	BELOW_THEORY
Slammer-class, N=8	8	8.5	0.540	[0.496, 0.583]	2.65	2.0	4.63	0.428	BELOW_THEORY
Slammer-class, N=16	16	8.5	0.620	[0.577, 0.661]	3.55	3.0	6.15	0.422	CONSISTENT_SIGN
Slammer-class, N=30	30	8.5	0.674	[0.632, 0.714]	4.96	5.0	7.57	0.346	PASS_THEORY
Slammer-class, N=100	100	8.5	0.730	[0.689, 0.767]	7.69	8.0	10.35	0.257	PASS_THEORY
Memcached-class, N=30	30	15.0	0.586	[0.542, 0.628]	5.57	4.0	13.37	0.583	CONSISTENT_SIGN
Mirai-class, N=30	30	30.0	0.566	[0.522, 0.609]	9.37	7.5	26.		

```
73    0.649   CONSISTENT_SIGN
      WannaCry-class,  N=30     30   120.0    0.472   [0.429, 0.516]    6.98   0.0   106.
93    0.935   BELOW_THEORY
      Code-Red-fast,   N=30     30   600.0    0.462   [0.419, 0.506]   14.99   0.0   534.
64    0.972   BELOW_THEORY
```

Verdict grid:

```
PASS_THEORY     (2 of 9):  Wilson_lo > 0.55  AND  rel.err <= 0.40
                           (sign + magnitude both confirmed).
CONSISTENT_SIGN (3 of 9):  Wilson_lo > 0.50; magnitude loose 40-65 %.
BELOW_THEORY    (4 of 9):  0.30 < Wilson_lo <= 0.50; weak lead,
                           closed form severely loose.
FALSIFIES       (0 of 9):  Wilson_lo <= 0.30; SIGN-CLAIM fails.
```

HEADLINE. The SIGN-CLAIM of L_ZD.2′ (positive expected lead) is empirically supported on ALL nine panel configurations (Wilson 95 % lower bound > 0.30 in every case, > 0.50 in five of nine, > 0.55 in two of nine). The MAGNITUDE-CLAIM (closed form within +-40 % of empirical) holds in 2 of 9, both at fast growth (Slammer T_d = 8.5 s) with N >= 30; loose by factor 1.4-1.6 in three additional CONSISTENT_SIGN configurations; loose by factor 5-30x in four BELOW_THEORY configurations (very small N or slow growth).

Operational reading. Fast-propagating events (T_d <= 30 s) with multi-vantage groups (N >= 30) are the operational sweet spot for MVPS zero-day-class lead-time. Slower events still give positive mean lead but the closed-form upper bound is loose; for those, operators SHOULD run the MC backtest at their specific (N, lambda) configuration to estimate realistic lead-time rather than relying on the closed-form value.

Receipt: evidence/zeroday_backtest_mc_<UTC>.json with full per-config payload, SHA-256 emitted on backtest stdout.

6. Conjecture T_ZD* and Falsification Protocol

CONJECTURE T_ZD* (Open, not yet tested on real historical data).

Let Z be a corpus of M historical, publicly-documented "propagating" network events satisfying ZD-1..ZD-4 of Section 2.1 on RIPE Atlas / RIPE RIS / Cloudflare Radar data in a window enclosing the event onset. For each event i in Z, let:

- t_IOC^(i) := first publicly available Indicator-of-Compromise timestamp.
- t_MVPS^(i) := timestamp at which D^2, computed on a holdout-calibrated RIPE Atlas / RIS / MRT archive window, first crosses q_chi(N, alpha). Holdout = [t_IOC - 14 d, t_IOC - 7 d]; detection = [t_IOC - 7 d, t_IOC + 1 d].

Lambda_ZD := | { i : t_MVPS^(i) < t_IOC^(i) } | / M.

T_ZD* (sufficiency). Lambda_ZD >= 1/3 with median observed lead at least (1/lambda_typ) * ln(ratio_typ) where ratio_typ is the L_ZD.2′ closed-form ratio at the operator's typical (N, alpha) (1.8545 at N = 30, alpha = 0.01).

STATUS. NOT YET CONFIRMED. Real-data extension of the synthetic-noise validation of Section 5.5.

6.1. Pre-registered corpus suggestion

i	event	approx t_IOC (UTC)
1	SQL Slammer worm	2003-01-25 05:30
2	Code Red v1 worm	2001-07-13 14:00
3	Code Red v2 worm	2001-07-19 22:00
4	Conficker initial wave	2008-11-21 12:00
5	Mirai (Krebs DDoS phase)	2016-09-20 02:00
6	Mirai (Dyn DNS DDoS)	2016-10-21 11:10
7	WannaCry (SMB propagation peak)	2017-05-12 07:00
8	NotPetya (initial wave)	2017-06-27 09:30
9	Memcached amplification (GitHub)	2018-02-28 17:21
10	Facebook BGP outage	2021-10-04 15:40
11	Cloudflare BGP leak (Verizon)	2019-06-24 10:30
12	Rostelecom BGP hijack (massive)	2020-04-01 16:00

6.2. Protocol P-ZD.1 .. P-ZD.6

P-ZD.1 Pre-register the corpus.

P-ZD.2 Fetch the BGP-update or RTT data covering [t_IOC - 14 d, t_IOC + 1 d] for each event from the appropriate archive (RIPE Atlas msm IDs, RIPE RIS MRT, Routeviews MRT, CAIDA BGPStream, Cloudflare Radar).

P-ZD.3 Compute D^2 on the BLIND HOLDOUT window [t_IOC - 14 d, t_IOC - 7 d] to set q_chi at the empirical 99-percentile. Calibration MUST NOT see any data later than t_IOC - 7 d.

P-ZD.4 Run D^2 forward through [t_IOC - 7 d, t_IOC + 1 d] and record t_MVPS = first time D^2 >= q_chi.

P-ZD.5 Compare to t_IOC, tabulate per-event lead, compute Lambda_ZD and Wilson 95 % CI per the convention of L_LT.A.

P-ZD.6 Apply the verdict:

FALSIFIES T_ZD*	if	Wilson 95 % CI upper bound on Lambda_ZD is below 1/3 OR if observed median lead is below the L_ZD.2′ closed-form prediction by a factor > 30. (Factor 30 chosen to match the BELOW_THEORY band of Section 5.5 MC backtest; tighter falsification thresholds are inappropriate given the closed-form is a known UPPER BOUND, not a tight prediction.)
CONSISTENT	if	Lambda_ZD >= 1/3 with Wilson lower bound > 0 but median lead is below the closed-form by 5-30x.
SUPPORTS T_ZD*	if	Lambda_ZD >= 1/3 with Wilson lower bound > 1/4 AND observed median lead matches the closed-form within a factor of 2.

6.3. Data-coverage gap (RIPE Stat smoke test)

A smoke test performed on 2026-05-25 (scripts/_smoke_ripestat_historical.py) confirmed that the free

RIPE Stat bgp-updates endpoint returns ZERO historical records
for the following events in the 2018-2021 window:

```
Facebook BGP outage 2021-10-04   (prefix 157.240.0.0/16): 0 updates
CF/Verizon leak    2019-06-24   (prefix 1.1.1.0/24):      0 updates
Rostelecom hijack  2020-04-01   (prefix 8.8.8.0/24):      0 updates
Memcached/GitHub   2018-02-28   (prefix 140.82.112.0/20): 0 updates
```

The same endpoint returned 3 records for the recent control
window 2026-05-24, confirming the endpoint is reachable but does
not retain history beyond a recent window.

IMPLICATION. Empirical execution of T_ZD* requires either
(a) RIPE RIS or Routeviews MRT-archive parsing (mrtparse / pybgpstream),
(b) cached Cloudflare Radar snapshots from public blog posts, or
(c) CAIDA BGPStream / Telescope archives for pre-2018 events.

This is the principal infrastructure gap to close before T_ZD*
can be tested empirically. It is identified as future work.

7. What This Profile Does NOT Claim

- MVPS does NOT find zero-day vulnerabilities in code. It finds the network-visible PROPAGATION SIGNATURE of an exploitation that produces coherent multi-vantage telemetry deviations. Single-host privilege escalations, passive memory leaks, cryptographic side-channels, and backdoors dormant before the first network beacon are INVISIBLE to MVPS by construction.

- MVPS does NOT name the responsible CVE or attack family.

- The closed-form lead-times of Sections 3.1 and 3.2 are FIRST-EXPECTED-CROSSING UPPER BOUNDS. Empirical MC (Section 5.5) shows the SIGN-CLAIM holds on the entire 9-configuration panel, but the MAGNITUDE-CLAIM (closed form tight within +-40 %) holds only on 2 of 9; on the remaining 7 the closed form overpredicts by factors of 1.4x (CONSISTENT_SIGN) to 30x (BELOW_THEORY, slow growth).

- Conjecture T_ZD* (Section 6) is NOT a theorem. It is the real-data extension of the synthetic-noise validation of Section 5.5 and depends on MRT-archive parsing infrastructure not yet in place (Section 6.3).

- Lemma L_ZD.3 (Section 3.3) PROVES that MVPS LOSES the lead-time race in the SPARSE-DIRECTION regime. Operators with predominantly local-jitter alarms SHOULD use a per-vantage detector, not MVPS.

- The Slammer / Code Red / WannaCry / Memcached / Mirai lead-time numbers in Section 5.2 are PREDICTIONS of the closed-form L_ZD.2’ evaluated at typical doubling times, NOT measurements on the corresponding actual events.

8. Operational Recommendations

This profile is RECOMMENDED to be deployed when:

- N >= 30 vantages observe coherent multi-AS or multi-prefix telemetry.

- The operational concern includes FAST-propagation events (doubling time T_d <= 30 s): mass scanning worms, novel DDoS amplification vectors, mass BGP misorigination.

- An empirical MC backtest per Section 5.5 has been performed at the operator's specific (N, lambda) configuration and the Wilson lower bound on Lambda_emp exceeds 0.5.

This profile is NOT RECOMMENDED when:

- N < 16 vantages are available (lead-time advantage is small even when present).

- Operational concern is dominated by single-vantage local jitter or last-mile microcuts (sparse-direction regime of L_ZD.3).

- Operational concern is dominated by slow-propagation events (T_d > 120 s) where the closed-form lead-time is severely loose; for these the closed form should not be used for operational planning.

- Code-level vulnerability discovery is the goal; MVPS operates strictly in the network-telemetry domain.

9. Reproducibility

All artefacts are public:

Lemma document and proofs:
docs/MVPS_ZERODAY_LEAD_TIME_LEMMA.txt

Numerical receipt (closed-form table, re-verified to 1e-6):
evidence/zeroday_lead_time_receipt.json
schema = com.catellix.mvps.zeroday_lead_time_receipt_v1

Monte Carlo empirical receipt (K = 500 trials per config):
evidence/zeroday_backtest_mc_<UTC>.json
schema = com.catellix.mvps.zeroday_backtest_mc_v1

Validator script:
scripts/validate_zeroday_lead_time.py

MC backtest script:
scripts/backtest_zeroday_mc.py

Historical-coverage smoke test:
scripts/_smoke_ripestat_historical.py

E[M_N] computation:
scripts/_compute_emax_table.py
(5e6 MC samples per N; Blom approximation verified to <2 %)

Catellix evidence page:
https://catellix.com/v11-evidence.html

Both receipts carry SHA-256 hashes emitted on script stdout and pinned in the v11 evidence manifest.

10. Security Considerations

This document defines no new wire formats and no new cryptographic

primitives. Two profile-specific considerations:

- ADVERSARIAL SPARSIFICATION (L_ZD.3). An adversary aware of the multi-vantage detector can deliberately shape the signal direction to be SPARSE and thereby defeat the lead-time advantage. The M-multiplier defence of [I-D.melegassi-coherence-bfd] Section 9.2 plus per-vantage cross-checks limit this evasion path.

- CORPUS POISONING. An adversary capable of injecting spurious holdout-window traffic that matches their later attack signature can degrade D^2's calibration. The BLIND holdout discipline of P-ZD.3 (calibration sees no data later than t_IOC - 7 d) limits but does not eliminate this risk; longer holdout windows (>= 7 d) and rolling per-week recalibration are RECOMMENDED.

11. IANA Considerations

This document has no IANA actions.

12. Privacy Considerations

The RIPE Atlas measurements used in the empirical conjecture protocol (Section 6) are public-target, public-probe measurements; no user-identifiable information is exposed.

13. References

13.1. Normative References

[I-D.melegassi-ippm-mvps-bundle]
Melegassi, L., "Multi-Vantage Path Synchrony Bundle Envelope and Vector Algebra", draft-melegassi-ippm-mvps-bundle-00, May 2026.

[I-D.melegassi-ippm-mvps-lead-time]
Melegassi, L., "Empirical Lead-Time Profile for Multi-Vantage Path Synchrony (MVPS): The T_LT Profile", draft-melegassi-ippm-mvps-lead-time-00, May 2026.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, May 2017.

13.2. Informative References

[I-D.melegassi-coherence-bfd]
Melegassi, L., "Coherence-BFD: Sub-Tick Coherence Detection over BFD Mechanisms", draft-melegassi-coherence-bfd-00, May 2026.

[I-D.melegassi-mvps-ddos-resilience]
Melegassi, L., "Volume-Independent DDoS Detection via Coherence-BFD: The MVPS DDoS Resilience Profile", draft-melegassi-mvps-ddos-resilience-00, May 2026.

[BLOM-1958]
Blom, G., "Statistical Estimates and Transformed Beta-Variables", Wiley, 1958, Chapter 5 (Blom approximation for E[M_N]).

[DAVID-NAGARAJA]
David, H. A. and H. N. Nagaraja, "Order Statistics", 3rd ed., Wiley, 2003, Section 4.4 (expected value of the maximum of N iid Normal variables).

[JOHNSON-KOTZ-BALAKRISHNAN]
Johnson, N. L., Kotz, S., and N. Balakrishnan, "Continuous Univariate Distributions", Volume 1, 2nd ed., Wiley, 1994, Chapter 18 (chi^2 tail expansions used in Lemma L_ZD.2').

[LEHMANN-ROMANO]
Lehmann, E. L. and J. P. Romano, "Testing Statistical Hypotheses", 3rd ed., Springer, 2005, Section 9.1 (Bonferroni-coordinated multiple-test thresholds).

[RIPE-ATLAS]
RIPE NCC, "RIPE Atlas", https://atlas.ripe.net/.

[RIPE-STAT]
RIPE NCC, "RIPE Stat", https://stat.ripe.net/.

[SLAMMER] Moore, D., Paxson, V., Savage, S., Shannon, C., Staniford, S., and N. Weaver, "Inside the Slammer Worm", IEEE Security & Privacy, vol. 1, no. 4, pp. 33-39, July 2003.

[WANNACRY] Symantec Security Response, "What you need to know about the WannaCry Ransomware", May 2017.

[MEMCACHED-AMP]
Cloudflare, "Memcached DDoS: The 1.7 Tbps attack against GitHub", March 2018.

[FACEBOOK-BGP]
Cloudflare, "Understanding How Facebook Disappeared from the Internet", October 2021.

Acknowledgements

The authors thank the IETF IPPM mailing list and the off-list reviewers of [I-D.melegassi-ippm-mvps-lead-time] for the honest-accounting discipline that this document inherits. The v0-to-v1 correction recorded in Section 1.4 follows the same pattern as the L_LT.A retraction of the original unconditional T_LT promise: replace the wrong claim with the conditional theorem that survives the data, retire the wrong claim explicitly, document the empirical artefact that caught the error.

The closed-form derivation of L_ZD.2' follows the chi^2 tail-expansion conventions of [JOHNSON-KOTZ-BALAKRISHNAN], the Bonferroni multiple-test framework of [LEHMANN-ROMANO], and the order-statistics tradition of [DAVID-NAGARAJA] and [BLOM-1958].

Author’s Address

Leonardo Melegassi
Catellix
Andradina, SP
Brazil

Email: melegassi@catellix.com
URI: https://catellix.com/