Tableau Server 10.0

고가용성: 규모에 맞는 필수 분석 기능 제공

Kitty Chou, 제품 관리자
Mike Klaczynski, 제품 마케팅 담당자

목차

업무에 중요한 셀프 서비스 분석 3

고가용성 이해 3

Tableau Server 확장성 4

즉각적인 고가용성 지원 4

Tableau Server 고가용성 이해 5

- 일반 5
- 게이트웨이 5
- 애플리케이션 서버 6
- 조정 서비스 6
- 클러스터 컨트롤러 7
- 리포지토리 7
- 백그라운더 9
- 데이터 서버 9
- 캐시 서버 9
- 데이터 엔진 9
- 파일 저장소 10
- 검색 및 찾아보기 11
- VizQL 서버 11

타사 모니터링 도구와 통합 11

주 서버 노드의 페일오버 12

주 서버의 백업 구성 13

클러스터 상태 모니터링 14

아키텍처 고려 사항 15

최적의 구성 선택 15

- 최소 3노드 HA 배포 16
- 3노드 이상의 배포 17

Tableau Server의 재해 복구 18

가장 간단한 DR 전략 18

타사 DR 솔루션 18

Tableau DR 기능 18

고가용성 및 기타 고려 사항 19

업무에 중요한 셀프 서비스 분석

오늘날 그 어느 때보다 셀프 서비스 분석과 데이터 기반 의사 결정이 전세계적으로 업계 표준이 되고 있습니다. 사용자와 의사 결정자들은 질문에 실시간으로 답하기 위해 데이터에 대한 즉각적인 액세스와 셀프 서비스 도구에 의존하게 되었습니다. 경영진은 회사의 데이터 기반 의사 결정의 중요성을 인식하고, 이러한 시스템을 매일 필요로 하고 있습니다. 데이터에 대한 이러한 의존도는 기초 시스템에 높은 수준의 가용성을 요구합니다. 플랫폼의 기능은 액세스가 더 쉬워지고, 기존 팀과 엔터프라이즈 도구에서 쉽게 구성할 수 있어야 합니다.

Tableau Server 10.0은 미래의 필수 셀프 서비스 분석 기능을 제공합니다. 신속한 셀프 서비스 데이터 탐색을 지원하고, 강력한 거버넌스로 콘텐츠와 데이터의 신뢰도를 높이며, 배포, 관리 및 확장이 용이합니다. 이 문서에서 Tableau Server 10.0이 어떻게 고가용성을 통해 규모에 맞는 셀프 서비스 분석을 제공하는지 알아보겠습니다.

고가용성 이해

가용성이 높은 시스템의 목표는 시스템 다운타임을 최소화하는 것입니다. 가용성은 흔히 '숫자 9의 개수'로 표현되며 실제 업타임 대비 예상 업타임의 비율로 측정됩니다. 아래 표는 연간 다운타임이 어떻게 숫자 9의 개수로 표시되는지 보여줍니다.

숫자 9의 개수	가용성 백분율	연간 총 중단 시간
1	90%	36.5일
2	99%	3일과 15시간
3	99.9%	8시간 45분
4	99.99%	52분 34초
5	99.999%	5분 15초

그림 1: 일반적인 가용성 지표 및 연간 다운타임

시스템 관리자는 비즈니스 사용자와 서비스 수준 계약(SLA)에 대해 협의하여 허용 가능한 다운타임 임계값을 정의하고, 협의된 SLA 목표를 달성할 수 있는 배포 아키텍처를 선택할 것입니다. 대부분의 시스템 관리자는 유지 관리, 업그레이드 및 패치를 위해 다운타임을 계획합니다. 또한, 예상치 못한 장애로 인해 계획되지 않은 다운타임이 발생할 수 있습니다. 계획되지 않은 다운타임을 최소화하기 위해서는 하드웨어 또는 소프트웨어에 필요한 유지 관리 작업을 계획하고 실행해야 합니다.

Tableau는 사용자가 순조롭게 데이터를 보고 이해하는 것이 얼마나 중요한지 잘 알고 있습니다. 또한, 하드웨어, 소프트웨어, 네트워크, 사용자 실수 등의 문제로 인해 비즈니스 인텔리전스 시스템의 가용성이 위협받는 경우가 늘 있을 것이라는 사실도 인지하고 있습니다. Tableau Server 10.0 프로세스는 구성 요소에 장애가 발생하는 경우 프로세스를 자동으로 다시 시작하여 시스템 실행 상태를 계속 유지합니다. 적절하게 구성된 멀티 노드 배포에서는 중복 프로세스를 통해 서버 고가용성(HA)을 실현합니다. 대부분의 시스템과는 달리 Tableau에서는 HA를 지원하는 분석 환경을 쉽게 구현할 수 있습니다.

Tableau Server 확장성

Tableau Server는 확장이 용이하도록 설계되어 있습니다. 대규모 조직에는 엔터프라이즈급의 배포 안정성을 제공하고, 소규모 팀에는 간단하고 사용이 손쉬운 기능을 제공하여 요구를 충족하고 있습니다. 사용자 요구 및 HA 요구 사항에 따라 Tableau Server를 여러 컴퓨터에서 실행하거나, 동일 노드에서 여러 개의 프로세스를 실행할 수 있습니다.

Tableau는 내부적으로 Tableau Public, Tableau Online 등의 여러 사이트에 클라우드급으로 Tableau Server를 배포하여 사용하고 있습니다. Tableau Server를 맞춤화하여 배포한 Tableau Public은 전 세계에서 매주 수백만 뷰를 지원하고 있습니다. Tableau에서는 개발 및 출시 과정의 일부로 Tableau Server의 베타 버전을 Tableau Public에 배포하여 공개 출시에 앞서 안정성과 품질을 점검합니다.

즉각적인 고가용성 지원

Tableau Server를 설치하고 고가용성 기능을 구현하는 작업은 쉽습니다. 기본 설치는 몇 분 만에 완료됩니다. Tableau Server를 주 컴퓨터와 클러스터의 모든 작업자 노드에 설치한 다음, 구성 유틸리티를 통해 고가용성 클러스터로 구성하여 운영할 수 있습니다.

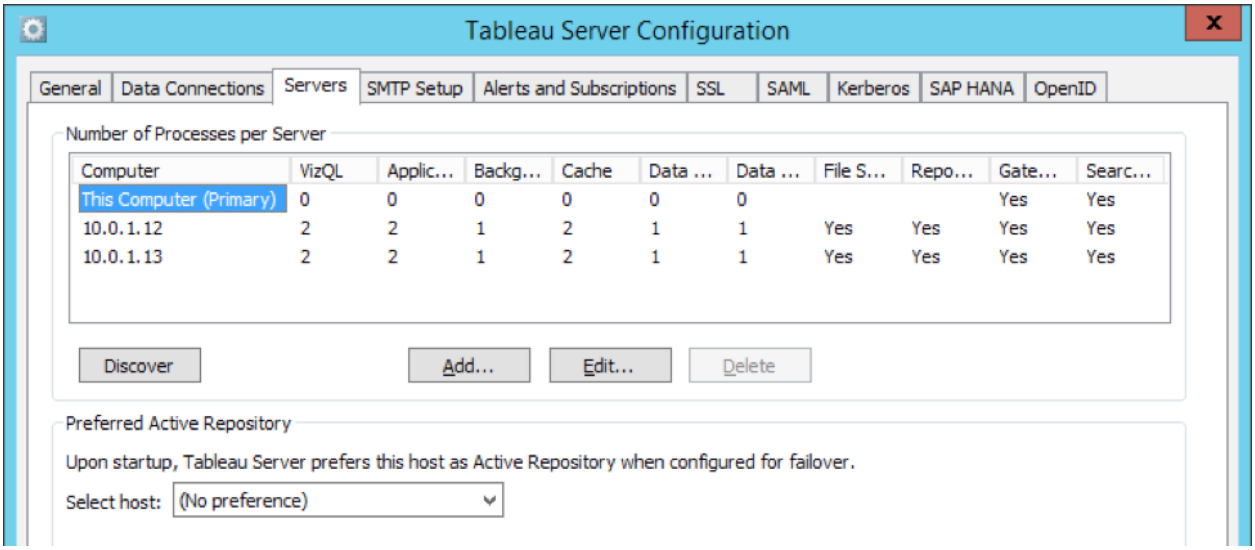

그림 2: Tableau Server 구성 유틸리티를 사용하면 유연하게 노드 수와 구성을 조정하여 클러스터의 노드 당 프로세스 수를 제어할 수 있습니다.

Tableau Server 고가용성 이해

Tableau Server에는 시스템 전반에서 최종 사용자의 액세스를 지원하는 여러 개의 전용 프로세스가 있습니다. 이 섹션은 Tableau Server의 각 구성 요소와 기능에 익숙하다는 가정 하에 준비되었습니다. Tableau Server의 구성 요소 및 기능에 대해 잘 모르는 경우 먼저 Tableau Server 관리 가이드를 읽어보시기 바랍니다.

Tableau Server를 고가용성 구조로 만드는 방법은 각 구성 요소를 고가용성으로 만드는 방법과 사실상 같습니다. 즉, 전체 Tableau Server 클러스터에 고가용성이 제공되려면 단일 장애 지점 모두에 중복성을 제공하여 모든 구성 요소의 고가용성을 보장해야 합니다. 각각의 구성 요소를 차례대로 살펴보겠습니다.

일반

Tableau Server는 내장 자동화 기능으로 모든 서버 프로세스를 다시 시작합니다. 이 자동화 기능은 중단된 서버 프로세스를 자동으로 다시 시작하고 관리자에 알려줌으로써 고가용성을 보장합니다. Tableau Server가 있는 하드웨어 또는 가상 컴퓨터는 이 자동 기능이 실행될 수 있도록 정상적인 상태여야 합니다.

전체 노드를 장애로부터 보호하려면, 클러스터의 서로 다른 노드 간에 중복성을 제공하도록 서버 프로세스를 개별적으로 구성해야 합니다. 이러한 중복 구성은 주 노드에서만 실행 가능한 라이선스 서비스를 제외하고 모든 프로세스에서 구현 가능합니다. 이 프로세스의 제한 사항은 *주 서버 노드의 페일오버* 섹션에 상세히 설명되어 있습니다.

게이트웨이

Tableau Server 8.1 버전부터 게이트웨이 프로세스는 주 Tableau Server 노드 뿐만이 아니라 Tableau Server의 모든 노드에서 실행될 수 있습니다. 이 사실은 고가용성에 중요한 영향을 미칩니다.

Tableau Server 8.1 이전 버전에서는 오직 하나의 노드에서만 게이트웨이 프로세스를 실행하도록 구성할 수 있었기 때문에, 해당 프로세스가 중단되면 Tableau Server는 클러스터 외부의 다른 구성 요소와 통신할 수 없으므로 내부 통신에도 장애가 발생했습니다. Tableau Server 10.0에서는 여러 노드가 게이트웨이 프로세스를 실행하도록 구성되면, 최소한 하나의 게이트웨이 프로세스가 실행 중이면 서버와의 통신은 계속 유지됩니다.

게이트웨이 장애 위험 경감

게이트웨이 고가용성은 Tableau Server 클러스터에 둘 이상의 노드를 만들어 둘 이상의 게이트웨이 프로세스가 실행되도록 구성하는 것이 핵심입니다. 각 노드에 하나의 게이트웨이 프로세스를 구성하는 것이 좋습니다. 그러면 게이트웨이 프로세스로 인한 단일 장애 지점의 위험이 줄어들고 서비스 가용성이 높아집니다.

게이트웨이 프로세스가 중단되면 어떤 문제가 발생할까요? 앞에서 언급한 바와 같이 실행 중인 게이트웨이 프로세스가 없으면 전체 Tableau Server 클러스터는 사용할 수 없게 됩니다. 계속 실행 중인 다른 게이트웨이 프로세스가 있다면, 실행 중인 게이트웨이로 보내진 요청이 정상적으로 처리됩니다. 그렇지만 장애가 있는 게이트웨이에서 수신한 모든 요청은 다른 정상적인 게이트웨이가 존재해도 다시 전달되지 않고 중단 상태로 계속 유지됩니다. 컴퓨터가 정상적으로 작동하는 경우 중단된 게이트웨이 프로세스는 자동으로 다시 시작되고, 요청 처리는 재개됩니다.

게이트웨이 장애에 보다 능동적으로 대응하도록 하려면, Tableau Server 클러스터 앞에 외부 로드 밸런서(ELB)를 두어 정상적으로 작동하는 게이트웨이 프로세스로만 요청이 전달되도록 구성하는 것이 좋습니다. 이러한 설정에 대한 상세한 내용은 Tableau Server 관리 가이드을 참조하시기 바랍니다.

애플리케이션 서버

애플리케이션 서버를 사용하면 고가용성 기능을 손쉽게 구현할 수 있습니다. Tableau Server 클러스터의 각 노드에 애플리케이션 서버 인스턴스를 구성하기만 하면 됩니다.

애플리케이션 서버 프로세스가 중단되면 어떤 일이 발생할까요? 해당 인스턴스가 처리 중인 요청은 실패하지만 후속 요청은 작동 중인 다른 애플리케이션 서버 프로세스에 전달됩니다. 장애가 있는 애플리케이션 서버가 포함된 노드가 계속 실행 중일 경우, 중단된 프로세스는 수 초 내에 자동으로 다시 시작됩니다.

조정 서비스

조정 서비스는 Tableau Server 기본 설치에 포함되어 있습니다. 조정 서비스가 설치되는 노드 수는 다음 표와 같이 클러스터의 노드 수에 따라 다릅니다.

클러스터의 노드 개수	조정 서비스가 있는 노드 개수
1 - 2	1
3 - 4	3
5개 이상	5

조정 서비스 프로세스는 주 노드를 포함하여 클러스터의 첫 번째 *n* 노드에 순차적으로 설치되며, 여기서 *n*은 위의 표와 같이 조정 서비스가 설치되는 노드 수입니다.

쿼럼 이해

실행 중인 조정 서비스 프로세스의 수(쿼럼 구성 수)가 충분하지 못할 경우 Tableau Server가 완전히 중단되며, 이 수는 구성된 전체 조정 서비스 프로세스의 수에 따라 다릅니다. 쿼럼은 절대다수를 이르는 다른 말입니다. 3~4대의 컴퓨터가 연결된 하나의 클러스터는 최대 하나의 노드, 즉 하나의 조정 서비스 인스턴스의 중단을 허용할 수 있으며, 5~6대의 컴퓨터가 연결된 클러스터는 두 개 까지의 조정 서비스 프로세스의 중단을 허용할 수 있습니다.

여기서, 단 두 개의 노드가 있는 클러스터는 단 하나의 조정 서비스 프로세스의 중단도 허용할 수 없는데, 이것이 바로 자동 페일오버를 포함한 완전한 고가용성 기능 지원에 최소 3개의 노드가 필요한 이유입니다.

조정 서비스 프로세스가 중단되면 어떤 일이 발생할까요? 남아 있는 조정 서비스 프로세스의 수가 쿼럼을 구성하기에 충분한 경우 아무런 문제가 없습니다. 정상 작동하는 조정 서비스 프로세스의 수가 쿼럼보다 적어지면 기반 Postgres 데이터베이스의 참조 무결성 보호를 위해 전체 Tableau Server 클러스터는 사용할 수 없게 됩니다.

중단된 조정 서비스 프로세스는 컴퓨터가 정상적으로 작동하면 자동으로 다시 시작됩니다.

클러스터 컨트롤러

클러스터 컨트롤러 또한 Tableau Server 10.0 기본 설치에 포함되어 있습니다. 설치 중에 별다른 구성 절차 없이, 클러스터의 각 노드 당 하나의 클러스터 컨트롤러 프로세스가 구성됩니다.

클러스터 컨트롤러 프로세스가 중단되면 어떤 일이 발생할까요? 동일 노드에 있는 다른 모든 Tableau Server 구성 요소가 사용할 수 없게 되어 Tableau Server 상태 페이지에 '사용 불가능'으로 표시됩니다. 해당 노드에서 실행되는 리포지토리 프로세스 또한 사용할 수 없게 됩니다. (이 결과에 대한 자세한 정보는 아래 *리포지토리* 섹션을 참조하십시오.) 클러스터 컨트롤러 프로세스 중단으로부터 시스템을 보호하려면, 각 개별 서버 구성 요소에 중복성을 제공하여 클러스터 내에 있는 최소 두 개의 다른 노드에서 프로세스가 실행되도록 구성합니다.

중단된 클러스터 컨트롤러 프로세스는 컴퓨터가 정상적으로 작동하면 자동으로 다시 시작됩니다. 클러스터 컨트롤러가 다시 시작되면, 해당 노드에 구성된 리포지토리 프로세스도 다시 시작됩니다.

리포지토리

리포지토리는 Tableau Server가 작동하는 데 필요한 데이터베이스입니다. 정상적으로 작동하는 리포지토리가 없는 경우, 전체 Tableau Server 클러스터는 사용할 수 없게 됩니다. 항상 정상적으로 작동하는 단일 리포지토리가 사용되고 있으며, 이 리포지토리는 '액티브' 리포지토리로 불립니다. 리포지토리와 연관된 모든 작업은 이 액티브 리포지토리에서 실행됩니다.

가용성 향상을 위해 '패시브' 리포지토리가 있는 Tableau Server를 클러스터의 다른 노드에 구성할 수 있습니다. 액티브 리포지토리의 콘텐츠는 패시브 리포지토리에 계속 전달됩니다. 액티브 리포지토리에 장애가 생기면 고가용성으로 구성된 클러스터에서 패시브 상태의 리포지토리가 액티브 상태의 리포지토리로 자동 전환되어 서버 가용성을 계속 유지합니다. 그러므로, 고가용성을 보장하려면 반드시 패시브 리포지토리를 구성해야 합니다. 전체 클러스터에는 액티브 및 패시브 상태인 최대 두 개의 리포지토리가 있을 수 있으며, 이들은 동일 노드에 위치할 수 없습니다.

클러스터 컨트롤러는 리포지토리의 시작 및 종료, 액티브에서 패시브로의 페일오버를 관리합니다. 그러므로 클러스터 컨트롤러 프로세스의 중단이 리포지토리 문제를 야기할 수도 있습니다. 리포지토리를 시작시킨 클러스터 컨트롤러 프로세스가 중단되면 해당 리포지토리에도 장애가 발생합니다.

리포지토리 프로세스가 중단되면 어떤 일이 발생할까요? 상황에 따라 다릅니다. 다음 목록에 다양한 사례가 정리되어 있습니다.

- 패시브 리포지토리가 중단되는 경우 사용자에게 미치는 영향은 없습니다. 액티브 리포지토리가 계속 작동하고 있으므로 모든 기능이 정상 작동합니다. 패시브 리포지토리가 액티브 리포지토리와 완전하게 동기화될 때까지 지연 시간이 있을 수 있기는 하지만 백그라운드로 패시브 리포지토리가 다시 시작되고, 데이터 복제가 재개될 것입니다.

- 액티브 리포지토리가 중단된 상태에서 완전히 동기화된 패시브 리포지토리가 없다면, Tableau Server는 액티브 리포지토리가 다시 시작될 때까지 사용할 수 없게 됩니다. 시스템에서 자동 복구를 시도하려 하겠지만, 장애 요인에 따라 가능하지 않을 수도 있습니다.
  참고: 현재 액티브 리포지토리만이 사용 가능한 경우, Tableau Server는 고가용성으로 구성되지 않은 것입니다. 동기화된 패시브 리포지토리가 없다면, 액티브 리포지토리는 전체 시스템의 단일 장애 지점이 됩니다.

- 액티브 리포지토리가 중단된 상태에서 완전히 동기화된 패시브 리포지토리가 있는 경우, 클러스터가 고가용성으로 구성되어 있다면 패시브 리포지토리로의 페일오버가 자동으로 시작됩니다. 페일오가 완료되면, 이전의 패시브 리포지토리가 새로운 액티브 리포지토리가 됩니다. 시스템에서는 중단된 이전 액티브 리포지토리가 다시 시작되어 새로운 패시브 리포지토리가 되고 동기화를 시작합니다. 또한 시스템에서 다른 연관 프로세스들이 자동으로 다시 시작되어 새로운 액티브 리포지토리를 인식하고 다시 연결하게 됩니다. 재시작이 일어나는 이 짧은 시간 동안 사용자에게 서비스 중단이 발생합니다. 이러한 모든 작업은 자동적으로 이루어지므로 지속되는 가용성 보장을 위한 관리자 개입은 필요하지 않습니다. 패시브 리포지토리가 액티브 리포지토리를 대체하도록 수동으로 전환하려면, tabadmin failoverrepository 명령어를 사용하면 됩니다.

백그라운더

백그라운더 서비스를 고가용성 구성으로 만들려면, 둘 이상의 백그라운더 프로세스가 클러스터의 여러 노드에서 실행되도록 구성해야 합니다. 실행할 백그라운더의 위치 및 개수를 결정하기 전에 설치된 다른 서버 프로세스를 확인한 후 각 컴퓨터의 가용 용량을 판단하는 것이 좋습니다.

백그라운더 프로세스가 중단되면 어떤 일이 발생할까요? 백그라운드에서 실행되는 작업이 중단되고 다시 수행되지 않습니다. 대부분의 백그라운드 작업은 정기적으로 실행되도록 예약되어 있으므로, 정상적으로 작동하는 백그라운더 프로세스는 다음 예약 시간에 동일한 백그라운드 작업을 선택하여 정상적으로 수행합니다.

컴퓨터가 정상적으로 작동하는 경우 중단된 백그라운더 프로세스는 자동으로 다시 시작되지만, 중단된 작업은 다시 수행되지 않습니다.

데이터 서버

데이터 서버를 고가용성으로 만들려면, 하나 이상의 추가 데이터 서버 프로세스를 구성하여 클러스터의 여러 노드에서 실행하면 됩니다.

데이터 서버 프로세스가 중단되면 어떤 일이 발생할까요? 데이터 서버 프로세스에서 프록시를 통해 실행되는 쿼리에 문제가 생기면 뷰 렌더링 작업이 실패합니다. 요청 경로를 재지정할 수 있는 정상적인 데이터 서버가 있는 경우 실패한 작업의 재시도를 포함한 후속 요청에 대한 작업은 수행됩니다.

Tableau Server는 데이터 서버와 독립적으로 실행되지만, 데이터 서버가 실행되지 않으면 클러스터에서 통합 문서의 외부 데이터 원본 연결을 중개하지 못합니다. 데이터 원본 사용에 데이터 서버가 사용되지 않는 뷰는 계속해서 올바르게 작동합니다.

캐시 서버

외부 쿼리 캐시를 공유하는 캐시 서버는 키/값 쌍 형식으로 이전 쿼리 정보를 유지하므로 향후 요청을 더욱 빠르게 처리할 수 있습니다. 이는 하나 또는 모든 캐시 서버 프로세스를 사용할 수 없는 경우에도 그 영향이 상대적으로 적다는 것입니다. Tableau Server는 계속 작동하지만 동작 속도는 사전에 캐시된 결과가 없으므로 상대적으로 느려질 수 있습니다. 쿼리가 다시 실행되면, 다시 시작된 캐시 서버가 캐시에 데이터를 다시 채우면서 처리 속도도 빨라집니다. 캐시 서버는 가용성에는 영향을 미치지 않지만, 최종 사용자의 성능에는 영향을 줍니다. 다른 모든 프로세스와 같이 컴퓨터가 정상적으로 작동하는 경우 캐시 서버도 자동으로 다시 시작됩니다.

데이터 엔진

데이터 엔진 구성 요소는 인메모리 분석에서 데이터 추출을 로드하고 쿼리를 수행합니다. 데이터 엔진을 고가용성으로 만들려면, 단순히 하나 이상의 추가 데이터 엔진 프로세스를 구성하여 클러스터의 여러 노드에서 실행하면 됩니다. 모든 데이터 엔진은 액티브/액티브 모드로 실행되며 정확히 같은 기능을 수행합니다. 여기에서 중요한 사항은 데이터 엔진 프로세스를 실행하도록 구성된 모든 노드는 또한 파일 저장소 프로세스를 실행하도록 구성된다는 점입니다. 파일 저장소 프로세스는 추출 파일의 저장 및 복제를 관리하며, 자세한 내용은 다음 섹션에 설명되어 있습니다.

데이터 엔진 프로세스가 중단되면 어떤 일이 발생할까요? 데이터 엔진 프로세스에서 실행되던 쿼리에 문제가 생기면 뷰 렌더링이 실패하거나 추출 새로 고침이 실행되지 않습니다. 동일 작업이 다시 실행되면 정상적인 다른 데이터 엔진으로 작업이 다시 할당됩니다.

컴퓨터가 정상적으로 작동하는 경우 중단된 데이터 엔진 프로세스는 자동으로 다시 시작되어 운영됩니다.

파일 저장소

앞서 언급했듯이, 하나 이상의 데이터 엔진 프로세스를 실행하는 노드에서 파일 저장소 프로세스의 인스턴스가 자동으로 생성됩니다. 파일 저장소 프로세스는 추출 파일의 노드 간 저장 및 복제를 관리합니다.

파일 저장소의 기능

추출 파일이 Tableau Server에 처음으로 게시되거나 추출 새로 고침이 발생하면 시스템에 추출 파일이 생성됩니다. 이러한 이벤트 중 하나가 발생하는 즉시 단일 노드의 단일 파일 저장소에 추출이 저장됩니다. 이 추출 파일은 중복하여 저장되지 않았으므로 단일 장애 지점이 되며, 고가용성으로 구성되지 않았습니다. 이후 해당 파일 저장소 프로세스는 다른 파일 저장소 프로세스와 통신하여 로컬 추출을 클러스터 내의 다른 모든 파일 저장소 노드에 신속하게 복제합니다. 파일 저장소 프로세스는 클러스터 네트워크 리소스가 허용하는 최대 속도로 파일을 복사하도록 설계되었지만, 추출의 크기에 따라 소요 시간은 달라질 수 있습니다. 클러스터의 여러 노드에서 복사본이 사용 가능하게 되면, 추출 파일은 고가용성으로 구성된 것입니다.

파일 저장소 프로세스가 중단되면 어떤 일이 발생할까요? 여기에서 다음 두 가지 결과를 예상할 수 있습니다.

- 영향을 받는 노드에서 추출 파일의 복사가 중단됩니다.
- 영향을 받는 노드에서 더 이상 필요하지 않은 추출 파일의 제거가 중단됩니다. 이 제거 프로세스는 보통 '추출 제거 작업'이라고도 합니다.

추출 제거 작업의 중단은 시스템에 별다른 영향이 없습니다. 단지, 원하지 않는 추출 파일이 누적되면 해당 노드의 디스크 공간이 줄어들 뿐입니다. 적절한 용량의 디스크 공간이 할당된 노드에는 일반적으로 이로 인해 문제가 야기되지는 않습니다.

파일 복제 실패는 장애가 있는 파일 저장소 노드에 추가된 새 추출 파일을 클러스터 내의 다른 모든 파일 저장소 노드에서 사용할 수 없게 된다는 뜻입니다. 파일 저장소 프로세스가 다시 시작되면, 해당 시스템에서 모든 노드의 파일 저장소 간에 동기화를 수행하여 자체적으로 문제를 수정합니다.

중단된 파일 저장소 프로세스는 컴퓨터가 정상적으로 작동하는 경우 자동으로 다시 시작됩니다. 파일 저장소 프로세스는 신속하게 재개되며, 모든 파일의 동기화를 포함하여 중단이 발생한 시점 중이나 그 이후에 추가된 모든 작업을 수행합니다.

검색 및 찾아보기

검색 및 찾아보기 프로세스의 고가용성 구성은 간단합니다. 검색 및 찾아보기 프로세스가 여러 컴퓨터에서 실행되도록 시스템을 구성하면 됩니다.

검색 및 찾아보기 프로세스가 중단되면 어떤 일이 발생할까요? 사용자가 여전히 시스템에 로그인할 수는 있지만 Tableau Server가 대부분 사용할 수 없게 되고, 통합 문서의 콘텐츠가 없다고 표시될 것입니다. 콘텐츠가 실제로 없는 것은 아닙니다. 다만 검색 결과에 표시되지 않으며, 검색 및 찾아보기 프로세스가 다시 시작되면 표시될 것입니다. 장애가 발생한 시점이 하나 이상의 검색 및 찾아보기 프로세스가 구성되어 실행되고 있는 경우라면, 중단된 검색 및 찾아보기 프로세스로 보내진 요청은 실패하고, 후속 요청의 경로가 작동 중인 검색 및 찾아보기 프로세스로 재지정됩니다. 각 검색 및 찾아보기 프로세스는 클러스터 내의 모든 노드의 색인을 작성하므로, 검색 및 찾아보기 프로세스 하나를 제외한 모두가 중단되는 경우라도 결과가 모든 노드에 전달됩니다.

VizQL 서버

VizQL 서버 프로세스를 고가용성으로 구현하려면 단순히 하나 이상의 추가 인스턴스를 구성하여 여러 컴퓨터에서 실행하면 됩니다.

VizQL 서버 프로세스가 중단되면 어떤 일이 발생할까요? 하나밖에 없는 VizQL 서버 프로세스가 중단되는 경우 Tableau Server는 더 이상 뷰를 렌더링하지 못하게 됩니다. 고가용성을 달성하려면 VizQL 프로세스에 중복성을 두도록 구성해야 합니다. 가장 일반적인 경우는 각 노드에 2~4개의 VizQL 서버 프로세스로 구성하는 것입니다. 이를 통해 고가용성 및 확장성 요구를 동시에 충족하게 됩니다. 여러 VizQL 서버 프로세스를 실행 중에 하나의 프로세스가 중단되는 경우, 중단된 프로세스로 보내진 요청은 처리되지 않고 세션 데이터는 손실됩니다. 향후 처리 요청의 경로는 Tableau Server 클러스터에 있는 정상적인 다른 VizQL 서버 프로세스로 지정됩니다.

지금까지 각 서버 프로세스에 장애가 발생할 때 어떻게 작동하는지와 Tableau Server 클러스터 전반에 고가용성 기능을 구현하여 장애 위험을 줄이는 방법을 살펴보았습니다. 관리자는 각각의 실패 시나리오에 대한 계획뿐만 아니라, 클러스터를 적극적으로 모니터링하여 이전 프로세스 중단도 확인해야 합니다.

타사 모니터링 도구와 통합

Tableau Server의 내장 메커니즘을 사용한 시스템 상태의 모니터링뿐만 아니라 시스템 판독 가능한 각 프로세스 상태의 XML 버전을 원격으로 받을 수도 있습니다. 이러한 정보를 받으려면 서버에 대한 원격 액세스를 허용하고 다음과 같이 해당 인스턴스에 맞는 URL을 사용해야 합니다. 예) http://<my_tableau_server>/admin/systeminfo.xml

이렇게 하면, 구문 분석을 통해 서버 상태를 식별하고 다른 시스템 또는 모니터링 도구에 통합할 수 있는 다음과 같은 status.xml 파일이 반환됩니다.

서버 상태 URL에 원격으로 액세스하여 받은 status.xml 파일의 시스템 판독 가능한 XML 샘플

```
<systeminfo xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
      <machines>
       <machine name="my_tableau_server">
       <repository worker="my_tableau_server:8060" status="Active"
preferred="false"/>
       <dataengine worker="my_tableau_server:27042" status="Active"/>
       <applicationserver worker="my_tableau_server:8600" status="Active"/>
       <apiserver worker="my_tableau_server:8000" status="Active"/
       <vizqlserver worker="my_tableau_server:9100" status="Active"/>
       <dataserver worker="my_tableau_server:9700" status="Active"/>
       <backgrounder worker="my_tableau_server:8250" status="Active"/>
        <gateway worker="my_tableau_server:80" status="Active"/>
         <searchandbrowse worker="my_tableau_server:11000" status="Active"/>
        <cacheserver worker="my_tableau_server:6379" status="Active"/>
       <filestore worker="my_tableau_server:9345" status="Active"
pendingTransfers="0"   failedTransfers="0" syncTimestamp="2015-02-
27T20:30:48.564Z"/>     <clustercontroller worker="my_tableau_server:12012"
status="Active"/>
        <coordination worker="my_tableau_server:12000" status="Active"/>
        </machine>
      </machines>
      <service status="Active"/>
</systeminfo>
```

주 서버 노드의 페일오버

Tableau Server가 첫 번째로 설치된 서버가 Tableau Server의 주 노드로 식별됩니다. 주 노드는 서버의 모든 가용한 기능뿐만 아니라 고유 라이선스 관리 프로세스 및 기타 관리 기능을 모두 갖춘 유일한 서버입니다. 대규모 배포에서는 작업자 노드에 중복되는 프로세스들은 빼고 핵심적인 고유 관리 기능만 포함하여 주 서버를 배포할 수 있습니다. Tableau Server는 많은 배포 유연성을 제공하며, 주지할 사항은 모든 프로세스들이 72시간마다 주 서버에 라이선스 확인을 요청한다는 것입니다. 라이선스 확인 중에 주 서버에 액세스할 수 없는 경우, Tableau Server는 '라이선스 없음' 상태가 되어 사용할 수 없게 됩니다.

하지만 실제 환경의 주 서버는 Tableau Server 소프트웨어가 아닌 하드웨어, OS, 네트워크 장애 등의 이유로 장애가 발생할 수 있습니다. 이러한 경우에 클러스터를 외부 로드 밸런서 뒤에 배포하고 클러스터에 있는 모든 노드에 게이트웨이를 설치하면, 클러스터의 다른 노드에서 지속적으로 서비스 요청에 대응할 수 있습니다. 하지만 라이선스 확인 가간이 만료될 때까지 주 노드에 접속할 수 없는 경우 전체 클러스터 라이선스가 취소됩니다. 이러한 문제를 방지하려면 주 서버의 백업 시스템을 준비하고 있어야 합니다.

주 서버의 백업 구성

주 서버에 장애가 발생하면 서버 장애 문제가 해결될 때까지 관리상의 다운타임이 발생합니다. 이러한 문제를 방지하려면 주 서버 노드에 장애가 발생할 경우를 대비하여 주 서버의 '백업' 기능을 수행할 전용 백업 시스템을 두어야 합니다. 구성 및 준비가 완료된 주 서버의 백업은 액티브 상태로 사용하거나 클러스터에 연결하면 안됩니다. 이를 통해 라이선스 및 관리 기능이 지속적으로 주 서버 노드에서 작동하도록 보장하는 것입니다. Tableau Server 관리 가이드의 지침에 따라 주 서버의 백업을 설정 및 운영하는 방법에 대해 알아보십시오.

주 서버에 장애가 발생할 때, 주 서버의 백업을 주 서버로 전환하는 단계도 Tableau Server 관리 가이드에 설명되어 있습니다. 이 단계에 따라 클러스터의 나머지 노드에 변경 사항을 전달하고, 관리 및 라이선스 요청 경로를 새로운 주 서버로 자동 또는 수동으로 재지정합니다.

클러스터 상태 모니터링

시스템 관리자는 Tableau Server 10.0 프로세스의 상태를 Tableau Server 상태 페이지에서 모니터링할 수 있습니다. 이 페이지에는 모든 노드의 서버 프로세스와 액티브/패시브 리포지토리의 상태가 표시되므로 관리자는 전체 클러스터의 상태를 손쉽게 파악하게 됩니다.

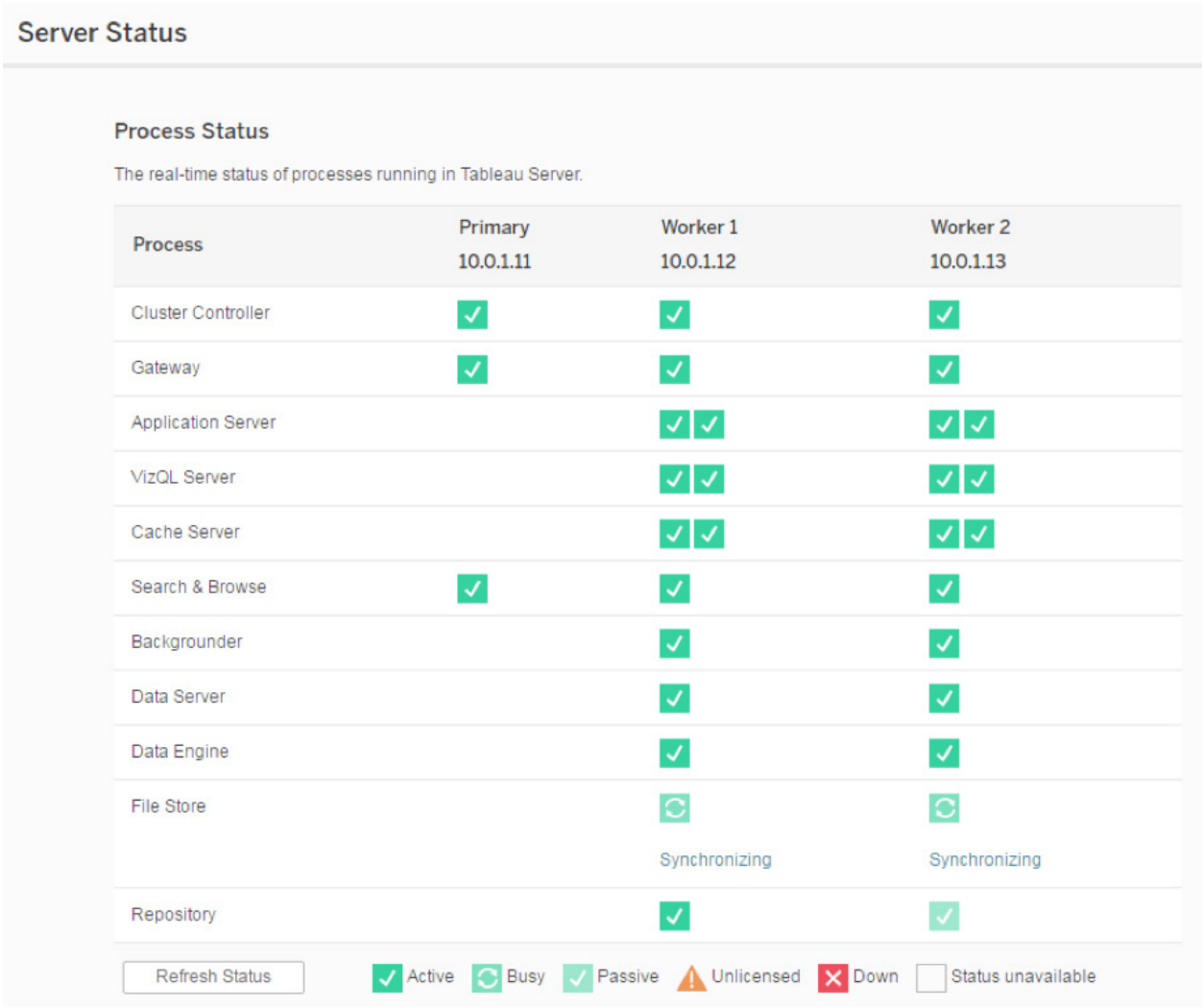

그림 3. 서버 상태 관리 페이지에 각 노드의 모든 프로세스 상태가 표시됩니다.

Tableau Server는 Tableau Server 구성 유틸리티를 통해 시스템 장애가 있을 때 서버 관리자에 알림 이메일을 보내도록 구성될 수도 있습니다. 또한 이 유틸리티를 사용하면 디스크 공간 문제가 발생하기 전에 미리 경고를 제공할 수 있습니다.

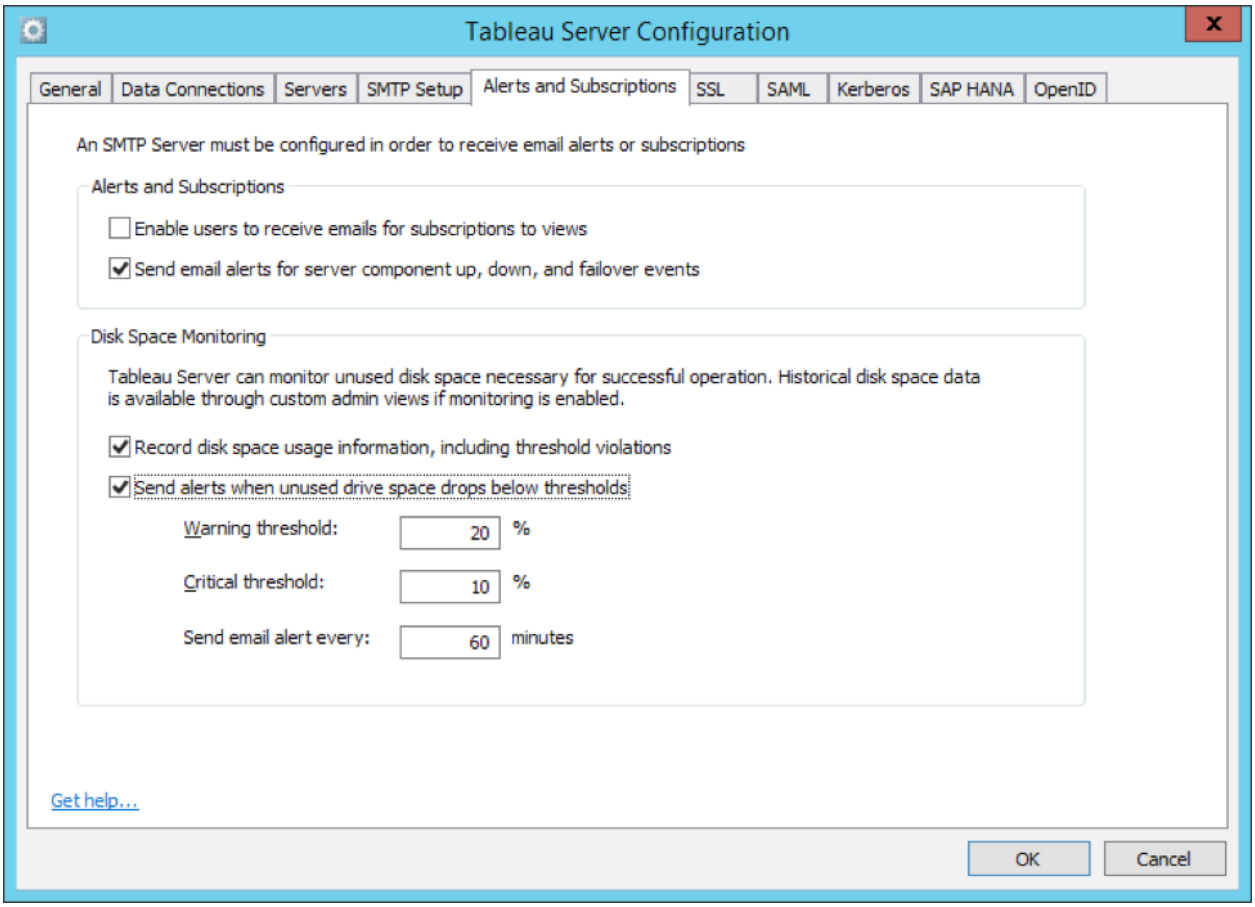

그림 4. Tableau Server 구성 유틸리티는 시스템 장애 시 알림을 설정하는 기능을 제공합니다.

아키텍처 고려 사항

고가용성을 성공적으로 구현하려면 업타임 목표와 조직에서 기대하는 서비스 수준을 파악하여 중복성에 대한 계획을 세워야 합니다.

최적의 구성 선택

최적의 클러스터 구성은 비즈니스의 요구사항 및 조직의 가용 리소스에 따라 결정됩니다. 고가용성을 성공적으로 구성하는 방법에는 여러 가지가 있겠지만, 다양한 크기의 클러스터 간의 장단점을 파악하여 환경에 맞는 최선의 구성 방법을 선택하는 것이 중요합니다. 또한, 각 작업자 노드가 모두 같은 구성 요소를 갖는 대칭형 구조와 각 노드가 다양한 유형 및 개수의 구성 요소를 갖는 비대칭형 구조 중에서 선택해야 합니다. 작업자 노드를 대칭형으로 구성하면 노드를 복제하거나 클러스터에 노드를 추가하는 작업이 간단해집니다. 하지만 유의할 사항은 이 구성에는 단일 액티브 및 단일 패시브 리포지토리라는 제약이 있다는 것입니다.

최소 3노드 HA 배포

이전에 설명한 바와 같이 완전한 고가용성 모드로 실행하기 위해서는 클러스터에 최소 세 개의 노드가 있어야 합니다. 3노드 클러스터로 시작하면 향후에 보다 큰 규모로 확장하기도 용이합니다. 한 노드에 장애가 발생하는 경우에도 여전히 두 노드가 액티브 상태여서 쿼럼 조건을 만족합니다. 아래 그림 5는 관리 및 라이선스 기능을 하는 주 노드의 역할을 보여줍니다. 또한 Tableau Server의 데이터 및 시각화 기능을 담당하는 두 개의 추가 작업자 노드도 도표로 예시하고 있습니다. 이러한 구성 요소의 실제 조합은 확장성 및 고가용성 요구 사항에 따라 달라집니다. 예를 들어 그림에는 작업자 노드당 VizQL 프로세스가 하나만 표시되어 있지만, 실제 구성에서는 작업자 노드에 두 개 이상의 VizQL 서버 프로세스가 실행되어 고가용성 및 확장성 요구 사항을 지원하게 됩니다. 또한 주 노드에 더 많은 코어 용량이 있다면, VizQL 서버 또는 데이터 엔진, 파일 저장소 등의 서버 프로세스를 추가하여 확장성을 높일 수 있습니다.

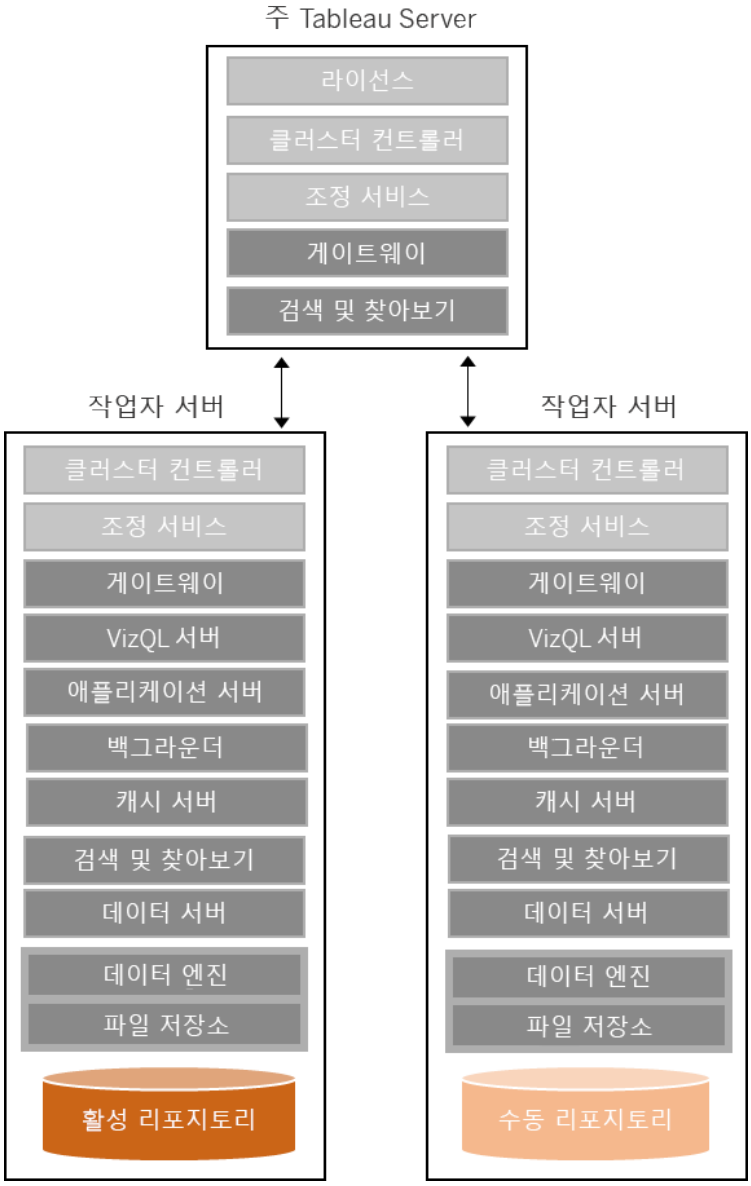

그림 5. 최소 3노드 배포의 예

여러 컴퓨터에 데이터 엔진을 설치하면 하드웨어 장애에 대한 중복성이 추가되고 백업 복구도 빨라집니다. 중복성을 더 추가할 필요가 있다면 각 작업자 노드에 추가 데이터 엔진을 배포하면 됩니다.

작업자 노드 하나에 장애가 발생하는 경우 주 서버의 프로세스는 영향을 받지 않습니다. 그러나, 작업자 노드는 주 서버 노드의 영향을 받습니다. 이러한 구성에서 액티브 리포지토리가 있는 작업자 노드에 장애가 발생하면 패시브 리포지토리가 있는 작업자 노드가 자동으로 액티브로 전환됩니다. 하지만 주 서버의 장애를 신속하게 복구하려면 주 서버의 백업을 구성하여 운용하는 것이 최선입니다. 또한, 조직에서 외부 로드 밸런서에 액세스할 권한이 있다면 해당 외부 로드 밸런서 뒤에 클러스터를 배치하여 최종 사용자의 가용성을 향상할 수 있습니다. 처음으로 고가용성을 구현하는 경우라면 이 구성으로 시작하는 것이 좋습니다.

3노드 이상의 배포

3개 이상의 노드를 사용하는 경우, 클러스터에 홀수 개수의 노드를 배포하는 것이 고가용성 기능 향상을 위해 좋은 토폴로지입니다. 짝수 개수의 노드가 있는 클러스터는 노드가 하나 적은 클러스터와 같은 쿼럼을 갖습니다. 두 개의 노드에 장애가 발생하는 경우의 예를 든다면, 세 개의 노드가 있는 클러스터와 네 개의 노드가 있는 클러스터 모두 쿼럼을 충족하지 못하므로 시스템이 중단됩니다.

3노드 클러스터에 짝수 개수가 되는 네 번째의 노드를 추가하면, 워크로드가 추가 노드로 분산되어 확장성이 개선되고, 중복 프로세스가 추가되어 전체 클러스터의 장애 위험을 줄입니다.

다른 고려 사항을 배제한 상태에서 최소 다운타임이 최우선 목표인 경우 5개 이상의 노드가 있는 배포 아키텍처를 고려하는 것이 좋습니다. 이 문서에서는 3노드 아키텍처로 시작하는 것을 권장하지만, 업무에 고가용성이 크게 요구되는 대규모의 엔터프라이즈 배포에서는 최소 5개 이상의 노드를 사용할 것을 권장합니다.

추출 및 추출 새로 고침에 많이 의존하는 조직에서 일반적으로 사용하는 배포 아키텍처는 전용 작업자 노드에 백그라운더 프로세스를 배포하는 것입니다. 추출 새로 고침 워크로드는 종종 VizQL 서버 프로세스의 비주얼라이제이션 워크로드에 영향을 줄 수 있으므로, 백그라운더 프로세스를 전용 작업자 노드에 분산하여 서로 다른 두 워크로드 간의 리소스 경합을 피할 수 있습니다. 클러스터에 있는 여러 노드 간의 프로세스 중복성을 보장하기 위해 백그라운더 전용 작업자 노드를 쌍으로 배포하기도 합니다.

앞서 설명한 바와 같이, 고가용성은 심각한 장애로 인한 다운타임을 줄여줍니다. 서비스 수준 요구 사항을 충족하고 고가용성 지원 토폴로지를 배포하는 것 이외에도 신속한 복구가 필요한 경우가 있을 수 있습니다. 그러므로 재해 및 하드웨어 장애가 발생하는 경우를 대비하여 재해 복구 목표에 따라 시스템 복구 계획을 세우는 것이 좋습니다.

Tableau Server의 재해 복구

Tableau 환경의 재해 복구(DR) 계획에는 다음의 두 가지 요인이 고려되어야 합니다. (1) 복구 프로세스의 일부로서 다운타임에 대한 비즈니스 허용 수준과 (2) 데이터 손실에 대한 비즈니스 허용 수준입니다. 이는 완전 복구까지 걸리는 시간에 대한 측정값인 목표 복구 시간(RTO) 및 데이터 손실 허용치에 대한 측정값인 목표 복구 시점(RPO)으로 정의됩니다.

RTO 및 RPO 임계값을 설정할 때는 항상 비용 대비 효율성을 고려하는 것이 좋습니다. 말할 필요도 없이 시스템을 온라인 상태로 신속하게 복구하는 데 소요되는 인프라와 인력을 고려할 때, 매우 짧은 목표 복구 시간을 달성하는 데는 많은 비용이 듭니다. 인프라와 대응력 측면에서 비용을 줄이려면 RTO를 늘리면 됩니다(허용 다운타임 증가). 이러한 사항은 Tableau Server에만 국한되는 것이 아니라 모든 서버 소프트웨어 배포에 해당됩니다.

데이터 손실에 대한 허용 수준은 무엇보다도 개별 애플리케이션과 해당 애플리케이션의 데이터 변동률에 따라 달라집니다. 예를 들어 24시간 RPO는 데이터 손실을 감당할 수 있는 한도가 최근 24시간 범위라는 것입니다.

가장 간단한 DR 전략

가장 간단하고 효과적인 DR 전략은 야간 백업을 실행하고, 예비 서버와 백업을 복구 데이터 센터에 두는 것입니다. Tableau Server는 기존 BI 도구에 비해 설치, 구성 및 실행이 빠르며, 대부분의 경우 예비 서버를 새 서버로 설치하고 백업에서 복원하는 것만으로 비즈니스의 목표을 충족할 수 있습니다. 서버 설치 및 백업 복원에 필요한 시간은 백업의 크기와 서버의 복잡성에 따라 다르며, 대개 몇 시간 내에 완료됩니다.

타사 DR 솔루션

15분 이상의 RPO를 제공하는 타사 솔루션이 많이 있으며, IT 팀에서 이미 사용하고 있을 것입니다. 백업 및 복구에 타사 백업/복구 및 DR 솔루션을 사용하는 경우, Tableau Server는 다른 비즈니스 애플리케이션과 동일하게 처리되므로 구성도 다른 애플리케이션과 유사하게 진행하면 됩니다.

Tableau DR 기능

타사 DR 기술을 사용할 수 없는 경우 Tableau Server에 내장된 기본 백업 및 복구 기능을 사용하면 됩니다. Tableau Server가 실행 중인 경우에도 백업을 수행할 수 있으며, 전체 백업 및 핫 백업을 만들 수 있습니다. Tableau Server의 효율적인 핫 백업에서는 추출 파일, 리포지토리 데이터베이스 및 관련 구성 파일만 백업합니다. 또한, Tableau Server가 실행되지 않는 계획된 다운타임 중에 백업을 실행할 수도 있습니다.

RPO가 24시간인 경우 24시간마다 서버를 백업하도록 설정하거나 RPO를 기준으로 백업 일정을 변경할 수 있습니다. 복구 시간은 장애 요인이 소프트웨어 또는 하드웨어인지 그리고 새 하드웨어를 현장에 바로 투입할 수 있는지에 따라 달라집니다. Tableau Server의 동일한 인스턴스가 이미 설치 및 구성된 다른 예비 노드가 있는 경우, Tableau Server를 복원하여 운영하는 작업은 단지 저장된 백업 파일로 빠르게 관리 명령어를 실행하는 것에 지나지 않습니다.

대다수 기업의 재해 복구 정책에는 주 서버가 실행되는 데이터 센터와 복구 시스템이 있는 데이터 센터 간의 거리 요구 사항도 명시되어 있습니다. 예를 들어, 미국에서는 DR 사이트가 최소한 250마일 떨어져 있어야 합니다. 상황에 따라 Tableau를 재해 복구가 필요한 비즈니스 애플리케이션으로 취급하여 관리할 수 있습니다. DR 요구 사항 충족에 Tableau의 내장 백업 및 복구 기술을 활용해 보십시오.

고가용성 및 기타 고려 사항

대규모 셀프 서비스 분석을 제공하도록 고가용성을 구성하려면 Tableau Server에 내장된 기능 외의 추가 기능이 필요합니다. 장애는 하드웨어 또는 네트워크 장애와 같이 Tableau Server의 소프트웨어 범위를 벗어나서 발생할 수 있습니다. 퍼블릭 또는 프라이빗 클라우드 상의 가상화된 하드웨어와 같이 다양한 유형의 장애에 보다 유연하게 대응할 수 있는 인프라에 Tableau Server를 배포하면 가용성을 높일 수 있습니다. 일부 업체는 VMware vMotion과 같은 라이브 마이그레이션 기능을 제공하기도 합니다.

진정한 고가용성은 사용자의 요구 사항을 파악하고 이러한 요구 사항을 충족하는 모범 사례 및 절차를 따름으로써 실현될 수 있습니다. Tableau의 강력한 기능을 통해 고가용성 기능을 손쉽게 구현하더라도 정기적인 유지 관리 및 백업과 같은 작업은 항상 필요합니다. Tableau는 비즈니스 의사 결정을 돕는 비즈니스 인텔리전스 애플리케이션의 중요성을 잘 알고 있습니다. Tableau Server는 조직의 중요 분석 질문에 답할 준비가 되어 있습니다.

Tableau 정보

Tableau를 사용하면 데이터를 비즈니스를 개선할 수 있는 실행 가능한 인사이트로 전환할 수 있습니다. 데이터가 저장된 위치나 형식에 상관없이 손쉽게 연결하고, 애드혹 분석을 빠르게 수행하여 숨겨진 기회를 발견하게 됩니다. 드래그 앤 드롭으로 고급 시각적 분석 기능을 갖춘 대화형 대시보드를 작성하여 조직 전체에서 공유하고, 팀원들이 데이터에 대한 고유한 관점을 탐색하도록 도와줍니다. 글로벌 기업부터 신생 기업 및 소규모 비즈니스에 이르기까지 전 세계 모든 사람들이 Tableau의 분석 플랫폼을 사용하여 데이터를 보고 이해하고 있습니다.

리소스

엔터프라이즈용 Tableau: IT 개요

Tableau Server 확장성: 서버 관리자를 위한 기술적 배포 가이드

서버 관리자 가이드

Amazon Web Services에서의 Tableau