Managing costs on Azure Databricks

The issues in DataOps contribute to direct organizational costs attributed to

- DevOps - 5 hours/week
- Big data engineers - 10 hours/week
- DBU - $200K/year
- Azure Linux VMs - $200K/year
- ADLS - $150K/year
- Manufacturing: overestimating inventory leading to millions in waste.
- Retail: lapses in care and millions in additional healthcare costs.
- Healthcare: leading to millions in theft and fines.
- Banking: Fraudulent transactions in banking not discovered in time, leading to costly maintenance calls.
- Equipment failure not detected with accuracy,
- Patient profiles not accurately described, leading to lapses in care and millions in additional healthcare costs.
- Customer demand not analyzed for a particular product, Year 1
  - $0
  - $5,000,000
  - $12,500,000
  - $20,000,000
  - $37,500,000
- Year 2
  - $0
  - $5,000,000
  - $12,500,000
  - $20,000,000
  - $30,000,000
- Year 3
  - $4,880,000
  - $9,200,000
  - $12,319,593
  - $20,540,800
  - $10,570,211
- Year 4
  - $0
  - $5,000,000
  - $12,500,000
  - $20,000,000
  - $37,500,000

Required team:
- DevOps - 10 hours/week
- Big data engineers - 40 hours/week
- DBU - $400K/year
- Azure Linux VMs - $500K/year
- ADLS - $150K/year

Required infrastructure:
- Days to weeks
- Premium/Standard levels within each workload.
- Choosing proper Azure VM topology
- Tuning Azure Databricks is not a one-time effort, but an ongoing process of monitoring, optimizing, and improving performance.

Example: Unravel banking customer TCO

- Before Unravel: $0
- After Unravel: $4,880,000
- Before Unravel: $0
- After Unravel: $9,200,000
- Before Unravel: $4,880,000
- After Unravel: $9,200,000
- Before Unravel: $4,880,000
- After Unravel: $9,200,000

- Before Unravel: $0
- After Unravel: $4,880,000
- Before Unravel: $0
- After Unravel: $9,200,000
- Before Unravel: $4,880,000
- After Unravel: $9,200,000
- Before Unravel: $4,880,000
- After Unravel: $9,200,000

- Before Unravel: $0
- After Unravel: $4,880,000
- Before Unravel: $0
- After Unravel: $9,200,000
- Before Unravel: $4,880,000
- After Unravel: $9,200,000
- Before Unravel: $4,880,000
- After Unravel: $9,200,000

Unravel techniques for Azure Databricks tuning

- Unravel will drive $6.5MM in Human Capital savings over 4 years.
- Unravel will help in reducing the number of missed SLAs and their associated costs.
- Unravel will help in optimizing resource utilization.
- Unravel will help in reducing the time to resolution of issues.
- Unravel will help in increasing the efficiency of data pipeline execution.

True cost go beyond data infra cost

- Azure Databricks pricing quick guide
- Tuning Azure Databricks is not a one-time effort, but an ongoing process of monitoring, optimizing, and improving performance.
- DevOps/ITOps blame game
- Workload identification
- Persistent resource requirements (and cost) than expected
- Data analytics, data engineering, data engineering light
- Inability to choose proper Azure VM topology
- Code optimization
- Choose proper Azure VM topology
- Automated workloads to run robust jobs via API or UI
- Automated workloads to run robust jobs via API or UI

Interested in learning more?

Contact us at hello@unraveldata.com

© Unravel. All rights reserved. Unravel and the Unravel logo are registered trademarks of Unravel.