Tribunal Session Summary

Session: conclave-20260803-222647-should-i-buy-jenbacher-af48

Depth: T6

Advocates: 6 | Judges: 5 | Cost: $14.4532 | Time: 41m 55s

THE PROMPT: Should I buy Jenbacher or Bergen for offline, medium based load power generation. We have to sign a purchase order in four hours.

THE RULING: Do not sign an unconditional purchase order for either vendor. Issue a conditional award to Bergen Engines for pipeline gas/LNG applications, subject to binding guarantees on delivery, site-rated efficiency, and island-mode performance. Select Jenbacher only if the fuel is difficult gas (landfill, sewage, coke-oven) or if Bergen cannot provide the required contractual commitments.

The Prompt

Should I buy Jenbacher or Bergen for offline, medium based load power generation. We have to sign a purchase order in four hours.

The Ruling

Do not sign an unconditional purchase order for either vendor. Issue a conditional award to Bergen Engines for pipeline gas/LNG applications, subject to binding guarantees on delivery, site-rated efficiency, and island-mode performance. Select Jenbacher only if the fuel is difficult gas (landfill, sewage, coke-oven) or if Bergen cannot provide the required contractual commitments.

The evidence shows Bergen's medium-speed architecture provides superior durability for continuous baseload operation, but only if backed by enforceable performance guarantees. Jenbacher's explicitly documented island-mode capability makes it the safer fallback option, particularly for difficult gases. The four-hour deadline necessitates a capped, conditional award rather than a binding PO to avoid unacceptable technical and commercial risks.

Core Analysis

Technical Assessment

Bergen's Durability Advantage

- Operates at 720-750 rpm vs. Jenbacher's 1,500 rpm, resulting in approximately half the firing cycles per operating hour
- Lower Brake Mean Effective Pressure (~21 bar vs. ~25 bar for Jenbacher J624) reduces mechanical stress
- Medium-speed design derived from marine applications, optimized for continuous high-load operation
- Documented in multiple true-island deployments (e.g., Liberty Energy's 500MW off-grid installation)

Efficiency Comparison

- Bergen B36:45V series: 48.9-50% nominal electrical efficiency (ISO conditions)
- Jenbacher portfolio: >45% efficiency, with J920 FleXtra at ~48.7%
- The 0.2-1.3 percentage point efficiency advantage favors Bergen but must be verified at site conditions

Island-Mode Capability

- Jenbacher: Explicitly validated for "grid-parallel and islanded/standalone operation in remote areas" in CURRENT FACTS

- Bergen: No explicit island-mode claim in CURRENT FACTS, but proven in operational projects (Svalbard, Australian outback)
- Both require contractual guarantees for transient performance (frequency nadir, voltage recovery, black start)

Mandatory Contractual Protections

1. **LD-Backed Delivery & COD** - Firm commercial operation date with liquidated damages - Jenbacher typical COD: 8-16 months for LNG-to-power - Bergen: "Short lead time" claim must be converted to specific date
2. **Site-Rated Performance Guarantees** - Net electrical efficiency at 50%, 75%, 100% load - Heat rate curves under actual dispatch profile - Ambient and altitude derating applied
3. **Island Transient Envelope** - Frequency stability during largest unit trip - Voltage dip/recovery for motor starts - Black-start sequence timing - Load acceptance/rejection performance
4. **N+1 Topology** - Largest single contingency ≤25% of online capacity - Multi-unit configuration to avoid excessive part-loading
5. **LTSA & Local Support** - Availability-guaranteed maintenance contract ($/MWh) - Consignment spares inventory - Maximum MTTR commitment - Local technician response time

Fuel-Specific Decision Gates

Fuel Type	Recommended Action
Pipeline Gas/LNG	Proceed with conditional Bergen award
Landfill/Sewage/Coke-oven Gas	Select Jenbacher immediately
Liquid/Dual-Fuel	Select Bergen (Jenbacher has no liquid platform)

Implementation Roadmap (4-Hour Window)

First 30 Minutes

- Confirm fuel specifications and methane number
- Validate required COD timeline
- Issue dual-track RFQ to both vendors with mandatory terms

Next 2 Hours

- Review vendor responses for completeness
- Verify N+1 topology compliance
- Validate site-net efficiency guarantees

Final 1.5 Hours

- Execute conditional award to vendor meeting all requirements
- If neither complies, issue capped LNTP (≤$50k) with 14-day diligence window

How We Got Here

Scorecard

Judge	Grok 4.5	Gemini 3.1 Pro	DeepSeek V4 Pro	GPT-5.6 Sol	Claude Opus 5	Magistral Medium
GPT-OSS 120B	STRONG	WEAK	WEAK	STRONG	STRONG	WEAK
DeepCogito v2.1	WEAK	WEAK	WEAK	STRONG	STRONG	WEAK
Kimi K2.6	STRONG	WEAK	WEAK	STRONG	STRONG	WEAK
Mistral Large 3	STRONG	WEAK	WEAK	STRONG	STRONG	WEAK
Nemotron 3 Ultra	STRONG	WEAK	WEAK	STRONG	STRONG	WEAK

Convergence Assessment

The convergence was primarily epistemic, driven by verified evidence rather than social pressure. Key evidence-driven shifts included:

- Bergen's 3.36MW unit discovery (portfolio document)
- RoCoF physics correction (textbook decomposition)
- Island-mode project verification (Liberty Energy, Svalbard)
- Service interval data (Jenbacher's 60k hours vs. Bergen's undocumented)

Magistral Medium's late concession on the unconditional PO was the only major affective shift, occurring after all other advocates had abandoned the position.

Closing Statements

Claude Opus 5: Abandoned all three original load-bearing arguments (granularity, delivery, Piller dependency) when evidence contradicted them. Concluded no brand-level technical winner exists for pipeline gas, with the decision turning on contractual guarantees and fuel type. Strongest surviving evidence: cycle-count physics and the documentation asymmetry (Jenbacher publishes TBO, Bergen doesn't).

GPT-5.6 Sol: Maintained consistent focus on contractual guarantees over brand preference. Refined position to acknowledge Bergen as a plausible tiebreaker (~60-65% confidence) for high-CF duty but

emphasized that neither vendor should receive an unconditional PO. Strongest surviving evidence: the N+1 capacity inequality and the distinction between startup time and transient stability.

Grok 4.5: Refined the durability argument from piston speed to cycle count/BMEP while maintaining Bergen as the conditional recommendation. Added specific gates for delivery, efficiency, and island-mode guarantees. Strongest surviving evidence: Bergen's 48.9-50% efficiency range and true-island project deployments.

DeepSeek V4 Pro: Completely reversed position from Jenbacher to Bergen conditional, abandoning the original "silence equals incapability" argument. Emphasized the need for binding guarantees on all critical parameters. Strongest surviving evidence: Bergen's liquid-fuel capability and the invalidity of arguments from silence.

Gemini 3.1 Pro: Abandoned the entire original thesis (fast-start criticality) and adopted Bergen based on cyclic durability. Conceded all physics arguments and granularity constraints. Strongest surviving evidence: the 2:1 firing-cycle advantage for medium-speed engines.

Magistral Medium: Reluctantly conceded most arguments (efficiency, startup, hydrogen) but maintained Jenbacher as fallback based on CURRENT FACTS wording. Showed the highest sycophancy risk, with position shifts lagging evidence. Strongest surviving evidence: Jenbacher's explicit island-mode validation in the provided facts.

Key Moments

- **Granularity Argument Collapse (Round 6)** Claude Opus 5 discovered Bergen's B33:45L6 3.36MW unit in the 2025 portfolio, disproving the claim that Bergen couldn't match Jenbacher's modularity. This eliminated Jenbacher's primary technical differentiator and forced advocates to focus on durability and contractual guarantees.
- **Durability Mechanism Correction (Round 5)** Claude Opus 5 challenged Grok 4.5's "lower piston speed" argument, proving Bergen actually operates at higher mean piston speed (11.25 m/s vs. Jenbacher's 11.0 m/s). The debate shifted to cycle count (2:1 advantage) and BMEP (21 bar vs. 25 bar) as the true durability differentiators.
- **Island-Mode Validation (Round 4)** External evidence revealed Bergen's Liberty Energy project (500MW off-grid) and Svalbard installations, disproving the argument that only Jenbacher had proven island capability. This neutralized Jenbacher's strongest documented advantage.
- **Cold-Start Physics Correction (Round 3)** GPT-5.6 Sol's timescale decomposition proved that sub-45-second startup cannot prevent sub-second frequency collapse, dismantling the argument that Jenbacher's fast-start capability was critical for island stability.

Reality Checks

Claim	Challenge	Result
Gemini 3.1 Pro: Jenbacher's <45s startup is critical for island stability	**Claude Opus 5:** Cold-start time cannot prevent sub-second frequency collapse (RoCoF)	Full concession; thesis abandoned
Magistral Medium: Bergen lacks explicit island-mode capability	**Grok 4.5:** Liberty Energy's 500MW off-grid Bergen installation	Full concession; argument from silence abandoned
DeepSeek V4 Pro: Jenbacher has superior granularity for N+1	**Claude Opus 5:** Bergen's B33:45L6 at 3.36MW matches Jenbacher's modularity	Full concession; position reversed
Gemini 3.1 Pro: Bergen's 6.8MW floor forces part-load inefficiency	**GPT-5.6 Sol:** 3.36MW Bergen units enable optimal N+1 topologies	Full concession; granularity argument retracted

Next Steps

Next Step	Why It Matters	Confidence
Confirm fuel specifications	Determines vendor eligibility (Bergen for liquid, Jenbacher for difficult gases)	95%
Secure LD-backed COD commitment	Current market shows >15GW backlogs; delivery certainty critical	90%
Obtain site-net heat rate curves	Brochure efficiency deltas may not hold under actual dispatch	85%
Validate island transient performance	Offline operation requires proven frequency/voltage control	80%
Negotiate LTSA with availability guarantee	MTTR more critical than MTBF for offline plants	75%

Dissenting Opinions

No dissenting opinions were filed in this session.

Glossary

Term	Definition
Baseload	Continuous, minimum level of electricity demand on the grid over a period of time
BMEP (Brake Mean Effective Pressure)	A measure of engine's efficiency in converting fuel into work, calculated from torque and displacement
COD (Commercial Operation Date)	The date when a power plant becomes operational and begins commercial service
HFO (Heavy Fuel Oil)	A type of fuel oil produced from petroleum, commonly used in large marine and industrial engines
LD (Liquidated Damages)	Contractual damages paid for failure to meet performance or delivery obligations
LTSA (Long-Term Service Agreement)	A maintenance contract covering parts, labor, and availability guarantees
MTBF (Mean Time Between Failures)	The average time between system failures during normal operation
MTTR (Mean Time To Repair)	The average time required to repair a failed component or system
N+1 Redundancy	A system design where capacity exists to handle the failure of any single component
RoCoF (Rate of Change of Frequency)	The speed at which power system frequency changes following a disturbance

How The Conclave Works

The Conclave is a structured multi-model deliberation system. Instead of asking one AI model a question and trusting its answer, the Conclave convenes a panel of independent models, forces them to argue, and subjects their conclusions to judicial review. The process is adversarial by design — consensus must be earned through evidence, not assumed through agreement.

Advocates are independent AI models (e.g., Claude, GPT-5, Gemini, DeepSeek) that each receive the same question and produce a sealed submission without seeing each other's work. Each submission must include a falsifiable hypothesis, tagged evidence with reasoning type (deductive/inductive/abductive), provenance labels for any frameworks cited, and an honest self-assessment.

Challenges follow submissions. Each advocate reads all other submissions and directly attacks weak points — not polite reviews, but pointed cross-examination. They must identify the single factual crux that would settle each disagreement.

Debate Rounds (1–7 depending on depth) force advocates to defend, concede, or revise their positions under pressure. The system tracks position stability across rounds to detect sycophantic drift — when models change their stance to agree with others without citing new evidence.

Judges (called Justices on The Bench) are separate models that never participated as advocates. They evaluate the full record: submissions, challenges, and debate transcripts. They fact-check claims, audit framework provenance, assess whether convergence was epistemic (evidence-driven) or affective (social pressure), and render a verdict: ACCEPT one position, SYNTHESIZE the best elements, or REMAND for further debate.

Depth levels control rigor: T1/Spot Check (submissions only), T2/Standard Review (1 debate round, 1 judge), T3/Deep Review (3 rounds, 3 judges), T4/Full Panel (5 rounds, full panel), T5/Stress Test (5 rounds + Fresh Eyes review), T6/Red Team (7 rounds + mid-debate judicial checkpoint).

The system is deterministic code — it cannot be sycophantic. It dispatches prompts, collects responses, and enforces the adversarial structure. The models argue; the code referees.