

AI를 활용한 커널 파라미터 최적화

0. Index

1. Kernel Parameter in Linux

2. Automatic Optimization of Linux Kernel Parameters with OpenAI

3. How to Build

4. Demo

5. Conclusion

0. Index

- 리눅스 시스템에서 커널 파라미터에 대한 소개
- kernel parameter가 서비스에 미치는 영향과 중요도에 대한 설명
- AI를 통한 kernel parameter 자동 최적화 방안
- AI를 통한 kernel parameter 자동 최적화 개발 및 로직
- Demo
- 결론

1. Kernel Parameter in Linux

- Kernel Parameter란
 - 커널 파라미터는 리눅스 커널이 시스템 리소스를 관리하고 동작하는 방식을 제어하는 설정값입니다. 이들은 /proc/sys 경로 아래에 존재하며, **네트워크, 메모리, 파일 시스템 등 다양한 커널 기능에 영향을 줍니다.**
- Kernel Parameter 튜닝의 중요성
 - 적절한 Kernel Parameter 튜닝은 시스템의 **성능과 안정성**에 큰 영향을 미칩니다.
 - ex)
 - **somaxconn**: TCP 연결 큐의 최대 길이를 설정합니다. 기본 값이 낮은 경우, 높은 트래픽에서 연결이 거부되거나 지연될 수 있습니다.

- **vm.swappiness**: 메모리 부족 시 스왑 사용 비율을 조절하는데, 기본값이 높을 경우 불필요한 디스크 I/O로 인해 성능 저하가 발생할 수 있음.

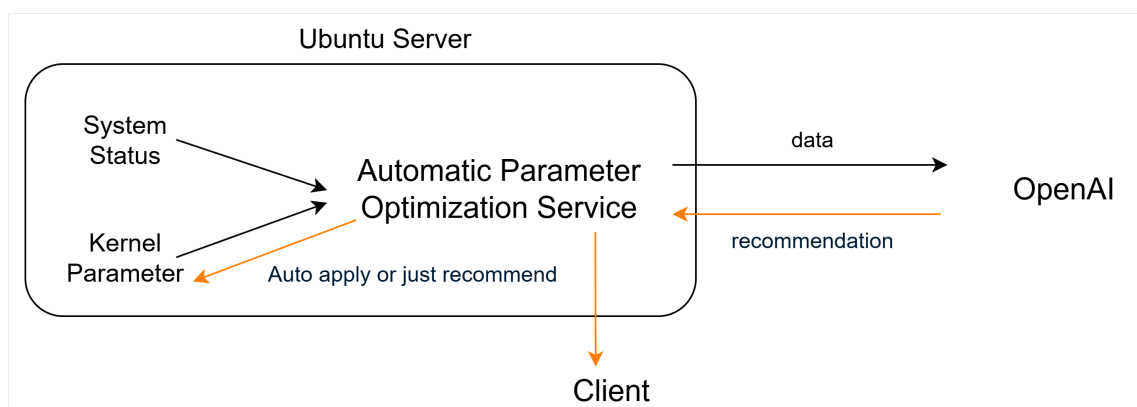
2. Automatic Optimization of Linux Kernel Parameters with OpenAI

- 운영자가 직접 Kernel Parameter를 관리할 때
 - 장점
 - 시스템에 대한 **이해와 경험을 바탕으로** 정확하게 필요한 parameter 값을 설정할 수 있음.
 - 단점
 - **실시간으로** 대응하기 어려움.
 - 어떤 kernel parameter를 변경해야 하는 지 **인식하지 못하는 경우도 있음**.



운영자가 직접 관리할 때의 단점을 보완하기 위해 AI를 활용해 데이터를 기반으로 최적의 커널 파라미터를 추천 및 자동 설정합니다.

- Base Architecture



3. How to Build

- 언어 : Golang
- 로직
 - 1분 간격으로 CPU, 메모리, 네트워크, I/O 등의 메트릭 수집
 - 10분 간격으로 수집 된 데이터를 OpenAI에 전송
 - 모델은 각 kernel parameter에 대해 아래와 같은 응답을 반환
 - 추천 값
 - 추천 이유
 - 현재 상태에서 자동 설정이 가능한지 여부
 - 반환 된 결과를 기반으로 다음과 같은 조치를 수행:
 - AI 판단이 안전할 경우 & 운영자가 미리 자동 설정을 허용한 parameter의 경우 → 자동 설정 적용
 - 위험할 경우 운영자에게 알림 전송 및 승인 대기 → 승인 완료 시 자동 적용

4. Demo

- **somaxconn default가 4096인 ubuntu server에서 Recv-q의 크기를 최대 값에 가까이 늘렸을 경우 → 늘리기를 추천 (자동 설정 가능하다고 판단)**
- **Recv-q의 크기를 다시 줄였을 경우 → 4096에서 변경할 필요 없다고 판단**

```

Metrics collected: {Timestamp:2025-05-16 04:17:18.001772222 +0900 KST m=+219.804011150 CPUIdlePercent:99.19986926331697 MemoryFreeMB:822 LoadAvg1Min:0.03 RecvQ:4050}
Metrics collected: {Timestamp:2025-05-16 04:17:38.010121115 +0900 KST m=+239.812360038 CPUIdlePercent:99.19983807448712 MemoryFreeMB:800 LoadAvg1Min:0.02 RecvQ:4050}
Metrics collected: {Timestamp:2025-05-16 04:17:58.0017064 +0900 KST m=+259.803945325 CPUIdlePercent:99.19980241764551 MemoryFreeMB:768 LoadAvg1Min:0.01 RecvQ:4050}
Metrics collected: {Timestamp:2025-05-16 04:18:18.010085193 +0900 KST m=+279.812324106 CPUIdlePercent:99.19976705448124 MemoryFreeMB:828 LoadAvg1Min:0.01 RecvQ:4050}
Metrics collected: {Timestamp:2025-05-16 04:18:38.004122576 +0900 KST m=+299.806361494 CPUIdlePercent:99.19974099359723 MemoryFreeMB:832 LoadAvg1Min:0.15 RecvQ:4050}
Metrics collected: {Timestamp:2025-05-16 04:18:58.010330333 +0900 KST m=+319.812569249 CPUIdlePercent:99.1997184487478 MemoryFreeMB:884 LoadAvg1Min:0.38 RecvQ:1000}
Metrics collected: {Timestamp:2025-05-16 04:19:18.001642542 +0900 KST m=+339.803881468 CPUIdlePercent:99.19970479608213 MemoryFreeMB:877 LoadAvg1Min:0.27 RecvQ:1000}
Metrics collected: {Timestamp:2025-05-16 04:19:38.000709088 +0900 KST m=+359.802948014 CPUIdlePercent:99.19969029460496 MemoryFreeMB:857 LoadAvg1Min:0.19 RecvQ:1000}
Metrics collected: {Timestamp:2025-05-16 04:19:58.019040035 +0900 KST m=+379.821278951 CPUIdlePercent:99.19967387651548 MemoryFreeMB:819 LoadAvg1Min:0.2 RecvQ:1000}
Metrics collected: {Timestamp:2025-05-16 04:20:18.014076187 +0900 KST m=+399.816315101 CPUIdlePercent:99.19966453951952 MemoryFreeMB:954 LoadAvg1Min:0.14 RecvQ:0}
OpenAI Response: [{Parameter:net.core.somaxconn RecommendedValue:8192 CanAutoApply:true Reason:Recv-Q 값이 지속적으로 높은 상태(4050)를 유지하고 있으며, 이는 대기 중인 연결 요청이 많음을 나타냅니다. 현재 설정된 somaxconn 값(4096)이 이 수요를 충족시키기에 충분하지 않을 수 있습니다. 따라서 SYN backlog을 방지하고 더 많은 연결 요청을 처리하기 위해 somaxconn 값을 8192로 증가시키는 것이 좋습니다.}]
Metrics collected: {Timestamp:2025-05-16 04:20:38.016670647 +0900 KST m=+419.818909572 CPUIdlePercent:99.19965326123494 MemoryFreeMB:937 LoadAvg1Min:0.1 RecvQ:0}
Metrics collected: {Timestamp:2025-05-16 04:20:58.000101791 +0900 KST m=+439.802340720 CPUIdlePercent:99.19963694648301 MemoryFreeMB:913 LoadAvg1Min:0.07 RecvQ:0}
Metrics collected: {Timestamp:2025-05-16 04:21:18.004039415 +0900 KST m=+459.806278336 CPUIdlePercent:99.19962078070277 MemoryFreeMB:943 LoadAvg1Min:0.05 RecvQ:0}
Metrics collected: {Timestamp:2025-05-16 04:21:38.000443372 +0900 KST m=+479.802682301 CPUIdlePercent:99.19959891351391 MemoryFreeMB:936 LoadAvg1Min:0.04 RecvQ:0}
Metrics collected: {Timestamp:2025-05-16 04:21:58.006678976 +0900 KST m=+499.808917895 CPUIdlePercent:99.19957469450588 MemoryFreeMB:904 LoadAvg1Min:0.02 RecvQ:0}
Metrics collected: {Timestamp:2025-05-16 04:22:18.002248106 +0900 KST m=+519.804487031 CPUIdlePercent:99.19955533252667 MemoryFreeMB:931 LoadAvg1Min:0.02 RecvQ:0}
Metrics collected: {Timestamp:2025-05-16 04:22:38.004765715 +0900 KST m=+539.807004645 CPUIdlePercent:99.19953847790717 MemoryFreeMB:908 LoadAvg1Min:0.16 RecvQ:0}
Metrics collected: {Timestamp:2025-05-16 04:22:58.02017311 +0900 KST m=+559.822412026 CPUIdlePercent:99.19951441344195 MemoryFreeMB:898 LoadAvg1Min:0.11 RecvQ:0}
Metrics collected: {Timestamp:2025-05-16 04:23:18.004674393 +0900 KST m=+579.806913323 CPUIdlePercent:99.19949588881902 MemoryFreeMB:915 LoadAvg1Min:0.08 RecvQ:0}
Metrics collected: {Timestamp:2025-05-16 04:23:38.014139392 +0900 KST m=+599.816378307 CPUIdlePercent:99.19947597422716 MemoryFreeMB:915 LoadAvg1Min:0.18 RecvQ:0}
OpenAI error: failed to parse recommendation: invalid character ''' looking for beginning of value
raw: ``json
[
{
  "parameter": "net.core.somaxconn",
  "recommended_value": 4096,
  "can_auto_apply": true,
  "reason": "현재 Recv-Q가 모두 0으로, 대기 중인 연결이 없으며, CPU와 메모리 사용률이 낮아 somaxconn 값을 변경할 필요가 없습니다."
}
]

```

5. Conclusion

- AI를 활용한 커널 파라미터 자동 최적화의 장점

- 데이터 기반의 튜닝

- 운영자의 경험에 의존하지 않고, 실시간으로 수집 되는 데이터를 기반으로 파라미터 설정 가능

- 실시간 대응

- 주기적으로 작동하기에 실시간 대응이 가능함

- 일관성 확보

- 동일한 상황일 때 동일한 판단을 내릴 수 있음

- 한계점

- 설정 실패 시 리스크 존재

- 자동 설정에 대한 리스크를 최대한 줄였지만, 이로 인해 서비스의 장애가 유발될 수 있다는 점은 피할 수 없음

- 해당 시스템에 대한 이해도 부족

- 오랜 시간 시스템에 대한 이해도를 쌓아온 운영자와 다르게 AI는 일반적인 판단 밖에 내릴 수 없음

- **추후 발전 가능성**

- 시스템 특성 별로 AI 모델을 분리 한 뒤, 실제 성능 개선 여부를 지속적으로 학습에 반영한다면 치명적인 시스템에 대한 이해도 부족을 점차 줄여갈 수 있을 것 같음.