RAMTERA TECHNICAL WHITE PAPER SERIES

WP-04

The Thousand-Violation Ceiling

Why result caps silently corrupt every cross-tool comparison

A truncated golden reference does not tell you it has been truncated. It just returns a number that looks like an answer.

1,000	100%	0%	Exclude
TYPICAL RESULT CAP	APPARENT RECALL	ACTUAL INFORMATION	CORRECT TREATMENT

Document	WP-04 — The Thousand-Violation Ceiling
Author	Sameer Pawanekar, Founder & Chief Executive Officer, Ramtera Inc.
Revision	v1.0
Date	August 2026
Classification	Public

ABSTRACT

Physical verification tools cap the number of results they will record for a single rule, typically at a round number such as one thousand. The cap is a sensible engineering decision and an entirely reasonable default. It is also the single most common way a tool-to-tool correlation study produces figures that are arithmetically correct and analytically meaningless. This paper explains the failure, shows why both the obvious treatments — counting the excess as false positives, or counting it as a match — are wrong, and sets out the correct handling.

01 What a result cap does

A design rule check that finds a systematic violation can find it a very large number of times. A single mistaken layer derivation can produce hundreds of thousands of markers, most of which are the same defect reported repeatedly.

Recording all of them serves nobody: the result database becomes enormous, the viewer becomes unusable, and the engineer needs the first few to diagnose the problem. Every serious verification tool therefore caps results per rule, and the default is often one thousand.

This is correct behaviour for the purpose it was designed for. The difficulty arises when the capped output is later used as a reference for comparison against another tool — which is precisely what a correlation or certification study does.

> **THE CENTRAL POINT**
> The essential property of a capped rule is that its recorded count is not a measurement of anything. It is the cap. A rule reporting exactly 1,000 violations tells you the true count is at least 1,000 and provides no information about what it actually is.

02 How the failure presents

Consider a correlation study comparing a tool under evaluation against a golden reference. For a capped rule, the reference records 1,000 markers. The tool under evaluation, run without a cap, records 47,000.

The apparent result

Metric	Naive computation	Apparent figure
Recall	1,000 reference markers, all covered by the tool	**100%**
Precision	1,000 matched of 47,000 reported	**2.1%**
Conclusion drawn	“Perfect recall, catastrophic false-positive rate”	**Entirely wrong**

Table 1 — A capped rule scored naively.

Both figures are arithmetically correct. Both are meaningless. The 46,000 unmatched results are not false positives: they are results for which the reference contains no information at all, because the reference stopped recording at one thousand. They have not been shown to be wrong. They have not been shown to be right. They are unverifiable.

The direction of the error

Note that the error is not symmetric. Capping inflates recall toward 100% and deflates precision toward zero. A study that reports pooled precision across a deck containing a handful of capped rules will show a precision figure dominated entirely by those few rules, regardless of how the other several hundred behaved.

> **WHAT TO DEMAND OF ANY REPORTED PRECISION FIGURE**
> This is why a pooled precision figure across a rule deck should always be accompanied by the number of capped rules it contains. Without that, the reader cannot tell whether they are looking at a measurement of the tool or a measurement of the cap.

03 Why both obvious treatments are wrong

Treatment one: count the excess as false positives

This is the default behaviour of a comparison script that has not considered caps, and it is wrong because it asserts something the data does not support. A false positive is a result shown to be incorrect. These results have not been examined; the reference simply stopped looking. Charging them as errors penalises the tool for the reference's truncation.

Treatment two: count the excess as correct

The opposite error, and equally unsupported. The excess results might be entirely correct, or might contain genuine false positives in unknown proportion. Assuming they are correct converts an unknown into a favourable known, which is the more dangerous of the two mistakes because it flatters the tool.

The correct treatment: exclude

Unmatched results on a capped rule are removed from the precision denominator entirely. They are neither correct nor incorrect; they are unmeasured, and a metric should not incorporate unmeasured quantities in either direction.

Quantity	Treatment on a capped rule
Reference markers matched by the tool	Counted normally. Recall on a capped rule is sound and quotable.
Reference markers not matched by the tool	Counted normally as recall failures. A miss is a miss regardless of the cap.
Tool results beyond the cap, unmatched	**Excluded from the precision denominator.** Unverifiable, not false.
Pooled precision across the deck	Computed with capped-rule excess excluded; number of capped rules stated alongside.
Pooled recall across the deck	Computed with the capped mass excluded, so that a few capped rules do not dominate.

Table 2 — Correct handling of each quantity on a capped rule.

> **WHAT SURVIVES THE CAP**
> Recall on a capped rule remains sound, which is worth stating clearly because it is the part people tend to discard along with the rest. If the reference recorded 1,000 markers and the tool covers all 1,000, that is a genuine and quotable result. The cap damages precision, not recall.

04 The partial case, which is the one that catches people

The clean case is a rule at exactly the cap. The awkward case is a rule *near* it, where the reference recorded, say, 1,000 markers of which 980 were matched, and the tool reported 1,400.

It is tempting to treat the 400 excess as excluded and the 20 misses as recall failures, which is correct, and then to treat a portion of the excess as false positives on the grounds that the rule was only partially capped. That last step is where phantom false positives are manufactured. There is no basis for apportioning the excess, because there is no information about it.

THE PARTIAL-CAP RULE

Once a rule has hit its cap, the entire unmatched excess is excluded. There is no partial credit and no partial penalty, because partiality would require knowing something about results the reference never recorded.

05 Practical recommendations

1. **Detect caps automatically.** A rule whose recorded count equals the configured maximum exactly should be flagged by the comparison tooling, not spotted by eye. Round numbers in a violation count are almost never natural.
2. **Re-run the reference uncapped where possible.** This is the only treatment that recovers the lost information. Where the reference is supplied by a third party, requesting an uncapped run for the affected rules is usually a small ask and removes the problem entirely.
3. **Report capped rules separately.** A correlation table should show capped rules as a distinct category with their own count, not folded into a pooled figure.
4. **State the cap value.** A correlation study that does not state the reference tool's result cap has omitted a parameter that materially affects every number in it.
5. **Separate measurement artifacts from genuine mismatches.** Caps are one of several classes of apparent disagreement that are not tool defects. Rule-deck version drift between the reference and the tool under evaluation is another. Both belong in a separate category from real errors.

06 Conclusion

Result caps are good engineering and a poor foundation for measurement. The failure they cause is particularly awkward because it produces numbers rather than errors: nothing crashes, no warning is issued, and the resulting figures are arithmetically defensible. The only defence is to treat a capped rule as what it is — a rule for which the reference has stopped providing information — and to exclude the unmeasured portion from the metrics rather than assigning it a value in either direction.

THE GENERAL PRINCIPLE

The general principle, of which this is one instance: when a reference stops measuring, the honest response is to narrow the claim, not to fill the gap with an assumption. An assumption that flatters the tool is worse than one that penalises it, because nobody goes looking for it.

About Ramtera

Ramtera Inc. develops **LVR**, a physical verification engine that executes foundry SVRF rule decks natively, covering design rule checking, layout versus schematic, electrical rule checking, layout XOR, density and fill, and parasitic extraction. LVR is written in modern C++ with a Qt6 interface, and has been validated against publicly available process design kits including SKY130 and IHP SG13G2.

Ramtera also develops **PRE**, a place-and-route engine sharing LVR's geometry kernel, and maintains a portfolio of granted patents, registered copyrights and trademarks across its verification and implementation technologies.

Ramtera Inc. is incorporated in the United States, with research and development conducted by Ramtera Design Automation India Private Limited in Bangalore, India.

Contact	
Web	https://ramtera.com
General enquiries	info@ramtera.com
Technical support	support@ramtera.com
Series	Ramtera Technical White Paper Series
This document	WP-04, revision v1.0, August 2026