WeScale
Augmented Dev & Ops

Principes fondamentaux de l'AIOps

Pourquoi l'AIOps ?

QUATRE DRIVERS QUI RENDENT LES OPS HUMAINES SEULES IMPOSSIBLES

01

Volume de données

Logs, métriques, traces et events : plusieurs To/jour dans les grandes organisations.
Au-delà de la lecture humaine.

02

Complexité distribuée

Microservices, conteneurs, multi-cloud.
Une requête traverse 10–50 services.
Les liens de cause à effet se diluent.

03

Alert fatigue

Milliers d'alertes/jour.
Majorité non actionnables.
Désensibilisation et bugs ratés.

04

MTTR insoutenable

Détection lente, RCA manuelle.
Dépendance aux experts rares.
Coût direct sur la disponibilité.

AIOps vs DevOps, SRE, MLOps.

QUATRE DISCIPLINES, QUATRE INTENTIONS — À NE PAS CONFONDRE

AIOPS

ML + big data sur la donnée opérationnelle pour automatiser corrélation, détection, RCA et remédiation.

DEVOPS

Mouvement culturel et pratiques d'unification Dev + Ops (CI/CD, IaC, ownership) pour livrer plus vite et plus sûr.

SRE

Discipline d'ingénierie de la fiabilité (Google) : SLI/SLO, budget d'erreur, automatisation du toil.

MLOPS

Pratiques DevOps appliquées au cycle de vie ML : entraînement, déploiement, monitoring et retraining des modèles.

MELT — les 4 piliers de la donnée ops.

MELT POPULARISÉ PAR NEW RELIC (2019) — METRICS, EVENTS, LOGS, TRACES

M

Metrics

Mesures numériques agrégées sur intervalles : CPU, latence, RPS, taux d'erreur. Faible coût stockage.

E

Events

Faits discrets datés : déploiement, change, alerte, incident. Souvent modélisés en logs structurés.

L

Logs

Traces textuelles horodatées des activités systèmes / applicatives. Riches mais coûteux à stocker.

T

Traces

Parcours d'une requête à travers des services distribués (spans + IDs). Standard W3C Trace Context / OTel.

Monitoring vs Observabilité.

DEUX POSTURES COMPLÉMENTAIRES — PRÉ-REQUIS POUR L'AIOPS

MONITORING

Connu / contrôlé.

Répond à des questions définies à l'avance.
Seuils, dashboards, KPIs prédéfinis.
Bon pour les pannes connues, faible pour l'inconnu.
Souvent silotés (réseau, infra, app).

OBSERVABILITÉ

Inconnu / exploratoire.

Permet d'investiguer ce qui n'a pas été prévu.
Données riches : MELT + topologie + contexte.
Pré-requis pour l'AIOps : sans signal, pas d'IA.
Charity Majors : known-unknowns vs unknown-unknowns.

Les 5 capacités de l’AIOps

MODÈLE DE RÉFÉRENCE — INGESTION, ML, ANOMALIES, PRÉDICTION, AUTOMATION

C1
Ingestion

Multi-sources :
metrics, logs, traces,
events, topologie.

C2
ML & analyse

Statistiques, patterns,
clustering, NLP
sur tickets et logs.

C3
Détection d'anomalies

Baselines adaptatives,
isolation forest,
auto-encoders.

C4
Prédiction

Capacité, défaillance,
scoring de risque
sur changes.

C5
Automation

Runbooks, remédiation
closed-loop,
self-healing.

Réduction du bruit & corrélation d'événements.

USE CASE #1 GARTNER — TRANSFORMER 1 000 ALERTES EN 10 INCIDENTS

TECHNIQUES

Comment l'AIOps regroupe.

Déduplication des alertes identiques.
Clustering temporel (fenêtres glissantes).
Corrélation topologique (services dépendants).
Inférence causale + signature pattern.

BÉNÉFICES

Ce que ça change.

–40 % à –70 % de bruit (Forrester TEI typique).
Une seule notif par incident, pas 200.
Priorisation par impact métier.
Concentration des Ops sur ce qui compte.

Détection d'anomalies.

DÉPASSER LES SEUILS STATIQUES — APPRENDRE LE COMPORTEMENT NORMAL

POURQUOI

Les seuils ne suffisent plus.

Comportements saisonniers (jour/nuit, week-end).
Lancements, soldes, pics commerciaux.
Microservices = baselines par service.
Trop de tuning manuel sur des systèmes vivants.

TECHNIQUES

ML non supervisé en tête.

Baselines statistiques adaptatives.
Isolation Forest, DBSCAN, auto-encoders.
Apprentissage continu (sliding window).
Faible sensibilité au labelling humain.

Root Cause Analysis (RCA).

PASSER DE « ÇA NE MARCHE PAS » À « VOICI POURQUOI » EN MINUTES

01

Topologie

CMDB, service map, dépendances.
Sans graphe : analyse à plat,
incapable de remonter la chaîne.

02

Corrélation temporelle

Aligner alertes, déploiements,
changes et anomalies sur le temps.
Premier suspect : le change.

03

Inférence causale

Au-delà de la corrélation.
Probabilité que A cause B,
ranking des hypothèses.

04

RAG opérationnel

Historique d'incidents résolus,
runbooks, patterns connus.
L'IA propose, l'humain valide.

Analyse prédictive & automation intelligente.

DEUX BRAS DE LEVIER — ANTICIPER ET AGIR SANS L'HUMAIN

PRÉDICTIF

Anticiper l'incident.

Prévision de capacité (CPU, mémoire, IOPS).
Scoring de risque sur changes / déploiements.
Détection de signaux faibles avant SLO breach.
Préventif > curatif sur les coûts cloud (FinOps).

AUTOMATION INTELLIGENTE

Boucle fermée et runbooks adaptatifs.

Runbooks déclenchés sur signature d'incident.
Closed-loop : action + vérification + rollback.
Self-healing sur incidents récurrents connus.
Distinction clé : automation = tâche, orchestration = workflow.

Golden Signals + SLO + budget d'erreur.

GOOGLE SRE BOOK — CHAPITRE MONITORING DISTRIBUTED SYSTEMS

S1

Latence

Temps de réponse,
succès et erreurs
séparés.

S2

Trafic

Charge appliquée :
RPS, sessions actives,
throughput.

S3

Erreurs

Taux d'échec :
5xx explicites,
4xx ciblés, soft errors.

S4

Saturation

Combien le système
est plein : queues,
CPU, IOPS, descripteurs.

SLO

Cible de fiabilité.

Objectif mesurable dérivé des SLI (ex. 99,9 %).

BUDGET D'ERREUR

100 % - SLO.

L'indispo « autorisée ». Consommé = ralentir, geler. Restant = innover, prendre risque.

Maturité AIOps — 5 niveaux.

MODÈLE WESCALE — RÉACTIF → PROACTIF → PRÉDICTIF → AUTONOME

N1
Découverte

Co-construction.
Validation systématique
humaine.

N2
Exploratoire

L'agent agit,
chaque action soumise
à validation.

N3
Apprentissage

Généralisation
des patterns,
supervision modérée.

N4
Autonomie guidée

Traite seul
le connu, escalade
l'inconnu.

N5
Maturité

Agit fiablement
et autonomement,
rend compte.

Pièges & métriques de succès.

CE QUI FAIT ÉCHOUER — ET CE QUI PROUVE LE ROI

ANTI-PATTERNS

Les 5 pièges récurrents.

Démarrer sans baseline MTTR / bruit / coût.
Données silotées non normalisées (« garbage in »).
POC qui restent en sandbox (pas de prod).
Confiance aveugle — pas de revue, pas de rollback.
Tool sprawl : observabilité + AIOps + ITSM redondants.

MÉTRIQUES

Comment mesurer le succès.

Opérationnel : MTTR, % alertes auto-traitées, faux positifs.
Adoption : tâches déléguées à un agent, % code assisté.
Qualité agent : score confiance, taux escalade.
Fiabilité : SLO tenus, error budget consommé.
Forrester TEI typique : –40 à –70 % MTTR & bruit.