



인시던트 사후 검토

2022 년 4 월 서비스 중단

이 번역본은 편의를 위해서만 제공됩니다. 번역본과 영문본 사이에 모호하거나 상충되는 사항이 있는 경우 기존 영문본이 우선합니다.

공동 설립자 겸 공동 CEO 가 드리는 편지

이달 초 고객 서비스에 차질을 빚은 운영 중단을 확인하고자 합니다. 귀하의 비즈니스에 있어 저희의 제품이 매우 중요하다는 것을 알고 있으며, 저희는 이에 대한 책임을 가볍게 여기지 않습니다. 그 책임은 저희에게 있습니다. 전적인 책임이라 말씀드립니다. 이로 인해 영향을 받은 고객 여러분, 저희는 귀하의 신뢰를 되찾기 위해 노력하고 있습니다.

i Atlassian 핵심 가치 중 하나는 “열린 컴퍼니, 허튼 소리 없음” 입니다. 우리는 사건에 대해 공개적으로 논의하고 학습 기회로 사용함으로써 이러한 가치를 부분적으로 실현합니다. 당사 고객, 당사 Atlassian 커뮤니티 및 광범위한 기술 커뮤니티를 위해 이 사후 인시던트 리뷰를 게시하고 있습니다. Atlassian 는 흠 없는 문화와 기술 시스템 및 프로세스를 개선하는 방법을 찾는 데 중점을 두는 것이 대규모의 신뢰할 수 있는 서비스를 제공하는 데 중요하다는 점을 강조하는 인시던트 관리 프로세스를 자랑스럽게 생각합니다. 우리는 모든 유형의 인시던트를 피하기 위해 최선을 다하지만 인시던트가 또한 개선될 수 있는 강력한 방법이라는 생각도 받아들입니다.

Atlassian 의 클라우드 플랫폼은 모든 산업 분야와 규모를 포괄하여 20 만 명 이상의 클라우드 고객의 다양한 요구를 충족할 수 있으므로 안심하셔도 좋습니다. 이번 인시던트 발생 이전에, 저희의 클라우드는 99.9%의 가동 시간을 지속적으로 제공하여 가동 시간 SLA 를 초과했습니다. 저희는 확장 가능한 인프라와 꾸준한 보안 향상을 통해 플랫폼 및 여러 중앙 집중식 플랫폼 기능에 장기적인 투자를 해 왔습니다.

지속적인 신뢰와 파트너십을 보여 주신 고객 및 파트너 여러분께 감사드립니다. Atlassian 은 본 문서에 요약되어 있는 세부 사항 및 조치를 통해 모든 팀의 요구를 충족하는 세계적 수준의 클라우드 플랫폼과 강력한 제품 포트폴리오를 계속해서 제공해 드리겠습니다.



-Scott 과 Mike

핵심 요약

2022년 4월 5일 화요일, 7:38 UTC 부터 Atlassian 고객 775 명이 Atlassian 제품을 이용할 수 없었습니다. 이러한 고객 중 일부는 최장 14 일까지 중단을 경험했으며, 이에 4월 8일에 일부 복원 후 점진적으로 복원 작업이 지속되어 4월 18일에는 나머지 모든 고객 사이트가 복원되었습니다.

이는 사이버 공격으로 인한 것이 아니며 고객 데이터에 대한 무단 액세스 또한 발생하지 않았습니다. Atlassian에는 공표되어 있는 SLA와 해당 SLA 초과에 대한 이력이 있는 포괄적인 [데이터 관리](#) 프로그램이 있습니다.

이는 주요 인시던트였지만, 5분을 초과하여 데이터를 손실한 고객은 없었습니다. 또한, 고객 및 사용자의 99.6% 이상이 복원 활동 중에도 중단 없이 계속해서 클라우드 제품을 사용했습니다.

i 이 문서 전체에서 이 인시던트의 일부로 사이트가 삭제된 고객을 “작용을 받는” 또는 “영향을 받는” 고객이라고 지칭합니다. 이 PIR은 인시던트의 정확한 세부 정보를 제공하고, 복구를 위해 취한 단계를 간략하게 설명하며, 향후 이와 같은 상황이 발생하지 않도록 방지하는 방법을 설명합니다. 이 섹션에서는 인시던트에 대한 높은 수준의 요약 제공하며 문서의 나머지 부분에 자세한 내용이 나와 있습니다.

어떤 일이 일어났을까요?

2021년, "Insight - 자산 관리"라고 불리는 Jira Service Management 및 Jira Software용 독립 실행형 Atlassian 앱에 대한 인수 및 통합을 완료했습니다. 이 독립 실행형 앱의 기능은 Jira Service Management에서 기본으로 제공되었지만 더 이상 Jira Software에서 사용할 수 없었습니다. 이로 인해 기존 앱을 설치한 고객 사이트에서 독립 실행형 레거시 앱을 삭제해야 했습니다. 저희 엔지니어링 팀은 기존 스크립트와 프로세스를 사용하여 이 독립 실행형 애플리케이션의 인스턴스를 삭제했지만, 두 가지 문제가 있었습니다.

- **커뮤니케이션 격차가 있었습니다.** 삭제를 요청한 팀과 이를 실행한 팀 간에 커뮤니케이션 격차가 있었습니다. 팀은 삭제 처리로 *표시된 앱 ID*를 제공하는 대신 앱을 삭제할 *전체 클라우드 사이트의 ID*를 제공했습니다.

- **시스템 경고가 부족했습니다.** 삭제를 수행하는 데 사용된 API 는 사이트 식별자와 앱 식별자를 모두 수락했으며 입력값을 정확한 것으로 가정했습니다. 즉, 사이트 ID 가 전달되면 사이트를 삭제하고 앱 ID 가 전달되면 앱을 삭제하게 된 것입니다. 요청된 삭제 유형(사이트 또는 앱)을 확인하는 경고 시그널의 부재가 있었습니다.

실행한 스크립트는 호출한 엔드포인트와 그 호출 방식에 중점을 둔 표준 동료 검토 프로세스를 따랐습니다. 제공된 클라우드 사이트 ID 가 Insight App 을 참조했는지, 아니면 전체 사이트를 참조했는지 교차 확인하지 않았으며, 문제는 해당 스크립트에 고객의 전체 사이트에 대한 ID 가 포함되었다는 것입니다. 그 결과 2022 년 4 월 5 일 화요일 UTC 기준 07:38 ~ 08:01 사이에 883 개의 사이트(775 명의 고객)가 즉시 삭제되었습니다. *“인시던트 개요” 참조*

저희는 어떻게 대응했을까요 ?

해당 인시던트가 4 월 5 일 UTC 기준 08:17 에 확인되자, 저희는 주요 인시던트 관리 프로세스를 트리거하고 부서 간 인시던트 관리 팀을 구성했습니다. 글로벌 인시던트 대응 팀은 모든 사이트가 복원 및 검증되고, 고객에게 반환될 때까지 인시던트 기간 동안 24 시간 중단 없이 근무했습니다. 또한, 인시던트 관리 리더들은 3 시간마다 회의를 진행하여 작업 흐름을 조정했습니다.

초기에, 저희는 여러 제품을 동시에 사용하는 수백 개의 고객 사이트를 복원하는 데 있어 여러 가지 어려움을 겪었습니다.

인시던트가 시작될 때, 저희는 영향을 받은 사이트에 대해 정확히 알고 있었고, 저희는 영향을 받은 각 사이트의 승인된 소유자와의 커뮤니케이션을 통해 중단에 대해 알리는 것을 우선 순위로 했습니다.

그러나, 일부 고객의 연락처 정보가 삭제되었습니다. 다시 말해, 고객은 평소처럼 지원 티켓을 제출할 수 없었습니다. 또한 저희는 주요 고객 연락처에 즉시 액세스할 수 없었습니다. *자세한 내용은 “복구 작업 흐름에 대한 간략한 개요” 참조*

향후 이러한 상황을 방지하기 위해 저희는 무엇을 하고 있을까요?

저희는 여러 가지 즉각적인 조치를 취했으며 향후 이러한 상황을 피하기 위한 변경에 최선을 다하고 있습니다. 다음은 중요한 변경을 수행했거나 수행할 네 가지 특정 영역입니다.

1. **모든 시스템에서 범용 '소프트 삭제'를 설정합니다.** 전반적으로 이러한 유형의 삭제는 금지하거나, 오류를 방지하기 위해 여러 계층의 보호 기능을 갖추어야 하고 이는 '소프트 삭제'에 대한 단계적 롤아웃 및 테스트된 롤백 계획을 포함합니다. 당사는 전세계에 대해 소프트 삭제 절차를 거치지 않은 고객 데이터 및 메타데이터 삭제를 금지할 것입니다.
2. **DR(재해 복구) 프로그램을 가속화하여 대규모 고객을 대상으로 다수의 사이트 또는 다수의 제품을 삭제하는 이벤트의 복원을 자동화합니다.** 이 인시던트를 통해 배운 자동화 및 학습을 활용하여 DR 프로그램을 가속화하고 이러한 규모의 인시던트에 대해 규정 정의된 대로 RTO(복구 시간 목표)를 달성할 것입니다. 정기적으로는 대규모 사이트에 대한 모든 제품을 복원하는 DR 연습을 진행할 것입니다.
3. **대규모 인시던트에 대한 인시던트 관리 프로세스를 수정합니다.** 당사는 대규모 사고에 대한 표준 운영 절차를 개선하고 해당 사고의 규모에 맞는 시뮬레이션을 통해 이를 실천할 것입니다. 많은 팀이 동시에 협업할 수 있도록 교육 및 도구도 업데이트 할 것입니다.
4. **대규모 인시던트 커뮤니케이션 플레이북을 마련합니다.** 당사는 여러 채널을 통해 조기에 사고를 발견하고 몇 시간 내에 사고에 대하여 공개적으로 소통할 것입니다. 영향을 받는 고객에게 더 잘 닿을 수 있도록 주요 연락처 백업을 개선하고 지원 도구를 개선하여 유효한 URL 또는 Atlassian ID 가 없는 고객이 기술 지원팀에 직접 연락할 수 있도록 지원할 것입니다.

조치 항목에 대한 전체 목록은 아래의 전체 인시던트 사후 검토에 자세히 설명되어 있습니다. *“개선 방식”* 참조

목차

Atlassian 클라우드 아키텍처 개요	7 페이지
<ul style="list-style-type: none">• Atlassian 의 클라우드 호스팅 아키텍처• 분산된 서비스 아키텍처• 다중 테넌트 아키텍처• 테넌트 프로비저닝 및 수명 주기• 재해 복구 프로그램<ul style="list-style-type: none">○ 복원력○ 서비스 스토리지 복원○ 다수 사이트, 다수 제품 자동 복원 기능	
일어난 일, 타임라인 및 복구	13 페이지
<ul style="list-style-type: none">• 인시던트 개요• 조정 방식• 인시던트 타임라인• 복구 작업 흐름에 대한 간략한 개요<ul style="list-style-type: none">○ 작업 흐름 1: 탐지, 복구 시작 및 접근법 식별○ 작업 흐름 2: 조기 복구 및 복원 1 접근법○ 작업 흐름 3: 복구 가속화 및 복원 2 접근법○ 삭제된 사이트 복원 후 데이터 손실 최소화	
인시던트 커뮤니케이션	21 페이지
<ul style="list-style-type: none">• 인시던트 개요	
지원 경험 & 고객 도달	23 페이지
<ul style="list-style-type: none">• 고객에 대한 지원은 어떤 영향을 받았습니까?• 저희는 어떻게 대응했을까요?	
어떻게 개선할 수 있을까요?	25 페이지
<ul style="list-style-type: none">• 학습 1: “소프트 삭제”는 모든 시스템에서 보편적이어야 함• 학습 2: DR 프로그램의 일환으로서, 다수의 고객을 대상으로 한 다중 사이트, 다중 제품 삭제 이벤트의 복원 자동화• 학습 3: 대규모 이벤트에 대한 인시던트 관리 프로세스 개선• 학습 4: 커뮤니케이션 프로세스 개선	
마무리하며 짚을 요소	31 페이지

Atlassian 클라우드 아키텍처 개요

본 문서 전반에 걸쳐 논의한 바와 같이 인시던트의 요인을 이해하려면, 먼저 Atlassian 의 제품, 서비스 및 인프라에 대한 배포 아키텍처를 이해하는 것이 좋습니다.

Atlassian 의 클라우드 호스팅 아키텍처

Atlassian 은 클라우드 서비스 공급자로 AWS(Amazon Web Services)를 선정하여 사용하고 있으며 [전 세계 여러 영역에 있는](#) AWS 의 고가용성 Data Center 시설을 활용합니다. 각 AWS 지역은 독립된 지리적 위치이며, 여기에는 가용성 영역(AZ)이라고 하는 격리 및 물리적으로 분리된 여러 Data Center 그룹이 있습니다.

AWS 의 컴퓨팅, 스토리지, 네트워크 및 데이터 서비스를 활용하여 저희의 제품 및 플랫폼 컴포넌트를 구축하므로, 가용성 영역 및 리전과 같은 AWS 에서 제공하는 이중화 기능을 활용할 수 있습니다.

분산된 서비스 아키텍처

이 AWS 아키텍처를 통해 솔루션 전반에 사용하는 다양한 플랫폼 및 제품 서비스를 호스팅합니다. 여기에는 미디어, ID, 상거래, 편집기와 같은 환경, Jira 이슈 서비스 및 Confluence 분석과 같은 제품별 기능 등 여러 Atlassian 제품에서 공유 및 사용하는 플랫폼 기능을 포함합니다.

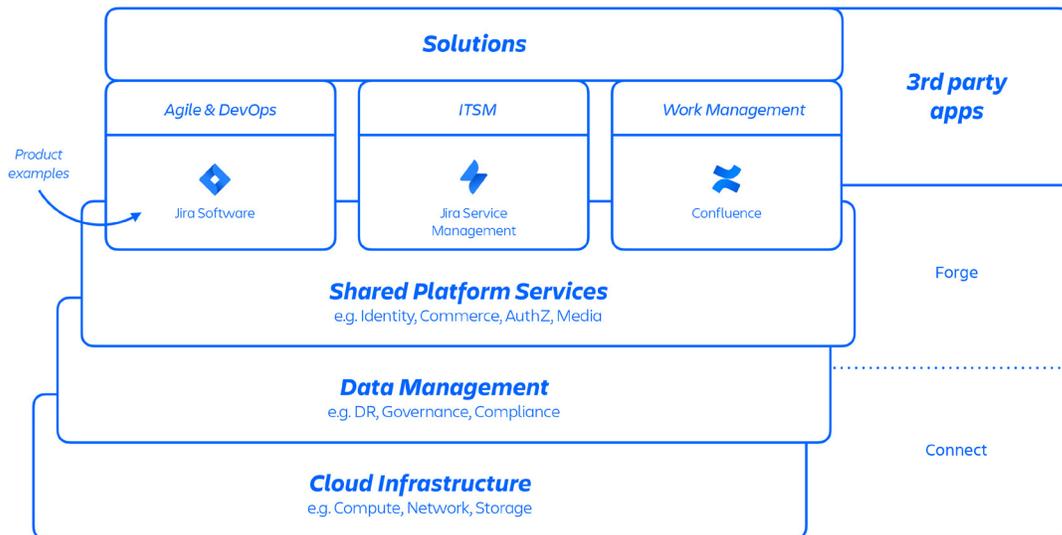


그림 1: Atlassian 플랫폼 아키텍처.

Atlassian 개발자는 Micros 라고 하는 내부적으로 개발된 PaaS(Platform-as-a-Service)를 통해 이러한 서비스를 프로비저닝합니다. 이 PaaS 는 공유 서비스, 인프라, 데이터 저장 및 보안 및 규정 준수 제어 요건(위의 그림1 참조)을 포함한 관리 기능의 배포를 자동으로 오케스트레이션합니다. 일반적으로 Atlassian 제품은 Micros 를 통해 AWS 에 배포하는 여러 '컨테이너화된' 서비스로 구성됩니다. Atlassian 제품은 요청 라우팅에서 이진 개체 저장소, 인증/권한 부여, 트랜잭션 UGC(사용자 생성 콘텐츠) 및 엔터티 관계 저장소, 데이터 레이크, 공통 로깅, 추적과 관찰 가능성 및 분석 서비스 요청에 이르기까지 중심 플랫폼 기능(아래 그림2 참조)을 사용합니다. 이러한 마이크로 서비스는 플랫폼 수준에서 표준화된 승인된 기술 스택을 사용하여 구축됩니다.

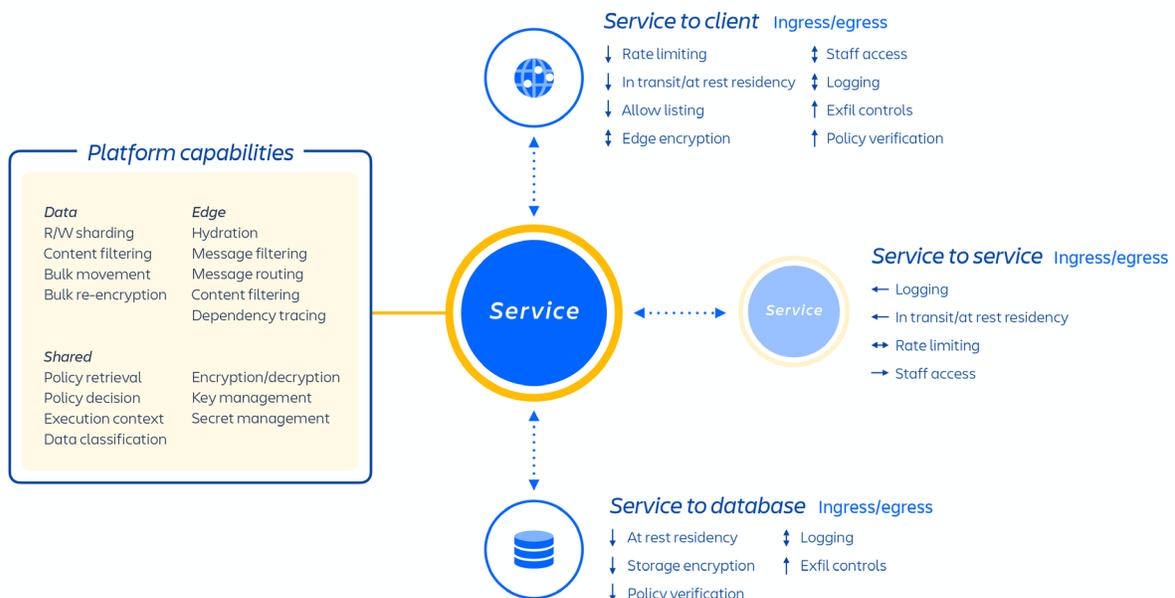


그림2: Atlassian 마이크로 서비스 개요

다중 테넌트 아키텍처

Atlassian 은 클라우드 인프라 외에도 제품을 지원하는 공유 플랫폼을 통해 다중 테넌트 마이크로 서비스 아키텍처를 구축하고 운영합니다. 다중 테넌트 아키텍처에서 단일 서비스가 여러 고객에게 서비스를 제공하며 여기에는 클라우드 제품을 실행하는 데 필요한 데이터베이스 및 컴퓨팅 인스턴스가 포함됩니다. 각 분할된 데이터베이스(본질적으로 컨테이너라고 볼 수 있는 데이터베이스 - 아래 그림3 참조)에는 여러 테넌트에 대한 데이터가 포함되어 있지만 각 테넌트의 데이터는 격리되어 다른 테넌트가 액세스할 수 없습니다.

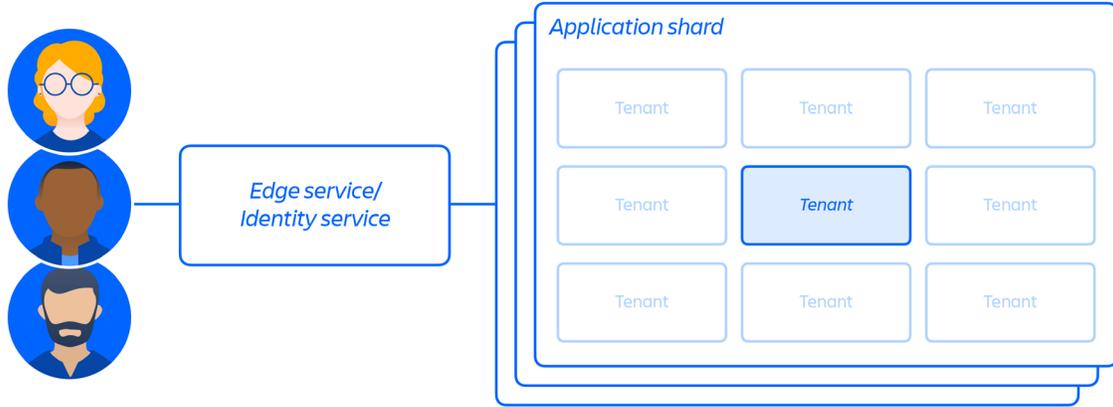


그림 3: 다중 테넌트 아키텍처에서 데이터를 저장하는 방법.

테넌트 프로비저닝 및 수명 주기

신규 고객을 프로비저닝하면, 일련의 이벤트가 분산된 서비스의 오케스트레이션 및 데이터 스토어의 프로비저닝을 트리거합니다. 이러한 이벤트는 일반적으로 수명 주기의 7 단계 중 하나에 매핑할 수 있습니다.

- 1 상거래 시스템은 해당 고객에 대한 최신 메타데이터 및 액세스 제어 정보로 즉시 업데이트되며, 프로비저닝 오케스트레이션 시스템은 일련의 테넌트 및 제품 이벤트를 통해 “프로비저닝한 리소스의 상태”를 라이선스 상태에 맞춥니다.

테넌트 이벤트

이러한 이벤트는 테넌트에 전체적으로 영향을 미치며 다음 중 하나가 될 수 있습니다.

- 생성: 테넌트가 생성되어 새로운 사이트에 사용됩니다.
- 파기: 전체 테넌트가 삭제됨

제품 이벤트

- 활성화: 라이선스가 부여된 제품 또는 타사 앱을 활성화한 후
- 비활성화: 특정 제품 또는 앱을 비활성화한 후
- 일시 중단: 지정된 기존 제품이 일시 중지된 후, 해당 제품이 소유한 특정 사이트에 대한 액세스를 사용 중지합니다.
- 일시 중지 해제: 특정 기존 제품의 일시 중지를 취소하여 제품이 소유 한 사이트에 액세스 할 수 있습니다.

라이선스 업데이트: 해당 제품의 라이선스 시트 수 및 상태(활성/비활성)에 대한 정보가 포함되어 있습니다.

- 2 고객 사이트를 생성하고 고객을 위해 올바른 제품군을 활성화합니다 사이트의 개념은 특정 고객에게 라이선스가 부여 된 여러 제품의 컨테이너라고 보면 됩니다(예: <사이트 이름>.atlassian.net 용 Confluence 및 Jira Software). 이는 이 보고서를 이해하는 데 중요한 부분입니다(아래 그림 4 참조). 이 사고에서 삭제된 내용이 사이트 컨테이너이며 사이트라는 개념은 이 문서의 전반에 걸쳐 논의되기 때문입니다.

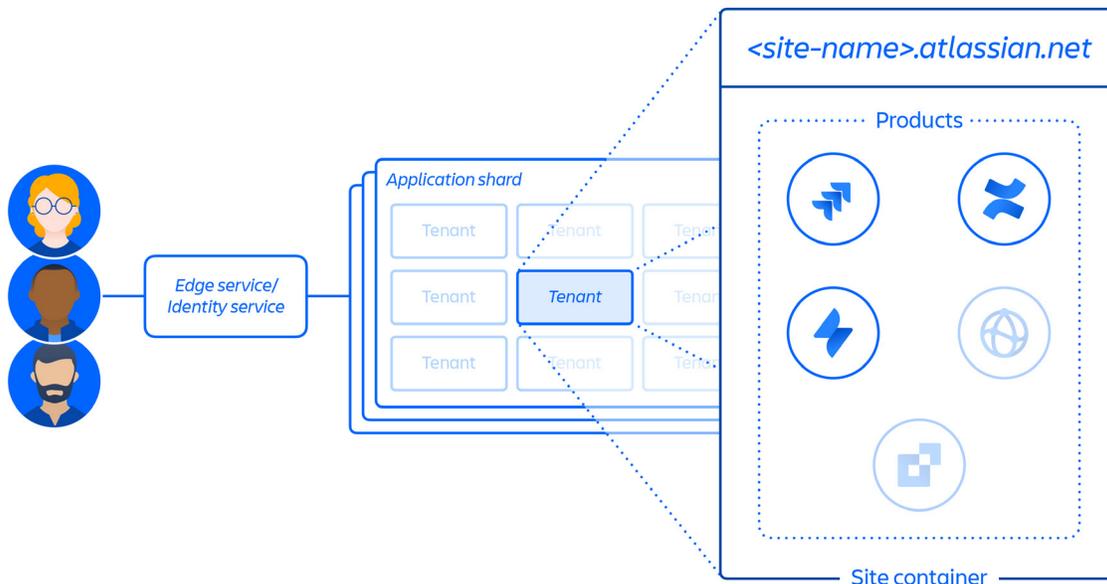


그림 4: 사이트 컨테이너 개요.

- 3 지정된 리전에 있는 고객 사이트 내에서 제품 프로비저닝.

제품을 프로비저닝하면 대부분의 콘텐츠는 사용자가 액세스하는 위치와 근접한 곳에서 호스팅됩니다. 전 세계로 호스팅할 때 제품 성능을 최적화하기 위해 데이터 이동을 제한하지 않으며, 필요에 따라 리전 간에 데이터를 이동할 수 있습니다.

일부 제품의 경우, 데이터 보존 또한 제공합니다. 데이터 보존을 통해 고객은 제품 데이터를 전 세계에 배포할 것인지 또는 정의된 지리적 위치 중 한 곳에 보관할 것인지 선택할 수 있습니다.

- 4 고객 사이트, 제품 핵심 메타데이터 및 구성의 생성 및 저장.

- 5 사이트 및 제품 ID 데이터(예: 사용자, 그룹, 권한 등)의 생성 및 저장.
- 6 사이트 내 제품 데이터베이스 프로비저닝, 예: Jira 제품군, Confluence, Compass, Atlas.
- 7 제품 라이선스가 부여된 앱의 프로비저닝.

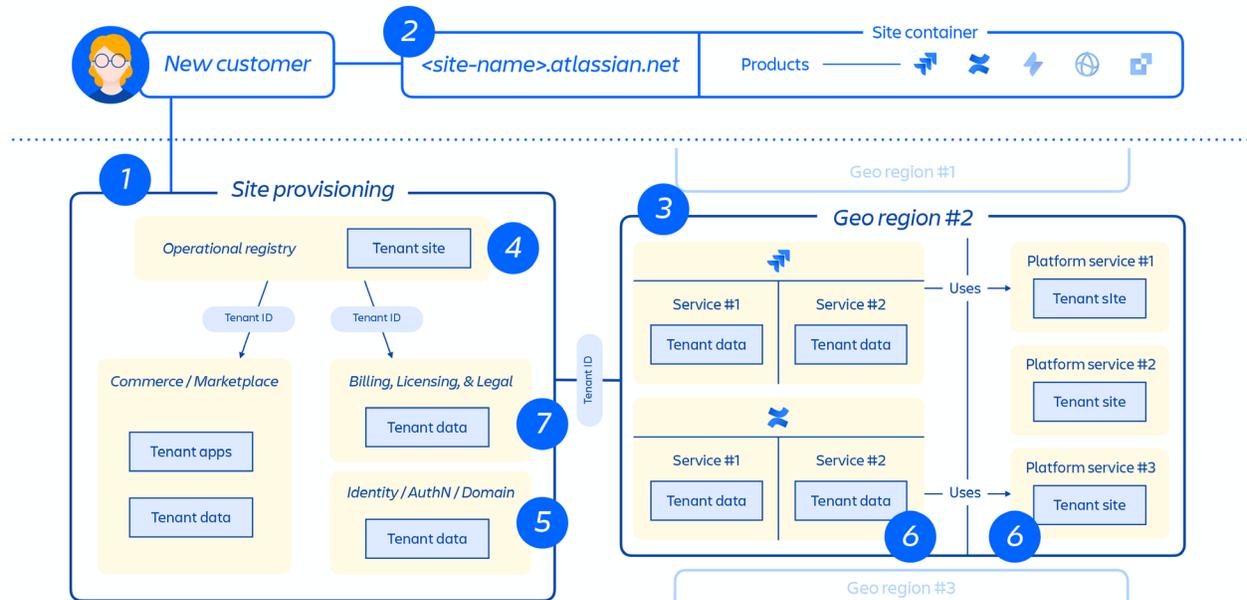


그림 5: 분산 아키텍처 전반에서 고객 사이트가 프로비저닝되는 방식에 대한 개요.

위의 그림 5 는 고객 사이트가 단일 데이터베이스 또는 스토어뿐만 아니라 분산 아키텍처 전반에 걸쳐 배포되는 방식을 보여줍니다. 여기에는 메타 데이터, 구성 데이터, 제품 데이터, 플랫폼 데이터 및 기타 관련 사이트 정보를 저장하는 여러 물리적 및 논리적 위치가 포함됩니다.

재해 복구 프로그램

당사의 [재해 복구](#)(DR) 프로그램에는 인프라 장애에 대한 복원력과 백업을 통한 서비스 스토리지의 복원 가능성을 제공하려는 모든 업무 활동이 포함됩니다. 재해 복구 프로그램을 이해하는 데 있어 중요한 두 가지 개념은 다음과 같습니다.

- **RTO(복구 시간 목표):** 재해 발생 중에 얼마나 빠르게 데이터를 복구하여 고객에게 반환할 수 있는가?
- **RPO(복구 지점 목표):** 백업을 통해 복구한 데이터는 얼마나 최신인가? 마지막 백업 이후 얼마나 많은 데이터가 손실되는가?

이번 인시던트에서 저희는 RTO 를 달성하지 못했지만 RPO 는 충족했습니다.

복원력

저희는 인프라 수준의 장애(예: 전체 데이터베이스, 서비스 또는 AWS 가용성 영역의 손실)에 대비합니다. 이러한 대비에는 여러 가용성 영역에 걸친 데이터 및 서비스의 복제와 정기적인 장애 조치 테스트가 포함됩니다.

서비스 스토리지 복원

또한 랜섬웨어, 악의적인 작업자, 소프트웨어 결함 및 운영 오류 등의 위험으로 인한 서비스 스토리지의 데이터 손상을 복구하기 위해 대비합니다. 이러한 대비에는 변경 불가능한 백업 및 서비스 스토리지 백업 복원 테스트가 포함됩니다. 개별 데이터 저장소를 가져와 이전 시점으로 복원할 수 있습니다.

다수 사이트, 다수 제품 자동 복원 기능

인시던트 발생 당시에는 고객 사이트를 대규모로 선택하고 백업을 통해 상호 연결된 모든 제품을 이전 시점으로 복원할 수 없었습니다.

저희의 역량은 인프라, 데이터 손상, 단일 서비스 이벤트 또는 단일 사이트 삭제에 중점을 두어 왔습니다. 과거에는 이러한 종류의 장애를 처리하고 테스트해야 했습니다. 사이트 수준의 삭제에는 이 이벤트의 규모에 맞게 신속하게 자동화할 수 있는 런북이 없었기 때문에 모든 제품 및 서비스 전반에 걸쳐 툴링 및 자동화를 조정된 방식으로 진행해야 했습니다.

다음 섹션에서는 Atlassian 이 이러한 복잡성과 아키텍처를 대규모로 유지 관리하는 능력을 발전시키고 최적화하기 위해 수행하는 작업에 대해 자세히 설명합니다.

일어난 일, 타임라인 및 복구

인시던트 개요

2021 년, "Insight - 자산 관리"라고 불리는 Jira Service Management 및 Jira Software 용 독립 실행형 Atlassian 앱에 대한 통합을 완료했습니다. 이 독립 실행형 앱의 기능은 Jira Service Management 에서 기본으로 제공되었지만 더 이상 Jira Software 에서 사용할 수 없었습니다. 이로 인해 기존 앱을 설치한 고객 사이트에서 독립 실행형 레거시 앱을 삭제해야 했습니다. 저희 엔지니어링 팀은 기존 스크립트와 프로세스를 사용하여 이 독립 실행형 애플리케이션의 인스턴스를 삭제했습니다.

그러나 두 가지 중요한 문제가 발생했습니다:

- **커뮤니케이션 격차가 있었습니다.** 삭제를 요청한 팀과 이를 실행한 팀 간에 커뮤니케이션 격차가 있었습니다. 팀은 삭제 처리로 표시된 앱 ID 를 제공하는 대신 앱을 삭제할 전체 클라우드 사이트의 ID 를 제공했습니다.
- **시스템 경고가 부족했습니다.** 삭제를 수행하는 데 사용된 API 는 사이트 식별자와 앱 식별자를 모두 수락했으며 입력값을 정확한 것으로 가정했습니다. 즉, 사이트 ID 가 전달되면 사이트를 삭제하고 앱 ID 가 전달되면 앱을 삭제하게 된 것입니다. 요청된 삭제 유형(사이트 또는 앱)을 확인하는 경고 시그널의 부재가 있었습니다.

실행된 스크립트는 호출한 엔드포인트와 그 호출 방식에 중점을 둔 표준 동료 검토 프로세스를 따랐습니다. 제공된 클라우드 사이트 ID 가 앱을 참조했는지, 아니면 전체 사이트를 참조했는지 교차 확인하지 않았습니다. 표준 변경 관리 프로세스에 따라 해당 스크립트를 스테이징에서 테스트했지만, 해당 ID 가 스테이징 환경에 없었기 때문에 ID 입력이 올바르지 않았음을 감지하지 못했습니다.

프로덕션에서 실행할 때, 처음에 30 개의 사이트에 대해 해당 스크립트를 실행했습니다. 첫 번째 프로덕션 실행은 성공적이었으며 다른 부작용 없이 30 개의 해당 사이트에 대한 Insight App 을 삭제했습니다. 그러나 30 개의 해당 사이트에 대한 ID 는 커뮤니케이션에 착오가 발생한 이벤트 이전에 제공되었으며 올바른 Insight App ID 를 포함했습니다.

후속 프로덕션 실행을 위한 스크립트에는 Insight 앱 ID 대신 사이트 ID 가 포함되어 있었으며 883 개 사이트에 실행되었습니다. 스크립트는 4 월 5 일 07:38 UTC 에 실행되기 시작했으며 08:01 UTC 에

완료되었습니다. 스크립트는 입력 목록에 따라 사이트를 순차적으로 삭제했기 때문에 첫 번째 고객의 사이트는 스크립트가 07:38 UTC 에 실행되기 시작한 직후 삭제되었습니다. 그 결과 엔지니어링 팀에 경고 신호 없이 883 개 사이트가 즉시 삭제되었습니다.

영향을 받은 고객은 Jira 제품군, Confluence, Atlassian Access, Opsgenie, Statuspage 등 Atlassian 제품을 사용할 수 없었습니다.

해당 인시던트를 파악하자마자, 저희 팀은 영향을 받은 모든 고객을 위한 복원에 집중했습니다. 당시, 저희는 영향을 받은 사이트의 수를 약 700 개로 추정했습니다(영향을 받은 사이트는 총 883 개이지만, Atlassian 소유의 사이트는 제외했습니다). 700 개 사이트 중 상당수는 비활성, 무료 계정 또는 활성 사용자 수가 적은 소규모 계정이었습니다. 이에 기반하여, 처음에는 영향을 받은 고객의 대략적인 수를 약 400 명으로 추정했습니다.

Atlassian 의 공식 고객 정의에 따른 완전한 투명성을 위해 말씀드리자면, 이제 저희는 훨씬 더 정확히 파악하고 있으며, 775 명의 고객이 중단의 영향을 받았습니다. 그러나 대부분의 사용자는 초기 추정치인 400 명의 고객에 해당합니다. 이러한 고객 중 일부에게는 중단이 최대 14 일까지 지속되었으며, 첫 번째 그룹의 고객 사이트는 4 월 8 일에 복원됐고, 모든 고객 사이트는 4 월 18 일자로 복원되었습니다.

조정 방식

4 월 5 일 07:46 UTC, 영향을 받은 고객 한 명이 첫 번째 지원 티켓을 생성했습니다. 사이트가 표준 워크플로를 통해 삭제되었기 때문에 내부 모니터링에서는 문제를 감지하지 못했습니다. 08:17 UTC, 당사는 주요 인시던트 관리 절차를 시작하여 여러 부서가 협업하는 인시던트 관리 팀을 구성했으며 7 분 만인 08:24 UTC 에 위기(Critical) 단계로 상향되었습니다. 08:53 UTC, 저희 팀은 고객 지원 티켓이 스크립트 실행과 관련이 있음을 확인했습니다. 12:38 UTC, 복원의 복잡성을 알게 되자 당사는 해당 사고를 가장 심각한 단계로 지정했습니다.

인시던트 관리 팀은 엔지니어링, 고객 지원, 프로그램 관리, 커뮤니케이션 등을 포함하여 Atlassian 내 여러 팀의 인원들로 구성되었습니다. 핵심 팀은 모든 사이트가 복원 및 검증되고, 고객에게 반환될 때까지 인시던트 기간 동안 3 시간마다 회의를 진행했습니다.

복원 진행 상황을 관리하기 위해 신규 Jira 프로젝트, SITE 및 여러 팀(엔지니어링, 프로그램 관리, 지원 등)에서 사이트별로 복원을 추적하는 워크플로를 만들었습니다. 이러한 접근 방식을 통해 모든 팀은 개별 사이트 복원과 관련된 이슈를 더 쉽게 식별하고 추적할 수 있었습니다.

또한 4 월 8 일 UTC 기준 03:30 에 인시던트 기간 동안 모든 엔지니어링 전반에 걸쳐 코드 동결을 구현했습니다. 이를 통해 고객 사이트 복원에 집중하고, 고객 데이터의 불일치를 야기하는 변경의 위험을 제거하고, 다른 중단의 위험을 최소화하고, 팀의 복구를 방해하는 관련 없는 변경의 가능성을 줄일 수 있었습니다.

인시던트 타임라인

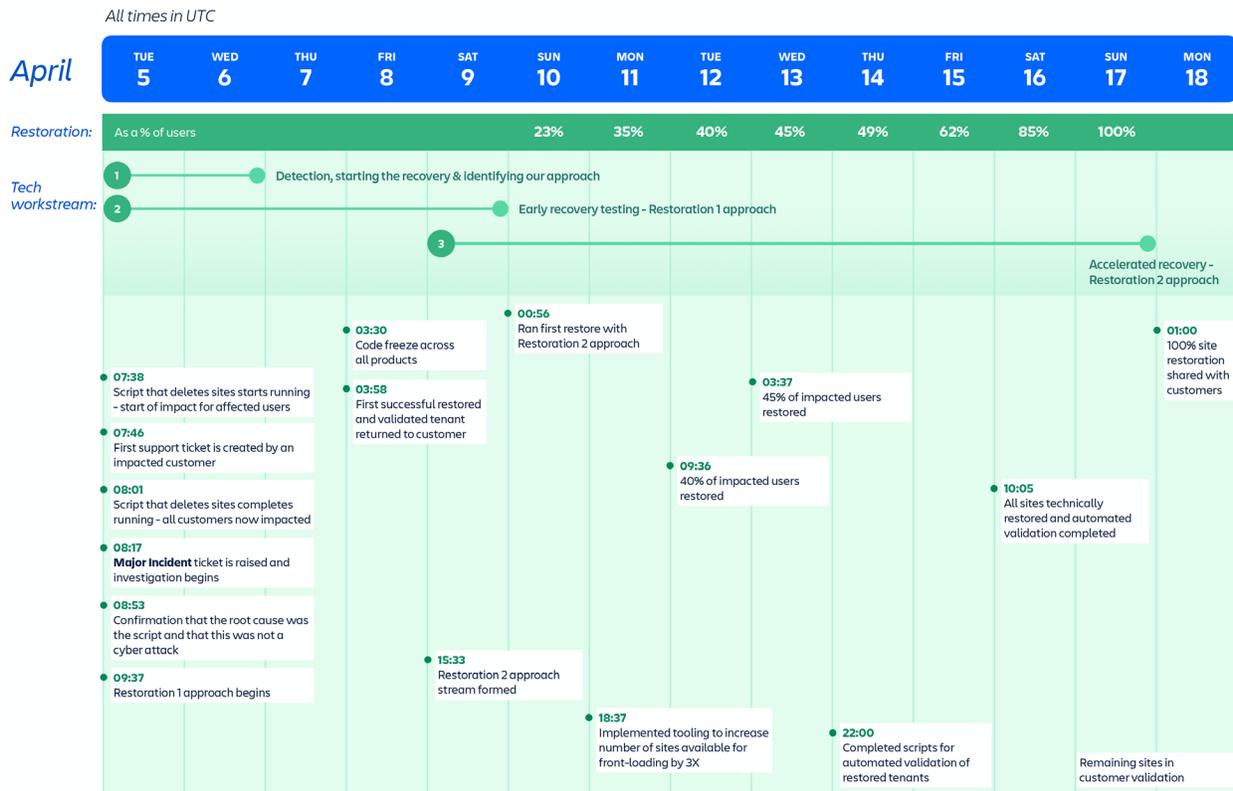


그림 6: 사건의 타임라인 및 주요 획기적인 복원 단계

복구 작업 흐름에 대한 간략한 개요

복구는 감지, 조기 복구 및 가속화의 세 가지 기본 작업 흐름으로 실행했습니다. 아래에 각 작업 흐름에 대해 하나씩 설명드렸지만, 복구 중에는 모든 작업 흐름에 걸쳐 작업을 동시에 진행했습니다.

작업 흐름 1: 탐지, 복구 시작 및 접근법 식별

타임스탬프: 1 일~2 일(4 월 5 일~6 일)

4 월 5 일 UTC 기준 08:53 에 Insight App 스크립트로 인한 사이트 삭제를 인식했습니다. 내부의 악의적인 행위나 사이버 공격으로 인한 것이 아니라는 점을 확인했습니다. 관련 제품 및 플랫폼 인프라 팀을 호출하여 해당 인시던트를 처리하도록 조치했습니다.

인시던트 초기에 저희는 다음 사항에 대해 인식했습니다.

- 삭제된 수백 개의 사이트를 복원하는 것은 복잡한 여러 단계의 프로세스로(위의 아키텍처 섹션 참조), 성공적으로 완료하려면 많은 팀이 필요하고 며칠이 걸립니다.
- 단일 사이트를 복구할 수 있었지만, 대규모 배치의 사이트를 복구하는 역량과 프로세스를 구축하지 못했습니다.

그 결과, 영향을 받은 고객이 최대한 빨리 Atlassian 제품에 다시 액세스할 수 있도록 복원 프로세스를 크게 병렬화 및 자동화해야 했습니다.

작업흐름 1 에는 다음과 같은 활동에 참여하는 수많은 개발 팀이 연계되어 있습니다.

- 파이프라인 내 여러 배치의 사이트에 대한 복원 단계 식별 및 실행.
- 관련 팀이 더 많은 사이트에 대한 복원 단계를 일괄적으로 실행할 수 있는 자동화 작성 및 개선.

작업 흐름 2: 조기 복구 및 복원 1 접근법

타임스탬프: 1 일~4 일(4 월 5 일~9 일)

스크립트 실행을 완료한 지 1 시간 이내인 4 월 5 일 UTC 기준 08:53 에 사이트 삭제의 원인을 파악했습니다. 또한 이전에 소수의 사이트를 프로덕션으로 복구하는 데 사용했던 복원 프로세스를 식별했습니다. 그러나 이러한 규모로 삭제된 사이트를 복원하는 복구 프로세스는 쉽게 정의되지 않았습니다.

신속하게 진행하기 위해 인시던트의 초기 단계를 다음과 같이 두 개의 작업 그룹으로 나누었습니다.

- 수동 작업 그룹은 필요한 단계를 검증하고 소수의 사이트에 대한 복원 프로세스를 수동으로 실행했습니다.
- 자동화 작업 그룹은 기존 복원 프로세스를 수행하고 자동화를 구축하여 대규모 배치의 사이트에서 해당 단계를 무사히 실행했습니다.

복원 1 접근법 개요(아래 그림 7 참조):

- 각각의 삭제된 사이트에 대해 신규 사이트를 만든 다음, 해당 사이트의 데이터를 복원해야 하는 모든 다운로드 제품, 서비스 및 데이터 저장소가 필요했습니다.
- 신규 사이트에는 클라우드 ID와 같은 신규 식별자가 제공됩니다. 이러한 식별자는 모두 변경 불가능한 것으로 간주됩니다. 다시 말해, 많은 시스템에서 이러한 식별자를 데이터 레코드에 포함합니다. 결과적으로 이러한 식별자가 변경되면, 대량의 데이터를 업데이트해야 했으며, 이는 타사 에코시스템 앱에 특히 문제가 됩니다.
- 삭제된 사이트의 상태를 복제하기 위한 신규 사이트의 수정에는 단계 간에 복잡하고 종종 예상치 못한 종속성이 있었습니다.

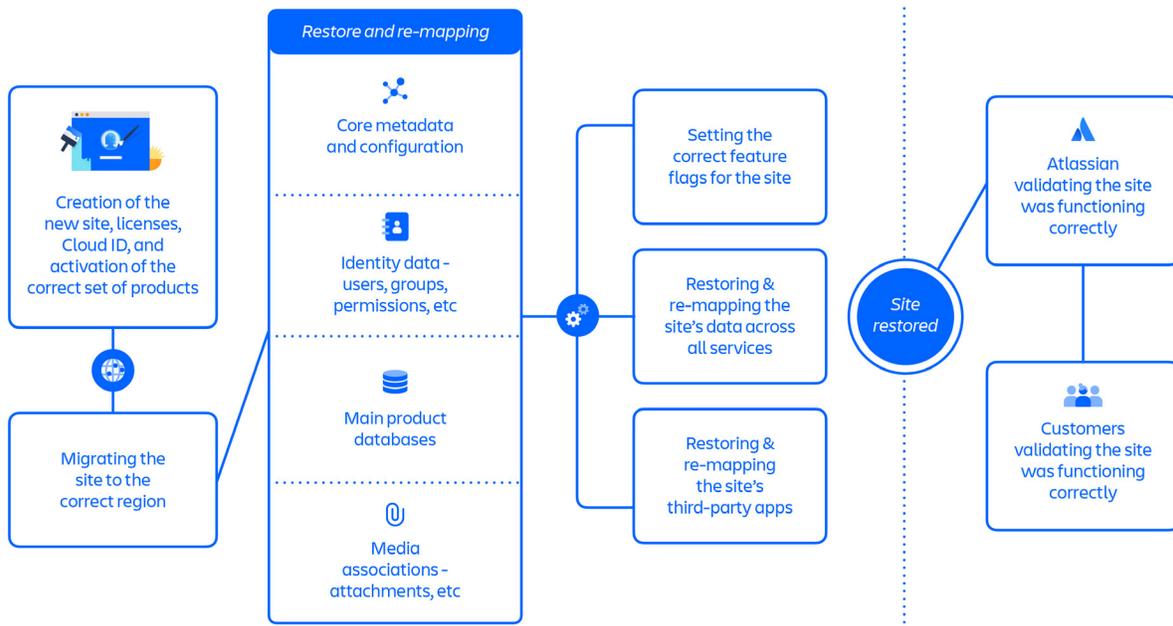


그림 7: 복원 1 접근법의 주요 단계.

복원 1 접근 방식에는 약 70 개의 개별 단계가 포함됐으며, 간략하게 종합하자면 주로 다음과 같은 흐름을 순차적으로 따랐습니다.

- 신규 사이트, 라이선스, 클라우드 ID 생성 및 올바른 제품 그룹 활성화
- 사이트를 올바른 영역으로 마이그레이션
- 사이트의 핵심 메타데이터 및 구성 복원 및 리매핑
- 사이트의 ID 데이터(사용자, 그룹, 권한 등) 복원 및 리매핑
- 사이트의 주요 제품 데이터베이스 복원
- 사이트의 미디어 연결(첨부 파일 등) 복원 및 리매핑
- 사이트에 대한 올바른 기능 플래그 설정
- 모든 서비스에서 사이트 데이터 복원 및 리매핑
- 사이트의 타사 앱 복원 및 리매핑
- Atlassian 에서 사이트가 올바르게 작동하고 있는지 검증
- 고객 측에서 사이트가 올바르게 작동하고 있는지 검증

한 번 최적화하면, 복원 1 접근 방식은 한 배치의 사이트를 복원하는 데 약 48 시간이 소요됐으며, 4 월 5 일에서 4 월 14 일 사이에 112 개의 사이트에서 영향을 받은 사용자의 53%를 복구하는 데 사용됐습니다.

작업 흐름 3: 복구 가속화 및 복원 2 접근법

타임스탬프: 4 일~13 일(4 월 9 일~17 일)

복원 1 접근법에서는 모든 고객을 복구하는 데 3 주가 예상되었습니다. 따라서 당사는 4 월 9 일에 모든 사이트의 복원 속도를 높이기 위해 새로운 접근법인 복원 2 접근법을 제안했습니다(아래 그림 8 참조).

복원 2 접근 방식은 복원 1 접근 방식에 존재하는 종속성의 수와 복잡성을 줄여서 복원 단계 간의 병렬성을 개선했습니다.

복원 2 에는 카탈로그 서비스 기록부터 시작하여 모든 각 시스템에서 사이트와 관련된 기록을 재생성(또는 삭제 취소) 하는 작업이 포함되었습니다. 이 새로운 접근 방식의 핵심 요소는 *이전 사이트 식별자를 모두 재사용*하는 것이었습니다. 이로 인해 각 사이트의 모든 타사 앱 공급업체와 협력해야 하는

필요성을 포함하여 이전 식별자를 새 식별자에 매핑하는 데 사용되었던 이전 프로세스의 단계 중 절반 이상이 제거되었습니다.

그러나 복원 1 에서 복원 2 접근 방식으로의 전환으로 인해 인시던트 대응에 상당한 간접비용이 추가되었습니다.

- 복원 1 접근 방식에서 설정된 많은 자동화 스크립트와 프로세스를 복원 2 에 맞게 바뀌야 했습니다.
- 복원을 수행하는 팀(인시던트 코디네이터 포함)은 두 접근 방식 모두에서 복원의 병렬 배치를 관리해야 했으며 복원 2 프로세스를 테스트하고 검증했습니다.
- 새로운 접근 방식을 사용한다는 것은 규모를 확장하기 전에 복원 2 프로세스를 테스트하고 검증해야 한다는 것을 의미했으며, 이를 위해서는 전에 복원 1 에 대해 완료된 중복 검증 작업이 필요했습니다.

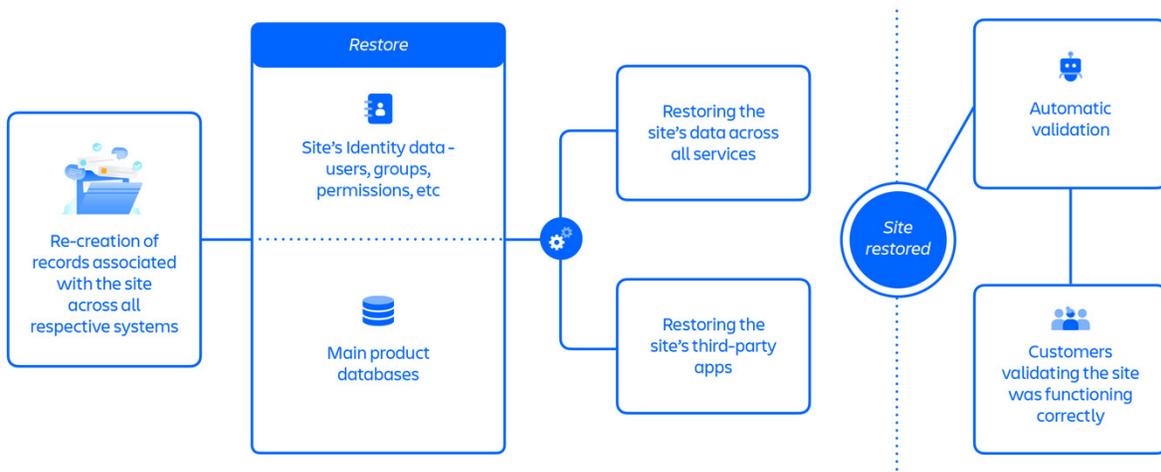


그림 8: 복원 2 접근법의 주요 단계.

위의 그래픽은 복원 2 접근 방식을 나타내며, 여기에는 다음과 같은 대체로 병렬화된 흐름을 따르는 30 개 이상의 단계가 포함되었습니다.

- 모든 각 시스템에서 사이트와 관련된 기록 재생성
- 사이트의 ID 데이터 복원 (사용자, 그룹, 권한 등)
- 사이트의 주요 제품 데이터베이스 복원
- 모든 서비스에서 사이트 데이터 복원
- 사이트의 타사 앱 복원

- 자동 유효성 검사
- 고객 측에서 사이트가 올바르게 작동하고 있는지 검증

복구 가속화의 일환으로 대규모 배치의 경우 수동 복원이 제대로 확장되지 않기 때문에 사이트 복원을 프런트로드 및 자동화하는 단계도 취했습니다. 복구 프로세스의 순차적 특성으로 인해 대규모 데이터베이스 복원 및 사용자 기반/권한 복원의 경우 사이트 복원이 느려질 수 있습니다. 구현한 최적화는 다음과 같습니다.

- 데이터베이스 복원 및 ID 동기화와 같은 초기 로드 및 장기 실행 단계에 필요한 도구 및 가드 레일을 개발하여 다른 복원 단계에 앞서 완료되도록 했습니다.
- 엔지니어링 팀은 개별 단계에 대한 자동화를 구축하여 대량의 복구작업을 안전하게 실행할 수 있도록 했습니다.
- 모든 복원 단계가 완료된 후 사이트가 올바르게 작동하는지 확인하기 위해 Automation 가 구축되었습니다.

가속화된 복원 2 접근 방식은 사이트를 복원하는 데 약 12 시간이 걸렸으며, 4 월 14 일에서 17 일 사이에 771 개 사이트에서 영향을 받는 사용자의 약 47% 를 복구하는 데 사용되었습니다.

삭제된 사이트 복원 후 데이터 손실 최소화

당사의 데이터베이스는 전체 백업과 증분 백업을 조합하여 백업되므로 백업 보존 기간(30 일) 내에 데이터 저장소를 복구할 특정 “시점”을 선택할 수 있습니다. 이 인시던트 발생 시 대부분의 고객은 당사 제품의 주요 데이터 저장소를 파악하여 안전한 동기화 지점으로 사이트를 삭제하기 5 분 전에 복원 지점을 사용하기로 결정했습니다. 기본이 아닌 데이터 저장소는 동일한 지점에 복원되거나 기록된 이벤트를 재생하여 복원되었습니다. 기본 스토어에 고정 복구 지점을 사용함으로써 모든 데이터 스토어에서 데이터의 일관성을 유지할 수 있었습니다.

인시던트 대응 초기에 복구된 고객 57 명에 대해 일관된 정책이 부족하고 데이터베이스 백업 스냅샷을 수동으로 검색하지 않아 일부 Confluence 및 Insight 데이터베이스가 사이트 삭제 5 분 이상전의 시점으로 복원되었습니다. 불일치는 복원 후 감사 프로세스 중에 발견되었습니다. 이후 나머지 데이터를 복구하고 이로 인해 영향을 받는 고객에게 연락했으며 변경 사항을 적용하여 데이터를 추가로 복원하도록 돕고 있습니다.

요약

- 이 인시던트 중에 한 시간이라는 복구 시점 목표 (RPO) 를 달성했습니다.
- 인시던트로 인한 데이터 손실은 사이트가 삭제되기 5 분 전에 제한됩니다.
- 소수의 고객이 Confluence 또는 Insight 데이터베이스를 사이트 삭제 5 분 전까지 복원했지만 데이터를 복구할 수 있으며 현재 이 데이터를 복원하기 위해 고객과 협력하고 있습니다.

인시던트 커뮤니케이션

인시던트 커뮤니케이션에 대해 이야기할 때 여기에는 고객, 파트너, 미디어, 업계 분석가, 투자자 및 광범위한 기술 커뮤니티와의 접점이 포함됩니다.

인시던트 개요

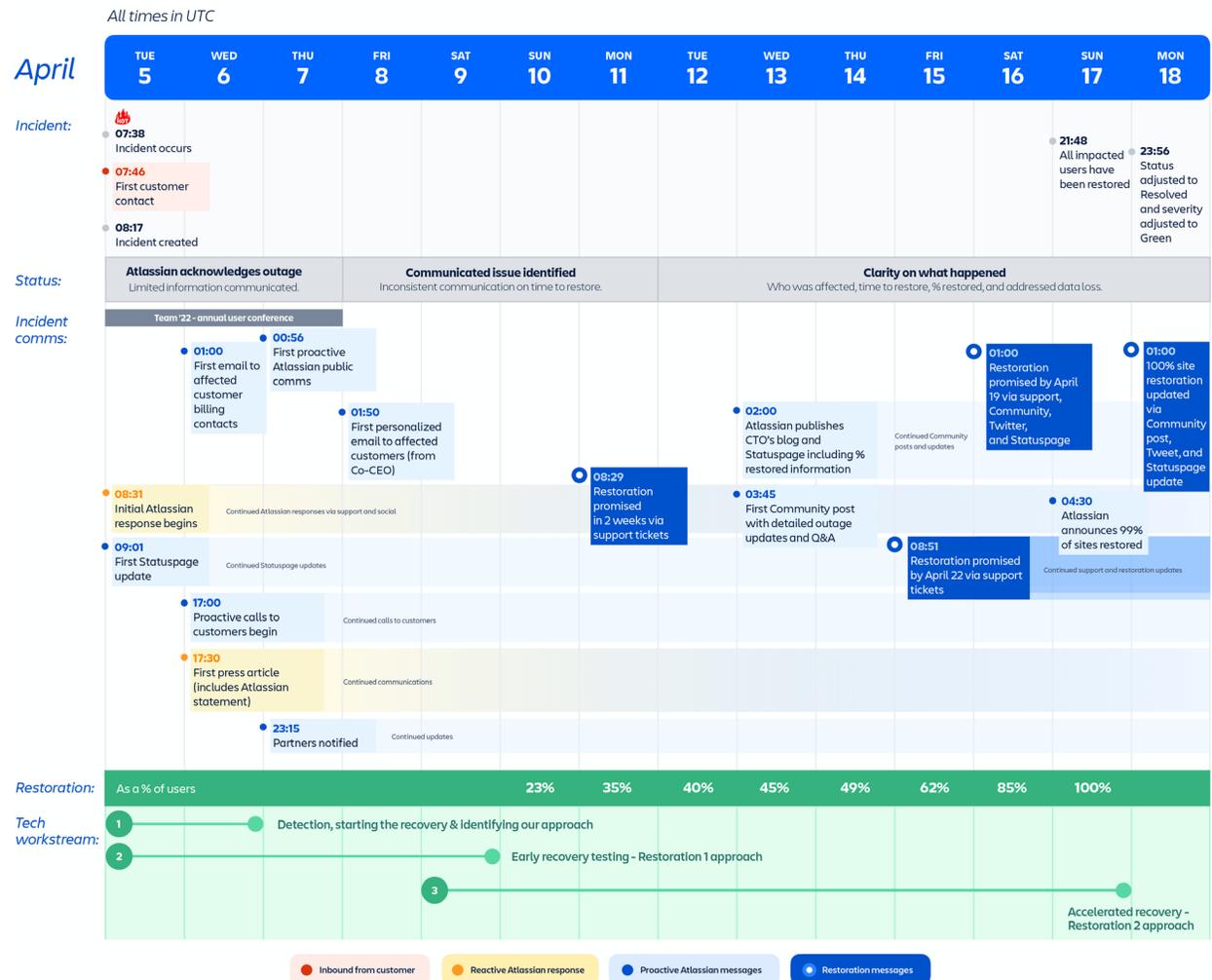


그림 9: 주요 인시던트 커뮤니케이션의 획기적인 단계 타임라인.

타임스탬프: 1 일 - 3 일 (4 월 5 일 ~ 7 일)

조기 응답

첫 번째 지원 티켓은 4 월 5 일 7:46 UTC 에 생성되었으며 Atlassian 고객 지원팀은 UTC 8:31 까지 인시던트를 확인하면서 응답했습니다. UTC 9:03 에 첫 번째 Statuspage 업데이트가 게시되어 고객에게 인시던트를 조사하고 있음을 알립니다. 그리고 11:13 UTC 에 Statuspage 를 통해 근본 원인을 식별했으며 수정 작업 중임을 확인했습니다. 4 월 6 일 1:00 UTC 까지 초기 고객 티켓 통신에서 중단은 유지 관리 스크립트로 인한 것이며 데이터 손실을 최소화할 것으로 예상했다고 밝혔습니다. Atlassian 4 월 6 일 17:30 UTC 에 성명을 통해 언론 문의에 응답했습니다. Atlassian 4 월 7 일 00:56 UTC 에 인시던트를 확인하는 첫 번째 광범위한 외부 메시지를 트윗했습니다.

타임스탬프: 4 일 - 7 일 (4 월 8 일 - 11 일)

보다 광범위하고 개인화된 지원 시작

4 월 8 일 1:50 UTC 에 Atlassian 영향을 받는 고객에게 공동 설립자이자 공동 CEO 인 스콧 파쿠하르로부터 사과를 이메일로 보냈습니다. 그 후 며칠 동안 우리는 삭제된 연락처 정보를 복원하고 아직 제출되지 않은 영향을 받는 모든 사이트에 대한 지원 티켓을 만들기 위해 노력했습니다. 그런 다음 지원 팀은 영향을 받는 각 사이트와 관련된 지원 티켓을 통해 복원 노력에 대한 정기적인 업데이트를 계속 보냈습니다.

타임스탬프: 8 일 - 14 일 (4 월 12 일 - 18 일)

명확성 향상 및 완벽한 복원

4 월 12 일, [Atlassian CTO 인 Sri Viswanath 의 업데이트를 게시하여 발생한 일](#), 영향을 받은 사람, 데이터 손실 여부, 복원 진행률 및 소요될 수 있는지에 대한 자세한 기술적 세부 정보를 제공합니다. 모든 사이트를 완전히 복구하는 데 2 주가 소요될 수 있습니다. 블로그에는 Sri 에 기인한 또 다른 언론 성명서가 수반되었습니다. 또한 [엔지니어링 책임자 스티븐 디시\(Stephen Deasy\)의 첫 번째 적극적인 Atlassian 커뮤니티 게시물에서 Sri 의](#) 블로그를 언급했습니다. 이 게시물은 이후 추가 업데이트 및

Q&A 를 위한 전용 장소가 되었습니다. 더 넓은 대중들. 이 게시물에 대한 4 월 18 일 업데이트는 영향을 받는 모든 고객 사이트의 전체 복원을 발표했습니다.

왜 더 빨리 공개적으로 응답하지 않았을까요?

1. Statuspage, 이메일, 지원 티켓 및 1:1 상호 작용을 통해 영향을 받는 고객과 직접 소통하는 것을 우선시했습니다. 하지만, 사이트가 삭제되었을 때 연락처 정보가 손실되어 많은 고객에게 연락할 수 없었습니다. 영향을 받는 고객과 최종 사용자에게 인시던트 대응 및 해결 일정을 알리기 위해 훨씬 더 광범위한 커뮤니케이션을 더 일찍 구현해야 했습니다.
2. 사건의 원인을 즉시 알았지만, 이 사건의 아키텍처 복잡성과 고유한 상황으로 인해 신속하게 범위 지정하고 해결 시간을 정확하게 예측하는 능력이 저하되었습니다. 전체 그림을 얻을 때까지 기다리지 않고 우리가 아는 것과 몰랐던 것에 대해 투명했을 것입니다. 일반적인 복원 추정치를 제공하고 (방향성이 있더라도) 보다 완전한 그림을 얻을 것으로 예상되는 시점을 명확하게 파악하면 고객이 사고에 대해 더 잘 계획을 세울 수 있었을 것입니다. 이는 조직 내 이해 관계자와 사용자를 관리하는 최전선에 있는 시스템 관리자 및 기술 담당자의 경우 특히 그렇습니다.

지원 경험 & 고객 도달

앞서 언급했듯이 고객 사이트를 삭제하는 동일한 스크립트에서 주요 고객 식별자와 연락처 정보도 삭제되었습니다 (예: 프로덕션 환경의 클라우드 URL, 사이트 시스템 관리자 연락처). 이는 핵심 시스템 (예: 지원, 라이선스, 청구) 은 모두 클라우드 URL 및 사이트 시스템 관리자 연락처의 존재를 보안, 라우팅 및 우선 순위 지정을 위한 기본 식별자로 활용합니다. 이러한 식별자를 분실했을 때 처음에는 고객을 체계적으로 식별하고 소통하는 능력을 상실했습니다.

고객에 대한 지원은 어떤 영향을 받았습니까?

첫째, 영향을 받는 고객의 대다수는 일반적인 [온라인 문의 양식을 통해 지원 팀에 연락할 수 없었습니다](#). 이 양식은 사용자가 자신의 Atlassian ID 로 로그인하고 유효한 클라우드 URL 을 제공하도록 하기 위해 고안되었습니다. 유효한 URL 이 없으면 사용자는 기술 지원 티켓을 제출할 수 없습니다. 정상적인 비즈니스 과정에서 이 확인은 사이트 보안 및 티켓 분류를 위해 의도된 것입니다. 그러나 해당 요구 사항으로 인해 이러한 중단의 영향을 받은 고객에게 의도하지 않은 결과가 발생했습니다. 고객은 우선 순위가 높은 사이트 지원 티켓을 제출할 수 없었습니다.

둘째, 인시던트로 인한 사이트 시스템 관리자 데이터의 삭제로 인해 영향을 받는 고객과 사전에 소통할 수 있는 능력에 격차가 생겼습니다. 인시던트가 발생한 첫 며칠 동안 Atlassian 에 등록된 해당 고객의 청구 및 기술 담당자에게 사전 대응적인 커뮤니케이션을 보냈습니다. 그러나 영향을 받는 고객에 대한 많은 청구 및 기술 연락처가 구식이라는 사실을 금방 확인했습니다. 각 사이트에 대한 시스템 관리자 정보가 없으면 참여할 활성 및 승인된 연락처의 전체 목록이 없었습니다.

저희는 어떻게 대응했을까요?

우리의 지원 팀은 인시던트 발생 첫 날에 현장 복원을 가속화하고 통신 채널의 파손을 수리하기 위해 똑같이 중요한 세 가지 우선 순위를 두었습니다.

먼저 검증된 고객 연락처의 신뢰할 수 있는 목록을 얻는 것입니다. 엔지니어링 팀이 고객 사이트를 복원하기 위해 노력하면서 고객 대면 팀은 검증된 연락처 정보를 복원하는 데 집중했습니다. 우리는 연락처 목록을 재구성하기 위해 모든 메커니즘(청구 시스템, 사전 지원 티켓, 기타 보안 사용자 백업, 직접 고객 지원 등)을 사용했습니다. 우리의 목표는 직접적인 지원 및 응답 시간을 간소화하기 위해 영향을 받는 각 사이트에 대해 하나의 인시던트 관련 지원 티켓을 보유하는 것이었습니다.

둘째, 이 인시던트와 관련된 워크플로, 큐 및 SLA 를 다시 설정하는 것입니다. Cloud ID 를 삭제하고 사용자를 올바르게 인증할 수 없는 경우에도 일반 시스템을 통해 인시던트 관련 지원 티켓을 처리하는 기능에 영향을 미쳤습니다. 티켓이 관련 우선 순위 및 에스컬레이션 큐 및 대시보드에 올바르게 표시되지 않았습니다. 로직, SLA, 워크플로 상태 및 대시보드를 설계하고 추가하기 위해 부서 간 팀(지원, 제품,

IT)을 신속하게 구성했습니다. 프로덕션 시스템 내에서 이 작업을 수행해야 했기 때문에 완전히 개발, 테스트 및 배포하데 며칠이 걸렸습니다.

셋째, 수동 검증을 대규모로 확장하여 사이트 복원을 가속화합니다. 엔지니어링이 초기 복원을 통해 발전함에 따라 수동 테스트 및 검증 검사를 통해 사이트 복구를 가속화하는 데 글로벌 지원 팀의 역량이 필요하다는 것이 분명해졌습니다. 이 검증 프로세스는 엔지니어링 팀이 데이터 복원을 가속화하면 고객에게 복원 사이트를 제공하는 데 중요한 경로가 될 것입니다. 우리는 표준 운영 절차(SOP), 워크플로우, 핸드오프 및 인력 명단의 독립적인 스트림을 만들어 450 명 이상의 지원 엔지니어를 동원하여 검증 점검을 실행해야 했으며, 교대 근무는 연중무휴 24 시간 서비스를 제공하여 고객의 손에 대한 복원을 가속화해야 했습니다.

첫 주 말까지 이러한 주요 우선 순위가 잘 확립되었지만 복원 프로세스의 복잡성으로 인한 인시던트 해결 일정에 대한 명확성이 부족하여 *의미있는* 업데이트를 제공 할 수 있는 능력이 제한적이었습니다. 우리는 사이트 복원 날짜를 더 빨리 제공하는데 대한 불확실성을 인정하고 고객이 그에 따라 계획을 세울 수 있도록 대면 대화에 더 일찍 참여할 수 있도록 했을 것입니다.

어떻게 개선할 수 있을까요?

당사는 변경 사항이 적절히 이루어질 때까지 즉시 대량 사이트 삭제를 차단했습니다.

이 인시던트에서 벗어나 내부 프로세스를 재평가하면서 사람들이 인시던트를 일으키지 않는다는 것을 인식하고자 합니다. 오히려 시스템은 실수를 저지를 수 있습니다. 이 섹션에는 이 인시던트에 기여한 요소가 요약되어 있습니다. 또한 이러한 약점과 문제를 해결하는 방법을 가속화하기 위한 계획에 대해서도 논의합니다.

학습 1: “소프트 삭제”는 모든 시스템에서 보편적이어야 함

전반적으로 이러한 유형의 삭제는 금지되거나 오류를 방지하기 위해 여러 계층의 보호를 적용해야 합니다. 주요 개선 사항은 소프트 삭제 프로세스를 거치지 않은 고객 데이터 및 메타데이터의 삭제를 전 세계적으로 방지하는 것입니다.

a) 데이터 삭제는 소프트 삭제로만 이루어져야 합니다.

전체 사이트의 삭제는 금지되어야 하며, 소프트 삭제는 오류를 방지하기 위해 다단계 보호가 필요합니다. 외부 스크립트나 시스템이 프로덕션 환경에서 고객 데이터를 삭제하지 못하도록 하는 “소프트 삭제” 정책을 시행합니다. 당사의 “소프트 삭제” 정책은 데이터 복구를 빠르고 안전하게 실행할 수 있도록 충분한 데이터 보존을 허용합니다. 데이터는 보존 기간이 만료된 후에만 운영 환경에서 삭제됩니다.

조치:

- ✔ **프로비저닝 워크플로 및 모든 관련 데이터 저장소에 “소프트 삭제” 구현:** 또한 테넌트 플랫폼 팀은 데이터 삭제가 비활성화 이후에만 발생할 수 있는지 확인하고 이 공간의 기타 보호 조치를 수행합니다. 장기적으로 테넌트 플랫폼은 테넌트 데이터의 올바른 상태 관리를 더욱 발전시키는 데 주도적인 역할을 할 것입니다.

b) 소프트 삭제는 표준화되고 검증된 검토 프로세스를 거쳐야 합니다.

소프트 삭제 작업은 위험도가 높은 작업입니다. 따라서 이러한 작업을 해결하기 위해 정의된 롤백 및 테스트 절차를 포함하는 표준화되거나 자동화된 검토 프로세스가 있어야 합니다.

조치:

- ✔ **일시 삭제 작업의 단계적 롤아웃 시행:** 삭제가 필요한 모든 새로운 작업은 먼저 자체 사이트 내에서 테스트되어 접근 방식을 검증하고 자동화를 검증합니다. 검증을 완료하면 고객을 동일한 프로세스로 점진적으로 이동시키고 선택한 전체 사용자 기반에 자동화를 적용하기 전에 불규칙성을 계속 테스트합니다.
- ✔ **일시 삭제 작업에는 테스트된 롤백 계획이 있어야 합니다. 데이터를 일시 삭제하는 모든 작업은:** 프로덕션에서 실행하기 전에 삭제된 데이터의 복원을 테스트하고 롤백 계획을 테스트해야 합니다.

학습 2: DR 프로그램의 일환으로서, 다수의 고객을 대상으로 한 다중 사이트, 다중 제품 삭제 이벤트의 복원 자동화

[Atlassian 데이터 관리](#)는 당사 데이터 관리 프로세스를 상세히 설명합니다. 고가용성을 보장하기 위해 여러 AWS 가용 영역(AZ)에 동기식 대기 복제본을 프로비저닝하고 유지 관리합니다. AZ 장애 조치는 자동화되며 일반적으로 60-120 초가 소요됩니다. 저희는 고객에게 영향이 가지 않도록 Data Center 가동 중단 및 기타 일반적인 중단을 정기적으로 처리합니다.

또한 데이터 손상 이벤트에 대한 복구력을 갖도록 설계된 변경 불가능한 백업을 유지 관리하여 이전 시점으로 복구할 수 있습니다. 백업은 30 일 동안 보관되며, Atlassian 복구를 위해 스토리지 백업을 지속적으로 테스트하고 감사합니다. 또한 필요한 경우, 모든 고객 사이트를 새로운 환경으로 한 번에 복구할 수 있습니다.

백업을 이용하여, 개인 고객 사이트 또는 실수로 자신의 데이터를 삭제한 소수의 고객 사이트를 정기적으로 롤백합니다. 사이트 수준의 삭제에는 이 이벤트의 규모에 맞게 신속하게 자동화할 수 있는 런북이 없었기 때문에 모든 제품 및 서비스 전반에 걸쳐 툴링 및 자동화를 조정된 방식으로 진행해야 했습니다.

아직 자동화하지 않은 것은 다른 고객 사이트에 영향을 미치지 않고 많은 고객 사이트를 기존(현재 사용 중인) 환경으로 복구하는 것입니다.

클라우드 환경 내에서 각 데이터 저장소에는 여러 고객 사이트의 데이터가 포함되어 있습니다. 이 인시던트에서 삭제된 데이터는 다른 고객이 계속 사용하는 데이터 저장소의 일부일 뿐이므로 백업에서 개별 조각을 수동으로 추출하고 복구해야 합니다. 각 고객 사이트 복구는 길고 복잡한 프로세스이므로 사이트 복구 시 내부 검증 및 최종 고객 확인이 필요합니다.

조치:



대규모 고객을 위해 멀티 제품, 멀티 사이트 복원 가속화: DR 프로그램은 현재의 RPO 표준인 1 시간을 충족합니다. 이 인시던트의 자동화 및 학습을 활용하여 DR 프로그램을 가속화하여 이러한 규모의 인시던트에 대한 정책에 정의된 대로 RTO 를 충족할 것입니다.

- ✔ **DR 테스트에 이 사례의 검증 자동화 및 추가:** 대규모 사이트에 대한 모든 제품을 복원하는 DR 연습을 정기적으로 실행합니다. 이러한 DR 테스트는 아키텍처가 발전하고 새로운 에지 사례가 발생할 때 운전 설명서가 최신 상태인지 확인합니다. 우리는 복원 접근 방식을 지속적으로 개선하고 더 많은 복원 프로세스를 자동화하며 복구 시간을 줄일 것입니다.

학습 3: 대규모 이벤트에 대한 인시던트 관리 프로세스 개선

당사의 인시던트 관리 프로그램은 수년 동안 발생한 주요 및 사소한 인시던트를 관리하는 데 적합합니다. 일반적으로 더 적은 인력과 팀이 참여하는 소규모의 단기 인시던트에 대한 인시던트 대응 시뮬레이션하는 경우가 많습니다.

이 사건이 가장 심각한 수준에 닿은 시점에는 고객 사이트 복원을 위해 수백 명의 엔지니어와 고객 지원 직원들이 동시에 작업했으나, 당사의 인시던트 관리 프로그램과 팀은 이러한 유형의 사고의 깊이, 범위 및 시간을 처리할 수 있도록 설계된 것이 아니었습니다(아래 그림 10 참조).

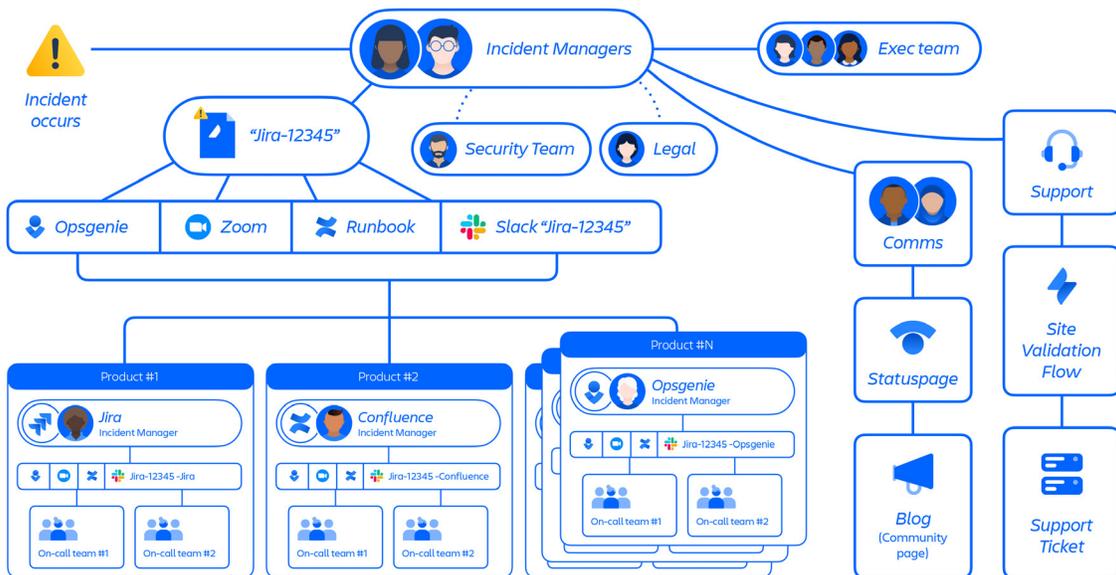


그림 10: 대규모 인시던트 관리 절차 개요

대규모 인시던트 관리 프로세스가 더 잘 정의되고 자주 실행됩니다.

제품 수준 인시던트에 대한 플레이북이 있지만 회사 전체에서 수백 명의 사람들이 동시에 작업하는 이러한 규모의 이벤트에 대한 플레이북은 없습니다. 인시던트 관리 도구에는 Slack, Zoom 및 Confluence doc 과 같은 통신 스트림을 생성하는 자동화 기능이 있지만 복원 스트림을 격리하기 위해 대규모 사고에 필요한 하위 스트림을 생성하지 못합니다.

조치:



대규모 인시던트에 대한 플레이북 및 도구 정의 및 시뮬레이션 연습 수행: 대규모로 간주될 수 있고 이 수준의 대응이 필요한 인시던트 유형을 정의하고 문서화합니다. 주요 조정 단계를 간략하게 설명하고 사고 관리자 및 기타 비즈니스 부서에서 대응을 간소화하고 복구를 시작하는 데 도움이 되는 도구를 구축합니다. 팀이 있는 인시던트 관리자는 지속적으로 개선하기 위해 정기적으로 시뮬레이션, 교육 및 도구 및 문서의 개선을 실행합니다.

학습 4: 커뮤니케이션 프로세스 개선

a) 중요한 고객 식별자를 삭제하여 영향을 받는 사람들에게 대한 커뮤니케이션 및 조치에 영향을 미쳤습니다.

고객 사이트를 삭제한 동일한 스크립트에서 주요 고객 식별자도 삭제했습니다(예: 프로덕션 환경의 클라우드 URL, 사이트 시스템 관리자 연락처). 그 결과 (1) 고객은 일반적인 지원 채널을 통해 기술 지원 티켓을 제출할 수 없었습니다. (2) 인시던트 대응으로 인해 서비스 중단으로 인해 영향을 받은 주요 고객 연락처(예: 사이트 시스템 관리자)의 신뢰할 수 있는 목록을 얻는 데 며칠이 걸렸습니다. (3) 인시던트의 고유한 특성 때문에 처음에는 워크플로우, SLA, 대시보드, 에스컬레이션 프로세스가 제대로 작동하지 않았습니다.

서비스 중단 기간 동안 고객 에스컬레이션은 여러 채널(이메일, 전화 통화, CEO 티켓, LinkedIn 및 기타 소셜 채널, 지원 티켓)을 통해서도 이루어졌습니다. 고객 응대 팀 전체의 서로 다른 도구와 프로세스로 인해 대응이 느려지고 이러한 에스컬레이션에 대한 전체적인 추적 및 보고가 더욱 어려워졌습니다.

b) 이러한 수준의 복잡성을 처리할 수 있을 만큼 철저한 인시던트 커뮤니케이션 플레이북이 없었습니다.

통합 교차 기능 인시던트 커뮤니케이션 팀을 충분히 신속하게 동원하기 위한 원칙과 역할 및 책임을 설명하는 인시던트 커뮤니케이션 플레이북이 없었습니다. 여러 채널, 특히 소셜 미디어를 통해 인시던트에 대한 승인을 신속하고 일관되게 제공하지 않았습니다. 데이터 손실이 없었고 이것이 사이버 공격의 결과가 아니라 중요한 메시지의 반복과 함께 중단 둘러싼 더 광범위한 공개 커뮤니케이션이 올바른 접근 방식이었을 것입니다.

조치:

- ✓ **주요 연락처 백업 개선:** 승인된 계정 연락처 정보를 제품 인스턴스 외부에 백업합니다.
- ✓ **지원 도구 개조:** 유효한 사이트 URL 또는 Atlassian ID 가 없는 고객을 위한 메커니즘을 만들어 기술 지원팀에 직접 연락할 수 있습니다.
- ✓ **고객 에스컬레이션 시스템 및 프로세스:** 여러 작업 객체(티켓, 작업 등)를 단일 고객 계정 객체 아래에 저장하여 당사 고객 응대 팀 간의 조정 및 가시성을 개선할 수 있는 통합 계정 기반 에스컬레이션 시스템 및 워크플로우에 투자합니다.
- ✓ **연중무휴 24 시간 에스컬레이션 관리 지원 촉진:** 에스컬레이션 관리 기능에 대한 글로벌 공간 확장 계획을 실행하여 각 주요 지역에 기반을 둔 지정된 직원과 필요한 제품 및 영업 주제 파트너 및 리더십을 지원하는 지원 역할을 통해 일관된 연중무휴 24 시간 지원을 제공합니다.
- ✓ **새로운 학습 내용으로 인시던트 커뮤니케이션 플레이북을 업데이트하고 정기적으로 다시 검토합니다:** 플레이북을 다시 방문하여 내부적으로 명확한 역할과 커뮤니케이션 라인을 정의합니다. 사고에 [DACI](#) 프레임워크를 사용하고 질병, 휴일 또는 기타 예상치 못한 사건이 발생할 경우 각 역할에 대해 연중무휴 24 시간 백업합니다.

조치(계속)

모든 커뮤니케이션에서 인시던트 커뮤니케이션 템플릿 따르십시오. 발생한 상황, 영향을 받은 사람, 복원 일정, 사이트 복원 비율, 예상 데이터 손실, 관련 신뢰 수준, 지원 팀에 문의하는 방법에 대한 명확한 지침을 제공합니다.

마무리하며 짚을 요소

서비스 중단이 해결되고 고객이 완전히 복원되는 동안 우리의 작업은 계속됩니다. 이 단계에서는 프로세스를 개선하고 복원력을 높이며 이와 같은 상황이 다시 발생하지 않도록 위에서 설명한 변경 사항을 구현하고 있습니다.

Atlassian 학습 조직이며, 우리 팀은 이 경험을 통해 많은 어려운 교훈을 확실히 배웠습니다. 우리는 비즈니스에 지속적인 변화를 만들기 위해 이러한 교훈을 실천하고 있습니다. 궁극적으로 우리는 이러한 경험을 통해 더 강해지고 더 나은 서비스를 제공 할 것입니다.

이 사건의 교훈이 고객에게 안정적인 서비스를 제공하기 위해 부지런히 노력하는 다른 팀에게도 도움이되기를 바랍니다.

마지막으로, 이 글을 읽고 우리와 함께 배우신 분들과 확장된 Atlassian 커뮤니티 및 팀의 일원인 분들께 감사드립니다.

-Sri Viswanath, CTO