



BizFirstAi

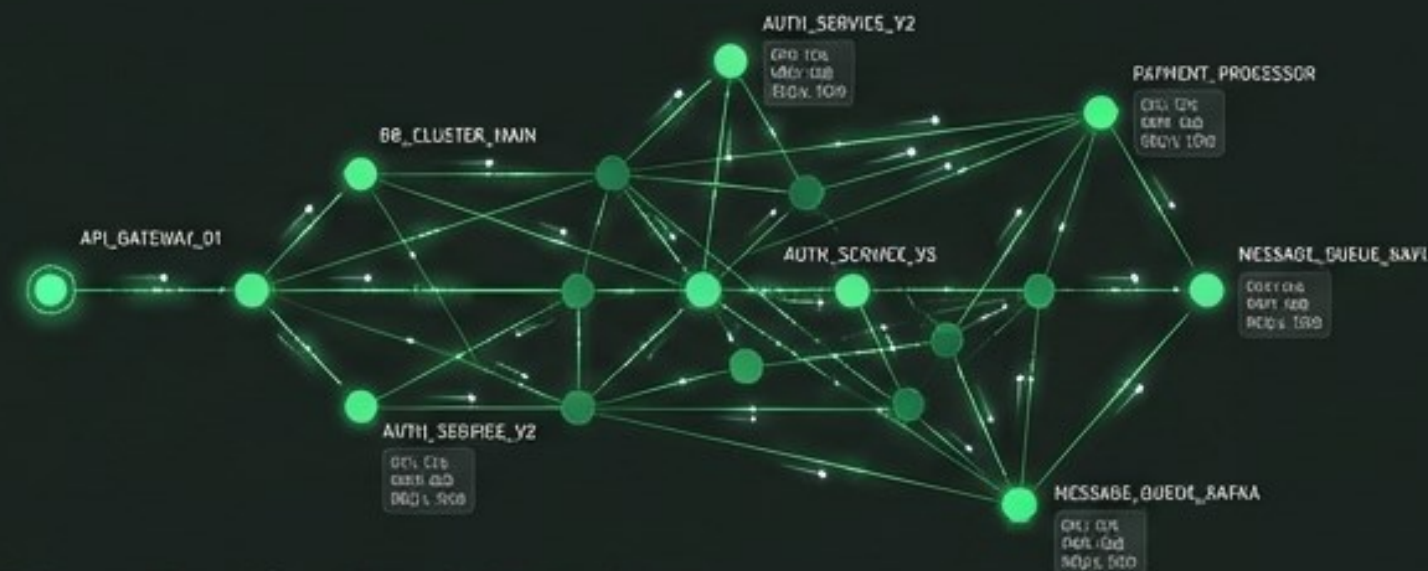
BizFirstAi Observability Architecture The Enterprise Command Center

Unifying Metrics, Logs, and Traces across a Multi-Tenant Ecosystem



90% Cost Reduction

Replaces expensive \$1M+ Datadog and Splunk legacy instances.



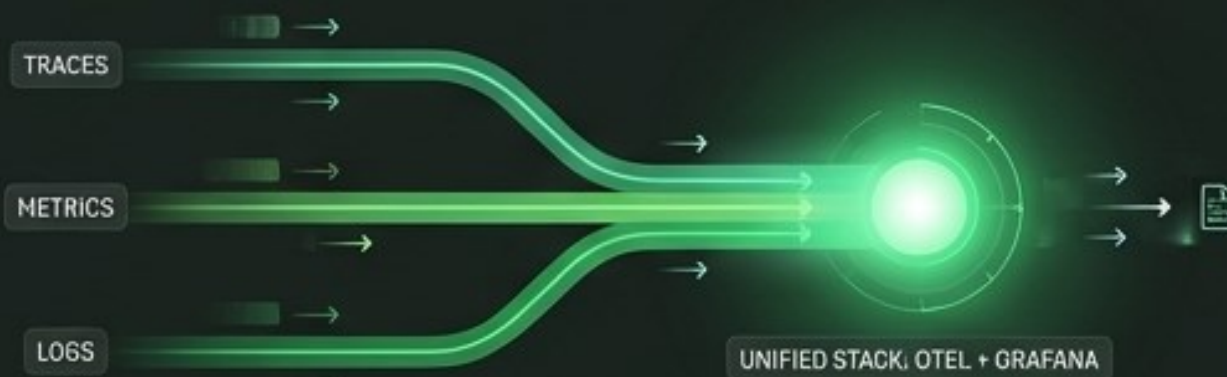
50+ Built-In Metrics

Simultaneously monitoring 10+ interconnected platform products.



Sub-Second Trace Lookup

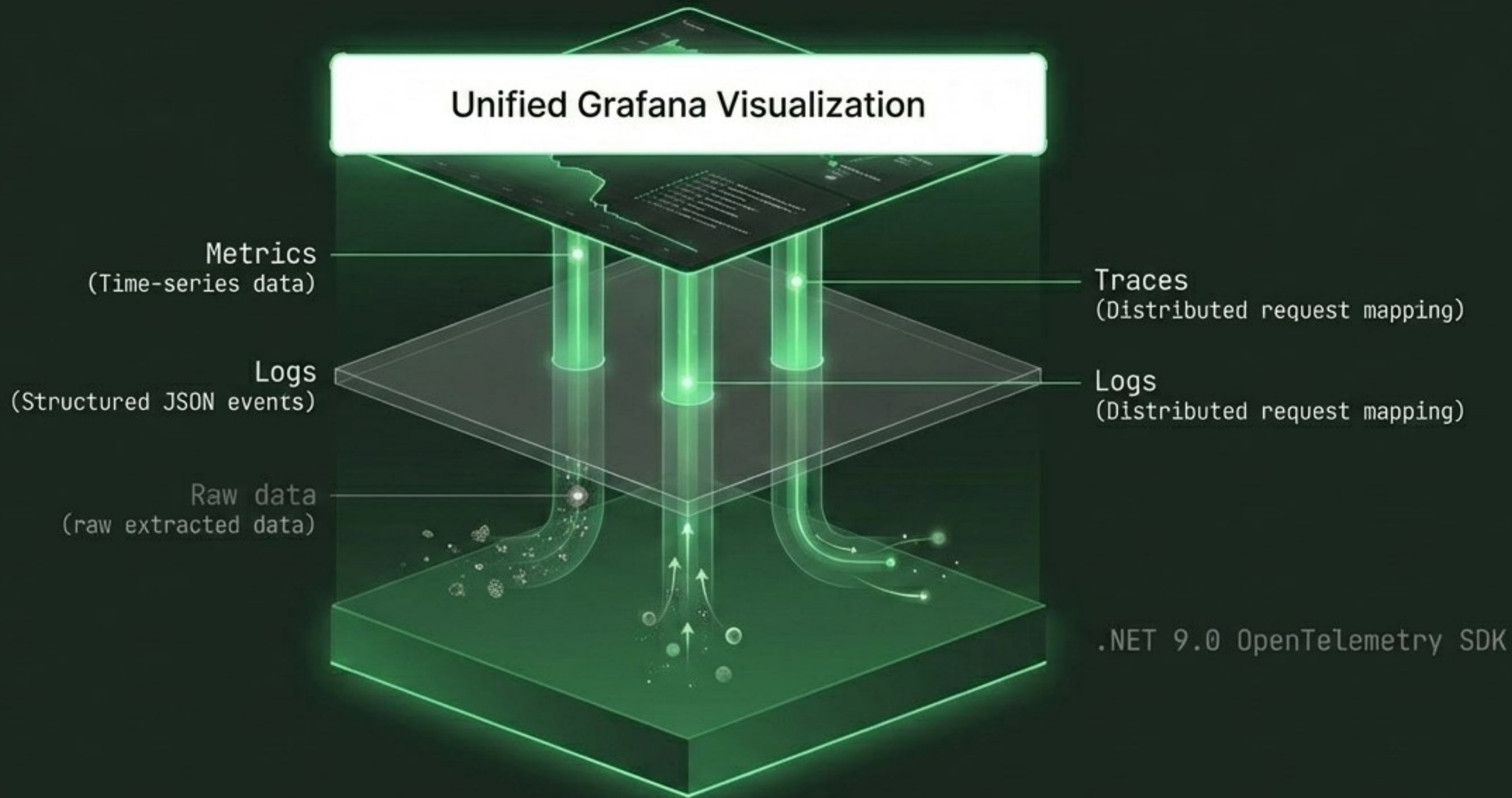
Reduces Mean Time To Resolution (MTTR) from 4+ hours to just 15 minutes.



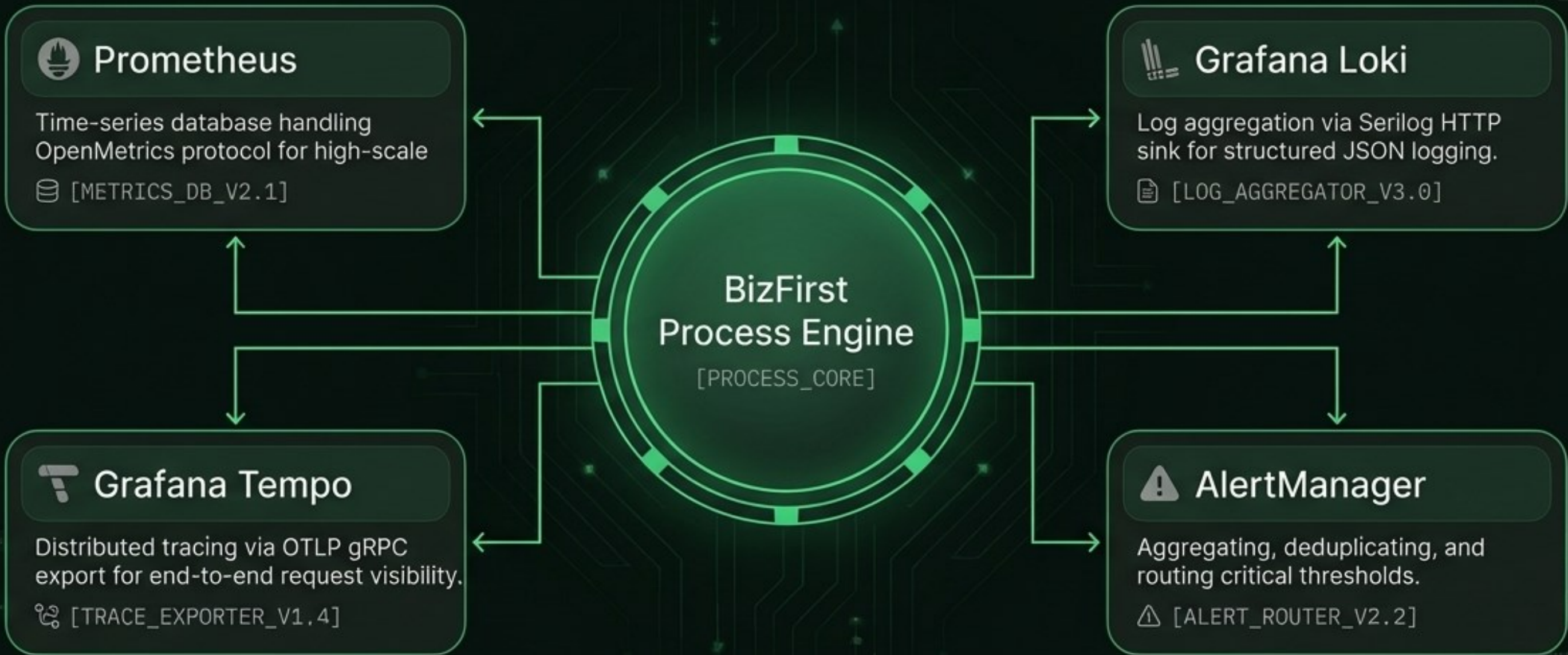
Unified 3-Signal Convergence

Bringing traces, metrics, and logs into a single OpenTelemetry and Grafana stack.

The Three Pillars of Observability



The Open-Source Engine Room



Zero Vendor Lock-in: Standardized entirely on CNCF open-source architecture.

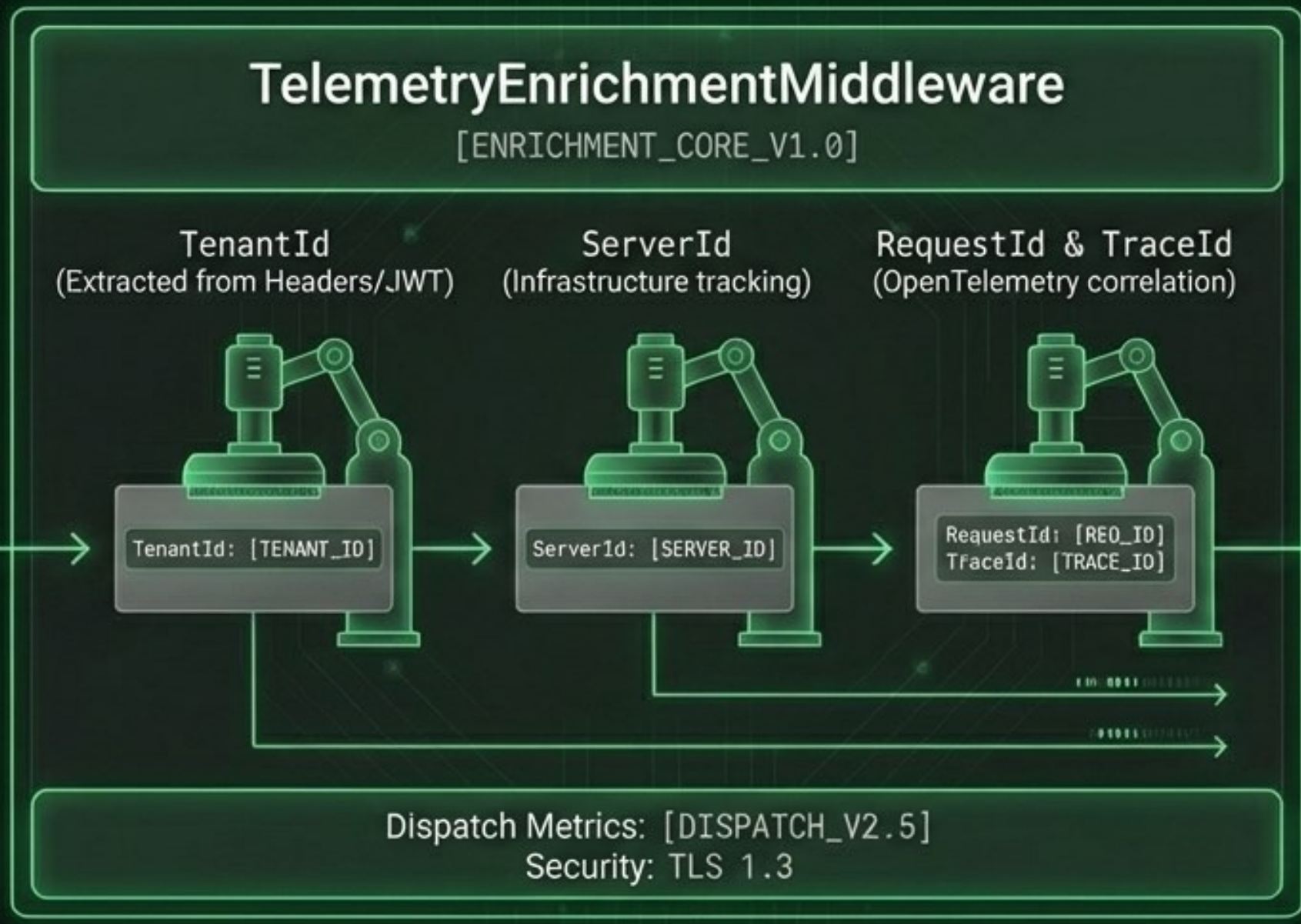
Telemetry Enrichment (The Secret Sauce)

Raw HTTP Request



(Unprocessed data)

```
Request: {  
  "headers": {  
    "Content-Type": "application/json",  
    "User-Agent": "Mozilla/5.0",  
    "Accept": "application/json",  
    "Host": "example.com",  
    "Authorization": "Bearer [JWT_TOKEN]"  
  },  
  "body": {  
    "id": 1,  
    "name": "John Doe",  
    "email": "john.doe@example.com",  
    "password": "secret123",  
    "role": "user"  
  }  
}
```



Enriched Data Packet



Securely isolated, enriched data packet ready for dispatch to Prometheus and Loki.

```
Request: {  
  "headers": {  
    "Content-Type": "application/json",  
    "User-Agent": "Mozilla/5.0",  
    "Accept": "application/json",  
    "Host": "example.com",  
    "Authorization": "Bearer [JWT_TOKEN]",  
    "X-Tenant-Id": "[TENANT_ID]",  
    "X-Server-Id": "[SERVER_ID]",  
    "X-Request-Id": "[REQ_ID]",  
    "X-Trace-Id": "[TRACE_ID]"  
  },  
  "body": {  
    "id": 1,  
    "name": "John Doe",  
    "email": "john.doe@example.com",  
    "password": "secret123",  
    "role": "user"  
  }  
}
```

Pillar 1: Time-Series Metrics

System Health



`bizfirst_health_check_status`

Monitors overall liveness. Dial shows '2' for fully Healthy.

Infrastructure



`edgestream_kafka_consumer_lag_messages`

Monitors distributed event streaming queue depth and consumer lag.

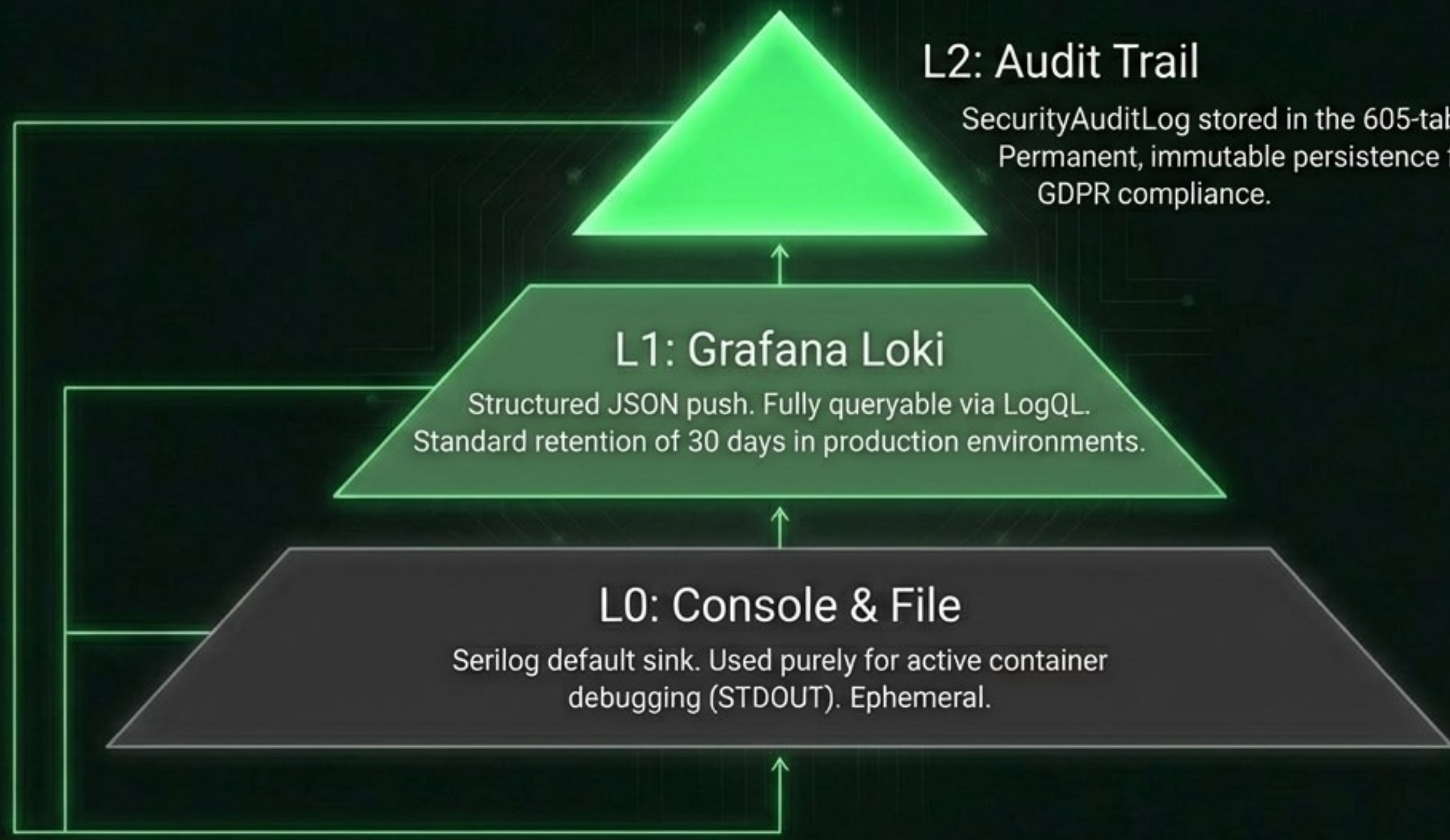
Application Latency



`http_server_request_duration_seconds`

Visualizes HTTP response percentiles at P50, P95, and P99 thresholds.

Pillar 2: Structured Logging



L2: Audit Trail

SecurityAuditLog stored in the 605-table DataOcean. Permanent, immutable persistence for SOC2 and GDPR compliance.

L1: Grafana Loki

Structured JSON push. Fully queryable via LogQL. Standard retention of 30 days in production environments.

L0: Console & File

Serilog default sink. Used purely for active container debugging (STDOUT). Ephemeral.

```
2024-07-10 10:00:00 [INFO] Application started successfully.
2024-07-10 10:00:01 [DEBUG] Database connection established.
2024-07-10 10:00:02 [WARN] Configuration file not found, using defaults.
2024-07-10 10:00:03 [INFO] User authentication successful for user: admin.
2024-07-10 10:00:04 [ERROR] Failed to process request: null reference exception.
2024-07-10 10:00:05 [INFO] System health check passed.
```

```
2024-07-10 10:00:00 [INFO] Application started successfully.
2024-07-10 10:00:01 [DEBUG] Database connection established.
2024-07-10 10:00:02 [WARN] Configuration file not found, using defaults.
2024-07-10 10:00:03 [INFO] User authentication successful for user: admin.
2024-07-10 10:00:04 [ERROR] Failed to process request: null reference exception.
2024-07-10 10:00:05 [INFO] System health check passed.
```

Pillar 3: Distributed Tracing

UI Request (Browser)

API Gateway routing

ProcessEngine Workflow Execution

Database Query execution (System.Data.SqlClient auto-instrumentation)

Key Insight: Diagnose N+1 query problems and user-visible latency bottlenecks instantly—without redeploying code or sifting through fragmented logs.

The Health Check Matrix

Liveness Probe

`/health/live`

Question: Does the application process exist and is it breathing?



Action if failing: Orchestrator (Kubernetes) automatically kills and restarts the container.

Readiness Probe

`/health/ready`

Question: Can it accept traffic? (Validates Redis, Kafka, and SQL DB connections).



Action if failing: Load balancer drains traffic and removes the node from rotation to prevent cascading failures.

Alerting & Incident Response

Stage 1: Detection

Prometheus evaluates Alert Rules continuously (e.g., InstanceDown or KafkaLagHigh).

Stage 2: Routing

Prometheus AlertManager groups related alerts, suppresses duplicates, and manages silence periods.

Stage 3: Escalation

System dispatches targeted webhooks to external services like **PagerDuty** or **Slack** channels.

Stage 4: Human-in-the- -Loop

Primary On-Call Engineer is alerted. If no response is logged within 5 minutes, the system automatically escalates to the **Backup Schedule**.

Enterprise Multi-Tenancy at Scale





Tenant Isolation at the Infrastructure Level

Every single log, metric, and distributed trace is strictly enforced and enriched with a permanent **'tenant_id'** label.

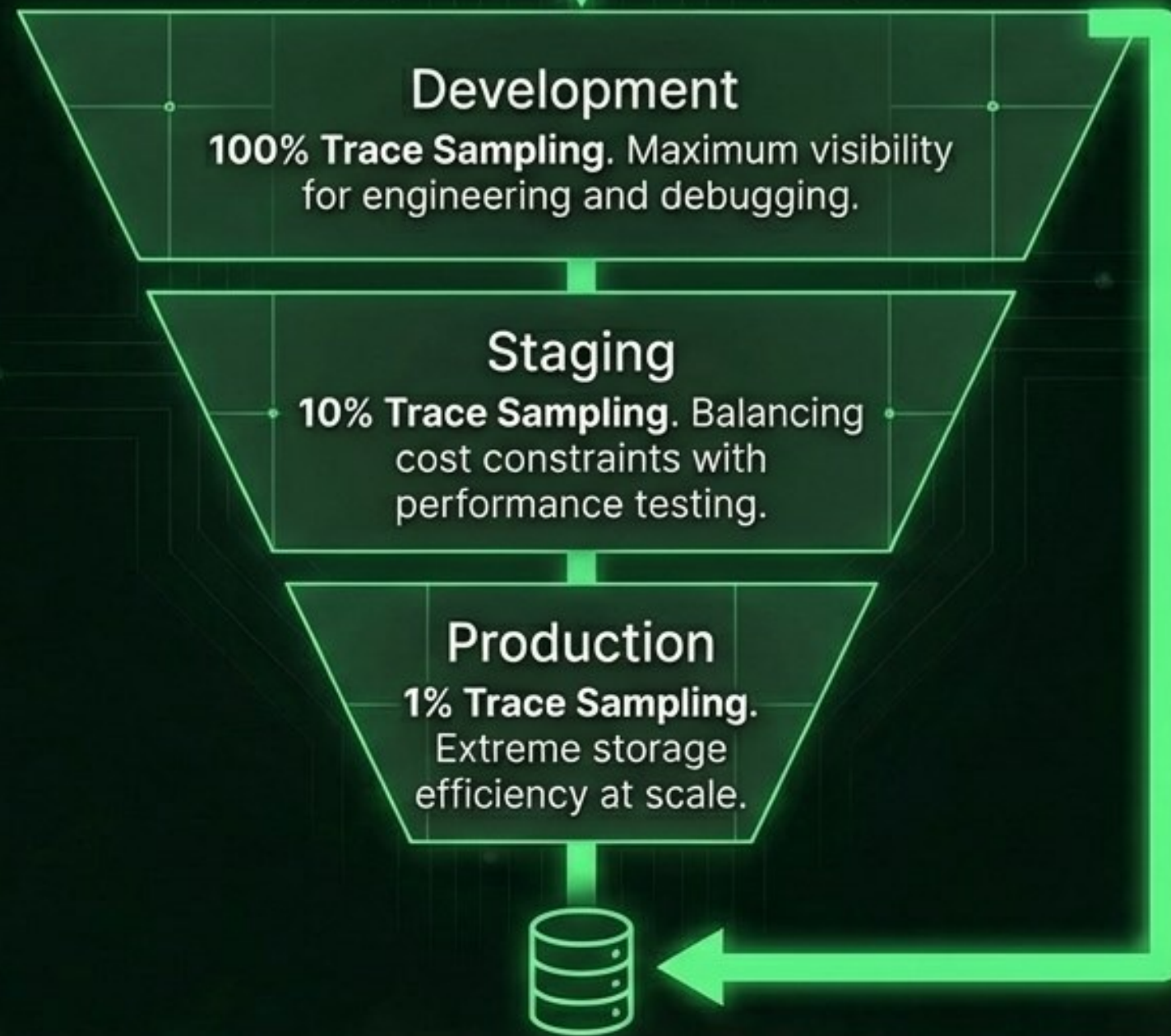
Grafana Dashboard Variables dynamically inject **'\$TenantId'** into all PromQL and LogQL queries. Users can exclusively visualize their authorized data scope.

This validated architectural isolation protects data integrity across all 605+ multi-tenant database tables within the BizFirst DataOcean.

The Incumbent Takedown Matrix

	Datadog	Splunk	BizFirst Observability
Cost Scaling	\$0.50+/GB ingested	Expensive linear scaling	Flat self-hosted cost 
Vendor Lock-in	High lock-in (proprietary)	Proprietary SPL language	Zero lock-in (Standard OTLP) 
Data Privacy	SaaS cloud only	Complex compliance mapping	Complete privacy (Stays on infra) 
Deployment	SaaS only	Complex implementation	On-prem / Hybrid / Cloud ready 

Implementation & Sampling Strategy



The Safety Net: Always-Sample-Errors

100% of failed traces are captured regardless of the environment, ensuring critical bugs are never lost to sampling logic.

The Future of Observability

Near Term

Service Mesh Integration

Native Istio integration enabling cross-service latency analysis and deep network policy observability.

Mid Term

eBPF-Based System Observability

Deep kernel-level system call tracing and advanced I/O performance analysis with zero instrumentation overhead.

Long Term

ML-Based Anomaly Detection

AI agents that automatically learn operational baselines, detect outliers, and forecast capacity needs without manual rule creation.

BizFirstAi: Illuminating the intelligent enterprise.