

Why BSA Is Needed for Traditional Music and AI-Generated Music

Introducing BSA-Compliant Endpoints for the Post-Stereo Era

Author: Raj Alur, Beyond Stereo Inc.

Date: June 9, 2026

Status: Working white paper draft

Executive Summary

Music is becoming stem-native, but playback is still mostly stereo-native.

For decades, the music industry has treated the final stereo master as the endpoint: a finished two-channel object optimized for distribution, radio, streaming, headphones, and speakers. That model will remain important. But it no longer describes the full reality of music production, remixing, fan interaction, spatial playback, automotive cabins, live/venue systems, creator tools, and AI music systems.

Traditional recordings already contain rich internal structure: vocals, drums, bass, guitars, keys, strings, percussion, ambience, effects, and mix decisions. AI tools are now making that structure easier to extract, generate, transform, and personalize. The bottleneck is no longer whether stems can exist. The bottleneck is that there is no common playback-native layer that tells devices, platforms, and systems what those stems are, what they are allowed to do, and how they should be rendered.

BSA — Beyond Stereo Audio — is needed because the industry is moving from finished stereo files toward structured musical assets, but the playback stack has no standard way to carry, control, route, reconstruct, or govern those assets.

BSA should be understood as a post-stereo format and operating layer for stem-native music. It can support traditional music, AI-assisted separation, AI-generated stems, authorized studio stems, immersive-renderer exports, cars, headphones, home systems, venues, and discrete speaker rigs.

The next step is not only defining BSA files. It is defining **BSA-compliant endpoints**: playback systems, devices, services, and APIs that can receive BSA metadata and stems, advertise their capabilities, enforce rules, and render music intelligently.

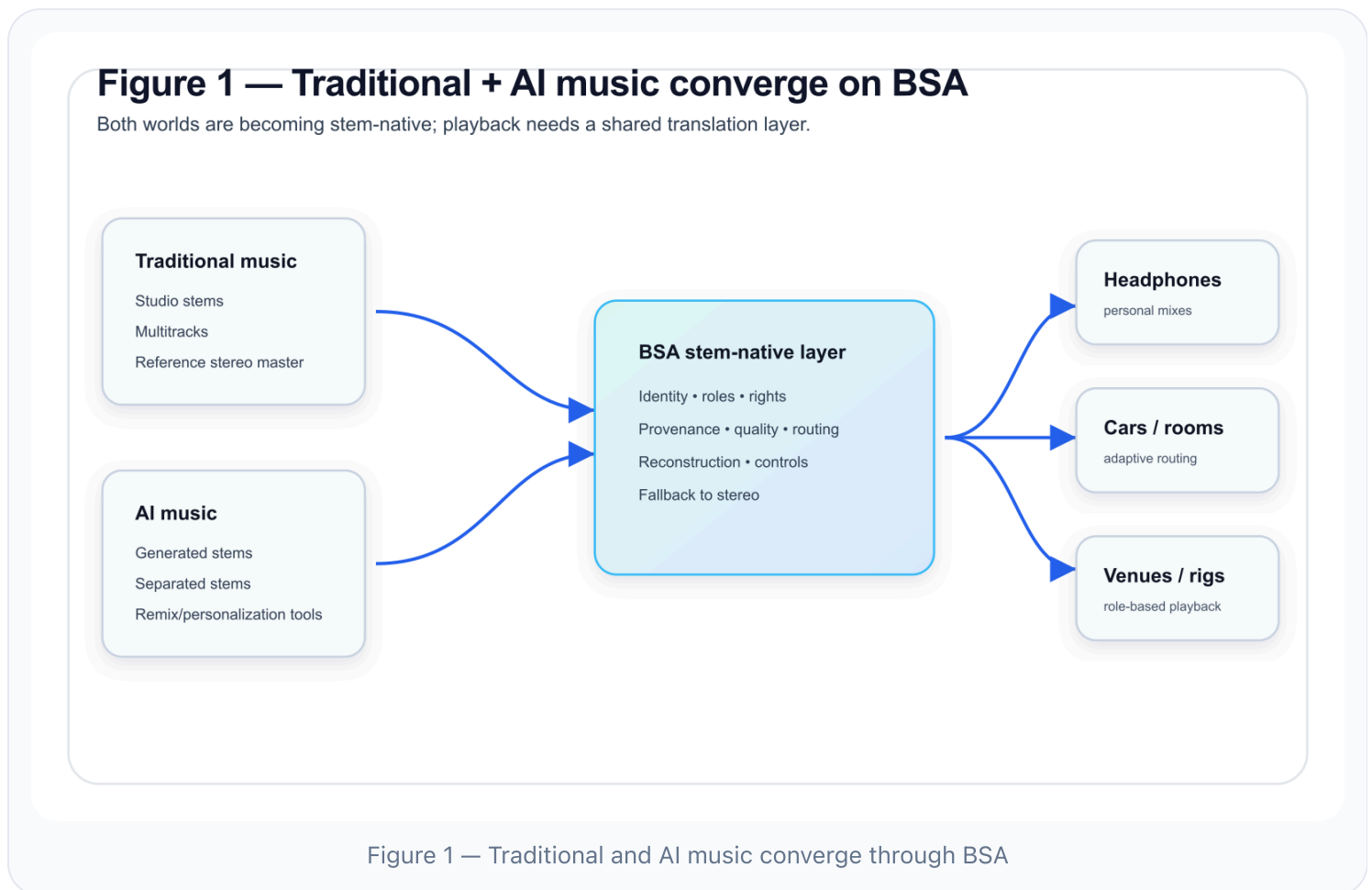


Figure 1: Traditional recordings and AI music are both becoming stem-native. BSA provides the missing translation layer from musical structure to real playback environments.

1. The Stereo Master Was Built for Distribution, Not Intelligence

Stereo solved the old distribution problem beautifully. A stereo file is compact, universal, predictable, and easy to play everywhere. It lets an artist, mixer, label, or platform deliver one finished experience across many environments.

But stereo also collapses musical structure.

Once vocals, bass, drums, instruments, ambience, and effects are mixed into two channels, the playback system can no longer reliably understand the song as a set of musical parts. It sees only left and right waveforms. It does not know which sound is a vocal, which sound is bass, which sound is a drum kit, which sound may be isolated for karaoke, which sound may be safely emphasized in a car, or which sound should be routed to a dedicated physical speaker.

That was acceptable when playback systems were passive. It is increasingly limiting now that playback systems are becoming computational, interactive, spatial, personalized, and AI-assisted.

The traditional stereo master remains essential as a reference, fallback, compatibility layer, and artistic anchor. But it is no longer enough as the only machine-readable representation of a song.

2. Traditional Music Already Wants a Stem-Native Layer

BSA is not only an AI music idea. It is needed for traditional music too.

Traditional recordings are created from stems and multitracks long before they become stereo masters. Artists and producers already make decisions around vocals, drums, bass, lead instruments, background vocals, effects, and ambience. These parts are then folded into a final mix.

A stem-native format gives traditional music several new capabilities without discarding the stereo master:

- **Better playback translation:** A song can adapt to headphones, cars, living rooms, soundbars, smart speakers, clubs, venues, and discrete speaker rigs while preserving the artistic intent of the original mix.

- **Controlled interactivity:** Listeners can experience authorized modes such as vocal-up, vocal-minus, drums-focus, bass-safe, practice mode, singalong, accessibility mode, or educational breakdowns.
- **Premium demos:** A car cabin or home speaker system can demonstrate obvious value in seconds by placing vocals, bass, drums, and instruments in ways that stereo cannot.
- **Catalog renewal:** Existing music catalogs can become interactive, adaptive, and format-forward without requiring every song to be re-recorded.
- **Artist and label control:** Rights holders can define what stems exist, what listeners may do with them, and what remains locked.
- **Compatibility:** The stereo mix can remain the reference mix and fallback path, while BSA adds structure and control where supported.

In other words: BSA does not replace traditional music production. It gives traditional music a delivery layer that matches how music is actually made.

3. AI Music Makes the Need Urgent

AI is accelerating the stem problem.

Modern AI systems can separate vocals, drums, bass, and other instruments from existing recordings. They can generate new performances, imitate production styles, create alternate arrangements, assist remixing, and produce stems directly. These capabilities are becoming faster, cheaper, and more accessible.

But AI-generated or AI-separated musical assets create a new question: *what happens after the stems exist?*

Without a standard layer, every platform must invent its own packaging, metadata, permissions, quality scoring, routing behavior, and playback logic. That creates fragmentation. It also makes it harder for artists, labels, devices, streaming services, car companies, speaker makers, and AI toolmakers to trust the output.

AI music needs BSA because stems alone are not a product. Stems need:

- identity: what each stem is;

- provenance: where it came from and how it was created;
- quality signals: whether it is usable, aligned, complete, and reconstructable;
- rights: what is allowed for listening, export, remixing, training, sharing, or monetization;
- reconstruction: how the original or reference mix is preserved;
- routing: how stems should map to playback environments;
- controls: what listeners, creators, platforms, or devices may adjust;
- fallback: how playback works when a system is not BSA-aware.

AI will produce more musical structure than the current playback stack can understand. BSA is the missing translation layer between AI stem creation and real listening experiences.

4. Why Current Approaches Are Not Enough

Stereo streaming

Stereo streaming is universal, but it has no knowledge of musical parts. It cannot natively expose authorized stem-level control, routing, or adaptive rendering.

Spatial audio and object renderers

Spatial audio systems can place sound in space, but they do not by themselves define a complete music-stem operating layer. They are renderers and delivery ecosystems. BSA can complement them by carrying stem identity, permissions, reconstruction logic, and export profiles.

Raw stem folders

A folder of WAV or MP3 stems is not a format. It lacks standard identity, permissions, routing, reference-mix reconstruction, capability negotiation, endpoint behavior, and consumer playback rules.

AI separation outputs

AI-separated stems can be useful, but quality varies. Some songs separate cleanly; others create artifacts, missing energy, phase issues, or incomplete parts. A BSA package can include quality

metadata and residual/reconstruction logic so playback systems know how to use the stems safely.

Proprietary platform formats

Closed formats may create good experiences inside one product, but they do not solve the industry-level problem. Music needs a common layer that can travel across tools, services, endpoints, and environments.

5. What BSA Adds

BSA is best understood as a stem-native music layer, not merely a codec.

A BSA-compliant asset can include:

- the original or reference stereo mix;
- authorized stems or AI-derived stems;
- stem identity and hierarchy;
- timing, alignment, loudness, and mix-preservation metadata;
- per-stem roles such as lead vocal, backing vocal, kick, snare, bass, guitar, keys, strings, ambience, effects, or residual;
- rights and usage metadata;
- quality metrics and separation provenance;
- reconstruction rules for preserving the intended mix;
- routing recommendations for different playback environments;
- fallback behavior for non-BSA players.

This creates a common bridge between production, AI processing, distribution, and playback.

The key shift is simple:

Stereo tells a system what sound to play. BSA tells a system what musical parts exist, what they mean, what rules apply, and how they can be rendered.

6. BSA for Traditional Music

For traditional catalogs, BSA can support both studio-authorized stems and AI-assisted conversion.

A label, artist, or distributor could release a BSA version of a track using official stems. Where official stems are unavailable, an authorized workflow could use AI separation to create a BSA-compatible experience, with metadata clearly identifying the source, model, quality, and limitations.

Traditional music use cases include:

- premium album editions;
- automotive demos and cabin-native listening;
- home theater and smart speaker experiences;
- karaoke and vocal-minus modes;
- creator-safe remix packs;
- music education and practice modes;
- accessibility mixes, such as vocal clarity or reduced crowd noise;
- venue and event playback;
- artist-approved interactive releases.

The value is not novelty for its own sake. The value is giving traditional music a modern delivery layer that can preserve the master while unlocking controlled new experiences.

7. BSA for AI Music and AI-Assisted Music

AI music makes BSA more necessary, not less.

As AI systems generate and separate stems, listeners will expect music to be interactive, personalized, and environment-aware. But rights holders and platforms will need clear boundaries.

BSA can provide those boundaries by packaging AI-created or AI-derived stems with:

- model/source provenance;
- generation or separation date;
- confidence and quality scores;
- human-authorized labels or approvals;
- restrictions on export, remixing, commercial use, and training;
- reference-mix reconstruction;
- endpoint-specific rendering rules.

This lets the industry avoid a false choice between uncontrolled AI remix chaos and locked-down stereo-only playback. BSA creates a middle path: structured, governed, playback-native music.

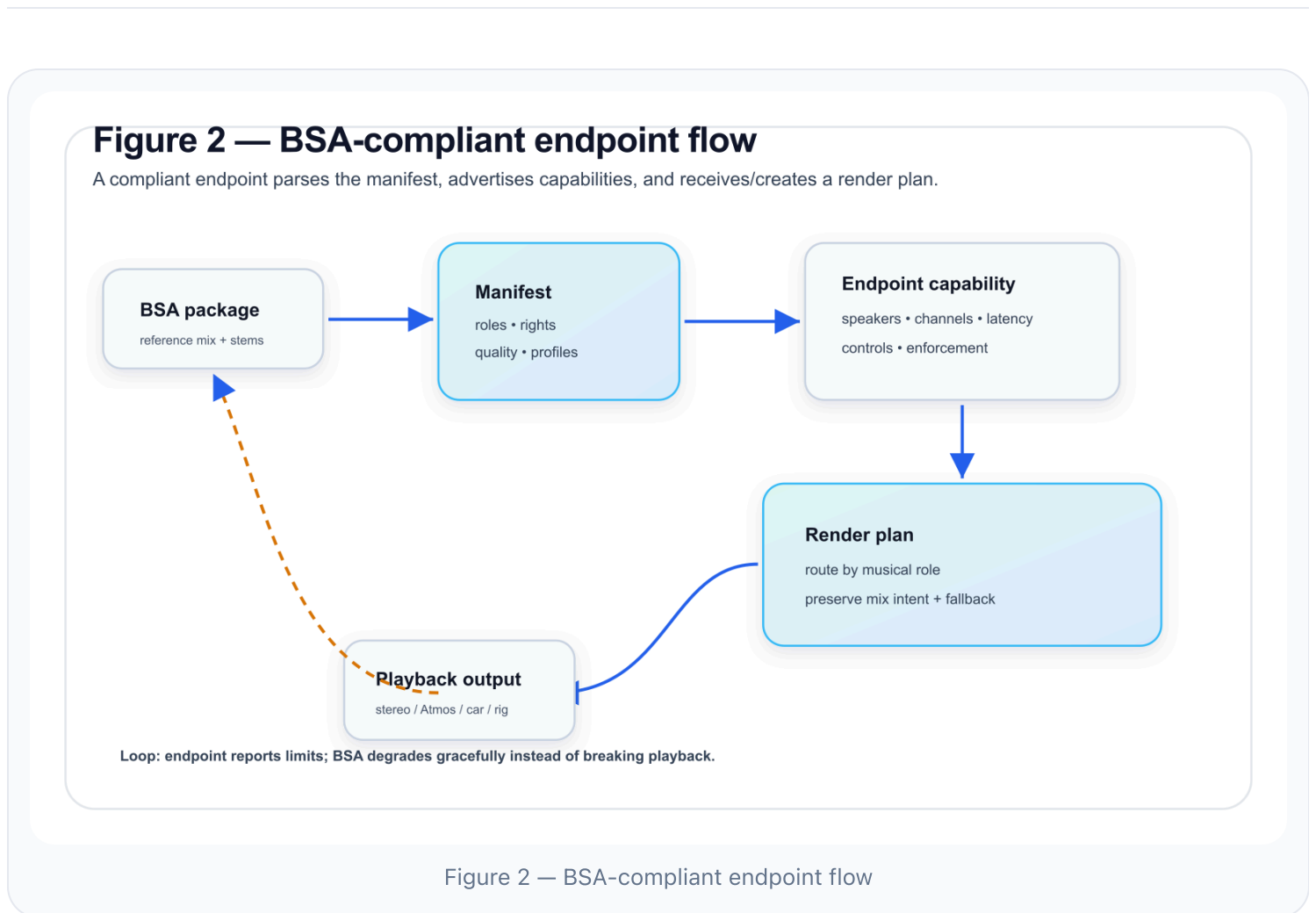


Figure 2: A BSA-compliant endpoint parses the manifest, advertises its capabilities, enforces rules, and renders a mix-preserving output plan.

8. The Next Layer: BSA-Compliant Endpoints

A format alone is not enough. The industry also needs endpoints that know how to receive and render BSA.

A **BSA-compliant endpoint** is any playback device, application, service, renderer, API, or hardware system that can interpret BSA metadata and behave according to BSA rules.

Examples include:

- a streaming app that can expose approved stem controls;
- a car audio system that can route vocals, bass, drums, and ambience intelligently across a cabin;
- a smart speaker or soundbar that can request the right mix profile for its speaker layout;
- a headphone app that can render a personal stem-aware mix;
- a venue processor that can map musical roles to zones;
- an AI music service that can export BSA-compliant assets;
- a cloud API that can validate, search, license, and deliver BSA packages;
- a discrete speaker rig where each physical speaker receives a musically meaningful role.

This is where BSA moves from “file format” to “music operating layer.”

9. What Makes an Endpoint BSA-Compliant

A practical BSA-compliant endpoint should support several core behaviors.

1. Capability advertising

The endpoint should describe what it can do:

- number of channels or speakers;
- speaker locations or roles;
- headphone, car, room, venue, or discrete-rig profile;

- supported codecs and container types;
- maximum simultaneous stems;
- latency and synchronization limits;
- supported control modes;
- whether it can enforce rights locally.

2. BSA metadata parsing

The endpoint should understand the BSA manifest:

- track identity;
- stem list;
- stem roles;
- timing and alignment;
- loudness and gain relationships;
- routing recommendations;
- allowed controls;
- provenance and quality metadata;
- fallback instructions.

3. Rights-aware behavior

The endpoint should respect permissions:

- play-only;
- export disabled;
- remix allowed or blocked;
- stems hidden, visible, soloable, or adjustable;
- subscription or license tier requirements;
- time-limited or territory-limited access where applicable.

4. Mix-preserving rendering

The endpoint should preserve the reference mix unless an allowed mode changes it. This matters because stem-native playback should not destroy the artistic balance of the song.

A compliant endpoint should understand default gain, allowed gain ranges, reconstruction rules, residual tracks, and safe fallback behavior.

5. Role-based routing

The endpoint should route stems based on musical role and playback environment. For example:

- lead vocal to center or near-field focus;
- bass to capable low-frequency speakers;
- drums to foundation or presence channels;
- ambience to surround or room-filling channels;
- guitars/keys/strings to spatial or supporting positions;
- residual/reference mix to preserve completeness.

6. Fallback and graceful degradation

If an endpoint cannot support full BSA playback, it should still be able to:

- play the reference stereo mix;
- request a reduced stem set;
- collapse to a safe mix profile;
- disable unsupported controls;
- report noncompliance clearly.

10. Proposed BSA-Compliant API Endpoints

The first BSA service layer can be simple and versioned from day one.

Suggested API surface:

- `GET /v1/tracks/{track_id}` — return track-level BSA metadata.

- `GET /v1/tracks/{track_id}/manifest` — return the BSA manifest without downloading full audio.
- `GET /v1/tracks/{track_id}/stems` — list available stems, roles, quality, and access permissions.
- `GET /v1/tracks/{track_id}/profiles` — return supported playback profiles such as stereo fallback, headphone, car cabin, soundbar, discrete rig, venue, or renderer export.
- `POST /v1/endpoints/register` — register or describe a playback endpoint and its capabilities.
- `POST /v1/render/plan` — submit endpoint capabilities and receive a stem routing/rendering plan.
- `POST /v1/licenses/verify` — verify whether a user, device, or service can access a track, stem, control mode, or export.
- `POST /v1/packages/validate` — validate whether a BSA package is structurally compliant.
- `POST /v1/packages` — upload or contribute a BSA package, subject to authentication and rights rules.
- `GET /v1/search` — search by artist, title, ISRC, MusicBrainz ID, model hash, stem role, or availability.
- `GET /v1/download/{package_id}` — download an authorized BSA package or selected profile.

This endpoint layer separates three concerns:

1. **What the music contains** — stems, roles, metadata, quality, rights.
 2. **What the playback endpoint can do** — speakers, channels, controls, latency, enforcement.
 3. **What should happen now** — routing, rendering, controls, fallback, or denial.
-

Figure 3 — BSA endpoint compliance levels

A staged model lets apps, devices, cars, and services adopt BSA incrementally.

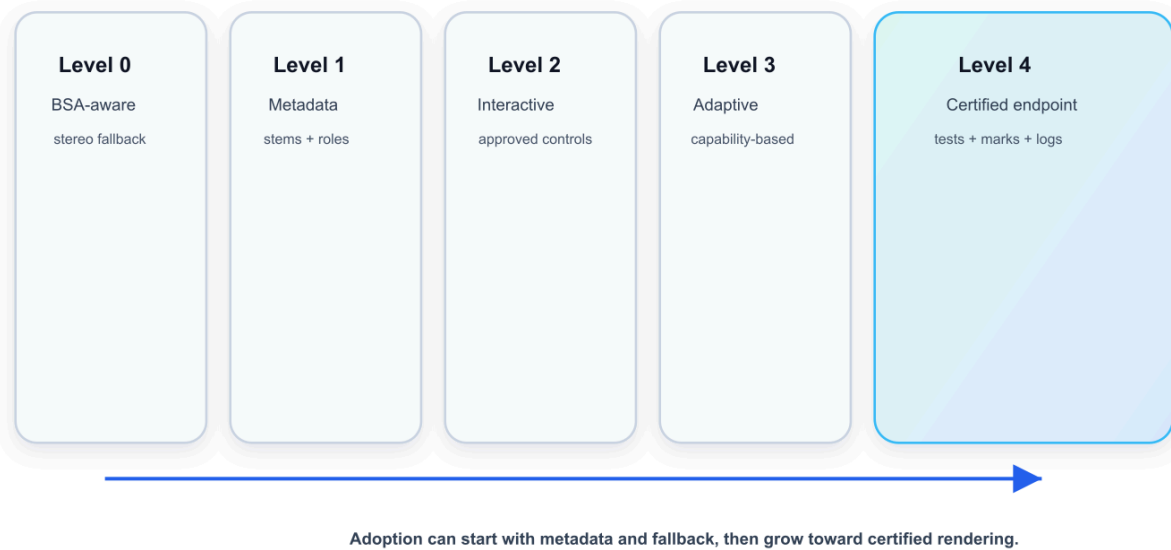


Figure 3 — BSA endpoint compliance levels

Figure 3: BSA compliance can be staged, from stereo fallback through certified adaptive endpoints.

11. Endpoint Compliance Levels

BSA can support staged adoption through compliance levels.

Level 0: BSA-aware fallback

- Reads basic metadata.
- Plays the reference stereo mix.
- Displays BSA availability.
- Does not expose stem control.

Level 1: Metadata-compliant playback

- Parses manifest.
- Lists stems and roles.

- Supports approved playback profiles.
- Uses safe default mix-preserving behavior.

Level 2: Interactive stem playback

- Enables approved user controls.
- Supports solo/mute/gain within allowed ranges.
- Supports role-based routing.
- Enforces rights at playback time.

Level 3: Adaptive endpoint rendering

- Advertises endpoint capabilities.
- Receives or computes render plans.
- Adapts to headphones, cars, rooms, venues, or discrete rigs.
- Preserves synchronization and mix intent.

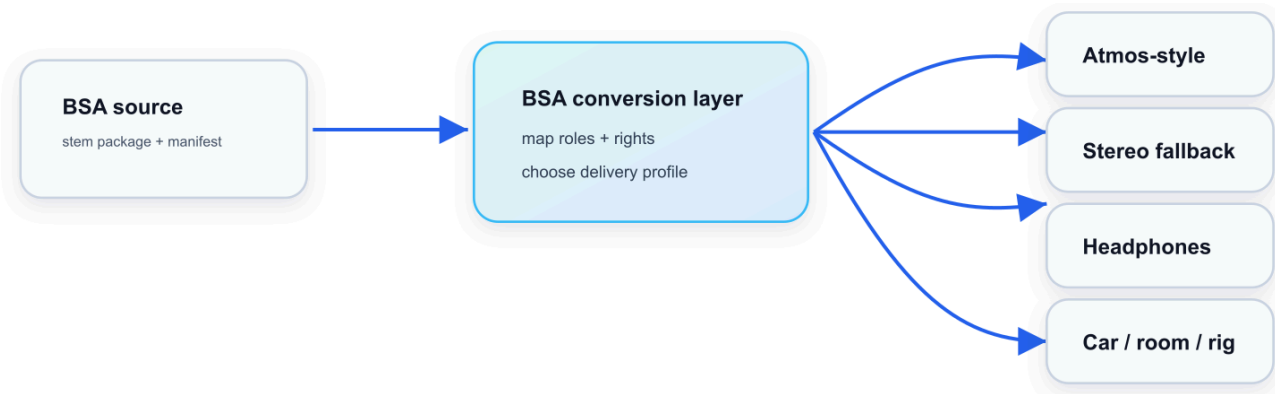
Level 4: Certified BSA endpoint

- Passes validation tests.
- Enforces rights and fallback rules.
- Reports compliance status.
- Supports robust logs, error handling, and security expectations.

This lets the market adopt BSA incrementally. A streaming app, car system, AI tool, speaker, and studio product do not all need the same implementation on day one.

Figure 4 — Atmos as a BSA-compatible output profile

BSA can map stem-native assets into spatial renderers, stereo, headphones, cars, or discrete rigs.



Positioning: Atmos can be a BSA-compatible output profile.
BSA carries identity, rights, provenance, reconstruction, and endpoint logic.

Figure 4 — Atmos as a BSA-compatible output profile

Figure 4: Atmos-style delivery can be treated as one BSA-compatible output profile. BSA carries stem identity, rights, provenance, reconstruction, and endpoint logic; the renderer handles final spatial delivery.

12. Why This Matters for the Music Industry

BSA gives different stakeholders a shared language.

For artists, it creates a way to release interactive music without losing control.

For labels and rights holders, it creates a structured path to monetize catalogs, authorize new experiences, and govern AI-assisted usage.

For streaming platforms, it creates differentiated listening modes beyond louder, higher-resolution, or more personalized recommendation feeds.

For AI companies, it creates a standard output target for generated or separated music.

For device makers, it creates a reason for better speakers, car systems, headphones, and local compute.

For listeners, it turns music from a fixed two-channel file into an experience that can adapt without becoming random or disrespectful to the original work.

13. Adoption Path

The practical adoption path should be simple:

1. **Start with BSA packages for selected demo tracks.** Include reference mix, stems, manifest, quality metadata, and fallback.
 2. **Define a minimal BSA manifest.** Keep it small enough for developers and partners to implement quickly.
 3. **Release BSA endpoint guidelines.** Define what apps, players, services, and devices must do to be BSA-compliant.
 4. **Build a validation tool.** Let developers test whether a package or endpoint follows the rules.
 5. **Create bridge profiles.** Support outputs to spatial renderers, car cabins, headphones, discrete speaker rigs, and stereo fallback.
 6. **Publish sample APIs.** Give partners a clean way to search, validate, license, download, and render-plan BSA assets.
 7. **Certify early endpoints.** Start with the Beyond Stereo player, demo rig, cloud API, and a small number of partner profiles.
-

14. Closing Claim

The future of music is not simply “AI-generated songs” or “better stereo.” The future is structured music: songs that retain their musical parts, provenance, rights, and intended relationships all the way to playback.

Stereo will continue to matter. But stereo alone cannot carry the next era of music.

BSA is needed because traditional music and AI music are both becoming stem-native, while playback remains trapped in a two-channel abstraction.

The next step is to make BSA real at the endpoint: apps, cars, speakers, headphones, AI services, cloud APIs, and discrete systems that can understand the song's parts and render them responsibly.

That is the purpose of BSA-compliant endpoints: to make post-stereo music playable, governable, and useful in the real world.

Appendix A: Minimal BSA Manifest Concepts

A minimal BSA manifest should describe:

- track identity;
 - artist/rightsholder identity where available;
 - reference mix location;
 - stem list;
 - stem role labels;
 - timing/alignment;
 - default gains and allowed gain ranges;
 - quality/provenance data;
 - rights/permissions;
 - playback profiles;
 - fallback behavior;
 - package/version information.
-

Appendix B: Minimal Endpoint Capability Descriptor

A minimal BSA endpoint descriptor should describe:

- endpoint type: app, headphones, car, soundbar, smart speaker, venue, discrete rig, cloud renderer, AI service;
 - speaker/channel count;
 - supported sample rates/codecs;
 - maximum stems;
 - latency constraints;
 - control surface;
 - rights enforcement capability;
 - supported playback profiles;
 - certification/compliance level.
-

Appendix C: Public-Safe One-Line Positioning

Beyond Stereo Audio is the stem-native format and endpoint layer for post-stereo music, enabling traditional and AI-generated songs to carry musical structure, rights, provenance, and playback intelligence across apps, devices, and environments.