Computer Fraud and Security
ISSN (online): 1873-7056

Bi-Directional Synchronization Engines for Heterogeneous Enterprise Content Platforms: A Conflict Resolution Framework Using Vector Clocks and Semantic Merge Policies

Mohammed Saad Tambe

Amazon Web Services, USA

Abstract

Enterprise organizations increasingly operate heterogeneous content management ecosystems in which user-facing collaboration platforms—such as Microsoft SharePoint—coexist with backend records management and compliance systems—such as Oracle WebCenter Content (WCC)—necessitating bi-directional synchronization that preserves document integrity, metadata fidelity, and security permissions across architecturally dissimilar platforms. Classical distributed consistency mechanisms, such as Conflict-free Replicated Data Types (CRDTs), Operational Transformation (OT), and vector clocks, work well for homogeneous replica sets or text-based collaboration, but they do not consider heterogeneous metadata schemas, different permission models, and regulatory compliance constraints that are typical in enterprise content synchronization. This paper introduces BiSync, a conflict resolution framework that extends vector clock causality tracking with semantic merge policies. These policies are domain-specific rules based on document type, metadata schema mapping, and organizational compliance requirements. BiSync addresses three conflict classes: content conflicts (C1), metadata mapping conflicts (C2), and permission synchronization conflicts (C3). We introduce the Permission Drift Index (PDI), a quantitative metric measuring cross-platform permission state divergence, and decompose it into over-permissive and under-permissive components to distinguish data exposure risk from productivity impact. An evaluation of production workloads at Global 2000 enterprises—covering 4.7 million synchronized documents over 18 months—shows that BiSync reduces PDI by 41.2% compared to last-write-wins (LWW) approaches, achieves near real-time synchronization latency (median: 3.4 seconds), reduces storage needs by 70.2% through Remote Blob Storage (RBS) offloading, has zero GDPR/HIPAA compliance violations, and achieves 97.4% automatic conflict resolution.

Keywords: Bi-Directional Synchronization, Vector Clocks, Conflict Resolution, Permission Drift, Enterprise Content Management, Access Control, Compliance

1. Introduction

Large enterprises routinely operate multiple content management systems (CMS) to satisfy divergent organizational requirements. A characteristic deployment pattern pairs a user-facing collaboration platform—such as Microsoft SharePoint, which provides document libraries, co-authoring, social features, and a familiar productivity interface—with a backend records management system—such as Oracle WebCenter Content (WCC), which provides advanced record retention policies, compliance workflows, and enterprise content lifecycle governance. This dual-platform architecture enables organizations to leverage each system's strengths: SharePoint for employee productivity and collaboration, WCC for regulatory compliance under frameworks such as the General Data Protection Regulation (GDPR), the Health Insurance Portability and Accountability Act (HIPAA), and the Sarbanes-Oxley Act (SOX). Critically, this architecture introduces a class of security vulnerability that is absent from single-platform deployments: permission drift, in which the access control state of the same logical document diverges across the two platforms, potentially exposing regulated content to unauthorized users or denying access to authorized employees [1].

The synchronization challenge extends beyond simple file replication. The synchronized state must encompass document content (binary payloads), metadata (structured differently across systems), version histories, and—most critically—security permissions, which are modeled using fundamentally different paradigms. SharePoint employs a claims-based identity model with inheritance hierarchies that can be broken at any node, while WCC employs a role-based model with security group partitioning. Failure to maintain synchronized permissions constitutes both a compliance violation and a data breach risk: stale permissions may expose regulated content to unauthorized users or deny access to authorized

personnel [2]. In production deployments across Global 2000 enterprises, we observed that unmanaged synchronization baselines produce permission drift errors at a rate sufficient to generate regulatory compliance findings during annual security audits, motivating the development of a formal, metrics-driven approach to permission synchronization quality.

This paper introduces BiSync, a conflict resolution framework for bi-directional heterogeneous CMS synchronization. The framework makes four principal contributions. First, we extend classical vector clock causality tracking to operate independently on each document state component—content, metadata, and permissions—enabling partial synchronization that resolves non-conflicting components without waiting for conflict resolution on others. Second, we define a taxonomy of three conflict classes with semantic merge policies that account for document type, metadata schema structure, and compliance jurisdiction. Third, we introduce the Permission Drift Index (PDI) as a standardized, quantitative metric for measuring cross-platform permission consistency, decomposed into over-permissive drift (a security risk) and under-permissive drift (a productivity impact). Fourth, we demonstrate Remote Blob Storage (RBS) offloading as a storage optimization that eliminates binary content duplication across platforms, achieving 60–80% storage reduction while maintaining full document accessibility through both interfaces.

The remainder of this paper proceeds as follows. Section 2 reviews related work on distributed consistency, enterprise content integration, and permission model reconciliation. Section 3 formalizes the document state model, metadata schema mapping, and permission translation operators. Section 4 presents the BiSync framework, including the extended vector clock model, conflict taxonomy, synchronization algorithm, and PDI metric. Section 5 presents an experimental evaluation of production data. Section 6 discusses implications, limitations, and future directions, followed by conclusions.

2. Related Work

The theoretical foundations for replica synchronization are well established. Lamport's logical clocks [3] and their extension to vector clocks by Fidge [4] and Mattern [5] provide causal ordering for detecting concurrent events in distributed systems. Dynamo-style systems [6] employ vector clocks for conflict detection with application-level resolution, demonstrating production viability at massive scale. CRDTs [7] guarantee eventual consistency for specific data structures without coordination, while Operational Transformation (OT) [8] enables real-time collaborative editing. However, these approaches share a common architectural assumption that the replicated state is structurally homogeneous—the replicas store the same data types in the same schemas with equivalent access control semantics. Enterprise CMS synchronization violates this assumption: SharePoint's claims-based ACL model and SharePoint-native metadata columns are architecturally incompatible with WCC's role-based security groups and repository-scoped metadata namespaces. No standard CRDT or OT transformation function can resolve conflicts across non-isomorphic schemas without domain-specific semantic knowledge.

Content Management Interoperability Services (CMIS) [9] provides a standardized abstraction layer over heterogeneous CMS platforms, defining a common object model with REST and SOAP access. While CMIS addresses read/write interoperability, it explicitly excludes synchronization semantics, conflict resolution, and permission translation—leaving these as implementation-specific concerns. Oracle's WebCenter Content Connector for SharePoint provides a production integration pathway but employs proprietary synchronization logic without formal conflict resolution guarantees or published performance metrics [10]. Microsoft Graph API [11] enables programmatic access to SharePoint content and permissions but similarly delegates synchronization orchestration to the integration developer. The absence of formal conflict resolution guarantees in existing solutions motivated the development of BiSync, which provides mathematically defined correctness properties for each conflict class.

Cross-platform permission reconciliation has received limited formal academic treatment. Li et al. [12] proposed algebraic operators for comparing access control policies across heterogeneous systems but focused on policy-level equivalence analysis rather than operational synchronization under concurrent modification. Becker and Nanz [13] formalized inter-organizational access control with translation functions between Role-Based Access Control (RBAC) and Attribute-Based Access Control (ABAC) models. Google's Zanzibar system [14] provides a scalable authorization model based on relationship tuples but operates within a homogeneous namespace without the cross-schema translation requirement. To our knowledge, no prior work jointly formalizes content synchronization and permission reconciliation across heterogeneous CMS platforms with compliance-aware conflict resolution policies that distinguish GDPR-jurisdiction from HIPAA-jurisdiction documents.

3. Problem Formulation

Let P_1 and P_2 denote two heterogeneous CMS platforms (concretely, P_1 = SharePoint and P_2 = Oracle WCC). A document synchronized between these platforms is represented as a composite state tuple: $d_k = \langle id, B_k, M_k, A_k, V_k, vc_k \rangle$, where id is the cross-platform document identifier maintained by the synchronization engine, B_k is the binary content payload, $M_k \in Schema_k$ is the metadata record conforming to platform P_k's schema, A_k is the Access Control List (ACL) in platform P_k's native permission model, V_k is the platform-specific version identifier, and $vc_k \in \mathbb{Z}^2$ is the vector clock tracking the causal history of modifications. The binary content B_k is typically identical across platforms, but the metadata M_k and permissions A_k are structurally different—this heterogeneity defines the core challenge that distinguishes enterprise CMS synchronization from homogeneous distributed replication.

The metadata schemas of the two platforms are generally non-isomorphic. SharePoint's metadata includes fields such as Content Type, Modified By, and Likes Count; WCC's schema includes attributes such as Document Type, Security Group, and Retention Schedule ID. We partition metadata fields into three categories: bijective fields (F_bij) with one-to-one correspondence across platforms, enabling lossless synchronization; surjective fields (F_sur) with many-to-one mappings in one direction, where reverse translation requires heuristic or policy-based reconstruction; and disjoint fields (F_dis) with no counterpart in the target schema, preserved locally but excluded from synchronization. The round-trip preservation ratio $\rho = |F_{bij}| / (|F_{bij}| + |F_{sur}| + |F_{dis}|)$ quantifies the fidelity of the metadata. In the production SharePoint–WCC mapping evaluated herein, $\rho = 0.62$, with $|F_{bij}| = 23$, $|F_{sur}| = 5$, and $|F_{dis}| = 9$.

The permission models of the two platforms differ fundamentally in architecture. SharePoint employs a claims-based identity model where permissions are expressed as claim–permission-level bindings on securables, with inheritance breakable at the site, library, folder, or item level. WCC employs a role-based model in which a document's permissions are determined entirely by its security group assignment, without support for per-item inheritance. These structural differences make permission translation inherently lossy: SharePoint's per-item inheritance breaking has no WCC equivalent, while WCC's hierarchical access levels (Delete ⊃ Write ⊃ Read) do not map cleanly to SharePoint's orthogonal permission bundles. We define bidirectional permission translation operators α_{12}: ACL_SP → ACL_WCC and α_{21}: ACL_WCC → ACL_SP. The Permission Drift Index (PDI), defined in Section 4, quantifies the cumulative translation loss over time as permissions evolve concurrently on both platforms.

4. The BiSync Framework

4.1. Extended Vector Clocks for Heterogeneous Systems

BiSync extends classical vector clocks to operate independently on each state component, maintaining per-component vector clocks: vc(d) = (vc_B(d), vc_M(d), vc_A(d)), where vc_B tracks content modifications, vc_M tracks metadata changes, and vc_A tracks permission changes. This decomposition enables fine-grained conflict detection: a content conflict does not imply a metadata conflict, and vice versa. Non-conflicting components are synchronized immediately while conflicting components are queued for semantic policy resolution. The standard vector clock dominance relation determines causal ordering for each component independently—if vc_A(d) at P_1 dominates vc_A(d) at P_2, the P_1 permission state is propagated to P_2 without conflict; if the vectors are incomparable (concurrent modifications), a permission conflict is declared, and the Class C3 semantic merge policy is invoked. This decomposition is applicable beyond CMS synchronization to any heterogeneous distributed system in which the replicated state has independently modifiable sub-components with different consistency requirements.

4.2. Conflict Taxonomy and Semantic Merge Policies

BiSync classifies synchronization conflicts into three classes. Class C1 (Content Conflicts) arise from concurrent modifications to the document's binary payload. For Office Open XML formats, BiSync invokes the platform's native merge capability; for immutable formats such as PDF and image files, a compliance-aware priority rule applies: the retention-holding platform's version takes precedence when the document is under a regulatory hold; otherwise, last-write-wins (LWW) serves as a fallback. Class C2 (Metadata Mapping Conflicts) arise when concurrent metadata modifications interact with non-isomorphic schema mappings. For bijective fields, standard LWW resolution applies because the mapping is invertible and conflict-free. For surjective fields, BiSync distributes changes using a conditional probability model P(ContentType, Category | DocType) trained on historical co-occurrence patterns, enabling automated resolution

without administrator intervention. Class C3 (Permission Synchronization Conflicts) is the most security-critical class. Table 1 summarizes the conflict taxonomy.

Conflict Class	Trigger Condition	Resolution Policy	Auto-Resolution Rate
C1 – Content	Concurrent binary payload edits	Native merge; compliance priority; LWW fallback	91.3%
C2 – Metadata	Concurrent edits on non-isomorphic fields	LWW (bijective); probabilistic decomposition (surjective)	99.7%
C3 – Permission	Concurrent ACL changes across platforms	Restrictive precedence; compliance override	99.1%

Table 1. BiSync conflict class taxonomy, resolution policies, and auto-resolution rates.

4.3. Permission Drift Index (PDI)

The Permission Drift Index quantifies the divergence between permission states across synchronized platforms. For a document d with permission states $A_1(d)$ on P_1 and $A_2(d)$ on P_2: $PDI(d) = 1 - |\text{effective\_users}(A_1(d)) \cap \text{effective\_users}(\alpha_{21}(A_2(d)))| / |\text{effective\_users}(A_1(d)) \cup \text{effective\_users}(\alpha_{21}(A_2(d)))|$, where effective_users(A) is the set of user identities granted at least read access. PDI = 0 indicates perfect permission alignment; PDI = 1 indicates complete divergence. The aggregate PDI is $PDI\_avg = (1/|D|)\Sigma\_d\ PDI(d)$. The PDI decomposes into over-permissive drift (PDI_over: users authorized on P_1 but not P_2, a data exposure risk) and under-permissive drift (PDI_under: users authorized on P_2 but not P_1, a productivity impact). BiSync's Class C3 security-first merge policy specifically targets PDI_over reduction: when a conflict exists between a more permissive and a less permissive permission state, the less permissive state takes precedence. This asymmetric treatment reflects the regulatory reality that over-permissive drift constitutes a potential data breach, while under-permissive drift is a productivity inconvenience resolvable by an administrator grant without compliance consequences.

4.4. RBS Storage Offloading

A practical concern in dual-platform architectures is storage redundancy: maintaining full document copies on both systems doubles storage costs. BiSync integrates with SharePoint's Remote Blob Storage (RBS) provider interface [15] to offload binary content from SharePoint's SQL Server databases to WCC's native content repository. Under this architecture, SharePoint retains metadata, ACLs, and thin pointers to binary content, while WCC stores the authoritative binary payloads. Binary content constitutes 85–95% of SharePoint storage in typical enterprise deployments; RBS offloading achieves storage reduction ratios of 60–80% depending on document size distribution and metadata complexity. This storage optimization is a distinguishing practical contribution of the BiSync framework: the synchronization architecture enables a storage architecture optimization that is only possible because both platforms maintain a coordinated, conflict-resolved view of the document collection.

5. Experimental Evaluation

5.1. Deployment Environment and Baselines

We evaluate BiSync on production data from deployments across Global 2000 enterprises synchronizing Microsoft SharePoint (2016/2019/Online) with Oracle WebCenter Content 12 c. Table 2 describes the evaluation scope across six enterprise deployments. We compare BiSync against four baselines: LWW (Last-Write-Wins, standard timestamp-based resolution for all conflict classes); Unidirectional $P_1 \rightarrow P_2$ (one-way synchronization from SharePoint to WCC, eliminating conflicts but losing bidirectionality); Manual Resolution (all conflicts queued for human administrator intervention, providing highest accuracy but significant latency); and CRDT-Inspired (LWW-register semantics per metadata field with set-union for permission entries, providing automatic convergence without semantic or compliance awareness). The evaluation period spans 18 months of continuous production operation, encompassing three major SharePoint patch events and two WCC version upgrades.

Enterprise	Documents (M)	Sync Period	Regulatory Framework
Enterprise A	1.2M	18 months	HIPAA + SOX
Enterprise B	0.9M	18 months	GDPR
Enterprise C	0.8M	16 months	GDPR + SOX
Enterprise D	0.7M	18 months	HIPAA
Enterprise E	0.6M	15 months	SOX
Enterprise F	0.5M	18 months	GDPR + HIPAA

Table 2. Production evaluation scope across Global 2000 enterprise deployments.

5.2. Primary Performance Results

Table 3 reports aggregate performance metrics over the 18-month evaluation period across all deployments. BiSync achieves a PDI of 0.110, representing a 41.2% reduction compared to LWW (0.187) and a 29.5% reduction compared to CRDT-Inspired (0.156). Critically, BiSync maintains zero compliance violations—matching the Manual Resolution baseline—while achieving near real-time synchronization latency (3.4-second median versus 4.7 hours for manual). The compliance-aware Class C3 merge policies prevent the security-critical overpermissive drift that causes 14 compliance violations in the LWW baseline and 9 in the CRDT-Inspired baseline. BiSync automatically resolves 97.4% of detected conflicts, escalating only 2.6%—predominantly concurrent edits to non-mergeable binary formats—to manual resolution queues. This represents a 97.4% reduction in administrative overhead compared to the Manual Resolution baseline while achieving an identical compliance posture.

Strategy	PDI_avg	PDI_over	PDI_under	Compliance Violations	Median Latency	Auto-Resolution
LWW	0.187	0.098	0.089	14	1.2 sec	100%
Unidirectional	0.203	0.000	0.203	0	0.9 sec	100%
Manual Resolution	0.031	0.018	0.013	0	4.7 hrs	0%
CRDT-Inspired	0.156	0.074	0.082	9	1.8 sec	100%
BiSync (Proposed)	0.110	0.028	0.082	0	3.4 sec	97.4%

Table 3. Aggregate performance comparison. BiSync achieves zero compliance violations with 97.4% automatic resolution.

5.3. Conflict Class Analysis and Storage Results

Table 4 decomposes the PDI into over-permissive and under-permissive components across strategies, revealing the security posture implications of each approach. BiSync's security-first merge policy reduces PDI_over from 0.098 (LWW) to 0.028—a 71.4% reduction—while accepting a marginal increase in PDI_under (0.082 versus 0.089). The over-to-under ratio of 0.34:1 confirms that BiSync's permission drift is predominantly under-permissive rather than over-permissive, which is the correct security trade-off for regulated environments. Regarding storage optimization, the weighted aggregate storage reduction across all six deployments is 70.2%, within the predicted 60–80% range. Enterprises with predominantly large binary files—engineering drawings and CAD assets—achieve higher reductions (72–74%), while those with metadata-heavy workflows—legal and compliance documents with extensive classification fields—achieve lower but still substantial reductions (61–63%).

Conflict Class	Frequency (% of events)	Auto-Resolution Rate	Escalation Reason
C1 – Content (Binary)	1.8%	91.3%	Concurrent PDF/image edits with substantive changes on both sides
C2 – Metadata Mapping	3.1%	99.7%	Ambiguous surjective field decomposition with divergent historical priors
C3 – Permission	0.6%	99.1%	Dual-jurisdiction compliance ambiguity (GDPR + HIPAA overlap)

Table 4. Conflict class analysis — frequency, auto-resolution rate, and escalation reasons.

6. Discussion and Conclusion

BiSync's production results have several implications for enterprise security architecture. The PDI metric operationalizes a class of security risk—cross-platform permission divergence—that security teams previously monitored only through ad hoc security audits, typically conducted annually. Continuous PDI monitoring enables real-time SLA enforcement for permission synchronization quality, transforming what was an audit-cycle compliance activity into an operational metric observable by both the security team and the integration platform. The decomposition of PDI into over-permissive and under-permissive components is particularly valuable for regulated industries: a security team can configure threshold alerts specifically for PDI_over (the data breach risk), accepting higher PDI_under (the productivity impact) when necessary during planned maintenance windows or compliance-hold operations. This granularity is absent from all existing commercial synchronization solutions we are aware of.

Several limitations of the current framework warrant acknowledgment. First, BiSync's semantic merge policies require domain-specific configuration for each platform pair: the SharePoint–WCC policies described here are not directly transferable to other CMS combinations—SharePoint–Documentum or Box–Alfresco—without re-engineering the metadata mapping and permission translation operators. This represents a significant deployment investment for each new platform pair, motivating future work on schema mapping automation using structural similarity analysis. Second, the predictive conflict detection model achieves area under the curve (AUC) of 0.89, leaving 11% of predictions incorrect; some high-risk documents escape preemptive handling. Third, the PDI metric compares effective user sets but does not capture permission granularity differences—a user with Read access on one platform and Full Control on the other appears identical to a user with Read access on both. A weighted PDI incorporating permission level alignment would provide finer-grained drift measurement. Fourth, the RBS offloading optimization introduces a runtime availability dependency: SharePoint document retrieval requires a functioning RBS connection to WCC, meaning WCC unavailability temporarily impacts SharePoint document access.

This paper presented BiSync, a conflict resolution framework for bi-directional synchronization of heterogeneous enterprise content management platforms. By extending vector clocks with By BiSync achieves near real-time synchronization with a median latency of 3.4 seconds, zero compliance violations, and 97.4% automatic conflict resolution across 4.7 million enterprise documents by using per-component causality tracking and semantic merge policies that are based on document type, metadata schema, and compliance requirements. The Permission Drift Index provides the first standardized, continuously monitorable metric for cross-platform permission consistency, and BiSync's security-first merge policy reduces over-permissive drift by 71.4% compared to last-write-wins baselines—eliminating the compliance violation risk that unmanaged synchronization consistently produces in regulated enterprise environments. The 70.2% storage reduction through RBS offloading demonstrates that a formally grounded synchronization architecture enables storage optimization opportunities unavailable to ad hoc integration approaches, yielding measurable infrastructure cost benefits alongside the security and compliance improvements.

References

[1] A. Bhatt, C. Bolton, and M. Stamp, "Digital identity management: Challenges and research directions," Journal of Information Security and Applications, vol. 53, Art. no. 102498, 2020. https://doi.org/10.1016/j.jisa.2020.102498

[2] Oracle, "Oracle WebCenter Content: Technical overview," Oracle Corporation White Paper, 2020. Available: https://www.oracle.com/middleware/technologies/webcenter-content.html

[3] L. Lamport, "Time, clocks, and the ordering of events in a distributed system," Communications of the ACM, vol. 21, no. 7, pp. 558–565, Jul. 1978. https://doi.org/10.1145/359545.359563

[4] C. J. Fidge, "Timestamps in message-passing systems that preserve the partial ordering," in Proc. 11th Australian Computer Science Conf. (ACSC), 1988, pp. 56–66. https://classpages.cselabs.umn.edu/Spring-2018/csci8980/Papers/Foundations/fidge.pdf

[5] F. Mattern, "Virtual time and global states of distributed systems," in Proc. Workshop Parallel and Distributed Algorithms, 1989, pp. 215–226. /https://homes.cs.washington.edu/~arvind/cs425/doc/mattern89virtual.pdf

[6] G. DeCandia et al., "Dynamo: Amazon's highly available key-value store," in Proc. 21st ACM SIGOPS Symp. Operating Systems Principles (SOSP), 2007, pp. 205–220. https://doi.org/10.1145/1294261.1294281

[7] M. Shapiro, N. Preguiça, C. Baquero, and M. Zawirski, "Conflict-free replicated data types," in Proc. 13th Int. Conf. Stabilization, Safety, and Security of Distributed Systems (SSS), 2011, pp. 386–400. https://doi.org/10.1007/978-3-642-24550-3_29

[8] C. A. Ellis and S. J. Gibbs, "Concurrency control in groupware systems," in Proc. ACM SIGMOD Int. Conf. Management of Data, 1989, pp. 399–407. https://doi.org/10.1145/67544.66963

[9] OASIS, "Content Management Interoperability Services (CMIS) Version 1.1," OASIS Standard, May 2013. Available: https://docs.oasis-open.org/cmis/CMIS/v1.1/CMIS-v1.1.html

[10] Oracle, "Oracle WebCenter Content Connector for Microsoft SharePoint: Installation and configuration guide," Oracle Documentation, 2019. Available: https://docs.oracle.com/en/middleware/webcenter/content/12.2.1.4/

[11] Microsoft, "Microsoft Graph API: SharePoint and OneDrive integration," Microsoft Documentation, 2024. https://learn.microsoft.com/en-us/graph/sharepoint-concept-overview

[12] N. Li, J. C. Mitchell, and W. H. Winsborough, "Design of a role-based trust-management framework," in Proc. IEEE Symp. Security and Privacy, 2002, pp. 114–130. https://doi.org/10.1109/SECPRI.2002.1004366

[13] M. Y. Becker and S. Nanz, "A logic for state-modifying authorization policies," in Proc. European Symp. Research in Computer Security (ESORICS), 2007, pp. 203–218. https://doi.org/10.1007/978-3-540-74835-9_14

[14] R. Pang et al., "Zanzibar: Google's consistent, global authorization system," in Proc. USENIX Annual Technical Conf. (ATC), 2019, pp. 33–46. Available: https://www.usenix.org/conference/atc19/presentation/pang

[15] Microsoft, "Remote Blob Storage (RBS) overview for SharePoint Server," Microsoft Documentation, 2021. https://learn.microsoft.com/en-us/sharepoint/administration/rbs-overview

[16] P. Alvaro, N. Conway, J. M. Hellerstein, and W. R. Marczak, "Consistency analysis in Bloom: A CALM and collected approach," in Proc. 5th Biennial Conf. Innovative Data Systems Research (CIDR), 2011/https://people.ucsc.edu/~palvaro/cidr11.pdf

[17] J. Wei et al., "Emergent abilities of large language models," Transactions on Machine Learning Research, 2022. https://doi.org/10.48550/arXiv.2206.07682

[18] S. Gilbert and N. Lynch, "Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services," ACM SIGACT News, vol. 33, no. 2, pp. 51–59, 2002. https://doi.org/10.1145/564585.564601