Google Technology Holdings LLC (Google) appeals the decision of the Patent Trial and Appeal Board (Board) sustaining the examiner’s final rejection of claims 1–9, 11,
Because we conclude that Google has forfeited its arguments put forth on appeal, we affirm.

BACKGROUND

The ’765 application relates to “distributed caching for video-on-demand systems, and in particular to a method and apparatus for transferring content within such video-on-demand systems.” J.A. 51. The application describes a three-level system that includes (i) a complete content library stored at a remote, master server referred to as a video server office (VSO), (ii) several smaller, partial-content libraries at local servers referred to as video home offices (VHOs), and (iii) individual users at set-top boxes (STBs). See id. at 44–45, 51–53. Different VHOs will store at least some different content so that when a particular VHO local to an STB lacks certain requested content, the system may stream that requested content from either a different VHO or the VSO. See id. at 54. The proposed invention presents a solution for determining how to stream content to STBs and where to store said content among the content servers, be it at one or more of the VHOs or at the VSO. See id. at 54–55.

Claim 1 describes a method of responding to requests to stream content to STBs from various content servers, i.e., VHOs and the VSO, based on the relative costs of the network impact of fetching the content from the various servers, which have the requested content. Claim 2 depends from claim 1 and further describes determining at which particular server/s to store the content depending on “a network penalty.”

1 All references to § 103 throughout the opinion refer to the pre-AIA version because the application’s priority date is before March 16, 2013.
Claims 1 and 2 of the ’765 application are representative. They recite in relevant part:

1. A method comprising:
   receiving, by a processing apparatus at a first content source, a request for content;
   in response to receiving the request, determining that the content is not available from the first content source;
   in response to determining that the content is not available from the first content source, determining that a second content source cost associated with retrieving the content from a second content source is less than a third content source cost associated with retrieving the content from a third content source, wherein the second content source cost is determined based on a network impact to fetch the content from the second content source to the first content source, . . .

2. The method of claim 1, further comprising:
   determining that there is not sufficient memory to cache the content at the first content source; and
   selecting one or more items to evict from a cache at the first content source to make available sufficient memory for the content, wherein the selection of the items to evict minimizes a network penalty associated with the eviction of the items, wherein the network penalty is based on sizes of the content and the items, and numbers of requests expected to be received for the content and the items.

’765 application at claims 1 and 2 (emphases added).
The examiner rejected the claims under § 103. The examiner determined that claim 1 would have been obvious in view of Costa and Scholl and claim 2 would have been obvious in view of Costa, Scholl, Allegrezza, and Ryu. The examiner found that Costa discloses all limitations of claim 1 except one, and Scholl discloses this remaining limitation, i.e., determination of the content route. The examiner also explained that Allegrezza and Ryu disclose the additional limitations in claim 2, because Allegrezza teaches caching content and minimizing a network penalty based on the number of requests expected to be received for the content and the items and Ryu teaches minimizing a network penalty based on the size of the content and the items.

Google appealed this final rejection to the Board. As to claim 1, Google contended that “the cited references do not teach or suggest [most of the claim limitations].” Largely quoting the claim language and references, Google suggested that Costa teaches only “randomly distributing a first set of videos for storage, distributing a second set of videos to all of the central offices for storage locally, and distributing a third set of videos to a particular central office based on at least one demographic of viewers,” and Scholl teaches only a “determination according to some rules,” (internal quotation marks omitted). Google similarly averred that the cited

2 The claims were also initially rejected for double patenting, but that rejection was resolved with Google’s filing of a terminal disclaimer. See J.A. 156.
4 U.S. Patent No. 8,023,319.
6 U.S. Patent No. 8,087,056.
references did not teach the limitations of claim 2. *Id.* at 211. Again, relying on block quotes from the claim language and the references, Google asserted that Allegrezza merely teaches the storage of a content file through comparison to a demand threshold, *id.* at 212, and Ryu “merely mentions classifying stored content into high-capacity and low-capacity content,” *id.* at 213.

The Board affirmed the examiner’s obviousness rejections. The Board interpreted Google’s arguments for claim 1 as suggesting that Costa and Scholl do not teach “distributing content based on ‘costs’ which are ‘based on a network impact.’” *J.A.* 7 (quoting claim 1). The Board was unpersuaded, finding the examiner’s broad reading of “costs,” as disclosed by Costa and Scholl, to be entirely consistent with the application’s specification. *Id.* at 8–10. Utilizing the broadest reasonable interpretation standard, the Board explained that “costs” “based on a network impact” encompassed costs “in terms of load, storage space, time/latency, predicted traffic (i.e., amount of data transmitted over time), or monetary costs.” *Id.* at 10 (citing the application’s written description). The Board, agreeing with the examiner, then found that Costa teaches costs based on “redundancy, distance, or congestion,” and Scholl teaches “costs based on proximity, time, and/or number of nodes,” *id.* at 9, and therefore, the “combination of Costa and Scholl teaches or suggests the method of distributed content management and fetching recited in claim 1,” *id.* Of note, the Board explained that “[Google] ha[d] not cited to a definition of ‘costs’ or ‘network impact’ in the [s]pecification that would preclude the [s]pecification that would preclude the [e]xaminer’s broader reading.” *Id.* at 10.

The Board then assessed Google’s arguments for claim 2, and again, agreed with the examiner. The Board concluded that the collective teachings of the combination of Costa, Scholl, Allegrezza, and Ryu disclose all the limitations of claim 2. *J.A.* 11–12. The Board summarized Google’s arguments as unpersuasive inasmuch as they
suggested only “the shortcomings in the teachings of Alle-
grezza and Ryu individually,” because the examiner “re-
lie[d] on a properly made combination of [all the
references].”  *Id.* at 11. Particularly, the Board explained
that “[Google’s] response to the [e]xaminer’s findings in
this regard [w]ere conclusory, fail[ed] to specifically ad-
dress the [e]xaminer’s findings . . . , and [w]ere thus unper-
suasive.”  *Id.* at 12. The Board concluded that Google failed
to rebut the collective teachings and suggestions of the ap-
plied references.  *Id.* Of consequence, the Board did not put
forth any explicit construction of the term “network pen-
alty” in its decision, nor did Google argue for one.  *Id.* at
10–13.

Google appeals to our court from the Board’s decision.
We have jurisdiction pursuant to 28 U.S.C. § 1295(a)(4)(A)

DISCUSSION

As an initial matter, we recognize that our court’s opin-
ions have not always been precise when discussing the doc-
trines of waiver and forfeiture. The court seemingly has
used the terms interchangeably at times.  *Compare In re
Watts,* 354 F.3d 1362, 1368 (Fed. Cir. 2004) (“Because the
appellant failed to argue his current interpretation of the
prior art below . . . we hold that appellant has waived [this
argument].”), and *In re Baxter Int’l, Inc.,* 678 F.3d 1357,
1362 (Fed. Cir. 2012) (“We agree with the PTO that Baxter
waived its arguments regarding the [means for] limitation
in claim 30 by failing to timely raise them before the
Board.”),  with *Nuvo Pharms. (Ireland) Designated Activity
Co. v. Dr. Reddy’s Labs. Inc.*, 923 F.3d 1368, 1378 (Fed. Cir.
2019) (concluding that “Nuvo’s argument was not raised
below and thus is forfeited” and citing in support of this
proposition  *TVIIM, LLC v. McAfee, Inc.*, 851 F.3d 1356,
1363 (Fed. Cir. 2017), which utilized the term “waiver”).

It is well established that “[w]aiver is different from
forfeiture.”  *United States v. Olano,* 507 U.S. 725, 733
“Whereas forfeiture is the failure to make the timely assertion of a right, waiver is the ‘intentional relinquishment or abandonment of a known right.’” Id. (quoting Johnson v. Zerbst, 304 U.S. 458, 464 (1938)) (additional citations omitted). The two scenarios can have different consequences for challenges raised on appeal, id. at 733–34, and for that reason, it is worth attending to which label is the right one in a particular case.

By and large, in reviewing this court’s precedent, it is evident that the court mainly uses the term “waiver” when applying the doctrine of “forfeiture.” 8 The parties in the case at hand, understandably, have done just the same. Specifically, the Patent Office contends that “Google never made [its claim construction lexicography] arguments to the Board, so it waived them.” Appellee’s Br. at 18. We interpret the Patent Office to be arguing that Google’s

---

7 We are reminded of this passage from Justice Scalia in Freytag v. Commissioner:

The Court uses the term “waive” instead of “forfeit.” The two are really not the same, although our cases have so often used them interchangeably that it may be too late to introduce precision.

501 U.S. 868, 894 n.2 (1991) (Scalia, J., concurring in part and concurring in judgment) (citation omitted). Two years after this observation from Justice Scalia, in an opinion in which he joined, the Supreme Court elucidated that this distinction is of some import. See Olano, 507 U.S. at 733 (discussing the differing consequences that attach to waiver and forfeiture and citing Justice Scalia’s footnote from Freytag).

8 Though previous cases may have used the term “waiver” instead of “forfeiture,” their holdings are good law for a case, like this one, involving the issue of forfeiture.
failure to raise its lexicography arguments, inadvertent or not, compels a finding of forfeiture. We agree.

On appeal, Google posits the following: “[T]he Board err[ed] when it construed the claim terms ‘cost associated with retrieving the content’ and ‘network penalty’ in contradiction to their explicit definitions in the specification.” Appellant’s Br. at 2. More specifically, Google argues that “cost” in claim 1 is “the cost of the path . . . defined by the bottleneck link in that path, which is the link that takes the longest to transfer item I,” id. at 39 (internal quotation marks omitted) (citing paragraph 67 of the written description), and “network penalty” in claim 2 is a specific formula based on the size of the content and items, expected requests for the content and items, and fetch cost for the content and items, see id. at 42 (quoting paragraphs 79–81 of the written description). Therefore, Google argues, because the Board relied on the incorrect constructions of these two terms, the Board’s decision upholding the rejections was incorrect.

Meritorious or not, Google never presented these arguments to the Board. And therein lies the problem. Because Google failed to present these claim construction arguments to the Board, Google forfeited both arguments. We have regularly stated and applied the important principle that a position not presented in the tribunal under review will not be considered on appeal in the absence of exceptional circumstances. See, e.g., Nuvo Pharm., 923 F.3d at 1378; Golden Bridge Tech., Inc. v. Nokia, Inc., 527 F.3d 1318, 1323 (Fed. Cir. 2008). We see no exceptional circumstances justifying a departure from that principle here.

Google presents two main reasons as to why this court should exercise its discretion to hear its forfeited arguments on appeal: (1) the Board in its decision sua sponte construed the term “cost” and thus, because it “passed upon” on the issue, Google is not barred from appealing that construction; and (2) the issue of the construction of
“network penalty” is one of law fully briefed (now) and consistent with Google’s position before the Board. We are unpersuaded.

In support of its first contention, Google cites United States v. Williams for the “traditional rule” that “permit[s] review of an issue not pressed so long as it has been passed upon.” Appellant’s Reply Br. at 3 (citing United States v. Williams, 504 U.S. 36, 41 (1992)). We agree with this premise. Yet even though review may be “permit[ted],” Williams, 504 U.S. at 41 (emphasis added), the ultimate decision about whether to hear a claim when it was not raised before the lower tribunal is one that remains within the discretion of the appellate court, see id. at 44–45 (explaining that the decision to undertake review of an issue not pressed by the appellant below is a “permissible exercise of [the court’s] discretion”).

In this instance, we decline to hear Google’s new arguments as to the construction of “cost.” Google has not provided any reasonable explanation as to why it never argued to the examiner during the iterative examination process or later to the Board for a particular construction of the term “cost”—an argument that is now the linchpin to its claims’ patentability. Accord Watts, 354 F.3d at 1368 (“Watts has shown no reason why we should excuse his failure to raise this argument before the Board.”). Allowing Google to press, on appeal, a specific claim construction that it did not present to the Board deprives the Board, an expert body, of its important role in reviewing the rejection of patent applications. See id. at 1367. Moreover, Google has not explained how the Board’s understanding of “cost” altered the theory underlying the examiner’s rejection or otherwise was a surprising, unexpected departure.

Google cites Conoco, Inc. v. Energy & Environmental International, L.C., 460 F.3d 1349 (Fed. Cir. 2006), as an example of our court reviewing a district court’s sua sponte claim construction. But nothing in that opinion contradicts
the Supreme Court’s articulation of the “passed upon” doctrine as permissive, discretionary, and context-driven. Whatever other differences there may be, the context of an infringement determination after a bench trial is quite different from the context of a Board unpatentability determination after examination—the latter context is iterative. Once an examiner has made a prima facie case for rejecting the application claims, the applicant is provided with an opportunity to submit any saving claim construction it believes may be grounds for reversing the rejection both to the examiner and potentially again, to the Board. The burden lies with the applicant to present this argument in the initial instances. An applicant who does not take those opportunities and is then further disappointed by a Board claim construction should be encouraged to avoid waste of appellate resources and instead take the intra-PTO route of filing new or amended claims (perhaps through a continuation application) containing language that makes the desired scope clear, thereby serving the goal of facial clarity of patent claims.

In sum: we are, as an appellate court, charged in this instance with reviewing the Board’s conclusions. “The very word ‘review’ presupposes that a litigant’s arguments have been raised and considered in the tribunal of first instance. To abandon that principle is to encourage the practice of ‘sandbagging’: suggesting or permitting, for strategic reasons, that the [lower tribunal] pursue a certain course, and later—if the outcome is unfavorable—claiming that the course followed was reversible error.” Freytag, 501 U.S. at 895 (Scalia, J., concurring in part and concurring in judgment). We decline to entertain Google’s effort here, for it would only encourage litigants to engage in more of this type of behavior.

As to Google’s second argument, that its construction of “network penalty” on appeal is consistent with its arguments below, we disagree. Put frankly, Google’s arguments before the Board were hardly more than a pastiche
of block quotes from the claims and references. See J.A. 211–13. Even under the most generous of readings, Google’s arguments below did not suggest any definition of “network penalty,” let alone the highly particularized definition it presents on appeal. Consistent with the aforementioned reasoning for Google’s claim 1 arguments, we conclude that Google forfeited its claim 2 argument as to the construction of “network penalty” because Google failed to raise this argument in proceedings before the Board. In the absence of exceptional circumstances, we decline to address the merits of Google’s proposed claim constructions.

CONCLUSION

We have considered Google’s remaining arguments and find them unpersuasive. For the reasons stated, we affirm the Board’s decision upholding the rejection of claims 1–9, 11, 14–17, 19, and 20 of the ’765 application.

AFFIRMED