Projections of global energy utilization ensure that combustion will continue to be the dominant mode of energy conversion for transportation, power generation, and industrial thermal processes for many decades. Considerations of energy and environmental security and sustainability, as well as economic competitiveness, demand accelerated development of advanced combustion technologies that combine high efficiency, low emissions, and the ability to reliably operate on an increasingly diverse range of fuels, including bio-derived and synthesized fuels, as well as an evolving feed of fossil fuels. Such technological developments are significantly constrained by the lack of robust, predictive computational design tools for advanced combustion systems. Combustion in practical devices covers myriad time and length scales, and so addressing this need ultimately requires a coordinated multi-scale approach for understanding and predicting combustion in turbulent environments. No single method is capable of handling such disparate ranges; hence, a multi-scale approach is adopted where experiments, theory, and simulations combine to treat overlapping scales from the atomistic to continuum regimes synergistically. Within this context the broad goal of the DOE/SC combustion programs is to create the basic science foundations that will enable predictive modeling in the design of new generations of combustion systems.

Owing to the inherent multi-scale, multi-physics nature of combustion – often spanning 10 decades or more of temporal and spatial scales – the combustion community has self-organized into sub-communities around distinct disciplines and scales. Broadly speaking, there are at least three sub-communities: the combustion chemistry research community, the turbulent reacting flow research community, and the applied combustion research and development community. Within each sub-community there are complicated dependencies for information and data flow between experimentalists, computational scientists and engineers, and theoreticians. Currently, there are good examples of long-term collaborative interactions involving tens of independent research groups, organized largely by grassroots efforts. However, the size and complexity of the data is increasing at such a rapid rate that the limiting factor in disseminating data, integrating results, and discovering new scientific knowledge is the lack of effective collaboration infrastructure, data sharing facilities, and tools to manipulate the data. There is also a need to communicate and share information and data between sub-communities, although to a lesser degree than within a given sub-community, for example, the turbulent reacting flow community is reliant on the chemistry community for the generation of accurate, computationally efficient chemical mechanisms for incorporation into computational fluid dynamics (CFD) simulations. Likewise, the applied industrial combustion R&D community depends upon both the combustion chemistry community and the turbulent reactive flow community for predictive models of key subprocesses in macroscopic combustion CFD simulations used to sweep large parameter spaces in optimizing the design of a combustor.

A centralized combustion science collaboration gateway is needed for providing users with discipline-specific tools and data through a web-hosted portal with the appropriate level of
authentication. The portal should provide access to the specialized simulation and analysis tools and data repositories from the various sub-communities performing combustion research. The remainder of this document gives an overview of the extreme scale science challenges and cyberinfrastructure needs of the combustion chemistry and turbulent reacting flow communities, which are strongly supported by DOE/SC. The applied combustion R&D community will also benefit from investment in a collaborative infrastructure for combustion science through access to benchmark data from experiments and simulations, validated chemical mechanisms and turbulent combustion models, and an array of shared toolsets.

**Combustion Chemistry**

Although the conceptual overall reaction of combustion (fuel + oxygen → water + carbon dioxide + heat) is simple, practical combustion in fact involves thousands of intermediate species and tens of thousands of associated reactions. In the context of predictive simulation of combustion, combustion chemistry research can be coarsely divided into elementary reaction kinetics, comprehensive combustion mechanism development and validation, and mechanism reduction for turbulent combustion simulation.

Advances in theoretical chemistry and elementary reaction kinetics make it possible in principle to calculate the details of all of the important combustion reactions. This chemical knowledge is essential to enable prediction of combustion performance (efficiency and emissions) for different fuels and for all conditions relevant to practical combustors. However, because of the size and complexity of the full chemical problem, first-principles calculation requires large-scale computational resources. Many of the remaining challenges for computational combustion chemistry are related to computing and require better integration of codes and parallel algorithm development to fully take advantage of current HPC capabilities. Elementary kinetics research provides experimental data to both guide and validate computational studies, and tools that enable comparison of all available data with uncertainty analysis are necessary for development of highest fidelity kinetic models. In addition to providing key insight to technically important phenomena such as ignition and extinction, flame propagation, and soot and pollutant formation, a rigorous understanding of key elementary reactions also serves as the foundation of a comprehensive combustion mechanism. Development of computational tools that generate appropriately reduced mechanisms from curated data and models are necessary to capture the relevant combustion chemistry and accurately simulate target observables in turbulent reacting flow simulations. Significant advances in experimental capabilities have led to the generation of large experimental benchmark flame datasets, and coordinated progress in this field will require an appropriate cyber infrastructure that provides means and tools to assimilate, share, analyze and compare computational and experimental data from the multitude of sources. This cyber infrastructure will enable coordination of efforts across research groups to avoid duplication and to move towards development of a validated and generally predictive mechanism.

**Turbulent Reacting Flows**

In the turbulent reactive flow community the approaches range from optically-based time-resolved measurements of multiple reactive scalars and velocity, to first principles direct numerical simulations (DNS) of turbulent flames and ignition in canonical flows, to large-eddy simulation (LES) where scales beneath the computational grid are modeled to allow calculations within realistic combustor geometries.
In both DNS and LES large systems of partial differential equations representing the nonlinear evolution and coupling of turbulence and chemistry are solved on supercomputers, generating large volumes of raw data that need to be analyzed, compared with experiments, and shared with a broader international modeling community. The amount of data is presently O(petabytes) and will continue to increase. This community requires software tools and infrastructure (middleware and hardware) for broad access to very large data sets from DNS and high-fidelity LES and for collaborative discovery of knowledge from those data sets. Similarly, there is a need for tools and infrastructure to enable broad access to large experimental data sets and to enable the collaborative development and application of methodologies and algorithms for quantitative comparison of multi-dimensional experimental data and simulations. The cyber infrastructure should include scientific workflow tools to orchestrate large-scale data movement, data transformations, and analysis & visualization of large, turbulent-combustion computations and experiments. These workflow tools need to operate in situ in a large-scale computation or experiment, in-transit as data is streamed off the scientific instrument or computer, or as a post-analysis step. The in situ and in-transit workflows are required for computational steering and to reduce the amount of salient data written to persistent storage for archival purposes. The infrastructure will provide digital archiving of key experimental and simulation benchmark validation data sets for model assessment and capturing the results of model calculations, with ease of access, to better document progress and avoid duplication. Curation and metadata tools are required to capture the community establishment of best practices, methods of analysis, baseline data for experiments and simulations, all based on quality metrics and uncertainty quantification (UQ) methods where possible. The infrastructure and toolsets will expand with time and evolve according to the needs of the research community.

In order to ensure that DNS and high-fidelity LES simulations in this community will run effectively on future exascale architectures, multi-disciplinary and multi-institutional collaboration and information exchange is required to facilitate a highly iterative, agile co-design process. Here, there is a need for collaboration technology that would enable computer scientists, applied mathematicians, computational combustion scientists, and high performance computing (HPC) vendors to share information, data, and software in a streamlined time-efficient manner. The exascale ecosystem involves multiple linkages and spans a wide range of topics in applied mathematics, numerical algorithms, and computer science. Hence there is a continual need for the translation of requirements and contributions across the fields. Translation is also necessity on the technical side of development to align data formats, vocabularies, and ontologies for interoperability and integration across all of the domains. Effective collaboration technology would help reduce the level of unwieldiness associated with effective communication of results, software, and data between all of the computer science & applied math research and industry stakeholders to co-design an optimal software stack and hardware architecture.