Distributed GraphLab: A Framework for Machine Learning and Data Mining in the Cloud

Yucheng Low  Carnegie Mellon University  ylow@cs.cmu.edu
Joseph Gonzalez  Carnegie Mellon University  jegonzal@cs.cmu.edu
Aapo Kyrola  Carnegie Mellon University  akyrola@cs.cmu.edu
Danny Bickson  Carnegie Mellon University  bickson@cs.cmu.edu
Carlos Guestrin  Carnegie Mellon University  guestrin@cs.cmu.edu
Joseph M. Hellerstein  UC Berkeley  hellerstein@cs.berkeley.edu

ABSTRACT

While high-level data parallel frameworks, like MapReduce, simplify the design and implementation of large-scale data processing systems, they do not naturally or efficiently support many important data mining and machine learning algorithms and can lead to inefficient learning systems. To help fill this critical void, we introduced the GraphLab abstraction which naturally expresses asynchronous, dynamic, graph-parallel computation while ensuring data consistency and achieving a high degree of parallel performance in the shared-memory setting. In this paper, we extend the GraphLab framework to the substantially more challenging distributed setting while preserving strong data consistency guarantees.

We develop graph based extensions to pipelined locking and data versioning to reduce network congestion and migrate the effect of network latency. We also introduce fault tolerance to the GraphLab abstraction using the classic Chandy-Lamport snapshot algorithm and demonstrate how it can be easily implemented by exploiting the GraphLab abstraction itself. Finally, we evaluate our distributed implementation of the GraphLab abstraction on a large Amazon EC2 deployment and show 1-2 orders of magnitude performance gains over Hadoop-based implementations.

1. INTRODUCTION

With the exponential growth in the scale of Machine Learning and Data Mining (MLDM) problems and increasing sophistication of MLDM techniques, there is an increasing need for systems that can execute MLDM algorithms efficiently in parallel on large clusters. Simultaneously, the availability of Cloud computing services like Amazon EC2 provide the promise of on-demand access to affordable large-scale computing and storage resources without substantial upfront investments. Unfortunately, designing, implementing, and debugging the distributed MLDM algorithms needed to fully utilize the Cloud can be prohibitively challenging requiring MLDM experts to address race conditions, deadlocks, distributed state, and communication protocols while simultaneously developing mathematically complex models and algorithms.

Nonetheless, the demand for large-scale computational and storage resources, has driven many high-level parallel abstractions (e.g. MapReduce [8, 9], Dryad [19] and Pregel [25]) to support new parallel and distributed MLDM systems targeted at individual models and applications. This time consuming and often redundant effort slows the progress of the field as different research groups repeatedly solve the same parallel/distributed computing problems. Therefore, the MLDM community needs a high-level distributed abstraction that specifically targets the asynchronous, dynamic, graph-parallel computation found in many MLDM applications while hiding the complexities of parallel/distributed system design. Unfortunately, existing high-level parallel abstractions (e.g. MapReduce [8, 9], Dryad [19] and Pregel [25]) fail to support these critical properties.

To help fill this void we introduced [24] GraphLab abstraction which directly targets asynchronous, dynamic, graph-parallel computation in the shared-memory setting.

In this paper we extend the multi-core GraphLab abstraction to the distributed setting and provide a formal description of the distributed execution model. We then explore several methods to implement an efficient distributed execution model while preserving strict consistency requirements. To achieve this goal we incorporate data versioning to reduce network congestion and pipelined distributed locking to mitigate the effects of network latency. To address the challenges of data locality and ingress we introduce the atom graph for rapidly placing graph structured data in the distributed setting. We also add fault tolerance to the GraphLab framework by adapting the classic Chandy-Lamport [6] snapshot algorithm and demonstrate how it can be easily implemented within the GraphLab abstraction.

We conduct a comprehensive performance analysis of our optimized C++ implementation on the Amazon Elastic Cloud (EC2) computing service. We show that applications created using GraphLab outperform equivalent Hadoop/MapReduce[9] implementations by 20-60x and match the performance of carefully constructed MPI implementations. Our main contributions are the following:

- A summary of common properties of MLDM algorithms and the limitations of existing large-scale frameworks. (Sec. [2])
- A modified version of the GraphLab abstraction and execution model tailored to the distributed setting. (Sec. [9])
- Two substantially different approaches to implementing the new distributed execution model (Sec. [4]).
2. MLDM ALGORITHM PROPERTIES

In this section, we describe several key properties of efficient large-scale parallel MLDM systems addressed by the GraphLab abstraction [24] and how other parallel frameworks fail to address these properties. A summary of these properties and parallel frameworks can be found in Table 1.

**Graph Structured Computation:** Many of the recent advances in MLDM have focused on modeling the dependencies between data. By modeling data dependencies, we are able to extract more signal from noisy data. For example, modeling the dependencies between similar shoppers allows us to make better product recommendations than treating shoppers in isolation. Unfortunately, data parallel abstractions like MapReduce [9] are not generally well suited for the dependent computation typically required by more advanced MLDM algorithms. Although it is often possible to map algorithms with computational dependencies into the MapReduce abstraction, the resulting transformations can be challenging and may introduce substantial inefficiency.

As a consequence, there has been a recent trend toward graph-parallel abstractions like Pregel [25] and GraphLab [24] which naturally express computational dependencies. These abstractions adopt a vertex-centric model in which computation is defined as kernels that run on each vertex. For instance, Pregel is a bulk synchronous message passing abstraction where vertices communicate through messages. On the other hand, GraphLab is a sequential shared memory abstraction where each vertex can read and write to data on adjacent vertices and edges. The GraphLab runtime is then responsible for ensuring a consistent parallel execution. Consequently, GraphLab simplifies the design and implementation of graph-parallel algorithms by freeing the user to focus on sequential computation rather than the parallel movement of data (i.e., messaging).

**Asynchronous Iterative Computation:** Many important MLDM algorithms iteratively update a large set of parameters. Because of the underlying graph structure, parameter updates (on vertices or edges) depend (through the graph adjacency structure) on the values of other parameters. In contrast to synchronous systems, which update all parameters simultaneously (in parallel) using parameter values from the previous time step as input, asynchronous systems update parameters using the most recent parameter values as input. As a consequence, asynchronous systems provides many MLDM algorithms with significant algorithmic benefits. For example, linear systems (common to many MLDM algorithms) have been shown to converge faster when solved asynchronously. Additionally, there are numerous other cases (e.g., belief propagation [13], expectation maximization [28], and stochastic optimization [35, 34]) where asynchronous procedures have been empirically shown to significantly outperform synchronous procedures. In Fig. 1(a), we demonstrate how asynchronous computation can substantially accelerate the convergence of PageRank.

Synchronous computation incurs costly performance penalties since the runtime of each phase is determined by the slowest machine. The poor performance of the slowest machine may be caused by a multitude of factors including: load and network imbalances, hardware variability, and multi-tenancy (a principal concern in the Cloud). Even in typical cluster settings, each compute node may also provide other services (e.g., distributed file systems). Imbalances in the utilization of these other services will result in substantial performance penalties if synchronous computation is used.

In addition, variability in the complexity and convergence of the individual vertex kernels can produce additional variability in execution time, even when the graph is uniformly partitioned. For example, natural graphs encountered in real-world applications have power-law degree distributions which can lead to highly skewed running times even with a random partition [36]. Furthermore, the actual work required for each vertex could depend on the data in a problem-specific manner (e.g., local rate of convergence).

While abstractions based on bulk data processing, such as MapReduce [9] and Dryad [19] were not designed for iterative computation, recent projects such as Spark [38] extend MapReduce and other data parallel abstractions to the iterative setting. However, these abstractions still do not support asynchronous computation. Bulk Synchronous Parallel (BSP) abstractions such as Pregel [25], Piccolo [33], and BPGL [16] do not naturally express asynchrony. On the other hand, the shared memory GraphLab abstraction was designed to efficiently and naturally express the asynchronous iterative algorithms common to advanced MLDM.

**Dynamic Computation:** In many MLDM algorithms, iterative computation converges asymmetrically. For example, in parameter optimization, often a large number of parameters will quickly converge in a few iterations, while the remaining parameters will converge slowly over many iterations [11, 10]. In Fig. 1(b), we plot the distribution of updates required to reach convergence for PageRank. Surprisingly, the majority of the vertices required only a single update while only about 3% of the vertices required more than 10 updates. Additionally, prioritizing computation can further accelerate convergence as demonstrated by Zhang et al. [39] for a variety of graph algorithms including PageRank. If we update all parameters equally often, we waste time recomputing parameters that have effectively converged. Conversely, by focusing early computation on more challenging parameters, we can potentially accelerate convergence. In Fig. 1(c) we empirically demonstrate how dynamic scheduling can accelerate convergence of loopy belief propagation (a popular MLDM algorithm).

Several recent abstractions have incorporated forms of dynamic computation. For example, Pregel [25] supports a limited form of dynamic computation by allowing some vertices to skip computation on each super-step. Other abstractions like Pearce et al. [12] and GraphLab allow the user to adaptively prioritize computation. While both Pregel and GraphLab support dynamic computation, only GraphLab permits prioritization as well as the ability to adaptively pull information from adjacent vertices (see Sec. 5.2 for more details). In this paper we relax some of the original GraphLab scheduling requirements described in [24] to enable efficient distributed FIFO and priority scheduling.

**Serializability:** By ensuring that all parallel executions have an equivalent sequential execution, serializability eliminates many challenges associated with designing, implementing, and testing parallel MLDM algorithms. In addition, many algorithms converge faster if serializability is ensured, and some even require serializability for correctness. For instance, Dynamic ALS (Sec. 5.1) is unstable when allowed to race (Fig. 1(d)). Gibbs sampling, a very popular MLDM algorithm, requires serializability for statistical correctness.
Table 1: Comparison chart of large-scale computation frameworks. (a) [38] describes and iterative extension of MapReduce. (b) [18] proposes an iterative extension for Dryad. (c) Piccolo does not provide a mechanism to ensure consistency but instead exposes a mechanism for the user to attempt to recover from simultaneous writes.

![GraphLab](image)

Figure 1: (a) Asy vs Sync PageRank
(b) Dynamic PageRank
(c) LoopyBP Conv.
(d) ALS Consistency

3. DISTRIBUTED GRAPHLAB ABSTRACTION

The GraphLab abstraction stores the program state as a directed graph called the data graph. The data graph \( G = (V, E, D) \) is a container that manages the user-defined data \( D \). Here we use the term data broadly to refer to model parameters, algorithm state, and even statistical data. Users can associate arbitrary data with each vertex \( D_v : v \in V \) and each edge \( D_u \rightarrow v : \{u, v\} \in E \) in the graph. However, as the GraphLab abstraction is not dependent on edge directions, we also use \( D_{u \rightarrow v} \) to denote the data on both edge directions \( u \rightarrow v \) and \( v \rightarrow u \). Finally, while the graph data is mutable, the structure is static and cannot be changed during execution.

### 3.1 Data Graph

The GraphLab abstraction defines the rank of a webpage as follows:

\[
R(v) = \frac{\alpha}{n} + (1 - \alpha) \sum_{u \text{ links to } v} w_{u,v} \times R(u)
\]

in terms of the weighted \( w_{u,v} \), ranks \( R(u) \) of the pages \( u \) that link to \( v \) as well as some probability \( \alpha \) of randomly jumping to that page. The PageRank algorithm iterates Eq. (1) until the PageRank changes by less than some small value \( \epsilon \).

### 3.2 Update Functions

Computation is encoded in the GraphLab abstraction in the form of update functions. An update function is a stateless procedure that modifies the data within the scope of a vertex and schedules the future execution of update functions on other vertices. The scope of vertex \( v \) (denoted by \( S_v \)) is the data stored in \( v \), as well as the data stored in all adjacent vertices and adjacent edges (Fig. 2(a)).

A GraphLab update function takes as an input a vertex \( v \) and its scope \( S_v \) and returns the new versions of the data in the scope as well as set vertices \( T \):

\[
\text{Update} : (v, S_v) \rightarrow (S_v, T)
\]

After executing an update function the modified data in \( S_v \) is written back to the data graph. The set of vertices \( u \in T \) are eventually
executed by applying the update function \( f(u, S_v) \) following the execution semantics described later in Sec. 3.3.

Rather than adopting a message passing or data flow model as in [25] [19], GraphLab allows the user defined update functions complete freedom to read and modify any of the data on adjacent vertices and edges. This simplifies user code and eliminates the need for the users to reason about the movement of data. By controlling what vertices are returned in \( T \) and thus to be executed, GraphLab update functions can efficiently express adaptive computation. For example, an update function may choose to return (schedule) its neighbors only when it has made a substantial change to its local data.

There is an important difference between Pregel and GraphLab in how dynamic computation is expressed. GraphLab decouples the scheduling of future computation from the movement of data. As a consequence, GraphLab update functions have access to data on adjacent vertices even if the adjacent vertices did not schedule the current update. Conversely, Pregel update functions are initiated by messages and can only access the data in the message, limiting what can be expressed. For instance, dynamic PageRank is difficult to express in Pregel since the PageRank computation for a given page requires the PageRank values of all adjacent pages even if some of the adjacent pages have not recently changed. Therefore, the decision to send data (PageRank values) to neighboring vertices cannot be made by the sending vertex (as required by Pregel) but instead must be made by the receiving vertex. GraphLab, naturally expresses the pull model, since adjacent vertices are only responsible for scheduling, and update functions can directly read adjacent vertex values even if they have not changed.

**Example 3** (\textsc{PageRank: Ex. 1}). The update function for \textsc{PageRank} (defined in Alg. 1) computes a weighted sum of the current ranks of neighboring vertices and assigns it as the rank of the current vertex. The algorithm is adaptive: neighbors are scheduled for update only if the value of current vertex changes by more than a predefined threshold.

**Algorithm 1: PageRank update function**

\[
\text{Input: Vertex data } R(v) \text{ from } S_v \\
\text{Input: Edge data } \{u_{uv}: u \in N[v]\} \text{ from } S_v \\
\text{Input: Neighbor vertex data } \{R(u) : u \in N[v]\} \text{ from } S_v \\
R(v) \leftarrow R(v) \leftarrow R(v) + \alpha/n \\
\forall u \in N[v] \text{ do} \quad \text{// Loop over neighbors} \\
\quad \begin{cases} 
R(v) \leftarrow R(v) + (1 - \alpha) \ast u_{uv} \ast R(u) \\
\end{cases} \\
\text{// If the PageRank changes sufficiently} \\
\text{if } |R(v) - R_{old}(v)| > \epsilon \text{ then} \\
\quad \begin{cases} 
\text{// Schedule neighbors to be updated} \\
\quad \forall u : u \in N[v]\} \\
\end{cases} \\
\text{Output: Modified scope } S_v \text{ with new } R(v)
\]

**3.3 The GraphLab Execution Model**

The GraphLab execution model, presented in Alg. 2 follows a simple single loop semantics. The input to the GraphLab abstraction consists of the data graph \( G = (V, E, D) \), an update function, an initial set of vertices \( T \) to be executed. While there are vertices remaining in \( T \), the algorithm removes (Line 1) and executes (Line 2) vertices, adding any new vertices back into \( T \) (Line 3). Duplicate vertices are ignored. The resulting data graph and global values are returned to the user on completion.

To enable a more efficient distributed execution, we relax the execution ordering requirements of the shared-memory GraphLab abstraction and allow the GraphLab run-time to determine the best order to execute vertices. For example, \textsc{RemoveNext}(T) (Line 1) may choose to return vertices in an order that minimizes network communication or latency (see Sec. 4.2.2). The only requirement imposed by the GraphLab abstraction is that all vertices in \( T \) are eventually executed. Because many MLDM applications benefit from prioritization, the GraphLab abstraction allows users to assign priorities to the vertices in \( T \). The GraphLab run-time may use these priorities in conjunction with system level objectives to optimize the order in which the vertices are executed.

**Algorithm 2: GraphLab Execution Model**

\[
\text{Input: Data Graph } G = (V, E, D) \\
\text{Input: Initial vertex set } T = \{v_1, v_2, \ldots\} \\
\text{while } T \text{ is not Empty do} \\
\quad v \leftarrow \text{RemoveNext}(T) \\
\quad \forall (T', S_v) \quad f(v, S_v) \\
\quad T \leftarrow T \cup T' \\
\text{Output: Modified Data Graph } G = (V, E, D')
\]

**3.4 Ensuring Serializability**

The GraphLab abstraction presents a rich sequential model which is automatically translated into a parallel execution by allowing multiple processors to execute the same loop on the same graph, removing and executing different vertices simultaneously. To retain the sequential execution semantics we must ensure that overlapping computation is not run simultaneously. We introduce several consistency models that allow the runtime to optimize the parallel execution while maintaining serializability.

The GraphLab run-time ensures a serializable execution. A serializable execution implies that there exists a corresponding serial schedule of update functions that when executed by Alg. 2 produces the same values in the data-graph. By ensuring serializability, GraphLab simplifies reasoning about highly-asynchronous dynamic computation in the distributed setting.

A simple method to achieve serializability is to ensure that the scopes of concurrently executing update functions do not overlap. In [24] we call this the full consistency model (see Fig. 2(b)). However, full consistency limits the potential parallelism since concurrently executing update functions must be at least two vertices apart (see Fig. 2(c)). However, for many machine learning algorithms, the update functions do not need full read/write access to all of the data within the scope. For instance, the PageRank update in Eq. 1 only requires read access to edges and neighboring vertices. To provide greater parallelism while retaining serializability, GraphLab defines the edge consistency model. The edge consistency model ensures each update function has exclusive read-write access to its vertex and adjacent edges but read only access to adjacent vertices (Fig. 2(b)). As a consequence, the edge consistency model increases parallelism by allowing update functions with slightly overlapping scopes to safely run in parallel (see Fig. 2(c)). Finally, the vertex consistency model allows all update functions to be run in parallel, providing maximum parallelism.

**3.5 Sync Operation and Global Values**

In many MLDM algorithms it is necessary to maintain global statistics describing data stored in the data graph. For example, many statistical inference algorithms require tracking global convergence estimators. To address this need, the GraphLab abstraction defines global values that may be read by update functions, but are written using sync operations. Similar to aggregates in Pregel, the sync operation is an associative commutative sum:

\[
Z = \text{Finalize} \left( \bigoplus_{v \in V} \text{Map}(S_v) \right)
\]
defined over all the scopes in the graph. Unlike Pregel, the sync operation introduces a finalization phase, `Finalize()`, to support tasks, like normalization, which are common in MLDM algorithms. Also in contrast to Pregel, where the aggregation runs after each super-step, the sync operation in GraphLab runs continuously in the background to maintain updated estimates of the global value.

Since every update function can access global values, ensuring serializability of the Sync operation with respect to update functions is costly and will generally require synchronizing and halting all computation. Just as GraphLab has multiple consistency levels for update functions, we similarly provide the option of consistent or inconsistent Sync computations.

## 4. DISTRIBUTED GRAPHLAB DESIGN

In this section we extend the shared memory system design of the GraphLab abstraction to the substantially more challenging distributed setting and discuss the techniques required to achieve this goal. An overview of the distributed design is illustrated in Fig. 2. Because of the inherently random memory access patterns common to dynamic asynchronous graph algorithms, we focus on the distributed in-memory setting, requiring the entire graph and all program state to reside in RAM. Our distributed implementation is written in C++ and extends the original open-sourced shared memory GraphLab implementation.

### 4.1 The Distributed Data Graph

Efficiently implementing the data graph in the distributed setting requires balancing computation, communication, and storage. Therefore, we need to construct balanced partitioning of the data graph that minimizes number of edges that cross between machines. Because the Cloud setting enables the size of the cluster to vary with budget and performance demands, we must be able to quickly load the data graph on varying sized Cloud deployments. To resolve these challenges, we developed a graph representation based on two-phased partitioning which can be efficiently load balanced on arbitrary cluster sizes.

The data graph is initially over-partitioned using domain specific knowledge (e.g., planar embedding), or by using a distributed graph partitioning heuristic (e.g., ParMetis [21], Random Hashing) into k parts where k is much greater than the number of machines. Each part, called an atom, is stored as a separate file on a distributed storage system (e.g., HDFS, Amazon S3). Each atom file is a simple binary compressed journal of graph generating commands such as `AddVertex(5000, vdata)` and `AddEdge(42 → 314, edata)`. In addition, each atom stores information regarding ghosts: the set of vertices and edges adjacent to the partition boundary. The connectivity structure and file locations of the k atoms is stored in an atom index file as a meta-graph with k vertices (corresponding to the atoms) and edges encoding the connectivity of the atoms.

Distributed loading is accomplished by performing a fast balanced partition of the meta-graph over the number of physical machines. Each machine then constructs its local portion of the graph by playing back the journal from each of its assigned atoms. The playback procedure also instantiates the ghost of the local partition in memory. The ghosts are used as caches for their true counterparts across the network. Cache coherence is managed using a simple versioning system, eliminating the transmission of unchanged or constant data (e.g., edge weights).

The two-stage partitioning technique allows the same graph partition computation to be reused for different numbers of machines without requiring a full repartitioning step. A study on the quality of the two stage partitioning scheme is beyond the scope of this paper, though simple experiments using graphs obtained from [23] suggest that the performance is comparable to direct partitioning.

### 4.2 Distributed GraphLab Engines

The Distributed GraphLab engine emulates the execution model defined in Sec. 3.3 and is responsible for executing update functions and sync operations, maintaining the set of scheduled vertices T, and ensuring serializability with respect to the appropriate consistency model (see Sec. 3.4). As discussed in Sec. 3.3 the precise order in which vertices are removed from T is up to the implementation and can affect performance and expressiveness. To evaluate this trade-off we built the low-overhead Chromatic Engine, which executes T partially asynchronously, and the more expressive Locking Engine which is fully asynchronous and supports vertex priorities.

#### 4.2.1 Chromatic Engine

A classic technique to achieve a serializable parallel execution of a set of dependent tasks (represented as vertices in a graph) is to construct a vertex coloring that assigns a color to each vertex
such that no adjacent vertices share the same color. Given a vertex coloring of the data graph, we can satisfy the edge consistency model by executing, synchronously, all vertices of the same color in the vertex set $T$ before proceeding to the next color. We use the term color-step, in analogy to the super-step in the BSP model, to describe the process of updating all the vertices within a single color and communicating all changes. The sync operation can then be run safely between color-steps.

We can satisfy the other consistency models simply by changing how the vertices are colored. The full consistency model is satisfied by constructing a second-order vertex coloring (i.e., no vertex shares the same color as any of its distance two neighbors). The vertex consistency model is satisfied by assigning all vertices the same color. While optimal graph coloring is NP-hard in general, a reasonable quality coloring can be constructed quickly using graph coloring heuristics (e.g., greedy coloring). Furthermore, many MLDM problems produce graphs with trivial colorings. For example, many optimization problems in MLDM are naturally expressed as bipartite (two-colorable) graphs, while problems based upon template models can be easily colored using the template.

While the chromatic engine operates in synchronous color-steps, changes to ghost vertices and edges are communicated asynchronously as they are made. Consequently, the chromatic engine efficiently uses both network bandwidth and processor time within each color-step. However, we must ensure that all modifications are communicated before moving to the next color and therefore we require a full communication barrier between color-steps.

### 4.2.2 Distributed Locking Engine

While the chromatic engine satisfies the distributed GraphLab abstraction defined in Sec. it does not provide sufficient scheduling flexibility for many interesting applications. In addition, it presupposes the availability of a graph coloring, which may not always be readily available. To overcome these limitations, we introduce the distributed locking engine which extends the mutual exclusion technique used in the shared memory engine.

We achieve distributed mutual exclusion by associating a readers-writer lock with each vertex. The different consistency models can then be implemented using different locking protocols. Vertex consistency is achieved by acquiring a write-lock on the central vertex of each requested scope. Edge consistency is achieved by acquiring a write lock on the central vertex, and read locks on adjacent vertices. Finally, full consistency is achieved by acquiring write locks on the central vertex and all adjacent vertices. Deadlocks are avoided by acquiring locks sequentially following a canonical order. We use the ordering induced by machine ID followed by vertex ID ($\text{owner}(v), v$) since this allows all locks on a remote machine to be requested in a single message.

Since the graph is partitioned, we restrict each machine to only run updates on local vertices. The ghost vertices/edges ensure that the update have direct memory access to all information in the scope. Each worker thread on each machine evaluates the loop described in Alg. until the scheduler is empty. Termination is evaluated using write locks on the central vertex and all adjacent vertices. Deadlocks (vertex ID) are satisfied by assigning all vertices the same color.

Pipelined Locking and Prefetching: Each machine maintains a pipeline of vertices for which locks have been requested, but have not been fulfilled. Vertices that complete lock acquisition and data synchronization leave the pipeline and are executed by worker threads. The local scheduler ensures that the pipeline is always filled to capacity. An overview of the pipelined locking engine loop is shown in Alg.

To implement the pipelining system, regular readers-writer locks cannot be used since they would halt the pipeline thread on contention. We therefore implemented a non-blocking variation of the readers-writer lock that operates through callbacks. Lock acquisition requests provide a pointer to a callback, that is called once the request is fulfilled. These callbacks are chained into a distributed continuation passing scheme that passes lock requests across machines in sequence. Since lock acquisition follows the total ordering described earlier, deadlock free operation is guaranteed. To further reduce latency, synchronization of locked data is performed immediately as each machine completes its local locks.

Example 4. To acquire a distributed edge consistent scope on a vertex $v$ owned by machine 2 with ghosts on machines 1 and 5, the system first sends a message to machine 1 to acquire a local edge consistent scope on machine 1 (write-lock on $v$, read-lock on neighbors). Once the locks are acquired, the message is passed on to machine 2 to again acquire a local edge consistent scope. Finally, the message is sent to machine 5 before returning to the owning machine to signal completion.

To evaluate the performance of the distributed pipelining system, we constructed a three-dimensional mesh of $300 \times 300 \times 300 = 27,000,000$ vertices. Each vertex is 26-connected (to immediately adjacent vertices along the axis directions, as well as all diagonals), producing over 375 million edges. The graph is partitioned using Metis into 512 atoms. We interpret the graph as a binary Markov Random Field and evaluate the runtime of 10 iterations of loopy Belief Propagation varying the length of the pipeline from 4 to 100, and the number of EC2 cluster compute instance (ccl.4xlarge) from 4 machines (32 processors) to 16 machines (128 processors). We observe in Fig. that the distributed locking system provides strong, nearly linear, scalability. In Fig. we evaluate the efficacy of the pipelining system by increasing the pipeline length. We find that increasing the length from 100 to 1000 leads to a factor of three reduction in runtime.
Figure 3: (a) Plots the runtime of the Distributed Locking Engine on a synthetic loopy belief propagation problem varying the number of machines with pipeline length = 10,000. (b) Plots the runtime of the Distributed Locking Engine on the same synthetic problem on 16 machines (128 CPUs), varying the pipeline length. Increasing pipeline length improves performance with diminishing returns.

Algorithm 5: Snapshot Update on vertex $v$

\[
\text{if } v \text{ was already snapshotted then} \\
\quad \text{Quit} \\
\text{Save } D_v \text{ // Save current vertex} \\
\text{foreach } u \in N[v] \text{ do} \\
\quad \text{if } u \text{ was not snapshotted then} \\
\quad \quad \text{Save data on edge } D_{uv} \\
\quad \quad \text{Schedule } u \text{ for a Snapshot Update} \\
\text{Mark } v \text{ as snapshotted}
\]

4.3 Fault Tolerance

We introduce fault tolerance to the distributed GraphLab framework using a distributed checkpoint mechanism. In the event of a failure, the system is recovered from the last checkpoint. We evaluate two strategies to construct distributed snapshots: a synchronous method that suspends all computation while the snapshot is constructed, and an asynchronous method that incrementally constructs a snapshot without suspending execution.

Synchronous snapshots are constructed by suspending execution of update functions, flushing all communication channels, and then saving all modified data since the last snapshot. Changes are written to journal files in a distributed file-system and can be used to restart the execution at any previous snapshot.

Unfortunately, synchronous snapshots expose the GraphLab engine to the same inefficiencies of synchronous computation (Sec. 2) that GraphLab is trying to address. Therefore we designed a fully asynchronous alternative based on the Chandy–Lamport [6] snapshot. Using the GraphLab abstraction we designed and implemented a variant of the Chandy-Lamport snapshot specifically tailored to the GraphLab data-graph and execution model. The resulting algorithm (Alg. 5) is expressed as an update function and guarantees a consistent snapshot under the following conditions:

- Edge Consistency is used on all update functions,
- Schedule completes before the scope is unlocked,
- the Snapshot Update is prioritized over other update functions,

which are satisfied with minimal changes to the GraphLab engine. The proof of correctness follows naturally from the original proof in [6] with the machines and channels replaced by vertices and edges and messages corresponding to scope modifications.

Both the synchronous and asynchronous snapshots are initiated at fixed intervals. The choice of interval must balance the cost of constructing the checkpoint with the computation lost since the last checkpoint in the event of a failure. Young et al. [37] derived a first-order approximation to the optimal checkpoint interval:

\[ T_{\text{Interval}} = \sqrt{2 T_{\text{checkpoint}} T_{\text{MTBF}}} \]

where $T_{\text{checkpoint}}$ is the time it takes to complete the checkpoint and $T_{\text{MTBF}}$ is the mean time between failures for the cluster. For instance, using a cluster of 64 machines, a per machine MTBF of 1 year, and a checkpoint time of 2 min leads to optimal checkpoint intervals of 3 hrs. Therefore, for the deployments considered in our experiments, even taking pessimistic assumptions for $T_{\text{MTBF}}$, leads to checkpoint intervals that far exceed the runtime of our experiments and in fact also exceed the Hadoop experiment runtimes. This brings into question the emphasis on strong fault tolerance in Hadoop. Better performance can be obtained by balancing fault tolerance costs against that of a job restart.

Evaluation: We evaluate the performance of the snapshotting algorithms on the same synthetic mesh problem described in the previous section, running on 16 machines (128 processors). We configure the implementation to issue exactly one snapshot in the middle of the second iteration. In Fig. 4(a), we plot the number of updates completed against time elapsed. The effect of the synchronous snapshot and the asynchronous snapshot can be clearly observed: synchronous snapshots stops execution, while the asynchronous snapshot only slows down execution.

The benefits of asynchronous snapshots become more apparent in the multi-tenancy setting where variation in system performance exacerbate the cost of synchronous operations. We simulate this on Amazon EC2 by halting one of the processes for 15 seconds after processing a section of the graph to be appended without reprocessing all the data.
Each experiment was based on large real-world problems (for instance, the movie “Harry Potter” connects to a very large number of users). Finally, many Pregel implementations of MLDM do not scale to even the smaller problems we considered. While ALS may not seem like a graph algorithm, it can be represented elegantly using the GraphLab abstraction. The sparse matrix...
Alltoall to the machines that need them in the next super-step.

This results in the generation of a large amount of unnecessary communication as there are ghosts of the vertex. In particular, only machines that require the vertex data for computation will receive it, and each machine receives each modified vertex data at most once, even if the vertex has many neighbors.

Our MPI implementation of ALS is highly optimized, and uses synchronous MPI collective operations for communication. The computation is broken into super-steps that alternate between recomputing the latent user and movies low rank matrices. Between super-steps the new user and movie values are scattered (using MPI_Alltoall) to the machines that need them in the next super-step. As a consequence our MPI implementation of ALS is roughly equivalent to an optimized Pregel version of ALS with added support for parallel broadcasts. Surprisingly, GraphLab was able to outperform the MPI implementation. We attribute the performance to the use of background asynchronous communication in GraphLab.

Finally, we can evaluate the effect of enabling dynamic computation. In Fig. 9(a) we plot the test error obtained over time using a dynamic update schedule as compared to a static BSP-style update schedule. This dynamic schedule is easily represented in GraphLab while it is difficult to express using Pregel messaging semantics. We observe that a dynamic schedule converges much faster, reaching a low test error in about half amount of work.

5.2 Video Co-segmentation (CoSeg)

Video co-segmentation automatically identifies and clusters spatio-temporal segments of video (Fig. 7(a)) that share similar texture and color characteristics. The resulting segmentation can be used in scene understanding and other computer vision and robotics applications. Previous co-segmentation methods have focused on processing frames in isolation. Instead, we developed a joint co-segmentation algorithm that processes all frames simultaneously and is able to model temporal stability.

We preprocessed 1,740 frames of high-resolution video by coarsening each frame to a regular grid of $120 \times 50$ rectangular super-pixels. Each super-pixel stores the color and texture statistics for all the raw pixels in its domain. The CoSeg algorithm predicts the best label (e.g., sky, building, grass, pavement, trees) for each super-pixel using Gaussian Mixture Model (GMM) in conjunction with Loopy Belief Propagation (LBP). The GMM estimates the best label given the color and texture statistics in the super-pixel. The algorithm operates by connecting neighboring pixels in time and space into a large three-dimensional grid and uses LBP to smooth the local estimates. We combined the two algorithms to form an Expectation-Maximization algorithm, alternating between LBP to compute the label for each super-pixel given the GMM and then updating the GMM given the labels from LBP.
The GraphLab update function executes the LBP local iterative update. We implement the state-of-the-art adaptive update schedule described by [11], where updates that are expected to change vertex values significantly are prioritized. We therefore make use of the locking engine with an approximate priority scheduler. The parameters for the GMM are maintained using the sync operation.

To the best of our knowledge, there are no other abstractions that provide the dynamic asynchronous scheduling as well as the sync (reduction) capabilities required by this application.

In Fig. 8(a) we demonstrate that the locking engine can achieve scalability and performance on the large 10.5 million vertex graph used by this application, resulting in a 10x speedup with 16x more machines. We also observe from Fig. 8(a) that the locking engine provides nearly optimal weak scaling: the runtime does not increase significantly as the size of the graph increases proportionately with the number of machines. We can attribute this to the properties of the graph partition where the number of edges crossing machines increases linearly with the number of machines, resulting in low communication volume.

While Sec. 4.2.2 contains a limited evaluation of the pipelining system on a synthetic graph, here we further investigate the behavior of the distributed lock implementation when run on a complete problem that makes use of all key aspects of GraphLab: both sync and dynamic prioritized scheduling. The evaluation is performed on a small 32-frame (192K vertices) problem using a 4 node cluster and two different partitioning. An optimal partition was constructed by evenly distributing 8 frame blocks to each machine. A worst case partition was constructed by striping frames across machines and consequently stressing the distributed lock implementation by forcing each scope acquisition is to grab at least one remote lock. We also vary the maximum length of the pipeline. Results are plotted in Fig. 8(b). We demonstrate that increasing the length of the pipeline increases performance significantly and is able to compensate for poor partitioning, rapidly bringing down the runtime of the problem.

Just as in Sec. 4.2.2, we observe diminishing returns with increasing pipeline length. While pipelining violates the priority order, rapid convergence is still achieved.

We conclude that for the video co-segmentation task, Distributed GraphLab provides excellent performance while being the only distributed graph abstraction that allows the use of dynamic prioritized scheduling. In addition, the pipelining system is an effective way to hide latency, and to some extent, a poor partitioning.

### 5.3 Named Entity Recognition (NER)

Named Entity Recognition (NER) is the task of determining the type (e.g., Person, Place, or Thing) of a noun-phrase (e.g., Obama, Chicago, or Car) from its context (e.g., “President”, “lives near”, or “bought a”). NER is used in many natural language processing applications as well as information retrieval. In this application we obtained a large crawl of the web from the NELL project [3], and we counted the number of occurrences of each noun-phrase in each context. Starting with a small seed set of pre-labeled noun-phrases, the CoEM algorithm labels the remaining noun-phrases and contexts (see Table 7(b)) by alternating between estimating the best assignment to each noun-phrase given the types of its contexts and estimating the type of each context given the types of its noun-phrases.

The data graph for the NER problem is bipartite with one set of vertices corresponding to noun-phrases and other corresponding to each contexts. There is an edge between a noun-phrase and a context if the noun-phrase occurs in the context. The vertices store the estimated distribution over types and the edges store the number of times the noun-phrase appears in a context. Since the graph is two colorable and relatively dense the chromatic engine was used with random partitioning. The lightweight floating point arithmetic in the NER computation in conjunction with the relatively dense graph structure and random partitioning is essentially the worst-case for the current Distributed GraphLab design, and thus allow us to evaluate the overhead of the Distributed GraphLab runtime.

From Fig. 6(a) we see that NER achieved only a modest 3x improvement using 16x more machines. We attribute the poor scaling performance of NER to the large vertex data size (816 bytes), dense connectivity, and poor partitioning (random cut) that resulted in substantial communication overhead per iteration. Fig. 8(b) shows for each application, the average number of bytes per second transmitted by each machine with varying size deployments. Beyond 16 machines, NER saturates with each machine sending at a rate of over 100MB per second.

We evaluated our Distributed GraphLab implementation against a Hadoop and an MPI implementation in Fig. 8(c). In addition to the optimizations listed in Sec. 5.1, our Hadoop implementation required the use of binary marshaling methods to obtain reasonable performance (decreasing runtime by 5x from baseline).

We demonstrate that GraphLab implementation of NER was able to a 20-30x speedup over Hadoop. The reason for the performance gap is the same as that for the Netflix evaluation. Since each vertex emits a copy of itself for each edge: in the extremely large CoEM graph, this corresponds to over 100 GB of HDFS writes occurring between the Map and Reduce stage.

On the other hand, our MPI implementation was able to outperform Distributed GraphLab by a healthy margin. The CoEM task requires extremely little computation in comparison to the amount of data it touches. We were able to evaluate that the NER update function requires 5.7x fewer cycles per byte of data accessed as compared to the Netflix problem at d = 5 (the hardest Netflix case evaluated). The extremely poor computation to communication ratio stresses our communication implementation, that is outperformed by MPI’s efficient communication layer. Furthermore, Fig. 6(b) provides further evidence that we fail to fully saturate the network (that offers 10Gbps). Further optimizations to eliminate inefficiencies in GraphLab’s communication layer should bring us up to parity with the MPI implementation.

We conclude that while Distributed GraphLab is suitable for the NER task providing an effective abstraction, further optimizations are needed to improve scalability and to bring performance closer to that of a dedicated MPI implementation.

### 5.4 EC2 Cost evaluation

To illustrate the monetary cost of using the alternative abstractions, we plot the price-runtime curve for the Netflix application in Fig. 8(b) in log-log scale. All costs are computed using fine-grained billing rather than the hourly billing used by Amazon EC2. The price-runtime curve demonstrates diminishing returns: the cost of attaining reduced runtimes increases faster than linearly. As a comparison, we provide the price-runtime curve for Hadoop on the same application. For the Netflix application, GraphLab is about two orders of magnitude more cost-effective than Hadoop.
### Table 2: Experiment input sizes. The vertex and edge data are measured in bytes and the $d$ in Netflix is the size of the latent dimension.

<table>
<thead>
<tr>
<th>Exp.</th>
<th>#Verts</th>
<th>#Edges</th>
<th>Vertex Data</th>
<th>Edge Data</th>
<th>Update Complexity</th>
<th>Shape</th>
<th>Partition</th>
<th>Engine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Netflix</td>
<td>0.5M</td>
<td>99M</td>
<td>$8d + 13$</td>
<td>16</td>
<td>$O(d^4 + deg.)$</td>
<td>bipartite</td>
<td>random</td>
<td>Chromatic</td>
</tr>
<tr>
<td>CoSeg</td>
<td>10.5M</td>
<td>31M</td>
<td>302</td>
<td>80</td>
<td>$O(deg.)$</td>
<td>3D grid</td>
<td>frames</td>
<td>Locking</td>
</tr>
<tr>
<td>NER</td>
<td>2M</td>
<td>200M</td>
<td>816</td>
<td>4</td>
<td>$O(deg.)$</td>
<td>bipartite</td>
<td>random</td>
<td>Chromatic</td>
</tr>
</tbody>
</table>

![Figure 8](image1.png)  
(a) CoSeg Weak Scaling  
(b) Pipelined Locking  
(c) NER Comparisons  
(d) Snapshot Overhead

Figure 8: (a) Runtime of the CoSeg experiment as data set size is scaled proportionately with the number of machines. Ideally, runtime is constant. GraphLab experiences an 11% increase in runtime scaling from 16 to 64 machines. (b) The performance effects of varying the length of the pipeline. Increasing the pipeline length has a small effect on performance when partitioning is good. When partitioning is poor, increasing the pipeline length improves performance to be comparable to that of optimal partitioning. Runtime for worst-case partitioning at pipeline length 0 is omitted due to excessive runtimes. (c) Runtime of the NER experiment with Distributed GraphLab, Hadoop and MPI implementations. Note the logarithmic scale. GraphLab outperforms Hadoop by about 80x when the number of machines is small, and about 30x when the number of machines is large. The performance of Distributed GraphLab is comparable to the MPI implementation. (d) For each application, the overhead of performing a complete snapshot of the graph every $|V|$ updates (where $|V|$ is the number of vertices in the graph), when running on a 64 machine cluster.

![Figure 9](image2.png)  
(a) Dynamic Netflix  
(b) EC2 Price/Performance

Figure 9: (a) Convergence rate when dynamic computation is used. Dynamic computation can converge to equivalent test error in about half the number of updates. (b) Price Performance ratio of GraphLab and Hadoop on Amazon EC2 HPC machine on a log-log scale. Costs assume fine-grained billing.

### 6. RELATED WORK

Section 2 provides a detailed comparison of several contemporary high-level parallel and distributed frameworks. In this section we review related work in classic parallel abstractions, graph databases, and domain specific languages.

There has been substantial work [1] in graph structured databases dating back to the 1980’s along with many recent open-source and commercial products (e.g., Neo4J [29]). Graph databases typically focus on efficient storage and retrieval of graph structured data with support for basic graph computation. In contrast, GraphLab focuses on iterative graph structured computation.

There are several notable projects focused on using MapReduce for graph computation. Pegasus [20] is a collection of algorithms for mining large graphs using Hadoop. Surfer [7] extends MapReduce with a propagation primitive, but does not support asynchronous or dynamic scheduling. Alternatively, large graphs may be “filtered” (possibly using MapReduce) to a size which can be processed on a single machine [22]. While [22] was able to derive reductions for some graph problems (e.g., minimum spanning tree), the techniques are not easily generalizable and may not be applicable to many MLDM algorithms.

### 7. CONCLUSION AND FUTURE WORK

Recent progress in MLDM research has emphasized the importance of sparse computational dependencies, asynchronous computation, dynamic scheduling and serializability in large scale MLDM problems. We described how recent distributed abstractions fail to support all three critical properties. To address these properties we introduced Distributed GraphLab, a graph-parallel distributed framework that targets these important properties of MLDM applications. Distributed GraphLab extends the shared memory GraphLab abstraction to the distributed setting by refining the execution model, relaxing the scheduling requirements, and introducing a new distributed data-graph, execution engines, and fault-tolerance systems.

We designed a distributed data graph format built around a two-stage partitioning scheme which allows for efficient load balancing and distributed ingress on variable-sized cluster deployments. We designed two GraphLab engines: a chromatic engine that is partially synchronous and assumes the existence of a graph coloring, and a locking engine that is fully asynchronous, supports general graph structures, and relies upon a novel graph-based pipelined locking system to hide network latency. Finally, we introduced two fault tolerance mechanisms: a synchronous snapshot algorithm and a fully asynchronous snapshot algorithm based on Chandy-Lamport snapshots that can be expressed using regular GraphLab primitives.

We implemented distributed GraphLab in C++ and evaluated it on three state-of-the-art MLDM algorithms using real data. The evaluation was performed on Amazon EC2 using up to 512 processors in 64 HPC machines. We demonstrated that Distributed GraphLab significantly outperforms Hadoop by 20-60x, and is competitive with tailored MPI implementations. We compared against BSP (Pregel) implementation of PageRank, LoopyBP, and ALS and demonstrated how support for dynamic asynchronous computation can lead to substantially improved convergence.

Future work includes extending the abstraction and runtime to support dynamically evolving graphs and external storage in graph databases. These features will enable Distributed GraphLab to continually store and processes the time evolving data commonly found in many real-world applications (e.g., social-networking and recom
mender systems). Finally, we believe that dynamic asynchronous graph-parallel computation will be a key component in large-scale machine learning and data-mining systems, and thus further research into the theory and application of these techniques will help define the emerging field of big learning.

Acknowledgments

This work is supported by the ONR Young Investigator Program grant N00014-08-1-0752, the ARO under MURI W911NF0810242, the ONR PECASE-N00014-10-1-0672, the National Science Foundation grant IIS-0803333 as well as the Intel Science and Technology Center for Cloud Computing. Joseph Gonzalez is supported by a Graduate Research Fellowship from the National Science Foundation and a fellowship from AT&T Labs.

8. REFERENCES