

Contenidos

| | |
|---------------|----|
| Prefacio..... | xi |
|---------------|----|

Parte I. Fundamentos de los sistemas de datos

| | |
|--|-----------|
| 1. Aplicaciones confiables, escalables y mantenibles..... | 3 |
| Reflexiones sobre los sistemas de datos..... | 4 |
| Confiabilidad..... | 6 |
| Fallos de <i>hardware</i> | 7 |
| Errores de <i>software</i> | 9 |
| Errores humanos..... | 9 |
| ¿Cuál es la importancia de la confiabilidad?..... | 11 |
| Escalabilidad..... | 11 |
| Descripción de la carga..... | 11 |
| Descripción del rendimiento..... | 14 |
| Enfoques para hacer frente a la carga..... | 18 |
| Mantenimiento..... | 20 |
| Operatividad: facilitar la vida a las operaciones..... | 20 |
| Simplicidad: gestión de la complejidad..... | 22 |
| Evolución: facilitar el cambio..... | 23 |
| Resumen..... | 24 |
| Referencias..... | 25 |
| | |
| 2. Modelos de datos y lenguajes de consulta..... | 29 |
| El modelo relacional frente al modelo de documentos..... | 30 |
| El nacimiento de NoSQL..... | 31 |
| El desajuste objeto-relacional..... | 32 |
| Relaciones de muchos a uno y muchos a muchos..... | 35 |
| ¿Están las bases de datos de documentos repitiendo la historia?..... | 38 |
| Bases de datos relacionales frente a las de documentos en la actualidad..... | 41 |
| Lenguajes de consulta de datos..... | 45 |
| Consultas declarativas en la web..... | 47 |

| | |
|--|------------|
| Consultas de MapReduce | 49 |
| Modelos de datos de tipo grafo..... | 52 |
| Grafos de propiedades..... | 53 |
| Lenguaje de consulta Cypher | 55 |
| Consulta de gráficos en SQL | 57 |
| Almacenes triples y SPARQL..... | 59 |
| Los fundamentos: Datalog..... | 63 |
| Resumen..... | 66 |
| Referencias..... | 67 |
| 3. Almacenamiento y recuperación | 73 |
| Estructuras de datos que potencian la base de datos | 74 |
| Índices <i>hash</i> | 76 |
| SSTables y árboles LSM..... | 80 |
| Árboles B..... | 85 |
| Comparación de los árboles B con los árboles LSM | 89 |
| Otras estructuras de indexación | 91 |
| ¿Procesamiento de transacciones o análisis?..... | 96 |
| Almacén de datos..... | 98 |
| Estrellas y copos de nieve: esquemas para el análisis | 100 |
| Almacenamiento orientado a columnas..... | 102 |
| Compresión de columnas | 104 |
| Orden de clasificación en el almacenamiento por columnas | 107 |
| Escritura en el almacenamiento orientado a columnas..... | 108 |
| Agregación: cubos de datos y vistas materializadas | 109 |
| Resumen..... | 111 |
| Referencias..... | 112 |
| 4. Codificación y evolución | 119 |
| Formatos de codificación de datos..... | 120 |
| Formatos específicos para cada lenguaje | 121 |
| JSON, XML y variantes binarias | 122 |
| Thrift y Protocol Buffers | 125 |
| Avro | 130 |
| Méritos de los esquemas | 136 |
| Modos de flujo de datos..... | 137 |
| Flujo de datos a través de bases de datos | 137 |
| Flujo de datos a través de servicios: REST y RPC..... | 140 |
| Flujo de datos por paso de mensajes | 146 |
| Resumen..... | 148 |
| Referencias..... | 150 |

Parte II. Datos distribuidos

| | |
|--|------------|
| 5. Replicación..... | 161 |
| Líderes y seguidores..... | 162 |
| Replicación síncrona frente a asíncrona..... | 163 |
| Cómo configurar nuevos seguidores..... | 165 |
| Gestión de las interrupciones de los nodos..... | 166 |
| Implementación de <i>logs</i> de replicación..... | 169 |
| Problemas de retardo en la replicación..... | 172 |
| Lectura de nuestras propias escrituras..... | 173 |
| Lecturas monotónicas..... | 175 |
| Lecturas de prefijos coherentes..... | 176 |
| Soluciones para el retardo en la replicación..... | 178 |
| Replicación multilíder..... | 178 |
| Casos de uso de la replicación multilíder..... | 179 |
| Gestión de conflictos de escritura..... | 182 |
| Topologías de replicación multilíder..... | 186 |
| Replicación sin líder..... | 188 |
| Escritura en la base de datos cuando un nodo no funciona..... | 189 |
| Limitaciones de la coherencia del <i>quorum</i> | 192 |
| <i>Quorum</i> descuidados y transferencias indirectas..... | 195 |
| Detección de escrituras simultáneas..... | 196 |
| Resumen..... | 204 |
| Referencias..... | 206 |
| | |
| 6. Particionado..... | 213 |
| Particionado y replicación..... | 214 |
| Particionado de datos clave-valor..... | 215 |
| Particionado por rangos de claves..... | 216 |
| Particionado por <i>hash</i> de claves..... | 217 |
| Cargas de trabajo desbalanceadas y mitigación de puntos calientes..... | 219 |
| Particionado e índices secundarios..... | 220 |
| Particionado de índices secundarios por documento..... | 220 |
| Particionado de índices secundarios por término..... | 222 |
| Rebalanceo de particiones..... | 223 |
| Estrategias de rebalanceo..... | 224 |
| Operaciones: rebalanceo automático o manual..... | 227 |
| Enrutamiento de solicitudes..... | 228 |
| Ejecución de consultas en paralelo..... | 230 |
| Resumen..... | 231 |
| Referencias..... | 232 |

| | |
|---|------------|
| 7. Transacciones..... | 237 |
| El resbaladizo concepto de «transacción» | 238 |
| El significado de ACID..... | 239 |
| Operaciones con un solo objeto y con varios objetos..... | 244 |
| Niveles de aislamiento débil..... | 249 |
| Lectura confirmada..... | 250 |
| Aislamiento de instantáneas y lectura repetitiva | 253 |
| Cómo evitar que se pierdan las actualizaciones..... | 259 |
| Escritura desviada y fantasmas | 263 |
| Serializabilidad..... | 269 |
| Ejecución en serie | 270 |
| Bloqueo en dos fases (2PL)..... | 275 |
| Aislamiento de instantáneas serializable (SSI)..... | 279 |
| Resumen..... | 284 |
| Referencias..... | 287 |
| | |
| 8. El problema de los sistemas distribuidos..... | 293 |
| Fallos y averías parciales | 294 |
| Computación en la nube y supercomputación..... | 295 |
| Redes poco fiables..... | 298 |
| Fallos de red en la práctica..... | 299 |
| Detección de fallos..... | 300 |
| Tiempos de espera y retardos ilimitados | 301 |
| Redes síncronas frente a asíncronas | 305 |
| Relojes poco fiables..... | 308 |
| Relojes monotónicos frente a relojes en tiempo real | 309 |
| Sincronización y precisión del reloj | 310 |
| Confianza en los relojes sincronizados | 312 |
| Pausas del proceso | 317 |
| Conocimiento, verdades y mentiras..... | 322 |
| La verdad la define la mayoría | 323 |
| Fallos bizantinos..... | 326 |
| Modelos de sistemas y realidad..... | 329 |
| Resumen..... | 333 |
| Referencias..... | 335 |
| | |
| 9. Coherencia y consenso..... | 345 |
| Garantías de coherencia..... | 346 |
| Linealizabilidad..... | 348 |
| ¿Qué hace que un sistema sea linealizable? | 349 |
| Confianza en la linealizabilidad | 354 |

| | |
|--|-----|
| Implementación de sistemas linealizables | 357 |
| El coste de la linealizabilidad | 360 |
| Garantías del ordenamiento | 364 |
| Ordenamiento y causalidad | 364 |
| Ordenamiento por números de secuencia | 369 |
| Difusión de orden total | 373 |
| Transacciones distribuidas y consenso | 378 |
| Confirmación atómica y confirmación en dos fases (2PC) | 380 |
| Transacciones distribuidas en la práctica | 386 |
| Consenso tolerante a fallos | 391 |
| Servicios de afiliación y coordinación | 397 |
| Resumen | 401 |
| Referencias | 403 |

Parte III. Datos derivados

| | |
|---|------------|
| 10. Procesamiento por lotes | 417 |
| Procesamiento por lotes con herramientas Unix | 419 |
| Análisis de un <i>log</i> sencillo | 419 |
| La filosofía Unix | 422 |
| MapReduce y sistemas de archivos distribuidos | 426 |
| Ejecución de trabajos MapReduce | 428 |
| Agrupaciones y uniones de lados reducidos | 432 |
| Uniones del lado del mapa | 438 |
| Resultado de los flujos de trabajo por lotes | 441 |
| Comparación de Hadoop con las bases de datos distribuidas | 445 |
| Más allá de MapReduce | 450 |
| Materialización del estado intermedio | 450 |
| Grafos y procesamiento iterativo | 455 |
| API y lenguajes de alto nivel | 458 |
| Resumen | 461 |
| Referencias | 463 |
| | |
| 11. Procesamiento de flujos | 471 |
| Transmisión de flujos de eventos | 472 |
| Sistemas de mensajería | 473 |
| <i>Logs</i> particionados | 479 |
| Bases de datos y flujos | 484 |
| Necesidad de mantener los sistemas sincronizados | 485 |
| Captura de datos de cambios | 487 |

| | |
|---|------------|
| Aprovisionamiento de eventos | 490 |
| Estado, flujos e inmutabilidad..... | 493 |
| Procesamiento de flujos..... | 498 |
| Usos del procesamiento de flujos | 499 |
| Razonamiento sobre el tiempo..... | 503 |
| Uniones de flujos | 507 |
| Tolerancia a fallos | 511 |
| Resumen..... | 515 |
| Referencias..... | 517 |
| 12. El futuro de los sistemas de datos | 525 |
| Integración de datos | 525 |
| Combinación de herramientas especializadas mediante la derivación de datos | 526 |
| Procesamiento por lotes y procesamiento de flujos | 530 |
| Desagregación de bases de datos | 535 |
| Composición de las tecnologías de almacenamiento de datos..... | 536 |
| Diseño de aplicaciones en torno al flujo de datos | 540 |
| Observación del estado derivado..... | 546 |
| En busca de la corrección | 552 |
| Argumento de las bases de datos de extremo a extremo..... | 553 |
| Aplicación de restricciones..... | 558 |
| Puntualidad e integridad | 562 |
| Confie, pero verifique..... | 567 |
| Hacer lo correcto | 571 |
| Análisis predictivo | 572 |
| Privacidad y seguimiento..... | 576 |
| Resumen..... | 583 |
| Referencias..... | 585 |
| Glosario | 593 |